ORIGINAL PAPER

Andrey V. Maximov · Olga G. Maximova
Yuli Ya. Gotlib

Long-scale conformational properties of peptide chains in β-sheets

Abstract We propose a sufficiently simple multichain model of ordered β-sheets, composed of extended macromolecules with rigid elements. The effective constants of intra- and interchain interactions describe primary and secondary structures of proteins, respectively. It is found that the long-range correlation of orientations of chain elements decreases with the separation along the same chain or between different chains according to the same asymptotic power law. The exponent in this law is determined by the ratio of the energy of thermal motion and the geometric mean of the energies of intra- and interchain interactions. The characteristic scale parameters are obtained, which define the crossover of the intra- and interchain correlation functions from the exponential law of decrease to the power one. The power law for intrachain correlations leads to a non-Gaussian behavior of the mean-square dimensions of chains. Several types of asymptotic dependences of mean-square dimensions of a chain in the β-sheet on the number of chain elements are found. Peptide chains may exist in different conformations: from extended ones to random Gaussian coils. Long-scale statistical properties of polymer systems with interchain interactions and those for polymer chains with excluded volume effects are compared.

Keywords β-Sheets · Intra- and interchain correlations · Mean-square chain dimensions · Excluded volume effects

A.V. Maximov (✉) · O.G. Maximova
Institute of Mathematics, Physics and Informatics,
Cherepovets State University, Lunacharskii Prospect 5,
Cherepovets, 162000 Russia
e-mail: a_v_maximov@mail.ru
Tel.: +7-8202-2892-60, Fax: +7-8202-5565-97

Yu.Ya. Gotlib
Institute of Macromolecular Compounds,
Russian Academy of Sciences, Bolshoi Prospect 31, V.O.,
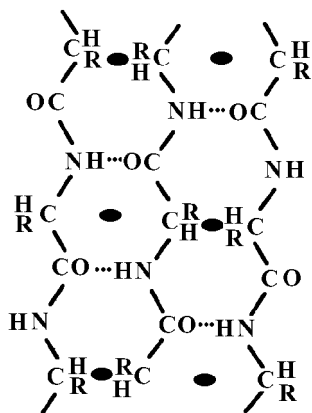St. Petersburg, 199004, Russia

Introduction

Investigation of self-assembled multichain systems composed of protein-like molecules is particularly interesting. They are biocompatible and biodegradable water-based systems, potentially highly responsive to moderate changes in the media properties (like pH, temperature, ionic composition, etc.). [1, 2, 3] The structures of natural peptides are very diverse, being determined by regular arrangements of hydrogen bonds as well as side-chain interactions [4].

The same peptide with the same primary structure can form various higher-order secondary structures with lability of the protein molecules: random coils, α-helices and β-sheets depending on solution conditions such as temperature, pH, ionic strength, and solvent composition [1, 2, 3]. In the last case strongly extended chain molecules are "stitched" together by hydrogen or other side-chain bonds or interactions (e.g. covalent, orientational and ionic interactions) forming independent structural units [4, 5, 6]. The rationally designed primary structure of peptides ensures that the β-sheet is nearly perfect: all molecules are aligned and connected in the same way; defects are extremely rare (Fig. 1) [1, 2, 3, 4]. Outstanding examples of such sheets are most simple fiber polypeptides with extended conformations of chains [4, 5, 6, 7, 8]. Configurations of parallel β-sheets are found in the extended forms of native proteins of keratin (hair, wool, horns etc.), in polyamides (nylon-6, nylon-66) and others polymers with a host of hydrogen interactions. Antiparallel β-sheets are formed, e.g. in silk fibroin [4, 7, 8]. The formation of such self-assembling β-sheet tapes was reported recently for several types of de novo designed oligopeptides [2]. When placed in their respective solvents, these oligopeptides aggregate into long semi-flexible tapes well below millimole concentrations. This self-assembly into tapes is often accompanied by LC ordering at higher concentrations for some oligopeptides. Formation of higher ordered self-assembled structures has also been observed [1, 2, 3].

The sheets are typically very long (up to several micrometers length at high enough concentrations). In na-

Fig. 1 The "stitching" of chain protein molecules in the β-config-uration by hydrogen bonds. The hydrogen interactions are indicat-ed by *dotted lines*

tive proteins, short β-sheets or barrels are widely dis-played, and there are also extended β-sheets in silk [4, 7, 8]. Where identical short peptide molecules are involved, these self-assembling sheets may be very long, i.e. tape-like [1, 2, 3].

Mathematical models previously used for the study of ordered protein structures formed by comparatively short peptide molecules are based on the methods of confor-mational analysis. These models include many energetic and geometrical parameters taking into account the real local structure of proteins and require laborious comput-er calculations of the conformational energy [8]. These difficulties increase significantly with the description of the statistical and dynamical characteristics of sufficient-ly extended structures consisting of long polypeptide molecules.

Therefore, a modern theoretical investigation of pro-tein structures requires not only the methods of confor-mational analysis, but also the development of physical models of these structures. One such well-known single-chain model of DNA is the model of a flexible elastic rod with bending and twist rigidity [7]. Recently the classical isodesmic one-dimensional model for equilibri-um polymerization was generalized in order to describe self-assembly in peptide solutions forming β-sheet tapes and stacks of various thickness: double tapes and fibrils [1, 2, 3]. The broad class of discrete and continuous sta-tistical multiparticle models with orientational interac-tions and fixed positional order (the Ising, plane-rotator and Heisenberg models) used early in the statistics of ferromagnetic systems [9] was proposed for the study of the conformational properties of single- [10, 11] or multichain [12, 13, 14] ordered polymer systems includ-ing liquid-crystalline (LC) ones.

The special study of two-dimensional statistical multichain models may be of practical interest in con-nection with the properties of thin films, membranes and other surface or interfacial mesophase polymeric struc-tures. Researchers have already tried to exploit the self-assembling properties of natural peptides by designing

new oligomeric peptide chains, which form macroscopic membranes [2]. The orientation of dipole groups in LC films induced by the electromagnetic field of a laser beam has been used for recording information [15].

The behavior of low-dimensional multiparticle sys-tems in physics has always been of special interest. Char-acteristic types of order and singularities at phase transi-tions in two-dimensional systems with orientational or other interactions differ essentially from those in one- and three-dimensional systems [9, 16, 17]. In particular, in the isotropic plane-rotator model (Vaks–Larkin model [17]) the power law of decay of the orientational correlation function was obtained. This decay is slower than the ex-ponential one for one-dimensional systems with nearest-neighbor interactions. On the other hand, a long-range or-der may not exist in the corresponding two-dimensional systems, in contrast to three-dimensional systems.

