Electronic Supplementary Information

Antiproliferative Activity of Bicyclic Benzimidazole Nucleosides:

Synthesis, DNA-Binding and Cell Cycle Analysis

Vyankat A Sontakke,[a] Pravin P Lawande,[a] Anup N Kate,[a] Ayesha Khan,[a] Rakesh Joshi,[b] Anupa A Kumbhar,[a] and Vaishali S Shinde[a*]

[a]Department of Chemistry, Savitribai Phule Pune University (formerly University of Pune) Pune-411007, India

[b]Institute of Bioinformatics and Biotechnology, Savitribai Phule Pune University Pune-411007, India

*Corresponding author. Fax: +91-20-25691728.

Email address: vsshinde@chem.unipune.ac.in; (Vaishali S Shinde).

Experimental procedure:

Single crystal

The single crystal data for compounds **2** and **12** were measured on a Bruker Apex II CCD diffractometer. Cell dimensions and the orientation matrix were initially determined from a least-squares refinement on reflections measured in three sets of 20 exposures, collected in three different ω regions, and eventually refined against all data. A full sphere of reciprocal space was scanned by 0.3^o ω steps. The software SMART[1] was used for collecting frames of data, indexing reflections and determination of lattice parameters. The collected frames were then processed for integration by the SAINT program and an empirical absorption correction was applied using SADABS.[2] The structures were solved by direct methods (SIR 97)[3] and subsequent Fourier syntheses and refined by full-matrix least-squares on F2 (SHELXTL)[4] using anisotropic thermal parameters for all non-hydrogen atoms. In the compound 2 a H_2O molecule is present in lattice. The aromatic hydrogen atoms were placed in calculated positions, constrained to ride on their parent atoms and refined isotropically with Uiso(H) = 1.2 Uiso(parent atom). The molecular graphics were generated by using Mercury 3.1 program.[5] The CCDC publication numbers for **2** and **12** is 140636 and 1430637 respectively.

Molecular docking

Molecular docking was performed in order to understand the molecular interaction of synthesized compounds with DNA. Docking studies were performed using AutoDock 4.2 version.[6] Crystal structure of the B form of DNA dodecamer (PDB ID: 1BNA) is retrieved from protein data bank (www.rcsb.org). Structures of compound 1-4 determined from NMR were used

for the further docking procedure. All the structures were energy minimized using GROMOS 43BI force field.[7] Nucleosides and receptor files were optimized for docking studies and converted from .pdb to .pdbqt format. AutoGrid module was used for calculating the grid maps and grid was set around the complete DNA molecule with dimensions of 24 x 28 x 42 Å. The docking parameters were configured on an LGA calculation of 10,000 runs. Energy evaluations were set to 1,500,000 and 27,000 generations. Population size was set to 150 and the rate of gene mutation and the rate of gene crossover were set to 0.02 and 0.8 respectively. The obtained conformations were later summarized, collected and extracted by using Autodock tool. Geometry of resulting complexes was studied using the PyMol Molecular Viewer utility (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC). Intermolecular interactions in nucleotide of DNA due to synthesized compounds were analyzed and represented using Accelery Discovery Visualizer.

DNA Binding

DNA binding ability of nucleosides **1-4** (50 μM) with CT-DNA (0-500 μM) was investigated by UV-visible and fluorescence spectroscopy. Binding experiments were carried out by measuring absorption and emission spectra of solutions in presence and absence of DNA. The binding constant (K_b) for the nucleoside-DNA interaction was measured as per reported method.[8] During this measurement, the total nucleoside concentration (50 μM) in the experimental solution was kept unchanged and the steady-state fluorescence spectra were recorded as a function of the CT-DNA concentration ((0-500 μM)). In these measurements, at any CT-DNA concentration, the observed fluorescence intensity I_{fl} was expressed as follows:[8a]

$$I_{fl} = I_L^0 \frac{[L]_{eq}}{[L]_0} + I_{H:L}^\infty \frac{[H:L]_{eq}}{[L]_0} \qquad (1)$$

3

Where, I_L^0 is the initial fluorescence intensity of nucleoside in the absence of DNA, $I_{H:L}^\infty$ is the final fluorescence intensity. When all the nucleosides were converted to DNA:nucleoside complex, [L]₀ is the total concentration of the nucleoside used and [L]$_{eq}$ is the equilibrium concentration of the free nucleoside present in the solution at any DNA concentration. In the fluorescence titration curve, the changes in fluorescence intensity ΔI_f of **1-4** was plotted against the total concentration of [DNA]⁰ used. Following eq. 2, the parameter ΔI_f was expressed in terms of measured emission intensity as:[8b]

$$\Delta I_{fl} = \left(1 - \frac{[L]_{eq}}{[L]_0}\right)(I_{H:L}^\infty - I_L^0) \tag{2}$$

Where, the equilibrium concentration of the free nucleoside [L]$_{eq}$ is expressed in terms of the total CT-DNA concentration [DNA]⁰ and the equilibrium constant K_b by the following equation.

$$[L]_{eq} = \frac{\left\{K_b[L]_0 - K_b[H]_0 - 1) + \sqrt{(Kb[L]_0 + K_b[H]_0 + 1)^2 - 4(K_b)^2[L]_0[H]_0}\right\}}{2K_b} \tag{3}$$

The fluorescence of ligand was corrected, according to equation (4)

$$F_{corr} = F_{obs} * [A_{(L)}/A_{(L+D)}] \tag{4}$$

where F_{corr} is the corrected fluorescence intensity, F_{obs} is the observed fluorescence intensity at 340 nm. $A_{(L)}$ is Absorbance of ligand at λ_{ex} and $A_{(L+D)}$ is absorbance of Ligand + DNA at λ_{ex}.

Ethidium bromide displacement assay:

Ethidium bromide displacement assay of **1-4** was carried out by using fluorescence spectroscopy. At room temperature varying concentrations of **1-4** (0-100 μM) were added to CT-DNA solution (10 μM) pretreated with ethidium bromide (EtBr) (10 μM) in phosphate buffer and the changes in fluorescence intensities at 585 nm (λ_{ex}=545 nm) were recorded.

Cell viability assay

The MCF-7 and MDA-MB-231 cell lines were obtained from National Centre for Cell Sciences Repository, University of Pune, Pune 411007. The cells were maintained in RPMI 1640 w/ 10% Fetal bovine serum (FBS), 2% P/S, 1.25% L-glut, and 1% sodium pyruvate at 37 $^{\circ}$C (5% CO_2) in the steri-cycle CO_2 incubator with HEPA Class 100 filters, Thermo Electron Corporation. The number of viable cells remaining after appropriate treatment was determined by 3-(4,5-dimethylthiazol-2-yl)- 2,5-diphenyltetrazolium bromide (MTT; Sigma Chemical Co.) assay.[8f] Briefly, cells were plated (4000 cells/well per 0.2 mL medium) in 96-well microtiter plates and incubated overnight. The compounds **1-4** were taken in the range from 0.2 to 2 μM (increments of 0.2) in 1 % DMSO and filtered through 0.2 μ filter and added to quadruplicate wells. After 24 h, MTT was added to each well at a final volume of 0.5 mg/mL and the microplates were incubated at 37 $^{\circ}$C for 3 h. After the supernatant was removed, the formazan salt resulting from the reduction of MTT was solubilized in dimethyl sulfoxide (DMSO, Sigma Chemical Co.) and the absorbance was read at 570 nm using an automatic plate reader (Thermo Corporation). The cell viability was extrapolated from optical density (OD) 570 nm values and expressed as percent survival. The percent cell death was calculated considering the untreated cells as 100 percent viable. The results allowed to establish dose-response curves and to calculate IC50 values, the concentration required to inhibit cell proliferation by 50%.

Confocal microscopy

MCF 7 cells were seeded on to glass cover slips at a density of 5×10^4 cells to achieve ~ 80 confluence. The cells were treated with the compounds **1-4** at the concentration obtained from the IC_{50} values obtained by the MTT assay for 48 h at 37 $^{\circ}$C. After 48 h of incubation the cells

were stained with Live/Dead® cell mediated cytotoxicity kit for animal cells procured from Thermo Fischer Scientific reagents, both at a dilution of 1:20 for 30 min at 37 °C. The excess staining solution was removed and fixed in 4% para-formaldehyde (PFA). The microscopic images were observed under LSM 780 confocal laser scanning microscope, Carl Zeiss (Magnification 40x) with a dual filter set excitation/ emission 484/501 green (live cells) and 536/617 red (dead cells).

Protocol for Flow Cytometry

For preparation of cells for Flow assisted cell cytometry (FACS), the APO-BrDU™ TUNEL Assay kit from Invitrogen was used. MCF-7 cells were grown on 6-well plates. The wells were designated as: unstained control, Control (without any treatment), only PBS added, and compounds **1-4** (IC_{50} values respectively) were added. The plates were incubated for 24 h and the cells were fixed after that.

Protocol used for Cell Fixation

Cells were trypsinized and suspended in 0.5 mL phosphate buffer saline (PBS). The cell suspension was added to 5 mL 1 % paraformaldehyde (w/v) in PBS and was placed on ice for 15 minutes. The cells were centrifuged at 300g for 5 minutes, and the supernatant was discarded. The cells were washed with 5 mL PBS and pelleted again by centrifugation. The cells were re-suspended in 0.5 mL PBS. The cells were added to ice-cold 70% ethanol (v/v). The cells were stored at -20 °C till use. The cells were stained as per given protocol in the Invitrogen APO-BrdU™ TUNEL Assay kit. The FACS analysis data was acquired using BD FACS Calibur. The data was analyzed using the CellQuest Pro software.

1′,2′-O-Isopropylidene-3′-C-(prop-1-enyl)-α-D-allo-1′,4′-furanose (6). To a stirring solution of **5** (5.00 g, 16.64 mmol, 1 eq.) in THF (50 mL), 30% $HClO_4$ (5 mL) was added at 0 °C and the reaction was stirred for additional 2 h at room temperature. After completion of reaction (monitored by TLC), the perchloric acid was quenched by using saturated solution of aq. K_2CO_3 and THF was removed under vacuum. Aqueous layer was extracted with chloroform (4 x 150 mL) and concentrated on rotary evaporator. Further, purification was done by column chromatography using ethylacetate/n-hexane (4:6) to afford triol **6** as a white solid (3.90 g, 90%); mp 100−103 °C; R_f 0.38 (ethylacetate/*n*-hexane = 1/1); $[\alpha]_D^{25}$ +81.9 (c 0.2, $CHCl_3$); ^{1}H-NMR (500 MHz, $CDCl_3$) δ 5.99-5.91 (m, 1H, HC=), 5.66 (d, J = 3.0 Hz, 1H, H1′), 5.21-5.14 (m, 2H, $=CH_2$), 4.33 (d, J = 3.6 Hz, 1H, H2′), 3.82-3.78 (m, 3H, H5′, H6′A, H6′B), 3.67 (d, J = 6.8 Hz, 1H, H4′), 3.16 (bs, 1H, OH), 3.12 (bs, 1H, OH), 2.69-2.65 (dd, J = 14.52 and 5.58 Hz, 1H, CH_2), 2.24-2.19 (dd, J = 14.3 and 8.5 Hz, 1H, CH_2), 2.03 (bs, 1H, OH), 1.57 (s, 3H, CH_3), 1.33 (s, 3H, CH_3); ^{13}C-NMR (125 MHz, $CDCl_3$) δ 132.3 (C=C), 119.2 (C=C), 112.6 (C3), 103.4 (C1), 80.8 (C2), 79.2(C4), 79.0 (C3), 69.7 (C5), 64.4 (C6), 36.1 (CH_2), 26.5 (CH_3), 26.4 (CH_3).

1′,2′-O-Isopropylidene-3′-C-(prop-1-enyl)-4′-vinyl-α-D-ribo-1′,4′-furanose (7). In a dry THF (100 mL), **6** (3 g, 11.52 mmol, 1 eq.) was dissolved. To this, PPh_3 (10.58 g, 40.34 mmol, 3.5 eq.), imidazole (2.74 g, 40.34 mmol, 3.5 eq.) was added sequentially and lastly iodine (5.84 g, 23.04 mmol, 2 eq) was added in two fractions. The reaction mixture was stirred at room temperature under nitrogen atmosphere for 2 h. After completion, reaction mass was filtered through a short pad of celite and residue was washed with cold ether in order to remove, partially, the triphenylphosphine oxide side product. Excess of iodine was quenched by using saturated solution of sodium thiosulphate. Further, this reaction mixture was evaporated partially

