

Supplementary Information

Discovery of a series of 1,2,3-triazole-containing Erlotinib derivatives with potent anti-tumor activities against non-small cell lung cancer (NSCLC)

Ge Sun^{a,#}, Longfei Mao^{b,#}, Wenjing Deng^a, Shuxiang Xu^a, Jie Zhao^b, Jianxue Yang^{c,d}, Kaitai Yao^{a*}, Miaomiao Yuan^{e*} and Wei Li^{b*}

^aCancer Research Institute, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, Guangdong, China

^bSchool of Chemistry and Chemical Engineering, Henan Normal University, Henan Engineering Research Center of Chiral Hydroxyl Pharmaceutical, Xinxiang 453007, China.

^cDepartment of Neurology, The First Affiliated Hospital of Henan University of Science and Technology 450052, Henan, China.

^dSchool of Nursing, Henan University of Science and Technology 450062, Henan, China

^eThe Eighth Affiliated Hospital, Sun Yat-sen University, Shenzhen 518033, Guangdong, China.

***Corresponding authors:**

Kaitai Yao, Southern Medical University, Guangzhou 510515, Guangdong, China. E-mail: ktyao1931@163.com

Miaomiao Yuan, The Eighth Affiliated Hospital, Sun Yat-sen University, Shenzhen

518033, Guangdong, China. E-mail: yuanmm2019@163.com

Wei Li, School of Chemistry and Chemical Engineering, Henan Normal University,

Henan Engineering Research Center of Chiral Hydroxyl Pharmaceutical, Xinxiang

453007, China. E-mail: liweigq@163.com

Table of Contents

Figure S1-1. ^1H NMR spectrum (600 MHz, DMSO-d_6) of compound e1

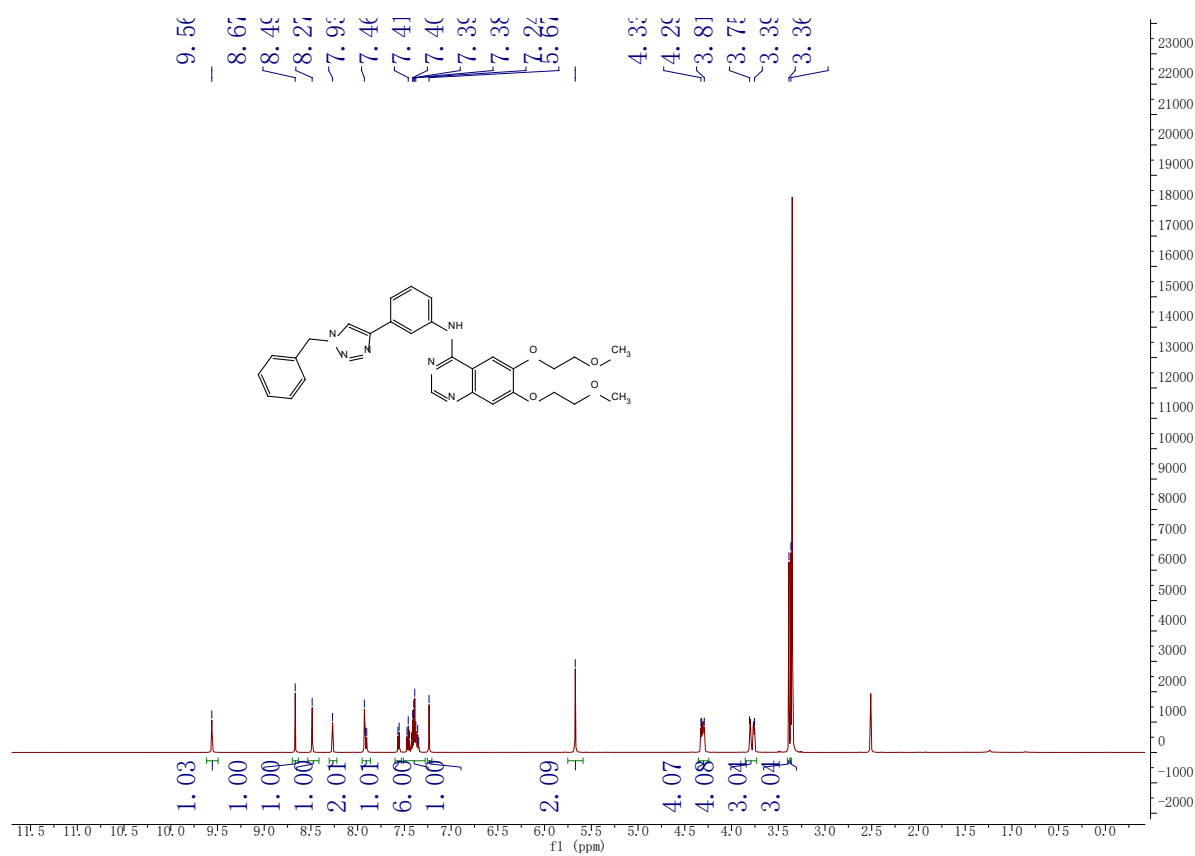


Figure S1-2. ^{13}C NMR spectrum (150 MHz, DMSO-d_6) of compound e1

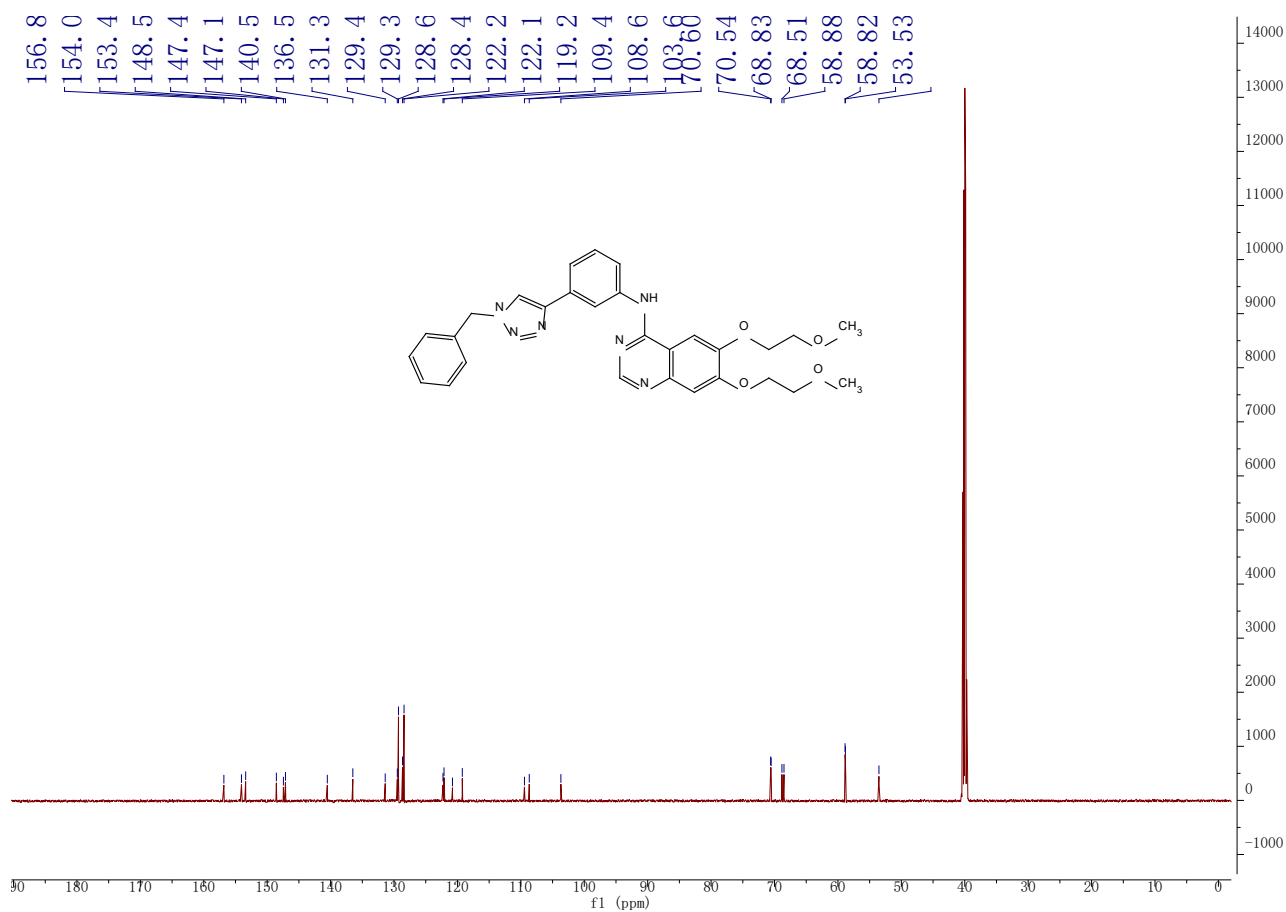


Figure S1-3. HR MS of compound e1

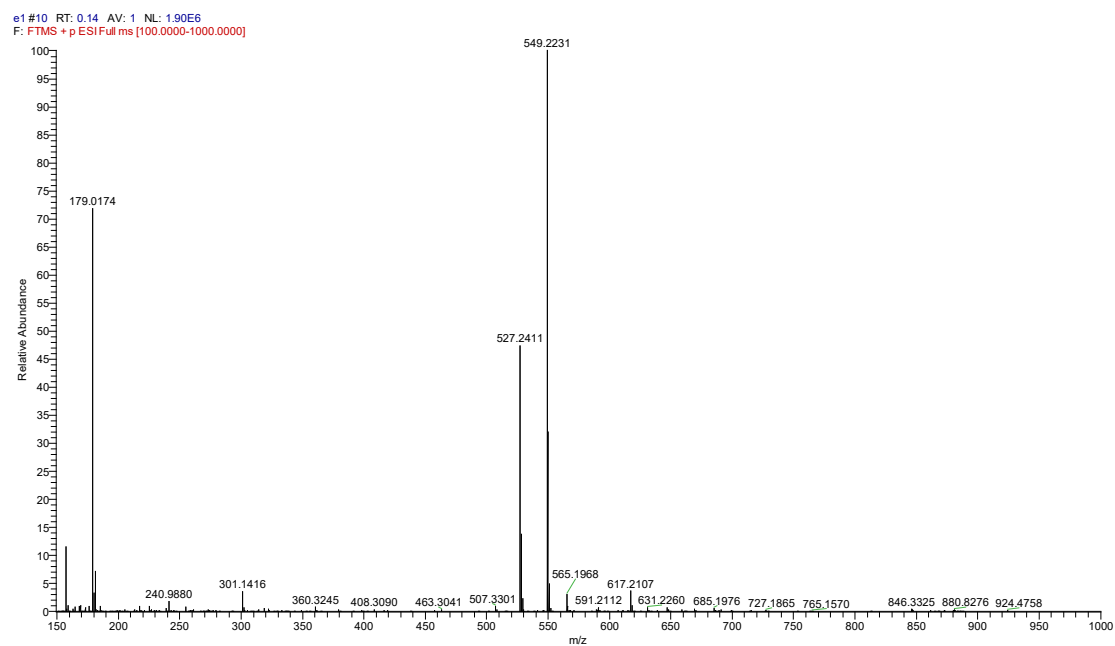


Figure S2-1. ^1H NMR spectrum (600 MHz, DMSO-d_6) of compound e2

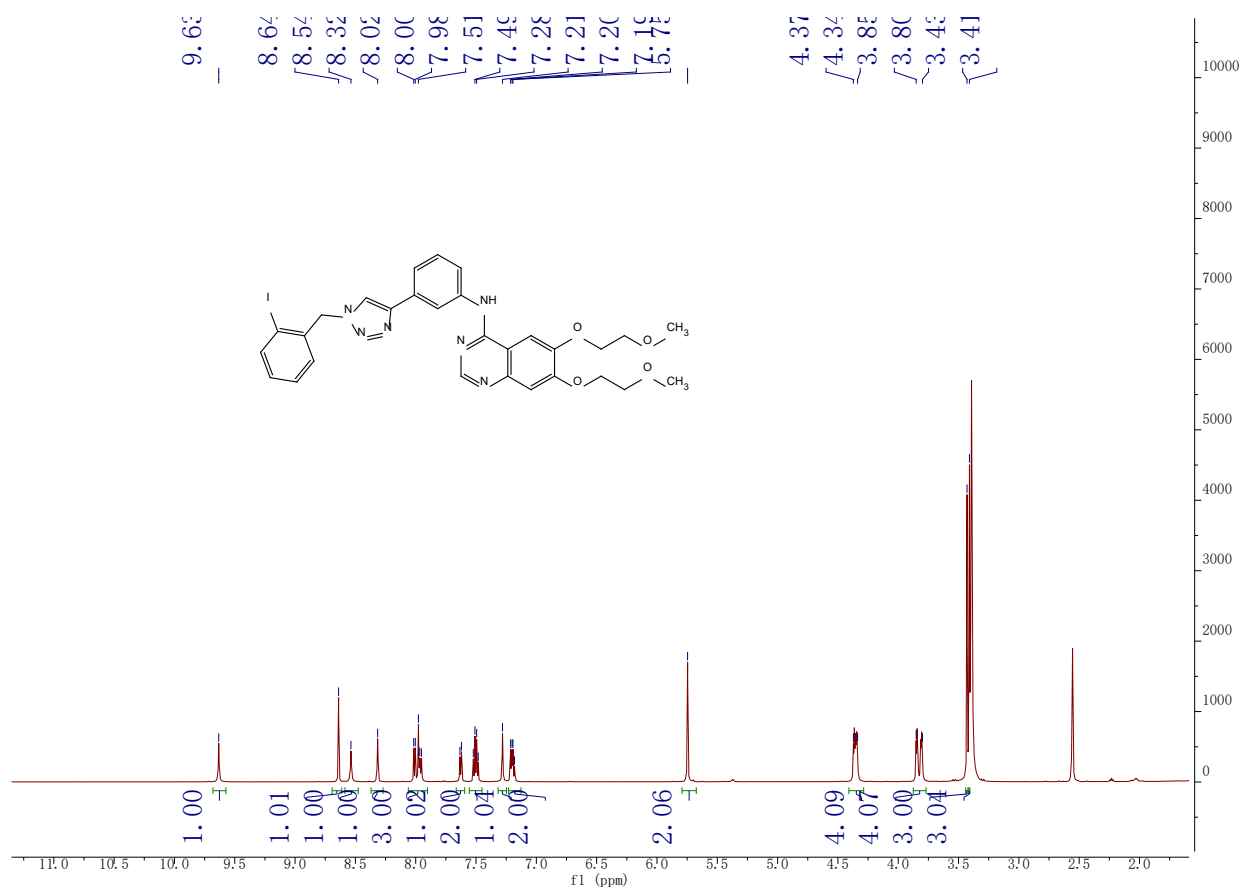


Figure S2-2. ^{13}C NMR spectrum (150 MHz, DMSO- d_6) of compound e2

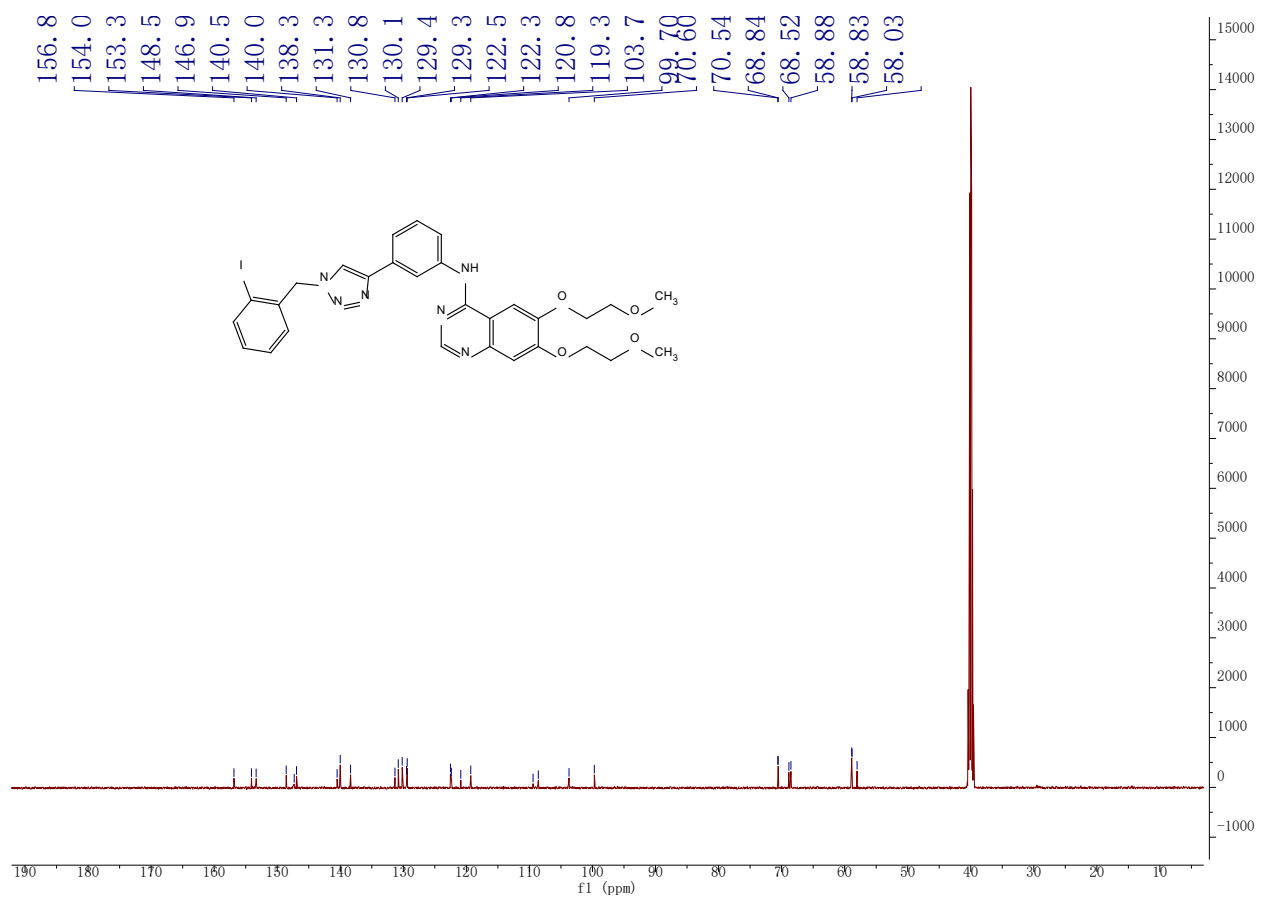
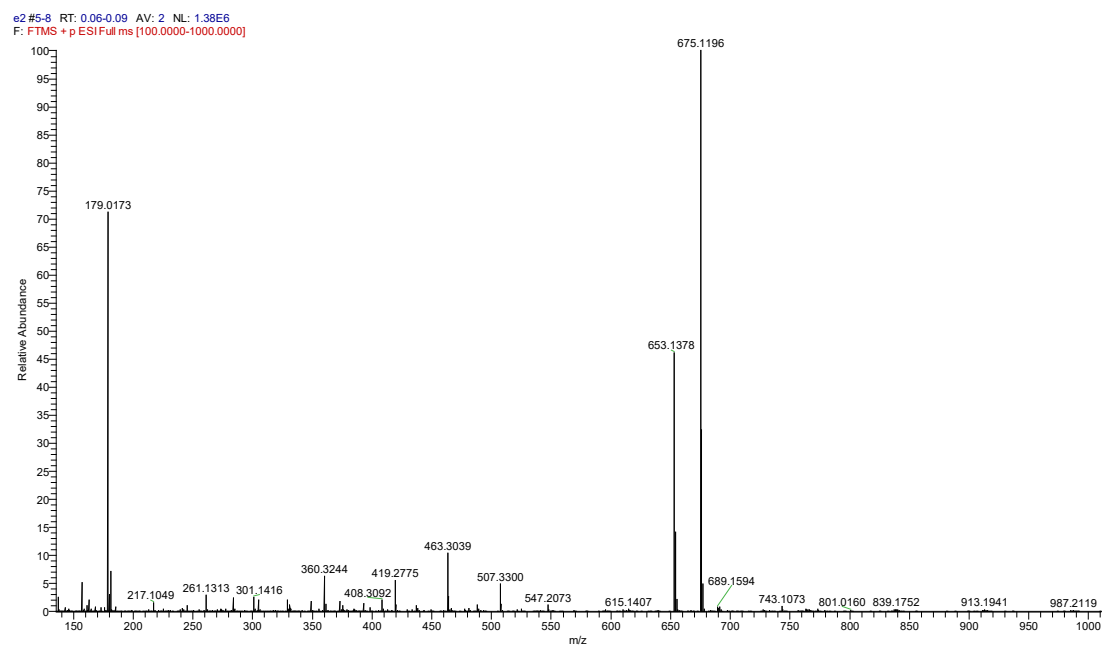


Figure S2-3. HR MS of compound e2



Chemical structure of compound 10: COCCOC1=CC=C2C(=C1)N=CN=C2C(=O)Nc3ccc(cc3)C4=NC=NC=C4Cc5ccccc5Br

¹H NMR spectrum (CDCl₃) of compound 10. The x-axis represents the chemical shift in ppm (f1), ranging from 1.2 to 10.5. The y-axis represents the intensity, ranging from -1000 to 13000. The spectrum shows several peaks with corresponding integration values.

