

Synthesis of [60]fullerene—glycopyranosylaminopyrimidin-4-one conjugates

Carina I.C. Jordão^a, Andreia S.F. Farinha^a, Roger F. Enes^a, Augusto C. Tomé^{a,*},
Artur M.S. Silva^a, José A.S. Cavaleiro^a, Catarina I.V. Ramos^a, M.G. Santana-Marques^a,
Filipe A. Almeida Paz^b, José M. de la Torre Ramirez^c, Maria D.L. de la Torre^c,
Manuel Nogueras^{c,*}

^a Departamento de Química, Universidade de Aveiro, 3810-193 Aveiro, Portugal

^b Departamento de Química, CICECO, Universidade de Aveiro, 3810-193 Aveiro, Portugal

^c Departamento de Química, Inorgánica y Orgánica, Facultad de Ciencias Experimentales, Universidad de Jaén, 23071 Jaén, Spain

Received 7 January 2008; received in revised form 12 February 2008; accepted 16 February 2008

Available online 21 February 2008

Abstract

The synthesis of several C₆₀ derivatives containing a 6-(β-D-glycopyranosylamino)pyrimidin-4-one unit and a C₆₀—uridine conjugate is described. The fullerene derivatives bearing a 4-(β-D-glycopyranosylamino)pyrimidin-4-one moiety were synthesised by 1,3-dipolar cycloaddition reactions of C₆₀ with azomethine ylides generated in situ from the corresponding 5-formylpyrimidin-4-one derivatives and *N*-methylglycine. The synthesis of the C₆₀—uridine conjugate involved the selective protection of the 2'- and 3'-hydroxyl groups of uridine, esterification, cyclopropanation of C₆₀ and, finally, the deprotection of the hydroxyl groups. One of the fullerene—glycopyranosylaminopyrimidin-4-one conjugates was characterised by single-crystal X-ray crystallography. Differentiation between pairs of diastereoisomers, for several fullerene derivatives, was achieved through the study of their gas-phase fragmentations.

© 2008 Elsevier Ltd. All rights reserved.

Keywords: Fullerene derivatives; Nucleosides; X-ray crystallography; Mass spectrometry

1. Introduction

Since the beginning of the large-scale production of C₆₀ in 1990, a wide variety of C₆₀ derivatives have been synthesised. Many of them exhibit interesting electronic properties¹ or biological activities. These compounds have potential biomedical application such as HIV-1 protease inhibitors, neuroprotective agents, photosensitisers for photodynamic therapy, X-ray and MRI contrast agents, etc.² The search for new biologically active fullerene derivatives is still a rewarding challenge.³ Following our interest in the synthesis of fullerene derivatives with

potential biological activity,⁴ we decided to prepare a series of fulleropyrrolidine—glycopyranosylaminopyrimidin-4-one conjugates and a C₆₀—uridine conjugate. 6-(β-D-Glycopyranosylamino)pyrimidin-4-ones, such as compounds **1–6**, are potential biologically active compounds.⁵ They can be considered as nucleoside analogues where the anomeric carbon of the carbohydrate is bonded to the 6-amino group and not to a pyrimidine nitrogen atom.

Reports on the synthesis of C₆₀—nucleoside conjugates are scarce but two examples have been published recently.^{6,7} In one case, nitrogen-bridged [60]fullerene—3'-deoxythymidine conjugates were prepared from the reaction of 3'-azido-3'-deoxythymidine (AZT) with C₆₀.⁶ Other studies involved the synthesis of a C₆₀—guanosine conjugate, which was used to build porphyrin—fullerene^{7a} and phthalocyanine—fullerene^{7b} dyads assembled through Watson—Crick hydrogen bonds.

* Corresponding authors. Tel.: +351 234 370 712; fax: +351 234 370 084.

E-mail addresses: actome@ua.pt (A.C. Tomé), mmontiel@ujaen.es (M. Nogueras).

In this study, we took advantage of the presence of a formyl group in the glycopyranosylaminopyrimidin-4-one derivatives to generate azomethine ylides, which were trapped with C_{60} yielding the corresponding fulleropyrrolidine derivatives.

2. Results and discussion

2.1. Fulleropyrrolidine–glycopyranosylaminopyrimidin-4-one conjugates

Compounds **7–12** were obtained in one step from 1,3-dipolar cycloaddition reactions of C_{60} with azomethine ylides **Y**, which were generated in situ by the reaction of the 5-formylpyrimidin-4-one derivatives **1–6**^{5a} with *N*-methylglycine in refluxing toluene (Scheme 1).⁸ Fullerene derivatives **7–12** were obtained in moderate yields as mixtures of two diastereoisomers, which were separated by column chromatography on silica gel. The diastereoisomers, which differ in the configuration of the pyrrolidine carbon C-2, were identified as **a** (higher R_f on silica) and **b** (lower R_f). Typically the main isomer is the one with lower R_f value on silica, except for compound **12** (Table 1). Each diastereoisomer was characterised by 1H and ^{13}C NMR and MS, and the structure of **10a** was unequivocally determined by means of single-crystal X-ray diffraction studies.

2.1.1. NMR studies

The 1H NMR spectra of the conjugates **7–12** show the typical resonance of the N(1)–Me protons as singlets at δ 2.68–2.79 ppm; compound **12b** is an exception because the resonance of those protons appears at δ 2.91 ppm (Table 2). The resonance corresponding to proton H-2 appears as a singlet at δ 5.51–5.61 ppm while the resonances of the two diastereotopic protons H-5⁹ appear as doublets at δ 4.19–4.46 (H-5a) and 4.91–5.00 ppm (H-5b), with a coupling constant of 9.5–9.7 Hz. Surprisingly, in the spectrum of diastereoisomer **8b** the resonance corresponding to H-5b appears at a much higher field (δ 4.24 ppm) than that observed for other compounds. The resonance of the 6'-NH proton appears as a doublet at δ 9.14–9.81 ppm with a coupling constant in the range of 8.9–10.8 Hz.

In the ^{13}C NMR spectra of the conjugates **7–12**, the resonance of the C-1'' carbon of the sugar moiety appears at δ 79.3–80.6 ppm while the resonances of the C_{60} -sp³ carbons C-3 and C-4 appear, respectively, at δ 75.8–76.3 and δ 76.2–77.6 ppm.

Table 1

Yields of the fulleropyrrolidine derivatives **7–12** and the monoisotopic masses of the corresponding $[M+H]^+$ ions

Compound	X	R ¹	R ²	Yield (%)	Diastereoisomers ratio (a/b)	$[M+H]^+$ ions (Da)
1	7	S	Me	38	2:3	1205
2	8	O	Me	24	4:5	1189
3	9	O	Me	25	5:6	1261
4	10	S	Me	31	2:3	1277
5	11	S	H	35	2:3	1263
6	12	S	H	27	3:2	1191

Table 2

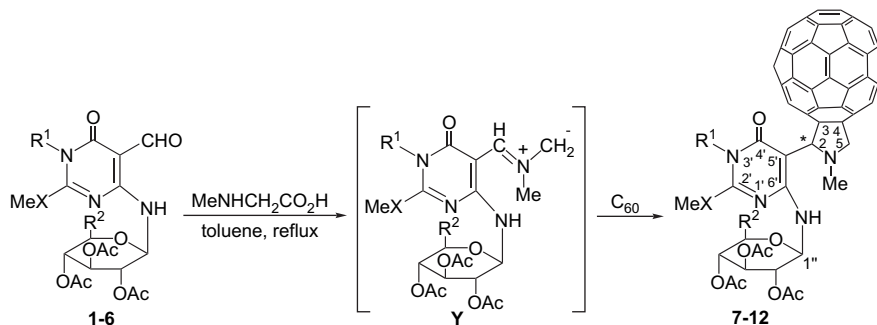
Selected 1H NMR spectroscopic data for the fulleropyrrolidine derivatives **7–12**

Compound	N(1)–Me	H-2	H-5a (J/Hz)	H-5b (J/Hz)	NH (J/Hz)
7a	2.79	5.53	4.25 (9.7)	4.99 (9.7)	9.15 (10.5)
7b	2.69	5.60	4.19 (9.5)	4.97 (9.5)	9.41 (9.2)
8a	2.78	5.51	4.24 (9.5)	4.99 (9.5)	9.14 (10.4)
8b	2.68	5.58	4.19 (9.5)	4.24 (9.5)	9.38 (9.2)
9a	2.75	5.51	4.24 (9.6)	4.99 (9.6)	9.19 (10.4)
9b	2.69	5.59	4.19 (9.5)	4.98 (9.5)	9.55 (9.5)
10a	2.77	5.53	4.25 (9.7)	5.00 (9.7)	9.21 (10.3)
10b	2.70	5.61	4.20 (9.6)	4.99 (9.6)	9.58 (9.2)
11a	2.77	5.52	4.26 (9.6)	4.99 (9.6)	9.36 (10.8)
11b	2.70	5.59	4.46 (9.6)	4.98 (9.6)	9.71 (9.3)
12a	2.86	5.57	4.19 (9.6)	4.97 (9.6)	9.55 (8.9)
12b	2.91	5.59	4.32 (9.7)	4.91 (9.7)	9.81 (9.8)

2.1.2. MS studies

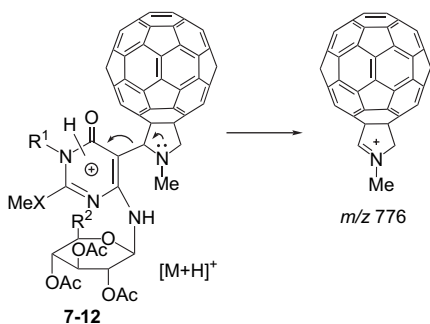
The molecular mass of compounds **7–12** was initially determined by Liquid Secondary Ion Mass Spectrometry in an EBEQ mass spectrometer. All spectra showed the $[M+H]^+$ ions and a fragment of m/z 720 corresponding to the intact fullerene moiety. However, in order to differentiate between each pair of diastereoisomers, as high collision energy conditions are usually insensitive to stereochemical differences,¹⁰ instruments with low energy collision regimes (Q-ToF and Ion Trap) and another ionisation technique (electrospray) were used. Using the latter instruments, the spectra obtained for all compounds show predominantly the protonated molecules $[M+H]^+$.

