

17. Cockroft S.L., Perkins J., Zonta C., Adams H., Spey S.E., Low C.M.R., Vinter J.G., Lawson K.R., Urch C.J., Hunter C.A. Substituent effects on aromatic stacking interactions. *Org. Biomol. Chem.*, 2007, vol. 5, pp. 1062-1080.
18. Martin R.B. Comparisons of indefinite self-association models. *Chem. Rev.*, 1996, vol. 96, pp. 3043-3064.
19. Mosunov A.A., Rybakova K.A., Rogova O.V., Evstigneev M.P. Binding polynomial in molecular self-assembly. *Phys. Rev. E*, 2014, vol. 89, pp. 062138-4.
20. Tobolsky A.V., Blatz P.J. Thermodynamics of linearly associated systems. *J. Chem. Phys.*, 1945, vol. 13, pp. 379-380.
21. Hernandez Santiago A.A., Buchelnikov A.S., Rubinson M.A., Yesylevskyy S.O., Parkinson J.A., Evstigneev M.P. Shape-independent model (SHIM) approach for studying aggregation by NMR diffusometry. *J. Chem. Phys.*, 2015, vol. 142, pp. 104202-13.
22. Tamura A., Privalov P.L. The entropy cost of protein association. *J. Mol. Biol.*, 1997, vol. 273, pp. 1048-1060.

### СТАТИСТИЧЕСКИЙ АНАЛИЗ ЭКСПЕРИМЕНТАЛЬНЫХ ЗНАЧЕНИЙ ХИМИЧЕСКИХ СДВИГОВ ПРОТОНОВ ДЕЗОКСИОЛИГОНУКЛЕОТИДОВ РАЗЛИЧНОГО СОСТАВА И ПОСЛЕДОВАТЕЛЬНОСТИ ОСНОВАНИЙ В ЦЕПИ

Завьялова О.С., Рогова О.В.

ФГАОУ ВО «Севастопольский государственный университет»

ул. Университетская, 33, г. Севастополь, 299053, РФ

e-mail:zavyalova.os@yandex.ru

**Аннотация.** Методом главных компонент проведен анализ экспериментальных значений химических сдвигов необменивающихся протонов само- и несамокомплементарных дезоксиолигонуклеотидов различного состава и последовательности оснований в цепи. Выявлены признаки, наиболее существенно влияющие на магнитное экранирование протонов оснований рассмотренных коротких фрагментов ДНК.

**Ключевые слова:** дезоксиолигонуклеотид, химический сдвиг, триплет, главные компоненты.

### STATISTICAL ANALYSIS OF EXPERIMENTAL VALUES OF PROTONS CHEMICAL SHIFTS OLIGONUCLEOTIDES OF VARIOUS CONTENT AND BASE SEQUENCE IN THE NUCLEOTIDE CHAIN

Zavyalova O.S., Rogova O.V.

Sevastopol State University

Universitetskaya St., 33, Sevastopol, 299053, Russia

e-mail:zavyalova.os@yandex.ru

**Abstract.** The analysis of experimental values of chemical shifts of non-exchanging protons of deoxyoligonucleotides of various content and base sequence in the nucleotide chain has been made using the method of Principal Components Analysis. Most important factors influencing the shielding of protons of the considered short fragments of DNA are revealed.

**Keywords:** deoxyoligonucleotide, chemical shift, triplet, the principal components.

В настоящей работе проведен статистический анализ экспериментальных значений химических сдвигов необменивающихся протонов само- и несамокомплементарных дезоксиолигонуклеотидов. В качестве статистического метода обработки данных использовался метод главных компонент (МГК). МГК является классическим методом снижения размерности данных путем определения незначительного числа линейных комбинаций исходных признаков, объясняющих большую часть изменчивости данных в целом.

