6282

MOLECULAR DYNAMICS STUDY OF MACROMOLECULAR MOTIONS AND POLYPEPTIDE HELICAL STABILITIES IN SOLUTION

by

PEMRA DORUKER

B.S. in Ch.E., Boğaziçi University, 1987
M.S. in Ch.E., Boğaziçi University, 1991

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor
in
Philosopy

Bogazici University Library

39001100024044

Boğaziçi University

1995

ACKNOWLEDGMENTS

I would like to express my gratitude to my thesis advisor, Prof. Dr. İvet Bahar, for her guidance, encouragement and helpful criticisms throughout my study. I am also grateful to Prof. Dr. Burak Erman for his interest in my work and constructive suggestions.

Special thanks are due to Prof. Dr. Salih Dinçer for the time he has devoted to reading and commenting on my thesis.

I would like to thank my friends at the Polymer Research Center for their help and patience, especially during the preparation of this thesis. I wish to remember my friends at the Chemical Engineering Department for the hope they have given me in all aspects.

This thesis is dedicated to my dear mother...

ABSTRACT

Molecular dynamics simulations have been performed to investigate the relative importance of the specific solvent effect and the intramolecular conformational potentials on the static and dynamic properties of polymers in dilute solution. Bead-spring, freely-rotating and polyethylene model chains, each composed of 30 beads, have been simulated in 738 solvent molecules. Equilibrium properties of chains, such as overall chain dimensions, pair-correlation functions and static scattering functions have been studied, as well as dynamic properties such as translational diffusivities, dynamic scattering characteristics and orientational correlation functions. Results indicate that chain statistics are equally sensitive to solvent effect and intramolecular constraints, whereas local chain dynamics are predominantly described by intramolecular potentials. Translational diffusivities of all chains exhibit a linear decrease with increasing polymer-solvent interactions. A power law relation of the form $\tau \sim n^a$ is observed between the orientational correlation times τ and the size n of chains segments for $n \leq 8$. *a* values increase with solvent quality.

Stability of short helical polypeptides have been studied by molecular dynamics simulations. 13-residue polypeptides, composed of alanine, valine, serine or glycine, have been simulated in water and in *vacuum*. It is observed that water destabilizes the helical structure of alanine and valine, whereas glycine and serine exhibit intrinsic helix-breaking propensities in the absence of water. Local solvation patterns around alanine C^β, valine C^γ and serine O^γ atoms are found to be similar in helical state. The locations of the first hydration shells of alanine C^β and valine C^γ are located closer to the atoms than those observed in protein crystals structures. Backbone hydrogen bonds are well protected by valine side chains, compared to alanine and serine. Distributions of water molecules around the backbone hydrogen bonds before and after these bonds break clearly shows that helical structure is destabilized as a result of water molecules competing to form alternative hydrogen bonds with the backbone.

ÖZET

Moleküler dinamik yöntemi kullanılarak, seyreltik çözeltilerde polimer-çözücü etkileşme parametresi ve polimerin iç yapısını belirleyen potansiyellerin, zincirin özellikleri üzerindeki etkisi araştırılmıştır. 30 birimden oluşan üç model zincir (boncuk-yay model zincir, serbest dönme açılarına sahip zincir ve polietilen) 738 çözücü molekülü içinde simüle edilmiştir. Zincir boyutları, ikili korelasyon ve statik saçılma fonksiyonları gibi denge özellikleri ve difüzyon katsayısı, dinamik saçılma fonksiyonu ve yerel yönlenme korelasyonları gibi dinamik özellikler incelenmiştir. Model zincirlerin statik davranışının, çözücü etkisi ve iç yapısal etkilere aynı oranda bağlı olduğu, oysa polimerin lokal dinamiğini zincirin iç yapısının belirlediği gözlenmiştir. Difüzyon katsayısı polimer-çözücü etkileşmesi arttıkça doğrusal olarak azalmaktadır. Yerel yönlenme korelasyonlarının karakteristik zamanı (τ) incelendiğinde, $n \leq 8$ birimden oluşan kısa segmanlar için $\tau \sim n^a$ şeklinde bir bağıntı bulunmuştur. Buradaki *a* katsayısı polimer-çözücü etkileşmesine orantılı olarak artmaktadır.

Sarmal yapıdaki kısa polipeptit zincirlerinin stabilitesi moleküler dinamik simülasyon yöntemi ile incelenmiştir. 13 birimden oluşan ve alanin, valin, serin veya glisin içeren polipeptitler suda ve vakum şartlarında incelenmiştir. Suyun alanin ve valinin sarmal yapısını bozduğu, bunun yanı sıra serin ve glisinin suyun etkisinden bağımsız olarak sarmal yapıyı bozma özellikleri olduğu gözlenmiştir. Alanin C^{β}, valin C^{γ} and serin O^{γ} atomları etrafındaki su molekülleri benzer bir dağılım sergilemektedir. Bu atomlar etrafındaki suların ilk hidrasyon düzeyi, protein kristal yapılarında gözlenenlere oranla daha yakında bulunmaktadır. Valinin yan grupları, sarmal yapıdaki hidrojen bağlarını, su moleküllerinden alanin ve serine oranla daha iyi korumaktadır. Sarmal yapıyı oluşturan hidrojen bağları bozulmadan önce ve sonra, etraflarında bulunan su moleküllerinin dağılımı incelendiğinde, su moleküllerinin alternatif hidrojen bağları oluşturarak sarmal yapıyı bozdukları gözlenmektedir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$a_1, a_2 ... a_5$	Coefficients of the dihedral potential
$\mathbf{a}_i$	Acceleration of particle i
b_{on}	Equilibrium bond length
b_i, l_i	Instantaneous lenght of ith bond
C_{12}	Coefficient of repulsive term of Lennard-Jones potential
C_6	Coefficient of attractive term of Lennard-Jones potential
C^{α}	Alpha carbon atom of amino acids
C^{β}	Beta carbon atom of amino acids
C^{γ}	Gamma carbon atom of amino acids
D	Translational diffusivity
E	Identity matrix
f	Fraction of helix in polypeptides
$\mathbf{f}_i$	Total force exerted on particle i
g_{bb}	Polymer-polymer pair correlation function
g_{bs}	Polymer-solvent pair correlation function
g_r	Radial distribution function of water molecules
g_{ss}	Solvent-solvent pair correlation function
g_{ϕ}	ϕ angle distribution of water molecules
g_{θ}	θ angle distribution of water molecules
g^+, g^-	*gauche*$^+$ and *gauche*$^-$ conformers of dihedral angles
k_b	Force constant of harmonic bond stretching potential
K_{bn}	Force constant of harmonic bond stretching potential
k_n	Multiplicity of sinusoidal dihedral potential
K_{unf}	Equilibrium constant of protein unfolding
k_o	Force constant of quasi-harmonic potential
k_{ϕ}	Force constant of dihedral angle torsion potential
$K_{\phi n}$	Force constant of harmonic improper dihedral potential
k_{θ}	Force constant of bond bending potential
$K_{\theta n}$	Force constant of harmonic bond bending potential
$K_{\xi n}$	Force constant of sinusoidal dihedral potential
L	Length of simulation box edges
l_o	Lenght parameter of quasi-harmonic potential
l_b	Lenght parameter of harmonic bond stretching potential
M_1, M_2	First and second orientational autocorretion functions

m_i	Mass of particle i
$\mathbf{m}$	Unit vector in motion
$\mathbf{m}_\perp$	Out-of-plane unit vector
$\mathbf{m}_{\|\|}$	Unit vector along bonds of the chain
N	Number of total molecules
[N]	Concentration of folded (native) proteins
N_{at}	Number of atoms in the system
N_b	Number of covalent bonds in the system
N_b	Number of beads in the polymer chain
N_s	Number of solvent molecules
N_θ	Number of bond angles in the system
N_ϕ	Number of dihedral angles in the system
N_ξ	Number of improper dihedral angles in the system
$O\gamma$	Gamma carbon atom of amino acids
P	Pressure
P_o	Reference pressure of external bath
P_α	Relative frequency of amino acids to occur in α-helices
q_i	Charge of ith atom
$\mathbf{q}$	Scattering wave vector
R	Gas constant
$\mathbf{r}_{cm}$	Position vector of the polymer center of mass
$\mathbf{r}_i$	Position vector (coordinates) of particle i
r_{ij}	Distance between atoms i and j
R_g	Radius of gyration of a chain ($< s^2 >^{1/2}$)
$< r^2 >$	Mean square end-to-end radius of a chain
$< r^2 >_o$	Unperturbed mean square end-to-end radius of a chain
$S(r_{ij})$	Switching function for nonbonded interactions
S(q)	Static scattering function
S(q,t)	Intermediate scattering function
t	Time
t	*trans* conformer of dihedral angles
T	Temperature
$\mathbf{T}$	Transformation matrix
T_o	Reference temperature of external bath
U	Total potential energy
[U]	Concentration of unfolded (denatured) proteins
U_b	Bonded interaction energy
U_{nb}	Non-bonded interaction energy

V	Volume of simulation box
V_b	Harmonic bond stretching potential
$\mathbf{v}_{cm}$	Velocity of the polymer center of mass
V_{ES}	Electrostatic interaction energy
$\mathbf{v}_i$	Velocity of particle i
V_{LJ}	Lennard-Jones potential energy
V_{SF}	Shifted-force potential energy
V_{QH}	Quasi-harmonic (FENE) potential energy
V_θ	Harmonic bond bending potential
V_ϕ	Dihedral angle torsion potential
V_ξ	Improper angle torsion potential
W_{QH}	FENE potential bond length distribution
α	Expansion coefficient of a chain
β	Compressibility
β_1, β_2	Exponents of the stretched exponential function
δ_n	Phase shift of sinusoidal dihedral potential
ΔG	Free energy difference
Δt	Time step size
ε_o	Dielectric permitivity in vacuum
ε_r	Relative dielectric permitivity
ε	Lennard-Jones energy parameter
θ_o	Supplemental bond angle
θ_{on}	Equilibrium bond angle
θ_i	Instantaneous angle of ith bond
λ	Wavelength of radiation, velocity scaling factor
ρ	Number density
ρ_m	Mass density
σ	Lennard-Jones length parameter
ξ_{on}	Equilibrium improper dihedral angle
ξ_i	Instantaneous value of improper dihedral angle
τ, τ_1, τ_2	Characteristic relaxation times
τ_P	Pressure scaling time constant
τ_T	Temperature scaling time constant
μ	Length scaling factor
(ϕ,ψ)	Backbone dihedral angle pairs in polypeptides
ϕ_i	Instantaneous value of ith dihedral angle
χ_1	Rotation angle of amino acid side chains
Ω	First cumulant

1. INTRODUCTION

Theoretical studies on polymers generally consider single polymer chains under *theta* solvent conditions, because the incorporation of the specific polymer-solvent interactions and the excluded volume effect in analytical models is quite difficult. However, it is known that these interactions significantly alter the properties of polymers in solution. In this respect, molecular dynamics (MD) simulation is an alternative and useful tool to investigate the behavior of polymers in various environment. MD simulations can produce realistic trajectories in time by precisely considering all the intra- and intermolecular interactions in the system, composed of the polymer and explicit solvent molecules.

In the first part of this thesis presented in Section 2, MD simulations are performed for model polymer chains of 30 monomers, immersed in 738 solvent molecules. Three chain models, which have been widely used in previous studies on polymers, are adopted, namely, the bead-spring[1], the freely-rotating[2] and the polyethylene[2] chain models. The polymer-solvent interaction energy parameter is systematically varied in the simulations to achieve distinct -poor, close to *theta* and good- solvent conditions. The solvent effect on the equilibrium properties of the model chains, such as pair correlation and static scattering functions is investigated. Two dynamic properties are of interest in this study: (i) the translational diffusivity of the overall chain, and (ii) the orientational mobility of chain segments of various sizes. The former is a global property which has been widely investigated for athermal solutions, in general. The second refers to internal conformational motions in polymers, which are critically important for the interpretation of experiments, such as NMR spectroscopy, fluorescence anisotropy and dielectric relaxation. This is the first detailed MD analysis of segmental relaxation processes in varying solvent conditions.

Small polypeptides, composed of less than 20 amino acid residues, are known to acquire α helical conformation in aqueous solution under physiological conditions. This secondary structure formation, which is not observed in synthetic polymers in solution, plays an important role in the initial stage of protein folding[3]. Amino acids display significant variation in propensity for α-helical, β-

sheet and other main chain conformational states in proteins and polypeptides. An important stage in the solution of the protein folding problem is to understand the factors that determine the tendency of short polypeptide sequences to populate different conformational states in solution. Several factors have been suggested to be responsible for the α-helix propensities of amino acids, such as steric factors, conformational entropy loss, hydrophobic effect and main-chain electrostatics. Meanwhile, it is accepted that water plays an important role in the formation of secondary structures because it competes to form hydrogen bonds with the main chain polar groups.

In the second part of this thesis, MD simulations are carried out for different model polypeptides in *vacuum* and in the presence of explicit water molecules. Model polypeptides, composed of 13 residues of alanine, valine, serine or glycine, are explcred. The dynamics of structural transitions are analyzed by monitoring the fractional helix contents of the peptides during the trajectories. Water distribution around the side chains and the backbone hydrogen bonds are investigated. The different solvation patterns of the four amino acids obtained in this study are correlated with their helix forming / breaking tendencies.

2. STATISTICS AND DYNAMICS OF POLYMER CHAINS IN DILUTE SOLUTION

Models based on single chain statistics[2] are applicable only in particular cases, such as dilute polymer-solvent systems under *theta* conditions or polymers in the bulk state, where a given chain does not distinguish between the surrounding molecules and intramolecular chain segments. Otherwise, the intrachain excluded volume effect and the specific solvent-polymer interaction, which are known to perturb chain equilibrium and dynamic properties[4, 5], need be considered for a realistic estimation of behavior in solution.

Simulation of polymer-solvent systems is a useful tool to investigate the behavior of polymers in various environment.[6, 7] Among various computational methods, Monte Carlo algorithms are suitable for studying equilibrium properties of large systems, such as long polymer chains or concentrated systems. These algorithms do not supply information on chain dynamics, except for the dynamic Monte Carlo method coupled with Metropolis algorithm. Brownian dynamics (BD) simulations reproduce most of time-dependent processes and are particularly useful for the study of relaxation phenomena with time scales on the order of nanoseconds. However, the local solvent structure and dynamics cannot be realistically represented in BD, because the solvent-polymer interactions are averaged by a normal stochastic noise and an adjustable friction coefficient is adopted for representing the effective frictional drag of the surroundings. In this respect, molecular dynamics (MD) simulations, which consider precisely the interaction between the individual constituents of a polymer-solvent system, are particularly suitable. It is noted that the intrachain excluded volume effect is also inherently present in MD simulations, as interactions between all nonbonded units along the chain are explicitly accounted for.

Molecular dynamics has proven to be a valuable tool for understanding the mechanism and evolution of several time-dependent processes in polymeric systems, such as orientational and translational motions in solution[8-17] or in the bulk state,[18, 19] freezing in of internal rotational motions near the glass transition,[20] and diffusion of simple gas molecules in polymer matrix.[21, 22] The major drawback of the MD method is that the complexity of the models and

the size of the systems to be studied are limited by the computational resources available.

The model chains used in MD simulations are selected at different levels of sophistication depending on the specific properties under study. Bead-rod model chains (or Kuhn equivalent freely jointed chain) and bead-spring chains are classical examples which have been adopted in early simulations of chain dynamics[8, 9, 11, 10, 12] and continue to be explored.[13-17] These simple models are particularly suitable for verifying various scaling arguments[23] and/or for establishing some concise analytical expressions relating chain size to static and dynamic characteristics. Another group of MD studies[24, 20, 25, 21, 22, 18, 19] incorporates the structure and energy parameters of real chains, based on the rotational isomeric state model[2] of equilibrium statistics; their approach is essential for establishing the physical connection between theory and experiments, and for rationalizing the distinct behavior of different macromolecules.

The first molecular dynamics simulations of polymer chains in solution have been performed by Bishop et. al.[8] and Rapaport[9]. Later, MD studies[11, 10, 12-14] have been carried out to investigate the role of interaction potentials, solvent density and quality on the various equilibrium and dynamic properties of polymer chains in solution. Bead-rod and bead-spring model chains of various lengths have been used in these particular studies.

On the same line of work, MD simulations for bead-spring model chains immersed in good solvents of varying quality have been performed[26] and the results indicate a systematic perturbation in static and dynamic properties of the chain due to the specific solvent effect. This study on bead-spring model chains, which will be presented in Section 2.2.1, invites attention to the importance of the interaction of polymer chains with their surroundings. Then, the question arises whether the solvent effect might have been overestimated due to the fact that bead-spring model chains do not carry any structural and conformational constraints other than chain connectivity. Thus, a complementary study has been carried out[27] to analyze the importance of solvent effect on polymer chains which are subject to more realistic intramolecular constraints, i.e. which possess some inherent stiffness in contrast to the fully flexible bead-spring model chains. The second part of the work on dilute polymer-solvent systems, which comprises Section 2.2.2 , summarizes the simulation results for two more realistic model

chains. The classical freely-rotating model chain[2], in which bond angles are constrained to be fixed around the tetrahedral value, whereas bonds are allowed to rotate freely about their own axis, is one of the models investigated. In the second model, the structure and energy parameters of polyethylene chains based on the rotational isomeric state model of equilibrium statistics[2] is adopted, where the CH_2 monomeric units are taken as collapsed on the chain backbone. The relative importance of the specific solvent effect and the intrachain potentials, i.e. bond stretching, bending and torsional potentials is analyzed by simulating both the freely-rotating (FRC) and the polyethylene-like(PE) model chains in solvents of varying polymer-solvent interactions. Details of the simulations, which are almost identical for the three model chains, will be explained in Section 2.1. However, the results of the simulations will be presented in separate sections, namely Section 2.2.1 on the bead-spring model chain and Section 2.2.2 on the freely-rotating and polyethylene-like model chains.

2.1. Simulation Model and Method

2.1.1. Molecular Models

The total potential energy U for the system of N particles, consisting of a polymer chain of N_b beads, i.e. monomer units collapsed on the backbone, surrounded by N_s solvent molecules, is obtained by summing up all the nonbonded (U_{nb}) and bonded (U_b) interaction energies.

$$U = U_{nb} + U_b \tag{2.1}$$

In all simulation systems, the nonbonded interactions between polymer bead-solvent, bead-bead, and solvent-solvent pairs are included in U_{nb} according to the following summation.

$$U_{nb} = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} V_{SF}(r_{ij}) \tag{2.2}$$

Here, the interaction between the nonbonded pair of particles *i* and *j* at a distance $r_{ij} \equiv | \mathbf{r}_j - \mathbf{r}_i |$ from each other is taken to be of the form of the shifted-force potential, $V_{SF}(r_{ij})$[6]

$$V_{SF}(r_{ij}) = \begin{cases} V_{LJ}(r_{ij}) - V_{LJ}(r_c) - (r_{ij} - r_c)\,(dV_{LJ}(r_{ij}) / dr_{ij})_{r_{ij}=r_c} & r_{ij} \le r_c \\ 0 & r_{ij} > r_c \end{cases} \tag{2.3}$$

where $V_{LJ}(r_{ij})$ is the Lennard-Jones (LJ) (6-12) potential given by

$$V_{LJ}(r_{ij}) = 4\,\varepsilon_{ij}\left[(\sigma_{ij} / r_{ij})^{12} - (\sigma_{ij} / r_{ij})^{6}\right] \tag{2.4}$$

and r_c is the cutoff distance beyond which the interaction vanishes. ε_{ij} and σ_{ij} in Equation (2.4) are the respective energy and distance parameters corresponding to the particular pair of particles *i* and *j*. The adoption of a cutoff separation introduces discontinuities in the absolute value and slope of the interaction energy at r_c, which are eliminated by the addition of the last two terms in Equation (2.3). Minimum image convention is used for the specification of r_{ij}.[6] The subscripts *i* and *j* of the above variables are later replaced by either *s*

(solvent) or *b* (polymer bead) depending on the type of interacting particles. The van der Waals radii and consequently the length parameters of polymer beads and solvent molecules are assumed to be equal to each other. Thus, the subscript ij in the LJ length parameter σ_{ij} can be omitted for brevity. The cutoff separation is taken as $r_c = 2.5\ \sigma$.

In bead-spring model chains, the elastic potential of the ith spring is given by the so-called finitely extendable nonlinear elastic or FENE potential[1]

$$\begin{aligned} V_{QH}(l_i) &= -0.5\ k_o l_o^2 \ln\left[1 - (l_i / l_o)^2\right] \qquad l_i < l_o \\ V_{QH}(l_i) &= \infty \qquad\qquad\qquad\qquad\qquad\quad l_i \geq l_o \end{aligned} \tag{2.5}$$

where l_i is the instantaneous length of the ith spring. The subscript QH refers to the quasiharmonic nature of the potential. k_o and l_o are the energy and length parameters of the quasiharmonic potential, respectively. The above form of connector potential has found widespread use in previous MD studies of polymer solutions.[11, 14, 15] This function is monotone strictly increasing with l_i, thus favoring small extensions and eventual overlap of the beads. Yet, the beads can not actually approach each other by more than some limiting separation, as inherently implemented by the shifted potential function of Equation (2.3), which applies to all pairs of nonbonded beads. Thus, the total bonded interaction energy for the bead-spring model chain is given by the following summation

$$U_b = \sum_{k=1}^{N_b - 1} V_{QH}(l_k) \tag{2.6}$$

In the polyethylene (PE) model chain, presented in Figure 2.1, the bonded interaction energy includes contributions from bond stretching, bond

angle bending and bond torsional motions, given by the first, second and third summations in the following equation, respectively.

$$U_b = \sum_{k=1}^{N_b - 1} V_b (l_k) + \sum_{m=1}^{N_b - 2} V_\theta (\theta_m) + \sum_{n=2}^{N_b - 2} V_\phi (\phi_n) \qquad (2.7)$$

Here, the spring-like character of the bonds connecting the beads are maintained by the harmonic potential, $V_b (l_i)$

$$V_b (l_i) = (k_b / 2) (l_i - l_b)^2 \qquad i = 1,... N_b - 1 \qquad (2.8)$$

where k_b is the bond-stretching force constant, l_i is the instantaneous length of the ith spring connecting atoms C_{i-1} and C_i, as shown in Figure 2.1, and l_b is the equilibrium bond length.

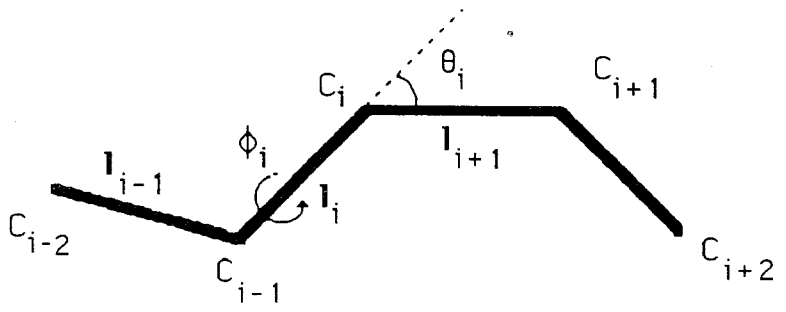


FIGURE 2.1. A section of the polyethylene-like model chain between atoms C_{i-2} and C_{i+2} The generalized coordinates are l_i, θ_i, ϕ_i, where l_i is the bond vector between atoms C_{i-1} and C_i, θ_i is the supplemental bond angle at C_i and ϕ_i is the torsional angle defined by bond vectors l_{i-1}, l_i and l_{i+1}.

The supplemental bond angle θ_i at the i th atom is constrained to fluctuate about θ_o by the quadratic potential, $V_\theta(\theta_i)$, with a force constant k_θ

$$V_\theta(\theta_i) = (k_\theta / 2)(\cos\theta_i - \cos\theta_o)^2 \qquad i = 1,\ldots N_b - 2 \qquad (2.9)$$

Finally, the torsional angle ϕ_i, defined by the spatial position of bond vectors $\mathbf{l}_{i-1}$,$\mathbf{l}_i$ and $\mathbf{l}_{i+1}$, is maintained by the rotational potential $V_\phi(\phi_i)$[24, 28]

$$V_\phi(\phi_i) = k_\phi \sum_{n=0}^{5} a_n \cos^n \phi_i \qquad i = 2,\ldots N_b - 2 \qquad (2.10)$$

where k_ϕ is the bond torsion constant, and the coefficients a_n ($0 \le n \le 5$), listed in Table 2.1, satisfy the relationship $\sum_{n=0}^{5} a_n = 0$. This potential ensures that the dihedral angles assume values in the neighborhood of three rotational minima, namely, *trans* (t, 0°), *gauche*$^+$ (g^+, 120°) and *gauche*$^-$ (g^-, -120°).

In the freely-rotating chain model, $V_\phi(\phi_i) = 0$ allowing the torsional angles to freely rotate about their own axis. In the freely-rotating and polyethylene-like model chains, nonbonded interactions among polymer beads which are separated by less than four bonds along the chain are not included in the summation of Equation (2.2) to avoid any bias of bond torsional and angle bending potentials.

2.1.2. Simulation System and Parameters

The system consists of $N_s = 738$ solvent molecules and a polymer chain of $N_b = 30$ beads, each of equal mass m. A cubic simulation box composed of body-centered cubic (bcc) lattice sites is adopted for describing the original coordinates of the set of $N_b + N_s = N$ particles.The box is subject to periodic boundary conditions.

The particles are originally placed at the centers and four of the corners along opposite diagonals of each bcc lattice site, thus leading to a tetrahedral arrangement for the polymer bonds in their original state. This structure will be certainly lost during simulation but such a symmetric arrangement of chain beads at start has the advantage of necessitating shorter equilibration period compared to randomly placed beads. Following this prescription, each site accommodates 1.5 particles. Thus, the set of $N = 768$ particles considered in the majority of simulations results from a cubic box of 8 bcc sites along each edge. The first bead of the polymer chain and the particular sequence of connected beads are selected by a random number generator subroutine. Accordingly, for the choice of the bead $i+1$ along the chain, the random number generator assigns one tetrahedral directional vector among those (three or less) accessible to bead i. The term accessible refers here to those directional vectors leading to sites which are not already occupied by the previously selected beads. The position vector of the ith particle (polymer bead or solvent) is denoted by $\mathbf{r}_i$, with respect to the laboratory-fixed frame whose origin is conveniently located at the center of the simulation box.

The initial velocities are assigned in conformity with the Boltzmann distribution at the simulation temperature (T = 378 K) and the net linear momentum is set equal to zero by subtracting the mean velocity from the initial velocities. Isothermal conditions are maintained by monitoring and rescaling the velocities of the particles at regular time intervals of 200 time steps. Inasmuch as the total energy is also conserved along the trajectory, the resulting averages approximate the behavior of a isenthalpic-canonical ensemble.

The values of the parameters used in simulations are listed in Table 2.1, as well as the expressions and values for their reduced counterparts, indicated with the subscript r. ε^*, σ^* and m*in Table 2.1 are the respective energy, length, and mass-scaling parameters used for obtaining the reduced quantities. The LJ parameter ε_{bs} is subjected to variations within the range $0.1 \leq \varepsilon_{bs} \leq 0.8$ kcal/mol, while the other two interaction potentials are kept fixed at $\varepsilon_{bb} = \varepsilon_{ss} = 0.5$ kcal/mol. The simulation box has a volume $V = L^3 = (24.3\ \text{Å})^3$, where L equals 9.0 in reduced units. Since the box accomodates 768 particles in total, the reduced number density $\rho_r = 1.054$ is equivalent to an absolute mass density ρ_m of 1.07 g / cm^3, as follows from the identity $\rho_m = N m / L^3$. The values $L = 9\ \sigma^*$ and $N_s = 738$ have been used in the simulations, except for a few runs performed with the aim of estimating the effect of system size and the extent of hydrodynamic interactions on the polymer properties, as will be indicated in the text.

2.1.3. Simulation Algorithm

The equation of motion for the ith particle, $1 \leq i \leq N$, is

$$m_i \mathbf{a}_i = \mathbf{f}_i = - \nabla_{\mathbf{r}_i} U \qquad (2.11)$$

where $m_i = m$ is the mass of particle *i*, $\mathbf{f}_i$ is the force exerted on particle *i*, $\mathbf{a}_i$ is its acceleration and $\nabla_{\mathbf{r}_i} U$ is the gradient of the potential U with respect to $\mathbf{r}_i$. For a given configuration at time t, $\mathbf{f}_i$ and $\mathbf{a}_i$ are calculated for each particle from the gradient of the potentials described in Section 2.1.1. The new positions of the particles at time $(t + \Delta t)$ are, then, computed by using a modified Verlet algorithm with the recurrence equations given by[6]

TABLE 2.1. Simulation Parameters in Absolute and Reduced Units

Real Variable	Absolute Value	Dimensionless Variable	Reduced Value
m	12 g mol^{-1}	$m_r = m / m^*$	1
T	378 K	$T_r = RT / \varepsilon^*$	0.75
$V = L^3$	$(24.3\ Å)^3$	$V_r = V / \sigma^{*3}$	9.0^3
$\rho = N / V$	0.0535 $Å^{-3}$	$\rho_r = \rho\ \sigma^{*3}$	1.054
Δt	4.57 fs	$\Delta t_r = \Delta t / (\sigma^{*2} m^* / \varepsilon^*)^{1/2}$	0.01
σ	2.7 Å	$\sigma_r = \sigma / \sigma^*$	1
ε_{bb}	0.5 kcal mol^{-1}	$\varepsilon_{bbr} = \varepsilon_{bb} / \varepsilon^*$	0.5
ε_{ss}	0.5 kcal mol^{-1}	$\varepsilon_{ssr} = \varepsilon_{ss} / \varepsilon^*$	0.5
ε_{bs}	0.1-0.8 kcal mol^{-1}	$\varepsilon_{bsr} = \varepsilon_{bs} / \varepsilon^*$	0.1-0.8
r_c	6.75 Å	$r_{cr} = r_c / \sigma^*$	2.5
k_o	$20\ \varepsilon_{bb} / \sigma^2$	$k_{or} = k_o\ \sigma^{*2} / \varepsilon^*$	10
l_o	5.265 Å	$l_{or} = l_o / \sigma^*$	1.95
k_b	82.7 kcal $Å^{-2}$ mol^{-1}	$k_{br} = k_b\ \sigma^{*2} / \varepsilon^*$	602.9
l_b	1.54 Å	$l_{br} = l_b / \sigma^*$	0.57
k_θ	43.5 kcal mol^{-1}	$k_{\theta r} = k_\theta / \varepsilon^*$	43.5
$\cos \theta_o$	1 / 3		
k_ϕ	2.22 kcal mol^{-1}	$k_{\phi r} = k_\phi / \varepsilon^*$	2.22
a_0	1.0000		
a_1	1.3108		
a_2	-1.4135		
a_3	-0.3358		
a_4	2.8271		
a_5	-3.3885		

Reference values: $m^* = 12$ g / mol; $\sigma^* = 2.7$ Å ; $\varepsilon^* = 1.0$ kcal / mol

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \Delta t\ \mathbf{v}_i(t) + \frac{2}{3}(\Delta t)^2\,\mathbf{a}_i(t) - \frac{1}{6}(\Delta t)^2\,\mathbf{a}_i(t-\Delta t)$$

$$\mathbf{v}_i(t+\Delta t) = \mathbf{v}_i(t) + \frac{1}{3}\Delta t\,\mathbf{a}_i(t+\Delta t) + \frac{5}{6}\Delta t\,\mathbf{a}_i(t) - \frac{1}{6}\Delta t\,\mathbf{a}_i(t-\Delta t) \qquad (2.12)$$

Here $\mathbf{v}_i(t)$ is the velocity of particle *i* at time t and Δt is the size of the time step. This algorithm, in which the acceleration $\mathbf{a}_i(t-\Delta t)$ is stored in addition to $\mathbf{r}_i(t)$, $\mathbf{v}_i(t)$ and $\mathbf{a}_i(t)$, allows for a more accurate estimation of velocities compared to the classical Verlet method, and consequently brings about an improvement in energy conservation.[6]

The recurrence Equations (2.12) may be identically rewritten in terms of reduced variables:.$\mathbf{r}_{ir} = \mathbf{r}_i / \sigma^*$; $\mathbf{v}_{ir} = \mathbf{v}_i\,(m^*/\varepsilon^*)^{1/2}$; $\mathbf{a}_{ir} = \mathbf{a}_i\,\sigma^* m^*/\varepsilon^*$. The expression for the reduced velocity $\mathbf{v}_{ir}$ follows from the broadening of the Gaussian distribution of velocities by a factor of $(m^*/\varepsilon^*)^{1/2}$ due to the use of reduced temperature and mass in the Maxwell-Boltzmann expression. A detailed summary of the molecular dynamics simulation algorithm, which is described in the present section, is given in Appendix A.

2. 2. Results and Discussion

2.2.1. Bead-Spring Model Chain

In the simulations with bead-spring model chains, ε_{bsr} is varied between 0.1 and 0.8. At each ε_{bsr} value, several simulations are performed starting from different initial configurations.

In general, an increase in ε_{ij} strengthens the interaction, both attractive and repulsive, between particles *i* and *j*. However, the attractive portion of the LJ

energy curves is more sensitive to ε_{ij} while the steep change in the repulsive regime is weakly affected. Thus, the effective change brought about by an increase in ε_{ij} is to enhance the favorable interaction between particle *i* and *j*. Accordingly, a lower value for ε_{bs} should give rise to a more compact chain configuration. In view of these qualitative features, the influence of solvent-polymer interaction on various polymer properties will be investigated in the following.

2.2.1.1. Mean Energies and Overall Chain Dimensions. Table 2.2 illustrates for three selected values of ε_{bs}, namely 0.2, 0.5 and 0.8 kcal/mol, the mean values and standard deviations of different types of energies, resulting from short runs of 10^4 steps.

TABLE 2.2. Mean energies (kcal / unit) and their standard deviations (S. D.)

ε_{bsr}		V_{QH}	V_{SF}	Kinetic	T_r
0.2	Mean	2.072	-0.096	1.125	0.750
	S. D.	0.051	0.009	0.026	0.017
0.5	Mean	2.070	-0.197	1.125	0.750
	S. D.	0.057	0.008	0.026	0.017
0.8	Mean	2.023	-0.343	1.125	0.750
	S. D.	0.062	0.010	0.027	0.018

Here, the quasiharmonic potential energies V_{QH} represent the average over the trajectory and over the N_b - 1 springs. The shifted force potentials V_{SF} and the kinetic energies correspond to averages over all units, i.e. polymer beads and solvents. The negative values for the mean V_{SF} indicates that, in general, favorable interactions, i.e. good solvent conditions, are brought about by the present choice of simulation parameters.

Table 2.3 summarizes the overall chain dimensions, namely, the mean-square end-to-end separation $< r^2 >$, the radius of gyration $R_g \equiv < s^2 >^{1/2}$, and the expansion coefficient α.

TABLE 2.3. Mean-square chain dimensions and expansion coefficients α as a function of solvent quality

ε_{bsr}	$< r^2 >$ (Å2)	$< s^2 >$ (Å2)	α
0.1	83.86 ± 5.81	19.57 ± 0.93	1.265
0.3	155.30 ± 19.19	29.96 ± 1.29	2.342
0.5	226.73 ± 12.90	40.34 ± 1.69	3.420
0.8	333.90 ± 11.55	55.91 ± 0.04	5.036

$< s^2>$ is given the expression

$$<s^2> = N_b^{-1} \sum_{i=1}^{N_b} < (|\mathbf{r}_i - \mathbf{r}_{cm}|)^2 > \qquad (2.13)$$

where $\mathbf{r}_{cm}$ is the instantaneous position vector of the chain center of mass, and the brackets refer to a time average over several snapshots in a given run. The expansion coefficient is defined as

$$\alpha = < r^2 > / < r^2 >_o \qquad (2.14)$$

where the unperturbed mean-square end-to-end separation $< r^2 >_o$ for bead-spring model chains is calculated from

$$< r^2 >_o = (N_b - 1)\, l_{av}^2 \qquad (2.15)$$

with the average bond-spring length taken as $l_{av} = 0.56\ \sigma^*$ (See Figure 2.5). All the bead-spring chain systems presently studied represent good solvent conditions, as demonstrated by the expansion coefficients $\alpha > 1$, listed in Table 2.3. It is clearly demostrated that the chains expand as the polymer-solvent interaction energy increases.

The ratio $<r^2> / <s^2>$ of the mean-square dimensions of the chains of 30 beads is found to be 6.5±0.7 in the present MD simulations. This follows from the slope of the two lines drawn in Figure 2.2, in which the results from several runs with various ε_{bs} are plotted. The least square fit to the data yields the line with slope 7.2. If the line is constrained to pass through the origin, on the other hand, the best fit leads to a value of 5.8, which is in agreement with the following expression for freely jointed chains of N_b units.

$$<r^2> / <s^2> = 6N_b / (N_b + 1) \qquad (2.16)$$

The slightly larger value presently obtained for $<r^2> / <s^2>$ may be attributed to the the finite persistence length of the present relatively short chains and the excluded volume effect, which is not included in the freely jointed chain model but is implicitly present in the simulation method. It is noted that a wide range of MD values is reported in literature for the ratio $<r^2> / <s^2>$ of short bead-spring model chains. These deviations may be understood in view of the large amplitude scatter observed here in independent MD runs of short duration.

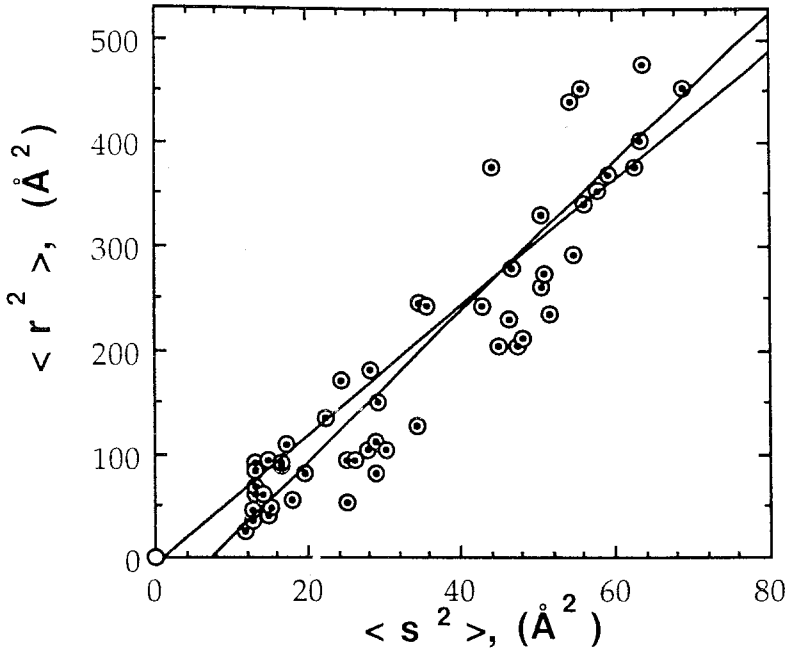


FIGURE 2.2. Dependence of the mean-square end-to-end separation $< r^2 >$ on the radius of gyration $< s^2 >$. Circles represent time averages from independent runs of 10^4 steps evaluated irrespective of solvent-polymer interaction. The best fitting line through the data points yields a slope equal to 7.2. The second line represents the best fit for the line constrained to cross the origin and leads to $< r^2 > / < s^2 > = 5.8$.

