

Supporting Information

Discovery and Structural Optimization of N5-Substituted 6,7-dioxo-6,7-dihydropteridines as Potent and Selective Epidermal Growth Factor Receptor (EGFR) Inhibitors Against L858R/T790M Resistance Mutation

Yongjia Hao^{1, #}, Xia Wang^{1, #}, Tao Zhang^{2, #}, Deheng Sun¹, Yi Tong¹, Yuqiong Xu¹, Haiyang Chen¹, Linjiang Tong², Lili Zhu¹, Zhenjiang Zhao¹, Zhuo Chen¹, Jian Ding², Hua Xie^{2, *}, Yufang Xu^{1, *} and Honglin Li^{1, *}

¹Shanghai Key Laboratory of New Drug Design, Shanghai Key Laboratory of Chemical Biology, State Key Laboratory of Bioreactor Engineering, School of Pharmacy, East China University of Science & Technology, Shanghai 200237, China;

²Division of Anti-tumor Pharmacology, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, 201203, China.

*To whom correspondence should be addressed. Email: hxie@jding.dhs.org, yfxu@ecust.edu.cn, hlli@ecust.edu.cn

Content

Superposition simulations	S2
Docking studies	S2
Binding mode analysis	S3
In vivo antitumor efficacy study	S4
Material and methods of kinase assay	S4
General preparation of the key intermediates 15f-j.	S5
NMR spectra.....	S8
REFERENCES	S22

Superposition simulations

We did molecular superposition simulations for compound **9** and the designed inhibitors (**17a-e**) by a fast 3D molecular similarity calculation algorithm, SHAFTS¹. All compounds were stored in a 2D SDF file and an original conformation would be generated for each compound, then a conformation generation program based on a multi-objective evolution algorithm, Cyndi², was used to generate the conformation ensembles (a maximum size of 200 conformers) for each compound. The template structure was the crystal conformation of compound **4** bound to the ATP-binding site of EGFR^{T790M} with the PDB code of 3IKA³. Then SHAFTS would do the molecular fitting considering the shape overlay and pharmacophore features to build more valid models. The entire calculation process was simply performed submitting a job to the ChemMapper⁴ server, and the structures and SHAFTS similarity scores for all the compounds can be accessed at the same online site using the job ID of 38880 (<http://lilab.ecust.edu.cn/chemmapper/result/getResult.html?jobId=38880>).

Docking studies

Glide docking. The X-ray structures of EGFR^{WT} (PDB 4G5J)⁵ and EGFR^{T790M} (PDB 3IKA) were initially used in docking studies. All EGFR structures were retrieved from the Protein Data Bank and the molecular docking was performed with Glide⁶ (Schrödinger, Inc., version 9.0) in extra precision (XP) mode. The designed compounds (**17c-e**) were first docked into the EGFR protein as a reversible inhibitor using Glide, which was applied following the next steps:

- 1) All water molecules were removed from the structure of the complex. Hydrogen atoms and charges were added during a brief relaxation that was performed using the “Protein Preparation Wizard” workflow. After optimizing the hydrogen bond network, the crystal structure was minimized using the OPLS_2005 force field with the maximum root mean square deviation (RMSD) value of 0.3 Å.
- 2) The grid-enclosing box was placed on the centroid of compound **4** in the optimized crystal structure as described above, and defined so as to enclose residues located within 20 Å around the ATP-binding site, and a scale factor of 1.0 was set to van der Waals (VDW) radii of those receptor atoms with the partial atomic charge less than 0.25.
- 3) The three ligands were prepared with LigPrep module in Maestro, including adding hydrogen atoms, ionizing at a pH range from 5.0 to 9.0, and generating stereoisomers and valid single 3D conformers.
- 4) XP approach of Glide was adopted to dock the molecules into the ATP-binding site with the default parameters, and the top-ranking poses of each molecule were retained.

Covalent bond formation. One of the best poses was chosen based on the distance and orientation of the electrophilic functional group for each compound.

Consequently, the covalent bond between the sulfur atom on the Cys797 thiol and the terminal carbon atom on the acrylamide group was formed, and the bond geometry was optimized through the minimization using OPLS_2005 with MacroModel module in Maestro. The default values of the optimization parameters and thresholds were kept. All torsion angles of compound were released to freely rotate. The modeling structure of EGFR^{T790M} was overlaid with the EGFR^{WT} structure using PyMol, and all the structures are saved in the form of PDB coordinate files with names beginning with PDB code and ending with compound No., such as “3IKA_complexed_with_17d.pdb”.

Binding mode analysis

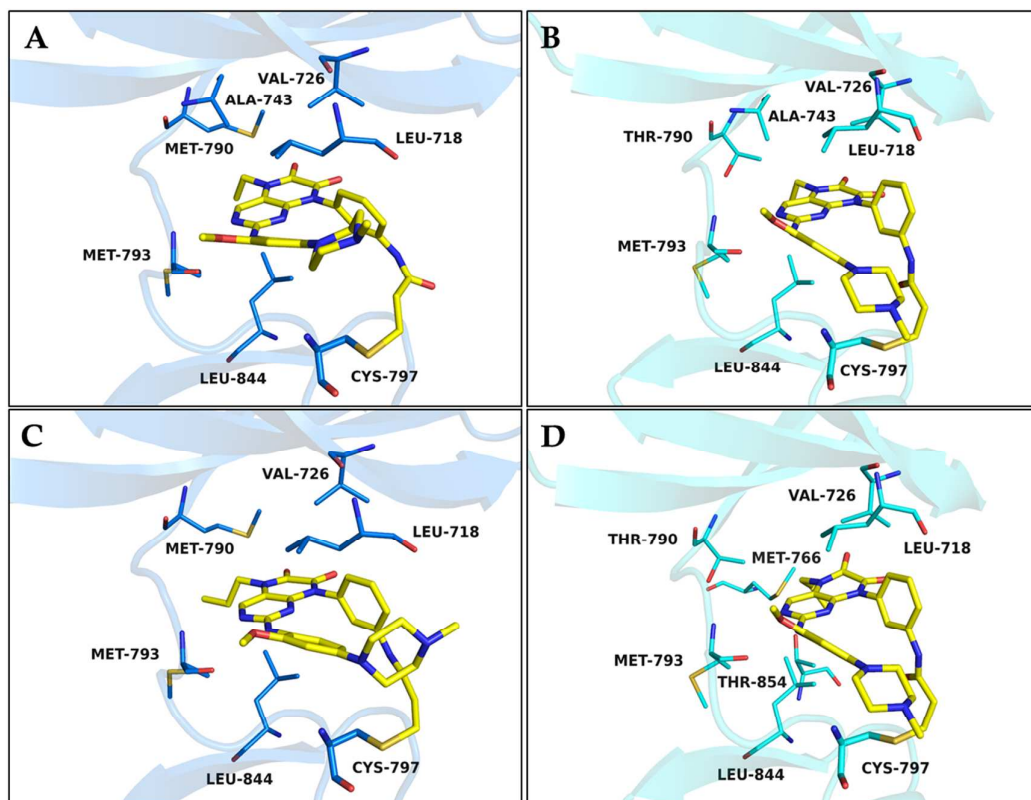


Figure S1. Structural modeling of compounds **17c** (A) and **17e** (C) bound to EGFR^{T790M} (PDB 3IKA) in the ATP-binding site. Predicted poses of compounds **17c** (B) and **17e** (D) against EGFR^{WT} (PDB 4G5J).

In vivo antitumor efficacy study

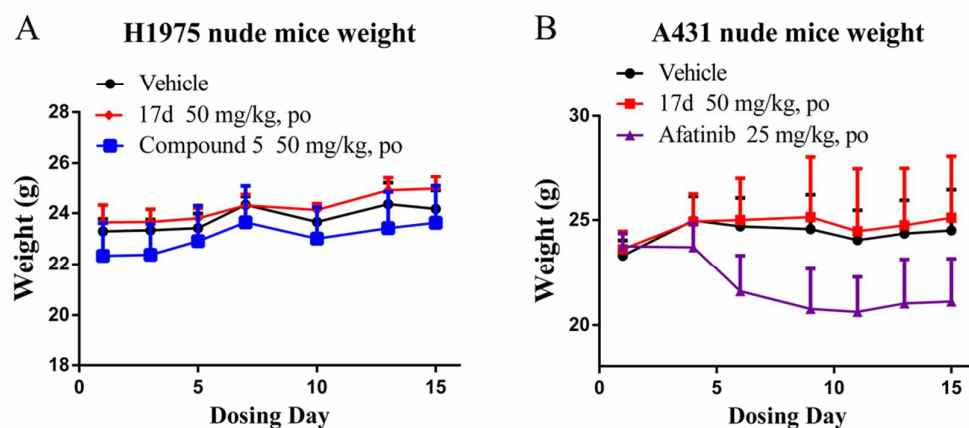


Figure S2. Preliminary *in vivo* antitumor efficacy of **17d**. (A) Body weights of H1975 xenograft mouse model (n = 3), (B) body weights of A431 xenograft mouse model (n = 6). All values represent mean \pm SEM.

