

Министерство образования и науки Российской Федерации  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра оптики и биофотоники

**Термодинамика связывания в комплементарных парах**

**ДНК и РНК и в молекулярных кластерах воды**

**АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ**

Студентки 2 курса 253 группы

03.04.02 «Физика» (Биофотоника)

Физического факультета

Бондяковой Анны Александровны

Научный руководитель

д.ф.-м.н, профессор \_\_\_\_\_

К.В. Березин

Заведующий кафедрой

д.ф.-м.н, профессор \_\_\_\_\_

В.В. Тучин

Саратов 2017

## **Введение**

### **Актуальность работы**

Стабильность нуклеиновых кислот ДНК, РНК взаимодействия канонических пар оснований за счет водородных связей и межплоскостного стекинг-взаимодействия [1]. В связи с этим исследованиям межмолекулярного взаимодействия оснований нуклеиновых кислот было посвящено много экспериментальных [2] и теоретических работ [3]. Квантово-механические расчеты биомолекул с использованием метода функционала плотности позволяют получать структурные и энергетические параметры молекулярных систем и их комплексов с водородными связями на уровне более трудоемких методов таких, как MP2 и CCSD(T). Они оказываются полезным дополнением к экспериментальным методам, а также для калибровки различных эмпирических силовых полей, применяемых для изучения более сложных комплексов, например, комплексов ДНК-белок-вода, ДНК-лекарство.

Обычно расчеты выполняют с применением ограниченных наборов базисных гауссовых функций, что приводит к необходимости учета суперпозиционной поправки в энергетических и структурных параметрах.

Молекулы воды образуют кластеры, посредством водородных связей, комплексы образованные молекулами воды представляют объект экспериментальных и теоретических исследований. Понимание их структуры важно для понимания процесса образования льда, химии растворов, а также большого количества биохимических процессов. Исследование структуры кластеров воды и расчет межмолекулярного взаимодействия является первым шагом к пониманию макроскопических свойств воды. Пентамеры молекул воды конфигурации кольца, широко распространены в природе, например, в решетках гидратов и в сольватации гидрофобных групп малых молекул, а также в белках и в молекулах ДНК.

Молекулы воды образуют кластеры, посредством водородных связей, комплексы, образованные молекулами воды представляют объект

экспериментальных и теоретических исследований. Понимание их структуры важно для понимания процесса образования льда, химии растворов, а также большого количества биохимических процессов. Исследование структуры кластеров воды и расчет межмолекулярного взаимодействия является первым шагом к пониманию макроскопических свойств воды. Трудность строгого всеобъемлющего описания структуры водных кластеров заключается в том, что существует возможность реализовать множество конфигураций комплексов. На поверхности потенциальной энергии водных кластеров есть множество локальных минимумов, число которых быстро растет с увеличением размера кластера, таким образом делая поиск глобального минимума вычислительно затратной работой. За последние годы был достигнут значительный прогресс в изучении структуры кластеров воды. Для понимания характеристик водородных связей теоретически изучено множество различных конфигураций водных кластеров.

Несмотря на то, что некоторые из расчетов [4, 5, 6] предполагают наиболее стабильной циклическую и призматическую структуры. В последнее время стало ясно, что большое число альтернативных трехмерных структур, таких, как стул, лодка, и клетки, вероятно, будут сравнимы по энергии [7, 8-10].

С увеличением размера кластера, число возможных структур быстро увеличивается, и становится трудно исследовать все возможные конформеры. Таким образом, мы решили исследовать возможность последовательного присоединения молекул воды с образованием двухрядной цепочки [11].

### **Цели и задачи выпускной квалификационной работы**

1. Расчет структурных и термодинамических параметров комплементарных пар оснований ДНК и РНК аденин-тимин (А-Т), аденин-урацил (А-У) и гуанин-цитозин (Г-Ц) с учетом базисной суперпозиционной ошибки методом функционала плотности в приближении гибридного функционала B3LYP, и ряда функционалов

(M062X, WB97XD, B97D) специально оптимизированных для расчетов Ван-дер-Ваальсовых взаимодействий.

2. Исследование поведения в структуре и термодинамике при последовательном образовании водного кластера в виде двухрядной цепочки.

### **Научная новизна работы:**

Впервые показана возможность применения некоторых функционалов, оптимизированных для расчетов слабых Ван-дер-Ваальсовых взаимодействий, к описанию водородо-связанных комплементарных пар оснований ДНК.

Впервые показано, что при наращивании водного молекулярного кластера в конфигурации двухрядная цепочка начиная с 17 молекул происходит изменение структуры кластера, которое выражается в сворачивании одного конца комплекса в призмоподобную конфигурацию.

### **Практическая значимость результатов**

Вычисленные значения молекулярных и термодинамических параметров комплементарных пар оснований ДНК и молекулярных кластеров воды могут быть использованы в дальнейшем при исследовании взаимодействия этих соединений с различными молекулярными системами.

