

## Regular article

# The glycosyl C<sub>1'</sub>–N<sub>9</sub> bond of deoxyadenosine and deoxyguanosine: response to electrophilic attacks on the purinic nitrogen atoms

E. L. Cavalieri<sup>1</sup>, E. C. Vauthier<sup>2</sup>, A. Cossé-Barbi<sup>2</sup>, S. Fliszár<sup>3</sup>

<sup>1</sup> University of Nebraska Medical Center, Eppley Institute, Omaha, NE 68198-6805, USA

<sup>2</sup> Université Paris VII, CNRS ESA 7086, Institut de Topologie et de Dynamique des Systèmes, 1, rue Guy de la Brosse, 75005 Paris, France

<sup>3</sup> Département de Chimie, Université de Montréal, C.P. 6128 Succ. Centre-ville, Montréal, Québec H3C 3J7, Canada

Received: 13 September 1999 / Accepted: 3 February 2000 / Published online: 21 June 2000

© Springer-Verlag 2000

**Abstract.** Self-consistent-field computations shed light on two relevant conformations of deoxyadenosine (dA) and deoxyguanosine (dG): one with a pseudoequatorial C<sub>1'</sub>N<sub>9</sub> glycosyl bond and the other, a slightly more stable one, with its C<sub>1'</sub>N<sub>9</sub> bond in a bisectonal orientation. In dA, both the N<sub>3</sub> and N<sub>7</sub> nitrogens are plausible sites for electrophilic attack, but only N<sub>7</sub> is a plausible site in dG. The addition of H<sup>+</sup>, CH<sub>3</sub><sup>+</sup>, C<sub>2</sub>H<sub>5</sub><sup>+</sup> or *tert*-C<sub>4</sub>H<sub>9</sub><sup>+</sup> onto N<sub>7</sub> does not provoke notable structural modifications and leaves the base of dA and dG in an antiperiplanar (or nearly antiperiplanar) position with respect to the sugar C<sub>1'</sub>O<sub>4'</sub> bond, but N<sub>3</sub> additions cause the base to adopt a synperiplanar or strongly chiral position. This produces strong interactions between the purine and deoxyribose moieties, whose relief could aid the eventual cleavage of the glycosyl bond of dA. Addition of a radical cation onto N<sub>7</sub> reduces the dissociation energy of the glycosyl bond by an estimated 8 kcal mol<sup>-1</sup> in dA and 4 kcal mol<sup>-1</sup> in dG – a bond weakening likely to concur to a depurination of DNA induced by radical cations.

**Key words:** Depurination – Glycosyl CN bond – Deoxyribonucleosides

## 1 Introduction

Our targets are two deoxyribonucleoside (dR) molecules, namely, deoxyadenosine (dA) and deoxyguanosine (dG) (Fig. 1), and the damage they may suffer as a result of reactions with electrophiles. Motivation is drawn from observations regarding the depurination of DNA [1] and the hope of getting insight into its physics.

More specifically, attacks by H<sup>+</sup>, CH<sub>3</sub><sup>+</sup>, C<sub>2</sub>H<sub>5</sub><sup>+</sup> and *tert*-C<sub>4</sub>H<sub>9</sub><sup>+</sup> at the N<sub>3</sub> and N<sub>7</sub> sites of dG and dA are evaluated. The underlying physics is rooted in a formula for chemical bonds, applied to the C–N link whose

cleavage causes depurination. Our procedure, in reasonable agreement with direct density functional theory computations, offers insight into details not found in conventional estimates of bond energies. Traditional conformational analyses are an integral part of this study.

