

X-ray resolution less than 0,2 nm was used for learning and prediction of protein secondary structure. Average accuracy of prediction for helix is 74 %, strand is 67 %, coil is 71 % and for three states simultaneously is 68 % of successful prediction.

#### СПИСОК ЛИТЕРАТУРЫ

1. Sternberg M. J. E., Islam S. A. Local protein sequence similarity does not imply a structural relationship // Prot. Engen.— 1990.— 4, N 2.— P. 125—131.
2. Гаск П., Гавранек Т. Автоматическое образование гипотез / Пер. с англ.— М.: Наука, 1984.
3. Мальченко С. Э., Чащин Н. А. Предсказание вторичной структуры белков // Биополимеры и клетка.— 1992.— 8, № 5.— С. 21—31.

Ин-т молекуляр. биологии и генетики  
АН Украины, Киев

Получено 26.04.93

УДК 547.963.3+577.323

**А. И. Егоренков, В. В. Король**

#### **ПАКЕТ ПРИКЛАДНЫХ ПРОГРАММ ДЛЯ ГРАФИЧЕСКОГО ИЗУЧЕНИЯ ТОПОЛОГИИ ПОВЕРХНОСТИ ПОТЕНЦИАЛЬНОЙ ЭНЕРГИИ, СООТВЕТСТВУЮЩЕЙ КОНФОРМАЦИОННОЙ ДИНАМИКЕ МОЛЕКУЛЫ ДНК**

*В работе описан пакет прикладных программ, предназначенный для визуального (графического) изучения топологии поверхности потенциальной энергии, соответствующей внутренним движениям двойной спирали ДНК. Пакет применен для анализа данных аналитического моделирования торсионной динамики ДНК, описывающей процесс локального раскрытия пар азотистых оснований. Обсуждаются возможности программ для решения задач численного моделирования динамики ДНК методами атом-атомных потенциалов.*

*Программы имеют удобный пользовательский интерфейс и современные алгоритмы работы с трехмерной графикой (построение карт поверхностей и изополос, динамическое картирование). Программы реализованы для персонального компьютера IBM PC AT/XT в стандартной конфигурации с видеоадаптером CGA/VGA, язык программирования С.*

**Введение.** Молекула ДНК как сложная биологическая система обладает большим разнообразием внутренних движений и их сложной иерархией. При изучении механизмов биологического функционирования ДНК необходимо учитывать динамические возможности молекулы. Классификация разных движений ДНК позволяет выделить следующие их типы, различающиеся по характерным временам и энергиям [1, 2]: малые колебания атомов около положения равновесия; ограниченные движения сахаров, фосфатов, азотистых оснований около положения равновесия; малоамплитудные торсионные и изгибные движения двойной спирали ДНК; движения, связанные с А→В→Z-переходами; движения большой амплитуды, зависящие от изменения суперспирального состояния ДНК; локальное расплетание двойной спирали. Один из возможных типов внутримолекулярных движений, а именно: локальное раскрытие пар оснований попадает в две сильно отличающиеся по характерным временам группы движений:  $10^{-7}$ — $10^{-5}$  с для раскрытия отдельной пары оснований и  $10^{-4}$ — $10^{-2}$  с для раскрытия пар оснований, связанного с расплетанием двойной спирали ДНК. Данный тип конформационной подвижности важен для понимания предполагаемых механизмов белково-нуклеинового взаимодействия и возможен при взаимодействии ДНК с лигандами разной природы [3], при флуктуационном локальном раскрытии пар азотистых оснований [4], при резонан-

