Carbon-14 decay as a source of non-canonical bases in DNA

Michel Sassi,[1] Damien J. Carter,[1] Blas P. Uberuaga,[2] Chris R. Stanek,[2] and Nigel A. Marks[1, 3, *]

[1]*Nanochemistry Research Institute, Curtin University, P.O. Box U1987, Perth, WA 6845, Australia*
[2]*Material Science and Technology Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[3]*Discipline of Physics & Astronomy, Curtin University, P.O. Box U1987, Perth, WA 6845, Australia*

Background: Significant experimental effort has been applied to study radioactive beta-decay in biological systems. Atomic-scale knowledge of this transmutation process is lacking due to the absence of computer simulations. Carbon-14 is an important beta-emitter, being ubiquitous in the environment and an intrinsic part of the genetic code. Over a lifetime, around 50 billion ^{14}C decays occur within human DNA.

Methods: We apply *ab initio* molecular dynamics to quantify ^{14}C-induced bond rupture in a variety of organic molecules, including DNA base pairs.

Results: We show that double bonds and ring structures confer radiation resistance. These features, present in the canonical bases of the DNA, enhance their resistance to ^{14}C-induced bond-breaking. In contrast, the sugar group of the DNA and RNA backbone is vulnerable to single-strand breaking. We also show that Carbon-14 decay provides a mechanism for creating mutagenic wobble-type mispairs.

Conclusions: The observation that DNA has a resistance to natural radioactivity has not previously been recognized. We show that ^{14}C decay can be a source for generating non-canonical bases.

General Significance: Our findings raise questions such as how the genetic apparatus deals with the appearance of an extra nitrogen in the canonical bases. It is not obvious whether or not the DNA repair mechanism detects this modification nor how DNA replication is affected by a non-canonical nucleobase. Accordingly, ^{14}C may prove to be a source of genetic alteration that is impossible to avoid due to the universal presence of radiocarbon in the environment.

Keywords: Carbon-14; beta-decay; DNA; bond-breaking; non-canonical bases; mutation

INTRODUCTION

The potential for radioactive Carbon-14 to have biological effects has long been recognized [1]. All the same, it is an empirical fact that in the human body is somehow capable of enduring 20–30 Carbon-14 decays per second within the base pairs and sugar groups of DNA [2]. The origin of this robustness has remained elusive, primarily due to the difficulty of separating the local effects of chemical transmutation from long-range stochastic effects associated with β-particle emission [3]. The local perturbation imposed by radioactive decay at first seems insurmountable for a biological molecule as four different atomic-scale disruptions arise from ^{14}C decay: the abrupt appearance of a nitrogen, the creation of a positively charged cation, secondary electron shake-off due to Coulombic excitation and physical recoil via momentum conservation. All of these effects have the potential to break bonds, create new chemical environments and generate molecules which do not belong to the genetic "alphabet".

Initial experiments showed the potential for ^{14}C→^{14}N transmutation to occur without necessarily breaking chemical bonds. Charge-spectrometry experiments involving radioactive carbon dioxide (^{14}CO$_2$) was found to yield stable ^{14}NO$^+$ molecule in 81% of cases [4], while in doubly-labelled ethane (H$_3$^{14}C–^{14}CH$_3$) the chance of bond-rupture was about 53% [5]. In methylbenzene (toluene) transmutation generated about 3% of C–N bond breaking, while in [α-^{14}C]ethylbenzene 15-20% of

bond rupture occurred [6]. These experiments provide a valuable reference point for validating the computational methods applied in this study.

From the early 1960's the impact of beta-decay on biomolecules attracted considerable experimental attention [see Refs. 3, 7–9 for reviews]. Many hundreds of radiolabelling experiments have since been performed and a large database of information has been collected on the biological response of viruses, bacteria and higher organisms to radioisotopes. While ^{3}H, ^{32}P and ^{125}I have been extensively studied, the number of biological experiments involving ^{14}C is relatively small due to the low specific activity associated with the long half-life (5730 years). When ^{14}C is incorporated into the DNA of *E. coli*, its decay induced lethalities with the principal cause being the transmutation rather than the radiation associated with the emitted β-particle [10]. Experiments on fruit fly sperm cells found that ^{14}C within DNA did not produce a detectable genetic effect [11] while similar experiments using ^{32}P produced mutations [12, 13]. Notably, these experiments eliminated stochastic β-irradiation effects via dilution. In Chinese hamster cells ^{14}C caused single-strand DNA breaks [14], but interpretation of the mechanism was hampered by difficulties in separating the transmutation and β-irradiation components. This is a general problem which plagues many experimental studies as discussed in Refs. 3 and 9.

The fundamental processes involved in ^{14}C beta-decay are summarized schematically in Figure 1a. The β-particle receives a maximum energy of 157 keV and

travels around 0.1 mm away [9] from the ^{14}N nucleus while the physical and chemical effects of the transmuted ^{14}N have effects on the nanometre scale. While Monte Carlo methods [15–18] are routinely applied to compute the stochastic effects induced by β-particles in biological systems, the local effects due to transmutation have not been studied using atomistic simulation. Indeed, aside from calculations of beta-decay in solids [19–21], and some early work on decay-induced excited states in molecules [22, 23], we are unaware of any remotely similar studies.

Figure 1b provides the starting point of our computational studies of transmutation. The blue curve shows the relative probability that the ^{14}N daughter recoils with a given energy, while the red curve shows the cumulative probability. The probability distribution is computed using the methodology in Ref. 21 and is based on the ^{14}C β-particle spectrum and the assumption of an uniform angular distribution between the β-particle and anti-neutrino. Figure 1b shows that the maximum recoil of the ^{14}N daughter is 7 eV and that 50% of the recoils have an energy below 2.4 eV. Given that bond enthalpies for carbon-carbon, carbon-nitrogen and carbon-hydrogen are typically 3–5 eV, this appears to suggest that ^{14}N recoil is capable of breaking bonds around 10–40% of the time. In practice we will show that the real situation is rather more subtle, with some molecules being highly resistant, while others are relatively easily broken.

In this work we employ *ab initio* molecular dynamics to study ^{14}C transmutation in DNA. By following the time evolution of recoiling ^{14}N species we study the dynamics of bond breaking and quantify the various factors which control it. Surprisingly, we find that many of the factors mitigating against ^{14}C-induced bond breaking are present in DNA. This observation suggests that non-canonical bases can be generated in DNA through ^{14}C transmutation.

METHODS

To study the recoil of the ^{14}N daughter, molecular dynamics simulations were employed using the VASP program [24, 25]. The majority of the calculations were carried out in the gas phase with a vacuum separation of at least 12 Å between the molecule and its periodic image. Simulations were performed in an NVE ensemble, using small timesteps (≤ 0.025 fs) and a tight energy convergence criteria (10^{-6} eV). The Perdew and Wang [26] exchange-correlation functional was used and the cutoff energies of the projector augmented wave [27] pseudopotentials were 400 eV. To correctly describe bond breaking, spin-polarization and the Vosko–Wilk–Nusair scheme [28] were used. Most calculations used a supercell with a net charge of $+1e$, reflecting the physics of beta-decay whereby the β-particle travels far from the

^{14}N nucleus (up to 0.1 mm) [9] leaving behind a cation. Some calculations used a supercell of net charge $+2e$ or $+3e$ to mimic shake-off which perturbs the electronic shells giving rise to possible additional ionization. Bader charges were computed to assess the spatial distribution of the electron density using the methodology developed by Henkelman *et al.* [29]. In this method the Bader partitioning is performed using the total of the valence and core charge. Once the Bader regions are determined, integration is performed using the valence charge. In calculations involving RNA nucleotides and DNA nucleosides two atoms at the extremity were held fixed to prevent net movement of the entire fragment. In the simulation of gas-phase molecules the linear and angular momentum components were removed and all atoms were free to move. Typical calculations followed 0.2 ps of motion, with the longest simulations running for 0.35 ps.

