

КВАНТОВО-МЕХАНІЧНИЙ КОНФОРМАЦІЙНИЙ АНАЛІЗ МОЛЕКУЛИ 5'-ТИМІДИЛОВОЇ КИСЛОТИ

Т. Ю. НИКОЛАЄНКО¹, Л. А. БУЛАВІН¹, Д. М. ГОВОРУН^{1,2}

¹Київський національний університет імені Тараса Шевченка, Україна;

²Інститут молекулярної біології і генетики НАН України, Київ;

e-mail: tim_mail@ukr.net

Методами обчислювальної квантової механіки на рівні теорії MP2/6-311++G(d,p) // B3LYP/6-31G(d,p) проведено повний конформаційний аналіз структурної ланки ДНК – нуклеотиду з тиміновою основою та електронейтральною фосфатною групою у 5'-положенні. Виявлено 660 його конформацій із відносними енергіями Гіббса 0–11,1 ккал/моль за нормальних умов. Серед них шість конформерів подібні до структури нуклеотиду ДНК у AI-, один – у AII- і сім – в VI-формах відповідно. Найнижчу енергію Гіббса серед ДНК-подібних конформерів ($\Delta G = 2,7$ ккал/моль) має VI-ДНК-подібна структура. Показано, що глікозидний зв'язок є найлабільнішим з усіх хімічних зв'язків молекули. Продемонстровано роль внутрішньомолекулярних водневих зв'язків типу $CN\cdots O$ у формуванні просторової структури конформерів 5'-тимідилової кислоти.

Ключові слова: нуклеотид, 2'-дезокситимідилова кислота, конформаційний аналіз, ДНК, ab initio розрахунки, водневі зв'язки.

У процесі функціонування ДНК, зокрема під час взаємодії з протеїнами у разі зчитування генетичної інформації, суттєвих конформаційних змін зазнають її структурні ланки [1]. Тому проведення повного конформаційного аналізу останніх є необхідним кроком для інтерпретації зміни структури ділянок ДНК як переходів між їхніми метастабільними конформаційними станами.

До мономерних ланок ДНК належить молекула 5'-тимідилової кислоти – нуклеотид із тиміновою основою та електронейтральною фосфатною групою у 5'-положенні (рис. 1). Дослідження конформаційних можливостей цієї молекули дозволить суттєво розширити дані [2] про конформаційні можливості піримідинових 2'-дезоксирибонуклеотидів.

Порівняння конформаційних властивостей нуклеотиду і нуклеозиду [3] з тією самою нуклеотидною основою (тиміном) дозволить визначити внутрішньомолекулярні чинники, що є важливим у визначенні основних структурних властивостей складових частин ДНК.

Матеріали і методи

Методи обчислювальної квантової механіки дозволяють провести аналіз усіх можливих конформерів молекули 5'-тимідилової кислоти незалежно від їхньої відносної енергії. Для проведення такого аналізу [2, 3] створюють якомога репрезентативніше сімейство стартових геометрій досліджуваної молекули і

проводять оптимізацію їхньої геометрії квантово-механічним методом функціоналу густини (DFT), після чого відсіюють однакові та нестійкі конформери і методом MP2 виконують розрахунок електронної енергії тих конформерів, що залишилися («унікальних»).

Для опису конформації молекули 5'-тимідилової кислоти (рис. 1) ми використовували 9 конформаційних параметрів: загальноприйняті торсійні кути ε , γ , β і χ , кут P і амплітуду ν_{\max} псевдообертання фуранозного кільця [4] та торсійні кути $\alpha' = C_5O_5PO_p$, $\zeta_1 = O_5PO_{p1}H_{p1}$ і $\zeta_2 = O_5PO_{p2}H_{p2}$. При цьому для трьох атомів кисню фосфатної групи використано такі позначення: не з'єднаний з атомом водню атом – O_p , а два інших – O_{p1} і O_{p2} , якщо дивитися вздовж зв'язку O_5P і рухатися від O_p за годинниковою стрілкою. Зазначимо, що кут α' відрізняється від номенклатурного кута α [4] на постійну величину. Кути ζ_1 і ζ_2 описують орієнтації гідроксильних груп $O_{p1}H_{p1}$ і $O_{p2}H_{p2}$ відповідно.

Першим кроком у створенні репрезентативного сімейства стартових геометрій молекули 5'-тимідилової кислоти було приєднання до кожного з 99 конформерів відповідного нуклеозиду – молекули 2'-дезокситимідину [3] – замість атома H_5 фосфатної групи $PO_pO_{p1}H_{p1}O_{p2}H_{p2}$ у кожному з трьох положень: $\alpha' = 173^\circ$, 113° та -127° (кути ζ_1 та ζ_2 мали при цьому фіксовані значення $-57,3^\circ$ та $115,2^\circ$ відповідно). Геометрії одержаних у такий спосіб 277 конформерів нуклеотиду було оптимізо-

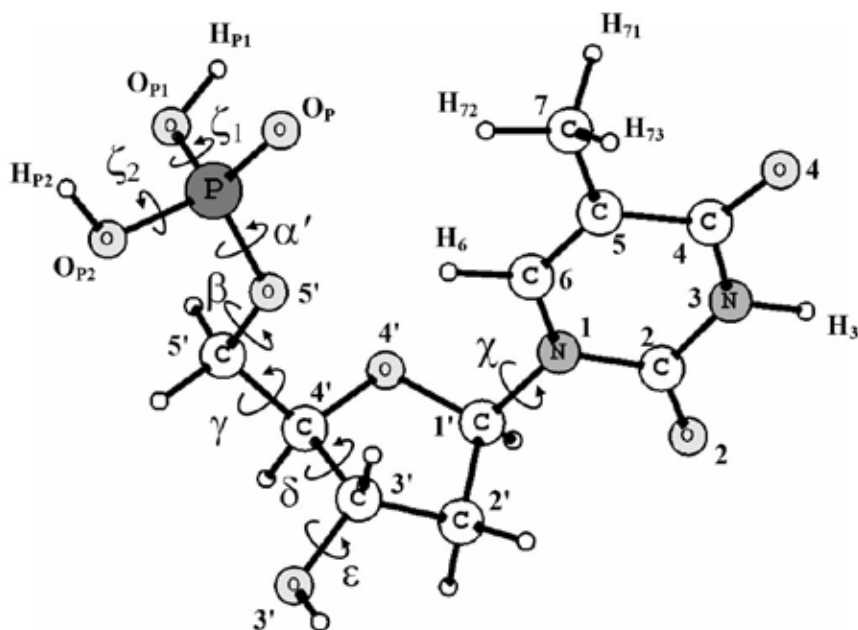


Рис. 1. Структура та позначення атомів молекули 5'-тимідилової кислоти

вано на рівні теорії DFT B3LYP/6-31G(d,p) з використанням програмного пакету Gaussian 03 [5] із налаштуванням точності обчислень за замовчуванням. Після цього для кожного конформера розраховували у гармонійному наближенні частоти нормальних коливань і використовували одержані дані для вилучення нестійких структур (в коливальних спектрах яких присутня хоча б одна уявна частота).

