Bi-imidazole nucleosides obtained by ring opening of etheno and substituted etheno derivatives of adenosine

Jukka Mäki, Rainer Sjöholm and Leif Kronberg *

Department of Organic Chemistry, Åbo Akademi University, Biskopsgatan 8, FIN-20500 Turku/Åbo, Finland

Received (in Cambridge, UK) 5th July 2000, Accepted 10th October 2000
First published as an Advance Article on the web 20th November 2000

Etheno- and substituted ethenoadenosines **1a**–**1d** are found to be converted into bi-imidazole nucleosides **2a**–**2d** in basic aqueous solutions. The formyl- and oxalo-substituted etheno derivatives of adenosine (**1b** and **1c**) are more labile than the other compounds. This difference can be explained by the electron-withdrawing effects of the formyl and oxalo groups that create a more pronounced partial positive charge on C-5 in the ethenoadenosine derivatives **1b** and **1c**. The structural characterisation of the bi-imidazoles is mainly carried out by ^{1}H and ^{13}C NMR spectroscopy. In **2b** and **2c** some of the carbon resonance signals are strongly broadened, most likely due to a prototropic effect in the outer imidazole ring. Addition of trifluoroacetic acid to the NMR samples suppresses the effect and sharp signals could be observed for all carbons in the compounds. The structural characterisation of the bi-imidazoles is largely dependent on multiple-bond proton–carbon coupling constants ($^{>1}J_{C,H}$), and therefore the constants have been assigned and are reported.

Introduction

It is generally believed that the genotoxic effects of chemical compounds are due to their reactions with the genetic material, the DNA, and especially with the base units of DNA. One of the most commonly encountered modification is the 1,N^6-ethenoadenosine adduct **1a**.[1] However, this adduct has been reported to be unstable and to decompose to a bi-imidazole derivative **2a**.[2] The decomposition takes place also in a synthetic oligonucleotide containing the ethenoadenosine unit. The formed bi-imidazole nucleoside is thought to be a more harmful lesion in DNA than is the ethenoadenosine lesion.[3]

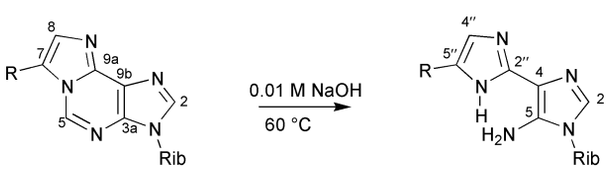
In previous work performed in our laboratory, we have shown that adenosine reacts with the genotoxic compound mucochloric acid [3,4-dichloro-5-hydroxyfuran-2(5*H*)-one] to produce ethenoadenosine **1a**,[4] formylethenoadenosine **1b**[5] and oxaloethenoadenosine **1c**.[6] Further, we found that the formyl-etheno adduct **1b** was formed in double-stranded DNA treated with mucochloric acid.[7] Therefore, we were interested in finding out whether also the formyl- and the oxaloetheno derivatives decompose to the corresponding substituted bi-imidazoles. At the same time, we studied also the degradation of the hydroxymethylethenoadenosine derivative **1d**. This derivative is obtained when glycidaldehyde (an animal carcinogen) is allowed to react with adenosine.[8]

This paper deals with the structural characterisation of the bi-imidazoles **2a**–**2d**. We report the full assignment of the proton and carbon chemical shifts. Also reported are the multiple bond proton–carbon coupling constants ($^{>1}J_{C,H}$), which are of critical importance in the structural characterisation of compounds containing only a few protons.

Results and discussion

Formation of bi-imidazoles

The compounds **1a**–**1d** were stored in 0.01 M NaOH solutions at 60 °C and the decomposition was followed by HPLC analyses of the mixtures. The major decomposition products were isolated and their structures were determined from spectroscopic and spectrometric data. The degradation was found to produce bi-imidazole derivatives **2a**–**2d** from the etheno compounds (Scheme 1). The mechanism depicted in



1a: R = H
1b: R = CHO
1c: R = COCOOH
1d: R = CH_2OH

2a: R = H
2b: R = CHO
2c: R = COCOOH
2d: R = CH_2OH

Rib = β-D-ribofuranosyl

Scheme 1 Ring-opening of ethenoadenosine and substituted etheno-adenosine derivatives.

Scheme 2 for the decomposition of ethenoadenosine has been proposed by Tsou *et al.*[2b] and is most likely to be valid for the decomposition of **1b**–**1d** as well. The C-5 in ethenoadenosine **1a** is attacked by a hydroxide anion, resulting in the cyclic intermediate. After the ring opening, another hydroxide ion is added to the formyl carbon. Finally, the C-5 carbon from the original ethenoadenosine moiety is lost as a formate anion to give the bi-imidazole product. Tsou and co-workers[2b] gained evidence for the mechanism from ^{1}H NMR spectroscopy studies and using tritium-labelled ethenoadenosine to monitor the degradation process.

