

Free energy of helix propagation in short polyalanine chains determined from peptide growth simulations of La³⁺-binding model peptides. Comparison with experimental data

Maciej Maciejczyk¹, Jan Hermans² and Andrzej Bierzyński¹✉

¹Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warszawa, Poland; ²Department of Biochemistry and Biophysics, School of Medicine, University of North Carolina, Chapel Hill, USA;

✉e-mail: ajb@ibb.waw.pl

Received: 19 April, 2005; revised: 14 November, 2005; accepted: 16 November, 2005
available on-line: 21 December, 2005

Molecular dynamics (MD) is, at present, a unique tool making it possible to study, at the atomic level, conformational transitions in peptides and proteins. Nevertheless, because MD calculations are always based on a more or less approximate physical model, using a set of approximate parameters, their reliability must be tested by comparison with experimental data. Unfortunately, it is very difficult to find a peptide system in which conformational transitions can be studied both experimentally and using MD simulations so that a direct comparison of the results obtained in both ways could be made. Such a system, containing a rigid α -helix nucleus stabilized by La³⁺ coordination to a 12-residue sequence taken from an EF-hand protein has recently been used to determine experimentally the helix propagation parameters in very short polyalanine segments (Goch *et al.* (2003) *Biochemistry* 42: 6840–6847). The same parameters were calculated here for the same peptide system using the peptide growth simulation method with, alternatively, charmm 22 and cedar potential energy functions. The calculated free energies of the helix-coil transition are about two times too large for cedar and even three times too large for charmm 22, as compared with the experimental values. We suggest that these discrepancies have their origin in the incorrect representation of unfolded peptide backbone in solution by the molecular mechanics force fields.

Keywords: molecular dynamics, thermodynamics, thermodynamic integration, peptide, helix-coil equilibrium

Molecular dynamics (MD) methods are a valuable tool in studies of conformational transitions in peptides and proteins. Only global kinetic and/or thermodynamic parameters of these processes can be determined from experiments, whereas they are precisely described, at the atomic level, by MD simulations. Consequently, the MD methods can give, in principle, answers to all possible questions concerning such processes. Unfortunately, they have one fundamental weakness: energy and forces are calculated on the basis of a *model* (a so-called molecular mechanics, or MM, model), and while the form and parameters of the model are based on our best understanding of the underlying physical reality, it is always uncertain to what extent the results of the calculations accurately describe the

physics of a particular system. To further confuse the issue, several implementations of the MM model have been developed, with equal claim to accuracy.

It is, therefore, essential to calibrate the MM model by establishing the reliability of MD results: only if the calculated global parameters are close enough to the experimental ones we can hope that the theoretical analysis is sound and that the detailed description given by simulations is close to reality. Many examples of calibration studies exist in the literature; however, the reliability of application of the MM model to helix-coil transitions of peptides remains inadequately documented.

An additional requirement of any test of reliability is that simulation results must be precise, in

Abbreviations: MD, molecular dynamics; MM, molecular mechanics; PDB, Protein Data Bank; PGS, peptide growth simulation; QM, quantum mechanics; RMSD, root mean square deviation; SCC-DFTB, self consistent charge density functional tight binding.

the sense that the effect of statistical noise has been eliminated (by sufficient averaging) from the model values. The results of experimental solution studies are averages over very large numbers of independent molecules, and hence have a negligible statistical error. Simulation studies are typically done with a single solute molecule; this gives a time-averaged behavior, which is then compared with the ensemble average obtained experimentally. The rate of convergence of a time average depends on the correlation over time of the averaged property: a property that fluctuates slowly requires a proportionally longer time for convergence to a given precision than one that fluctuates rapidly (with the RMSD approximately proportional to the square root of the simulation time). Thus, precision requires simulations extending over many times a property's correlation time (Straatsma *et al.*, 1986).

Unfortunately, it is very difficult to find a peptide system in which conformational transitions can be studied both experimentally and using MD simulations, so that a direct comparison of the results obtained in both ways could be made. A major difficulty is the long relaxation time of the helix-coil transition, as is typical for any nucleation-growth process (Schellman, 1955). This time is much shorter for helix formation in more flexible peptides composed of β -amino acids, and this has permitted precise studies of such systems by van Gunsteren and coworkers (Daura *et al.*, 1997; 1998; 1999). For peptides composed of α -amino acids, calculation results have frequently been compared with experimental data obtained for systems not identical with, but only more or less similar to, those studied by MD methods. The similarity is often so remote that the one conclusion that can be drawn from such comparisons is that the calculation results are not at odds with the experiment.

Because of its relevance for the protein folding process, the helix-coil transition has been studied very extensively. The experimental work concentrated mainly on determination of thermodynamic parameters of the transition in peptides of various lengths and sequences (Doig, 2002). Only peptides with a sufficiently large helix content (in practice >10%) and well soluble in water could be studied. Therefore, they must be long enough — built of at least thirteen, properly chosen amino-acid residues. Such long peptides present a major challenge for MD methods.

The fundamental problem met in MD simulations of helix-coil transitions is adequate probing of the peptide conformational space which is exceedingly large even for medium-sized peptides, built of about ten amino-acid residues. Moreover, the transition of each residue between the helical and coil states is a rare event because of a large en-

ergy barrier separating the extended and the helical conformations corresponding, respectively, to the second and the third quadrant of the Ramachandran plot.

Therefore, MD methods have been used mainly to characterize the structural and dynamical properties of the peptide helical conformation and the mechanism of its unfolding (DiCapua *et al.*, 1990; Daggett *et al.*, 1991; Soman *et al.*, 1991; Tirado-Rives & Jorgensen, 1991; Daggett & Levitt, 1992; Lee *et al.*, 2000; Armen *et al.*, 2003). The results of these studies were shown to be compatible with what was known from experimental investigations but quantitative comparisons with any experimental data were not possible.