### **Достоверность результатов работы**

Достоверность полученных результатов и выводов обеспечивается адекватностью используемых физических моделей и квантово-механических вычислительных методов.

### **Основные результаты и положения, выносимые на защиту**

1. Энтальпии образования комплементарных пар оснований ДНК, РНК и молекулярных кластеров воды, вычисленные с учетом базисной суперпозиционной ошибки.

2. Относительно большое значение константы равновесия комплекса гуанин-цитозин обусловлено значительным преобладанием энтальпийного фактора над энтропийным.
3. При наращивании молекулярного кластера начиная с 17 молекул происходит изменение структуры кластера, которое выражается в сворачивании одного конца комплекса в призмоподобную конфигурацию
4. Наиболее энергетически выгодными являются кластеры с числом молекул кратным четырем.

## **Структура и объём ВКР**

Магистерская выпускная квалификационная работа состоит из введения, двух глав, заключения и списка литературы, состоящего из 48 наименований. Материалы работы изложены на 31 странице, содержащих 11 рисунков.

## **Содержание**

### **Термодинамика связывания в комплементарных парах ДНК и РНК**

#### *Пары А-Т и А-У*

Пространственное строение комплементарных пар А-Т и А-У показано на рисунке 1. Как видно из рисунка 1, эти пары стабилизируются посредством двух водородных связей, показанных пунктирными линиями.

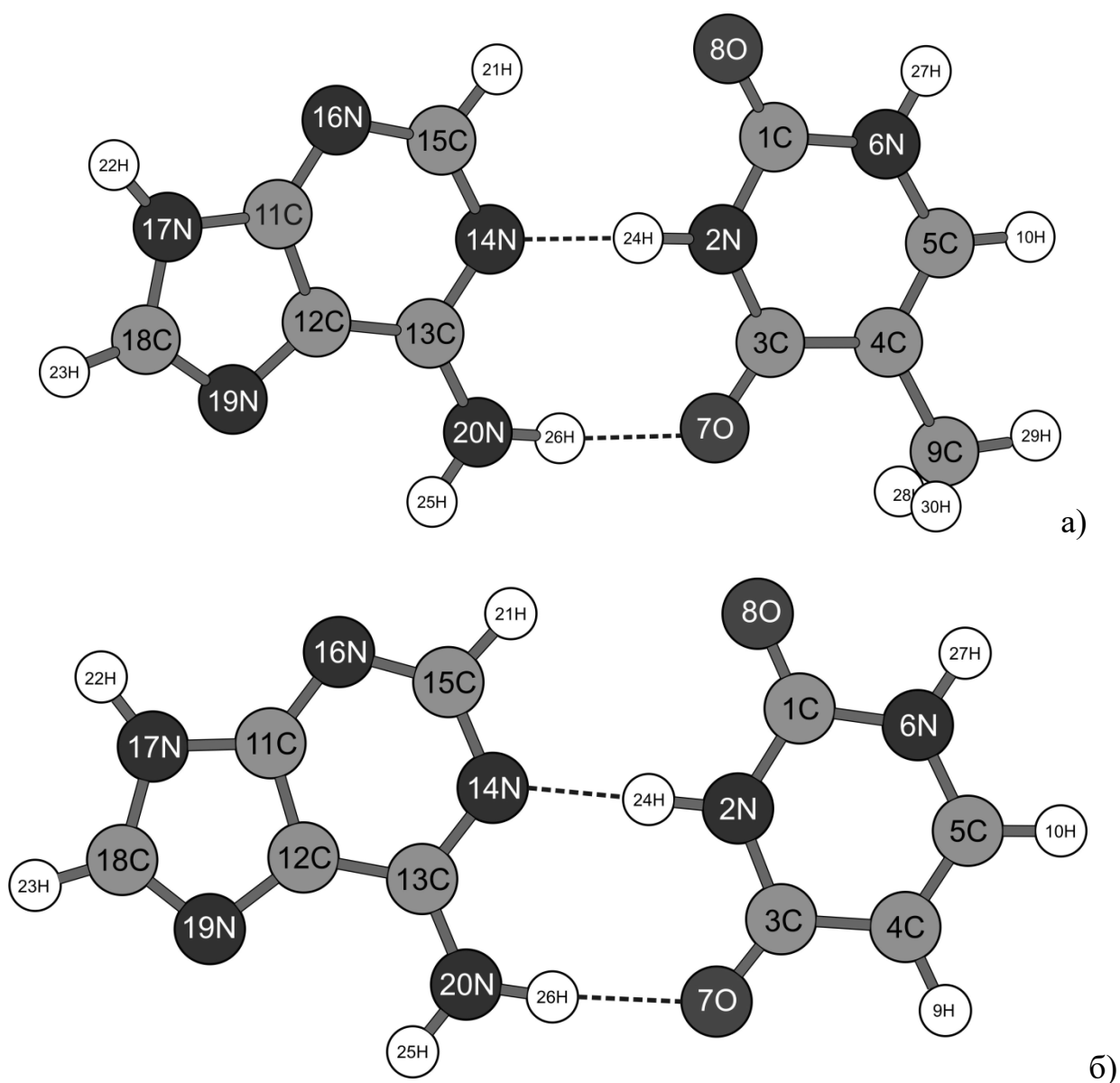


Рис. 1 – Пространственное строение комплексов: а) аденин-тимин и б) аденин-урацил