2.2.1.2. Pair-Correlation Functions. The pair-correlation function $g_{ij}(r)$, also referred to as radial distribution function, gives the unnormalized probability of finding a pair of particles *i* and *j* at a distance r apart. $g_{ij}(r)$ is evaluated from

$$g_{ij}(r) = (\rho^{-1} / N) < n_{ij}(r) > / (4\pi r^2 \Delta r) \qquad (2.17)$$

where the number $< n_{ij}(r) >$ of particles located at a distance r from each other is found from the time average of the following expression.

$$n_{ij}(r) = 2\int dr \sum_{i} \sum_{j \neq i} \delta(\mathbf{r} - \mathbf{r}_i + \mathbf{r}_j) . \qquad (2.18)$$

This integration is performed over a thin spherical shell of thickness Δr, about **r**, and the summations of the delta functions include all pair of particles of type *i* and *j* with $j \neq i$, changing in the range $[1, N_b]$ for polymer beads and $[1, N_S]$ for solvent molecules. In the limit as r becomes infinitely large, using image particles convention throughout space, $g_{ij}(r)$ approaches the ratio $N_i N_j / N^2$, for $i \neq j$ and $N_i (N_i - 1)/N^2$ for pairs of the same type. Thus, the radial distribution curves should approach unity provided that they are normalized with respect to these asymptotic values. Both time and ensemble averages are used in evaluating $g_{ij}(r)$, leading to curves with minimal noise.

The computation of pair correlation functions from MD simulations is performed here for an assessment of the compliance of our results with previous work,[14] and for a clear visualization of the change in intrachain separations as a function of solvent type. Figures 2.3 and 2.4 illustrate the dependences of $g_{bs}(r)$, $g_{bb}(r)$ and $g_{ss}(r)$ on solvent quality. The grid size is taken as $\Delta r = 0.01\sigma^*$. Short durations of simulation ($\sim 10^4$ time steps) are found to be sufficient to reproduce -indistinguishably- each of the curves displayed in the figures. This tendency of the chain to assume more expanded configurations in good solvent, i.e. with increasing ε_{bs}, is manifested in Figure 2.3 by the larger $g_{bs}(r)$ values attained at short separations in media with higher ε_{bs}, in agreement with the MD results obtained by Freire and collaborators[14] for a smaller system. The successive peaks in the distribution function occur at locations $r / \sigma^* \sim 1.1$, 2.0 and 3.0, in perfect agreement with previous work, in which the same model and comparable energy and length parameters have been adopted.[14] The gradual shifting of the first peak to larger r / σ^* values with increasing ε_{bs} is also in quantitative conformity with previous MD results.[14] The oscillations of $g_{bs}(r)$ corresponding to successive shells of neighbors are observed to be stronger with increasing ε_{bs}. Weaker ε_{bs} values, on the other hand, induce less pronounced peaks associated with more randomized relative positions of particles. This feature is indicative of the increased diffusional mobility of the particles in the presence of weak intermolecular associations, which is further exploited below.

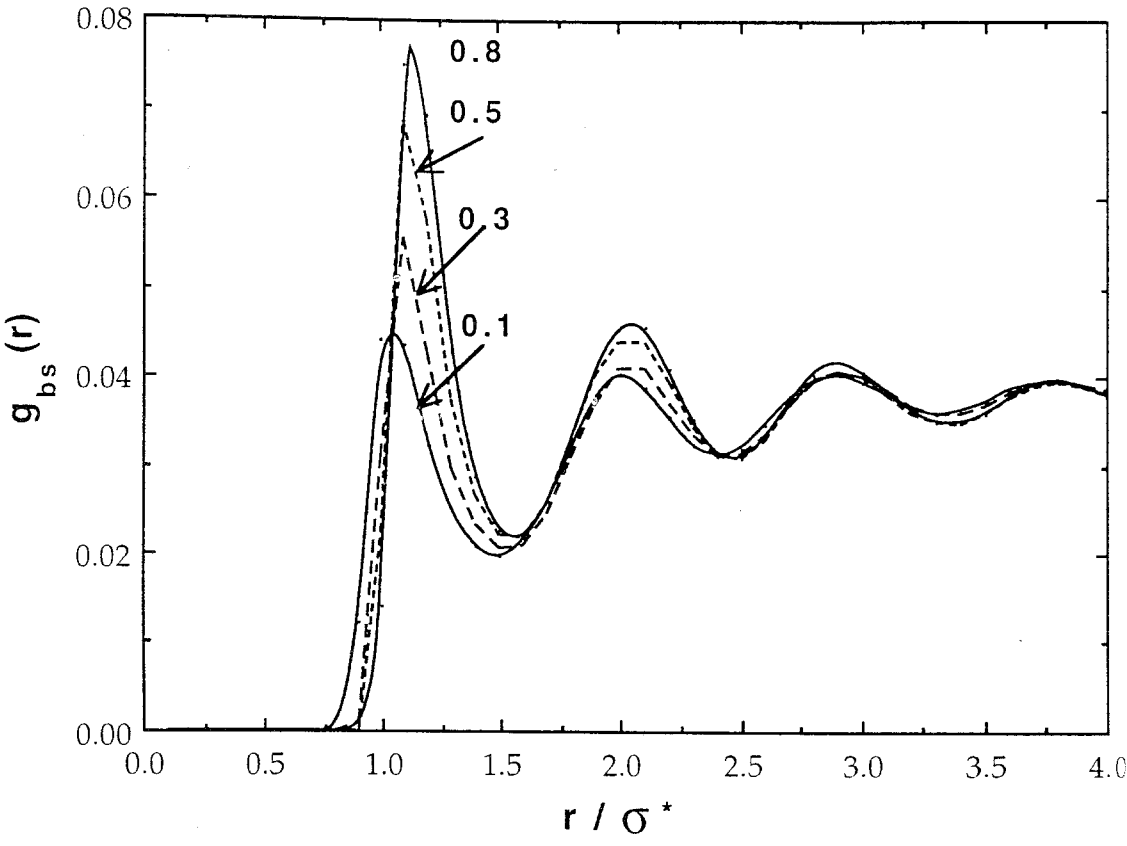


FIGURE 2.3. Radial distribution function $g_{bs}(r)$ for polymer-solvent pair as a function of the reduced separation r / σ^*. Results are shown for systems subject to $\varepsilon_{bsr} = 0.1, 0.3, 0.5$ and 0.8.

The intrachain radial distribution function $g_{bb}(r)$ shown in Figure 2.4 has been evaluated on the basis of nonbonded units along the chain. First neighboring units are not included in this analysis inasmuch as their separation is predominantly determined by the quasiharmonic potential and attention here is primarily focused on the influence of solvent quality on the spatial arrangement of the atoms/beads. The distribution of bond lengths l in response to the changes in solvent quality will be separately considered next. A strong dependence on ε_{bs} is observable in Figure 2.4, in conformity with the behavior described above: The beads come closer to each other as the solvent becomes poorer. The diagram on the upper right part of Figure 2.4 shows the solvent-solvent radial distribution functions $g_{ss}(r)$ which have been obtained in MD simulations, for various ε_{bs}. The insensitivity of $g_{ss}(r)$ to ε_{bs} is legitimate for the highly dilute system presently investigated.

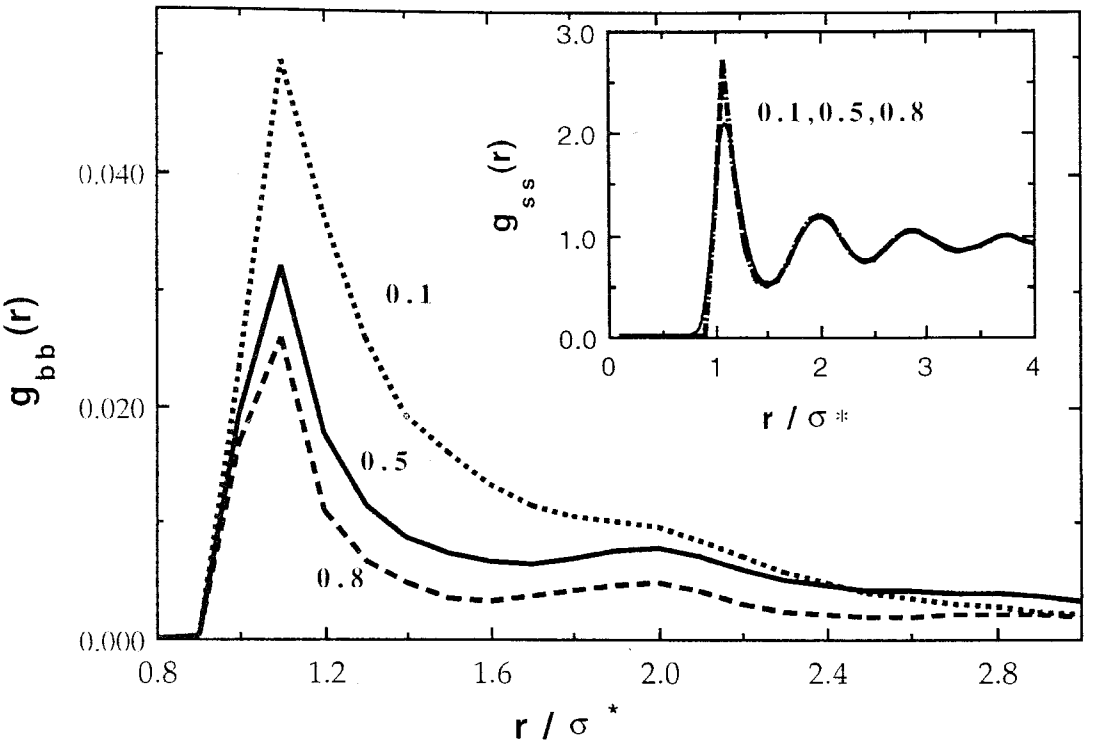


FIGURE 2.4. Radial distribution function $g_{bb}(r)$ for non-bonded beads pairs of the polymer chain as a function of the reduced separation r / σ^*. Results are shown for ε_{bsr} = 0.1, 0.5 and 0.8. The upper small diagram on the right is obtained with the same set of data and illustrates the insensitivity of solvent-solvent pair correlation function $g_{ss}(r)$ to polymer-solvent interaction.

2.2.1.3. <u>Equilibrium Distribution of Bond/Spring Lengths</u>. Figure 2.5 displays the probability distribution $W(l / \sigma^*)$ of the reduced length l / σ^* of bonds for various ε_{bs} values. The zero level of the distribution curves are vertically shifted by a value of 1.0 for clarity. With increasing ε_{bs} the approximately Gaussian shape of the distribution is distorted into a bimodal structure.

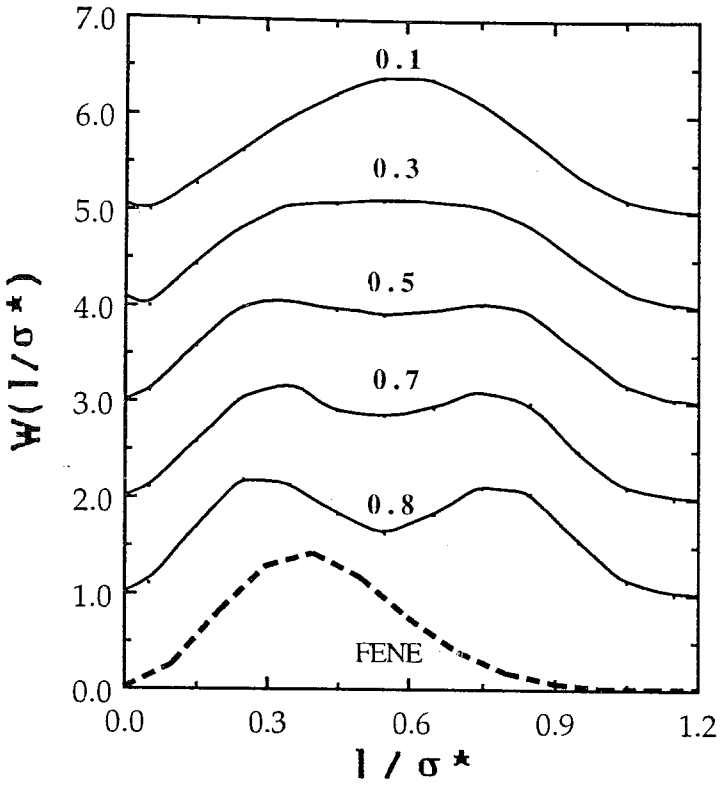


FIGURE 2.5. Distribution W(l) of bond/spring lengths l in chains subject to interactions of various strength with the surroundings. ε_{bsr} values are 0.1, 0.3, 0.5, 0.7 and 0.8 as indicated by the labels on each curve.

For isolated dumbbells subject to the FENE potential, the connector length distribution is given by[1]

$$W_{QH}(l_i) = J\,[\,1 - (l_i / l_o)^2\,]^{k_o l_o^2 / 2kT} \qquad l_i < l_o$$
$$W_{QH}(l_i) = 0 \qquad l_i \geq l_o \tag{2.19}$$

where J is the normalization constant. The lowest dashed curve in Figure 2.5 is calculated by inserting the FENE parameters of Table 2.1, adopted in simulations, into Equation (2.19) multiplied by $4\pi l_i^2$. All of the curves are normalized. Comparison of the dashed curve with those resulting from MD simulations

indicates that the chain connectivity and the interaction with the surroundings substantially modify the distribution corresponding to isolated dumbbells. In general, the mean of the distribution function is shifted to larger values (i.e. from ~ 0.40 to 0.56) when the dumbbell belongs to a chain. This shift is presumably forced by the intrachain Lennard-Jones type potential, which requires second neighboring bonds to be separated by a distance of about $2^{1/6}\sigma^*$. This requirement leads to an average separation of $0.56\sigma^*$ between adjacent beads, in exact conformity with the MD results. At $\varepsilon_{bs} = 0.1\varepsilon^*$ the attractive potential of the environment is relatively weak to perturb the shape of the distribution function and a unimodal distribution centered about $0.56\sigma^*$ is observed. However, as the interaction between polymer and solvent is strengthened, the distribution is gradually modified into a bimodal shape. The peak to the right may be attributed to an increased separation between adjacent beads due to their simultaneous interaction with a single solvent molecule. These two beads that are strongly attracted to a single molecule are being forced apart by that interaction. However, when the corresponding bond assumes this increased length, its first neighbor has to assume a relatively contracted configuration in order to comply with the requirements of intrachain LJ potential between second neighbors, and hence the appearance of the accompanying peak at the left.

2.2.1.4. Translational Diffusivity as a Function of Solvent-Polymer Interaction. The diffusion coefficient is computed from the time decay of the chain center-of-mass velocity autocorrelation function $\langle \mathbf{v}_{cm}(t_o) \cdot \mathbf{v}_{cm}(t_o+t) \rangle$ according to the Green-Kubo relationship

$$D = \frac{1}{3}\int_0^\infty \langle \mathbf{v}_{cm}(t_o) \cdot \mathbf{v}_{cm}(t_o+t) \rangle \, dt \qquad (2.20)$$

where the angular brackets refer to the time average over initial times t_o. This average relies on the independence of the dynamics of stationary processes upon the time origin. The velocity $\mathbf{v}_{cm}(t)$ of the polymer mass center at a given

time t is found from the average of the instantaneous velocities of the N_b beads as $\mathbf{v}_{cm}(t) = N_b^{-1} \sum_i \mathbf{v}_i(t)$.

Figure 2.6 displays the time decay of the normalized velocity autocorrelation function of the polymer mass center, for various ε_{bs} values as indicated by the labels. The abscissa is the ratio t / Δt of the elapsed time t to the time step Δt of simulation. Full relaxation of $\mathbf{v}_{cm}(t)$ is observed to occur in the range of picoseconds. The weak statistical noise of the decay curves in the long time range is minimized by taking averages over simulations durations of about 10^5 time steps, for each ε_{bs}. The fluctuations occurring in the long time portions of the curves do not arise from any statistical uncertainties but are identically reproduced in repetitive runs, indicating the particular structure of the time decay of $< \mathbf{v}_{cm}(t_o) \cdot \mathbf{v}_{cm}(t_o+t) >$. Performing a set of short runs with different original configurations rather than a single run of long duration has proven to be a computationally efficient method of determining the decay curves.

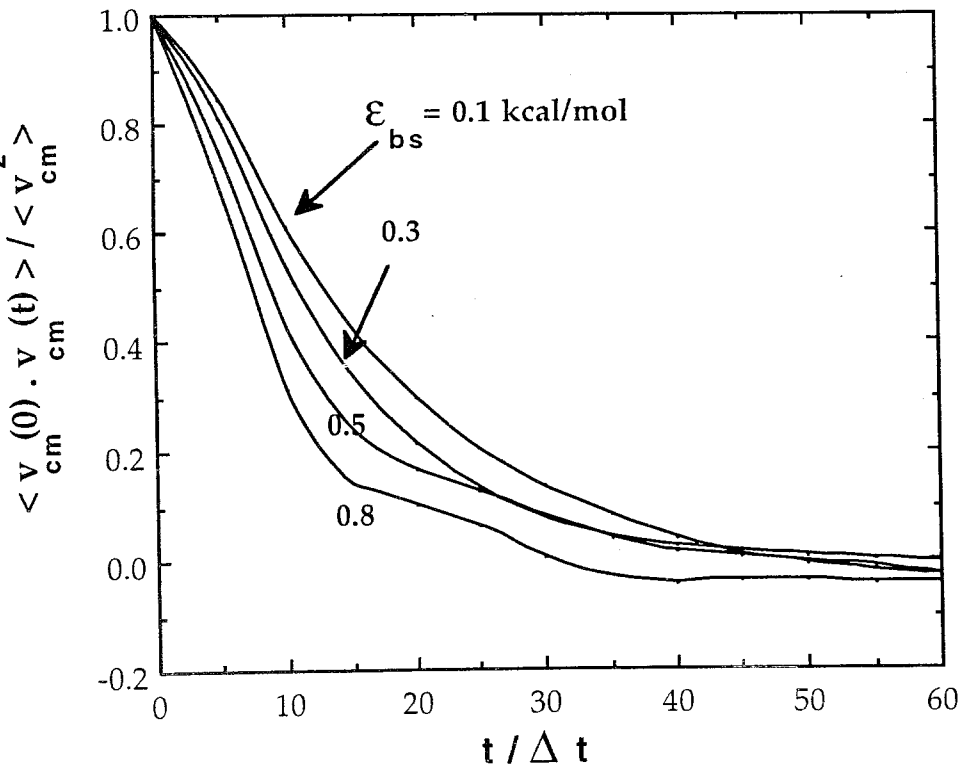


FIGURE 2.6. Time decay of the autocorrelation function $<\mathbf{v}_{cm}(0) \cdot \mathbf{v}_{cm}(t)> / <\mathbf{v}_{cm}^2>$ of the polymer mass center velocity $\mathbf{v}_{cm}$. The abscissa represents the reduced time t / Δt or the number of MD time steps. Curves are given for ε_{bs} = 0.1, 0.3, 0.5 and 0.8 kcal / mol.

The resulting translational diffusion coefficients from the decay curves are shown by the filled circles in Figure 2.7 as a function of ε_{bs}. Results are displayed in reduced units $D_r = D\ (m^* / \varepsilon^*)^{1/2}/\sigma^*$. The diffusion coefficients calculated from short runs of 10^4 time steps are plotted in the upper small diagram in the figure. This diagram gives an estimate of the scatter of D_r values resulting from successive runs of different original configurations. However, by taking the averages for each given ε_{bs}, the smooth linear decrease of D_r with increasing solvent quality emerges, which has been also pointed out by Luque et al.[14] for shorter chains (N_b = 12). The present study indicates that numerical uncertainties are eliminated only if sufficiently long simulations with a variety of different original configurations are performed.

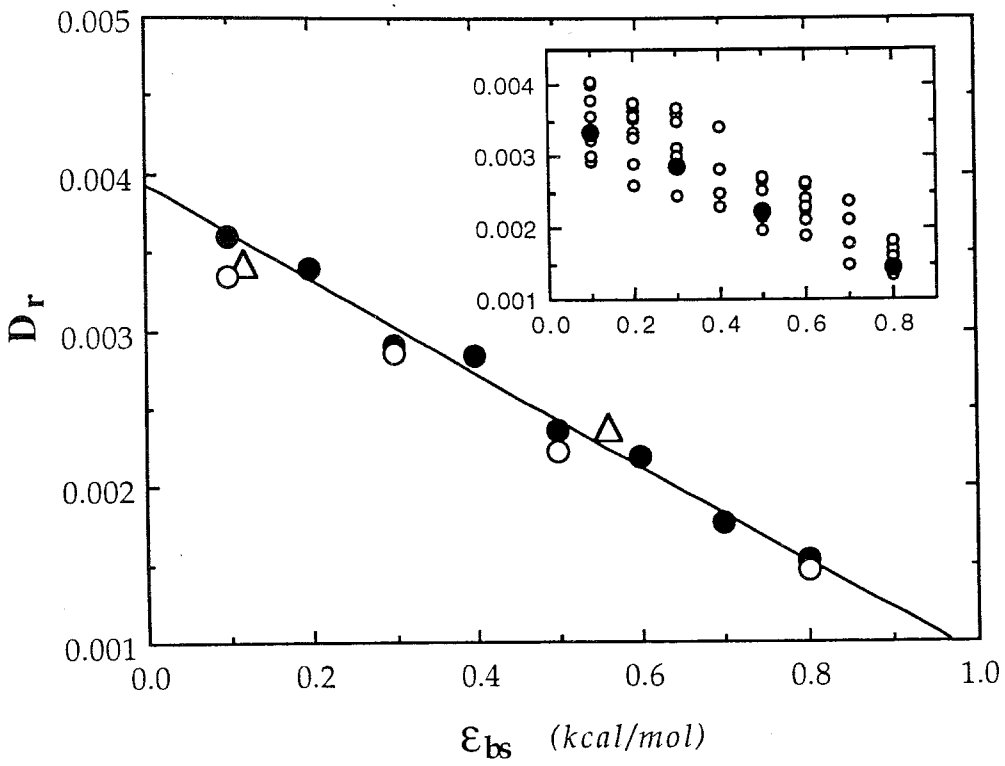


FIGURE 2.7. Dependence of reduced translational diffusivity D_r on solvent-polymer interaction energy ε_{bs}. The filled circles represent the averages of several runs of 10^4 time steps for each ε_{bs} value, which are illustrated in the upper small diagram. The empty triangles represent the results obtained by keeping R_g/L fixed as $R_g/L = 0.21$, and rescaling the size of the simulation box accordingly. The empty circles in the figure (and the filled circles in the upper small diagram) represent results obtained from Einstein plots. The best fitting line through the MD results is given by the equation $D_r = 3.931 \times 10^{-3} - 3.020 \times 10^{-3}\ \varepsilon_{bs}$ with a correlation coefficient of 0.988.

As a further check, the diffusion coefficients are computed from the long time slope of the mean-square displacement of the mass center as a function of time, using the Einstein relationship

$$6Dt = < [\, \mathbf{r}_{cm}(t_o + t) - \mathbf{r}_{cm}(t_o) \,]^2 > = < [\, \Delta \mathbf{r}_{cm}(t) \,]^2 > \qquad (2.21)$$

where the brackets refer to average over various initial times t_o. Figure 2.8 displays the change in $< [\, \Delta \mathbf{r}_{cm}(t) \,]^2 >$ as a function of time for ε_{bs} = 0.1, 0.5 and 0.8 kcal/mol. The resulting diffusion coefficients, calculated from the slopes according to Equation (2.21), are displayed by the empty circles in Figure 2.7, and by the filled circles in the upper small diagram of the same figure.

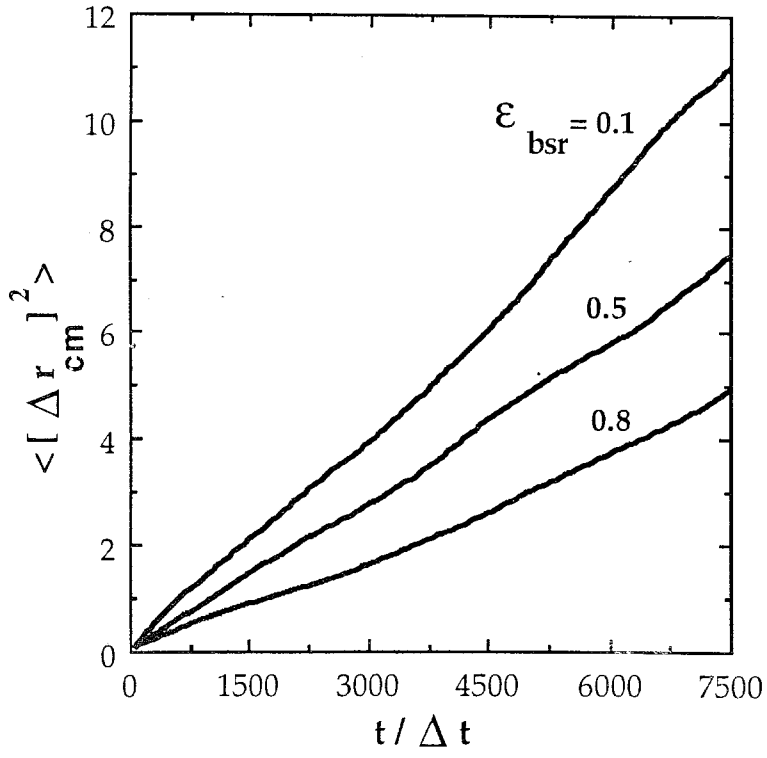


FIGURE 2.8. Mean-square displacement of the mass center $< [\, \Delta \mathbf{r}_{cm}(t) \,]^2 >$ as a function of time, for polymer-solvent interaction energies ε_{bsr} = 0.1, 0.5 and 0.8.

2.2.1.5. Translational Diffusivity in Relation to Chain Dimensions. The change in the diffusivity of the polymer with the solvent quality is an effect which should be rather attributed to the change in the equilibrium spatial distribution of chain atoms in response to their interaction with the surroundings. In fact, Figure 2.4 already confirmed that the interatomic separations are significantly perturbed by specific polymer-solvent interactions. These perturbations are directly reflected on overall chain dimensions, such as $< r^2 >$, and $< s^2 >$, as presented in Table 2.3, which in turn affects the chain diffusivity. Figure 2.9 displays the change in diffusivity with radius of gyration, resulting from runs performed for various ε_{bs}. The reduced diffusion coefficient is plotted against the reciprocal of the reduced mean-square radius of gyration, $<s^2>_r^{1/2} = R_g / \sigma^*$. The filled circles in Figure 2.9 represent the average values obtained from runs of 10^4 steps and organized on the basis of grids of size $\Delta<s^2>_r^{-1/2} = 0.05$ in the range $0.1 \leq \varepsilon_{bs} \leq 0.8$ kcal/mol. A smooth linear dependence of D_r on $1/R_g$ follows upon consideration of simulation durations $\geq 10^5$ steps, as shown in the figure. The best fitting line is drawn. It is noted that here the variations in radius of gyration do not arise from any change in chain length, but from the change in the environment for a fixed chain length. Yet, the translational diffusion coefficient exhibits a linear dependence on the radius of gyration in conformity with the implications of Zimm's theory[29] of chain dynamics.

It has been recently pointed out[15] that in standard MD simulations of finite-size systems subject to periodic boundary conditions, inasmuch as the infinite dilution hypothesis does not strictly hold, an *effective* hydrodynamic radius R_h, incorporating the interaction between the chain and all its images, need be computed for a correct comparison of the results with the Kirkwood theory. A practical approach to circumvent such periodic boundary artifacts in finite-size systems is asserted to perform the MD simulations at constant R_g/L ratios, when comparing the diffusivity of chains of various length.[15] This approach has recently proven useful for verifying the scaling law $D \sim N_b^{-\nu}$ or $D \sim R_g^{-1}$.[17] In view of these arguments, the influence of the finite-system-size effect on the presently obtained MD results might be questioned.

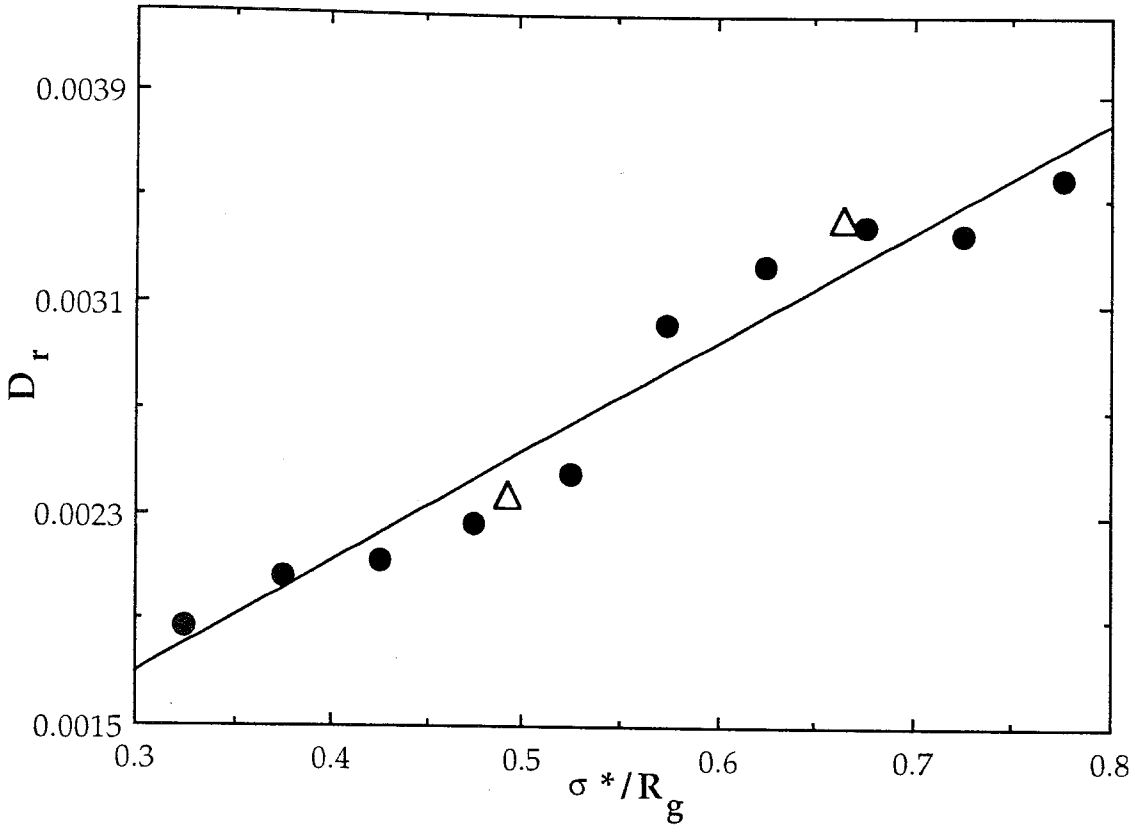


FIGURE 2.9. Dependence of D_r on the radius of gyration R_g, irrespective of the specific solvent-polymer interaction. The abscissa is the reciprocal of the reduced radius of gyration, written as $1/ < s^2 >_r^{1/2}$. Filled circles represent the averages over grids of $1/ < s^2 >_r^{1/2} = 0.05$. The empty triangles are found by keeping the ratio R_g/L fixed at 0.21. The best fitting line through the MD results is drawn. The corresponding equation is $D_r = 4.391 \times 10^{-4} + 4.190 \times 10^{-3} (\sigma^* / R_g)$ with a correlation coefficient of 0.950.

In the present simulations, although the chain length is kept fixed, the radius of gyration changes depending on solvent quality. All simulations being carried out in the same simulation box of size $L_r^3 \equiv (L / \sigma^*)^3 = 9.0^3$ accommodating an ensemble of 738 solvent molecules, an implicit increase in the ratio R_g/L is implemented with increasing ε_{bs}. Here, the ratio R_g/L is found to increase from 0.13 to 0.29, as ε_{bs} varies from 0.1 to 0.8 kcal/mol.This increase may be partly responsible for the observed decrease of D with increasing ε_{bs}. Likewise, the observed linear dependence of D on R_g^{-1} in Figure 2.9 might be biased with the same size effect, although these results are in satisfactory agreement with the predictions of the Zimm theory. As an inspection of this

effect, simulations have been repeated with different box sizes, keeping on the other hand the density of the system fixed at ρ_r = 1.054. As an illustrative example, let us consider the reference value of R_g/L = 0.21 corresponding to the case ε_{bs} = 0.4 kcal/mol = 0.4ε^*. Now, if one is interested in performing the simulations at fixed R_g/L and ρ_r, this implies a choice of L= 6.7 and $N_s = \rho_r L_r^3 - N_b$ = 294 for the case $\varepsilon_{bs} = 0.12\varepsilon^*$, for example. Performing the simulations for $\varepsilon_{bs} = 0.12\varepsilon^*$ with those parameters, yields the result displayed by the empty triangles at $D_r = 3.34 \times 10^{-3}$ in Figures 2.7 and 2.9. A weak decrease in D, almost ranging within statistical error limits of Green-Kubo integration method, is observable. The close agreement with the original results, which have been obtained with different R_g/L ratios, indicates that the size effect is practically inconsequential in the investigated R_g/L range and chain length. A further check would be to perform the simulations for the opposite case, say $\varepsilon_{bs} = 0.56\ \varepsilon^*$ adopting now a cubic box of edge L_r = 11.2, and N_s = 1470, and control whether the polymer in a larger simulation box enjoys higher mobility. The result is shown by the empty triangles located at $D_r = 2.28 \times 10^{-3}$ in Figures 2.7 and 2.9. Only a slight tendency for increased mobility in larger box is discernible in this latter simulation, which again supports the idea that the size effect is of secondary importance in the present simulation data. It could be desirable to check the extreme case of $\varepsilon_{bs} = 0.8\ \varepsilon^*$, as well. However, this task is not undertaken in view of (a) the high computational cost and/or less precise data collected within reasonable CPU time in this case, and (b) the fact that the results obtained for ε_{bs}=0.12 and 0.56 kcal/mol do not invoke a further search in that direction.

2.2.1.6. Orientational Autocorrelation Functions. The first and second orientational autocorrelation functions (OACF) for a unit vector **m** in motion are defined by the following expressions for $M_1(t)$ and $M_2(t)$, respectively.

$$M_1(t) = \langle \mathbf{m}(0) \cdot \mathbf{m}(t) \rangle \qquad (2.22)$$

$$M_2(t) = \frac{1}{2} \langle 3[\mathbf{m}(0) \cdot \mathbf{m}(t)]^2 - 1 \rangle \qquad (2.23)$$

Angular brackets in Equations (2.22) and (2.23) refer to the ensemble average over all configurational transitions, evaluated here on the basis of snapshots at various starting times. The first and second OACFs are normalized, i.e. start from the value of 1 at t = 0, because **m** is a unit vector. Representation of local dynamics in terms of OACFs is of interest, inasmuch as $M_1(t)$ is measured in dielectric relaxation experiments, and $M_2(t)$ in fluorescence anisotropy, NMR and ESR. In principle, the reorientational motion of a unit vector rigidly embedded in any direction on a chain segment can be extracted from the MD trajectories. Here, the first OACF is evaluated for bond vectors and chain segments of various sizes for the bead-spring model chain. In Section 2.2.2 on realistic model chains, both $M_1(t)$ and $M_2(t)$ will be analyzed.

The influence of solvent-polymer interaction on local motion of the chain may be surveyed from the comparative analysis of the first orientational autocorrelation functions obtained for bond vectors under various ε_{bs}. Figure 2.10 displays the time decay of the bond autocorrelations obtained from MD runs of 2×10^5 time steps duration, for each of the cases $\varepsilon_{bs} = 0.2$, 0.4 and 0.6. A significant change in local orientational mobility is observed, the motion becoming slower in the environment with larger ε_{bs}, as is conceivable in a system subject to stronger polymer-solvent attractive interactions.

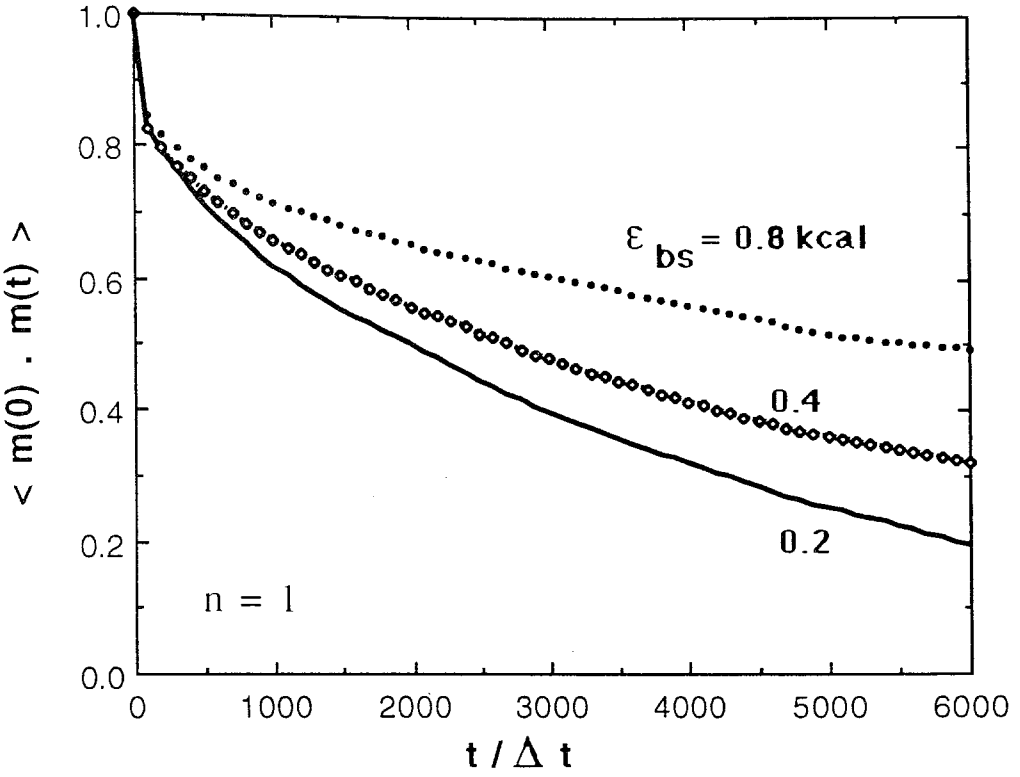


FIGURE 2.10. Influence of solvent-polymer interaction on the orientational relaxation of bond vectors. The time decays of < **m**(0) . **m** (t) > of unit vectors along backbone bonds are shown for ε_{bs} = 0.2, 0.4 and 0.6 kcal / mol.

