

Epigenetic effects of cytosine derivatives are caused by their tautomers in Hoogsteen base pairs

Denis Semyonov.

E-mail address: dasem@mail.ru

Summary. *Deoxycytidine in solution exists as two tautomers one of which is an “uncanonical” imino one. The latter can dominate with such derivatives as 5-methyl, 5-hydroxymethyl- and 5-formylcytosine. The imino tautomer potentially is able to form a Hoogsteen GC base pair. To detect such pair, it is suggested to use ¹H¹⁵N NMR. Formation of GC-Hoogsteen base pair with imino tautomer of cytosine can be a reason for epigenetic effects of 5-methyl- and , 5-hydroxymethylcytosine.*

Introduction. Tautomers of heterocyclic bases of nucleic acids are formed via intramolecular transfer of a single proton, and cytosine can exist as the amino- (canonical) and imino- (rare) tautomer. NMR data on deoxycytidine solution (unpublished results) show the presence of both tautomers in comparable amounts. Derivatives of cytosine, such as 5-methylcytosine and 5-hydroxymethylcytosine, are expected to form the imino tautomer with higher probability. It is known that these derivatives actively participate in regulation of genes expression since DNA regions containing derivatized cytosines are recognized by special regulatory proteins. Participation in regulation of genes is considered as epigenetic effect of methylation and hydroxymethylation of cytosines. Tautomeric properties of cytosine derivatives can underlie their epigenetic effects.

The imino tautomer, which potentially can form a GC-Hoogsteen base pair. The ability of the imino-tautomer to form such pair has not been observed so far, and it is possible that it explains epigenetic effects of cytosine derivatives

1. It is known that deoxycytidine in solution at neutral pH exists in two tautomeric forms present in comparable amounts. These forms can be experimentally observed by two-dimensional ¹H¹⁵N NMR spectra of the solutions (natural ¹⁵N content is sufficient to obtain the spectra) if protons at the C5 and C6 atoms are not rapidly exchangeable with the solvent. Such spectra have been obtained in 2012 by Ilya El'tsov in the Novosibirsk State University (Novosibirsk, Russia) and can be reproduced by everybody who has NMR spectrometer and

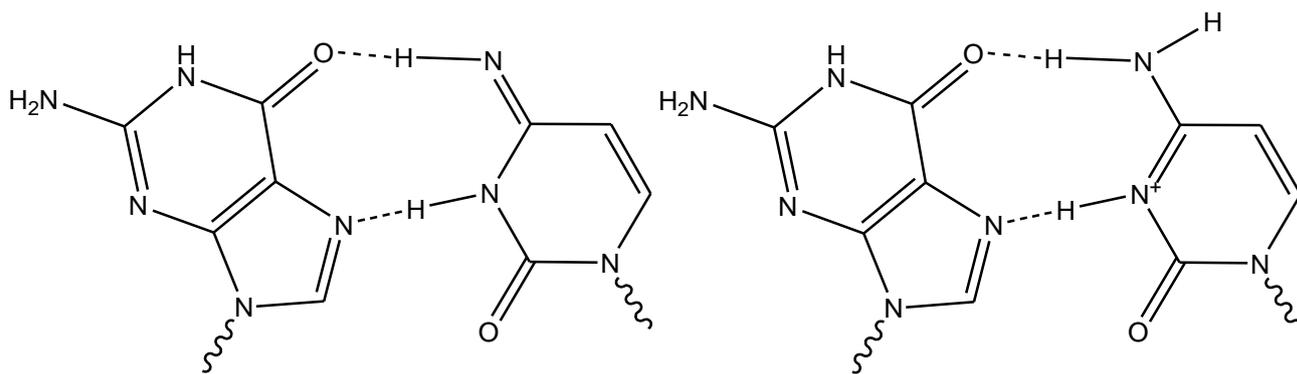


Figure 3. (a) GC-Hoogsteen pair formed by imino tautomer of cytosine as likely variant of the base pair formation at pH near 7. (b) Classical GC^+ -Hoogsteen base pair

