

THEORETICAL STUDY OF EQUILIBRIUM STRUCTURES
AND KINETIC PHENOMENA IN MACROMOLECULAR
SOLUTIONS

A THESIS SUBMITTED TO
THE NATIONAL UNIVERSITY OF IRELAND
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

by
Yuri A. Kuznetsov



Based on research carried out in
the Department of Chemistry, University College Dublin

under the direction of **Professor Kenneth A. Dawson**
and the supervision of **Professor Kenneth A. Dawson**
and **Dr Edward G. Timoshenko**

Contents

| | |
|--|-------------|
| Abstract | xii |
| Acknowledgements | xiii |
| 1 Introduction | 1 |
| 2 The Gaussian Self-Consistent Method | 25 |
| 2.1 Model of Copolymer Solutions at Finite Concentrations | 25 |
| 2.2 The Gibbs-Bogoliubov Variational Principle | 30 |
| 2.3 General Form of the Gaussian Self-Consistent Equations | 35 |
| 2.4 The Self-Interaction Energy Term | 40 |
| 2.5 Techniques for Numerical Solution of the GSC Equations | 41 |
| 2.6 GSC Equations for Homopolymer Solutions | 45 |
| 3 Simulation Techniques | 49 |
| 3.1 Lattice Monte Carlo Model of Copolymer Solutions | 49 |
| 3.2 Monte Carlo Model of the Ring Homopolymer in Continuous Space | 69 |
| 4 Infinitely Dilute Homopolymer Solutions | 76 |
| 4.1 The Good Solvent Regime | 76 |
| 4.2 Collapse Transition of the Homopolymer | 83 |
| 4.3 Kinetic Laws at the Collapse Transition of the Homopolymer . . . | 98 |

| | | |
|----------|---|------------|
| 4.3.1 | Linearised GSC equations for early times | 100 |
| 4.3.2 | Linearised GSC equations for late times | 102 |
| 4.3.3 | Kinetics of folding | 104 |
| 4.3.4 | Specific features in kinetics of folding of an open chain . . . | 117 |
| 4.4 | Structure of the Homopolymer Globule | 122 |
| 4.5 | Conformations of Stiff Macromolecules | 131 |
| 4.6 | Orientationally Ordered Phases of the Homopolymer | 141 |
| 4.6.1 | Equilibrium phases | 143 |
| 4.6.2 | Kinetics of folding | 148 |
| 5 | Infinitely Dilute Heteropolymer Solutions | 152 |
| 5.1 | Phase Diagrams of Amphiphilic Heteropolymers | 153 |
| 5.2 | Kinetics of Folding | 168 |
| 6 | Polymer Solutions at Finite Concentrations | 180 |
| 6.1 | Homopolymer Solutions | 182 |
| 6.1.1 | Thermodynamic limit | 182 |
| 6.1.2 | Equilibrium phase diagram | 185 |
| 6.1.3 | Metastable states | 188 |
| 6.1.4 | Phase separation: a Monte Carlo study | 190 |
| 6.2 | Phase Separation in Copolymer Solutions | 199 |
| 7 | Conclusion | 210 |
| | Bibliography | 215 |
| | Appendix A | 230 |
| | Appendix B | 245 |

List of Tables

| | | |
|-----|---|-----|
| 4.1 | Values of the mean squared radius of gyration, $3R_g^2$, vs the degree of polymerisation, N , for different simulation procedures. | 78 |
| 4.2 | The behaviour of various observables across the collapse transition vs the degree of polymerisation, N | 86 |
| 4.3 | The main kinetic exponents of the homopolymer in near equilibrium regimes deduced analytically from the GSC method. | 104 |
| 4.4 | Numerical values of the exponent α_i in (4.32) vs the degree of polymerisation, N | 113 |
| 4.5 | Values of the characteristic collapse times τ_m ($c_3 = 0$), τ_t ($c_3 = 0$) and τ_f vs the degree of polymerisation, N , for different viscosities, η_s | 115 |
| 4.6 | Numerical values of parameters α , τ_S , Z , $R_g^2(\infty)$, A_1 , τ_1 from formulae (4.58, 4.59, 4.60) for polymers of different degrees of polymerisation, N | 150 |
| 6.1 | Values of the specific free energy, $a = \mathcal{A}/MN$, at various minima for the system of 12 open homopolymers in volume $L = 10$ for various values of the second virial coefficient, $u^{(2)}$ | 189 |
| 6.2 | Values of the specific free energy, $a = \mathcal{A}/MN$, at various minima for the system of 12 $(ab)_6$ copolymers for various values of the mean second virial coefficient, $\bar{u}^{(2)}$ | 202 |

| | | |
|-----|---|-----|
| 6.3 | Values of the specific free energy, $a = \mathcal{A}/MN$, at various minima for the system of 12 $(a_2b_2)_3$ copolymers for various values of the mean second virial coefficient, $\bar{u}^{(2)}$ | 202 |
| 6.4 | Values of the specific free energy, $a = \mathcal{A}/MN$, at various minima for the system of 12 $(a_3b_3)_2$ copolymers for various values of the mean second virial coefficient, $\bar{u}^{(2)}$ | 203 |

List of Figures

| | | |
|-----|--|----|
| 3.1 | An example of a self-avoiding walk on a two-dimensional square lattice. | 65 |
| 3.2 | An illustration of the nonphantomness of chains in the lattice Monte Carlo simulation. | 67 |
| 3.3 | The two-body interaction potential used in the model of a homopolymer ring in continuous space. | 70 |
| 4.1 | Plot of the effective monomer size, $a_{e-e} \equiv R_{e-e}/N^{0.588}$, vs the degree of polymerisation, N , for different simulation procedures. | 80 |
| 4.2 | Plot of the number of loops consisting of k segments, L_k , vs their length, k , for different simulation procedures. | 82 |
| 4.3 | Plot of the mean squared radius of gyration, R_g^2 , vs the Flory interaction parameter, χ , in the lattice model of the homopolymer. | 84 |
| 4.4 | Plot of the specific mean energy, $\langle H \rangle/N$, vs the Flory interaction parameter, χ | 84 |
| 4.5 | Plot of the specific heat capacity, c_V/k_B , vs the Flory interaction parameter, χ | 85 |
| 4.6 | Plot of the mean squared radius of gyration, $3R_g^2$, vs the two-body interaction parameter, V_0 , in the homopolymer model in continuous space. | 88 |

| | | |
|------|--|-----|
| 4.7 | Plot of the specific mean energy, $\langle H \rangle / N$, vs the two-body interaction parameter, V_0 | 88 |
| 4.8 | Plot of the specific heat capacity, c_V / k_B , vs the two-body interaction parameter, V_0 | 89 |
| 4.9 | Phase diagram of the ring homopolymer with the degree of polymerisation $N = 100$ in terms of the second, $u^{(2)}$, and third, $u^{(3)}$, virial coefficients. | 91 |
| 4.10 | Plot of the mean squared distances, D_k , vs the chain index, k , in the off-lattice Monte Carlo simulation. | 92 |
| 4.11 | Plot of the two-body correlation function $h_5^{(2)}(r)$ (Fig. a) and $h_{50}^{(2)}(r)$ (Fig. b) vs the separation, r , for the homopolymer with $N = 100$ and $V_0 = 0$ | 94 |
| 4.12 | Plot of the two-body correlation function $h_5^{(2)}(r)$ (Fig. a) and $h_{50}^{(2)}(r)$ (Fig. b) vs the separation, r , for the homopolymer with $N = 100$ and $V_0 = 5$ | 96 |
| 4.13 | Evolution of a particular polymer conformation during kinetics at the collapse transition (a, b). | 105 |
| 4.14 | Evolution of a particular polymer conformation during kinetics at the collapse transition (c, d). | 106 |
| 4.15 | Evolution of a particular polymer conformation during kinetics at the collapse transition (e, f). | 107 |
| 4.16 | Evolution of a particular polymer conformation during kinetics at the collapse transition (g, h). | 108 |
| 4.17 | Evolution of a particular polymer conformation during kinetics at the collapse transition (a, b). | 109 |
| 4.18 | Evolution of a particular polymer conformation during kinetics at the collapse transition (c, d). | 110 |

| | | |
|------|---|-----|
| 4.19 | Evolution of a particular polymer conformation during kinetics at the collapse transition (e, f). | 111 |
| 4.20 | Evolution of a particular polymer conformation during kinetics at the collapse transition (g, h). | 112 |
| 4.21 | Diagrams of matrix of the mean squared distances, D_{ij} , for the open homopolymer with $N = 150$, in kinetics after the quench, $u^{(2)} = 30 \rightarrow -50$ | 121 |
| 4.22 | Plot of the partially summated mean squared distances, $D_{\hat{k}}$, vs the normalised chain index, $\hat{k} = k/N - 1$, for homopolymer globules of different sizes and different simulation procedures. | 124 |
| 4.23 | Plot of the probability, $P_i^{(surf)}$, for the i -th monomer in the chain to appear on the globule surface vs the monomer index, i , for the homopolymer with the degree of polymerisation, $N = 400$, for different simulation procedures. | 126 |
| 4.24 | Diagrams of the ratios of the mean squared distances with and without reptations, $D_{ij}^{(r)}/D_{ij}^{(n)}$ for different values of the degree of polymerisation. | 127 |
| 4.25 | Plots of the number of loops per monomer, $l_{\hat{k}} = L_{\hat{k}}/N$, (Fig. a) and the ratio of the number of loops in simulations with, $l_{\hat{k}}^{(r)}$, $p_r = 1/N$, and without, $l_{\hat{k}}^{(n)}$, $p_r = 0$, reptations (Fig. b) vs the normalised chain index, \hat{k} , for the homopolymer globule. | 129 |
| 4.26 | Plots of the mean squared radius of gyration, R_g^2 , vs the second virial coefficient, $u^{(2)}$, upon a quasistatic change of the latter. . . . | 132 |
| 4.27 | Plot of the mean squared distances between monomers, D_k , vs the chain index, k , for various conformational states. | 134 |
| 4.28 | Plots of the mean squared radius of gyration, R_g^2 , vs the stiffness parameter, λ | 136 |

| | | |
|------|--|-----|
| 4.29 | Phase diagram of the stiff homopolymer in terms of the stiffness parameter, λ , and the second virial coefficient, $u^{(2)}$ | 137 |
| 4.30 | Typical conformations of a stiff homopolymer with the degree of polymerisation $N = 100$ from the off-lattice Monte Carlo simulation. | 139 |
| 4.31 | Phase diagram of the homopolymer with the degree of polymerisation $N = 120$ in the model including three-body interactions. | 143 |
| 4.32 | Conformations of the collapsed polymer at equilibrium for the degree of polymerisation $N = 768$ in the model with three-body interactions. | 144 |
| 4.33 | Plot of $\langle m_r^{(l,m)} \rangle$ vs the two-body interaction parameter, $\chi^{(2)}$ | 146 |
| 4.34 | Plot of $\langle m_r^{(l,m)} \rangle$ vs the three-body interaction parameter, $\chi^{(3)}$ | 147 |
| 4.35 | Plots of the mean squared radius of gyration, $3R_g^2$, vs time, t , for different collapse transitions. | 151 |
| 5.1 | Phase diagrams of “short“ blocks copolymer sequences $'(ab)_{15}'$ (Fig. a) and $'(ab)_{30}'$ (Fig. b) in terms of the mean second virial coefficient, $\bar{u}^{(2)}$, and the amphiphilicity, Δ | 154 |
| 5.2 | Phase diagrams of “long” blocks copolymer sequences $'(a_3b_3)_5'$ (Fig. a) and $'(a_3b_3)_{10}'$ (Fig. b) in terms of the mean second virial coefficient, $\bar{u}^{(2)}$, and the amphiphilicity, Δ | 155 |
| 5.3 | Phase diagrams of copolymers with the “random” sequence $'babca_2cbac_2acb_3cac_3a_2b_2cac_2b'$ (Fig. a) and the sequence twice as in Fig. a $'(babca_2cbac_2acb_3cac_3a_2b_2cac_2b)_2'$ (Fig. b) in terms of the mean second virial coefficient, $\bar{u}^{(2)}$, and the amphiphilicity, Δ | 156 |
| 5.4 | Plots of the mean squared radius of gyration, R_g^2 , (Fig. a) and the parameter of micro-phase separation, Ψ , (Fig. b) vs the amphiphilicity, Δ | 157 |

| | | |
|------|---|-----|
| 5.5 | Typical conformations of $(a_6b_6)_{20}$ hydrophobic–neutral copolymer at equilibrium from the lattice Monte Carlo simulation. | 159 |
| 5.6 | Plots of the mean squared radius of gyration, R_g^2 , (Fig. a) and the parameter of micro–phase separation, Ψ , (Fig. b) vs the mean second virial coefficient, $\bar{u}^{(2)}$, for $'(a_3b_3)_{10}'$ copolymer. | 160 |
| 5.7 | Plots of the mean squared radius of gyration, R_g^2 , (Fig. a) and the parameter of micro–phase separation, Ψ , (Fig. b) vs the mean second virial coefficient, $\bar{u}^{(2)}$, for $'(ab)_{30}'$ copolymer. | 163 |
| 5.8 | Plots of the mean squared radius of gyration, R_g^2 , (Fig. a) and the parameter of micro–phase separation, Ψ , (Fig. b) vs the mean second virial coefficient, $\bar{u}^{(2)}$, for $'(babca_2cbac_2acb_3cac_3a_2b_2cac_2b)_2'$ copolymer. | 164 |
| 5.9 | Diagrams of the mean squared distances matrix, D_{ij} , for $'(ab)_{30}'$ copolymer. | 165 |
| 5.10 | Plot of the mean squared radius of gyration, R_g^2 , vs time, t , during kinetics after the quench from $\bar{u}^{(2)} = 15$, $\Delta = 0$ to $\bar{u}^{(2)} = -35$, $\Delta = 30$ for copolymer sequences with $N = 50$ | 169 |
| 5.11 | Plot of the MPS order parameter, Ψ , vs time, t , during kinetics after the same quench and for the same copolymer sequences as in Fig. 5.10. | 170 |
| 5.12 | Plot of the instantaneous free energy, \mathcal{A} , vs time, t , during kinetics after the same quench and for the same copolymer sequences as in Fig. 5.10. | 171 |
| 5.13 | Diagrams of the mean squared distances matrix, $D_{ij}(t)$, for $'(ba)_{25}'$ copolymer in kinetics after the same quench as in Figs. 5.10, 5.11 and 5.12. | 173 |

| | | |
|------|--|-----|
| 5.14 | Typical conformations of $(a_6b_6)_{20}$ hydrophobic–neutral copolymer during folding kinetics. | 175 |
| 5.15 | Diagram of the mean squared distances matrix, $D_{ij}(t)$, for copolymer sequence $'(ba)_3b_2(ba)_9a_2(ba)_5b_2a_2(ba)_4'$ at the time moment $t = 45$, at which the system is close to the final equilibrium. . . . | 178 |
| 6.1 | Phase diagram of the homopolymer solution in variables of the concentration, c , (in the logarithmic scale) and the second virial coefficient, $u^{(2)}$ | 184 |
| 6.2 | Plots of the mean squared inter–chain distance, Γ^2 , and the mean squared radius of gyration of a chain, R_g^2 , vs the second virial coefficient, $u^{(2)}$ | 185 |
| 6.3 | Plot of the quantity γ vs the concentration, c , for $N = 100$, $L = 100$ and $u^{(2)} = -15$ | 186 |
| 6.4 | Plot of the mean squared cluster size, R_{clus}^2 , vs the second virial coefficient, $u^{(2)}$, for different cluster states. | 187 |
| 6.5 | Plot of the mean number of clusters, $\langle n_{clus} \rangle$, vs the Flory interaction parameter, χ , in the lattice Monte Carlo model of homopolymer solutions. | 189 |
| 6.6 | Plot of the mean squared total radius of gyration, R_{tot}^2 , vs the Flory interaction parameter, χ | 190 |
| 6.7 | Plot of the specific mean energy, $\langle H \rangle / NM$, vs the Flory interaction parameter, χ | 191 |
| 6.8 | Plot of the specific heat capacity, c_V / k_B , vs the Flory interaction parameter, χ | 191 |
| 6.9 | Typical conformations in the homopolymer solution at a finite concentration (a, b). | 192 |

| | | |
|------|--|-----|
| 6.10 | A typical conformation in the homopolymer solution at a finite concentration corresponding to the precipitate ($\chi = 0.5$). | 193 |
| 6.11 | Plot of the number of clusters, n_{clus} , vs the Monte Carlo time, t , after quenches near the transition line. | 195 |
| 6.12 | Typical conformations in the homopolymer solution during the kinetics of precipitation after a quench near the transition line. . . | 196 |
| 6.13 | Plot of the average number of clusters, $\langle n_{clus} \rangle$, vs the Monte Carlo time, t , after quenches well beyond the transition line. | 197 |
| 6.14 | Plot of the specific mean energy, $\langle H \rangle / MN$, vs the Monte Carlo time, t , after quenches well beyond the transition line. | 198 |
| 6.15 | Plot of the mean squared radius of gyration of a chain, R_g^2 , vs the Monte Carlo time, t , after quenches well beyond the transition line. | 198 |
| 6.16 | Phase diagram for solution of 4 $(ab)_9$ copolymers in terms of the amphiphilicity, Δ , and the mean second virial coefficient, $\bar{u}^{(2)}$. . . | 200 |
| 6.17 | Phase diagram for solution of 4 $(a_3b_3)_3$ copolymers in terms of the amphiphilicity, Δ , and the mean second virial coefficient, $\bar{u}^{(2)}$. . . | 200 |
| 6.18 | Typical conformations in $(a_3b_3)_2$ copolymer solution at a finite concentration (a, b) from the Monte Carlo simulation. | 204 |
| 6.19 | Typical conformations in $(a_3b_3)_2$ copolymer solution at a finite concentration (c, d). | 205 |
| 6.20 | Typical conformations in diblock copolymer, (a_6b_6) , solution at a finite concentration (a, b). | 206 |
| 6.21 | Typical conformations in diblock copolymer, (a_6b_6) , solution at a finite concentration (c, d). | 207 |

Abstract

This thesis is devoted to studying the equilibrium structures and kinetic phenomena in coarse-grained statistical mechanical models of macromolecular solutions using a combination of analytical and computational techniques. For this the Gaussian Self-Consistent method and Monte Carlo techniques on a lattice and in continuous space have been developed and applied to the problem. The systems under study include homopolymers with varying flexibility and specific interactions, and amphiphilic heteropolymers with periodic and random sequences in a wide range of solution concentrations.

The equilibrium phase diagrams for these systems are determined. The conformational structures and various observables are examined in different equilibrium and metastable states. Kinetics of conformational transitions after an instantaneous quench, which brings the system from one equilibrium state to another, is studied. A special attention is paid to the kinetics of folding and precipitation of polymer chains. Attendant kinetic laws are elucidated for the homopolymer, and the kinetic exponents for near-equilibrium conditions are deduced. Also, various aspects related to topological restrictions of nonphantom chains are investigated. A number of novel conformational structures, such as the stable mesoglobules in copolymer solutions and toroidal globules with distinct winding numbers for the stiff polymer, are discovered.

The thesis is based on the papers Refs. [I-IX], which have been published or accepted for publication.

Acknowledgements

First and foremost, I would like to thank my research director and supervisor Professor K.A. Dawson for the choice of a very interesting subject for the thesis and continuous support. I am grateful to the external examiner Professor P.G. Wolynes for the attentive reading of the thesis. I would like to thank especially my second supervisor and colleague Dr E.G. Timoshenko, in collaboration with whom most of this work was done, for his help and guidance.

I am grateful to Professors A.Yu. Grosberg, A.R. Khokhlov, R. Lynden–Bell, P. Pincus, R.V. Polozov and Drs A.V. Gorelov, A. Moskalenko, Yu.A. Rochev, D.A. Tikhonov for numerous fruitful discussions. I also would like to thank Dr W.E. Waghorne and Mrs A. Farrell for their support and encouragement, as well as A. Byrne, D. Hegarty, D. Jennings and P. Kiernan for miscellaneous assistance and useful advice.

Yuri A. Kuznetsov

“All theoretical chemistry is really physics; and all theoretical chemists know it.”

Richard P. Feynman

Chapter 1

Introduction

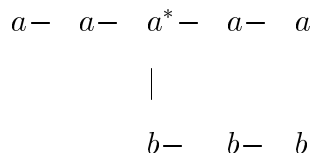
The term *macromolecule*, as well as the macromolecular hypothesis, were suggested by H. Staudinger in 1920-30th. Many significant ideas in this area are also associated with the name of prominent American physical chemist P.J. Flory. Macromolecules present great interest to a number of different areas of science, especially synthetic and physical chemistry, molecular biology, biochemistry and biophysics. Industrial applications are as diverse: chemical, pharmaceutical and food industries, as well as the emerging biotechnology. Among main applications of synthetic and semi-synthetic polymers one can mention plastics (polyethylene, PVC, polystyrene), fibres (nylons, polyester and acrylic fibres), elastomers or rubbers, as well as coatings and adhesives (PMMA and PVAC) [1,2].

Examples of practically useful polymeric materials are numerous: natural polymers such as cellulose, gums and natural rubber, starch, dextran, proteins and nucleic acids; industrial derivatives of natural polymers such as sodium carboxymethyl cellulose, rayon and vulcanised rubber; and purely synthetic polymers such as polyethylene, Teflon (polytetrafluoroethylene), polystyrene $[-\text{CH}_2\text{CH}(\text{C}_6\text{H}_5)-]_N$, Perspex (poly (methacrylate)), terylene and the nylons (poly (hexamethylene adipamide)).

Long polymer molecules are built from a large number of repeated low molar

mass base units, i.e. *monomers*, linked by primary valence bonds [1] (Greek $\pi\omicron\lambda\nu\mu\epsilon\rho\sigma$ means “many parts”). If ‘ a ’ is a monomer molecule and ‘ $-a-$ ’ is the base unit, then the polymer molecule is represented by the formula $\dots a - a - a - a - a \dots$, or $[-a-]_N$, where N is called the degree of polymerisation. For example, poly (vinyl chloride) (PVC) can be presented as, $[-\text{CHClCH}_2-]_N$. Such a polymer consisting of identical monomers is called *homopolymer*.

Polymer chains can form not necessarily linear structures, but also can be branched (e.g. star or comb polymer), form networks (3-dimensional or platelets) and ladders (by repetitions of cross-links between two chains). If a polymer is formed from two types of monomers ‘ a ’ and ‘ b ’ it is called *copolymer*, if from three types of monomers ‘ a ’, ‘ b ’ and ‘ c ’ it is called *terpolymer*. We shall call all of these generally *heteropolymers*. Examples of copolymers are: alternating $[-a - b-]_N$, diblock $[-a-]_N - [-b-]_M$, random $a - a - a - b - a - a - b - b - a - b - a - a$, and graft



copolymers. Proteins are examples of random copolymers consisting of 20 possible monomers, DNA and RNA consist of possible 4 types of nucleotides [3].

High polymers are synthesised from simple molecular elements by series of chemical reactions of two basic types: polymerisation (radical, anionic, cationic or stereospecific), $[-a-]_N + [a-] \rightarrow [-a-]_{N+1}$, and polycondensation, $[-a-]_N + [-a-]_M \rightarrow [-a-]_{N+M}$, in the bulk monomer material or in solution. Biological polymers are assembled *in vivo* by specialised processes: double stranded DNA is replicated by the DNA polymerase according to the Watson-Crick pairing; proteins are assembled from peptide units inside the ribosome by translation of the genetic code in the messenger RNA obtained by transcription of the DNA sequence.

The sequence of monomers forming a polymer is referred to as the *primary* structure. The term *configuration* refers to the structure of chemical bonds and electronic shells, so a configuration can be changed only by breaking bonds or forming new ones. The term *conformation* refers to the spatial arrangement of different parts of the chain, and one conformation can be changed to another e.g. by rotation around bonds. Thus, most generally, we can identify conformation with the set of spatial coordinates of all atoms (or monomers in a coarse-grained description) in the macromolecule, subject to bond, steric and other constraints.

The primary structures of biopolymers are products of a long biological evolution and express hereditary features. Conformations of biopolymers are conveniently divided into a few levels of structure. On a small scale one can distinguish elements of the *secondary* structure. These include alpha-helices, beta-sheets, U-turns, loops and random coils. The relative spatial positioning of groups of atoms in a biopolymer is called the *tertiary* structure. Tertiary structure of a globular protein is essentially fixed and predetermined by its primary structure. And lastly, one sometimes uses the term *quaternary* structure to describe complexes of a number of distinct biopolymers or complex multimeric proteins. Moreover, a given biopolymer is a *monodispersed* compound, e.g. a given protein presents always the same sequence of monomer residues (up to 20 amino acids types).

On the contrary, synthetic polymer systems are generally *polydisperse*, i.e. a given sample would contain polymers of various size resulting from a particular synthesis procedure. The polydispersity parameter is defined as, $p \equiv \langle N^2 \rangle / \langle N \rangle^2 - 1$. In order to prepare a relatively monodisperse sample various fractionation procedures (precipitation fractionation or exclusion chromatography) should be adopted. Synthetic polymers also suffer from isomerism defects. The origins of those are the following.

The carbon has 4 valence bonds of the tetrahedral type. Consider 4 successive

carbon atoms C_1, \dots, C_4 . Such a configuration is characterised by the angle φ between the two planes passing through the common line $C_2 C_3$, and $C_1 C_4$ respectively. The torsional potential has the absolute minimum at $\varphi = \pi$, that corresponds to the ‘trans’ configuration t , and two other minima at $\varphi = \pm\pi/3$ called ‘gauche’ configurations g^+ and g^- . The difference between the depths of the gauche and trans configurations is somewhat lower than $k_B T$ at the room temperature. Thus, due to thermal motions, transforming one rotational state to another, a long chain would have considerable flexibility. Any existing side group would add to the total number of conformations via combination of their own rotational states.

If a monomer unit is asymmetrical, the ‘head–tail’ isomerism generates periodicity defects. Another type of isomerism is related to the chirality. This happens if e.g. there is an asymmetric carbon and the side groups may appear indifferently on either side of the skeleton plane. The setting of these groups with respect to the skeleton plane is called *tacticity* (stereo isomerism): if all side groups are in the same side the chain is called isotactic; if they alternate the chain is syndiotactic; and if they are disordered the chain is atactic.

Due to polydispersity and isomerism defects the elements of structure are not as strictly expressed for synthetic polymers as they are for biological, for which the chirality, tacticity and the primary structure are strictly determined.

The structure of covalent bonds is fixed at room temperature since the energy for breaking such a bond is rather high. However, a linear macromolecule is hardly a very rigid object. First, even rather small harmonic fluctuations in bending of sections of a very long chain lead to its coiling. This type of flexibility is very general and is called *persistent flexibility* since it is uniform along the whole chain. For many stiff and helical polymers persistent flexibility is most significant. Second, a linear polymer molecule exhibits considerable flexibility due to rotation around

carbon–carbon and other bonds. This rotational–isomeric flexibility is also very important for many polymers.

Although the detailed mechanisms of flexibility may be very specific, one is really interested only in the overall effect and ways of characterising it quantitatively. Generally, the directional correlation of two segments of a linear chain decreases approximately exponentially with the separation, s_{ij} , along the chain, $\langle \mathbf{u}_i \mathbf{u}_j \rangle = \exp(-s_{ij}/\lambda)$, where the constant λ is called the *persistent length* of the chain. Such persistent length can be considered as a maximum length at which the chain remains straight, at larger lengths bending fluctuations being too significant. For example, for polystyrene $\lambda \simeq 1.4$ nm, which corresponds to about 5 chain bonds, whereas for the double helix DNA $\lambda \simeq 50$ nm, i.e. about 150 base pairs.

Quite a lot of polymers have various applications while in bulk state since they are characterised by high viscosity and elasticity. Many polymers are also soluble in liquids composed of small molecules. There are three major types of solvents: nonpolar solvents like benzene, polar solvents like nitrobenzene, and solvents with hydrogen bonds like water. While the microstructure of the water molecule and hydrogen bonds contribute significantly into the hydration properties of biomacromolecules, the microstructure of the solvent is not as essential for conformational structure of many synthetic macromolecules.

Long polymer chains are generally soluble in low molecular mass liquid solvents. This means that polymer chains in solution repel each other. Weak van der Waals forces between polymer chains are counterbalanced here by entropic contribution. We refer to this situation as the *good solvent*, in which each chain is very swollen and is called *Flory coil*. On changing temperature, or adding something into the solvent, the repulsive entropic contribution can diminish and each polymer experiences contraction forces that cause its *collapse* in the *poor solvent*.

An important parameter, which characterises the size of a polymer in solution, is the *statistical (Kuhn) length*. Such characteristics was introduced for the *ideal* macromolecule by W. Kuhn, E. Guth and G. Mark. For a long chain it is formally defined as,

$$\langle \mathbf{R}_{0N}^2 \rangle = l^2 N, \quad (1.1)$$

where \mathbf{R}_{0N} is the end-to-end distance. However, the problem with this definition is that a polymer can be considered as ideal in solution only in a rather narrow region of parameter space around the so-called *theta-point*, at which repulsive and attractive interactions compensate each other. This notion is analogous to that of the Boyle point, $u_2 = 0$, in the van der Waals gas.

In solution the shape (conformation) of macromolecules changes continuously due to thermal motion. A reasonable theoretical model should treat a polymer molecule as a kind of random coil. However, such a description is too simplistic. Rotation around bonds restricts flexibility, whereas steric and excluded volume effects do not permit many of possible conformations. Therefore, dissolved linear molecules tend to be more extended than the random coil. The actual magnitudes of monomer-monomer and monomer-solvent interactions determine the resulting conformational state. If parts of polymer chain experience overall attraction a compact conformation, and possibly precipitation, results, whereas strong solvation and electrical repulsion would favour the extended coil conformation.

Polymer chains can exhibit an enormous number of conformations in solution. Such a large number of accessible conformations makes it essentially impossible to use Molecular Dynamics simulations for sufficiently long chains. Typical times are about 0.1 picoseconds for local oscillations, requiring MD time steps of order 10 femtoseconds, whereas a typical diffusion time of a chains could be up to seconds or hours, and collective phenomena and relaxations near the glass transition may take much longer. Conformational transitions of a polymer, such as its collapse,

also involve rather spectacular changes. For example, a double helix DNA is packed into the cell nucleus, which is only about 10^3 nm in size, but while in the ideal state it has size about $0.1\text{mm} = 10^5$ nm, and if unwound into a straight line it takes about a meter long.

This presence of very small and very huge numbers for spatial and time scales and an astronomically large number of accessible conformations in real polymeric systems naturally favours the use of statistical mechanical methods as the only adequate means. Thus, one can expect that many properties of solutions are sufficiently universal to permit description in terms of simpler coarse grained models instead of the over-detailed atomistic description. Such a coarse graining procedure is supported by the concepts of the renormalisation group, but in practice any systematic application of the latter turns out very complicated for polymers. Instead, one attempts to build semi-phenomenological models inspired by the common wisdom of the statistical mechanics and thermodynamics and to study them by approximate and computational methods available at the disposal of that science.

Such approach permits to make a considerable progress as regards the most general thermodynamic and many conformational characteristics of relatively simple polymers. General behaviour of models of extremely complex polymers can be also addressed in part. However, one should clearly realise that in order to be able to predict any *specific* properties of concrete biopolymers, e.g. the tertiary structure of a folded protein with a given primary sequence, or hydration of a given DNA sequence, a more detailed molecular description is inevitably required. Moreover, some conformational changes, such as switching of an enzyme during a reaction that it catalyses, do require computations based on Quantum Chemistry.

* * *

Now let us briefly discuss the main classical methods and models used for simulating macromolecular solutions. Since the literature in this area is rather extensive only a number of key references and techniques will be mentioned, without pretending on any comprehensive coverage or keeping exact chronological precedence.

Molecular Dynamics simulation [4–6] samples phase space of the system by numerically integrating the Newton’s equations of motion. Such calculations can account for thermal motion as molecules may possess enough kinetic energy to cross potential barriers. Molecular Dynamics calculations provide information about possible conformations, thermodynamic properties and dynamic behaviour of molecules. Unfortunately, as we have discussed above, their application to study of long macromolecules is essentially impossible due to the huge computational expenses involved.

Instead of explicitly including solvent molecules, the method of Langevin Dynamics [4–6] models the effect of solvent molecules through random forces applied to a macromolecule to simulate collisions, and frictional forces are added to model the energy losses due to these collisions. Langevin Dynamics simulation can be used to study time dependent properties of solvated systems at a nonzero temperature. Because of the implicit treatment of the solvent, this method is better suited for studying macromolecules in solution. However, the computational expenses for large macromolecules are still quite extensive here, especially if one is interested in reaching equilibrium.

Since the invention of the Metropolis algorithm [7] in 1953 Monte Carlo techniques have become probably the most widely used tool of computer simulations for many problems of condensed matter [5, 8, 9] and polymers in particular [6]. The main idea of the Monte Carlo method is to sample phase space of the system by generating random conformations with the Boltzmann distribution at a

given temperature. This method can be used to find the equilibrium statistical averages and thermodynamic potentials, and, with some reservations, even some dynamic characteristics [10], of rather complex interacting systems. Based on the detailed atomistic representation of even simplest polymers, such as polyethylene, equilibration times are still very long even in Monte Carlo simulation.

An alternative approach, which reduces the computational task is based on eliminating the fastest degrees of freedom on very small scales, i.e. getting rid of the local structure and focusing on the long wave-length behaviour. In this way one arrives at a coarse-grained model with effective monomers and effective bonds as we discussed above.

The simplest type of models here, either based on a lattice [11,12] or the bead-and-spring chain in continuum [13] — are precisely the types of models adopted in standard polymer theories. Monte Carlo simulation can be straightforwardly realised in continuous space for the latter using the harmonic, or slightly anharmonic springs to represent finite extendibility [14], either with hard spheres [15], or a soft repulsive potential [16], and a Lennard-Jones attractive tail [17, 18].

For studying properties of athermal solutions in the repulsive regime even a simpler approach based on self-avoiding walks (SAWs) on a lattice is possible [6, 8, 19]. In this approach averaging is performed over all randomly generated SAWs from an acceptable class without the need to perform the Metropolis check. This approach, however, is too idealised and only applicable to the most universal properties of long flexible chains in the good solvent.

The Verdier-Stockmayer algorithm [20–24] is based on Monte Carlo moves on a lattice with each site occupied by at most one bead, so that the chain cannot cross itself and excluded volume restrictions are automatically satisfied. Polymer links in that model lie along the Cartesian axes and elementary Monte

Carlo moves include: end–bond rotation, kink–jump motion, and, in the generalised Verdier–Stockmayer algorithm, 90° crankshaft rotation. Later other more sophisticated moves were introduced to improve convergence of the system to equilibrium. These are the “slithering snake” (reptational) motion [25] and the pivot motion [26, 27], which involves a rotation of an end chain section. It has been realised, however, that the algorithm without pivot motions is not, strictly speaking, ergodic since there are certain locked–in conformations that cannot be neither reached nor left. Nevertheless, the resulting error in estimates for statistical averages due to the disregard of such conformations seems to be very small for sufficiently long chains. On the other hand, reptations and pivot motions do not correspond to any realistic chain motion and, strictly speaking, cannot be used in dynamics.

There is also a model that, in some sense, is intermediate between the lattice SAW and the off–lattice bead–and–spring model, known as the bond fluctuation model [28–31]. Although the effective bond length in it may vary over a wide range, this approach has turned out very useful for studying melts and the glass transition in various systems.

One should emphasise that development of an efficient Monte Carlo algorithm for a particular problem is often *art* as much as *science*, otherwise large computational expenses would make its study beyond one’s practical reach. Among the principal drawbacks of any Monte Carlo method one should mention the following ones.

First, even though ergodicity may be formally satisfied, the time required to go from one conformation to another may be much longer than any realistic computational time attainable. Such *quasi–nonergodicity* poses serious problems for finding the proper equilibrium distribution at low temperatures, or if there are strong competing interactions. Typically, the system can be trapped for a long

time in some low energy local minimum and does not explore the phase space enough to achieve a good sampling. This makes study of the glass transition particularly difficult, even using various tricks such as simulated annealing [32,33]. Second, for extended objects it may be practically very hard to satisfy the *detailed balance* condition. If that is the case sampling would not reproduce the Boltzmann distribution, but should start to converge to it only asymptotically. Even though the detailed balance is not the necessary, but rather the sufficient condition for producing equilibrium averages, one still arrives at a difficult problem. Indeed, the global *balance condition* is hard to satisfy even approximately without including some collective moves. Practically, the accuracy of such a treatment is very hard to determine, and collective moves are generally computationally expensive. Third, the Markov process generated by a Monte Carlo algorithm with local, and possibly a few collective moves, is hardly a very good representation of the Master equation. Thus, application of Monte Carlo techniques to study of dynamic and kinetic properties is justified only in some cases if the moves included in the scheme are the only relevant motions in the system.

Thus, although computer simulation is a very valuable tool in this arena, one cannot underestimate the importance of analytical methods of statistical mechanics, which do not have these particular problems and limitations. Let us briefly overview such most important approaches here.

The equilibrium properties of polymers in solution have been thoroughly studied for many years using well founded scaling [11,12,34–36] and more sophisticated statistical mechanical methods [13,37–39]. The most fundamental problem here is to describe various macromolecular conformational changes, such as the coil-to-globule, helix-to-coil, and transformations between the dense globular states including the micro-phase separation, crystallisation, renaturation of protein-like biomolecules and formation of toroidal and even more complicated structures.

Thus, although these phenomena do not qualify as true phase transitions for finite-length chains in dilute solution, there is reason to believe that they possess certain generic features, characteristic laws, and even, in some circumstances, universal exponents.

Generally, polymer solutions are characterised by delicately balanced interactions, so that for some range of conditions the polymer chains behave as Flory coils, in others the chains are quite well packed, and there are some intermediate situations [40–44]. The transition between these regimes of Flory coil and condensed collapsed state [45] is of considerable interest to scholars of statistical mechanics due to its intrinsic beauty, as well as a simple model of biopolymer folding.

Probably the most well understood is the collapse transition of a flexible homopolymer. Thus, as a function of temperature, pH, salinity or other control parameters it is possible to cause the polymer to undergo a precipitous change from an extended coil state to a collapsed globular form. Since the introduction of the concept of the coil-to-globule transition in polymer physics [46] the phenomenon has attracted widespread interest, both theoretical and experimental [47–49].

Most of the early experimental work has been carried out near the upper critical solubility temperature of polystyrene in cyclohexane. In recent years there are numerous works devoted to the behaviour of water-soluble polymers near a lower critical solubility temperature (LCST). The most common systems are poly-oxyethyleneoxypropylene block copolymers and poly-N-isopropylacrylamide (PNIPAM) homopolymer, both of which have a LCST near room temperature. The phase behaviour of PNIPAM has been widely investigated in recent years [50, 51].

Although the contraction of chains beyond the theta-temperature has been

observed many times [47–49], compact isolated globules have probably never been observed at equilibrium for homopolymers in pure solvent. The apparent extent of contraction of the coil that is achieved varies with molecular weight and experimental conditions. It has been argued, however, that the compact state has been obtained in kinetic experiments in the presence of large aggregates [47, 52].

Various aspects of the problem have been theoretically studied using simple Flory–type mean–field theory [12, 53, 54], Lifshitz theory of the globule [38], variational method in terms of the Fourier modes [55–57] and, most prominently, the renormalisation group techniques [58–60]. The latter, in conjunction with the relation of self–avoiding walks to a ferromagnet with an n -component magnetisation in the limit $n \rightarrow 0$ [11] proposed by de Gennes, has led to elucidation of the set of critical exponents for polymers [37].

Other motifs in monomer interactions, such as stiffness or the local hydrophobicity [61–65], can not be attributed to the self–avoidance and hence analysed in the framework of the above field theoretical approaches. Study of block and random copolymers therefore presents a more complicated technical task. These chains can exhibit various types of the micro–phase separation [64, 66–68] and undergo freezing transition [64, 65], which is characteristic of most of disordered systems such as spin glasses [32] possessing frustration due to competing interactions. Detailed understanding of the equilibrium properties of heteropolymers and in particular the ability to predict the structure of the globule for arbitrary monomer sequence is a great challenge in this area, for it should provide valuable information for synthesis and design of copolymers carrying out specific functions.

Since behaviour of macromolecular solutions changes significantly across the range of interaction parameters, different types of theories are best suited for different regimes. For solutions at high concentration and melts the description

in terms of density variables [13, 37, 69] is well justified. There has been a significant amount of theoretical work carried out on concentrated copolymer solutions, mixtures and blends [11, 13, 37, 38, 66–68, 70–78] using various types of the density formalism.

Stiffness could be introduced into the theory in a number of physically equivalent ways. There is a significant amount of theoretical [79–83] and experimental [84–88] literature dealing with various questions about equilibrium properties of rigid chains, the most important practical example of which is DNA. Experimentally, it is well known that DNA can acquire a torus-like shape in its globular state [38], and that condensation of DNA induced by various agents could lead to even more complicated phases [84–88]. In theory the torus shape has been predicted in Ref. [89]. The physical reason for a torus is clear — a persistent chain has no desire to bend, so it tends to have as large a radius of curvature as possible, consistent with quite close packing of the chain. Thus, it wraps around itself forming a hole in the middle of the torus. Nevertheless, to construct a good qualitative theory of such states is not very simple [80–82].

In the framework of the bead-and-spring model to account for the stiffness it is sufficient to introduce the bending energy, $\sum_i (d^2 \mathbf{X}_i / di^2)^2$, as first proposed by Kratky and Porod in 1949 in Ref. [90] and used by Harris and Hearst in Ref. [91] for the description of DNA. Alternatively, one can use the model of the freely-rotating chain discussed e.g. in Refs. [92–94].

These equilibrium studies being of utmost importance in themselves for characterising the structural properties of polymeric systems, serve only as a preliminary for study of the dynamic and kinetic phenomena, a matter to which we currently turn. Dynamic information is vital for extracting rheological properties of such materials, i.e. their response to external perturbations and applied stress,

which determine stability and specific functionality of materials. In some practical situations the equilibrium state of the system may be difficult to maintain. Therefore, one would like to understand the system behaviour upon an instantaneous, or rather quick, change of external parameters. This might leave the polymer in the same conformational state, or initiate a kinetic process of phase transformation.

Kinetic processes involving biopolymers, such as protein folding [62, 65, 95–104], occur incessantly in cells maintaining life. They frequently accompany technological processes in polymer chemistry whenever temperature or concentration of various agents changes. Depending on the application, a conformational change of the polymer could be desirable, in which case one would like to control it in such a way as to lead the polymer to the required final state, or undesirable, in which case one would like to suppress it.

Although computer simulation of polymer kinetics has proved very fruitful [105–109], analytical techniques that permit to study dynamic and kinetic phenomena are much less well developed, and the traditional methods of non-equilibrium statistical mechanics [110–112] are not so readily applied to polymers, networks and other such materials.

The kinetics of the conformational changes of a single polymer chain is presently understood only in part. The experimental situation in this area is gradually improving [47, 113] providing valuable inputs for the theorists. Most popular techniques for experimental study of polymer solutions include electron and fluorescent microscopy, dynamic and static light scattering [114], small angle neutron and X-ray scattering, and NMR. Even at equilibrium to avoid a side-effect of aggregation one has to deal with a very dilute solution, making precision measurements difficult for the dynamic light scattering techniques [114]. In kinetics there are a number of other difficult issues involved [47] and still much

remains to be done.

Important theoretical progress in attacking the equilibrium issues has been made using various mean-field type approaches. However, it is worthwhile to emphasise that such approaches are not valid neither at infinitely low dilution, where the fundamental interactions of the individual macromolecule determine its conformational state, nor at nonequilibrium conditions. Moreover, one may seriously doubt whether they are adequate for description of complex inhomogeneous systems. Relatively little is also known about the structures that could be formed within the two phase separation region for copolymer solutions at sufficiently low concentrations.

One interesting theoretical method for study of kinetic phenomena relies on the Gaussian self-consistent approach [115–117], that replaces the exact non-linear Langevin equation by a linear stochastic ensemble with unknown time-dependent parameters determined in a self-consistent way. Such an approximation allows one to obtain closed equations for the self-consistent potentials and friction, that are suitable for further analytical and numerical solution. The latter can be used for evaluation of observable quantities at equilibrium and during kinetics, e.g. the dynamic structure factor of light scattering, and therefore the theory may be verified by experiment [118].

At equilibrium this method reduces exactly to the Gibbs–Bogoliubov variational estimate widely used in many different fields. For polymers the variational treatment was proposed by des Cloizeaux [119] and Edwards [120], and later generalised in various works such as those of Ref. [55]. Unfortunately, it is known to yield an incorrect Flory exponent for the good solvent condition due to an improper probability distribution at small distances. Nevertheless, this drawback of the Gaussian theory is currently well understood and may be resolved, either fundamentally by improving the distribution function [37], or *ad hoc* by enforcing

a short range cut-off [121]. Despite this limitation, even the original version of the method turns out quite adequate not only around the theta-point, where the monomer distribution is Gaussian by definition, but also in the dense globular state. Generalisations of the method to kinetics of the homopolymer have been considered by our Group [116,122] and others [123–126].

In papers [116, 122] kinetics at the collapse of the homopolymer has been discussed in some detail and the notion of several kinetic stages governed by particular universal laws has been introduced. The necklace mechanism of the early kinetic stage, describing the formation and growth of a quasiperiodic array of clusters along the chain, has been explained and studied there using the concept of the spinodal decomposition in the internal metric of the chain. This phenomenon has been also observed in some Monte Carlo simulations [105–107]. Another interesting conclusion is that the characteristic collapse time scales as the degree of polymerisation squared in the absence of the hydrodynamic interaction [13]. This agrees with the prediction of the phenomenological “sausage” model of collapse proposed by de Gennes [127]. These considerations can be also extended to various types of heteropolymers, and the simplest alternating ‘*ab*’ sequence has been considered in Ref. [128].

Apart from the intrinsic theoretical and experimental interest in the heteropolymer solutions there has long been an underlying belief that such studies would lead one closer to the understanding of biopolymer folding as observed in proteins, DNA and other biomolecules.

It is well established that the one-dimensional sequence of amino acid residues, the primary structure, encodes the information about the unique three-dimensional conformation, the tertiary structure, of a folded protein [3, 129–139]. The *native* state, in which a protein exhibits such properties, is thermodynamically preferable under normal conditions within a living cell [140]. In this state

a protein is usually a compact globule with size reaching 10 nanometers. The compactness of the globule in the native state is predominantly maintained by the hydrophobic effect [61–63] that forces the hydrophobic units to be mainly located inside the globule and the hydrophilic ones on the surface. It is further stabilised by other specific interactions such as hydrogen bonding and disulfide bridges. Charges of amino acid residues and steric constraints do play an important role as well. The stability of the spatial structure of proteins is believed to be the source of their ability to reliably function while they are in the biologically active form *in vivo* [140].

Another intriguing problem appears in considering the kinetics of protein folding. There are several different characteristic kinetic folding pathways for the globular proteins studied at present. For example in Ref. [129, 130] the authors consider proteins for which folding starts by a rapid collapse from a random coil state to a semi-compact molten globule, and then proceeds by a slow search to a state from which the chain will eventually access the native state. The number of available conformations of a polypeptide chain grows exponentially with the chain length and is therefore extremely large. Nevertheless, it is known experimentally that proteins do fold into a unique native state in just a matter of seconds. The Levinthal paradox is that the chain is somehow able to find its native configuration without extensively exploring all possible conformations.

It is widely accepted that amongst the essential features required for modelling biopolymers is the presence of *frustration* due to a combination of monomer units with some type of opposing tendency to associate or repel each other, along with the connectivity constraints. It has been recognised that the rugged free energy profile of proteins has an analogy with other disordered systems similar to spin glasses [141].

These frustrated systems have been much studied in the last decade [32, 142–

144]. As relaxation times of glasses are typically very large, the system is trapped in a low energy state separated from others by high potential barriers. The analytical methods that have been traditionally used for studying such problems originated from the theory of spin glasses, and rely mainly on the replica formalism [32].

However, these are more difficult to apply to polymers, particularly if dynamical properties and kinetics of conformational transitions are concerned. Such transitions may be accompanied by a change of not only the characteristic order parameters describing phase separation, freezing and so on, but also of the fractal dimensions of the system, both in the extrinsic space, in which it is embedded, and in the internal metric of the chain.

One popular view is that the low energy states of proteins correspond to minima of the free energy, and the protein may be considered as “kinetically arrested” in one of its dominant conformations. This *freezing transition* occurs upon reducing the “effective temperature” of the polymer system. Finally, the protein reaches the unique global minimum of the free energy.

The statistical–mechanical approach to protein folding is based on the investigation of the properties of simple models of heteropolymers. The simplest approximation for a protein is a *random heteropolymer*. In the sequence model [64] monomer types are represented by random variables with a given distribution of disorder (e.g. binary or Gaussian), that determine the excluded volume interactions between pairs of amino acids. There have been extensive studies of the equilibrium properties of the model carried out using the replica trick of the spin glass theory [64, 65, 77, 102, 145].

A limitation of these approaches lies in their validity only for globular states, where it is possible to use the ground state dominance and constant density approximations. We would like to emphasise that to really understand the protein

folding problem, one needs a method which is equally reliable for both the extended coil and compact globular states.

The kinetic mechanism responsible for the restriction of accessible conformations during folding is probably also related to the kinetics of spin glass-like systems [129, 130]. However, in theory this connection has not as yet been really explained. There have been proposed several approaches to the explanation of the Levinthal paradox dealing with folding kinetics along the so called “preferred pathways” on the free energy landscape [95, 96, 98, 99, 101, 103, 146, 147]. These approaches share some common features in that they explain the tendency of folding towards the native conformation, but they differ in many details. Although there are numerous analytical works on the equilibrium aspects of the problem [64, 65, 77, 102, 145], the knowledge about kinetics of the process is mainly restricted to Monte Carlo simulations on lattice [95, 96, 98, 104, 146, 148, 149].

One recent analytical approach to kinetics of random heteropolymers is based on a version of the Gaussian self-consistent method with a disorder-dependent effective potential for the model with monomers possessing a Gaussian distribution of quenched disorder. It yields a set of self-consistent equations, which are further directly averaged over the quenched disorder in the lowest order of the dispersion keeping the “fully dressed” quantities and applying an enforced closure for the infinite chain of equations. This allows one to avoid the replica trick.

The approach leads to a set of nonlinear and highly coupled differential equations for the main observables of the system. In Ref. [150] the resulting self-consistent equations have been solved numerically for a few examples of kinetics after a quench from the extended coil to globular states. At equilibrium the phase diagram of the model contains five different states separated by the collapse, glass and folding transitions [151, 152].

An important reservation in trying to apply the methods of statistical mechanics to proteins is that protein sequences are very special since they have been carefully selected by the long biological evolution. It has become clear that to ensure kinetic reliability of folding sequences should be more special than simply random. This is important in order to make the native conformation stable even in a model with noisy distorted potentials [153]. The interactions of a kinetically foldable protein should obey the “minimal frustration principle” [99, 101] requiring minimisation of frustration, or in other words the ratio of glass to folding temperatures. One of possibilities to satisfy this principle is to consider “selected” [154], “designed” [155] or “imprinted” [156] sequences. The principles of the energy landscape analysis have been studied and confirmed in many simulations of simple models [95, 96, 98, 103, 146]. It was found that for fast folding sequences there should be a special interplay between the energy frustration and entropic barriers.

These suggestions are certainly interesting and perhaps quite important for explaining the role of evolutionary selection within the “primordial soup”. The final resolution of the protein folding puzzle and elucidation of the kinetic laws that govern this process would shed some light on the questions related to the origin of life in the pre-biotic and early biotic environments and might have implications for the theory of evolution. Practical consequences could also emerge for protein engineering.

The thesis is based on the papers Refs. [I-IX], which have been published or accepted for publication. Further progress will be reported in a number of papers in preparation [X-XIII]. Although the methods and results presented in this thesis have been in part inspired by the underlying interest to understand some generic properties of biopolymers [157], potential applications to biopolymers will not be discussed here. Instead, we shall concentrate on the in-depth study of simple

examples of different types of homo- and heteropolymers. These are more relevant for description of synthetic polymers and further complications would be required to model more complex biological macromolecules.

* * *

The main objective of the current thesis is to study equilibrium structures and kinetic phenomena in coarse-grained statistical mechanical models of macromolecular solutions using a combination of analytical and computational techniques.

In particular, the Gaussian Self-Consistent method is used as the main analytical tool, in conjunction with the standard bead-and-spring model of polymers with volume interactions represented by a virial expansion with site-dependent virial coefficients. The method results in a set of nonlinear first-order in time differential equations for the set of the mean squared distances $D_{AA'}(t)$ between all pairs of monomers. Such equations, either with or without account of hydrodynamics, are solved numerically using a specially designed program (see Appendix B), which implements the fifth order adaptive timestep Runge-Kutta technique in an optimal way and takes the symmetries of the equations into account to improve efficiency. In the stationary limit the GSC equations are reduced to the extrema conditions for the free energy in the variational principle, so that the phase diagrams and metastable states can be determined. In some simple regimes analytical analysis is also possible, which permits to derive some universal laws.

These analytical studies are complemented by lattice (see Appendix A) and continuous space Monte Carlo simulations based on the original C code. On lattice a discrete model is used, which accounts for the excluded volume, connectivity and links integrity constraints, and has the Hamiltonian including local interactions up to the second nearest neighbours. The important advantage of this technique is in its speed and that the time required for one Monte Carlo

step does not depend on the polymer length, so that rather long chains up to $N \simeq 1000$ monomers can be studied. For off lattice continuous space simulation the bead-and-spring model with hard spheres and Lennard-Jones interactions is adopted. This has certain advantages as it is free of the lattice artefacts, but it is more computationally expensive as the time per one Monte Carlo step scales approximately as N here. Thus, although for a chain of about 30 monomers both lattice and off lattice schemes have comparable speeds, systems of more than 200 monomers have not been considered in the latter.

Although molecular architectures of macromolecules may be rather intricate, only the following particular examples of polymers, which are of the main interest in applications, are to be studied: flexible and stiff homopolymers with open and ring topology, homopolymers with three-body short ranged interactions, hydrophilic-hydrophobic heteropolymers with periodic and random sequences of monomers. These are considered as an isolated chain, i.e in the infinitely dilute solution, and in solution at finite concentration in which many chains interact with the solvent and each other.

We study the equilibrium phase diagrams of these systems in terms of the main thermodynamic parameters such as the second $u_2(T)$ and third virial $u_3(T)$ coefficients, the persistent length λ , degree of amphiphilicity Δ , concentration c and some others. Conformational structures corresponding to different equilibrium and metastable states are examined in terms of the $D_{AA'}$ matrix and visually. The main relevant observables such as the mean squared radius of gyration R_g^2 , the mean degree of the micro-phase separation (MPS) Ψ , the mean energy \mathcal{E} , and the Helmholtz free energy \mathcal{A} , as well as some others are obtained.

Kinetics of various conformational transitions after an instantaneous quench that brings the system from one equilibrium state to another is also considered. Special interest is devoted to the kinetics of folding and precipitation of polymers

chains on going from the good to the poor solvent. Attendant kinetics laws are elucidated for the simplest case case of the homopolymer, and the kinetic exponents for near-equilibrium conditions are obtained analytically. Some of these exponents are in addition supported by data from the Monte Carlo simulation. Whenever possible both analytical and computational methods are used and results obtained from both are compared. However, some issues are difficult or too computationally expensive to study reliably by simulation, while for others adequate analytical techniques simply do not exist as yet. For instance, Monte Carlo simulation allows us to make a progress in understanding the role of topological constraints and effects, which are quite significant for the dense compacted states. An important role of the continuous space Monte Carlo is also in that it permits to address the applicability and limitations of the analytical GSC method by determining the actual two-body probability distribution function of monomers.

Overall, the systems under study here are rather complex and they exhibit most interesting behaviour, particularly at nonequilibrium conditions. Although this thesis studies only a limited number of examples of model polymeric systems and regimes, there is quite a number of novel and unexpected phenomena encountered. Finally, the methods, techniques and their numerical implementation developed here are quite general and can be adopted for other applications, which are beyond the scope of this thesis. For instance, gels, both chemically cross-linked and soft [158–160], as well polymers with more sophisticated molecular architecture would certainly be interesting to study along those lines.

“We are all agreed that your theory is crazy. The question which divides us is whether it is crazy enough to have a chance of being correct. My own feeling is that it is not crazy enough.”

Niels Bohr

Chapter 2

The Gaussian Self–Consistent Method

In this chapter we introduce the model and derive the Gaussian Self–consistent (GSC) equations for study of the equilibrium and kinetic phenomena in copolymer solutions at finite concentration. Also we present some details about techniques for solving numerically the GSC equations.

2.1 Model of Copolymer Solutions at Finite Concentrations

Let us consider the canonical ensemble of M macromolecules in a box of finite volume V at temperature T . We denote by \mathbf{r}_A the spatial coordinates of atoms constituting a macromolecule and by \mathbf{R}_a the spatial coordinates of the solvent molecules. Here the index a enumerates the solvent molecules and $A = \{i, m\}$ is the multi–index, where i counts atoms in a given macromolecule and the index m is the number of the macromolecule. The partition function of the system can

be written as the following integral of the molecular interaction potential, H_{mol} ,

$$Z = Z_0 \int \prod_{Aa} d\mathbf{r}_A d\mathbf{R}_a \exp(-\beta H_{mol}[\mathbf{r}, \mathbf{R}]), \quad (2.1)$$

where $\beta \equiv (k_B T)^{-1}$ and Z_0 is the momentum part of the partition function. Here, however, we do not pretend on such a detailed description of the system. The procedure of coarse-graining of the model consists in introducing the collective coordinates, $\mathbf{X}_A = \mathbf{f}_A(\mathbf{r})$, describing coordinates of larger atomic groups, which we call *monomers*. In the simplest case, for example, the collective coordinate of a monomer can be taken as the coordinate of the centre of mass of atoms constituting the monomer. Importantly, despite the molecular interaction potentials are rather complicated and specific, the coarse-grained coordinates permit a simpler and more universal, though approximate, description. Inserting the unity,

$$1 = \int \prod_A d\mathbf{X}_A \delta(\mathbf{X}_A - \mathbf{f}_A(\mathbf{r})), \quad (2.2)$$

we can integrate out the initial coordinates,

$$Z/Z_0 = \int \prod_{Aa} d\mathbf{X}_A d\mathbf{R}_a \exp(-\beta H_{eff}[\mathbf{X}, \mathbf{R}]). \quad (2.3)$$

Note that the multi-index A has a different meaning for the coarse-grained coordinates and the chain indices i, j change in the range $0, \dots, N - 1$, where N is called the degree of polymerisation. The effective free energy functional (EFEF) includes contributions describing the solvent and monomer degrees of freedom and the solvent–monomer interaction term,

$$H_{eff}[\mathbf{X}, \mathbf{R}] = H_{sol}[\mathbf{R}] + H_{mon}[\mathbf{X}] + H_{int}[\mathbf{X}, \mathbf{R}]. \quad (2.4)$$

The effective free energy functional of monomers can be written as a virial-type series in powers of the monomer density function,

$$H_{mon}[\mathbf{X}] = \sum_{J=2}^{\infty} u^{(J)} \int d\mathbf{y} \rho_{mon}(\mathbf{y})^J + H'[\mathbf{X}], \quad (2.5)$$

$$\rho_{mon}(\mathbf{y}) = \sum_A \delta(\mathbf{y} - \mathbf{X}_A), \quad (2.6)$$

where $u^{(J)}$ are the virial coefficients which can be calculated as functions of the temperature and parameters of molecular interaction potentials using the Mayer diagram technique [111,161]. However, for our purposes we do not need to know these explicit dependencies here as we shall keep only a few first terms. Appropriate coefficients then may be viewed as independent phenomenological parameters, which could be directly related to experimentally measurable quantities. The term H' contains other interaction terms,

$$H'[\mathbf{X}] = \frac{k_B T}{2L^2} \sum_m (\mathbf{Y}^m - \mathbf{Y})^2 + \frac{k_B T}{2l^2} \sum_{i,m} (\mathbf{X}_i^m - \mathbf{X}_{i-1}^m)^2 + \quad (2.7)$$

$$+ \frac{k_B T \lambda}{2l^3} \sum_{i,m} (\mathbf{X}_{i+1}^m + \mathbf{X}_{i-1}^m - 2\mathbf{X}_i^m)^2, \quad (2.8)$$

where we have introduced \mathbf{Y}^m and \mathbf{Y} as the centre-of-mass of the m -th chain and the whole system respectively,

$$\mathbf{Y}^m \equiv \frac{1}{N} \sum_i \mathbf{X}_i^m, \quad \mathbf{Y} \equiv \frac{1}{M} \sum_m \mathbf{Y}^m. \quad (2.9)$$

The first term in (2.8) effectively introduces a soft spatial cut-off, L , onto the system. Without loss of generality we may assume the condition $\mathbf{Y} = 0$, i.e. we fix the reference point in the centre-of-mass of the whole system. As for the use of a soft Gaussian cut-off, it should not matter for sufficiently large system which type of the boundary conditions is imposed. The second term in Eq. (2.8) represents the energy of harmonic springs between neighbouring monomers along the chain. It provides the connectivity of each chain with l being the characteristic statistical length. The last term in (2.8) in the continuous limit becomes the integral of the squared curvature of the polymer chain. This is called the bending energy with λ being the persistent length of the chain.

The monomer-solvent interactions can be well approximated via contact interactions,

$$H_{int}[\mathbf{X}, \mathbf{R}] = - \sum_{Aa} I_A \delta(\mathbf{X}_A - \mathbf{R}_a), \quad (2.10)$$

where I_A is the coupling constant of the A -th monomer with the a -th solvent molecule. Using the solution incompressibility condition [162],

$$\rho_{sol}(\mathbf{y}) + \rho_{mon}(\mathbf{y}) = \rho_0 = \text{const}, \quad (2.11)$$

$$\rho_{sol}(\mathbf{y}) = \sum_a \delta(\mathbf{y} - \mathbf{R}_a), \quad (2.12)$$

the solvent degrees of freedom can be integrated out. Thus, the partition function of the system can be written as,

$$Z = Z_0 Z_{sol} \int \prod_A d\mathbf{X}_A \exp(-\beta H[\mathbf{X}]), \quad (2.13)$$

where the EFEF of monomers after exclusion of the solvent is,

$$\begin{aligned} H[\mathbf{X}] &= \sum_{A \neq A'} (u^{(2)} + (I_A + I_{A'})/2) \delta(\mathbf{X}_A - \mathbf{X}_{A'}) \\ &+ \sum_{J=3}^{\infty} u^{(J)} \int d\mathbf{y} \rho_{mon}^J(\mathbf{y}) + H'[\mathbf{X}] + \text{const}. \end{aligned} \quad (2.14)$$

As we can see, the effect of the solvent influence on the monomer degrees of freedom appears via monomer dependent shifts of the second virial coefficient in the EFEF. The average second virial coefficient is, thus, related to the quality of the solvent: positive values correspond to the good solvent, where effective two-body repulsion of monomers results in the extended coil conformation, and the negative values correspond to the poor solvent condition, where the effective two-body attraction tends to compact the chain. The third and higher virial coefficients mainly characterise the repulsion between monomers due to excluded volume interactions and should be kept positive. In a more general situation the coefficients $u_{AA'...}^{(J)}$ are allowed in principle to have any dependence on monomer indices.

Let us consider the following choice of monomer-dependent second virial coefficients in Eq. (2.14),

$$u_{AA'}^{(2)} = \bar{u}^{(2)} + \Delta \frac{\sigma_A + \sigma_{A'}}{2}, \quad (2.15)$$

and $u_{\{A\}}^{(J)} = u^{(J)}$ for $J > 2$. This corresponds to the case of amphiphilic heteropolymers, for which monomers differ only in the monomer–solvent coupling constants, I_A . Then the *mean second virial coefficient*, $\bar{u}^{(2)}$, is associated with the quality of the solvent and the parameter Δ is called the *degree of amphiphilicity* of the chain. The set $\{\sigma_i\}$ expresses the chemical composition, or the *primary sequence* of a heteropolymer. Here we consider the case in which variables σ_A can take only three values: $-1, 1$ and 0 corresponding to hydrophobic ‘ a ’, hydrophilic ‘ b ’ and “neutral” ‘ c ’ monomers respectively.

The above equilibrium construction can be generalised to the realm of kinetics, but it is a rather complicated procedure. Instead, we consider the phenomenological Langevin equation which describes the coarse–grained dynamic of the system on timescales much larger than typical collision times [13]. In a strongly dissipative medium the Langevin equation takes the form,

$$\frac{d}{dt}X_A^\alpha(t) = \sum_{\alpha', A'} \mathcal{H}_{AA'}^{\alpha\alpha'}[\mathbf{X}(t)] \left(-\frac{\partial H[\mathbf{X}]}{\partial X_{A'}^{\alpha'}(t)} + \eta_{A'}^{\alpha'}(t) \right), \quad (2.16)$$

where the Greek letters denote the spatial indices ($\alpha, \alpha' = 1, 2, 3$) and $\mathcal{H}_{AA'}^{\alpha\alpha'}$ is the Oseen hydrodynamic tensor. It describes the backflow effect of the solvent relating the generalised velocity and the force, $v_A^\alpha = \sum_{\alpha' A'} H_{AA'}^{\alpha\alpha'} F_{A'}^{\alpha'}$. Using the Navier–Stokes equation one can obtain the following approximate expression for the hydrodynamic tensor [13],

$$\mathcal{H}_{AA'}^{\alpha\alpha'}[\mathbf{X}] = \frac{\delta_{AA'}}{\zeta_b} + \frac{1 - \delta_{AA'}}{8\pi\eta_s} \frac{\mathcal{P}_{\alpha\alpha'}[\hat{\mathbf{X}}_{AA'}]}{|\mathbf{X}_{AA'}|}, \quad (2.17)$$

$$\begin{aligned} \frac{\mathcal{P}_{\alpha\alpha'}[\hat{\mathbf{X}}_{AA'}]}{8\pi|\mathbf{X}_{AA'}|} &= \int \frac{d\mathbf{w}}{(2\pi)^3} \frac{\mathcal{P}_{\alpha\alpha'}(\hat{\mathbf{w}})}{\mathbf{w}^2} \exp(-i\mathbf{w}\mathbf{X}_{AA'}), \\ \mathcal{P}_{\alpha\alpha'}(\hat{\mathbf{w}}) &\equiv \delta_{\alpha\alpha'} - \hat{w}_\alpha \hat{w}_{\alpha'}, \quad \hat{w}_\alpha = w_\alpha/w, \end{aligned} \quad (2.18)$$

where $\mathbf{X}_{AA'} = \mathbf{X}_A - \mathbf{X}_{A'}$ and η_s is the viscosity of the solvent. Note that the additional term proportional to the bare friction of a monomer, ζ_b , coming from the diagonal elements of the hydrodynamic tensor, may be neglected for large

values of N . The random force, $\eta_{A'}^{\alpha'}(t)$, in the Langevin equation (2.16) is the Gaussian random noise with zero average and the second momentum proportional to the temperature in the system,

$$\langle \eta_A^\alpha(t) \eta_{A'}^{\alpha'}(t') \rangle = (\mathcal{H}^{-1})_{AA'}^{\alpha\alpha'}[\mathbf{X}(t)] 2k_B T \delta(t - t'). \quad (2.19)$$

It represents the effect of random collisions of monomers with molecules of the solvent. This form of the second momentum in (2.19) guarantees the relaxation of the system towards the equilibrium at $t \rightarrow \infty$ independently of the choice of the initial condition [163].

2.2 The Gibbs–Bogoliubov Variational Principle

In this section we give the formulation of the Gibbs–Bogoliubov variational principle. Then choosing the trial Hamiltonian as a most general quadratic form we derive the variational free energy of the system with the EFEF given by Eqs. (2.14, 2.8).

First, let us give a simple derivation of the Gibbs–Bogoliubov variational principle for classical systems. Let us denote by H the total Hamiltonian and by H_0 the trial one and introduce the charging parameter g such that,

$$\exp(-\beta \mathcal{A}(g)) = \text{Tr} \left(\exp(-\beta (H_0 + gH_1)) \right), \quad H_1 \equiv H - H_0, \quad (2.20)$$

where $\mathcal{A}(g)$ is the exact free energy of the system with the Hamiltonian H_g . First differentiation of this expression by g gives,

$$\frac{d\mathcal{A}}{dg} \exp(-\beta \mathcal{A}) = \text{Tr} \left(\exp(-\beta H_g) H_1 \right) = \exp(-\beta \mathcal{A}) \langle H_1 \rangle_g, \quad (2.21)$$

and the second differentiation gives,

$$\left(\frac{d^2 \mathcal{A}}{dg^2} - \beta \left(\frac{d\mathcal{A}}{dg} \right)^2 \right) \exp(-\beta \mathcal{A}) = -\beta \text{Tr} \left(\exp(-\beta H_g) H_1^2 \right). \quad (2.22)$$

Thus, for the second derivative we obtain that it is nonpositive,

$$\frac{d^2 \mathcal{A}}{dg^2} = -\beta \left(\langle H_1^2 \rangle_g - \langle H_1 \rangle_g^2 \right) \leq 0. \quad (2.23)$$

Now by setting $g = 1$ and expanding the total free energy, $\mathcal{A} \equiv \mathcal{A}(1) = \mathcal{A}(0) + \mathcal{A}'(0) + (1/2)\mathcal{A}''(0)$ in Taylor series around $g = 0$ we derive the required inequality,

$$\mathcal{A} \leq \mathcal{A}_{trial} \equiv \mathcal{A}_0 + \langle H - H_0 \rangle_0. \quad (2.24)$$

Thus, the trial free energy, \mathcal{A}_{trial} , gives an upper bound for the exact free energy and by minimising it with respect to the variational parameters of the trial Hamiltonian one obtains the best possible estimate for the free energy.

* * *

Before proceeding with the derivation of the variational free energy for the system with EFEF given by Eqs. (2.14, 2.8) let us first introduce some important observables. The equal time pair correlation function of four monomers is defined in the spatially isotropic case as,

$$D_{AA',BB'}(t) \equiv \frac{1}{3} \langle (\mathbf{X}_A(t) - \mathbf{X}_{A'}(t)) (\mathbf{X}_B(t) - \mathbf{X}_{B'}(t)) \rangle, \quad (2.25)$$

The pair correlations of two and three monomers can be obtained from these as,

$$D_{AA'B} \equiv D_{AA',BA'}, \quad (2.26)$$

$$D_{AB} \equiv D_{AB,AB} = \frac{1}{3} \langle (\mathbf{X}_A - \mathbf{X}_B)^2 \rangle. \quad (2.27)$$

Using these definitions one can derive the following reduction formulae,

$$D_{AA',BB'} = -\frac{1}{2} (D_{AB} + D_{A'B'} - D_{AB'} - D_{A'B}), \quad (2.28)$$

$$D_{AA'B} = -\frac{1}{2} (D_{AB} - D_{AB'} - D_{A'B}), \quad (2.29)$$

Now, let us define the trace as the integration over all monomer coordinates subject to the constraint that the reference point is fixed in the centre-of-mass of the system,

$$\text{Tr} \equiv \int \prod_A d\mathbf{X}_A \delta \left(\sum_A \mathbf{X}_A \right). \quad (2.30)$$

Then the partition function is obtained as $Z_{tot} = \text{Tr} \exp(-H/k_B T)$ with H given by Eqs. (2.14, 2.8). The delta-function in (2.30) removes the trivial divergence in Z_{tot} due to the translational invariance.

We choose the trial Hamiltonian, H_0 , as a generic quadratic form,

$$\frac{H_0}{k_B T} = \frac{1}{2} \sum_{AA'} K_{AA'} \mathbf{X}_A \mathbf{X}_{A'} + \sum_A \mathbf{J}_A \mathbf{X}_A, \quad (2.31)$$

where we have also introduced arbitrary sources \mathbf{J}_A . Using the delta-function one can exclude, say, \mathbf{X}_0 and derive,

$$\begin{aligned} Z_0[\mathbf{J}] &= (2\pi)^{3(MN-1)/2} (\det \tilde{K}^{(MN-1)})^{-3/2} \\ &\times \exp \left(\frac{1}{2} \sum_{AA'} '(\mathbf{J}_A - \mathbf{J}_0) (\tilde{K}^{-1})_{AA'} (\mathbf{J}_{A'} - \mathbf{J}_0) \right), \end{aligned} \quad (2.32)$$

where the $(MN - 1)$ dimensional matrix \tilde{K} is

$$(\tilde{K})_{AA'} = K_{AA'} + K_{00} - K_{A0} - K_{0A'}. \quad (2.33)$$

Then the averages can be calculated as,

$$\mathcal{F}_{AA'} = \frac{1}{3} \langle \mathbf{X}_A \mathbf{X}_{A'} \rangle_0 = \frac{1}{3} \frac{\partial^2 \log Z_0[\mathbf{J}]}{\partial \mathbf{J}_A \partial \mathbf{J}_{A'}} = (\tilde{K}^{-1})_{AA'}, \quad (2.34)$$

where again the latter identity holds only for $A, A' \neq 0$. From these quantities it is straightforward to obtain the mean squared distances,

$$D_{AA'} = \mathcal{F}_{AA} + \mathcal{F}_{A'A'} - 2\mathcal{F}_{AA'}. \quad (2.35)$$

which also holds only for $A, A' \neq 0$.

The ‘‘entropic’’ part of the variational free energy can be calculated from the partition function (2.32) as follows,

$$\mathcal{A}_0 \equiv -T\mathcal{S} = -k_B T \log Z_0[0] = \frac{3k_B T}{2} \log(\det \tilde{K}^{(MN-1)}). \quad (2.36)$$

Let us now express all independent parameters of the matrix \tilde{K} via the mean squared distances $D_{AA'}$. For this we compute the following sums applying the

above delta-function constraint,

$$\sum_{AA'} D_{AA'} = 2MN \sum_A \mathcal{F}_{AA}, \quad \sum_A D_{AB} = \frac{1}{2MN} \sum_{AA'} D_{AA'} + MN \mathcal{F}_{BB}. \quad (2.37)$$

Substituting \mathcal{F}_{AA} from the second formula here to Eq. (2.35) and recalling the relation (2.25) we derive the following expression for the variational entropy via the mean squared distances between monomers,

$$\mathcal{S} = \frac{3}{2} k_B \log \det R^{(MN-1)}, \quad (2.38)$$

where we have used the determinant of the $(MN - 1)$ -dimensional major submatrix of the matrix,

$$R_{AA'} = \frac{1}{M^2 N^2} \sum_{BB'} D_{AB, A'B'}. \quad (2.39)$$

The reason for appearance of the truncated matrix is that we have excluded the zero eigenvalue of R related to the translational invariance.

It can be shown easily that the quadratic form for the trial Hamiltonian H_0 in Eq. (2.31) produces the Gaussian distribution for the monomer-monomer distances, $\mathbf{X}_{AA'}$. Thus, the two-body monomer-monomer correlation function, which is the probability density to find the monomer A' at distance \mathbf{r} from the monomer A , is,

$$h_{AA'}^{(2)}(\mathbf{r}) \equiv \langle \delta(\mathbf{r} - \mathbf{X}_A + \mathbf{X}_{A'}) \rangle_0 = \frac{1}{(2\pi D_{AA'})^{3/2}} \exp\left(-\frac{\mathbf{r}^2}{2D_{AA'}}\right). \quad (2.40)$$

The Fourier transform of this two-body correlation function gives the generating function, which is also called the partial structure factor, of the Gaussian shape too,

$$C_{AA'}(\mathbf{k}) \equiv \langle \exp(i\mathbf{k}(\mathbf{X}_A - \mathbf{X}_{A'})) \rangle_0 = \exp\left(-\frac{1}{2}\mathbf{k}^2 D_{AA'}\right). \quad (2.41)$$

Using the integral representation for the δ -function,

$$\delta(\mathbf{r}) = (2\pi)^{-3} \int d\mathbf{k} \exp(i\mathbf{k}\mathbf{r}) \quad (2.42)$$

and the generating function (2.41) one can calculate the averages of the products of δ -functions appearing in the virial-type expansion in Eq. (2.14). Thus, for the mean energy of the system, $\mathcal{E} = \langle H \rangle$, we derive in a rather straightforward manner,

$$\begin{aligned} \mathcal{E} = & \sum_m \left(\frac{3k_B T}{2l^2} \sum_i D_{ii-1}^{mm} + \frac{3k_B T \lambda}{2l^3} \sum_i (2D_{i+1i}^{mm} + 2D_{i-1i}^{mm} - D_{i-1i+1}^{mm}) \right) \\ & + \frac{3k_B T}{2L^2} M (R_{tot}^2 - R_g^2) + \sum_{J=2}^{\infty} \sum_{\{A\}} \hat{u}_{\{A\}}^{(J)} (\det \Delta_{\{A\}}^{(J-1)})^{-3/2} + \mathcal{E}_{si}, \end{aligned} \quad (2.43)$$

where $\hat{u}^{(J)} \equiv (2\pi)^{-3(J-1)/2} u^{(J)}$ and we have introduced the total mean squared radius, R_{tot}^2 , and the mean squared radius of gyration of a chain, R_g^2 , as follows,

$$R_{tot}^2 = \frac{1}{2N^2 M^2} \sum_{AA'} D_{AA'}, \quad R_g^2 = \frac{1}{M} \sum_m R_m^2, \quad R_m^2 = \frac{1}{2N^2} \sum_{ij} D_{ij}^{mm}. \quad (2.44)$$

In Eq. (2.43) matrices $\Delta_{\{A\}}^{(J-1)}$ depend on indices A_1, \dots, A_J and their elements are the pair correlations of three monomers (2.26),

$$\Delta_{ij}^{(J-1)} = D_{A_1, A_{i+1}, A_1, A_{j+1}}. \quad (2.45)$$

Let us write here the explicit formulae for the determinants appearing in the two- and three-body interaction terms,

$$\begin{aligned} \det \Delta_{AA'}^{(1)} &= D_{AA'} & (2.46) \\ \det \Delta_{AA'A''}^{(2)} &= \begin{vmatrix} D_{AA'} & \frac{1}{2}(D_{AA'} + D_{AA''} - D_{A'A''}) \\ \frac{1}{2}(D_{AA'} + D_{AA''} - D_{A'A''}) & D_{AA''} \end{vmatrix} = \\ &= \frac{1}{2} (D_{AA'} D_{AA''} + D_{AA'} D_{A'A''} + D_{AA''} D_{A'A''}) - \\ &= -\frac{1}{4} (D_{AA'}^2 + D_{AA''}^2 + D_{A'A''}^2). \end{aligned} \quad (2.47)$$

The positiveness of the first determinant requires all mean squared distances to be positive, $D_{AA'} > 0$, which is true as this is the average of a quantity squared. The positiveness of the three-body determinant in Eq. (2.47), on the other hand, is insured by the triangle inequality,

$$D_{AA''}^{1/2} < D_{AA'}^{1/2} + D_{A'A''}^{1/2}. \quad (2.48)$$

It is interesting to note that the quantity $(\det\Delta_{A_1, \dots, A_J}^{(J-1)})^{3/2}$ characterises the volume of the phase space occupied by J particles if the mean squared distances between all pairs are specified. Moreover, the logarithm of this characteristic volume of all MN particles is equal to the variational entropy of the system up to a trivial constant,

$$\det R^{(MN-1)} = \frac{1}{M^2 N^2} \det \Delta^{(MN-1)}. \quad (2.49)$$

We discuss the appearance of the last term, \mathcal{E}_{si} in the mean energy of the system (2.43) later in Sec. 2.4.

It is also important to introduce the micro-phase separation (MPS) order parameter,

$$\Psi = \frac{1}{M} \sum_m \Psi_m \quad \Psi_m = \frac{1}{N^2 R_m^2} \sum_{ii'} \frac{\sigma_i + \sigma_{i'}}{2\Delta_\sigma} D_{ii'}^{mm}, \quad (\Delta_\sigma)^2 = \frac{1}{N} \sum_i \sigma_i^2. \quad (2.50)$$

The MPS parameter describes the degree of correlation between matrices of the relative two-body interaction, $(\sigma_i + \sigma_{i'})/2$, and the mean squared distances, $D_{ii'}$.

Finally, note that the term $\langle H_0 \rangle_0$ is a trivial constant which can be neglected. The Gibbs–Bogoliubov variational principle is then based on minimising this variational free energy $\mathcal{A} = \mathcal{E} - T\mathcal{S}$ in Eqs. (2.38, 2.43) with respect to the $MN(MN - 1)/2$ independent variational parameters $D_{AA'}$.

2.3 General Form of the Gaussian Self-Consistent Equations

In this section we derive the most general form of the Gaussian self-consistent kinetic equations proceeding from the Langevin equation (2.16) for the model of polymer solutions presented above.

The main idea of the Gaussian self-consistent method is to replace the non-linear Langevin equation (2.16),

$$\frac{d}{dt}X_A^\alpha(t) = \sum_{A', \alpha'} \mathcal{H}_{AA'}^{\alpha\alpha'}[\mathbf{X}(t)] \phi^{\alpha', A'} \quad \phi^{\alpha, A} \equiv -\frac{\partial H[\mathbf{X}]}{\partial X_A^\alpha} + \eta_A^\alpha(t), \quad (2.51)$$

where the noise has second momentum given by Eq. (2.19) and the Oseen tensor is given by (2.17), by a linear stochastic ensemble,

$$\frac{d}{dt}\mathbf{X}_A(t) = \sum_{A'} \xi_{AA'}(t) \boldsymbol{\phi}_0^{A'}(t), \quad \boldsymbol{\phi}_0^A(t) \equiv -\frac{\partial H_0[\mathbf{X}(t)]}{\partial \mathbf{X}_A(t)} + \boldsymbol{\eta}_0^A(t). \quad (2.52)$$

Here the trial Hamiltonian, H_0 , is a generic quadratic form, with matrix coefficients called the time-dependent effective potentials,

$$H_0[\mathbf{X}(t)] = \frac{1}{2} \sum_{AA'} \mathbf{X}_A(t) V_{AA'}(t) \mathbf{X}_{A'}(t), \quad (2.53)$$

and $\boldsymbol{\eta}_0^A(t)$ is the Gaussian noise with the second momentum,

$$\langle \eta_0^{\alpha, A}(t) \eta_0^{\alpha', A'}(t') \rangle_0 = 2k_B T \delta_{\alpha\alpha'} \delta(t-t') \xi_{AA'}^{-1}(t). \quad (2.54)$$

The exact Langevin equation (2.16), which includes the hydrodynamic interaction, is rather nonlinear. To make some progress one has to use the preaveraged approximation for the Oseen tensor [13], which is the adoption of the following factorisation,

$$\langle \mathcal{H}_{AA'}^{\alpha\alpha'} f_A^\alpha[\mathbf{X}] \phi_{A'}^{\alpha'} \rangle_0 = \langle \mathcal{H}_{AA'}^{\alpha\alpha'} \rangle_0 \langle f_A^\alpha[\mathbf{X}] \phi_{A'}^{\alpha'} \rangle_0, \quad (2.55)$$

for any function $f_A^\alpha[\mathbf{X}]$.