The $[M+H]^+$ ions were selected and subjected to low energy collisions in the same experimental conditions, both with the Q-ToF and with the Ion Trap instruments. For each compound, two diagnostic fragment ions are obtained: the ion formed by elimination of acetic acid from the glycosidic



Scheme 1.

moiety, $[(M+H)-AcOH]^+$, and the ion with m/z 776 formed by loss of a neutral fragment through heterolytic cleavage of the C-2–C-5' carbon bond (Scheme 2). Under these conditions, the intact fullerene moiety is not observed in almost all spectra and, when present, its relative abundance is low.



Scheme 2.

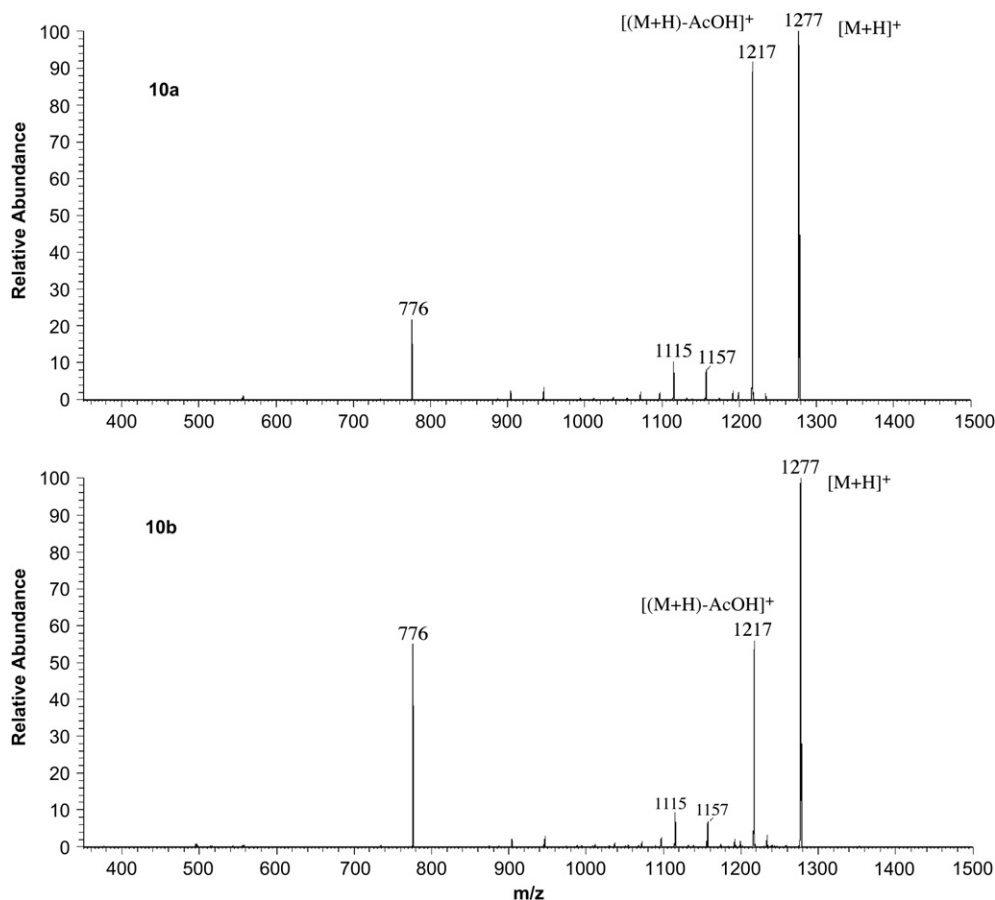
Differentiation of diastereoisomers by collision induced fragmentation in mass spectrometry is not very common and it usually relies on the formation of adducts, complexes or derivatised species, either in situ^{11,12} or prior to the introduction in the mass spectrometer^{13,14} and, when applied directly to protonated molecules, multiple stage mass analysis MS^n ($n > 2$)¹⁵

is usually involved. In the present case, the relative abundances of the two diagnostic ions allow a rapid and easy differentiation of the diastereoisomers as it can be seen in Figure 1 for the pair **10a** and **10b**.

For the diastereoisomers **7a–12a**, the most abundant fragment ions are formed by losses of AcOH, whereas for the diastereoisomers **7b–12b**, the formation of the ion with m/z 776 competes favourably with acetic acid loss. To our knowledge, this is the first case where diastereoisomers of fullerene derivatives are differentiated by MS^2 of their protonated molecules. The increased relative abundance of the m/z 776 ion for the diastereoisomers **7b–12b**, when compared with their **7a–12a** counterparts, may be due to a less favourable configuration of the former due to the bulky leaving neutral group being closer to the fullerene cage.

2.1.3. Single-crystal X-ray crystallographic studies

To the best of our knowledge, and according to a search in the Cambridge Structural Database (CSD, Version 5.28—August 2007),^{16,17} compound **10a** is the first C_{60} derivative with a nucleoside-type unit. Moreover, only a couple of related nucleoside crystal structures are available in the literature, namely those described by Ferguson et al. for 6-methyl-5-methoxy-3-(2,3,4,6-tetra-*O*-acetyl- β -D-glucopyranos-1-yl)-3*H*-1,2,3-triazolo[4,5-*d*]pyrimidin-7(6*H*)-one¹⁸ and 2-methylthio-5-nitroso-

Figure 1. Product ion spectra (MS^2) of the $[M+H]^+$ ions of diastereoisomers **10a** and **10b**.

6-*N*-(2,3,4-tri-*O*-acetyl- β -D-xylopyranosyl)aminopyrimidin-4(3*H*)-one.¹⁹

Compound **10a** crystallises in the orthorhombic non-centrosymmetric and chiral $P2_12_12_1$ space group, in a similar way as described for the couple nucleoside units available in the literature.^{18,19} The asymmetric unit is composed of a single molecule of **10a** plus an undetermined number of solvent molecules, which were found to be composed of partially occupied chloroform and toluene. Indeed, even though difference Fourier maps allowed to partially unveil the location of these crystallisation chemical moieties, full structural refinement was found to be rather unstable and their corresponding contribution to the electron density was ultimately removed in order to reach the sensible structural model reported here (see details in the appropriate section). Figure 2 depicts a view of the molecular unit of **10a**, with all non-hydrogen atoms being represented as thermal ellipsoids (50% of probability). The β configuration of the anomeric carbon atom of the substituted glucopyranosyl ring was known a priori, and in the crystal structure of **10a** this ring significantly approaches the near-ideal 4C_1 chair conformation with all the carbonyl groups of the acetyl substituents being *cis* with respect to the adjacent C–H groups, either of the six-membered ring or of the –CH₂– substituent (Fig. 2). The puckering parameters calculated after Cremer and Pople²⁰ for this six-membered

ring are: $Q=0.619$ Å, $q_2=0.089$ Å, $q_3=0.613$ Å, $\varphi_2=6.7^\circ$ and $\theta=8.3^\circ$. These findings further support the presence of a slightly distorted chair conformation for the substituted glycopyranosyl ring since, on the one hand, the puckering amplitude q_3 is significantly higher than q_2 and, on the other, the total puckering amplitude Q is only slightly lower than the calculated for the ideal cyclohexane chair (0.63 Å based on an ideal C–C interatomic distance of 1.54 Å). Interestingly, the magnitude of the distortion in **10a** reflects the crystal packing arrangement of individual molecular units (see below), imposing significant strain in the glycopyranosyl ring, which is ultimately described by a value of θ (8.3°) greater than the one for the ideal cyclohexane chair conformation (calculated as ca. 5°).²⁰ The pyrimidine ring is, as expected, almost planar with the six defining atoms being in average 0.013 Å displaced from the plane of the ring. The substituent methyl and methylthio groups are approximately co-planar with the pyrimidine ring, exhibiting dihedral angles of about 1.9 and 5.7° , respectively.

Following the assumptions from the synthesis, the C₆₀ derivatives **10** were obtained as a mixture of two diastereoisomers (see Scheme 1) with carbon 2 exhibiting distinct configurations. The crystal structural determination of compound **10a** clearly supports the presence of configuration *R* for this chirality centre.

As also pointed out by Cobo et al.,¹⁹ in **10a** the bridging –NH– moiety of the glucopyranosylaminopyrimidin-4-one unit establishes a rather strong [$d(D\cdots A)$ of 2.733(3) Å] and reasonably directional [$\angle(DHA)$ of $140(2)^\circ$] intramolecular N–H $\cdots$ N hydrogen-bonding interaction with the –NMe– moiety belonging to the C2 $\rightarrow$ C5 ring bound to C₆₀ (not shown). However, in the particular structure of **10a**, it is also of considerable importance to emphasise that the –NH– moiety is further involved in a secondary, but weaker, hydrogen-bonding interaction with a neighbouring substituent acetyl group of the glucopyranosyl ring: N–H $\cdots$ O with $d(D\cdots A)$ of 3.054(3) Å and $\angle(DHA)$ of $119(2)^\circ$. These two interactions can also be described as *S*(5) and *S*(6), respectively, graph set motifs.²¹ It is thus feasible to assume that if the –NH– group was not directly involved in bifurcated hydrogen bonds, the former interaction could be significantly more linear and, consequently, stronger. In summary, the mutual spatial arrangement between the glucopyranosyl and pyrimidine rings, which ultimately subtend an approximate dihedral angle of about 54° between the average planes containing the rings, is such that, on one hand, it maximises the two aforementioned hydrogen-bonding interactions and, on the other hand, it minimises the steric repulsion between all functional groups composing the molecular unit of **10a**.

