Supplementary Information

A Novel DNA Sequence-Selective, Guanine Mono-Alkylating ADC Payload Suitable for Solid Tumour Treatment

Paolo Andriollo[1], Daniella di Mascio[2], Paul J. M. Jackson[1], Md. Mahbub Hasan[1], Ilona Pysz-Hosey[1], George Procopiou[1], Keith R. Fox[2], Khondaker Miraz Rahman[1]* and David E. Thurston[1]

[1]School of Cancer & Pharmaceutical Sciences, King's College London, WC2R 2LS, United Kingdom

[2]School of Biological Sciences, University of Southampton, SO17 1BJ, United Kingdom.

*To whom correspondence should be addressed

KRM: e-mail k.miraz.rahman@kcl.ac.uk

Contents

Purity determination of synthesised final compounds

The level of purity of the compounds for biological testing has been evaluated through LC-MS analysis, using three different gradient methods, reported hereafter. Liquid Chromatography Mass Spectroscopy (LCMS) (**Methods A & B**) analysis was performed on a Waters Alliance 2695 with water (A) and acetonitrile (B) comprising the mobile phases. Formic acid (0.1%) was added to both acetonitrile and water to ensure acidic conditions throughout the analysis. Function type: Diode array (535 scans). Column type: Monolithic C18 50 X 4.60 mm. Mass spectrometry data were collected using a Waters Micromass ZQ instrument coupled to a Waters 2695 HPLC with a Waters 2996 PDA. Waters Micromass ZQ parameters used were: Capillary (kV), 3.38; Cone (V), 35; Extractor (V), 3.0; Source temperature (°C), 100; De-solvation Temperature (°C), 200; Cone flow rate (L/h), 50; De-solvation flow rate (L/h), 250. LCMS gradient conditions are described below. Ultra-Performance Liquid Chromatography Mass Spectroscopy (UPLC-MS) (**Methods C**) analysis was performed on a Waters Acquity H-class UPLC with water (A) and acetonitrile (B) comprising the mobile phases. Trifluoracetic acid (0.1%) was added to both acetonitrile and water to ensure acidic conditions throughout the analysis. Function type: Photo Diode array (502.93 n). Column type: Acquity UPLC BEH C18 1.7µm 2.1 X 50 mm. Mass spectrometry data were collected using a Waters SQ Detector 2 coupled to a Waters Acquity H Class UPLC with ACQ-PDA. Waters SQ Detector 2 parameters used were: Capillary (kV), 3.00; Cone (V), 30; De-solvation Temperature (°C), 600; Cone flow rate (L/h), 50; De-solvation flow rate (L/h), 600. UPLC-MS gradient conditions are described below. Sample preparation: samples were dissolved in methanol at 1-10 μg/mL, then filtered through a 0.22 μm filter membrane. Injection volume: 1-10 μL. Gradient conditions are described below.

Method A (10 min): from 95% A/5% B to 50% B over 3 min. Then from 50% B to 80% B over 2 min. Then from 80% B to 95% B over 1.5 min and held constant for 1.5 min. This was then reduced to 5% B over 0.2 min and maintained to 5% B for 1.8 min. The flow rate was 0.5 mL/min, 200 µL was split via a zero dead volume T piece which passed into the mass spectrometer. The wavelength range of the UV detector was 220-400 nm.

Method B (5 min): from 95% A/5% B to 90% B over 3 min. Then from 90% B to 95% B over 0.5 min and held constant for 1 min. This was then reduced to 5% B over 0.5 min. The flow rate was 1.0 mL/min, 100 µL was split via a zero dead volume T piece which passed into the mass spectrometer. The wavelength range of the UV detector was 220-500 nm.

Method C (7 min): from 90% A/10% B to 50% B over 1.5 min. Then from 50% B to 75% B over 1.5 min. Then from 75% B to 90% B over 1 min. This was then reduced to 10% B over 1 min. The flow rate was 0.6 mL/min, 5 µL was split via a zero-dead volume T piece which passed into the mass spectrometer. The wavelength range of the UV detector was 230-280 nm.

Final compounds purity

	Purity %		
Compound	LC-MS Method A	LC-MS Method B	UPLC-MS Method C
18	≥ 95	≥ 95	≥ 95
19	≥ 95	≥ 95	≥ 95
20	≥ 95	≥ 95	≥ 95

Table S1. Purity of **18,19** and **20.**

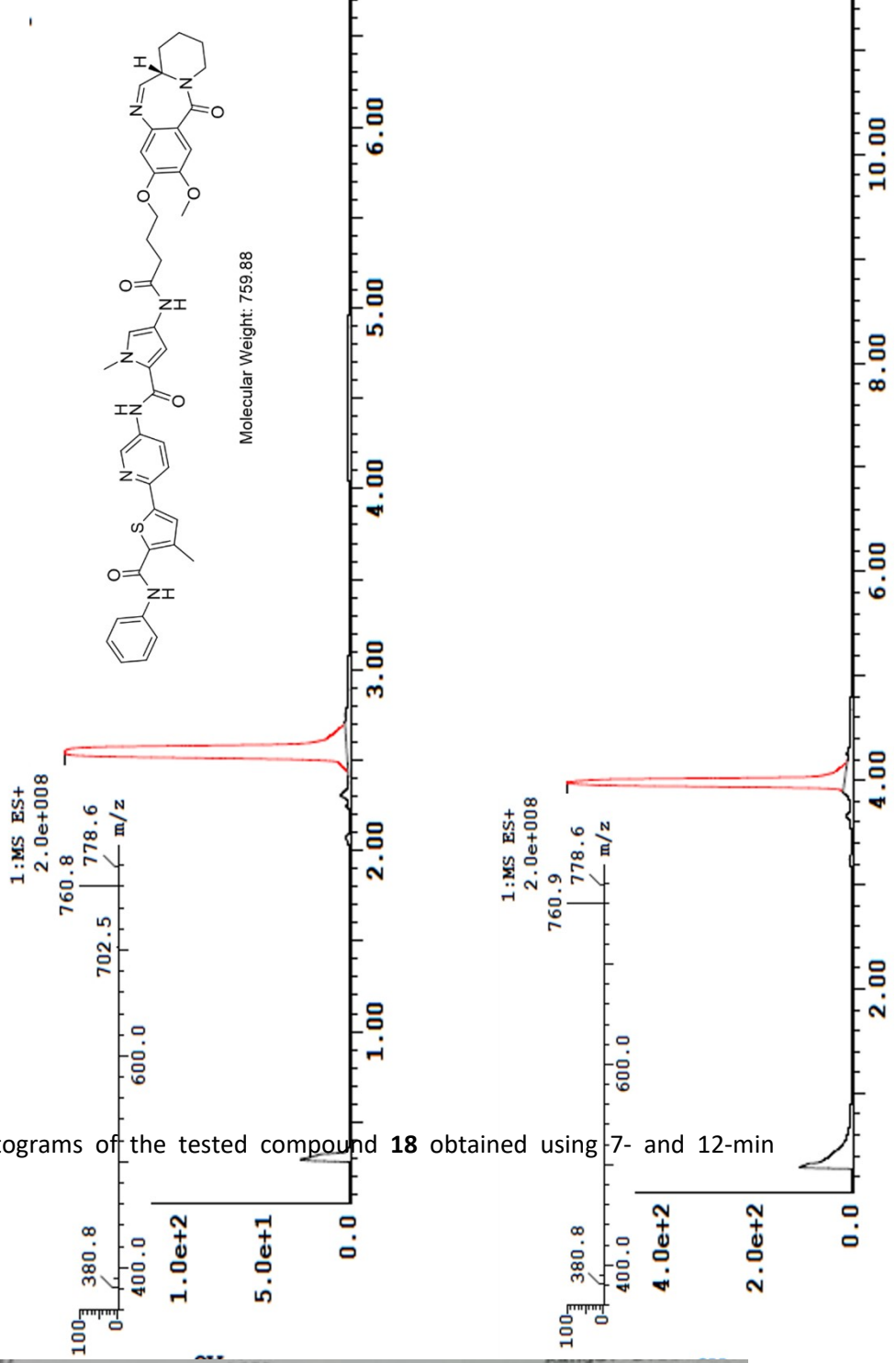
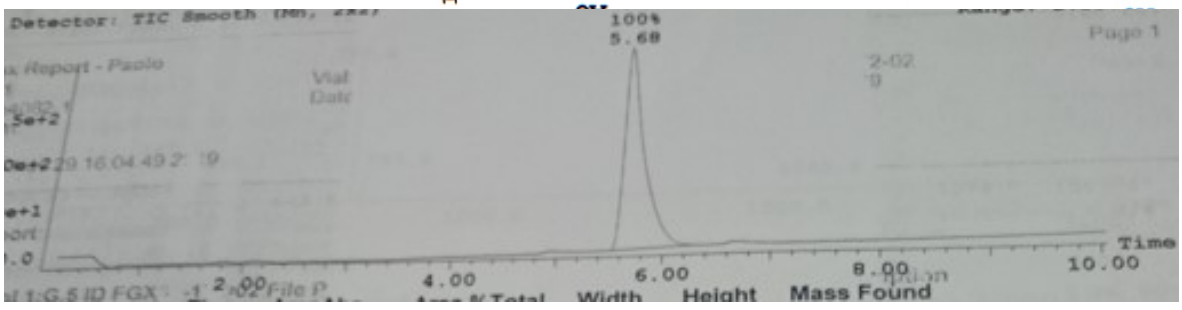
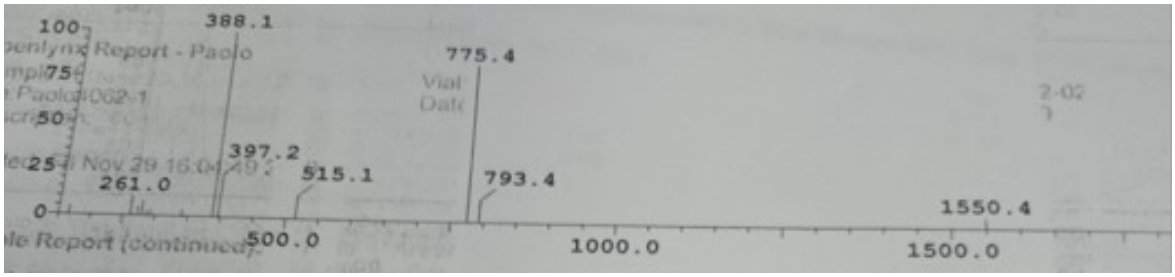