The Gaussian self-consistent equations are obtained by requiring that the pair correlation $\langle X_A^\alpha(t) X_{A'}^{\alpha'}(t') \rangle_0$ found from the exact (2.16) and the trial (2.52) dynamics coincide. To ensure this it is sufficient to provide that the time derivative of these correlations found from Eqs. (2.16, 2.52) are the same. Using the preaveraged approximation Eq. (2.55) and the Fluctuation-Dissipation Theorem,

$$\langle X_A^\alpha(t) \eta_0^{\alpha' A'}(t) \rangle_0 = \delta_{\alpha\alpha'} \delta_{AA'} k_B T, \quad (2.56)$$

we obtain the following self-consistency conditions,

$$\langle \mathcal{H}_{AA'}^{\alpha\alpha'} \rangle_0 = \delta^{\alpha\alpha'} \xi_{AA'} \quad (2.57)$$

$$\sum_B V_{AB} \mathcal{F}_{BA'} = \left\langle \mathbf{X}_A \frac{\partial H_0[\mathbf{X}]}{\partial \mathbf{X}_{A'}} \right\rangle_0 = \left\langle \mathbf{X}_A \frac{\partial H[\mathbf{X}]}{\partial \mathbf{X}_{A'}} \right\rangle_0, \quad (2.58)$$

where we have introduced the equal-time pair correlations,

$$\mathcal{F}_{AA'}(t) = \frac{1}{3} \langle \mathbf{X}_A(t) \mathbf{X}_{A'}(t) \rangle_0, \quad (2.59)$$

and the effective potentials, $V_{AA'}(t)$. In addition, we have the requirement that the second momenta of the noise from the exact and the trial dynamics are the same. This gives,

$$\left\langle \left(\mathcal{H}_{AA'}^{\alpha\alpha'} \right)^{-1} \right\rangle_0 = \delta^{\alpha\alpha'} (\xi^{-1})_{AA'} \quad (2.60)$$

$$\left\langle X_A^\alpha(t) \eta^{\alpha' A'}(t) \right\rangle = \delta_{\alpha\alpha'} \delta_{AA'} k_B T. \quad (2.61)$$

The latter equation is the Fluctuation–Dissipation Theorem for the exact dynamics.

From Eq. (2.52) one can obtain the following kinetic equations for the equal-time pair correlation $\mathcal{F}_{AA'}(t)$ in Eq. (2.59),

$$\begin{aligned} \frac{d}{dt} \mathcal{F}_{AA'}(t) &= \sum_{BB'} \left(\xi_{AB}(t) (k_B T \delta_{BA'} - V_{BB'}(t) \mathcal{F}_{B'A'}(t)) \right. \\ &\quad \left. + \xi_{A'B}(t) (k_B T \delta_{BA} - V_{BB'}(t) \mathcal{F}_{B'A'}(t)) \right). \end{aligned} \quad (2.62)$$

The matrix convolution $\sum_B V_{AB} \mathcal{F}_{BA'}$ can be calculated from the exact EFEF using the self-consistency equation (2.58). Here we use the technique similar to that of Sec. 2.2 for calculation of the mean energy and the non-equilibrium analog of the generating function (2.41),

$$C_{AA'}(\mathbf{k}, t) \equiv \langle \exp(i\mathbf{k}(\mathbf{X}_A(t) - \mathbf{X}_{A'}(t))) \rangle_0 \quad (2.63)$$

Now let us obtain the expression for this generating function $C_{AA'}(\mathbf{k}, t)$ via the mean squared distances. Solution of Eq. (2.52) may be presented as follows,

$$\mathbf{X}_A(t) = G_{AA'}(t, 0) \mathbf{X}_{A'}(0) + \int_0^t dt' G_{AA'}(t, t') \xi_{A'A''}(t') \boldsymbol{\eta}_{A''}(t'), \quad (2.64)$$

$$G(t, t') \equiv \text{T-exp} \left(\int_{t'}^t dt'' (\xi \bullet V)(t'') \right), \quad (2.65)$$

where we assume summation over repeated indices, bullet stands for the matrix multiplication and T-exp is the Dyson time-ordered exponential. Using (2.64) and the noise second momentum (2.54) one can easily find the correlation,

$$\begin{aligned} \frac{1}{3} \langle \mathbf{X}_A(t) \mathbf{X}_B(t) \rangle_0 &= G_{AA'}(t, 0) G_{BA''}(t, 0) \frac{1}{3} \langle \mathbf{X}_{A'}(0) \mathbf{X}_{A''}(0) \rangle_0 \\ &+ 2k_B T \int_0^t dt' G_{AA'}(t, t') G_{BA''}(t, t') \xi_{A'A''}(t'). \end{aligned} \quad (2.66)$$

Now, using (2.64) let us calculate the generating function (2.63),

$$\begin{aligned} C_{AB}(\mathbf{k}) &= C_{AB}^{(0)}(\mathbf{k}) C_{AB}^{(\eta)}(\mathbf{k}), \\ C_{AB}^{(0)}(\mathbf{k}) &= \left\langle \exp \left[i\mathbf{k} (G_{AA'}(t, 0) - G_{BA'}(t, 0)) \mathbf{X}_{A'}(0) \right] \right\rangle_{\mathbf{X}_0}, \\ C_{AB}^{(\eta)}(\mathbf{k}) &= \left\langle \exp \left[i\mathbf{k} \int_0^t dt' (G_{AA'}(t, t') - G_{BA'}(t, t')) \xi_{A'A''}(t') \boldsymbol{\eta}_{A''}(t') \right] \right\rangle_{\boldsymbol{\eta}(t)}. \end{aligned} \quad (2.67)$$

Calculating simple Gaussian integrals over $\mathbf{X}(0)$ and $\boldsymbol{\eta}(t)$ we derive,

$$C_{AB}^{(0)}(\mathbf{k}) = \exp \left[- \frac{\mathbf{k}^2}{2} (G_{AA'}(t, 0) - G_{BA'}(t, 0)) (G_{AA''}(t, 0) - G_{BA''}(t, 0)) \frac{1}{3} \langle \mathbf{X}_{A'}(0) \mathbf{X}_{A''}(0) \rangle_0 \right], \quad (2.68)$$

$$C_{AB}^{(\eta)}(\mathbf{k}) = \exp \left[- \frac{\mathbf{k}^2}{2} \int_0^t dt' (G_{AA'}(t, t') - G_{BA'}(t, t')) 2k_B T \xi_{A'A''}(t') (G_{AA''}(t, t') - G_{BA''}(t, t')) \right]. \quad (2.69)$$

From Eq. (2.66) one easily finds the explicit expression for $D_{AA'}(t)$ and thus the previous formula becomes simply,

$$C_{AB}(\mathbf{k}, t) = \exp \left(- \frac{\mathbf{k}^2}{2} D_{AB}(t) \right). \quad (2.70)$$

One can calculate the average of the Oseen tensor straightforwardly,

$$\langle H_{AA'}^{\alpha\alpha'} \rangle = \delta^{\alpha\alpha'} \xi_{AA'}, \quad \xi_{AA'}(t) = \frac{\delta_{AA'}}{\zeta_b} + \frac{1 - \delta_{AA'}}{3(2\pi^3)^{1/2} \eta_s D_{AA'}^{1/2}(t)}. \quad (2.71)$$

We do not write down the explicit form for all effective potentials, $V_{AA'}$, at this point, presenting some of them later in the section 2.5. A useful observation

can be made, however, about their connection to the *instantaneous* mean energy, $\mathcal{E}(t) = \mathcal{E}[D_{AA'}(t)]$, which we define as having the same functional form in terms of the mean squared distances as at equilibrium in Eq. (2.43), but now also functions of time. Namely, the effective potentials can be expressed as the partial derivatives of the instantaneous mean energy by the mean squared distances,

$$V_{AA'}(t) = -\frac{2}{3} \frac{\partial \mathcal{E}(t)}{\partial D_{AA'}(t)}. \quad (2.72)$$

We note that the equations for the pair correlations (2.59) contain some redundancies and also are intermixed with the diffusive degree of freedom describing the translation of the centre-of-mass of the whole system. The latter can be easily decoupled from the rest of the degrees of freedom by rewriting the GSC equations in terms of the mean squared distances between monomers (2.27), which are simple linear combinations of pair correlations (2.35). Thus, we obtain the following kinetic equations for the mean squared monomer distances,

$$\frac{d}{dt} D_{AA'} = 2 \sum_B (\xi_{AB}(t) - \xi_{A'B}(t)) (\Gamma_{AB}(t) - \Gamma_{A'B}(t)), \quad (2.73)$$

$$\Gamma_{AA'}(t) = k_B T \delta_{AA'} + \sum_B D_{AB}(t) V_{A'B}(t). \quad (2.74)$$

The GSC equations (2.73, 2.74) together with the effective potentials given by Eq. (2.72) and explicit expression for the matrix ξ in Eq. (2.71) express in a closed form the evolution in time of the mean squared distances between monomers.

At the end of this section we would like to establish connection of the GSC kinetic equations with the variational free energy. Namely, by direct tedious differentiation it can be shown that the matrix Γ appearing in the right-hand-side of the kinetic equations (2.73) can be expressed via the derivatives of the instantaneous free energy $\mathcal{A}(t)$,

$$\Gamma_{AA'}(t) = -\frac{2}{3} \sum_B D_{AB}(t) \frac{\partial \mathcal{A}[D(t)]}{\partial D_{A'B}(t)}. \quad (2.75)$$

Thus, GSC kinetic equations describing time evolution of the statistical ensemble

may be viewed as a motion in the phase space of the mean squared distances $D_{AA'}(t)$ against the gradient of the instantaneous free energy.

2.4 The Self–Interaction Energy Term

The appearance of the last term in Eq. (2.43), \mathcal{E}_{si} , is somewhat more nontrivial. In fact, in the EFEF of the model (2.14) we have discarded terms with two or more coinciding indices in the three– and higher body contributions. These terms come formally from the virial–type expansion, but each of them gives a singular contribution to the mean energy (2.43). It turns out, however, that on suppressing these terms there appear additional pathological solutions of the GSC equations with singular free energy if at least one element of the two–body interaction matrix, $u_{AA'}^{(2)}$, becomes negative. This is easy to see. Indeed, consider volume interactions of just three monomers under condition that the mean squared distances from monomers '0' and '1' to '2' are equal to each other, $D_{0,2} = D_{1,2} = D$. These interactions produce the mean energy contribution,

$$\mathcal{E}_3 = \frac{u_{0,2}^{(2)} + u_{1,2}^{(2)}}{(2\pi D)^{3/2}} + \frac{1}{(2\pi D_{0,1})^{3/2}} \left(u_{0,1}^{(2)} + \frac{6u^{(3)}(2\pi)^{-3/2}}{(D - D_{0,1}/4)^{3/2}} \right) \quad (2.76)$$

In the case when $u_{0,1}^{(2)} < 0$ and monomer '2' is placed away from monomers '0' and '1', $D > D_{0,1}/4 + (6u^{(3)}(2\pi)^{-3/2}/|u_{0,1}^{(2)}|)^{2/3}$, obviously in the limit $D_{0,1} \rightarrow 0$ the energy possesses a singular minimum, $\mathcal{E}_3 \rightarrow -\infty$. As for the free energy, the logarithmic divergence of the entropy could not change the situation, thus $\mathcal{A} \rightarrow -\infty$ as well. One can show that the inclusion of more monomers in the chain or of higher than three–body interactions does not improve the situation, but produces more and more of such pathological solutions. Note, however, that the additional symmetry properties of monomer–monomer distances, which come from the symmetry of the EFEF (2.14) prevent the problem from happening. For example, in the case of the ring homopolymer, due to the inverse symmetry [128],

we have assumed that for any indices i, i' the following mean squared distances are equal, $D_{i,i'} = D_{i,2i-i'} = D_{2i'-i,i'}$. This provides sufficient repulsion coming from three-body term to preclude pathological solutions.

Thus, in a more general case, where no symmetry properties could be assumed for an arbitrary sequence, the standard procedure of suppressing terms with coinciding indices is not satisfactory. Fortunately, it could be remedied by using another prescription — replacing the terms with coinciding indices by the so-called self-interaction terms. Namely, we propose the following prescription for three-body interaction sufficient for the current purpose,

$$\mathcal{E}^{(3')} = c_3 u^{(3)} \sum_{A \neq A'} \left\langle \delta(\mathbf{X}_A - \mathbf{X}_{A'}) \right\rangle^2 = c_3 \hat{u}^{(3)} \sum_{A \neq A'} D_{AA'}^{-3}, \quad (2.77)$$

where $c_3 = 3$ is a combinatorial factor related to the three possible ways of having coinciding pairs of indices in a triple summation. Obviously, the higher negative power of $D_{AA'}$ in (2.77) compared to the two-body term in (2.43) prevents one monomer from falling onto another.

2.5 Techniques for Numerical Solution of the GSC Equations

In this section we discuss various details concerning techniques for numerical integration of the Gaussian self-consistent kinetic equations (2.73) representing a set of the first-order nonlinear differential equations in time.

In what follows we fix the units of temperature, size and time by choosing $k_B T = 1$, $l = 1$ and $\zeta_b = 1$. We account for the volume interactions up to three-body terms, i.e. we assume that $u_{\{A\}}^{(J)} = 0$ for $J > 3$. Taking into consideration the self-interaction term (2.77) it is sufficient for stability if the third virial coefficient is positive. In most cases below we keep the third virial coefficient equal to, $u^{(3)} = 10$.

First, we consider techniques of calculating the right-hand-side of the GSC equations in Eq. (2.73). Since the entropic contribution in combinations $\Gamma_{AA'}$ can be calculated exactly, the representation of $\Gamma_{AA'}$ in Eq. (2.74) via the effective potentials $V_{AA'}$ given by Eq. (2.72) is the most convenient for numerical integration. Also, since the effective potentials are partial derivatives of the mean energy over the mean squared distances, one can reduce the order of summation in the virial contribution, which is, obviously, the most numerically expensive part of the procedure. Let us write down the explicit formulae for the two- and three-body effective potentials, as well as the three-body self-interaction term,

$$V_{AA'}^{(2)} = (1 - \delta_{AA'}) \frac{\hat{u}_{AA'}^{(2)}}{D_{AA'}^{5/2}} - \delta_{AA'} \sum_{\substack{B \\ B \neq A}} \frac{\hat{u}_{AB}^{(2)}}{D_{AB}^{5/2}}, \quad (2.78)$$

$$\begin{aligned} V_{AA'}^{(3)} &= (1 - \delta_{AA'}) 3\hat{u}^{(3)} \sum_{\substack{B \\ B \neq A, A'}} \frac{D_{AB, A'B}}{(\det \Delta^{(2)})_{AA'B}^{5/2}} - \\ &\quad - \delta_{AA'} \frac{3}{2} \hat{u}^{(3)} \sum_{\substack{BB' \\ B \neq B' \neq A}} \frac{D_{BB'}}{(\det \Delta^{(2)})_{ABB'}^{5/2}}, \end{aligned} \quad (2.79)$$

$$V_{AA'}^{(3')} = (1 - \delta_{AA'}) 6\hat{u}^{(3)} D_{AA'}^{-4} - \delta_{AA'} 6\hat{u}^{(3)} \sum_{\substack{B \\ B \neq A}} D_{AB}^{-4}. \quad (2.80)$$

These are obtained by differentiation of the appropriate terms in the mean energy, \mathcal{E} , over $D_{AA'}$. Note here that elements of each of the effective potential matrices are not independent, since $D_{AA} = 0$. It can be easily shown that sum of every column or row in the matrix vanishes, i.e. $\sum_A V_{AA'} = 0$. Thus, the diagonal elements of the effective potential matrices (2.78-2.80), which contain an additional summation over all monomers, can be expressed through the non diagonal ones,

$$V_{AA}^{(2,3,3')} = - \sum_{\substack{A' \\ A' \neq A}} V_{AA'}^{(2,3,3')}. \quad (2.81)$$

Note also that the number of calls of the power (or square root) function to calculate the inverse characteristic volume of three-particles, $\sigma_{AA'A''}^{(2)} \equiv (\det \Delta^{(2)})_{AA'A''}^{-5/2}$, may be also reduced since the three-body determinant (2.47) is a symmetric

function of $D_{AA'}$, $D_{AA''}$ and $D_{A'A''}$. Thus, one needs to calculate $\sigma_{AA'A''}^{(2)}$ only for $A < A' < A''$, thus reducing the number of calculations at this stage by 6 times. Finally, we remark that the calculation of all elements in $\xi_{AA'}$, $V_{AA'}^{(2)}$, $V_{AA'}^{(3')}$ requires the number of floating point operations of order M^2N^2 , whilst calculation of $V^{(3)}$, $\Gamma_{AA'}$ and the right-hand-side of Eq. (2.73) requires that of order M^3N^3 .

The GSC equations (2.73) are best numerically integrated using the fifth order Runge–Kutta algorithm with adaptive time step [33]. Let us describe briefly this particular integration routine. Thus, consider the system of ordinary differential equations,

$$\frac{d}{dt} \vec{D} = \vec{F}[\vec{D}], \quad (2.82)$$

where the vector stands for both multi-indices A , A' , $A < A'$ and \vec{F} is the right-hand side of Eq. (2.73). Generally, the folding kinetics of copolymer sequences proceeds in a rather unpredictable manner with slow evolution alternating with periods, when some or even all of the mean squared distances change rapidly during a short period of time. The optimal timestep needed for the system to be stable may vary as much as 100 times during kinetic process. So in this case it is almost impossible to achieve a desired accuracy if solving (2.82) by an integration scheme with a fixed timestep, such as the Euler or Runge–Kutta low-order methods. The purpose of the adaptive timestep control is to achieve some pre-determined precision in solution by adjusting the size of timestep depending on how quickly the solution changes at the current moment in time. The step size alternative algorithm is based on the *embedded* Runge–Kutta formulae, which allow one to obtain two estimates, corresponding to different orders of integration precision, in the same way. Comparison of these two estimates gives the magnitude of deviation of the estimate from the exact solution. Having the value of the desired precision and knowing the precision order of the integration scheme, one can predict the optimal time step for the next iteration in time.

The general form of the fifth–order Runge–Kutta formula is,

$$\vec{D}^{(5)}(t_{n+1}) = \vec{D}(t_n) + \sum_{i=1}^6 c_i \vec{K}_i, \quad (2.83)$$

$$\vec{K}_i = \Delta t \vec{F} \left[\vec{D}(t_n) + \sum_{j=1}^{i-1} b_{ij} \vec{K}_j \right], \quad i = 1, \dots, 6. \quad (2.84)$$

The embedded fourth–order Runge–Kutta formula is,

$$\vec{D}^{(4)}(t_{n+1}) = \vec{D}(t_n) + \sum_{i=1}^6 c_i^* \vec{K}_i. \quad (2.85)$$

The coefficients b_{ij} , c_i and c_i^* called the Cash–Karp parameters are given in Ref. [33] (p. 717).

Since all off–diagonal elements of the mean squared distances matrix, $D_{AA'}$, are essentially nonnegative the relative deviation of the estimate from the exact solution is a natural way of introducing the predetermined precision parameter, ϵ_0 . The desired accuracy ϵ_0 is compared to the maximal relative difference between the fourth– and fifth–order estimates,

$$\epsilon_1 = \max_{AA'} \frac{|D_{AA'}^{(5)}(t_{n+1}) - D_{AA'}^{(4)}(t_{n+1})|}{D_{AA'}(t_n)}, \quad (2.86)$$

where $t_{n+1} = t_n + \Delta_1 t$. The desired timestep size, $\Delta_0 t$, which should produce an error about or smaller than ϵ_0 during the next iteration in time, can be estimated as,

$$\Delta_0 t = C_{sav} \Delta_1 t \left(\frac{\epsilon_0}{\epsilon_1} \right)^{1/5}, \quad (2.87)$$

where C_{sav} is some constant about or smaller than unity, e.g. $C_{sav} = 0.85$, and the exponent $1/5$ is the inverse power of the Runge–Kutta algorithm precision. If the algorithm failed to produce the desired precision on a given step, i.e. $\epsilon_0 < \epsilon_1$, the time step size should be rescaled according to Eq. (2.87) without doing an actual increment of $D_{AA'}$ given by Eq. (2.83).

In studying the equilibrium we consider only stationary points of Eqs. (2.73), i.e. the limit $t \rightarrow \infty$. If for some set of interaction parameters, $u_{AA'}^{(2)}$ one obtains

several stationary states, the values of the variational free energy, \mathcal{A} should be compared for them. The deepest minimum of the free energy corresponds to the thermodynamically stable state and the rest of the solutions to metastable states. For systems with a rugged free energy landscape such a procedure for finding minima turns out very efficient.

2.6 GSC Equations for Homopolymer Solutions

In the Gaussian self-consistent kinetic equations (2.73) we consider all mean squared distances between monomers independent. However, in some cases there are additional symmetries for some of the elements of $D_{AA'}$ matrix coming from the symmetries of the effective free energy functional in Eq. (2.14). Thus, one can use these symmetries explicitly and rewrite the GSC equations in a more convenient form. In this section we present such a simplified formalism for solution of identical ring homopolymers.

There are two different symmetry properties present in the system of M identical homopolymer rings of the degree of polymerisation N . First, the mean squared distance between monomers from different chains does not depend neither on their chain indices, i_1, i_2 nor on the chains numbers, m_1, m_2 . Second, the intra-chain mean squared distances depend only on the relative distance between monomers along the chain, $|i_1 - i_2|$, and do not depend on the chain number, m .

For ring polymers we have also the cyclicity condition $\mathbf{X}_{i+N}^m = \mathbf{X}_i^m$ for any i (counting monomers in a chain) and m (counting chains). In addition, we can assume the formal cyclicity for the radii of the centre of mass, $\mathbf{Y}^{m+M} = \mathbf{Y}^m$ for any m , which in fact is not an additional restriction since different chains are unconnected. This allows us to introduce the Fourier transforms as follows,

$$\bar{\mathbf{X}}_i^m \equiv \mathbf{X}_i^m - \mathbf{Y}^m = \sum_{q=0}^{N-1} f_i^{(-q)} \mathbf{x}_q^m, \quad \mathbf{x}_q^m = \frac{1}{N} \sum_{i=0}^{N-1} f_i^{(q)} \bar{\mathbf{X}}_i^m, \quad (2.88)$$

$$\bar{\mathbf{Y}}^m \equiv \mathbf{Y}^m - \mathbf{Y} = \sum_{Q=0}^{M-1} g_m^{(-Q)} \mathbf{y}^Q, \quad \mathbf{y}^Q = \frac{1}{M} \sum_{m=0}^{M-1} g_m^{(Q)} \bar{\mathbf{Y}}^m, \quad (2.89)$$

$$f_j^{(q)} \equiv \exp\left(\frac{2\pi i q j}{N}\right), \quad g_m^{(Q)} \equiv \exp\left(\frac{2\pi i Q m}{M}\right). \quad (2.90)$$

Henceforth we shall call \mathbf{x}_q^m ($q \neq 0$) the intra-molecular and \mathbf{y}^Q ($Q \neq 0$) the inter-molecular degrees of freedom respectively, with the zero modes being identically zero due to our earlier convention.

Now we have to take into account the symmetries in the system. There are only the following non-vanishing equal-time correlations,

$$\frac{1}{3} \langle \mathbf{x}_q^m(t) \mathbf{x}_{q'}^{m'}(t) \rangle \equiv \delta^{m,m'} \delta_{-q,q'} \mathcal{F}_q(t), \quad (2.91)$$

$$\frac{1}{3} \langle \mathbf{y}^Q(t) \mathbf{y}^{Q'}(t) \rangle \equiv \delta^{-Q,Q'} \frac{\Gamma^2(t)}{M}. \quad (2.92)$$

The mean squared distances between i_1 -th monomer in m_1 -th chain and i_2 -th monomer from m_2 -th chain, $D_{i_1, i_2}^{m_1, m_2}$, may be reduced via,

$$D_{i_1, i_2}^{m_1, m_2} = \delta^{m_1, m_2} D_{i_1 - i_2} + 2(1 - \delta^{m_1, m_2}) (R_g^2 + \Gamma^2), \quad (2.93)$$

where we have introduced the notations for the correlations of monomer positions, D_k , and the mean squared radius of gyration, R_g^2 , of a chain,

$$D_k \equiv D_{0k}^{mm} = \sum_{q=1}^{N-1} d_k^{(q)} \mathcal{F}_q, \quad R_g^2 \equiv \sum_{q=1}^{N-1} \mathcal{F}_q, \quad d_k^{(q)} \equiv 2 \left(1 - \cos \frac{2\pi q k}{N}\right). \quad (2.94)$$

Finally, the total radius of gyration is simply,

$$R_{tot}^2 \equiv \frac{1}{6N^2 M^2} \sum_{(i,m),(i',m')} \langle (\mathbf{X}_i^m - \mathbf{X}_{i'}^{m'})^2 \rangle = R_g^2 + \frac{M-1}{M} \Gamma^2. \quad (2.95)$$

Equation (2.93) expresses simplifications in the intra- and inter-chain mean squared monomer distances due to the EFEF symmetries discussed above.

Performing derivations analogous to those of Sec. 2.2 and keeping the excluded volume terms up to the three-body interactions, we obtain the explicit expression

for the mean energy,

$$\begin{aligned}
\frac{\mathcal{E}}{MN} &= \frac{3k_B T}{2l^2} D_1 + \frac{3k_B T \lambda}{2l^3} (4D_1 - D_2) + \frac{3k_B T}{2L^2} \frac{M-1}{M} \Gamma^2 + \\
&+ \hat{u}^{(2)} \left(\sum_k D_k^{-3/2} + (M-1)N(2R_g^2 + 2\Gamma^2)^{-3/2} \right) + \\
&+ \hat{u}^{(3)} \left(\sum_{k_1 \neq k_2} (D_{k_1} D_{k_2} - D_{k_1, k_2}^2)^{-3/2} + \right. \\
&+ 3(M-1)N \sum_k \left(D_k(2R_g^2 + 2\Gamma^2) - D_k^2/4 \right)^{-3/2} + \\
&+ (M-1)(M-2)N^2 \left(\frac{3}{4}(2R_g^2 + 2\Gamma^2)^2 \right)^{-3/2} \left. \right) + \\
&+ 3\hat{u}^{(3)} \left(\sum_k D_k^{-3} + (M-1)N(2R_g^2 + 2\Gamma^2)^{-3} \right), \tag{2.96}
\end{aligned}$$

where

$$D_{k_1, k_2} \equiv \frac{1}{2} (D_{k_1} + D_{k_2} - D_{k_1 - k_2}). \tag{2.97}$$

Similar calculations for the entropy give,

$$\mathcal{S} = \frac{3}{2} k_B M \left(\sum_{q=1}^{N-1} \log \mathcal{F}_q + \frac{M-1}{M} \log \Gamma^2 \right). \tag{2.98}$$

Thus, we have derived the variational free energy, $\mathcal{A} = \mathcal{E} - T\mathcal{S}$, as a function of the variables \mathcal{F}_q and Γ . Its minimisation with respect to them gives the equilibrium free energy, \mathcal{A}_{eq} . Note here also that only half of the Fourier amplitudes \mathcal{F}_q are independent since $\mathcal{F}_q = \mathcal{F}_{N-q}$ due to the inverse symmetry, i.e. one can count monomers in any direction along the chain.

It is also straightforward to extract the equilibrium thermodynamic functions such as the chemical potential and the osmotic pressure,

$$\mu_{eq} = \left(\frac{\partial \mathcal{A}}{\partial M} \right)_{eq}, \quad \Pi_{eq} = - \left(\frac{\partial \mathcal{A}}{\partial L^3} \right)_{eq} = k_B T \frac{M}{L^5} \Gamma^2. \tag{2.99}$$

Proceeding from the Langevin equation (2.16) in the way described in detail in Sec. 2.3 it is possible to derive the kinetic equations for our dynamic variables

(2.91, 2.92). These may be written as follows,

$$\frac{\zeta}{2} \frac{d}{dt} \mathcal{F}_q = -\frac{2}{3} \mathcal{F}_q \frac{\partial \mathcal{A}}{\partial \mathcal{F}_q} = k_B T - V_q \mathcal{F}_q, \quad \zeta \frac{d}{dt} \Gamma = -\frac{1}{3} \frac{\partial \mathcal{A}}{\partial \Gamma}, \quad (2.100)$$

$$V_q = -\frac{2}{3} \frac{\partial \mathcal{E}}{\partial \mathcal{F}_q}, \quad \zeta \equiv NM\zeta_b. \quad (2.101)$$

Compared to the general formalism in Eq. (2.73) here we have only $1 + N/2$ independent variational parameters and the number of chains appears only as parameter in the free energy (2.96, 2.98) and in the GSC kinetic equations (2.100). Integration of the kinetic equations can be carried out in the same way as described in Sec. 2.5. Since differentiation over D_k reduces the order of summation in virial expansion terms, the effective potentials V_q should be calculated as,

$$V_q = \sum_k V_k \frac{d D_k}{d \mathcal{F}_q} = \sum_k V_k d_k^{(q)}, \quad V_k = -\frac{2}{3} \frac{\partial \mathcal{E}}{\partial D_k}, \quad (2.102)$$

where coefficients $d_k^{(q)}$ are given in Eq. (2.94). Thus, the CPU time expenses per time step here are proportional to N^2 compared to $M^3 N^3$ in the general formalism.

The generation of random numbers is too important to be left to chance.

A random message from fortune

Chapter 3

Simulation Techniques

In this chapter we present models which are used in Monte Carlo simulations. These include the model of copolymer solutions on a lattice and the model of a single homopolymer ring in continuous space. Also, the Monte Carlo Metropolis algorithm for simulating the equilibrium is overviewed in some detail and its applicability for dynamics and kinetics is discussed. For both considered models we discuss the procedures for calculating the energy change, algorithms for performing various Monte Carlo moves, and prescriptions for ensuring the topological integrity of the polymer chain.

3.1 Lattice Monte Carlo Model of Copolymer Solutions

In this section we describe the model and some technical details of Monte Carlo simulations for equilibrium and kinetics of polymer solutions. In Appendix A we present the C code of the kernel of the program which has been used for Monte Carlo simulation of several (co)polymer chains of identical lengths on a cubic lattice with reflective boundary conditions.

We do not intend to work at the level of detailed atomistic description, but use a coarse-grained notion of monomers instead. Thus, let us consider the canonical ensemble of M identical polymers of N building blocks, monomers, in a finite volume V at temperature T . In a lattice model we allow only some discrete values for spatial positions of monomer and solvent molecules. Thus, here the volume is represented by a cubic lattice of linear size Lb , where L is number of possible values for each spatial component and b is lattice spacing, which we set $b = 1$ to fix the unit of size in the system. As in the analytical treatment we only consider a model of incompressible solution. In other words, each lattice site is occupied by either a monomer or a solvent molecule. The polymer concentration is thus given by,

$$c = NM/L^3. \quad (3.1)$$

The polymer dynamics can be implemented by permutations of monomer and solvent units on the lattice. We shall call such a permutation an elementary move.

In a coarse-grained model there are at least two obvious restrictions on how monomers can possibly be arranged in the lattice for each particular conformation. Namely, we must ensure polymer connectivity, and excluded volume. We restrict our model by making the following particular choices of elementary moves. The maximal distance between the nearest neighbours along the chain (NNC) is equal to $r_{max} = \sqrt{3}$. Thus, for every bead its NNC are located in the nearest lattice sites along the vertices of the lattice, or on second or third lattice neighbours. This condition provides for connectivity of the chain. Furthermore, excluded volume is incorporated by ensuring that only NNC are permitted in the nearest neighbour lattice sites, i. e. the minimum distance between beads is $r_{min} = 1$ for NNC beads (NNC cannot overlap), and $r_{min} = \sqrt{2}$ otherwise. We shall comment later in this section about such a choice for connectivity and

excluded volume parameters, r_{max} and r_{min} .

Next, in our scheme there are “weak” pair-wise interactions between lattice sites depending on the separation between the site and their contents. By “weak” we mean interactions of order $k_B T$, which do not result in formation of strong chemical bonds. These represent e.g. the hydrophobic and van der Waals interactions, but after an appropriate coarse-graining procedure. Such interactions can be described by the following Hamiltonian,

$$H = \frac{1}{2} \sum_{i \neq j} w(r_{ij}) \mathcal{I}_{s_i s_j}, \quad (3.2)$$

where i, j enumerate lattice sites; s_i labels the contents of site i , $\mathcal{I}_{s_i s_j}$ is some symmetric matrix of monomer and solvent interactions, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the separation between the two sites and $w(r_{ij})$ is some function giving the form of the potential interaction.

Heterogeneous chains with different types of monomers can be straightforwardly incorporated into the model Hamiltonian (3.2). Let the polymers consist of K various types of monomers, which we call ‘a’, ‘b’ etc, with the index ‘s’ standing for the solvent. This system can be described by the Hamiltonian (3.2) with the following matrix of size $K + 1$ of monomer and solvent interactions,

$$\mathcal{I} = \begin{pmatrix} \mathcal{I}_{ss} & \mathcal{I}_{sa} & \mathcal{I}_{sb} & \dots \\ \mathcal{I}_{as} & \mathcal{I}_{aa} & \mathcal{I}_{ab} & \dots \\ \mathcal{I}_{bs} & \mathcal{I}_{ba} & \mathcal{I}_{bb} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (3.3)$$

Thus, in our model building blocks of copolymers have the same internal geometry, i.e. the connectivity and excluded volume restrictions are the same for all types of monomers, but differ only in their “weak” interaction properties. In principle, the structure of the Hamiltonian (3.2) implies that each lattice site interacts with all others. Thus, calculation of the energy change due to an elementary move in the Metropolis algorithm would involve about L^3 operations,

which is rather unacceptable for lattice sizes of interest in the range $L = 100-250$. The situation can be improved dramatically by introduction of a cut-off at the interaction range, i. e. $w(r) = 0$, for $r > R_{max}$. This procedure is well justified for the short ranged Lennard-Jones potential, where one can introduce a spatial cut-off at the distance of a few van der Waals radii. In our model we chose $w(1) = w(\sqrt{2}) = 1$, $w(\sqrt{3}) = 0.7$, $w(2) = 1/2$ and $w(r) = 0$ for $r > 2$. Thus, the range of interaction includes 6 nearest, 12 second lattice, 8 third lattice and 6 second-nearest neighbours, so that in total each lattice site interacts with 32 of its nearest neighbours.

* * *

The system is simulated based on the Metropolis algorithm [7,8] for calculation of the statistical averages in the system at temperature T . Let us briefly overview the algorithm and its meaning here. In principle, having the expression for the energy of the system in Eq. (3.2) one can calculate the average of any observable, A , in the canonical ensemble precisely according to the formula,

$$\langle A \rangle = \frac{1}{Z} \sum_i A_i \exp\left(-\frac{H_i}{k_B T}\right), \quad (3.4)$$

$$Z = \sum_i \exp\left(-\frac{H_i}{k_B T}\right), \quad (3.5)$$

where index i passes through all microstates consistent with the connectivity and excluded volume requirements. Here A_i and H_i are the values of the observable and the energy (3.2) in the i -th microstate, and Z is called the canonical partition function. This, however, is not feasible since the total number of allowed microstates in the system grows exponentially quickly with increasing the system size. For example, the number of microstates of a single chain of size N is of order $\mathcal{N}_{microstates} \approx 20^N$. For disconnected units this number grows even faster. So, the number of microstates for the lattice gas of M particles on a cubic lattice of linear size L and $r_{min} = 1$ is, $\mathcal{N}_{microstates} = L^3(L^3 - 1) \dots (L^3 - M + 1)$, as gas molecules cannot occupy the same lattice site. Therefore we have the base

close to L^3 , if at least $L^3 \gg M$, instead of 20 in the case of a single chain. Thus, based on Eqs. (3.4,3.5) the existing computational powers allows one to consider systems of rather miserable sizes and no revolutionary advance in the microchip technology would ever permit to study a system of any reasonable size in this way.

A possible way-out of the deadlock is provided by so-called Monte Carlo methods. The main idea of the simplest of them, called the “simple sampling” Monte Carlo method [8] is the following. One can still use the exact formula of Eqs. (3.4,3.5), but instead of summing over *all* microstates one restricts the sums to only *some* randomly picked microstates, $\{\mathbf{X}_A\}$, so that they sample the phase space of the system uniformly. Theoretically, the result calculated over such a random subset, $\langle A \rangle_{random}$, even though different each time, should converge to the exact value, $\langle A \rangle$ with increasing the number of statistical measurements, Q , as

$$|\langle A \rangle_{random} - \langle A \rangle| \sim Q^{-1/2}. \quad (3.6)$$

Such a naive Monte Carlo scheme works well at high temperatures as all microstates give approximately comparable contribution to the partition function (3.5). However, at low temperatures, and especially if the system undergoes a sort of condensation phenomenon, only a negligibly small fraction of all microstates, which still may be a large number in the absolute value, contribute to the partition function. So there it is essentially impossible to obtain any reasonable estimate for the thermodynamic averages by simple sampling.

Such a Monte Carlo scheme can be considerably advanced by applying the so-called *importance sampling* method. Its idea is to generate a random subset of microstates, which are in some sense representative of the system behaviour and thus important. Basically, we would like to generate a sequence of microstates which would converge to the Boltzmann distribution in the phase space as the

number of microstates increases.

Though there are several variations of Monte Carlo importance sampling techniques we shall use the most popular scheme known as the Metropolis algorithm. Its main idea is to construct a Markov process in the phase space,

$$\mathbf{X}_A(1) \rightarrow \mathbf{X}_A(2) \rightarrow \dots \rightarrow \mathbf{X}_A(i-1) \rightarrow \mathbf{X}_A(i) \rightarrow \dots, \quad (3.7)$$

where each state $\mathbf{X}_A(i)$ follows from the previous one $\mathbf{X}_A(i-1)$ according to the transition probability, $\Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A)$.

Let $P_0(\mathbf{X}_A)$ denotes the initial probability with any distribution, e.g. uniform. At each time step we update these probabilities by a certain rule. The updated probabilities set may be obtained as, $P_1(\mathbf{X}_A) = \sum_{\mathbf{X}'_A} \Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A) P_0(\mathbf{X}'_A)$, where $\Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A)$ is the corresponding transition probability. We then repeat the above procedure many times using the results from the previous step as input.

Let us introduce the *balance condition*, which expresses the stationarity of the equilibrium distribution, $P_{eq}(\mathbf{X}_A) = Z^{-1} \exp(-\beta H(\mathbf{X}_A))$, with respect to the action of the Markov process:

$$P_{eq}(\mathbf{X}_A) = \sum_{\mathbf{X}'_A} \Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A) P_{eq}(\mathbf{X}'_A). \quad (3.8)$$

Next, the *ergodicity condition* means that the phase space does not factorise into disjoint parts, i.e. for any \mathbf{X}'_A and \mathbf{X}_A there exists a finite sequence of $\mathbf{X}_A(i)$ $i = 1, \dots, m$ such that $\Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A) \neq 0$ and $\mathbf{X}_A(1) = \mathbf{X}'_A$ and $\mathbf{X}_A(m+1) = \mathbf{X}_A$. In other words, the system can go from any state to any other state by a finite number of steps.

These two conditions are sufficient for convergence of the ensemble produced by a Markov process to the equilibrium distribution [112,164]. To show that we need to introduce the notion of the “distance” between two ensembles E and E' via the norm [165],

$$\|E - E'\| = \sum_{\mathbf{X}_A} |P(\mathbf{X}_A) - P'(\mathbf{X}_A)|. \quad (3.9)$$

Now, if the ensemble E' is obtained from E by action of one step in the Markov chain we have,

$$P'(\mathbf{X}_A) = \sum_{\mathbf{X}'_A} \Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A) P(\mathbf{X}'_A). \quad (3.10)$$

This allows us to obtain the following estimate,

$$\begin{aligned} \|E' - E_{eq}\| &= \sum_{\mathbf{X}_A} \left| \sum_{\mathbf{X}'_A} \Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A) (P(\mathbf{X}'_A) - P_{eq}(\mathbf{X}'_A)) \right| \\ &\leq \sum_{\mathbf{X}_A, \mathbf{X}'_A} \Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A) |P(\mathbf{X}'_A) - P_{eq}(\mathbf{X}'_A)| = \|E - E_{eq}\|, \end{aligned} \quad (3.11)$$

where we have used the non-negativeness of the transition probability and the normalisation condition, $\sum_{\mathbf{X}'_A} \Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A) = 1$. Thus, application of such a step moves the ensemble closer to the equilibrium. Finally, note that the inequality in Eq. (3.11) becomes strict thanks to the ergodicity, and so the equality is only possible when the ensemble has reached the equilibrium.

Instead of the balance condition (3.8) one can apply a more restrictive *detailed balance condition*, which relates the transition probabilities of the forward and backward transitions,

$$\exp\left(-\frac{H[\mathbf{X}_A]}{k_B T}\right) \Pi(\mathbf{X}_A \rightarrow \mathbf{X}'_A) = \exp\left(-\frac{H[\mathbf{X}'_A]}{k_B T}\right) \Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A), \quad (3.12)$$

which obviously produces the latter balance condition by summing over \mathbf{X}'_A .

In the Metropolis algorithm one step of the Markov chain is generated in two steps:

1. Given $\mathbf{X}_A(n)$ one generates a trial conformation $\mathbf{X}_A(T)$ by a random algorithm with a symmetric transition probability,

$$\Pi_S(\mathbf{X}_A(n) \rightarrow \mathbf{X}_A(T)) = \Pi_S(\mathbf{X}_A(T) \rightarrow \mathbf{X}_A(n)). \quad (3.13)$$

This process looks like a simple random walk in the phase space and does not depend on the particular Hamiltonian H of the system. For instance, for the Ising model it would mean to pick up a randomly chosen spin.

2. Perform a transition with the conditional probability $\Pi_M(\mathbf{X}_A(n) \rightarrow \mathbf{X}_A(T))$, so that the total transitional probability is,

$$\Pi(\mathbf{X}_A(n) \rightarrow \mathbf{X}_A(T)) = \Pi_S(\mathbf{X}_A(n) \rightarrow \mathbf{X}_A(T)) \Pi_M(\mathbf{X}_A(n) \rightarrow \mathbf{X}_A(T)). \quad (3.14)$$

A simple choice of this transition probability was proposed by Metropolis et al [7],

$$\Pi_M(\mathbf{X}_A \rightarrow \mathbf{X}'_A) = \begin{cases} \exp\left(-\frac{\Delta H}{k_B T}\right), & \text{for } \Delta H \equiv H[\mathbf{X}'_A] - H[\mathbf{X}_A] > 0, \\ 1, & \text{for } \Delta H \leq 0. \end{cases} \quad (3.15)$$

Due to the explicit definition of Π_M and symmetricity of the matrix Π_S the total transition probability Π still satisfies the detailed balance condition.

The related Metropolis algorithm in our particular case follows the procedure:

- 0) establish an initial conformation;
- 1) select a monomer (or a group of monomers) at random;
- 2) select the type and direction of the trial move at random;
- 3) calculate the energy change due to that move, $\Delta E = E_{fin} - E_{ini}$;
- 4) compare $\eta = \exp(-\Delta E/k_B T)$ with a newly generated random number, r , uniformly distribution between 0 and 1;
- 5) if $\eta \geq r$ accept the move, i.e. change to the new conformation, otherwise reject it and keep the old conformation;
- 6) repeat steps 1) - 5) over and over again.

In step 4) it is worthwhile to check first, if $\Delta E \leq 0$, i.e. the move has caused a decrease of the energy; then one can accept such a move without the unnecessary check and, thus, save a few microseconds of valuable CPU time per elementary move. So basically, the Metropolis algorithm selects microstates by suppressing moves which are accompanied by an unreasonably large increase of the energy of the system, still allowing them to happen rarely, which implicitly accounts for the entropy. An infinite set of microstates obtained by the Metropolis algorithm contains all microstates with the weight equal to their Boltzmann factor.

The practical procedure for obtaining equilibrium values of thermodynamic quantities in the Metropolis algorithm can be divided in two parts:

i) Equilibration of the system.

During equilibration steps 1) to 5) are repeated enough times for the system to reach the region of the phase space preferable for given temperature and set of interaction parameters. It is worthwhile here to decrease the temperature by small steps, starting from a sufficiently high value. In fact, sufficiently far from a transition point, i.e. where typical system conformations change slightly with the temperature, the equilibration procedure is not really necessary. Talking about the temperature, one should bear in mind that the set of the phenomenological interaction constants \mathcal{I} in Eq. (3.3) is not temperature independent. In what follows we shall set $k_B T = 1$, which fixes the unit of measurements of energy in the system. So instead of changing the temperature itself we shall change the values of the interaction constants between the solvent and monomers, or a their linear combination of those constants.

ii) Collecting statistics.

After the system is believed to be at equilibrium one can proceed with the procedure of collecting statistics. Each particular system conformation after certain number of elementary moves is recorded and subsequently analysed. Importantly, since the Metropolis algorithm already produces microstates with the Boltzmann distribution estimates of thermodynamic quantities here are made by using the simple arithmetic average instead of Eq. (3.4),

$$\langle A \rangle \approx \frac{1}{Q} \sum_i A_i, \quad (3.16)$$

where Q is number of states obtained by the Metropolis algorithm and numerated by the index i . We should also mention here that states in Eq. (3.16), over which statistical averaging is performed, should be uncorrelated as much as possible. It is reasonable to somewhat increase the number of elementary moves made

between subsequent events in Eq. (3.16), though this obviously increases the computational expenses.

Now, let us consider the procedure of performing statistical averaging at nonequilibrium conditions. The dynamic interpretation of the Monte Carlo averaging is obtained by associating a “time” with the label ‘ i ’ of successive conformations, so that A_i in Eq. (3.16) is interpreted as $A(t)$. If we interpret Π in Eq. (3.15) as the transition probability per unit “time” the Markov chain of phase space points (3.7) can be considered as a numerical realisation of the following *Master equation*,

$$\frac{dP(\mathbf{X}_A, t)}{dt} = \sum_{\mathbf{X}'_A} \left(P(\mathbf{X}'_A) \Pi(\mathbf{X}'_A \rightarrow \mathbf{X}_A) - P(\mathbf{X}_A) \Pi(\mathbf{X}_A \rightarrow \mathbf{X}'_A) \right). \quad (3.17)$$

This has the following physical meaning: the probability of a microstate \mathbf{X}_A decreases by all processes that lead away from this state to other state \mathbf{X}'_A . Such a “loss” of probability is counteracted by the inverse process from \mathbf{X}'_A to \mathbf{X}_A leading to a “gain” of probability. At equilibrium, due to the detailed balance (3.12), both the gain and loss processes compensate each other, i.e. $dP(\mathbf{X}_A, t)/dt = 0$. Thus, the Monte Carlo averaging can be viewed as a “time averaging” along a stochastic trajectory in the phase space. We shall call the sequence of NM attempts of elementary moves the Monte Carlo sweep (MCS). The latter is assumed to play the role of a discrete time unit. During 1 MCS each monomer in average is moved once, and 1 MCS is thus proportional to the real time with a coefficient depending on the effective friction of monomers, but independent of the system size.

Let us mention here one particularly interesting nonequilibrium problem. Consider our system at equilibrium under some set of interaction parameters, \mathcal{I} . We record the state of the system, and then make a quench, i.e. instantaneously change values of the interaction parameters. The typical conformations of the system start to change during the Metropolis evolution. After some time the system reaches its final equilibrium state and observables start to fluctuate around

their new equilibrium values. Each particular system state before the quench plus its particular evolution due to random moves towards the new equilibrium represents an ensemble element. In this problem, however, we are interested in the time evolution of the total statistical ensemble, i.e. how the average of each observable of interest changes in time. For this one should perform averaging over the initial conditions and random noise and finally obtain the average of an observable at each moment in time. Practically, this is achieved by performing for a number of times, typically a few hundreds, the following procedure: an equilibration for the initial interaction parameters, an instantaneous quench, and recording of the consequent time evolution.

Finally, note that there is a considerable choice in performing moves to execute the transition $\mathbf{X}_A(i-1) \rightarrow \mathbf{X}_A(i)$ in the Markov chain (3.7). We should emphasise that in the limit of vanishing two-body interactions in Eq. (3.2) moves should sample the phase space of the system, $\{\mathbf{X}_A\}$, essentially uniformly. In particular, starting from any point in the phase space \mathbf{X}_A one should reach another point \mathbf{X}'_A in a finite number of steps to ensure ergodicity.

* * *

Now we turn our attention to the question of calculating the energy change, ΔE , due to an elementary move. Let us assume for the moment that the nearest, second lattice, third lattice and second nearest lattice neighbours have the same weights, i.e. $w(1) = w(\sqrt{2}) = w(\sqrt{3}) = w(2) = 1$. If, moreover, we restrict ourselves to the case of just two types of monomers, 'a' and 'b' the Hamiltonian (3.2 can be rewritten in the equivalent form,

$$H = \mathcal{I}_{aa}C_{aa} + \mathcal{I}_{bb}C_{bb} + \mathcal{I}_{ss}C_{ss} + \mathcal{I}_{ab}C_{ab} + \mathcal{I}_{as}C_{as} + \mathcal{I}_{bs}C_{bs}, \quad (3.18)$$

where C_{xy} is the total number of xy -contacts on the lattice. The xy -contact is defined as a pair formed by lattice sites occupied by 'x' and 'y' species, such that these sites are in the range of interaction. Since the number of both types of

monomers and solvent molecules are fixed during simulation, there are additional constraints on the number of contacts. Indeed, by considering contacts formed by 'a'-monomers, we can write,

$$2C_{aa} + C_{ab} + C_{as} = 32N_a, \quad (3.19)$$

where N_a is the total number of 'a' monomers on the lattice and the factor 32 is the total number of contacts per lattice site. The C_{aa} contacts should be multiplied by 2 since homogeneous interactions appear twice in the sum. Analogously, by considering 'b' and 's' lattice sites we can write,

$$2C_{bb} + C_{ab} + C_{bs} = 32N_b, \quad (3.20)$$

$$2C_{ss} + C_{as} + C_{bs} = 32(L^3 - N_a - N_b). \quad (3.21)$$

Using relations (3.19,3.20,3.21) one can exclude the monomer-solvent contacts from consideration and rewrite Eq. (3.18) in the following form,

$$H = H_0 - k_B T (\chi_{aa} C_{aa} + \chi_{ab} C_{ab} + \chi_{bb} C_{bb}), \quad (3.22)$$

$$H_0 = 16L^3 \mathcal{I}_{ss} + 32N_a (\mathcal{I}_{as} - \mathcal{I}_{ss}) + 32N_b (\mathcal{I}_{bs} - \mathcal{I}_{ss}) \quad (3.23)$$

$$\chi_{aa} = \frac{2\mathcal{I}_{sa} - \mathcal{I}_{aa} - \mathcal{I}_{ss}}{k_B T}, \quad (3.24)$$

$$\chi_{bb} = \frac{2\mathcal{I}_{sb} - \mathcal{I}_{bb} - \mathcal{I}_{ss}}{k_B T}, \quad (3.25)$$

$$\chi_{ab} = \frac{\mathcal{I}_{sa} + \mathcal{I}_{sb} - \mathcal{I}_{ab} - \mathcal{I}_{ss}}{k_B T}. \quad (3.26)$$

The first term in (3.22) is just a constant which does not depend on the system conformation and can be neglected. The combinations of interaction parameters in Eqs. (3.24-3.26) describe the degree of corresponding monomer-monomer attraction and they are the only relevant thermodynamic parameters characterising interactions in the system for a given polymer sequences and lattice size L .

Calculation of the energy change is a rather simple procedure now. While moving a monomer of the type 'x' ('a' or 'b') one needs to calculate the change in the

number of homogeneous, ΔC_{xx} , and heterogeneous, ΔC_{ab} , monomer–monomer contacts associated with the given monomer. Then, the change of energy is given simply by,

$$\Delta E = -\chi_{ab}\Delta C_{ab} - \chi_{xx}\Delta C_{xx}. \quad (3.27)$$

This procedure can be easily generalised for a larger number of monomer types. Here one has $K(K + 1)/2$ independent interaction parameters, akin to those in Eqs. (3.24-3.26), between all possible pairs of monomers. Note, that for the homopolymer the interactions are characterised by only one parameter,

$$\chi = \frac{2\mathcal{I}_{sm} - \mathcal{I}_{mm} - \mathcal{I}_{ss}}{k_B T}, \quad (3.28)$$

called the Flory interaction parameter.

Also the expression for the energy change is easy to generalise for a more complicated form of the weight function, $w(r_{ij})$. One would still have the same number of χ -parameters, but the formula for calculation of the energy change (3.27), becomes a little bit more complicated,

$$\Delta E = -\chi_{ab} \sum_l w(l)\Delta C_{ab}(l) - \chi_{xx} \sum_l w(l)\Delta C_{xx}(l), \quad (3.29)$$

where l passes through various ranges of the interaction (nearest, second lattice, third lattice, second nearest neighbours), and $\Delta C(l)$ is the change of the number of contacts at the l -th interaction range.

Now let us describe the procedure of how to update system conformations. The most important moves, which are applied most frequently in a sequence of Metropolis updates, are the local monomer moves. Here a monomer picked randomly from a randomly picked chain is attempted to randomly change its spatial position, so that the connectivity and excluded volume constraints are preserved. Practically, we first try to find a position satisfying the connectivity restriction

and then, by looking up through the lattice, check the excluded volume condition. The first procedure is slightly different for the end and internal monomers, since they have a different number of NNC. Note also that due to the specific choice of the connectivity condition (the NNC of a monomer should be found in the cube formed by nearest, second and third lattice neighbours) each of spatial projections can be treated independently. Thus, for example, to find the new random X -coordinate of the end monomer one takes the X -coordinate of the penultimate monomer and adds to it $r\{-1, 0, 1\}$, i.e. one of the numbers, -1 , 0 or 1 picked randomly. In the case of an internal monomer one should consider the X -coordinate of its NNCs, which we denote as X_1 and X_2 , $X_1 \leq X_2$. Then there are only three possibilities:

- 1) if $X_2 - X_1 = 2$, then $x = X_1 + 1$;
- 2) if $X_2 - X_1 = 1$, then $x = X_1 + r\{0, 1\}$;
- 3) if $X_2 = X_1$, then $x = X_1 + r\{-1, 0, 1\}$.

The Y - and Z -projections are treated in precisely the same way. Then excluded volume and Metropolis checks are made.

Another important type of elementary move is a snake-like reptational move of a chain along its path. Reptations are collective motions of the whole chain that can be realised in the case of homopolymer by transfer of one of the homopolymer ends to another one. Importantly, a trial move of a monomer here can be done in exactly the same way as the local end-monomer move. Also in the case of the homopolymer one does not need to calculate the change of energy for the whole system. Reptations are made with some predefined probability, p_r . In Sections 4.1 and 4.4 we will discuss how inclusion of reptational moves affects the system conformations. Note that at equilibrium one can mix various types of moves to improve convergence of averages to the equilibrium value, but one should be careful whether the conditions of ergodicity and detailed balance are satisfied. In

kinetics, global moves such as reptations are not permitted. Moreover, reptations are only relevant for the homopolymer since in case of heteropolymers they would violate the local interactions.

Also, in a system of two or more chains at equilibrium it is highly desirable to apply translational moves of the whole chains as well. This can be motivated as follows. After a polymer has collapsed the chain becomes rather immobile if only local moves are involved. This is because monomers forming the core of the globule can hardly move at all, and movements of the globular shell do not provide the required mobility of the object as a whole. Thus, the aggregation or similar phenomena would be oversuspended in a scheme with only local moves. The situation improves dramatically by introducing translational moves representing diffusion of chains.

Importantly, in the multichain case condensates involving several polymers can be formed and they should be treated by the scheme of translational moves in exactly the same manner as single chains themselves. Thus, translational moves are applied to all connected *clusters* of interacting chains with a probability inversely proportional to their size (Stokes law in the absence of the hydrodynamic interaction). Such a move results in shifting the cluster in a random direction along lattice vertices (6 possible directions).

Now let us consider now a situation of clusters collision. Imagine that due to a translation a cluster of chains enters the interaction range of another cluster(s). This results in formation of a larger cluster incorporating the initial ones. In fact, the question about a collision of clusters of chains is rather subtle, since the scheme of moves becomes, in a certain sense, disbalanced. Indeed, if we do allow clusters to collide, we must also implement a type of moves which would allow clusters to split onto two or more parts constituent parts. There are two problems in doing that. First, it is difficult to make such a scheme effective as it needs to

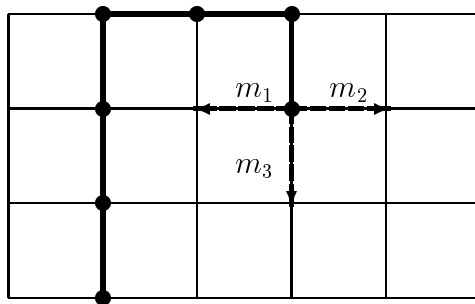
pick the right chain or several chains, which are close to the cluster boundaries and can be separated. Second, in a rather dense cluster there are many contacts between monomers from distinct chains, and thus most of attempts would be rejected by the Metropolis check due to a large unfavourable energy change. In Subsec. 6.1.4 we estimate the Metropolis probability to remove a chain from a large cluster at $\chi = 1/2$ and find it to be roughly, $\eta \approx \exp(-2N)$.

One simple practical suggestion would be to perform only those translational moves which do not result in a collision. We call this the *reduced translational moves scheme*. Such translational moves increase mobility of chains and clusters of chains and help them to find each other on the lattice. The final decision of whether to stay together or break apart is made on the level of local monomer and reptational moves. Another nice simplification due to making only reduced translational moves is that there is no need to perform the Metropolis check for any interaction parameters. Indeed, the reduced moves cannot change the number of monomer contacts. We should emphasise, however, that generalisation of this procedure to kinetics is beset with many subtle problems and will not be considered here.

Next, it is worthwhile to discuss briefly the choice of the boundary conditions. Reflective boundaries mean that we simply do not allow polymers to move beyond the first and last lattice sites in all three directions. In case of the periodic boundary condition each spatial direction is considered to have the ring topology, i.e. next to the last lattice site is also the first. One of the disadvantages in using the reflective boundary conditions is the spatial inhomogeneity. Monomer movements near the lattice edges are somewhat suspended compared to the lattice center. Thus, on decreasing temperature in the system a condensation may start prematurely somewhere near the lattice edges, which is mainly caused by the lack of mobility there. This effect, however, is of concern only for rather

Figure 3.1: An example of a self-avoiding walk on a two-dimensional square lattice.

Only a walk along the lattice axes, which does not cross itself is allowed, so that $r_{max} = 1$ and $r_{min} = 1$.



dense systems. In our case of a dilute solution chains and their clusters are quite mobile due to translational moves and this artificial condensation probably may only take place in the near vicinity of the phase transition curve. Also, with increasing the lattice size the contribution of the boundary becomes negligible compared to the bulk. On the other hand, the periodic boundary conditions may bring some irregularities for considering elongated objects such as semi-stiff polymers, films or membranes, especially if the linear size of the lattice is comparable or shorter to the object size. In many cases the choice of the boundary condition is irrelevant.

Let us also briefly discuss how to generate random initial polymer conformations, which satisfy the connectivity and excluded volume constraints. This technique is based on generating a self-avoiding walk (SAW) on a lattice [6, 9] and consists of the following steps:

1. The first monomer in the chain is placed randomly on the lattice;
2. Starting from the position of the previous monomer all possible directions for the next monomer, i.e. those which satisfy the excluded volume constraints, are analysed;

3. If no possible direction is left, i.e. the system is trapped, the procedure is repeated starting from step 1;
4. Otherwise, the position for the next monomer is chosen randomly among all possible directions;
5. To produce a SAW of N monomers long steps 2-4 should be performed successfully N times.

In case of several chains in the system the procedure is carried out sequentially for each chain.

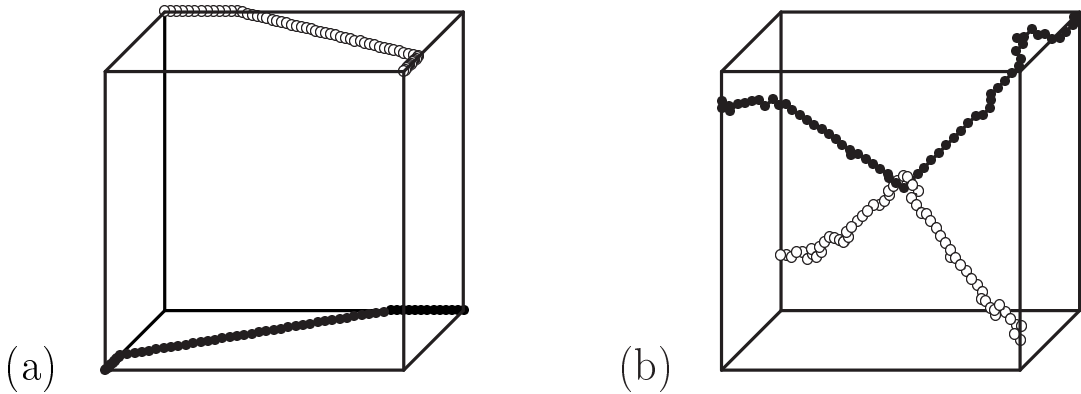
Let us comment on the step 4 in more detail. An example of a SAW on a 2-dimensional lattice is shown in Fig. 3.1. Here there are three possible directions, denoted as m_1 , m_2 and m_3 , to extend the walk. In the simplest case one can choose any of these directions uniformly randomly, i.e. with the probability equal to $p = 1/3$. To reduce the frequency of the system being trapped and to increase the number of extended walks in the ensemble of SAWs, one can apply more sophisticated algorithms for generating SAWs [6, 9].

These are based on the following procedure. First, one should calculate the number of possible distinct SAWs for each direction. Then, a direction for the SAW is chosen proportionally to that its number. However, since computational expenses increase exponentially with the length of walk, the procedure of calculating the number of distinct SAWs can include only a few steps. We shall call this as the number of steps to look ahead, n_s , with $n_s = 0$ corresponding to choosing each walk with equal probability. In the particular example in Fig. 3.1, the scheme with $n_s = 1$ gives: $p_{m_1} = 1/7$ and $p_{m_2} = p_{m_3} = 3/7$, and with $n_s = 2$ gives: $p_{m_1} = 1/9$ and $p_{m_2} = p_{m_3} = 4/9$. Thus, the larger the number of steps to look ahead, the larger probabilities for extended walks.

It is important to stress that the present model does not permit local and,

Figure 3.2: An illustration of the nonphantomness of chains in the lattice Monte Carlo simulation.

Here $N = 62$, $M = 2$, $L = 50$, the external field is, $E = 0.1k_B T$. The conformation depicted in Fig. (a) corresponds to the initial state, and Fig. (b) corresponds to the conformation after 100,000 MCS (6,200,000 Monte Carlo steps).



of course, translational moves that violate integrity of links, i.e. the chain is *nonphantom*. This is ensured automatically due to our particular parameters of the connectivity and the excluded volume restrictions. Thus, no additional procedure to check link integrity, which may slow down significantly the execution of the code, is really needed. This can be established from rather direct, but quite intricate geometrical considerations as well as purely ‘experimentally’ using the following simple routine.

Let us consider two chains on the lattice, $L < N < 2L$, forming a cross in the projection to the xy -plane (see Fig. 3.2a). First chain is at the bottom of the lattice (black monomers), and the second one is at the top (white monomers). We switch on a weak external field which pushes the first chain upwards and the second chain downwards. End monomers of both chains are restricted to move only along the z -direction (along the field). Then we initiate the Metropolis iterations and check if the chains could pass through each other, in which case they would reach the opposite sides of the lattice. Such a system conformation presented in

Fig. 3.2b is obtained after a rather long evolution. So we find in all cases that the chains are indeed *nonphantom*, as polymers in nature. Therefore, this model is suitable for study of the effects of topological constraints and entanglements.

* * *

To conclude this section we would like to summarise the main features, possibilities and disadvantages of the current lattice Monte Carlo model.

1. It allows one to study equilibrium and dynamics, including nonequilibrium dynamics (kinetics), of several polymers of arbitrary chemical composition in a finite volume. It is an important advance compared with SAW (self-avoiding walks) methods, which cannot address dynamical properties. “Weak” pair-wise interactions can be easily included.
2. Using discrete rather than continuous space allows us to reduce significantly the computational expenses in calculation of two-body potential interactions. Thus, the time required to perform a Monte Carlo step does not depend on size of the system. It is equal to approximately $10\mu s$ of CPU time for a processor like Pentium PRO or similar in productivity.
3. Such productivity allows one to resolve, typically, the following questions in a matter of days of CPU time:
 - a. Build approximate phase diagrams of systems of sizes up to $NM = 200$;
 - b. Make very precise calculations of averages ($Q \approx 10^5$ and more) for a given temperature for systems of sizes up to $NM = 1000$;
 - c. Study general behaviour of systems of sizes $NM \geq 10^3$ and larger.
4. Importantly, since the current scheme of polymer movements does not brake the polymer links integrity, one can study effects of topological constraints, the problem which can hardly be treated by any analytical or mean-field approach.
5. The model is free from the artefacts of mean-field theories.

As disadvantages of the current model we should mention the following ones:

1. The lattice itself does not allow one to advance in the degree of complexity of local interactions; the model is appropriate for only a coarse-grained description, which allows one to study the most generic properties of macromolecules in solution;
2. The lattice model can not produce the correct structure factor for wave lengths shorter than the lattice spacing ;
3. The lattice slightly overemphasises crystallisation phenomena;
4. It is hard to implement sophisticated collective moves in the Monte Carlo method generally, and in the lattice model especially. These would be important for addressing such phenomena as glassy dynamics.

3.2 Monte Carlo Model of the Ring Homopolymer in Continuous Space

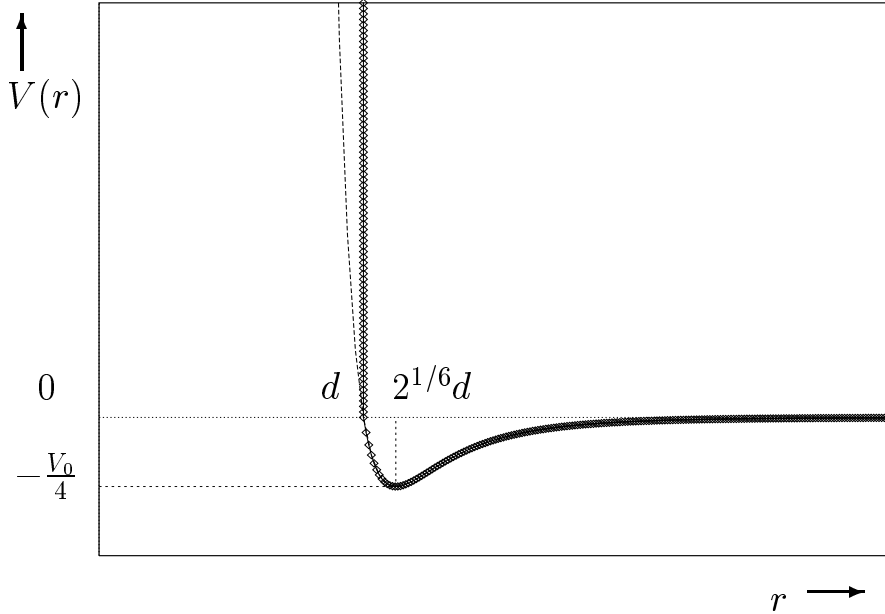
In this section we describe a Monte Carlo model for the ring homopolymer in continuous space. In some cases the lattice model described in section 3.1 may produce some artefacts due to the discreteness of space and the lack of complete rotational invariance. Thus, the main purpose of introducing this model is to compare the results here from those from the lattice Monte Carlo model and the Gaussian self-consistent theory.

The model is implemented for a single homopolymer consisting of N monomers connected by springs or sticks in a ring, which additionally interact with each other via a pair-wise short ranged spherically symmetric potential. In the case of springs, which provide a soft connectivity condition between nearest monomers along the chain, the model Hamiltonian may be written as,

$$H = \frac{k_B T}{2l^2} \sum_i (\mathbf{X}_i - \mathbf{X}_{i-1})^2 + \frac{k_B T \lambda}{2l^3} \sum_i (\mathbf{X}_{i+1} + \mathbf{X}_{i-1} - 2\mathbf{X}_i)^2 +$$

Figure 3.3: The two-body interaction potential used in the model of a homopolymer ring in continuous space.

Dashed line denotes the hard-core repulsion of the Lennard-Jones 6-12 potential.



$$+\frac{1}{2} \sum_{ij, i \neq j} V(|\mathbf{X}_i - \mathbf{X}_j|), \quad (3.30)$$

where analogously to Eq. (2.8) the first term is the elastic energy of springs and the second term in Eq. (3.30) is the bending energy, with l and λ being the statistical and the persistent lengths of the chain respectively. In the bead-and-stick model instead of the first term in Eq. (3.30) we keep the distances between nearest monomers along the chain fixed and equal to l . Unlike the GSC theory, where one has to introduce a virial-type expansion representing the pair-wise potential, here we use the following two-body interaction potential explicitly,

$$V(r) = \begin{cases} +\infty & \text{for } r < d \\ V_0 \left(\left(\frac{d}{r}\right)^{12} - \left(\frac{d}{r}\right)^6 \right) & \text{for } r > d \end{cases}, \quad (3.31)$$

which is exhibited in Fig. 3.3. Thus, monomers are represented by hard spheres of the diameter d , having a weak short ranged Lennard-Jones attraction of characteristic strength V_0 .

Similarly to the lattice model, during simulation we change the strength of the two-body attraction V_0 , which can be viewed as basically the “inverse temperature”, rather than changing the temperature T itself. The procedures of collecting statistical data for equilibrium and kinetics are completely analogous to those of the lattice model in section 3.1.

Let us describe the procedure of finding a new trial position for a monomer. In the case when monomers are connected by springs the new coordinate of a monomer can be sought as,

$$q^{new} = q^{old} + r_{\Delta}, \quad (3.32)$$

where q stands for X , Y and Z spatial projections and r_{Δ} is the random number uniformly distributed in the interval $[-\Delta, \Delta]$. Here Δ is some additional parameter of the Monte Carlo scheme, which, in a sense, characterises the timescale involved in the Monte Carlo sweep (MCS). For small values of Δ most of Monte Carlo trials would be accepted, since the corresponding energy change is rather small. But in this case the MSC involves a small change in the system conformation, thus, corresponding to time evolution on short timescales. Making the parameter Δ large is also not so good since the corresponding energy change would be unfavourably high, leading to most of Monte Carlo trials being rejected by the Metropolis check. The parameter Δ may be adjusted during a simulation to follow the “golden rule” by keeping the ratio of acceptance of Monte Carlo moves approximately at a constant value, 1/2 being not a bad choice. However, one should be careful when adjusting the parameter Δ in dynamics and, especially, kinetics since a change of Δ leads to a rescaling of the time unit in the system.

In the case of sticks of fixed length l we should keep fixed the distance between the monomer which we would like to move and its chain neighbours. Thus,

possible positions of a monomer lie on a circle in the perpendicular direction to the vector connecting the chain neighbours. If the spatial positions of the $(i-1)$ -th and $(i+1)$ -th monomers are given by triplets $X_{i-1}, Y_{i-1}, Z_{i-1}$ and $X_{i+1}, Y_{i+1}, Z_{i+1}$ respectively we introduce their relative distance projections, $\Delta X = X_{i+1} - X_{i-1}$, $\Delta Y = Y_{i+1} - Y_{i-1}$ and $\Delta Z = Z_{i+1} - Z_{i-1}$. The circle, where the new position of the i -th monomer should be sought, is parametrised by the angle variable φ , in the range, $-\pi \leq \varphi < \pi$, as follows,

$$X = \frac{1}{2}(X_{i-1} + X_{i+1}) - \sqrt{\frac{l^2 - (\Delta X^2 + \Delta Y^2 + \Delta Z^2)/4}{\Delta X^2 + \Delta Y^2}} \times \left(\Delta Y \cos \varphi + \frac{\Delta X \Delta Z}{\sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2}} \sin \varphi \right), \quad (3.33)$$

$$Y = \frac{1}{2}(Y_{i-1} + Y_{i+1}) + \sqrt{\frac{l^2 - (\Delta X^2 + \Delta Y^2 + \Delta Z^2)/4}{\Delta X^2 + \Delta Y^2}} \times \left(\Delta X \cos \varphi + \frac{\Delta Y \Delta Z}{\sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2}} \sin \varphi \right), \quad (3.34)$$

$$Z = \frac{1}{2}(Z_{i-1} + Z_{i+1}) + \sqrt{\frac{l^2 - (\Delta X^2 + \Delta Y^2 + \Delta Z^2)/4}{\Delta X^2 + \Delta Y^2 + \Delta Z^2}} \times \sqrt{\Delta X^2 + \Delta Y^2} \sin \varphi. \quad (3.35)$$

In the case when $\Delta X^2 + \Delta Y^2 \leq \varepsilon^2$, where ε is some small number, i.e. the previous and the next monomers are nearly along Z direction, the parametrisation of the circle is just the following,

$$X = X_{i-1} + \sqrt{l^2 - (\Delta X^2 + \Delta Y^2 + \Delta Z^2)/4} \cos \varphi, \quad (3.36)$$

$$Y = Y_{i-1} + \sqrt{l^2 - (\Delta X^2 + \Delta Y^2 + \Delta Z^2)/4} \sin \varphi, \quad (3.37)$$

$$Z = \frac{1}{2}(Z_{i-1} + Z_{i+1}). \quad (3.38)$$

A local monomer move here hence results in a pivot movement of the i -th monomer in respect to its neighbours. Similarly to the case of springs in Eq. (3.32) the new angle can be found as,

$$\varphi^{new} = \varphi^{old} + r_{\Delta}, \quad (3.39)$$

where $\Delta \leq \pi$. In the case $\Delta = \pi$, i.e. r_π is a random number uniformly distributed in the interval $[0, 2\pi]$, and the new position of a monomer is not correlated with the previous one. During Metropolis evolution the centre-of-mass of the polymer performs a random walk, and, generally speaking, is moving away from the zero coordinate. However, the precision with which coordinates of monomers are kept in the computer memory (we use the single precision floating numbers of 4 byte length keeping around 7 decimal digits in the mantissa) deteriorates on moving monomers away from the zero coordinate. Thus, after some number of elementary steps the centre-of-mass of the system is moved back to the zero coordinate.

The change of energy of the i -th monomer due to a local move is then calculated according to the following formula,

$$\begin{aligned}
\Delta E &= E(\mathbf{X}_i^{fin}) - E(\mathbf{X}_i^{ini}), & (3.40) \\
E(\mathbf{X}_i) &= \frac{k_B T}{2l^2} \left((\mathbf{X}_i - \mathbf{X}_{i-1})^2 + (\mathbf{X}_i - \mathbf{X}_{i+1})^2 \right) + \\
&\quad + \frac{k_B T \lambda}{2l^3} \left((\mathbf{X}_i + \mathbf{X}_{i-2} - 2\mathbf{X}_{i-1})^2 + \right. \\
&\quad \left. + (\mathbf{X}_{i+1} + \mathbf{X}_{i-1} - 2\mathbf{X}_i)^2 + (\mathbf{X}_{i+2} + \mathbf{X}_i - 2\mathbf{X}_{i+1})^2 \right) + \\
&\quad + \sum_{j, j \neq i} V(|\mathbf{X}_i - \mathbf{X}_j|). & (3.41)
\end{aligned}$$

Thus, as one can see from the last term in Eq. (3.41), here the CPU time required to perform an elementary move of a monomer grows proportionally to the system size, N .

We should mention also that the procedure for calculation of the two-body interactions can be optimised. Since monomer two-body interactions are short ranged in nature we can introduce a cut-off from some value of radius r , e.g. $V(r) = 0$ for $r > 2d$. Then one can create a cubic lattice with the lattice spacing equal to the interaction cut-off. Each lattice site here addresses the appropriate space cubic cell and contains the list of the monomers which appear in this part of

space. Then, the sum over all monomers in the chain can be replaced by the sum over monomers which belong to the current lattice cell and its neighbours, which gives 27 lattice cells. So the memory consumption and CPU expenses to keep and update the lattice can be made rather low. In fact, as the lattice spacing not necessarily has to be equal to the interaction cut-off, one passes through some other regions of lattice space to reach all possible monomers in contact. Introduction of a lattice requires the introduction of a finite volume with, analogously to the lattice Monte Carlo, periodic or reflective boundary conditions. This is a standard procedure normally called the domain decomposition and it is rather efficient especially for low density states where few monomers interact with each other simultaneously. In the worst case scenario of a high density state for the interaction cut-off equal to $2d$, i.e. if each lattice cell contains about $2^3 = 8$ monomers, we may expect that the domain decomposition technique reduces the number of summations in Eq. (3.41) to $8 \times 27 - 1 = 215$ monomers. Thus, this technique is not well justified for either rather small dense systems, where the traditional technique works fine, or for very long-ranged potentials.

Finally, let us discuss the question of the polymer link integrity in respect to the local monomer movements. For the model with springs the answer is trivial: thermodynamical fluctuations can pull springs enough to allow one link to pass through another. The same concerns monomer-monomer hard-core potentials, $V(r)$, for which $V(r) < +\infty$ for any $r > 0$, e.g. repulsive part of the Lennard-Jones potential, $(r/d)^{-12}$. The question is somewhat more delicate for the case of hard spheres of diameter d , the centre-of-masses of which are connected by inelastic sticks of length l . Here three possibilities exist:

1) for $l/d \geq \sqrt{2}$ there are local monomer moves which break the integrity of links, thus the chain is phantom;

- 2) for $l/d < 2/\sqrt{3}$ the chain is nonphantom;
- 3) for $2/\sqrt{3} \leq l/d < \sqrt{2}$ the answer about the integrity of links depends on how far we allow a monomer to move in respect to its initial position, namely the parameter Δ in Eq. (3.39).

These statements can be established from either a direct geometrical analysis, or by the experimental procedure similar to that used in the lattice model (see arguments around Fig. 3.2). The case 3) is most interesting in a sense that here one can control the property of the link integrity, chain phantomness, by an additional parameter which does not directly affect any of the equilibrium static observables. Since the parameter Δ somehow represents the time in the system, it does, however, affect the dynamic and kinetic properties.

“A fractal is by definition a set for which the Hausdorff–Besicovitch dimension strictly exceeds the topological dimension.”

Benoit Mandelbrot

Chapter 4

Infinitely Dilute Homopolymer Solutions

In this chapter we consider various aspects of the equilibrium and kinetic behaviour of the simplest polymer system, that is the homopolymer in infinitely dilute solution. Here one can neglect the inter-chain interactions and consider a single macromolecule. We study the equilibrium coil-to-globule transition, as well as various characteristics and scalings of the extended Flory coil and condensed globular states. The kinetic laws at the collapse transition of the flexible homopolymer are elucidated. We also consider how additional interactions such as the local rigidity and short ranged three-body repulsions affect the conformational structure of the homopolymer at equilibrium and in kinetics.

4.1 The Good Solvent Regime

In this section we would be interested in the most universal equilibrium properties of a flexible polymer chain in the good solvent obtained from a lattice Monte Carlo simulation. In particular, a special attention is paid to the swelling exponent ν

of the chain defined as,

$$R_g^2 = a^2 N^{2\nu}, \quad (4.1)$$

where the mean squared radius of gyration defined is defined by Eq. (2.44). The inverse to the swelling exponent ν has the meaning of the *fractal dimension* of the polymer chain.

The value of the swelling exponent has been studied by a variety of techniques. The value predicted by the simple Flory mean-field theory [12, 13] is $\nu_F = 3/5$. Theories based on integral equations techniques give a slightly higher prediction $\nu_F = 0.63 - 0.66$, depending on the closure relation involved [166–169]. The most accurate prediction is believed to be given by the renormalisation group technique using the de Gennes relation to the field theory $\nu_F \simeq 0.588$ [11, 37].

On a lattice the repulsive regime at $\chi = 0$ corresponds to the so-called athermal situation, which in our model formally means that all Monte Carlo moves consistent with the connectivity and excluded volume constraints should be accepted, i.e. there is no need to perform the Metropolis check.

Application of the self-avoiding walk (SAW) technique does not allow one to obtain the correct value of the Flory exponent easily. Increasing the number of steps to look ahead, n_s , (see explanations to Fig. 3.1 in Sec. 3.1) improves the agreement for rather short chains only (few first lines in Tab. 4.1). Comparing values of the effective exponents ν_1 and ν_2 from Tab. 4.1 one can also conclude that the effective Flory exponent decreases with increasing polymer length, approaching the exponent of the ideal coil, $\nu_i = 1/2$, rather quickly. The reason for this deficiency is quite clear. Indeed, it is known that the SAW technique could give a satisfactory accuracy if the parameter n_s is only a few times less than the polymer length. However, in this model the computational expenses for creating a SAW grow approximately as 20^{n_s} , so that for chains of a few hundreds of monomers generating SAWs with $n_s \simeq 20$, which is still very small compared

Table 4.1: Values of the mean squared radius of gyration, $3R_g^2$, vs the degree of polymerisation, N , for different simulation procedures.

Here Q is the number of statistical measurements, p_r is the probability of making reptations in the dynamical scheme, with $1 - p_r$ being the probability of making local monomer moves, and n_s is the number of steps to look ahead in the SAW static simulations. The exponents ν_1 and ν_2 were obtained by a least square fit of $\log R_g$ vs $\log N$ in the ranges $100 - 1000$ and $500 - 1000$ respectively.

| N | $p_r = 0$ | $p_r = 1/N$ | $p_r = 0.1$ | $p_r = 1$ | $n_s = 0$ | $n_s = 1$ | $n_s = 2$ | $n_s = 3$ |
|---------|-----------|-------------|-------------|-----------|-----------|-----------|-----------|-----------|
| Q | 80,000 | 60,000 | 40,000 | 40,000 | 100,000 | 40,000 | 40,000 | 40,000 |
| 7 | 3.323 | 3.278 | 3.285 | 3.196 | 3.160 | 3.301 | 3.317 | 3.321 |
| 10 | 5.356 | 5.209 | 5.209 | 5.001 | 4.916 | 5.252 | 5.331 | 5.335 |
| 15 | 9.086 | 8.738 | 8.636 | 8.193 | 8.006 | 8.746 | 8.959 | 9.038 |
| 20 | 13.14 | 12.55 | 12.29 | 11.50 | 11.21 | 12.39 | 12.80 | 13.03 |
| 30 | 22.21 | 20.74 | 19.88 | 18.44 | 17.82 | 19.98 | 20.85 | |
| 50 | 42.16 | 38.50 | 35.73 | 32.63 | 31.54 | 35.64 | 37.62 | |
| 70 | 64.89 | 57.66 | 52.54 | 48.45 | 45.50 | 52.01 | 55.11 | |
| 100 | 98.17 | 88.15 | 77.75 | 72.54 | 66.52 | 76.42 | 81.48 | |
| 150 | 162.2 | 142.7 | 122.6 | 112.9 | 102.1 | 118.0 | 126.3 | |
| 200 | 227.2 | 201.2 | 169.0 | 156.8 | 138.0 | 160.8 | 171.8 | |
| 300 | 381.2 | 322.7 | 261.8 | 245.9 | 211.2 | 244.9 | 263.7 | |
| 500 | 696.5 | 588.1 | 460.0 | 435.9 | 358.0 | 415.9 | 450.5 | |
| 700 | 1029 | 856.1 | 667.2 | 637.6 | 505.5 | 586.9 | 637.8 | |
| 1000 | 1643 | 1329 | 991.5 | 947.7 | 728.6 | 849.7 | 921.6 | |
| ν_1 | 0.608 | 0.587 | 0.551 | 0.559 | 0.519 | 0.522 | 0.526 | |
| ν_2 | 0.619 | 0.588 | 0.554 | 0.560 | 0.512 | 0.515 | 0.516 | |

to the degree of polymerisation, would be an unachievable task.