Структура этих комплексов практически одинакова, так как тимин является метил замещенным урацилом. Исходя из этого, можно ожидать, что такие характеристики, как межмолекулярные расстояния и энтальпия комплексообразования, будут близки между собой.

Термодинамические характеристики комплексообразования для пар А-Т и А-У представлены в таблице 1.

Таблица 1 – Термодинамические характеристики (кДж/моль) образования водородосвязанных комплексов аденин-тимин и аденин-урацил

|               | $\Delta E$ | BSSE | $\Delta ZPE$ | T $\Delta S$ | $\Delta G$ | $\Delta H_0$ | $\Delta H_{298}$ | k     |
|---------------|------------|------|--------------|--------------|------------|--------------|------------------|-------|
| A-T           |            |      |              |              |            |              |                  |       |
| Exp. [14]     |            |      |              |              |            |              | -50,63           |       |
| B3LYP         |            |      |              |              |            |              |                  |       |
| 6-311+G(d,p)  | -52.92     | 3.00 | 5.48*        | -49.64       | 5.60       | -44.45       | -44.04           | 0.10  |
| 6-311++G(d,p) | -52.96     | 3.04 | 5.51*        | -52.32       | 8.21       | -44.40       | -44.10           | 0.04  |
| aug-cc-pVDZ   | -54.80     | 4.85 | 6.16         | -49.87       | 6.32       | -43.80       | -43.55           | 0.08  |
| M062X         |            |      |              |              |            |              |                  |       |
| 6-311+G(d,p)  | -62.28     | 3.46 | 4.23         | -47.70       | -6.18      | -54.59       | -53.58           | 12.1  |
| WB97XD        |            |      |              |              |            |              |                  |       |
| 6-311+G(d,p)  | -68.13     | 3.28 | 5.17         | -47.41       | -11.52     | -59.67       | -58.93           | 104.1 |
| B97D          |            |      |              |              |            |              |                  |       |
| 6-311+G(d,p)  | -65.04     | 3.45 | 3.91         | -46.56       | -10.17     | -57.67       | -56.73           | 60.60 |
| A-Y           |            |      |              |              |            |              |                  |       |
| B3LYP         |            |      |              |              |            |              |                  |       |
| 6-311+G(d,p)  | -53.60     | 2.94 | 5.54*        | -49.46       | 4.68       | -45.12       | -44.78           | 0.15  |
| 6-311++G(d,p) | -53.64     | 2.99 | 5.60*        | -52.19       | 7.35       | -45.05       | -44.84           | 0.05  |
| aug-cc-pVDZ   | -55.69     | 4.87 | 6.26         | -49.89       | 5.47       | -44.56       | -44.43           | 0.11  |
| M062X         |            |      |              |              |            |              |                  |       |
| 6-311+G(d,p)  | -62.85     | 3.41 | 4.21         | -46.98       | -7.53      | -55.23       | -54.52           | 20.86 |
| WB97XD        |            |      |              |              |            |              |                  |       |
| 6-311+G(d,p)  | -68.55     | 3.25 | 5.20         | -47.24       | -12.17     | -60.09       | -59.42           | 135.6 |
| B97D          |            |      |              |              |            |              |                  |       |

|              |        |      |      |        |        |        |        |       |
|--------------|--------|------|------|--------|--------|--------|--------|-------|
| 6-311+G(d,p) | -65.50 | 3.38 | 4.00 | -46.48 | -10.80 | -58.13 | -57.28 | 77.93 |
|--------------|--------|------|------|--------|--------|--------|--------|-------|

\*При расчете нулевой колебательной энергии вычисленные значения волновых чисел корректировались по методике линейного масштабирования [12].

Как видно из таблицы 1 и таблицы 2, суперпозиционная ошибка не превысила 9% для пар А-Т, А-У и 6процентов для Г-Ц, что говорит о применимости данных базисных наборов для оценки термодинамических характеристик комплексов такого типа. Наименьшая величина базисной суперпозиционной ошибки получилась для базисного набора 6-311+G(d,p). Результаты полученные в рамках метода V3LYP с базисными наборами более широкими, чем 6-311+G(d,p) не дают улучшения согласия вычисленных параметров с экспериментальными как для величин межмолекулярных расстояний, так и для энтальпии комплексообразования.