Figure S1. UPLC-MS chromatograms of the tested compound **18** obtained using 7- and 12-min methods.

A



B



C

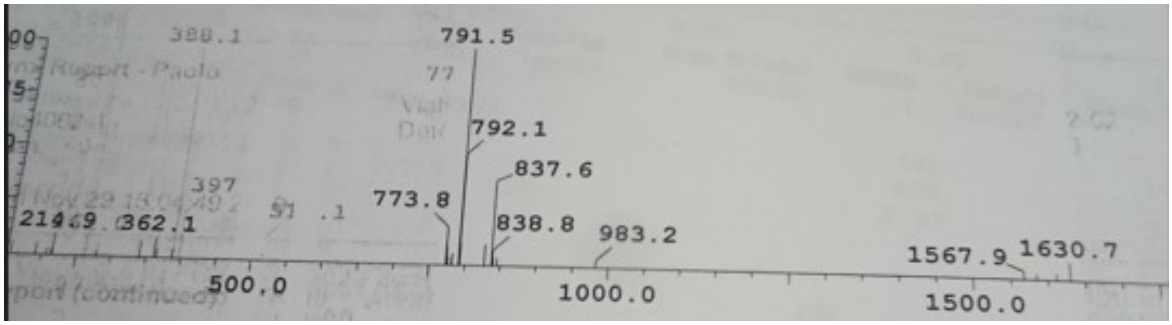


Figure S2. HPLC-MS chromatograms of the tested compound **18** obtained using 10-min (**A**) method.

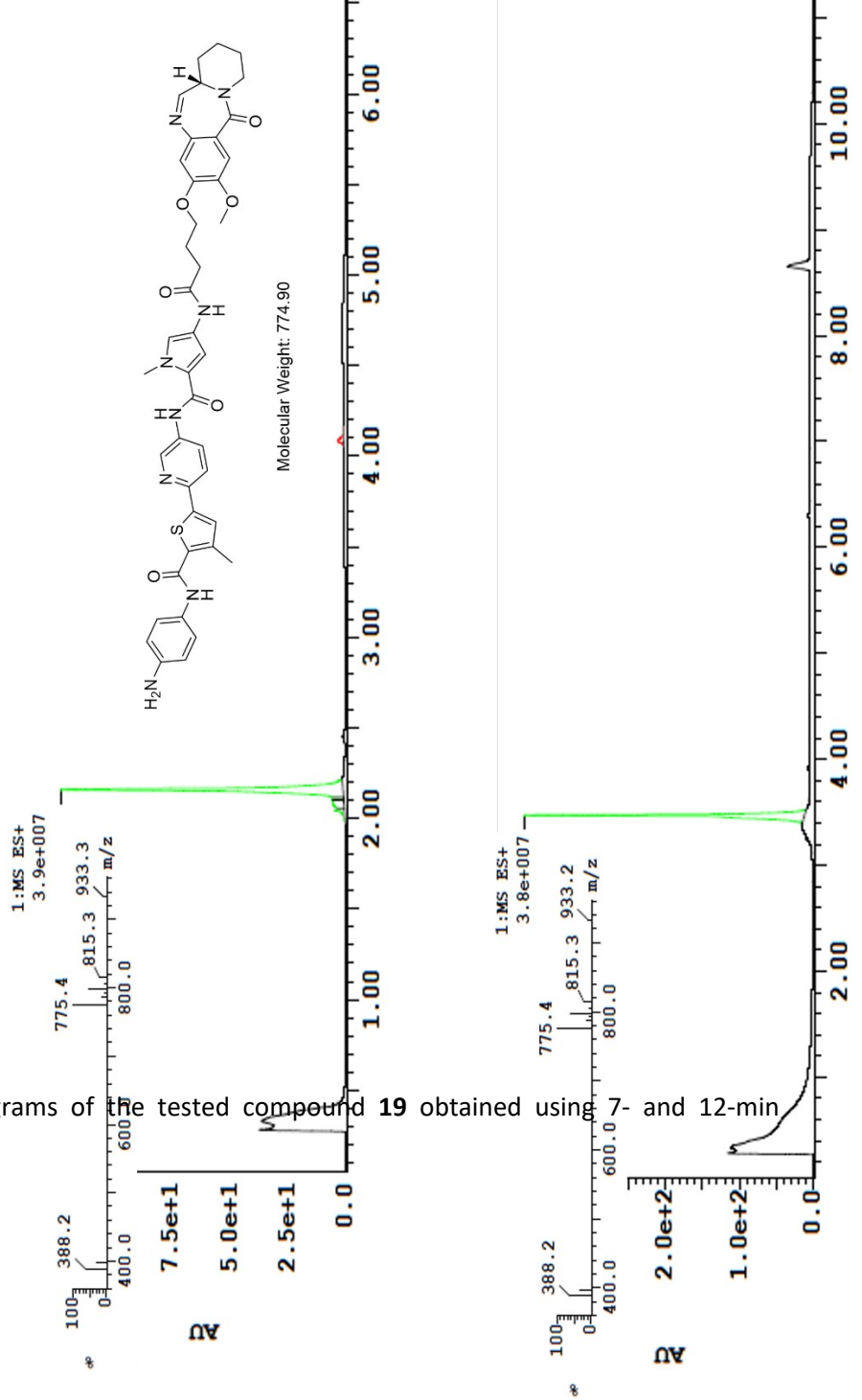


Figure S3. UPLC-MS chromatograms of the tested compound **19** obtained using 7- and 12-min methods.

