



УДК 547.963.32:539.26

5'-ДЕЗОКСИЦИТИДИЛОВАЯ КИСЛОТА: СТРУКТУРА
И КОНФОРМАЦИЯ

Булева В. Н., Воробьев Ю. Н., Гатиллов Ю. В.

Институт органической химии Сибирского отделения
Академии наук СССР, Новосибирск

Проведено рентгеноструктурное исследование dCMP. Найдено, что кристалл dCMP — моногидрат. Кристаллы ромбические: a 6,778(3); b 11,315(10); c 16,776(10) Å; Z 4. Пространственная группа — $P2_12_12_1$, асимметрической частью структуры является одна молекула dCMP и одна молекула H_2O . Структура установлена прямым методом определения фаз и уточнена в анизотропном полноматричном приближении, R 0,075. Найдены следующие конформации для составляющих dCMP: *анти* — для нуклеозида ($\varphi_{CN} = -7,2^\circ$), $C3'$ -*экзо* — для $2'$ -дезоксирибозы и *гош-гош* — для экзациклической связи $C4'$ — $C5'$. Упаковка молекул dCMP и H_2O в кристалле характеризуется большим числом водородных связей (10 на одну молекулу dCMP).

Вопрос о структуре и конформации «мономерной единицы» полинуклеотида — мононуклеотида важен для определения конформации и для понимания строения и функций нуклеиновых кислот.

Рентгеноструктурное исследование кристаллов нуклеозидов, нуклеотидов и динуклеотидов составляет базу для построения моделей нуклеиновых кислот и полинуклеотидов. Методом рентгеноструктурного анализа изучена структура и конформация следующих нуклеотидов: цитидин-3'-фосфата [1—3], цитидин-2'-фосфата [3], аденозин-3'-фосфата [4], аденозин-5'-фосфата [5], уридин-5'-фосфата [6], тимидин-5'-фосфата [7] и дезоксигуанозин-5'-фосфата [8].

Цель настоящей работы — установление структуры и конформации 5'-дезоксипитидиловой кислоты $C_9H_{14}O_7N_3P$.

Строение молекулы, длины связей и валентные углы dCMP представлены на рис. 1. Найденные длины связей и углы удовлетворительно согласуются с литературными [1—13]. Стандартные отклонения для длин связей и углов составляют 0,02 Å и $1,5^\circ$ соответственно. Сравнение длин связей с данными для $2'$ -дезоксирибозы с $3'$ -*экзо*-конформацией [12] показывает, что связь $C1'$ — $C2'$ в нашем случае на 0,08 Å длиннее.

Конформация гетероциклического основания является практически плоской с выведением атома $C2$ на 0,05 Å из средней плоскости основания. Заместители кольца значительно смещены из средней плоскости: $N4$ на 0,1; $O2''$ на 0,14 и $C1'$ на 0,13 Å, аналогично литературным данным [14], причем все атомы ($C2$, $N4$, $O2''$ и $C1'$) смещены в сторону, противоположную рибозофосфату.

Найденная в данной работе конформация $2'$ -дезоксирибозы соответствует $C3'$ -*экзо*-конформации, довольно редко реализующейся в кристаллах.

Сокращение: dCMP — 5'-дезоксипитидиловая кислота.

