

сообщения  
объединенного  
института  
ядерных  
исследований  
дубна

35.39 / 2-80

28/7-80

P11-80-229

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

ПОСТРОЕНИЕ ПРОСТРАНСТВЕННОЙ МОДЕЛИ  
МОЛЕКУЛЫ БЕЛКА В РЕЖИМЕ ДИАЛОГА  
НА ОСНОВЕ КАРТ ЭЛЕКТРОННОЙ ПЛОТНОСТИ

1980

## 1. ВВЕДЕНИЕ

Построение моделей молекул сложных соединений - тяжелая и трудоемкая

работа, требующая нескольких месяцев труда высококвалифицированных специалистов. Обычно построение модели происходит при помощи метода оптического сравнения, предложенного Ричардсом<sup>/1/</sup>. Модель строится на основе трехмерной карты электронной плотности с помощью полупрозрачного зеркала. При этом модель оптически совмещается с картой таким образом, что построение ее происходит как бы внутри карты электронной плотности. Однако само построение модели осуществляется вручную, что требует значительных затрат времени; построенная таким образом модель достаточно громоздка и сложна в обращении; возникают большие трудности при снятии пространственных координат атомов, используются двумерные сечения карты электронной плотности только по одной из координатных осей, что создает определенные неудобства при уточнении пространственного положения отдельных частей молекулы.

Данная работа проводилась в направлении автоматизации процесса интерпретации карты и построения модели молекулы. С этой целью использовались возможности современных диалоговых систем, обеспечивающих представление на экране дисплея графической информации и ее редактирование.

Работа выполнена на языке ФОРТРАН в Лаборатории вычислительной техники и автоматизации ОИЯИ с использованием удаленной дисплейной станции /УДС/<sup>/2/</sup> ЭВМ БЭСМ-6 применительно к задаче построения модели белковой молекулы пепсина<sup>/3/</sup>.

## 2. ПОДГОТОВКА ВХОДНЫХ ДАННЫХ

Для расчета предварительных данных о пространственной структуре скелета молекулы /предварительный скелет/ был

создан специальный комплекс программ<sup>/4/</sup>. Данные о предварительном скелете и о сечениях по одной из координатных осей трехмерной карты электронной плотности рассчитываются в пакетном режиме на ЭВМ CDC-6500. Карта электронной плотности и предварительный скелет молекулы задаются положительными действительными числами в узлах равномерных трехмерных сеток и записываются в форме матриц на внешних носителях. Простой расчет дает возможность вычислить произвольное двумерное сечение по любой из выбранных координатных осей. Программа построения

изолиний <sup>15/</sup> позволяет получить изображение заданного сечения на графическом дисплее, а программа для построения изображения графа дает возможность вывести рисунок предварительного скелета на экран дисплея.

В связи с тем, что дальнейшая обработка данных, рассчитанных на CDC-6500, происходит на ЭВМ БЭСМ-6, были написаны программы перезаписи числовой информации с одной машины на другую.

**3. ОРГАНИЗАЦИЯ РАБОТЫ** Математическое обеспечение /МО/ УДС <sup>16/</sup> позволяет создать удобный диалоговый аппарат <sup>17/</sup> для решения задачи моделирования молекулы. Пользователь может быть абсолютно незнаком ни с МО УДС, ни с МО задачи моделирования. Ему достаточно знать небольшое количество простых приказов и реакцию ЭВМ на выполнение этих приказов. Например, приказ:  $x, i$  - высветит на экране дисплея заданный фрагмент  $i$ -ой плоскости сечения карты электронной плотности по оси X, а приказ /S, Y2, 30. - позволит присвоить новое значение переменной, определяющей одну из границ фрагмента по оси Y. Все приказы задаются с пульта дисплея и высвечиваются на экране. В случае неверного приказа или тогда, когда его невозможно выполнить по каким-либо причинам, на экране появляется соответствующая диагностика. Результат выполнения приказа также высвечивается на экране.

