Reviews

Enzyme active sites: bioinformatics, architecture, and mechanisms of action

S. D. Varfolomeev★ *and K. G. Gurevich*

Chemistry Department, M. V. Lomonosov Moscow State University, Leninskie Gory, 119899 Moscow, Russian Federation. Fax: +7 (095) 939 5417. E-mail: sdvarf@enzyme.chem.msu.ru

A comparative analysis of the amino acid sequences of some enzymes which comprise superfamilies of enzymes belonging to different classes was carried out. Based on the amino acid sequence alignment for enzymes belonging to different classes with the use of the information entropy as a criterion, the amino acid residues involved in the catalytic portion of the active site are demonstrated to be most conservative. The rating scale for conservativeness of amino acids in enzymes is created. Glycine and aspartic acid are the most commonly occurring conservative amino acids essential for the catalysis. The role of aspartic acid and histidine in the mechanism of molecule activation in the catalytic site is considered using hydrolases as examples. The role of glycine, proline, and cysteine in the structural organization of the active sites is discussed.

Key words: bioinformatics, enzyme active site, conservative amino acid residues, aspartic acid, glycine.

Bioinformatics in analysis of enzyme active sites

Amino acid sequences are responsible for the structures and properties of proteins. Presently, it is evident that despite the enormous (virtually infinite) variability of proteins, particular structural elements are rather conservative and these elements govern to a large extent the function of the protein molecule. This is most pronounced in the case of proteins possessing the catalytic functions. For example, only five types of sites forming the catalytic structure were found for hydrolases, which comprise about one-third of all known enzymes (approximately 1100 enzymes of 3700).[1]

In the discussion of the structures of enzyme active sites, two structural components of the active site should be distinguished:

1) the sorption subsite responsible for binding, fixation, and orientation of substrates; the properties of this site determine the enzyme specificity;[2—4]

2) the catalytic subsite performing the chemical transformation of substrate molecules, as a rule, through the general acid-base catalysis.[5—9]

Within the framework of one superfamily, the sorption subsites responsible for the enzyme specificity would be expected to have diversified structures corresponding to a variety of substrate structures. At the same time, only a limited number of catalytic sites are known and these sites are actually conservative structural elements. To confirm this statement, we used the bioinformatical approach based on comparison of the amino acid sequences of proteins belonging to one large family.[10,11] We analyzed the results of the sequence alignment for several large enzyme families available in the HSSP

Published in Russian in *Izvestiya Akademii Nauk. Seriya Khimicheskaya,* No. 10, pp. 1629—1637, October, 2001.

1066-5285/01/5010-1709 $25.00

database (http://www.sander.embl-heidelberg.de/hssp/). The choice of the enzyme families was governed by the following criteria:

1) the number of the representatives of the family must be larger than 100, which assures the statistical reliability of the results;

2) enzyme families belonging to different classes (oxidoreductases, hydrolases, isomerases, *etc.*) should be analyzed;

3) it is desirable, where possible, to choose enzymes with the known structures of the active sites and with the catalysis mechanisms determined with a high degree of assurance.

Generally, the results of sequence alignments are presented as large tables obtained by superimposing the protein sequences on the sequence taken as the reference. The conservative elements of the amino acid sequence can be revealed by visual comparison. Evidently, this approach is extremely inefficient and is of little use in the case of comparison of more than three—five proteins. This procedure can be automated by obtaining quantitative characteristics of the conservativeness of the amino acid sites in the sequence. The Shannon entropy statistical criterion can serve as a quantitative criterion for the conservativeness of each amino acid in the protein sequences.[12—14] It should be noted that Shannon's entropy is one of the most important functions in informatics. This function was proposed as a measure of uncertainty characterizing a particular event occurring with certain probability. The information can be determined as a measure of uncertainty, which is refined once the information is obtained. The information quantity is formally represented as the difference between the information entropies before and after the experiment (after the information acquisition). The information entropy (Shannon's entropy) is a very convenient function for comparing related proteins with different amino acid sequences. The procedure for the sequence alignment relative to a reference protein involves positioning of the sequences one above the other with the fixation of homologous regions and determination of deletions and insertions. When performing such a comparison for a large number of proteins, one can obtain a sufficiently accurate estimate of the probability of the occurrence of a particular amino acid in each site of the protein sequence. This probability is determined as the relative frequency of occurrence of the amino acid j in the site i. For each site in the protein sequence, the entropy function can be calculated for all 20 amino acids[12]

$$H_j = -\sum_i p_i^j \log_2 p_i^j .$$

An important characteristic feature of this function is the fact that it is close to zero for events with high probability ($p_i^j \to 1$) as well as for events with low probability ($p_i^j \to 0$). Hence, calculations of Shannon's entropy can reveal sites in the sequence which are common (absolutely conservative) to the j-th amino acid throughout the large protein family. In such sites, the probability of the occurrence of a particular acid is close to unity, whereas the probabilities of the occurrence of all other acids are close to zero. High Shannon's entropies are characteristic of sites in which various amino acids can be located, whereas low Shannon's entropies are typical of conservative amino acid sites. In the limit $p_i^j \to 1$ (absolute conservativeness), $H_j \to 0$.

The Shannon entropy criterion was used in the analysis of protein superfamilies based on the following enzymes: trypsin/chymotrypsin, subtilisin, lipase, cytochrome-c oxidase, pepsin/pepsinogen, enolase, lactate dehydrogenase, DNase, aminoacyl-tRNA synthase, and lysozyme. The families include proteins with diversified structures and of various origin. For example, the trypsin family consists of more than 500 proteins, including such enzymes as chymotrypsin, kallikrein, plasmin, hypostatin, neuropsin, coagulation factors IX and X, thrombocyte aggregation protease, hepatocyte growth factor activator, elastase, transmembrane triptase, thrombin, and many other proteins. The data on Shannon's entropies for the multiple sequence alignments for the subtilisin and lipase families are presented in Figs. 1 and 2, respectively.