To capture the effect of the aqueous environment, we performed additional simulations in which an adenine-thymine nucleoside was surrounded by a complete first hydration shell and a partial second hydration shell. The arrangement of water molecules around the DNA structure, especially around the bases, has been widely investigated experimentally and theoretically [30–33]. Previous
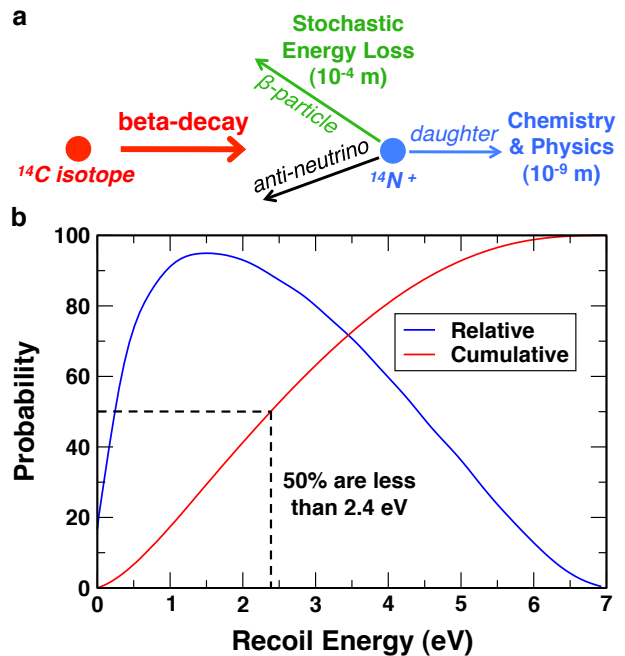

FIG. 1. Radioactive beta-decay for the carbon-14 isotope. (a) Schematic of the processes involved in radioactive beta-decay of ^{14}C. (b) Recoil energy spectrum (blue line) of the ^{14}N daughter nucleus following ^{14}C beta-decay. The spectrum was derived from the energy spectrum of the emitted β-particle and presumes an isotropic angular distribution between the β-particle and the anti-neutrino. The red line is the corresponding cumulative probability distribution, from which it can be seen that 50% of the ^{14}N recoils are below 2.4 eV.

molecular dynamics simulations have shown that water molecules around the canonical bases are structured in three hydration shells [31], while analysis of DNA crystal structure data showed that the first hydration shell is local and the position of water molecules can be predicted [30]. We used these predicted hydration sites for the first shell to set up an initial configuration of water molecules around the adenine and thymine bases. The locations of the water molecules for the second hydration shell were extracted from the BDJ025 crystal structure [34] for an adenine-thymine base pair. As the second hydration shell extends from 4 to 7 Å [31], we removed all of the water oxygen atoms with a distance greater than 7 Å from the canonical bases.

The structural optimization of the hydrated adenine-thymine nucleoside was performed in three steps. During the first step all the atoms were fixed except the hydrogen and oxygen atoms constituting the second hydration shell. In the second step, all of the water molecules (i.e. both hydration shells) were allowed to relax but the structure of the nucleoside was kept fixed. For the final step, we performed a full relaxation of every atom. During the *ab initio* molecular dynamic simulations, all atoms were allowed to move except the water oxygen atoms of the second hydration shell.

RESULTS

Ab initio molecular dynamics has not previously been applied to beta-decay in molecules and so we introduce our computational methodology via a series of steps. First we study two prototypical molecules, radioactive methane and carbon dioxide, and examine the threshold recoil energy (E_T) required to break bonds. Next we consider a set of hydrocarbon and alcohol molecules with an increasing degree of structural complexity. These studies highlight the structural features that provide resistance to ^{14}C decay and validate our methodology where experimental data is available. Finally, we consider the effect of ^{14}C decay for carbon atoms present in the ribose group of RNA and within the four canonical base pairs in DNA.

Decay in ^{14}CH$_4$

One of the simplest conceptual systems is the beta-decay of radioactive methane to ammonium (Figure 2a). We investigated three different recoil directions, as indicated by the arrows in Figure 2b. For direction (1), the recoil is initiated along the N–H bond and in the direction of the hydrogen, while direction (2) has the opposite displacement. Along direction (3) the recoil is initiated towards the midpoint of two hydrogen atoms. Figure 2c shows the time evolution of the N–H distance for a recoil in direction (1) and energies above and below the
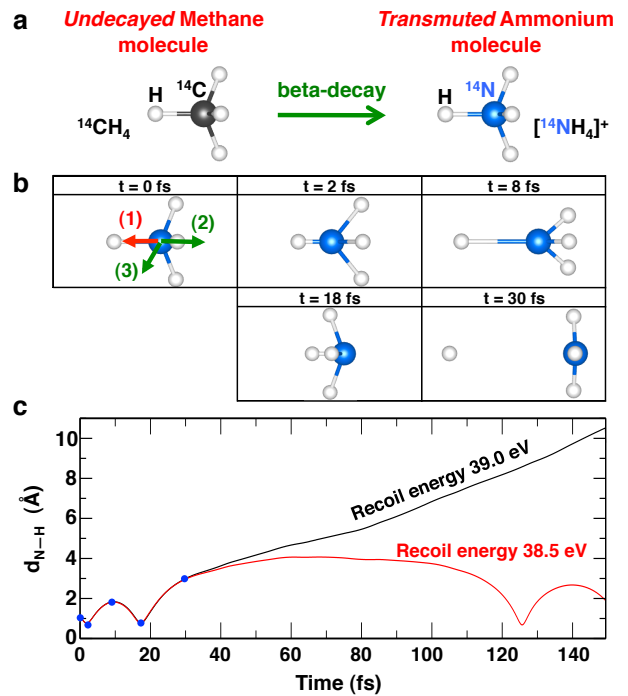

FIG. 2. Transmutation and effect of recoil in the CH$_4$ molecule. (a) Ball-and-stick model of the radioactive (parent) methane molecule and its transmuted (daughter) form following ^{14}C beta-decay. The carbon-14 in methane and the daughter cation nucleus in ammonium are shown as black and blue spheres respectively. (b) Evolution of the molecular structure over time, after a recoil is directed along the N–H bond, shown by the red arrow. Green arrows show the two other directions investigated. (c) N–H distance as function of time for recoil energies above and below the threshold value of 38.7 eV. The blue dots highlight the position of the snapshots shown in (b).

threshold value of E_T=38.7 eV. The blue dots in the figure corresponding to successive snapshots shown in Figure 2b. In Figure 2c, we can see that a recoil energy of 39 eV leads to a bond breaking, while for a recoil energy of 38.5 eV the hydrogen atom moves as far as 4 Å away from the NH$_3$ group before returning to reform the ammonium molecule. Threshold energies for breaking an N–H bond along directions (2) and (3) (green arrows) were computed to be E_T=34 eV and E_T=42 eV respectively.

These high values of the threshold energy have two origins: the first is due to the degrees of freedom of the molecule, indeed, when the recoil is initiated, molecular translations and rotations leave less energy to break the chemical bond. The internal kinetic energy available (E_{int}) to break bonds can be deduced by the simple kinematic formula:

$$E_{int} = \frac{M_{tot} - M_{rec}}{M_{tot}} \times E_{rec}, \qquad (1)$$

where M_{tot}, M_{rec} and E_{rec} are respectively the mass of

the molecule, the mass of the recoiling atom and its recoil energy. For directions (1), (2) and (3), the internal kinetic energies corresponding to the computed E_T values are 8.6 eV, 7.5 eV and 9.3 eV respectively. The second aspect is induced by the electron that travels far away from the local environment of the daughter molecule. A positive charge is then associated with the conversion of the neutron to a proton leading to a daughter ^{14}N iso-electronic with ^{14}C. However, extra ionization of the daughter ^{14}N can also be obtained via shake-off. To mimic this additional loss of electrons, we performed simulations of recoil for a $[NH_4]^{2+}$ molecule. For this system, the threshold energy drops below 0.1 eV indicating that $[NH_4]^{2+}$ cannot survive recoil and always breaks into NH_3 and H fragments. Table I lists the E_T values obtained for $[NH_4]^+$ and $[NH_4]^{2+}$. Unphysical simulations employing a neutral ^{14}N find that a recoil energy of 7 eV is already enough to break a N–H bond. Although experiments have not been performed, we predict that beta-decay in methane would result in ammonium daughters, and if shake-off is involved, NH_3 and H fragments would also be detected.

