Supplementary data

Dual-targeted anti-CMV/anti-HIV heterodimers

Anastasia Khandazhinskaya[1*], Vincenzo Mercurio[2], Anna Maslova[1], Rogers Alberto Ñahui Palomino[2], Mikhail Novikov[3], Elena Matyugina[1], Maria Paramonova[3], Marina Kukhanova[1], Natalya E. Fedorova[4], Kirill I. Yurlov[4], Alla A. Kushch[4], Olga Tarasova[5], Leonid Margolis[2], Sergey N. Kochetkov[1], Christophe Vanpouille[2]*

[1]*Engelhardt Institute of Molecular Biology, Vavilova Str., 32, Moscow, 119991, Russia;*

[2]*Eunice Kennedy-Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892, USA;*

[3]*Department of Pharmaceutical & Toxicological Chemistry, Volgograd State Medical University, Paversushikh Bortsov Sq., 1, Volgograd 400131, Russia;*

[4]*Ivanovsky Institute of Virology, Gamaleya National Research Center of Epidemiology and Microbiology, Gamaleya Str., 16, Moscow, 123098, Russia;*

[5]*Institute of Biomedical Chemistry, Pogodinskaya Str., 10, Moscow, Russia, 119121, Russia.*

*To whom correspondence should be addressed:

Anastasia Khandazhinskaya, e-mail: khandazhinskaya@bk.ru, tel.+79104401271

Christophe Vanpouille, e-mail: vanpouic@mail.nih.gov, tel.301-594-0826

Contents

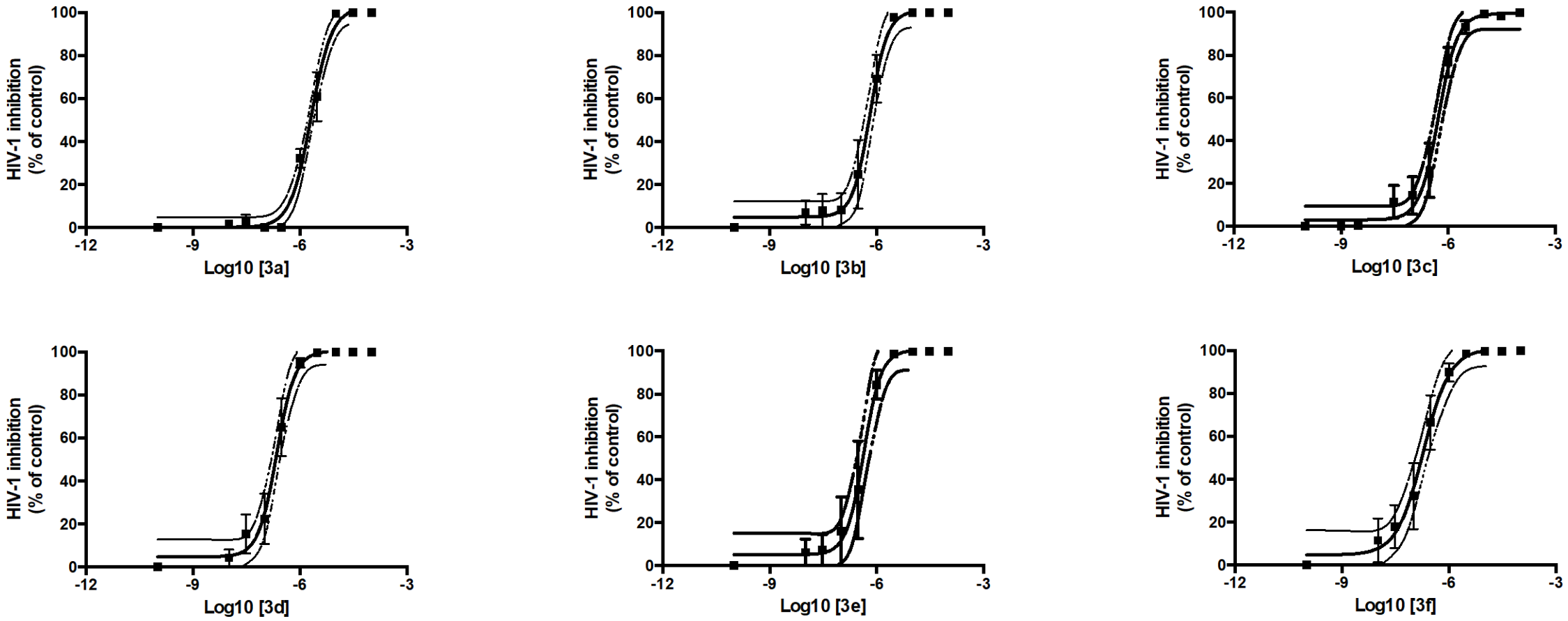


Figure S1. Anti-HIV activity of heterodimers in MT-4 cell cultures

10^6 MT-4 cells were inoculated with 100 µL of HIV-1 viral stock of $X4_{LAI.04}$ at a p24gag concentration of 50 ng/mL for 1 h 30 min at 37°C. After incubation, excess of HIV was washed off with PBS and 10^5 infected cells were transferred to 24-well plates, mixed with 1 mL of medium containing each of the 6 test compounds at concentrations ranging from 100 pM to 100 µM (in triplicate). After 3 days of culture at 37°C, the anti-HIV activity was evaluated from the suppression of viral replication compared with untreated MT-4 cell cultures. EC_{50}s or effective concentrations that inhibited HIV-1 replication by 50% and the 95% confidence intervals were estimated by fitting the data points to a sigmoidal dose-response curve, using Prism software (Table 1).