**4. АЛГОРИТМ ПОСТРОЕНИЯ МОДЕЛИ МОЛЕКУЛЫ** Работа по построению модели молекулы начинается с выбора координатной оси, вдоль которой будет осуществляться просмотр структуры молекулы /рис. 1 и 2/, и просмотра различных плоскостей координатных сечений на экране дисплея. Это позволит быстро найти плоскости сечений той части молекулы, которая является в данный момент наиболее интересной для исследователя.

Графический дисплей предоставляет пользователю большие возможности для выбора начального уровня и шага, с которым изолинии карты электронной плотности будут получены на экране. В режиме диалога осуществляется задание различных величин начального уровня и шага /рис. 3 и 4/, что позволит в дальнейшем эффективнее определить особенности структуры. Неправильный подбор этих величин может привести к ошибкам интерпретации карты электронной плотности.

Часто бывает неудобно рассматривать целиком плоскости сечений. Поэтому в соответствии с выбранной для исследования частью молекулы задаются границы фрагмента карты электронной

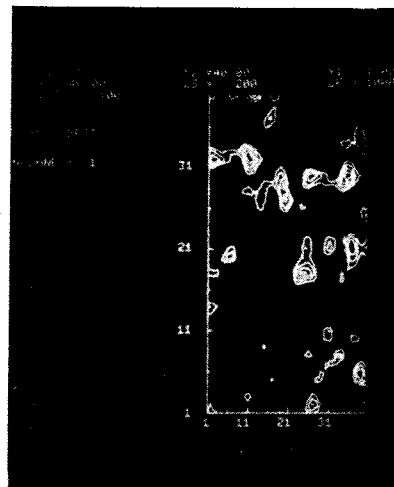


Рис.1. Первое сечение по оси X /фрагмент/.

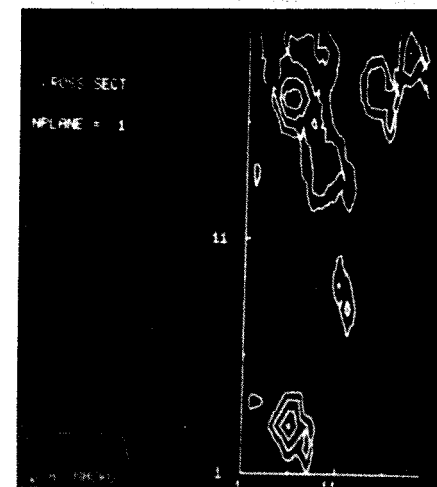


Рис.2. Фрагмент первого сечения по оси Y.

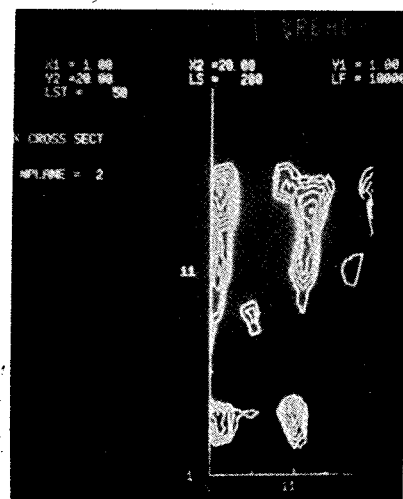


Рис.3. Фрагмент второго сечения по оси X. Изолинии проведены начиная с 200 уровня до 10000 с шагом 50.

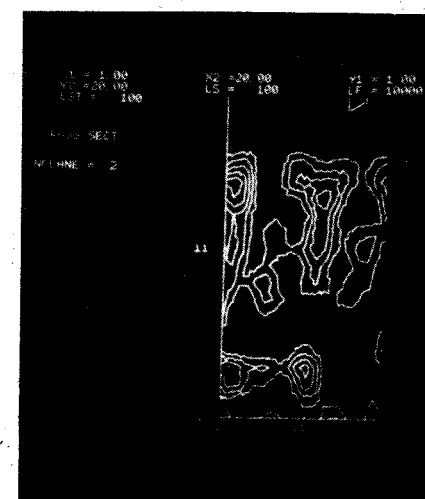


Рис.4. Фрагмент второго сечения по оси X. Изолинии проведены начиная с 100 уровня до 1000 с шагом 100.