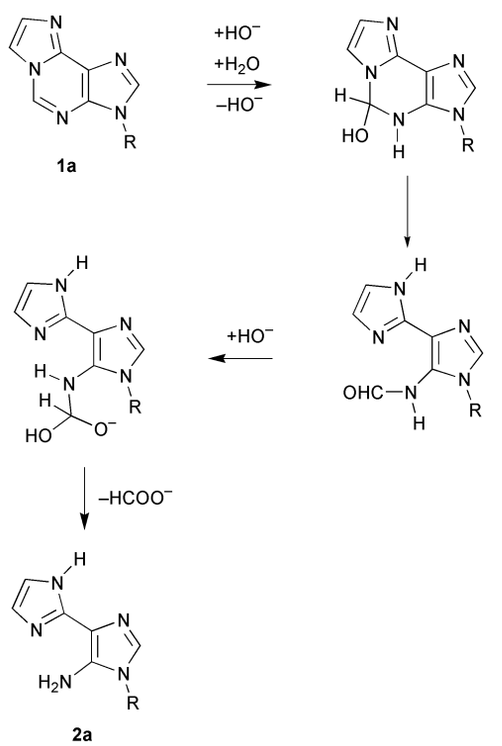
We observed that **1b** and **1c** formed the bi-imidazoles under slightly basic conditions (pH 9.0, 37 °C), whereas **1a** and **1d** were almost stable under these conditions. The most likely reason for the difference in stability is the presence of electron-withdrawing groups on the etheno bridge of **1b** and **1c**. The formyl and oxalo groups draw electrons from the pyrimidine ring and make the carbon C-5 more susceptible to attack of the nucleophilic hydroxide ion. This electron-withdrawing effect of the substituents is also reflected in the difference in the proton chemical shift of H-5 in **1b** and **1c** compared with the shift of H-5 in **1a**.[6]

DOI: 10.1039/b005407k

Table 1 The ^{1}H chemical shifts (δ) and spin–spin coupling constants ($J_{H,H}$) of compounds **2a**–**2d**

	Compound							
	2a		**2b**		**2c**		**2d**	
Proton[a]	δ (ppm)[b]	J/Hz[c]	δ (ppm)[d]	J/Hz[c]	δ (ppm)[d]	J/Hz[c]	δ (ppm)[b]	J/Hz
H-2	7.68 (1H)		8.28 (1H)		8.01 (1H)		7.63 (1H)	
5-NH2	6.73 (2H)		[e]		[e]		6.48 (2H)	
H-1″	[e]		[e]		[e]		[e]	
H-4″	7.40 (2H)		8.22 (1H)		8.18 (1H)		7.16 (1H)	
CHO			9.71 (1H)					
CO_2H					[e]			
CH_2OH							4.47 (2H)	
CH_2OH							[e]	
H-1′	5.59 (1H)	6.1 (H1′–H2′)	5.73 (1H)	5.2 (H1′–H2′)	5.64 (1H)	6.2 (H1′–H2′)	5.57 (1H)	6.2 (H1′–H2′)
H-2′	4.29 (1H)	5.3 (H2′–H3′)	4.33 (1H)	5.4 (H2′–H3′)	4.36 (1H)	5.1 (H2′–H3′)	4.31 (1H)	5.3 (H2′–H3′)
H-3′	4.10 (1H)	3.3 (H3′–H4′)	4.13 (1H)	3.9 (H3′–H4′)	4.11 (1H)	3.2 (H3′–H4′)	4.09 (1H)	3.1 (H3′–H4′)
H-4′	3.95 (1H)	3.0 (H4′–H^{a}5′) (H4′–H^{b}5′)	4.00 (1H)	3.1 (H4′–H^{a}5′) (H4′–H^{b}5′)	3.98 (1H)	3.2 (H4′–H^{a}5′) (H4′–H^{b}5′)	3.94 (1H)	3.2 (H4′–H^{a}5′) (H4′–H^{b}5′)
H^a-5′	3.61 (1H)	2.9 (H^{a}5′–H^{b}5′)	3.67 (1H)	3.3 (H^{a}5′–H^{b}5′)	3.65 (1H)	2.8 (H^{a}5′–H^{b}5′)	3.62 (1H)	3.1 (H^{a}5′–H^{b}5′)
H^b-5′	3.60 (1H)	−12.0	3.63 (1H)	−11.6	3.62 (1H)	−12.1	3.60 (1H)	−12.2
HO[f]	[e]		[e]		[e]		[e]	

[a] H-1′–H^b-5′ Protons in the ribosyl units. [b] In DMSO-d_6. [c] The linewidths used in the simulation were optimised for each spin particle. The couplings are given only once for each involved nucleus. [d] In DMSO-d_6 + TFA. [e] Not observed. [f] Hydroxy groups in the ribosyl units.



Scheme 2 The proposed mechanism for the formation of bi-imidazole nucleoside.[2b]

Spectroscopic characterisation of 2a †

The UV spectrum of **2a** recorded in water displayed absorption maxima at 213 and 278 nm. A UV minimum was observed at 230 nm.

