J Mol Model (2009) 15:817–827
DOI 10.1007/s00894-008-0434-5

ORIGINAL PAPER

Electrostatic potential maps of damaged DNA studied by image analysis tools. 8-Oxoguanine and abasic site lesions

Rafał A. Bachorz · Giovanni Lupica · Maciej Gutowski · Maciej Haranczyk

Received: 29 September 2008 / Accepted: 5 December 2008 / Published online: 7 January 2009

Abstract Changes of electrostatic potential around the DNA molecule resulting from chemical modifications of nucleotides may play a role in enzymatic recognition of damaged sites. The electrostatic potential around the DNA fragments containing either the intact guanine-cytosine pair or 8-oxoguanine-cytosine or the guanine-abasic site was projected on a cylindrical surface around the double helix. The 2D maps of EP of intact and damaged DNA fragments were compared using image analysis methods. Occurrence of abasic site and 8-oxoguanine lesions were found to be reflected in the EP maps. In the case of the 8-oxoguanine lesion, the two phosphate groups and countercations of the damaged strand are moved away from the lesion in opposite directions, whereas they are moved in the same direction in the case of the abasic site lesion. The characteristic features of 8-oxoguanine lesion might be identified in the major groove, whereas the features of abasic site lesion the minor groove.

Keywords Abasic site · Cylindrical projection · DNA lesions · Electrostatic potential · 8-oxoguanine

Introduction

DNA is continuously exposed to a number of environmental agents acting on living cells, which include UV light, ionizing radiation as well as various chemical species. In many cases, the mechanism of action of these agents involves generation of free radicals that attack DNA and

Electronic supplementary material The online version of this article (doi:10.1007/s00894-008-0434-5) contains supplementary material, which is available to authorized users.

R. A. Bachorz · M. Gutowski · M. Haranczyk
Chemical Sciences Division,
Pacific Northwest National Laboratory,
Richland, WA 99352, USA

G. Lupica
Department of Electronic and Electrical Engineering,
University of Sheffield,
S1 3JD Sheffield, UK

M. Gutowski · M. Haranczyk
Department of Chemistry, University of Gdańsk,
Sobieskiego 18,
80-952 Gdańsk, Poland

M. Gutowski
Chemistry-School of Engineering and Physical Sciences,
Heriot-Watt University,
Edinburgh EH14 4AS, UK

M. Haranczyk
Department of Information Studies,
University of Sheffield,
Sheffield S1 4DP, UK

Present address:
R. A. Bachorz
Lehrstuhl für Theoretische Chemie Karlsruhe Universität (TH),
D-76128 Karlsruhe, Germany

Present address:
M. Haranczyk (✉)
Computational Research Division,
Lawrence Berkeley National Laboratory,
1 Cyclotron Road, MS 50F-1650,
Berkeley, CA 94720, USA
e-mail: mharanczyk@lbl.gov

produce a variety of lesions, including sugar and base modifications, strand breaks, and DNA-protein cross-links [1]. Some reactive oxygen species (ROS), e.g., O_2^-, •OH, are also generated endogenously during cellular aerobic metabolism, and the damage they cause may be an important factor in aging and age-dependent diseases, including cancer [2–5]. In other cases, oxygen radicals may be introduced exogenously to the living cells, e.g., with tobacco smoke. They have been recognized as a major etiological factor for cancers of the upper aerodigestive tract [6].

An example of a ROS related, mutagenic modification of DNA is an oxidative derivative of guanine purine base, 8-oxoguanine (8oG, see Fig. 1). The genomes of aerobic organisms suffer from the presence of 8oG in DNA as the replicative DNA polymerases misread residues with 8oG and place nucleotides with adenine instead of cytosine opposite to the oxidized base [7]. Both bases in the resulting A-8oG mispair are mutagenic lesions, and both must undergo a base-specific replacement to restore the original C-G pair [8].

Another common type of lesions in DNA are abasic sites, denoted as (AP), which correspond to these modifications of DNA, in which nucleic acid bases are removed from the nucleotides. Loss of a base can take place as a spontaneous process [9], but typically it happens as a result of base modification, or it can be observed during an enzymatic repair process [10–12]. If such damage is not repaired, the resulting abasic site has a high potential for mutagenicity and might led to cell death. Due to the biological relevance and importance, there is strong interest in methods of recognizing abasic sites in DNA both for diagnostics and for pharmaceutical use. From the structural point of view, an abasic site represents a discontinuity of the DNA stack, and clearly leads to a deviation from the regular duplex structure [13]. Recent studies showed that within the human cell the spontaneous generation of abasic site can occur 10,000 times per one cell per day [14]. Abasic sites are very base-labile and can also spontaneously further a DNA fragment to form cytotoxic single-strand breaks.

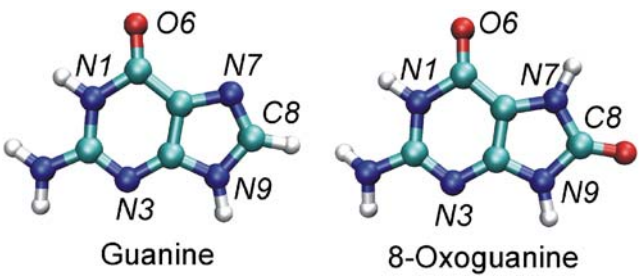

Fig. 1 Labeling of atoms in guanine and 8-oxoguanine

The DNA lesions occur as an inseparable part of life processes. However, to sustain the "state of living", the cell's repair machinery has to constantly detect and correct the damaged DNA. The damage detection and repair is a complicated process in practice. The idea of damage detection is, however, fairly simple: the repair enzymes have to locate the damaged sites based on features that make those sites different from intact fragments of DNA. The DNA-enzyme interaction is determined by hydrogen bonds, electrostatic interactions and dispersion forces. The two first have similar origins and result from the interaction of effective atomic charges of molecule A with the polarizable charge distribution of molecule B. The dispersion interaction, on the other hand, can be associated with the interaction of instantaneous multipoles of A and B. In our recent studies we analyzed electrostatic potential around an intact DNA fragment and around the same fragment but with guanine being replaced with 8-oxoguanine (8oG, see Fig. 1) [15, 16], adenine being replaced with 8-oxoadenine (8oA) [17], and thymine being replaced with thymine glycol (Tg) [18]. Our goal was to develop and employ a proper approach to identify differences in the electrostatic potential that might be relevant for enzymatic recognition of alternated nucleic acid bases.