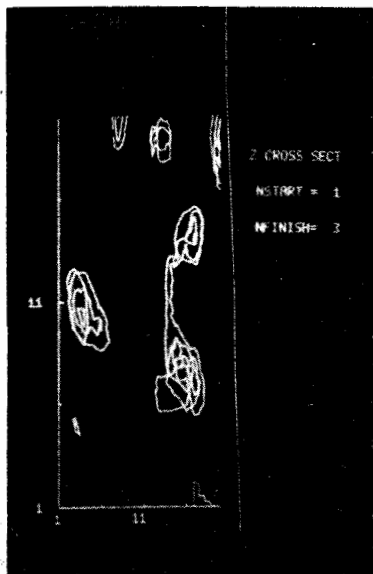
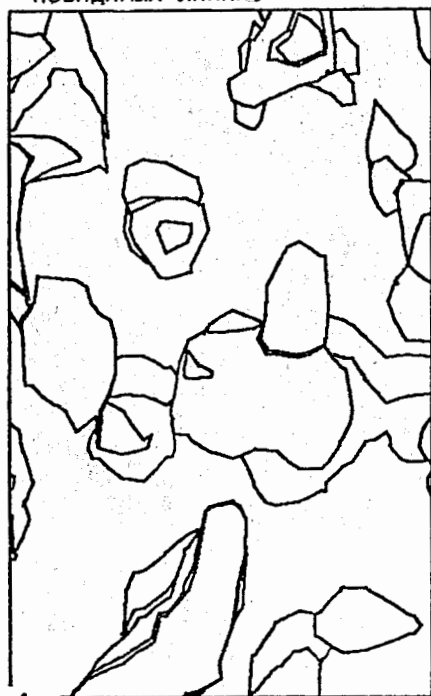


Рис. 5. Последовательное наложение трех фрагментов по оси Z без стирания невидимых линий.



плотности, с которым будет проводиться дальнейшая работа. Заданный фрагмент выделяется из карты и изображается на экране. При этом полезно бывает получить изображения линий только одного уровня плотности, а при наложении различных плоскостей сечений воспользоваться программой стирания невидимых линий /рис. 5 и 6/. Затем изображение этого фрагмента и изображение соответствующей части скелета молекулы совмещаются на экране.

В связи с тем, что предварительный скелет молекулы вследствие неоднозначности интерпретации карты в некоторых частях может не удовлетворять исследователя, возникает потребность редактирования рассчитанного ранее скелета. Цель редактирования состоит в том, чтобы добиться наибольшего соответствия предварительного скелета карте электронной плотности, на основе которой он был вычислен. В результате уточняются пространственное положение главной цепи молекулы, ее боковых ветвей, определяются особенности структуры. Более подробно алгоритм редактирования описан ниже в разделе 5. Если исследователь считает, что достигнуто желаемое соответствие, то он может выбранную точку рассматриваемой части пространства объявить  $C\alpha$ -атомом /см.

Рис. 6. Последовательное наложение пяти фрагментов со стиранием невидимых линий.

раздел 6/. Пространственная структура главной цепи молекулы определяется пространственными координатами  $C\alpha$ -атомов.

Имеется также возможность работы с поименованными скелетными объектами.

После того, как работа с выбранной частью модели молекулы завершена, информация о соответствующем отредактированном куске скелета записывается в виде файлов на диск для дальнейшего использования /анализа, редактирования и т.п./. При этом вся работа организована так, что в любой момент времени она может быть прервана и продолжена через произвольный интервал времени. Все промежуточные результаты будут сохранены.

Схема алгоритма автоматического моделирования дана на рис. 7.