Враховуючи невідворотні обчислювальні похибки, що виникають у процесі оптимізації геометрії, вважали однаковими будь-які два конформери, в яких значення кута ν_{\max} відрізнялися не більше ніж на 2° та інші кути (α' , β , γ , ε , ζ_1 , ζ_2 , P і χ) не більше ніж на 5° (відмінність між електронними енергіями таких однакових конформерів виявилася не більшою 0,01 ккал/моль, що відповідає величині обчислювальної похибки). Після вилучення однакових структур залишилося 197 «унікальних» конформерів.

Далі одержане сімейство конформерів розширювали за допомогою спеціальної процедури доповнення, яку виконували декілька разів. Ця процедура полягає в тому, що для кожного із наявних оптимізованих конформерів шляхом зміни в ньому одного із конформаційних параметрів (торсійного кута тощо) створюють певну кількість нових стартових конформацій, виконують оптимізацію їхніх геометрій, відсіюють нестійкі структури і долучають ті, що залишилися до наявного сімейства конфор-

мерів, після чого з нього видаляють однакові структури.

Першу процедуру доповнення було виконано для кута χ . Виявлено, що у одержаних на першому етапі 197-ми конформерів розподіл значення кута χ є бімодальним, тобто усі його значення потрапляють у досить вузьку *syn*- ($\chi \in 75^\circ \pm 20^\circ$) та значно ширшу *anti*- ($\chi \in -144^\circ \pm 40^\circ$) області. З огляду на це, доповнення наявного сімейства конформерів новими стартовими геометріями було виконано наступним чином. Увесь діапазон значень кута χ розбито на 3 сектори: $s = \{X_s \pm 20^\circ\}$, $a_1 = \{X_{a1} \pm 15^\circ\}$ і $a_2 = \{X_{a2} \pm 25^\circ\}$, з центрами $X_s = 75^\circ$, $X_{a1} = -120^\circ$ і $X_{a2} = -160^\circ$ відповідно. Для кожного із оптимізованих конформерів, у яких кут χ попадав у сектор s , створювали 2 нові стартові геометрії, в яких за незмінних інших конформаційних параметрів кут χ встановлювався рівним X_{a1} та X_{a2} . Аналогічно для кожного із конформерів, в яких кут χ попадав у сектор a_1 , створювали дві нові стартові геометрії з $\chi = X_s$ і $\chi = X_{a2}$, а для конформерів, в яких $\chi \in a_2$ — дві нові стартові геометрії з $\chi = X_s$ і $\chi = X_{a1}$.

Після оптимізації геометрії створених у такий спосіб 394-х нових структур, відсіювання нестійких та однакових конформерів загальна кількість унікальних стійких конформерів склала 278.

Далі таку саму процедуру доповнення сімейства було виконано щодо торсійного кута γ .

Зважаючи на чітко виражену секторну структуру у розподілі його значень, область можливих значень кута γ розбили на 3 сектори: g^+ , g^- і t з центрами 60° , -60° і 180° відповідно та завширшки 120° кожен.

Після оптимізації 556-ти новостворених геометрій та відсіювання нестійких і однакових конформерів загальна кількість унікальних конформерів склала 394.

Далі це сімейство було доповнене щодо кутів ζ_1 і ζ_2 , необхідність чого пов'язана з тим, що при побудові вихідних стартових геометрій молекули 5'-тимідилової кислоти із 2'-дезокситимідину та фосфатної групи орієнтації гідроксильних груп $O_{p1}H_{p1}$ і $O_{p2}H_{p2}$ (кути ζ_1 і ζ_2) були фіксованими. Оскільки у 95% конформерів із наявного сімейства значення кута ζ_1 виявилися майже неперервно розподіленими в діапазоні $-90^\circ \pm 90^\circ$, а значення кута ζ_2 у 90% конформерів – в діапазоні $90^\circ \pm 90^\circ$, то замість «секторного» алгоритму, застосованого вище для кутів χ та γ , задля доповнення конформаційного сімейства щодо ζ_1 і ζ_2 для кожного із наявних оптимізованих конформерів створювали єдину нову стартову структуру шляхом заміни значень кутів $\zeta_1 \leftrightarrow \zeta_2$ у ньому. Оптимізація новостворених структур збільшила загальну кількість стійких унікальних конформерів до 554, а повторення цієї процедури доповнення щодо ζ_1 і ζ_2 ще раз збільшило її неістотно – до 572.

Останнім кроком до розширення цього сімейства було виконання процедури доповнення щодо кута ε з використанням того ж секторного алгоритму і області значень (g^+ , g^- і t сектори), що і для кута γ . Внаслідок виконання цієї процедури загальна кількість стійких унікальних конформерів молекули 5'-тимідилової кислоти склала 727.

Їхні геометрії було додатково оптимізовано на рівні теорії DFT B3LYP/6-31G(d,p) з використанням сітки інтегрування підвищеної точності (*int=ultrafine*)¹ та жорстких умов збіжності (*opt=tight*)², що є необхідним для надійного опису нормальних коливань з найбільш низькими частотами [6] і підвищення точності у визначенні енергії Гіббса, що припадає на молекулу у газовій фазі. Внаслідок цього кількість унікальних стійких конформерів склала 660. У рамках цієї роботи одержане сімейство конформерів можна вважати повним.

Електронні енергії всіх структур цього сімейства було обчислено методом MP2, який

адекватно передає енергії внутрішньомолекулярних нековалентних (зокрема вандерваальсових) взаємодій, з використанням базисного набору 6-311++G(d,p).

Для виявлення внутрішньомолекулярних водневих зв'язків використовували метод QTAІМ аналізу топології просторового розподілу густини електронного заряду [7].