## 2 Outline of methods: the rupture of the glycosyl C–N bond

The expression [2, 3]

$$\varepsilon_{kl} = \varepsilon_{kl}^{\circ} + a_{kl}\Delta q_k + a_{lk}\Delta q_l \quad (1)$$

indicates how the intrinsic energy of a chemical bond linking atoms *k* and *l* depends on the electronic charges carried by the bond-forming atoms:  $\varepsilon_{kl}^{\circ}$  is for a reference bond with net charges  $q_k^{\circ}$  and  $q_l^{\circ}$  at atoms *k* and *l*, respectively, whereas  $\varepsilon_{kl}$  corresponds to modified charges  $q_k = q_k^{\circ} + \Delta q_k$  and  $q_l = q_l^{\circ} + \Delta q_l$ . ‘Intrinsic energy’ refers here to ground-state molecules at their potential minimum. The  $a_{kl}$  and  $a_{lk}$  parameters, “measuring” the changes in bond energy accompanying unit charge variations at atoms *k* and *l*, respectively, are readily deduced from theory, which gives [2, 3]

$$\varepsilon_{\text{CN}} = \varepsilon_{\text{CN}}^{\circ} - 0.603\Delta q_{\text{C}} - 0.448\Delta q_{\text{N}} \quad \text{kcal mol}^{-1}, \quad (2)$$

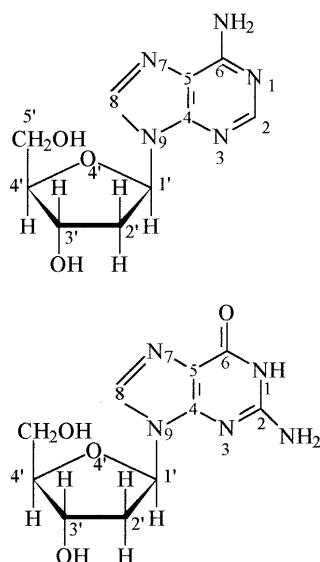
where  $\Delta q_{\text{C}}$  and  $\Delta q_{\text{N}}$  are expressed in millielectron units. Here we take the original parent dR molecule as a reference and select  $\varepsilon_{\text{CN}}^{\circ}$  as the energy of its sugar–purine C–N bond. For molecules that have reacted with a radical cation (R<sup>+</sup>), we write  $\varepsilon_{\text{CN}}$  with  $\Delta q_{\text{C}}$  and  $\Delta q_{\text{N}}$  indicating how the charges at carbon and nitrogen, respectively, differ from those of the reference C–N bond. Finally, the difference  $\Delta\varepsilon_{\text{CN}} = \varepsilon_{\text{CN}} - \varepsilon_{\text{CN}}^{\circ}$  measures the change in C–N bond energy on going from the reference molecule to that modified by a reaction with a radical cation, so

$$\Delta\varepsilon_{\text{CN}} = -0.603\Delta q_{\text{C}} - 0.448\Delta q_{\text{N}} \quad \text{kcal mol}^{-1}. \quad (3)$$

The intrinsic bond energy described by Eq. (1) represents the part of the total atomization energy (at the potential minimum of the molecule) contributed by a particular *k*–*l* bond. This contribution is dictated by actual properties of the molecule (internuclear distance and local charge densities) and should not be mistaken for the bond dissociation energy,  $D_{kl}$ , required to break that bond. The exact relationship between  $D_{kl}$  and  $\varepsilon_{kl}$  is known [3, 4], but for the problem at hand we can safely use

$$D_{\text{CN}} \simeq \varepsilon_{\text{CN}} - E_{\text{nb}}(\text{S} \cdot \text{B}) + \text{RE}(\text{S} \cdot) + \text{RE}(\text{B} \cdot), \quad (4)$$

where  $E_{\text{nb}}(\text{S} \cdot \text{B})$  is the nonbonded interaction energy between the sugar (S) and the base (B) moieties in the molecule, while  $\text{RE}(\text{S} \cdot)$



**Fig. 1.** Atom numbering in deoxyadenosine (*top*) and deoxyguanosine (*bottom*)

and  $RE(B\cdot)$  are the reorganizational energies of the corresponding free radicals in Sanderson's approximation [5]. ( $RE$  measures the relaxation of a fragment as found in a molecule to give the corresponding ground-state free radical [4].) Under certain circumstances these reorganizational energies can be treated as constants. In our problem, addition of  $R^+$  to a dR leaves, at least to a good approximation, the purine free radical and the purine moiety in the molecule in what is an essentially planar geometry [6], so  $\Delta RE(B\cdot) \simeq 0$ . For the deoxyribose part we write  $\Delta RE(S\cdot) \simeq 0$  because the energies of the pertinent sugar conformers are practically all the same [6]. In that situation, Eq. (4) leads to  $\Delta D_{CN} \simeq \Delta \varepsilon_{CN} - \Delta E_{nb}$ , i.e.,

$$\Delta D_{CN} \simeq -0.603\Delta q_C - 0.448\Delta q_N - \Delta E_{nb} \text{ kcal mol}^{-1} \quad (5)$$

This is our working formula for use in comparisons between C—N bonds.