The helix-coil transitions were simulated in very short peptides built of three (Tobias & Brooks, 1991) or five (Hummer *et al.*, 2000; 2001) amino-acid residues. Parallel experimental studies of such peptides are not feasible. Therefore, the calculated kinetic and/or thermodynamic parameters could be only quantitatively compared with information derived from experimental data obtained for much longer peptides.

A first attempt to estimate the free energy of the helix-coil transition in a peptide with a length comparable to that of the shortest peptide systems studied using experimental methods (thirteen amino-acid residues) was made by Daggett and Levitt (1992) from calculated waiting times for the $h \rightarrow c$ and $c \rightarrow h$ transitions. A similar approach was used recently by Armen *et al.* (2003). Because in both cases few such transitions were observed during the simulations, the free energy estimates can be treated, at best, as semi-quantitative.

Subsequently, Young and Brooks (1996) were able to calculate the free energy of helix propagation using the umbrella sampling method in a peptide containing as many as fifteen alanine residues. However, the authors met yet another problem: the calculation results strongly depended on the accepted, arbitrary definitions of the helix and coil states. Moreover, the calculated microscopic parameters were, as admitted by the authors, related to, but not identical with, the experimental parameters.

These two fundamental problems met in MD simulations of helix formation, i.e. adequate sampling of the peptide conformational space and the calculation of sufficiently precise thermodynamic parameters of reasonably long oligomers, were apparently solved in a satisfactory way by Wang and coworkers (1995) using the peptide growth simulation (PGS) method and, recently, by García and Sanbonmatsu (2002) using the replica exchange method (Sugita & Okamoto, 1999). Those two studies address the problem of convergence in a completely different manner: the PGS method does not

depend on the (slow) equilibration between helix and coil states, while the replica exchange method exploits the much more rapid equilibration that occurs in the high-temperature replicas. Furthermore, the free energy of helix propagation in long polyalanine chains (up to fourteen (Wang *et al.*, 1995) and twenty one (Garcia & Sanbonmatsu, 2002) Ala residues) calculated using these methods should, and was shown to be, close to the analogous parameter determined from experimental studies of polyalanine-based model peptide systems. Nevertheless, the compared theoretical and experimental systems, although close to each other, were still not identical. (These alanine oligomers are poorly soluble in water, and the experimental system is an extrapolation from experiments with molecules containing also amino-acid residues other than alanine.)

A peptide system suitable for both experimental and MD studies of the helix-coil transition has been proposed by Siedlecka *et al.* (1999). The calcium-binding peptide Ac-Loop-NH₂ with the sequence AcDKDGDGYISAAENH₂ strongly coordinates lanthanum ions and as a result assumes a rigid conformation with the φ angles of Ala10, Ala11, and Glu12 and the ψ angles of Ser9, Ala10 and Ala11 confined to the third quadrant of the Ramachandran plot. The La³⁺-saturated Loop sequence provides, therefore, a helix nucleus for the polyalanine segments A_n attached to its C-terminus in peptides PA_n with the sequence AcDKDGDGYISAAE(A)_nNH₂.

The conformation of metal-free peptides Ac-Loop-NH₂ (alternatively denoted as PA₀) and PA_n, where $n = 1 \div 4$, is close to random coil. Upon La³⁺ coordination the A_n part of the peptides (reduced to the NH₂ group in the case of PA₀) assumes, to a large extent (about 70%), the helical conformation. Recently, the free energies of these partial helix-coil transitions (ΔG_{An}) have been determined from measurements of La³⁺-binding constants to PA_n peptides, complemented by NMR measurements of proton exchange rates of their C-terminal NH₂ group (Goch *et al.*, 2003).

In their metal-bound form, PA_n peptides are excellent objects for MD studies. They are short, built of at most 16 residues, with twelve of them restricted to a very small fraction of conformational space. Moreover, their conformation is very compact, keeping at the minimum the number of solvent molecules that must be used in MD simulations.

Here we have calculated ΔG_{An} values for a series of peptides PA_n using the PGS method (see below) with, alternatively, charmm 22 and cedar potential energy functions (force fields) and compared them with exactly the same thermodynamic parameters determined experimentally.

MATERIALS AND METHODS

PGS method. PGS (Wang *et al.*, 1995) determine free energies of the transformation of shorter peptides into longer ones, in terms of the Thermodynamic Integration method (Allen & Tildesley, 1987). An atom (group of atoms) is slowly transformed into a different atom (group of atoms) and the free energy of this transformation is calculated according to the equation:

$$\delta G = \int_0^1 \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda \quad (1)$$

where V is a potential energy function depending on the transformation parameter λ in the following way: $V(\lambda=0)$ is the potential energy function of the starting component, and $V(\lambda=1)$ is the potential energy function of the final component. The brackets imply the averaging over a Boltzmann distribution. During the simulation the parameter λ is slowly changed either from 0 to 1 (growth process) or from 1 to 0 (annihilation process) and the integral described by Eqn. 1 is evaluated. For an infinitely long simulation process $\delta G_{\text{growth}} = -\delta G_{\text{annihilation}}$. In real calculations Eqn. 1 is approximated by:

$$\delta G \approx W = \sum \frac{\partial V}{\partial \lambda} \delta \lambda \quad (2)$$

where W represents the work performed on the system in order to force its state to change.

