J. Modern Technology & Engineering
Vol.2, No.2, 2017, pp.140-145

IMPROVEMENT OF PROGRAM OF CALCULATION OF MOLECULAR CONFORMATION

G.A. Akverdieva[1*], N.M.Godjayev[1]

[1]Institute for Physical Problems, Baku State University, Baku, Azerbaijan

Abstract. The computer program of semiempirical calculation of the biopolymer conformation has been improved. The algorithm of the call of the necessary information about amino acid residues of the peptide molecule from the direct access file was introduced to the appropriate subroutine of the modified version of the calculation program. The created file is a database of the naturally occurring amino acid residues and their analogs, modified in view of chirality and of their arrangement in the terminal positions of the amino acid sequences with different atomic groups at the terminals. The information on the chemical structure and geometry of each amino acid residue, namely, the signs of atoms, their arrangement, the phases of atoms, the atomic charges, the values of the bond lengths and of the bond angles, the signs of rotation around certain chemical bonds, the parameters of the torsion potentials have been included in this database. The program allows us to calculate the energy and geometrical parameters of the both isolated molecules and molecular complex. The innovation allows avoiding heaps of inputs used in the construction of molecules carry the unmistakable assembling of the desired amino acid sequence.

Keywords: *conformation, calculation, program, direct-access file*

Corresponding Author: *Gulnara, Akverdieva, Ph.D., Institute for Physical Problems, Baku State University, Z.Khalilov str.23, AZ1148, Baku, Azerbaijan, e-mail: hagverdigulnara@gmail.com*

Manuscript received: *4 May 2017*

1. Introduction

All biological processes occurring in living organisms are related to the functioning of biomolecules. The study of the spatial structure of such molecules makes it possible to understand the mechanism of their action, their reactivity. It is known that conformational changes play an important role in the activity of biomolecules; therefore, the study of the relationship between the structure of a molecule and its properties is associated with the definition of its dynamic capabilities, the potency of this structure to changes. Of the available theoretical approaches to the study of the conformations of such complicated molecules as oligopeptides, hormones and proteins, the semiempirical method is the most developed and promising one. This method allows us to successfully solve the problems of studying the structural and functional relationship of molecules. A special service program for the semiempirical calculation of molecular conformations was developed in the research laboratory of the Molecular Biophysics of the Baku State University [16]. This program allows you to analyze the structure of both an isolated molecule and a molecular complex, taking into account the intermolecular interaction at various stages of complexation - the non-valent complex, the Michaelis complex, the valence complex important in the process of enzymatic catalysis and also at ligand-receptor binding. The program

is based on the matrix method for determining of the coordinates of atoms, proposed by Hermans and Ferro [17], and potential functions with parametrization, developed under the leadership of Sheraga [18-20]. The presented program allows us to calculate the energy and geometrical of the molecules under study. In the current version of this program, the energy is calculated as the sum of the independent contributions of the energies of the nonvalent, electrostatic interactions, and the energies of torsion barriers and hydrogen bonds. The value of the contribution of the energy of non-valence interactions is determined using the Lenard-Johnson potentials, the energy of electrostatic contacts is calculated in the monopole approximation according to Coulomb's law with $\varepsilon = 10$, which corresponds to the conditions of the aqueous environment. If necessary, a value of ε is taken as 4, which corresponds to a low polarity of the medium in the contact area of the active center and the substrate. The hydrogen bond energy is estimated using the Morse potential with a hydrogen bond dissociation energy value of 1.5 kcal/mol corresponding to the NH ... OC r = 1.8Å bond distance for aqueous solutions. The energy of the sorbed molecule is minimized by the Powell-Davidson-Fletcher method [12,13]. The search for a productive orientation of the substrate relative to the enzyme or receptor is made on the basis of the calculation of the energy gradient of the intermolecular interaction. The program also provides for a simultaneous change of the conformations of interacting molecules during the formation of the complex. With the help of this program, scientific research has been carried out in various directions [1-11, 14-15]. Thus, it has been tested on a number of both isolated molecules and intermolecular interactions at various stages of complexation. This program is implemented in the leading scientific centers of Azerbaijan, Turkey, Russia and Ukraine.

To determine the stable states of both an isolated molecule and each of the interacting molecules in the complex, it is necessary to input the initial data required for their construction, the so-called "molecule assembling". We note that in the case of studying enzyme-substrate or receptor-substrat interactions, the structure of the enzyme or receptor is taken from X-ray diffraction analysis, information on the productive conformation of the substrate can also be used effectively.