Our calculations were carried out with the help of the **MONSTERGAUSS** program [7], using the familiar **STO-3G** minimal basis set [8]. All scale factors were optimized following Davidson's optimally conditioned minimization technique [9]. These optimizations are of utmost importance in order to obtain consistent sets of charge results [3, 10]. The minimal basis is well suited to do this job. The net (nuclear minus electronic) charges were deduced by means of Mulliken population analysis [11]. The generalization [3, 10] not involving the halving of all overlap population terms is not required in the present situation, in which only atoms engaged in the same type of bonding are compared to one another. Briefly, Mulliken net charges are appropriate in evaluations of  $\Delta q_C$  and  $\Delta q_N$ , with C in its deoxyribose environment and N always flanked by two  $sp^2$  carbons. Numerous analyses of self-consistent field (SCF) charges [3, 10] and detailed comparisons involving extended basis sets and configuration interaction calculations [12] indicate that **STO-3G** charge differences, such as those obtained here, are systematically overestimated. In line with earlier work [3, 10, 12], our results for  $\Delta q_C$  and  $\Delta q_N$  are scaled by a factor of 35.1/69.4 as the results obtained from Eq. (3) are known to be generally valid within experimental accuracy [3]. It is understood that not every situation can be handled as simply as that.

### 3 Results

The crystallographic data of dRs, compared with A- and B-type DNA fragments [6], indicate that the former

represent valid models for helical polynucleotides; hence their selection. Geometry optimizations were carried out in the usual way. All the internal and dihedral angles, as well as all the internuclear distances, were allowed to change and this optimization also included that of the scale factors. A free rotation around the glycosyl C—N bond was allowed. The pertinent glycosyl torsion angle,  $\chi$ , and the endocyclic furanose torsion angles,  $v_0, v_1, v_2, v_3$  and  $v_4$  (Table 1), were calculated as well as the pseudorotation phase angle,  $P$

$$\tan P = \frac{(v_4 + v_1) - (v_3 + v_0)}{2v_2(\sin 36^\circ + \sin 72^\circ)}$$

For each molecule, two conformers were found, namely, one with a pseudoequatorial  $C_1'N_9$  bond, with  $C_4'O_4'C_1'N_9 \simeq 200^\circ$  (Table 2) and the other with a bisectonal  $C_1'N_9$  bond, for  $C_4'O_4'C_1'N_9 \simeq 240^\circ$  (Table 3). (Common use identifies the pseudoequatorial conformer through its nearly  $C_2'$ -endo sugar puckering mode and the second through its  $C_3'$ -endo pucker.) Here we indicate the total SCF energy, the length,  $R_{CN}$ , of the glycosyl  $C_1'N_9$  bond as well as the **STO-3G** Mulliken net atomic charges,  $q_C$  on  $C_1'$  and  $q_N$  on  $N_9$ . Finally, we also report  $P$  and  $\chi$ . The former throws light on the conformation ( $C_3'$ -endo,  $C_2'$ -endo,  $O_4'$ -endo, etc.) of the furanose cycle, while  $\chi$  reveals the syn or anti orientation of the purine base with respect to the  $C_1'O_4'$  bond.