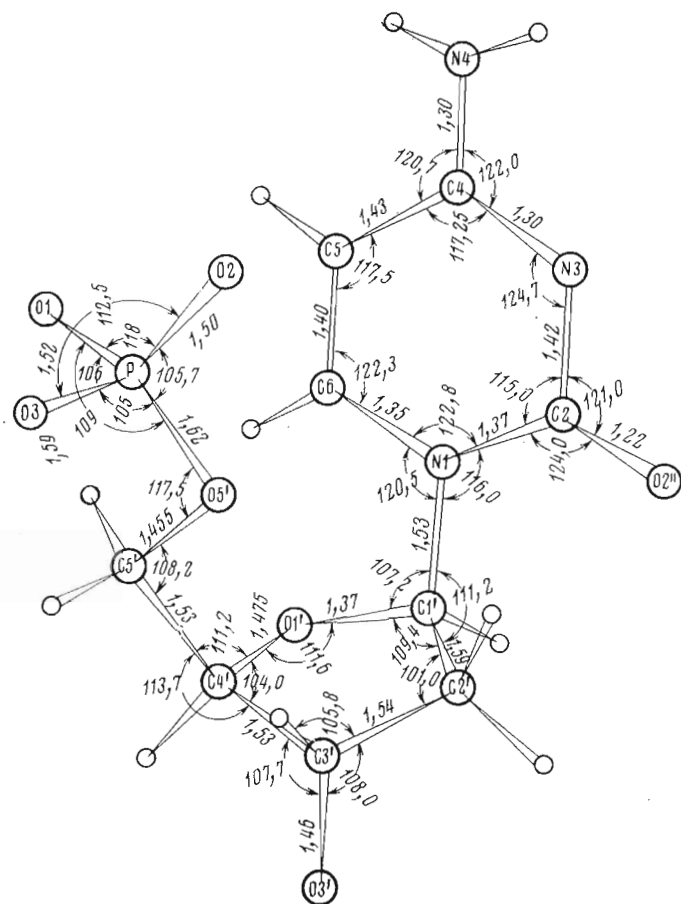


Рис. 1. Длины связей (Å) и валентные углы (град) в dCMP;

Отклонения от базисной плоскости, проходящей через $C1'$, $O1'$ и $C4'$, для $C3'$ равно $0,54 \text{ \AA}$, а для $C5'$ — $1,39 \text{ \AA}$. Значения торсионных углов в кольце дезоксирибозы (табл. 1) согласуются с ранее установленными для $C3'$ -экзо-конформации [15—17]. Согласно проведенным рентгеноструктурным исследованиям, $C2'$ -эндо- и $C3'$ -эндо-конформации найдены в большинстве изученных рибонуклеозидов и нуклеотидов [1—9, 11, 13—15], $C3'$ -экзо-конформация обнаружена только в дезокси-нуклеозидов и В-ДНК [12, 14, 17, 18]. Полуэмпирические конформационные расчеты свидетельствуют о наличии менее глубокого минимума энергии в области $C3'$ -экзо-конформации, чем в $C3'$ -эндо- и $C2'$ -эндо-конформациях, хотя относительная глубина минимумов энергии, по-видимому, существенно зависит от корректности учета кулоновских взаимодействий [16, 19, 20]. Важную роль в устойчивости той или иной конформации рибозы играет также ее участие в образовании водородных связей [9, 12].

Взаимное расположение пиримидинового кольца и остатка дезоксирибозы характеризуется прежде всего двугранным углом между их средними плоскостями, равным 55° . По данным работы [14], он находится в пределах $55—90^\circ$ для различных исследованных нуклеозидов и нуклеотидов. Второй важнейшей характеристикой расположения является угол вращения остатка сахара относительно гетероциклического основания вокруг связи $C1'—N1$, т. е. торсионный угол φ_{CN} ($O1'—C1'—N1—C6$). Найденное значение φ_{CN} равно $-7,2^\circ$, что соответствует анти-конформации нуклеотида. Согласно теоретическим представлениям [1, 21], существ-

Сравнение торсионных углов (град) 2'-дезоксирибозы dCMP с литературными данными для C3'-экзо-конформации рибозы

Торсионный угол	dCMP	[11] *	[10] **		[12] ***
			DPFUD ¹	DA ²	
C4'-O1'-C1'-C2'	4,9	0,0	9,9	-2,2	-4,0
O1'-C1'-C2'-C3'	13,2	24,8	8,0	23,9	24,7
C1'-C2'-C3'-C4'	-25,2	-38,1	-21,3	-34,4	-34,8
C2'-C3'-C4'-O1'	28,8	38,9	27,7	33,6	33,3
C3'-C4'-O1'-C1'	-21,1	-25,0	-23,5	-19,9	-18,6

* Теоретический расчет значений геометрических параметров для C3'-экзо-конформации β-D-рибозы.

** Значения углов, полученные из анализа рентгеноструктурных данных для: ¹ 5-[1-(2'-деокси-β-D-рибофуранозил)-урацилил]дисульфида, ² дезоксиаденозина.

*** Данные уточненной молекулярной модели В-ДНК.

вуют две наиболее вероятные области этого угла: $30 \pm 45^\circ$ (*анти*-область) и $150 \pm 45^\circ$ (*син*-область), причем для пиримидиновых нуклеозидов более выгодна *анти*-конформация [14, 19]. Возможность и преимущественность *анти*- или *син*-конформации зависит также от конформации рибозы. В работе [12] конформационным анализом было показано, что в случае C3'-экзо-рибозы пиримидиновые нуклеотиды могут находиться только в *анти*-конформации, что подтверждается экспериментальными данными [12, 14, 18].

Для экзоциклической углерод-углеродной связи C4'-C5' нами установлена *гош-гош*-конформация, т. е. *гош*-расположение связи C5'-O5' относительно связей C4'-O1' и C4'-C3', торсионные углы C3'-C4'-C5'-O5' и O1'-C4'-C5'-O5' равны 57 и -60° соответственно.