Decay in $^{14}CO_2$

Our study of bond breaking in $^{14}CO_2$ is supported by charge spectrometry experiments [4] in which a variety of charged decay products have been detected, the most common being $[NO_2]^+$ with an abundance of 81%. We first studied the singly ionized daughter, i.e. $[NO_2]^+$, corresponding to the case of no shake-off. We performed simulations for 11 different angles at the maximum recoil energy of 7 eV and bond-breaking was never observed. This result demonstrates that $[NO_2]^+$ is robust against recoil-induced fragmentation and is consistent with the experimental result in which a high fraction of $[NO_2]^+$ is observed. The experiments also detected NO^+, O^+ and N^+, providing evidence of fragmentation due to shake-off. Our simulations of recoil in $[NO_2]^{2+}$, shown in Figure 3a, and $[NO_2]^{3+}$ confirm this interpretation. The removal of electrons destabilizes the molecule, making it vulnerable to fragmentation due to recoil. A typical reaction in which $[NO_2]^{2+}$ breaks into NO and O is shown in Figure 3b. In this case, Figure 3c shows the time-evolution of the distance between the NO and O fragments, with the blue dots corresponding to successive snapshots shown in Figure 3b. The total energy, potential energy and magnetization are also shown in Figure 3c. The abrupt change in the total energy at 83 fs is associated with a sudden change in the magnetization. Indeed, when the NO–O bond breaks, the magnetization increases by 2 μ_B. This electronic redistribution induces a change in the potential energy, where the drop at 83 fs corresponds to that of the total energy. Bader charge analysis [29, 35] shows that the NO and O fragments each possess a charge of $+1e$,
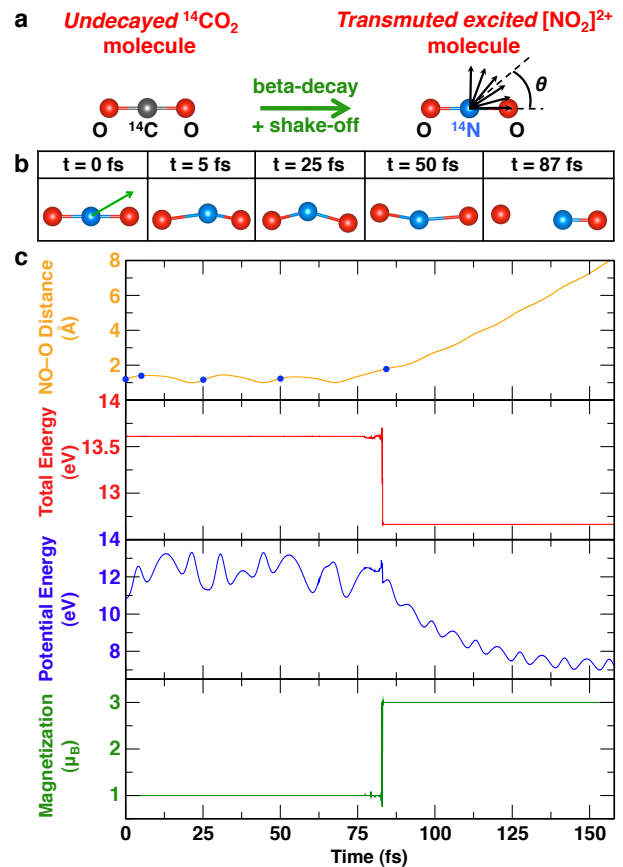

FIG. 3. Transmutation and effect of recoil in the CO_2 molecule. (a) Ball-and-stick model of beta-decay in $^{14}CO_2$ in which shake-off increases the daughter charge from $+1e$ to $+2e$. (b) Time-series of snapshots illustrating bond breaking in $[NO_2]^{2+}$ for $\theta=30°$ and a recoil energy of 4.0 eV. (c) Evolution of the [NO]–O distance, total and potential energies and of the magnetization over the time. The NO–O bond breaking occurs at 83 fs. The blue dots in the curve of [NO]–O distance highlight the position of snapshots shown in (b).

consistent with the experiments. Using a large number of simulations (11 angles and 5 energies) we find a threshold energy to break a N–O bond in $[NO_2]^{2+}$ of 3–4 eV, depending on the angle θ (Figure 3a). Further simulations for $[NO_2]^{3+}$ show that E_T drops below 1 eV. As the molecular fragments obtained for singly- and doubly-ionized molecules account for over 95% of the observed experimental data, we limit our following studies to these two cases.

Hydrocarbon and alcohol molecules

To quantify bond rupture in more complex molecules we simulated recoil in a series of hydrocarbons and alcohols containing single bonds, double bonds and aromatic rings. Table I lists the molecules considered, along with the corresponding daughter molecule, molecular mass

and threshold energy for the case of single- and double-ionization. For reasons of computational cost, recoils were only initiated along a bond axis, either towards a neighboring atom or in the opposite direction. While multiple leaving groups are possible (e.g. NH_3, OH and CH_3), we never observed N–H bond breaking in any of our simulations. The general trend is that doubly ionized molecules have lower threshold energies than singly ionized molecules, indicating that shake-off increases the vulnerability of the daughter. Regardless of the charge (q), ethene and benzene are very stable, with no bond breaking observed for a 7 eV recoil in any direction. This shows that transmutations involving shake-off have a strong impact on parent molecules with only single (σ) bonds, while π bonds play an important stabilizing role. Table I also shows that molecular mass is significant, with lighter molecules having higher thresholds (e.g. ethane vs. butane and methanol vs. ethanol). A ring structure also contributes to stabilization, with cyclohexane being more stable than butane despite a higher molecular mass. These principles are supported by ^{14}C experiments on ethane [5], methylbenzene [6] ($^{14}CH_3$–C_6H_5) and [α-^{14}C]ethylbenzene [6] ($^{14}CH_3$–CH_2–C_6H_5), in which fragmentation was interpreted as a shake-off effect. The experiments also showed that the aromatic ring of benzene increases the stability of the daughter, with fragmentation observed in 3% of cases for methylbenzene and 15–20% of cases for [α-^{14}C]ethylbenzene. Our results on ethane and methyl/ethylbenzene show the same trend, with only the doubly-ionized ethane molecule being vulnerable to fragmentation and with methylbenzene being more stable than [α-^{14}C]ethylbenzene. However, our results on methyl/ethylbenzene also show that the cause of fragmentation is essentially due to the structure of the

Parent	Daughter Molecule	M	$q=+1e$	$q=+2e$
methane	NH_4	18	>7.0	<0.1
ethene	NH_2–CH_2	30	>7.0	>7.0
ethane	NH_3–CH_3	32	>7.0	4.5–4.7
methanol	NH_3–OH	34	>7.0	3.3–3.9
ethanol	NH_3–CH_2–OH	48	4.7–6.5	3.7–5.5
	CH_3–NH_2–OH		5.4–7.0	1.5–2.3
butane	NH_3–$(CH_2)_2$–CH_3	60	3.9–6.4	2.5–4.5
	CH_3–NH_2–CH_2–CH_3		4.1–7.0	1.5–2.7
benzene	C_5H_5NH	80	>7.0	>7.0
cyclohexane	$C_5H_{10}NH_2$	86	5.9–7.0	2.2–4.2
methylbenzene	NH_3–C_6H_5	94	4.3–5.0	4.3–5.0
ethylbenzene	NH_3–CH_2–C_6H_5	108	2.5–3.5	2.5–3.5
	CH_3–NH_2–C_6H_5		4.5–7.0	3.3–7.0

TABLE I. Molecular mass (M in g/mol) and threshold energies (in eV) for singly- and doubly-ionized molecules.

molecule rather than an effect of electronic excitation, as inferred in Ref. 6. Indeed, shake-off has almost no effect on the threshold energies of methyl/ethylbenzene, as shown in Table I, indicating that the benzene ring confers extra resistance to the loss of electron.

