

# Computer simulation of polypeptides in a confinement

Andrzej Sikorski · Piotr Romiszowski

Received: 30 November 2005 / Accepted: 27 June 2006 / Published online: 15 September 2006  
© Springer-Verlag 2006

**Abstract** A coarse-grained model of polypeptide chains confined in a slit formed by two parallel impenetrable surfaces was studied. The chains were flexible heteropolymers (polypeptides) built of two kinds of united atoms—hydrophobic and hydrophilic. The positions of the united atoms were restricted to the vertices of a [310] lattice. The force field consisted of a rigorous excluded volume, a long-distance potential between a pair of amino-acid residues and a local preference for forming secondary structure (helices). The properties of the chains were studied at a wide range of temperatures from good to bad solvent conditions. Monte-Carlo simulations were carried out using the algorithm based on the chain's local changes of conformation and employing the Replica Exchange technique. The influence of the chain length, the distances between the confining surfaces, the temperature and the force field on the dimension and the structure of chains were studied. It was shown that the presence of the confinement chain complicates the process of the chain collapse to low-temperature structures. For some conditions, one can find a rapid decrease of chain size and a second transition indicated by the rapid decrease of the total energy of the system.

**Keywords** Coil-to-globule transition ·  $\alpha$ -helical globular proteins · Lattice models · Molecular confinement · Monte Carlo method · Replica Exchange method

Proceedings of “Modeling Interactions in Biomolecules II,” Prague, September 5th–9th, 2005.

A. Sikorski (✉) · P. Romiszowski  
Department of Chemistry, University of Warsaw,  
Pasteura 1,  
02-093 Warsaw, Poland  
e-mail: sikorski@chem.uw.edu.pl

## Introduction

The question of the folding transition in globular proteins is still far from being understood in spite of extensive experimental and theoretical work [1, 2]. Protein folding in a solution is rather a complicated phenomenon because of the chains' internal complexity: polypeptides consisting of thousands of atoms can adopt an enormous number of conformations but under certain conditions adopt only one conformation known as the native state. The folding of protein chains usually takes place in a crowded environment because of the presence of other biomolecules in cells. This kind of environment must restrict the chains' conformational states, especially expanded ones, which are characteristic for unfolded states, and thus, to stabilize the native state. The formation of protein aggregates is also affected by the effect of crowding [3, 4]. Recent experiments were able to encapsulate a protein in pores of silica glass [5], between layers of  $\alpha$ -zirconium [6], in a gel [7] and in a reversed micelle [8]. All confined proteins were found to be more stable and the melting temperatures appear to be considerably higher.

The problem of the influence of confinement on the properties of proteins and polymers is also interesting from the theoretical point of view. Simple theoretical considerations concerning the influence of the confinement on the stability of proteins in cages were given by Zhou and Dill [9]. Minton calculated changes in the reactivity of macromolecules using hard-particle partition theory [10]. Kinjo and Takada developed a density-functional theory of proteins in crowded solutions [11]. Recent computer simulations were focused on the stability of folded chains and the changes of the folding temperatures and kinetics of folding. Betancourt and Thirumalai [12], studied the influence of chaperonins of *Escherichia coli* GroEL on

the folding rates, Klimov et al. [13] simulated the folding of  $\beta$ -hairpin proteins in a spherical pore, and Takada et al. [14] simulated similar models of proteins confined in a cylindrical cage by means of the Brownian Dynamics. All these simulations concerned very simple Go-like models of proteins. The above-mentioned theoretical considerations and computer simulations showed that this kind of inert confinement stabilized the fold due to the reduction of the entropy of the unfolded state.

In this work, we present a different approach to the study of polypeptide chains in confinement [15–17]. The confinement effect on the properties of the chain can be of importance in view of the thin layer and nanotechnology techniques. Also the ‘natural’ confinement (originating from the walls of a biological cell) can change, to some extent, the properties of macromolecules. This study is also a stepping-stone as well as the state of reference for the investigation of the ‘macromolecular crowding’ on the properties of model chains. Contrary to other simulation studies based usually on Go-like models, we did not introduce any bias towards ‘native’ structures. Therefore, our chains do not have native-like structures but a series of conformations with low energy. The main goal of this work was the study of the interplay between the two potentials employed in the force field: a long-distance contact potential and a local potential that favors certain conformations characteristic for real  $\alpha$ -helical proteins. The balance between these two potentials was expected to be distorted by the presence of the confinement.