In the ^{1}H NMR spectrum, two signals originating from the imidazole rings were observed (Table 1). The signal at δ_H 7.68 was assigned to H-2 due to the long-range H–H correlation (long-range COSY) with the distinct doublet of H-1′ in the ribosyl unit observed at δ_H 5.59. The other signal, a singlet corresponding to the area of two protons, was observed at δ_H 7.40. This signal was assigned to both H-4″ and H-5″. This observation could be due to either a protonation of the N-3″ (pK_a = 7.00 for protonated imidazole[9]), or a rapid prototropic equilibrium (on the NMR time-scale) in the 'symmetrical' imidazole ring (*i.e.* the proton transfer between N-1″ and N-3″) and only one resonance signal arising from both H-4″ and H-5″ can be observed in the spectrum. Previously it had been reported that in methyl-substituted aminobi-imidazole (the methyl group being bound to N-1″) the H-4″ and H-5″ protons were magnetically equivalent when the NMR spectrum was recorded in DCl or D_2O.[10] Whether it is due to protonation or prototropy (or both), the process that is taking place in **2a** is presumably enhanced by the residual amount of water present in the NMR sample.

In addition to the ribosyl carbons, the ^{13}C NMR spectrum displayed five signals originating from the imidazole moieties (Table 2). The methine carbons were assigned from the one-bond C–H correlation spectra (HETCOR) and the fully proton-coupled ^{13}C spectrum (the coupling constants $^1J_{C,H}$ and $^{>1}J_{C,H}$ are reported in Table 3). The signal at δ_C 132.0 was assigned to C-2 based on the one-bond correlation to H-2. The signal displayed a long-range correlation (COLOC) and coupling ($^3J_{C,H}$ 3.9 Hz) to H-1′. Further, two long-range correlations were observed from H-2. One correlation was observed to the signal at δ_C 105.1. This signal was attributed to C-4 due to the vicinal coupling ($^3J_{C,H}$ 11.5 Hz) to H-2 mediated *via* the unsaturated CNCH fragment. The other correlation was observed to the signal at δ_C 139.7 and this was assigned to C-5. The C-5 signal was split into two doublets ($^3J_{C,H}$ 4.1 and 3.2 Hz) based on the couplings with H-2 and H-1′, respectively. The signal at δ_C 117.6 was assigned to C-4″ and C-5″ based on the one-bond correlation with the signal of H-4″ and H-5″. The signal was split into two doublets due to the direct one-bond coupling ($^1J_{C,H}$ 199.8 Hz) and the geminal coupling ($^2J_{C,H}$ 11.9 Hz) to the adjacent proton. Finally, the signal at δ_C 140.4 was assigned to C-2″ based on the long-range correlation with the signal of H-4″ and H-5″. The splitting of the signal into a

† The assignment of the NMR signals from the ribosyl units of compounds **2a**–**2d** is omitted from the discussion. However, the chemical shifts and coupling constants are reported in the Tables.

triplet ($J_{C,H}$ 6.9 Hz) because of the coupling to H-4″ and H-5″ confirmed the assignment.

Spectroscopic characterisation of 2b

Compound **2b** exhibited UV absorption maxima at 245 and 326 nm, whereas UV minima were observed at 212 and 297 nm.

When the ^{13}C NMR spectrum of **2b** was recorded in DMSO-d_6, it displayed only two observable signals (C-2 and C-4), other signals being hidden in the baseline noise. However, after the addition of trifluoroacetic acid (TFA) to the NMR sample, all the signals in the ^{13}C spectrum were observable. Therefore, the ^{1}H NMR spectrum of **2b** was recorded in the presence of TFA. Three proton signals were found to arise from the imidazole moieties (Table 1). The signal resonating at the lowest field, at δ_H 9.71, was assigned to the formyl proton (CHO). The signal at δ_H 8.28 was assigned to H-2 based on the long-range H–H correlation with the doublet signal of H-1′ (at δ_H 5.73). The third signal, observed at δ_H 8.22, was attributed to H-4″.‡ In the sample recorded in the absence of TFA, the proton signals shifted upfield by 0.4 to 0.9 ppm (data not shown). Similar solvent-dependence of chemical shifts has been observed for 4(5)-formylimidazole.[11]

The ^{13}C NMR spectrum displayed seven signals from the bi-imidazole moiety (Table 2). The carbon signal at δ_C 131.0 was assigned to C-2 based on the one-bond correlation and coupling to H-2 ($^1J_{C,H}$ 219.7 Hz). The signal displayed also a long-range correlation and vicinal coupling to H-1′ ($^3J_{C,H}$ 3.4 Hz). The most upfield signal from the imidazole moieties observed at δ_C 103.0 was assigned to C-4. Splitting of the signal into a doublet ($^3J_{C,H}$ 9.2 Hz) is due to the coupling with H-2 transmitted through the unsaturated CNCH fragment. In the ^{13}C NMR spectrum recorded in the absence of TFA, the C-4 signal shifted downfield and was observed at δ_C 110.2. The signal observed at δ_C 140.5 was split into two doublets exhibiting couplings with H-2 ($^3J_{C,H}$ 5.0 Hz) and H-1′ ($^3J_{C,H}$ 2.8 Hz). Consequently, this signal was attributed to C-5. In the heteronuclear long-range correlation spectra, the signal of H-1′ displayed a correlation to both C-2 and C-5. The C-4″ signal ($^1J_{C,H}$ 197.0 Hz) was observed at δ_C 131.3. The signals assigned to C-2″ and C-5″ observed at δ_C 142.8 and δ_C 133.9, respectively, exhibited long-range correlations and couplings to H-4″ ($^3J_{C\text{-}2''\text{–}H\text{-}4''}$ 9.6 Hz; $^2J_{C\text{-}5''\text{–}H\text{-}4''}$ 11.0 Hz). The C-5″ signal was split into another doublet ($^2J_{C,H}$ 29.8 Hz) based on the coupling with the formyl proton (CHO). The doublet signal ($^1J_{C,H}$ 181.9 Hz) observed at the lowest field, δ_C 180.9, was assigned to the formyl carbon (*C*HO).