DNA Fluorescence Melting Curves on sequence 2, 3, 4 and 5

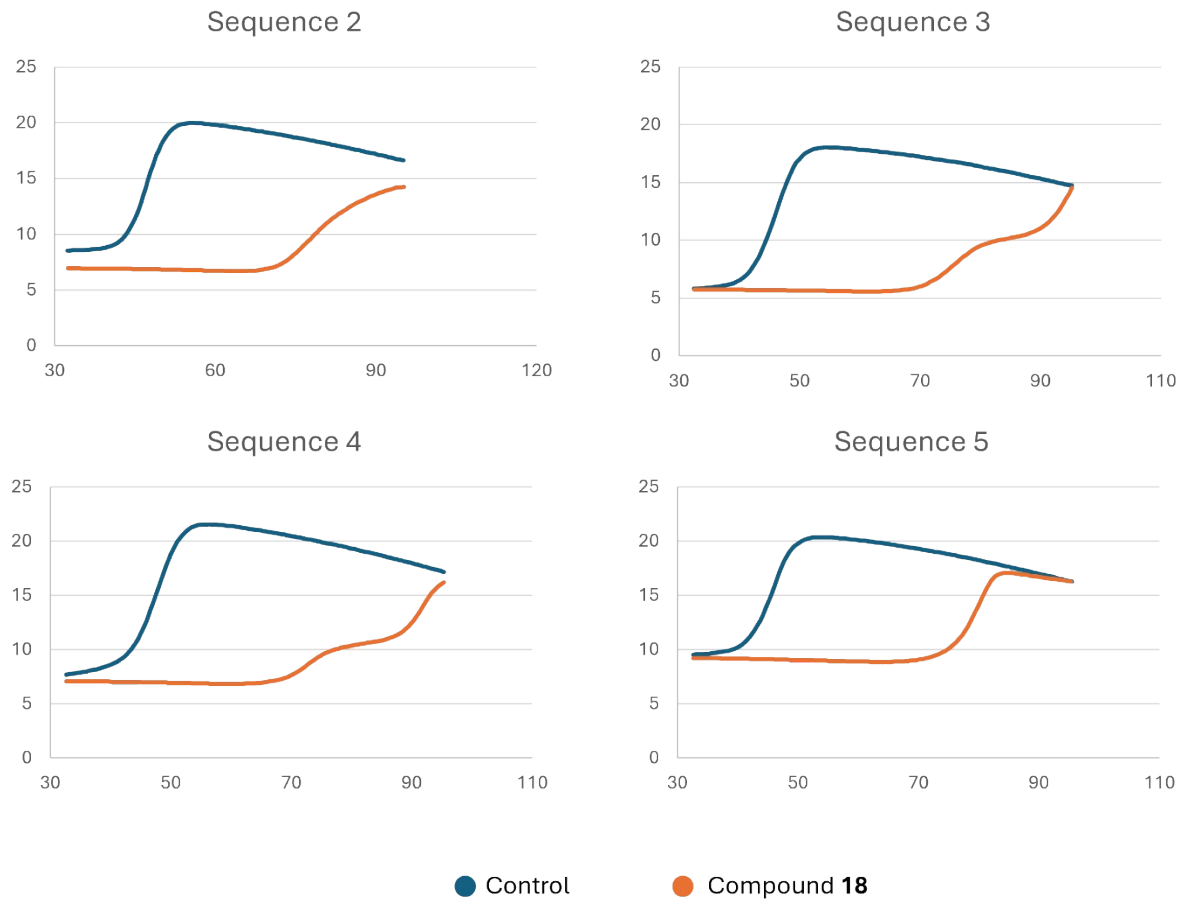


Figure S4. Melting profile of sequences 2-5 treated with compound **18**.

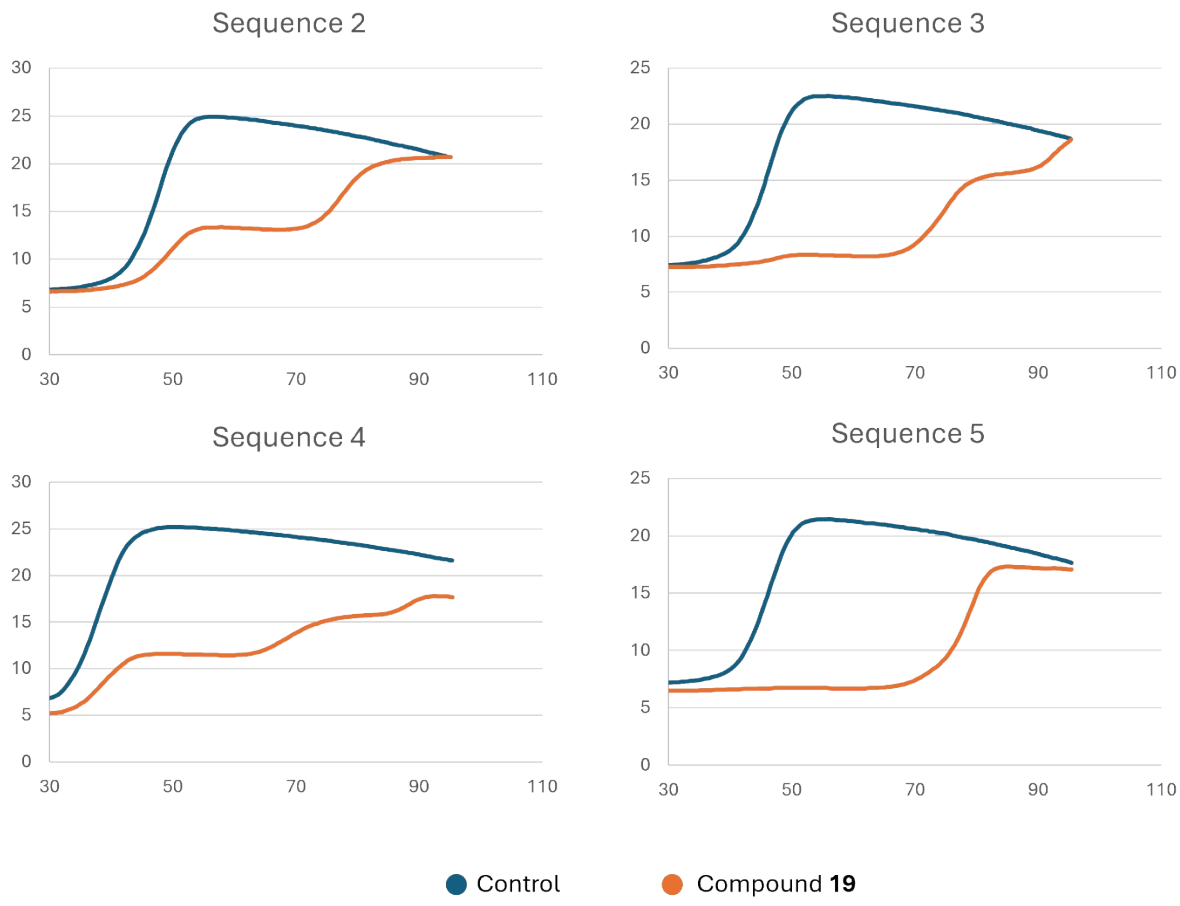


Figure S5. Melting profile of sequences 2-5 treated with compound **18**.

