Journal of
Synchrotron Radiation

ISSN 0909-0495

Received 30 October 2007
Accepted 8 January 2008

The attachment of amino fragment to purine: inner-shell structures and spectra

Saumitra Saha,[a] Feng Wang,[a]* Janay B. MacNaughton,[b] Alex Moewes[c] and Denalo P. Chong[d]

[a]Centre for Molecular Simulation, Swinburne University of Technology, Hawthorn, Melbourne, Victoria 3122, Australia, [b]Stanford Synchrotron Radiation Laboratory, 2575 Sand Hill Road, MS 69, Menlo Park, CA 94025, USA, [c]Department of Physics and Engineering Physics, University of Saskatchewan, Canada S7N 5E2, and [d]Department of Chemistry, University of British Columbia, Vancouver, 2036 Main Mall, BC, Canada V6T 1Z1. E-mail: fwang@swin.edu.au

The impact of the amino fragment ($-NH_2$) attachment on the inner-shell structures and spectra of unsubstituted purine and the purine ring of adenine are studied. Density functional theory calculations, using the LB94/TZ2P//B3LYP/TZVP model, reveal significant site-dependent electronic structural changes in the inner shell of the species. A condensed Fukui function indicates that all of the N and C sites, except for $N_{(1)}$ and $C_{(5)}$, demonstrate significant electrophilic reactivity ($f^- > 0.5$ in $|e|$) in the unsubstituted purine. Once the amino fragment binds to the $C_{(6)}$ position of purine to form adenine, the electrophilic reactivity of these N and C sites is greatly reduced. As expected, the $C_{(6)}$ position experiences substantial changes in energy and charge transfer, owing to the formation of the $C-NH_2$ bond in adenine. The present study reveals that the N1*s* spectra of adenine inherit the N1*s* spectra of the unsubstituted purine, whereas the C1*s* spectra experience significant changes although purine and adenine have geometrically similar carbon frames. The findings also indicate that the attachment of the NH_2 fragment to purine exhibits deeply rooted influences to the inner-shell structures of DNA/RNA bases. The present study suggests that some fragment-based methods may not be applicable to spectral analyses in the inner shell.

Keywords: purine; adenine; inner-shell binding-energy spectra; Hirshfeld charge; condensed Fukui function.

1. Introduction

A large fraction of biologically relevant damage in living cells can be traced to structural and chemical modifications of cellular DNA and RNA. Studies have revealed that radiation can induce damage in living cells and create exceeding amounts of low-energy electrons along the ionization track (Sanche, 2003). Electrons with energies of the order of 10 eV can immediately induce reactions, both directly in DNA and indirectly through the environment (*e.g.* medium). Recent high-resolution synchrotron X-ray spectroscopic studies include X-ray absorption spectra (XAS) of pyridine (Kolczewski *et al.*, 2001) and DNA bases (Fujii *et al.*, 2004; MacNaughton *et al.*, 2005; Harada *et al.*, 2006), the photoelectron spectrum (PES) of linear alkanes (Karlsen *et al.*, 2002; Svensson, 2005) and other biologically important species such as amino acids (Powis *et al.*, 2003; Plekan *et al.*, 2007*a*).

Core orbitals are essentially localized and they exhibit properties of theoretical and practical significance (Svensson, 2005). Different core energies in the molecular environment are indicated in core binding-energy spectra (Siegbahn *et al.*, 1969), which provide valuable information for studying charge distribution and chemical bonding of molecules (Abu-samha *et al.*, 2007). The binding energy of inner-shell electrons depends on the charge distribution in a molecule, and the ability of the neighboring atoms to reorganize upon the local perturbation of the positive charge introduced through ionization (Siegbahn *et al.*, 1969; Karlsen *et al.*, 2002). Hence, core-ionization energies express the ability of a molecule to accept charge at a specific site (Karlsen *et al.*, 2002; Wang, 2005). In a recent investigation, Wang (2005) demonstrated significant changes in the inner-shell orbitals as a result of the C=C double-bond saturation in norbornadiene (C_7H_8), compared with norbornene (C_7H_{10}) and norbornane (C_7H_{12}). The study indicates that the core orbital chemical shift and wavefunction distortion of the strained organic molecular species are not subtle in the core shell. These differences in inner-shell orbitals exist even though norbornadiene, norbornene and norbornane are structurally similar bicyclic hydrocarbons.

Accurate core-electron binding energies (CEBEs) for many small molecules are known both experimentally and theoretically. A comprehensive compilation of the observed CEBEs can be found by Jolly *et al.* (1984). Reliable CEBEs for larger molecules such as amino acids, and nucleobases like alanine (Powis *et al.*, 2003) in the gas phase are technically more difficult. Reliable 1*s* CEBEs for B to F atoms have been calculated by density functional theory with the method denoted by $\Delta E(PW86\text{-}PW91) + C_{rel}$ (Chong, 1995). The last term, C_{rel}, is a semi-empirical correction for relativistic effects (Chong, 1995). Other groups of systems, such as carbonyls (and H_2O), small hydrocarbons, nitriles (and fluorobenzene) and various core centers in the methyl-methacrylate molecule were studied by the Swedish group (Triguero *et al.*, 1999) using the $\Delta E(B88\text{-}P86) + C_{rel}$ method.