The ^{13}C chemical shifts of C-2, C-4 and C-5 in **2b** correlate well with the values observed for **2a** and the differences in shifts between the carbons C-2″, C-4″ and C-5″ are easily explained by the effects of the formyl group at C-5″. The ^{13}C chemical shifts observed for C-2″, C-4″ and C-5″ are close to those reported by Kim *et al.*[12] for 4(5)-formylimidazole and do not agree with the data reported by Aulaskari *et al.*[13]

Table 2 The ^{13}C chemical shifts (δ_C) of compounds **2a**–**2d**

	Compound			
	2a[b]	**2b**[c]	**2c**[c]	**2d**[b]
Carbon[a]	δ (ppm)	δ (ppm)	δ (ppm)	δ (ppm)
C-2	132.0	131.0	130.8	131.6
C-4	105.1	103.0	105.5	106.3
C-5	139.7	140.5	140.3	139.2
C-2″	140.4	142.8	145.2	141.0
C-4″	117.6	131.3	133.3	114.6
C-5″	117.6	133.9	131.1	133.1
CHO		180.9		
CO			175.9	
CO_2H			163.6	
CH_2OH				53.9
C-1′	88.0	89.1	88.4	87.9
C-2′	73.1	73.5	73.1	73.1
C-3′	70.1	69.8	70.1	70.1
C-4′	85.6	86.1	85.9	85.5
C-5′	60.9	60.7	60.9	61.0

[a] C-1′–C-5′ Carbons in the ribosyl units. [b] In DMSO-d_6. [c] In DMSO-d_6 with two drops of TFA.

‡ Note: For simplicity, the explicit numbering of the atoms of **2b**–**2d** (as depicted in Scheme 1) in the unsymmetrically substituted imidazole moieties is used throughout the text without taking into account the other possible tautomeric form.

Table 3 The observed one- ($^1J_{C,H}$) and multiple-bond ($^{>1}J_{C,H}$) proton–carbon coupling constants of compounds **2a**–**2d**. The coupled protons are given in parentheses

	Compound											
	2a[b]			**2b**[c]			**2c**[c]			**2d**[b]		
Carbon[a]	m[d]	$^1J_{C,H}$/Hz	$^{>1}J_{C,H}$/Hz (proton)	m[d]	$^1J_{C,H}$/Hz	$^{>1}J_{C,H}$/Hz (proton)	m[d]	$^1J_{C,H}$/Hz	$^{>1}J_{C,H}$/Hz (proton)	m[d]	$^1J_{C,H}$/Hz	$^{>1}J_{C,H}$/H (proton)
C-2	dd	215.1	3.9 (H-1′)	dd	219.7	3.4 (H-1′)	dd	217.7	4.1 (H-1′)	dd	214.7	4.4 (H-1′)
C-4	d		11.5 (H-2)	d		9.2 (H-2)	d		9.4 (H-2)	d		11.5 (H-2)
C-5	dd		4.1 (H-2)	dd		5.0 (H-2)	dd		4.9 (H-2)	dd		4.2 (H-2)
			3.2 (H-1′)			2.8 (H-1′)			3.1 (H-1′)			3.1 (H-1′)
C-2″	t		6.9 (H-4″ and H-5″)	d		9.6 (H-4″)	d		10.3 (H-4″)	d		7.1 (H-4″)
C-4″	dd	199.8	11.9 (H-5″)	d	197.0		d	199.1		dm	195.9	
C-5″	dd	199.8	11.9 (H-4″)	dd		29.8 (CHO) 11.0 (H-4″)	d		10.5 (H-4″)	dm		12.6 (H-4″)
CHO				d	181.9							
CO							s					
CO_2H							s					
CH_2OH										t	143.2	
C-1′	d	162.2		d	166.3		d	164.3		dm	163.2	
C-2′	d	149.4		d	149.4		d	149.2		dm	147.6	
C-3′	d	146.6		d	148.5		d	149.8		dm	151.4	
C-4′	d	148.5		d	147.5		d	149.2		dm	148.0	
C-5′	t	140.4		t	141.1		t	141.0		tm	141.4	

[a] C-1′–C-5′ Carbons in the ribosyl units. [b] In DMSO-d_6. [c] In DMSO-d_6 with two drops of TFA added. [d] Multiplicity.