В исследованных методом рентгеноструктурного анализа нуклеотидах [9, 14] также найдена *гош-гош*-конформация экзоциклической углеродной связи. Теоретические работы [23, 24], в которых наиболее полно учитывалось взаимодействие всех компонентов нуклеотида, установили предпочтительность *гош-гош*-конформации для нуклеотидов в *анти*-конформации, особенно для 5'-нуклеотидов, содержащих цитозин.

Торсионный угол C4'-C5'-O5'-P, по нашим данным, равен 165° . Это соответствует *транс*-положению связей O5'-P и C4'-C5', что согласуется с литературными данными для всех исследованных нуклеотидов [11, 14, 15, 25]. Торсионные углы C5'-O5'-P-O1, C5'-O5'-P-O2 и C5'-O5'-P-O3 в фосфатной группе равны 316, 189 и 69° соответственно и хорошо соотносятся с известными из литературы [15, 19, 27].

Для упаковки молекул в кристаллах нуклеозидов и нуклеотидов характерно наличие водородных связей как между молекулами основного вещества, так и с присутствующими в кристалле молекулами воды [1, 12].

В нашем случае установлено, что каждая молекула dCMP имеет 11 коротких (длиной менее 3 Å) контактов с окружающими молекулами dCMP и воды (рис. 2 и табл. 2). Так как эти контакты короче простой суммы ван-дер-ваальсовых радиусов, они могут быть идентифицированы как некоторые связи. Анализ показал, что 10 из 11 связей водородные: 3 связи с молекулами воды и 7 связей с четырьмя молекулами dCMP.

Таблица 2
Водородные связи в кристалле dCMP

Тип связи	Длина связи, Å
N-H...O	
N4 O2	2,923
N4 O1	2,674
O-H...O	
O3 O3'	2,569
O3' H ₂ O	2,686
H ₂ O O1	2,687
H ₂ O O2''	2,902

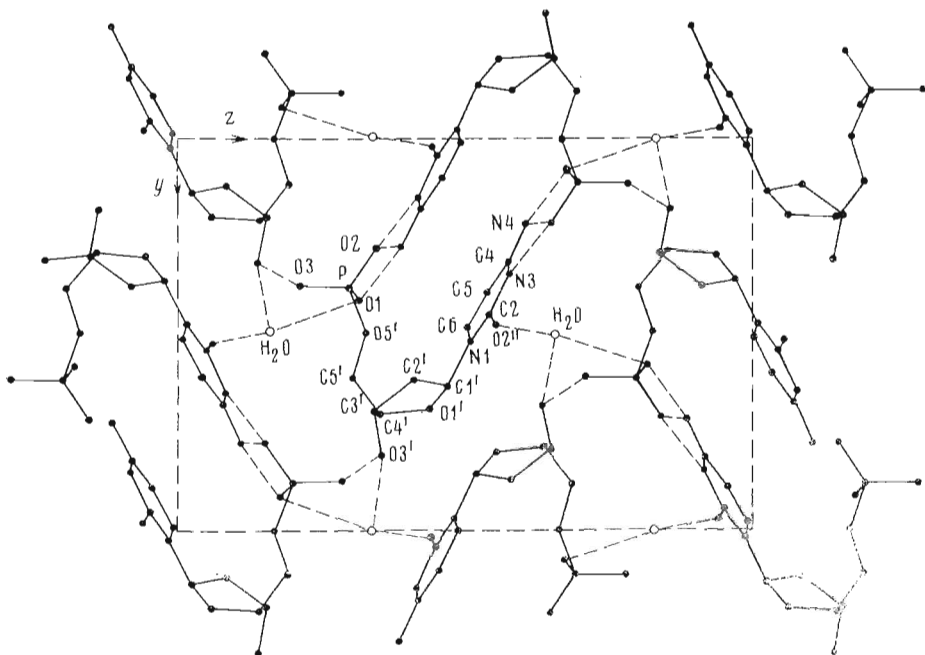


Рис. 2. Упаковка молекул dCMP в кристалле и короткие контакты (проекция на плоскость yz)

Контакт N3—O2 длиной 2,858 Å может быть водородной связью, если N3 протонирован либо у O2 имеется водород. Однако наличие водорода у N3 или O2 достоверно установить не удалось, поэтому можно предположить, что контакт N3—O2, который на 0,09 Å короче простого ван-дер-ваальсового контакта, является либо слабой ионно-солевой связью, либо напряженным межмолекулярным контактом.