### Пара Г-Ц

Пространственное строение комплементарной пары Г-Ц показано на рисунке 2. Как видно из рисунка 2, данный комплекс в отличие от предыдущих стабилизируется уже посредством трех водородных связей.

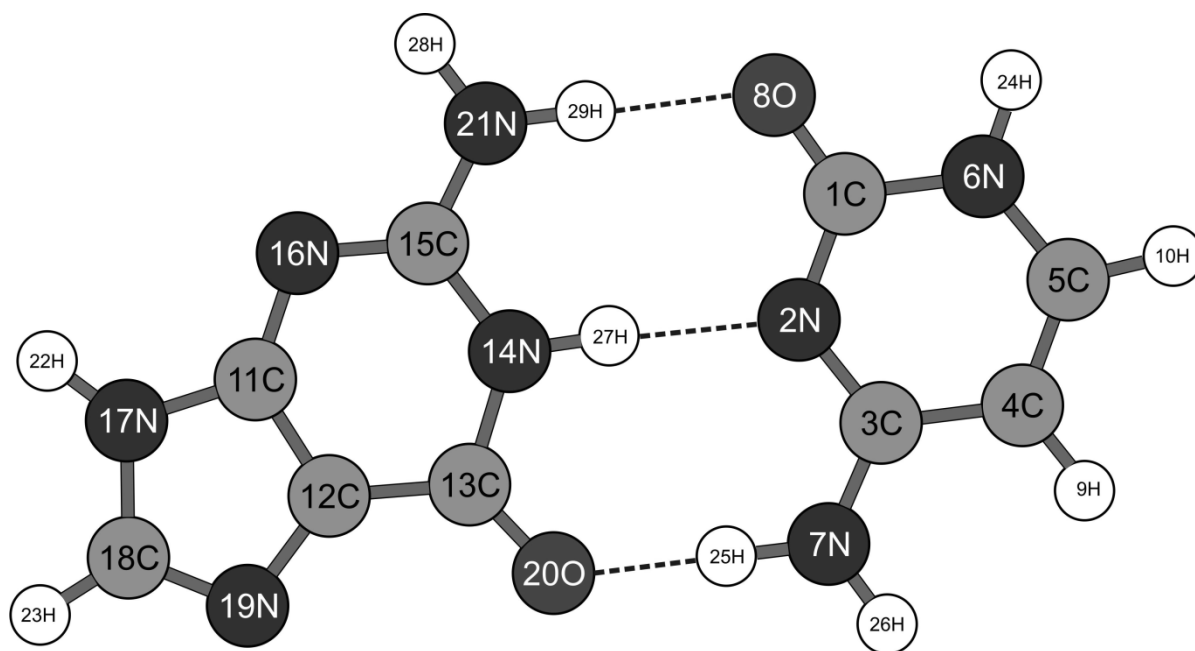


Рисунок 2 – Пространственное строение комплекса гуанин-цитозин



В рамках метода B97D/6-311+G(d,p) нам не удалось вычислить равновесную электронную конфигурацию этого комплекса, что подтверждалось наличием в значениях волновых чисел одного отрицательного значения, которое соответствует деформационному межмолекулярному колебанию.

Лучшее согласие с экспериментом получено в рамках методов B3LYP/6-311+G(d,p)/6-311++G(d,p) и M062X/6-311+G(d,p), при этом лучшего всего описывается расстояние 14N-2N. Коррекция значений на базисную суперпозиционную ошибку в большинстве случаев улучшает согласие с экспериментальными данными.

Значение термодинамических параметров комплексообразования пары Г-Ц показаны в таблице 2.

Таблица 2 – Термодинамические характеристики (кДж/моль) образования водородо-связанного комплекса Г-Ц

|               | $\Delta E$ | BSSE | $\Delta ZPE$ | $\Delta S$ | $\Delta G$ | $\Delta H_0$ | $\Delta H_{298}$ | k                   |
|---------------|------------|------|--------------|------------|------------|--------------|------------------|---------------------|
| Exp. [13]     |            |      |              |            |            |              | -87.87           |                     |
| B3LYP         |            |      |              |            |            |              |                  |                     |
| 6-311+G(d,p)  | -107.61    | 3.98 | 7.20*        | -49.82     | -46.72     | -95.45       | -95.64           | $10.6 \times 10^7$  |
| 6-311++G(d,p) | -107.62    | 4.02 | 6.58*        | -48.26     | -48.54     | -97.02       | -96.80           | $31.9 \times 10^7$  |
| aug-cc-pVDZ   | -110.12    | 6.01 | 9.73         | -55.85     | -40.01     | -94.38       | -95.87           | $1.02 \times 10^7$  |
| M062X         |            |      |              |            |            |              |                  |                     |
| 6-311+G(d,p)  | -118.03    | 4.48 | 5.85         | -45.71     | -61.68     | -107.7       | -107.39          | $6392 \times 10^7$  |
| WB97XD        |            |      |              |            |            |              |                  |                     |
| 6-311+G(d,p)  | -126.39    | 4.35 | 7.89         | -51.84     | -63.00     | -114.16      | -114.84          | $10896 \times 10^7$ |
| B97D**        |            |      |              |            |            |              |                  |                     |
| 6-311+G(d,p)  | -117.53    | 4.46 | -            | -          | -          | -            | -                | -                   |