Material and methods of kinase assay

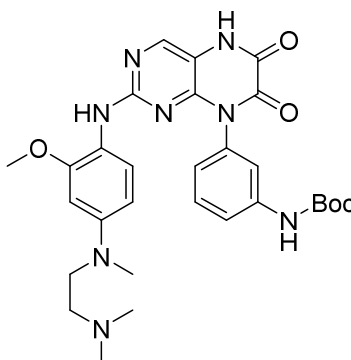
Kinases domain of EGFR^{WT}, EGFR^{L858R/T790M} were expressed using the Bac-to-BacTM baculovirus expression system (Invitrogen, Carlsbad, CA, USA) and purified in Ni-NTA columns (QIAGEN Inc., Valencia, CA, USA). The kinase activity was evaluated with enzyme-linked immunosorbent assay (ELISA). Briefly, 20 μ g/mL Poly (Glu, Tyr) 4:1 (Sigma, St. Louis, MO) was precoated in 96-well ELISA plates as substrate. Added 50 μ L of 10 μ mol/L ATP solution which was diluted in kinase reaction buffer (50 mM HEPES pH7.4, 20 mM MgCl₂, 0.1 mM MnCl₂, 0.2 mM Na₃VO₄, 1 mM DTT), the plate was treated with 1 μ L of indicated concentrations of compounds (dissolved in DMSO) per well. Experiments at each concentration were performed in duplicate. Reaction was initiated by adding tyrosine kinase diluted in kinase reaction buffer. After incubation at 37°C for 1 h, the wells were washed three times with phosphate buffered saline (PBS) containing 0.1% Tween 20 (T-PBS). 100 μ L anti-phosphotyrosine (PY99) antibody (1:1000, Santa Cruz Biotechnology, Santa Cruz, CA) diluted in T-PBS containing 5 mg/mL BSA was added and the plate was incubated at 37°C for 30 min. After the plate was washed three times, 100 μ L horseradish peroxidase-conjugated goat anti-mouse IgG (1:2000, Calbiochem, San Diego, CA) was added and the plate was incubated at 37°C for 30 min. The plate was washed, added with 100 μ L citrate buffer (0.1 M, pH 5.5) containing 0.03% H₂O₂. Then 2 mg/mL o-phenylenediamine was added, and samples were incubated at room temperature until color emerged. The reaction was terminated immediately by adding 50 μ L of 2M H₂SO₄. Plate was read using a multiwell spectrophotometer (VERSAmaxTM, Molecular Devices, Sunnyvale, CA, USA) at 492 nm. The inhibitory

rate (%) was calculated with the formula: $[1-(A_{492} \text{ treated}/A_{492} \text{ control})] \times 100\%$. IC₅₀ values were calculated from the inhibitory curves.

General preparation of the key intermediates 15f-j.

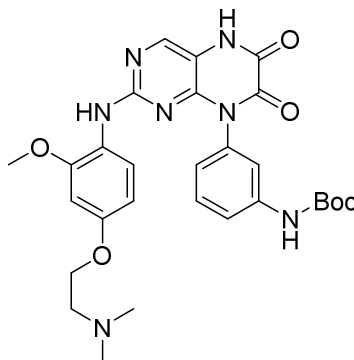
The following compounds **15f-j** were prepared by a method similar to that for compound **15**.

tert-butyl(3-(2-((4-((2-(dimethylamino)ethyl)(methyl)amino)-2-methoxyphenyl)amino)-6,7-dioxo-6,7-dihydropteridin-8(5H)-yl)phenyl)carbamate (15f)



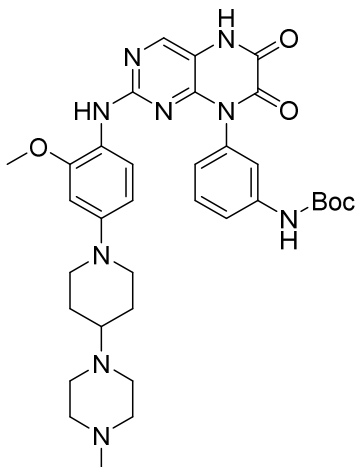
¹H NMR (400 MHz, DMSO-*d*₆) δ 9.61 (s, 1H), 8.12 (s, 1H), 7.57 (s, 1H), 7.54-7.52 (m, 1H), 7.41 (t, *J* = 8.0 Hz, 1H), 7.18 (d, *J* = 8.8 Hz, 1H), 6.95 (d, *J* = 8.0 Hz, 1H), 6.26 (d, *J* = 2.4 Hz, 1H), 5.86 (d, *J* = 7.6 Hz, 1H), 3.75 (s, 3H), 2.84 (s, 3H), 2.49 (t, *J* = 7.2 Hz, 2H), 2.24 (t, *J* = 7.2 Hz, 2H), 2.19 (s, 6H), 1.45 (s, 9H). LC-MS: *m/z*: 577.4 (M+H)⁺.

tert-butyl(3-(2-((4-(2-(dimethylamino)ethoxy)-2-methoxyphenyl)amino)-6,7-dioxo-6,7-dihydropteridin-8(5H)-yl)phenyl)carbamate (15g)



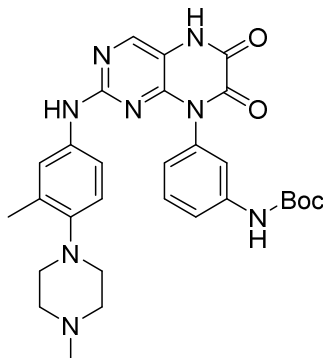
¹H NMR (400 MHz, DMSO-*d*₆) δ 9.64 (s, 1H), 8.15 (s, 1H), 7.63-7.61 (m, 2H), 7.52 (d, *J* = 8.0 Hz, 1H), 7.44 (t, *J* = 8.0 Hz, 1H), 7.31 (d, *J* = 8.8 Hz, 1H), 6.96 (d, *J* = 7.2 Hz, 1H), 6.53 (s, 1H), 6.08 (d, *J* = 8.8 Hz, 1H), 3.96 (t, *J* = 7.2 Hz, 2H), 3.77 (s, 3H), 2.59 (t, *J* = 7.2 Hz, 2H), 2.22 (s, 6H), 1.45 (s, 9H). LC-MS: *m/z*: 564.3 (M+H)⁺.

tert-butyl(3-(2-((2-methoxy-4-(4-(4-methylpiperazin-1-yl)piperidin-1-yl)phenyl)amino)-6,7-dioxo-6,7-dihydropteridin-8(5H)-yl)phenyl)carbamate (15h)



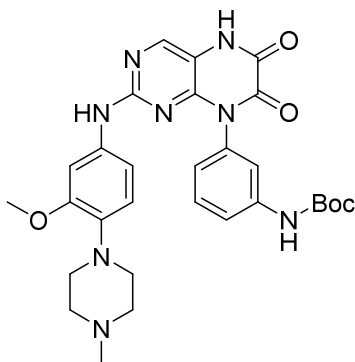
^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 9.64 (s, 1H), 8.15 (s, 1H), 7.58-7.56 (m, 2H), 7.55 (s, 1H), 7.43 (t, $J = 8.0$ Hz, 1H), 7.25 (d, $J = 8.8$ Hz, 1H), 6.95 (d, $J = 7.6$ Hz, 1H), 6.52 (d, $J = 2.0$ Hz, 1H), 6.07 (d, $J = 8.8$ Hz, 1H), 3.76 (s, 3H), 3.59-3.56 (m, 2H), 3.45-3.41 (m, 4H), 2.58-2.53 (m, 6H), 2.34-2.32 (m, 2H), 2.25 (s, 3H), 1.83 (d, $J = 11.2$ Hz, 2H), 1.54-1.49 (m, 1H), 1.46 (s, 9H). LC-MS: m/z : 658.5 ($\text{M}+\text{H}$) $^+$.

***tert*-butyl(3-(2-((3-methyl-4-(4-methylpiperazin-1-yl)phenyl)amino)-6,7-dioxo-6,7-dihydropteridin-8(5H)-yl)phenyl)carbamate (15i)**



^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 9.64 (s, 1H), 9.29 (s, 1H), 8.18 (s, 1H), 7.60 (s, 1H), 7.54 (d, $J = 6.4$ Hz, 1H), 7.44 (t, $J = 7.2$ Hz, 1H), 7.08 (s, 1H), 7.05-7.03 (m, 1H), 6.95 (d, $J = 4.4$ Hz, 1H), 6.66 (d, $J = 7.6$ Hz, 1H), 2.70 (t, $J = 4.0$ Hz, 4H), 2.43 (t, $J = 4.0$ Hz, 4H), 2.22 (s, 3H), 1.99 (s, 3H), 1.45 (s, 9H). LC-MS: m/z : 559.3 ($\text{M}+\text{H}$) $^+$.

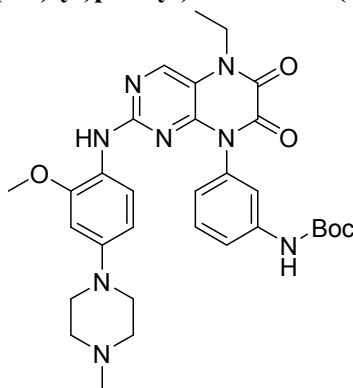
***tert*-butyl(3-(2-((3-methoxy-4-(4-methylpiperazin-1-yl)phenyl)amino)-6,7-dioxo-6,7-dihydropteridin-8(5H)-yl)phenyl)carbamate (15j)**



^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 9.64 (s, 1H), 9.23 (s, 1H), 8.19 (s, 1H), 7.60 (s, 1H), 7.54 (d, $J = 6.0$ Hz, 1H), 7.43 (t, $J = 7.6$ Hz, 1H), 6.94 (t, $J = 6.4$ Hz, 1H), 6.87 (s, 1H), 6.45 (d, $J = 7.2$ Hz, 1H), 3.59 (s, 3H), 2.83 (t, $J = 4.0$ Hz, 4H), 2.42 (t, $J = 4.0$ Hz, 4H), 2.21 (s, 3H), 1.45 (s, 9H). LC-MS: m/z : 575.3 ($\text{M}+\text{H}$) $^+$.

The 2D HMBC spectrum of compound 16c.

