

Аннотация

В настоящей статье мы обсуждаем некоторые геометрические подходы к задачам молекулярной биологии.

In the present paper we discuss some geometrical approaches to molecular biology problems.

Геометрическое моделирование в биологии:
ПОИСК ВОЗМОЖНЫХ ПОДХОДОВ
Geometrical modelling in biology: discovering
possible approaches

В. К. Боженко[†], А. О. Иванов[‡], А. С. Мищенко[§],
А. А. Тужилин[¶], А. М. Шишкин^{||}

Цель настоящей статьи — проиллюстрировать возможности геометрического моделирования на примере его приложения в биологии. Помимо огромной сложности живых структур, мотивирующей необходимость наглядного их представления для осознания скрытых закономерностей, управляющих их функционированием, важность геометрического подхода объясняется еще и тем, что многие жизненно важные свойства того или иного организма существенно зависят как от его формы в целом, так и от формы составляющих его частей. При этом геометрические особенности проявляют себя на всех уровнях структуры, начиная с макроуровня и заканчивая взаимодействием отдельных молекул. Даже социальные взаимодействия, в которых важную роль играют априорные симпатии, во многом определяются внешним видом контактирующих индивидов, что объясняется геометрической визуализацией свойств, отвечающих за жизнеспособность и адекватность возникающих контактов. В данной работе мы сосредоточим основное внимание на проявлениях геометрии в микромасштабе, а именно, в мире биологических молекул и совокупностей клеток. Дополнительную информацию можно найти в [1]–[5].

Начнем с краткого обсуждения того, как геометрическое представление помогает в научных исследованиях. Отметим, что основная задача при изучении объектов — догадаться и сформулировать гипотезы, описывающие их свойства и взаимосвязи. Использование визуальных компьютерных

*Эта заметка написана специально для сборника статей, посвященного юбилею Ректора Московского государственного университета, академика В. А. Садовниченко.

[†]e-mail: vbojenko@mail.ru, тел:(495)1201114

[‡]e-mail: aoiva@mech.math.msu.su, тел:(495)9393940

[§]e-mail: asmish@mech.math.msu.su, тел:(495)9393798

[¶]e-mail: tuz@mech.math.msu.su, тел:(495)9393940

^{||}e-mail: schy@mail.ru, тел:(495)1201114

моделей позволяет перевести набор формальных определений в наглядные геометрические свойства, за которыми можно наблюдать, динамически меняя параметры. Последнее дает возможность непосредственно увидеть свойства и взаимосвязи, что приводит к формулировке гипотез, далеко не очевидных без компьютерного эксперимента и геометрической визуализации. Конечно же, далеко не все гипотезы превращаются в теоремы. Чтобы проверить справедливость найденных связей, ученые все больше и больше используют компьютерный эксперимент для того, чтобы или построить контрпример, или найти подсказки, которые позволят доказать гипотезу. При этом контрпримеры также являются важным этапом исследования, потому что позволяют более глубоко осознать изучаемые объекты и уточнить формулировки гипотез.

Ниже мы покажем, как можно использовать геометрическое моделирование на трех примерах их биологии: моделировании “химической лаборатории” живой клетки; изучении формы (конформации) биополимеров таких, как ДНК, РНК и белки; исследование формообразования различных тканей в живом организме.

1 Симулирование внутриклеточного взаимодействия

Одной из важнейших задач современной молекулярной биологии является изучение взаимодействия белков и нуклеиновых кислот, регулирующего важные внутриклеточные процессы. Такие взаимодействия обычно моделируются с помощью дискретных динамических систем на графах с весами на ребрах и вершинах, которые пересчитываются по некоторым правилам. Сложность такого моделирования объясняется большим количеством молекул, участвующих во взаимодействии, а также недостатком информации о природе и законах взаимодействия. Одним из возможных подходов является построение самонастраивающейся многопараметрической модели, в которой бы реализовались хорошо известные глобальные свойства жизнедеятельности клетки такие, как последовательность и продолжительность фаз клеточного цикла и др. При работе с такими моделями особую роль играет наглядность представления поведения весов при динамическом изменении параметров системы.