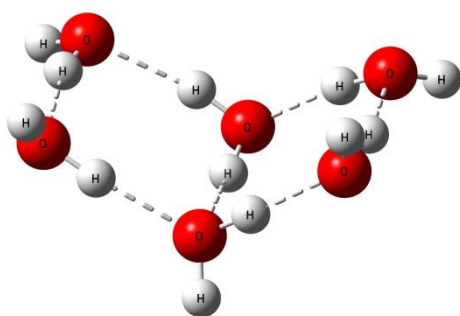
\*При вычислении значений нулевой колебательной энергии волновые числа масштабировались по методике [12].

\*\*Для данного метода приведены только те параметры, в расчете которых не используются значения волновых чисел.

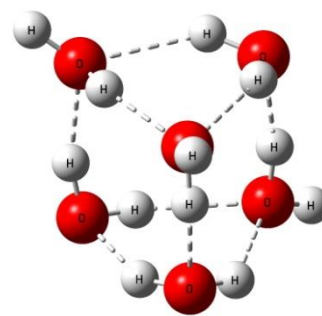
Как видно из таблицы 2, вычисленные в рамках всех теоретических методов значения энтальпии комплексообразования пары Г-Ц оказались завышенными относительно экспериментального значения. Наименьшая ошибка составляет около 9 процентов для функционала B3LYP. Если сравнить между собой термодинамические параметры образования комплексов А-Т и Г-Ц, то сразу видно, что энтальпия образования последнего значительно выше по абсолютному значению, а энтропийный фактор в обоих случаях примерно одинаков. Таким образом, мы имеем для пары Г-Ц большую отрицательную величину свободной энергии Гиббса, что в свою очередь приводит к большому значению константы равновесия.

### **Структурные параметры и термодинамика образования молекулярных кластеров воды**

На рисунке 6 приведены структуры двух гексамеров молекул воды в конфигурации кресло и призма. Структура гексамера в конфигурации кресло получена разрывом связи пентмера такой же конфигурации и добавлением одной молекулы воды. Структура гексамера призмоподобной конфигурации получена соответствующим размещением двух тримеров молекул воды друг над другом.



а)



б)

Рисунок 3 – Структуры гексамера молекул воды в конфигурациях кресло а) и призма б)  
вычисленные методом DFT/B3LYP/6-311+G(d,p)

Дальнейшее увеличение цепочки шло последовательным добавлением молекул, аналогично пентамеру и гексамеру в конфигурации кресло. По мере наращивания кластера в конфигурации кресло происходит постепенный изгиб всей структуры комплекса. Конфигурация кресло сохранялась вплоть до гексадекамера.

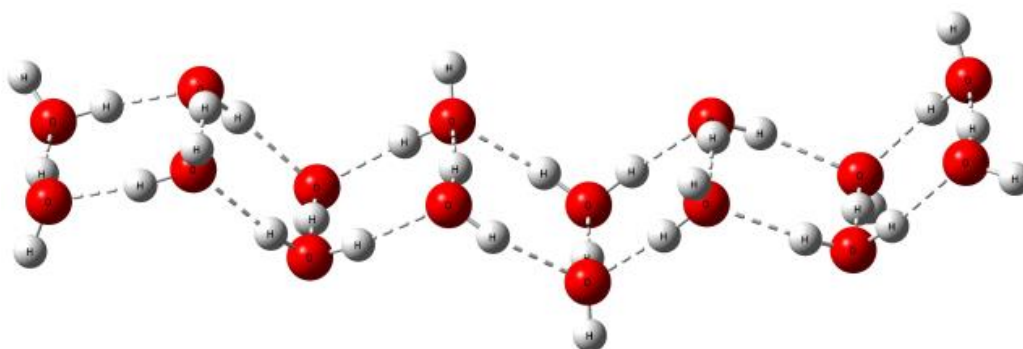


Рисунок 4 – Структура гексадекамера молекул воды вычисленная методом DFT/B3LYP/6-311+G(d,p)

При присоединении к гексадекамеру молекулы воды, одна сторона комплекса в процессе оптимизации геометрической структуры свернулась в призмоподобную конфигурацию гексамера и дальнейшее наращивание комплекса производилось со свободной стороны, противоположной гексамерному окончанию.

