

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРОГНОСТИЧЕСКИХ МОДЕЛЕЙ РАСЧЕТА ХИМИЧЕСКИХ СДВИГОВ НЕОБМЕНИВАЮЩИХСЯ ПРОТОНОВ ДЕЗОКСИОЛИГОНУКЛЕОТИДОВ

Завьялова О.С.

Севастопольский государственный университет

ул. Университетская, 33, г. Севастополь, 299053, РФ; e-mail: zavyalova.os@mail.ru

Поступила в редакцию: 25.08.2018.

**Аннотация.** Проведен сравнительный анализ значений химических сдвигов необмениваемых протонов дезоксиолигонуклеотидов, рассчитанных по двум прогностическим моделям. Методы прогнозирования основаны на использовании эталонных значений химических сдвигов и поправочных коэффициентов, учитывающих влияние «ближайших соседей» в нуклеотидной цепи. Анализ проводился для дезоксигептануклеотида 5'-d(GCGAAGC), образующего шпильчатую структуру в растворе. Теоретические модели показали достаточно высокую точность для значений химических сдвигов протонов оснований, находящихся в стебле шпильки и соответствующих канонической двухспиральной форме.

**Ключевые слова:** ЯМР, химический сдвиг, нуклеотид, триплет, шпилька.

Известно, что вторичная структура нуклеиновых кислот и конформационные параметры молекулы ДНК существенно зависят от нуклеотидной последовательности. Детальное исследование пространственной структуры двойной спирали, в значительной степени, стало возможным благодаря возможности синтезировать олигонуклеотиды с заданной последовательностью оснований. Важную роль в определении конформационных параметров также играют различные расчетно-теоретические и прогностические методы. В настоящее время их использование стало составной частью методик построения моделей молекулярных систем, основанных на результатах экспериментальных данных, полученных, в частности, с помощью методов ЯМР.

Химический сдвиг, как одна из характеристик ЯМР, содержит важную информацию о вторичной структуре ДНК. В настоящее время, существуют различные методы прогнозирования значений химических сдвигов, как для одноцепочечных нуклеотидных последовательностей, так и двойных спиральных форм. Эти методы основаны на анализе экспериментальных значений химических сдвигов протонов различных последовательностей оснований и поправочных коэффициентов, полученных путем статистического анализа и полумэмпирических расчетов. Величины поправок определяются вкладом от «ближайших соседей» в нуклеотидной цепи и могут быть рассчитаны по различным методикам.

Целью настоящей работы являлось проведение сравнительного анализа экспериментальных значений химических сдвигов необмениваемых протонов дезоксиолигонуклеотидов, со значениями, полученными по прогностическим моделям, приведенным в работе Altona C. [1] и Lam S.L. [2, 3]. Для сравнительного анализа значений химических сдвигов была рассмотрена гептануклеотидная последовательность состава 5'-d(GCGAAGC). Дезоксигептануклеотид 5'-d(GCGAAGC), содержит палиндромную последовательность в цепи и способен образовывать шпильчатую структуру с короткой GAA- петлей в растворе [4]. Отнесение резонансов протонов дезоксиолигонуклеотида было проведено методом двумерной гомоядерной ЯМР спектроскопии (2M-TOCSY, 2M-NOESY/ROESY) при следующих экспериментальных условиях: 0,1 М фосфатный буфер, рD=7,1 (табл. 1). Шпильку данного гептамера можно условно подразделить на двуспиральный стебель d(GC)<sub>2</sub> и короткую петлю, состоящую из трех нуклеотидов d(GAA). Как оказалось, такая структура шпильки, имеет высокий уровень перекрытия плоскостей ароматических оснований (существенный выигрыш в энергии за счет стэкинг-взаимодействия), что обеспечивает высокую термостабильность (температура плавления 350 К) и стойкость к действию нуклеаз, содержащихся в экстрактах *E. coli*.

**Таблица 1.** Экспериментальные значения химических сдвигов протонов азотистых оснований и сахарных остатков, входящих в триплеты, гептамера 5'-d(GCGAAGC)

нуклеотид	Химический сдвиг (ppm)*							
	H6/H8	H1'	H2'	H2''	H3'	H4'	H5'	H5''
C2	7,01	6,04	1,68	2,31	4,82	4,24	4,07	4,07
G3	8,09	5,38	2,68	2,55	4,92	4,47	4,12	4,12
A4	8,12	5,99	2,33	2,26	4,59	2,12	3,08	3,40
A5	8,05	6,31	2,93	2,92	4,84	4,37	3,82	-
G6	8,02	5,55	2,60	2,57	4,92	4,40	4,17	4,31

В модели 1, предложенной в работе Altona С. [1] протонный химический сдвиг рассчитывается на основе триплетной схемы. Методика предусматривает разбиение нуклеотидных последовательностей на триплеты с последующим анализом химических сдвигов ароматических протонов и протонов дезоксирибозы центрального нуклеотида в триплете. При этом вводится поправка на местоположение триплета (5'- или 3'-конец цепи), а также учитывается влияние «ближайшего соседа» аденина, гуанина, тимина либо цитозина, на соответствующие химические сдвиги.