***tert*-butyl(3-(5-ethyl-2-((2-methoxy-4-(4-methylpiperazin-1-yl)phenyl)amino)-6,7-dioxo-6,7-dihydropteridin-8(5H)-yl)phenyl)carbamate (16c).**



^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ 9.68 (s, 1H), 8.54 (s, 1H), 7.69 (s, 1H), 7.65 (s, 1H), 7.60 (d, $J = 6.4$ Hz, 1H), 7.49 (t, $J = 6.4$ Hz, 1H), 7.30 (d, $J = 6.8$ Hz, 1H), 7.00 (d, $J = 6.4$ Hz, 1H), 6.61 (s, 1H), 6.12 (d, $J = 6.0$ Hz, 1H), 4.18 (q, $J = 5.2$ Hz, 2H), 3.82 (s, 3H), 3.19 (t, $J = 4.4$ Hz, 4H), 2.88 (t, $J = 4.4$ Hz, 4H), 2.54 (s, 3H), 1.49 (s, 9H), 1.30 (t, $J = 5.6$ Hz, 3H). ^{13}C NMR (125 MHz, $\text{DMSO-}d_6$) δ 155.52, 154.59, 152.72, 152.65, 148.97, 147.14, 146.10, 143.59, 140.58, 135.97, 129.33, 120.04, 121.29, 119.38, 118.04, 113.67, 106.63, 100.16, 79.27, 55.77, 53.73, 47.78, 44.25, 36.85, 28.11, 21.08, 12.00. HRMS(ESI) (m/z): ($\text{M}+\text{H}$) $^+$ calcd for $\text{C}_{31}\text{H}_{39}\text{N}_8\text{O}_5$ 603.3043, found, 603.3047. The obtained 2D HMBC spectrum was shown in Figure S3, which was consistent with the description in the reference.⁷

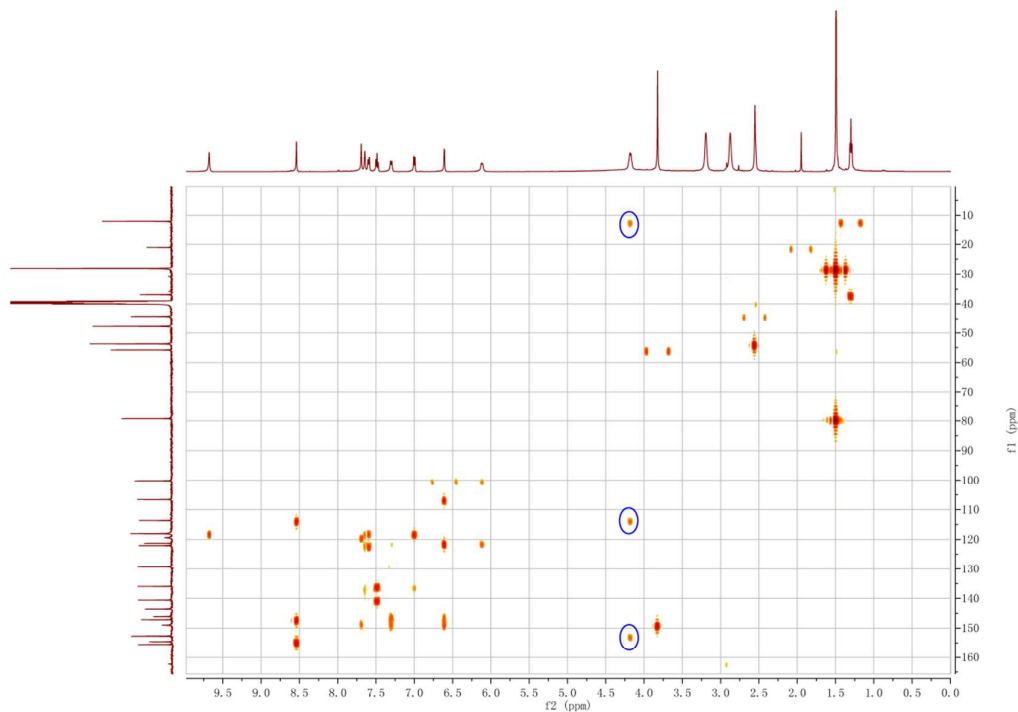
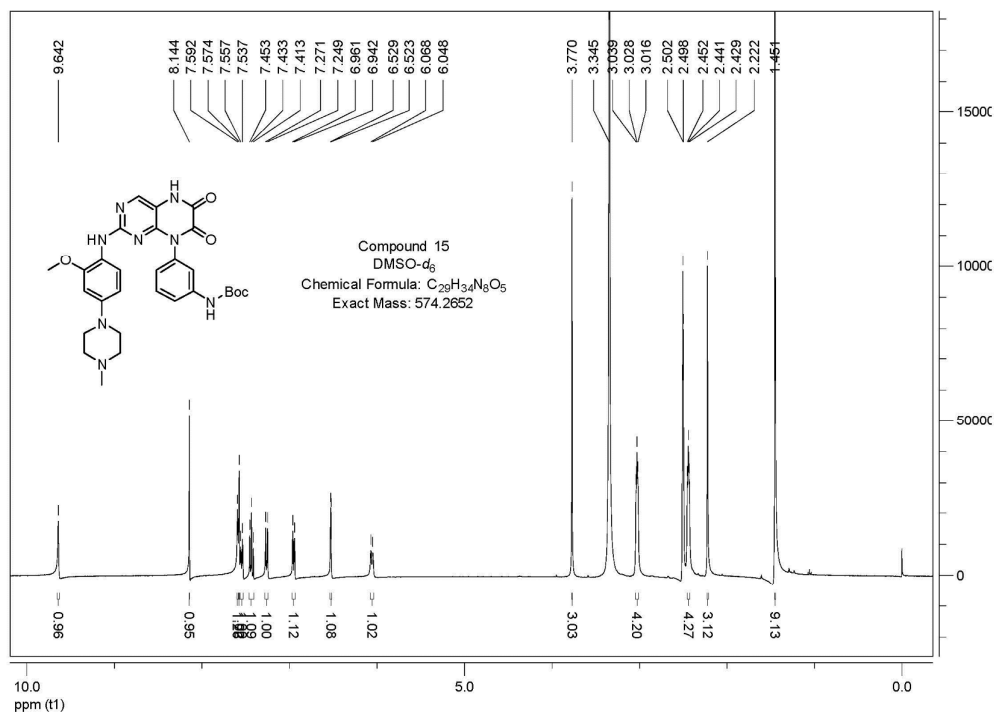
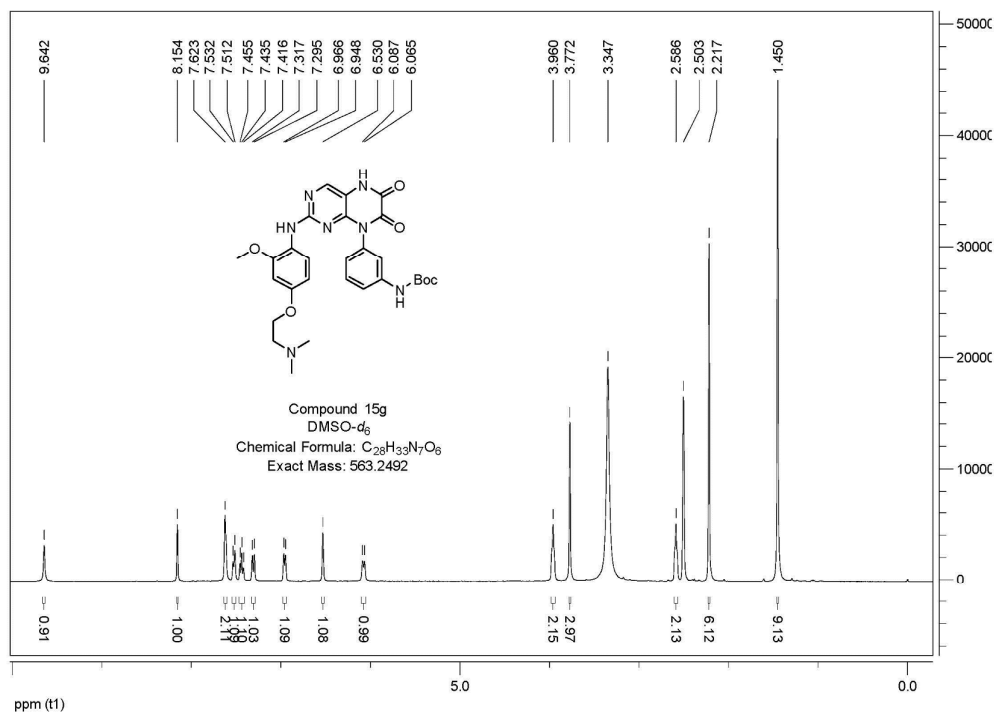
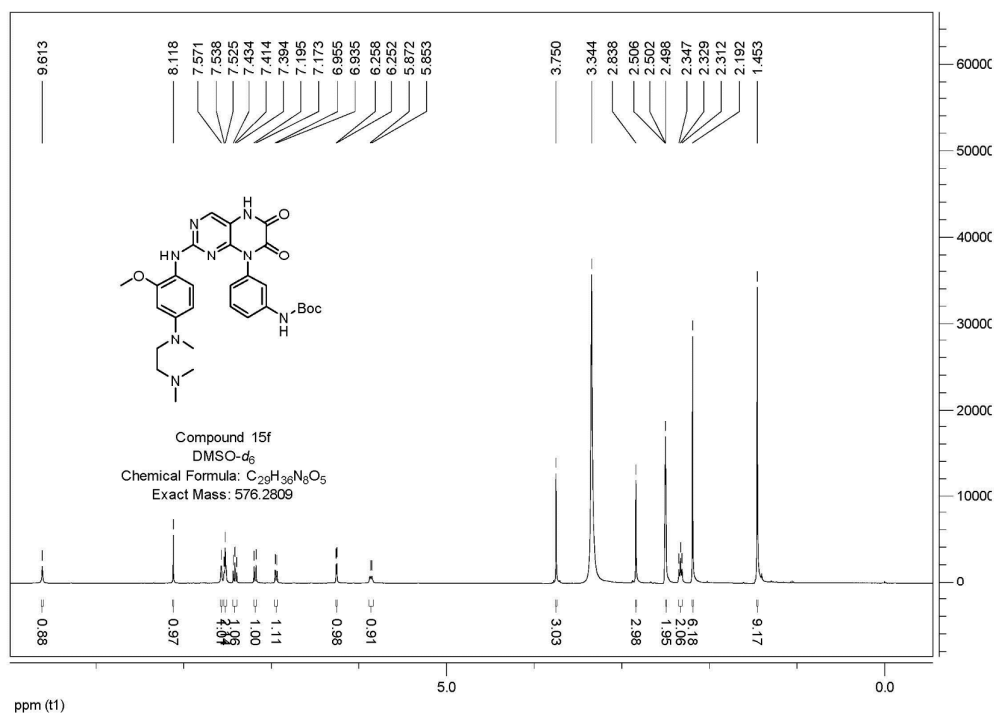
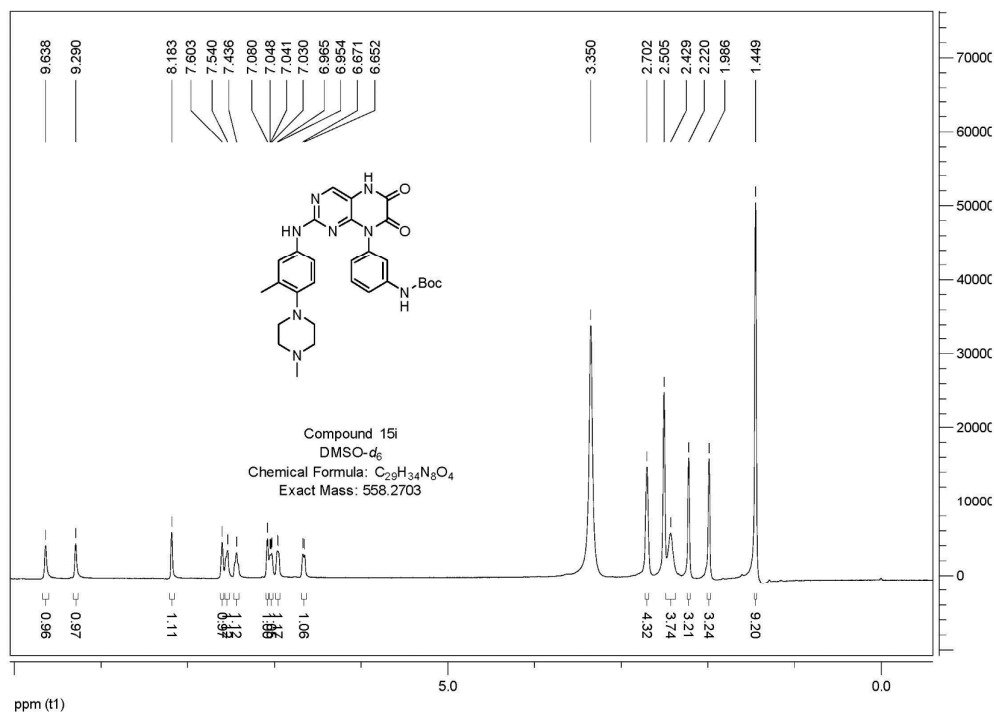
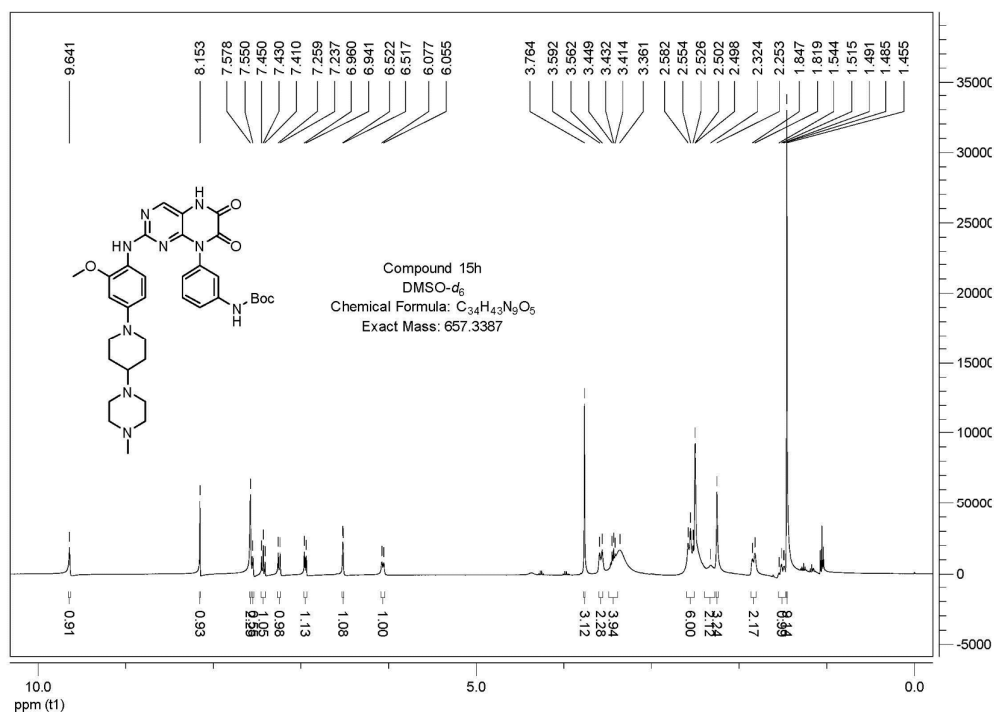