#### 3.1 Conformation of the furanose cycle

The crystallographic data of dRs and of A- and B-type DNA indicate [6] that the oses adopt  $C_3'$ -endo ( $0^\circ \leq P \leq 36^\circ$ ) or  $C_2'$ -endo ( $144^\circ \leq P \leq 180^\circ$ ) conformations; however, the flexibility of the  $C_3'$ -endo furanoses is somewhat reduced (as found in the A-type dRs and DNA) while that of the  $C_2'$ -endo furanoses, namely in B-type dRs and DNA, is considerable. The latter can adopt conformations that are close to  $C_2'$ -endo, namely, the  $C_1'$ -exo ( $108^\circ \leq P \leq 144^\circ$ ) or the  $C_3'$ -exo ( $180^\circ \leq P \leq 216^\circ$ ), as well as the  $O_4'$ -endo ( $72^\circ \leq P \leq 108^\circ$ ) conformations. This is indeed what we observe. The conformers with a bisectonal  $C_1'N_9$  bond orientation exhibit only one ose (with *tert*- $C_4H_9^+$  on the  $N_3$  atom of dG) that departs considerably from the  $C_3'$ -endo conformation. In contrast, the furanoses with a pseudoequatorial  $C_1'N_9$  bond readily adopt  $C_1'$ -exo or  $O_4'$ -endo conformations.

**Table 1.** Definition of pertinent torsion angles

Angle	Atoms involved
$\chi$	$O_4'-C_1'-N_9-C_4$
$v_0$	$C_4'-O_4'-C_1'-C_2'$
$v_1$	$O_4'-C_1'-C_2'-C_3'$
$v_2$	$C_1'-C_2'-C_3'-C_4'$
$v_3$	$C_2'-C_3'-C_4'-O_4'$
$v_4$	$C_3'-C_4'-O_4'-C_1'$

**Table 2.** Deoxyribonucleosides with their  $C_1'N_9$  bond in pseudoequatorial orientation

Base	Reagent	$E$ (au)	$R_{CN}$ (Å)	$q_C$ (me)	$q_N$ (me)	$P$	$\chi$
Adenine	None	-871.92434	1.4488	316.54	-201.28	115	153
	H <sup>+</sup> on N <sub>1</sub>	-872.40535	1.4661	293.41	-143.79	115	163
	CH <sub>3</sub> <sup>+</sup> on N <sub>1</sub>	-910.98841	1.4634	294.63	-147.97	117	147
	H <sup>+</sup> on N <sub>3</sub>	-872.39839	1.4659	298.39	-179.65	104	157
	CH <sub>3</sub> <sup>+</sup> on N <sub>3</sub>	-910.98366	1.4618	285.13	-178.23	89	52
	C <sub>2</sub> H <sub>5</sub> <sup>+</sup> on N <sub>3</sub>	-949.56877	1.4631	287.36	-180.32	90	45
	H <sup>+</sup> on N <sub>7</sub>	-872.39014	1.4734	287.62	-127.25	124	176
	CH <sub>3</sub> <sup>+</sup> on N <sub>7</sub>	-910.97893	1.4714	294.11	-134.75	115	170
C <sub>2</sub> H <sub>5</sub> <sup>+</sup> on N <sub>7</sub>	-949.56286	1.4699	287.52	-128.91	113	167	
Guanine	None	-945.77575	1.4490	313.47	-177.82	119	153
	H <sup>+</sup> on N <sub>3</sub>	-946.24308	1.4622	291.01	-156.28	99	162
	CH <sub>3</sub> <sup>+</sup> on N <sub>3</sub>	-984.82503	1.4567	280.87	-160.09	89	60
	C <sub>2</sub> H <sub>5</sub> <sup>+</sup> on N <sub>3</sub>	-1023.40798	1.4580	282.75	-163.53	89	66
	H <sup>+</sup> on N <sub>7</sub>	-946.26303	1.4716	291.46	-127.13	144	180
	CH <sub>3</sub> <sup>+</sup> on N <sub>7</sub>	-984.85370	1.4678	294.32	-129.52	134	181
	C <sub>2</sub> H <sub>5</sub> <sup>+</sup> on N <sub>7</sub>	-1023.44113	1.4663	295.96	-131.01	128	193