In our opinion, it is of interest to consider highly conservative amino acids for which $H_j = 0$ (or close to zero). Analysis allowed us to reveal the following regularities.

1. Amino acids comprising the catalytically active site always manifest themselves as conservative elements in the amino acid sequence alignment of enzymes. It is known that the catalytic site of acid proteases of the pepsin type involves carboxy groups of two aspartic acid residues Asp32 and Asp215.[15—17] These aspartic acid residues reveal themselves as conservative sites with minimum Shannon's entropies in the sequences alignment of the proteins belonging to the pepsin family.

The subtilisin and lipase families[18,19] (see Figs. 1 and 2, respectively) possess the Ser-His-Asp triad (the residues Ser214, His63, and Asp32 in the subtilisin family) as a catalytically active group, which performs the proton transfer (the charge relay system) and catalytic water activation according to the nucleophilic mechanism.[6]

2. Comparison of the amino acid sequences of the enzymes reveals that glycine and aspartic acid most often occur as absolutely conservative amino acids. The set of the enzyme families under examination (trypsin, subtilisin, lipases, alcohol dehydrogenase, lactate dehydrogenase, cytochrome-c oxidase, pepsin/pepsinogen, enolase, lysozyme, alkaline phosphatase, aminoacyl-tRNA synthase, and DNase) includes 2281 proteins and is represented by 749252 amino acid sites. The fact that glycine is the most conservative amino acid is somewhat unexpected. Aspartic acid is the second most conservative amino acid. Glycine and aspartic acid taken together account for ~50% of all conservative amino acids.

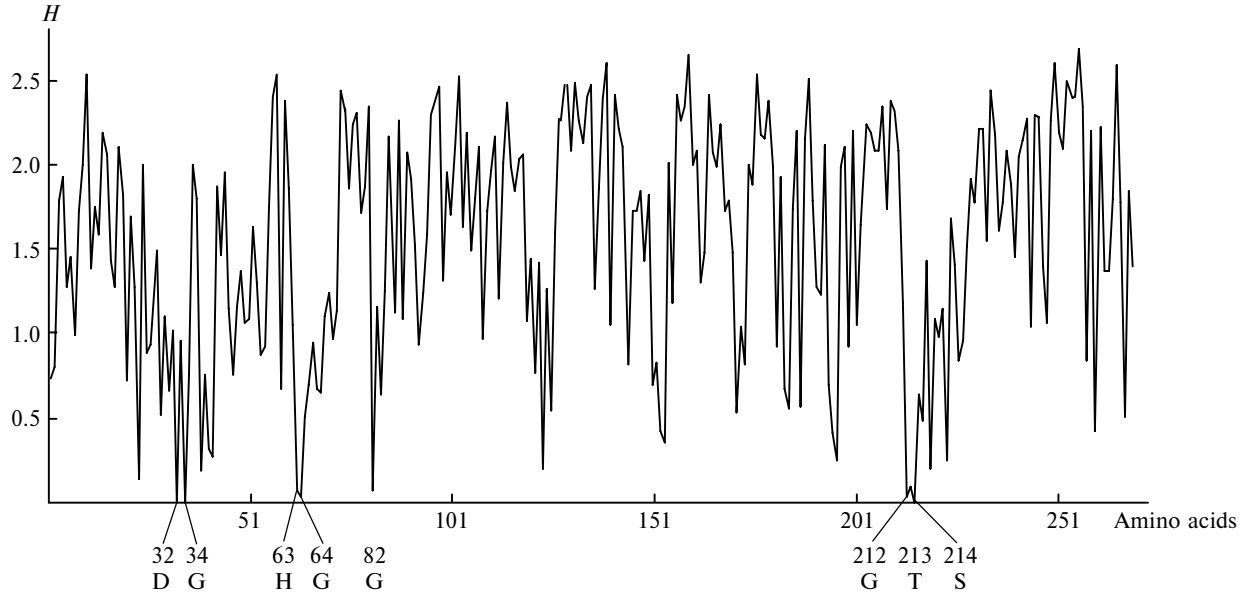


Fig. 1. Subtilisin family. Shannon's entropy (H) was calculated for the sites of the multiple sequence alignment relative to subtilisin (286 proteins were compared).