сном поглощении микроволнового излучения [6]. Дополнительный интерес к данному типу внутренней подвижности двойной спирали ДНК связан с дискуссией о существовании и механизме долгоживущих «открытых состояний» [4, 5]. Раскрытие пар азотистых оснований изучалось различными экспериментальными методами: методом спинных меток [7], рентгеновской дифракции [8], ЯМР [9], реакции с формальдегидом [10, 11], водородно-третиевого обмена [12]. В результате исследований стало ясно, что характер раскрытия пар азотистых оснований определяется вторичной структурой ДНК, зависимостью этой структуры от последовательности оснований. Тем не менее на многие вопросы, связанные с изучением процесса раскрытия пар оснований, не получено однозначных ответов. Неизвестны конформации ДНК в «открытых состояниях», промежуточные состояния ДНК при переходе ее в раскрытое состояние, динамика движения азотистых оснований и сахаро-фосфатного остова, условия возникновения «открытых состояний». В частности, не выяснены возможность и условия существования долгоживущих раскрытых пар азотистых оснований. Для решения этих вопросов и интерпретации экспериментальных данных необходима разработка адекватных теоретических моделей и методов.

В Борн-Оппенгеймеровском приближении каждому мгновенному положению ядер молекулярной системы соответствует свое распределение электронной плотности и единственное значение потенциальной энергии системы. Совокупность таких значений образует многомерную поверхность потенциальной энергии (ППЭ), содержащую всю информацию о поведении системы атомов в заданном электронном состоянии. Поэтому анализ конформационной динамики, соответствующей внутренним движениям, можно представить как анализ топологических особенностей многомерной ППЭ. Многомерную ППЭ для такой структурно-нежесткой молекулярной системы, как ДНК, трудно представить наглядно. Поэтому при топологическом методе изучения конформационной динамики возникает проблема замены гиперповерхности некоторым наглядным объектом, доступным для графического анализа. В стереохимии существует метод гомоморфного отображения ППЭ на граф-дерево. Такой конформационный граф отражает основные топологические особенности ППЭ и может быть использован для оценки конформационных свойств молекулярной системы. В конформационном анализе макромолекул традиционно гиперповерхность изучается при помощи конформационных карт, являющихся сечением гиперповерхности по двум переменным. Имея набор таких сечений и разумно выбрав пары переменных, можно получить представление об исследуемом конформационном процессе.

В данной работе для анализа гиперповерхности потенциальной энергии авторы используют метод сечения ППЭ по трем существенным переменным-координатам. В этом случае, применяя динамическое картирование, визуализируется зависимость ППЭ молекулярной системы от любых трех координат (третьей координате соответствует временная переменная), по набору сечений делаются выводы обо всей гиперповерхности и описываются конформационные возможности молекулы. Практически значимое описание конформационной динамики с помощью топологического анализа следует проводить исходя из понятия стационарных (критических) точек: локальные минимумы отвечают устойчивым конформациям, седловые точки соответствуют переходным состояниям. В критических точках гиперповерхности производные энергии по всем независимым переменным обращаются в ноль. В общем случае классификация этих точек проводится по значениям матрицы вторых производных.

Метод сечений и его компьютерная визуализация, разработанные авторами, имеют ряд преимуществ по сравнению с другими методами исследований ППЭ. Для расчета ППЭ таких многоатомных систем, как ДНК, размер пространства  $3 \cdot N - 6$  (где  $N$  — число взаимодействующих

ших атомов;  $b$  — число степеней свободы, относящихся к системе в целом, для линейной молекулы размер пространства  $3 \cdot N - 5$ ). В методе сечений отпадает необходимость расчета всей ППЭ, что существенно для проведения вычислительного эксперимента на персональной ЭВМ с такой молекулярной системой, как ДНК. Кроме этого, известные компьютерные методы работы с трехмерной графикой позволяют создать удобную интерактивную систему для анализа потенциальной поверхности.

**Метод.** В качестве метода визуализации трехмерного представления ППЭ и построения карт изополюсов используются модификации алгоритма «плавающего горизонта» [13, 14] и алгоритма, описанного в работе [15].

