ISSN 0006-2979, Biochemistry (Moscow), 2012, Vol. 77, No. 5, pp. 469-484. © Pleiades Publishing, Ltd., 2012.
Published in Russian in Biokhimiya, 2012, Vol. 77, No. 5, pp. 583-602.

Correlation between Biological Activity and Conformational Dynamics Properties of Tetra- and Pentapeptides Derived from Fetoplacental Proteins

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

[1]Russian State Medical University, ul. Ostrovityanova 1, 117997 Moscow, Russia; E-mail: nmoldogazieva@mail.ru; aaterent@inbox.ru
[2]Lomonosov Moscow State University, 119991 Moscow, Russia; E-mail: shaitan@moldyn.org

Received August 22, 2011
Revision received December 9, 2011

Abstract—In this work, using molecular dynamics simulation, we study conformational and dynamic properties of biologically active penta- and tetrapeptides derived from fetoplacental proteins such as alpha-fetoprotein, pregnancy specific β1-glycoprotein, and carcinoembryonic antigen. Existence of correlation between flexibility of peptide backbone and biological activity of the investigated peptides was shown. It was also demonstrated that flexibility of peptide backbone depends not only on its length, but also on the presence of reactive functional groups in amino acid side chains that participate in intramolecular interactions. Peptides that demonstrate similar biological effects in regulation of proliferation of lymphocytes and expression of differentiation antigens on their surface (LDSYQCT, PYECE, YECE, and YVCE) are characterized by rigidity of their peptide backbone. Increased backbone flexibility in peptides PYQCE, YQCE, SYKCE, YQCT, YQCS, YVCS, YACS, and YACE is correlated with decreased biological activity. Conformational mobility of amino acid residues does not depend on physicochemical properties only, but also on intramolecular interactions. So, evolutionary restrictions should exist to maintain such interactions in the environment of functionally important sites.

DOI: 10.1134/S0006297912050070

Key words: alpha-fetoprotein, PSG, CEA, tetra- and pentapeptides, analogs, molecular dynamics, biological activity, flexibility of the peptide backbone

Over the last two decades it has been demonstrated that along with globular domains, there are other types of functionally important structural modules in proteins. Among such modules, short linear motifs (of 3-10 amino acid residues) have been revealed in many regulatory proteins [1, 2]. Short peptide motifs may mediate protein–protein interactions and participate in signal transduction, cell cycle regulation, etc. Revealing of similar peptide motifs in different homologous and non-homologous proteins may indicate similarity of their functions. These proteins may function in a cooperative and coordinated manner.

Our earlier investigations revealed biologically active tetra- and pentapeptide motifs in structures of several proteins of the fetoplacental complex, including pregnancy specific β1-glycoproteins (PSGs) and carcinoembryonic antigen (CEA) [3, 4]. These are some pentapeptides in human PSGs (Table 1) (PYECE, PYQCE, LYVCS, and LYACS) and also SYKCE (a.a. 213-217) in human CEA. They were revealed as sites with high degree of similarity with part of the biologically active site of human alpha-fetoprotein (AFP), namely, the heptapeptide LDSYQCT (a.a. 14-20), designated as AFP_{14-20} [5, 6]. These pentapeptides and some of their tetrapeptide fragments were synthesized by solid-phase chemistry with subsequent study of their biological activity in comparison with that of human AFP. It was shown that these pentapeptides and their tetrapeptide analogs are able to stimulate proliferation of peripheral blood lymphocytes and to influence cell surface differentiation.

In particular, human AFP and its synthetic heptapeptide LDSYQCT (AFP_{14-20}), along with synthetic pentapeptide fragments of PSG and CEA were shown to possess immunosuppressive activity. They decrease total

Abbreviations: AFP, alpha-fetoprotein; CEA, carcinoembryonic antigen; 2D, two-dimensional; 3D, three-dimensional; EGF, epidermal growth factor; MD, molecular dynamics; PSG, pregnancy specific β1-glycoprotein.

* To whom correspondence should be addressed.

Table 1. Functionally important linear motifs in human PSGs

Types of PSG	Amino acid compositions of linear motifs and their residue numbers			
	RGD	PYECE	PYQCE	LYVCS/LYACS
PSG1	127-129 (GDD)	214-218	307-311	391-395 (LYVCS)
PSG2	127-129	214-218	–	298-302 (LYVCS)
PSG3	127-129	214-218	307-311	391-395 (LYACS)
PSG4	127-129 (RRD)	214-218	307-311	391-395 (LYACS)
PSG5	127-129	214-218	307-311	298-302 (LYTCS)
PSG6	126-128	213-217	306-310	390-394 (LYACS)
PSG7	127-129	214-218	307-311	391-395 (LYACS)
PSG8	127-129 (GGD)	214-218	307-311	391-395 (LYACS)
PSG9	127-129	214-218	307-311	391-395 (LYACS)
PSG10	126-128	213-217	306-310	390-394 (LYACS)
PSG11	127-129	214-218	307-311	298-302 (LYACS)

amount of T- and B-lymphocytes and their subpopulations and reduce the number of natural killer (NK) cells in patients with infectious allergic diseases [7-10]. AFP decreases the number of activated lymphocytes increased by inflammation and decreases by 2.5 times exposure (i.e. normalizes the level) of differentiation antigens CD25, CD71, and HLA-DR and increases by 1.4 times the level of $CD95^+$-lymphocytes in patients with atopic bronchial asthma. In these patients, the peptides AFP_{14-20} and PYECE significantly decrease expression of early (CD25 and CD71) and late (HLA-DR and CD95) differentiation antigens. In patients with rheumatoid arthritis, these peptides demonstrate similar effects on expression of the above-mentioned antigens. The heptapeptide AFP_{14-20} also significantly (by 3 times) increases expression of CD95 antigen on the surface of lymphocytes in patients with infectious and allergic myocarditis in the period of aggravation and decreases by 1.7 times the level of HLA-DR antigen, while it does not influence level of other antigens [7, 8].

In general, the pentapeptides under study were divided into two groups according to their influence on expression of activation antigens on the surface of lymphocytes in patients with infectious allergic diseases: 1) peptides AFP_{14-20} and PYECE that demonstrate marked influence on number of lymphocytes carrying early and late activation antigens on their surface; 2) peptides PYQCE, SYKCE, and LYVCS that do not show any effects.

Also, the biological activity of some tetrapeptide fragments has been studied. It was shown that peptides YECE and YVCE do not influence expression of CD25, CD71, HLA-DR, and CD95 antigens on the surface of lymphocytes of healthy donors. However, at concentration of 10^{-7} M they were able to decrease by 1.4 times expression of activation antigen CD95 on the surface of lymphocytes in patients with atopic bronchial asthma and by 1.2 times in patients with infectious and allergic myocarditis [9, 10]. Peptide YECE was also able to increase expression of CD95 antigen in patients with combined pathology. Peptides YQCE and YACE did not influence expression of any activation antigens of lymphocytes [9]. Thus, it can be concluded that penta- and tetrapeptides under study possess different activities in the same test. However, molecular mechanisms underlying these differences remain unrevealed.

Interestingly, the peptides PYQCE and PYECE are analogs that differ by a single amino acid substitution Q/E. Despite this fact, these two peptides are sufficiently different in their biological activity. The same is true for the pair of peptides YQCE and YECE.

Moreover, the peptides YVCE and YECE are also analogs, but at the second position they contain amino acid residues with significantly different physicochemical properties (size of side chains, hydrophobicity degree, and charge) [11]. However, the latter peptides possess similar biological activity (the same is true for the pair of peptides YQCE and YACE). Thus, difference or similarity in biological activity of the studied peptides cannot be explained by different or similar physicochemical properties of residues located at the same position.

It is recognized that conformational and dynamic properties of proteins and peptides play an important role in their functioning. For example, conformational and dynamic properties influence interaction of proteins with other molecules, such as ligands or receptors, by changing thermodynamics of binding [12, 13]. So, study of effects of amino acid substitutions that are made in functionally important sites of proteins and peptides on their conformational and dynamic properties along with

revealing of correlation between changes in conformational–dynamic properties and biological activities are of a particular interest.

Methods of molecular dynamics (MD) are now widely used to study conformational and dynamic properties of different molecules and mechanisms of their functioning [14, 15]. These methods allow studying detailed conformational changes that take place in molecules as well as modeling intra- and intermolecular interactions as a result of a force field action at a definite atom.

We showed earlier that conformational and dynamic properties of amino acid residues in peptides depend on intramolecular interactions between functional groups of side chains or between these groups and the peptide backbone. We also demonstrated that peptides differing by single amino acid substitution (with similar physicochemical properties of both amino acid residues — to be substituted and the substitute) could significantly differ in their biological activities.

In the present work we computationally constructed the above-mentioned peptide fragments derived from AFP, PSGs, and CEA and some of their analogs. Conformational and dynamic properties of the peptides were studied using the method of equilibrium MD in an implicit water model. Changes in dynamic and conformational properties of amino acid residues in the peptides was assessed using two-dimensional (2D) and three-dimensional (3D) maps of free energy change (Poincare sections) and also of curves of autocorrelation functions of dihedral angles φ, ψ, and χ. Clustering (dispersion) analysis was performed to compare conformational and dynamic behavior of amino acid residues located at the same positions in different analogs.

METHODS OF INVESTIGATION

Computer-based modeling of heptapeptide AFP_{14-20} and tetra- and pentapeptide fragments. All the studied peptides were constructed with the HyperChem molecular modeling program [16]. Along with heptapeptide AFP_{14-20} (LDSYQCT), seven pentapeptides were constructed, four of which were derived from human PSGs. Two of these have sequences PYQCE and PYECE (designated P26 and P27, respectively) and correspond to biologically active sites of human PSGs (Table 1). Two other pentapeptides (LYACS and LYVCS) contain leucine residues at their *N*-termini (peptides P41 and P42, respectively). The fifth pentapeptide (SYKCE, P28) is a fragment of human CEA. Also, two more pentapeptides were constructed, one of which is a fragment of heptapeptide AFP_{14-20} (SYQCT) and the other is its analog (SYQCE) (peptides P44 and P43, respectively).