Reliable XAS of small molecules have been determined experimentally and theoretically. However, theoretical XAS are more challenging than simple $\Delta E(PW86\text{-}PW91)$ or $\Delta E(B88\text{-}P86)$ calculations because Rydberg-like levels require careful basis set design and the relative intensities are not simple to determine. Different groups use various methods with varying degrees of success, such as $\Delta\varepsilon(B88\text{-}P86)$ with transition potential (Triguero *et al.*, 1998; Karlsen *et al.*, 2002), GSCF3 (MacNaughton *et al.*, 2005), $\Delta E(PW86\text{-}PW91)$ combined with time-dependent DFT for *f* values (Chong, 2005), real-space Green's function formalism and curved-wave multiple-scattering theory (Rehr & Ankudinov, 2001), $\Delta E(B88\text{-}P86)$ with transition potential for *f* values (Kolczewski *et al.*, 2001), Hartree–Fock with static exchange (Mochizuki *et al.*, 2001), and the method of CVR-B3LYP, developed very recently (Nakata *et al.*, 2006).

Quantitative treatment of the near-edge X-ray absorption fine-structure (NEXAFS) spectra for comparatively large molecules such as the DNA bases and their derivatives are not yet fully understood, since calculations far from trivial are necessary (Mochizuki *et al.*, 2001). Recent state-of-the-art measurements such as XAS and X-ray emission spectroscopy (MacNaughton *et al.*, 2005; Harada *et al.*, 2006) confirmed earlier experimental conclusions (Barber & Clark, 1970) that the O-*K*, N-*K* and C-*K* inner-shell chemical shifts reveal the specific chemical environments of the individual DNA bases. A theoretical and experimental study of NEXAFS and X-ray photoelectron spectra (XPS) of adenine and thymine in the gas phase was reported recently (Plekan *et al.*, 2007*b*), unraveling the core-level chemical shifts and valence electron relaxation. Theoretical binding energies were also reported for the XPS spectra of nucleobases (Takahata *et al.*, 2006).

Correlation of the inner-shell electronic structure of the DNA purine bases such as adenine with the unsubstituted purine molecule provides important insight into the understanding of their structures and functions. Chemical shifts of the inner-shell binding energies and electron distributions differentiate atoms of the same element located at the non-equivalent sites in the purine rings, producing a simpler but novel picture of the molecular structures than the largely delocalized valence space information. For example, the reasons underlining the double spectral peaks appearing in the N1*s* NEXAFS and XAS spectra of adenine obtained in the experiments (Harada *et al.*, 2006; Fujii *et al.*, 2004; Mochizuki *et al.*, 2001) are revealed by amino (–N–) and imino (=N–) bonding. In this work, inner-shell responses to the purine ring resulting from the attachment of amino fragment in adenine are studied from electronic structural information of both adenine and purine.

2. Method and computational details

The molecular structure of adenine in the gas phase is slightly non-planar, which is mainly caused by the pyramidalization of the $-NH_2$ group (Dong & Miller, 2002; Sponer & Hobza, 2003; Wang *et al.*, 2005; Downton & Wang, 2006; Guerra *et al.*, 2006). For simplicity, both adenine and purine are treated as if they were planar molecules, by imposing the C_s point group symmetry (σ-plane) in the present work. Fig. 1 gives the chemical structures and numbering nomenclature of adenine and purine. This idealization to planarity not only simplifies the electronic calculations because the a' and a'' states are treated in separated and independent Hamiltonian matrices, but also provides insight into chemical bonding by separating the in-plane σ bonds and out-of-plane π bonds. Such treatment is particularly advantageous in valence space applications (Wang *et al.*, 2005; Saha *et al.*, 2005).

The geometry optimizations are based on the B3LYP/TZVP model, which were detailed previously (Wang *et al.*, 2005) using the GAMESS-US02 (Schmidt *et al.*, 1993) package of computational chemistry programs. The core vertical ionization energy calculations of the base pair employed the LB94/TZ2P model (van Leeuwen & Baerends, 1994), where the TZ2P basis set is a Slater-type triple-ζ plus double-polarization basis set (van Lenthe & Baerends, 2003), the closest basis set to the DGauss TZVP basis set (Godbout *et al.*, 1992) employed in the B3LYP/TZVP model. The LB94/TZ2P model is embedded in the *ADF* (Baerends *et al.*, 2004) package of computational chemistry programs. In addition, the ΔKohn–Sham calculations, which employed the gradient correlated Becke (B88) (Becke, 1988) for exchange functional and Perdew (P86) (Perdew, 1986) for the correlation functional, together with a triple-ζ valence plus polarization for N and C but double-ζ valence plus polarization for H basis set, are generated using the *StoBe* computational chemistry programs

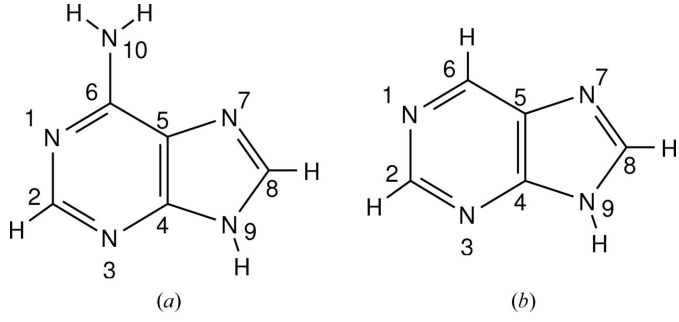

Figure 1
Structure and numbering convention of (*a*) adenine and (*b*) purine.

(Hermann *et al.*, 2005). Computational details using the *StoBe* method were detailed previously (MacNaughton *et al.*, 2005).

Calculations for molecular properties such as Hirshfeld atomic charges (Hirshfeld, 1977), molecule dipole moments and condensed Fukui functions (Fukui, 1982), based on the LB94/TZ2P model, are generated using the same procedure as described previously (Saha *et al.*, 2006).