The growth processes analyzed in this work are presented in Fig. 1. The molecule is transformed to one containing an additional Ala residue at the end of the chain, and this is done twice, once with the added residue in an extended conformation, and again with the residue in the α -helical conformation, the free energy difference being accumulated for both processes. The difference between the two free energies is then the free energy difference for the transition of the added residue from extended (i.e., coil) to helical conformation. One sees that no transition between helix and coil state is directly simulated; consequently, the relaxation time of the system is not the considerable relaxation time of the helix-coil equilibrium, but the (apparently much) shorter time associated with rearrangement of the solvent structure around the growing residue.

The starting, artificial peptide P* has the same sequence as peptide PA₀ but with the C-terminal CO-NH₂ group substituted by the neutral carbon atom. The conformation of the metal-binding loop coordinated to an Ln³⁺ ion was taken from the PDB (code 1NKF). The lanthanum ion was modeled as a soft sphere with the Lennard-Jones parameters of the Ca²⁺ ion and the charge set to +3 elementary charges.

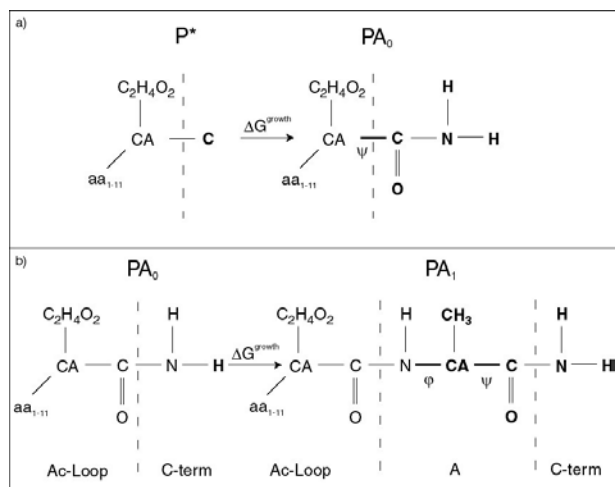


Figure 1. Illustration of PGS.

a) In $P^* \rightarrow PA_0$ transformation the starting, artificial peptide is terminated with a neutral carbon which is slowly replaced by $CONH_2$ group; b) The C-terminal amide NH_2 group of peptide PA_0 is replaced by terminally amidated alanine in the $PA_0 \rightarrow PA_1$ transformation. The segment built of amino acids 1–11 is denoted as aa_{1-11} . The transformed atoms are shown in bold characters. Flat-bottom restraints placed on the indicated ϕ and ψ dihedrals keep the growing groups in the desired (helical or extended) conformation.

A more faithful representation of this system is not possible because the force field parameters for lanthanide ions are not known. For the cedar force field additional restraints on distances between the Ln^{3+} ion and the groups coordinating it were applied, preventing their dissociation from the ion.

To keep the dihedral angles of the peptide backbone groups in the desired region of conformational space “flat-bottom” restraints (Hermans *et al.*, 1992) were applied. Such restraints, when placed correctly, effectively divide the probed conformational space and allow one to simulate exclusively one particular conformation, but to not perturb it locally.

Positions of the alanine free energy barriers were located with the potential of mean force calculations of the dihedrals ϕ and ψ of the C-terminal alanine in peptide PA_2 (results not shown). Similar positions of energy barriers have been obtained from free dynamics simulations of a system modeling the alanine dipeptide (Smith, 1999; Hu *et al.*, 2003). Each restraint consists of two ramp potentials placed on the free energy barriers on either side of the free energy minimum. These ramp consisted of a penalty energy changing linearly from 0 to 20 kcal/mol over a 10° interval. For the helical and extended conformations the backbone dihedrals were restrained to the regions $-180^\circ < \phi < 0^\circ$, $-140^\circ < \psi < 20^\circ$ and $140^\circ < \phi < 0^\circ$, $20^\circ < \psi < -140^\circ$ of the Ramachandran plot, respectively.

Simulation procedure. All simulations were carried out with the program Sigma (Mann *et al.*, 2002). Bond lengths were constrained to their equilibrium values with the SHAKE algorithm (Ryckaert *et al.*, 1977). Electrostatic interactions were calculated with a multiple time step method (Tuckerman *et al.*, 1992) with the dual cut-off set to 6 Å and 10 Å. Short range, medium range, and long range (*via* the particle-mesh Ewald summation method (Darden *et al.*, 1993; Schlick *et al.*, 1999)) interactions were calculated, respectively, at 2, 4, and 12 fs intervals. Simulations were carried out at constant pressure (1 atm) and temperature (300 K) (Berendsen *et al.*, 1984).

Each peptide was placed in the center of a water-filled cubic box and periodic boundary conditions were applied. The box dimensions ranged from 38 up to 44 Å depending on peptide length and conformation (extended or helical) and always exceeded the peptide dimensions by at least 20 Å.

Each calculation of the growth free energy consisted of the following steps: 1) minimization of water energy with fixed solute coordinates; 2) 50 ps equilibration of water with fixed solute conformation; 3) minimization of water energy with fixed solute conformation; 4) minimization of solute energy with fixed solvent conformation; 5) energy minimization of the whole system; 6) 9.6 ps MD simulations at the temperature of 50 K with starting velocities assigned from Boltzmann distribution; 7) five 5.8 ps MD simulations at temperatures increasing from 100 to 300 K in 50 K steps; 8) two peptide growth/annihilation cycles, each consisting of a) 29 ps equilibration at $\lambda=0$, b) 576 ps growth simulation, c) 29 ps equilibration at $\lambda=1$, d) 576 ps annihilation simulation.

Two force-fields were tested: 1) charmm 22 all-atom force field (MacKerell *et al.*, 1998) distributed with the XPLOR program (Brunger, 1992) combined with the TIP3P (Jorgensen *et al.*, 1983) water model and 2) cedar/gromos all-atom force field distributed in the Sigma package (Ferro *et al.*, 1980; Hermans *et al.*, 1984) combined with the SPC water model. Protein structure files were prepared with the XPLOR program (Brunger, 1992).