Figure 2.11 displays the time decays for a series of unit vectors **m** along the end-to-end separation of segments of n bonds, for $1 \leq n \leq 12$. ε_{bs} is taken as 0.2 kcal/mol. The autocorrelation functions represent the averages over all internal chain segments, excluding the terminal four atoms at both ends of the chain, in order to eliminate the bias arising from end effects.

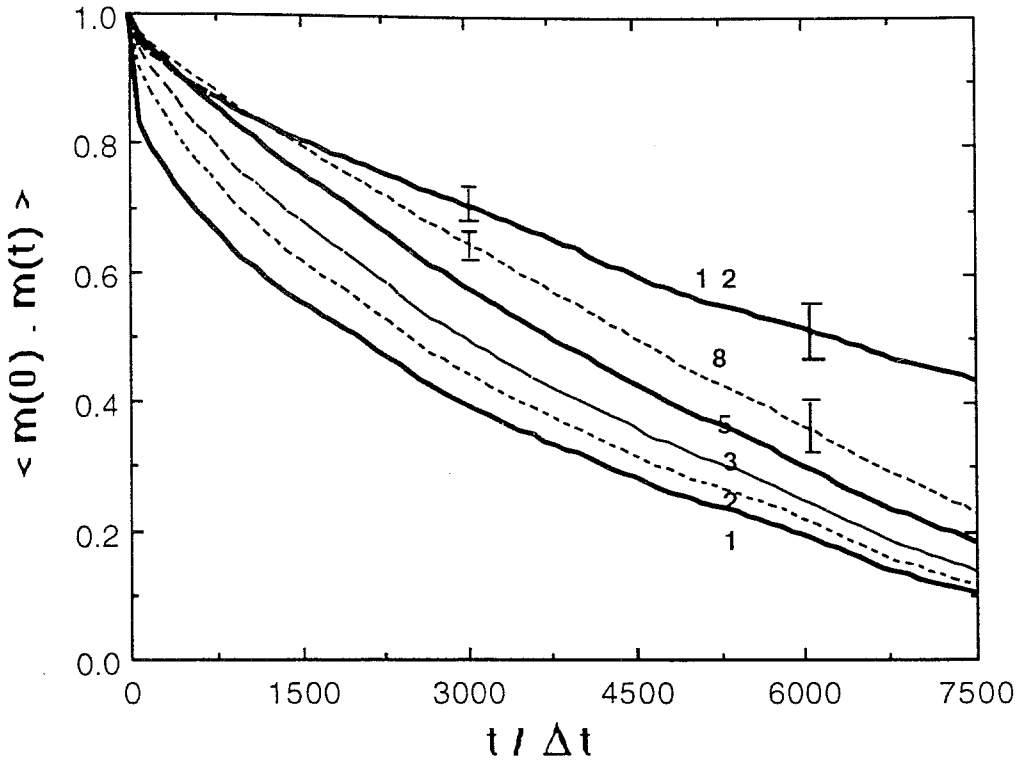


FIGURE 2.11. Time decay of orientational correlation functions < **m**(0) . **m** (t) > for unit vectors along the end-to-end separation of chain segments of n bonds, for n = 1, 2, 3, 5, 8 and 12, as indicated. Curves are obtained from MD simulations with ε_{bs} = 0.2 kcal/mol.The error bars in the cases of highest noise are shown.

Figure 2.11 shows that relaxation rates depend strongly on the size of the chain segment, being faster in the case of more localized motions i.e. shorter segments, as expected. The slowest relaxation process would occur in the case of $n = N_b - 1$. (The latter is not shown since it decays to about 0.8 only, within the time scale of the figure, and is subject to relatively large fluctuations.)

Simulations of 2 x 10^5 time steps (~1ns) have been performed to obtain the decay curves of Figure 2.11. Yet, the curves from independent runs exhibit considerable fluctuations, increasing with time and segment size. The statistical error due to finite time averaging in computer experiments has been analyzed by Zwanzig and Ailawadi.[30, 6] Accordingly, the MD results for a given normalized correlation function C(t) differ from the exact value R(t) by the equation

$$R(t) = C(t) \pm (2\tau / T)^{1/2} [1 - C(t)] \qquad (2.24)$$

where T is the total duration of simulation and τ is the relaxation time associated with the particular correlation function. Thus, the statistical error is proportional to $T^{-1/2}$ in general, and increases from 0 at t = 0 to $(2\tau / T)^{1/2}$ at long times. On the basis of the characteristic times τ of the orientational relaxation of segments of various sizes, which will be considered in the next paragraph, the highest statistical errors in Figure 2.11 are estimated to be R(t) - C(t) = 0.140 ± 0.004 for n ≥ 8 at t = 6000 Δt , and decreases to 0.081 ± 0.002 at t =3000 Δt. It is pointed out, however, that the error is reduced by a factor of $N^{1/2}$ when an average over N identical, separate particles/functions is possible.[30, 6] Although the precision introduced by this approach depends upon the range of correlations in the chain, the extra averaging over several internal chain segments of a given length, as presently performed, reduces further the statistical noise of the correlation functions. For example, the statistical errors at t = 6000 Δt decrease to 0.039 and 0.043 for n = 8 and 12, respectively, when the extra averaging over (N_b -n-7) internal segments of the chain is performed.These statistical uncertainties are indicated by the error bars in Figure 2.11.

2.2.1.7. <u>Dependence of Correlation Time on the Size of Chain Segment in Motion</u>. The dependence of the rate of orientational relaxation on the size of chain segments involved in local motions has been a subject of interest in previous studies. A correlation time τ may be assigned to each size segment either from the integral of the above time decay curves or from the inverse of their initial slope.[31] The second approach is presently undertaken, inasmuch as the correlation function for large n becomes statistically unreliable at long times. The resulting correlation times are plotted in Figure 2.12 for n < 8 and ε_{bs} = 0.1, 0.3, 0.5, 0.6 and 0.8 kcal/mol. An almost linear dependence between τ and n is observed in the logarithmic plot of Figure 2.12, supporting the presence of a power law of the form $\tau \sim n^a$ between the size and the orientational relaxation time of short chain segments in a given environment. For ε_{bs} = 0.1 kcal/mol, the best fitting line indicates an exponent of a = 1.00 for the

variation $\tau \sim n^a$. The exponent increases with increasing quality of the solvent and equates to 1.46 for ε_{bs} =0.8 kcal/mol.

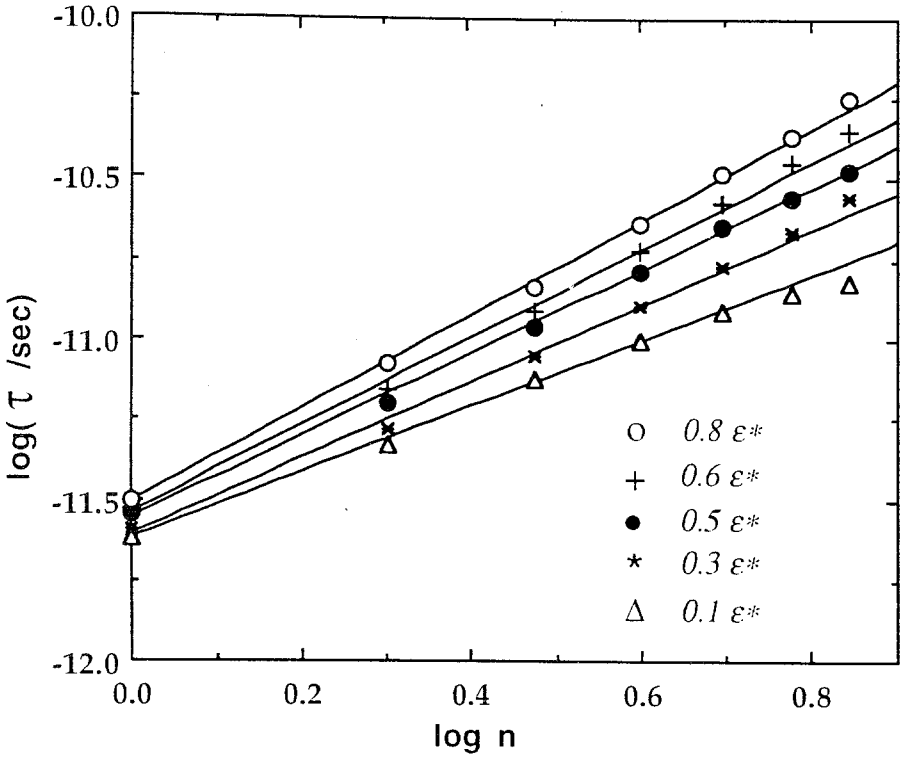


FIGURE 2.12. Dependence of orientational correlation times τ on the size of the chain segment involved in local motion. log(τ / sec) is plotted against log n where n is the number of bonds in the segment. The straight lines are drawn by least square fit.

In Rouse dynamics, the relaxation times scale as n^2, whereas the Zimm theory yields an exponent of 3v in the presence of hydrodynamic interactions, v assuming the values 0.5 and 0.6 in theta and good solvents, respectively. The Rouse-Zimm model is developed for long chains consisting of several Gaussian subchains, and gives a good description of low frequency motions in polymers. In the so-called sub-Rouse regime, on the other hand, which involves small-wavelength or intermediate frequency range motions, a weaker dependence on the size of the moving unit is predicted.[32-34] Both analytical[34] and numerical[32] analysis of the dispersion of normal relaxational modes for a segment of n = 16 coupled bonds indicate that, except for a few slowest modes, a

plateau value is approached for the frequency of all relaxational modes, and an almost linear dependence of terminal mode relaxation rate on n is observed up to n =150. This weaker dependence of relaxation rate on the size of the kinetic unit, compared to classical Rouse-Zimm model, has been thoroughly discussed in previous work.[34] Arguments on the reduced mobility of short constrained chain segments associated with their smaller number of degrees of freedom, go back to the original work of Kuhn.[23] The present simulations lend support to those arguments. The exponent *a* lies below the Rouse-Zimm values, in general. Furthermore, the present simulations indicate that the exponent, not only assumes lower values in high frequency regime, but also changes depending on the quality of the solvent. Interestingly enough, the lowest exponent is observed here in the case of the most compact chains (i.e. for ε_{bs} =0.1 kcal/mol), which is intrinsically subject to strongest intramolecular constraints. In analogy with the increase of the Zimm exponent from 1.5 to 1.8 with solvent quality, the exponent *a* shows a systematic increase as the chain assumes more expanded configurations.

2.2.2. Freely-Rotating Chain and Polyethylene Chain

Two competing effects prescribe the static and dynamic properties of the polymer chain in solution: (a) the intramolecular potentials and (b) the intermolecular nonbonded interactions between the polymer beads and surrounding solvent molecules. In the following, simulation results for the freely rotating chain (FRC) and the polyethylene chain (PE) will be examined in three different solvent environment with ε_{bs} = 0.2, 0.5, and 0.8 kcal / mol and the influence of these two competing effects on the properties of the polymers will be discussed. Simulations are carried out for 127,000 steps of total duration at each ε_{bs} value.

2.2.2.1. Overall Chain Dimensions. Table 2.4 summarizes the simulation results for the overall molecular chain dimensions $\langle r^2 \rangle$ and $\langle s^2 \rangle$, and the expansion coefficient α. Simulations carried out at different ε_{bs} = 0.2, 0.5, and 0.8 kcal / mol with the FRC and PE models will be named FRC-2, FRC-5, FRC-8, PE-2 , PE-5 and PE-8, as indicated in the table.

TABLE 2.4. Mean-square chain dimensions and expansion coefficients α as a function of solvent quality

Simulation	ε_{bsr}	$\langle r^2 \rangle$ (Å2)	$\langle s^2 \rangle$ (Å2)	α
FRC-2	0.2	39.65 ± 16.48	11.08 ± 0.44	0.544
FRC-5	0.5	70.64 ± 26.83	15.38 ± 1.17	0.726
FRC-8	0.8	291.24 ± 82.60	39.58 ± 5.54	1.475
PE-2	0.2	35.65 ± 18.95	12.17 ± 0.87	0.351
PE-5	0.5	214.98 ± 33.02	29.96 ± 6.49	0.861
PE-8	0.8	495.14 ± 143.76	59.34 ± 6.49	1.307

In the calculation of α by Equation (2.14), the unperturbed mean-square end-to-end separation $\langle r^2 \rangle_o$ is given by[2]

$$< r^2 >_o = n\, l_o^2 \left[(1 + \alpha)(1 - \alpha)^{-1} - (2\alpha / n)(1 - \alpha^n)(1 - \alpha)^{-2} \right] \qquad (2.25)$$

for the FRC model of $n = N_b - 1$ bonds, with $\alpha \equiv \cos \theta_o = 1/3$, and by

$$< r^2 >_o = n l_o^2 \left[(\mathbf{E} + <\mathbf{T}>)(\mathbf{E} - <\mathbf{T}>)^{-1} - (2 <\mathbf{T}> / n)(\mathbf{E} - <\mathbf{T}>^n)(\mathbf{E} - <\mathbf{T}>)^{-2} \right]_{11} \qquad (2.26)$$

for the PE model, using the rotational isomeric state formalism. Here **E** is the identity matrix of order 3, **T** is the transformation matrix operating on conventional bond-based frames,[2] the angular brackets refer to the averaging of the elements of **T** using the potential given in Equation (2.10), and the subscript 11 refer to the element 11 of the matrix in square brackets. The term $<\mathbf{T}>^n$ is conveniently calculated by transformation of $<\mathbf{T}>$ into the diagonal matrix of its eigenvalues. Application of Equations (2.25) and (2.26) yield $<r^2>_o / \sigma^{*2} = 18.36$ and 39.76 for FRC and PE models, respectively. The expansion coefficients calculated on the basis of these values, indicate the occurrence of poor solvent conditions when $\varepsilon_{bs} = 0.2$ kcal/mol, approximately *theta* conditions at $\varepsilon_{bs} = 0.5$ kcal/mol, and finally relatively good solvent environment with $\varepsilon_{bs} = 0.8$ kcal/mol. Thus, the present choice of polymer-solvent interaction energy parameters allows for the examination of the dynamics of both model chains under three distinct states, i.e. contracted, unperturbed, and expanded configurations. It should be noted that these specific conditions result from the counterplay between intermolecular and intrachain potentials, adopted in the present study, and should not be uniquely attributed to the choice of the parameter ε_{bs}. In fact, adoption of the same polymer-solvent interaction parameters is observed to lead to good solvent conditions for the bead-spring model chain in Section 2.2.1.1., where the FENE potential between bonded units, and the shifted-force potential among beads separated by *two* or more bonds are used. The latter interaction, in particular, places severe constraints as to the spatial location of near neighboring (separated by 2 or 3 bonds) beads, forcing them to be sufficiently

apart from each other, and consequently leading to more expanded chain configurations.

2.2.2.2. Translational Diffusivity as a Function of Solvent-Polymer Interaction. Figure 2.13 displays the reduced translational diffusivities ($D_r \equiv D (m^* / \varepsilon^*)^{1/2} / \sigma^*$) of the two model chains as a function of ε_{bs}. The circles are the simulation results, obtained from the Einstein relationship, given by Equation (2.21). In this figure, the FRC model exhibits larger diffusivity compared to PE, as might be expected from the increased intrinsic flexibility of the former. The dependence of D on solvent quality is more significant, however. The slower diffusion of the chains with increasing ε_{bs} is mainly attributed to the expansion of the chain as the interactions with the solvent get more favorable. Similar trends, i.e. an inverse linear relationship between D and ε_{bs}, have been obtained for the bead-spring model chain in Section 2.2.1.4, in conformity with the present results. Careful examination of the possible changes in D due to artificial increases in hydrodynamic radius (resulting from the physically unrealistic interaction of polymer beads with image beads), has indicated that this periodic boundary artifact is negligibly small for the bead-spring model chains of 30 beads in the adopted simulation boxes. Except for PE-8, the chains investigated in the present section assume even less expanded configurations, compared to those of bead-spring model chains, and such periodic boundary artifacts should be further reduced. Thus, the linear decrease of D with ε_{bs} appears as a general property change associated with solvent quality, irrespective of the particular structural characteristics of the polymer.

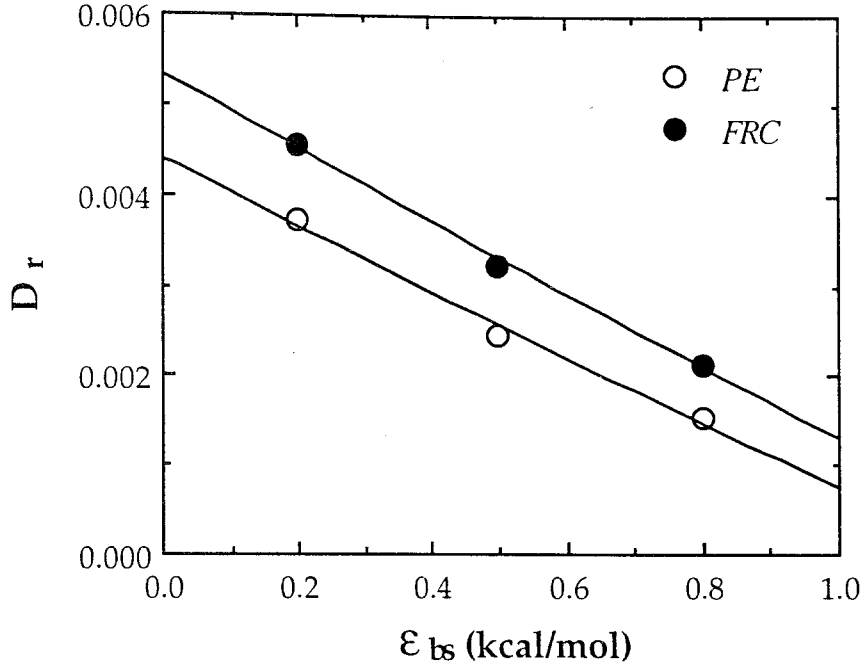


FIGURE 2.13. Dependence of polymer translational diffusivity on polymer-solvent interaction. Reduced diffusion coefficient D_r is plotted against polymer-solvent interaction energy ε_{bs}. Filled and empty circles represent the average D_r values for the respective FRC and PE models, calculated by the Einstein formula.

2.2.2.3. Pair-Correlation Functions. Figure 2.14 illustrates polymer-solvent pair correlation function $g_{bs}(r)$ for the model chains FRC-5 and PE-5 (See Section 2.2.1.2 for definitions). A plateau region, which is not observed in bead-spring chains at the same ε_{bs} value, appears between the first two peaks when the polymer chain is constrained by bond bending and torsional potentials. The larger $g_{bs}(r)$ values of PE at short separations indicate its more expanded configuration relative to the FRC at the same ε_{bs}.

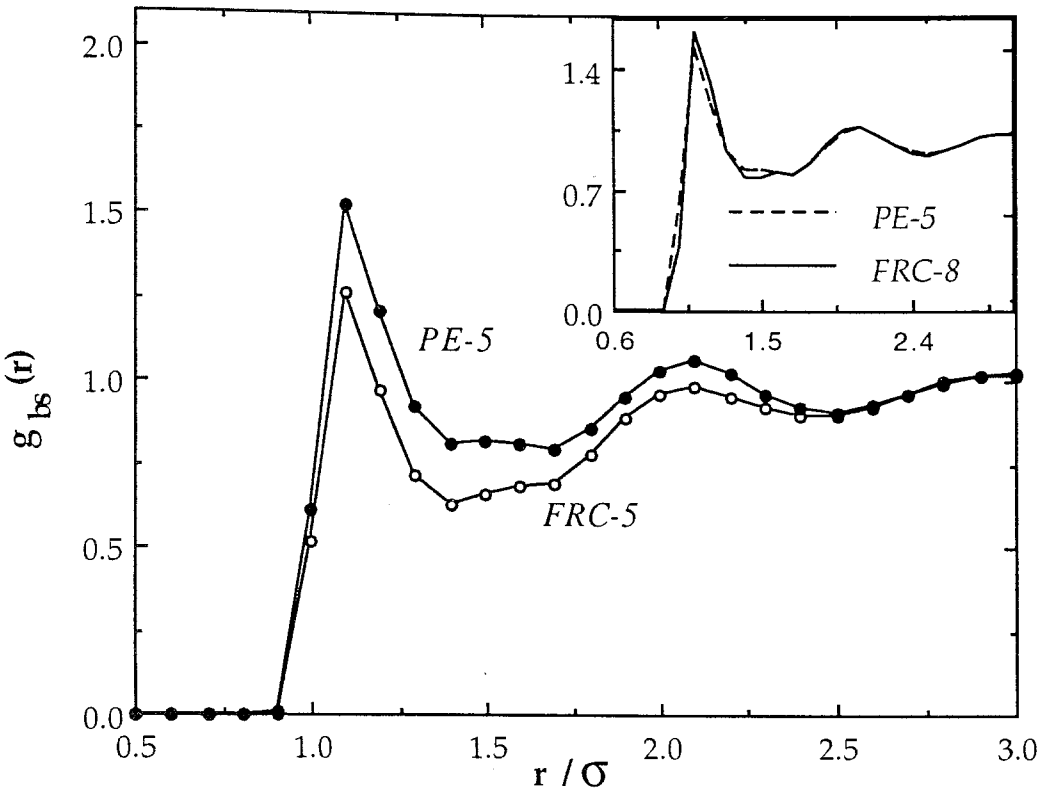


FIGURE 2.14. Radial distribution function $g_{bs}(r)$ for polymer bead-solvent pairs as a function of the reduced separation r / σ. Empty and filled circles represent $g_{bs}(r)$ for FRC-5 and PE-5, respectively. The upper small diagram compares $g_{bs}(r)$ for FRC-8 and PE-5.

Pair-correlation functions analyzed for various ε_{bs} (not shown) indicated the following features: The oscillations of $g_{bs}(r)$ in the FRC model are found to become more pronounced as the quality of the solvent is improved, in conformity with prior observations on bead-spring model chains. Although a similar trend is observed with PE in poor solvent, an asymptotic behavior is reached at $\varepsilon_{bs} = 0.5$ kcal/mol, and the shape of the radial distribution function does not change upon further increase in solvent quality. More interestingly, the pair correlation function of FRC-8 closely resembles that of PE-5, as illustrated in the insertion of Figure 2.14. FRC-5 and PE-2 also exhibit similar patterns. This suggests that in the presence of relatively favorable interactions with the surroundings, leading to lower energy or more stable conformations on a local scale, the decreased flexibility of FRC is manifested by the same behavior as that of a polyethylenelike chain which, itself, is locally stiff due to torsional potentials hindering bond

rotations, but enjoys some freedom due to weaker interaction with the medium. Thus, irrespective of the origin -intramolecular or intermolecular- of local reduced flexibility, the same functional form is obtained in two distinct cases for the pair correlation function.

Figure 2.15 displays the intrachain pair correlation functions $g_{bb}(r)$ for FRC and PE models at ε_{bs}=0.5 kcal/mol, computed on the basis of non-bonded beads separated by at least two bonds along the polymer backbone.The first sharp peak appears at about $r / \sigma^* = 0.9$ on both curves. This corresponds to the reduced distance between atoms C_i and C_{i+2} which are constrained around the tetrahedral angle in both cases. The particular choice of parameters in the torsional potential $V_\phi(\phi_i)$ restricts the dihedral angle values around 0° and ±120° and thereby constrains the separation of atoms C_i and C_{i+3}. In fact, the second peak at $r / \sigma^* = 1.43$ on the curve of the PE-5 corresponds to the *trans* placement of the bond (C_{i+1} -C_{i+2}). Likewise, two consecutive bonds at the *trans* state place backbone atoms separated by 4 bonds at $r / \sigma^* = 1.86$, which coincides with the third peak observed in PE-5. A similar pattern can be observed on the $g_{bb}(r)$ of FRC-8 as shown on the inlet of Figure 2.15. Thus, in parallel with the relative behavior of PE-5 and FRC-8, displayed in the inset of Figure 2.14, the existence of more favorable polymer-solvent interactions is manifested as if a rotational barrier were imposed on the dihedral angles of FRC. For comparative purposes, $g_{bb}(r)$ of FRC-2 is also given in the insertion of Figure 2.15.

These observations stipulate the competing contributions of inter- and intramolecular effects in prescribing the particular shapes of pair-correlation functions. And attempts to attribute a particular shape of $g_{ij}(r)$ to entirely intrachain constraints or, alternatively, to environmental effects might fail, if one considers exclusively $g_{ij}(r)$ as a measure of characterizing chain behavior.

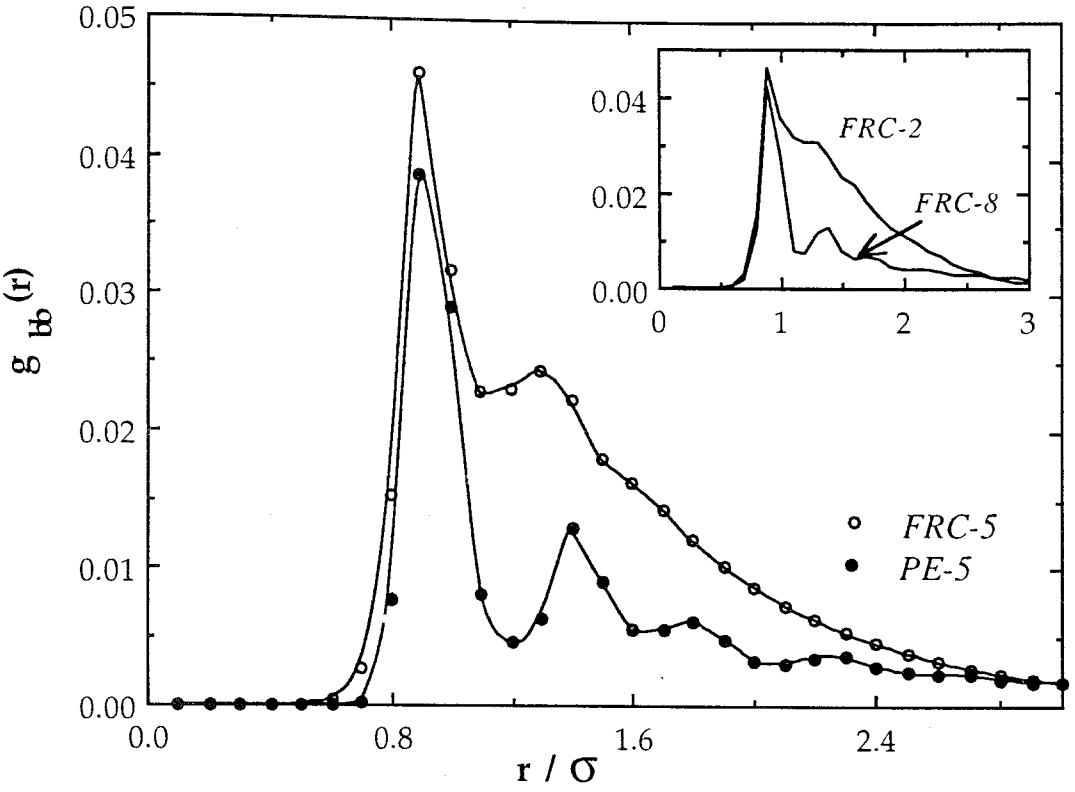


FIGURE 2.15. Radial distribution function $g_{bb}(r)$ for pairs of non-bonded units along the polymer chain as a function of the reduced separation r / σ. Empty and filled circles represent $g_{bb}(r)$ for FRC-5 and PE-5, respectively. The upper small diagram displays $g_{bb}(r)$ for FRC-2 and FRC-8.

2.2.2.4. Static Scattering Functions. For a chain composed of N_b scattering units, the static scattering function S(q), also called the particle scattering function, is defined as

$$S(q) = \frac{1}{N_b^2} < \sum_{j=1}^{N_b} \sum_{k=1}^{N_b} \exp [- i\mathbf{q} \cdot (\mathbf{r}_j - \mathbf{r}_k)] > \qquad (2.27)$$

where $\mathbf{r}_j$ and $\mathbf{r}_k$ are the respective position vectors of particles j and k, $\mathbf{q}$ is the scattering wavevector with magnitude $q = (4\pi / \lambda) \sin(\theta / 2)$, depending on the wavelength λ of the radiation and the scattering angle θ. The angular brackets refer to a space and ensemble average over all equilibrium configurations of the chain. Averaging over all orientations of the vector $\mathbf{q}$ in space, and taking the real part of the exponential function yields the expression

$$S(q) = \frac{1}{N_b^2} < \sum_{j=1}^{N_b} \sum_{k=1}^{N_b} \frac{\sin qr_{jk}}{qr_{jk}} > \qquad (2.28)$$

where $r_{jk} = | \mathbf{r}_j - \mathbf{r}_k |$. The ensemble average represented by angular brackets is evaluated here as a time average over several snapshots in a given run.

The static scattering functions for the investigated model chains in different solvents are presented in the form of Kratky plots in Figures 2.16(a) and 2.16(b). Figure 2.16(a) displays the MD results for FRC-2, FRC-5 and FRC-8. The oscillations are observed to damp out with increasing strength of polymer-solvent interactions. In Figure 2.16(b), PE model chains exhibit a similar behavior at ε_{bs} =0.2 and 0.5 kcal/mol, but the shape of the scattering function does not undergo further change upon increase of the polymer-solvent interaction to ε_{bs} = 0.8 kcal/mol. This behavior, in fact, is in conformity with the previous results on pair correlation functions $g_{bb}(r)$ and $g_{bs}(r)$. Moreover, the Kratky plot of FRC-8 is almost coincident with the plots for the PE-5 or PE-8. The fact that the local behavior of FRC in the presence of stronger interactions with the surroundings resembles that of a chain with constrained rotational angles in a less interactive medium, has been already commented on in the preceding section.

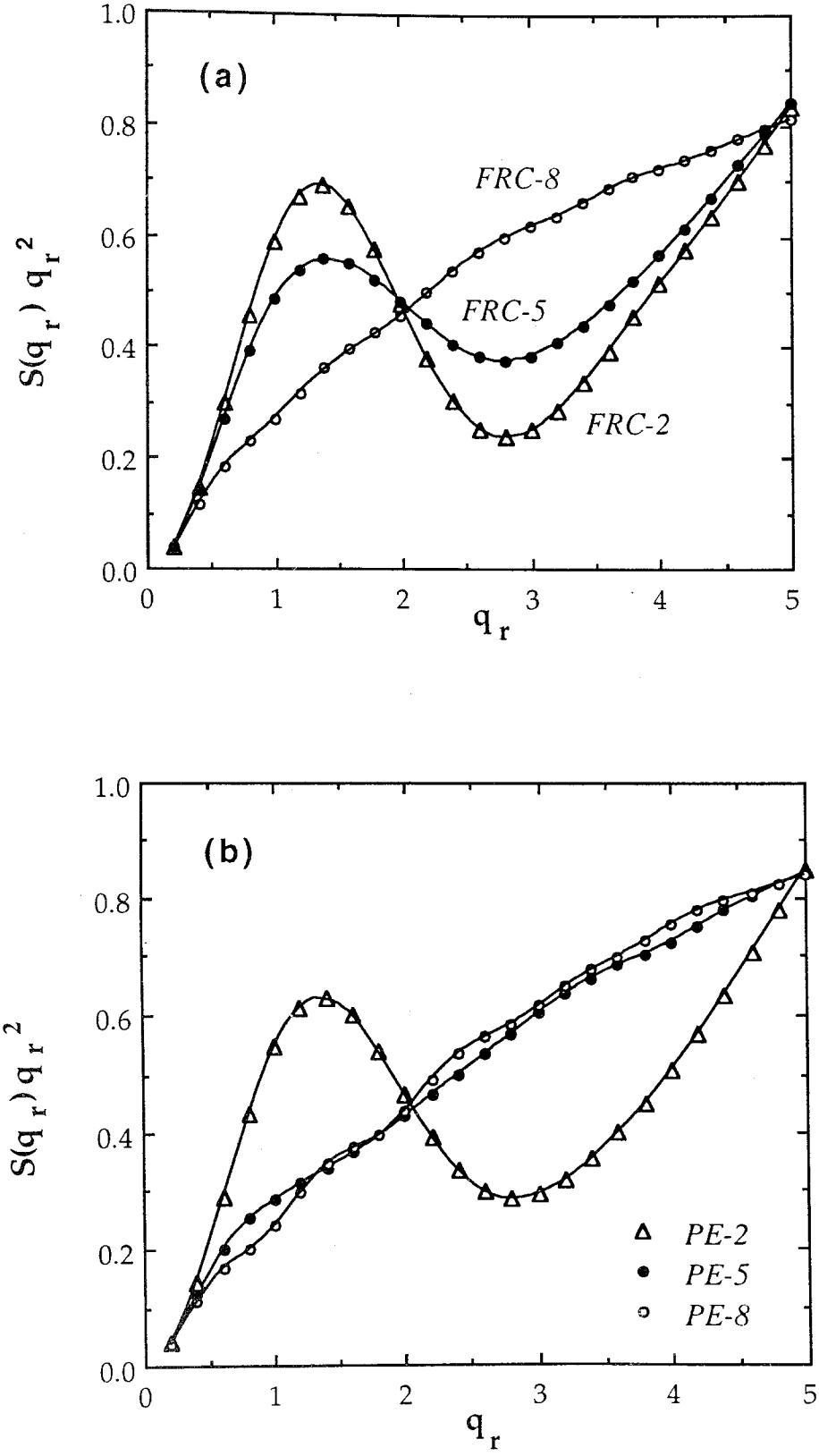


FIGURE 2.16. Kratky plots of static scattering function, as $S(q_r)q_r^2$ against $q_r = q\sigma^*$ (a) for FRC-2, FRC-5 and FRC-8. (b) for PE-2, PE-5 and PE-8.

A wider range of Kratky plot, covering highly localized domains up to $q_r \equiv q\,\sigma^* = 10$, is displayed in Figure 2.17. The curves for ε_{bs} = 0.5 kcal/mol are plotted only, as indicated by the labels. At large q_r values, domains of the size of small amplitude fluctuations about minimum energy states are observed, and gradually differences between FRC and PE model chains vanish. This may be explained by the fact that in this range, mostly stretching and bending potentials, which are taken to be the same for both chains, are operative.Torsional motions of backbone bonds occur through vibrational motions only in this q-range, and no contribution from large jumps between rotational isomeric states is observable.

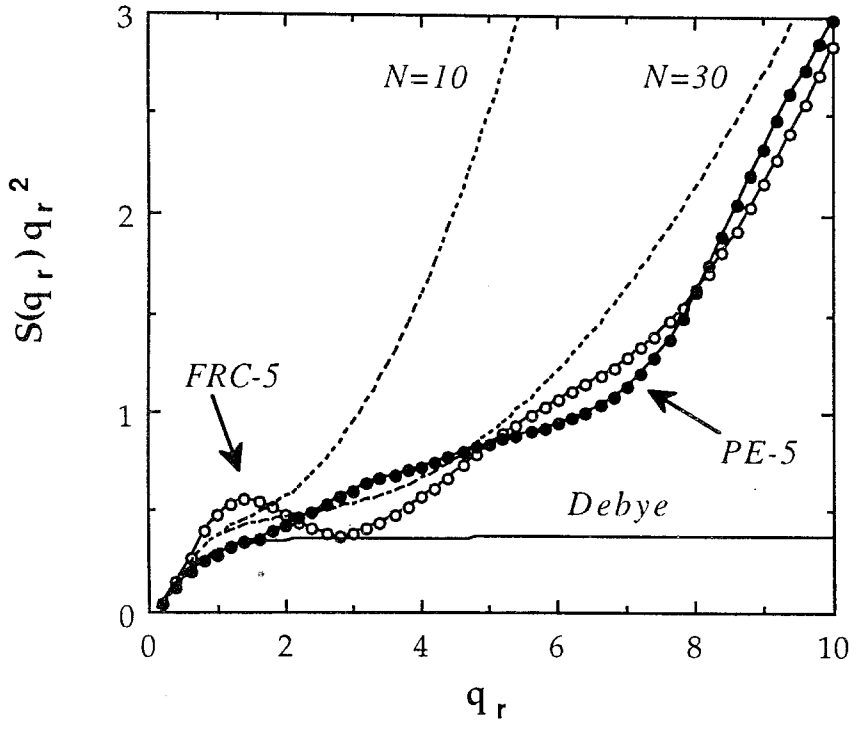


FIGURE 2.17. Representation of Kratky plots on a wider q_r (= 0-10) range. The curves for FRC-5 (empty circles) and PE-5 (filled circles) are shown. $S(q_r)q_r^2$ vs q_r curves theoretically predicted for freely jointed Gaussian chains of N = 10 and 30 units, and the Debye limit are also drawn.