3. Results and discussion

Structurally, the attachment of the amino fragment ($-NH_2$) to the unsubstituted purine generates the molecule adenine. The attachment of the fragment causes a total energy reduction of $55.366136E_h$ and a dipole moment reduction of 1.29 Debye, based on the B3LYP/TZVP model. However, geometrical changes with respect to the unsubstituted purine are rather small. For example, the overall purine ring length, $R_p = R_6 + R_5$, where R_6 and R_5 are the perimeters of the hexagon and pentagon rings (Wang *et al.*, 2005), respectively, remains almost unchanged: the R_p of purine is given by 15.010 Å and in adenine by 15.012 Å. The hexagon ring length (R_6) increases slightly, from 8.147 Å (purine) to 8.162 Å (adenine), which is compensated by the small decrease of the pentagon ring length (R_5), from 6.863 Å (purine) to 6.850 Å (adenine) (Wang *et al.*, 2005).

The condensed Fukui function, f_c^-, of the pair indicating the capacity of an electrophilic attack is presented in Fig. 2 for the N and C atoms. The electrophilic reactivity of N and C sites in purine is significantly different, depending on the type (N or C) and site locations. However, in adenine, the condensed Fukui functions for all sites, regardless of the types and site locations, are significantly reduced to well below 0.2. The reduction of the condensed Fukui function in adenine indicates the reduction of the site-related electrophilic reactivity from their purine counterparts. Moreover, in both the N and the C cases shown in Fig. 2, the site-dependent electrophilic reactivity of the unsubstituted purine is very different. For example, the order of the condensed Fukui function, f_c^-, in the unsubstituted purine is $C_{(5)} < C_{(2)} < C_{(4)} < C_{(6)} < C_{(8)}$ for the C sites and $N_{(1)} < N_{(7)} < N_{(9)} < N_{(3)}$ for the N sites. In particular, the condensed Fukui function variations of the $N_{(3)}$ and $C_{(8)}$ sites are the most significant in response to the attachment of the amino group $-NH_2$.

The distribution of the charge density according to the Hirshfeld scheme (Hirshfeld, 1977) based on the LB94/TZ2P wavefunctions is analyzed next. The atomic Hirshfeld charge is a hypothetical 'promolecule' with electron density $\Sigma\rho_B$ constructed by superposition of spherically symmetric charge densities ρ_B of an isolated atom B. The Hirshfeld atomic charge (Q_A^H) is obtained by subtracting the resulting partial electron density associated with atom A from the corresponding nuclear charge Z as (Bickelhaupt & Baerends, 2000; Hirshfeld, 1977) $Q_A^H = Z - \int w_A(\mathbf{r})\rho(\mathbf{r})\,d\mathbf{r}$, generated using the *ADF* package (Baerends *et al.*, 2004). Table 1 reports the Hirshfeld charges and the site-based dipole moment analysis. These charges and dipole moments are important anisotropic properties (Wang *et al.*, 2006) and indications of fragments-in-molecules (Gazquez *et al.*, 2006). Fig. 3 compares the atomic Hirshfeld charges between adenine (dark columns) and purine (light columns). It is apparent that all nitrogen sites exhibit negative charges, indicating that the nitrogen sites can be served as electron donors to accept protons. It is not a surprise that the significant increase in Q^H on the $C_{(6)}$ site of adenine is 'compensated' by a large decrease in Q^H on the $N_{(1)}$ site, which is the nearest N site of the $C_{(6)}$ in the aromatic ring.

Table 1
Hirshfield charges and dipole moments of adenine and purine obtained using the LB94/TZ2P//B3LYP/TZVP model.

	Adenine			Purine		
Site	Q^H	μ_x (D)	μ_y (D)	Q^H	μ_x (D)	μ_y (D)
$N_{(1)}$	−0.22	0.25	0.04	−0.19	−0.03	0.23
$N_{(3)}$	−0.22	0.14	0.02	−0.19	−0.20	−0.11
$N_{(7)}$	−0.22	−0.12	0.07	−0.21	0.23	0.01
$N_{(9)}$	−0.08	−0.20	−0.12	−0.08	0.08	0.11
$N_{(10)}$	−0.16	0.19	0.11			
$C_{(2)}$	0.07	0.01	0.04	0.08	0.05	−0.02
$C_{(4)}$	0.09	−0.22	−0.27	0.10	0.04	0.38
$C_{(5)}$	0.02	0.18	−0.16	0.03	−0.07	0.09
$C_{(6)}$	0.12	0.20	0.28	0.04	0.00	−0.20
$C_{(8)}$	0.07	−0.18	0.04	0.08	0.19	0.07

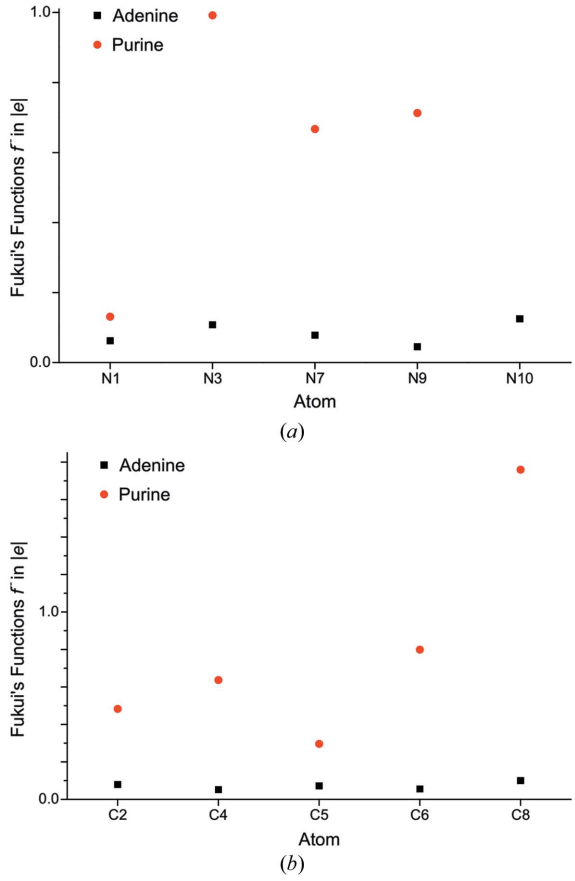

Figure 2
Condensed Fukui functions at the nitrogen and carbon atoms in adenine and purine. (*a*) Condensed Fukui function for the N sites. (*b*) Condensed Fukui function for the C sites.