The broadening of the ^{13}C NMR signals when the spectrum of **2b** was recorded in the absence of TFA might be due to a prototropic equilibrium between N-1″ and N-3″ in the unsymmetrically substituted outer imidazole unit. This prototropy is possible since the pK_a-value for protonated 4(5)-formylimidazole is 2.90[9] and thus the primary form of **2b** should be the unprotonated form. Upon addition of TFA to the NMR sample, the 4″(5″)-formylimidazole is protonated, the prototropic effect is suppressed and sharp ^{13}C resonance signals are obtained.

Spectroscopic characterisation of 2c

The compound **2c** exhibited UV absorption maxima at 254 and 339 nm. UV minima were observed at 214 and 300 nm.

Although the ^1H NMR spectrum of **2c** gave sharp resonance signals, some of the signals were severely broadened in the ^{13}C spectrum. Therefore, TFA was added to the NMR sample. The signal broadening is presumably due to the factors discussed in the case of **2b**. And as for **2b**, the proton chemical shifts were dependent on the solvent (0.4 to 0.6 ppm downfield shifts upon the addition of TFA).

The ^1H NMR spectrum (recorded after the addition of TFA) displayed two singlet resonance signals attributable to the non-ribosyl protons. The signal at δ_H 8.01 was assigned to H-2 due to correlation with the doublet signal of H-1′ in the ribosyl unit observed at δ_H 5.64. The other non-ribosyl proton, assigned as H-4″, was observed at δ_H 8.18.

In the ^{13}C NMR spectrum, five signals from the ribosyl unit and eight signals from the modified imidazole moieties were observed. The two singlet signals resonating at lowest field, δ_C 175.9 and 163.6, were assigned to the carbonyl and carboxyl carbons in the oxalo group, respectively. The signal resonating at δ_C 130.8 was assigned to C-2 from the one-bond heteronuclear correlation spectra. The signal was split into two doublets due to the one-bond coupling with H-2 ($^1J_{C,H}$ 217.7 Hz) and the vicinal coupling with H-1′ ($^3J_{C,H}$ 4.1 Hz). In addition, a long-range correlation was observed between the signals of C-2 and H-1′. The signal at δ_C 140.3 was assigned to C-5 due to the three-bond correlations and couplings with H-2 ($^3J_{C,H}$ 4.9 Hz) and H-1′ ($^3J_{C,H}$ 3.1 Hz). The doublet signal at δ_C 105.5 was assigned to C-4 due to the long-range correlation with H-2. The coupling of the signal to H-2 ($^3J_{C,H}$ 9.4 Hz) is mediated *via* the unsaturated CNCH fragment. Similarly to **2b**, this carbon was sensitive to the solvent since, without TFA, the signal was observed at δ_C 110.7. The signal at δ_C 133.3 was assigned to C-4″. The observed one-bond coupling constant ($^1J_{C,H}$ 199.1 Hz) and correlation to H-4″ confirmed the assignment. Finally, the signals at δ_C 145.2 and 131.1 were assigned to C-2″ and C-5″, respectively. Both signals were split into doublets due to coupling with H-4″. For C-2″, the coupling is mediated *via* the CNCH fragment, which gives rise to the vicinal coupling constant ($^3J_{C,H}$ 10.3 Hz), whereas for C-5″ the coupling constant ($^2J_{C,H}$ 10.5 Hz) is due to the geminal coupling with H-4″. Furthermore, long-range correlations were observed from the signal of H-4″ to both the signals of C-2″ and C-5″.

Spectroscopic characterisation of 2d

Compound **2d** exhibited UV absorption maxima at 212 and 270 nm. UV minimum was observed at 228 nm.

In the ^1H NMR spectrum of **2d**, the doublet resonance signal of H-1′ was observed at δ_H 5.57. Based on the long-range correlation from H-1′, the singlet signal at δ_H 7.63 was assigned to H-2. The amino-group protons were observed at δ_H 6.48 as a broad singlet. The signal for H-4″ was observed at δ_H 7.16, shifted upfield by 0.24 ppm compared with the value for 4(5)-(hydroxymethyl)imidazole hydrochloride (in D_2O–1,4-dioxane).[13] Finally, the two methylene protons from the hydroxymethyl substituent were observed as a singlet at δ_H 4.47.