Alternatively, one can use the Monte Carlo simulation with the monomer movements scheme described in Sec. 3.1. It seems that to improve convergence of the system to equilibrium some global moves such as reptations may be included. It would be interesting to study how the frequency of reptations, p_r , affects the averages. Thus, in Tab. 4.1 we present values of the mean squared radius of gyration for different N and the swelling exponent found by fitting this data for different probabilities of doing reptations, p_r . The value obtained without reptations (at the bottom of the left column) is only a little higher than the mean-field prediction. As we increase the probability of doing reptations it starts to

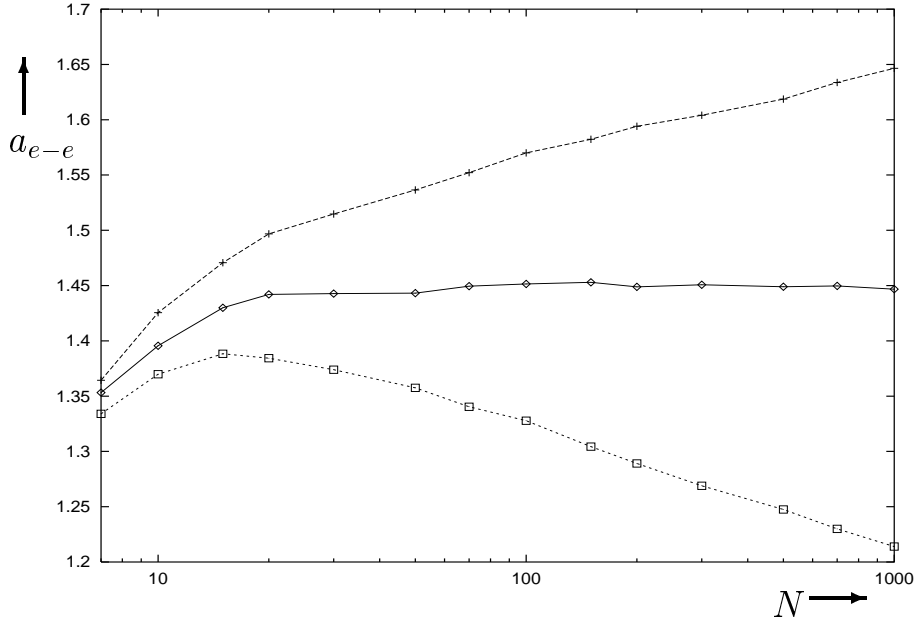
decrease a little, approaching the value from the renormalisation group approach. However, it keeps on decreasing to some value on half way from the Flory exponent to that of the ideal coil $\nu_i = 1/2$ in the limit of pure reptations, $p_r = 1$.

This result is quite unexpected from the point of view of the standard Monte Carlo paradigm. What we have here for the scheme with local moves only is a rather subtle disagreement — just a bit more than the difference between the Flory exponents from the mean-field and renormalisation group. However, the scheme with reptations only seems to lead to a totally wrong result, and thus something must be seriously wrong there. Moreover, by choosing different reptation probabilities one can get almost any result for ν between 0.62 and 0.56 — a situation completely unacceptable for a reasonable simulation.

However, we should emphasise here that if the chain was phantom we would not have this problem. Thus, for a nonphantom chain there must be some violation of the conditions which ensure the convergence of estimates for averages produced by Monte Carlo to the equilibrium values. This violation is very weak for local moves and it exists entirely due to global topological constraints, whereas it is rather serious for reptations. In order to understand the problem better let us examine the behaviour of some relevant observables for different reptation probabilities p_r .

In Fig. 4.1 we draw the dependence of the effective monomer size $a_{e-e} \equiv \langle |\mathbf{X}_{N-1} - \mathbf{X}_0| \rangle / N^{\nu_F}$, which is defined as the end-to-end distance normalised to the degree of polymerisation raised to the power of the exact theoretical value of the Flory exponent, versus N . The effective monomer size, a_{e-e} increases for rather small values of the degree of polymerisation, $N \lesssim 70$, i.e. the effective Flory exponent here is somewhat larger than expected. For larger values of N one can see that in the scheme without reptations the swelling exponent seems to be overestimated as a_{e-e} grows with N ; in the scheme without local moves it is quite

Figure 4.1: Plot of the effective monomer size, $a_{e-e} \equiv R_{e-e}/N^{0.588}$, vs the degree of polymerisation, N , for different simulation procedures. Lines correspond respectively to the following ratios of the reptational moves (from top to bottom): $p_r = 0$ (pluses), $p_r = 1/N$ (diamonds) and $p_r = 1$ (quadrangles).



underestimated as a_{e-e} falls with N ; but if $p_r = 1/N$ the effective monomer size a_{e-e} remains remarkably constant with N . The value of the swelling exponent in Tab. 4.1, found for such a rate of reptations, also seems to be in almost perfect agreement with that of the renormalisation group — an agreement good enough to be attributed to a mere coincidence!

Let us introduce an observable, which indirectly characterises entanglements of the polymer chain. We say that two monomers i and j form a *loop* of length $|i - j|$ if they are closer to each other than some distance b , which in our case was chosen as $3/2$, i.e. more than the distance between the nearest and second diagonal nearest lattice neighbours. Then the quantity L_k is defined as the average number of loops of a fixed length, k , and it may be formally written as follows,

$$L_k \equiv \frac{1}{N - k} \sum_{i=0}^{N-1-k} \langle \theta(b - |\mathbf{X}_i - \mathbf{X}_{i+k}|) \rangle, \quad l_k \equiv L_k/N, \quad (4.2)$$

where we have also introduced the specific number of loops, l_k , and $\theta(x)$ is the standard step-function. Note, that, in principle, a loop can be a part of a larger loop, or include a smaller loop, or even partially overlap with another loop, and so on.

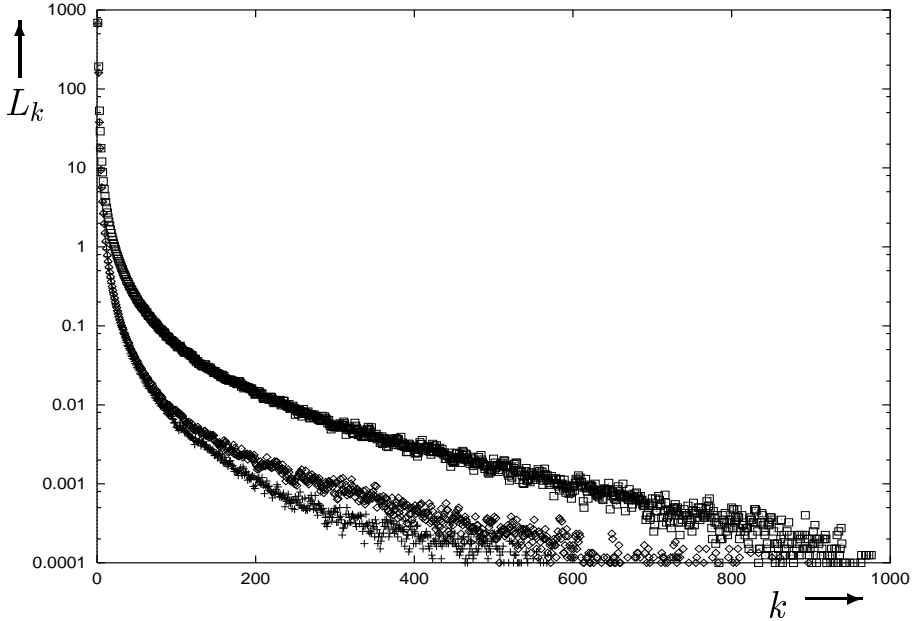
In Fig. 4.2 we draw the k -dependence of the mean number of loops L_k for $p_r = 0, 1/N$ and 1 . At small values of the loop length k all three plots seem to almost coincide and they obey a power law, $L_k \sim k^{-\beta_l}$, with the exponent rather close to the quadruple value of the swelling exponent, $\beta_l = 4\nu_F$. However, one can see that in the scheme with local moves only we obtain a smaller number of long loops, and in the scheme with reptations only we get a larger number of long loops compared to the number of long loops for the scheme with $p_r = 1/N$. In other words, in the scheme with reptations only the ensemble has more compact entangled conformations, leading to a smaller radius of gyration and underestimation of the swelling exponent. Thus, applying reptational moves increases chain entanglements and the average number of loops.

On the other hand, the scheme with local monomer moves only favours topologically simple conformations as it is rather improbable to create a knot by local movements. If some knots already exist in the initial conformation, the local monomer movements would tend to disentangle them. Indeed, one can imagine that simple ‘shaking’ of an entangled boot strap would most likely disentangle it rather than entangle more. This weak topological effect reduces the number of entanglements, and thus the average number of loops, which leads to a larger radius of gyration and a slight overestimation of the swelling exponent in the scheme with local moves only.

Now that we have understood the physical reason for the problem, we may ask what goes wrong with the current Monte Carlo scheme formally. If we return to our discussion of the Metropolis algorithm in Sec. 3.1, one could see that

Figure 4.2: Plot of the number of loops consisting of k segments, L_k , vs their length, k , for different simulation procedures.

Here the degree of polymerisation is equal to $N = 1000$ and lines correspond respectively to the following ratios of the reptational moves (from bottom to top): $p_r = 0$ (pluses), $p_r = 1/N$ (diamonds) and $p_r = 1$ (quadrangles).



the Metropolis check itself is not sufficient for satisfying the detailed balance condition. We also have to ensure the condition that the phase space of the system is sampled uniformly by attempted Monte Carlo moves, so that the transition probability $\Pi_S(old \rightarrow new)$ is symmetric. This may be simple enough to ensure for point-like objects by simply picking them at random, however, in our case due to topology this is not quite so. The above observations do indicate that our sampling procedure is not uniform, but biased. The bias is present in both schemes with reptations only and local moves only, but has the opposite effect.

It might be possible to modify the Metropolis check to recover the detailed balance, but that would require global considerations, which means a slow and inefficient algorithm. Purely experimentally, we have discovered that if about one reptation is performed per Monte Carlo sweep (i.e. per N local moves) the topological effects of entanglements and disentanglements balance each other,

making the sampling of the phase space essentially uniform and the swelling exponent correct.

We should emphasise, however, that this problem of the bad influence of reptations is only present for the Flory coil and it is irrelevant for the ideal coil. Both schemes with local moves only and reptations only would give $\nu = 1/2$ for the ideal solution. That is why reptation techniques are extremely popular and well justified for studying melts and concentrated solutions, in which reptations also may be the only physically relevant motions.

Even though the above discussed effect for the Flory coil is clearly topological, it only presents a problem for the simulation procedure and has no implications for real polymers. As we shall see in Sec. 4.4, for the globule there are different kinds of topological effects, which do have truly physical consequences for the structure of the homopolymer globule.

4.2 Collapse Transition of the Homopolymer

In this section we study the equilibrium coil-to-globule transition in infinitely dilute homopolymer solution. We present results from different simulation models, as well as from the Gaussian self-consistent theory. Thus, increasing the effective monomer-monomer attraction causes polymer to transform its conformational state from the extended Flory coil, which is entropically favourable in the weak attraction regime, to that of the compact globule, where increasing number of monomer-monomer contacts leads to a smaller energy. The transition occurs in a comparatively narrow interval in interaction parameters, though it is not a true phase transition due to finite chain lengths involved. It is believed [11,37,38] that the collapse transition of a flexible polymer becomes second order as $N \rightarrow \infty$.

First, let us consider the collapse transition in the lattice Monte Carlo model described in Sec. 3.1. As we have already discussed, the Flory parameter χ

Figure 4.3: Plot of the mean squared radius of gyration, R_g^2 , vs the Flory interaction parameter, χ , in the lattice model of the homopolymer. Lines correspond to the following values of the degree of polymerisation (from bottom to top): $N = 100$ (diamonds), $N = 150$ (pluses) and $N = 200$ (quadrangles).

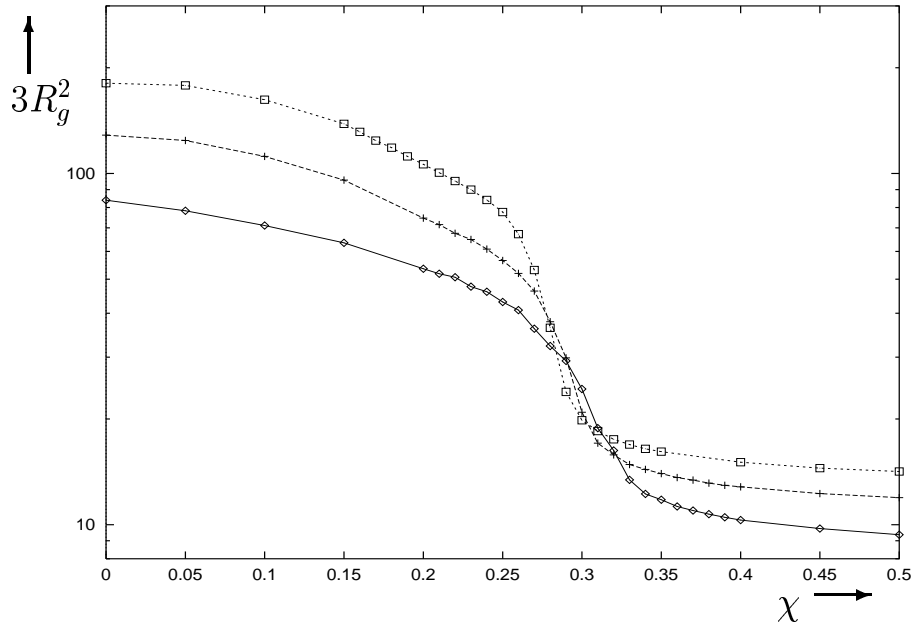


Figure 4.4: Plot of the specific mean energy, $\langle H \rangle / N$, vs the Flory interaction parameter, χ . Lines correspond to the same values of N as in Fig. 4.3.

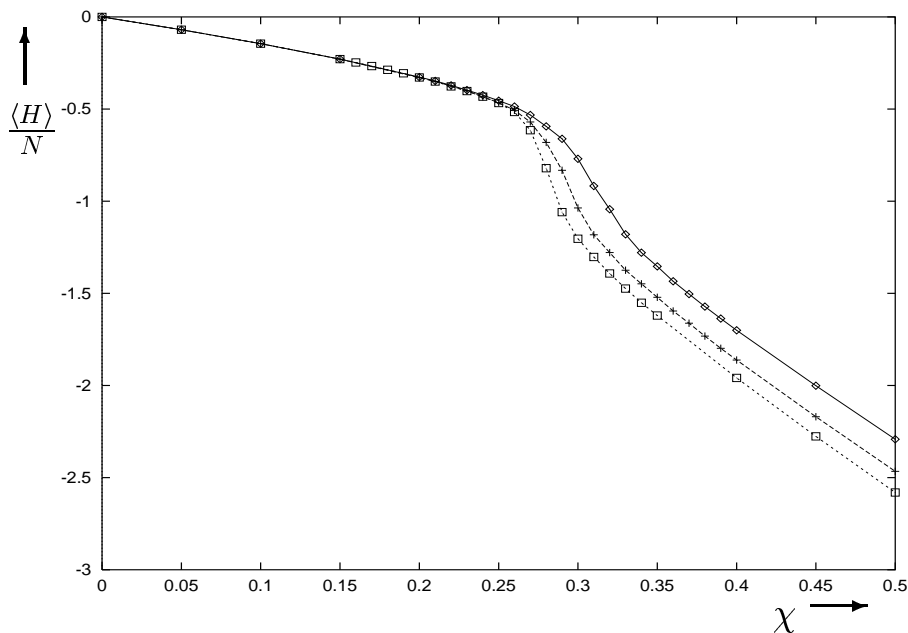
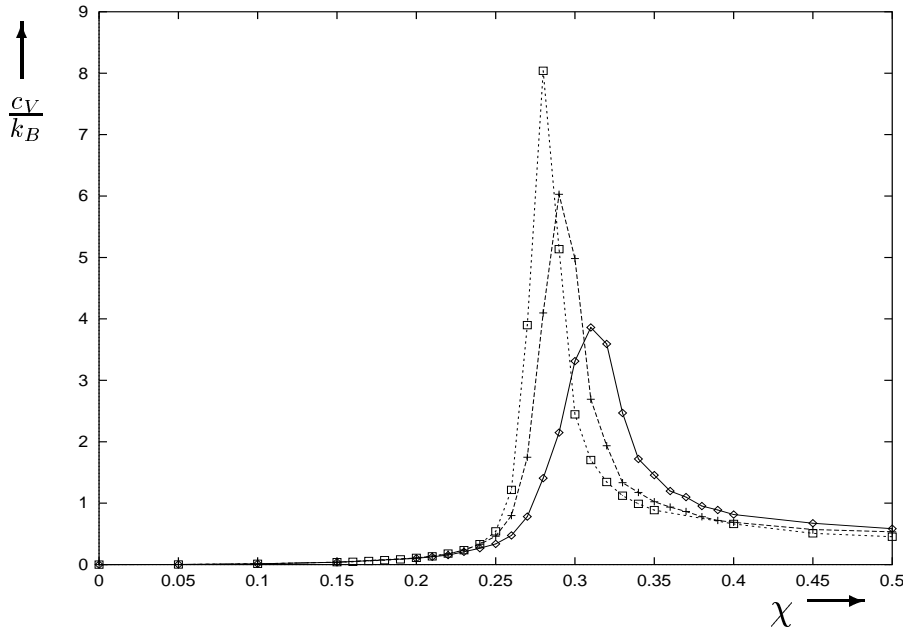


Figure 4.5: Plot of the specific heat capacity, c_V/k_B , vs the Flory interaction parameter, χ .

Lines correspond to the same values of N as in Fig. 4.3.



defined by Eq. (3.28), which plays the role of the inverse temperature, along with the degree of polymerisation N and the lattice size L , uniquely determine the thermodynamic properties of the system. Thus, increasing χ causes the polymer to collapse. In Figs. 4.3, 4.4 and 4.5 we present the equilibrium plots of the mean squared radius of gyration, R_g^2 (see Eq. (2.44)), the specific mean energy, $\langle H \rangle/N$, where H is given by Eq. (3.2), and the specific heat capacity defined as,

$$\frac{c_V}{k_B} = \frac{\langle (\Delta E)^2 \rangle}{N(k_B T)^2}. \quad (4.3)$$

Here $\langle (\Delta E)^2 \rangle$ denotes the mean squared dispersion calculated by averaging over the statistical ensemble.

Thus, the coil-to-globule transition occurs in some interval in χ , being sharper for larger polymer chains. Across the collapse transition the mean squared radius of gyration drops dramatically, changing much slower away from the transition. As one can see from Figs. 4.4 and 4.5 the collapse transition is also characterised

Table 4.2: The behaviour of various observables across the collapse transition vs the degree of polymerisation, N .

Here $3R_g^2$ is the mean squared radius of gyration in the collapsed state at $\chi = 1$. χ_c and $c_V(\chi_c)$ denote the position and the height of the peak in the specific heat capacity. In the last three columns we present the mean squared radii of gyration from the numerical solution of the GSC equations at equilibrium for the Flory coil state, $R_g^2(F)$ at $u^{(2)} = 5$, $u^{(3)} = 0$ and for the collapsed globule, $R_g^2(G)$ at $u^{(2)} = -25$, $u^{(3)} = 10$ and $R_g^2(G')$ at $u^{(2)} = -40$, $u^{(3)} = 10$. In the last row ν contains the effective exponent of the appropriate quantity in terms of the degree of polymerisation, N .

| N | $3R_g^2$ | χ_c | $c_V(\chi_c)$ | $R_g^2(F)$ | $R_g^2(G)$ | $R_g^2(G')$ |
|-------|----------|----------|---------------|------------|------------|-------------|
| 30 | 3.88 | 0.45 | 1.5 | 7.202 | 0.91 | 0.637 |
| 50 | 5.40 | 0.37 | 2.2 | 14.10 | 1.32 | 0.917 |
| 70 | 6.71 | 0.33 | 2.8 | 21.99 | 1.68 | 1.163 |
| 100 | 8.45 | 0.31 | 3.9 | 35.22 | 2.17 | 1.494 |
| 150 | 11.05 | 0.292 | 6.1 | 60.06 | 2.90 | 1.983 |
| 200 | 13.36 | 0.280 | 8.1 | 87.36 | 3.55 | 2.422 |
| 300 | 17.2 | 0.27 | 14.2 | 146.2 | 4.72 | 3.206 |
| ν | 0.33 | | 1.12 | 0.652 | 0.346 | 0.339 |

by a decrease of the energy and a pronounced peak in the specific heat capacity. On increasing χ in the collapsed state the radius of gyration continues to fall, asymptotically reaching the value corresponding to a nearly spherical globule with maximal monomer packing. Obviously, in our case such a conformation possesses the minimal possible energy. Importantly, the swelling exponent ν of the polymer chain (see Eq. (4.1)) changes during the transition, becoming $\nu = 1/3$ in the collapsed phase, which corresponds to a compact 3-dimensional object. This is illustrated in Tab. 4.2 (first column), where we present the values of the mean squared radius of gyration in the collapsed state at $\chi = 1$.

Also in Tab. 4.2 we exhibit the value of the critical point, χ_c , and the height of the peak in the heat capacity, $c_V(\chi_c)$, at this point for various values of the degree of polymerisation, N . Thus, with increasing N the collapse transition point shifts towards lower values of the interaction parameter χ , which reflects the collective nature of the transition. This effect is most pronounced for short

chains, $N < 100$. According to Tab. 4.2, the value of the peak in the specific heat capacity, c_V , grows with the degree of polymerisation somewhat faster than linearly. Away from the collapse transition the specific heat capacity depends weakly on N (see Fig. 4.5).

As one can see from Fig. 4.4 the specific energy in the Flory coil state decreases almost linearly with the interaction parameter χ , with the slope being independent of the degree of polymerisation, N . It means that the average number of monomer–monomer contacts per monomer, c_{mm} , changes weakly with χ and N . Across the collapse transition c_{mm} increases dramatically, being nearly constant in the collapsed region, i.e. the energy continues to decrease linearly in the right–hand–side of Fig. 4.4. However, the value of c_{mm} is N -dependent here, being somewhat larger for longer polymer chains. Indeed, the number of monomers on the surface of the globule possessing approximately half of monomer–monomer contacts compared to the number of those inside the globule grows with N as $N^{2/3}$. Thus, their relative contribution in c_{mm} decreases approximately as $N^{-1/3}$.

Now let us discuss briefly the collapse transition using data from Monte Carlo simulation in the bead–and–spring model of a ring homopolymer in continuous space (see Sec. 3.2). In series of graphs in Figs. 4.6, 4.7 and 4.8 we exhibit plots of the same observables, R_g^2 , $\langle H \rangle/N$ and $c_V/k_B T$ as in Figs. 4.3–4.5. One can see that here the collapse transition is much weaker. Although the mean squared radius of gyration in Fig. 4.6 drops strongly across the transition, the energetical characteristics behave less pronounsly. The specific mean energy in Fig. 4.7 changes its behaviour across the collapse transition from one linear law to another, rather than decreases quickly as in the lattice Monte Carlo simulation. Moreover, the transition is accompanied by a significant increase, rather than a peak, in the specific heat capacity in Fig. 4.8. However, away from the collapse transition the behaviour of the main observables remains quite consistent between

Figure 4.6: Plot of the mean squared radius of gyration, $3R_g^2$, vs the two-body interaction parameter, V_0 , in the homopolymer model in continuous space. Lines correspond to the following values of the degree of polymerisation (from bottom to top): $N = 50$ (diamonds) and $N = 100$ (pluses).

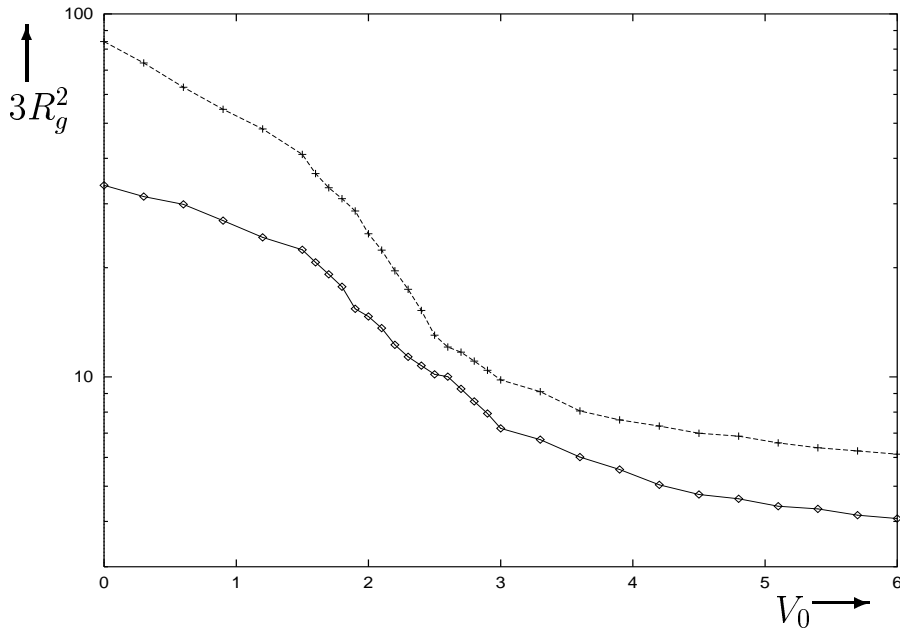


Figure 4.7: Plot of the specific mean energy, $\langle H \rangle / N$, vs the two-body interaction parameter, V_0 . Lines correspond to the same values of N as in Fig. 4.6.

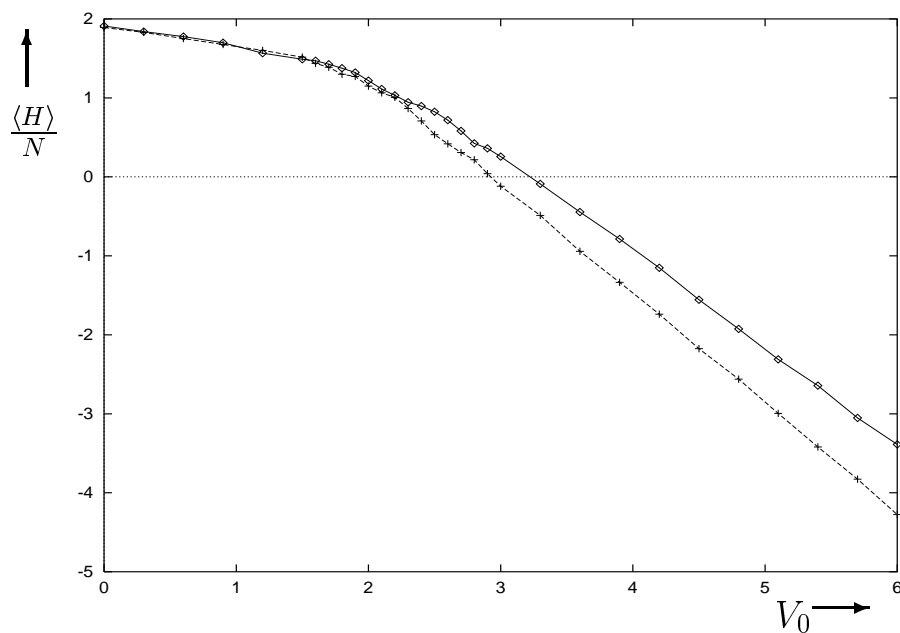
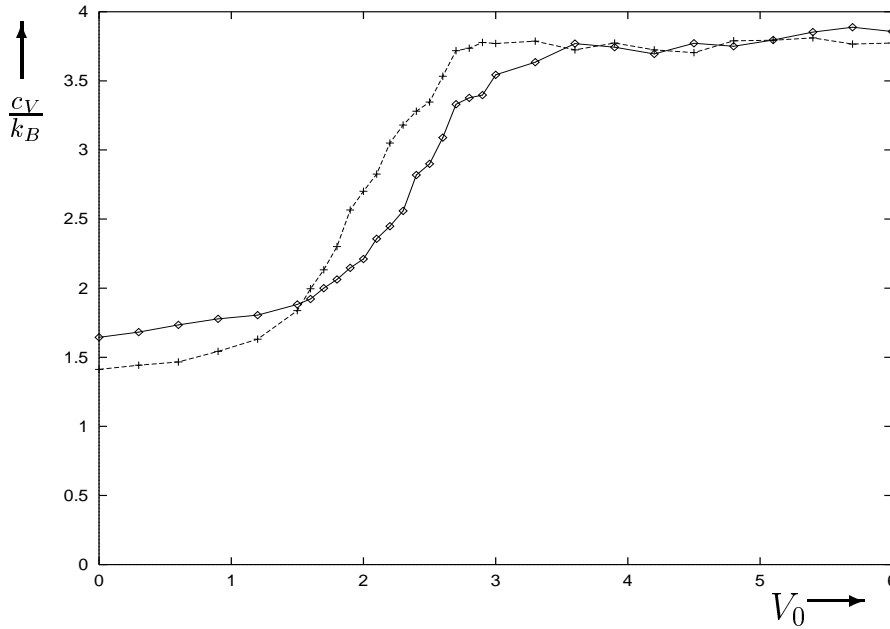


Figure 4.8: Plot of the specific heat capacity, c_V/k_B , vs the two-body interaction parameter, V_0 .

Lines correspond to the same values of N as in Fig. 4.6.



both treatments. In both continuous and discrete space simulations the collapse transition occurs earlier for larger systems.

In fact, behaviour of the main observables at the transition from the continuous space simulation is more consistent with what is expected of the second order phase transition. Indeed, one may expect that for an infinitely long polymer chain the plot of the mean specific energy would remain continuous, with the slope of the curve changing discontinuously at the point corresponding to the transition. At the same point the specific heat capacity increases discontinuously. The plots for finite-size systems in Figs. 4.7 and 4.8 look quite consistent with such expectations. A phase transition, which is characterised by a discontinuous behaviour of the specific heat capacity and continuous behaviour of the specific energy, is indeed characterised as second order [110].

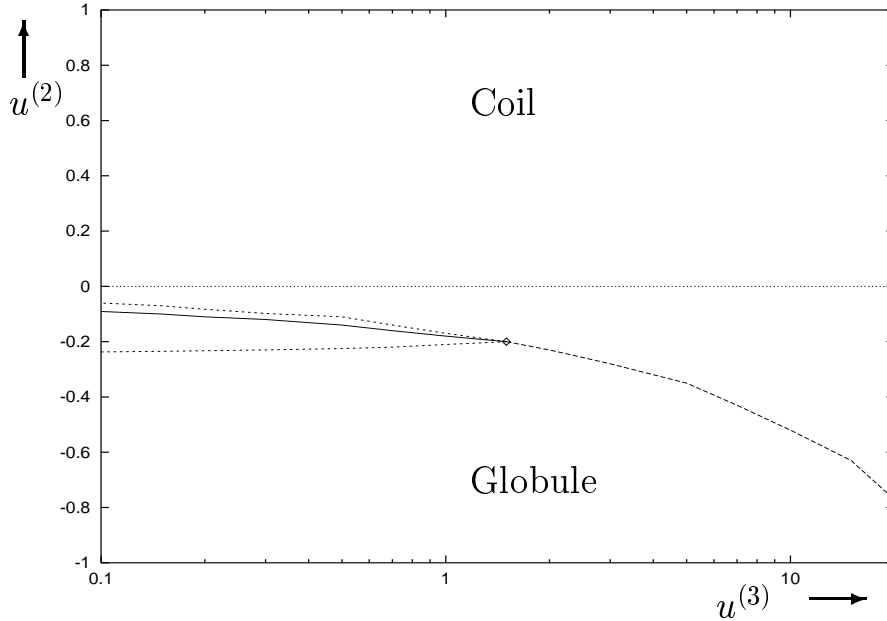
Thus, the lattice Monte Carlo model overemphasises the sharpness of the collapse transition, which looks there more like a λ -type phase transition rather than

a second order one. For the former the specific mean energy in the thermodynamic limit should still change continuously, falling significantly around the transition, with an infinite derivative over the temperature T at the transition point. Hence the specific heat capacity diverges at this point as $c_V \sim -\ln|1 - T/T_c|$.

Such a type of critical behaviour in the lattice model is not too surprising. The condensed phase (the globule) on a lattice is an analog of a crystal rather than of a liquid state. To produce a more adequate description of the liquid phase on a lattice a set of additional and rather cumbersome collective Monte Carlo moves should be included to describe thermal motions in the liquid phase. Thus, on a lattice the collapse transition leads the system to a sort of a crystalline state instead of the liquid-like globule. This condensation is accompanied by formation of a spatial order in the system, which leads to stronger fluctuations, and thus a peak in the heat capacity at the transition point. Generally speaking, both situations may occur in nature depending on the specificity of the polymer and its interactions with the solvent. Often polymer collapse may be accompanied by a crystallisation, possibly even involving some chemical transformation of the monomers.

Now let us consider the collapse transition and the equilibrium phases based on the numerical analysis of the Gaussian self-consistent equations (2.100). In the limit of infinite dilution one does not need to consider the inter-chain distance, Γ , defined by Eq. (2.92). The phase diagram in terms of the second and third virial coefficients for the homopolymer with $N = 100$ is exhibited in Fig. 4.9. At small values of the third virial coefficient the collapse transition becomes discontinuous, being second order for larger $u^{(3)}$. This coefficient mainly characterises the effective monomer-monomer interactions originating from the hard-core repulsion. The larger $u^{(3)}$ the larger is repulsion between monomers i.e. their excluded volume sizes. So the collapse transition in the bead-and-spring model

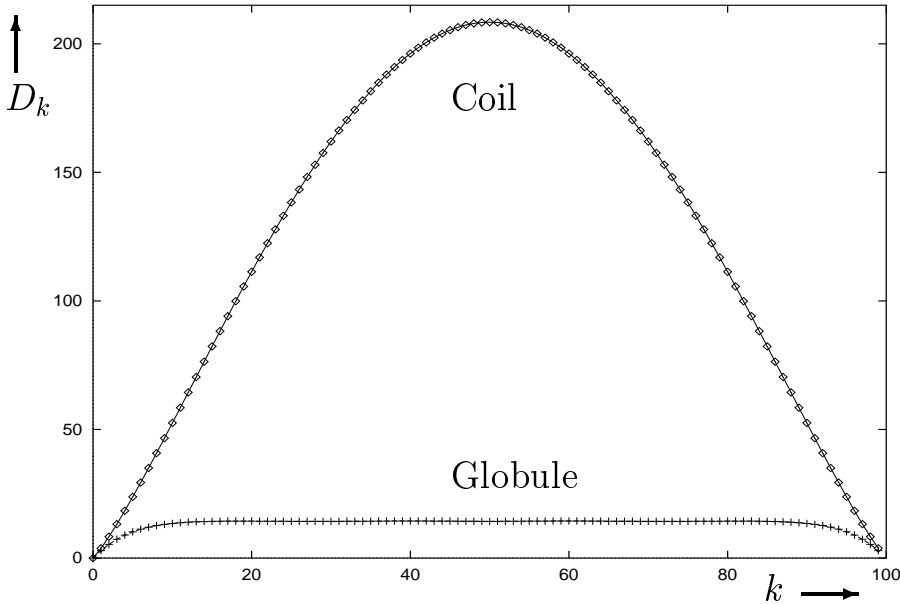
Figure 4.9: Phase diagram of the ring homopolymer with the degree of polymerisation $N = 100$ in terms of the second, $u^{(2)}$, and third, $u^{(3)}$, virial coefficients. Solid long dashed and short dashed lines correspond respectively to discontinuous (first order), continuous (second order) phase transitions and metastability boundaries.



with monomers excluded volume size small enough compared to the statistical length, i.e. a typical distance between nearest neighbours along the chain, becomes first order. The entropic barrier separating the system from the extended to the condensed phase is rather high here and such a system is more an analog of a dilute gas rather than a polymer. Quite naturally, for values of the third virial coefficients $u^{(3)}$, for which the collapse transition is second order, the transition occurs around $u^{(2)} = 0$, where the repulsive regime changes to the attractive. Behaviour of the main observables such as the mean squared radius of gyration and the mean energy across the collapse transition from the GSC method at large $u^{(3)}$ is quite consistent with the results from continuous space simulation: the system shrinks significantly, whilst the mean energy changes continuously from the asymptotic behaviour $\mathcal{E} \sim (u^{(2)})^{2/5}$ in the good solvent to $\mathcal{E} \sim -|u^{(2)}|^2$ in the poor solvent regime.

Figure 4.10: Plot of the mean squared distances, D_k , vs the chain index, k , in the off-lattice Monte Carlo simulation.

Here the polymer length is equal to $N = 100$, the interaction parameter is $V_0 = 0$ for the Flory coil and $V_0 = 5$ for the globule.



Now let us consider the mean squared monomer–monomer distances. For the ring homopolymer there are additional symmetry properties for the elements of the D_{ij} matrix discussed at the beginning of Sec. 2.6. In Fig. 4.10 we present plots of the mean squared distances between monomers, D_k , versus the chain index k , i.e. the relative distance along the chain. These plots from the GSC method are quite similar to those from the Monte Carlo simulation for both the extended and the globular states. In the Flory coil state function D_k increases monotonically until separation of half of the chain, $k = N/2$, and for small chain indices it can be approximated by a power law, $D_k = a^2 k^{2\nu}$, where ν is the swelling exponent and the coefficient a does not depend on the degree of polymerisation. In the collapsed state the function D_k first increases approximately linearly, $D_k = b^2 k$, with the parameter b being almost N -independent, and then saturates reaching a constant value proportional the globule size.

In Tab. 4.2 in the last three columns we present the radii of gyration for

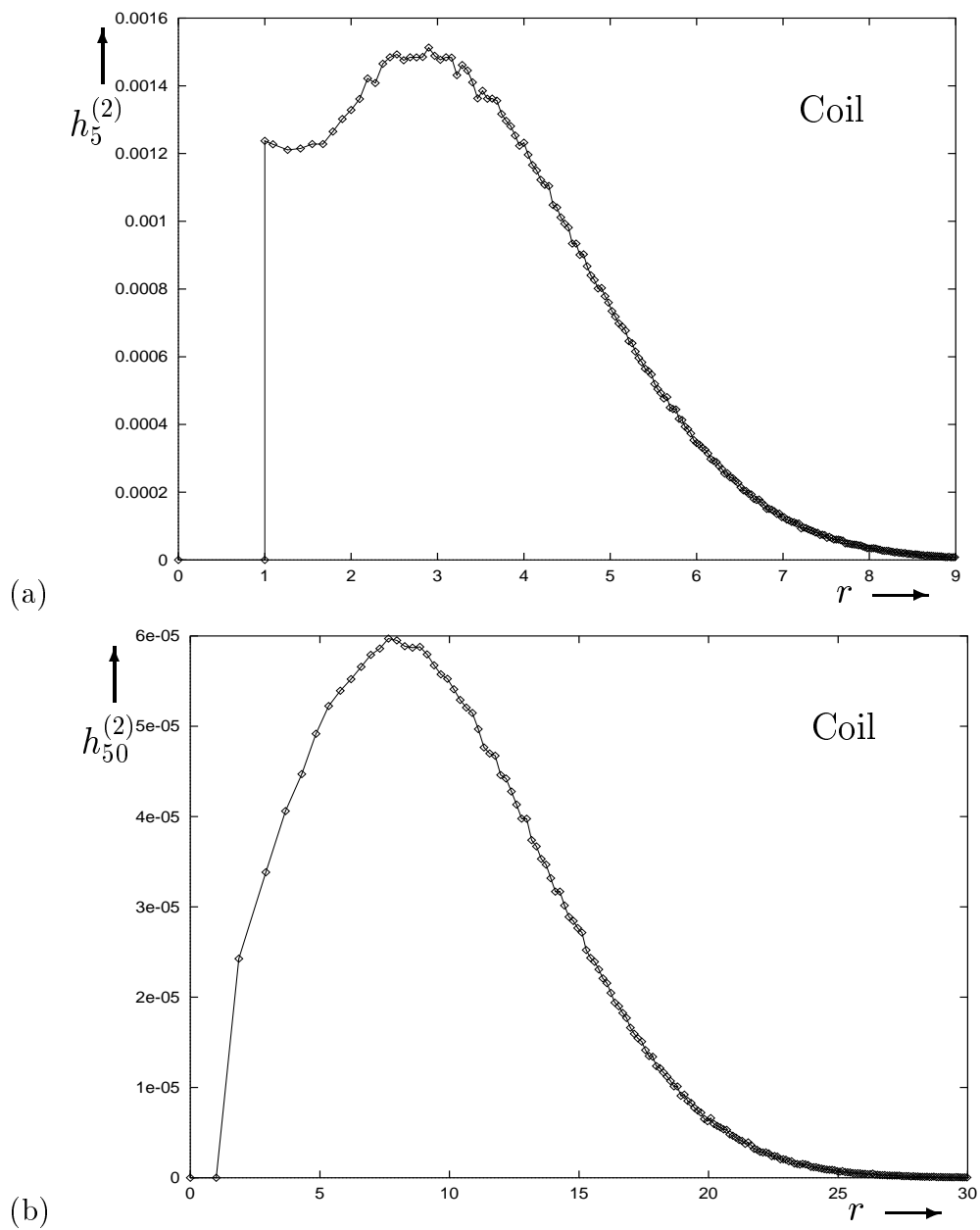
the extended coil and the globule. The swelling exponent for the globular state coincides almost perfectly with $\nu = 1/3$. However, this exponent for the Flory coil state in the GSC method is in a worse agreement with the expected result $\nu = 3/5$ obtained from the mean-field theories [11,38] and the more precise result $\nu \approx 0.588$ from the renormalisation group theory [11,37]. The Flory exponent in the GSC method tends to the so-called Reiss exponent $\nu = 2/3$ as $N \rightarrow \infty$, a deficiency of the method that was recognised and understood some time ago [13, 37, 121, 122]. Indeed this limitation was sufficiently serious that a better equilibrium theory was sought by numerous authors. It is known [115, 116, 121] that by *ad hoc* application of a cut-off the problem at equilibrium is fixed, and one can obtain the normal Flory exponent. We may note here that the problem of incorrect Flory exponent in the GSC method appears for comparatively large N only; as we have observed in Sec. 4.1 the effective swelling exponent for short chains $N \lesssim 100$ is higher than the expected asymptotic value.

Now let us consider the monomer–monomer correlation function from the Monte Carlo simulation in continuous space, calculated from the definition,

$$h_k^{(2)}(r) \equiv \langle \delta(\mathbf{X}_0 - \mathbf{X}_k - \mathbf{r}) \rangle = \frac{1}{4\pi r^2 N} \sum_i \langle \delta(|\mathbf{X}_i - \mathbf{X}_{i+k}| - r) \rangle, \quad (4.4)$$

and which has the meaning of the probability density to find two monomers separated by k links at the distance r from each other. In Fig. 4.11 we exhibit the correlation function for a small value of the chain index k (see Fig. 4.11a) and for the chain index equal to the half of the chain length (Fig. 4.11b) in the Flory coil state. Since monomers in the model are represented by hard spheres of diameter d all two-body correlations are equal to zero on distances smaller than the diameter, i.e. $h_k^{(2)}(r < d) = 0$. At $r = d$ the correlation function changes discontinuously. Note also that for large- k correlation functions, such as in Fig. 4.11b, this jump is almost invisible and the correlation function starts to grow from nearly zero. Small- k correlation functions either grow insignificantly as

Figure 4.11: Plot of the two-body correlation function $h_5^{(2)}(r)$ (Fig. a) and $h_{50}^{(2)}(r)$ (Fig. b) vs the separation, r , for the homopolymer with $N = 100$ and $V_0 = 0$.



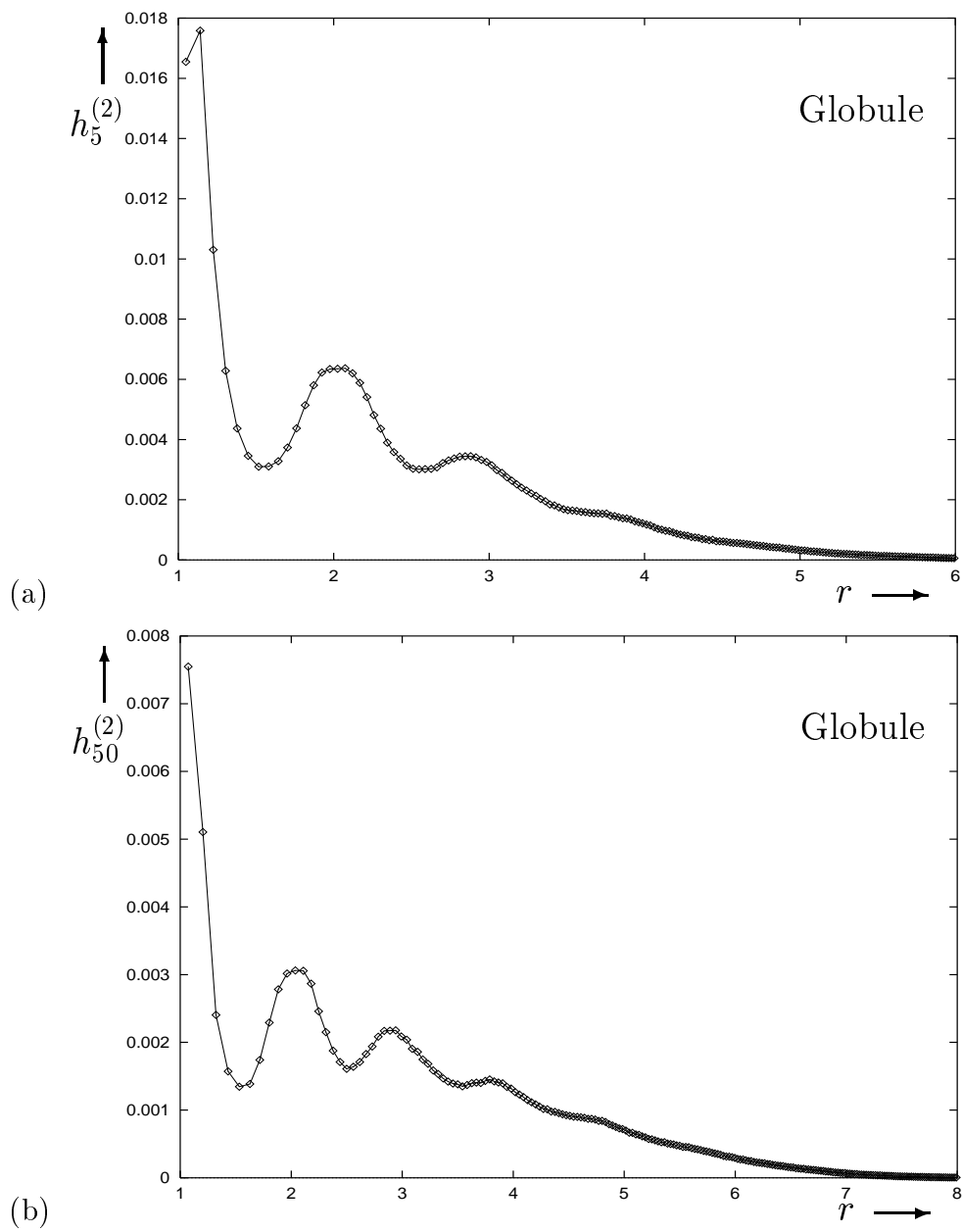
in Fig. 4.11a, or do not grow at all for smaller values of the chain index k . At larger separations these correlation functions start to decrease for distances greater than some value, which depends on the index k , and have an almost perfect Gaussian tail there. We may conclude therefore that in the Flory coil state the correlation function depends strongly on the chain index, k . However, one may expect that it depends on the degree of polymerisation N much weaker, i.e. adding more monomers to the chain does not change significantly the distributions for the existing monomers.

The problem of the GSC method in producing a correct Flory exponent likely appears due to a wrong form of the correlation function for large values of the chain index k . The correlation function should grow as a power first, reaching a maximum at some characteristic scale, which is close to the mean distance between the appropriate monomers. Thus, for distances smaller than this scale the Gaussian shape of the correlation function, which is used in the GSC method in Eq. (2.40), is a poor approximation.

On the other hand, the reason for an incorrect Flory exponent might be more nontrivial. Methods based on closures of chains of the integral equations for the correlation functions [166–169] do overestimate the Flory exponent as well as the GSC method: $\nu = 0.64$ for the Kirkwood hierarchy [167] and $\nu = 0.635$ for the BBGKY hierarchy [169], whilst $\nu = 0.65$ for an open homopolymer chain [168]. These, however, seem much more accurate theories than the GSC method and they are capable of finding the right shape for large- k correlation functions. Also, as we have seen in Sec. 4.1 it is by no means straightforward to produce the correct value for the Flory exponent using direct Monte Carlo simulations.

Similarly, in Fig. 4.12 we present the small- k and large- k correlation functions for the liquid globule. As in the Flory coil state these correlation functions possess

Figure 4.12: Plot of the two-body correlation function $h_5^{(2)}(r)$ (Fig. a) and $h_{50}^{(2)}(r)$ (Fig. b) vs the separation, r , for the homopolymer with $N = 100$ and $V_0 = 5$.



an essentially Gaussian tail. Here the characteristic distance, at which the Gaussian regime appears, is determined by either the statistical length for small values of the chain index k , or by the radius of the whole globule for large- k correlation functions. At smaller distances the correlation functions exhibit an oscillating behaviour typical for a liquid state [111]. These oscillations appear due to formation of a quasi-order in liquids, with neighbouring liquid molecules being either in contact or quite close to each other. Thus, the first maximum in the correlation function coincides approximately with the diameter of hard spheres, d , the second maximum can be found at the distance of approximately $2d$, and so on. Due to increasing disorder in location of molecules further away from each other in the liquid state each consequent peak in the correlation function is lower, and each consequent minimum is higher than the previous ones. Finally, for a macroscopic liquid droplet the correlation function becomes a constant with oscillations vanishing at distances of several diameters of the constituting particles. For example, one can see four such oscillations in the correlation function in Fig. 4.12b. In our case as the globule is formed by collapse of a single chain, the correlation function does not become a constant at large distances, but rather possesses a sort of a Gaussian shape after smoothening the oscillations. This might be the reason why the GSC method works reasonably well in the collapsed state. As we have already mentioned the mean squared distances between monomers in the globule become constant starting from some value of the chain index, k (see Fig. 4.10 and below). The same is correct for the correlation function: starting from the same value of the chain index, k , the shape of the correlation function becomes k -independent. The shape of the correlation function, however, is rather N -dependent on large distances of order of the globule size since adding monomers to the globule increases its size. At smaller values of the chain index the correlation functions possess a shorter Gaussian tail.

4.3 Kinetic Laws at the Collapse Transition of the Homopolymer

The purpose of this section is to present the main kinetic laws that govern the collapse transition of the homopolymer. We would be interested in the laws describing conformational changes after the system has been subject to an instantaneous quench which brings it from the initial equilibrium to the final one. Such changes are normally caused by a temperature jump. However since the second virial coefficient changes most rapidly as a function of temperature $u^{(2)}(T) \simeq a(1 - T_0/T)$ we shall assume that only this parameter is changed $u_i^{(2)} \rightarrow u_f^{(2)}$ and all others remain the same.

The full time evolution of the system from the old equilibrium state to the new one is described by the GSC equations. For a single ($M = 1$) ring homopolymer these equations (2.100) may be rewritten in terms of the mean squared amplitudes of Fourier modes, $\mathcal{F}_q \equiv (1/3)\langle \mathbf{x}_{-q} \mathbf{x}_q \rangle$, as follows,

$$\frac{\zeta_q(t)}{2} \frac{d}{dt} \mathcal{F}_q = k_B T - V_q(t) \mathcal{F}_q(t), \quad V_q(t) = -\frac{2}{3} \frac{\partial \mathcal{E}}{\partial \mathcal{F}_q(t)} \quad (4.5)$$

$$\frac{1}{\zeta_q(t)} = \frac{1}{N\zeta_b} + \frac{1}{3(2\pi)^{3/2}\eta_s N^2} \sum_{j \neq j'} \frac{2 \cos(2\pi q(j - j')/N)}{D_{jj'}^{1/2}(t)}. \quad (4.6)$$

These are extremely complicated nonlinear differential equations for $(N - 1)/2$ independent parameters. Their solution can be found numerically for systems of several hundreds of monomers.

There are two simple regimes amenable to completely analytical treatments. These are close to equilibrium, whether the initial or the final, and thus the behaviour is mainly determined by the properties of that equilibrium state. At equilibrium Eq.(4.5) becomes a simple relation,

$$\mathcal{F}_q^{(eq)} = \frac{k_B T}{V_q^{(eq)}}, \quad (4.7)$$

where $V_q^{(eq)}$ is expressed via $\mathcal{F}_q^{(eq)}$ according to Eqs. (2.78, 2.79, 2.80). These equilibrium GSC equations can be analysed using the dominant balance method [115]. Its main idea is to substitute a power law *Ansatz*,

$$V_q^{(eq)} = A^{-1} N \underline{q}^{2\beta}, \quad \underline{q} = \frac{2\pi q}{N}, \quad (4.8)$$

where A and β are unknown parameters, for which simple equations are then deduced by converting sums into integrals.

After rather straightforward derivations one can obtain the following q -dependence of the effective friction and the normal modes,

$$\zeta_q^{(eq)} \sim \eta_s N \underline{q}^{1-\nu} \quad (\zeta_b N), \quad (4.9)$$

$$\mathcal{F}_q^{(eq)} \sim N^{-1} \underline{q}^{-1-2\nu} \quad \text{for the Flory and ideal coils}, \quad (4.10)$$

with (without) hydrodynamics respectively. Here ν is the swelling exponent (inverse fractal dimension) of the equilibrium state: $\nu = 3/5$ for the Flory coil, $\nu = 1/2$ for the ideal coil, and $\nu = 1/3$ for the collapsed globule. Such a power dependence of the Fourier modes then leads to the power dependence of the mean squared distances, $D_{jj'} = k_B T A |j - j'|^{2\nu}$.

However, it is important to emphasise that for the collapsed globule formula (4.10) does not hold. This is because $D_{jj'}$ looks differently due to screening and thus the dominant balance does not work. A good alternative approximation for the function D_{0m} in the liquid-like globule would be the following one: D_{0j} is linearly increasing function of j until some value $D_{0j_1} = D$, where the N -dependence is, $D_{0j} \sim j_1 \sim N^{2/3}$ and then it remains constant. This approximation can also be obtained from the Lifshitz theory [38] and has been seen by us in Sec. 4.2. Now using the inverse Fourier transformation,

$$\mathcal{F}_q = -\frac{1}{2N} \sum_j \cos\left(\frac{2\pi q j}{N}\right) D_{0j}, \quad (4.11)$$

one can see that $\mathcal{F}_1 \sim N^{1/3}$, instead of the usual scaling $\mathcal{F}_1 \sim N^{2\nu}$ which follows

from the dominant balance. Thus, on the contrary to the coil state, for the liquid-like globule the first normal mode, \mathcal{F}_1 , neither gives the main contribution to the mean squared radius of gyration, nor even has the same scaling law in the degree of polymerisation, N .

Now that we have presented necessary preliminaries let us return to discussion of the linearised kinetics at close to equilibrium conditions.

4.3.1 Linearised GSC equations for early times

First, at early times after the quench $u_i^{(2)} \rightarrow u_f^{(2)}$ the deviation $\Delta\mathcal{F}_q(t) \equiv \mathcal{F}_q(t) - \mathcal{F}_q^{(eq)}$ from the initial equilibrium values is small and we can consider the linearised equations,

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \Delta\mathcal{F}_q &= -\Delta V_q[u_f^{(2)}](t) \frac{\mathcal{F}_q^{(eq)}}{\zeta_q^{(eq)}} - \frac{k_B T}{\mathcal{F}_q^{(eq)} \zeta_q^{(eq)}} \Delta\mathcal{F}_q(t) \\ &- \Delta_2 V_q^{(eq)} (\mathcal{F}_q^{(eq)} + \Delta\mathcal{F}_q(t)), \\ \Delta_2 V_q^{(eq)} &\equiv \Delta \hat{u}^{(2)} \sum_{jj'} \frac{d_{j-j'}^{(q)}}{(D_{jj'}^{(eq)})^{5/2}}, \\ \Delta V_q[u_f^{(2)}](t) &\equiv \sum_{q'} \frac{\partial V_q^{(eq)}[u_f^{(2)}]}{\partial \mathcal{F}_{q'}^{(eq)}} \Delta\mathcal{F}_{q'}(t), \end{aligned} \quad (4.12)$$

where the effective potentials V_q are expressed via the normal modes \mathcal{F}_q by the explicit expression (2.78-2.80) at $M = 1$. Note that although $\mathcal{F}_q(t)$ change smoothly with time, the effective potential experiences a discontinuity at the initial time $t = 0$, $\Delta_2 V_q^{(eq)}$, which is proportional to the variation of the second virial coefficient.

The deviation from the equilibrium value may be sought in the form,

$$\Delta\mathcal{F}_q = \mathcal{A}_q (1 - e^{-\lambda_q t}) \simeq \frac{1}{\zeta_q^{(eq)}} (B_q t + C_q \frac{t^2}{2} + \dots), \quad (4.13)$$

$$\lambda_q \simeq -\frac{C_q}{B_q}, \quad \mathcal{A}_q = \frac{B_q}{\lambda_q \zeta_q^{(eq)}}, \quad (4.14)$$

because the deviation at the initial moment $\Delta\mathcal{F}_q(0)$ must vanish. Here λ_q has

the meaning of the Lyapunov exponent. We find the following expressions,

$$\frac{1}{2}B_q = \Delta \hat{u}^{(2)} \mathcal{F}_q^{(eq)} \sum_{jj'} \frac{d_{j-j'}^{(q)}}{D_{jj'}^{(eq) 5/2}}, \quad (4.15)$$

$$\begin{aligned} \frac{1}{2}C_q &= -V_q^{(eq)} B_q + \Delta \hat{u}^{(2)} B_q \sum_{jj'} \frac{d_{j-j'}^{(q)}}{D_{jj'}^{(eq) 5/2}} - \frac{5}{2} \hat{u}_f^{(2)} \mathcal{F}_q^{(eq)} \sum_{jj'} \frac{d_{j-j'}^{(q)} B_{jj'}}{D_{jj'}^{(eq) 7/2}} \\ &- \hat{u}^{(3)} \mathcal{F}_q^{(eq)} \left[\frac{5}{2} \sum_{jj'j''} \frac{d_{j-j'}^{(q)} D_{j'j''}^{(eq)} + d_{j'-j''}^{(q)} D_{jj'}^{(eq)} - 2d_{jj'j''}^{(q)} D_{jj'j''}^{(eq)}}{(D_{jj'}^{(eq)} D_{j'j''}^{(eq)} - D_{jj'j''}^{(eq) 2})^{7/2}} \right. \\ &\quad \left. (D_{j'j''}^{(eq)} B_{jj'} + D_{jj'}^{(eq)} B_{j'j''} - 2D_{jj'j''}^{(eq)} B_{jj'j''}) \right. \\ &\quad \left. - \sum_{jj'j''} \frac{d_{j-j'}^{(q)} B_{j'j''} + d_{j'-j''}^{(q)} B_{jj'} - 2d_{jj'j''}^{(q)} B_{jj'j''}}{(D_{jj'}^{(eq)} D_{j'j''}^{(eq)} - D_{jj'j''}^{(eq) 2})^{5/2}} \right], \end{aligned} \quad (4.16)$$

where the standard notations $B_{jj'j''} = \sum_q d_{jj'j''}^{(q)} B_q$ have been used.

At this point it is possible to make a crude estimate of the asymptotic behaviour at small q from power counting. The advantage of this is that it exhibits plainly the competing terms that compose the spectrum of modes, and aids in the physical understanding the problem. Using Eqs. (4.15, 4.16, 4.9, 4.10) we obtain,

$$\frac{1}{2}B_q \propto \frac{\Delta \hat{u}^{(2)}}{b^3 N} \underline{q}^{3\nu-2}, \quad (4.17)$$

where $b = \sqrt{k_B T A}$ and ν are parameters from the dominant balance method and for the inverse decay rate, λ_q , up to some numerical constants I, I' , we find,

$$\lambda_q = k_B T b^{-2} \underline{q}^{2\nu+1} - I' |u_f^{(2)}| b^{-5} \underline{q}^{5\nu-1} + I \hat{u}^{(3)} b^{-8} \underline{q}^{8\nu-2}. \quad (4.18)$$

The presence of attractive two-body interactions causes the instability at \underline{q} for values that are smaller than a critical value \underline{q}_c . This value \underline{q}_c may be most easily found in the case of very large $-\Delta \hat{u}_f^{(2)}$

$$\underline{q}_c \propto \left(-\frac{\hat{u}_f^{(2)}}{\hat{u}^{(3)}} b^3 \right)^{5/4}. \quad (4.19)$$

Evidently, the value \underline{q}_{max} at which $|\lambda_q|$ is maximal and the critical value \underline{q}_c are related in this small q limit as follows $\underline{q}_{max} = (5/7)^{5/4} \hat{u}^{(qc)}$.

The few first internal modes with $q < q_c$ are unstable, i. e. $\lambda_q < 0$ for small times when the linear representation (4.14) is valid. Their contribution to the radius of gyration becomes significant at that stage, but for very early times, about $t \leq 1/|\lambda_{max}|$, where λ_{max} is the maximal unstable value, the main change of R_g^2 is due to the first Taylor term in (4.14). Therefore R_g^2 mainly changes due to the large q (stable) modes and we can ignore the contributions of unstable modes $q < q_c$.

Then the functions B_q and λ_q scale for sufficiently large q as,

$$B_q \propto \underline{q}^{3\nu-2}, \quad \lambda_q \propto \underline{q}^{2\nu+1}, \quad (4.20)$$

The deviation of the radius of gyration square $\Delta R_g^2(t)$ may be estimated as,

$$\Delta R_g^2(t) = \sum_{q>q_c} \mathcal{A}_q (1 - e^{-\lambda_q t}) \simeq \frac{N}{2\pi} \int_{\underline{q}_c}^{\infty} d\underline{q} \mathcal{A}_{\underline{q}} (1 - e^{-\lambda_{\underline{q}} t}) \quad (4.21)$$

and using the scalings (4.20) one finds the law for the evolution away from the equilibrium at early times,

$$R_g^2(t) = R_g^2(0) - A_i t^{\alpha_i}, \quad t < 1/|\lambda_{max}| \quad (4.22)$$

$$\alpha_i = \frac{2 - 2\nu}{2\nu + 1} \left(\frac{2 - \nu}{2\nu + 1} \right). \quad (4.23)$$

Here we identify the crossover time for the first stage to be determined by $1/|\lambda_{max}|$. The power law appears basically due to a strong mode coupling after an instantaneous quench.

4.3.2 Linearised GSC equations for late times

For relaxation towards the final equilibrium the linearised equations is,

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \Delta \mathcal{F}_q &= -\Delta V_q(t) \frac{\mathcal{F}_q^{(eq)}}{\zeta_q^{(eq)}} - \frac{k_B T}{\mathcal{F}_q^{(eq)} \zeta_q^{(eq)}} \Delta \mathcal{F}_q(t) \\ \Delta V_q(t) &\equiv \sum_{q'} \frac{\partial V_q^{(eq)}}{\partial \mathcal{F}_{q'}^{(eq)}} \Delta \mathcal{F}_{q'}(t). \end{aligned} \quad (4.24)$$

This is a linear differential equation also involving mode coupling and it may be written as,

$$\frac{d}{dt}\Delta\tilde{\mathcal{F}}_q = -\sum_{q'}\Lambda_{qq'}\Delta\tilde{\mathcal{F}}_{q'}(t), \quad (4.25)$$

Here the new variables are $\Delta\tilde{\mathcal{F}}_{q'} \equiv \sqrt{V_q^{(eq)}\zeta_q^{(eq)}}\Delta\mathcal{F}_q$, so that the matrix $\Lambda_{qq'}$ is symmetric and has the form,

$$\Lambda_{qq'} = \frac{V_q^{(eq)}}{\zeta_q^{(eq)}}\delta_{qq'} + (V_q^{(eq)}\zeta_q^{(eq)}V_{q'}^{(eq)}\zeta_{q'}^{(eq)})^{-1/2} \quad (4.26)$$

$$\frac{5}{2}k_B T \left(\hat{u}^{(2)} \sum_{jj'} \frac{d_{j-j'}^{(q)} d_{j-j'}^{(q')}}{(D_{jj'}^{(eq)})^{7/2}} + \hat{u}^{(3)} \dots \right), \quad (4.27)$$

where we have omitted the very cumbersome expression for the three-body term.

The solution of (4.25) then may be written in the matrix form,

$$\tilde{\mathcal{F}}(t) = \tilde{\mathcal{F}}(\infty) + \exp(-\Lambda t) \bullet A_f, \quad (4.28)$$

with some column A_f . Thus, the main asymptotics at large times is determined by the smallest eigenvalue λ_1 of the matrix Λ .

The analytic form of this eigenvalue can be easily found from the integral equation which is equivalent to for the exact kinetic equation (4.5),

$$\mathcal{F}_q(t) = \frac{k_B T}{V_q(t)} + \left(\mathcal{F}_q(0) - \frac{k_B T}{V_q(t)} \right) \exp \left(-2 \int_0^t dt' \frac{V_q(t')}{\zeta_q(t')} \right). \quad (4.29)$$

Thus, by taking into account that the system is close to the final equilibrium we obtain,

$$\lambda_1 = \frac{1}{\tau_f} = \frac{2k_B T}{\mathcal{F}_1^{(eq)}\zeta_1^{(eq)}}. \quad (4.30)$$

Therefore, the law for relaxation towards the equilibrium at large times is,

$$R_g^2(t) = R_g^2(\infty) + A_f \exp(-t/\tau_f), \quad \tau_f \sim N^{\gamma_f}, \quad (4.31)$$

$$\gamma_f = \nu(3 - \delta_{\nu,1/3}) \left(1 + \nu(2 - \delta_{\nu,1/3}) \right),$$

where we have taken into account the above scalings for $\mathcal{F}_1^{(eq)}$ and $\zeta_q^{(eq)}$.

The numerical values of the main kinetic exponents corresponding to the kinetic laws (4.22,4.31) are summarised in Tab. 4.3.

Table 4.3: The main kinetic exponents of the homopolymer in near equilibrium regimes deduced analytically from the GSC method.

| conform. state | with hydrodynamics | | without hydrodynamics | |
|------------------------|--------------------|----------------------|-----------------------|----------------------|
| | α_i (away) | γ_f (towards) | α_i (away) | γ_f (towards) |
| Flory $\nu = 3/5$ | 9/11 | 9/5 | 7/11 | 11/5 |
| Ideal $\nu = 1/2$ | 1 | 3/2 | 3/4 | 2 |
| Globule $\nu = 1/3$ | 7/5 | 2/3 | 1 | 4/3 |

4.3.3 Kinetics of folding

In this subsection we consider in more detail kinetics after a quench from the Flory coil ($u_i^{(2)} > 0$) to the collapsed globule ($u_f^{(2)} < 0$) since this folding process is of primary importance in many applications. We shall also compare the predictions of the GSC theory, using both the analytical linearised and the full numerical treatments, with data from Monte Carlo simulations.

The folding process can be divided into a number of distinct kinetic stages.

1. Necklace formation stage

The earliest kinetic stage is characterized by rapid formation of numerous small collapsed globules [II, 105, 108, 170]. This process is essentially local and its duration τ_i is quite independent of the degree of polymerisation. Formally, we can define it as the time τ_i at which the critical mode $\mathcal{F}_{q_c}(t)$ reaches its minimum.

An example of particular polymer conformations during this stage is drawn in Figs. 4.13, 4.17. Pictures 4.13a, 4.17a correspond to an equilibrium polymer conformation in the Flory coil state. One can see numerous small and comparatively low density accumulations of monomers present in that state and after the

Figure 4.13: Evolution of a particular polymer conformation during kinetics at the collapse transition (a, b).

The degree of polymerisation, the linear lattice size and quench are equal to $N = 1000$, $L = 150$ and $\chi_i = 0 \rightarrow \chi_f = 0.36$ respectively. Here and in Figs. 4.14, 4.15 and 4.16 pictures (a)-(h) correspond respectively to the following moments in time: $t = 0$ (before the quench), $t = 5,000$ MCS (approximately the end of the first kinetic stage), $t = 30,000$ MCS, 130 KMCS, 230 KMCS, 420 KMCS, 530 KMCS and 635KMCS.

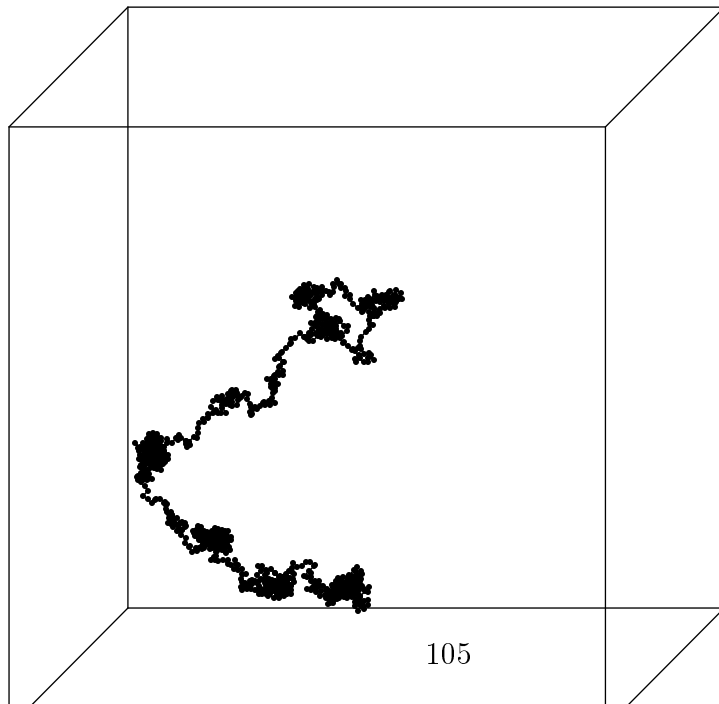
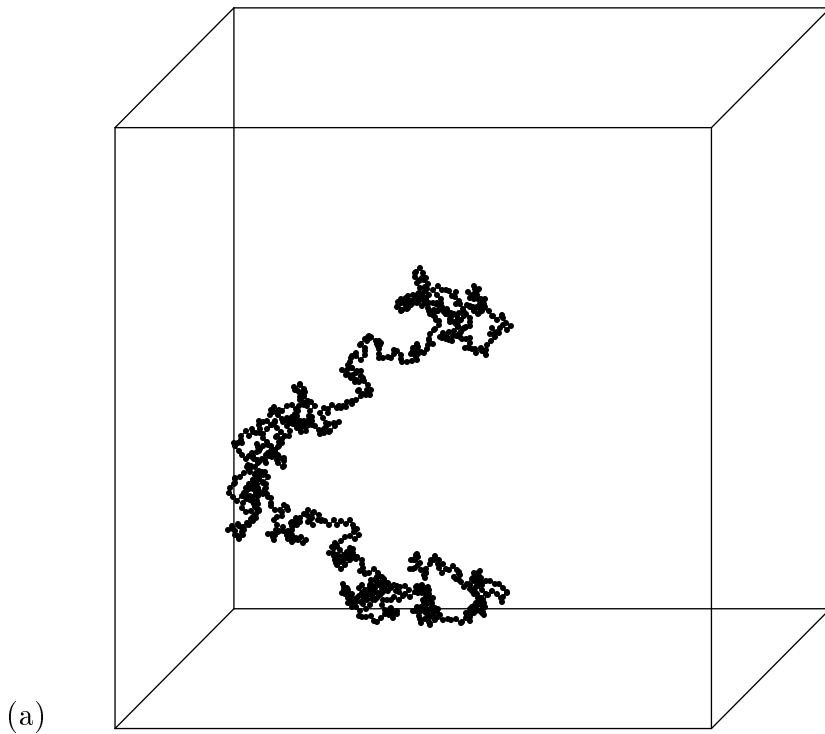


Figure 4.14: Evolution of a particular polymer conformation during kinetics at the collapse transition (c, d).

See caption of Fig. 4.13 for more details.

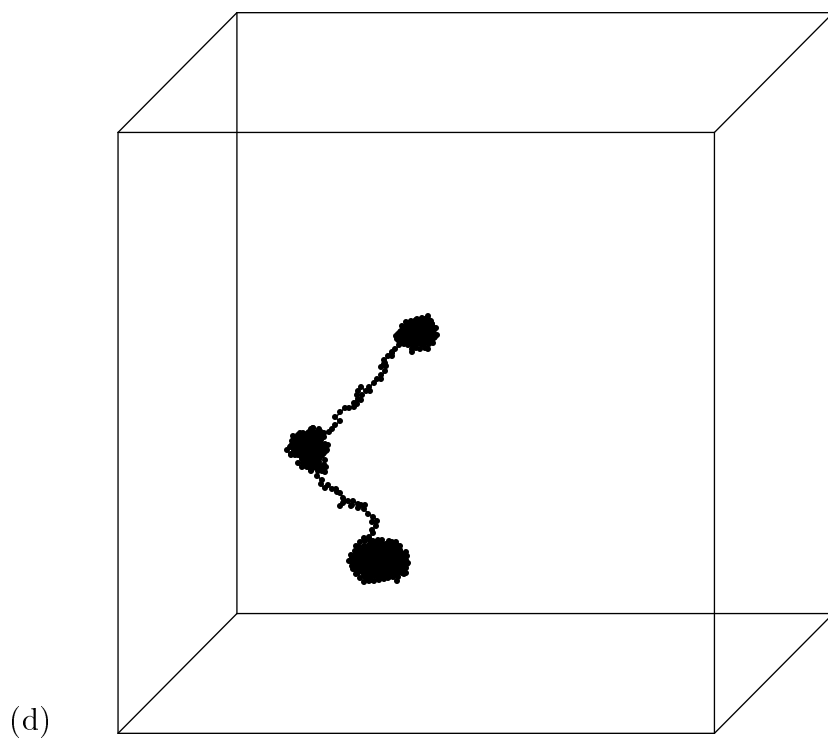
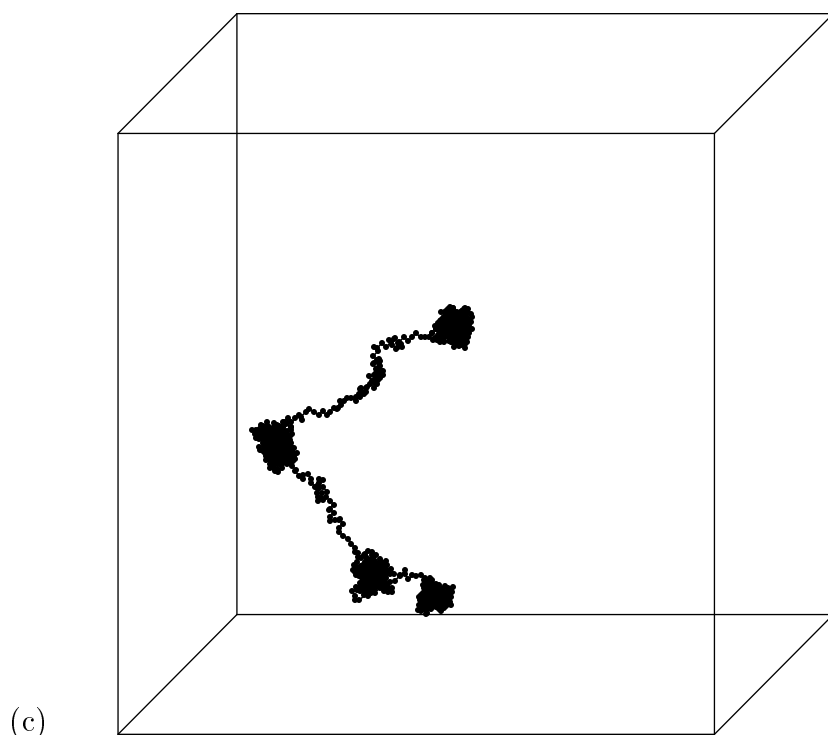


Figure 4.15: Evolution of a particular polymer conformation during kinetics at the collapse transition (e, f).

See caption of Fig. 4.13 for more details.

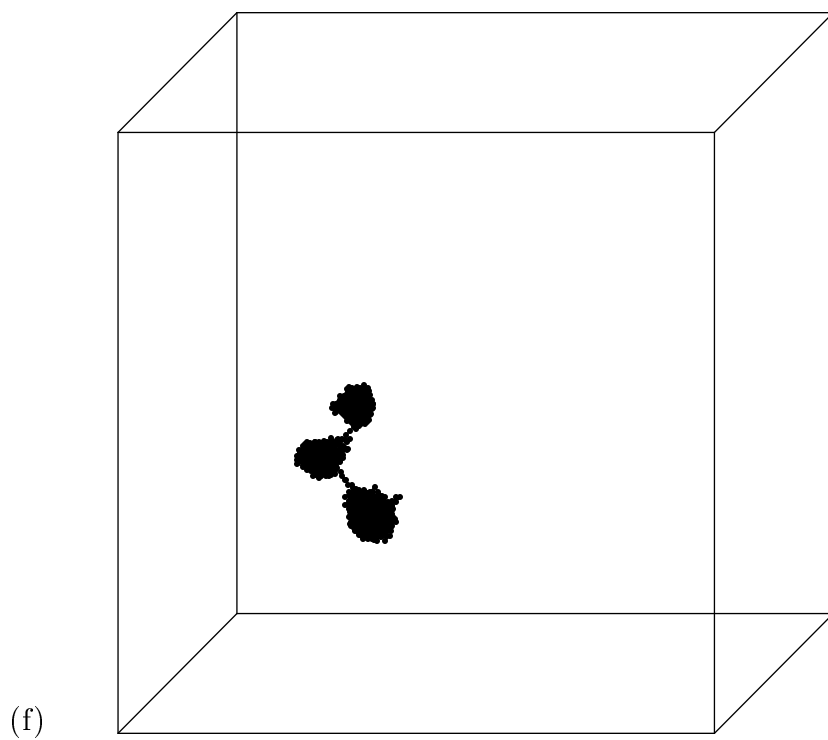
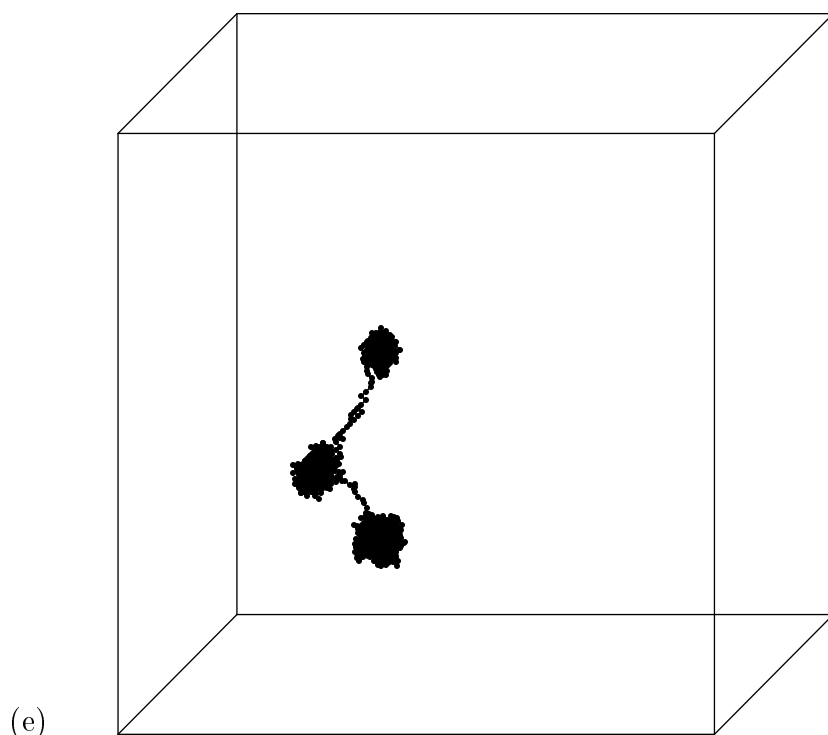


Figure 4.16: Evolution of a particular polymer conformation during kinetics at the collapse transition (g, h).

See caption of Fig. 4.13 for more details.

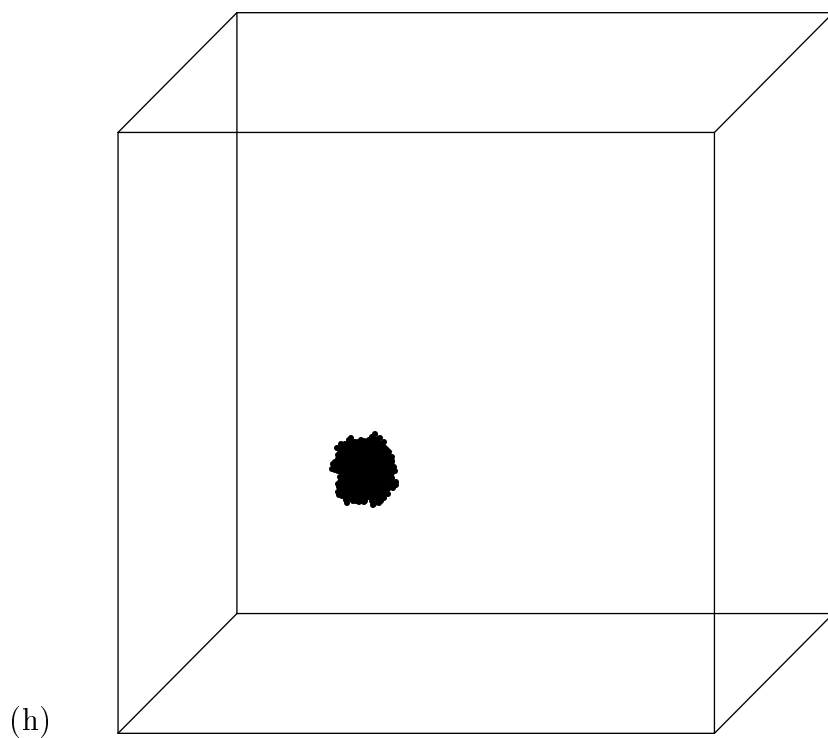
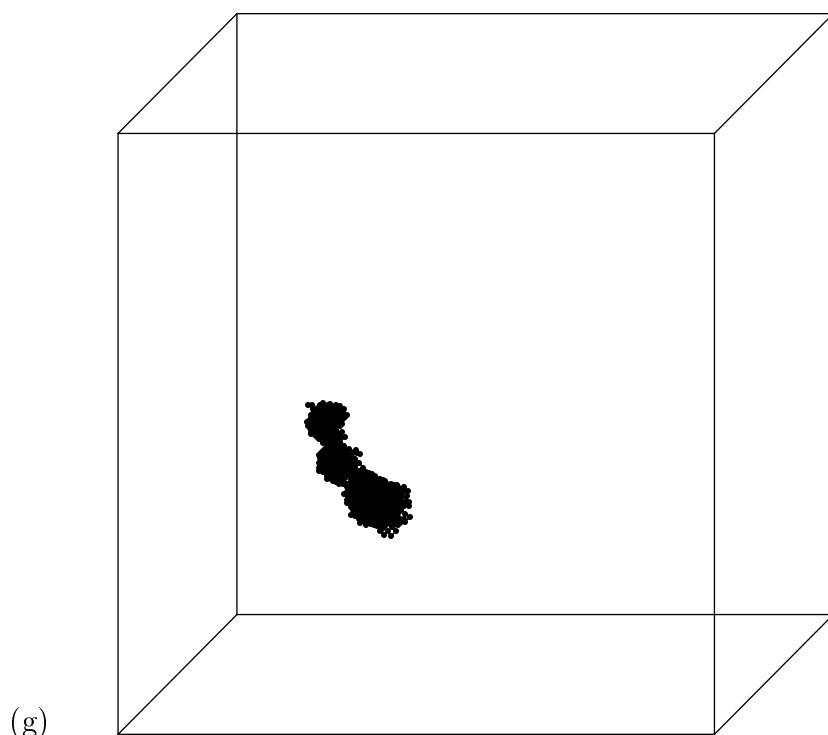


Figure 4.17: Evolution of a particular polymer conformation during kinetics at the collapse transition (a, b).

The degree of polymerisation, the linear lattice size and the quench here are the same as in Figs. 4.13-4.16. Here and in Figs. 4.18, 4.19 and 4.20 pictures (a)-(h) correspond respectively to the following moments in time: $t = 0$, 5 KMCS, 20 KMCS, 90 KMCS, 105 KMCS, 300 KMCS, 470 KMCS and 505KMCS.