Table 2
Comparison of the inner-shell vertical ionization energies (eV) of adenine.

Molecular orbital	Site	LB94 TZ2P $-\varepsilon_i$	Mean† TZVP $-\varepsilon_i$	ΔKS‡ TZVP IP	Chong *et al.*§ TZP IP	Obs¶ ADC(4) $-\varepsilon_i$	Dev.††	Dev.‡‡	RASSCF§§ IP	Exp.¶¶ IP
1a′	$N_{(9)}$	405.00	408.36	407.60	406.74	406.7†††	0.0	1.70	406.1	406.7
2a′	$N_{(10)}$	403.91	407.14	406.65	405.86	405.7†††	0.2	1.79	405.9	405.7
3a′	$N_{(7)}$	403.25	406.66	405.52	404.78	405.0	0.2	1.75	404.2	–
4a′	$N_{(3)}$	403.03	406.36	404.87	404.37	404.4	0.0	1.37	403.8	–
5a′	$N_{(1)}$	402.87	293.19	404.78	404.23	404.3	0.1	1.43	403.6	404.4
6a′	$C_{(6)}$	292.05	293.92	292.70	292.64	293.0	0.4	0.95	–	–
7a′	$C_{(8)}$	291.77	293.54	292.31	292.13	292.7	0.5	0.93	–	–
8a′	$C_{(4)}$	291.66	293.40	292.37	292.15	292.5	0.4	0.84	–	–
9a′	$C_{(2)}$	291.23	293.10	291.86	291.66	291.9	0.2	0.67	–	292.5
10a′	$C_{(5)}$	290.57	291.89	291.09	290.96	291.0†††	0.0	0.43	–	291.0

† Mean of B3LYP/TZVP and RHF/TZVP based on *Gaussian03* (Frisch *et al.*, 2004). ‡ BP functional (Becke, 1988; Perdew, 1986) in *StoBe-deMon* (Hermann *et al.*, 2005). § CEBEs calculated by the method of Chong *et al.* (Takahata & Chong, 2003; Takahata *et al.*, 2006) based on *ADF* (Baerends *et al.*, 2004). ¶ Deviation of CEBEs of Chong *et al.* from ADC(4) energies. †† Deviation of LB94/TZ2P energies from ADC(4) energies. ‡‡ *Ab initio* RASSCF calculations for planar adenine (12 × 12-RAS) (Mochizuki *et al.*, 2001) for N-*K*. §§ XPS experiment at Elettra Synchrotron laboratory (Plekan *et al.*, 2007*b*). ¶¶ *Ab initio* ADC(4) calculation with basis set 6-31G (Plekan *et al.*, 2007*b*). ††† Observed XPS energies of the respective band maxima (Plekan *et al.*, 2007*b*).

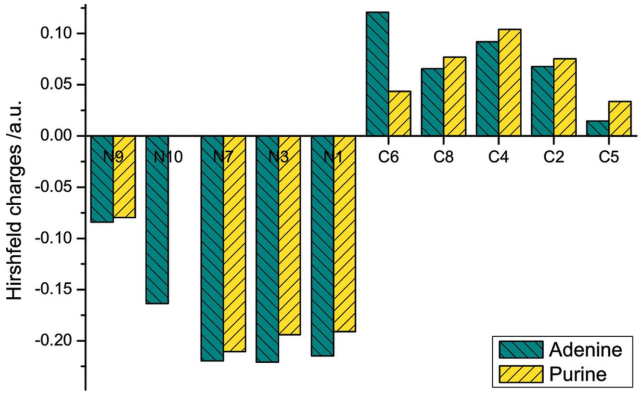

Figure 3
Comparison of atomic Hirshfeld charges of adenine (dark) and purine (light).

Table 3
Comparison of the inner-shell vertical ionization energies (eV) of purine.

	Site	LB94† TZ2P $-\varepsilon_i$	Mean‡ TZVP $-\varepsilon_i$	ΔKS§ TZVP IP	Chong *et al.*¶ TZP IP	(Chong *et al.* LB94)†† IP-$(-\varepsilon_i)$
1a′	$N_{(9)}$	405.34	408.73	407.42	407.16	1.82
2a′	$N_{(3)}$	403.63	40.708	405.53	405.19	1.56
3a′	$N_{(7)}$	403.49	406.92	405.41	405.09	1.60
4a′	$N_{(1)}$	403.32	406.76	405.19	404.84	1.52
5a′	$C_{(4)}$	292.06	294.01	293.10	292.73	0.67
6a′	$C_{(8)}$	292.13	293.89	292.99	292.67	0.54
7a′	$C_{(2)}$	291.52	293.88	292.44	292.09	0.57
8a′	$C_{(6)}$	291.13	293.43	292.07	291.71	0.58
9a′	$C_{(5)}$	290.95	292.44	291.93	291.58	0.63