In the ^{13}C NMR spectrum, the signal of C-2 was observed at δ_C 131.6 exhibiting couplings to H-2 ($^1J_{C,H}$ 214.7 Hz) and to H-1′ ($^3J_{C,H}$ 4.4 Hz). The most upfield signal originating from the imidazole rings observed at δ_C 106.3 was assigned to C-4. The signal displayed a long-range correlation and coupling to H-2 ($^3J_{C,H}$ 11.5 Hz). The signal for C-5 observed at δ_C 139.2 was split into two doublets due to the coupling with H-2 ($^3J_{C,H}$ 4.2 Hz) and H-1′ ($^3J_{C,H}$ 3.1 Hz). The signal of C-5 displayed also a long-range correlation to H-2. The other methine signal, at δ_C 114.6, was assigned to C-4″. The signal displayed a one-bond coupling to H-4″ ($^1J_{C,H}$ 195.9 Hz) and further small long-range couplings giving rise to the doublet-multiplet splitting of the signal. The small couplings were not assigned. Based on the long-range correlation and coupling ($^3J_{C,H}$ 7.1 Hz) from H-4″, the most downfield signal at δ_C 141.0 was assigned to C-2″. Finally, the last imidazole ring carbon C-5″ was observed at δ_C 133.1 displaying a correlation with H-4″. The signal showed also a doublet-multiplet splitting, from which the doublet ($^2J_{C,H}$ 12.6 Hz) was due to the geminal coupling with H-4″. The smaller couplings originate presumably from the protons of the hydroxymethyl group. The signal of the hydroxymethyl carbon was observed at δ_C 53.9 ppm, split into a triplet ($^1J_{C,H}$ 143.2 Hz). The hydroxymethyl protons displayed a long-range correlation to both C-4″ and C-5″. Altogether, the chemical shifts for C-4″, C-5″ and the hydroxymethyl carbon are close to those observed for 4(5)-(hydroxymethyl)imidazole.[13,14]

Analogously to **2a** and **2b** and by taking into account the pK_a-value for 4(5)-(hydroxymethyl)imidazole (6.54),[9] **2d** should exist primarily in the protonated form and the prototropic equilibrium would be suppressed.

Mass spectrometric characterisation of 2a–2d

In the positive-mode ESI mass spectra, the protonated molecular ion together with a characteristic fragment ion corresponding to the loss of the ribosyl unit (followed by protonation at N-1) were observed for all the compounds **2a**–**2d**. The elemental composition of the compounds was finally verified by exact mass measurement. The observed protonated molecular ions were consistent with the calculated values and these are reported in the Experimental section.

Conclusions

The results of our work show that etheno- and substituted ethenoadenosines are unstable when stored in slightly basic solutions and that the major degradation products are the corresponding bi-imidazoles. In particular, compounds **1b** and **1c** are very easily converted into the corresponding bi-imidazoles. This shows that in double-stranded DNA, simultaneously with the adducts **1a**–**1d** the corresponding bi-imidazoles might also occur. This has to be taken into account when the mutagenic effects of the adducts are considered.

In this work, the structural identification of the products was of crucial importance. Thus, great emphasis was laid on the complete interpretation of the spectra of all products. The results show that long-range correlation spectroscopy and fully proton-coupled ^{13}C NMR spectroscopy are invaluable tools for the elucidation of proton-poor structures.

Experimental

Chromatographic methods and isolation of products

The starting materials **1a**–**1c** were obtained by the reaction of adenosine with chloroacetaldehyde (**1a**),[15] bromomalonaldehyde (**1b**)[16] and mucochloric acid (**1c**).[6] Compound **1d** was obtained by reduction of **1b** with $NaBH_4$ in ethanol solution according to the method described by Nair *et al.*[16] Compounds

1a–**1d** were separated and purified by column chromatography on a 4 cm × 4 cm column of preparative C_{18} bonded silica grade (40 μm, Bondesil, Analytichem International, Harbor City, CA, USA). Water and mixtures of water–acetonitrile were used as the mobile phases. Analytical HPLC was performed on a Kontron Instruments liquid chromatographic system consisting of a model 322 pump, a 440 diode-array detector (UV), and a KromaSystem 2000 data-handling program (Kontron Instruments S.P.A., Milan, Italy). The reaction mixtures were chromatographed on a 5 μm, 4 mm × 125 mm reversed-phase C_{18} analytical column (Spherisorb ODS2, Hewlett-Packard, Espoo/Esbo, Finland). The column was eluted isocratically for 5 min with water or potassium dihydrogen phosphate (0.01 M; pH 4.6) and then with gradient of acetonitrile from 0% to 30% over the course of 25 min at a flow rate of 1 ml min^{-1}. The final products (**2a**–**2d**) were precipitated from the concentrated reaction mixtures by the addition of ethanol or mixtures of ethanol–diethyl ether. Final purification was performed with column chromatography on C_{18} reversed-phase material as described above. Analytical samples of **2a**–**2d** were reprecipitated from water by the addition of ethanol–diethyl ether or acetone.

Spectroscopic and spectrometric methods

^1H and ^{13}C NMR spectra were recorded at 30 °C on a JEOL JNM-A500 Fourier transform NMR spectrometer at 500 and 125 MHz, respectively. The samples were dissolved in DMSO-d_6 or in DMSO-d_6 containing two drops of TFA. Signals were referenced internally to DMSO-d_6 (2.50 ppm for ^1H and 39.50 ppm for ^{13}C). The ^1H NMR signal assignments were based on proton chemical shifts and 2D homonuclear H–H correlation (COSY) experiments. The chemical shifts and the coupling constants of the multiplets of the proton signals in the ribose units of **2a**–**2d** were calculated using the Perch program.[17] The assignment of ^{13}C signals was based on carbon chemical shifts, 2D heteronuclear ^1H–^{13}C correlation experiments (HETCOR and COLOC) and fully proton-coupled carbon spectra. The multiple-bond proton–carbon couplings were obtained by selective decoupling of the protons in the imidazole moieties and the H-1′ in the ribosyl unit. All the coupling constants (J) are given in Hz. The positive-mode electrospray ionisation (ESI) mass spectra were recorded on a Fisons ZABSpec-oaTOF spectrometer (Manchester, UK). Ionisations were carried out using nitrogen as both nebulising and bath gas. A potential of 8.0 kV was applied to the ESI needle. The temperature of the pepperpot counter-electrode was 90 °C. The samples were introduced by loop injection at a flow rate of 20 μl min^{-1} (water–CH_3CN–AcOH 80 : 20 : 1). PEG 200 was used as standard for the exact mass determinations. The mass spectrometer had a resolution of 7000. UV spectra were recorded with a Shimadzu UV-160 A spectrophotometer (Shimadzu Europe, Dusseldorf, Germany).

