*ISSN 0006-2979, Biochemistry (Moscow), 2007, Vol. 72, No. 5, pp. 529-539. © Pleiades Publishing, Ltd., 2007.*
*Original Russian Text © N. T. Moldogazieva, A. A. Terentiev, A. N. Kazimirsky, M. Yu. Antonov, K. V. Shaitan, 2007, published in Biokhimiya, 2007, Vol. 72, No. 5, pp. 655-667.*

# Conformational Dynamics of Human α-Fetoprotein-Derived Heptapeptide LDSYQCT Analogs

**N. T. Moldogazieva[1]*, A. A. Terentiev[1], A. N. Kazimirsky[1], M. Yu. Antonov[2], and K. V. Shaitan[2]**

*[1]Russian State Medical University, ul. Ostrovityanova 1, 117997 Moscow, Russia;*
*fax: (495) 434-0588; E-mail: nmoldogazieva@mail.ru; aaterent@mtu-net.ru*
*[2]Department of Bioengineering, Lomonosov Moscow State University, 119992 Moscow, Russia;*
*fax: (495) 939-2374; E-mail: shaitan@moldyn.org*

Received August 11, 2006
Revision received January 18, 2007

**Abstract**—Conformational dynamics of a biologically active fragment of α-fetoprotein, the heptapeptide LDSYQCT, and its analogs obtained by site-directed substitutions of amino acid residues were studied. The conformational dynamics of the peptide were conservative under the substitutions Y17F, Y17S, and D15E. Substitutions C19A and S16V resulted only in local changes in the dynamic behavior of the peptide. Chemical modification of cysteine (C19) or dimerization of the peptide by producing a disulfide bond between cysteine residues of two parallel peptide chains, as well as the substitutions C19G, C19S, Q18E, and D15N changed a set of possible conformations and dynamic behavior of all amino acid residues. The most significant changes were caused by substitution of uncharged amino acid residues by charged ones, and *vice versa*.

**DOI:** 10.1134/S0006297907050094

*Key words*: molecular dynamics, α-fetoprotein, site-directed substitutions, peptides

α-Fetoprotein (AFP) is a major oncofetal protein not only of all mammals but, as it is currently suggested, also of all vertebrates [1-3]. Intensive investigations carried out during the last four decades have revealed detailed physicochemical and immunochemical characteristics of AFP. Primary structures of AFP are completely determined for ten biological species, including human, chimpanzee, gorilla, horse, dog, pig, marmot, mouse, rat, and chicken (Swiss-Prot/TrEMBL and GenBank databases). However, the biological role of AFP, including the mechanism of its functioning, is still unclear.

Comparison of the amino acid sequences of AFP and of some physiologically active proteins revealed functionally important sites in AFP. Determination of the location of these sites and their amino acid composition and biological activity allowed us to design a structural–functional map of AFP [4]. Some of human AFP-derived peptides have been synthesized and tested for biological activity. One of these peptides is fragment LDSYQCT (aa 14-20) [5]. This peptide was detected as an AFP fragment similar to a part of the receptor-binding site of human epidermal growth factor (EGF). Later the heptapeptide LDSYQCT was shown to be a biologically active region of human AFP [6]. However, the mechanism of action of this heptapeptide, including the role of individual amino acid residues, remains unclear.

The purpose of the present work was to study the possible role of individual aa in conformational dynamics and functioning of the heptapeptide LDSYQCT using molecular dynamics (MD) simulation. For this purpose, some amino acid residues of this peptide were substituted, and the effects of these substitutions on conformational properties and dynamic behavior of amino acid residues in the resulting analogs were studied.

Site-directed mutagenesis based on substitution of a single amino acid in proteins and peptides followed by testing for biological activity of the resulting analogs is a commonly accepted approach for investigation of the structural and functional role of individual amino acid residues. Substitution of a single amino acid in the polypeptide chain may considerably change its biological activity, including ligand or receptor binding [7, 8]. Functional changes may be accompanied by conformational changes altering the availability of amino acid

---

*Abbreviations*: AFP) α-fetoprotein; MD) molecular dynamics.
* To whom correspondence should be addressed.

residues during their interaction with other groups (see review [9]).

The intrinsic dynamics of biomacromolecules are also important for their functioning. Changes in the intrinsic dynamics may be promoted by protein interactions with other molecules and result in changes in the binding thermodynamics [10]. MD simulation study is now widely used for investigation of conformational and dynamic properties of various biomacromolecules and mechanism of their functioning [11-13]. In the present work, conformational properties and dynamic behavior of amino acid residues in the heptapeptide LDSYQCT and its analogs were studied by the method of equilibrium MD. Conformational dynamics of the peptides were evaluated using two-dimensional (2D) and three-dimensional (3D) maps of free energy levels (Poincare cross-sections) and autocorrelation functions of dihedral angles φ, ψ, and χ. Dynamics of the AFP-derived peptide and its analogs were studied using an implicit solvent model with parameters (effective viscosity, etc.) corresponding to characteristics of aqueous medium [14, 15]. This was dictated by AFP being a secreted protein active in aqueous medium.

Using the substitutions Y17F and Y17S, it was shown that neither aromatic ring of tyrosine nor its OH-group affected the conformational dynamics of the peptide LDSYQCT. The substitution S16V demonstrated that this was also true for the OH-group of serine. Hydrophobicity of cysteine (C19) seems to be crucial for maintaining the peptide conformation and dynamics, because the substitution C19A caused changes only in the position of C19, whereas the substitution C19S led to changes in the conformational properties and dynamics of all amino acid residues in the peptide. Chemical modification of cysteine or dimerization of the peptide due to formation of disulfide bond between cysteine residues in two parallel peptide chains also changed the conformational properties and dynamic behavior of all amino acid residues. Presence or absence of negative charge on the COO-group of dicarbon amino acids was the most important for maintaining the conformational dynamics of the peptide under conditions of MD calculations. This was concluded because the substitutions Q18E and D15N caused the most significant changes in the conformational properties and dynamic behavior of amino acid residues in the peptide.

