Education Program for Macromolecules Structure Examination

A. S. Ivanov,* A. B. Rumjantsev, V. S. Skvortsov, and A. I. Archakov

Laboratory "Computers in Biochemistry," Institute of Biomedical Chemistry,
Russian Academy of Medical Sciences, Pogodinskaya Street 10, Moscow, 119832, Russia

Received September 14, 1995[⊗]

A new program ONIX v.2.00 was specially designed with adaptation for the purposes of biochemist's education. ONIX works with proteins and nucleic acids, contained in PDB, and has the program interface based on molecular structure hierarchy. ONIX also has multimolecules and multiwindows environment, error control of PDB files, 3D presentation and space handling of models, calculation of solvent-accessible surface, macro-language, and saving of work results (working history). The program is oriented on widely accessible PC computers running Windows 3.1xx.

INTRODUCTION

Students' education in biochemistry and molecular biology requires the evident illustration of complicated structures of biological macromolecules, like proteins and nucleic acids. Such macromolecules consist of plenty of structural elements (subunits, sequences of residues, helixes, turns, prosthetic groups, etc.). The understanding of molecular function of biological macromolecules in many cases depends on the understanding of its structure. The most effective education can be achieved by using computer methods, enabling one to execute 3D visualization of macromolecules and their structural elements in the interactive mode and real time.

The program ONIX v.1.03,[1] enabling one to use 3D visualization and interactive analysis of protein structures from the Protein Data Bank[2] (PDB), was developed earlier. The given program is oriented on a widely accessible computer base (PC with processor i386sx or higher, 8Mb RAM, Windows 3.1xx, 256 colors video mode). A distinctive feature of this program is the principle of its construction, based on hierarchy of protein structural elements. In the present work the new version of this program ONIX v.2.00 for Windows 3.1xx/Win32 is described. This version was specially adapted for students' education in biochemistry and molecular biology.

SYSTEMS AND METHODS

The first version of ONIX (v. 1.03) operated only with protein structures from PDB. The program had the following properties: (1) multimolecules and multiwindows environment; (2) interactive error control of PDB files; (3) presentation, rotation, moving, and zooming macromolecules in 3D space; (4) copying, deleting, and transferring molecules between the windows; (5) direct access to any structural element; (6) disassembling molecule on any fragments and saving results; and (7) calculation of solvent-accessible surface. We have added in ONIX v.2.0 some new functions: integrated lists of structural elements of molecules; examination of protein and nucleic acid 3D structures; macrolanguage; and saving of work history. ONIX v.2.0 was written using a modern programming tool—MS Visual C/C++ v.2.1. and MS Foundation Classes v.2.5.[3] The program has been implemented on different types of PC and other platforms including MIPS and DEC under all versions of Windows including NT and 95.

ONIX requires the following configuration: PC 386/387 or higher; at least 4 Mb RAM; about 1 Mb of free disk space; SVGA 512 Kb graphics or higher; standard keyboard and mouse. It requires also the following software configuration: Windows 3.1 or higher; 640 × 480 or higher with 256 colors' graphics mode.

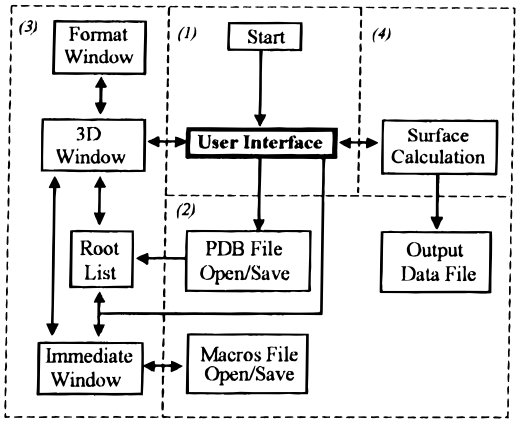

Figure 1. General flowchart of ONIX v.2.0. Program sections: (1) user interface; (2) Files read/write; (3) hierarchic visualization of macromolecule structure; and (4) surface calculation.

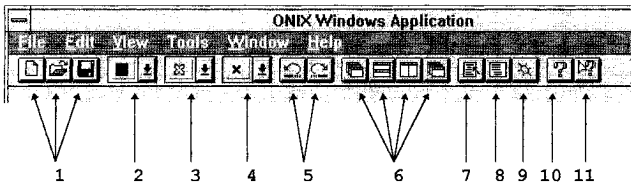

Figure 2. Main Menu of ONIX 2.0: (1) File operations commands (New, Open, Save); (2) palette command; (3) mode command; (4) check command; (5) undo/redo commands; (6) window commands (Cascade, Horizontal, Vertical, New); (7) immediate window; (8) New list window; (9) new View window; (10) Help menu; and (11) Help tool.