The electrostatic potential (EP) around DNA is nonuniform because the molecule possesses many polar and nonpolar groups specifically arranged in space. The analysis of electrostatic potential in three dimensions might be a very complex task. Assuming that the most important interactions take place near the surface of DNA, we proposed to focus our analysis on the values of the electrostatic potential at the "molecular surface" [16, 17]. To facilitate such analysis we developed and implemented a cylindrical projection of EP, in which the values of the electrostatic potential at a complicated molecular surface of DNA are projected onto the walls of a cylinder, which is built around an approximate DNA axis. The resulting two dimensional (2D) electrostatic maps were then analyzed for the intact fragments of DNA (consisting typically of three (5′-TGT-3′ [15], 5′-AAC-3′ [17]) or five (5′-ACTAG-3′ [18]) nucleic acid base pairs) and the corresponding fragments containing alterations (8oG in the place of guanine, 8oA in the place of adenine, Tg in the place of thymine). The main new features in the EP map resulting from the presence of the lesions were reorganizations of countercations and phosphate groups [16, 17].

In the current study, we conduct the analysis of the electrostatic potential around the 5′-AGT-3′ trimer and around the analogous systems where the central nucleic base pair was alternated. We have investigated two damaged trimers: (i) the guanine is replaced by 8-oxoguanine, and (ii) the cytosine is removed to create an abasic site lesion. The resulting systems are denoted as 5′-A

(8oG)T-3′ and 5′-A(AP)T-3′, respectively. This study represents an advancement over previous studies as now we include abasic sites to the portfolio of studied lesions and we compare the relative structural and EP changes caused by different lesions. We demonstrate that the presence of these lesions leads to the reorganization of the DNA fragment which is especially pronounced in the case of the abasic site lesion. The reorganization of the DNA includes shifts of the C and G pairs as well as movements of the countercations and sugar-phosphate backbones. In addition, and in contrast to the initial studies [16, 17], we improved the analysis of EP maps by employing advanced image analysis methods [18, 19]. Here we demonstrate how important features on EP maps can be automatically detected and measured. Our results shows that image analysis techniques might be very useful in chemical applications, where they can automate some part of the results analysis process.

Methods

Molecular structure and introduction of DNA lesions To obtain the structure of intact and damaged DNA fragments, we used the same approach as described before [15–17, 20]. From a starting molecule, the 5′-GGGAACAACTAG-3′ DNA dodecamer described by Arnott et al. [21], we cut a 5′-AGT-3′ fragment (a trimer containing three nucleic base pairs). To limit the size of the system, the phosphate groups were removed from the 5′ and 3′ ends, and the strands were saturated with OH groups. The sodium countercations were initially placed at each of four phosphate groups at the position found to be optimal in our calculations of the gas-phase NaH_2PO_4 obtained in our previous study [18].

The resulting structure was a starting point for density functional theory (DFT) optimizations using a B3LYP exchange-correlation functional [22–24]. In order to suppress computational artifacts resulting from: (i) using a trimer as a model of the DNA double strand, and (ii) a failure of B3LYP to reproduce intermolecular dispersion interactions [25], we introduced additional constraints on the trimer. The atoms of the top and bottom nucleic acid base pairs were fixed in space during the optimization [15–17, 20]. This approach guarantees that the B3LYP failure to reproduce the π-π stacking interaction does not lead to unphysical geometries of the trimer.

The 6–31G(d,p) basis set was used for all geometrically optimized atoms and the 6–31G basis set [26, 27] for all geometrically fixed atoms resulting in 1925 basis functions for the 193 atom system. In the next step, we introduced lesions. The guanine of the intact 5′-AGT-3′ trimer was replaced with 8oG by substituting hydrogen with oxygen at the C8 position and adding the hydrogen atom to the N7 atom (see Fig. 1). The (AP) site lesion was created by removing the cytosine and saturating the dangling bond with a hydrogen atom. The geometries of the resulting damaged structures were optimized in the analogous way to the intact form.

All DFT geometry optimizations were preformed using the NWChem 4.5–4.7 program package [19] on a cluster of dual Intel Itanium2 nodes with Quadrics ELAN4 interconnect. Structural optimizations required approximately 200 geometrical steps with the convergence criteria being set to default values (the maximum and root mean square gradient in the coordinates being set to 0.00045 and 0.00030 a.u., respectively). The NWChem calculations used 128 Gb of memory and up to 1Tb of disk space. We also tried to use 64 and 256 processors and found out that the calculation time scales nearly linearly with the number of processors involved. The detailed benchmarking of different tasks performed with the NWChem code on various machines is presented on the NWChem website [28]. The Extensible Computational Chemistry Environment (ECCE) [29, 30], which provides a sophisticated graphical user interface and scientific visualization tools, was particularly useful to set up calculations for systems with "frozen" atoms and mixed-basis sets. The optimized geometries of all studied systems are provided as Supporting Information.

Regular cylindrical projection of electrostatic potential Analysis of electrostatic potential for a molecule is a complicated process. One needs to look at the potential created by charges in the 3D space around the molecule. A simplification is provided by a technique that is called a regular cylindrical projection, which is similar to gnomonic projection methods [31–35]. The details of this technique and our implementation are provided in [18]. Here we only briefly summarize how the EP maps are obtained.

First, a cylinder is built around the DNA axis. The axis is defined by a point and a vector. For consistency of the EP analysis, they both are kept unchanged when analyzing both the intact and damaged DNA fragments. The point is defined by the geometrical center of the intact DNA fragment. The axis is derived by diagonalization of the tensor of inertia of the starting DNA dodecamer. The eigenvector corresponding to the largest eigenvalue defines the axis of the DNA molecule.

Next, we uniformly distribute points on the side wall of the cylinder. It is done in two loops:

(i.) over the height of the cylinder along the z axis (Fig. 2), and
(ii.) over the angle α in the xy plane (Fig. 2).

For each step in height (along the z axis), the **R** vector (a radius of the cylinder is set to 20 Å) is rotated around the z axis and the resulting points on the cylinder are recorded

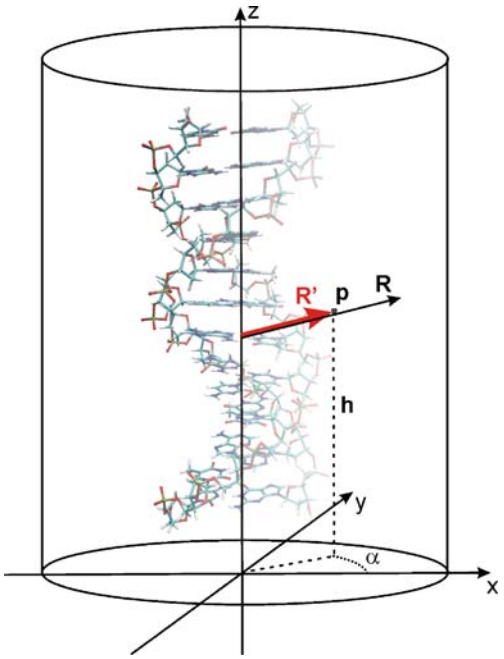

Fig. 2 Details of cylindrical regular projection

(Fig. 3a). The angular coordinate was scanned from 0 to 359.5 degrees with a step of 0.5 degree. The z coordinate was scanned with 101 steps of 0.1 Å each. The starting point of the z coordinate scan was set to 5.0 Å below the geometrical center of the intact trimer. The resulting cylinder consists of 72,720 points, which coordinates are provided as Supporting Information.