## MATERIALS AND METHODS

**Constructing the heptapeptide LDSYQCT analogs.** Analogs of the LDSYQCT peptide were obtained by site-directed substitutions of amino acid residues taking into account differences in physicochemical features of the amino acids and their conservation during evolution. The AFP regions of ten biological species corresponding to human AFP-derived peptide LDSYQCT, were revealed using multiple alignment program ClustalW (version 1.82). As shown in Table 1, five (L, D, S, Q, and C) of seven amino acid residues are preserved in all species of mammals, and two amino acid residues (Y and T) are substituted. Y is substituted by an amino acid residue with similar physicochemical properties, namely, by an aromatic aa (F) or a hydrophilic one (S), which contains OH-group similarly to Y. The last aa (T) is substituted by those with different physicochemical properties in different biological species, but in the species evolutionarily

**Table 1.** Amino acid sequences corresponding to the heptapeptide LDSYQCT in α-fetoproteins of different biological species

| Names of biological species | Amino acid sequences | Numbers of amino acid residues in polypeptide chain | Identity degree, % | SwissProt/TrEMBL database accession numbers |
|---|---|---|---|---|
| Human | L D S Y Q C T | 14-20 | 7/7 (100) | P02771 |
| Chimpanzee | L D S Y Q C T | 14-20 | 7/7 (100) | Q27789 |
| Gorilla | L D S Y Q C T | 14-20 | 7/7 (100) | P28050 |
| Pig | L D S **S** Q C **S** | 32-38 | 5/7 (71) | Q8MJ76 |
| Dog | L D S **S** Q C **S** | 32-38 | 5/7 (71) | Q8MJU5 |
| Horse | L D S **F** Q C **S** | 14-20 | 5/7 (71) | P49066 |
| Marmot | L D S **S** Q C **S** | 32-38 | 5/7 (71) | Q7TSF3 |
| Mouse | L D S **S** Q C **V** | 14-20 | 5/7 (71) | P02772 |
| Rat | L D S **S** Q C **P** | 14-20 | 5/7 (71) | P02773 |
| Chicken | L D **V Q** Q **I Q** | 28-34 | 3/7 (43) | P84407 |

Note: The amino acid residues are numerated for mature molecules, except AFP of pig, dog, and marmot, for which there are no data on the signal peptide length. The amino acid residue substitutions are indicated in bold print. The sites were revealed by multiple alignment of full polypeptide chains using the ClustalW program (version 1.82).

**Table 2.** Analogs of human AFP-derived heptapeptide LDSYQCT

| Peptide designation | Substitution performed | Amino acid composition of peptide | Changes in physicochemical properties of peptide due to aa substitutions | Conformational and dynamic behaviour of peptides |
|---|---|---|---|---|
| P5 | original peptide | LDSYQCT | – | – |
| P1 | Y17F | LDS**F**QCT | hydrophilic aromatic/ hydrophobic aromatic | no essential changes |
| P2 | Y17S | LDS**S**QCT | hydrophilic aromatic/ hydrophilic aliphatic | no essential changes |
| P3 | Q18E | LDSY**E**CT | hydrophilic uncharged/ hydrophilic charged | changes for all amino acid residues |
| P4 | C19A | LDSYQ**A**T | hydrophobic uncharged/ hydrophobic uncharged | changes in positions of amino acid residue substitutions |
| P6 | C19G | LDSYQ**G**T | hydrophobic uncharged/ lacking side chain | changes for all amino acid residues |
| P7 | C19S | LDSYQ**S**T | hydrophobic uncharged/ hydrophilic uncharged | changes for all amino acid residues |
| P8 | S16V | LD**V**YQCT | hydrophilic uncharged/ hydrophobic uncharged | changes in positions of amino acid residue substitutions |
| P9 | D15E | L**E**SYQCT | hydrophilic charged/ hydrophilic charged | no essential changes |
| P10 | D15N | L**N**SYQCT | hydrophilic charged/ hydrophilic uncharged | changes for all amino acid residues |
| P11 | modified C19 | LDSYQ**Cm**T | cysteine with a free SH-group/ cysteine with a protected SH-group | changes for all amino acid residues |
| P12 | dimer | 2LDSYQ**Cx**T | cysteine with a free SH-group/ cystine of disulfide bond | changes for all amino acid residues |

Note: The amino acid residues are numerated for the mature molecule of human AFP. Substitutions of amino acid residues are indicated in bold print. Hydrophobicity of amino acid residues was assessed according to Kyte and Doolittle [23].

related to human there is a conservative substitution T/S. Amino acid compositions of the resulting analogs and their designations are presented in Table 2. Analogs of the peptide LDSYQCT (peptide P5) were obtained using the following substitutions: Y17F, Y17S, Q18E, C19A, C19G, C19S, S16V, D15E, D15N (peptides P1-P4 and P6-P10). In addition, the P11 peptide containing a modified cysteine residue was also constructed. In this peptide, the sulfhydryl group of cysteine was protected by S-acetamidomethyl as it was done in the chemically synthesized peptide used for biological testing. Moreover, to study the effect of dimerization of the peptide on its conformational dynamics, a dimer of the LDSYQCT peptide (P12) was also constructed by formation of an intermolecular disulfide bond between the cysteine residues (C19) of two parallel peptide chains.

In all peptides, aspartic and glutamic acid residues were left in a deprotonated (charged) form. The N- and C-ends of the peptides were blocked by acetyl and N-methyl groups, respectively.

**Molecular dynamics protocol.** Models of molecules were studied in a full-atom approximation using an Amber 99 force field [16, 17] and periodical boundary conditions with the box dimensions of 100 × 100 × 100 Å. MD calculations were performed at the trajectory length of 10 nsec and temperature of 2000K in order to accelerate the scanning of the conformational space by representative point and to achieve quasiergodicity of the trajectories [18]. To maintain the constant temperature, in addition to a Berendsen thermostat, a collision thermostat was used, based on interaction of atoms of the molecular system under study with an equilibrium ensemble of particles of definite masses at a given temperature, according to elastic collision laws [19, 20]. Using the collision thermostat ensured a statistically valid energy distribution on degrees of freedom [21]. Masses of the particles were taken as 18 amu, the average frequency of their collisions with atoms of the molecule under study was taken as 55 $\text{psec}^{-1}$. Viscosity of such a medium is close to the viscosity of water under normal conditions. MD calculations were performed at a dielectric constant $\varepsilon = 1$. For numerical integration the Verle algorithm was used. Initial velocities of the atoms were set by a random number generator using Maxwell's distribution.