Error estimates. The calculated growth free energy, δG for each process is based on the estimates of energy changes, W calculated (with Eqn. 2) for at least two simulations in which a residue was inserted and two simulations in which the same residue was deleted. These values were used to give independent mean and mean-square deviation for the insertion and deletion processes. Because the simulation time is finite, the net energy change for a cyclic process of insertion followed by deletion is not zero, $\langle W_i \rangle + \langle W_d \rangle > 0$; however, the mean value, W_i of the energy for insertion and of $-W_d$ (where W_d is the mean energy for deletion) can be taken as a

best estimate of the growth free energy. In addition, an error estimate follows by taking the mean of the mean-square deviations of the insertion and deletion energies (Hermans *et al.*, 1992; Hu *et al.*, 2002).

RESULTS

Simulation results

The calculated growth free energies $\delta G_n^{i,e}$, cumulative free energies $\Delta G_n^{i,e}$, and free energies of single residue transitions from the extended to the helical conformation $\Delta\Delta G_n^{i,e} = \Delta G_n^{i,e} - \Delta G_n^{i-1,e}$ defined in Fig. 2 are listed in Table 1. Cumulative free energies $\Delta G_n^{i,e}$ are sums of growth free energies $\delta G_n^{i,e}$ along the appropriate path (see Fig. 2).

$$\Delta G_n^{i,e} = \sum_{path} \delta G_n^{i,e} \quad (3)$$

For example, $\Delta G_2^{1,e} = \delta G_0^{1,e} + \delta G_1^{1,e} + \delta G_2^{1,e}$ and $\Delta G_2^{2,e} = \delta G_0^{1,e} + \delta G_1^{2,e} + \delta G_2^{2,e}$.

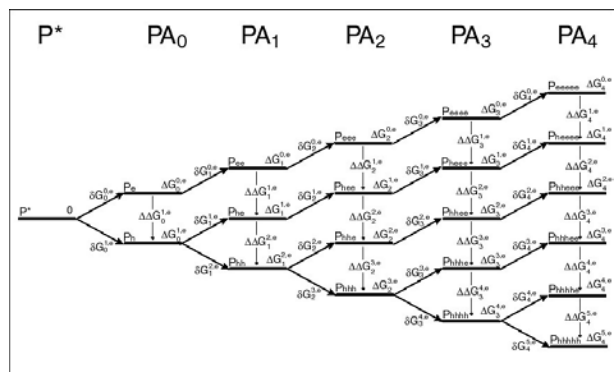


Figure 2. Schematic presentation of PGS performed in this work.

Extended, helical, and partly helical conformations of C-terminal segments of PA_n peptides are described by chains of e and/or h subscripts. E.g., the symbol P_{hhhe} denotes peptide PA_3 with its loop conformation restricted by Ln^{3+} ion binding, Ala13 and Ala14 in the helical conformation, and Ala15 and the terminal NH_2 group in the extended conformation. Growth simulations are indicated by thick arrows and growth free energies corresponding to them are denoted by $\delta G_n^{i,e}$ where n is the number of alanine residues created by the PGS procedure, i denotes the total number of peptide C-terminal groups, including NH_2 , in the helical conformation, and e indicates that the remaining groups are in the extended conformation, even if their number is zero, as in the case of $\delta G_n^{n+1,e}$ symbols. Similarly, the cumulative free energy of a series of growth steps transforming P^* into peptide PA_n with i terminal groups in the helical conformation is denoted by $\Delta G_n^{i,e}$ symbol. Thin vertical arrows indicate the conformational transitions of a single alanine residue or NH_2 group from the extended to the helical conformation. The free energies of these transitions $\Delta\Delta G_n^{i,e} = \Delta G_n^{i,e} - \Delta G_n^{i-1,e}$.

Free energies of helix-coil transitions

The helix-coil transitions observed in experimental studies do not correspond precisely with the transitions between the helical and the extended conformations simulated by the PGS method. Although the coil state is dominated by the extended conformation it includes also some residues with the φ and/or ψ angles characteristic for the helical conformation (Wang *et al.*, 1995), if these residues are not a part of a helix nucleus and the amide protons of the peptide groups following their C_α carbons do not form the helical ($i \rightarrow i - 4$ or $i \rightarrow i - 3$) hydrogen bonds. These states are excluded from the extended conformation by the restraints imposed during the simulations.

The non-helical, coil state of PA_0 peptide (Pc) contains only one, extended conformation (Pe) with the ψ angle of Glu12 kept outside the helical region, because otherwise the terminal NH_2 group would be involved in a helical hydrogen bond. Consequently, the coil state is identical with the extended conformation simulated by the PGS procedure.

This is not the case for longer peptides. E.g., the coil state of PA_2 peptide (Pccc), consists not only of the extended conformation Peee but also of three additional conformations: Pehe, Peeh, and Pehh. Consequently, the free energy of this state $\Delta G_2^{2,c} < \Delta G_2^{2,e}$. Similarly, $\Delta G_2^{1,c} < \Delta G_2^{1,e}$ because not only the Phee conformation but also the Pheh one is populated in the Phcc state.

Corrections needed to transform $\Delta G_2^{1,e}$ parameters into $\Delta G_2^{1,c}$ ones are small (see Table 1). They can be calculated if populations of all conformations including "h" alanines in the respective coil states are known. We estimated them from the data obtained by Hu *et al.* (2003) from MD simulations of the alanine dipeptide. For both the charmm 22 and cedar force fields they calculated the equilibrium constants K between the alanine residue conformations with helical and non-helical φ and ψ dihedrals: $K = 1$ and 0.29 for charmm 22 and cedar, respectively.