5-Amino-4-(imidazol-2″-yl)-1-(β-D-ribofuranosyl)-1*H*-imidazole 2a

Ethenoadenosine **1a** (0.2238 g, 0.768 mmol) was dissolved in 100 ml of 0.01 M NaOH solution. The flask was set in an oil-bath (60 °C) and the reaction was followed by HPLC. After 8 h no peak for the starting material could be detected; however, the reaction was allowed to proceed for an additional 2 h. After cooling of the reaction mixture, the pH was adjusted to 5. The mixture was concentrated by rotatory evaporation to a volume of ≈20 ml. Ethanol–diethyl ether (1 : 3) was added until the formation of a greenish precipitate. The flask was stored in a refrigerator for 5 h and the green precipitate was filtered off. The slightly green solution was concentrated again to a volume of 5 ml and ethanol–diethyl ether (1 : 4) was added until the solution became cloudy. The solution was stored overnight in a refrigerator and the precipitate was collected by filtration and dried in a vacuum desiccator over diphosphorus pentaoxide to give title compound **2a** (0.1094 g, 50.6%) as a white powder. An analytical sample was reprecipitated from water–ethanol–diethyl ether (1 : 2 : 4) to give **2a** as a white powder, mp 224–228 °C (decomp.) [lit.,[2b] 225–227 °C (decomp.)]; $\lambda_{max}(H_2O)$/nm 213, 278; λ_{min}/nm 230; δ_H(500 MHz; DMSO-d_6), see Table 1; δ_C(125 MHz; DMSO-d_6), see Table 2; *m*/*z* (ESI) 282 (100%, MH^+), 150 (7%, MH^+ − ribosyl + H) [HRMS: Calc. for ($C_{11}H_{15}N_5O_4$ + H): *m*/*z*, 282.1202. Found: *m*/*z*, 282.1202].

5-Amino-4-[4″(5″)-formylimidazol-2″-yl]-1-(β-D-ribofuranosyl)-1*H*-imidazole 2b

Formylethenoadenosine **1b** (0.0837 g, 0.262 mmol) was decomposed in 40 ml of 0.01 M NaOH as described for **1a**. After 2 h, the reaction had gone to completion and the pH of the mixture was adjusted to 5. The mixture was concentrated to a volume of 10 ml and ethanol–diethyl ether (1 : 3) was added until a dark powder started to precipitate. The mixture was stored in a refrigerator overnight and the precipitate was collected by filtration. The crystalline material was re-dissolved in water and purified by column chromatography. Fractions containing the product were combined, evaporated to dryness and dried in a vacuum desiccator over diphosphorus pentaoxide to give the title compound **2b** (0.0640 g, 78.9%) as a yellow powder. An analytical sample was reprecipitated from water–acetone (1 : 3) to give **2b** as yellow powder, mp > 230 °C; $\lambda_{max}(H_2O)$/nm 245 and 326; λ_{min}/nm 212 and 297; δ_H(500 MHz; DMSO-d_6 + TFA), see Table 1; δ_C(125 MHz; DMSO-d_6 + TFA), see Table 2; *m*/*z* (ESI) 310 (72%, MH^+), 178 (100, MH^+ − ribosyl + H) [HRMS: Calc. for ($C_{12}H_{15}N_5O_5$ + H): *m*/*z*, 310.1151. Found: *m*/*z*, 310.1162].

5-Amino-4-[4″(5″)-oxaloimidazol-2″-yl]-1-(β-D-ribofuranosyl)-1*H*-imidazole 2c

Oxaloethenoadenosine **1c** (0.0505 g, 0.139 mmol) was decomposed in 40 ml of 0.01 M NaOH as described for **1a**. After 3 h, the reaction had gone to completion and the pH of the mixture was adjusted to 5. The mixture was concentrated to a volume of 10 ml and acetone was added until a yellow powder started to precipitate. The mixture was stored in a refrigerator overnight and the solution was decanted. The yellow precipitate was re-dissolved in water and purified by column chromatography. The fractions containing the product were combined, evaporated to dryness and dried in a vacuum desiccator over diphosphorus pentaoxide to give the title compound **2c** as a yellow powder (0.0482 g, 98.1%). An analytical sample was reprecipitated from water–ethanol–diethyl ether (1 : 2 : 4) to give **2c** as a yellow powder, mp > 230 °C; $\lambda_{max}(H_2O)$/nm 254 and 339; λ_{min}/nm 214 and 300; δ_H(500 MHz; DMSO-d_6 + TFA), see Table 1; δ_C(125 MHz; DMSO-d_6 + TFA), see Table 2; *m*/*z* (ESI) 354 (53%, MH^+), 222 (100, MH^+ − ribosyl + H) [HRMS: Calc. for ($C_{13}H_{15}N_5O_7$ + H): *m*/*z*, 354.1050. Found: *m*/*z*, 354.1076].