Chemical Shift (ppm)	Integration
9.60	1.00
8.62	1.03
8.27	1.01
7.95	1.01
7.72	1.99
7.72	1.00
7.58	1.00
7.47	2.01
7.45	1.01
7.38	1.01
7.28	2.00
7.28	2.03
7.26	4.00
7.26	4.12
7.26	3.01
7.26	3.11
7.26	3.38
7.26	3.38
2.50	1.00

Figure S3-2. ^{13}C NMR spectrum (150 MHz, DMSO- d_6) of compound e3

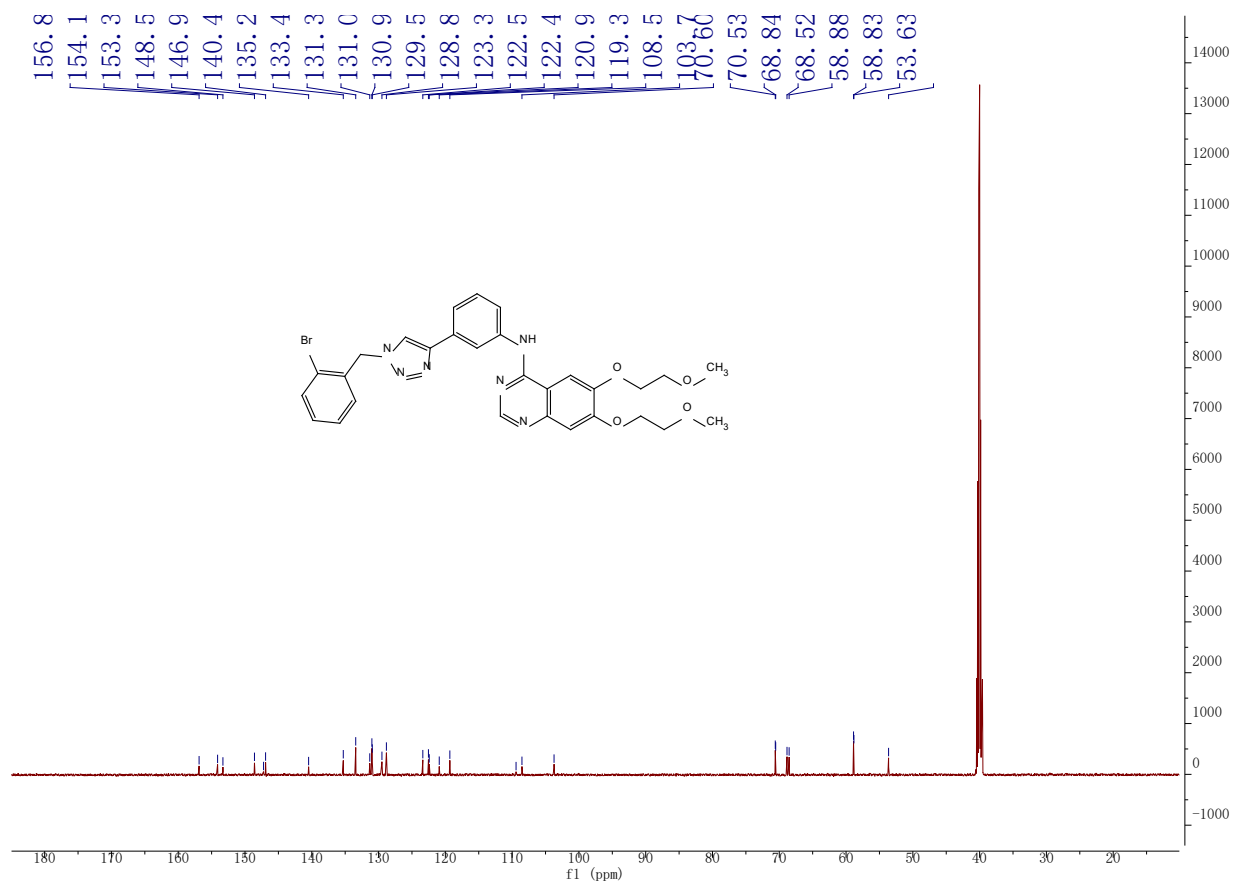


Figure S3-3. HR MS of compound e3

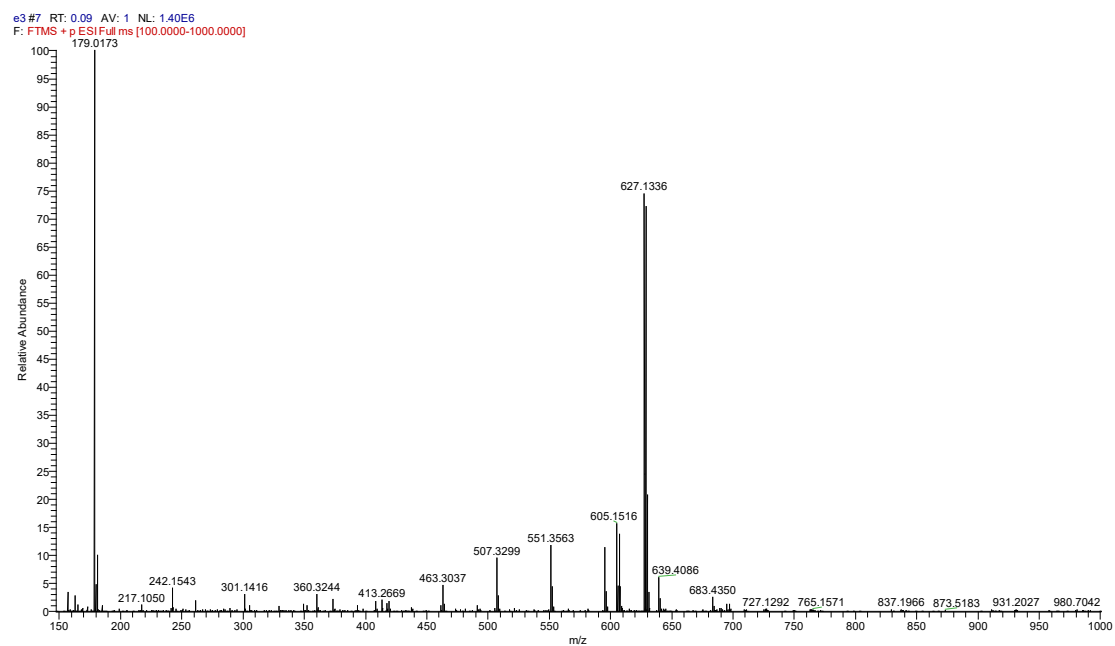


Figure S4-1. ^1H NMR spectrum (600 MHz, DMSO- d_6) of compound e4

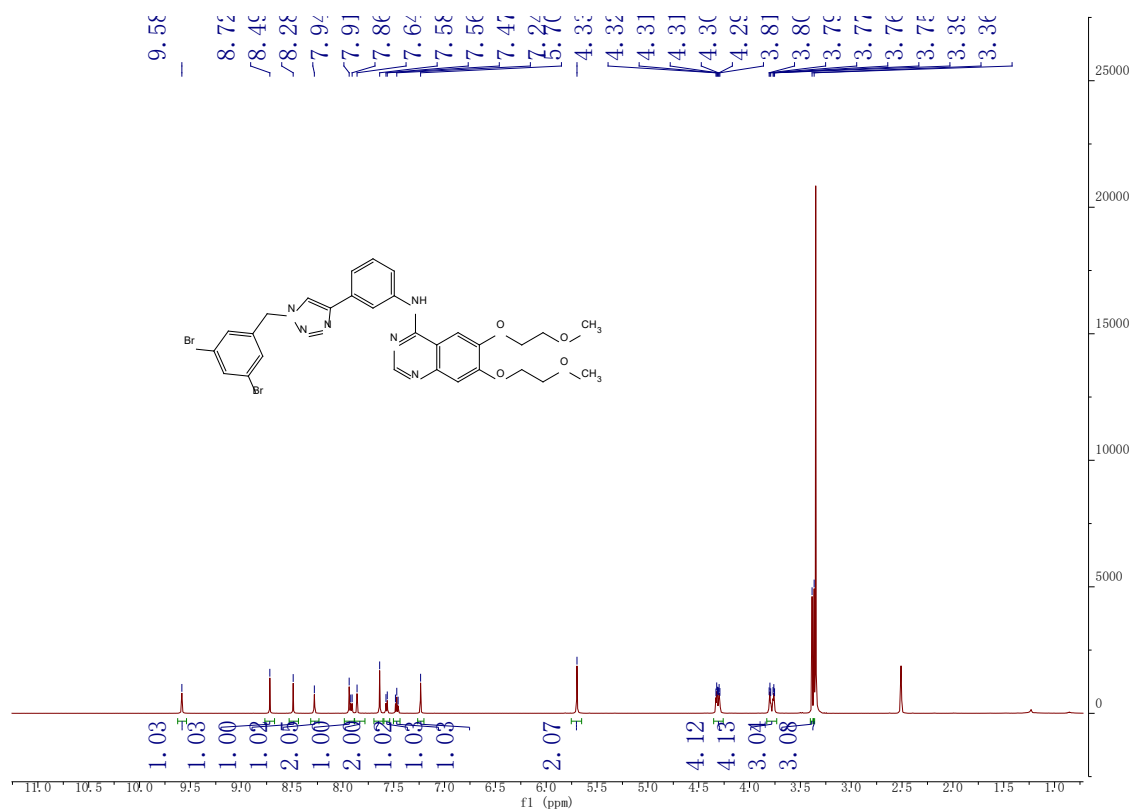


Figure S4-2. ^{13}C NMR spectrum (150 MHz, DMSO- d_6) of compound e4

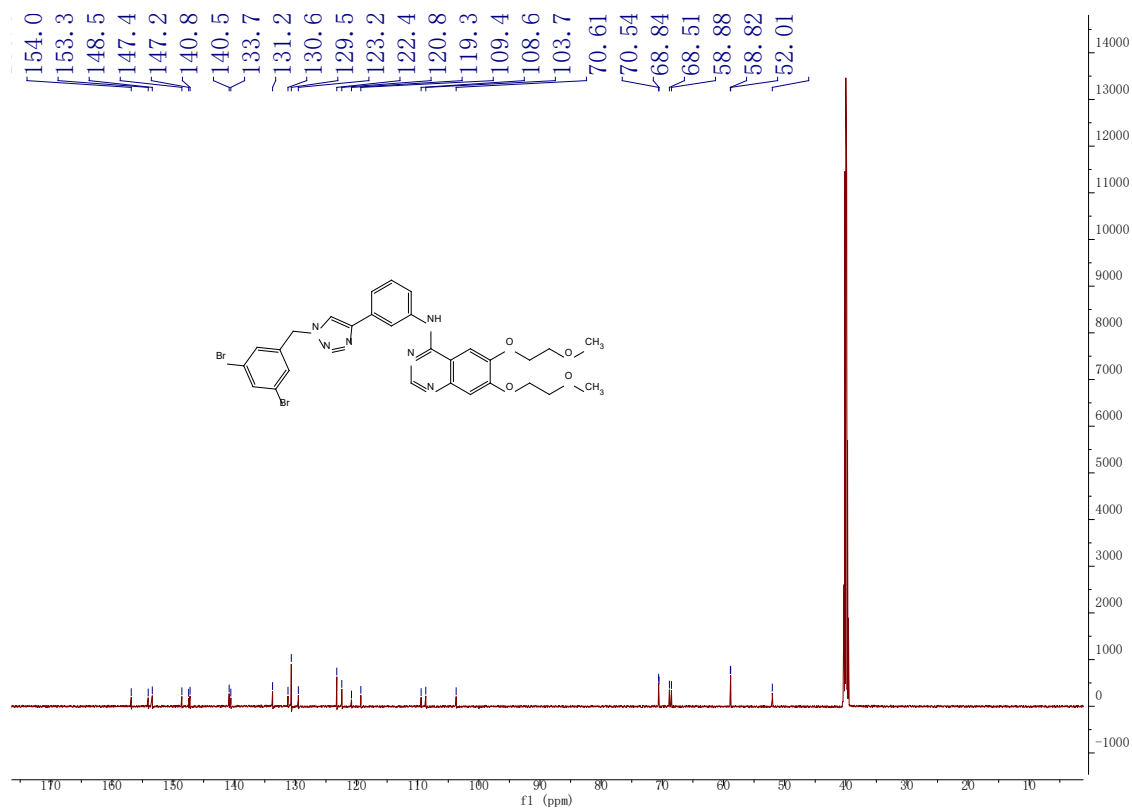


Figure S4-3. HR MS of compound e4

e4 #7 RT: 0.08 AV: 1 NL: 1.55E7
F: FTMS + p ESI Full ms [100.0000-1000.0000]