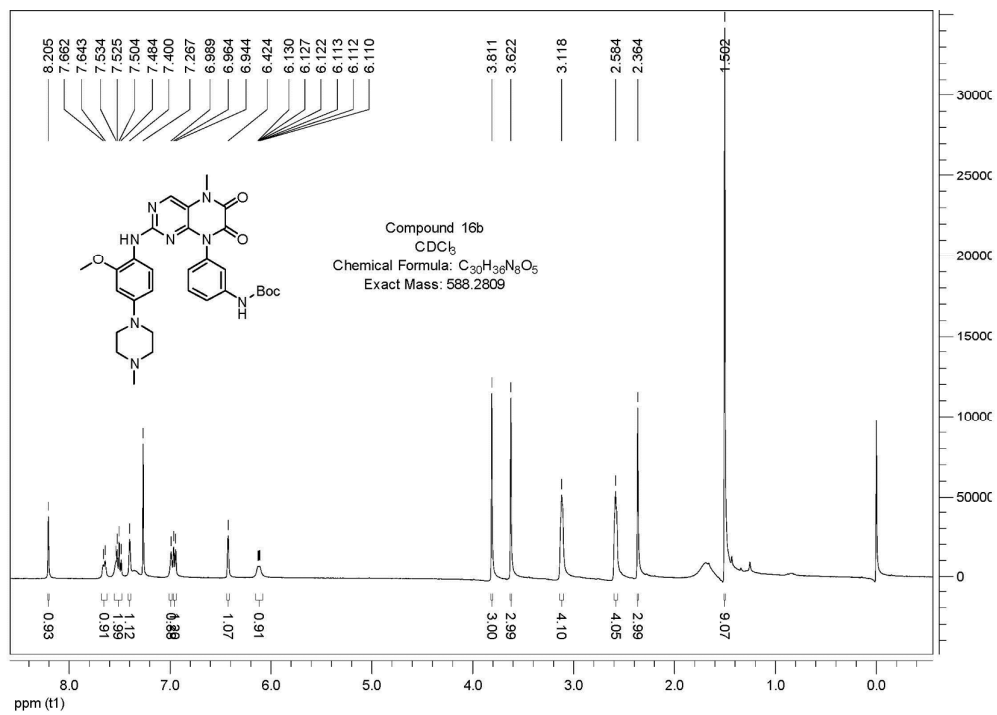
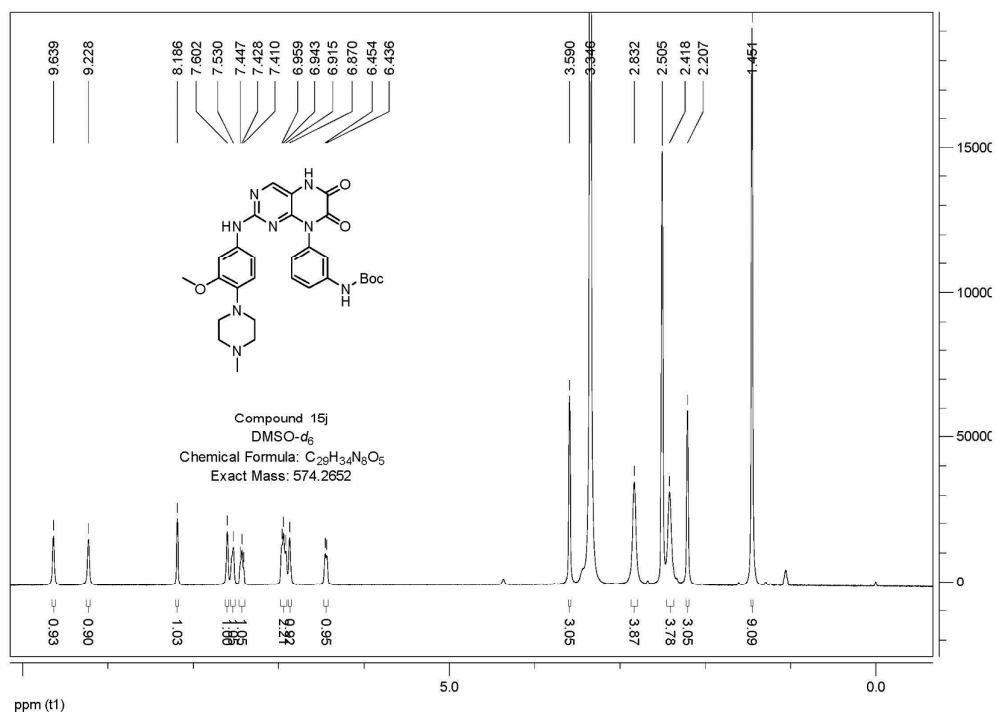
Figure S3. The 2D HMBC NMR spectrum of compound **16c**.