In the present paper we propose a sufficiently simple two-dimensional plane-rotator model for describing long-scale conformational properties of sufficiently long chains in self-assembled multichain β-sheets in proteins. This paper is organized as follows. At first a general de-scription of the anisotropic version of the plane-rotator model with only two energetic parameters is considered. At low temperatures or strong intra- and interchain inter-actions the harmonic (quadratic) approximation for the potential energy is used. The behavior of intra- and inter-chain orientational correlation functions and mean-square dimensions of the chains in the β-sheet are inves-tigated in detail. The conditions of existence of different conformations, i.e. extended ones, Gaussian coils and other states with intermediate types of behavior of mean-square chain dimensions, are obtained. In conclusion, long-scale statistical properties of two-dimensional poly-mer systems with intra- and interchain nearest-neighbor interactions and those for polymer chains with excluded volume effects are compared.

General description of the model

In the present paper a two-dimensional multichain sheet is considered as an intermediate type of macromolecular aggregates (between isolated polymer chains and three-dimensional multichain ensembles). Macromolecules are essentially planar when compared to three-dimensional multichain layers, in which parallel chains with a finite contour length are arranged perpendicular to this plane, e.g. in smectic LC structures [18].

We use a simplified two-dimensional multichain mod-el of an infinitely extended sheet composed of chains with rigid elements – plane rotators with the average length l, simulating bonds of amino acid residues in pro-teins. In this model the N elements, arranged along one "zigzag" line in a selected "longitudinal" curvilinear di-rection (**l** on Fig. 2), form an entire polypeptide chain. The index **n** characterizes the position of an element in a chain ($n=1...N$) and the index **m** numbers a chain itself in the other, "transverse" curvilinear direction (**t** on the
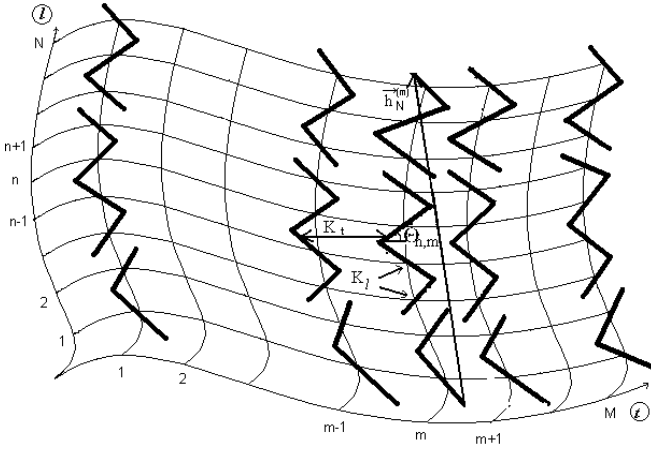
Fig. 2 The two-dimensional model of the anisotropic multichain sheet, composed of M chains with N rigid elements. The energetic constant K_l characterizes the bending rigidity of a chain, and K_t is the constant of interchain interactions. The angle $\theta_{n,m}$ determines an orientation of the n-th element in the m-th chain respect to a selected in the plane $(\mathbf{l}, \mathbf{t})$ direction
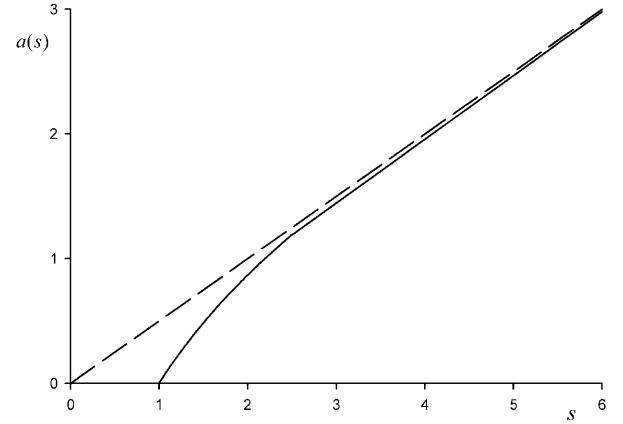
Fig. 3 The dimensionless persistent length $a(s)=2K_l(s)/k_BT$ (in a number of chain elements) versus the number of elements $s=A/l$ in the statistical segment. The asymptotic dependence $a(s)$ is represented by the *dashed line*

Fig. 2) of the sheet: $m=1...M$. This multichain model is similar to a two-dimensional version of the "curvilinear crystal" [4] used for describing of mesophase ordered polymer systems with extended chains (e.g. in cellulose films with LC ordering type etc.). However, we assume in this model that a local mobility may occur in multichain β-sheets as in LC structures, [18] i.e. we consider the case of higher temperatures than in model structures of the crystalline domains [8].

The energy of intra- and interchain interactions of nearest-neighbor chain elements of plane chains is represented in the form

$$H = -K_l \cdot \sum_{n,m} \cos(\theta_{n,m} - \theta_{n-1,m})$$
$$\qquad -K_t \cdot \sum_{n,m} \cos(\theta_{n,m} - \theta_{n,m-1}) \qquad (1)$$

where $\theta_{n,m}$ is the rotation angle of the n-th element in the m-th chain in the $(\mathbf{l}, \mathbf{t})$ plane (Fig. 2). It is difficult to relate $\theta_{n,m}$ to standard protein angles in the Ramachandran method of analysis of local chain conformations [5, 6, 7, 8]. The rotation angle $\theta_{n,m}$ in this simplified model characterizes average orientations of links (e.g. amino acid residues) in the plane of the sheet.

The energetic constants in Eq. (1) describe interactions of chain elements in the "longitudinal" direction, K_l (along the chains), and "transverse" direction, K_t (between neighboring chains, Fig. 2). These constants may have a different origin and physical meaning and differ strongly each from one another. Let us carry out some estimation of these constants.

For homogeneous polymer chains the intrachain energetic constant K_l is the same for all elements of the chains. The value of K_l determines the mean cosine of angle between neighboring chain elements (bonds or monomer units)

$$\eta_0 = \exp(-k_B T/2K_l) \qquad (2)$$

which is an important conformational characteristic of the isolated semirigid chain ($K_t=0$) [19]. In Eq. (2) T is the absolute temperature and k_B is the Boltzmann constant.

