

Ц8412
К-238

СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

ДУБНА

3996/2-76



11/x-76

P11 - 9881

А.А.Карлов, А.Д.Польницев, Т.Ф.Смолякова, Г.Б.Щенкова

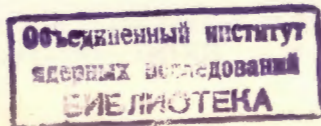
АНАЛИЗ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ БЕЛКА
В РЕЖИМЕ ДИАЛОГА С ИСПОЛЬЗОВАНИЕМ
ГРАФИЧЕСКОГО ДИСПЛЕЯ

1976

P11 - 9881

А.А.Карлов, А.Д.Польнцев, Т.Ф.Смолякова, Г.Б.Щенкова

АНАЛИЗ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ БЕЛКА
В РЕЖИМЕ ДИАЛОГА С ИСПОЛЬЗОВАНИЕМ
ГРАФИЧЕСКОГО ДИСПЛЕЯ



Карлов А.А., Польшцев А.Д., Смолякова Т.Ф.,
Шенкова Г.Б.

P11 - 9881

Анализ пространственной структуры белка в режиме диалога
с использованием графического дисплея

В работе рассматриваются возможности программ, которые разработаны совместно ОИЯИ и Институтом молекулярной биологии АН СССР и предназначены для представления на экране дисплея и анализа структур молекул сложных соединений. Пользователю предоставлена возможность осуществлять построение на экране дисплея изображений молекулы и отдельных ее фрагментов, вращение полученного изображения в пространстве вокруг произвольной оси, получение стереопроекций и т.п.

Весь комплекс программ написан на языке ФОРТРАН.

Работа выполнена в Лаборатории вычислительной техники
и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований
Дубна 1976

Karlov A.A. et al.

P11 - 9881

The Interactive Analysis of Space Protein
Structure with the Use of Graphical Display

Some programme features developed by JINR in cooperation with the Institute for Molecular Biology of the USSR Academy of Sciences and intended for presentation on the display screen and for the analysis of complex molecular structures are considered. A programme user is provided with the opportunity of constructing molecular images and its segments on display screen, rotation of the image obtained in space round an arbitrary axis and constructing stereoviews.

The programme complex is written in FORTRAN for CDC-1604A computer.

Communication of the Joint Institute for Nuclear Research
Dubna 1976

Элементарные носители биологических функций - молекулы белков и нуклеиновых кислот. Способность выполнять соответствующие биологические функции обусловлена сложной пространственной и химической структурой молекул этих соединений. Насколько сложна белковая молекула, можно судить из следующего факта. Белки представляют собой линейные полимеры, состоящие из аминокислот, связанных так называемыми пептидными связями. В каждой молекуле белка обычно содержится 100-400 аминокислотных остатков. Молекулярные веса белков колеблются от нескольких тысяч до десятков и сотен тысяч. Пример общего вида модели молекулы показан на *рис. 1.*^{1/1}

Полную модель молекулы обычно строят из молекулярных моделей Кендрию^{2/}. Однако такая модель не всегда удобна для изучения молекулы, так как по ней трудно различим скелет пептидной цепи, затруднительно выделение для изучения отдельного фрагмента молекулы, определение его пространственной структуры, а также его положения относительно других фрагментов. При исследовании структуры молекулы часто бывает необходимо повернуть модель или ее часть на некоторый угол вокруг какой-либо оси, что с моделями Кендрию в большинстве случаев довольно затруднительно.

Появление в последние годы графических диалоговых систем предоставило в руки исследователей гибкий инструмент для изображения и анализа сложных пространственных структур в режиме активного взаимодействия человека с ЭВМ.