Химический сдвиг  $\delta_i$   $i$ -го протона, принадлежащего центральному нуклеотиду  $N_x$  в триплетах вида  $N_5N_xN_3$ ,  $X_5N_xN_3$  и  $N_5N_xX_3$ , моделируется как сумма:

$$\begin{aligned}\delta_i &= N_{xi} + N_{5i} + N_{3i} \\ \delta_i &= N_{xi} + X_{5i} + N_{3i}, \\ \delta_i &= N_{xi} + N_{5i} + X_{3i}\end{aligned}$$

где  $N_{xi}$  – эталонное значение;  $N_{5i}$  и  $N_{3i}$  представляют собой поправки на добавочное экранирование/дезэкранирование протонами одного из четырех оснований, соответствующих 5'- и 3'-концу нуклеотидной цепи. Значения поправок  $X_{5i}$  и  $X_{3i}$  указывают на отсутствие фосфатной группы в соответствующем фланкирующем нуклеотиде.

Несколько иная концепция реализуется в модели 2, разработанной Lam S.L. [2]. В этом случае в качестве структурной единицы выступает пентамер вида  $N_2^5N_1^5XN_1^3N_2^3$ . В центре пентамерной последовательности располагается триплет  $N_1^5XN_1^3$ , фланкированный парой  $N_2^5N_2^3$  с 5'- и 3'-конца соответственно. Таким образом в этой модели учитывается влияние на  $X$  не только «ближайших соседей», но и «next nearest neighbor». Прогнозируемый химический сдвиг  $i$ -го протона задается выражением:

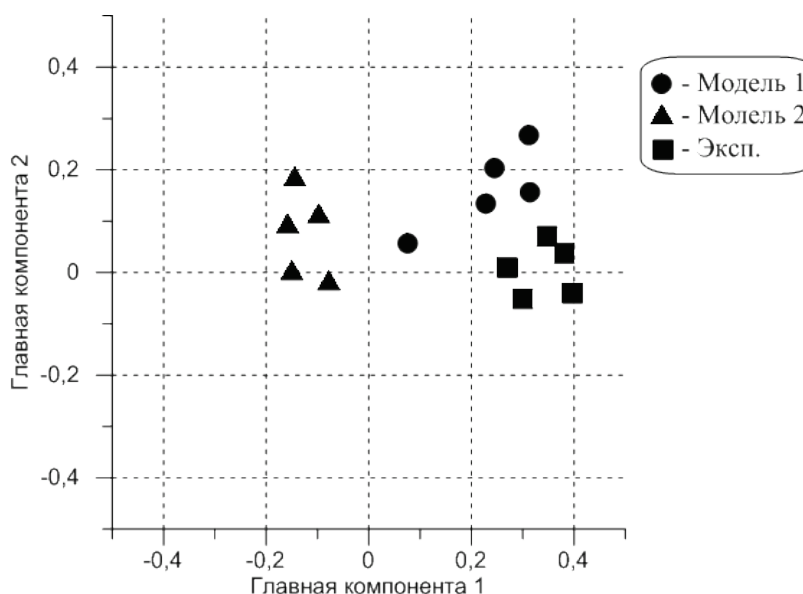
$$\delta_i = \delta_{\text{penta}}(N_1^5XN_1^3) - \Delta_2^{5T} - \Delta_2^{3T} + \Delta_2^{5N} + \Delta_2^{3N},$$

где  $\Delta_2^{5T}$ ,  $\Delta_2^{3T}$  представляют собой поправки на 5'- и 3'-фланкирующие триплет тимина [3], а значения  $\Delta_2^{5N}$ ,  $\Delta_2^{3N}$  учитывают влияние эффекта «next nearest neighbor». Согласно данным, приведенным авторами, эта модель дает точность прогноза порядка 0,02–0,03 м.д. для протонов H6 / H8, H1', H2', H2'' и H3'. Кроме этого необходимо отметить, что модель 2 позволяет рассчитать не только значения резонансных частот ароматических протонов и протонов дезоксирибозы, но и отдельных атомов углерода и фосфатных групп.

В таблице 2 приведены значения химических сдвигов, рассчитанные по двум моделям (номера моделей обозначены цифрами 1 и 2). Число знаков после запятой соответствует значениям, приведенным в [1, 2]. Из сравнения данных таблиц 1 и 2 видно, что для протона H4' аденина, находящегося в петле шпильки, отклонение значения химического сдвига от рассчитанного достигает двух м.д. Это можно объяснить тем, что прогностические модели позволяют рассчитать химические сдвиги, характерные для стандартной двойной спирали ДНК. В нашем случае протон H4', вследствие конформационных особенностей формирования петли шпильки, испытывает экранирующее влияние соседнего ароматического азотистого основания гуанина (G<sub>3</sub>) [4]. С другой стороны, значения химических сдвигов протонов центрального цитозина и гуанина (GCG и AGC), находящихся в стебле шпильки, показывают хорошее совпадение с расчетными. Это позволяет предположить, что стебель находится в конформации, близкой к В-форме двойной спирали.