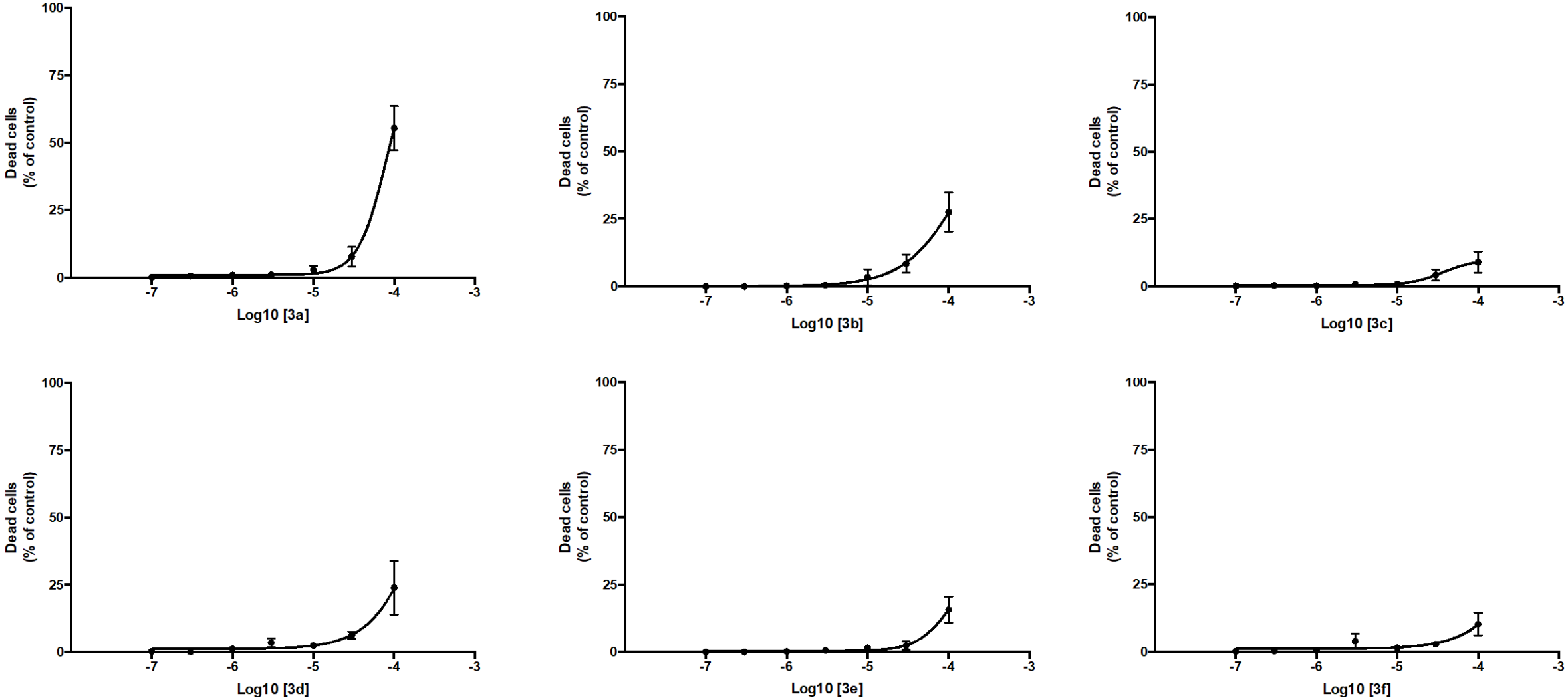
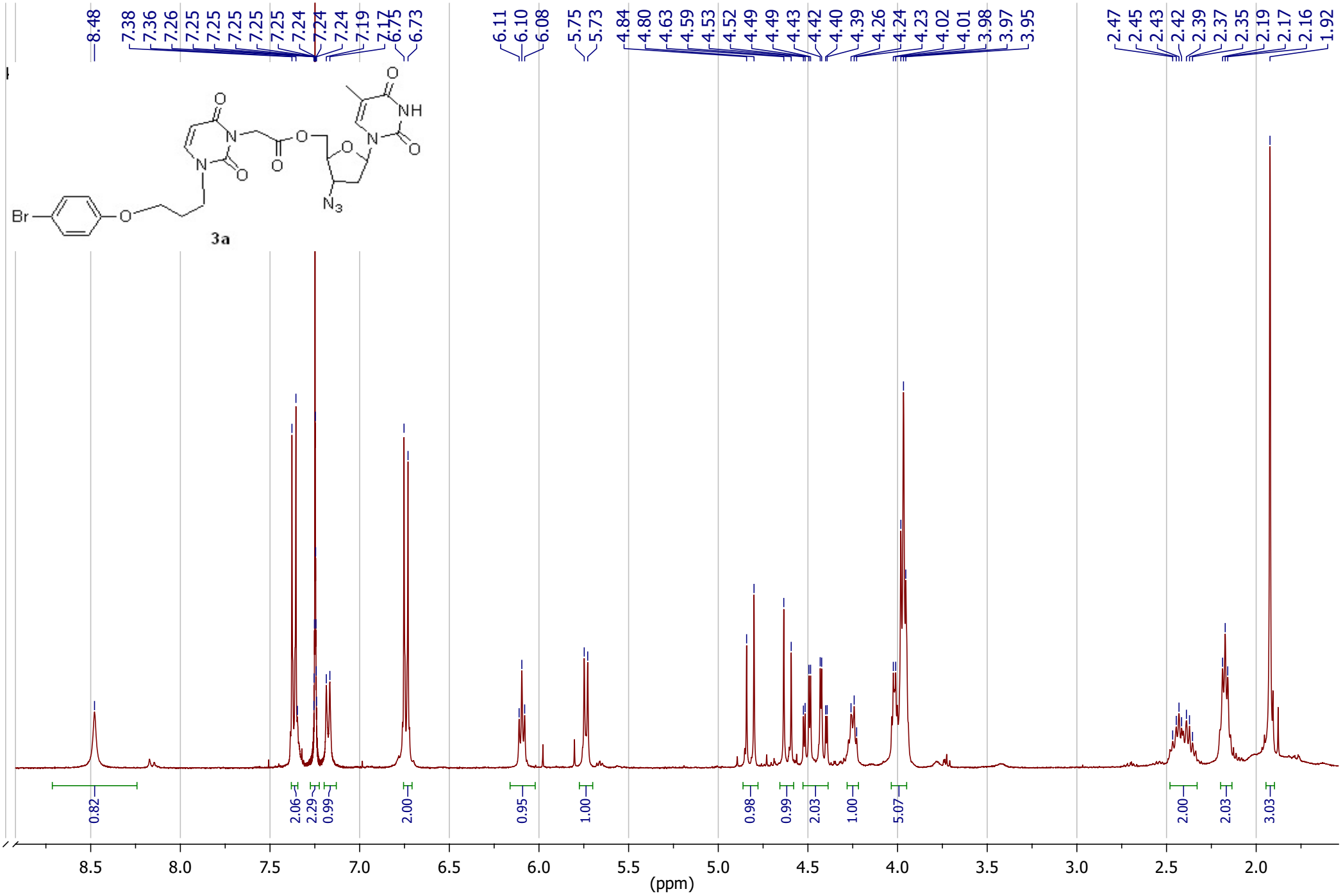
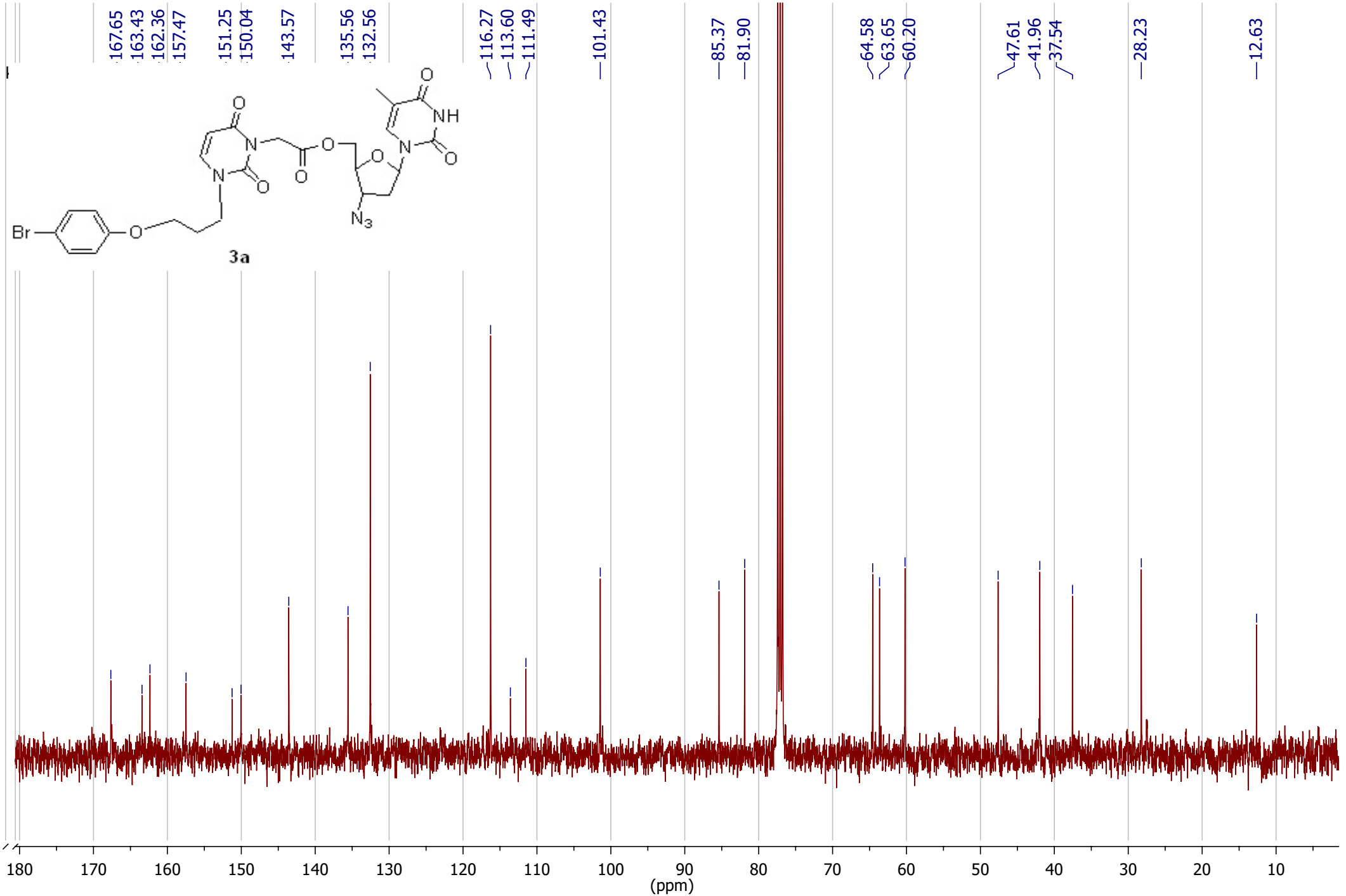


Figure S2. Low cytotoxicity of heterodimers in MT-4 cell cultures

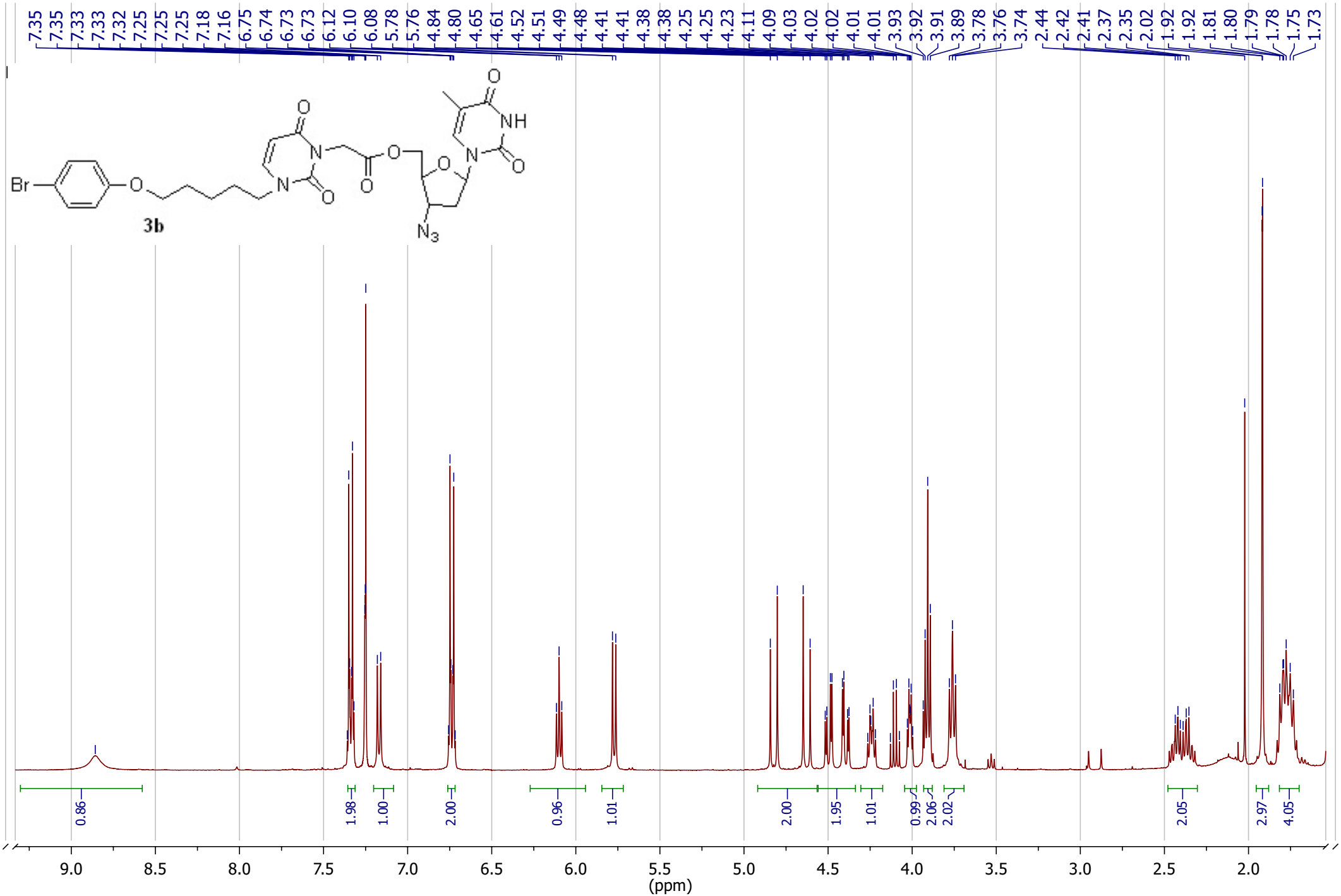
10^5 cells MT-4 cells were incubated at 37°C with increasing concentrations of heterodimers and cultured for 3 days. Cells were counted using a Nucleocounter® and cell death (expressed as percent of control) was estimated from five independent experiments, each of them testing seven different concentrations (ranging from 100 nM to 100 µM).



