Pergamon

Biophysics, Vol. 42, No. 1, pp. 45–51, 1997

0006–3509/97 $24.00+.00

PII: S0006-3509(97)00069-O

MOLECULAR DYNAMICS OF OLIGOPEPTIDES. I. THE USE OF LONG TRAJECTORIES AND HIGH TEMPERATURES TO DETERMINE THE STATISTICAL WEIGHT OF CONFORMATIONAL SUBSTATES*

K. V. SHAITAN, N. K. BALABAYEV, A. S. LEMAK, M. D. YERMOLAYEVA, A. G. IVAIKINA, M. V. ORLOV and Ye. V. GEL'FAND

Biological Institute of the Lomonosov State University, Moscow
Institute of Mathematical Problems of Biology,
Russian Academy of Sciences, Pushchino (Moscow region)

(*Received* 19 *August* 1996)

A technique has been developed for studying the physicochemical properties of biomacromolecules, as illustrated by some modified dipeptides, using the method of molecular dynamics. The results obtained at different temperatures, trajectory lengths and ways of setting the thermostat are compared. It is shown that for the peptide group studied the conditions of a collisional medium, trajectory lengths of not less than 5000 ps and temperatures of the order of 1000 K are optimal. In this case the figurative point is able to scan the configurational space of the molecule a sufficient number of times to obtain a statistically significant result. © 1997 Elsevier Science Ltd. All rights reserved.

Calculations of the molecular dynamics of proteins within particular approximations are now highly popular [1, 2]. The use of the methods of molecular dynamics and graphics for the graphic visual perception of the detailed structural and energy features of the behaviour and interaction of biomacromolecules also proves to be very useful for studying the mechanisms by which proteins function [3, 4]. So far a very limited amount of work has been published on a comparative study of the dynamic properties of relatively simple biopolymers with the aim of establishing the main physicochemical patterns or rules determining the dynamic behaviour of amino acid residues and other elementary "building blocks" forming the macromolecule (see, for example, [5–8]). This is made even more interesting by the fact that for systems with conformational degrees of freedom the concept of normal modes turns out to be meaningless and different approaches are needed to predict their dynamic behaviour (see, for example, [8–10]). The present paper is the first in a series devoted precisely to a comparative study of the dynamics of amino acid residues under different conditions, a closer definition of the influence of different approximations and which develop methods of treating the molecular dynamics of trajectories in order to obtain quantitative characteristics describing the dynamic structure of the biomacromolecule.

---

* *Biofizika*, **42**, No. 1, 47–53, 1997.

METHODS

The test object was a series of dipeptides: Asp–Asp, Asp–Lys, Lys–Lys, Ala–Tyr and Tyr–Trp and also a set of dipeptides containing a glycine residue and each of 20 biologically important amino acid residues. To lessen the influence on the dynamics of marginal effects from the strongly polar terminal groups, the N-terminals of the dipeptides were closed by acetyl, and the C-terminals by N-methylamine with the formation of two extra peptide bonds. The general structure of the molecules is shown in Fig. 1. The dynamics of the modified dipeptides were modelled numerically using the standard molecular-dynamic technique [1, 2]. In our numerical experiments the heavy atom model of the molecule was used. The dynamics of the molecules was modelled at constant temperature using the Berendsen method [11] and the method of collisional dynamics [12]. The potential energy $U(r)$, determining the force field, had components corresponding to the deformations of the valence lengths and valence angles, the torsional angles, the van der Waals and the Coloumb non-valent interactions and also the hydrogen bonds. The parameters of the force interactions and the partial charges on the atoms of the polypeptide chain are taken from [5], having regard to all paired interactions. The permittivity of the medium corresponded to vacuum conditions. The procedure of fixing the pseudotorsional angles was also used to block transitions leading to a change in the chirality of the amino acid residues. For the calculations used in the Berendsen thermostat [11], the set temperature (the mean critical kinetic energy) was maintained by introducing sign-variable friction with a coefficient $\gamma(T(t)/T_0 - 1)$, where T_0 is the temperature of the thermostat and $T(t)$ is the current "temperature" calculated via the mean kinetic energy of the molecule. The characteristic relaxation time of the temperature fluctuations $1/\gamma$ was taken to be 1 ps.