DNA base pairs

Next we consider E_T in molecules of adenine, guanine, thymine and cytosine, the four nucleobases of the genetic code. For each carbon site ^{14}N recoils were initiated either along the bond axis or perpendicular to the molecular plane. For the adenine molecule (Figure 4a) we found that both singly- and doubly-ionized molecules are highly resistant to ^{14}C. Only three recoil directions lead to bond breaking, all of which lie in the molecular plane [see green arrows in Figure 4a]. Direction (1) results in an NH_2 leaving group, while directions (2) and (3) lead to irreversible ring opening. Each E_T value was converted to an equivalent probability using the energy spectrum of the β-particle (Figure 1b) and assuming an isotropic angular distribution between the electron and neutrino [3, 9]. It should be noted that these probabilities are an upper limit since every recoil is initiated along the bond axis. Being close to $E_{\max}$, the probability of bond breaking is very small. For [adenine]$^{2+}$ the same three directions led to bond breaking, with very similar thresholds for directions (2) and (3). A lower threshold of 4.5 eV was obtained for recoil direction (1). This shows that single bonds are weakened by transmutation and shake-off, while the purine-type structure of adenine increases the stability of the daughter.

Simulations on guanine (Figure 4e) showed similar behavior to that of adenine. The singly ionized [guanine]$^+$ daughter molecule is highly resistant to ^{14}C decay, as shown by the E_T values in Figure 4f. Only three recoil directions [see green arrows in Figure 4e] led to bond breaking: direction (1) and (3) give an NH_2 and O leaving group, while direction (2) leads to ring opening. As for the [adenine]$^{2+}$ molecule, the E_T value for ring opening in [guanine]$^{2+}$ remains very high, 6.7 eV, while the E_T for bond breaking have a threshold reduced to 4.5–4.7 eV, as shown in Figure 4f.

The thymine molecule is also very resistant to ^{14}C (Figure 4d). The lowest E_T value, 4.7 eV, is obtained for direction (1) in thymine, leading to an NH_3 leaving group and a deprotonated uracil molecule. No ring opening was observed for [thymine]$^+$ under any 7 eV recoils. The thresholds for [thymine]$^{2+}$ show that thymine is more sensitive to shake-off than adenine. However, when the nucleobase is attached to the backbone of RNA or DNA the effect of shake-off is minimal due to the large number of electrons present in the system. Accordingly, most of the remaining calculations are singly-charged.

For the cytosine molecule, shown in Figure 4g, we

Recoil direction	[adenine]+ E_T [eV]	[adenine]+ (Prob.)	[adenine]2+ E_T [eV]	[adenine]2+ (Prob.)
(1)	6.3	(0.4%)	4.5	(12%)
(2)	6.9	(0.0%)	6.5	(0.2%)
(3)	6.5	(0.2%)	6.5	(0.2%)

Recoil direction	[thymine]+ E_T [eV]	[thymine]+ (Prob.)	[thymine]2+ E_T [eV]	[thymine]2+ (Prob.)
(1)	4.7	(10%)	4.7	(10%)
(2)	6.9	(0.0%)	5.4	(4%)
(3)	6.3	(0.4%)	4.5	(12%)
(4)	6.0	(1.2)	4.7	(10%)
(5)	> 7.0	(0.0%)	4.5	(12%)

Recoil direction	[guanine]+ E_T [eV]	[guanine]+ (Prob.)	[guanine]2+ E_T [eV]	[guanine]2+ (Prob.)
(1)	6.3	(0.4%)	4.5	(12%)
(2)	6.7	(0.0%)	6.7	(0.0%)
(3)	6.0	(1.2%)	4.7	(10%)

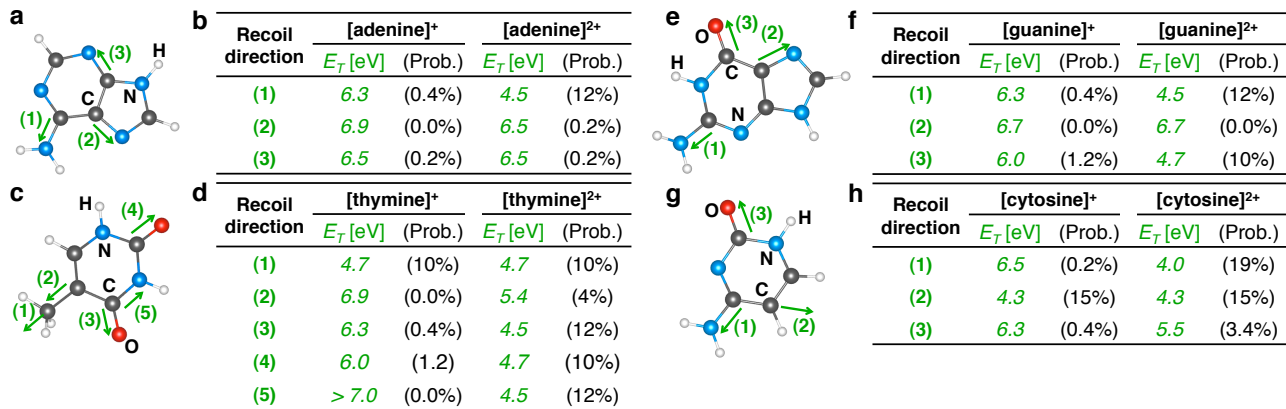| Recoil direction	[cytosine]+ E_T [eV]	[cytosine]+ (Prob.)	[cytosine]2+ E_T [eV]	[cytosine]2+ (Prob.)
(1)	6.5	(0.2%)	4.0	(19%)
(2)	4.3	(15%)	4.3	(15%)
(3)	6.3	(0.4%)	5.5	(3.4%)

FIG. 4. Recoil directions, threshold energies and probabilities in the four DNA base pairs. Recoil directions (green arrows) leading to bond breaking in (a) adenine, (c) thymine, (e) guanine and (g) cytosine molecules in the gas-phase. The tables in (b), (d), (f) and (h) list the threshold energies (E_T) and associated ^{14}C probabilities derived from the cumulative probability spectrum in Figure 1b.

found three recoil directions leading to bond breaking in the transmuted [cytosine]+ and [cytosine]2+ molecules. A recoil in the direction (1) and (3) lead to a NH_2 and O leaving group, while a recoil along the direction (2) lead to ring opening. The corresponding E_T values are listed in Figure 4h.

Interestingly, the E_T value for ring opening in doubly ionized adenine [direction (2) and (3)], guanine [direction (2)] and cytosine [direction (2)] is very similar to that of their respective singly ionized molecules, suggesting that ring opening is not very sensitive to shake-off. Regarding the NH_2 and O leaving groups, common characteristics across the four bases can be observed. For the singly ionized molecules, the E_T value for an NH_2 leaving group in adenine, guanine and cytosine is about 6.3–6.5 eV, while for the doubly ionized molecules, E_T drops to 4.0–4.5 eV. The E_T value for creating an O leaving atom in singly ionized thymine, guanine and cytosine is about 6.0–6.3 eV, while for the doubly ionized molecules E_T is mainly around 4.5–4.7 eV, with a slightly higher value of 5.5 eV for [cytosine]2+ (Figure 4h).