under reduced pressure and extracted with dichloromethane (3 x 100 mL). The organic layer was dried using sodium sulpfate and concentrated under reduced pressure to afford crude product. Purification by column chromatography using 1:9 ethylacetate/*n*-hexane gave **7** as white solid in 68% yield 1.78 g; mp 96-98 °C; R_f 0.4 (40% ethylacetate in *n*-hexane), $[\alpha]_D^{27}$ +31.6 (c 0.1, CHCl3). ^{1}H-NMR (500 MHz, CDCl$_3$) δ 5.94-5.79 (m, 2H, H5′, HC=), 5.74 (d, *J* = 3.95 Hz, 1H, H1′), 5.43-5.39 (dt, *J* = 1.60 and 1.62 Hz, 1H, H6′A), 5.31-5.28 (dt, *J* = 1.52 and 1.50 Hz, 1H, H6′B), 5.17-5.09 (m, 2H, =CH$_2$), 4.33 (d, *J* = 3.95 Hz, 1H, H2′), 4.27 (d, *J* = 4.43 Hz, 1H, H4′), 2.68 (brs, 1H, OH), 2.39-2.35 (m, 1H, CH$_2$), 2.08-2.03 (m, 1H, CH$_2$), 1.57 (s, 3H, CH$_3$) 1.37 (s, 3H, CH$_3$); ^{13}C-NMR (125 MHz, CDCl$_3$) δ 133.0 (C5′), 131.7 (HC=), 119.1 (=CH$_2$), 118.7 (C6), 112.5 (spiro C), 103.7 (C1), 83.5 (C4), 80.8 (C2), 79.6 (C3), 36.7 (CH$_2$), 26.57 (CH$_3$), 26.54 (CH$_3$); HRMS (ESI) *m/z* [M + Na]$^+$ calcd. for C$_{12}$H$_{18}$NaO$_4$, 249.1103, found 249.1099.
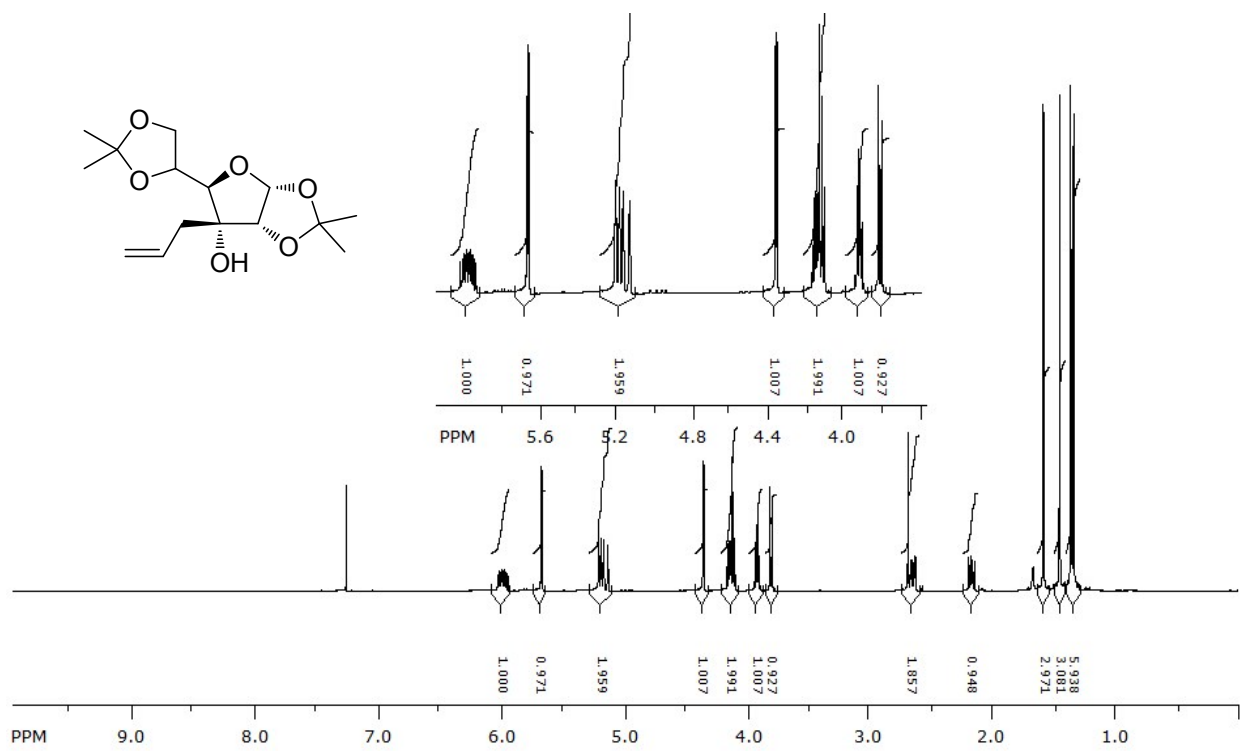
Fig. 1 ^{1}H NMR (500MHz; CDCl$_3$) of compound **5**.
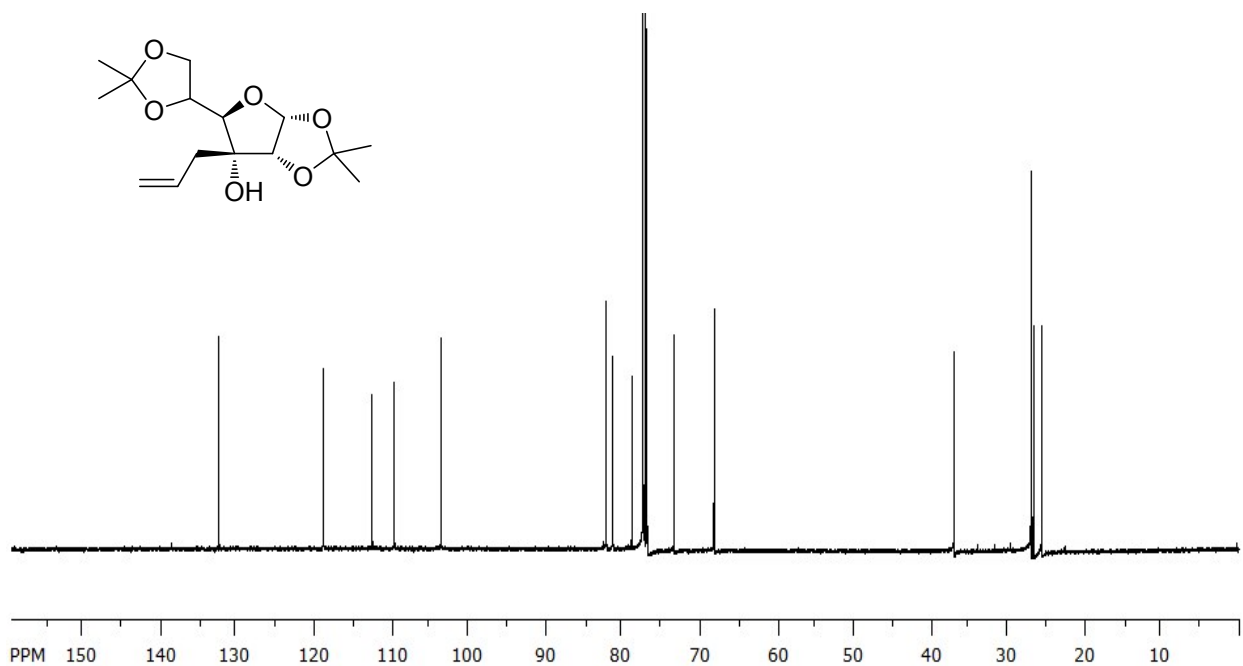
Fig. 2 ^{13}C NMR (125MHz; CDCl$_3$) of compound **5**.
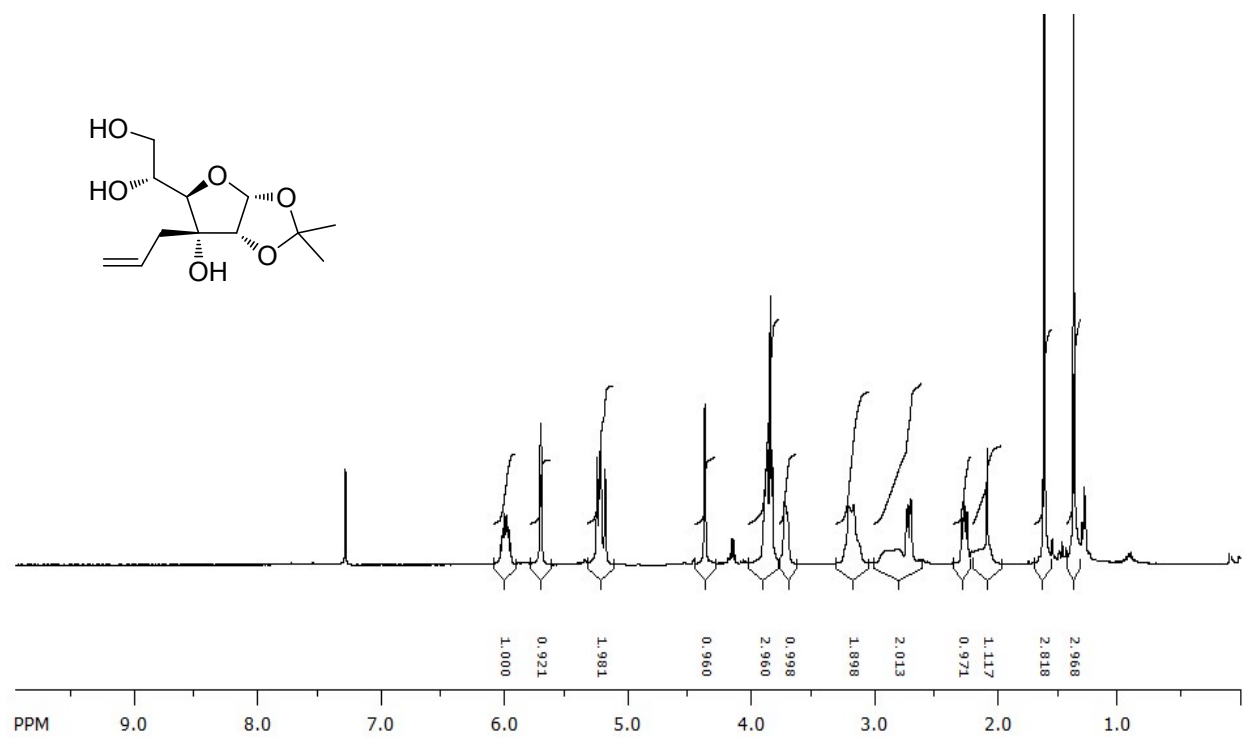
Fig. 3 ^{1}H NMR (500MHz; CDCl$_3$) of compound **6**.
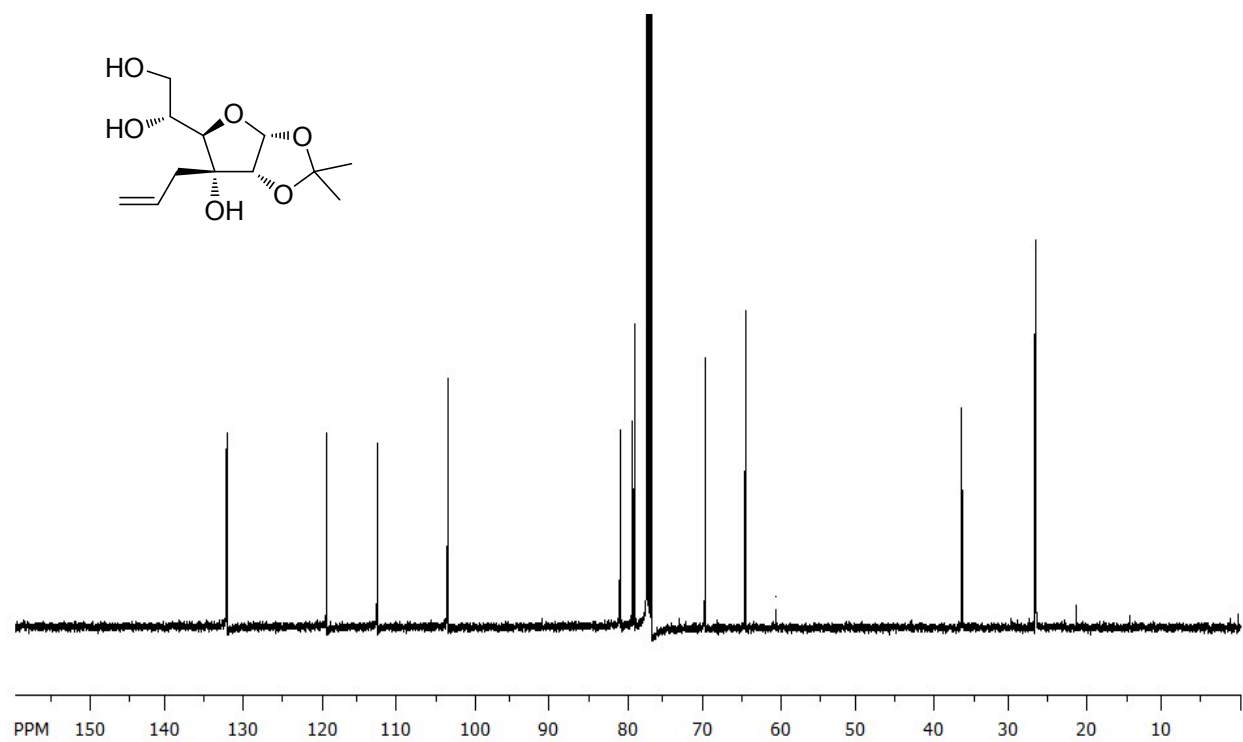
Fig. 4 ^{13}C NMR (125MHz; CDCl$_3$) of compound **6**.
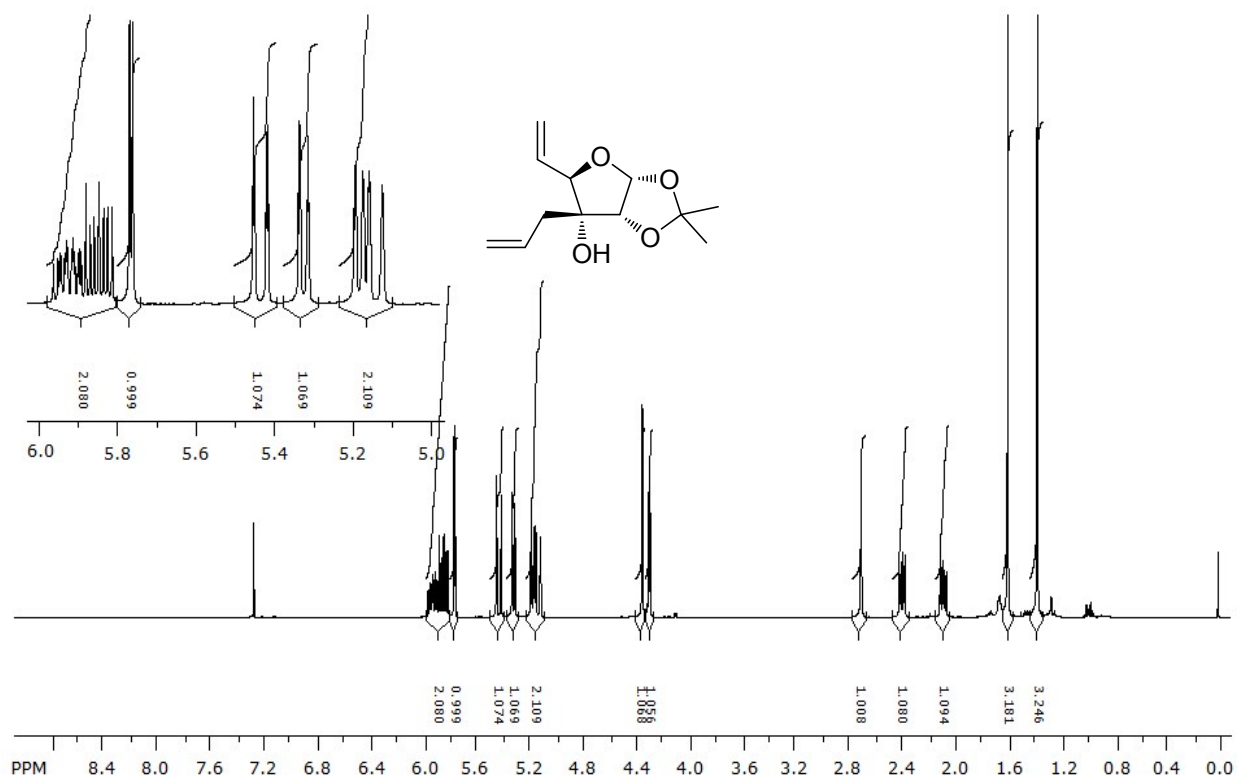
Fig. 5 ^{1}H NMR (500MHz; CDCl$_3$) of compound **7**.
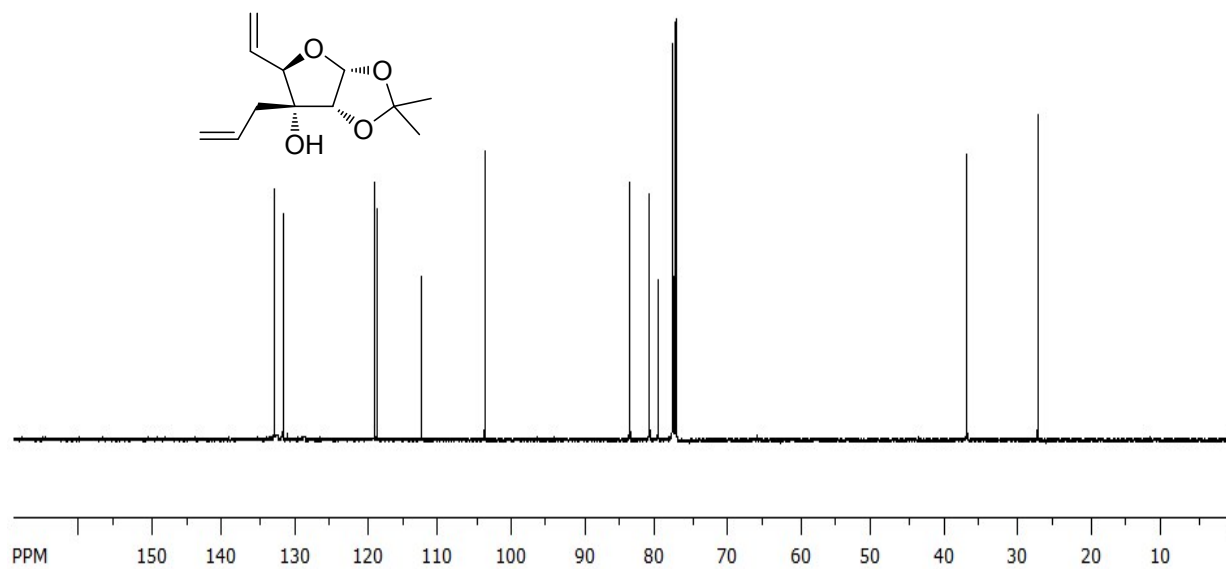
Fig. 6 ^{13}C NMR (125MHz; CDCl$_3$) of compound **7**.
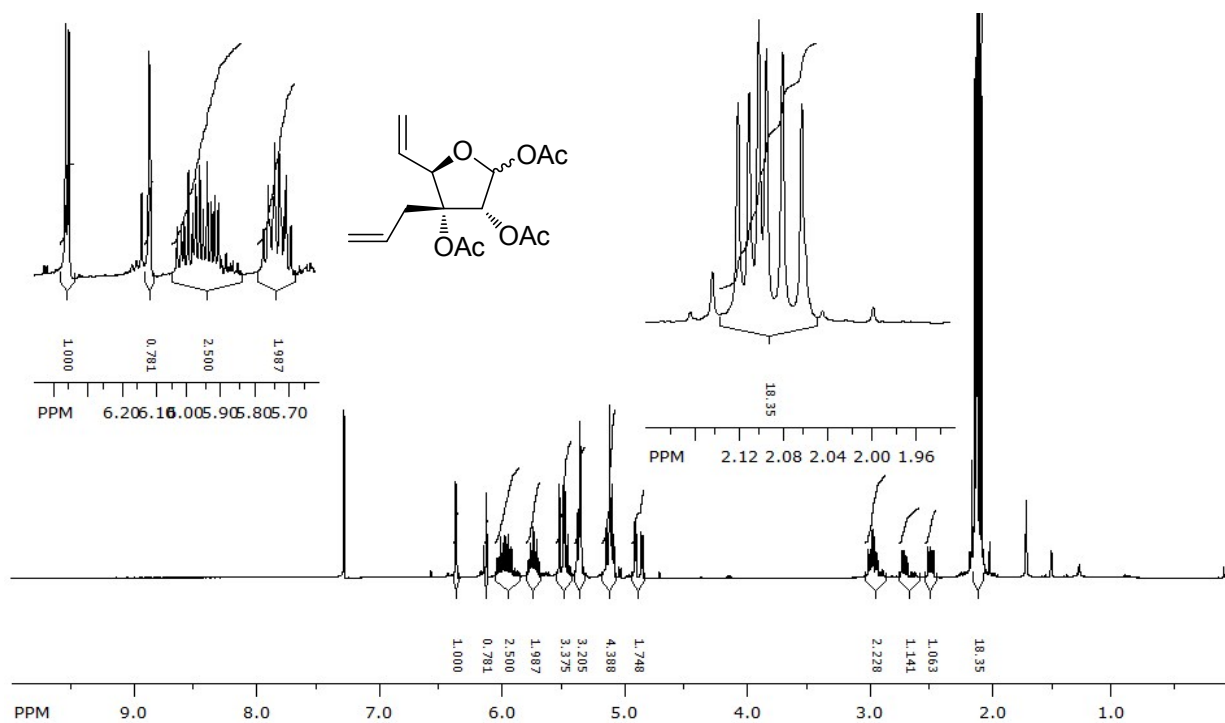
Fig. 7 ^{1}H NMR (500MHz; CDCl$_3$) of compound **8**.
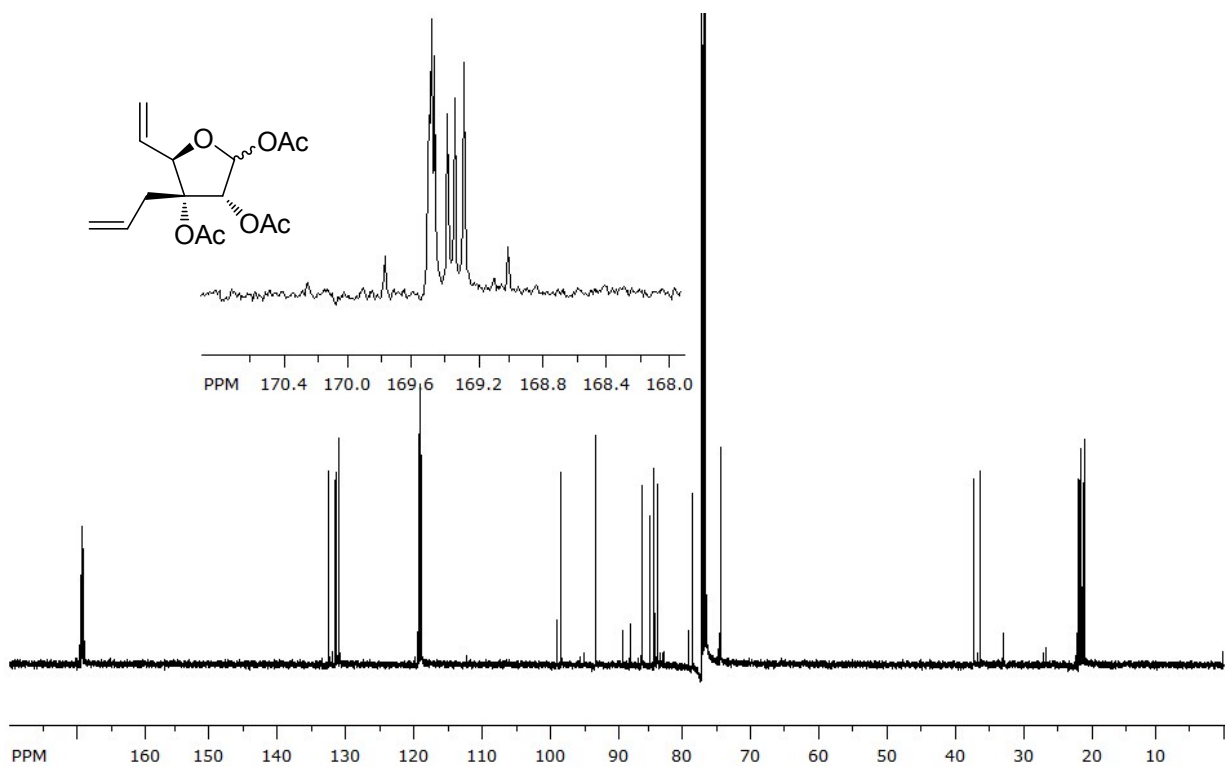
Fig. 8 ^{13}C NMR (125MHz; CDCl$_3$) of compound **8.**
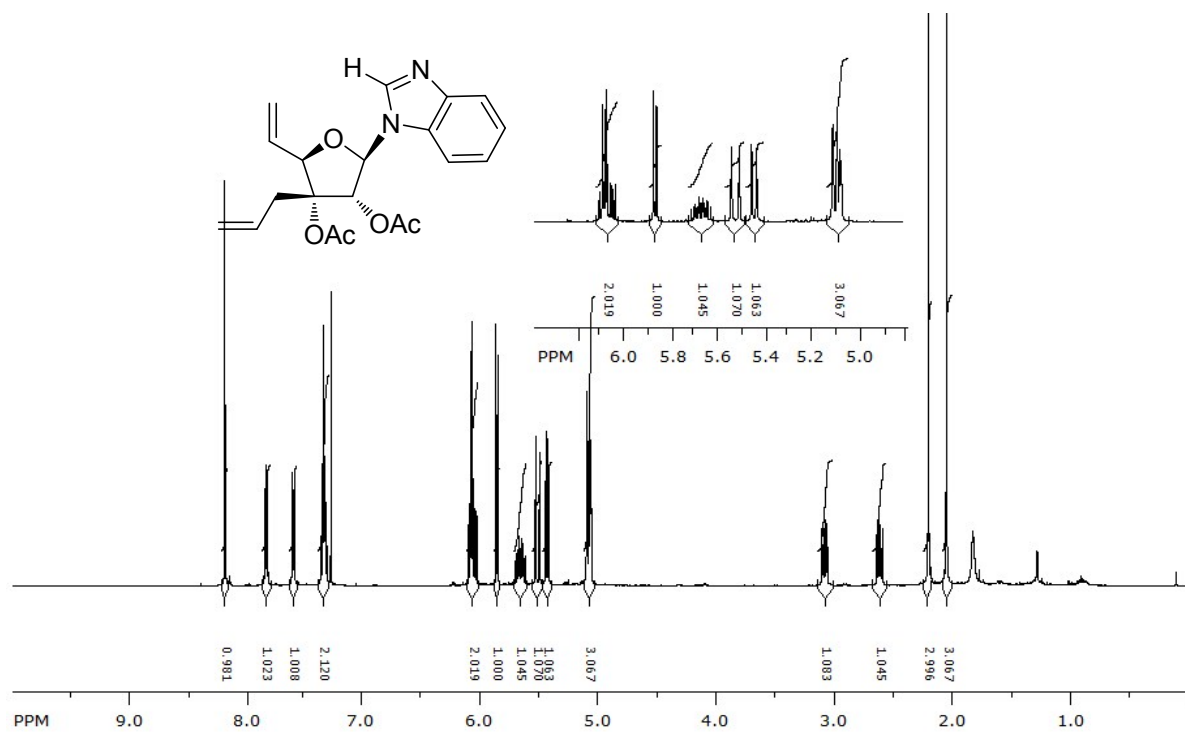
Fig. 9 ^{1}H NMR (500MHz; CDCl$_3$) of compound **9**.
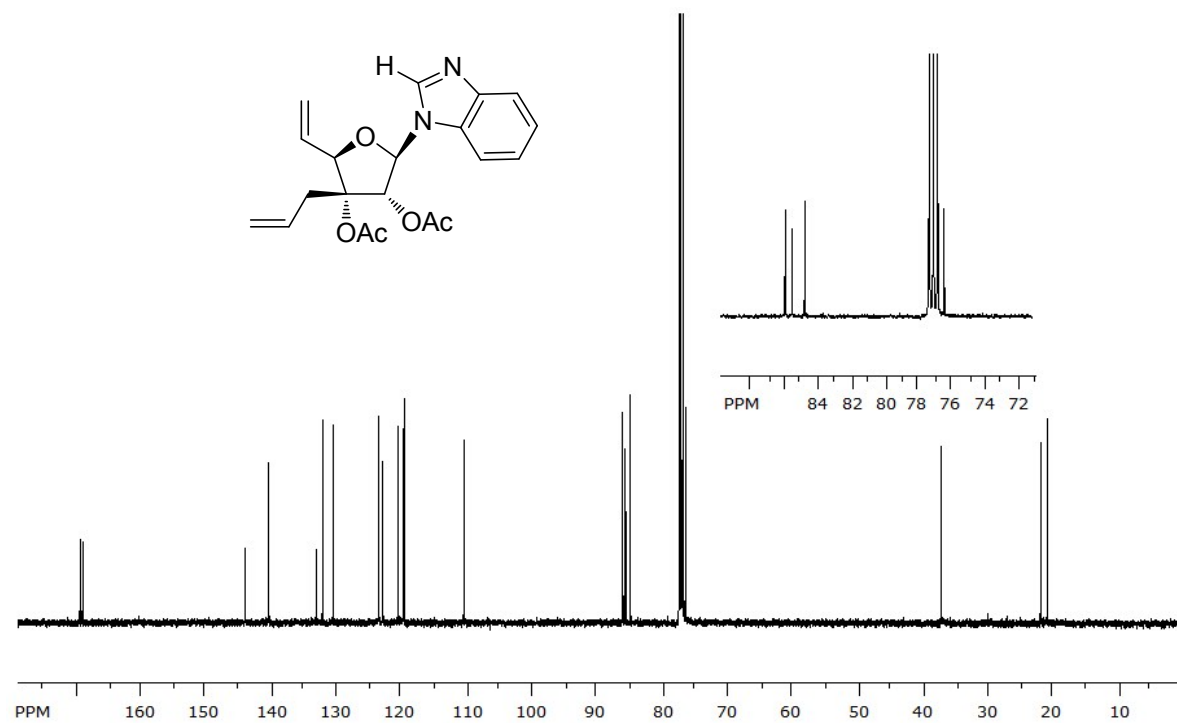
Fig. 10 ^{13}C NMR (125MHz; CDCl$_3$) of compound **9**.
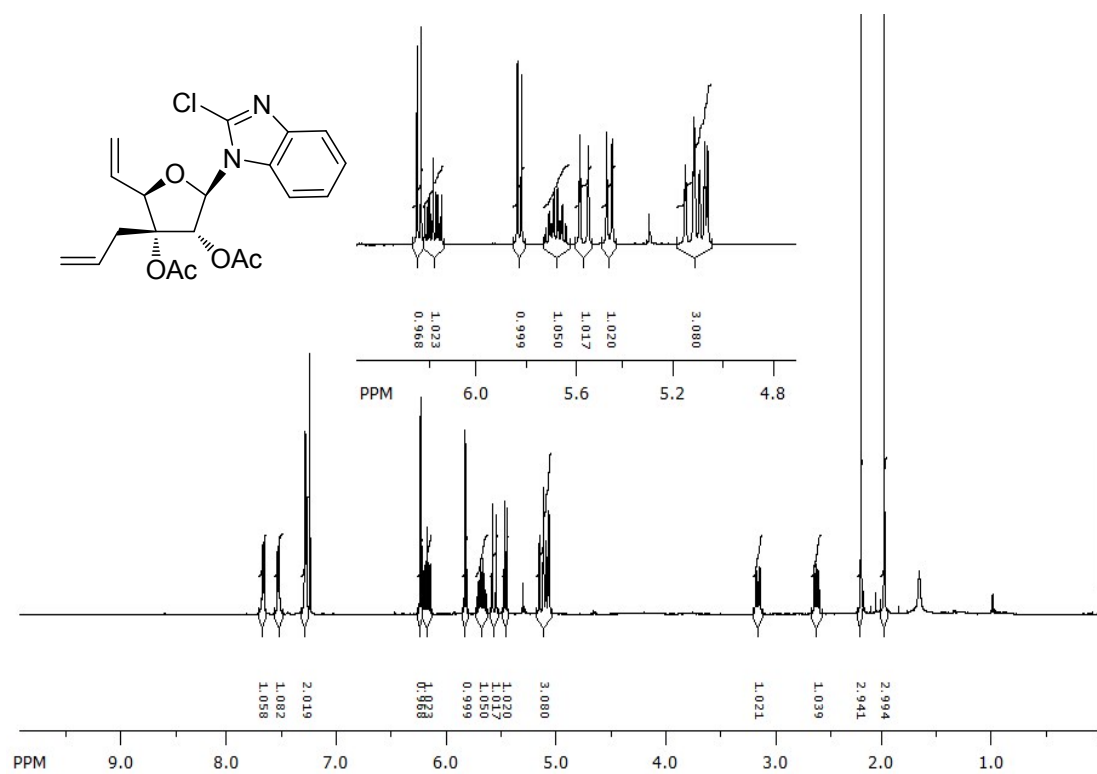
Fig. 11 ^{1}H NMR (500MHz; CDCl$_3$) of compound **10**.
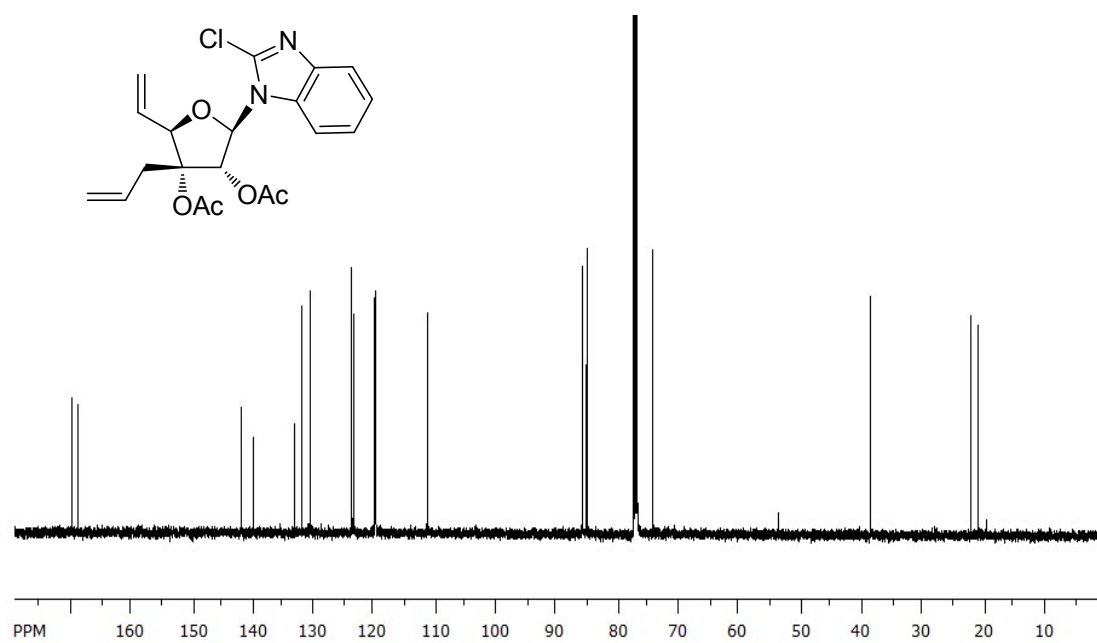
Fig. 12 ^{13}C NMR (125MHz; CDCl$_3$) of compound **10.**