For this purpose we designed and developed a coarse and simplified model of heteropolymers. In this model all atomic details of polypeptide chains were suppressed and we employed a united-atom representation of chains, where chains were constructed of  $\alpha$ -carbons only. These united atoms were differentiated in order to mimic hydrophobic and hydrophilic properties of real amino-acid residues. We also introduced the lattice approximation of our model chains in order to study the phenomena that appear on larger time scales, like the entire process of the chain’s collapse. We used a Replica-Exchange Monte-Carlo simulation algorithm to determine the properties of the model chains at a wide range of temperatures, from these corresponding to a random coil state to those at which dense packed globules resembling globular proteins were formed. The bulk properties of such chains, the influence of the sequence pattern and the local potential on the secondary-structure formation and thermodynamic descrip-

tion have already been studied in detail for these models [15–17]. The influence of partial confinement on the polypeptides within the frame of similar models was also studied: a polypeptide chain was threaded through a hole of comparable size to that of the chain in an impenetrable wall [18]. Scaling properties of the translocation times and the structure of chains during the threading were determined and discussed.

## Model and methods

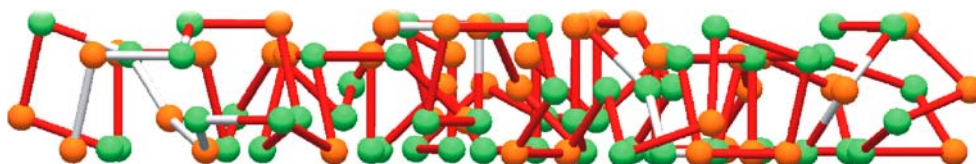
We studied model chains built of  $N=100$  residues. Each amino-acid residue in the chain was represented by a united atom located at the  $\alpha$ -carbon position. The united atoms were allowed to be put in vertices of a quasi-crystalline lattice only. For the purpose of the polypeptide chain representation, we chose lattices of a [310] type with the coordination number  $\zeta=90$  [15]. Chains constructed on this lattice resemble conformations of real proteins: one can represent real structures with an accuracy (RMSD) close to 0.7 Å. This best fit was found for a lattice unit equal to 1.22 Å. The chains were put into a slit formed by two parallel surfaces. These surfaces were impenetrable for amino-acid residues. No other interactions between the surfaces and the chains were assumed. The surfaces were parallel to the  $xy$  plane and periodic boundary conditions were imposed in the  $x$  and  $y$  directions. The scheme of the model system is shown in Fig. 1. The size of the Monte-Carlo cell was chosen large enough to avoid interactions of the chain with itself, and hence its edge was  $L=200$  lattice units (244 Å).

The model chain consisted of two kinds of residues: hydrophilic (polar, P) and hydrophobic (non-polar, H). The distinction between these two kinds of residues can be made by their interaction potential. This potential had the following form:

$$V_{ij} = \begin{cases} \varepsilon_{rep} & \text{for } r_{ij} < r_1 \\ \varepsilon_a & \text{for } r_1 \leq r_{ij} \leq r_2 \\ 0 & \text{for } r_{ij} > r_2 \end{cases} \quad (1)$$

where  $r_{ij}$  was a distance between a pair of residues,  $r_1=3$  (in lattice units) and  $r_2=5$  (in lattice units). The  $\varepsilon_a$  potential took the following values for three different types of residue–residue contacts:  $\varepsilon_{HH}=-2k_B T$  and  $\varepsilon_{PP}=-1k_B T$ ,  $\varepsilon_{HP}=0$ , where  $k_B=1$  is the Boltzmann constant [15, 19]. The rigorous excluded volume effect was softened in order