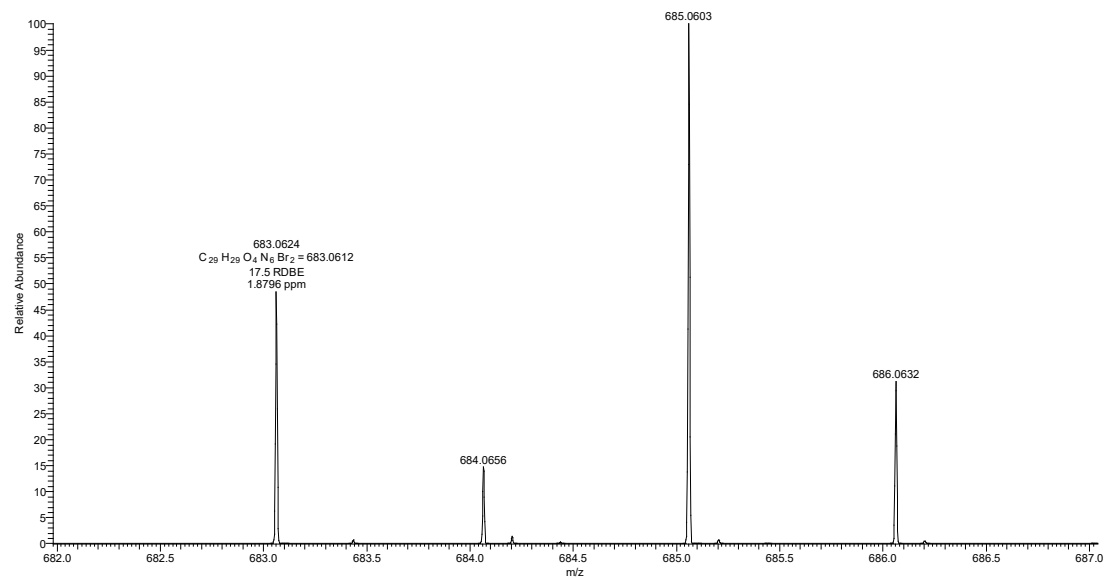


Figure S5-1. ^1H NMR spectrum (600 MHz, DMSO- d_6) of compound e5

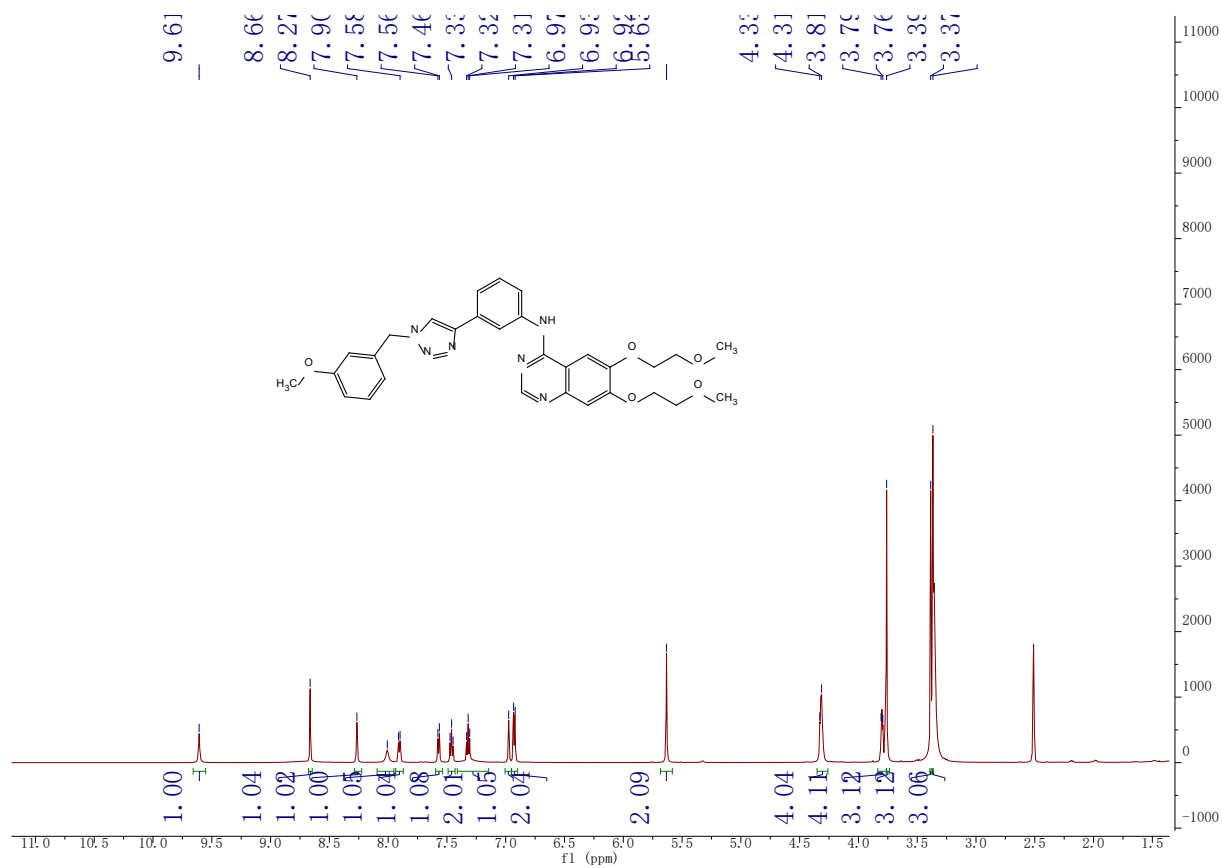


Figure S5-2. ^{13}C NMR spectrum (150 MHz, DMSO-d_6) of compound e5

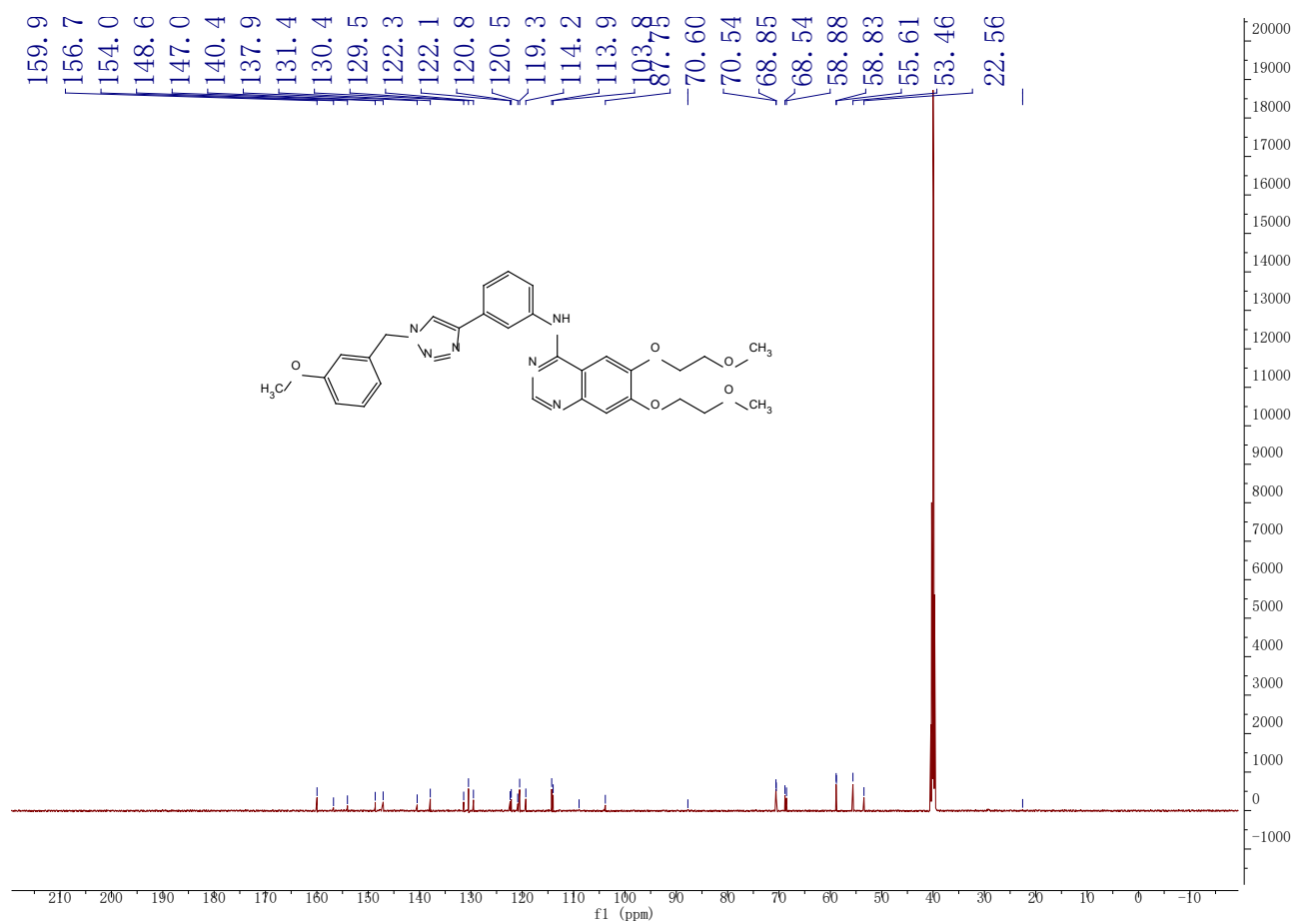


Figure S5-3. HR MS of compound e5

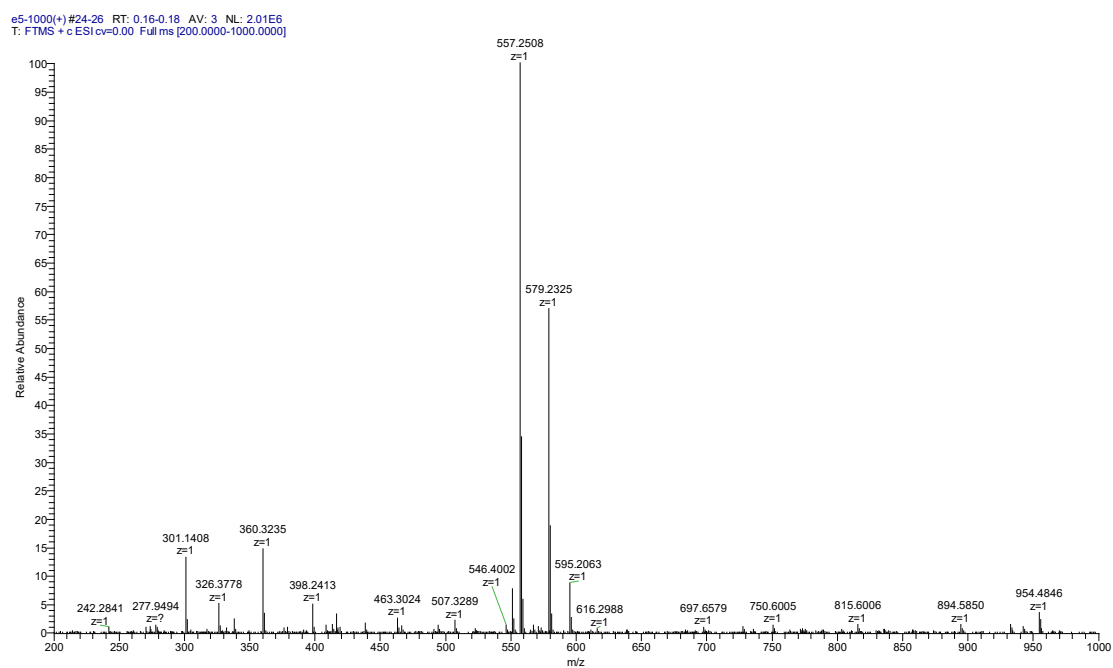


Figure S6-1. ^1H NMR spectrum (600 MHz, DMSO-d_6) of compound e6

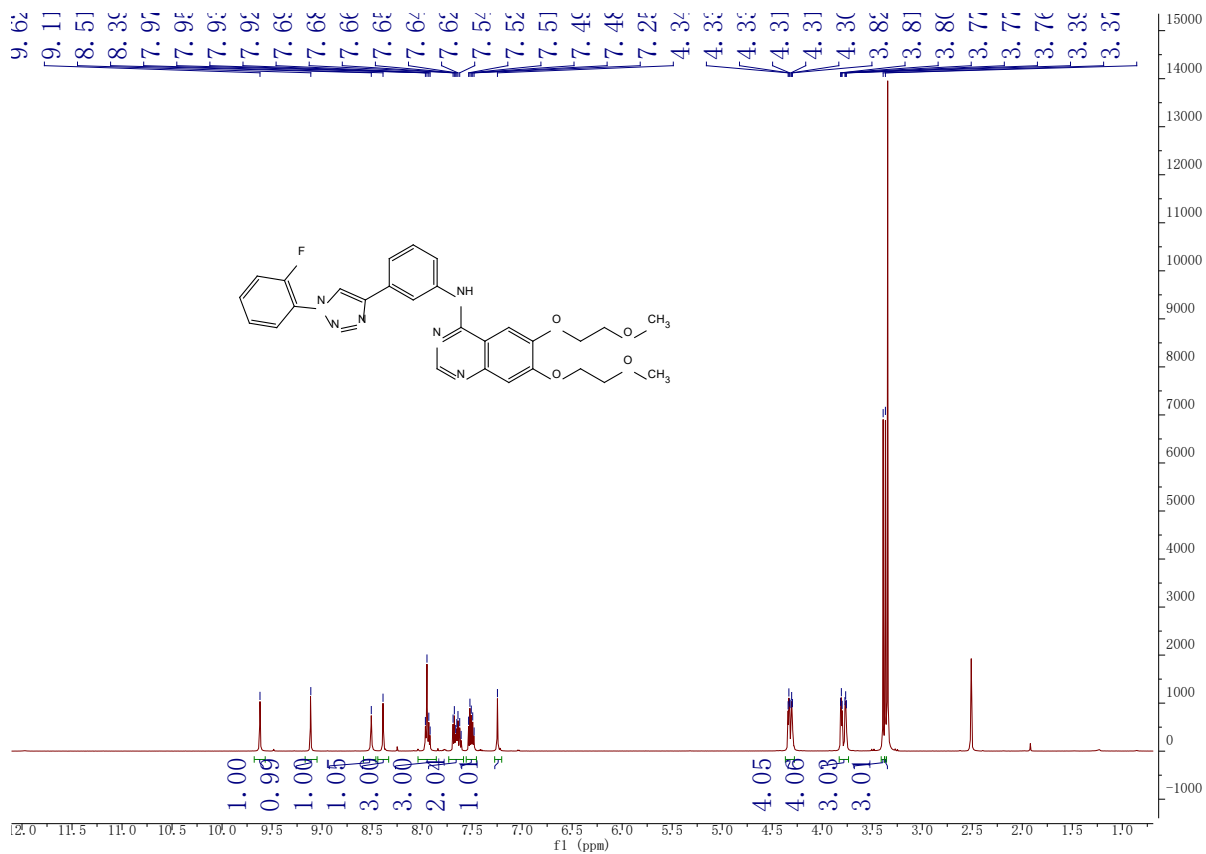


Figure S6-2. ^{13}C NMR spectrum (150 MHz, DMSO- d_6) of compound e6

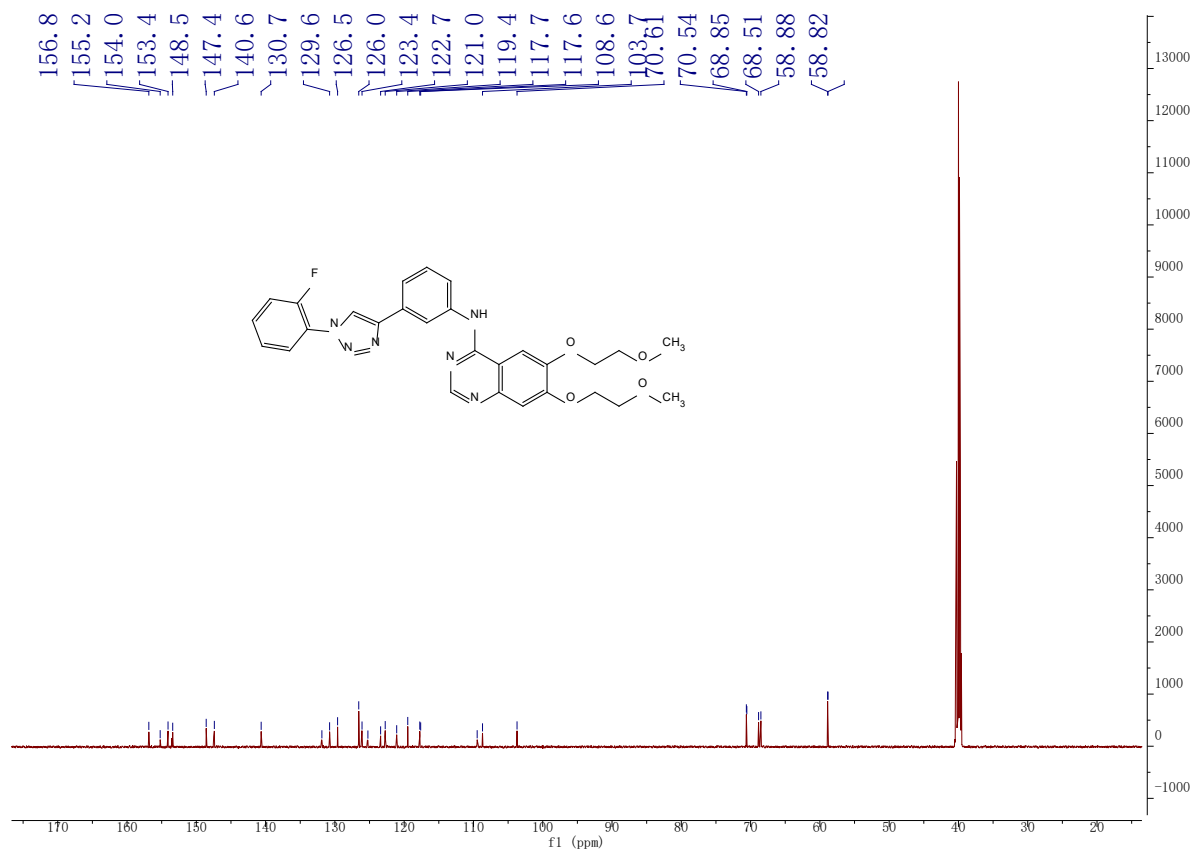
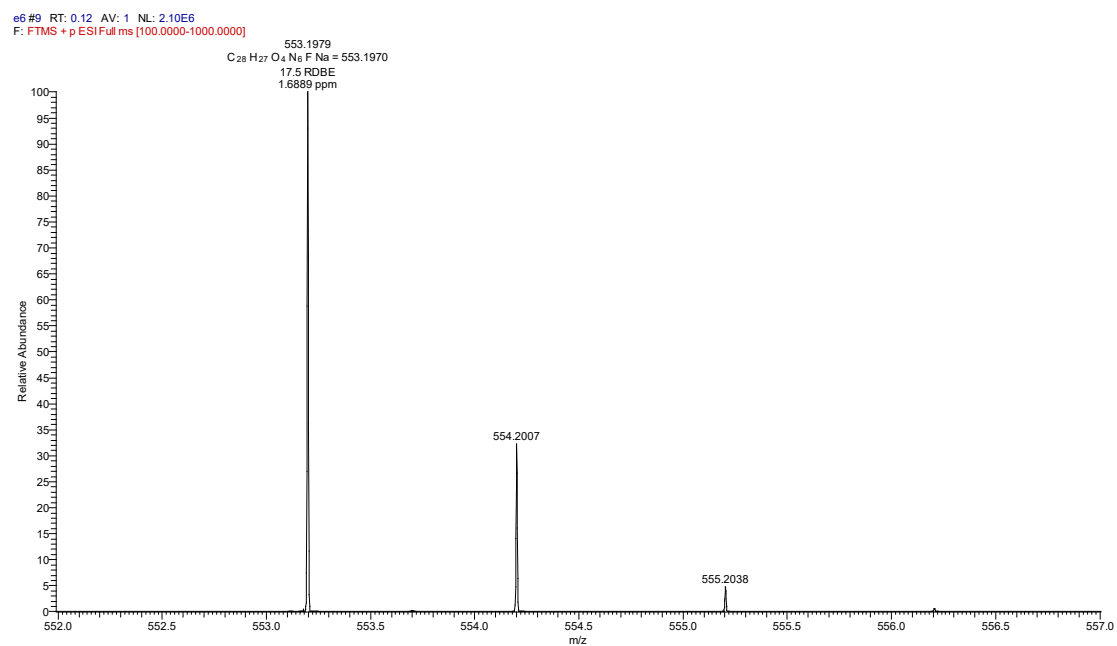


Figure S6-3. HR MS of compound e6



Chemical structure of compound 10: COCCOC1=CC=C2C(=C1)N=CN=C2c3nc(NC4=CC=CC=C4)nn3-c5ccc(F)cc5

¹H NMR spectrum (CDCl₃) of compound 10. The x-axis represents the chemical shift in ppm (f1), ranging from 1.5 to 10.0. The y-axis represents the intensity, ranging from -500 to 5000. The spectrum shows several peaks with corresponding integration values (top) and peak labels (bottom).

Chemical Shift (ppm)	Integration	Label
~9.75	1.00	9.75
~9.35	1.02	9.35
~8.35	0.91	8.35
~8.05	1.04	8.05
~7.95	4.00	7.95
~7.85	1.02	7.85
~7.75	3.03	7.75
~7.65	1.02	7.65
~4.35	4.03	4.35
~4.05	4.05	4.05
~3.85	3.03	3.85
~3.75	3.04	3.75