14
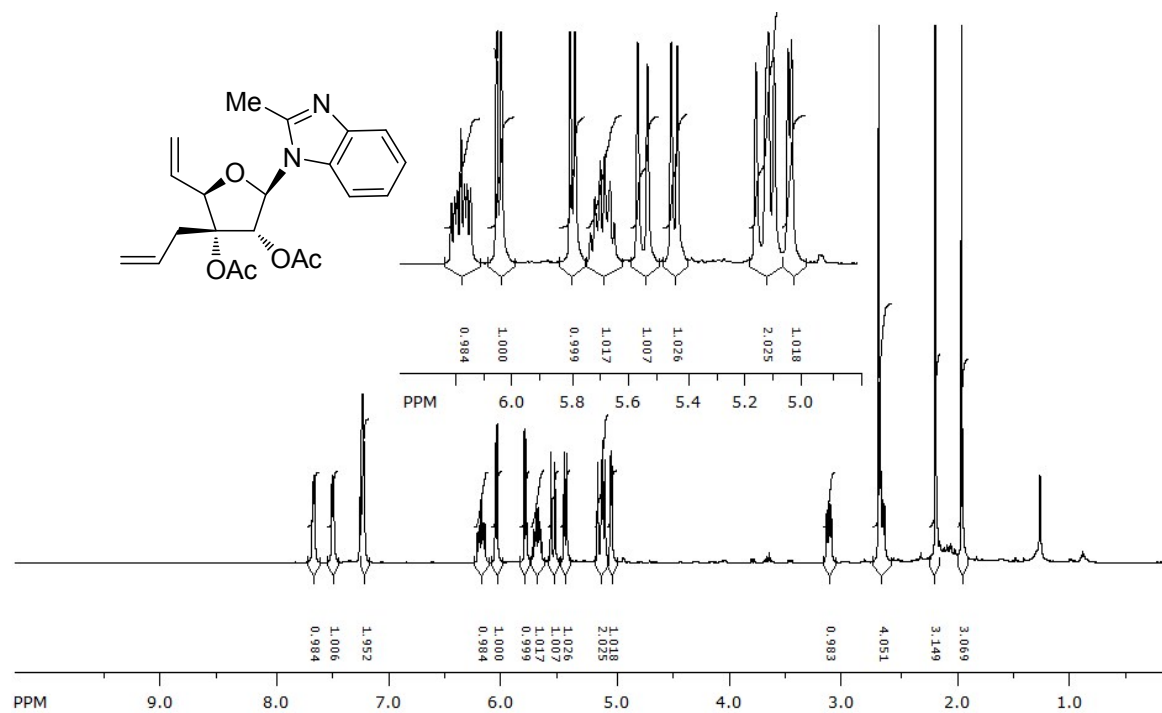
Fig. 13 ^{1}H NMR (500MHz; CDCl$_3$) of compound **11**.
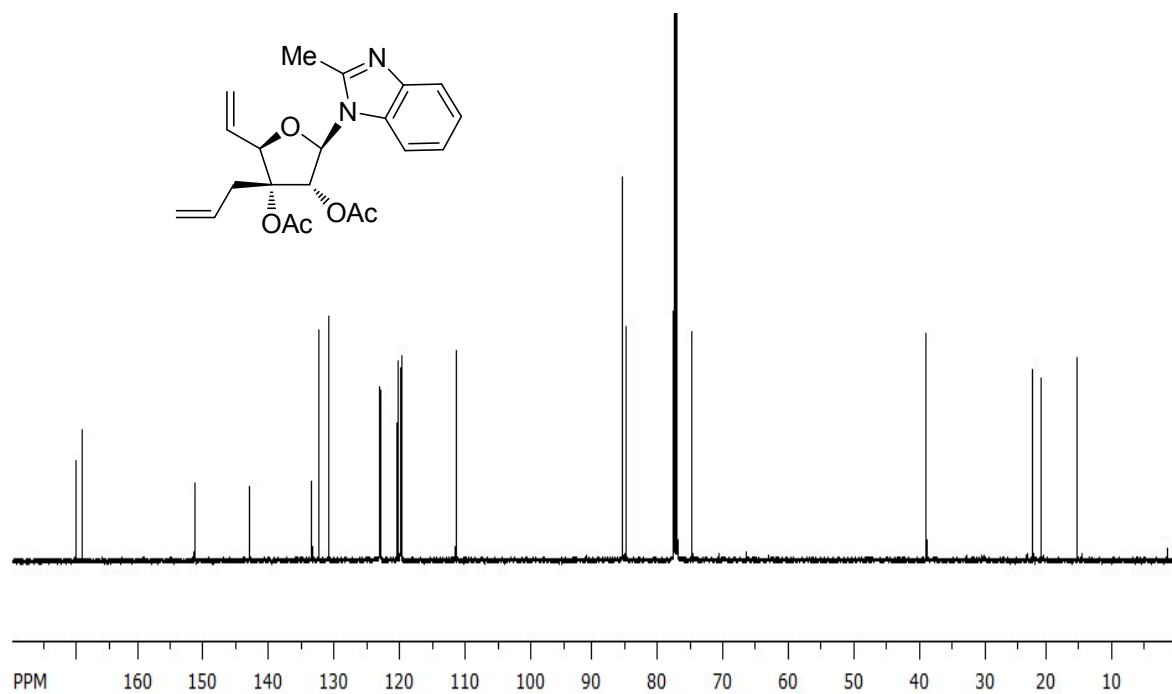
Fig. 14 ^{13}C NMR (125MHz; CDCl$_3$) of compound **11.**
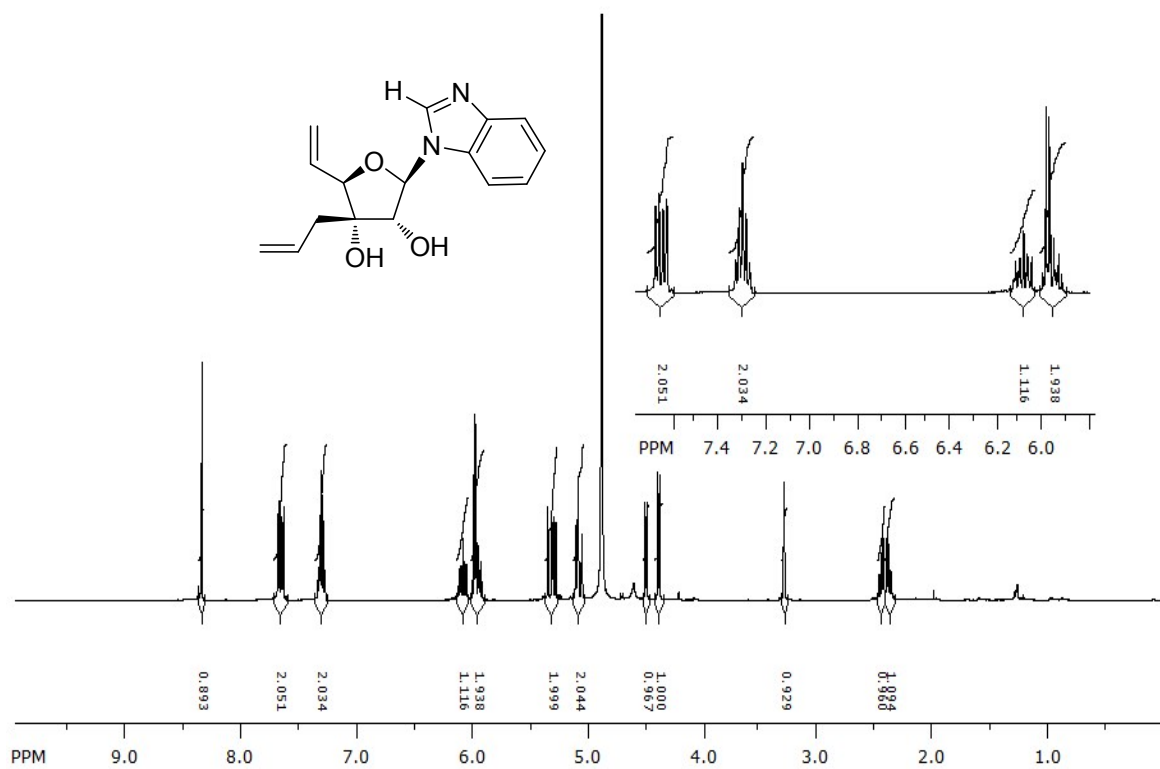
Fig. 15 ^{1}H NMR (500MHz; CD$_3$OD) of compound **12**.
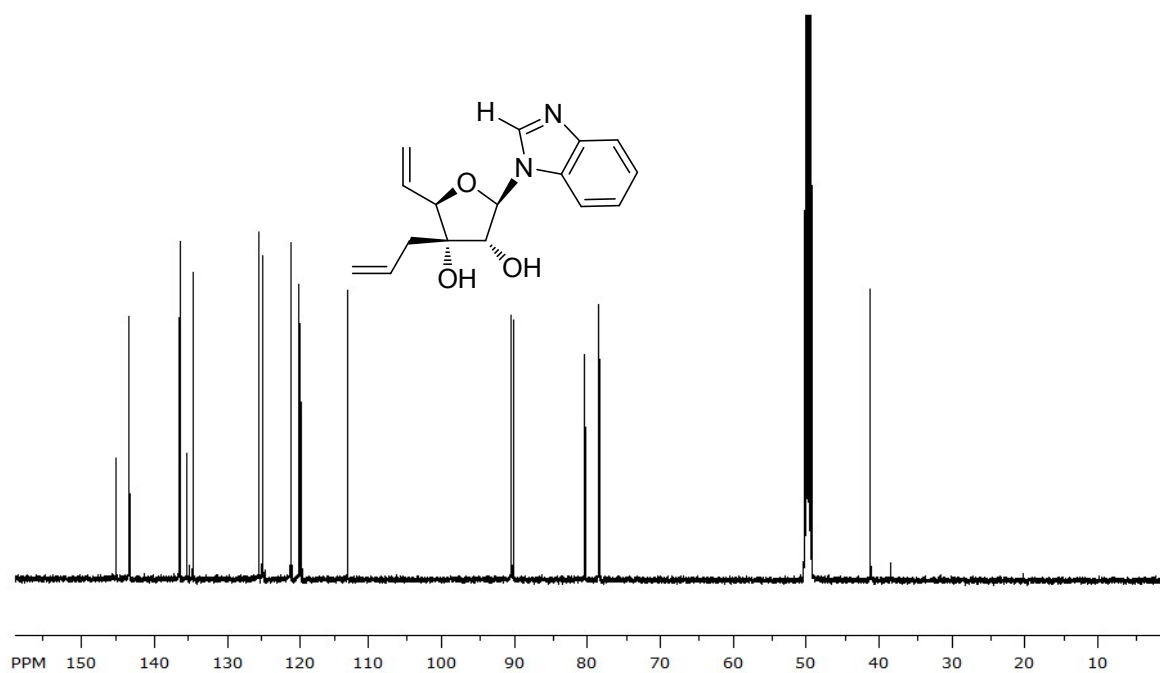
Fig. 16 ^{13}C NMR (125MHz; CD$_3$OD) of compound **12**.

