

## ВЛИЯНИЕ ГИДРАТАЦИИ НА ТАУТОМЕРИЮ ОСНОВАНИЙ НУКЛЕИНОВЫХ КИСЛОТ: КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ МЕТОДОМ МОНТЕ-КАРЛО

В.И. Данилов

Институт молекулярной биологии и генетики, Национальная Академия Наук Украины,  
252143, Киев, ул. Академика Заболотного, 150

Поступила в редакцию 17 ноября 1999 г.

Проведено экстенсивное компьютерное моделирование гидратации оснований нуклеиновых кислот и их основных таутомерных форм методом Монте-Карло. Показано, что в воде лактамная форма гуанина и аминная форма цитозина энергетически более предпочтительны, чем их лактимная и иминная формы соответственно. Большая стабильность в водном кластере обычных форм этих оснований по сравнению с их редкими таутомерными формами целиком обусловлена энергетически более выгодным взаимодействием молекул воды друг с другом. В то же время обнаружено, что иминная форма аденина и лактимная формы тимина и урацила более стабильны, чем соответствующие аминная и лактамная формы. Этот факт определяется более выгодным взаимодействием редких форм оснований с водой.

**КЛЮЧЕВЫЕ СЛОВА:** нуклеотидные основания, таутомерия, гидратация, метод Монте-Карло, компьютерное моделирование

Возникновение точечных спонтанных мутаций частично связывают с существованием редких таутомерных форм пуриновых и пиримидиновых оснований нуклеиновых кислот и, как следствие этого, образованием ошибочных (не уотсон-криковских) пар оснований. Поэтому весьма важным является знание относительных стабильностей таутомерных форм нуклеотидных оснований и родственных модельных систем.

Для того, чтобы оценить относительные стабильности таутомерных форм оснований в последние годы было выполнено большое количество экспериментальных и теоретических работ (см., например, обзоры [1-4]). Теоретические исследования были проведены для изолированных оснований (оснований в вакууме), т.е. в отсутствие межмолекулярного взаимодействия между основаниями и их окружением. Результаты таких расчетов следует сравнивать с данными экспериментального изучения ИК-спектров оснований, изолированных в инертных матрицах [3]. В то же время подавляющая часть экспериментальных данных была получена в растворах и твердой фазе [1]. В этом случае изучение действия окружения на относительные стабильности таутомеров обычно проводилось в рамках взаимодействия полярной системы с непрерывной диэлектрической средой (см., например, [5]).

Однако для исследования стабильности таутомеров в растворах необходимо изучить взаимодействия растворитель-растворитель, растворитель-растворенное вещество и растворенное вещество-растворенное вещество на молекулярном уровне. Правильное описание влияния сольватации на таутомерию оснований должно рассматривать взаимодействие между молекулами растворителя и растворенного вещества в явном виде. Таких теоретических данных до настоящего времени не имеется.

### МАТЕРИАЛЫ И МЕТОДЫ

В связи с этим мы провели компьютерное моделирование гидратации пуриновых и пиримидиновых оснований Ade, Gua, Cyt, Ura, Thy и их основных таутомерных форм. Была изучена амино-иминная и лактам-лактимная таутомерия этих соединений. При этом были рассмотрены оба возможных положения имино- и гидроксильной групп по отношению к атому N(1) пуриновых и атому N(3) пиримидиновых оснований - *anti*- и *syn*-конформации. В дальнейшем эти таутомерные формы для оснований обозначаются как *imino (anti)*, *imino (syn)*, *lactim (anti)*, *lactim (syn)*.

Расчеты были выполнены методом Монте-Карло в каноническом (T, V, N) ансамбле, используя выборку Метрополиса [6]. В каждой изученной системе (основание + водный кластер) число молекул воды составляло 400. Температура системы была равной 298K. В качестве граничных условий мы воспользовались кластерным приближением Абрахама и др. [7-10], так как оно хорошо описывает образование ядра жидкости и жидких растворов из пара.

