

## НЕКАНОНИЧЕСКИЕ ПАРЫ GC И ЭПИГЕНЕТИКА

Нечипуренко Ю.Д.<sup>1</sup>, Семёнов Д.А.<sup>2</sup><sup>1</sup> Институт молекулярной биологии им. В.А. Энгельгардта РАН

г. Москва, РФ; e-mail: nech99@mail.ru

<sup>2</sup> Институт биофизики СО РАН

г. Красноярск, РФ

Поступила в редакцию: 20.07.2020

**Аннотация.** В течение последних десяти лет работы в области ЯМР ДНК позволили регистрировать не только канонические Уотсон-Криковские пары в ДНК, но и существующие несколько миллисекунд неканонические пары. Оказалось, что почти все возможные способы, которыми можно соединить нуклеотиды, реализуются при физиологических условиях в обычной ДНК. Мы проанализировали свойства неканонических пар оснований ГЦ и их связь с механохимическим расщеплением ДНК. Предложена гипотеза об участии переходных воббл пар оснований ГЦ (GC wobble) в механизмах механохимического расщепления ДНК и эпигенетических процессах с участием 5-метилцитозина. Гипотеза объясняет увеличение частоты разрывов сахарофосфатного остова ДНК после цитозинов, асимметричный характер этих разрывов и увеличение частоты разрывов CpG после метилирования цитозина. Эпигенетические свойства 5-гидроксиметилцитозина, 5-формилцитозина и N4-метилцитозина также находят своё объяснение в рамках нашей гипотезы. Появляется представление о том, что регуляторные белки узнают не метильную группу на цитозине, а гораздо более масштабное искажение структуры ДНК, вызванное GC wobble парой. GC wobble пара позволяет объяснить повышенную частоту дезаминирования цитозина, то есть высокую частоту мутаций в регуляторных участках.

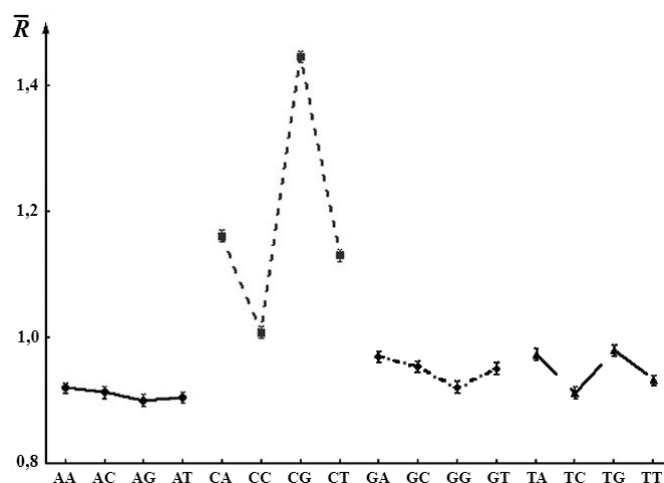
**Ключевые слова:** эпигенетика, GC воббл пары ДНК, механохимическое расщепление ДНК, таутомерия цитозина, переходные неканонические пары оснований.

Два направления работ в современной молекулярной биофизике близко подошли к фундаментальным основам эпигенетики. ЯМР исследования переходных неканонических пар нуклеотидов в ДНК развиваются последние 10 лет, прежде всего в работах Ai-Hashimi и соавторов [1,2]. Изучение селективности механохимического расщепления ДНК началось с работы Гроховского 2006 года и связано с отечественными исследователями [3-10]. Мы обнаружили, что два этих направления стыкуются, но при этом требуют предположить, что GC пара в ДНК может переходить в воббл (wobble) форму. Наше предположение позволяет объяснить существующие результаты механохимического расщепления ДНК. Гипотеза позволяет предложить простые способы проверки и при этом предсказывает системные связи между таутомерией, эпигенетикой, мутагенезом и конформационной динамикой нуклеотидов ДНК.

Работы С.Л. Гроховского и соавторов продемонстрировали неслучайный характер фрагментации ДНК ультразвуком: частота разрывов сахарофосфатного остова зависела от последовательности нуклеотидов. Механохимическая фрагментация ДНК является этапом подготовки к процедуре секвенирования нового поколения (NGS), это позволило подтвердить результаты на больших объемах экспериментальных данных. Одновременно большие объемы данных вскрыли сильное влияние метилирования цитозина на частоту разрыва ДНК после цитозина в динуклеотиде CpG [7-10]. Возможность объяснить эпигенетические эффекты метилирования цитозина выводят работы по механохимическому расщеплению ДНК на передний край современной молекулярной биологии.