16
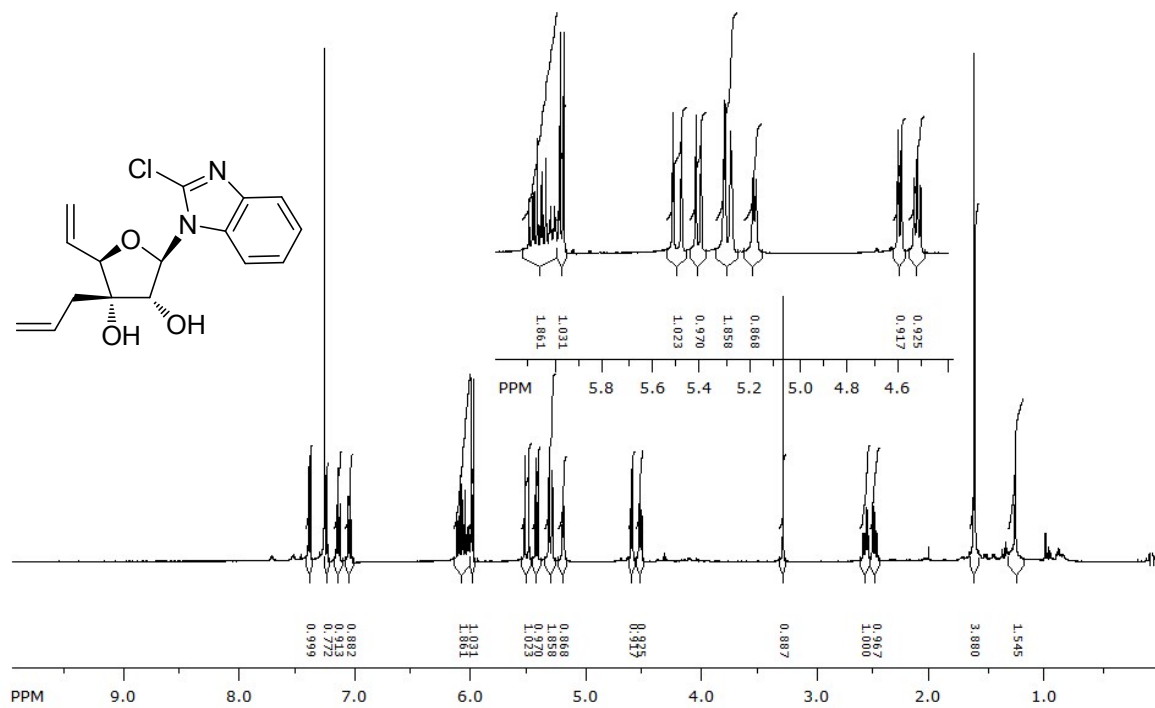
Fig. 17 ^{1}H NMR (500MHz; CDCl$_3$) of compound **13**.
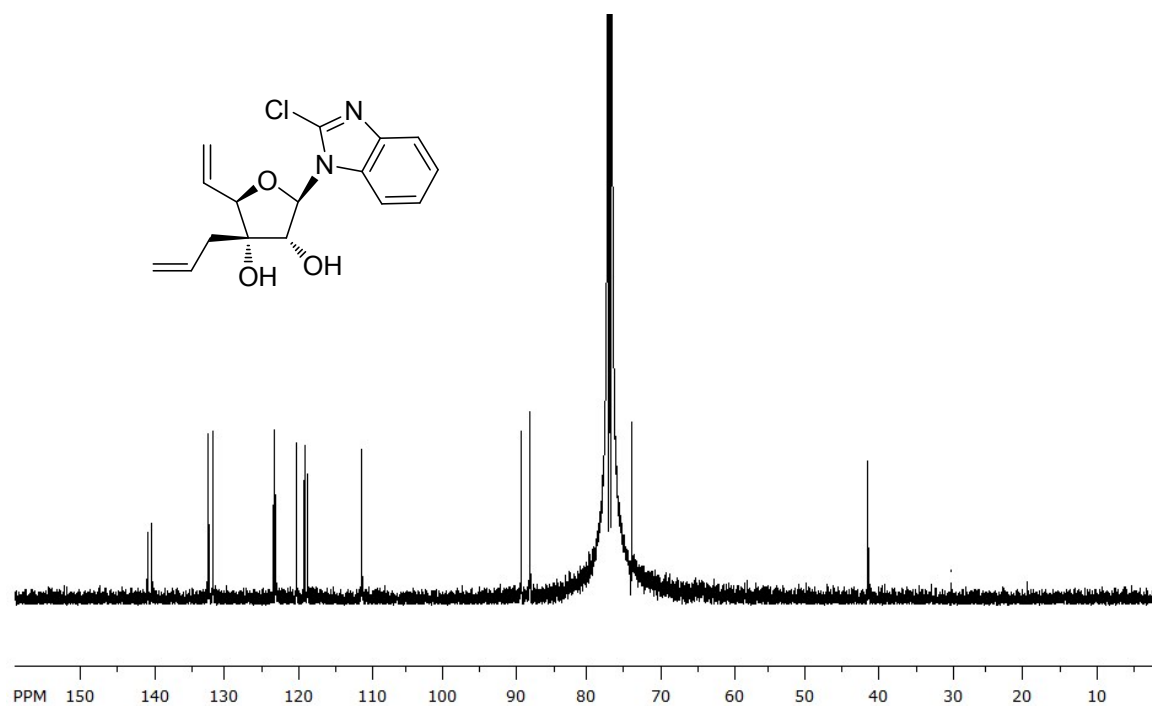
Fig. 18 ^{13}C NMR (125MHz; CDCl$_3$) of compound **13**.
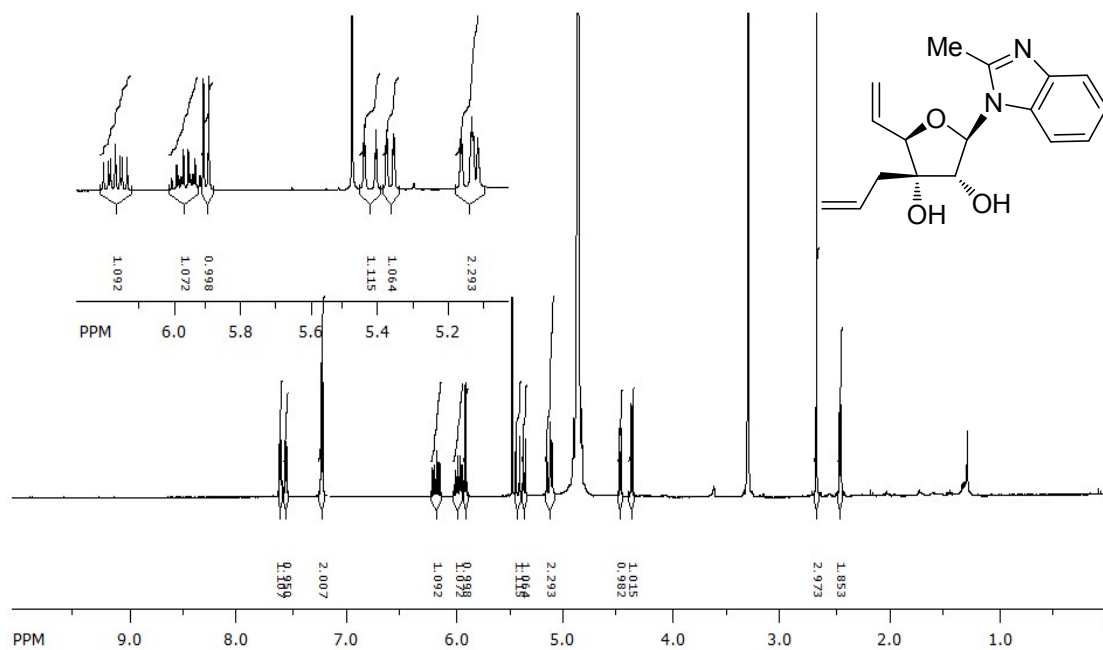
Fig. 19 ^{1}H NMR (500MHz; CD$_3$OD) of compound **14**.
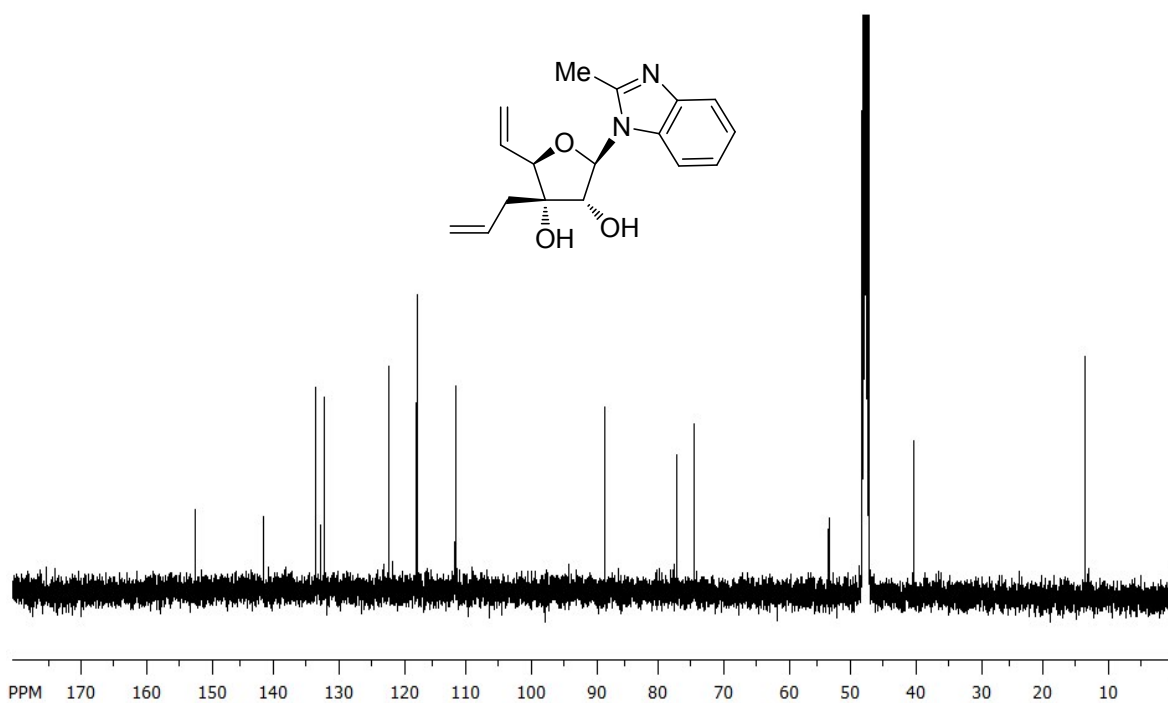
Fig. 20 ^{13}C NMR (125MHz; CD$_3$OD) of compound **14.**
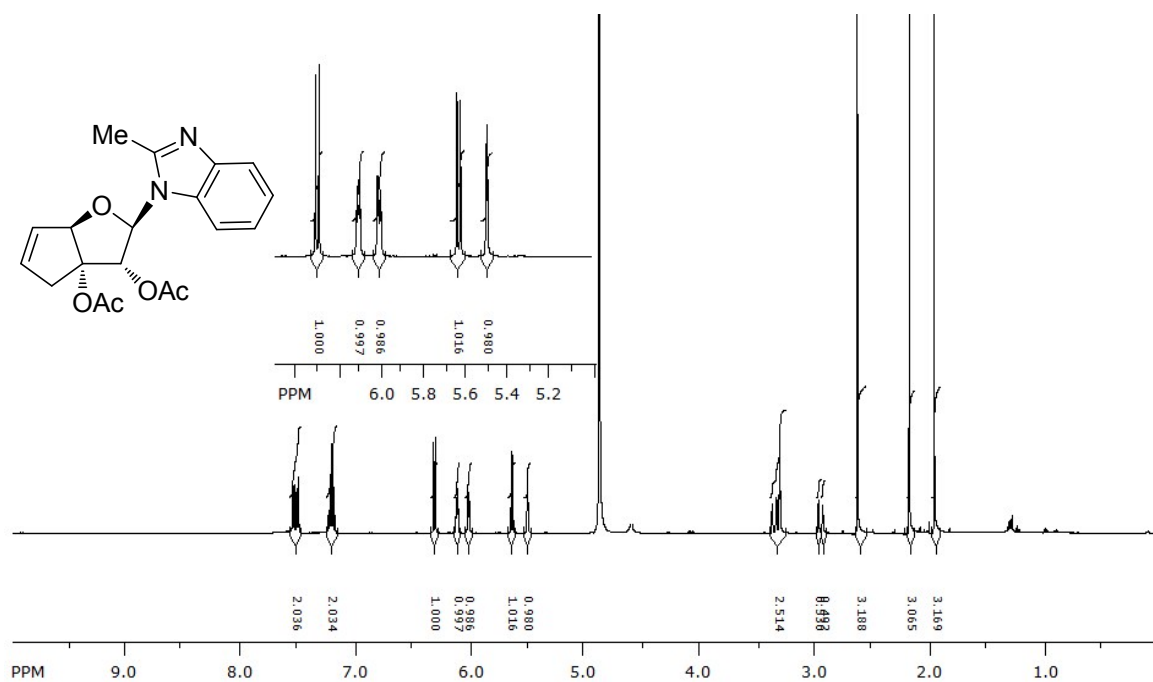
Fig. 21 ^{1}H NMR (500MHz; CD$_3$OD) of compound **15**.
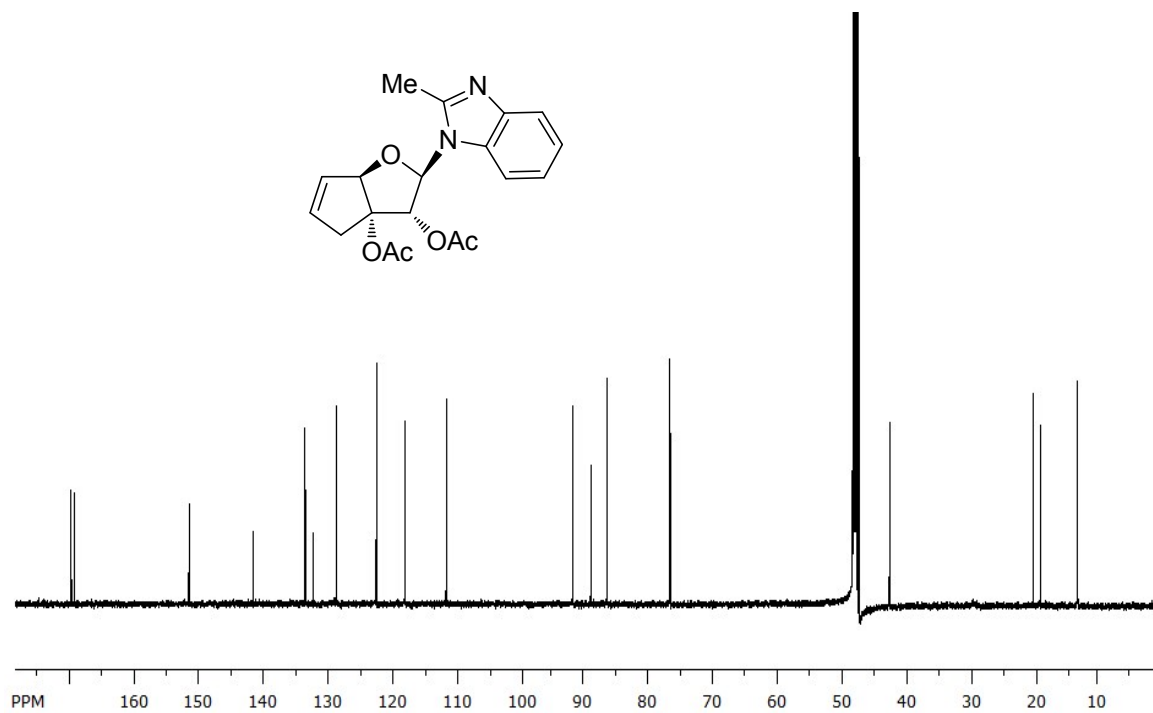
Fig. 22 ^{13}C NMR (125MHz; CD$_3$OD) of compound **15.**
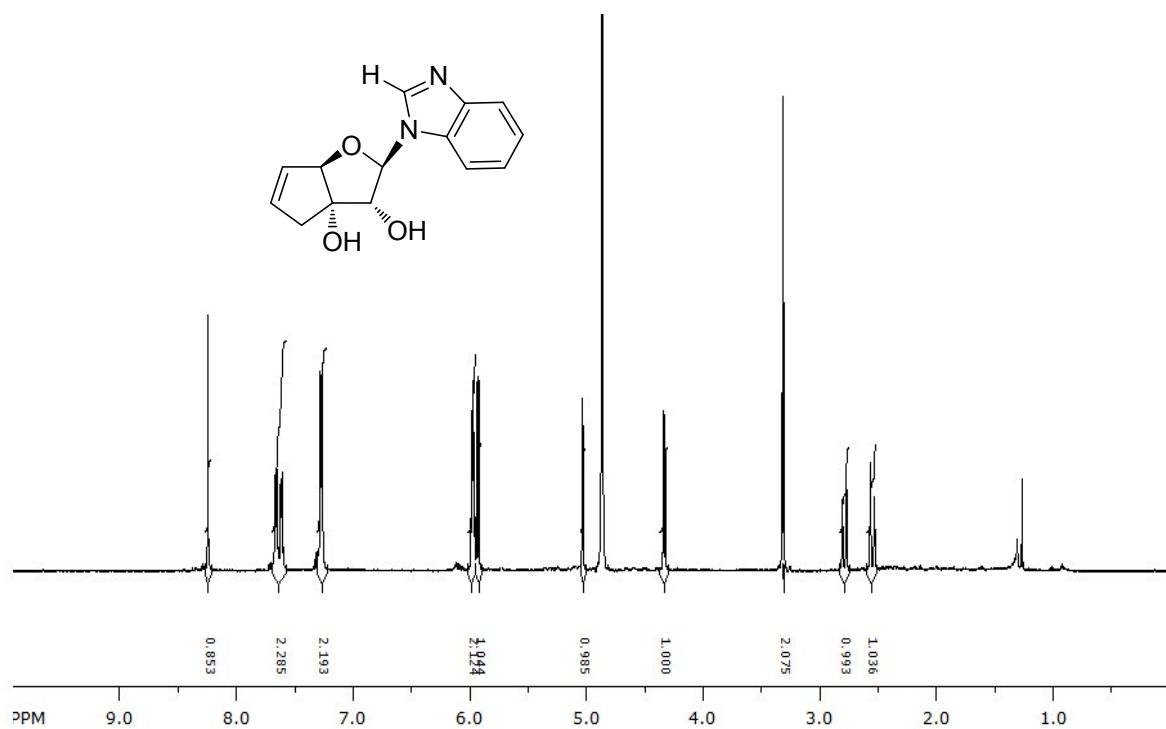
Fig. 23 ^{1}H NMR (500MHz; CD$_3$OD) of compound **1**.
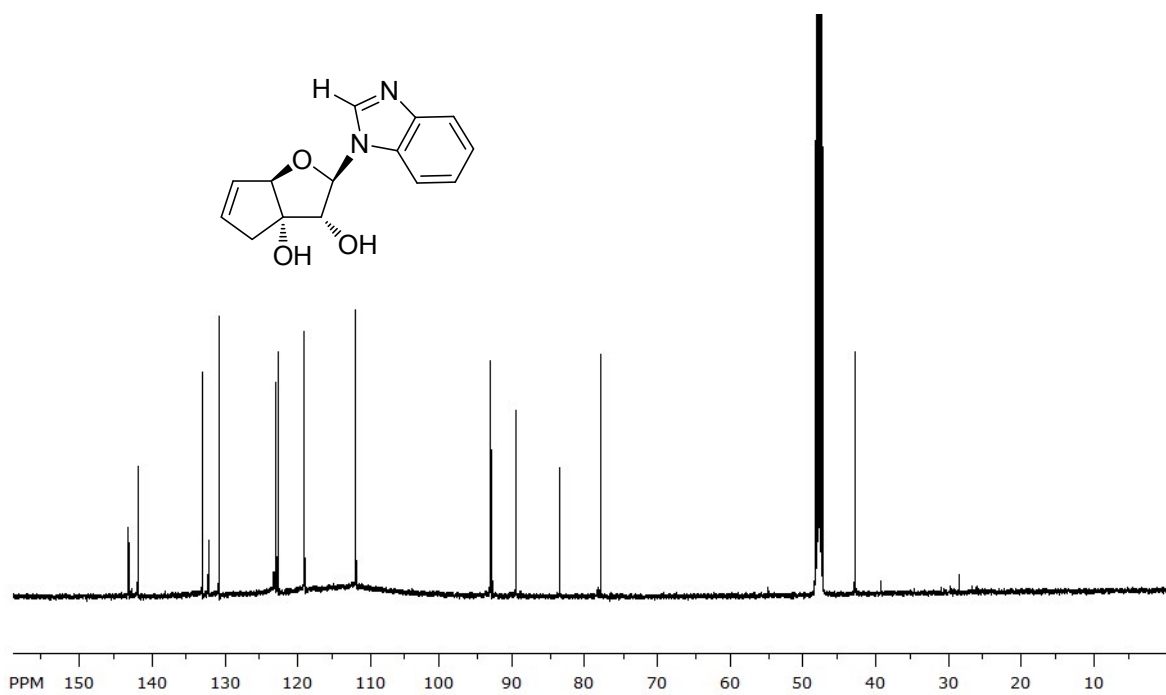
Fig. 24 ^{1}H NMR (125MHz; CD$_3$OD) of compound **1**.
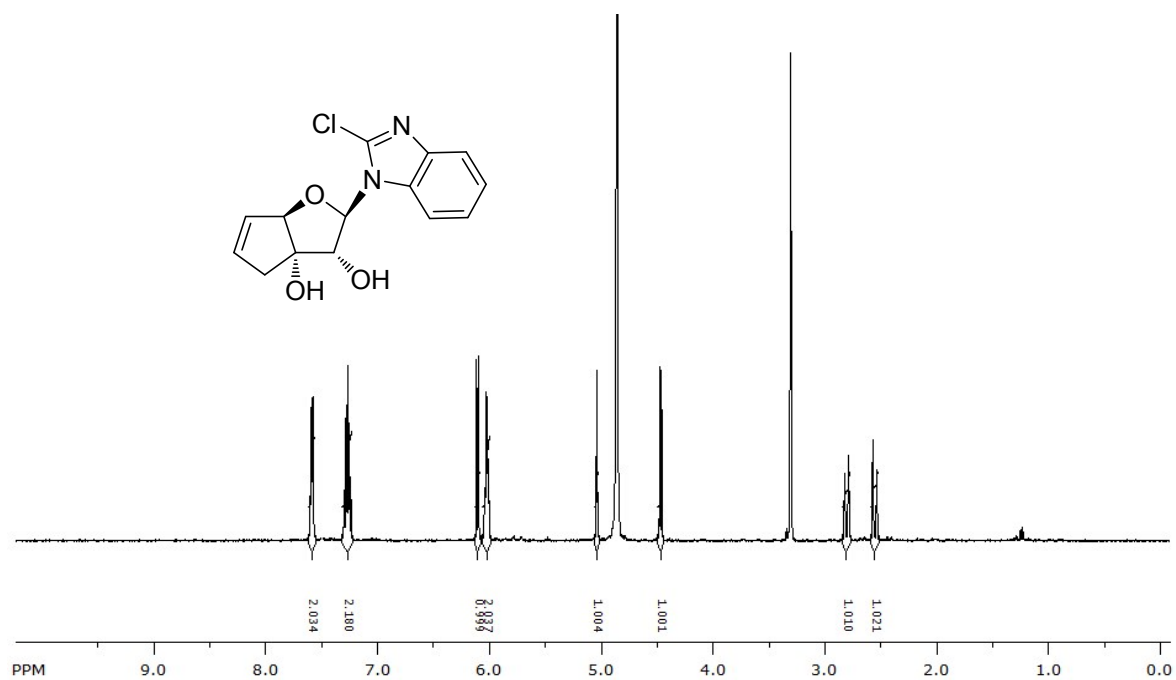
Fig. 25 ^{1}H NMR (500MHz; CD$_3$OD) of compound **2**.
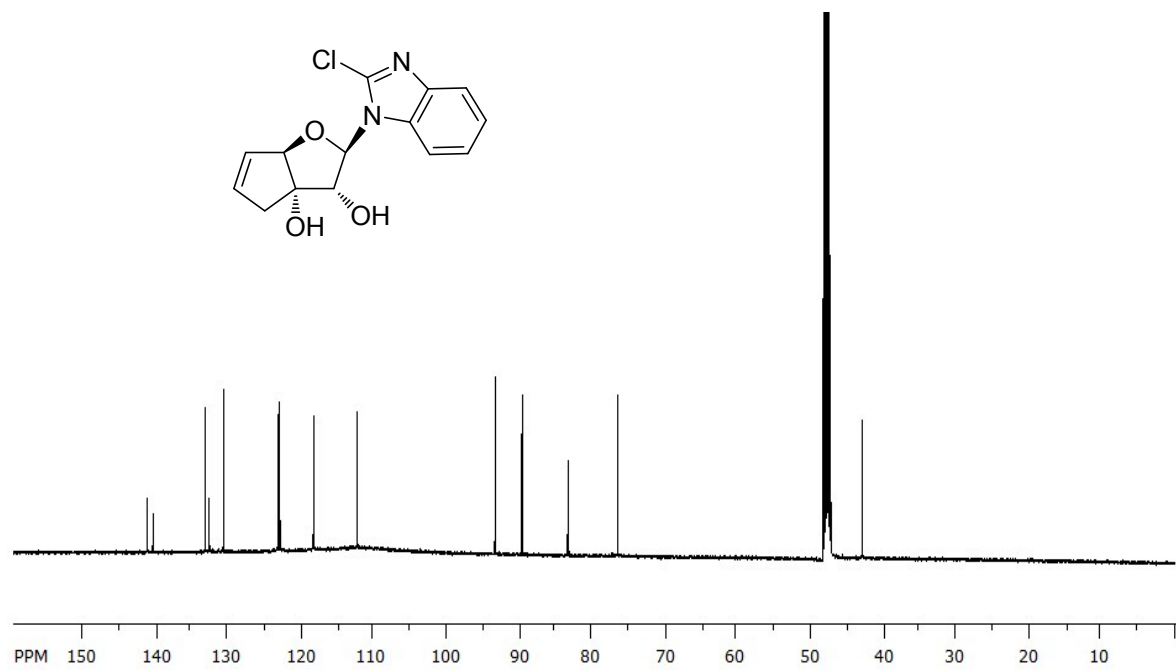
Fig. 26 ^{13}C NMR (125MHz; CD$_3$OD) of compound **2.**
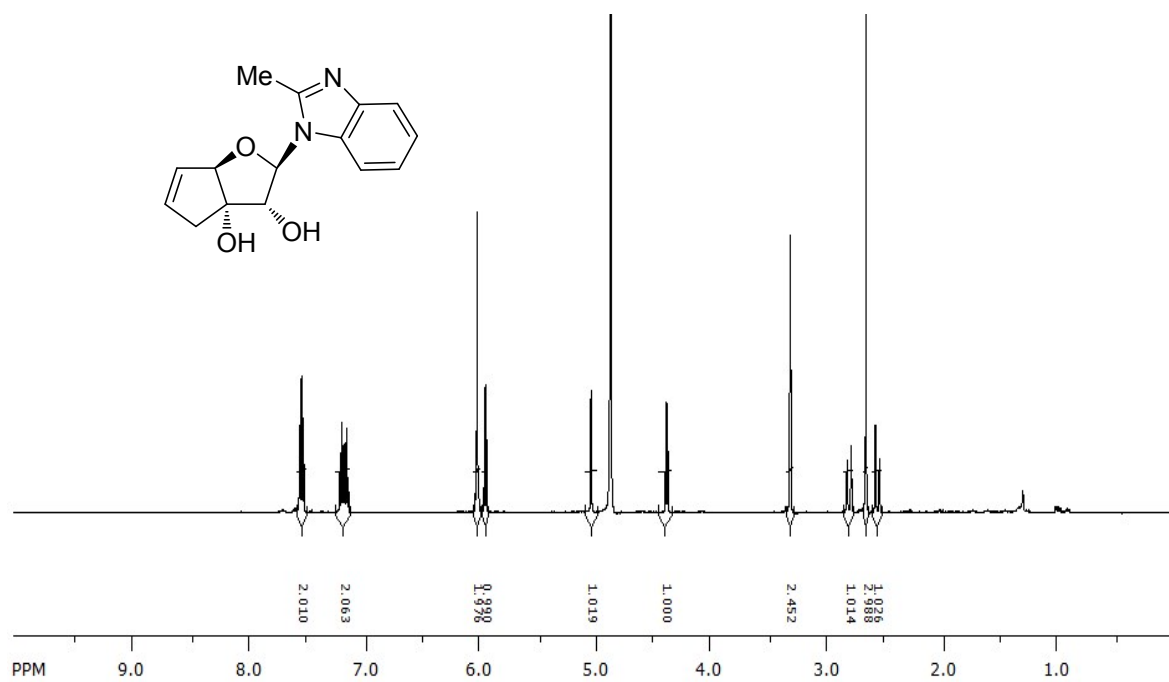
Fig. 27 ^{1}H NMR (500MHz; CD$_3$OD) of compound **3**.
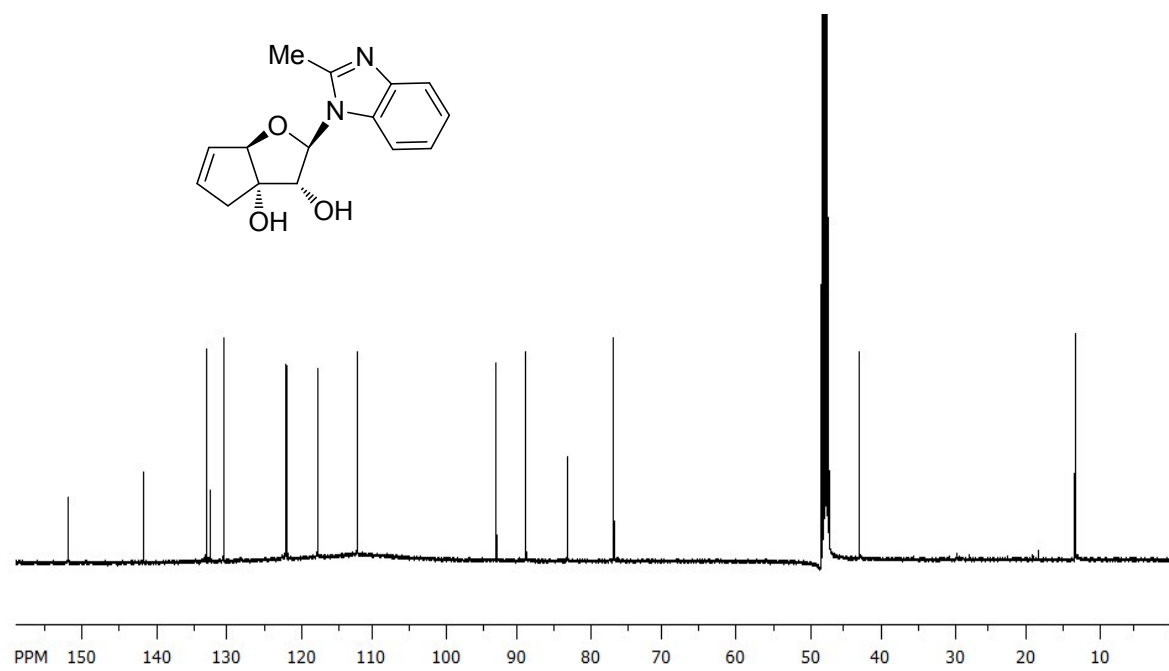
Fig. 28 ^{13}C NMR (125MHz; CD$_3$OD) of compound **3.**
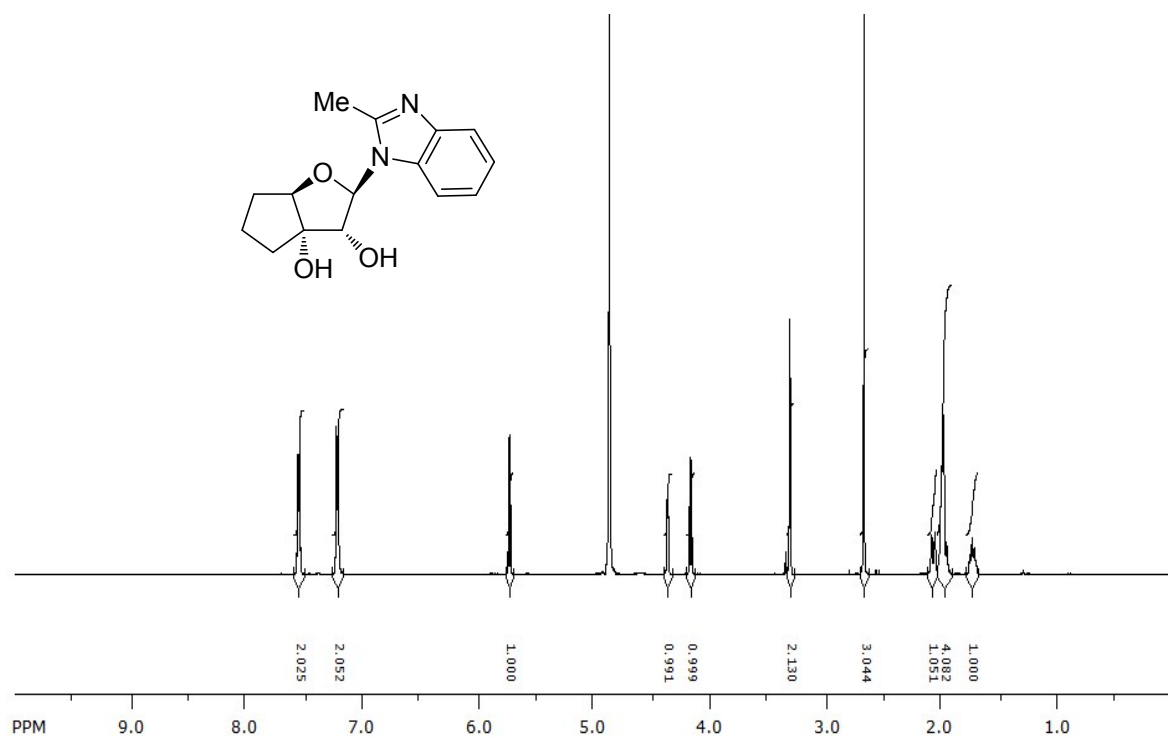
Fig. 29 ^{1}H NMR (500MHz; CD$_3$OD) of compound **4**.
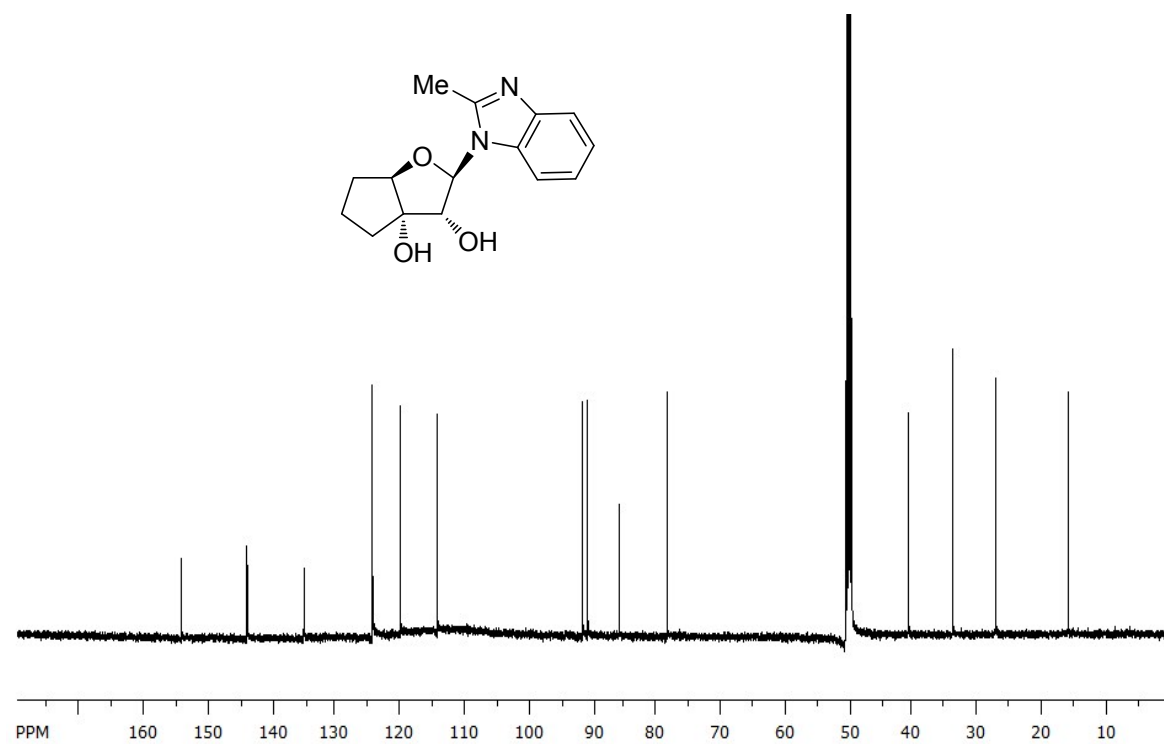
Fig. 30 ^{13}C NMR (125MHz; CD$_3$OD) of compound **4.**