For heterogeneous macromolecules of proteins the constant K_l and correspondingly the parameter η_0 depend on the type of interacting neighboring amino acid residues. However, in the frame of the given model we propose that the effective energetic constant K_l is the same for all chain elements, as for homogeneous polymer chains. The corresponding quantity η_0 is connected with the length A of the statistical Kuhn segment (or persistent length) of the real heterogeneous polypeptide chain by the well-known expression in theory of conformational properties of homogeneous chains [19]

$$\frac{A}{l} = S = \frac{1+\eta_0}{1-\eta_0} \qquad (3)$$

In Eq. (3) the dimensionless parameter $s=A/l$ determines the number of rigid elements in the statistical Kuhn segment. The increase of bending rigidity of the chain (i.e. the length A of the statistical Kuhn segment) leads to the increase of the intrachain constant K_l (see Fig. 3). The length of a rigid element (bond or monomer unit) for proteins, nucleic acids and organic molecules l~$(1.5–5)$ Å [20, 21]. For flexible long polymer chains the value A~$(10–15)$ Å, [21] and the value $s=A/l$≤10. Then, the quantity η_0≤0.7, and correspondingly the constant K_l≤$2k_BT$~10^{-20} J (at $T=300$ K). On the contrary, for very stiff peptide chains A≥100 Å, [21] and the values s≥70, η_0≥0.9, and correspondingly, K_l≥$20k_BT$~10^{-19} J (at the same temperature $T=300$ K). In this case the ratio $k_BT/2K_l$ ≪1 in the exponent of Eq. (2), and the Eq. (3) leads to the relation

$$K_l \sim (k_B \cdot T/4)s \qquad (4)$$

The energetic constants of the intrachain orientational interaction W_Θ for shorter molecules (proteins, nucleic acids and other rod-like organic molecules) were calculated in [20]. The values $W_\Theta \sim (60–140)$ kcal mol^{-1} result in the following estimation of the energy of orientational interaction per bond or monomer unit $K_l \sim 10^{-20}$ J, that is a quantity of the same order as for the corresponding intrachain constant for long polymer chains.

The effective energetic constant K_t in Eq. (1) describes interchain interactions caused by hydrogen and other interchain bonds or side interactions in real proteins, e.g. interactions of orientational type. Its value is also defined by the real chemical structure of the polymer. As is known, [4, 19] hydrogen bonds are directive. Their force depends strongly on an orientation of bonds relative to the backbone of a chain. These bonds usually lead to a stretching of chains like orientational interactions or excluded volume effects. Therefore, in spite of the different nature of hydrogen, orientational and other interactions, we propose in this model that they may cause analogous conformational effects and interchain ordering.

As far as the considered two-dimensional multichain model must describe condensed polymer systems (e.g. the β-sheet), the value of the constant K_t may not be obtained from statistical theory developed for weakly concentrated systems, e.g. from the known Onsager self-consistent mean-field approach [18]. Interchain interactions for polypeptide chains with polar groups are defined by interactions of dipoles in neighboring chains. The constant K_t of interchain interactions in polar peptides may be estimated from the energy of dipole–dipole interaction between two dipole moments μ placed in neighboring chains [22]

$$K_t = \frac{1}{4\pi\varepsilon_0} \frac{\mu^2}{r^3} \tag{5}$$

where the quantity r characterizes a mean distance between neighboring chains, and ε_0 is dielectric constant.

The dipole moment of a bond for flexible polymer chains is μ~10 D [21]. At r~10 Å, Eq. (5) gives the value $K_t \sim 10^{-20}$ J. On the contrary, at the same distance between dipoles (r~10 Å) in very stiff chains, e.g. in peptides $A \geq 100$ Å, μ~1 D [21] and the constant of interchain interactions $K_t \sim 10^{-21}$ J.

Both parallel and antiparallel β-sheets may be described in the framework of the model. The value of the constant K_t is related to the lattice constant in the transverse direction in the β-sheet. Therefore, the difference between describing parallel and antiparallel β-sheets in this model manifests itself in values of the energetic constant of interchain interactions (K_t). The lattice constant in the transverse direction in the antiparallel Pauling–Corey β-sheets is equal to twice the distance between polar polypeptide chains (r_1~10 Å) compared with parallel β-sheets (r_2~5 Å), the corresponding constants K_t in parallel and antiparallel β-sheets differ from one another ($r_1/r_2)^3$~8 times, in agreement with Eq. (5).

For nonpolar macromolecules, forming e.g. systems with a LC ordering, the value of the constant of interchain interactions K_t may be obtained from the modulus of elasticity in Frank–Oseen–Zocher elasticity theory [18]. It may be estimated roughly from the typical energy of intermolecular interactions in low-molecular liquid crystals W~2 kcal mol^{-1}, which gives the value K_t~10^{-20} J. This value is adjusted with the preceding estimation of the constant K_t obtained above for stiff chains with dipole groups. Therefore, the energetic constants K_l and K_t have the same order of magnitude for flexible polymer chains or differ significantly for stiff ones: $K_l = (10–10^2)K_t$. In the last case an anisotropy of intra- and interchain interactions exists, as in β-sheets formed from peptide chains with a sufficiently large bending rigidity.

The intra- and interchain orientational correlation functions

Many statistical properties of multichain polymer systems, e.g. mean-square dimensions of chains, the dipole moment of the system etc., are determined by long-range correlations between orientations of chain elements along and between chains (intra- and interchain correlations, respectively). The corresponding correlation functions were defined in references [12, 13, 14] as the mean cosine of angle between chain elements

$$g(p, q) \equiv \langle \cos(\theta_{n+p, m+q} - \theta_{n, m}) \rangle \tag{6}$$

The integers p and q in Eq. (6) determine the number of elements between chain elements in the longitudinal direction l, i.e. along chains, and in the transverse ones t, i.e. between chains, respectively (Fig. 2).