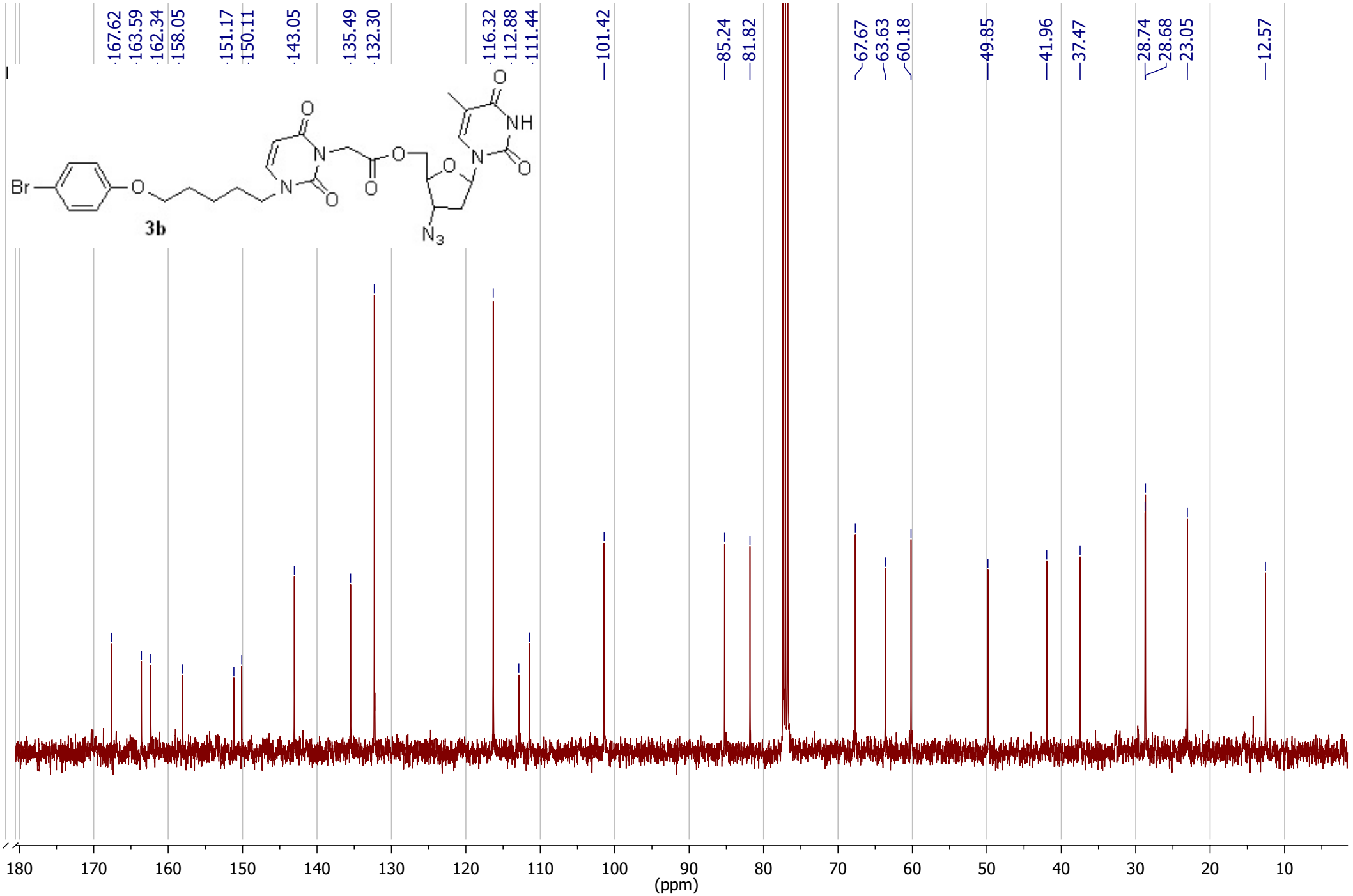
S5 ^{1}H NMR spectrum of **3a** in $CDCl_3$



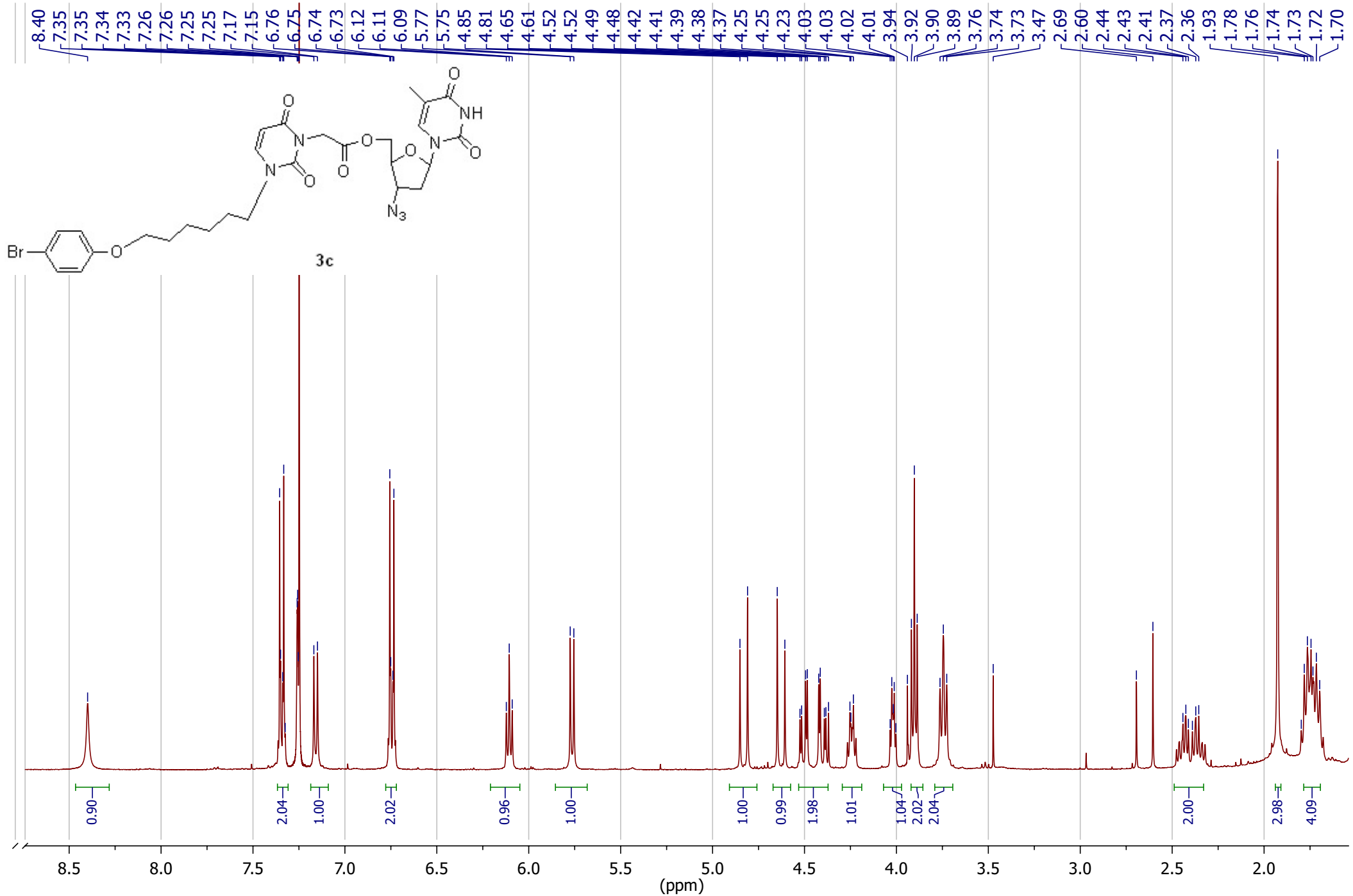
S6 ^{13}C NMR spectrum of **3a** in $CDCl_3$