Работа с пакетом организована как работа с меню, включающим следующие разделы: создание файла данных или вызов готового файла для последующей визуализации; визуализация данных в виде поверхности или карты изополюсов; поворот и масштабирование поверхности; запоминание поверхности или карты изополюсов для последующего динамического картирования; динамическое картирование в двух режимах (быстрый и медленный) записанных изображений; просмотр отдельных кадров «фильма»; вывод полученных изображений на принтер (без цветовой градаций). Режим создания файлов имеет три подрежима: вызов файла, полученного из расчетов методом полуэмпирического силового поля или по заданной с клавиатуры формуле; ручной ввод двухмерного массива данных; расчет двухмерного массива по введенной с клавиатуры аналитической функции, описывающей потенциал взаимодействия. Интерпретатор, реализующий лексический и синтаксический разборы формулы, вводимой в интерактивном режиме, написан на основе алгоритма, указанного в работе [16]. Получающееся в результате визуализации на экране монитора изображение ППЭ имеет девятицветную окраску, соответствующую уровням энергии молекулярной системы. Такая цветовая окраска облегчает графический анализ полученных сечений ППЭ.

Приведем основные технические характеристики пакета программ. Размер записи на диске (без массивов данных для визуализации) — 100 Kb; требуемый для работы объем оперативной памяти около 300 Kb; адаптер типа EGA (640×350); 16 цветов; принтер типа EPSON; язык программирования Си.

Для процесса визуализации в случае динамического картирования (включение в модель третьей существенной переменной) важны временные характеристики. Хронометрирование, осуществленное на компьютере IBM PC с процессором INTEL 80286 в режиме Турбо, даст следующие параметры, достаточные для получения информативных «фильмов»: время визуализации одного кадра с изображением поверхности или карты изополюсов — 1 с, время расчета массива данных (матрица 51×51) по заданной с клавиатуры формуле около 40 с.

В отличие от имеющихся отечественных и зарубежных пакетов программ работы с трехмерной графикой, разработанный авторами пакет ориентирован на задачи теоретической молекулярной биологии и биофизики, объединяет в себе ряд необходимых для топологического анализа ППЭ процедур.

**Результаты и обсуждение.** Возможности пакета использованы для решения задач, связанных с одним из новых и во многом еще спорных направлений исследования конформационной динамики ДНК, в частности, изучения раскрытия пар азотистых оснований. В настоящее время для объяснения экспериментальных данных по водородно-третичному обмену, которые не укладываются в схему линейной волновой теории, ряд авторов предложил различные нелинейные модели, описывающие механизм «open states» в ДНК (см. обзоры [17, 18]). Необходимость адекватного отражения в моделях конформационных свойств реальной молекулы ДНК требует предварительного анализа потенциалов, моделирующих взаимодействие оснований внутри пары. Методика такого

анализа реализуется с использованием графического изучения ППЭ и состоит из двух взаимосвязанных этапов:

а) сравнительный анализ потенциалов разных феноменологических моделей;

б) сравнение модельных потенциалов с таковыми, полученными на основе расчетов с использованием атом-атомных потенциальных функций.

Пакет программ «Топология» дает возможность реализовать оба этапа. Кроме того, он позволяет провести подбор численных значений, входящих в модельные потенциалы параметров, характеризующих различные компоненты энергии (водородная связь, диполь-дипольное взаимодействие, диполь-индуцированное взаимодействие), и привести их в соответствие с данными численного моделирования методами атом-атомных потенциалов.

В работе [19] показано, что решение системы нелинейных дифференциальных уравнений, описывающих торсионную динамику ДНК, можно представить в виде некоторой траектории движения механической частицы-аналога в плоскости  $\{\varphi_1, \varphi_2\}$  (где  $\varphi_1, \varphi_2$  играют роль «координат-частицы», аналогичных существенным переменным-координатам, по которым можно проводить сечение ППЭ), под действием потенциала  $[-U(\varphi_1, \varphi_2)]$ . Данный потенциал входит в уравнение движения вида:

$$\frac{d^2\varphi}{dx^2} = -\frac{\Delta U(\varphi_1, \varphi_2)}{\Delta\varphi_i}, \quad i = 1, 2.$$

Такая аналогия между решением системы дифференциальных уравнений и траекториями в плоскости позволяет дать новую трактовку уже известных результатов и качественно проанализировать другие возможные решения. Поэтому графический анализ поверхности  $U(\varphi_1, \varphi_2)$  для разных моделей дает исследователю дополнительную информацию.