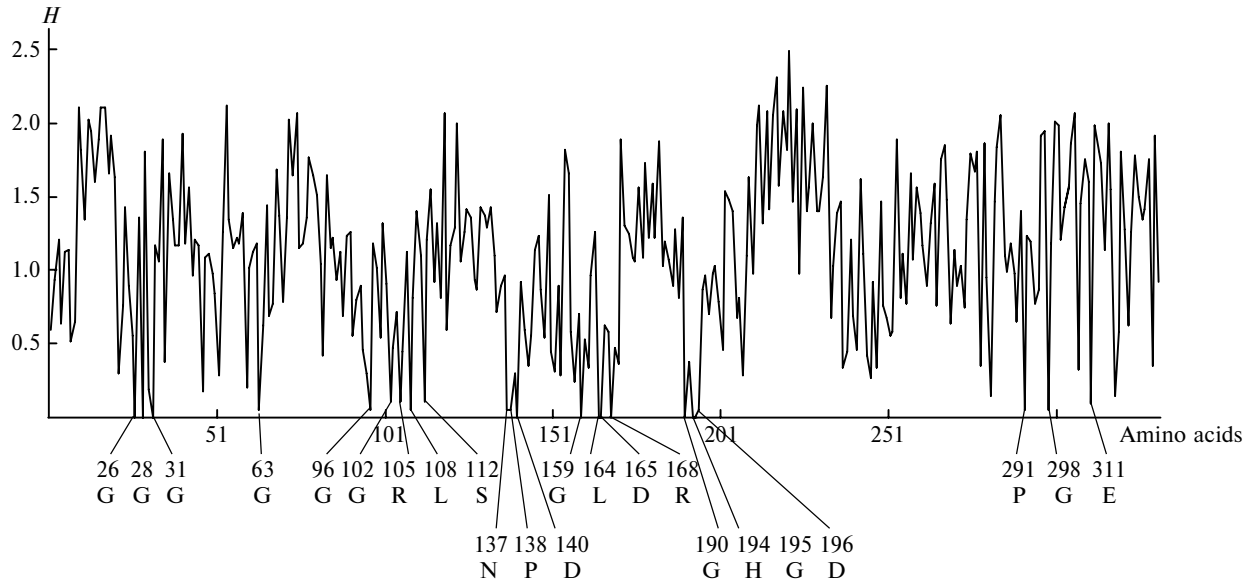


Fig. 2. Lipase family. Shannon's entropy (H) was calculated for the sites of the multiple sequence alignment relative to lipase (120 proteins were compared).

The rating scale of conservativeness for amino acids in the families under study was created. For this purpose, we estimated the frequency of occurrence of each amino acid as a conservative element ($H_j \approx 0$) referred to the total number of conservative sites for all amino acids in the families under examination. The rating scale of conservativeness for amino acids is shown in Fig. 3.

It can be seen that glycine, aspartic acid, cysteine, proline, and histidine are the most commonly occurring conservative amino acids in the enzyme sequences; these amino acids account for ~70% of all conservative sites in enzymes. Methionine and isoleucine are virtually never present as conservative amino acids.

It is reasonable to divide the most conservative amino acids into two radically different groups:

1) amino acids involved in elementary events of substrate activation as acids and bases (aspartic acid and histidine);

2) amino acids responsible for the architectonics of the active site (glycine, cysteine, and proline).

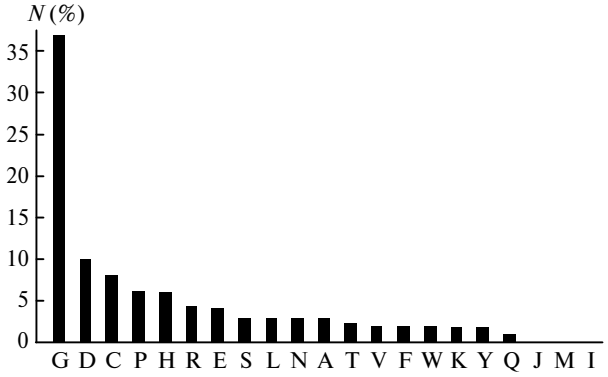


Fig. 3. Rating scale of conservative amino acids. *N* is the frequency of occurrence of amino acids as conservative elements (%).

Aspartic acid and histidine in the enzyme catalytic cycle

The bioinformatical approach employed demonstrated the great role of aspartic acid and histidine in the functioning of the enzyme active sites. Let us consider this role in more detail based on the analysis of the mechanisms of action of hydrolases. The choice of hydrolases is governed, on the one hand, by the fact that hydrolases represent the largest enzyme class (about one-third of all known enzymes) and, on the other hand, by the fact that the molecular mechanisms of catalysis by hydrolases are best understood. For most of hydrolases, the groups comprising the catalytically active sites are well identified and the justified concepts of the role of interactions between these groups in the mechanism of the catalytic cycle are available.

Hydrolases can arbitrarily be divided into five main types according to the structures of the active sites and the mechanisms of action:

1) hydrolases whose active sites contain aspartic or glutamic acid (the lysozyme-pepsin type);

2) hydrolases whose active sites contain the hydroxy group of serine or threonine or the sulfhydryl group of cysteine and the charge relay system activating this group (the chymotrypsin type);

3) hydrolases which use the imidazole group directly for water activation (the type of pancreatic ribonuclease);

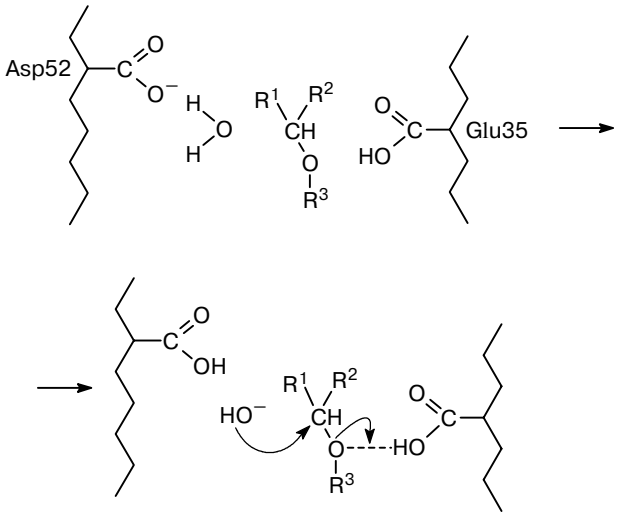
4) hydrolases which use Zn^{2+} or Co^{2+} complexes for activation of water and the substrate (the types of alkaline phosphatase, carboxypeptidase A, or organophosphate hydrolase);

5) hydrolases which use Mg^{2+} or Mn^{2+} ions for activation of water and the substrate (the pyrophosphatase type).

Analysis of the catalysis mechanisms revealed that aspartic acid and histidine are of primary importance in most types of catalytic sites. Let us exemplify the mechanisms of catalysis by hydrolases using a number of chemical schemes.