**Processing of trajectories.** The relative positions of atomic groups in a polypeptide chain are described using torsion angles $\varphi$, $\psi$, and $\chi$ (or angles of rotations around the valence bonds N–$C_\alpha$, $C_\alpha$–C′, and $C_\alpha$–$C_\beta$, respectively), changes of which make the major contribution to flexibility of the polypeptide chain. We consider here only the angle $\chi_1$ designated further as $\chi$. To evaluate densities of conformation probabilities under changes in the torsion angles $\varphi$, $\psi$, and $\chi$, 2D and 3D maps of the free energy levels (Poincare cross-sections) were used which, unlike the maps of potential energy levels (Ramachandran maps) [22], take into account contribution of the entropy factor to stabilization of conformations. The probability densities were determined with fixed values of two or three angular variables, and the others were averaged.

The 2D- and 3D-distributions of the probability densities of conformations for all combinations of the angles $\varphi$, $\psi$, and $\chi$ were calculated by the formula:

$$p(\alpha_n, \alpha_m, \ldots) = \int_{-\pi}^{\pi} \ldots \int_{-\pi}^{\pi} p(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_N) \prod_{\substack{i=1 \\ i \neq n, m, \ldots}}^{N} d\alpha_i, \quad (1)$$

where $\alpha_n, \alpha_m$ is a set of dynamic variables and $p(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_N)$ is the density of probability to find the system in a given point of the conformational space.

Dynamic behavior of individual conformational degrees of freedom of a molecule was assessed using a special type of autocorrelation functions of the torsion angles. Normalized autocorrelation functions were calculated according to formula:

$$F_{xx} = \left\langle e^{i\varphi(t)} e^{-i\varphi(t+\tau)} \right\rangle - \left| \left\langle e^{i\varphi(t+\tau)} \right\rangle \right|^2, \quad (2)$$

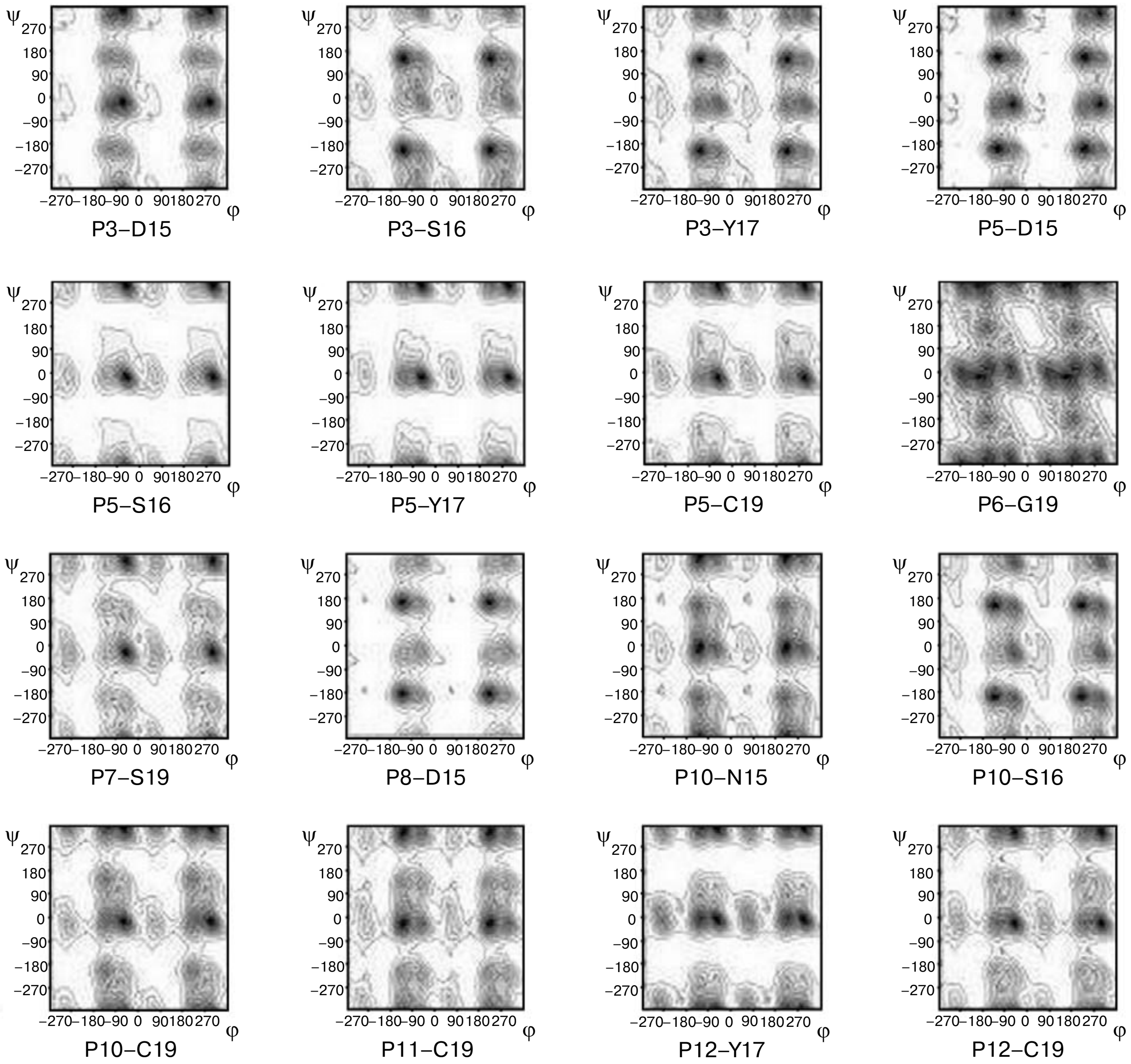
where $\varphi$ are the torsion angle values at the time points $t$ and $t + \tau$. The necessary information is contained in the time dependence of the real part of the autocorrelation function $\mathrm{Re}(F_{xx})$. Autocorrelation functions reveal characteristic movement times and dynamic behavior of the torsion angles.