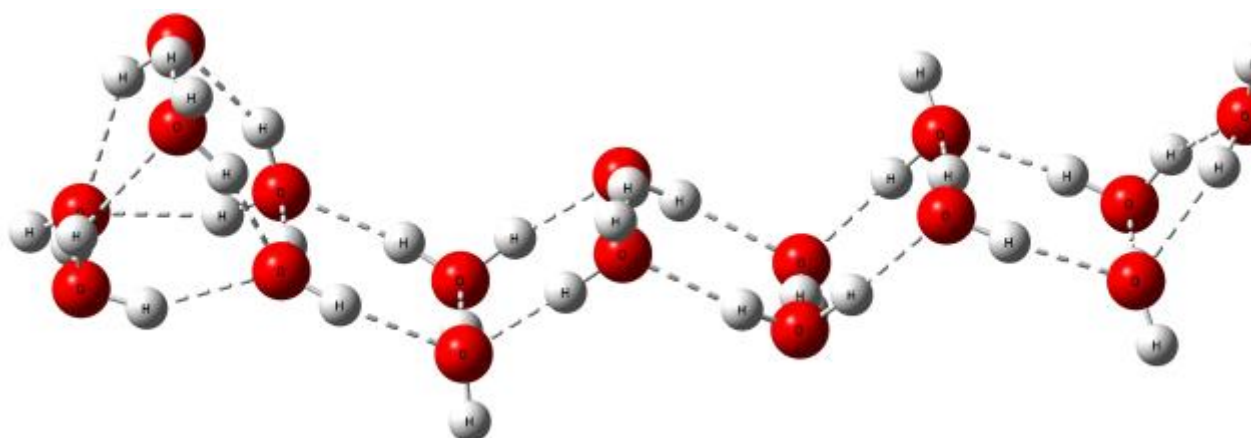


Рисунок 5 – Структура гептадекамера молекул воды вычисленная методом DFT/B3LYP/6-311+G(d,p)

Из рисунка 6 видно, что изменение энтальпии при изменении количества молекул воды в комплексе носит периодический характер. Димер и комплексы с нечетным количеством молекул самые высокоэнергетические, исключением является тример молекул воды. Стоит отметить, что наименьшая энтальпия комплексообразования получена, для первого комплекса с гексамерным окончание кластера. Комплексы с четным количеством молекул (кроме димера) лежат по энергии ниже. Самые низкоэнергетические комплексы были получены с количеством молекул кратным четырем.

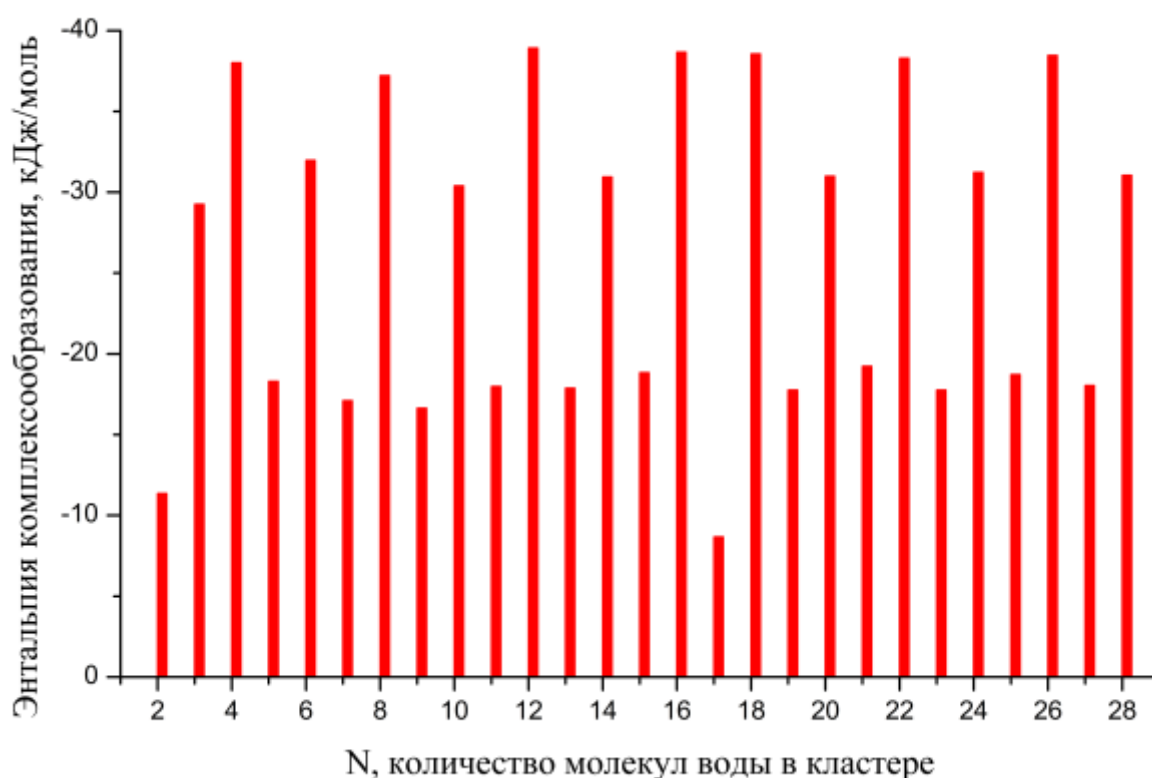


Рисунок 6 – Изменение вычисленных энтальпий образования комплексов в зависимости от числа молекул воды в них

