ISSN 0006-2979, Biochemistry (Moscow), 2007, Vol. 72, No. 4, pp. 392-397. © Pleiades Publishing, Ltd., 2007.
Original Russian Text © P. V. Kostetsky, 2007, published in Biokhimiya, 2007, Vol. 72, No. 4, pp. 486-492.

Location and Volume of the Active Site of Chymotrypsin

P. V. Kostetsky

Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, 117997 Moscow, Russia; fax: (495) 335-7103; E-mail: artem1@ibch.ru

Received February 17, 2006
Revision received September 1, 2006

Abstract—The active site of chymotrypsin molecule (approximated by a sphere with radius of 20 Å) was taken as the largest cavity on the enzyme surface. The volume inside the approximating sphere is sufficient for placement of 95% of non-hydrogen atoms of the enzyme. The active site cavity is localized in a spherical sector with solid angle of 80° whose axis passes through the CB-atom of the Ser195 residue. The volume of the active site cavity is about 2700 Å^3 (8% of the volume of the approximating sphere) as computed by the Monte-Carlo method from known X-ray data. The size and shape of the active site cavity is sufficient for entrance of significantly large fragments (more than 60 non-hydrogen atoms) of the substrate molecule. At the active site cavity bottom, there is a narrow compartment adjacent to an oxy-anion hollow and accessible to water but not to substrate molecules. The water molecules inside this narrow compartment can take part in heat exchange with the external medium during different steps of the enzymatic process.

DOI: 10.1134/S0006297907040050

Key words: approximating sphere, oxy-anion hollow, heat exchange

The active sites in most enzymes are localized inside large cavities on the surfaces of the protein molecules. To date, 67 single-chain enzyme molecules have been studied [1, 2] for which spatial structures are known at high resolution. The cavities and inner hollows in the enzyme molecules can be filled with spheres of 1 to 4 Å in radius between each pair of atoms. The spheres striking neighboring atoms were excluded. The surfaces thus formed bounded the cavities and hollows, the volume of which has been estimated interactively using standard software for molecular graphics.

For 83% of these enzymes, the active site is localized in the largest cavities, whose volumes often exceed 3000 Å^3, and the average volume of these cavities comprise about 17% of total enzyme molecule. The exact volumes of active sites are only reported for 10 enzymes [1], among which, however, the chymotrypsin (CT) molecule is absent, although it is an important classical object of enzymology [3].

We previously studied the space surrounding the side chain of the His57 residue in active site of the CT molecule [4]. We have shown that within a radius of 7 Å from imidazole ring, a hollow more than 300 Å^3 in volume is present, in which atoms comprising the side chain of the His57 residue can move. The atoms of main chain of the enzyme molecule, as well as CB-atoms of amino acid residues firmly attached to it, form the molecule backbone, which is an inflexible element [5] of enzyme machine. On the other hand, side chains of His57 residue and some other residues in the active site of the enzyme are movable elements of this machine.

In the present work, a method is proposed for determination of localization and volume of the largest cavity on the surface of the CT molecule approximated with a sphere 20 Å in radius using the Monte-Carlo method [6] based on known X-ray structure data. This cavity is localized within a spherical sector with solid angle of 80°, the axis of which intersects the CB-atom of the Ser195 residue, which is the most important residue of the enzyme active site. The volume of the active site cavity is 2700 Å^3, and the atoms of the catalytic triad His57-Asp102-Ser195 are on the bottom of the cavity at the distance of 7-12 Å from the center of gravity of the CT molecule.

We have also shown that a substantial part of a substrate molecule of arbitrary form and size not exceeding 1500 Å^3 in volume can be tucked into the cavity of the CT active site. This figure is formed by the intersection of approximating sphere and a smaller sphere of 9.5 Å in radius and with its center localized on the surface of the larger sphere near the axis of the spherical sector containing the active site cavity.

Abbreviations: CT) chymotrypsin.