**Fig. 1** A scheme of a polypeptide chain built on a [310] lattice and confined to a slit formed by a pair of two parallel impenetrable surfaces



to help the chain to rearrange its conformation in a collapsed state and the repulsive potential between a pair of residues was assumed to be  $\epsilon_{rep}=10k_B T$ . This large value of the repulsion potential was introduced in order to avoid a larger number of conformations with self intersections of the chain, when the polypeptide is highly squeezed by the slit. The next component of the force field was the local potential  $\epsilon_{loc}$ , which expressed the tendency of the chain to form  $\alpha$ -helical structures. In the model chain,  $N$  united atoms were connected with  $N-1$  vectors  $\nu_i$ . In order to identify the presence of a helical structure in the chain, one has to check the mutual orientation of three consecutive vectors  $\nu_{i-1}$ ,  $\nu_i$ ,  $\nu_{i+1}$ . The following expression was calculated for this purpose:

$$r_{i-1,i+2}^{*2} = (\nu_{i-1} + \nu_i + \nu_{i+1})^2 \cdot \text{sign}((\nu_{i-1} \times \nu_i) \cdot \nu_{i+1}) \quad (2)$$

A right-handed  $\alpha$ -helical state corresponds to values of  $r_{i-1,i+2}^{*2}$  between 9 and 25 [15]. The appearance of a right-handed  $\alpha$ -helical state in the chain during the simulation run was associated with the energy loss  $\epsilon_{loc}$ .

The properties of the model chains were determined by means of Monte-Carlo simulations. We used the Replica Exchange-Monte-Carlo method (REMC) [17]. In this technique simulations of the  $m=20$  chains (replicas) are performed simultaneously at different temperatures [17, 20]. Each replica chain was a subject to classical Metropolis sampling and the changes of the chain conformations were made using a set of local motions. This set consisted of one-residue motions, two-residue motions and two-residue end reorientations. A Monte-Carlo time unit was defined as one attempt of each micromodification per residue. Each simulation run lasted for  $10^7$  time units and the exchanges between replicas were performed for every 10 time units. The results for different simulation runs were repeatable, with the standard error not exceeding 3%. The distribution of temperatures was linear and during the simulation the replicas at the neighboring temperatures were exchanged with the probability [17]:

$$p_{ij} = \min(1, \exp(-\Delta)) \quad (3)$$

$$\Delta = \left( \frac{1}{k_B T_i} - \frac{1}{k_B T_j} \right) \cdot (E_j - E_i) \quad (4)$$

where  $T_i$  and  $T_j$  are the temperatures of the  $i$ -th and  $j$ -th replica respectively while  $E_i$  and  $E_j$  are their total energies.

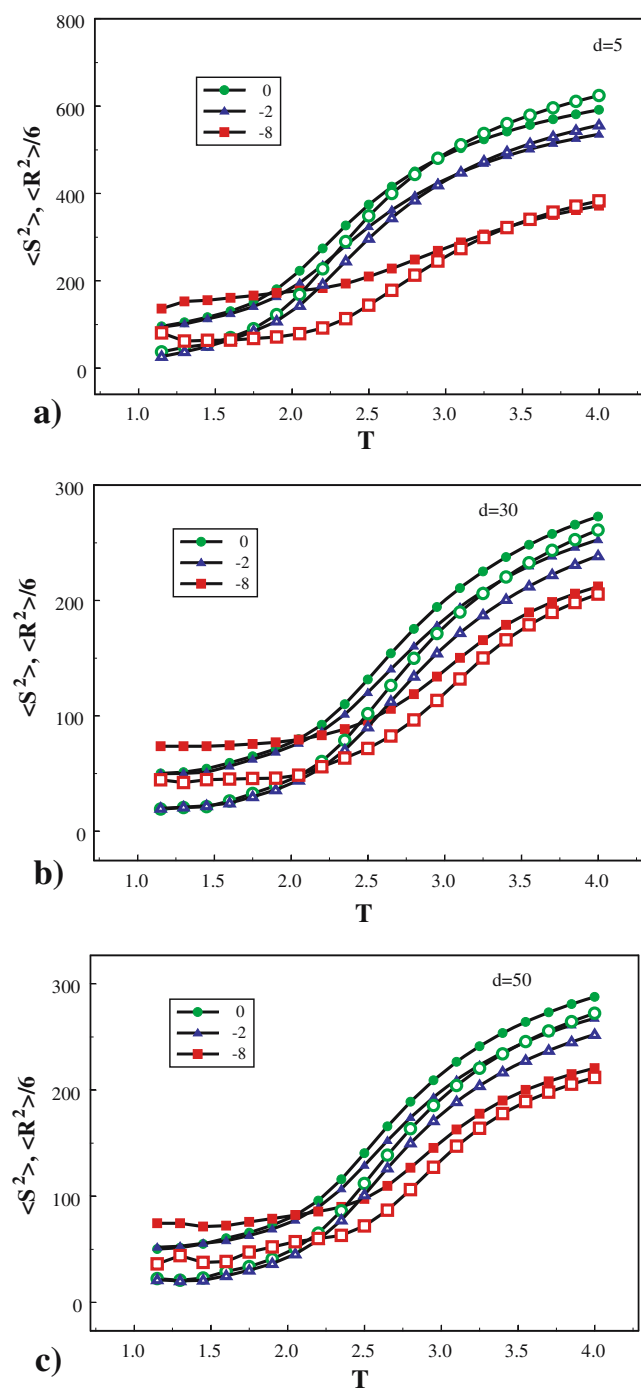
The Replica-Exchange simulation method was chosen since this algorithm is capable of crossing energy barriers, which is important for protein simulations, where the energy landscape is especially rugged [20]. The simulations started from a random conformation of the chain for each