**Table 3.** Deoxyribonucleosides with their  $C_1'N_9$  bond in bisectonal orientation

Base	Reagent	$E$ (au)	$R_{CN}$ (Å)	$q_C$ (me)	$q_N$ (me)	$P$	$\chi$
Adenine	None	-871.93113	1.4546	296.50	-205.59	7	220
	H <sup>+</sup> on N <sub>1</sub>	-872.41758	1.4725	285.27	-153.67	5	218
	CH <sub>3</sub> <sup>+</sup> on N <sub>1</sub>	-911.00147	1.4718	285.69	-156.67	6	218
	H <sup>+</sup> on N <sub>3</sub>	-872.41379	1.4711	285.76	-173.52	28	343
	CH <sub>3</sub> <sup>+</sup> on N <sub>3</sub>	-910.98967	1.4762	277.14	-176.07	37	301
	C <sub>2</sub> H <sub>5</sub> <sup>+</sup> on N <sub>3</sub>	-949.57256	1.4767	278.06	-174.09	40	303
	<i>tert</i> -C <sub>4</sub> H <sub>9</sub> <sup>+</sup> on N <sub>3</sub>	-1026.72401	1.4817	293.85	-196.72	21	219
	H <sup>+</sup> on N <sub>7</sub>	-872.40869	1.4804	278.74	-145.12	1	214
	CH <sub>3</sub> <sup>+</sup> on N <sub>7</sub>	-910.99679	1.4792	278.75	-147.46	1	214
	C <sub>2</sub> H <sub>5</sub> <sup>+</sup> on N <sub>7</sub>	-949.58352	1.4788	278.82	-146.82	1	215
	<i>tert</i> -C <sub>4</sub> H <sub>9</sub> <sup>+</sup> on N <sub>7</sub>	-1026.74533	1.4781	278.57	-147.23	3	215
	Guanine	None	-945.78268	1.4535	292.78	-178.95	7
H <sup>+</sup> on N <sub>3</sub>		-946.25939	1.4661	283.50	-155.36	25	340
CH <sub>3</sub> <sup>+</sup> on N <sub>3</sub>		-984.82950	1.4700	278.66	-156.78	38	291
C <sub>2</sub> H <sub>5</sub> <sup>+</sup> on N <sub>3</sub>		-1023.41164	1.4698	280.43	-157.51	38	284
<i>tert</i> -C <sub>4</sub> H <sub>9</sub> <sup>+</sup> on N <sub>3</sub>		-1100.55983	1.4683	290.87	-160.76	-27	262
H <sup>+</sup> on N <sub>7</sub>		-946.28025	1.4769	281.39	-143.53	1	214
CH <sub>3</sub> <sup>+</sup> on N <sub>7</sub>		-984.86992	1.4759	281.94	-145.42	2	214
C <sub>2</sub> H <sub>5</sub> <sup>+</sup> on N <sub>7</sub>		-1023.45771	1.4757	281.72	-147.32	2	215
<i>tert</i> -C <sub>4</sub> H <sub>9</sub> <sup>+</sup> on N <sub>7</sub>		-1100.62581	1.4747	282.05	-148.62	3	214

### 3.2 Syn or anti purine base

$\chi$  reveals the syn ( $-90^\circ \leq \chi \leq 90^\circ$ ) or anti ( $90^\circ \leq \chi \leq 270^\circ$ ) position of the base, by reference to the sugar  $C_1'O_4'$  bond.

Most of the molecules investigated here, whether their sugar puckering is in  $C_2'$ -endo or in  $C_3'$ -endo mode, exhibit an antiperiplanar ( $150^\circ \leq \chi \leq 210^\circ$ ) or nearly antiperiplanar base with respect to the  $C_1'O_4'$  bond. Exceptions to this anti orientation of the purinic base are offered by dAs and dGs with  $R^+$  attached to the N<sub>3</sub> nitrogen, as with H<sup>+</sup>, for example (Table 3). This forces the base into an almost synperiplanar position with respect to the  $C_1'O_4'$  bond. This also holds true for CH<sub>3</sub><sup>+</sup>, C<sub>2</sub>H<sub>5</sub><sup>+</sup> and *tert*-C<sub>4</sub>H<sub>9</sub><sup>+</sup> on N<sub>3</sub>. Then, the base finds itself in a strongly clinal situation ( $\chi \simeq 270^\circ$ ). Down the line, major structural changes mark the attacks at the N<sub>3</sub>, in lieu of the N<sub>7</sub>, sites of dA and dG.