ALGORITHM AND IMPLEMENTATION

ONIX v.2.0 deals with protein and nucleic acid structures from PDB. The general flowchart of ONIX v.2.0 is shown in Figure 1. The program consists of four principal sections:

[⊗] Abstract published in *Advance ACS Abstracts,* April 1, 1996.

S0095-2338(95)00127-2 CCC: $12.00

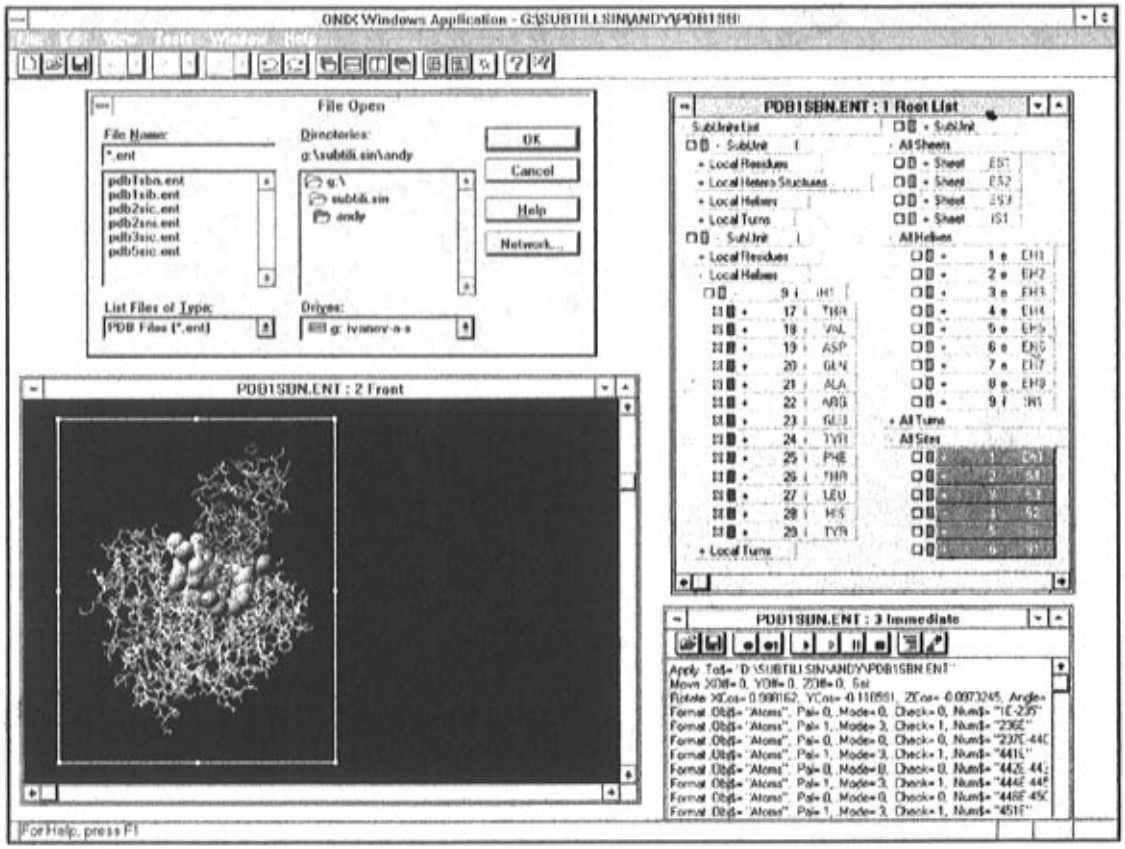

Figure 3. General view of working table of ONIX 2.0: *File Open window*—loading and saving files with PDB format; *Root List window*—hierarchic lists of structural elements and management of their properties; *3D Window*—visualization macromolecules in 3D space according their properties that were set in Root List window; and *Immediate window*—set of tools for macros reading, saving, creation and execution.

Table 1. Protein Structure Hierarchy, Which was Utilized for Hierarchic Organization of the Program ONIX

no.	level	types of molecular objects
5	molecular	molecule
4	quaternary	subunit (chain), heterostructure
3	tertiary	functional active domain (site)
2	secondary	helix, turn, sheet
1	primary	residues sequence, nonprotein substance
0	atomic	atoms

(1) User interface, realized in Main Menu of the program (Figure 2) and in commands and toolbars of certain number of working windows;

(2) Files Read/Write, includes PDB file Open/Save (reading files in PDB format with 3D coordinates of macromolecule as well as saving user results in file of the same format, see Figure 3, File Open window), Output Data File (saving of results of surface calculation), Macros File Open/Save (reading and writing files with macro commands, see Figure 3, Immediate window);

(3) Hierarchic visualization of macromolecule structure, consists of 3D Window (window of 3D visualization, see Figure 3), Root List (window with hierarchic lists of structural elements, see Figure 3, Root List window), Format Window (a set of special tools for changing color, type and visualization of group of elements, see Figures 6 and 7);

(4) Surface calculation for analysis of solvent accessibility of surface elements.