Ribose in an RNA nucleotide

The effect of ^{14}C on the ribose ring of the RNA backbone was investigated for five different recoil directions, as shown in Figure 5. We found that ^{14}C can lead to serious fragmentation: e.g. rupture of the backbone, ring opening and loss of the adenine–backbone link. With E_T in the range 2.7–4.7 eV, these events occur for 10–43% of decays, depending on the position and recoil direction of [14]C. Proton transfer to the phosphate unit was also seen but we did not classify this as bond-breaking as it can be reversed. In a biological context ^{14}C is therefore an unavoidable source of strand breakage in RNA. Equivalent simulations in DNA observed similar E_T values.

DNA nucleoside

The next step in our calculations was to consider complete DNA nucleosides (i.e. a base pair and associated deoxyribose rings) for adenine-thymine (Figure 6a) and guanine-cytosine (Figure 7a). Unexpectedly, we found that the environment around the nucleobase slightly reduces the threshold energies for bond breaking and ring opening (Figures 6b, 6c and Figures 7b, 7c). On average, the E_T values for the adenine-thymine and guanine-thymine base pairs are respectively 1.3 eV and 0.75 eV lower than their gas-phase counterparts (Figures 4b, 4d and Figures 4f, 4h); even so, the probability of ^{14}C-induced bond breaking is generally less than 10%. The comparison of E_T obtained for the adenine-thymine (Figures 6b, 6c) and guanine-cytosine (Figures 7b, 7c) nucleosides show that the guanine-cytosine base pair is slightly more resistant to ^{14}C decay than the adenine-thymine base pair, with an average E_T value higher by 0.3 eV.

One of the contributing factors that reduces the thresh-

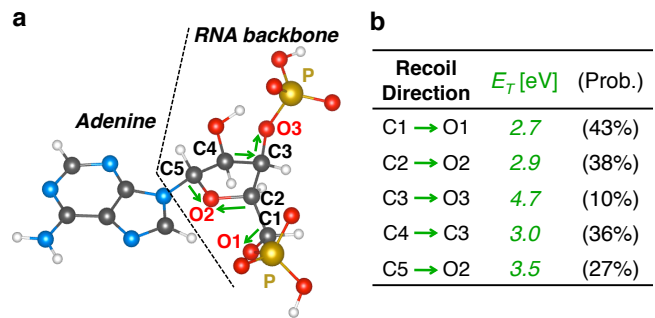| Recoil Direction	E_T [eV]	(Prob.)
C1 → O1	2.7	(43%)
C2 → O2	2.9	(38%)
C3 → O3	4.7	(10%)
C4 → C3	3.0	(36%)
C5 → O2	3.5	(27%)

FIG. 5. Recoil directions, threshold energies and probabilities in RNA nucleotide. (a) Recoil directions leading to bond breaking (shown by green arrows) in an RNA nucleotide. (b) Threshold energies and associated ^{14}C probabilities.
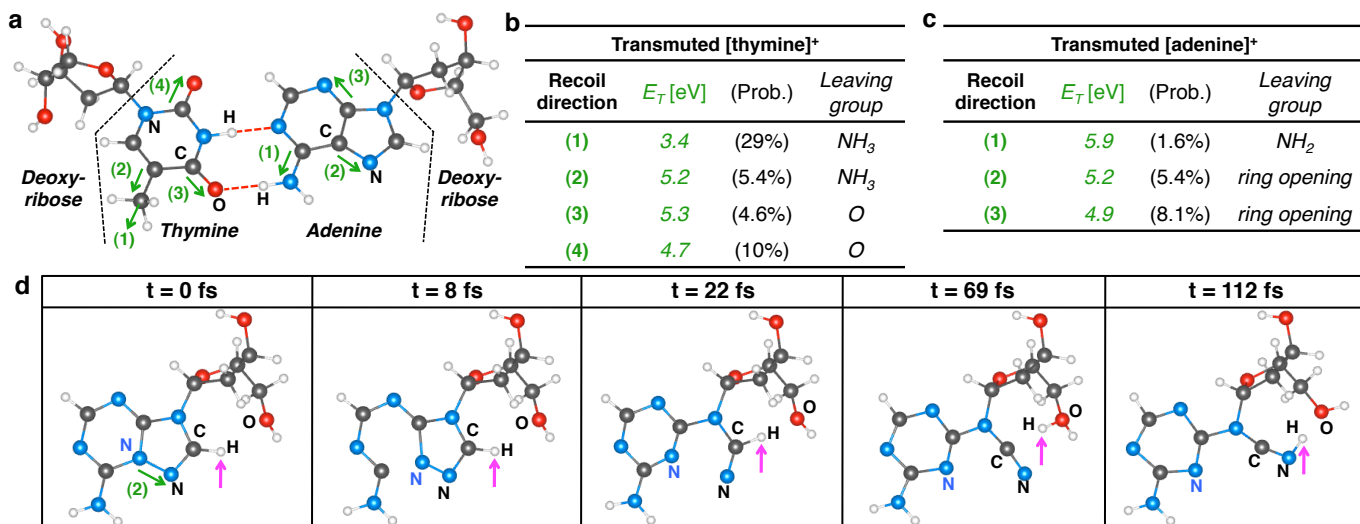
FIG. 6. Recoil directions, threshold energies and probabilities in the adenine-thymine DNA nucleoside. (a) Recoil directions leading to bond breaking (shown by green arrows) in a adenine-thymine DNA nucleoside. The tables in (b) and (c) list the threshold energies, associated ^{14}C probabilities and leaving groups. Panel (d) shows a 6 eV recoil along direction (2) in adenine for which ring-opening occurs. A proton transfer reaction (C→O→N) is indicated by the purple arrow.

old energy relative to the gas-phase values is the proximity of surrounding oxygen atoms which facilitate proton transfer and ring opening for adenine and guanine. Figure 6d shows an example of such a process in an adenine-thymine base pair in which the hydrogen atom (indicated by the purple arrow) is shuttled from a carbon to a nitrogen via an oxygen in the phosphate backbone. A similar process occurs in the guanine-cytosine base pair, whereby a recoil in the direction (2) of the guanine leads to ring opening (Figure 7a, 7b). For the guanine-cytosine nucleoside, the sequence of images in Figure 7d shows the oxygen atom of the cytosine leaving after a recoil energy of 4.7 eV in the direction (3). This event irreversibly breaks one of the three hydrogen bonds linking the guanine and cytosine molecules, and is one of the most likely sources of damage from ^{14}C, with an equivalent probability of 10%. Ring opening was observed for three of the four base pairs, with thymine being the only base where ring opening did not occur for 7 eV recoils in any direction. We also found that transmutations in the base pairs did not affect the deoxyribose rings. To confirm the insensitivity of the nucleoside to shake-off we simulated additional recoils in a doubly ionized supercell. The threshold energy was lowered by a maximum of 0.2 eV, and in some cases was identical to the singly ionized case.

The threshold energies presented to this point have only discussed leaving groups associated with the initial direction of the recoil as this is the primary pathway for recoil-induced bond-breaking. However, there is an additional indirect reaction which can occur when recoil shakes the ring. In this indirect reaction, even when there is no bond-breaking along the recoil direction, a proton can be shuttled from one base pair to another, effectively inverting the N–H···N hydrogen bond. In thymine this reaction was frequently observed, even when the recoil energy was as low as 3.5 eV, well below the threshold for direct bond-breaking. Proton shuttling was less frequent in guanine and was never seen in adenine and cytosine. The absence of proton shuttling in adenine and cytosine can be understood on the basis that the hydrogen bonds they form involve NH_2 groups which isolates the proton from the ring. In contrast, the vulnerable proton in thymine and guanine is directly linked to the hexagonal pyrimidine ring. This suggests that structural distortion associated with shaking of the ring weakens the N–H bond and hence facilitates proton shuttling. The same interpretation also explains the observed difference between thymine and guanine, where the latter is less vulnerable to proton shuffling due to the pentagonal imidazole ring which absorbs some of the vibration.