For comparative purposes, theoretical curves are plotted in Figure 2.17 using the expression

$$S(q) = N^{-1} \{1 + 2(e^{\alpha} - 1)^{-1} [1 - N^{-1} (1 - e^{-\alpha})^{-1} (1 - e^{-\alpha N})] \} \qquad (2.29)$$

for freely jointed Gaussian chains containing N segments.[35] Here the variable α is given by $\alpha = q^2a^2/6$, where a is the statistical segment length evaluated from $R_g^2 = Na^2/6$. The curves are drawn for chains of N = 10 and 30 units, adopting R_g of PE-5 from simulations. S(q) for the Debye limit [35] with $\kappa^2 = \alpha N$ is also shown in the figure. The latter is evaluated from

$$S(q) = (2 / \kappa^4) (e^{-\kappa^2} + \kappa^2 - 1) \qquad (2.30)$$

These results indicate that Equation (2.29), with suitable choice of parameters, may provide a qualitative representation of the q-dependence of the static scattering function in the high q regime. Yet, a rigorous theoretical interpretation or reproduction of the detailed form of the scattering function, including its dependence on intrachain and solvent effects, is not possible with current analytical approaches, and MD simulations appear to be of use in estimating best fitting parameters to be conveniently adopted in existing closed form expressions.[36, 35]

2.2.2.5. Intermediate Scattering Functions. The intermediate scattering function S(q,t) given by

$$S(q,t) = \frac{1}{Nb^2} < \sum_{j=1}^{Nb} \sum_{k=1}^{Nb} \exp \{ - i\mathbf{q} . [\mathbf{r}_j(0) - \mathbf{r}_k(t)] \} > \qquad (2.31)$$

is transformed into a form similar to Equation (2.28), by replacing r_{jk} with $r_{jk} = |\mathbf{r}_j(0) - \mathbf{r}_k(t)|$. The angular brackets here refer to an average over different initial times. The dynamic scattering behavior of a polymer chain can be divided into three different regimes. [35, 37] At low frequencies, identified by $qR_g \leq 1$, the diffusion of the whole polymer chain is observed. At large q values ($ql_o \geq 1$), local motions dominate. And in the intermediate region defined by $qR_g \geq 1$ and $ql_o \leq 1$, the coupled conformational changes chain segments operate. The time decays of S(q,t) for FRC and PE in different solvents are plotted as $\ln[S(q_r, t/\Delta t)/S(q_r, 0)]$ vs $t/\Delta t$ in Figure 2.18, using $q_r = 4$ which corresponds to the regime of local motions. Clearly, at t=0, S(q, 0) is simply the static scattering function, S(q). The increased flexibility of FRC on a local scale, compared to PE, is verified from the faster decay of the curves for FRC. On the other hand, decay curves for the intermediate region (not shown), using $q_r = 1$ for instance, indicate that the differences in the dynamic behavior of the respective chains diminish when the chains are observed from a sufficiently wide time window.

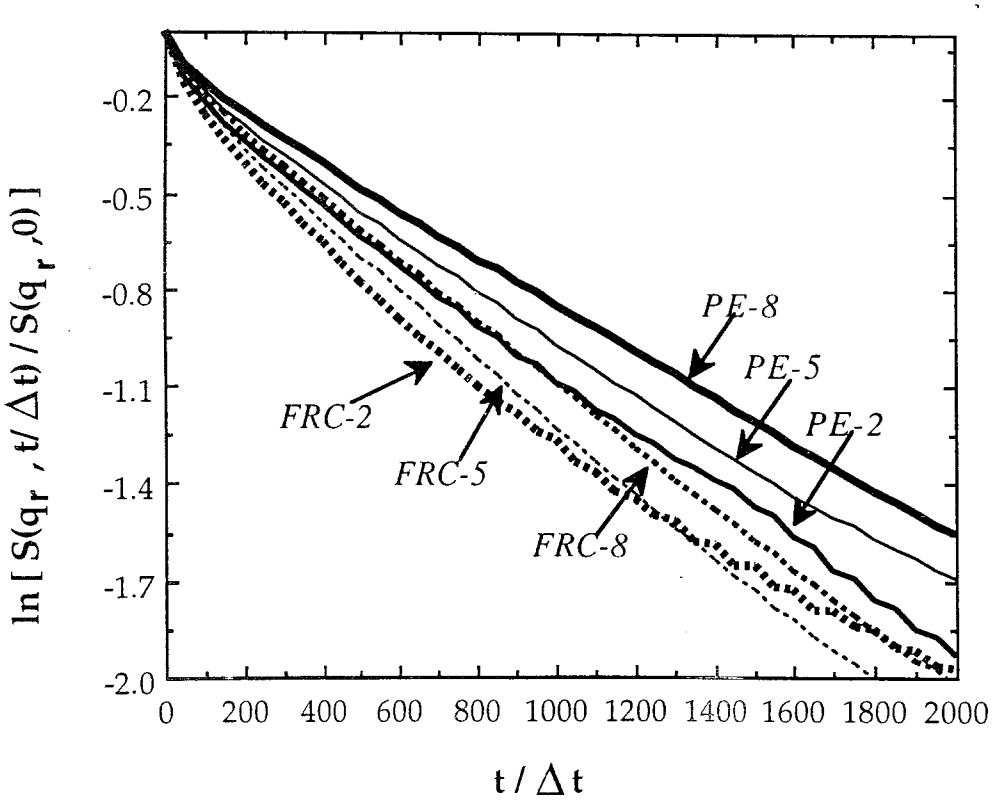


FIGURE 2.18. Time decay of $\ln [S(q_r, t/\Delta t) / S(q_r, 0))$ at $q_r = q\,\sigma^* = 4.0$ showing local dynamics for FRC-2, 5, 8 and PE-2, 5, 8.

The first cumulant, Ω, which is defined as

$$\Omega = - \lim \left[\frac{\partial \ln [S(q,t) / S(q,0)]}{\partial t} \right]_{t \to 0} \qquad (2.32)$$

permits the representation of the dispersion of the intermediate scattering function over a wide frequency range.[35, 37] Figure 2.19 displays the q-dependence of Ω_r/q_r^2 for FRC-5, where Ω_r is the reduced cumulant in which t and q are replaced by $t / \Delta t$ and q_r, respectively. The evaluation of first cumulants from the time decay of intermediate scattering functions is a tedious and somewhat arbitrary task, inasmuch as the results are quite sensitive to the time interval selected for the estimation of Ω. For this reason, two different methods have been adopted for determining Ω_r: (a) the initial slopes of $\ln[S(q_r, t / \Delta t) /S(q_r,0)]$ decay curves, and (b) the coefficient of the first order term in fifth order polynomials fitted by least squares to $\ln[S(q_r, t / \Delta t) /S(q_r,0)]$ curves. The corresponding results are shown by the empty and filled circles, respectively. The method (b) might slightly overestimate the rates, as apparent from the vertical shift between the results from the two methods. Yet, the same *qualitative* trend is followed in both cases, lending support to the generality of the observed dependence of first cumulants on q. The three different q regimes can be tentatively distinguished on the graph, conforming with the particular shape predicted by Akcasu et al.[35] Comparison of the Ω_r/q_r^2 against q_r curves drawn for FRC-5 and PE-5 (not shown)have indicated that the structural differences between PE and FRC do not appear explicitly on logarithmic plot. But, the verification of the specific dependence of Ω_r/q_r^2 on q_r , which, to our knowledge, has not been achieved in previous MD studies, is interesting enough. Furthermore, the oscillations in the curves in the higher q-regime might be also consistent with previous theoretical and experimental results.[38]

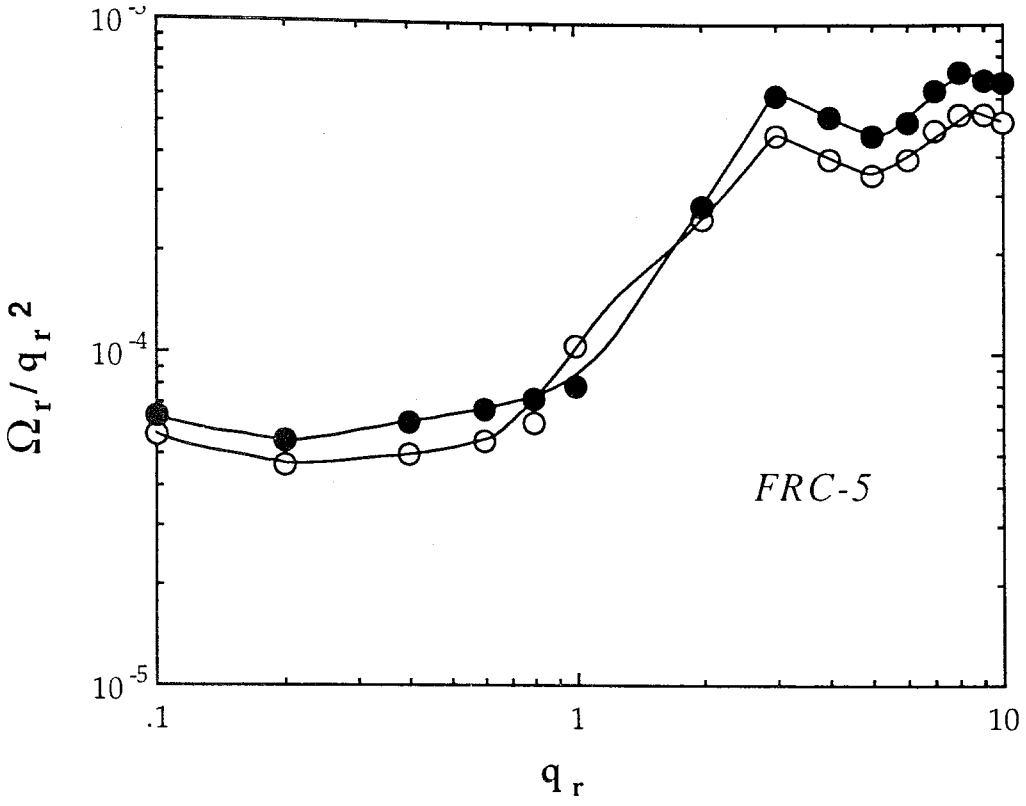


FIGURE 2.19. Dependence of first cumulant on q, represented in reduced form $\Omega_r (q_r, t/\Delta t) / q_r^2$ vs q_r for FRC-5. Ω_r values are determined from: (a) the initial slopes of $\ln[S(q_r, t / \Delta t) / S(q_r,0)]$ decay curves (empty circles), and (b) the coefficient of the first order term in fifth order polynomials fitted by least squares to $\ln[S(q_r, t / \Delta t) / S(q_r,0)]$ curves (filled circles).

2.2.2.6. Orientational Autocorrelation Functions. Here, the first and second OACFs, given by the respective Equations (2.22) and (2.23), are evaluated for the following two types of unit vectors: (a) $\mathbf{m}_{||}$, representing unit vectors along the bonds of the chain and (b) $\mathbf{m}_{\perp}$, the out-of-plane vectors perpendicular to the local planes defined by successive pairs of bonds. The autocorrelations are computed from the average over unit vectors appended to all atoms along the chain, excluding the terminal four atoms at both ends of the chain. Figures 2.20(a) and 2.20(b) give the first and second OACFs, respectively, obtained for the vectors $\mathbf{m}_{||}$ in FRC and PE models at various ε_{bs}. The relative importance of intra- and intermolecular factors in controlling local chain dynamics can be reckoned from these graphs. FRC and PE exhibit quite distinct relaxation time scales in both figures, indicating that intramolecular constraints are of primary importance in determining the rate of local relaxation

processes. Intermolecular forces due to interactions with the solvent molecules affect the rates of OACF decays in a given type of chain towards the long time portions of the curves, in particular. PE appears to be more strongly affected by solvent effect, compared to FRC. Yet, the solvent interaction is not sufficient to compensate for effects due to specific chain structure: In contrast to the equilibrium characteristics of the two types of chains in different solvents, revealing for example close similarities between PE-5 and FRC-8 as discussed above, the dynamic behaviors of PE and FRC are quite distinct and this difference is preserved irrespective of the solvent quality. This suggests that intramolecular conformational characteristics play a predominant role in prescribing local chain dynamics compared to solvent effects.

It should be noted that in the absence of bending and torsional constraints, the local dynamics of the bead-spring model chains in solution is found to be more significantly affected by the specific solvent effect in Section 2.2.1.6. Inclusion of intramolecular conformational potentials decreases the apparent effect of solvent. This is understandable, since the solvent effect now competes with effects from intrachain conformational potentials.

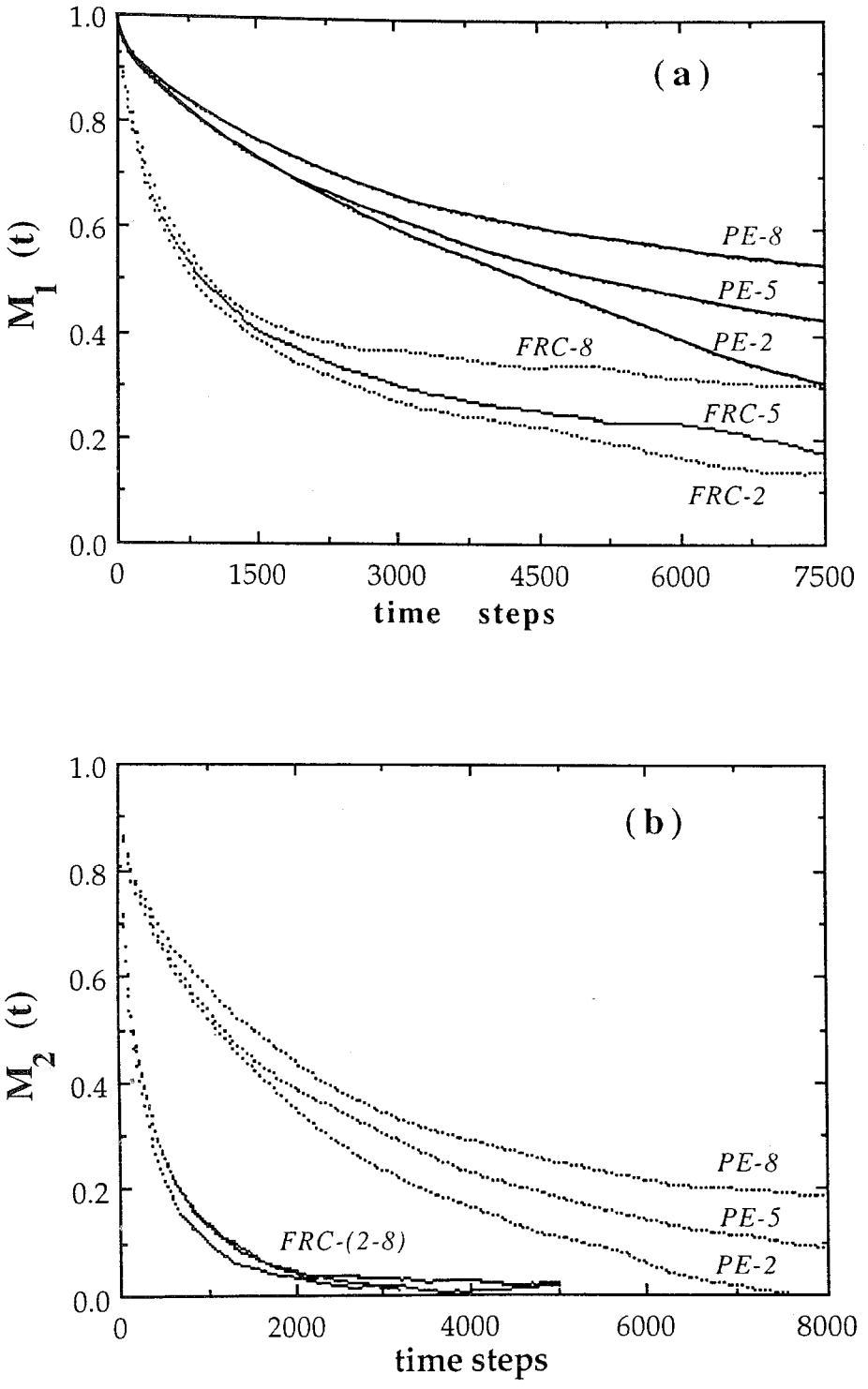


FIGURE 2.20. Time decay of (a) the first orientational autocorrelation function $M_1(t)$ and (b) the second orientational autocorrelation function $M_2(t)$ for unit vectors $\mathbf{m}_{||}$ along the bonds, for FRC-2, 5, 8 and PE-2, 5, 8.

2.2.2.7. Streched Exponetial Forms of OACFs. The stretched exponential expression

$$M_i(t) = \exp \{ - (t / \tau_i)^{\beta} \} \qquad (2.33)$$

with $0 \leq \beta \leq 1$ has found widespread use in literature for representing data on $M_1(t)$ and/or $M_2(t)$.[39] τ_i is the characteristic relaxation time at which $M_i(t)$ decays to 1/e, with i = 1, 2. Equation (2.33) may be rewritten as

$$\log [-\ln M_i(t)] = \beta \log t - \beta \log \tau_i \qquad (2.34)$$

which permits easier graphical evaluation of β and τ_i. Such log [-ln $M_i(t)$] vs log t plots are displayed in Figures 2.21(a) and 2.21(b), for i = 1 and 2, respectively. Results for FRC-2, FRC-8, PE-2 and PE-8 are shown in the figures, as indicated, those for FRC-5 and PE-5 being not displayed for clarity. Again, the specific intramolecular characteristics of the chain are more important in prescribing the time scale of motion, rather than the solvent effect. This is obvious from the time scale difference between FRC and PE, which is definitely large compared to that occurring between FRC-2 and FRC-8 (or PE-2 and PE-8). The solvent effect appears in these logarithmic plots as a weak perturbation, whereas intramolecular conformational potentials exert a dramatic effect on the time scale of the local motions.

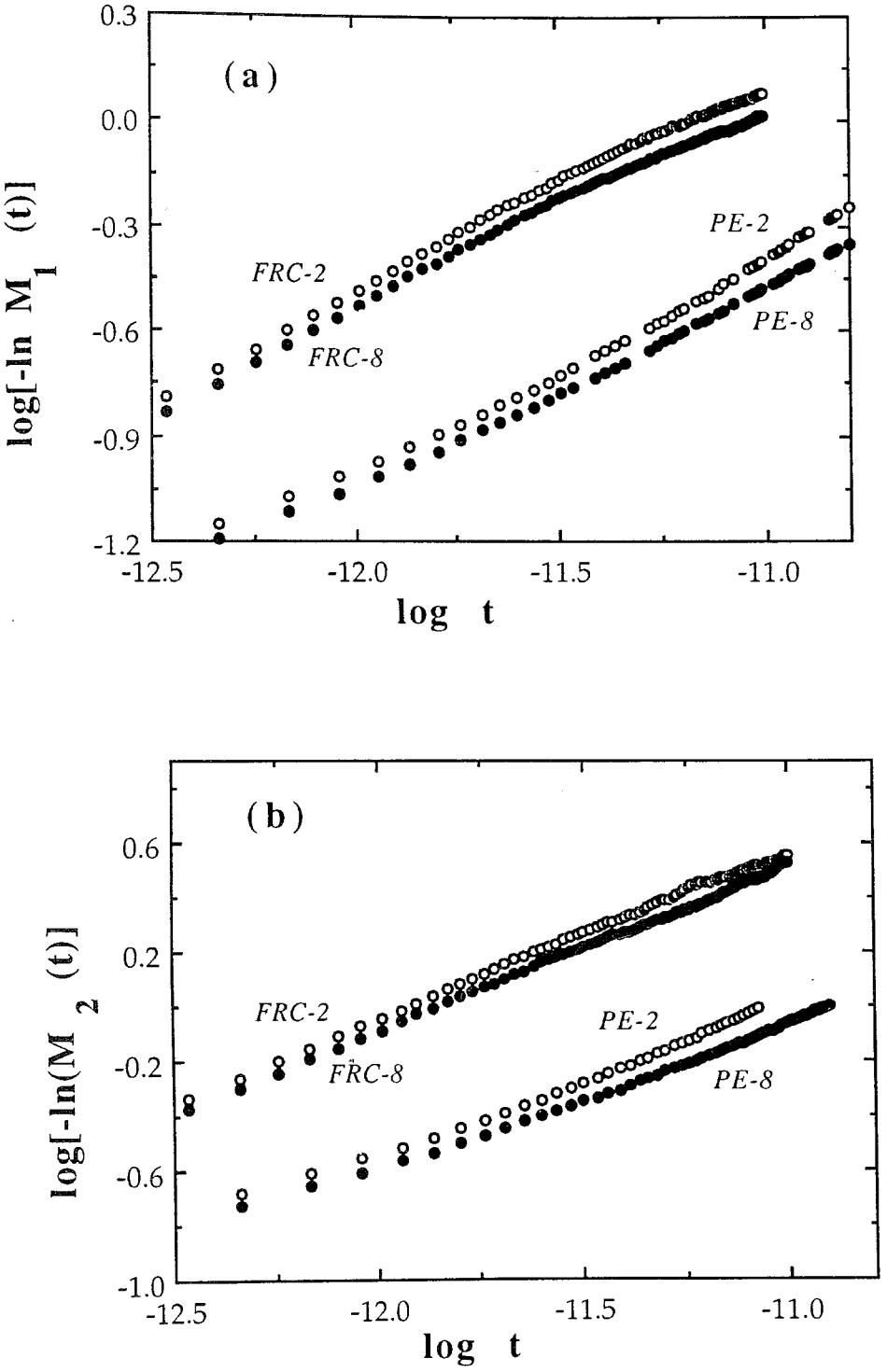


FIGURE 2.21. Stretched exponential representations of (a) $M_1(t)$ and (b) $M_2(t)$ for unit bond vectors $m_{||}$ in FRC-2, 8 and PE-2, 8.

Tables 2.5 and 2.6 summarize the stretched exponential parameters τ_i and β evaluated for the first and second OACFs of vectors $\mathbf{m}_{||}$ and $\mathbf{m}_{\perp}$, repectively

TABLE 2.5. Stretched exponential parameters for OACF $m_{||}$

Model	τ_1 (ps)	β_1	τ_2 (ps)	β_2	τ_1 / τ_2
FRC-2	6.52	0.60	1.18	0.62	5.53
FRC-5	8.28	0.58	1.40	0.61	5.91
FRC-8	8.39	0.58	1.42	0.61	5.91
PE-2	38.0	0.65	7.30	0.53	5.21
PE-5	45.7	0.61	10.7	0.54	4.27
PE-8	74.8	0.56	14.1	0.51	5.30

TABLE 2.6. Stretched exponential parameters for OACF $m_{\perp}$

Model	τ_1 (ps)	β_1	τ_2 (ps)	β_2	τ_1 / τ_2
FRC-2	1.88	0.72	0.61	0.55	3.09
FRC-5	2.18	0.70	0.69	0.54	3.17
FRC-8	2.06	0.71	0.70	0.47	2.94
PE-2	23.6	0.67	8.68	0.48	2.72
PE-5	16.3	0.66	6.61	0.47	2.47
PE-8	16.7	0.63	6.51	0.43	2.57

Several features emerge from the analysis of the results presented in these tables:

(a) In agreement with the above observations, the solvent effect is significantly weaker compared to the effect of intrinsic chain structure, except for the relaxation times of $\mathbf{m}_{||}$ in PE. This general feature is inferred from the close values of the parameters τ_i (or β) obtained for a given type of chain subject to different interactions with the surroundings. In particular, the exponents β are quite insensitive to solvent effect.

(b) In general, the exponents β decrease with increasing intrachain constraints. This final property has lead to an interpretation of the exponent β as a measure of cooperativity between chain segments during relaxational motions. [40]

(c) The relative time scale of the decay of the OACFs associated with the two types of reference vectors needs attention. For the vector $\mathbf{m}_{||}$ in FRC and PE, the correlation times of the first and second OACFs differ by a factor of $\tau_1 / \tau_2 =$ 5.78±0.25 and 4.93±0.66, respectively, whereas for $\mathbf{m}_{\perp}$ in FRC and PE the respective ratios are 3.07±0.13 and 2.59±0.13, as may be deduced from the last column in Tables 2.5 and 2.6. The latter turn out to be relatively close to the values 2.14 and 2.10 obtained for τ_1 / τ_2 of $\mathbf{m}_{\perp}$ in MD simulations of bulk FRC and PE, respectively.[18] For $\mathbf{m}_{||}$, on the other hand, departing from the behavior in solution, bulk polymers were shown to exhibit significant increase in relaxation times, and a sharp difference ($\tau_1 / \tau_2 \sim 10^2$) between the time scales of the two OACFs.[18] That this ratio is decreased to about 5.0 in solution, in agreement with present results, is also confirmed by recent BD simulations of PE chains.[39]

(d) The anisotropy of motion for chain segments, which is a well known phenomena for polymers in solution or in the bulk state,[41, 18] is affected by solvent effect. In general, the orientational relaxation of vectors lying along the chain axis ($\mathbf{m}_{||}$) is recognized to be slower than that of the out-of plane vector ($\mathbf{m}_{\perp}$), this being a consequence of chain connectivity. This behavior is, in fact, verified here by the relative values of the relaxation times for $\mathbf{m}_{||}$ and $\mathbf{m}_{\perp}$. For example, the mean ratios $\tau_{1,||} / \tau_{1,\perp}$ and $\tau_{2,||} / \tau_{2,\perp}$ are larger than unity. They are calculated as 3.78±0.30 and 2.00±0.07, respectively, for FRC model chains. As to the PE model, it would be inaccurate to report mean values for $\tau_{1,||} / \tau_{1,\perp}$ or $\tau_{2,||} / \tau_{2,\perp}$, due to large fluctuations arising from solvent effect. In particular, for

PE chains in the highly contracted state (PE-2) $\tau_{2,\parallel} < \tau_{2,\perp}$, in contrast to the general trend. Thus, the motions along the backbone and those perpendicular to it, which might be identified as the axial and lateral motions, are of comparable rate in the highly contracted state of PE. This is understandable in view of the gradual weakening of the constraints from connectivity along the chain contour, with contraction of the chain configuration. Accordingly, $\tau_{1,\parallel}$ and $\tau_{2,\parallel}$ increase as PE expands from $\alpha = 0.351$ to $\alpha = 1.307$ whereas an opposite change occurs in $\tau_{i,\perp}$. This invites attention to the role of solvent in perturbing the qualitative nature of the motion anisotropy, even though its effect on the absolute time scales of relaxation is not as significant as that of intramolecular constraints.

2.2.2.8. Dependence of Correlation Time on the Size of Chain Segment in Motion. The orientational relaxation of chain segments containing more than a single bond are analyzed in various ε_{bs}. Equation (2.22) is used to calculate the first OACF for unit vectors lying along chain segments of different sizes, characterized by the number of bonds n. A correlation time τ can be assigned to each segment size from the inverse of the initial slopes of the OACF curves. Figure 2.22 displays log τ vs log n curves for FRC and PE model chains in various solvents with indicated ε_{bs} values.

In accordance with the results obtained for bead-spring model chains in Section 2.2.1.7, a linear dependence is observed on the logarithmic scale. Thus, a power law of the form $\tau \sim n^a$ seems applicable for short segments with $n \leq 8$. The exponent *a* is found to be equal to 0.66, 0.89 and 1.13 for PE-2, PE-5 and PE-8, respectively; whereas it changes between 0.90 and 0.98 for FRC2-8. For bead-spring model chains in solution, the exponent *a* has been found to increase in the range $1.0 \leq a \leq 1.5$ with increasing solvent quality. In summary, for all model chains studied by MD simulations,the value of *a* increases as the intermolecular interactions get more favorable. Yet, the values of the exponent *a* obtained by MD simulations in various solvent lie below the Rouse and Zimm exponents, which equate to 2 and 3v, respectively, v assuming the respective values 0.5 and 0.6 in theta and good solvent conditions. This latter result lends support to the existence of a so-called sub-Rouse regime, in which an exponent *a* lower than the Rouse-Zimm values applies.[33, 34] As a final remark, a definite time scale difference is observed in the orientational relaxation of FRC and PE

in parallel with the behavior discussed in the previous section, and PE seems to be more affected by the changes in solvent quality.

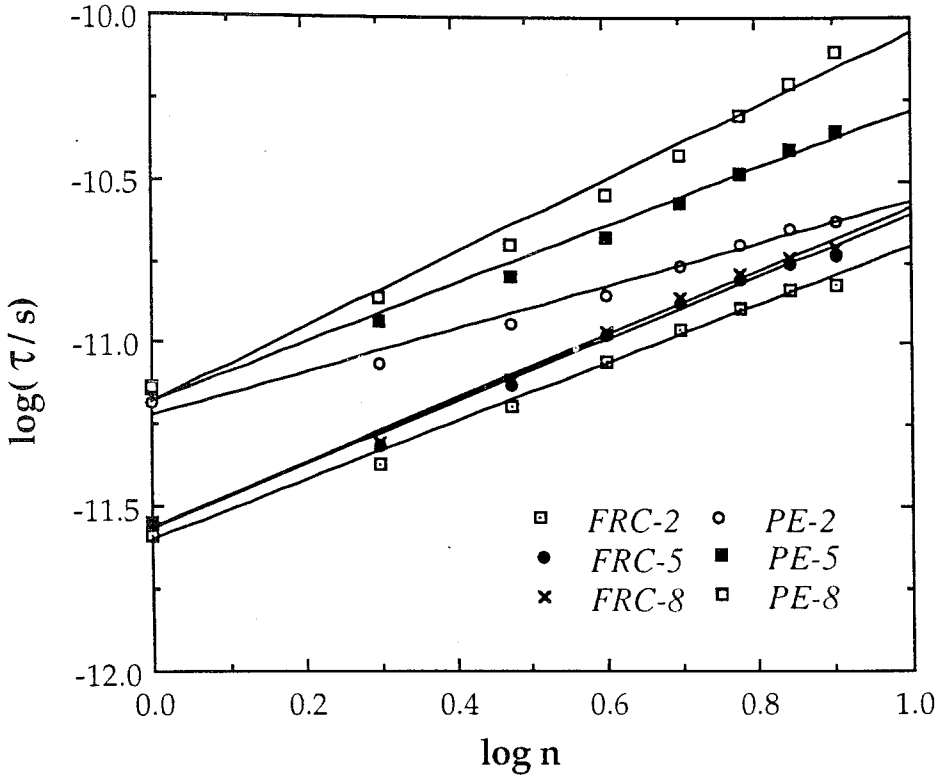


FIGURE 2.22. Dependence of the orientational correlation times τ on the size of the chain segment involved in local motion. log(τ /s) vs. logn is plotted for FRC-2, 5, 8 and PE-2, 5, 8, where n is the number of bonds in the chain segment. The straight lines are drawn by least square fit.

3. DYNAMICS OF HELICAL POLYPEPTIDES IN *VACUUM* AND IN WATER

3.1. Theoretical Background

3.1.1. Biological Activity of Proteins

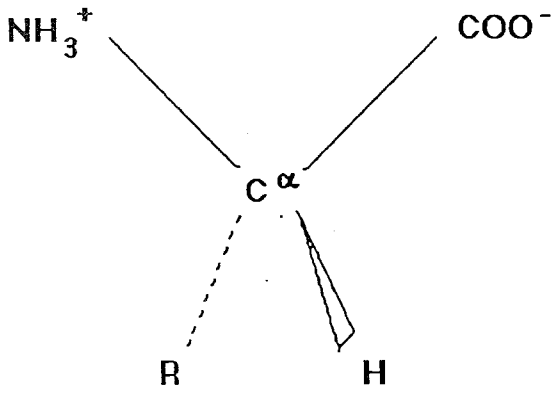
Proteins play important roles in virtually all biological processes. [42] They may be classified mainly into three types: fibrous, membrane and globular.[43] Fibrous proteins form the major component of structures, such as muscle, skin, tendon, bone and hair. Membrane proteins reside in cellular membranes, where they mediate the exchange of small molecules and information across cell boundaries. Enzymes, which are globular proteins, are responsible for the catalysis and regulation of the chemical reactions that take place in biological systems. Chemical transformations *in vivo* rarely proceed at perceptible rates in the absence of enzymes. Other biological processes in which proteins play a major role are the immune protection, the generation and transmission of nerve impulses and the control of growth and differentiation of cells.[42] Thousands of different types of proteins exist to participate in this great number and variety of biological events.

3.1.2. Structure of Proteins

Proteins are linear unbranched polymer chains made up of tens to thousands of monomer units. The monomers are the naturally occurring amino acids, which are successively linked to each other by peptide bonds. A chain composed of amino acids is called a polypeptide or simply a peptide. If a

polypeptide consists of more than 50 amino acids, it is placed in the protein category. [44]

Naturally occurring amino acids have the following general chemical structure



where the amino (NH_3^+) and carboxyl (COO^-) functional groups attached to the alpha carbon are protonated and deprotonated, respectively, at neutral pH. Thus, the ionic form of the amino acids predominates at neutral pH. Each amino acid is distinguished by the structure of the R group. The R group can vary from a single hydrogen atom in glycine to a more complex structure such as the guanidine group in arginine. The chemical structures of the R groups, the names and the commonly used three-letter codes for the twenty amino acids found in proteins are given in Appendix B. The side chain units are successively designated by the symbols β, γ, δ, ε and ζ (e.g. Cβ, Cγ, etc.), starting from the group connected to the Cα. In all twenty natural amino acids, except glycine, the α carbon is an asymmetric center, attached to four different groups. Among the two stereoisomers, that are called the L and D configurations, only L-amino acids are known to occur in proteins. The carboxyl group of one amino acid is joined to the amino group of another and a water molecule is lost by the formation of a peptide bond. Figure 3.1 displays a section of an α-L-polypeptide chain in all-*trans* conformation.[45]

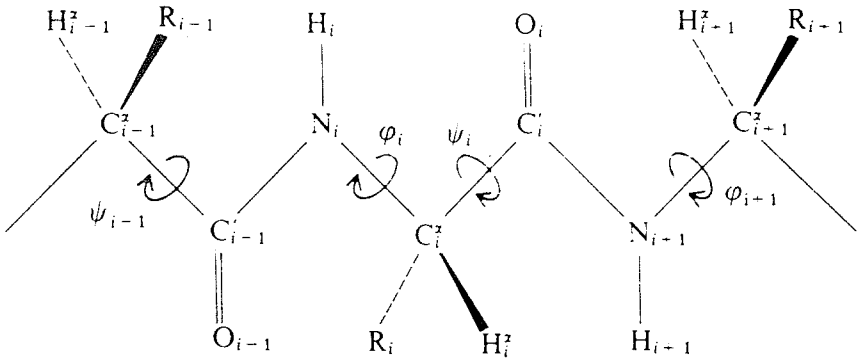


FIGURE 3.1 A section of an α-L-polypeptide chain in all-*trans* conformation.[45]

A polypeptide chain has a direction due to the two distinct terminal groups existing at the ends, namely the α amino group (N-terminal) and the α carboxyl group (C-terminal). Residues are conventionally indexed serially from 1 to n, starting from the N-terminus. The individual amino acids differ according to their R-group side chains. The primary structure of the chain is determined by specifying the identity and sequence of all amino acids in the polypeptide.

Considering all bond lengths and angles to be fixed at their equilibrium values, the conformation of a polypeptide is determined by the rotations of the backbone and side chain dihedral angles. An important aspect of the polypeptide structure is that the peptide bond usually occurs in planar *trans* conformation due to its partial double-bond character. Therefore, the specification of all ϕ_i and ψ_i rotations about the respective N-C^{α} and C^{α}-C bonds is sufficient to uniquely determine the conformation of the polypeptide backbone. The specification of rotation angles in biopolymers is different from the Flory notation[2], which is adopted in the Section 2 on polymer dynamics. In Flory notation, the *trans* state is assigned a rotational angle of 0°, whereas in Bio-notation, which will be adopted in current section of this thesis on polypeptides, the *cis* conformation is taken as 0° and *trans* is 180° (or -180°).

Polypeptides can fold into regular structures. Patterns of hydrogen bonding between carbonyl oxygens and amide hydrogens of the backbone, combined with repeating values of ϕ_i and ψ_i angles define the regular secondary structures observed in polypeptides. α-helix and β-sheet are the two common secondary structures observed in proteins. The β-sheet structure involves repeating patterns of hydrogen bonds between distant parts of the backbone, whereas helices involve local H-bonding.

The right-handed α-helix, shown in Figure 3.2, is formed with $\phi = -57°$ and $\psi = -47°$. The tightly coiled polypeptide main chain forms the inner part of the rod, and the side chains extend outward. The α-helix is stabilized by the H-bonds between the CO group of residue i and the NH group of residue (i+4) along the backbone. The H-bonds are nearly parallel to the helix axis, with all CO groups pointing toward the C-terminal. The peptide bond has a substantial dipole moment and in helix structure the peptide dipole moments add end-to-end across the H-bonds to generate a macrodipole, with positive and negative poles at the N- and C-terminus, respectively.[46] The pitch, or repeat, of an ideal α-helix is 3.6 residues per turn. For that pitch, the rise per residue along the helix axis is 1.5 Å, or 5.4 Å per turn. Thus, amino acids that are placed three or four residues apart in the linear sequence are spatially quite close to one another, whereas amino acids two residues apart are situated on opposite sides of the helix. [47] Three NH and three CO groups at the respective N- and C-terminus of the polypeptide do not have H-bonding partners along the backbone. They often form H-bonds with the solvent or participate in helix-capping interactions with the side chains of nearby residues. The α helices observed in proteins are right-handed, because in the left-handed α-helix ($\phi = 57°$ and $\psi = 47°$) each C^β atom would collide with the following turn on the helix. However, isolated residues with left-handed helix torsions are fairly common, especially glycine.