Результати та обговорення

Енергії Гіббса всіх 660 унікальних структур молекули 5'-тимідилової кислоти за нормальних умов (температура $T = 298,15$ К (25°C) і тиск $p = 1$ атм) лежать у межах $0 \div 11,1$ ккал/моль, причому найнижчою є енергія Гіббса конформера **1** (табл. 1, рис. 2), який має одночасно і найнижчу з усіх конформерів електронну енергію. Енергія Гіббса конформера **2**, структура якого відрізняється від структури конформера **1** лише орієнтацією гідроксильної групи $O_{p2}H_{p2}$, близька до мінімальної (табл. 1), тож можна говорити, що глобальний мінімум енергії Гіббса є квазивиродженим. Цікаво зазначити, що у 2'-дезокситимідині [3] теж має місце близькість енергії Гіббса двох енергетично найвигідніших конформерів, проте їхні структури відрізняються значенням кута ε .

Відзначимо, що структури енергетично найвигідніших конформерів **1** і **2** стабілізовані внутрішньомолекулярними водневими зв'язками $O_{p1}H_{p1} \cdots O_2$, $O_3H_3 \cdots O_p$, $C_3H \cdots O_p$ і $C_3H \cdots O_2$, енергії перших двох з яких, оцінені за зсувом частот валентних *ОН* коливань [8], становлять в обох конформерах 5,6 та 3,1 ккал/моль відповідно.

Наступний за енергією Гіббса конформер **3** відділений від конформерів **1** і **2** енергетичною «щилиною» у $\Delta G = 0,4$ ккал/моль (і $\Delta E = 1,4$ ккал/моль за електронною енергією).

Сумарна імовірність $w = \frac{1}{Z} \cdot \sum_{i=1}^{30} \exp\left(-\frac{\Delta G_i}{k_B T}\right)$ (k_B – стала Больцмана, ΔG_i – енергія Гіббса i -го конформера при $T = 298,15$ К (25°C), нормувальний множник $Z = \sum_{j=1}^{660} \exp\left(-\frac{\Delta G_j}{k_B T}\right)$) виявлення усіх конформерів **1–30** (табл. 1) одночасно у газовій фазі (так звана «заселеність») складає 90%, із яких 42% припадає на енерге-

¹Така сітка складається із 99 радіальних оболонок навколо кожного ядра та 590 точок на кожній із них;

²Для контролю збіжності використовується порогове значення, рівне 10^{-5} ат.од. для середньоквадратичної величини похідних електронної енергії за координатами ядер

Таблиця 1. Структурні та енергетичні параметри вибраних³ конформерів молекули 2'-дезокситимідилової кислоти (квантово-механічний розрахунок на рівні теорії MP2/6-311++G(d,p) // B3LYP/6-31G(d,p))

№ к-ра	Конформаційні параметри, град.									Енергія ⁴ , ккал/моль	
	α'	β	γ	ε	ζ_1	ζ_2	χ	P	v_{max}	ΔE	ΔG
1	9,9	-90,6	55,0	-91,0	-44,4	-70,5	67,0	32,5	34,7	0,00	0,00
2	9,1	-90,6	54,5	-91,0	-41,4	121,7	66,9	32,2	34,5	0,10	0,02
3	-118,2	77,5	51,5	-174,5	-63,9	133,5	-110,8	127,6	42,9	1,41	0,42
4	-100,4	63,5	49,8	-170,0	-72,1	-50,8	-111,6	114,7	41,6	2,12	0,65
5	-30,4	-88,1	175,0	177,4	-54,4	-79,1	-131,5	155,1	39,2	2,17	1,12
6	-32,0	-87,9	175,5	177,3	-52,5	106,3	-132,2	155,3	39,0	2,18	1,20
7	-120,6	79,3	51,0	-74,2	-62,5	132,1	-108,9	129,6	42,6	2,62	1,39
8	48,0	-104,4	47,2	-89,9	-103,0	162,2	-159,2	15,7	39,7	3,50	1,41
9	-106,2	66,6	44,4	-99,9	-69,3	135,7	-116,4	64,1	40,3	3,21	1,58
10	-31,8	-86,6	177,2	-163,7	-53,7	107,3	79,5	315,0	20,7	2,47	1,58
11	-32,7	-88,5	175,7	-64,9	-52,3	106,7	-128,3	158,0	38,9	2,72	1,65
12	-31,0	-88,7	174,9	-65,7	-54,0	-81,2	-127,6	157,5	39,1	2,81	1,66
13	46,9	141,9	42,9	-175,8	-110,8	-36,7	65,3	147,5	39,6	4,20	1,66
14	-16,8	132,6	48,7	-176,2	-134,4	56,7	70,4	146,1	38,6	3,65	1,70
15	51,5	-103,2	45,4	-93,0	52,2	153,4	-154,8	20,8	39,5	4,27	1,71
16	-6,9	110,6	44,5	-152,9	-123,0	46,6	65,1	50,3	31,9	3,48	1,71
17	-30,3	-86,1	177,2	-163,6	-56,4	-71,5	79,2	314,9	20,7	2,75	1,74
18	-32,2	-86,7	177,8	-84,9	-54,1	107,9	79,1	312,2	20,2	2,88	1,81
19	-5,4	110,7	41,5	-94,8	-123,7	49,1	64,6	48,2	33,1	3,68	1,84
20	-115,6	76,1	51,3	65,3	-65,9	133,8	-111,4	123,5	41,8	3,11	1,91
21	48,6	146,7	40,3	-68,9	-112,4	-37,2	64,7	153,1	39,1	4,40	1,93
22	-30,8	-86,3	177,6	-86,3	-56,4	-73,6	78,8	312,7	20,3	3,22	2,03
23	-23,2	129,7	50,1	173,0	-169,4	91,1	-169,3	169,7	33,4	3,86	2,06
24	-18,1	135,6	48,0	-68,5	-135,0	55,4	70,8	150,4	38,2	4,05	2,06
25	-14,2	129,9	49,2	-176,2	75,2	56,6	70,4	144,4	38,9	3,97	2,13
26	45,2	-103,6	47,8	-89,6	-141,3	-54,9	-155,5	18,4	39,5	4,55	2,15
27	84,8	106,6	50,2	-149,8	48,8	-21,3	68,1	46,0	31,3	2,48	2,24
28	62,8	-121,1	56,8	-179,7	-124,7	155,7	-128,8	157,1	33,4	6,61	2,26
29	17,9	-96,7	49,2	-84,4	164,3	112,2	-155,0	20,6	39,4	4,23	2,27
30	85,7	106,6	48,7	-101,3	50,3	-22,2	67,6	45,0	32,3	2,72	2,29
41	68,8	-173,0	55,4	176,1	-123,4	153,8	-126,7	158,3	34,5	6,42	2,68
48	45,4	170,3	48,9	174,4	-104,8	160,5	-136,4	159,4	35,5	6,30	2,80
55	69,5	166,0	57,6	173,2	51,0	149,5	-169,2	184,3	31,2	5,00	3,11
60	44,5	167,8	51,1	-169,9	-101,7	162,2	-165,1	14,1	34,3	7,37	3,21
61	-162,5	169,0	55,7	174,6	-167,4	104,0	-122,0	158,1	34,2	6,74	3,22
73	42,0	163,5	48,9	173,0	-126,7	-54,1	-175,4	188,3	31,5	6,85	3,43
78	-171,8	169,4	56,3	-167,6	-102,6	175,0	-151,8	21,0	34,4	7,01	3,52