Each selected point on the cylinder defines a point on the DNA surface (Fig. 3b). To identify the point on the DNA surface one walks along **R** toward the cylinder axis until one encounters the DNA surface. This procedure defines a DNA surface point **p** marked by **R′** in Fig. 2. The DNA surface is spanned by spheres centered on atomic nuclei with excluded cavities, which are not accessible to solvent molecules (simulated by a probe of a 1.3 Å radius). The spheres' radii for the C, N, O, H, P atoms [36] are based on van der Waals radii used in the MM3 force field [37]. The radius for Na is based on its ionic radius [38]. Finally, at each point on the molecular surface, we calculated the EP value directly from the B3LYP/6-31G(d,p) electron density. These calculations were done with the Gaussian03 program [39] on an SGI Altix machine with 128 Gb of shared-memory.

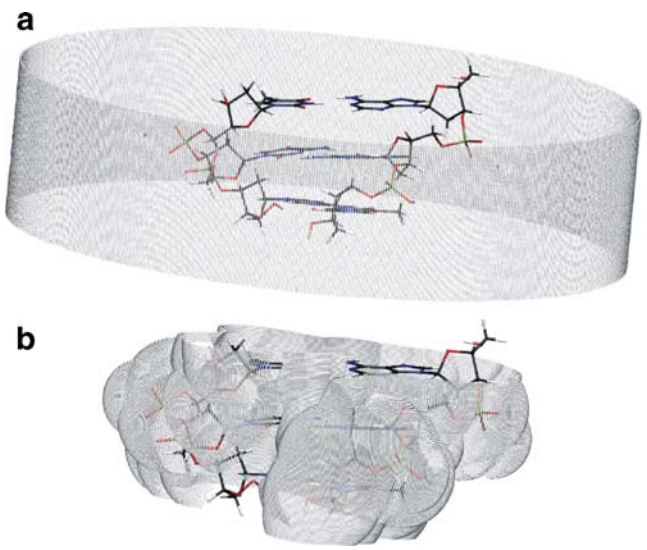

Fig. 3 Points on the projection cylinder (**a**) and the molecular surface (**b**)

The final 2D electrostatic potential maps are presented in Fig. 4. The angle α and the z coordinate of each point are displayed on the horizontal and vertical axis, respectively. The color is assigned according to the value of the electrostatic potential. Red and blue points have positive and negative values of the EP, respectively. EP maps are stored in standard bitmap files (bmp format). A color of each pixel of a bitmap is a superposition of three colors and it is stored in 24-bits of data: 8-bits (256 possible values) to code each of the red, green and blue channels (RGB). However, in the presented results only two channels are used to store information (one channel for positive values of EP, the other for negative values of EP), therefore the values of EP changing from –0.53 to 0.53 au are represented by 511 colors.

The maps generated with (720×101) points are scaled too a higher resolution (720×801) using a bi-directional cubic spline interpolation, as implemented in the FreeImage library [40]. One should have in mind that such scaling distorts the original ratio between height and length of the map and it does not reflect the true ratio between the height of the DNA fragment and its circumference. The origin of important features of the resulting EP maps was interpreted based on geometrical structures of the molecules and their accompanying molecular surfaces (Fig. 3b), as well as on topographical maps described later in this text. The differences between two maps can be highlighted in difference maps (Fig. S-1) obtained by subtracting the values of corresponding points of two maps. A quantitative analysis of the EP maps was done using the image analysis techniques described in the following section.

Image analysis methods employed in analysis of EP maps We have created specific computer vision tools to quantitatively measure differences in the EP maps presented in Fig. 4a, c, d. These tools provide the following general facilities:

1. Segmentation (detection) of red features corresponding to counteractions in the EP map images (Fig. 4b);

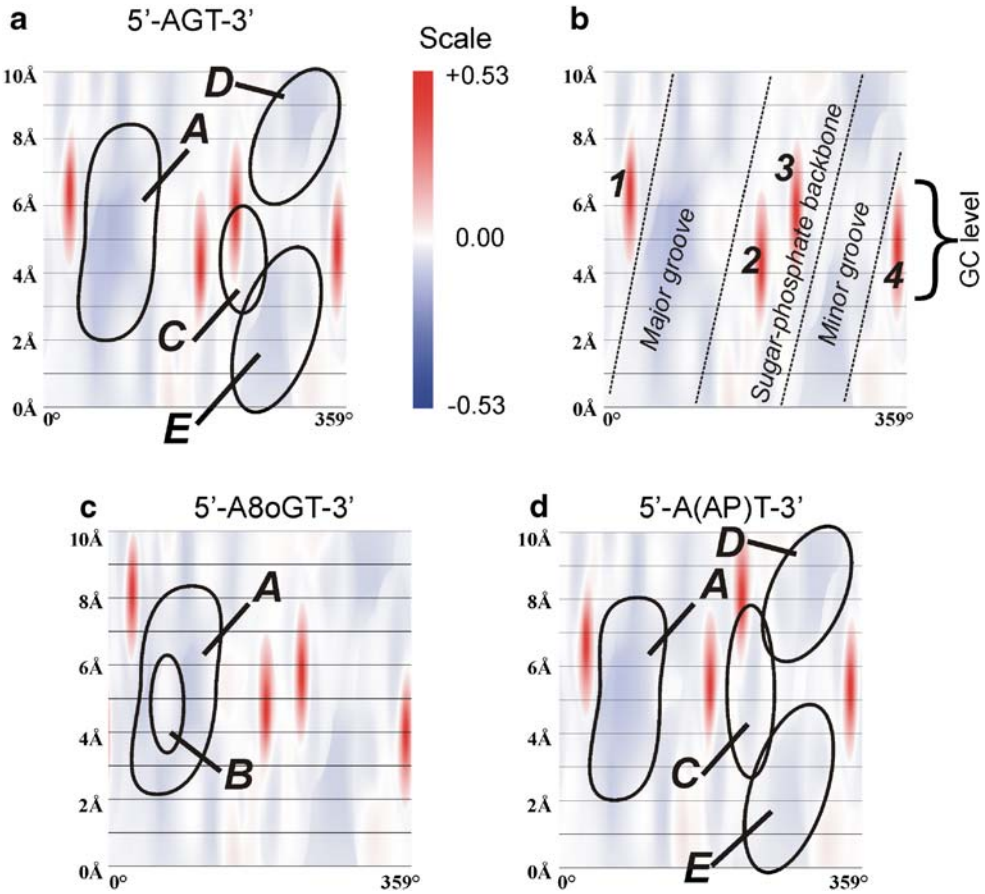

Fig. 4 Electrostatic potential maps of 5′-AGT-3′ (**a**), A(8oG)T-3′ (**c**) and 5′-A(AP)T-3′ (**d**) DNA trimers. The angular coordinate α is placed on the horizontal axis and the height z is on the vertical axis with zero set to the upmost AT pair. The dominant structural features of DNA fragments are marked in the b map, with the sodium countercations labeled ***1***–***4***. Important features of the considered DNA fragments are encircled on the maps a, c and d and labeled ***A***–***E***. One should keep in mind that the 0° to 359.5° range on the horizontal axis corresponds to the circumference of the DNA double helix, which is approximately 60 Å. Thus the maps are "compressed" in the horizontal dimension

2. Measurement of the displacement of red features (counteractions) among different EP maps;
3. Segmentation (detection) of fuzzy features, A–E in Fig. 4a, c, d, in the EP map images;
4. Description of the features previously segmented in terms of size and intensity.

The segmentation of the red features (counteractions) has been achieved by applying a threshold based segmentation technique (P-tile method [41]) to the blue channel of the EP map images. Once segmented, the countercations are automatically labeled using a morphological operator [42] (***1***–***4*** in Fig. S-2b), and their positions are computed as the center of gravity (centroids) of the detected object. The displacements of countercations having the same labels but belonging to different EP maps are finally computed as the distances between their centroids computed along the axes α and z. The results are presented in Table 1.

The segmentation technique applied to detect the fuzzy features labeled as ***A***–***E*** in Fig. 4a, c, d, exploits the *extended h-minima transform* [42, 43], a morphological filter that adopts a contrast criterion in the selection of minima from an image. Once the fuzzy features ***A***–***E*** are segmented, see Figs. S-2 and S-3, the EP values are calculated by

Table 1 Positions of Na countercations in the reference EP map of the intact DNA (5′-AGT-3′) and corresponding displacement vectors for the damaged DNA containing 8oG and (AP) lesions obtained using image analysis algorithms. Positions and vectors in "(x, y)" format with x coordinate in Å and y coordinate in °. Displacements of the Na^+ counteractions is 3D and are in Å

Countercation	Initial position in the intact 5′-AGT-3′ trimer	Displacement vectors		Displacement in 3D	
		A(8oG)T-3′	5′-A(AP)T-3′	A(8oG)T-3′	5′-A(AP)T-3′
1	(6.16, 32.0)	(1.90, −3.0)	(0.29, 1.0)	1.95	0.30
2	(4.44, 187.5)	(0.50, 0.5)	(1.24, −8.0)	0.56	1.91
3	(5.84, 229.0)	(−0.09, 1.0)	(2.43, −11.5)	0.26	3.22
4	(4.64, 350.0)	(−0.89, 3.5)	(0.65, −3.5)	1.23	0.90

Table 2 The mean electrostatic potential and the size of the important ***A*** and ***B*** features identified in the major groove of intact and damaged DNA trimers

Feature	Intensity of EP (in a.u.)		Shape			
Object ***A*** in the major groove	Mean EP Value	Standard Deviation	Width (in °)		Length (in Å)	Area (in ° * Å)
			At 3.5 Å	At 7 Å		
5′-AGT-3′	−0.106	0.018	12.5°	12°	3.73	146.76
5′-A(8oG)T-3′	−0.066	0.014	44. 5°	40°	6.84	276.36
5′-A(AP)T-3′	−0.087	0.016	30.5°	27°	4.37	202.63
Object ***B*** in the major groove	Mean EP Value	Standard Deviation	Width (in °) At 5 Å		Length (in Å)	Area (in ° * Å)
5′-A(8oG)T-3′	−0.018	0.007	20.5°		2.53	41.18

measuring the mean intensity value, the standard deviation and the shape of each object detected. The results are presented in Tables 2 and 3.

The image analysis was preformed using the MATLAB (version R2006b) [44] and functions available in the image processing toolbox.

Topographical maps of the DNA surface Similarly to the EP maps, we have generated topographical maps of the DNA surface (Fig. 5). The angle α and the z coordinate of each point are displayed on the horizontal and vertical axis, respectively. For each point on the surface we found the closest atom. The color is assigned according to the value of the corresponding atomic number. The maps were generated with 720x801 points, similarly as in the scaled EP maps.

Results

First, the geometrical differences between the intact and damaged DNA fragments were analyzed. Since both fragments had the same top and bottom nucleic acid base pairs fixed in space during the geometry optimization, it is possible to superimpose these pairs and compare the differences in the remaining molecular parts (Fig. 6). To clarify the picture, the top and bottom base pairs of trimers were removed from Fig. 6. The structural differences between the DNA fragments containing G and 8oG are mostly limited to the strand with the lesion. The 8oG is slightly displaced in the direction of the major groove in comparison with the position of G in the intact strand. Some reorganization of the phosphate group with sodium countercations next to the lesion might also be observed.

