

ПРОБЛЕМЫ МОДЕЛИРОВАНИЯ ФЕНОМЕНА САМОСБОРКИ (ч. I)

А.В. Махиборода*, А.А. Подобин, А.В. Ильичёв

*Департамент прикладной математики МИЭМ,
Национальный исследовательский университет «Высшая школа экономики»*

makhiboroda@yandex.ru

Поступила 01.03.2018

Долгосрочные программы технологического развития и перспективы освоения природоподобных технологий требуют форсированного исследования и создания математических основ моделирования феномена самосборки. В настоящей статье анализируются факты экспериментальных наблюдений явлений самосборки и обосновывается вывод о том, что самосборка представляет собой реализацию алгоритмических форм поведения на молекулярном уровне. Известно, что алгоритмические формы поведения порождаются знаковыми системами определённых типов. Проблема состоит в том, что знаковые системы, функционирующие в компьютерной технике и порождаемая ими алгоритмическая динамика, не соответствуют фактам наблюдения самосборки. Задача математического моделирования самосборки формулируется как разработка специфических знаковых систем, порождающих алгоритмическую динамику, адекватную явлению самосборки. В статье излагаются требования к знаковым системам с распределённым управлением, определяется их состав и общее устройство.

УДК 51-76

Введение

В ноябре 2016 года в Москве был проведен Первый Российский кристаллографический конгресс и издан сборник тезисов докладов Конгресса [1]. Конгресс был задуман как широкий междисциплинарный форум, на котором было представлено шесть тематических направлений, таких как современная кристаллография, методы и техника структурных исследований, кристаллография в биологии и медицине, кристаллография в когнитивных исследованиях, кристаллографические методы в гуманитарных науках и

* Работа выполнена в 2017-2018 годах при поддержке Программы фундаментальных исследований НИУ ВШЭ

образовательные аспекты современной кристаллографии. При этом в пределах каждого из перечисленных тематических направлений был представлен свой максимально широкий список тем, направлений исследований и технологических разработок. При подготовке Конгресса было зарегистрировано более полутора тысяч участников и представлено около 840 докладов. **Девиз форума – «От конвергенции наук к природоподобным технологиям».** Следует признать, что данное мероприятие стало важным событием интеллектуальной жизни страны, поскольку явилось адекватным ответом на цивилизационный вызов в области технологического развития. Научное сообщество России ясно осознаёт, что предшествующий путь технологического развития, основанный на узкой специализации производства и последовательном углублении разделения труда, приобрёл глобальный характер и исчерпал свой идейный, экономический и производственный потенциал. Так, например, при современной технологии для производства одного ноутбука затрачивается 36 тонн различных веществ и расходных материалов. Если производственная техносфера не будет изменена кардинальным образом, на удовлетворение растущих потребностей человека как биологического вида не хватит вещества литосферы планеты. Современное производство предполагает извлечение из недр концентрированных моноэлементных сырьевых компонентов и дальнейшее изготовление из них технических изделий путём удаления лишнего материала из исходного образца. При этом львиная доля добытого вещества при огромных энергозатратах идёт в отвал и захламляет поверхность планеты ядовитыми отходами.

В целом ряде пленарных докладов была озвучена цель переориентации программ технологического развития в направлении создания природоподобных технологий, которые строятся на принципиально иной основе – агрегации и структурообразовании. Это так называемые аддитивные технологии, в основе которых лежат процессы формирования технических изделий из некоего стартового ядра путём запуска управляемых физико-химических процессов структурообразования и роста.

Грандиозные цели, декларированные Конгрессом, останутся не более чем благими пожеланиями, если не будут предприняты форсированные попытки исследования феномена самосборки. Десятилетний опыт разработок в области нанотехнологий в большинстве случаев привёл на практике к производству нанопорошков. Создание сложных структур и функционально способных платформ в наноразмерном диапазоне невозможно без понимания и практического освоения принципов самосборки. Можно без преувеличения сказать, что в свете целей, заявленных Конгрессом, исследование феномена самосборки становится важнейшим направлением фундаментальных исследований, поддерживающих процессы технологического развития.

Следует уточнить, что именно мы понимаем под термином самосборка. Рассмотрим наглядный пример из области эмбрионального роста. Это известный эксперимент по reagregации и пересборке зародыша морского ежа, разрушенного на стадии гастрюлы. Эксперимент описан, например, в учебнике биологии К. Вили и В. Детье [2]. В ходе клеточных делений число клеток зародыша последовательно удваивается. При этом одновременно идёт процесс формообразования зародыша в целом и дифференцировка или специализация каждой отдельной клетки. На стадии гастрюлы зародыш содержит несколько тысяч клеток и имеет достаточно сложную геометрию. Начальная стадия роста структуры гастрюлы морского ежа в схематическом изображении приведена на рис. 1. На рис. 2 приведены реальные изображения общего вида и срезов гастрюлы, полученные на сканирующем электронном микроскопе. Гастрюла представляет собой двухслойный мешок с вдавленным внутрь бластопором. Клетки внешней поверхности

называются эктодермой, клетки, составляющие вдавленный бластопор называются энтодермой. Кроме того, во внутренней полости имеется группа мигрирующих клеток, которые называются мезодермой.

Стадия гастролы это важный этап процесса эмбрионального роста, в ходе которого складывается презумптивная разметка клеточной массы зародыша. В ходе процесса дифференцировки клетки специализируются и разбиваются на три типа, каждый из которых закладывает основу для развития важнейших подсистем и органов будущего организма. Из клеток эктодермы формируется внешний покров и все сопутствующие ему ткани и органы, а также органы чувств и нервная система. Энтодерма преобразуется в выстилку пищевого тракта и ряд внутренних органов, возникающих в результате вдавливания внутрь и удлинения определённых фрагментов бластопора, в том числе органы дыхания. Мезодерма даёт начало кровеносной и опорно-двигательной подсистемам.

Смысл эксперимента состоит в том, что гастрולה разрушается путём диссоциации таким образом, чтобы были устранены межклеточные связи, а сами клетки при этом не

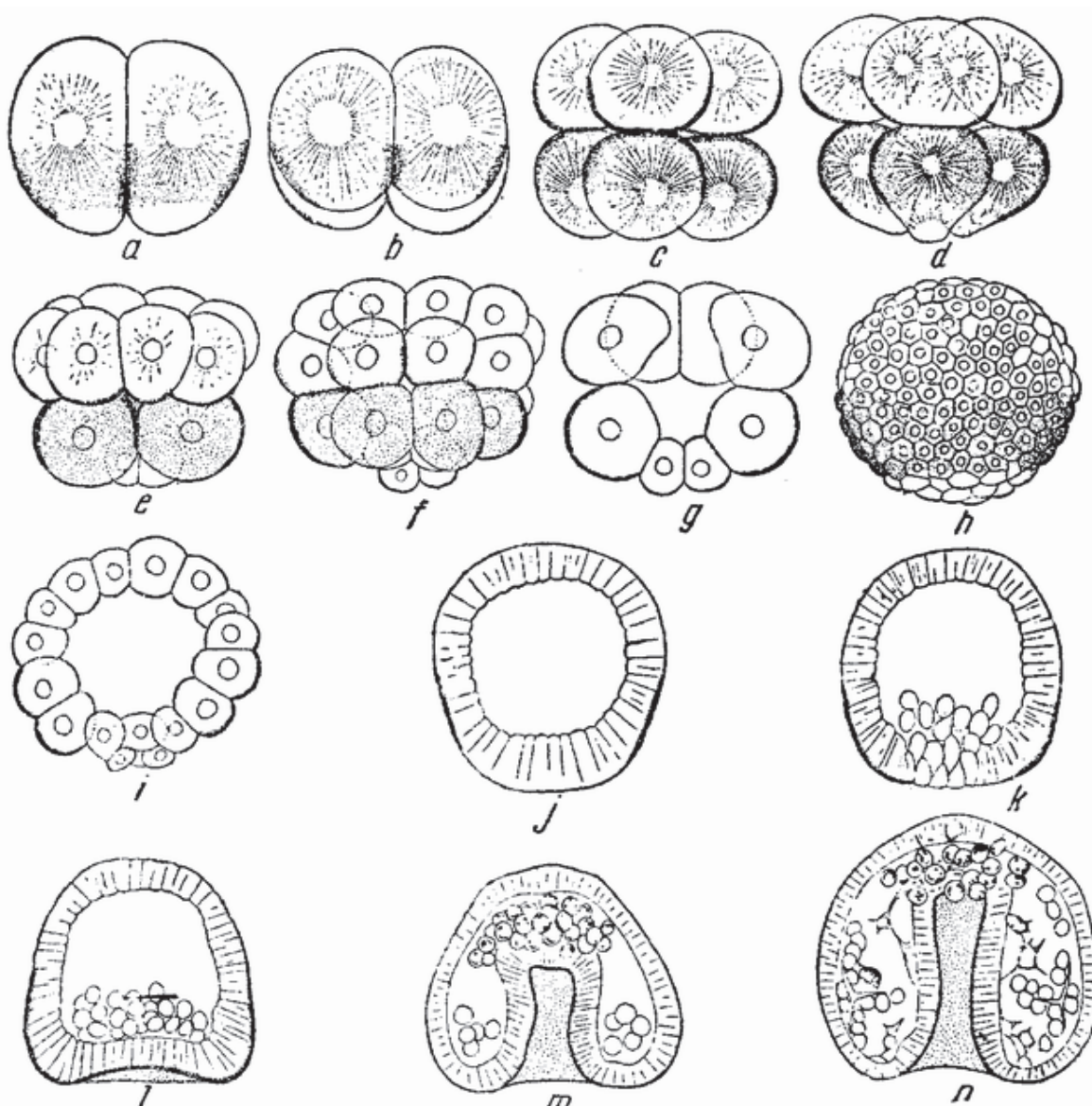


Рис. 1. Начальная последовательность роста зародыша морского ежа, стадия формирования гастролы

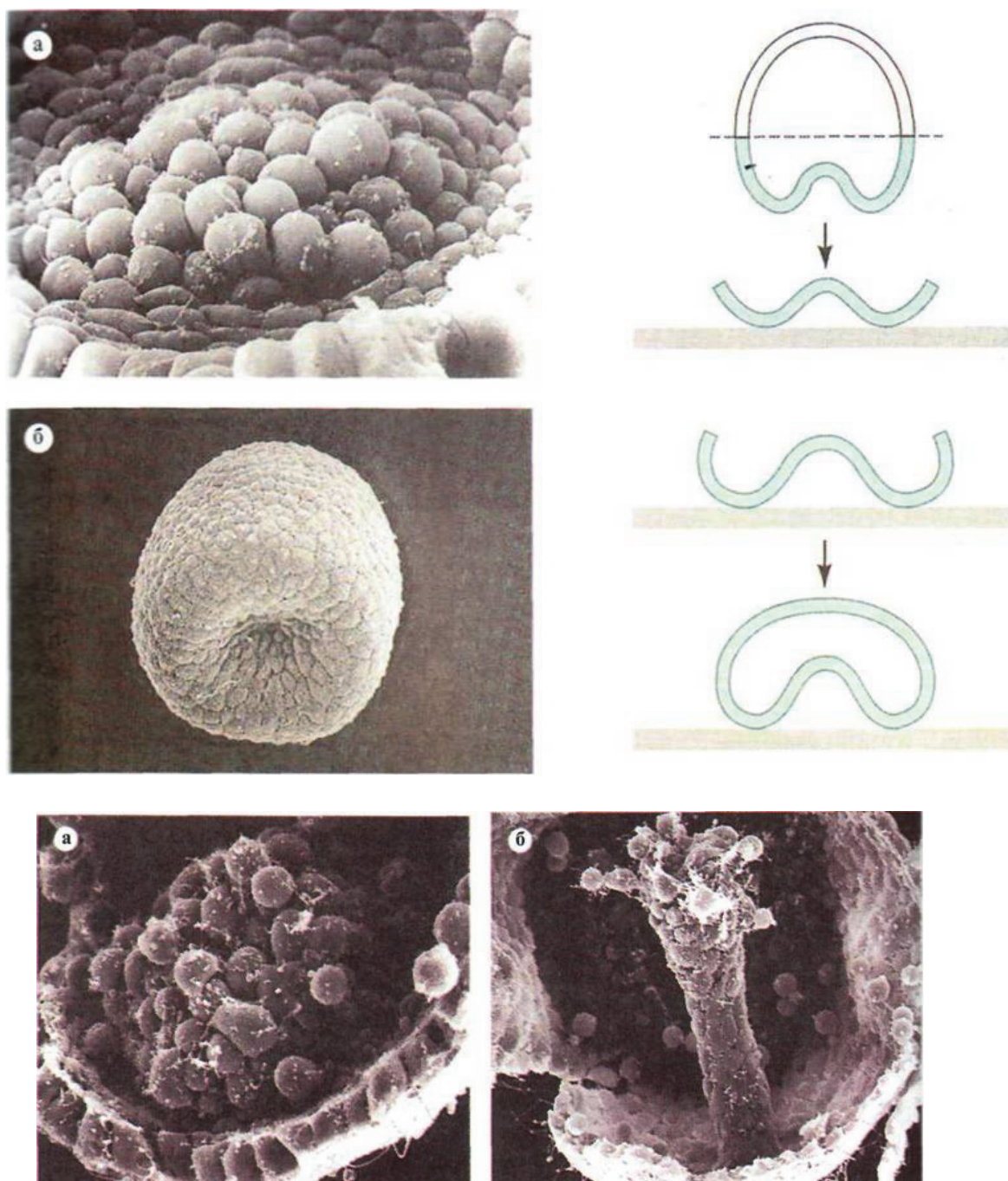


Рис. 2. Изображения гастролы, полученные на сканирующем электронном микроскопе

пострадали и сохранили свою целостность и жизнеспособность. Гастрола перестаёт существовать как функционально-способная структура и превращается в бесформенный ком клеточной массы. Будучи предоставленной самой себе и при отсутствии агрессивных факторов внешней среды, клеточная масса приходит в движение, клетки мигрируют, сортируются, и через некоторое время гастрола полностью восстанавливается. При этом восстанавливается не только геометрия гастролы, восстанавливается и презумптивная разметка. Разные типы клеток, прошедшие свой индивидуальный путь специализации, в ходе клеточной дифференцировки занимают свои места в структуре гастролы. Восстановленный зародыш продолжает прерванный процесс эмбрионального

роста. Логика процесса реагрегации и пересборки до сих пор не установлена, и целостная картина явления самосборки пока отсутствует, что является вызовом современному естествознанию.

Рассмотренный эксперимент позволяет уточнить определение феномена самосборки. Существует обширный класс явлений агрегации и структурообразования, который характеризуется значительным разнообразием условий протекания, механизмов взаимодействия элементов и процессов сборки. Самосборка является специфическим подклассом со своими характерными особенностями, которые не следует путать с другими явлениями спонтанной агрегации. В первую очередь самосборка это многостадийный протяжённый процесс, состоящий из целого ряда строго согласованных последовательностей процедур, который в целом носит детерминированный алгоритмический характер. Однако все, что нам известно об алгоритмах и методах управления процессами из вычислительной техники не применимо к данной ситуации. Многочисленными наблюдениями надёжно установлено, что при самосборке централизованное управление процессом агрегации отсутствует. В процессе самосборки не наблюдается манипулятор, который управляет перемещениями и прикреплёнными пассивных элементов. Участники самосборки это достаточно сложные элементы, обладающие собственной поведенческой активностью и способностью самостоятельного выполнения избирательных монтажных действий. Предопределённость и согласованность развития процесса в целом, является следствием локальных избирательных взаимодействий отдельных участников самосборки. Управление процессом самосборки носит виртуальный характер и распределено по всем участникам процесса. Вся необходимая для управления процессом информация содержится в элементах-участниках. Важнейшим свойством процесса самосборки также является его обратимость, что продемонстрировано рассмотренным выше примером пересборки зародыша морского ежа.

Термин самосборка наиболее активно используется в вирусологии при обсуждении процессов формирования вирусных капсидов. (Собственно термин самосборка был введён в научный оборот именно вирусологами.) Вирусный капсид это замкнутая белковая оболочка, в которой сохраняется и транспортируется наследственный материал вируса. В функции капсида входит избирательная адсорбция вируса на определённых тканях и клетках, а также внедрение наследственного материала вирусной частицы внутрь заражаемой клетки. Вирусный капсид является надмолекулярной структурой, обладающей строгой геометрией и определённой функцией. Капсид формируется в процессе самосборки, участниками самосборки являются вирусные белки, обладающие достаточно сложным поведением и способностью к избирательному взаимодействию. Таким образом, процедуры формирования вирусных капсидов полностью соответствуют данному нами определению явления самосборки. В вирусологии накоплен огромный массив данных экспериментальных наблюдений процессов самосборки вирусных частиц. Результаты кропотливого труда нескольких поколений вирусологов образуют обширную литературу по проблеме самосборки и наше изложение будет в значительной степени базироваться на результатах, полученных в вирусологии. Но это не означает, что мы ограничиваемся исключительно только изучением процессов сборки вирусных частиц. Наша цель построение логики самосборки, применимой в самом широком контексте.

Целью настоящей публикации является обсуждение и решение двух проблем:

1. Обоснование тезиса о том, что самосборка – это алгоритмический процесс, который управляется знаковой системой определённого типа.
2. Разработка специфических алгоритмических систем с распределённым управлением, адекватных явлению самосборки и не имеющих аналогов в современных информационных технологиях.

1. СТРОЕНИЕ И ФУНКЦИИ ВИРУСНЫХ КАПСИДОВ

По современным представлениям вирусы – это определенная форма существования и функционирования нуклеиновых кислот преимущественно во внутриклеточной среде. Вирусные нуклеиновые кислоты могут быть представлены либо ДНК, либо РНК, и в зависимости от этого различают ДНК-содержащие и РНК-содержащие вирусы. Существует определённое разнообразие форм представления и структурной организации вирусных нуклеиновых кислот, в детали которого мы не станем углубляться. О роли вирусов и их жизненных циклах можно прочесть в [3].

Вирусы, как новая форма существования жизни, были открыты В.И. Ивановским в 1892 году. Это был легендарный вирус табачной мозаики (ВТМ). Первоначальное представление о вирусах как о паразитных формах, вызывающих заболевания за счёт поражения клеток быстро устаревает. Обнаружено огромное разнообразие вирусов, большая часть из которых сосуществует с другими биологическими видами, не вызывая никаких болезненных проявлений. Существуют экспертные оценки массы органического вещества, сосредоточенного в вирусных формах, обитающих в воздухе, водной среде и донных отложениях, и оно сопоставимо с объёмами биомассы клеточных форм [4]. Видов вирусов чрезвычайно много, это наиболее многочисленные организмы на Земле (до 250 миллионов в одном миллилитре воды). Они различаются между собой свойствами, которые зависят не только от строения их генов (вида генов), но и от числа генов в каждом вирусе – от 1-го до нескольких десятков, сотен и даже тысяч генов. Современные представления об эволюционной динамике отрицают Дарвиновскую модель эволюции, осуществляемой за счёт спонтанных точечных мутаций генома и последующего отбора удачных комбинаций. В современных моделях эволюция осуществляется в процессе направленного крупноблочного редактирования генома, и основными агентами этого процесса являются вирусы [5,6,7]. Предполагается, что вирусы являются предшественниками и создателями клеточных форм существования биомассы, а также основными двигателями эволюции и видообразования. Вирусам отводится роль создателей генов и их межвидового горизонтального переноса.

Вирусы могут встраиваться в геном клетки хозяина и существовать в нем в скрытом латентном либо активном состоянии. Встроенные в геном вирусы участвуют в репликации генома при клеточных делениях и таким образом могут наследоваться. Вирусные нуклеиновые кислоты, находящиеся в клеточной среде могут осуществлять собственную репликацию, захватывать клеточные механизмы синтеза белков и таким образом воспроизводить свои специфические вирусные белки. Собственно, вирус это есть симбиоз нуклеиновой кислоты и некоторых сопутствующих белков, которые осуществляют, например, встраивание в геном заражённой клетки или захват клеточной системы синтеза белков и запуск процесса экспрессии вирусного генома. В результате перестройки системы синтеза белков заражённой клетки начинается процесс синтеза

перестройки системы синтеза белков заражённой клетки начинается процесс синтеза вирусных белков с использованием клеточных ресурсов.

Вирусные белки, в свою очередь, могут менять метаболизм клетки и в том числе осуществлять построение вирусных капсидов и наполнять их внутреннюю полость копиями вирусных нуклеиновых кислот. В результате во внутренней среде клетки воспроизводятся вирусные частицы, называемые вирионами. Вирион – это форма внеклеточного существования вирусов, в которой нуклеиновая кислота заключена в защитную оболочку, состоящую из белковых молекул и называемую вирусный капсид. Существует несколько вариантов выхода вирионов из заражённой клетки. Самый простой вариант запускает массовый процесс репродукции вирионов до истощения ресурсов клетки, разрыва клетки в результате осмотического шока и высвобождения сотен вирионов одновременно. Существует более сложный вариант поэтапного выхода небольших порций вирионов с сохранением жизнеспособности клетки. И, наконец, так называемый латентный вирус, который встраивается в геном поражённой клетки и воспроизводится в нём при клеточных делениях, последовательно захватывая ткань или клеточную популяцию.