В данной работе рассматривается дисплейная методика, разработанная совместно Институтом молекулярной биологии АН СССР и Лабораторией вычислительной техники и автоматизации ОИЯИ и предназначенная для

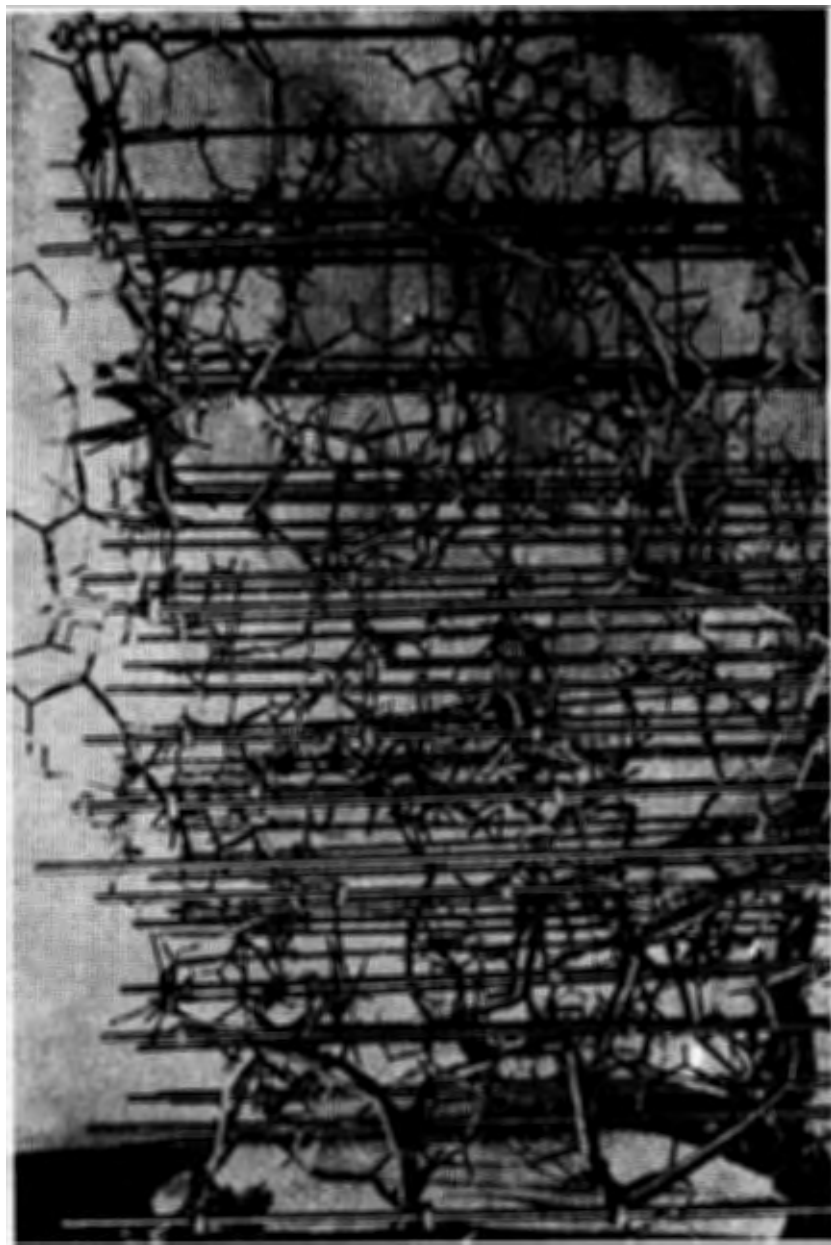


Рис. 1.

представления на экране и анализа структур молекул сложных соединений. Эта методика позволяет в диалоговом режиме осуществлять построение на экране дисплея изображений как полной молекулы, так и отдельных ее фрагментов, вращение полученного изображения в пространстве вокруг произвольной оси, последовательное прослеживание заданного фрагмента относительно всей молекулы с помощью специального маркера, а также построение стереопроекций. Сочетание вычислительных мощностей ЭВМ для проведения расчетов и построения сложных изображений с научной интуицией и творческими способностями человека для оперативного анализа получаемых результатов позволяет не только существенно ускорить по сравнению с традиционными методами процесс исследований, но и повысить их эффективность за счет появления принципиально новых возможностей анализа.

Весь комплекс программ, реализующий данную методику, написан на языке ФОРТРАН и используется в настоящее время для изучения структуры белка пепсина, содержащего 327 аминокислотных остатков. Работа проводилась на ЭВМ CDC-1604А ОИЯИ, имеющей в составе внешних устройств графический дисплей ^{/3/} с развитым математическим обеспечением ^{/4/}. Специфика рассматриваемой задачи потребовала также модификации ряда подпрограмм дисплейной библиотеки с целью расширения их возможностей.

