

01

## Математическое моделирование образования комплекса белковых молекул с учетом их доменной структуры

© Т.В. Кошлан,<sup>1</sup> К.Г. Куликов<sup>2,¶</sup>

<sup>1</sup> Санкт-Петербургский государственный университет,  
19034 Санкт-Петербург, Россия

<sup>2</sup> Санкт-Петербургский политехнический университет Петра Великого,  
195251 Санкт-Петербург, Россия

¶ e-mail: kulikov.kirill.g@gmail.com

(Поступило в Редакцию 21 сентября 2016 г.)

Представлена физическая модель взаимодействия белковых молекул и проведен анализ их способности образовывать сложные биологические комплексы, а также изучены их реакционные способности с использованием методов электростатики на примере белков гистонового шаперона Nap1 и гистонов H2A и H2B. Проанализирована способность белков образовывать устойчивые биологические комплексы с учетом различных участков аминокислотных последовательностей. Проведен анализ способности белковых молекул образовывать соединения путем расчета матриц потенциальной энергии электростатического взаимодействия аминокислотных остатков полипептидной белковой цепи. Для анализа способности белковых молекул образовывать сложные биологические соединения применен метод блочных матриц.

DOI: 10.21883/JTF.2017.04.44306.2040

### Введение

Настоящая работа посвящена разработке математической модели, которая позволит теоретически предсказать прохождение биохимической реакции в выбранном направлении с заданными белками, имеющими известные аминокислотные последовательности.

Отметим ряд работ, в которых приводится анализ различных аминокислотных последовательностей выбранных нами белков. Так, в статье [1] приведен обзор различных функций белка Nap1, преимущественно его участие в сборке и разборке нуклеосомы, взаимодействие белка Nap1 с различными факторами ремоделирования хроматина, приведена информация о различных участках связывания белка Nap1 с другими белками. В [2] авторы использовали водно-дейтериевый обмен в сочетании с масс-спектрометрией для нахождения участков связывания димера (H2A–H2B) с белком Nap1. В результате выполненной работы авторы установили, что при низкой ионной силе гистоны в димере (H2A–H2B) могут находиться в неупорядоченной конформации, однако связывание димера (H2A–H2B) с Nap1 уменьшает эту структурную неупорядоченность. В настоящей работе указывается, что две копии (H2A–H2B) связываются с гомодимером Nap1, также анализируются различные участки белков H2A и H2B, которые отвечают за связывание с различными участками белка Nap1.

В [3] авторы разработали количественный метод изучения аффинности Nap1 с гистонами. Было установлено, что Nap1 связывается с (H2A–H2B) с наномолярной аффинностью. Отмечено, что каждый димер Nap1 связывает 2 димера (H2A–H2B), а также показано, что концы молекулы Nap1 вносят синергетический вклад в связывание с гистонами.

Отметим, что, согласно ранее выполненным работам, исследующим взаимодействие гистонового шаперона с гистоновыми белками, нет полной ясности в вопросе об участках связывания данных биологических молекул при образовании комплекса гистонов с гистоновым шапероном. Так в [2,4] указывается, что концы белка Nap1 не принимают непосредственного участия в организации комплекса с гистоновыми димерами. В работах [1,3,5] указывается на участие C-конца белка Nap1 в организации комплекса с гистоновыми димерами.

В приведенных работах не указаны четкие критерии нахождения реакционной активности различных доменов белков, которые ответственны за вступление целой молекулы в различные биохимические реакции.

Таким образом, настоящая работа посвящена разработке математической модели, которая поможет обработать имеющиеся экспериментальные данные и теоретически предскажет прохождение биохимической реакции в выбранном направлении с заданными аминокислотными последовательностями, а также выявит участки белков, отвечающие за взаимодействия между различными белковыми молекулами, что и определяет актуальность поставленной задачи.

Работа состоит из нескольких частей. В первой части приведены основные положения формирования нуклеосомы, описаны принципы взаимодействия белка Nap1 с гистонами H2A и H2B, представлены вторичные структуры исследуемых белков. Вторая часть посвящена разработке физической модели взаимодействия белков с образованием биологических комплексов на основании электростатического взаимодействия белковых молекул с учетом их аминокислотной последовательности. В третьей части подробно рассмотрены биологические системы: образование гомодимеров (Nap1–Nap1), образова-