Целью данной работы являлось проведение сравнительного анализа резонансных сигналов (химических сдвигов) протонов дезоксиолигонуклеотидов, отличающихся длиной и последовательностью оснований в цепи. Химический сдвиг, являющийся одним из экспериментальных параметров ЯМР, непосредственно связан с электронным и магнитным окружением молекул и позволяет изучить конформационные особенности соединений непосредственно в растворе. Тетрануклеотиды, рассмотренные в работе, представлены пятью самокомплементарными последовательностями: 5'-d(GCGC), 5'-d(CGCG), 5'-d(ACGT), 5'-d(AGCT), 5'-d(TGCA), гексануклеотидами - 5'-d(CGTACG), 5'-d(CGCGCG), 5'-d(TACGTA) и октамером - 5'-d(GACATGTC) [1,2]. К несамокомплементарным последовательностям, изученным в работе, относятся тетрамеры 5'-d(AAGC), 5'-d(CTGA), 5'-d(CGAA), 5'-d(GAAG) и гептамер - 5'-d(GCGAAGC) [3]. Каждый из дезоксиолигонуклеотидов был разбит на триплеты, отличающиеся составом и локализацией оснований, что позволило проанализировать особенности магнитного экранирования протонов центрального в триплете основания [4].

В зависимости от длины последовательности триплеты одного состава могут быть как терминальными, так и внутренними; терминальные триплеты, в свою очередь, могут принадлежать как к 5'-, так и к 3'- концу цепи; центральное основание может быть окружено двумя пуринами, пиримидинами или какой-либо чередующейся комбинацией оснований. Все эти «признаки» могут быть использованы для выявления особенностей экранирования протонов азотистых оснований методом главных компонент.

В таблице 1, в качестве примера, приведены химические сдвиги протонов центрального аденина и сахарных остатков, входящих в триплеты, содержащие только пуриновые основания. Триплеты взяты как из само-, так и несамокомплементарных дезоксиолигонуклеотидов. Вероятность образования молекулярных дуплексов для

таких олигомеров достаточно мала и в растворе они находятся, преимущественно, в виде одноцепочечных структур.

На рисунке 1 приведен результат первичного анализ химических сдвигов протонов Н1', Н2', Н2'', Н3', Н4', Н5' аденина в триплете состава ААГ. Из рисунка видно, что в случае принадлежности триплета к 5'-концу цепи, диапазон отклонения химических сдвигов протонов от стандартных значений значительно больше, чем для 3'-конца цепи. Химические сдвиги протонов сдвинуты в область более высокого поля, тогда как в триплете ААГ-3' имеют практически те же значения, что и для внутренней последовательности ААГ.

Таблица 1 – Химические сдвиги необменивающихся протонов (м.д.) дезоксиолигонуклеотидов 5'-d(A<sub>1</sub>A<sub>2</sub>G<sub>3</sub>C<sub>4</sub>), 5'-d(G<sub>1</sub>A<sub>2</sub>A<sub>3</sub>G<sub>4</sub>), 5'-d(G<sub>1</sub>C<sub>2</sub>G<sub>3</sub>A<sub>4</sub>A<sub>5</sub>G<sub>6</sub>C<sub>7</sub>)

Протоны	Н8	Н2	Н1'	Н2'	Н2''	Н3'	Н4'	Н5'	Н5''
<b>5'-ААГ</b>	<b>8,12(4)</b>	<b>7,65(4)</b>	<b>5,99(4)</b>	<b>2,70(7)</b>	<b>2,83(3)</b>	<b>4,99(5)</b>	<b>4,37(8)</b>	<b>4,24</b>	-
5'- А <sub>1</sub> А <sub>2</sub> G <sub>3</sub>	7,91	8,01	6,00	2,14	2,42	4,8	4,2	3,7	3,7
<b>ААГ-3'</b>	<b>8,00(4)</b>	<b>7,57(4)</b>	<b>5,99(4)</b>	<b>2,60(7)</b>	<b>2,77(3)</b>	<b>4,99(5)</b>	<b>4,37(8)</b>	<b>4,24</b>	-
А <sub>2</sub> А <sub>3</sub> G <sub>4</sub> -3'	8,08	7,68	5,96	2,57	2,52	4,95	4,34	4,12	4,12
<b>ААГ</b>	<b>8,00(4)</b>	<b>7,47(3)</b>	<b>5,92(4)</b>	<b>2,60(7)</b>	<b>2,77(3)</b>	<b>4,99(5)</b>	<b>4,37(8)</b>	<b>4,24</b>	-
А <sub>4</sub> А <sub>5</sub> G <sub>6</sub>	8,05	8,15	6,31	2,93	2,92	4,84	4,37	3,82	-
<b>ГАА-3'</b>	<b>8,10(2)</b>	<b>7,34(4)</b>	<b>6,04(3)</b>	<b>2,66(3)</b>	<b>2,90(2)</b>	<b>5,05(2)</b>	<b>4,42(2)</b>	<b>4,19(3)</b>	<b>4,14(6)</b>
G <sub>2</sub> A <sub>3</sub> A <sub>4</sub> -3'	8,12	7,83	6,02	2,62	2,51	4,8	4,22	4,02	4,02
<b>5'-ГАА</b>	<b>8,22(4)</b>	<b>7,42(4)</b>	<b>6,04(3)</b>	<b>2,76(3)</b>	<b>2,96(3)</b>	<b>5,05(2)</b>	<b>4,42(2)</b>	<b>4,19(3)</b>	<b>4,14(6)</b>
5'- G <sub>1</sub> A <sub>2</sub> A <sub>3</sub>	8,12	7,78	5,95	2,45	2,33	4,94	4,66	3,65	4,11