На рис. 2а,б изображена общая блок-схема программы GALJA. Входными данными программы являются пространственные координаты атомов молекулы, подготовленные на перфокартах. Так как координаты атомов заданы в системе координат, связанной с пространственной ориентацией кристалла белка, производится преобразование аффинной системы координат в декартову. Кроме того, для получения изображения молекулы в центре экрана выполняется также линейное преобразование координат. По запросу от ЭВМ с пультовой пишущей машинки /ПМ/ вводятся значения переменных, определяющих режимы работы программы /здесь, как и во всех других случаях ввода с ПМ, предусмотрена

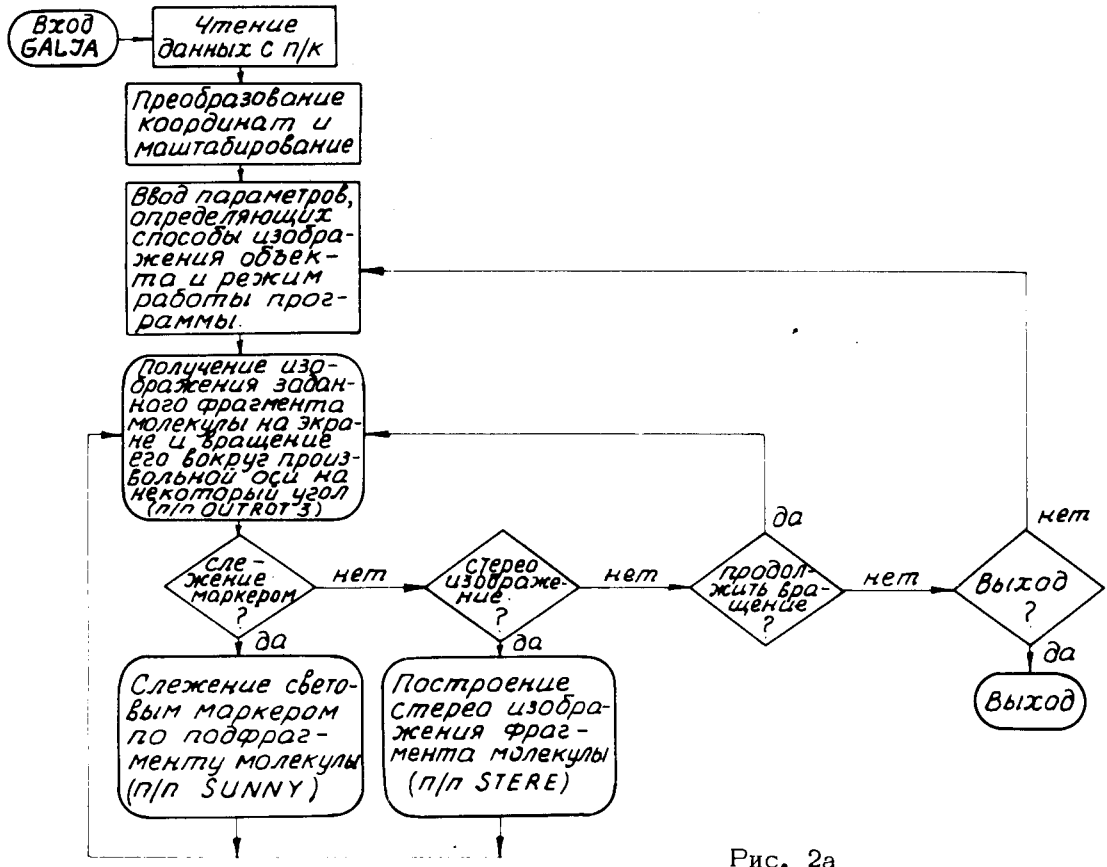


Рис. 2а

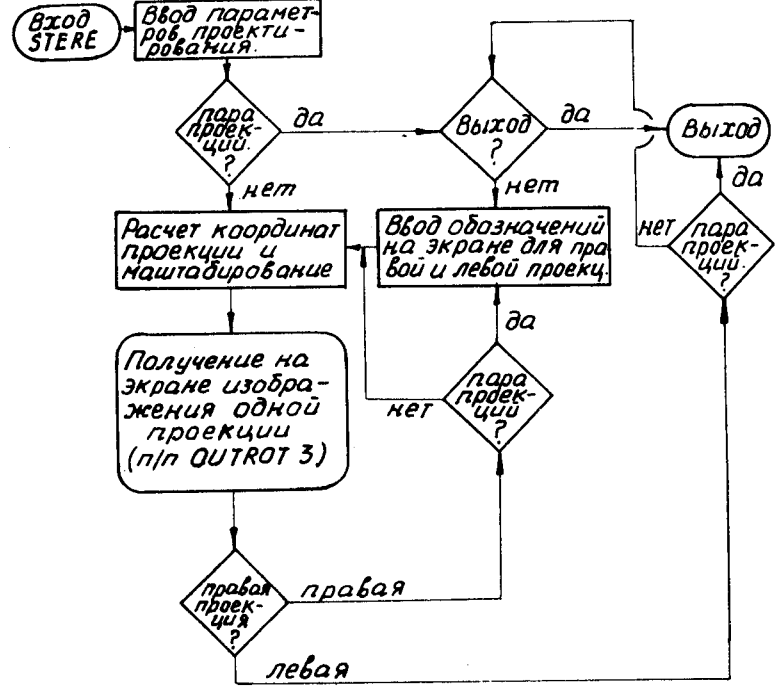
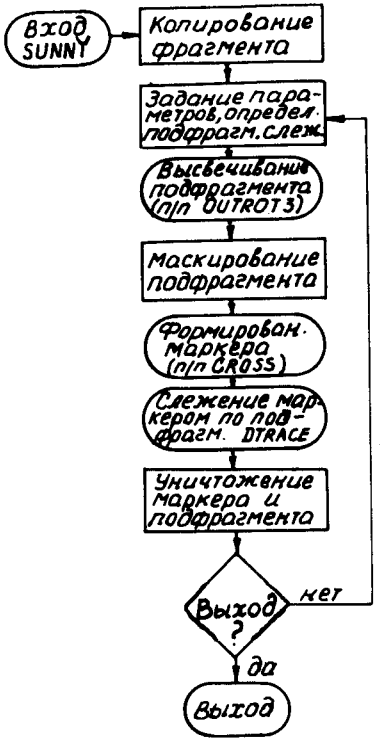


Рис. 2б.

проверка вводимых переменных на корректность принимаемых ими значений - при неправильном вводе запрос повторяется/. Переменные K1 и KN определяют номера начального и конечного атомов для выводимого на экран фрагмента молекулы, переменная STEP - шаг, с которым высвечиваются атомы фрагмента /при STEP=1 высвечиваются все атомы фрагмента, при STEP=2 - каждый второй и т.д./. Значения этих переменных изображаются на экране /рис. 3/ вместе с переменными, которые указывают число высвеченных атомов фрагмента (NCA), а также общее число атомов в молекуле (NC). Задание других переменных с ПМ определяет способ изображения молекулы или ее фрагмента. Так, например, можно представить на экране или отдельные атомы в виде точек, или соединить их отрезками прямых линий. Можно также определить способ выдачи изображения на экран: получить сразу полное изображение или, задавая временной интервал, через который на экране должен появиться очередной атом или отрезок, соединяющий атомы, обеспечить последовательное рисование молекулы на экране. Пример изображения молекулы и ее фрагмента представлен на рис. 3 и 4.

После завершения ввода с ПМ производится переход на модифицированную библиотечную подпрограмму OUTROT3 /4/, которая предназначена для построения на экране дисплея трехмерного изображения по топологии, задаваемой пользователем и выполнения однократных и непрерывных поворотов полученного изображения вокруг произвольной оси. Этой подпрограммой предусмотрен вывод на экран световых клавиш DTX, DPY, DFZ, а также клавиш R, X, Y, Z, RES, OUT. Выбор световым карандашом /СК/ одной из первых трех клавиш вызывает со стороны ЭВМ выдачу запроса на ввод с ПМ значения угла для однократного поворота вокруг оси X, Y или Z, соответственно. В исходном состоянии $DTX = DPY = DFZ = 0$. Однократный поворот выполняется после завершения ввода значений углов с помощью световой клавиши R. Световые клавиши X, Y, Z служат для плавного /через 1 градус/ вращения вокруг соответствующей оси. Повторный выбор клавиши X, Y, Z пре-