Free energy levels corresponding to additional h conformations are shifted down by the value equal to $l k_B T \ln K$ below the extended conformation. l is the number of 'e' alanines replaced by 'h' ones and k_B is Boltzmann's constant. From statistical thermodynamics (Hill, 1986) it then follows that

$$\Delta G_n^{i,c} = \Delta G_n^{i,e} - kT \ln \left(1 + \sum_{l=1}^{p-1} \binom{p-1}{l} K^l \right) \quad (4)$$

Free energies of single residue helix-coil transitions $\Delta\Delta G_n^{i,c} = \Delta G_n^{i,c} - \Delta G_n^{i-1,c}$, analogous to the $\Delta\Delta G_n^{i,e}$ parameters, were calculated from the $\Delta G_n^{i,e}$ values using Eqn. 4 and are listed in Table 1.

Table 1. Results of PGS obtained with charmm 22 and cedar force fields. $\delta G_n^{i,e}$, $\Delta G_n^{i,e}$, and $\Delta\Delta G_n^{i,e}$ and parameters are defined in Fig. 2. $\Delta\Delta G_n^{i,e}$ parameters have been transformed into the free energies of the respective helix-coil transitions ($\Delta\Delta G_n^{i,c}$) as described in the text.

Charmm 22							Cedar				
n	i	#	$\delta G_n^{i,e}$	$\Delta G_n^{i,e}$	$\Delta\Delta G_n^{i,e}$	$\Delta\Delta G_n^{i,c}$	#	$\delta G_n^{i,e}$	$\Delta G_n^{i,e}$	$\Delta\Delta G_n^{i,e}$	$\Delta\Delta G_n^{i,c}$
0	0	4	-7.1 ± 0.2	-7.1			4	-35.5 ± 0.2	-35.5		
		1	-8.5 ± 0.1	-8.5	-1.4	-1.4	4	-36.0 ± 0.1	-36.0	-0.5	-0.5
1	0	4	11.5 ± 0.5	4.4			4	6.1 ± 0.4	-29.4		
		1	12.1 ± 0.2	3.6	-0.8	-0.4	2	6.0 ± 0.2	-30.0	-0.6	-0.4
		2	10.4 ± 0.4	1.8	-1.8	-1.8	2	5.9 ± 0.2	-30.1	-0.1	-0.1
2	0	4	10.9 ± 0.4	15.3			2	7.1 ± 0.3	-22.3		
		1	11.3 ± 0.3	14.9	-0.4	-0.1	2	6.7 ± 0.3	-23.3	-1.0	-0.8
		2	10.5 ± 0.3	12.4	-2.5	-2.1	2	6.1 ± 0.1	-24.0	-0.7	-0.4
		3	9.7 ± 0.2	11.6	-0.8	-0.8	2	5.8 ± 0.3	-24.3	-0.3	-0.3
3	0	2	11.2 ± 0.1	26.5			2	6.7 ± 0.3	-15.6		
		1	11.4 ± 0.1	26.3	-0.2	0.2	2	6.7 ± 0.3	-16.6	-1.0	-0.8
		2	11.6 ± 0.2	24.0	-2.3	-2.0	2	6.6 ± 0.3	-17.4	-0.8	-0.7
		3	10.8 ± 0.3	22.3	-1.7	-1.3	2	6.0 ± 0.4	-18.3	-0.9	-0.7
		4	10.2 ± 0.3	21.8	-0.5	-0.5	2	5.7 ± 0.3	-18.6	-0.3	-0.3
4	0	2	11.3 ± 0.2	37.8			2	7.1 ± 0.1	-8.5		
		1	11.3 ± 0.2	37.6	-0.2	0.1	2	7.0 ± 0.2	-9.6	-1.1	-1.0
		2	11.4 ± 0.2	35.4	-2.2	-1.7	2	7.2 ± 0.2	-10.3	-0.7	-0.5
		3	11.7 ± 0.2	34.1	-1.3	-1.0	2	6.6 ± 0.2	-11.7	-1.4	-1.3
		4	10.9 ± 0.4	32.7	-1.4	-1.0	4	6.5 ± 0.3	-12.1	-0.4	-0.2
		5	9.9 ± 0.2	31.7	-1.0	-1.0	2	6.0 ± 0.3	-12.6	-0.5	-0.5

Free energies of conformational transitions induced in C-terminal segments of PA_n peptides by La³⁺ binding

The conformation of C-terminal segments of apo-PA_n peptides is close to random coil (Goch *et al.*, 2003). Upon La³⁺ binding a large fraction of their residues assume the helical conformation. The helix builds up from its nucleus formed in the Loop sequence by coordination of the metal ion. The conformational state of A_n segments can be described as a dynamical equilibrium between the coil conformation and an assembly of helical segments containing from 1 up to $n + 1$ backbone groups (n alanine residues and the C-terminal NH₂ group). The free energy of such a multiple-equilibrium system ΔG_{An}^{met} (see Fig. 3) is given by the equation (Poland & Scheraga, 1970):

$$\Delta G_{An}^{met} = -kT \ln \left[\sum_{i=0}^{n+1} \exp(-\Delta G_n^{i,c} / kT) \right] \quad (5)$$

Since the free energies of the coil conformation of A_n segments of apo-PA_n peptides equal $\Delta G_n^{0,c}$ the free energy changes ΔG_{An} related with the conformational transitions induced in them by Ln³⁺ coordination are described by the equation:

$$\Delta G_{An} = \Delta G_{An}^{met} - \Delta G_n^{0,c} = -kT \ln \left[\frac{\sum_{i=0}^{n+1} \exp(-\Delta G_n^{i,c} / kT)}{\exp(-\Delta G_n^{0,c} / kT)} \right] = -kT \ln \left[1 + \sum_{j=1}^{n+1} \prod_{i=1}^j \exp(-\Delta\Delta G_n^{i,c} / kT) \right] \quad (6)$$

The ΔG_{An} values calculated from Eqn. 6 and the $\Delta\Delta G_n^{i,c}$ parameters listed in Table 1 are presented in Table 2.

Helix populations in PA_n peptides

The population of helical segments built of j residues is given by the equation:

$$p_j = \frac{\exp\left(-\Delta G_n^{j,c}/kT\right)}{\sum_{i=0}^{n+1} \exp\left(-\Delta G_n^{i,c}/kT\right)} \quad (7)$$

Using Eqn. 6 the molar fractions of full helices including the C-terminal NH_2 groups ($x_{\text{nNh}} = p_{j+1}$) and the average populations of amino-acid residues in the helical conformation $x_{\text{nh}} = \sum_{j=1}^{n+1} jp_j/(n+1)$ were calculated and compared with the experimental values in Table 2.

DISCUSSION

The importance of a “direct” comparison of MD results with experimental data was stressed in the introduction. Let us specify what it actually means. Raw data of any experimental or theoretical study must be analyzed and interpreted to give physically meaningful results. Such a procedure is based on accepted theoretical models and on more or less solid assumptions. Although indispensable, such an analysis in some way always reduces and distorts the information contained in the data. Therefore, the data obtained using different methods should be compared at the lowest possible level of interpretation.

The crucial merit of our peptide system is that after the very first steps of data analysis the experimental and theoretical results can be compared. Only one assumption, i.e., that the interaction of the An and Loop segments with each other in metal-free PA_n peptides is negligible, was necessary to determine the $\Delta G_{\text{An}}^{\text{exp}}$ values from the experimental data (Goch *et al.*, 2003) without using any explicit helix-coil transition theory. On the other hand, the model’s ΔG_{An} parameters (see Table 1) have been calculated using Eqn. 6, based only on fundamental

principles of statistical thermodynamics, from δG_n^{ie} the values obtained from PGS, with only a slight correction being needed to transform these into the δG_n^{ic} parameters.

The statistical errors of δG_n^{ic} (see Table 1) lead to errors in the determined ΔG_{An} values that can be evaluated at about 0.3 kcal/mol (PA_0) to 1 kcal/mol (PA_4), which is of the same order of magnitude as the experimental values. Clearly, the peptide systems studied by us, although very small, need much longer simulation times, which, at present, are unacceptable because of the limits on available computing power. However, the systematic errors due to the use of imperfect force field parameters are much larger: the ΔG_{An} values calculated using cedar and charmm 22 exceed the experimental ones by as much as 3 kcal/mol.

The free energy is, of course, the fundamental parameter describing the helix-coil transition. Nevertheless, the precision of its determination does not correlate linearly with the precision of the prediction of the helix population which is often used, in practice, to estimate the value of theoretical calculations. Whereas the helix populations (x_{nNh} and x_{nh}) determined using the cedar force field are very close to the experimental data, much larger deviations are observed when charmm 22 is used (see Table 2).

We suggest that the serious discrepancies between the results of MD simulations and experimental data have their origin in the incorrect representation of unfolded peptide backbone in solution (Hu *et al.*, 2003) by the MM force fields. Hu and coworkers presented results of MD simulations of Ace-Ala-Nme and Ace-Gly-Nme “dipeptides” in solution, using several force fields including charmm22 and cedar, and also with a mixed QM/MM force field, which they performed with the SCC-DFTB method (Elstner *et al.*, 1998). One result was the great variation of the conformational distributions obtained with the dif-

Table 2. Comparison of simulated and experimental results: free energies of the helix-coil transitions induced in C-terminal segments of peptides PA_n by La^{3+} binding calculated using charmm 22 and cedar force fields (ΔG_{An}) and determined from experiments ($\Delta G_{\text{An}}^{\text{exp}}$), full-helix populations in An segments of La^{3+} -saturated peptides PA_n calculated (x_{nNh}) and determined experimentally ($x_{\text{nNh}}^{\text{exp}}$), fractional population of helical hydrogen bonds calculated from the MD simulations (x_{nh}) and the experimental data ($x_{\text{nh}}^{\text{exp}}$).

n	ΔG_{An} charmm	ΔG_{An} cedar	$\Delta G_{\text{An}}^{\text{exp}}$	x_{nNh} charmm	x_{nNh} cedar	$x_{\text{nNh}}^{\text{exp}}$	x_{nh} charmm	x_{nh} cedar	$x_{\text{nh}}^{\text{exp}}$
0	-1.5	-0.7	$-0.3 \pm 0.1^{(a)}$	0.91	0.70	$0.61 \pm 0.04^{(a)}$	0.91	0.70	0.61
1	-2.2	-1.0	$-0.7 \pm 0.3^{(b)}$	0.93	0.45	$0.56 \pm 0.04^{(a)}$	0.95	0.64	0.67
2	-3.2	-2.1	$-1.0 \pm 0.2^{(b)}$	0.78	0.52		0.92	0.77	0.73
3	-3.7	-2.9	$-1.3 \pm 0.2^{(b)}$	0.69	0.54	$0.45 \pm 0.05^{(a)}$	0.91	0.84	0.73
4	-4.8	-3.8	$-1.5 \pm 0.2^{(b)}$	0.83	0.56	$0.43 \pm 0.05^{(a)}$	0.96	0.86	0.74

^(a)Determined from NMR measurements of the exchange rates of C-terminal NH_2 protons in peptides PA_n . See Table 2 in Goch *et al.* (2003). ^(b)Obtained from $\Delta \Delta G_{\text{tn}}$ values determined from La^{3+} -binding constants to peptides PA_n given in the last column of Table 1 in Goch *et al.* (2003). $\Delta G_{\text{An}}^{\text{exp}} = \Delta G_{\text{A0}}^{\text{exp}} + \Delta \Delta G_{\text{tn}}$. The error limits have been estimated from the errors of binding constant determinations listed in the first column of Table 1 in Goch *et al.* (2003).