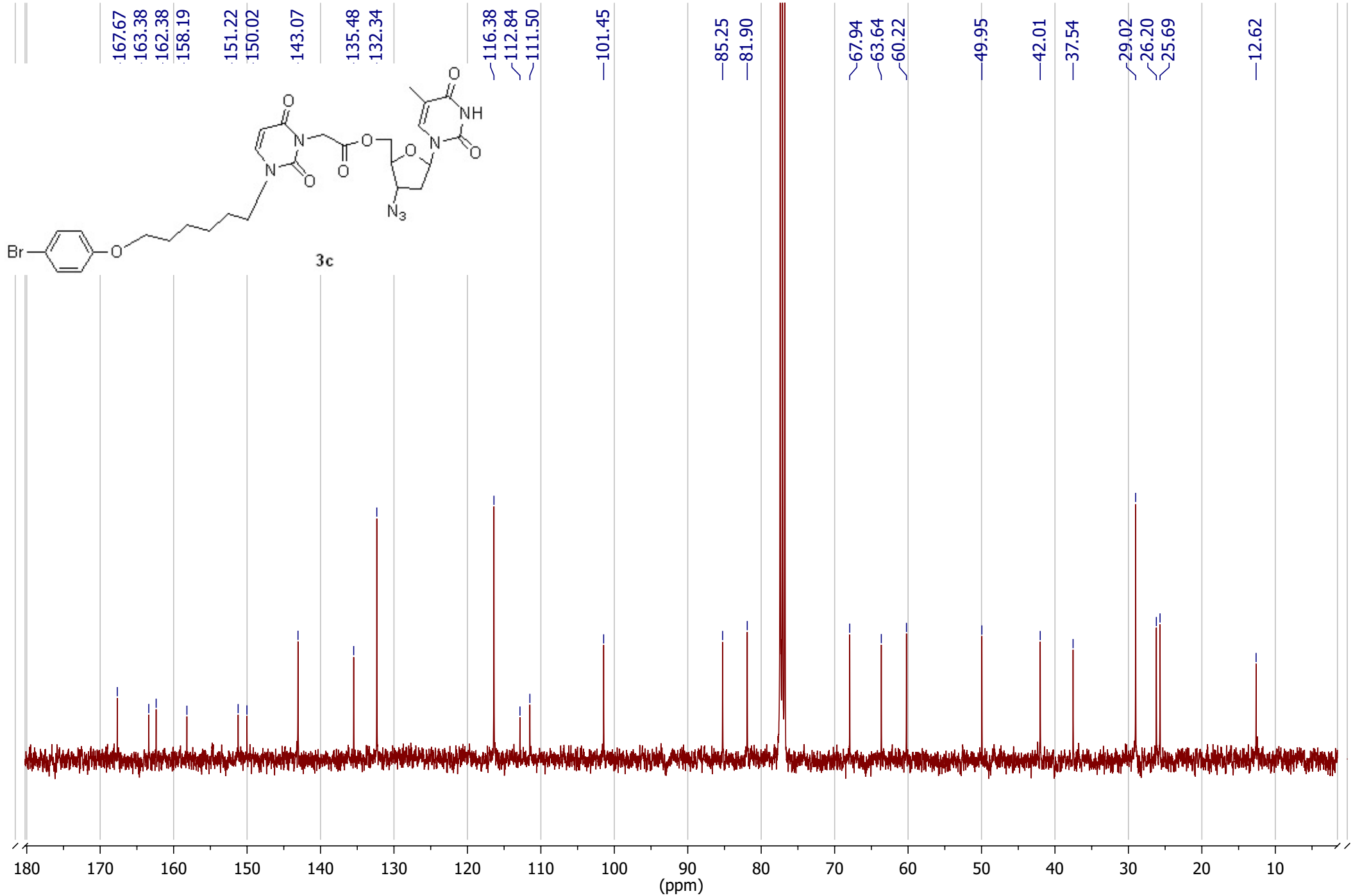
S7 ^{1}H NMR spectrum of **3b** in $CDCl_3$



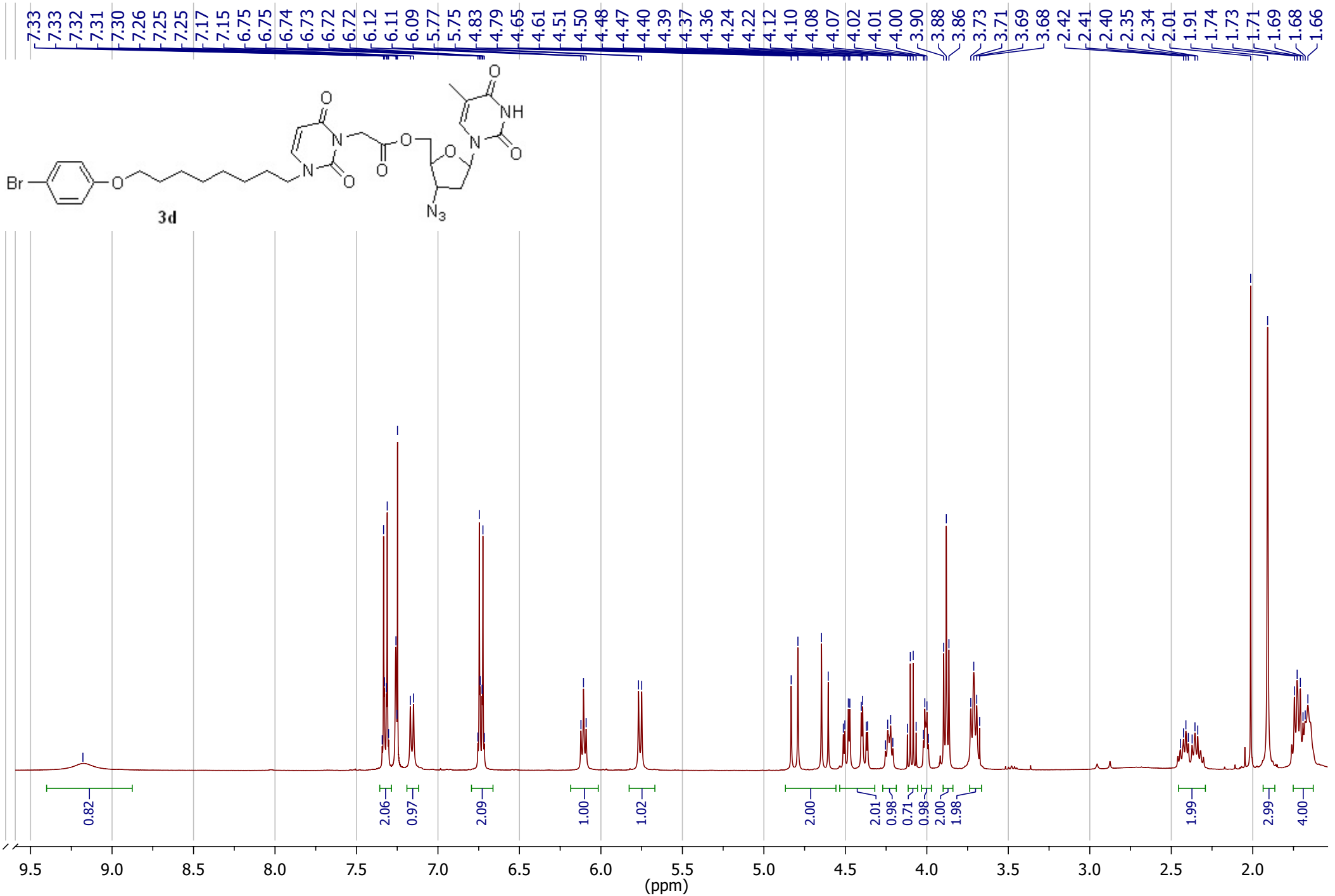
S8 ^{13}C NMR spectrum of **3b** in $CDCl_3$