95	61,7	-169,4	57,6	-166,8	-112,8	157,8	-159,9	18,4	34,7	7,52	3,84
101	-176,2	177,1	55,9	-171,8	-171,4	103,4	-156,5	17,9	35,0	7,37	3,93
103	169,0	164,9	47,3	173,9	-179,6	108,2	-136,8	159,5	35,4	7,20	3,99
129	65,8	-172,1	57,6	177,5	-137,9	-61,8	-123,5	156,9	34,3	7,98	4,27
149	72,6	-57,5	-54,5	-57,4	-111,0	156,7	69,8	353,6	9,9	6,66	4,54
154	59,7	164,6	51,6	-165,6	41,2	151,5	-176,9	356,5	31,6	6,72	4,58
156	58,3	-175,5	58,0	-165,7	-133,4	-60,2	-152,9	22,5	34,6	8,82	4,63
161	68,0	-57,9	-53,4	-57,8	36,6	150,3	69,6	352,5	9,2	6,61	4,71
199	-102,8	68,7	26,2	-167,2	-105,7	144,5	59,2	124,5	43,3	6,26	5,07
303	167,3	-162,0	170,9	51,1	-162,0	84,6	-178,9	216,8	37,6	8,15	6,02
311	69,8	-58,3	-54,1	-55,5	-132,5	-62,3	69,4	355,6	9,4	8,18	6,06
399	-55,2	-164,1	-179,7	-173,1	-158,0	103,9	-174,0	357,8	32,6	10,66	6,96
639	-97,0	60,6	-120,5	-170,8	-83,4	135,9	70,4	41,2	35,6	11,38	9,51
660	-87,3	55,7	-115,6	175,4	-90,2	-39,1	-167,5	23,4	39,0	13,65	11,1

³ Файли структур повного сімейства конформерів можна одержати у авторів за запитом

⁴ ΔE (електронні енергії) та ΔG (енергії Гіббса, які припадають на одну молекулу у газовій фазі) відраховані від відповідних енергій конформера **1**: $\Delta E_i = E_i^{MP2} - E_1^{MP2}$, $\Delta G_i = G_i - G_1$ (i – номер конформера). E_i^{MP2} розраховували методом MP2/6-311++G(d,p) за геометріями, оптимізованими на рівні теорії DFT B3LYP/6-31G(d,p), а G_i обчислювали як $G_i = G_i^{DFT} - E_i^{DFT} + E_i^{MP2}$, де внесок $G_i^{DFT} - E_i^{DFT}$ трансляційних, обертальних і коливальних ступенів вільності молекули розраховували на тому ж рівні теорії, на якому проводили оптимізацію геометрії

тично найвигідніші конформери **1** і **2**. Таким чином, саме ці 30 конформерів визначатимуть, наприклад, ІЧ-спектр молекули 5'-тимідилової кислоти у газовій фазі або в інертній матриці.

На рис. 3 наведено так звані «конформаційні кільця», які унаочнюють розподіл значень конформаційних параметрів в усіх можливих конформерах молекули 5'-тимідилової кислоти. Так, зокрема «північну» ($P \in 0^\circ \pm 90^\circ$) конформація фуранозного кільця мають 330 конформерів (з яких 215 – класичну C3'endo ($P \in 18^\circ \pm 18^\circ$)), «південну» – 330 конформерів (з яких 202 – класичну C2'endo ($P \in 162^\circ \pm 18^\circ$)). За температури 25 °C у газовій фазі заселеності «північної» і «південної» конформацій фуранозного кільця складають 59 і 41% відповідно. При цьому у жодного конформера P не потрапляє в «західну» область ($P \in 225^\circ \div 330^\circ$); у 26 конформерів P потрапляє у «східну» область $P \in 60^\circ \div 120^\circ$. Відомо [4], що дві останні області відповідають бар'єрам у циклі псевдообертання фуранозного кільця.

Амплітуда ν_{\max} псевдообертання фуранозного кільця у 657 конформерів лежить у межах $\nu_{\max} \in 29^\circ \pm 16^\circ$, тоді як у трьох конформерів (**149**, **161** і **311** в табл. 1) вона не перевищує 10°. Найменшого значення (9,2°) ν_{\max} досягає у конформері **161** (рис. 2), а найбільшого (43,3°) – у **199**. Структурним чинником стабілізації нетипової

пласкої конформації фуранозного кільця у конформері **161** є внутрішньомолекулярний водневий зв'язок $O_3H \cdots O_p$ із енергією 6,5 ккал/моль (визначеною методом [9]). Відзначимо, що такий самий водневий зв'язок може реалізуватися і у молекулі 2'-дезокситидилової кислоти [2], в якій він також стабілізує пласку конформацію фуранозного кільця.

Серед усіх можливих конформерів 5'-тимідилової кислоти 276 мають *syn*-орієнтацію нуклеотидної основи ($\chi \in 73^\circ \pm 20^\circ$) і 384 – *anti*-орієнтацію ($\chi \in -144^\circ \pm 50^\circ$). У газовій фазі за температури 25 °C заселеності *syn*- і *anti*- конформерів складають 57,7 та 42,3% відповідно.

Розподіли значень торсійних кутів γ і ε (рис. 3) виявляють чітку секторну структуру: у 95% конформерів значення цих кутів потрапляють у сектори $60^\circ \pm 30^\circ$ (*gauche+*), $-60^\circ \pm 30^\circ$ (*gauche-*) і $180^\circ \pm 30^\circ$ (*trans*). Разом із цим, у конформерів **639** і **660** значення кута γ є близькими до -120° .

У цілому, розподіли усіх торсійних кутів подібні до аналогічних у молекулі 2'-дезокситидилової кислоти [2].