Вирус в форме вириона обладает высокой устойчивостью по отношению к внешней среде и может транспортировать и сохранять вирусную нуклеиновую кислоту неопределенно долго. Капсид обладает способностью к избирательной абсорбции на определенных тканях и типах клеток, аффилированных с данным вирусом и в конечном итоге обеспечивает доставку нуклеиновых кислот во внутреннюю среду клетки. Экспрессия вирусной нуклеиновой кислоты, производство капсидных белков и сборка вириона это одна из форм репродукции вируса. В дальнейшем нас будет интересовать только этот этап в жизненном цикле вирусов, поскольку построение вирусного капсида осуществляется путем самосборки при участии капсидных белков.

Исследование процессов самосборки вирусных капсидов имеет большой практический и теоретический интерес по нескольким направлениям. Борьба с вирусными инфекциями окажется наиболее эффективной в случае, если удастся подорвать репродуктивную способность вирусов путём нарушения процесса сборки вирусных капсидов. Завершённые вирусные капсиды обладают значительной прочностью и устойчивы к разрушительным воздействиям. Мероприятия по расстройству процесса самосборки вирусных капсидов могут оказаться более избирательными и ограничатся меньшим числом побочных эффектов, вредных для организма в целом.

В настоящее время активно разрабатывается перспективное направление создания методами генной инженерии модифицированных вирусных капсидов, которые предполагается использовать для адресной доставки лекарственных препаратов к конкретным тканям и клеткам [8,9,10]. При этом используется способность вирионов обходить иммунную защиту и сорбироваться на строго определённых тканях и клетках. Капсид используется как контейнер для лекарств, и важно уметь управлять не только его сборкой и наполнением, но и разборкой и высвобождением доставленного вещества в конце пути.

Известны также чисто технические разработки использования вирусных частиц для создания развитых поверхностей, например, при производстве литий-полимерных аккумуляторов или топливных элементов [11]. Создание развитых контактных поверхностей осуществляется путём осаждения металлов на каркас из вирусных частиц с после-

дующим удалением белковой основы. Методы технологического применения вирусов описаны в [12,13].

Исследования процессов самосборки вирусных капсидов также могут оказаться актуальным для решения фундаментальных проблем теоретической биологии. В работах Уоддингтона (см. сборник докладов «На пути к теоретической биологии» [14]) излагается острая дискуссия по проблеме взаимоотношения генотипа и фенотипа. Это достаточно сложная проблема даже для одноклеточных организмов. В этой ситуации вирусы представляют собой удобные для исследования простейшие эталонные объекты. Генотип вирусов имеет наименьший объём, это короткие молекулы вирусных нуклеиновых кислот, а фенотип представлен простейшей белковой конструкцией в виде вирусного капсида.

Вирусные капсиды строятся из молекул белков, которые являются продуктами одного или нескольких генов, содержащихся в вирусном геноме. Вирусные капсиды представляют собой надмолекулярные структуры, при образовании которых, как правило, отсутствуют химические связи. Относительно природы агрегации и механизмов прикрепления капсидных белков пока нет обоснованного единого мнения. Вопрос о природе связи и характере прикрепления белков в вирусных капсидах находится в стадии изучения и разработки. По нашему мнению определение механизмов взаимодействия капсидных белков является ключевым в понимании всего процесса самосборки. Во всяком случае, вопрос о контактах и силах сцепления не может быть сведён к простейшим моделям кулоновских или Ван-дер-Ваальсовых взаимодействий заряженных частиц. В литературе, обсуждающей данные экспериментальных наблюдений процессов самосборки, встречаются утверждения о том, что на поверхности капсидных белков имеются так называемые сайты прикрепления или связывания, имеющие вполне определённую дислокацию и достаточно сложное устройство и функции. При этом следует учесть, что вирион имеет высокую устойчивость к воздействию радиации, высоких и низких температур и к разрушающему действию химических растворителей.

Важнейшей функцией капсида является распознавание определённых клеток, прикрепление вириона к поверхности клетки и инъецирование вирусного генома внутрь клетки. После введения вирусного генома происходит перестройка биосинтетических машин заражённой клетки на производство вирусных белков и копий вирусных нуклеиновых кислот. Одним из проявлений этой активности является производство капсидных белков и запуск процесса самосборки вирионов.

Геометрические формы вирусных капсидов сводятся к нескольким основным типам: палочковидные, нитевидные, сферические и сочетающие сферу с палочковидным или нитевидным отростком. Линейные размеры вирионов варьируют от 20 до 300 нм. В среднем вирусы в 50 раз меньше бактерий.

1.1. Капсид вируса табачной мозаики

Простейшей геометрией обладает вирус табачной мозаики (ВТМ). ВТМ относится к РНК-содержащим вирусам. Геном ВТМ насчитывает 6300-6500 пар оснований и кодирует один капсидный белок, который содержит 158 аминокислотных остатков. Капсид ВТМ имеет палочковидную форму, что зафиксировано на многочисленных микрофотографиях, например, приведенных на рис. 3. Непосредственное наблюдение вирусных частиц невозможно в оптической микроскопии, поскольку размеры вирусов меньше длины световых волн. (Длины волн света в видимой области спектра лежат в диапа-

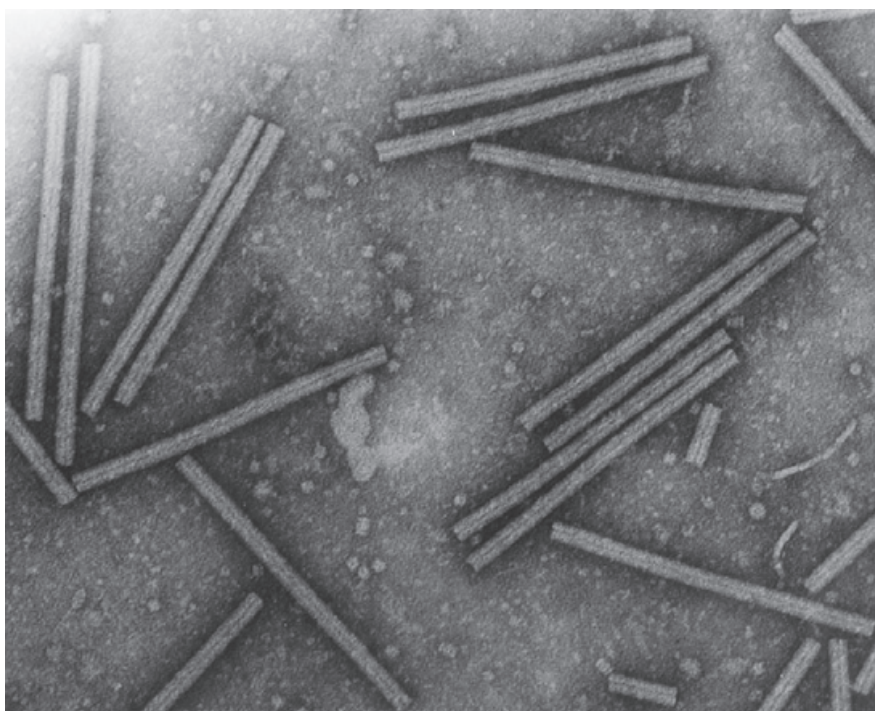
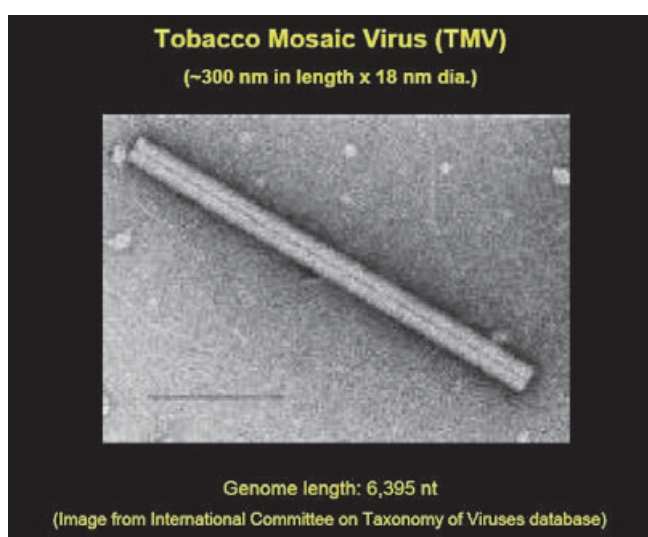


Рис. 3. Микрофотографии вируса табачной мозаики

зоне от 380 нм до 750 нм). Получение изображений вирусных частиц стало возможным только с появлением техники сканирующей электронной микроскопии.

При более детальном рассмотрении капсид ВТМ представляет собой микротрубку, сформированную в результате круговой спиральной укладки множества однотипных молекул капсидного белка, имеющих симметричную форму, близкую к эллипсоиду. Детальная структура капсида ВТМ в схематическом представлении изображена на рис. 4.

Общая длина капсида ВТМ составляет 300 нм, внешний диаметр трубки 18 нм, диаметр внутреннего канала 4 нм. В один виток спирали укладывается 16 и 1/3 молекулы капсидного белка, в целом капсид ВТМ содержит 130.4 витков спирали, в образовании капсида участвует 2130 молекул. Капсидный белок ВТМ содержит 158 аминокислотных остатков.

Важно отметить, что приведенные метрические характеристики капсида носят строго детерминированный характер и устойчиво сохраняются. Во всяком случае, мы не встретили в литературе и при личном общении каких-либо упоминаний о наблюдении отклонений от этой нормы.

Молекула капсидного белка ВТМ имеет на одном из концов вдавленную ложбинку, которая используется для укладки вирусной РНК. При сборке капсида ложбинки образуют внутренний спиралевидный канал, в котором располагается вирусная РНК, зажата прилегающими слоями спиральной укладки. Таким образом, капсид ВТМ обеспечивает эффективную спиральную укладку и надёжное закрепление вирусной РНК в своей внутренней полости (см. рис. 5).

Такая регулярная укладка множества однотипных молекул в монослойный капсид наблюдается только у простейших вирусов. Сложность устройства и размер капсида пропорциональны сложности и размеру вирусного генома, который может варьировать по объёму от сотен до сотен тысяч пар оснований и может кодировать от одного до нескольких сотен белков. Так, например, бактериофаг Т4 содержит в своём геноме около 280 генов, в то время как геном сателлитного вируса некроза табака (СВНТ) кодирует

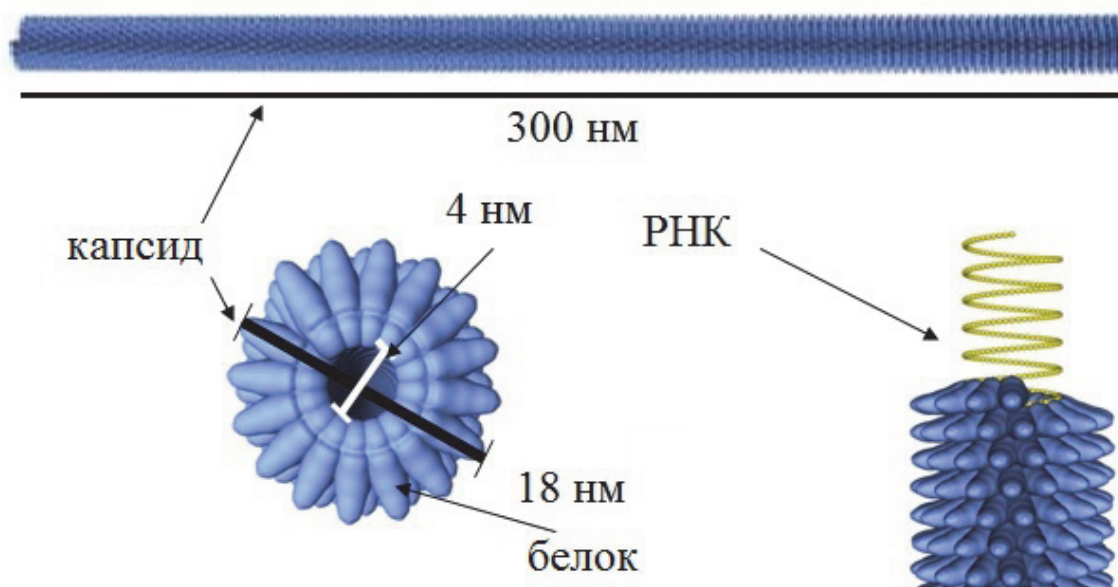


Рис. 4. Структура капсида вируса табачной мозаики

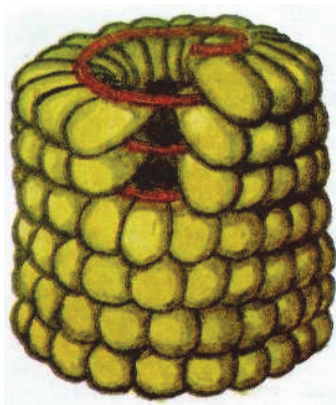


Рис. 5. Иллюстрация принципа укладки РНК в капсиде ВТМ

лишь один белок. (Данные приведены в [15]). Капсидные белки, в свою очередь, могут содержать в первичной структуре от полутора сотен до полутора тысяч аминокислотных остатков.

Как правило, при построении вирусных капсидов наблюдается постадийная агрегация капсидных белков в субблоки или капсомеры, которые представляют собой ассоциации от нескольких до нескольких десятков белков. Это могут быть кольца, усики или относительно плоские многоугольники. Капсомеры не всегда однородны по составу и могут агрегироваться, как из одного, так и из нескольких типов различных белков. Далее капсомеры ассоциируются в более сложные агрегаты, из которых в свою очередь строятся более сложные геометрические конструкции. Капсиды некоторых сферических вирусов имеют многослойную структуру и могут содержать внешнюю липидную оболочку, построенную из липидов, заимствованных из клеточной или ядерной мембраны заражённой клетки [15].

1.2. Капсиды сферических вирусов

Самой распространённой формой вирусных капсидов является замкнутая квазисферическая оболочка. Первые данные о форме сферических капсидов были получены в ходе рентгеноструктурного анализа в 50-е годы (см., например, работы Уотсона и Крика [16]). Авторы исходили из предположения о том, что геном вируса слишком мал для фиксации детерминантов, определяющих всю структуру капсида в явном виде. Скорее всего, вирусный геном предопределяет структуру небольшого числа базовых элементов и правил их взаимодействия, которые затем многократно повторяются и образуют регулярные конструкции, обладающие определёнными видами симметрии. Понятно, что для реконструкции данных рентгеноструктурного анализа гипотеза о симметрии имела решающее значение. Эксперимент эту гипотезу полностью подтвердил. Почти все сферические вирусы имеют икосаэдрическую симметрию. На рис. 6 изображена структура икосаэдра и дана его плоская развёртка, состоящая из 20 равносторонних треугольников.

Пространственная форма икосаэдра образуется замкнутым пояском, состоящим из 10 равносторонних треугольников, и двумя крышечками, представляющими собой пятигранные треугольные пирамидки. Таким образом, икосаэдр содержит 20 идентичных граней в виде равносторонних треугольников, имеет 12 вершин и 30 рёбер. Икосаэдр обладает осями симметрии 5, 3 и 2 порядков. Достаточно полный обзор по архитектуре

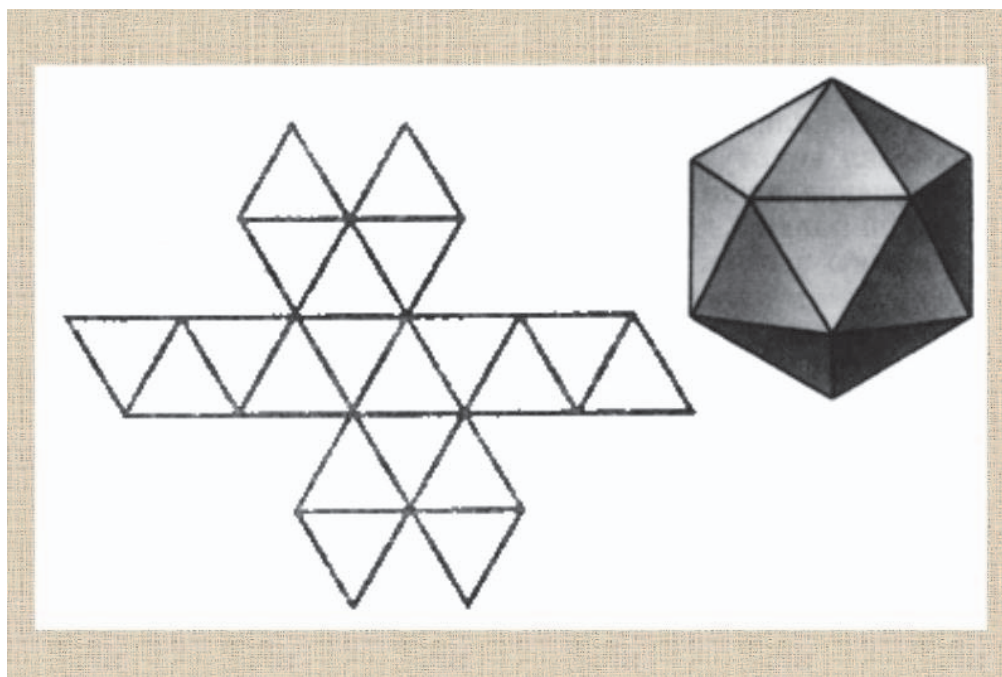


Рис. 6. Структура икосаэдра

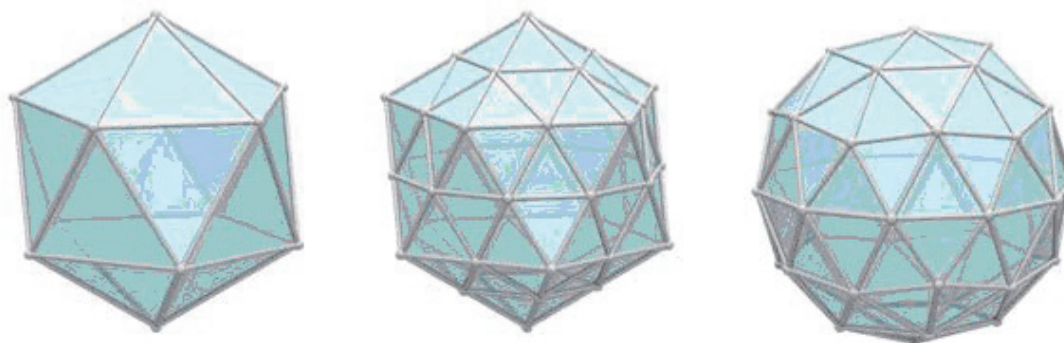


Рис. 7. Примеры построения икосаэдрических многогранников

сферических вирусных капсидов содержится в [15]. В частности, в работе [15] обсуждается проблема так называемой псевдосимметрии. Дело в том, что симметрия с индексом 5:3:2 в классической кристаллографии считается невозможной. Требуемая укладка элементов вирусных капсидов достигается за счёт конформационной подвижности белков, приводящей к частичной деформации граней в соответствующем контекстном окружении.

На рис. 7 приводится ряд примеров построения икосаэдрических многогранников, которые наглядно показывают, как могут строиться вирусные сферические капсиды, обладающие икосаэдрической симметрией.

На рис. 8 приведена подборка структурных реконструкций различных сферических вирусов в диапазоне размеров от 60 нм до 26 нм.

На рис. 9 приведен ряд примеров более детального изображения результатов реконструкции сферических вирусов.