Анализ ППЭ проведен для трех разных моделей торсионной динамики ДНК, включающих локальное раскрытие пар оснований. Визуализировалась зависимость потенциальной энергии  $U$  от «координат»  $\{\varphi_1, \varphi_2\}$ . В исследуемых потенциалах отсутствует третья переменная-«координата», поэтому возможности динамического картирования (третья переменная) были использованы для визуализации характера изменения ППЭ в зависимости от изменения численного соотношения силовых констант, характеризующих различные взаимодействия между азотистыми основаниями в паре. Отметим, что данные потенциалы не учитывают различия во взаимодействии в парах А—Т и G—C.

Из результатов исследований следует, что для потенциала Йомосы [20] существуют дополнительные области устойчивых конформационных состояний (локальные минимумы на ППЭ), кроме Уотсон-Криковского (центрального минимума на ППЭ), при отклонении оснований на  $\pm 180^\circ$  от центрального равновесного положения. Траектории перехода из центрального минимума в дополнительные одинаковы и соответствуют самосогласованным отклонениям оснований в сторону одной или разных бороздок двойной спирали ДНК. Расчеты показали, что необходимым условием существования указанных дополнительных минимумов является определенное соотношение численных значений силовых констант, характеризующих водородную связь и диполь-дипольное взаимодействие оснований. А именно: силовая константа, которая характеризует водородную связь, должна быть по значению на порядок меньшей, чем константа диполь-дипольного взаимодействия. При сближении численных значений этих констант локальные минимумы на ППЭ исчезают и в случае их равенства остается только один локальный минимум на ППЭ, соответствующий Уотсон-Криковскому положению оснований нуклеотидной пары. С дальнейшим увеличением численного значения силовой константы, характеризующей Н-связь, в сравнении с константой диполь-дипольного взаимодействия становится энер-

гетически выгодным только такое движение оснований в паре, когда одно из оснований не отклоняется от положения равновесия, а другое — движется в сторону большой или малой бороздки ДНК.

Аналогичный анализ изменения ППЭ был проведен для потенциала Гочева—Лисы—Федяшина—Якушевич (ГЛФЯ) [21], который учитывает водородную связь и диполь-индуцированное взаимодействие. Для него, в отличие от потенциала Йомосы, в области значений углов отклонения оснований пары на  $\pm 180^\circ$  не существует дополнительных устойчивых состояний (локальных минимумов на ППЭ), кроме Уотсон-Криковского. Условием существования Уотсон-Криковского равновесного состояния (центрального минимума на ППЭ) является превышение численного значения силовой константы водородной связи по сравнению с константой диполь-индуцированного взаимодействия. Этим потенциал ГЛФЯ отличается от потенциала Йомосы.

При графическом анализе топологии ППЭ, соответствующей потенциалу Жаанга [22], необходимо учитывать изменения численных соотношений трех силовых констант, характеризующих Н-связь, диполь-дипольное и диполь-индуцированное взаимодействие. Расчеты ППЭ для потенциала Жаанга показали, что условием существования дополнительных (кроме Уотсон-Криковского) положений конформационного равновесия (локального минимума на ППЭ) в области значений углов отклонения на  $\pm 180^\circ$  является превышение на порядок численного значения константы диполь-дипольного взаимодействия по сравнению с константой Н-связи (при изменении указанного соотношения дополнительные локальные минимумы исчезают), что согласуется с результатами изучения поверхности потенциала Йомосы. При этом в зависимости от численного значения константы диполь-индуцированного взаимодействия возможны две ситуации. Если константа диполь-индуцированного взаимодействия близка по значению к константе Н-связи, то траектории раскрытия оснований пары (пути наименьшей энергии на ППЭ) при движении оснований в сторону разных бороздок или в сторону одной бороздки одинаковы и соответствуют самосогласованным отклонениям. При увеличении значения этой силовой константы и приближении к значению константы диполь-дипольного взаимодействия высота энергетического барьера для самосогласованного раскрытия оснований в сторону одной (любой) бороздки увеличивается, седловая точка такого конформационного перехода становится менее выраженной и для раскрытия пары остается доступным самосогласованное движение оснований в сторону разных бороздок.