Transcription factors down-regulation and up-regulation

TRANSCRIPTION FACTOR	% CHANGE
ROR	-30.79
HIF	-29.24
NKX2-5	-26.74
RUNX	-23.34
GLI-1	-21.73
PIT1	-21.50
NFKB	-21.31
PIT	-20.85
CDP	-19.72
ELK	-19.05
TCF/LEF	-18.13
PAX8	-18.12
HEN	-17.88
EGR	-15.89
HOX A-5	-15.60
OCT-01	-15.17
PAX2	-14.15
FREAC2	-13.64
COUP TF	-12.52
NFAT	-12.05
NF-1	-11.76
SNAIL	-11.55
FOX01,FKHR	-11.53
PLAG1	-11.37
HSF	-11.26
RXR	-10.86
CREB	-10.73
SF-1	-10.51
YY1	-10.45
XBP	-9.94
PPAR	-9.87
GATA	-9.69
FOXC1	-9.53
AP1	-9.49
PBXL	-9.15
TR	-8.90
SRY	-8.58
MYC MAX	-8.33
AP3	4.02
NRF2-ARE	4.17
GR/PR	4.66
FOXG1	4.99

PROX1	5.06
E2F1	5.49
KLF-4	5.50
HOX4C	5.84
PXR	6.18
ATF2	6.28
SATB1	6.42
STAT3	6.48
OCT-01	6.57
BM-3	6.69
NF-E2	7.07
NRF-1	7.37
AP2	12.67
STAT4	13.01
RB	13.58
FOXD3	13.64
SPI	13.72
ETS	13.72
AR	13.81
SOX18	14.18
AP4	14.25
SOX9	14.32
VDR	14.32
SOX2	14.37
ER	14.41
GAS/ISRE	14.41
PAX3	14.43
STAT1	14.61
TFIID	14.65
GFI1	14.66
WTI	14.71
SMUC	14.85
STAT5	15.01
STAT6	15.30
MEF-2	15.47
MEF1	15.51
SRF	15.65
MZF	16.04
MYOD	16.05
IRF	16.36
CAR	16.59
HNF-4	16.64
SMAD	16.66
HNF-1	16.74
PAX-5	16.91
C/EBP	17.44
CBF	18.11
FOXA1	18.39

MYB	18.99
TFE3	19.10
FAST -1	20.61
P53	20.67
NKX3-2	21.07
USF1	21.77

Table S2. Transcription factors down-regulation and up-regulation expressed as percentage, in HeLa cells after treatment with after 6-hour treatment with 100 mM of PDD analogue **18**.

Characterization of compounds 18-20

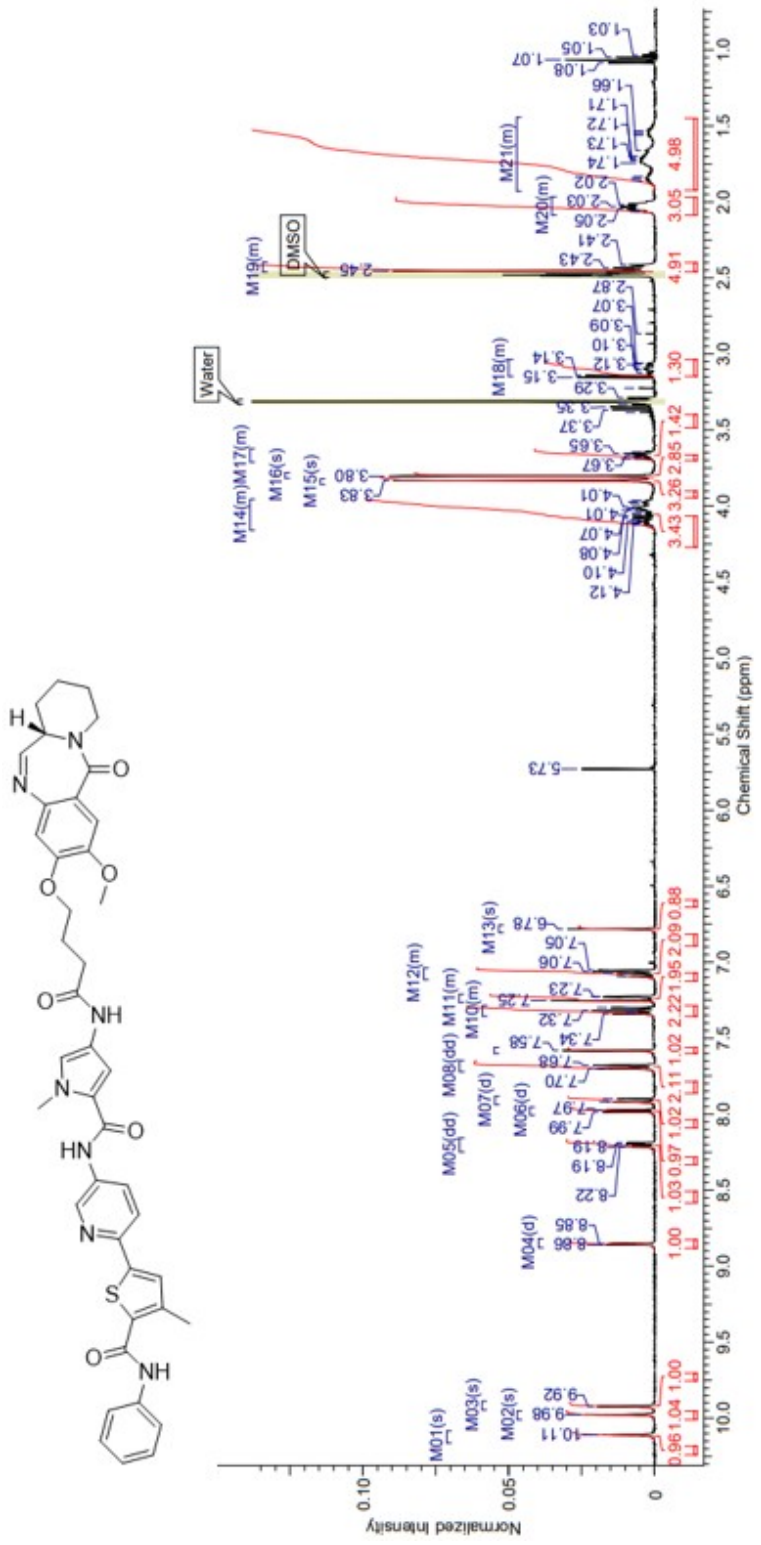


Figure S6. Proton NMR of compound **18**.