Нами была разработана программа, позволяющая изучать внутриклеточные взаимодействия. Фактически, с помощью этой программы можно строить графы, приписывать ребрам и вершинам векторные веса, визуализировать эти веса цветными столбиками, расположенными рядом с вершинами и ребрами, задавать интерактивно законы пересчета этих весов во времени, наблюдать изменение во времени векторных характеристик графов. Особенностью нашей компьютерной программы является возмож-

ность изменения всех ее параметров, включая структуру графа, прямо во время работы динамического модуля, что позволяет эффективно управлять динамикой, добиваться требуемого поведения системы. На рис. 1 показано одно из рабочих окон программы.

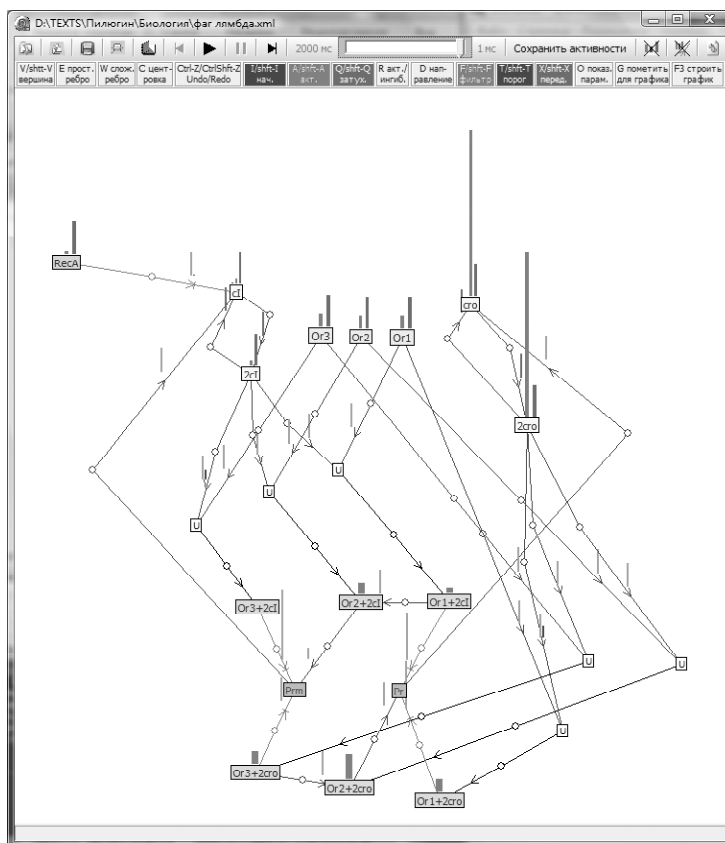


Рис. 1: Моделирование внутриклеточных взаимодействий

2 Геометрические эвристики для изучения конформации белков

Глобальная проблема современной биологии - понять, как функционируют живые клетки. Основными объектами живых клеток являются белки и нуклеиновые кислоты. Одно из приложений - исследование пространственной структуры (так называемой, конформации) этих биополимеров в терминах их химических формул. Важность изучения конформации объясняется су-

шественным влиянием последней на каталитические свойства. Непосредственное математическое моделирование, основанное на законах классической или квантовой физики, описывающих взаимодействие составляющих молекулы атомов, оказывается невозможным благодаря огромному количеству составляющих атомов. Один из подходов — использование геометрических эвристик, тесно связанных с конформацией. Для исследования таких эвристик требуется обширный компьютерный эксперимент, в частности, визуализация различных геометрических свойств приближений и изучение проявления этих свойств в разных биомолекулах.

Рассмотрим частный случай задачи — изучение конформации белка. Напомним, что основу белков живых клеток составляют аминокислоты, соединенные в линейную цепь. В подавляющем большинстве клеток имеется ровно 20 типов аминокислот, формирующих белки. Пространственная структура белка представляет собой углеродно-азотную линейную цепь, к которой крепятся атомы водорода H , кислорода O , а также аминокислотные остатки R_i , характеризующие ту или иную аминокислоту (по одному остатку на кислоту), см. рис. 2.