Hydrolases of the lysozyme-pepsin type. The catalytic subsites of this type of enzymes contain two or several carboxy groups.[6,9] This enzyme group involves enzymes cleaving esters and amides as well as glycosidic bonds. This group of enzymes is generally characterized by the rather low pH optimum. Pepsin and lysozyme are typical representatives of this group of hydrolases. The protonated carboxy group is involved in the substrate activation as a general acid and the deprotonated carboxy group serves as a base in the water activation. The concerted mechanism of action of two carboxy groups serving as a nucleophile (the water activator) and an electrophile (the substrate activator) have been the subject of comprehensive consideration.[5—8] In lysozyme, the protonated carboxy group of Glu35 activates the reaction site and induces the electron density deficiency through the proton donation. The deprotonated carboxy group of Asp52 activates water according to the general base mechanism, the reactivity of water approximating that of the hydroxyl ion (Scheme 1).

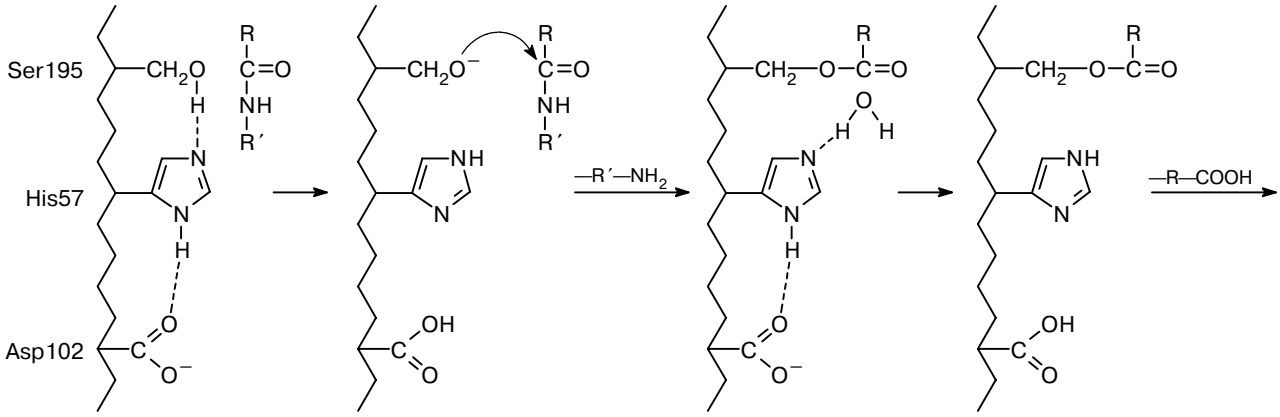
Scheme 1



Reactions of this enzyme class proceed under the concerted action of two carboxy groups one of which in the ionized form acts as a nucleophilic activator of water, while another group in the protonated form serves as an electrophilic activator of the substrate.

Hydrolases of the chymotrypsin type. Chymotrypsin, trypsin (serine proteinases), and papain (cysteine proteinase) are typical representatives of this type of enzymes. The catalysis mechanism of these enzymes involves intermediate acylation of the hydroxy or the thiol group of the active site to form acyl-enzyme intermediates.[2,3,20,21] Hydrolases of the chymotrypsin type contain the charge relay system involving the carboxy group which participates in the activation of either water or the hydroxy group of serine in the acylation stage. The mechanism of catalysis by this type of enzymes has been

Scheme 2



surveyed in detail in a number of monographs.[2,6,7] In the case of α-chymotrypsin, water is activated by the imidazole ring of His57 linked with Asp102 by a hydrogen bond. The catalysis mechanism of chymotrypsin is shown in Scheme 2. The hydroxy group of serine and the water molecule are activated by the charge relay system involving the imidazole ring and the carboxy group. It should be noted that the possibility of isomorphous replacements of the catalytic groups by structurally related groups are very characteristic of the enzyme active sites. Thus serine can be replaced by threonine without substantial changes in the catalysis mechanism. This problem has been thoroughly examined.[11]

Hydrolases of the type of pancreatic ribonuclease. The mechanism of catalysis by this type of enzymes is provided by the concerted action of the imidazole groups of two histidine residues. The imidazole group of His12 activates the carbohydrate hydroxy group through the proton transfer in the stage of formation of a cyclophosphate intermediate. The imidazole group of His119 is involved in water activation in the stage of hydrolysis of cyclophosphate.[22]

The type of alkaline phosphatase or organophosphate hydrolase. The active sites of enzymes of this group contain metal ions (Zn^{2+}, Co^{2+}, or Ni^{2+}), which form complexes with the protein. These complexes have a unified structure. The functional groups of histidine (imidazole) and aspartic or glutamic acid (the carboxy group) serve as ligands for the metal ion. Carboxypeptidase A and thermolysin are typical representatives of this enzyme group.[6] The metal ions (Zn^{2+} or Co^{2+}) form structurally similar complexes with the proteins. The imidazole group of histidine and the carboxy group of aspartic or glutamic acid serve as analogs of metal ions. It is known that metal ions in the enzyme active sites can perform the dual function:[23] 1) act as an electrophilic agent activating the reaction site; and 2) serve as an electrophilic activator of water, *viz.*, as a generator of hydroxyl ions. The complex of the metal ion in the active sites of carboxypeptidase and thermolysin serves as an electrophilic agent activating the reaction site of the substrate subjected to the attack. Water is activated by the base, *viz.*, by the deprotonated carboxy group. Organophosphate hydrolase, which catalyzes hydrolysis of organophosphorus pesticides and chemical warfare agents (sarin, soman, and VX), exemplifies this type of substrate activation.[24—28] Catalytic hydrolysis under the action of this type of enzymes proceeds with the concerted participation of two complexes of the Zn^{2+} (or Co^{2+}) ions. The catalytic action of organophosphate hydrolase performing hydrolysis of phosphoric esters can be represented by Scheme 3.[29]