Примечание. Жирным шрифтом представлены данные из работы [4], где в скобках указаны величины погрешностей.

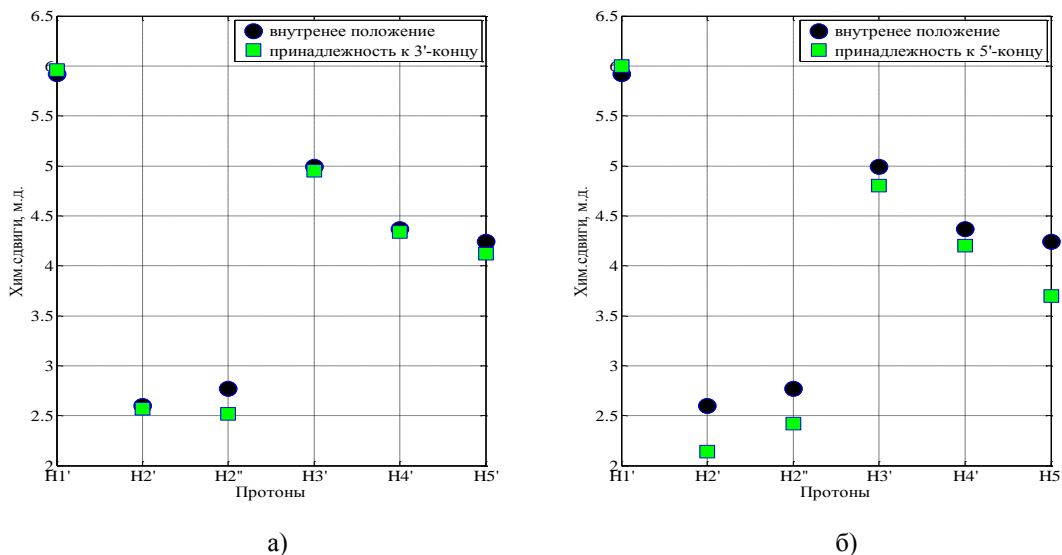


Рисунок 1 – Зависимость химических сдвигов протонов дезоксирибозы центрального аденина от положения триплета ААГ в цепи: а) внутреннее (А<sub>4</sub>А<sub>5</sub>G<sub>6</sub>), терминальное (А<sub>2</sub>А<sub>3</sub>G<sub>4</sub>-3'); б) внутреннее (А<sub>4</sub>А<sub>5</sub>G<sub>6</sub>), терминальное (5'- А<sub>1</sub>А<sub>2</sub>G<sub>3</sub>)