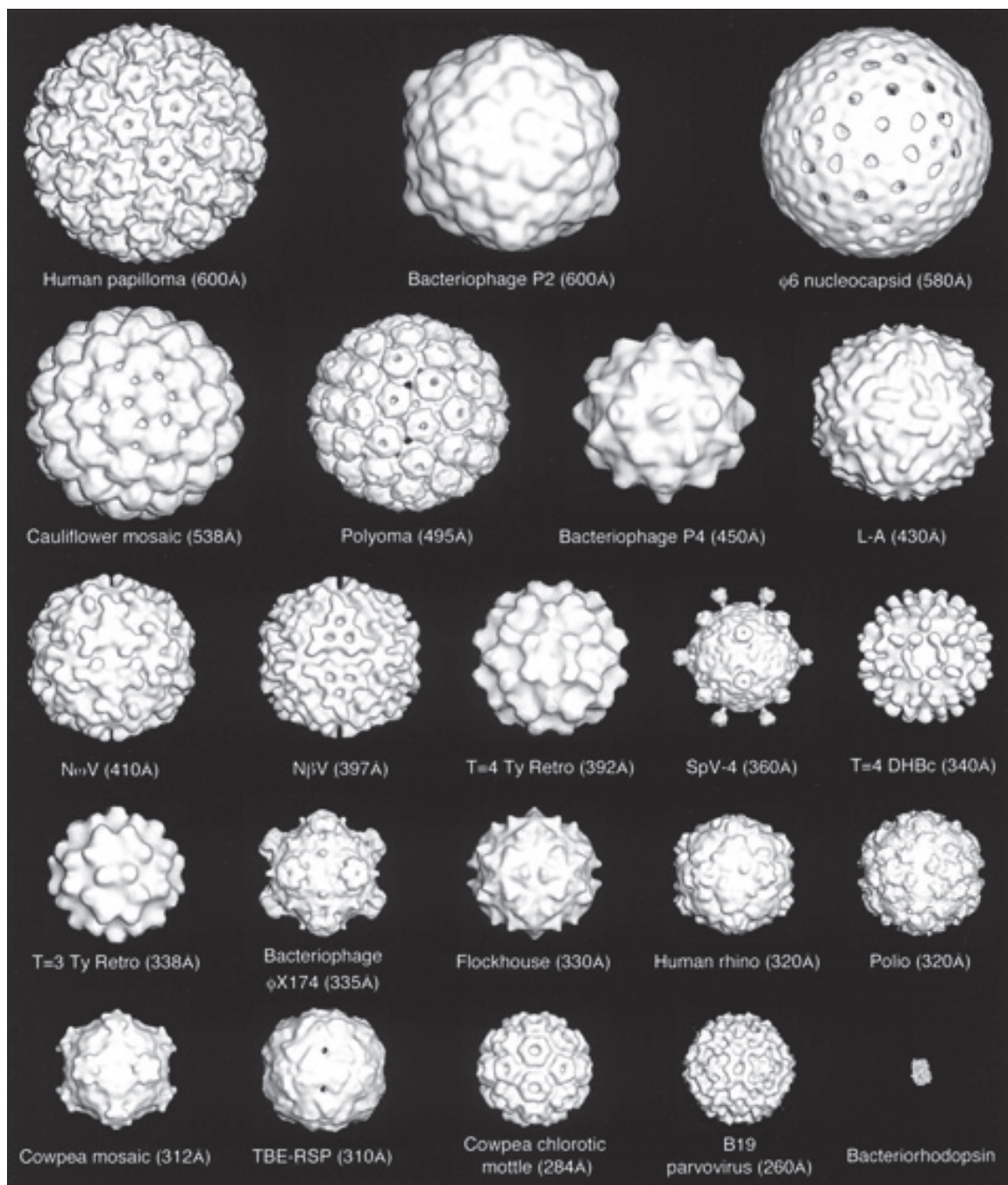


Рис. 8. Результаты структурной реконструкции различных сферических вирусов.

Приведенные капсиды имеют размер от 60 нм до 26 нм. ($1\text{\AA} = 0,1\text{ нм}$)

1.3. Капсид бактериофага T4

Одним из наиболее сложных вирусных капсидов является капсид бактериофага T4, который относится к классу ДНК-содержащих вирусов. Геном фага T4 состоит из 168903 пар оснований и имеет 289 рамок считывания, кодирующих белки. (Данные по составу генома T4 не окончательные и постоянно уточняются и пополняются. В целом геном фага T4 значительно избыточен, содержит множество интронов и далеко не все гены существенны для построения капсида).

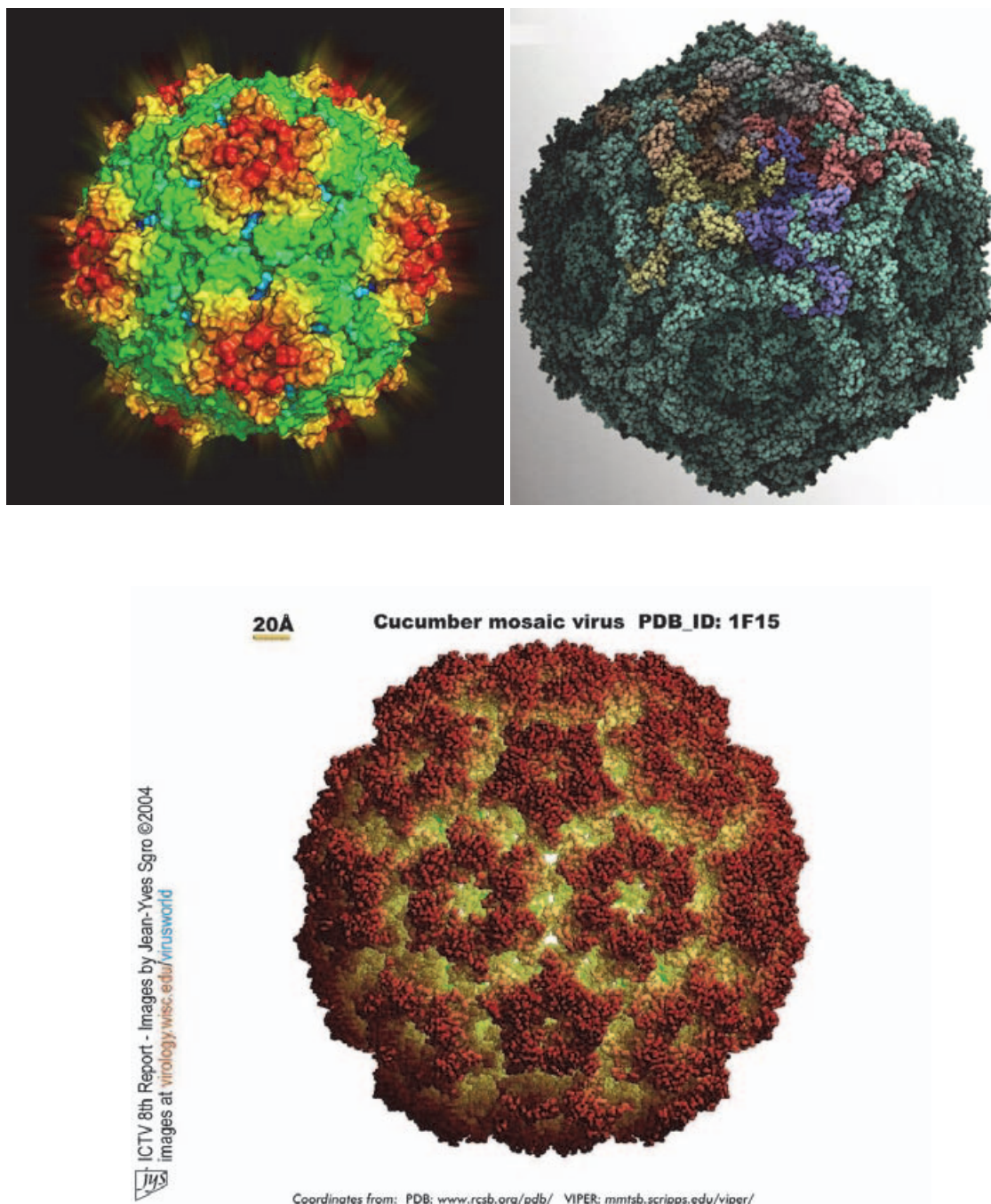


Рис. 9. Примеры более детального изображения результатов реконструкции структуры сферических вирусов

Капсид бактериофага Т4 представляет собой квазисферическую головку и пристыкованный к ней палочковидный хвостовой отросток. Сферическая головка – это двухслойная полость, обладающая икосаэдрической симметрией. Более точная интерпретация предполагает, что головка вируса Т-4 представляет собой удлинённый икосаэдр. В полости головки размещается вирусная ДНК. Хвостовой отросток представляет собой двухслойную трубку, состыкованную с коннектором, который образует достаточно сложную шестигранную платформу со множеством вертикальных стоек и шестью остроконечными опорами или шипами. (В разных источниках коннектор именуют ещё ба-

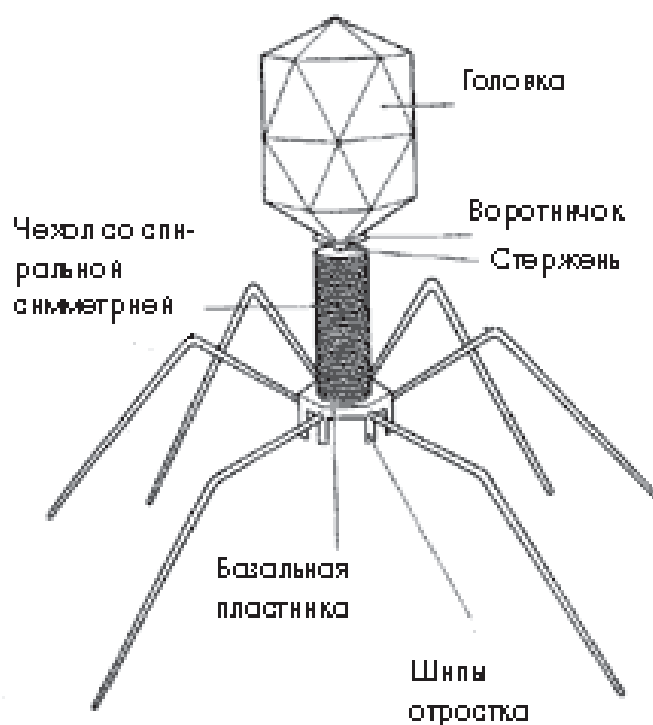


Рис. 10. Общий вид структурной реконструкции бактериофага T4

зальной пластинкой, монтажным столиком, опорной плитой и т. п. При цитировании разных работ мы вынуждены использовать разные обозначения этого элемента.) Кроме того, в состав вирусного капсида входят длинные усики или фибриллы. Всего имеется шесть длинных и шесть коротких фибрилл. Короткие фибриллы крепятся к воротничку по окружности стыковки хвостового отростка со сферической головкой, длинные крепятся у основания коннектора. Общий вид структурной реконструкции бактериофага T4 изображен на рис. 10. Следует отметить, что мы приводим одну из версий реконструкции, в литературе существует множество, несколько отличающихся и постоянно уточняющихся версий.

Жизненный цикл бактериофага T4 аффилирован с бактерией *Escherichia coli* (*E. Coli*, эшерихия коли), именуемой также кишечной палочкой. Это означает, что вирус T4 может осуществлять свои репродуктивные функции только путём проникновения во внутриклеточную среду бактерии *E. Coli*, захвата её биосинтетических машин и перестройки их на производство собственных вирусных белков. В связи с этим функции вирусного капсида T-4 достаточно сложные и разнообразные и не могут быть ограничены только ролью защитной оболочки. По сути вирусный капсид T4 состоит из нескольких молекулярных машин, которые обеспечивают сохранение и транспорт вирусной ДНК, поиск и распознавание клеток бактерии *E. Coli*, абсорбцию и закрепление на поверхности бактерии и, наконец, инъектирование вирусной ДНК и ряда сопутствующих белков внутрь заражённой клетки.

На рис. 11 приведена более детальная схема устройства капсида T4. Общий размер капсида 200 нм, головка 86 нм в ширину и 120 нм в высоту, длина хвостового отростка 92.5 нм, диаметр 24 нм, длина фибрилл 145 нм.

Учитывая, что общая длина вирусной ДНК намного превосходит линейные размеры сферической головки, следует отметить как один из феноменов устройства и функ-

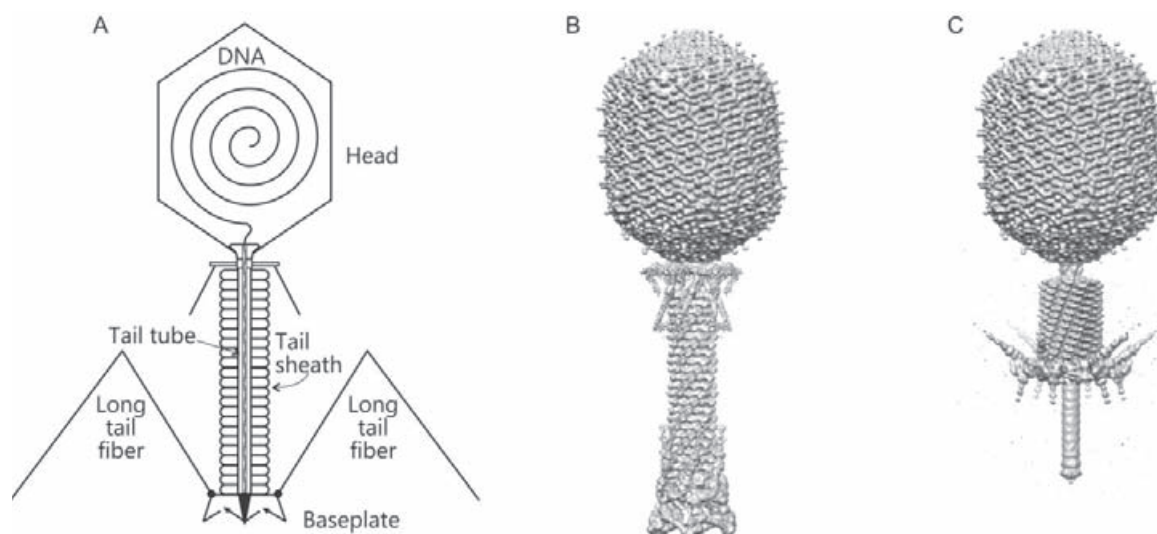


Рис. 11. Детальная структура капсида бактериофага Т4
 (А – продольный разрез и внутренне устройство;
 В – общий вид в начальном состоянии;
 С – общий вид после сжатия мягкого чехла хвостового отростка)

ционирования капсида процедуру укладки вирусной ДНК и её сверхплотной упаковки внутри сферической головки.

На рис. 11А даётся продольный разрез вирусного капсида, на котором видно, что хвостовой отросток состоит из двух трубок – жёсткой внутренней трубки, изображённой как сплошная структура и внешнего мягкого чехла, в виде трубки, образованной отдельными кольцами, собранными в вертикальную стопку. Каждое такое кольцо состоит из 6 молекул белка одного типа, содержащего 659 аминокислотных остатков в первичной структуре. Всего 23 кольца образуют завершённый внешний чехол. Хвостовой отросток крепится к коннектору и располагается на нём как на монтажной платформе. К основанию хвостового отростка крепятся длинные фибриллы, которые выполняют роль рецепторов, осуществляющих распознавание клеток, аффилированных с данным вирусом. Функцию закрепления на поверхности клетки выполняет коннектор. Фибриллы также принимают участие в закреплении на поверхности клетки и последующем инъецировании содержимого головки капсида внутрь клетки.

Иньектирование осуществляется за счёт сжатия мягкого внешнего чехла хвостового отростка. Сжатие осуществляется за счёт того, что молекулы, образующие мягкий чехол, уложены неплотно, и, в целом, чехол действует как растянутая пружина, имеющая запас энергии, достаточный для одного акта сжатия. В случае если коннектор надёжно закреплён на поверхности клеточной мембраны, а жёсткая внутренняя трубка не деформируется, при сжатии внешнего мягкого чехла внутренняя трубка совершает поступательное движение и погружается в толщу клеточной мембраны. Молекулы белка, образующие внешний мягкий чехол, уложены по вертикали не соосно, а со сдвигом и образуют спиральную конфигурацию. Это приводит к тому, что при сжатии мягкого чехла возникает вращательный момент, который передаётся внутренней жёсткой трубке. Жёсткая внутренняя трубка, погружаясь в клеточную мембрану, производит одновременно поступательное и вращательное движение подобно буру и таким образом

обеспечивает сквозной канал внутрь клетки. По этому каналу содержимое головки капсида транспортируется во внутреннюю среду клетки.

На рис. 11В и рис. 11С изображены состояния вирусного капсида до и после сжатия мягкого внешнего чехла хвостового отростка. Таким образом, можно утверждать, что вирусный капсид бактериофага Т4 представляет собой специализированный одноразовый шприц, осуществляющий избирательную инъекцию вирусной ДНК внутрь аффилированной с ним бактериальной клетки.

На рис. 12 приводится более детальная схема устройства и функционирования хвостового отростка Т4. Дается общий план агрегата и срез одного кольца мягкого внешнего чехла. В верхнем ряду зафиксировано исходное состояние, а в нижнем состоянии после сжатия. Белки, составляющие кольца мягкого чехла изображены с полной детализацией укладки полипептидной цепи. В электронной версии текста существует возможность изменения масштаба и более подробного просмотра структуры белка. Поскольку весь мягкий чехол образуют белковые молекулы одного типа, на рисунке применена условная раскраска, позволяющая визуальнo представить спиральную форму, которая образуется в результате сдвига колец при их вертикальной укладке.

До сих пор мы рассматривали графику, изображающую схемы реконструкции внешнего вида вирусных частиц, далее предлагается рассмотреть ряд реальных изображений бактериофага Т4, полученных в технике сканирующей электронной микроскопии. На рис. 13 приводится общий вид бактериофага.

На рис. 14 изображён препарат, содержащий вирусный капсид и молекулу вирусной ДНК, высвобожденную из головки вируса.

На рис. 15 приводится изображение кишечной палочки *E. Coli*, атакованной бактериофагами Т4. Изображение на рис. 15 получено на основе данных сканирующей элек-

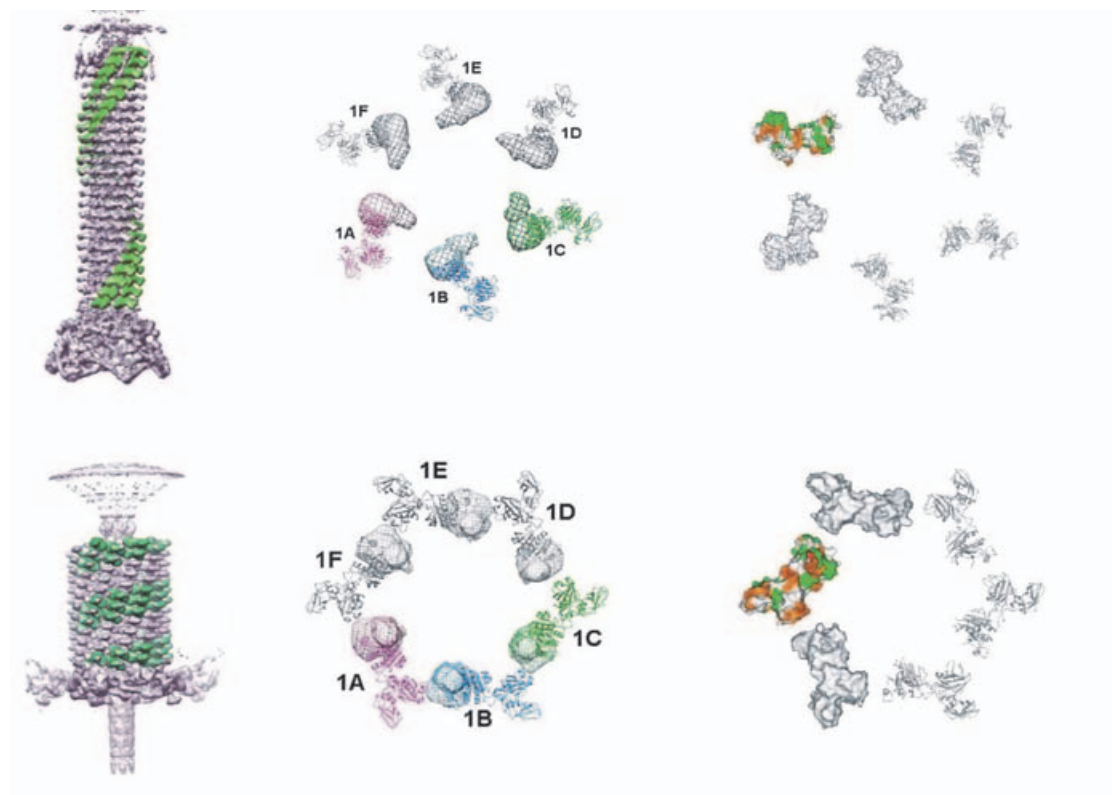


Рис. 12. Детальная схема устройства хвостового отростка бактериофага Т4

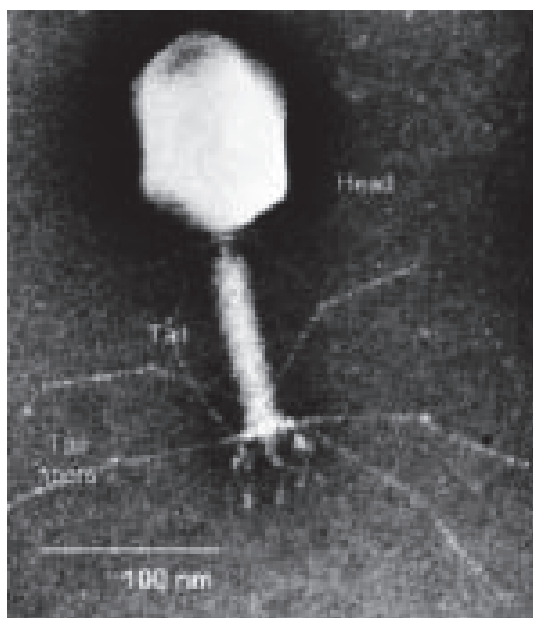


Рис. 13. Изображение бактериофага Т4, полученное на сканирующем электронном микроскопе

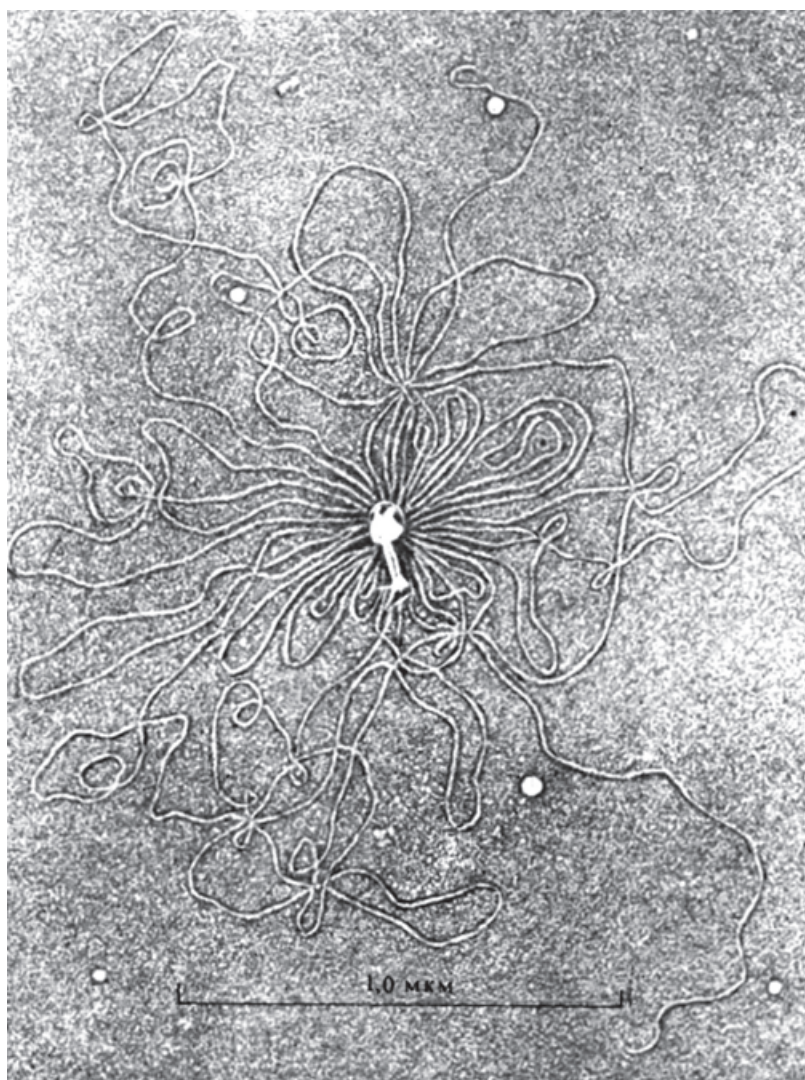


Рис. 14. Вирусная ДНК, высвобожденная из головки капсида бактериофага Т4.