Проведений за методикою [10] статистичний аналіз не виявив істотних кореляцій між значеннями торсійних кутів у сукупності одержаних конформерів молекули 5'-тимідилової кислоти. Так, обчислені згідно з [10]

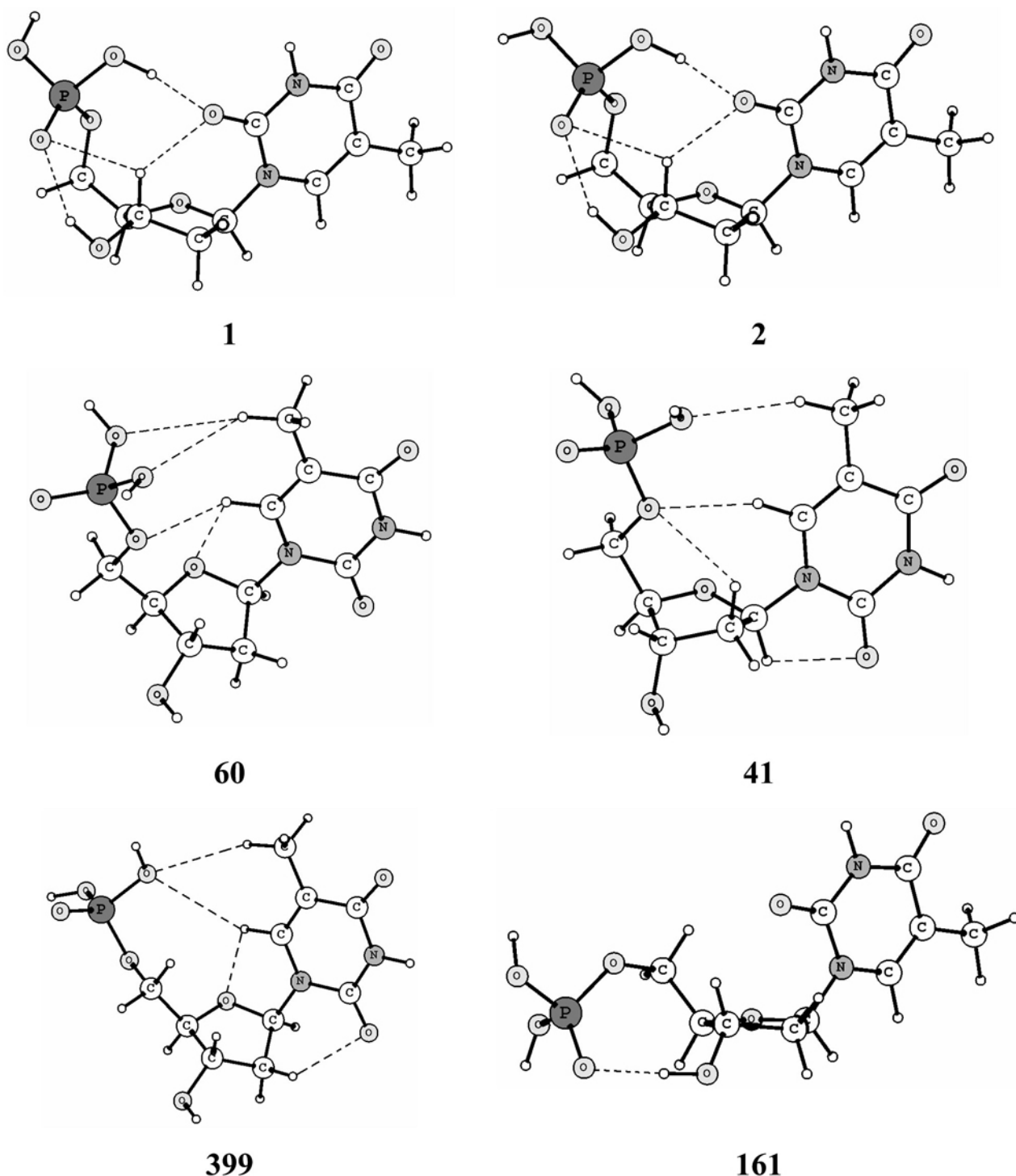


Рис. 2. Структури вибраних конформерів молекули 5'-тимідилової кислоти: енергетично найвигідніших (1 і 2), біологічно важливих AI- (60), AII- (399) і VI- (41) ДНК-подібних та структури із плоским цукровим кільцем (161)

коефіцієнти лінійної кореляції r виявилися більшими за 0,2 лише для трьох пар кутів: $r(\zeta_1, \zeta_2)=0,29$, $r(P, \nu_{\max})=0,32$ і $r(\chi, \nu_{\max})=0,52$, причому незначна кореляція між χ і ν_{\max} обумовлена тим, що плоскі ($\nu_{\max} < 22^\circ$) цукри завжди

супроводжуються *syn*-орієнтацією нуклеотидної основи (а саме $-\chi \in 70^\circ \pm 15^\circ$). Значення цих коефіцієнтів кореляції є близькими до аналогічних величин для молекули 2'-дезокситидилової кислоти [2].