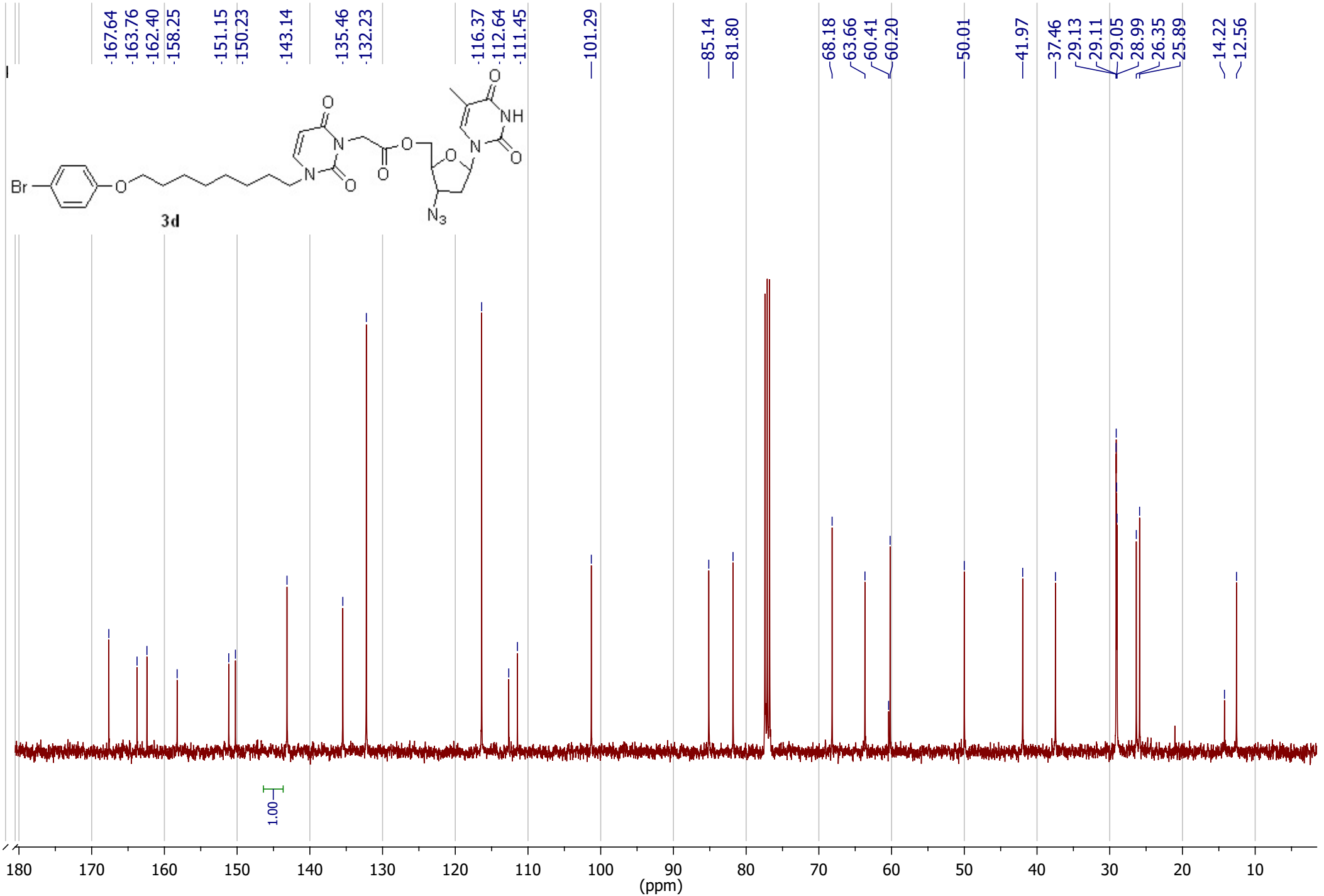
S9 ^{1}H NMR spectrum of **3c** in $CDCl_3$



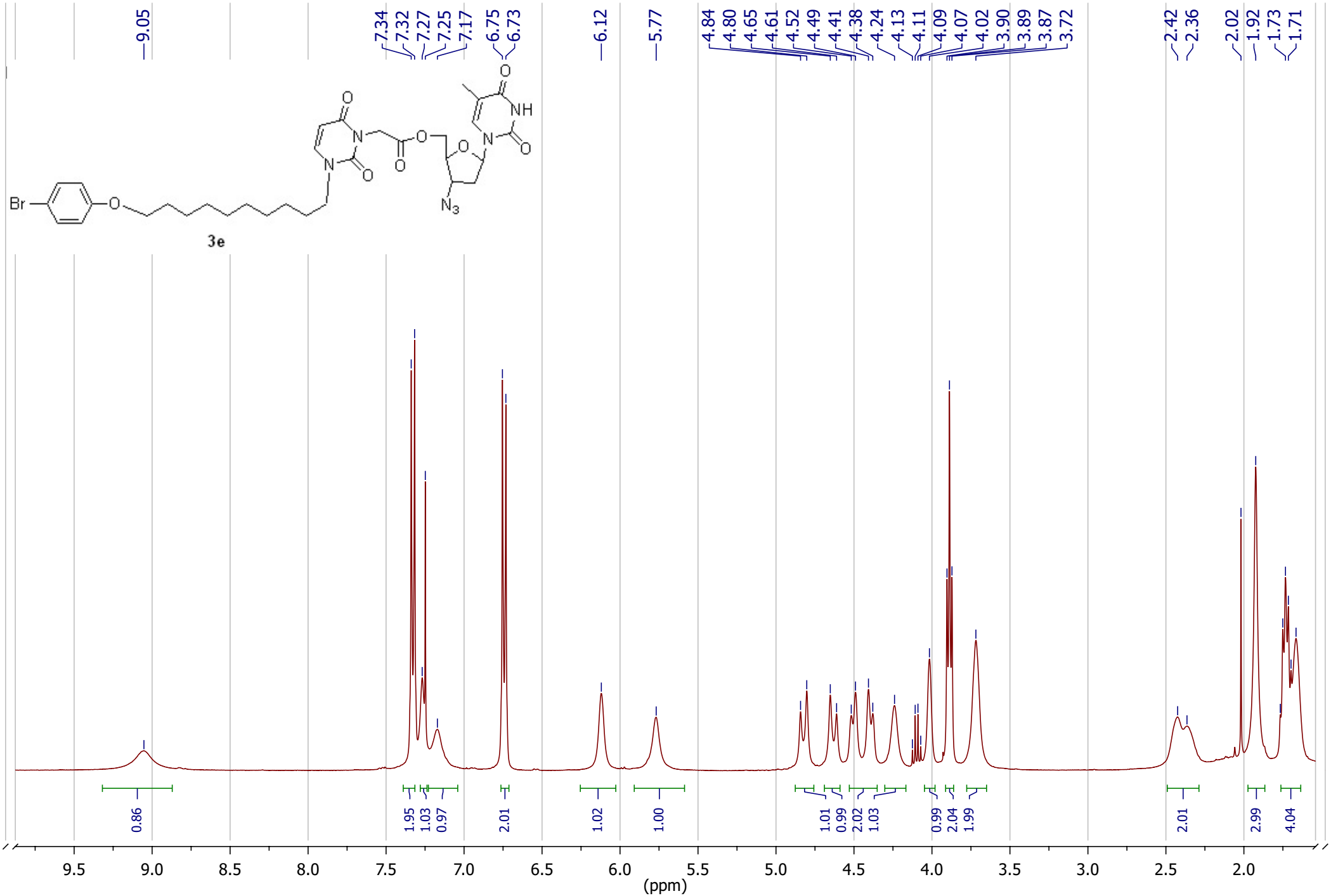
S10 ^{13}C NMR spectrum of **3c** in $CDCl_3$



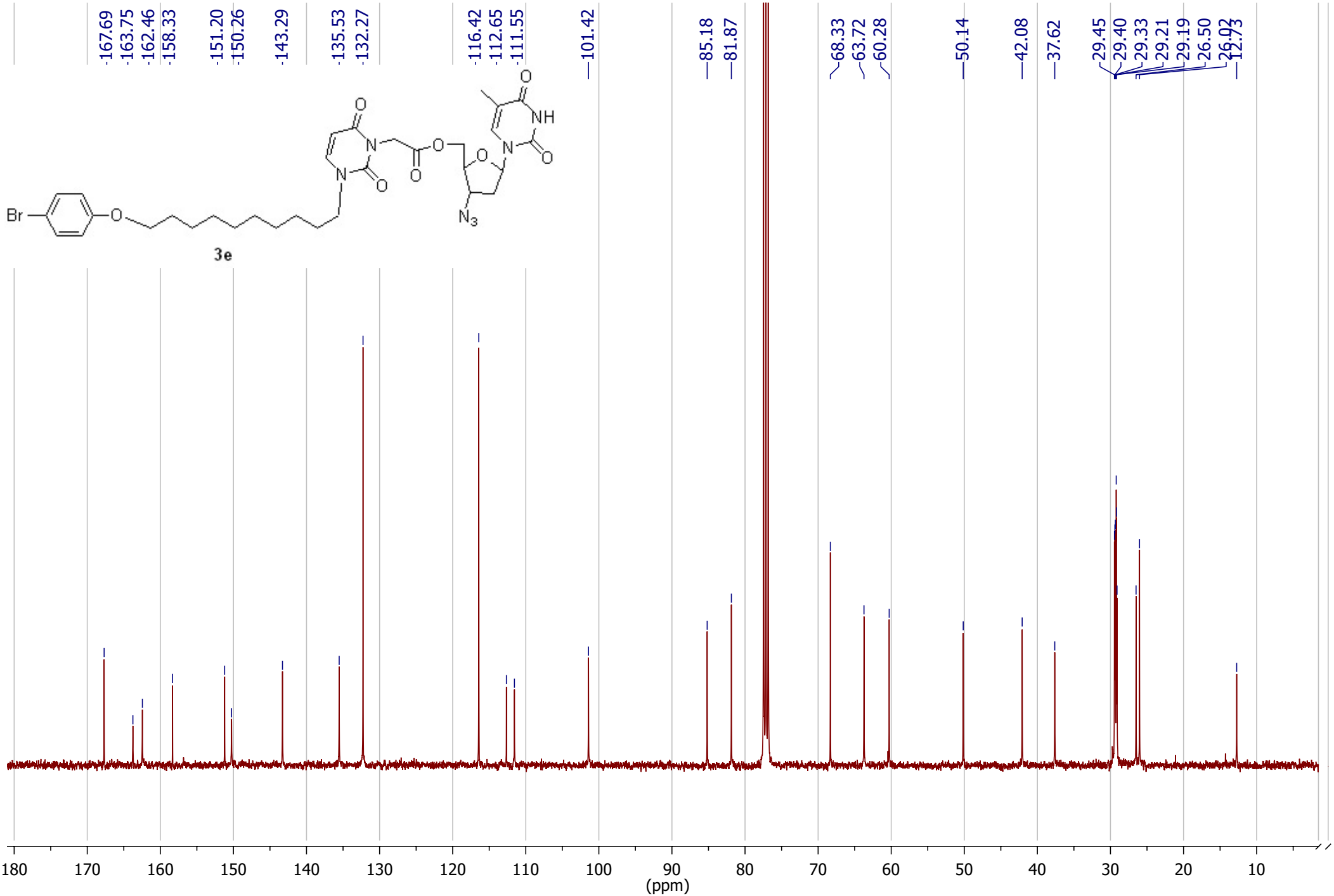
S11 ^{1}H NMR spectrum of **3d** in $CDCl_3$