Для вычисления энергий межмолекулярного взаимодействия были использованы полуэмпирические атом-атомные потенциальные функции, предложенные Полтевым и его коллегами (см. сводку данных в работах [11-13]). В соответствии с этими функциями заряды на различных атомах оснований и их таутомерных формах вычислены так же, как в наших недавних работах [11-13]. При проведении расчетов для всех изученных молекул была использована

геометрия, полученная с помощью многоконфигурационной теории самосогласованного поля с оптимизацией геометрии в рамках гамильтониана метода AM1 (см. [14]). Атомы водородов метильных групп были учтены в явном виде.

Первоначально радиус сферы был равен радиусу сферы для найденной начальной конфигурации водного кластера вокруг каждой молекулы растворенного вещества ( $R_c=23.0$  Е). Для создания в такой сфере «водоподобной» структуры были проведены вычисления, включающие  $2.4 \times 10^6$  конфигураций. Затем согласно кластерной теории [7-10] радиус сферы для рассматриваемых систем был увеличен до 25.0 Е и выполнены дальнейшие вычисления, которые содержали  $2.4 \times 10^6$  конфигураций. Таким образом, для того, чтобы уравновесить каждую из систем, были использованы  $4.8 \times 10^6$  конфигураций, которые при вычислении средних свойств были отброшены.

В наших вычислениях статистическая ошибка (значение дисперсии) была вычислена с точностью  $\pm 0.005$ . Для того, чтобы при вычислении термодинамических данных достичь указанной точности, длины генерируемых марковских цепей для изучаемых систем (основание + водный кластер) были следующими: Ade –  $1.6 \times 10^7$  конфигураций, Ade imino (*anti*) –  $1.68 \times 10^7$  конфигураций, Ade imino (*syn*) –  $1.76 \times 10^7$  конфигураций, Gua –  $2.08 \times 10^7$  конфигураций, Gua imino (*anti*) –  $1.72 \times 10^7$  конфигураций, Gua imino (*syn*) –  $1.68 \times 10^7$  конфигураций, Cyt –  $2.32 \times 10^7$  конфигураций, Cyt imino (*anti*) –  $1.32 \times 10^7$  конфигураций, Cyt imino (*syn*) –  $2.04 \times 10^7$  конфигураций, Thy –  $1.44 \times 10^7$  конфигураций, Thy imino (*anti*) –  $1.76 \times 10^7$  конфигураций, Thy imino (*syn*) –  $1.2 \times 10^7$  конфигураций, Ura –  $2.44 \times 10^7$  конфигураций, Ura imino (*anti*) –  $1.96 \times 10^7$  конфигураций, Ura imino (*syn*) –  $2.28 \times 10^7$  конфигураций, вода –  $2.64 \times 10^7$  конфигураций.

Все расчеты были выполнены в Киеве на группе рабочих станций UltraSparc-II Ракового института им. Розвел парк в Баффало (США) с помощью глобальной сети INTERNET.

### РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Результаты расчетов средних значений потенциальной энергии системы  $U$ , энергии взаимодействия вода-вода  $U_{ww}$  и энергии взаимодействия вода-основание  $U_{wb}$  даны в табл. 1. Величины энергии гидратации  $\Delta U_{hyd}$  для всех систем, полученные из приведенных данных, также представлены в этой таблице. Здесь же приведены значения статистической ошибки (стандартного отклонения) для  $U$  и  $\Delta U_{hyd}$ .

Таблица 1. Энергетические характеристики гидратации оснований нуклеиновых кислот и их основных таутомерных форм<sup>a</sup>