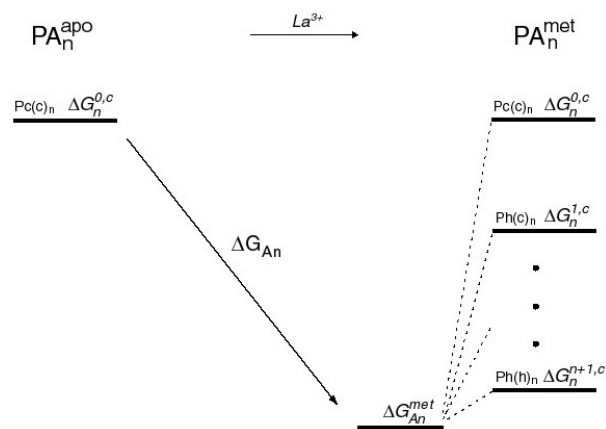


Figure 3. Schematic presentation of free energy changes (ΔG_{An}) related to conformational transitions induced in the C-terminal segments of peptides PA_n by La^{3+} ion coordination.

$\Delta G_n^{i,c}$ parameters were calculated from the corresponding $\Delta G_n^{i,e}$ values (see Fig. 2 and Table 1) using Eqn. 3. ΔG_{An}^{met} describe free energies of multiple equilibrium systems composed of $n + 2$ conformational states.

ferent force fields, which clearly indicated a general lack of accuracy of the MM energy representation. The distributions obtained with the QM/MM simulation method were in striking agreement with the distributions of well-defined Ala and Gly residues in globular proteins and was judged by them to be the more reliable result. For the charmm22 force field the equilibrium constant K between the α_R and β conformations of the Ala dipeptide obtained from classical simulation ($K = 1$) is larger than the equilibrium constant obtained with the QM representation ($K = 0.69$). This result suggests that with this force field the helical conformation is too stable, and this agrees well with the results of our thermodynamic test. For the cedar force field the alanine dipeptide's α_R -to- β ratio, ($K = 0.31$) is closer to that found with the QM/MM force field ($K = 0.44$). Also, the results presented in this paper suggest that the cedar force field produces a better description of the thermodynamics of helix formation than does charmm 22, as free energies ΔG_{An} as well as full (x_{nNh}) and fractional (x_{nh}) helix populations obtained with cedar force field are smaller and closer to the experiment than those obtained with charmm 22 (see Table 2).

One can see qualitatively that the conformational distribution of the Ala dipeptide with a particular MM potential function (if Ala is taken as a representative of all amino acids with a β -carbon) is relevant to the problem of helix stability predicted with that same potential, in two ways. First, if the population of dipeptide conformations with the dihedral angles similar to those found in the α -helix is high, this reflects a low energy state that can be expected to carry over to also favor the α -helical structure in a longer peptide. Second, if the distribution

of the dipeptide is broad, then this corresponds to high conformational entropy, that can be expected to carry over to favor the coil structure in a longer peptide.

Backbone potentials can be quite easily altered and tested by simulation of Ala and Gly dipeptides. Guided by the results such as those from simulations with QM energy and forces, by distributions of amino acids in proteins and by any available results of experimental studies of short peptides in solution, one can proceed to check one or several of these altered potentials and see how these affect the free energies of helix formation in a system such as the La^{3+} -binding peptides.

Recently MacKerell *et al.* (2004) developed grid corrections for the charmm 22 force field. They showed that simulation of Ace-Ala-NMe with this improved force-field led to a better agreement of the generated distributions of backbone dihedrals with the results of QM calculations and with distributions obtained from a PDB survey. One may expect that this corrected force-field should provide us also with a better thermodynamic description of helix formation. The peptide system proposed by us is an excellent tool for performing such thermodynamic tests of new, improved force-fields.

Acknowledgements

This work was supported by the State Committee for Scientific Research Grant (Poland) 6 P04 009 16, a research grant from the U.S. National Institutes of Health and a computational grant from the Interdisciplinary Centre for Mathematical and Computational Modeling (Warszawa, Poland).

REFERENCES

- Allen MP, Tildesley DJ (1987) *Computer Simulation of Liquids*. Oxford University Press, New York.
- Armen R, Alonso DO, Daggett V (2003) The role of α , 3_{10} , and π -helix in helix-to-coil transitions. *Protein Sci* **12**: 1145–1157.
- Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* **81**: 3684–3690.
- Brunger AT (1992) *X-PLOR, a System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.
- Daggett V, Levitt M (1992) Molecular dynamics simulations of helix denaturation. *J Mol Biol* **223**: 1121–1138.
- Daggett V, Kollman PA, Kuntz ID (1991) A molecular dynamics simulation of polyaniline: an analysis of equilibrium motions and helix-coil transitions. *Biopolymers* **31**: 1115–1134.
- Darden TA, York DM, Pedersen LG (1993) Particle mesh Ewald: an $N \cdot \log(N)$ method for Ewald sums in large systems. *J Chem Phys* **98**: 10089–10092.
- Daura X, van Gunsteren WF, Rigo D, Juan B, Seebach D (1997) Studying the stability of a helical β -heptapep-