Figure S7-2. ^{13}C NMR spectrum (150 MHz, DMSO- d_6) of compound e7

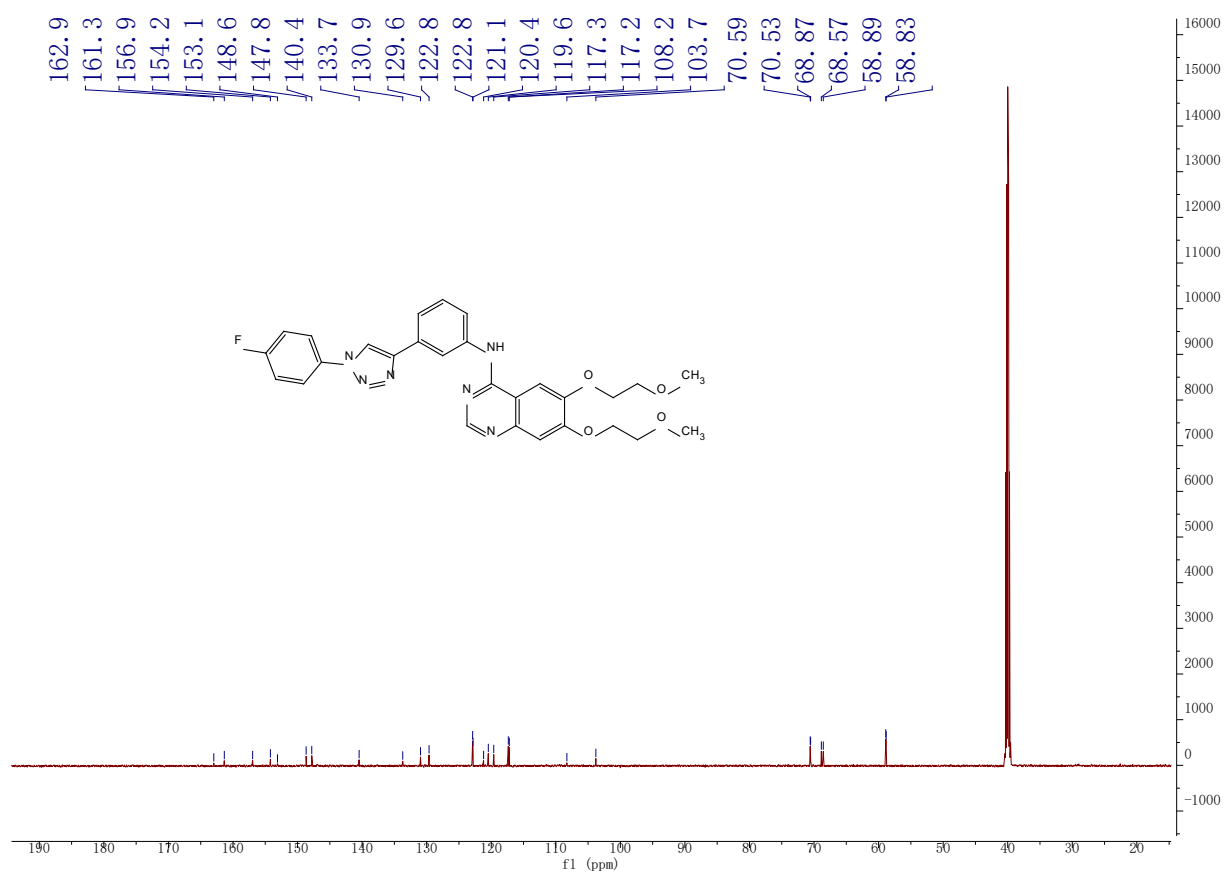


Figure S7-3. HR MS of compound e7

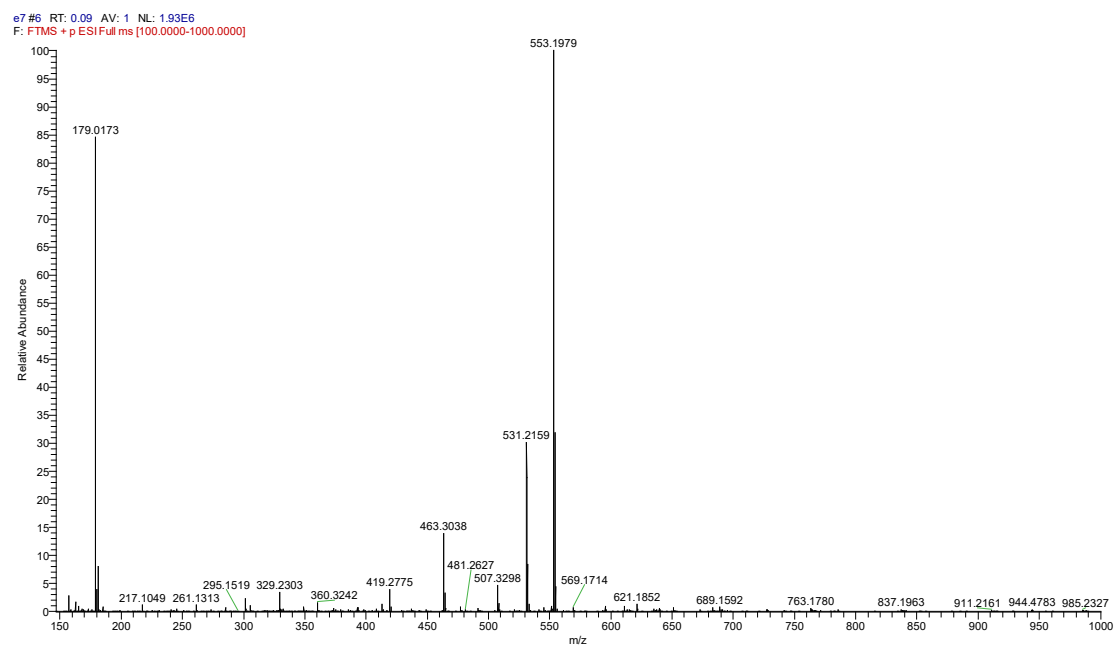


Figure S8-1. ^1H NMR spectrum (600 MHz, DMSO-d_6) of compound e8

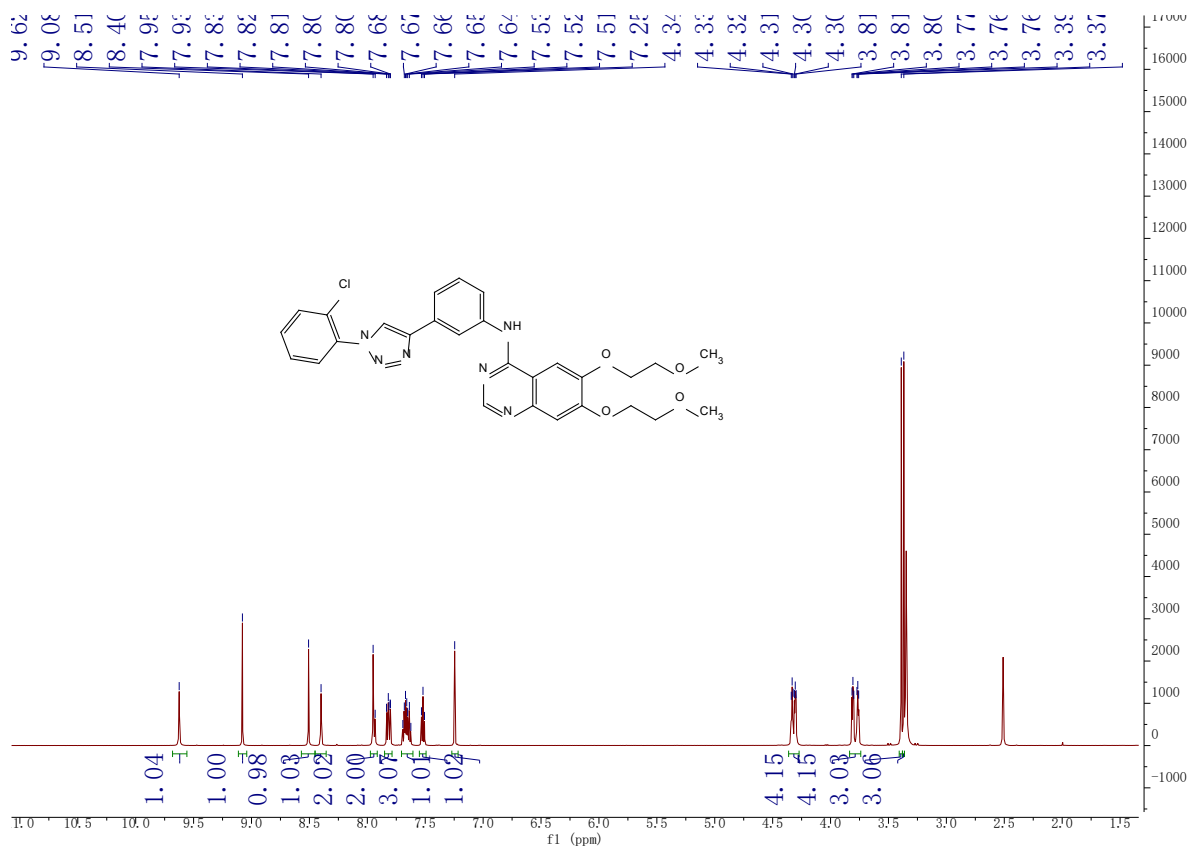


Figure S8-2. ^{13}C NMR spectrum (150 MHz, DMSO-d_6) of compound e8

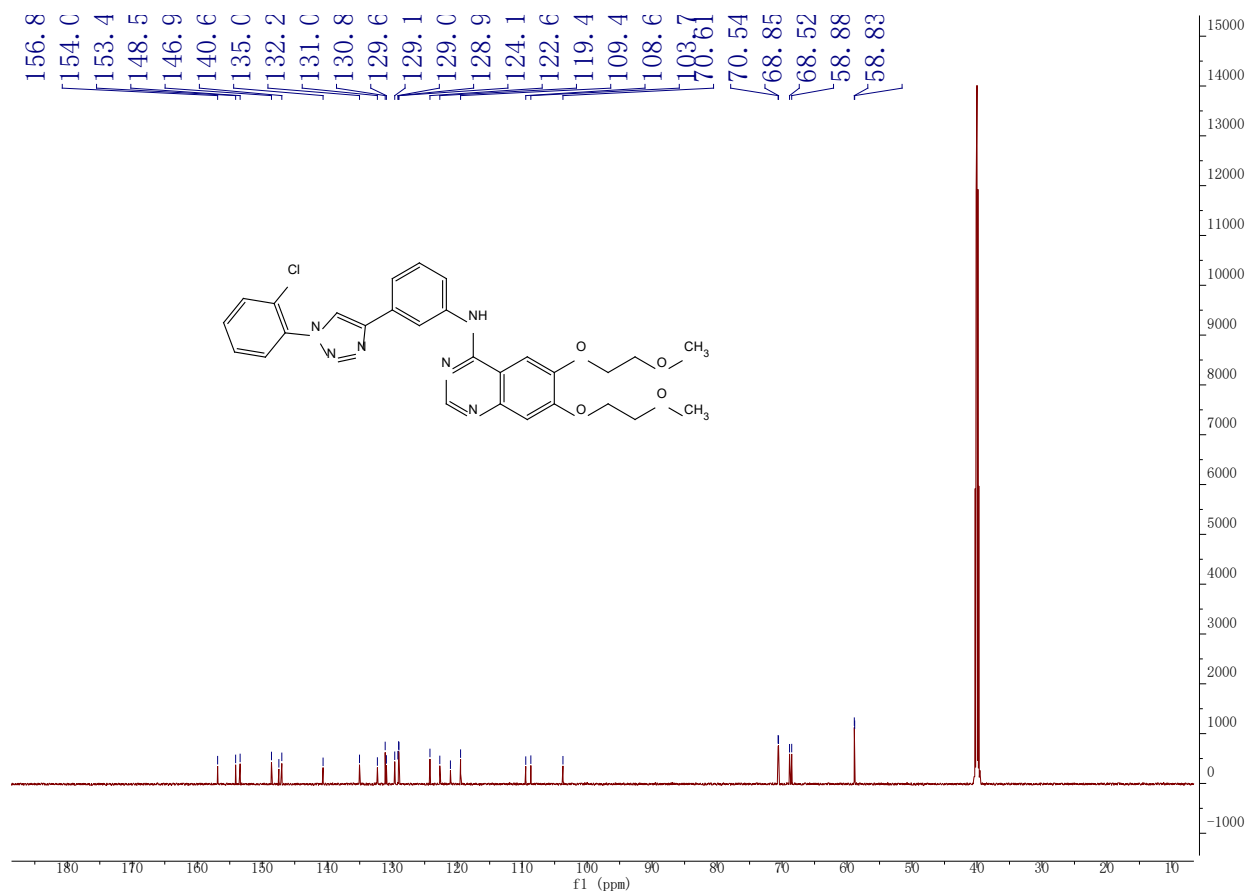


Figure S8-3. HR MS of compound e8

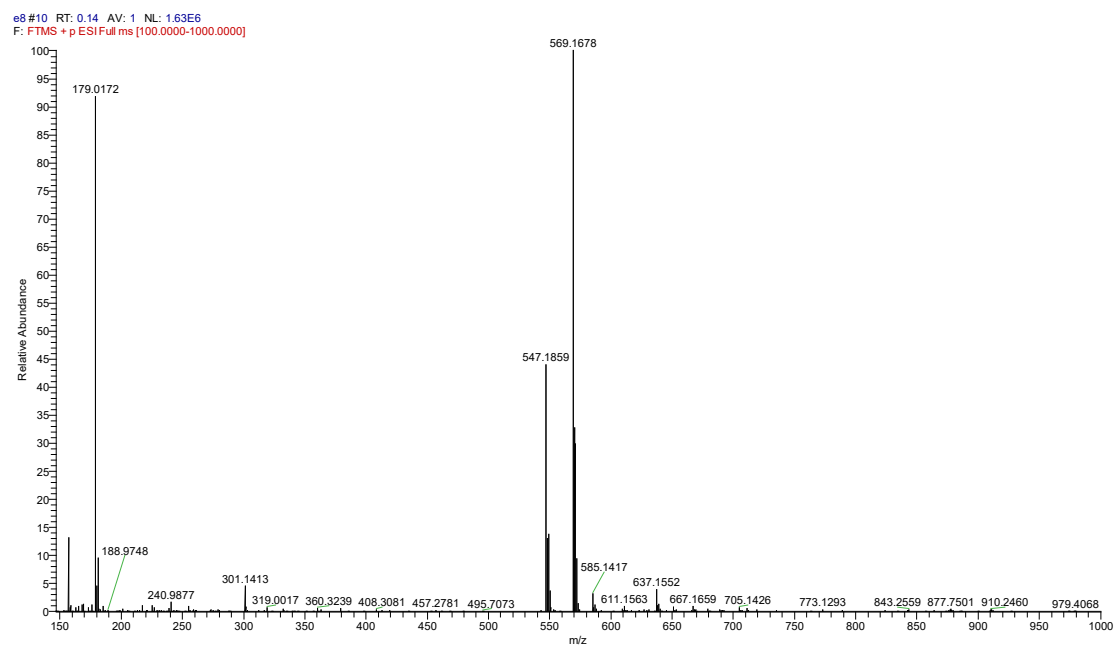


Figure S9-1. ^1H NMR spectrum (600 MHz, DMSO-d_6) of compound e9

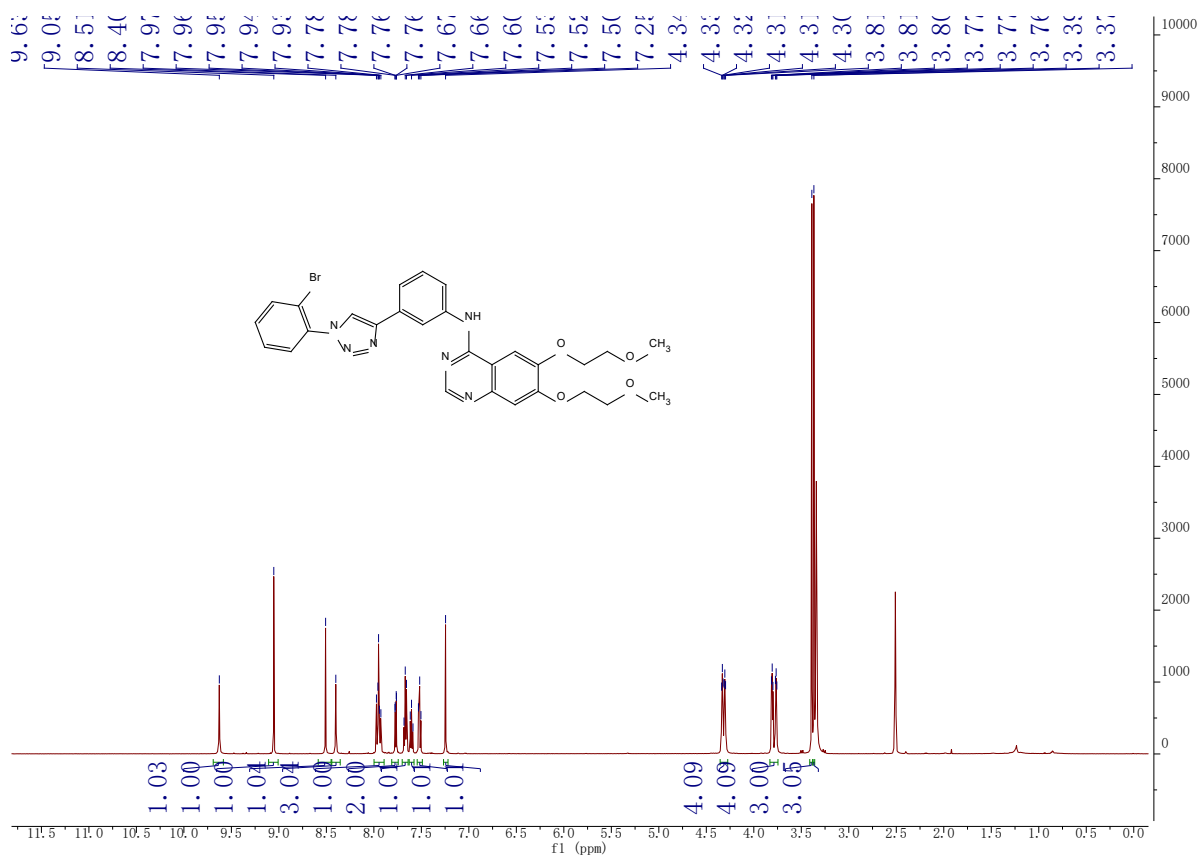


Figure S9-2. ^{13}C NMR spectrum (150 MHz, DMSO-d_6) of compound e9

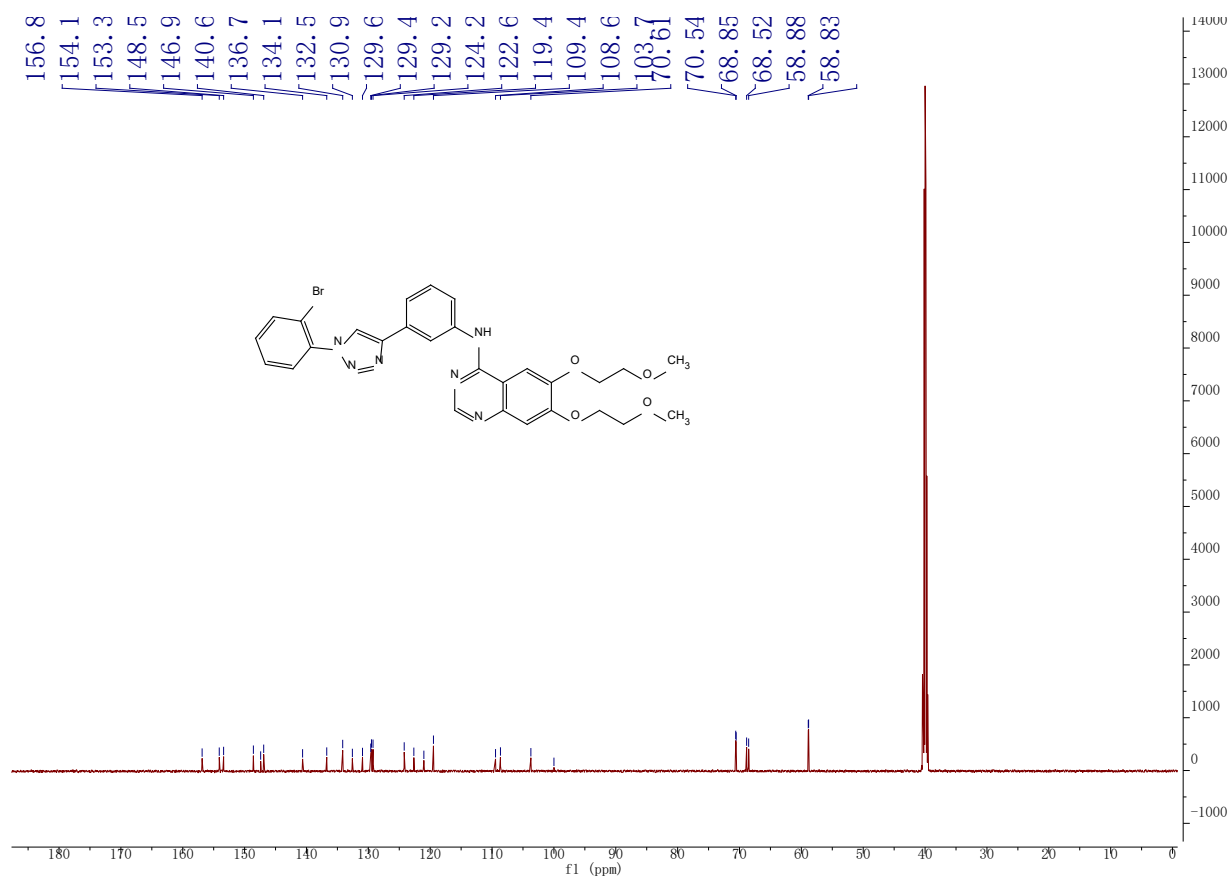


Figure S9-3. HR MS of compound e9

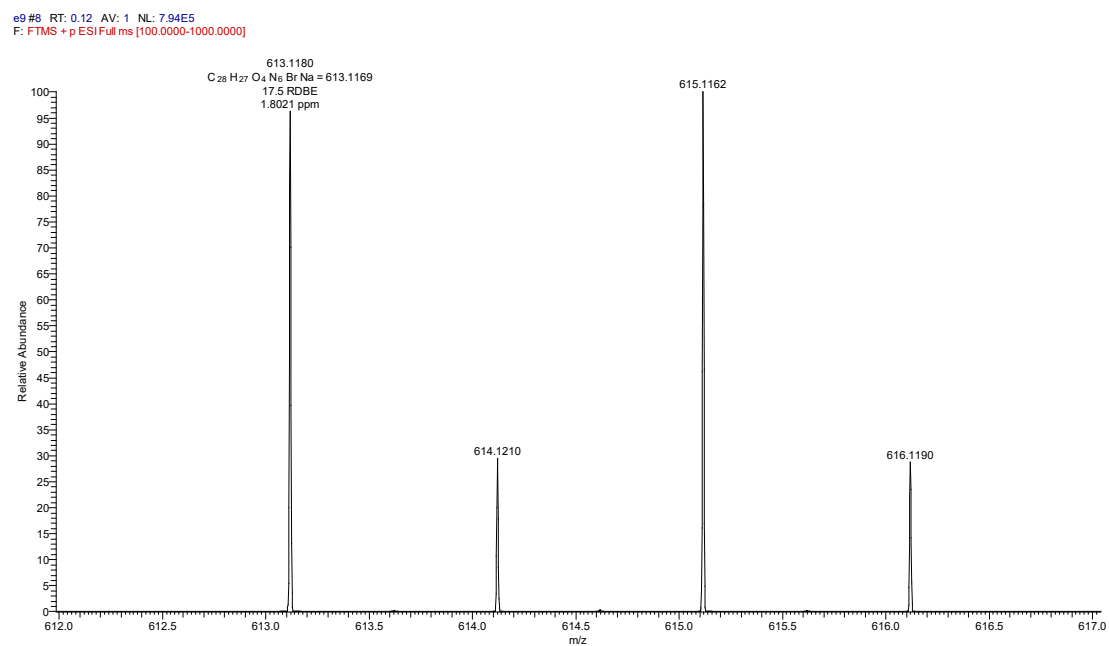


Figure S10-1. ^1H NMR spectrum (600 MHz, DMSO- d_6) of compound e10



Figure S10-2. ^{13}C NMR spectrum (150 MHz, DMSO- d_6) of compound e10

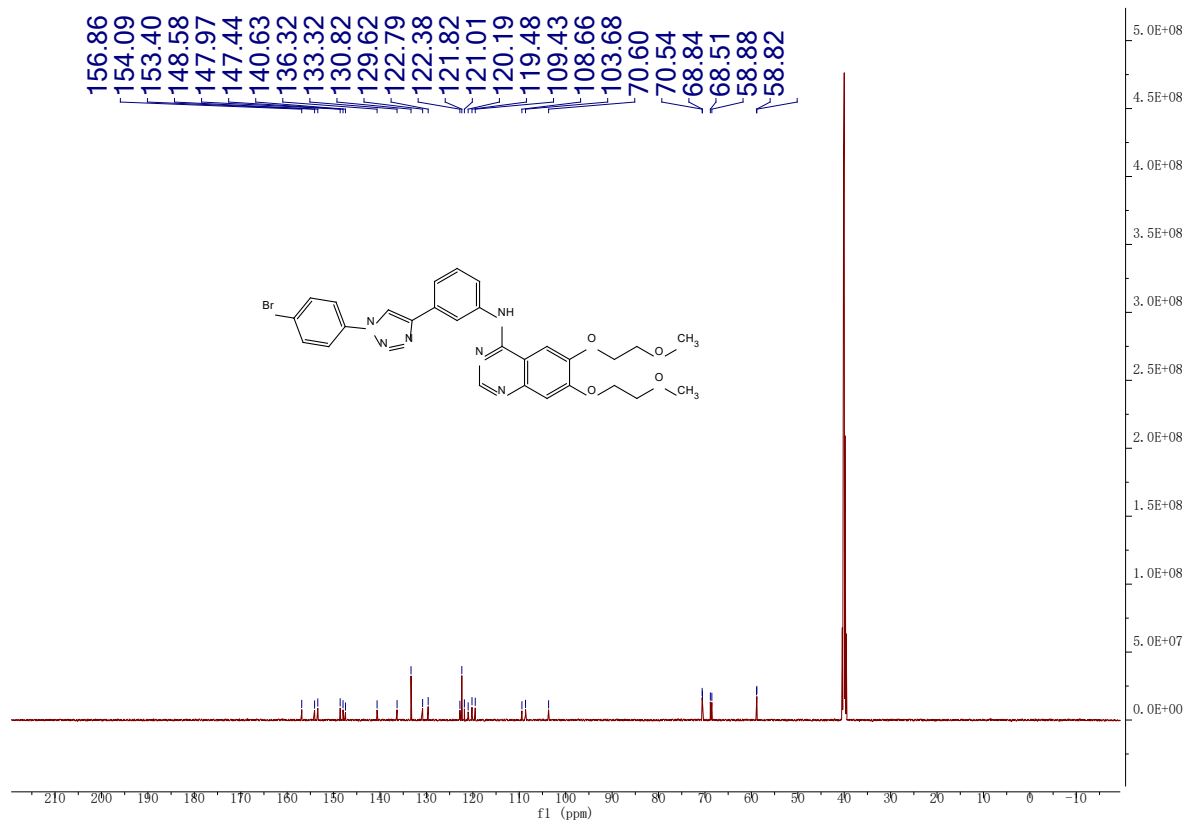


Figure S10-3. HR MS of compound e10

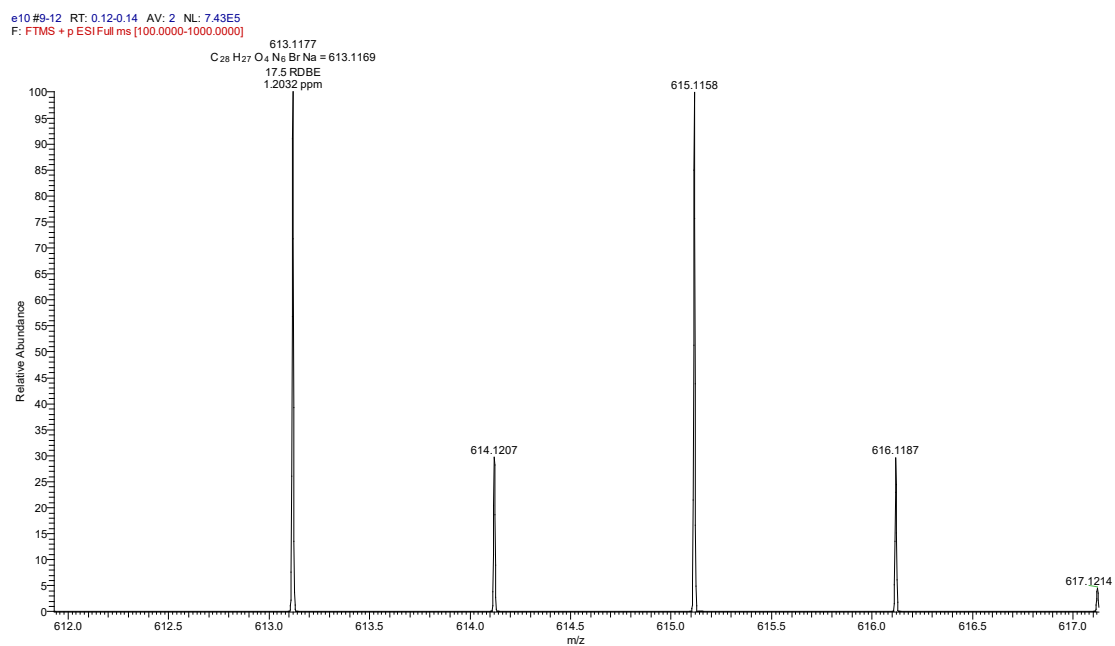


Figure S11-1. ^1H NMR spectrum (600 MHz, DMSO- d_6) of compound e11