A narrow compartment accessible for water but not substrate molecules is on the bottom of the active site of the CT molecule near the His57 and Ser195 residues. This compartment is a part of the oxyanionic hollow of the enzyme molecule and may serve as a heat exchanger for energy exchange between the active site zone and the environment.

METHODS OF INVESTIGATION

The survey of the sphere approximating the CT molecule was performed using exploratory spherical sectors with the apex at the sphere center and with the apical angle of 50°, enabling the finding of hollows up to 900-1100 Å^3 in volume. The directions of axes (vectors) of exploratory spherical sectors were given by means of stochastic values of two angles in spherical coordinates at the fixed (20 Å) radius value. This facilitated the comparison of either spherical sector location by operating already rectangular coordinates of the corresponding vector endings disposed at the equal distances from the center of the sphere.

Calculations associated with the evaluation of cavity volumes were carried out using the Monte-Carlo method [6] by generating 30,000 stochastic points in the sphere and analyzing those falling in the exploratory spherical sector. The portion of points falling in the sector and distant no less than by the sum of corresponding atomic radii from non-hydrogen atoms of neighboring amino acid residues was counted in every numerical experiment. This allows water molecules, which were approximated by balls 3 Å in diameter, to be in these points of space. The radii of carbonic and sulfuric atoms in CT molecule were taken as 1.9 Å, and radii of oxygen and nitrogen were taken as 1.4 Å [7].

Knowing the volume of the whole sphere, total number of stochastic points (30,000), and the number of points filling the requirements, it is easy to estimate approximately the volume of hollow falling inside the exploratory spherical sector. The most successful exploratory spherical sectors contain hollows of 1000-1200 Å^3 and more in volume, and all of them are closely situated within the approximating sphere, which is well seen from a comparison of Cartesian coordinates of sector axis ends. So, the vector whose end was given by the mean of 20 exploratory spherical sectors with hollow volumes exceeding 1000 Å^3 was taken as the best direction for the spherical sector axis. This axis goes near the CB-atom of the Ser195 residue comprising the CT active site.

A range from 20 to 180° was considered to study the dependence of the hollow volume inside the spherical sector with the best orientation of its axis (hereafter referred to as spherical sector) on the value of solid angle. This dependence is well approximated by two linear areas with sharply different slopes: 28 Å^3/deg within the range of 20-80°, and only 6 Å^3/deg within the range of 80-180°. Both straight lines intersect at a point at the angle of 80°, this value being fixed for further calculation of the active site volume. The hollow volume reached 1670 Å^3 when the solid angle of the spherical sector was 80°; thereafter, the growth of the volume became insignificant.

The obtained value 1670 Å^3 characterized only the volume accessible for the centers of exploratory balls 3 Å in diameter. Naturally, this volume should be distinctly smaller than the complete "Archimedean" volume of the hollow, because the exploratory balls completely, but not only with their centers, fall into it. It was necessary to evaluate the volume of additional layer, the thickness of which was equal to 1.5 Å radii of exploratory balls. With this aim, a numerical experiment was carried out using balls of radii decreased to 1 Å.

The new value of 2020 Å^3 allowed evaluation of an additional 0.5-Å layer volume. This enabled extrapolation of the found augmentation in 350 Å^3 onto the whole additional layer of 1.5 Å, that is 1050 Å^3. Thus, the total volume of the active site cavity of 2700 Å^3 was calculated approximately. Note, due to the absence of hydrogen atoms the exploratory balls penetrate outside the borders of the real hollow, thus enlarging its volume, so further decrease of dimensions of the exploratory balls was considered inappropriate [7].