The averaging in Eq. (6) with the distribution function $\rho(H) \sim \exp[-H/k_B T]$ may be carried out in an exact analytic form at sufficiently strong interactions or at low temperatures. In both cases the relations $K_l/k_B T \gg 1$ and $K_t/k_B T \gg 1$ are valid. Therefore, the quadratic (harmonic) approximation for Eq. (1) may be used [16, 17]

$$H^* = \frac{1}{2} K_l \sum_{n,m} (\theta_{n,m} - \theta_{n-1,m})^2 + \frac{1}{2} K_t \sum_{n,m} (\theta_{n,m} - \theta_{n,m-1})^2 \tag{7}$$

Intrachain correlations

Let us consider at first the behavior of the intrachain correlation function

$$g_l(p) \equiv g(p, 0) = \langle \cos(\theta_{n+p, m} - \theta_{n, m}) \rangle \tag{8}$$

that will be necessary later for analysis of mean-square dimensions of the chains in the sheet. In the low-temperature approximation (7) the expression for the intrachain correlation function $g_l(p)$ is reduced to the form

$$g_l(p) = \exp\left[-\frac{k_B T}{2\pi K_l} \int_0^\pi \frac{(1 - \cos px)\,dx}{\sin\frac{x}{2}\sqrt{\sin^2\frac{x}{2} + \varepsilon}} \right] \tag{9}$$

Here $\varepsilon=K_t/K_l$ is the parameter of anisotropy of inter- and intrachain interactions. The behavior of the intrachain correlation function $g_l(p)$ may be described by the expression [13, 14]

$$g_l(p) \cong (1+\frac{p}{p^*})^{-\gamma} \qquad (10)$$

where the characteristic scale parameter $p^* = 1/\pi\sqrt{K_l/K_t}$. The exponent γ in Eq. (10) is equal to the ratio of the energy of thermal motion ($k_B T$) to the geometric mean of intra- and interchain interactions energies of chain elements (K_l and K_t, accordingly)

$$\gamma = k_B T/2\pi\sqrt{K_l K_t} \qquad (11)$$

At sufficiently small distances between chain elements ($p<p^*$) the right-hand side of Eq. (10) is represented as

$$\exp\left[-\gamma\cdot\left(1+\frac{p}{p^*}\right)\right] \sim \exp\left(-\gamma\cdot\frac{p}{p^*}\right)$$

and may be presented in the exponential form

$$g_{l0}(p) = \exp(-p/a) = \eta_0^p \qquad (12)$$

where the parameter $a=p^*/\gamma=2K_l/k_B T$ defines the corresponding dimensionless persistent length of chain $a=A/2l$ in the continuous model of a worm-like isolated chain ($K_t=0$) with length A of the statistical Kuhn segment. [19] At sufficiently large distances ($p>p^*$) between elements along the given chain the intrachain correlation function (10) diminishes to zero as $p\rightarrow\infty$ according to the power law

$$g_l(p) \simeq (p^*/p)^\gamma. \qquad (13)$$

At $K_l=K_t=K$ Eq. (13) is reduced to the power law with the exponent $\gamma=k_B T/2\pi K$. This result was obtained earlier for the isotropic plane-rotator model [17]. Therefore, for a chain in two-dimensional polymer systems with interchain interactions ($K_t\neq 0$) besides the persistent length (a) another characteristic parameter exists

$$p^* = 1/\pi\sqrt{K_l/K_t} = \sqrt{\varepsilon}/\pi \qquad (14)$$

The quantity $\varepsilon=K_t/K_l$ characterizes an anisotropy of intra- and interchain interactions. The parameter p^* defines the crossover of the intrachain correlation function $g_l(p)$ from the exponential law (12) to the power ones (13). At $p<p^*$ intrachain interactions are still prevailing over interchain ones, although the chain is included in the two-dimensional multichain sheet. In this sense p^* characterizes a manifestation of statistical properties of the individual chain. At $p>p^*$ intrachain correlations of chain elements decrease by the slower power law (13) due to interchain interactions in the multichain sheet.

Therefore, for rigid chains, e.g. for peptides, the parameter $p^*\gg 1$, i.e. interchain interactions are important at larger distances than those for flexible chains. For a chain with the given bending rigidity at sufficiently strong interchain interactions such that $\gamma=k_B T/2\pi\sqrt{K_l K_t} < 1$, the parameter $p^*=\gamma a<a$, i.e. interchain interactions occur

even at smaller distances than the persistent length a of an isolated (individual) chain. In this case the bending rigidity of a chain does not manifest itself even in the two-dimensional ensemble of such chains. On the contrary, in the limiting case $K_t\rightarrow 0$, when the parameters p*, $\gamma\rightarrow\infty$, Eq. (10) transforms to the exponential law (12) that is valid for all distances p between chain elements in the individual chain ($K_t=0$). In this case only the persistent length a determines intrachain correlations.

Interchain correlations

The interchain correlations for elements of different chains in the multichain sheet may be described by the function

$$g_t(q)\equiv g(0, q)=\langle\cos(\theta_{n, m+q}-\theta_{n, m})\rangle \qquad (15)$$

which determines the mean cosine of the angle between chain elements disposed at the distance q (in the number of elements) in the transverse to chains direction (between chains on Fig. 2, cf. Eq. (8)). In the low-temperature approximation (7) the interchain orientational correlation function $g_t(q)$ is defined by the expression

$$\begin{aligned} g_t(q) &= \exp\left[-\frac{k_B T}{2\pi K_t}\int_0^\pi\frac{(1-\cos qx)dx}{\sin\frac{x}{2}\sqrt{\sin^2\frac{x}{2}+\varepsilon^{-1}}}\right]\\ &\cong \left(1+\frac{q}{q^*}\right)^{-\gamma} \end{aligned} \qquad (16)$$

Equation (16) differs from Eq. (9) by the corresponding parameters in the exponent. It may be obtained from Eq. (9) by replacing $K_l\rightarrow K_t$ (see Table 1). The corresponding characteristic scale parameter $q^* = 1/\pi\sqrt{K_t/K_l}$ defines the crossover from the power law (16) of decrease of the interchain correlation function to an exponential one

$$g_{t0}(q) = \exp(-q/b) = \zeta_0^q \qquad (17)$$

where the quantity $b=2K_t/k_B T$. The parameter $\zeta_0=\exp(-1/b)$ in Eq. (17) has the meaning of mean cosine of the angle between elements in neighboring free-jointed chains ($q=1$, $K_l=0$), interacting with the same interchain constant K_t as in the considered model (see Table 1).

Thus, the asymptotic behavior of intra- and interchain correlation functions depends on both the parameter of anisotropy of inter- and intrachain interactions $\varepsilon=K_t/K_l$ and the exponent $\gamma = k_B T/2\pi\sqrt{K_l K_t}$ in the power laws (10) and (16) (see Table 1). As shown above, for stiff chains, e.g. for peptide chains, the relation $K_l\gg K_t$ is valid, and consequently, $p^*\gg q^*$. In such systems the crossover to the power law for long-range correlations between chain elements in the "transverse" direction (between chains) manifests itself at smaller distances than the corresponding crossover in the "longitudinal" direction (along the chains, Fig. 2).