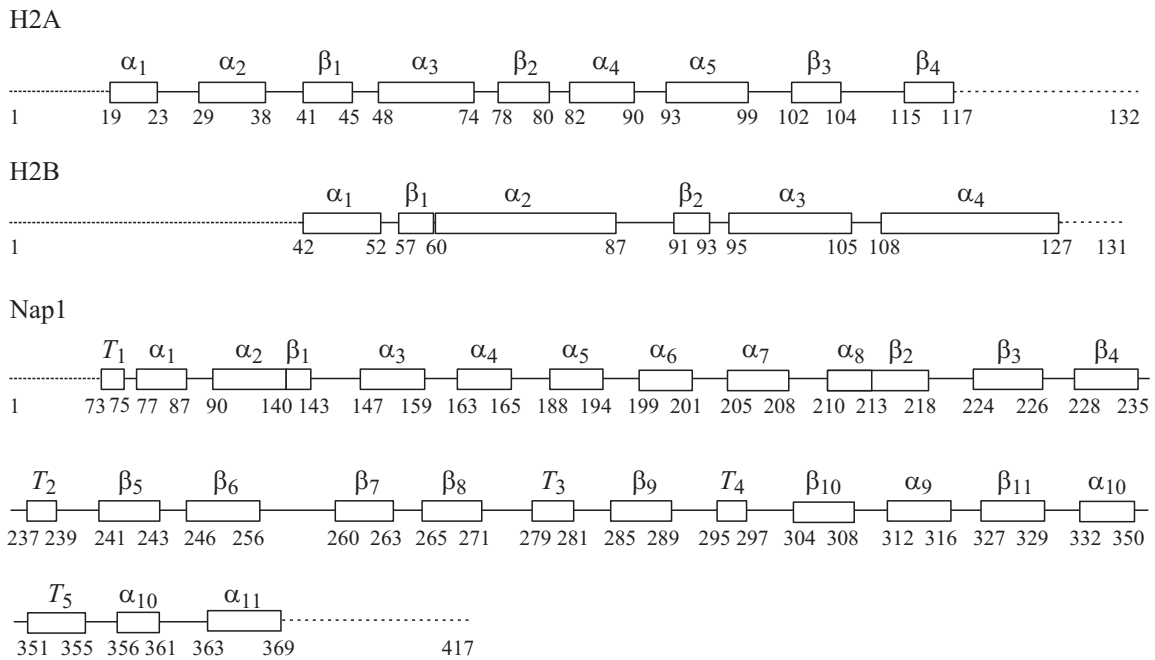


Рис. 1. Схемы белков H2A, H2B, Nap1 с учетом их вторичной структуры.

ние гетеродимеров (H2A–H2B), образование комплексов из четырех белков Nap1<sub>2</sub>(H2A–H2B) и представлены результаты численного моделирования взаимодействия выбранных биологических систем.

Все расчеты выполнены с учетом различных участков аминокислотных последовательностей, участвующих в реакции исследуемых белков.

### 1. Общие принципы формирования биологических комплексов

Рассмотрим общие принципы образования биологических комплексов, которые имеют отношение к нашей работе.

Мы будем анализировать белки H2A и H2B, которые принимают активное участие в формировании нуклеосомы. Нуклеосома представляет собой октамер белков гистонов, который несет на себе 145–147 пар нуклеотидов. В ядро нуклеосомы входит четыре семейства гистонов H3, H4 и H2A, H2B, каждый из которых представлен дважды. Восемь гистонов нуклеосомы собраны в четыре гетеродимера: два (H2A–H2B) и два (H3–H4) [6].

В естественных условиях сборке нуклеосомы способствуют специальные белки, такие как гистоновые шапероны Nap1 [1,6]. Белок Nap1 образует комплекс с гистоновыми димерами (H2A–H2B) и (H3–H4) [1,3,4,7]. Гистоновые шапероны связывают гистоны и вовлекают гистоны в упорядоченную сборку и разборку нуклеосомы [1]. Белок Nap1 был хорошо исследован in vitro [1,3,5].

В представленной работе будет проведен анализ различных белковых последовательностей с учетом их

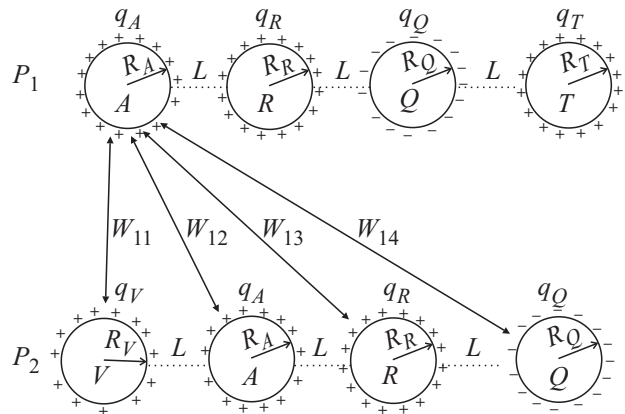


Рис. 2. Схема взаимодействия аминокислотных остатков двух взаимодействующих белков P<sub>1</sub> и P<sub>2</sub>. Каждая аминокислота представлена в виде равномерно заряженной сферы заданного диаметра.

доменной структуры с целью определения возможности образовывать устойчивые соединения (H2A–H2B), (Nap1–Nap1), Nap1<sub>2</sub>(H2A–H2B).

Отметим, что домены создаются объединением и чередованием  $\alpha$ -спиралей и  $\beta$ -слоев, между которыми открываются более рыхлые структуры [8], схемы белков с учетом их вторичной структуры  $\alpha$ -спиралей и  $\beta$ -слоев, T-поворотов, приведены на рис. 1.

Аминокислотные последовательности гистоновых белков H2A, H2B и гистонического шаперона Nap1 и их вторичные структуры были взяты в [9], номера белков в [9] составили соответственно P04911, P02293, P25293.