The program organization is based on hierarchy of protein structural elements (Table 1). Macromolecule consists of the following types of objects, described in the PDB file:

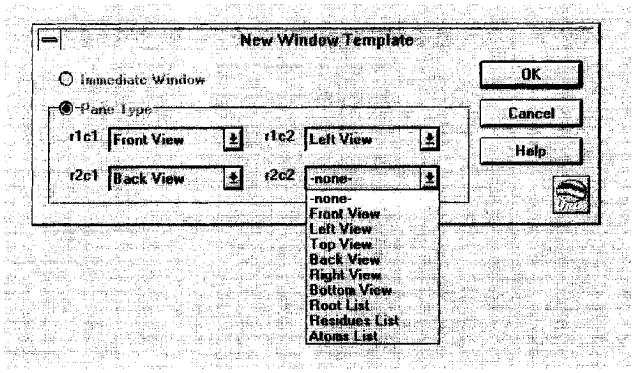

Figure 4. Creation of new 3D Window with parallel visualization of the same molecule. In this figure we show the creation of new 3D Window with four frames. Any frame can visualize a macromolecule from different points of view. For frame **r1c1** we choose *"Front View"*, for **r2c1**—*"Back View"*, for **r2c2**—*"Left View"*. For the last frame **r2c2** type of *View* is not defined yet. One can see the complete list of settings. From this picture one can clearly see that in any frame the user can set instead of the 3D View one of the following lists—Root List, Residues List, or Atoms List.

atoms, sequence of residues, heterostructure, SS-bond, secondary structure elements, functionally active site. These objects can be classified into five groups based on structural hierarchy of proteins: (1) atoms; (2) elements of primary structure (residues, heterostructures, SS-bonds), consisting from atoms; (3) elements of secondary structure (HELIX, TURN, SHEET, and SITE), consisting from residues; (4) elements of tertiary structure (SITE, combination of secondary structures); and (5) elements of quaternary structure (subunits, protein chains). The relationship between elements