23
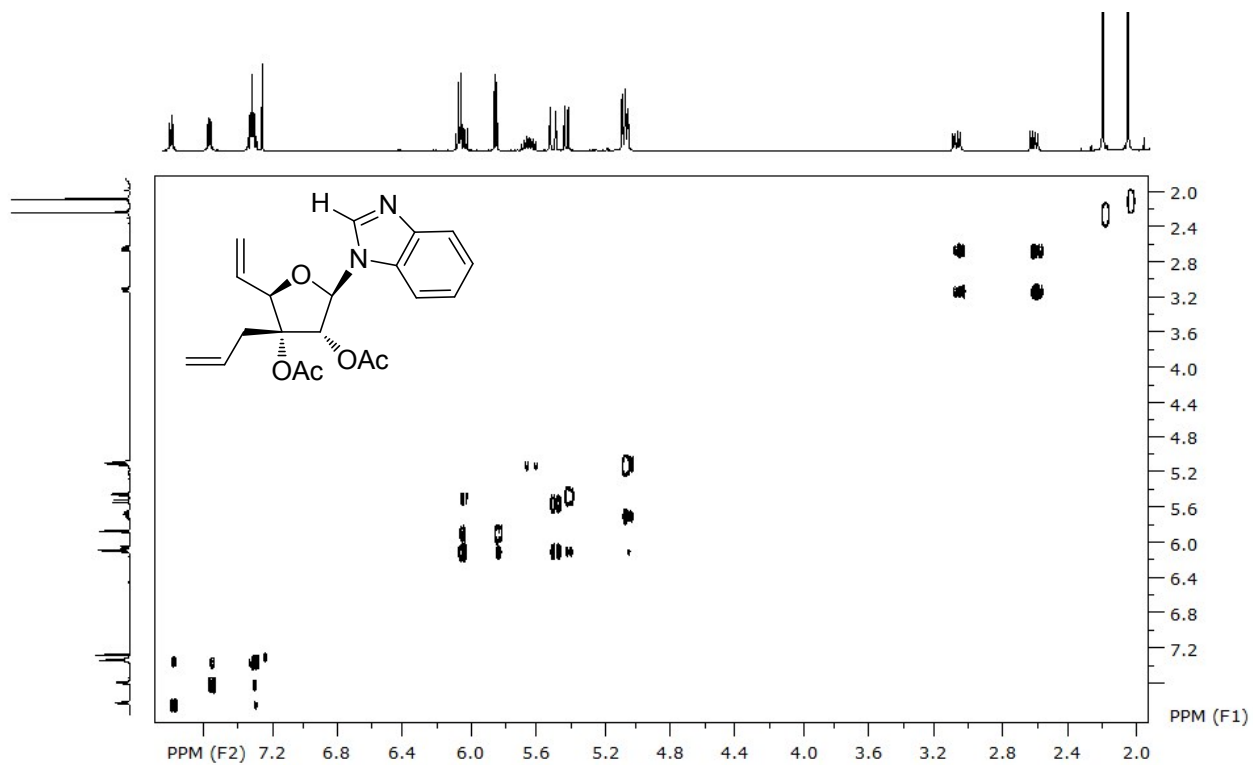
Fig. 31 COSY spectrum (CDCl$_3$) of compound **9**
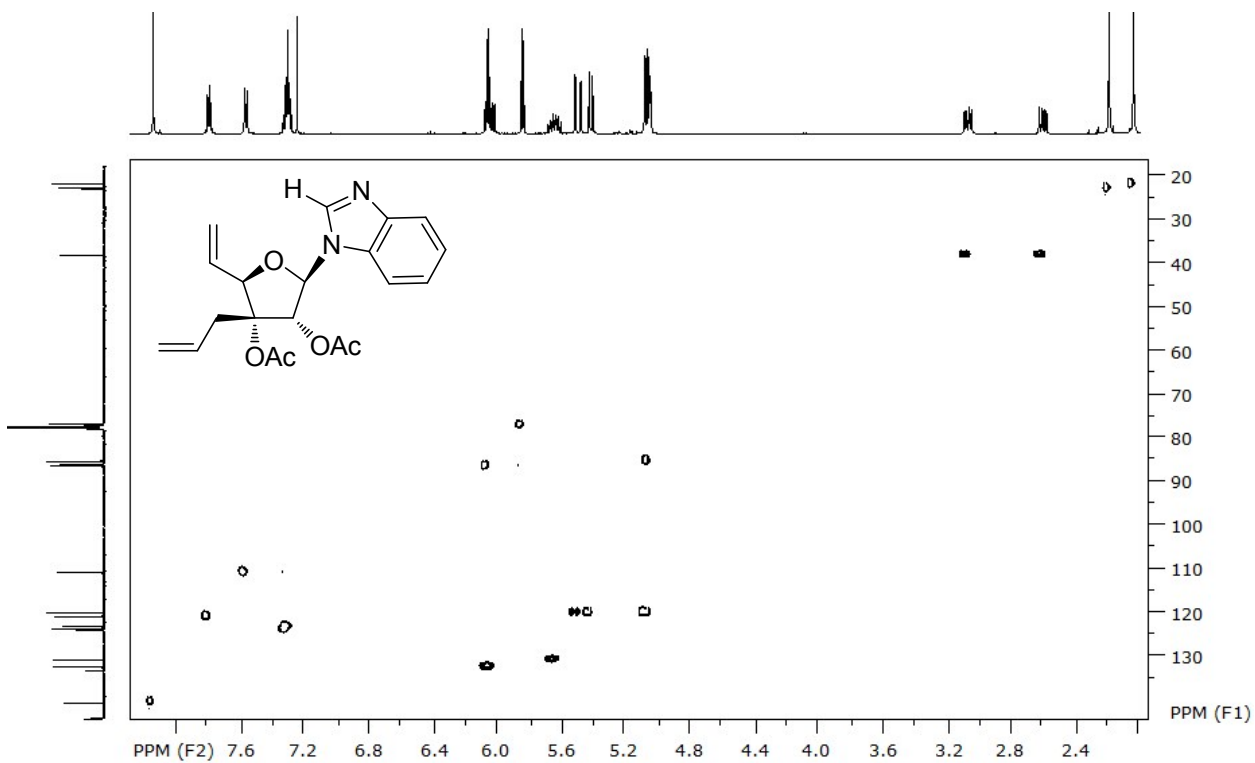
Fig. 32 HSQC spectrum (CDCl$_3$) of compound **9**