Table 3 The mean electrostatic potential and the size of the important ***C***–***E*** features identified in the minor groove of intact and damaged DNA trimers

Feature	Intensity of EP (in a.u.)		Shape			
Object ***C*** in the minor groove	Mean EP Value	Standard Deviation	Width (in °)		Length (in Å)	Area (in ° * Å)
			At 5 Å	At 6 Å		
5′-AGT-3′	−0.037	0.003	8°	0°	3.55	33.66
5′-A(8oG)T-3′	−0.037	0.003	8°	0°	3.62	33.73
5′-A(AP)T-3′	−0.045	0.007	18.5°	15°	3.55	47.81
Object ***D*** in the minor groove	Mean EP Value	Standard Deviation	Width (in °)		Length (in Å)	Area (in ° * Å)
			At 7.5 Å	At 9 Å		
5′-AGT-3′	−0.072	0.010	12.5°	47.5°	2.63	95.02
5′-A(8oG)T-3′	−0.070	0.010	19.5°	49.5°	2.79	103.06
5′-A(AP)T-3′	−0.060	0.011	22.5°	45.5°	3.65	111.47
Object ***E*** in the minor groove	Mean EP Value	Standard Deviation	Width (in °)		Length (in Å)	Area (in ° * Å)
			At 1.5 Å	At 3 Å		
5′-AGT-3′	−0.064	0.011	49.5°	30.5°	4.68	176.58
5′-A(8oG)T-3′	−0.062	0.011	48.5°	31.0°	4.79	178.39
5′-A(AP)T-3′	−0.047	0.008	42.5°	36.0°	4.52	157.70

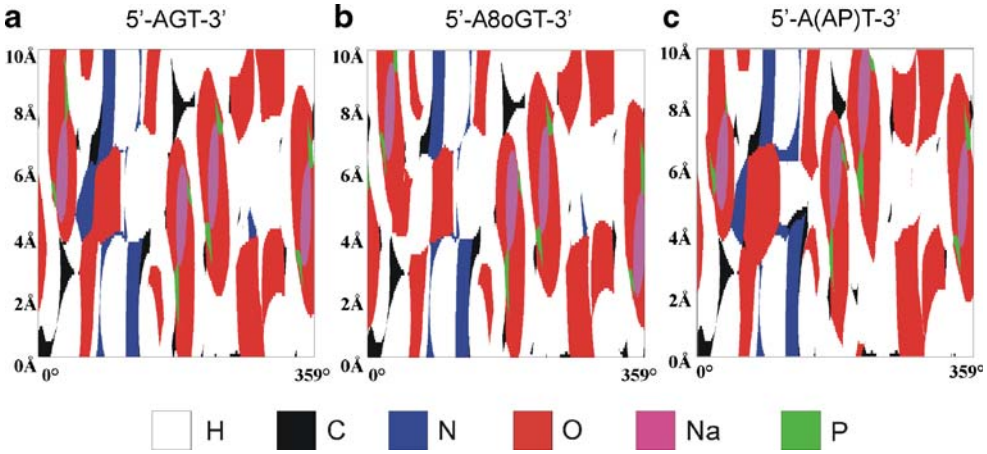

Fig. 5 Topographical maps of the intact (**a**) and damaged (**b**–**c**) DNA surfaces

Those atoms are moved along the long axis of DNA because of the repulsive interaction between the negatively polarized O8 in 8oG and a negatively charged phosphate group. Similar geometry changes were observed when G was substituted with 8oG in the 5′-TGT-3′ trimer studied before [16], suggesting that such geometrical features might be typical to the 8oG lesion and not only limited to the selected DNA trimer. The extent of structural changes

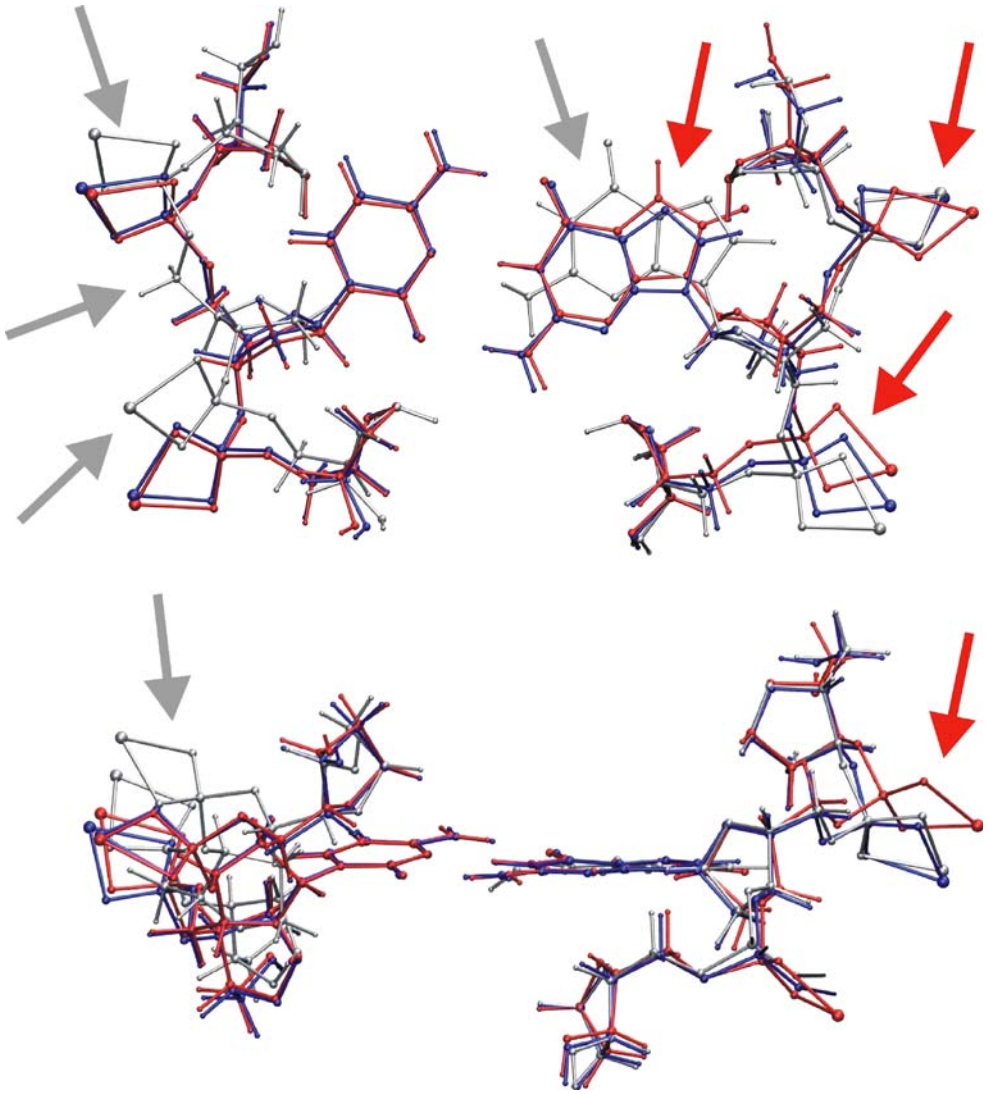

Fig. 6 Top and side views of the superimposed 5′-AGT-3′ (blue), A(8oG)T-3′ (red) and 5′-A(AP)T-3′ (gray) DNA trimers demonstrating differences in their geometries. The largest changes for 8-oxoguanine and abasic site lesions are marked with red and gray arrows, respectively

caused by introduction of 8oG lesion reported here is in good agreement with the crystallographic study by Lipscomb et al. [45]. The reported experimental root mean squared deviation (RMSD) between the coordinates of intact and damaged DNA decamer is 0.27 Å, whereas it is 0.40 Å for DNA trimers in our study. The fact that the RMSD value obtained for our DNA trimers is larger corresponds with the fact, the geometry alternations are fairy local. Including contributions from atoms located further from the lesion in the case of DNA decamer lowers the RMSD value. This fact is also reflected in description of results – the authors of the experimental study describe the structural perturbations induced by 8oG as "subtle", whereas we provide a more detail description, important from the perspective of shorter DNA trimer.