The method of collisional dynamics [12] uses a virtual collisional medium, the set effective temperature being maintained by pulses of random strength acting on each heavy atom. The mean frequency of the collisions with each atom was taken as equal to 10 ps, which roughly corresponds to the situation of an aqueous medium. Virtual collisions occur by the law of elastic spheres. The distribution of the pulses of the particles of the "medium" with a mass $m = 18$ (the mass of H_2O) corresponds to the set temperature. The set of equations of motion was integrated with a step 1 ps and the trajectory file was recorded every 0.1 ps. The trajectories obtained were then processed by an original technique which enabled us to determine the one-dimensional probability distribution $P(\alpha_n)$ and the two-dimensional probability distribution $P(\alpha_n, \alpha_m)$ for finding the system in different regions of the configurational space of the molecule. A strict definition of these quantities is given by the following integrals:

$$\begin{aligned} P(\alpha_n) &= \int \ldots \int P(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_N) \prod d\alpha_i (i \neq n), \\ P(\alpha_n, \alpha_m) &= \int \ldots \int P(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_N) \prod d\alpha_i (i \neq n, m), \end{aligned} \tag{1}$$

where α_i is the set of dynamic variables and $P(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_N)$ is the probability density of

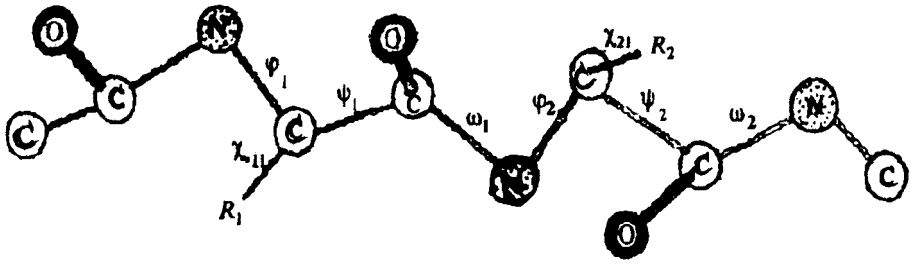

Fig. 1. Molecule of modified dipeptide (the hydrogen atoms are not shown).

detecting the system at a set point of configurational space. These distributions are ultimately determined by the energy and entropy characteristics of the corresponding regions of the configurational space. The reliability and accuracy of reproduction of these distributions is, in fact, one of the criteria of the suitability of the calculated trajectories for determining the physicochemical parameters and dynamic properties of biomacromolecules. Such trajectories may be obtained for a given group of molecules respecting the procedure discussed below.

DISCUSSION

Figure 2 presents the situation typical of all torsional angles in the given group of molecules in the form of a comparison of the two-dimensional angle distributions, for example, φ and ψ of the first amino acid residue in the modified Asp–Asp dimer. It will be seen that for the Berendsen thermostat the long (5000 ps) trajectory at a temperature of 1000 K does not enable us to speak of any complete scanning for the motion of the figurative point of the energetically accessible regions of configurational space. Squeezing of the molecule close to one of the local minima is observed. For trajectories with a length of the order of several hundreds of picoseconds the distribution over the conformations was substantially wider. However, in times of the order of 1000 ps there is quite a sharp decrease in the total energy of the molecule from the magnitude when the same value of the effective temperature is maintained. In the case of a molecule with an unfixed mass centre this is due to the previously observed effect [12] of the transfer of part of the internal energy to the forward motion of the molecule, which is a feature of the Berendsen thermostat [12]. In the case of a system with a fixed mass centre the internal energy is transferred to the rotation of the molecule as a whole. Similarly, on fixing one (extreme) atom, energy is transferred to the degree of freedom corresponding to rotation of the molecule about the point of fixation. And only on fixation of the complete moment of the pulse of the molecule do the effects of this type weaken. A similar picture is also observed when, for example, two right extreme atoms are fixed in the modified dimer, which is connected with braking of rotation about the peptide bond (Fig. 1). But in these cases, too, effects of serious disturbance of the equidistribution of thermal energy over the degrees of freedom are noted (Fig. 2*b*). This applies particularly to the angles φ and to a somewhat lesser degree χ_1 in both amino acid residues where considerable squeezing of the system close to one of the local minima is observed. And only an increase in temperature to 3000 K will lead to sufficiently complete scanning of the configurational space due to an increase in the interaction between the degrees of freedom as the energy of the molecule increases. The physical nature of the feature of the Berendsen thermostat described is linked with the incorrect introduction of friction into the mechanical equations, not taking into account the known relationships between the intensity of the fluctuations and dissipation (the fluctuational–dissipative theorem) [13]. As a result when the system moves in phase space over the energy surface within a narrow energy layer), the kinetic energy is so redistributed over the degrees of freedom that the interaction between modes becomes minimal. Since the interaction of the internal degrees of freedom with the forward motion of the molecule or its rotation as a whole hardly depends on the size of the total pulse or the moment of the pulse of the system, all the internal energy of the unfixed molecule in the Berendsen thermostat is gradually transferred to these isolated degrees of freedom. When the molecule is fixed, there are no completely isolated degrees of freedom and the scenario of the development of events is somewhat different. When there is motion in the Berendsen thermostat there is effective "freezing"; to be more precise, localization of those degrees of freedom, which, close, for example, to one of the positions of