In search of atoms of amino acid residues localized near the narrow compartment on the bottom of the active site cavity accessible for water, but not substrate, molecules, we determined first the coordinates of the arbitrary center of the compartment. For this, we selected from all stochastic points inside all cavities of the approximating sphere only those that were no more than 7 Å distant from OG-atom of the Ser195 residue and did not belong to an additional minor sphere. The minor sphere, 9.5 Å in radius, with the center on the crossing of the spherical sector axis with the surface of the approximating sphere, does not contain non-hydrogen atoms and bounds the bottom of the active site of the CT molecule. Coordinates obtained by averaging were taken as the arbitrary center of the compartment. The atoms of twelve amino acid residues localized at the border of the heat-exchanging compartment around the found point at a range of 6 Å were identified.

When evaluating the volume of the narrow compartment, we considered only stochastic points around its arbitrary center at the distance less than 6 ± 1 Å, wherein the distance from the surface of the approximating sphere was presumed more than 4 ± 1 Å (within the given limits, the contribution of extraneous points is insignificant). The value of total "Archimedean" volume of the compartment, calculated similarly to the volume of active site cavity, comprised 260-340 Å^3.

Coordinates of non-hydrogen atoms of the CT molecule and oxygen atoms of inner water molecules were obtained from the PDB database of protein structures ([8], code 4cha). For designations of atoms of basic and

side chains of CT molecule the PDB nomenclature was used. When searching for necessary projection of inner parts of the CT molecule, the origin of coordinates was matched with the center of gravity of the object. Further transformation of coordinates of atoms and stochastic points was made via usual operations by rotation of coordinate axes following the rule on counter-clockwise reading of turning angle. The OZ axis was matched with the axis of the spherical sector containing the cavity of the active site of the molecule, and the YZ-projection of the sphere central disk approximating the CT molecule was displayed on the monitor.

RESULTS AND DISCUSSION

Because the CT crystal structure at the resolution of 1.68 Å is available, it is possible to determine the site of localization of cavities on the enzyme surface. Sizes of these cavities must depend on the sphere radius approximating the protein globule in the simplest way. The center of such sphere was matched with the center of gravity of the protein molecule, and the radius of the sphere was chosen from simple physical considerations so that not less than 95% of non-hydrogen atoms of the enzyme molecule could fit into the sphere.

In fact, 1476 (84%) of all non-hydrogen atoms of the enzyme go inward, approximating a sphere 20 Å in radius and 33,500 Å^3 in volume. The density of non-hydrogen atoms is 50 atoms/nm^3 in depth of the sphere at a distance up to 14 Å from the center of gravity of CT molecule, which is in good agreement with the atom density 1-1.2 g/cm^3 known for other globular proteins [9]. It is easy to guess that the volume of an approximating sphere 20 Å in radius is enough to fit 95% from 1757 heavy atoms of the CT molecule with the density of 50 atoms/nm^3.

The deviation of CT molecular shape from spherical cannot be more than the portion of non-hydrogen atoms outside of the sphere, i.e. 16%. Moreover, when the active site cavity is considered as a part of total molecular volume of the enzyme, the deviation of the globule shape from spherical is limited to 10% (see below). Note that the volume of the sphere differs insignificantly from the volume of the CT molecule – 31,300 Å^3 given in [10].

It was possible to reveal some sites accessible for water molecules inside the approximating sphere. Figure 1 shows that about one-half of the inner surface of the sphere is covered with water molecules, whereas in the depth of the sphere at a distance of 3 Å and more from its surface, the space accessible for water molecules rapidly decreases (Fig. 1, a and b). The most vast and deep cavi-