Hydrated adenine-thymine nucleoside

To assess the impact of water on recoil-induced bond breaking we considered an adenine-thymine nucleoside which was hydrated by explicit water molecules. These calculations contained 24 water molecules, with 6 and 18 water molecules in the first and second hydration shells, respectively. Figure 8 shows a top view of the system with the oxygen of the first and second hydration shells represented by green and yellow spheres, respectively.

The recoil directions investigated were the same as those indicated by the green arrows in Figure 6a. The threshold energies for hydrated thymine along directions (1), (3) and (4) are about 0.5 eV higher than those in
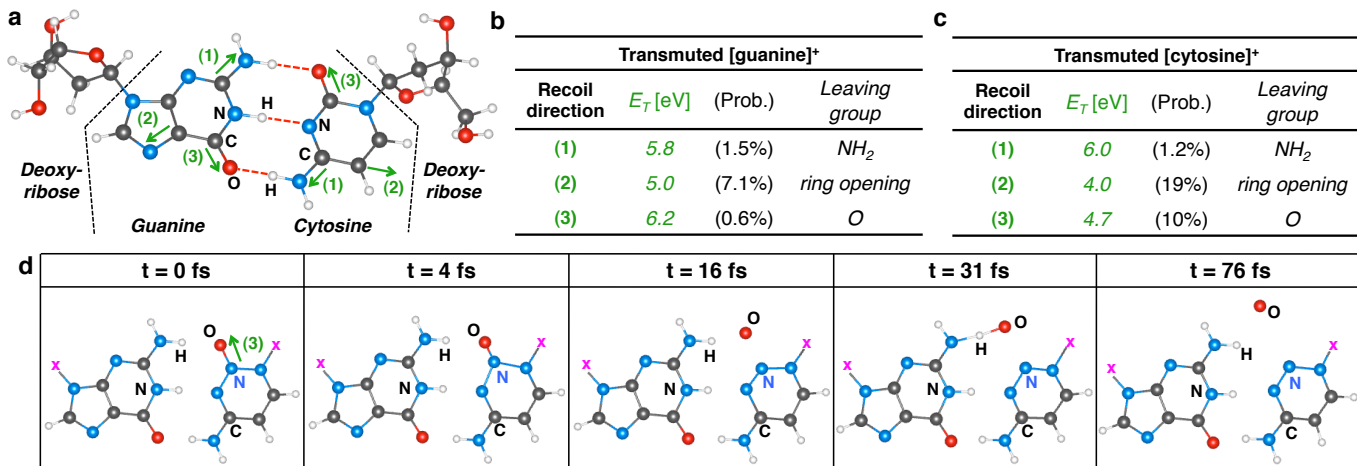
FIG. 7. Recoil directions, threshold energies and probabilities in the guanine-cytosine DNA nucleoside. (a) Recoil directions leading to bond breaking (shown by green arrows) in a guanine-cytosine DNA nucleoside. The tables in (b) and (c) list the threshold energies, associated ^{14}C probabilities and leaving groups. Panel (d) shows a 4.7 eV recoil along direction (3) in cytosine for which a leaving oxygen atom is observed. The purple cross locate the link to the DNA backbone.

the gas-phase (Figure 6b), while in direction (2) the hydrated and gas-phase values are essentially the same. This suggests that nearby water molecules (labelled 1 and 5 in Figure 8) help to prevent fragmentation by creating additional O–H$\cdots$O hydrogen bonds with the thymine molecule.

Compared to the gas-phase, the simulations of recoil in hydrated adenine show that along directions (1) and (2) [see Figure 6a] the threshold for bond breaking decreases by 0.6 eV and 0.5 eV, respectively, while the threshold
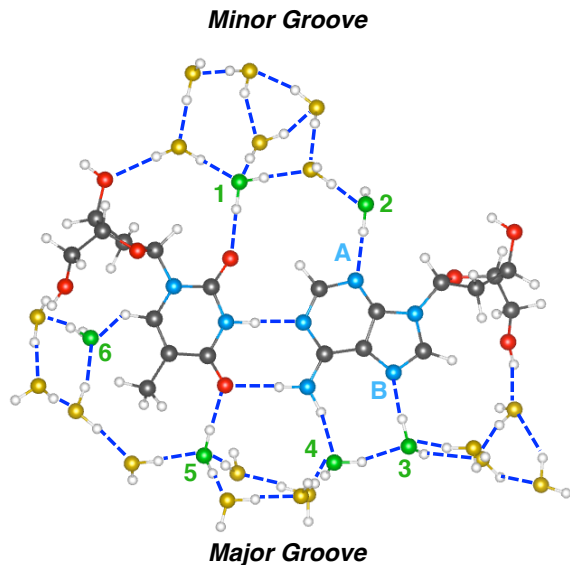

FIG. 8. Top view of a hydrated adenine-thymine nucleoside. Water molecules of the first and second hydration shells are represented by green and yellow oxygen atoms respectively. The latter were held fixed during the recoil simulations. Hydrogen bonds are symbolized by blue dashed lines.

along the direction (3) increases by 0.4 eV. The higher threshold energy of the hexagonal ring is due to the presence of the water molecule, labelled 2 in Figure 8. This molecule sits on the trajectory of the nitrogen atom, labelled A, and occupies volumes needed for the ring to open, and therefore a slightly greater recoil energy is required to overcome this obstruction. In contrast, the pentagonal ring opens more easily when water is present. This is due to the N$\cdots$H–O hydrogen bond between the adenine and the water molecule, labelled 3. This helps to open up the pentagon ring by attracting the nitrogen atom, labelled B. Opening of the pentagonal ring is also assisted by the proton transfer process identified in the gas-phase calculations (see Figure 6d).

Altogether, these simulations show that the qualitative effect of radiocarbon decay is in fact well represented by the gas-phase calculations. In some instances the presence of water slightly increases the threshold energy for bond breaking, while in other cases the threshold is slightly reduced. The precise details depend on the direction of the recoil and the position of the water molecules at the time of recoil. The variation between the gas-phase and hydrated calculations are at most 0.6 eV and typically much less. The reason for this similarity is that beta-decay recoil is a highly non-equilibrium process, and hence the surrounding water molecules play only a second-order effect with regards threshold energies for bond breaking.

DISCUSSION

We are now in a position to make some general comments regarding the effect of ^{14}C in biological systems. The first observation is that the simulations show that

^{14}C decay always leads to some type of structural or chemical change at the molecular level. Sometimes the result will be broken bonds (perhaps $\sim$10% of the time), while in other cases ^{14}C decay in a base can create a molecule foreign to the DNA "alphabet". The result of ^{14}C decay can be more finely classified into three categories: (i) the creation of open rings and leaving groups which will almost certainly distort the double helix and change the stacking interactions, (ii) the special case of an NH_3 leaving group in thymine which transforms the base into (deprotonated) uracil, and (iii) molecules remain intact but their chemical structure is modified. The first category can likely be corrected by DNA repair mechanisms through which enzymes detect distortion to the double helix and repair the damaged region using information derived from the complementary base. The second category of uracil arising from thymine can be repaired by one of several pathways which excise uracil from a DNA strand [36]. An interesting aside is that uracil in DNA usually arises from either deamination of cytosine, creating a mutagenic uracil-guanine (mis)pair, or from misincorporation of deoxyuridylate which creates a uracil-adenine pair. Our simulations highlight a new pathway to create uracil in DNA via the decay of ^{14}C in the methyl group of thymine, thereby creating a uracil-adenine pair. The third category of damage involving C$\rightarrow$N substitution without bond-breaking raises intriguing and open questions. It is not immediately obvious that this alteration will trigger the DNA repair mechanism as the chemical change is subtle and may not generate structural distortion of the major and minor grooves. We are not aware of any biological process for dealing with three N atoms (rather than the usual two) in the ring of a nitrogenous base, nor of the consequences during DNA replication or RNA transcription. The posing of these questions is an important outcome of this study.