5-Amino-4-[4″(5″)-(hydroxymethyl)imidazol-2″-yl]-1-(β-D-ribofuranosyl)-1*H*-imidazole 2d

(Hydroxymethyl)ethenoadenosine **1d** (0.0734 g, 0.228 mmol) was decomposed in 30 ml of 0.01 M NaOH as described for **1a**. After 5 h, the reaction had gone to completion and the pH of the mixture was adjusted to 5. The mixture was concentrated and the product was precipitated by the addition of acetone. The precipitate was filtered off, re-dissolved in water and purified by column chromatography. The fractions containing the product were combined, evaporated to dryness and dried in a vacuum desiccator over diphosphorus pentaoxide to give the title compound **2d** as a pale yellow powder (0.0487 g, 68.5%). An analytical sample was reprecipitated from water–ethanol–diethyl ether (1 : 2 : 4) to give **2d** as a pale yellow powder, mp > 230 °C; $\lambda_{max}(H_2O)$/nm 212 and 270; λ_{min}/nm 228; δ_H(500

MHz; DMSO-d_6), see Table 1; δ_C(125 MHz; DMSO-d_6), see Table 2; *m*/*z* (ESI) 312 (100%, MH^+), 180 (22, MH^+ − ribosyl + H) [HRMS: Calc. for ($C_{12}H_{17}N_5O_5$ + H): *m*/*z*, 312.1308. Found: *m*/*z*, 312.1311].

Acknowledgements

We thank Mr Markku Reunanen for performing the mass spectrometric analyses.

References

1 N. J. Leonard, *CRC Crit. Rev. Biochem.*, 1984, **15**, 125.
2 (*a*) K. F. Yip and K. C. Tsou, *Tetrahedron Lett.*, 1973, 3087; (*b*) K. C. Tsou, K. F. Yip, E. E. Miller and K. W. Lo, *Nucleic Acids Res.*, 1974, **1**, 531; (*c*) K. F. Yip and K. C. Tsou, *J. Org. Chem.*, 1975, **40**, 1066.
3 A. K. Basu, M. L. Wood, L. J. Niedernhofer, L. A. Ramos and J. M. Essigmann, *Biochemistry*, 1993, **32**, 12793.
4 L. Kronberg, R. Sjöholm and S. Karlsson, *Chem. Res. Toxicol.*, 1992, **5**, 852.
5 L. Kronberg, S. Karlsson and R. Sjöholm, *Chem. Res. Toxicol.*, 1993, **6**, 495.
6 J. Mäki, R. Sjöholm and L. Kronberg, *J. Chem. Soc.*, *Perkin Trans. 1*, 1999, 2923.
7 F. Le Curieux, T. Munter and L. Kronberg, *Chem. Res. Toxicol.*, 1997, **10**, 1181.
8 S. Steiner, A. E. Crane and W. P. Watson, *Carcinogenesis*, 1992, **13**, 119.
9 M. R. Grimmett, in *Comprehensive Heterocyclic Chemistry. The Structure, Reactions, Synthesis and Uses of Heterocyclic Compounds*, ser. ed. A. R. Katritzky and C. W. Rees, Pergamon Press, Oxford, 1984, vol. 5 (ed. K. T. Potts), p. 373.
10 P. D. Sattsangi, J. R. Barrio and N. J. Leonard, *J. Am. Chem. Soc.*, 1980, **102**, 770.
11 K.-E. Stensiö, K. Wahlberg and R. Wahren, *Acta Chem. Scand.*, 1973, **27**, 2179.
12 J.-W. Kim, S. M. Abdelaal, L. Bauer and N. E. Heimer, *J. Heterocycl. Chem.*, 1995, **32**, 611.
13 P. Aulaskari, M. Ahlgrén, J. Rouvinen, P. Vainiotalo, E. Pohjola and J. Vepsäläinen, *J. Heterocycl. Chem.*, 1996, **33**, 1345.
14 A. R. Katritzky and K. W. Law, *Magn. Reson. Chem.*, 1988, **26**, 129.
15 J. R. Barrio, J. A. Secrist III and N. J. Leonard, *Biochem. Biophys. Res. Commun.*, 1972, **46**, 597.
16 V. Nair, R. J. Offerman and G. A. Turner, *J. Org. Chem.*, 1984, **49**, 4021.
17 The PERCH Software is distributed by PERCH Project, Department of Chemistry, University of Kuopio, Kuopio, Finland. For details of the program, see: R. Laatikainen, M. Niemitz, U. Weber, J. Sundelin, T. Hassinen and J. Vepsäläinen, *J. Magn. Reson.*, *Ser. A*, 1996, **120**, 1.