replica. For each set of parameters 25–30 simulations were carried out using different starting conformations.

### Results and the discussion

We studied models of  $\alpha$ -helical proteins and the chains were therefore constructed of amino-acid residue sequences typical for this kind of protein: an idealized helical septet — HHPHPP — was repeated through the entire chain. The temperature was varied in the range from  $T=1$  to 4 according to the previous findings: one could find chain in a random coil state at temperatures above  $T=3$  (no helical potential) or above  $T=3.5$  (very strong helical potential) while collapsed (folded) structures of the chain were found to be stable at  $T < 2.5$  for all values of the helical potential [16, 17]. The helical potential  $\epsilon_{loc}$  was assumed to take the values 0,  $-2$ , and  $-8$  as it has been shown that for  $\epsilon_{loc}$  between 0 and  $-8$  one can observe an interplay between the two parts of potential in forming a dense globule and in forming secondary structures (helices) [17]. The size of the slit was changed between  $d=5$  and 50 in order to study highly squeezed chains and chains that are almost unaffected by the confining surfaces [21].

The location of the coil-to-globule transition of a polymer chain can usually be identified by the analysis of the changes of its size with temperature. However, one has to remember that in our model of confined chains the effect of the transition from three-dimensional to two-dimensional chain caused by the presence of the slit can influence the course of the coil-to-globule transition under certain conditions. The size of a chain is described as usually by the mean-squared radius of gyration  $\langle S^2 \rangle$  and the mean-squared end-to-end distance  $\langle R^2 \rangle$ . Figure 2a–c show the changes of these parameters with temperature for some values of the local potential and for some widths of the slit. Values of  $\langle R^2 \rangle / 6$  were added to the Fig. 2 in order to visualize the deviation from the theoretical values of the ratio  $\langle S^2 \rangle / \langle R^2 \rangle = 1/6$ . The size of the slit was chosen as  $d=50$  (the chain is not squeezed but it ‘feels’ the presence of the confinement),  $d=30$  (size of the slit is approximately equal to the free chain diameter) and  $d=5$  (the chain is highly squeezed). In all cases shown, the curves had the typical S-shape similar to that of free chain, although the rapid decrease of chain’s size is shifted towards lower temperatures [17]. This behavior suggests the presence of a transition from a coil to a more compact and dense globule for confined chains too. The increase of the local potential shifts the transition point towards higher temperatures. The same behavior was observed for unconfined chains, for which the coil-to-globule transition temperature also depended on the magnitude of the local potential [17]. The reason of this behavior will be discussed below. The