The proton shuttling mechanism observed in thymine and guanine offer an additional route to create mutation by promoting wobble-type mispairings between bases. In the case of proton transfer from thymine to adenine, the resulting protonated adenine can pair with a cytosine base during DNA replication [37], thereby creating a mutation at the thymine site. In contrast, the strand containing the deprotonated thymine (i.e. the location of the transmuted ^{14}C) cannot pair to any of the canonical bases. A similar situation applies for the less vulnerable guanine, whereby the protonated cytosine can form mispairs and thus generate mutation. These wobble mispairs are well-known and are thought to be the dominant source of DNA replication errors [38], being far more prevalent than tautomeric errors initially suggested by Watson and Crick [39] in which proton shift within a base creates a higher-energy "enol" form which has different pairing characteristics.

One aspect of the ^{14}C decay process which is difficult to assess is the importance of transmutation relative to other forms of DNA damage. UV light, oxidation and hydrolysis are just some of the many sources of damage which need to be dealt with by DNA repair mechanisms. These repair mechanisms are highly efficient, detecting and correcting tens of thousands of DNA modifications per cell each day [40]. In contrast, around 20–30 ^{14}C decays occur in the entire human body each second, amounting to around 2 million ^{14}C decays in DNA per day. On the basis of 50 trillion cells in the body, the probability of a ^{14}C decay occurring on a given day in a given cell is accordingly very small, of order 1 in 25 million. The critical question therefore is whether the DNA repair mechanism is sensitive to non-canonical bases. If it is not, then ^{14}C decay is a potential source of genetic damage and/or mutation which despite its low frequency can persist and accumulate from one cell generation to the next.

The next set of observations concern the genetic code itself. The first, almost anthropic, observation is that the Q-value (or more precisely, the β spectrum end-point) is such that some carbon-based molecules break under ^{14}C decay while others don't. If Q had been larger, then the presence of ^{14}C may have significantly decreased the prospects of carbon-based life. For example, tripling the Q-value for carbon gives a maximum recoil energy of 27 eV, which greatly increases the probability of bond breaking. Comparison of ^{14}C with its neighbours in the periodic table reveals a surprising fact, namely that its maximum recoil energy is very low. Figure 9 presents data for all of the second-row elements, showing the maximum recoil energy for the lightest unstable isotope with an excess of neutrons. All of these isotopes are β^- emitters, with most having 2 more neutrons than protons, as in ^{14}C. It is apparent from the figure that ^{14}C is unusual compared to its elemental neighbours, most of which have maximum recoil energies above 500 eV. Were such recoil energies to occur within DNA the disruption would likely be irreversible due to the large number of displaced atoms. From the viewpoint of organic life which must accommodate radiocarbon, it is therefore rather fortunate that the constants of nature are such that ^{14}C has a much lower Q-value (and hence much lower recoil energy) than its surrounding elements. This situation is reminiscent of another coincidence involving carbon anticipated by Fred Hoyle [41]. He proposed that an excited state of ^{12}C exists with very particular properties that enable the nucleosynthesis of carbon in stars, and several years later this prediction was confirmed by experiments [42].

Another observation specifically related to the molecular structure of the genetic code concerns the presence of purines and pyrimidines in DNA. One of the important findings of this study is the demonstration that double bonds and ring structures increase the resistance of a molecule to ^{14}C decay, both of which are present in purines and pyrimidines. This observation that nucleobases are largely resistant to radiocarbon is likely a co-
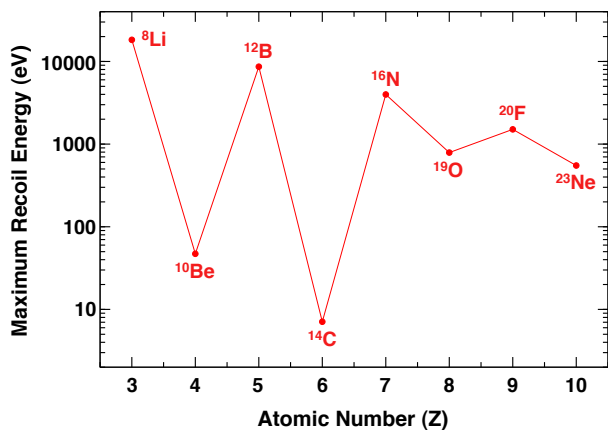
FIG. 9. Maximum recoil energy for second-row unstable isotopes analogous to ^{14}C. Each isotope has one more neutron than its heaviest stable counterpart. The recoil energies were computed using relativistic kinematics and experimental Q-values.

incidence of evolutionary history. In the early Earth the absence of oxygen in the atmosphere exposed molecules to deep UV light, with energies as high as 6-10 eV [43]. These conditions provided a strong selection pressure for molecules that can resist photo-ionization. While ionization and recoil-induced bond-breaking are clearly different processes, the similarity in the respective energies suggests that the physical and chemical processes which selected purines and pyrimidines simultaneously resulted in molecules which resist radiocarbon.

From the viewpoint of the backbone, the simulations reveal a vulnerability to ^{14}C decay which has not previously been discussed in the literature. All ^{14}C radiolabelling experiments that we are aware of have employed ^{14}C in the nucleobase, rather than in the sugar group. Our simulations show that the ribose group in RNA and deoxyribose group in DNA are both highly vulnerable to ^{14}C-induced fragmentation due to the presence of single-bonds in the pentagonal ring. In a biological context this means that a ^{14}C atom present in a sugar group will have a significant chance to create a single strand break. In RNA this damage would be irreversible, while in DNA the repair mechanism can restore the nucleoside using information from the complementary base. It would be instructive for experimentalists to study this process by radiolabelling the sugar group in molecules which substitute for nucleosides. We predict a large number of single strand breaks (as compared to ^{14}C in a nucleobase) but relatively few double strand streaks since the recoil process is extremely localised and shake-off is minimal.

To summarize, we have applied *ab initio* molecular dynamics to measure the recoil radiation stability of a variety of hydrocarbon and alcohol molecules, including DNA base pairs. We find that the recoil energy of the daughter ^{14}N is sometimes sufficient to break chemical bonds and that the presence of water has only a minor effect on the fundamental processes. Whereas chain-like molecules are relatively vulnerable to fragmentation, features such as double bonds and ring structures enhance the molecular resistance to ^{14}C decay. The presence of these characteristics in the canonical nucleobases confers DNA with a resistance to natural radioactivity that has not previously been recognized. Our findings also raise questions as to how the genetic apparatus deals with the appearance of an extra nitrogen, as in most situations the radioactive decay process will not lead to bond breaking. We also identify proton shuttling as a mechanism for creating wobble-type mispairs through the formation of a protonated base. Over the lifetime of a typical human approximately 50 billion such events will occur within DNA, amounting to a radioactive decay of ^{14}C in the DNA of one every thousand cells of the human body. It is not obvious whether or not the DNA repair mechanism detects this modification nor how DNA replication is affected by a non-canonical nucleobase. Accordingly, ^{14}C may prove to be a source of genetic alteration that is impossible to avoid due to the universal presence of radiocarbon in the environment.

* mail:N.Marks@curtin.edu.au

[1] C. E. Purdom, Biological hazards of carbon-14, New Scientist 298 (1962) 255–257.

[2] Calculation based upon 50 trillion cells, two copies of the genome per cell, 2.91 billion base pairs in the genome, an average of 19.5 carbon atoms per nucleoside pair, an atomic mass of carbon of 12.01 amu and a specific activity for radiocarbon of 0.25 Bq/g.