The only other common type of repeating secondary structure in proteins is the extended β-pleated sheet. Figure 3.3 displays a polypeptide in β-sheet conformation. This is almost fully extended rather than being tightly coiled as in the α-helix. The axial distance between adjacent amino acids is 3.5 Å, in contrast to 1.5 Å for the α-helix. The nonlocal H-bonds are between different strands of the same polypeptide chain or of different chains. The H-bonded strands in a β-sheet can run in the same direction (parallel β-sheet with $\phi = -119°$ and $\psi = 113°$) or in opposite directions (antiparallel β-sheet with $\phi = -139°$, $\psi = 135°$).[45]

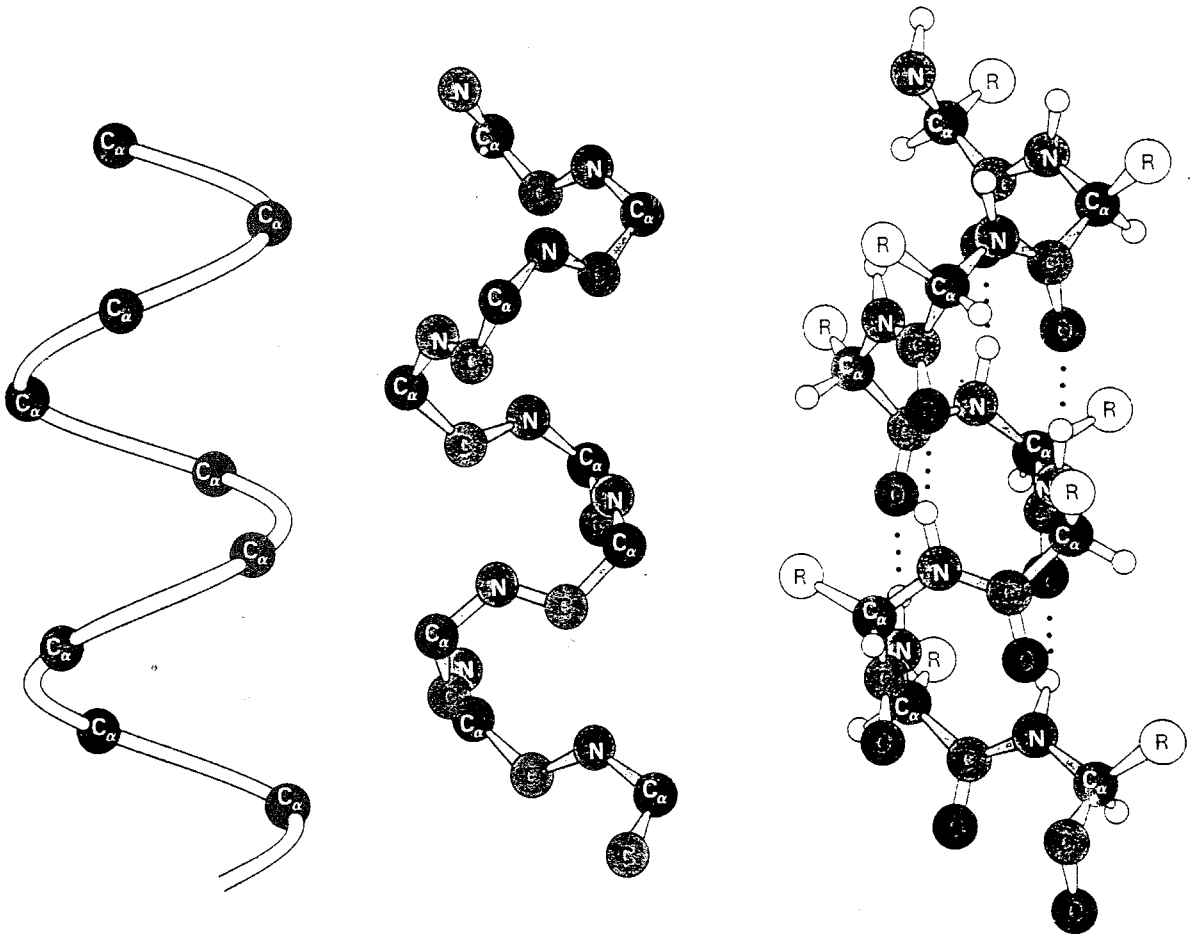


FIGURE 3.2. Right-handed α-helix structure[42]

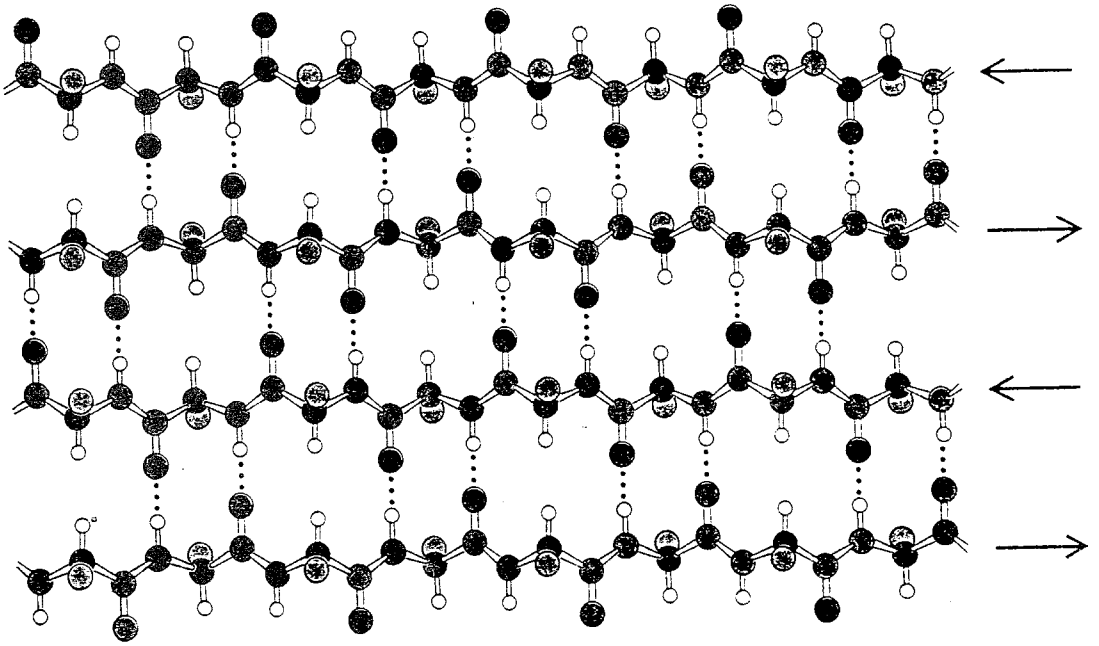


FIGURE 3.3. β-pleated sheet structure[42]

Two dimensional plots of the backbone ϕ and ψ angles, known as Ramachandran plots, show clearly the regions of repeating secondary structures and regions of unfavorable steric overlaps. Figure 3.4 displays the Ramachandran plot for all amino acid residues, except glycine and proline. This plot is obtained from the crystal structures of 310 proteins.[48] The major regions observed on the plot are the right-handed α-helical cluster in the lower left, the broad region of extended β strands in upper left , and the sparsely populated left-handed α-helical region in the upper right quadrants. Vacant areas are the conformations that place atoms unfavorably close to each other within the dipeptide unit. The asymmetry of the plot results from the steric clashes with the C^{β} atoms of the side chains. The Ramachandran diagram for glycine is centrally symmetric, i.e. symmetric with respect to any line that passes through the center ($\phi = 0^{o}$ and $\psi = 0^{o}$). This is a consequence of the symmetry of the glycine residue.[45]

Other secondary structures, such as turns, connections and compact loops, exist which are composed of well-ordered but non-repeating conformations. These structures are different from random coil conformation, because they are compact and stabilized by backbone H-bonds and side-chain-to-main-chain H-bonds. For example, β-turns are formed when a polypeptide chain makes a sharp 180^{o} reversal, which is generally stabilized by a local $CO_{i}....NH_{i+3}$ hydrogen bond.[48]

The secondary structural elements assemble in various ways into larger modules called supersecondary structures, such as β-hairpin, the $\beta/\alpha/\beta$ unit, and the α/α hairpin. These in turn form the basis of larger domain structures. The high-level architecture of a protein, in which helices, sheets and other secondary structures fit together, is called the tertiary structure. The unique tertiary structure of a globular protein, known as its native or folded state, is extremely compact and determines its biological activity.

There is still a higher level of organization in some proteins, which is called the quaternary structure.This structure is found in oligomeric proteins that exist as aggregates of two or more polypeptide chains, which are linked together by noncovalent forces.

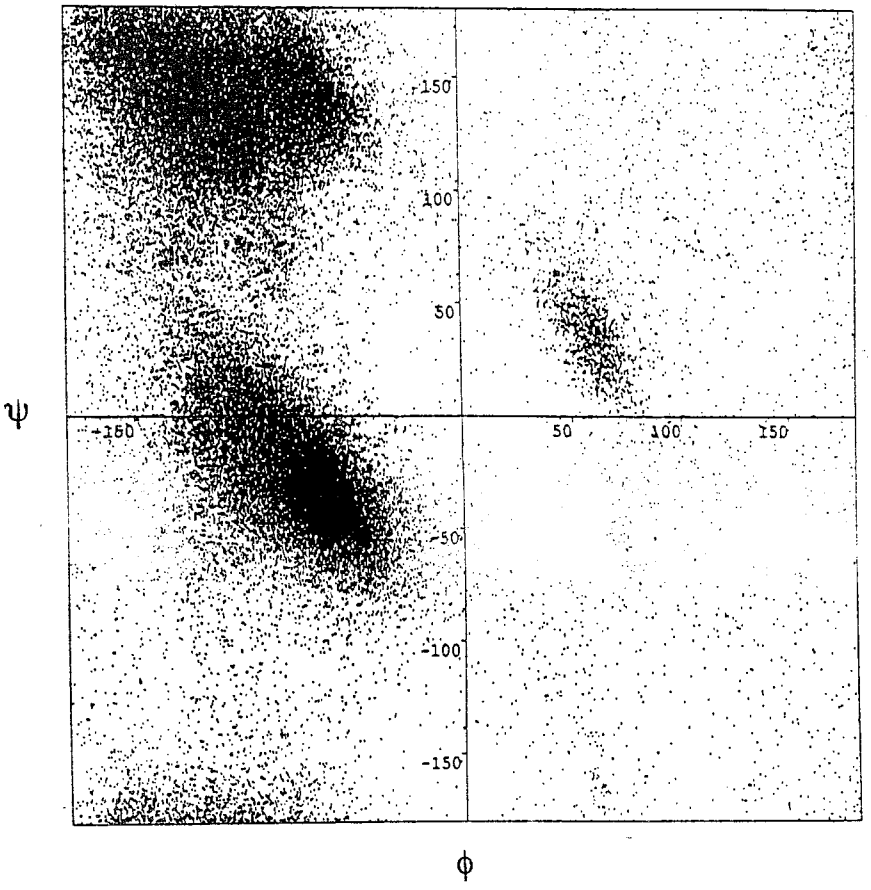


FIGURE 3.4. Ramachandran plot for all amino acid residues, except proline and glycine, obtained from the crystal structures of 310 proteins[48]

3.1.3. The Protein Folding Problem

A protein molecule adopts its unique, three-dimensional structure spontaneously under physiological conditions, i.e. in aqueous solution near neutral pH and at 20-40°C. [43, 49, 3] This native structure can be denatured often reversibly by elevated temperature, acidic or basic pH and some nonaqueous solvents, and the protein unfolds to an ensemble of more expanded conformations. The protein in coil conformation may refold to the unique, native state *in vitro* conditions, with full restoration of its biological activity.

The transition of the polypeptide chain from a disordered state to the ordered, native state is called protein folding.[43, 49, 3] It is well established that this transition is directed primarily by the amino acid sequence of the chain. The protein folding problem is to understand how the primary structure of the chain determines its three-dimensional structure. The solution of this fundamental problem will be a breakthrough in biomedicine, which will enable the design of novel proteins and drugs.

Several important experimental observations have been made about the native states of proteins and the folding process, which will be summarized below:

(a) The protein folding transition of small globular proteins exhibits two-state kinetics, in which the folded and unfolded states with respective concentrations of [N] and [U] interconvert with negligible population of stable intermediates.[49] The difference in free energy between the native and unfolded states is given by

$$\Delta G = - RT \ln K_{unf} \qquad (3.1)$$

where $K_{unf} = [N] / [U]$, R is the gas constant and T is the absolute temperature. For globular proteins, typical values of ΔG are quite small, lying in the range of about -5 to -15 kcal / mol.[49]

(b) Amino acids can be classified as nonpolar, polar or ionic. In globular proteins, nonpolar residues tend to be buried inside the core of the globule, implying that proteins are driven to compactness due to the hydrophobic effect.[43] Polar and ionic residues tend to reside on the surface of the globule, although exceptions are common. However, even residues with hydrophobic side chains have H-bonding capacity due to backbone NH and CO groups. If these groups were not able to make H-bonds in the protein interior through secondary structure formation, they would drive the protein to denaturation.[3]

(c) The interior of globular proteins are tightly packed, with packing densities close to crystals of small organic molecules, i.e. solids rather than hydrophobic oil aggregates. The protein core is comprised primarily of α and β structures, with connecting elements of nonrepetitive turns and loops existing on the outside. Side chains in the interior fit together with striking complementarity, like pieces of a three-dimensional jigsaw puzzle.[3] These specific packing interactions within the globule are believed to be the major source of structural specificity in proteins. The formation of this unique fold instead of an ensemble of micelle-like structures cannot be explained by the hydrophobic effect alone.

(d) Although static measurements suggest tight packing with constrained tolerances inside the globule, both dynamics and mutagenesis studies indicate that the protein molecule can tolerate a broad diversity of residue substitutions with minor effects on structure, stability and function.[3] These conflicting properties of protein structure, i.e. tight versus malleable, could coexist if tight packing of the protein is attained primarily by local arrangements, without global rearrangements.

In view of these observations, a hierarchic framework model of protein folding has been proposed[3], in which secondary structural elements associate in step-wise fashion, leading to larger modules and further association to tertiary structure. The experimental evidence of short, stable helices in water, which will be explained in the next section, have suggested that one of the earliest steps in protein folding is the formation of secondary structural elements. Stable, isolated turns and loops in solution have also been reported[3] However studies on isolated β-sheets have not been possible because sheets tend to aggregate in solution.

3.1.4. Experimental Work on α-Helices

The α-helix is the most abundant form of secondary structure found in globular proteins. Yet, it has long been regarded that short helices of less than 20 residues should not exhibit measurable helix formation in water at any temperature, according to the Zimm-Bragg helix-coil transition theory.[50] The N-terminal C- and S-peptide fragments of ribonuclease A (residues 1-13 and 1-20, respectively) are the first examples with significant α-helix formation in aqueous solution near 0°C.[50] Further studies on the C-peptide[51, 46, 52]show that helix stability is pH dependent. In α helices, acidic and basic residues occur preferentially near the respective N- and C-terminus, which permits favorable interactions between charged groups and the macrodipole of the helix. The studies on the C-peptide[46, 52] support the existence of this charged group effect on helix stability termed as the 'helix dipole model'. Marqusee and Baldwin[53] have designed alanine-based peptides, which are 16-17 residues long and contain three glutamic acid/lysine residue pairs. These peptides exhibit high helix content (up to 80 per cent) due to the formation of (Glu^-, Lys^+) ion pair interactions.

The examples of short peptide helices in water stabilized by specific side chain interactions have raised the question of what helix content these short peptide sequences would exhibit in the absence of these interactions. It is shown by Marqusee et al.[50] that short 16-residue, alanine based peptides, which are solubilized in water by insertion of charged residues of a single type(Glu^- or Lys^+), form stable helices in water. These surprising results, which cannot be explained by specific interactions, indicated that individual alanine residues have a high helical potential and can stabilize α helices. Later other *de novo* designed peptides[54-56] have been reported to exhibit helical structure at low temperature.

3.1.5. α-Helix Propensities of Amino Acids

Amino acids differ in their intrinsic helical propensities. Statistical analysis of the frequency of occurrence of amino acids in different secondary structures have suggested that different side-chains have distinct conformational preferences.[57] Different experiments have been carried out to estimate α-helix propensities of amino acids. Scheraga and co-workers measured helix stability by introducing the amino acid of interest as a 'guest' into water-soluble helix-forming synthetic copolymers.[58] Small model peptides have also been used to determine the changes in free energy associated with the substitution of different amino acids in the host systems.[59-61] On the other hand, directed mutagenesis experiments have been carried out by substituting the amino acids at specific solvent-exposed α-helical sites of proteins, such as lysozyme and barnase.[62, 63] Table 3.1 summarizes the helix propensity scales obtained from the statistical and experimental studies. In experimental studies, free energy differences ΔG for amino acids are calculated from Equation (3.1), by considering a two-state model between helix and non-helix states (or native and unfolded proteins). In this equation, K_{unf} = f / (1-f), where f is the fraction of helix determined from circular dichroism (CD) or fluorescence spectroscopy measurements of the polypeptides. The free-energies of helix formation relative to glycine are tabulated as ΔΔ G values so that the helix propensity of Gly is 0.[59]

It appears that the rank order of helix-forming tendencies depend on the system in which replacements are made. One major difference between helical peptides and α helices in proteins is that the ends of the peptides fray. There are also recent experiments,[64, 65] which give scales of β-sheet propensities of amino acids.

Different factors have been suggested as the reason for the relative helical propensities of amino acids, including conformational entropy, steric factors, hydrophobic effect and main-chain electrostatics.

Helix formation imposes restrictions on the side chain conformations of residues because of steric clashes with the bulky helix backbone. In extended conformations, side chain dihedral angles, χ_1, populate the g^+ (-60°), g^- (60°) and t (180°) conformers for optimal staggering of substituents on the Cα and Cβ

TABLE 3.1. Helix Propensity Values for Amino Acids.[62, 63]

Amino Acid	ΔΔG (kcal/mol)				Percent helix (E)	P_α (F)
	(A)	(B)	(C)	(D)		
Ala	-0.96	-0.91	-0.77	-0.79	78	1.42
Leu	-0.92	-0.56	-0.62	-0.62	80	1.21
Met	-0.86	-0.60	-0.50	-0.57		1.45
Ile	-0.84	-0.10	-0.23	-0.39	41	1.08
Gln	-0.80	-0.43	-0.33	-0.48		1.11
Arg	-0.77	-0.77	-0.68			0.98
Lys	-0.73	-0.72	-1.23			1.16
Tyr	-0.72	-0.09	-0.17			0.69
Val	-0.63	-0.03	-0.14	-0.34	17	1.06
Phe	-0.59	-0.22	-0.41		23	1.13
Trp	-0.58	-0.07	-0.45			1.08
His	-0.57	-0.13	-0.06			1.00
Thr	-0.54	-0.12	-0.11	-0.23		0.83
Glu	-0.53	-0.36	-0.27			1.51
Ser	-0.53	-0.50	-0.35	-0.28		0.77
Asp	-0.42	-0.20	-0.15			1.01
Cys	-0.42	-0.09	-0.23			0.70
Asn	-0.39	0.25	-0.07	-0.18		0.67
Gly	0.00	0.00	0.00	0.00		0.57
Pro	2.50	3.17	3.00			0.57

(A) Results of directed mutagenesis of T4 lysozyme at residue 44[63]

(B) Results of directed mutagenesis of barnase at residue 32[62]

(C) Results obtained with a noncovalent α-helical dimer[60]

(D) Data from a peptide containing blocks of Glu and Lys residues[59]

(E) Data from a 17-residue alanine-based peptide[61]

(F) Statistical survey of Chou and Fasman.[57] P_α represents the relative frequency with which amino acids occur in α helices.

atoms. For most residues in helix conformation, the *g*⁻ conformer becomes effectively forbidden by steric clashes between Cγ of residue i and the carbonyl oxygen of residue i-3 on the helix.[48] Alanine has no Cγ atom, which may be one factor affecting its high helical propensity. In the most common conformer, g^+, in α helices, the average χ_1 value is shifted by 8° from -60°, which is the value observed in β-sheets or coils.[48] This must introduce a strain energy, which will be absent in Gly and Ala residues. The branched β-carbon amino acids (Val, Ile and Thr) are special because they cannot relieve clashes with the previous turn by changing χ_1. Consequently their χ_1 values remain close to -60°, but they must be under considerable strain. This may also explain why these amino acids are relatively poor helix formers.

Creamer and Rose[66] have shown by Monte Carlo simulations that the loss of side chain conformational entropy upon helix formation correlates with the experimental propensity scales for the eight nonpolar amino acids. The stability of small nonpolar amino acid residues in helix conformation has also been studied by free energy simulations.[67] The results correlate well with the experimental propensities and qualitatively suggest that the conformational entropy and steric strain affect the α-helix propensities of amino acid residues.

The site-directed mutagenesis experiments of Blaber et al.[63] on T4 lysozyme show that the side chain hydrophobic surface area buried against the protein correlates well with the relative free energies of unfolding of the mutant helices. They therefore propose that hydrophobicity primarily determines the preferences of residues for helix conformation.

Avbelj and Moult[68] have proposed main chain electrostatics to be the determining factor of residue conformational preferences. The α-helical, β-sheet and other main chain conformational states of residues have quite distinct electrostatic energies of interaction between adjacent peptide groups. Specifically, the CO and NH dipoles are aligned antiparallel in β conformation and interact favorably , whereas in α-helices, parallel orientation of neighboring dipoles lead to unfavorable interactions. These local orientation effects extend over the whole polypeptide. Screening of these electrostatic interactions by surrounding water molecules depends on the type of amino acid side chains. The electrostatic model of Avbelj and Moult, which is based on the distributions of conformations in experimental protein structures, explains the conformational preferences of residues in terms of the degree of screening of main chain

electrostatics by different side chain types. The model seems to explain the experimental α-helix and β-sheet propensity scales of amino acids, at the same time.

3.1.6. MD Simulations of Helical Peptides

Molecular dynamics simulations have been carried out to investigate the dynamics of small peptides in various solvents. Daggett et al.[69] have performed MD simulations of a 17-residue peptide at 278 K, experimentally designed by Bradley and co-workers.[55] Three different dielectric models, namely a linear distance dependent dielectric function, a sigmoidal distance dependent dielectric function and a constant dielectric with explicit water molecules around the peptide, are used in the simulations with the aim of determining how well these models reproduce the experimentally observed helical characteristics of the peptide. Simulations of 1.2 ns total duration, carried out at low pH using the sigmoidal function, lead to comparable results with experimental data. In contrast, the results obtained with the linear dielectric function are not satisfactory. Although the simulations with explicit water molecules (100 ps) are the most realistic, they have the disadvantage of computational ineffiency. In this study, the sigmoidal function is shown to be a reasonable alternative, since the peptide samples both helical and nonhelical regions of the conformational space efficiently.

Simulations have been performed to analyze the dynamics of helix-coil transitions, which have a time scale on the order of 10^{-9} - 10^{-6} sec. However, it is impossible to cover the whole time-scale range of such transitions by simulations, that can go up to several nanoseconds (10^{-9}) by the currently available computer resources. McCammon and Northrup[70] have performed the first relatively long simulations (12 ns) of a 15-residue polyvaline using a simplified model where each residue is represented by a soft sphere and virtual bonds. Stochastic dynamics simulations start from an initial α-helix conformation and the dynamics of unwinding of the last four residues are analyzed by fixing all other residues in space.

Daggett et al.[71] have simulated a 20-residue polyalanine chain at high temperature (400 K) using a linear dielectric constant to make a detailed analysis of the helix-coil transitions. The results of the 4 ns MD trajectory indicate that the peptide spends most of its time fluctuating between different conformations with intermediate helix contents. Transitions between high and low helical contents are rare and rapid. These results are in accordance with the experimental findings that small peptides are only marginally stable as helices in solution.

In a survey of protein crystal structures that contain hydrated α helices, Sundaralingam and Sekharudu[72] have found a number of instances in which a water molecule is inserted into the backbone H-bond, acting as a bridge between the carbonyl and amide groups and suggested that these structures may be intermediates in the folding / unfolding pathway. Later, MD simulations of polyalanine peptides[73] have revealed instances of local helix destabilization by this specific water insertion process.

Daggett and Levitt[74] have investigated the helix-coil transitions in a 13-residue polyalanine both in vacuum and in the presence of explicit water molecules by MD simulations. The denaturation of the peptide is monitored as a function of temperature (ranging from 5 to 200 ^{o}C). In vacuum, the helical state predominates at all temperatures, whereas in solution the helix melts with increasing temperature. At high temperature (473 K), the peptide unfolds and adopts various collapsed unstructured states. The intrachain H-bonds that break at high temperature are not fully compensated by H-bonds with water molecules. Numerous water molecules inserted into the intrachain hydrogen bonds have been observed, the occurrence of which increased with temperature.

Molecular dynamics simulations[75] have been performed on the C-terminal α-helix part of myoglobin at 300 K in two different solutions, namely water and water with 30 per cent trifluoroethanol (TFE). In the course of 200 ps trajectory, unfolding of the helix is observed in water. In TFE solution, two stable parts of the helix exist till the end of the trajectory, supporting that TFE is a structure-forming solvent.

In practice, it is almost impossible to study the conformational equilibrium in peptides by a single simulation, since the system generally gets trapped in wells on the potential energy surface and transitions across high energy barriers occur rarely during a sufficiently long simulation. Tobias and Brooks[76] have

overcome this problem by using a specialized technique known as 'umbrella sampling', and used molecular dynamics simulations to study the folding / unfolding of one turn of an α-helix in alanine and valine tripeptides with methyl-capped ends. Analysis of the peptide conformations during the simulation show that α helices, reverse turns and extended conformations correspond to minima on the free energy surfaces of both peptides. The free energy difference between α-helix and extended conformations is determined to be approximately -1 kcal / mol and -5 kcal / mol for the alanine and valine peptides, respectively. Reverse turns appear as important intermediates along the helix folding / unfolding pathway, in both peptides. Their results indicate that the large differences between the helical propensities of Ala and Val cannot be explained simply by concepts of side-chain rotamer restriction or unfavorable steric interactions. Rather, the origin of the difference seems to be more complicated due to difference in the solvation of the two peptides.

Brooks and Nilsson[77] have used the same sampling technique to study the conformational free energy surfaces of the same alanine tripeptide in methanol, TFE, water and their mixed solvents. Simulation results indicate the helix-promoting abilities of the particular solvents in the following order: TFE-water ≈ methanol > methanol-water > water ≈ TFE. Differences in the solvation thermodynamics of peptides in alcoholic and aqueous solvents have been observed.

3.2. Simulation Method and Parameters

Molecular dynamics simulations of polypeptides are carried out using GROMOS87 Simulation Package, which has been developed for the dynamic modeling of biomolecules. (W.F. van Gunsteren and H.J.C. Berendsen, GROMOS: Groningen Molecular Simulation Program, University of Groningen, 1987) Appendix C gives information on the basic concepts, equations and force fields utilized in GROMOS87, together with a description of its application to the specific polypeptide systems considered in this work.

Four different model polypeptides, each composed of 13 residues, are simulated in *vacuum* and in the presence of explicit water molecules at 350 K and 1 bar. The primary structures of the polypeptides are Ala_{13} (polyalanine), Val_{13} (polyvaline), Ser_{13} (polyserine) and $Ala_4Gly_5Ala_4$ (polyglycine). It should be noted that the peptide named as polyglycine also contains alanine residues. Table 3.2 summarizes the simulations performed with the four polypeptides, which will be designated as pAla, pVal, pSer and pGly, for simplicity, in the following. Simulations in water and in *vacuum* will be differentiated by the capital letters W and V, respectively. Each run in Table 3.2 represents a simulation of 300 ps duration. Independent runs starting from the same initial configuration of each peptide and water molecules but with different initial Gaussian velocity assignments are numbered as 1, 2 and 3. The simulations in water are performed in simulation boxes subject to periodic boundary conditions.

The initial conformation of each polypeptide is generated as a full right-handed α-helix by assigning ϕ = -57.5°, ψ = -47.5° for the backbone dihedral angles. (Bio-notation is used, see Figure 3.1) The N-terminal and C-terminal groups of each chain are NH_3^+ and COO^- and the molecule is taken electrically neutral overall. Each helical polypeptide is then placed in the middle of an orthogonal simulation box and the box is filled with equilibrated SPC (Simple Point Charge Model)[78] water molecules. The water molecules are placed with the requirements that the minimum distance between any atom of the peptide and the walls of the box is 7 Å and that no water molecule exists within 2.3 Å of any peptide atom. The lengths of the edges of the orthogonal box along x, y and z directions and the number of the generated solvent molecules are given in Table 3.2 for each peptide. The helix axis lies along the z-direction. The box edges along x and y direction and the number of solvent molecules generated for polyvaline and polyserine are considerably greater than those for polyalanine and polyglycine due to the bulky side groups of valine and serine. For polyalanine, a box with longer edges along x and y directions is also generated for exploring the effect of the size of the simulation box, which is indicated by pAla-WL.

The energy of the original configuration is minimized by applying 50 steps of steepest descent algorithm. One Na^+ and one Cl^- ions are added to the box by randomly replacing two water molecules. A brief energy minimization of this configuration is performed (about 30 steepest descent cycles). All atoms in the box are assigned velocities from a Gaussian distribution at 350 K. Two

different equilibration procedures are employed. Runs indicated by W1 are equilibrated by adopting the procedure of van Buuren and Berendsen.[75] In this procedure, a 20 ps equilibration period is adopted, during which the system is heated by coupling to an external heat and pressure bath[79] at 350 K and 1 bar. In this external bath system, the temperature of the system, T, is kept around the reference temperature T_o by rescaling the velocities at each time step. Accordingly, the velocities are multiplied by a scaling factor λ, of the form

$$\lambda = \left[1 + \frac{\Delta t}{\tau_T}\left(\frac{T_o}{T} - 1\right)\right]^{\frac{1}{2}} \qquad (3.2)$$

where Δ t is the time step size and τ_T is the temperature scaling constant. τ_T is a measure of the strength of coupling to the heat bath. If τ_T is infinitely large, there is no temperature scaling and if $\tau_T = \Delta$ t, the temperature of the system is kept at T_o at all times. Similarly, the pressure P of the system is kept around the reference pressure, P_o, by multiplying all the atomic coordinates by μ and the box volume by μ^3 at each time step. μ is given by

$$\mu = \left[1 + \frac{\Delta t}{\tau_P}\beta\,(P - P_o)\right]^{\frac{1}{3}} \qquad (3.3)$$

where τ_P is the pressure scaling constant and β is the compressibility of the system. τ_T and τ_P are taken as 0.01 and 0.05 ps, respectively, during equilibration periods. During the first 5 ps of the equilibration period of W1 runs, the peptide is harmonically constrained to its initial helical conformation. The runs designated by W2 and W3 involve short equilibration periods of 1 ps by coupling to the same external bath at 350 K and 1 bar and applying no constraints on the peptide. MD simulations of 300 ps are performed after the equilibration periods by coupling to an external bath at 350 K and 1 bar with τ_T = 0.1 and τ_P = 0.5 ps. The structures are saved every 0.25 ps during the trajectory

resulting in 1200 structures for analysis. Covalent bond lengths and the water angle are constrained by the SHAKE algorithm[80] (tolerance 0.0001Å), which allows the use of 2 fs time steps. Nonbonded interactions are cut off at 0.8 nm.

In *vacuum* simulations, the energy minimization is performed on the initially generated polypeptide in full α-helix conformation by applying 50 cycles of conjugate gradients. A brief equilibration procedure of 1 ps is applied as in water simulations indicated by W2 and W3. MD simulations of 300 ps follow after the equilibration periods with 2 fs time steps. The external bath at 350 K and 1 bar with the same time constants as those of water simulations is used. No nonbonded cutoff is applied in *vacuum* simulations.

TABLE 3.2. Summary of Simulation Experiments

Run	Polypeptide	Model	No.of Solvents	Simulation Box Sizes (nm)		
				x	y	z
1	Ala_{13}	pAla-W1	540	2.13913	2.27856	3.61099
2	Ala_{13}	pAla-W2	540	2.13913	2.27856	3.61099
3	Ala_{13}	pAla-W3	540	2.13913	2.27856	3.61099
4	Ala_{13}	pAla-WL	764	2.53913	2.67856	3.61099
5	Ala_{13}	pAla-V1	*vacuum*			
6	Ala_{13}	pAla-V2	*vacuum*			
7	Val_{13}	pVal-W1	677	2.40030	2.53279	3.61099
8	Val_{13}	pVal-W2	677	2.40030	2.53279	3.61099
9	Val_{13}	pVal-W3	677	2.40030	2.53279	3.61099
10	Val_{13}	pVal-V1	*vacuum*			
11	Val_{13}	pVal-V2	*vacuum*			
12	$Ala_4Gly_5Ala_4$	pGly-W1	544	2.13913	2.27856	3.61099
13	$Ala_4Gly_5Ala_4$	pGly-W2	544	2.13913	2.27856	3.61099
14	$Ala_4Gly_5Ala_4$	pGly-V1	*vacuum*			
15	$Ala_4Gly_5Ala_4$	pGly-V2	*vacuum*			
16	Ser_{13}	pSer-W1	627	2.34443	2.42446	3.62330
17	Ser_{13}	pSer-W2	627	2.34443	2.42446	3.62330
18	Ser_{13}	pSer-W3	627	2.34443	2.42446	3.62330
19	Ser_{13}	pSer-V1	*vacuum*			
20	Ser_{13}	pSer-V2	*vacuum*			

3.3. Simulation Results and Discussion

In this section, the simulation results for the four model polypeptides will be presented and discussed. Structural properties are analyzed with emphasis on the changes in the helix contents of the peptides during each trajectory. Patterns of water distribution around the side chains and the hydrogen bonds are investigated to explain the differences between the helix-forming / breaking tendencies of the four amino acids.

3.3.1. Structural Properties of the Polypeptides

The (ϕ, ψ) dihedral angle pairs of the residues making up each model chain are plotted in the form of Ramachandran maps in Figure 3.5. Each plot represents the results from an independent run and each point on the plot represents a (ϕ, ψ) pair of a specific residue at a snapshot of the simulation. The initial 100 ps period of the simulations is discarded to remove the bias resulting from the starting full α-helix conformation. In the pAla, pVal and pSer, the end residues are excluded, whereas only glycine residues are considered for pGly. Figures 3.5 (a), (c), (e) and (g) give the results from pAla, pVal, pSer and pGly simulations, respectively, in explicit water molecules. The α-helix (lower left quadrant) and β–sheet (upper left quadrant) regions of pAla-WL, pVal-W2 and pSer-W2 can be clearly observed in these plots. Figure 3.5 (g) represents a typical symmetric plot of glycine. In *vacuum*, the α-helix regions of pAla and pVal are highly populated (Figures 3.5 (b) and (d)), which indicates that the initial helical conformation is highly stable. The plot of pSer-V2 shows more transitions to the β-sheet region compared to pAla and pVal. The Ramachandran map of pGly-V1 is not symmetric as the map of pGly-W1 in solution, because the upper-right hand quadrant is not populated.

Changes in the helical content of the polypeptides are monitored during the course of the simulations. The broad α-helical regions observed in the

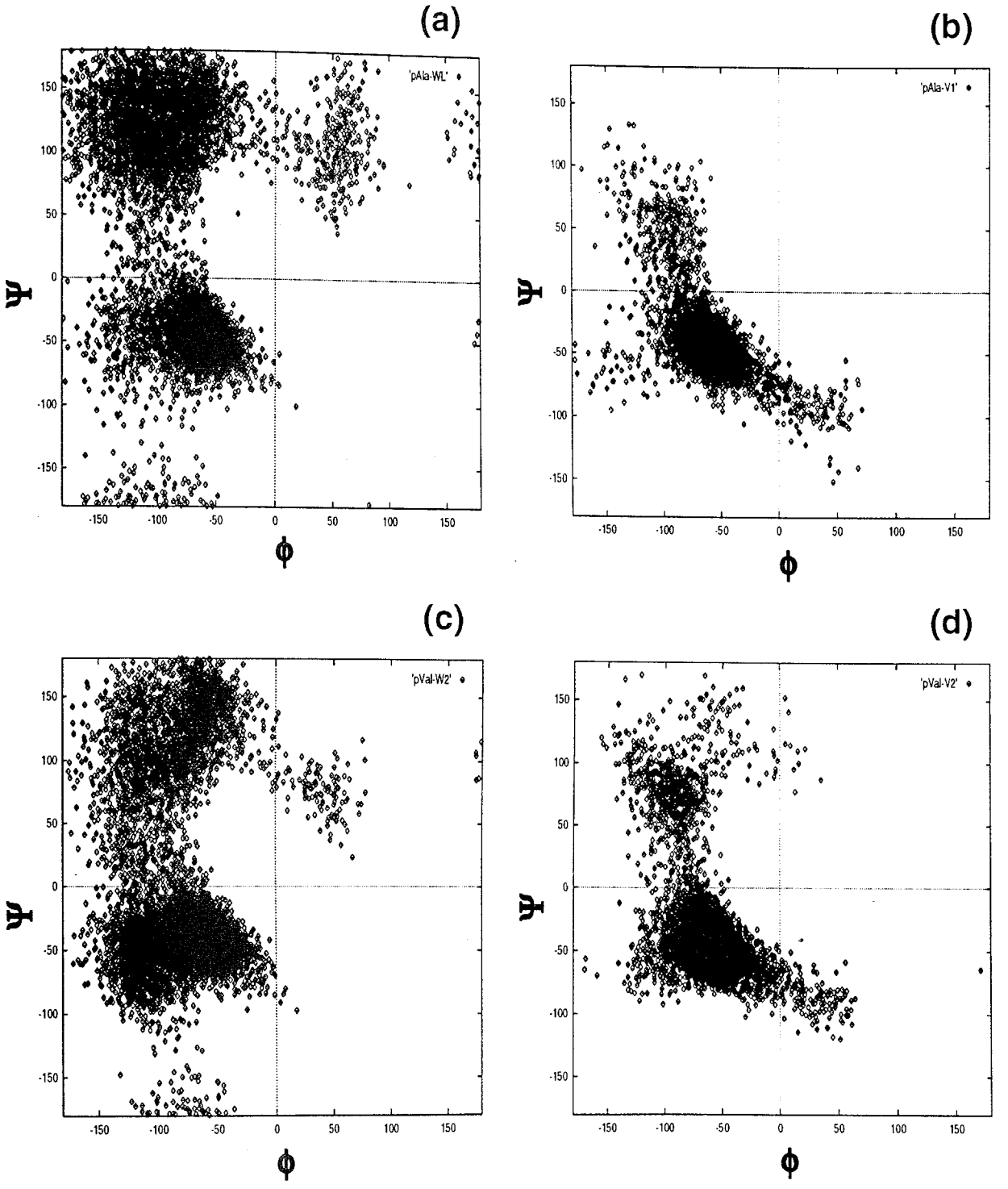


FIGURE 3.5. Ramachandran plots showing all dihedral pairs during the final 200 ps of the simulations: (a) pAla–WL, (b) pAla–V1, (c) pVal–W2, (d) pVal–V2, (e) pSer–W2, (f) pSer–V2, (g) pGly–W1 and (h) pGly–V1.