Since the intra- and interchain correlation functions $g_l(p)$ and $g_t(q)$ slow down to zero as $p, q\rightarrow\infty$, it follows

Table 1 The intra- and inter-chain orientational correlation functions

	Intrachain correlation function	Interchain correlation function
The low-temperature behavior	$g_l(p) \cong \left(1 + \frac{p}{p^*}\right)^{-\gamma}$	$g_t(q) \cong \left(1 + \frac{q}{q^*}\right)^{-\gamma}$
The exponent $\gamma = k_B T / 2\pi \sqrt{K_l K_t}$		
Characteristic scale parameter	$p^* = 1/\pi \sqrt{K_l/K_t}$	$q^* = (1/\pi)\sqrt{K_t/K_l}$
The behavior at small distances	$\exp(-p/a) = \eta_0^p; p < p^*$	$\exp(-q/b) = \zeta_0^q; q < q^*$
The parameters	$\eta_0 = \exp(-1/a)$ $a = 2K_l/k_B T$	$\zeta_0 = \exp(-1/b)$ $b = 2K_t/k_B T$
The behavior at large distances	$(p^*/p)^\gamma; p > p^*$	$(q^*/q)^\gamma; q > q^*$

that a long-range orientational order in the considered multichain sheet is absent as in one-dimensional systems with only short-range interactions, [16, 17] e.g. in an individual linear chain ($K_t = 0$). This result corresponds to the general Mermin–Wagner theorem [23]. According to this theorem, the long-range order does not exist in one- or two-dimensional systems, where the energy of interaction of nearest-neighbor particles is invariant under a continuous transformation group of symmetry, e.g. the group of rotations in the plane, as in this model.

Thus, intra- and interchain correlations at sufficiently large distances between chain elements in the two-dimensional multichain sheet decrease by the power law, i.e. slowly compared with the isolated chain. The power law for the correlation function is a general statistical property of low-temperature behavior for most two-dimensional systems. Other typical examples include the continuum model of a nematic LC film in the single-constant approximation for the Frank–Oseen–Zocher potential [18], the magnetic with an easy magnetization axis, the Bose liquid, the isotropic plane-rotator model and two-dimensional Heisenberg ferromagnetic etc. [16, 17].

Mean-square dimensions of chains in the sheet

In this model the chain is defined as a finite "zigzag" sequence of N elements – rotators arranged along the selected "longitudinal" curvilinear direction of the infinitely extended sheet (**l** on Fig. 2). The end-to-end vector of the chain consisting of N rigid elements is given by the expression

$$\vec{h}_N = l \cdot \sum_{n=1}^{N} \vec{e}_n \tag{18}$$

where l is the length of a rigid element. The values $\vec{e}_n$ in Eq. (18) are unit vectors directed along corresponding chain elements. The mean-square dimensions $\langle h_N^2 \rangle$ of a chain are defined as follows

$$\langle h_N^2 \rangle = l^2 \sum_{n=1}^{N} \sum_{n'=1}^{N} \langle \vec{e}_n \cdot \vec{e}_{n'} \rangle \tag{19}$$

For sufficiently long chains, included in an infinitely extended sheet, the values $\langle \vec{e}_n \cdot \vec{e}_{n'} \rangle = \langle \cos(\theta_n - \theta_{n'}) \rangle$ in Eq. (19) depend only on the distance $p = |n - n'|$ between elements along the chains. The quantities $\langle \vec{e}_n \cdot \vec{e}_{n'} \rangle$ are defined by the intrachain function $g_l(p)$ for a selected chain (see Eq. (8), where the quantity $p = |n - n'|$). According to reference [19] Eq. (19) may be transformed to the form

$$\langle h_N^2 \rangle = l^2 \left[N + 2 \sum_{p=1}^{N-1} (N - p) g_l(p) \right] \tag{20}$$

For the model considered the function $g_l(p)$ may be described by the power law (10). The relation between the power and Γ function

$$(1/p)^\gamma = \int dx \, x^{\gamma-1} e^{-px} / \Gamma(\gamma) \tag{21}$$

allows the expression (21) to be reduced to the specific form

$$\langle h_N^2 \rangle = \int_0^\infty dx \rho(x) \langle h_N^2 \rangle_0 (x) \tag{22}$$

The function $\rho(x)$ in Eq. (22) is defined by the expression

$$\rho(x) = e^{-x} x^{\gamma-1} / \Gamma(x) \tag{23}$$

This function satisfies the normalization condition

$$\int_0^\infty dx \rho(x) = 1 \tag{24}$$

Thus, the function $\rho(x)$ has the meaning of the distribution function of the quantity $\langle h_N^2 \rangle_0 (x)$ as a function of the parameter x. Substituting Eqs. (10) and (21) in Eq. (20), we get the following expression for the quantity $\langle h_N^2 \rangle_0 (x)$ in Eq. (22)

$$\langle h_N^2 \rangle_0 (x) = l^2 \left[N \frac{1 + \tilde{\eta}(x)}{1 - \tilde{\eta}(x)} - \frac{2\tilde{\eta}(x)}{[1 - \tilde{\eta}(x)]^2} - \frac{2[\tilde{\eta}(x)]^{N+1}}{[1 - \tilde{\eta}(x)]^2} \right] \tag{25}$$

Equation (25) is similar to the well-known expression for the mean-square dimensions of an isolated chain in the theory of conformational properties of single chains [19]. The quantity $\tilde{\eta}(x)$ is analogous to the mean cosine of the angle between neighboring chain elements. It is connected with the corresponding quantity η_0 of an isolated chain ($K_t=0$, see Eq. (2)) by the relationship $\tilde{\eta}(x) = \eta_0^{x/\gamma}$, where the parameter γ is defined by Eq. (11). The function $\tilde{\eta}(x)$ may be represented in the form $\tilde{\eta}(x) = \exp[-1/\tilde{a}(x)]$ is the effective persistent length, and the parameter p^* is determined by Eq. (14). The quantity $\tilde{\eta}(x)$ is a monotonically decreasing function of x and the parameter of anisotropy $\varepsilon = K_t/K_l$, and $\tilde{\eta}(x) \to 1$ as $x \to 0$ (Fig. 4).