- tide by molecular dynamics simulation. *Chem Eur J* **3**: 1410–1417.
- Daura X, Jaun B, Seebach D, van Gunsteren WF, Mark AE (1998) Reversible peptide folding in solution by molecular dynamics simulation. *J Mol Biol* **280**: 925–932.
- Daura X, van Gunsteren WF, Mark AE (1999) Folding-unfolding thermodynamics of a β -heptapeptide from equilibrium simulations. *Proteins* **34**: 269–280.
- DiCapua FM, Swaminathan S, Berveridge DL (1990) Theoretical evidence for destabilization of an α -helix by water insertion: Molecular dynamics of hydrated decaalanine. *J Am Chem Soc* **112**: 6768–6771.
- Doig AJ (2002) Recent advances in helix-coil theory. *Biophys Chem* **101/102**: 281–293.
- Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G (1998) Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys Rev B* **58**: 7260–7268.
- Ferro DR, McQueen JE Jr, McCown JT, Hermans J (1980) Energy minimizations of rubredoxin. *J Mol Biol* **136**: 1–18.
- Garcia AE, Sanbonmatsu KY (2002) α -Helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc Natl Acad Sci USA* **99**: 2782–2787.
- Goch G, Maciejczyk M, Oleszczuk M, Stachowiak D, Malicka J, Bierzynski A (2003) Experimental investigation of initial steps of helix propagation in model peptides. *Biochemistry* **42**: 6840–6847.
- Hermans J, Berendsen HJC, van Gunsteren WF, Postma JPM (1984) A consistent empirical potential for water-protein interactions. *Biopolymers* **23**: 1513–1518.
- Hermans J, Yun RH, Anderson AG (1992) Precision of free energies calculated by molecular dynamics simulations of peptides in solution. *J Comp Chem* **13**: 429–442.
- Hill TL (1986) *An Introduction to Statistical Thermodynamics*. Dover Publications Inc., New York.
- Hu H, Yun RH, Hermans J (2002) Reversibility of free energy simulations: Slow growth may have a unique advantage. *Mol Simul* **28**: 67–80.
- Hu H, Elstner M, Hermans J (2003) Comparison of a QM/MM force field and molecular mechanics force fields in simulations of alanine and glycine “dipeptides” (Ace-Ala-Nme and Ace-Gly-Nme) in water in relation to the problem of modeling the unfolded peptide backbone in solution. *Proteins* **50**: 451–463.
- Hummer G, Garcia AE, Garde S (2000) Conformational diffusion and helix formation kinetics. *Phys Rev Lett* **85**: 2637–2640.
- Hummer G, Garcia AE, Garde S (2001) Helix nucleation kinetics from molecular simulations in explicit solvent. *Proteins* **42**: 77–84.
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**: 926–935.
- Lee KH, Benson DR, Kuczera K (2000) Transitions from α to π helix observed in molecular dynamics simulations of synthetic peptides. *Biochemistry* **39**: 13737–13747.
- MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fisher S, Gao J, Guo H, Ha S, Nguyen DT, Prodhom B, Reiher III WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **102**: 3586–3616.
- MacKerell AD Jr, Feig M, Brooks CL 3rd (2004) Improved Treatment of the protein backbone in empirical force fields. *J Am Chem Soc* **126**: 698–699.
- Mann G, Yun RH, Nyland L, Prins J, Board J, Hermans J (2002) Sigma MD Program and a generic interface applicable to multifunctional programs with complex, hierarchical command structure. In *Computational Methods for Macromolecules: Challenges and Applications* (Schlick T, Gan HH, eds). *Proc 3rd Int Workshop: Algorithms for Macromolecular Modeling*. October 12–14, 2000, Springer-Verlag, Berlin, New York.
- Poland D, Scheraga HA (1970) *Theory of Helix-coil Transitions in Biopolymers*. Academic Press, New York and London.
- Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comp Phys* **23**: 327–341.
- Schellman JA (1955) The stability of hydrogen-bonded peptide structures in aqueous solution. *Compt Rend Trav Carlsberg (Ser Chim)* **29**: 230–259.
- Schlick T, Skeel RD, Brunger AT, Kale LV, Board JA, Hermans J, Schulten K (1999) Algorithmic challenges in computational molecular biophysics. *J Comp Phys* **151**: 9–48.
- Siedlecka M, Goch G, Ejchart A, Sticht H, Bierzynski A (1999) α -Helix nucleation by a calcium-binding peptide loop. *Proc Natl Acad Sci USA* **96**: 903–908.
- Smith PE (1999) The alanine dipeptide free energy surface in solution. *J Chem Phys* **111**: 5568–5579.
- Soman KV, Karimi A, Case DA (1991) Unfolding of an α -helix in water. *Biopolymers* **31**: 1351–1361.
- Straatsma TP, Berendsen HJC, Stam A (1986) Estimation of statistical errors in molecular simulation calculations. *Mol Phys* **57**: 89–95.
- Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* **314**: 141–151.
- Tirado-Rives J, Jorgensen WL (1991) Molecular dynamics simulations of the unfolding of an α -helical analogue of ribonuclease A S-peptide in water. *Biochemistry* **30**: 3864–3871.
- Tobias DJ, Brooks CL 3rd (1991) Thermodynamics and mechanism of α helix initiation in alanine and valine peptides. *Biochemistry* **30**: 6059–6070.
- Tuckerman ME, Berne BJ, Martyna GJ (1992) Reversible multiple time scale molecular dynamics. *J Chem Phys* **97**: 1990–2001.
- Wang L, O’Connell T, Tropsha A, Hermans J (1995) Thermodynamic parameters for the helix-coil transition of oligopeptides: molecular dynamics simulation with the peptide growth method. *Proc Natl Acad Sci USA* **92**: 10924–1098.
- Young WS, Brooks CL 3rd (1996) A microscopic view of helix propagation: N and C-terminal helix growth in alanine helices. *J Mol Biol* **259**: 560–572.