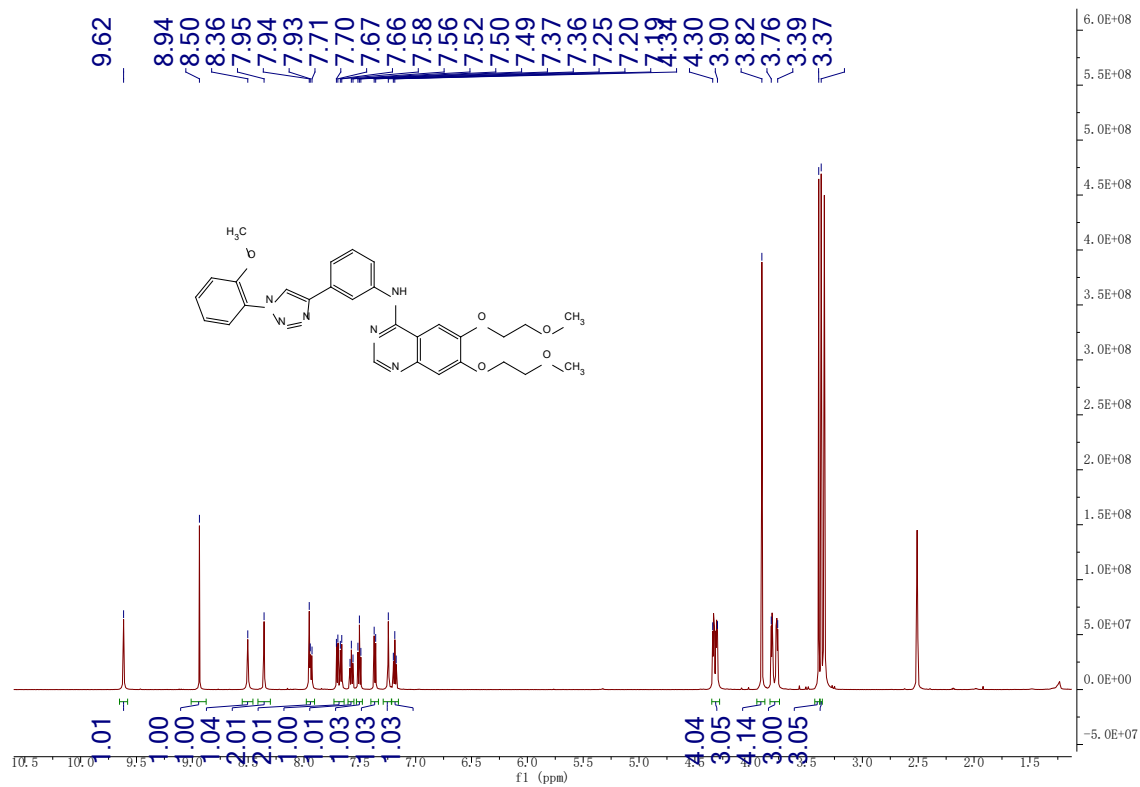


Figure S11-2. ^{13}C NMR spectrum (150 MHz, DMSO- d_6) of compound e11

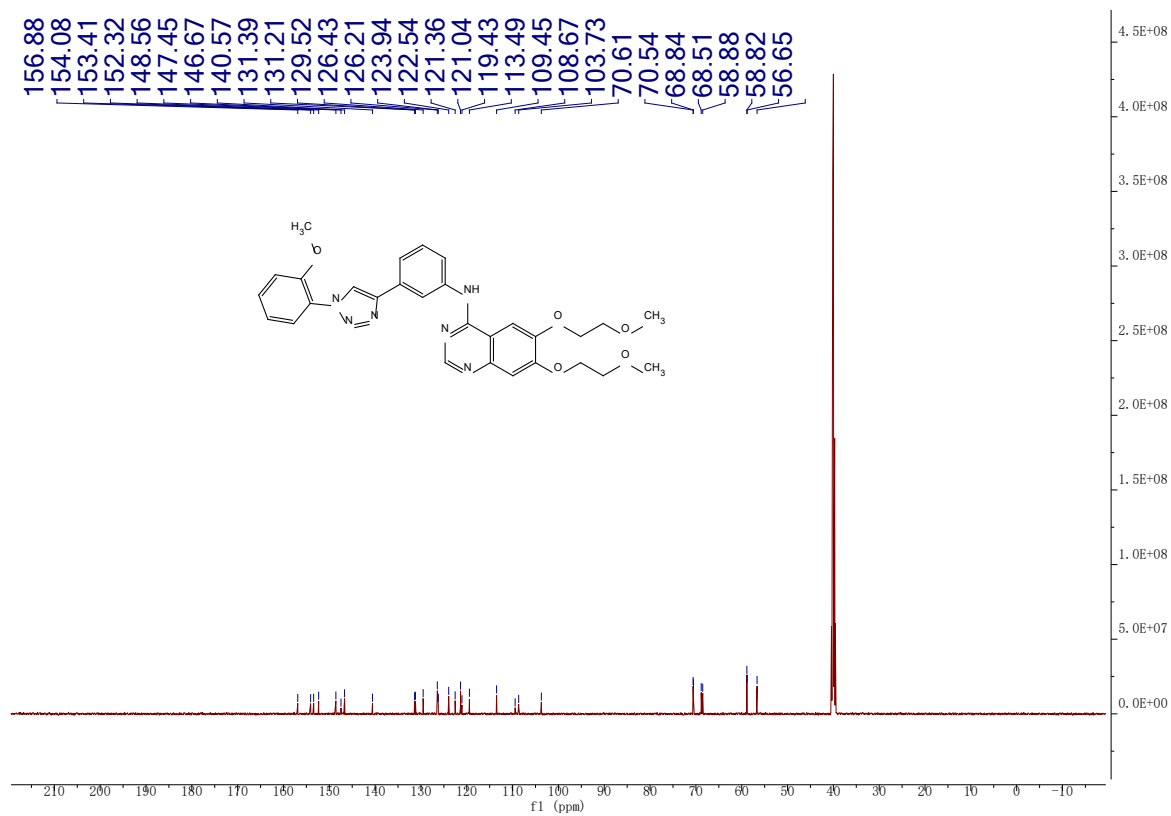
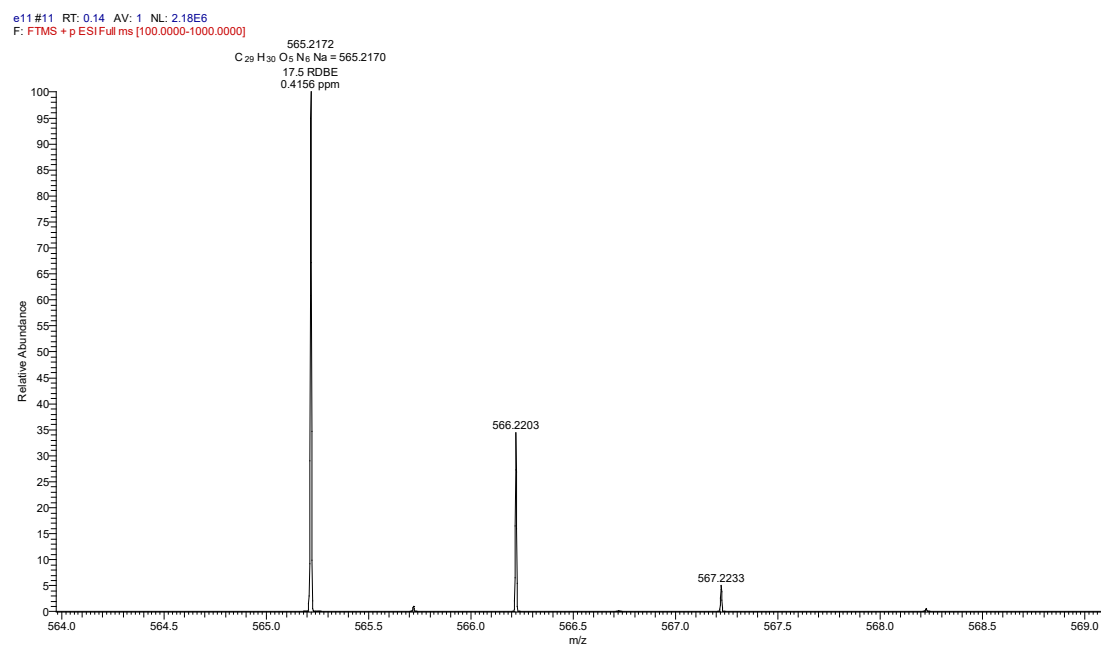


Figure S11-3. HR MS of compound e11



Chemical structure of compound 10 is shown above the spectrum. The structure is a benzimidazole derivative with a 4-methylphenyl group at position 2, a 4-aminophenyl group at position 4, and a 2,2-bis(methoxy)ethyl group at position 6.

¹H NMR spectrum (CDCl₃) of compound 10. The x-axis represents the chemical shift in ppm (f1), ranging from 10.5 to 1.0. The y-axis represents the intensity, ranging from -2.0E+07 to 3.8E+08. The spectrum shows several peaks, with the following chemical shifts (ppm) and integration values (area) labeled below the peaks:

Chemical Shift (ppm)	Integration (Area)
9.63	0.01
9.28	1.00
8.51	0.93
8.37	1.01
7.95	2.00
7.94	2.01
7.88	1.00
7.86	1.00
7.67	1.01
7.66	1.02
7.54	0.01
7.52	0.05
7.51	0.06
7.45	0.09
7.44	3.00
7.25	0.00
4.34	0.01
4.30	0.05
3.82	0.06
3.76	0.09
3.39	3.00
3.37	0.00
2.51	0.00