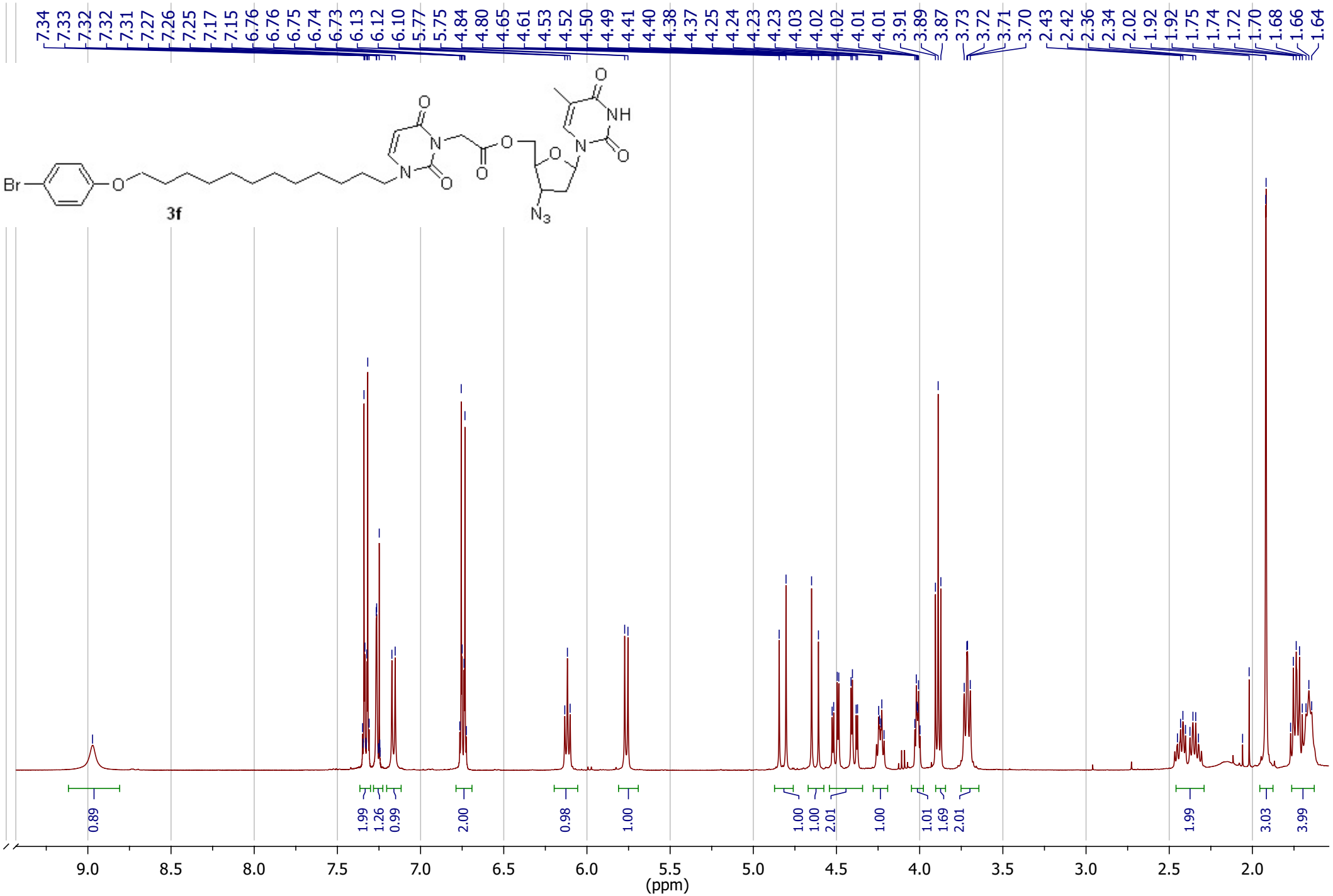
S12 ^{13}C NMR spectrum of **3d** in $CDCl_3$



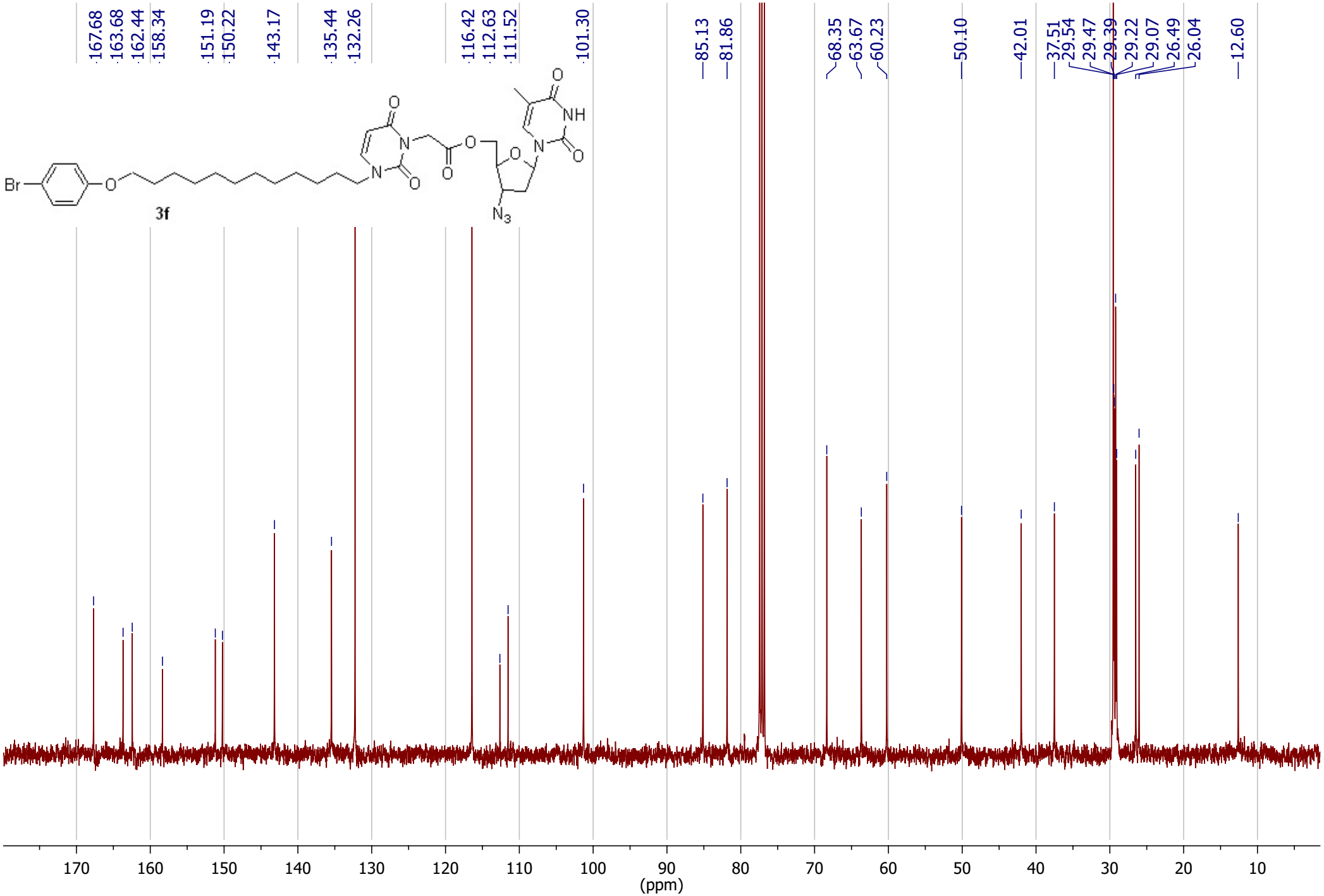
S13 ^{1}H NMR spectrum of **3e** in $CDCl_3$