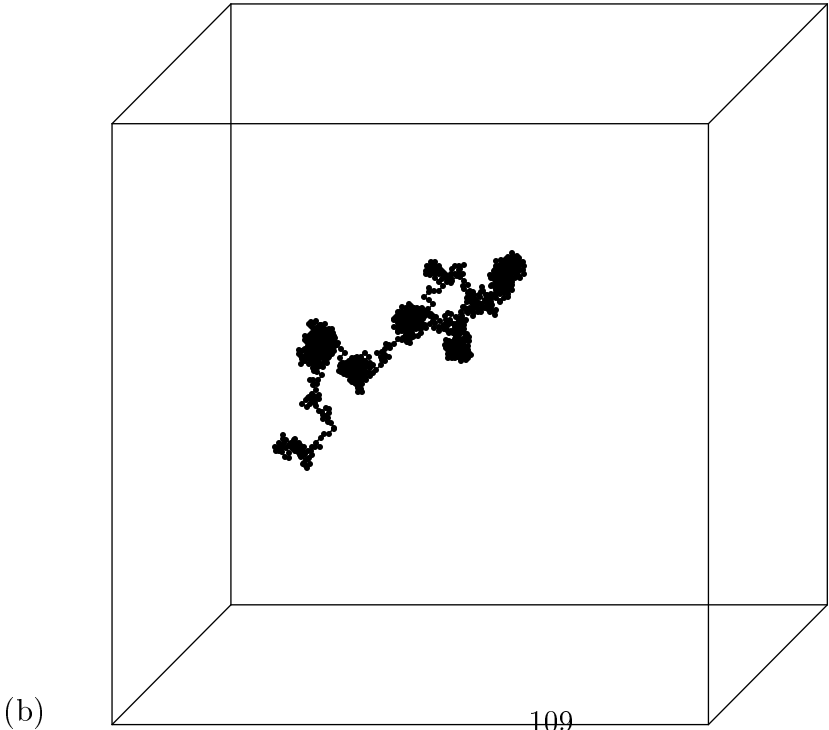
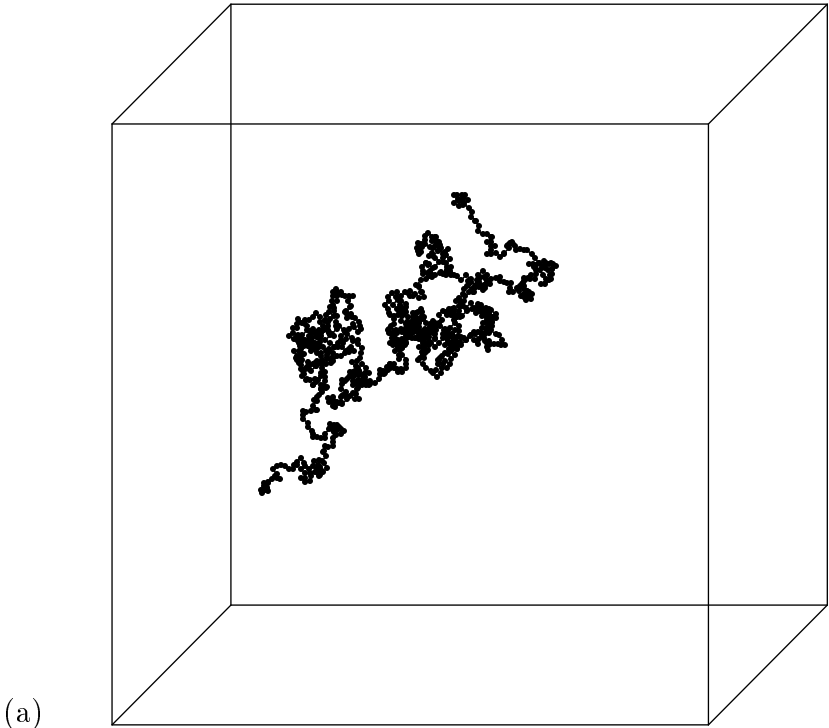


Figure 4.18: Evolution of a particular polymer conformation during kinetics at the collapse transition (c, d).

See caption of Fig. 4.17 for more details.

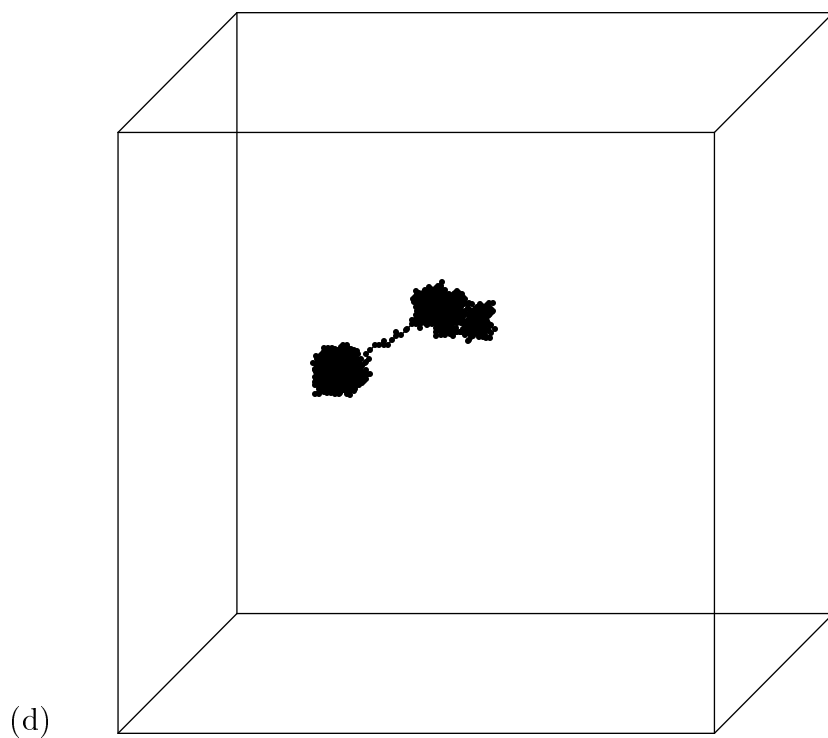
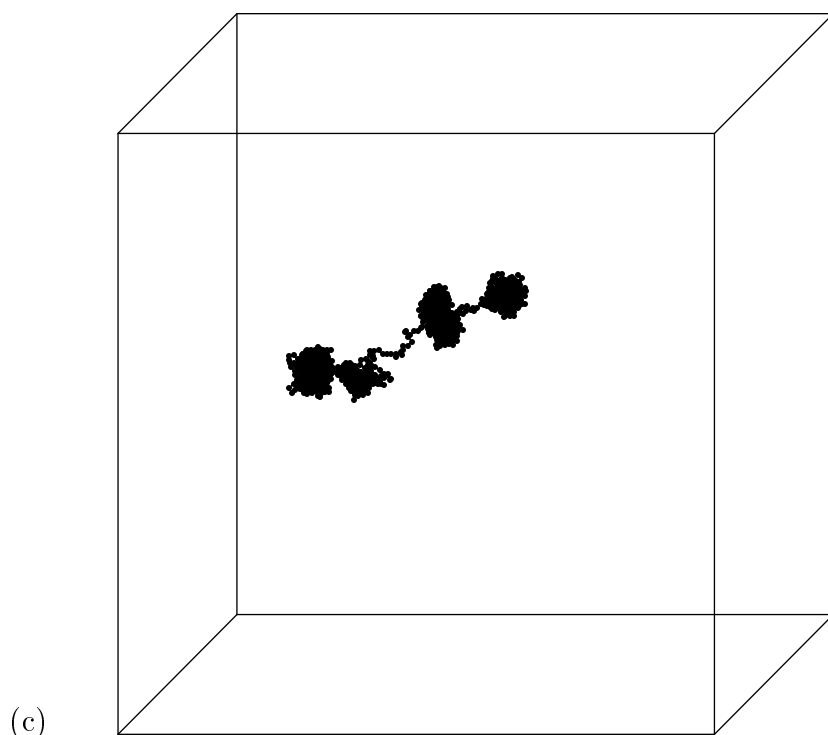


Figure 4.19: Evolution of a particular polymer conformation during kinetics at the collapse transition (e, f).

See caption of Fig. 4.17 for more details.

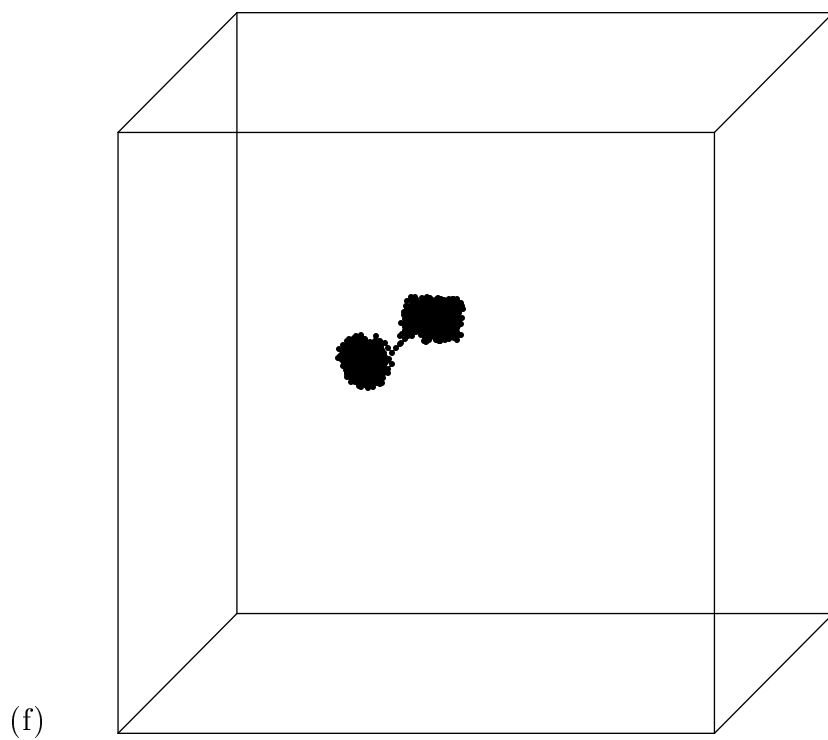
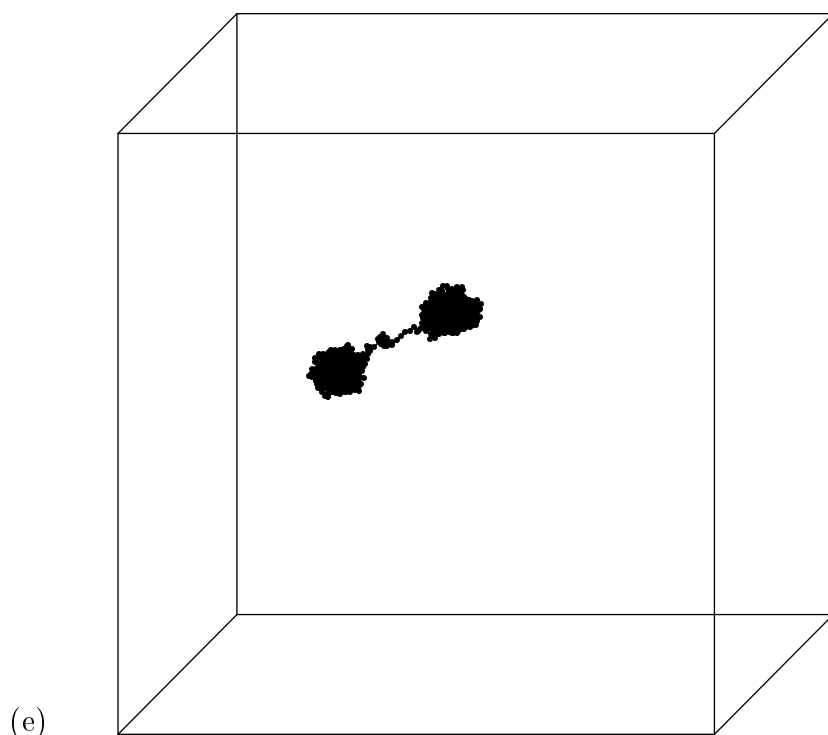


Figure 4.20: Evolution of a particular polymer conformation during kinetics at the collapse transition (g, h).

See caption of Fig. 4.17 for more details.

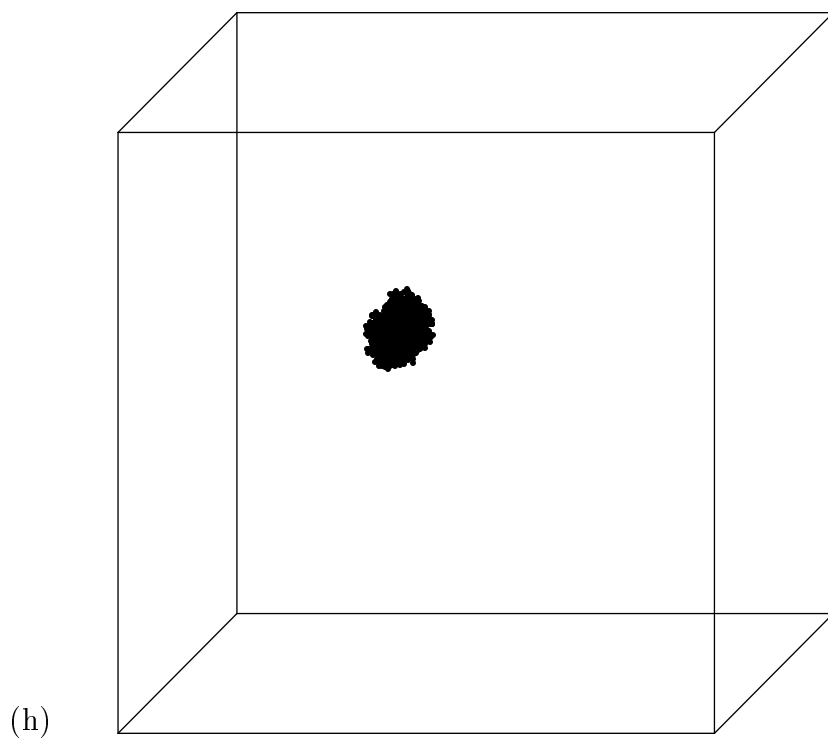
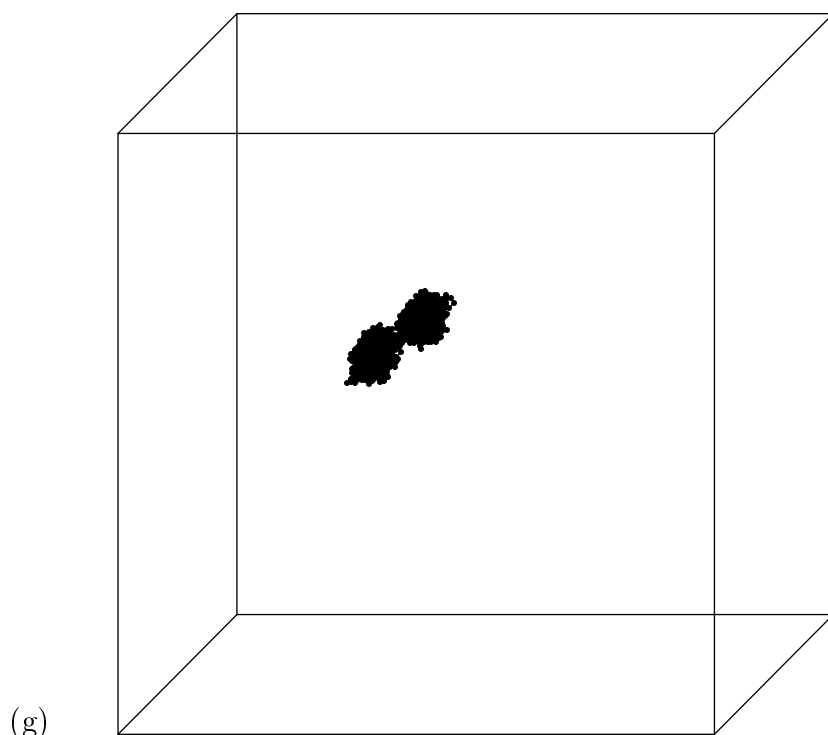


Table 4.4: Numerical values of the exponent α_i in (4.32) vs the degree of polymerisation, N .

This exponent is obtained by a fitting of the change of the mean squared radius of gyration during the first 100 MCS.

| | | | | | | | |
|------------|------|------|------|------|------|------|------|
| N | 96 | 160 | 256 | 384 | 512 | 768 | 1024 |
| α_i | 0.65 | 0.64 | 0.69 | 0.66 | 0.64 | 0.67 | 0.66 |

quench these become sites for the formation of locally collapsed globules. This situation is depicted in Figs. 4.13b, 4.17b and conformations there correspond, approximately, to the end of the first kinetic stage.

The first kinetic stage corresponds to a rapid initial growth of the average number of clusters $\langle n \rangle$ and average number of monomers in a cluster S . It may be observed in Figs. 4.13b, 4.17b, that after the earliest stage of kinetics, the polymer conformation contains several quite distinguishable clusters.

Now, the mean squared radius of gyration decreases very sharply during the earliest kinetic stage and can be approximated by the formula of type (4.22),

$$R_g^2(t) = R_0^2 - At^{\alpha_i}, \quad \alpha_i = 0.66 \pm 0.03, \quad (4.32)$$

for polymers of chain lengths about 100-1,000. Some particular values of the exponent α_i for different degrees of polymerisation are given in Table 4.4. The theoretical value for this exponent $\alpha_i = 7/11 \approx 0.636$ (see Tab. 4.3) is in quite good agreement with simulation results.

Thus, during this early stage, the main change of the radius of gyration arises from the large- q internal modes in (2.91), even though their amplitude is a decreasing function of q . The small- q modes hardly change at early time [122]. In this description [116] the high- q modes decrease exponentially at small time. In any case, we can see that the decrease in amplitude of high- q modes corresponds to shrinking of the polymer at small distances, i.e. formation of small dense clusters.

Clearly, for rather deep quenches the contribution of the unstable modes to the mean squared radius of gyration becomes significant and the power law (4.32) would be violated.

At the end of this stage most of the slack polymer has been gathered up into small more condensed clusters with relatively tight chain between them (see Figs. 4.13b, 4.17b). The effective Flory exponent has risen to be unity, characterising a rigid rod of clusters, and thereafter begins to fall rapidly until it reaches a near-ideal Gaussian value.

2. Coarsening stage

The second stage involves cluster coarsening, the analog of Lifshitz–Slyozov cluster growth after a quench into the liquid–gas coexistence region.

In series of pictures in Figs. 4.14, 4.15, 4.18 and 4.19 we present typical conformations from this stage of kinetics, where the initial conformations were in the Flory coil state of Figs. 4.13a, 4.17a. From the simulation it is clear that growth of clusters proceeds mainly by their unification with smaller ones. Polymer segments linking different globules are quite tight and accretion of monomer units from such segments by the neighbouring globules gives a much less significant contribution. After a few clusters are left, e.g. three clusters in Fig. 4.14d and two clusters in Fig. 4.18d, the following evolution is characterised by their approach to each other due to the force from linking segments. However, we emphasise that this chain of clusters is essentially an ideal coil at long lengths with a set of topologically connected locally collapsed clusters. The evidence is that the average effective swelling exponent during this stage corresponds to an ideal coil $\nu_T = 1/2$, and we therefore have a set of clusters growing against the tension of an ideal coil elasticity. The characteristic time expected from elementary arguments

Table 4.5: Values of the characteristic collapse times τ_m ($c_3 = 0$), τ_t ($c_3 = 0$) and τ_f vs the degree of polymerisation, N , for different viscosities, η_s .

In this table $u_i^{(2)} = 15$, $u_f^{(2)} = -25$ and $u^{(3)} = 10$. Here the quantity γ denotes the exponent of the appropriate time in terms of the degree of polymerisation, i.e. $\tau \sim N^\gamma$. Note that the limit $\eta_s = \infty$ corresponds to the absence of hydrodynamics, whilst $\eta_s = 0.1$ describes the regime of strong hydrodynamic effect, and $\eta_s = 0.5$ corresponds to the crossover region.

| N | 30 | 50 | 70 | 100 | 150 | 200 | 300 | γ |
|--------------------------------|------|------|------|------|------|-------|------|-----------------|
| τ_m ($\eta_s = \infty$) | 1.8 | 4.4 | 8.3 | 16.6 | 38 | 70 | 156 | 2.00 ± 0.03 |
| τ_m ($\eta_s = 0.5$) | 1.4 | 3.2 | 5.6 | 10.2 | 20.7 | 34.5 | 71.7 | 1.71 ± 0.03 |
| τ_m ($\eta_s = 0.1$) | 0.93 | 1.4 | 2.6 | 4.2 | 7.5 | 11.8 | 22.3 | 1.48 ± 0.04 |
| τ_t ($\eta_s = \infty$) | 3.9 | 10.5 | 20.3 | 40.8 | 90.7 | 160.3 | 357 | 1.96 ± 0.01 |
| τ_t ($\eta_s = 0.5$) | 3.0 | 7.1 | 12.0 | 22.0 | 42.6 | 67.6 | 129 | 1.63 ± 0.02 |
| τ_t ($\eta_s = 0.1$) | 1.7 | 3.2 | 5.0 | 8.1 | 14.0 | 21.0 | 37.3 | 1.34 ± 0.03 |
| τ_f ($\eta_s = \infty$) | 0.73 | 1.41 | 2.66 | 4.25 | 7.34 | 9.92 | 16.8 | 1.27 ± 0.05 |

is,

$$\tau_m = AN^{3\nu_T}, \quad (\tau_m = A'N^{2\nu_T+1}) \quad (4.33)$$

for the model with (without) hydrodynamic interaction. Thus, we get the cluster growth law without hydrodynamics,

$$S(t) = A_s t^Z, \quad Z = 1/2. \quad (4.34)$$

In other words, τ_m scales as the Zimm (Rouse) relaxation time. In this sense, one can say that the coarsening stage of collapse is dominated by the behaviour of the ideal Gaussian coil. This is in agreement with the phenomenological theory of de Gennes [127], and with our Monte Carlo calculations [II].

Alternatively, τ_m can be defined as the time when the second internal mode \mathcal{F}_2 reaches its minimum amplitude. In Tab. 4.5 we present values of τ_m for different degrees of polymerisation and viscosities obtained from the numerical solution of the GSC equations (4.5).

This is the predominant stage of kinetics, for beyond this stage the polymer coil is space-filling, though not yet compacted.

3. Shape optimisation stage

The final relaxation is characterised by a slow exponential relaxation with the time-scale,

$$\tau_f \sim N^{2/3} \quad (N^{4/3}). \quad (4.35)$$

It corresponds to a process of shape optimisation and further compaction of the globule. This process is presented in Figs. 4.16 and 4.20.

Clearly it is, strictly speaking, impossible to distinguish precisely between the second and the final stages. Let us introduce what we call the “total” collapse time, τ_t , as the time, at which a single exponential approximation begins to work reliably well for the radius of gyration. For example, for the chain lengths we have considered, this may be chosen as that time, when the squared radius of gyration has passed through 99 per cent of its overall change: $R_g^2(\tau_t) = 0.01 R_g^2(0) + 0.99 R_g^2(\infty)$. In Tab. 4.5 we show the final relaxation time τ_f determined from the mean squared radius of gyration $R_g^2(t) = R_g^2(\infty) + A_f e^{-t/\tau_f}$, or equivalently the time-scale deduced from the first internal mode $\mathcal{F}_1(t)$. Thus, we find $\tau_f \sim N^{\gamma_f}$, where $\gamma_f = 2/3$ ($\gamma_f = 4/3$) with (without) hydrodynamics. Naturally, the “total” collapse time should obey a bound $\tau_t > \tau_m$, and scaling of the exponent γ_t , where $\tau_t \sim N^{\gamma_t}$, is some cross-over between the exponents γ_m and γ_f (see Tab. 4.5) with $\gamma_t \rightarrow \gamma_m$ when $N \rightarrow \infty$. Thus, for sufficiently long chains, the main collapse takes a finite time about $\tau_t \sim N^{\gamma_m}$, and then the process is exponentially slow with the time characterized by universal exponent γ_f .

3. Reptational stage

As we have already mentioned in the Introduction some collective polymer moves, such as the “slithering snake” (reptational) motions [25] and the pivot motions [26, 27], do not correspond to any realistic chain motion and, strictly speaking, cannot be used in dynamics and kinetics. However, they can be applied when

considering the system near equilibrium on a much longer time scales.

Due to topological restrictions the structure of the globule after the late stages of kinetics for a non-phantom chain is different than that for a phantom one. For an open polymer topological restrictions may be removed via self-reptations of the chain, and it has been argued that this leads to an even longer final kinetic stage with the time scale $\tau_{reptations} \sim N^3$ [171]. The resulting structural changes in the globular state due to applying reptations are discussed in more details in Sec. 4.4.

4.3.4 Specific features in kinetics of folding of an open chain

Here we generalise our treatment of Chapter 2 to describe both ring and open polymers in the GSC method. Thus, let us consider in parallel two homopolymers, one of which is a ring constituted of N_r beads and another one is open with N_o beads. For a ring polymer one has cyclic boundary condition $\mathbf{X}_{j+N_r} = \mathbf{X}_j$ for all n , whilst there is no special restriction on the conformations of open polymer. However, to treat the open polymer it is useful to formally introduce the fictitious beads with the chain indices $j = -1$ and $j = N_o$, such that their positions satisfy the discreet Neumann boundary condition $\mathbf{X}_{-1} = \mathbf{X}_0$ and $\mathbf{X}_{N_o} = \mathbf{X}_{N_o-1}$.

We begin with a mathematical question about the proper definition of the independent modes. Let us introduce the following notations:

$$\begin{array}{ll}
 \text{ring} & \text{open} \\
 f_j^{(q)} = \exp(i2\pi qj/N_r), & f_j^{(q)} = \cos(\pi q(j + 1/2)/N_o), \\
 \varepsilon^{(q)} = -4 \sin^2(\pi q/N_r), & \varepsilon^{(q)} = -4 \sin^2(\pi q/2N_o), \\
 \gamma_q = 1, & \gamma_q = 2 - \delta_{q0}.
 \end{array} \tag{4.36}$$

Then the Fourier transform may be defined in the general form

$$\mathbf{x}_q = \frac{\gamma_q}{N} \sum_{j=0}^{N-1} f_j^{(q)} \mathbf{X}_j, \quad \mathbf{X}_j = \sum_{q=0}^{N-1} f_j^{(-q)} \mathbf{x}_q, \quad (4.37)$$

where $N = N_r, N_o$ for ring and open polymer respectively. Indeed, $f_j^{(q)}$ may be treated as eigenfunctions and as $\varepsilon^{(q)}$ eigenvalues of the discrete second derivative operator Δ ,

$$\Delta f_j^{(q)} \equiv f_{j+1}^{(q)} + f_{j-1}^{(q)} - 2f_j^{(q)} = \varepsilon^{(q)} f_j^{(q)} \quad (4.38)$$

with corresponding boundary conditions. Thus, the set $\{\sqrt{\frac{\gamma_q}{N}} f_j^{(q)}\}$ is a complete and orthonormal basis. For ring polymers this basis is well known, whilst for open polymers its properties can be easily established using the relation connecting it with the Chebyshev polynomials

$$f_j^{(q)} = U_j(f_1^{(q)}) - U_{j-1}(f_1^{(q)}), \quad (4.39)$$

$$U_j(x) = (1-x^2)^{-1/2} \sin((j+1) \arccos x). \quad (4.40)$$

There is a fundamental distinction between ring and open polymers due to topology. This however is irrelevant as long as we consider only phantom chains. The Gaussian self-consistent method is applicable to open polymers as well as to rings. It is well established that, at equilibrium, the values of the universal exponents are different in the middle of the chain and near the ends [37]. The difference however is rather small and could not be deduced at the mean-field level, which our approach actually represents. Thus, in the GSC method it is sufficient to keep only the diagonal elements of the effective potential for sufficiently long chains. Physically, this is equivalent to the property that the spatial correlations $\langle(\mathbf{x}_j - \mathbf{x}_{j'})^2\rangle$ depend only on the difference $|j - j'|$. It is known that this is valid for finite chains when $|j - j'|$ is sufficiently large. In other words, neglect of the non-diagonal elements affects only properties of the chain at short distances and does not change any of its global characteristics such as the radius of gyration. The bigger the chain length the better the precision of the diagonal method.

Bearing in mind these reservations, the self-consistency equations (4.5) remain valid for both cases. However, formulae for the friction, ζ_q , and connection (2.94) between the spatial correlations and the Fourier modes should be generalised,

$$\zeta_q = \frac{N}{\gamma_q} \zeta_b, \quad (4.41)$$

$$D_{jj'} = \sum_{q=1}^{N-1} d_{jj'}^{(q)} \mathcal{F}_q, \quad d_{jj'}^{(q)} = |f_j^{(q)} - f_{j'}^{(q)}|^2. \quad (4.42)$$

Similarly, the squared radius of gyration is given by,

$$R_g^2 = \frac{1}{\gamma_1} \sum_{q=1}^{N-1} \mathcal{F}_q. \quad (4.43)$$

The crucial observation is that the spatial correlations for open polymer may be decomposed,

$$D_{jj'} = \frac{1}{2} D_{|j-j'|} + \frac{1}{2} C_{jj'}, \quad (4.44)$$

$$D_{|j-j'|} = 4 \sum_{q=1}^{N_o-1} \sin^2 \frac{2\pi q(j-j')}{N_o} \mathcal{F}_q, \quad (4.45)$$

$$C_{jj'} = \sum_{q=1}^{N_o-1} \left(2 \cos \frac{2\pi q(j+j'+1)}{N_o} - \cos \frac{2\pi q(2j+1)}{N_o} \right. \quad (4.46)$$

$$\left. - \cos \frac{2\pi q(2j'+1)}{N_o} \right) \mathcal{F}_q. \quad (4.47)$$

The first term $D_{|j-j'|}$ is analogous to the expression for the spatial correlation of ring polymer. The sum in the second term is rapidly oscillating and for large N_o is negligible. Hence at the level of our mean-field approximation open and ring chains become equivalent for sufficiently large degrees of polymerisation.

Note that for ring polymers only half of the Fourier modes are independent due to the relation $\mathcal{F}_{N_r-q} = \mathcal{F}_q$. This, together with explicit formulae (4.36, 4.36), allows one to establish a connection between ring and open polymers for large degrees of polymerisation satisfying condition $N_r = 2N_o$. Given that virial coefficients are in the relations,

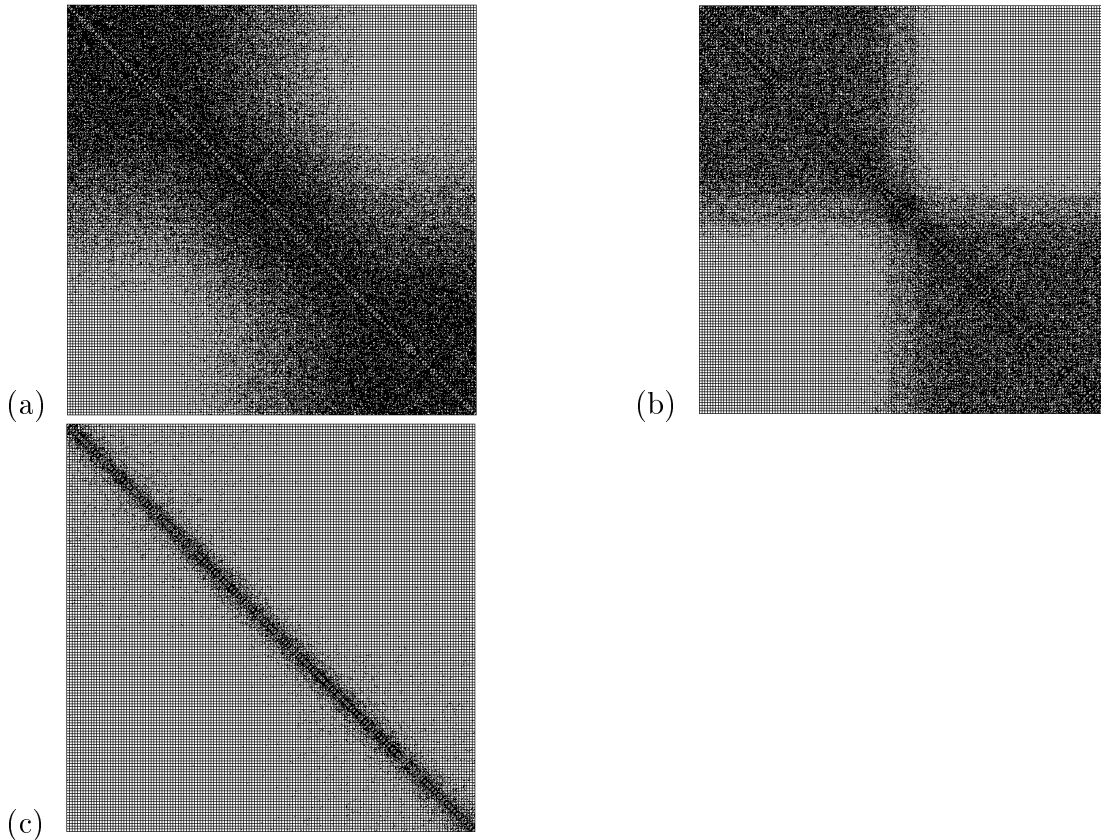
$$u_{ring}^{(J)} = 2^{1-J} u_{open}^{(J)}, \quad (4.48)$$

the radii of gyration and spatial correlations of both chains will be equal. This statement is valid at equilibrium as well as for kinetics. It generalises the well known observation that ideal ring chains are more compact than open ones with the equal number of units [38], $R_g^{ring} = R_g^{open} / \sqrt{2}$. As we have seen it is more natural to compare ring chains with twice as many units as open ones and appropriately redefined virial coefficients (4.48).

Thus, for large N the GSC equations for an open homopolymer become equivalent to those of a ring with $N/2$ monomers and rescaled virial coefficients. However, this is merely an approximation since, strictly speaking, we can not use the set of diagonal Fourier modes, but have to include the set of all D_{ij} . We would like to briefly discuss here the differences in the folding kinetics between rings and open polymers which we could study in such a more accurate treatment. Formally, we only need to remove the spring between the first and the last monomers in the effective free energy functional (2.14) to describe an open chain.

In Fig. 4.21 we present the typographic map of the D_{ij} matrix at various times during folding kinetics. Starting from some moment deeply inside the middle stage one clearly observes that the ends of the chain become quite different from the middle. Namely, in Fig. 4.21b at long times after the quench there are two distinct end clusters and the conformation here reminds that of a dumbbell. The connecting link slowly shrinks until it eventually produces a normal spherical globule in Fig. 4.21c. This is quite consistent with results from the Monte Carlo simulation presented in Figs. 4.13-4.20, where about 3/4 of the nonequilibrium conformations at late stages take form of a dumbbell as in Figs. 4.18d and 4.19e, f. Such effects are present for all values of the viscosity at sufficiently late times. However, the end clusters grow with about the same rate as the middle ones during the first and a considerable part of the middle kinetic stages. Interestingly enough, we find that the collapse time of an open chain is about 2-2.5 times longer

Figure 4.21: Diagrams of matrix of the mean squared distances, D_{ij} , for the open homopolymer with $N = 150$, in kinetics after the quench, $u^{(2)} = 30 \rightarrow -50$. Here $\eta_s = 0.1$. Diagrams (a-c) correspond respectively to the following moments in time: $t = 27, 52$ and 73 . Indices i, j start counting from the upper left corner. Each matrix element, D_{ij} is denoted by a quadratic cell with varying degree of black colour, the darkest and the lightest cells corresponding respectively to the smallest and to the largest mean squared distances. The diagonal elements are not painted since by definition, $D_{ii} = 0$.



than for the equivalent ring for chain lengths that we have studied.

Thus, although one may say that the ends play a special role in folding due to their higher freedom to move, the process of folding is not sped up by that, but is slowed down. This can be simply interpreted by observing an evolving polymer chain in the Monte Carlo simulation presented in Figs. 4.13-4.20. Indeed, the chain is very tight during the middle stage when the end domination becomes significant. A ring chain pulls itself uniformly towards the centre thus helping the chain to collapse into a ball. On the contrary, an open chain has free end clusters which tend to be far from each other and it takes quite a long time for them to come into a contact and coalesce eventually.

For considered system sizes we have also found that the kinetic exponents for an open homopolymer obtained by fitting remain in good agreement with the scaling laws discussed above for rings. This implies that the end effects do not change the dominant collapse mechanism — midchain clusters grow mostly on the expense of absorbing the slack and neighbouring clusters along the chain.

4.4 Structure of the Homopolymer Globule

As we have already discussed for simulating dynamic or kinetic phenomena on rather short timescales of order of nanoseconds and microseconds using the Monte Carlo technique one cannot apply all set of moves. Instead, only those moves which are consistent with the dynamic behaviour of the system must be applied. In case of polymers these include local monomer moves and moves of groups of strongly interacting monomers, i.e. clusters. The latter should be performed with their probability appropriately lowered. Since the kinetics of polymer collapse proceed through formation and growth of locally collapsed clusters along the chain, with their final unification into a single globule, one may expect that

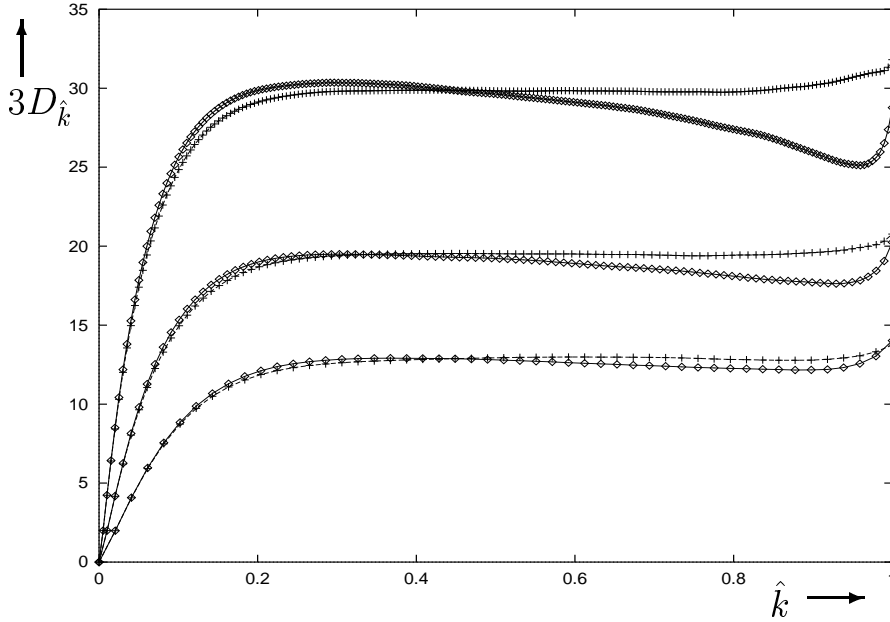
after the shape optimisation stage (see Subsec. 4.3.3) the homopolymer globule possesses a comparatively simple topological structure. Indeed, typical coil conformations before the quench possess a statistically small number of entanglements, or knots, and the folding kinetics adds virtually nothing to that quantity. Further relaxation of the globule towards the equilibrium requires participation of the chain ends and, thus, can be viewed as auto-reptation. This auto-reptation time can be estimated as, $\tau_{rept} = \tau_0 N^3$ [38], which yields a time of order 10^3 sec under usual experimental conditions [171].

In this section we study the structural differences between the two globules of the open homopolymer: one with relatively simple topological structure corresponding to the late stage of folding kinetics, and another with topological entanglements. In practice these simulations have been carried out in the following way. Using the lattice Monte Carlo method a large set of homopolymer globules was produced by independent kinetic processes starting from initial coil conformations. Since reptational moves were not included during this simulation the globules in this set correspond to the late stage of the folding kinetics. To produce entangled globules an additional simulation was applied to the set with some amount of reptational moves, typically $p_r = 1/N$, with N being the degree of polymerisation. For considered chain lengths the typical number of statistically independent globules in each set was about 15,000-30,000.

It is important to note here that study of this problem requires a model in which the polymer chain integrity is not violated, i.e. the polymer is strictly nonphantom — the property automatically provided by the lattice Monte Carlo model (see Fig. 3.2 and explanations below). Otherwise, chain entanglements would be produced during polymer collapse and any excess of knots would be destroyed by nonphantom moves during the reptational stage of simulation. In the extreme case when a typical time required for the chain to pass through

Figure 4.22: Plot of the partially summated mean squared distances, $D_{\hat{k}}$, vs the normalised chain index, $\hat{k} = k/N - 1$, for homopolymer globules of different sizes and different simulation procedures.

Pairs of lines correspond respectively to the following values of the degree of polymerisation (from bottom to top): $N = 50$, $N = 100$ and $N = 200$. Lines denoted by diamonds and pluses in each pair correspond to the simulation procedures with, $p_r = 1/N$, and without reptations, $p_r = 0$, respectively.



itself is much smaller than the time needed to produce a knot by reptations two different simulation procedures would produce no difference at all.

To study the difference one needs to consider some simple observables. First, let us introduce the partially summated mean squared distances along the chain,

$$D_k \equiv \frac{1}{3(N-k)} \sum_{i=0}^{N-1-k} \langle (\mathbf{X}_i - \mathbf{X}_{i+k})^2 \rangle, \quad (4.49)$$

as a function of the chain index, i.e. number of polymer links in between, k . Another observable we shall consider here is the number of loops L_k defined by Eq. (4.2). For comparing polymers with different degrees of polymerisation it is convenient to introduce the mean squared distanced, $D_{\hat{k}}$, and the number of loops, $L_{\hat{k}}$ as functions of the normalised chain index, $\hat{k} \equiv k/(N-1)$, which changes in the range, $0 \leq \hat{k} \leq 1$, and has a meaning of the fraction of the number

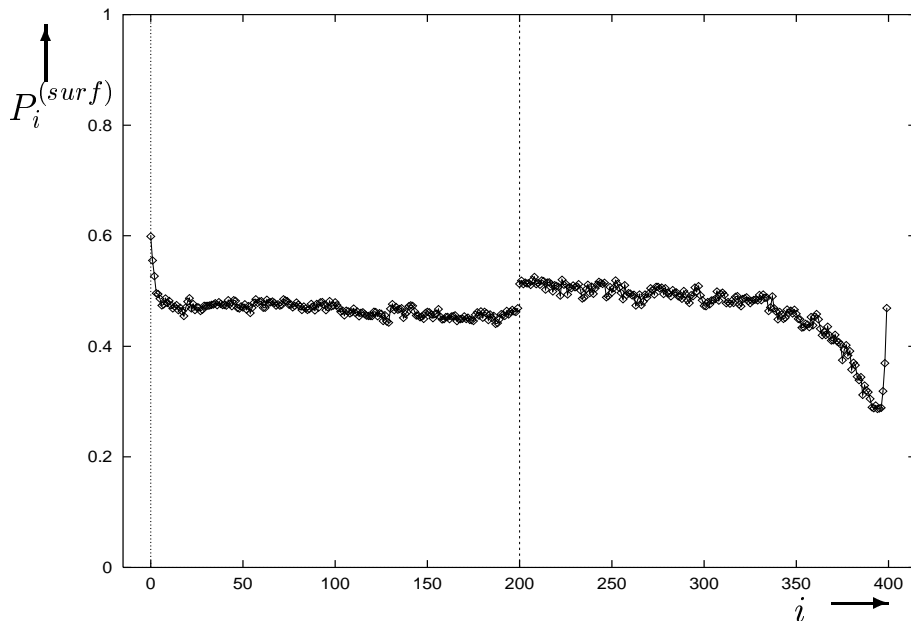
of links, k , in the polymer which has $N - 1$ links.

In Fig. 4.22 we present the mean squared distances, $D_{\hat{k}}$ versus the normalised chain index, \hat{k} , for polymers of different degrees of polymerisation. For very small chain indices function $D_{\hat{k}}$ does not depend on N and on the particular simulation procedure, which reflects local packing of monomers in the dense globule. Let us consider first the behaviour of this function for globules prepared without reptations. For small values of the chain index the function is almost linear, up to some cross-over value that scales as $N^{2/3}$. Then it saturates to some level, which is proportional to the radius of the globule and, thus, also scaling as $N^{2/3}$. Interestingly enough, for values of the chain index in the vicinity of the chain ends the function $D_{\hat{k}}$ increases once again. This phenomenon actually reflects the mechanism of the polymer collapse during late coarsening stage. The globule is usually formed by a final unification of two end clusters (see Fig. 4.20) and, sometimes, a few middle clusters (see Fig. 4.16). Thus, the chain ends possess a somewhat higher probability to appear on opposite sides of the globule than the rest of monomers. This effect in the mean squared distances is not very strong as the function experiences about 10% increase towards the ends.

In fact, this observation is related to that the chain ends are more exposed to the globule surface, and it can be better justified by considering the probability of the i -th monomer in the chain to appear on the surface of the globule, $P_i^{(surf)}$. This is presented in the left-hand-side of Fig. 4.23. Thus, the function $P_i^{(surf)}$ for kinetic simulation is nearly constant except for a few monomers at the very chain ends. In particular, for the end monomer this probability is about 1.3 times larger than that for a monomer in the centre of the chain.

The effect due to applying reptational moves during later kinetic stages is quite distinguishable in observables in Figs. 4.22 and 4.23. For sufficiently large values of the chain indices, $\hat{k} \gtrsim 0.35-0.4$, the function of mean squared distances obtained

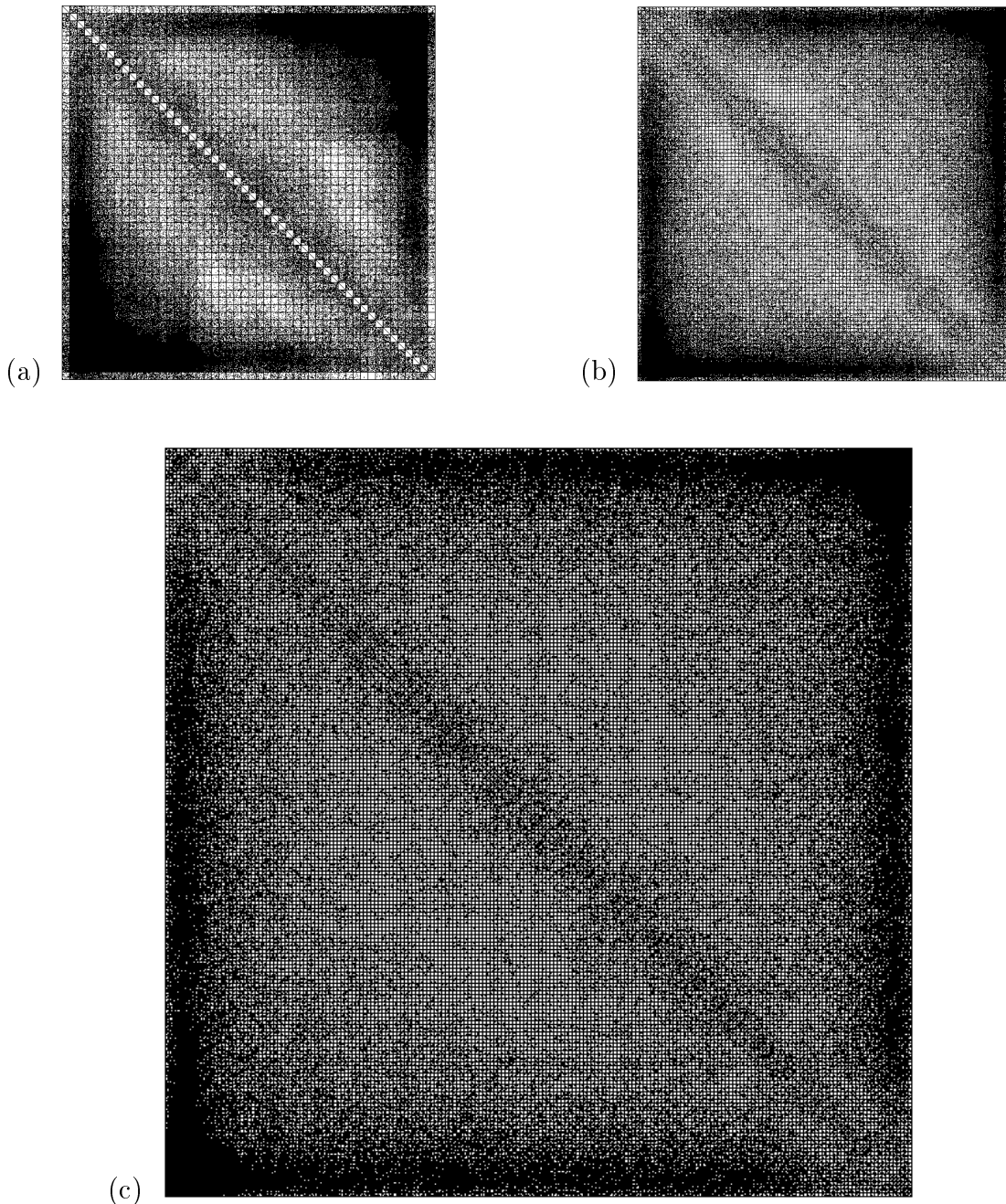
Figure 4.23: Plot of the probability, $P_i^{(surf)}$, for the i -th monomer in the chain to appear on the globule surface vs the monomer index, i , for the homopolymer with the degree of polymerisation, $N = 400$, for different simulation procedures. The left- and right-hand-side curves correspond to the kinetic (no reptations) and reptational simulation procedures. Both distributions are symmetric in monomer index, i , and for convenience they are presented only on halves of the interval.



from simulation with reptations (see lines denoted by diamonds in Fig. 4.22) lies below the appropriate curves from the kinetic simulation. The situation is just the reverse for smaller k , with the difference slowly vanishing towards $k = 0$. Thus, the reptational curve for D_k behaves in the following way: it is linear for small k , reaches a maximum at around $\hat{k} = 0.3$ and then slowly decreases reaching a minimum in the vicinity of the chain ends. At the very end the function increases strongly, still being significantly smaller than for the function D_k without reptations.

Note, however, that the mean squared distances expressed as functions of the chain index, in some sense, are not fully representative due to the role of the end effects in the open chain. These do not distinguish the position of monomers in the chain, but only their relative distance. Thus, it is interesting to consider the mean

Figure 4.24: Diagrams of the ratios of the mean squared distances with and without reptations, $D_{ij}^{(r)}/D_{ij}^{(n)}$ for different values of the degree of polymerisation. Diagrams (a-c) correspond respectively to the following values of the degree of polymerisation: $N = 50$, $N = 100$ and $N = 200$. Indices i, j start counting from the upper left corner. Each matrix element is denoted by a quadratic cell with varying degree of black colour, the darkest and the lightest cells corresponding respectively to the smallest and to the largest ratios of the mean squared distances. Intensity of black colour in the near diagonal elements corresponds approximately to the ratio $D_{ij}^{(r)}/D_{ij}^{(n)} = 1$.



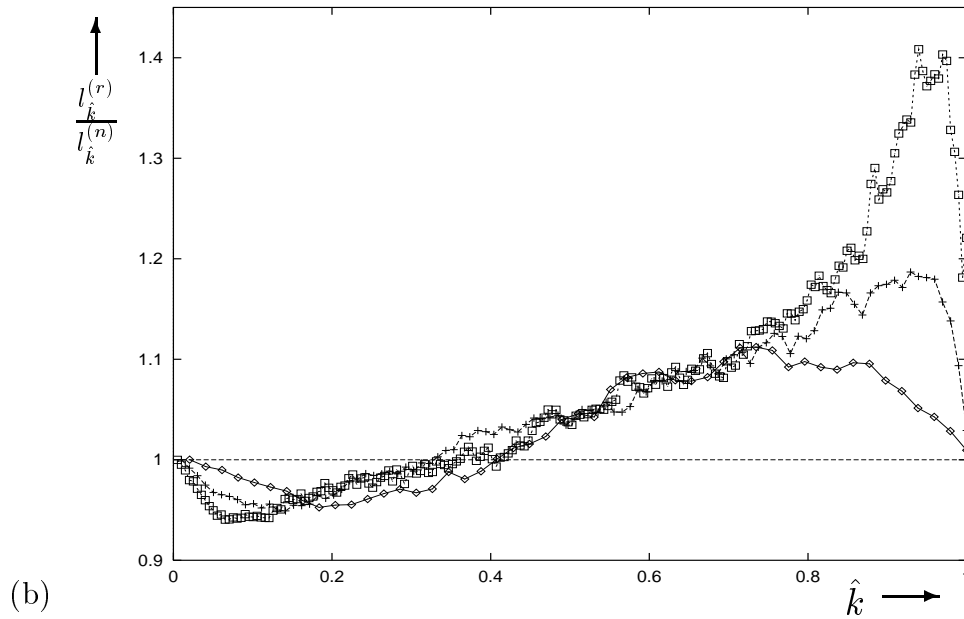
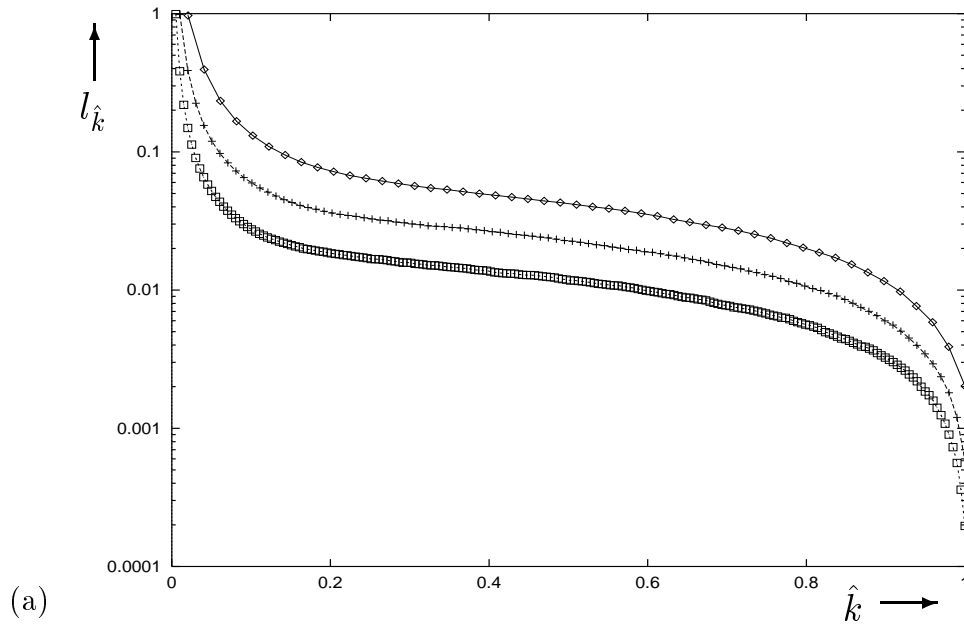
squared distances as functions of two indices, representing the monomer positions along the chain. In Fig. 4.24 we exhibit the ratios of mean squared distances with and without reptations, $D_{ij}^{(r)}/D_{ij}^{(n)}$. Depending on the behaviour of this ratio, the monomers in the chain may be roughly divided into three groups: for monomers in the centre of the chain, $0.1 \lesssim i/N \lesssim 0.9$, the mean squared distances significantly increase due to reptations; for groups of penultimate to the end monomers, $3 \lesssim i, N - i - 1 \lesssim 0.1N$, the distances decrease due to reptations, with the effect becoming more pronounced on approaching the ends of the chain; and for a few end monomers the distances decrease but much more weakly. Thus, we can say that reptations push the end groups a bit and the penultimate groups of monomers more towards the centre of the globule, thus reducing their mean squared distances between each other and monomers from the central group. As the density of the globule does not change during this stage this leads to some increase of the mean squared distances for monomers in the centre of the chain.

Such behaviour of the mean squared distances is quite consistent with the plot of the surface probability $P_i^{(surf)}$ after reptational stage (right-hand-side of Fig. 4.23). It increases for monomers in the central group by about 10% and, instead of being constant, decreases slowly towards the ends of the chain. The drop in the probability is most pronounced for monomers from the penultimate groups. The function rapidly increases for a few monomers at the very ends of the chain, although, they still possess a lower probability to be found on the surface than monomers from the central group. This rapid increase of the probability for a few end monomers may be interpreted as an effect of a single end seeking to maximise its entropy without a complete loss of the attractive interaction. Indeed, the chain ends are most free to explore the surface, whilst the rest of monomers are restricted due to their connectivity.

In Fig. 4.25a we exhibit the number of loops per monomer, $l_{\hat{k}}$, defined by

Figure 4.25: Plots of the number of loops per monomer, $l_{\hat{k}} = L_{\hat{k}}/N$, (Fig. a) and the ratio of the number of loops in simulations with, $l_{\hat{k}}^{(r)}$, $p_r = 1/N$, and without, $l_{\hat{k}}^{(n)}$, $p_r = 0$, reptations (Fig. b) vs the normalised chain index, \hat{k} , for the homopolymer globule.

Lines correspond to the following values of the degree of polymerisation (from top to bottom in Fig. a): $N = 50$ (diamonds), $N = 100$ (pluses) and $N = 200$ (quadrangles).



Eq. (4.2), versus the normalised chain index, \hat{k} for the globule obtained after the reptational stage. The specific number of loops for small values of the chain index behaves as follows. The first point is $l_1 = 1$, which reflects close monomer packing inside a dense globule. Really, since here monomers are packed in a 3-dimensional chess board manner the distance between nearest neighbours along the chain is equal to $\sqrt{2}$, and thus the number of loops of unit length is equal to $L_1 = N$. For the values of the chain index in some range, which corresponds approximately to the region of linear increase in the mean squared distances (see left-hand-side of Fig. 4.22), the number of loops falls in k as a power law $l_k \sim k^{-\beta_l}$. The numerical value of this exponent is found to be $\beta_l = 1.34$, which is rather close to the quadruple value of the swelling exponent of the globule, $4\nu_g = 4/3$. For chain indices in the middle of the chain the number of loops decreases approximately exponentially according to Fig. 4.25a. For large values of the chain index corresponding to the penultimate and end groups of monomers the number of loops falls even faster than exponentially. However, it is interesting to note that for the ring homopolymer globule one would expect a saturation of this quantity at certain level. Indeed, since the monomer–monomer correlation function, $h_k^{(2)}(r)$ defined by Eq. (4.4), is k -independent for large chain indices, and as the number of loops is proportional to (see Eq. (4.2)),

$$L_k \sim \int_0^b h_k^{(2)}(r) r^2 dr, \quad (4.50)$$

with some k -independent coefficient, the number of loops is k -independent in this region of the chain indices as well.

Analogously to Fig. 4.24, in Fig. 4.25b we exhibit the relative change in the number of loops during the reptational kinetic stage. For small values of the chain index, $0 < \hat{k} \lesssim 0.4$, the reptational stage leads to a certain decrease in the number of loops. It is natural to expect that parts of the chain which are pushed towards the surface of the globule would possess a smaller number of loops. On

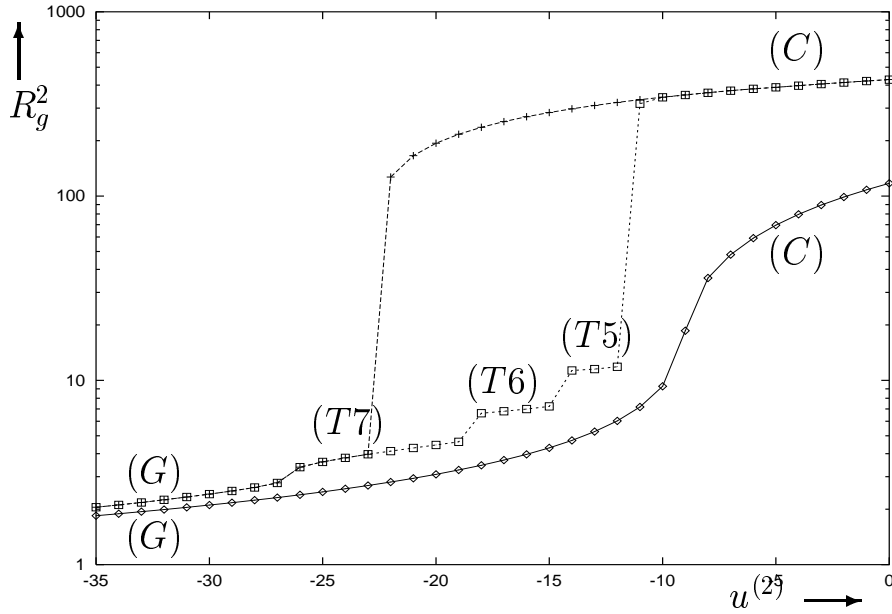
the contrary, for large values of the chain index, $0.4 \lesssim \hat{k} \leq 1$, the number of loops increases due to reptations. The ratio $l_{\hat{k}}^{(r)}/l_{\hat{k}}^{(n)}$ increases most significantly for the end-penultimate monomer groups since their escape from the surface of the globule due to reptations is most pronounced (see the right-hand-side of Fig. 4.23) and, thus, the probability to form a loop is higher. For the end monomers, according to the above study of different observables, the number of loops increases due to reptations in a much weaker fashion. Interestingly enough, the effect in the relative change of the number of loops for the penultimate and the end groups increases with the degree of polymerisation.

We may conclude therefore that reptations do play an important role in changing the structure of the globule of the open nonphantom chain. The globule after the late kinetic stage is characterised by a uniform monomer distribution except for the end monomers, which have a higher probability to be on the surface. Due to reptations the end, and even more so, the penultimate monomer groups are pushed more towards the centre of the globule, with the central monomer group thus more exposed to the surface. A typical magnitude of this effect in observables expressing the k -dependent properties is about a few dozens percents in the relative change, while the global characteristics such as the mean energy and radius of gyration are practically unaffected by this.

4.5 Conformations of Stiff Macromolecules

In this section we study how inclusion of local rigidity of the polymer chain affects the equilibrium phase diagram and kinetic phenomena at the collapse transition of the homopolymer. We shall present results from the numerical studies of the GSC equations (2.73) as well as from the off-lattice Monte Carlo simulation. In both cases the local rigidity is introduced in the models as the bending energy contribution, $\sum_i (d^2 \mathbf{X}_i / di^2)^2$, according to Kratky and Porod [90].

Figure 4.26: Plots of the mean squared radius of gyration, R_g^2 , vs the second virial coefficient, $u^{(2)}$, upon a quasistatic change of the latter. Solid line (diamonds) corresponds to the value of the stiffness parameter $\lambda = 5$, where the collapse transition is second order and dashed lines correspond to $\lambda = 25$ (pluses from positive $u^{(2)}$ to negative one, and quadrangles from negative to positive). Labels (C), (G), (T5)-(T7) correspond to the coil, globule and various toroidal states respectively.



First, let us consider the system behaviour upon a quasistatic change of interaction parameters. In Fig. 4.26 we present the plot of the mean squared radius of gyration, R_g^2 , versus the second virial coefficient at a fixed stiffness parameter, λ . The regime of repulsion and comparatively weak attraction in the right-hand-side of the figure corresponds to the extended coil conformations of the polymer with a large radius of gyration scaling as $R_g \sim N^{\nu_{coil}}$, where the exponent ν_{coil} is close to the Flory value $\nu_F = 3/5$ for a flexible chain, $\lambda = 0$, becomes a rigid rod exponent $\nu_{rod} = 1$ for a very stiff chain, with a cross-over in between. Since increasing the stiffness increases effective repulsion between monomers, the extended phase expands in the region of negative second virial coefficient for higher

values of the stiffness parameter.

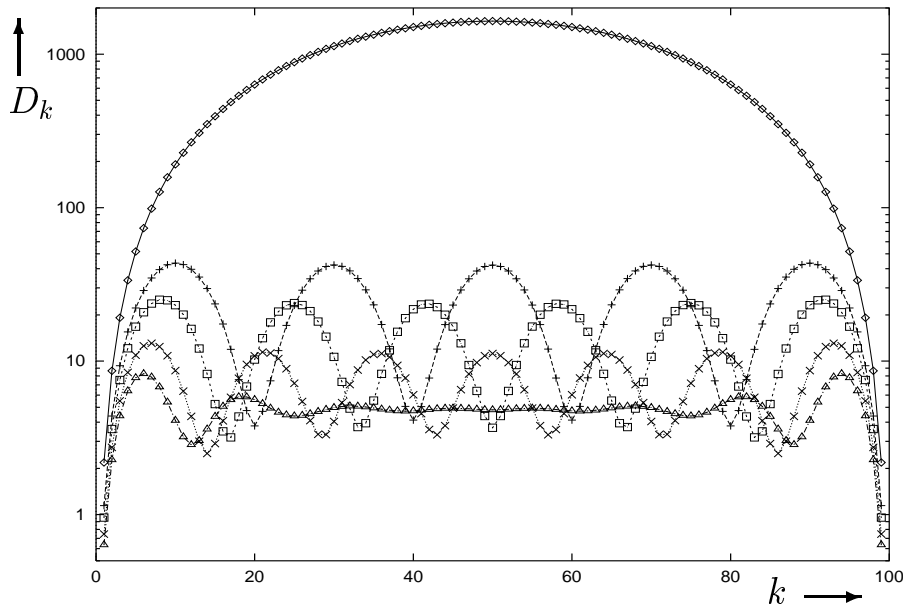
For comparatively small values of λ the plot of the radius of gyration (solid line and diamonds in Fig. 4.26) is quite similar to that of the flexible homopolymer for the equilibrium coil-to-globule transition (see e.g. Fig. 4.6), which is second order.

However, at higher values of the stiffness parameter the collapse transition becomes first order. In this case after the system has been quasistatically quenched to the region of a higher monomer attraction (line denoted by pluses in Fig. 4.26), the local minimum corresponding to the coil suddenly disappears becoming an inflexion point somewhere in the interval, $-23 < u^{(2)} < -22$, and the system passes to another free energy minimum with a much lower value of the radius of gyration. We shall call the boundary of existence of a minimum of the free energy a *spinodal*. Similarly, upon changing $u^{(2)}$ quasistatically towards monomer repulsion (line denoted by quadrangles in Fig. 4.26), the free energy minimum disappears in the interval, $-12 < u^{(2)} < -11$ and the system transforms into the coil state. If at least two minima of the free energy can coexist in some interval of interaction parameters, the transition point is defined by the condition that the current minimum of \mathcal{A} becomes the deepest one. Observables such as the energy, \mathcal{E} , the radius of gyration, R_g , or the mean squared distances between monomers, D_{ij} , experience a discontinuous jump at such a transition.

It is important to note that upon a quasistatic change of the second virial coefficient towards repulsion (line denoted by quadrangles in Fig. 4.26) the mean squared radius of gyration increases in a three-step-like fashion before the homopolymer expands to the coil. This is a manifestation of some additional, similar to the globule, condensed phases, which we denoted by labels $(T5)$, $(T6)$ and $(T7)$. To understand the distinction between these phases and the spherical globule let us consider their monomer-monomer mean squared distances, D_k as functions

Figure 4.27: Plot of the mean squared distances between monomers, D_k , vs the chain index, k , for various conformational states.

Lines denoted by diamonds, triangles, pluses, quadrangles and crosses correspond respectively to the coil, globule, and toroidal states with the winding number equal to 5, 6 and 7.



of the chain index (number of links between monomers), k , which we exhibit in Fig. 4.27. As we have discussed in Sec. 4.2 (see Fig. 4.10) for the state of the extended coil this function monotonically increases towards half of the chain. The situation does not change much for the coil of a stiff homopolymer (line denoted by diamonds in Fig. 4.27).

However, the function D_k changes for the state of the liquid-like globule (line denoted by triangles in Fig. 4.27), especially for small values of the chain index (compare this with the plot of the function D_k for the globule in Fig. 4.10). At small distances along the chain it is nearly parabolic, $D_k \sim |k|^2$, i.e. the chain represents almost a rigid rod, reaching a maximum at some value of the chain index k^* . In the range of intermediate values of the chain index, $0 < k < 6k^*$, one can see 2-3 oscillations in D_k with the amplitude decreasing quickly to stabilise at some level. At higher values of the chain index towards half of the chain the

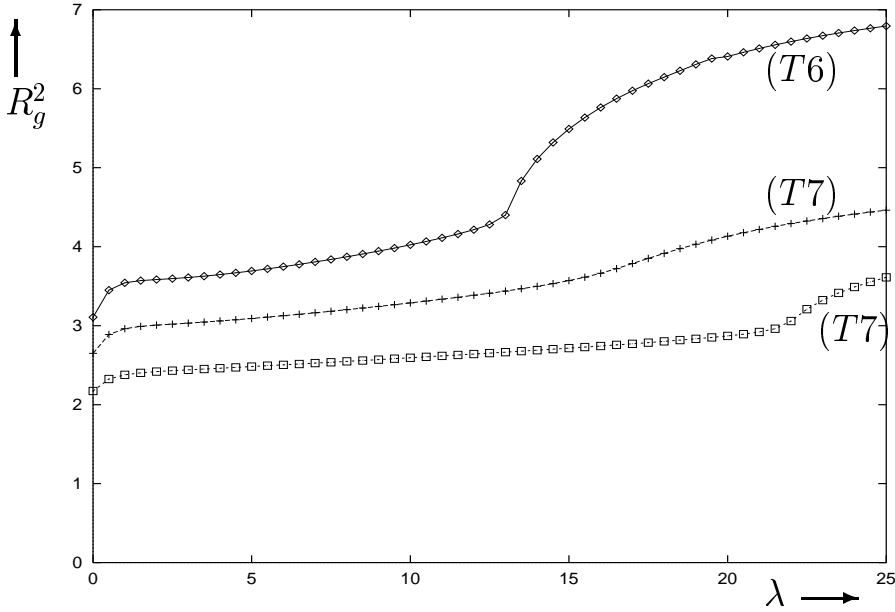
function remains constant. Thus, for large values of the stiffness parameter the structure of the globule is different on small chain distances. Chain segments in the globule are locally straightened on a characteristic scale, which is smaller than the globule size. When the length of these straightened segments becomes comparable with the size of the globule, a few oscillations can be seen in the mean squared distances.

Transition of the conventional globule to the phase labelled as (*T7*) is accompanied by a spectacular change of this function of the mean squared distances (line denoted by crosses in Fig. 4.27). In this phase the function strongly oscillates with the amplitude decreasing slowly towards half of the chain. During such oscillations the mean squared distances change 5-6 times from a minimum to a maximum around half of the chain. In phase (*T7*) one can see 7 such oscillations, while in the phases labelled by (*T6*) and (*T5*) there are 6 and 5 oscillations respectively. The amplitude of oscillations in these phases is even higher.

Such properties can be naturally attributed to the toroidal conformation. Thus, a stiff chain, having no point of “easy bending”, is nearly straight locally and it wraps around itself forming a hole in the middle of the torus. The chain index k^* , where the function D_k reaches its first maximum, is equal to the number of monomers forming half-period of the first winding starting from the zeroth monomer. Therefore, $D_{max} = D_{k^*}$ may be interpreted roughly as the mean squared *external* diameter of the torus. By moving from the monomer number k^* to $2k^*$ the first winding is completed. However, because of the excluded volume interaction the chain cannot return to the same coordinate, giving rise in the average to the value $D_{min} = D_{2k^*}$, which may be considered as the mean squared *internal* diameter of the torus. The number of local maxima in the function of mean squared distances is then equal to the winding number of the torus. The physical reason for a torus is clear — a persistent chain has no desire to bend,

Figure 4.28: Plots of the mean squared radius of gyration, R_g^2 , vs the stiffness parameter, λ .

Lines correspond to (from top to bottom) $u^{(2)} = -17$ (diamonds), $u^{(2)} = -20$ (pluses) and $u^{(2)} = -25$ (quadrangles).

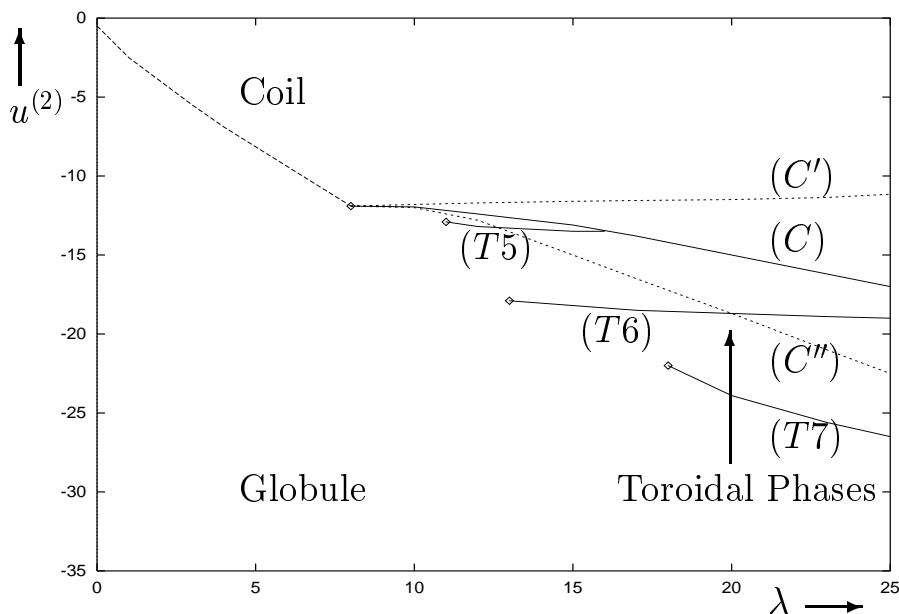


so it tends to have as large a radius of curvature as possible, however, two-body attraction tends to keep quite close packing of the chain.

Thus, a quasistatic increase of repulsion (see line denoted by quadrangles in Fig. 4.26) results in transformation of the liquid-like globule to the toroidal globule with 7 windings in the interval, $-27 < u^{(2)} < -26$. This is a rather weak discontinuous phase transition, during which the oscillations reach the region of chain index, $k = N/2$, and a hole in the centre of the globule is formed. Transitions between various toroidal states are also first order, but much stronger since they are accompanied by a global restructuring of the polymer conformation.

In Fig. 4.28 we present the change of the mean squared radius of gyration, R_g^2 , upon a quasistatic change of the stiffness parameter at a fixed negative second virial coefficient, $u^{(2)}$, corresponding to the globule. The radius of gyration increases monotonically during this change. The most significant change occurs

Figure 4.29: Phase diagram of the stiff homopolymer in terms of the stiffness parameter, λ , and the second virial coefficient, $u^{(2)}$. Here (C) is the transition line between the coil and globular phases, lines (C') and (C'') are spinodals of that transition. Lines $(T5)$ - $(T7)$ represent transitions to toroidal states with corresponding winding numbers. The spinodals of these transitions are not depicted. The phase diagram has been obtained for the polymer with the degree of polymerisation $N = 100$.



in the region of rather small stiffness parameter, $0 < \lambda < 1/2$, and in the region where the liquid-like globule transforms to the toroidal state. The latter is a weak first order transition. Note that the final toroidal state may depend on the value of the second virial coefficient. At weaker attraction the globule is transformed to a toroidal state with a smaller winding number. However, the transition from the coil to toroidal phases upon a quasistatic change of $u^{(2)}$ (line denoted by pluses in Fig. 4.26) is quite suppressed due to a huge potential barrier, which makes the coil a metastable state in a large region of the phase diagram.

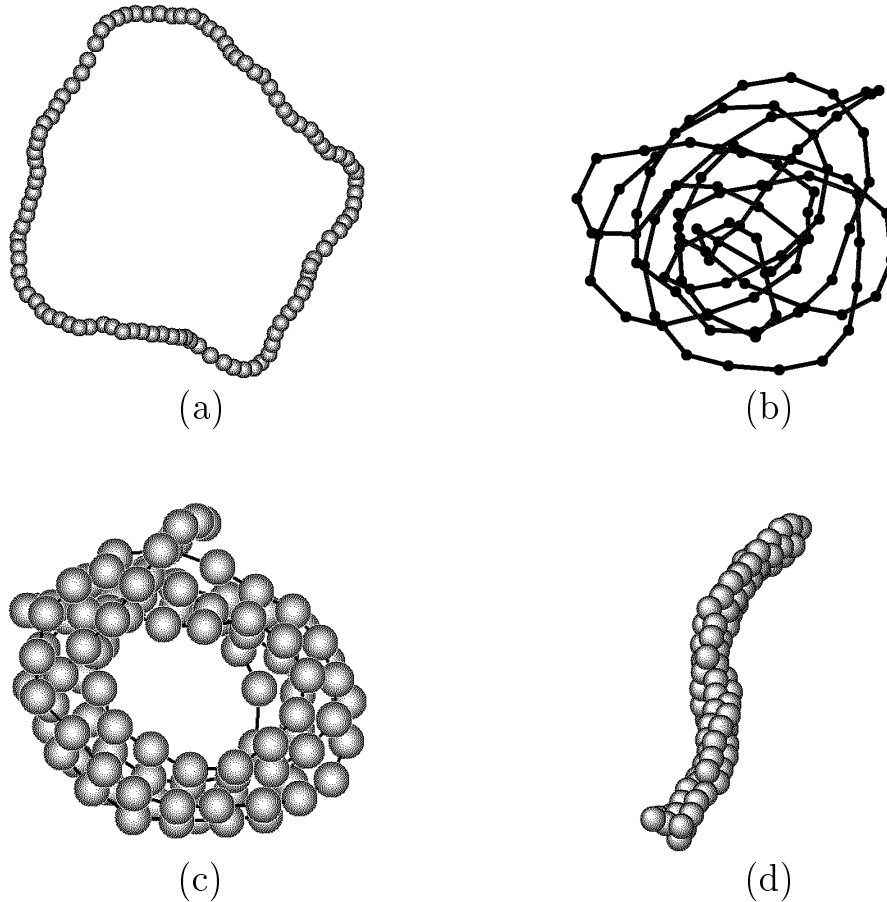
In Fig. 4.29 we present the resulting phase diagram of the persistent homopolymer in terms of the second virial coefficient, $u^{(2)}$, and the stiffness parameter, λ . It contains phases of coil, where monomer attraction is insufficient to form more

compact states, spherical globule in the region of either low rigidity or strong monomer attraction, and a number of toroidal phases characterised by distinct winding numbers. As we have already discussed the collapse transition changes its behaviour from continuous to discontinuous starting from some value of the stiffness. The spherical globule is somewhat different here in its local structure compared to that of the flexible homopolymer, although global scaling characteristics are the same for both states. The toroidal phases lie in the intermediate region in $u^{(2)}$ starting from some critical value of the stiffness parameter. Note that these toroidal states become thermodynamically stable starting from some value of the winding number. For a large fixed value of the stiffness parameter the number of toroidal phases increases approximately linearly with the degree of polymerisation, N . For example, the maximal winding numbers, n_w , at $\lambda = 25$ for polymers with the degree of polymerisation $N = 50, 100, 150$ and 200 are equal to: $n_w = 4, 7, 11$ and 13 respectively.

Now let us discuss results from the Monte Carlo simulation. Since the lattice Monte Carlo model (see Sec. 3.1) is not particularly good for study of the effects of chain rigidity for several reasons, we have carried out simulation in continuous space instead (see Sec. 3.2). The disadvantages of the lattice model include the rotational anisotropy and tendency for condensed phases to form some kind of a crystalline structures on a lattice. Thus, to produce toroidal states on a lattice for comparatively short polymer chains is rather difficult. Another disadvantage of the lattice model is that the persistence of the polymer chain reduces dramatically the acceptance ratio of the local monomer moves and, thus, other types of moves performing shifts and rotations of chain segments as a whole are needed.

On the other hand, in simulation in continuous space one can reduce the parameter of relative trial change of coordinates, Δ , although at a cost of longer

Figure 4.30: Typical conformations of a stiff homopolymer with the degree of polymerisation $N = 100$ from the off-lattice Monte Carlo simulation. Figures (a)-(d) correspond respectively to the coil, $V_0 = 0$, $\lambda = 5$; the globule, $V_0 = 5$, $\lambda = 5$; the toroidal globule, $V_0 = 5$, $\lambda = 15$; and the hairpin kinetically arrested state, for the quench $V_0 = 0 \rightarrow 5$, $\lambda = 8$.



equilibration. The latter, along with the problem of quasi-nonergodicity, which is inevitable for any Monte Carlo (and any other) simulation for systems with strong interactions, make it rather difficult to obtain a good statistics for estimates of averages. Thus, the results from Monte Carlo simulation are presented at a qualitative level mainly.

In series of pictures in Fig. 4.30 we exhibit typical conformations in various phases from the off-lattice Monte Carlo simulation. Fig. 4.30a corresponds to the extended conformation, which for large values of the stiffness parameter takes a

form close to a ring. In Fig. 4.30b we present the backbone of typical globular conformation for the values of the rigidity parameter not large enough to form toroidal conformation. The globule structure here is quite different from that of a flexible homopolymer (compare it with e.g. with the globule in Fig. 4.32a). One can see that the globule consists of entangled loops of the radius close to the size of the globule. The function of mean squared distances between monomers here possesses a typical form presented by triangles in Fig. 4.27 with significant oscillations on small chain distances, which quickly saturate due to varying number of monomers in each loop. Increasing the stiffness parameter transforms the globule to the toroidal conformation exhibited in Fig. 4.30c.

It is important to note that to produce the globular state as in Fig. 4.30b, as well as the toroidal conformations as in Fig. 4.30c in Monte Carlo simulation, it is easier to bring the system quasistatically first to the globule of the flexible homopolymer and then to increase the stiffness parameter quasistatically. As we have mentioned earlier, the region of the metastable coil is quite wide for large values of the stiffness parameter. Another problem is that the system becomes trapped in metastable states which appear during transformation of the persistent chain from the extended to the compact states. Polymer conformations in these states have typical form of a *hairpin* exhibited in Fig. 4.30d. Here the chain folds several times along nearly straight line and bends only few times at the ends. However, these ends contribute more significantly to the bending energy than a uniform slow bending, thus, such conformations possess larger energy and hence are metastable. Typically, the process of forming hairpins is the following. The ring-like conformation as in Fig. 4.30a due to rigidity is quite stable with respect to the increase of monomer attractions. However, due to thermal fluctuations some distant parts of the chain meet each other occasionally, i.e. the chain takes a form of the digit ‘8’. If monomers are attractive enough, parts of the chain would

arrange along each other starting from the centre towards edges. The process can repeat if the persistent length is shorter than half of the chain. Without performing quite sophisticated collective movements of the chain segments in the hairpin state it is virtually impossible to proceed further in the collapsed phase.

Finally, let us note that the hairpin conformations are not observed in the Gaussian Self-consistent method neither at equilibrium, nor even as intermediate kinetic states. In fact, the origin of their formation lies in a process which starts due to local fluctuations in the monomer density. However, in the GSC method the monomer–monomer correlation function is represented by only one parameter characterising the mean squared distance between them. Proper account for the density fluctuations is beyond the scope of even an optimised mean–field type such as the GSC method. On the other hand, the lack of collective motions in the Monte Carlo scheme, quite likely, overestimates the role of such conformations as both metastable states and kinetics intermediates.

4.6 Orientationally Ordered Phases of the Homopolymer

In this section we consider the consequences of having various mechanisms of local in space interactions between monomers on the collapse of the homopolymer. This may lead, as we shall see, to a global refolding of the polymer in its globular state forming anisotropic globules. These studies may be interesting in respect to various crystallisation phenomena occurring at the collapse of polymers.