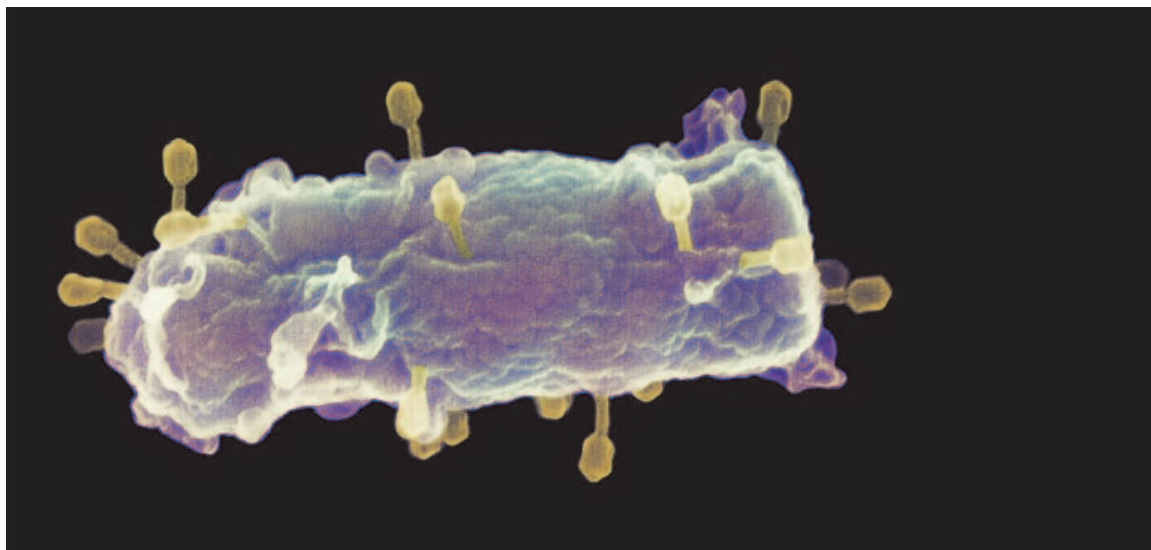


Рис. 15. Бактерия E. Coli, атакованная бактериофагами T4

тронной микроскопии с применением компьютерной технологии условной раскраски и представляет собой прекрасный образец современной техники визуализации объектов микромира.

В результате вирусной атаки внутренняя среда бактерии наполняется многими экземплярами вирусной ДНК и сопутствующих белков, которые осуществляют захват бактериальной системы синтеза белков. Начинается интенсивный процесс экспрессии вирусной ДНК, производства вирусных белков и копирования вирусной ДНК. В результате, спустя чуть более чем 2 часа происходит так называемый лизис или взрыв заражённой клетки и высвобождается более 200 копий бактериофага T4. Причиной лизиса клетки является осмотический шок вследствие резкого изменения pH её внутренней среды.

Здесь важно отметить, что в процессе самосборки вирусных капсидов «выход годных» близок к 100%. В литературе отсутствуют упоминания об обнаружении заметного количества дефектных или недостоенных капсидов.

2. ДАННЫЕ НАБЛЮДЕНИЙ ПРОЦЕССОВ САМОСБОРКИ ВИРУСНЫХ КАПСИДОВ

2.1. Сборка вируса табачной мозаики

В 1955 году Френкель–Конрат и Вильямс опубликовали данные экспериментов по самосборке вируса табачной мозаики (ВТМ) [17]. Никто не верил, что такая сложная структура как вирус табачной мозаики может собраться самостоятельно без участия каких-либо внешних факторов и применения специальных технологий синтеза. Эксперимент был очень прост, смешивались два раствора – раствор капсидного белка ВТМ и раствор вирусной РНК ВТМ. Через некоторое время, без каких-либо дополнительных действий и реагентов в растворе возникали полноценные вирулентные (способные к заражению) экземпляры ВТМ. Эксперимент убедительно демонстрировал явление самосборки. Все необходимые и достаточные факторы формирования ВТМ содержались

в исходных компонентах, а сам процесс формирования вирионов ВТМ демонстрировал полную обратимость, поскольку исходные компоненты были получены в результате диссоциации полноценных вирусов природного происхождения.

Далее следовала умозрительная реконструкция процесса самосборки. Поскольку структура капсида ВТМ была известна как спиральная укладка однотипного белка вдоль цепи РНК, логично было предположить, что при сборке, молекулы капсидного белка последовательно стыкуются друг к другу вдоль цепи РНК наподобие ступеней винтовой лестницы. Такая модель процесса до сих пор ещё излагается в некоторых учебниках. Однако в действительности всё оказалось значительно сложнее. Более полная и правдоподобная картина изложена в нобелевской лекции Клуга (русский перевод опубликован в [18]). Клуг занимался рентгеноструктурным анализом ВТМ и по мере освоения техники высокого разрешения решил исследовать более мелкие, чем завершённый капсид, структуры предшественников. Это оказались кольцевые структуры, содержащие 17 молекул капсидного белка и объединённые в двухслойные диски (всего по 34 молекулы). Более ранние предшественники представляли собой двухэтажные секторальные отрезки.

В результате появилась реконструкция последовательности сборки капсида ВТМ, изображенная на диаграмме рис. 16.

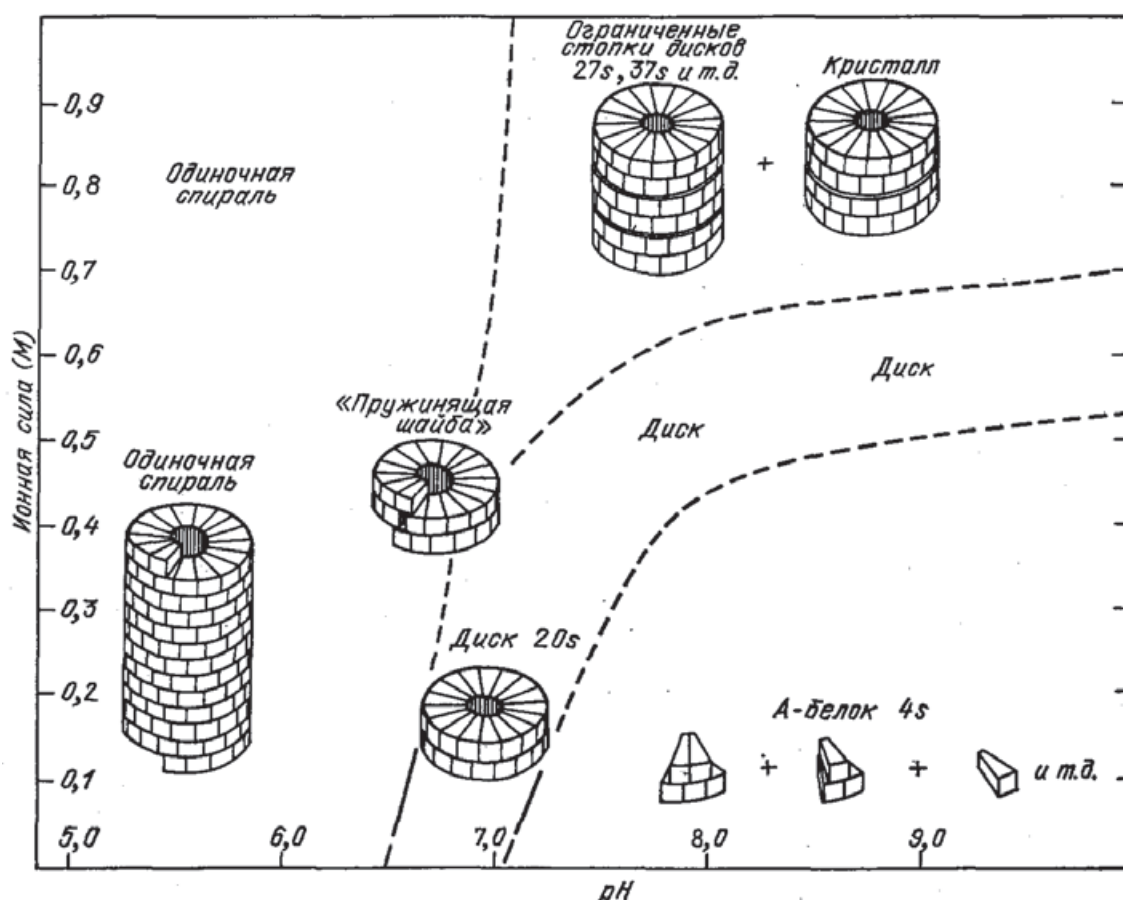


Рис. 16. Последовательность сборки капсида ВТМ.

Фазовая диаграмма содержит распределение различных агрегатов и субагрегатов с привязкой к определённым значениям рН среды

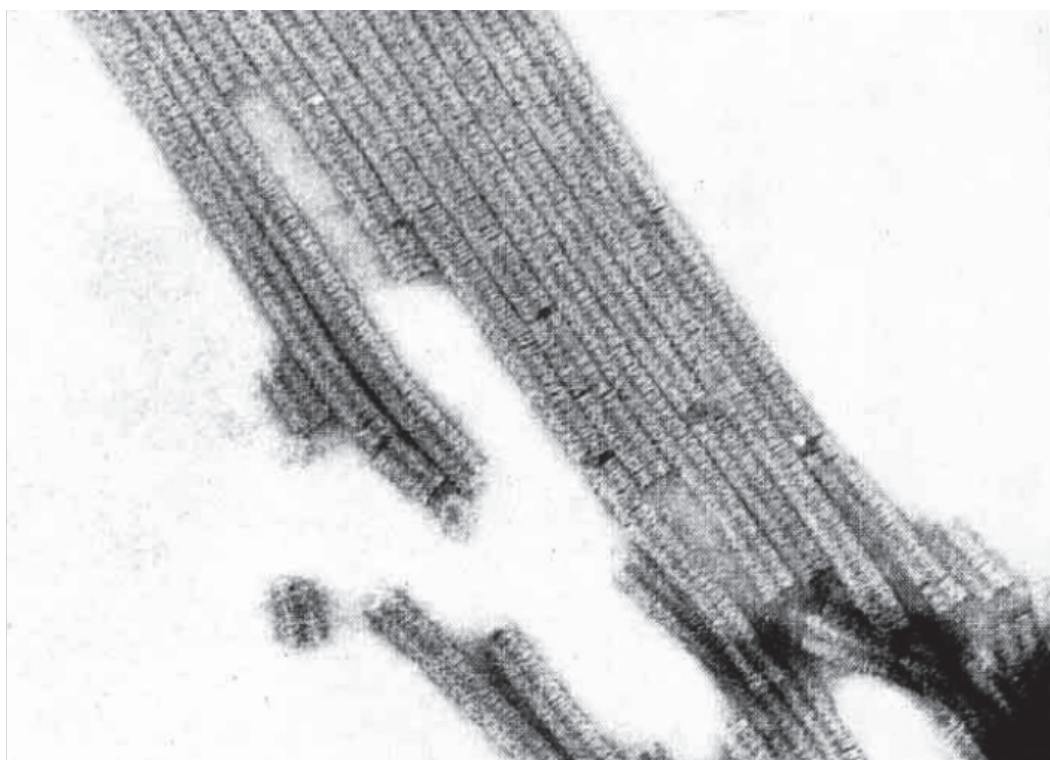
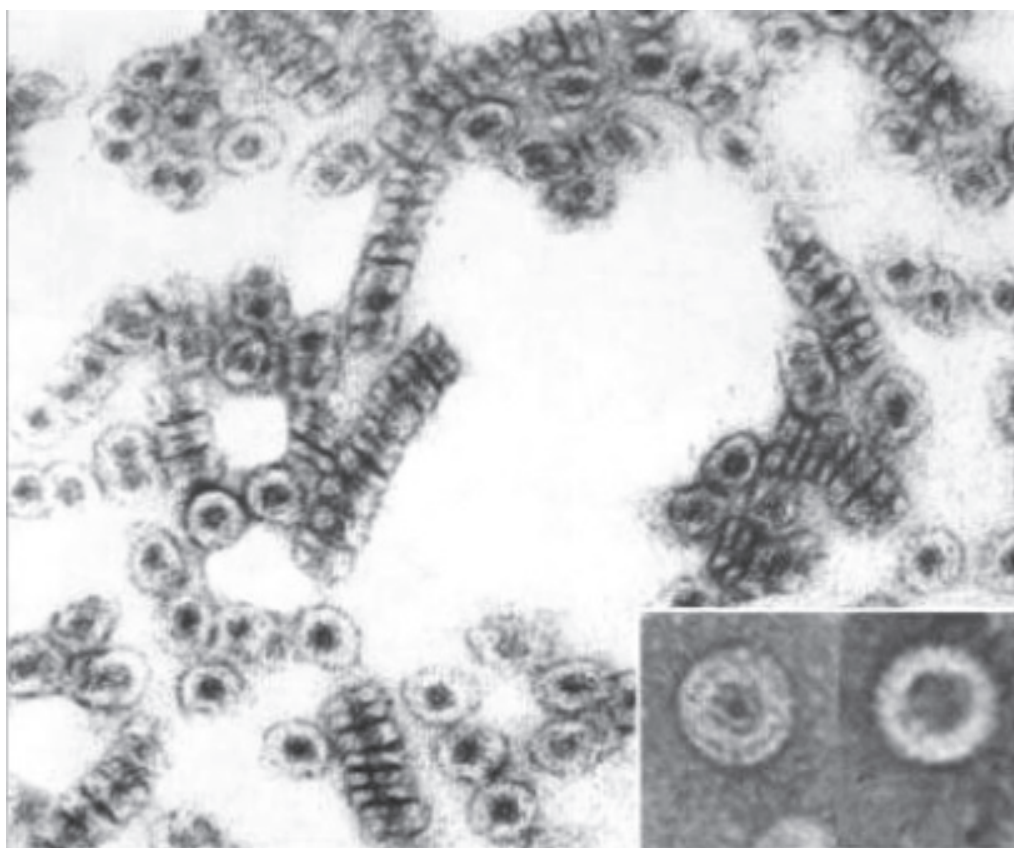


Рис. 17. Результаты спонтанной агрегации капсидных белков в отсутствии РНК

После спонтанного образования двухслойных дисков, в котором участвуют только молекулы капсидного белка, дальнейшая сборка ВТМ существенно зависит от участия вирусной РНК. В отсутствие РНК диски могут собираться в стопки и образовывать микротрубки произвольной длины. На рис. 17 приведены изображения результатов спонтанной агрегации капсидных белков вируса штриховатой мозаики ячменя, который имеет палочковидную форму, во многом подобную капсиду ВТМ.

Взаимодействие двухслойного диска с вирусной РНК приводит к его трансформации в два витка спирали как показано на рис. 18.

Взаимодействие диска с РНК носит кодоспецифический характер. На поверхности диска имеется возможность наложения по окружности достаточно длинного участка РНК для того, чтобы произошло узнавание. Только после этого происходит процедура трансформации диска в спираль. Собранный таким образом комплекс является функционально полным «зародышем» капсида. Далее осуществляется рост путём многократного повторения описанной процедуры, когда к растущей структуре пристыкуются новые диски. Процедура роста изображена на рис. 19.

Проблема заключается в том, что маркерный участок, по которому происходит опознавание, находится в средней части цепи РНК. Это приводит к тому, что РНК должна складываться пополам и продеваться во внутреннее отверстие диска в виде шпильки, что изображено на рис. 19. В ходе роста капсида РНК должна протягиваться дальше, подставляя маркерный участок в соприкосновение со следующим диском. Эта гипотеза

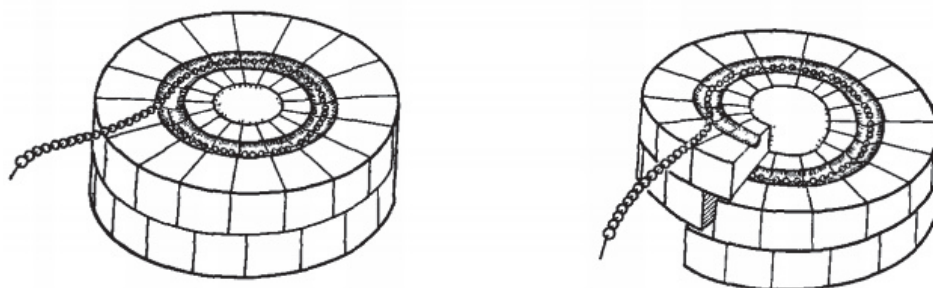


Рис. 18. Взаимодействие двухслойного диска с РНК

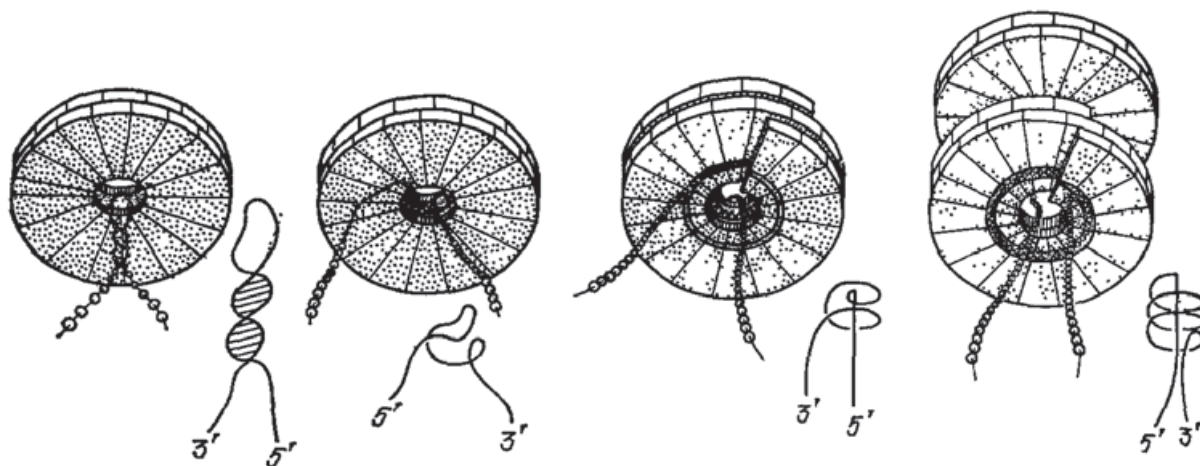


Рис. 19. Зарождение исходного элемента и последующая сборка вируса происходит благодаря проникновению шпильки РНК

получила надёжное подтверждение по данным микроскопии. При получении изображений ранних фаз роста капсида ВТМ отчётливо наблюдались два отростка цепи РНК, торчащие из одного конца недостроенного капсида. В целом длина капсида регулируется размером РНК.