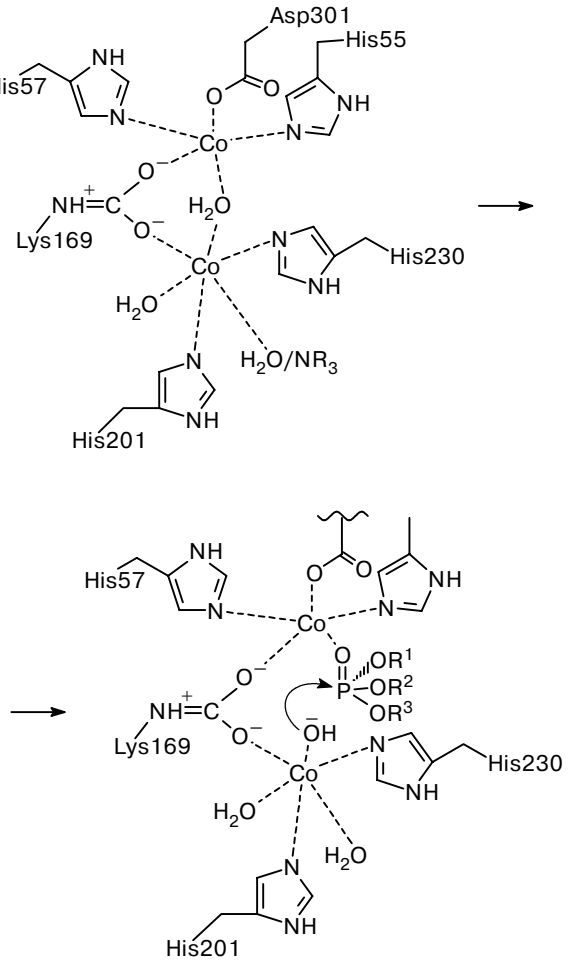
The pyrophosphatase type (Mg^{2+}- and Mn^{2+}-dependent hydrolases). The metal complexes in the active sites of hydrolases typically act as strong electrophilic agents activating the substrate by inducing the electron density deficiency on the reactive site. In addition, the metal complex can be involved in the electrophilic activation of water. This allocation of the functions of the metal ions is most clearly evident from analysis of the structures and catalysis mechanisms of Mg^{2+}(Mn^{2+})-dependent enzymes, in particular, of inorganic pyrophosphatase.[30—34] The substrate is fixed in the enzyme active site through numerous ionic interactions involving positively charged groups of Lys193, Arg78, and Lys56. The major catalytic events occur, apparently, with the participation of two other Mg^{2+} or Mn^{2+} ions. The carboxy groups of Asp65, Asp70, and Asp102 and three water molecules serve as ligands for the high-affinity complex-forming Mg^{2+} or Mn^{2+} site. The low-affinity metal site is formed by the carboxy groups of aspartic acid residues, five water molecules, and the carboxy group of Glu20. Formally, the complex is positively charged. The assumption that the high-affinity site serves as the electrophilic activator of water (the generator of hydroxyl ions) and the low-affinity site acts as an electrophilic agent (the substrate activator) seems to be reasonable (Scheme 4).

The above-considered examples evidence the role of aspartic acid and histidine in the mechanisms of action of enzymes. The functioning of the enzyme active sites is based on the principle of the concerted action of the

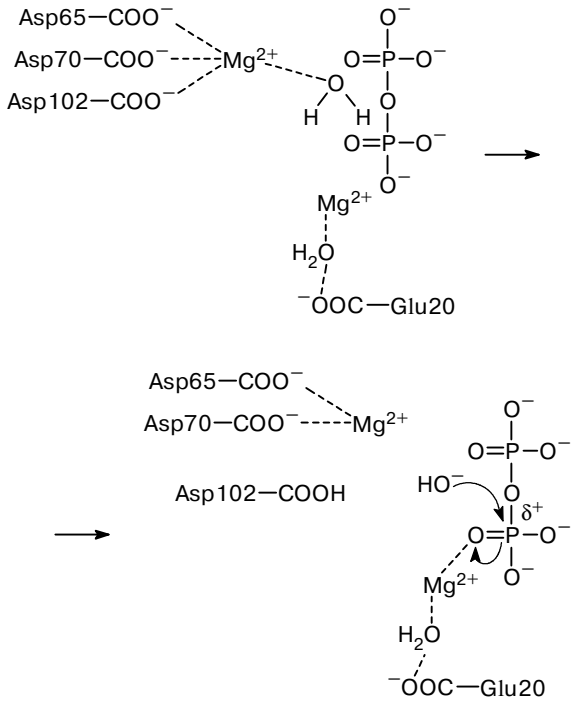
nucleophilic and electrophilic components of the active site, which provides essential acceleration of the reactions.[1] Apparently, aspartic acid is of fundamental importance in these processes. The ionized form of the carboxy group of aspartic acid acts as a powerful nucleophilic reagent in the water activation and in the charge relay systems (pepsin, lysozyme, and α-chymotrypsin). Aspartic acid is also of basic importance in the formation of metal complexes giving rise to the active sites of metal-dependent enzymes. In the protonated form, the carboxy group of aspartic acid serves as a proton donor, thus acting as an electrophilic agent.