## RESULTS

**Analysis of 2D and 3D maps of free energy levels.** Figure 1 presents 2D maps of free energy levels (Poincare cross-sections) for amino acid residues of the heptapeptide LDSYQCT (peptide P5) and its analogs (peptides P1-P4 and P6-P12), which demonstrate the probability densities of conformations under changes of the dihedral angles $\varphi$ and $\psi$. The darkest areas correspond to free energy minima with the maximum probability densities of conformations. In the figure, 2D maps of the aspartic acid residue (D15) in the original peptide P5 have an elongated profile with two local free energy minima which correspond to $\varphi = -60$, $\psi = -30$ and $\varphi = -120$, $\psi = 150$ connected by a neck at $\varphi = -90$, $\psi = 60$. The first minimum corresponds to the conformation of the right $\alpha$- and $3_{10}$-helices, and the other one is located in the $\beta$-structure region. The transition between these conformations takes place through forming of the right $2_7$-helix. The 2D maps of serine (S16), tyrosine (Y17), and cysteine (C19) residues have one deep minimum at $\varphi = -60$, $\psi = -30$ (the right $\alpha$- and $3_{10}$-helices) connected with an additional shallow locus at $\varphi = 60$, $\psi = -30$ corresponding to the left $3_{10}$-helix.

The peptides P1 and P2 were obtained by tyrosine substitution by phenylalanine or serine, i.e., Y17F and Y17S, respectively. These substitutions do not significantly change the distribution pattern of free energy levels for all residues under consideration, except arising of necks between the free energy minima in 2D maps of cysteine (C19) in the P1 peptide at $\varphi = -120$, $\psi = -120$, and in the case of the peptide P2 at $\varphi = 120$, $\psi = -45$. This suggests a possibility of conformational transitions. Note that in the P2 peptide there are double serines, i.e. two serine residues, one after the other, namely S16 and S17. The 2D maps of these residues are little different between themselves, and also from the 2D map of the S16 residue in the P1 peptide and the original P5 peptide.

In the peptide P3 glutamine is substituted by glutamic acid, i.e. Q18E, which results in two negatively charged amino acid residues in the peptide: D15 and E18. In total, this increases the set of probable conformations for the residues located between these two negatively charged amino acids, possibly because of their electrostatic repulsion (Fig. 1). In 2D maps of serine (S16) and tyrosine (Y17) residues in peptide P3 the minimum at $\varphi = -60$, $\psi = -30$ is less pronounced than in the maps of these residues in peptide P5, but a new deep locus appears at $\varphi = -150$, $\psi = 150$, which corresponds to formation of $\beta$-structure. Consequently, the substitution Q18E results in changes of conformational possibilities of amino acid residues at positions 16 and 17. The 2D map for E18 is substantially similar to the 2D map of aspartic acid D15 with a reduced depth of the minima compared to D15 in the original peptide P5. This is illustrated by 3D Poincare cross-sections (Fig. 2). The lower limit of probability is here 1%, which corresponds to the change in free energy by 2.74 kcal/mol. The darkest areas indicate the maximum probability densities of conformations with free energy minima.


P3–D15 P3–S16 P3–Y17 P5–D15

P5–S16 P5–Y17 P5–C19 P6–G19

P7–S19 P8–D15 P10–N15 P10–S16

P10–C19 P11–C19 P12–Y17 P12–C19

**Fig. 1.** Two-dimensional (2D) Poincare cross-sections for the dihedral angles φ and ψ. The darkest areas correspond to free energy minima with the maximum densities of conformation probabilities. The 2D maps are presented for different amino acid residues numerated for the mature polypeptide chain of human AFP. Thus, P3–D15 means the aspartic acid residue in the peptide P3, etc. Designations of the peptides are given in Table 2.

The peptides P4, P6, and P7 were obtained by cysteine substitutions at position 19, namely, C19A, C19G, and C19S, respectively. The cysteine substitution by alanine (peptide P4) does not considerably change the 2D maps of all residues, but the 3D maps demonstrate enhanced conformational properties of alanine as compared to those of cysteine (Fig. 2). The 2D map of the residue at position 19 in peptide P6 with cysteine substituted by glycine (C19G) is almost fully occupied, which is characteristic for glycine possessing greater conformational mobility due to absence of a side chain. However, in this peptide the 2D maps of the adjacent residues are also changed, which suggests an increase in their conformational mobilities due to absence of steric restraints. Thus, as compared with the original peptide P5, the 2D map of the Y17 residue in peptide P6 contains an additional minimum at $\varphi = -120$, $\psi = -30$ on the boundary between the right α- and π-helices.

The substitution of cysteine by serine (peptide P7) is accompanied by a slight but noticeable increase in the set of probable conformations for all residues. Note that in peptide P7 (LDSYQST) the 2D maps of two serine residues, S16 and S19, are different due to presence in the latter of two additional minima corresponding to the β-structure and P-helix conformations and also a neck between the neighboring minima (Fig. 1). Thus, S19 is