The geometrical changes in the trimer containing abasic site lesion have been studied for the first time and are more severe than for 8oG [16] and 8oA [17]. The structural differences between the intact and the 5′-A(AP)T-3′ fragments are seen for both stands, however, the strand with the (AP) lesion seems to be displaced to a much greater extent. The counteractions are dragged by the phosphate groups as the whole backbone moves along the z axis. The guanine base on the strand opposite the (AP) lesion is not hydrogen-bonded with the missing cytosine and becomes displaced toward the major groove.

The geometrical features characterizing the 8oG and (AP) lesions can be compared. For both lesions significant reorganizations of the phosphate groups with sodium countercations are observed (Fig. 6) and the alternated central nucleic acid pairs are moved toward the major groove. However, the above structure changes are much more evident for the (AP) lesion.

The cylindrical projections of EP around the 5′-AGT-3′, A(8oG)T-3′ and 5′-A(AP)T-3′ DNA fragments show multiple features (Fig. 4a, c, d). The differences between the EP maps are analyzed in the context of chemical differences between the intact and damaged DNA fragments, which are hidden under the DNA surfaces presented in Fig. 3b. The most important structural features of a DNA fragment are marked in Fig. 4b. The most distinctive features are the sodium countercations that are represented by four big red spots of a positive electrostatic potential and labeled ***1***–***4*** in Fig. 4b. The positions of countercations in the reference map for 5′-AGT-3′ and the displacement vectors for DNA fragments with lesions are summarized in Table 1. The sugar-phosphate backbones are represented by two elongated shapes of a neutral potential (Fig. 4b). The minor and major grooves might be seen between the two strands (Fig. 4b). The grooves are characterized by a negative potential. A contribution from the intact G in 5′-AGT-3′, labeled ***A*** in Fig. 4a, can be found in an encircled area. This contribution is represented as a spot of negative potential, which is ca. 3.73 Å long and up to 12.5° wide (Table 2). It is associated with a nitrogen atom and an oxygen atom, as shown in Fig. 5, and the identity of these atoms are N7 and O6 of guanine, respectively (Fig. 1).

At the first glance, the EP maps of the 5′-A8oGT-3′ and 5′-A(AP)T-3′ DNA fragments (Fig. 4c-d) look very much like the EP map of the intact fragment (Fig. 4a). The sodium countercations, strands, and grooves might be identified in both cases. The intensity of colors, reflecting the values of EP, is similar except in the middle part of the plot. The visual inspection might suggest that the lesions did not much affect the DNA structure. However, the difference maps (Fig. S-1) expose the regions on the EP maps affected by the lesions. The characteristic changes are measured in detail using the image analysis tools and the results are summarized in Tables 1, 2, 3.

The most characteristic feature of the 8-oxoguanine lesion can be identified in the major groove. A spot of the negative potential, which is encircled and denoted ***A*** in Fig. 4a, c is now interrupted by a positive feature ***B*** (Fig. 4c). This feature corresponds to a hydrogen atom (see a topography map, Fig. 5) and can be associated with H connected to N7 in 8oG (Fig. 1). Interestingly, although the feature ***B*** disrupts feature ***A***, the latter is still much larger (ca. 60%) in terms of the area of the major groove it occupies. The larger area of ***A*** results from the presence of an additional oxygen atom (O8) in 8-oxoguanine. The feature ***A*** is however less negative (Table 2) than in the intact fragment, which is the effect of a hydrogen atom giving the origin to ***B***.

The incorporation of 8oG into the 5′-AGT-3′ trimer does not affect the electrostatic potential in the minor groove but it does induce a geometrical relaxation of the phosphate backbone, which is followed by a displacement of sodium counteractions. These differences are clearly seen in Fig. S-1. The relocations along the DNA axis are the most profound for two countercations of the damaged DNA strand, labeled ***4***. They displace away from the lesion by as much as 1.90 Å and 0.89 Å, respectively, with both moving also along the angular coordinate by −3.0° and 3.5°, respectively (Table 1). These displacements on the EP maps correspond to displacements in 3D by 1.95 and 1.23 Å, respectively. The geometrical relaxation of the countercation ***2*** is also noticeable as it moves up to 0.5 Å along the DNA axis.

Next, we turn attention to the 5′-A(AP)T–3′ fragment. A difference map (Fig. S-1) corresponding to the intact and 5′-A(AP)T-3′ suggests that the abasic site lesion has influenced almost all regions of the map. The most profound is the reorganization of phosphate groups and the related counteractions of the DNA strand containing the lesion. The countercations ***2*** and ***3*** were displaced by, respectively, 1.24 and 2.43 Å along the DNA axis and by −8.0° and

−11.5°, respectively, along the angular coordinate. These displacements on the EP map correspond to displacements in 3D of 1.90 and 3.22 Å, respectively. The geometrical relaxation of the countercation ***4*** is also noticeable as it moves up by 0.65 Å along the DNA axis and −3.5° along the angular coordinate. The difference map (Fig. S-1) also suggests extensive but not profound changes of features in both the major and minor groove regions. These features, marked in encircled areas ***A***, ***C*** - ***E*** in Fig. 4a, d, have been measured and a summary is presented in Tables 2 and 3. The feature ***A,*** corresponding to a negative potential of guanine's O6 and N7 atoms, has ca. 40% larger area (its length is 0.7 Å larger and its width ca. 20° wider) for a trimer with the lesion. At the same time the mean value of EP of this feature in 5′-A(AP)T-3′ is ca. 20% less negative than in the intact trimer.

The features ***C***–***E*** located in the minor groove are also affected by the introduction of the abasic site lesion. The feature ***C*** becomes more negative and its area expands by ca. 40% for the damaged fragment. The feature ***D*** also expands its area as its length and width grow by, respectively, ca. 1 Å and 10° (at 7.5 Å). At the same time ***D*** becomes less negative. In turn, a negative feature ***E*** slightly decreases its area (ca. 10%) and it becomes less negative.