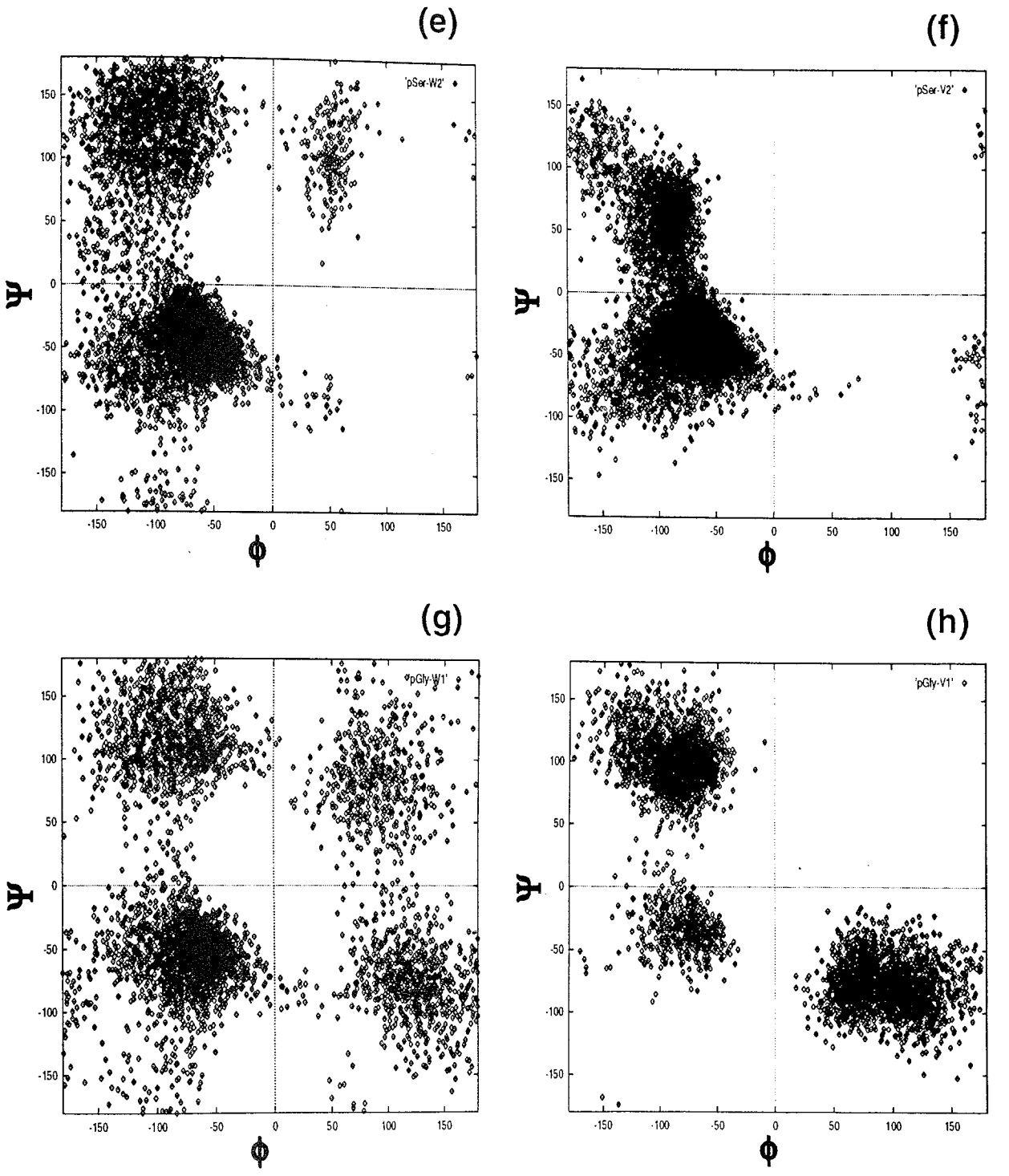


FIGURE 3.5. (continued)

Ramachandran maps of polyalanine, polyvaline and polyserine can be described by the region of $-100^{o} \leq \phi \leq -25^{o}$ and $-80^{o} \leq \psi \leq -5^{o}$, in conformity with the work of Daggett and Levitt.[74] The percentage of residues that are locally helical are plotted as a function of time in Figure 3.6 (a)-(p) for each run. The N-terminal and C-terminal residues are not considered in the calculation of the percentages because their respective ϕ and ψ angles cannot be uniquely determined. A 100 per cent helix conformation refers to a structure with 11 residues with ϕ and ψ angles in the defined helical region. All simulations carried out in water except pVal-W1 shows drops in local helix content with time.

Figures 3.6 (a)-(d) give helix content trajectories for polyalanine in water. The first three simulations (W1,W2 and W3) start from the same initial configuration and exhibit completely different behavior in time, after the equilibration periods which are not shown. In the run pAla-W3, the local helix content drops drastically at about 100 ps and stays below 20 per cent for the rest of the trajectory. This behavior seems unexpected for polyalanine, since alanine is a good α-helix former. Analysis of the three-dimensional structures of pAla in the simulation box indicates that the helix axis of polyalanine changes its direction in time and does not lie along the z-axis during the whole trajectory. The edges of the box along x and y directions are not sufficiently large to hydrate the helix and it interacts with its periodic images. To observe the effect of box size, i.e. the number of solvent molecules, a box with longer edges along x and y directions is generated and the simulation pAla-WL is performed. Starting with the same random number of pAla-W3, which determines the initial Gaussian velocity distribution, pAla-WL exhibits higher helix content compared to pAla-W3.

Figures 3.6 (e)-(g) represent the simulation results of polyvaline in water. Run pVal-W1 shows a steady local helix percentage of 80 to 90 throughout the whole trajectory, whereas the helix contents of pVal-W2 and pVal-W3 decrease in time. pVal-W1 seems to get trapped in the helical minimum conformation. The helix content trajectories of polyserine in water given in Figures 3.6 (h)-(j) display various degrees of unfolding with time. pSer-W1 exhibits a similar dynamical behavior to pVal-W1. pGly-W1 and pGly-W2 (Figures 3.6 (k) and (l)) show large decreases in local helix percentage with time. The helix content of pGly falls to around 20 per cent at the end of both simulations. This behavior of polyglycine in solution can be rationalized by the fact that glycine is a helix breaker.

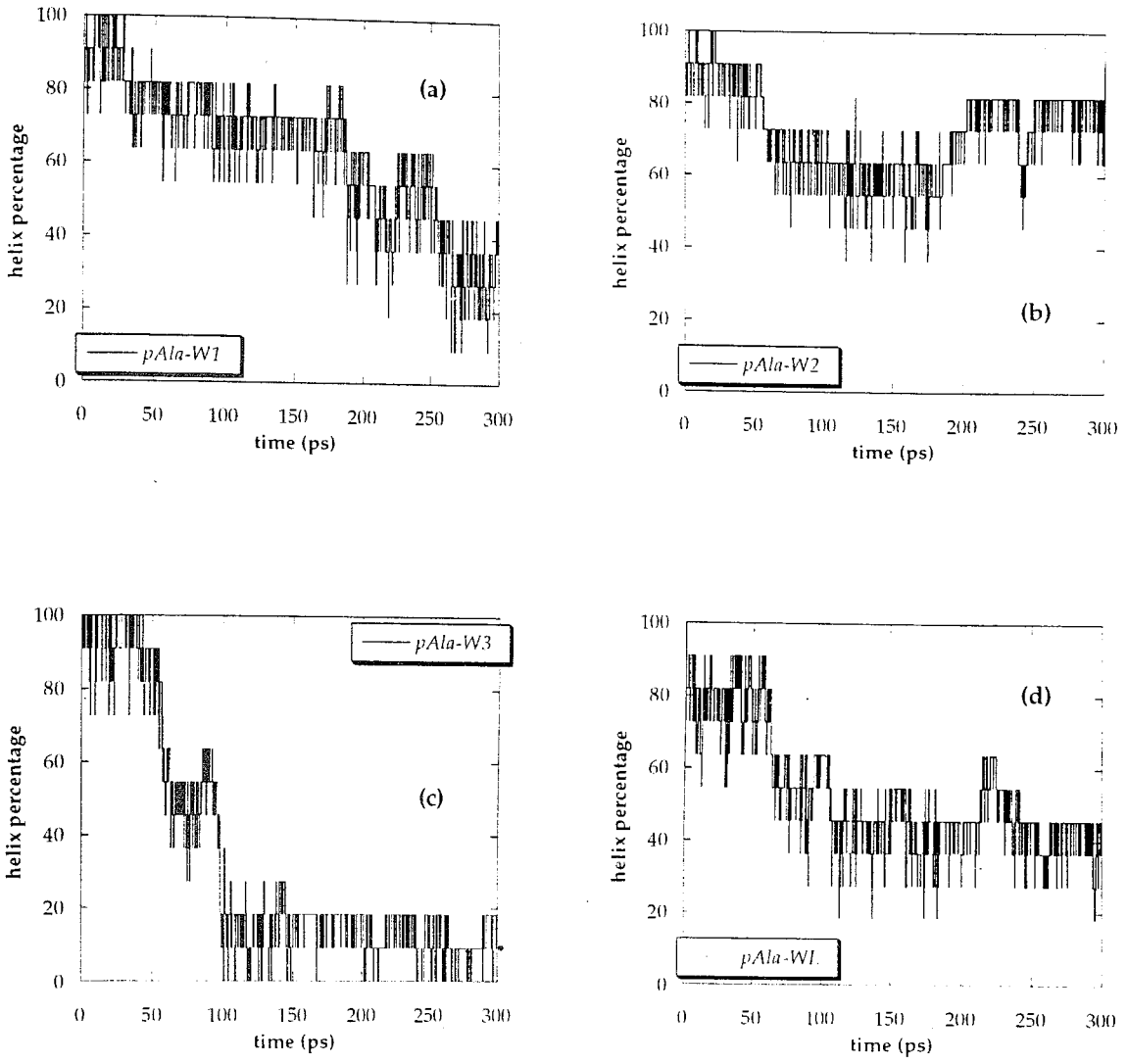
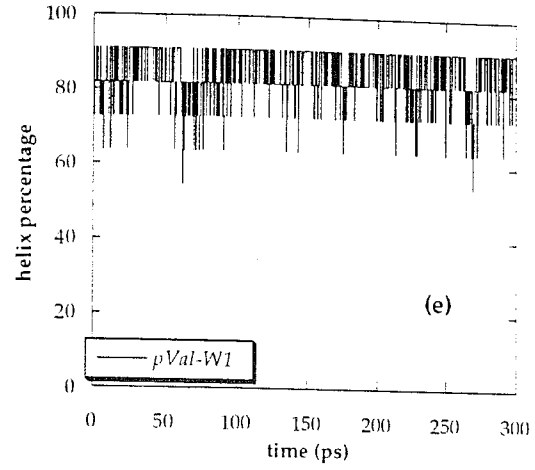
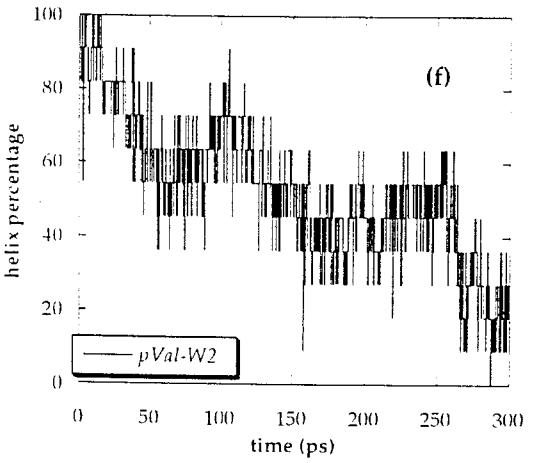
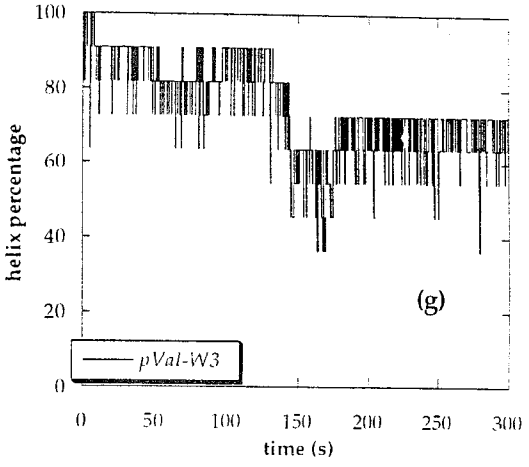
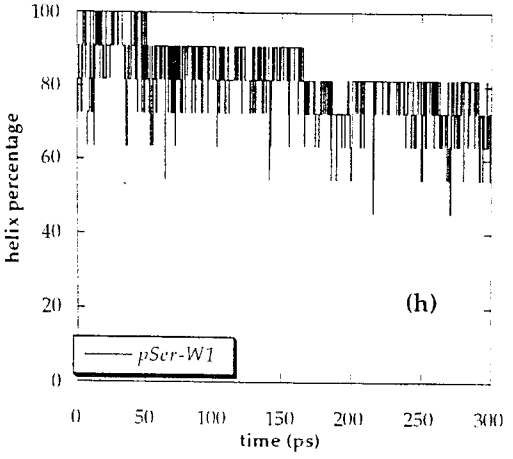


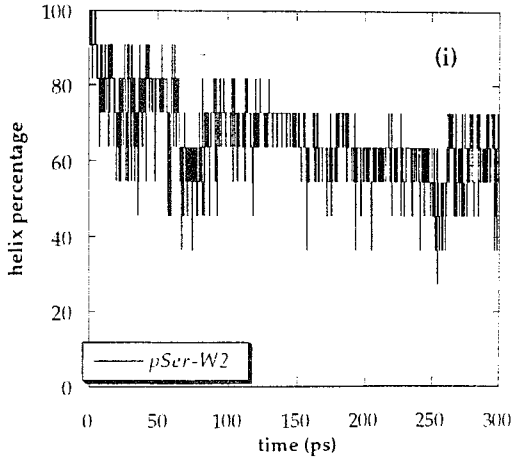
FIGURE 3.6. Local helix percentage trajectories of polypeptides pAla, pVal, pSer and pGly in water and in *vacuum*. Results of different runs in Table 3.2 are displayed, as indicated in the legends of (a)-(p). (A residue is considered locally helical, if $-100° < \phi < -25°$ and $-80° < \psi < -5°$)











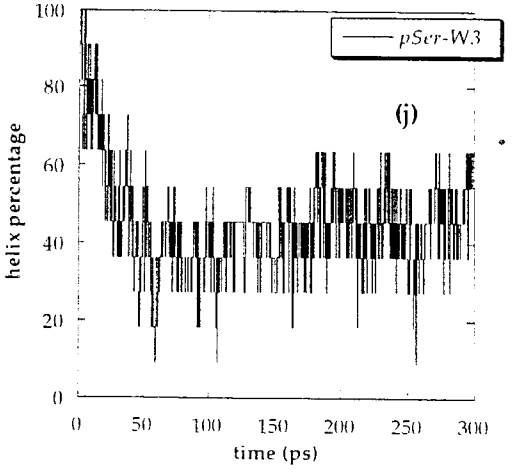


FIGURE 3.6. (continued)

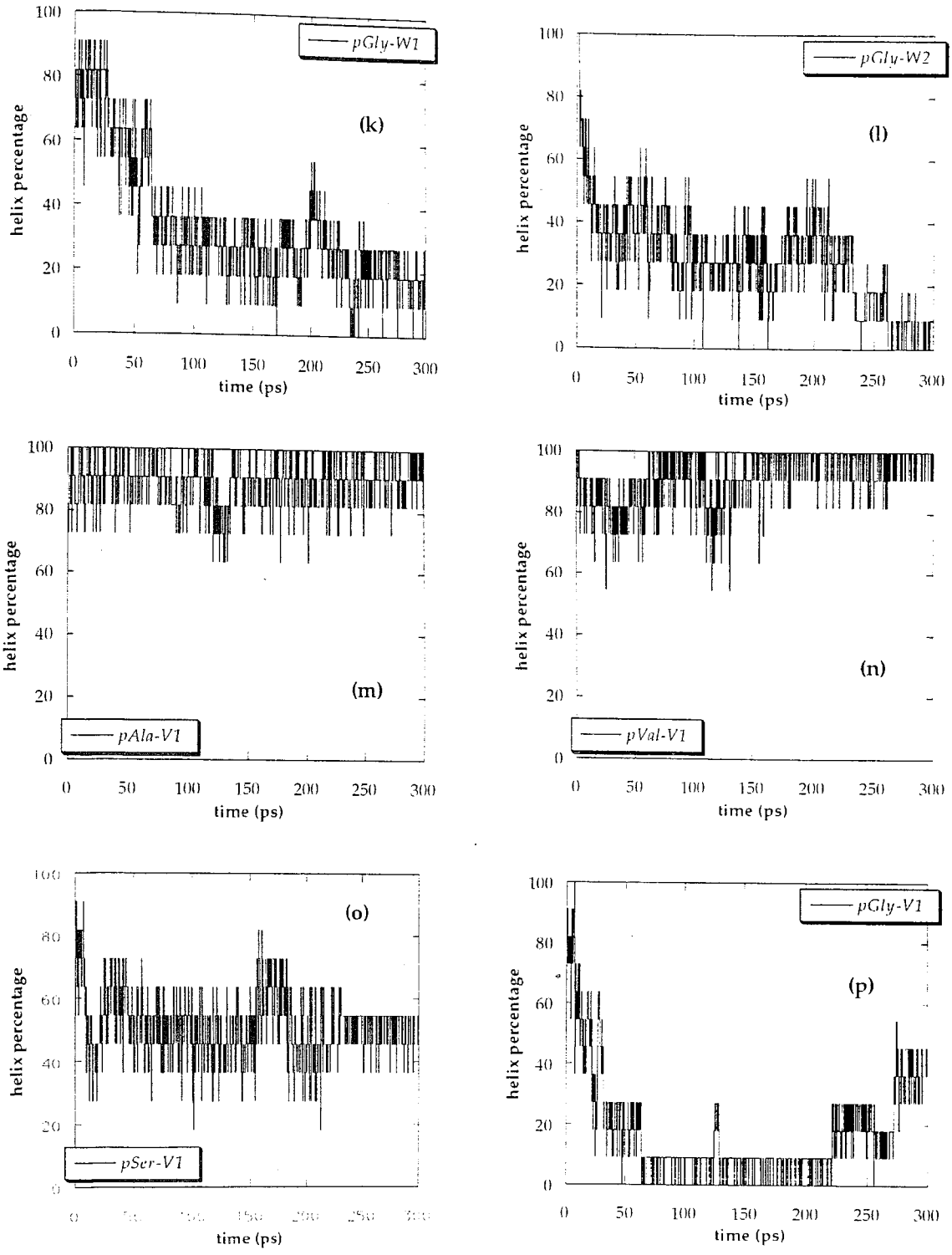


FIGURE 3.6. (continued)

Figures 3.6 (m)-(p) represent the results of *vacuum* trajectories. Only the runs indicated by V1 are presented for each model peptide, inasmuch as reproducible trends are observed in V1 and V2 simulations. pAla and pVal stay close to full helical conformation, whereas pSer and pGly show decrease in helix content with time.

To get an overall view of helical structure evolutions for each model peptide, the helix fractions of the different trajectories are averaged out. The results are displayed in Figures 3.7 (a)-(d) for water and vacuum simulations, with a time spacing of 1.5 ps between successive data points. pAla-W3, pVal-W1 and pSer-W1 are not included in the averages. When the average helix content trajectories of peptides are analyzed, pAla and pVal show a significant decrease in helix content in water compared to *vacuum*. In contrast, pSer and pGly exhibit similar helix contents in solution and in *vacuum*.

As a summary of water simulation results, a helix content scale of pAla-W $\approx$ pVal-W $\approx$ pSer-W > pGly-W is obtained. In vacuum, pAla and pVal display the highest helix contents and the following scale is observed: pAla-V $\approx$ pVal-V > pSer-V > pGly-V.

The fraction of time a residue stays locally helical during a complete trajectory is presented as a function of residue number along the peptide in Figure 3.8 (a)-(f). The same criteria of ϕ,ψ angle pairs used in determining the helix content trajectories is applied. Helicity of end residues cannot be determined as mentioned before. Figures 3.8 (a)-(d) display results for all the simulations in water. A common feature of pAla, pVal and pSer is that the end residues of the peptides spend less time in helix conformation than the middle residues, with the C-terminal fraying more compared to the N-terminal. On the contrary, the middle residues of pGly-W exhibit lower helix percentage. This behavior of pGly is expected since it contains four alanine residues at each end, which have helical propensity. Averages of the different runs for each peptide model are presented in Figures 3.8 (e) and (f) for water and *vacuum* simulations, respectively. As in the calculations of average helix content trajectories, simulations pAla-W3, pVal-W1 and pSer-W1 are not included while taking the averages.

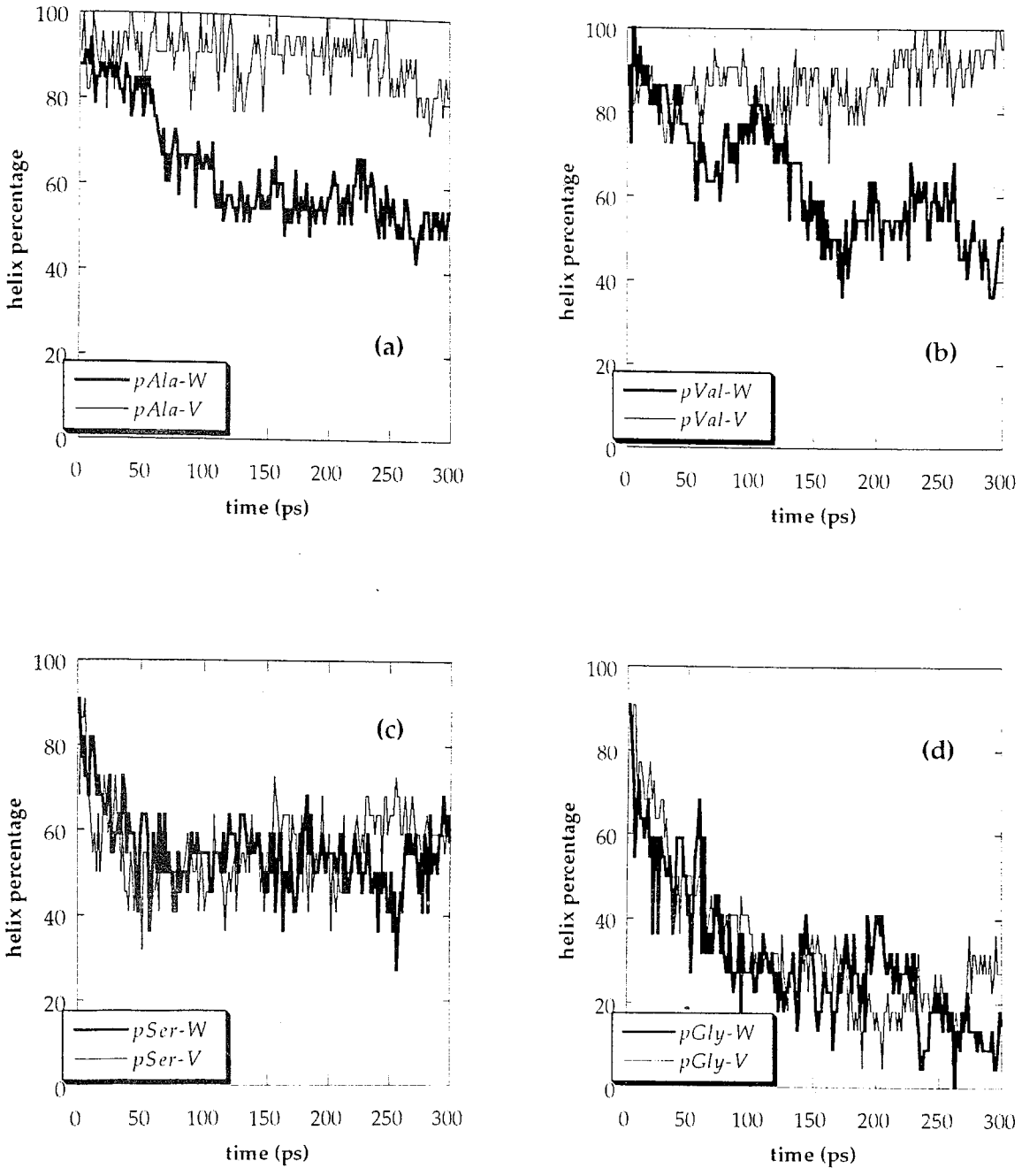


FIGURE 3.7. Comparison of average helix content trajectories in water and in *vacuum* for (a) pAla, (b) pVal, (c) pSer, and (d) pGly.

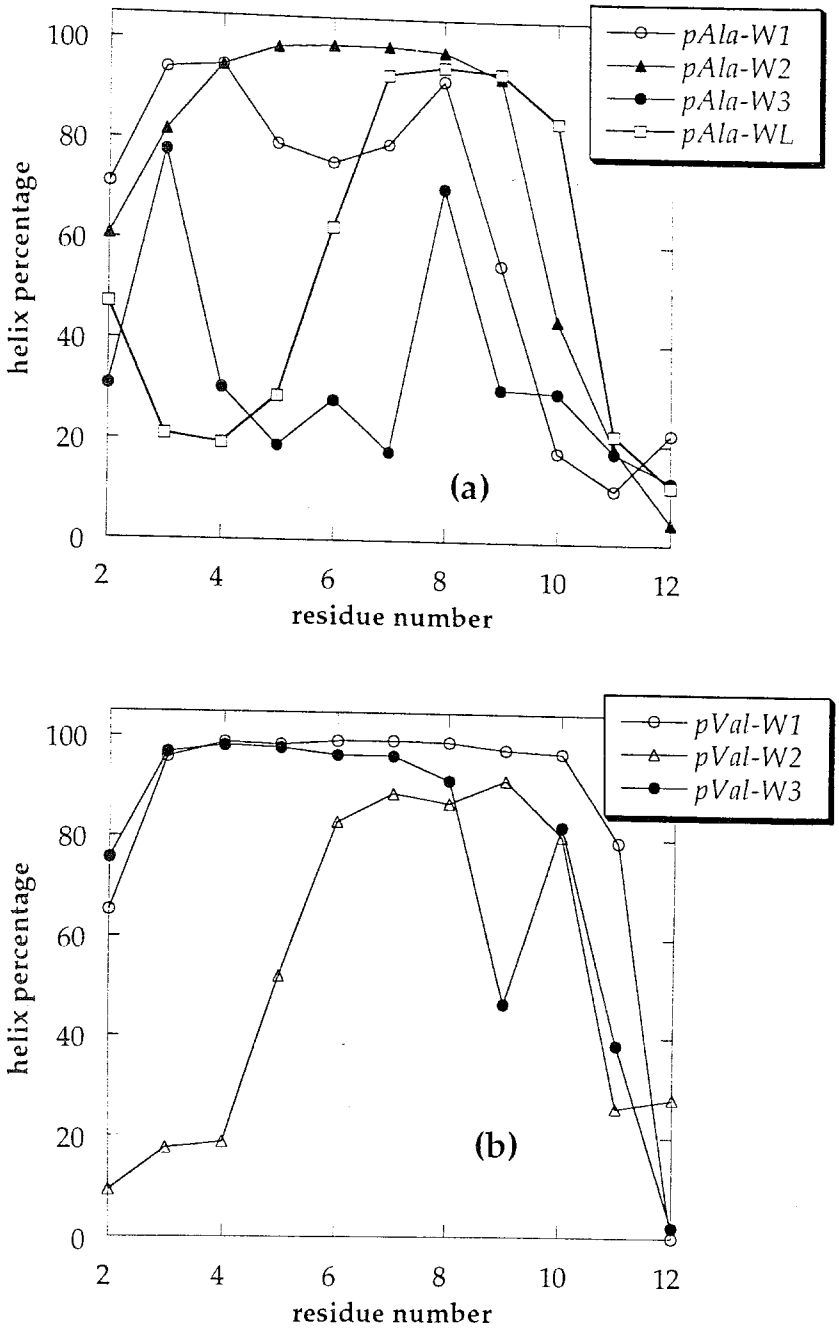
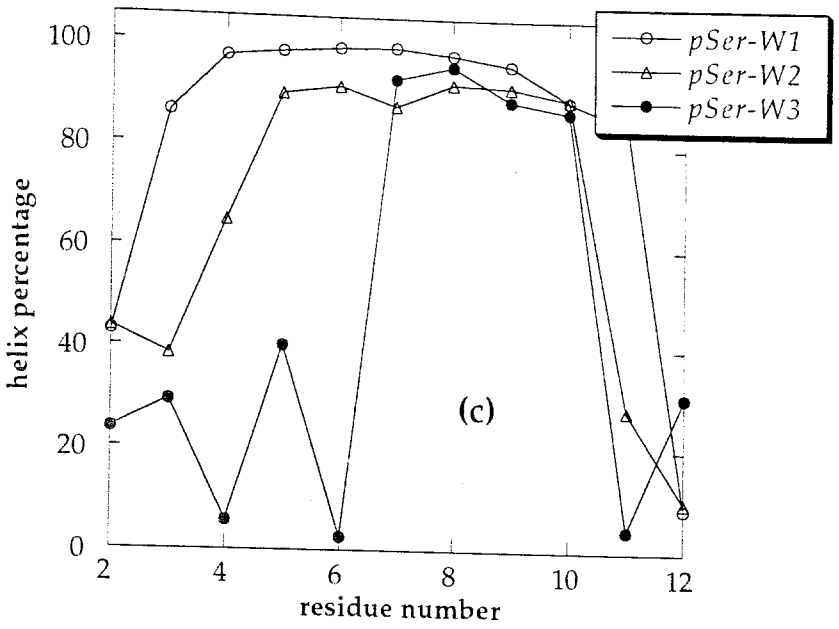


FIGURE 3.8. Percentage of local helical structure as a function of residue number during 300 ps trajectories. Results from different simulations of (a) pAla-W, (b) pVal-W, (c) pSer-W, (d) pGly-W and comparison of averages for all model peptides (e) in water and (f) in *vacuum* are displayed. (A residue is considered locally helical if $-100° < \phi < -25°$ and $-80° < \psi < -5°$)



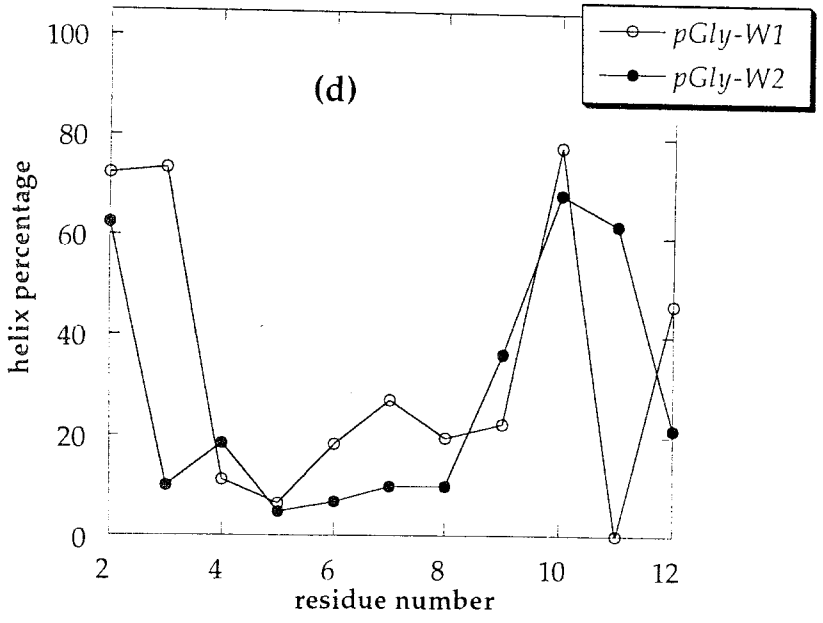


FIGURE 3.8. (continued)

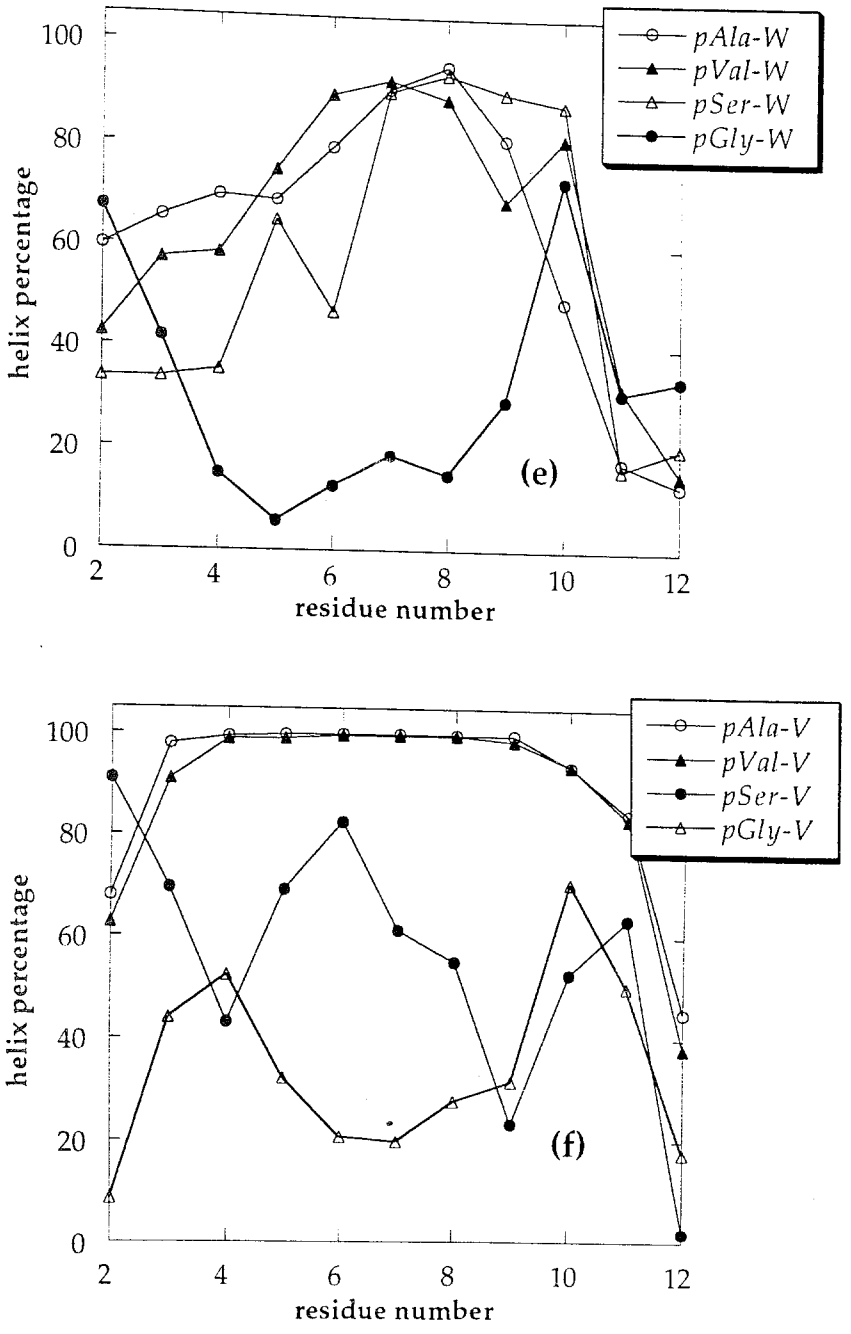


FIGURE 3.8. (continued)

3.3.2. Equilibrium Distribution of Water around Side-chains

The equilibrium distribution of water molecules around the side chain atoms of alanine, valine and serine are investigated to elucidate the differences in the hydration patterns around these side chains. The geometry of side-chain-water interactions is adopted from the work of J.M. Goodfellow and co-workers[81], who have analyzed the solvation of all amino acids[82] and main chain atoms[83] using more than 20 high resolution protein crystal structures from the Brookhaven Databank.[84] Figure 3.9 is a representation of the spherical polar coordinates (r, θ, ϕ) used in the analysis of water distribution around atom A, placed at the origin. In this geometry, atoms A, B, and C lie in the xz-plane. B and C are the atoms connecting the side-chain and the backbone (see Table 3.3)

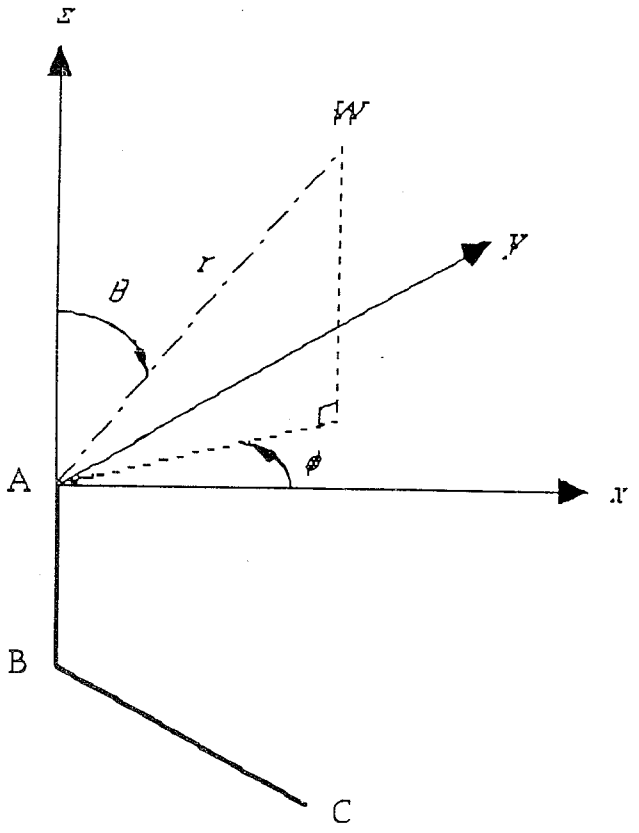


FIGURE 3.9. Geometrical representation of the spherical polar coordinates (r, θ, ϕ) used in calculating the distributions of solvent molecules (W) around atom A.

The position of the oxygen atoms of water molecules (represented by point W in Figure 3.9) is used in calculating the (r, θ, ϕ) distributions around atom A. The x, y and z coordinates of the (water) oxygen atom W are $x = r \sin\theta \cos\phi$, $y = r \sin\theta \sin\phi$ and $z = r \cos\theta$. In this geometrical representation, $\phi = 0^o$, 90^o, 180^o and 270^o lie along the x-axis, y-axis, - x-axis and - y-axis, respectively.

Table 3.3 summarizes the atoms that are fixed as A, B and C in the analysis of water distribution around the beta carbon ($C\beta$) of alanine, valine and serine, gamma carbons ($C\gamma1$ and $C\gamma2$) of valine and the oxygen atom ($O\gamma$) of serine. Only the residues that are locally helical are considered in the analysis. The normalized equilibrium distributions of r, θ and ϕ from independent simulations give reproducible results for each residue type. Two runs (indicated in Table 3.3) are found to be satisfactory for the analysis of average distributions around each amino acid.

TABLE 3.3. The Atoms Fixed as A, B and C in the Analysis of Water Distributions around Amino Acids

Residue	Atom A	Atom B	Atom C	Runs Analyzed
alanine	$C\beta$	$C\alpha$	N	pAla-W1, pAla-W2
valine	$C\beta$	$C\alpha$	N	pVal-W1, pVal-W2
valine	$C\gamma1$	$C\gamma2$	$C\beta$	pVal-W1, pVal-W2
valine	$C\gamma2$	$C\gamma1$	$C\beta$	pVal-W1, pVal-W2
serine	$C\beta$	$C\alpha$	N	pSer-W2, pSer-W3
serine	$O\gamma$	$C\beta$	$C\alpha$	pSer-W2, pSer-W3

The radial distribution function $g_r(r)$, is proportional to the probability of finding a solvent oxygen atom at a distance r from atom A. $g_r(r)$ is calculated from

$$g_r(r) = < n(r) > / r^2 N_{asy} \qquad (3.4)$$

where n(r) is the number of oxygen molecules located in a thin spherical shell of thickness Δr = 0.2 Å at a distance r from each side-chain atom A of the locally helical residues at each snapshot of the trajectory. Angular brackets represent a time average over different snapshots. Normalization of the distributions is accomplished first by division by r^2 and then by the asymptotic value N_{asy} of the curves at r = 10 Å.