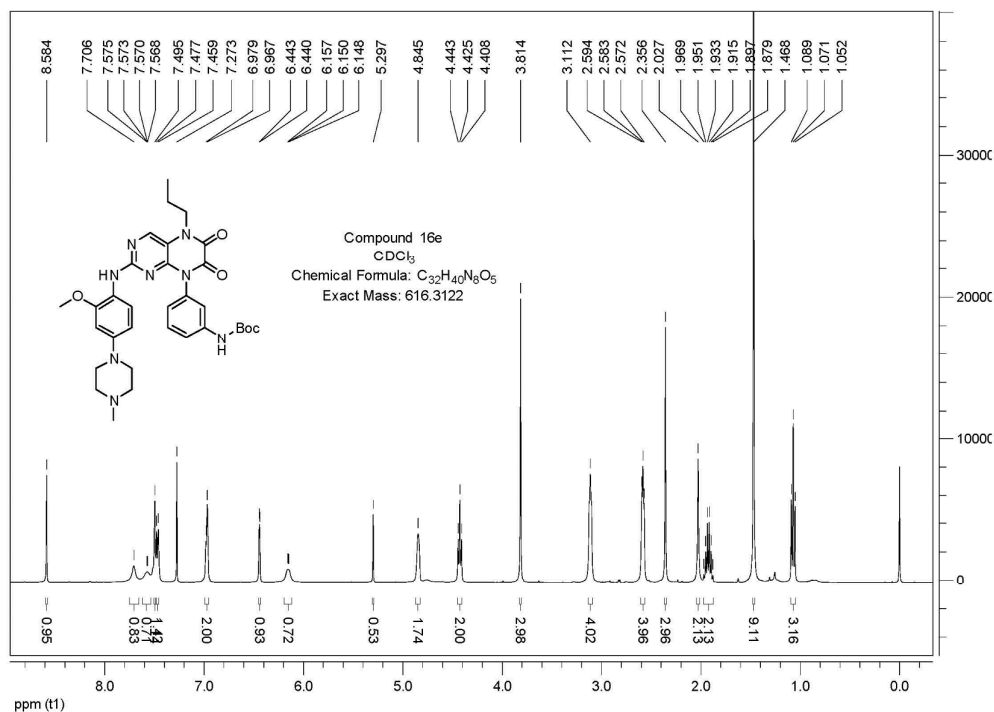
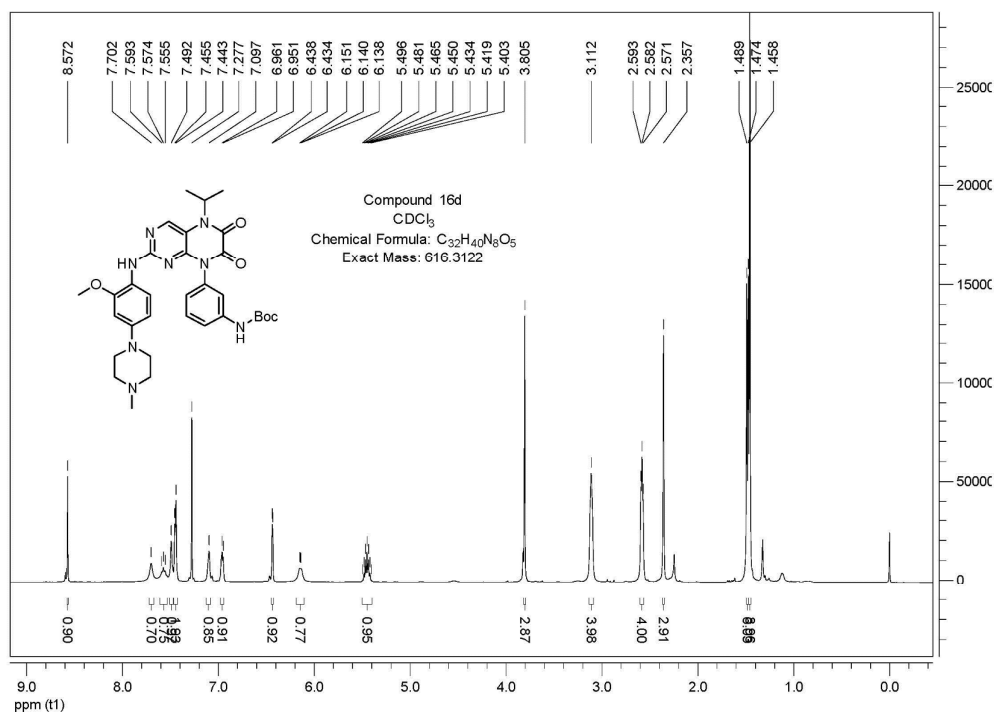
NMR spectra

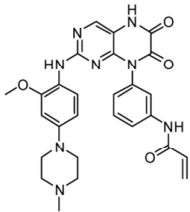
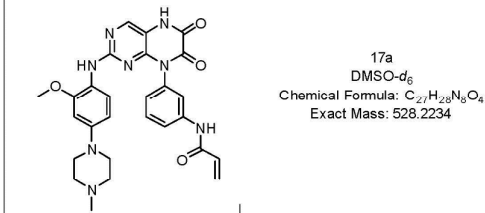


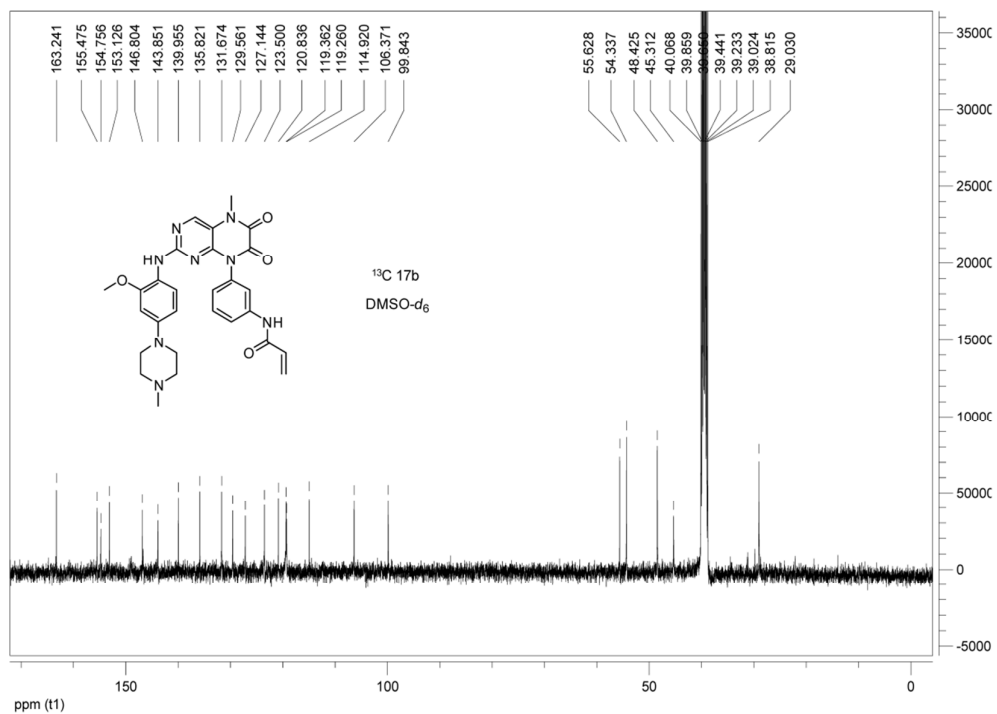
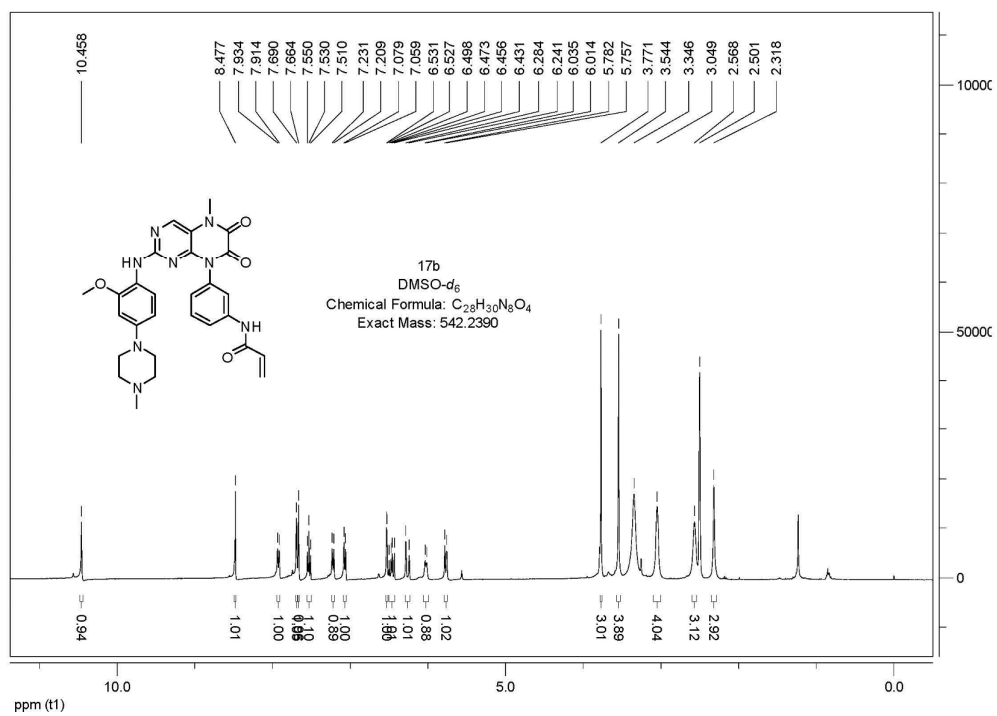