Tetrapeptides were considered as analogs of fragments of the heptapeptide AFP_{14-20}, namely, YQCT, in which successive point substitutions of amino acid residues were made. This was done taking into account data on similar tetra- and pentapeptide motifs in proteins from human fetoplacental complex and also differences in physicochemical properties of natural amino acids [11].

The amino acid sequences of the analogs are presented in Table 2. Conformational and dynamic properties of these analogs are studied in comparison with those of heptapeptide LDSYQCT (AFP_{14-20}) that was designated as in our previous works as peptide P5 [17, 18].

In all peptides residues of acidic amino acids (aspartic and glutamic acids) were in deprotonated (negatively charged) form while residues of basic amino acids (lysine, histidine, and arginine) were in protonated (positively charged) form. To avoid terminal effects, the *N*- and *C*-ends of all peptides were blocked by acetyl- and *N*-methyl-groups, correspondingly.

MD protocol. Models of the molecules were studied in full-atomic approximation in potential field Amber [19, 20]. Parameters of the potential field were completed by experimental data [21] and quantum chemistry calculations performed by the GAMESS program package [22]. The restricted Hartree–Fock method with basis set 6-31GF was applied.

MD calculations were performed at trajectory length of 10 nsec in a periodic box of 100 × 100 × 100 Å. To accelerate scanning of conformational space by a representative point and to reach quasi-ergodicity of trajectories [23], temperature was given equal to 2000 K. Both Berendsen and collision thermostats were used to maintain constant temperature. Use of the collision thermostat provides statistically correct distribution of energy on degrees of freedom [24]. Effects of the collision thermostat are based on interactions between atoms of the molecular system studied and the equilibrium ensemble of particles with a definite mass at a given temperature according to the elastic collision law [25, 26]. Masses of the particles were given to be equal to 18 Da, and the mean frequency of collisions was equal to 55 $psec^{-1}$. Viscosity of the collision medium is close to that of water at normal conditions. Dielectric constant ε was equal to 1. The Verlet algorithm was used for numerical integration. Initial atomic velocities were chosen by a random number generator with Maxwell distribution.

Treatment of trajectories obtained by MD calculations. *2D and 3D maps of free energy levels.* Conformational and dynamic properties of amino acid residues in the peptides were studied using the dihedral angles φ, ψ, and χ (or angles of rotations around single covalent bonds N–Cα, Cα–C and Cα–Cβ, correspondingly). 2D and 3D maps of free energy levels (Poincare sections) were used to estimate probability densities of conformations during change in values of dihedral angles φ, ψ, and χ [27]. The probability densities were determined with fixed values of two or three angular variables and with averaging of the others. The 2D and 3D distributions of probability

Table 2. Changes in conformational and dynamic properties of amino acid residues in penta- and tetrapeptides derived from fetoplacental proteins in comparison with those in the peptide AFP_{14-20}

Designation of peptide	Substitution of amino acid residue	Amino acid sequence of peptide	Changes in physicochemical properties of residues at a definite position	Changes in conformational mobility of residues in analogs
1	2	3	4	5
P5	initial peptide	LDSYQCT	–	–
P26	S16P T20E	PYQCE	hydrophilic uncharged/hydrophobic cyclized hydrophilic uncharged/hydrophilic negatively charged	restricted for P16, increased for Y17 and Q18, decreased for C19 and E20
P27	S16P Q18E T20E	PYECE	hydrophilic uncharged/hydrophobic cyclized hydrophilic uncharged/hydrophilic negatively charged hydrophilic uncharged/hydrophilic negatively charged	restricted for P16, decreased for Y17 and E18, not changed for E20, increased for C19
P28	Q18K T20E	SYKCE	hydrophilic uncharged/hydrophilic positively charged hydrophilic uncharged/hydrophilic negatively charged	increased for S16, Y17 and K18, decreased for C19 and E20
P29	–	YQCT	tetrapeptide fragment of AFP_{14-20}	increased for Y17, Q18 and C19, not changed for T20
P30	T20E	YQCE	hydrophilic uncharged/hydrophilic negatively charged	increased for Y17 and Q18, decreased for C19 and E20
P31	Q18E T20E	YECE	hydrophilic uncharged/hydrophilic negatively charged hydrophilic uncharged/hydrophilic negatively charged	decreased for Y17 and E18, not changed for E20, increased for C19
P32	Q18A T20E	YACE	hydrophilic uncharged/hydrophobic aliphatic hydrophilic uncharged/hydrophilic negatively charged	increased for Y17 and A18, decreased for C19 and E20
P33	Q18A T20S	YACS	hydrophilic uncharged/hydrophobic aliphatic hydrophilic uncharged/hydrophilic uncharged	increased for Y17, A18, C19 and S20
P34	T20S	YQCS	hydrophilic uncharged/hydrophilic uncharged	increased for Y17, Q18, C19 and S20

Table 2. (Contd.)

1	2	3	4	5
P35	Q18V T20S	YVCS	hydrophilic uncharged/hydrophobic aliphatic hydrophilic uncharged/hydrophilic uncharged	increased for Y17, C19 and S20, decreased for V18
P36	Q18V T20E	YVCE	hydrophilic uncharged/hydrophobic aliphatic hydrophilic uncharged/hydrophilic negatively charged	increased for Y17, decreased for V18, C19 and E20
P41	S16L Q18A T20S	LYACS	hydrophilic uncharged/hydrophobic aliphatic hydrophilic uncharged/hydrophobic aliphatic hydrophilic uncharged/hydrophilic uncharged	increased for L16, Y17, A18, C19 and S20
P42	S16L Q18V T20S	LYVCS	hydrophilic uncharged/hydrophobic aliphatic hydrophilic uncharged/hydrophobic aliphatic hydrophilic uncharged/hydrophilic uncharged	increased for L16, Y17, C19 and S20, decreased for V18
P43	T20E	SYQCE	hydrophilic uncharged/hydrophilic negatively charged	increased for S16, Y17 and Q18, decreased for C19 and E20
P44	–	SYQCT	pentapeptide fragment of AFP_{14-20}	increased for S16, slightly increased for Y17, Q18 and C19, not changed for T20

Notes: Numerations are given for amino acid residues in mature polypeptide chains.

densities of conformations for all angular combinations were calculated according to Eq. (1):

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

Here α_n, α_m is a set of dynamic variables, p (α_1, ..., α_i, ... α_N) is density of probability to find the system at a given point of conformational space.

Autocorrelation functions. Dynamic behavior of an individual conformational degree of freedom of molecules was evaluated using autocorrelation functions of dihedral angles of a special type. Autocorrelation functions allow judging about characteristic times of quenching and types of dynamic behavior of dihedral angles. Normalized autocorrelation functions were calculated according to Eq. (2):

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

Here φ is dihedral angle value at time intervals t and $t + \tau$. Information for analysis is contained in dependence of the real part of the autocorrelation function (F_{xx}) on time.

Clustering analysis. We performed clustering (dispersion) analysis for comparative study of conformational and dynamic behavior of amino acid residues located at the same positions in different peptide analogs. This method of statistical data processing reveals relatively

similar subjects and classifies conformational degrees of freedom. We used it to evaluate the degree of difference (or similarity) between free energy level maps or autocorrelation functions with construction of a clustering tree. Euclidian distance was used as a measure to differentiate clusters. An algorithm of selection of minimum distances is used to construct the clustering tree. Degree of similarity between 2D maps of free energy levels was calculated according to Eq. (3):

$$d_{sr} = a^2 \sqrt{\sum_i (p_{r,i}(\varphi,\psi) - p_{s,i}(\varphi,\psi))^2}. \qquad (3)$$

Here *r* and *s* correspond to different amino acid residues, *a* is clustering parameter, and *p* is probability density.

RESULTS

Analysis of 2D and 3D maps of free energy levels (Poincare sections). Evaluation of conformational changes in molecules of proteins and peptides is usually performed using a map of potential energy levels (Ramachandran map) [27]. Changes in potential energy during rotation around covalent bonds within the same residue provide determining areas of sterically allowed and prohibited conformations for a pair of dihedral angles [28].