The normalized θ distribution around atom A, indicated by $g_\theta(\theta)$, is calculated by counting the number of oxygen atoms which have r < 5 Å and fall into $\Delta\theta = 10^o$ grids. Normalization is done by division by the total number of occurrences of oxygen atoms. Similarly, the normalized ϕ distributions $g_\phi(\phi)$ are calculated for oxygen atoms with r < 5 Å by considering $\Delta\phi = 10^o$ grids and by using the same normalization as in $g_\theta(\theta)$. Figure 3.10 displays the $g_r(r)$ distributions around the side chain atoms of different residues. Figures 3.11 and 3.12 give the results of $g_\theta(\theta)$ and $g_\phi(\phi)$ distributions, respectively. Solvation of each amino acid will be separately discussed in the following.

3.3.2.1. Alanine. Figure 3.10 (a) displays two distinct radial distribution functions around C^β atoms of alanine residues. Curve indicated by pAla is the average distribution when all the solvent molecules are considered in the analysis. The second distribution, pAla(np), is calculated by excluding all water molecules which are located within 3.5 Å of a polar atom, i.e. a carbonyl oxygen or backbone nitrogen. This latter representation gives a nonpolar distribution by excluding the water molecules that may be hydrogen bonded to the polar atoms.

Both r distributions show peaks at about r = 3.3 Å, which corresponds to the first hydration shell around C^β. These peaks were found to be located between 3.8 and 4.0 Å in the crystal structures studied by Walshaw et al.[81] This difference might arise from the potentials of GROMOS87, in which the CH_3 groups are taken as collapsed.

The θ distributions in Figure 3.11 (a) exhibit an excluded region between $\theta = 140^o$ and 180^o, which may result from steric overlap of the water molecules with the backbone atoms. The distributions have peaks at 85^o and 65^o when all and nonpolar atoms are considered, respectively. Similar peaks have been observed in the analysis of crystal structures.[81]

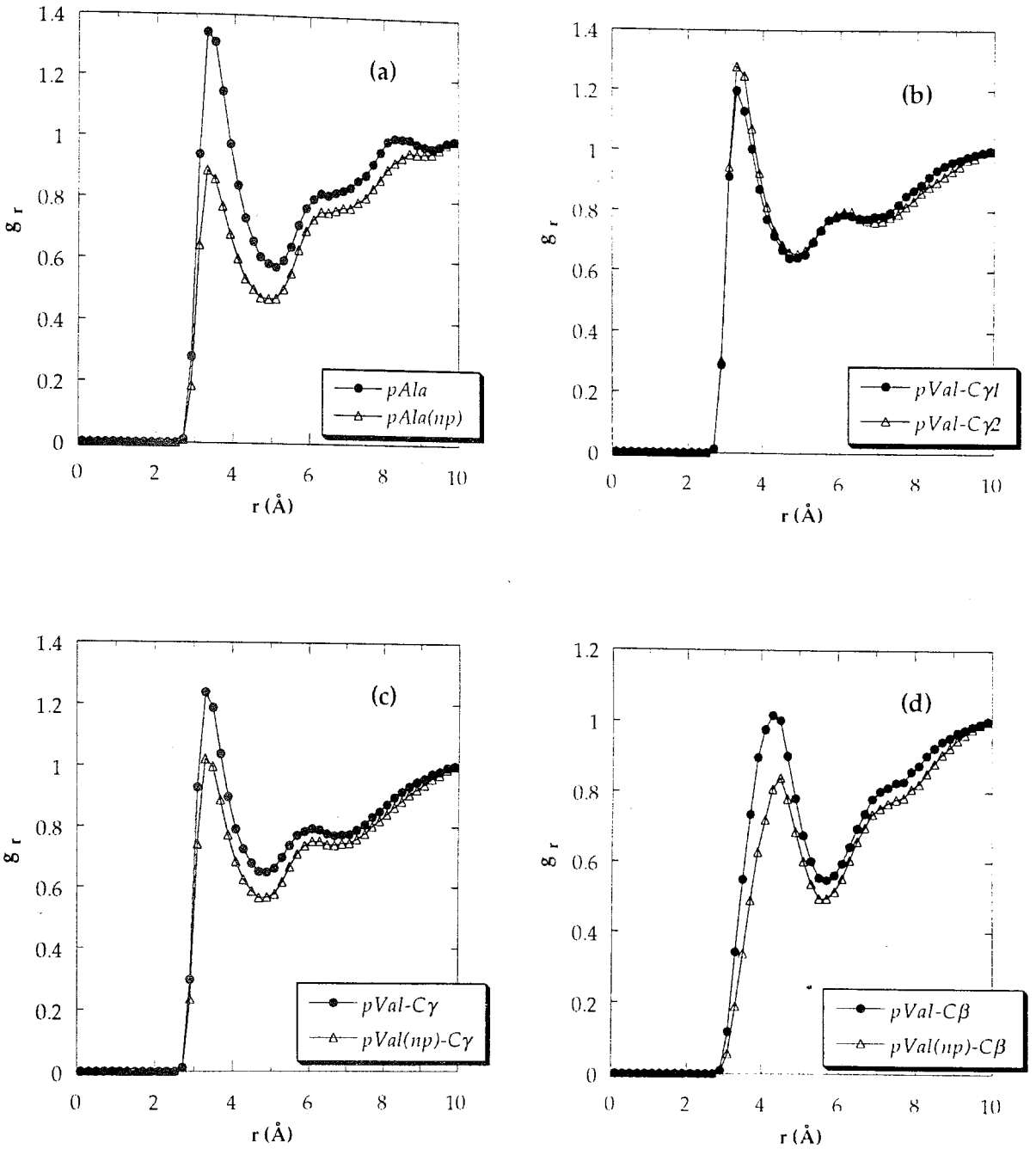


FIGURE 3.10. Normalized radial distribution functions, $g_r(r)$, of water molecules around side chain atoms: (a) alanine Cβ, (b) valine Cγ1 and Cγ2, (c) valine Cγ (average of Cγ1 and Cγ2), (d) valine Cβ, (e) serine Oγ, (f) serine Cβ, (g) alanine, valine and serine Cβ, and (h) alanine Cβ, valine Cγ, and serine Oγ. (*np* indicates the distribution of nonpolar water molecules, which are not in the 3.5 Å neighborhood of any backbone nitrogen or carbonyl oxygen)

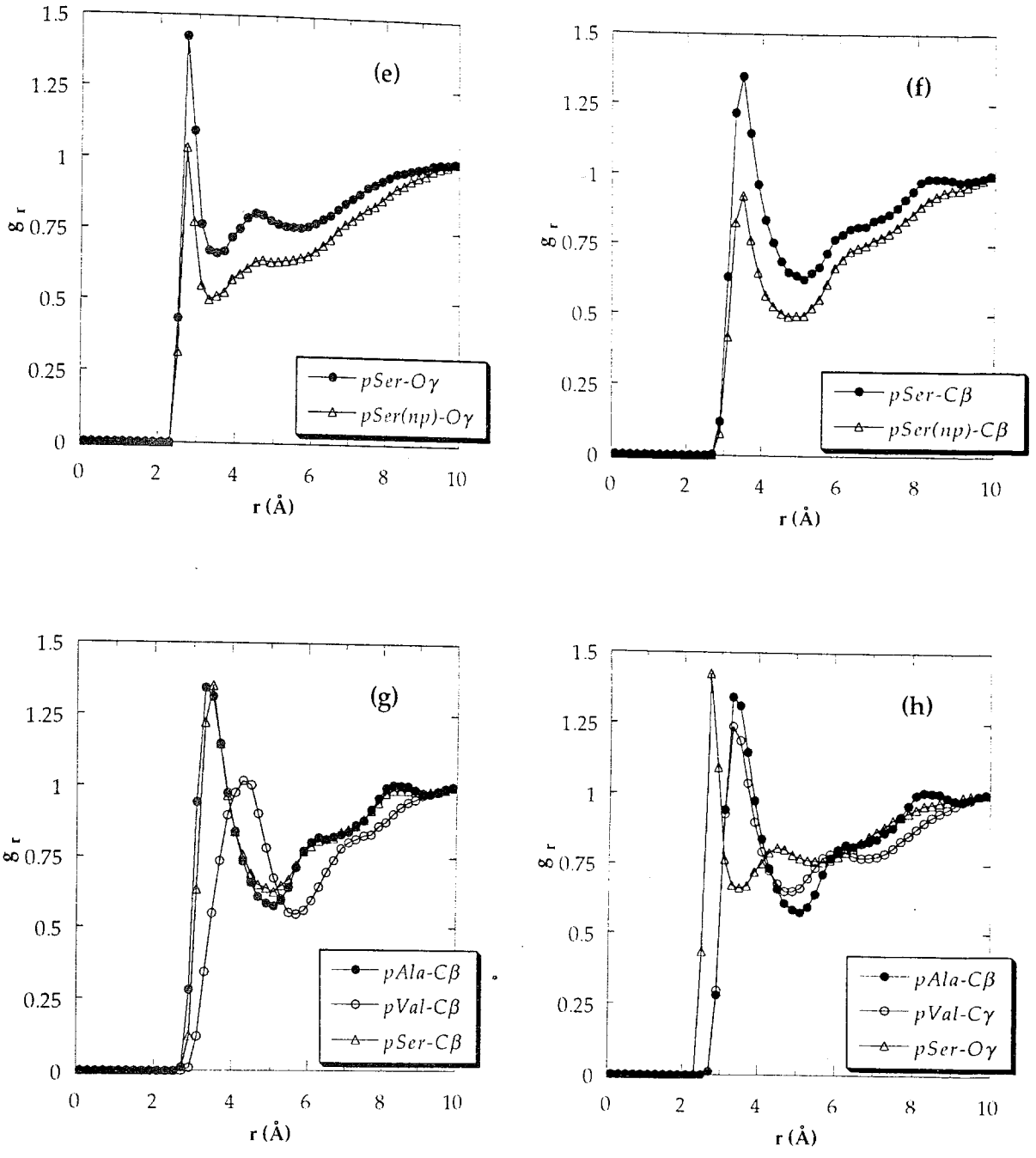


FIGURE 3.10. (continued)

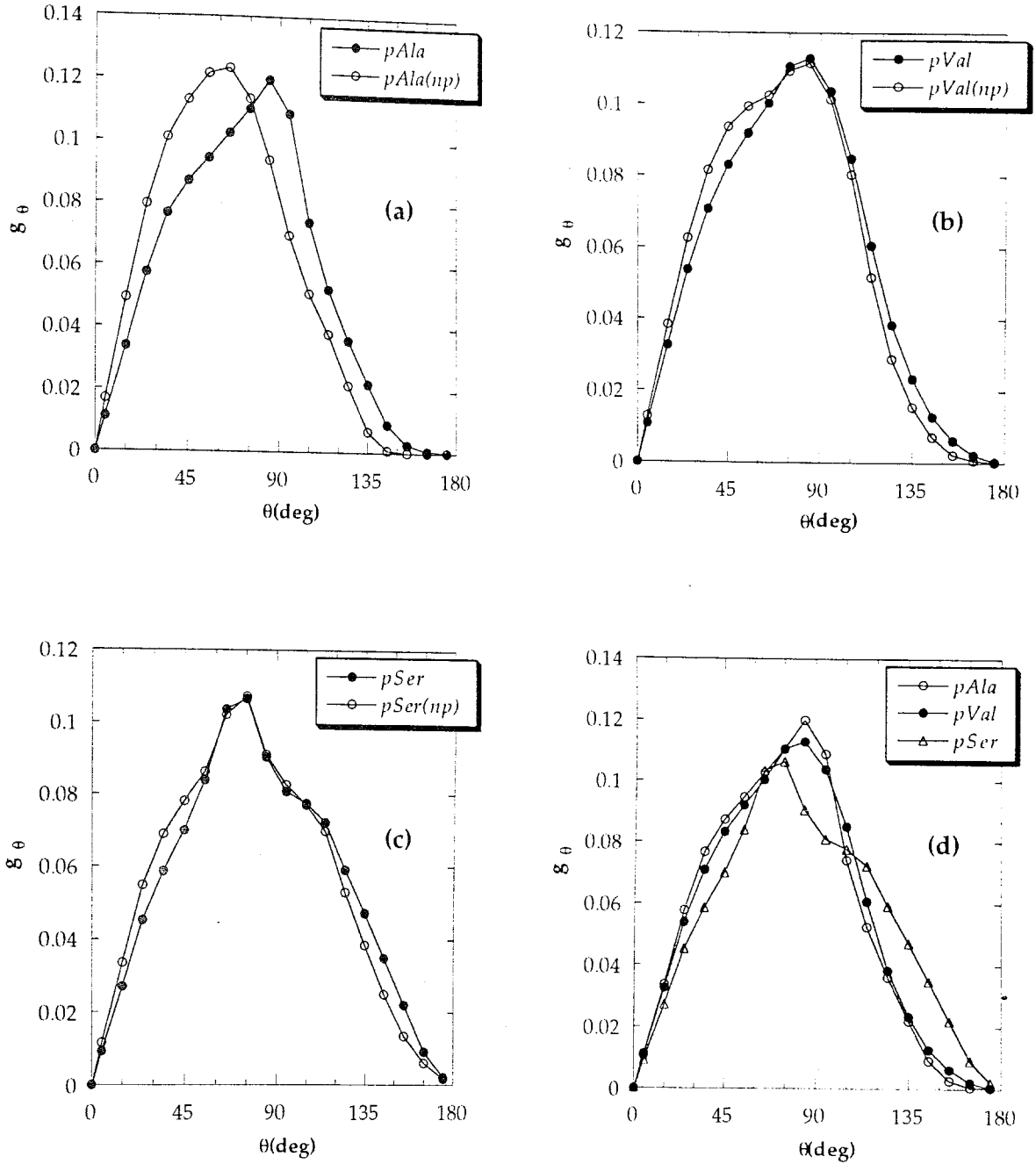


FIGURE 3.11. Normalized θ angle distributions , $g_\theta(\theta)$, of water molecules around side chain atoms: (a) alanine Cβ, (b) valine Cγ, (c) serine Oγ, (d) valine Cβ, valine Cγ, and serine Oγ. (*np* indicates the distribution of nonpolar water molecules, which are not in the 3.5 Å neighborhood of any backbone nitrogen or carbonyl oxygen.)

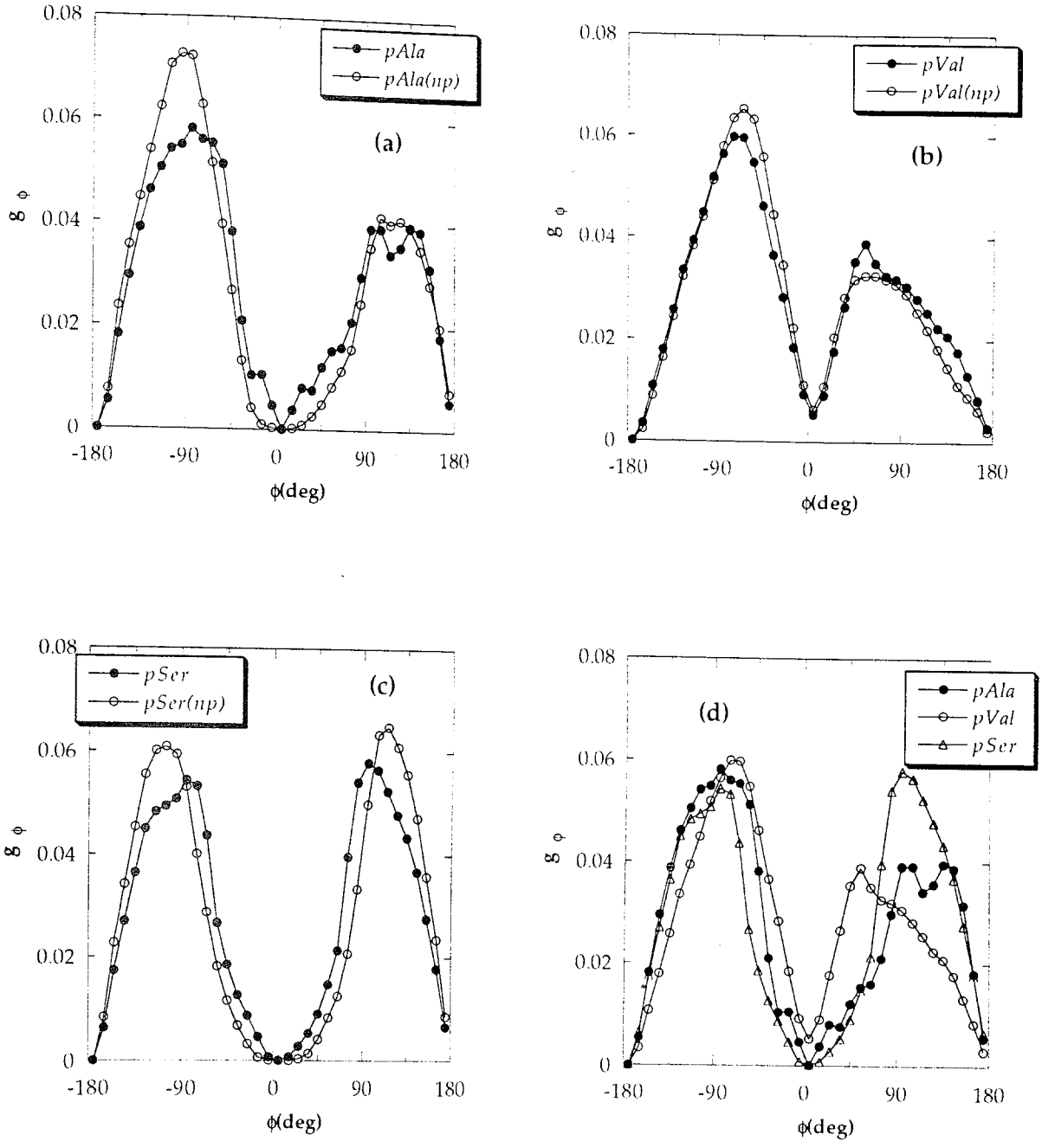


FIGURE 3.12. Normalized ϕ angle distributions , $g_\phi(\phi)$, of water molecules around side chain atoms: (a) alanine Cβ, (b) valine Cγ, (c) serine Oγ, (d) valine Cβ, valine Cγ, and serine Oγ. (*np* indicates the distribution of nonpolar water molecules, which are not in the 3.5 Å neighborhood of any backbone nitrogen or carbonyl oxygen.)

Distribution of the azimuthal angle ϕ in Figure 3.12 (a) shows more water molecules at negative values, i.e. in the front plane defined by C^{β}, C^{α} and N atoms, in conformity with the results from crystal structures. However, the excluded regions around $\phi_i = 0^o$ and 180^o (or -180^o) in our $g_{\phi}(\phi)$ plots, which apparently reflects the steric overlap with the $(NH)_{i+1}$ and $(CO)_{i-1}$ groups, respectively, cannot be clearly observed in the crystal structures.

3.3.2.2. Valine. In the analysis of crystal structures[81], the r, θ and ϕ distributions were calculated for $C^{\gamma 1}$ and $C^{\gamma 2}$ separately and differences were shown to exist among the two structures. Our results, on the contrary, indicate no significant difference between the solvation of the two gamma carbons. Figure 3.10 (b) shows the r distribution and the θ and ϕ distributions (not shown) are also identical. The $g_{\phi}(\phi)$ curves are mirror images of each other due to the definitions of ϕ angle around $C^{\gamma 1}$ and $C^{\gamma 2}$. The reason for the differences observed in the solvation of $C^{\gamma 1}$ and $C^{\gamma 2}$ atoms in crystal structures may be due to the inadequate number of water molecules analyzed.

Figures 3.10 (c) and (d) show the pair distributions around C^{γ} and C^{β} atoms of valine, respectively. The peak of first hydration shell of gamma carbons is located at r = 3.3 Å as in C^{β} of alanine. In crystal structures, this peak has been determined around 3.8 - 4.0 Å.[81] The first hydration shell around the beta carbon of valine is shifted to higher values of r, namely to r = 4.3 - 4.5 Å. This indicates that the C^{β} of valine is relatively protected, i.e. is less accessible to solvent, compared to C^{β} of alanine.

The $g_{\theta}(\theta)$ and $g_{\phi}(\phi)$ distributions (Figures 3.11(b) and 3.12 (b)) around the C^{γ} atoms of valine display similar features irrespective of the choice of all or nonpolar water molecules. $g_{\theta}(\theta)$ plots show peaks at 85^o and excluded regions between θ = 150 and 180^o.

3.3.2.3. Serine. The radial distribution functions around the O^{γ} and C^{β} atoms of serine in Figures 3.10 (e) and (f) display first hydration shells at 2.7 and 3.3 Å, respectively. The peak at 2.7 Å is a result of hydrogen bonding with water. The crystal structure data[82] show a broad r distribution between 2.5 and 3.5 Å around the O^{γ} atom, which is in contrast with the sharp peaks observed in Figure 3.10 (e).

The θ distribution around Oγ in Figure 3.11(c) displays a peak at 75° and no significant excluded region. This peak corresponds to the expected value of 70° for tetrahedral hybridization(interaction with water molecules). The distribution from crystal structures[82] shows a similar peak at 65° and a broad excluded region above 120°.

The ϕ distribution around Oγ in Figure 3.12 (c) is symmetric since a water molecule above the xz-plane cannot be distinguished from one below the plane if their θ value is the same. In contrast, the distributions around alanine and valine have been found to be nonsymmetric due to the chirality of the neighboring carbon atoms. A peak at 180° is observed in the symmetric distribution of crystal structures.

Figures 3.10 (g, h), 3.11 (d) and 3.12 (d) compare the $g_r(r)$, $g_\theta(\theta)$ and $g_\phi(\phi)$ distributions of alanine, valine and serine. The distributions around the Cβ atoms indicate that the beta carbons of alanine and serine are solvated in a similar pattern, whereas the Cβ of valine is less accessible to water. The gamma carbons of valine also show a similar hydration pattern to that of Cβ of alanine, whereas the water molecules are located quite close to the Oγ of serine due to hydrogen bonding.

3.3.3. Water Distribution around Hydrogen Bond Forming Groups

The stability of the backbone hydrogen bonds is inversely proportional to their accessibility to water molecules, which compete to form alternative hydrogen bonds with the backbone carbonyl and amide groups. The hydrogen bonds that are protected from water molecules are expected to be more stable which implies that the four residues in between the hydrogen bonding partners stay locally helical.

To observe the differences between the solvation of hydrogen bonds in the four peptides, pair distribution functions $g_r(r)$ of water molecules around the backbone carbonyl oxygen are drawn before and after the hydrogen bond

breaks. The closest hydrogen atom of each water molecule to the carbonyl oxygen is considered in the calculation of distributions. Backbone hydrogen bonds are considered intact if the distance between the i^{th} residue oxygen and the $(i+4)^{th}$ residue hydrogen atoms forming the H-bond is less than 2.3 Å, and broken otherwise. Seven carbonyl oxygens have $(i+4)^{th}$ bonding partners in each peptide. Thus, in pAla, pVal and pSer, seven hydrogen bonds are analyzed at each snapshot. In pGly, only five bonds that have at least two glycine residues in between are considered, i.e. the two hydrogen bonds that may form at the two ends of the peptide are not taken into account.

Pair distribution functions are calculated according to Equation (3.4), where N_{asy} is taken as the sum of instances that the hydrogen bonds between the carbonyl oxygens and the (i+4) th hydrogens are intact or not during the trajectory. Total number of instances for pAla, pVal and pSer is 7 bonds x 1200 snapshots = 8400 in each 300 ps simulation, and 5 bonds x 1200 snapshots = 6000 for pGly. These total numbers are divided into helical and non-helical (coil) instances, which are listed in Table 3.4 for the runs considered in this analysis.

TABLE 3.4. Instances of Helical and Coil Conformations in the Simulations

Run Analyzed	Helical Instances	Coil Instances	Total Instances
pAla-W1	3147	5253	8400
pAla-WL	2229	6171	8400
pVal-W2	1847	6553	8400
pVal-W3	3613	4787	8400
pSer-W1	4759	3641	8400
pSer-W2	3087	5313	8400
pGly-W1	0	6000	6000
pGly-W2	0	6000	6000

Figures 3.13 (a)-(d) compare the hydration patterns around the carbonyl oxygen before and after the hydrogen bond breaks, which are termed as helical and coil conformations, respectively, and are represented by the letters h and c, in the figure legends. Intact hydrogen bonds in pGly are not observed during the whole trajectory. Thus, only the equilibrium distribution for the coil conformation

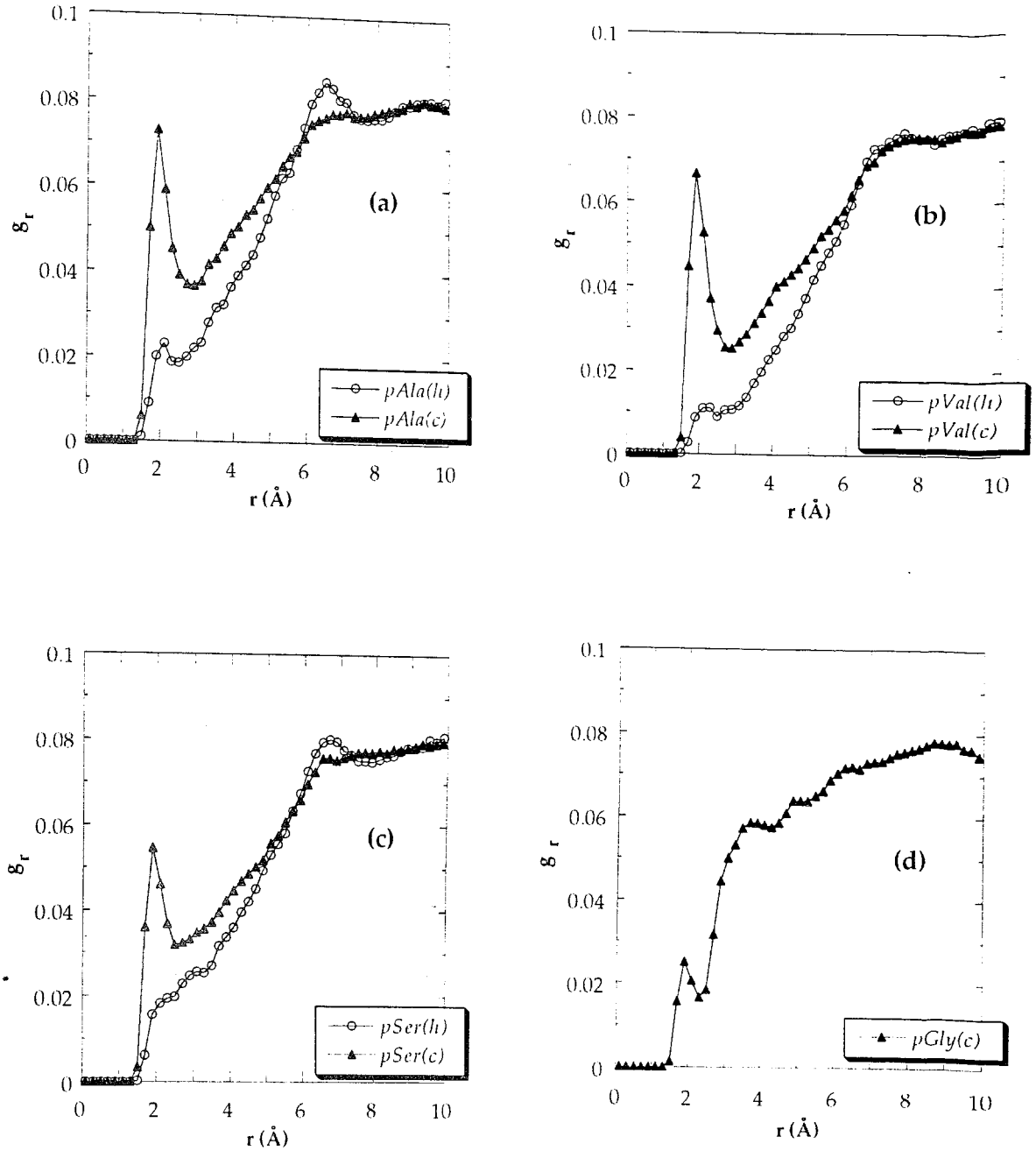
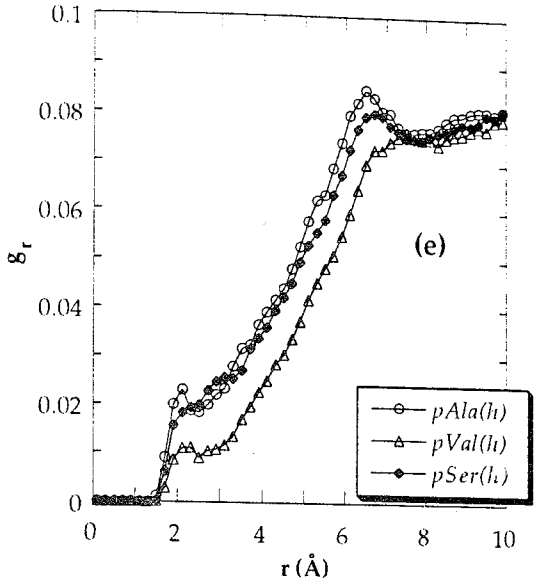
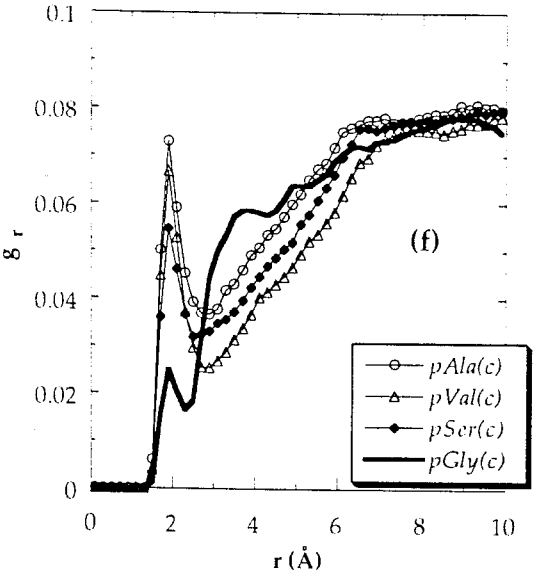
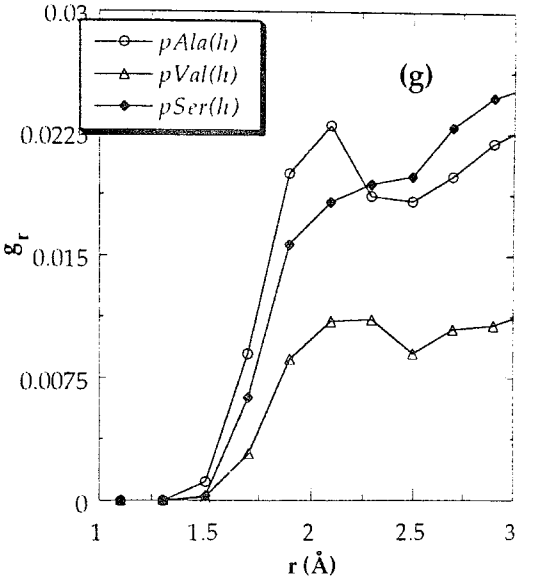


FIGURE 3.13. Comparison of normalized radial distribution functions, $g_r(r)$, of water molecules around the carbonyl oxygens of backbone hydrogen bonds, before and after the hydrogen bond breaks. *h* and *c* represent helix (intact H-bonds) and coil (broken H-bonds) states, respectively. Results are displayed for the helix and coil states of (a) pAla, (b) pVal, (c) pSer, and (d) pGly, and for comparison of (e, g) pAla, pVal and pSer in helix state and (f, h) pAla, pVal, pSer and pGly in coil state.







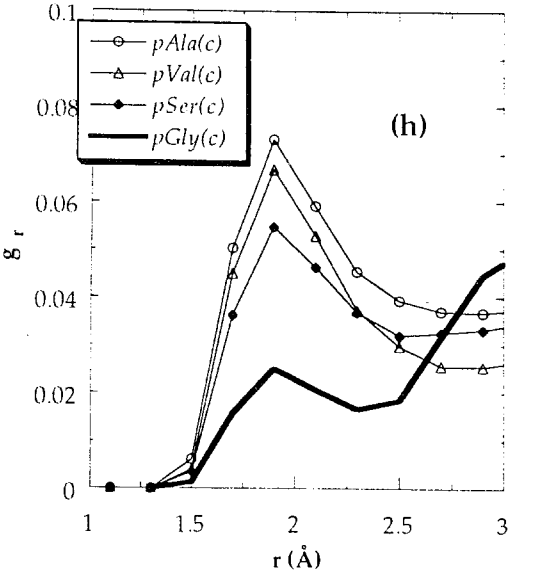


FIGURE 3.13. (continued)

is represented. Each curve is an average of the two simulations chosen for that peptide, given in Table 3.4. It can be observed in pAla, pVal and pSer that more hydrogen bonds with water molecules form after the backbone hydrogen bonds break and the carbonyl oxygens become more exposed to solvent. The small peak at r = 2.1 Å in pAla(h) moves to 1.9 Å and becomes highly pronounced when pAla(c) is considered. A similar trend is observed for pVal and pSer.

The helix and coil distributions of the peptides are plotted separately for comparison in Figures 3.13 (e) and (f). All distributions reach the same $g_r(r)$ value at r = 10 Å, which indicates that the normalization procedure applied is suitable for comparison of results from runs carried out with different peptide models. Figures 3.13 (g) and (h) give the same comparison curves for $1 < r < 3$ Å. In these short range plots, it can be observed that intact hydrogen bonds in pAla are more exposed to water than in pSer and pVal, and bonds are the most hindered to solvent in pVal. Water distribution around carbonyl oxygen after the hydrogen bonds break shows the following order of hydrogen bond formation with water: pAla > pVal > pSer > pGly.

The angular distribution of the water hydrogens around the carbonyl oxygen is plotted considering only those hydrogen atoms that are less than 3.5 Å from the oxygens. Angle θ is defined as in Figure 3.9, provided that A, B and W atoms are replaced by carbonyl oxygen, carbon and water hydrogen atoms, respectively. Figures 3.14 (a) and (b) summarize the $g_\theta(\theta)$ distributions. It can be clearly observed that there are less water molecules in the region $r < 3.5$ Å around the carbonyl oxygens when the hydrogen bonds are intact. This difference is the most pronounced for pVal(h). The peaks in θ distributions are located at 82.5° and 67.5° for helical and coiled residues, respectively.

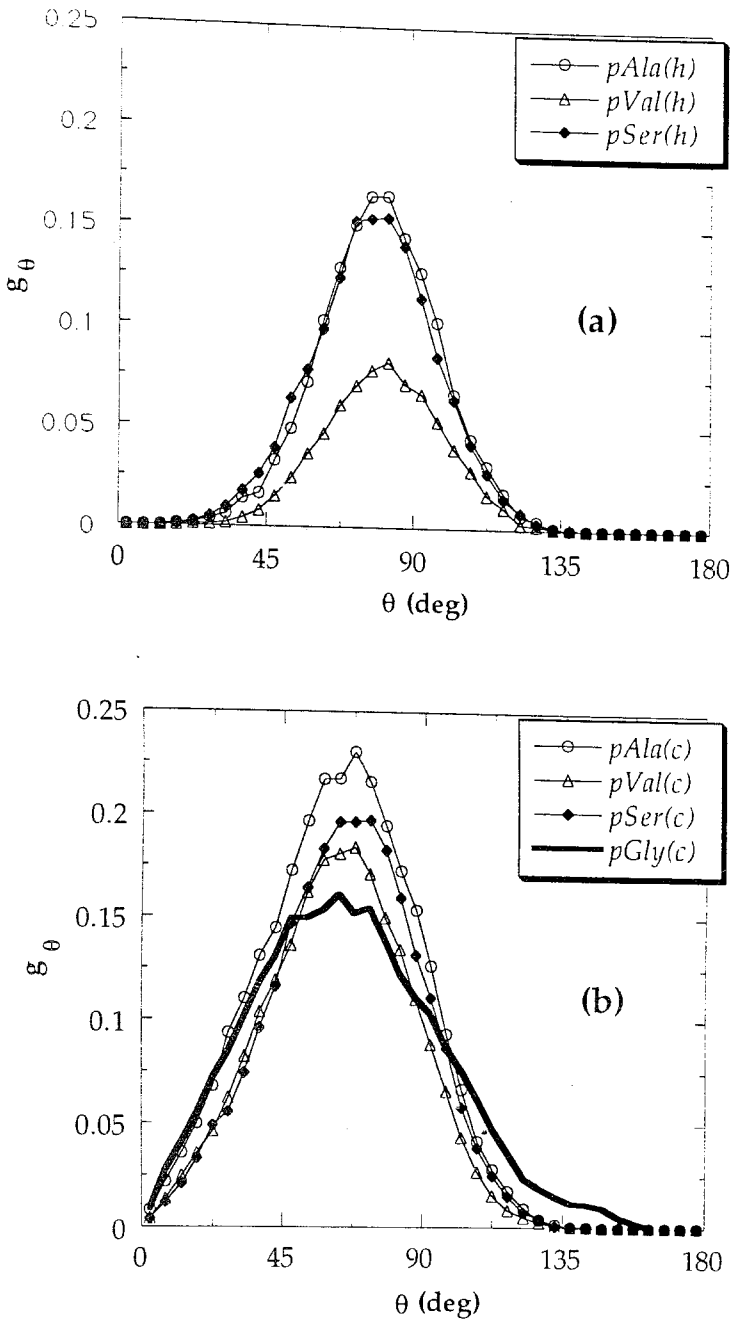


FIGURE 3.14. Comparison of normalized θ distributions, $g_\theta(\theta)$, of water molecules around carbonyl oxygens of backbone hydrogen bonds, before and after the hydrogen bond breaks. *h* and *c* represent helix (intact H-bonds) and coil (broken H-bonds) states, respectively. Results are displayed for (a) pAla, pVal and pSer in helix state and (b) pAla, pVal, pSer and pGly in coil state.