### 3.3 Nonbonded interactions

In the anti conformation there is no particular steric hindrance between sugar and base, but in the syn conformation the bulky part of the base is located over the sugar, giving rise to close interatomic contacts [6]. With  $R^+$  added onto N<sub>1</sub> or N<sub>7</sub> no notable change, i.e., no notable steric hindrance, is to be expected, so  $\Delta E_{nb} = E_{nb}^+ - E_{nb}^{reference} \simeq 0$  can be regarded as a reasonable approximation; however, with N<sub>3</sub> things are different. Addition at that position causes energetically unfavorable interactions with the sugar, which are in part relieved by a change in conformation, to adopt a nearly synperiplanar or strongly gauche position with respect to the  $C_1'N_9$  link; nonbonded interactions are minimized, but not at the level attained by the reference in its anti conformation, thus suggesting that  $\Delta E_{nb} > 0$ .

Now, as indicated by Eq. (5), a positive  $\Delta E_{nb}$  “assists” the cleavage of the glycosyl C—N bond because it

renders  $\Delta D_{\text{CN}}$  more negative. This qualitative assessment suggests that one should not judge the stability of the glycosyl bonds solely in terms of charge effects, but that nonbonded interactions may occasionally play a decisive role.

### 3.4 The dissociation of the glycosyl C—N bond

The net atomic charges of the pertinent  $C_{1'}$  and  $N_9$  atoms of dA and dG (Tables 2, 3) serve as references for the calculation of the Mulliken charge differences  $\Delta q_{\text{C}}$  and  $\Delta q_{\text{N}}$  and of the rescaled  $\Delta q_{\text{k}} = (35.1/69.4)\Delta q_{\text{k}}^{\text{Mulliken}}$  results indicated in Table 4, for use in Eq. (5).

The reported  $\Delta D_{\text{CN}}$  results do not include nonbonded terms for the  $N_3$  adducts. Moreover, we take  $\Delta \text{RE}(\text{S}\cdot) \simeq 0$  and  $\Delta \text{RE}(\text{B}\cdot) \simeq 0$ . By and large, this situation has undeniable merits: it isolates the part of  $\Delta D_{\text{CN}}$  that is exclusively due to the changes in electronic charges at the bond-forming atoms, as shown in Table 4. Now, as regards  $\Delta E_{\text{nb}}$ , strong interatomic contacts between the base and sugar moieties in  $N_3$ -substituted dRs are expected to assist the  $C_{1'}$  $N_9$  bond cleavage, i.e., to render  $\Delta D_{\text{CN}}$  more negative than shown in Table 4.

Another point deserves attention. As revealed by Tables 2 and 3, a reaction of any  $\text{R}^+$  at the  $N_7$  site of dG is energetically overwhelmingly favored over a reaction at  $N_3$ .  $\text{H}^+$ , for example, in  $C_{3'}$ -endo molecules, favors  $N_7$  by 13 kcal mol<sup>-1</sup>, while  $\text{C}_2\text{H}_5^+$  and *tert*- $\text{C}_4\text{H}_9^+$  favor it by about 29 and about 41 kcal mol<sup>-1</sup>, respectively. Similar results are found for the pseudoequatorial conformers of dG.  $N_3$  adducts are highly improbable and depurination of dG due to attacks at that position is not a likely event. For dA the trends differ. In the bisectinal form,  $\text{H}^+$  favors  $N_3$  by about 3.2 kcal mol<sup>-1</sup>, but  $\text{C}_2\text{H}_5^+$  favors  $N_7$  by about 6.9 kcal mol<sup>-1</sup>, whereas the pseudoequatorial

form of dA consistently favors  $N_3$  by about  $4 \pm 1$  kcal mol<sup>-1</sup>. This renders attacks on  $N_3$  plausible.

## 4 Conclusions

The computation of relatively small dR motifs reveals important aspects of DNA chemistry. Two structures merit attention: one with a pseudoequatorial  $C_{1'}$  $N_9$  glycosyl bond and the other, a slightly more stable form, with a bisectinal  $C_{1'}$  $N_9$  bond. Electrophiles reacting with DNA surely affect these bonds.