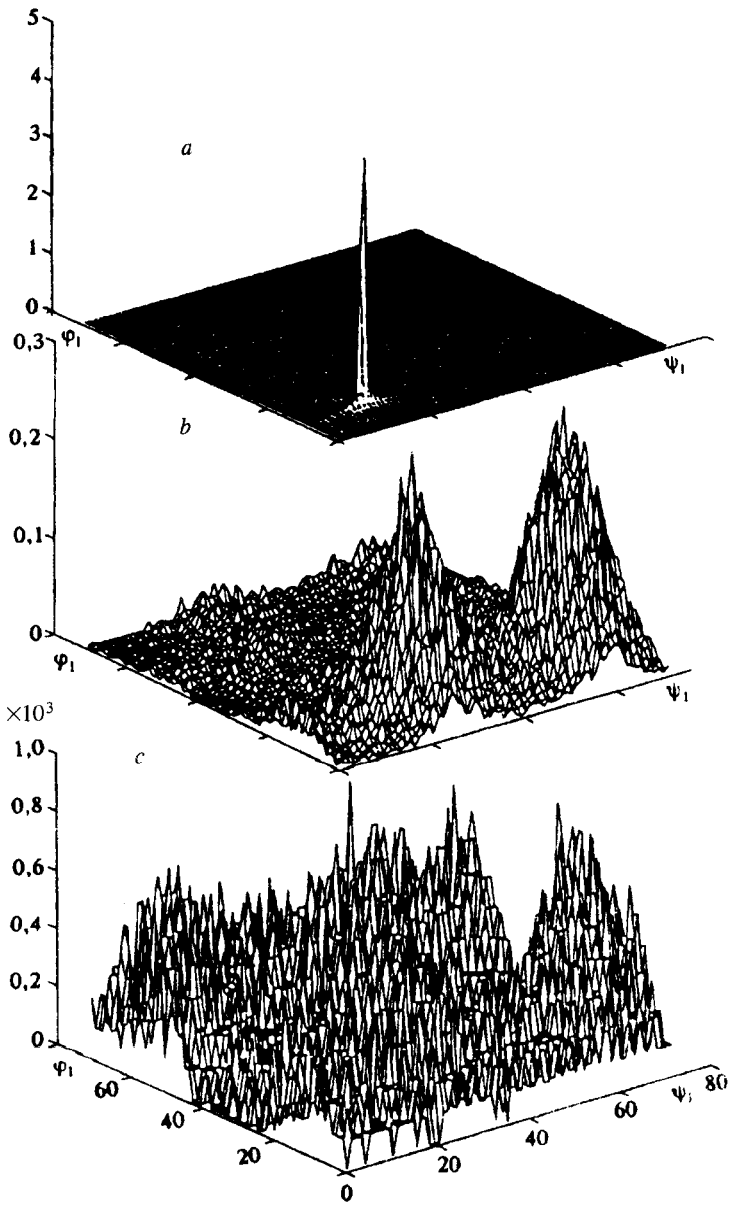

Fig. 2. Two-dimensional distribution functions of the conformational substates of the angles φ and ψ of the first residue in the modified Asp–Asp dipeptide. The range of variation in the angles from $-180°$ to $180°$ are given on a linear scale. Length of trajectory 5000 ps, $T = 1000$ K: *a*, calculation by the Berendsen method, the mass centre is not fixed (a similar result is also obtained in the case of a fixed mass centre and fixation of one (extreme) atom); *b*, same, total moment of the pulse is fixed (a similar result is obtained on fixation, for example, of the last two atoms); *c*, collisional dynamics.

equilibrium, turn out to be far less bound to the other modes. The probabilities of the events discussed in this context are proportional to the volume of the corresponding relatively narrow regions of phase space and the corresponding transition becomes observable for a sufficient increase in the length of the trajectory (Fig. 3). In the versions of the Berendsen thermostat and microcanonical ensemble for not too high energies, there is practically no mechanism for ejecting