[3] A. Halpern, G. Stöcklin, Chemical and biological consequences of β-decay. Part II, Rad. and Environm. Biophys. 14 (1977) 257–274.

[4] A. H. Snell, F. Pleasonton, The atomic and molecular consequences of radioactive decay, J. Phys. Chem. 62 (1958) 1377–1382.

[5] R. L. Wolfgang, R. C. Anderson, R. W. Dodson, Bond rupture and nonrupture in the beta decay of carbon-14 studied by double isotopic labeling, J. Chem. Phys. 24 (1956) 16–23.

[6] P. G. Manning, C. B. Monk, 496. beta energy recoil effects of carbon-14 in toluene and ethylbenzene, J. Chem. Soc. (1962) 2573–2576.

[7] IAEA, Biological Effects of Transmutations And Decay Of Incorporated Radioisotopes, IAEA: International Atomic Energy Agency, Vienna, 1968.

[8] R. E. Krisch, M. R. Zelle, Biological effects of radioactive decay. The role of transmutation effect, Advanc. Rad. Biol. 3 (1969) 177–213.

[9] A. Halpern, G. Stöcklin, Chemical and biological consequences of β-decay. Part I, Rad. and Environm. Biophys. 14 (1977) 167–183.

[10] S. Apelgot, R. Latarjet, Labelling of bacterial deoxyribonucleic acid with radioactive phosphorus, carbon and tritium; comparison of the lethal effects, Biochim. Biophys. Acta 55 (1962) 40–55.

[11] W. R. Lee, G. A. Sega, E. S. Benson, Transmutation of carbon-14 within DNA of drosophila melanogaster spermatozoa, Mutat. Res. 16 (1972) 195–201.

[12] W. R. Lee, et al., Stability of drosophila chromosomes to radioactive decay of incorporated phosphorus-32, Genetics 53 (1966) 807–822.

[13] W. R. Lee, G. A. Sega, C. F. Alford, Mutations produced by transmutation of phosphorus-32 to sulfur-32 within drosophila DNA, Proc. Natl Acad. Sci. USA 58 (1967) 1472–1479.

[14] J. E. Cleaver, H. J. Burki, Biological damage from intranuclear carbon-14 decays: DNA single-strand breaks and repair in mammalian cells, Int. J. Radiat. Biol. 26 (1974) 399–403.

[15] D. E. Charlton, J. Booz, A Monte Carlo treatment of the decay of ^{125}I, Radiat. Res. 87 (1981) 10–23.

[16] E. Pomplun, J. Booz, D. E. Charlton, A Monte Carlo simulation of Auger cascades, Radiat. Res. 111 (1987) 533–552.

[17] D. E. Charlton, J. L. Humm, A method of calculating initial DNA strand breakage following the decay of incorporated ^{125}I, Int. J. Radiat. Biol. 53 (1988) 353–365.

[18] G. Raisali, L. Mirzakhanian, S. F. Masoudi, F. Semsarha, Calculation of DNA strand breaks due to direct and indirect effects of auger electrons from incorporated ^{123}I and ^{125}I radionuclides using the Geant4 computer code, Int. J. Radiat. Biol. 89 (2013) 57–64.

[19] C. Jiang, C. R. Stanek, N. A. Marks, K. E. Sickafus, B. P. Uberuaga, Predicting from first principles the chemical evolution of crystalline compounds due to radioactive decay: The case of the transformation of cscl to bacl, Phys. Rev. B 79 (2009) 132110–132114.

[20] C. Jiang, B. P. Uberuaga, K. E. Sickafus, F. M. Nortier, J. J. Kitten, N. A. Marks, C. R. Stanek, Using "radioparagenesis" to design robust nuclear waste forms, Energy Environ. Sci. 3 (2010) 130–135.

[21] N. A. Marks, D. J. Carter, M. Sassi, A. L. Rohl, K. E. Sickafus, B. P. Uberuaga, C. R. Stanek, Chemical evolution via beta decay: a case study in strontium-90, J. Phys.: Condens. Matter 25 (2013) 065504.

[22] C. Raadschelders-Buijze, C. L. Roos, P. Ros, A theoretical study of the effect of the molecular structure on he electronic transition probabilities after the ^{14}C β-decay, Chem. Phys. 1 (1973) 463–475.

[23] S. Ikuta, S. Iwata, M. Imamura, Ab initio studies of the β-decay in OHT, NH_2T, CH_3T, and $^{14}CH_4$, J. Chem. Phys. 66 (1977) 4671–4676.

[24] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B 54 (1996) 11169–11186.

[25] G. Kresse, J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Comput. Mater. Sci. 6 (1996) 15–50.

[26] Y. Wang, J. P. Perdew, Correlation hole of the spin-polarized electron gas, with exact small-wave-vector and high-density scaling, Phys. Rev. B 44 (1991) 13298–13307.

[27] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B 50 (1994) 17953–17979.

[28] S. H. Vosko, L. Wilk, M. Nusair, Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis, Can. J. Phys. 58 (1980) 1200–1211.

[29] G. Henkelman, A. Arnaldsson, H. Jónsson, A fast and robust algorithm for Bader decomposition of charge density, Comput. Mater. Sci. 36 (2006) 354–360.

[30] B. Schneider, H. M. Berman, Hydration of the DNA bases is local, Biophys. J. 69 (1995) 2661–2669.

[31] M. Feig, B. M. Pettitt, Modeling high-resolution hydration patterns in correlation with DNA sequence and conformation, J. Mol. Biol. 286 (1999) 1075–1095.

[32] P. Auffinger, E. Westhof, Water and ion binding around RNA and DNA (C,G) oligomers, J. Mol. Biol. 300 (2000) 1113–1131.

[33] S. Arai, T. Chatake, T. Ohhara, K. Kurihara, I. Tanaka, N. Suzuki, Z. Fujimoto, H. Mizuno, N. Niimura, Complicated water orientations in the minor groove of the B-DNA decamer d(CCATTAATGG)$_2$ observed by neutron diffraction measurements, Nucleic Acids Res. 33 (2005) 3017–3024.

[34] K. Grzeskowiak, K. Yanagi, G. G. Prive, R. E. Dickerson, The structure of B-helical G-G-A-T-C-G-A-T-C-G and comparison with C-C-A-A-C-G-T-T-G-G, J. Biol. Chem. 266 (1991) 8861–8883.

[35] R. F. W. Bader, Atoms in Molecules: A Quantum Theory, The International Series of Monographs on Chemistry, Clarendon Press, Oxford, 1990.

[36] H. E. Krokan, F. Drabløs, G. Slupphaug, Uracil in DNA - occurrence, consequences and repair, Oncogene 21 (2002) 8935–8948.

[37] B. A. Pierce, Genetics: A Conceptual Approach, 2nd edition, W.H. Freeman and Company, New York, 2005.

[38] L. A. Pray, DNA replication and causes of mutation, Nature Education 1 (1) (2008).

[39] J. D. Watson, F. H. C. Crick, Genetical implications of the structure of deoxyribonucleic acid, Nature 171 (1953) 964–967.

[40] S. L. Wolfe, Molecular and Cellular Biology, Wadsworth Inc., Wadsworth Publishing Company, Belmont, California, 1993.

[41] F. Hoyle, On nuclear reactions occurring in very hot STARS.I. the synthesis of elements from carbon to nickel, Astrophys. J. Suppl. Ser. 1 (1954) 121–146.

[42] C. Cook, W. A. Fowler, C. C. Lauritsen, T. Lauritsen, B^{12}, C^{12}, and the Red Giants, Phys. Rev. 107 (1957) 508–515.

[43] J. A. Ratcliffe, Physics of the Upper Atmosphere, Academic Press, New York, 1960.