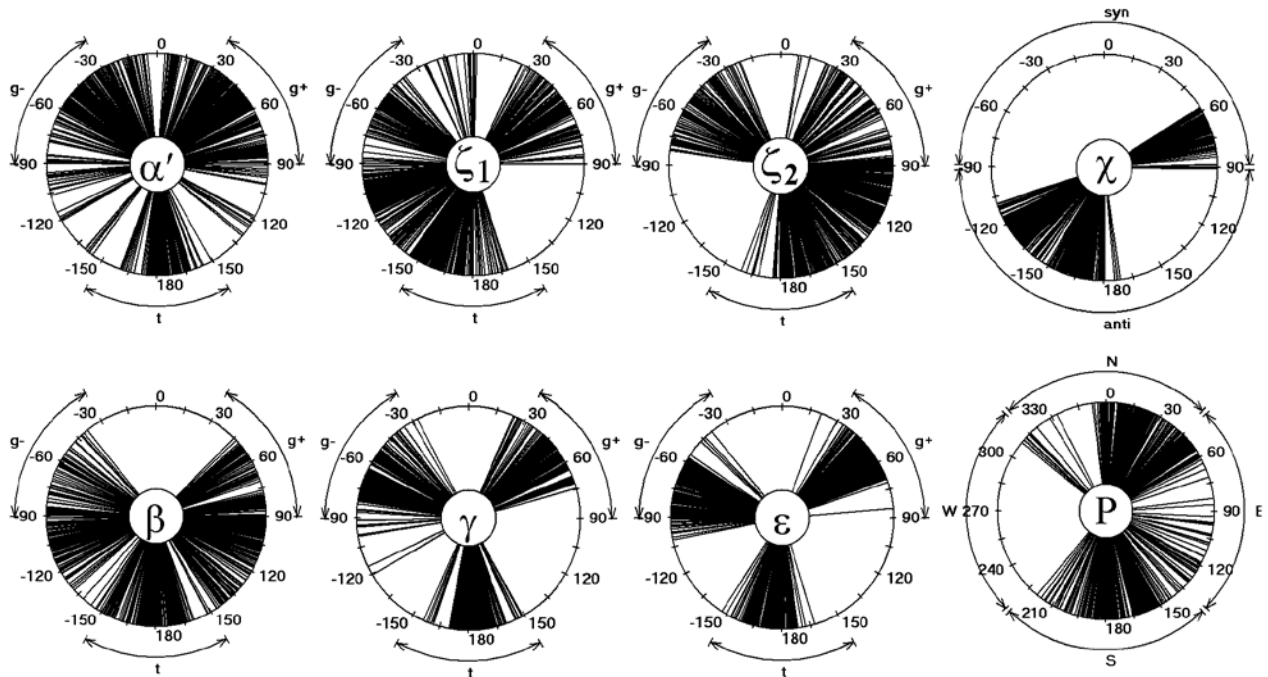


Рис. 3. Розподіл значень конформаційних параметрів в усіх можливих конформерах молекули 5'-тимідилової кислоти

Структурна мінливість. Значна кількість одержаних конформерів молекули 5'-тимідилової кислоти дозволила нам дослідити статистичні характеристики (середні значення і середньоквадратичні відхилення від них) «тонких» конформаційних ступенів вільності: ендо- та екзоциклічних торсійних кутів нуклеотидної основи (останні визначають положення її бічних радикалів), а також валентних кутів і довжин хімічних зв'язків. Порівняння мінливості значень цих структурних параметрів дає змогу виділити менш та більш лабільні («м'які») ступені вільності молекули.

Середньоквадратичне відхилення σ_l довжини l вибраного хімічного зв'язку обчислювали за формулою:

$$\sigma_l = \sqrt{\frac{1}{N_c} \sum_{i=1}^{N_c} (l_i - \langle l \rangle)^2},$$

де $N_c = 660$ – кількість конформерів, а середнє значення:

$$\langle l \rangle = \frac{1}{N_c} \sum_{i=1}^{N_c} l_i.$$

У табл. 2 наведено одержані $\langle l \rangle$ та σ_l для довжин тих хімічних зв'язків, у яких $\sigma_l / \langle l \rangle \geq 0,005$, а також найбільші відхилення $\Delta_{\max} = \max_k \{l_k - \langle l \rangle\}$ ($k = 1 \dots N_c$ – індекс, що ну-

мерує конформери) довжини l хімічного зв'язку від її середнього значення $\langle l \rangle$.

Одержані дані свідчать, що серед хімічних зв'язків найлабільнішим є глікозидний зв'язок C_1N_1 , що можливо обумовлено здатністю нуклеотидної основи утворювати численні водневі зв'язки із атомами фосфатної групи або фуранозного кільця (див., наприклад, конформери **41** та **60** на рис. 2). Решта порівняно «м'яких» хімічних зв'язків (в яких $\sigma_l / \langle l \rangle \geq 0,005$) належать до фосфатної групи, що дозволяє віднести фосфатну групу ДНК до її структурно нежорстких ланок.

Середнє значення $\bar{\theta}$ ($|\bar{\theta}| < \pi$) кутових змінних (валентних і торсійних кутів) визначали зі співвідношення [10] $\bar{\omega} \cdot e^{i\bar{\theta}} = \frac{1}{N_c} \sum_{k=1}^{N_c} e^{i\theta_k}$, з якого випливає, що

$$\bar{\theta} = \arctan(\bar{s} / \bar{c}),$$

де $\bar{s} = \frac{1}{N_c} \sum_{k=1}^{N_c} \sin \theta_k$ і $\bar{c} = \frac{1}{N_c} \sum_{k=1}^{N_c} \cos \theta_k$, а відхилення від середнього характеризували величиною $\delta_\theta \in [0; \pi/2]$

$$\delta_\theta = \arccos \bar{\omega},$$

де $\bar{\omega} = \sqrt{\bar{c}^2 + \bar{s}^2}$. Окрім того, визначали $\Delta_{\max} = \max_k \{\theta_k - \bar{\theta}\}$, враховуючи при обчислен-

Таблиця 2. Характеристики структурної мінливості хімічних зв'язків

Зв'язок	C_1N_1	C_5P	$H_{P1}O_{P1}$	$H_{P2}O_{P2}$	$O_{P1}P$	$O_{P2}P$
$\langle l \rangle$, Å	1,473	1,612	0,971	0,971	1,610	1,611
σ_p , Å	0,013	0,012	0,006	0,006	0,009	0,009
Δ_{\max} , Å	0,031	0,030	0,020	0,017	0,028	0,027

Таблиця 3. Характеристики структурної мінливості ендоциклічних торсійних кутів тиміну

Кут	$N_1C_2N_3C_4$	$C_2N_3C_4C_5$	$N_3C_4C_5C_6$	$C_4C_5C_6N_1$	$C_5C_6N_1C_2$	$C_6N_1C_2N_3$
$\bar{\theta}$, град.	0,7	-0,6	0,2	-0,1	0,2	-0,5
δ_θ , град.	2,2	0,9	0,5	0,9	2,7	3,1
Δ_{\max} , град.	8,5	5,7	1,7	2,4	9,2	11,0

Таблиця 4. Структурна мінливість торсійних кутів, що характеризують відхилення бічних радикалів від площини нуклеотидної основи (тиміну)

Радикал	O_2	H_3	O_4	H_6	C_7	H_{72}
Кут	$O_2C_2N_3C_4$	$H_3N_3C_4C_5$	$O_4C_4C_5C_6$	$H_6C_6C_5C_4$	$C_7C_5C_4N_3$	$H_{72}C_7C_5C_6$
$\bar{\theta}$, град.	-179,4	-179,7	-180,0	-179,1	-179,5	0,7
δ_θ , град.	1,6	1,5	0,3	1,3	1,3	2,8
Δ_{\max} , град.	7,5	6,4	1,0	2,5	6,0	14,1

ні різниці кутів $\theta_k - \bar{\theta}$ періодичність кутових змінних.