In our work we used 2D and 3D maps of free energy levels (Poincare maps). These maps, unlike the Ramachandran maps, take into account the contribution of entropy factor into stabilization of conformations. The 2D and 3D maps demonstrate probability densities of

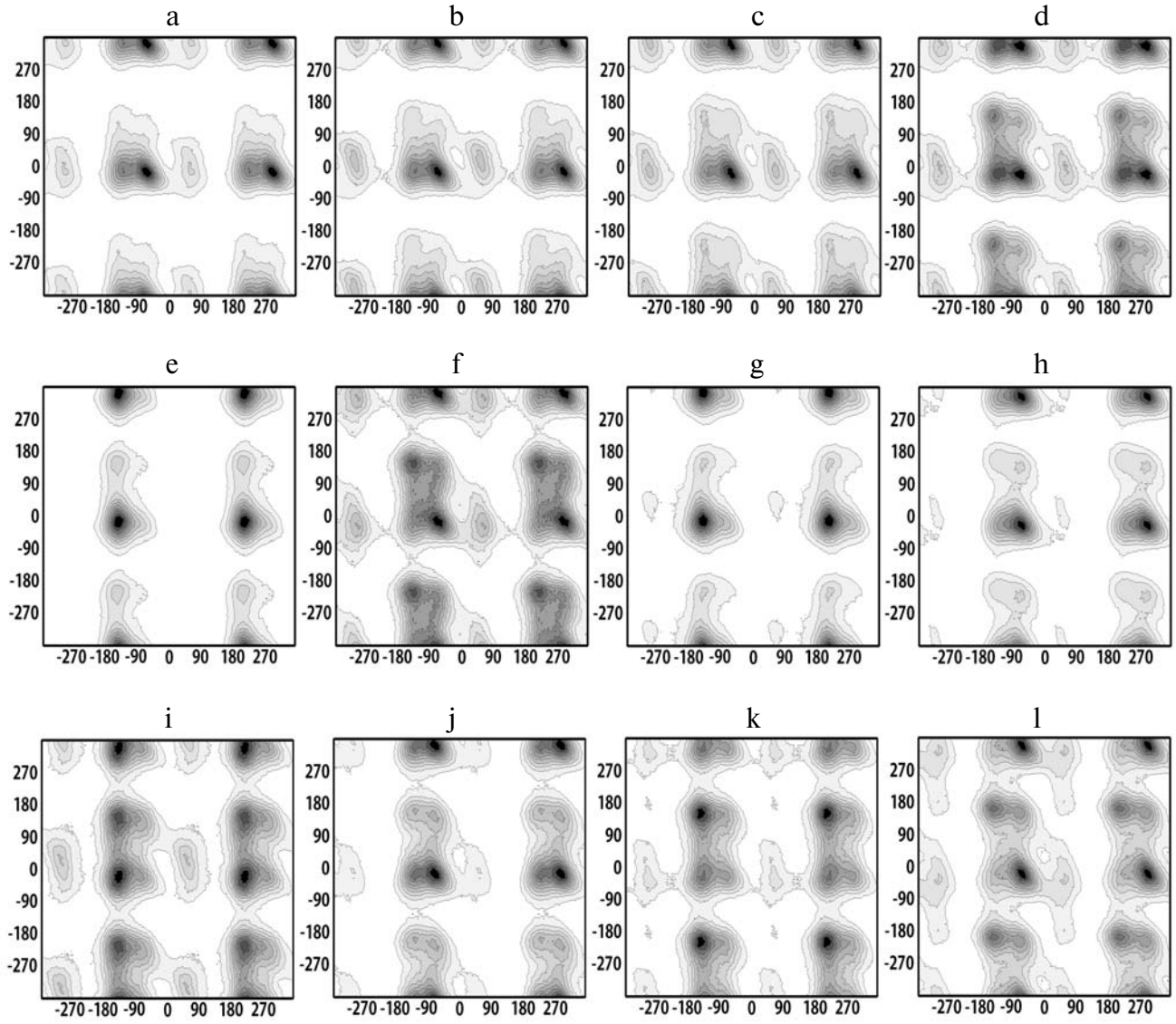

Fig. 1. Two-dimensional (2D) maps of free energy levels in coordinates (φ, ψ) for residues at positions 17 (a, e, i), 18 (b, f, j), 19 (c, g, k), and 20 (d, h) in peptides P5 (a-d), P26 (i), P27 (e), P29 (f), P30 (g), P31 (j), P32 (h), and P33 (k) and also at position 16 (l) in peptide P28.

conformations by changing values of dihedral angles φ, ψ, and χ. The darkest areas on the maps correspond to minima of free energy levels with the highest probability densities of conformations (Figs. 1 and 2).

Here we studied conformational and dynamic properties of seven pentapeptides that are similar to the part of heptapeptide LDSYQCT (AFP_{14-20}), namely, SYQCT (residues 16-20). These are peptides PYQCE, PYECE, SYKCE, LYACS, LYVCS, SYQCE, and SYQCT designated as P26, P27, P28, P41, P42, P43, and P44, respectively. Also, the short tetrapeptide fragment YQCT (residues 17-20) and its seven analogs (Table 2) obtained with consecutive point substitutions of amino acid residues (peptides P29-P36) were studied. Preference for these peptides was supported by discovery of similar amino acid sequences in other proteins of the fetoplacental complex, in particular, in human pregnancy specific β1-glycoproteins (PSGs) and carcinoembryonic antigen (CEA). We compared conformational and dynamic properties of residues in the mentioned analogs with those in heptapeptide AFP_{14-20} (peptide P5) (Fig. 1, a-d).

Pentapeptides P26 and P27 (PYQCE and PYECE, correspondingly) differ from the other peptides by the presence of a proline residue at their *N*-termini (at the position that corresponds to a.a. 16 in human AFP). On 2D and 3D maps, the proline residue demonstrates significant limitation in the set of probable conformations, and its conformational properties are similar to those in peptides PDSYQCT and PGSYQCT (peptides 22 and P23, correspondingly) that were studied earlier [29].

Pentapeptides P28, P43, and P44 retain serine residue *at position 16*. The 2D map of this residue in peptide P28 (SYKCE) contains two loci of energy minima at φ = −60°, ψ = −30° (right-handed α-helix) and φ = −135°, ψ = 150° (β-structure) (Fig. 1l). As a whole this peptide is characterized by significant increase in conformational mobility in comparison with S16 in the initial peptide P5. In this case, conformational properties of S16 in peptide P28 are similar to those in peptide P13 (IMSYICS) that was studied earlier [29]. This similarity, seemingly, is due to loss of ability of S16 in peptides P13 and P28 to form a hydrogen bond with the residue at position 18, as occurs in the initial peptide P5 (with Q18).

The 2D map of the serine residue in peptide P43 (SYQCE) also contains two loci that correspond to right-handed α-helix and β-structure and is similar to that of S16 in peptide P28. Nevertheless, the 3D maps show increasing conformational mobility of S16 in P43 in comparison with that in peptide P28. Although the 2D map of residue S16 in peptide P44 (SYQCT) contains the only deep minimum at φ = −60°, ψ = −30° along with shallow loci of β-structure and polyproline helix, its 3D map shows increase in conformational mobility if compared to that in the initial peptide P5.

Peptides P41 (LYACS) and P42 (LYVCS) contain a leucine residue at position 16 that like in the initial peptide P5 is the *N*-terminal residue. While the 2D map of residue L14 in peptide P5 contains a single locus at φ = −135°, ψ = −30° (right-handed 3_{10}-helix), in peptides P41 and P42 there are two loci that correspond to right-handed α-helix and β-structure. Some increase in conformational mobility can also be seen on 3D maps of L16 in peptides P41 and P42 in comparison with L14 in P5.

At position 17 all of the studied peptides contain a tyrosine residue (Y17). Let us first consider pentapeptides. The 2D maps of this residue in peptides P26 and P27 show change in location of the energy minimum in comparison with that in the initial peptide P5. Y17 in the initial peptide (Fig. 1a) shows a locus at φ = −60°, ψ = −30° (right-handed α-helix), when in peptide P27 at φ = −135°, ψ = −30° (right-handed 3_{10}-helix) (Fig. 1e), and in peptide P26 one more deep locus (Fig. 1i) at φ = −135°, ψ = 150° (β-structure). Significant decrease in conformational mobility of Y17 in peptide P27 and increase in the set of probable conformations in peptide P26 in comparison with that in the initial peptide P5 is observed. These differences can be explained by the presence of two negatively charged and, consequently, repulsing residues of glutamic acid (E18 and E20) in peptide P27. As a result, the tyrosine residue, being squeezed between P16 and E18, has restricted mobility.

On 2D maps of the residue Y17 in peptide P41 two loci corresponding to right-handed α-helix and β-structure may be observed, and in peptide P42 – a single locus of β-structure. The latter is, probably, explained by the presence of valine residue with a large side chain after the tyrosine residue. Nevertheless, 3D maps show significantly higher conformational mobility of Y17 in both peptides than that in the initial peptide P5.

The 2D map of Y17 in peptide P43 (SYQCE) contains two loci that correspond to right-handed 3_{10}-helix and β-structure. The 3D map shows marked increase in conformational mobility in comparison with that in peptide P5. In peptide P44 (SYQCT), the tyrosine residue shows a single locus that corresponds to right-handed α-helix and a small increase in mobility.

Let us now consider changes in conformational and dynamic properties of the tyrosine residue at position 17 in the tetrapeptides – in peptide P29 (YQCT) and its analogs P30-P36 (Table 2). Analysis of 2D and 3D maps shows that Y17 in all the studied peptides, except for P31, possesses higher than in peptide P5 (Fig. 2, a-d) set of probable conformations (Fig. 2e). This could be explained by the *N*-terminal position of the tyrosine residue in all these tetrapeptides and possible terminal effects. However, this suggestion is disproved by the decrease in mobility of Y17 in peptide P31 (Fig. 2i). Besides, our data demonstrate that there is a pairwise similarity in conformational properties of Y17 in peptides P26 (PYQCE) and P30 (YQCE) and in peptides P27 (PYECE) and P31 (YECE) (Fig. 2i). This excludes termi-