The relationship between enzymes hydrolyzing individual bonds and enzymes forming these bonds is of interest. The structure of the active site of exo-nucleases hydrolyzing DNA was established.[35—39] The complex of magnesium ions which is formed with the participation of the carboxy groups of aspartic and glutamic acids serves as the active site of the pyrophosphatase type. The active site of the DNA-synthesizing enzyme, *viz.*, DNA polymerase, is highly homologous to the active site of this nuclease.[35—39] The active sites of both

Scheme 3



Scheme 4



enzymes are structurally very similar. In DNA polymerase, the reaction site and the nucleophile are activated by the same functional groups; however, in DNA polymerase the hydroxy group of the sugar fragment rather than the water molecule (by abstracting the proton) is activated. The nucleophilic attack at the reaction center of the phosphate group provides the DNA-chain growth.

It should be noted that the structural elements comprising the active sites of hydrolases can be found in the active sites of other enzyme classes. For example, the charge relay system containing the imidazole group together with the carboxy group of aspartic acid is often observed in the active sites of oxidoreductases. The structures of peroxidases and the protein environments about the heme were considered.[40—42] It was demonstrated that the imidazole—carboxy group pair in this type of enzymes is directly involved in the structure of the catalytic site and apparently acts as a common nucleophilic agent activating the iron ion of the heme. Analogously, the structural element of the charge relay system, *viz.*, the intraprotein complex of imidazole with the carboxy group, can be identified in the active sites of dehydrogenases.[43—45]

The role of glycine in the formation of the architecture and assistance in the conformational flexibility of the active site

The role of glycine in the formation and functioning of the active sites is less well understood compared with

the role of aspartic acid. Clearly, conservative glycine residues are of little importance in the chemical events of molecule activation in the catalytic cycle. Since glycine possesses no substituents at the α-carbon atom, it is devoid of the pronounced chemical function.

Nevertheless, the glycine residues are significant in the protein structure.[46—51] The fact that conservative glycine residues are of fundamental importance in the enzyme catalysis was evidenced by experiments in which the site-specific replacements of these conservative residues by different amino acids were carried out. Generally, such replacements led to the complete loss or a substantial decrease in the enzyme activity (see, for example, Ref. 49).

Apparently, conservative glycine residues are of fundamental importance for two reasons.

1. Being a unique amino acid characterized by the low energy barrier to rotation about the C—N and C—C bonds of the polypeptide chain (the φ and ψ angles according to Ramachandran), glycine can serve as a node, which makes possible the change of the polypeptide-chain folding upon the "assembly" of amino acid residues to form the active site. Therefore, the presence of conservative glycine residues accounts for the structural paradox of the enzyme catalysis, namely, the fact that the identical active sites are "assembled" of quite different polypeptide chains. Common to these chains are the presence of conservative glycine residues and the factors stabilizing the structure, for example, through disulfide bonds (cysteine is also a highly conservative residue, being the third most conservative amino acid). Interestingly, conservative glycine residues in enzymes are generally "twisted" with respect to the ψ angle (the rotation around the C—C bond in the amino acid). Figure 4 shows the Ramachandran plot where the φ and ψ angles for conservative glycine residues in a large group of enzymes are indicated by points. It can be seen that the angles of rotation about the C—N bonds have arbitrary values and vary from −180° to +180°, whereas the angles of rotation about the C—C bonds are grouped about the maximum "twisted" values for glycine residues (shaded zones).

2. Conservative glycine residues can act as conformational "hinges" providing certain conformational flexibility of the active site. This is supported by the fact that conservative glycine residues are observed in many proteins in the vicinity of the catalytically active groups. For example, the following conservative fragments were found in hydrolases belonging to different families: Asp215XGly217 (pepsin); Asp170XXGly173 (thermolysin); Asp32XGly34, His63Gly64, and Gly119XSer221 (subtilisin); Gly173XSer177 (trypsin); and His76Gly77, Ser153XGly155, and Gly175XAsp177 (lipases). The amino acids Asp, Ser, and His in the above-mentioned enzymes are involved in the active sites.

It should be noted that the φ and ψ angles for amino acids involved in some catalytically active sites do not fall within the ranges of energetically "relaxed" values.

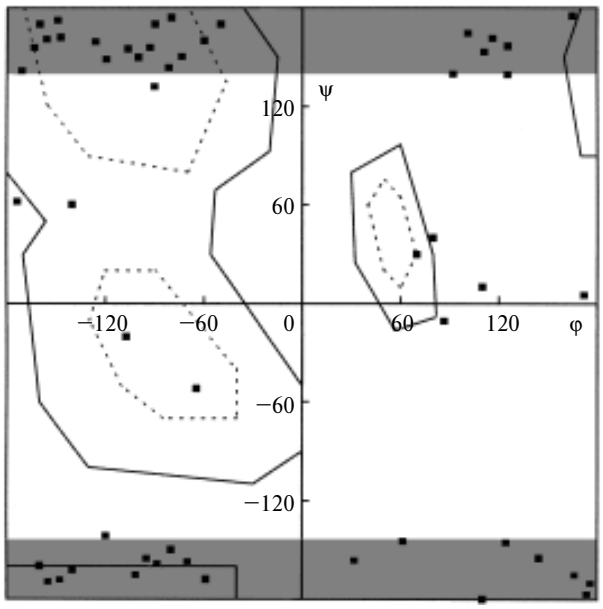


Fig. 4. The Ramachandran angles for conservative glycine residues (the selection of enzymes with the known structures of the active sites includes α-chymotrypsin, pepsin, subtilisin, thermolysin, lipase, DNA polymerase, cytochrome-c oxidase, lactate dehydrogenase, and carboxypeptidase A). The isoenergetical contours for the Ramachandran angles: the regions bounded by the dashed and solid lines correspond to the minimum and medium energies, respectively. The Ramachandran angles characterized by the optimum energy are located within the region enclosed by the dashed line.*

For example, this is evident from the constructions of the Ramachandran plot for the amino acids comprising the active site of α-chymotrypsin (His57, Asp102, and Ser195). The active site of this enzyme is conformationally strained (the φ and ψ angles are located in the energetically unfavorable region).