The expression (22) may be considered as a result of averaging the mean-square dimensions $\langle h_N^2 \rangle_0 (x)$ over an infinitely large ensemble of isolated chains, characterized by the exponential law (12) for intrachain correlations. However, the chains from this ensemble are different by the effective persistent length $\tilde{a}(x)$ and the corresponding bending rigidity $(\tilde{\eta}(x))$. This averaging is performed with the distribution function $\rho(x)$ defined by Eq. (23).

The behavior of the function $\rho(x)$ in Eq. (23) depends strongly on the value of the parameter γ (Fig. 5). At sufficiently low temperatures or strong intra- and interchain interactions of chains such that the relation $0 < \gamma \gg 1$ is valid (see Eq. (11)), the distribution function $\rho(x)$ behaves as $(1/x)^{1-\gamma}$. In this case the function $\rho(x)$ has a singularity at the point $x=0$, i.e. $\rho(0)=\infty$ (see Eq. (23)). Therefore, the main contribution to the value of $\langle h_N^2 \rangle$ in Eq. (22) is determined by $\rho(x)$ at small values x. This range corresponds to chains with a high effective bending rigidity $(\tilde{\eta}(x) = \eta_0^{x/\gamma} \sim 1$, see Fig. 4). Consequently, at $\gamma \gg 1$ the second and third terms in Equation (25) lead to the dependence $\langle h_N^2 \rangle_0 (x) \sim N^2$. In this case, according to Eq. (22), the mean-square dimensions $\langle h_N^2 \rangle \sim N^2$ correspond to sufficiently extended conformations of chains, characteristic for β-sheets. Of course, it is valid, if the contour length of the chain is smaller than the persistent length.

In contrast, at sufficiently high temperatures or small intra- and interchain interactions such that $\gamma \gg 1$, the singularity of the function $\rho(x)$ disappears (see Eqs. (11) and (23)), and $\rho(0)=0$ at the point $x=0$. In this case the dependence $\rho(x)$ already has a non-monotonic character, and Equation (22) is determined generally by values of x near the point $x_m = \gamma - 1$. The point x_m corresponds to a maximum of the function $\rho(x)$ (e.g. at $\gamma=1.5$ the value $x_m=0.5$, see Fig. 5). This maximum moves to large values of x if the parameter γ increases, i.e. interchain interactions decrease or the temperature increases (Eq. (11)). At sufficiently large values $x > p^*$ such that the corresponding effective persistent length $\tilde{a}(x) = p^*/x < 1$ and $\tilde{\eta}(x) \ll 1$ (Fig. 4), the main contribution for long chains ($N \gg \tilde{a}(x)$) is determined by the first term of Eq. (25), i.e. $\langle h_N^2 \rangle_0 (x) \sim N$. In this case, according to Eq. (22), the mean-square dimensions $\langle h_N^2 \rangle \sim N$. This behavior corre-
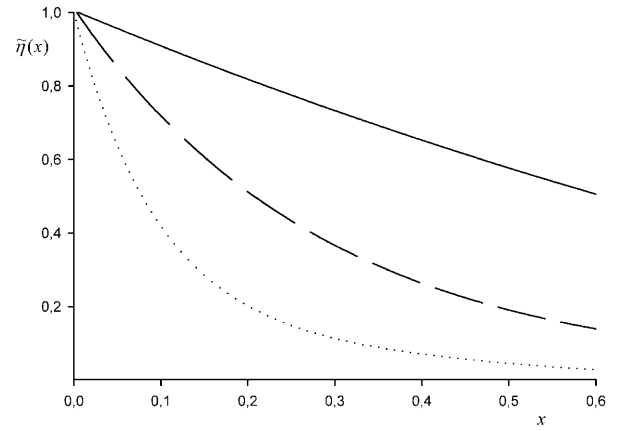

Fig. 4 The mean cosine $\tilde{\eta}(x)$ of the angle between neighboring elements of the chain with the effective dimensionless persistent length $\tilde{a}(x)$ versus the quantity x at different values of the parameter of anisotropy $\varepsilon = K_t/K_l = 10$ (*solid line*), 1 (*dashed line*), 0.1 (*dotted line*)
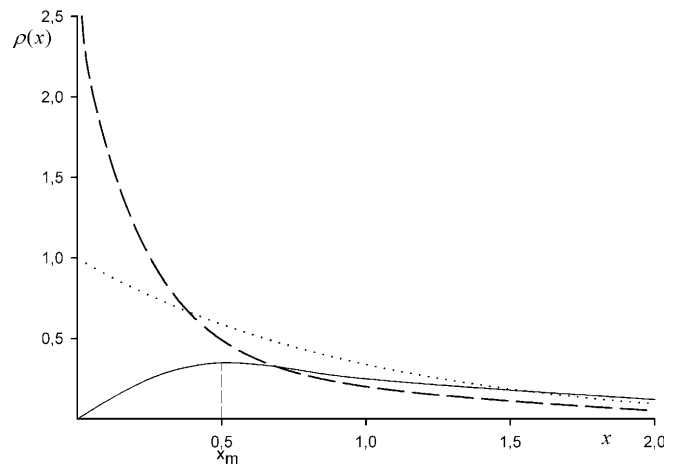

Fig. 5 The distribution function $\rho(x)$ versus the quantity x for different values of the parameter $\gamma=0.5$ (*dashed line*), 1 (*dotted line*), 1.5 (*solid line*)

sponds to Gaussian chains, characteristic for long flexible chains in the conformation of random coils.

The complete dependence of the mean-square dimensions $\langle h_N^2 \rangle$ in Eq. (22) on the number of chain elements N for different values of the parameter g may be obtained on the basis of Eqs. (22)–(25). For sufficiently large values of N the asymptotic expression is obtained

$$\langle h_N^2 \rangle / l^2 \simeq C(\eta_0, \gamma) N^{2\nu(\gamma)} + R_N(\gamma) \qquad (26)$$

Therefore, in the general case the dependence of the quantity $\langle h_N^2 \rangle$ on the number of chain elements N is non-Gaussian. This property is a consequence of the power law (10) for intrachain correlations. A general form of the dependence of the characteristic exponent $\nu(\gamma)$, the multiplier $C(\eta_0, \gamma)$, and the remainder $R_N(\gamma)$ in Eq. (26) on the parameter γ was discussed in reference [14] The different ranges of behavior of the mean-square dimensions $\langle h_N^2 \rangle = f(N)$ are shown in Table 2.