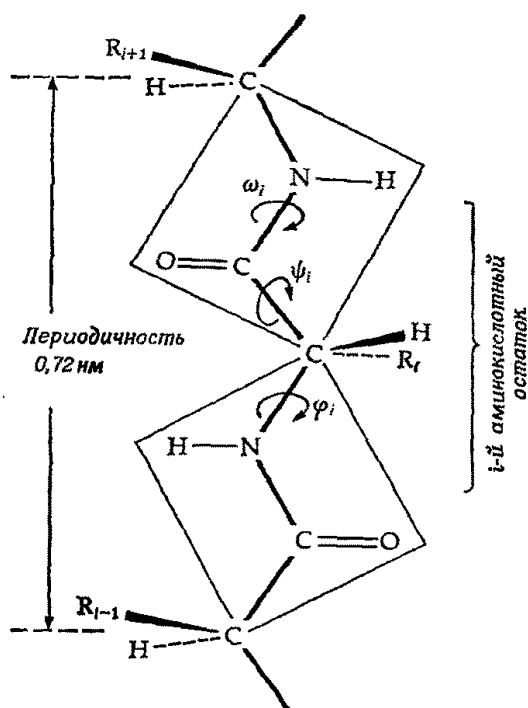


Рис. 2: Фрагмент структуры белка

Углероды из углеродно-азотной цепи, к которым крепятся аминокислот-

ные остатки, традиционно называются альфа-углеродами.

Соединим последовательные вдоль цепи альфа-углероды прямолинейными отрезками (в масштабах задачи атомы можно считать точками). Полученную ломаную будем называть *характеристической*. Конформацию белка будем описывать геометрией именно такой ломаной.

Отметим, что в белках живых клеток расстояние между соседними альфа-углеродами практически одинаково. Тем самым, характеристическая ломаная является равнозвенной и, значит, при известной длине ребра, для ее описания достаточно знать углы между последовательными звеньями (углы первого типа), а также между плоскостями (углы второго типа), натянутыми, построенными для каждого некрайнего звена e так: первая из плоскостей натянута на e и предыдущее звено, а вторая — на e и последующее звено, см. рис. 3: здесь α и β — углы первого типа, а ψ — угол второго типа. Отметим, что иногда углы первого типа называют валентными, а второго — торсионными.

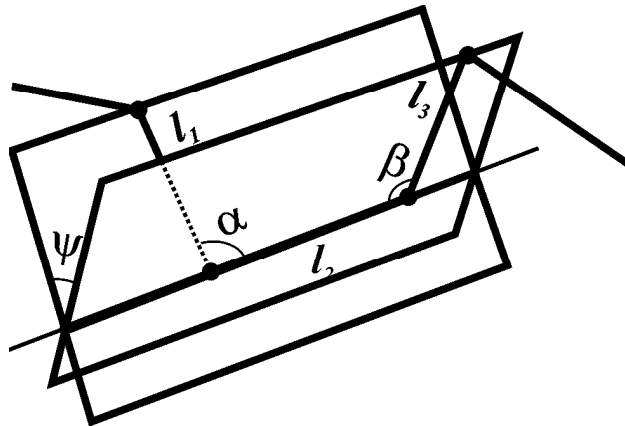


Рис. 3: Углы первого и второго типов

Чтобы лучше понять, как свойства углов связаны с геометрией ломаной, полезно провести аналогию между ломаными и кривыми. Хорошо известно, что бигулярные пространственные кривые $\gamma(s)$ определяются однозначно (с точностью до движения) двумя функциями параметра s — кривизной $k(s)$ и кручением $\kappa(s)$. Напомним определение этих функций.

Пусть $\gamma(s) = (x(s), y(s), z(s))$ — натурально параметризованная гладкая пространственная кривая, т.е. функции $x(s)$, $y(s)$ и $z(s)$ непрерывно дифференцируемы нужное число раз, и абсолютная величина скорости $\gamma'(s) = (x'(s), y'(s), z'(s))$ движения по кривой равна 1.