В работе используются два графических дисплея, имеющих в составе УДС. Один из них, регенеративного типа, полезен для оперативного редактирования изображения. Другой, запоминающего типа, в отличие от первого, обладает практически неограниченной информационной емкостью экрана и высоким качеством изображения и поэтому используется для представления сложной графической информации /например, полного изображения одной или нескольких карт электронной плотности/ и получения высококачественных фотокопий. Такая организация работы на двух дисплеях разного типа оказывается весьма полезной в сложных задачах, подобных описанной.

## 5. РЕДАКТИРОВАНИЕ

### 5.1. Описание входных данных

Граф в памяти ЭВМ представляется двумерным массивом /массив соединений/ и сохраняется постранично на внешнем носителе. В трех первых и трех последних столбцах записываются действительные числовые величины координат двух соединяемых отрезком точек. Программа построения графа заносит в массив необходимую информацию.

### 5.2. Алгоритм редактирования

Существуют четыре режима операции редактирования: построение точки, построение отрезка, удаление точки и удаление отрезка. По приказу с пульта графического дисплея задается необходимый режим.

#### 5.2.1. Построение точки

С клавиатуры дисплея перемещаемый маркер подводится в нужное место экрана дисплея. После задания с

### Вызов головной программы моделирования молекулы.

- 1. Задание оси сечения.
- 2. Задание границ фрагмента.
- 3. Задание начального, конечного уровня и шага, с которым изолинии изображаются на экране.
- 4. Изображение на экране отредактированных переменных.
- 5. Просмотр фрагмента одного сечения.
- 6. Просмотр фрагмента нескольких сечений без стирания невидимых линий /наложение вперед/.
- 7. Просмотр фрагмента нескольких сечений без стирания невидимых линий /наложение назад/.
- 8. Просмотр фрагмента нескольких сечений со стиранием невидимых линий.
- 9. Построение на экране выбранного фрагмента молекулы.
- 10. Изображение отредактированного скелета выбранного фрагмента молекулы.
- 11. Редактирование скелета, изображенного на экране.
  - 11.1. Построение точки.
  - 11.2. Удаление точки.
  - 11.3. Построение отрезка.
  - 11.4. Удаление отрезка.
  - 11.5. Построение изображения отредактированного скелета выбранного фрагмента молекулы.
  - 11.6. Окончание редактирования.
- 12. Объявление точки  $C_{\alpha}$ -атомом.
- 13. Работа с поименованными скелетными объектами.
- 14. Завершение сеанса.

Рис.7. Схема алгоритма автоматического моделирования молекулы.

пульты дисплея режима редактирования, на экране отмечается точка и программе становятся доступны ее пространственные координаты. В первую свободную строку массива соединений в первые три столбца заносятся координаты построенной точки. Точка, не имеющая соединений ни с какой другой точкой, называется свободной.

#### 5.2.2. Удаление точки

В этом режиме пространственные координаты точки, на которую указал маркер, сравниваются со всеми коор-

динатами свободных точек, находящихся в массиве соединений. Если координаты совпали, она исчезает на экране, и соответствующая строка в массиве соединений "обнуляется". Если совпадения не произошло или не были удалены все соединения этой точки, на экране дисплея появляется диагностика об ошибке.

#### 5.2.3. Построение отрезка

В режиме построения отрезка маркером указываются две точки. Их координаты становятся доступны программе. Сначала определяется, присутствуют ли эти точки в редактируемом графе. В случае отсутствия одной из них на экран выдается соответствующая диагностика. Затем проверяется, является ли какая-нибудь из указанных точек свободной. Если одна из точек - свободная точка, то в последние три столбца заносятся координаты другой точки, и на экране появляется отрезок, их соединяющий. Если же ни одна из указанных точек свободной не является, то в первую свободную строку массива соединений заносятся координаты обеих точек, и на экране дисплея высвечивается требуемый отрезок. При этом порядок занесения координат точек несуществен.