В процессе сборки ВТМ различаются две принципиально различные стадии. Первая стадия – это спонтанная сборка двухслойных дисков, которая протекает в соответствии с определением процесса самосборки т. е. управляется и регулируется поведенческой активностью капсидных белков. Капсидный белок ВТМ это очень небольшая полипептидная цепь, содержащая всего 158 аминокислот. По-видимому, этим объясняются её ограниченные возможности управления и регуляции. На первой стадии формируются исходные капсомеры для дальнейшей сборки. Двойные диски самостоятельно могут только неуправляемо агрегироваться в микротрубки произвольной длины, что не соответствует нужному сценарию сборки капсида ВТМ. Вторая стадия сборки капсида ВТМ протекает в присутствии и под управлением вирусной РНК. РНК выполняет роль направляющей матрицы, на которой происходит трансформация диска в спираль и регулируемая по длине сборка полноценного капсида. Таким образом, в процессе сборки капсида ВТМ наблюдается сосуществование и взаимодействие двух принципиально различных методов управления процессом сборки.

2.2. Сборка бактериофага Т4

Бактериофаг Т4 аффилирован с бактерией *Escherichia coli*, которая является эталонным объектом исследований в области молекулярной генетики. Подавляющее большинство экспериментальных данных по молекулярной генетике были получены в ходе исследований бактерии *E. Coli*. (см., например, [19]). В связи с этим бактериофаг Т4 оказался вовлечённым в интенсивный поток экспериментальных исследований и в настоящий момент является одним из самых изученных вирусов. Геном бактериофага Т4 полностью картирован, почти все вирусные белки идентифицированы, а в области экспериментальных наблюдений процессов самосборки вирусного капсида имеется самая полная и достоверная картина. По данному вопросу существует обширная литература, как в периодических изданиях, так и в монографиях и учебниках. При этом картина экспериментальных данных очень динамична, в высоком темпе публикуются новые факты, быстро заполняются пробелы, меняются и уточняются спорные интерпретации. В нашем изложении мы будем опираться как на классические источники (например, [3, 19]), так и на относительно недавние обзоры [20, 21, 22].

В последнее время при публикации данных по исследованию самосборки успешно применяются современные достижения компьютерной графики, публикуются красочные анимации. Мы считаем полезным дать электронную сетевую ссылку на некоторые анимационные сюжеты [23]. Изготовителями приведенных сюжетов являются Department of Biological Sciences, Lilly Hall of Life Sciences, Purdue University.

Следует отметить, что среди специалистов отношение к анимации неоднозначное, поскольку она воздействует на воображение и может способствовать стиранию граней между достоверными и воображаемыми представлениями. К анимации следует относиться критически и не поддаваться иллюзии несуществующей достоверности. Анимация отличается тем, что вынуждает авторов изобразить в явном виде то, что в тексте маскируется по умолчанию. Так что при критическом восприятии анимация даёт гораздо больше поводов для оппонирования авторам, чем текстовое изложение.

Общий сценарий сборки капсида бактериофага Т4 распадается на три независимые и полностью обособленные ветви, продукты которых затем в определённом порядке агрегируются в полноценный вирион. Общая схема сценария сборки Т4 приведена на рис. 20. Цифрами на схеме обозначены гены, кодирующие белки, участвующие в сборке соответствующих субагрегатов. Напоминаем, что приведенная схема упрощенная и неполная.

На рис. 21 приводится более детальная схема сценария сборки бактериофага Т4, позаимствованная из обзора [22]. Мы приводим обе схемы и не считаем это избыточным, поскольку обе схемы носят взаимодополняющий характер. Схема, приведенная на рис. 20 более грубая и упрощенная, но даёт чёткое общее представление о процессе в целом. Схема на рис. 21 более громоздкая, но полезная в том отношении, что демонстрирует большую полноту достигнутых представлений о процессах самосборки капсида Т4.

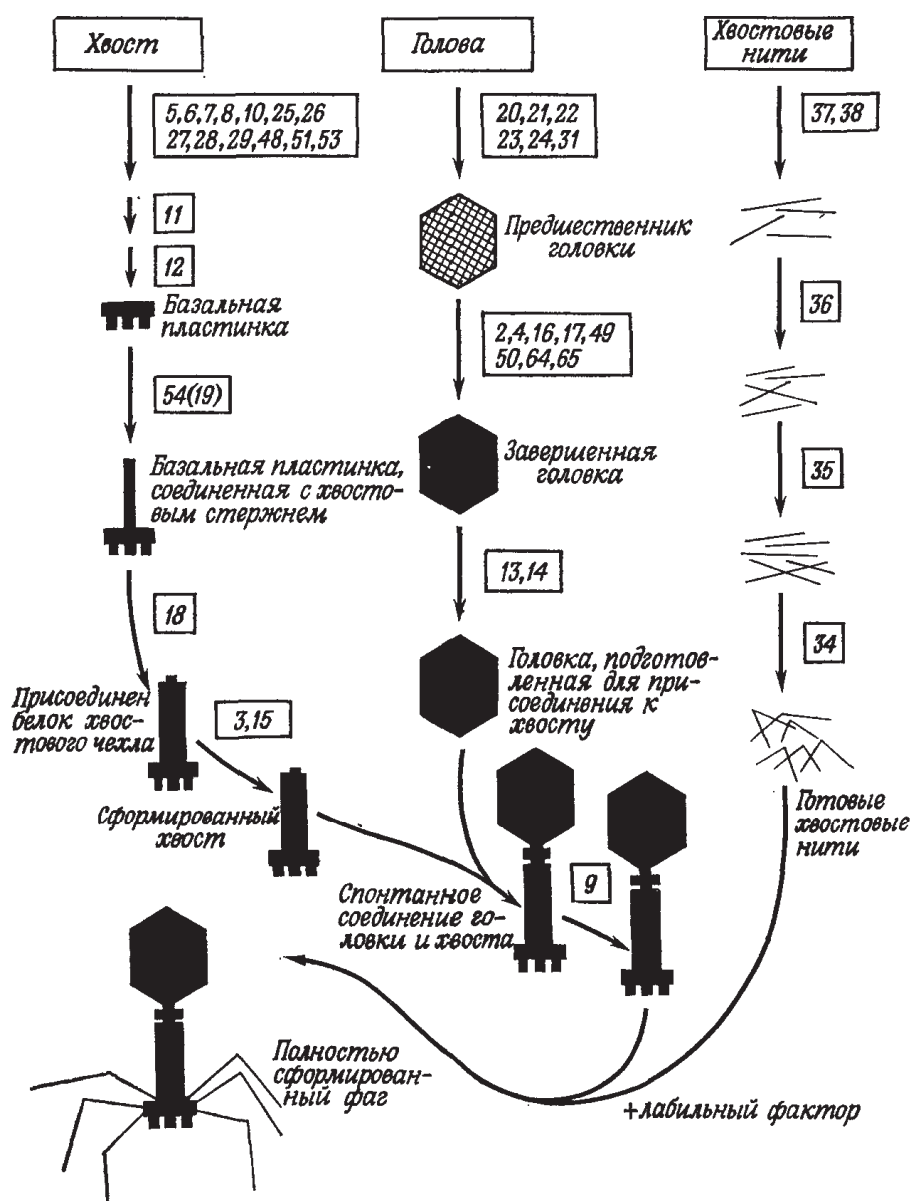


Рис. 20. Общая схема сценария сборки бактериофага Т4 (цифры обозначают белки, участвующие в сборке)

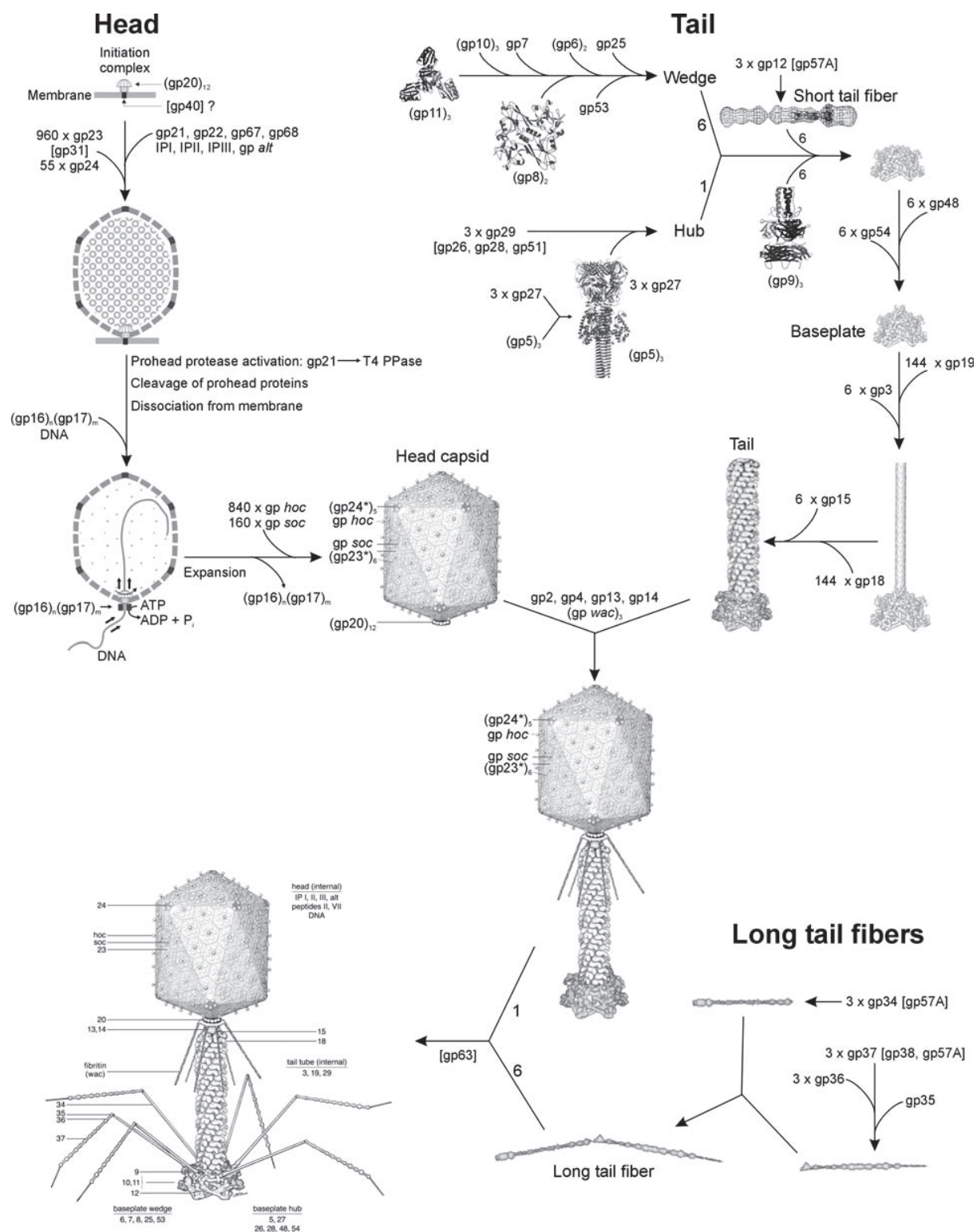


Рис. 21. Более детальная схема сценария сборки бактериофага Т4 (Схема заимствована из обзора [22], схема сопровождается обозначениями белков, участвующих в сборке соответствующих субагрегатов)

По независимым ветвям сценария происходит самосборка сферической головки, хвостового отростка и фибрилл. Экспериментально проверено, что каждая из этих трёх ветвей может осуществляться независимо от других. Более того, полноценный вирион может быть собран на последней стадии агрегации из головки, хвостового отростка и фибрилл, полученных в результате экспрессии генома разных экземпляров фага Т4. В ранних экспериментах независимая сборка фибрилл и хвостового отростка были осуществлены *in vitro* при наличии полного набора капсидных белков, входящих в данные агрегаты. Трудности возникли при попытке собрать *in vitro* сферическую головку. Белки, составляющие головку капсида, в пробирке не собирались в замкнутый икосаэдр, а образовывали отдельные плёнки [19]. Позднее выяснилось, что в сборке головки участвует вспомогательный белок, из которого сначала собирается выпуклый предшественник головки, так называемый прокапсид. Прокапсид служит матрицей, на которой собирается первый слой головки. По завершении сборки первого слоя головки прокапсид уничтожается.

Сферическая головка в собранном виде не является полностью замкнутой симметричной статической конструкцией наподобие кристалла. В основании головки отсутствует один капсомер, и это отверстие – функционально необходимая часть конструкции. Отверстие используется сначала для подачи и укладки во внутренней полости головки вирусной ДНК, а затем для стыковки с хвостовым отростком.

После завершения сборки первого слоя головки собирается звёздообразная ассоциация белков, которая является молекулярной машиной, выполняющей ряд необходимых функций. (Эта звёздообразная ассоциация напоминает известные структуры, называемые шаперонами и ответственные за репарацию повреждений полипептидов.) Детальное описание упомянутой ассоциации содержится в обзоре [22]. Машина осуществляет поиск вирусной ДНК, определяет её начальный конец, прикрепляется к нему и транспортирует ДНК к отверстию сферической головки. Машина закрепляется у основания отверстия сферической головки и приступает к операции проталкивания ДНК внутрь сферической полости. По-видимому, машина работает согласованно с другими белками, находящимися во внутренней полости головки и осуществляющими уникальную сверхплотную укладку ДНК. Операция проталкивания требует энергозатрат, поэтому машина осуществляет энергетическую трансформацию, потребляя и расщепляя АТФ. Описанная процедура сборки сферической головки и её заполнения вирусной ДНК наглядно показана в анимации [23].

После завершения описанных процедур заполнения внутренней полости сферической головки запускается процедура построения второго слоя головки, придающего ей дополнительную прочность. На рис. 22, приводятся отдельные кадры из анимационного фильма, иллюстрирующие определённые этапы сборки сферической головки фага Т4 и его заполнение вирусной ДНК.

Опубликованный в 2010 году фундаментальный обзор [20] суммирует множество отдельных данных по сборке хвостового отростка бактериофага Т4. Обзор представляет впечатляющую целостную и обоснованную картину структуры, функций и динамики построения одного из важнейших элементов сложного вирусного капсида. Данные, приведенные в обзоре, обладают полнотой взаимосвязей явлений и глубокой детализацией структурной организации компонентов носителей распределённого управления самосборкой. Важнейшая особенность данного обзора состоит в том, что вирусный капсид рассматривается как молекулярная машина, объединяющая совокупность по-

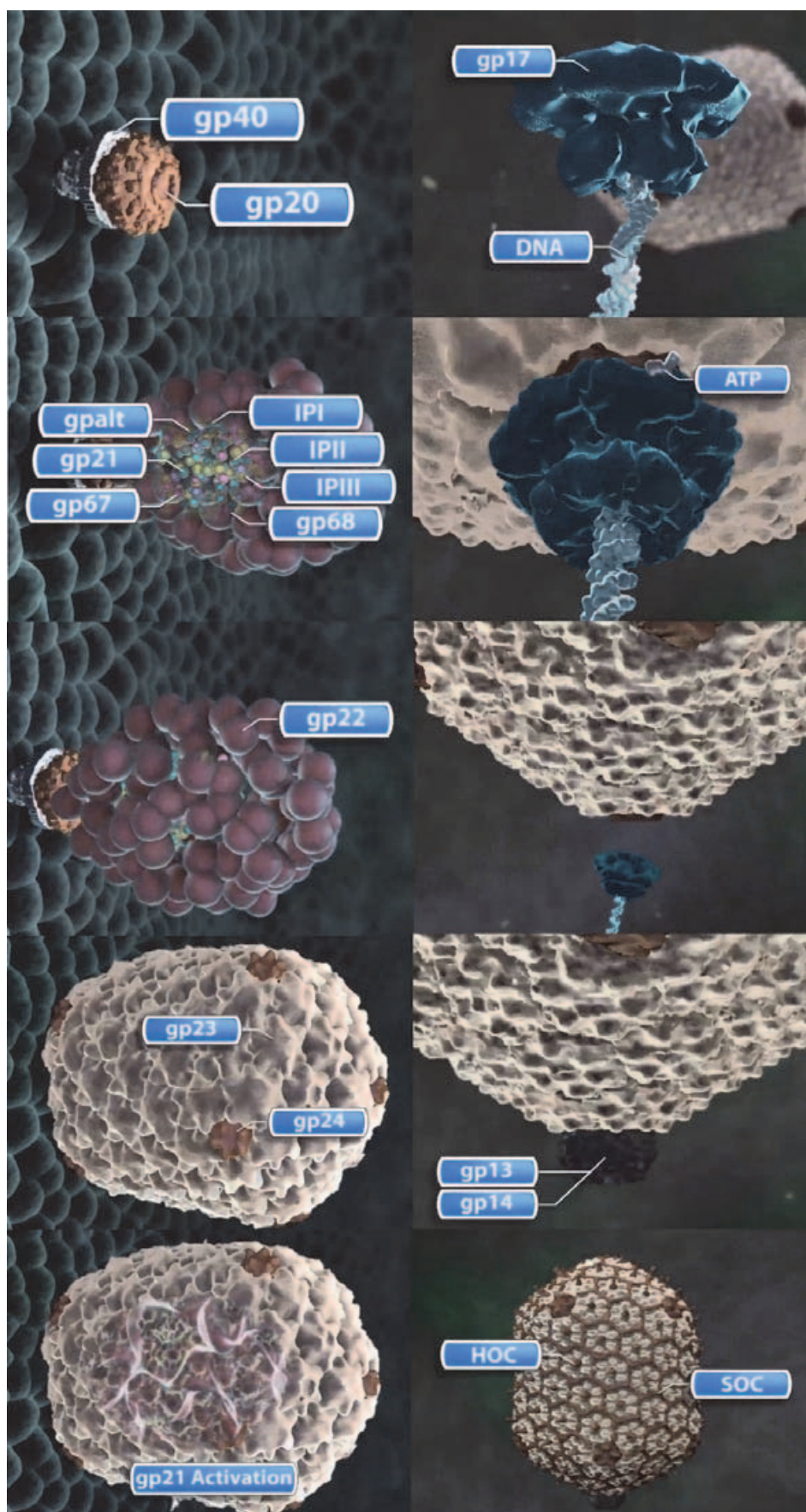


Рис. 22. Скриншоты из анимационного фильма, демонстрирующие сборку сферической головки фага T4

движных механизмов, обладающих алгоритмической динамикой. Материалы данного обзора могут быть использованы как исходные данные для построения математических моделей самосборки. При этом имеется возможность отображения в моделях наиболее типичных основополагающих механизмов распределённого управления, поддержания устойчивости и сходимости последовательностей монтажных актов. В обзоре обозначены домены кристаллической структуры белков потенциально ответственные за реализацию сайтов связывания и приводятся соображения по их возможной конформационной подвижности.

В данной статье мы приведём и прокомментируем только некоторые фрагменты из упомянутого обзора, подтверждающие нашу гипотезу об алгоритмической природе процесса самосборки, а также позволяющие выделить основополагающие механизмы управления самосборкой. Мы будем активно использовать данные обзора [20] в последующем изложении при построении моделей самосборки. Далее по тексту следует выборочное цитирование [20].

На рис. 23 изображён общий сценарий сборки хвостового отростка бактериофага Т4 с детализацией по основным субблокам отростка и по каждому белку участнику самосборки. Всего в сборке хвостового отростка фага Т4 участвует 22 гена, а завершённый агрегат состоит из 430 полипептидных цепей или молекул капсидных белков. В целом процесс самосборки хвостового отростка представляет собой строго упорядоченную и жёстко управляемую последовательность процедур. При этом экспериментально доказано, что последовательность сборки определяется порядком следования актов взаимодействия белков, а не порядком экспрессии вирусного генома.

На рисунке представлены три строки **А**, **В** и **С**. В строке **А** изображается последовательность сборки стойки или клина. Всего собирается шесть таких клиньев. В строке **В** представлена процедура сборки коннектора или опорной плиты, которая состоит из шести клиньев и нескольких дополнительных элементов. В строке **С** изображается процесс сборки двухслойной трубки хвостового отростка.

Опорная плита состоит примерно из 140 молекул белков по меньшей мере 16 типов. Продукты двух генов *gp51* и *gp57A* участвуют в сборке, но не присутствуют в завершённой конструкции. Опорная плита в собранном виде является необходимым условием для сборки двухслойной хвостовой трубки. Начало процесса сборки хвостовой трубки возможно только на опорной плите. Белки, образующие хвостовую трубку и прочно связанные между собой в составе завершённой конструкции, в изолированном состоянии и в отсутствии опорной плиты не взаимодействуют между собой. Это касается не только участников сборки хвостовой трубки. Аналогичные эксперименты были проведены с некоторыми изолированными белками, входящими в состав клиньев, образующих опорную плиту. Результаты экспериментов показали отсутствие взаимодействия либо слабые взаимодействия между белками, жёстко связанными в составе завершённой конструкции. Этот факт объясняется тем, что в ходе выполнения сборочных актов в сайтах связывания на поверхности белков происходят конформационные изменения. Кроме того, в ряде случаев в результате выполнения определённых шагов самосборки на поверхности белков образуются новые сайты связывания. Таким образом, экспериментально зафиксирован важнейший факт, свидетельствующий о том, что белки, прошедшие через определённые этапы процесса самосборки не идентичны своим первичным формам.