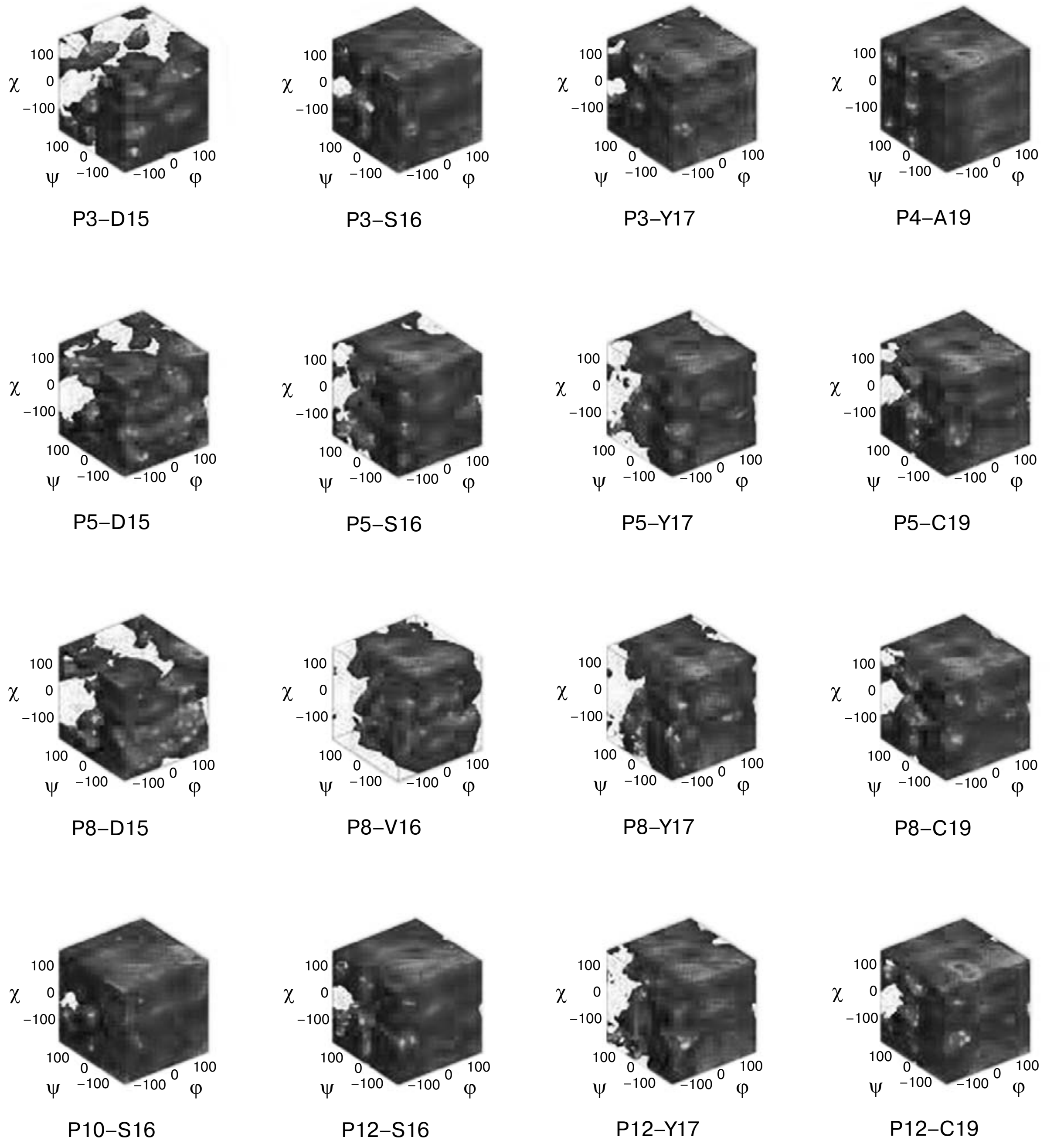


**Fig. 2.** Three-dimensional (3D) Poincare cross-sections for the dihedral angles φ, ψ, and χ. The darkest areas correspond to free energy minima with the maximum densities of conformational probabilities. The 3D maps are presented for different amino acid residues numerated for the mature polypeptide chain of human AFP. Designations of the peptides are given in Table 2.

characterized by a greater number of possible conformations than S16, and this seems to be caused by influence of the microenvironment.

Peptide P8 obtained by the site-directed substitution S16V is characterized by the presence of a large side chain. In this case, the substitution of serine by valine reduces conformational possibilities of amino acid residues at positions 15 and 16, which suggests an emergence of steric restraints (Figs. 1 and 2). However, the substitution S16V has virtually no effect on the 2D and 3D maps of other residues in this peptide.

Peptides P9 and P10 were obtained by substitution of aspartic acid at position 15 by glutamic acid or asparagine, D15E and D15N, respectively. The substitution by glutamic acid, which has a longer side chain due to an additional $CH_2$-group and the negative charge, does not markedly change the set of possible conformations of amino acid residues at position 15 and the other positions. Considerable changes in the 2D maps of nearly all residues are observed in peptide P10 obtained by substitution of the negatively charged amino acid (D) by the uncharged one (N) with the side chain of the same length. Peptide P10 differs from the other peptides by the absence of charged amino acid residues, and this increases conformational mobilities of all aa.

Chemical modification of cysteine (peptide P11) and also formation of the disulfide bond (peptide P12) enhance a set of probable conformations under variations of the dihedral angles φ and ψ for the cysteine residue C19 (Fig. 1). The conformational mobilities of other amino acid residues are virtually unchanged, except for Y17 in the dimeric peptide P12, which is characterized by an increased set of probable conformations.

**Analysis of autocorrelation functions.** Plots of autocorrelation functions calculated by formula (2) contain information about dynamic parameters of amino acid residues and are characterized by the relaxation time (τ) and the residual correlation value. The characteristic relaxation time of the autocorrelation function is the time of conformational transition under a dihedral angle change, and the residual correlation value contains information about the mobility restraints in the potential well. Figure 3 demonstrates time dependences of the real part of the autocorrelation function for the dihedral angles, which indicate that the residual correlation value depends on both the types of residue at a particular position and the dihedral angle. For the residue at position 15, the maximum residual correlation value is observed under variation of angle φ (Fig. 3a). For all peptides, the residual correlation value is close to 0.4, except P10 (with the value of 0.2), which contains the substitution D15N. Consequently, in this peptide asparagine has a higher dynamic mobility than aspartic acid in the original peptide LDSYQCT. The characteristic relaxation time for all of the peptides varies from 10 to 20 psec.

For the dihedral angle ψ, the residual correlation values for residues at position 15 (Fig. 3b) are approximately zero (this evidences the possibility of virtually full rotation on this angle) in all peptides except P8 (with the value of 0.2). Thus, the substitution of serine by valine results in limitation of mobility of the adjacent aspartic acid residue (D15). The residual correlation value is the least for peptides P7 and P10 that contain the substitutions C19S and D15N, respectively. Variations of the angle ψ also suggest that asparagine in peptide P10 has higher mobility than aspartic acid in peptide P5. Moreover, the substitution C19S increases the mobility of aspartic acid (D15). Under changes in the angle χ, the residual correlation values are close to zero for all peptides except P3, which has the substitution Q18E and the maximum residual correlation value of 0.2. The characteristic relaxation time for angles ψ and χ is 20-30 psec for all of the peptides.

For residues at position 16, the residual correlation values of the autocorrelation function for dihedral angle ψ in all of the peptides are close to 0.2, except P8 with V16 which is characterized by the greatest residual correlation value of 0.3. This confirms the limited mobility of valine compared to that of serine, which occupies position 16 in all other peptides. The characteristic relaxation time for angle ψ for the residues at position 16 varies from 20 to 40 psec. Under variation of angle ψ, pronounced scattering of residual correlation values from zero in P3 to 0.5 in P8 is observed (Fig. 3c). This confirms the limited mobility of valine and the increased mobility of serine in peptide P3 caused by the substitution Q18E. The dynamic similarity of all aa in this position is demonstrated by autocorrelation functions of angle χ with residual correlation value close to zero and characteristic relaxation time of 3-5 psec. An exception is presented only by peptide P8 (Fig. 3d) with a delayed conformational transition for valine with the characteristic relaxation time of 30 psec.