† LB94 functional (van Leeuwen & Baerends, 1994) in *ADF* package (Baerends *et al.*, 2004). ‡ Mean of B3LYP/TZVP and RHF/TZVP based on *Gaussian03* (Frisch *et al.*, 2004). § BP functional (Becke, 1988; Perdew, 1986) in *StoBe-deMon* (Hermann *et al.*, 2005). ¶ CEBEs calculated by the method of Chong *et al.* (Takahata & Chong, 2003; Takahata *et al.*, 2006) based on *ADF* (Baerends *et al.*, 2004). †† Deviation of LB94/TZ2P energies from CEBEs of Chong *et al.*

Tables 2 and 3 compare the inner-shell ionization potentials (IPs) for adenine and purine site-by-site, using various theoretical methods, such as the LB94 (Van Leeuwen & Baerends, 1994), ΔKohn–Sham BP/TZVP (*StoBe*) (Hermann *et al.*, 2005), methods developed by Chong *et al.* for CEBEs (Cavigliasso & Chong, 1999; Takahata & Chong, 2003; Takahata *et al.*, 2006), the mean of the B3LYP/TZVP and RHF/TZVP models, *ab initio* IPs (Mochizuki *et al.*, 2001) using the RASSCF method, *ab initio* calculation using fourth-order algebraic diagrammatic construction ADC(4) method (Plekan *et al.*, 2007*b*) and the recently experimentally observed XPS. The core orbital energies obtained from the LB94 DFT functional were found to generate good estimated IPs for the C1*s* orbitals of camphor enantiomers (Stener *et al.*, 2006) using the meta-Koopman's theorem (Chong, 2005). Our ΔKohn–Sham BP/TZVP (*StoBe*) calculations use the same model in the study of XAS spectra for the DNA bases (MacNaughton *et al.*, 2005). The present CEBEs obtained for purine employ the same method as that for adenine (Takahata *et al.*, 2006). The simple mean IPs generated from the B3LYP and RHF orbital energies do not possess rigorous scientific basis; however, it was found that in the case of L-alanine (Powis *et al.*, 2003) the experimental values are rather symmetrically bracketed by the pair estimates and the relative splittings are in close accord with experiment.

CEBEs calculated using the method of Chong *et al.* show an average absolute deviation of 0.2 eV for adenine from the observed or best-estimated IPs based on ADC(4) (Plekan *et al.*, 2007*b*) (see Table 2). It should be mentioned that the ADC(4) energies have been corrected by 1.27 and 1.32 eV for C1*s* and N1*s* binding energies, respectively, to compensate for a uniform error of the theoretical approach mainly caused by the use of a smaller basis set in the ADC(4) method.

The IPs produced using the present model (LB94/TZ2P) are the results of a simple single calculation in accordance with the meta-Koopman's theorem, without taking account of a number of effects such as orbital relaxation, relativistic energy corrections (Chong, 2005) and systematic energy error corrections (Plekan *et al.*, 2007*b*) *etc.* For example, to obtain the CEBEs using the method of Chong *et al.*, one needs to perform calculations for $N + 1$ states which include the individual site-specific ionized core hole state and the ground

electronic state, whereas the present model calculates only the ground electronic state. In addition, the present model does not consider the empirical relativistic energy corrections which, according to Chong *et al.*, amount to 0.1 eV and 0.05 eV for N and C, respectively (Takahata *et al.*, 2006). The deviations of the core IPs of adenine calculated using the LB94/TZ2P model from the IPs based on the ADC(4) method are also presented in Table 2. Thus the site-dependent IPs generated using the present model can be considered as a good estimation over those higher-level calculations such as the method of Chong *et al.* and ADC(4).

Fig. 4 provides a clear correlation diagram of the relative core IPs of the N1*s* and C1*s* sites between purine and adenine, based on our LB94/TZ2P calculations. All N-*K* and C-*K* sites within purine and adenine are non-equivalent so that the binding energies are site dependent with small splits. Together with Figs. 5 and 6, which provide the N1*s* and C1*s* core level density of state spectra, respectively, a comprehensive picture correlating the core spectra and the electronic structures is presented. From the N1*s* spectra of adenine and purine, certain similarities lead to the anticipation of a correlation between the corresponding N1*s* sites of purine and adenine.

In Fig. 5, it is quite clear that the insertion of a peak at approximately 404 eV in the adenine N1*s* spectra (upper) is due to the additional amino group, $-NH_2$, which contributes to narrow the energy gap between the amino N1*s* (–N–) sites and the imino N1*s* (–N=) sites in adenine. A small relaxation of other peaks corresponding to the N1*s* sites towards the higher energy end in the purine ring of adenine also exist, which causes the spectral line position exchange of $N_{(3)}$ and $N_{(7)}$, when adenine is formed from purine.

The N1*s* binding energy shift in purine and adenine is of interest. Unlike the C sites, the N sites in both purine and adenine exhibit a similar pattern with a bond-type dependency. For example, the N1*s* binding energies in purine are clearly separated between $N_{(9)}$ and the other N sites of $N_{(3)}$, $N_{(7)}$ and $N_{(1)}$. In adenine spectra, the peak at 403.91 eV of the amino group is inserted between the peaks of $N_{(9)}$ and $N_{(7)}$. As demonstrated in Fig. 4, the N1*s* binding energies of purine can be grouped into the imino N atoms of $N_{(1)}$, $N_{(7)}$ and $N_{(3)}$, which are associated with double N═C bonds, and the amino nitrogen $N_{(9)}$ which connects with a single N—C bond. The amino and imino N1*s* sites are separated in the binding energy spectrum of purine by an energy gap of 1.87 eV. The amino–