Table 2 The behavior of mean-square dimensions $\langle h_N^2 \rangle$ of chains depending on the value of the parameter $\gamma = k_B T / 2\pi \sqrt{K_l K_t}$

The value of parameter γ	The asymptotic dependence of $\langle h_N^2 \rangle$ on N
$\gamma \gg 1$	$2a \cdot N \cdot l^2 + o\left(\eta_0^{N+1}\right)$
$\gamma \geq 2$	$2a \cdot \gamma \cdot N \cdot l^2/(\gamma-1)+o(p^*/N)$
$\gamma = 2$	$2(p^*l)^2[(N/p^*)-\ln(N/p^*)]$
$1<\gamma<2$	$\frac{2p^* \cdot l^2 \cdot \gamma}{\gamma-1}N - O(N/p^*)^\delta, 0 < \delta = 2-\gamma < 1$
$\gamma = 1$	$2(p^*l)^2(N/p^*)[\ln(N/p^*)-1]$
$\gamma < 1$	$l^2 N^{2-\gamma}+O(N/p^*)$

Fig. 6 The quantity $G(N) = \ln\left\lfloor\langle h_N^2 \rangle/l^2\right\rfloor$ versus $\ln N$ for chains with the given bending rigidity ($a=2K_l/k_B T=0.1$) at different values of the parameter $\gamma=0$ (*solid line*), 0.5 (*dotted line*), 1 (*long dashed line*), 1.5 (*medium dashed line*), ∞ (*dash–dot–dot line*). The cases $\gamma=0$ and $\gamma=\infty$ correspond to systems of fully extended chains and isolated Gaussian chains, respectively

In the low-temperature approximation (at $\gamma \gg 1$) the mean-square dimensions of the chain are represented in the form

$$\langle h_N^2 \rangle/l^2 \simeq N^{2-\gamma}+O(N/p^*) \quad (\gamma << 1). \quad (27)$$

In this range the following relation between the exponents ν and γ is valid (cf. Eqs. (26) and (27))

$$\nu \simeq 1 - \gamma/2. \quad (28)$$

On changing the parameter γ from 0 to 1, the quantity ν changes from 1 to 0.5. The value $\nu=1$ corresponds to fully extended chains in the state of absolute long-range orientational order (at $T=0$). The value $\nu=0.5$ corresponds to conformations with Gaussian coils.

At $\gamma=1$ there is a break in the function $\nu(\gamma)$ in Eq. (26) in the thermodynamic limit as $N\rightarrow\infty$ (see Fig. 7). In this case the quantity $\langle h_N^2 \rangle$ is defined by the following expression (see Table 2)

$$\langle h_N^2 \rangle/l^2 \simeq 2(p^*)^2(N/p^*)\left[\ln(N/p^*) - 1\right]. \quad (29)$$

Using the method of high temperature expansion [13] for values $\gamma \gg 1$, we get the expression for mean-square dimensions

$$\langle h_N^2 \rangle/l^2 \simeq 2aN\gamma/(\gamma-1)+o(p^*/N) \quad (30)$$

All asymptotic dependencies for the quantity $\langle h_N^2 \rangle$ in Table 2 are valid at sufficiently large N such that $N \geq p^*$. For small values $N \leq p^*$ the corresponding expression for $\langle h_N^2 \rangle$ is sufficiently complicated [13]. In this case the relation between the exponents ν and γ may not be described by Eq. (28). At values $N \leq p^*$ we assume that the power dependence (26) also holds; however, the exponent ν is a function of the parameter γ and N. Then, the exponent ν may be defined from the slope of the plot $\langle h_N^2 \rangle = f(N)$ drawn on a log–log scale (Fig. 6). The dependences of the exponent $\nu(\gamma)$ at different finite values of N are represented in Fig. 7.

Conclusion

Therefore, the proposed "quasi-crystalline" multichain model of β-sheets may be applied not only to a very spe-
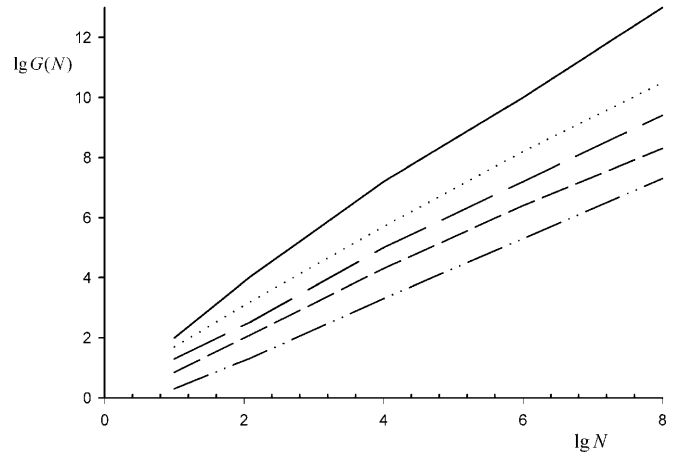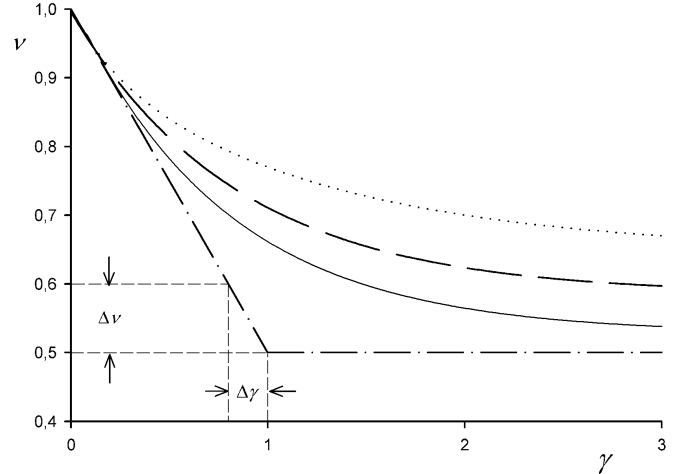

Fig. 7 The logarithmical exponent $\nu = \lim_{N\rightarrow\infty}\left[\ln\langle h_N^2 \rangle/\ln N^2\right]$ versus the ratio γ for the chains with the given bending rigidity ($a=2K_l/k_B T=1$) and the number of chain elements: $N=10^2$ (dotted line), 10^4 (dashed line), 10^6 (solid line), ∞ (dotted-dashed line). The value $\nu=0.5$ is related to isolated Gaussian chains. The intervals $\Delta\nu$ and $\Delta\gamma$ show the ranges of change of the exponents ν and γ, where an analogy of the two-dimensional systems with orientational interactions and systems with excluded volume effects exists [14]

cial case of protein structure, namely extended conformations of chains. In the considered harmonic approximation (7) polypeptide chains may exist in different conformations: from extended ones, as in β-sheets (at $0 \leq \gamma \leq 1$), to random Gaussian coils (at $\gamma \geq 2$). In the range $\gamma<1$ the mean-square dimensions of the chain in the sheet have an non-Gaussian behavior, i.e. $\langle h_N^2 \rangle \simeq N^{2\nu(\gamma)}$ ($0.5<\nu<1$). In this range interchain interactions are strong enough that a power law (13) for the intrachain correlation function exists.