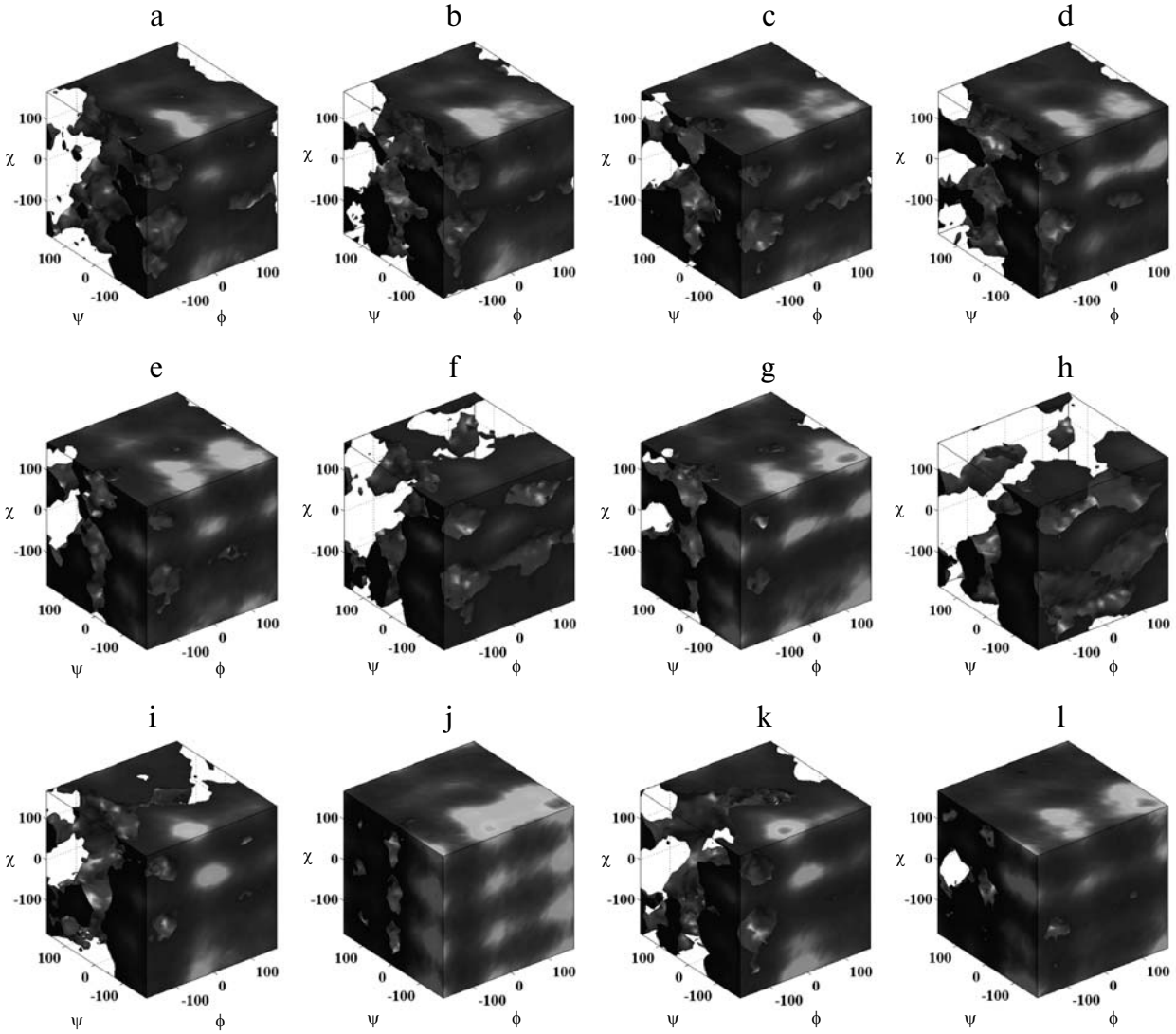

Fig. 2. Three-dimensional (3D) maps of free energy levels for residues at positions 17 (a, e, i), 18 (b, f, j), 19 (c, g, k), and 20 (d, h, l) in peptides P5 (a-d), P27 (f), P29 (e), P30 (h), P31 (i), P32 (k), P33 (g, j), and P34 (l).

nal effects. Also, the conformational mobility of Y17 in peptide P28 (SYKCE) is similar to that in the tetrapeptides despite the fact that Y17 in P28 is not the *N*-terminal residue.

Thus, our data suggest that the proline residue located before the tyrosine residue in peptides P26 and P27 does not influence conformational and dynamic properties of Y17. Increase in its mobility in most of the peptides (except P27 and P31, in which mobility of Y17 is decreased) can be explained by loss of hydrogen bonding between residues at positions 16 and 18.

At position 18 some analogs retain glutamine residue as in the initial peptide P5, while others contain substitutions Q18E, Q18A, Q18V, or Q18K (Table 2). In peptide P26 the glutamine residue was kept, but in peptide P27 the substitution Q18E was made. The 2D and 3D maps of the glutamine residue (Q18) in peptide P26 show increase in conformational mobility in comparison with that in the initial peptide P5, while for glutamic acid (E18) in peptide P27 decrease in mobility is observed (Fig. 2f). Previously [12] we studied a heptapeptide LDSYECT (P3) that contains glutamine-to-glutamic acid substitution (Q18E), which leads to change in conformational and dynamic properties for all residues in this peptide. This may be explained by the presence of two negatively charged, i.e. repulsing, residues (D15 and E18) in the peptide P3. This leads to increase in conformational mobility of residues located between them (S16 and Y17) due to decrease in steric restrictions.

In peptide P28 (SYKCE) substitution Q18K was made. This leads to increase in conformational mobility of residue at position 18. Two heptapeptides—P16 (LDKYQCT) and P25 (LDKYACN)—similarly to peptide P28 contain a lysine residue but at position 16 [17,

29]. The set of probable conformations for K16 in P16 and P25 significantly differs from that for K18 in peptide P28. This is probably a consequence of electrostatic attraction arising between oppositely charged residues D15 and K16 in peptides P16 and P25, which decreases mobility of K16.

Peptides P41 and P42 have substitutions Q18A and Q18V, respectively. Conformational mobility of the alanine is high and of the valine is low, and this is a consequence of difference in side chain sizes of these residues.

Peptides P43 and P44 retain the glutamine residue at position 18. However, 2D maps of Q18 in these peptides differ from each other: in P44, similarly to that in P5, there is a single locus of α-helix, whereas in P43 it is of β-structure. The 3D maps show increase in mobility of Q18 in both peptides in comparison with that in the initial peptide P5.

Let us consider how the picture of free energy levels is changed for amino acid residues at position 18 in the tetrapeptides. Three of them (P29, P30, and P34) retain the glutamine residue at this position (Table 2). In comparison with the initial peptide P5, change in contours of energy minima on 2D and 3D maps of Q18 in all peptides and, on the whole, increase in set of probable conformations is observed (Fig. 1f). In peptide P29 along with a deep locus at $\varphi = -60^\circ$, $\psi = -30^\circ$ (right-handed α-helix), a new deep locus at $\varphi = -150^\circ$, $\psi = 150^\circ$ (β-structure) is observed. In peptide P30 (YQCE), a locus of α-helix disappears, but two new loci arise at $\varphi = -135^\circ$, $\psi = -30^\circ$ (right-handed 3_{10}-helix) and at $\varphi = -150^\circ$, $\psi = 150^\circ$. All three loci are observed in peptide P34 (YQCS).

In peptide P31 (YECE) that contains two substitutions (Q18E and T20E), the glutamic acid residue E18 demonstrates decrease in conformational mobility (Fig. 1j) in comparison with Q18 in peptides P5, P29, P30, and P34. Interestingly, the 2D and 3D maps demonstrate pairwise similarity in distribution of free energy levels for residue E18 in peptides P27 and P31, and for residue Q18 in peptides P26 and P30, like in the case of residue Y17. This indicates that the presence of proline at *N*-termini of peptides P26 and P27 does not influence the conformational and dynamic properties of any residue located after it in the studied peptides.

In peptides P32 and P33, the glutamine residue is substituted to alanine (Q18A). These two peptides differ from each other by a second substitution (T20E and T20S, correspondingly). The alanine residue (A18) in both peptides shows significant increase in set of probable conformations as compared to glutamine or glutamic acid at the same position in the other peptides (Fig. 2j). In peptides P35 and P36, the glutamine residue at position 18 is substituted to valine (Q18V), which is characterized by a limited set of probable conformations.

At position 19 all the studied peptides contain cysteine residue (C19). On the 2D map of C19 in the initial peptide P5 (Fig. 1c) there is a single deep locus at $\varphi = -60^\circ$, $\psi = -30^\circ$ (α-helix). In tetra- and pentapeptide fragments of the heptapeptide P5, namely, YQCT (P29) and SYQCT (P44) and also in peptides P27 (PYECE), P31 (YECE), P33 (YACS), P34 (YQCS), P35 (YVCS), P41 (LYACS), and P42 (LYVCS) a shift of energy minimum to $\varphi = -150^\circ$, $\psi = 150^\circ$ takes place, and conformational mobility of C19 in all these peptides on variations of dihedral angles φ and ψ increases (Fig. 1k). As for 2D maps of the cysteine residue in peptides P26 (PYQCE), P28 (SYKCE), P30 (YQCE), P32 (YACE), P36 (YVCE), and P43 (SYQCE), only one locus at $\varphi = -135^\circ$, $\psi = -30^\circ$ is observed, and the mobility of C19 is significantly decreased (Fig. 1g). It is seen that the presence of serine residue after the cysteine residue (at *C*-terminus) leads to increase, and the presence of glutamic acid leads to decrease in conformational mobility of C19, except the situation when C19 is located between two negatively charged residues E18 and E20 (in peptide P27 and P31). These data are confirmed by 3D maps of the cysteine residue in these peptides (Figs. 2g and 2k).

At position 20 the threonine residue is kept only in peptides P29 and P44, and its conformational and dynamic properties are changed only a little. Other tetra- and pentapeptides contain substitutions T20S or T20E (Table 2). In peptides P26, P27, P28, P30, P31, P32, P36, and P43 threonine residue is substituted to glutamic acid (T20E). The glutamic acid residue E20 in peptides P26 (PYQCE), P28 (SYKCE), P30 (YQCE), P32 (YACE), P36 (YVCE), and P43 (SYQCE) demonstrates significant decrease in set of probable conformations in comparison with that for T20 in the initial peptide P5 (Figs. 1h and 2h). These data suggest that the glutamic acid residue on the whole is characterized by a limited set of probable conformations. However, its microenvironment and existence of intramolecular interactions may influence conformational mobility. For example, in the case of peptide P31 (YECE), the mutual repulsion of two negatively charged residues of glutamic acid leads to increase in mobility of the *C*-terminal residue E20.