Проведений аналіз засвідчив, що лише для трьох валентних кутів характерне відхилення δ_θ від середнього перевищувало $2,5^\circ$: $C_2C_3O_3$ ($\delta_\theta = 2,9^\circ$), $C_4C_3O_3$ ($\delta_\theta = 2,8^\circ$) і $H_3C_3O_3$ ($\delta_\theta = 2,8^\circ$), причому мінливість цих кутів обумовлена втягуванням атома O_3 гідроксильної групи O_3H у внутрішньомолекулярні водневі зв'язки, тоді як група атомів $C_2C_3C_4$ є порівняно жорсткою (так, для валентного кута $C_2C_3C_4$ характерне відхилення від середнього складає $\delta_\theta = 0,4^\circ$).

Достатньо жорсткими є і ендоциклічні торсійні кути нуклеотидної основи (табл. 3).

Одержані дані вказують, що в цілому відхилення ендоциклічних торсійних кутів нуклеотидної основи у 5'-тимідилової кислоти від їхніх середніх значень менші, ніж у 2'-дезокситидилової кислоти [2]. В обох молекулах нуклеотидні основи є пласкими в середньому, оскільки середні значення усіх ендоциклічних торсійних кутів близькі до нуля.

Згідно з даними табл. 3 кут $C_6N_1C_2N_3$ є «найм'якшим» з усіх ендоциклічних торсійних

кутів – його значення найбільше відхиляється від середнього, а кут $N_3C_4C_5C_6$ є «найжорсткішим», оскільки його максимальне відхилення від середнього значення не перевищує 2° . Це означає, що в усіх можливих конформерах молекули 5'-тимідилової кислоти атоми N_3 , C_4 , C_5 і C_6 завжди знаходяться в одній площині.

Ці атоми вважали «точкою відліку» під час вибору торсійних кутів для кількісного опису виходу бічних радикалів нуклеотидної основи з її площини (табл. 4).

Серед них в усіх торсійних кутів, окрім $H_{72}C_7C_5C_6$, характерні відхилення від середніх значень в цілому менші, ніж у нуклеотидної основи 2'-дезокситидилової кислоти [2], що також свідчить про вищу механічну жорсткість тиміну порівняно з цитозином. Відзначимо, що найбільше відхилення кута $C_1N_1C_6C_5$, який характеризує відхилення атома C_1 від площини нуклеотидної основи, від його середнього значення $\bar{\theta} = 178,5^\circ$ складає $\Delta_{\max} = 8,5^\circ$ (при $\delta_\theta = 3,0^\circ$), що також значно менше, ніж у 2'-дезокситидилової кислоти [2], в окремих конформерах якої це відхилення досягає 20° .

Разом із цим, відхилення від середнього для кута $H_{72}C_7C_5C_6$, який характеризує поворот метильної групи тиміну навколо одинарного зв'язку C_7C_5 , у конформера **303** (див. табл. 1) досягає 14° , що є прямим структурним проявом внутрішньомолекулярного водневого зв'язку $C_7H_{72}\cdots O_p$ з енергією 1,2 ккал/моль (визначено методом [9]).

ДНК-подібні конформери. Серед усіх одержаних конформерів молекули 5'-тимідилової кислоти нами виявлено такі, в яких номенклатурні торсійні кути α , β , γ , ε , P та χ близькі до тих, що реалізуються у нуклеотидах у складі ДНК [1]: структура шести конформерів (**60**, **78**, **95**, **101**, **154** і **156** у табл. 1) подібна до ДНК в АІ-формі і семи конформерів (**41**, **48**, **55**, **61**, **73**, **103** і **129** у табл. 1) – до ВІ-ДНК; структура конформера **399** є АІІ-ДНК-подібною.

Серед виявлених нами А- і В-ДНК-подібних конформерів 5'-тимідилової кислоти найнижчу енергію Гіббса за нормальних умов ($\Delta G=2,7$ ккал/моль) має конформер **41** (рис. 2), структура якого близька до ДНК у ВІ-формі. Разом із тим, з усіх В-ДНК-подібних конформерів найближчою до В-ДНК є структура конформера **129**, відносна енергія Гіббса якого складає $\Delta G=4,3$ ккал/моль. Найбільш близькою до АІ-ДНК є структура конформера **156** ($\Delta G=4,6$ ккал/моль), тоді як найнижчу енергію Гіббса ($\Delta G=3,2$ ккал/моль) серед АІ-ДНК-подібних має конформер **60** (рис. 2).

Структури енергетично найвигідніших ДНК-подібних конформерів **41** і **60** стабілізують чотири внутрішньомолекулярні водневі зв'язки типу $C-H\cdots O$, із яких найбільшу енергію (визначену за [9]) у ВІ-ДНК-подібного конформера **41** має зв'язок $C_1H\cdots O_2$, (4,7 ккал/моль), а у АІ-ДНК-подібного конформера **60** – зв'язок $C_6H\cdots O_4$, (4,7 ккал/моль).

Зупинимося докладніше на тому, чим відрізняються один від одного конформери, подібні до ДНК у АІ- чи ВІ-формі. Ці відмінності зводяться до трьох чинників: локалізація протона у фосфатній групі (на атомах O_{1p} або O_{2p}), який забезпечує її електронейтральність; орієнтація відповідної гідроксильної групи (кута ζ_1 та/або ζ_2) та орієнтація фосфатної групи (кут α'). Остання визначається перш за все конфігурацією водневих зв'язків між метильною групою тиміну і атомами кисню фосфатної групи. Так, у ДНК-подібних конформерів **41**, **95**, **103**, **129**, **156** і **399** виявлено внутрішньомолекулярний водневий зв'язок $C_7H\cdots O_{1p}$,

у конформерів **48**, **55** і **154** – $C_7H\cdots O_{2p}$, у **61**, **78** і **101** – $C_7H\cdots O_p$; у конформера **60** присутні обидва водневі зв'язки: $C_7H\cdots O_{1p}$ і $C_7H\cdots O_{2p}$, а у конформера **73** водневі зв'язки між метильною групою тиміну і фосфатною групою відсутні.

Локалізація вільного протона на атомі O_{1p} фосфатної групи забезпечує більше можливостей для утворення внутрішньомолекулярних $C-H\cdots O$ водневих зв'язків, ніж за його локалізації на атомі O_{2p} . Такі зв'язки стабілізують структури відповідно 4 із 6 та 5 із 7 АІ- і ВІ-ДНК-подібних конформерів 5'-тимідилової кислоти.