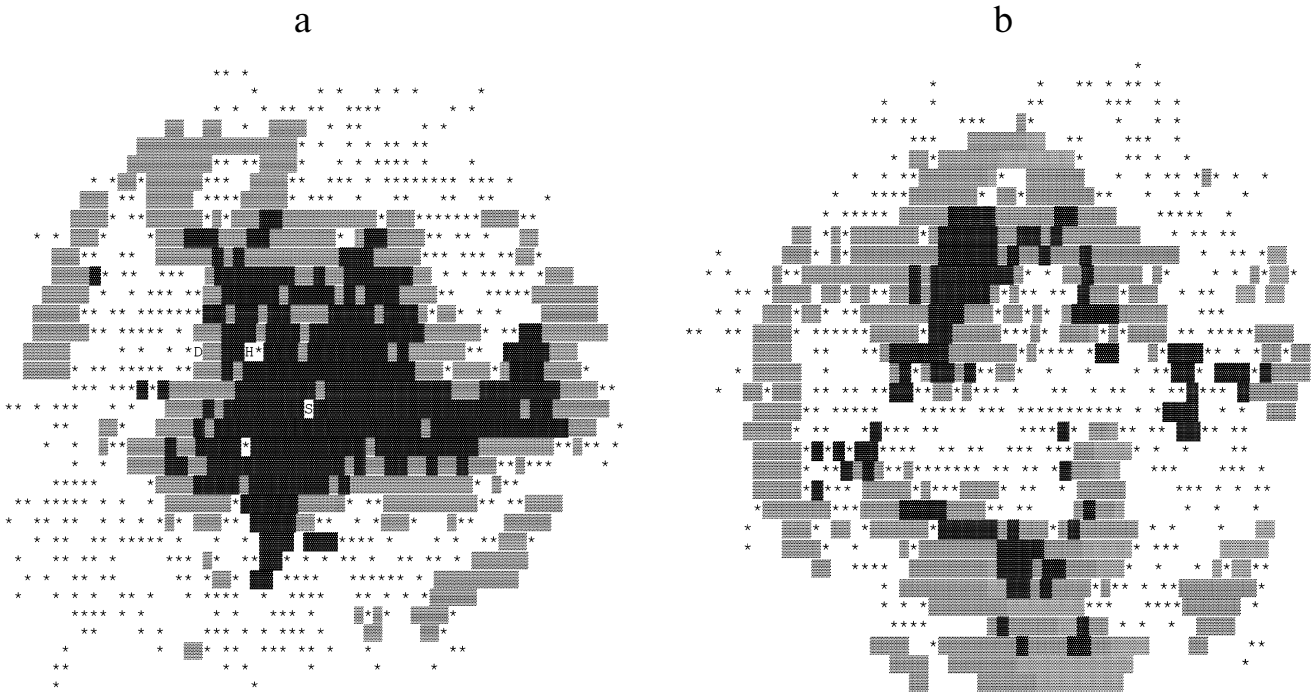

Fig. 1. The sphere approximating the CT molecule. The radius of the sphere is 20 Å. The centers of non-hydrogen atoms correspond to asterisks (*). The compartments accessible to water molecules correspond to large gray symbols, and the compartments accessible for water molecules at the depth of 3 Å and more from the surface of sphere correspond to the dark symbols. a) Projection along OZ axis of upper part of the approximating sphere; the atoms comprising catalytic triad His57 (CG-atom), Asp102 (CG-atom), and Ser195 (OG-atom) localized at the bottom of the enzyme active site are given as letters H, D, and S; b) projection along OZ axis of lower hemisphere, within which only small islands accessible for water molecules are at the depth of 3 Å and more.

ty is localized near atoms of His57, Asp102, and Ser195 residues forming the catalytic triad of CT. In contrast, in other parts of the sphere the space accessible for water molecules at the depth more than 3 Å consists of small islands and is three times smaller than the main active site cavity (Fig. 1b).

To reveal localization of cavities on the surface of the CT molecule, we scanned the approximating sphere 20 Å in radius with test ball sectors, the apex of which was matched with the center of the sphere, the solid angle at the apex taken a 50°, and the axis oriented randomly. Most fitted vectors specifying the direction of axes of the test spherical sectors proved to be directed (come) to the same cavity 1000-1200 Å^3 in volume. The coordinates of the best-fit vector were equal to the average from coordinates of 20 most fitted vectors. The axis of the best-fit vector passed through (crossed) the CB-atom of Ser195 residue.