Особое положение среди рассматриваемых олигонуклеотидов занимает несамодополнительный дезоксигептануклеотид 5'-d(GCGAAGC), содержащий палиндромную последовательность в цепи и способный образовывать шпильчатую структуру в растворе [3]. Гептамер d(GCGAAGC) образует шпильку с короткой ГАА-петлей. На рисунке 2 изображена диаграмма размаха, отражающая отклонения величин химических сдвигов протонов от усредненных значений по всем триплетным последовательностям. Из рисунка 2 видно, что для протона Н4' аденина, находящегося в петле шпильки, отклонение значения химического сдвига от наблюдаемого для аналогичных протонов в стандартной двойной спирали ДНК, достигает двух м.д. По всей видимости, это связано с тем, что протон Н4', вследствие конформационных особенностей формирования петли шпильчатой структуры, испытывает экранирующее влияние соседнего ароматического азотистого основания гуанина (G) [3]. Подобные особенности экранирования характерны и для протонов Н5', Н5'' аденина, находящегося в петле шпильки, что отображается на диаграмме в виде «выбросов» значений, смещенных от медианы в область сильного поля.

Для каждого из протонов Н1', Н2'', Н3' распределение величин химических сдвигов относительно значения медианы практически равновероятно. Это свидетельствует о том, что совокупность значений химических

сдвигов протонов, составляющих выборку, подчиняется нормальному распределению, а значит соответствует конформации близкой к В-форме двойной спирали.

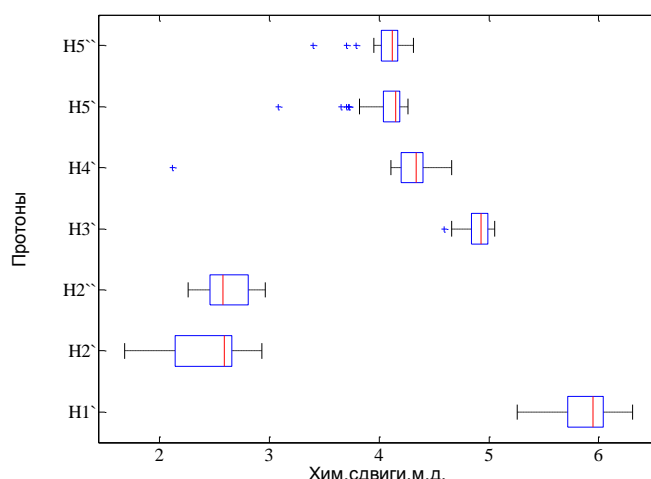


Рисунок 2 – Диаграмма размаха для протонов центрального аденина последовательности ААG

Использование МГК, позволяет представить полезную информацию в более компактном виде, удобном для визуализации и интерпретации [5]. Для проведения численного эксперимента методом главных компонент был использован массив данных, включающий химические сдвиги протонов азотистых оснований всех триплетных последовательностей рассмотренных в работе. Проанализируем графики счетов, которые показывают, как расположились «образцы» в новом подпространстве. На графике счетов PC1-PC2, показанном на рисунке 3, видны отдельные группы с тесной внутренней корреляцией. Относительно оси первой компоненты PC1 группы расположены практически вертикально, что говорит о наличии преобладающего свойства, характеризующего ось этой компоненты. Если проанализировать график счетов совместно с исходным массивом данных, то можно выяснить, что таким свойством является длина последовательности, которой принадлежат триплеты. Из рисунка 3 видно, что в левую вертикальную группу вошли триплеты, выделенные из тетра-нуклеотидов, в среднюю – из гептануклеотида, в правую – из октамера. Так, например, триплет состава ААG входит в «тетра-область» под номерами 24, 26 (выделен из 5'-d(A<sub>1</sub>A<sub>2</sub>G<sub>3</sub>C<sub>4</sub>) и 5'-d(G<sub>1</sub>A<sub>2</sub>A<sub>3</sub>G<sub>4</sub>) соответственно), а в «гепта-область» – под номером 16 (выделен из 5'-d(G<sub>1</sub>C<sub>2</sub>G<sub>3</sub>A<sub>4</sub>A<sub>5</sub>G<sub>6</sub>C<sub>7</sub>). Такая же картина наблюдается и при анализе положения триплетов состава CGA, CGC и GCG. Это свидетельствует о существенном влиянии длины олигонуклеотидной последовательности на положение резонансных пиков протонов. По всей видимости такой факт объясняется тем, что в более длинных последовательностях, эффект плавления концевых нуклеотидных пар оказывает существенно меньшее влияние на экранирование, что связано с большей термической стабильностью таких структур.