24
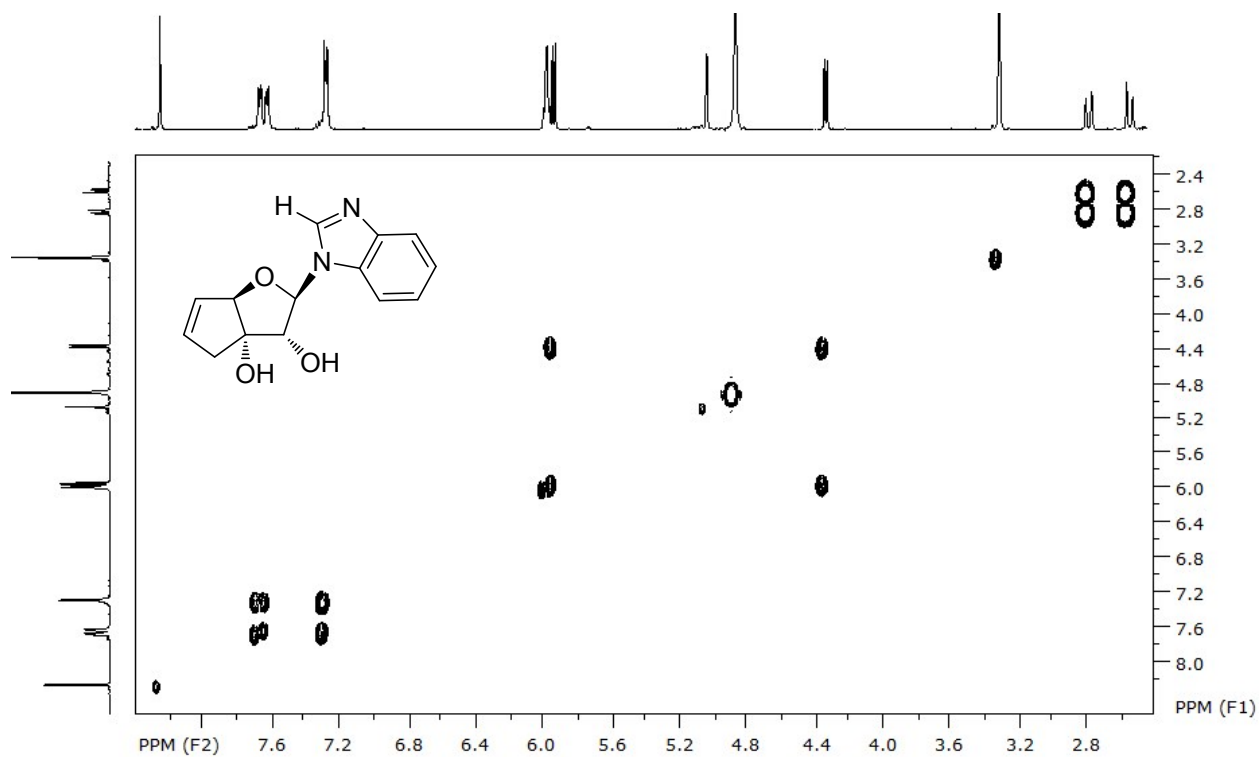
Fig. 33 COSY spectrum (CD$_3$OD) of compound **1**
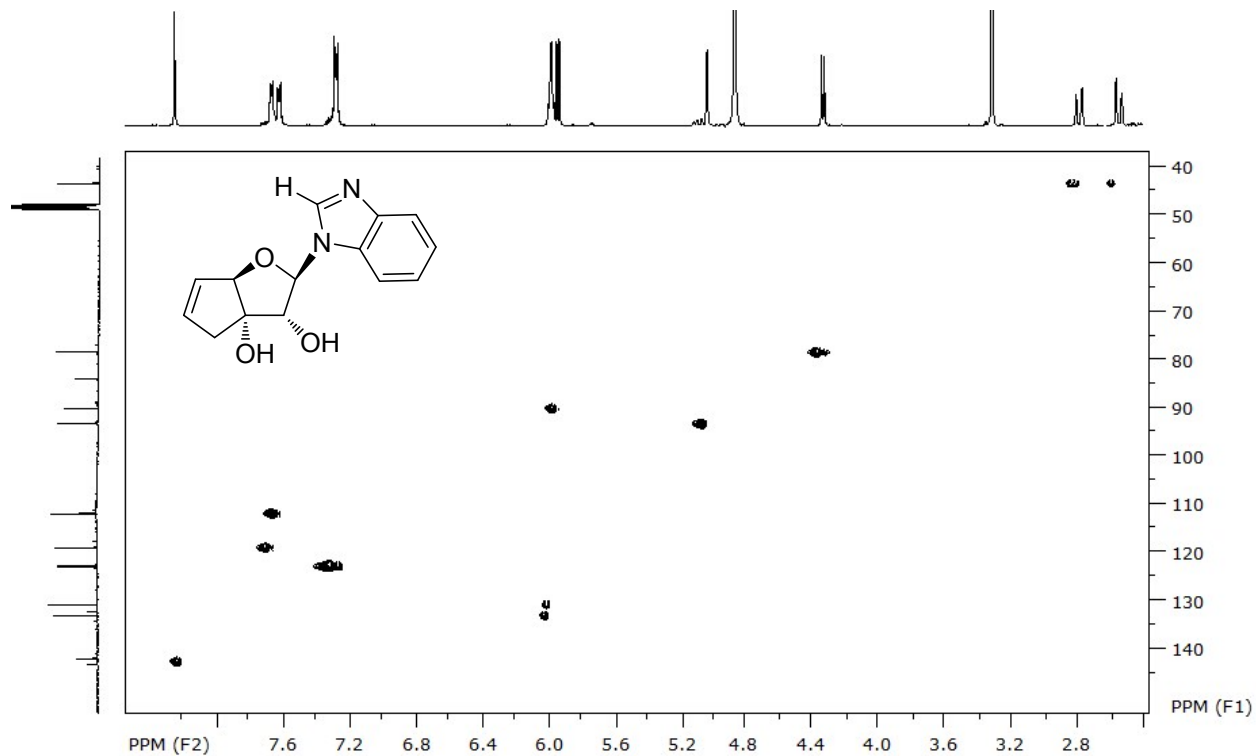
Fig. 34 HSQC spectrum (CD$_3$OD) of compound **1**