## 2. Описание физической модели

Опишем кратко физическую модель электростатического взаимодействия между аминокислотными последовательностями различных белков (рис. 2). Каждая аминокислота представлена в виде равномерно заряженной сферы со своим значением радиуса. В данной модели мы не учитываем сворачивания полипептидной цепи белка в правильную пространственную структуру. Белок мы представляем, как свободно-сочлененную полиаминокислотную последовательность [10]. Расстояния, заряды и более подробное описание модели приведены в [11]. Далее решается задача электростатического взаимодействия двух заряженных шаров и полученные в результате решения этой задачи значения потенциальной энергии электростатического взаимодействия между соответствующими аминокислотными остатками, которые мы представляем в виде равномерно заряженных сфер, записываются в матрицу (рис. 3).

При расчете потенциальной энергии электростатического взаимодействия между двумя белками, мы разработали два способа формирования матрицы для различных случаев:

1) Количество аминокислотных остатков одного белка равняется количеству аминокислотных остатков другого белка.

2) Количество аминокислотных остатков одного белка не равняется количеству аминокислотных остатков другого белка.

В первом случае мы формируем квадратную матрицу потенциальной энергии взаимодействия белков.

Во втором случае мы дописываем нулевые элементы к аминокислотной последовательности белка, у которого число аминокислотных остатков в полипептидной последовательности меньше. Это позволит сделать матрицу потенциальной энергии электростатического взаимодействия более разреженной, что приближает данную модель к реальным биологическим системам, в которых далеко не все аминокислотные остатки могут вступать во взаимодействия с аминокислотными остатками другого белка по причине наличия конформации у биологических молекул.

	<i>A</i>	<i>R</i>	<i>Q</i>	<i>T</i>
<i>V</i>	$W_{11}$	$W_{12}$	$W_{13}$	$W_{14}$
<i>A</i>	$W_{21}$	$W_{22}$	$W_{23}$	$W_{24}$
<i>R</i>	$W_{31}$	$W_{32}$	$W_{33}$	$W_{34}$
<i>Q</i>	$W_{41}$	$W_{42}$	$W_{43}$	$W_{44}$

**Рис. 3.** Представление матрицы потенциальной энергии электростатического взаимодействия  $W_{i,j}$ ,  $i = \overline{1,4}$ ,  $j = \overline{1,4}$  двух белков  $P_1$  и  $P_2$ .

	$P_1$			
	$W_{11}$	$W_{12}$	...	$W_{1n}$
	$W_{21}$	$W_{22}$	...	$W_{2n}$
$P_1$	...	...	...	...
	$W_{n1}$	$W_{n2}$	...	$W_{nn}$

**Рис. 4.** Квадратная матрица потенциальной энергии электростатического взаимодействия  $W_{i,j}$ ,  $i = \overline{1,n}$ ,  $j = \overline{1,n}$  для одинаковых белков с образованием гомодимера  $P_1-P_1$ .

	$P_1$	$P_2$
$P_1$	0	$W_{(P_1-P_2)}$
$P_2$	$W_{(P_2-P_1)}$	0

**Рис. 5.** Блочная матрица потенциальной энергии электростатического взаимодействия для двух различных белков при образовании гетеродимера белками различной длины  $P_1$  и  $P_2$ .

Если мы моделируем образование гомодимера, т.е. димера, образованного двумя одинаковыми белками, например  $P_1$  и  $P_1$ , то мы составляем квадратную матрицу потенциальной энергии электростатического взаимодействия двух одинаковых белков (рис. 4).

При моделировании взаимодействия двух различных белков с образованием гетеродимера, в котором составляющие его белки имеют различное количество аминокислотных остатков, мы формируем блочную матрицу (рис. 5).

Так, при образовании гетеродимера, например белками  $P_1$  и  $P_2$ , блочная матрица имеет вид (рис. 5).

Так как белки  $P_1$  и  $P_2$  представлены в димере по одному разу, блочные диагональные матрицы потенциальной энергии электростатического взаимодействия будут нулевыми.

Поскольку в нашей модели мы не учитываем конформацию биологических молекул, в этом случае будем считать, что каждый аминокислотный остаток одной молекулы белка может взаимодействовать с любым другим аминокислотным остатком другого белка.

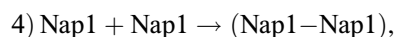
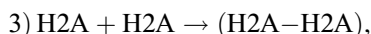
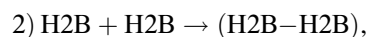
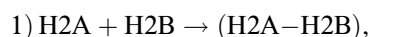
Для анализа биохимических процессов используем понятие числа обусловленности, которое будет характеризовать в данной физической постановке степень устойчивости конфигурации биологического комплекса (см. Приложение). При этом для выбора более устойчивого биохимического соединения между белками мы выбираем матрицу потенциальной энергии электростатического взаимодействия с наименьшим значением числа обусловленности.

**Таблица 1.** Числа обусловленности матриц потенциальной энергии электростатического взаимодействия пар димеров

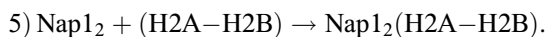
Наименование белка № 1	Наименование белка № 2	Число обусловленности
H2A	H2A	$1.85 \cdot 10^{12}$
H2B	H2B	$4.43 \cdot 10^{12}$
H2A	H2B	$4.31 \cdot 10^{10}$
H2A <sub>(19–117)</sub>	H2B <sub>(42–127)</sub>	$2.49 \cdot 10^{10}$

### 3. Численное моделирование взаимодействия биологических систем

Нами были промоделированы взаимодействия различных аминокислотных последовательностей белков, которые составляют вторичные структуры, входящие в различные домены белков, и проведен анализ возможности образования димеров различными участками аминокислотной последовательности белков:



а также проведен анализ их способности образовывать комплексы из четырех белков



Вторичные структуры моделируемых белков представлены на рис. 1.