The volume of the cavity increased rapidly with increasing solid angle and reached 1600 Å^3 at 80°, and thereafter the growth of volume became insignificant. The volume of the active site measured in this way was bounded by the surface formed with random points — the centers of balls 1.5 Å in radii. Hence, the volume of these balls partially extends beyond the bounding surface, and the virtual total volume of active site includes an additional layer 1.5 Å in width. The value of the total volume of the active site of the CT molecule calculated from a number of measurements was 2700 A^3.

This value comprises only 8% of the approximating sphere volume, which is substantially smaller than the average for 67 enzymes (17%) studied in work [1]. Nevertheless, for carboxypeptidase A (29,400 Å^3), an enzyme close in these dimensions to CT, the active site volume comprises only 1330 Å^3, i.e. less than 5% of the total enzyme.

As seen from Fig. 1, there are several hollows on the CT molecule surface directed inward into the approximating sphere. More than 40% of the volume of all these hollows form one of the deepest cavities, where the active site of the enzyme is localized (Fig. 1a). Other rather vast places accessible for water molecules are localized on the surface of the protein globule and do not form any significant cavity (Fig. 1b). The volume of hollows in the test spherical sectors falling in these places of the approximating sphere is not more than 300-400 Å^3.

The resulting spherical sector covers most of the atoms of the His57 and Ser195 residues of the catalytic triad. It is interesting that the active site cavity has no septa. This is based on the fact that a substantial part of an additional sphere of 9.5 Å in radius with its center placed on the surface of the approximating sphere 20 Å in radius passes freely inward (inside) this sphere and forms a total site 1240 Å^3 in volume.

Note that the calculated volume of the active site cavity of the CT molecule cannot be taken as an absolute magnitude, and this volume depends on the radius of the approximating sphere. We have mentioned above that the chosen radius of the sphere (20 Å) provides its volume enough for placement of 95% of the non-hydrogen atoms of the CT molecule at atomic density of 50 atoms/nm^3.

The structure of the active site of the CT molecule can be visually examined using projections of central disks of small width when the OZ axis is aligned with the axis of the best-fit spherical sector. One of these projections obtained by rotation of the image around the OZ axis is presented in Fig. 2. One can see a compartment accessible for water, but not for larger molecules of substrate or its fragments, which is oriented from the main part of the active site cavity inward toward the CT molecule. The length and volume of this compartment are 6-7 Å and 260-340 Å^3, respectively. The atoms of His57 and Ser195 residues are localized in close proximity to the compartment, whereas atoms of Asp102 residue, which also comprises the catalytic triad, are localized apart from the compartment and even do not fall into the 5 Å-thick central disk (Fig. 2). Forty non-hydrogen atoms of 12