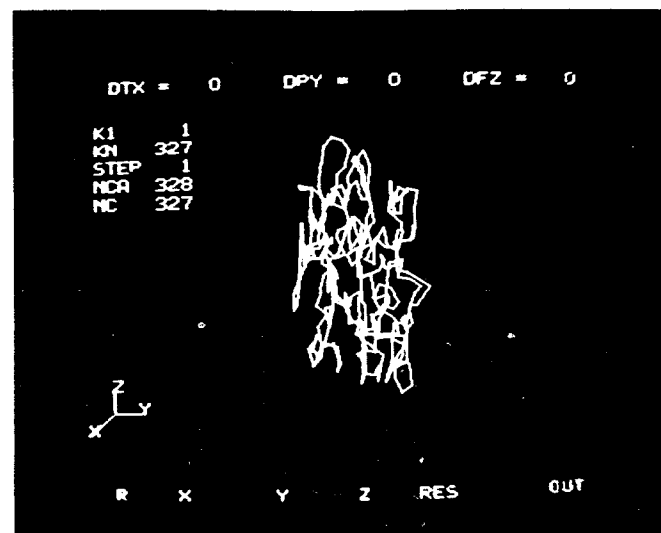


Рис. 3. Пример изображения молекулы пепсина.

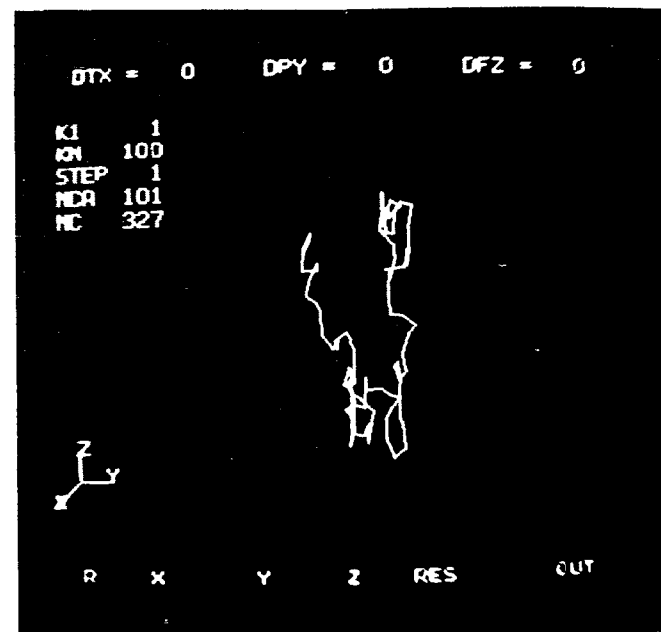


Рис. 4. Пример изображения фрагмента молекулы.

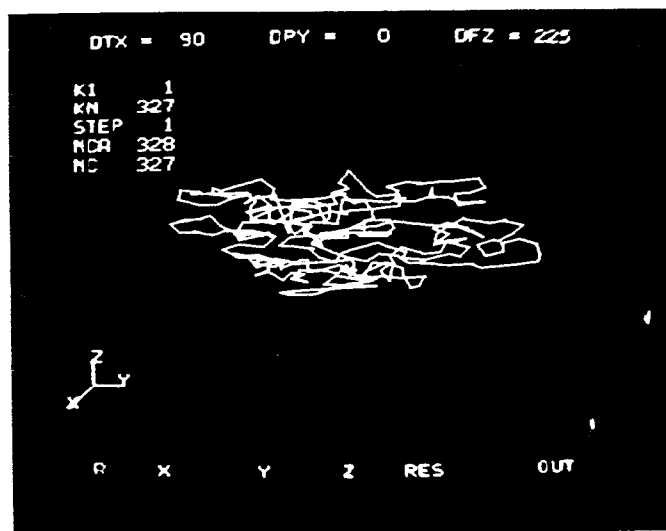
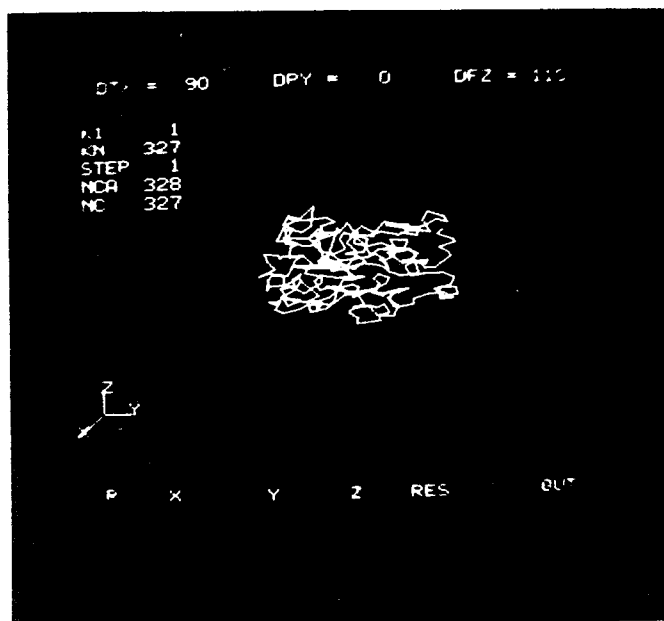