In peptides P33 (YACS), P34 (YQCS), P35 (YVCS), P41 (LYACS), and P42 (LYVCS) the threonine residue is substituted to serine (T20S), which has the largest conformational mobility (Figs. 1l and 2l). Moreover, conformational and dynamic properties of serine residue S20 in peptides P33, P34, and P35 are very similar to each other as well as with S16 in peptide P28 (SYKCE) and with S20 in peptide P13 (IMSYICS). However, conformational mobility of S16 in the initial peptide P5 is notably smaller than those of S20 in all the above-mentioned peptides. These differences may be explained by intramolecular hydrogen bonding between S16 and Q18 in the initial peptide P5 and by the absence such a bond in the other peptides.

Analysis of autocorrelation function curves. Curves of autocorrelation function calculated according to Eq. (2) contain information about dynamic parameters of amino acid residues: characteristic time of quenching (τ) and

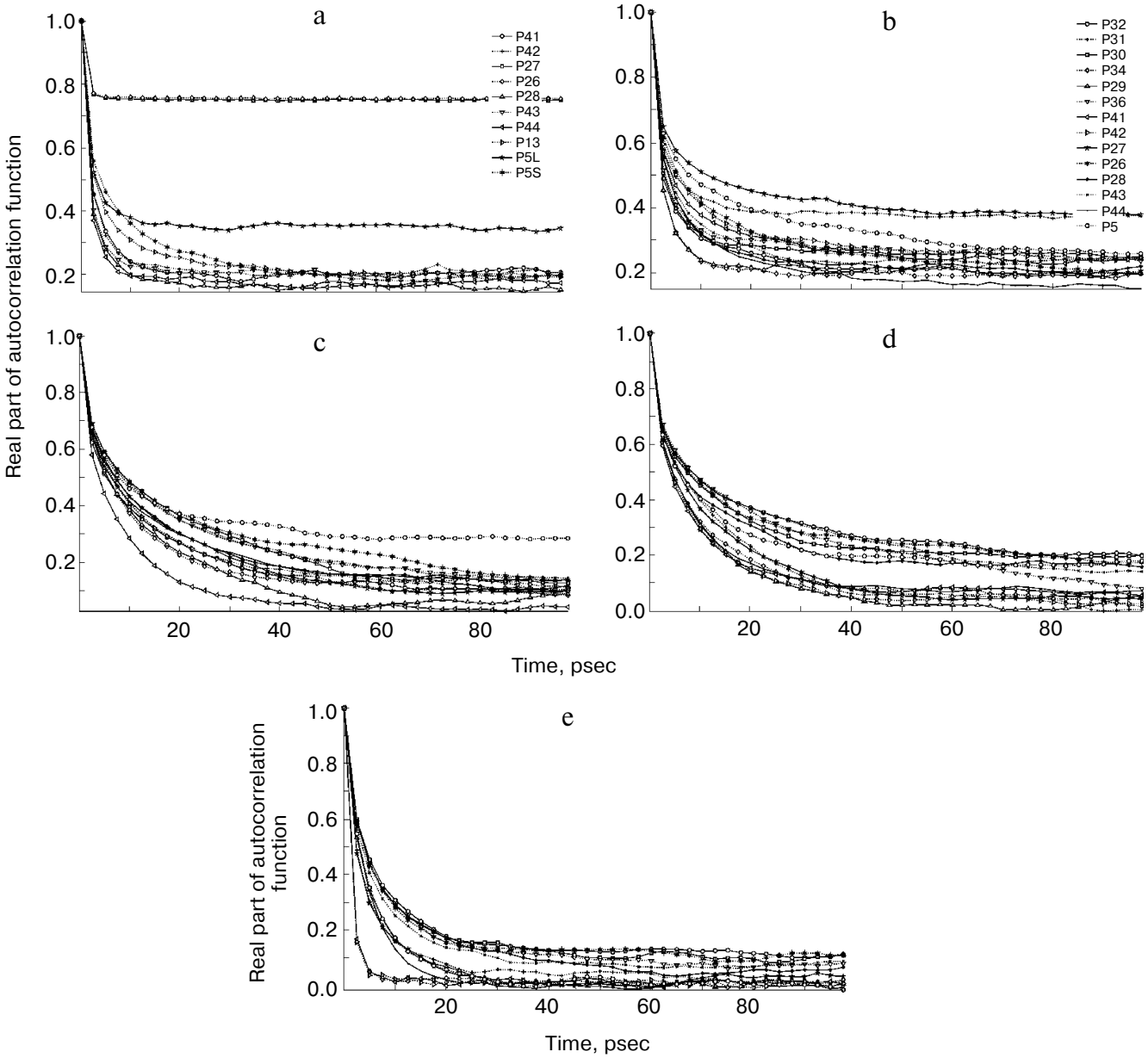

Fig. 3. Curves of autocorrelation functions for residues at positions 16 (a), 17 (b), 18 (c), 19 (d), and 20 (e) in changing of dihedral angles φ (a, b), ψ (c, d), and χ (e). Here P5L is the leucine residue in peptide P5, and P5S is the serine residue in peptide P5.

residual correlation. Characteristic time of quenching of the autocorrelation function is the time of conformational transition during change in dihedral angle, while residual correlation contains information about possible restraints in rotation around covalent bonds.

Let us first consider curves of autocorrelation functions in variations of *dihedral angle φ*. At position 16 two pentapeptides, namely, P26 (PYQCE) and P27 (PYECE), contain a proline residue, three peptides P28 (SYKCE), P43 (SYQCE), and P44 (SYQCT) contain a serine residue, and two peptides P41 (LYACS) and P42 (LYVCS) contain a leucine residue.

The proline residue in peptides P26 and P27 demonstrates high value of residual correlation (0.75) at τ = 5 psec (Fig. 3a). High residual correlation may be explained by the participation of the N and Cα atoms of proline residue in formation of the 5-member heterocyclic ring, and rotation around the covalent bond between these atoms is greatly restricted.

Residual correlation values for serine residue S16 in peptides P28, P43, and P44 as well as for leucine L16 in peptides P41 and P42 vary from 0.15 to 0.20 (in the initial peptide P5 it is equal to 0.20). The characteristic time of quenching (τ = 5 psec) is significantly lower than that in

the initial peptide P5 (20 psec), and this indicates higher mobility of S16 and L16 in comparison to S16 in the initial peptide.

At position 17 all the studied peptides kept the tyrosine residue (Y17). It is necessary to note that the tyrosine residue in peptides P27 and P31 are located before glutamic acid residue E18 and demonstrates the highest value of residual correlation, close to 0.40 (Fig. 3b). Electrostatic repulsion between two similarly charged amino acid residues (E18 and E20) seemingly decreases Y17 mobility.

The characteristic time of quenching is maximal (τ = 30 psec) for Y17, which is located after the proline residue in peptides P26 and P27. Hence, the proline residue increases the characteristic time of quenching and does not influence residual correlation value for the residue following it.

The residual correlation values for residue Y17 in the other peptides are rather close to each other and vary from 0.20 (P34 and P35) to 0.30 (P36) at low τ = 5-10 psec (Fig. 3b).

At position 18 some of the peptides keep the glutamine residue, but others contain substitutions – Q18E (in peptides P27 and P31), Q18K (in peptide P28), Q18A (in peptides P32, P33 and P41), and Q18V (in peptides P35, P36 and P42). Residual correlation values for residues at this position vary from 0.2 (for Q18 in peptides P29 and P34, and also for A18 in peptides P33 and P41) to 0.32 (for E18 in peptides P27 and P31) at τ = 10-20 psec. So, peptides P27 and P31 differ from others by having highest values of residual correlation for E18. In the initial peptide P5, the dynamic parameters are also high (0.25 at τ = 20 psec).

At position 19 there is a cysteine residue (C19) in all the peptides, for which a wide variation of residual correlation is observed – from 0.2 (in peptides P28 and P29) to 0.30-0.37 (in peptides P26, P30, P32, P36, and P43). Thus, the presence of a glutamic acid immediately after the cysteine residue increases the value of residual correlation for the latter. With the exception of peptides P27 and P31, in which cysteine residue is located between two negatively charged (and, so, mutually repulsing) residues E18 and E20, it has residual correlation value between 0.20 and 0.25.

At position 20 in most of the peptides substitutions were made – T20E (in peptides P26-P28, P30-P32, P36, and P43) and T20S (in peptides P33-P35, P41, and P42). All of the studied peptides can be divided into two groups – with residual correlation close to 0.2 (for S20 in peptides P33-P35, P41, and P42 and also for T20 in the initial peptide P5 and its fragments P29 and P44) and from 0.3 to 0.4 (for E20 in the other peptides). Characteristic time of quenching varies from 10 to 20 psec.

On variation of *dihedral angle* ψ residues at position 16 show the highest value of residual correlation for S16 in the initial peptide P5 (0.3) at τ = 30 psec. In the pentapeptides the values of residual correlation are smaller and vary from zero for S16 in peptides P28, P43, and P44 at τ = 5 psec to 0.10 for L16 in peptides P41 and P42 and also for the proline residue in peptides P26 and P27, τ = 40-60 psec. It is seen that the serine residue in the pentapeptides has higher mobility than in the initial peptide P5. This may not be explained by *N*-terminal position of serine residue in pentapeptides because dynamic parameters of the same residue at the same position in heptapeptide P13 are similar (i.e. residual correlation of 0.05) to those in pentapeptides and are different from those in the initial heptapeptide P5.