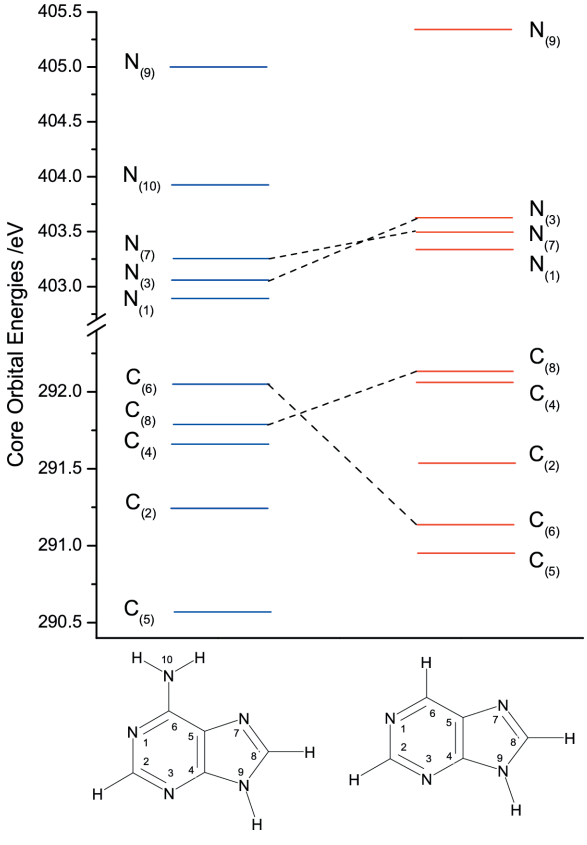

Figure 4
Energy correlation diagram of adenine and purine in the inner shell.

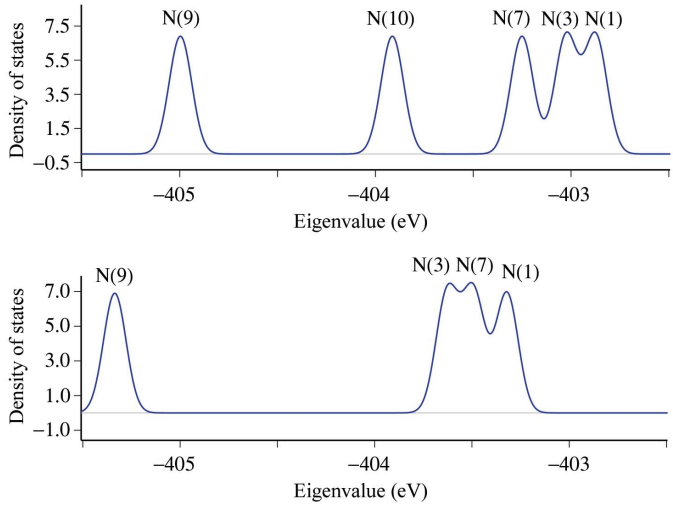

Figure 5
N-*K* ionization energy spectra of adenine (upper) and purine (lower).

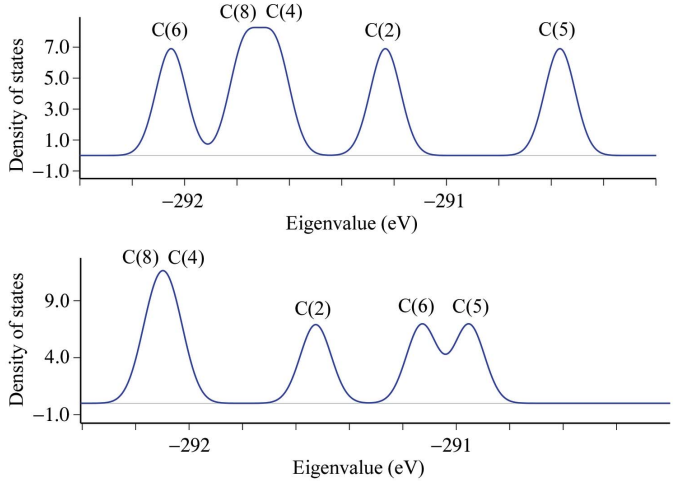

Figure 6
C-*K* ionization spectra of adenine (upper) and purine (lower).

imino energy gap of adenine reduces to 1.13 eV owing to the additional N site of the amino group, $N_{(10)}$. It results in the two peaks in the XAS spectra of adenine as observed by Harada *et al.* (2006).

A significant C1*s* spectral change is observed when adenine is formed from unsubstituted purine, even though the number of carbon atoms is conserved in purine and adenine in the same geometric position, as shown in Fig. 6. The order of the spectral peaks of $C_{(8)}$, $C_{(4)}$ and $C_{(2)}$ in adenine and purine indicates that the species are somehow related. However, the striking differences in the C1*s* spectra reveal that the electronic structure and therefore properties of adenine and purine are significantly different, which largely results from their C1*s* configuration. In Fig. 6 the close 'twin peaks' resulting from $C_{(6)}$ and $C_{(5)}$ in purine at low IP are split into the higher [$C_{(6)}$] and lower [$C_{(5)}$] extremes of the C1*s* spectra of adenine, as a response to the $-NH_2$ attachment. Therefore, the C1*s* spectra may yield the key information differentiating purine and adenine.