Here, in addition to the lattice Monte Carlo model described in Section 3.1, we include a short ranged three–body term. This allows us to have attractive two–body interactions in competition with repulsive three–body interactions, and we

focus our discussion on the new features due to these effects. The monomer–solvent interaction is now described by the Hamiltonian,

$$H = \frac{1}{2} \sum_{i \neq j} w(r_{ij}) \mathcal{I}_{s_i s_j} + \frac{1}{6} \sum_{i \neq j \neq k} v(r_{ij}, r_{jk}, r_{ik}) \mathcal{J}_{s_i s_j s_k}, \quad (4.51)$$

where \mathcal{J}_{abc} is a 2x2x2 symmetric tensor of three–body interaction constants and $v(r, r', r'')$ is a function representing the form of the three–body potential interactions. We chose the interaction functions $w(r)$ and $v(r, r', r'')$ so that the range of the three–body interaction is shorter than the range of two–body interaction,

$$w(r) = \begin{cases} 1 & r \leq 2 \\ 0 & r > 2 \end{cases},$$

$$v(r, r', r'') = \begin{cases} 1 & r, r', r'' \leq \sqrt{2} \\ 0 & r, r', r'' > \sqrt{2} \end{cases}.$$

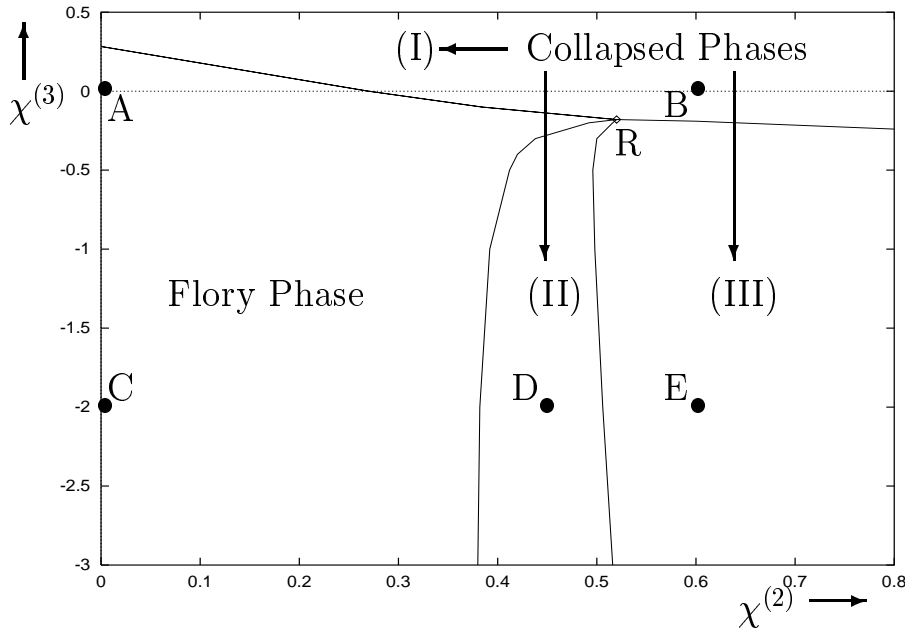
Similar to our previous discussion it can be shown that the probability of an update of the system conformation may be expressed as an exponent of a linear combination of two– and three–body interaction parameters. These are defined as follows,

$$\begin{aligned} \chi^{(2)} &\equiv \frac{2\mathcal{I}_{ms} - \mathcal{I}_{mm} - \mathcal{I}_{ss}}{k_B T}, \\ \chi^{(3)} &\equiv \frac{2\mathcal{J}_{mms} - \mathcal{J}_{mmm} - \mathcal{J}_{mss}}{k_B T}, \\ \chi^{(3)'} &\equiv \frac{2\mathcal{J}_{mss} - \mathcal{J}_{mms} - \mathcal{J}_{sss}}{k_B T}. \end{aligned} \quad (4.52)$$

For simplicity we restrict ourselves by the additional condition $\mathcal{J}_{sab} = 0$, which leads to $\chi^{(3)'} = 0$. Positive or negative values of the interaction parameters in (4.52) correspond to attraction or repulsion respectively. Thus, we have only two parameters, those given in equation (4.52), that determine the thermodynamical properties of the system defined by Hamiltonian (4.51).

Figure 4.31: Phase diagram of the homopolymer with the degree of polymerisation $N = 120$ in the model including three-body interactions.

Solid lines correspond to the peaks in the heat capacity, which determine the transition values of the interaction parameters $\chi^{(2)}$ and $\chi^{(3)}$ defined by equation (4.52). Region (I) of the phase diagram corresponds to isotropic liquid compacted state. Region (III) corresponds to anisotropic collapsed state, which is analogous to nematic state of liquid crystal. In the region (II) the single globule consists of several locally ordered regions. The point R appears to represent the joining of all transition curves.

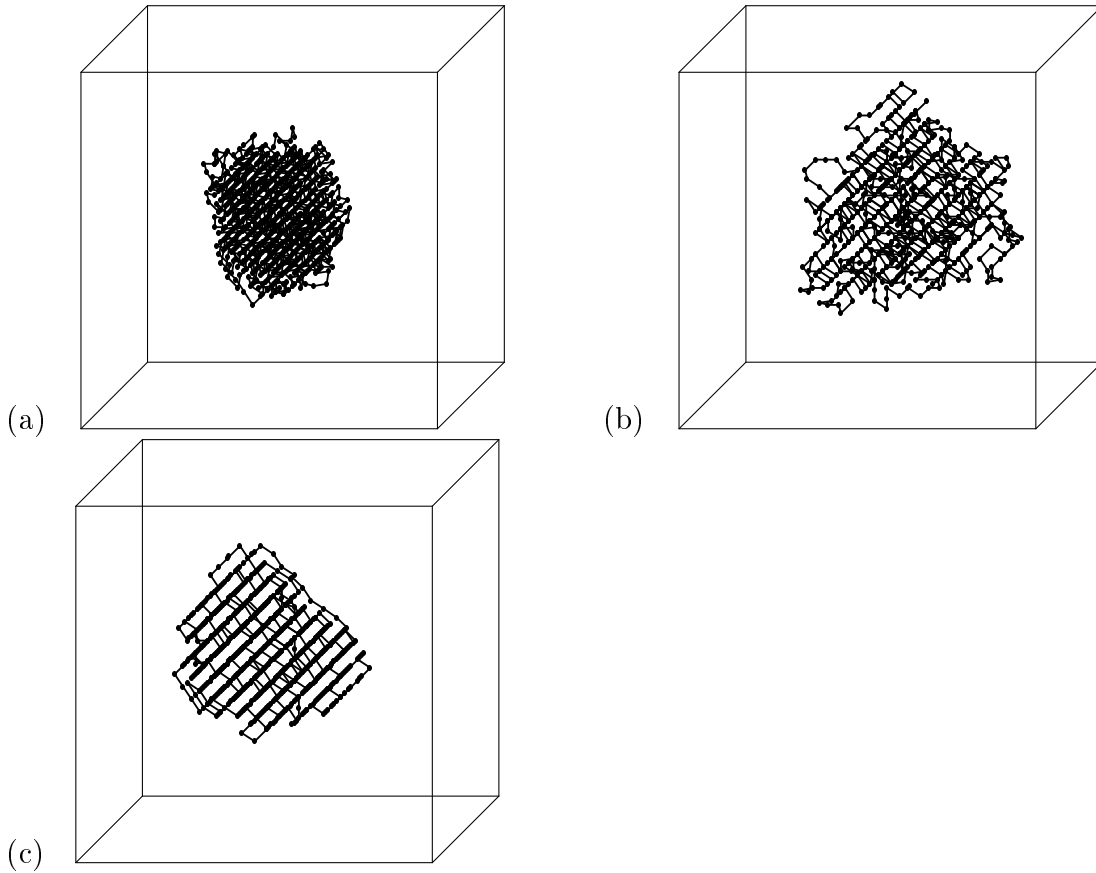


4.6.1 Equilibrium phases

We aim to demonstrate in this subsection that the phase diagram of the model appears to contain more phases than it is usually expected. We are apparently able to distinguish three different collapsed phases divided by sharp peaks in the heat capacity. These may be characterised by differing degrees of orientational order of the polymer segments.

The phase diagram of the model described by the Hamiltonian (4.51) is exhibited in Fig. 4.31. The lower left part of the diagram corresponds to the good solvent condition and there the Flory coil is preferred. Of course we expect no real changes in the Flory coil since for a fluctuating state only the effective Kuhn

Figure 4.32: Conformations of the collapsed polymer at equilibrium for the degree of polymerisation $N = 768$ in the model with three-body interactions. The size of the framing box is equal to $L = 32$. Pictures (a), (b) and (c) correspond respectively to states found in the regions of the phase diagram in Fig. 4.31 labelled by (I), (II) and (III).



length will be modified. Other regions labelled I, II and III correspond to compacted globule phases of different types.

In Fig. 4.32 we exhibit examples of particular polymer conformations representative of these phases. In phase I (see Fig. 4.32a) the packing is space-isotropic and it is determined only by the connectivity and excluded volume restrictions. The conformation does not possess any order along the chain. This is what is conventionally called the collapsed phase. We prefer to call it the isotropic phase since there are other collapsed phases.

In phase III (see Fig. 4.32c) a typical conformation possesses a preferred direction. Most chain segments are aligned along this direction forming rods that are arranged in a 2-dimensional chess-board pattern in the plane perpendicular to the rods. One can see that such a structure minimises the three-body interaction term of the energy. This phase may be considered as analogous to the uniaxial nematic phase of a liquid crystal [172]. At first this conclusion may seem somewhat surprising because originally the coil does not consist of any elongated objects. However the appearance of rod-like segments clearly arises because the three-body repulsion makes long straight pieces of chain more favourable, whilst the two-body effects cause compaction. We should note also, that the radius of gyration in the region I is somewhat smaller than in the region III.

The phase II (see Fig. 4.32b) lies between the Flory phase and orientationally ordered phase III. Again, this state is stable because of the balance between long ranged two-body attraction and short ranged effective three-body repulsion. A typical conformation here consists of locally ordered domains of segments of the chain aligned in different directions. Thus, there is no overall alignment between domains and the phase is globally isotropic.

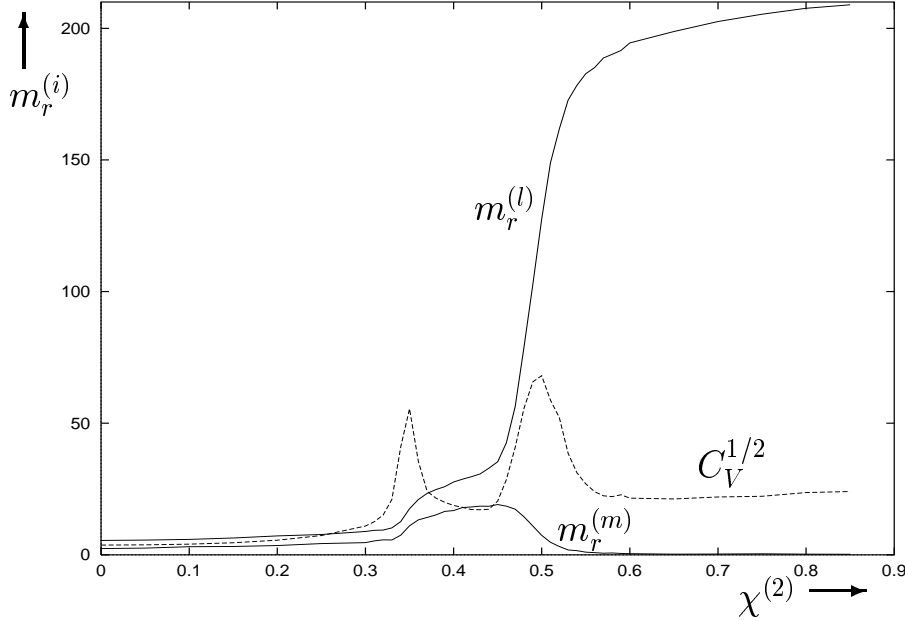
To make our discussion more formal, let us introduce the orientational order parameters. We denote by $l_r^{(i)}$ the number of consecutive linearly ordered monomer beads along one of the axes, i , (x , y or z). We take into account only rods for which $l_r^{(i)} > 2$. The number of rods per polymer chain along the i axis we denote as $n_r^{(i)}$. Now let us consider the following quantities,

$$\begin{aligned} m_r^{(i)} &\equiv l_r^{(i)} n_r^{(i)}, \\ m_r &\equiv m_r^{(x)} + m_r^{(y)} + m_r^{(z)}. \end{aligned} \quad (4.53)$$

These give the number of monomers that are aligned along the appropriate directions and their sum. Obviously, space isotropy implies,

$$\langle m_r^{(x)} \rangle = \langle m_r^{(y)} \rangle = \langle m_r^{(z)} \rangle. \quad (4.54)$$

Figure 4.33: Plot of $\langle m_r^{(l,m)} \rangle$ vs the two-body interaction parameter, $\chi^{(2)}$. Here the degree of polymerisation is $N = 240$ and value of the three-body interaction parameter $\chi^{(3)} = -2$ (locus between C and E in Fig. 4.31). The plot of heat capacity is labelled by C_V .



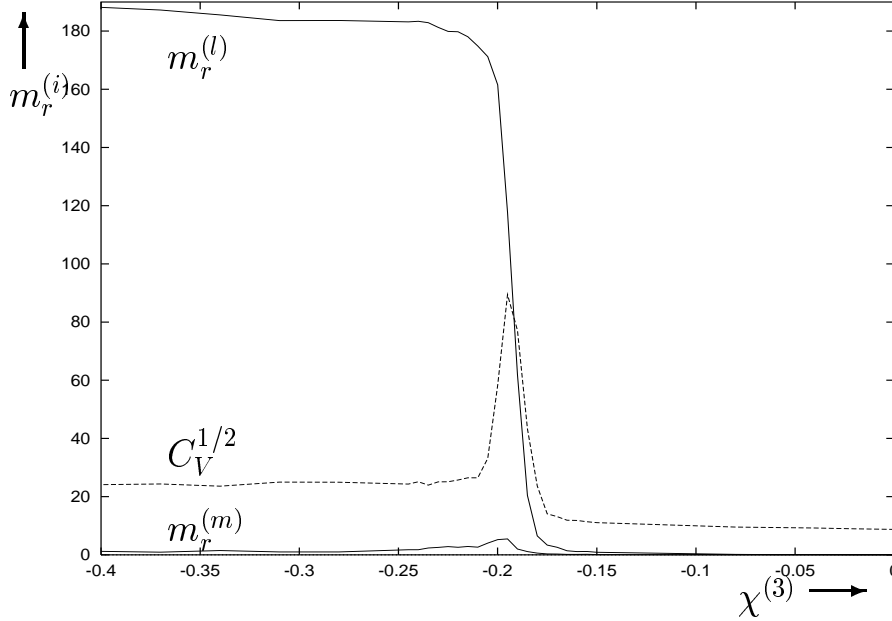
When this isotropy is broken the values of the three averages are different. In our case there is no explicitly symmetry breaking term in the Hamiltonian (4.51). A procedure analogous to the introduction of an infinitesimal magnetic field in Ising model would be hard to adopt here. Let us instead consider the triple $m_r^{(l)}$, $m_r^{(m)}$ and $m_r^{(s)}$, these being respectively the largest, middle and smallest values of the components $m_r^{(i)}$ defined before the statistical averaging.

The behaviour of the quantities $\langle m_r^{(l,m,s)} \rangle$ across different lines of the phase diagram (Fig. 4.31) is shown in Figs. 4.33 and 4.34. Simulations show that the dependence of the quantity m_r on the degree of polymerisation in the Flory coil state may be well described by a linear law,

$$\langle m_r \rangle = bN, \quad b = 0.0348 \pm 0.0006, \quad \chi^{(2)} = \chi^{(3)} = 0. \quad (4.55)$$

The value of the coefficient b is close to that obtained by approximating the

Figure 4.34: Plot of $\langle m_r^{(l,m)} \rangle$ vs the three-body interaction parameter, $\chi^{(3)}$. Here the degree of polymerisation is $N = 240$ and value of the two-body interaction parameter $\chi^{(2)} = 0.6$ (locus between B and E in Fig. 4.31). The plot of heat capacity is labelled by C_V .



excluded volume effect for only the first two nearest neighbours along the chain,

$$b = \frac{1}{91} \sum_{j=0}^{\infty} \frac{3+j}{21^j} \approx 0.035186. \quad (4.56)$$

We derive this formula as follows. The probability of a straight walk of length k along one of the axes x, y, z in some direction starting from an arbitrary point on the lattice is,

$$P_k = \frac{6}{26} \left(\frac{1}{21} \right)^{k-1}.$$

Here the first multiplier represents the probability of the first step, and the second of any further one. By definition, then, b can be written as,

$$b = \sum_{k=3}^{\infty} k P_k.$$

Although this is obviously a quite rough estimate, it appears to be in fair agreement with (4.55).

From Fig. 4.34 one can see that all the quantities $\langle m_r^{(l,m,s)} \rangle$ are negligibly small in the isotropic phase I due to the absence of elongated objects. In the region II the values of all $\langle m_r^{(l,m,s)} \rangle$ appear to be much larger than in the Flory state. This is consistent with visual observations where one sees local orientational ordering in some regions of the globule. Across the transition between phases II and III, the values of $\langle m_r^{(m,s)} \rangle$ fall to zero, whilst the value of $\langle m_r^{(l)} \rangle$ essentially increases. In the ideal case when one can neglect the noise, the quantity $\langle m_r^{(l)} \rangle$ scales with the degree of polymerisation as,

$$\langle m_r^{(l)} \rangle \rightarrow N, \quad N \rightarrow \infty, \quad (4.57)$$

i.e. all chain segments come to be distributed in a rod-like manner if the surface effects are negligible.

Thus, by modification of the Hamiltonian of the system (4.51) to include repulsive short ranged three-body interactions one can produce rather complicated structure of the globular state. There a set of various liquid, crystal and other phases can be found. The distinction between the Flory and phase II is consistent with a type of glass transition, whilst that between I and III is consistent with crystallisation.

4.6.2 Kinetics of folding

In this section we present results for kinetics at the transition from the Flory to the various collapsed states. The kinetic processes at the collapse transition into phase I (line AB in Fig. 4.31) in the present model have been already discussed and kinetic laws during different kinetic stages have been obtained. Here we discuss the possible applicability of those laws to the new transitions. Thus we will present kinetics results for the rapid quenches Flory \rightarrow phase II (line CD in Fig. 4.31) and Flory \rightarrow phase III (line CE). Kinetic calculation for quenches

across different collapsed phases is a significant computational task due to the large characteristic times of rearrangements in a dense globular state.

Consistent with previous calculations, during the earliest kinetic stages we find there is rapid formation of small locally compacted globules. A mode coupling mechanism for the internal modes leads to a decrease of the squared radius of gyration by a power law [II],

$$\Delta R_g^2(t) = -At^\alpha, \quad \alpha = 7/11 \approx 0.636. \quad (4.58)$$

The value of the exponent α obtained from simulations of the kinetics to the phase I is a little larger $\alpha = 0.66 \pm 0.03$. Numerical values of the exponent α obtained from simulations of early-stage kinetics to the other collapsed states for different degrees of polymerisation are exhibited in Table 1. One can see that the values of the exponent α for kinetics to different regions of the phase diagram are nearly the same. Thus, whatever the explanation, this early-stage results does indeed seem to be universal. Nevertheless, the internal structures of those locally collapsed globules obtained at early times as we pass to phases II and III are different and seem to be similar to the structure of the final collapsed state at equilibrium.

The second kinetic regime of the transition from Flory coil to the phase I proceeds mainly by unification of smaller clusters as we have earlier argued [II]. It was established that clusters grow according to a power law,

$$S = N(t/\tau_S)^Z, \quad Z = 1/2, \quad (4.59)$$

where S is the average number of monomers in a cluster. We call the parameter τ_S the total collapse time. The evolution of the squared radius of gyration during that stage may also be approximated by a few terms of the series,

$$R_g^2(t) = R_g^2(\infty) + \sum_{\alpha} A_{\alpha} \exp(-t/\tau_{\alpha}). \quad (4.60)$$

Table 4.6: Numerical values of parameters α , τ_S , Z , $R_g^2(\infty)$, A_1 , τ_1 from formulae (4.58, 4.59, 4.60) for polymers of different degrees of polymerisation, N . The exponent α has been obtained by a fitting procedure for times up to 500 MCS.

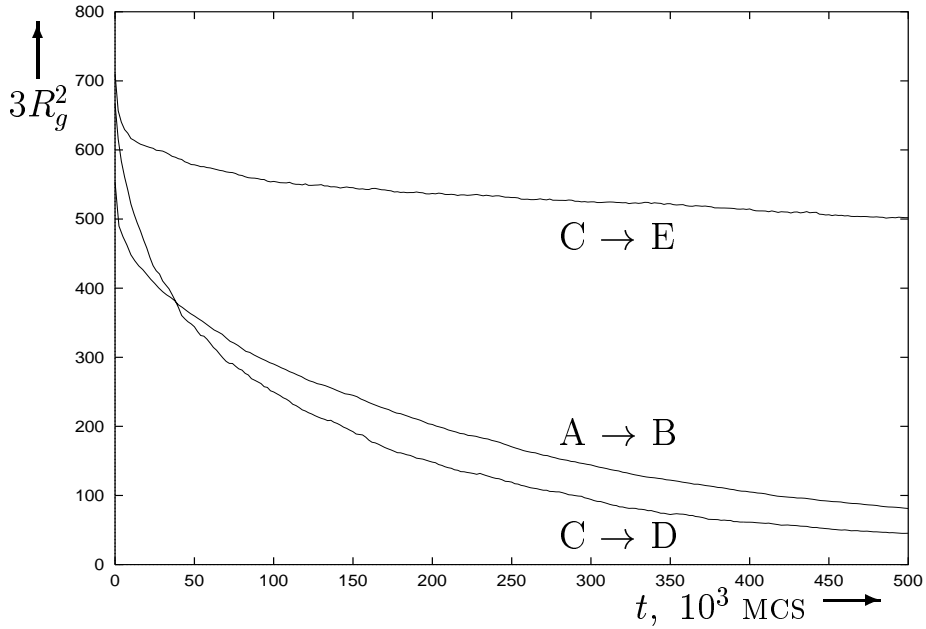
| N , line | α | Z | τ_S , 10^3 MCS | $R_g^2(\infty)$ | A_1 | τ_1 , 10^3 MCS |
|------------|-----------------|-----------------|-----------------------|-----------------|-------|-----------------------|
| 120, CD | 0.66 ± 0.02 | 0.46 ± 0.03 | 28 | 15.6 | 92 | 5.4 |
| 120, CE | 0.65 ± 0.04 | 0.26 ± 0.03 | 67 | 22 | 50 | 18 |
| 240, CD | 0.66 ± 0.03 | 0.43 ± 0.04 | 135 | 19.1 | 220 | 31 |
| 240, CE | 0.64 ± 0.04 | 0.18 ± 0.02 | 8920 | 135 | 78 | 83 |
| 360, CD | 0.65 ± 0.05 | 0.42 ± 0.03 | 371 | 20.6 | 330 | 89 |
| 360, CE | 0.67 ± 0.03 | 0.17 ± 0.03 | 68200 | 246 | 115 | 433 |
| 480, CD | 0.68 ± 0.03 | 0.43 ± 0.03 | 678 | 22.2 | 404 | 170 |
| 480, CE | 0.65 ± 0.04 | 0.20 ± 0.02 | 157000 | 432 | 136 | 779 |

For convenience we arrange the relaxation times τ_α in decreasing order $\tau_\alpha \geq \tau_{\alpha'}$ for $\alpha < \alpha'$. We call the relaxation time τ_1 the characteristic collapse time. For kinetics $A \rightarrow B$ both characteristic and total collapse times scale with the degree of polymerisation as $\tau \propto N^2$. The fact that τ_S and τ_1 have the same scaling with N is as one would expect from a continuous transition.

Our attempts to apply the middle kinetic stage laws (4.59) and (4.60) to the kinetics at other collapse transitions are presented in Table 4.6. The typical evolutions of the squared radii of gyration R_g^2 during that kinetic stage for different kinetic processes are presented in Fig. 4.35. One can see that the relaxation processes in the phase III are very slow compared to those in other collapse regions. After small anisotropic globules have been created during the first kinetic stage, their further unification is inhibited due to different orientations in neighbouring globules. Thus, for this case the final value of the squared radius of gyration $R_g^2(\infty)$ exhibited in Table 4.6 is still far enough from the one obtained from equilibrium calculations. Furthermore, the numerical values of the total collapse times τ_S are very large. We must point out that this orientational frustration of locally collapsed globules is exacerbated by the underlying lattice. However, we expect this motif to persist even in continuum calculations.

Figure 4.35: Plots of the mean squared radius of gyration, $3R_g^2$, vs time, t , for different collapse transitions.

Here the degree of polymerisation is $N = 480$. Labels denote the initial and final point in the phase diagram according to Fig. 4.31.



On the contrary, the kinetic process in the phase II does not lead to metastable states. The reason seems to lie in the fact that the anisotropic accumulations are quite unstable in that region of the interaction parameters. Thus, some rearrangements of the internal structure make possible further unification of clusters. The simulation data appear to be consistent with the kinetic laws proposed earlier, but with new exponents. Without offering any theoretical interpretation we present them below,

$$\begin{aligned} \tau_1 &\propto N^\gamma, & \gamma &= 2.51 \pm 0.06 \approx 5/2, \\ \tau_S &\propto N^{\gamma'}, & \gamma' &= 2.32 \pm 0.05 \approx 7/3. \end{aligned}$$

It is interesting to note that the characteristic and total collapse times (see Eqs. (4.59) and (4.60)) scale with the degree of polymerisation in a different manner here.

Chapter 5

Infinitely Dilute Heteropolymer Solutions

In this chapter we study the equilibrium and kinetics of amphiphilic heteropolymers in infinitely dilute solutions, i.e. we consider a single heteropolymer chain and neglect the inter-chain interactions. Our studies here are mainly based on the numerical solution of the Gaussian self-consistent equations (2.73) with the matrix of the second virial coefficients, $u_{ij}^{(2)}$, parametrised by Eq. (2.15). The knowledge of the free energy allows us to determine the equilibrium phase diagram of various hydrophilic–hydrophobic copolymers consisting of alternating monomers, alternating blocks and random sequences. We also analyse the behaviour of the main observables in different phases. The folding kinetics after an instantaneous quench is studied in terms of the changes in nonequilibrium conformations and observables.

5.1 Phase Diagrams of Amphiphilic Heteropolymers

In this section we study the equilibrium properties of amphiphilic heteropolymers. For simplicity let us consider copolymers with the ring topology, though the current treatment may be easily extended for study of copolymers with any other topology just by changing the spring term in Eqs. (2.8, 2.43).

Typical phase diagrams in terms of the mean second virial coefficient, $\bar{u}^{(2)}$, and the amphiphilicity, Δ , in Eq. (2.15) for some “random” and periodic sequences are presented in Figs. 5.1, 5.2 and 5.3. In the region $\bar{u}^{(2)} > 0$ and for small values of amphiphilicity, $\Delta < 5$, typical conformations of copolymers are akin to the homopolymer extended coil. By decreasing $\bar{u}^{(2)}$ to the negative region the chain undergoes the continuous collapse transition, similarly to what we observed in Sec. 4.2. The collapse transition is characterised by a rapid fall of the radius of gyration, R_g^2 , (2.44) and the change of the inverse fractal dimension, ν (see Sec. 4.2).

The globular state for large values of Δ is different from the liquid-like globule. Proceeding from the collapsed globule at a fixed negative $\bar{u}^{(2)}$ the increase of the amphiphilicity Δ causes a continuous transition to the micro-phase separated (MPS) globule. It is characterized by a somewhat higher value of the radius of gyration (see Fig. 5.4a), lower energy and extremely large value of the MPS order parameter (2.50), as depicted in Fig. 5.4b. Thus we call this state the MPS globule. The change in radius of gyration during the MPS transition is most pronounced for sequences consisting of long alternating blocks, as e.g. sequence $(a_3b_3)_10$ (see Fig. 5.4a). For other sequences the connectivity constraints impede formation of structures with a pronounced hydrophilic shell and hydrophobic core. The MPS order parameter Ψ in Fig. 5.4b increases almost linearly with Δ

Figure 5.1: Phase diagrams of “short” blocks copolymer sequences $(ab)_{15}$ (Fig. a) and $(ab)_{30}$ (Fig. b) in terms of the mean second virial coefficient, $\bar{u}^{(2)}$, and the amphiphilicity, Δ .

Curves (C) and (S) correspond respectively to the collapse and the micro-phase separation continuous transitions. Curves (I) and (II) correspond to discontinuous transitions to the frustrated phases. “Spinodal” curves (I’) and (II’’) bound the regions of metastability of the frustrated states. Transition curves and boundaries distinguishing different frustrated states are not depicted.

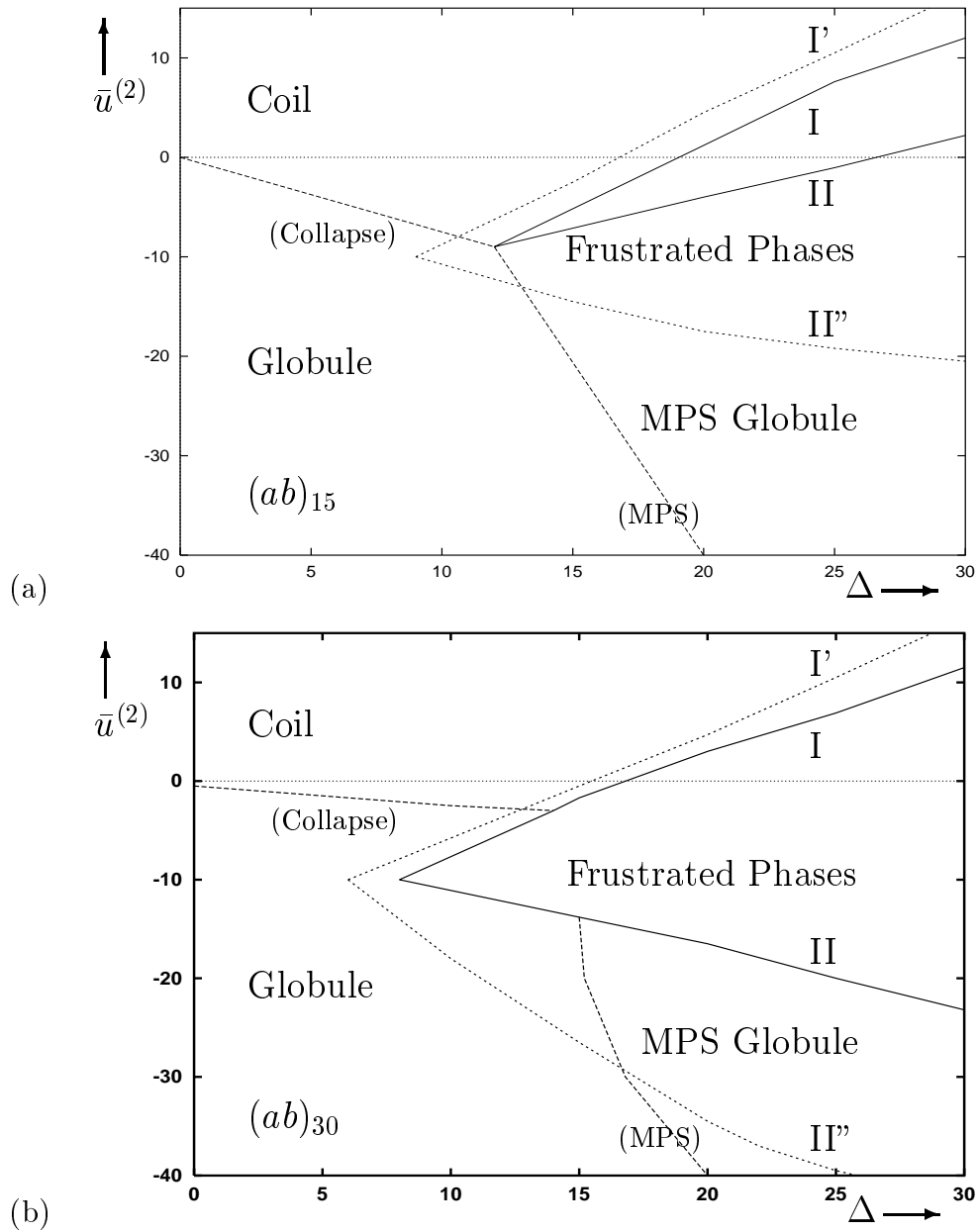


Figure 5.2: Phase diagrams of “long” blocks copolymer sequences $'(a_3b_3)_5'$ (Fig. a) and $'(a_3b_3)_{10}'$ (Fig. b) in terms of the mean second virial coefficient, $\bar{u}^{(2)}$, and the amphiphilicity, Δ .

For large values of Δ the collapse transition becomes discontinuous (curve (I)) and it is accompanied by micro-phase separation (see also Fig. 5.6). Curves (I') and (I'') are spinodals.

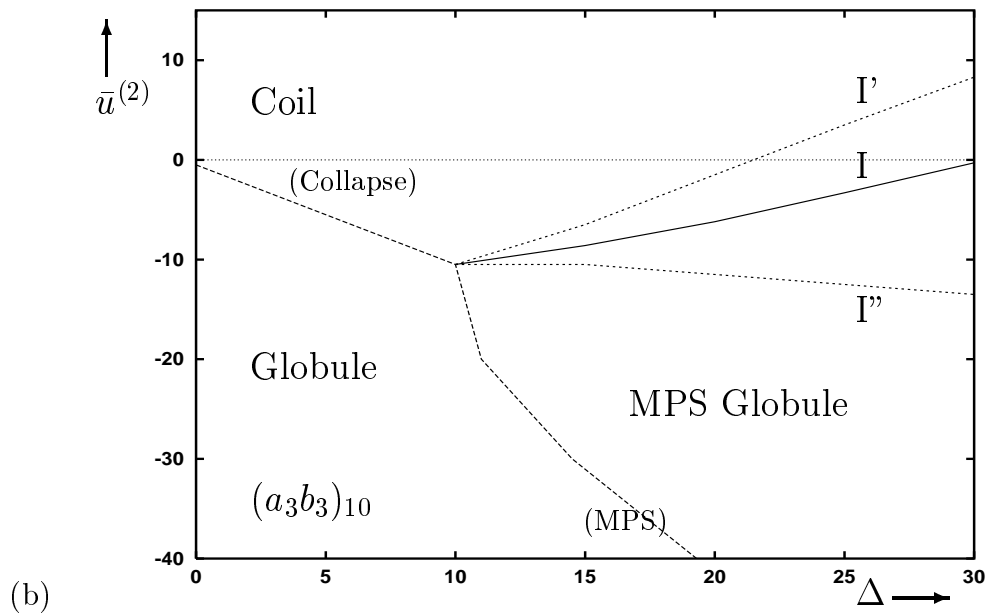
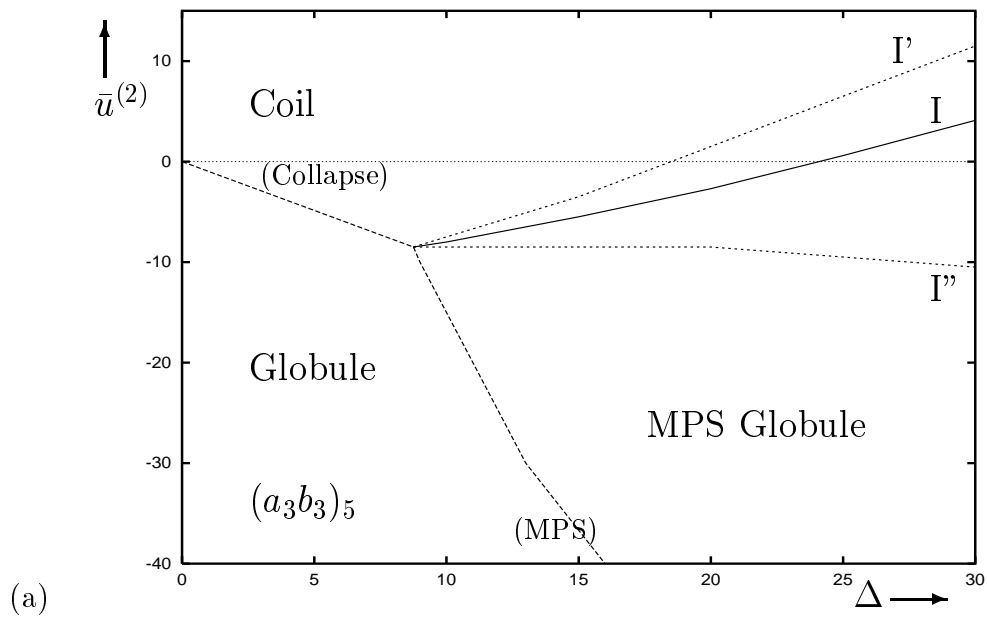


Figure 5.3: Phase diagrams of copolymers with the “random” sequence $'babca_2cbac_2acb_3cac_3a_2b_2cac_2b'$ (Fig. a) and the sequence twice as in Fig. a $'(babca_2cbac_2acb_3cac_3a_2b_2cac_2b)_2'$ (Fig. b) in terms of the mean second virial coefficient, $\bar{u}^{(2)}$, and the amphiphilicity, Δ . See caption to Fig. 5.1 for more details.

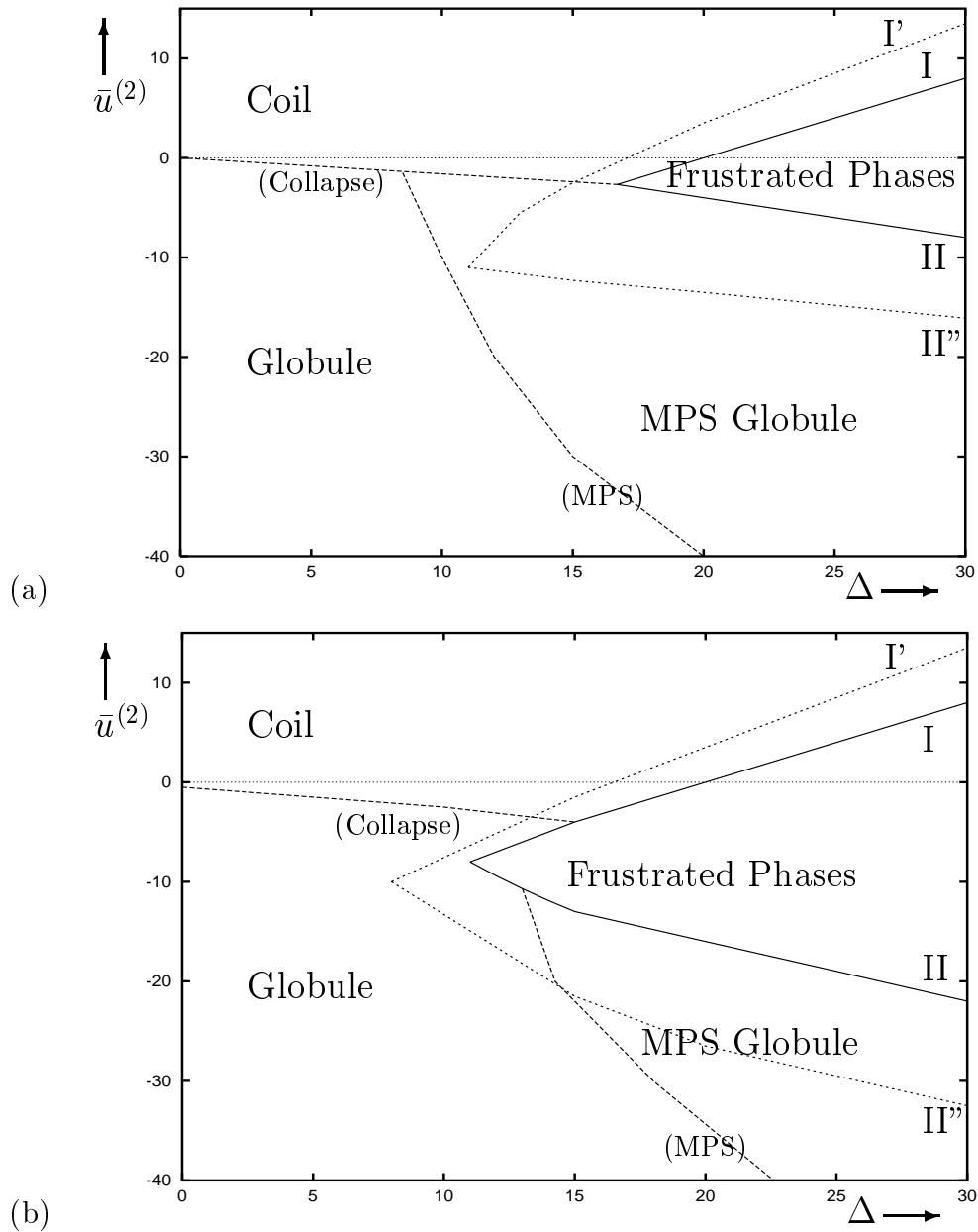
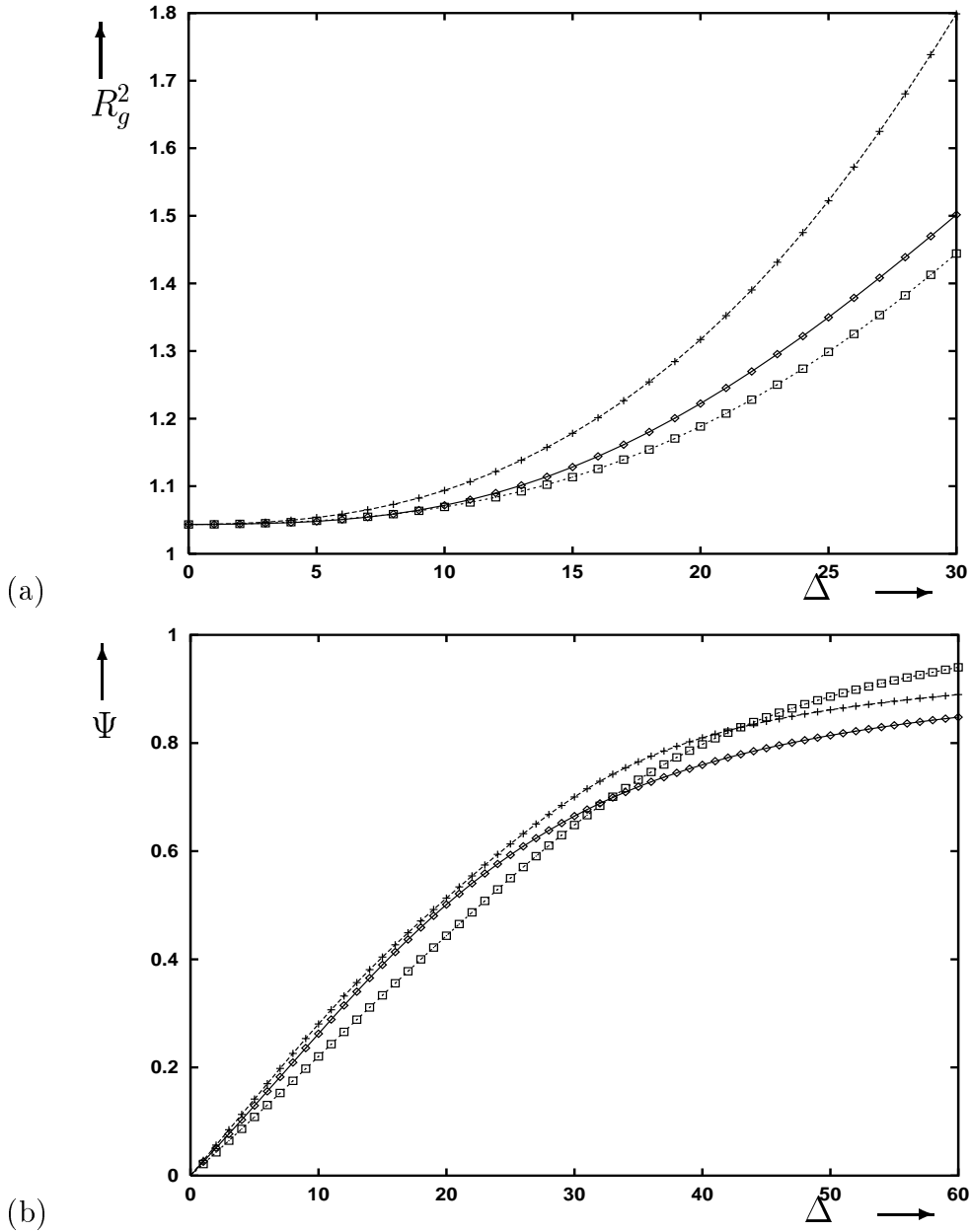


Figure 5.4: Plots of the mean squared radius of gyration, R_g^2 , (Fig. a) and the parameter of micro-phase separation, Ψ , (Fig. b) vs the amphiphilicity, Δ . Here we have fixed $\bar{u}^{(2)} = -40$ and the lines correspond respectively to the following copolymer sequences: “long” blocks $(a_3b_3)_{10}$ (pluses), “short” blocks $(ab)_{30}$, (diamonds) and the “random” $(babca_2cbac_2acb_3cac_3a_2b_2cac_2b)_2$ sequence (squares).



for small values of amphiphilicity, and then after the transition asymptotically saturates.

In series of pictures in Fig. 5.5 we exhibit typical copolymer conformations from the lattice Monte Carlo simulation for various values of the degree of hydrophobicity. In the hydrophobic–neutral copolymer model we set $\chi_{ab} = \chi_{aa}/2$ and $\chi_{bb} = 0$, i.e. a change in the number of bb contacts on the lattice does not change the energy. At low amphiphilicity in Fig. 5.5a the hydrophobic forces are too weak to cause polymer collapse and typical copolymer conformations are quite similar to those of the homopolymer Flory coil (see Figs. 4.13a, 4.17a). The micro–phase separation here is extremely weak and can hardly be observed visually. At higher values of the amphiphilicity, when ‘ a ’ monomers become hydrophobic enough, the copolymer chain folds into the MPS globule in Fig. 5.5b, where the separation into a hydrophobic core and hydrophilic shell is quite distinguishable.

The transition from the coil to the MPS globule for larger values of amphiphilicity turns out to be more complicated, and essentially dependent on the sequence. In the case of long blocks (Fig. 5.2) the collapse transition to the MPS globule becomes discontinuous (first–order–like). The spinodals I’ and I’’ designate the region where two distinct states corresponding to the coil and the MPS globule can be found. The former minimum is characterised by large size, R_g^2 , (on the right of Fig.5.6a) and small MPS order parameter Ψ in Fig. 5.6b, while for the latter minimum the situation is reversed (see left hand side of Figs. 5.6a,b). The depths of the free energy minima become exactly equal on the transition curve I in Fig. 5.2.

However, for a wide class of sequences, for example for periodic in Fig. 5.1 and aperiodic sequences in Fig. 5.3, the phase diagram at large amphiphilicity, Δ , is much more complicated. Starting from some value of Δ in some intermediate

Figure 5.5: Typical conformations of $(a_6b_6)_{20}$ hydrophobic-neutral copolymer at equilibrium from the lattice Monte Carlo simulation.

Figs. (a) and (b) correspond to the extended Flory coil, $\chi_{aa} = 0.2$, and the MPS globule, $\chi_{aa} = 1.2$. In Fig. (b) the globule is magnified 3 times compared to Fig. (a) and Fig. 5.14.

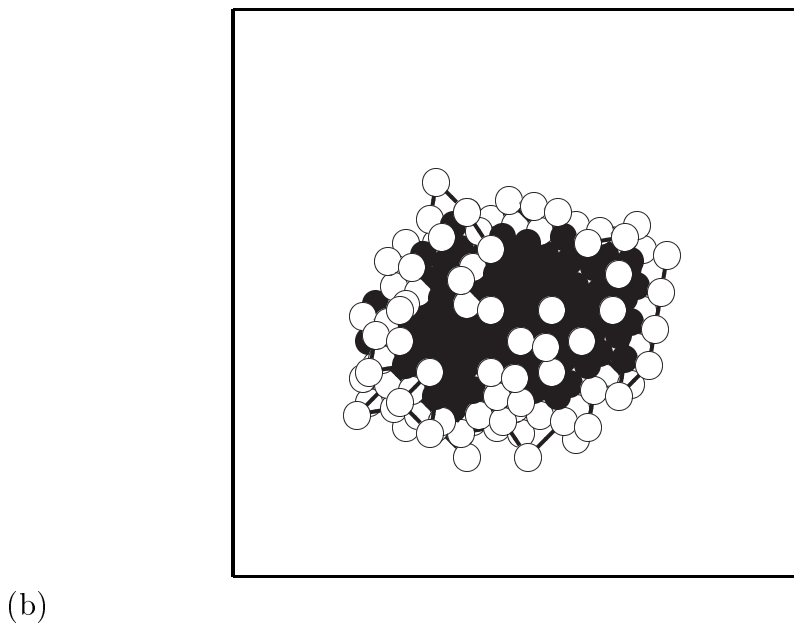
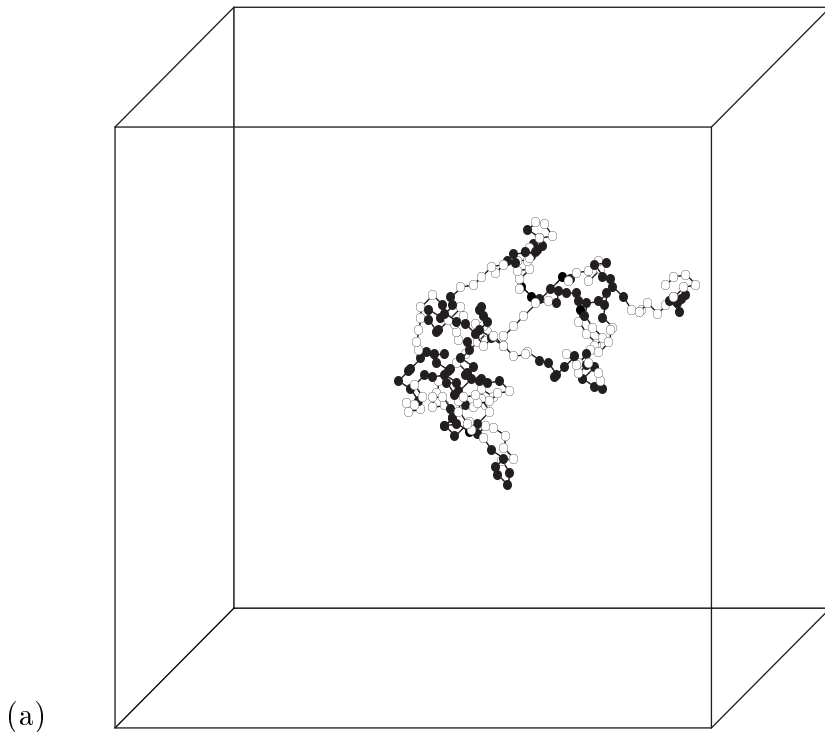
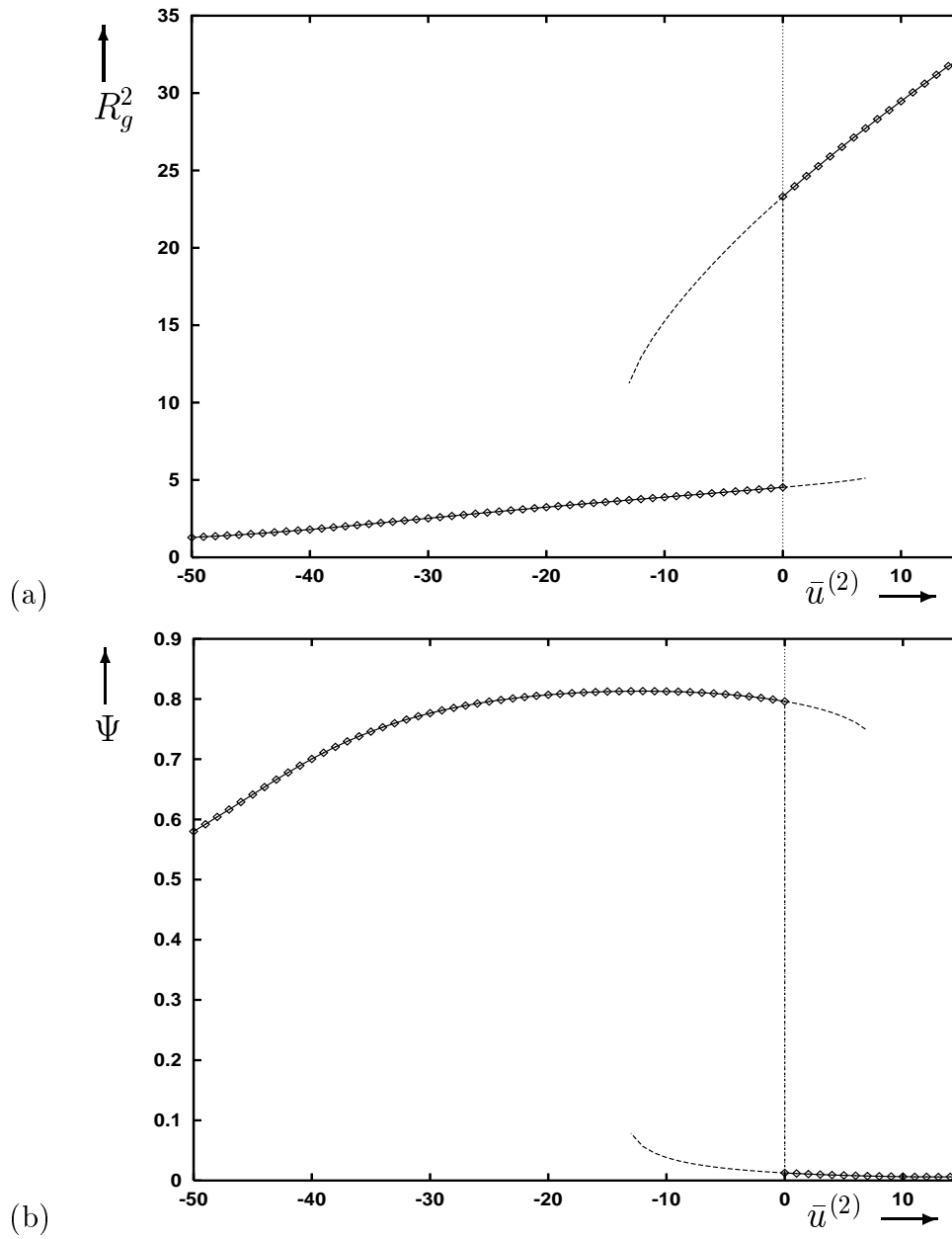


Figure 5.6: Plots of the mean squared radius of gyration, R_g^2 , (Fig. a) and the parameter of micro-phase separation, Ψ , (Fig. b) vs the mean second virial coefficient, $\bar{u}^{(2)}$, for '(a₃b₃)₁₀' copolymer. See caption to Fig. 5.7 for more details.



region of $\bar{u}^{(2)}$ there appear additional solutions corresponding to local minima of the free energy. The broad region where this could take place is bounded by the curves I' and II'' in Figs. 5.1, 5.3. With increasing Δ the number of such solutions grows quickly. Significantly, in the region of the phase diagram, between curves I and II in Figs. 5.1, 5.3, some of these possess the lowest free energy value, thus being the thermodynamically stable state. Since the number of such solutions is rather high even for short sequences and their number grows quickly with the chain length, we do not attempt to draw all their boundaries of (meta)stability. We shall call them collectively *frustrated phases*, explaining this terminology below.

Now let us compare the phase diagrams in Figs. 5.3a and 5.3b, the latter corresponding to the sequence twice longer than the former. An interesting observation is that the region between spinodals I' and II'', designating where the frustrated phases can exist, expands dramatically with increasing chain length. The same concerns the region of thermodynamically stable frustrated phases between curves I and II. More exactly, these regions expand downwards and to the left, so that the position of curves I' and I change slightly with increasing N , whilst curves II and II'' depend significantly on the size of the system. This observation is also confirmed by comparing phase diagrams of periodic copolymers, consisting of '(ab)' blocks in Figs. 5.1a, b. For rather long chains we may expect that the regions of stability and metastability of the frustrated phases will continue to expand downwards and to the left, so that the lines II and II'' will become nearly vertical, displacing the region of stable MPS globule. Probably, for most of long heteropolymer chains the MPS globule does not exist as thermodynamically stable state, becoming stable only for some special sequences. Unfortunately, we can not proceed with numerical solution for much larger system sizes, N , since the calculational expenses grow in N as N^3 per iteration and also the total number

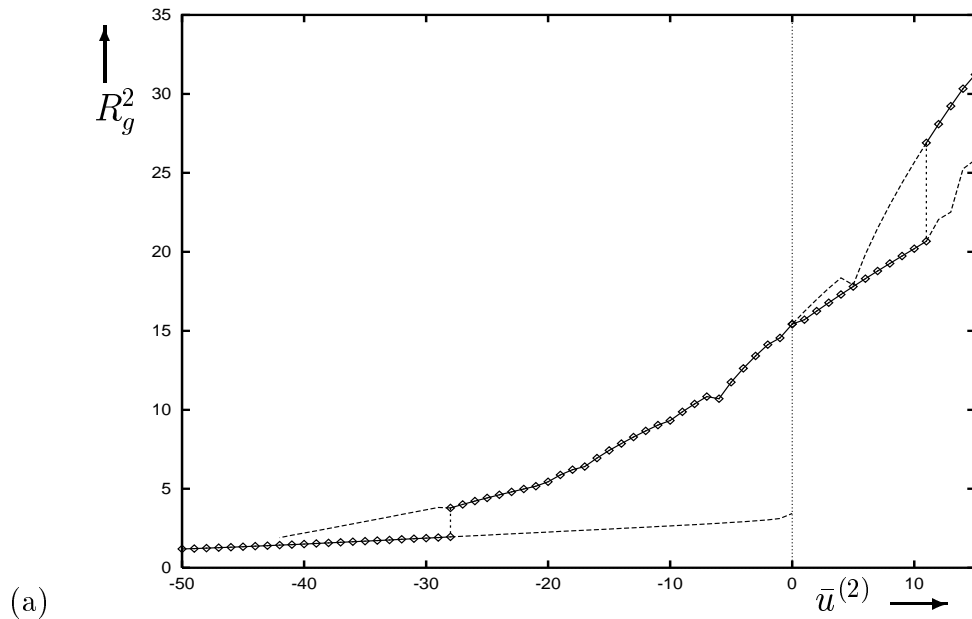
of frustrated states becomes huge for large system sizes. This diversity and a special foliating structure of various branches leads in the thermodynamic limit to what is known as a spin glass *frozen* phase [32] of random copolymers.

Let us consider the conformational structure of the frustrated states for the copolymer consisting of repeating 'ab' blocks, the phase diagram of which also exhibits the thermodynamically stable frustrated phases (see Figs. 5.1a, b). In Fig. 5.7 we present the dependence of the observables in (2.44, 2.50) at a quasi-static change of the mean second virial coefficient, $\bar{u}^{(2)}$, from the coil state to the MPS globule and back. From these pictures one can see in the intermediate region only a few (seven to be precise) of all possible frustrated phases. The values of the radius of gyration and the MPS order parameter are intermediate for these solutions, lying between those of the coil and the MPS globule. In this sense, we can call them non-fully compacted and misfolded states. The general behaviour of the size R_g^2 and MPS parameter Ψ , which in a sense characterise the system in total, are rather similar in the case of aperiodic sequences (see Fig. 5.8). However, the boundaries, where the frustrated states exist and are stable, number as well as particular local structure of those frustrated phases are quite sensitive to the sequence.

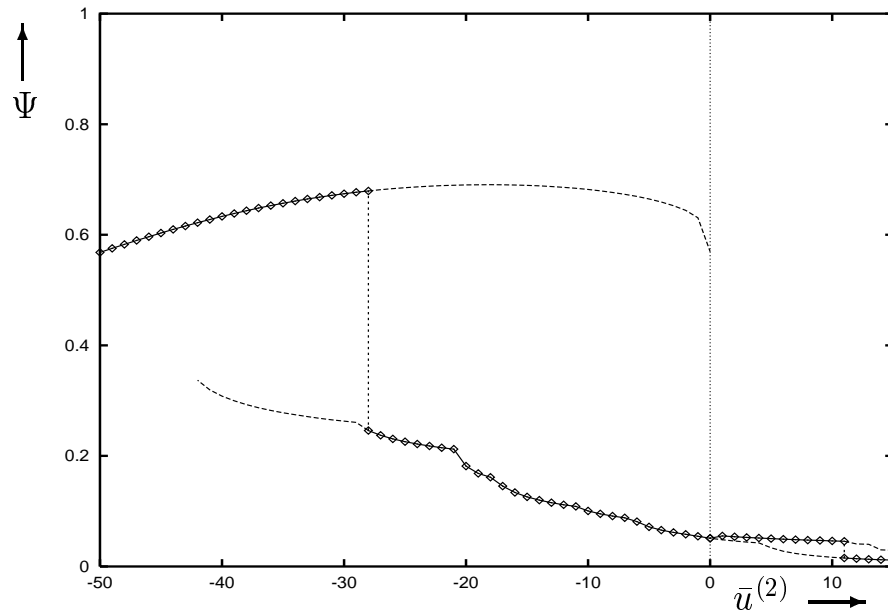
In the series of pictures in Fig. 5.9 we present the matrix of mean squared distances, D_{ij} , for the copolymer, consisting of 'ab' blocks at some values of the mean second virial coefficient, $\bar{u}^{(2)}$. The set of D_{ij} in the GSC method completely determines the conformational structure of any, equilibrium or kinetic, state. For positive $\bar{u}^{(2)}$ (see Fig. 5.9a) the mean squared distances possess the structure typical for the extended coil. The elements of the matrix, D_{ij} , increase monotonically on moving away from its diagonals towards the distance of half-ring along the chain. Thus, the D_{ij} matrix may be approximated here by a monotonically increasing function of the chain distance, $|i - j|$. Decreasing the mean second

Figure 5.7: Plots of the mean squared radius of gyration, R_g^2 , (Fig. a) and the parameter of micro-phase separation, Ψ , (Fig. b) vs the mean second virial coefficient, $\bar{u}^{(2)}$, for '(ab)₃₀' copolymer.

Here and in Figs. 5.6 and 5.8 $\Delta = 30$, solid lines correspond to values of observables in the main free energy minimum and dashed lines — in metastable minima.



(a)



(b)

Figure 5.8: Plots of the mean squared radius of gyration, R_g^2 , (Fig. a) and the parameter of micro-phase separation, Ψ , (Fig. b) vs the mean second virial coefficient, $\bar{u}^{(2)}$, for '(*babca₂cbac₂acb₃cac₃a₂b₂cac₂b*)₂' copolymer. See caption to Fig. 5.7 for more details.

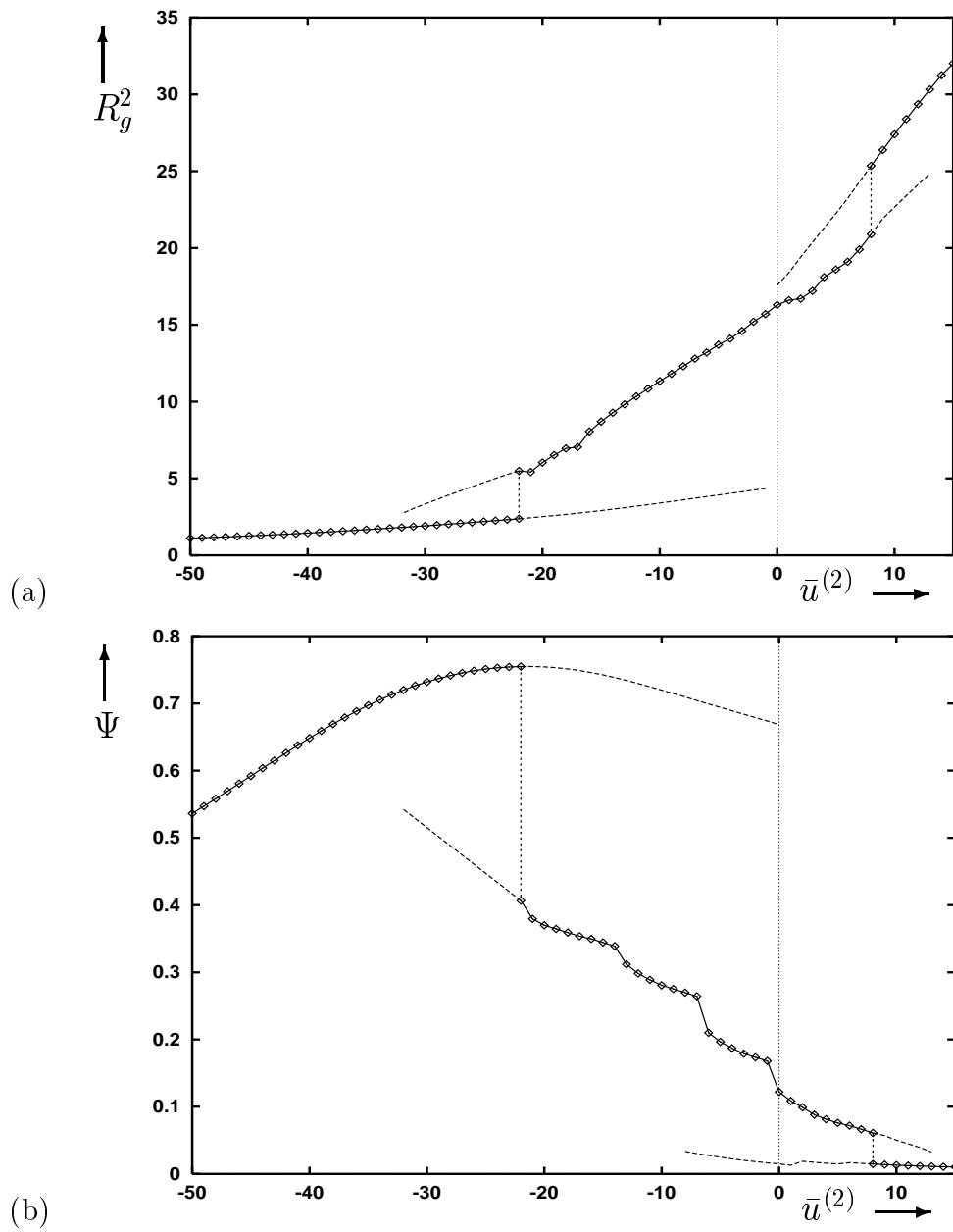
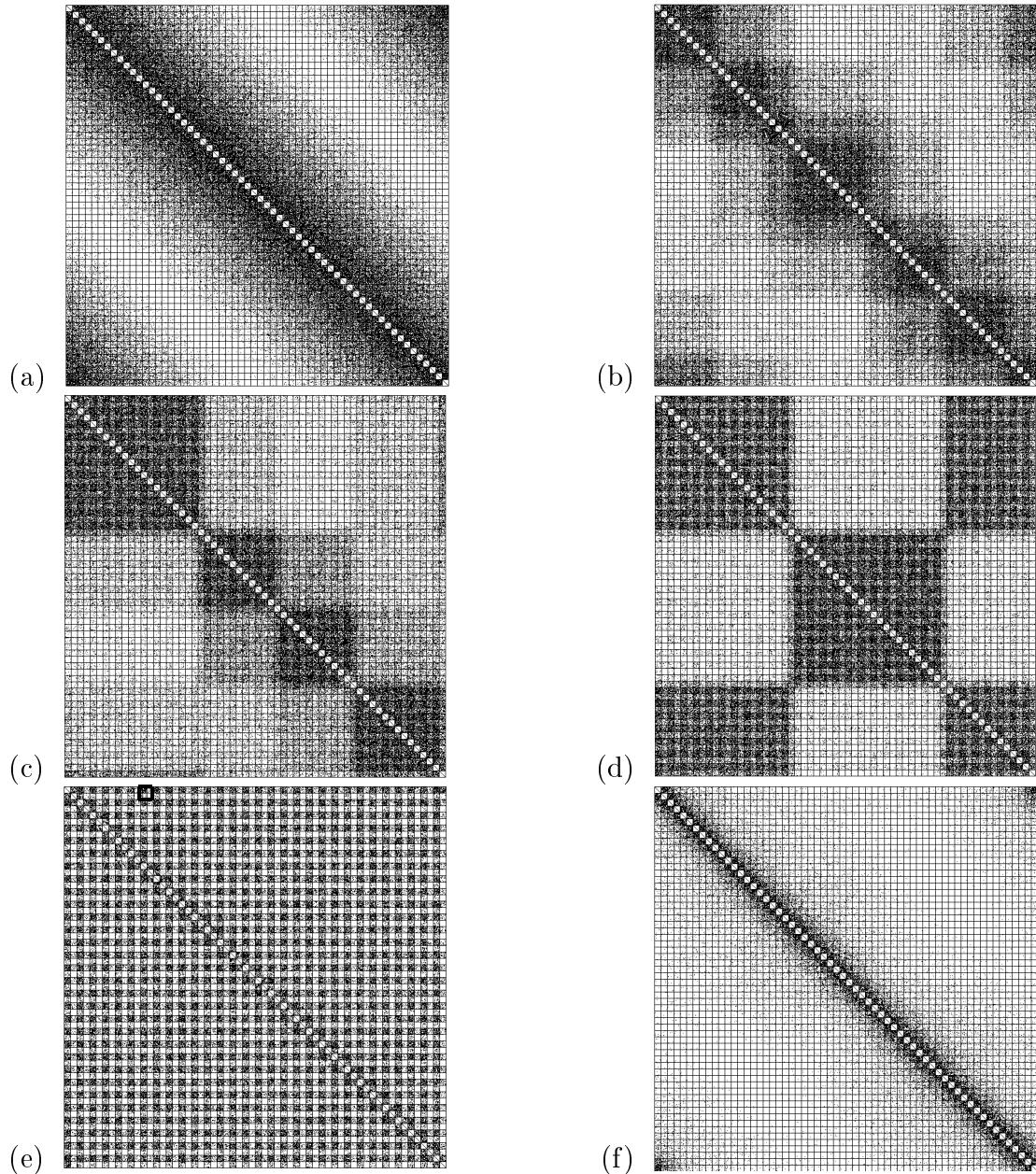


Figure 5.9: Diagrams of the mean squared distances matrix, D_{ij} , for '(ab)₃₀' copolymer.

Diagrams (a-e) correspond respectively to the values of the mean second virial coefficient, $u^{(2)} = 15, -11, -21, -30$ and -40 and amphiphilicity $\Delta = 20$. Diagram (f) corresponds to the homopolymer globule, for which $u^{(2)} = -25$ and $\Delta = 0$. Monomer indices i, j start counting from the upper left corner. Each matrix element, D_{ij} is denoted by a quadratic cell with varying degree of black colour, the darkest and the lightest cells corresponding respectively to the smallest and to the largest mean squared distances.



virial coefficient, $\bar{u}^{(2)}$, causes the copolymer to pass through frustrated states, in Figs. 5.9b-5.9d, finally reaching the MPS globule state (see Fig. 5.9e). The characteristic feature of the D_{ij} matrix in a frustrated state is that it possesses some number of monomer groups having smaller distances between each other than between monomers from other groups. Clearly, such a group represents a cluster of monomers, so that here the copolymer chain forms a set of clusters (approximately 8 clusters in Fig. 5.9b, 4 clusters in Fig. 5.9c and 2 clusters in Fig. 5.9d), each consisting of monomers nearest along the chain.

The internal structure of each cluster is similar to the structure of the final MPS globule. For the simplest copolymer sequence presented here, the D_{ij} matrix away from the diagonal can be divided into sub-matrices of the size 2×2 (see top of the Fig. 5.9e), each of which possesses approximately the same structure: small values in the upper left corner correspond to the mean squared distances between hydrophobic monomers; large values in the lower right corner, which are the mean squared distances between hydrophilic monomers; and off-diagonal elements are nearly equal and correspond to the distances between different species. Thus, for the MPS globule, such structure of the D_{ij} matrix reflects the structure of the two-body interaction matrix, $u_{ij}^{(2)}$. Obviously, the higher correlation between these two matrices is manifested in the higher value of the MPS order parameter. For another, more complicated, sequences the D_{ij} matrix in the MPS globular state has a different structure, but still it resembles in some way the interaction matrix, $u_{ij}^{(2)}$ away from the diagonal, being distorted by the spring interactions in the elements close to the diagonal. Note that for the liquid-like globule the pattern of the D_{ij} matrix looks sufficiently trivial. Namely, it has a narrow band of dark color along the main diagonal with its intensity quickly decreasing, as one can see from Fig. 5.9f.

From Fig. 5.9 one can see that the interaction symmetries imposed by the

EFEF (2.14) are spontaneously broken in the region of the phase diagram corresponding to the frustrated phases. Obviously, in our case the number of possible ways to break the symmetry is enormously huge and moreover it grows exponentially with increasing system size. Despite the kinematic symmetries are not present for arbitrary sequences, the structure of the phase diagram (see Fig. 5.3) and behaviour of main observables remain very similar. It is the particular structure of D_{ij} matrix, number and the shape of boundaries of frustrated phases that are quite sensitive on the sequence. The symmetry that may be broken in this case has a subtler meaning and may be expressed in terms of the replica formalism [32]. Consider, for example, two nearly identical blocks with nearly identical surroundings. In the coil and the MPS globular states one should expect the conformational matrices of these blocks to be nearly equal to each other. On the contrary, the small difference in the interactions in the region of frustrated phases may lead to a huge difference in the conformations. Thus, it is by no means surprising that the replica symmetry breaking in the case of periodic systems takes such an explicit manifestation in the breaking of the block translational symmetry.

An important point here is that the frustrated phases become dominant in an intermediate region of the phase diagram not due to a low mean energy, but mostly due to a higher entropy. The MPS globule is entropically unfavourable there because the overall shrinking force is insufficiently strong. Also, in the regime of nearly compensating repulsion and attraction between monomers it is more preferable to achieve phase separation on a smaller, than the globular, scale by forming clusters.

Finally, let us comment on the phase diagram of the “long” blocks (Fig. 5.2). The micro-phase separation is obviously easier in this case and it dominates for large values of Δ , so that the frustrated phases are suppressed. We found that,

qualitatively, in order to form a frustrated phase, in a finite range of Δ values, the number of (not necessarily identical) pieces of the chain with competing interactions should be larger than some critical number, in principle, weakly dependent on N . Here are a few examples conforming to this qualitative criterion: the phase diagram of $'(ab)_{10}'$ behaves roughly as for $'(a_3b_3)_{10}'$, but for $'(ab)_{15}'$ it behaves similarly to $'(ab)_{30}'$; while for $'(a_2b_2)_{15}'$ the phase diagram looks like for $'(a_3b_3)_{15}'$ at small and moderate values of Δ , it becomes as for $'(ab)_{30}'$ for much larger values of the amphiphilicity. It is reasonable to conjecture therefore that for a large number of $'(a_3b_3)'$ blocks, as well as for extremely high values of Δ and just 10 blocks, the frustrated phases may be found again.

5.2 Kinetics of Folding

Here we shall consider the time evolution of the conformational state of a copolymer away from the initial equilibrium after it has been subjected to an instantaneous temperature jump that causes the two-body interaction parameters in Eq. (2.15), mean second virial coefficient, $\bar{u}^{(2)}$, and amphiphilicity, Δ , to change. The composition of copolymer $\{\sigma\}$ and the rest of parameters remain the same and do not change with time. We are interested in quenches from the homopolymer coil, where all monomers are equally hydrophilic ($\bar{u}^{(2)} > 0$ and $\Delta = 0$), to the region of parameters corresponding to the MPS globular state, so that the 'a' species became strongly hydrophobic and the 'b' species remained nearly neutral ($\bar{u}^{(2)} \ll 0$ and $\Delta \approx |\bar{u}^{(2)}|$). Here we consider some binary copolymer systems, consisting of the same number of 'a' and 'b' monomers and of the same size, $N = 50$, and quench, $(\bar{u}^{(2)} = 15, \Delta = 0) \rightarrow (\bar{u}^{(2)} = -35, \Delta = 30)$.

Let us first discuss the general features of the kinetics of the copolymer folding. The time evolution of the mean squared radius of gyration, $R_g^2(t)$, the MPS order parameter, $\Psi(t)$, and the instantaneous free energy, $\mathcal{A}(t)$, are presented

Figure 5.10: Plot of the mean squared radius of gyration, R_g^2 , vs time, t , during kinetics after the quench from $\bar{u}^{(2)} = 15$, $\Delta = 0$ to $\bar{u}^{(2)} = -35$, $\Delta = 30$ for copolymer sequences with $N = 50$.

Lines A-F in the figures (a) and (b) correspond respectively to the following sequences: ' c_{50} ' (homopolymer), ' $(ba)_{25}$ ', ' $a_2b_3a_3b_2a_3b_2a_2b_3(ba)_{15}$ ', ' $(ba)_3b_2(ba)_9a_2(ba)_5b_2a_2(ba)_4$ ', ' $(b_5a_5)_5$ ' and ' $b_6a_4b_5a_5b_4a_6b_3a_7b_7a_3$ '.

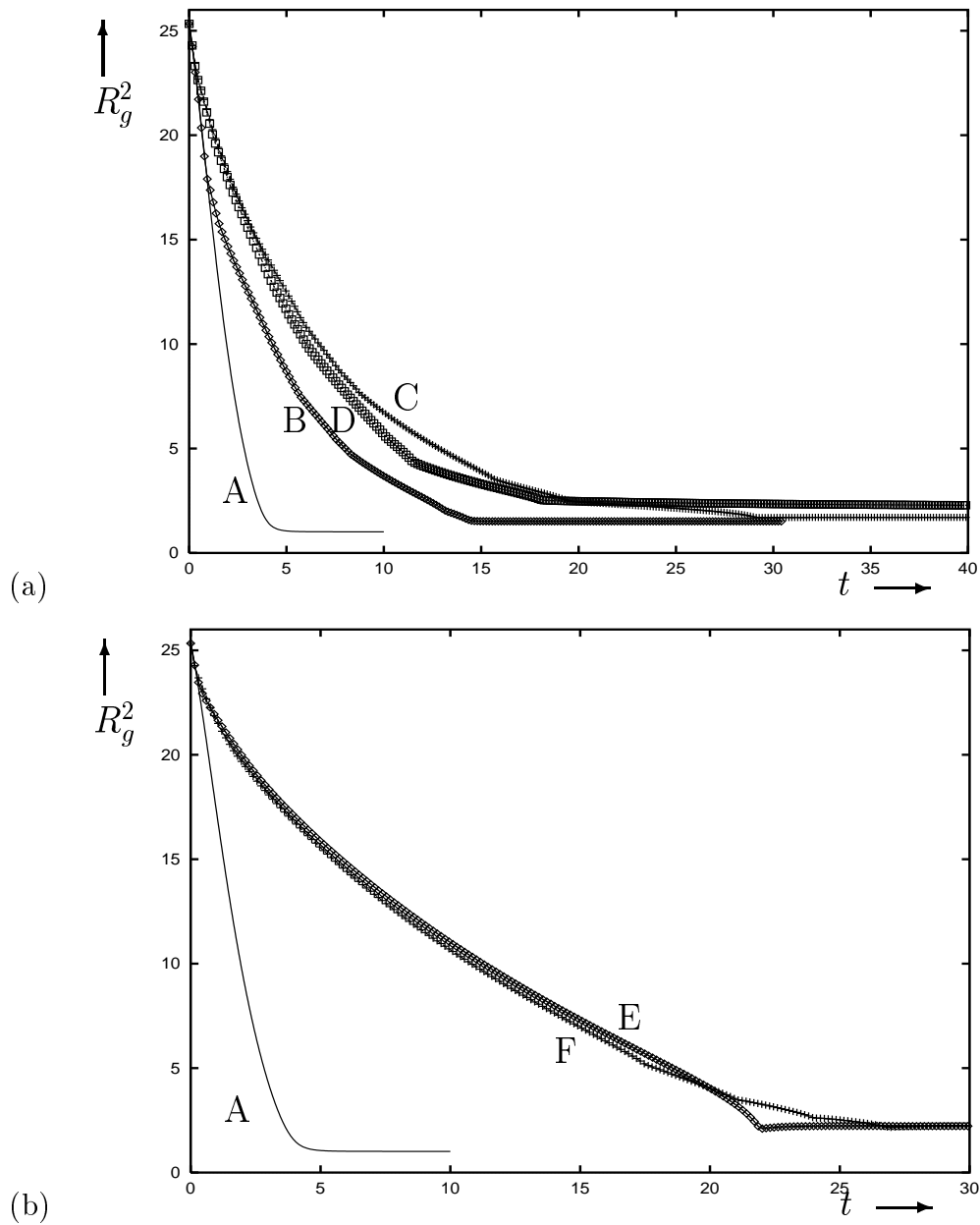


Figure 5.11: Plot of the MPS order parameter, Ψ , vs time, t , during kinetics after the same quench and for the same copolymer sequences as in Fig. 5.10.