Положим $v(s) = g'(s)$ — скорость (ее величина равна 1), $k(s) = |g''(s)|$ — кривизна, $n(s) = g''(s)/|g''(s)|$ — главная нормаль (мы предполагаем, что кривизна отлична от нуля, т.е. кривая бигулярна), $b(s) = [v(s), n(s)]$ — би-

нормаль (здесь $[\cdot, \cdot]$ обозначает векторное произведение). Легко показать, что векторы $b'(s)$ и $n(s)$ пропорциональны. Коэффициент пропорциональности $\kappa(s)$ между этими векторами, взятый со знаком минус, называется *кручением* кривой $\gamma(s)$. Таким образом, $b'(s) = -\kappa(s)n(s)$.

Теорема 2.1 *Для каждой положительной гладкой функции $f(s)$ и произвольной гладкой функции $g(s)$ существует пространственная натурально параметризованная кривая $\gamma(s)$ с кривизной $k(s)$, равной $f(s)$, и кручением $\kappa(s)$, равным $g(s)$. Все такие кривые отличаются друг от друга на движение пространства. Плоские кривые однозначно определяются своей кривизной.*

Приведем примеры плоских и пространственных кривых с заданными функциями кривизны и кручения (для пространственных кривых).

Если $k(s) = 0$, то $g(s)$ — прямая линия. В случае пространственной кривой кручение $\kappa(s)$ не определено.

Если $\kappa(s) = 0$, то $\gamma(s)$ — плоская кривая.

Если $k(s) = \text{const} > 0$, то плоская кривая является окружностью. Для пространственной кривой окружность получается при добавлении условия $\kappa(s) = 0$.

Если $k(s) = \text{const} > 0$ и $\kappa(s) = \text{const} \neq 0$, то $\gamma(s)$ — винтовая линия, см. рис. 4.

$$k(s) = 8$$

$$\kappa(s) = 1$$

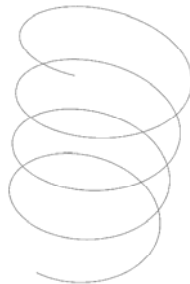


Рис. 4: Винтовая линия

На самом деле, соотношение между видом функций кривизны и кручения, с одной стороны, и формой кривой с другой стороны оказывается крайне нетривиальным: функциям достаточно простого вида могут соответствовать сложно устроенные кривые и обратно. Проиллюстрируем это, приведя еще несколько примеров. Начнем с плоских кривых.

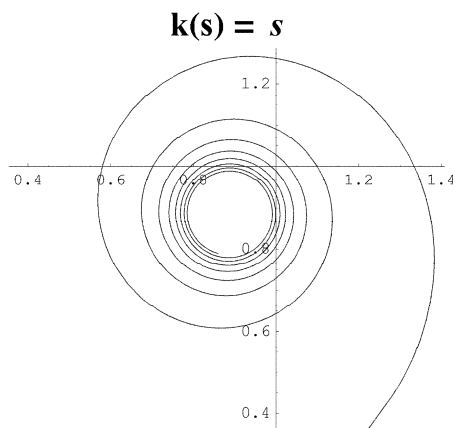


Рис. 5: Спираль

В случае пространственных кривых особенно наглядно сложность зависимости между кривизной, кручением и формой кривой можно увидеть на примере того, как деформируется кривая при изменении этих функций. Ниже приведена серия кадров анимации. Слева на каждом кадре изображена кривая, а справа — графики ее кривизны и кручения. Обращаем внимание на то, что кривизны и кручения в данном случае задаются синусами и косинусами с меняющимися периодами, т.е. как сами функции, так и их изменение устроены достаточно просто. Тем не менее, соответствующие кривые выглядят крайне нетривиально (при просмотре анимационного ролика возникает впечатление, что кривая — это змейка, извивающаяся под звуки флейты факира).

Таким образом, кривизна и кручение являются удобными характеристиками кривой, которые могут существенно упростить описание и исследование пространственной формы.