For residues at position 17, the residual correlation values for dihedral angle ψ in all of the peptides vary from 0.2 to 0.3 with the minimum value of 0.1 for peptide P12. This suggests an increase in the tyrosine mobility associated with the substitution of cysteine possessing the free sulfhydryl group by half-cystine in the disulfide bridge. The characteristic relaxation time for all of the peptides is 20-30 psec, except P10 (10 psec). Increasing of the tyrosine mobility is caused by the substitution D15N. The autocorrelation functions of angle ψ manifest the dynamic uniformity of residues at position 17 in all of the peptides, except glutamic acid, which occupies this position in peptide P3, and glutamine in peptide P10, which contains the substitution D15N (Fig. 3e). These data suggest the highest mobility of these residues.

Plots of the autocorrelation functions of dihedral angles φ and ψ for residues at position 19 reveal a specific feature of the dynamic behavior of glycine and serine (peptides P6 and P7, respectively), which demonstrate the highest mobilities. The autocorrelation functions of angle χ indicate that, in addition to glycine and serine, alanine (Fig. 3f) also has the highest mobility at this position (in peptide P4).

## DISCUSSION

Considering the capability of changing the conformational potentialities and dynamic behavior of amino acid residues, site-directed substitutions in the heptapeptide LDSYQCT may be grouped as follows: i) having no considerable effect on realization of possible conformations and mobilities of amino acid residues – these are Y17F (P1), Y17S (P2), D15E (P9); ii) changing a set of possible conformations and dynamic behavior of amino acid residues at positions subjected to substitutions and also at the adjacent positions – these are C19A (P4), S16V

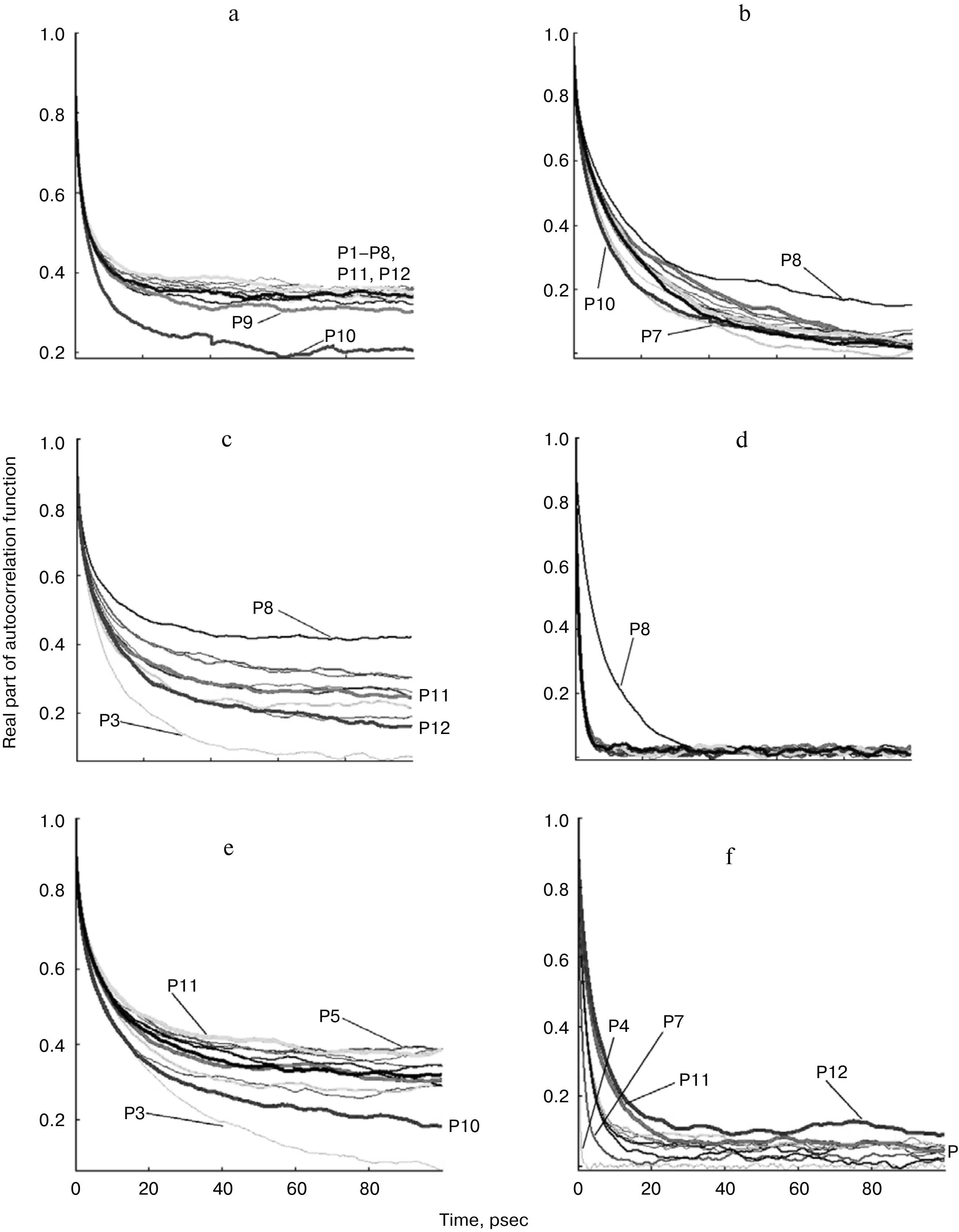


**Fig. 3.** Autocorrelation functions of the dihedral angles φ (a), ψ (b, c, e), and χ (d, f) for the residues at positions 15 (a, b), 16 (c, d), 17 (e), and 19 (f) in the peptide LDSYQCT and its analogs.

(P8). Thus, in peptide P8 the substitution S16V decreases the set of possible conformations and mobility of the residue at position 16 and the adjacent aspartic acid residue, not effecting amino acid residues at other positions; iii) changing the set of probable conformations and mobilities of amino acid residues at all positions, not only at those subjected to substitutions and the adjacent positions, but also at the remote ones. The aa may be structurally and possibly functionally bound with each other. These are Q18E (P3), C19G (P6), C19S (P7), D15N (P10) and also a chemical modification of cysteine C19 (P11) and the peptide dimerization (P12).