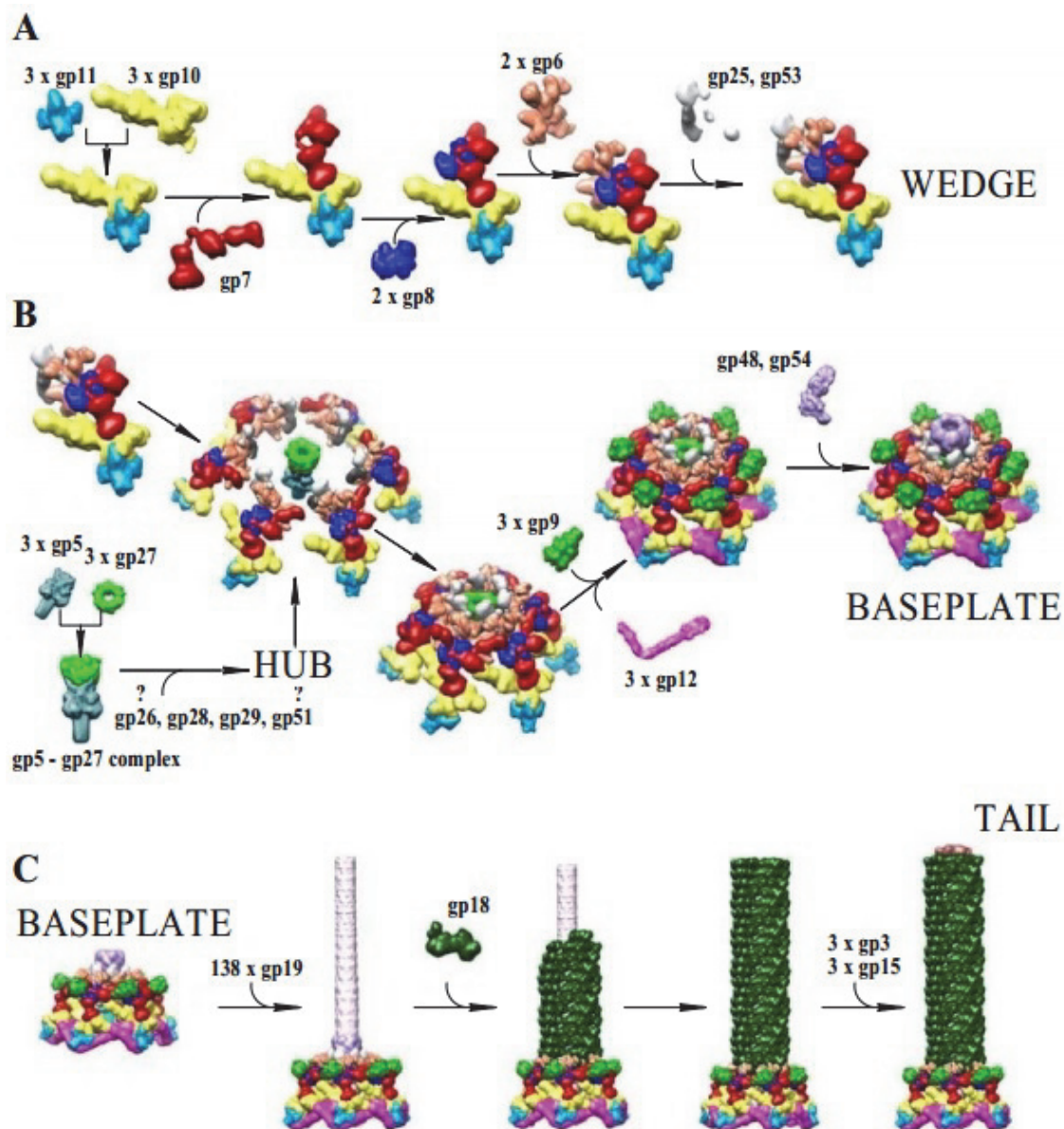


Рис. 23. Сценарий сборки хвостового отростка бактериофага Т-4

В завершающей стадии сборки коннектора (опорной плиты) в центре верхней части купола белки gp48 и gp54 образуют платформу для начала сборки жёсткой внутренней трубки. Внутренняя трубка образуется 138 молекулами белка gp19. Похоже, что трубка представляет собой стопку из 23 колец, а каждое кольцо состоит из 6 молекул gp19. Механизм регуляции длины трубки пока не выяснен. Если верить изображению рис. 22С, сборка внешнего мягкого чехла начинается после завершения сборки жёсткой внутренней трубки. Внешний чехол образуется укладкой белка gp18 в виде вертикальной стопки из 23 колец, каждое из которых состоит из 6 молекул gp18. По завершении сборки внешнего чехла 3 белка gp15 и 3 белка gp3 образуют гексамерное кольцо, которое накрывает последний слой трубки, состоящий из молекул белка gp19 и gp18. Таким образом создаётся активное завершение хвостового отростка в виде переходной муфты, несущей сайты связывания комплиментарные поверхности крепления к сферической головке капсида.

На рис. 24 изображена детализация горизонтальных срезов двухслойной трубки, соответствующих её вытянутому и сокращённому состоянию, а также структура белка gp18, из которого строится внешний мягкий чехол.

На рис. 24 также приводится изображение двух состояний мягкого чехла растянутое и сокращённое, с плотной укладкой белков и разреженной. При этом выясняется на первый взгляд противоречивая ситуация. Если внешний чехол находится в вытянутом состоянии, образующие его молекулы белка gp18 на горизонтальном срезе демонстрируют плотную укладку, а их вертикальные контакты растянуты. При сокращении чехла вертикальные контакты уплотняются, но горизонтальный срез претерпевает утолщение

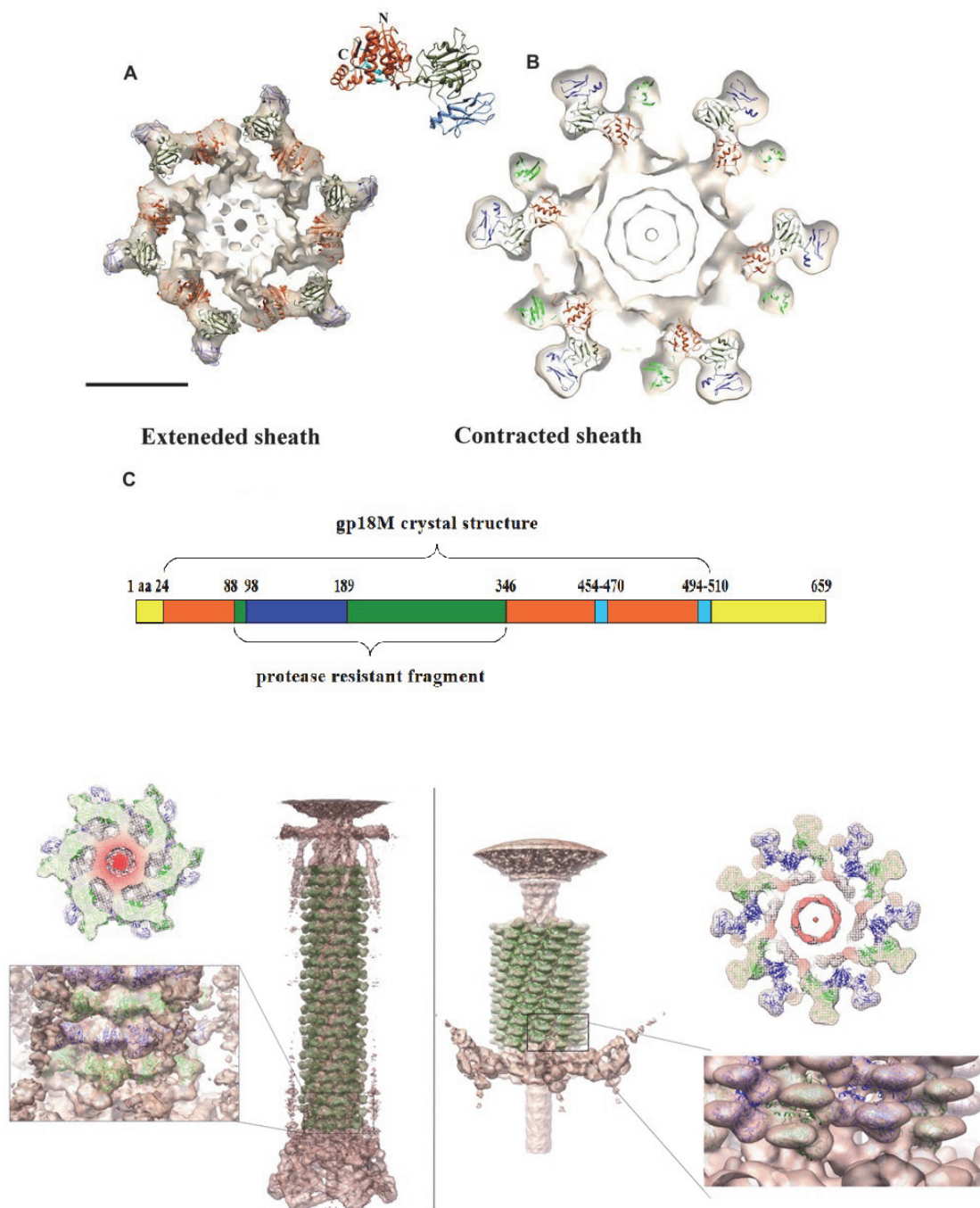


Рис. 24. Горизонтальные срезы двухслойной трубки хвостового отростка в вытянутом и сокращённом состояниях, структура белка gp 18

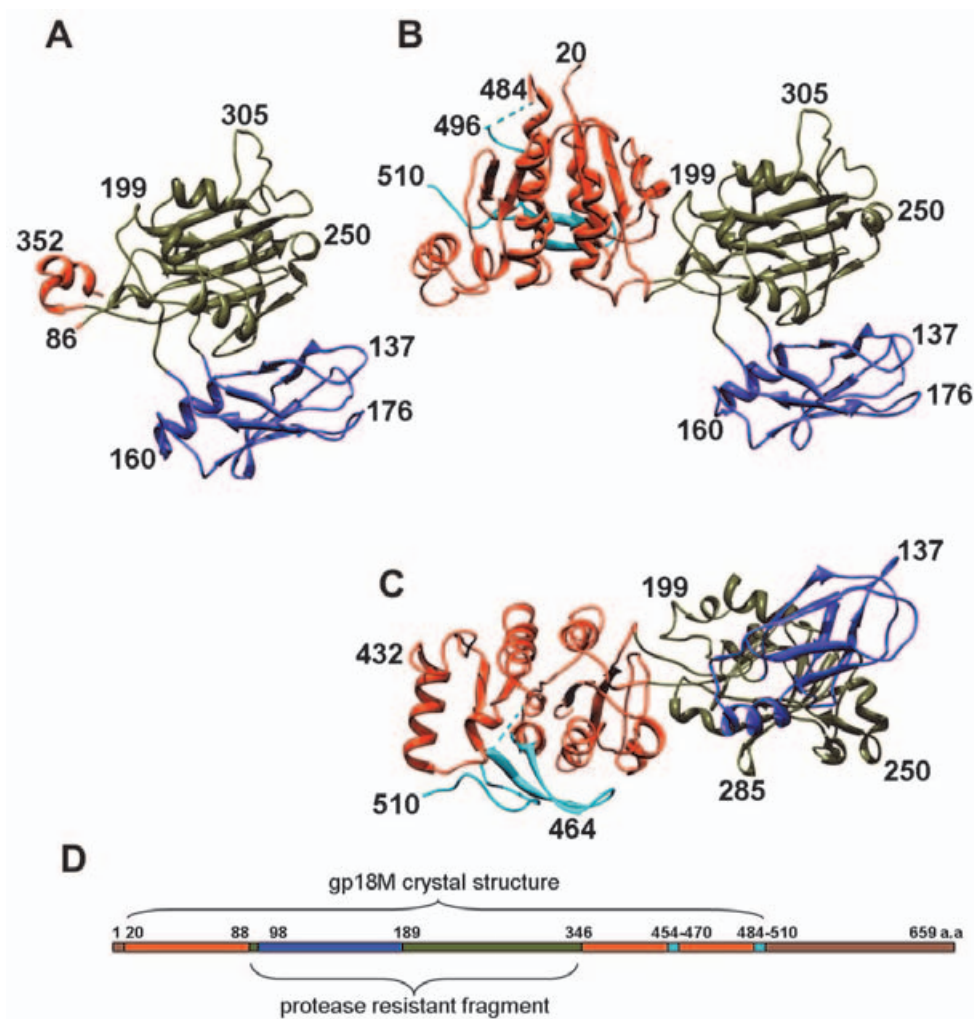


Рис. 25. Детализация структуры белка gp18 и его изображения в разных проекциях

и переходит в состояние разрежённой укладки. Похожим образом ведёт себя спиральная пружина при растяжении и сокращении. В обзоре очень детально излагается механизм сжатия мягкого чехла с привязкой к доменной структуре белка gp18 и указаниями дислокации сайтов связывания.

На рис. 25 приводится детализация структуры белка gp18 и варианты его изображения в разных проекциях.

Мы приводим эти отрывочные фрагменты обзора [20] с целью демонстрации уровня детализации и полноты накопленных экспериментальных данных о строении и функциях конкретного субагрегата фага T4, а именно хвостового отростка.

К основанию хвостового отростка на границе сочленения хвостовой трубки и опорной плиты крепятся длинные фибриллы. Всего насчитывается шесть длинных фибрилл, каждая из которых представляет собой тонкий ус, длиной 144 нм и толщиной 5 нм. Каждый ус разделён на две половинки, сочленённые под углом около 20°. Дальний конец фибриллы содержит молекулярный рецептор, способный распознавать типичные мембранные белки на внешней поверхности бактерии *E. Coli*. Экспериментом установлено, что для надёжного распознавания в составе вириона фага T4 должно быть не ме-

нее трёх фибрилл. В противном случае фаг теряет вирулентность (способность к заражению).

Крепление фибриллы к опорной плите и сочленение двух её половинок представляют собой шарнирные механизмы, обеспечивающие подвижность фибрилл. Кроме распознавания мембранных белков бактерии фибриллы участвуют в механизме закрепления капсида на её поверхности и инъектирования вирусной ДНК внутрь клетки.

На рис. 26 изображена последовательность сборки длинной фибриллы.

В сборке фибриллы участвуют продукты четырёх генов *gp34*, *gp35*, *gp36* и *gp37*. Так называемое «бедро» фибриллы формируется из продуктов гена *gp34*, содержащих 1289 аминокислот. Ближний конец фибриллы, образованный белком *gp34* крепится к белку *gp9* на опорной плите. Так называемое «колено» фибриллы формируется мономером *gp35* (372 аминокислотных остатка) и димером *gp36* (221 аминокислотный остаток). И, наконец, «голень» фибриллы собирается из продуктов гена *gp37*, состоящих из 1026 аминокислот.

На рис. 27 даётся иллюстрация механики функционирования капсида фага Т4 в ходе выполнения процедуры внедрения вирусной ДНК внутрь клетки.

Мы не станем пересказывать подробные описания действий, изображённых на рис. 27, рисунок достаточно наглядный и даёт представление о сути происходящего. Длинные фибриллы кроме распознавания мембранных белков бактерии *E.Coli* устанавливают правильную ориентацию капсида, закрепляются на поверхности бактерии и при-

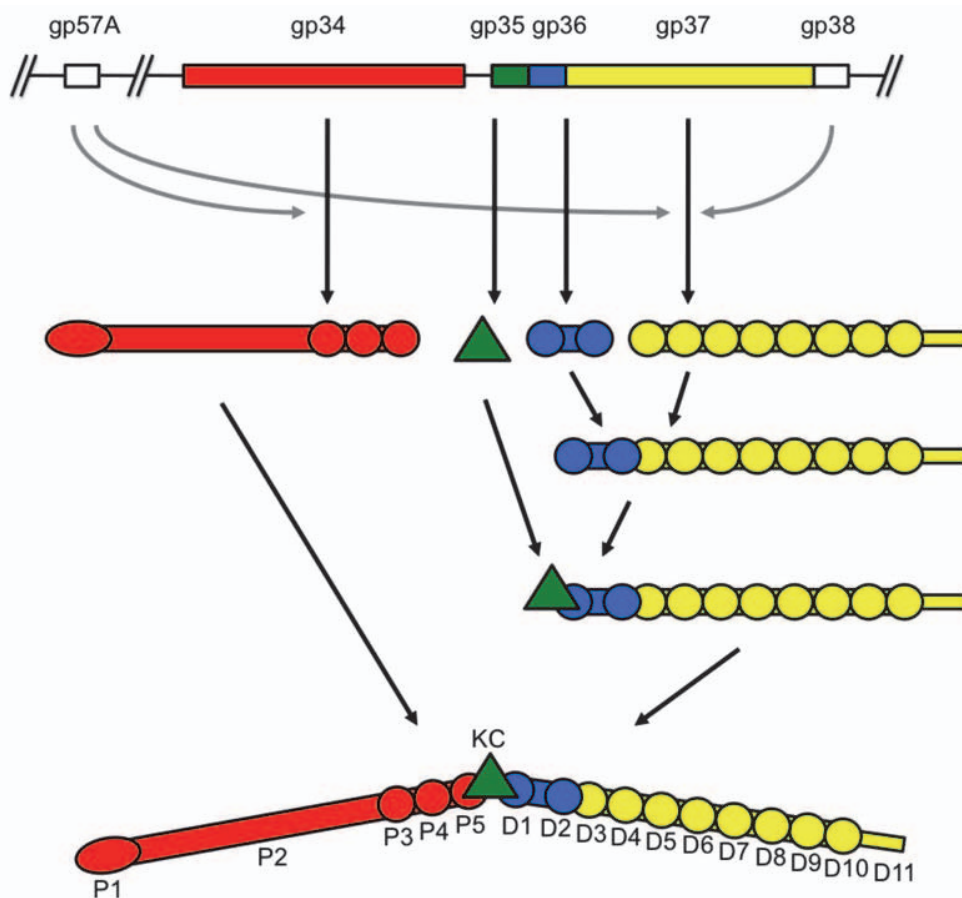


Рис. 26. Последовательность сборки длинной фибриллы

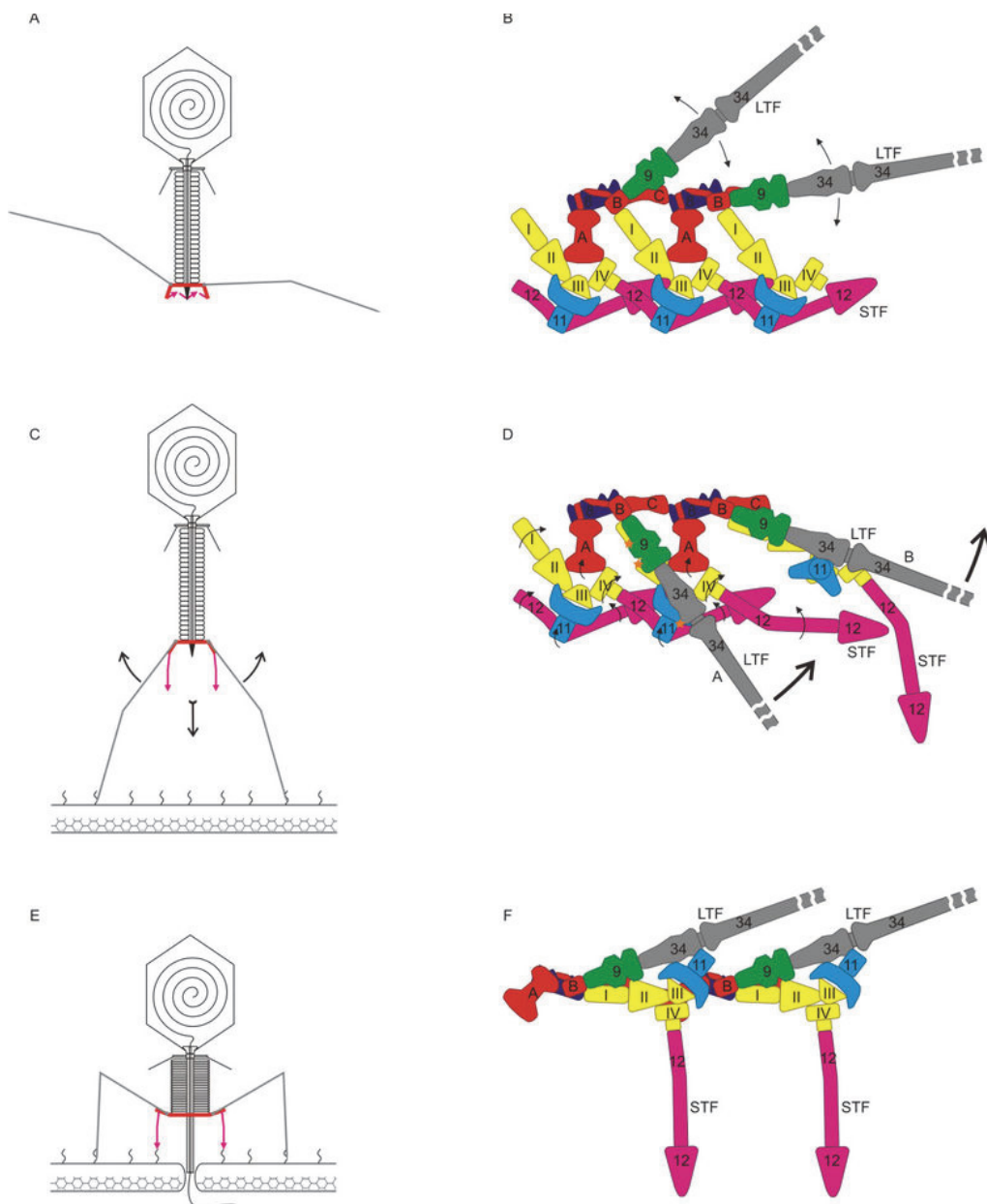


Рис. 27. Диаграмма работы механизма закрепления капсида и инъектирования вирусной ДНК

двигают опорную плиту к мембране, подготавливая условия для её прокалывания хвостовой трубкой. На рис. 27 проиллюстрирована динамика работы сайтов прикрепления фибрилл к опорной плите, которые работают как шарнирные поворотные механизмы. В результате закрепления фибрилл на поверхности бактерии запускается перестройка опорной плиты, в ходе которой происходит выпуск усиков, направленных вниз. Далее работает узел крепления фибрилл к опорной плите и производится поворот «бедра», приводящий к опусканию капсида. В результате соприкосновения усиков плиты с поверхностью мембраны бактерии вырабатывается сигнал, запускающий процедуры сокращения мягкого чехла хвостового отростка.