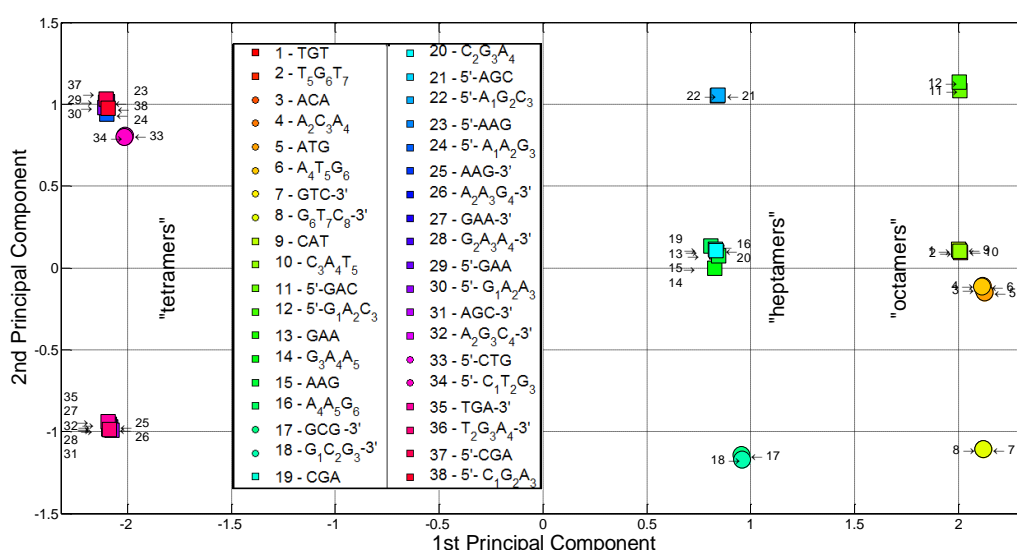


Рисунок 3 – График счетов PC1-PC2

Что касается второй компоненты, она отвечает за местоположение триплета в цепи, а именно – примыкание к 5', 3'- концу, либо внутреннее положение. На «карте образцов» видно, что в нижней части располагаются

триплеты, примыкающие к 3'-концу цепи (7,8,17,18,25,26,27,28,), выше по оси второй компоненты PC2 — внутренние триплеты (9,10,13,14,15,16,19,20), а затем - триплеты принадлежащие 5'-концу (23,24,33,34,37,38).

На рисунке 4 приведена «карта образцов» в подпространстве компонент PC1-PC3. Во-первых, она подтверждает нагрузку первой главной компоненты PC1 по принадлежности к последовательностям различной длины. Во-вторых, из графика видно, что направление оси третьей компоненты PC3 отвечает за принадлежность центрального основания к пуринам или пиримидинам. Верхняя полуплоскость соответствует центральному пурину (аденину или гуанину) в триплете, а нижняя – центральному пурину (цитозину, тимину). Последующее выделение «более старших» компонент PC4, PC5 и т.д. в данном массиве данных, не позволило соотнести с ними какие-либо существенные признаки.