Як А-, так і В-формам ДНК відповідають декілька конформацій ізольованого нуклеотиду. Цей висновок разом із даними [2] дозволяє стверджувати, що досліджуючи нуклеотид як структурну ланку ДНК недостатньо розглядати лише одну його конформацію як модель тієї чи іншої форми ДНК (як, наприклад, це зроблено у [11]).

Таким чином, виявлено 660 конформерів молекули 5'-тимідилової кислоти, відносні значення енергії Гіббса, яких за нормальних умов знаходяться у межах $0\div 11,1$ ккал/моль. Серед них структури шести, одного та семи конформерів подібні до структур ДНК у АІ, АІІ і ВІ формах відповідно, причому найнижчу відносну енергію Гіббса ($\Delta G=2,7$ ккал/моль) має ВІ-ДНК-подібний конформер. В усіх одержаних конформерах значення глікозидного кута χ потрапляють у одну із областей: $73^\circ\pm 20^\circ$ або $-144^\circ\pm 50^\circ$, межі яких дають природне визначення для понять *syn*- і *anti*-орієнтацій нуклеотидної основи відповідно. Показано, що в усіх можливих конформерах ізольованої 2'-тимідилової кислоти відсутні значимі кореляції між більшістю торсійних кутів за винятком пари (χ, ν_{\max}) з коефіцієнтом лінійної кореляції 0,52. Аналіз структурної мінливості «тонких» конформаційних ступенів вільності показав, що кільце тиміну більш жорстке, ніж кільце цитозину. Продемонстровано роль внутрішньомолекулярних водневих зв'язків типу $C-H\cdots O$ у формуванні просторової структури молекули 2'-дезокситимідилової кислоти.

Автори висловлюють щирю вдячність Інституту теоретичної фізики ім. М. М. Боголюбова НАН України за надане програмне забезпечення і обчислювальні ресурси для проведення розрахунків.

**КВАНТОВО-МЕХАНИЧЕСКИЙ
КОНФОРМАЦИОННЫЙ АНАЛИЗ
МОЛЕКУЛЫ 5'-ТИМИДИЛОВОЙ
КИСЛОТЫ**

Т. Ю. Николаенко¹, Л. А. Булавин¹,
Д. Н. Говорун^{1,2}

¹Киевский национальный университет
имени Тараса Шевченко, Украина;

²Институт молекулярной биологии
и генетики НАН Украины, Киев;
e-mail: tim_mail@ukr.net

Методами вычислительной квантовой механики на уровне теории MP2/6-311++G(d,p) // B3LYP/6-31G(d,p) проведен полный конформационный анализ структурного элемента ДНК – нуклеотида с тиминовой основой и электронейтральной фосфатной группой в 5'-положении. Выявлены 660 его конформаций с относительной энергией Гиббса в пределах 0–11,1 ккал/моль при нормальных условиях. Среди них структура шести конформеров подобна структуре нуклеотида ДНК в AI-, одного – в AII- и семи – в BI-форме соответственно. При этом наименьшую энергию Гиббса среди ДНК-подобных конформеров ($\Delta G = 2,7$ ккал/моль) имеет BI-ДНК-подобная структура. Показано, что гликозидная связь является наиболее лабильной из всех химических связей молекулы. Продемонстрирована роль внутримолекулярных водородных связей типа $CH\cdots O$ в формировании пространственной структуры конформеров 5'-тимидиловой кислоты.

Ключевые слова: нуклеотид, 5'-тимидиловая кислота, конформационный анализ, ДНК, *ab initio* расчеты, водородные связи.

**QUANTUM-MECHANICAL
CONFORMATIONAL ANALYSIS OF THE
5'-THYMIDILIC ACID MOLECULE**

T. Yu. Nikolaienko¹, L. A. Bulavin¹,
D. M. Hovorun^{1,2}

¹Taras Shevchenko Kyiv National University, Ukraine;

²Institute of Molecular Biology and Genetics,
National Academy of Sciences of Ukraine, Kyiv;
e-mail: tim_mail@ukr.net

S u m m a r y

The conformational analysis of the DNA structural unit – the nucleotide with thymine base and electroneutral phosphate group at 5'-position – has been carried out with the applied quantum mechanics methods at the MP2/6-311++G(d,p) // B3LYP/6-31G(d,p) theory level. As many as 660 conformations with relative Gibbs energies under standard conditions from 0 to 11.1 kcal/mole have been found. Among them, six conformations are similar to the structure of the nucleotide of AI-DNA, one – to AII- and seven – to the DNA in BI-form. The lowest Gibbs energy among the DNA-like conformations ($\Delta G = 2.7$ kcal/mole) belongs to BI-DNA-like structure. It is shown that the glycoside chemical bond is the most labile one. The role of intramolecular $CH\cdots O$ hydrogen bonds in formation of the 5'-thymidilic acid molecule structure is demonstrated.

Key words: nucleotide, 5'-thymidilic acid, conformational analysis, DNA, *ab initio* calculations, hydrogen bonds.

1. Борискина Е. П., Ткаченко М. Ю., Шестопалова А. В. // Biopolymers and Cell. – 2010. – № 5. – С. 360–372.
2. Николаенко Т. Ю., Говорун Д. М. // Доп. НАНУ. – 2010. – № 9. – С. 173–184.

3. Жураківський Р. О., Говорун Д. М. // *Ukrainica Bioorganica Acta*. – 2006. – № 2. – С. 56–63.
4. Зенгер В. Принципы структурной организации нуклеиновых кислот. – Москва: Мир, 1987. – 584 с.
5. Frisch M. J., Trucks G. W., Schlegel H. B., et al. Gaussian 03, Revision E.01 / Gaussian, Inc., Wallingford CT, 2004.
6. Ochterski J. W. Vibrational Analysis in Gaussian // Режим доступу: http://www.gaussian.com/g_whitepap/vib.htm
7. Бэйдер Р. Атомы в молекулах. Квантовая теория. – Москва: Мир, 2001. – 532 с.
8. Иогансен А. В. // Водородная связь / Под ред. Соколова Н. Д. – Москва: Химия, 1981. – С. 112–155.
9. Espinosa E., Molins E., Lecomte C. // *Chem. Phys. Lett.* – 1998. – **285**. – P. 170–173.
10. Kitamura K., Wakahara A., Mizuno H. et al. // *J. Am. Chem. Soc.* – 1981. – **103**. – P. 3899–3904.
11. Shishkin O. V., Gorb L., Zhikol O. A., Leszczynski J. // *J. Biomol. Struct. Dyn.* – 2004. – **21**, N. 4. – P. 537–554.

Отримано 06.12.2010