В течение последних десяти лет работы в области ЯМР ДНК позволили регистрировать не только канонические Уотсон-Криковские пары в ДНК, но и существующие несколько миллисекунд неканонические пары [1,2]. Оказалось, что почти все возможные способы, которыми можно соединить нуклеотиды, реализуются при физиологических условиях в обычной ДНК. Удивительным результатом этого направления оказалась экспериментальная демонстрация высокой частоты встречаемости таких неканонических пар. Например, пара АТ при комнатной температуре в одном проценте случаев образует не Уотсон-Криковскую, а Хугстиновскую структуру. Такая распространенность позволяет высказать предположение о биологической значимости этих структур. Например, Хугстиновские пары рассматривались как связанные с эпигенетическими свойствами 5-метилцитозина в теоретической работе [11], но в дальнейшем эксперимент не подтвердил данную гипотезу [2]. Необходимо отметить, что сам метод, развиваемый Ai-Hashimi и соавторами, привлекает значительный интерес, так как открывает возможность использования ЯМР для изучения динамики ДНК и РНК. В основе этого метода лежит вычислительная модель подвижности нуклеотидов в ДНК, но эта модель требует долгой калибровки с привлечением данных множества экспериментов. В результате получается метод дорогостоящий, хотя потенциально мощный. Оправданием затрат может быть изучение биологически очень важных характеристик, например, эпигенетических свойств 5-метилцитозина.

**Гипотеза о роли GC wobble пар в эпигенетике и расщеплении ДНК ультразвуком.** Разрывы ДНК ультразвуком чаще всего происходят после цитозина и в значительной степени зависят от ближайшего



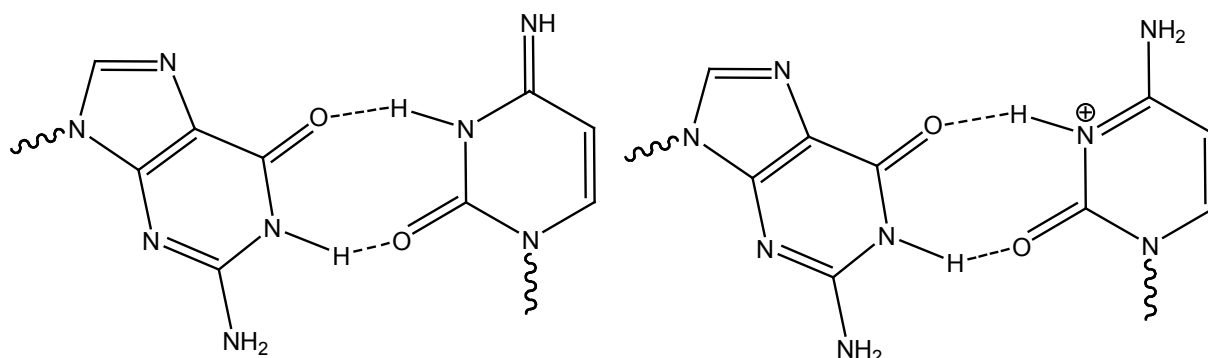
**Рисунок 1.** Средние значения относительных интенсивностей ультразвукового расщепления для динуклеотидов (данные из работы [6])

нуклеотидного окружения. В этом ряду выделяется динуклеотид CpG, для которого частота разрыва сахарофосфатного остова в цепи ДНК после цитозина (в направлении 5' – 3') в примерно в полтора раза выше, чем для других динуклеотидов, не содержащих в первом положении цитозина (рис. 1).