2. Discussion

The program includes 15 subroutines. The description of all program identifiers and input of initial data is carried out in the MAINPGM main subroutine, this subroutine is associated with all other subroutines (See Figure). In the ARBRE subroutine of this calculation program, it is planned to build a molecule according to the given information for each atom as a building block of the constituent aminoacid residues. In the last version of the calculation program, this subroutine worked on the basis of information obtained by accessing the main routine MAINPGM, in which the solution of each individual task envisaged the reading of data on the amino acid residues that make up the molecule being studied. This was creating a heap of input data.

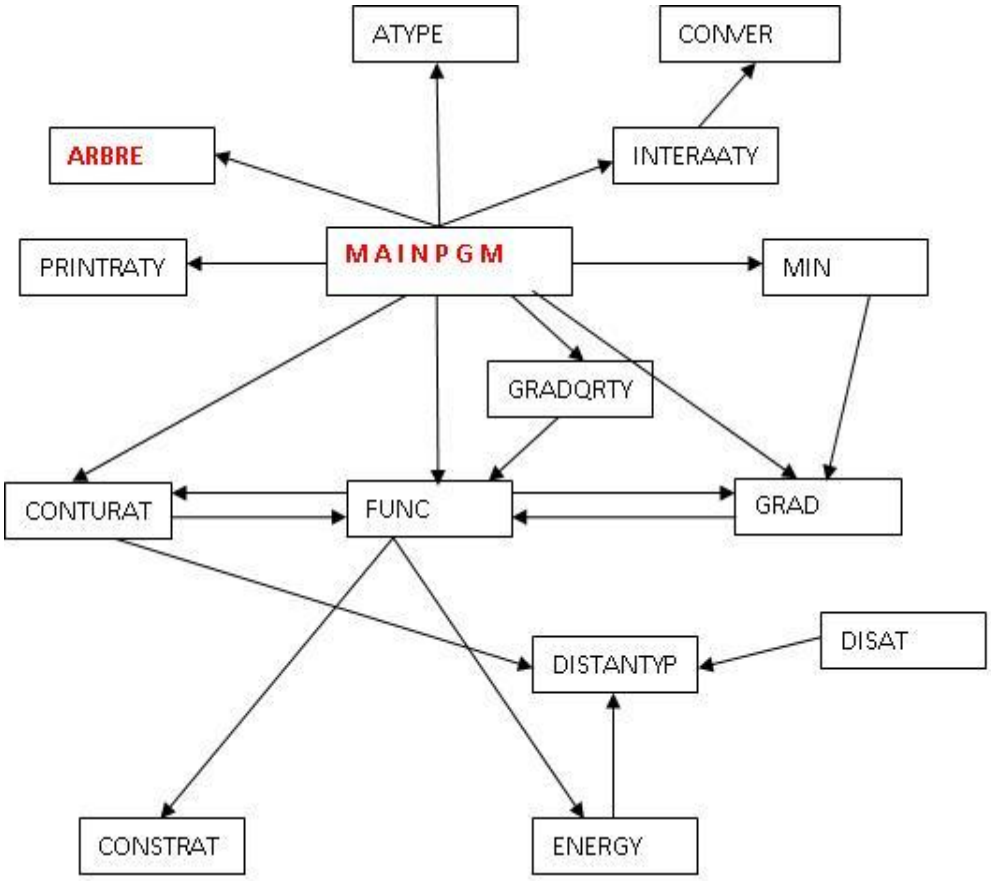

Figure. The diagram of the subroutines of program of calculation of molecular conformation

So, here the name of the amino acid residue of the substrate and the complete information about each atom of the given residue were introduced. In the modified version of the calculation program, a new algorithm for calling of the necessary information about an arbitrary amino acid residue from the direct access file was introduced into the ARBRE subroutine. The innovation allows to accurately assemble the desired amino acid sequence, avoiding the accumulation of input data during the construction of the molecule. Note that the created file is a database and can be placed on direct access devices, information media of both internal and external computer memory, which are may be disks, magnetic, or laser, flash drives, etc. The contents of the direct-access file can be scattered across different blocks of the disk. In contrast to the sequential access file, which is a sequence of entries, information from the direct access file can be read in any order. In this file each information block has its own identifier, the so-called key, which allows you to find the desired record. Note that the relative number that specifies this block among all blocks of this file is uniquely determined by the position within the file. For this reason, it is not necessary to view the file from the very beginning to access the information from the middle of the file. Therefore, the files bytes of which are read in any order are called direct access

files. The files consisting of single-byte entries on direct access devices are the most common way to organize of a file. For such files, the basic operations are reading or writing of the symbol to the current position. In the programming languages, the operators for the symbol-by-symbol transfer of data to or from a file are provided. The files in many OC Unix and MS-DOS file systems have the similar logical structure.

