Computational mapping reveals dramatic effect of Hoogsteen breathing on duplex DNA reactivity with formaldehyde

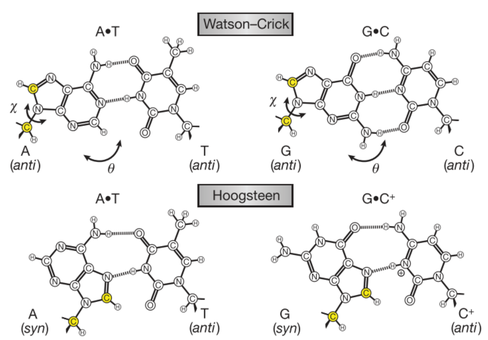
Tanggis Bohnuud, Dmitri Beglov, Chi Ho Ngan, Brandon Zerbe, David R. Hall, Ryan Brenke, Sandor Vajda, Maxim D. Frank-Kamenetskii and Dima Kozakov
Boston University

Abstract

Formaldehyde has long been recognized as a hazardous environmental agent highly reactive with DNA. Recently, it has been realized that, due to the activity of histone demethylation enzymes within the cell nucleus, formaldehyde is produced endogenously, in direct vicinity of genomic DNA. Should it lead to extensive DNA damage? We address this question with the aid of a computational mapping method, analogous to X-ray and NMR techniques for observing weakly specific interactions of small organic compounds with a macromolecule in order to establish important functional sites. We concentrate on the leading reaction of formaldehyde with free bases: hydroxymethylation of cytosine amino groups. Our results show that in B-DNA cytosine amino groups are totally inaccessible for the formaldehyde attack. Then we explore the effect of recently discovered transient flipping of Watson-Crick (WC) pairs into Hoogsteen (HG) pairs ("Hoogsteen breathing"). Our results show that the HG base pair formation dramatically affects the accessibility for formaldehyde of cytosine amino nitrogens within WC base pairs adjacent to HG base pairs. The extensive literature on DNA interaction with formaldehyde is analyzed in light of the new findings. The obtained data emphasize the significance of DNA HG breathing.

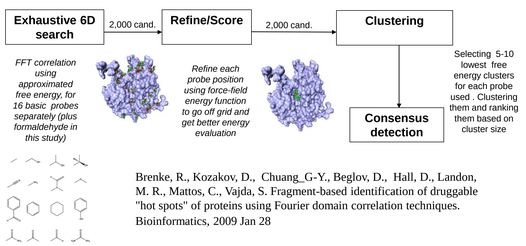
Bohnuud T *et al.*, Computational mapping reveals dramatic effect of Hoogsteen breathing on duplex DNA reactivity with formaldehyde. Nucleic Acids Res. 2012 Jun 16. [Epub ahead of print].

Motivation

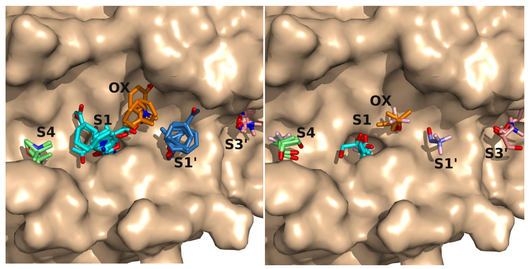


Nikolova EN *et al.*, Transient Hoogsteen base pairs in canonical duplex DNA. Nature. 2011

Method



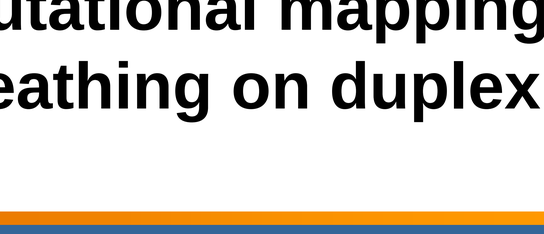
Brenke, R., Kozakov, D., Chuang_G-Y., Beglov, D., Hall, D., Landon, M. R., Mattos, C., Vajda, S. Fragment-based identification of druggable "hot spots" of proteins using Fourier domain correlation techniques. Bioinformatics, 2009 Jan 28



Comparison of computational and experimental results. A. Computational mapping of 1ela B. Experimental mapping superimposed on 1ela structure

Results

Identification of Binding hot spots in B-DNA



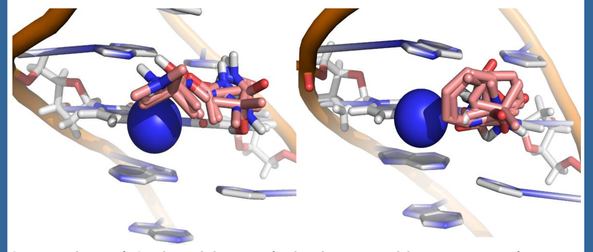
(A) The largest consensus site CS1 (with 23 probe clusters, shown as cyan mesh) overlaps with a fragment of the drug distamycin A (shown as sticks in green) in the minor groove. (B) Consensus sites CS1 (cyan), CS3 (yellow), and CS5 (purple) cover the entire drug binding site. Consensus sites CS2, CS4, and CS6-CS8 (all shown in wheat) extend the site in both directions. Consensus site C9 (with 10 probe clusters, shown in blue), is the largest hot spot located in the major groove. (C) Close-up of consensus sites CS1, CS3, and CS5, overlapping the bound structures of three drugs shown as sticks. The names and color codes of the drugs and the PDB codes of the structures are as follows: distamycin A, green, 267D; netropsin, salmon, 101D; and diamidine, blue, 1VZK.