Besides, earlier [17] we showed that existence of negatively charged residue (of aspartic or glutamic acids) decreases conformational mobility of neighboring residues (provided by they are not between two negatively charged residues). In peptide P44 serine residue has higher than in the initial peptide P5 mobility because of absence of aspartic residue in the neighborhood. This might be also explained by participation of the serine residue in the initial peptide P5 in formation of a hydrogen bond with glutamine Q18, which creates restrictions for rotation. So, participation of residue S16 in intramolecular interaction leads to increase in τ value in variation of dihedral angle φ and increase in residual correlation value in variation of dihedral angle ψ.

In peptides P26 and P27 there is a proline residue at position 16 for which very high value of time of quenching (τ = 70 psec) and residual correlation close to 0 is characteristic.

As for residues at position 17, there are significant variations in residual correlation values (from 0.05 to 0.40). The greatest value is observed for the tyrosine residue in the initial peptide P5, which might be explained by squeezing of Y17 between S16 and Q18 that form a hydrogen bond. The smallest value is characteristic for peptides Y17 in peptides P32, P33, and P41, where alanine residue A18 is located immediately after a serine residue, and also in P34 and P35 that contain a *C*-terminal serine residue (substitution T20S). Values of dynamic parameters are close to the highest one also for Y17 residue in peptides P27 (PYECE) and P31 (YECE), and this indicates its low mobility caused by the presence of neighboring negatively charged glutamic acid residue.

For residues at position 18, values of residual correlation vary from 0.0-0.05 (for Q18 and A18 in peptides P29, P33, and P41) to 0.30 (for Q18 in the initial peptide P5) at τ = 30-40 psec, with the exception of peptides P33 and P41 with relatively low value of τ = 10 psec (Fig. 3c).

At position 19 (C19), all the studied peptides can be divided into two groups – those with residual correlation values close to zero (peptides P27, P29, P31, P33, P34, P35, P41, P42, and P44) and those with values close to 0.2 (the other peptides). The characteristic time of quenching varies from 20 to 40 psec (Fig. 3d). Hence, the curves of autocorrelation functions confirm the data

obtained for the 2D and 3D maps and show that the presence of glutamic acid following a cysteine residue creates conformational restraints for mobility of the latter. There is an exception for peptides P27 and P31 that contain two glutamic acid residues (E18 and E20), mutual repulsion of which increases mobility of residue C19.

Residual correlation values for residues at position 20 vary from 0 (for S20 in peptides P33-P35, P41, P42) to 0.1 (for E20 in peptides P26, P30, P32, and P43) at τ = 10-30 psec.

On variation of *dihedral angle* χ, the residual correlation values for Y17 in all peptides approach zero at sufficiently high value of τ = 30-60 psec. For residues at position 18, residual correlation values vary from 0 (for A18 in peptides P32, P33, and P41 and Q18 in peptide P34) at τ = 0 psec to 0.1 (for E18 in peptides P27 and P31) at τ = 20 psec. For C19 residual correlation values vary in the range from 0 to 0.05 at very low time of quenching (10 psec). At position 20 (Fig. 3e), the lowest values of residual correlation (near 0) and characteristic time of quenching (τ = 2-5 psec) are observed for serine residue (S20) in peptides P33, P34, P35, P41, P42, and the highest ones (0.10) are for E20 in peptides P26, P30, P32, P43 (at τ = 20-30 psec). So, the lowest mobility is observed for *C*-terminal glutamic acid with the exception of peptides P27 and P31 in which E20 demonstrates dynamic parameters similar to those for T20.

Clustering analysis. A goal of clustering analysis is to divide a given set of subjects into groups (clusters or classes) of comparatively homogeneous (similar in properties) subjects. Subjects that belong to different clusters (nonoverlapping sets of subjects) must differ greatly. Results of dispersion analysis are represented as hierarchically organized tree-like diagrams or clustering trees (Fig. 4) in which the horizontal axis shows number of a subject and vertical axis – the linkage distance. Upon constructing a clustering tree, the subjects that differ from each other by value $d_{sr} \leq 0.004$ were placed in one group. In our work, the subjects were amino acid residues at the same positions in different peptides. The 2D maps of free energy levels in coordinates (φ,ψ) were chosen for evaluation.

Residues at position 16 in all penta- and two heptapeptides (P5 and P13) are clustered as follows: the smallest distance ($1.0 \cdot 10^{-3}$) is observed for S16 in peptides P28 and P44 with joining of P13 and P43 (Fig. 4a). Interestingly, S16 in P5 is not clustered with that in other peptides. Minimal distances are also observed for L16 in peptides P41 and P42. Leucine residues L14 in peptide P5 and L16 in peptides P41 and P42 have significant difference in their 2D maps despite the fact that all are *N*-terminal ones.

The 2D maps for the tyrosine residue at position 17 (Y17) are clustered as follows (Fig. 4b): 1) P28 (SYKCE), P29 (YQCT), P33 (YACS), P34 (YQCS), P35 (YVCS), P41 (LYACS), and P42 (LYVCS); peptide P44 is outgroup for this cluster; 2) P26 (PYQCE), P30 (YQCE), P32 (YACE), and P36 (YVCE); peptide P43 is out-group for the above two clusters; 3) initial peptide P5 (LDSYQCT); 4) P27 (PYECE) and P31 (YECE). It is seen that peptides that contain serine or glutamic acid residue are clustered differently. Peptides SYKCE and YQCT are clustered together and this confirms the absence of hydrogen bonding between S16 and K18 in peptide P28.

It is also seen that the third group is represented by a single peptide (the initial peptide P5) that demonstrates unusual properties – it does not form a cluster with other peptides. A hierarchical structure is observed in which two clusters unite with formation of a larger cluster with subsequent joining of peptide P5. This circumstance confirms the similar character of changes in conformational and dynamic properties of Y17 in these peptides in comparison with the initial peptide P5 (increase in probable conformations set). Peptides P27 and P31, as expected, form a separate cluster at the greatest distance from other clusters – analysis of 2D and 3D maps shows that Y17 in these peptides demonstrates, unlike the other peptides, decrease in the set of probable conformations.

Five types of residues with different physicochemical properties are located at position 18. These are glutamine, glutamic acid, lysine, valine, and alanine. The smallest distance ($<1.0 \cdot 10^{-3}$) is observed for residue A18 in peptides P33 and P41. So, the presence of leucine residue at the *C*-terminus does not influence conformational and dynamic properties of residues that follow it. Peptides P28 (SYKCE) and P43 (SYQCE) are clustered together (Fig. 4c). A separate cluster is formed by peptides P26 (PYQCE), P30 (YQCE), P32 (YACE), and P36 (YVCE). In the third group peptides P34 (YQCS) and P35 (YVCS) show the minimum linkage distance with hierarchical joining of P29 (YQCT), P42 (LYVCS), and P44 (SYQCT). It is seen here that the 2D maps for E18 in peptides P27 and P31 demonstrate, on one side, the greatest degree of similarity with each other and, on the other side, maximal distance from the other peptides. So, clustering of peptides takes place depending not on type of residue at position 18 or similarity in their physicochemical properties, but on type of *C*-terminal residue: peptides that contain glutamic acid or serine (threonine) residues at their *C*-termini are independently clustered.

At position 19, all of the peptides keep their cysteine residue. However, there is notably different character of free energy level distribution (Fig. 4d) for this residue in different peptides. The 2D maps for C19 are clustered as follows: 1) peptides P33 (YACS), P34 (YQCS), P35 (YVCS), P42 (LYVCS), and P41 (LYACS); 2) the previous cluster is enlarged by successive addition of two groups of peptides, one of which includes P29 (YQCT) and P44 (SYQCT) and the other – P27 (PYECE) and P31 (YECE) with P28 (SYKCE); 3) P26 (PYQCE), P30 (YQCE), P32 (YACE), P36 (YVCE), and P43 (SYQCE); 4) the initial peptide P5. Thus, conformational properties