Рис. 5. Результат поворота молекулы вокруг оси X(90°) и Z(110° и 225°).

кращает вращение вокруг данной оси. Выбором световой клавиши OUT обеспечивается возврат на начало подпрограммы OUTROT3. Световая клавиша RES служит для выхода из подпрограммы OUTROT3.

После выхода из подпрограммы OUTROT3 со стороны ЭВМ на ПМ формируется запрос на ввод параметров, определяющих характер дальнейшей работы.

На этом этапе имеется возможность:

- осуществить последовательное прослеживание заданного фрагмента молекулы относительно всего изображения с помощью специального светового маркера /переход на подпрограмму SUNNY /;
- получить стереопроекции для изображения, которое было сформировано ранее подпрограммой OUTROT3 /переход на подпрограмму STERE /;
- переопределить значения переменных, которые задают режимы работы программы;
- завершить работу с программой.

Последовательное прослеживание заданного участка молекулы на фоне всего изображения состоит в том, что подпрограммой SUNNY реализована возможность наблюдать перемещение по этому участку специального маркера. Границы участка /фрагмента/ задаются при входе в подпрограмму SUNNY с клавиатуры ПМ и высвечиваются на экране.

С точки зрения техники программирования возможность прослеживания заданного фрагмента реализована следующим образом. Среди подпрограмм дисплейной библиотеки имеется подпрограмма, которая позволяет поместить произвольный графический объект /в данном случае маркер/ в указанную точку на экране. Последовательно выбирая /с помощью другой дисплейной подпрограммы/ координаты точек рассматриваемого фрагмента и используя их для указания положения маркера, можно получить эффект перемещения маркера вдоль исследуемого участка молекулы. Скорость движения маркера можно изменять, задавая с ПМ значение соответствующей переменной. После завершения процедуры слежения можно перейти к заданию нового фрагмента или выйти из подпрограммы.

Получение стереопрооекций обеспечивается подпрограммой STERE.

Эффект восприятия наблюдателем пространственной структуры молекулы на плоском экране графического дисплея может быть получен с помощью пары двумерных стереопрооекций. При этом необходимо, чтобы в тот момент, когда человек рассматривает изображение, каждый глаз наблюдателя видел только одну соответствующую проекцию. Проекция получается в результате проектирования атомов молекулы на выбранную плоскость относительно двух точек наблюдения /левого и правого глаза/. Вопросы, связанные со стереовидением и получением стереофотографий, более подробно изложены в /5,6/.

Принятая в данном случае процедура проектирования отдельной точки $A(x, y, z)$ на плоскость P относительно точки наблюдения $C(c_x, c_y, c_z)$ видна из рис. 6. Плоскость проектирования P задается уравнением $x=p$, где p — константа. Координаты точки $B(b_x, b_y, b_z)$, которая является проекцией точки A на плоскость P относительно точки C , определяются из подобия треугольников по формулам:

$$b_x = p$$

$$b_y = y - (p - x)(y - c_y) / (c_x - x)$$

$$b_z = z(c_x - p) / (c_x - x).$$

При входе в подпрограмму STERE на ПМ формируется запрос на ввод переменных, определяющих значение величин c_x , c_y , p . Введенные значения проверяются на корректность с точки зрения получения стереоэффекта /например, c_x должно быть достаточно большим, чтобы точка наблюдения была вне молекулы: принято также, чтобы плоскость проектирования располагалась между молекулой и точкой наблюдения и т.п./.