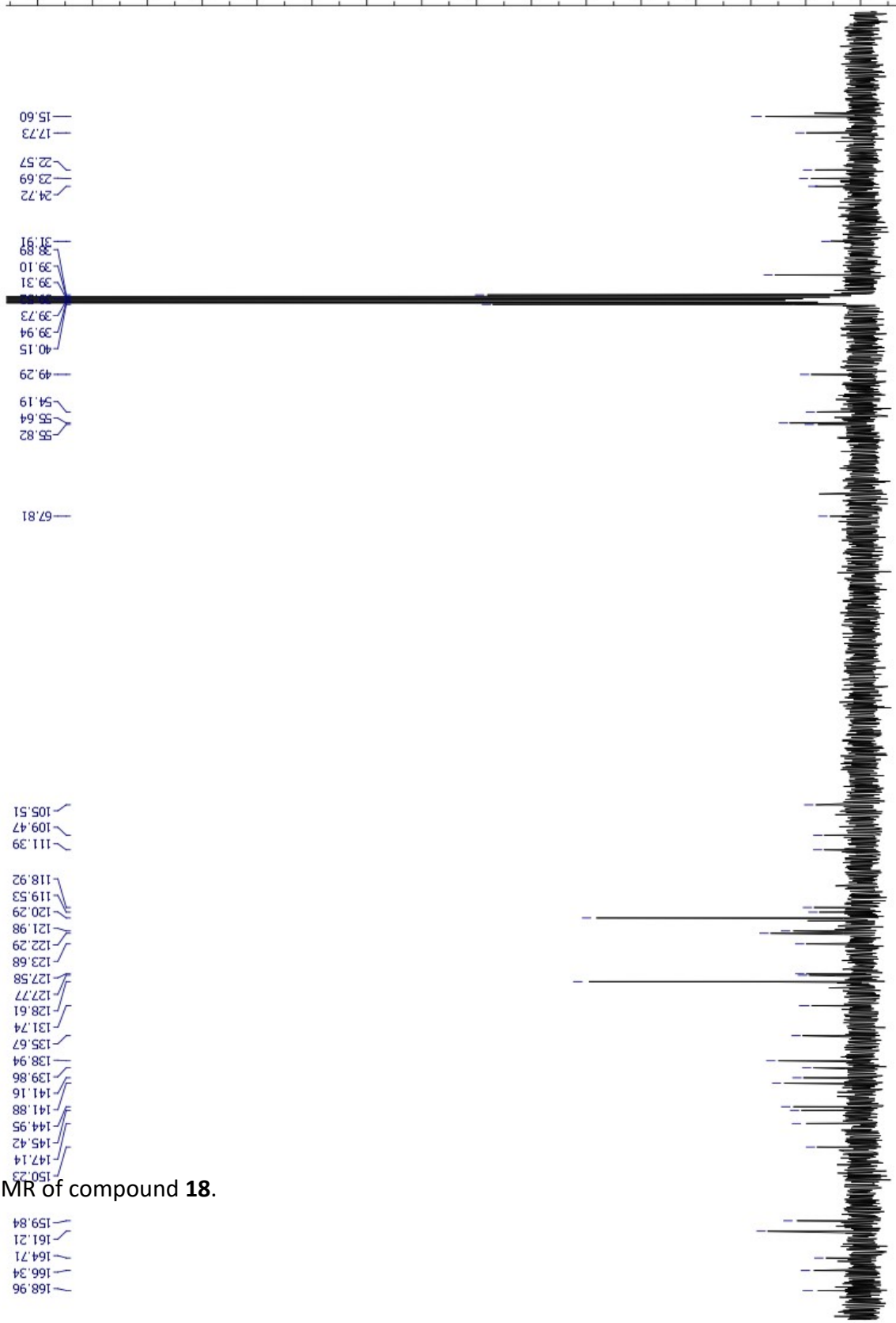


Figure S7. Carbon NMR of compound **18**.

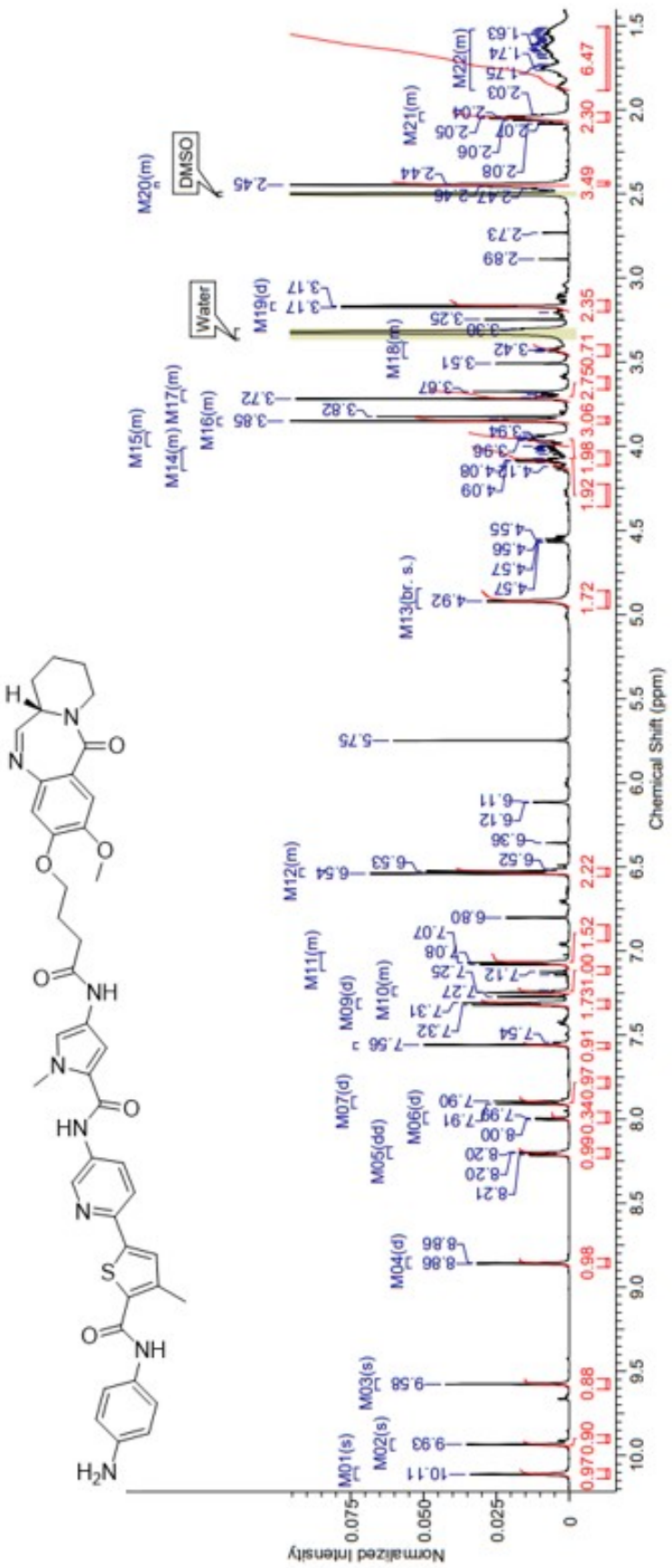


Figure S8. Proton NMR of compound **19**.