The created by us direct access file is an information database for naturally occurring amino acid residues, as well as for their various modifications, taking into account the chirality (L- and D-stereoisomers) and taking into account their location at the N- and C-terminal regions of the aminoacid sequences with different terminal atomic groups, for example, with charged groups COO^- or $NH3^+$, with amidated terminal parts, etc. This database includes the information on the chemical structure and geometry of each amino acid residue, namely, the attributes of atoms, the sequence of their location, the phases of atoms, the charges on atoms, the values of bond lengths and valence angles, the signs of rotation around certain chemical bonds, the parameters of torsional potentials. For each amino acid residue in the database its own identifier is assigned. It is a key that allows access to the information corresponding to certain atom. The new algorithm of the ARBRE subroutine of the calculation program provides for the use of an array of amino acid residue identifiers (FASEQ (I), I = 1, NPR) used to call information on an arbitrary amino acid residue from the direct access file, where NPR is the number of amino acid residues that are not repetitive in the molecule. The corresponding changes are entered in the main subroutine MAINPGM, where all program identifiers are described and data input is performed. It is here that the description of the extension of the FASEQ array is introduced, commands and formats for reading and printing of this array are introduced too. The corresponding changes of the parameters that provide for the use of a direct access file are introduced to program. Note that the repeated residues in the aminoacid sequence are accounted by using the array (SEQ (I), I = 1, NPRM), where NPRM is the number of all amino acid residues in the molecule. The input and description of the specified array is also carried out in the main subroutine. For example, in order to assemble the amino acid sequence of the gamma-melanotropin molecule H-Tyr1-Val2-Met3-Gly4-His5-Phe6-Arg7-Trp8-Asp9-Arg10-Phe11-Gly12-OH, consisting of twelve amino acid residues, it is necessary to call data of only ten residues, of Tyr, Val, Met, Gly, His, Phe, Arg, Trp, Asp, Gly, since the two residues, Phe and Arg, occur twice in this sequence. In spite of the fact that there are two glycines in this molecule, we do not consider it as a repeating residue, since the glycine located at the C-terminal part of the molecule is modified taking into account the negatively charged atomic group of COO^-. Note that in this case, the amino acid residue tyrosine, located at the N-terminal part of the molecule, is also modified, taking into account the presence of a positively charged atomic group NH_3^+. To call the desired aminoacid residue will be involved array FASEQ: non-recurring amino acid residues will be called using their keys from the direct access file. Note that the identification numbers assigned to the amino acid residues or any atomic groups should be known in advance to the user. Then, using the array SEQ, in which the repeated amino acid

residues take into account, the entire sequence of the molecule is collected. In our case, to collect the amino acid sequence, the SEQ array will be a sequence of numbers 1 2 3 4 5 6 7 8 9 7 6 10. As you can see, the numbers 6 and 7 in the array SEQ occur twice, since the Phe residue is present in the sixth and eleventh positions, and the residue Arg - in the seventh and tenth positions of this aminoacid sequence.

3. Conclusion

The algorithm of the call of the necessary information about amino acid residues of the peptide molecule from the direct access file was introduced to the appropriate subroutine of the program of semiempirical calculation of the biopolymer conformation. The innovation allows avoiding heaps of inputs used in the construction of molecules carry the unmistakable assembling of the desired amino acid sequence. Note that with the help of basic system calls it is possible to arbitrarily structure of the direct access files. For example, it is possible to enter data for arbitrary amino acid residues into the created database, and also make certain changes concerning both the sequence of available residues and the parameters describing each of them. Note that with the help of this program it is possible to calculate the conformational profiles of not only peptide compounds, but also nucleic acids, sugars, lipids and other molecular compounds. The direct access file provides for storing information data for building blocks of such compounds. This scheme of data storage and data using provides flexibility and versatility and facilitates the work of users of the calculation program.