Вернемся теперь к изучению ломаных. Обсудим для начала, как можно определить характеристики ломаной, которые связаны с геометрией кривых? Одним из критериев естественности определения является требование того, чтобы при для каждой фиксированной бигулярной кривой кривизна и кручение ломаной, вписанной в кривую, стремились бы к кривизне и кручению кривой измельчении ломаной. В действительности, существует много таких функций. Вот некоторые примеры.

1) В обозначениях из рис. 3, определим кривизну в вершине с углом α как отношение $\alpha/(\ell_1 + \ell_2)$, где ℓ_i — длины соответствующих звеньев ломаной, а кручение на ребре длины ℓ_2 — как отношение $\psi/(\ell_1 + \ell_2 + \ell_3)$.

2) Определим кручение во внутренней вершине как величину, обратно пропорциональную радиусу окружности, описанной вокруг треугольника,

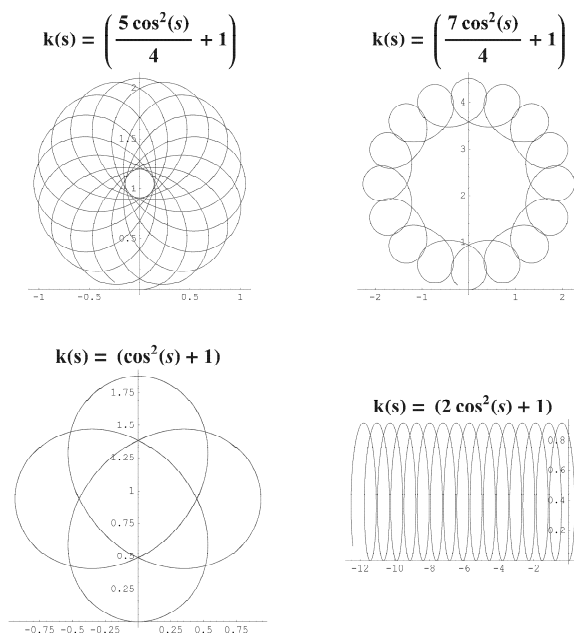


Рис. 6: Серия плоских кривых

вершины которого — предыдущая, данная и последующая вершины ломаной. Кривизну определим как величину, пропорциональную объему тетраэдра, натянутого на три последовательных звена ломаной, деленного на произведение граней этого тетраэдра, примыкающих к среднему звену.

3) В случае равнозвенных ломаных первое определение, без деления на длины ребер, т.е. когда в качестве кривизны и кручения используются углы первого и второго типов, дает функции, отличающиеся от определенных в 1) на некоторые постоянные множители. Поэтому при исследовании свойств ломаной, которые определяются по кривизнам и кручениям с точностью до постоянного множителя, можно воспользоваться определением 2), которое более просто для вычислений.

Приведем результаты численного эксперимента, демонстрирующего как конформацию фрагмента гемоглобина-мутанта, так и соответствующие функции кривизны и кручения.

Отметим, что качественное поведение полученных функций практически одинаково при каждом из приведенных выше трех их определений. Это, как уже отмечалось, является следствием того, что здесь ломаные равнозвенные.

Как можно использовать построенные функции? Напомним, что в гладком случае кривая с постоянными ненулевыми кривизной и кручением является винтовой линией. Заметим, что полученные графики имеют несколько