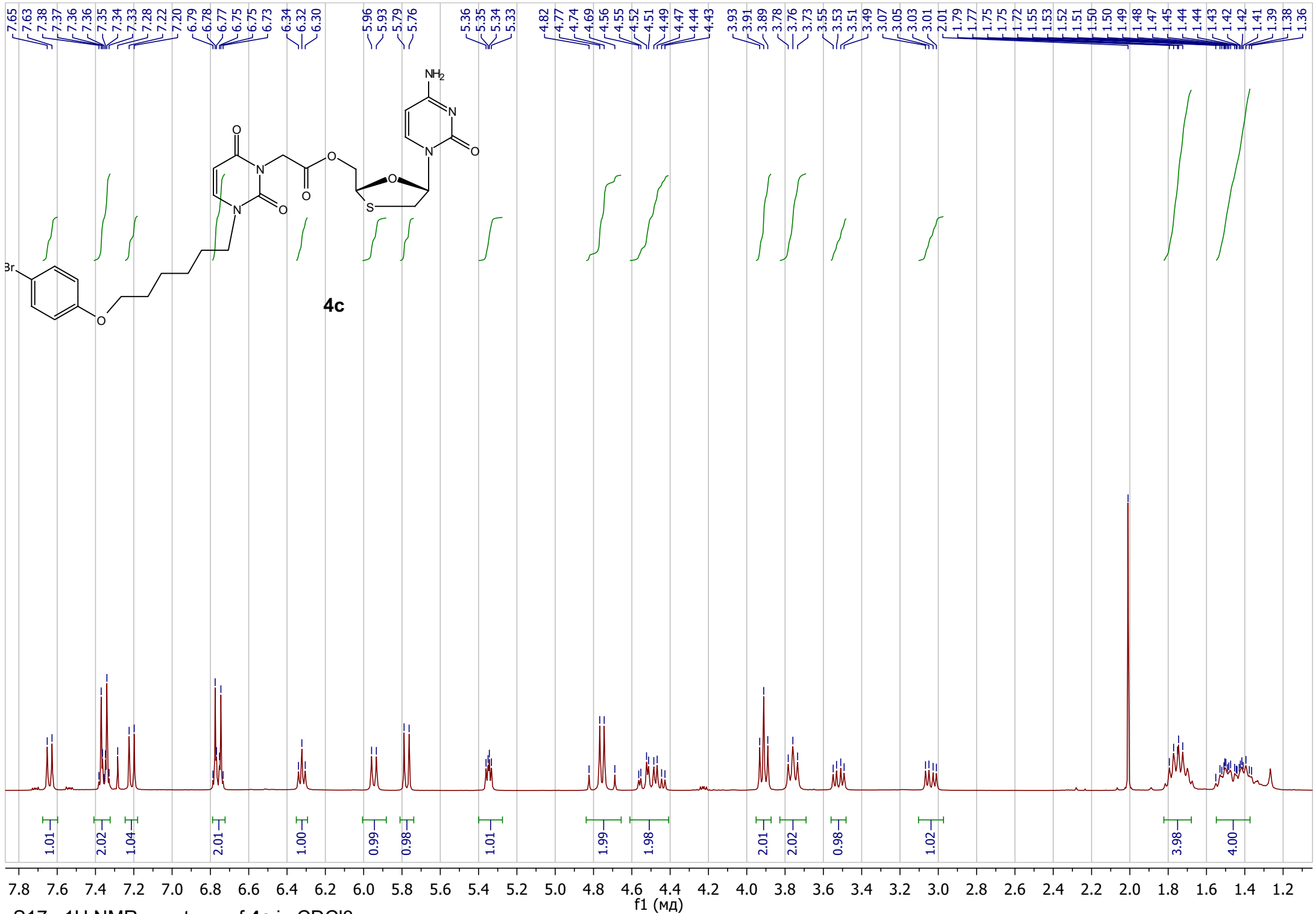
S14 ^{13}C NMR spectrum of **3e** in $CDCl_3$



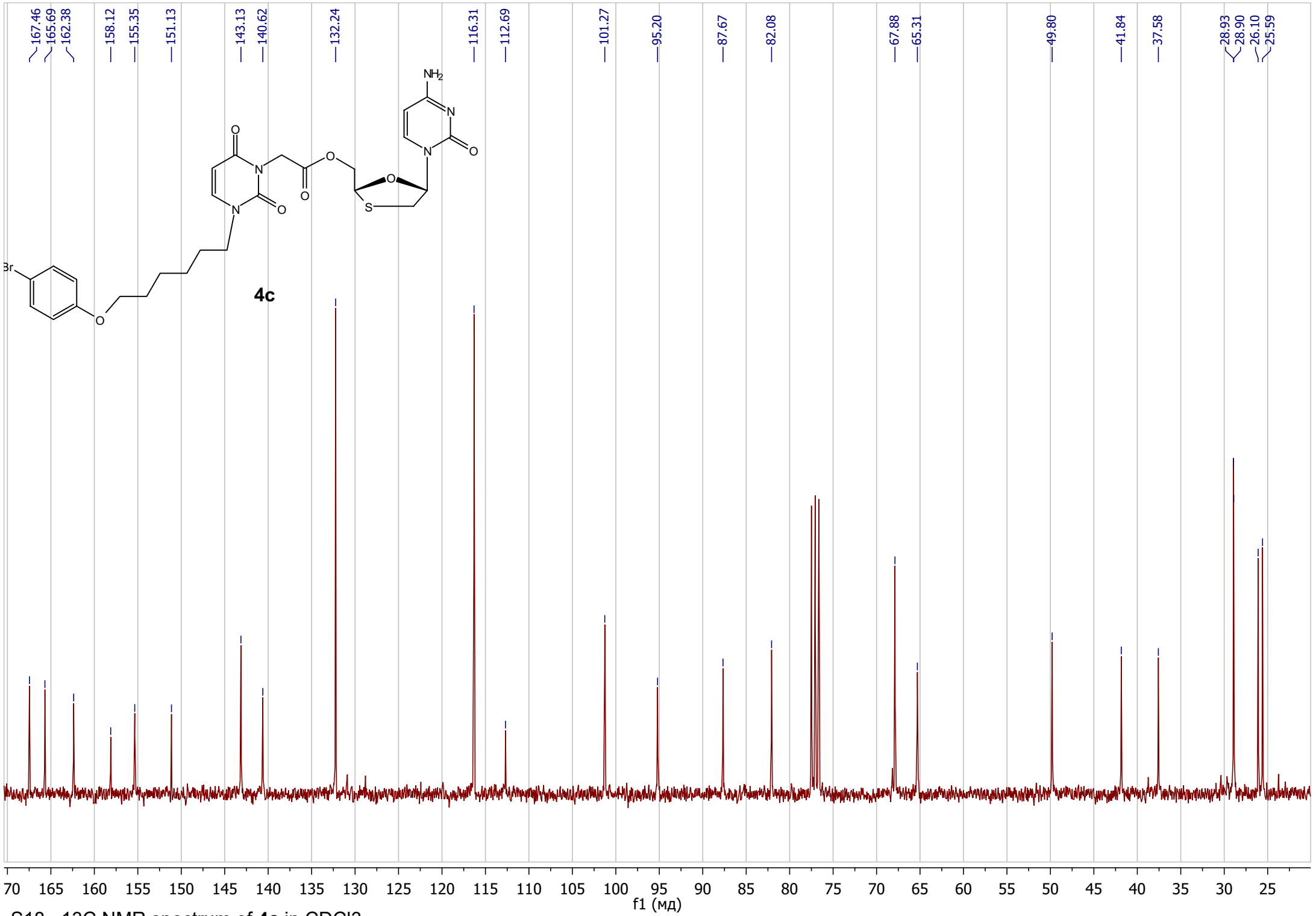
S15 ^{1}H NMR spectrum of **3f** in $CDCl_3$



S16 ^{13}C NMR spectrum of **3f** in $CDCl_3$



S17 1H NMR spectrum of **4c** in CDCl3



S18 13C NMR spectrum of **4c** in CDCl3

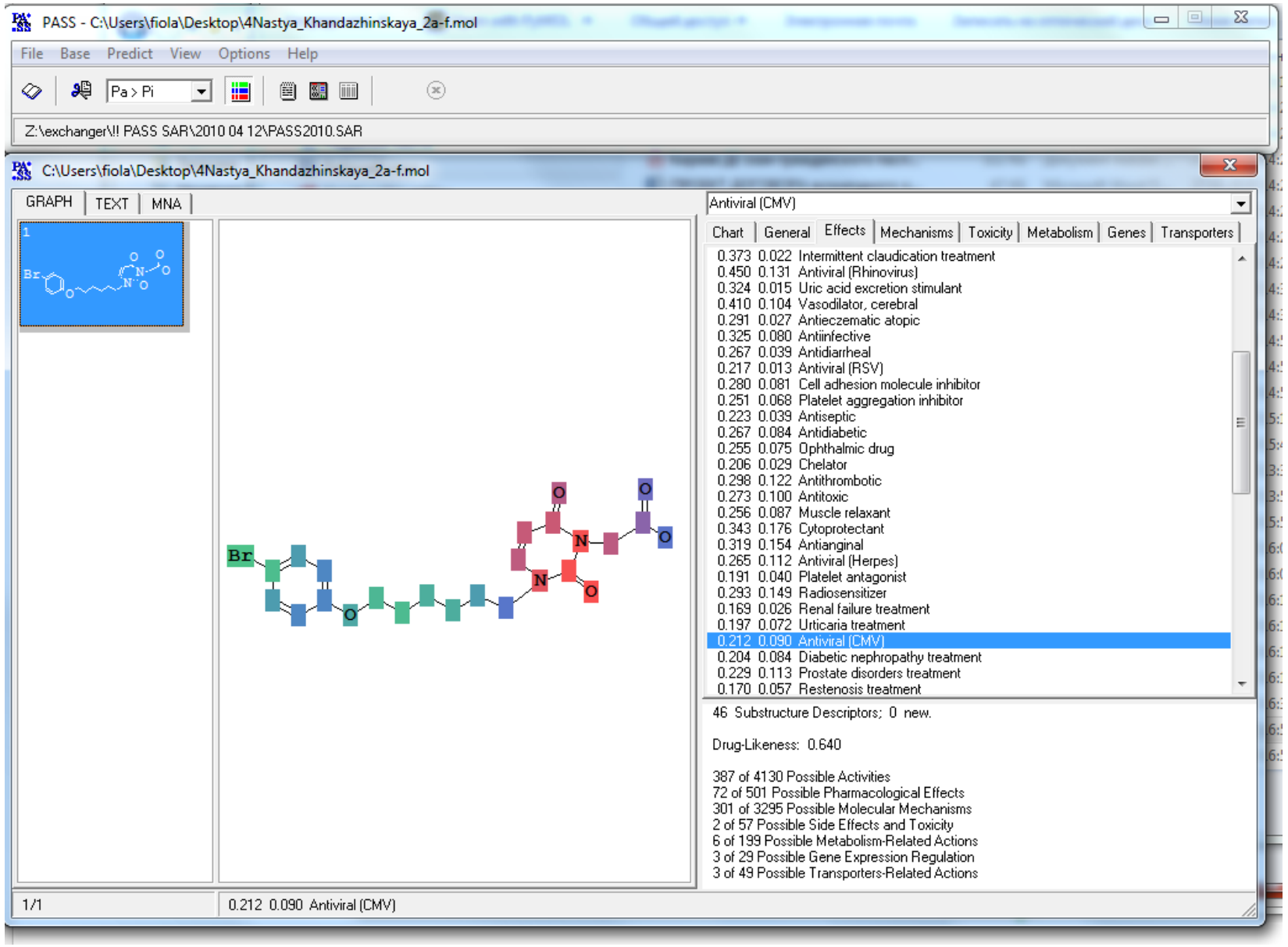
1. Description-of-PASS-prediction-for-antiviral-compounds