В подпрограмме GALJA предусмотрена возможность получения на экране как одновременно обеих проекций, так и изображения отдельной /левой или правой/ проекции /рис. 7а, б/.

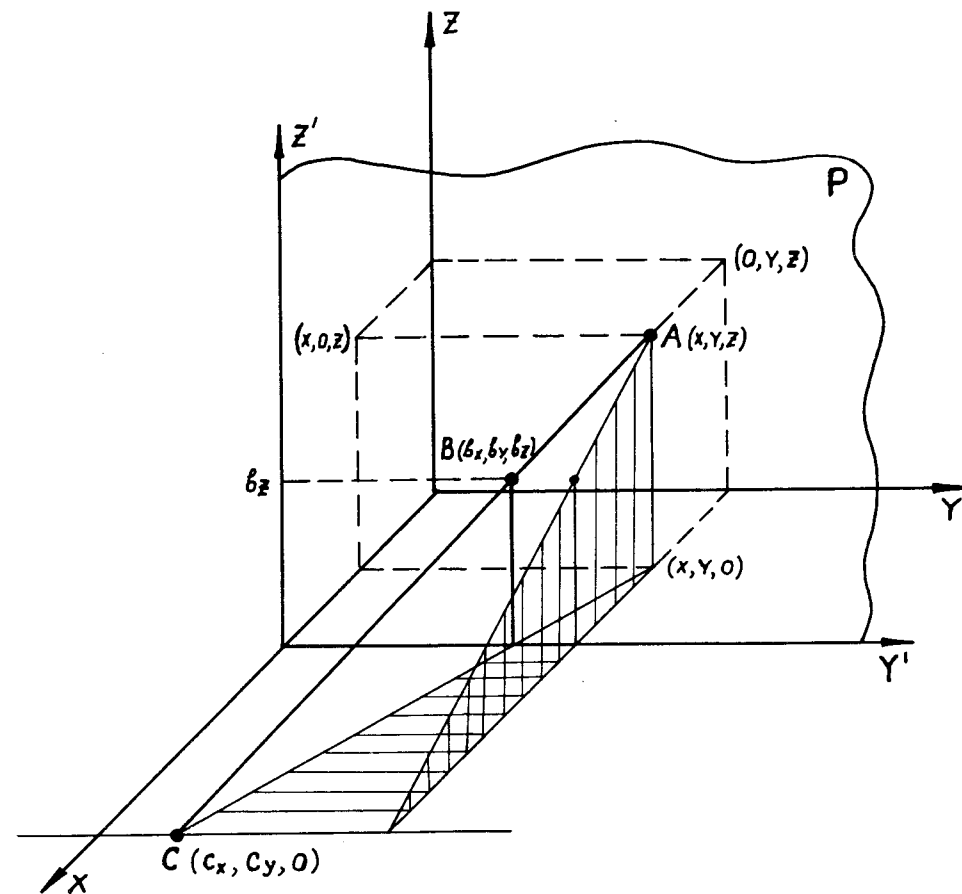
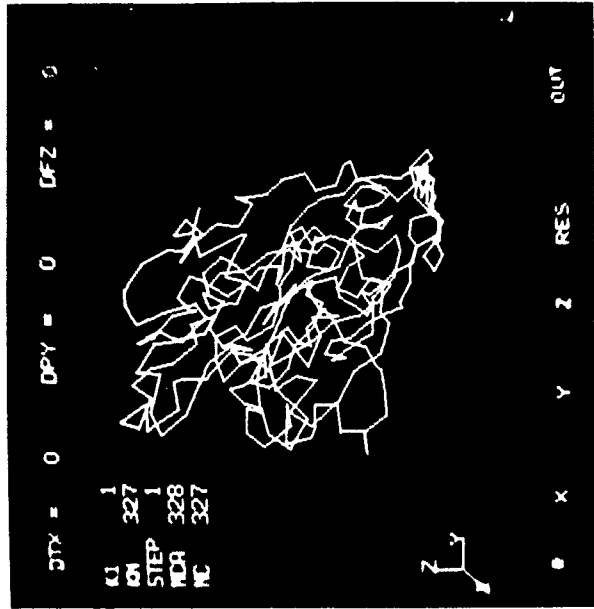