There exists another important class of polymer systems with a power-law decrease of the intrachain correlation function and correspondingly with a non-Gaussian dependence of the mean-square dimensions of the chain. The power law occurs in chains with excluded volume effects [19, 24]. As first pointed out by Flory, [19] the power law for intrachain correlations exists even for individual isolated chains due to long-range intrachain steric interactions, i.e. interactions of the elements distant along the chain.

In the considered multichain model the power law (13) manifests itself as a consequence of sufficiently strong intra- and interchain nearest-neighbor interactions of chain elements in the corresponding directions of the sheet (Fig. 2). In spite of the different physical meaning of parameters in the single-chain theory of excluded volume effects and the considered two-dimensional multichain model, the existence of the asymptotic power law in these systems allows us to compare their statistical properties. The comparison of these systems was considered and discussed in reference [14]. Therefore, some effects of intra- and interchain interactions (steric, orientational etc.) in the low-temperature range of behavior of polymer systems, e.g. in polypeptides, may be described by a harmonic potential (7) of the orientational type.

We might include in addition a term with a small constant in Eq. (1) that will take into account non-nearest-neighbor interactions of chain elements. This term must lead to inherent twist effects for chains in the β-sheet (formed, e.g. by any chiral residues). However, it may be shown that the quadratic (harmonic) approximation (7) used in this paper leads only to a renormalization of the bending rigidity constant (K_l).

As stated above, there is a problem of more detailed consideration of the crossover of the intra- and interchain orientational correlation functions from the exponential law to the power one. Such a crossover is usually related to the existence of a phase transition [17, 25]. This transition may be considered strongly only in the high-temperature approach. [16] At higher temperatures or weak interactions one must also take into account anharmonic corrections to the potential (7). These corrections may be made by the methods used for analysis of the high-temperature behavior of isotropic low-molecular systems, e.g. by the variation method in the theory of anharmonic crystals or by the renormalization group treatment in the theory of liquid crystals [17]. By drawing on the correspondence between statistical properties of two-dimensional polymer systems with orientational interactions and those for polymeric systems with excluded volume effects it is hoped that a similar clarification can be achieved by means of the theory the well-known "coil-helices" or "coil-globule" transitions, [26] developed in detail for systems with excluded volume effects, e.g. for peptides.

The other interesting problem is a generalization of the proposed approach for describing statistical properties of chains in self-assembled peptide three-dimensional ordered structures, e.g. stacks composed of β-sheet tapes of various thickness: double tapes, fibrils (several double tapes stacked together) etc. [1, 2, 3] The generalization of the harmonic approach (7) to three-dimensional ordered systems might allow us to obtain some twist effects of β-sheets similar to those in chiral (cholesteric) LC structures. However, these problems will be considered in detail in the future with the special model of interacting sheets, composed of plane chains [27].

References

1. Aggely A, Bell M, Boden N, Keen JN, McLeish TCB, Nyrkova IA, Radford SE, Semenov AN (1997) J Mater Chem 7:1135
2. Aggely A, Bell M, Boden N et al. (1997) Nature 386:259
3. Nyrkova IA, Semenov AN, Joanny JF (1997) J Phys II France 7:847
4. Vainstein BK (1963) Difraction of X-rays on chain molecules. Academy of Science of SU, Moscow
5. Branden C, Tooze J (1991) Introduction to protein structure. Garland, New York London
6. Creighton TE (1993) Proteins. Structures and molecular properties. Freeman and Co, New York
7. Rubin AB (1987) Biophysics. Vishaja Shkola, Moscow
8. Fossey SA, Nemethy G, Gibson KD, Scheraga HA (1991) Biopolymers 31:1529
9. Ziman JM (1982) Models of disorder. Cambridge University Press, Cambridge
10. Volkenstein MV (1963) Configurational statistics of polymer chains. Interscience, New York
11. Nikomarov ES, Obukhov SP (1981) Zh Eksp Teor Fiz 80:650
12. Gotlib YY, Maximov AV (1992) Polym Sci Russ 34:902
13. Gotlib YY, Maximov AV, Maximova OG (1996) Polym Sci Russ 38:650
14. Maximov AV, Gotlib YY, Maximova OG, Egorova EM (1999) Polym Sci Russ 41:732
15. Shibaev VP, Kostromin SG, Plate NA et al. (1983) Polym Comm 24:363
16. Landau LD, Lifshits EM (1980) Statistical physics. Pergamon, New York
17. Patashinsky AS, Pokrovsky VL (1975) Fluctuation theory of phase transitions. Nauka, Moscow
18. De Gennes PG, Prost J (1995) The physics of liquid crystals. Clarendon Press, Oxford
19. Flory P (1969) Statistical mechanics of chain molecules. Interscience, New York
20. Cornell WD et al. (1995) J Am Chem Soc 117:5179
21. Zvetkov VN (1986) Stiffness-chain polymer molecules. Nauka, Leningrad
22. Buckingham E (1978) The basis of theory of intermolecular forces. The application to small molecules. In: intermolecular interactions: from diatomics to biopolymers. Wiley, New York
23. Mermin N, Wagner H (1966) Phys Rev Lett 17:1133
24. Domb C (1969) Adv Chem Phys 15:229
25. Ruelle D (1969) Statistical Mechanics. Rigorous Results. Benjamin, New York
26. Peliti L (1988) Statistics of polymers and self-avoiding walk. In: Fractals in physics. Elsevier, New York
27. Maximov AV, Maximova OG (1999) Applied Mechanics & Engineering 4:79