#### 5.2.4. Удаление отрезка

При помощи маркера на экране дисплея указываются две граничные точки удаляемого отрезка. В массиве соединений ищется этот отрезок. Если он не найден, на экране дисплея появляется сообщение об ошибке. В противном случае указанный отрезок на экране исчезает, и соответствующая строка в массиве соединений "обнуляется". Далее осуществляется проверка на принадлежность какой-либо из этих граничных точек правой или левой части массива соединений. Если такая принадлежность обнаружена, т.е. эта точка имеет соединение с еще какой-то точкой, то в массив соединений не вносятся никаких изменений, если же точка не имела соединений, то ее координаты записываются в первую свободную строку массива соединений /свободная точка/.

#### 5.3. Завершение редактирования

После того, как сеанс редактирования закончен, в массиве соединений могут остаться пропущенные строки. По приказу с пульта графического дисплея о завершении редактирования происходит обращение к подпрограмме уплотнения записи. Эта программа ликвидирует пропуски в массиве соединений путем сдвига на пустые места заполненных строк.

## 6. ОПРЕДЕЛЕНИЕ ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ ГЛАВНОЙ ЦЕПИ МОЛЕКУЛЫ

Для запоминания пространственной структуры главной цепи молекулы в памяти ЭВМ отводится место для массива  $\text{Ca}$ -атомов, имеющего такую же структуру, как и массив сохранений. При объявлении очередной точки скелета  $\text{Ca}$ -атомом координаты этой точки последовательно сравниваются с координатами всех уже ранее построенных  $\text{Ca}$ -атомов, и вычисляется расстояние между ними. Расстояние между двумя соседними  $\text{Ca}$ -атомами в молекуле равно 3,8 Å. Если вычисленное расстояние отлично от 3,8 Å, то на экране появляется сообщение о том, что указанная точка не может быть объявлена  $\text{Ca}$ -атомом. Происходит также проверка на величину плоского угла между тремя соседними  $\text{Ca}$ -атомами /рис.8/. Если угол  $\rho$  меньше или равен  $90^\circ$ , то данная точка также не может быть объявлена  $\text{Ca}$ -атомом.

После того, как получены координаты всех  $\text{Ca}$ -атомов молекулы, можно приступить к анализу исследуемой структуры<sup>/8/</sup>.

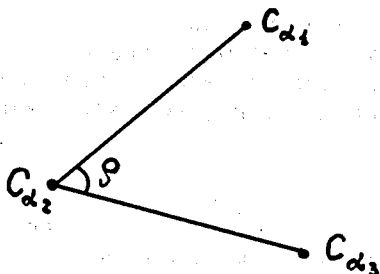


Рис. 8

## 7. ЗАКЛЮЧЕНИЕ

Использование диалоговой методики в задаче построения модели молекулы белка пепсина показало ее эффективность и универсальность. Данная методика позволяет быстро и экономно получить пространственный скелет исследуемой молекулы.

В заключение авторы выражают искреннюю благодарность члену-корреспонденту АН СССР проф. Н.Н. Говоруну и проф. Н.С. Андреевой за постановку задачи и постоянное внимание, а также А.П. Сапожникову и С.Каданцеву за помощь в работе.

## ЛИТЕРАТУРА

1. Richards F.M. J.Mol.Biol., 1968, 37, p.225-230.
2. Кавченко А.В. и др. ОИЯИ, Р10-9325, Дубна, 1975.
3. Андреева Н.С. и др. ДАН СССР, 1976, 228, №2, с.480-483.
4. Щенкова Г.Б. Препринт ИТЭФ №146, м., 1979.
5. Щенкова Г.Б. ОИЯИ, 10-11343, Дубна, 1978.
6. Карлов А.А. и др. ОИЯИ, Б1-10-11999, Дубна, 1978.
7. Карлов А.А., Смолякова Т.Ф. ОИЯИ, Р11-12837, Дубна, 1979.
8. Карлов А.А. и др. ОИЯИ, Р11-9881, Дубна, 1976.

Рукопись поступила в издательский отдел

20 марта 1980 года.