(A)

(B)
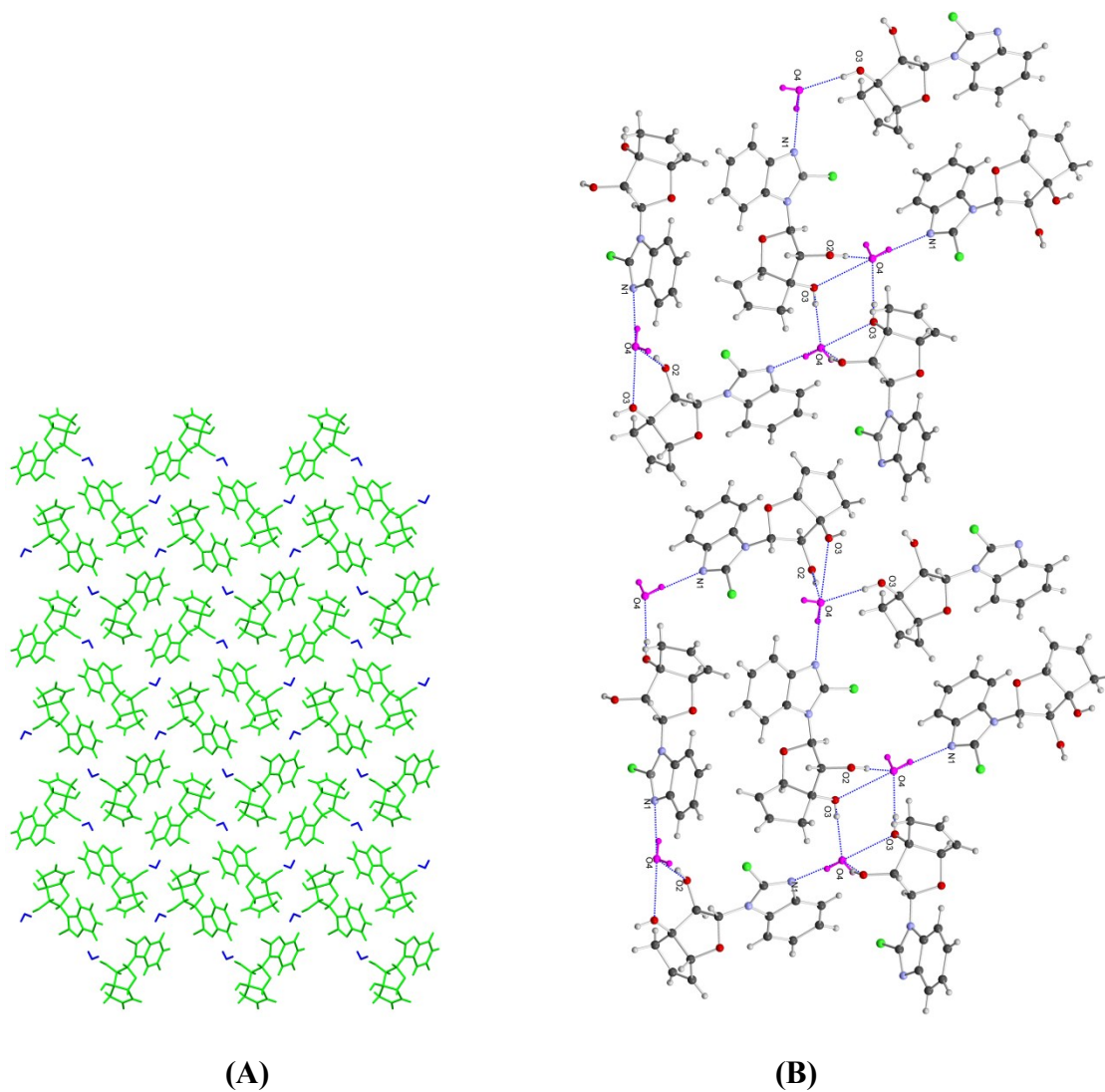
Fig. 35S: (A) Crystal packing of **2** viwed dawn c-axis. (B) H-bonding interactions of **2** in crystal

packing
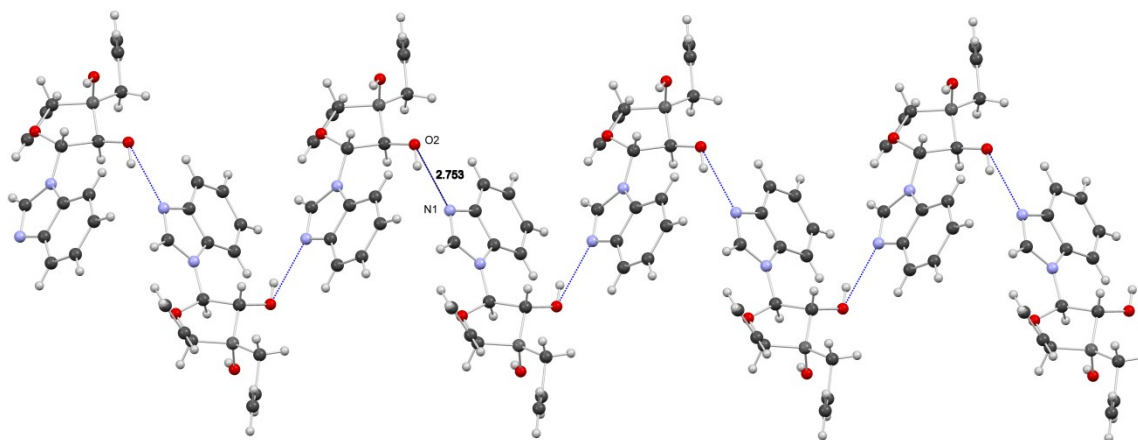
Fig. 36S: H-Bonding interactions in **12**
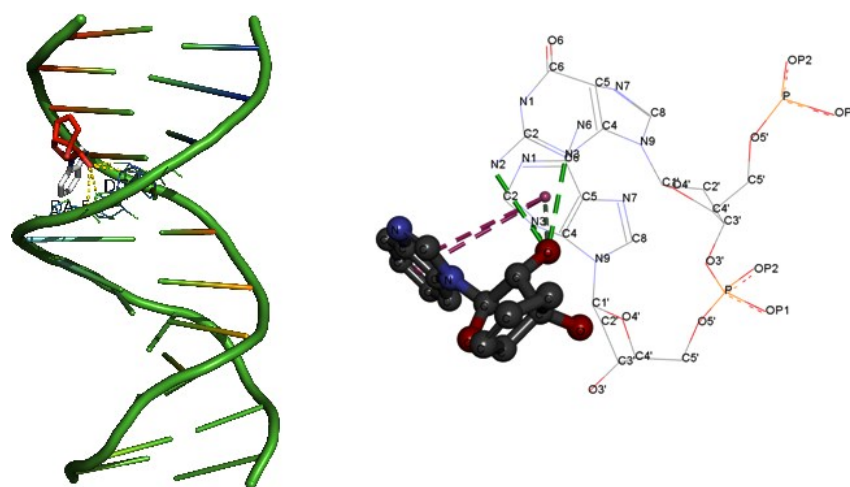
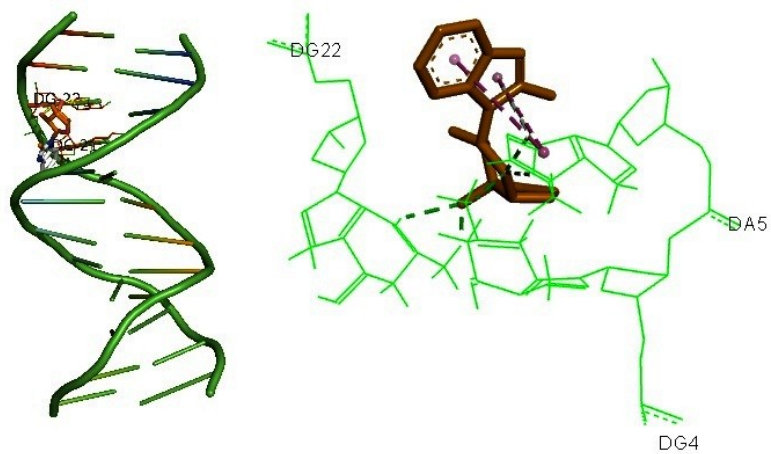**Fig. 37S:** Docked complex of compound **1**-DNA and H-bond interactions.

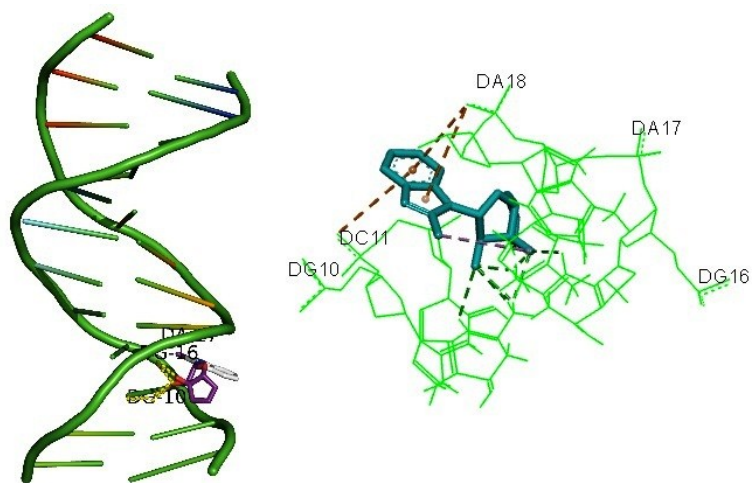**Fig. 38S** Docked complex and interaction of compound **3** with DNA.

Fig. 39S Docked pose of compound **4** and H-bond interactions with DNA.

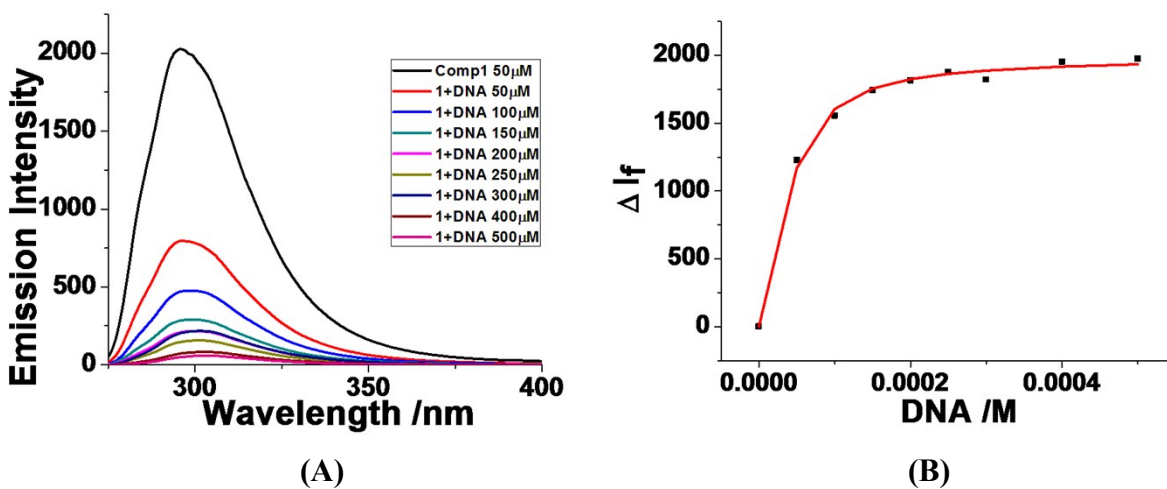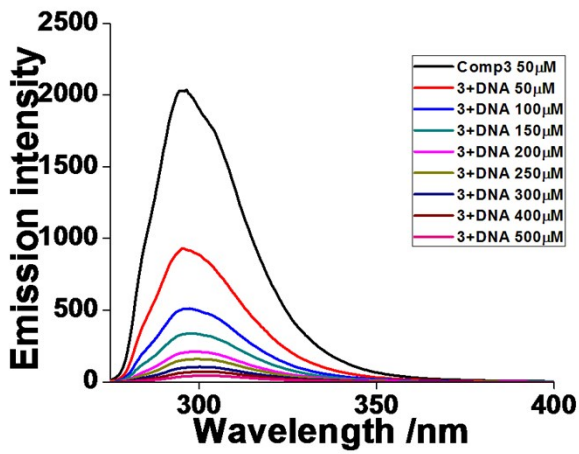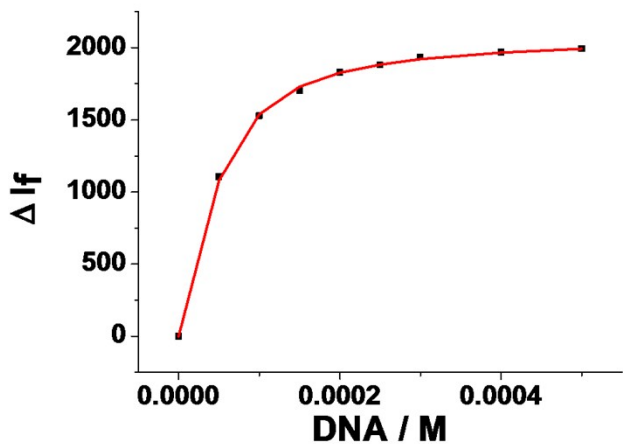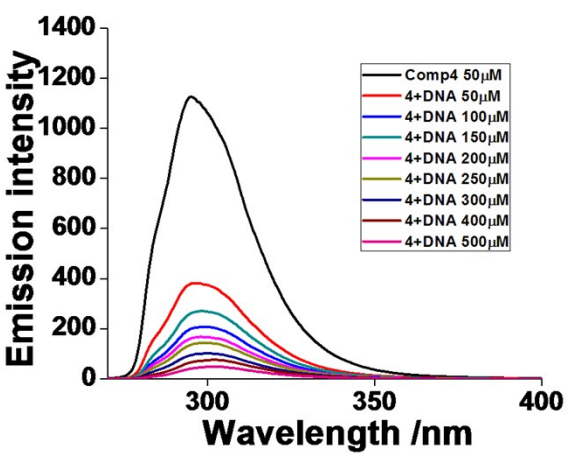**(A)**   **(B)**

Fig. 40S (A) Changes in emission of **1**(50µM) with increasing concentration of CT-DNA (0-500 µM, λex = 250 nm), (B) binding curve for **1** with CT-DNA.