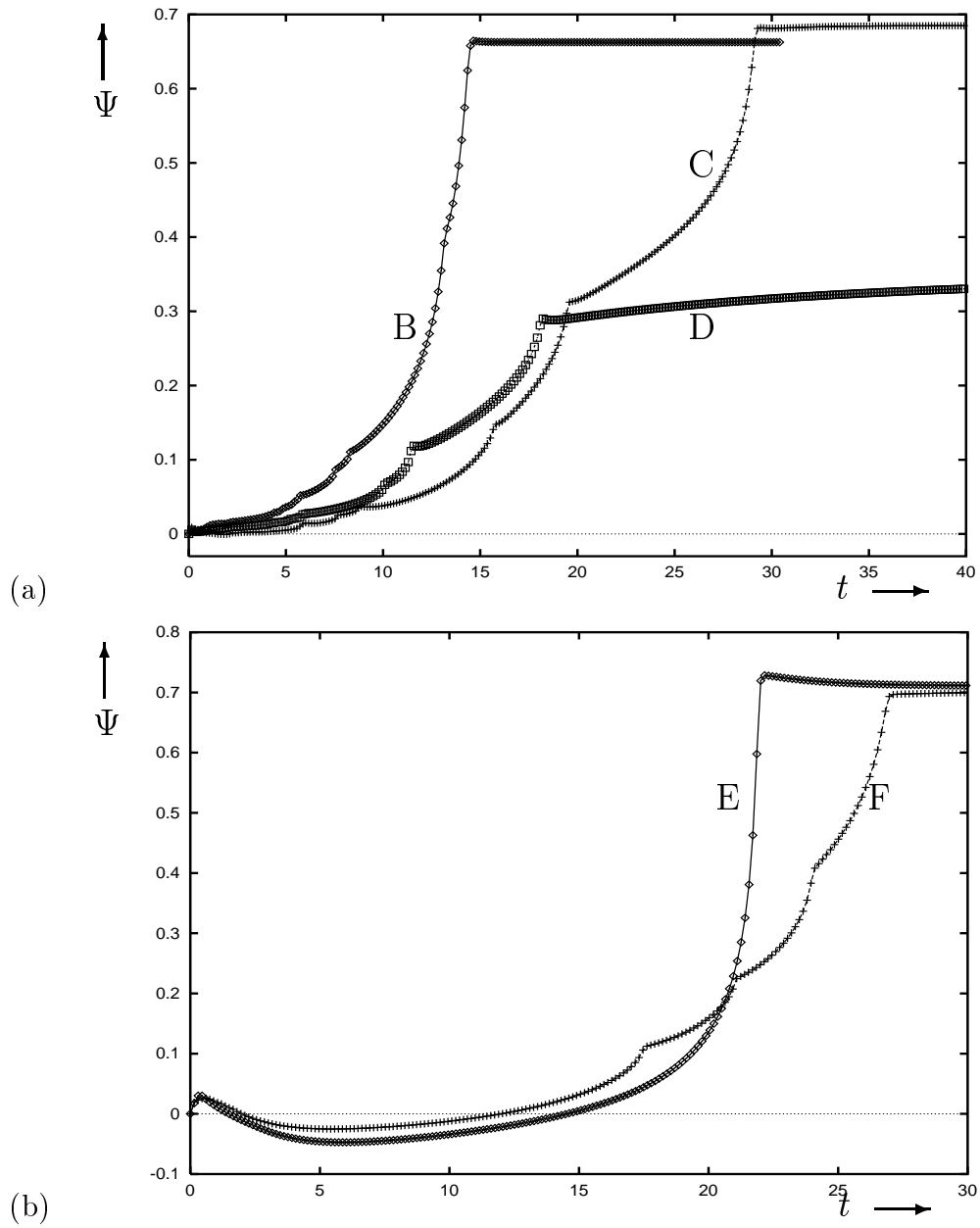
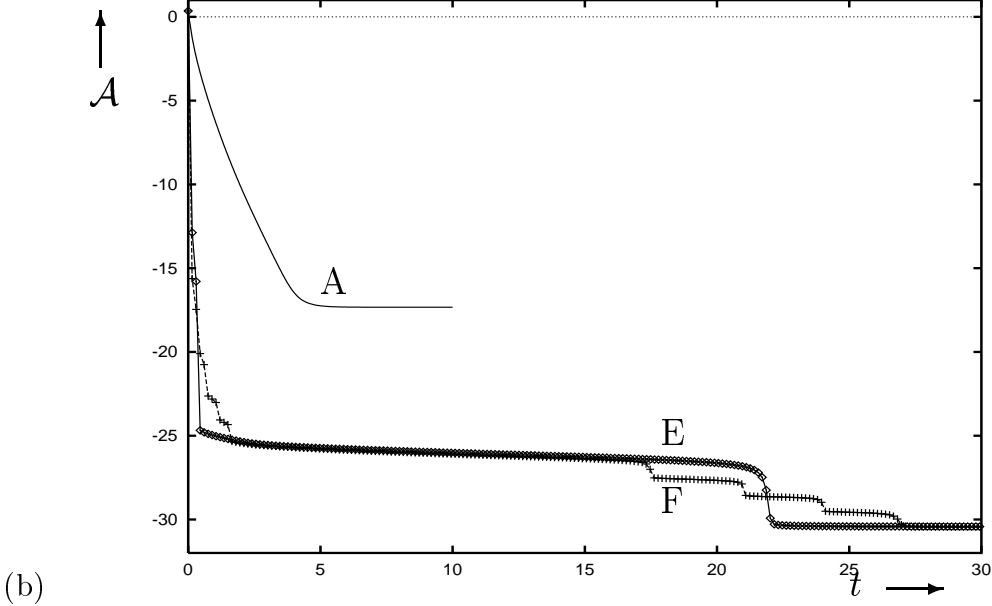
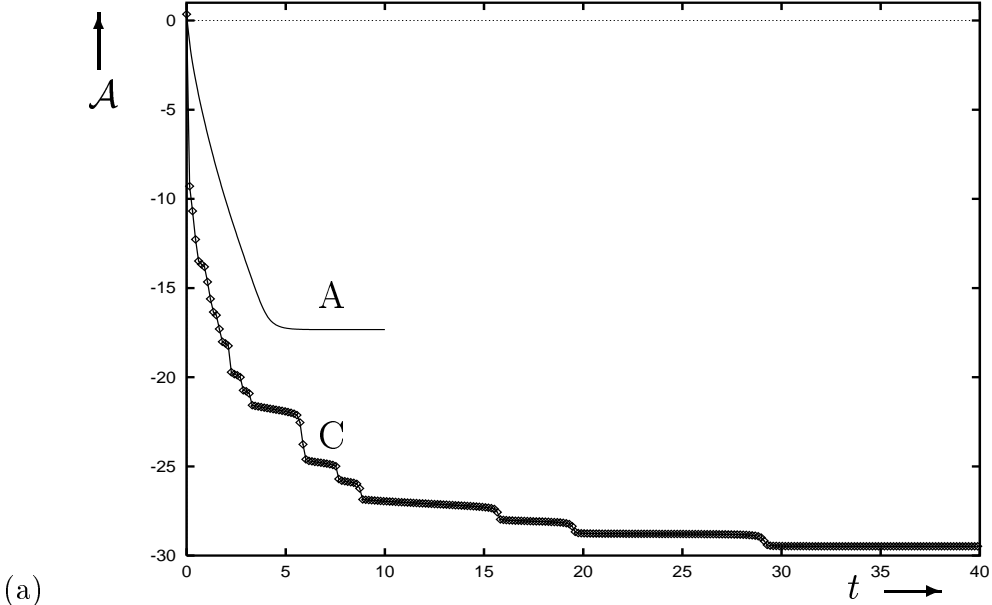


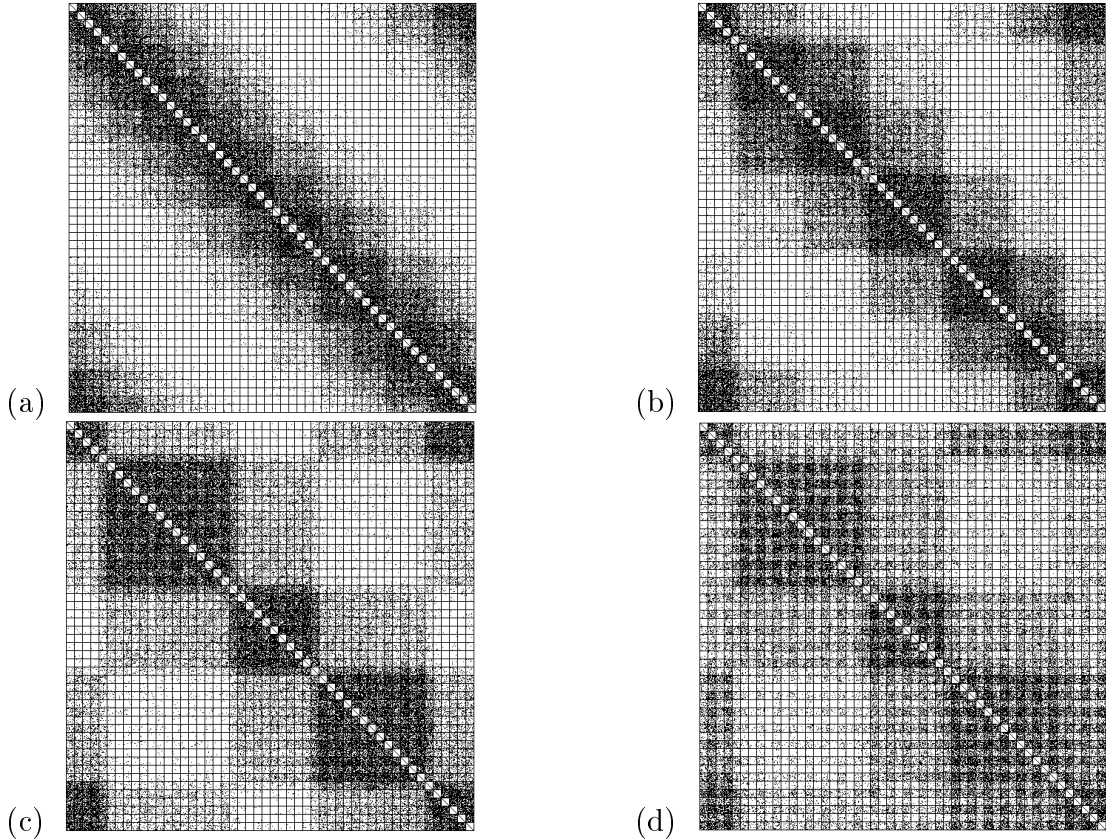
Figure 5.12: Plot of the instantaneous free energy, \mathcal{A} , vs time, t , during kinetics after the same quench and for the same copolymer sequences as in Fig. 5.10.



in Figs. 5.10, 5.11 and 5.12. Here lines A correspond to the homopolymer and serve for reference purposes. In case of the homopolymer one can see that R_g^2 and \mathcal{A} decrease monotonically to their equilibrium values corresponding to the liquid-like globule state, while Ψ remains identically zero. As for the copolymer kinetics, the first observation consists in that it proceeds much slower than for the homopolymer. For example, for the simplest copolymer sequence, $'(ba)_{25}'$, the total collapse time (see Sec. 4.3) is more than 3 times longer than that of the homopolymer, other copolymer sequences collapsing even longer. The total collapse time also seems to be quite sensitive on the copolymer sequence. Since the micro-phase separation is one of the main factors in collapse of copolymers, the MPS order parameter, Ψ , grows in kinetics most of the time for all considered sequences, though rather nonmonotonically. Nevertheless, for some sequences it might be negative during some time in kinetics (see e.g. sequences E and F in Fig. 5.11b), something we never observed at equilibrium.

The evolution of the instantaneous free energy, $\mathcal{A}(t)$, depicted in Figs. 5.12a, 5.12b, is most unusual. Typically it proceeds through multiple accelerations and decelerations. The flat regions of a staircase-like function correspond to temporary kinetic arrest of the system in transient nonequilibrium conformations, i.e. to transient trappings of various members of the ensemble in their local shallow energy minima. Since such minima are encountered at different moments in time for different members of the ensemble, their influence on the overall time evolution of averaged observables is manifested in a smooth characteristic slowing down. Note here that the number of steps in kinetics process hardly can be guessed from the given primary sequence. Typically, the kinetics for sequences with smaller number of blocks proceeds through smaller number of steps. For example, folding of the periodic sequence consisting of long blocks, $'(b_5a_5)_5'$ (see line E in Fig. 5.12b), proceeds through only one, though rather long, kinetically

Figure 5.13: Diagrams of the mean squared distances matrix, $D_{ij}(t)$, for '(ba)₂₅' copolymer in kinetics after the same quench as in Figs. 5.10, 5.11 and 5.12. Diagrams (a-d) correspond respectively to the following moments in time: $t = 3.5$, 5.75, 8.60 and 13.0. See also caption to Fig. 5.9 for more details.



arrested step. However, for the aperiodic sequence of long blocks (line F in Fig. 5.12b) we can see at least six such steps, the third one being the longest in time.

Now let us consider the nonequilibrium conformations, given by the matrix of the mean squared distances between monomers, D_{ij} . In Fig. 5.13 we exhibit the D_{ij} matrix for '(ba)₂₅' sequence at different times in the folding kinetics. We can see that kinetic process proceeds through formation of locally collapsed and phase-separated clusters. The initial conformation is similar to Fig. 5.9a, then at time $t = 3.6$ we can see approximately 10 clusters (Fig. 5.13a), at time

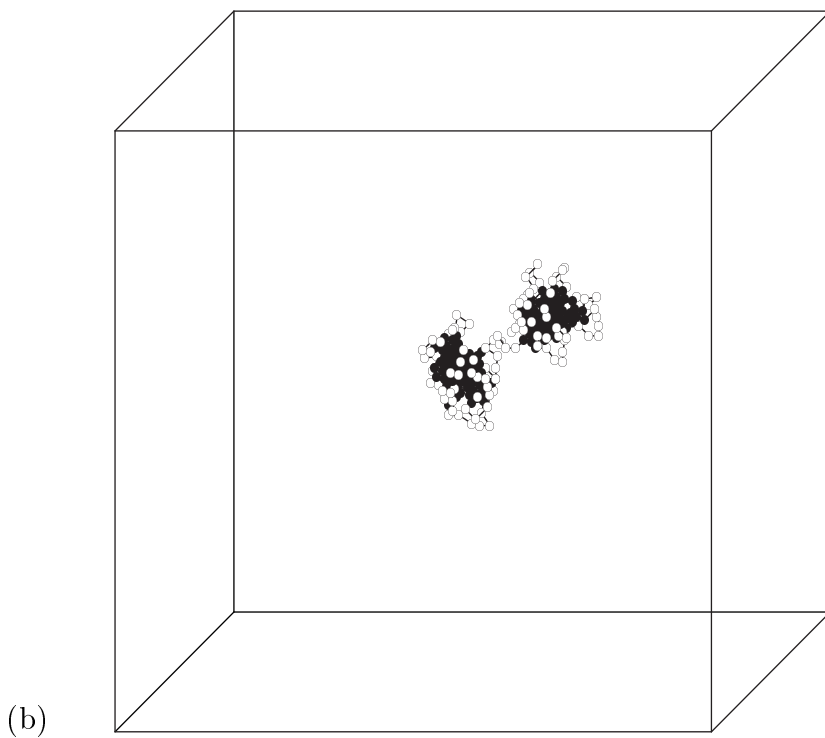
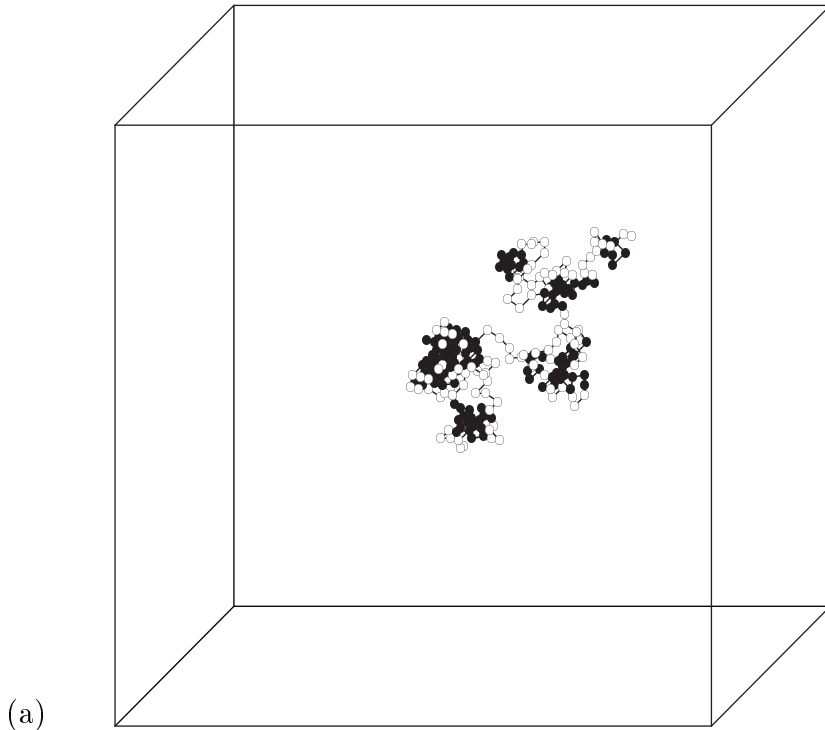
$t = 5.75$ — 6 clusters (Fig. 5.13b) and at $t = 8.6$ — 4 clusters (Fig. 5.13c). During later evolution these four clusters approach each other since, as one can see from Fig. 5.13d, the inter-cluster distances are much smaller than in Fig. 5.13c. Finally, the clusters unify forming the MPS globule, the D_{ij} matrix of which is quite similar to that presented in Fig. 5.9e. We should note here also that these nonequilibrium states do not possess the translational block symmetry, present for the system in the EFEF and the initial conditions. After some time in kinetics, $t \approx 1$ this symmetry breaks down, and restores only during final kinetic stages at time moment $t \approx 15$.

In Fig. 5.14 we present typical copolymer conformations during the folding kinetics after the instantaneous change of the interaction parameters as in Fig. 5.5a. The first kinetic stage (Fig. 5.14a corresponds to approximately the end of this stage) is characterised by formation of small globules along the chain as for the homopolymers. However, new cluster now form mostly in accumulations of hydrophobic monomers. As a result, a few clusters consisting mainly from the hydrophobic material and connected mainly by hydrophilic links are formed. During later kinetic stages the number of clusters decreases and hydrophilic surroundings around hydrophobic cores start to form and stabilise. The formation of hydrophilic shells deters unification of clusters since it is becoming less and less energetically favourable for them to collide and higher barrier should be crossed. This produces what seems to be a kinetic arrest. The conformation in Fig. 5.14b, consisting of two well separated clusters, is obtained after quite a long evolution, the duration of which is comparable with a typical collapse time of the homopolymer four times longer than this copolymer (see Figs. 4.16, 4.20). This conformation may be viewed as representative of the (meta)stable frustrated phases.

We would like to make a general comment here on the nature of spontaneous

Figure 5.14: Typical conformations of $(a_6b_6)_{20}$ hydrophobic-neutral copolymer during folding kinetics.

Parameters of the quench are equal to: $\chi_{aa} = 0.2 \rightarrow 0.8$ ($\chi_{ab} = 0.1 \rightarrow 0.4$).
Fig. 5.5a and Figs. (a), (b) correspond respectively to the following moments in time: $t = 0$, $t = 1,000$ MCS and $t = 500,000$ MCS.



symmetry breaking in kinetics. Thus, normally in such situations at equilibrium there exist a thermodynamically unstable symmetric free energy minimum and a disjoint set of symmetry broken minima, which may be transformed to each other by the residual subgroup of symmetry transformations. These states may also be obtained kinetically as infinite time limits of the time evolution starting from any, for example, the symmetric initial state, which happens to be the main free energy minimum before the quench.

However, the formal structure of the GSC kinetic equations (2.73) is such that they yield a symmetric solution at any moment in time if one proceeds from the symmetric initial condition. What we observe here is a spontaneous symmetry breaking in kinetics. Namely, at some moment in time the symmetric solution of the kinetic equations becomes unstable with respect to perturbations, whether of the initial condition, or of the interaction matrix, $u_{ij}^{(2)}$. In the exact theory there are fluctuations that can transform between different spontaneously broken states in kinetics, thus destroying the unstable symmetric state. To describe such phenomena strictly in the framework of the GSC method, which presents an improved, but still a mean-field type of theory, we should include explicitly an infinitesimal symmetry breaking term ε_{ij} to the two-body interaction matrix, $u_{ij}^{(2)}$, and consider the limit of vanishing perturbation in the solution. In fact, in numerical integration such a regularisation procedure is not even necessary, since there is always an intrinsic perturbation due to computer round off and numerical integration errors. Thus, if the symmetry is favourable to be kinetically broken somewhere, numerically one obtains some spontaneously broken solutions there, rather than the unstable symmetric solution, unless the symmetry conditions have been imposed by hand. In fact, adding a rather small perturbation matrix, ε_{ij} , changes the behaviour of the global observables, such as R_g^2 , Ψ , \mathcal{A} and characteristic kinetic times rather insignificantly. However, what can be changed by

including of a perturbation is the centres along the chain around which clusters form and grow.

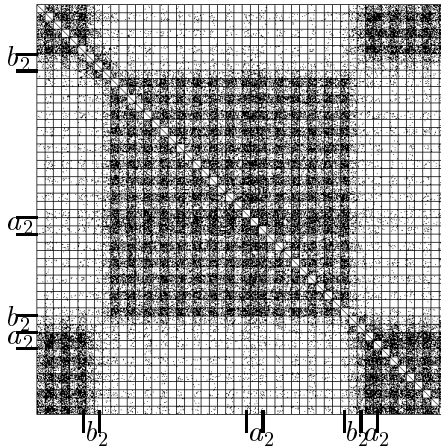
It is worthwhile to make a comment here on the notion of kinetic stages introduced in earlier works [II,116,122]. Those we associated with typical structures of the conformation and accompanying kinetic laws. We distinguished at least the following kinetic stages at the collapse of the homopolymer: early time necklace formation, middle time coarsening and a number of final relaxational processes. The multistep character of folding observed in this work affects the middle kinetic stage, resulting in its considerable complication and splitting into multiple sub-stages with respective complex kinetic laws that are determined by the sequence. Universality of such kinetic laws is doubtful, but probably it can be recovered by averaging over certain classes of sequences with similar folding properties.

Apparently, in the GSC method the kinetics is a motion in the space of $N(N - 1)/2$ averaged dynamic variables, $D_{ij}(t)$, and it is determined by the profile of the free energy. It may be instructive, using Figs. 5.10 and 5.12, to present the kinetics via a parametric plot of \mathcal{A} vs R_g , the latter being the main, though not the only, relevant “coordinate” of the system. That would produce a kind of “bottleneck” picture that was much discussed by P. Wolynes and others [101] in relation to the protein folding problem. This indicates that our complicated method produces the behaviour that may permit interpretation in terms of phenomenological energy landscape models.

Let us utilise once again the connection between the kinetic evolution of the free energy and the ruggedness of the potential energy landscape. This would allow us to shed some light on the general structure of the energy landscape for complex heteropolymers. Thus, a typical example of the “random” sequence kinetics (curve (C) in Fig. 5.12a) shows first a fast drop, then appearance of short steps that are growing longer with time, until the last step becomes infinite. This

Figure 5.15: Diagram of the mean squared distances matrix, $D_{ij}(t)$, for copolymer sequence $'(ba)_3b_2(ba)_9a_2(ba)_5b_2a_2(ba)_4'$ at the time moment $t = 45$, at which the system is close to the final equilibrium.

Parameters of the quench here are the same as in Figs. 5.10-5.12. See also caption to Fig. 5.9, 5.13 for more details.



translates to the following equivalent energy landscape versus some collective coordinate: first there is a rapid drop away from the unstable coil state, then the surface flattens and small wrinkles appear on it, they are growing larger in the amplitude gradually becoming high mountains and deep valleys, until there is eventually a very deep ravine corresponding to the “ground” state of the system separated by a very high barrier from other minima. In our view, the latter picture bears a remarkable resemblance to a typical mean field energy landscape in spin glass systems [32] recently discussed by G. Parisi [173]. This observation seems encouraging to us and indicates that the GSC method is capable of describing very complex systems, though it is too detailed and expensive in the complete form for description on a macroscopic scale. Nevertheless, by taking the quenched disorder averaging (similarly to Ref. [150]) it is possible to achieve an adequate description for heteropolymers that is alternative to the replica formalism.

Finally, let us discuss how the folding kinetics depends on the sequence. In series of pictures in Figs. 5.10a-5.12a we present kinetic processes for the simplest

copolymer consisting of 'ab' blocks (sequence B), and also for some sequences obtained by certain modifications in it. In the sequence denoted by C in Figs. 5.10a-5.12a we have replaced ten short blocks by four of longer size. In the sequence D we have inserted only two hydrophobic and two hydrophilic fragments into the sequence, i.e. we have made only two permutations. For the former sequence the total kinetic time became approximately 1.7 times longer than that for the B sequence. The final values of the free energy for both sequences B and C are nearly the same, whilst the micro-phase separation is somewhat better for the modified sequence C, due to longer blocks. Thus, for that sequence the kinetic properties do not deteriorate very much, except that kinetic process takes essentially longer. However, the kinetic foldability of the D sequence is much more poor than for B and C sequences. In fact, here not only the kinetic process takes much longer, but the final state is different from the MPS globule. As one can see from Fig. 5.15 the final kinetic state for the D sequence consists of two clusters. These clusters are connected by two links, formed by nearly neutral fragments, 'b₂'. Further collapse of this conformation is unfavourable due to the entropy and partial screening of hydrophobic monomers by nearly neutral species 'b'. Size of such misfolded state is larger and the MPS order parameter is approximately twice smaller than what we could expect for the MPS globule. As we have already seen kinetics of sequences consisting of longer blocks is also quite sensitive to the sequence. Some modifications of the periodic sequence E in Figs. 5.10b-5.12b to form aperiodic long blocks sequence, F, result in somewhat longer, but much more complicated, kinetic processes.

*If you're not part of the solution, you're
part of the precipitate.*

Source unknown

Chapter 6

Polymer Solutions at Finite Concentrations

In recent experiments with PNIPAM copolymers with small number of ionomers [174, 175] it has been found that such systems can form stable nanoparticles in aqueous solution, instead of simply aggregating on heating above the lower critical solution temperature (LCST). Such novel type of mesoscopic aggregates (mesoglobules) have been earlier reported for the homopolymer PNIPAM [VI], in which case they are very long-lived, but not truly stable.

These metastable structures for the homopolymer cannot be envisaged in the framework of the conventional Flory–Huggins type theories. Therefore, in Ref. [IV] we have extended the GSC method to multiple chains and argued the existence of mesoglobules in dilute solution from the thermodynamic considerations. Although the conventional Flory–Huggins theory can be indeed derived from our method in the thermodynamic limit using some approximation as we shall see in Subsec. 6.1.1, that approximation breaks down at low concentrations. In Ref. [VII] we have also mentioned the possibility that the micro-phase separation can additionally stabilise the mesoglobules in heteropolymers, which as we consequently learned from work [175] by Qiu et al., is indeed the case. Here,

based on the complete form of the GSC method we shall support that conjecture by a direct numerical evidence. Thus, it is possible to consider several sufficiently short heteropolymer chains of a given composition in a box of finite volume V and to analyse the values of the Helmholtz free energy \mathcal{A} at all its possible local minima, which are hence the stationary points of the GSC equations.

At this point let us make some comments on the general aspects of the two phase coexistence in statistical mechanics. The standard procedure would be the following. First, one has to obtain the dependence of the specific free energy $a = \lim \mathcal{A}/K$ (or possibly other equivalent thermodynamic potential) on the concentration $c = K/V$ (or chemical potential) in the thermodynamic limit when $V \rightarrow \infty$ and the number of particles K also diverges so that c remains finite. Second, one has to build the convex hull of the function $a[c]$ by applying the Maxwell construction [110, 161]. Within the two phase separation region, where a straight line is drawn connecting a section of the $a[c]$ plot, the pure equilibrium states are unstable with respect to fluctuation as indicated by the wrong sign of $\partial^2 a / \partial c^2$. The stability of the *mixed* two-phase state in this case relies heavily on the thermodynamic additivity properties and that one can neglect the surface contribution of the interface between two phases.

Polymers, however, are pretty much finite, though normally long, chains. Due to the connectivity of each chain there is no well defined interface between the high and low density phases, and, moreover, the surface entropic contributions do seem significant too. These factors should be carefully accounted for while trying to study possible metastable states. So, we shall consider a finite number of finite length polymer chains, for which one, strictly speaking, cannot use the Maxwell construction. However, in this context the two-phase coexistence still means that the single phase pure states are thermodynamically unstable. We can expect here that there are other metastable states present at the same values of

thermodynamic parameters and that strong fluctuations may bring the system from one of those states to another. Thus, the equilibrium turns out to be a mixed state including many local minima of the free energy for a finite system, rather than just two — the high and low density states — as in the thermodynamic limit. Increasing the system size would lead to an expected view of the polymer precipitate as consisting of many various sized aggregated and a few single globules, instead of the oversimplified picture of one huge and some single globules based on application of the Maxwell construction in a mean-field treatment.

In heteropolymers the situation is even more dramatic. Due to the competing hydrophobic and hydrophilic interactions the free energy profile is very rugged. There may appear some new global minima due to a specific compensation of the interactions and the entropy. Such a situation corresponds to stabilisation of the mesoglobules which we shall study in this chapter. For this we perform numerical study of the complete GSC equations (2.73) for all elements of the matrix of mean squared distances, $D_{AA'}$. Comparing depths of various minima of the free energy allows us to determine the equilibrium phase diagrams. Some preliminary studies of the kinetics of precipitation based on Monte Carlo simulation on a lattice are also presented here.

6.1 Homopolymer Solutions

In this section we consider the phase diagram of the homopolymer solution and kinetics of precipitation.

6.1.1 Thermodynamic limit

In section 2.6 we have derived the GSC equations (2.100) for solution of identical ring homopolymers. In this subsection we would like to establish connection of

these GSC equations in the thermodynamic limit $M, L \rightarrow \infty$, $M/L^3 = \text{const}$ with the Flory–Huggins theory [11, 53, 54].

The specific free energy given by Eqs. (2.96, 2.98) may be decomposed into the three parts,

$$a \equiv \lim_{L, M \rightarrow \infty} \frac{\mathcal{A}}{L^3} = a_{inter}[\Gamma] + a_{inter-intra}[\Gamma, \mathcal{F}_q] + a_{intra}[\mathcal{F}_q]. \quad (6.1)$$

To obtain a finite answer we should keep the concentration, $c = NM/L^3$, and the ratio $\gamma \equiv L^{-2}\Gamma^2$ finite. The inter–intra–molecular term nontrivially renormalises the interaction parameters in a_{inter} , and we obtain,

$$\frac{a'_{inter}}{k_B T} = -\frac{3}{2} \frac{c}{N} \log \left((MN/c)^{2/3} \gamma \right) + \frac{3}{2} \frac{c\gamma}{N} + v^{(2)} c^2 \gamma^{-3/2} + v^{(3)} c^3 \gamma^{-3}, \quad (6.2)$$

where we have introduced the rescaled variables,

$$v^{(J)} \equiv J^{-3/2} \frac{\hat{u}^{(J)}}{k_B T} \alpha_J, \quad \alpha_2 \equiv \left(1 + 3 \frac{\hat{u}^{(3)}}{\hat{u}^{(2)}} \sum_k D_k^{-3/2} + \dots \right), \quad \alpha_3 = 1. \quad (6.3)$$

Minimisation of (6.2) with respect to γ yields,

$$\frac{1}{N\gamma} = \frac{1}{N} - v^{(2)} \frac{c}{\gamma^{5/2}} - 2v^{(3)} \frac{c^2}{\gamma^4}. \quad (6.4)$$

We can now consider two simple limiting regimes:

(a) If $|v^{(2)}|c \ll \frac{1}{N}$, $v^{(3)}c^2 \ll \frac{1}{N}$ we obtain the solution $\gamma \simeq 1 + N(v^{(2)}c + 2v^{(3)}c^2)$.

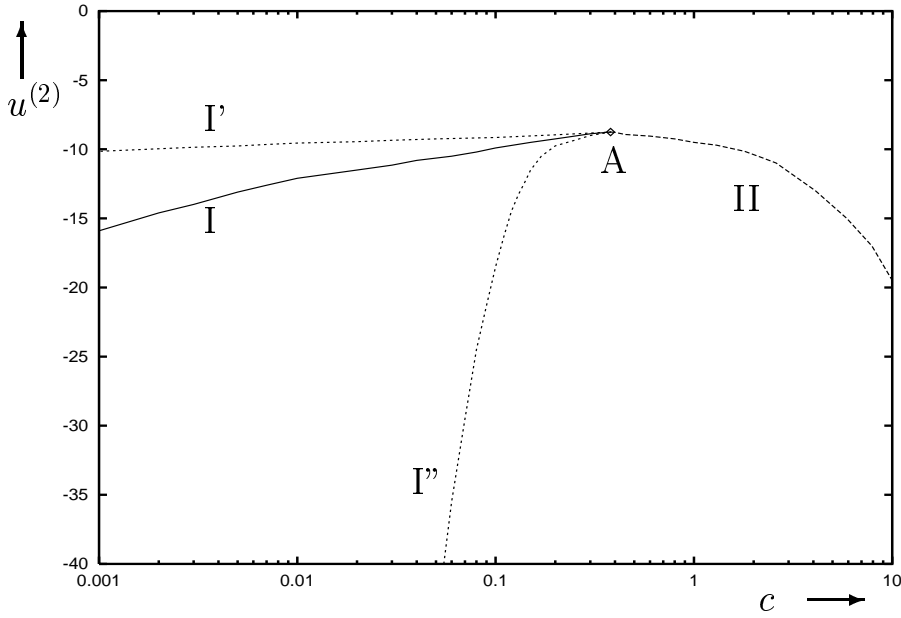
In this regime we can solve the equations for \mathcal{F}_q first, ignoring the coupling with Γ , and substitute the result into Eq. (6.3). Then we obtain the Flory–Huggins expressions for the osmotic pressure and the free energy,

$$\frac{\Pi_{FH}}{k_B T} = \frac{c}{N} \gamma = \frac{c}{N} + v^{(2)} c^2 + 2v^{(3)} c^3, \quad (6.5)$$

$$\frac{a_{FH}}{k_B T} = \frac{c}{N} \log \frac{c}{c_0} + v^{(2)} c^2 + v^{(3)} c^3, \quad (6.6)$$

where $c_0 = MNe^{-3/2}$ is an irrelevant constant removed by consideration of the free energy of mixing. Now let us go back to the equations for \mathcal{F}_q for a moment.

Figure 6.1: Phase diagram of the homopolymer solution in variables of the concentration, c , (in the logarithmic scale) and the second virial coefficient, $u^{(2)}$. Curve I corresponds to first order phase transition and curve II — to continuous transition. Point A is a critical point. Curves I' and I'' are the regions of existence of corresponding pure phases. The region of phase coexistence determined by the Maxwell construction lies below curves I' and II but not explicitly marked. Here $N = 100$ and $L = 100$.



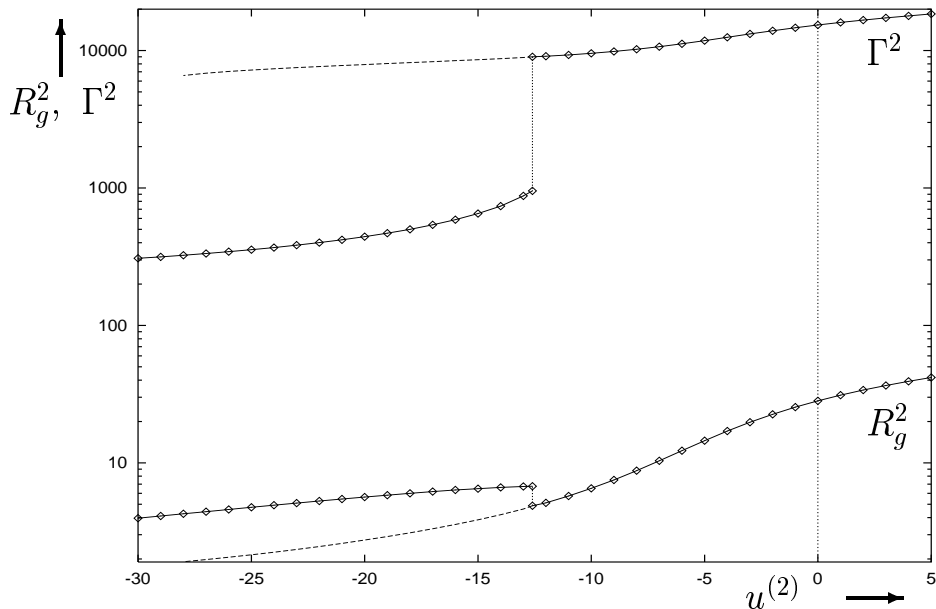
These have the same form as the equations for one chain with the substitution $\hat{u}^{(2)} \rightarrow \hat{u}^{(2)} + 2^{-3/2} 3 \hat{u}^{(3)} c \gamma^{-3/2}$. Since we have $c \hat{v}^{(2)} \ll \hat{v}^{(3)}$ the coupling between both types of the variables indeed may be neglected there.

(b) If on the contrary $|v^{(2)}|c > \frac{1}{N}$, i.e. N is large, we can solve (6.4) first ignoring the coupling in (6.3). The solution is $\gamma \simeq \left(2 \frac{v^{(3)}}{|v^{(2)}|} c\right)^{2/3}$. This corresponds to the formation of an insoluble precipitate with the free energy and osmotic pressure,

$$\frac{a_{prec}}{k_B T} = -\frac{|v^{(2)}|^2}{4v^{(3)}} c + O\left(\frac{1}{N}\right), \quad \frac{\Pi_{prec}}{k_B T} = \frac{c^{5/3}}{N} \left(\frac{2v^{(3)}}{|v^{(2)}|}\right)^{2/3} + O\left(\frac{1}{N^2}\right). \quad (6.7)$$

If we substitute the above γ back to the equations for \mathcal{F}_q we shall see that the inter-chain repulsion is noticeably weakened.

Figure 6.2: Plots of the mean squared inter-chain distance, Γ^2 , and the mean squared radius of gyration of a chain, R_g^2 , vs the second virial coefficient, $u^{(2)}$. Here and below solid curves represent observables in the global free energy minimum and dashed curves — in metastable minima. Here $N = 100$, $L = 100$ and $c = 0.12$.



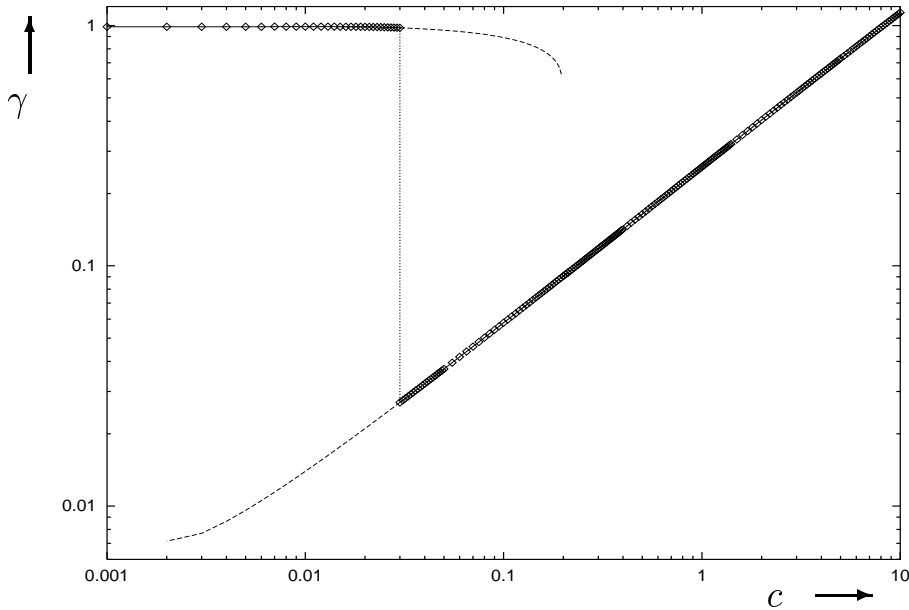
6.1.2 Equilibrium phase diagram

In this subsection we present results from the numerical study of the equilibrium limit of the GSC equations (2.100) for solution of identical ring homopolymers.

At equilibrium we obtain the phase diagram presented in Fig. 6.1. For positive $u^{(2)}$ chains exist in the extended coil state. In the attractive regime, $u^{(2)} < 0$, in a very narrow region at low concentrations (above and to the left of curve I' in Fig. 6.1) chains collapse individually, forming a gas of practically non-interacting globules. The narrowness of this region is responsible for the difficulties in experimental observation of the single polymer collapse. The area below the curves I' and II corresponds to the two phase coexistence region where $\partial^2 \mathcal{A} / \partial c^2$ is negative, and point A denotes the critical point. We agree with the Flory–Huggins theory with respect to the laws describing the boundary of the coexistence region.

Figure 6.3: Plot of the quantity γ vs the concentration, c , for $N = 100$, $L = 100$ and $u^{(2)} = -15$.

For the aggregate, in the right hand side of the figure, γ scales as, $\gamma \sim c^{2/3}$, corresponding to space filling.

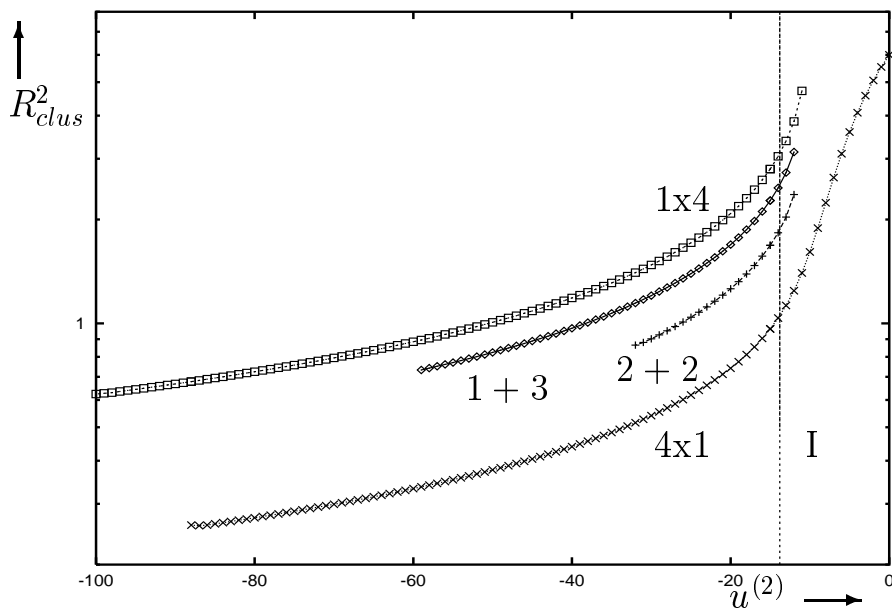


We may remark that the variational method allows one to obtain the pure thermodynamically dominant states which could correspond either to stable (global minimum of \mathcal{A}) or metastable (local minima) states. Thus, we have two such states here: a gas of globules above curve I, and a dense macroscopic precipitate of aggregated chains below and to the right of the first order transition curve I. Beyond curve II only the macroscopic precipitate can exist, which may be viewed simply as a macroscopic dense globular state filling most of the available space. In the region between I and the curve I' at the same concentration the gas may exist as a metastable state, as well the metastable precipitate may exist between I and I'. The crucial observation is that $\partial^2 \mathcal{A} / \partial c^2$ is negative at all of these branches in the coexistence region.

In Fig. 6.2 we draw the quantities Γ^2 and R_g^2 across the transition curve I at

Figure 6.4: Plot of the mean squared cluster size, R_{clus}^2 , vs the second virial coefficient, $u^{(2)}$, for different cluster states.

This data is obtained for open homopolymers with $N = 18$, $M = 4$, $L = 10$ and $u^{(3)} = 10$.



different free energy branches. At I the inter-chain distance, Γ , drops dramatically from the magnitude of order L to a much smaller value corresponding to the dense precipitate. In Fig. 6.3 we present $L^{-2}\Gamma^2$ as function of the concentration. This figure confirms that the two limits discussed in subsection 6.1.1 are satisfactory, and also that for the precipitate $\gamma \sim c^{2/3}$ reaffirming the idea of the macroscopic dense state. On the contrary, the intra-chain distance, R_g^2 , grows markedly at the transition curve I. Really, chains tend to swell once they have formed a larger globule approaching a nearly ideal coil conformation. It is interesting to comment that R_g is very weakly dependent of the concentration c at both branches as may be established from our analysis in Subsec. 6.1.1.

6.1.3 Metastable states

In fact, the GSC equations (2.100) in the simplified treatment of Sec. 2.6) are capable to describe only the “symmetric” phases where all polymers appear apart from each other or collapse to a single precipitate. Numerical analysis of the complete set of equations (2.73) shows that there are additional solutions related to various local free energy minima. These minima correspond to conformations where polymers in solution form several clusters, each consisting of one or a few chains. Obviously, for a large number of chains there may be many such states. The situation is illustrated in Fig. 6.4, in which we present the mean squared radii of gyration of clusters formed by various number of chains in solution of just 4 homopolymers. The lines are drawn as long as the corresponding minimum of the free energy exists. One can see from Fig. 6.4 that for rather small number of chains such “asymmetric” minima exist in the region of the phase diagram bounded by the curves I' and I'' in Fig. 6.1. As can be seen from Fig. 6.4 the upper bounds in the second virial coefficient $u^{(2)}$ for the existence of such solutions are approximately the same. Indeed, since the curve I' goes nearly horizontally in Fig. 6.1 the critical value of $u^{(2)}$, at which a cluster splits into separate chains, does not depend significantly on the number of chains forming such a cluster.

Now let us compare the values of the free energy at minima corresponding to various possibilities to divide the system into subsets. In Tab. 6.1 we present values of the free energy for some cluster sizes, corresponding to either a symmetric division of the system into clusters of equal size, or a large precipitate plus one or two single globules. One can see from Tab. 6.1 that all asymmetric minima for the homopolymer are indeed metastable. Note also that the system divided into a large aggregate plus a few single globules possesses the free energy value close enough to that at the global minimum, whilst the system divided into several clusters of equal size has a higher free energy. In fact, one would expect

Table 6.1: Values of the specific free energy, $a = \mathcal{A}/MN$, at various minima for the system of 12 open homopolymers in volume $L = 10$ for various values of the second virial coefficient, $u^{(2)}$.

The value of the global minimum of the free energy in each line is printed in **bold**.

| $u^{(2)}$ | 12×1 | 6×2 | 4×3 | 3×4 | 2×6 | 1×12 | 1, 11 | 1, 1, 10 |
|-----------|---------------|--------------|--------------|--------------|--------------|---------------|--------|----------|
| -12 | -1.028 | -0.843 | -0.815 | -0.814 | -0.832 | -0.886 | -0.890 | -0.896 |
| -13 | -1.378 | -1.206 | -1.185 | -1.188 | -1.210 | -1.270 | -1.271 | -1.274 |
| -14 | -1.759 | -1.604 | -1.589 | -1.596 | -1.623 | -1.688 | -1.685 | -1.685 |
| -15 | -2.172 | -2.035 | -2.027 | -2.038 | -2.069 | -2.140 | -2.132 | -2.127 |
| -16 | -2.615 | -2.499 | -2.498 | -2.513 | -2.548 | -2.626 | -2.615 | -2.606 |
| -18 | -3.595 | -3.523 | -3.539 | -3.562 | -3.607 | -3.695 | -3.676 | -3.659 |
| -20 | -4.696 | -4.676 | -4.709 | -4.742 | -4.797 | -4.897 | -4.867 | -4.840 |
| -25 | -7.982 | -8.113 | -8.198 | -8.258 | -8.341 | -8.472 | -8.416 | -8.386 |

Figure 6.5: Plot of the mean number of clusters, $\langle n_{clus} \rangle$, vs the Flory interaction parameter, χ , in the lattice Monte Carlo model of homopolymer solutions.

Here the degree of polymerisation and the linear lattice size are equal to $N = 12$ and $L = 100$. Lines correspond to the following numbers of polymer chains (from bottom to top in the left hand side of the figure): $M = 25$ (diamonds), $M = 50$ (pluses) and $M = 100$ (quadrangles).

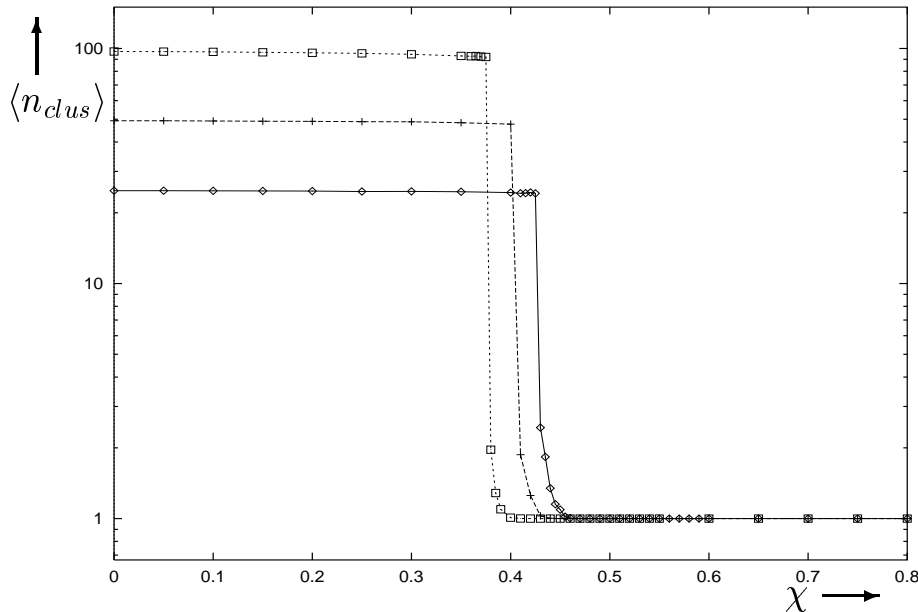
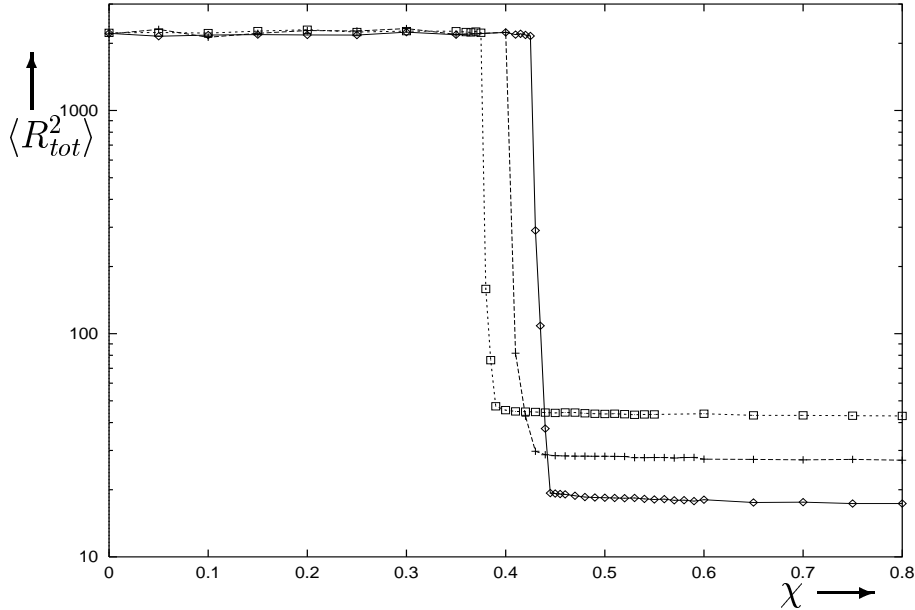


Figure 6.6: Plot of the mean squared total radius of gyration, R_{tot}^2 , vs the Flory interaction parameter, χ .

Lines correspond to the same values of M as in Fig. 6.5.



that for the homopolymer solution of a huge number of chains the minimum of the free energy should be at one of the asymmetric solutions consisting of one large precipitate plus a gas of single globules. Such behaviour would be consistent with the standard picture of a two-phase coexistence in the thermodynamic limit. However, due to high computational expenses we have not been able to test it properly based on the GSC equations yet.

6.1.4 Phase separation: a Monte Carlo study

Now let us consider the phase separation at equilibrium and the kinetics of precipitation in the lattice Monte Carlo model introduced in Sec. 3.1.

In series of plots in Figs. 6.5, 6.6, 6.7 and 6.8 we present the behaviour of various observables versus the Flory interaction parameter, χ , defined by Eq. (3.28), for three different values of the concentration. In the region of weak attraction typical system conformations are those of a gas of swollen coils, as depicted in

Figure 6.7: Plot of the specific mean energy, $\langle H \rangle / NM$, vs the Flory interaction parameter, χ .
 Lines correspond to the same values of M as in Fig. 6.5.

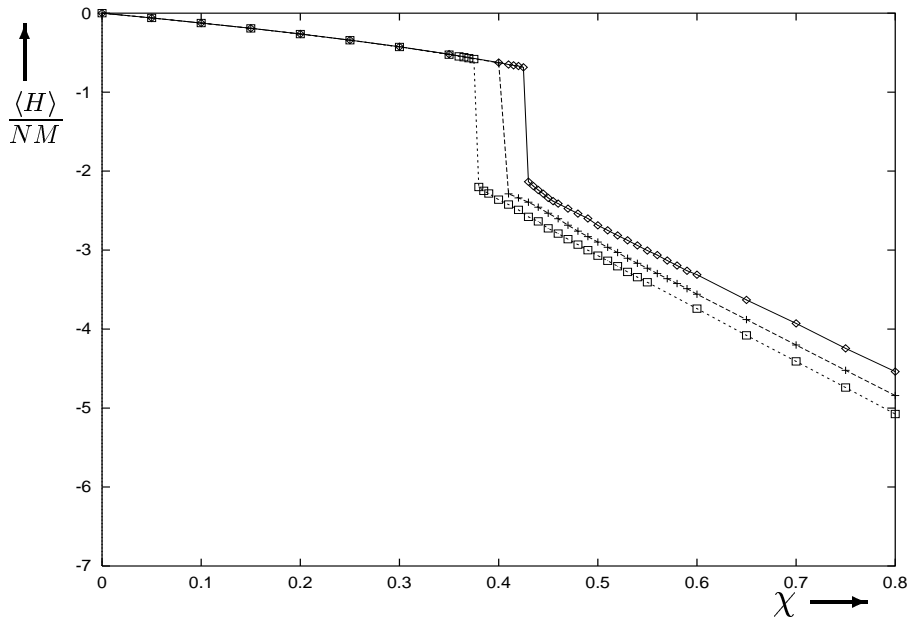


Figure 6.8: Plot of the specific heat capacity, c_V / k_B , vs the Flory interaction parameter, χ .
 Lines correspond to the same values of M as in Fig. 6.5.

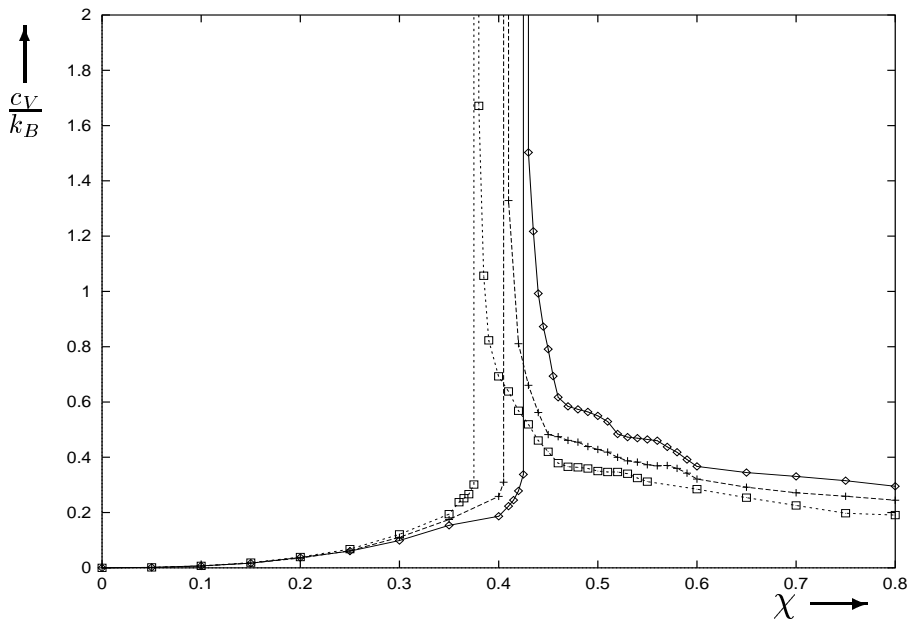


Figure 6.9: Typical conformations in the homopolymer solution at a finite concentration (a, b).

Here Figs. (a) and (b) correspond respectively to the good solvent condition, $\chi = 0$, and the poor solvent near the transition point, $\chi = 0.42$. Other parameters are: $N = 12$, $M = 50$, $L = 100$.

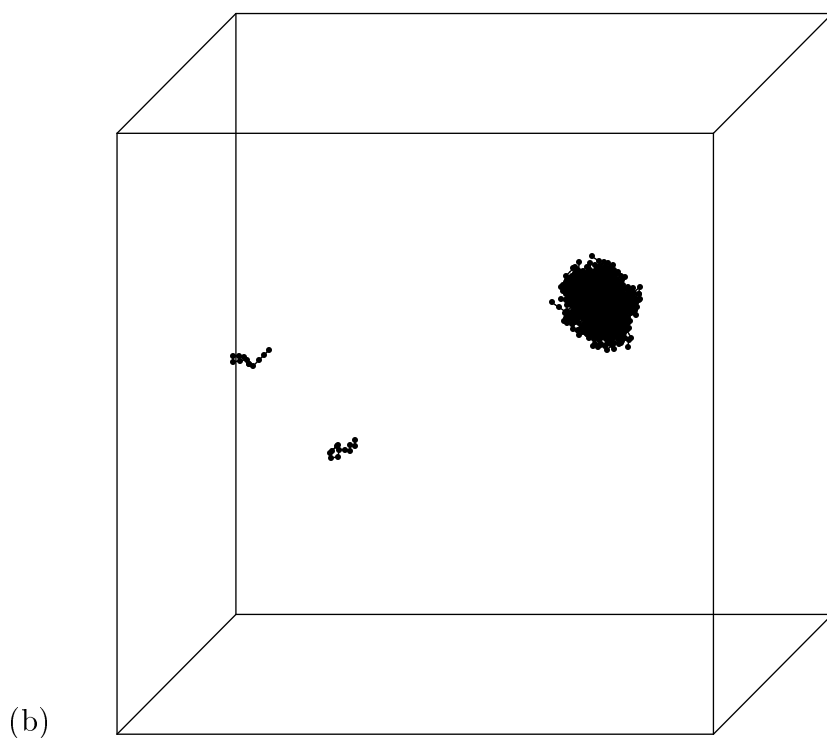
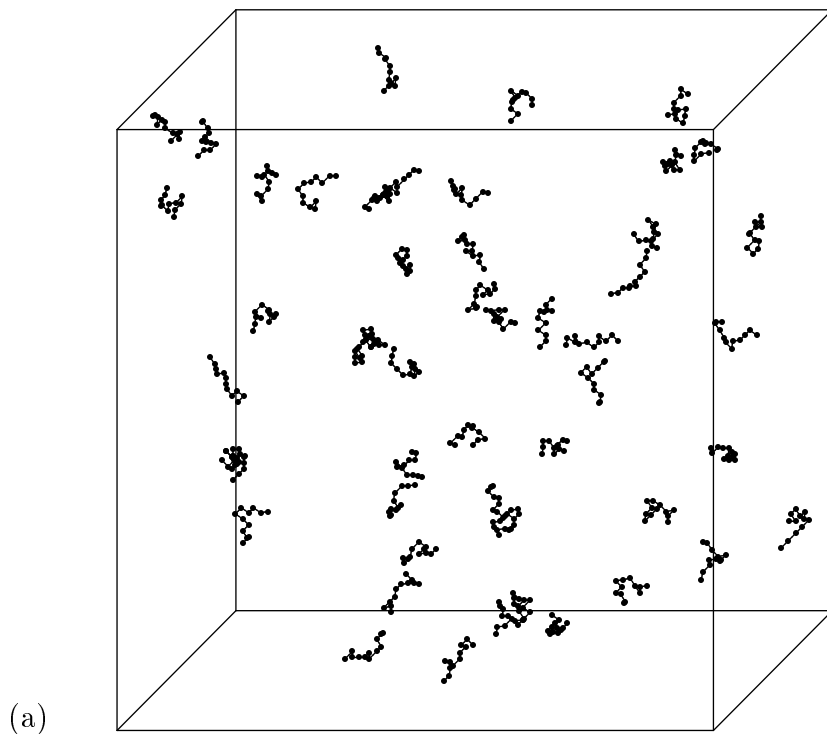


Figure 6.10: A typical conformation in the homopolymer solution at a finite concentration corresponding to the precipitate ($\chi = 0.5$). Other parameters here are the same as in Fig. 6.9.

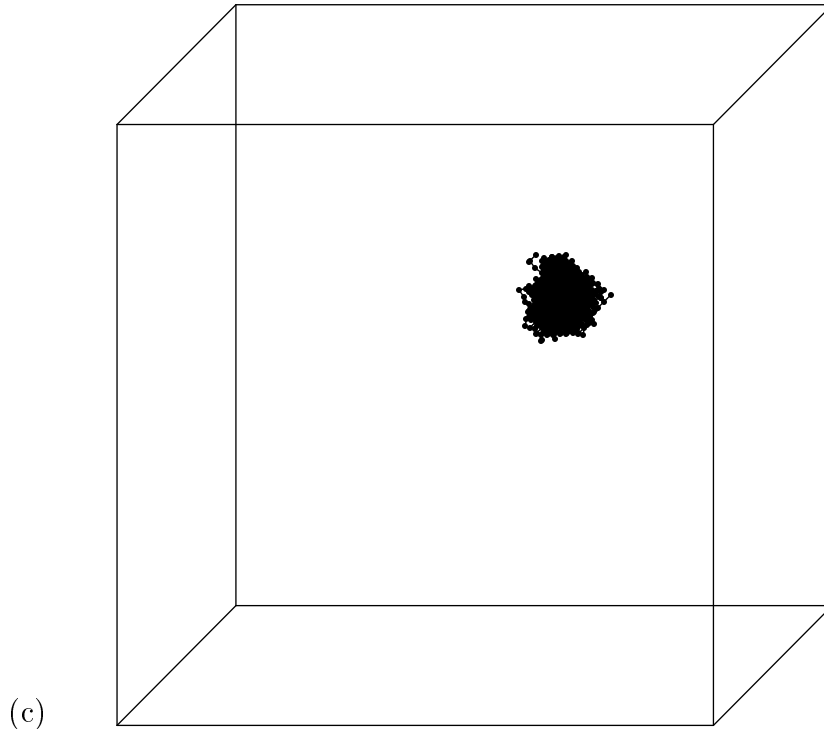


Fig. 6.9a. In the mean number of clusters, $\langle n_{clus} \rangle$, shown in Fig. 6.5, clusters are defined as the subsets of noninteracting with each other chains, each consisting of interacting chains. For small values of χ this quantity is insignificantly smaller than the total number of chains on the lattice, M , i.e. chains prefer to stay apart from each other in essentially dilute solution, experiencing rare collisions. The number of clusters, as well as the total radius of gyration, R_{tot}^2 , presented in Fig. 6.6, depend weakly on χ in the good solvent. The energetic characteristics in Figs. 6.7 and 6.8 behave quite similarly to those of the single homopolymer coil (see Figs. 4.4, and 4.5).

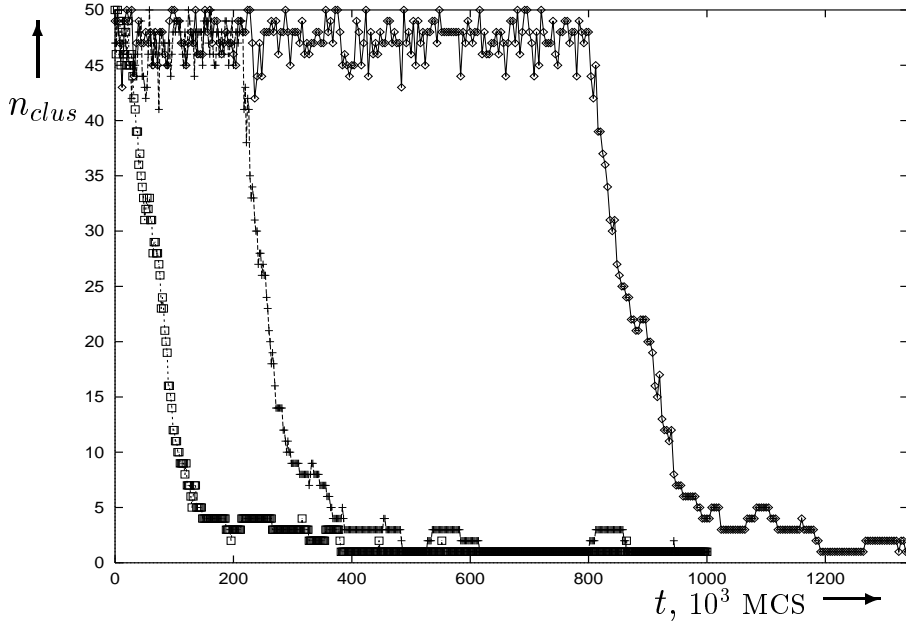
At some critical value of the Flory parameter, $0.3 < \chi < 0.5$, depending on the polymer concentration, the mean number of clusters drops significantly, which corresponds to the phase separation transition. In a region just beyond

the transition curve there exist one large aggregate plus a few distinct chains (see Fig. 6.9b). This represents the two-phase coexistence of phases of the high (the aggregate) and the low (a few chains apart from the aggregate) density states. With increasing χ the mean number of clusters quickly reaches the unity value, $\langle n_{clus} \rangle = 1$, corresponding to the single macroaggregate, such as that presented in Fig. 6.10. In fact, this situation still corresponds to the two-phase coexistence with the low density state being practically unobservable in a finite system. Indeed, as one can see from Fig. 6.7, a typical gain of the energy for a monomer to be inside the aggregate rather than in a single globule is about, $e_{0.4} \approx 1.6k_B T$ at $\chi \approx 0.4$. Thus, for the same value of the Flory interaction parameter, χ , as in Fig. 6.10, where the system prefers to be in a single macroaggregate, the energy required to remove a chain may be roughly estimated as, $\Delta E_{0.5} \approx 0.5N e_{0.4}/0.4 = 24k_B T$, and the appropriate Metropolis probability is extremely small, $\eta = \exp(-\Delta E_{0.5}/k_B T) \approx 3.8 \cdot 10^{-11}$. In fact, for simulating coexistence in simple liquid-gas systems, which exhibit states with very high and very low density, the canonical ensemble is quite inefficient and the isobaric ensemble Monte Carlo simulation is more popular [5, 9]. In simulation the main observables, such as mean energy (Fig. 6.7) and the mean squared total radius (Fig. 6.6), change across the phase separation transition in a quite discontinuous manner, which is consistent with the transition being first order.

Now let us examine the nonequilibrium process after a quench from the gas of swollen coils to the region of the interaction parameters just beyond the phase separation curve. In Fig. 6.11 we present the kinetic evolution of a single member of the ensemble for three different quenches. Right after the quench the system remains in the gaseous state for some time, which strongly depends on the quench depth. As one can see from Fig. 6.11, this time may be up to hundreds of thousands and millions of Monte Carlo sweeps. On approaching the transition

Figure 6.11: Plot of the number of clusters, n_{clus} , vs the Monte Carlo time, t , after quenches near the transition line.

Lines correspond to the following quenches: $\chi_f = 0.41$ (diamonds), $\chi_f = 0.42$ (pluses) and $\chi_f = 0.43$ (quadrangles). Other parameters are: $N = 12$, $M = 50$, $L = 100$.



curve this trapping time diverges as a power law. Thus, in a sense, it is extremely difficult to locate the transition point using simulations since times required for the system to start evolving towards the equilibrium are enormous. For this reason, in Fig. 6.8 we draw the peak in the heat capacity only symbolically.

The situation here is completely analogous to an arrest in kinetics of the gas-liquid transition after a quench to the region of nucleation and growth [176,177]. A long trapping time here is related to a low probability of forming a droplet of the new phase of greater than the critical size. This probability may be estimated roughly as, $\exp(-\Delta\mathcal{A}/k_B T)$, where $\Delta\mathcal{A}$ is the barrier height in the free energy. In our case the barrier is mostly of the entropic nature and the probability for different chains to form a sufficiently large cluster is very low indeed. In Fig. 6.12 we present snapshots of the system at different times after a cluster of the critical size has been formed. This process of cluster growth is relatively fast and

Figure 6.12: Typical conformations in the homopolymer solution during the kinetics of precipitation after a quench near the transition line. Here Figs. (a) and (b) correspond respectively to the following moments in time $t = 816,000$ MCS and $t = 880,000$ MCS. Other parameters are: $N = 12$, $M = 50$, $L = 100$ (line denoted by diamonds in Fig. 6.11).

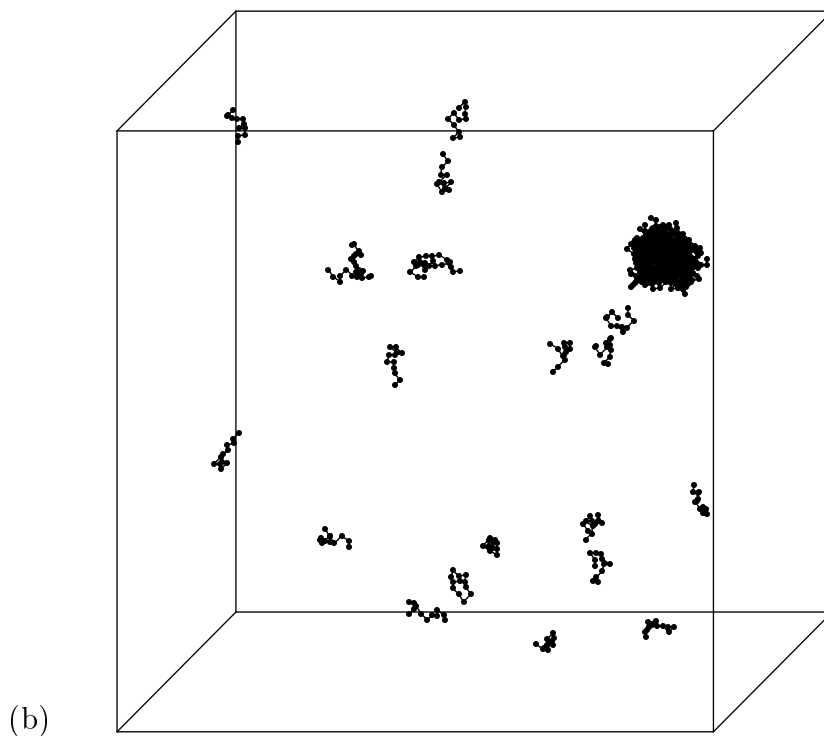
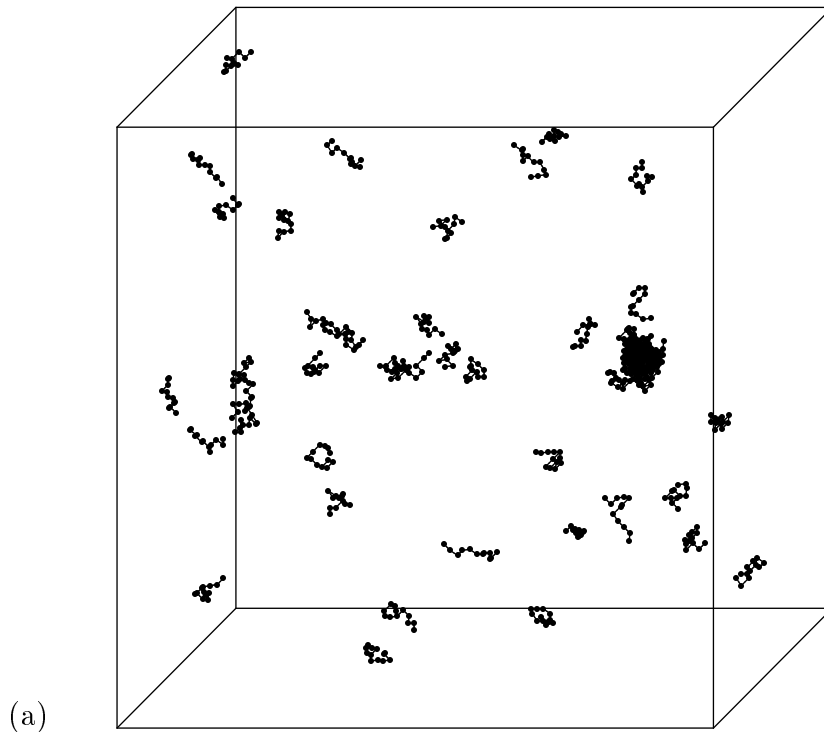
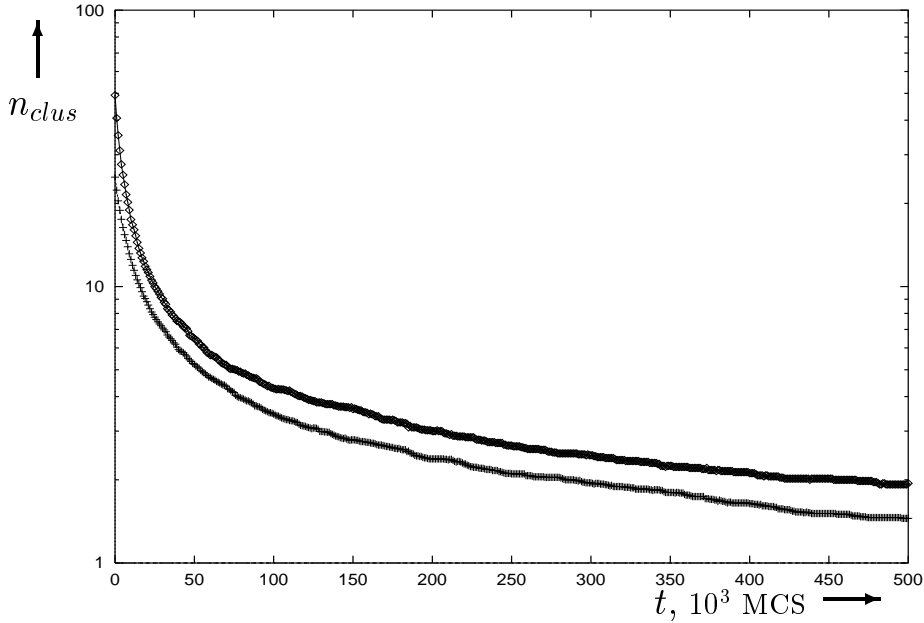


Figure 6.13: Plot of the average number of clusters, $\langle n_{clus} \rangle$, vs the Monte Carlo time, t , after quenches well beyond the transition line.

Lines correspond to the following numbers of polymer chains: $M = 25$ (diamonds) and $M = 50$ (pluses). Here $N = 12$, $L = 100$, $\chi_f = 0.8$.



is probably governed by a power law in analogy with the Lifshitz–Slyozov theory [178]. The over-critical cluster quickly absorbs more and more chains until one macroglobule, possibly with a rather few single globules left, is reached. After the equilibrium coexistence state has been reached, fluctuations, which correspond to a dynamic equilibrium between the macroglobule and the low density solution, continue (see the right-hand side in Fig. 6.11).

On increasing the quench depth the trapping time decreases until the activation barrier vanishes at the spinodal curve, at which the critical size of clusters becomes equal to zero [179]. In our particular case exhibited in Fig. 6.11 this corresponds approximately to $\chi = 0.44$. In series of plots in Figs. 6.13, 6.14 and 6.15 we present the kinetic evolution of the mean number of clusters, the mean specific energy and the mean squared radius of gyration of an individual chain after the quench deeply inside the coexistence region. One can see that the kinetic

Figure 6.14: Plot of the specific mean energy, $\langle H \rangle / MN$, vs the Monte Carlo time, t , after quenches well beyond the transition line. Here parameters and notations are the same as in Fig. 6.13.

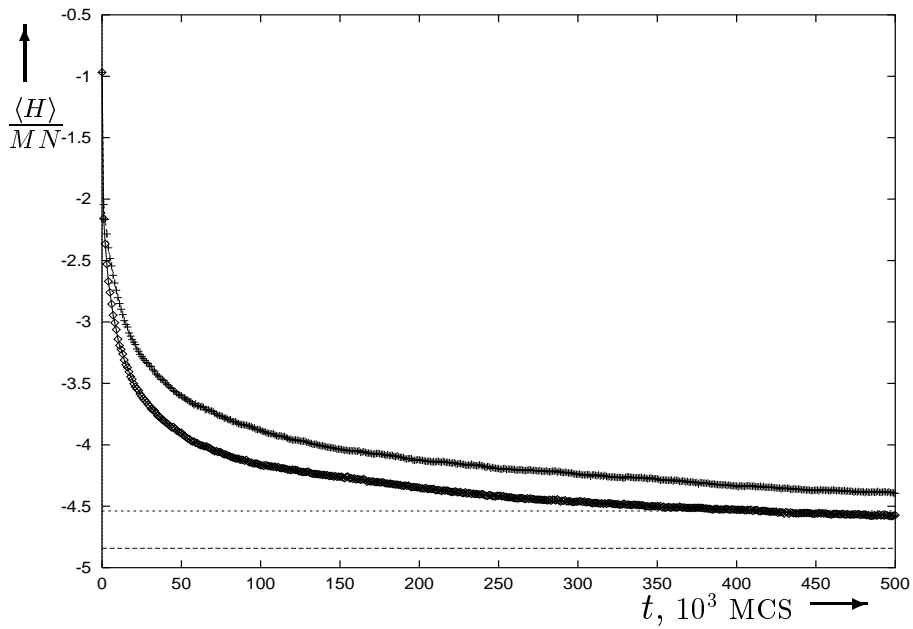
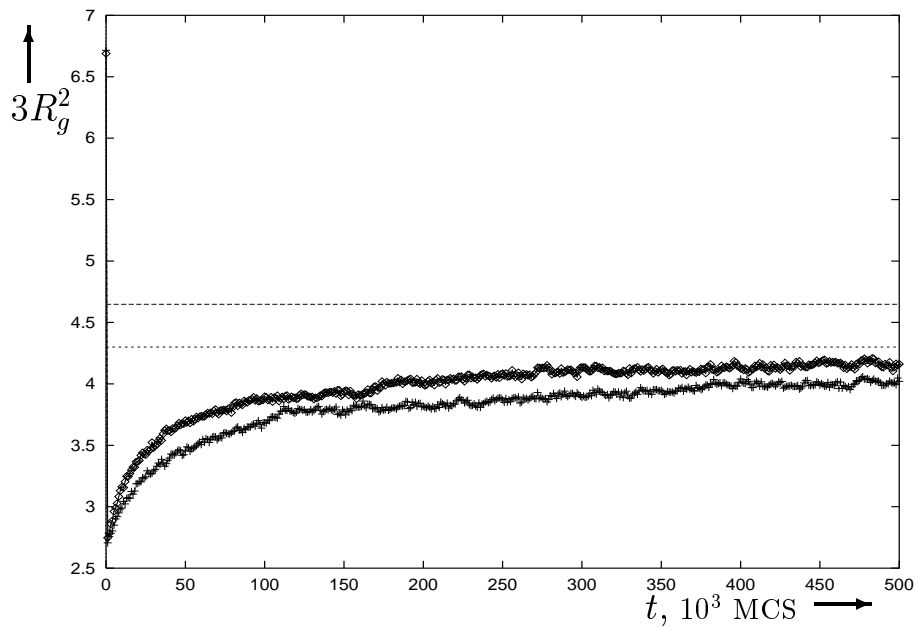


Figure 6.15: Plot of the mean squared radius of gyration of a chain, R_g^2 , vs the Monte Carlo time, t , after quenches well beyond the transition line. Here parameters and notations are the same as in Fig. 6.13.



behaviour has changed dramatically indeed.

The kinetics starts by a quick decrease in the mean number of clusters and the mean energy. A typical system conformation after this stage contains many clusters of various sizes characteristic of the spinodal decomposition process of the unstable state. Consequent kinetics proceeds through unification of these clusters, on the contrary to the nucleation and growth we saw above, where one supercritical cluster absorbs the rest. The process slows down with time due to the two main factors. First, the concentration of clusters gradually decreases and they tend to collide less. Second, the diffusion constant of larger clusters decreases inversely proportional to the number of monomers forming them according to the Stokes law, and thus such larger clusters are less mobile. At this stage, while there is still a sufficient number of clusters left, the mean number of clusters decreases with time as a power law, $\langle n_{clus} \rangle \sim t^{-\alpha_s}$ with value of the exponent found from simulations $\alpha_s \simeq 0.52 - 0.54$ close to $1/2$.

The final approach to the equilibrium is exponentially slow. It is interesting to note also that the radius of gyration of an individual chain after first decreasing as each chain collapses (see Sec. 4.3), starts to grow slowly as each chain slightly expands within a growing cluster. Thus, these observations confirm that clusters formed during the spinodal decomposition in the homopolymer solution, corresponding to local minima of the free energy in the GSC method of Subsec. 6.1.3, are indeed mean-field unstable and decay due to fluctuations.

6.2 Phase Separation in Copolymer Solutions

This section is devoted to studying the phase separation in dilute solution of periodic copolymers.

First, let us overview the main results obtained by numerical analysis of the complete set of the GSC equations (2.73). In Figs. 6.16 and 6.17 we present the

Figure 6.16: Phase diagram for solution of 4 $(ab)_9$ copolymers in terms of the amphiphilicity, Δ , and the mean second virial coefficient, $\bar{u}^{(2)}$. Here the linear box size is $L = 5$ and $u^{(3)} = 10$.

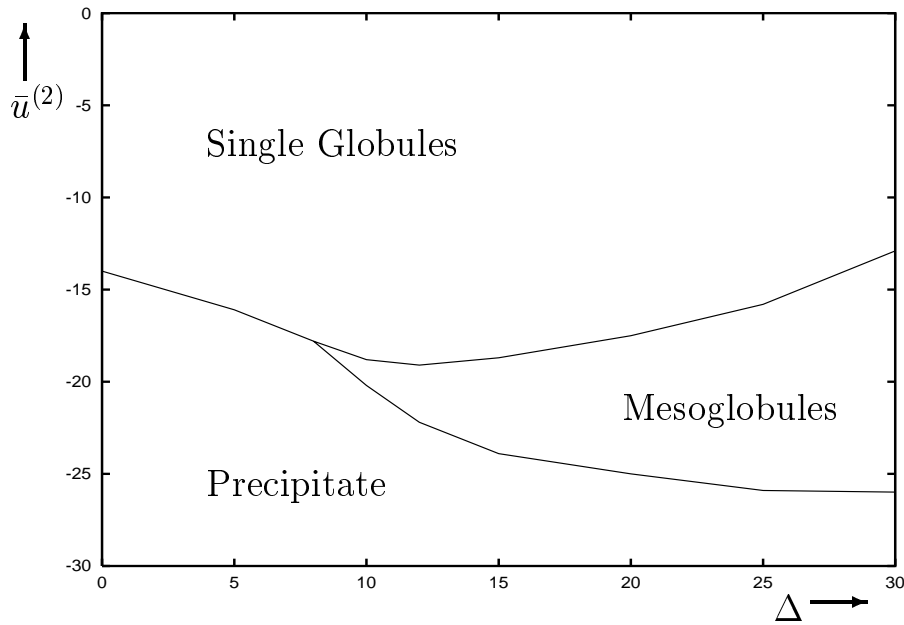
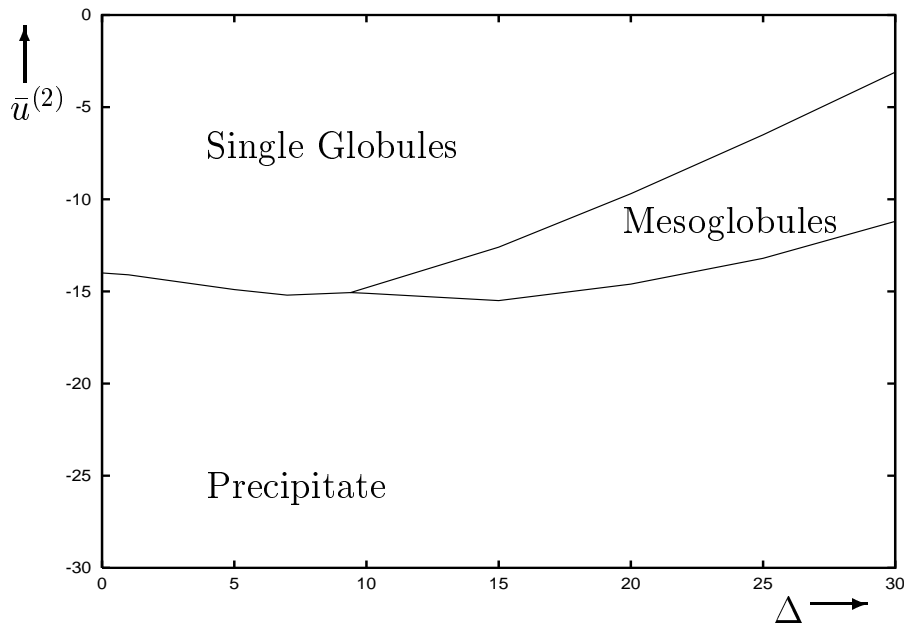


Figure 6.17: Phase diagram for solution of 4 $(a_3b_3)_3$ copolymers in terms of the amphiphilicity, Δ , and the mean second virial coefficient, $\bar{u}^{(2)}$. Here the linear box size is $L = 5$ and $u^{(3)} = 10$.



equilibrium phase diagrams for solution of $M = 4$ copolymers consisting of short and long blocks respectively. These diagrams are drawn at a fixed concentration in terms of the mean second virial coefficient $\bar{u}^{(2)}$ and the amphiphilicity Δ , which parametrise the matrix of the two-body virial coefficients by Eq. (2.15). For small values of the amphiphilicity Δ the phase diagrams of copolymers are essentially the same as for the homopolymer. Thus, there are the low-density phase of individual globules (or coils in the repulsive regime) and the high-density macroglobule, as well as the region of their coexistence.

Let us now discuss how the situation changes with increasing Δ at a fixed low concentration. For the two copolymers under consideration there appears an intermediate region, in which the state corresponding to two clusters of two chains each possesses the lowest free energy. Such a minimum appears starting from some critical value of the amphiphilicity and is bound to a narrow range in the mean second virial coefficient for any fixed Δ . As it is clear from Figs. 6.16 and 6.17, this region designated as the ‘Mesoglobules’ expands with increasing the amphiphilicity. The location and shape of this region turn out very sensitive on the copolymer sequence. For long blocks copolymers this region is narrower and appears at a weaker attraction, characterised by a smaller $|\bar{u}^{(2)}|$ compared to the case of short blocks (alternating monomers). Indeed, for the former the micro-phase separation, which stabilises the mesoglobules, proceeds easier, i.e. it requires a weaker attraction to occur. It is important to emphasise that the asymmetric clusters ‘3 + 1’ and ‘1 + 1 + 2’ always possess a higher free energy than other minima (the macroglobule ‘4’, the single globules ‘1 + 1 + 1 + 1’ or the mesoglobules ‘2 + 2’), and thus are merely metastable.

Now then, let us examine somewhat larger systems composed of $M = 12$ chains of length $N = 12$ with varying block length. Values of the free energy at various local minima, corresponding to symmetric and some asymmetric clusters,

Table 6.2: Values of the specific free energy, $a = \mathcal{A}/MN$, at various minima for the system of 12 $(ab)_6$ copolymers for various values of the mean second virial coefficient, $\bar{u}^{(2)}$.