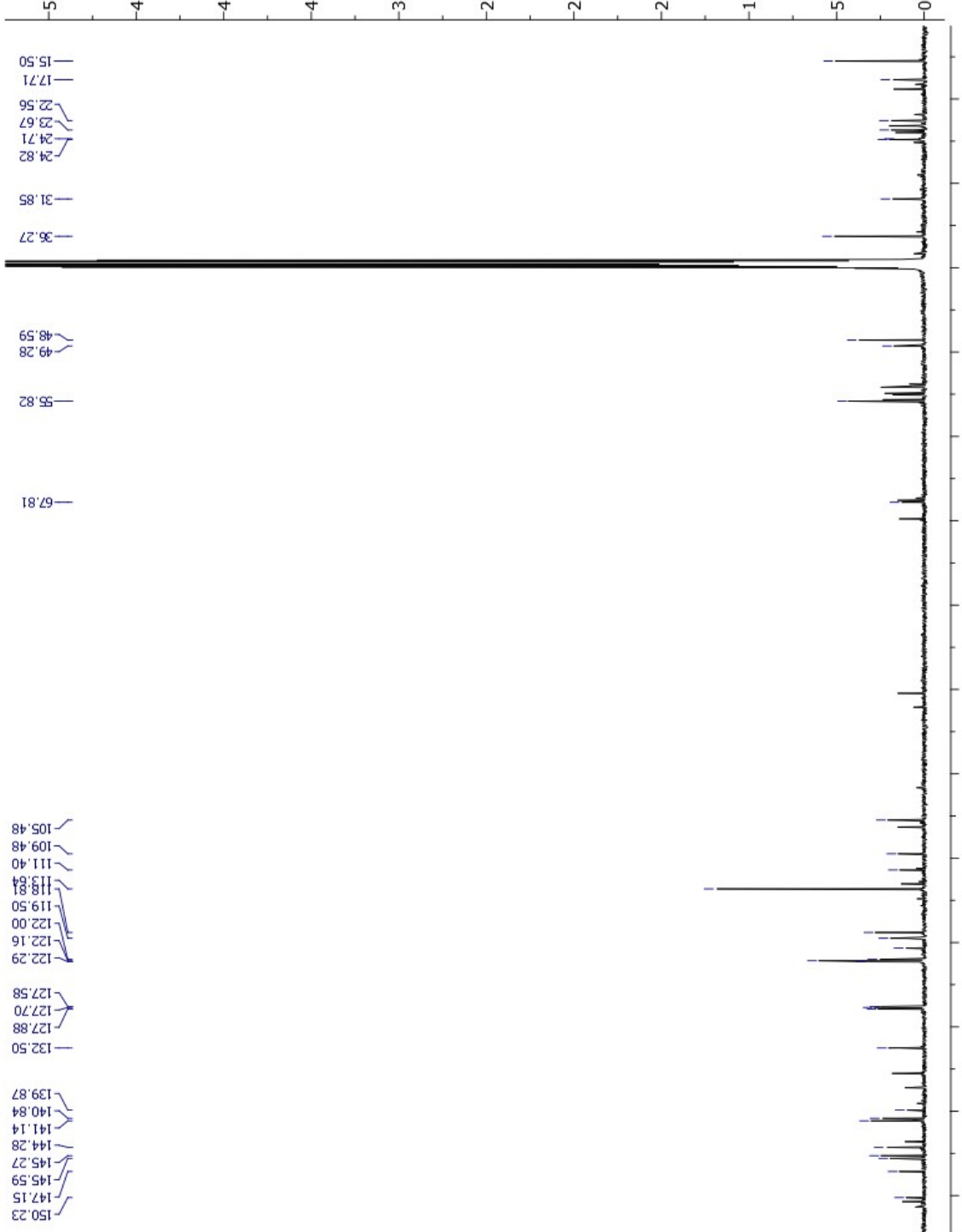


Figure S9. Carbon NMR of compound **19**.

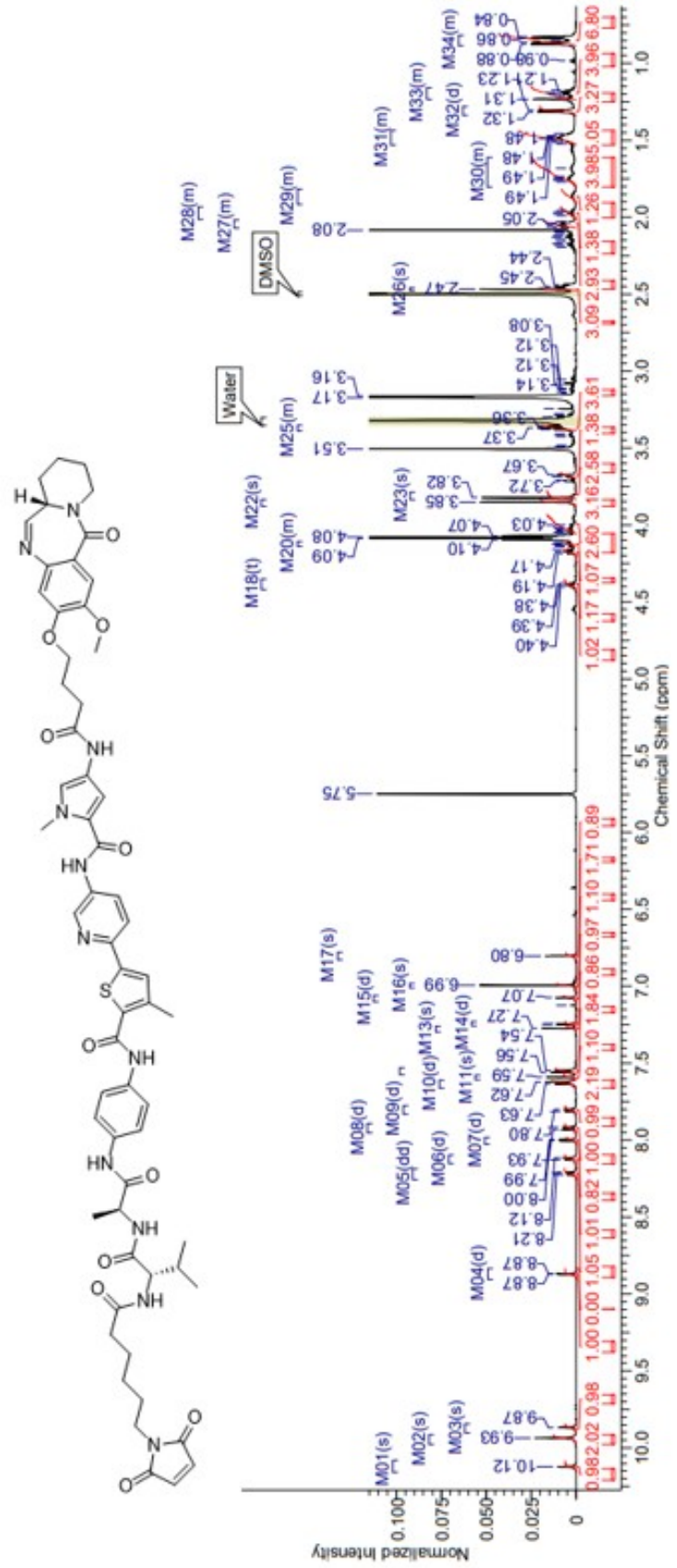


Figure S10. Proton NMR of compound **20**.