Графические демонстрации указанных изменений ППЭ, трудно воспроизводимые на страницах журнала, могут быть высланы заинтересованным исследователям в виде записанных на дискете динамических «фильмов» или набора распечаток на принтере.

Исследованное графическим методом различие в характере изменений ППЭ для разных потенциалов показывает степень их несогласованности с точки зрения моделирования взаимодействий оснований в паре при ее раскрытии и при возможной реакции на изменения численных значений силовых констант. В данной работе не обсуждаются возможности существования тех физических условий, при которых могут быть указанные выше численные соотношения силовых констант. Все же отметим, что, хотя раскрытие пар оснований может проходить в разных условиях (частичное расплетание двойной спирали ДНК и т. д.), требуемое в некоторых потенциалах соотношение численных значений силовых констант для существования дополнительных минимумов на ППЭ физически нереально.

Окончательный ответ на вопрос об адекватности описания раскрытия пар оснований аналитическими потенциалами может дать численное моделирование методами молекулярной динамики на основе атомных потенциальных функций [23—25]. Использование последних позволит в явном виде учесть особенности взаимодействия в А—Т, G—C-парах и влияние на локальное раскрытие пары соседних нуклеотид-

ных пар. Графический метод изучения топологических свойств энергетической поверхности, без сомнения, будет полезен в этом случае как предварительный этап при численном моделировании конформационной подвижности ДНК [26]. В частности, свойства окрестностей равновесных критических точек могут помочь в нахождении амплитуд и частот малых колебаний, а свойства окрестностей седловых точек ППЭ — определить характер конформационного перехода.

Авторы признательны Л. В. Якушевич (Ин-т биофизики клетки, Пущино) за внимание к работе и полезные замечания.

**Резюме.** В роботі описано пакет прикладних програм «Топологія», призначений для візуального (графічного) вивчення топології поверхні потенціальної енергії, яка відповідає конформаційній динаміці внутрішнього руху молекули ДНК. Пакет застосовується для аналізу даних аналітичного моделювання торсіонної динаміки ДНК, яка описує локальне розкриття пар азотистих основ. Обговорюється можливість програм для вирішення задач з чисельного моделювання динаміки ДНК методами атом-атомних потенціалів. Програма має зручний інтерфейс і використовує сучасні алгоритми роботи з тривимірною графікою (побудова карт поверхонь та ізосмуг).

Програма реалізована для персонального комп'ютера IBM PC AT/XT, мова програмування — С.

**Summary.** The package of applied programs «Topology» specially designed visual (graphical) investigations of the potential energy surface topology corresponding to the conformational dynamics of the internal movements of DNA molecule is described. The package is used to analyse the data of the analytical simulation DNA torsion dynamics.

Possibilities of the program to solve problems of numerical simulation of DNA dynamics by the method of atom-atomic potentials are discussed. The program has an adaptable user interface and it uses modern algorithms of three-dimensional graphic (surfaces and isolayers maps construction, dynamic mapping). The program has been realized for the IBM PC AT/XT personal computer, C being the programming language.