Individual molecules of **10a** close pack in the solid state along the [100] crystallographic direction, forming infinite tapes following the symmetry of the 2₁ screw axes (Fig. 3a). Noteworthy is the fact that while the C₆₀ moieties point outwards, the glucopyranosylaminopyrimidin-4-one units constitute the main core of these tapes, with a significant number of weak C–H $\cdots$ (N,O) hydrogen-bonding interactions establishing the physical links between neighbouring moieties. These supramolecular tapes are distributed in the crystal structure of compound **10a** in a typical herringbone fashion (Fig. 3c), most certainly mediated by a significant number of

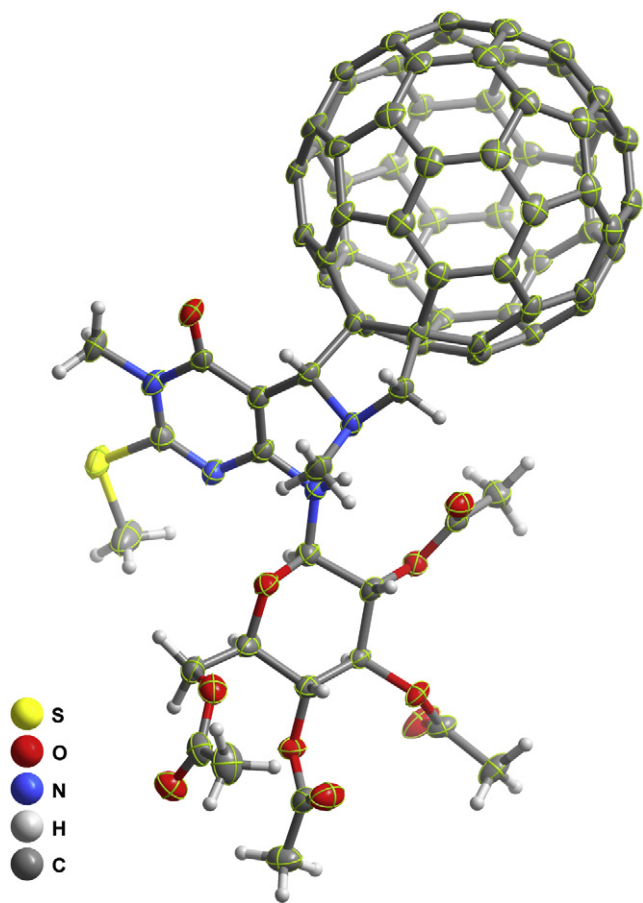


Figure 2. Molecular unit of the crystal structure of compound **10a**. Non-hydrogen atoms are represented as thermal ellipsoids drawn at the 50% probability level and hydrogen atoms as small spheres with arbitrary radii.

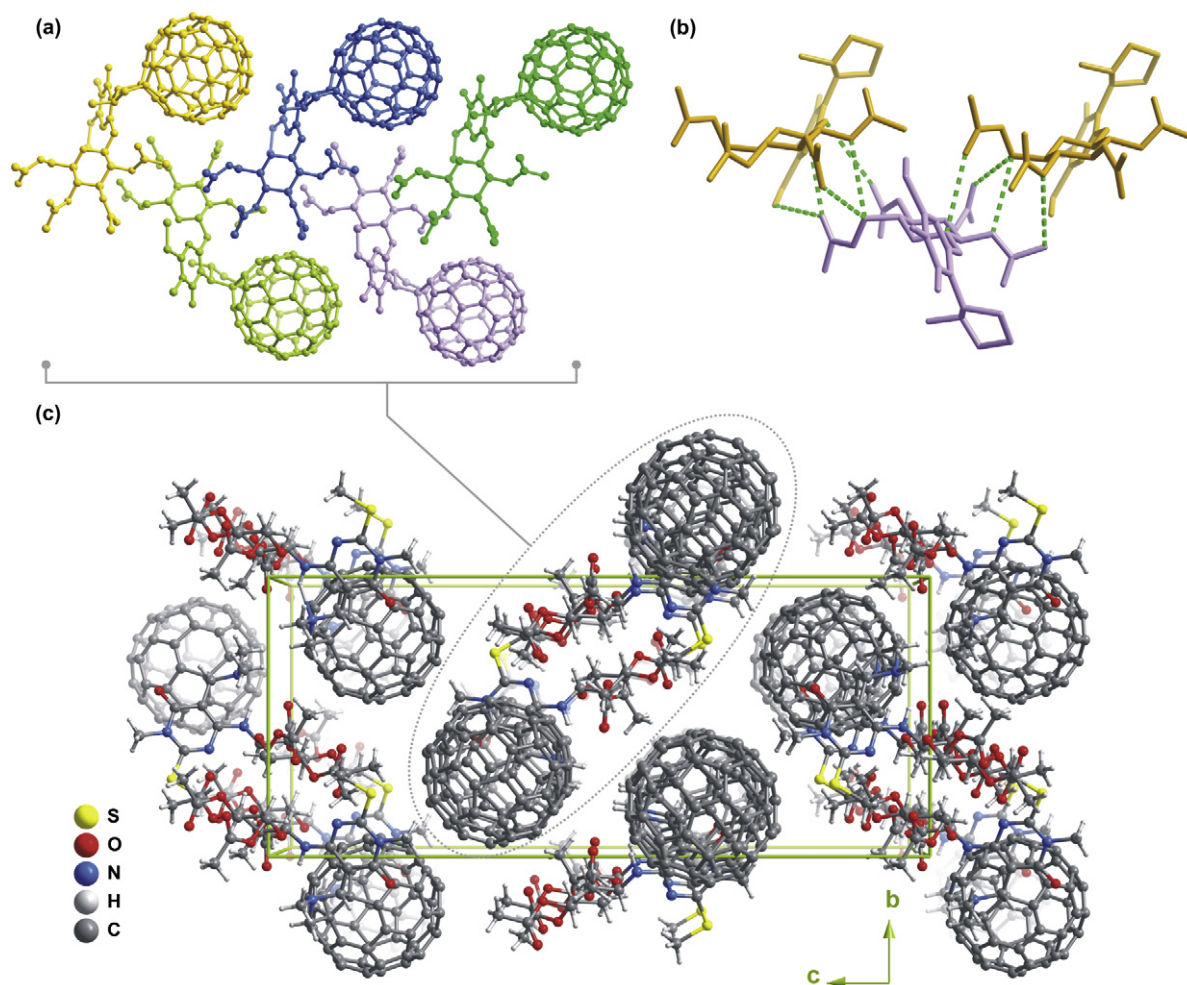


Figure 3. (a) and (b) Parallel distribution of individual molecular units of **10a**, leading to the formation of a supramolecular tape assembled by a series of weak but cooperative C–H⋯(N,O) hydrogen-bonding interactions (green dashed lines) involving adjacent glucopyranosylaminopyrimidin-4-one units. (c) Crystal packing viewed in perspective along the [100] crystallographic direction.

weak interactions involving the solvent molecules of crystallisation (which could not be located), plus C–H⋯ π contacts. For further details on crystal solution and refinement see Section 4.3 below.

2.2. Fullerene–uridine conjugates

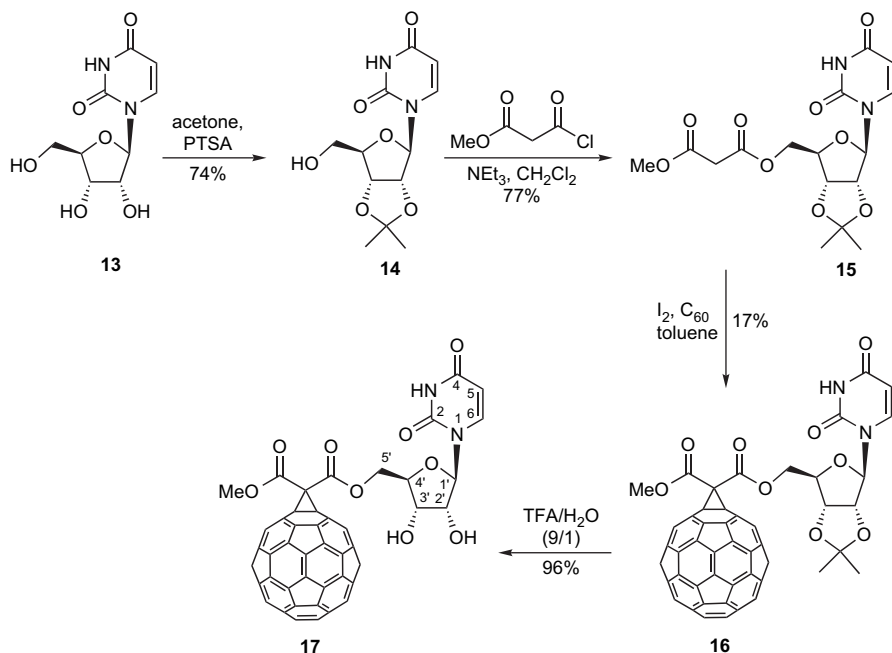
The synthesis of the fullerene–uridine conjugates **16** and **17** involved, as the first step, the protection of the 2'- and 3'-hydroxyl groups of uridine (**13**) by reaction with acetone and using *p*-toluenesulfonic acid (PTSA) as the catalyst (Scheme 3). 2',3'-*O*-Isopropylideneuridine (**14**),²² obtained in 74% yield, was then reacted with methyl 3-chloro-3-oxopropanoate to afford malonate **15** in 77% yield. This malonate was covalently linked to C₆₀, following a modification of Bingel's procedure,²³ to yield the conjugate **16**. This compound was purified by flash chromatography (silica gel) using toluene as the eluent and crystallised from chloroform/methanol yielding a black solid (17% yield). The yields of the cyclopropanation reactions of fullerene with malonates are typically low but, in this particular case, the reactivity of uridine derivatives with iodine and nucleophiles²⁴ may also be contributing

to the low yield of this reaction. Deprotection of the hydroxyl groups in dyad **16** with trifluoroacetic acid afforded dyad **17** in excellent yield (96%).

Compounds **14–17** were characterised by ¹H and ¹³C NMR and MS. The NMR spectra of these compounds confirm their structures and the mass spectra show the expected molecular ions. In the ¹H NMR spectra of compounds **14–17**, the resonances of protons H-5 and H-6 appear as doublets at δ 5.63–5.78 ppm, with a coupling constant of 8.0–8.4 Hz. The resonance corresponding to H-1' appears as a doublet at δ 5.67–5.82 ppm with a coupling constant of 2.0–2.3 Hz (for compound **17**, a higher coupling constant (5.7 Hz) is observed). In the ¹³C NMR spectra of compounds **14–16**, the resonances of C-5 and C-6 appear at δ 102.4–102.9 and 142.2–142.5 ppm, respectively. The resonance of carbon C-1' appears typically at δ 94.62–95.48 ppm.

3. Conclusions

Fullerene derivatives bearing a 4-(β -D-glycopyranosyl-amino)pyrimidin-4-one moiety can be easily synthesised from the corresponding 5-formylpyrimidin-4-one derivatives,



Scheme 3.

C_{60} and *N*-methylglycine. These reactions afford mixtures of two diastereoisomers, which can be separated by column chromatography. The study of the gas-phase fragmentations of these diastereoisomers allows their differentiation. Single-crystal X-ray crystallography of one of the fullerene–glycopyranosylaminopyrimidin-4-one conjugates (**10a**) shows that individual molecules close pack in the solid state forming infinite tapes. The main core of these tapes is constituted by the glycopyranosylaminopyrimidin-4-one units while the C_{60} moieties point outwards. These supramolecular tapes are distributed in the crystal structure of the compound in a typical herringbone fashion. The first C_{60} –uridine conjugate was prepared in five steps from commercial uridine.