Conformational and dynamic changes observed in the LDSYQCT peptide analogs depend on changes in the intramolecular interactions (electrostatic, van der Waals, hydrogen bonds, etc.) of side chains with each other or with NH- and CO-groups of the polypeptide chain backbone [15]. Substitutions of uncharged amino acid residues by charged ones and *vice versa* in the heptapeptide LDSYQCT analogs are the most sensitive and significantly change a set of possible conformations and dynamic behavior of all amino acid residues in the heptapeptide.

Similar changes are observed in peptide P3 where the substitution Q18E, i.e. of glutamine with uncharged side chain by glutamic acid with the same length of the side chain but with charged group, increases the conformational mobilities of amino acid residues S16 and Y17. It seems that glutamine Q18 is involved in formation of hydrogen bonds with the polypeptide chain backbone or the side chain of another amino acid residue and this stabilizes the peptide conformation. The substitution of glutamine in peptide P3 by deprotonated glutamic acid, which is unable to participate in formation of hydrogen bonds, increases the conformational flexibility of the peptide. Moreover, substitution of the negatively charged aspartic acid by the uncharged asparagine residue, namely, D15N in peptide P10 considerably increases the set of possible conformations for all amino acid residues. Intramolecular interactions of D15 with involvement of negatively charged COO-group seem to play an important role.

In the peptides P4 and P7 the substitution of cysteine by an amino acid with comparable hydrophobicity [23], namely C19A (P4), changes the set of probable conformations only for alanine at position 19, whereas the substitution of cysteine by a more hydrophilic amino acid residue, namely C19S (P7), noticeably increases the set of probable conformations for all residues. The conformational freedom of the side chain of serine is close to that of alanine, as is shown by the plot of autocorrelation functions of the dihedral angle χ (Fig. 3f). However, in natural proteins the serine (and threonine) residue destabilizes α-helix because of tendency for formation of hydrogen bonds between the side chain of serine and the peptide group -NH [15].

Cysteine residue C19 has the highest mobility in peptides P10 (D15N), P11 (modified cysteine), and P12 (dimer). Causes of changes in the probable conformations for all amino acid residues in peptide P10 have been discussed above. The increase in mobilities of amino acid residues in peptides P11 and P12 under the cysteine modification seems to be caused by involvement of the free sulfhydryl group of cysteine in intramolecular interactions that reduces flexibility of the polypeptide chain in the original peptide P5. The SH-group of cysteine is known to be highly reactive, and in proteins cysteine is characterized by more contacts than other amino acid residues with the same length of side chain [24]. The SH-group of cysteine is also shown to form a thioester bond with the carboxyl group of glutamic acid and produce β-cysteinyl-γ-glutamyl ester [25, 26]. Human AFP contains 32 cysteine residues which form 15 regular disulfide bonds. The first two cysteine residues, one of which is in the fragment LDSYQCT, are not involved in formation of disulfide bonds and retain free SH-groups [4]. These residues can be involved in formation of disulfide bonds during AFP dimerization. Their involvement in the formation of thioester bonds is also not excluded.

The increase in the conformational mobilities of all amino acid residues in peptides P11 and P12 can be also explained by disorders in the intramolecular interactions between cysteine and aromatic amino acids. The sulfur atom in the methylsulfonyl group of methionine (possibly, also in the chemically modified cysteine) or half-cystines from the disulfide bond is involved in intramolecular interactions with aromatic rings of phenylalanine, tyrosine, and tryptophan [27, 28]. These interactions are conservative, i.e., preserved in the course of evolution and can be functionally important. Thus, in the transforming growth factor TGF-β1 there is a segment LDTNYCN (amino acid residues 2-8), which is similar to the peptide LDSYQCT and contains the disulfide bond C7-C16, half-cystines of which interact with tyrosine Y6 [27]. In the peptide LDSYQCT and its analogs, such an interaction can occur between Y17 and C19; however, these interactions are very weak and have a stabilizing effect only due to their multiplicity. Our findings are consistent with data of other authors [29] which showed that substitution by alanine of two cysteine residues involved in formation of the disulfide bond reduces the flexibility of the polypeptide chain.

Data obtained on two- and three-dimensional maps of the free energy levels are confirmed by autocorrelation functions of the dihedral angles φ, ψ, and χ. These data indicate that the substitutions Q18E and D15N (peptides P3 and P10, respectively), i.e., of an uncharged residue by a charged one with the same length of the side chain and *vice versa*, most strongly affect the dynamic behavior of all amino acid residues in a particular peptide. Peptides P7, P11, and P12 are also characterized by increase in mobilities of amino acid residues. Thus, substitutions resulting in disorders of intramolecular interactions enhance the conformational flexibility of peptides and destabilize its secondary structure. On the other hand, an increase in the conformational mobilities of amino acid residues can promote changes in functional activity of the peptide.

Comparison of autocorrelation functions of dihedral angles of amino acid residues in the heptapeptides under study and amino acid residues in monopeptides (i.e. N-methyl- and acetyl-derivatives of amino acids) studied in a collision medium [30] revealed that the residual correlation values are comparable, whereas the specific relaxation time is higher for amino acid residues in oligopeptides than in monopeptides. This may be explained by presence in oligopeptides of steric restraints for mobilities of amino acid residues and also by stabilizing intramolecular bonds.

Note that amino acid residue substitutions similar to those, which occurred during the evolution of α-fetoproteins, do not markedly change a set of possible conformations and dynamic behavior of all amino acid residues in the LDSYQCT peptide analogs. Thus, in peptides P1 and P2 the substitutions Y17F and Y17S, respectively, fail to considerably change conformational mobilities of the whole peptides. Note that the Y17F substitution takes place in horse AFP and the Y17S substitution is present in AFP of all other mammals (Table 1). Consequently, the substitutions Y17F and Y17S are permissible, despite the differences in physicochemical properties of tyrosine, phenylalanine, and serine (Table 2).