Figure S12-2. ^{12}C NMR spectrum (150 MHz, DMSO-d_6) of compound e12

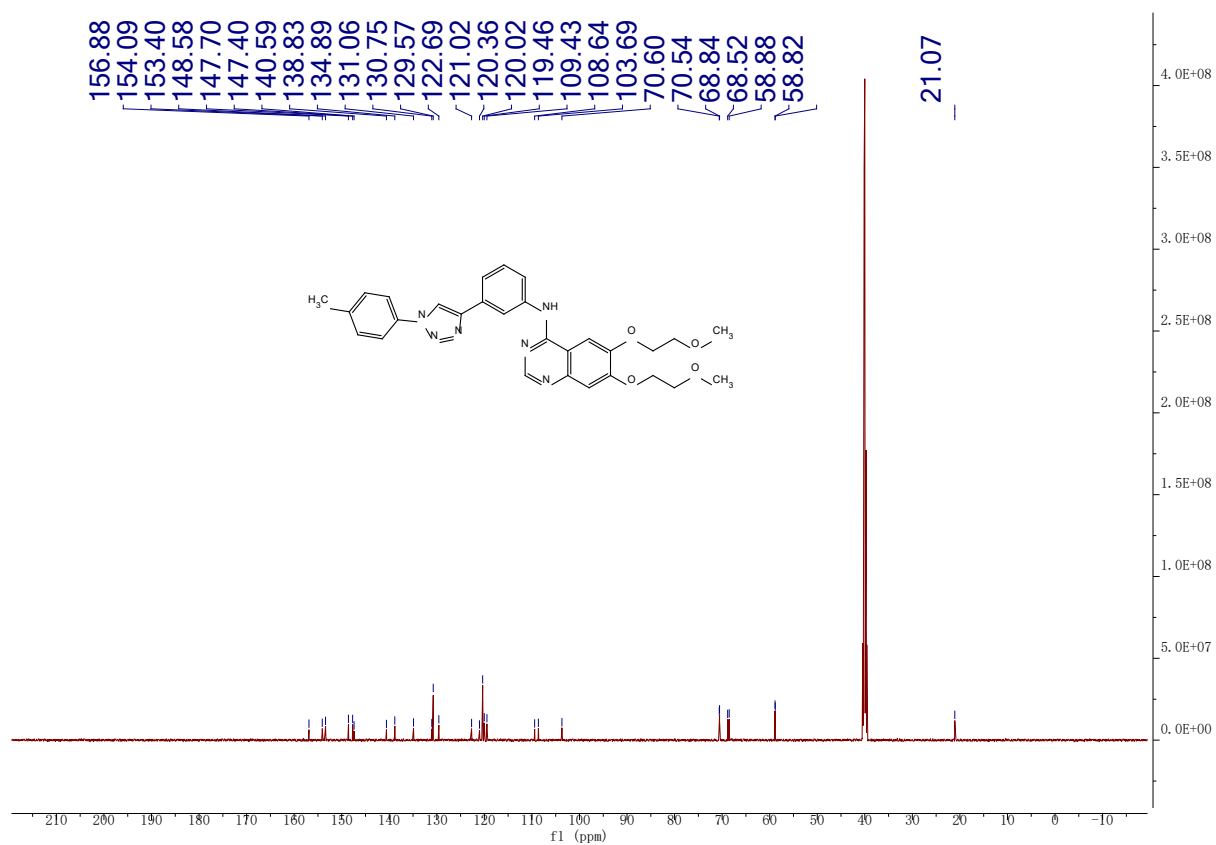


Figure S12-3. HR MS of compound e12

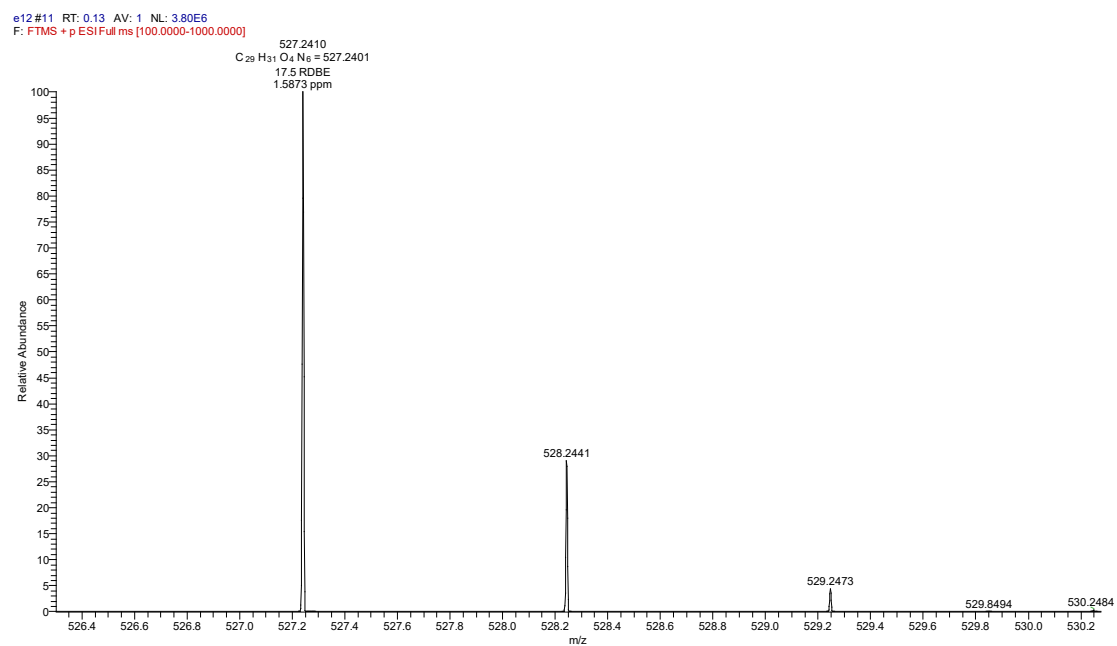


Figure S13-1. ^1H NMR spectrum (600 MHz, DMSO- d_6) of compound a13

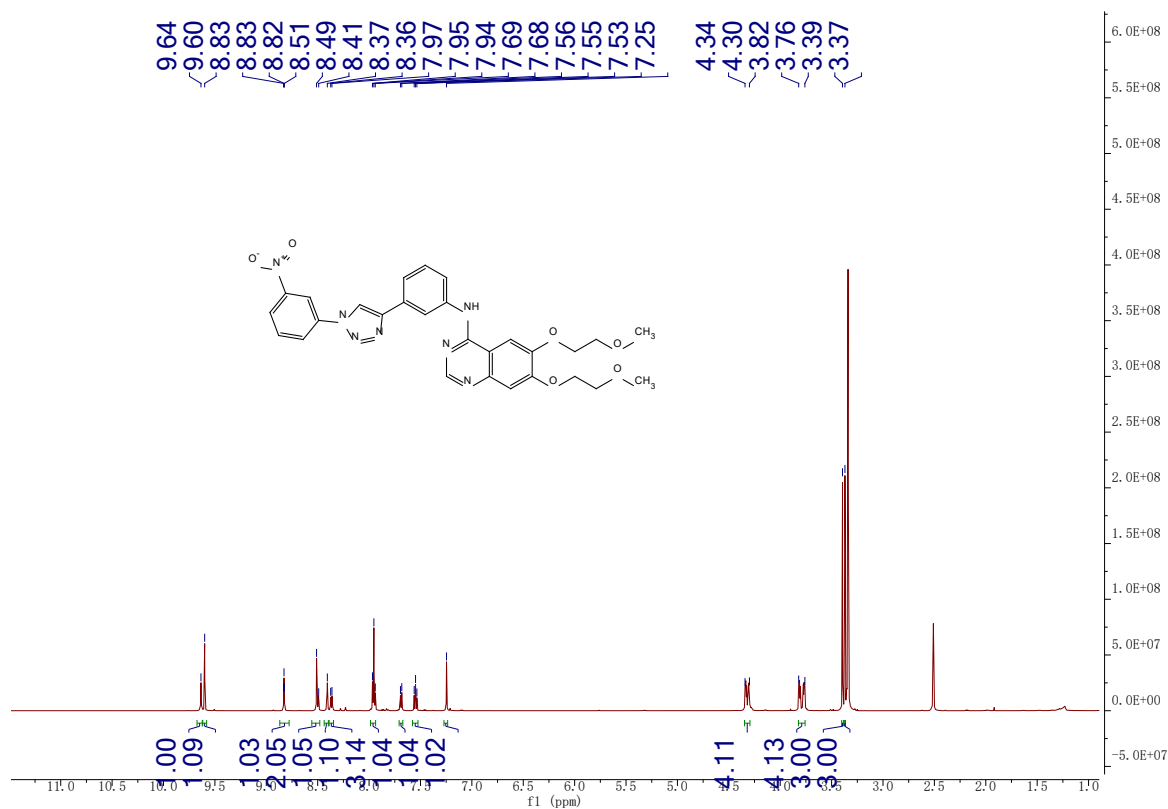


Figure S13-2. ^{13}C NMR spectrum (150 MHz, DMSO-d_6) of compound e13

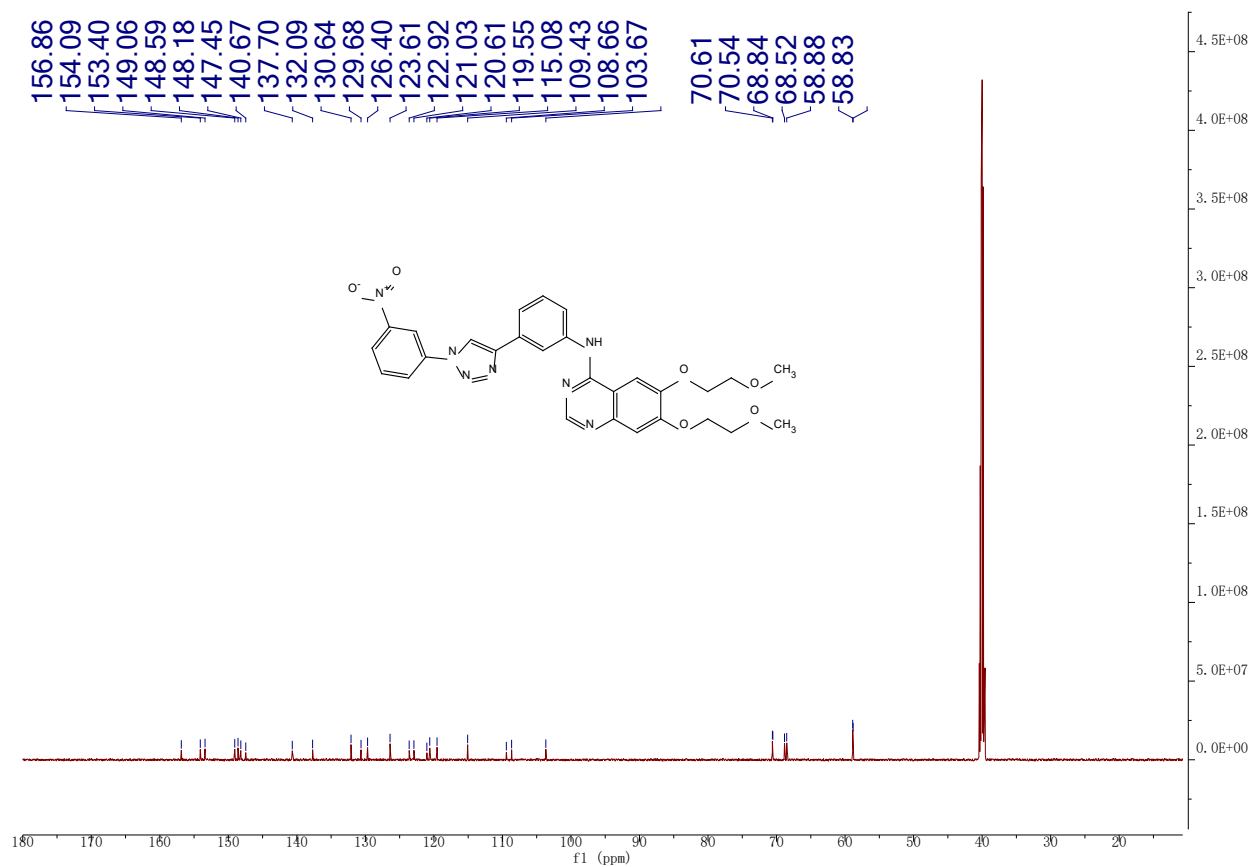


Figure S13-3. HR MS of compound e13

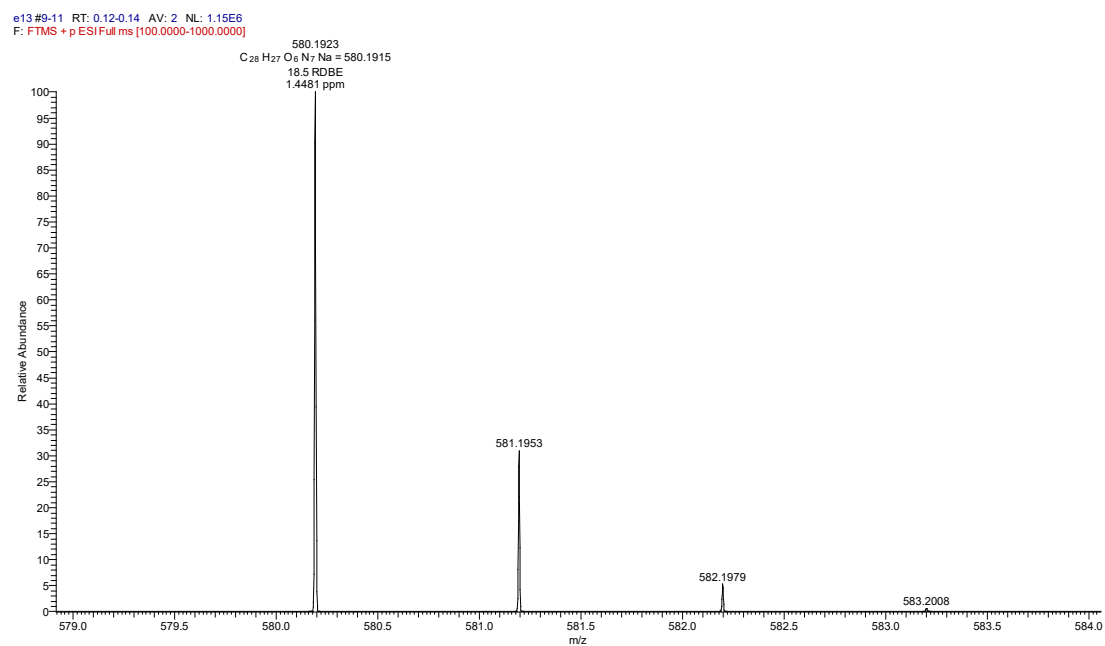


Figure S14-1. ^1H NMR spectrum (600 MHz, DMSO- d_6) of compound e14

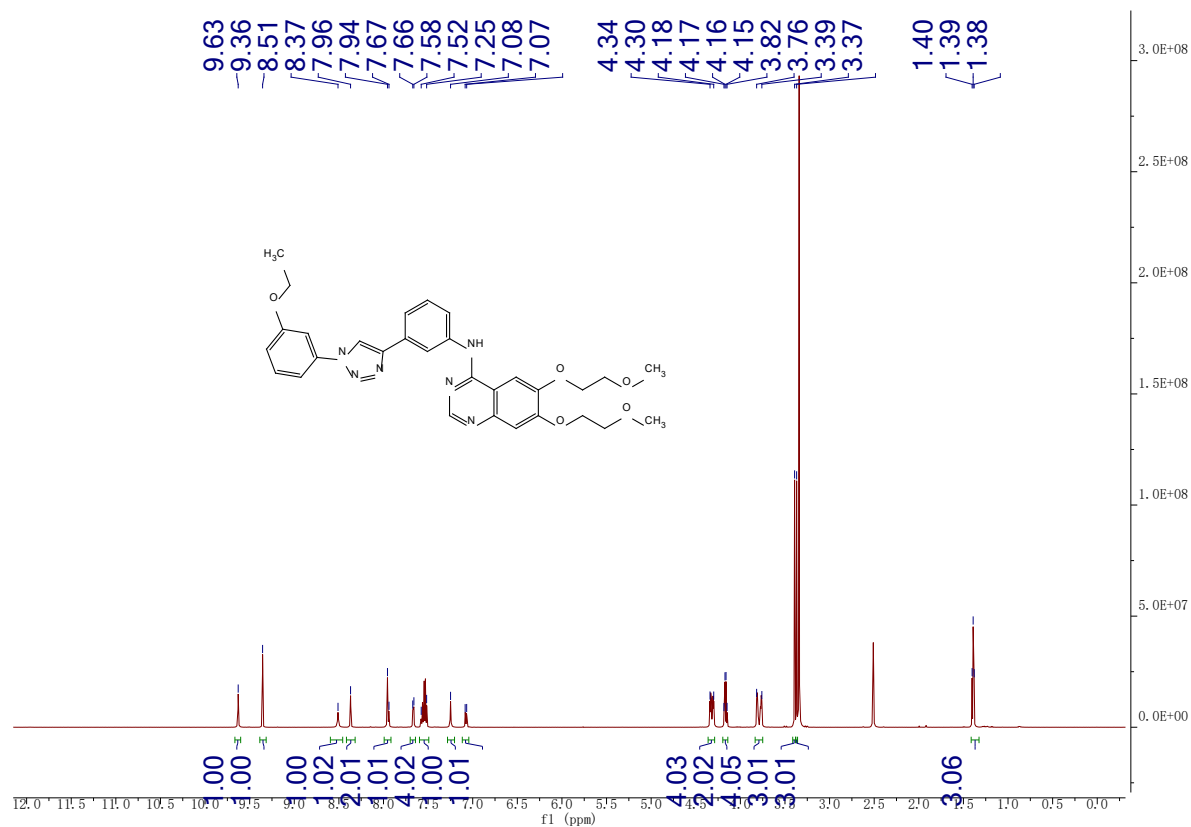


Figure S14-2. ^{13}C NMR spectrum (150 MHz, DMSO- d_6) of compound e14

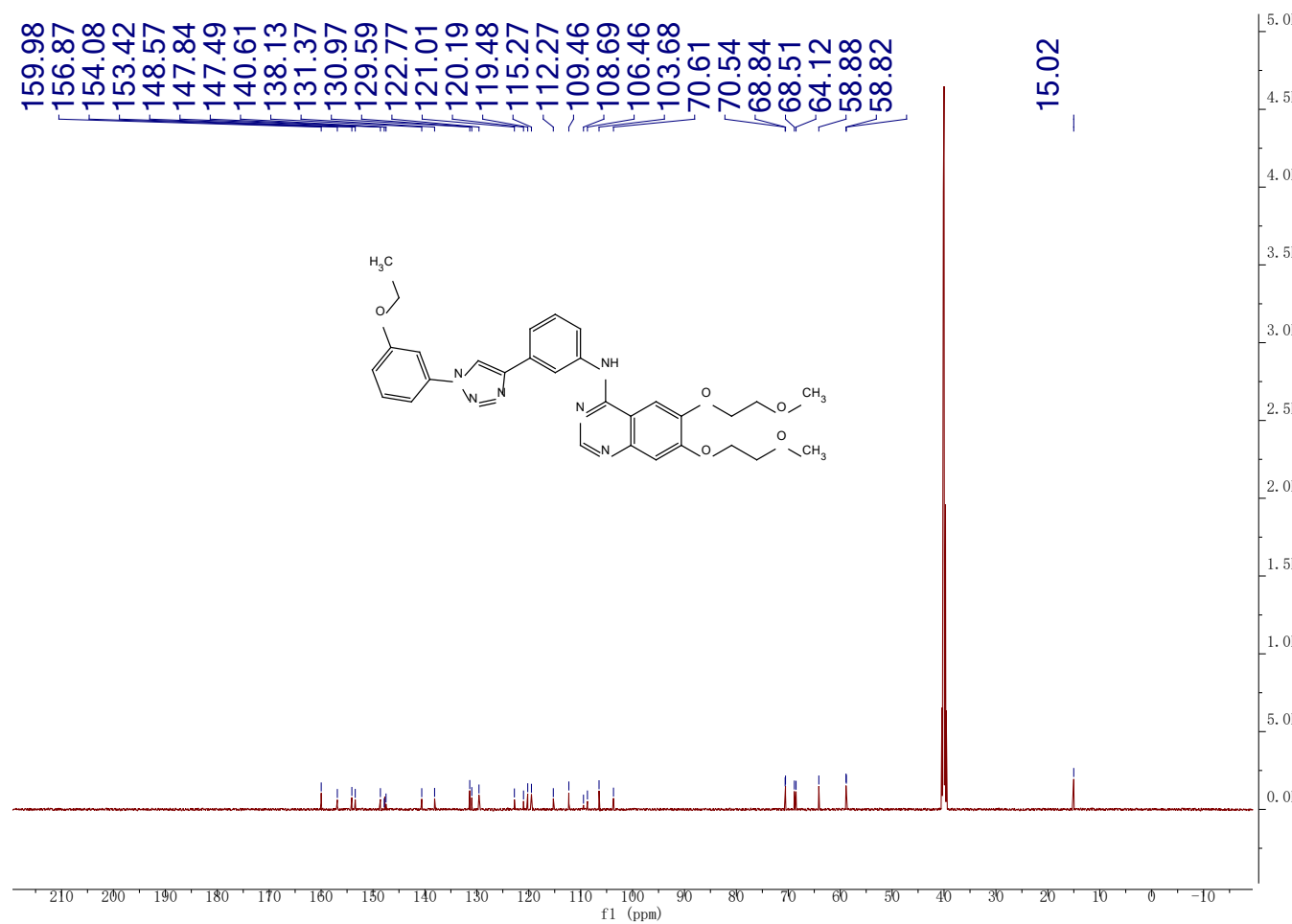


Figure S14-3. HR MS of compound e14

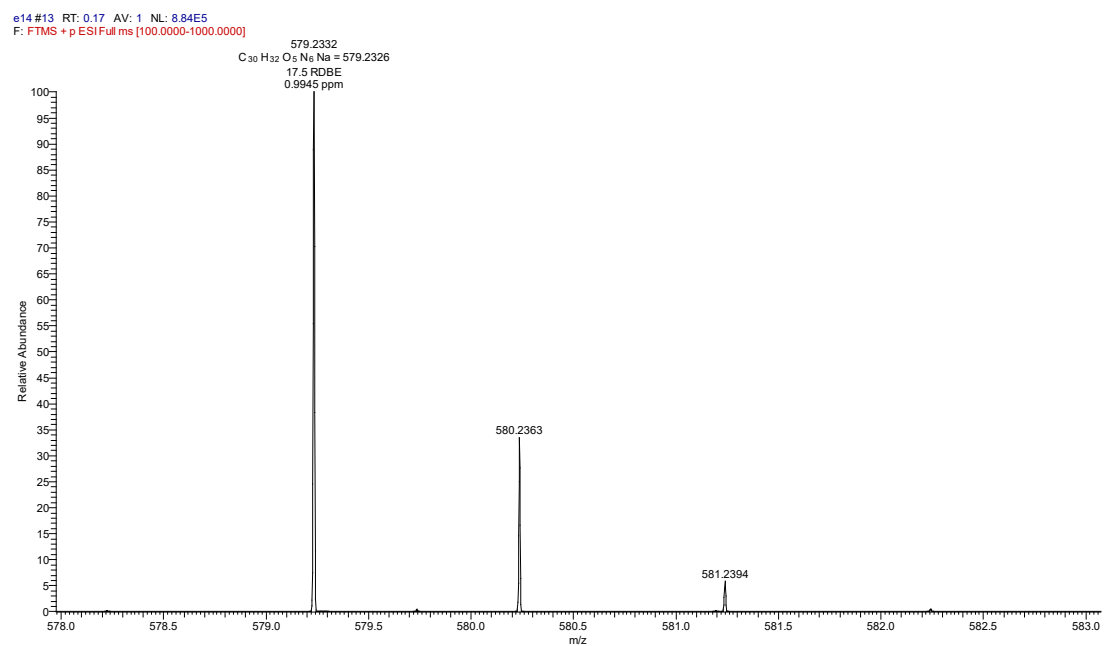


Figure S15-1. ^1H NMR spectrum (600 MHz, DMSO- d_6) of compound e15

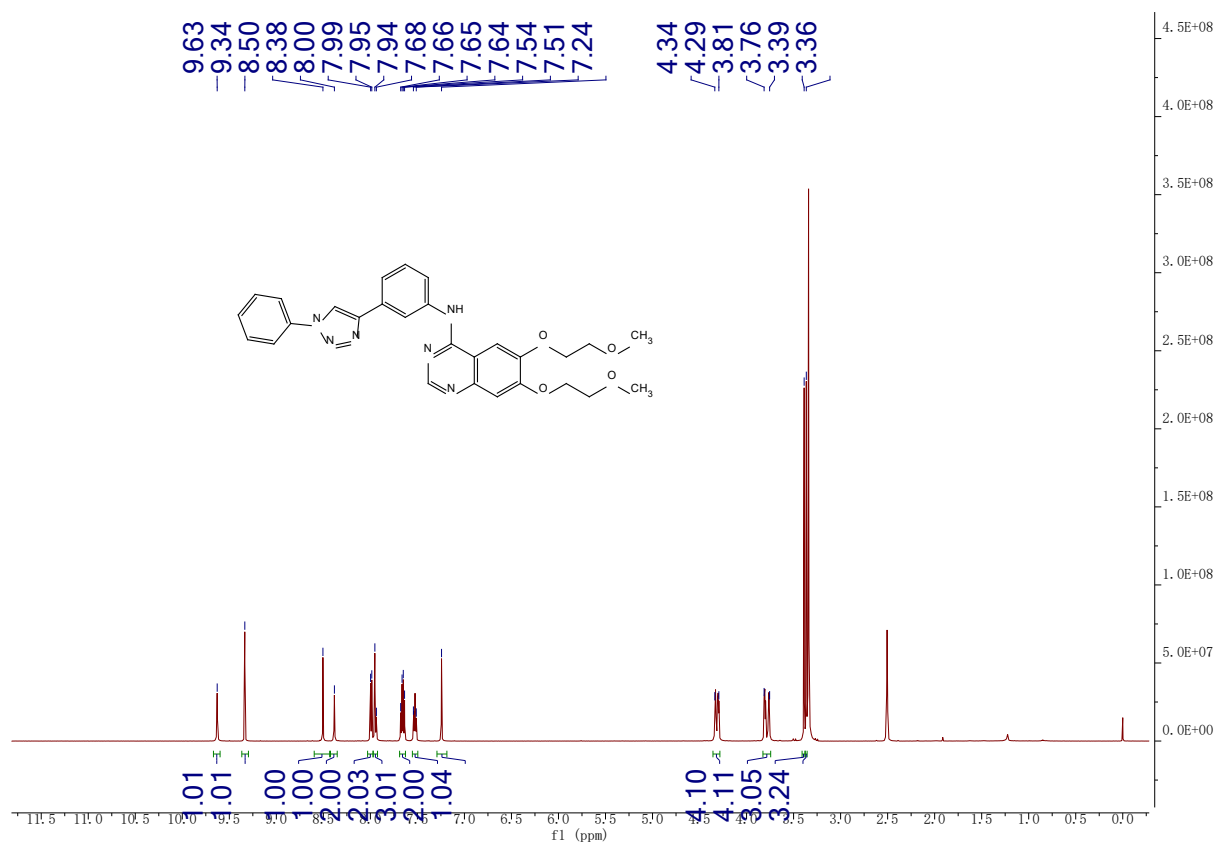


Figure S15-2. ^{13}C NMR spectrum (150 MHz, DMSO- d_6) of compound e15

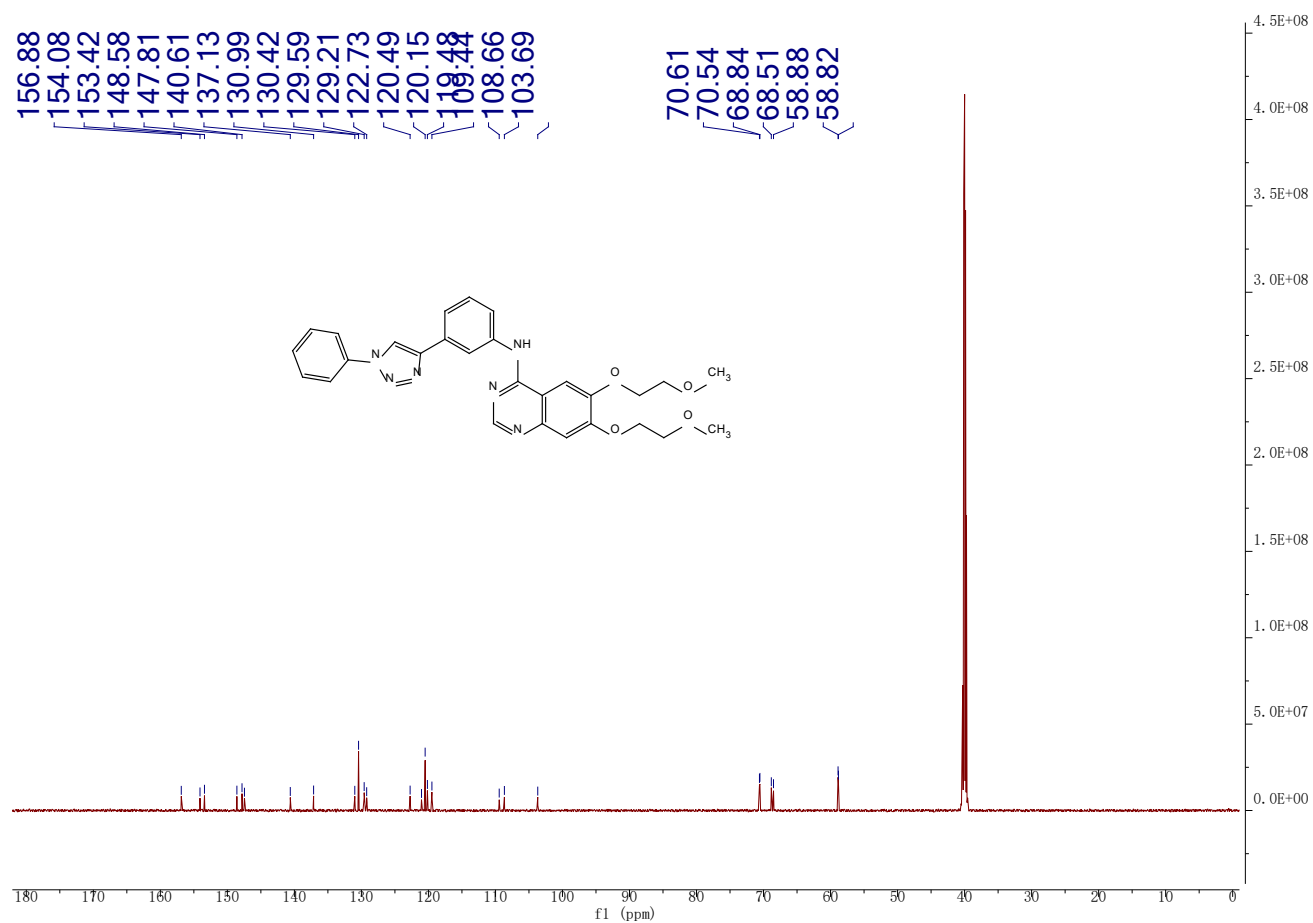


Figure S15-3. HR MS of compound e15

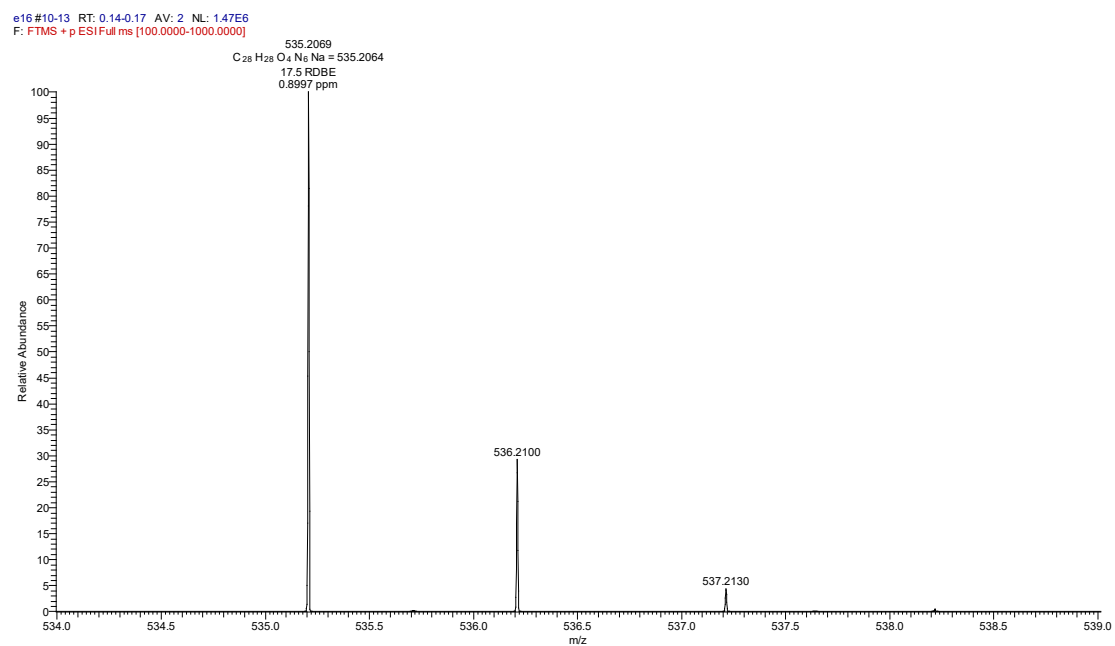


Figure S16-1. ¹H NMR spectrum (400 MHz, DMSO-d₆) of compound e16

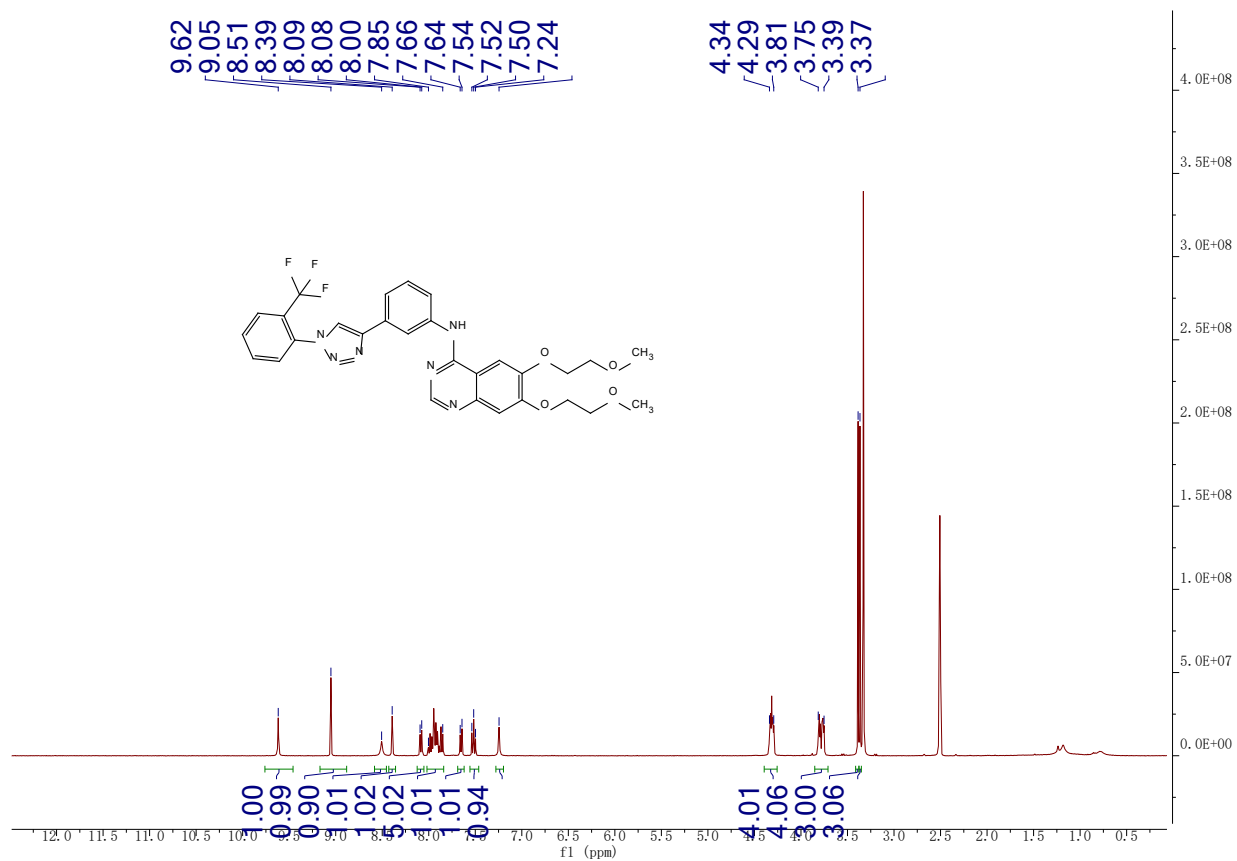