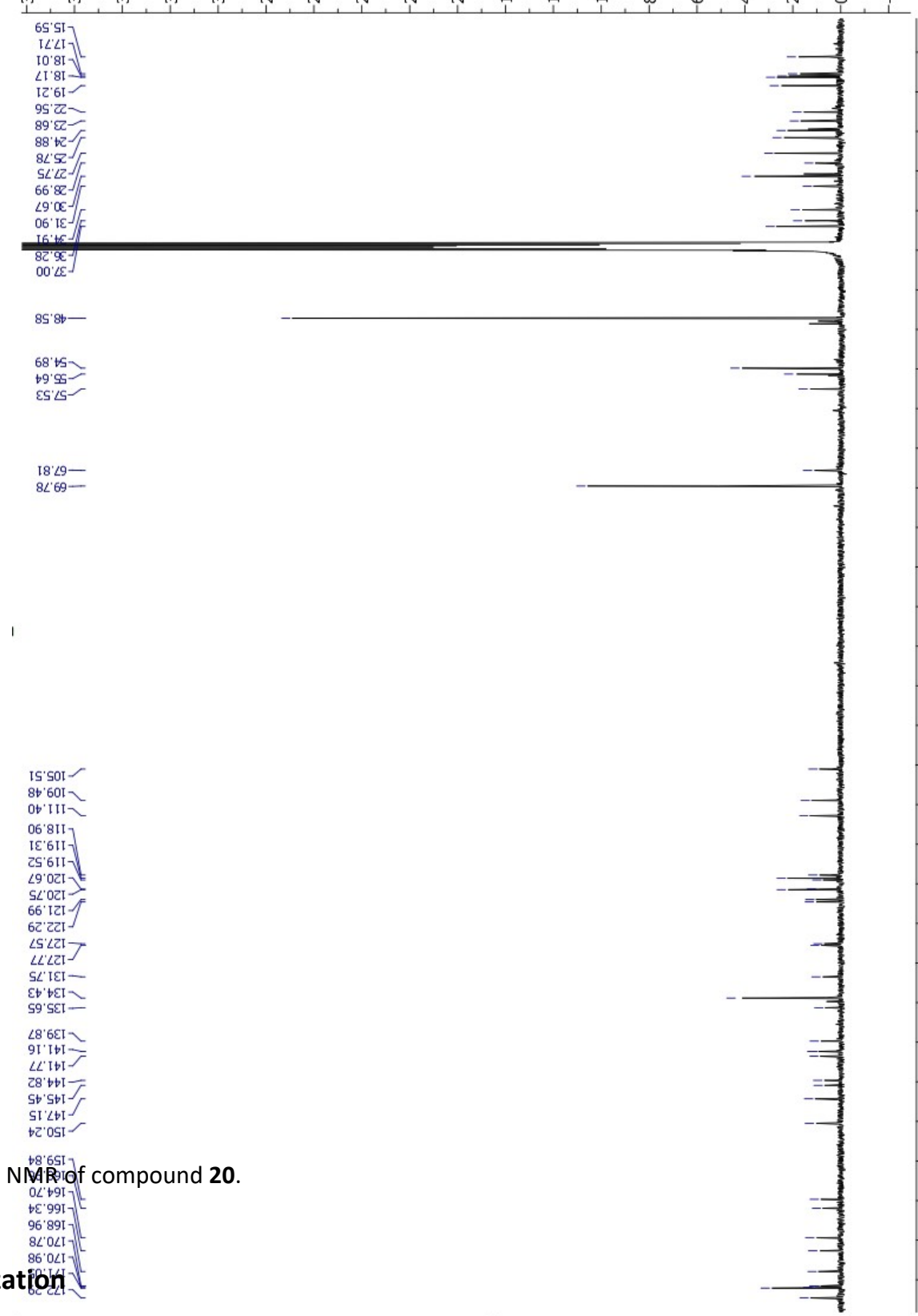


Figure S11. Carbon NMR of compound **20**.

ADC Characterization

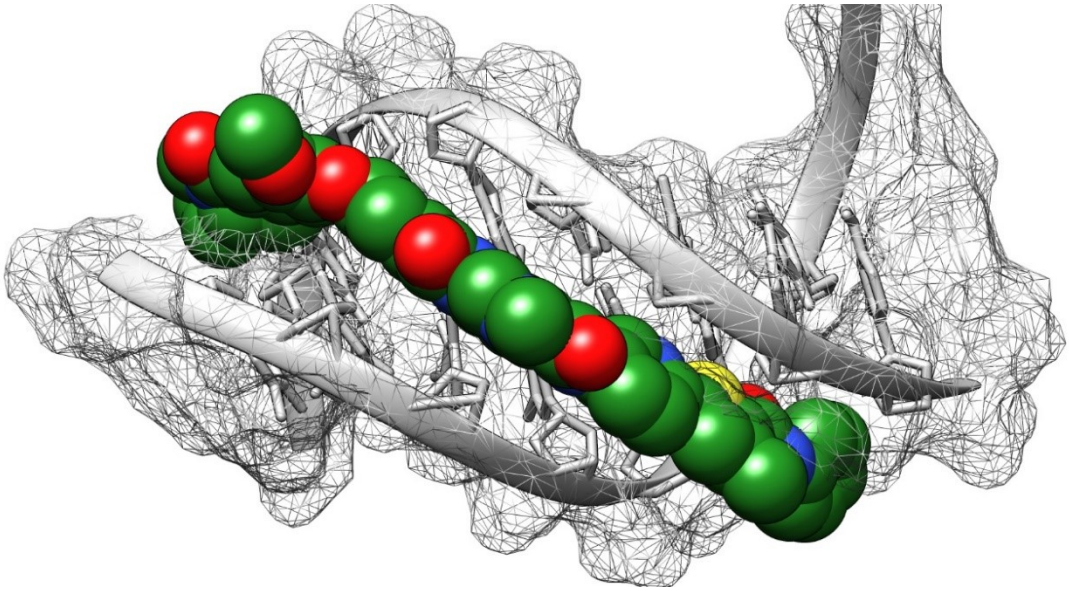


Figure S12. Low energy snapshot of **18** covalently bound to a 5'-AGAAAGAA-3 sequence.

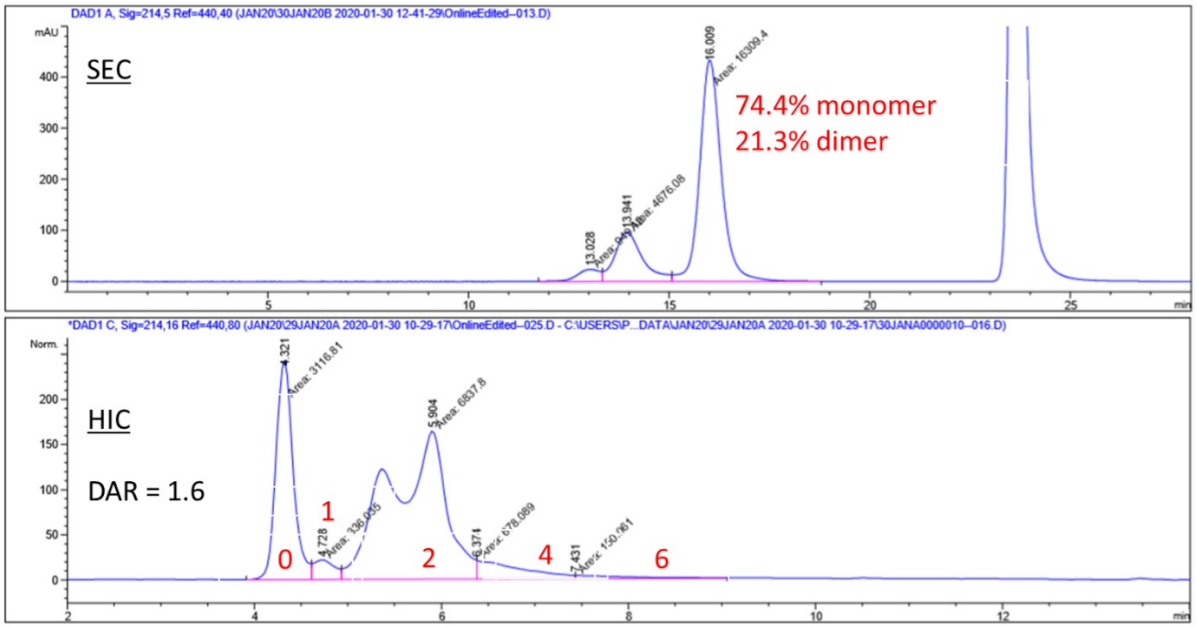


Figure S13. Size Exclusion Chromatography (SEC) chromatogram (top panel) showing percentage of monomer of Trastuzumab-(**20**) and Hydrophobic Interaction Chromatography (HIC) chromatogram (bottom panel) showing its DAR species distribution.