The differences between the two lesions considered here, namely 8-oxoguanine and abasic site, can be observed in Fig S-1 and quantitatively described by the image analysis results of Tables 1, 2, 3. The most noteworthy difference between these lesions is the pattern of the geometrical relaxation of countercations and the accompanying phosphate groups. In both cases the counteractions corresponding to the strand with the lesion are the most affected. However, the direction of the displacements depends on the nature of the lesion. In the case of the 5′-A8oGT-3′ DNA trimer, the countercations move away from the damaged site with each of them moving in the opposite direction (the countercation ***1*** moves toward larger z values while the countercation ***4*** toward lower z values). In the case of the 5′-A(AP)T-3′ fragment, both countercations, namely ***2*** and ***3***, move toward larger values of z. The mean displacement in 3D is also larger in the case of the abasic site lesion.

As far as the features in the major groove are concerned, both damaged DNA fragments are characterized by larger areas of the negative spots (features ***A*** in Fig. 4c, d) than the intact DNA trimer. However, the 5′-A8oGT-3′ fragment has a characteristic feature, marked ***B*** in Fig. 4c, which is related to a hydrogen atom at guanine's N7. This new feature appears in the center of feature ***A*** (Fig. 4c). The 5′-A8oGT-3′ and 5′-A(AP)T-3′ DNA trimers are also different in the region of the minor groove. The 5′-A8oGT-3′ practically reassembles the intact DNA, whereas the 5′-A (AP)T-3′ trimer has larger areas of features ***C*** and ***D***, and a smaller area of feature ***E***.

Summary

We have studied a fragment of intact DNA, 5′-AGT-3′, and its counterparts containing 8-oxoguanine and abasic site lesions at the central nucleic acid base pairs (5′-A8oGT-3′ and 5′-A(AP)T-3′, respectively) using the B3LYP/6-31G(d,p) method. The geometries of 5′-AGT-3′, 5′-A8oGT-3′ and 5′-A(AP)T-3′ were optimized with a constraint of the fixed top and bottom base pairs. The electrostatic potential around these DNA fragments was studied using the regular cylindrical projection technique. The values of electrostatic potential at the molecular surfaces were projected onto the side walls of cylinders surrounding the DNA fragments. The resulting maps of electrostatic potential and optimized geometries of the intact and damaged DNA fragments were compared using image analysis techniques. The results can be summarized as follows:

1. The introduction of either 8-oxoguanine or abasic site lesion distorts primarily the damaged strand of DNA. The complementary strand is less distorted.
2. In the case of 8-oxoguanine, the damaged nucleic acid base is slightly displaced in the direction of the major groove in comparison with the position of the corresponding undamaged base in the intact DNA fragment.
3. A spot of negative electrostatic potential in the major groove changes its character when guanine is replaced by 8oG. Although the area of the spot becomes larger, at the same time it is disrupted by a positive electrostatic potential, which results from a hydrogen atom connected to N7 of 8oG. The EP in the minor groove is not affected by the presence of 8oG.
4. A reorganization of countercations and the neighboring phosphate groups along the axis of the DNA fragment is clearly reflected in the electrostatic potential maps. The countercations of the damaged strand move away from 8-oxoguanine by 1.9 and 0.9 Å.
5. In the case of the abasic site lesion, the guanine opposite to the removed cytosine is displaced in the direction of the major groove in comparison with its position in the corresponding intact DNA fragment.
6. The occurrence of the abasic site lesion is reflected in the EP map in the areas of both the major and minor groove. A spot of negative potential in the major groove increases by ca. 40%. The EP in the minor groove is less affected.
7. A reorganization of counteractions and the neighboring phosphate groups is also characteristic for the abasic

site lesion. The countercations of the damaged strand move by 2.4 and 1.2 Å along the axis of the DNA fragment.

8. The main differences between the 8-oxoguanine and abasic site lesions are: (a) the presence of the abasic site lesion has much larger consequences for the structure of DNA than the presence of 8-oxoguanine. In the latter cases the changes were mostly limited to the damaged site; (b) two counteractions of the damaged strand are displaced in a different way along the DNA axis: they move in opposite directions in 5′-A8oGT-3′ and in the same direction in 5′-A(AP)T-3′); (c) the patterns of EP in the major groove is different as 5′-A8oGT-3′ has a distinctive positive spot of EP; (d) the minor groove is affected in the 5′-A(AP)T-3′ trimer but not in the 5′-A8oGT-3′ fragment.
9. The identified differences in the electrostatic potential of the intact and damaged DNA fragments provide hints on how enzymatic recognition of alternated nucleic acid bases may take place. As in the cell environment, in the presence of water, the counteractions are not tightly connected with phosphaste groups, the differences in EP in the major groove become the dominant features distinguishing damaged fragments from intact ones. Therefore, repair enzymes are likely to exploit these features by scanning the major groove for alternated EP features pattern. The act of recognition can usually be explained in terms of reorganization energy [46]. The enzyme is preoganized (the organization means a specific orientation of charged and polar groups in the folder state of protein) to facile operation on the intact DNA. However, when damaged DNA fragment is encountered, different EP of DNA will force structural changes which correspond to damage recognition.

We demonstrated how the electrostatic potential maps obtained by the cylindrical projection technique can be automatically analyzed with appropriate image tools. We believe that this approach can be used to build (semi-) automatic tools to study large number of structures. For example the image analysis/motion detection algorithm can be used to track and report the changes of EP as a function of simulation time in molecular dynamics simulations or detect characteristic features of particular compounds in database screening experiments.

Acknowledgements Helpful discussions with John H. Miller and Michel Dupuis are gratefully acknowledged. M.H. acknowledges the support from Chemical Structure Association Trust (CSAT) that covered the cost of visit at the University of Sheffield through the Jacques-Emile Dubois Grant. This work was supported by the: (i) US DOE Office of Biological and Environmental Research, Low Dose Radiation Research Program, (ii) Polish State Committee for Scientific Research (KBN) Grant DS/8221-4-0140-8 (M.G. and M.H.). The calculations were performed at the Academic Computer Center in Gdańsk (TASK) and at the Molecular Science Computing Facility (MSCF) in the William R. Wiley Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research and located at the Pacific Northwest National Laboratory, which is operated by Battelle for the US Department of Energy. The MSCF resources were available through a grand challenge project.