Figure S16-2. ^{13}C NMR spectrum (100 MHz, DMSO-d_6) of compound e16

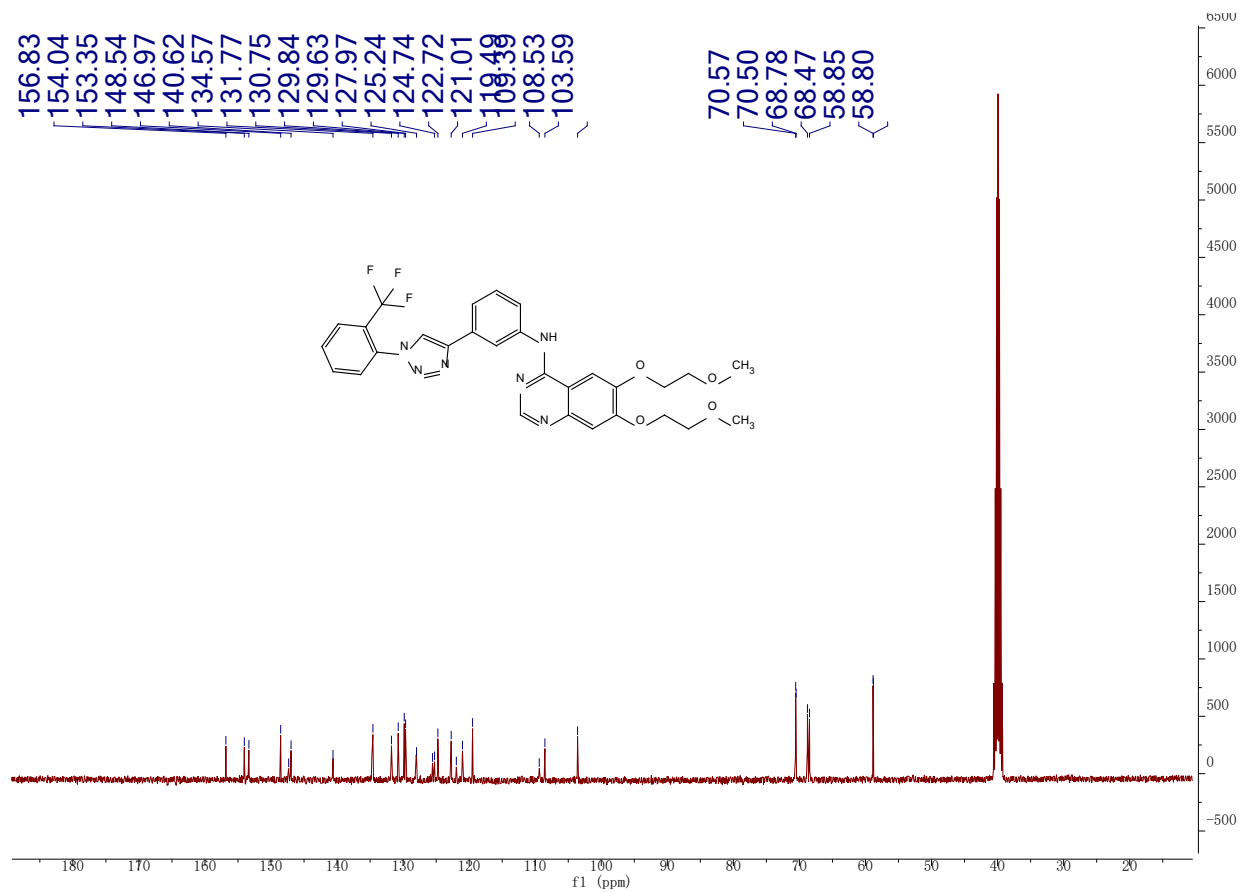
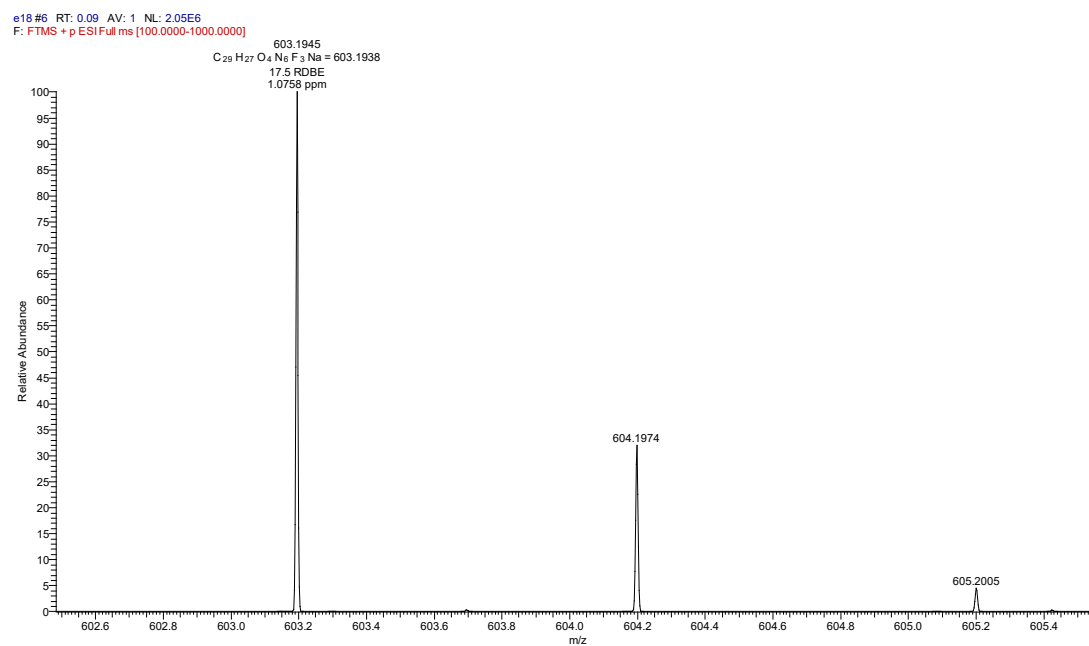


Figure S16-3. HR MS of compound e16



Chemical structure of the compound is shown above the spectrum. The structure is a complex molecule featuring a quinoline core, a pyrazole ring, and a benzimidazole moiety, with various substituents including methoxy groups and a dimethoxyethyl ether.

¹H NMR spectrum (CDCl₃) showing peaks from 1.0 to 10.0 ppm. The x-axis is labeled f1 (ppm) and the y-axis is labeled intensity. The spectrum displays several sharp peaks, with integration values provided below the baseline. The chemical structure is shown above the spectrum.

Chemical structure of the compound is shown above the spectrum. The structure is a complex molecule featuring a quinoline core, a pyrazole ring, and a benzimidazole moiety, with various substituents including methoxy groups and a dimethoxyethyl ether.

¹H NMR spectrum (CDCl₃) showing peaks from 1.0 to 10.0 ppm. The x-axis is labeled f1 (ppm) and the y-axis is labeled intensity. The spectrum displays several sharp peaks, with integration values provided below the baseline. The chemical structure is shown above the spectrum.