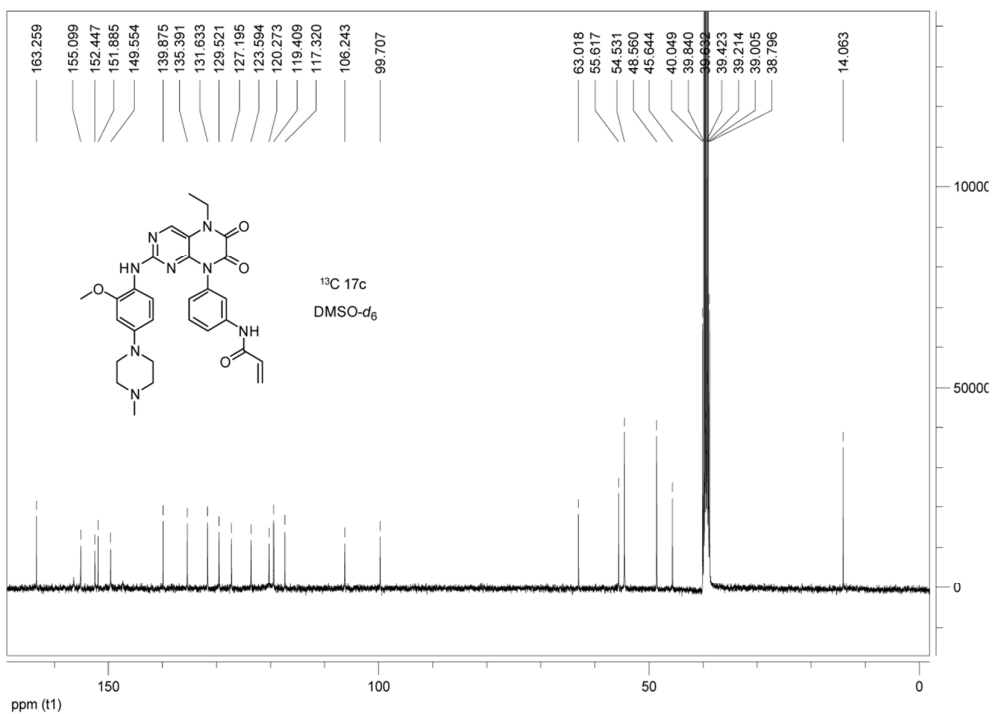


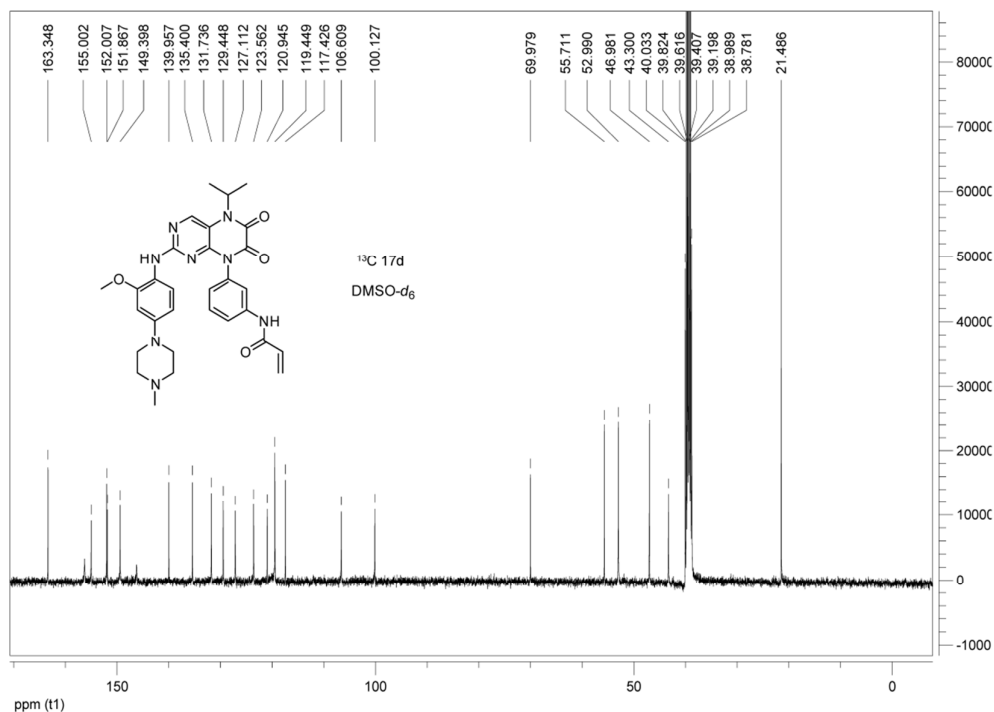
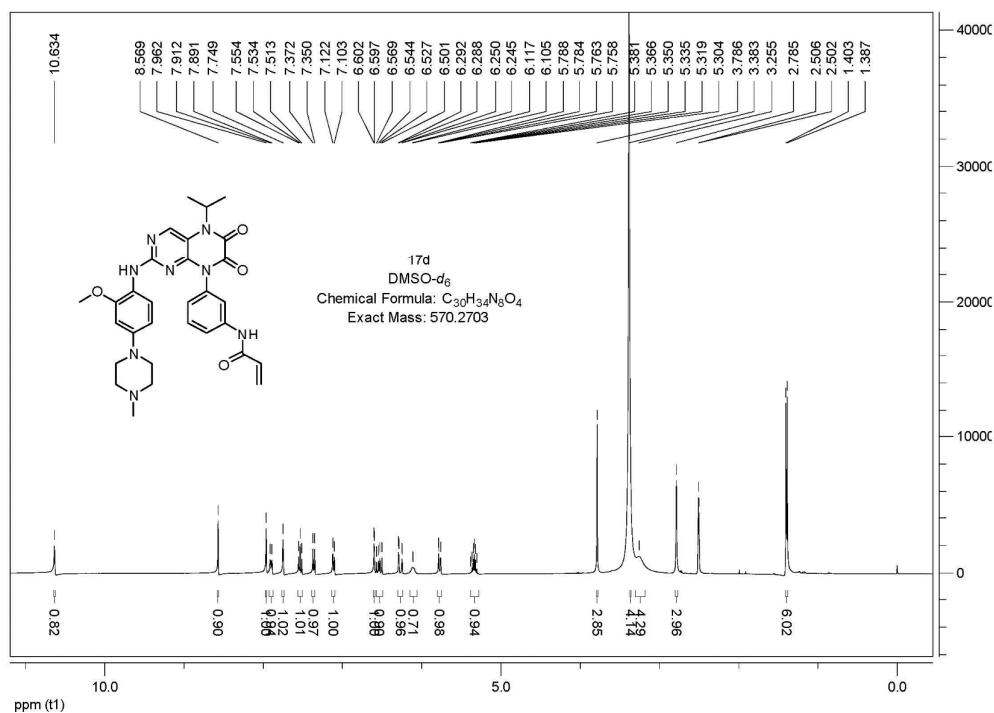


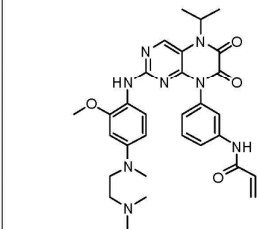
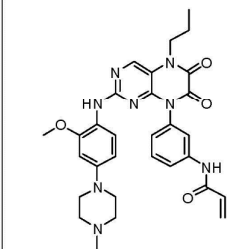