The conversion of the initial substrate into the final products in the enzyme catalysis proceeds through a large number of intermediates, which differ in the structure from the initial substrate. The glycine residues of the active site can act as "relaxing" elements ensuring the conformational tuning of the active site to the subsequent elementary event.

Cysteine and proline (the fourth and fifth most conservative amino acids, respectively) play a great role in the architecture of the active site. It is known that proline is a unique imino acid providing the turn of the polypeptide chain.

Apparently, cysteine has the function of forming disulfide bridges, which fix the required structure of the active site composed of various and often very remote regions of the polypeptide chain. For many enzymes, this completes the formation of the architecture of the active site.

* The data from the Protein Data Bank were used (http://www.rcsb.org/pdb).

Conclusion

Amino acid sequence comparison and alignment of a large number of related enzymes established a series of amino acid sites which are reproduced in all representatives of the family. Comparison of the proteins is most convenient to perform using Shannon's entropy as the statistical information characteristic.

The bioinformatical approach employed made it possible to reveal conservative amino acids for each of the related enzymes under study and to construct the rating scale of conservative amino acids throughout the enzyme catalysis. It appeared that conservative residues of glycine and aspartic acid are of primary importance for the structures of enzymes and the manifestation of the catalytic function. It was found that amino acids involved in the catalytic subsite of the enzyme active site are always highly conservative.

Comparison of the bioinformatical characteristics with the data on the mechanisms of action of the catalytic subsites for a series of well-studied enzymes demonstrated that the carboxy group of aspartic acid is the most commonly occurring nucleophilic (in the deprotonated form) and electrophilic (in the protonated form) agent involved in the molecule activation according to the mechanism of general acid-base catalysis in the enzyme catalytic sites.

The role of conservative glycine residues is less well understood. Glycine is a unique amino acid characterized by the enhanced capabilities for rotating about the C—N and C—C bonds of the polypeptide chain. Probably, the conservative fixation of glycine in the polypeptide chains of related enzymes provides the polypeptide-chain folding to form the structure of the catalytic subsite. It is known that the catalytic subsite is composed of amino acid residues located in different regions of the polypeptide chain. The compositions and the three-dimensional structures of the catalytic subsites are identical for various polypeptides forming the protein molecules. Apparently, conservative glycine, proline, and cysteine residues play a decisive role in the architecture of the active site. It is also probable that conservative glycine residues ensure conformational flexibility of proteins and their active sites.

We thank V. V. Poroikov, B. N. Sobolev, and A. E. Fomenko (the Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences) for helpful discussion and valuable advice.

References

1. S. D. Varfolomeev and A. E. Pozhitkov, *Vestn. Mosk. Univ., Ser. 2, Khim.*, 1999, **3**, 147 [*Moscow Univ. Chem. Bull.*, 1999, **3** (Engl. Transl.)].
2. I. V. Berezin and K. Martinek, *Osnovy fizicheskoi khimii fermentativnogo kataliza* [*Physicochemical Basics of Enzyme Catalysis*], Vysshaya Shkola, Moscow, 1977, 280 pp. (in Russian).
3. S. D. Varfolomeev and S. V. Zaitsev, *Kineticheskie metody v biokhimicheskikh issledovaniyakh* [*Kinetic Methods of Biochemical Investigations*], Izd. MGU, Moscow, 1982, 344 pp. (in Russian).
4. A. A. Klesov, *Uspekhi biologicheskogo kataliza* [*Comprehensive Biological Catalysis*], Izd. MGU, Moscow, 1979, 7 (in Russian).
5. O. M. Poltorak and E. S. Chukhrai, *Fiziko-khimicheskie osnovy fermentativnogo kataliza* [*Physicochemical Basics of Enzyme Catalysis*], Vysshaya Shkola, Moscow, 1971, 311 pp. (in Russian).
6. V. K. Antonov, *Khimiya proteoliza* [*Chemistry of Proteolysis*], Nauka, Moscow, 1991, 504 pp. (in Russian).
7. M. L. Bender, *Mechanisms of Homogeneous Catalysis from Proton to Proteins*, Wiley Interscience, New York, 1971, 362 pp.
8. W. P. Jencks, *Catalysis in Chemistry and Enzymology*, McGraw Hill, New York—London—Mexico, 1969.
9. M. Dixon and E. C. Webb, *Enzymes*, 3d Ed., Longman, London, 1979.
10. W. R. Taylor, *J. Biotechnol.*, 1994, **35**, 281.
11. G. Dodson and A. Wlodawer, *Trends Biochem. Sci.*, 1998, **23**, 347.
12. A. M. Yaglom and I. M. Yaglom, *Veroyatnost′ i informatsiya* [*Probability and Informatics*], Nauka, Moscow, 1973, 511 pp. (in Russian).
13. C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Illinois, 1949.
14. A. G. Wilson, *Entropy in Urban and Regional Modelling*, Pion Limited, London, 1970.
15. N. S. Andreeva, A. E. Gurtliana, and A. Fedorov, *Adv. Exp. Med. Biol.*, 1977, **95**, 23.
16. N. S. Andreeva, A. S. Zhdanov, and A. A. Fedorov, *Mol. Biol.*, 1984, **18**, 313 [*Mol. Biol. USSR*, 1984 (Engl. Transl.)].
17. A. R. Shielecki, A. A. Fedorov, and A. Boodhoo, *J. Mol. Biol.*, 1990, **214**, 148.
18. *Lipases*, *Their Structure*, *Biochemistry and Application*, Eds. P. Wooley and S. Petersen, Cambridge University Press, Cambridge, 1994, 360 pp.
19. F. Dratlas and E. B. Peterson, in *Methods in Enzymology*, Eds. J. N. Atelson and M. I. Simon, 1997, **284**, 28.
20. S. D. Varfolomeev and K. G. Gurevich, *Biokinetika* [*Biokinetics*], Fair-press, Moscow, 1999, 720 pp. (in Russian).
21. K. Martinek, V. N. Dorovska, S. D. Varfolomeev, and I. V. Berezin, *Biochim. Biophys. Acta*, 1972, **271**, 8086.
22. D. G. Knorre and S. D. Myzina, *Biologicheskaya khimiya* [*Biological Chemistry*], Vysshaya Shkola, Moscow, 1992, 420 pp. (in Russian).
23. *Inorganic Biohemistry*, Ed. G. L. Eichhorn, Elsevier, Amsterdam, Oxford, New York, 1973.
24. J. K. Crimsky, J. M. Scholtz, C. N. Pace, and J. R. Wild, *Biochemistry*, 1992, **36**, 14366.
25. K. Lay, K. I. Dave, and J. R. Wild, *J. Biol. Chem.*, 1994, **269**, 16579.
26. I. G. Danilova, A. D. Ryabov, and S. D. Varfolomeev, *J. Mol. Cat.*, 1997, **118**, 161.
27. V. S. Sergeeva, E. N. Efremenko, G. M. Kazankov, A. K. Gladilin, and S. D. Varfolomeev, *Biotechnol. Techniques*, 1999, **13**, 479.
28. G. A. Omburo, J. M. Kuo, L. S. Mullins, and F. M. Rushel, *J. Biol. Chem.*, 1992, **267**, 13278.
29. V. S. Sergeeva, E. N. Efremenko, G. M. Kazankov, and S. D. Varfolomeev, *J. Mol. Cat. Part B: Enzymatic*, 2000, **10**, 571.
30. P. H. Burton, D. C. Hall, and J. Josse, *J. Biol. Chem.*, 1970, **245**, 4346.