4. Experimental

4.1. General remarks

^1H NMR spectra were recorded on a Bruker Avance 300 spectrometer at 300.13 MHz. ^{13}C NMR spectra were recorded on Bruker Avance 300 or Bruker Avance 500 spectrometer at 75.47 or 125.77 MHz, respectively. Chemical shifts (δ) are quoted in parts per million relative to TMS and the coupling constants (J) are expressed in hertz (Hz). Melting points were measured on a Reichert Thermovar apparatus fitted with a microscope and are uncorrected.

4.2. Mass spectrometric studies

LSIMS mass spectra were obtained with a VG Autospec Q (VG Analytical, Manchester, UK) instrument, equipped with a caesium ion gun (2 μA at 20 kV). The accelerating voltage was 8 kV. *m*-Nitrobenzyl alcohol was used as matrix. ESI

mass spectra were acquired with two different instruments: a Micromass Q-ToF 2 (Micromass, Manchester, UK) and a Finnigan LXQ ion trap (Thermo Finnigan, San Jose, CA, USA). Electrospray mass spectra were obtained from methanol/1% formic acid solutions. For the Q-ToF 2, source and desolvation temperatures were 80 and 150 $^\circ\text{C}$, respectively, and capillary voltage was 3000 V. The spectra were acquired at cone voltages of 35 V. Nebulisation and collision gases were N_2 and Ar, respectively. MS^2 spectra were acquired by selecting the $[\text{M}+\text{H}]^+$ ions with the quadrupole and using the hexapole as collision cell (between 40 and 80 eV). Ion trap experimental parameters for the ESI were: capillary temperature and voltage 350 $^\circ\text{C}$ and 1 V, respectively, tube lens voltage 40 V and nitrogen sheath gas 30 psi. MS^2 spectra were acquired by selecting and exciting the $[\text{M}+\text{H}]^+$ ions by standard isolation and excitation procedures.

4.3. Single-crystal X-ray crystallographic studies

A suitable single crystal of compound **10a** was mounted on a Hampton Research CryoLoop using FOMBLIN Y perfluoropolyether vacuum oil (LVAC 25/6) purchased from Aldrich,²⁵ with the help of a Stemi 2000 stereomicroscope equipped with Carl Zeiss lenses. Data were collected at 100(2) K on a Bruker X8 Kappa APEX II charge-coupled device (CCD) area-detector diffractometer (Cu $K\alpha$ graphite-monochromated radiation, $\lambda=1.54178$ Å) controlled by the APEX2 software package,²⁶ and equipped with an Oxford Cryosystems Series 700 cryostream monitored remotely using the software interface Cryopad.²⁷ Images were processed using the software package SAINT+,²⁸ and data were corrected for absorption by the multi-scan semi-empirical method implemented in SADABS.²⁹ The structure was solved using the direct methods implemented in

SHELXS-97,³⁰ which ultimately allowed the immediate location of the majority of the heaviest atoms, including those belonging to the C₆₀ moiety. The remaining non-hydrogen atoms were directly located from difference Fourier maps calculated from successive full-matrix least squares refinement cycles on F^2 using SHELXL-97.³¹ All non-hydrogen atoms were successfully refined using anisotropic displacement parameters. Friedel pairs (a total of 6590) have not been merged and were counted as independent data during the refinement stages.

Even though the vast majority of the hydrogen atoms attached to carbon were markedly visible in difference Fourier maps, they have been instead located at their idealised positions using appropriate *HFIX* instructions in SHELXL-97:³¹ 13 and 23 for the tertiary and secondary carbon atoms, respectively; 137 for the terminal methyl groups. In the case of the N–H moiety, the hydrogen atom was instead added at the position of the observable *Q* peak, with the corresponding N–H distance being restrained to 0.95(1) Å in order to ensure a chemically reasonable geometry for this moiety. All these hydrogen atoms were included in the final structural model in subsequent refinement cycles in riding-motion approximation with isotropic thermal displacement parameters (U_{iso}) fixed at 1.2 or 1.5 (only for the –CH₃ and –NH moieties) times U_{eq} of the carbon (or nitrogen) atom to which they are attached.

During crystal refinement using the original integrated and scaled dataset, it was clearly visible from difference Fourier maps that individual molecules of **10a** close pack in the solid state with a number of highly disordered solvent molecules, in particular an undetermined number of chloroform and toluene molecules. Attempts to model these solvent chemical moieties using a battery of crystallographic constraints proved to be unfruitful. Indeed, the significant smeared-out electron density arising from the extensive structural disorder associated with these moieties created instability during refinement, with corresponding very large *R* factors: $R1=0.1433$ and $wR2=0.3966$. The original dataset was then treated using the *SQUEEZE*³² subroutines provided with the software package *PLATON*^{33,34} in order to eliminate the contribution of the molecules present in the solvent-accessible area to the overall electron density. *PLATON* estimated that the unit cell contains ca. 1025 Å³ of potential solvent-accessible area (ca. 17% of the total volume), distributed mainly across two large (but symmetry-related) cavities centred at (0.0.934 3/4) and at (1/2 0.240 1/4), and containing a total of ca. 220 electrons. The resulting solvent-free reflection file produced by *PLATON* was used for subsequent refinement by full-matrix least squares on F^2 using SHELXL-97,³¹ which ultimately allowed a much more reliable and stable convergence. The Flack parameter was refined to 0.04(8),^{35,36} thus supporting the assumption of the presence of an enantiopure compound.

Information concerning crystallographic data collection and structural refinement details using the *SQUEEZE* (solvent-free) data is summarised in Table 3. Crystallographic data (excluding structural factors) for the structure reported in this paper have been deposited with the Cambridge Crystallographic Data Centre as supplementary publication no. CCDC-641402. Copies of the data can be obtained free of charge on application to CCDC,

Table 3

Crystal and *SQUEEZE* structural refinement data for fulleropyrrolidine derivative **10a**

Formula	C ₈₃ H ₃₂ N ₄ O ₁₀ S
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
<i>a</i> /Å	12.5427(4)
<i>b</i> /Å	14.1930(5)
<i>c</i> /Å	33.3711(10)
Volume/Å ³	5940.7(3)
<i>Z</i>	4
<i>D_c</i> /g cm ^{−3}	1.428
μ (Mo K α)/mm ^{−1}	0.129
Crystal size/mm	0.45×0.45×0.40
Crystal type	Black blocks
θ Range	3.52–29.13
Index ranges	−15≤ <i>h</i> ≤15, −19≤ <i>k</i> ≤19, −45≤ <i>l</i> ≤45
Reflections collected	102,789
Independent reflections	14,844 ($R_{\text{int}}=0.0609$)
Completeness to $\theta=29.13^\circ$	94.2%
Final <i>R</i> indices [$I>2\sigma(I)$] ^{a,b}	$R1=0.0546$, $wR2=0.1163$
Final <i>R</i> indices (all data) ^{a,b}	$R1=0.0806$, $wR2=0.1250$
Weighting scheme ^c	$m=0.0640$, $n=0$
Largest diff. peak and hole	0.441 and −0.231 e Å ^{−3}

$$^a R1 = \sum ||F_o| - |F_c|| / \sum |F_o|.$$

$$^b wR2 = \sqrt{\sum [w(F_o^2 - F_c^2)^2] / \sum [w(F_o^2)^2]}.$$

$$^c w = 1/[\sigma^2(F_o^2) + (mP)^2 + nP] \text{ where } P = (F_o^2 - 2F_c^2)/3.$$

12 Union Road, Cambridge CB2 2EZ, UK. Fax: +44 1223 336033; E-mail: deposit@ccdc.cam.ac.uk.

4.4. Synthesis of the fulleropyrrolidine derivatives 7–12: general procedure

A solution of C₆₀ (100 mg, 0.14 mmol), *N*-methylglycine (18 mg, 0.20 mmol) and 5-formylpyrimidin-4-one^{5a} (0.068 mmol) in toluene (100 mL) was heated at reflux under N₂ for 10 h. A new portion of *N*-methylglycine (9 mg, 0.10 mmol) was added to the reaction mixture and it was refluxed for 10 h more. This last procedure was repeated twice.

The reaction mixture was concentrated and purified by column chromatography using a gradient of toluene to toluene/ethyl acetate (8:2) as the eluent. The first fraction was the unchanged C₆₀ and the second and third fractions were the two diastereoisomeric cycloadducts. Some unchanged 5-formylpyrimidin-4-ones were also recovered.

4.5. Characterisation of the fulleropyrrolidine–glycopyranosylaminopyrimidin-4-one conjugates

4.5.1. Compound 7a

Mp >300 °C. ¹H NMR (300.13 MHz, CHCl₃): $\delta=1.98$, 2.03, 2.08 (3s, 3×3H, 3×AcO), 2.58 (s, 3H, SCH₃), 2.79 (s, 3H, N(1)–CH₃), 3.38 (t, $J=10.8$ Hz, 1H, H-sugar), 3.44 (s, 3H, N(3')–CH₃), 4.16–4.21 (m, 1H, H-sugar), 4.25 (d, $J=9.7$ Hz, 1H, H-5), 4.99 (d, $J=9.7$ Hz, 1H, H-5), 5.10–5.14 (m, 1H, H-sugar), 5.27–5.38 (m, 2H, H-sugar), 5.53 (s, 1H, H-2), 5.70 (t, $J=8.9$ Hz, 1H, H-sugar), 9.15 (d, $J=10.5$ Hz, 1H, NH). ¹³C NMR (75.47 MHz, CDCl₃): $\delta=15.0$ (CH₃), 20.7, 20.8, 21.5

($3 \times \text{CH}_3\text{C}=\text{O}$), 30.4 (N(3')–CH₃), 40.1 (N(1)–CH₃), 64.2, 69.0, 69.2, 69.5, 70.9, 73.7, 75.8, 76.2 ($2 \times \text{C}_{60}\text{-sp}^3$), 80.6 (C-1''), 92.9 (C-5'), 105.3, 137.2, 139.7, 141.5, 141.9, 142.0, 142.2, 142.3, 142.6, 142.9, 144.3, 145.1, 145.3, 145.5, 145.7, 146.1, 146.2, 146.6, 147.1, 153.2, 153.4, 155.2, 157.9, 162.0, 169.9, 170.3 ($3 \times \text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1205 [M+H]⁺, 720 [C₆₀]⁺⁺.