The coordinates of structures **2c** were saved in SDF format. PASS software was used to predict the spectra of biological activity for this structure. PASS uses the naïve Bayes approach [29755970]. A chemical structure is represented as a set of the so-called multilevel neighborhoods of atoms (MNA descriptors). MNA descriptors allow taking into account each atom of the molecule along with its nearest neighbors. In the last version of PASS, MNA-descriptors of the second level (MNA-2) are used. Estimated values Pa and Pi are calculated based on the frequencies of MNA-descriptors in the sets of active and inactive compounds of the training set. Pa and Pi reflect the probabilities for the chemical compound to be either active (Pa) or inactive (Pi) in certain biological activities. A compound can be active in case Pa exceeds Pi. During training process MNA descriptors along with the data on compounds, they are associated with and experiments activity spectra for these compounds are stored in SAR (structure-activity relationship) knowledge base.

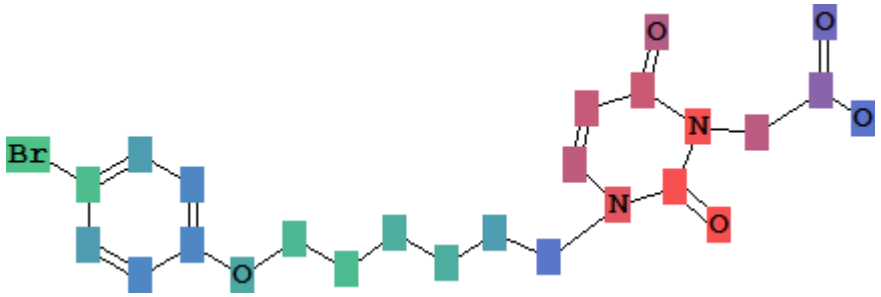
To predict antiviral activity of the compound **2c** and AZT we used PASS (v.2010) specially developed for predicting antiviral properties. SAR base included 266697 compounds with information about 5825 activities. There are data on over 10000 compounds with overall antiviral properties, over 6600 compounds active against human immunodeficiency virus HIV. Also this specifically compiled SAR base have information about chemical compounds which selectively inhibit herpesviruses, including cytomegalovirus (CMV). It also includes information about compounds active against respiratory syncytial virus (RSV), adenovirus and some others.

According to the results of PASS prediction, the compound **2c** must have antiviral activity. In particular, it should be active against herpesviruses, including specific activity against CMV. It can also be active against RSV, adenoviruses, rhinovirus. In other words, it may have a wide spectrum of antiviral activity. According to the PASS results, AZT must have antiviral activity against HIV, which corresponds to the experimental results. The full predicted spectra on biological activity is provided in *.CSV format along with SDF file containing 2D coordinates of the structure.

PASS for compound 2c



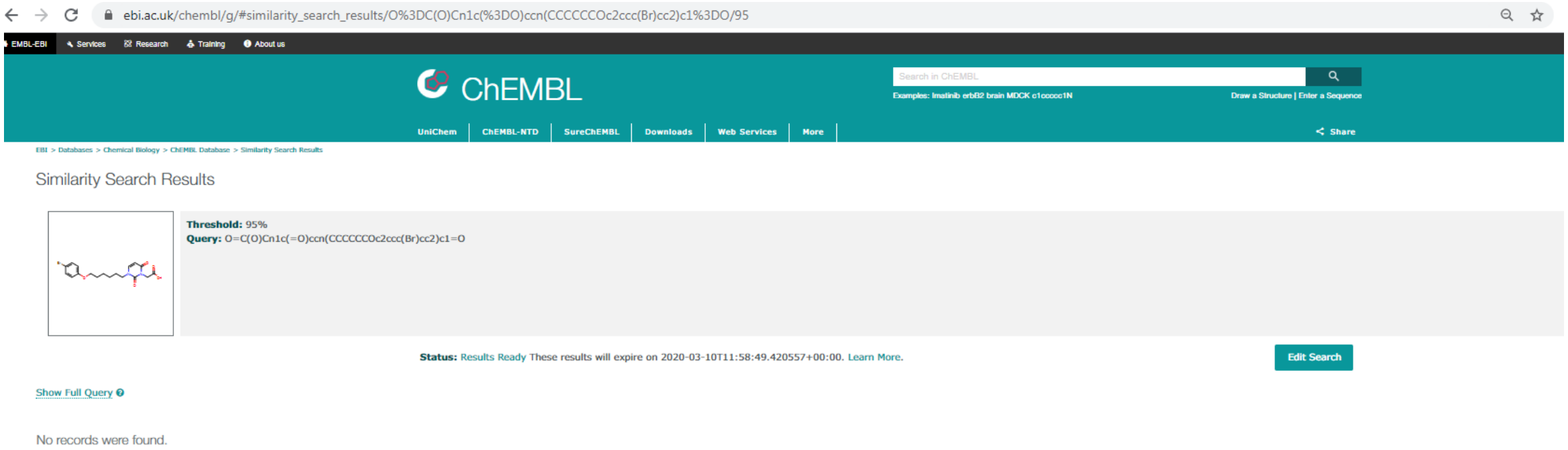
Prediction of the toxicity for compound 2c by PASS



2 of 57 Possible Side Effects and Toxicity at Pa > Pi

0.341 0.183 Cardiotoxic
0.324 0.280 Emetic

Search results for compound 2c in ChEMBL



Search results with similarity cut-off threshold> 95% - no similar structures found

Search results for compound 2c in ChEMBL

Molecule ChEMBL ID	Compound Key	Standard Type	Standard Relation	Standard Value	Standard Units	pChEMBL Value	Comment	Assay ChEMBL ID	Assay Description	BAO Label	Assay Organism	Target ChEMBL ID	Target Name	Target Organism	Target Type	Document ChEMBL ID	Source Description	Cell ChEMBL ID
 CHEMBL3623864	7k	MCC	=	4000	nM	No Data	No Data	CHEMBL3626040	Cytotoxicity against human cytomegalovirus infected HEL cells assessed as alteration of cell morphology incubated for 3 days by microscopy	cell-based format	Homo sapiens	CHEMBL613888	HEL	Homo sapiens	CELL-LINE	CHEMBL3621200	Scientific Literature	CHEMBL3308090
CHEMBL3623860	7g	Inhibition	No Data	No Data	No Data	No Data	Not Active	CHEMBL3626155	Antiviral activity against Human coxsackievirus B4	organism-based format	Human coxsackievirus B4	CHEMBL612320	Human coxsackievirus B4	Human coxsackievirus B4	ORGANISM	CHEMBL3621200	Scientific Literature	No Data
CHEMBL3623865	7l	Inhibition	No Data	No Data	No Data	No Data	Not Active	CHEMBL3626150	Antiviral activity against HSV1 KOS	organism-based format	Human herpesvirus 1 strain KOS	CHEMBL2367034	Human herpesvirus 1 strain KOS	Human herpesvirus 1 strain KOS	ORGANISM	CHEMBL3621200	Scientific Literature	No Data
CHEMBL3623857	7d	CC50	=	4900	nM	No Data	No Data	CHEMBL3626041	Cytotoxicity against human cytomegalovirus infected HEL cells assessed as cell growth inhibition incubated for 3 days by coulter counter method	cell-based format	Homo sapiens	CHEMBL613888	HEL	Homo sapiens	CELL-LINE	CHEMBL3621200	Scientific Literature	CHEMBL3308090
 CHEMBL4074053	33	Ratio EC50	No Data	No Data	No Data	No Data	Not Determined	CHEMBL4036856	Ratio of EC50 for agonist activity at CB1 receptor (unknown origin) expressed in CHO cells to EC50 for agonist activity at CB2 receptor (unknown origin) expressed in CHO cells	cell-based format	Homo sapiens	CHEMBL612545	Unchecked	No Data	UNCHECKED	CHEMBL4033762	Scientific Literature	No Data

Search results with similarity cut-off> 80% - examples of biological activity