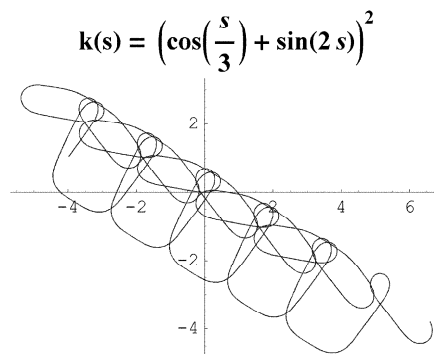


Рис. 7: Еще один пример плоской кривой

фрагментов, где обе функции почти постоянны (их графики горизонтальны). С другой стороны, на левой части рисунка отчетливо видны спиральные фрагменты. В биологии белка такие фрагменты называются альфа-спиралями и являются важным структурным элементом. Другой важный класс элементов белка — это плоские бета-слои, которые тоже отражаются на графиках функций в виде фрагментов, где функция, соответствующая кручению, осциллирует в окрестности нуля (напомним, что бигулярная кривая лежит в плоскости, если и только если ее кручение равно нулю). Таким образом, построенные аналоги кривизны и кручения для ломаных позволяют выделять важные структурные фрагменты белка без апелляции к трехмерному изображению белка. Более того, выделение соответствующих фрагментов функций легко автоматизируется с помощью компьютера (скажем, найти фрагмент, где функция почти постоянна, — простая алгоритмическая задача). Более того, подобное представление позволяет обнаружить новые структурные фрагменты белков, в которых реализуются кривизны и кручения, задающие в случае гладких кривых хорошо известные классы последних.

Еще одним примером использования геометрического подхода является построение механической модели пространственной структуры белка, которая основана на модели упругой фермы, составленной из узлов и жестких стержней, их соединяющих. Предполагается, что в узлах расположены атомы белка (альфа-углероды), а связи интерпретируются как состояние стержней, соединяющих узлы. Предполагается, что деформации фермы зависят от сил, приложенных к узлам, и подчиняются некоторому закону Гука. Силы, приложенные к узлам, порождаются силами притяжения или отталкивания некоторых атомов. Все связи делятся на валентные, пептидные, которые формируют структуру фермы, и водородные, слабые связи, которые деформируют стержни, а в конечном счете и всю ферму.

Расчет статического состояния упругой фермы приводит к нелинейным

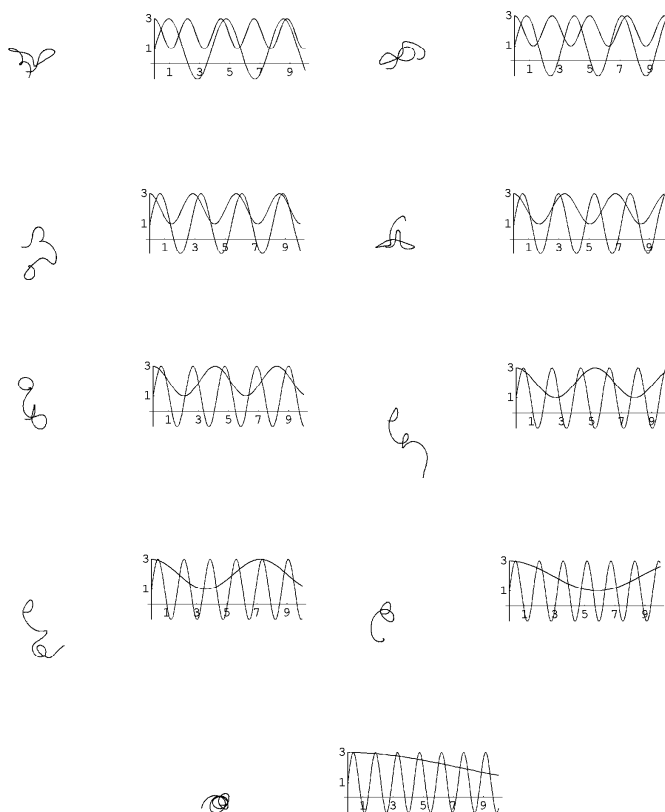


Рис. 8: Зависимость формы пространственной кривой от кривизны и кручения

уравнениям, что приводит к методу линеаризации задачи и, тем самым, к расчету динамики формирования пространственной структуры молекулы.

Отличие от существующих методов вычисления пространственной структуры белка заключается в том, что предлагаемая модель позволяет проводить вычисления при помощи быстрых алгоритмов, сложность которых оценивается порядком $O(n^2)$, где n — длина сиквенса пептида.