The addition of  $\text{R}^+$  onto the  $N_1$  and  $N_7$  nitrogen atoms does not provoke any notable conformational modification, neither at the level of the preferred bisectinal orientation of the  $C_{1'}$  $N_9$  bond, which is weakened by the anomeric effect due to atom  $\text{O}_{4'}$ , nor as regards the preferred  $C_{3'}$ -endo conformation of the furanose moiety, nor as concerns the antiperiplanar (or nearly antiperiplanar) position of the base with respect to the  $C_{1'}$  $\text{O}_{4'}$  bond. In contrast, the addition of  $\text{H}^+$ ,  $\text{CH}_3^+$ ,  $\text{C}_2\text{H}_5^+$  or *tert*- $\text{C}_4\text{H}_9^+$  onto  $N_3$  provokes drastic conformational changes, where the base adopts a synperiplanar or strongly clinal position. For the  $N_3$ -substituted dAs, we tentatively submit that their depurination is prompted by the relief of nonbonded contacts between sugar and base in nearly synperiplanar or highly clinal positions.

In dA, both  $N_3$  and  $N_7$  are plausible targets for electrophilic attack. With dG, however,  $N_7$  is consistently favored over  $N_3$ . For all practical purposes, there seems no point in considering the formation – and thus the decomposition – of  $N_3$ -substituted dGs.

The sizeable weakening of the glycosyl C—N bond accompanying an attack on  $N_1$  in dA is of only marginal interest since attacks of that sort do not take place in

**Table 4.** The glycosyl C—N bond response to selected electrophiles:  $\Delta q$  charges (*me*) and  $\Delta D_{\text{CN}}$  energies (kcal mol<sup>-1</sup>) relative to those of the parent deoxyribonucleosides

Base	Reagent	Pseudoequatorial CN			Bisectinal CN		
		$\Delta q_{\text{C}}$	$\Delta q_{\text{N}}$	$\Delta D_{\text{CN}}$	$\Delta q_{\text{C}}$	$\Delta q_{\text{N}}$	$\Delta D_{\text{CN}}$
Adenine	$\text{H}^+$ on $N_1$	-11.7	29.1	-6.0	-5.7	26.3	-8.3
	$\text{CH}_3^+$ on $N_1$	-11.1	27.0	-5.4	-5.5	24.7	-7.8
	$\text{H}^+$ on $N_3$	-9.2	10.9	0.6	-5.4	16.2	-4.0
	$\text{CH}_3^+$ on $N_3^{\text{a}}$	-15.9	11.7	4.4	-9.8	14.9	-0.8
	$\text{C}_2\text{H}_5^+$ on $N_3$	-14.8	10.6	4.2	-9.3	15.9	-1.5
	<i>tert</i> - $\text{C}_4\text{H}_9^+$ on $N_3$				-1.3	4.5	-1.2
	$\text{H}^+$ on $N_7$	-14.6	37.4	-8.0	-9.0	30.6	-8.3
	$\text{CH}_3^+$ on $N_7$	-11.3	33.7	-8.2	-9.0	29.4	-7.8
	$\text{C}_2\text{H}_5^+$ on $N_7$	-14.7	36.6	-7.6	-8.9	29.7	-7.9
	<i>tert</i> - $\text{C}_4\text{H}_9^+$ on $N_7$				-9.1	29.5	-7.8
Guanine	$\text{H}^+$ on $N_3$	-11.4	10.9	2.0	-4.7	11.9	-2.5
	$\text{CH}_3^+$ on $N_3^{\text{a}}$	-16.5	9.0	5.9	-7.1	11.2	-0.7
	$\text{C}_2\text{H}_5^+$ on $N_3$	-15.5	7.2	6.1	-6.3	10.8	-1.1
	<i>tert</i> - $\text{C}_4\text{H}_9^+$ on $N_3$				-1.0	9.2	-3.5
	$\text{H}^+$ on $N_7$	-11.1	25.6	-4.8	-5.8	17.9	-4.6
	$\text{CH}_3^+$ on $N_7$	-9.7	24.4	-5.1	-5.5	16.9	-4.3
	$\text{C}_2\text{H}_5^+$ on $N_7$	-8.9	23.7	-5.3	-5.6	16.0	-3.8
	<i>tert</i> - $\text{C}_4\text{H}_9^+$ on $N_7$				-5.4	15.3	-3.6