4.5.2. Compound 7b

Mp >300 °C. ¹H NMR (300.13 MHz, CHCl₃): δ =2.01, 2.07, 2.18 (3s, $3 \times 3\text{H}$, $3 \times \text{AcO}$), 2.55 (s, 3H, SCH₃), 2.69 (s, 3H, N(1)–CH₃), 3.25 (t, J =9.2 Hz, 1H, H-sugar), 3.45 (s, 3H, N(3')–CH₃), 3.88–3.94 (m, 1H, H-sugar), 4.19 (d, J =9.5 Hz, 1H, H-5), 4.97 (d, J =9.5 Hz, 1H, H-5), 4.94–5.00 (m, 1H, H-sugar), 5.24 (t, J =9.2 Hz, 1H, H-sugar), 5.35 (t, J =9.2 Hz, 1H, H-sugar), 5.52 (t, J =9.2 Hz, 1H, H-sugar), 5.60 (s, 1H, H-2), 9.41 (d, J =9.2 Hz, 1H, NH). ¹³C NMR (75.47 MHz, CDCl₃): δ =14.9 (SCH₃), 20.69, 20.73, 20.8 ($3 \times \text{CH}_3\text{C}=\text{O}$), 30.3 (N(3')–CH₃), 39.7 (N(1)–CH₃), 64.1, 69.2, 69.4, 70.8, 73.1, 75.1, 76.0, 76.2 ($2 \times \text{C}_{60}\text{-sp}^3$), 80.2 (C-1''), 91.2 (C-5'), 134.8, 136.0, 136.3, 136.9, 140.0, 140.1, 141.2, 141.6, 142.0, 142.2, 142.28, 142.34, 142.5, 142.45, 142.54, 142.9, 143.1, 144.4, 145.0, 145.3, 145.4, 145.7, 145.9, 146.0, 146.1, 146.3, 146.4, 146.5, 147.0, 147.3, 151.8, 153.3, 155.0, 156.9, 157.3, 162.9, 169.7, 169.8, 170.3 ($3 \times \text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1205 [M+H]⁺, 720 [C₆₀]⁺⁺.

4.5.3. Compound 8a

Mp >300 °C. ¹H NMR (300.13 MHz, CDCl₃): δ =1.97, 2.03, 2.07 (3s, $3 \times 3\text{H}$, $3 \times \text{AcO}$), 2.78 (s, 3H, N(1)–CH₃), 3.34 (s, 3H, N(3')–CH₃), 3.38–3.49 (m, 1H, H-sugar), 4.03 (s, 3H, OCH₃), 4.19 (dd, J =11.4, 5.6 Hz, 1H, H-sugar), 4.24 (d, J =9.6 Hz, 1H, H-5), 4.99 (d, J =9.6 Hz, 1H, H-5), 5.05–5.14 (m, 1H, H-sugar), 5.27–5.39 (m, 2H, $2 \times \text{H-sugar}$), 5.51 (s, 1H, H-2), 5.62–5.69 (m, 1H, H-sugar), 9.14 (d, J =10.4 Hz, 1H, NH). ¹³C NMR (125.77 MHz, CDCl₃): δ =20.7, 20.8, 21.5 ($3 \times \text{CH}_3\text{C}=\text{O}$), 27.9 (N(3')–CH₃), 40.0 (N(1)–CH₃), 55.4 (OCH₃), 64.2, 69.0, 69.2, 69.5, 70.9, 73.7, 75.9, 76.3, 77.6 ($2 \times \text{C}_{60}\text{-sp}^3$), 80.6 (C-1''), 91.0 (C-5'), 134.4, 136.1, 136.7, 137.2, 139.4, 139.8, 140.06, 140.13, 141.45, 141.49, 141.7, 141.9, 142.0, 142.2, 142.3, 142.48, 142.53, 142.56, 142.58, 142.88, 142.92, 144.3, 144.4, 144.9, 144.99, 145.09, 145.2, 145.3, 145.37, 145.40, 145.46, 145.53, 145.7, 145.9, 145.95, 145.99, 146.0, 146.1, 146.2, 146.55, 146.64, 147.3, 153.3, 153.5, 155.3, 155.9, 156.2, 158.9, 163.6, 169.9, 170.0, 170.3 ($3 \times \text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1189 [M+H]⁺, 720 [C₆₀]⁺⁺.

4.5.4. Compound 8b

Mp >300 °C. ¹H NMR (300.13 MHz, CDCl₃): δ =2.01, 2.07, 2.17 (3s, $3 \times 3\text{H}$, $3 \times \text{AcO}$), 2.68 (s, 3H, N(1)–CH₃), 3.28 (t, J =11.0 Hz, 1H, H-sugar), 3.34 (s, 3H, N(3')–CH₃), 3.91 (dd, J =11.4, 5.7 Hz, 1H, H-sugar), 4.01 (s, 3H, OCH₃), 4.19 (d, J =9.5 Hz, 1H, H-5), 4.95–5.03 (m, 2H, H-5 and H-sugar), 5.24 (t, J =9.2 Hz, 1H, H-sugar), 5.36 (t, J =9.2 Hz, 1H, H-sugar), 5.49 (t, J =9.2 Hz, 1H, H-sugar), 5.58 (s, 1H, H-2), 9.38 (d, J =9.2 Hz, 1H, NH). ¹³C NMR (125.77 MHz, CDCl₃): δ =20.7, 20.8, 20.9

($3 \times \text{CH}_3\text{C}=\text{O}$), 27.8 (N(3')–CH₃), 39.7 (N(1)–CH₃), 55.3 (OCH₃), 64.1, 69.2, 69.3, 69.4, 70.9, 73.1, 75.2, 76.1, 77.6 ($2 \times \text{C}_{60}\text{-sp}^3$), 80.1 (C-1''), 89.2 (C-5'), 134.8, 136.1, 136.3, 137.9, 139.3, 139.96, 140.00, 141.3, 141.6, 141.7, 142.0, 142.05, 142.07, 142.2, 142.25, 142.29, 142.3, 142.36, 142.44, 142.6, 143.1, 144.3, 144.4, 144.5, 144.6, 145.0, 145.18, 145.22, 145.25, 145.27, 145.3, 145.4, 145.7, 145.88, 145.90, 145.98, 146.0, 146.1, 146.3, 146.5, 147.1, 147.2, 147.3, 152.0, 153.4, 155.2, 155.9, 157.0, 158.4, 163.5, 169.78, 169.82, 170.3 ($3 \times \text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1189 [M+H]⁺, 720 [C₆₀]⁺⁺.

4.5.5. Compound 9a

Mp >300 °C. ¹H NMR (300.13 MHz, CDCl₃): δ =1.96, 2.00, 2.06, 2.12 (4s, $4 \times 3\text{H}$, $4 \times \text{AcO}$), 2.76 (s, 3H, N(1)–CH₃), 3.35 (s, 3H, N(3')–CH₃), 3.81 (ddd, J =9.9, 4.4, 2.0 Hz, 1H, H-sugar), 4.03 (s, 3H, OCH₃), 4.19–4.33 (m, 3H, H-5 and $2 \times \text{H-sugar}$), 4.99 (d, J =9.6 Hz, 1H, H-5), 5.21–5.30 (m, 1H, H-sugar), 5.32–5.40 (m, 2H, H-sugar), 5.51 (s, 1H, H-2), 5.70–5.76 (m, 1H, H-sugar), 9.19 (d, J =10.4 Hz, 1H, NH). ¹³C NMR (75.47 MHz, CDCl₃): δ =20.6, 20.7, 20.8, 21.4 ($4 \times \text{CH}_3\text{C}=\text{O}$), 27.9 (N(3')–CH₃), 39.8 (N(1)–CH₃), 55.4 (OCH₃), 62.0 (C-6''), 68.5, 69.0, 69.4 (C-5), 70.7, 73.0 (C-5''), 74.2, 75.8 (C-2), 76.3 (C₆₀-sp³), 80.1 (C-1''), 91.1 (C-5'), 134.4, 136.1, 136.7, 137.2, 139.4, 139.8, 140.08, 140.13, 141.46, 141.49, 141.7, 141.9, 142.0, 142.1, 142.2, 142.27, 142.27, 142.30, 142.48, 142.53, 142.6, 142.90, 142.92, 144.36, 144.43, 144.6, 144.94, 145.00, 145.1, 145.3, 145.4, 145.46, 145.54, 145.7, 145.9, 145.96, 145.99, 146.1, 146.2, 146.5, 146.7, 147.18, 147.23, 153.3, 153.4, 155.3, 155.8, 156.2, 158.9, 163.6, 169.5, 169.8, 170.4, 170.6 ($4 \times \text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1261 [M+H]⁺, 720 [C₆₀]⁺⁺.

4.5.6. Compound 9b

Mp >300 °C. ¹H NMR (300.13 MHz, CDCl₃): δ =1.89, 2.00, 2.05, 2.16 (4s, $4 \times 3\text{H}$, $4 \times \text{AcO}$), 2.69 (s, 3H, N(1)–CH₃), 3.35 (s, 3H, N(3')–CH₃), 3.70 (ddd, J =9.9, 4.4, 1.9 Hz, 1H, H-5''), 3.82 (dd, J =12.2, 1.9 Hz, 1H, H-6''), 4.00 (s, 3H, OCH₃), 4.20 (d, J =9.5 Hz, 1H, H-5), 4.98 (d, J =9.5 Hz, 1H, H-5), 5.12 (t, J =9.5 Hz, 1H, H-sugar), 5.31 (t, J =9.5 Hz, 1H, H-sugar), 5.38 (t, J =9.5 Hz, 1H, H-sugar), 5.59 (s, H-2, 1H), 5.65 (t, J =9.5 Hz, 1H, H-sugar), 9.55 (d, J =9.5 Hz, 1H, NH). ¹³C NMR (125.77 MHz, CDCl₃): δ =20.6, 20.67, 20.70, 20.8 ($4 \times \text{CH}_3\text{C}=\text{O}$), 27.9 (N(3')–CH₃), 39.8 (N(1)–CH₃), 55.3 (OCH₃), 62.2, 68.6, 69.1, 69.3, 70.7, 73.1, 73.6, 75.2, 76.2, 76.6 ($2 \times \text{C}_{60}\text{-sp}^3$), 79.3 (C-1''), 89.0 (C-5'), 139.9, 140.05, 140.10, 141.47, 141.52, 141.6, 141.8, 141.96, 142.04, 142.1, 142.2, 142.27, 142.29, 142.4, 142.5, 142.6, 142.8, 142.9, 144.3, 144.4, 144.5, 144.6, 144.9, 145.0, 145.22, 145.24, 145.31, 145.32, 145.4, 145.66, 145.70, 145.9, 145.97, 146.01, 146.15, 146.17, 146.18, 146.45, 146.54, 147.0, 147.2, 147.3, 152.1, 153.3, 155.1, 155.9, 156.8, 158.0, 163.5, 169.4, 169.7, 170.3, 170.7 ($4 \times \text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1261 [M+H]⁺, 720 [C₆₀]⁺⁺.