В обзоре [22] описывается ещё более детальная картина процедуры инъектирования вирусной ДНК внутрь заражаемой бактерии. В качестве иллюстрации мы приведём на рис. 28 диаграмму, позаимствованную из [22].

Для наших целей в данном случае важно отметить, что в ходе экспериментальных наблюдений накоплен значительный объём фактических данных и составлены детальные последовательности взаимосвязанных событий и процедур, установлено понимание важнейших свойств и функций вирусного капсида как совокупности молекулярных машин. Сайты связывания не являются только статическими замками, обеспечивающими жёсткое прикрепление капсидных белков при построении замкнутых поверхностей. Формирование защитных оболочек, это только одно из назначений капсида. Роль сайтов связывания намного сложнее и разнообразнее. После прикрепления капсида к поверхности бактерии опорная плита претерпевает значительные изменения и перестройку геометрии, сложившейся в ходе самосборки. Компактный куполообразный коннектор распрямляется и раскрывается как цветок. Длинные фибриллы разворачиваются, обеспечивая правильную ориентацию капсида и нужное положение опорной плиты для прокалывания мембраны. При этом сайты связывания производят сложные последовательности взаимообусловленных движений, капсидные белки осуществляют определённые конформационные трансформации. И всё это в целом представляет собой алгоритмические формы поведения на молекулярном уровне.

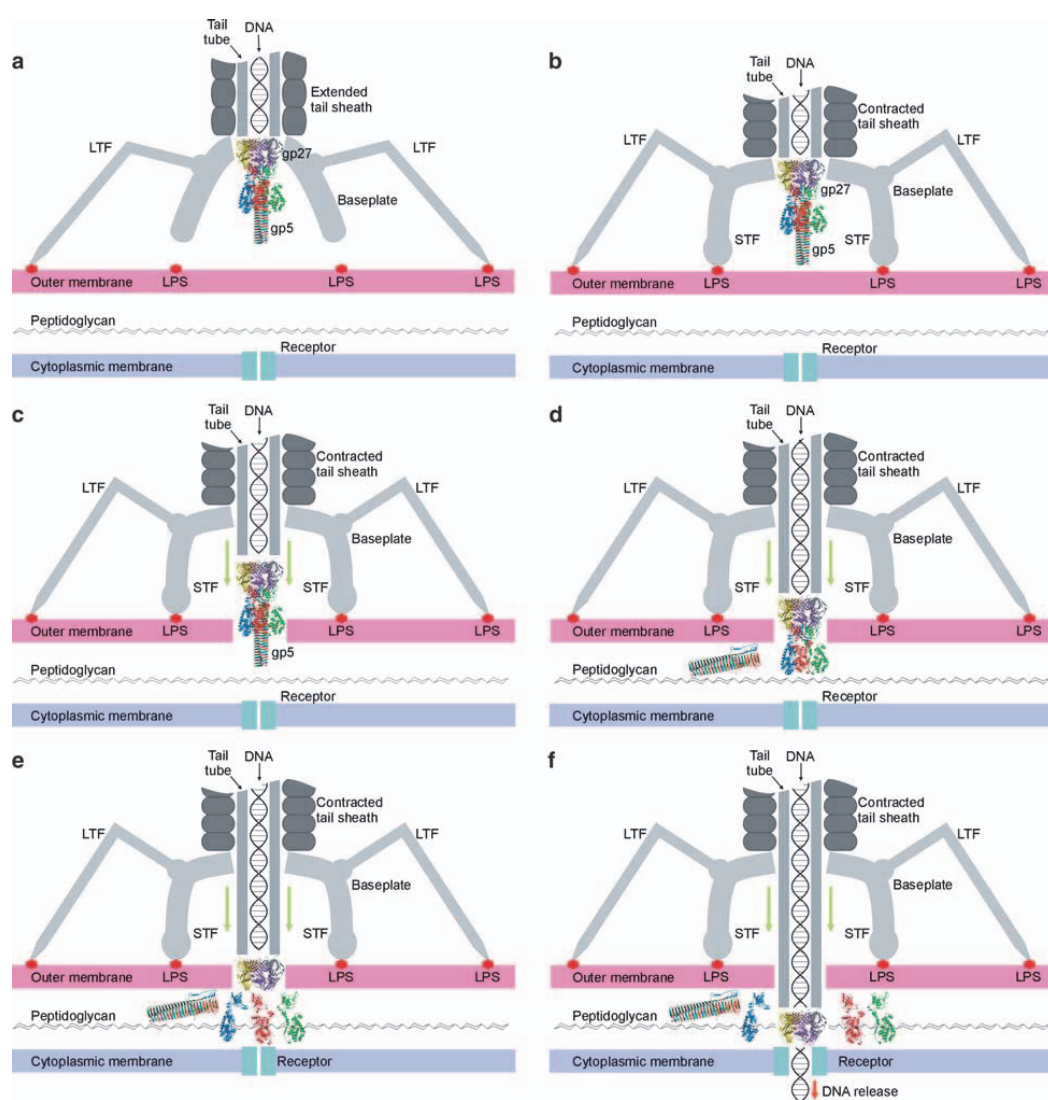


Рис. 28 Диаграмма процедуры инъектирования вирусной ДНК бактериофага Т4

3. АЛГОРИТМИЧЕСКИЕ ФОРМЫ ПОВЕДЕНИЯ И ЗНАКОВЫЕ СИСТЕМЫ

Рассмотренные выше данные о структуре вирусных капсидов и процессах их самосборки позволяют сделать некоторые важные выводы и заключения. Общая картина динамики процессов самосборки может быть получена только в результате реконструкции некоторой совокупности наблюдений промежуточных состояний. Это положение носит принципиальный характер. Техника экспериментов и наблюдений не позволяет зафиксировать процесс непосредственно в виде фильма. Наблюдаются в статике препараты, полученные из «живых» объектов и фиксирующие некоторые промежуточные состояния. Общая картина динамики процесса формируется в результате длительного накопления фактов и проявления интуиции в процессе реконструкции последовательностей событий. В этих условиях необходимым инструментом исследования должны стать математические модели, позволяющие увидеть скрытые от непосредственного наблюдения взаимосвязи и явления, что в целом должно повысить обоснованность, достоверность и продуктивность реконструкций.

Для принятия решения о том, что же именно является предметом математического моделирования, мы должны ответить на вопрос, что является главным сущностным признаком явления самосборки. Приведенные данные наблюдений и опубликованные рассуждения специалистов мирового уровня убеждают нас в том, что главным сущностным признаком самосборки является алгоритмический характер протекающих процессов избирательного взаимодействия участников самосборки. К настоящему времени трудом нескольких поколений вирусологов построены достоверные, многократно выверенные сценарии самосборки некоторых вирусов или их отдельных компонентов. Нет сомнений в том, что это устойчивые, точно воспроизводимые упорядоченные последовательности взаимосвязанных процедур, которые являются не чем иным как алгоритмами.

Необходимо отметить важнейший факт – алгоритмические формы поведения не являются исключительной прерогативой техногенных систем, созданных человеком. При исследовании процессов самосборки вирусных капсидов наблюдаются алгоритмические формы поведения в природных системах на молекулярном уровне. Вирусологи накопили значительные массивы экспериментальных данных, в которых зафиксированы многочисленные взаимосвязанные последовательности событий сборки. Для целого ряда объектов получен ответ на вопрос – как это происходит. Математическое моделирование процессов самосборки должно ответить на вопрос – почему это происходит именно так, а не иначе. Для полноценного моделирования самосборки необходима математическая модель алгоритмической динамики.

При постановке задачи моделирования самосборки сферических вирусов возникает соблазн провести аналогию, например, с фуллеренами и применить богатый арсенал математических моделей кристаллографии. Фуллерен это статическая, полностью замкнутая квазисферическая оболочка, обладающая определённой структурой симметрии. Это очень похоже на капсид сферического вируса. Однако, сферический вирус отличается тем, что его оболочка никогда не строится до конца в один приём. У сферического вируса на предоконечной стадии сборки всегда отсутствует завершающий капсомер, а к образовавшейся прорехе пристыковывается специфическая ассоциация белков, которая доставляет вирусную нуклеиновую кислоту и проталкивает её внутрь капсида. А во внутренней полости размещается другая ассоциация белков, которая плотно укла-

дыдает длинную цепь вирусной нуклеиновой кислоты. И только после завершения укладки нуклеиновой кислоты машина проталкивания отстыковывается от капсида, в нужную точку прибывает последний капсомер, и занимает своё место, заполняя прореху в сферическом капсиде. Так что вирусный капсид это молекулярная машина, которая осуществляет циклическую последовательность процедур, которая также является ничем иным как алгоритмом. То же можно утверждать о хвостовом отростке и фибриллах бактериофага Т4. Мягкий внешний чехол хвостового отростка собирается вполне определённым образом, в котором фиксируется неплотная упаковка капсидных белков по вертикальным контактам. Затем в нужный момент после закрепления вириона на поверхности клетки происходит вертикальное сжатие и перестройка структуры внешнего чехла, обеспечивающая поступательно-вращательное движение внутренней трубки, необходимое для прокалывания клеточной мембраны. Алгоритмическая динамика – это сущность исследуемого нами явления, которая проявляется на всех стадиях существования вириона – и при его самосборке, и при его самовосстановлении, и при выполнении функций транспорта и заражения клетки.

Из опыта построения и функционирования компьютерной техники мы знаем, что любой алгоритм – это продукт функционирования знаковой системы определённого типа. Если мы наблюдаем алгоритмические формы поведения, значит где-то, за пределами нашего наблюдения есть знаковая система, управляющая этим поведением. Вычислительная машина – это физическая реализация знаковой системы. Машина может быть механической, гидравлической, пневматической, электронной, может быть сделана из дерева, из металла, из пластика, может быть ламповой, транзисторной, твёрдотельной, но сущность её остаётся единой в разных реализациях и в разные времена на протяжении сотен и тысяч лет – это разные формы физического воплощения знаковой системы определённого типа. (Мы не ошиблись, упоминая столетние и тысячелетние периоды в истории существования программируемых автоматов. К программируемым автоматам кроме современного компьютера относятся музыкальные шкатулки XVII века, Жаккардов ткацкий станок XVIII века, вычислительная машина Бэббиджа XIX века, а также кукольные театры-автоматы, существовавшие в древней Александрии в античный период её истории.)

Знаковая система – это абстрактное умозрительное построение, которое определяет набор символов или алфавит, правила агрегации символов и построения записей, правила трансформации и редактирования записей, а также правила интерпретации записей. Современный компьютер – это техническая реализация простейшей примитивной знаковой системы, которая называется линейный императивный язык. Исходный набор символов этого языка называется система команд, в каждом типе компьютера имеется свой набор команд примерно из нескольких сотен позиций. Правила агрегации символов это их объединение в линейные последовательности или цепочки. Линейная последовательность символов образует запись, называемую программой. Правила редактирования – это копирование, слияние цепочек и их удаление или вырезание. Символы или команды компьютерного языка обозначают арифметические, логические и транспортные операции, выполняемые над данными. Правила интерпретации – это запуск аппаратуры, выполняющей арифметические, логические и транспортные операции с данными. Динамика работы машины определяется механизмом чтения записи путём последовательной выборки символов из цепочки, называемой программой. Выбранный символ или команда запускает соответствующую интерпретацию, в ходе которой про-

изводится транспорт и преобразование данных. По завершении выполнения текущей команды выбирается следующая команда из цепочки и так до исчерпания всей записи.

Точно такая же знаковая система функционирует на молекулярном уровне во внутриклеточной среде – это система синтеза белков. Совершенно справедливо в работе [15] авторы называют её биосинтетической машиной. Мы уточним, что совокупность процедур и алгоритмов синтеза белков управляется знаковой системой, построенной на базе линейного императивного языка. Биохимическая реализация этой знаковой системы охватывает взаимодействия целого ряда биомолекулярных комплексов от ДНК и сопутствующих белков из класса ДНК-полимераз до рибосомы и согласованных с нею транспортных РНК, манипулирующих аминокислотами. Система синтеза белков подробно описана в многочисленных учебниках (см., например, [19]). Нет смысла повторять изложение общеизвестных фактов. Тем не менее, мы намерены обратить внимание на то, что все компоненты, декларированные в составе понятия знаковая система, присутствуют в системе синтеза белков. Есть набор символов, это генетический код, состоящий из 64 символов, и есть носитель генетического кода – молекула ДНК. Есть правила агрегации символов в последовательные линейные цепочки, образующие записи, называемые генами, которые в свою очередь могут агрегироваться в более сложные конструкции, такие как группы генов, объединённые в опероны. Опероны, в свою очередь, содержат свои служебные записи, поддерживающие функции их идентификации и управления активностью генов. Есть процедуры копирования записей и их переноса на другой носитель, с ДНК на матричную РНК. Есть процедуры редактирования записей в ходе так называемого сплайсинга. При выполнении сплайсинга белки рестриктазы находят нужные участки символьной цепи матричной РНК и вырезают их. И наконец, есть система интерпретации записей, это ассоциация белков, называемая рибосомой, на которой происходит сборка первичной структуры белка. С точки зрения построения информационных технологий мы наблюдаем полноценную знаковую систему со всеми необходимыми атрибутами и базовыми функциями.

Мы свободно манипулируем терминами знак, символ, знаковая система, редактирование записей, не требуя пояснений, когда излагаются принципы действия компьютерной техники. Можно не вникать в детали реализации, важно, что в данном случае не возникает проблем происхождения знаков. Происхождение знаков известно по определению – так определили конструкторы, создавшие эту технику. В случае описания функций биохимических явлений требуется уточнить природу и статус понятия знак.

В природе нет специфических информационных веществ и соединений, содержащих знаки и символы. ДНК состоит из химических элементов, объединённых ковалентными и водородными связями. Откуда там появились знаки и символы. Знаки или символы появились не в молекулах ДНК, а в научных текстах, описывающих явления молекулярной генетики. Справочники говорят, что древнегреческое слово *σύμβολον* (symbolon) первоначально означало «соответствие». Со временем укоренилось распространённое употребление слова символ в ситуации, когда некий материальный предмет выступает как представитель другого материального предмета, свойства или отношения. В технических системах и в культуре установление отношений соответствия и выбор материальных предметов, которые будут выступать в роли символов или знаков в определённой степени произвольно и диктуется авторскими предпочтениями либо инженерной целесообразностью. Для природных знаковых систем факт соответствия устанавливается в ходе эксперимента и является выражением объективной реальности.

Введение в состав биохимических объектов символов и знаков это интеллектуальный приём, позволяющий представлять целые комплексы взаимосвязанных реакций на языке описания знаковых систем. При этом на уровне функционирования знаковой системы явление может быть представлено как некая целостность задолго до того, как будут экспериментально установлены все детали и подробности его химической природы. Понятие «знаковая система» в данном случае выступает как мощная синтетическая абстракция, интегрирующая целый комплекс взаимосвязанных реакций и биохимических явлений и позволяющая судить о нём в терминах построения знаковых систем и проявления их специфических закономерностей. А специфические закономерности знаковых систем объективно существуют и неумолимо проявляются как в техногенных системах, так и в природных явлениях, представленных в виде знаковых систем.

Далее мы будем рассматривать и сравнивать две знаковые системы: компьютер и систему синтеза белков. Оба явления могут быть представлены как знаковые системы, принадлежащие к одному типу – как знаковые системы, построенные на базе линейного императивного языка. Базовое соответствие, которое позволяет представить молекулу ДНК в качестве носителя некой символьной записи это соответствие комбинаторики состояний пар оснований в составе молекулы ДНК и расположения нуклеиновых кислот в полипептидной цепи первичной структуры белка. Соответствие это было установлено экспериментально и завершилось созданием таблицы генетического кода. Появление понятия символа применительно к молекуле ДНК произошло исключительно в контексте такого явления как синтез белков. Вне этого контекста, при изолированном исследовании молекулы ДНК можно было бы отметить только, что молекула обладает определённой вариабельностью и при разных вариантах взаимного расположения четырёх допустимых состояний реализации пар оснований она остаётся одним и тем же химическим объектом – молекулой дезоксирибонуклеиновой кислоты.

Специфические закономерности присущие знаковым системам проявляются незамедлительно. Трёхэлементная структура генетического кода была установлена умозрительно исходя из необходимости кодирования 20 – 23 аминокислот одиночными связями с четырьмя состояниями. В любом учебнике можно найти рассуждения о том, что двух пар нуклеотидов недостаточно, это будет только $4^2 = 16$, а три нуклеотида достаточно, хотя и с избытком, поскольку $4^3 = 64$ и таким образом остановились на триплете. Далее занялись установлением соответствия состояний триплета конкретным аминокислотам, что завершилось построением таблицы генетического кода. Был уточнён состав аминокислот, равный 20, установлена избыточность и многозначность кодирования, при которой одной аминокислоте могут соответствовать несколько значений кода. Были выделены коды, играющие роль знаков препинания и не соответствующие никаким аминокислотам. Важно отметить, что первоначально эти коды были названы нон-сенс кодами. Но самое интересное случилось несколько позднее – несколько позднее было установлено экспериментом, что на рибосоме есть механизм поддержания рамки чтения, который правильно фиксирует точку начала кодовой последовательности и далее синхронизирует продвижение цепи матричной РНК ровно на три нуклеотида при подготовке каждого следующего шага синтеза. Этот факт поражает воображение современного образованного человека. Мы узнали о существовании триплетного генетического кода из учебников, которые были написаны в 50-тые годы XX века. А откуда об этом узнали рибосомы миллионы лет тому назад? И тут мы должны признать, что

знаковые системы существуют не только в нашем воображении – они существуют в природе независимо от нас и нашего знания или незнания.

Разработчики компьютерной техники озабочены проблемой повышения производительности процессора. В результате длительного накопления опыта разработок сложились стандартные приёмы и методы повышения производительности. Одним из наиболее эффективных приёмов повышения продуктивности работы аппаратуры является метод конвейеризации, который хорошо себя зарекомендовал и широко используется в современной компьютерной технике. Заметим, что метод конвейеризации эффективен именно в контексте конструктивной специфики современного компьютера как носителя знаковой системы, построенной на базе линейного императивного языка. Тот же метод конвейеризации мы наблюдаем и в системах синтеза белков. В учебниках по молекулярной биологии описаны принципы работы рибосом как единичных объектов. В действительности это встречается относительно редко. В реальности, рибосомы, как правило, объединяются в пакеты, содержащие от 3 до 6 экземпляров, и работают в режиме конвейера. Конвейеризация работы рибосом показана на рис. 29.

Одна цепь м-РНК протягивается через последовательную упаковку рибосом и на каждой из них собирается своя полипептидная цепь первичной структуры белка. Таким образом одна молекула м-РНК поддерживает сборку сразу нескольких молекул белка одновременно в конвейерном режиме.