<sup>a</sup> Direct density functional theory computations indicate that addition of  $\text{CH}_3^+$  onto the  $N_7$  and  $N_3$  nitrogen atoms provoke  $\Delta D_{\text{CN}}$  differences between the two of  $-5.0$  kcal mol<sup>-1</sup> in deoxyadenosine and of  $-3.9$  kcal mol<sup>-1</sup> in deoxyguanosine, the C—N bond weakening being more important by these amounts for additions occurring at  $N_7$ , in acceptable agreement with the results given by Eq. (3)

authentic DNA structures because of the inaccessibility of this particular site.

Finally, as concerns the attack on N<sub>7</sub>, our results clearly point to a significant weakening of the glycosyl C—N bond. Now, most of the substituents whose attack on N<sub>7</sub> has been studied [1], polycyclic aromatic hydrocarbons and estrogens, were found to provoke rapid cleavage of the C—N bond, resulting in instantaneous depurination.

It seems reasonable to argue that at least part of the explanation addressing the depurination of DNA has something to do with the bond weakening, first and foremost due to local charge effects; however, these are not the only contributors. Steric relief promoting the glycosyl bond cleavage is a strong argument for explaining the depurination of N<sub>3</sub>-substituted dA in DNA.

Surely, it would be desirable to go beyond the present gas-phase model and to consider its condensed-phase analog, namely, by assessing the role of the appropriate zero-point energy (ZPE) and heat-content ( $H_T - H_0$ ) energy. The assumption is that attacks on N<sub>3</sub> or N<sub>7</sub> do not differ significantly as regards the changes,  $\Delta(\text{ZPE} + H_T - H_0)$ , accompanying the cleavage of the CN bond. We hope that this work suggests a useful framework for tackling the specific problem of bond breaking leading to depurination.

*Acknowledgements.* We are grateful to R. A. Poirier for his assistance with the MONSTERGAUSS set of programs and to the Center of Computational Research of the University of Paris VII (CCR Jussieu) for the large amount of computer time made available.

## References

1. Cavalieri E, Rogan E (1998) In: Neilson AH (ed) The handbook of environmental chemistry vol 3J: PAH and related compounds. Springer, Berlin Heidelberg New York, pp 81–117
2. Fliszár S (1980) J Am Chem Soc 102: 6946
3. Fliszár S (1994) Atoms, chemical bonds and bond dissociation energies, Lecture Notes in Chemistry 63, Springer, Berlin Heidelberg New York
4. (a) Fliszár S, Minichino C (1987) Can J Chem 65: 2495; (b) Fliszár S, Minichino C (1987) J Phys (Paris) Colloq C4 (Suppl 9): 367
5. (a) Sanderson RT (1982) J Org Chem 47: 3835; (b) Sanderson RT (1976) Chemical bonds and bond energy, 2nd edn. Academic, New York
6. Saenger W (1984) Principles of nucleic acid structure. Springer, Berlin Heidelberg New York
7. Peterson MR, Poirier RA (1981) University of Toronto, Ontario
8. Hehre WJ, Stewart RF, Pople JA (1969) J Chem Phys 51: 2657
9. Davidon WC (1975) Math Program 9: 1
10. (a) Fliszár S, Kean G, Macaulay R (1974) J Am Chem Soc 96: 4353; (b) Fliszár S, Goursot A, Dugas H (1974) J Am Chem Soc 96: 4358; (c) Kean G, Fliszár S (1974) Can J Chem 52: 2772; (d) Henry H, Fliszár S (1978) J Am Chem Soc 100: 3312
11. (a) Mulliken RS (1955) J Chem Phys 23: 1833; (b) Mulliken RS (1955) J Chem Phys 23: 1841; (c) Mulliken RS (1955) J Chem Phys 23: 2338; (d) Mulliken RS (1955) J Chem Phys 23: 2343
12. (a) Mijoule C, Leclercq JM, Comeau M, Fliszár S, Picard M (1992) Can J Chem 70: 68; (b) Mijoule C, Leclercq JM, Odiot S, Fliszár S (1985) Can J Chem 63: 1741