Here the linear system size and the amphiphilicity are equal to $L = 10$ and $\Delta = 30$.

| $\bar{u}^{(2)}$ | 12×1 | 6×2 | 4×3 | 3×4 | 2×6 | 1×12 | 1, 11 | 2, 10 | 1, 1, 10 |
|-----------------|---------------|---------------|---------------|---------------|---------------|---------------|--------|--------|----------|
| -5 | -4.25 | -3.70 | -3.25 | 2.89 | -2.33 | — | -4.17 | -3.46 | -3.56 |
| -10 | -7.08 | -6.78 | -6.42 | -6.09 | -5.57 | -4.54 | -4.88 | -5.15 | -5.20 |
| -15 | -10.31 | -10.31 | -10.05 | -9.79 | -9.32 | -8.34 | -8.63 | -8.90 | -8.90 |
| -20 | -13.96 | -14.29 | -14.16 | -13.97 | -13.58 | -12.69 | -12.91 | -13.17 | -13.12 |
| -25 | -18.04 | -18.74 | -18.75 | -18.64 | -18.35 | -17.58 | -17.72 | -17.97 | -17.85 |
| -30 | -22.61 | -23.69 | -23.86 | -23.83 | -23.65 | -23.02 | -23.07 | -23.29 | -23.12 |
| -35 | -27.73 | -29.18 | -29.51 | -29.57 | -29.50 | -29.03 | -28.99 | -29.19 | -28.94 |
| -40 | -33.87 | -35.27 | -35.75 | -35.91 | -35.95 | -35.64 | -35.51 | -35.68 | -35.37 |
| -45 | -40.41 | -42.47 | -42.65 | -42.90 | -43.04 | -42.88 | -42.67 | -42.81 | -42.45 |
| -50 | -47.64 | -50.05 | -50.65 | -50.68 | -50.83 | -50.81 | -50.52 | -50.63 | -50.21 |
| -55 | -55.57 | -58.36 | -58.69 | -59.05 | -59.34 | -59.46 | -59.08 | -59.17 | -58.69 |

Table 6.3: Values of the specific free energy, $a = \mathcal{A}/MN$, at various minima for the system of 12 $(a_2b_2)_3$ copolymers for various values of the mean second virial coefficient, $\bar{u}^{(2)}$.

Here the linear system size and the amphiphilicity are equal to $L = 10$ and $\Delta = 30$.

| $\bar{u}^{(2)}$ | 12×1 | 6×2 | 4×3 | 3×4 | 2×6 | 1×12 | 1, 11 | 2, 10 | 1, 1, 10 |
|-----------------|---------------|---------------|---------------|---------------|---------------|---------------|--------|--------|----------|
| -5 | -4.98 | -4.77 | -4.53 | 4.31 | -3.94 | -3.20 | -3.44 | -3.63 | -3.67 |
| -10 | -7.75 | -7.80 | -7.65 | -7.47 | -7.16 | -6.47 | -6.66 | -6.85 | -6.85 |
| -15 | -10.91 | -11.25 | -11.21 | -11.09 | -10.85 | -10.23 | -10.37 | -10.55 | -10.50 |
| -20 | -14.47 | -15.14 | -15.21 | -15.17 | -15.00 | -14.48 | -14.55 | -14.73 | -14.62 |
| -25 | -18.45 | -19.48 | -19.69 | -19.72 | -19.64 | -19.24 | -19.24 | -19.39 | -19.22 |
| -30 | -22.91 | -24.30 | -24.66 | -24.77 | -24.79 | -24.42 | -24.43 | -24.57 | -24.34 |
| -35 | -27.93 | -29.66 | -30.16 | -30.36 | -30.48 | -30.35 | -30.18 | -30.29 | -30.00 |
| -40 | -34.02 | -36.03 | -36.27 | -36.55 | -36.76 | -36.77 | -36.52 | -36.51 | -36.27 |
| -45 | -40.51 | -42.82 | -43.27 | -43.42 | -43.57 | -43.85 | -43.53 | -43.49 | -43.20 |

Table 6.4: Values of the specific free energy, $a = \mathcal{A}/MN$, at various minima for the system of 12 $(a_3b_3)_2$ copolymers for various values of the mean second virial coefficient, $\bar{u}^{(2)}$.

Here the linear system size and the amphiphilicity are equal to $L = 10$ and $\Delta = 30$.

| $\bar{u}^{(2)}$ | 12×1 | 6×2 | 4×3 | 3×4 | 2×6 | 1×12 | 1, 11 | 2, 10 | 1, 1, 10 |
|-----------------|---------------|--------------|---------------|----------------|---------------|---------------|--------|--------|----------|
| 0 | -3.02 | -2.81 | -2.61 | 2.45 | -2.18 | -1.65 | -1.83 | -1.97 | -2.01 |
| -5 | -5.38 | -5.39 | -5.27 | 5.14 | -4.91 | -4.40 | -4.55 | -4.69 | -4.69 |
| -10 | -8.12 | -8.37 | -8.34 | -8.26 | -8.08 | -7.63 | -7.73 | -7.89 | -7.83 |
| -15 | -11.23 | -11.78 | -11.85 | -11.83 | -11.71 | -11.33 | -11.38 | -11.51 | -11.42 |
| -20 | -14.74 | -15.60 | -15.80 | -15.840 | -15.79 | -15.51 | -15.49 | -15.60 | -15.46 |
| -25 | -18.66 | -19.87 | -20.19 | -20.31 | -20.34 | -20.16 | -20.07 | -20.17 | -19.97 |
| -30 | -23.05 | -24.61 | -25.070 | -25.26 | -25.39 | -25.32 | -25.16 | -25.24 | -24.98 |
| -35 | -28.00 | -29.89 | -30.48 | -30.75 | -30.97 | -31.03 | -30.78 | -30.85 | -30.54 |

are presented in the series of Tabs. 6.2, 6.3 and 6.4 for different values of $\bar{u}^{(2)}$ and at a fixed sufficiently high Δ . All considered asymmetric clusters have been found to possess a higher value of the free energy than the symmetric ones. The main conclusion from the above case of a smaller system that the mesoglobules are thermodynamically stable in some intermediate region, remains valid here as well. However, in the current case a few different mesoglobules sizes are possible, namely ‘ 6×2 ’, ‘ 4×3 ’, ‘ 3×4 ’ and ‘ 2×6 ’. We also find that at a given mean second virial coefficient, amphiphilicity, concentration and the sequence, only one of these is thermodynamically stable. With all other parameters fixed, the size of stable mesoglobules increases with the concentration and decreases with $|\bar{u}^{(2)}|$. Thus, the equilibrium transition from the gas of single globules to the macroaggregate on increasing $|\bar{u}^{(2)}|$ proceeds in a few steps. Clearly, the number of various possible clusters grows exponentially with the system size, and for a sufficiently large system it is impossible to enumerate all possible divisions.

It is interesting to check these predictions of the GSC method by the Monte Carlo simulation on a lattice. Here we shall consider amphiphilic copolymers consisting of rather hydrophobic and slightly hydrophilic units. In Figs. 6.18, 6.19,

Figure 6.18: Typical conformations in $(a_3b_3)_2$ copolymer solution at a finite concentration (a, b) from the Monte Carlo simulation.

Here Figs. (a) and (b) correspond respectively to the good ($\chi_{aa} = 0.7$, $\chi_{ab} = 0.28$, $\chi_{bb} = -0.14$) and poor solvent conditions ($\chi_{aa} = 0.8$, $\chi_{ab} = 0.32$, $\chi_{bb} = -0.16$) near the critical point.

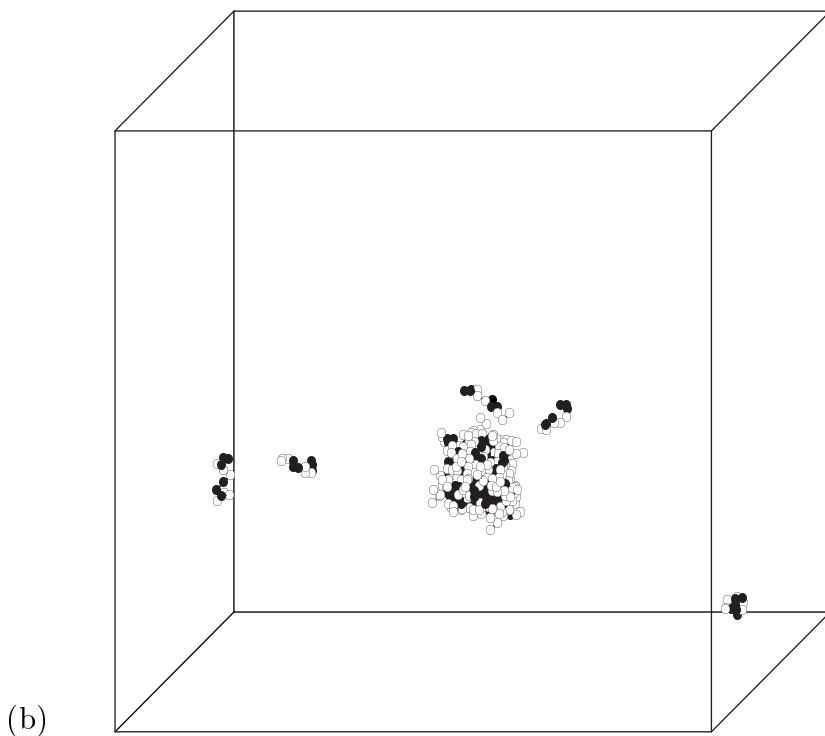
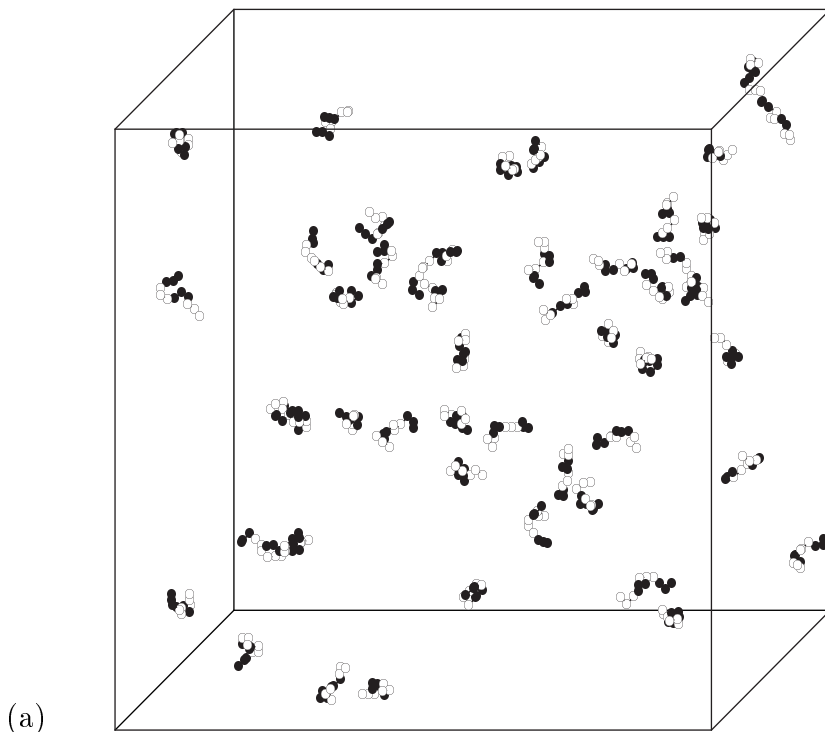


Figure 6.19: Typical conformations in $(a_3b_3)_2$ copolymer solution at a finite concentration (c, d).

Here Figs. (c) and (d) correspond respectively to the macroglobule and mesoglobules at the same interaction parameters, $\chi_{aa} = 1.1$, $\chi_{ab} = 0.44$, $\chi_{bb} = -0.22$.

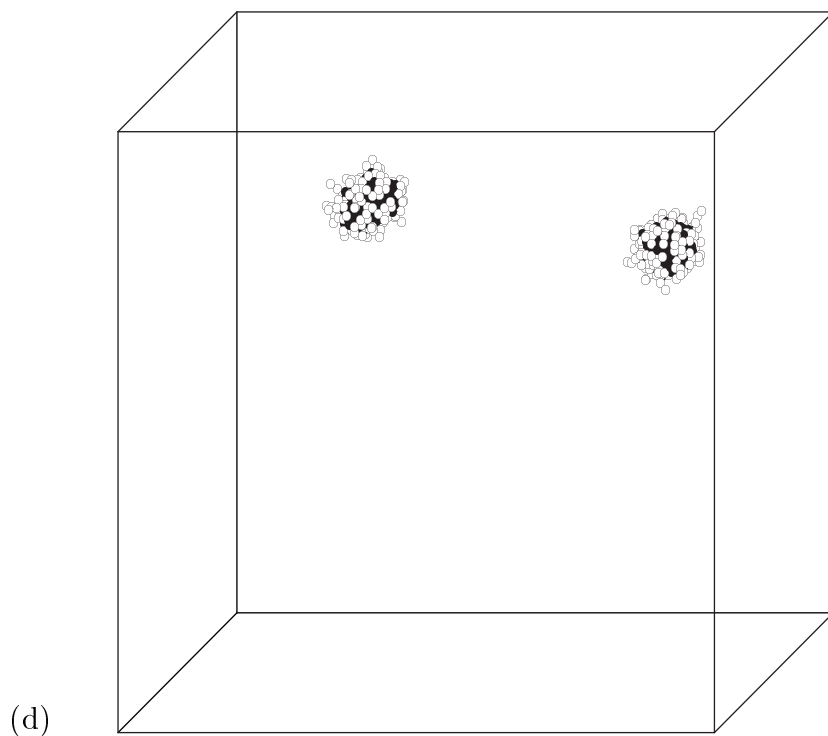
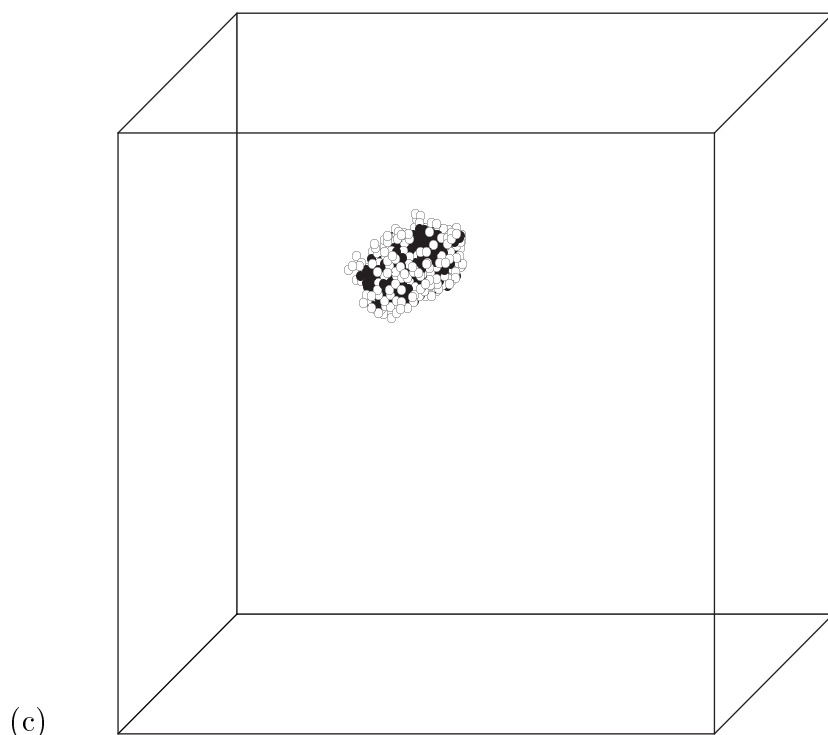


Figure 6.20: Typical conformations in diblock copolymer, (a_6b_6) , solution at a finite concentration (a, b).

Here Figs. (a) and (b) correspond respectively to the good ($\chi_{aa} = 0.5$, $\chi_{ab} = 0.2$, $\chi_{bb} = -0.1$) and poor solvent ($\chi_{aa} = 0.6$, $\chi_{ab} = 0.24$, $\chi_{bb} = -0.12$) near the critical point.

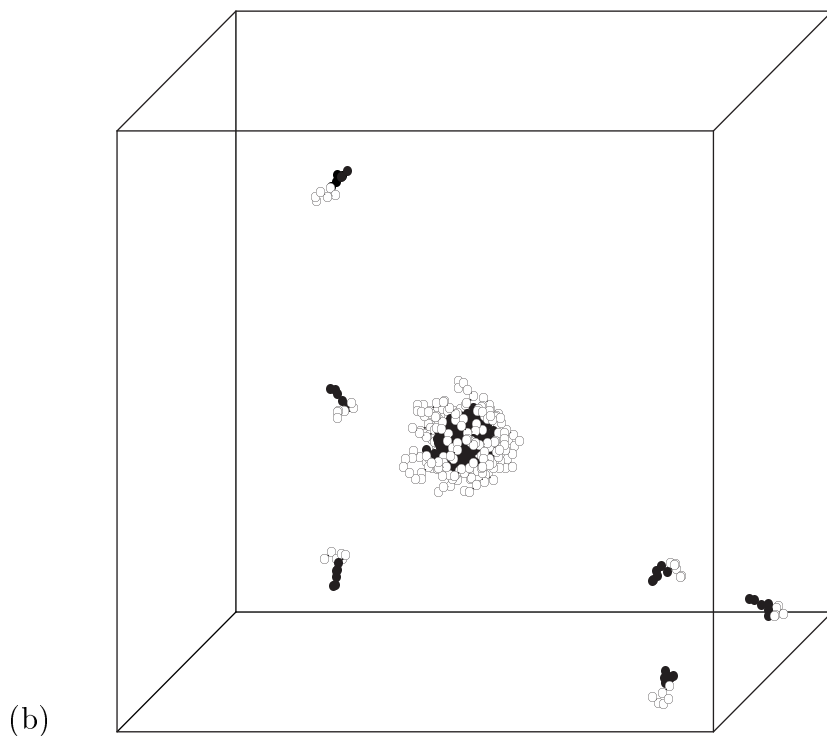
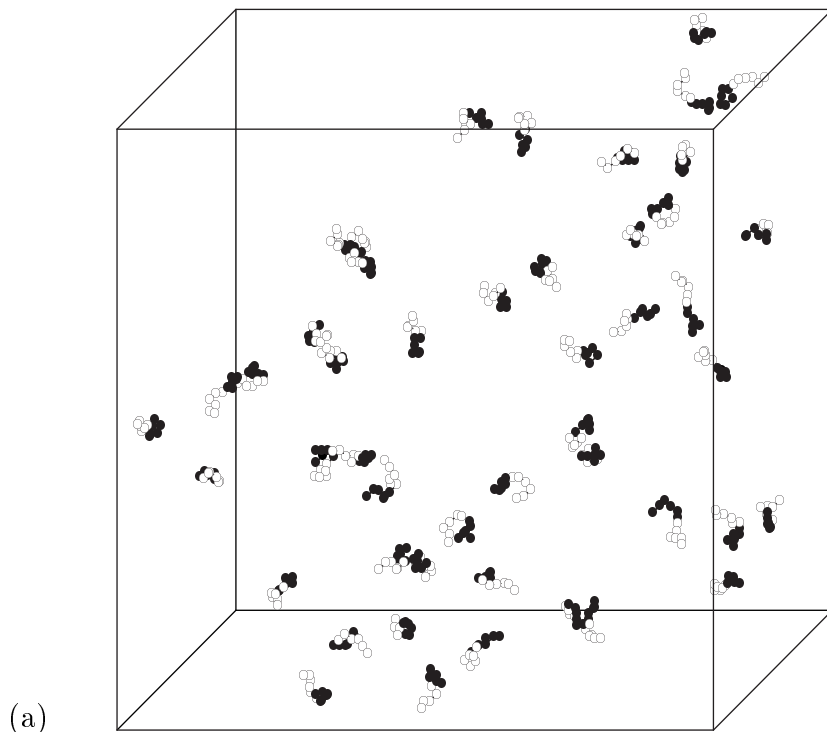
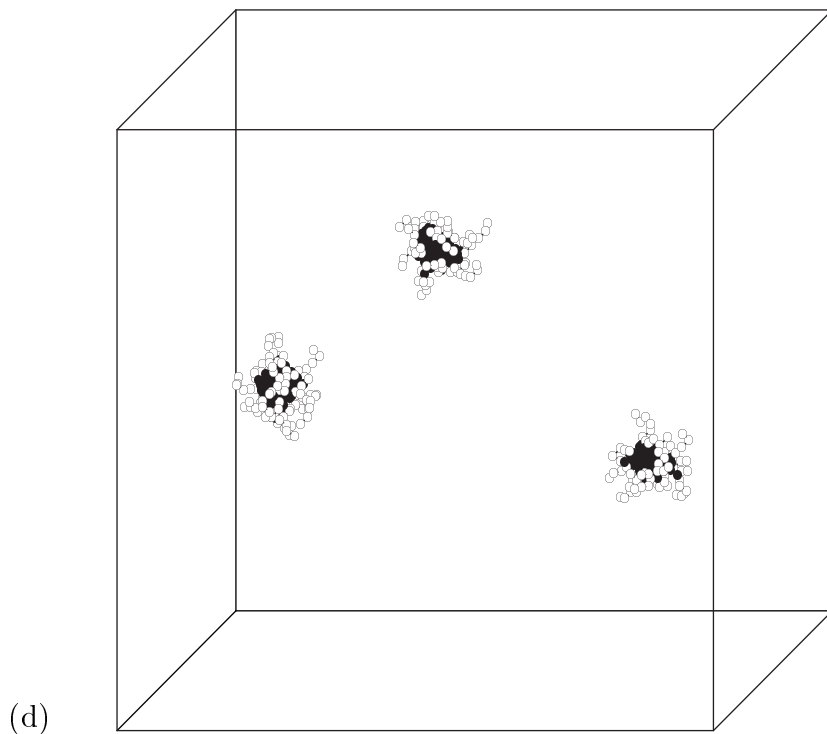
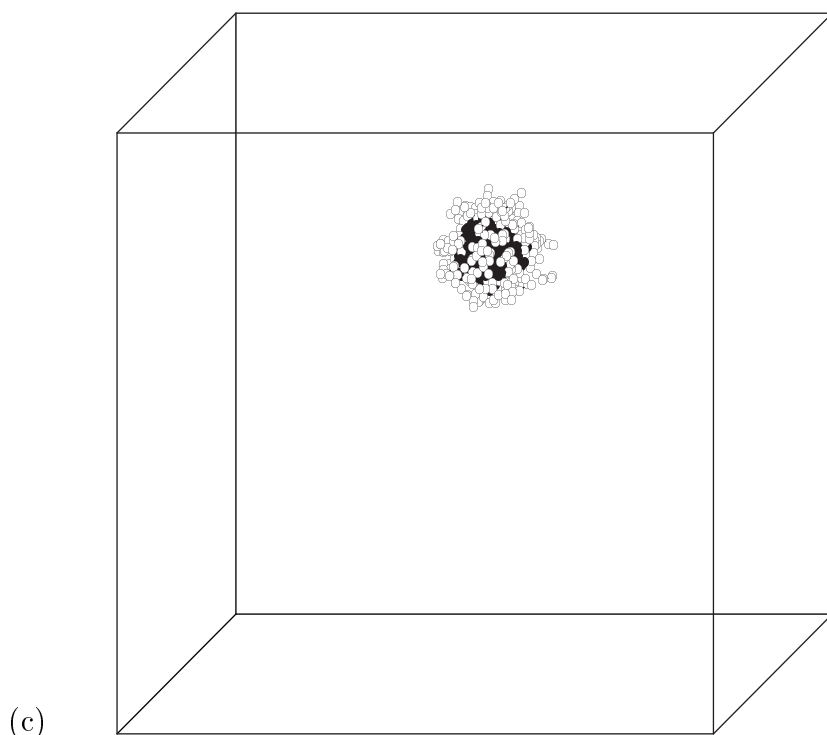


Figure 6.21: Typical conformations in diblock copolymer, (a_6b_6) , solution at a finite concentration (c, d).

Here Figs. (c) and (d) correspond respectively to the macroglobule and mesoglobules at the same interaction parameters, $\chi_{aa} = 0.8$, $\chi_{ab} = 0.32$, $\chi_{bb} = -0.16$.



6.20 and 6.21 we exhibit snapshots of typical system equilibrium conformations for copolymers with two different block lengths at different values of the interaction parameters. In Figs. 6.18a and 6.20a we have snapshots in the regime of a sufficiently weak attraction, for which the system forms a gas of single chains. Similarly to the homopolymer, increasing the attraction χ_{aa} and repulsion χ_{bb} parameters proportionally leads to the phase separation, and the snapshots right after this transition are depicted in Figs. 6.18b and 6.20b. Here we have an aggregate containing most of the chains and a few single globules just as expected. We note that the phase separation transition in copolymers is sequence dependent and occurs earlier for longer blocks, in agreement with a similar conclusion from the above studies of the GSC equations. A further proportional increase of the parameters χ_{aa} and χ_{bb} leads to coalescence of all chains into a simple macroglobule, which is depicted in Figs. 6.18c and 6.20c.

Figs. 6.18d and 6.20d have been obtained by a long kinetic process after the quench from the gas of swollen chains to the same values of the interaction parameters as in Figs. 6.18c and 6.20c. Here we obtain mesoglobules with a quite monodispersed size, which are found to remain stable after a few dozens of millions of Monte Carlo sweeps. A direct check shows that these mesoglobules possess a somewhat higher energy than the macroglobule at the same interaction parameters. However, their entropy is, obviously, higher as well, and the net result in the free energy favours the mesoglobules somewhat, which is manifested by their stability. This interplay of different contributions is rather subtle, and, clearly, the micro-phase separation, which leads to a repulsive shell on the surface of the mesoglobules, does help their stability.

Practically, although the mesoglobules are rather monodispersed, their size can vary a little due to fluctuations. This situation is analogous to the GSC theory, in which clusters of slightly unequal sizes have close, but somewhat higher free

energy than the respective symmetric clusters. This means also that the barriers separating such slightly different minima make it hard for the system to transform from one of the metastable states to the true free energy minimum. Such kinetic arrest is responsible for that we have obtained distinct conformations in Figs. 6.18c and 6.18d by using different preparation procedures. In Monte Carlo simulation fluctuations permit to move a single chain from one cluster to another occasionally, but the average mesoglobules size does not really fluctuate. Transitions between states with different sized mesoglobules are strictly suppressed due to rather high activation barriers. Finally, we remark that adequate description of the nucleation process would require an introduction of collective moves, which can split clusters and form new ones, to the Monte Carlo scheme.

*On a paper submitted by a physicist
colleague:
“This isn’t right. This isn’t even wrong.”
Wolfgang Pauli*

Chapter 7

Conclusion

The main conclusion of this thesis is that by using relatively simple analytical and computational methods of statistical mechanics in conjunction with simple coarse-grained models of macromolecular solutions it is possible to study very complex systems at nonequilibrium conditions.

Now, let us briefly summarise the main results obtained here.

In Chapter 2 the Gaussian Self-consistent method is presented in the most complete form. In Sec. 2.1 the origin of the coarse-grained model, which is used throughout, is explained. In Sec. 2.4 it is shown how the model can be improved to avoid problems with the virial type expansion for dense states. In Sec. 2.3 the earlier derived kinetic equations are reformulated in terms of the mean squared distances between monomers and the relation to the Gibbs–Bogoliubov variational principle of Sec. 2.2 is established. Then in Sec. 2.5 we propose efficient numerical techniques for solution of the GSC equations.

Chapter 3 is devoted to presentation of the Monte Carlo simulation techniques. Sections 3.1 and 3.2 describe the lattice and continuous space Monte Carlo models and methods in detail.

Chapter 4 is devoted to the study of a single macromolecule, i.e infinitely dilute solution. In Sec. 4.1 the extended coil state is considered, and as it is

well understood, the validity of our Monte Carlo techniques is tested and various subtle topological questions associated with that are addressed. Sec. 4.2 is devoted to the study of the milestone phenomenon — the equilibrium coil-to-globule transition for the homopolymer. Again, since this transition has been well studied, we can use it as a test ground for our analytical and computational techniques. The results from the GSC method, and the on lattice and off lattice Monte Carlo simulations are compared. In addition, the pair monomer probability distribution functions are measured, which allows us to understand the limitations of the GSC treatment.

In Sec. 4.3 we study the kinetic laws at the collapse transition of a flexible homopolymer. The linearised GSC equations for the early and late times are derived, which allows us to deduce a set of the near-equilibrium kinetic exponents with and without hydrodynamics. Some of these 12 kinetic exponents have been also reproduced by a somewhat different approach in a recent work [180] and an agreement was found. Then, the overall folding process is discussed in detail and the specific features of folding of an open chain are described. Sec. 4.4 is devoted to the study of the structure of the homopolymer globule and the role of topological restrictions in particular. In Sec. 4.5 we study the conformations and the phase diagram of a stiff macromolecule. The nontrivial toroidal states distinguished by the topological winding number are discovered. These are stable thermodynamic states in strips of the phase diagram, while hairpin conformations, which are also observed, are merely metastable. In Sec. 4.6 we discover that the introduction of the short ranged three-body interactions leads to the appearance of orientationally ordered globular phases of the homopolymer with crystalline and glassy structure.

In chapter 5 we study in detail the equilibrium phase diagrams (Sec. 5.1) and folding kinetics (Sec. 5.2) of various amphiphilic heteropolymers in infinitely

dilute solution. The particular cases of alternating monomers, alternative blocks and random sequences are examined. The states of the liquid-like, micro-phase separated globules and a large number of semi-compacted frustrated phases are found for concrete sequences. The relation of the rugged free energy profile of such systems to the spin glass models is discussed. We discover that the kinetics of folding proceeds in a complicated multi-step manner for alternating and random heteropolymers, and that it is extremely sensitive on the primary sequence. The mechanism of spontaneous symmetry breaking of the kinematic symmetry related to the primary sequence is explained.

In Chapter 6 we extend our considerations to higher concentrations. In Sec. 6.1 it is shown that for solution of identical homopolymer rings in part of the parameter space the GSC equations can be reduced in some approximation to those of the conventional Flory-Huggins theory. In addition, the metastable states corresponding to various clusters obtained by association of distinct chains are examined. In Sec. 6.2 we study the equilibrium phase diagrams of copolymer solutions. In addition to the regions of the swollen coils and the two-phase separation, a number of new thermodynamically stable states are discovered. These correspond to the mesoglobules obtained from a number of chains, that are stabilised and have strong preference to be of a fixed size due to the micro-phase separation. This novel phenomenon predicted by the GSC theory is supported by lattice Monte Carlo simulation, in which the mesoglobules are clearly visible. They also have been recently discovered experimentally [118, 174, 175].

We may conclude that the above described kinetic and equilibrium phenomena are very complex and the accessible thermodynamic and metastable states are extremely diverse even in simple models of macromolecules studied by relatively simple methods that have been used here. The agreement between the results from quite different analytical and computational techniques is encouraging and it

allows us to believe that the observed phenomena are indeed genuine. However, computational difficulties and limitations of the theory have permitted us to describe many complex features only at a qualitative level.

One may hope that better and more accurate methods can be developed for the systems under study in time. This is, of course, a rather complicated task and the efforts required for their development may be considerable. Would such improved methods be really justified in terms of the efforts/achievements ratio? On the one hand, probably not. One is tempted to speculate that the behaviour of these complex systems is somehow dominated by their complexity itself.

On the other hand, there are well founded methods of the statistical mechanics that have been successfully developed for simple liquids. Most important among them are the methods formulated in terms of the structure factors (correlation functions). One can mention here e.g. the approach based on site-site extensions of the Ornstein–Zernike equations. These methods are known to be very accurate for the correlation functions of monoatomic and diatomic liquids. However, unfortunately they are not as good for the thermodynamic characteristics. For instance, the RISM method [111], although pretty accurate inside each phase separately, has serious problems already in trying to describe the simple two-phase coexistence at a first order phase transition. But, clearly, for heteropolymers we do know that the thermodynamics is most crucial. There are just too many metastable states, and perhaps they are all one can observe in experiment in some cases. From a somewhat different perspective the traditional polymer methods based on the density variables methods do not account properly for the connectivity, and thus are only valid for dense states. They also do not allow to extract explicitly the conformational structure of the chain. This limits their applicability to melts and concentrated solutions.

What realistic ways to make a further progress may be foreseen? One view

is that variational methods such as the GSC method can be improved. For this more complicated trial Hamiltonians, i.e. more realistic (non-Gaussian) trial probability distributions should be considered. This, however, appears technically very difficult at the moment as one has to manage the averaging over such a trial distribution exactly. Extension to kinetics is even less clear. Thus, at the moment perhaps the Gaussian self-consistent method, complemented by conventional Monte Carlo techniques and powerful modern computers, may be viewed probably as a reasonable practical tool for studying complex systems, as well as a prototype for improved computational techniques of the future.

Bibliography

Papers on which this thesis is based

- [I] E.G. Timoshenko, Yu.A. Kuznetsov, K.A. Dawson. “Conformational transitions of heteropolymers in dilute solutions.” *Phys. Rev.*, Vol **E 57**, No 6, pp. 6801-6811 (1998).
- [II] Yu.A. Kuznetsov, E.G. Timoshenko, K.A. Dawson. “Kinetics at the collapse transition of homopolymers and random copolymers.” *J. Chem. Phys.*, Vol. **103**, No 11, pp. 4807-4818 (1995).
- [III] Yu.A. Kuznetsov, E.G. Timoshenko, K.A. Dawson. “New orientationally ordered phases of a homopolymer.” *J. Chem. Phys.*, Vol. **104**, No 1, pp. 336-341 (1996).
- [IV] E.G. Timoshenko, Yu.A. Kuznetsov, K.A. Dawson. “Gaussian self-consistent approach to multichain polymer systems.” *Physica*, Vol. **A 240**, No 3-4, pp. 432-442 (1997).
- [V] E.G. Timoshenko, Yu.A. Kuznetsov, K.A. Dawson. “Simple Models of Biological Macromolecules.” To be published in *the Proceedings of “XII International HEPQFT Workshop.”* Samara, Russia, 4 - 10 September 1997.
- [VI] K.A. Dawson, A.V. Gorelov, E.G. Timoshenko, Yu.A. Kuznetsov, A. Du

- Chesne. "Formation of mesoglobules from phase separation in dilute polymer solutions; A study in experiment, theory, and applications." *Physica*, Vol. **A 244**, pp. 68-80 (1997).
- [VII] E.G. Timoshenko, Yu.A. Kuznetsov. "A non-equilibrium approach for polymer solutions." *Il Nuovo Cimento*, Vol. **20D**, No 12bis, pp. 2359-2364 (1998).
- [VIII] Yu.A. Kuznetsov, E.G. Timoshenko. "Self-consistent treatment of copolymers with arbitrary sequences." *Il Nuovo Cimento*, Vol. **20D**, No 12bis, pp. 2265-2288 (1998).
- [IX] Yu.A. Kuznetsov, E.G. Timoshenko, K.A. Dawson. "Conformational structure of amphiphilic copolymers in dilute solutions." *Physica A*, Vol. **257**, No 1-4, pp. 61-76 (1998).

Papers in preparation (updated)

- [X] E.G. Timoshenko, Yu.A. Kuznetsov. "Analysis of stability of macromolecular clusters in dilute heteropolymer solutions." *J. Chem. Phys.*, Vol. **112**, No 18, pp. 8163-8175 (2000).
- [XI] E.G. Timoshenko, Yu.A. Kuznetsov. "Kinetic laws at the collapse transition of the homopolymer revisited." To be submitted to *J. Chem. Phys.*
- [XII] Yu.A. Kuznetsov, E.G. Timoshenko. "Influence of reptations on conformations of the homopolymer in Monte Carlo simulation." *Progr. Colloid Polym. Sci.*, Vol. **115**, No 30, pp. 146-150 (2000).
- [XIII] Yu.A. Kuznetsov, E.G. Timoshenko. "On the conformational structure of a stiff homopolymer." *J. Chem. Phys.*, Vol. **111**, No 8, pp. 3744-3752 (1999).

Other references

- [1] G. Chala, *Polymer Chemistry: an Introduction* (Ellis Horwood, 1993).
- [2] D. Shaw, *Introduction to Colloid and Surface Chemistry* (Butterworth–Heinemann, Oxford, 1992).
- [3] H. Frauenfelder, in *Structure and dynamics of nucleic acids, proteins and membranes*, Ed. by E. Clementi, S. Chin (Plenum, 1986).
- [4] H. Gould and J. Tobochnik, *An Introduction to Computer Simulation Methods* (Addison–Wesley, 1988).
- [5] M.P. Allen and D.J. Tildesley (Ed.), *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987).
- [6] K. Binder (Ed), *Monte Carlo and Molecular Dynamics Simulation in Polymer Science* (Oxford University Press, Oxford, 1995).
- [7] N. Metropolis, A.W. Rosebluth, M.N. Rosenbluth, A.N. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [8] K. Binder (Ed), *Applications of Monte Carlo Method in Statistical Physics* (Springer–Verlag, Berlin, 1984).
- [9] M.P. Allen and D.J. Tildesley (Ed.), *Computer Simulation in Chemical Physics. NATO ASI Series C: Math. and Phys. Sci.* **397** (Kluwer Academic, 1993).
- [10] K. Binder and W. Paul, *J. Polym. Sci.* **B 35**, 1 (1997).
- [11] P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, NY, 3rd printing, 1988).

- [12] P. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, NY, 1971).
- [13] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford Science, New York, 1989).
- [14] M. Bishop, M.H. Kalos, and H.L. Frisch, *J. Chem. Phys.* **70**, 1299 (1979).
- [15] A. Baumgartner and K. Binder, *J. Chem. Phys.* **71**, 2541 (1979).
- [16] I. Webman, J.L. Lebowitz, and M.H. Kalos, *Phys. Rev.* **21**, 5540 (1980).
- [17] A. Baumgartner and K. Binder, *J. Chem. Phys.* **75**, 2994 (1981).
- [18] W. Bruns and R. Bansal, *J. Chem. Phys.* **75**, 5149 (1981).
- [19] K. Kremer and K. Binder, *Comput. Phys. Repts.* **7**, 259 (1988).
- [20] P.H. Verdier and W.H. Stockmayer, *J. Chem. Phys.* **36**, 227 (1962).
- [21] P.H. Verdier, *J. Chem. Phys.* **45**, 2122 (1966).
- [22] P.H. Verdier, *J. Chem. Phys.* **52**, 5512 (1970).
- [23] P.H. Verdier, *J. Chem. Phys.* **59**, 6119 (1973).
- [24] H.J. Hilhorst and J.M. Deutsch, *J. Chem. Phys.* **63**, 5153 (1975).
- [25] F.T. Wall and F. Mandel, *J. Chem. Phys.* **63**, 4542 (1975).
- [26] M. Lal, *Molec. Phys.* **17**, 57 (1969).
- [27] A. Sariban and K. Binder, *Macromolecules* **21**, 711 (1988).
- [28] I. Carmesin and K. Kremer, *Macromolecules* **21**, 2819 (1988).
- [29] H.-P. Deutsch and K. Binder, *J. Chem. Phys.* **94**, 2294 (1991).

- [30] W. Paul, K. Binder, D.W. Heermann, and K. Kremer, *Macromolecules* **24**, 6332 (1991).
- [31] J.S. Shaffer, *J. Chem. Phys.* **101**, 4205 (1994).
- [32] M. Mezard, G. Parisi, and M. Virasoro, *Spin glass theory and beyond* (World Scientific, Singapore, 1987).
- [33] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, 1992).
- [34] M.V. Volkenstein, *Configurational Statistics of Polymeric Chains* (Interscience, N.Y., 1963).
- [35] F. Tanaka, *J. Chem. Phys.* **82**, 4707 (1985).
- [36] M. Daoud and G. Janninck, *J. Phys. (Paris)* **37**, 973 (1976); M. Daoud, *J. Polym. Sci. Polym. Symp.* **61**, 305 (1977).
- [37] J. des Cloizeaux and G. Jannink, *Polymers in Solution* (Clarendon Press, Oxford, 1990).
- [38] A. Yu. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (AIP, NY, 1994).
- [39] B. Duplantier, *J. Phys. (Paris)* **43**, 991 (1982); *Europhys. Lett.* **1**, 491 (1986).
- [40] H. Yamakawa, *Modern Theory of Polymer Solutions* (Harper & Row, New York, 1971).
- [41] S.F. Edwards and K.F. Freed, *J. Chem. Phys.* **61**, 1189 (1974).
- [42] B.H. Zimm, G.M. Roe, and L.F. Epstein, *J. Chem. Phys.* **37**, 2547 (1962).

- [43] A.J. Peterlin, *J. Chem. Phys.* **23**, 2464 (1955).
- [44] Y. Oono and M. Kohmoto, *J. Chem. Phys.* **78**, 520 (1983).
- [45] O.B. Ptitsyn, A.K. Kron, and Y.Y. Eizner, *J. Polym. Sci., Polym. Phys. Ed.* **16**, 3509 (1968).
- [46] W. H. Stockmayer, *Makromol. Chem.* **35**, 54 (1960).
- [47] B. Chu and Z. Wang, *Macromolecules* **21**, 2283 (1988); B. Chu, R. Xu, and J. Zhuo, *Macromolecules* **21**, 273 (1988); J. Yu, Z. Wang, and B. Chu, *Macromolecules* **25**, 1618 (1992); B. Chu, Q. Ying, and A.Yu. Grosberg, *Macromolecules* **28**, 180 (1995).
- [48] S.-T. Sun, I. Nishio, G. Swislow, and T. Tanaka, *J. Chem. Phys.* **73**, 5971 (1980).
- [49] M. Meewes, J. Rička, M. de Silva, R. Nyffenegger, and T. Binkert, *Macromolecules* **24**, 5811 (1991).
- [50] H.G. Schild, *Prog. Polym. Sci.* **17**, 163 (1992).
- [51] S. Fujishige, K. Kubota, and I. Ando, *J. Phys. Chem.* **93**, 3311 (1989).
- [52] C. Wu and S. Zhou, *Macromolecules* **28**, 8381 (1995).
- [53] P.J. Flory, *J. Chem. Phys.* **9**, 660 (1941).
- [54] M.L. Huggins, *J. Chem. Phys.* **9**, 440 (1941).
- [55] G. Allegra and F. Ganazzoli, *J. Chem. Phys.* **83(1)**, 397 (1985); *ibid* **87**, 1817 (1987).
- [56] G. Allegra and F. Ganazzoli, *Adv. Chem. Phys.* **75**, 265 (1989).

- [57] G. Allegra, E. Colombo, and F. Ganazzoli, *Macromolecules* **26**(2), 330 (1993).
- [58] P.G. de Gennes, *Phys. Lett.* **38 A**, 339 (1972).
- [59] J. des Cloizeaux, *Phys. Rev.* **A 10**, 1665 (1974); *J. Physique* **42**, 635 (1981).
- [60] H.C. Öttinger and Y. Rabin, *J. Non-Newt. Fluid Mech.* **33**, 53 (1989).
- [61] C. Tanford, *The Hydrophobic Effect* (Wiley, New York, 1980).
- [62] K. A. Dill, *Biochemistry*, **29**, 7133 (1990).
- [63] S.P. Obukhov, *J. Phys.* **A 19**, 3655 (1986).
- [64] C. D. Sfatos, A. M. Gutin, and E. I. Shakhnovich, *Phys. Rev.* **E 48**, 465 (1993); *Phys. Rev.* **E 51**, 4727 (1995).
- [65] V.S. Pande, A.Yu. Grosberg, and T. Tanaka, *Proc. Natl. Acad. Sci. USA* **91**, 12976 (1994); *J. Chem. Phys.* **101** (9), 8246 (1994); *Phys. Rev.* **E 51** (4), 3381 (1995).
- [66] T. Garel and H. Orland, *Europhys. Lett.* **6**, 597 (1988).
- [67] G. H. Fredrickson and E. Helfand, *J. Chem. Phys.* **87**, 697 (1987).
- [68] G.H. Fredrickson, S.T. Milner, and L. Leibler, *Macromolecules* **25**, 6341 (1992).
- [69] J. Garcia-Ojalvo, J.M. Sancho, and L. Ramirez-Piscina, *Phys. Rev.* **A 46**, 4670 (1992).
- [70] D. R. Paul and S. Neuman (Ed.), *Polymer Blends* (Academic Press, 1987).
- [71] I.S. Miles and S. Rostami (Ed.), *Multicomponent Polymer Systems* (Longman Scientific and Technical, 1992).

- [72] D. J. Meier, *J. Polym. Sci.* **C26**, 81 (1969); *J. Polym. Prepr.* **11**, 400 (1970); **15**, 171 (1974).
- [73] A. N. Semenov, *Sov. Phys. JETP* **61**, 733 (1985).
- [74] T. Ohta and K. Kawasaki, *Macromolecules* **19**, 2621 (1986).
- [75] L. Leibler, *Macromolecules* **13**, 1602 (1980).
- [76] T. Hashimoto, *Macromolecules* **20**, 465 (1987).
- [77] A.V. Dobrynin and I.Y. Erukhimovich, *Pis'ma Zh. Eksp. Teor. Fiz.* **53** (**11**), 545 (1991); *J. Phys. (France) I* **5**, 677 (1995).
- [78] S.V. Panyukov and S.I. Kuchanov, *Sov. Phys. JETP* **72** (**2**), 368 (1991).
- [79] H. Yamakawa, *Ann. Rev. Phys. Chem.* **35**, 23 (1984).
- [80] V.A. Bloomfield, *Biopolymers* **31**, 1471 (1991).
- [81] J. Ubbink and T. Odijk, *Biophys. J.* **68**, 54 (1995); *ibid Europhys. Lett.* **33** (**5**), 353 (1996).
- [82] N.V. Hud, K.H. Downing, and R. Balhorn, *Proc. Natl. Acad. Sci. USA* **92**, 3581 (1995).
- [83] Yu. A. Kuznetsov, E. G. Timoshenko, and K. A. Dawson, *J. Chem. Phys.* **105**(**16**), 7116 (1996).
- [84] L.S. Lerman, *Proc. Natl. Acad. Sci. USA* **68**, 1886 (1971).
- [85] U.K. Laemmli, *Proc. Natl. Acad. Sci. USA* **72**, 4288 (1975).
- [86] Yu.M. Evdokimov *et al*, *Nucl. Acids Res.* **3**, 2353 (1976).
- [87] G.E. Plum, P.G. Arscott, and V.A. Bloomfield, *Biopolymers* **30**, 631 (1990).

- [88] V.V. Vasilevskaya, A.R. Khokhlov Y. Matsuzawa, and K. Yoshikawa, *J. Chem. Phys.* **102**, 6595 (1995).
- [89] A.Yu. Grosberg, *Biophysics* **24**, 30 (1979); A.Yu. Grosberg and A.R. Khokhlov, *Adv. Polym. Sci.* **41**, 53 (1981).
- [90] O. Kratky and G. Porod, *Rec. Trav. Chim.* **68**, 1106 (1949).
- [91] R.A. Harris and J.E. Hearst, *J. Chem. Phys.* **44**, 2595 (1966).
- [92] B.H. Zimm, *J. Chem. Phys.* **24**, 269 (1956).
- [93] M. Bixon and R. Zwanzig, *J. Chem. Phys.* **68**, 1896 (1978).
- [94] A. Perico, S. Bisio, and C. Cuniberti, *Macromolecules* **17**, 2686 (1984).
- [95] K.F. Lau and K.A. Dill, *Macromolecules* **22**, 3986 (1989).
- [96] K.A. Dill, K.M. Fiebig, and H.S. Chan, *Proc. Natl. Acad. Sci. USA* **90**, 1942 (1993).
- [97] K.A. Dill, S. Bromberg, K. Yue, K.M. Fiebig, D.P. Yee, P.D. Thomas, and H.S. Chan, *Protein Sci.* **4**, 561 (1995).
- [98] C.J. Camacho and D. Thirumalai, *Phys. Rev. Lett.* **71**, 2505 (1993); *Proc. Natl. Acad. Sci. USA* **90**, 6369 (1993); *ibid* **90**, 1277 (1995).
- [99] J.D. Bryngelson and P.G. Wolynes, *Proc. Natl. Acad. Sci. USA* **84**, 7524 (1987).
- [100] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, and P.G. Wolynes, *Proteins Struct. Funct. Genet.* **21**, 167 (1995).
- [101] P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, *Science* **267**, 1619 (1995).

- [102] E.I. Shakhnovich and A.M. Gutin, *J. Phys.* **A 22**, 1647 (1989); *J. Phys. (France)* **50**, 1843 (1989).
- [103] J. D. Honeycutt and D. Thirumalai, *Biopolymers* **32**, 695 (1992); D. Thirumalai, *J. de Phys. I* **5**, 1457 (1995).
- [104] P. E. Leopold, M. Montal, and J. Onuchic, *Proc. Natl. Acad. Sci. USA* **89**, 9721 (1992); N. D. Socci and J. N. Onuchic, *J. Chem. Phys.* **103** (11), 4732 (1995).
- [105] B. Ostrovsky and Y. Bar-Yam, *Comp. Polym. Sci.* **3**, 9 (1993); M. A. Smith, Y. Bar-Yam, Y. Rabin, B. Ostrovski, C. A. Bennett, N. Margolus, and T. Toffoli, *Comp. Polym. Sci.* **2**, 165 (1992).
- [106] B. Ostrovsky and Y. Bar-Yam, *Europhys. Lett.* **25**, 409 (1994).
- [107] Y. Bar-Yam, Y. Rabin, and M. A. Smith, *Macromolecules Reprints*, **25**, 2985 (1992).
- [108] A. Byrne, P. Kiernan, D. Green, and K. A. Dawson, *J. Chem. Phys.* **102**, 573 (1995).
- [109] T. A. Kavassalis and P. R. Sundrarajan, *Macromolecules* **26**, 4144 (1993); G. Tanaka and W. L. Mattice, *Macromolecules* **28**, 1049 (1995).
- [110] S. Ma, *Statistical Mechanics* (World Scientific, 1985).
- [111] J.-P. Hansen and I.R. McDonald, *Theory of Simple Liquids* (Academic Press, London, 1986).
- [112] G. Parisi, *Statistical field theory* (Addison-Wesley, 1988).
- [113] I.C. Sanchez, *J. Phys. Chem.* **93**, 6983 (1989);

- [114] K.S. Schmitz, *An Introduction to Dynamic Light Scattering by Macromolecules* (Academic Press, Boston, 1990).
- [115] E. G. Timoshenko and K. A. Dawson, *Phys. Rev. E* **51(1)**, 492 (1995).
- [116] E. G. Timoshenko, Yu. A. Kuznetsov, and K. A. Dawson, *J. Chem. Phys.* **102(4)**, 1816 (1995).
- [117] K.A. Dawson, E.G. Timoshenko, and Yu.A. Kuznetsov. *Physica A* **236(1-2)**, 58 (1997).
- [118] A.V. Gorelov, L.N. Vasil'eva, A. du Chesne, E.G. Timoshenko, Yu.A. Kuznetsov, and K.A. Dawson. *Il Nuovo Cimento D* **16 (7)**, 711 (1994).
- [119] J. des Cloizeaux, *J. Phys. (France)* **31**, 715 (1970).
- [120] S.F. Edwards and P. Singh, *J. Chem. Soc. Faraday Trans. II* **75**, 1001 (1979).
- [121] D. Bratko and K.A. Dawson, *J. Chem. Phys.* **99(7)**, 5352 (1993).
- [122] Yu. A. Kuznetsov, E. G. Timoshenko, and K. A. Dawson, *J. Chem. Phys.* **104(9)**, 3338 (1996).
- [123] F. Ganazzoli, *Macromolecules* **25(26)**, 7357 (1992).
- [124] G. Raos, G. Allegra, and F. Ganazzoli, *J. Chem. Phys.* **100(10)**, 7804 (1994).
- [125] F. Ganazzoli, R. La Ferla, and G. Allegra, *Macromolecules* **28(15)**, 5285 (1995).
- [126] G. Raos and G. Allegra, *J. Chem. Phys.* **104(4)**, 1626 (1996).

- [127] P. G. de Gennes, *J. Phys. Lett.* **46**, L639 (1985).
- [128] E. G. Timoshenko, Yu. A. Kuznetsov, and K. A. Dawson, *Phys. Rev. E* **53(4)**, 3886 (1996).
- [129] E. Shakhnovich and A.M. Gutin, *Nature* **346**, 773 (1990).
- [130] A. Sali, E. Shakhnovich, and M. Karplus, *Nature* **369**, 248 (1994).
- [131] T.E. Creighton (Ed.), *Protein Folding* (Wiley, New York, 1992).
- [132] R. Elber (Ed.), *New Developments in Theoretical Studies of Proteins* (World Scientific, Singapore, 1994).
- [133] H. A. Scheraga, *Pure Appl. Chem.* **36**, 1 (1973).
- [134] N. Go and H. Taketomi, *Proc. Natl. Acad. Sci. USA* **75 (2)**, 559 (1978).
- [135] O. Ptitsyn and A. Finkelstein, *Q. Rev. Biophys.* **13**, 339 (1980).
- [136] P. Privalov, *Adv. Protein Chem.* **35**, 1 (1982).
- [137] O. Ptitsyn, *J. Protein Chem.* **6 (4)**, 273 (1987).
- [138] K.A. Dill, *Biochemistry* **24**, 1501 (1985).
- [139] H. Frauenfelder, S.G. Sligar, and P.G. Wolynes, *Science* **254**, 1598 (1991).
- [140] J. Darnell, H. Lodish, and D. Baltimore, *Molecular Cell Biology* (Scientific American Books, 1990).
- [141] P. G. Wolynes, in *Spin glass ideas in biology*, edited by D. Stein (World Scientific, Singapore, 1991).
- [142] M. Mezard and G. Parisi, *J. Phys.* **I 1**, 809 (1991).

- [143] H. Sompolinsky and A. Zippelius, *Phys. Rev. Lett.* **45**, 359 (1981); *Phys. Rev.* **B 25**, 274 (1982).
- [144] W. Götze and L. Sjögren, *Z. Phys.* **B 65**, 415 (1987); *J. Phys.* **C 21**, 3407 (1988).
- [145] G. Z. Archonts and E. I. Shakhnovich, *Phys. Rev.* **E 49 (4)**, 3109 (1994).
- [146] H.S. Chan and K.A. Dill, *Macromolecules* **22**, 4559 (1989); *J. Chem. Phys.* **95**, 3775 (1991); *ibid* **96**, 768 (1992); *ibid* **99**, 2116 (1993).
- [147] E.I. Shakhnovich, *Phys. Rev. Lett.* **72**, 3907 (1994).
- [148] A. Dinner, A. Sali, M. Karplus, and E. Shakhnovich, *J. Chem. Phys.* **101 (2)**, 1444 (1994).
- [149] E. I. Shakhnovich and A. M. Gutin, *J. Phys.* **A 22**, 1647 (1989); *J. Chem. Phys.*, **93 (8)**, 5967 (1990).
- [150] E. G. Timoshenko, Yu. A. Kuznetsov, and K. A. Dawson, *Phys. Rev.* **E 54(4)**, 4071 (1996).
- [151] E. G. Timoshenko, Yu. A. Kuznetsov, and K. A. Dawson, *Phys. Rev.* **E 55(5)**, 5750 (1997).
- [152] E.G. Timoshenko, Yu.A. Kuznetsov, and K.A. Dawson, *J. Stat. Phys.* **89(1-2)**, 347 (1997).
- [153] J.D. Bryngelson, *J. Chem. Phys.* **100**, 6038 (1994).
- [154] K. Yue and K.A. Dill, *Proc. Natl. Acad. Sci. USA* **89**, 4163 (1992).
- [155] E.I. Shakhnovich and A.M. Gutin, *Proc. Natl. Acad. Sci. USA* **90**, 7195 (1993).

- [156] V.S. Pande, A.Yu. Grosberg, and T. Tanaka, *Macromolecules* **28**, 2218 (1995); *J. Chem. Phys.* **103** (21), 9482 (1995).
- [157] Yu.A. Kuznetsov, *Kinetics of Conformational Transitions of Biopolymers in Infinitely Dilute Solution*. PhD thesis in Biophysics, Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, Pushchino, Moscow Reg., Russia (in Russian, 1996).
- [158] S. Panyukov and Y. Rabin, *Phys. Rep.* **269** (1,2), 1 (1996).
- [159] S. Panyukov, Y. Rabin, and A. Feigel, *Europhys. Lett.* **28**, 149 (1994).
- [160] A.R. Khokhlov and E.Yu. Kramarenko, *Macromol. Theory Simul.* **3**, 45 (1994).
- [161] L.I. Reichl, *A Modern Course in Statistical Physics* (Univ. of Texas Press, Austin, 1980).
- [162] T. Garel, and H. Orland, unpublished, preprint (1995).
- [163] H. Risken, *The Fokker–Planck Equation* (Springer–Verlag, Berlin, 1989).
- [164] G. Parisi, in *Les Houches Session XLIII*, ed. by K. Osterwalder and R. Stora, North–Holland, Amsterdam, 1986.
- [165] M. Creutz, *Quarks, gluons and lattices* (Cambridge University Press, 1983).
- [166] H.H. Gan and B.C. Eu, *J. Chem. Phys.* **99**(5), 4084 (1993).
- [167] H.H. Gan and B.C. Eu, *J. Chem. Phys.* **99**(5), 4103 (1993).
- [168] M.P. Taylor and J.E.G. Lipson, *J. Chem. Phys.* **104**(12), 4835 (1996).
- [169] M.P. Taylor and J.E.G. Lipson, *J. Chem. Phys.* **106**(12), 5181 (1997).

- [170] A. Buguin, P.G. de Gennes, and F. Brochart–Wyart, *C. R. Acad. Sci. Paris* **322**, 741 (1996).
- [171] A. Yu. Grosberg and D. V. Kuznetsov, *Macromolecules* **26**, 4249 (1993).
- [172] P.G. de Gennes and J. Prost, *The physics of liquid crystals* (Clarendon Press, Oxford, 1993).
- [173] G. Parisi, in *Proceedings of the International Conference on Morphology and Kinetics of Phase Separating Complex Fluids*, edited by F. Mallamace (Messina, Italy, 1997).
- [174] Y. Deng and R. Pelton, *Macromolecules* **28**, 4617 (1995).
- [175] X. Qiu, C.M.S. Kwan, and C. Wu, *Macromolecules* **30**, 6090 (1997).
- [176] J.S. Langer and A.J. Schwartz, *Ann. Phys.* **54**, 258 (1969).
- [177] J.S. Langer, *Phys. Rev. A* **21**, 948 (1980).
- [178] I. Lifshitz, and V. Slyozov, *J. Phys. Chem. Solids* **19**, 35 (1961).
- [179] J.D. Gunton and M. Droz, *Lectures Notes in Physics* **183: Introduction to the Theory of Metastable and Unstable States (Springer–Verlag, 1983).**
- [180] H. Orland and E. Pitard, *Europhys. Lett.* **41(4)**, 467 (1998).

Real programmers don't comment their code. It was hard to write, it should be hard to understand.

A common wisdom

Appendix A

In this appendix we present listing of the C code for the kernel of the program called `many_cop` version 0.0.4alpha, which performs Monte Carlo simulation for several copolymer chains of identical length on a cubic lattice with reflective boundary conditions. However, this code is not complete, and about the same number of lines as presented, which contain additional routines for run-time calculation of various observables and data integrity checks, are suppressed.

The distribution consists of four files:

| | |
|------------------------------|--|
| <code>many_cop.cfg</code> | An example of input configuration file |
| <code>many_cop.h</code> | Header file with data and function prototypes |
| <code>many_cop.c</code> | C code of the general interface and initialisation and deinitialisation routines |
| <code>many_cop_move.c</code> | C code of routines performing various Monte Carlo moves described in detail in Sec. 3.1 |

These C files can be compiled using the following commands:

```
cc -c -I. -O2 many_cop.c
```

```
cc -c -I. -O2 many_cop_move.c
```

To make the program some additional routines (see bottom of Page 1 and four top lines on Page 2 of the listing) in the beginning of file `many_cop.h` should be added: the program uses two random number generators, `simple_*` being somewhat faster, and `*sRandom` routines are used to set starting random seeds at initialisation. The Schrage's and the Bays-Durham algorithms [33] for these

routines can be used. Next four functions provide error processing, messaging and dynamic memory allocation. The last four functions provide read/write routines for the header and polymer configurations.

Let us briefly discuss the structure of the input configuration file. The meaning of the first seven fields in the input is the following: number of monomers in each chain, N ; linear size of lattice, L ; number of polymers on the lattice, M ; starting seed of the random number generators, if zero the current time is used; probability of making reptations for homopolymers, p_r ; probability of moving a chain as a whole; a flag specifying the scheme of how to move clusters of several chains.

Initial conformations of the system may be taken from the `InputBinFile` file, or else the primary sequence and conformation are generated randomly. In the latter case one can set primary sequences in the file `InputTxtFile`, as `ab`, `AB` or `01`, without setting the conformation. If `InputTxtFile` is not specified or does not exist the primary sequences of polymers are generated randomly with the probability of ‘ a ’ monomers given by the field `Probability_of_(a)_Monomers_in_Random_Sequence_When_No_InTxT`. The output system conformations and messages are written to the files `OutputBinFile` and `LogOutFile` respectively.

The next seven lines set the initial two-body interaction constants, $\mathcal{I}_{s_i s_j}$ and weights for nearest, 2D, 3D and second nearest lattice neighbours, $w(r_{ij})$ (see Eq. (3.2)). In fact, program `many_cop` allows one to go along a line in the phase diagram of the system, i.e. to perform Monte Carlo calculations and measurements for several values of interaction parameters which lie on a line. The number of points in the phase diagram, for which measurements are made, is set by `Number_of_Steps_Alone_Interaction_Parameters`. For each value of interaction parameters the system is initially equilibrated for `Number_of_Sweeps_to_Equilibrate` Monte Carlo sweeps (1 Monte Carlo sweep is equal to NM Monte Carlo steps), and then `Number_of_Series` measurements are performed with

`Number_of_Sweeps_in_Seria` Monte Carlo sweeps between them to reduce correlations between system conformations. Here a measurement consists in writing the current system conformation to the output file `OutputBinFile`. After measurements for the current set of interaction parameters have been performed the latter are incremented by values given in the last six lines of the configuration file, `dIaa-dIww`. If one requires different steps in interaction parameters and statistics at different intervals in interactions, e.g. to pass through a phase transition having smaller steps and larger number of sweeps for equilibration of the system, ten last lines may be repeated with somewhat different values of parameters. The procedure of measurements from the initial interactions to the final is repeated `Ensemble_Size` times.

Manipulating with the last ten lines allows one to study kinetics as well as equilibrium phenomena. For example, to study kinetics after an instantaneous quench one should write these lines two times to the input configuration file and set parameters as follows: in the first structure set `Number_of_Series` to 1 and `Iaa-Iww` and `dIaa-dIww` to the initial interactions and the quench value respectively; in the second structure set all `dIaa-dIww` to zero. Then each measurement may be considered as a point in Monte Carlo time with `Number_of_Sweeps_in_Seria` being the time step. The first conformation in the output then corresponds to the conformation exactly before the quench.

```

/* ===== File:      many_cop.cfg      =====
25      Size_of_Each_Chain
16      Linear_Size_of_Lattice
4       Number_of_Chains
0       RandomNumberGeneratorSeed(0-Current_Time)
0.0    Reptation_Probability
0.1    Probability_to_Move_A_Chain
0      cc11(0-prop,1-hydro,2-nohydro)

-      InputBinFile
-      InputTxtFile
many_cop.bin      OutputBinFile
many_cop.LOG      LogOutFile
0.5      Probability_of_(a)_Monomers_in_Random_Sequence_When_No_InTxt

2.1      Iaa
2.2      Iab
2.3      Ibb
1.0      Iaw
1.0      Ibw
0.0      Iww

1.0  1.0  0.7  0.5  Weights_for_Inters_of_NN

1      Ensemble_Size

21      Number_of_Steps_Along_Interaction_Parameters
1000    Number_of_Sweeps_to_Equilibrate
10      Number_of_Series
100     Number_of_Sweeps_in_Seria
0.0    dIaa
0.0    dIab
0.0    dIbb
0.0    dIaw
-0.05  dIbw
0.0    dIww

===== End of File:      many_cop.cfg      ===== */
/* ===== File:      many_cop.h      ===== */

#ifndef  MANY_COP_h
#define  MANY_COP_h

/* --- !!!! These Functions should be provided --- */

extern void  sRandom(rndType);
extern void  simple_sRandom(int);
extern fpType Random(void);
extern fpType simple_Random(void);

extern void  fatal_error(int, char *, ... );
extern void  warning(char *, ... );
extern void  *emalloc(size_t);
extern void  *erealloc(void *, size_t);

```

```

extern int  read_bin_header(FILE *, int *, int *, int *);
extern int  read_one_state(FILE *, plm_site *, int, int, int);
extern int  write_bin_header(FILE *, int, int, int);
extern int  write_one_state(FILE *, plm_site ***, int, int, int);

/* ----- */

/* !!!! Comment this define to allow full cluster moves */
#define REDUCED_CLUSTER_MOVES

#include <stdio.h>
#include <stdlib.h>
#include <math.h>

#define WATER 0
#define MONOMER1 1
#define MONOMER2 2
#define LIMIT_L 5
#define INTR_LEVEL 4
#define INTR_NUM 6
#define efree(a) free(a)

typedef unsigned int LatType;
typedef float fpType;
typedef unsigned int UInt;
typedef unsigned int idxType;
typedef long int rndType;

typedef struct {
    int    x, y, z;
    LatType type;
} plm_site;

typedef struct {
    int    psize;          /* Polymer size          */
    int    lsize;         /* Lattice size          */
    int    pnum;          /* Number of chains      */
    fpType rept;          /* Reptation Probab for Homopols only */
    fpType diffuse;       /* Probability to MoveChain */
    int    cc11;          /* Flag how to move M chains*/
    fpType wgt[INTR_LEVEL]; /* Weights for Interactions */
    fpType sInter[INTR_NUM]; /* Current Interactions  */
    fpType pChi[3][INTR_LEVEL]; /* Initial Values      */
    fpType *cl_prob;      /* Probs of moving clusters */
    FILE  *ifp_bin, *ifp_txt; /* Input FILE structures */
    FILE  *ofp_bin;      /* Output FILE structures */
    fpType prob_ab;      /* Probability of (a) monomers in sequence */
} pnetw_datastruct;

typedef struct {
    int    n_inter;
    int    eq;
    int    n_ser;
    int    sweep;
    fpType dInter[INTR_NUM]; /* shifts of Interactions */
    void  *next;

```

```

} pnetw_runstruct;

/* ----- Function Prototypes ----- */

extern int pnetw_ctor(FILE *);
extern void pnetw_eng_reset(void);
extern void pnetw_pos_reset(void);
extern void pnetw_dtor(void);
extern int pnetw_run(pnetw_runstruct *);

/* ----- External Variables ----- */

extern LatType      ***E_lattice;
extern pnetw_datastruct *E_pdata;
extern plm_site     ***E_link;
extern plm_site     ***E_lnkbuf;
extern plm_site     *E_try;
extern plm_site     *S_sites;
extern int          *E_clusters;
extern int          *E_clus_fls;
extern int          *E_clus_szs;
extern int          *E_homoflag;

#endif /* MANY_COP_h */

/* ===== End of File:    many_cop.h    ===== */

/* ===== File:    many_cop.c    ===== */

#include <many_cop.h>
#include <time.h>

#define TRIESNUM      64
#define DEF_CFG_FILE "many_cop.cfg"

char E_version[] = "0.0.4a";

LatType      ***E_lattice;
pnetw_datastruct *E_pdata = NULL;
plm_site     ***E_link = NULL;
plm_site     ***E_lnkbuf = NULL;
plm_site     *E_try = NULL;
plm_site     *S_sites = NULL;
int          *E_clusters = NULL;
int          *E_clus_fls = NULL;
int          *E_clus_szs = NULL;
int          *E_homoflag = NULL;
int          E_inter_fl;
int          E_nice = -1;
static pnetw_datastruct S_data;
static int      (*S_make_MC_sweep)(void);

static void proc_one(char *);
static int  chk_neib(plm_site *, int);
static void default_init(void);
static void pnetw_rand_init_one_chain(int);
static int  read_monomer_types(FILE *, plm_site *);

```

```

static void default_type_init(plm_site *);
static int  is_homo(int);

extern int  MC_sweep_lhom(void);
extern int  MC_sweep_lcop(void);
extern int  MC_sweep_Many(void);

/* ----- */

int main(int argc, char **argv)
{
    char *cfg_file = DEF_CFG_FILE;

    warning("ManyCop v. %s\t\twritten by Yuri.Kuznetsov@ucd.ie\n", E_version);

    if ( argc < 2 ) { proc_one(cfg_file); }
    else if ( argv[1][0] == '-' && argv[1][1] == 'h' ) {
        warning("Usage: %s [cfg_file1=%s cfg_file2 ...]\n"
              "%s -h      - to see this help\n", *argv, cfg_file, *argv);
        exit(0);
    }
    else {
        if ( argv[1][0] == '-' && argv[1][1] == 'n' ) {
            argv++;
            E_nice = atoi(argv[0]+2);
        }
        for ( argv++; *argv; argv++ ) {
            proc_one(*argv);
        }
    }
    return 0;
}

/* ----- */

static void proc_one(char *cfg_file)
{
    static char iform[] = "%d%s", fform[] = "%e%s";
    pnetw_runstruct tmprun, *prun, *start, *prev;
    FILE *fp = fopen(cfg_file, "r");
    int ens_sz, e, i, j;

    warning("\ncfg_file: %s\n", cfg_file);

    if (!fp) {
        warning("Unable open config file: '%s'; using stdin\n", cfg_file);
        fp = stdin;
    }
    if ( pnetw_ctor(fp) == -1 || fscanf(fp, "%d%s", &ens_sz) == EOF ) {
        warning("inpfile: %s: ctor failure\n", cfg_file);
        return;
    }

    /* Reading run structs from fp */
    for ( prev=NULL; fscanf(fp, iform, &(tmprun.n_inter) ) > 0 &&
          fscanf(fp, iform, &(tmprun.eq) ) > 0 &&
          fscanf(fp, iform, &(tmprun.n_ser) ) > 0 &&

```

```

        fscanf(fp, iform, &(tmprun.sweep) ) > 0 &&
        fscanf(fp, fform, &(tmprun.dInter[0])) > 0 &&
        fscanf(fp, fform, &(tmprun.dInter[1])) > 0 &&
        fscanf(fp, fform, &(tmprun.dInter[2])) > 0 &&
        fscanf(fp, fform, &(tmprun.dInter[3])) > 0 &&
        fscanf(fp, fform, &(tmprun.dInter[4])) > 0 &&
        fscanf(fp, fform, &(tmprun.dInter[5])) > 0; ) {
    pnetw_runstruct *nrun=(pnetw_runstruct*)emalloc(sizeof(pnetw_runstruct))
;
    nrun->n_inter = tmprun.n_inter;
    nrun->eq      = tmprun.eq;
    nrun->n_ser   = tmprun.n_ser;
    nrun->sweep   = tmprun.sweep;
    nrun->dInter[0] = tmprun.dInter[0];
    nrun->dInter[1] = tmprun.dInter[1];
    nrun->dInter[2] = tmprun.dInter[2];
    nrun->dInter[3] = tmprun.dInter[3];
    nrun->dInter[4] = tmprun.dInter[4];
    nrun->dInter[5] = tmprun.dInter[5];
    if (!prev) {
        start = prev = nrun;
    }
    else {
        prev->next = nrun;
        prev = nrun;
    }
}
if (prev) prev->next = NULL;
if (fp != stdin) fclose(fp);

/* Main Loop */
for ( e = 0; e < ens_sz; e++ ) {
    pnetw_pos_reset();
    for ( i=1, prun=start; prun; i++, prun=prun->next ) {
        int steps = pnetw_run(prun);
        warning("File='%s', ens=%d, run=%d -> steps=%d\n",cfg_file,e,i,steps);
    }
}

/* Clearing Stuff */
pnetw_dtor();
for ( prun = start; prun; prun = start ) {
    start = prun->next;
    efree(prun);
}

/* ----- */
int pnetw_ctor(FILE *fp)
{
    static char ib_fname[128], it_fname[128], ob_fname[128], lo_fname[128];
    static char iform[] = "%d%s", sform[] = "%s%s", fform[] = "%e%s";
    time_t rseed, tmm;
    int i, pnum;
    LatType ***lattice;
    FILE *fpt;

```

```

    pnetw_datastruct *pl = &s_data;

    E_pdata = pl;

    /* Reading File */
    if (fscanf(fp, iform, &pl->psize) == EOF ||
        fscanf(fp, iform, &pl->lsize) == EOF ||
        fscanf(fp, iform, &pl->pnum) == EOF ||
        fscanf(fp, iform, &rseed) == EOF ||
        fscanf(fp, fform, &pl->rept) == EOF ||
        fscanf(fp, fform, &pl->diffuse) == EOF ||
        fscanf(fp, iform, &pl->cc11) == EOF ||
        fscanf(fp, sform, ib_fname) == EOF ||
        fscanf(fp, sform, it_fname) == EOF ||
        fscanf(fp, sform, ob_fname) == EOF ||
        fscanf(fp, sform, lo_fname) == EOF ||
        fscanf(fp, fform, &pl->prob_ab) == EOF) goto read_error;
    for ( i = 0; i < INTR_NUM; i++ ) {
        if (fscanf(fp, fform, &(pl->sInter[i]) ) == EOF) goto read_error;
    }
    for ( i = 0; i < INTR_LEVEL; i++ ) {
        if (fscanf(fp, "%e", &(pl->wgt[i]) ) == EOF) goto read_error;
    }
    fscanf(fp, "%s");

    /* Checking Parameters */
    if ( pl->psize <= 0 ) {
        warning("Illegal Polymer Size %d\n", pl->psize);
        return -1;
    }
    if ( pl->lsize <= 0 ) {
        warning("Illegal Lattice Size %d\n", pl->lsize);
        return -1;
    }
    if ( (pnum=pl->pnum) <= 0 ) {
        warning("Illegal Number of Polymers: %d\n", pl->pnum);
        return -1;
    }

    /* Random Generator Settings */
    if (!rseed) {
        time(&rseed);
        warning("Using default seed for random number generator: r=%d\n", rseed);
    }
    sRandom(rseed);
    simple_sRandom(rseed);

    /* Space Lattice Allocation & Initialization */
    lattice = (LatType ***)emalloc(pl->lsize*sizeof(void *));
    for ( i = 0; i < pl->lsize; i++ ) {
        int j;
        lattice[i] = (LatType **)emalloc(pl->lsize*sizeof(void *));
        for ( j = 0; j < pl->lsize; j++ ) {
            int k;
            lattice[i][j] = (LatType *)emalloc(pl->lsize*sizeof(LatType));
            for ( k = 0; k < pl->lsize; k++ )
                lattice[i][j][k] = WATER;

```

```

    }
}
E_lattice = lattice;

/* Opening Files */
if ( (fpt=fopen(ib_fname, "rb")) != NULL ) {
    int psz1, lsz1, pnum1;
    if (read_bin_header(fpt, &psz1, &lsz1, &pnum1) != EOF &&
        psz1 == pl->psize && lsz1 == pl->lsize && pnum1 == pnum) {
        pl->ifp_bin = fpt;
        warning("Using '%s' for initial configurations\n", ib_fname);
    }
    else {
        warning("Illegal bin input; using default\n");
        pl->ifp_bin = NULL;
    }
}
else {
    warning("No bin input; using default\n");
    pl->ifp_bin = NULL;
}
if ( (pl->ifp_txt=fopen(it_fname, "r")) == NULL ) {
    warning("Using random sequences\n");
}
else {
    warning("Using sequences from '%s'\n", it_fname);
}
if ( (fpt=fopen(ob_fname, "rb")) != NULL ) {
    fclose(fpt);
    warning("Binary output file already %s exists. Delete|Move it\
then restart...\n", ob_fname);
    return -1;
}
if ( (fpt = fopen(ob_fname, "wb+")) == NULL ) {
    warning("Cannot open binary output file %s\n", ob_fname);
    return -1;
}
write_bin_header(fpt, pl->psize, pl->lsize, pnum);
pl->ofp_bin = fpt;

/* Reopening stderr */
if (lo_fname[0] != '-' || lo_fname[1] != 0) {
    freopen(lo_fname, "a+", stderr);
}
time(&tmm);
warning("PID=%d\nstart=%u\nhostname=", getpid(), tmm);
system("hostname -s l>&2");

if (E_nice >= 0) {
    static char nice[128];
    sprintf(nice, "/usr/bin/renice %d %d", E_nice, getpid());
    system(nice);
}
E_clusters = (int *)emalloc( 2*(pnum+1) * sizeof(int) );
E_clus_fls = (int *)emalloc( pnum*sizeof(int) );
E_clus_szs = (int *)emalloc( pnum*sizeof(int) );

```

```

for ( i = 0; i < pnum; i++ ) {
    E_clus_szs[i] = 1;
}
pl->cl_prob = (fpType *)emalloc( (pnum+1) * sizeof(fpType) );
pl->cl_prob[0] = 0.0; pl->cl_prob[1] = 1.0;
for ( i = 2; i <= pnum; i++ ) {
    fpType prob = (!pl->cc11) ? 1.0 : ( (pl->cc11 == 1) ?
        pow((double)i, -4.0/3.0) : 1.0/(double)(i*i) );
    pl->cl_prob[i] = prob;
}
E_homoflag = (int *)emalloc( pnum * sizeof(int) );
return 0;

read_error:
warning("Cannot Read Input\n");
return -1;
}

/* ----- */

void pnetw_eng_reset(void)
{
    pnetw_datastruct *pl = E_pdata;
    int i;
    fpType chi[3], *inter = pl->Inter, *wgt = pl->wgt;

    chi[0] = inter[1] - inter[3] - inter[4] + inter[5]; /* chi_AB */
    chi[1] = inter[0] - 2.0 * inter[3] + inter[5]; /* chi_AA */
    chi[2] = inter[2] - 2.0 * inter[4] + inter[5]; /* chi_BB */
    E_inter_fl = (chi[0] == 0.0 && chi[1] == 0.0 && chi[2] == 0.0) ? 0 : 1;

    for ( i = 0; i < INTR_LEVEL; i++ ) {
        pl->pChi[0][i] = wgt[i]*chi[0];
        pl->pChi[1][i] = wgt[i]*chi[1];
        pl->pChi[2][i] = wgt[i]*chi[2];
    }
}

/* ----- */

void pnetw_pos_reset(void)
{
    LatType ***lattice = E_lattice;
    plm_site *sites = S_sites;
    pnetw_datastruct *pl = E_pdata;
    int i, pnum = pl->pnum, psize = pl->psize, totsites = pnum*psize+1;

    for ( i = 0; i < INTR_NUM; i++ ) { /* Restoring initial interaction pars */
        pl->Inter[i] = pl->sInter[i];
    }
    if (!sites) { /* Allocating */
        sites = (plm_site *)emalloc( totsites*sizeof(plm_site) );
        E_link = (plm_site ***)emalloc( pnum*sizeof(plm_site **) );
        E_lnkbuf = (plm_site ***)emalloc( pnum*sizeof(plm_site **) );
        S_sites = sites;
        for ( i = 0; i < pnum; i++ ) {
            E_link[i] = (plm_site **)emalloc( 3*psize*sizeof(plm_site *) );

```

```

        E_lnkbuf[i] = E_link[i];
    }
}
else { /* Clearing Lattice */
    plm_site ***cur = E_link;
    for ( i = 0; i < pnum; i++, cur++ ) {
        int j;
        plm_site **sit = cur[0];
        for ( j = 0; j < psize; j++, sit++ )
            lattice[sit[0]->z][sit[0]->y][sit[0]->x] = WATER;
    }
}

if (pl->ifp_bin) { /* Initialization */
    if (read_one_state(pl->ifp_bin, S_sites, psize, pl->lsize, pnum) == EOF)
    {
        warning("Cannot read next configuration from input; using default\n");
    }
    fclose(pl->ifp_bin);
    pl->ifp_bin = NULL;
    default_init();
}
else {
    for ( i = 0, sites = S_sites; i < totsites; i++, sites++ ) {
        sites->type += (i / psize) << 8;
    }
}
else {
    default_init();
}

for ( i = 0, sites = S_sites; i < pnum; i++ ) {
    int j;
    for ( j = 0; j < psize; j++, sites++ )
        E_link[i][j] = sites;
}
E_try = S_sites+totsites-1;

for ( i = 0; i < pnum; i++ ) {
    E_homoflag[i] = is_homo(i);
}
if (pnum == 1) {
    S_make_MC_sweep = (!E_homoflag[0] || pl->rept <= 0.0) ?
        MC_sweep_lcop : MC_sweep_lhom;
}
else {
    S_make_MC_sweep = MC_sweep_Many;
}
}

/* ----- */

static void default_init(void)
{
    plm_site *sites = S_sites;
    pnetw_datastruct *pl = E_pdata;

```

```

    int i, j;

    for ( i = 0; i < pl->pnum; i++, sites += pl->psize ) {
        pnetw_rand_init_one_chain(i);
    }

    /* ----- */

static void pnetw_rand_init_one_chain(int pno)
{
    pnetw_datastruct *pl = E_pdata;
    LatType ***lattice = E_lattice;
    int lsize = pl->lsize, psize = pl->psize;
    plm_site *bases = S_sites + pno*psize, *curs = bases;
    int i, tries;
    FILE *ifp = pl->ifp_txt;

    /* Setting copolymer site types */
    if (ifp) {
        if (read_monomer_types(ifp, bases) == EOF) {
            warning("Cannot read types; using default\n");
            fclose(ifp);
            pl->ifp_txt = NULL;
            default_type_init(bases);
        }
    }
    else {
        default_type_init(bases);
    }

    /* Setting 3 higher bytes of LatType to pno */
    for ( i = 0, curs = bases, pno *= 256; i < psize; i++, curs++ ) {
        curs->type += pno;
    }

    for ( i = 0, tries = 0, curs = bases; i < psize; ) {
        fpType x = Random(), y = Random(), z = Random();
        if (i) {
            curs->x = (int)(x*(fpType)lsize);
            curs->y = (int)(y*(fpType)lsize);
            curs->z = (int)(z*(fpType)lsize);
        }
        else {
            curs->x = curs[-1].x + (int)(x*3.0) - 1;
            curs->y = curs[-1].y + (int)(y*3.0) - 1;
            curs->z = curs[-1].z + (int)(z*3.0) - 1;
        }
        if (!chk_neib(curs, i)) { /* Success */
            curs++; i++;
            tries = 0;
        }
        else { /* Failure */
            if ( ++tries >= TRIESNUM ) { /* Reset Current sequence */
                int j;
                plm_site *cc = bases;
                for ( j = 0; j < i; j++, cc++)
                    lattice[cc->z][cc->y][cc->x] = WATER;
            }
        }
    }
}

```

```

        i = 0;
        curs = bases;
    }
} /* ?chk_neib */
} /* End of for */
}

/* ----- */

static int chk_neib(plm_site *cur, int flag)
{
    pnetw_datastruct *pl = E_pdata;
    LatType ***lattice = E_lattice;
    register int x = cur->x, y = cur->y, z = cur->z, lsize = pl->lsize;
    int ret;

    if (x < 0 || x >= lsize || y < 0 || y >= lsize ||
        z < 0 || z >= lsize || lattice[z][y][x] != WATER) return -1;

    /* Look up over Nearest Neighbours */
    if (flag) {
        lattice[cur[-1].z][cur[-1].y][cur[-1].x] = WATER;
    }

    if (x < 0 && lattice[z][y][x-1] != WATER) ret = -1;
    else if (x >= lsize && lattice[z][y][x+1] != WATER) ret = -1;
    else if (y < 0 && lattice[z][y-1][x] != WATER) ret = -1;
    else if (y >= lsize && lattice[z][y+1][x] != WATER) ret = -1;
    else if (z < 0 && lattice[z-1][y][x] != WATER) ret = -1;
    else if (z >= lsize && lattice[z+1][y][x] != WATER) ret = -1;
    else ret = 0;

    if (flag) {
        lattice[cur[-1].z][cur[-1].y][cur[-1].x] = cur[-1].type;
    }
    if (!ret)
        lattice[z][y][x] = cur->type;
    return ret;
}

/* ----- */

static int read_monomer_types(FILE *fp, plm_site *sites)
{
    pnetw_datastruct *pl = E_pdata;
    int i, psz = pl->psize;

    for (i = 0; i < psz; i++, sites++) {
        int c;
        while (isspace(c = fgetc(fp))) {}
        if (c == EOF) return EOF;
        sites->type = (c == '0' || c == 'a' || c == 'A') ?
            MONOMER1 : MONOMER2;
    }
    return 0;
}