Figure S17-2. ^{13}C NMR spectrum (100 MHz, DMSO-d_6) of compound e17

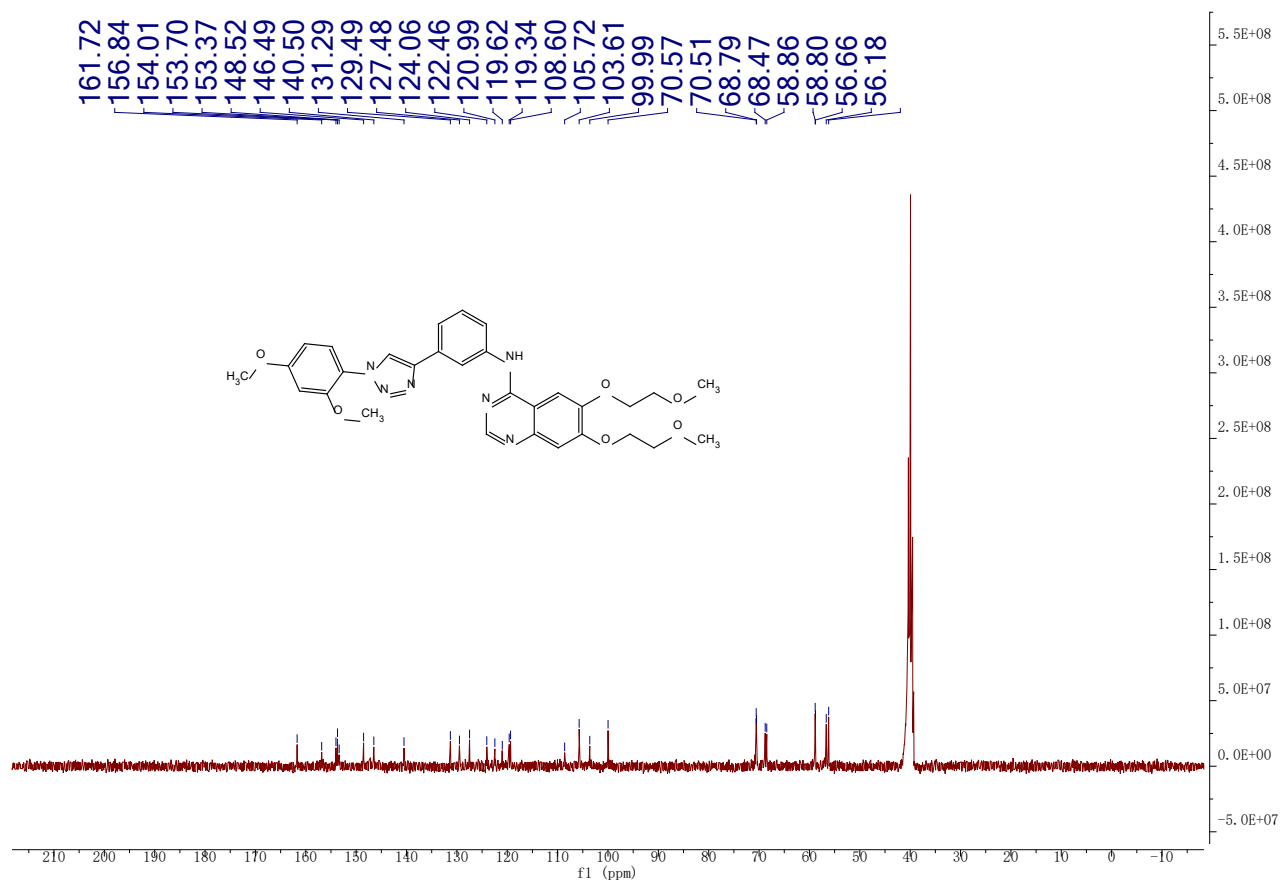


Figure S17-3. HR MS of compound e17

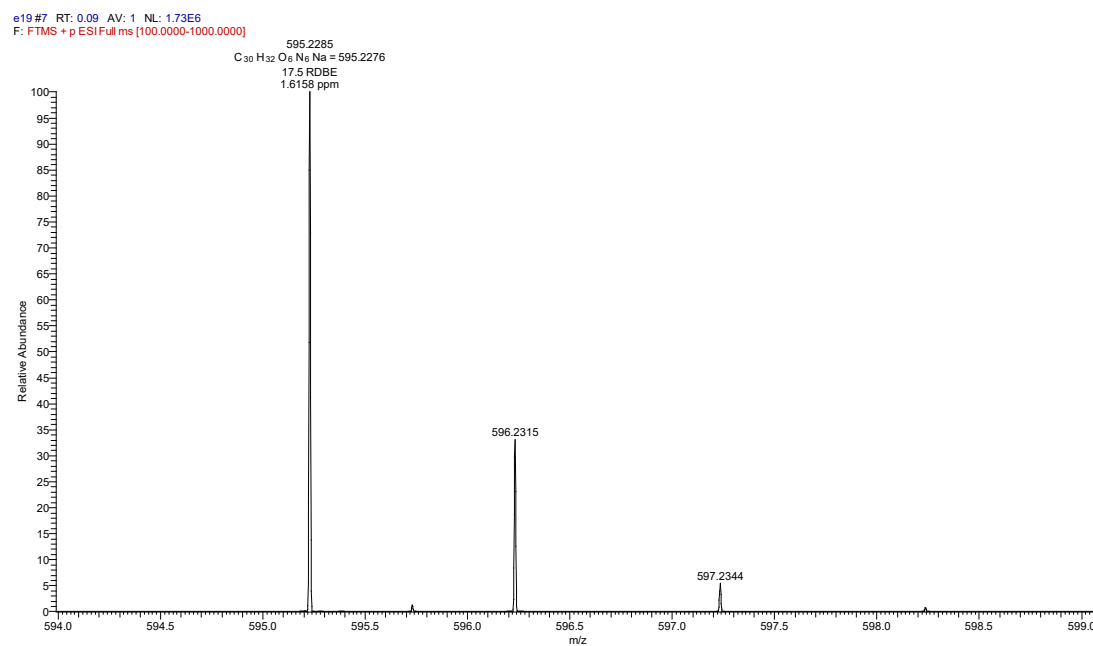


Figure S18-1. ^1H NMR spectrum (400 MHz, DMSO- d_6) of compound e18

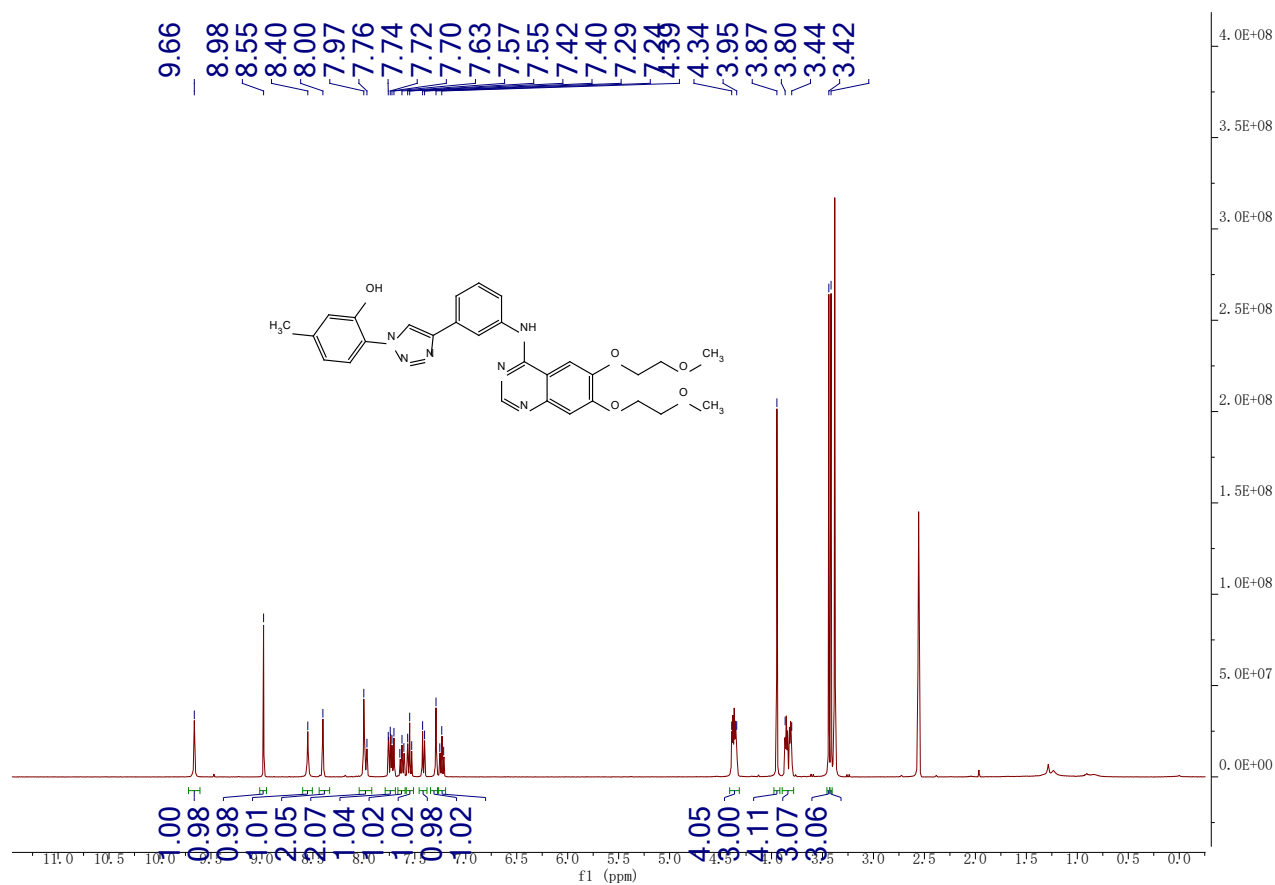


Figure S18-2. ^{13}C NMR spectrum (100 MHz, DMSO-d_6) of compound e18

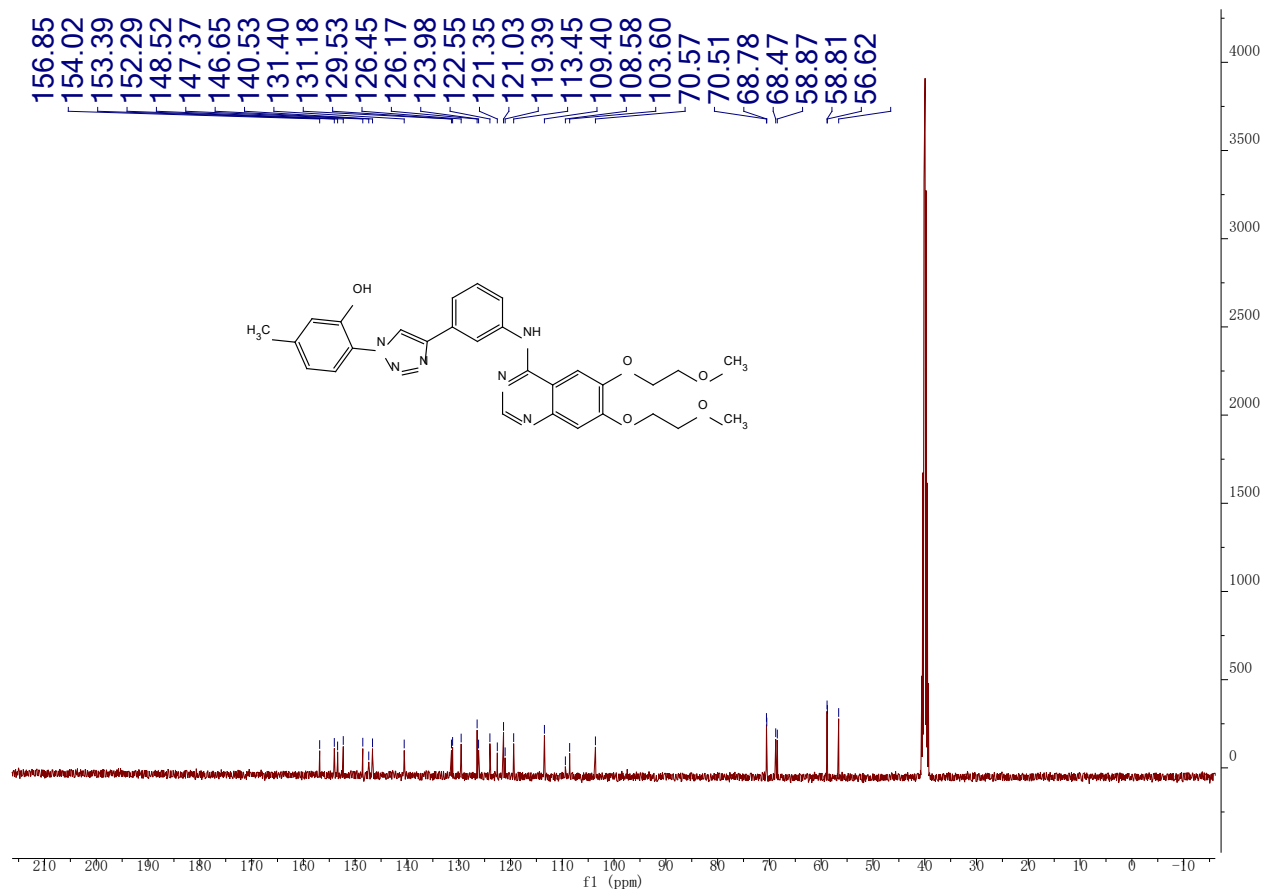


Figure S18-3. HR MS of compound e18

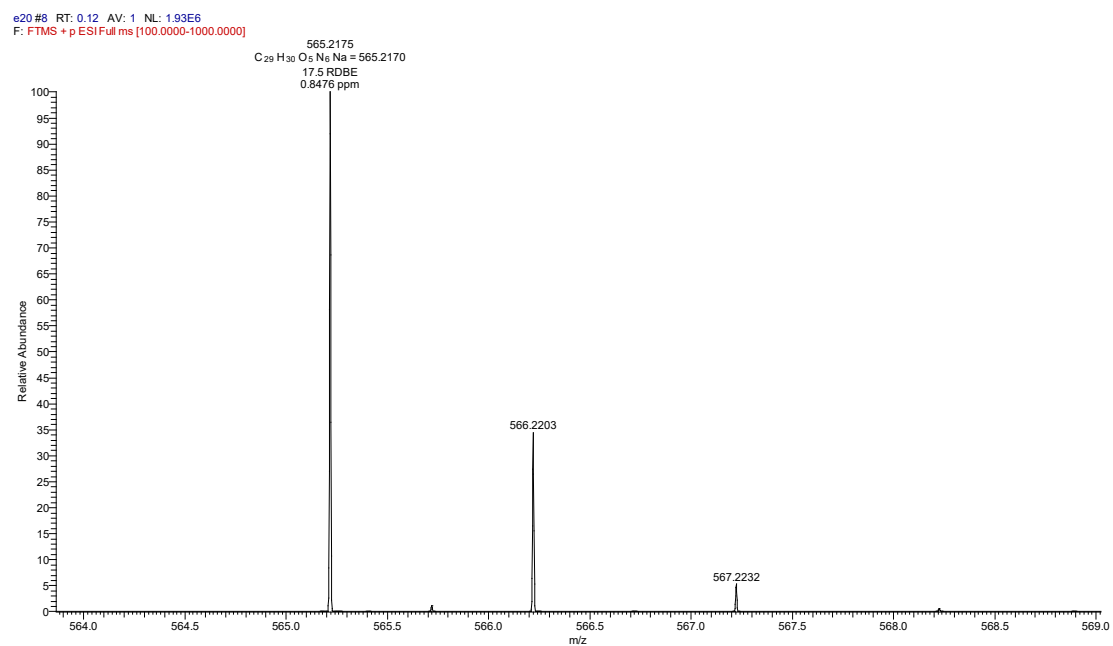


Figure S19-1. ^1H NMR spectrum (400 MHz, DMSO- d_6) of compound e19

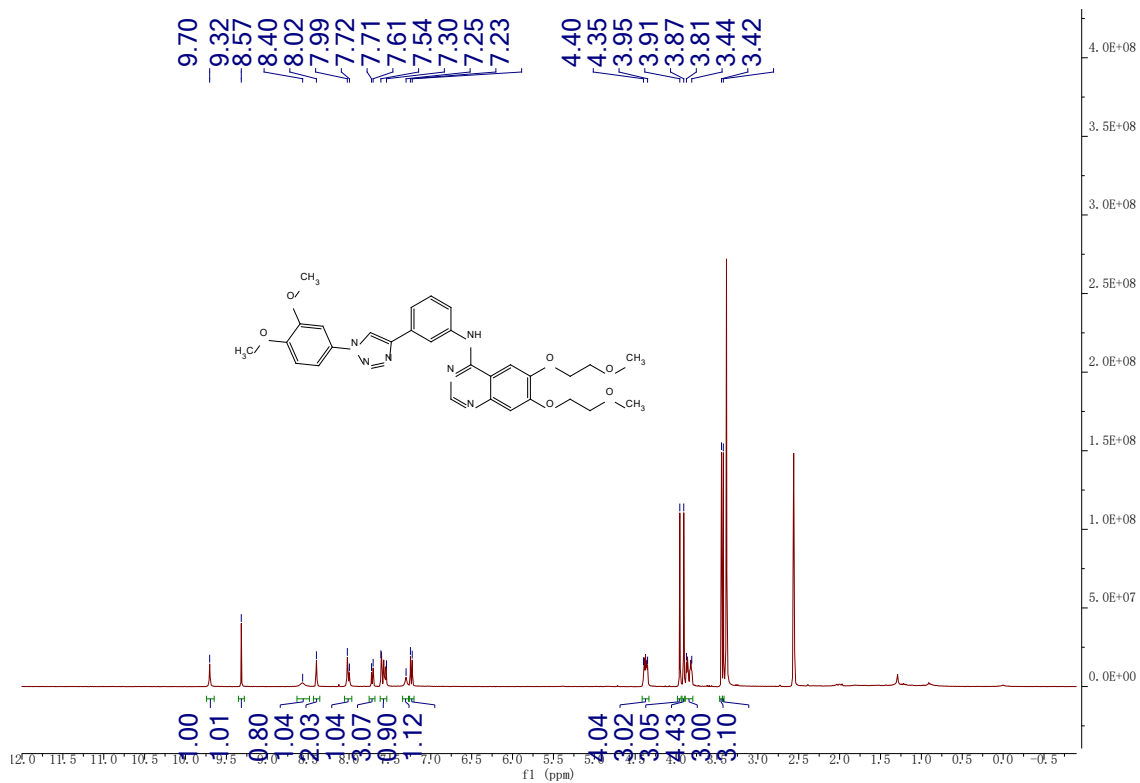


Figure S19-2. ^{13}C NMR spectrum (100 MHz, DMSO-d_6) of compound e19

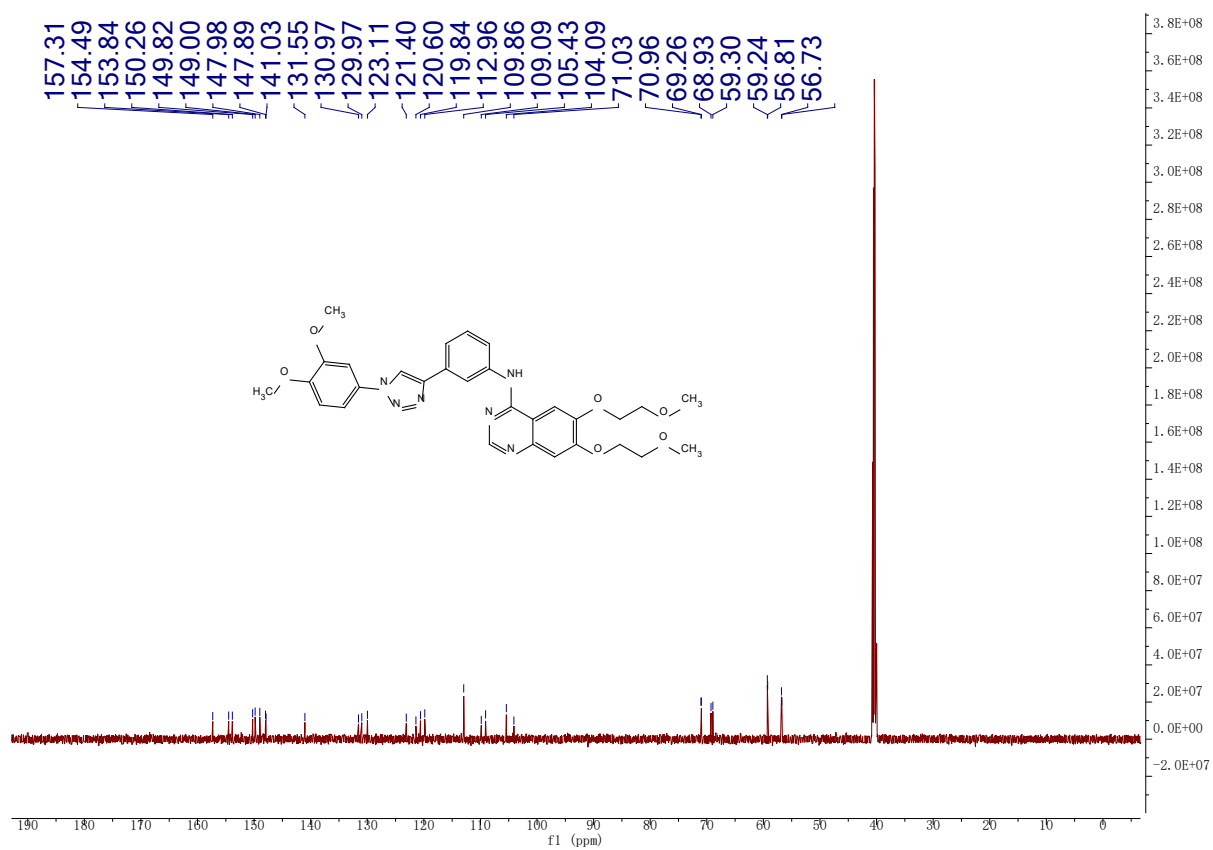


Figure S19-3. HR MS of compound e19

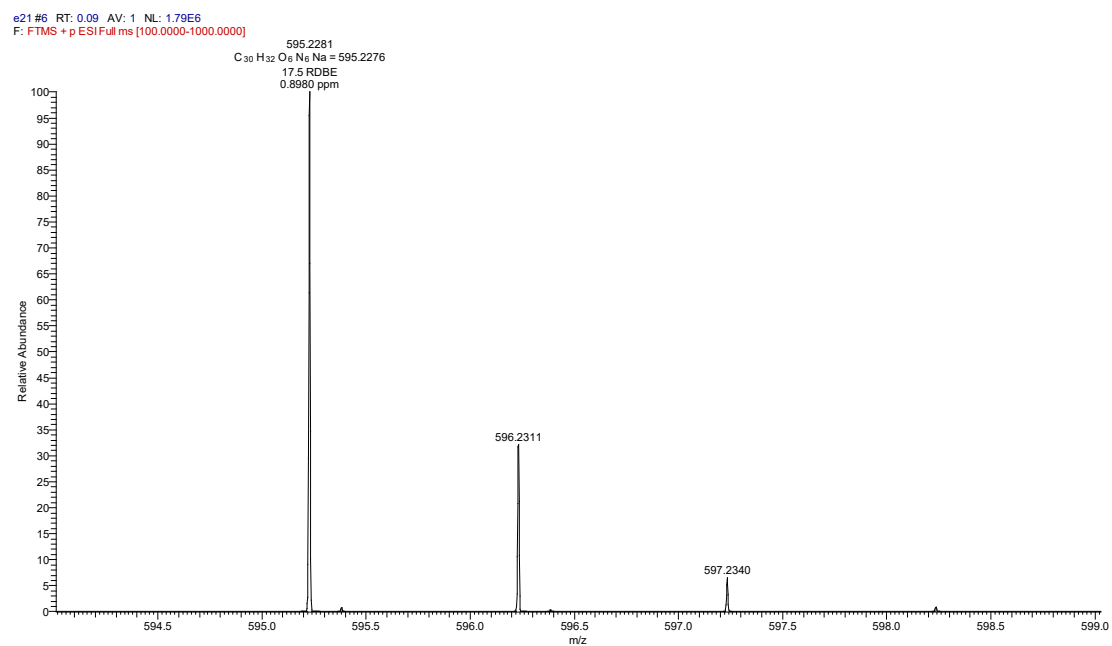


Figure S20-1. ^1H NMR spectrum (600 MHz, DMSO- d_6) of compound e20

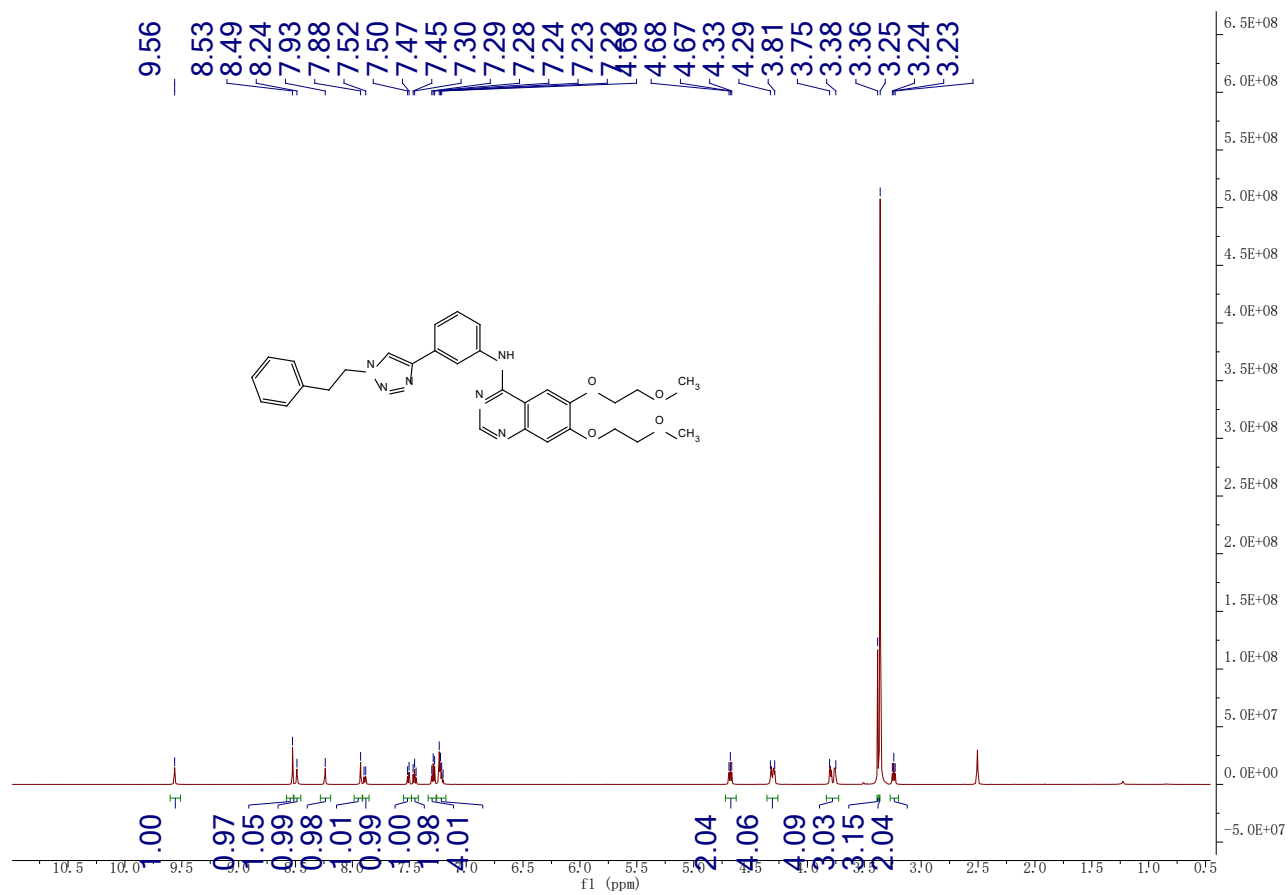


Figure S20-2. ^{13}C NMR spectrum (150 MHz, DMSO- d_6) of compound e20

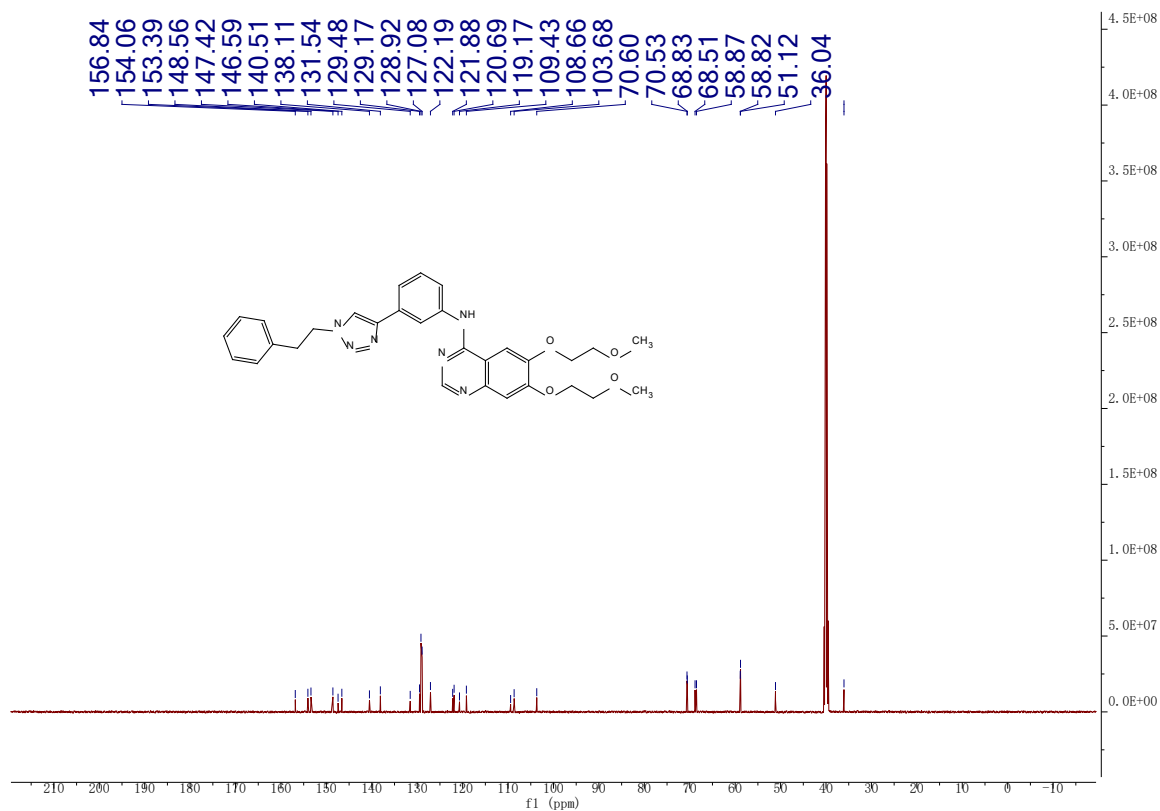
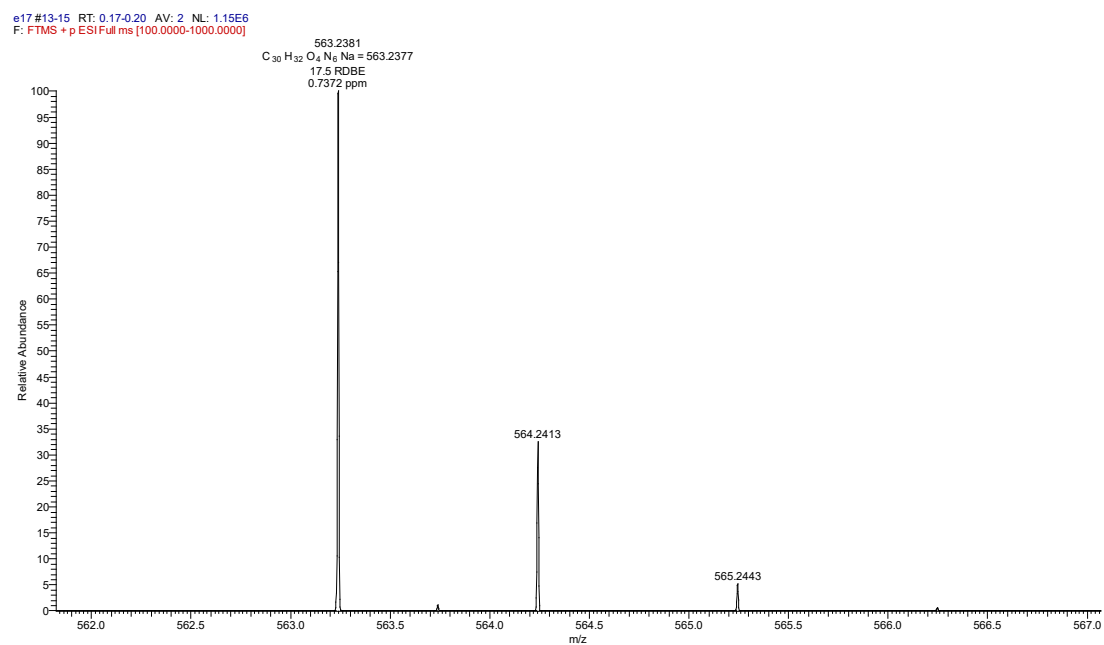


Figure S20-3. HR MS of compound e20



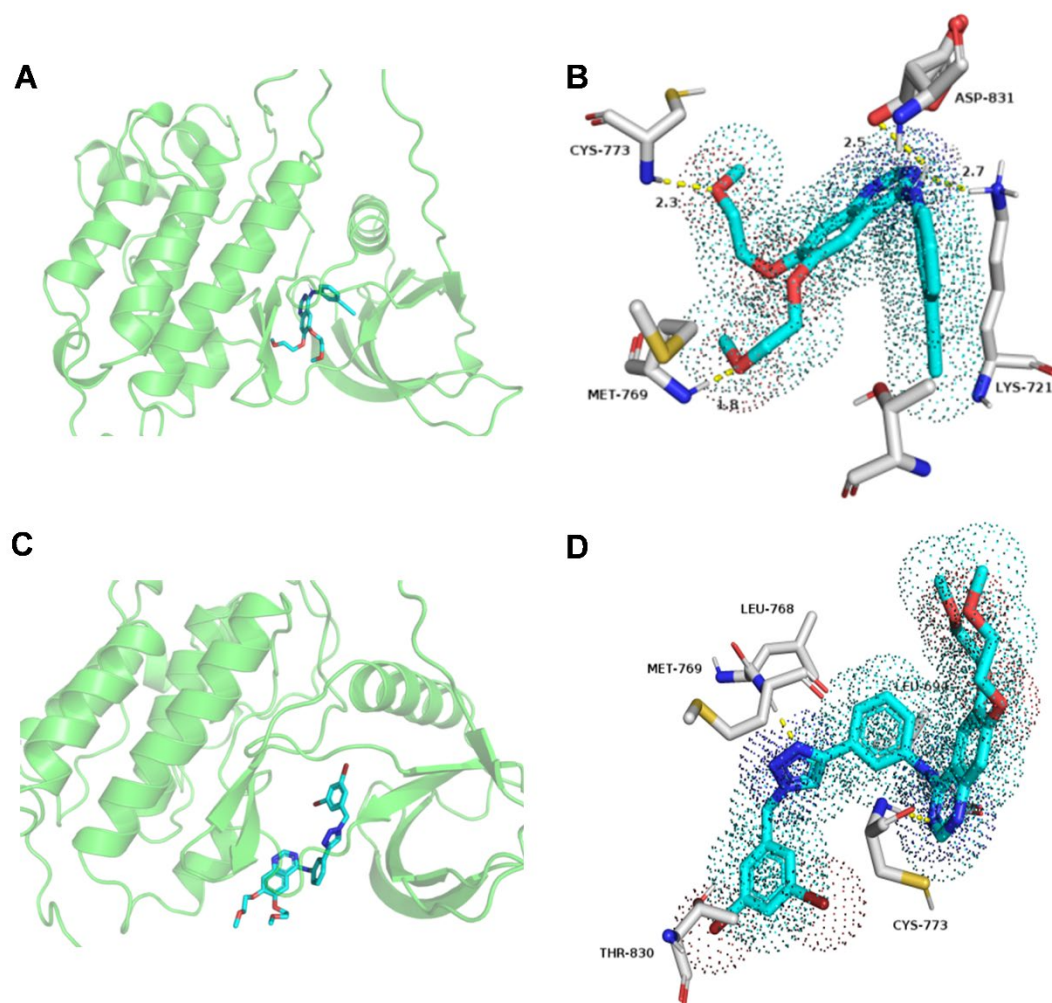


Fig. S21. The binding modes of Erlotinib and compound e4 in complex with EGFR (PDB: 1M17). (A) and (B) The binding mode of Erlotinib in the ATP binding site of EGFR. (C) and (D) The binding mode of compound e4 in the ATP binding site of EGFR.

We conducted molecular docking studies to explore the potential binding modes of compound e4 in the active site of EGFR. The docking studies states that the binding energy of erlotinib is -7.4 kcal/mol, and Elotinib formed hydrogen bonds with the target amino acid residues including LYS721, MET769, CYS773 and ASP831, and form hydrophobic interactions with THR766. Results show that e4 binds to EGFR in the ATP binding site mainly composed of hydrophobic residues with docking score of -9.6 kcal/mol. We found that the 3,5-dibromobenzy group introduced to the N3 position of 1,2,3-triazole could occupy the hydrophobic pocket containing LEU 694 and THR 830, and the backbone amino group of MET 769 formed a hydrogen bond with one amino of the 1,2,3-triazole group. The amino group of CYS 773 formed hydrogen

bonds with the quinazoline.

The docking results show that e4 has a desirable binding affinity with EGFR but a different binding mode from that of Erlotinib. However, our results show that the EGFR inhibitory activities of Erlotinib is better, but e4 demonstrates preferable anticancer activities in vitro. Thus, more mechanisms about the anticancer activities of these new compounds should be further studied.

Methods

In silico docking was carried out using AutoDock 4.2¹. Crystal structure of human epidermal growth factor receptor (EGFR) (PDB ID: 1M17) were downloaded from the protein data bank (<https://www.rcsb.org/>). Pymol (The PyMOL Molecular Graphics System) programs was used to remove all waters, ligands and co-factors. AutoDock Tools² was used to add hydrogens, calculate Gasteiger charges, and generate PDBQT files of compounds and receptor. The grid of 54, 54, and 54 points in x, y, and z directions of EGFR were built with a grid spacing of 0.375 Å and a distance-dependent function of the dielectric constant were used for the calculation of the energetic map. The default settings were used for all other parameters. Lamarckian genetic algorithm method³ was employed for docking simulations. The standard docking procedure was used for a rigid protein and a flexible ligand whose torsion angles were identified (for 200 independent runs per ligand).

Reference

1. Rizvi, S. M.; Shakil, S.; Haneef, M., A simple click by click protocol to perform docking: AutoDock 4.2 made easy for non-bioinformaticians. **2013**, (1611-2156 (Print)).
2. Morris, G. M.; Huey R Fau - Lindstrom, W.; Lindstrom W Fau - Sanner, M. F.; Sanner Mf Fau - Belew, R. K.; Belew Rk Fau - Goodsell, D. S.; Goodsell Ds Fau - Olson, A. J.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. **2009**, (1096-987X (Electronic)).
3. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J., Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry* **1998**, 19 (14), 1639-1662.