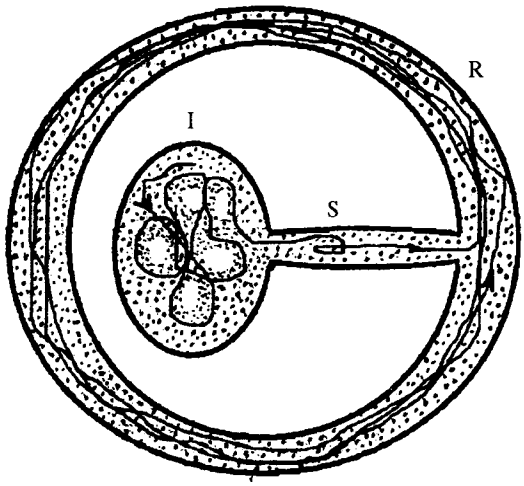

Fig. 3. Schematic representation of the trajectory and section of the hypersurface of the level E of the potential energy of the molecule ($U(q)<E$) [10, 14]. The system begins to move in the region I roughly corresponding to an equi-energy distribution over the internal degrees of freedom. Because of collisions with the potential barriers (unshaded regions) there is a small but finite probability of passage through the narrow neck S to the low-entropic state due, for example, to transfer of the internal kinetic energy to rotation of the molecule (motion in the "rift" R). The reverse transition in the version of the Berendsen thermostat is practically impossible in view of the absence of collisions or other factors for a sharp change in the direction of motion.

the system from such a narrow rift, since the temperature fluctuations for the relaxed trajectory are low and the system turns out to be in a state with low internal energy and very low entropy.

However, in the case of collisional dynamics, the fluctuations of the pulses of atoms on collisions with the particles of the medium lead to ejection of the system from the states with low entropy already at temperatures of the order of 1000 K and the probability distribution becomes far smoother and physically plausible (Fig. 2*c*).

Figure 4 shows a typical comparison, for example, of the one-dimensional distributions for the torsional angles φ obtained at different temperatures and lengths of the trajectories within the method of collisional dynamics. It can be seen that in times of the order of hundreds of picoseconds at room temperatures (often used in calculations of the condition!), even a relatively low molecular system does not have time to scan the obviously accessible regions of configurational space. At the same temperatures even 5000 ps is not enough for this. A rise in temperature to 1000 K for such a length of trajectory leads for such systems to a statistically reliable result. Likewise, one may increase, for example, the temperature to 1500 K and somewhat shorten the length of the trajectory (Fig. 4*d*). However, in this case the result statistically will be worse, since the sampling volume decreases, another consequence of which is obviously high asymmetry of the distribution (Fig. 4*d*). It is important to note that the ratio of the probabilities in the distribution maxima (Fig. 4*c*) corresponding to the potential energy minimum in respect of the angle φ to the value of the probability at the minimum of this distribution agrees well with the expected value $\exp(\varepsilon/k_B T)$ at $T=1000$ K, where $\varepsilon \sim 4$–5 kcal·mole is the height of the potential barrier for rotation about the torsional angle. This indicates the manifestation of the properties of ergodicity in the system considered in the conditions chosen. Thus, the choice of the method of introducing the thermostat is of great importance and may crucially affect the dynamics of the