3 Формообразование тканей

Интересна способность многих живых низших организмов восстанавливать свою структуру после непоправимого ее нарушения. Например, взяв губку и протерев ее через сито, мы получим аморфную массу клеток. Но

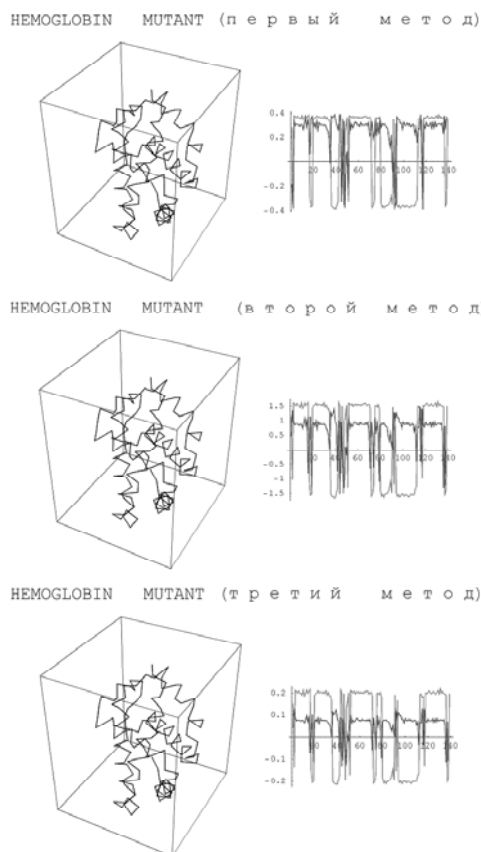


Рис. 9: Разные способы вычисления кривизны и кручения ломаных

если поместить суспензию этих клеток в воду, то через некоторое время из них вновь образуется губка. Кроме того, каноническое формообразование характерно для всех многоклеточных организмов и может естественно наблюдаться как при формировании и развитии организма в процессе эмбриогенеза, так и при всевозможных репарациях тканей. Описаны разные модели, в которых коллектив элементов, связанных между собой локальным взаимодействием, оказывается способным сохранять и восстанавливать некоторую определенную структуру [6].

Мы предположили, что подобное поведение клеток может быть, в первом приближении, обусловлено наличием достаточно жесткой структуры точек крепления (плотных контактов), расположенных на поверхности клеток. Тем самым возникла задача исследования связи пространственного расположения точек крепления (плотных контактов на поверхности клетки) и формирующимися пространственными конгломератами клеток - трех-

мерной структуры ткани.

В рамках этой гипотезы нами была написана программа, позволяющая моделировать различные ткани в зависимости от структуры точек крепления. Простейшая версия нашей модели такова.

Клетки моделируются сферами некоторого радиуса (в простейшем случае все радиусы одинаковы). Структура точек крепления задается конечным набором точек на единичной сфере (модельной сфере). В начальный момент времени мы задаем случайные положения некоторого количества клеток координатами центров соответствующих сфер. Кроме того, для каждой такой клетки мы задаем случайно ориентированную структуру точек крепления, выполняя отображение подобия с модельной сферы на сферу, задающую клетку.

Затем запускается модуль, имитирующий деление клеток для формирования ткани. Программа перебирает имеющиеся клетки и для каждой из них просматривает свободные точки крепления. Если такая точка найдена, то в систему добавляется новая клетка, которая касается данной по выбранной точке крепления, при условии, что эта новая клетка не пересекает внутренностей уже построенных клеток. Если новая клетка добавлена, структура ее точек крепления порождается симметрией относительно общей касательной плоскости в точке касания исходной и построенной клеток.

Мы описали лишь ключевую идею программы. На самом деле, имеется много разных параметров, которые управляют процессом тканеобразования. Так, например, чтобы отразить в модели неточечное касание клеток, форма которых может отличаться от сферы, мы допускаем пересечение клеток с заранее заданным ограничением на глубину. Это делает картину более реалистической.

Преимущество программы состоит в том, что в процессе тканеобразования разрешается менять структуру точек касания, что мгновенно модифицирует изображенную ткань и, тем самым, позволяет эффективно увидеть соответствие между локальной геометрией точек крепления и глобальными свойствами соответствующей ткани. Приведем несколько результатов компьютерного эксперимента рис. 10 (численный эксперимент был проведен М. А. Тужилиным, учеником 10 класса лицея “Вторая школа”, г. Москва).