Приведем для примера схему блочной матрицы потенциальной энергии электростатического взаимодействия для димера (H2A–H2B) (рис. 6).

Вначале мы выполнили расчет чисел обусловленности матриц потенциальной энергии электростатического взаимодействия для полных последовательностей небольших белков, таких как H2A и H2B, и для укороченных последовательностей H2A<sub>(19–117)</sub> и H2B<sub>(42–127)</sub> без учета неструктурированных концов молекул. Результаты расчетов для различных пар димеров (H2A–H2A), (H2B–H2B), (H2A–H2B) представлены в табл. 1.

	H2A	H2B
H2A	0	$W_{(\text{H2A-H2B})}$
H2B	$W_{(\text{H2B-H2A})}$	0

**Рис. 6.** Блочная матрица потенциальной энергии электростатического взаимодействия для соединений (H2A–H2B).

Полученные нами результаты указывают на то, что наиболее устойчивым соединением является димер (H2A–H2B). Отметим также, что выполненные расчеты для укороченных аминокислотных последовательностей H2A<sub>(19–117)</sub> и H2B<sub>(42–127)</sub>, где (19–117), (42–127) — аминокислотные остатки, показали, что происходит гораздо более устойчивое соединение центральных участков данных белковых молекул с образованием гетеродимера (H2A–H2B). Отметим, что данный результат согласуется с ранее полученными экспериментальными данными [12,13].

#### 3.1. Образование гетеродимера (H2A–H2B)

Нами было выполнено моделирование взаимодействия гистоновых белков при связывании их в димер (H2A–H2B) с учетом их вторичной структуры и проанализированы различные участки белков H2A и H2B с целью определения их способности образовывать устойчивые биологические комплексы. Отметим, что димер (H2A–H2B) образуется путем объединения двух гистонов в структуру, именуемую „рукопожатие“, в направлении „голова“ к „хвосту“ (рис. 7, а) [6,14]. Таким образом, было рассчитано взаимодействие между различными доменами белков H2A и H2B с учетом того, что при связывании в димер они связываются в направлении голова к хвосту, путем анализа их потенциальной энергии электростатического взаимодействия. Так же была проанализирована возможность связываться данных гистонов в димер в направлении „голова“ к „голове“ (рис. 7, б).

Для решения данной задачи нами брались различные участки аминокислотных последовательностей белков и выполнялся анализ матрицы их электростатического взаимодействия.

При образовании димера (H2A–H2B) белки H2A и H2B представлены в соединении по одному разу, следовательно, при образовании димера (H2A–H2B) из мономеров H2A и H2B диагональные блочные матрицы будут нулевыми.

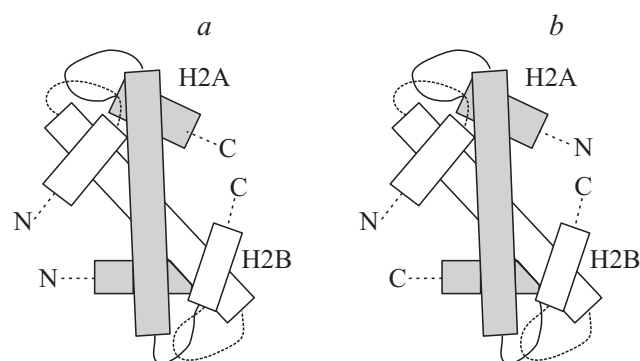
Каждый блок в матрице представляет собой матрицу потенциальной энергии электростатического взаимодействия двух соответствующих белков. В расчетах были использованы центральные участки аминокислотной последовательности белков H2A с 19 по 117 аминокислотный остаток и H2B с 42 по 127 аминокислотный

**Таблица 2.** Числа обусловленности матриц потенциальной энергии электростатического взаимодействия белков (H2A–H2B) с учетом структуры голова–хвост

Наименование белка	Аминокислотная последовательность	Число обусловленности
H2A <sub>(19–45)</sub> H2B <sub>(101–127)</sub>	RSAKAGLTFPVGRVHRLRRGNYAQRI AVRLILPGELAKHAVSEGTRAVTKYSS	5790
H2A <sub>(93–117)</sub> H2B <sub>(42–68)</sub>	DELNKLGNVTIAQGGVLPNIHQNL SSYIYKVLKQTHPDTGISQKSMSILNS	25.6
H2A <sub>(48–74)</sub> H2B <sub>(61–87)</sub>	GAPVYLTAVLEYLAAEILELAGNAARD KSMSILNSFVNDIFERATEASKLAAY	1220

**Таблица 3.** Числа обусловленности матриц потенциальной энергии электростатического взаимодействия белков (H2A–H2B) с учетом структуры голова к голове

Наименование белка	Аминокислотная последовательность	Число обусловленности
H2A <sub>(19–45)</sub> H2B <sub>(42–68)</sub>	RSAKAGLTFPVGRVHRLRRGNYAQRI SSYIYKVLKQTHPDTGISQKSMSILNS	$7.78 \cdot 10^9$
H2A <sub>(90–117)</sub> H2B <sub>(87–114)</sub>	RNDDELNKLGNVTIAQGGVLPNIHQNL YNKKSTISAREIQTAVRLILPGELAKHA	$5.04 \cdot 10^5$
H2A <sub>(93–120)</sub> H2B <sub>(95–123)</sub>	DELNKLGNVTIAQGGVLPNIHQNL REIQTAVRLILPGELAKHAVSEGTRAVT	$2.74 \cdot 10^3$



**Рис. 7.** Схематическое представление образования димера (H2A–H2B) в случае образования структуры голова–хвост (a) и соответственно в случае голова–голова (b). Полипептидная цепь нумеруется в направлении от N-конца к C-концу цепи.