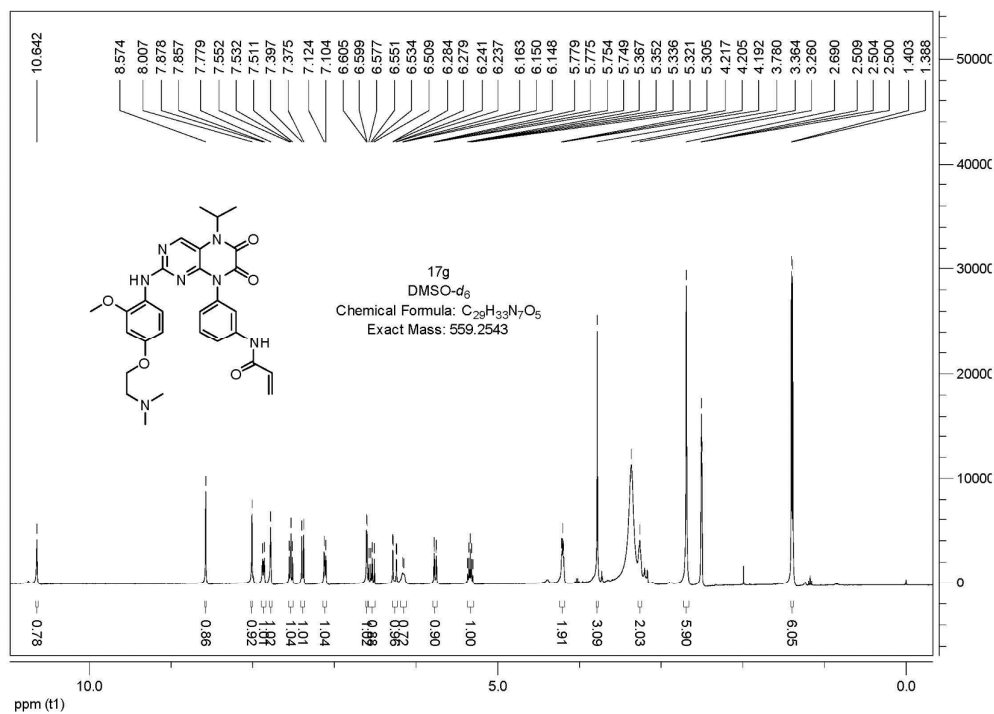
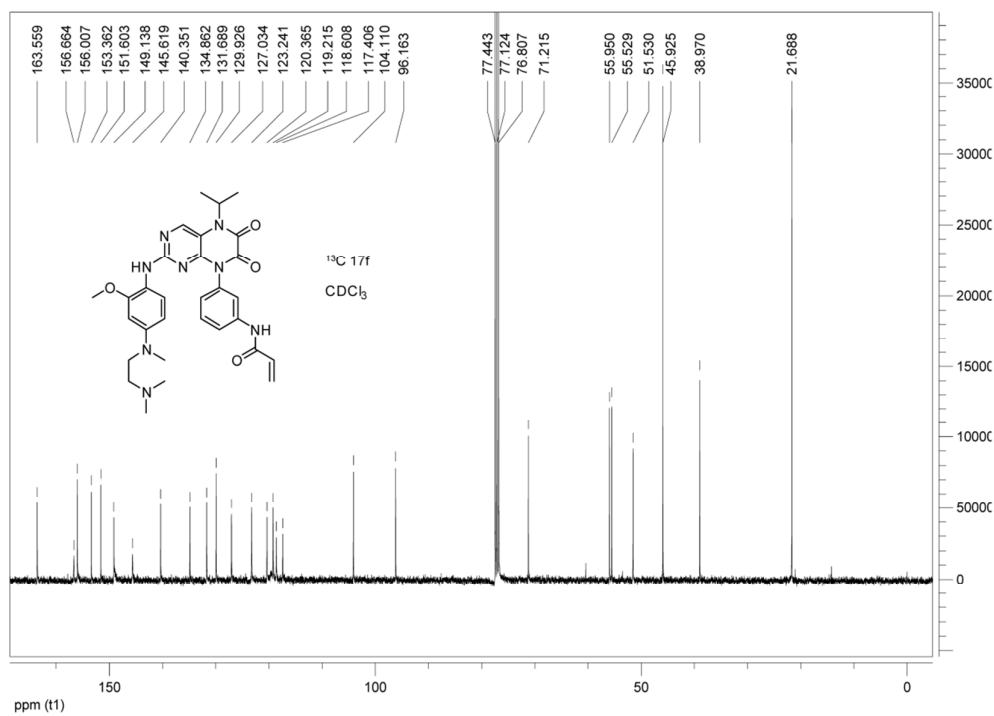


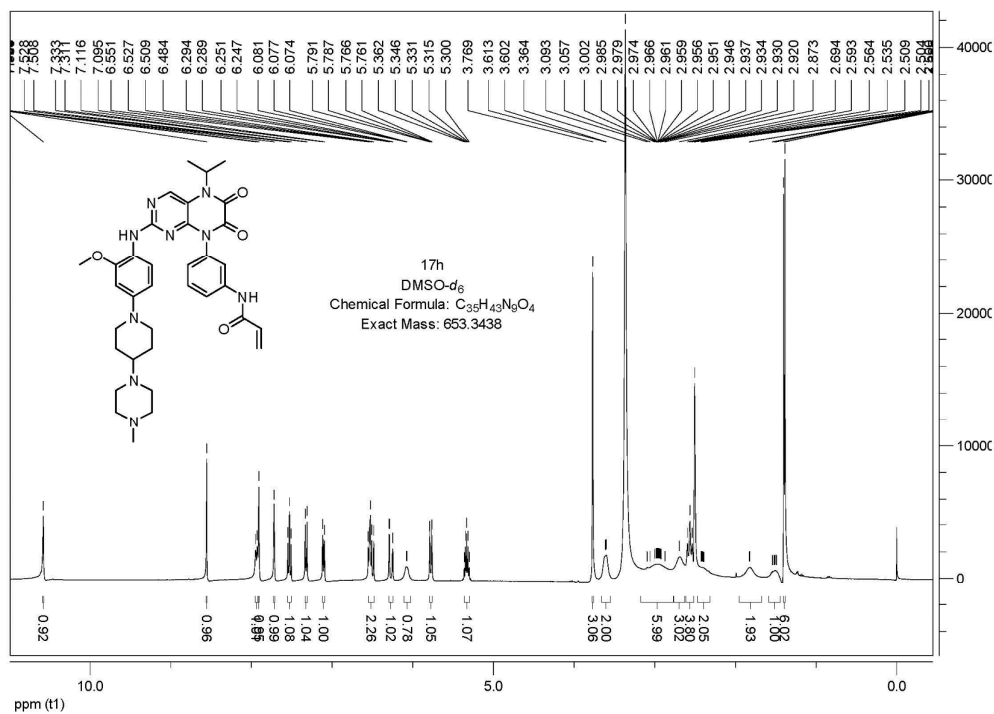
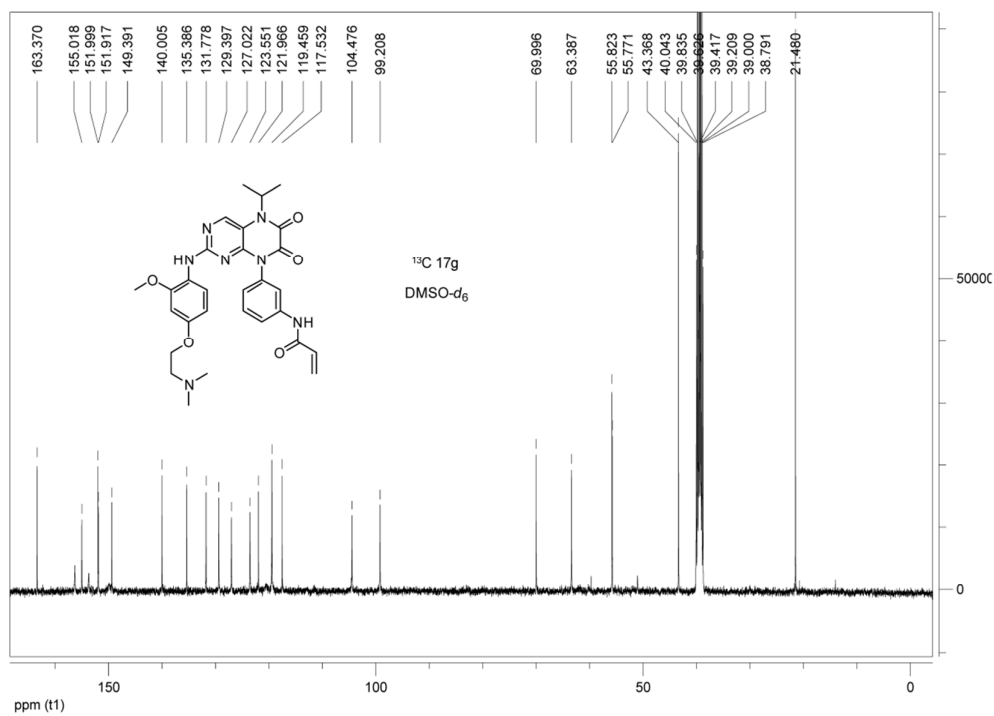