При построении систем программного обеспечения компьютеров широко используется метод макрогенерации. Программисты давно заметили, что огромное разнообра-

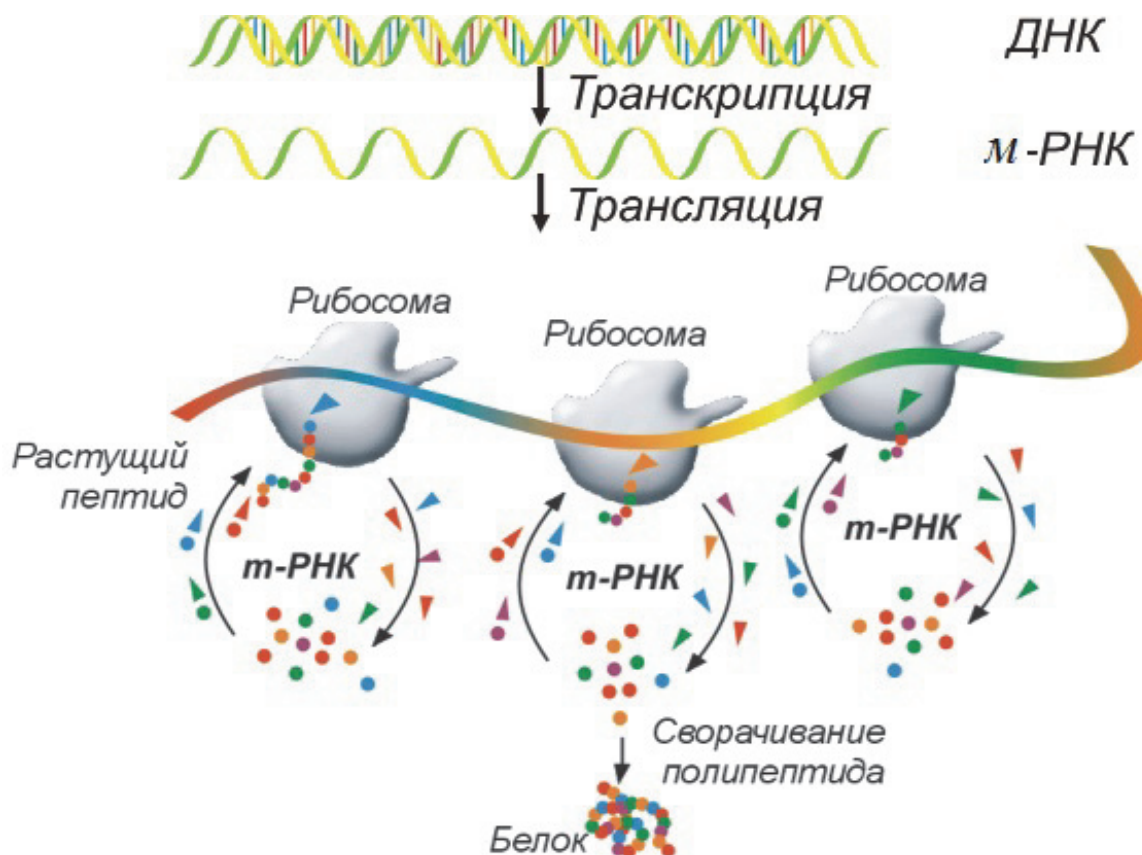


Рис. 29. Конвейеризация синтеза белков

зие реальных программ, имеет в своём составе ограниченный набор стереотипных часто повторяемых фрагментов. Разрабатывать и хранить для каждой рабочей программы полный загрузочный код при этом не всегда целесообразно. Гораздо эффективнее создавать библиотеки базовых стандартных подпрограмм для определённых классов задач и далее компоновать исполняемые программы из библиотечных элементов. При этом достигается значительная экономия ресурсов, как при программировании, так и при эксплуатации компьютерной техники. Для воплощения этой идеи в системах программирования были созданы и широко применяются так называемые средства макрогенерации. Библиотечным элементам присваиваются имена, а их программный код заносится в специальные сегменты памяти. Таким образом создаются макроопределения. Далее программирование ведётся уже на более высоком уровне в терминах макроопределений, где записи программ представляют собой последовательности имён макроопределений. При формировании и исполнении программ макрогенератор воспринимает имена макроопределений и осуществляет сборку загрузочного кода путём извлечения из памяти полных текстов макроопределений. Этот процесс называется макрорасширением. При этом существует два варианта макрорасширения – долговременное и оперативное макрорасширения. При долговременном макрорасширении загрузочный код собирается из макроопределений путём их копирования из памяти и встраивания в текст исполняемой программы. При оперативном макрорасширении текст макрорасширения не копируется, а макрогенератор в ходе выполнения программы передаёт управление библиотечному элементу, и он исполняется процессором.

Метод макрогенерации вероятнее всего работает и в системе синтеза белков. Первоначально на базе данных исследования генома прокариот, в частности классической *E. Coli* был провозглашён принцип «один ген – один белок». Это соответствует фактам, биосинтетическая машина прокариот устроена именно так. Но это самый нижний и примитивный уровень организации генома. У эукариот всё значительно сложнее. У эукариот была обнаружена ощутимая задержка выхода матричной РНК из ядра после её формирования в ходе считывания с ДНК. Далее было обнаружено, что в это время происходит процесс редактирования записи м-РНК, который получил название сплайсинг (от англ. splice-соединять, сращивать). Дело в том, что запись, называемая геном у эукариот носит прерывистый характер и состоит из чередующихся участков, называемых интронами и экзонами. Экзоны несут генетическую информацию и воплощаются в последовательности аминокислот в структуре белка, а интроны при синтезе белка прямо не используются и предположительно выполняют роль разделительных сегментов. Поэтому в процессе созревания м-РНК называемом сплайсингом, происходит удаление интронов. Однако этим дело не ограничилось. Далее было обнаружено явление очень сложного преобразования цепи м-РНК, в ходе которого меняется порядок следования экзонов. Это явление назвали альтернативным сплайсингом. Кроме того, существует ещё так называемый белковый сплайсинг. Это процесс редактирования первичной полипептидной цепи уже в ходе внутренней трансформации синтезированных белков. Похоже, что исследование явления сплайсинга продолжается и ещё не обрело окончательного завершённого состояния.

По результатам выполнения программы «Геном человека» приводятся некоторые данные по его организации и объёму, которые постоянно меняются и уточняются. Так на ранних стадиях считалось, что в геноме человека содержится примерно 50 тысяч генов, затем появилась цифра чуть более 30 тысяч и, наконец, в [24, 25] приводятся

оценки в 22 тысячи генов. В тоже время число действующих в организме человека белков оценивается порядком сотен тысяч или как близкое к 1 миллиону. Т.е. белков примерно в 30 – 40 раз больше чем генов. Реальное разнообразие белков создаётся в результате сплайсинга базовых записей, размещённых в геноме. Любопытно отметить, что число генов в геноме примитивного червя нематоды почти такое же как и у человека около 30 тысяч. Похоже, что доминирующим методом создания разнообразия белков и как следствие морфологического разнообразия является сплайсинг. Таким образом, можно считать, что подтверждается фундаментальная закономерность функционирования знаковых систем, построенных на базе линейного императивного языка. С точки зрения эффективности систем программирования, построенных на базе линейного императивного языка целесообразно создавать библиотеки стандартных стереотипных, часто повторяемых базовых фрагментов и далее комбинировать их в разных последовательностях средствами макрогенерации при формировании реального разнообразия белков.

Мы уточнили происхождение и статус понятия знак в природных системах как интеллектуальный приём фиксации экспериментально установленного соответствия двух разных физических объектов – последовательности состояний пар оснований в ДНК и последовательности аминокислотных остатков в первичной структуре белка. Таким же образом необходимо уточнить статус знаковой системы. Знаковая система – это символическое представление конструкции, которая поддерживает сложный и многоступенчатый процесс сборки белков на направляющей матрице, в роли которой выступает матричная РНК. Можно говорить, что запись, хранимая в ДНК и её копия в матричной РНК, являются символьными копиями первичной структуры белка. Таким образом проявляется роль знаковой системы. Знаковая система позволяет программировать огромное разнообразие белков путём формирования разнообразия символьных записей, сопряжённых с первичной структурой белков. При этом символьные записи являются программируемыми символьными копиями белков. В стихийный энтропийный процесс молекулярных взаимодействий вносится строгий порядок и программируемый детерминизм. Вероятность сборки заданной полипептидной цепи без направляющей матрицы близка к нулю и, похоже, практически неосуществима.

Роль знаковой системы можно проиллюстрировать по аналогии на одном умоглядном примере. Предлагается ответить на вопрос – может ли в результате массового камнепада спонтанно сложиться свод. Свод это базовая архитектурная конструкция, отличающаяся высокой устойчивостью и представляющая собой стабильную самоподдерживающуюся структуру. Чисто теоретически можно утверждать, что вероятность такого события не равна нулю, но практически это не реально. А допустим, что на пути камнепада располагается выпуклый глиняный купол. Самопроизвольное образование выпуклых симметричных глиняных куполов достаточно частое и естественное явление. Тогда в результате стихийного камнепада камни уложатся на глиняном куполе в строгом порядке с вероятностью близкой к единице. Далее мягкая порода глиняного купола со временем размывается, выветривается и мы имеем спонтанно сложившийся каменный свод.

Знаковая система – это метод описания и исследования явлений, представляющих собой локальные ниши, отграниченные от мира стихийного энтропийного поведения и поддерживающие внутренние упорядоченные взаимодействия и детерминированную алгоритмическую динамику.

Поскольку мы установили, что в процессах самосборки вирусных капсидов наблюдаются алгоритмические формы молекулярного поведения, логично предположить, что эффективным методом исследования самосборки может стать её представление в виде знаковой системы. Однако проблема моделирования явления самосборки заключается в том, что при его исследовании действительно наблюдаются алгоритмические формы поведения, а атрибуты описанных выше знаковых систем, основанных на линейных императивных языках, отсутствуют. В явлении самосборки отсутствует символьная копия объекта и манипулятор, который эту символьную запись интерпретирует. Это приводит нас к следующему важному выводу – процессы самосборки управляются знаковыми системами другого типа, отличного от линейных императивных языков. Решение проблемы математического моделирования явления самосборки требует расширения понятия знаковая система и построения других типов знаковых систем, адекватных явлению самосборки. Накопленные данные экспериментальных наблюдений самосборки позволяют определить облик таких знаковых систем и сформировать техническое задание на их разработку.

В литературе неоднократно встречаются рассуждения о том, что процесс самосборки подобен так называемым каскадным реакциям, в которых произошедшие текущие события молекулярных взаимодействий подготавливают условия для следующего шага, и в итоге развивается предопределённая последовательность реакций, где каждый текущий шаг является функцией от предыдущего состояния. В линейных императивных языках символьная копия объекта представлена экстенционально в полном объёме в виде символьной цепи. Далее символьная цепь протягивается через рибосому таким образом, что на каждом текущем шаге сборки в рамке чтения в активной зоне находится только один знак, один триплет и именно он определяет происходящие на этом шаге события. Не меняя порядок работы знаковой системы в целом в модельном эксперименте можно заменить явно заданную символьную цепь рекуррентным процессом, который на каждом шаге порождает текущий символ. Вместо символьной цепи необходимо устроить преобразователь символов и запустить его работу в режиме самообращения, при котором каждый текущий символ, порождаемый выходом преобразователя, используется как входной аргумент для порождения следующего символа. Такой самообращающийся рекуррентный генератор порождает определённую символьную последовательность, которая вполне может быть использована как символьная копия объекта в составе знаковой системы. При этом текущий знак в рамке чтения будет появляться не в результате протяжки символьной цепи, а в результате трансформации предыдущего символа в последующий функциональным преобразователем.

Замена громоздкой символьной цепи на компактный рекуррентный генератор позволяет многократно продублировать его и разместить символьные копии объекта непосредственно на множестве участников самосборки. Таким образом, знаковая система становится динамической и распределённой и может функционировать параллельно и без единого управляющего центра. При этом возникают определённые требования к участникам самосборки. Из реализации знаковой системы удаляются носитель символьной копии и манипулятор сборщик, но вместо этого их функции распределяются по участникам самосборки.

Участники самосборки становятся носителями символики знаковой системы, но теперь они обозначают сами себя, то есть несут на себе индикаторные метки, необходимые для осуществления избирательных взаимодействий. Важнейший атрибут само-

сборки это избирательные взаимодействия участников процесса. Следовательно, роль символов в данном случае заключается в индикаторной разметке всех участников процесса.

Поскольку самосборка это многошаговый протяжённый во времени процесс, в состав исходных элементов участников самосборки должны входить кодовые преобразователи, осуществляющие преобразования индикаторных кодовых меток, что необходимо для рекуррентного порождения знаковых последовательностей, поддерживающих многошаговый процесс сборочных операций.

Участники самосборки, осуществляющие избирательные акты взаимодействия должны «видеть» друг друга и распознавать значения кодовых меток. Это значит, что при построении моделей самосборки необходимо определённым образом решать проблему межмолекулярной коммуникации. Здесь не удастся отделаться умолчаниями. В математических моделях самосборки необходимо на абстрактном уровне постулировать правила взаимодействия кодовых меток, а при физической интерпретации формулировать правдоподобные гипотезы межмолекулярной коммуникации и молекулярного распознавания.

И, наконец, в рамках физической интерпретации необходима продуктивная гипотеза о работе механизмов прикрепления и сил удержания взаимодействующих участников самосборки, совместимая с принципами распознавания и избирательного взаимодействия.

ЗАКЛЮЧЕНИЕ И ВЫВОДЫ

Мы показали, что процесс самосборки представляет собой определённую форму алгоритмического поведения, наблюдаемую на молекулярном уровне. Этот тезис подтверждается многочисленными экспериментами, устанавливающими порядок следования и взаимосвязи процедур взаимодействия вирусных белков в ходе построения вирусных капсидов.

Алгоритмические формы поведения поддерживаются и управляются знаковыми системами определённых типов. Так, например, современный компьютер и система синтеза белков управляется идентичными знаковыми системами, построенными на базе линейных императивных языков. Знаковая система реализует символьную копию объекта, и содержит средства интерпретации, которые строят заданный объект по символической копии. При этом в знаковых системах, построенных на базе линейных императивных языков символическая копия и строящийся объект связаны отношением изоморфизма. Символическая цепь, записанная в геноме, представляет собой изоморфную копию первичной структуры белка.

Анализ данных экспериментальных наблюдений процессов построения вирусных капсидов позволяет сделать вывод о том, что процессы самосборки управляются знаковыми системами других типов, отличных от знаковых систем, построенных на базе линейных императивных языков и не имеющих аналогов в существующей компьютерной технике. Накопленные данные исследования процессов построения вирусных капсидов позволили определить общий облик и конструктивные требования, необходимые для разработки знаковых систем адекватных данным наблюдениям процессов самосборки. Сущность процессов самосборки достаточно точно определили Крик и Уотсон в работе [16]. Авторы считают, что геном вируса слишком мал, и его объём недостаточен для

записи полной изоморфной символьной копии вирусного капсида. Следовательно, вирусный геном определяет набор базовых элементов и правил взаимодействия, которые многократно повторяются при самосборке, образуя регулярные конструкции, обладающие определёнными видами симметрии. На языке построения знаковых систем это означает, что символьная копия и объект построения не связаны отношением изоморфизма. Символьная копия, реализует в знаковой системе сжатое компрессивное описание объекта, а алгоритмы самосборки разворачивают процесс декомпрессии сжатого описания в полноценный объект.

Выбор математических средств построения знаковых систем, управляющих процессами самосборки и разбор примеров моделирования самосборки типичных структур и субагрегатов вирусных капсидов, будут изложены во второй части статьи.

Литература

1. Первый Российский кристаллографический конгресс. От конвергенции наук к природоподобным технологиям. Сборник тезисов // Москва, ВДНХ, 21-26 ноября 2016, 452 стр.
2. Вили К., Детье В. Биология // Мир, Москва, 1975, 822 стр.
3. Лурия С., Дарнелл Дж., Балтимор Д., Кэмпбелл Э. Общая вирусология. // М.: Мир 1981, 680 стр.
4. Whon T. W. et al. Metagenomic characterization of airborne viral DNA diversity in the near-surface atmosphere // Journal of virology, 2012, Т. 86, №. 15, 8221-8231.
5. Зуссман М. Биология развития // М. Мир 1977, 301 с.
6. А.Лима-де-Фариа Эволюция без отбора: Автоэволюция формы и функции // М. Мир 1991, 455 с.
7. Агол В.И. Вирусы: корневая система древа жизни? // Природа 9 (2009), 3–11.
8. Атабеков И. Г. Применение вирусных структур в качестве инструментов нанотехнологий. Рос. Нанотехнологии 3 (2008), (1—2), 130–139.
9. Atabekov J., Nikitin N., Arkhipenko M., Chirkov S., Karpova O. (2011). Thermal transition of native Tobacco mosaic virus and RNA free viral proteins into spherical nanoparticles. J. // Gen. Virol, 92, 453–456.
10. Karpova O., Nikitin N., Chirkov S., Trifonova E., Sheveleva A., Lazareva E., Atabekov J. (2012). Immunogenic compositions assembled from tobacco mosaic virus-generated spherical particle platforms and foreign antigens // J. Gen. Virol. 93(2), 400–407.
11. Gerasopoulos K. et al. Microbatteries with tobacco mosaic virus templated electrodes // Micro Electro Mechanical Systems, 2008. MEMS 2008. IEEE 21st International Conference on. – IEEE, 2008, 960-963.
12. Газит Э. (2011). Нанобиотехнология: необъятные перспективы развития. // М.: Научный мир. 152 стр.
13. Архипенко М. В. и др. (2011). Искусственные вирусоподобные частицы, полученные in vitro из белка оболочки Х-вируса картофеля и чужеродных вирусных РНК // Acta naturae, 3(10), 42–48.
14. Под редакцией Уоддингтона К.Х. На пути к теоретической биологии // М. Мир. 1970 180 с.
15. Костюченко В.А. Месяндржинов В.В. Архитектура сферических вирусов // Успехи биологических наук 2002 г. т. 42 с. 177-192.
16. Crick F.H., Watson J.D. Virus Structure: General principles // Nature. 1956 Vol. 177, P. 473 – 475.
17. Fraenkel-Conrat H., Williams R. C. Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components // Proceedings of the National Academy of Sciences., 1955, Т. 41, №. 10, 690-698.
18. Клуг А. От молекул к биологическим ансамблям // Успехи биологических наук, 1984, Т. 142, вып. 1 с. 3 – 30.
19. Стент Г. Кэлиндар Р. Молекулярная генетика // М. Мир, 1981, 648 стр.

20. *Leiman P. G., Arisaka F., van Raaij M. J., Kostyuchenko V. A., Aksyuk A. A., Kanamaru S., Rossmann M. G.* Morphogenesis of the T4 tail and tail fibers // *Virology journal*, 2010, 7 (1), 355 pp.
21. *Iwasaki K., Trus B. L., Wingfield P. T., Cheng N., Campusano G., Rao V. B., Steven A. C.* Molecular Architecture of Bacteriophage T4 Capsid: Vertex Structure and Bimodal Binding of the Stabilizing Accessory Protein, *Soc // Virology* 271, (2000) pp. 321-333.
22. *Leiman P. G., Kanamaru S., Mesyanzhinov V.V., Arisaka F., Rossmann M. G.* Structure and morphogenesis of bacteriophage T4 // *CMLS, Cell. Mol. Life Sci.* 60 (2003) pp.2356–2370.
23. *Leiman, P. G., Kostyuchenko V. A., Rossmann M. G.* Bacteriophage T4 Assembly // Department of Biological Sciences Lilly Hall of Life Sciences, Purdue University, https://bilbo.bio.purdue.edu/~viruswww/Rossmann_home/movies/t4_final_02.php.
24. *Баранов В.С.* Генетический паспорт – основа индивидуальной и предикативной медицины // СПб, издательство Н-Л, 2009, 582 с.
25. *J. C. Venter, M. D. Adams et al.* The Sequence of the Human Genome. // *Science*, 2001, 291 (5507), pp 1304 – 1351. DOI: 10.1126/science.1058040

PROBLEMS OF MODELING OF THE SELF-ASSEMBLING PHENOMENON (pt. I)

A.V. Makhiboroda, A.A. Podobin, A.V. Ilichev

*Department of Applied Mathematics MIEM,
National Research University «Higher School of Economics»*

makhiboroda@yandex.ru

Received 01.03.2018

The long-term programs of technological development and the perspectives of adoption of nature-like technologies require intensive study and creation of mathematical foundation of self-assembling phenomenon. In this paper, the facts of experimental observations are analyzed and the hypothesis that the self-assembling is a realization of algorithmic forms of behavior at a molecular level is justified. It is well known that the algorithmic forms of behavior are generated by symbolic systems of different types. The problem is that the symbolic systems functioning in computer technology and the algorithmic dynamics generated by them do not agree well with the observed facts of self-assembling. The problem of mathematical modeling of self-assembling is formulated as the development of specific symbolic systems which generate the algorithmic dynamics adequate to the self-assembling phenomenon. The paper describes requirements to symbolic systems with distributed control, their compound, and general structure.