остаток. Возможные взаимодействия концов гистонов при образовании димеров нами не учитывались.

В табл. 2 представлены результаты выполненных расчетов для различных доменов белков H2A и H2B с учетом двух анализируемых структур связывания белков в димеры.

Результаты выполненных расчетов для структуры связывания белков голова к хвосту демонстрируют образование более устойчивого биологического комплекса для всех трех выбранных участков:

(H2A<sub>(19–45)</sub>–H2B<sub>(101–127)</sub>), (H2A<sub>(93–117)</sub>–H2B<sub>(42–68)</sub>), (H2A<sub>(48–74)</sub>–H2B<sub>(61–87)</sub>).

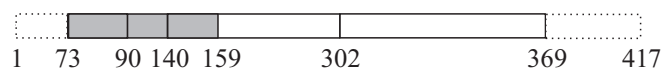
Результаты, полученные для структуры голова к голове, на примере связывания димера H2A–H2B, показывают менее устойчивый биологический комплекс между последовательностями (H2A<sub>(19–45)</sub>–H2B<sub>(42–66)</sub>) и (H2A<sub>(90–117)</sub>–H2B<sub>(87–114)</sub>) и возможность более устойчивого взаимодействия для последовательностей (H2A<sub>(93–117)</sub>–H2B<sub>(95–123)</sub>).

Результаты табл. 3 показывают, что наличие одного участка, демонстрирующего образование устойчивого биологического комплекса (H2A<sub>(93–117)</sub>–H2B<sub>(95–123)</sub>), не имеет решающего значения в образовании структуры димера, поскольку связывание других двух участков в структуре голова к голове очень слабое.

Полученные результаты находятся в хорошем согласии с ранее выполненными работами [6,14], в которых указывается, что гетеродимер (H2A–H2B) образуется при связывании в направлении голова к хвосту.

### 3.2. Анализ образования гомодимера гистоновым шапероном Nap1

Как известно из ранее выполненных работ [1,3,4], белок Nap1 образует в растворе гомодимер. Этот белок имеет определенные участки связывания, отвечающие за гомодимеризацию [1,4,15]. Отметим, что в этих работах был указан один из таких участков, который оказался одинаковой аминокислотной последовательностью для двух одинаковых белков Nap1.



**Рис. 8.** Схема гистонового шаперона с учетом анализируемых в работе участков, принимающих участие в образовании гомодимера и в связывании с гетеродимером (H2A–H2B).

**Таблица 4.** Числа обусловленности матриц потенциальной энергии электростатического взаимодействия различных участков Nap1

Наименование белка	Число обусловленности
Nap1 <sub>(90–140)</sub> Nap1 <sub>(90–140)</sub>	$1.68 \cdot 10^6$
Nap1 <sub>(145–195)</sub> Nap1 <sub>(145–195)</sub>	$8.33 \cdot 10^{11}$
Nap1 <sub>(190–249)</sub> Nap1 <sub>(190–249)</sub>	$2.11 \cdot 10^{11}$
Nap1 <sub>(260–310)</sub> Nap1 <sub>(260–310)</sub>	$2.54 \cdot 10^{11}$
Nap1 <sub>(340–390)</sub> Nap1 <sub>(340–390)</sub>	$2.96 \cdot 10^8$

На рис. 8 изображена схема белка Nap1 с учетом различных участков, принимающих участие в связывании гистоновых белков H2A и H2B в образовании гомодимера Nap1<sub>2</sub>. Пунктирной линией на схеме обозначены неструктурированные концы молекулы, которые являются участками с аминокислотной последовательностью 1–74 и 369–417 для белка Nap1<sub>2</sub> [1,15].

На рисунке показан отрезок аминокислотной последовательности 73–159, ответственный за гомодимеризацию белка Nap1. Участок 302–417 ранее был указан в работе [1], как участок электростатического связывания с гистоновыми белками. В [1,3,5] указывается на роль С-конца белка Nap1 в образовании комплекса с гистоновыми димерами. В настоящей работе было выполнено моделирование с целью выяснения, какие из одинаковых последовательностей могут принимать участие в процессе гомодимеризации белка Nap1, т.е. был выполнен анализ взаимодействия между одинаковыми последовательностями на предмет образования устойчивого комплекса (Nap1–Nap1). В работе [15] указывается на важность структуры альфа-спирали на участке с 90 по 140 аминокислотный остаток в организации комплекса гомодимера белка Nap1.