Основание	$U^b$	$U_{ww}^b$	$U_{wb}^b$	$U_{hyd}^c$
Ade	$-8.331 \pm 0.005$	-8.091	-0.240	$-43.6 \pm 2.8$
Ade imino ( <i>anti</i> )	$-8.374 \pm 0.005$	-8.053	-0.321	$-60.8 \pm 2.8$
Ade imino ( <i>syn</i> )	$-8.429 \pm 0.005$	-8.102	-0.327	$-82.8 \pm 2.8$
Gua	$-8.407 \pm 0.005$	-8.111	-0.296	$-74.0 \pm 2.8$
Gua lactim	$-8.381 \pm 0.005$	-8.085	-0.296	$-63.6 \pm 2.8$
Gua lactim ( <i>syn</i> )	$-8.390 \pm 0.005$	-8.083	-0.307	$-67.2 \pm 2.8$
Cyt	$-8.341 \pm 0.005$	-8.128	-0.213	$-47.6 \pm 2.8$
Cyt imino ( <i>anti</i> )	$-8.301 \pm 0.005$	-8.087	-0.214	$-31.6 \pm 2.8$
Cyt imino ( <i>syn</i> )	$-8.333 \pm 0.005$	-8.118	-0.215	$-44.4 \pm 2.8$
Thy	$-8.321 \pm 0.005$	-8.177	-0.144	$-39.6 \pm 2.8$
Thy lactim ( <i>anti</i> )	$-8.335 \pm 0.005$	-8.104	-0.231	$-45.2 \pm 2.8$
Thy lactim ( <i>syn</i> )	$-8.342 \pm 0.005$	-8.130	-0.212	$-48.0 \pm 2.8$
Ura	$-8.333 \pm 0.005$	-8.194	-0.139	$-44.4 \pm 2.8$
Ura lactim ( <i>anti</i> )	$-8.383 \pm 0.005$	-8.133	-0.250	$-64.4 \pm 2.8$
Ura lactim ( <i>syn</i> )	$-8.376 \pm 0.005$	-8.162	-0.214	$-61.6 \pm 2.8$
Water	$-8.222 \pm 0.005$	-8.222	–	–

<sup>a</sup>Некоторые отличия энергетических характеристик обычной формы оснований от аналогичных данных, полученных в [11, 12], обусловлены использованием в данной работе иной геометрии (см. выше).

<sup>b</sup>Ккал/моль растворителя. <sup>c</sup>Ккал/моль системы.

Видно, что величины энергии гидратации для всех изученных систем отрицательны. Это находится в согласии с экспериментальными данными по энтальпии гидратации обычных таутомерных форм оснований, которые легко могут быть вычислены из измеренных теплот растворения и сублимации [15-21].

Сравнение теоретических и экспериментальных значений  $\Delta U_{hyd}$  для А, U, Т, G и С (см. Табл. 1 и данные [15-21]) показывает неудовлетворительное согласие между ними. Экспериментальные значения существенно ниже значений, приведенных в Табл. 1.

Можно предположить, что потенциальные функции, которые мы использовали для описания взаимодействия вода-растворенное вещество, приводят к переоценке этого взаимодействия. Другими словами, основания нуклеиновых кислот гидратируются в водном кластере намного сильнее, чем в реальной воде. Тогда следует ожидать, что в более реалистической воде член, ответственный за взаимодействие вода-основание, будет несколько меньше.

Однако параметры этих потенциальных функций были выбраны так, чтобы длины Н-связей между молекулой воды и основаниями и экспериментальные положения молекул воды в кристаллогидратах оснований, нуклеозидов, нуклеотидов и родственных соединений могли быть воспроизведены. В результате потенциальные функции для взаимодействия вода - растворенное вещество позволили нам найти схемы моногидратации нуклеотидных оснований и их пар и значения энергии взаимодействия вода-основание [22, 23], которые очень близки к значениям, полученным с помощью *ab initio* атом-атомных потенциальных функций [24, 25] и уточненных полуэмпирических потенциальных функций [26]. Эти потенциальные функции также позволили нам получить значения энтальпии образования моно-, ди- и тригидратов Ade [27] и метилпроизводных Uga и Thy [28], которые находятся в количественном согласии с экспериментальными данными, определенными в вакууме на основании масс-спектрометрического изучения [15, 16].

В то же время следует подчеркнуть, что моделирование гидратации метилпроизводных Uga и Thy в кластере, состоящем из 200 молекул воды [28], с теми же потенциальными функциями, что и в данной работе, правильно воспроизводит последовательность их экспериментальных значений  $\Delta U_{\text{solv}}$ . Аналогичное моделирование гидратации кофеина [29, 30] показало, что значения энтальпии гидратации, вычисленные на основе теоретических и экспериментальных данных, находятся в хорошем согласии.