```

```

/* ----- */

static void default_type_init(plm_site *sites)
{
    pnetw_datastruct *pl = E_pdata;
    int i, psz = pl->psize;
    fpType prob_ab = pl->prob_ab;

    for (i = 0; i < psz; i++, sites++) {
        sites->type = (Random() <= prob_ab) ? MONOMER1 : MONOMER2;
    }
}

/* ----- */

static int is_homo(int pno)
{
    pnetw_datastruct *pl = E_pdata;
    int i, psize = pl->psize;
    plm_site *bases = S_sites + pno*psize, *curs = bases;
    int ret = curs->type;

    for (i = 1, curs++; i < psize; i++, curs++) {
        if (curs->type != ret) { ret = 0; break; }
    }
    return ret;
}

/* ----- */

void pnetw_dtor(void)
{
    pnetw_datastruct *pl = E_pdata;
    int i, j;
    time_t tmm;

    fclose(pl->ofp_bin);
    if (pl->ifp_bin) fclose(pl->ifp_bin);
    if (pl->ifp_txt) fclose(pl->ifp_txt);

    S_make_MC_sweep = NULL;
    efree(pl->cl_prob);
    efree(E_clus_szs);
    efree(E_clus_fls);
    efree(E_clusters);
    efree(E_homoflag);

    for (i = 0; i < pl->pnum; i++)
        efree(E_lnkbuf[i]);
    efree(E_lnkbuf);
    efree(E_link);
    efree(S_sites);

    for (i = 0; i < pl->lsize; i++) {
        for (j = 0; j < pl->lsize; j++)
            efree(E_lattice[i][j]);
        efree(E_lattice[i]);
    }
}

```

```

}
efree(E_lattice);

E_pdata = NULL;
E_lattice = NULL;
E_link = NULL;
E_lnkbuf = NULL;
E_try = NULL;
S_sites = NULL;

time(&tmm);
warning("finish=%u\n", tmm);
}

/* ----- */

int pnetw_run(pnetw_runstruct *prun)
{
    pnetw_datastruct *pl = E_pdata;
    int n_int=0, steps=0, sweep, ser, i;

    for ( n_int = 0; n_int < prun->n_inter; n_int++ ) {
        pnetw_eng_reset();
        for ( sweep = 0; sweep < prun->eq; sweep++ ) {
            steps += S_make_MC_sweep();
        }
        for ( ser = 0; ser < prun->n_ser; ser++ ) {
            for ( sweep = 0; sweep < prun->sweep; sweep++ ) {
                steps += S_make_MC_sweep();
            }
            if ( write_one_state(pl->ofp_bin, E_link, pl->psize,
                               pl->lsize, pl->pnum) == EOF ) {
                warning("Cannot write binary configuration\n");
            }
            else {
                fflush(pl->ofp_bin);
            }
        }
        for ( i = 0; i < INTR_NUM; i++ ) {
            pl->Inter[i] += prun->dInter[i];
        }
    }
    return steps;
}

/* ===== End of File:    many_cop.c    ===== */
/* ===== File:    many_cop_move.c    ===== */

#include <many_cop.h>

extern int E_inter_fl;

typedef struct {
    int plevel2[3*INTR_LEVEL];
} neib_pop;

```

```

int MC_sweep_lcop(void);
int MC_sweep_lhom(void);
int MC_sweep_Many(void);

static int make_pmove(int, int, plm_site **);
static int make_diffuse(int);
static int make_prept(int, int);

static plm_site *inside_try(plm_site **, plm_site *);
static plm_site *bound_try(plm_site *, plm_site *, plm_site *);
static void get_localexenergy(plm_site *, neib_pop *);
static fpType clc_prob(neib_pop *, neib_pop *);
static void is_other_chain(plm_site *, int, int *);

/* ----- Makes Local Moves only ----- */

int MC_sweep_lcop(void)
{
    pnetw_datastruct *ps = E_pdata;
    int psize = ps->psize, i, ret = 0;

    for ( i = 0; i < psize; i++ ) {
        int mno = (int)((fpType)simple_Random()*(fpType)psize);
        int neib = (mno <= 0) ? 1 : ( (mno >= psize-1) ? psize-2 : -1 );
        int rc = make_pmove(mno, neib, E_link[0]);
        ret += rc;
    }
    return ret;
}

/* ----- Makes Local and Reptational Moves ----- */

int MC_sweep_lhom(void)
{
    pnetw_datastruct *ps = E_pdata;
    int psize = ps->psize, i, ret = 0;
    fpType rept = ps->rept;

    for ( i = 0; i < psize; i++ ) {
        fpType rnd = simple_Random();
        int rc;

        if ( rnd < rept ) {
            int t2h_flag = (rnd < 0.5*rept) ? 1 : 0;
            rc = make_prept(0, t2h_flag);
        }
        else {
            int mno, neib;
            mno = (int)((fpType)psize * (rnd-rept) / (1.0-rept));
            neib = (mno <= 0) ? 1 : ( (mno >= psize-1) ? psize-2 : -1 );
            rc = make_pmove(mno, neib, E_link[0]);
        }
        ret += rc;
    }
    return ret;
}

```

```

}
/* ----- */
int MC_sweep_Many(void)
{
  pnetw_datastruct *ps = E_pdata;
  int pnum = ps->pnum, psize = ps->psize,
  tot = pnum*psize, i, ret = 0;
  fpType rept = ps->rept, clus = ps->diffuse;

  for ( i = 0; i < tot; i++ ) {
    int pno = (int)((fpType)simple_Random()*(fpType)pnum), rc;
    fpType rnd = simple_Random();
    fpType rept_1 = ((!E_homoflag[pno]) ? 0.0 : rept), r_p_cl = rept_1 + clus;

    if (rnd < rept_1) {
      int t2h_flag = (rnd < 0.5*rept) ? 1 : 0;
      rc = make_prept(0, t2h_flag);
    }
    else if (rnd < r_p_cl) {
      rc = make_diffuse(pno);
    }
    else {
      int mno = (int)((fpType)psize * (rnd-r_p_cl) / (1.0-r_p_cl));
      int neib = (mno <= 0) ? 1 : (mno >= psize-1) ? psize-2 : -1;
      rc = make_pmove(mno, neib, E_link[pno]);
    }
    ret += rc;
  }
  return ret;
}
/* ----- */

static int make_pmove(int site, int neib, plm_site **pplink)
{
  static neib_pop num_init, num_end;
  register plm_site *old_pos = pplink[site];
  register plm_site *new_pos;
  register int ox = old_pos->x, oy = old_pos->y, oz = old_pos->z;
  int cur_type = old_pos->type;
  fpType probab;

  /* Attempt to Move */
  E_lattice[oz][oy][ox] = WATER;
  new_pos = (neib == -1) ? inside_try(pplink+site, E_try)
  : bound_try(pplink[site], pplink[neib], E_try);
  if (!new_pos || (new_pos->x==ox && new_pos->y==oy && new_pos->z==oz)) {
    E_lattice[oz][oy][ox] = cur_type;
    return 0;
  }

  /* Energy Approx */
  if (E_inter_fl) {
    get_localexenergy(old_pos, &num_init);
  }
}

```

```

    get_localexenergy(new_pos, &num_end);
    probab = clc_prob(&num_init, &num_end);
  }
  else
    probab = 1.0;

  if (probab >= 1.0 || probab >= Random()) { /* Accept */
    E_try = old_pos;
    pplink[site] = new_pos;
    E_lattice[new_pos->z][new_pos->y][new_pos->x] = cur_type;
    return 1;
  }
  else {
    E_lattice[oz][oy][ox] = cur_type;
    return 0;
  }
}
/* ----- */

#define MK_INSIDE(cur, prev, next) \
switch (prev-next) { \
  case 1: cur = next + (int)(simple_Random()*2.0); break; \
  case -1: cur = prev + (int)(simple_Random()*2.0); break; \
  case 2: cur = next + 1; break; \
  case -2: cur = prev + 1; break; \
  default: if ((cur=next+(int)(simple_Random()*3.0)-1) < 0 || cur>=lsize)\
    cur=next; break; \
}

#define MK_BOUND(cur, next) \
if ((cur=next+(int)(simple_Random()*3.0)-1) < 0 || cur >= lsize)\
  cur = next;

/* ----- */

static plm_site *inside_try(plm_site **pcur, plm_site *try)
{
  register plm_site *cur = pcur[0], *pr = pcur[-1], *ne = pcur[1];
  register int x, y, z, lsize = E_pdata->lsize, plsz = lsize-1;
  register int prevx = pr->x, prevy = pr->y, prevz = pr->z;
  register int nextx = ne->x, nexty = ne->y, nextz = ne->z;
  register LatType ***lattice = E_lattice, c;
  register int pcnt;
  int limit;

  for ( limit = LIMIT_L; ; ) {
    pcnt = 0;
    if ( --limit < 0 ) return NULL;
    MK_INSIDE(x, prevx, nextx);
    MK_INSIDE(y, prevy, nexty);
    MK_INSIDE(z, prevz, nextz);
    if (lattice[z][y][x] != WATER) continue;
    lattice[prevz][prevy][prevx] = WATER;
    lattice[nextz][nexty][nextx] = WATER;
    if ( x > 0 && lattice[z][y][x-1] != WATER ) pcnt++;
    else if ( x < plsz && lattice[z][y][x+1] != WATER ) pcnt++;
  }
}

```

```

else if ( y > 0    && lattice[z][y-1][x] != WATER ) pcnt++;
else if ( y < plsz && lattice[z][y+1][x] != WATER ) pcnt++;
else if ( z > 0    && lattice[z-1][y][x] != WATER ) pcnt++;
else if ( z < plsz && lattice[z+1][y][x] != WATER ) pcnt++;
lattice[prevz][prevy][prevx] = pr->type;
lattice[nextz][nexty][nextx] = ne->type;
if (!pcnt) break;
}

try->x = x;
try->y = y;
try->z = z;
try->type = cur->type;
return try;
}

/* ----- */

static plm_site *bound_try(plm_site *cur, plm_site *next, plm_site *try)
{
register int x, y, z, lsize = E_pdata->lsize, plsz = lsize-1;
register int limit, pcnt;
register LatType ***lattice = E_lattice, c;

for ( limit = LIMIT_L+1; ; ) {
pcnt = 0;
if ( --limit == 0 ) return NULL;
MK_BOUND(x,next->x);
MK_BOUND(y,next->y);
MK_BOUND(z,next->z);
if (lattice[z][y][x] != WATER) continue;
lattice[next->z][next->y][next->x] = WATER;
if ( x > 0    && lattice[z][y][x-1] != WATER ) pcnt++;
else if ( x < plsz && lattice[z][y][x+1] != WATER ) pcnt++;
else if ( y > 0    && lattice[z][y-1][x] != WATER ) pcnt++;
else if ( y < plsz && lattice[z][y+1][x] != WATER ) pcnt++;
else if ( z > 0    && lattice[z-1][y][x] != WATER ) pcnt++;
else if ( z < plsz && lattice[z+1][y][x] != WATER ) pcnt++;
lattice[next->z][next->y][next->x] = next->type;
if (!pcnt) break;
}

try->x = x;
try->y = y;
try->z = z;
try->type = cur->type;
return try;
}

/* ----- */

static void get_localexpenegy(plm_site *cur, neib_pop *populat)
{
pnetw_datastruct *pl = E_pdata;
int lsz = E_pdata->lsize;
register int type = cur->type & 255;
register int x = cur->x, y = cur->y, z = cur->z;

```

```

register int ufl, dfl, nfl, sfl, wfl, efl;
register LatType ***lattice = E_lattice;
register LatType c;
register int i;
register int *pnum = populat->plevel2, *pp=pnum;

#define CHK_DIR(cond, x, y, z) \
if ( cond && ((c=(lattice[z][y][x]&255))==MONOMER1 || c==MONOMER2) \
pnum[(c == type) ? type : 0]++;

for ( i = 0; i < 3*INTR_LEVEL; i++, pp++ ) { *pp = 0; }

efl = (x==lsz-1) ? 0 : 1;
wfl = (x==0) ? 0 : 1;
sfl = (y==lsz-1) ? 0 : 1;
nfl = (y==0) ? 0 : 1;
ufl = (z==lsz-1) ? 0 : 1;
dfl = (z==0) ? 0 : 1;

CHK_DIR(efl, x+1, y, z);
CHK_DIR(wfl, x-1, y, z);
CHK_DIR(sfl, x, y+1, z);
CHK_DIR(nfl, x, y-1, z);
CHK_DIR(ufl, x, y, z+1);
CHK_DIR(dfl, x, y, z-1);

pnum += 3;
CHK_DIR(efl&&sfl, x+1, y+1, z);
CHK_DIR(efl&&nfl, x+1, y-1, z);
CHK_DIR(wfl&&sfl, x-1, y+1, z);
CHK_DIR(wfl&&nfl, x-1, y-1, z);
CHK_DIR(efl&&ufl, x+1, y, z+1);
CHK_DIR(efl&&dfl, x+1, y, z-1);
CHK_DIR(wfl&&ufl, x-1, y, z+1);
CHK_DIR(wfl&&dfl, x-1, y, z-1);
CHK_DIR(sfl&&ufl, x, y+1, z+1);
CHK_DIR(sfl&&dfl, x, y+1, z-1);
CHK_DIR(nfl&&ufl, x, y-1, z+1);
CHK_DIR(nfl&&dfl, x, y-1, z-1);

pnum += 3;
CHK_DIR(efl&&sfl&&ufl, x+1, y+1, z+1);
CHK_DIR(efl&&sfl&&dfl, x+1, y+1, z-1);
CHK_DIR(efl&&nfl&&ufl, x+1, y-1, z+1);
CHK_DIR(efl&&nfl&&dfl, x+1, y-1, z-1);
CHK_DIR(wfl&&sfl&&ufl, x-1, y+1, z+1);
CHK_DIR(wfl&&sfl&&dfl, x-1, y+1, z-1);
CHK_DIR(wfl&&nfl&&ufl, x-1, y-1, z+1);
CHK_DIR(wfl&&nfl&&dfl, x-1, y-1, z-1);

efl = (x>=lsz-2) ? 0 : 1;
wfl = (x<=1) ? 0 : 1;
sfl = (y>=lsz-2) ? 0 : 1;
nfl = (y<=1) ? 0 : 1;
ufl = (z>=lsz-2) ? 0 : 1;
dfl = (z<=1) ? 0 : 1;

```

```

pnum += 3;
CHK_DIR(efl, x+2, y, z);
CHK_DIR(wfl, x-2, y, z);
CHK_DIR(sfl, x, y+2, z);
CHK_DIR(nfl, x, y-2, z);
CHK_DIR(uf1, x, y, z+2);
CHK_DIR(df1, x, y, z-2);
}

/* ----- */

static fpType clc_prob(neib_pop *ipopul, neib_pop *epopul)
{
register int il;
register fpType ret = 0.0;
register pnetw_datastruct *pl = E_pdata;
register int *init = ipopul->plevel2, *end = epopul->plevel2;

for ( il = 0; il < INTR_LEVEL; il++, init += 3, end += 3 ) {
ret += pl->pChi[0][il] * (end[0]-init[0]) +
pl->pChi[1][il] * (end[1]-init[1]) +
pl->pChi[2][il] * (end[2]-init[2]);
}

return ( ret >= 0.0 ) ? 1.0 : exp(ret);
}

static int S_last_cluster;

/* ----- */
/* Move polymer as a whole along any of x,y,z axis
if it is noninteracting with any other chain;
for that one need to check the interaction volume
as for monomers from other chains. Absence of them
guarantee the excluded volume */

static int make_diffuse(int pno)
{
pnetw_datastruct *pl = E_pdata;
plm_site **bases;
register int psize = pl->psize, pnum = pl->pnum, dir, i, tot;
register int dx = 0, dy = 0, dz = 0;
register LatType **lattice = E_lattice;
int avail_dirs[6] = { 1, 1, 1, 1, 1, 1 }; /* All avail by default */
int *clusters = E_clusters, *clus_fls = E_clus_fls,
cur_cl, cl_size, cur_pno;
fpType prob;

/* Move with appropriate probability; probability dependence on
number of chains is in pl->cl_prob[M]; number of chains in
cluster in which pno is involved is taken from previous
steps and stored in E_clus_szs[pno]; less precise scheme,
but much more efficient in the case of nohydrodynamics */

cl_size = E_clus_szs[pno];
prob = pl->cl_prob[cl_size];
if ( prob < 1.0 && prob < Random() ) {

```

```

return 0;
}

/* Initialization of cluster arrays */
for ( i = 0, clusters[0] = pno, S_last_cluster = 1; i < pnum; ) {
clus_fls[i] = (i == pno) ? 1 : 0;
i++;
clusters[i] = -1;
}

/* Get cluster of chains; E_clusters array contains finally numbers
of all chains participating in current cluster starting from pno;
E_clus_fls are flags for all chains, are they in this cluster or not */

for ( cur_cl = 0; (cur_pno=clusters[cur_cl]) != -1; cur_cl++ ) {
bases = E_link[cur_pno];
for ( i = 0; i < psize; i++ ) {
is_other_chain(bases[i], cur_pno, avail_dirs);
}
}

/* Storing the last information about cluster size in E_clus_szs;
will be used later at estimation of the subsequent move probability */

cl_size = cur_cl;
for ( cur_cl = 0; (cur_pno=clusters[cur_cl]) != -1; cur_cl++ ) {
E_clus_szs[cur_pno] = cl_size;
}

/* Skipping unavailable directions */
for ( i = 0, tot = 0; i < 6; i++ ) { tot += avail_dirs[i]; }
if (!tot) { return 0; }
dir = (int)( (fpType)tot * simple_Random() );

for ( i = 0; i < 6; i++ ) {
if (!avail_dirs[i]) continue;
if (!dir) break;
else dir--;
}

dir = i;
if (dir == 0) dx = -1;
else if (dir == 1) dx = 1;
else if (dir == 2) dy = -1;
else if (dir == 3) dy = 1;
else if (dir == 4) dz = -1;
else if (dir == 5) dz = 1;

/* Clearing Lattice and reseting plm_sites */

for ( cur_cl = 0; (cur_pno=clusters[cur_cl]) != -1; cur_cl++ ) {
bases = E_link[cur_pno];
for ( i = 0; i < psize; i++ ) {
register int *x = &(bases[i]->x), *y = &(bases[i]->y),
*z = &(bases[i]->z);
lattice[*z][*y][*x] = WATER;
*x += dx; *y += dy; *z += dz;
}
}

```

```

    }

/* Updating Lattice */
for ( cur_cl = 0; (cur_pno=clusters[cur_cl]) != -1; cur_cl++ ) {
    bases = E_link[cur_pno];
    for ( i = 0; i < psize; i++ ) {
        register int x = bases[i]->x, y = bases[i]->y, z = bases[i]->z,
                    t = bases[i]->type;
        lattice[z][y][x] = t;
    }
}

#ifdef REDUCED_CLUSTER_MOVES

if (cl_size < pnum) {
    E_clusters += pnum+1;
    clusters = E_clusters;
    for ( i = 0, clusters[0] = pno, S_last_cluster = 1; i < pnum; ) {
        clus_fls[i] = (i == pno) ? 1 : 0;
        i++;
        clusters[i] = -1;
    }
    /* Get cluster of chains */
    for ( cur_cl = 0; (cur_pno=clusters[cur_cl]) != -1; cur_cl++ ) {
        bases = E_link[cur_pno];
        for ( i = 0; i < psize; i++ ) {
            is_other_chain(bases[i], cur_pno, avail_dirs);
        }
        E_clusters -= pnum+1;
        clusters = E_clusters;
    }
    if (cur_cl > cl_size) { /* get to another cluster; go back */
        for ( cur_cl = 0; (cur_pno=clusters[cur_cl]) != -1; cur_cl++ ) {
            bases = E_link[cur_pno];
            for ( i = 0; i < psize; i++ ) {
                register int *x = &(bases[i]->x), *y = &(bases[i]->y),
                            *z = &(bases[i]->z);
                lattice[*z][*y][*x] = WATER;
                *x -= dx; *y -= dy; *z -= dz;
            }
        }
        for ( cur_cl = 0; (cur_pno=clusters[cur_cl]) != -1; cur_cl++ ) {
            bases = E_link[cur_pno];
            for ( i = 0; i < psize; i++ ) {
                register int x = bases[i]->x, y = bases[i]->y,
                            z = bases[i]->z, t = bases[i]->type;
                lattice[z][y][x] = t;
            }
        }
    }
}

#endif /* REDUCED_CLUSTER_MOVES */

return 1;
}

```

```

/* ----- */

static void is_other_chain(plm_site *cur, int pno, int *avail_dirs)
{
    pnetw_datastruct *pl = E_pdata;
    int lsz = E_pdata->lsz;
    register int x = cur->x, y = cur->y, z = cur->z;
    register int ufl, dfl, nfl, sfl, wfl, efl;
    register LatType ***lattice = E_lattice;
    register LatType c;
    register int cpno;
    int *clusters = E_clusters, *clus_fls = E_clus_fls;

#define CHK_DIR_DIFF(cond, x, y, z) \
    if ( cond && (c=lattice[z][y][x]) != WATER && \
        (cpno=(c>>8)) != pno && !clus_fls[cpno] ) { \
        clus_fls[cpno] = 1; \
        clusters[S_last_cluster++] = cpno; \
    }

    if (x == 0) { wfl = avail_dirs[0] = 0; } else { wfl = 1; }
    if (x == lsz-1) { efl = avail_dirs[1] = 0; } else { efl = 1; }
    if (y == 0) { nfl = avail_dirs[2] = 0; } else { nfl = 1; }
    if (y == lsz-1) { sfl = avail_dirs[3] = 0; } else { sfl = 1; }
    if (z == 0) { dfl = avail_dirs[4] = 0; } else { dfl = 1; }
    if (z == lsz-1) { ufl = avail_dirs[5] = 0; } else { ufl = 1; }

    CHK_DIR_DIFF(efl&&sfl, x+1, y+1, z);
    CHK_DIR_DIFF(efl&&nfl, x+1, y-1, z);
    CHK_DIR_DIFF(wfl&&sfl, x-1, y+1, z);
    CHK_DIR_DIFF(wfl&&nfl, x-1, y-1, z);
    CHK_DIR_DIFF(efl&&ufl, x+1, y, z+1);
    CHK_DIR_DIFF(efl&&dfl, x+1, y, z-1);
    CHK_DIR_DIFF(wfl&&ufl, x-1, y, z+1);
    CHK_DIR_DIFF(wfl&&dfl, x-1, y, z-1);
    CHK_DIR_DIFF(sfl&&ufl, x, y+1, z+1);
    CHK_DIR_DIFF(sfl&&dfl, x, y+1, z-1);
    CHK_DIR_DIFF(nfl&&ufl, x, y-1, z+1);
    CHK_DIR_DIFF(nfl&&dfl, x, y-1, z-1);

    CHK_DIR_DIFF(efl&&sfl&&ufl, x+1, y+1, z+1);
    CHK_DIR_DIFF(efl&&sfl&&dfl, x+1, y+1, z-1);
    CHK_DIR_DIFF(efl&&nfl&&ufl, x+1, y-1, z+1);
    CHK_DIR_DIFF(efl&&nfl&&dfl, x+1, y-1, z-1);
    CHK_DIR_DIFF(wfl&&sfl&&ufl, x-1, y+1, z+1);
    CHK_DIR_DIFF(wfl&&sfl&&dfl, x-1, y+1, z-1);
    CHK_DIR_DIFF(wfl&&nfl&&ufl, x-1, y-1, z+1);
    CHK_DIR_DIFF(wfl&&nfl&&dfl, x-1, y-1, z-1);

    efl = (x>=lsz-2) ? 0 : 1;
    wfl = (x<=1) ? 0 : 1;
    sfl = (y>=lsz-2) ? 0 : 1;
    nfl = (y<=1) ? 0 : 1;
    ufl = (z>=lsz-2) ? 0 : 1;
    dfl = (z<=1) ? 0 : 1;
}

```

```

CHK_DIR_DIFF(efl, x+2, y, z);
CHK_DIR_DIFF(wfl, x-2, y, z);
CHK_DIR_DIFF(sfl, x, y+2, z);
CHK_DIR_DIFF(nfl, x, y-2, z);
CHK_DIR_DIFF(uf1, x, y, z+2);
CHK_DIR_DIFF(df1, x, y, z-2);
}

/* ----- */
/* reptation of pno chain tail to head or head to tail */

static int make_prept(int pno, int tail2head)
{
    static neib_pop num_init, num_end;
    pnetw_datastruct *pl = E_pdata;
    plm_site **pplink = E_link[pno], **ppbuf = E_lnkbuf[pno];
    register plm_site *old_pos, *new_pos;
    register int ox, oy, oz;
    register LatType ***lattice = E_lattice;
    int cur_type;
    fpType probab;
    int last = pl->psize - 1;

    if (tail2head) { /* Tail to Head */
        old_pos = pplink[last];
        ox = old_pos->x;
        oy = old_pos->y;
        oz = old_pos->z;
        cur_type = old_pos->type;

        lattice[oz][oy][ox] = WATER;
        new_pos = bound_try(old_pos, pplink[0], E_try);
        if (!new_pos || (new_pos->x==ox && new_pos->y==oy && new_pos->z==oz)) {
            lattice[oz][oy][ox] = cur_type;
            return 0;
        }
        /* Energy Approx */
        if (E_inter_fl) {
            get_localexpenenergy(old_pos, &num_init);
            get_localexpenenergy(new_pos, &num_end);
            probab = clc_prob(&num_init, &num_end);
        }
        else {
            probab = 1.0;
        }

        if (probab >= 1.0 || probab >= Random()) { /* Accept */
            if (pplink <= ppbuf) {
                register int i;
                register plm_site **from = pplink, **to = ppbuf+pl->psize+1;
                for ( i = last; i > 0; i--, to++, from++ ) { *to = *from; }
                E_link[pno] = E_lnkbuf[pno] + pl->psize;
            }
            else E_link[pno]--;
            E_link[pno][0] = new_pos;
            E_try = old_pos;
            E_lattice[new_pos->z][new_pos->y][new_pos->x] = cur_type;

```

```

        return 1;
    }
    else {
        E_lattice[oz][oy][ox] = cur_type;
        return 0;
    }
}
else { /* Head to Tail */
    old_pos = pplink[0];
    ox = old_pos->x;
    oy = old_pos->y;
    oz = old_pos->z;
    cur_type = old_pos->type;

    lattice[oz][oy][ox] = WATER;
    new_pos = bound_try(old_pos, pplink[last], E_try);
    if (!new_pos || (new_pos->x==ox && new_pos->y==oy && new_pos->z==oz)) {
        lattice[oz][oy][ox] = cur_type;
        return 0;
    }
    /* Energy Approx */
    if (E_inter_fl) {
        get_localexpenenergy(old_pos, &num_init);
        get_localexpenenergy(new_pos, &num_end);
        probab = clc_prob(&num_init, &num_end);
    }
    else {
        probab = 1.0;
    }

    if (probab >= 1.0 || probab >= Random()) { /* Accept */
        if (pplink >= ppbuf+2*pl->psize) {
            register int i;
            register plm_site **from = pplink+1, **to = ppbuf+pl->psize;
            for ( i = last; i > 0; i--, to++, from++ ) *to = *from;
            E_link[pno] = E_lnkbuf[pno] + pl->psize;
        }
        else E_link[pno]++;
        E_link[pno][last] = new_pos;
        E_try = old_pos;
        lattice[new_pos->z][new_pos->y][new_pos->x] = cur_type;
        return 1;
    }
    else {
        lattice[oz][oy][ox] = cur_type;
        return 0;
    }
}
}

/* ===== End of File: many_cop_move.c ===== */

```

Appendix B

In this appendix we present listing of the C code for the program called `gsck_plarb` (Gaussian Self-Consistent equations for Kinetics of PLural ARbitrary heteropolymers) version 0.0.1alpha, which performs numerical solution of the GSC equations (2.73) in kinetics using the fifth order Runge-Kutta algorithm with adaptive time step for several heteropolymer chains of arbitrary length and composition and either ring or open topology.

The distribution consists of five files:

| | |
|------------------------------|---|
| <code>gsck_plarb.cfg</code> | An example of the input configuration file |
| <code>gsck_plarb.h</code> | Header file with data and function prototypes |
| <code>gsck_plarb.c</code> | C code of <code>main</code> and few other functions |
| <code>gsck_plarb_f.c</code> | C code of general interface, ODE integrator and initialisation and deinitialisation routines |
| <code>gsck_plarb_vx.c</code> | C code of routines performing calculation of the effective potentials described in Sec. 2.5 |

C code files can be compiled in the same way as described in Appendix A. Four more functions prototyped at the beginning of the header file, which perform error processing, messaging and dynamic memory allocation, should be provided to build the program.

Let us describe the structure of the input configuration file. First ten lines set up the following parameters of the effective free energy functional after the

quench respectively: number of polymer chains, M ; temperature, $k_B T$; soft spatial cut-off in the system, L ; elasticity of springs between monomers, $k_B T/l^2$; bare friction per monomer, ζ_b ; viscosity of the solvent, η_s ; mean second virial coefficient, $\bar{u}^{(2)}$; third virial coefficient, $u^{(3)}$; coefficient in the self-interaction energy term, c_3 ; amphiphilicity Δ . Setting viscosity $\eta_s \leq 0$ forces the program to neglect the hydrodynamic interaction. The input consists of two files: polymer primary sequences $\{\sigma_A\}$ and the system conformation before the quench are taken from files `Wgt_FileName` and `Inp_FileName` respectively. System evolution during kinetics, $D_{AA'}(t)$ and few other observables, are written to file `Out_FileName` and messages are written to `log_FileName`. The next line specifies the lengths and topologies of each of M polymers. After the program has been started it integrates the GSC equations and writes `Number_of_Measurements` conformations to output with the timestep approximately equal to `Minimal_Time_between_Measurements`. Two other values in the last four lines set the initial timestep and the relative error between estimates from the fourth and fifth order Runge-Kutta schemes, ϵ_0 (see Sec. 2.5). If one requires different precision or frequency of measurements at different time intervals in kinetics four last lines may be repeated with somewhat different values of parameters.

```

/* ===== File:      gsck_plarb.cfg      =====
5          Number_of_Chains
1.0        Temperature
5.0e+01    LinearCutOff
1.0        Spring_Const
1.0        Friction
0.0        Viscosity
-25.0      Aver_U2_Value
10.0       U3_Value
1.0        C3_Value_in_Self_Inter
0.0        Amphiphilicity_Delta
0.0        Initial_Time

0          Input_File_Mode(1/0-Binary/Text)
wgt.cfg    Wgt_FileName
ini.txt    Inp_FileName
out_030.bin Out_FileName
log        Log_FileName

10o 10r 10o 10r 10o    Lengthes_&_Topologies_Of_Chains(o-open,r-ring)

0.01       Initial_TimeStep
1.0e-06    Relative_Error
100        Number_of_Measurements
0.1        Minimal_Time_between_Measurements

===== End of File:      gsck_plarb.cfg      ===== */
/* ===== File:      gsck_plarb.h      ===== */
/*          Gsc Method for Kinetic of a Few Arbitrary CoPolymers          */

#ifndef GSKC_PLARB_H
#define GSKC_PLARB_H

#include <stdio.h>
#include <stdlib.h>
#include <math.h>

/* --- !!!! These Functions should be provided --- */

extern void fatal_error(int, char *, ... );
extern void warning(char *, ... );
extern void *emalloc(size_t);
extern void *erealloc(void *, size_t);

/* ----- */

#define BIG_VAL    1.0e15
#define SMALL_VAL  1.0e-15
#define efree(a)   free(a)

#define OPEN_TOP  'o'
#define RING_TOP  'r'

typedef float fpType;

```

```

typedef unsigned int UInt;
typedef unsigned int idxType;
typedef long int rndType;

typedef struct {
    double ini_dt;
    double error;
    int tnum;
    double mint;
    void *next;
} gsck_run;

typedef struct {
    int top, psz;
} plm_desc;

typedef struct {
    int pnum, psize;
    plm_desc *poly;
    double temper;
    double lco;
    double spring;
    double frict;
    double visc;
    double u2;
    double u3;
    double C3;
    double disp; /* Coefficient in u^{(2)}_{mm}' */
    int mode; /* HplmText=0, binary=1, GenText=2 */
    char *inp_fname;
    char *out_fname;
    char *wgt_fname;
    char *log_fname;
    double time0;
    gsck_run *pkin;
} gsck_inparam;

typedef struct {
    double **Dmn;
} RK_type; /* Vars integrated by RK scheme */

typedef struct {
    double t;
    double rg2;
    double mps;
    double eng;
    double ent;
    double cnt[3];
    double *Rga;
    double *Mpa;
} observ;

typedef struct {
    int pnum, psize;
    plm_desc *poly;
    double temp;
    double lco;

```

```

double    sph;
int       hd_flag;
double    m1, v1;
double    **pp_u2h;
double    u3h, C3;
double    curtime;
RK_type   *p_Xq;
RK_type   *p_Xqn;
RK_type   *p_deriv[7];
RK_type   *p_cerr;
double    ***ppp_l;
double    **pp_D52;
double    **pp_D40;
double    **pp_V1;
double    **pp_V2;
double    **pp_V3;
double    **pp_Vsi;
double    **pp_Val;
double    **pp_Wmn;
double    **pp_Mob;
FILE      *cfp;
observ    *p_obs;
double    **pp_ent, *p_Rgm, Rg2;
} gsck_workval;

/* File: gsck_gen.c */
extern int  gsck_read_param_file(char *, gsck_inparam *);

/* File: gsck_gen_f.c */
extern int  gsck_ctor(gsck_inparam *);
extern int  gsck_process(gsck_run *);
extern void gsck_print(void);
extern void gsck_dtor(void);

/* File: gsck_gen_vx.c */
extern void gsck_clc_spring(void);
extern int  gsck_clc_poten(RK_type *, double);
extern void gsck_clc_observ(void);

extern gsck_workval *E_workvals;

#endif /* GSCK_PLARB_H */

/* ===== End of File:   gsck_plarb.h   ===== */
/* ===== File:         gsck_plarb.c   ===== */

#include <gsck_plarb.h>

#define DEF_CFG_FILENAME  "gsck_plarb.cfg"

void proc_one_cfg_file(char *);

/* ----- */

int main(int argc, char **argv)
{

```

```

char *in_fname = DEF_CFG_FILENAME;

if (argc < 2) {
    proc_one_cfg_file(in_fname);
}
else if (argv[1][0] == '-' && argv[1][1] == 'h') {
    fprintf(stderr, "Usage: %s [cfgfile1=%s cfgfile2 ...]\n",
        *argv, in_fname);
}
else {
    for ( argc--, argv++; argc; argc--, argv++ )
        proc_one_cfg_file(*argv);
}
return 0;
}

/* ----- */

void proc_one_cfg_file(char *fname)
{
    gsck_inparam inpars;
    gsck_run *prun;
    int i;

    if ( gsck_read_param_file(fname, &inpars) == -1 )
        return;
    if ( gsck_ctor(&inpars) == -1 ) {
        warning("inpfile: %s: ctor failure\n", fname);
        return;
    }

    for ( i = 1, prun = inpars.pkin; prun; i++, prun = prun->next ) {
        int steps = gsck_process(prun);
        if ( steps >= 0 ) {
            warning("File: %s, run: %d -> steps proceed: %d\n", fname, i, steps);
        }
        else {
            warning("Cannot proceed configuration in %s\n", fname);
            break;
        }
    }
    gsck_dtor();
}

/* ----- */

int gsck_read_param_file(char *fname, gsck_inparam *pinpar)
{
    FILE *fp;
    char *p = (!fname) ? "" : fname;
    static char inp_fname[32], out_fname[32], wgt_fname[32], log_fname[32];
    gsck_run tmprun, *prev;
    plm_desc *pp;
    int i, size, tot;
    char top;

    pinpar->inp_fname = inp_fname;

```

```

pinpar->out_fname = out_fname;
pinpar->wgt_fname = wgt_fname;
pinpar->log_fname = log_fname;

if ( !fname || (fp=fopen(fname,"r") == NULL ) {
    fprintf(stderr, "Cannot open input file: %s. Using stdin\n", p);
    fp = stdin;
}
if ( fscanf(fp, "%d%s", &(pinpar->pnum )) == EOF ||
    fscanf(fp, "%lf%s", &(pinpar->temper)) == EOF ||
    fscanf(fp, "%lf%s", &(pinpar->Lco )) == EOF ||
    fscanf(fp, "%lf%s", &(pinpar->spring)) == EOF ||
    fscanf(fp, "%lf%s", &(pinpar->frict )) == EOF ||
    fscanf(fp, "%lf%s", &(pinpar->visc )) == EOF ||
    fscanf(fp, "%lf%s", &(pinpar->u2 )) == EOF ||
    fscanf(fp, "%lf%s", &(pinpar->u3 )) == EOF ||
    fscanf(fp, "%lf%s", &(pinpar->C3 )) == EOF ||
    fscanf(fp, "%lf%s", &(pinpar->disp )) == EOF ||
    fscanf(fp, "%lf%s", &(pinpar->time0 )) == EOF ||
    fscanf(fp, "%d%s", &(pinpar->mode )) == EOF ||
    fscanf(fp, "%s%s", wgt_fname ) == EOF ||
    fscanf(fp, "%s%s", inp_fname ) == EOF ||
    fscanf(fp, "%s%s", out_fname ) == EOF ||
    fscanf(fp, "%s%s", log_fname ) == EOF ) {
    fclose(fp);
    fprintf(stderr, "Cannot read input; file: %s aborted\n", p);
    return -1;
}

if ( pinpar->pnum <= 0 ) {
    warning("Illegal pnum: %d\n", pinpar->pnum);
    return -1;
}
if ( pinpar->temper <= 0.0 ) {
    warning("Illegal Temperature: %lf\n", pinpar->temper);
    return -1;
}
if ( pinpar->Lco <= 0.0 ) {
    warning("Illegal Linear CutOff: %lf\n", pinpar->Lco);
    return -1;
}
if ( pinpar->frict <= 0.0 ) {
    warning("Illegal friction: %lf\n", pinpar->frict);
    return -1;
}
if (log_fname[0] != '-' || log_fname[1])
    freopen(log_fname, "a+", stderr);

/* Reading description of polys */

pinpar->poly = pp = (plm_desc *)emalloc( pinpar->pnum*sizeof(plm_desc) );
for ( i = 0, tot = 0; i < pinpar->pnum; i++ ) {
    if (fscanf(fp, "%d%c", &size, &top) <= 0 || size <= 0) {
        warning("Illegal description of %d poly\n", i);
        return -1;
    }
    pp[i].top = (top == 'o' || top == 'O') ? OPEN_TOP : RING_TOP;
}

```

```

    pp[i].psz = size;
    tot += size;
}
fscanf(fp, "%s");
pinpar->psize = tot;

/* Reading gsck_run structures */

for ( prev=NULL; fscanf(fp, "%le%s", &(tmprun.ini_dt)) > 0 &&
    fscanf(fp, "%le%s", &(tmprun.error)) > 0 &&
    fscanf(fp, "%d%s", &(tmprun.tnum) ) > 0 &&
    fscanf(fp, "%le%s", &(tmprun.mint) ) > 0; ) {
    gsck_run *nrun = (gsck_run *)emalloc( sizeof(gsck_run) );
    nrun->ini_dt = tmprun.ini_dt;
    nrun->error = tmprun.error;
    nrun->tnum = tmprun.tnum;
    nrun->mint = tmprun.mint;
    if (!prev) {
        pinpar->pkin = nrun;
        prev = nrun;
    }
    else {
        prev->next = nrun;
        prev = nrun;
    }
}
if (prev) prev->next = NULL;
fclose(fp);
return 0;
}

/* ===== End of File: gsck_plarb.c ===== */
/* ===== File: gsck_plarb_f.c ===== */

#include <gsck_plarb.h>

gsck_workval *E_workvals;

static int runge_kutt_step(RK_type*, RK_type*, RK_type *,double,double);
static double make_step_in_time(double, double *, double);
static int clc_derivs(RK_type *, RK_type *, double);
static int clc_derivs_hd(RK_type *, RK_type *, double);
static int clc_derivs_nohd(RK_type *, RK_type *, double);
static RK_type *RK_ct(void);
static void RK_dt(RK_type *);
static int size_of_one_record(int, int);

static observ S_obs;

/* ----- */

int gsck_ctor(gsck_inparam *inpars)
{
    gsck_workval *pwp = (gsck_workval *)emalloc( sizeof(gsck_workval) );
    int psize, i, j;
    FILE *wfp, *ifp, *ofp;
}

```

```

double cch = pow(2.0*M_PI, -1.5), *pwgt;

/* Initialization of Interaction & Other Parameters */
E_workvals = pwp;
pwp->p_obs = &S_obs;
pwp->psize = psize = inpars->psize;
pwp->pnum = inpars->pnum;
pwp->poly = inpars->poly;
pwp->temp = inpars->temper;
pwp->Lco = inpars->Lco;
pwp->sph = inpars->spring;
pwp->m1 = 1.0/inpars->frict;
if (inpars->visc <= 0.0) {
    pwp->hd_flag = 0;
    pwp->v1 = 0.0;
}
else {
    pwp->hd_flag = 1;
    pwp->v1 = 1.0/( 3.0*sqrt(2.0*M_PI*M_PI*M_PI)*inpars->visc );
}
pwp->curtime = inpars->time0;
pwgt = (double *)emalloc(psize*sizeof(double));
for ( i = 0; i < psize; i++ )
    pwgt[i] = 0.0;
if ( (wfp=fopen(inpars->wgt_fname, "r")) == NULL )
    warning("cannot open wgt_file '%s'; Using zeroes\n", inpars->wgt_fname);
else {
    for ( i = 0; i < psize; i++ ) {
        if (fscanf(wfp, "%le", pwgt+i) == EOF) {
            warning("zeroes starting from m=%d\n", i);
            break;
        }
    }
    fclose(wfp);
}
pwp->pp_u2h = (double **)emalloc(psize*sizeof(double *));
for ( i = 0; i < psize; i++ ) {
    pwp->pp_u2h[i] = (double *)emalloc(psize*sizeof(double));
    for ( j = 0; j < psize; j++ ) {
        pwp->pp_u2h[i][j] =
            cch * (inpars->u2 + inpars->disp*(pwgt[i]+pwgt[j])/2.0);
    }
}
free(pwgt);
pwp->u3h = inpars->u3*cch*cch;
pwp->C3 = inpars->C3;

/* RuKut Array's Initialization */
for ( i = 0; i < 7; i++ )
    pwp->p_deriv[i] = RK_ct();
pwp->p_Xq = RK_ct();
pwp->p_Xqn = RK_ct();
pwp->p_cerr = RK_ct();

/* ppp_l Arrays Initialization */
pwp->ppp_l = (double ***)emalloc( psize*sizeof(double ** ) );
for ( i = 0; i < psize; i++ ) {

```

```

    int j;
    pwp->ppp_l[i] = (double **)emalloc( psize*sizeof(double * ) );
    for ( j = 0; j < psize; j++ ) {
        pwp->ppp_l[i][j] = (double *)emalloc( psize*sizeof(double) );
    }
}

/* ppp_Vx Arrays Initialization */
pwp->pp_D52 = (double **)emalloc( psize*sizeof(double * ) );
pwp->pp_D40 = (double **)emalloc( psize*sizeof(double * ) );
pwp->pp_V1 = (double **)emalloc( psize*sizeof(double * ) );
pwp->pp_V2 = (double **)emalloc( psize*sizeof(double * ) );
pwp->pp_V3 = (double **)emalloc( psize*sizeof(double * ) );
pwp->pp_Vsi = (double **)emalloc( psize*sizeof(double * ) );
pwp->pp_Val = (double **)emalloc( psize*sizeof(double * ) );
pwp->pp_ent = (double **)emalloc( psize*sizeof(double * ) );
pwp->p_Rgm = (double *)emalloc( psize*sizeof(double) );
for ( i = 0; i < psize; i++ ) {
    pwp->pp_D52[i] = (double *)emalloc( psize*sizeof(double) );
    pwp->pp_D40[i] = (double *)emalloc( psize*sizeof(double) );
    pwp->pp_V1[i] = (double *)emalloc( psize*sizeof(double) );
    pwp->pp_V2[i] = (double *)emalloc( psize*sizeof(double) );
    pwp->pp_V3[i] = (double *)emalloc( psize*sizeof(double) );
    pwp->pp_Vsi[i] = (double *)emalloc( psize*sizeof(double) );
    pwp->pp_Val[i] = (double *)emalloc( psize*sizeof(double) );
    pwp->pp_ent[i] = (double *)emalloc( psize*sizeof(double) );
    pwp->pp_D52[i][i] = pwp->pp_D40[i][i] = 0.0;
}
if (pwp->hd_flag) {
    pwp->pp_Wmn = (double **)emalloc( psize*sizeof(double * ) );
    pwp->pp_Mob = (double **)emalloc( psize*sizeof(double * ) );
    for ( i = 0; i < psize; i++ ) {
        pwp->pp_Wmn[i] = (double *)emalloc( psize*sizeof(double) );
        pwp->pp_Mob[i] = (double *)emalloc( psize*sizeof(double) );
        pwp->pp_Mob[i][i] = pwp->m1;
    }
}

pwp->p_obs->Rga = (double *)emalloc( pwp->pnum*sizeof(double) );
pwp->p_obs->Mpa = (double *)emalloc( pwp->pnum*sizeof(double) );
gsce_clc_spring();

/* Output File Opening */
if ( (ofp=fopen(inpars->out_fname, "rb")) != NULL ) {
    int pnl, szl;
    double tml, lc1, spl, fr1, vc1, u21, u31, dsp;
    fread(&pnl, sizeof(int), 1, ofp);
    fread(&szl, sizeof(int), 1, ofp);
    fread(&tml, sizeof(double), 1, ofp);
    fread(&lc1, sizeof(double), 1, ofp);
    fread(&spl, sizeof(double), 1, ofp);
    fread(&fr1, sizeof(double), 1, ofp);
    fread(&vc1, sizeof(double), 1, ofp);
    fread(&u21, sizeof(double), 1, ofp);
    fread(&u31, sizeof(double), 1, ofp);
    fread(&dsp, sizeof(double), 1, ofp);
}

```

```

if ( pnl != inpars->pnum      || sz1 != inpars->psize  ||
    tml != inpars->temper    || lcl != inpars->Lco    ||
    spl != inpars->spring    || frl != inpars->frict  ||
    vc1 != inpars->visc     || u21 != inpars->u2     ||
    u31 != inpars->u3       || dsp != inpars->disp ) {
    warning("Interaction parameters are different\n");
    return -1;
}
else {
    fclose(ofp);
    if ( (ofp=fopen(inpars->out_fname,"ab+")) == NULL ) {
        warning("Cannot reopen outfile: %s\n", inpars->out_fname);
        return -1;
    }
    else {
        warning("Tables coincide. Appending to '%s'\n", inpars->out_fname);
    }
}
}
else {
    if ( (ofp=fopen(inpars->out_fname,"wb+")) == NULL ) {
        warning("Cannot open outfile: %s\n", inpars->out_fname);
        return -1;
    }
    else {
        fwrite(&(inpars->pnum ), sizeof(int), 1, ofp);
        fwrite(&(inpars->psize ), sizeof(int), 1, ofp);
        fwrite(&(inpars->temper), sizeof(double), 1, ofp);
        fwrite(&(inpars->Lco ), sizeof(double), 1, ofp);
        fwrite(&(inpars->spring), sizeof(double), 1, ofp);
        fwrite(&(inpars->frict ), sizeof(double), 1, ofp);
        fwrite(&(inpars->visc ), sizeof(double), 1, ofp);
        fwrite(&(inpars->u2 ), sizeof(double), 1, ofp);
        fwrite(&(inpars->u3 ), sizeof(double), 1, ofp);
        fwrite(&(inpars->disp ), sizeof(double), 1, ofp);
    }
}
pwp->ofp = ofp;

/* Reading Starting Configuration */
if (inpars->mode) {
/* ==== Initialization from Binary Configuration ==== */
    int pnl, sz1;
    double tml, lcl, spl, frl, vc1, u21, u31, dsp;

    if ( (ifp=fopen(inpars->inp_fname, "rb")) == NULL ) {
        warning("Cannot open binary inpfile: %s\n", inpars->inp_fname);
        return -1;
    }
    fread(&pnl, sizeof(int), 1, ifp);
    fread(&sz1, sizeof(int), 1, ifp);
    fread(&tml, sizeof(double), 1, ifp);
    fread(&lcl, sizeof(double), 1, ifp);
    fread(&spl, sizeof(double), 1, ifp);
    fread(&frl, sizeof(double), 1, ifp);
    fread(&vc1, sizeof(double), 1, ifp);
}

```

```

    fread(&u21, sizeof(double), 1, ifp);
    fread(&u31, sizeof(double), 1, ifp);
    fread(&dsp, sizeof(double), 1, ifp);
    if (pnl != pwp->pnum) {
        warning("PolyNum is %d in input file: %s\n", pnl, inpars->inp_fname);
        return -1;
    }
    if (sz1 != psize) {
        warning("PolySize is %d in input file: %s\n", sz1, inpars->inp_fname);
        return -1;
    }
    if (tml != inpars->temper || frl != inpars->frict ||
        spl != inpars->spring || u21 != inpars->u2 ||
        vc1 != inpars->visc || lcl != inpars->Lco ||
        u31 != inpars->u3 || dsp != inpars->disp ) {
        warning("Interaction Parameters are different\n");
    }

    fseek(ifp, -size_of_one_record(psize, pwp->pnum), SEEK_END);
    fread(&S_obs, sizeof(observ) - 2*sizeof(double *), 1, ifp);
    fread(S_obs.Rga, sizeof(double), pwp->pnum, ifp);
    fread(S_obs.Mpa, sizeof(double), pwp->pnum, ifp);
    for (i = 0; i < psize-1; i++) {
        fread(pwp->p_Xq->Dmn[i]+i+1, sizeof(double), psize-i-1, ifp);
        for ( j = i+1; j < psize; j++ )
            pwp->p_Xq->Dmn[j][i] = pwp->p_Xq->Dmn[i][j];
    }
    fclose(ifp);
    if (pwp->p_obs->t != pwp->curtime) {
        warning("Assuming last time from binary data as initial time\n");
        pwp->curtime = pwp->p_obs->t;
    }
}
else {
/* ==== Initialization from Text Configuration of 'gsce_gen' ==== */
    int pnl, m, sz1;
    if ( (ifp=fopen(inpars->inp_fname, "r")) == NULL ) {
        warning("Cannot open text inpfile: %s\n", inpars->inp_fname);
        return -1;
    }
    /* Header & Observables */
    fscanf(ifp, "%d*%le*%le*%le*s", &pnl);
    for (m = 0, sz1 = 0; m < pnl; m++) {
        int pp;
        fscanf(ifp, "%d*%c", &pp);
        sz1 += pp;
    }
    if (pnl != pwp->pnum) {
        warning("PolyNum is %d in input file: %s\n", pnl, inpars->inp_fname);
        return -1;
    }
    if (sz1 != psize) {
        warning("PolySize is %d in input file: %s\n", sz1, inpars->inp_fname);
        return -1;
    }
}
}

```

```

fscanf(ifp, "%*le%*le%*le%*le"); /* u2 u3 c3 disp */
fscanf(ifp, "%*le%*le%*le%*le"); /* rg2 mps eng ent */
fscanf(ifp, "%*le%*le%*le"); /* e0 e1 e2 */

for (m = 0; m < pnl; m++) {
    fscanf(ifp, "%*le%*le"); /* rg2_a mps_a */
}

for (i = 0; i < psize; i++) {
    for (j = i+1; j < psize; j++) {
        if (fscanf(ifp, "%*d%*d%le",
            &(pwp->p_Xq->Dmn[i][j])) <= 0) {
            warning("Cannot read inpfile: %s\n", inpars->inp_fname);
            return -1;
        }
        pwp->p_Xq->Dmn[j][i] = pwp->p_Xq->Dmn[i][j];
    }
}
fclose(ifp);
gsck_print();
}
return 0;
}

/* ----- */

void gsck_dtor(void)
{
    gsck_workval *pwp = E_workvals;
    int i, j, psz = pwp->psize;

    for (i = 0; i < psz; i++) {
        for (j = 0; j < psz; j++) {
            free(pwp->ppp_l[i][j]);
        }
        free(pwp->ppp_l[i]);
        free(pwp->pp_D52[i]);
        free(pwp->pp_D40[i]);
        free(pwp->pp_V1[i]);
        free(pwp->pp_V2[i]);
        free(pwp->pp_V3[i]);
        free(pwp->pp_Vsi[i]);
        free(pwp->pp_Val[i]);
        free(pwp->pp_ent[i]);
    }
    free(pwp->ppp_l);
    free(pwp->pp_D52);
    free(pwp->pp_D40);
    free(pwp->pp_V1);
    free(pwp->pp_V2);
    free(pwp->pp_V3);
    free(pwp->pp_Vsi);
    free(pwp->pp_Val);
    free(pwp->pp_ent);
    free(pwp->p_Rgm);
    free(pwp->p_obs->Rga);
    free(pwp->p_obs->Mpa);
}

```

```

if (pwp->hd_flag) {
    for (i = 0; i < psz; i++) {
        free(pwp->pp_Wmn[i]);
        free(pwp->pp_Mob[i]);
    }
    free(pwp->pp_Wmn);
    free(pwp->pp_Mob);
}

for (i = 0; i < 7; i++) {
    RK_dt(pwp->p_deriv[i]);
}
RK_dt(pwp->p_Xq);
RK_dt(pwp->p_Xqn);
RK_dt(pwp->p_cerr);

free(pwp->poly);
fclose(pwp->ofp);
free(pwp);

E_workvals = NULL;
}

/* ----- */

void gsck_print(void)
{
    gsck_workval *pwp = E_workvals;
    int m, psz = pwp->psize, pnum = pwp->pnum, ll = psz-1;
    FILE *ofp = pwp->ofp;

    gsck_clc_observ();
    fwrite(pwp->p_obs, sizeof(observ) - 2*sizeof(double *), 1, ofp);
    fwrite(pwp->p_obs->Rga, sizeof(double), pnum, ofp);
    fwrite(pwp->p_obs->Mpa, sizeof(double), pnum, ofp);
    for (m = 0; m < ll; m++) {
        fwrite(pwp->p_Xq->Dmn[m]+m+1, sizeof(double), ll-m, ofp);
    }
    fflush(ofp);
}

/* ----- */

int gsck_process(gsck_run *prun)
{
    gsck_workval *pwp = E_workvals;
    int ex_step, in_step;
    double curt = pwp->curtime, dt = prun->ini_dt, mint = prun->mint,
        error = prun->error;

    /* Main Loop */
    for (ex_step = 0; ex_step < prun->tnum; ex_step++) {
        double t_last = 0.0;
        for (in_step = 0; t_last < mint; in_step++) {
            double dt_max = mint-t_last, /* what's left */
                dt_new = make_step_in_time(curt, &dt, error);

```

```

    if (dt < 0.0) {
        warning("\nFail at estep = %d, istep = %d\n", ex_step, in_step);
        return -1;
    }
    fputc(' ', stderr); fflush(stderr);
    curt += dt;
    t_last += dt;
    dt = (dt_new > dt_max) ? dt_max : dt_new;
}
pwp->curtime = curt;
gsck_print();
fputc('\n', stderr); fflush(stderr);
}
return 0;
}

/* ----- */
/* Returns proposed timestep, put in pdt curently performed one*/

#define MIN_DTIME 1.0e-12
#define SAFETY 0.9
#define PGROW -0.2
#define PSHRINK -0.25
#define ERRCON 1.89e-04
#define MAX_INCREASE 5.0
#define MIN_DECREASE 0.1

static double make_step_in_time(double curt, double *pdt, double error)
{
    gsck_workval *pwp = E_workvals;
    double dt = *pdt;
    register int m, n, psz = pwp->psize;

    if (clc_derivs(pwp->p_Xq, pwp->p_deriv[0], curt) == -1) {
        warning("\nclc_deriv failed at ini point\n");
        return (*pdt = -1.0);
    }

    for ( ; ; ) {
        register double max_rel_error = 0.0, **Dmn = pwp->p_Xq->Dmn,
            **Err = pwp->p_cerr->Dmn;
        if (runge_kutt_step(pwp->p_Xq, pwp->p_Xqn, pwp->p_cerr, curt, dt) == -1) {
            if ( (dt * = MIN_DECREASE) < MIN_DTIME) {
                warning("\nmake_step: timestep underflow\n");
                return (*pdt = -1.0);
            }
            warning("\nrk_failed resetting dt to %le\n", dt);
            continue;
        }
        /* derivs OK, checking errors & stuff */
        for ( m = 0; m < psz; m++ ) {
            for ( n = m+1; n < psz; n++ ) {
                register double rel_error = fabs(Err[m][n]/Dmn[m][n]);
                if (rel_error > max_rel_error) max_rel_error = rel_error;
            }
        }
        max_rel_error /= error;
    }
}

```

```

    if (max_rel_error > 1.0) { /* Failed to keep errors in integration */
        double new_dt = SAFETY*dt*pow(max_rel_error, PSHRINK);
        dt = (new_dt > MIN_DECREASE*dt) ? new_dt : MIN_DECREASE*dt;
        if (dt < MIN_DTIME) {
            warning("\nmake_step: timestep underflow\n");
            return (*pdt = -1.0);
        }
    }
    else { /* Everything is OK; estimating new time step */
        double new_dt = (max_rel_error > ERRCON) ?
            SAFETY*dt*pow(max_rel_error, PGROW) : MAX_INCREASE*dt;
        RK_type *p = pwp->p_Xq;
        pwp->p_Xq = pwp->p_Xqn; pwp->p_Xqn = p; /* Pointers swap */
        *pdt = dt;
        return new_dt;
    }
}

/* ----- */
/* Warning!!! do not calculate the derivative at initial point */

static int runge_kutt_step(RK_type *rk_in, RK_type *rk_out, RK_type *rk_err,
    double curt, double dt)
{
    {
        static char *mesg = "runge_kutt_step fails at iteration %d\n";
        /* Cash-Karp Parameters */
        static double aa2 = 0.2, aa3 = 0.3, aa4 = 0.6, aa5 = 1.0, aa6 = 0.875;
        static double cc1 = 37.0/378.0, cc3 = 250.0/621.0, cc4 = 125.0/594.0,
            cc6 = 512.0/1771.0, dc5 = 277.0/14336.0;
        double dc1 = 2825.0/27648.0 - cc1, dc3 = 18575.0/48384.0 - cc3,
            dc4 = 13525.0/55296.0 - cc4, dc6 = 1.0/ 4.0 - cc6;
        static double b21 = 0.2, b31 = 0.075, b41 = 0.3, b51 = -11.0/54.0,
            b61 = 1631.0/55296.0, b32 = 0.225, b42 = -0.9, b52 = 2.5,
            b62 = 175.0/512.0, b43 = 1.2, b53 = -70.0/27.0,
            b63 = 575.0/13824.0, b54 = 35.0/27.0,
            b64 = 44275.0/110592.0, b65 = 253.0/4096.0;

        gsck_workval *pwp = E_workvals;
        RK_type *k1 = pwp->p_deriv[0], *k2 = pwp->p_deriv[1], *k3 = pwp->p_deriv[2],
            *k4 = pwp->p_deriv[3], *k5 = pwp->p_deriv[4], *k6 = pwp->p_deriv[5],
            *kt = pwp->p_deriv[6];
        register double **D1=k1->Dmn, **D2=k2->Dmn, **D3=k3->Dmn, **D4=k4->Dmn,
            **D5=k5->Dmn, **D6=k6->Dmn, **Dt=kt->Dmn, **Di=rk_in->Dmn,
            **Do=rk_out->Dmn, **De=rk_err->Dmn;
        register int m, n, psz = pwp->psize;

        for ( m = 0; m < psz; m++ )
            for ( n = m+1; n < psz; n++ )
                Dt[m][n] = Dt[n][m] = Di[m][n] + dt*b21*D1[m][n];
        if ( clc_derivs(kt, k2, curt+aa2*dt) == -1 ) { /* 1st step */
            warning(mesg, 1);
            return -1;
        }
    }

    for ( m = 0; m < psz; m++ )

```

```

for ( n = m+1; n < psz; n++ )
    Dt[m][n] = Dt[n][m] = Di[m][n] + dt*(b31*D1[m][n]+b32*D2[m][n]);
if ( clc_derivs(kt, k3, curt+aa3*dt) == -1 ) { /* 2nd step */
    warning(mesg, 2);
    return -1;
}

for ( m = 0; m < psz; m++ )
    for ( n = m+1; n < psz; n++ )
        Dt[m][n] = Dt[n][m] = Di[m][n] + dt*(b41*D1[m][n]+b42*D2[m][n]+
        b43*D3[m][n]);
if ( clc_derivs(kt, k4, curt+aa4*dt) == -1 ) { /* 3rd step */
    warning(mesg, 3);
    return -1;
}

for ( m = 0; m < psz; m++ )
    for ( n = m+1; n < psz; n++ )
        Dt[m][n] = Dt[n][m] = Di[m][n] + dt*(b51*D1[m][n]+b52*D2[m][n]+
        b53*D3[m][n]+b54*D4[m][n]);
if ( clc_derivs(kt, k5, curt+aa5*dt) == -1 ) { /* 4th step */
    warning(mesg, 4);
    return -1;
}

for ( m = 0; m < psz; m++ )
    for ( n = m+1; n < psz; n++ )
        Dt[m][n] = Dt[n][m] = Di[m][n] + dt*(b61*D1[m][n]+b62*D2[m][n]+
        b63*D3[m][n]+b64*D4[m][n]+
        b65*D5[m][n]);
if ( clc_derivs(kt, k6, curt+aa6*dt) == -1 ) { /* 5th step */
    warning(mesg, 5);
    return -1;
}

for ( m = 0; m < psz; m++ ) {
    for ( n = m+1; n < psz; n++ ) {
        register double newD = Di[m][n] + dt*(cc1*D1[m][n]+cc3*D3[m][n]+
        cc4*D4[m][n]+cc6*D6[m][n]),
        errD = dt*(dc1*D1[m][n] + dc3*D3[m][n] + dc4*D4[m][n]
        + dc5*D5[m][n] + dc6*D6[m][n]);
        if (newD < SMALL_VAL) {
            warning(mesg, 6);
            return -1;
        }
        Do[m][n] = Do[n][m] = newD;
        De[m][n] = De[n][m] = errD;
    }
}
return 0;
}
/* ----- */
static int clc_derivs(RK_type *lval, RK_type *deriv, double t)
{

```

```

return (E_workvals->hd_flag) ?
    clc_derivs_hd(lval, deriv, t) : clc_derivs_nohd(lval, deriv, t);
}
/* ----- */
static int clc_derivs_nohd(RK_type *lval, RK_type *deriv, double curt)
{
    gsck_workval *pwp = E_workvals;
    register int psz = pwp->psize, m, n, k;
    double temp = 2.0*pwp->temp, mobil = 2.0*pwp->ml;
    double **Dmn = lval->Dmn, **rvDmn = deriv->Dmn, **Val = pwp->pp_Val;

    if (gsck_clc_poten(lval, curt) == -1) return -1;

    for ( m = 0; m < psz; m++ ) {
        register double *pDm = Dmn[m], *pVm = Val[m];
        for ( n = m+1; n < psz; n++ ) {
            register double Veng_mn = 0.0, *pDn = Dmn[n], *pVn = Val[n];
            for ( k = 0; k < psz; k++ ) {
                Veng_mn += (pDm[k] - pDn[k]) * (pVm[k] - pVn[k]);
            }
            rvDmn[m][n] = rvDmn[n][m] = mobil * (temp + Veng_mn);
        }
    }

    return 0;
}
/* ----- */
static int clc_derivs_hd(RK_type *lval, RK_type *deriv, double curt)
{
    gsck_workval *pwp = E_workvals;
    register int psz = pwp->psize, m, n, k;
    double **Dmn = lval->Dmn, **rvDmn = deriv->Dmn,
        **Wmn = pwp->pp_Wmn, **Mob = pwp->pp_Mob;
    double temp = 4.0*pwp->temp, MDiag = Mob[0][0];

    if (gsck_clc_poten(lval, curt) == -1) return -1;

    for ( m = 0; m < psz; m++ ) {
        register double *pDm = Dmn[m], *pWm = Wmn[m];
        for ( n = m+1; n < psz; n++ ) {
            register double Veng_mn = 0.0, *pDn = Dmn[n], *pWn = Wmn[n];
            for ( k = 0; k < psz; k++ ) {
                Veng_mn += (pDm[k] - pDn[k]) * (pWm[k] - pWn[k]);
            }
            rvDmn[m][n] = rvDmn[n][m] = temp*(MDiag - Mob[m][n]) + 2.0*Veng_mn;
        }
    }

    return 0;
}
/* ----- */

```

```

static RK_type *RK_ct(void)
{
    gsck_workval *pwp = E_workvals;
    int i, psize = pwp->psize;
    RK_type *ret = (RK_type *)emalloc(sizeof(RK_type));

    ret->Dmn = (double **)emalloc( psize*sizeof(double * ) );
    for ( i = 0; i < psize; i++ ) {
        ret->Dmn[i] = (double *)emalloc( psize*sizeof(double) );
        ret->Dmn[i][i] = 0.0;
    }

    return ret;
}

/* ----- */

static void RK_dt(RK_type *rk)
{
    gsck_workval *pwp = E_workvals;
    int i, psz = pwp->psize;

    for ( i = 0; i < psz; i++ ) {
        free(rk->Dmn[i]);
    }
    free(rk->Dmn);
    free(rk);
}

/* ----- */

static int size_of_one_record(int psize, int pnum)
{
    return sizeof(observ) - 2*sizeof(double *) + 2*pnum*sizeof(double) +
        (psize-1)*psize*sizeof(double)/2;
}

/* ===== End of File:    gsck_plarb_f.c    ===== */
/* ===== File:         gsck_plarb_vx.c    ===== */
/* ===== Calculation of Effective Potentials etc... ===== */

#include <gsck_plarb.h>

static int    clc_D52(RK_type *);
static int    clc_D52_Mob(RK_type *);
static int    clc_lam(RK_type *);
static void    clc_V2 (RK_type *);
static void    clc_V3 (RK_type *);
static void    clc_Vsi(RK_type *);
static double clc_mps_rg2(double *);
static void    clc_mps_rg2_one(double *, double *, int, int);
static double clc_eng(double *);
static double clc_ent(void);
static double determinant(int, double **);

```

```

#define CLC_52_USE_SQRT

/* ----- */

int gsck_clc_poten(RK_type *Xq, double cur_time)
{
    gsck_workval *pwp = E_workvals;
    register int m, n, psz = pwp->psize;
    double **V1 = pwp->pp_V1, **V2 = pwp->pp_V2,
        **V3 = pwp->pp_V3, **Vsi = pwp->pp_Vsi;
    register double **Val = pwp->pp_Val;
    int hd_flag = pwp->hd_flag;

    if (hd_flag) {
        if ( clc_D52_Mob(Xq) == -1 || clc_lam(Xq) == -1 ) return -1;
    }
    else {
        if ( clc_D52(Xq) == -1 || clc_lam(Xq) == -1 ) return -1;
    }
    clc_V2(Xq); clc_V3(Xq); clc_Vsi(Xq);

    for ( m = 0; m < psz; m++ ) {
        register double *pV1=V1[m], *pV2=V2[m], *pV3=V3[m], *pVsi=Vsi[m];
        for ( n = m; n < psz; n++ ) {
            Val[m][n] = Val[n][m] = pV1[n] + pV2[n] + pV3[n] + pVsi[n];
        }
    }

    if (hd_flag) {
        register double **Wmn = pwp->pp_Wmn;
        for ( m = 0; m < psz; m++ ) {
            register double *pMob = pwp->pp_Mob[m];
            for ( n = 0; n < psz; n++ ) {
                register double sum = 0.0, *pVal = Val[n];
                register int k;
                for ( k = 0; k < psz; k++ ) {
                    sum += pMob[k] * pVal[k];
                }
                Wmn[m][n] = sum;
            }
        }
    }

    return 0;
}

/* ----- */

void gsck_clc_observ(void)
{
    gsck_workval *pwp = E_workvals;
    observ *obs = pwp->p_obs;

    obs->t = pwp->curtime;
    obs->mps = clc_mps_rg2(&obs->rg2);
    obs->eng = clc_eng(obs->cnt);
    obs->ent = clc_ent();
}

```

```

/* ----- */
void gsce_clc_spring(void)
{
    gsck_workval *pwp = E_workvals;
    plm_desc *ppd = pwp->poly;
    int m1, m2, pn1, pn2, all, al2, pnum = pwp->pnum, psize = pwp->psize;
    double **V1 = pwp->pp_V1, spr = 0.5*pwp->sph,
           spl = 0.5*pwp->temp/(pwp->Lco*pwp->Lco);

    for ( m1 = 0, pn1 = 0, all = 0; m1 < psize; m1++ ) {
        int sz1 = ppd[pn1].psz, top = ppd[pn1].top;
        for ( m2 = 0, pn2 = 0, al2 = 0; m2 < m1; m2++ ) {
            int sz2 = ppd[pn2].psz;
            register int vai = 0;
            register double vae = ( (pn1 == pn2) ? 1.0/(double)(sz1*sz1) : 0.0)
                - (double)pnum / (double)(psize*psize), rrr=0.0;

0;
            if (pn1 == pn2) {
                if (al2+1==all || (top==RING_TOP && !al2 && all==sz1-1)) vai = 1
;
            }
            V1[m1][m2] = V1[m2][m1] = -spr*(double)vai + spl*vae - rrr;
            if (++al2 == sz2) { al2 = 0; pn2++; }
        }
        if (++all == sz1) { all = 0; pn1++; }
    }

    for ( m1 = 0; m1 < psize; m1++ ) {
        register double val = 0.0;
        for ( m2 = 0; m2 < psize; m2++ ) {
            if (m1 == m2) continue;
            val -= V1[m1][m2];
        }
        V1[m1][m1] = val;
    }
}

/* ----- */

static int clc_D52(RK_type *Xq)
{
    gsck_workval *pwp = E_workvals;
    register int m, n, psz = pwp->psize;
    double **Dmn = Xq->Dmn;
    register double **D52 = pwp->pp_D52, **D40 = pwp->pp_D40;

    for ( m = 0; m < psz; m++ ) {
        register double *p_Dm = Dmn[m];
        for ( n = m+1; n < psz; n++ ) {
            register double val = p_Dm[n];
            if (val < SMALL_VAL) {
                warning("D[%d][%d] = %le\n", m, n, val);
                return -1;
            }
            D52[m][n] = D52[n][m] =
#ifdef CLC_52_USE_SQRT

```

```

1.0/(val*val*sqrt(val));
        #else
            pow(val, -2.5);
        #endif
        D40[m][n] = D40[n][m] = 1.0/(val*val*val*val);
    }
}
return 0;
}

/* ----- */

static int clc_D52_Mob(RK_type *Xq)
{
    gsck_workval *pwp = E_workvals;
    register int m, n, psz = pwp->psize;
    double **Dmn = Xq->Dmn;
    register double **D52 = pwp->pp_D52, **D40 = pwp->pp_D40,
           **Mob = pwp->pp_Mob, v1 = pwp->v1;

    for ( m = 0; m < psz; m++ ) {
        register double *p_Dm = Dmn[m];
        for ( n = m+1; n < psz; n++ ) {
            register double val = p_Dm[n], v12;
            if (val < SMALL_VAL) {
                warning("D[%d][%d] = %le\n", m, n, val);
                return -1;
            }
            v12 = 1.0/sqrt(val);
            Mob[m][n] = Mob[n][m] = v1*v12;
            D52[m][n] = D52[n][m] = v12/(val*val);
            D40[m][n] = D40[n][m] = 1.0/(val*val*val*val);
        }
    }
    return 0;
}

/* ----- */

static int clc_lam(RK_type *Xq)
{
    gsck_workval *pwp = E_workvals;
    register int m, n, k, psz = pwp->psize;
    register double **pp_Dmn = Xq->Dmn, ***lam = pwp->ppp_l;

    for ( m = 0; m < psz; m++ ) {
        for ( n = m+1; n < psz; n++ ) {
            register double Dmn = pp_Dmn[m][n];
            for ( k = n+1; k < psz; k++ ) {
                register double Dmk = pp_Dmn[m][k], Dnk = pp_Dmn[n][k],
                       Dmnk = 0.5*(Dnk-Dmn-Dmk),
                       val = Dmn*Dmk - Dmnk*Dmnk;
                if (val < SMALL_VAL) {
                    warning("DD[%d][%d][%d] = %le\n", m, n, k, val);
                    return -1;
                }
                lam[m][n][k] = lam[m][k][n] =

```

```

        lam[n][m][k] = lam[n][k][m] =
        lam[k][n][m] = lam[k][m][n] =
#ifdef CLC_52_USE_SQRT
        1.0/(val*val*sqrt(val));
#else
        pow(val, -2.5);
#endif
    }
}
return 0;
}
}
/* ----- */

static void clc_V2(RK_type *Xq)
{
    gsck_workval *pwp = E_workvals;
    register int m1, m2, n, psz = pwp->psize;
    double **Dmn = Xq->Dmn, **D52 = pwp->pp_D52, **pp_u2 = pwp->pp_u2h;
    register double **V2 = pwp->pp_V2;

    for ( m1 = 0; m1 < psz; m1++ ) {
        register double *p_d52 = D52[m1], *p_u2m1 = pp_u2[m1];
        for ( m2 = m1+1; m2 < psz; m2++ ) {
            V2[m1][m2] = V2[m2][m1] = p_u2m1[m2] * p_d52[m2];
        }
    }

    for ( m1 = 0; m1 < psz; m1++ ) {
        register double val = 0.0;
        for ( n = 0; n < psz; n++ ) {
            if ( n == m1 ) continue;
            val -= V2[m1][n];
        }
        V2[m1][m1] = val;
    }
}
/* ----- */

static void clc_V3(RK_type *Xq)
{
    gsck_workval *pwp = E_workvals;
    register int m1, m2, n1, n2, psz = pwp->psize;
    register double **pp_Dmn = Xq->Dmn, ***ppp_l = pwp->ppp_l,
        **V3 = pwp->pp_V3, u3 = 1.5 * pwp->u3h;

    for ( m1 = 0; m1 < psz; m1++ ) {
        register double *p_Dm1 = pp_Dmn[m1];
        for ( m2 = m1+1; m2 < psz; m2++ ) {
            register double val=0.0, *p_lam = ppp_l[m1][m2],
                *p_Dm2 = pp_Dmn[m2], Dm1m2 = pp_Dmn[m1][m2];
            for ( n1 = 0; n1 < psz; n1++ ) {
                if ( n1 == m1 || n1 == m2 ) continue;
                val += p_lam[n1] * ( p_Dm1[n1] + p_Dm2[n1] - Dm1m2 );
            }
            V3[m1][m2] = V3[m2][m1] = u3 * val;
        }
    }
}

```

```

    }
}
for ( m1 = 0; m1 < psz; m1++ ) {
    register double val = 0.0;
    for ( n1 = 0; n1 < psz; n1++ ) {
        if ( n1 == m1 ) continue;
        val -= V3[m1][n1];
    }
    V3[m1][m1] = val;
}
}
/* ----- */

static void clc_Vsi(RK_type *Xq)
{
    gsck_workval *pwp = E_workvals;
    register int m1, m2, n, psz = pwp->psize;
    double **D40 = pwp->pp_D40, u3 = 2.0 * pwp->u3h * pwp->C3;
    register double **Vsi = pwp->pp_Vsi;

    for ( m1 = 0; m1 < psz; m1++ ) {
        register double *p_d40 = D40[m1];
        for ( m2 = m1+1; m2 < psz; m2++ ) {
            Vsi[m1][m2] = Vsi[m2][m1] = u3 * p_d40[m2];
        }
    }

    for ( m1 = 0; m1 < psz; m1++ ) {
        register double val = 0.0;
        for ( n = 0; n < psz; n++ ) {
            if ( n == m1 ) continue;
            val -= Vsi[m1][n];
        }
        Vsi[m1][m1] = val;
    }
}
/* ----- */

static double clc_eng(double *contrib)
{
    gsck_workval *pwp = E_workvals;
    register int psz = pwp->psize, m, n, k;
    register double e1 = 0.0, e2 = 0.0, e3 = 0.0, e3_si = 0.0,
        **pp_Dmn = pwp->pp_Dmn, **u2 = pwp->pp_u2h,
        **spr = pwp->pp_V1;

    for ( m = 0; m < psz; m++ ) { /* Spring term */
        for ( n = m+1; n < psz; n++ ) {
            e1 += spr[m][n]*pp_Dmn[m][n];
        }
    }
    e1 *= -3.0/(double)psz;

    for ( m = 0; m < psz; m++ ) { /* 2-body terms */
        for ( n = m+1; n < psz; n++ ) {
            register double d = pp_Dmn[m][n],

```

```

        d32 = (d <= SMALL_VAL) ? 0.0 : pow(d, -1.5);
        e2 += u2[m][n]*d32;
        e3_si += d32*d32; /* Self Interaction */
    }
}
e2 *= 2.0/(double)psz;
e3_si *= 2.0*pwp->u3h*pwp->C3/(double)psz;

for ( m = 0; m < psz; m++ ) { /* 3-body terms */
    for ( n = m+1; n < psz; n++ ) {
        register double Dmn = pp_Dmn[m][n];
        for ( k = n+1; k < psz; k++ ) {
            register double Dmk = pp_Dmn[m][k], Dnk = pp_Dmn[n][k],
                Dmnk = 0.5*(Dnk-Dmn-Dmk),
                dd = Dmn*Dmk - Dmnk*Dmnk;
            register double dd32 = (dd <= SMALL_VAL) ? 0.0 : pow(dd, -1.5);
            e3 += dd32;
        }
    }
}
e3 *= 6.0*pwp->u3h/(double)psz;

contrib[0] = e1;
contrib[1] = e2;
contrib[2] = e3+e3_si;
return e1 + e2 + e3 + e3_si;
}

/* ----- */

#define LOG_SMALL_VAL 1.0e-200

static double clc_ent(void)
{
    gsck_workval *pwp = E_workvals;
    register int psz = pwp->psize, i, j;
    register double **Dmn = pwp->p_Xq->Dmn, **ent = pwp->pp_ent,
        *Rgm = pwp->p_Rgm, Rg2 = pwp->Rg2, det;

    for ( i = 0; i < psz; i++ ) {
        for ( j = i; j < psz; j++ ) {
            register double s_ij = (Rgm[i] + Rgm[j]) / (2.0*(double)psz) -
                Rg2 - 0.5*Dmn[i][j];
            ent[i][j] = ent[j][i] = s_ij;
        }
    }

    if ( (det = determinant(psz-1, ent)) < LOG_SMALL_VAL ) {
        warning("clc_ent: det = %le\n", det);
        return 0.0;
    }

    return 1.5*pwp->temp*log(det)/(double)psz;
}

/* ----- */

static double clc_mps_rg2(double *pRg2)
{

```

```

    gsck_workval *pwp = E_workvals;
    register int m, n, psz = pwp->psize, pn = pwp->pnum, start, end;
    register double au2 = 0.0, mps = 0.0, rg2 = 0.0,
        **Dmn = pwp->p_Xq->Dmn, **pp_u2 = pwp->pp_u2h, *Rgm = pwp->p_Rgm;

    for ( m = 0; m < psz; m++ ) {
        register double r_m = 0.0;
        for ( n = 0; n < psz; n++ ) {
            register double u2_mn = pp_u2[m][n], d_mn = Dmn[m][n];
            mps += u2_mn*d_mn;
            au2 += u2_mn;
            r_m += d_mn;
        }
        rg2 += r_m;
        Rgm[m] = r_m;
    }

    for ( m = 0, start = 0; m < pn; m++, start = end ) {
        end = start + pwp->poly[m].psz;
        clc_mps_rg2_one(pwp->p_obs->Rga+m, pwp->p_obs->Mpa+m, start, end);
    }

    rg2 /= (double)(2.0*psz*psz);
    mps /= (double)(2.0*psz*psz);
    au2 /= (double)(psz*psz);
    mps -= au2*rg2;
    *pRg2 = rg2;
    pwp->Rg2 = rg2;
    return mps;
}

/* ----- */

static void clc_mps_rg2_one(double *pRg2, double *pMps, int start, int end)
{
    gsck_workval *pwp = E_workvals;
    register int m, n, psz = end-start;
    register double au2 = 0.0, mps = 0.0, rg2 = 0.0,
        **Dmn = pwp->p_Xq->Dmn, **pp_u2 = pwp->pp_u2h;

    for ( m = start; m < end; m++ ) {
        au2 += 0.5*pp_u2[m][m];
        for ( n = m+1; n < end; n++ ) {
            register double u2_mn = pp_u2[m][n], d_mn = Dmn[m][n];
            mps += u2_mn*d_mn;
            au2 += u2_mn;
            rg2 += d_mn;
        }
    }

    rg2 /= (double)(psz*psz);
    mps /= (double)(psz*psz);
    au2 /= 0.5*(double)(psz*psz);
    *pRg2 = rg2;
    *pMps = mps - au2*rg2;
}

/* ----- */

static double determinant(int msize, double **matrix)

```

```
{
register int i, k, n = msize, nl = n-1, p = 1;
register double **a = matrix, ret;

if (n < 1) return 0.0;
if (n == 1) return a[0][0];

for ( k = 0; k < nl; k++ ) {
register double s = fabs(a[k][k]), t;
register int j = k, kl = k+1, m;

for ( m = kl; m < n; m++ )
if ( s < (t=fabs(a[m][k])) ) {
s = t;
j = m;
}
if (s < 1.0e-50) return 0.0;
if (j != k) {
double *ptmp = a[k];
a[k] = a[j];
a[j] = ptmp;
p *= -1;
}
for ( i = kl; i < n; i++ ) {
register double alf = a[k][i]/a[k][k];
for ( j = kl; j < n; j++ )
a[j][i] -= alf*a[j][k];
}
}

for ( ret = (double)p, k = 0; k < n; k++ ) {
ret *= a[k][k];
}
return ret;
}

/* ===== End of File: gsck_plarb_vx.c ===== */
```

Summary

PhD Thesis

Theoretical study of equilibrium structures and kinetic phenomena in macromolecular solutions

Yu.A. Kuznetsov

Under the direction of Professor K.A. Dawson

and supervision of Professor K.A. Dawson and Dr E.G. Timoshenko

The thesis consists of **259** pages, **7** chapters including Introduction and Conclusion, **75** figures, **10** tables, Table of Contents, List of Tables, List of Figures, Acknowledgements, Abstract, Bibliography containing **180** items, and **2** Appendices containing listings of the C code of programs `Many_Cop 0.0.4a` of **1359** lines and `Gsck_P1Arb 0.0.1a` of **1399** lines. The material of the thesis is based on **9** papers, which have been published or already accepted for publication, with **8** of which in international refereed journals.

The thesis is devoted to studying the equilibrium structures and kinetic phenomena in coarse-grained statistical mechanical models of macromolecular solutions using a combination of analytical and computational techniques. For this the Gaussian Self-Consistent method and Monte Carlo techniques on a lattice and in continuous space have been developed and applied to the problem. The systems under study include homopolymers with varying flexibility and specific interactions, and amphiphilic heteropolymers with periodic and random sequences in a wide range of solution concentrations.