Анализировались следующие одинаковые последовательности на предмет поиска более устойчивого биологического комплекса:

1. Nap1<sub>(90–140)</sub> KNVKEKLLSLKTLQSELFVEVEKEFEQVEMFELENKFLQKYKPIWEQRSRIIS;
2. Nap1<sub>(145–195)</sub> PKPEQIAKQGEIVESLNETELLYDDEEKAQNDSSEEEQVKGIPSWLTALEN;
3. Nap1<sub>(190–249)</sub> PIVCDTITDRDAEVLLEYLQDIGLEYLTDGRPGFKLLFRFDSSANPFFTNDILC;
4. Nap1<sub>(260–310)</sub> YSGDFIYDHAEGCEISWKDNAHNVTVDLEMRKQRNKTTKQVRIPEKITPIE;
5. Nap1<sub>(340–390)</sub> LDYSIGEQLKDKLPIRAVDWFTGAALFEFEFEDEEEADEDEDEEEDDDHGL.

Из табл. 4 следует, что первая последовательность Nap1<sub>(90–140)</sub> вносит наибольший вклад в образование гомодимера Nap1–Nap1, поскольку данные аминокислотные последовательности образуют более устойчивый биологический комплекс. Остальные анализируемые аминокислотные последовательности не могут образовывать такой же устойчивый биологический комплекс, как последовательность Nap1<sub>(90–140)</sub>, что хорошо согласуется с ранее выполненными работами [2,15].

### 3.3. Анализ образования комплекса Nap1<sub>2</sub>(H2A–H2B) с учетом различных участков белка Nap1

Гистоновый шаперон Nap1 в растворе обычно существует как гомодимер (Nap1–Nap1), который связывается с гистоновым димером (H2A–H2B) [1,5].

Отметим, что гистоновый шаперон имеет гораздо более длинную последовательность, чем гистоновые белки. Предполагается, что за связывание с димерами гистоновых белков отвечают определенные участки аминокислотной последовательности гистонового шаперона [5,3].

Так, в работах [2,3] указано, что с гомодимером Nap1<sub>2</sub> связывается два гетеродимера H2A–H2B. В [4] проведенный экспериментальный анализ взаимодействия гистонового шаперона с гистоновыми белками показал, что с гомодимером Nap1<sub>2</sub> связывается только один гетеродимер (H2A–H2B). В [5] отмечено, что на данный момент существуют различные гипотезы о характере связывания димера Nap1<sub>2</sub> с гистоновым димером (H2A–H2B).

Нами был произведен анализ способности различных последовательностей гистонового шаперона Nap1 связываться с одним димером гистоновых белков (H2A–H2B). Для этого мы разделили гистоновый шаперон на несколько отрезков с целью выявления наиболее устойчивого участка, образующегося в процессе образования сложного комплекса с гистоновыми белками Nap1<sub>2</sub>(H2A–H2B).

Были рассчитаны значения потенциальной энергии электростатического взаимодействия для следующих пар белков: (Nap1–H2A), (Nap1–H2B), (H2A–H2B). Для каждой пары обозначенных белков нами были получены соответствующие матрицы потенциальной энергии

	Nap1	H2A	H2B
Nap1	$W_{(Nap1-Nap1)}$	$W_{(Nap1-H2A)}$	$W_{(Nap1-H2B)}$
H2A	$W_{(H2A-Nap1)}$	0	$W_{(H2A-H2B)}$
H2B	$W_{(H2B-Nap1)}$	$W_{(H2B-H2A)}$	0

**Рис. 9.** Блочная матрица потенциальной энергии электростатического взаимодействия для соединения Nap1<sub>2</sub>(H2A–H2B).

**Таблица 5.** Числа обусловленности потенциальной энергии электростатических взаимодействий для различных последовательностей белка Nap1 и целых белковых последовательностей гистоновых белков H2A и H2B

Последовательность белка Nap1	Наименование димера	Число обусловленности
Nap1 <sub>(74–417)</sub>	H2A–H2B	$4.49 \cdot 10^{14}$
Nap1 <sub>(74–365)</sub>	H2A–H2B	$2.58 \cdot 10^{16}$
Nap1 <sub>(302–417)</sub>	H2A–H2B	$7.12 \cdot 10^6$

электростатического взаимодействия. Используя полученные матрицы для каждой пары белков, мы сформировали блочную матрицу (рис. 9).

Следует отметить, что белок Nap1 представлен в соединении дважды, поэтому первая блочная матрица будет содержать матрицу электростатического взаимодействия белка Nap1 с Nap1. Остальные диагональные блочные матрицы нулевые, так как белки H2A и H2B представлены только по одному разу.

В результате моделирования мы получили число обусловленности блочной матрицы потенциальной энергии электростатического взаимодействия для различных аминокислотных последовательностей белка Nap1, при этом нами брались целые последовательности гистоновых белков H2A и H2B. Отметим, что при расчете взаимодействия Nap1 с гистоновым димером мы использовали следующие последовательности белка Nap1: Nap1<sub>(74–417)</sub>, Nap1<sub>(74–365)</sub>, Nap1<sub>(302–417)</sub>.

Из табл. 5 следует, что число обусловленности, полученное для аминокислотной последовательности Nap1<sub>(74–417)</sub>, указывает на возможность образования более устойчивого биологического комплекса с димером (H2A–H2B) по сравнению с результатом, полученным для аминокислотной последовательности гистонового шаперона Nap1<sub>(74–369)</sub>. Данные результаты указывают на возможность участия в образовании комплекса С-конца гистонового шаперона Nap1<sub>(369–417)</sub> с гистоновым гетеродимером (H2A–H2B). Отметим, что полученный результат согласуется с работой [3], в которой указывается на вклад во взаимодействие С-конца гистонового шаперона при образовании комплекса с гетеродимером гистонов (H2A–H2B).