На графике (рисунок 7) отображено изменение дипольных моментов комплексов при изменении их конфигурации. Изменение дипольных

моментов имеет также периодический характер для комплексов с количеством молекул меньше семнадцати. Видно, что минимальный дипольный момент имеют комплексы с количеством молекул кратным четырем ( $N = 4, 8, 12, 16$ ), так как это наиболее симметричные комплексы. Однако для них наблюдается возрастание дипольного момента с ростом числа  $N$ . Так гексадекамер имеет уже заметно изогнутую конфигурацию. Димер и кластеры с числом молекул 6, 10, 14 и имеют максимальный дипольный момент 3,17Д, 2,53Д, 2,72Д, 2,59Д соответственно. Дипольный момент комплексов с числом молекул 7, 11, 15 растет с увеличением их числа. Для комплексов с числом молекул от семнадцати до двадцати восьми возможно наличие более сложной периодической зависимости, кратной восьми.

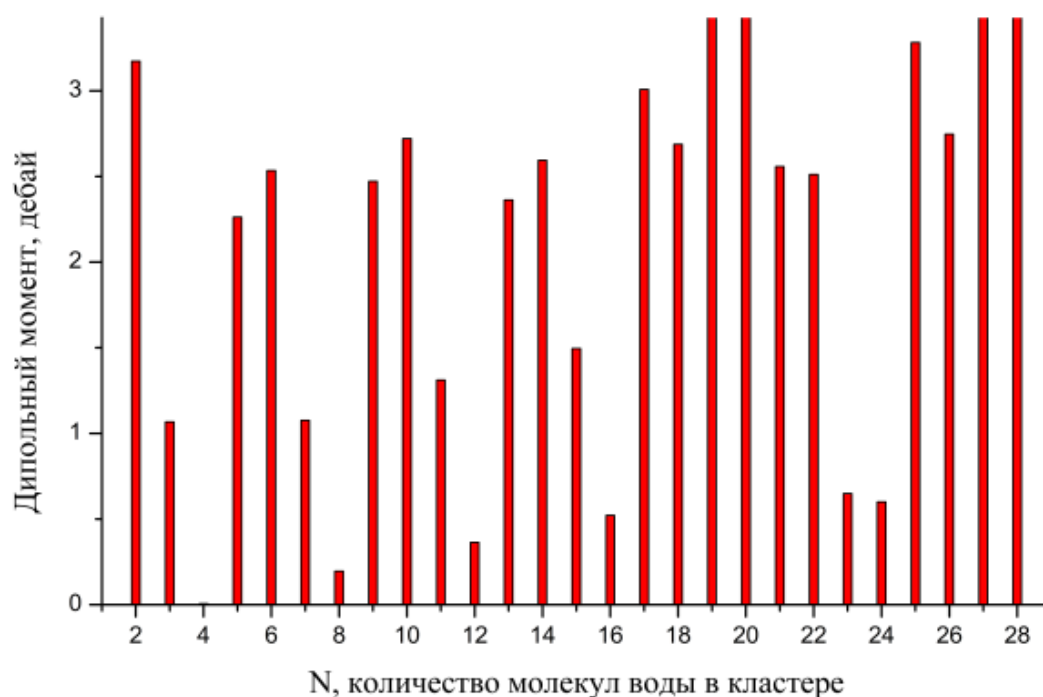


Рисунок 7 – Вычисленные методом DFT/B3LYP/6-311+G(d,p)  
дипольные моменты комплексов

Изменение энтальпии, приходящиеся на одну водородную связь ( $\Delta H_H$ ) носит периодический характер в зависимости от четности, комплексы с четным количеством молекул низкоэнергетические. Сначала рассмотрим

комплексы с числом молекул менее семнадцати. Наибольшая энтальпия приходящиеся на одну водородную связь у тетрамера, далее она постепенно уменьшается для комплексов с номерами 6, 8, 10, -23.2 кДж/моль, -21.8 кДж/моль, -21.6 кДж/моль, -21.1 кДж/моль соответственно, далее  $\Delta H_H$  для комплексов с номерами 12, 14, 16 практически не изменяется.

Рассмотрим комплексы с числом молекул от 17 до 28. Энтальпии, приходящиеся на одну водородную связь для комплексов с четным количеством молекул, лежит в пределах -19.1 кДж/моль – -19.5 кДж/моль.  $\Delta H_H$  для комплексов с нечетным количеством молекул незначительно выше и лежит в пределах -18.5 кДж/моль – -19.1 кДж/моль.