**Таблица 2.** Значения химических сдвигов, рассчитанные по прогностическим моделям.

Нуклеотид	Химический сдвиг (м.д.)											
	H6/H8		H1'		H2'		H2''		H3'		H4'	
	1	2	1	2	1	2	1	2	1	2	1	2
G1	7.92	-	5.93	-	2.62	-	2.73	-	4.83	-	4.22	-
C2	7.41	7.558	5.72	6.089	1.99	1.963	2.37	2.237	4.79	4.749	4.14	-
G3	7.86	7.735	5.53	5.712	2.66	2.429	2.74	2.439	4.99	4.865	4.31	-
A4	8.08	8.058	5.93	6.025	2.64	2.512	2.88	2.623	5.04	4.945	4.41	-
A5	7.99	8.060	5.87	6.022	2.56	2.502	2.8	2.578	5.00	4.939	4.39	-
G6	7.62	7.843	5.80	5.978	2.47	2.621	2.6	2.644	4.93	4.944	4.35	-
C7	7.4	-	6.17	-	2.16	-	2.21	-	4.48	-	4.04	-



**Рисунок 1.** График счетов для химических сдвигов протонов

Для более наглядной интерпретации, значения химических сдвигов были проанализированы с помощью метода главных компонент (МГК) [5], реализованного в среде Matlab. На графике счетов отчетливо видно, что данные расположились тремя компактными группами, каждая из которых соответствует своему набору данных (рис. 1).

По всей видимости, такое расположение можно объяснить тем, что каждая из прогностических моделей, основываясь на собственной методике, дает хоть и небольшую, но систематическую ошибку по отношению к значениям, полученным экспериментально. Кроме того, на рисунке видно, что значения, полученные по первой модели фактически образовали единую группу с экспериментальными химическими сдвигами. Этот факт говорит о достаточно высокой корреляции между соответствующими величинами и указывает на то, что первая модель дает более «реалистичное» предсказание.

Проанализировав полученные результаты можно сделать вывод о том, что эмпирические модели, разработанные для прогнозирования химических сдвигов, могут быть полезными для исследований конформационных особенностей ДНК и позволяют выявить структуры, обладающие отклонениями от стандартной геометрии двойной спирали, в частности, шпильчатые структуры и разветвления различного вида.

#### **Список литературы / References:**

1. Altona C., Faber D.H., Westra Hoekzema A.J.A. Double-helical DNA  $^1\text{H}$  chemical shifts: an accurate and balanced predictive empirical scheme. *Magn. Reson. Chem.*, 2000, vol. 38, pp. 95-107.
2. Lam S.L. DSHIFT: a web server for predicting DNA chemical shifts. *Nucleic Acids Res.*, 2007, no. 35, pp. 713-717.
3. Lam S.L., Cui X., Ho C.N. Random coil proton chemical shifts of deoxyribonucleic acids. *J. Biomol. NMR*, 2002, no. 24 (4), pp. 329-337.
4. Веселков А.Н., Пахомов В.И., Итон Р., Дэвис Д. ЯМР-исследование самоассоциации молекул дезоксигептануклеотида 5'-d(GCGAAGC) в водном растворе. *Биофизика*, 2000, т. 45, с. 20-26. [Veselkov A.N., Pahomov V.I., Iton R., Davies D. NMR study of the self-association of the deoxyheptanucleotide 5'-d(GCGAAGC) molecules in an aqueous solution. *Biophysics*, 2000, vol. 45, pp. 20-26. (In Russ.)]
5. Эсбенсен К. *Анализ многомерных данных*. Черноголовка: ИПХФ РАН, 2005, 160 с. [Esenssen K.H., *Multivariate Data Analysis*. Chernogolovka: IPCP RAS, 2005, 160 p. (In Russ.)]

**COMPARATIVE ANALYSIS OF PREDICTION MODELS OF CHEMICAL SHIFTS OF  
NON-EXCHANGING PROTONS OF DEOXYOLIGONUCLEOTIDES****Zavyalova O.S.**

Sevastopol State University

*Universitetskaya St., 33, Sevastopol, 299053, Russia; e-mail:zavyalova.os@mail.ru*

**Abstract.** Analysis of the chemical shifts of exchange protons of deoxyoligonucleotides has been made using two predictive models. The prediction methods are based on sets of published reference chemical shift values and correction factors, which account for shielding or deshielding effects from neighboring nucleotides. The analysis has been made for deoxyheptanucleotide 5'-d (GCGAAGC) forming a hairpin in solution. Theoretical models have shown high accuracy for the values of chemical shifts of protons for double helical B-DNA.

**Key words:** *HMR, chemical shift, nucleotide, triplet, hairpin.*