The largest binding energy split for the C1*s* sites of purine is between $C_{(8)}$ and $C_{(5)}$, whose core IPs appear at both ends [upper for $C_{(8)}$ and lower for $C_{(5)}$] with an energy gap of 1.17 eV based on the LB94/TZ2P calculation. The largest energy split in adenine is between $C_{(6)}$, which moves up to the top of the energy band after the addition, and $C_{(5)}$, which drops down. The carbon core ionization energy band in adenine spreads more extensively in energy (bandwidth of the C sites increases to 1.61 eV) than that in purine. A similar trend is also observed in the N sites of the species. The most significant binding-energy change in the C1*s* sites is the large energy blue-shift (higher IPs) for the $C_{(6)}$ site in adenine, while all of the other C1*s* sites exhibit red-shifts (lower IPs), owing to the attachment of the amino fragment at the $C_{(6)}$ site. As a result, the C1*s* binding-energy blue-shift at the $C_{(6)}$ site can be a signature of the amino fragment attachment.

Fig. 7 further indicates the 'anti-correlation' between the Hirshfeld charge difference and the inner-shell chemical shift. It also indicates that the corresponding atomic site IP changes of adenine with respect to purine are due to the electronic charge redistributions. All atomic sites in adenine exhibit various degrees of blue-shift (energy increase) in IPs compared with the corresponding sites in purine, except for the $C_{(6)}$ site, which indicates a red-shift in IPs [see Fig. 7(lower)]. The IP shift in the opposite direction on the $C_{(6)}$ site is because of its bond with the amino fragment. As a result, this site shows a substantial increase in positive-charge concentration in adenine as indicated in Fig. 7(upper).

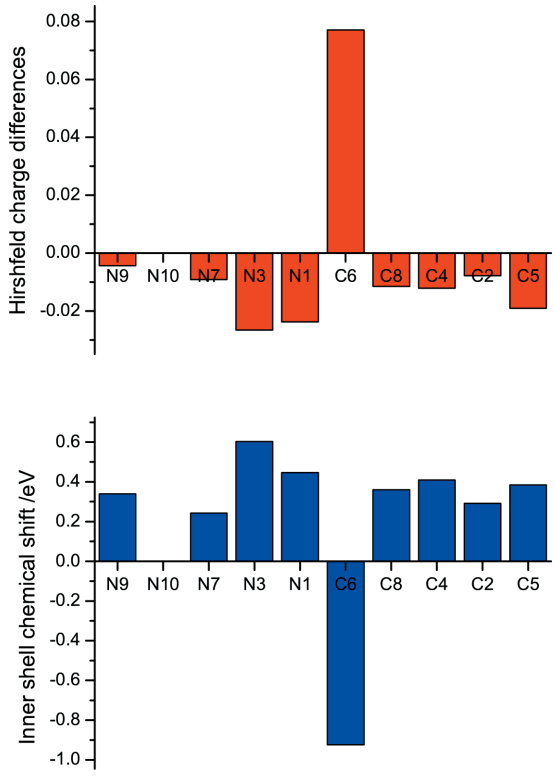

Figure 7
Changes in Hirshfeld charges (upper) and chemical shift in the inner shell (lower) of adenine and purine (Δ = adenine − purine).

4. Conclusion

The results of the present study indicate a significant inner-shell electronic structural change in response to the addition of the amino fragment ($-NH_2$) to the unsubstituted purine molecule in the binding energy spectra, in particular the C1*s* spectra. All the vertical ionization energies related to the C1*s* sites exhibit red-shift (lower ionization energies) in adenine in comparison with the C1*s* of purine, except for $C_{(6)}$, which demonstrates a significant blue-shift to accommodate the attachment of $-NH_2$. While the C1*s* orbitals' response to the attachment of amino fragment is site specific, the N1*s* orbitals show dependencies on the bonding mechanism, which exhibit a red-shift in binding energy. Moreover, the amino–imino nitrogen separation in the binding-energy spectrum of purine has been reduced by more than a half in the energy owing to the insertion of the amino $N_{(10)}$ site. Results from the core binding energy spectra of the DNA base pair agree well with the calculated Hirshfeld charge distribution of the molecules.

The present study also indicates that the degree of electronic structural changes of purine and adenine does not depend on the numbers of the element (*e.g.* N1*s*) alone, nor on the geometric similarities, but on their electron charge distribution. As a result, studies of the inner-shell X-ray spectra of these DNA bases need sufficiently accurate quantum mechanical support to assist with the correct assignment of spectral features. The C1*s* and N1*s* spectra of purine and adenine indicate that some fragment-based methods may not be applicable to spectral analyses in the inner shell of DNA/RNA bases.

This work is partly supported by a Vice-Chancellor's Strategic Research Initiative Grant, the Australian Research Council (ARC) and the Natural Sciences and Engineering Research Council (NSERC) of Canada. SS and FW acknowledge the Australian Partnership for Advanced

Computing (APAC) for use of the National Supercomputing Facilities.

References

Abu-samha, M., Borve, K., Harnes, J. & Bergersen, H. (2007). *J. Phys. Chem. A*, **111**, 8903–8909.

Baerends, E. J. *et al.* (2004). *ADF2004.01*, http://www.scm.com/.

Barber, M. & Clark, D. T. (1970). *Chem. Commun.* p. 23–24.

Becke, A. D. (1988). *Phys. Rev. A*, **38**, 3098–3100.

Bickelhaupt, F. M. & Baerends, E. J. (2000). *Theor. Chim. Acta*, **15**, 1–86.

Cavigliasso, G. & Chong, D. P. (1999). *J. Chem. Phys.* **111**, 9485–9492.

Chong, D. P. (1995). *J. Chem. Phys.* **103**, 1842–1845.

Chong, D. P. (2005). *J. Electron Spectrosc. Relat. Phenom.* **148**, 115–121.

Dong, F. & Miller, R. E. (2002). *Science*, **298**, 1227–1230.

Downton, M. & Wang, F. (2006). *Mol. Simul.* **32**, 667–673.