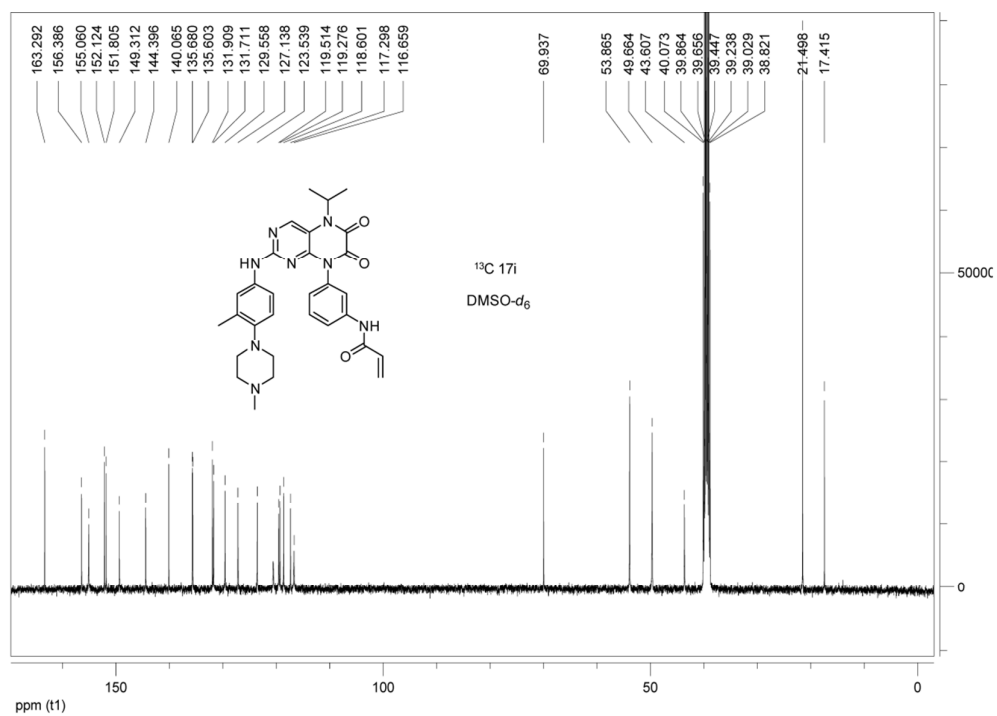
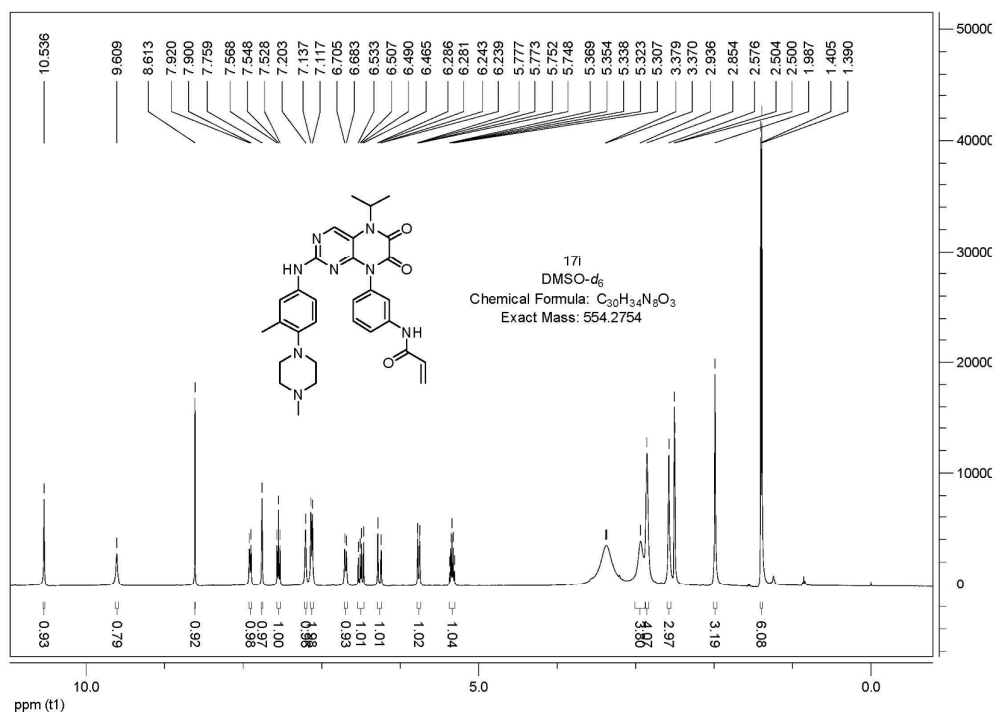


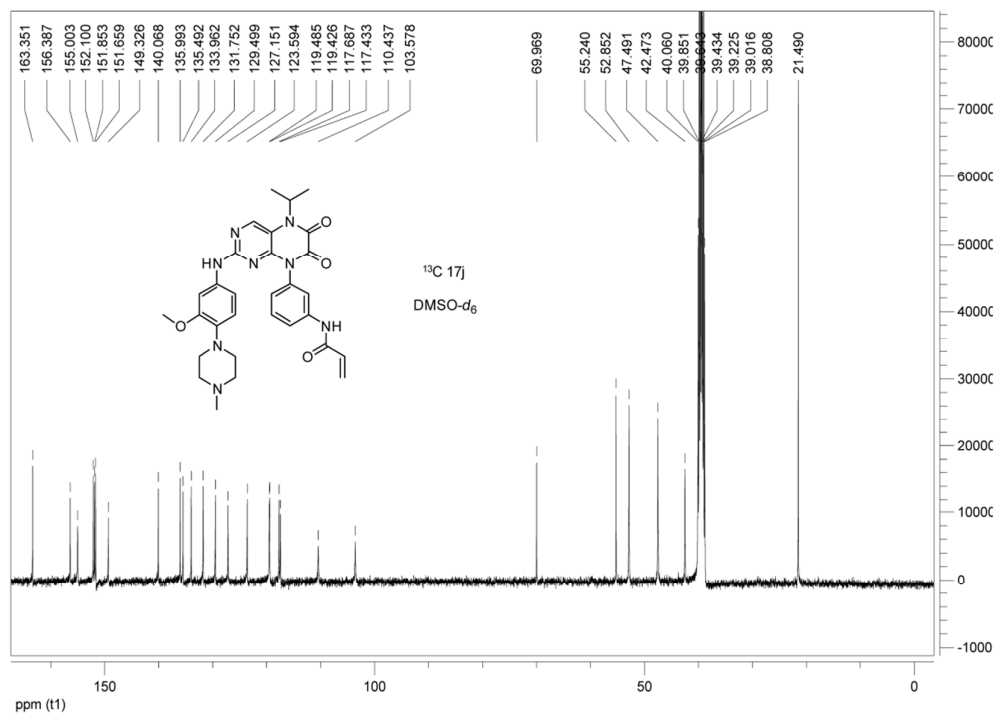
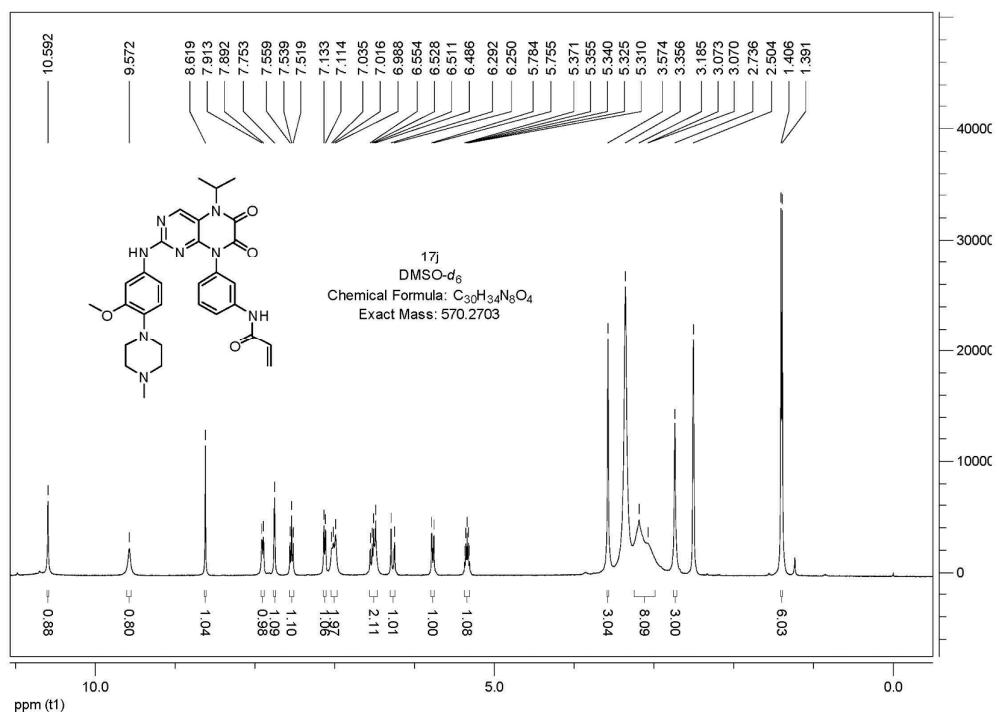












REFERENCES

- (1) Liu, X.; Jiang, H.; Li, H. SHAFTS: A Hybrid Approach for 3D Molecular Similarity Calculation. 1. Method and Assessment of Virtual Screening. *J. Chem. Inf. Model.* **2011**, *51*, 2372-2385.
- (2) Liu, X.; Bai, F.; Ouyang, S.; Wang, X.; Li, H.; Jiang, H. Cyndi: A multi-objective evolution algorithm based method for bioactive molecular conformational generation. *BMC Bioinformatics* **2009**, *10*, 101.
- (3) Zhou, W.; Ercan, D.; Chen, L.; Yun, C.; Li, D.; Capelletti, M.; Cortot, A. B.; Chirieac, L.; Iacob, R. E.; Padera, R.; Engen, J. R.; Wong, K.; Eck, M. J.; Gray, N. S.; Jänne, P. A. Novel Mutant-Selective EGFR Kinase Inhibitors against EGFR T790M. *Nature* **2009**, *462*, 1070-1074.
- (4) Gong J.; Cai C.; Liu X.; Ku X.; Jiang H.; Gao D.; Li H. ChemMapper: A Versatile Web Server for Exploring Pharmacology and Chemical Structure Association Based on Molecular 3D Similarity Method. *Bioinformatics* **2013**, *29*, 1827-1829.
- (5) Solca, F.; Dahl, G.; Zoephel, A.; Bader, G.; Sanderson, M.; Klein, C.; Kraemer, O.; Himmelsbach, F.; Haaksma, E.; Adolf, G. R. Target Binding Properties and Cellular Activity of Afatinib (BIBW2992), an Irreversible ErbB Family Blocker. *J. Pharmacol. Exp. Ther.* **2012**, *343*, 342-350.
- (6) Friesner, R.; Banks, J.; Murphy, R.; Halgren, T.; Klicic, J.; Mainz, D.; Repasky, M.; Knoll, E.; Shelley, M.; Perry, J. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*(7), 1739-1749.
- (7) Ghomsi, N. T.; Ahabchane, N. E. H.; Es-Safi, N. E.; Garrigues, B.; Essassi, E. M. Synthesis and Spectroscopic Structural Elucidation of New Quinoxaline Derivatives. *Spectrosc. Lett.* **2007**, *40*, 741-751.