## Основные результаты

1. Методами функционала плотности, с использованием функционала B3LYP с базисными наборами 6-311+G(d,p), 6-311++G(d,p), aug-cc-pVDZ и функционалов M062X, WB97XD, B97D с базисным набором 6-311+G(d,p) проведен теоретический анализ структурных и термодинамических параметров канонических комплементарных пар оснований аденин-тимин, аденин-урацил, гуанин-цитозин.
2. Показано, что учет базисной суперпозиционной ошибки в большинстве случаев улучшает согласие вычисленных значений с соответствующими экспериментальными.
3. Сделан вывод о возможности применения функционалов, оптимизированных для расчетов слабых Ван-дер-Ваальсовых взаимодействий, к описания водородо-связанных комплексов. Исключение составляет функционал B97D, в рамках которого не удалось получить оптимальную конфигурацию для пары Г-Ц.
4. Установлено, что плоская относительно пиримидинового кольца конфигурация аминогруппы в молекуле цитозина получается в рамках метода DFT/B3LYP начиная с базисного набора 6-311++G(df,p).

5. Показано, что стабилизация комплекса гуанин-цитозин обусловлена значительным преобладанием энтальпийного фактора над энтропийным.
6. Выявлена периодическая зависимость изменения энтальпии комплексообразования от числа молекул воды в кластере.
7. Установлено, что наиболее энергетически выгодными являются кластеры с числом молекул кратным четырем.
8. При наращивании молекулярного кластера начиная с 17 молекул происходит изменение структуры кластера, которое выражается в сворачивании одного конца комплекса в призмоподобную конфигурацию.
9. В силу наибольшей симметрии минимальный дипольный момент имеют комплексы с количеством молекул кратным четырем ( $N = 4, 8, 12, 16$ ).

#### **Список используемых источников**

1. В.Зенгер / Принципы структурной организации нуклеиновых кислот М.: Мир, 1987. - 584с.
2. P. Yakovchuk, E. Protozanova, M.D. Frank-Kamenetskii/ Base-stacking and base-pairing contributions into thermal stability of the DNA double helix / /Nucleic Acids Research, 2006, Vol.34, № 2, P.564-574.
3. J.Sponer, P.Jurecka, P.Hobza / Accurate Interaction Energies of Hydrogen-Bonded Nucleic Acid Base Pairs // J.Am.Chem.Soc., 2004, V.126, № 32, P.10142-10151.
4. Laasonen K., Parinello M., Car R., Lee C., Vanderbilt D. Structures of Small Water Clusters using Gradient-Corrected Density Functional Theory // Chem. Phys. Lett. – 1993. - № 207. – С. 208.
5. Krishnan P. N., Jensen J. O., Burke L. A. Theoretical study of water clusters. II. Hexamer //Chem. Phys. Lett. – 1994. - № 217. – С. 311-318.
6. Tsai C. J., Jordan K. D. Theoretical study of the  $(\text{H}_2\text{O})_6$  cluster //Chem. Phys. Lett. – 1993. - № 213. – С. 181-188.
7. Kim K. S., Dupuis M., Lie G. C., Clementi E. Revisiting small clusters of water molecules //Chem. Phys. Lett. – 1986. - № 131. – С. 451-456.

8. Kim K. S., Jordan K. D., Zwier T. S. Low-Energy Structures and Vibrational Frequencies of the Water Hexamer: Comparison with Benzene-(H<sub>2</sub>O)<sub>6</sub> //J. Am. Chem. Soc. – 1994. - № 116. – C. 11568-11569.
9. Mhin B. J., Kim J., Lee S., Lee J. Y., Kim K. S. What is the global minimum energy structure of the water hexamer? The importance of nonadditive interactions //Chem. Phys. – 1994. - № 100. – C. 4484-4487.
10. Lee C., Chen H., Fitzgerald G. J. Structures of the water hexamer using density functional methods //Chem. Phys. – 1994. - № 101. – C. 4472-4476.
11. Shruti Maheshwary, Nitin Patel, and Narayanasami, Sathyamurthy Structure and Stability of Water Clusters (H<sub>2</sub>O)<sub>n</sub>, *n* = 8-20: An Ab Initio Investigation // *J. Phys. Chem. A* 2001, 105, 10525-10537.
12. Yoshida H., Ehara A., Matsuura H. Density functional vibrational analysis using wavenumber-linear scale factors. // Chem. Phys. Lett. 2000. V. 325, № 4. P. 477–483.
13. Célia Fonseca Guerra, F. Matthias Bickelhaupt, Jaap G. Snijders, and Evert Jan Baerends. Hydrogen Bonding in DNA Base Pairs: Reconciliation of Theory and Experiment// J. Am. Chem. Soc. 2000, 122, p. 4117-4128.