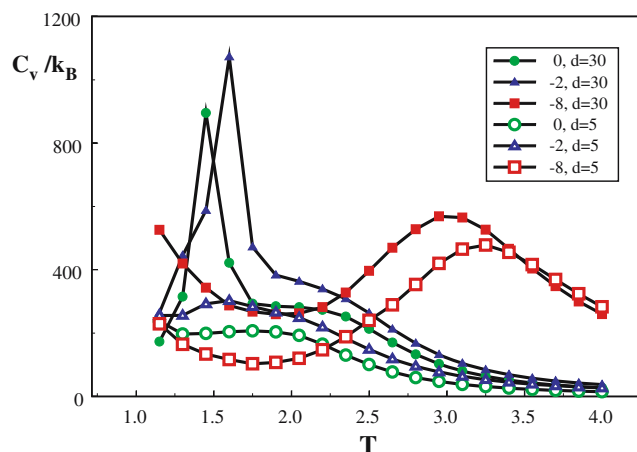


**Fig. 2** The mean-square radius of gyration  $\langle S^2 \rangle$  (solid symbols) and the mean-squared end-to-end vector  $\langle R^2 \rangle$  (open symbols) versus the temperature  $T$ . The values of  $\langle R^2 \rangle$  are divided by the factor 6. The values of local potential  $\epsilon_{loc}$  are given in the inset. The case of the slit  $d=5$  (a),  $d=30$  (b) and  $d=50$  (c)

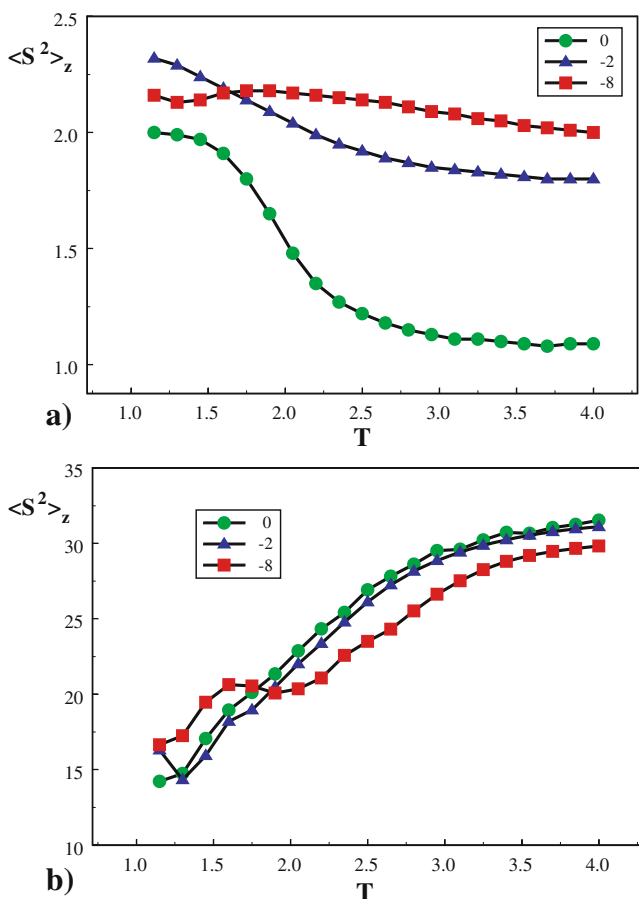
changes of the chain's size are much slower compared with those for the unconfined case. The influence of the confinement was more pronounced for narrow slits than for wider ones. The impact of the confinement on the size of the chain was more pronounced for narrow slits than for wider ones. The differences between the case of the chain

that is not squeezed ( $d=50$ ) and the chain in the slit of size two radii of gyration ( $d=30$ ) are rather small and thus we will discuss the latter case only. The above behavior of the model polypeptide is different from that found in a study of homopolymer molecules confined between adsorbing surfaces but under good solvent conditions. It was found that in the latter case macromolecules expand their dimensions for wider slits and remain almost constant for narrow slits [22].