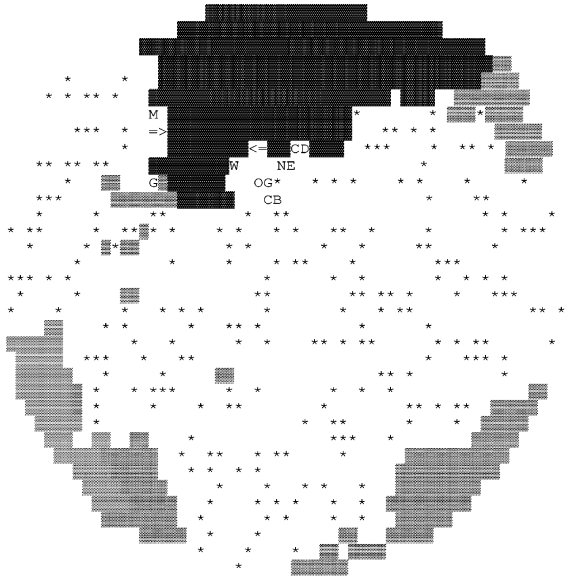

Fig. 2. Central disk of a sphere approximating the chymotrypsin molecule. The sphere radius is 20 Å and the disk width is 5 Å. The centers of non-hydrogen atoms of the enzyme molecule correspond to asterisks (*), the inner sphere compartments accessible to water molecules correspond to the large gray symbols, and the compartments accessible for water molecules in the active site cavity correspond to the large dark symbols. The arrows => and <= indicate entry into the heat exchange compartment of the active site, inaccessible for substrate molecules. Atoms of Ser195 residue are given as letters OG and CB, and atoms of His57 (CD2- and NE2-atoms) presenting inside the central disk and localized on the bottom of the enzyme active site are given as letters CD and NE. The atoms of Gly216 (CA-atom), Met192 (CE-atom), and Trp215 (CA-atom) residues localized at the border of the heat-exchanging compartment but outside the central disk are given as letters G, M, and W.

amino acid residues (His57, Ser190, Cys191, Met192, Gly193, Asp194, Ser195, Val213, Ser214, Trp215, Gly216, and Ser217) are localized on the bottom of the CT active site near the borders of the narrow compartment at distance less than 6 Å from its center. Note that the central disk of the approximating sphere (Fig. 2) contains two surface sites accessible for water molecules.

The active site of the CT molecule is known to have an oxyanionic hollow localized between the OG- and N-atoms of the Ser195 residue and the N-atom of the Gly193 residue [3, 11]. As it turned out, these three atoms localized at short range from each other are right up to the narrow compartment. Thus, the space occupied by the narrow compartment can be considered a part of the oxyanionic hollow, while the hydrophobic pocket, which serves for specific binding of aromatic rings of substrate molecules [3], is localized apart from the narrow compartment. For instance, atoms of residue Ile16 localized at the boundary of hydrophobic pocket are spaced at 7.5 Å and more from the narrow compartment.

As seen from the present results, a substantial part of a substrate molecule of arbitrary shape can be free to tuck into the space of the active site of the CT molecule within the spatial figure of 1240 Å^3 in volume formed by the intersection of approximating and additional spheres. This volume can contain the centers of 62 non-hydrogen atoms, when the atom density in substrate molecule is taken as 50 atoms/nm^3, the value corresponding to an average density of atoms in the depth of CT molecule at the distance of 14 Å from its center of gravity.

Note also that, according to our calculations, among 80 water molecules found in CT crystals [8], almost 20 molecules are in the active site of the enzyme. Eight of them are on the bottom of the active site at a distance more than 5 Å from the surface of the approximating sphere. They are disposed within a narrow compartment 260-340 Å^3 in volume.

One of eight water molecules lies at the minimal distance of 3 Å from the OG-atom of the Ser195 residue and seems to be directly involved in the reaction of enzymatic hydrolysis. One can suppose that the other seven water molecules localized on the bottom of the active site cavity serve as transmitters of heat energy released or absorbed at different stages of the enzymatic reaction. In fact, the volume of the narrow compartment of the CT molecule (260-340 Å^3) is adequate for placement of seven water molecules capable of exchange with the environment. Such possibility for internal water molecules to exchange with the environment and temporal parameters of this exchange were earlier studied on the example of bovine pancreatic inhibitor and other proteins [12, 13].

Seven inner water molecules are enough to accept or give up 1.63 kJ in the hydrolysis reaction of 1 mole of methyl ester of N-acetyl phenylalanine involving CT [14]. One can approximately evaluate the temperature difference (ΔT) between water molecules inside the active site after splitting of the ester bond and external water molecules:

$$\Delta T = Q/(m \cdot c),$$

where Q = 1.63 kJ is the enthalpy of the hydrolysis reaction, m = 126 g (7 moles) is the mass of exchanged water, and c = 4.2 J/g·K is the heat capacity of water. This value is 3.5°C and it does not depend on substrate concentration and amount, because the ratio Q/m is constant for this reaction. It is self-evident that the mass of all external water is far greater than the exchanged mass, and the total temperature effect is barely perceptible.