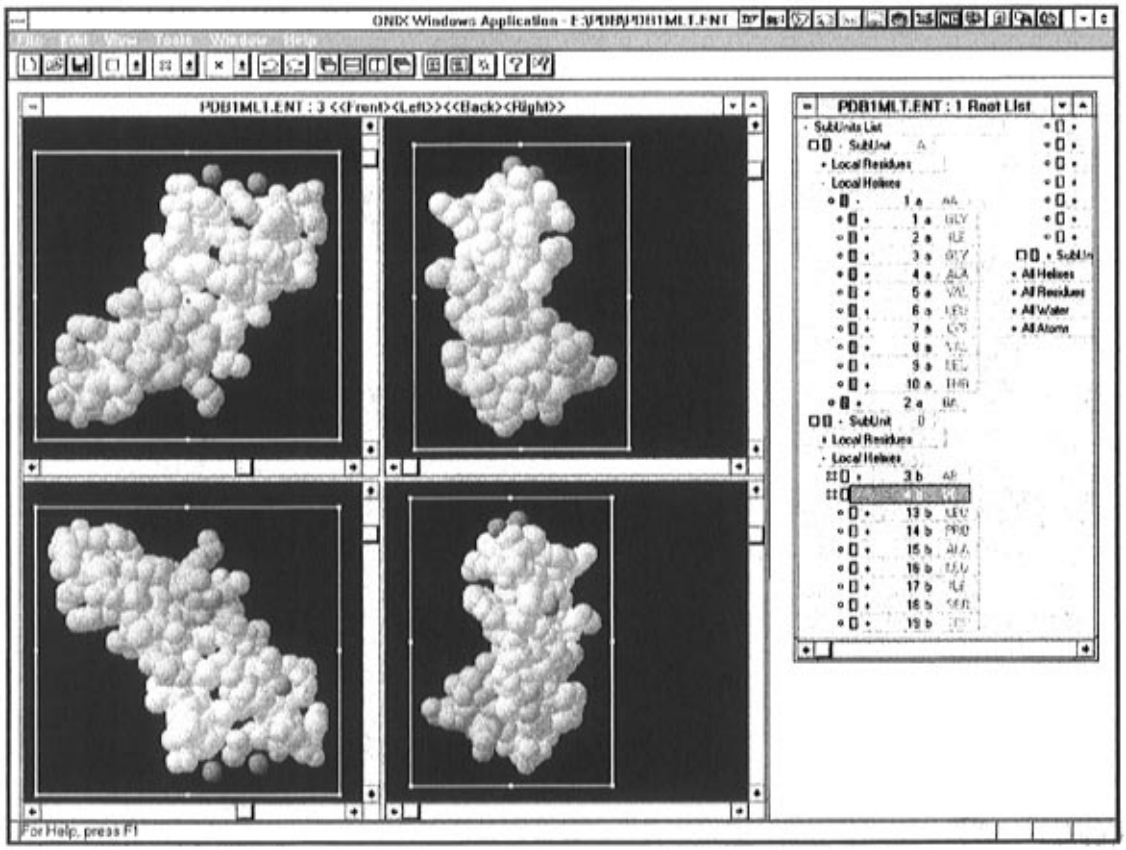

Figure 5. Example of parallel visualization of one molecule from four different view points. One molecule was visualized from <Front>, <Left>, Back>, and <Right> sides. All frames are interactively connected one to another, so any manipulation with a molecule in one of them leads to the changing of molecular images in all frames.

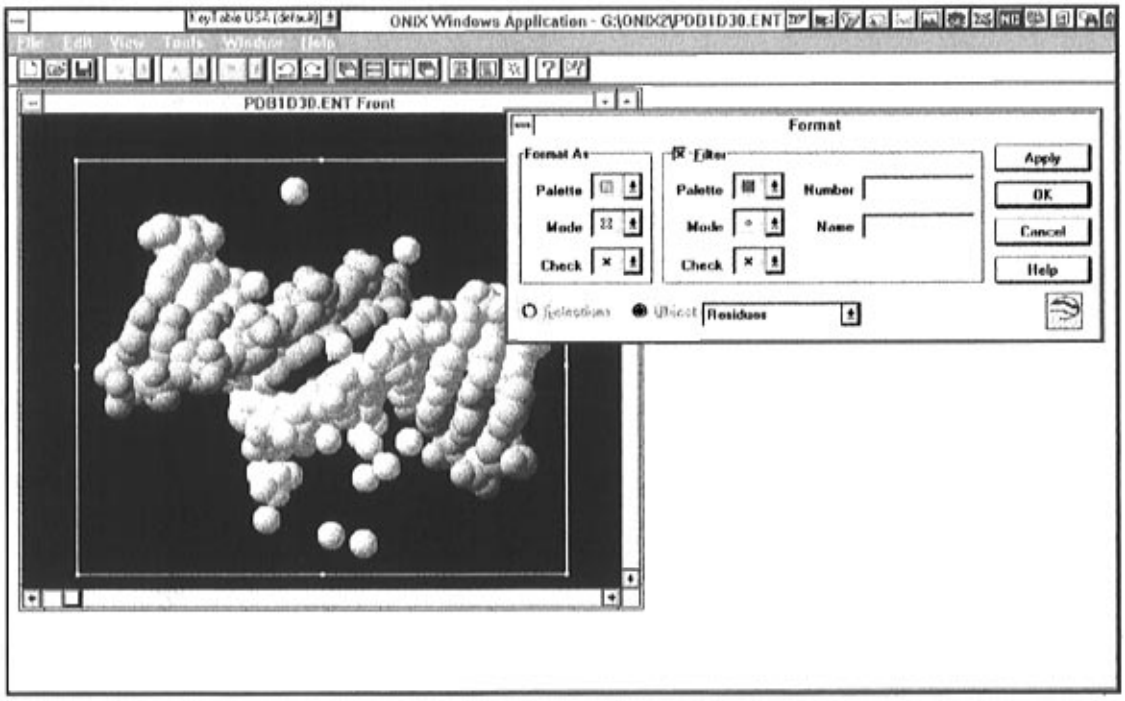

Figure 6. Window "FORMAT" with Filter tool. Using this tool an investigator can produce a group changing of elements' properties. For example, the current figure shows the following filter procedure: For elements type "Residues" with color as shown in the Palette box, in Mode VDW, and Check "ON"—their properties will change (Format as) in color as shown in the Palette box, Mode WIRE, and Check "ON". The user can test the results of this action by pressing the button Apply or run this procedure and close Format window by pressing the OK button.

from different groups can be described with structural hierarchy. The relationship "upwards" or "from bottom to top" is based on the fact that each atom is associated with exact amino acid residue. The last one is associated with an element of secondary structure, etc. Similarly the relationship "from above" or "from top to bottom" can be