Экспериментальная часть

dCMP легко кристаллизуется из водно-этанольного раствора в ромбической форме с параметрами ячейки: a 6,778(3); b 11,315(10); c 16,776(10) Å; Z 4. Пространственная группа $P2_12_12_1$, асимметрической частью структуры является одна молекула dCMP и одна молекула H_2O .

Трехмерный набор интенсивностей получен с монокристалла dCMP размерами 0,3 × 0,3 × 0,6 мм на автоматическом четырехкружном дифрактометре P2₁ фирмы Sintex (MoK α с графитовым монохроматором) θ — 2θ -сканированием. Обработка массива интенсивностей (1393 отражения) проведена по программам XTL. Массив значений структурных факторов $|F_{hkl}| > 3\sigma_F$, найденный из массива интенсивностей без учета поглощения, составляет 1018 независимых отражений.

Структура расшифрована прямым методом по программам Multan XTL. Расчет синтеза Фурье по найденным фазам позволил идентифицировать положение почти всех атомов (R на этой стадии равен 0,28).

Для дальнейшей расшифровки проведено 2 цикла уточнения в изотропном блок-диагональном приближении ($R = 0,19$). Повторным расчетом синтеза Фурье определены координаты всех атомов C, N, O, P и наличие одной молекулы воды в асимметрической части ячейки. При уточнении структуры была предпринята попытка ввести атомы водорода в соответствующие водородсодержащие группы. Однако это не дало заметного снижения R , а температурный фактор атомов водорода возрастал до ~ 12 — 15 Å², поэтому атомы водорода в дальнейшем не учитывались.

Координаты атомов и их стандартные отклонения

Атом	x/a	y/b	Z/c	Атом	x/a	y/b	Z/c
P	0,0830 (5)	0,3875 (3)	0,3043 (2)	O1'	0,2263 (11)	0,6947 (6)	0,4388 (5)
O1	-0,4332 (11)	0,4173 (7)	0,3188 (4)	N1	0,3009 (15)	0,5207 (8)	0,5074 (5)
O2	0,1699 (12)	0,2835 (7)	0,3478 (5)	C2	0,4331 (19)	0,4516 (11)	0,5482 (7)
O3	0,4072 (13)	0,3774 (7)	0,2104 (4)	O2''	0,6075 (12)	0,4759 (7)	0,5557 (5)
O5'	0,2468 (12)	0,5002 (7)	0,3294 (4)	N3	0,3546 (14)	0,3455 (8)	0,5799 (6)
C5'	0,1486 (19)	0,6163 (10)	0,3042 (8)	C4	0,1586 (18)	0,3158 (10)	0,5796 (7)
C4'	0,2648 (18)	0,7088 (11)	0,3525 (7)	N4	0,0921 (18)	0,2233 (8)	0,6172 (5)
C3'	0,4864 (18)	0,6983 (11)	0,3442 (8)	C5	0,0265 (18)	0,3936 (12)	0,5384 (7)
O3'	0,5690 (13)	0,8160 (6)	0,3577 (5)	C6	0,1066 (20)	0,4958 (12)	0,5441 (7)
C2'	0,5535 (17)	0,6178 (10)	0,4133 (6)	H ₂ O	0,3516 (12)	0,0029 (6)	0,3035 (5)
C1'	0,3806 (18)	0,6385 (9)	0,4757 (6)				

Таблица 4

Анизотропные (тензоры B_{ij} в диагональном* виде) и изотропные температурные факторы**

Атом	B_{11}	B_{22}	B_{33}	$B_{изотр}$	Атом	B_{11}	B_{22}	B_{33}	$B_{изотр}$
P	2,53	2,16	1,52	2,16	O1'	3,96	1,96	0,55	2,41
O1	4,63	2,35	1,47	3,03	N1	2,14	1,98	0,68	1,75
O2	2,97	2,52	1,58	2,48	C2	2,98	1,59	0,85	1,96
O3	3,73	2,69	1,98	2,93	O2''	4,12	2,21	1,01	2,67
C5'	3,06	2,19	1,45	2,36	N3	2,68	1,65	0,97	1,89
O5'	7,67	2,25	0,77	4,01	C4	3,22	1,58	0,73	2,01
C4'	2,62	1,75	1,38	2,0	N4	4,68	1,62	1,22	2,71
C3'	3,67	3,13	1,05	2,86	C5	3,37	1,98	1,37	2,38
O3'	3,48	2,67	1,22	2,65	C6	3,46	2,37	0,93	2,46
C2'	3,42	2,32	1,33	2,40	H ₂ O	3,91	3,42	1,47	3,15
C1'	2,6	1,39	0,40	1,63					

* $B_{12}=B_{13}=B_{23}=0$.** Поправка на анизотропные температурные колебания рассчитывается по уравнению $\exp\{-1/4(B_{11}h^2a^{*2}+B_{22}h^2b^{*2}+B_{33}h^2c^{*2})\}$, где a^* , b^* , c^* — параметры обратной решетки.