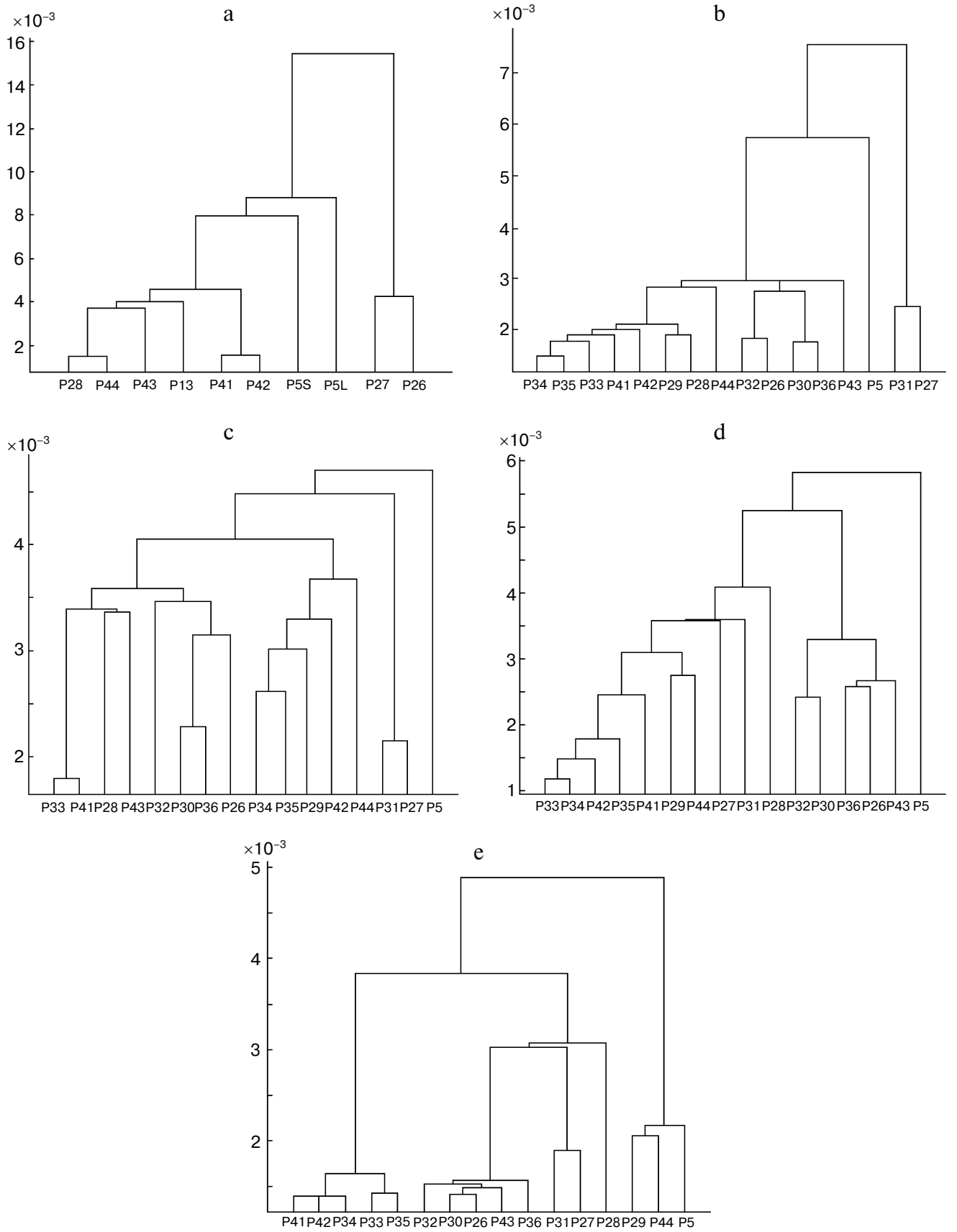

Fig. 4. Clustering trees for residues at positions 16 (a), 17 (b), 18 (c), 19 (d), and 20 (e).

of the cysteine residue depend on the type of the residue following it: peptides that contain serine and glutamic acid residues at position 20 are clustered differently. This is true for all peptides except for P27 and P31, in which the 2D maps for cysteine residue are similar to those in peptides that contain serine residue at their *C*-terminus.

At position 20, there are the following types of residues – threonine, serine, and glutamic acid. Clustering of the 2D maps for these residues takes place depending strictly on type of residue at this position (Fig. 4e). The lowest distance ($1.4 \cdot 10^{-3}$) is observed between: 1) glutamic acid residues in peptides P26 (PYQCE), P30 (YQCE), P32 (YACE), P36 (YVCE), P43 (SYQCE), and 2) serine S20 residues in peptides P33 (YACS), P34 (YQCS), P35 (YVCS), P41 (LYACS), and P42 (LYVCS). Glutamic acid residue E20 in peptides P27 (PYECE) and P31 (YECE), and P28 (SYKCE) join after. Peptides that contain threonine (T20) residue (P5, P29, P44) are at maximal distance from the others and form the fourth cluster. Similar results were obtained for clustering analysis performed for autocorrelation functions.

Thus, all of the studied peptides in similarity of their 2D maps for amino acid residues at positions 17-20 can be divided into four groups. The first group is composed of peptides that contain serine residue at their *C*-termini: P33 (YACS), P34 (YQCS), P35 (YVCS), P41 (LYACS), and P42 (LYVCS). The second group includes peptides that contain a glutamic acid residue at the *C*-terminus: this is P26 (PYQCE), P30 (YQCE), P32 (YACE), P36 (YVCE), and P43 (SYQCE). Peptides P27 (PYECE) and P31 (YECE) are clustered separately from the other peptides. The fourth group is represented by a single peptide – heptapeptide P5 that is not clustered with other peptides except for residue T20. Peptides P28 (SYKCE), P29 (YQCT), and P44 (SYQCT) can be clustered either with the first (by residues at positions 17, 18, and 19), or with the second group (peptide P28 by residues C19 and E20), or with the heptapeptide P5 (P29 by residue at position 20).

DISCUSSION

In general, all tetra- and pentapeptides studied in this work can be divided into three groups according to alterations in conformational mobility of their amino acid residues. The first group includes peptides in which conformational mobility of residues higher than that of corresponding residues at the same positions in the initial peptide P5. The majority of the peptides of the first group contain a serine residue at their *C*-termini (Table 2). These are tetrapeptides P29 (YQCT), P33 (YACS), P34 (YQCS), and P35 (YVCS) and also pentapeptides P28 (SYKCE), P41 (LYACS), P42 (LYVCS), and P44 (SYQCT). Notably, all of the studied peptides differ in physicochemical properties of the residues at position 18 that have similar picture of free energy level distribution. For example, the alanine residue with small hydrophobic side chain is clustered with lysine residue K18 with large hydrophilic positively charged side chain. Moreover, Q18 with hydrophilic side chain is clustered with V18 that possesses a hydrophobic side chain.

The second group includes peptides that contain a glutamic acid residue at the *C*-terminus. These are P26 (PYQCE), P30 (YQCE), P32 (YACE), P36 (YVCE), and P43 (SYQCE). In these peptides, the conformational mobility of some residues is higher (at positions 17 and 18), and others are lower (at positions 19 and 20) than of residues at the same positions in the initial peptide P5. The glutamic acid residue has a large negatively charged side chain and restricted mobility. It decreases (unlike the serine residue that increases) mobility of the cysteine residue located before it. These data are in agreement with results obtained by other authors [30] and can be explained by the presence of a large side chain as well as its high reactivity in glutamic acid. It is possible that a glutamic acid residue can interact with the side chain of a cysteine residue that in some proteins and peptides leads to formation of β-cysteinyl-γ-glutamyl ether [31].

Peptides P27 (PYECE) and P31 (YECE) form the third group and are distinguished by peculiarities of conformational and dynamic behavior. In these peptides conformational mobility of residues in general is lower than of those (except for C19) at corresponding positions in peptide P5, despite the presence of glutamic acid residue at *C*-terminus. Distinguished conformational-dynamic behavior of these peptides is dictated by the presence of two negatively charged glutamic acid residues that themselves have restricted mobility and decrease mobility of neighboring residues. Increase in a set of probable conformations for C19 can be explained by electrostatic repulsing between two negatively charged residues E18 and E20. This decreases steric constraints in rotation around covalent bonds in residues located between them.

Remarkably, peptides PYQCE and PYECE as well as YQCE and YECE are pairwise analogs that differ by a single amino acid substitution, namely, glutamic acid amide to glutamic acid (Q18E). These amino acids are similar in their physicochemical properties (values of hydrophobicity and size of side chains) but differ due to the presence or absence of a charge. Despite the conservative character of the substitution, there is a significant difference in biological activity in each pair of peptides [7-10]. We suggest that exactly electrostatic repulsion between two negatively charged glutamic acid residues in peptides PYECE (P27) and YECE (P31) provides peculiarities of conformational and dynamic behavior of the peptides, and consequently their biological activity.

Previously we showed that synthetic peptides LDSYQCT (AFP_{14-20} or peptide P5) and PYECE significantly decrease expression of early (CD25 and CD71) as well as late (HLA-DR and CD95) activation antigens on

lymphocytes of patients with atopic bronchial asthma and rheumatoid arthritis. It was also shown that peptides YECE and YVCE decrease expression of the activation antigen CD95 on the surface of lymphocytes in patients with ABA and infectious and allergic myocarditis [8, 9]. Peptides PYQCE, SYKCE, YQCE, YACE, and YVCS do not influence the expression of activation antigens of lymphocytes.

So, peptides that possess the greatest biological activity are characterized by relative rigidity of their backbone. These are peptides PYECE, YECE, and YVCE as well as the initial peptide LDSYQCT in which amino acid residues have restricted conformational mobility due to the presence of negatively charged aspartic acid along with hydrogen bonding between serine and glutamine residues [17]. Peptides in which amino acid residues have high conformational mobility are characterized by low biological activity. So, there is reverse correlation between biological activity and flexibility of the peptide backbone.

Our data confirm the suggestion that was made earlier that conformational and dynamic properties of amino acids depend on their participation in intramolecular interactions arising between functionally active groups of side chains or with CO- and NH-groups of the peptide backbone [29]. For example, conformational mobility of residues S16 and K18 is higher and of residue C19 is lower in peptide P28 than that of residues at corresponding positions in the initial peptide P5. This can be explained by formation of hydrogen bond between S16 and Q18 in peptide P5 that leads to restriction in their mobility and also by possible electrostatic attraction between K18 and E20 in peptide P28. Formation of a hydrogen bond between side chains of S16 and Q18 can serve as an explanation for the fact that conformational mobility of S16 in heptapeptide P5 is lower than that of S20 in all the studied tetrapeptides.

Intramolecular interactions play an important role in maintaining structure and folding of a protein. However, our MD calculations showed that not all residues able to take part in intramolecular interactions realize this ability in molecules of peptides or proteins. Moreover, only some of these interactions are crucial for maintaining spatial structure of a protein. This may explain the fact that even at 88% identity of amino acid sequences of two proteins they may have absolutely different structure and functions [32]. Fine balance of intramolecular interactions can be impaired if residues that participate in these interactions are substituted with no compensatory substitution [33]. Thus, our results show that MD methods can be used to reveal amino acid residues participating in intramolecular interactions. This, in turn, is necessary to elucidate the detailed mechanisms underlying protein folding. Moreover, MD methods can be used to predict relative flexibility of various parts of a polypeptide backbone and biological activity of a peptide.