It is worth noting that total reaction mass containing many CT, substrate, and water molecules forms a complex thermodynamic system. Description of its behavior is a difficult task [15]. Nevertheless, regarding the spatial structure of the active site of the CT molecule and its acceptability for external water molecules can be useful for understanding the characteristics of the enzymatic reaction.

For instance, one could suggest that a substitution of some amino acid residues near the narrow heat-exchange compartment by means of protein engineering might be significant for its physical properties and functioning of the whole CT molecule. In particular, a substitution of the Gly216 residue, whose CA-atom is localized near the middle of the compartment, by larger and more hydrophobic residues Ala, Val, or Leu would remarkably decrease the volume of the compartment and hamper the heat-exchange process.

Thus, in the present study a definition is given for the active site region of the CT molecule as the largest cavity on the protein surface bounded by an approximation sphere 20 Å in radius. This cavity contains the active site of the enzyme and is localized in a spherical sector with solid angle of 80°, whose axis passes near the CB-atom of Ser195 residue. The calculated volume of the active site cavity is about 2700 Å^3. Its size and shape — the intersection of approximating and additional spheres — enables the entrance of large fragments of substrate molecules containing more than 60 non-hydrogen atoms into the active site. A narrow compartment 260-340 Å^3 in volume, which is a part of oxyanionic hollow, is on the bottom of active site. Its space is inaccessible for substrate molecules, whereas the water molecules localized inside the compartment can participate in heat energy exchange with the environment at different stages of the enzymatic reaction.

REFERENCES

1. Laskowski, R. A., Luscombe, N. M., Swindells, M. B., and Thornton, J. M. (1996) *Protein Sci.*, **5**, 2438-2452.
2. Glaser, F., Morris, R. J., Najmanovich, R. J., Laskowski, R. A., and Thornton, J. M. (2006) *Proteins: Structure, Function, and Bioinformatics*, **62**, 479-488.

3. Stryer, L. (1988) *Biochemistry*, 3rd Edn., Vol. 1, W. H. Freeman and Co., New York.
4. Kostetsky, P. V. (2004) *Biophysics (Moscow)*, **49**, 552-556.
5. Artobolevsky, I. I. (1975) *Theory of Machines and Mechanisms* [in Russian], Nauka, Moscow.
6. Sobol, I. M. (1975) *Little Mathematics Library. The Monte-Carlo Method* [in Russian], Nauka, Moscow.
7. Williams, M. A., Goodfellow, J. M., and Thornton, J. M. (1994) *Protein Sci.*, **3**, 1224-1235.
8. Berman, H. M., Westbook, J., Feng, Z., Gilliand, G., Bhat, T. N., Weissig, H., Shendyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.*, **28**, 235-242.
9. Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry*, Vol. 1, *The Conformation of Biological Macromolecules*, Freeman, San Francisco.
10. Gedeon, V. F. (2000) *Biophysics*, **45**, 756-765.
11. Finkelstein, A. V., and Ptitsyn, O. B. (2002) *Protein Physics. A Course of Lectures*, Academic Press, An Imprint of Elsevier Science, Amsterdam-Boston-London-New York-Oxford-Paris-San Diego-San Francisco-Singapore-Sydney-Tokyo.
12. Otting, G., Liepinsh, G., and Wuetrich, K. (1991) *J. Am. Chem. Soc.*, **113**, 4363-4364.
13. Abseher, R., Schreiber, H., and Steinhauser, O. (1996) *Proteins*, **25**, 366-375.
14. Rekharsky, M. V., Rumsh, L. D., Antonov, V. K., and Gal'chenko, G. L. (1984) *Thermochim. Acta*, **81**, 67-174.
15. Marshell, A. (1978) *Biophysical Chemistry*, Vol. 1, Wiley, New York.