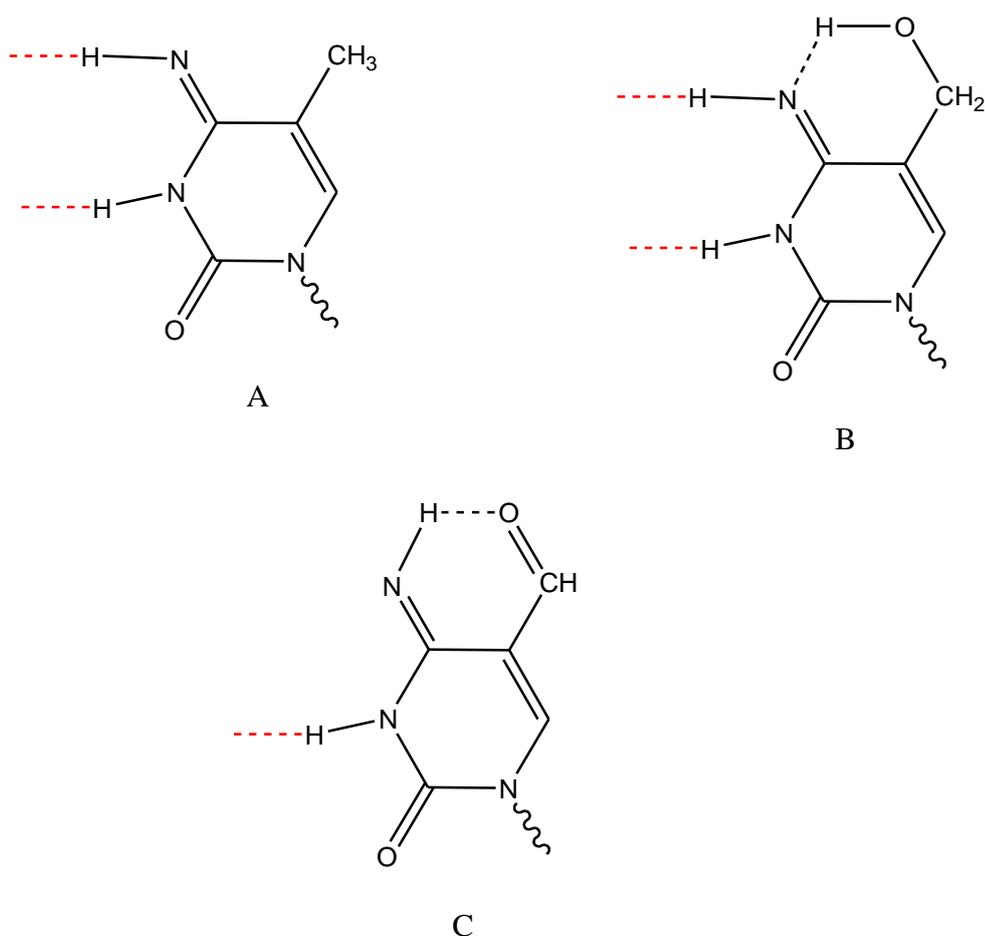


Figure 4. GC-Hoogsteen pair potentially could be formed with participation of 5-methyl- and 5-hydroxymethylcytosine but 5-formylcytosine is unable to form such type pair because of formation of the intramolecular hydrogen bond.

GC-Hoogsteen base pair was suggested to be the reason for epigenetic processes in the works of Al-Hashimi et al. [3, 4], where spontaneous formation of these pairs was demonstrated to occur in the DNA double helix. In these works only protonated GC⁺-Hoogsteen base pair was considered since the work has been carried out at pH<5,5. Under these conditions, protonated form of cytosine dominates, and the presence imino-tautomer is unlikely. In general, the experiments under acidic conditions do not completely explain epigenetic properties of cytosine derivatives. Protonated forms of cytosine and all his derivatives mentioned above should not have significant dissimilarities upon formation of pairs similar to the GC⁺-Hoogsteen base pair. It is worth to mention here that protonated form of 5-formylcytosine should be able to form GC⁺-Hoogsteen base pair. However, this cytosine derivative has no epigenetic effects, which indicates that these effects are not caused by protonated forms of cytosine derivatives.