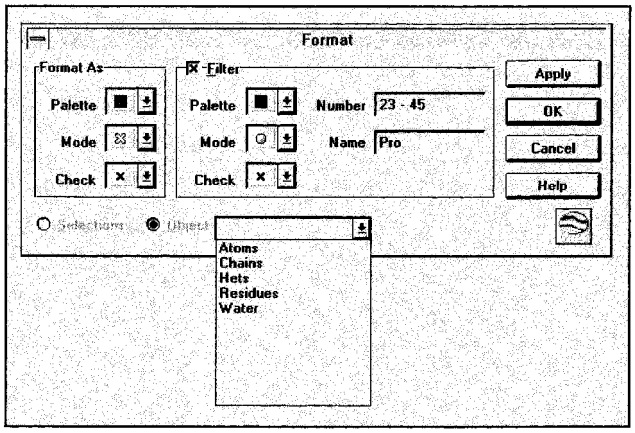

Figure 7. . Choosing of objects type in Format window. Different possibilities of the Format procedure are illustrated in Figure 7. The user can apply this tool to many types of objects: Atoms, Chains, Hets, Residues, and Water. Besides this, Format can be done only for specified elements. For example, in the current example for Residues with Name "Pro" and Numbers among 23 and 45.

recognized. Using the lists of considered above objects the user can visualize all diversity of their hierarchical relationships.

The working table of ONIX has main menu (Figure 2). Some of the most frequently used functions of the main menu (Open File, Write File, Color, Mode, Visualization, Undo, Redo, Help, etc.) are doubled by special buttons of the management panel.

A general view of a working table of ONIX 2.0 is shown in Figure 3. File Open interface was designed as a standard Windows file management interface. The program contains the macrolanguage for automatization of the working process (see Immediate window). It enables one to save the results and steps of actions (image formation, moving, scaling, etc.) as well as the execution of the commands Undo/Redo. A 3D visualization window serves for 3D presentation of molecules with zooming, rotation, moving, etc. This window can contain from one to four frames with different views of the same molecule or lists of structural elements (Figure 4). An example of parallel visualization of one molecule from four different view points is shown in Figure 5. In program ONIX 2.0 a special window "Format" was designed for visualization management by using a special tool—Filter (Figure 6).

DISCUSSION

ONIX 2.0 was specially designed for the purposes of students' education in biochemistry and molecular biology. This program may also be useful in science for the task of visualization of experimental data, for example surface structural elements such as protein antigenic determinants.[4] Program is oriented on widely accessible computer base (PC computers under Windows 3.1xx.). ONIX works with proteins and nucleic acids, contained in PDB. The program permits the visualization of 3D structure of a macromolecule and its structural elements in any combinations. Structural elements from atoms up to chains and subunits are systematized in the several hierarchical lists and can be displayed in bonds (wire) form, van der Waals spheres (VDW), and Richards' surface with various coloring ways. The original algorithm of 3D visualization enables the work with the image in real time. For example, creation of the image (1024 × 1280 pixels) of protein molecule (about 3500 atoms in VDW) on PC 486 DX4-100 takes about 3 s. We have 32 bit version ONIX 2.00pro designed for powerful PC Pentium running Windows 95 or Windows NT.

A free copy of ONIX v.1.03 is distributed by the authors. Readers can send their requests to Prof. Alexis S. Ivanov by e-mail: ivanov@ibmh.msk.su.

ACKNOWLEDGMENT

This work was supported in part by the Russian Ministry of Science and Technical Policy (Scientific State Program "Designing of new drugs by methods of chemical and biological synthesis" branch 04 "Computer designing of new drugs," grants 04.01.03 and 04.01.04).

REFERENCES AND NOTES

(1) Rumyantsev, A. B.; Skvortsov, V. S.; Lyulkin, Yu. A.; Ivanov, A. S.; Archakov, A. I. Refinement of ligand binding site structure in protein by computer interactive 3D visualization. *3rd IUBMB Conference "Molecular Recognition*; Singapore, 1995; p 144.
(2) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M.D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. *J. Mol. Biol.* **1977**, *112*, 535−542.
(3) Introducing Visual C++. Microsoft Visual C++, Development System for Windows and Windows NT, Version 2.1, 1995.
(4) Kolesanova, E. F.; Kozin, S. A.; Lemeshko, A. O.; Archakov, A. I. Epitope Mapping of Cytochrome-P450-2B4 by Peptide Scanning. *Biochem. Mol. Biol. Int.* **1994**, *32*, 465−473.

CI950127L