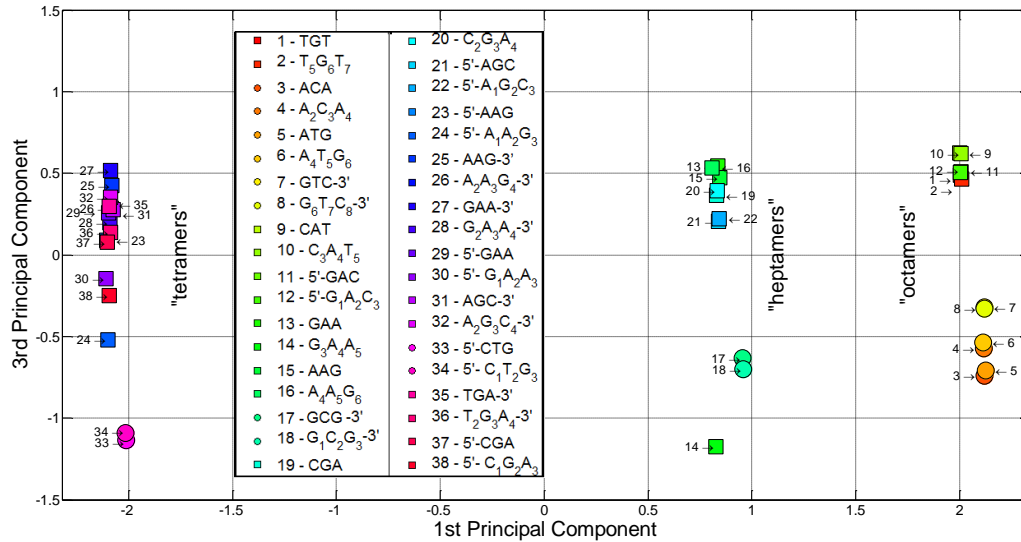


Рисунок 4 – График счетов PC1-PC3

Матрица нагрузок главных компонент на исходные признаки, соответствующая данному численному эксперименту, приведена в таблице 2. Из таблицы видно, что на первые три компоненты приходится порядка 92% суммарной вариации данных.

Таблица 2 – Матрица нагрузок главных компонент на исходные признаки

Главная компонента PC <sub>j</sub>	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
Вклад j-главной компоненты, %	71,841	13,903	6,100	3,941	1,887	1,289	0,493	0,254	0,1799	0,079	0,032
Суммарный вклад, %	71,841	85,744	91,844	95,785	97,672	98,961	99,454	99,708	99,888	99,967	100

Это означает, что из всех признаков, которыми обладает использованная экспериментальная база данных, наиболее существенными, влияющими на значения экранирования протонов, оказываются следующие: длина нуклеотидной последовательности (PC1), местоположение триплета в нуклеотидной цепи (PC2), его принадлежность к пуринам или пиримидинам (PC3). При этом наибольшую нагрузку на главные компоненты несет такой признак, как длина олигонуклеотидной последовательности. Это, по всей видимости, связано с эффектом плавления концевых нуклеотидных пар, который в случае коротких последовательностей дает более существенный вклад в экранирование по сравнению с последовательностями большей длины.

Следует отметить, что для более детального выявления закономерностей экранирования протонов ароматических оснований и дезоксирибозы, необходимо рассмотрение более широкого спектра триплетных последовательностей, принадлежащих олигонуклеотидам различного состава и последовательности оснований в цепи. Представленная в настоящей работе методика статистического анализа протонных химических сдвигов методом главных компонент может быть использована в дальнейшем для решения различных задач структурного анализа в молекулярной биофизике, в частности, для выявления структур, обладающих отклонениями от стандартной геометрии двойной спирали ДНК.

**Список литературы / References:**

1. Веселков А.Н., Дэвис Д., Дымант Л.Н., Паркес Х. 1М- и 2М- <sup>1</sup>Н-ЯМР исследование самоассоциации дезокситетрануклеозидтрифосфатов различной последовательности оснований в водном растворе. *Биополимеры и клетка*, 1991, т. 7, № 5, с. 15-22. [Veselkov A.N., Davies D., Djimant L. N., Parkes H. 1D- and 2D- <sup>1</sup>H NMR investigation of self-association of deoxytetranucleotides of different base sequence in aqueous solution. *Biopolymers and Cell*, 1991, vol. 7, no. 5, pp. 15-22. (In Russ.)]
2. Веселков А.Н., Пахомов В.И., Итон Р., Дэвис Д. ЯМР-исследование самоассоциации молекул дезоксигексануклеотидов 5'-d(CpGpTpApCpG) и 5'-d(CpGpCpGpCpG) в водном растворе. *Биофизика*, 2000, т. 45, с. 20-26. [Veselkov A.N., Pahomov V.I., Eaton R., Davies D. NMR investigation of self-association of hexadeoxynucleotides 5'-d(CpGpTpApCpG) and 5'-d(CpGpCpGpCpG) in aqueous solution. *Biophysics*, 2000, vol. 45, pp. 20-26. (In Russ.)]
3. Davies B.D., Pahomov V.I., Veselkov A.N. NMR determination of the conformation and drug binding properties of the DNA heptamer d(GpCpGpApApGpC) in aqueous solution. *Nucleic Acids Res.*, 1997, vol. 25, pp. 4523-4531.
4. Altona C, Faber D.H., Westra Hoekzema A.J.A. Double-helical DNA 1H chemical shifts: an accurate and balanced predictive empirical scheme. *Magn. Reson. Chem.*, 2000, vol.38, pp.95-107.
5. Эсбенсен К. *Анализ многомерных данных*, сокр. пер. с англ. под ред. О. Родионовой, Изд-во ИПХФ РАН, 2005, 204с. [Esbensen K.H.. *Multivariate Data Analysis*, abbr. per. from English. ed. O. Rodionova, ICP RAS, 2005, 204 p. (In Russ.)]