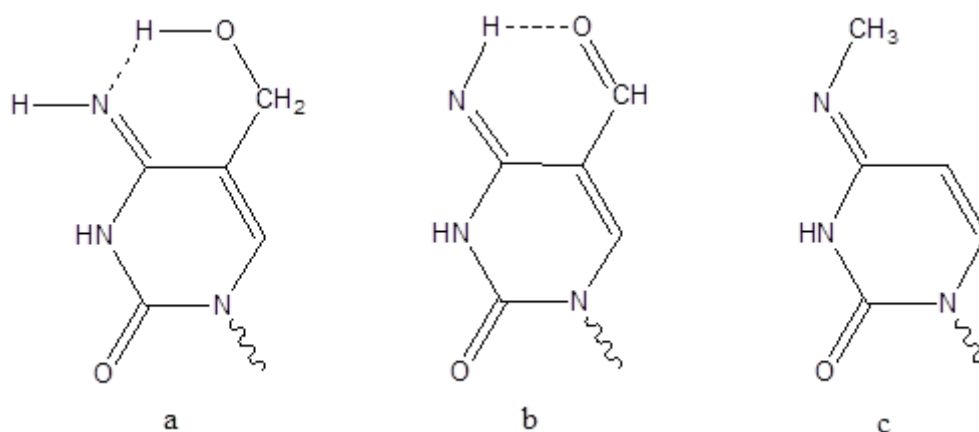
Причину повышения частоты разрыва можно искать в свойствах цитозина. Цитозин может участвовать в образовании неканонических пар оснований в различных структурах на основе ДНК. Например, кроме известных «параллельной ДНК» [12,13] и G-квадруплекса обнаружена ещё одна структура – I форма ДНК, которая образована четырьмя цитозинами [14]. Для цитозина характерна повышенная способность образовывать таутомерную (имино) форму. Таутомерная форма цитозина потенциально способна формировать GC wobble пары (рис. 2). Кроме склонности цитозина к таутомеризации, необходимо также указать на возможность протонирования азота N3 в слабокислой среде. Следует отметить, что протонированная форма цитозина также потенциально способна к формированию подобной пары (рис. 2, справа). Гроховский в пионерной работе по расщеплению ДНК ультразвуком [3] обратил внимание, что понижение pH до 5 значительно увеличивает фрагментацию, что как раз соответствует протонированию цитозина. Таким образом, предлагаемая пара объясняет один из экспериментально наблюдаемых феноменов.

**GC wobble пары и эпигенетика.** Для того, чтобы понять, как на формирование GC wobble пар влияет метилирование цитозина, необходимо принять во внимание, что метилирование способствует переходу цитозина в форму имино таутомера. Подобное влияние заместителя можно пронаблюдать еще у трех производных цитозина, демонстрирующих эпигенетические свойства: 5-гидроксиметилцитозина, 5-формилцитозина и N4-метилцитозина (рис. 3). Таким образом, мы получаем возможность привести все эпигенетические модификации цитозина к общему объяснению: модификации сдвигают таутомерное равновесие и управляют созданием неканонических GC wobble пар.

**Некоторые следствия предлагаемой гипотезы.** Возможно, метилирование цитозина в ДНК в составе динуклеотидов CpG облегчает появление переходных GC wobble пар. Такие пары могли бы узнаваться регуляторными белками. GU и GT пары могут появляться в ДНК под действием мутации дезаминирования цитозина и 5-метилцитозина, соответственно. Ферменты репарации узнают эти пары, то есть они взаимодействуют с локальным нарушением спирали ДНК. Вероятно, схожий механизм использует цитозин(C5)-ДНК-метилтрансфераза при поиске места метилирования ДНК для обнаружения переходных GC wobble пар.



**Рисунок 2.** Переходная GC wobble пара (слева) и GC+ wobble пара (справа)



**Рисунок 3.** Эпигенетические производные цитозина в форме имино таутомеров: 5-гидроксиметилцитозин (а), 5-формилцитозин (b) и N4-метилцитозин (с)

Регуляторные белки, узнающие метилированные цитозины в CpG в составе ДНК, могут использовать схожий механизм узнавания. Сравнение структур этих трех классов белков: ферментов репарации, метилтрансфераз и регуляторных белков, узнающих 5mCpG, может дать материал для дополнительного анализа предлагаемого эпигенетического механизма [15,16]. Заметим также, что в работах [15,16] приведены данные ЯМР цитозина, которые свидетельствуют в пользу нашей гипотезы.

Современные представления о ДНК отличаются от «статичной» конструкции Уотсона и Крика. ДНК «дышит» и «волнуется», её пары «качаются» и переходят в разные состояния. Управление неканоническими парами в ДНК - это управление не только структурой, но и динамикой, так как эти пары короткоживущие и запасенная при их создании энергия через короткое время будет преобразована в колебания ДНК. Эпигенетика всё больше становится объектом биофизики.