Работа выполнена при поддержке РФФИ (проекты 07–01–00648, 05–01–22002 НЦНИ), гранта Президента РФ поддержки ведущих научных школ РФ (проект НШ–660.2008.1), программы Эйлера ДААД, а также проекта РНП 2.1.1.7988.

Список литературы

- [1] Fold fully forecast. Computer predicts protein's shape from sequence

- alone. Nature News (23 Sep 2002), doi: 10.1038/news020923-4, News <http://www.nature.com/news/2002/020923/full/news020923-4.html>
- [2] К.В. Шайтан. Молекулярная динамика пептидов. <http://www.moldyn.org/library/moldyn/moldyn.htm>
- [3] К.В. Шайтан, С.С. Сарайкин, Метод молекулярной динамики, 1999, <http://www.moldyn.ru/library/md/default.htm>
- [4] А. В. Финкельштейн. Введение в физику белка. Курс лекций 1999-2000.
- [5] А.В. Финкельштейн, О.Б. Птицын “Физика белка”, Москва, “Книжный дом “Университет””, 2002.
- [6] А.М. Леонтович, И.И. Пятецкий-Шапиро, О.Н. Ставская. Некоторые математические задачи, связанные с формообразованием. Автоматика и телемеханика 4 (1970), стр. 94-107.

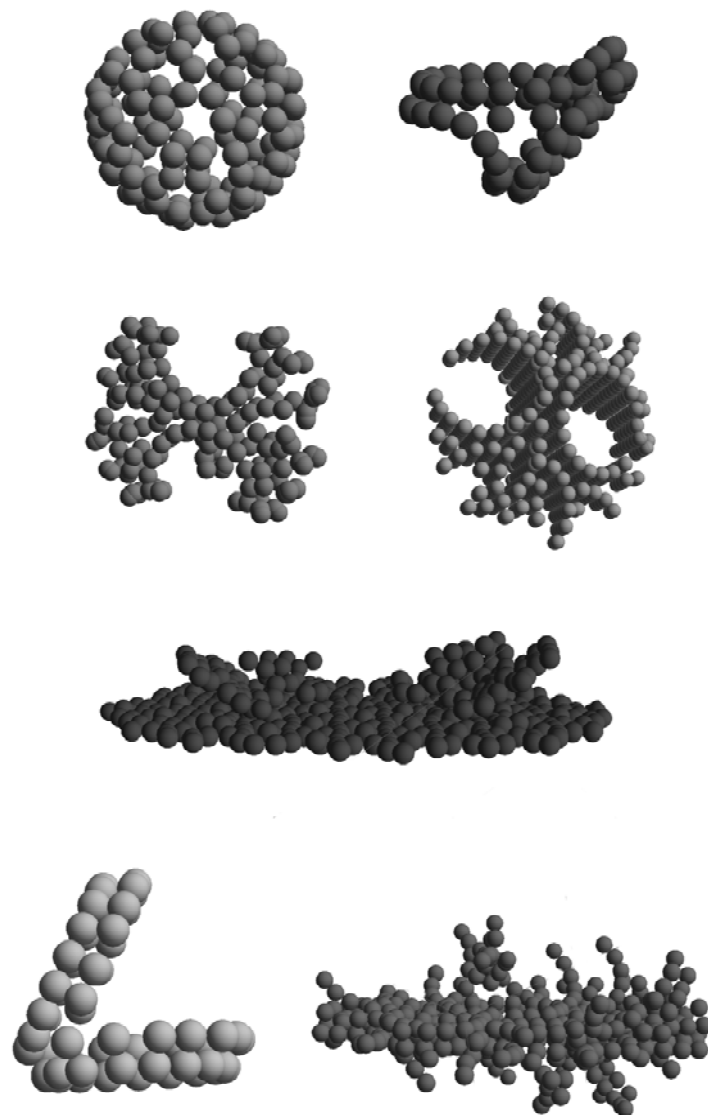


Рис. 10: Тканеобразование