4.5.7. Compound 10a

Mp >300 °C. ^1H NMR (300.13 MHz, CDCl_3): δ =1.98, 2.01, 2.07, 2.12 (4s, 4 $\times$ 3H, 4 $\times$ AcO), 2.58 (s, 3H, SCH_3), 2.77 (s, 3H, N(1)– CH_3), 3.46 (s, 3H, N(3')– CH_3), 3.78 (ddd, J =9.8, 4.7, 2.7 Hz, 1H, H-sugar), 4.18–4.32 (m, 3H, H-5 and 2 $\times$ H-6''), 5.00 (d, J =9.7 Hz, 1H, H-5), 5.19–5.26 (m, 1H, H-sugar), 5.31–5.40 (m, 2H, H-sugar), 5.53 (s, 1H, H-2), 5.73–5.79 (m, 1H, H-sugar), 9.21 (d, J =10.3 Hz, 1H, NH). ^{13}C NMR (125.77 MHz, CDCl_3): δ =20.60, 20.64, 20.7, 21.4 (4 $\times$ $\text{CH}_3\text{C}=\text{O}$), 29.7 (N(3')– CH_3), 39.9 (N(1)– CH_3), 62.1, 68.6, 69.0, 69.5, 70.7, 73.1, 74.2, 75.8, 77.6 (2 $\times$ $\text{C}_{60}\text{-sp}^3$), 80.1 (C-1''), 93.1 (C-5'), 134.4, 136.1, 136.7, 137.2, 137.4, 139.5, 139.7, 139.9, 140.10, 140.14, 141.4, 141.5, 141.7, 142.0, 142.16, 142.18, 142.2, 142.3, 142.5, 142.6, 142.90, 142.94, 143.1, 144.4, 144.5, 144.6, 145.0, 145.06, 145.11, 145.2, 145.3, 145.35, 145.41, 145.46, 145.48, 145.7, 145.9, 146.0, 146.1, 146.2, 146.6, 147.1, 147.2, 147.3, 153.1, 153.4, 155.2, 156.1, 157.9, 161.9, 162.9, 169.5, 169.8, 170.4, 170.6 (4 $\times$ $\text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1277 $[\text{M}+\text{H}]^+$, 720 $[\text{C}_{60}]^{++}$.

4.5.8. Compound 10b

Mp >300 °C. ^1H NMR (300.13 MHz, CDCl_3): δ =1.89, 2.00, 2.06, 2.17 (4s, 4 $\times$ 3H, 4 $\times$ AcO), 2.54 (s, 3H, SCH_3), 2.70 (s, 3H, N(1)– CH_3), 3.46 (s, 3H, N(3')– CH_3), 3.67 (ddd, J =10.0, 4.6, 1.9 Hz, 1H, H-sugar), 3.84 (dd, J =12.1, 1.9 Hz, 1H, H-6''), 4.17 (dd, J =12.1, 4.7 Hz, 1H, H-6''), 4.22 (d, J =9.6 Hz, 1H, H-5), 4.99 (d, J =9.6 Hz, 1H, H-5), 5.09 (t, J =9.2 Hz, 1H, H-sugar), 5.31 (t, J =9.2 Hz, 1H, H-sugar), 5.37 (t, J =9.2 Hz, 1H, H-sugar), 5.61 (s, 1H, H-2), 5.68 (t, J =9.2 Hz, 1H, H-sugar), 9.58 (d, J =9.2 Hz, 1H, NH). ^{13}C NMR (125.77 MHz, CDCl_3): δ =20.6, 20.7, 20.8, 22.7 (4 $\times$ $\text{CH}_3\text{C}=\text{O}$), 30.3 (N(3')– CH_3), 39.8 (N(1)– CH_3), 62.3, 68.7, 69.2, 69.3, 70.6, 73.1, 73.5, 75.1, 76.1, 77.6 (2 $\times$ $\text{C}_{60}\text{-sp}^3$), 79.4 (C-1''), 91.0 (C-5'), 134.8, 136.0, 136.5, 137.7, 139.3, 139.9, 140.07, 140.13, 140.14, 141.56, 141.61, 141.7, 142.0, 142.1, 142.19, 142.23, 142.3, 142.4, 142.5, 142.6, 142.8, 143.1, 144.3, 144.4, 144.5, 144.6, 144.9, 145.2, 145.26, 145.32, 145.4, 145.65, 145.68, 145.9, 145.98, 146.02, 146.1, 146.19, 146.20, 146.5, 146.9, 147.2, 147.3, 151.9, 153.2, 154.9, 156.8, 157.0, 162.1, 162.8, 169.4, 169.8, 170.3, 170.6 (4 $\times$ $\text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1277 $[\text{M}+\text{H}]^+$, 720 $[\text{C}_{60}]^{++}$.

4.5.9. Compound 11a

Mp >300 °C. ^1H NMR (300.13 MHz, CDCl_3): δ =1.93, 1.99, 2.05, 2.12 (4s, 4 $\times$ 3H, 4 $\times$ AcO), 2.46 (s, 3H, SCH_3), 2.77 (s, 3H, N(1)– CH_3), 3.79 (ddd, J =9.8, 4.7, 2.3 Hz, 1H, H-sugar), 4.17 (dd, J =12.2, 2.3 Hz, 1H, H-6''), 4.26 (d, J =9.6 Hz, 1H, H-5), 4.30 (dd, J =12.2, 4.7 Hz, 1H, H-6''), 4.99 (d, J =9.6 Hz, 1H, H-5), 5.16–5.25 (m, 1H, H-sugar), 5.29–5.36 (m, 2H, 2 $\times$ H-sugar), 5.52 (s, 1H, H-2), 5.73–5.79 (m, 1H, H-sugar), 9.36 (d, J =10.8 Hz, 1H, NH), 12.86 (br s, 1H, N(3')–H). ^{13}C NMR (75.47 MHz, CDCl_3): δ =13.4 (SCH_3), 20.5, 20.6, 20.8, 20.3 (4 $\times$ $\text{CH}_3\text{C}=\text{O}$), 40.0 (N(1)– CH_3), 62.2 (C-6''), 68.6, 69.4 (C-5), 70.7, 73.2 (C-5''), 74.1, 74.8 (C-2), 76.2, 76.3 (2 $\times$ $\text{C}_{60}\text{-sp}^3$), 79.9 (C-1''), 93.2 (C-5'), 135.9, 136.7, 137.0, 139.6, 140.1, 140.2, 141.48, 141.54, 141.7, 141.8, 141.97, 142.01, 142.1, 142.2, 142.3, 142.5, 142.56, 142.64, 142.9, 144.3, 144.4,

144.6, 145.02, 145.04, 145.1, 145.20, 145.24, 145.3, 145.45, 145.49, 145.6, 145.9, 146.0, 146.06, 146.11, 146.2, 146.5, 146.78, 146.81, 147.2, 147.3, 152.8, 153.4, 155.5, 155.9, 160.3, 161.2, 165.3, 169.5, 169.8, 170.3, 170.6 (4 $\times$ $\text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1263 $[\text{M}+\text{H}]^+$, 720 $[\text{C}_{60}]^{++}$.

4.5.10. Compound 11b

Mp >300 °C. ^1H NMR (300.13 MHz, CDCl_3): δ =1.85, 1.99, 2.04, 2.16 (4s, 4 $\times$ 3H, 4 $\times$ AcO), 2.43 (s, 3H, SCH_3), 2.70 (s, 3H, N(1)– CH_3), 3.67 (ddd, J =10.0, 4.6, 1.7 Hz, 1H, H-sugar), 3.80 (dd, J =12.2, 1.7 Hz, 1H, H-6''), 4.15 (dd, J =12.2, 4.6 Hz, 1H, H-6''), 4.20 (d, J =9.5 Hz, 1H, H-5), 4.98 (d, J =9.5 Hz, 1H, H-5), 5.08 (t, J =9.3 Hz, 1H, H-sugar), 5.29 (t, J =9.3 Hz, 1H, H-sugar), 5.35 (t, J =9.3 Hz, 1H, H-sugar), 5.59 (s, 1H, H-2), 5.67 (t, J =9.3 Hz, 1H, H-sugar), 9.71 (d, J =9.3 Hz, 1H, NH), 12.77 (br s, 1H, N(3')–H). ^{13}C NMR (125.77 MHz, CDCl_3): δ =13.3 (SCH_3), 20.57, 20.64, 20.7, 20.8 (4 $\times$ $\text{CH}_3\text{C}=\text{O}$), 39.9 (N(1)– CH_3), 62.2 (C-6''), 68.5, 69.1 (C-5), 70.7, 73.1 (C-5''), 73.5, 74.2 (C-2), 76.0, 78.0 (2 $\times$ $\text{C}_{60}\text{-sp}^3$), 79.3 (C-1''), 91.3 (C-5'), 134.8, 135.8, 136.5, 137.6, 138.5, 139.3, 139.8, 140.1, 140.2, 141.35, 141.39, 141.6, 141.97, 142.02, 142.1, 142.20, 142.22, 142.3, 142.49, 142.54, 142.6, 142.9, 144.2, 144.3, 144.4, 144.5, 145.0, 145.15, 145.19, 145.24, 145.3, 145.4, 145.59, 145.64, 145.9, 145.96, 146.01, 146.04, 146.1, 146.2, 146.4, 146.67, 146.69, 147.2, 147.3, 151.6, 153.2, 155.2, 156.6, 159.4, 161.3, 165.1, 169.4, 169.8, 170.3, 170.6 (4 $\times$ $\text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1263 $[\text{M}+\text{H}]^+$, 720 $[\text{C}_{60}]^{++}$.