References

1. Akhmedov N.A., Akverdieva G.A., Godjayev N.M., Popov E.M., (1986) Theoretical conformational analysis of brain peptides. β-melanosytostimulating hormone, *International Journal of Peptide and Protein Res.*, 27(2), 95-112.
2. Akhmedov N.A., Tagiyev Z.H., Hasanov E.M., Akverdieva G.A., (2003) Theoretical conformational analysis of the bovine adrenal medulla 12 residue peptide molecule, *J. Molecular Structure* , 646, 75-80.
3. Akverdieva G.A., (2016) Insights into bioactive conformation of dermorphin, *Azerbaijan Journal of Physics*, XXII(2), Section:En, 12-20.
4. Akverdieva G.A., Godjayev N.M., Nabiyev N.S, (2005) Conformational changes of kyotorphin and its biologically active analog, *Biophysics*, 50(1), 1-7. Translated from *Biofizika*, 50(1), 5-12. (in Russian)
5. Akverdieva G.A., Godjayev N.M., Akyuz S. , Akyuz T., Dogan N.E., (2010) Molecular Mechanics Studies of Antihypertensive Val-Tyr Dipeptide, *Asian Journal of Chemistry*, 22(1), 311-318.
6. Akverdieva G.A., Godjayev N.M., Akyuz S., (2009) Comparative conformational analysis of peptide T analogs, *Journal of Molecular Structure*, 917, 21-26.
7. Akverdieva G.A., Godjayev N.M., Akyuz S., (2002) Conformational dynamics of peptide T molecule, *Journal of Molecular Structure*, 609, 115-128.
8. Akverdieva G.A., (2016) Insights into spatial structure of deltorphins, *Journal of Qafqaz University*, 4(1), 13-20.

9. Aliyeva I.N., Godjayev N.M., (1985) Conformational dynamics of the ionophore lasalocid A (X-537A) at the complexation with catecholamines, *Studia Biophysica*, 108(2), 125-132.
10. Alieva I.N., Velieva L.I., Musayev M.A., Godjayev N.M., (2006) Conformational features of the Dippu-Ast8 neuropeptide from the cockroach Diploptera punctata, *Journal of Protein and Peptide Letters*, 13(10), 1007-1015.
11. Alieva I.N., Velieva L.I., Aliev D.I., Godjayev N.M., (2002) Spatial organization and conformational peculiarities of the gallatostatin family of neuropeptides, *Journal of Peptide Science,* 8(8), 385-397.
12. Davidon W.C., (1959) Variable Metric Method for Minimization, AEC Research and Development Report, ANL-5990.
13. Fletcher R., Powell M.J.D., (1963) A rapidly convergent descent method for minimization, *J. Computer*, 6, 163-168.
14. Godjayev N.M., (1986) Conformatianol aspects of interactions of the trypsyn with substrates and inhibitors, *Studia Biophysica*, 113(1-2), 89-94.
15. Godjayev N.M., Akyuz S., Akverdieva G.A., (1997) A molecular mechanics conformational study of peptide T, *Journal of Molecular Structure*, 403, 95-110.
16. Godjayev N.M., Maksumov I.S., Ismayilova L.I., (1983) Proqramma poluempiricheskoqo rascheta konphormatsiy molekulyarnich kompleksov, *Jurnal Strukturnoy Chimii*, 24, 147-148 (in Russian).
17. Hermans J.Jr., Ferro D., (1971) Representation of a protein molecule as a tree and application to modular computer programs which calculate and modify atomic coordinates, *Biopolymers*, 10, 1121-1138.
18. McGuire R.F., Momany F.A., Scheraga H.A., (1972) Energy parameters in polypeptides.V. Empirical hydrogen bond potential function based on molecular orbital calculations, *J. Phys. Chem.*, (76), 375-393.
19. Momany F.A., Carruthers L.M., McGuire R.F., Scheraga H.A., (1974) Intermolecular potentials from crystal data. III. Determination of empirical potentials and application to the packing configurations and lattice energies in crystals of hydrocarbons, carboxylic acids, amines, and amides, *J. Phys. Chem.*, 78, 1595-1620.
20. Momany F.A., McGuire R.F., Burgess A.W. and Sheraga H.A., (1975) Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids, *J. Phys. Chem*., 79, 2361-2381.