*Авторы считают своим приятным долгом поблагодарить Д.М. Грайфера за плодотворные дискуссии.*

*Работа поддержана Программой фундаментальных исследований государственных академий наук на 2013 - 2020 годы (темы №01201363818 и № 01201363820) и также грантом Президиума РАН по Молекулярной и клеточной биологии.*

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

1. Nikolova E.N., Goh G.B., Brooks C.L., Al-Hashimi H.M. Characterizing the Protonation State of Cytosine in Transient G•C Hoogsteen Base Pairs in Duplex DNA. *J. Am. Chem. Soc.*, 2013, vol. 135, no. 6766-6769. DOI: 10.1021/ja400994e.
2. Rangadurai A., Kremser J2., Shi H., Kreutz C., Al-Hashimi H.M. Direct evidence for (G)O6•••H2-N4(C)+ hydrogen bonding in transient G(syn)-C+ and G(syn)-m5C+ Hoogsteen base pairs in duplex DNA from cytosine amino nitrogen off-resonance R1ρ relaxation dispersion measurements. *J Magn Reson.*, 2019, vol. 308, p. 106589.
3. Гроховский С.Л. Специфичность расщепления ДНК ультразвуком. *Молекуляр. Биология*, 2006, т. 40, с. 317. [Grokhovskii S.L. The specific cleavage of DNA with ultrasound. *Molekuliarnaia biologii*, 2006, vol. 40, pp. 317-325. (In Russ.)]
4. Гроховский С.Л., Ильичева И.А., Нечипуренко Д.Ю., Панченко Л.А., Полозов Р.В., Нечипуренко Ю.Д. Локальные неоднородности структуры и динамики двухспиральной ДНК: исследование при помощи ультразвука. *Биофизика*, 2008, т. 53, с. 417-425. [Grokhovsky S.L., Il'icheva I.A., Nechipurenko D.Y., Panchenko L.A., Polozov R.V., Nechipurenko Y.D. Ultrasonic cleavage of DNA: quantitative analysis of sequence specificity. *Biophysics*, 2008, vol. 53, p. 250.
5. Нечипуренко Ю.Д., Головкин М.В., Нечипуренко Д.Ю., Ильичева И.А., Панченко Л.А., Полозов Р.В., Гроховский С.Л. Характерные особенности расщепления ДНК ультразвуком. *Журнал структурной химии*, 2009, т. 50, с. 1040-1047. [Nechipurenko Yu.D., Golovkin M.V., Nechipurenko D.Yu., Il'icheva I.A., Panchenko L.A., Polozov R.V., Grokhovskii S.L. Characteristics of ultrasonic cleavage of DNA. *Journal of Structural Chemistry*, 2010, vol. 50, pp. 1007-1013. (In Russ.)]
6. Grokhovsky S.L., Il'icheva I.A., Nechipurenko D.Yu., Golovkin M.V., Panchenko L.A., Polozov R.V., Nechipurenko Yu.D. Sequence-Specific Ultrasonic Cleavage of DNA *Biophysical Journal*, 2011, vol. 100, pp. 117-125.
7. Poptsova M.S., Il'icheva I.A., Nechipurenko D.Y., Panchenko L.A., Khodikov M.V., Oparina N.Y., Polozov R.V., Nechipurenko Y.D., Grokhovsky S.L. Non-random DNA fragmentation in next-generation sequencing. *Sci Rep.*, 2014, vol. 4, p. 4532.
8. Нечипуренко Ю.Д., Ильичева И.А., Абдуллаев Э.Т., Урошлев Л.А. и Гроховский С.Л. Физические характеристики регуляторных участков генома, эпигенетика и канцерогенез. *Актуальные вопросы биологической физики и химии*, 2018, т. 3, с. 884-887. [Nechipurenko Y.D., Il'icheva I.A., Abdullaev E.T., Uroshlev L.A., Grokhovsky S.L. Physico-chemical properties of DNA in regulatory sites of genomes, epigenetics and cancerogenesis *Russian Journal of Biological Physics and Chemistry*, 2018, vol. 3, pp. 884-887. (In Russ.)]