Как следует из табл. 5, участок, который формирует более устойчивый биологический комплекс гомодимера Nap1<sub>2</sub> с гистоновым гетеродимером (H2A–H2B), будет являться аминокислотной последовательностью Nap1<sub>(302–417)</sub>.

Результаты, полученные для аминокислотной последовательности Nap1, указывают на образование устойчивого биологического комплекса между димером (H2A–H2B) и выбранным участком белка Nap1<sub>(302–417)</sub>.

Отметим, что в работе [1] указывается, что ответственным участком белка Nap1 за связывание со всеми гистонами является аминокислотная последовательность с 302 по 417 аминокислотный остаток, а в [4] установлено, что главный контакт между димерами

(H2A–H2B) и белком Nap1 приходится на аминокислотные остатки 328–336 белка Nap1.

## Заключение

В настоящей работе мы провели анализ взаимодействия различных аминокислотных последовательностей белков H2A, H2B и Nap1 образовывать устойчивые биологические комплексы. В результате проведенного моделирования мы ввели три градации числа обусловленности для различных длин исследуемых белков. Для этих трех данных групп мы установили следующие критерии, которые позволят нам определить степень устойчивости между выбранными аминокислотными последовательностями.

1. При моделировании взаимодействия полных белковых последовательностей до 140 аминокислотных остатков мы полагаем, что образованию устойчивого биологического комплекса отвечает число обусловленности  $10^{10}$ . Если число обусловленности больше  $10^{10}$ , то это будет означать, что сформированный биологический комплекс образует менее устойчивые связи.

2. При моделировании взаимодействия аминокислотных последовательностей, содержащих примерно 50 аминокислотных остатков, мы считаем, что образованию устойчивого биологического комплекса отвечает показатель числа обусловленности  $10^6$ . При этом если число обусловленности меньше выбранного предела, то это будет указывать на усиление связывания выбранных аминокислотных последовательностей. Если число обусловленности большего введенного критерия, то это будет указывать на ухудшение взаимодействия между аминокислотными последовательностями.

3. При моделировании белковых последовательностей, содержащих 25–35 аминокислотных остатков, мы считаем, что образованию устойчивого биологического комплекса отвечает число обусловленности  $10^3$ . Числа обусловленности меньше выбранного значения указывают на более прочное связывание выбранных биологических объектов, а числа обусловленности больше данного значения будут указывать на менее прочное связывание биологического комплекса.

Таким образом, настоящая работа позволяет сделать вывод о взаимодействии между различными аминокислотными последовательностями на основании использованного критерия (числа обусловленности) и опре-

делять, какие из приведенных биологических объектов образуют более устойчивые соединения. Так, к примеру, если известна конформация молекул и известно, какими доменами структуры могут взаимодействовать, то наш расчет внесет ясность в экспериментальные работы и выявит наиболее устойчивые участки взаимодействия.

Модель также продемонстрировала хорошую чувствительность к аминокислотному составу исследуемых белков, что позволит в дальнейшем прогнозировать теоретически аминокислотные последовательности с заданными физическими свойствами, и тем самым ускорить и повысить эффективность проводимых экспериментов, а так же снизить их стоимость за счет уменьшения числа проводимых опытов.

Отметим, что мы выполнили моделирование с учетом идентификации уже известных мест связывания и получили достаточно хорошее согласие с экспериментом, однако, возможно, в исследуемых нами биологических комплексах есть и другие участки, отвечающие за связывание белков между собой, для которых необходимо выполнять моделирование отдельно при необходимости проведения эксперимента по заданным аминокислотным последовательностям.

Авторы считают своим долгом выразить глубокую благодарность профессору кафедры „Физическая электроника“ СПбГПУ доктору физ.-мат. наук А.П. Головицкому за консультации и конструктивное обсуждение результатов.

## Приложение. Физическая интерпретация числа обусловленности

Пусть даны две базисные функции  $\varphi_1$ ,  $\varphi_2$  и два параметра  $c_1$ ,  $c_2$ . Рассмотрим невязку в МНК

$$\tilde{p}^2(c_1, c_2) = \sum_{i=1}^k [y_i - c_1\varphi_1(x_i) - c_2\varphi_2(x_i)]^2.$$

Тогда величина  $\tilde{p}^2(c_1, c_2)$  будет характеризовать потенциальную энергию деформированного биологического комплекса (конгломерата), т. е. подвергнутого внешнему воздействию (например, вследствие термических столкновений при  $T > 0$ ). Роль  $c_1$ ,  $c_2$  играют длины связей.

Отметим, что в пространстве параметров  $c_j$ , где  $j = \overline{1, 2}$ , т. е. в координатах  $c_1$  и  $c_2$ , функция  $\tilde{p}^2(c_1, c_2)$  при малых воздействиях представляет собой параболоид (рис. 10, а).