Surface accessibility of Cytosine-N4 atom in DNA structures with HG or WC pairing

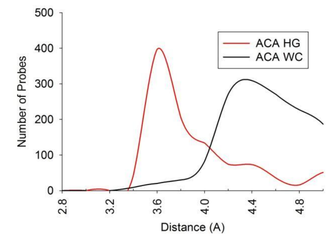
Sequence[a,b]	SA of C-N4 (Å^2)		
	HG	WC	HG-WC
G**C**A$_H$	19.71	14.93	4.78
A**C**A$_H$	19.38	15.05	4.32
C**C**A$_H$	15.60	12.27	3.33
T**C**A$_H$	14.50	8.59	5.90
T**C**$_H$A	7.63	8.46	-0.83
A**C**$_H$A	12.39	11.75	0.63
G**C**$_H$A	12.24	11.63	0.61
C**C**$_H$A	9.92	9.23	0.69
C**C**$_H$A	14.87	16.51	-1.63

[a] Subscript H indicates base with HG pairing.
[b] SA is calculated for C underlined.

Consensus sites around Cytosine N4

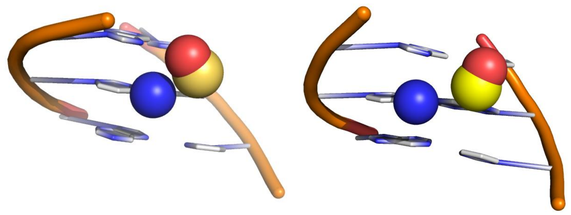


Consensus clusters of 16 probes and clustering of carbonyl group around the amino nitrogen of cytosine of DNA with an ACA triplet. The blue sphere indicates the cytosine's amino nitrogen atom. Left: HG pairing; right: all-WC paring.

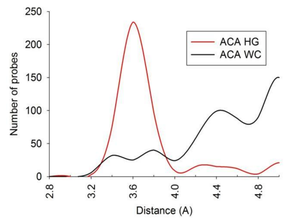


Number of probe atoms in spherical intervals around the amino nitrogen atom.

Substructure clustering on carboxyl groups

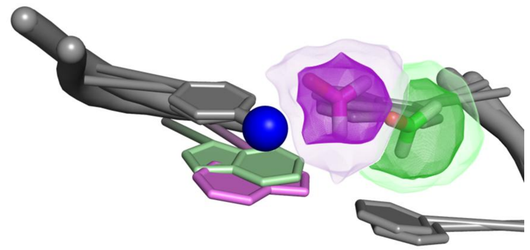


Carbonyl moiety (spheres in red and yellow) closest to the geometric center of the cluster formed by all CO groups in probes acetone, acetaldehyde, acetamide, and urea for DNA with HG pairing (left) and all-WC pairing (right). The triplet shown here is ACA.

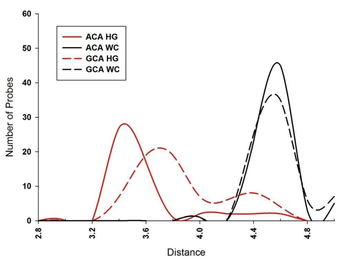


Number of CO groups in spherical intervals around the amino nitrogen atom.

Shift in docked formaldehyde density

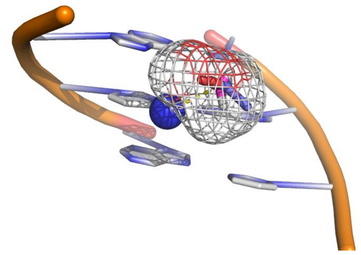


Superimposed formaldehyde (sticks) positioning near the amino nitrogen atom (blue sphere) of cytosine in DNA with Hoogsteen pairing (purple) versus Watson-Crick pairing (green) at the adjacent base pair. The density map shows distribution of low energy formaldehyde positions.



Distributions of formaldehyde positions around the amino nitrogen atom for DNA with HG and WC pairing.

Docking position of acetaldehyde



Lowest energy acetaldehyde pose (in purple stick) bound in a similar manner with that of formaldehyde, overlapping the highest population cluster of CO groups (in mesh) for the GCA triplet with HG pairing.

Please try out our server: **ftmap.bu.edu/beta**

FTMap
computational solvent mapping