9. Ильичева И.А., Ходыков М.В., Панченко Л.А., Полозов Р.В., Нечипуренко Ю.Д. Ультразвуковое расщепление ДНК: анализ структурно-динамических характеристик регуляторных участков генома и ошибок секвенирования. *Биофизика*, 2020, т. 65, с. 504-511. [Il'icheva I.A., Khodikov M. V., Panchenko L.A., Polozov R.V., Nechipurenko Yu.D. Ultrasonic DNA Cleavage: Analysis of Conformational-Dynamic Features of Regulatory Regions in the Genome and Sequencing Errors *Biophysics*, 2020, vol. 65, pp. 504-511. (In Russ.)]
10. Uroshlev L.A., Abdullaev E.T., Umarova I.R., Il'icheva I.A., Panchenko L.A., Polozov R.V., Kondrashov F.A., Nechipurenko Y.D., Grokhovsky S.L. A Method for Identification of the Methylation Level of CpG Islands from NGS Data. *Sci. Rep.*, 2020, vol. 10, p. 8635.
11. Semenov D.A. *Epigenetic effects of cytosine derivatives are caused by their tautomers in Hoogsteen base pairs.* arXiv:1410.6763
12. Чуриков Н.А., Чернов Б.К., Голова Ю.Б., Нечипуренко Ю.Д. Параллельные ДНК – возможность существования. *Доклады Академии Наук СССР*, 1988, т. 303, с. 1254-1258. [Tchurikov N.A., Chernov B.K., Golova Y.B., Nechipurenko Y.D. Parallel DNA-possibility of existence. In Proc. *Acad. Sci. USSR*, 1988, vol. 303, pp. 1254-1258. (In Russ.)]
13. Tchurikov N.A., Chernov B.K., Golova Yu.B., Nechipurenko Yu.D. Parallel DNA: generation of a duplex between two Drosophila sequences in vitro. *FEBS Lett.*, 1989, vol. 257, pp. 415-418.
14. Zeraati M., Langley D.B., Schofield P., Moye A.L., Rouet R., Hughes W.E., Christ D. I-motif DNA structures are formed in the nuclei of human cells. *Nature chemistry*, 2018, vol. 10, pp. 631-637.
15. Semyonov D.A., Nechipurenko Y.D. *Non-Canonical GC Base Pairs and Mechanochemical Cleavage of DNA.* arXiv:2001.03561
16. Semyonov D.A., Eltsov I.V., Nechipurenko Y.D. A new bias site for epigenetic modifications: How Non-Canonical GC Base Pairs favour Mechanochemical Cleavage of DNA. *BioEssay*, 2020. DOI: 10.1002/bies.202000051.

#### NON CANONICAL BASE PAIRS & EPIGENETICS

Nechipurenko Y.D.<sup>1</sup>, Semyonov D.A.<sup>2</sup>

<sup>1</sup> Engelhardt Institute of Molecular Biology, Russian Academy of Sciences  
Moscow, Russia; e-mail: nech99@mail.ru

<sup>2</sup> Institute of Biophysics, Siberian Branch of Russian Academy of Science  
Krasnoyarsk, Russia

**Abstract.** Over the past ten years, work in the field of DNA NMR has made it possible to record not only canonical Watson-Crick pairs in DNA, but also non-canonical pairs that exist several milliseconds. It turned out that almost all possible ways to connect nucleotides are realized under physiological conditions in ordinary DNA. Properties of non-canonical GC Base Pairs and their relation with mechanochemical cleavage of DNA are analyzed here. A hypothesis of the involvement of the Transient GC Wobble Base Pairs in the mechanisms of the mechanochemical cleavage of DNA and epigenetic mechanisms with participation of 5-methylcytosine is proposed. The hypothesis explains the increase in the frequency of the breaks of the sugar-phosphate backbone of DNA after cytosine, asymmetric character of these breaks, and an increase in the frequency of breaks in CpG after cytosine methylation. The epigenetic properties of 5-hydroxymethylcytosine, 5-formylcytosine, and N4-methylcytosine are also explained within our hypothesis. There is an idea that regulatory proteins do not recognize the methyl group on the cytosine, but a much larger distortion of the DNA structure caused by the GC Wobble pair. GC Wobble pair allows to explain the increased frequency of cytosine deamination, that is, the high frequency of mutations in regulatory sites.

**Key words:** mechanochemical cleavage of DNA, transient Base Pairs in DNA, Wobble Base Pair, tautomer, 5-methylcytosine, molecular basis of epigenetics