31. C. M. Avaeva and T. I. Nazarova, *Usp. Biol. Khim.*, 1985, **26**, 42 [*Adv. Biol. Chem.*, 1985, **26** (in Russian)].
32. Yu. V. Oganessyan, S. A. Kurilova, N. N. Vorobyova, T. I. Nazarova, A. N. Popov, A. A. Lebedev, S. M. Avaeva, and E. H. Harutyunyan, *FEBS Lett.*, 1974, **348**, 301.
33. E. G. Arutyunyan, V. Yu. Oganesyan, N. N. Oganesyan, S. S. Terzyan, A. N. Popov, S. V. Rubinskii, B. K. Vainshtein, T. I. Nazarova, S. A. Kurilova, N. N. Vorob′eva, and S. M. Avaeva, *Kristallografiya*, 1996, **41**, 84 [*Crystallogr. Repts.*, 1996, **41** (Engl. Transl.)].
34. E. H. Harutyunyan, V. Yu. Oganessyan, N. N. Oganessyan, S. M. Avaeva, T. I. Nazarova, N. N. Vorobyova, S. A. Kurilova, and R. Huber, *Biochemistry*, 1996, **35**, 7754.
35. T. A. Steitz, *J. Biol. Chem.*, 1999, **274**, 17395.
36. D. Sam and G. G. Perona, *Biochemistry*, 1999, **38**, 6576.
37. G. S. Baldwin, R. B. Sessions, S. G. Erskine, and S. E. Halford, *J. Mol. Biol.*, 1999, **288**, 87.
38. H. Viadiu and A. K. Aggarwal, *Nature Struct. Biol.*, 1998, **5**, 910.
39. L. Holm and C. Sandek, *Science*, 1996, **273**, 595.
40. I. G. Gazaryan, *Itogi nauki i tekhniki. Biotekhnologiya* [*Results of Science and Engineering. Biotechnology*], VINITI, Moscow, 1992, **36**, 4 (in Russian).
41. B. C. Finzel, T. L. Poulos, and J. Kraut, *J. Biol. Chem.*, 1984, **259**, 13027.
42. N. Kunishima, K. Fukuyama, H. Matsubara, H. Hatanaka, Y. Shibano, and T. Amachi, *J. Mol. Biol.*, 1994, **235**, 331.
43. J. J. Birktoft and L. J. Banaszak, *Biol. Chem.*, 1983, **258**, 472.
44. A. R. Clarke, H. M. Wilks, D. A. Barstow, T. Atkinson, N. N. Chia, and J. J. Holbrook, *Biochemistry*, 1988, **27**, 1617.
45. H. Taguchi, T. Ohta, and H. Matsuzawa, *J. Biochem. (Tokyo)*, 1997, **122**, 802.
46. N. G. Esipova and V. G. Tumanyan, *Mol. Biol.*, 1972, **6**, 840 [*Mol. Biol. USSR*, 1972, **6** (Engl. Transl.)].
47. M. Sasai, *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 8438
48. I. Friedberg, T. Kaplan, and H. Margalit, *ISMB*, 2000, **8**, 162.
49. Z. Q. Cheng and B. A. McFadden, *Protein Eng.*, 1998, **11**, 457.
50. A. C. Martin and J. M. Thornton, *J. Mol. Biol.*, 1996, **263**, 800.
51. A. V. Efimov, *FEBS Lett.*, 1999, **355**, 213.

Received March 5, 2001;
in revised form June 15, 2001