The proline residue in peptides P26 and P27 demonstrates in variation of dihedral angle φ high value of residual correlation (0.75) and $\tau = 5$ psec, where as in variation of dihedral angle ψ, on the contrary, very high value of time of quenching ($\tau = 70$ psec) and residual correlation close to 0. This is provided by formation of the heterocyclic ring with participation of its side chain and N and Cα atoms. As a result, rotation around covalent bond N–Cα (dihedral angle φ) is restrained; however, rotation around the Cα–C′ bond (dihedral angle ψ) is the same as for alanine. Comparative study of two pairs of peptides (P26 and P30, P27, and P31) that are differed by presence or absence of a proline residue at the N-terminus showed that the proline residue does not influence residual correlation value, but it increases characteristic time of quenching for the residue following it. So, a proline residue narrows the area of allowed conformations for residues located before it [27] and does not change the set of probable conformations with increase in time of conformational transition for residues following it.

The polypeptide chain of a protein can be considered as a set of evolutionarily selected short peptide sequences. It is quite possible that short peptide motifs (not only a definite amino acid residue) contain information about spatial structure and functions of a protein. This is confirmed by data obtained on revealing so-called “protein grammar” [34-38]. According to these data, the hierarchically organized structural elements can be distinguished in a protein molecule. Such elements may be of different size and may be responsible for a definite protein function.

The molecule of a protein may be composed of different types of structural elements – from motifs to domains. This was shown on the basis of calculations of positional entropy as a function of difference in distances between amino acid residues in a polypeptide chain [36, 37]. This approach revealed the low and constant values of positional entropy at small distances between amino acids (≤5 amino acid residues). The suggestion was made that short pentapeptide segments might serve as informational and structural units of the lowest level in proteins, while domains – of the highest level.

The MD calculations showed that pentapeptide fragments can have the most preferable conformation and, hence, may serve as rigid elements responsible for maintaining the protein structure. However, analysis of primary structures of proteins available in databases showed that not all theoretically possible pentapeptides (unlike tripeptides) might be found in real proteins [34, 35]. Besides, the existence of functionally significant tripeptide fragments (such as RGD) suggests that tripeptides (not pentapeptides) may serve as structural and functional elements of minimal length.

We suggest that the shortest elements of the first level may be represented by tripeptide segments such as RGD [39] and the longest ones – by EGF-like modules of 30-40 residues [40]. Besides, short peptide segments of the same sequence (or similar linear motifs) can be found in different proteins and may be responsible for similar

functions of those proteins (for example, PxxP or LLXXL motifs) [41, 42]. One protein molecule can contain few or many repeats of such segments (motifs).

Earlier, we propose that conformational and dynamic properties of proteins play an important role in molecular evolution [17]. Analysis of protein–protein interaction interfaces performed by Nekrasov et al. suggested that for effective interactions, one polypeptide chain must have strictly determined topology and another one must be able to undergo adaptive conformational changes [37]. This property of interacting proteins is a result of molecular evolution. It was also shown that in protein molecules the functionally important sites (for example, antigenic epitopes) contain short peptide segments with rigid conformation [38]. Consequently, a polypeptide chain can be composed of a short peptide segment of rigid conformation that is provided by intramolecular interactions. So, molecular evolution should keep intramolecular interactions that maintain protein structure, especially in the environment of functionally important sites.

REFERENCES

1. Neduva, V., and Russel, R. B. (2005) *FEBS Lett.*, **579**, 3342-3345.
2. Zaretsky, J. Z., and Wreschner, D. H. (2008) *Translat. Oncogenomics*, **3**, 99-136.
3. Terentiev, A. A., Mokhosoev, I. M., and Moldogazieva, N. T. (2010) in *Human Placenta: Structure and Development, Circulation and Function*, Chap. 4 (Berven, E., and Freberg, A., eds.) Nova Science Publishers, Inc., pp. 125-143.
4. Streydio, C., Lacka, K., Swellens, S., and Vassart, G. (1988) *Biochem. Biophys. Res. Commun.*, **154**, 130-137.
5. Terentiev, A. A. (1997) *Vestnik RGMU*, **1**, 76-79.
6. Terentiev, A. A., and Moldogazieva, N. T. (2006) *Biochemistry (Moscow)*, **71**, 120-132.
7. Kazimirsky, A. N., Tagirova, A. N., Salmasi, J. M., Moldogazieva, N. T., Gurina, A. E., and Terentiev, A. A. (2010) *Astrakhan. Med. Zh.*, **1**, 106-110.
8. Kazimirsky, A. N., Salmasi, J. M., Kuptsova, N. V., Tagirova, A. N., Moldogazieva, N. T., and Terentiev, A. A. (2010) *Int. J. Exper. Educat.*, **7**, 28-29.
9. Tagirova, A. K., Terentiev, A. A., Alexandrova, I. A., Moldogazieva, N. T., Salmasi, J. M., Porjadin, G. V., and Kazimirsky, A. N. (2007) *Tumor Biol.*, **28** (Suppl. 1), S106.
10. Terentiev, A. A., Tagirova, A. K., Tutcheva, O., Moldogazieva, N. T., Salmasi, J. M., Kazimirsky, A. N., Riabinina, Z., and Makarov, T. (2010) *Tumor Biol.*, **31** (Suppl. 1), S78.
11. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.*, **157**, 105-132.
12. Gohlke, H., Kiel, C., and Case, D. A. (2003) *J. Mol. Biol.*, **330**, 891-913.
13. Case, D. A., Cheatham, T. E., III, Darden, T., Gohlke, H., Lou, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) *J. Comput. Chem.*, **26**, 1668-1688.
14. Shaitan, K. V., Vasiliev, A. K., Saraikin, S. S., and Michailiuk, M. G. (1999) *Biofizika (Moscow)*, **44**, 668-675.
15. Shaitan, K. V., Tourleigh, Ye. V., Golik, D. N., and Kirpichnikov, M. P. (2006) *J. Drug Deliv. Sci. Technol.*, **16**, 253-258.
16. Froimowitz, M. (1993) *Biotechniques*, **14**, 1010-1013.
17. Moldogazieva, N. T., Terentiev, A. A., Kazimirsky, A. N., Antonov, M. Yu., and Shaitan, K. V. (2007) *Biochemistry (Moscow)*, **5**, 529-539.
12. Moldogazieva, N. T., Shaitan, K. V., Tereshkina, K. B., Antonov, M. Yu., and Terentiev, A. A. (2007) *Biofizika (Moscow)*, **4**, 611-624.
19. 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.
20. 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.
21. Dashevsky, V. G. (1974) *Conformation of Organic Molecules*, Khimiya, Moscow.
22. Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S., Windus, T. L., Dupuis, M., and Montgomery, J. A., Jr. (1993) *J. Comput. Chem.*, **14**, 1347-1352.
23. Vesterkhoff, H., and van Dam, K. (1992) *Thermodynamics and Regulation of Transformations of Free Energy in Biosystems* [Russian translation], Mir, Moscow.
24. Lemak, A. S., and Balabaev, N. K. (1995) *Mol. Simulation*, **15**, 223-231.
25. Lemak, A. S., and Balabaev, N. K. (1996) *J. Comput. Chem.*, **17**, 1685-1695.
26. Golo, V. L., and Shaitan, K. V. (2002) *Biofizika (Moscow)*, **47**, 611-617.
27. Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963) *J. Mol. Biol.*, **7**, 75-102.
28. Govariker, V. R., Visvanathan, N. V., and Shridhar, J. (1990) *Polymers* [Russia translation], Nauka, Moscow.
29. Moldogazieva, N. T., Shaitan, K. V., Antonov, M. Yu., Vainogradova, I. K., and Terentiev, A. A. (2011) *Biochemistry (Moscow)*, **76**, 1321-1336.
30. Tereshkina, K. B., Shaitan, K. V., Levtsova, O. V., and Golik, D. N. (2005) *Biofizika (Moscow)*, **50**, 974-985.
31. Pangbum, M. K. (1992) *FEBS Lett.*, **308**, 280-282.
32. Alexander, P. A., He, Y., Chen, Y., Orban, J., and Bryan, P. A. (2007) *Proc. Natl. Acad. Sci. USA*, **104**, 11963-11968.
33. Chelia, V., and Blundel, T. (2005) *Biochemistry (Moscow)*, **70**, 835-840.
34. Otaki, J. M., Ienaka, S., Gotoh, T., and Yamamoto, H. (2005) *Protein Sci.*, **14**, 617-625.
35. Otaki, J. M., Tsutsumi, M., Gotoh, T., and Yamamoto, H. (2010) *J. Chem. Inf. Model.*, **50**, 690-700.
36. Nekrasov, A. N. (2004) *J. Biomol. Struct. Dyn.*, **21**, 615-623.
37. Nekrasov, A. N., and Zinchenko, A. A. (2010) *J. Biomol. Struct. Dyn.*, **28**, 85-96.
38. Svirshchevskaya, E., Alekseeva, L., Marchenko, A., Benevolenskii, N., Berzhec, D. M., and Nekrasov, A. N. (2006) *J. Bioinform. Comput. Biol.*, **4**, 389-402.
39. Ruoslahti, E., and Pierschbacher, M. D. (1987) *Science*, **238**, 491-497.
40. Terentiev, A. A., and Moldogazieva, N. T. (2006) *Uspekhi Biol. Khim.*, **46**, 99-148.
41. Sadana, P., and Park, E. A. (2007) *Biochem. J.*, **403**, 511-518.
42. Solomaha, E., Szeto, F. L., Youself, M. A., and Palfrey, H. C. (2005) *J. Biol. Chem.*, **280**, 23147-23156.