Несмотря на все сказанное взаимодействие вода-основание действительно может быть переоценено из-за недооценки взаимодействия вода-вода в кластерной теории, так как между этими взаимодействиями всегда существует конкуренция. Фактически, это было прямо подтверждено изучением гидратации обычных форм оснований методом Монте-Карло с использованием периодических граничных условий [31]. Применение в этой работе уточненных полуэмпирических атом-атомных потенциальных функций привело к энергетическим и структурным характеристикам жидкой воды и значениям энергии образования моногидратов оснований, которые подобны, полученным нами данным. В то же время это позволило авторам [31] вычислить энергии гидратации для Uga, Thy и Ade, которые хорошо согласуются с экспериментальными значениями.

Таким образом, различие между теоретическими и экспериментальными значениями энтальпии гидратации определяется не выбором потенциалов взаимодействия для систем вода-основание, а связано с использованием кластерной теории. Недооценка взаимодействия растворитель-растворитель, приводящая к переоценке взаимодействия растворитель-растворенное вещество, является причиной завышенных значений энергии гидратации для изученных оснований. Однако, так как дальнейшее рассмотрение основано на энергетических различиях, кажется невероятным, чтобы данные, приведенные в Табл. 1, могли существенно измениться.

Как видно из этой таблицы, из двух рассмотренных конформаций (*anti* и *syn*) каждой из редких форм более стабильной является *syn*-конформация. Исключение составляет Uga. Однако для Thy и Uga этот вывод недостоверен, поскольку разность энергий между *anti*- и *syn*-конформациями равна статистической ошибке.

Из этой же таблицы следует, что лактамная форма Gua и аминная форма Cyt энергетически более выгодны, чем каждая из конформаций лактимной и иминной таутомерных форм соответственно. Этот вывод для Cyt согласуется с результатом работы [32], в которой были вычислены разности свободных энергий между аминной и иминной таутомерными формами Cyt методом молекулярной динамики. Из Табл. 1 видно также, что большая стабильность в воде обычных форм этих оснований по сравнению с их редкими формами целиком обусловлена энергетическим членом  $U_{\text{ww}}$ , который описывает взаимодействие молекул воды друг с другом. Иначе говоря, вода вокруг обычных форм оснований Gua и Cyt более структурирована, чем вокруг редких форм.

В то же время обнаружено (см. Табл. 1), что в воде каждая из конформаций иминоформы Ade и лактимной формы Thy и Uga более стабильна, чем их amino- и лактамная формы соответственно. Это определяется более выгодным взаимодействием редких форм оснований с водой, т.е. членом  $U_{\text{wb}}$ .

Здесь следует отметить, что, большая (меньшая) выгодность энергетического члена  $U_{\text{ww}}$  и, следовательно, большая (меньшая) структурированность воды вокруг оснований не всегда обусловлены большим (меньшим) числом водородных связей, образуемых между молекулами воды. Об этом свидетельствует анализ функций распределения для воды, которые здесь не приводим.

Хотя сопоставление теоретических результатов по энергетической выгодности обычных и редких форм с экспериментальными данными невозможно из-за отсутствия последних для

большинства моделируемых систем, однако, как легко вычислить из Табл. 1, количественное значение энергетической разности значительно превосходит статистическую ошибку, возникающую при вычислениях методом Монте-Карло. Только в случае Cyt разность энергий между *суз*-конформацией его иминоформы и аминоконформой составляет 3.2 ккал/моль, что лишь незначительно превосходит статистическую ошибку, равную 2.8 ккал/моль.