The more precise location of the coil-to-globule transition is usually determined by analysis of the heat capacity of the system. The heat capacity of the model chain  $C_v/k_B$  was calculated as a variance of the mean total energy. In Fig. 3, we show the heat capacity as a function of the temperature  $T$  for various sizes of the slit and varied values of the local potential  $\epsilon_{loc}$ . One can observe that the curves are almost smooth with well-defined maxima on each. However, one can notice a significant difference in the location of peaks as the results for different values of the local potential are concerned. For all sizes of the slit under consideration, one can observe that for the helical potential  $\epsilon_{loc}=0$  and  $-2$  the peaks are located at lower temperatures (close to  $T=1.5$ ) while for  $\epsilon_{loc}=-8$  the peak appeared between  $T=3$  and  $3.3$ . The latter peak corresponds apparently to the rapid changes of the chain's size visible in Fig. 2. The shift of the folding transition was also found in real experiments [7] as well as in computer simulations of Go-like protein models [14]. The location of the peaks for lower helical potential is rather surprising. The size of the slit almost does not change the location of peaks, although for the smaller slit the peaks for  $\epsilon_{loc}=0$  and  $-2$  are considerably less pronounced and wider. The location of these peaks quite below the temperatures where the size of the chain diminishes rapidly is apparently caused by further rearrangement of the chain structure. During this rearrange-



**Fig. 3** The heat capacity  $C_v/k_B$  versus the temperature  $T$ . The size of the slit  $d$  and the values of local potential  $\epsilon_{loc}$  are given in the inset



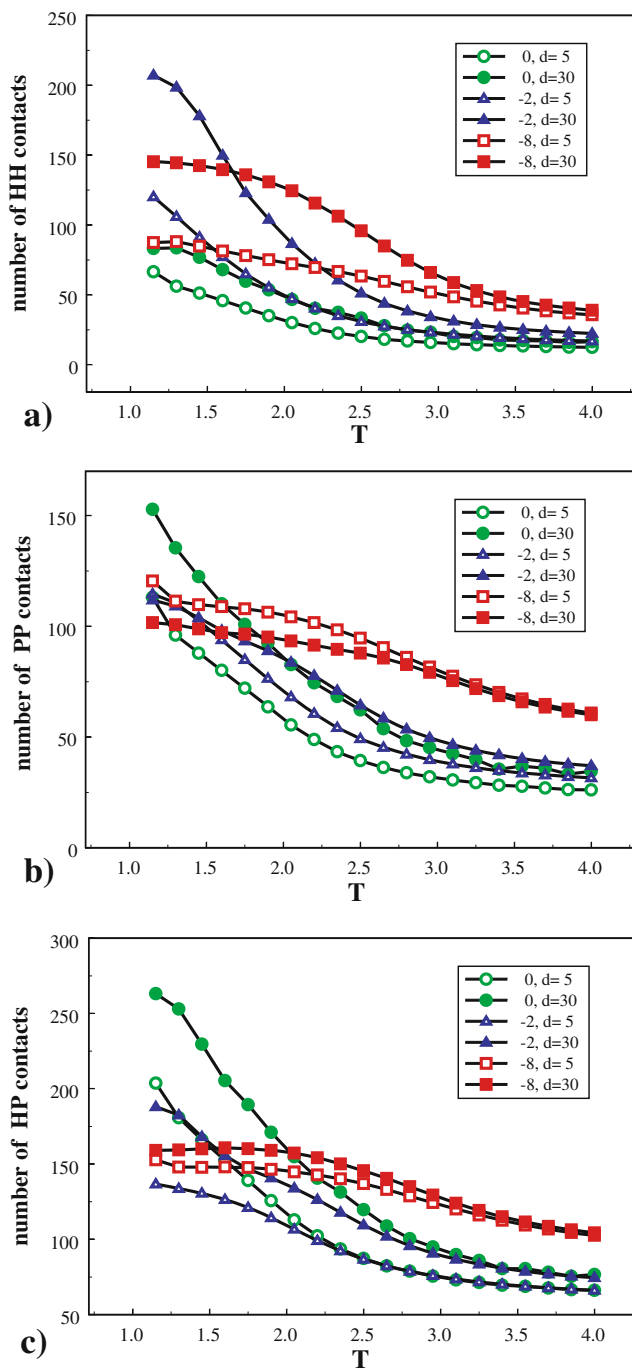
**Fig. 4** The perpendicular contribution to the radius of gyration  $\langle S^2 \rangle_z$  versus the temperature  $T$ . The values of local potential  $\epsilon_{loc}$  are given in the inset. The size of the slit  $d=5$  (a) and  $d=30$  (b)

ment, the global size of the chains did not change significantly.