4.5.11. Compound 12a

Mp >300 °C. ^1H NMR (300.13 MHz, CDCl_3): δ =1.88, 2.10, 2.16 (3s, 3 $\times$ 3H, 3 $\times$ AcO), 2.46 (s, 3H, SCH_3), 2.86 (s, 3H, N(1)– CH_3), 3.78–3.88 (m, 2H, H-sugar), 4.32 (d, J =9.8 Hz, 1H, H-5), 4.95–5.25 (m, 3H, H-5 and 2 $\times$ H-sugar), 5.63 (s, 1H, H-2), 5.70 (t, J =9.3 Hz, 1H, H-sugar), 6.09 (dd, J =7.0, 4.5 Hz, 1H, H-sugar), 10.07 (d, J =7.1 Hz, 1H, NH), 12.69 (br s, 1H, N(3')–H). MS (LSIMS) m/z 1191 $[\text{M}+\text{H}]^+$, 720 $[\text{C}_{60}]^{++}$.

4.5.12. Compound 12b

Mp >300 °C. ^1H NMR (300.13 MHz, CDCl_3): δ =2.13, 2.16, 2.18 (3s, 3 $\times$ 3H, 3 $\times$ AcO), 2.49 (s, 3H, SCH_3), 2.91 (s, 3H, N(1)– CH_3), 3.96 (dd, J =13.4, 2.1 Hz, 1H, H-sugar), 4.08–4.13 (m, 1H, H-sugar), 4.32 (d, J =9.7 Hz, 1H, H-5), 4.73–4.79 (m, 2H, H-sugar), 4.91 (d, J =9.7 Hz, 1H, H-5), 5.27–5.29 (m, 1H, H-sugar), 5.59 (s, 1H, H-2), 5.88 (dd, J =9.7, 1.7 Hz, 1H, H-sugar), 9.81 (d, J =9.8 Hz, 1H, NH), 12.03 (br s, 1H, N(3')–H). ^{13}C NMR (125.77 MHz, CDCl_3): δ =13.1 (SCH_3), 20.9, 21.07, 21.10, 22.7 (4 $\times$ $\text{CH}_3\text{C}=\text{O}$), 40.1 (N(1)– CH_3), 64.7, 66.2, 67.0, 68.8, 69.0, 69.7, 74.6, 75.6, 76.3, 78.0 (2 $\times$ $\text{C}_{60}\text{-sp}^3$), 91.1 (C-5'), 134.7, 135.7, 136.0, 136.1, 136.2, 138.78, 138.82, 139.5, 139.6, 139.8, 140.1, 140.2, 141.0, 141.3, 141.6, 141.8, 141.96, 141.98, 142.0, 142.1, 142.2, 142.3, 142.4, 142.51, 142.53, 142.7, 142.9, 143.0, 144.3, 144.4, 144.7, 145.09, 145.14, 145.16, 145.19, 145.25, 145.29, 145.33, 145.5, 145.6, 145.8, 145.98, 145.99, 146.0, 146.05, 146.12, 146.3, 146.5, 147.3, 152.96, 153.01, 154.9, 155.9, 159.0, 161.6, 165.0, 167.8, 168.4,

169.6, 169.7 ($4 \times \text{CH}_3\text{C}=\text{O}$). MS (LSIMS) m/z 1191 $[\text{M}+\text{H}]^+$, 720 $[\text{C}_{60}]^+$.

4.5.13. 2',3'-O-Isopropylideneuridine **14**

PTSA (0.016 mg, 0.1 equiv) was added to a solution of uridine (200 mg, 0.82 mmol) in acetone (6 mL) and the resulting mixture was stirred at room temperature for 48 h. One drop of NEt_3 was added to the reaction mixture and then it was concentrated. Compound **14** was purified by preparative TLC (silica gel) using a 95:5 mixture of $\text{CHCl}_3/\text{MeOH}$ as the eluent. Evaporation of the solvent afforded a white solid (173 mg, 74% yield) with mp 146–147 °C. ^1H NMR (300.13 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$): $\delta=1.39$, 1.59 (2s, $2 \times 3\text{H}$, $2 \times \text{CH}_3$), 3.76 (dd, $J=12.0$, 3.5 Hz, 1H, H-5'), 3.87 (dd, $J=12.0$, 2.5 Hz, 1H, H-5'), 4.29 (q, $J=2.5$ Hz, 1H, H-4'), 4.88–4.93 (m, 2H, H-2' and H-3'), 5.71 (d, $J=8.1$ Hz, 1H, H-5), 5.75 (d, $J=2.3$ Hz, 1H, H-1'), 7.60 (d, $J=8.1$ Hz, 1H, H-6). ^{13}C NMR (75.47 MHz, $\text{CDCl}_3/\text{CD}_3\text{OD}$): $\delta=25.1$, 27.1 ($2 \times \text{CH}_3$), 62.3 (C-5'), 80.3 (C-3'), 83.9 (C-2'), 86.8 (C-4'), 95.0 (C-1'), 102.4 (C-5), 114.6 ($\text{C}(\text{CH}_3)_2$), 142.6 (C-6), 150.42 (C-2), 163.7 (C-4). MS (ESI) m/z 285 $[\text{M}+\text{H}]^+$, 307 $[\text{M}+\text{Na}]^+$. Elemental analysis (%) calcd for $\text{C}_{12}\text{H}_{16}\text{N}_2\text{O}_6$: C 50.70, H 5.67, N 9.85; found C 50.41, H 5.79, N 9.93.

4.5.14. Methyl uridiny malonate **15**

A solution of methyl 3-chloro-3-oxopropanoate (0.034 mL, 2 equiv) diluted in CH_2Cl_2 was added dropwise to an ice cooled solution of compound **14** (50 mg, 0.18 mmol) and anhydrous NEt_3 (0.097 mL, 0.7 mmol) in CH_2Cl_2 (7 mL). The ice bath was then removed and the reaction mixture was stirred at room temperature for 3 h. It was then washed with an aqueous solution of 0.5% HCl (2×30 mL) and with a saturated aqueous solution of NaHCO_3 (2×30 mL). The organic phase was dried (Na_2SO_4) and the solvent was evaporated under vacuum. Compound **15** was purified by TLC (silica gel) using an 8:2 mixture of ethyl acetate/petroleum ether as the eluent. Evaporation of the solvent afforded a white solid (53 mg, 77% yield) with mp 128–129 °C. ^1H NMR (300.13 MHz, CDCl_3): $\delta=1.36$, 1.57 (2s, $2 \times 3\text{H}$, $2 \times \text{CH}_3$), 3.44 (s, 2H, malonate CH_2), 3.75 (s, 3H, OCH_3), 4.34–4.35 (m, 3H, H-4' and $2 \times \text{H-5'}$), 4.84 (dd, $J=6.5$, 3.7 Hz, 1H, H-3'), 4.99 (dd, $J=6.5$, 2.0 Hz, 1H, H-2'), 5.67 (d, $J=2.0$ Hz, 1H, H-1'), 5.76 (d, $J=8.4$ Hz, 1H, H-5), 7.30 (d, $J=8.4$ Hz, 1H, H-6). ^{13}C NMR (75.47 MHz, CDCl_3): $\delta=27.0$, 25.18 (2CH_3), 41.1 (malonate CH_2), 52.7 (OCH_3), 64.8 (C-5'), 80.8 (C-3'), 84.4 (C-2'), 85.0 (C-4'), 94.6 (C-1'), 102.6 (C-5), 114.6 ($\text{C}(\text{CH}_3)_2$), 142.3 (C-6), 149.9 (C-2), 163.4 (C-4), 165.9, 166.8 ($2 \times \text{malonate C}=\text{O}$). MS (ESI) m/z 407 $[\text{M}+\text{Na}]^+$. Elemental analysis (%) calcd for $\text{C}_{16}\text{H}_{20}\text{N}_2\text{O}_9$: C 50.00, H 5.25, N 7.24; found C 50.05, H 5.36, N 7.26.

4.5.15. Fullerene–uridine conjugate **16**

To a solution of malonate **15** (35 mg, 0.091 mmol) and C_{60} (98 mg, 1.5 equiv) in anhydrous toluene (50 mL) were added iodine (34 mg, 1.5 equiv) and DBU (0.042 mL, 3 equiv). The reaction mixture was stirred for 8 h at room temperature, under a nitrogen atmosphere. It was then washed with an aqueous

solution of 5% $\text{Na}_2\text{S}_2\text{O}_3$ (2×50 mL). The organic phase was concentrated and purified by flash chromatography (silica gel) using toluene as the eluent. Compound **16** was crystallised from $\text{CHCl}_3/\text{MeOH}$ giving brown crystals (17 mg, 17% yield). Mp >300 °C. ^1H NMR (300.13 MHz, CDCl_3): $\delta=1.36$, 1.57 (2s, $2 \times 3\text{H}$, $2 \times \text{CH}_3$), 3.44 (s, 2H, malonate CH_2), 3.75 (s, 3H, OCH_3), 3.76 (dd, $J=11.8$, 4.5 Hz, 1H, H-5'), 3.87 (dd, $J=11.8$, 6.0 Hz, 1H, H-5'), 4.52–4.57 (m, 1H, H-4'), 4.84 (dd, $J=6.5$, 3.7 Hz, 1H, H-3'), 4.99 (dd, $J=6.5$, 2.0 Hz, 1H, H-2'), 5.67 (d, $J=2.0$ Hz, 1H, H-1'), 5.78 (d, $J=8.0$ Hz, 1H, H-5), 7.34 (d, $J=8.0$ Hz, 1H, H-6), 8.80 (br s, 1H, NH). ^{13}C NMR (75.47 MHz, CDCl_3): $\delta=27.1$, 25.2 ($2 \times \text{CH}_3$), 54.2 (OCH_3), 66.3 (C-5'), 71.23 ($\text{C}_{60}\text{-sp}^3$), 81.5 (C-3'), 84.4 (C-2'), 85.1 (C-4'), 95.6 (C-1'), 102.9 (C-5), 114.7 ($\text{C}(\text{CH}_3)_2$), 142.2 (C-6), 142.4, 143.0, 143.9, 144.8, 144.9, 145.2, 145.3, 149.7 (C-2), 162.6 (C-4), 163.1, 163.9 ($2 \times \text{malonate C}=\text{O}$). MS (MALDI-TOF) m/z 1102 $[\text{M}]^+$.