4. CONCLUSIONS AND RECOMMENDATIONS

4.1. Conclusions

The conclusions that can be drawn from the molecular dynamics simulation results presented in Sections 2 and 3 will be summarized separately in the following.

4.1.1. Bead-Spring, Freely-Rotating and Polyethylene Model Chains

1. The MD study of bead-spring model chains indicates that the solvent quality affects the properties of polymers in dilute solution at two distinct levels. First, chain properties are affected on a local scale, i.e. at the level of individual chain units and segments. Second, an impact on the chain overall statistics and dynamics is observable.

2. A linear decrease in translational diffusion coefficients is observed with the enhancement of more favorable polymer-solvent interactions, for all model chains.

3. The local flexibility of a chain may be decreased by either introducing barriers to intrachain bond rotations which impart certain chain stiffness, or by increasing the strength of polymer-solvent interactions. The reduction of local flexibility by either means leads to almost indistinguishable behavior. Specifically, the local equilibrium characteristics of a FRC subject to strong interactions with the surroundings closely resemble that of a PE chain in the presence of less favorable interactions.

4. Orientational autocorrelation functions for bond vectors and their stretched exponential forms indicate that the intramolecular constraints are predominant in determining the rates of local relaxational processes. This is apparent from the

quite different time scales of relaxations of FRC and PE. Thus, the solvent effect seems to cause relatively small perturbations on highly localized dynamic phenomena, in contrast to its significant effect on the equilibrium properties. The local dynamics of bead-spring model chains seem to be more efficiently affected by solvent quality, compared to FRC and PE.

5. The stretched exponential functions fitting the time decay of orientational autocorrelation functions yield almost constant exponents β for a given type of chain in various environment. β assumes the respective values of 0.61 ± 0.01 and 0.53 ± 0.02 for bond vectors in FRC and PE. β is rather affected by the intrinsic conformational features of the chain.

6. In contrast to the widely accepted view that vectorial quantities perpendicular to the chain backbone ($\mathbf{m}_{\perp}$) undergo faster relaxation compared to those along the axis ($\mathbf{m}_{\parallel}$), it is shown that under suitable choice of polymer-solvent interaction energy parameters leading to highly contracted chain conformations this behavior may be inverted.

7. A power law relation of the form $\tau \sim n^a$ is found to exist for short segments of n bonds ($n \leq 8$). The exponent *a* increases with solvent quality and remains always lower than 1.5, in contrast to Rouse regime dynamics where $a = 2$. For bead-spring model chains, $1.00 < a < 1.46$, when the polymer-solvent interaction energy is varied in the range $0.1 < \varepsilon_{bs} < 0.8$ kcal / mol. It remains almost fixed at $a = 0.94 \pm 0.04$ in the case of FRC model and varies in the interval $0.66 < a < 1.13$ in the case of polyethylene chains in various solvent environment of $0.1 < \varepsilon_{bs} < 0.8$ kcal / mol. These results are in accordance with sub-Rouse regime dynamics of short polymer chains.

4.1.2. Polypeptide Model Chains

Comparison of the results of *vacuum* and water trajectories indicates that water has a destabilizing effect on helices, especially of alanine and valine residues. Glycine and serine, on the other hand, have some intrinsic helix-breaking propensities, which can be observed in the absence of water. Simulations

averaged over the trajectories indicate the following scales of helix stabilities: pAla-W ≈ pVal-W ≈ pSer-W > pGly-W and pAla-V ≈ pVal-V > pSer-V > pGly-V in water and in *vacuum*, respectively. These tentative scales agree with the experimentally determined fact that glycine is a helix-breaker. Table 3.1 clearly shows that alanine is a good-helix former. Valine and serine are generally known to occur less frequently in α-helices compared to alanine. However, no clear difference between the helix propensities of alanine, valine and serine can be observed in water simulations. Two reasons may be given for this observation. First, the differences in the experimentally observed propensities of amino acids in Table 3.1 are quite small to be detected accurately by MD simulations. A second reason might be that the model parameters of the simulation program do not satisfactorily represent the peptides in solution. For instance, valine whose bulky side groups are expected to be highly strained in a helical structure seems to sustain this compact environment without any unfavorable interactions in simulations.

Equilibrium distribution of water molecules around the side chains of the locally helical alanine, valine and serine residues indicate that the beta carbon of alanine and serine and the gamma carbon of valine have similar solvation patterns. The locations of the first hydration shells around alanine C^{β} and valine C^{γ} are quite different from those observed in crystal structures. This might indicate that the collapsed CH_3 groups of alanine and valine side chains are not adequate to realistically model these amino acids. Specifically, the helix-breaking characteristics of valine due to steric overlap might not be realized with such collapsed CH_3 groups. Furthermore, the torsional mobility of the C^{α}-C^{β} bonds is rather reduced in helices and from the entropic point of view valine should dislike a tight environment hindering C^{α}-C^{β} rotations In fact, exploratory calculations for valine indicate some lowering in helix stability provided that the intrinsic torsional barriers are not as high as those used in the simulation program.

Water distributions around intact hydrogen bonds of the backbone indicate that the hydrogen bonds are well protected by valine side chains, compared to alanine and serine. The fact that hydrogen bonds in alanine are quite open to attact by water molecules gives an explanation for the helix propensity of alanine, which is found to be lower than expected. It is clear that water molecules compete to form hydrogen bonds with the backbone and , as a result, the helical structure gets disrupted.

4.2. Recommendations

As computer resources permit in the future, it would be complementary to study system size effects in more detail by using larger simulation boxes for all the systems considered in this thesis. The problem of hydrodynamic interactions due to periodic boundary artifacts could be alternatively overcome by the method of Ewald summation. It would be interesting to analyze the solvent effect on real polymers by considering solvents of varying quality, such as polar and nonpolar solvents.

Other polar and ionic amino acid residues might be analyzed for helix forming / breaking tendencies. Specifically, the stability of some experimentally designed peptides, some of which are mentioned in Table 3.1, and other short protein segments can be investigated by MD. For example, the experimentally designed peptides with Asp-Arg, Glu-Arg and Glu-Lys [53, 56] residues in different orientations and spacings and specific peptide fragments of lysozyme[85] may be simulated. β-sheet and other secondary structures, which are more difficult to study experimentally, remain other alternatives for future MD simulations.

APPENDIX A

Algorithm of the molecular dynamics simulation program used in Section 2 is summarized as follows:

1. Input of data
2. Initial data manipulation
3. Calculation of initial positions of all atoms (Subroutine BCC)
4. Assignment of initial velocities (Subroutine CONVEL)
5. Generation of polymer chain
6. Calculation of initial forces using classical Verlet algorithm to evaluate **a**(t).
7. Beginning of main loop
8. First part of the Modified Verlet Algorithm (Subroutine MOVEAP)
9. Storing forces in previous step as old forces
10. Calculation of forces to evaluate **a**(t+dt) (Subroutine FORCEP)
11. Second part of modified Verlet algorithm (Subroutine MOVEBP)
12. Calculation of relevant quantities, such as kinetic energy, center of mass, overall chain dimensions, e. t. c., at regular time intervals.
13. Printing of atomic coordinates at regular time intervals
14. End of main loop
15. Calculation of averages, standard deviations
16. Calculation of velocity autocorrelation functions, diffusion coefficients, pair distribution functions e.t.c.
17. Printing of the results

APPENDIX B

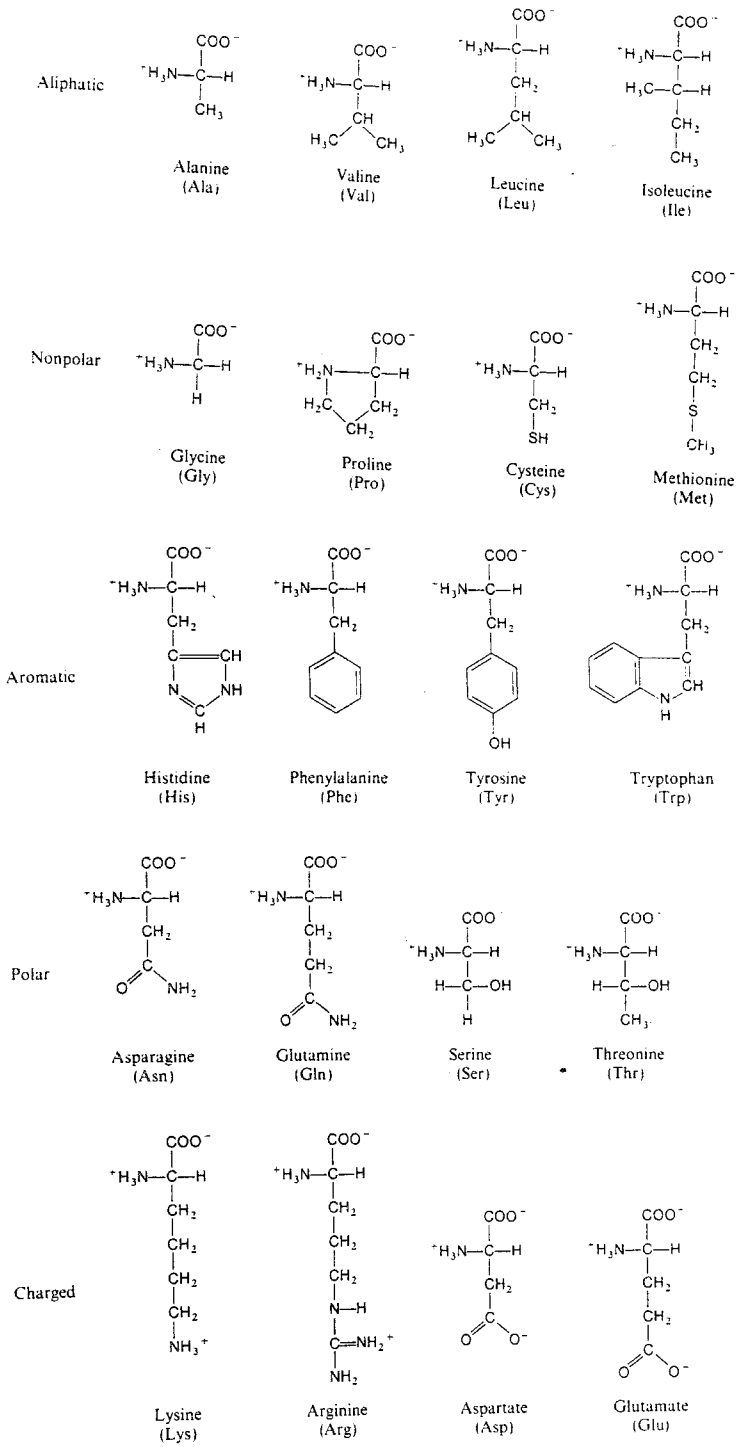


FIGURE B.1. Structures of the twenty amino acids at pH 7 [45]

APPENDIX C

GROMOS87 (W.F. van Gunsteren and H.J.C. Berendsen, GROMOS: Groningen Molecular Simulation Program, University of Groningen, 1987) is a computer simulation package, which has been developed for the dynamic modelling of biomolecules. GROMOS can perform the following operations:

1. Simulation of proteins or arbitrary molecules using the molecular dynamics (MD) or stochastic dynamics (SD) method.
2. Energy minimization (EM) of these molecules.
3. Analysis of molecular conformations obtained by experiment (X-ray, 2D-NMR), by model building or by computer simulation.

Different applications of GROMOS in chemistry and physics are:

1. Prediction of the dependence of a molecular conformation on the type of environment (water, apolar solvent, crystal, etc.)
2. Calculation of relative binding constants by evaluating free energy differences between various molecular complexes
3. Prediction of energetic and structural changes caused by modification of amino acids in enzymes or base pairs in DNA
4. Derivation of three-dimensional (3D) molecular structure on the basis of 2D-NMR data by using restrained MD techniques
5. Dynamic modelling of molecular complexes by searching configuration space by MD
6. Prediction of properties of materials under extreme conditions of temperature and pressure, which may be experimentally inaccessible

GROMOS consists of about 45000 lines of standard FORTRAN code and is a batch oriented packege. It contains about 100 independent building blocks (programs and subroutines), which can be combined in a great variety ways to perform different tasks. The building block philosphy of GROMOS provides great flexibility to the user.

GROMOS contains six types of programs:

1. Programs that build a molecular topology.

2. Programs that transform a given atom coordinate sequence to the atom sequence in the molecular topology, and to the GROMOS coordinate format
3. Programs that generate atom coordinates
4. Programs that perform simulations (EM, MD and SD)
5. Programs that analyze configurations or sequences of configurations.
6. Programs that merge or reduce coordinate files or transform atom coordinates to a special format for interfacing.

The total energy, U, in GROMOS is expressed in the form:

$$U = \sum_{n=1}^{N_b} V_b(b_n) + \sum_{n=1}^{N_\theta} V_\theta(\theta_n) + \sum_{n=1}^{N_\xi} V_\xi(\xi_n) + \sum_{n=1}^{N_\phi} V_\phi(\phi_n) + \sum_{i<j}^{N_{at}} [\, V_{LJ}(r_{ij}) + V_{ES}(r_{ij}) \,]\, S(r_{ij})$$

$$+ \text{special terms} \qquad \text{(C.1)}$$

where the equation for each of the summation terms is presented below:

Bond stretching potential:

$$V_b(b_n) = \frac{1}{2} K_{bn} [\, b_n - b_{on} \,]^2 \qquad \text{(C.2)}$$

Bond angle potential:

$$V_\theta(\theta_n) = \frac{1}{2} K_{\theta n} [\, \theta_n - \theta_{on} \,]^2 \qquad \text{(C.3)}$$

Improper angle torsion potential:

$$V_{\xi}(\xi_n) = \frac{1}{2} K_{\xi n} [\, \xi_n - \xi_{on} \,]^2 \qquad \text{(C.4)}$$

Dihedral angle torsion potential:

$$V_{\phi}(\phi_n) = \frac{1}{2} K_{\phi_n} [\, 1 + \cos (k_n \phi_n - \delta_n) \,] \qquad \text{(C.5)}$$

Lennard-Jones potential:

$$V_{LJ}(r_{ij}) = \frac{C_{12}(i,j)}{r_{ij}^{12}} - \frac{C_6(i,j)}{r_{ij}^{6}} \qquad \text{(C.6)}$$

Electrostatic potential:

$$V_{ES}(r_{ij}) = \frac{q_i \, q_j}{4\pi\varepsilon_o\varepsilon_r r_{ij}} \qquad \text{(C.7)}$$

All the parameters in the above equations are described in the List of Symbols Section. First and second neighbors are excluded from the summation of nonbonded interactions, i.e. van der Waals (Lennard-Jones) and electrostatic interactions.

Two different potential energy parameter sets are used in GROMOS. The C-versions are the basic force fields designed for molecules in solution or in crystalline form. The D-versions are derived from the C-versions for simulating molecules in *vacuum*. In these force fields, all parameters in Equation (C.1) are specified for 37 different types of atoms that exist in amino acids, DNA, sugars, phospholipids and other molecules.

The non-bonded interaction energies and forces can be efficiently calculated in GROMOS by the method of charge groups. A charge group is a bonded group of atoms whose partial atomic charges add up to zero. The leading term of the electrostatic interactions between two charge groups of atoms is of dipolar ($1/r^3$) character, because the 1/r monopole contributions of the various atom pairs to the group-group electrostatic interaction will be zero. The range of electrostatic interactions is considerably reduced when atoms are assembled into charge groups. Consequently, the errors due to the application of a cut-off distance for nonbonded interactions are also reduced. The switching function, $S(r_{ij})$ is added to the nonbonded interactions to obtain a smooth potential energy function when a cut-off radius is used.

Interaction potential may contain special terms for position restraining and/or distance restraining. When simulating a molecular system, it may be desirable to fix specific atoms or parts of the system to given reference positions. In GROMOS, the motions of the atoms are restrained around specific positions by applying a harmonic force. The distance between specific atoms may similarly be restrained to a specified value. It is also possible to restrain bond angles and dihedral angles to specified values.

The method of constrained dynamics can be applied in GROMOS. In this method [80], also called the SHAKE algorithm, the degrees of freedom with the highest frequencies in the system are frozen at equilibrium values, which allows the use of larger time step sizes, Δt. It turns out that the application of bond-length constraints saves abour a factor of 2-3 in computer time. All solvents in GROMOS are modeled with bond length and bond angle constraints.

Energy minimization, which is used to reach low energy configurations in a system, is accomplished by two different methods, namely steepest descents and conjugate gradients minimization methods.

In MD simulations, an external heat and/or pressure bath are utilized to keep the temperature and/pressure of the system at specified values. Details on this external bath are given in Section 3.2 and by Equations (3.2) and (3.3).

Two different boundary conditions are used in the simulations of finite size systems. The vacuum boundary condition, which corresponds to the gas phase at zero pressure, gives best results for relatively large globular macromolecules. In the simulation of liquids or solutions, periodic boundary conditions are used to minimize edge or wall effects.Different types of periodic boxes may be used in GROMOS, namely rectangular, monoclinic or truncated octahedron.

In GROMOS, the information concerning a molecular system is distributed over two distinct data files: a molecular topology (MT) file and an atomic coordinate (AC) file. An MT file contains information about the topology of a molecular system, i.e. data on the covalent structure, atomic masses, charges, van der Waals parameters, etc. An AC file specifies the configurations of a molecular system, such as Cartesian coordinates, velocities and sizes of the periodic box. These two types of information are stored separately, because configurations change continuously, whereas the molecular topology generally remains unchanged during simulation. Only in the application of thermodynamic integration formalism, used for the determination of free energy differences between two states, the molecular topology of a system is perturbed by a perturbation potential. Most programs of GROMOS use a binary molecular topology file containing the topological and force field data of the system. There are several different ways of forming MT files. Specifically, the MT files for large biomolecules can be easily generated by the use of the existing molecular topology building blocks in GROMOS, which are molecules or parts of molecules like amino acid residues, nucleotides, etc.

The GROMOS programs are run by the use of shell command files, which specify the names and locations of input and output files and the programs and subroutines to be used to perform a specific task. An example of a shell command file to carry out a molecular dynamics run and its sample input file defining the simulation parameters are given in the following.

SHELL COMMAND FILE

```
#!/bin/sh
# This file created from the vms original on
# Fri Jun 17 16:39:24 MET DST 1994

# Start by removing old fortran links
rm -f fort.?; rm -f fort.??

# Identify the gromos directory for input files
eg=../examples
dt=../data

input_f=$eg/imdalai1.dat                  # input polyalanine (+ions-startup)
ln -s     mtalaic4.bin     fort.20        # binary molecular topology (pAla+ions)
ln -s     alaixem1.dat     fort.21        # initial coordinates (pAla+ions)
# ln -s   xxxx             fort.23        # old option, not used
# ln -s   xxxx             fort.24        # restraining reference positions
# ln -s   xxxx             fort.25        # sequence numbers of restrained atoms
# ln -s   xxxx             fort.26        # distance restraint atom pairs
# ln -s   xxxx             fort.27        # perturbation of molecular topology
# ln -s   xxxx             fort.28        # restrained dihedral angles
ln -s     alaixmd1.dat     fort.31        # final coordinates (pAla+ions+water)
ln -s     alairmd1.dat     fort.12        # md trajectory coordinates (idem)
# ln -s   xxxx             fort.13        # md trajectory velocities
# ln -s   xxxx             fort.15        # md trajectory energies, etc.
output_f=outmdalai1.lis
../promd.64 <  $input_f  >  $eg/out/$output_f
rm -f  fort.20 fort.21 fort.31 fort.12
```

INPUT FILE 'imdalai1.dat'

```
 13residue poly(ALA)+NA+CL+540H2O,MD(STARTUP),ALAXEM1,ALAXMD1,pemra,jan1995
         1       540         1         0    280686 350.0         .0
         2   2.13913   2.27856   3.61099  90.0
         1         2 350.0      10.0        .01       .01
         2         3    .06102  .0007476    .05
        81        82        83
         3         1    100000
       500         1         3   0.0         .002
         3         0    .0001
         3         2         2         0         81
         1        10    .8      10.0      10.0         .2         .8
        10         1
        25    100000        25         0         1         0         2
         0        19000.0
         0    1000.0  10.0        .1        .153
         0         0         2    .0        .0
         0   1.0
```

REFERENCES

1. Bird, R. B., C. F. Curtiss, R. C. Armstrong and O. Hassager, *Dynamics of Polymeric Liquids. Volume 2 Kinetic Theory*, Wiley-Interscience, New York, 1987.

2. Flory, P. J., *Statistical Mechanics of Chain Molecules*, Interscience, New York, 1969.

3. Rose, G. D. and R. Wolfenden, "Hydrogen Bonding, Hydrophobicity, Packing, and Protein Folding," *Annual Review of Biophysics and Biomolecular Structure*, Vol. 22, pp. 381-415, 1993.

4. Yamakawa, H., *Modern Theory of Polymer Solutions*, Harper & Row, New York, 1971.

5. Bahar, I., B. M. Baysal and B. Erman, "Local Solvent Effects on Configurational Characteristics of Polymer Chains: Poly(*p*-Chlorostyrene) in Benzene," *Macromolecules*, Vol. 19, pp. 1703-1709, 1986.

6. Allen, M. P. and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon, Oxford, 1989.

7. Roe, R-J., (Ed.), *Computer Simulation of Polymers*, Prentice Hall, New York, 1990.

8. Bishop, M., M. H. Kalos and H. L. Frisch, "Molecular Dynamics of Polymeric Systems," *Journal of Chemical Physics*, Vol. 70, pp. 1299-1304, 1979.

9. Rapaport, D. C., "Molecular Dynamics Study of a Polymer Chain in Solution," *Journal of Chemical Physics*, Vol. 71, pp. 3299-3303, 1979.

10. Bruns, W. and R. Bansal, "Molecular Dynamics Study of a Single Polymer Chain in Solution. II. Bead-Spring Model," *Journal of Chemical Physics*, Vol. 75, pp. 5149-5152, 1981.

11. Bruns, W. and R. Bansal, "Molecular Dynamics Study of a Single Polymer Chain in Solution," *Journal of Chemical Physics*, Vol. 74 (3), pp. 2064-2072, 1981.

12. Bishop, M., M. H. Kalos and H. L. Frisch, "The Influence of Attractions on the Static and Dynamic Properties of Simulated Single and Multichain Systems," *Journal of Chemical Physics*, Vol. 79, pp. 3500-3504, 1983.

13. Smit, B., A. van der Put, C. J. Peters, J. de Swaan Arons and J. P. J. Michels, "The Influence of the Density of the Solvent on the Static and Dynamic Properties of Star Polymers. A Molecular Dynamics Study," *Journal of Chemical Physics*, Vol. 88, pp. 3372-3375, 1988.

14. Luque, J., J. Santamaria and J. J. Freire, "Molecular Dynamics of Chain Molecules in Solution. Static and Dynamic Properties," *Journal of Chemical Physics*, Vol. 92 (1), pp. 584-589, 1989.

15. Dunweg, B. and K. Kremer, "Microscopic Verification of Dynamic Scaling in Dilute Polymer Solutions: A Molecular Dynamics Simulation," *Physical Review Letters*, Vol. 66, pp. 2996-2999, 1991.

16. Pierleoni, C. and J-P.Ryckaert, "Relaxation of a Single Chain Molecule in Good Solvent Conditions by Molecular-Dynamics Simulation," *Physical Review Letters*, Vol. 66, pp. 2992-2995, 1991 .

17. Pierleoni, C. and J-P.Ryckaert, "Molecular Dynamics Investigation of Dynamic Scaling for Dilute Polymer Solutions in Good Solvent Conditions," *Journal of Chemical Physics*, Vol. 96(11), pp. 8539-8551, 1992.

18. Takeuchi, H. and R-J. Roe, "Molecular Dynamics Simulation of Local Chain Motion in Bulk Amorphous Polymers. I. Dynamics Above the Glass Transition," *Journal of Chemical Physics*, Vol. 94, pp. 7446-7457, 1991.

19. Takeuchi, H. and R-J. Roe, "Molecular Dynamics Simulation of Local Chain Motion in Bulk Amorphous Polymers. II. Dynamics at Glass Transition," *Journal of Chemical Physics*, Vol. 94, pp. 7458-7465, 1991.

20. Rigby, D. and R-J. Roe, "Molecular Dynamics Simulation of Polymer Liquid and Glass. I. Glass Transition," *Journal of Chemical Physics*, Vol. 87, pp. 7285-7292, 1987.

21. Takeuchi, H. and K. Okazaki, "Molecular Dynamics Simulation of Diffusion of Simple Gas Molecules in a Short Chain Polymer," *Journal of Chemical Physics*, Vol. 92, pp. 5643-5652, 1990.

22. Takeuchi, H., R-J. Roe and J. E. Mark, "Molecular Dynamics Simulation of Diffusion of Small Molecules in Polymers. II. Effect of Free Volume Distribution," *Journal of Chemical Physics*, Vol. 93, pp. 9042-9048, 1990.

23. de Gennes, P-G., *Scaling Concepts in Polymer Physics*, Ithaca, Cornell University, 1979.

24. Ryckaert, J-P. and A. Bellemans, "Molecular Dynamics of Liquid *n*-Butane Near Its Boiling Point," *Chemical Physics Letters*, Vol. 30, pp. 123-125, 1975.

25. Rigby, D. and R-J. Roe, "Molecular Dynamics Simulation of Polymer Liquid and Glass. II. Short Range Order and Orientation Correlation," *Journal of Chemical Physics*, Vol. 89, pp. 5280-5290, 1988.

26. Bahar, I., B. Badur and P. Doruker, "Solvent Effect on Translational Diffusivity and Orientational Mobility of Polymers in Solution: A Molecular Dynamics Study," *Journal of Chemical Physics*, Vol. 99, pp. 2235-2246, 1993.

27. Doruker, P. and I. Bahar, "Effect of Intrachain Constraints and Polymer-Solvent Interactions on Chain Statistics and Dynamics," *Computational Polymer Science*, Vol. 3, pp. 87-99, 1993.

28. Helfand, E., Z. R. Wasserman and T. A. Weber, "Brownian Dynamics Study of Polymer Conformational Transitions," *Macromolecules*, Vol. 13, pp. 526-533, 1980.

29. Doi, M. and S. F. Edwards, *The Theory of Polymer Dynamics*, Clarendon, Oxford, 1986.

30. Zwanzig, R. and N. K. Ailawadi, 'Statistical Error Due to Finite Time Averaging in Computer Experiments," *Physical Review Letters*, Vol. 182, pp. 280-283, 1969.

31. Bahar, I., B. Erman and L. Monnerie, "Local Orientational Motions in Flexible Polymer Chains," *Macromolecules*, Vol. 23, pp. 1174-1180, 1990.

32. Allegra, G., "Elasticity of a Real Polymer Chain for Different Modes of Motion, " *Journal of Chemical Physics* ,Vol. 68: 3600-3606, 1978.

33. Fixman, M., "Simulation of Polymer Dynamics. II. Relaxation Rates and Dynamic Viscosity," *Journal of Chemical Physics*, Vol. 69, pp. 1527-1538, 1978.

34. Bahar, I., B. Erman and L. Monnerie, "Stochastic Treatment of Conformational Transitions of Polymer Chains in the Sub-Rouse Regime," *Macromolecules*, Vol. 24, pp. 3618-3626, 1991.

35. Akcasu, A. Z., M. Benmouna and C. C. Han, "Interpretation of Dynamic Light Scattering from Polymer Solutions," *Polymer*, Vol. 21, pp. 866-890, 1980.

36. Akcasu, A. Z., personal communication.

37. Nicholson, L. K., J. S. Higgins, J. B. Hayter, "Dynamics of Dilute Polymer Solutions," *Macromolecules*, Vol. 14, pp. 836-843, 1981.

38. Akcasu, A. Z. and J. S. Higgins, "Quasielastic Scattering of Neutrons from Freely Jointed Polymer Chains in Dilute Solution," *Journal of Polymer Science, Polymer Physics Edition*, Vol. 15, pp. 1745-1756, 1977.

39. Haliloglu, T., I. Bahar and B. Erman, "Orientational and Conformational Correlations in Deformed Polymer Chains with Fixed End-to-End Separation: A Brownian Dynamics Simulation Study," *Journal of Chemical Physics*, Vol. 97, pp. 4428-4437, 1992.

40. Bahar, I., B. Erman and L. Monnerie, "Effect of Molecular Structure on Local Chain Dynamics: Analytical Approaches and Computational Methods," *Advances in Polymer Science*, Vol. 116, pp. 145-206, 1994.

41. Bahar, I. and B. Erman, "Anisotropy of Static and Dynamic Orientational Correlations in *n*-Alkanes," *Journal of Chemical Physics*, Vol. 88, pp. 1228-1233, 1988.

42. Stryer, L., *Biochemistry*, W.H. Freeman and Company, New York, 1988.

43. Chan, H. S. and K. A. Dill, "The Protein Folding Problem," *Physics Today*, pp. 24-32, 1993.

44. Bohinski, R. C., *Modern Concepts in Biochemistry*, Allyn and Bacon, Inc., Boston, 1983.

45. Cantor, C. R. and P. R. Schimmel, *Biophysical Chemistry. Part I: The Conformation of Biological Macromolecules*, W.H. Freeman and Company, New York, 1980.

46. Shoemaker, K. R., P. S. Kim, D. N. Brems, S. Marqusee, E. J. York, I. M. Chaiken, J. M. Stewart and R. I. Baldwin, "Nature of the Charged-Group Effect on the Stability of the C-peptide Helix," *Proceedings of the National Academy of Sciences*, Vol. 82, pp. 2349-2353, 1985.

47. Fasman, G. D., (Ed.), *Prediction of Protein Structure and the Principles of Protein Conformation*, Plenum Press, New York, 1989.

48. Thornton, J. M., "Protein Structures: The End of the Folding Pathway," in T. E. Creighton (Ed.) , *Protein Folding*, pp. 59-81, New York, W.H. Freeman and Company. 1992.

49. Lattman, E. E. and G. D. Rose, "Protein Folding- What's the Question?," *Proceedings of the National Academy of Sciences, USA*, Vol. 90, pp. 439-441, 1993.

50. Marqusee, S., V. H. Robbins, and R. L. Baldwin, "Unusually Stable Helix Formation in Short Alanine-Based Peptides," *Proceedings of the National Academy of Sciences*, Vol. 86, pp. 5286-5290, 1989.

51. Bierzynski, A., P. S. Kim and R. L. Baldwin, "A Salt Bridge Stabilizes the Helix Formed by Isolated C-peptide of RNase A," *Proceedings of the National Academy of Sciences*, Vol. 79, pp. 2470-2474, 1982.

52. Shoemaker, K. R,. P. S. Kim, E. L. York, J. M. Stewart and R. L. Baldwin, "Tests of the Helix Dipole Model for Stabilization of α-Helices," *Nature*, Vol. 326, pp. 563-567, 1987.

53. Marqusee, S. and R.L. Baldwin,. "Helix Stabilization by $Glu^{-}...Lys^{+}$ Salt Bridges in Short Peptides of *de novo* Design," *Proceedings of the National Academy of Sciences*, Vol. 84, pp. 8898-8902, 1987.

54. Lyu, P. C., L. A. Marky and N. R. Kallenbach, "The Role of Ion Pairs in α-Helix Stability: Two New Designed Helical Peptides," *Journal of American Chemical Society*, Vol. 111, pp. 2733-2734, 1989.

55. Bradley, E.K., J. F. Thomason, F. E. Cohen, P. A. Kosen and I. D. Kuntz, "Studies of Synthetic Helical Peptides Using Circular Dichroism and Nuclear Magnetic Resonance," *Journal of Molecular Biology*, Vol. 215, pp. 607-622, 1990.

56. Huyghues-Despointes, B. M. P., J. M. Scholtz and R. L. Baldwin, "Helical Peptides with Three Pairs of Asp-Arg and Glu-Arg Residues in Different Orientations and Spacings," *Protein Science*, Vol. 2, pp. 80-85, 1993.

57. Chou, P. Y., and G. D. Fasman, "Empirical Predictions of Protein Conformation," *Annual Review of Biochemistry*, Vol. 47, pp. 251-276, 1978.

58. Wojcik, J., K. H. Altmann and H. A. Scheraga, "Helix-Coil Stability for the Naturally Occuring Amino Acids in Water," *Biopolymers*, Vol. 30, pp. 121-134, 1990.

59. Lyu, P. C., M. I. Liff, L. A. Marky and N. R. Kallenbach, "Side Chain Contributions to the Stability of Alpha-Helical Structure in Peptides," *Science*, Vol. 250, pp. 669-673, 1990.

60. O'neil, K. T. and W. F. DeGrado, "A Thermodynamic Scale for the Helix-Forming Tendencies of the Commonly Occuring Amino Acids," *Science*, Vol. 250, pp. 646-651, 1990.

61. Padmanabhan, S., S. Marqusee, T. Ridgeway, T. M. Laue and R. L. Baldwin, "Relative Helix-Forming Tendencies of Nonpolar Amino Acids," *Nature*, Vol. 344, pp. 268-270, 1990.

62. Horovitz, A., J. M. Matthews and A. R. Fersht, "α-Helix Stability in Proteins. II. Factors that Influence Stability at an Internal Position," *Journal od Molecular Biology*, Vol. 227, pp. 560-568, 1992.

63. Blaber, M., X. Zhang and B. W. Matthews, "Structural Basis of Amino Acid α Helix Porpensity," *Science*, Vol. 260, pp. 1637-1640, 1993.

64. Kim, C. A. and J. M. Berg, "Thermodynamic β-Sheet Propensities Measured Using a Zinc-Finger Host Peptide," *Nature*, Vol. 362, pp. 267-270, 1993.

65. Bai, Y. and S. W. Englander, "Hydrogen Bond Strength and β-Sheet Propensities: The Role of a Side Chain Blocking Effect," *Proteins: Structure, Function and Genetics*, Vol. 18, pp. 262-266, 1994.

66. Creamer, T. P. and G. D. Rose, "Side-Chain Entropy Opposes α-Helix Formation but Rationalizes Experimentally Determined Helix-Forming Porpensities," *Proceedings of the National Academy of Sciences*, Vol. 89, pp. 5937-5941, 1992.

67. Hermans, J., A. G. Anderson and R. H. Yun, "Differential Helix Propensity of Small Apolal Side Chains Studied by Molecular Dynamics Simulations," *Biochemistry*, Vol. 31, pp. 5646-5653, 1992.

68. Avbelj, F. and J. Moult, "The Role of Electrostatic Screening in Determining Protein Mail Chain Conformational Preferences," *Biochemistry*, (in press).

69. Daggett, V., P. A. Kollman and I. D. Kuntz, "Molecular Dynamics Simulations of Small Peptides: Dependence on Dielectric Model and pH," *Biopolymers*, Vol. 31, pp. 285-304, 1991.

70. McCammon, J. A. and S. H. Northrup, "Helix-Coil Transitions in a Simple Polypeptide Model," *Biopolymers*, Vol. 19, pp. 2033-2045, 1980.

71. Daggett, V., P. A. Kollman and I. D. Kuntz, "A Molecular Dynamics Simulation of Polyalanine: An Analysis of Equilibrium Motions and Helix-Coil Transitions," *Biopolymers*, Vol. 31, pp. 1115-1134, 1991.

72. Sundaralingam, M. and Y. C. Sekharudu, "Water-Inserted α-Helical Segments Implicate Reverse Turns as Folding Intermediates," *Science* , Vol. 244, pp. 1333-1337, 1989.

73. DiCapua, F. M., S. Swatminathan and D. L. Beveridge, "Theoretical Evidence for Water Insertion in α-Helix Bending: Molecular Dynamics of Gly_{30} and Ala_{30} in Vacuo and in Solution," *Journal of American Chemical Society*, Vol. 113, pp. 6145-6155, 1991.

74. Daggett, V. and M. Levitt, "Molecular Dynamics Simulations of Helix Denaturation," *Journal of Molecular Biology*, Vol. 223, pp. 1121-1138, 1992.

75. van Buuren, A. R and H. J. C. Berendsen, "Molecular Dynamics Simulation of the Stability of a 22-Residue α-Helix in Water and 30% Trifluoroethanol," *Biopolymers* , Vol. 33, pp. 1159-1166, 1993.

76. Tobias, D. J. and C. L. Brooks III, "Thermodynamics and Mechanism of α Helix Initiation in Alanine and Valine Peptides," *Biochemistry*, Vol. 30, pp. 6059-6070, 1991.

77. Brooks III, C. L. and L. Nilsson, "Promotion of Helix Formation in Peptides Dissolved in Alcohol and Water-Alcohol Mixtures," *Journal of American Chemical Society*, Vol. 115, pp. 11034-11035, 1993.

78. Berendsen, H. J. C., J. P. M. Postma, W. F van Gunsteren and J Hermans, "Interaction Models in Relation to Protein Hydration," in B. Pullman (Ed.) *Intermolecular Forces* , pp. 331-342, D. Reidel Publishing Company, Holland, 1981.

79. Berendsen, H. J. C., J. P. M. Postma, W. F van Gunsteren, W. F. D. Nola and J. R. Haak, "Molecular Dynamics with Coupling to an External Bath," *Journal of Chemical Physics*, Vol. 81, pp. 3684-3690, 1984.

80. van Gunsteren, W. F., and H. J. C. Berendsen, "Algorithms for Macromolecular Dynamics and Constraint Dynamics," *Molecular Physics*, Vol. 34, pp. 1311-1327, 1977.

81. Walshaw, J. and J. M. Goodfellow, "Distribution of Solvent Molecules Around Apolar Side-chains in Protein Crystals," *Journal of Molecular Biology*, Vol. 231, pp. 392-414, 1993.

82. Thanki, N., J. M. Thornton and J. M. Goodfellow, "Distributions of Water Around Amino Acid Residues in Proteins," *Journal of Molecular Biology*, Vol. 202, pp. 637-657, 1988.

83. Thanki, N., Y.Umrania., J. M. Thornton and J. M. Goodfellow, "Analysis of Main-chain Solvation as a Function of Secondary Structure," *Journal of Molecular Biology*, Vol. 221, pp. 669-691, 1991.

84. Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R.Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi, "The Protein Data Bank: a Computer Based Archival File for Macromolecular Structures," *Journal of Molecular Biology*, Vol. 112, pp. 535-542, 1977 .

85. Segawa, S., T. Fukuno, K. Fujiwara and Y. Noda, "Local Structures in Unfolded Lysozyme and Correlation with Secondary Structure Propensities in the Native Conformation: Helix- Forming or -Breaking Propensity of Peptide Fragments," *Biopolymers* , Vol. 31, pp. 497-509, 1991.