Cytotoxicity curves

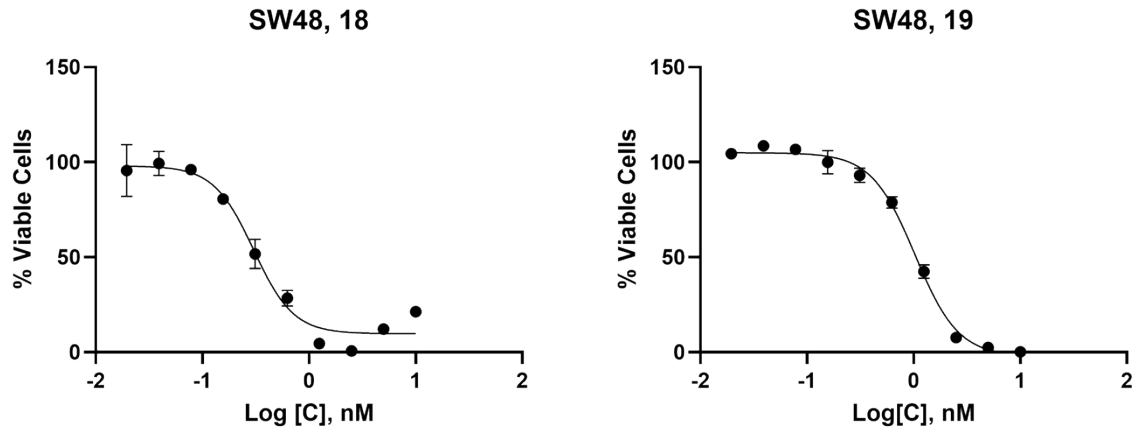


Figure S14. Cytotoxicity of **18** and **19** in SW48 cell line

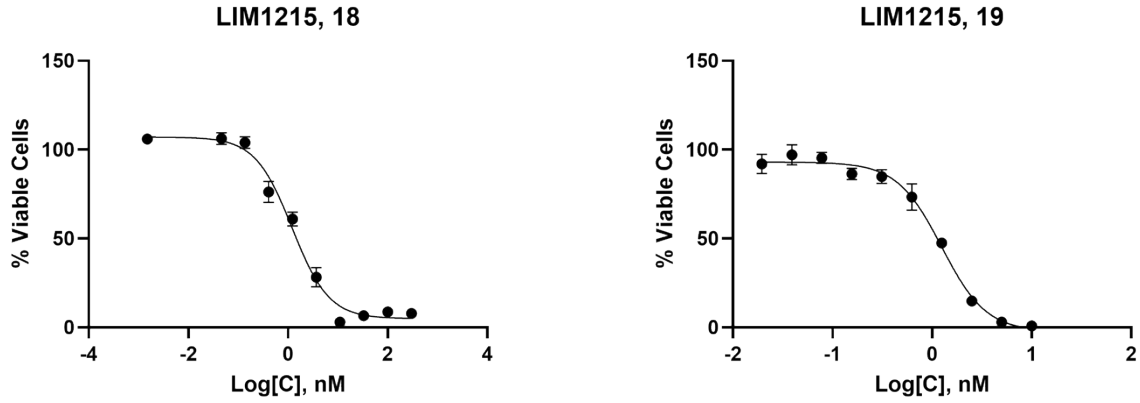


Figure S15. Cytotoxicity of **18** and **19** in LIM1215 cell line

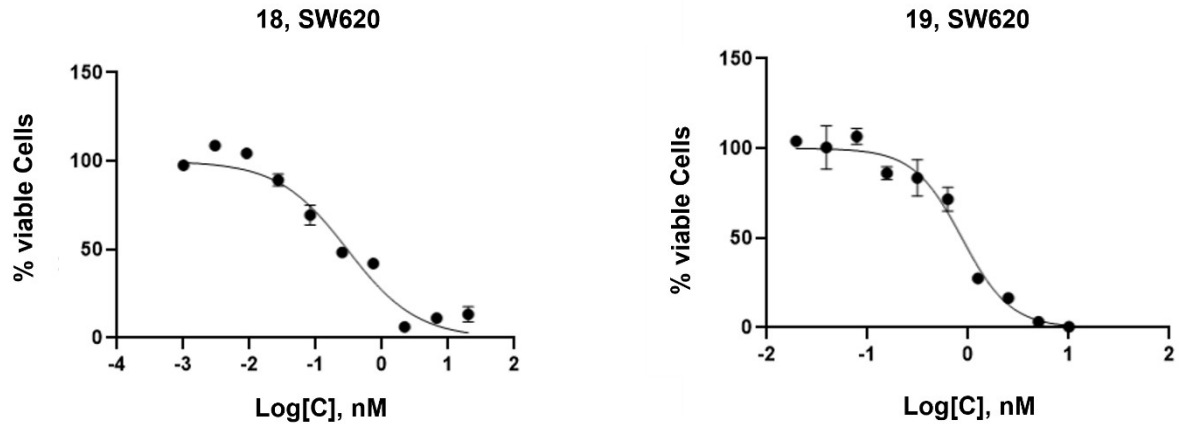


Figure S16. Cytotoxicity of **18** and **19** in SW620 cell line

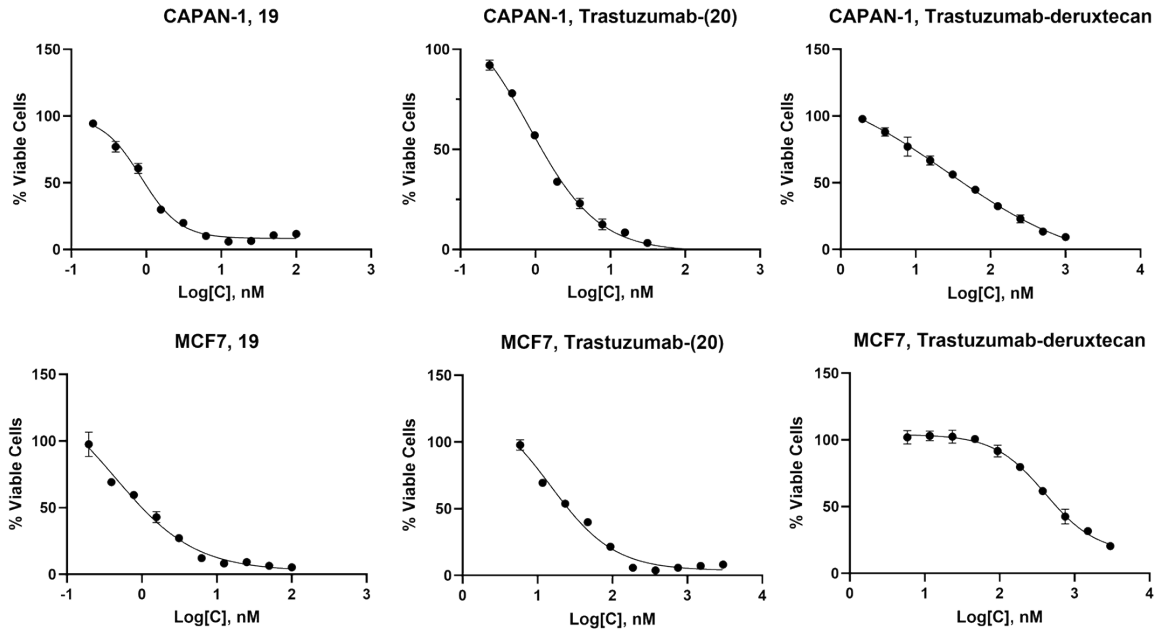


Figure S17. Cytotoxicity of **19**, trastuzumab-**20** and trastuzumab-deruxtecan in CAPAN1 and MCF-7 cell lines