One of methods used by Al-Hashimi et al. Could be applied for experimental testing of the suggestion on existence of GC-Hoogsteen pairs formed by imino tautomer of cytosine. To stabilize GC⁺-Hoogsteen pairs in [3] 1-methylguanidine was used, which prevented formation of the GC Watson-Crick base pair. At pH 7 this approach is expect to make possible detection of the suggested base pair (Fig. 6). To obtain a proof for the tautomeric structure of cytosine participating in the base pair formation, it is necessary to apply ¹H¹⁵N NMR. In these experiments chemical shifts of the nitrogen atoms should exactly show which forms of cytosine are involved, protonated or tautomeric. Information concerning relative stability of the tautomers originated from various cytosine derivatives can be obtained from the comparison of T_m of the corresponding duplexes.

There is an experimental indication for the increased probability of tautomeric transition with 5-formylcytosine. The latter can form a stable complementary pair with adenine wobble position of codon [5]. Formation of the imino tautomer of 5-formylcytosine in this pair was shown by X-ray crystallography [5], and the geometry of this pair is similar to that of the AU pair (Fig. 5). One can expect that studying pair 5-formylC-A by NMR approaches will make it possible confirm the tautomer existence and besides provide data on intramolecular hydrogen bond that deprives imino tautomer of 5-formylcytosine from ability to form GC-Hoogsteen base pair.

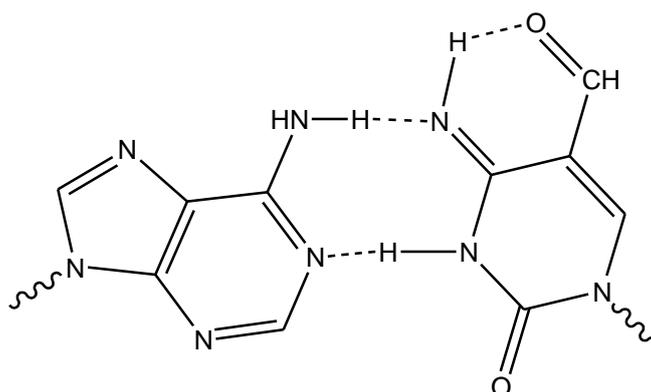


Figure 5. Pair adenine-5-formylcytosine formed by imino tautomer of the cytosine derivative. The tautomer can be stabilized by the intramolecular hydrogen bond that should be well recognizable on the NMR spectra.

The main conclusion: methylation and hydroxymethylation of cytosine can shift tautomer equilibrium to the imino tautomer, and the latter can form GC-Hoogsteen base pair at pH near 7, which can be the reason for epigenetic effects of 5-methylcytosine and 5-hydroxymethylcytosine.

Acknowledgements:

The author is grateful to D.M. Graifer for helpful discussion and his help in the text preparation.

References:

1. M.R. Branco, G. Ficiz, W. Reik. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet.* 2011 13(1):7-13
2. D.M. Franchini, K.M. Schmitz, S.K. Petersen-Mahrt. 5-Methylcytosine DNA Demethylation: More Than Losing a Methyl Group *Annual Review of Genetics* 2012 46: 419-441
3. E.N. Nikolova, G.B. Goh, C.L. Brooks, H.M. Al-Hashimi Characterizing the Protonation State of Cytosine in Transient G•C Hoogsteen Base Pairs in Duplex DNA. *J. Am. Chem. Soc.* 2013 135(18): 6766-9
4. E.N. Nikolova, H. Zhou, F.L. Gottardo, H.S. Alvey, I.J. Kimsey, H.M. Al-Hashimi. A Historical Account of Hoogsteen Base Pairs in Duplex DNA. *Biopolymers* 2012 99(12): 955-68
5. W.A. Cantara, F.V. Murphy, H. Demirci, P.F. Agris. Expanded use of sense codons is regulated by modified cytidines in RNA. *2013 Proc Natl Acad Sci USA* 110(27): 10964-9