Под  $\tilde{p}_{\min}^2$  мы будем понимать потенциальную энергию стационарного состояния биологического комплекса, т. е. когда реализуется минимум его потенциальной энергии, тогда под  $p$  понимается средняя потенциальная энергия биологического комплекса при наличии воздействия, т. е.

$$p - \tilde{p}_{\min}^2 \approx kT,$$

где  $k$  — постоянная Больцмана,  $T$  — абсолютная температура.

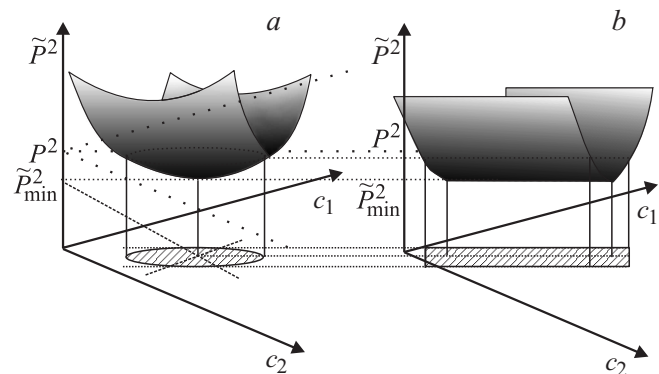


Рис. 10. Геометрическая интерпретация поведения невязки МНК в зависимости от величин параметров  $c_1$  и  $c_2$ .

Пусть  $p \geq \tilde{p}_{\min}^2$  (если имеет место обратное неравенство, то это означает, что вид аппроксимационной функции выбран неудачно). Тогда сечение параболоида плоскостью  $\tilde{p}^2 = p^2$  (поскольку  $p = \text{const}(c_1, c_2)$ ) дает связь между параметрами  $c_1$ ,  $c_2$ :

$$\tilde{p}^2(c_1, c_2) = \sum_{i=1}^k [y_i - c_1\varphi_1(x_i) - c_2\varphi_2(x_i)]^2,$$

т. е. эллипс в координатах  $c_1$ ,  $c_2$  (рис. 10, б). Отношение длин полуосей этого эллипса (большой к малой) равно мере обусловленности матрицы [16]. Чем хуже обусловлена матрица, тем более вытянутым является этот эллипс. Например, если представить, что ширина эллипса  $\approx 1$  mm, то его длина при числе обусловленности  $\approx 2009$  — более 2 m. Таким образом, в этом случае параболоид будет похож на желоб параболического сечения с осью, направленной вдоль большой оси эллипса (рис. 10).

В работе [16] этот эллипс назван характеристическим. При этом форма характеристического эллипса для малых воздействий не будет зависеть от  $kT$ . В направлении большой оси сопротивление растяжению наименьшее, и там наиболее вероятен разрыв связей. Следовательно, число обусловленности характеризует степень устойчивости данной конфигурации биологического комплекса.

## Список литературы

- [1] Zlatanova J., Seebart C., Tomschik M. // FASEB J. 2007. Vol. 21. N 7. P. 1294–1310.
- [2] D'Arcy S., Martin K.W., Panchenko T., Chen X., Bergeron S., Stargell L.A., Black B.E., Luger K. // Mol. Cell. 2013. Vol. 51. N 5. P. 662–677.
- [3] Andrews A.J., Downing G., Brown K., Park Y.J., Luger K. // J. Biol. Chem. 2008. Vol. 283. N 47. P. 32412–32418.
- [4] Aguilar-Gurrieri C. // EMBO J. 2016. Vol. 35. N 13. P. 1465–1482.
- [5] Aguilar-Gurrieri C. // Structural studies of nucleosome assembly. 2013.



- [6] *Dias R., Lindman B.* DNA Interactions with Polymers and Surfactants. John Wiley & Sons, Inc. 2007. P. 135–172.
- [7] *Bowman A., Ward R., Wiechens N., Singh V., El-Mkami H., Norman D.G., Owen-Hughes T.* // *Mol. Cell.* 2011. Vol. 41. N 4. P. 398–408.
- [8] *Березов Т.Т.* Биологическая химия. М.: Медицина, 1998.
- [9] *The Universal Protein Resource* // <http://www.uniprot.org/> (дата обращения 30.07.2016).
- [10] *Семчиков Ю.Д.* Высокомолекулярные соединения. М.: Академия, 2010.
- [11] *Куликов К.Г., Кошлан Т.В.* // *ЖТФ.* 2016. Т. 86. Вып. 10. С. 131–138.
- [12] *Shimoyama S., Nagadoi A., Tachiwana H., Yamada M., Sato M., Kurumizaka H., Nishimura Y., Akashi S.* // *J. Mass. Spectrom.* 2010. Vol. 45. N 8. P. 900–908.
- [13] *Moss T., Cary P.D., Abercrombie B.D., Crane-Robinson C., Bradbury E.M.* // *Eur. J. Biochem.* 1976. Vol. 71. N 2. P. 337–350.
- [14] *Ramaswamy A., Bahar I., Ioshikhes I.* // *Proteins Structure Function and Bioinformatics.* 2004. Vol. 58. N 3. P. 683–696.
- [15] *Park Y.J., Luger K.* // *Proc. Natl. Acad. Sci. U.S.A.* 2006. Vol. 103. N 5. P. 1248–1253.
- [16] *Демиденко Е.З.* Линейная и нелинейная регрессии. М.: Финансы и статистика, 1981.