4.5.16. Fullerene–uridine conjugate **17**

To a solution of dyad **16** (15 mg) in CHCl_3 (3 mL) was added a 9:1 mixture of $\text{CF}_3\text{CO}_2\text{H}/\text{H}_2\text{O}$ (5 mL) and the reaction mixture was stirred for 14 h at room temperature. It was then neutralised with a saturated aqueous solution of Na_2CO_3 and dyad **17** was extracted with a mixture of $\text{CHCl}_3/\text{MeOH}$ (3×30 mL). The organic layer was dried (Na_2SO_4), concentrated, and compound **17** was precipitated by the addition of MeOH affording a brown solid (14 mg, 96% yield). Mp >300 °C. ^1H NMR (300.13 MHz, $\text{DMSO}-d_6$): $\delta=4.05$ (s, 3H, OCH_3), 4.19 (m, 2H, H-5'), 4.66–4.73 (m, 2H, H-2' and H-3'), 5.45–5.49 (m, 1H, H-4'), 5.63 (d, $J=8.3$ Hz, 1H, H-5), 5.82 (d, $J=5.7$ Hz, 1H, H-1'), 7.72 (d, $J=8.3$ Hz, 1H, H-6), 11.39 (br s, 1H, NH). MS (ESI) m/z 1063 $[\text{M}+\text{H}]^+$, 1085 $[\text{M}+\text{Na}]^+$.

Acknowledgements

Thanks are due to the Fundação para a Ciência e a Tecnologia (FCT, Portugal) and FEDER for funding the Organic Chemistry Research Unit and the project POCI/QUI/58515/2004 and also for funding towards the purchase of the single-crystal diffractometer. A.S.F.F. and R.F.E. also thank FCT for their doctoral (SFRH/BD/32219/2006) and post-doctoral (SFRH/BPD/24213/2005) grants, respectively.

References and notes

- Recent reviews: (a) Bendikov, M.; Wudl, F.; Perepichka, D. F. *Chem. Rev.* **2004**, *104*, 4891–4945; (b) Sánchez, L.; Herranz, Á.; Martín, N. J. *Mater. Chem.* **2005**, *15*, 1409–1421; (c) Martín, N. *Chem. Commun.* **2006**, 2093–2104; (d) Figueira-Duarte, T. M.; Gégout, A.; Nierengarten, J.-F. *Chem. Commun.* **2007**, 109–119; (e) *Fullerenes: Principles and Applications*; Langa, F., Nierengarten, J.-F., Eds.; RSC: Cambridge, UK, 2007.
- Reviews: (a) Jensen, W.; Wilson, S. R.; Schuster, D. I. *Bioorg. Med. Chem.* **1996**, *4*, 767–779; (b) Da Ros, T.; Prato, M. *Chem. Commun.* **1999**, 663–669; (c) Tagmatarchis, N.; Shinohara, H. *Mini-Rev. Med. Chem.* **2001**, *1*, 339–348; (d) Nakamura, E.; Isobe, H. *Acc. Chem. Res.* **2003**, *36*, 807–815; (e) Bosi, S.; Da Ros, T.; Spalluto, G.; Prato, M. *Eur. J. Med. Chem.* **2003**, *38*, 913–923; (f) Bianco, A.; Da Ros, T. *Biological Applications of Fullerenes*. In *Fullerenes: Principles and Applications*; Langa, F., Nierengarten, J.-F., Eds.; RSC: Cambridge, UK, 2007.

3. Recent articles: (a) Zakharian, T. Y.; Seryshev, A.; Sitharaman, B.; Gilbert, B. E.; Knight, V.; Wilson, L. J. *J. Am. Chem. Soc.* **2005**, *127*, 12508–12509; (b) Ashcroft, J. M.; Tsybouski, D. A.; Hartman, K. B.; Zakharian, T. Y.; Marks, J. W.; Weisman, R. B.; Rosenblumb, M. G.; Wilson, L. J. *Chem. Commun.* **2006**, 3004–3006; (c) Troshina, O. A.; Troshin, P. A.; Peregodov, A. S.; Kozlovskiy, V. I.; Balzarini, J.; Lyubovskaya, R. N. *Org. Biomol. Chem.* **2007**, *5*, 2783–2791; (d) Liu, J.; Ohta, S.; Sonoda, A.; Yamada, M.; Yamamoto, M.; Nitta, N.; Murata, K.; Tabata, Y. *J. Controlled Release* **2007**, *117*, 104–110.
4. (a) Enes, R. F.; Tomé, A. C.; Cavaleiro, J. A. S.; Amorati, R.; Fumo, M. G.; Pedulli, G. F.; Valgimigli, L. *Chem.—Eur. J.* **2006**, *12*, 4646–4653; (b) Enes, R. F.; Tomé, A. C.; Cavaleiro, J. A. S.; El-Agamey, A.; McGarvey, D. J. *Tetrahedron* **2005**, *61*, 11873–11881; (c) Enes, R. F.; Tomé, A. C.; Cavaleiro, J. A. S. *Tetrahedron* **2005**, *61*, 1423–1431; (d) De la Torre, M. D. L.; Rodrigues, A. G. P.; Tomé, A. C.; Silva, A. M. S.; Cavaleiro, J. A. S. *Tetrahedron* **2004**, *60*, 3581–3592.
5. (a) Negrillo, J.; Nogueras, M.; Sánchez, A.; Melgarejo, M. *Chem. Pharm. Bull.* **1988**, *36*, 386–393; (b) Nogueras, M.; Cobo, J.; Quijano, M. L.; Melguizo, M.; Sanchez, A.; Melgarejo, M. *Nucleosides Nucleotides* **1994**, *13*, 447–457.
6. (a) Ungureanu, C.; Pinteala, M.; Simionescu, B. C. *Synthesis* **2005**, 361–363; (b) Greisch, J.-F.; Pauw, E. D. *J. Mass Spectrom.* **2007**, *42*, 304–311.
7. (a) Sessler, J. L.; Jayawickramarajah, J.; Gouloumis, A.; Torres, T.; Guldi, D. M.; Maldonado, S.; Stevenson, K. J. *Chem. Commun.* **2005**, 1892–1894; (b) Torres, T.; Gouloumis, A.; Sanchez-Garcia, D.; Jayawickramarajah, J.; Seitz, W.; Guldi, D. M.; Sessler, J. L. *Chem. Commun.* **2007**, 292–294.
8. Maggini, M.; Scorrano, G.; Prato, M. *J. Am. Chem. Soc.* **1993**, *115*, 9798–9799.
9. For the ^1H NMR discussion, protons H-5 were indicated as H-5a (the one with a cis relationship with proton H-2) and H-5b.
10. Laprévote, O.; Ducrot, P.; Thal, C.; Serani, L.; Das, B. P. *J. Mass Spectrom.* **1996**, *31*, 1149–1155.
11. Madhusudanam, K. P. *J. Mass Spectrom.* **2006**, *41*, 1096–1104.
12. Tsunematsu, H.; Ikeda, H.; Hanazono, H.; Inagaki, M.; Isobe, R.; Higuchi, R.; Goto, Y.; Yamamoto, M. *J. Mass Spectrom.* **2003**, *38*, 188–195.
13. Gaucher, S. P.; Leary, J. A. *Anal. Chem.* **1998**, *70*, 3009–3014.
14. Perreault, H.; Costello, C. E. *J. Mass Spectrom.* **1999**, *34*, 184–197.
15. Zhu, X.; Sato, T. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 191–198.
16. Allen, F. H. *Acta Crystallogr., Sect. B* **2002**, *58*, 380–388.
17. Allen, F. H.; Motherwell, W. D. S. *Acta Crystallogr., Sect. B* **2002**, *58*, 407–422.
18. Ferguson, G.; Low, J. N.; Nogueras, M.; Cobo, J.; Lopez, M. D.; Quijano, M. L.; Sanchez, A. *Acta Crystallogr., Sect. C* **1998**, *54*, IUC9800031. doi:10.1107/S0108270198099521
19. Cobo, J.; Melguizo, M.; Sanchez, A.; Nogueras, M.; Low, J. N.; Ferguson, G. *Acta Crystallogr., Sect. C* **1996**, *52*, 148–150.
20. Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358.
21. Bernstein, J.; Davis, R. E.; Shimon, L.; Chang, N. L. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 1555–1573.
22. (a) Levene, P. A.; Tipson, R. S. *J. Biol. Chem.* **1934**, *106*, 113–124; (b) Scheit, K. H. *Chem. Ber.* **1966**, *99*, 3884–3891; (c) Fromageot, H. P. M.; Griefin, B. E.; Reese, C. B.; Sulston, J. E. *Tetrahedron* **1967**, *23*, 2315–2331.
23. Bingel, C. *Chem. Ber.* **1993**, *126*, 1957–1959.
24. Bradshaw, T. K.; Hutchinson, D. W. *Chem. Soc. Rev.* **1977**, *6*, 43–62.
25. Kottke, T.; Stalke, D. *J. Appl. Crystallogr.* **1993**, *26*, 615–619.
26. APEX2. *Data Collection Software Version 2.1-RC13*; Bruker AXS: Delft, The Netherlands, 2006.
27. *Cryopad. Remote Monitoring and Control, Version 1.451*; Oxford Cryosystems: Oxford, United Kingdom, 2006.
28. SAINT+. *Data Integration Engine v. 7.23a*®; Bruker AXS: Madison, Wisconsin, USA, 1997–2005.
29. Sheldrick, G. M. *SADABS v.2.01, Bruker/Siemens Area Detector Absorption Correction Program*; Bruker AXS: Madison, Wisconsin, USA, 1998.
30. Sheldrick, G. M. *SHELXS-97, Program for Crystal Structure Solution*; University of Göttingen: Göttingen, Germany, 1997.
31. Sheldrick, G. M. *SHELXL-97, Program for Crystal Structure Refinement*; University of Göttingen: Göttingen, Germany, 1997.
32. van der Sluis, P.; Spek, A. L. *Acta Crystallogr., Sect. A* **1990**, *46*, 194–201.
33. Spek, A. L. *Acta Crystallogr., Sect. A* **1990**, *46*, C34.
34. Spek, A. L. *PLATON, A Multipurpose Crystallographic Tool*; Utrecht University: Utrecht, The Netherlands, 2003.
35. Flack, H. D.; Bernardinelli, G. *J. Appl. Crystallogr.* **2000**, *33*, 1143–1148.
36. Flack, H. D. *Acta Crystallogr., Sect. A* **1983**, *39*, 876–881.