Fig. 41S (A) Changes in emission of **3** (50μM) with increasing concentration of CT-DNA (0-500 μM, λex = 250 nm), (B) Binding curve for **3** with CT-DNA, (C) Changes in emission of **4** (50μM) with increasing concentration of CT-DNA (0-500 μM, λex = 250 nm), (D) Binding curve for **4** with CT-DNA.

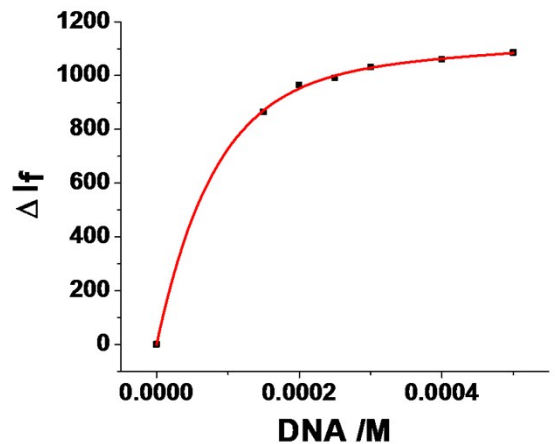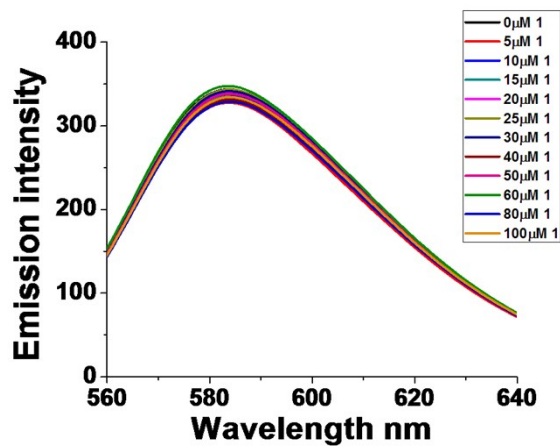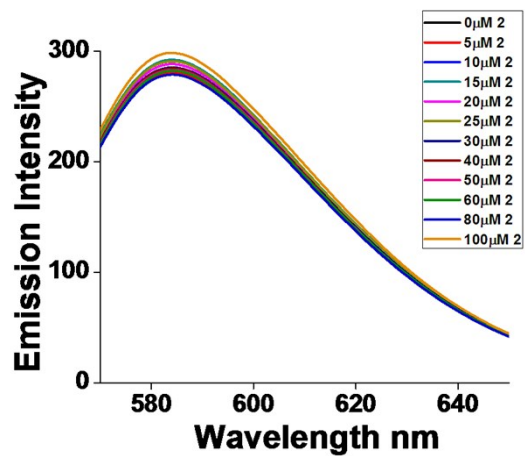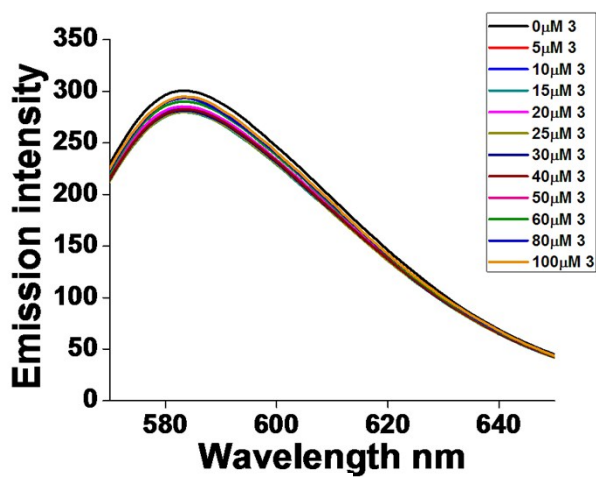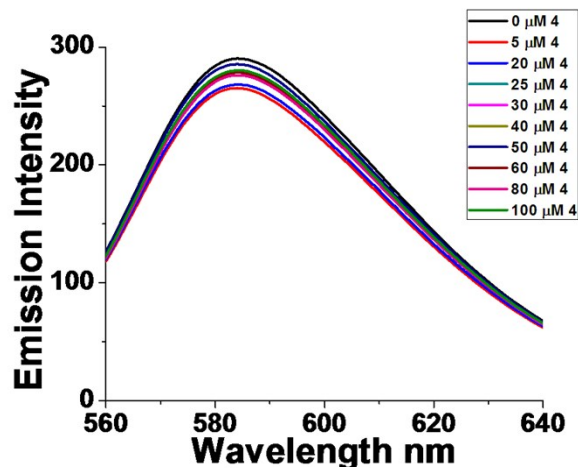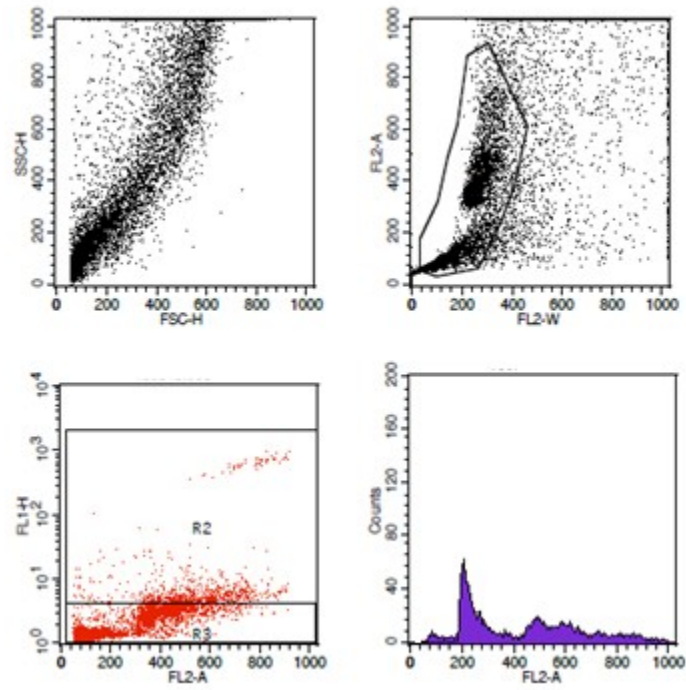**Fig. 42S** Effect of addition of (A) **1**(0-100μM), (B) **2**(0-100μM), (C) **3**(0-100μM), and (D) **4**(0-100μM) on the emission intensity of the EtBr (10 μM) bound CT-DNA (10 μM) at different concentrations.
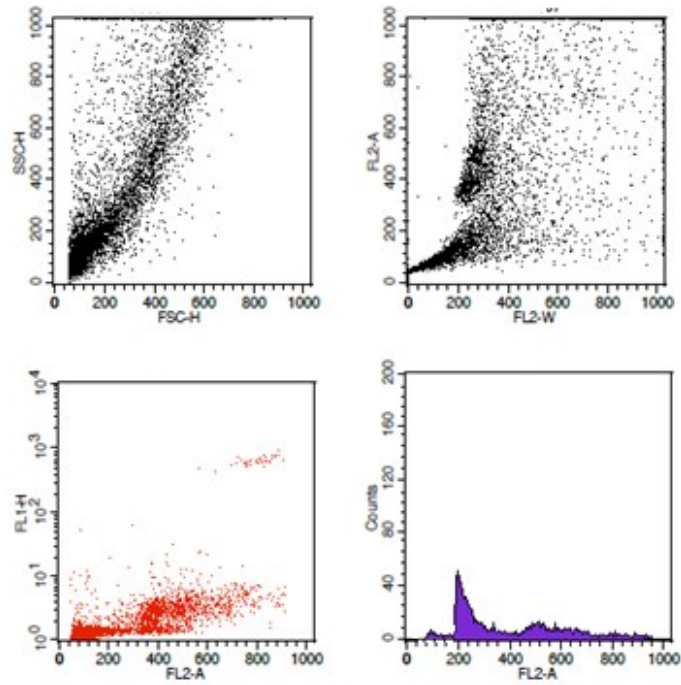
Table 1. Parameters for single crystal structure of compound **12** and **2**.

| Parameters	Compound 12	Compound 2
Empirical formula	$C_{16}H_{18}N_2O_3$	$C_{14}H_{13}ClN_2O_3$
Mol. Wt.	286.32	292.71
Wavelength (Å)	0.71073	0.71073
T	298(2)K	273 (2)
Crystal system	Orthorhombic	Orthorhombic
Space group	P 21 21 21	P 2 2ab
a(Å)	5.7178(3)	13.7554 (4)
b(Å)	10.4187(5)	18.6455 (5)
c(Å)	24.9522(11)	5.38910 (10)
α(deg)	90º	90º
β(deg)	90º	90º
γ(deg)	90º	90º
V(Å^3)	1506.26(6)	1382.16 (6)
Z	4	4
μ (mm^{-1})	0.077	0.285
F(000)	608	608
$D_{calcd.}$(mg/cm^3)	1.279	1.407
total reelections	9079	12954
$R_{int.}$	0.0285	0.0487
No. of parameters refined	208	206
Final R (I>2σ)	0.0580	0.0388
WR$_2$ (I>2σ)	0.1795	0.1111

Compound 1

Compound 2

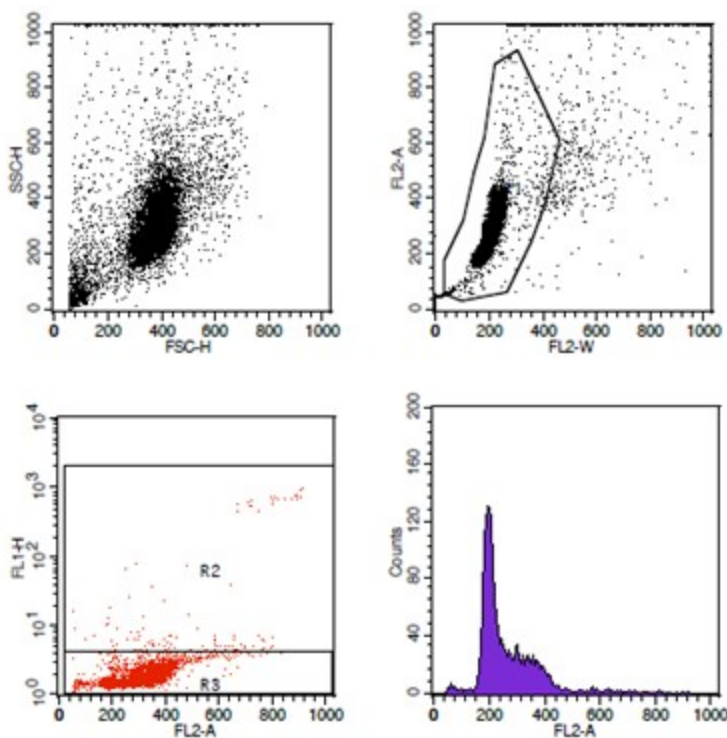**Compound 3**
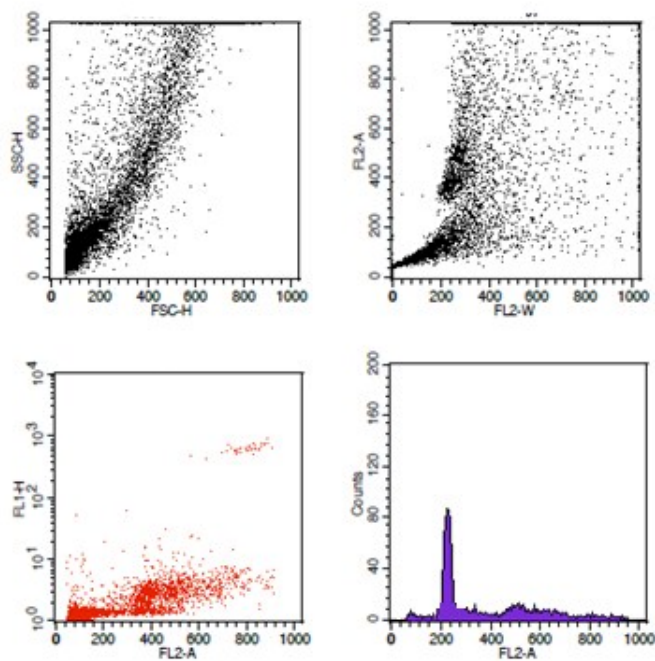
Compound 4
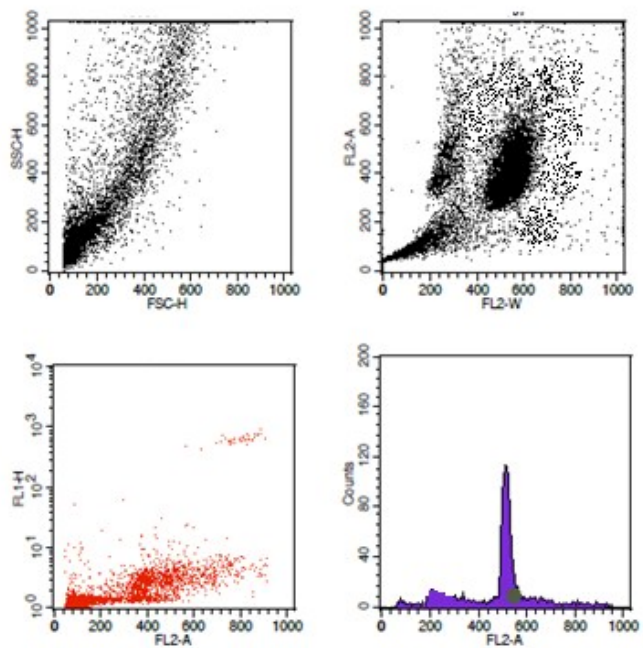
DOX

Fig: 43S: Cell cycle analysis of nucleosides **1-4.**

References

1   SMART & SAINT Software Reference Manuals, version 5.051 (Windows NT Version), Bruker Analytical X-ray Instruments Inc., Madison, Wi, 1998.

2   Sheldrick, G.M. (1996) SADABS, Program for Empirical Absorption Correction, University of Göttingen, Germany.

3   A. Altomare, M. C. Burla, M. Camalli, G. L. Cascarano, C. Giacovazzo, A. Guagliardi, A. G. G. Moliterni, G. Polidori and R. Spagna, *J. Appl. Crystallogra.*, 1999, **32**, 115-119.

4   G. M. Sheldrick, SHELXTLplus (Windows NT Version) Structure Determination

Package, Version 5.1. Bruker Analytical X-ray Instruments Inc.: Madison, WI, USA. 1998.

5   C. D. Nicola, F. Garau, A. Lanza, M. Monari, L. Pandolfo, C. Pettinari and A. Zorzi, *Inorg. Chim. Acta*., 2014, **416**, 186-194.

6   G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comp. Chem*., 2009, **30**, 2785-2791.

7   W. F. Van-Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott and I. G. Tironi, The GROMOS96 manual and user guide, 1996, 1–1042.

8   a) M. K. Singh, H. Pal, A. S. R. Koti and A. V. Sapre, *J. Phys. Chem. A.,* 2004, **108**, 1465-1474; b) N. Kandoth, S. D. Choudhury, J. Mohanty, A. C. Bhasikuttan and H. Pal, *J. Phys. Chem. B,* 2010, **114**, 2617-2626; c) N. Kandoth, S. D. Choudhury, T. Mukherjee and H. Pal, *Photochem. Photobiol. Sci*., 2009, **8**, 82-90; d) T. N. Burai, N. Bag, S. Agarwal, E. S. S. Iyer and A. Datta, *Chem. Phys. Lett*., 2010, **495**, 208-213; e) T. N. Burai, D. Panda and A. Datta, *Chem. Phys. Lett*., 2008, **455**, 42-46; f) J. S. K. Chen, M. Konopleva, M. Andreeff, A. Multani, S. Pathak and K. Mehta, *J. Cell. Physiol.,* 2004, **200**, 223-234.