Окончательное уточнение было проведено в полноматричном анизотропном приближении, полученное значение R равно 0,075. Координаты атомов и их стандартные отклонения приведены в табл. 3, анизотропные и изотропные температурные факторы — в табл. 4.

ЛИТЕРАТУРА

- Sundaralingam M., Jensen L. H. (1965) J. Mol. Biol., 13, 914—929.
- Bugg C. E., March R. E. (1967) J. Mol. Biol., 25, 67—82.
- Kartha G., Ambady G., Viswamitra M. A. (1973) Science, 179, 495—496.
- Sundaralingam M. (1966) Acta Crystallogr., 21, 495—506.
- Kraut J., Jensen L. H. (1963) Acta Crystallogr., 16, 79—88.
- Sheffer E., Trueblood K. N. (1965) Acta Crystallogr., 18, 1067—1077.
- Trueblood K. N., Horn P., Luzzati V. (1961) Acta Crystallogr., 14, 965—982.
- Young D. W., Tollin P., Wilson H. R. (1974) Nature, 248, 513—514.
- Sundaralingam M., Jensen L. H. (1965) J. Mol. Biol., 13, 930—943.
- Arnott S. A. (1970) Progr. in Biophys. and Mol. Biol., 21, 265—319.
- Туманин В. Г. (1973) Молекулярная биология, т. 2 (Итоги науки и техники, ВИНИТИ АН СССР), с. 132—182, М.
- Watson D. G., Sutor D. J., Tollin P. (1965) Acta Crystallogr., 19, 111—119.
- Voet D., Rich A. (1970) Progr. Nucl. Acid. Res. Mol. Biol., 10, 183—265.
- Преображенская Н. А., Шабарова З. А. (1969) Успехи химии, 38, 222—247.
- Sundaralingam M. (1969) Biopolymers, 7, 821—860.
- Луговской А. А., Дашевский В. Г. (1972) Молекулярн. биология, 6, 440—448.
- Arnott S., Hukins D. W. (1972) Biochem. and Biophys. Res. Commun., 47, 1504—1509.

18. Tollin P., Wilson H. E., Young D. W. (1968) *Nature*, **217**, 1148—1149.
19. Дашевский В. Г. (1974) Конформация органических молекул, с. 400—423, «Химия», М.
20. Кистер А. А., Дашевский В. Г. (1975) Молекулярн. биология, **9**, 443—451.
21. Donohue J., Trueblood K. N. (1960) *J. Mol. Biol.*, **2**, 363—371.
22. Lakshminarayanan A. V., Sasisekharan V. (1969) *Biopolymers*, **8**, 475—488.
23. Vathindra N., Sundaralingam M. (1973) *Biopolymers*, **12**, 2261—2277.
24. Vathindra N., Sundaralingam M. (1973) *Biopolymers*, **12**, 279—314.
25. Lakshminarayanan A., Sasisekharan V. (1969) *Biopolymers*, **8**, 489—503.
26. Lakshminarayanan A., Sasisekharan V. (1969) *Biopolymers*, **8**, 505—514.

Поступила в редакцию
17.XII.1976

5'-DEOXYCYTIDYLIC ACID: STRUCTURE AND CONFORMATION

BUNEVA V. N., VOROB'EV Yu. N., GATILOV Yu. V.

*Institute of Organic Chemistry, Siberian Branch of
Academy of Sciences of the USSR, Novosibirsk*

The X-ray diffraction analysis of the dCMP crystal was performed. The crystal, which was found to be a monohydrate, is orthorhombic: a 6,778(3); b 11,315(10); c 16,776(10) Å; Z 4. Space group is $P2_12_12_1$, asymmetric part of the structure consists of one dCMP molecule and one water molecule. The structure was elucidated using direct methods of phase determination and refined by anisotropic method of full matrix to $R = 0,075$. The following conformations were determined for dCMP constituents: C3'-*exo* for 2'-deoxyribose, *anti* for nucleoside ($\varphi_{\text{CN}} = -7,2^\circ$) and *gauche-gauche* for exocyclic C4'—C5' bond. The molecules of dCMP and H₂O in the crystal lattice are firmly held together by the network of hydrogen bonds (10 H-bonds per one molecule of dCMP).