#### СПИСОК ЛИТЕРАТУРЫ

1. Полозов Р. В., Якушевич Л. В. Теоретический анализ внутренних движений ДНК // Межмолекуляр взаимодействия и конформация молекул.— Пущино: ОНТИ НЦБИ, 1987.— С. 186—191.
2. Якушевич Л. В. Динамика ДНК // Молекуляр. биология.— 1989.— 23, № 3.— С. 652—661.
3. Von Hippel P. H., Berg O. G. On the specificity of DNA-protein interactions // Proc. Nat. Acad. Sci. USA.— 1986.— 83, N 6.— P. 1608—1612.
4. Франк-Каменецкий М. Д. Флуктуационная подвижность ДНК // Молекуляр. биология.— 1983.— 17, № 3.— С. 639—652.
5. Frank-Kamenetskii M. D. How the double helix breathes // Nature.— 1987.— 238, N 2.— P. 17—18.
6. Zhang Chun-Ting. Harmonic and subharmonic resonances of microwave absorption in DNA // Phys. Rev.— 1989.— A 40, N 4.— P. 2148—2153.
7. Shih-Chung Kao, Bodsi A. M. Local base dynamics and local structural features in RNA and DNA duplexes // Biochemistry.— 1985.— 24.— P. 5465—5469.
8. Holbrook S., Sung-Hou Kim. Local mobility of nucleic acids as determined from crystallographic data // J. Mol. Biol.— 1984.— 173.— P. 361—388.
9. Gueron M., Kochoyan M., Leroy J.-L. A single mode of DNA base-pair opening drives imino proton exchange // Nature.— 1987.— 328, N 6125.— P. 89—92.
10. Von Hippel P., Wong K.-Y. Dynamic aspects of native DNA structure: Kinetics of the formaldehyde reaction with calf thymus DNA // J. Mol. Biol.— 1971.— 61.— P. 587—613.
11. Frank-Kamenetskii M. D. Structure and motion: membranes, nucleic acids and proteins // Eds E. Clementi, G. Corongli, M. Sarma, R. Sarma.— New York: Adenine press, 1985.— P. 417—432.
12. Kochoyan M., Leroy J. L., Gueron M. Processes of base-pair opening and proton exchange in Z-DNA // Biochemistry.— 1990.— 29, N 20.— P. 7799—7805.
13. Роджерс Д. Алгоритмические основы машинной графики.— М.: Мир, 1989.— 480 с.

14. Баяковский Ю. М., Галактионова В. А., Михайлова Т. Н. ГРАФОР. Графическое расширение ФОРТРАНА.— М.: Наука, 1985.— 288 с.
15. Сегерлинд Л. Применение метода конечных элементов.— М.: Мир, 1979.— 312 с.
16. Донован Дж. Системное программирование.— М.: Мир, 1975.— 514 с.
17. Zhou Gen-fa, Zhang C.-T. Short review on the nonlinear motion in DNA // Phys. Scr.— 1991.— 43, N 3.— P. 347—352.
18. Стариков Е. Б. Нелинейная физика нуклеиновых кислот: солитоны или автосолитоны // Молекуляр. биология.— 1990.— 24, № 6.— С. 1504—1524.
19. Якушевич Л. В. Исследование системы двух нелинейных дифференциальных уравнений, имитирующих торсионную динамику ДНК.— Пущино, 1989.— 13 с.— (Препринт / Науч. центр биол. исследований).
20. Yomosa S. Solitary excitations in deoxyribonucleic acid (DNA) double helices // Phys. Rev.— 1984.— A 30, N 1.— P. 474—480.
21. Fedyanin V. K., Gochev I., Lisy V. Nonlinear dynamics of bases in a continual model of DNA double helices // Stud. biophys.— 1986.— 116, N 1.— P. 59—64.
22. Zhang C.-T. Soliton excitations in deoxyribonucleic acid (DNA) double helices // Phys. Rev.— 1987.— A 35, N 2.— P. 886—891.
23. Keepers J., Kollman P., Weiner P., James T. Molecular mechanical studies of DNA flexibility: Coupled backbone torsion angles and base-pair openings // Proc. Nat. Acad. Sci. USA.— 1982.— 79.— P. 5537—5541.
24. Ramstein J., Lavery R. Base pair opening pathways in B-DNA // J. Biomol. Struct. and Dyn.— 1990.— 7, N 4.— P. 915—933.
25. Briki F., Ramstein J., Lavery R., Genest D. Evidence for the stochastic nature of base pair opening in DNA: A brownian dynamics simulation // J. Amer. Chem. Soc.— 1991.— 113, N 7.— P. 2490—2493.
26. Егоренков А. И., Полозов Р. В. Графический метод изучения раскрытия пар азотистых оснований ДНК.— М., 1988.— 27 с.— Рукопись деп. в ВИНТИ, № 4152—В88.

Укр. мед. ун-т, Киев

Получено 09.03.92