#### СПИСОК ЛИТЕРАТУРЫ

1. Kwiatkowski J.S., Pulman B. // *Adv. Heterocyclic Chem.* 1975. V.18. P.199-335
2. Kwiatkowski J.S., Zielinski T.J., Rein R. // *Adv. Quant. Chem.* 1986. V.18. P.85-131
3. Nowak M. J., Lapinski L., Kwiatkowski J.S., Leszczynski J. // In "Computational Chemistry: Reviews of Current Trends". World Scientific Publication. 1997. V.II. P.140-216
4. Leszczynski J. // In "The Encyclopedia of Computational Chemistry", John Wiley & Sons. 1998
5. Tapia O., Silvi B. // *J. Phys. Chem.* 1980. V.84. P.2646-2652
6. Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.N., Teller E. // *J. Chem Phys.* 1953. V.21. P.1087-1092
7. Lee J.K., Barker J.A., Abraham F.F. // *J. Chem. Phys.* 1973. V.58. P.3166-3180
8. Abraham F.F., Lee J.K., Barker J.A. // *J. Chem. Phys.* 1974. V.60. P.246-250
9. Abraham F.F. // *J. Chem. Phys.* 1974. V.61. P.1221-1225
10. Mruzik M.R., Abraham F.F., Schreiber D.E., Pound G.M. // *J. Chem. Phys.* 1976. V.64. P. 481-491
11. Данилов В.И., Желтовский Н.В., Слюсарчук О.Н., Альдерфер Дж.Л. // *Биополимеры и клетка.* 1997. V.13. С.46-54
12. Danilov V.I., Zheltovsky N.V., Slyusarchuk O.N., Poltev V.I., Alderfer J.L. // *J. Biomol. Struct. Dyn.* 1997. V.15. P.69-80
13. Alderfer J.L., Danilov V.I., Poltev V.I., Slyusarchuk O.N. // *J. Biomol. Struct. Dyn.* 1999. V.16. P.1107-1117
14. В.И. Данилов, Дж.Л. Альдерфер // *Биополимеры и клетка.* 1999. Т.15. №6.
15. Sukhodub L.F., Yanson I.K., Shelkovski V.S., Wierzchowski K. // *Biophys. Chem.* 1982. V.15. P.149-155
16. Веркин Б.И., Янсон И.К., Суходуб Л.Ф., Теплицкий А.Б. *Взаимодействия биомолекул: Новые экспериментальные подходы и методы.* К. Наукова Думка, 1985. 164 с.
17. Teplitsky A.B., Yanson I.K., Glukhova O.T., Zielenkiewicz A., Zielenkiewicz W., Wierzchowski K.L. // *Biophys. Chem.* 1980. V.11. P.17-21
18. Глухова О.Т., Киселев В.Д., Теплицкий А.Б., Устюгов А.Н., Янсон И.К. // *Биофизика.* 1981. Т.26. С.351-352
19. Teplitsky A.B., Glukhova O.T., Sukhodub L.F., Yanson I.K., Zielenkiewicz A., Zielenkiewicz W., Kosinski J., Wierzchowski K.L. // *Biophys. Chem.* 1982. V.15. P.139-147
20. Zielenkiewicz A., Zielenkiewicz W., Sukhodub L.F., Glukhova O.T., Teplitsky A.B., Wierzchowski K.L. // *J. Solution. Chem.* 1984. V.13. P. 757-765
21. Kilday M.V. // *J. Res. Nat. Bur. Stand. (U.S.).* 1981. V.86. P. 367-382
22. Danilov V.I., Shvartsman A.S. // *Молекулярная генетика и биофизика.* 1982. N 7. С.28-35
23. Danilov V.I., Shestopalova A.V. // *Молекулярная генетика и биофизика.* 1984. N 9. С.47-54
24. Scordamaglia R., Cavallone F., Clementi E. // *J. Am. Chem. Soc.* 1977. V.99. С.5545-5550
25. Clementi E., Corongiu G. // *J. Chem. Phys.* 1980. V.72. P.3979-3992
26. Poltev V.I., Grokhlina T.I., Malenkov G.G. // *J. Biomol. Struct. Dyn.* 1984. V.2. P.413-429
27. Шарафутдинов М.Р., Данилов В.И., Полтев В.И. // *Доклады Академии Нук Украинской ССР.* 1983. N 12. С.69-72
28. Danilov V.I., Tolokh I.S. // *J. Biomol. Struct. Dyn.* 1990. V.7. С.1167-1183
29. Шестопалова А.В., Данилов В.И., Малеев В.Я. // *ДАН СССР.* Т.282. С.1000-1003
30. Danilov V.I., Shestopalova A.V. // *Int. J. Quant. Chem.* 1989. V.35. P.103-112
31. Poltev V.I., Malenkov G.G., Gonzalez E.J., Teplukhin A.V., Rein R., Shibata M., Miller J.H. // *J. Biomol. Struct. Dyn.* 1996. V.13. P.717-725
32. Cieplak P., Bash P., Singh U.C., Kollman P.A. // *J. Am. Chem. Soc.* 1987. V.109. P.6283-6289