Frisch, M. J. *et al.* (2004). *Gaussian03*, revision C.02. Gaussian, Wallingford, CT, USA.

Fujii, K., Akamatsu, K. & Tokoya, A. (2004). *J. Phys. Chem. B*, **108**, 8031–8035.

Fukui, K. (1982). *Science*, **218**, 747–784.

Gazquez, J. L., Cedillo, A., Gomez, B. & Vela, A. (2006). *J. Phys. Chem. A*, **110**, 4535–4537.

Godbout, N., Salahub, D. R., Andzelm, J. & Wimmer, E. (1992). *Can. J. Chem.* **70**, 560–571.

Guerra, C. F., Bickelhaupt, F. M., Saha, S. & Wang, F. (2006). *J. Phys. Chem. A*, **110**, 4012–4020.

Harada, Y., Takeuchi, T., Kino, H., Fukushima, A., Takaura, K., Hieda, K., Nakao, A., Shin, S. & Fukuyama, H. (2006). *J. Phys. Chem. A*, **110**, 13227–13231.

Hermann, K. *et al.* (2005). *Stobe-demon*, version 2.1. http://w3.rz-berlin.mpg.de/~hermann/StoBe/index.html.

Hirshfeld, F. L. (1977). *Theor. Chim. Acta*, **44**, 129–138.

Jolly, W., Bomben, K. & Eyermann, C. J. (1984). *Atom. Data Nucl. Data Tables*, **31**, 433–493.

Karlsen, T., Borve, K., Saerthre, L. J., Wiesner, K., Bassler, M. & Svensson, S. (2002). *J. Am. Chem. Soc.* **124**, 7866–7873.

Kolczewski, C., Puttuer, C., Plashkerych, O., Agren, H., Staemmler, V., Martins, M., Snell, G., Schlachter, A. S., SantAnna, M., Kaindl, G. & Pettersson, L. G. M. (2001). *J. Chem. Phys.* **115**, 6426–6437.

Leeuwen, R. van & Baerends, E. J. (1994). *Phys. Rev. A*, **49**, 2421–2431.

Lenthe, E. van & Baerends, E. J. (2003). *J. Comput. Chem.* **24**, 1142–1156.

MacNaughton, J., Moewes, A. & Kurmaev, E. Z. (2005). *J. Phys. Chem. B*, **109**, 7749–7757.

Mochizuki, Y., Koide, H., Imamura, T. & Takemiya, H. (2001). *J. Synchrotron Rad.* **8**, 1003–1005.

Nakata, A., Imamura, Y. & Nakai, H. (2006). *J. Chem. Phys.* **125**, 064109.

Perdew, J. P. (1986). *Phys. Rev. B*, **33**, 8822–8824.

Plekan, O., Feyer, V., Richter, R., Coreno, M., de Simone, M., Prince, K. C. & Carravetta, V. (2007*a*). *J. Electron Spectrosc. Relat. Phenom.* **155**, 47–53.

Plekan, O., Feyer, V., Richter, R., Coreno, M., de Simone, M., Prince, K. C., Trofimov, A. B., Gromov, E. V., Zaytseva, I. L. & Schirmer, J. (2007*b*). *Chem. Phys.* In the press.

Powis, I., Rennie, E. E., Hergenhahn, U., Kugeler, O. & Bussy-Socrate, R. (2003). *J. Phys. Chem. A*, **107**, 25–34.

Rehr, J. J. & Ankudinov, A. L. (2001). *J. Synchrotron Rad.* **8**, 61–65.

Saha, S., Wang, F. & Brunger, M. J. (2006). *Mol. Simul.* **32**, 1261–1270.

Saha, S., Wang, F., Falzon, C. T. & Brunger, M. J. (2005). *J. Chem. Phys.* **123**, 124315.

Sanche, L. (2003). *Mass Spectrom. Rev.* **21**, 349–369.

Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S. J., Windus, T. L., Dupuis, M. & Montgomery, J. A. (1993). *J. Comput. Chem.* **14**, 1347–1363.

Siegbahn, K., Nordling, C., Johansson, G., Hedman, J., Hedén, P. F., Hamrin, K., Gelius, U., Bergmark, T., Werme, L. O., Manne, R. & Baer, Y. (1969). *ESCA Applied to Free Molecules.* Amsterdam: North-Holland.

Sponer, J. & Hobza, P. (2003). *Collec. Czech. Chem. Commun.* **68**, 2231–2282.

Stener, M., Tommaso, D. D., Fronzoni, G. & Decleva, P. (2006). *J. Chem. Phys.* **124**, 024326.

Svensson, S. (2005). *J. Phys. B*, **38**, S821–S838.

Takahata, Y. & Chong, D. P. (2003). *J. Electron Spectrosc. Relat. Phenom.* **133**, 69–76.

Takahata, Y., Okamoto, A. K. & Chong, D. P. (2006). *Int. J. Quant. Chem.* **106**, 2581–2586.

Triguero, L., Pettersson, L. G. M. & Agren, H. (1998). *Phys. Rev. B*, **58**, 8097–8110.

Triguero, L., Plashkevych, O., Pettersson, L. G. M. & Agren, H. (1999). *J. Electron Spectrosc. Relat. Phenom.* **104**, 195–207.

Wang, F. (2005). *J. Mol. Struct. (Theochem)*, **728**, 31–42.

Wang, F., Downton, M. & Kidwani, N. (2005). *J. Theor. Comput. Chem.* **4**, 247–264.

Wang, F., Pang, W. N. & Huang, M. (2006). *J. Electron Spectrosc. Relat. Phenom.* **151**, 215–223.