The explanation of the behavior of the changes of chain’s size described above and the heat capacity can be elucidated by studying the  $z$ -contribution (perpendicular to the surfaces) to the radius of gyration. In Fig. 4a–b we show the changes of this component of the radius of gyration. One can easily observe the difference between the cases for the weak and strong local potentials. For the narrow slit, the size of a chain along the  $z$ -axis increases considerably for the weak local potential. This rapid increase occurs at a temperature close to  $T=1.5$  and thus corresponds to the heat-capacity peak. The changes of  $\langle S^2 \rangle_z$  for the strong local potential are less pronounced. For the wide slit, the changes of the size component perpendicular to unconfined dimensions are almost the same as for all values of the helical potential.

Further insight into the phenomenon of the coil-to-globule transition of confined polypeptides during the annealing process can be gained from analysis of the number of residue–residue contacts. Figure 5a–c show the mean number of hydrophobic–hydrophobic, hydrophilic–hydrophilic, and hydrophobic–hydrophilic contacts, respectively,

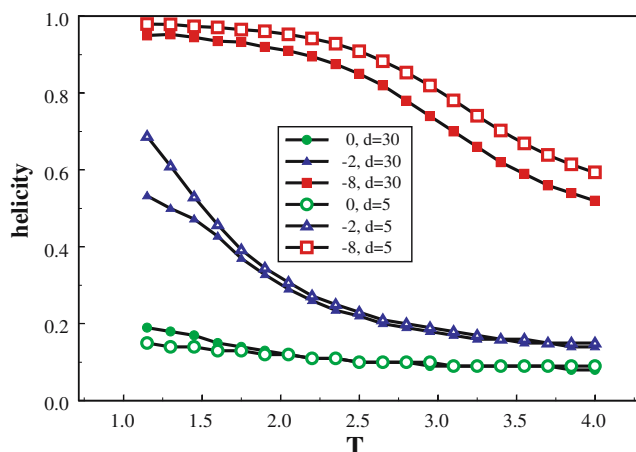
as a function of the temperature. One could expect that the number of contacts in general should increase during the annealing and this effect is observed in our simulations. However, there are quantitative differences between the number of contacts for chains with weak and strong helical potentials, even at high temperatures: for a weak helical potential the number of all types of contacts is considerably lower when compared to the case of the strong potential.



**Fig. 5** The mean number of HH (a), PP (b) and HP (c) contacts versus the temperature  $T$ . The size of the slit  $d$  and the values of local potential  $\epsilon_{loc}$  are given in the inset

For chains with a weak helical potential, the increase of the number of  $HH$  contacts is the most pronounced, while for the second case the process of chain collapse competes with the stronger preference to form helical conformations. The increase of the number of contacts for  $\epsilon_{loc}=0$  corresponds to the increase of  $\langle S^2 \rangle$  and to the location of  $C_v$  peak. Diminishing the slit's size reduced the number of contacts in all cases. One can observe that, for the weak helical potential, the most rapid increase in the number of contacts is located near the corresponding heat-capacity peak (see Fig. 3).

The last question concerned the influence of the confinement on the formation of  $\alpha$ -helices in polypeptide chains. It was previously shown for unconfined chains that in the absence of specific longrange interactions, low-temperature states are not unique and, moreover, helical fragments in formed dense globules are quite unstable [2, 15, 16]. In Fig. 6 the changes of the mean fraction of  $\alpha$ -helical states with the temperature  $T$  are shown for some widths of the slit. One has to remember that the annealing of model polypeptide chains causes an increase in the number of helical states even if the helical potential is turned off, because the compact structure of the chains generates some secondary structure [2, 15, 16]. The same effect was found for the confined fully flexible chains ( $\epsilon_{loc}=0$ ), where for both confinements the helicity increases slightly from 0.14 to 0.16 ( $d=5$ ) or from 0.15 to 0.19 ( $d=30$ ). For chains with non-zero local potential, the increase of the number of helical states is considerably higher. For a weak helical potential  $\epsilon_{loc}=-2$  the increase of the helicity is more pronounced: from 0.2 to 0.6, but the influence of the confinement is also small. For chains with a strong local potential ( $\epsilon_{loc}=-8$ ) the differences in helicity for both confinements can be found in the entire range of temperatures but these differences are also rather small.