**КОНФОРМАЦИОННЫЙ АНАЛИЗ НЕЙРОПЕПТИДА [HYP2]МЕТ-ГАЛЛАТОСТАТИНА**

Велиева Л.И.

Бакинский государственный университет  
ул. З. Халилова, 23, Баку, AZ1148, Азербайджан  
e-mail: Lala\_Veliyeva@rambler.ru

**Аннотация.** В работе методом молекулярной механики в приближении атом-атомных потенциальных функций исследовано пространственное строение и конформационные свойства нейропептида из семейства аллатостатинов, получившего название [Hyp2]Met-галлатостатина. Исследования основывались на минимизации энергии внутримолекулярных взаимодействий (невалентных, электростатических, торсионных) в окрестностях углов внутреннего вращения, соответствующих всем комбинациям стабильных конформаций монопептидных остатков, образующих молекулу. Установлены десять энергетически стабильных конформаций нейропептида, реализуемых в условиях неявно заданного водного окружения, относительная конформационная энергия которых варьирует в интервале значений 0-10 ккал/моль. Показано, система водородных связей, несмотря на небольшой вклад в общую энергию молекулы, весьма важна для сохранения устойчивой структуры и ограничения подвижности [Hyp2]Met-галлатостатина.

Согласно результатам исследования, основной вклад в стабилизацию низкоэнергетических конформаций молекулы вносят невалентные взаимодействия.

**Ключевые слова:** нейропептиды; структура; конформационный анализ

**CONFORMATIONAL ANALYSIS NEUROPEPTIDE [HYP2]MET-CALLATOSTATIN**

Veliyeva L.I.

Baku State University  
Z. Khalilov str.23, Baku, AZ 1148, Azerbaijan  
e-mail: Lala\_Veliyeva@rambler.ru

**Abstract.** In work by the method of molecular mechanics in approximation of atom-atom potential functions is studied the spatial structure and conformational properties of the neuropeptide from allatostatin family, got the name of [Hyp2]Met-callatostatin. The researches were based on minimizing of the energy of intramolecular interactions ( non-bonded, electrostatic, torsion ) in the vicinity of the internal rotation angles corresponding to all combinations of stable conformations of mono-peptide residues forming a molecule. There determined ten energetically stable conformations of neuropeptide, implemented under the conditions of implicitly given aqueous environment , the relative conformational energy of which varies in the range of 0-10 kcal / mol. It was shown that the system of hydrogen bonds, despite a small contribution to the total energy of the molecule is essential for the preservation of a stable structure and restriction of [Hyp2]Met-callatostatin .

According to the research results, the main contribution to the stabilization of low-energy conformations of the molecule are made by non - valent interactions.

**Keywords:** neuropeptides, structures, conformational analysis

**Введение.**

Одной из актуальных проблем в современной науке является поиск и целенаправленный синтез соединений, используемых для регуляции численности вредителей сельскохозяйственных культур. К числу таких соединений относятся нейропептиды, синтезируемые нейросекреторными клетками мозга различных видов насекомых, в частности, *Calliphora Vomitoria* [1-3]. Нейропептиды ингибируют синтез и выделение ювенильных гормонов в