References

1. Dizdaroglu M, Jaruga P, Birincioglu M, Rodriguez (2002) Free Radical Biol and Med 32:1102–1115
2. Lutz WK (1990) Mutation Res 238:287–295
3. Ames BN (1989) Mutation Res 214:41–46
4. Ames BN (1989) Environ Mol Mutagen 14:66–77
5. Ames BN, Gold LS (1991) Mutation Res 250:3–16
6. Jaloszynski P, Jaruga P, Olinski R, Biczysko W, Szyfter W, Nagy E, Moller L, Szyfter K (2003) Free Radical Res 37:231–240
7. Lindhal T (1993) Nature 362:709–714
8. Fromme JC, Banerjee A, Huang SJ, Verdine GL (2004) Nature 427:652–656
9. Lindhal T, Nyberg B (1972) Biochemistry 11:3610–3618
10. Demple B, Harrison L (1994) Annu Rev Biochem 63:915–948
11. Scharer OD (2003) Angew Chem 115:3052–3082
12. Scharer OD (2003) Angew Chem Int Ed 42:2946–2974
13. Lhomme J, Constant JF, Demeunynck M (1999) Biopolymers 52:65–83
14. Nakamura J, Walker VE, Upton PB, Chiang YW, Swenberg JA (1998) Cancer Res 58:222
15. Haranczyk M, Bachorz RA, Dabkowska I, Dupuis M, Miller JH, Gutowski MS, Computational Characterization of Lesions in DNA, 228th ACS National Meeting, U175–U175, 025–BIOL Part 1, Philadelphia, PA, August 22–26, 2004.
16. Haranczyk M, Gutowski M (2007) Theo Chem Acc 117:291–296
17. Haranczyk M, Miller JH, Gutowski M (2007) J Mol Graph Model 26:282–289
18. Haranczyk M, Lupica G, Dąbkowska I, Gutowski M (2008) J Phys Chem B 112:2198–2206
19. Gonzales RC, Woods EW (1992) Digital Image Processing. Addison–Wesley, Boston
20. Miller JH, Aceves-Gaona A, Ernst MB, Haranczyk M, Gutowski M, Vorpagel ER, Dupuis M (2005) Rad Res 164:582–585
21. Arnott S, Hukins DWL (1972) Biochem Biophys Res Commun 47:1504–1509
22. Becke AD (1988) Phys Rev A 38:3098–3100
23. Becke AD (1993) J Chem Phys 98:5648–5652
24. Lee C, Yang W, Paar RG (1988) Phys Rev B 37:785–789
25. Kohn W, Meir Y, Makarov DE (1998) Phys Rev Lett 80:4153–4156
26. Straatsma TP, Apra E, Windus TL, Bylaska EJ, de Jong W, Hirata S, Valiev M, Hackler M, Pollack L, Harrison R, Dupuis M, Smith DMA, Nieplocha J, Tipparaju V, Krishnan M, Auer AA, Brown E, Cisneros G, Fann G, Früchtl H, Garza J, Hirao K, Kendall R, Nichols J, Tsemekhman K, Wolinski K, Anchell J, Bernholdt D, Borowski P, Clark T, Clerc D, Dachsel H, Deegan M, Dyall K, Elwood D, Glendening E, Gutowski M, Hess A, Jaffe J, Johnson B, Ju J, Kobayashi R, Kutteh R, Lin Z, Littlefield R, Long X, Meng B, Nakajima T, Niu S, Rosing M, Sandrone G, Stave M, Taylor H, Thomas G, van Lenthe J, Wong A, Zhang Z *NWChem, A Computational Chemistry Package for Parallel Computers, Version 4*, 2004, Pacific Northwest National Laboratory, Richland, WA 99352–0999, USA

27. Kendall RA, Aprà E, Bernholdt DE, Bylaska EJ, Dupuis M, Fann GI, Harrison RJ, Ju J, Nichols JA, Nieplocha J, Straatsma TP, Windus TL, Wong AT (2000) Computer Phys Comm 128:260–283
28. NWChem, http://www.emsl.pnl.gov/docs/nwchem/benchmarks/index.html
29. Black G, Didier B, Bisethagen T, Feller D, Gracio D, Hackler M, Havre S, Jones D, Jurrus E, Keller T, Lansing C, Matsumoto S, Palmer B, Peterson M, Schuchardt K, Stephan E, Sun L, Taylor H, Thomas G, Vorpagel E, Windus T, Winters C, Ecce (2006) A Problem Solving Environment for Computational Chemistry, Software Version 3.2.5. Pacific Northwest National Laboratory, Richland, WA 99352–0999, USA
30. Black GD, Schuchardt KL, Gracio DK, Palmer B, The Extensible Computational Chemistry Environment: A Problem Solving Environment for High Performance Theoretical Chemistry, Computational Science - ICCS 2003, International Conference Saint Petersburg, Russian Federation, Melbourne, Australia, Proceedings 2660, vol. 81, ed. Sloot PMA, Abramson D, Bogdanov AV, Dongarra J, Springer, Berlin, 2003, pp 122–131
31. Van Geerestein VJ, Perry NC, Grootenhuis PG, Haasnoot CAG (1990) Tetrahedron Computer Methodology 3:595–613
32. Perry NC, van Geerestein VJ (1992) J Chem Inf Comp Sci 32:607–616
33. Chau P-L, Dean PM (1987) J Mol Graph 5:97–100
34. Blaney FE, Edge C, Phippen RW (1995) J Mol Graph 13:165–174
35. Leach AR, Villet VJ (2003) An Introduction to Chemoinformatics, Kluwer, Dordrecht, The Netherlands, p 118
36. Used radii in Å: H - 1.62, C - 2.04, N - 1.93, O - 1.82, P - 2.22, Na - 1.02
37. Allinger NL, Zhou X, Bergsma J (1994) J Mol Struct (Theochem) 312:69–83
38. Shriver DF, Atkins PW, Langford CH (1995) Inorganic chemistry (2nd edition), Oxford University Press, p 35
39. Gaussian 03, Revision C.02, Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr. JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong, MW, Gonzalez C, Pople JA, Gaussian, Inc., Wallingford CT, 2004.
40. FreeImage Library, http://freeimage.sourceforge.net
41. Yasuno M, Yasuda N (2004) Aoki Computer Vision and Pattern Recognition Workshop, 27-02 June 2004, Conference proceedings, p 125
42. Soille P (1999) Morphological Image Analysis: Principles and Applications, Springer-Verlag
43. Wählby C, Sintorn IM (2004) J Microscopy 215:67–76
44. MatLab, http://www.mathworks.com/products/matlab/
45. Lipscomp LA, Peek ME, Morningstar ML, Verghis SM, Miller EM, Rich A, Essigmann JM, Williams LD (1995) Proc Natl Acad Sci USA 92:719–723
46. Xiang Y, Goodman MF, Beard WA, Wilson SH, Warshel A (2008) Proteins Struct Funct Bioinfo 70:231–247