**Fig. 6** The mean fraction of  $\alpha$ -helical states versus the temperature  $T$ . The size of the slit  $d$  and the values of local potential  $\epsilon_{loc}$  are given in the inset

Therefore, the influence of confinement on the presence of secondary structures is significant only for model chains with a moderate value of the helical potential ( $\epsilon_{loc}=-2$ ) at low temperatures. In this case, one can consider two possible folding mechanisms in the annealing process: in the first, the formation of the secondary structure is followed by the subsequent rearrangement of the structural elements. In the second, the formation of the secondary structure takes place after the dense hydrophobic core is formed. A clue to the answer can be obtained after analyzing the results shown in Figs. 2 and 3; the position of the peak on the  $C_v$  curve is located at  $T=1.6$ , while the collapse of the molecule takes place at significantly higher temperatures. This result suggests that the final number of secondary structures is achieved as an effect of rearrangements taking place in the collapsed structure.

## Conclusions

We have studied the properties of simplified models of confined polypeptide chains. The chains were modeled as linear sequences of united atoms located at the positions of the  $\alpha$ -carbons. The chains were put into a slit formed by a pair of two parallel impenetrable surfaces. The model was studied by means of Monte-Carlo simulations using the Replica-Exchange method. It was shown that the size of the slit had a significant influence on the size of the chains. Changes in the strength of the local potential also led to the different size and the local structure of chains. Introduction of the helical potential leads to the formation of significantly higher numbers of helical structures during the annealing of the system but the resulting low-temperature structures were not unique.

The most important conclusion is that under certain conditions the polypeptide model chain collapses into two stages. In the first stage, the size of the confined chain diminishes smoothly. In the second stage, the size contribution perpendicular to unconfined dimensions increases. In this stage, the change of the size along the  $z$ -axis is coupled with an increase of the number of residue–residue contacts, i.e. a decrease in energy.

**Acknowledgment** The computational part of this work was done using the computer cluster at the Computing Center of the Department of Chemistry, University of Warsaw.

## References

1. Pande VS, Grosberg AY, Tanaka T (2000) *Rev Mod Phys* 72:259–314
2. Mirny L, Shakhnovich E (2001) *Annu Rev Biophys Biomol Struct* 30:361–396

3. Minton AP (2001) *J Biol Chem* 276:10577–10580
4. Rivas G, Ferrone F, Hertzfeld J (2003) *EMBO Rep* 5:23–27
5. Eggers DK, Valentine JA (2001) *Protein Sci* 10:250–261
6. Kumar CV, Chaudhari A (2000) *J Am Chem Soc* 122:830–837
7. Bolis D, Politou AS, Kelly G, Pastore A, Temussi PA (2004) *J Mol Biol* 336:203–212
8. Peterson RW, Anbalagan K, Tommos C, Wand AJ (2004) *J Am Chem Soc* 126:9498–9499
9. Zhou H-X, Dill KA (2001) *Biochemistry* 40:11289–11293
10. Minton AP (1992) *Biophys J* 63:11090–11100
11. Kinjo AR, Takada S (2002) *Phys Rev E* 66, art no 031811
12. Betancourt MR, Thirumalai D (1999) *J Mol Biol* 287:627–644
13. Klimov DK, Newfield D, Thirumalai D (2002) *Proc Natl Acad Sci USA* 99:8010–8024
14. Takagi F, Koga N, Takada S (2003) *Proc Natl Acad Sci USA* 100:11367–11372
15. Kolinski A, Madziar P (1998) *Biopolymers* 42:537–548
16. Romiszowski P, Sikorski A (2000) *Biopolymers* 54:262–272
17. Sikorski A, Romiszowski P (2003) *Biopolymers* 69:391–398
18. Sikorski A, Romiszowski P (2005) *J Mol Model* 11:379–384
19. Shakhnovich EI, Gutin AM (1993) *Proc Natl Acad Sci USA* 32:7195–7199
20. Hansmann UHE, Okamoto Y (1999) In: Stauffer D (ed) *Annual reviews of computational physics VI*. World Scientific, Singapore, p 129
21. Sikorski A, Romiszowski P (2002) *J Chem Phys* 116:1731–1736
22. Romiszowski P, Sikorski A (2005) *J Chem Phys* 123, art no 104905