Рис. 6

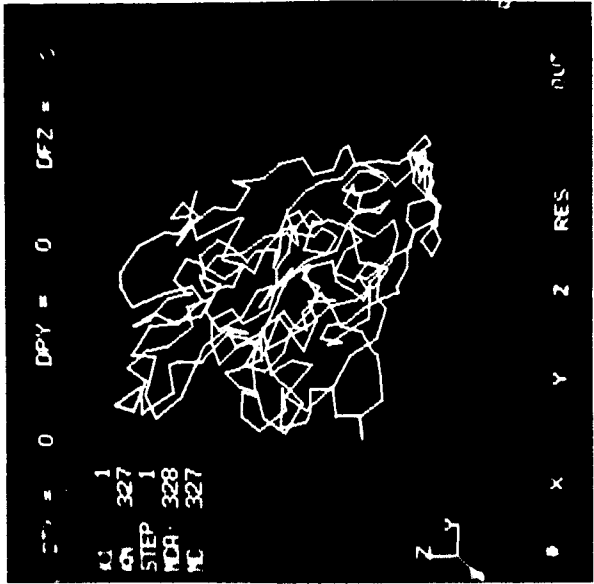
Возврат на начало программы GALJA позволяет установить новое значение переменных, определяющих режимы работы программы, и провести повторный анализ.

Для получения координат второй проекции в вышеприведенных формулах производится замена c_y на $-c_y$.

Опыт использования рассмотренной методики при анализе структуры пепсина показал, что она является достаточно универсальной, позволяет более эффективно



а/



б/

Рис. 7. Изображение левой /а/ и правой /б/ стереопроекций молекулы.

по сравнению с традиционными методами анализа выделять особенности структуры, получать графический материал в виде фотографий, слайдов, кинофильма. Режим диалога, при котором от человека требуются лишь указания о режимах работы программы и значения отдельных величин в виде сообщений, вводимых через клавиатуру ПМ, позволяет легко привлечь к работе на ЭВМ, специалистов в области структурного анализа, которые не имеют опыта программирования. Таким образом, круг пользователей может быть значительно расширен.

С помощью дисплея удобно наблюдать особенности пространственной структуры молекулы. Так, на рис. 5а,б легко заметить асимметрию молекулы пепсина. Кроме того, графический дисплей является хорошим инструментом для сопоставления гомологичных белков. При повороте молекулы пепсина относительно исходного положения вокруг осей x, y, z на $90^\circ, 90^\circ$ и 225° , соответственно, удалось установить большое сходство полученного изображения /рис. 7а,б/ с изображением кислой протеазы из бактерий *Physopus chinensis*.

В заключение авторы выражают благодарность члену-корреспонденту АН СССР, профессору Н.Н.Говоруну и профессору Н.С.Андреевой за постановку задачи, постоянный интерес и помощь в работе, а также Н.Н.Втюрину за полезные обсуждения.

Литература

1. Н.С.Андреева, А.А.Федоров, А.Е.Гущина, Н.Е.Шуцкевер, Р.Р.Рискулов, Т.В.Вольнова. ДАН, 228, №2, стр. 224-227, 1976.
2. J.C.Kendrew, R.E.Dickerson, B.E.Stranberg, R.G.Hart, D.R.Davies, D.C.Phillips, V.C.Shore. Nature, 185, 442 /1960/.
3. А.И.Ефимова, Г.И.Забиякин, А.А.Карлов, А.П.Кретов, И.Н.Кухтина, Ф.В.Левчановский, В.И.Приходько, В.Р.Трубников, Э.В.Шарапова. ПТЭ, 4, 91-96, 1971.
4. А.А.Карлов, А.В.Кавченко, А.Д.Полынцев, Т.Ф.Смолякова. ОИЯИ, Б1-11-6439, Дубна, 1972.
5. Б.У.Барщевский, Б.Т.Иванов. Объемная фотография, М., Искусство, 1970.

6. *Е.Г.Шаер. Применение фотографии в медицине. М., Медицина, 1964.*

*Рукопись поступила в издательский отдел
16 июня 1976 года.*