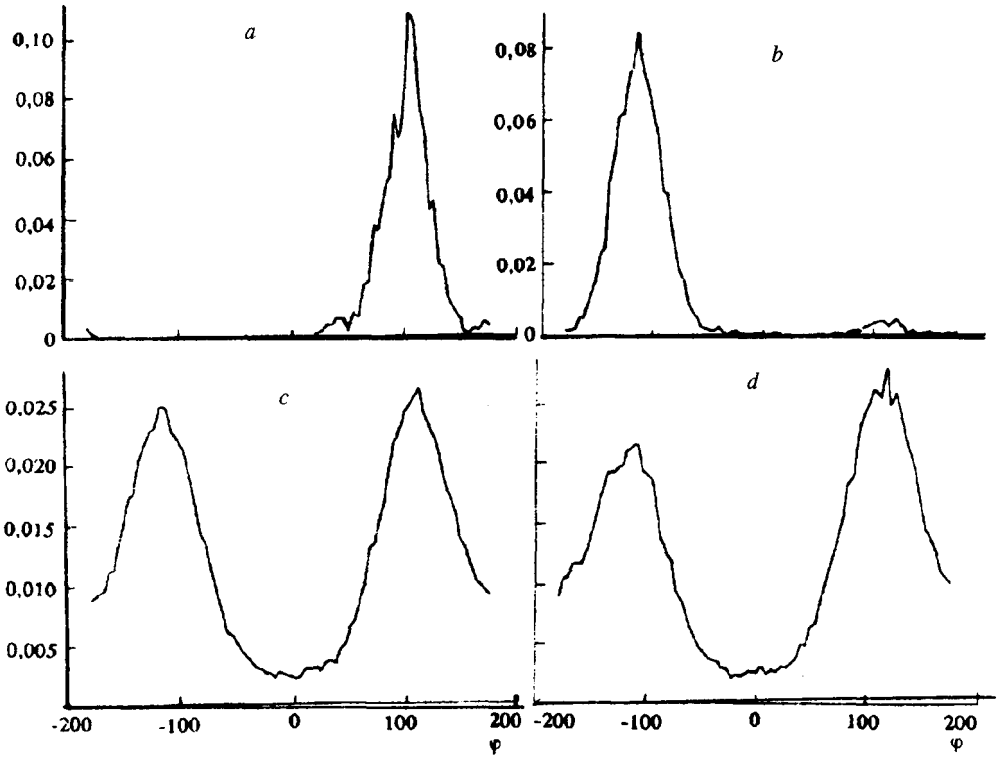

Fig. 4. Distribution functions of the conformational substates over the angle φ of the first residue in the modified Asp–Asp dipeptide. Collisional dynamics. Range of variation in the angle φ from −180° to 180°: *a*, length of trajectory 100 ps; $T=300$ K; *b*, length of trajectory 5000 ps; $T=300$ K; *c*, length of trajectory 5000 ps; $T=1000$ K; *d*, length of trajectory 3000 ps; $T=1500$ K.

system. For a series of modified dipeptides the regime that is optimal in terms of obtaining statistically significant results necessary for calculating the physicochemical properties is the regime of collisional dynamics and the choice of the lengths of trajectories of the order 5000 ps at $T=1000$ K. Conversion of the results to the case of lower temperatures may be based on scale conversions using the corresponding Boltzmann factors. The regime of collisional dynamics is undoubtedly preferable for long trajectories necessary for studying the physical characteristics of macromolecules. The Berendsen thermostat can obviously be used when searching for the local minima (conformational substates) on the hypersurface of the potential energy of the macromolecule and what is more relevant for detecting quasi-isolated internal degrees of freedom.

In conclusion, we wish to express our gratitude to the Russian Foundation for Basic Research and the Siberian Commercial Bank (grant No. 95 04 12197a) and also the "Fizmat" Programme of Goskomvuz of the Russian Foundation for partial financial support of this research.

REFERENCES

1. J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge (1987).
2. Ch. L. Brooks III, M. Karplus and B. M. Pettitt, *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*, John Wiley, New York (1987).
3. I. Schlichting, J. Berendzen, G. N. Phillips Jr and R. M. Sweet, *Nature*, **371**, 808 (1994).
4. N. K. Balabayev, A. S. Lemak and K. V. Shaitan, *Molek. Biol.* **30**, (in press).
5. B. Gelin and M. Karplus, *Biochemistry*, **18**, 1256 (1979).

6. N. K. Balabayev, A. A. Nurislamov, I. V. Uporov and K. V. Shaitan, *Mathematical and Computational Methods in Biology. Biomolecular Systems*, Pushchino (1987).
7. R. Czerminski and R. Elber, *Proc. Nat. Acad. Sci. USA*, **86**, 6963 (1989).
8. R. Susnov, C. Schutt, H. Rabitz and S. Subramaniam, *J. Comp. Chem.*, **15**, 947 (1994).
9. T. C. Beutler and W. F. van Gunsteren, *Chem. Phys. Lett.*, **237**, 308 (1995).
10. K. V. Shaitan, *Biofizika*, **39**, 949 (1994).
11. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, **81**, 3684 (1984).
12. A. S. Lemak and N. K. Balabaev, *Molecular Simulation*, **13**, 177 (1994).
13. P. Rezibua and M. De. Lerner, *The Classical Kinetic Theory of Liquids and Gases*, 423 pp., Mir, Moscow (1980).
14. M. D. Yermolayeva and K. V. Shaitan, *Biofizika*, **41**, 1168 (1996).