There are structural and functional restrictions for amino acid substitutions in proteins during their evolution [31]. These restrictions are determined by preservation of the secondary structure, availability of amino acid residues for solvents, and hydrogen bonds (mainly between side chains and the NH-group of the polypeptide chain backbone). The preservation of such structural properties of proteins in local environment of functionally important sites is necessary for maintaining their biological activity. This is ensured by those amino acid residues, which promote the safety of α-helix and β-sheets in proteins, are not accesible for solvents, and are involved in formation of hydrogen bonds. Thus, mutations resulting in such substitutions of amino acid residues, which lead to changes in the protein conformation, disorders of spatial organization of amino acid residues in the functionally important sites. They are not secured in the population, and are eliminated during evolution. Because dynamic features of proteins are essential for their functioning, it was suggested that dynamic restrictions of amino acid substitutions during evolution should also exist. This means that there should be restrictions preventing amino acid residue substitutions resulting in changes in conformational properties and dynamic behavior of amino acid residues surrounding functionally important regions.

Calculation of frequences of amino acid substitutions during evolution in natural proteins reveals that cysteine and also tryptophan and glycine are substituted less frequently than other amino acids [32]. Amino acid residues with charged groups in the side chains are also rarely substituted. Substitutions of uncharged, polar, and hydrophobic amino acid residues are more permissible. The most permissible are substitutions between the groups of physicochemically similar amino acid residues, especially between uncharged polar ones. The substitutions alanine–serine, serine–threonine, and isoleucine–valine, both direct and opposite, are the most neutral. The permissibility of a neutral hydrophilic amino acid (serine) substitution by a hydrophobic one (alanine or valine) and also impermissibility of amino acid substitution with similar size side chains and hydrophobicity value can be explained by specific features of their conformational dynamics. These data also seem to support the hypothesis about existence of not only structural but also dynamic restrictions in amino acid substitutions.

This work was supported by the Russian Foundation for Basic Research (project Nos. 04-04-49645 and 06-04-08136), the Russian Ministry of Science and Education (National Project of the Innovation Education in Moscow State University), and CRDF (2803).

## REFERENCES

1. Abelev, G. I. (1971) *Adv. Cancer Res.*, **14**, 295-358.
2. Tatarinov, Yu. S. (1994) *Human Alpha-Fetoprotein. 30 Years from Basic Research to Clinical Application*, RSMU, Moscow.
3. Mizejewski, G. J. (2004) *Exp. Biol. Med.*, **229**, 439-463.
4. Terentiev, A. A., and Moldogazieva, N. T. (2006) *Biochemistry (Moscow)*, **71**, 120-132.
5. Terentiev, A. A. (1997) *Vestnik RGMU*, **1**, 76-79.
6. Tatarinov, Yu. S., Terentiev, A. A., Moldogazieva, N. T., Tagirova, A. K., Kasimirskaya, V. A., and Pugacheva, O. G. (2004) *Med. Immunol. (Moscow)*, **6**, 255-256.
7. Koide, H., Muto, Y., Kasai, H., Hoshi, K., Takusari, H., Kohri, K., Takahashi, S., Sasaki, T., Tsukumo, K., and Miyake, T. (1992) *FEBS Lett.*, **302**, 39-42.
8. Murray, M. B., Tadaki, D. K., Campion, S. R., Lamerdin, J. A., Srepersu, E. H., Bradrick, T. D., and Niyogi, S. K. (1998) *Protein Eng.*, **11**, 1041-1050.
9. Moldogazieva, N. T., Terentiev, A. A., and Shaitan, K. V. (2005) *Biomed. Khim.*, **51**, 127-151.
10. Gohlke, H., Kiel, C., and Case, D. A. (2003) *J. Mol. Biol.*, **330**, 891-913.
11. Shaitan, K. V., Mikhailyuk, M. G., Leontiev, K. M., Saraikin, S. S., and Belyakov, A. A. (2003) *Biofizika*, **48**, 210-216.
12. Shaitan, K. V., Mikhailyuk, M. G., Leontiev, K. M., Saraikin, S. S., and Belyakov, A. A. (2002) *Biofizika*, **47**, 411-419.
13. Shaitan, K. V., Vasil'ev, A. K., Saraikin, S. S., and Mikhailyuk, M. G. (1999) *Biofizika*, **44**, 668-675.
14. Case, D. A., Cheatham, III, T. E., Darden, T., Golhke, H., Lou, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C.,

Wang, B., and Woods, R. J. (2005) *J. Comput. Chem.*, **26**, 1668-1688.

15. Popov, E. M. (1997) *The Problem of Protein* [in Russian], Vol. 3, Nauka, Moscow.
16. Pearlman, D. A., Case, D. A., Caldwell, J. W., Seibel, G. L., Singh, U. C., Weiner, P., and Kollman, P. A. (1991) *Amber 4.0*, University of California, San Francisco.
17. Cornell, W. D., Cieplak, P., Bayly, C., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc.*, **117**, 5179-5188.
18. Westerhoff, H., and van Dam, K. (1992) *Thermodynamics and Control of Biological Free-Energy Transduction* [Russian translation], Mir, Moscow.
19. Lemak, A. S., and Balabaev, N. K. (1995) *Mol. Simulation*, **15**, 223-231.
20. Lemak, A. S., and Balabaev, N. K. (1996) *J. Comput. Chem.*, **17**, 1685-1695.
21. Golo, V. L., and Shaitan, K. V. (2002) *Biofizika*, **47**, 611-617.
22. Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963) *J. Mol. Biol.*, **7**, 75-102.
23. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.*, **157**, 105-132.
24. Samanta, U., Bahadur, R. P., and Chakrabarti, P. (2002) *Protein Eng.*, **15**, 659-667.
25. Pangbum, M. K. (1992) *FEBS Lett.*, **308**, 280-282.
26. Zhabin, S. G., Terentiev, A. A., and Gorin, V. S. (1999) *Blood Plasma Macroglobulins: Structure*, *Biological Activity*, *Clinical Application* [in Russian], Novosibirsk.
27. Bhattacharyya, R., Pal, D., and Chakrabarti, P. (2004) *Protein Eng. Design Selec.*, **17**, 795-808.
28. Pal, D., and Chakrabarti, P. (2001) *J. Biomol. Struct. Dynam.*, **19**, 115-128.
29. Rizzuti, B., Sportelli, L., and Guzzi, R. (2001) *Biophys. Chem.*, **94**, 107-120.
30. Tereshkina, K. B., Shaitan, K. V., Levtsova, O. V., and Golik, D. N. (2005) *Biofizika*, **50**, 974-985.
31. Chelliah, V., and Blundell, T. L. (2005) *Biochemistry (Moscow)*, **70**, 835-840.
32. Majewski, J., and Ott, J. (2003) *Gene*, **305**, 167-173.