Supporting Information

Carborane as an Alternative Efficient Hydrophobic Tag for Protein Degradation

Yasunobu Asawa,[a] Kei Nishida,[b] Kazuki Kawai,[a] Kiyotaka Domae,[a] Hyun Seung Ban,[c] Akihiro Kitazaki,[d] Hiroya Asami,[d] Jun-Ya Kohno,[d] Satoshi Okada,[a,e] Hiraku Tokuma,[a] Daisuke Sakano,[a] Shoen Kume,[a] Masaru Tanaka,[b] Hiroyuki Nakamura [a,e,*]

[a] *School of Life Science and Technology, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8503, Japan*

[b] *Institute for Materials Chemistry and Engineering, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan*

[c] *Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, South Korea*

[d] *Department of Chemistry, Facility of Science, Gakushuin University, 1-5-1 Mejiro, Toshima-ku, Tokyo 171-8588, Japan*

[e] *Laboratory for Chemistry and Life Science, Institute of Innovative Research, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8503, Japan*

*E-mail: hiro@res.titech.ac.jp

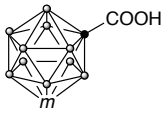
Content

1. Synthesis of Carborane Tags

General

NMR spectra were recorded on a Bruker biospin AVANCE II (400 MHz for ^{1}H, 100 MHz for ^{13}C) or a Bruker biospin AVANCE III (500 MHz for ^{1}H, 125 MHz for ^{13}C, 470 MHz for ^{11}B) instrument in the indicated solvent. Chemical shifts are reported in units per million (ppm) relative to the signal (0.00 ppm) for internal tetramethylsilane solutions in $CDCl_3$ (7.26 ppm for ^{1}H, 77.16 ppm for ^{13}C). Multiplicities are reported using the following abbreviations: s; singlet, d; doublet, dd; doublet of doublets, t; triplet, q; quartet, m; multiplet, br; broad, J; coupling constants in Hertz. Mass spectra were measured using a JMS-700 Mstation. IR spectra were recorded on a JASCO FT/IR-4100 spectrometer. IR spectrum was recorded on a JASCO Corporation FT/IR-4100 FT-IR Spectrometer. ATR PRO ONE was attached to the FT/IR-4100 in measuring solid IR spectroscopy by single reflection attenuated total reflection. Only the strongest and/or structurally important peaks were reported as the IR data given in cm^{-1}. High-resolution mass spectra (HRMS) were recorded on a Bruker ESI-TOF-MS (micrOTOF II). All reactions were monitored by thin-layer chromatography carried out on 0.2 mm E. Merck silica gel plates (60F-254) with UV light (254 nm) and were visualized using Hanessian's stain solution. Column chromatography was performed on Silica Gel 60 N, purchased from Fuji Silysia Chemical Ltd. HyT13, HyT55 and 2-(2-((6-chlorohexyl)oxy)ethoxy)-ethanamine (HaloTag linker **6**) were known and synthesized according to the literature procedures.[1,2]

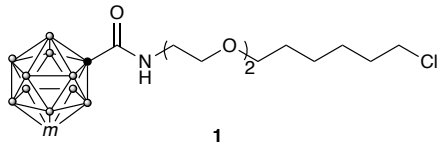
Synthesis of 1-carboxyl-1,7-dicarba-*closo*-dodecaborane (3)



3

To a solution of *m*-carborane (488 mg, 3.38 mmol) in THF (15 mL), was slowly added *n*-BuLi 1.6 M solution in hexane (2 mL, 3.21 mmol) at -78 ℃. After the resulting mixture was stirred room temperature under argon atmosphere for 1 h, the CO_2 balloon was attached. Then, the resulting mixture was stirred at room temperature CO_2 atmosphere for 3 h. After that, the reaction mixture was concentrated under pressure. The resulting mixture was added and Hexane after which the product was partitioned between the aqueous and organic layers. The aqueous layer was washed with Hexane. Then, the aqueous layer was added 6 M HCl and Hexane after which the product was partitioned between the aqueous and organic layers. The aqueous layer was extracted with Hexane and combined organic layers were dried over sodium sulfate and concentrated under vacuum to afford **3** (444 mg, 2.36 mmol, 74%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.03 (brs, 1H), 4.10-1.60 (m, 10H). This spectrum is consistent with the reported values.[3]

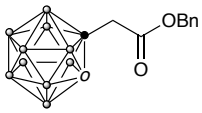
Synthesis of *N*-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-1,7-dicarba-*closo*-dodecaborane-1-carboxamide (1)



1

To a solution of **3** (6.1 mg, 0.031 mmol.) and HaloTag linker **6** (40 mg, 1.78 mmol) in DCM (1.2 mL) were added HOBt (15 mg, 0.11 mmol) and EDCI (17.9 mg, 0.093 mmol). After the reaction mixture was cooled to 0 ℃, was slowly added DIEA (40 μL, 3.09 mmol). The resulting mixture was stirred at room temperature for 8.5 h and quenched with H_2O. The mixture was extracted twice with ethyl acetate and the combined extracts were washed with brine, dried over Na_2SO_4, filtered, and concentrated. The residue was purified by column chromatography on silica gel (70% EtOAc in Hexane) afforded **1** as slight yellow oil (7.5 mg, 0.019 mmol, 60%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.32 (bs, 1H), 3.61-1.59 (m, 2H), 3.57-3.51 (m, 6H), 3.47 (t, *J* = 6.7 Hz, 2H), 3.39 (q, *J* = 5.2 Hz, 2H), 3.02 (s, 1H), 2.80-1.80 (m, 10H), 1.82-1.76 (m, 2H), 1.65-1.59 (m, 2H), 1.48-1.44 (m, 2H), 1.41-1.36 (m, 2H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 160.6, 71.7, 70.8, 70.4, 69.3, 55.2, 45.4, 40.8, 32.9, 30.0, 29.8, 27.0, 25.8. ^{11}B NMR (160 MHz, $CDCl_3$): δ -5.68, -7.46, -10.9, -11.6, -13.3, -15.7; IR (NaCl disc): 2926, 2856, 2604, 1686, 1519, 1276, 1216, 1114, 760 cm^{-1}; HRMS (ESI, positive) for $C_{13}H_{32}B_{10}ClNO_3$ (m/z): calculated 418.2927 $(M+Na)^+$, found 418.2918.

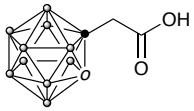
Synthesis of 1-(2-(benzyloxy)-2-oxoethyl)-1,2-dicarba-*closo*-dodecaborane (4)



4

To a solution of *o*-carborane (757 mg, 5.25 mmol) in THF (20 mL), was added *n*-BuLi 1.6 M solution in hexane (3.28 mL, 5.25 mmol) at -78 ℃. After the resulting mixture was stirred at 0 ℃ under argon atmosphere for 1 h, benzyl bromoacetate (784 μL, 5.0 mmol) was added. Then, the resulting mixture was stirred at room temperature for 3 h. After that, the reaction mixture was quenched with H_2O. The mixture was extracted twice with ethyl acetate and the combined extracts were washed with brine, dried over Na_2SO_4, filtered, and concentrated. The residue was purified by column chromatography on silica gel (10% EtOAc in Hexane) afforded **6** as colorless oil (691 mg, 2.36 mmol, 47%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.40-7.33 (m, 5H), 5.14 (s, 2H), 4.35 (s, 1H), 3.27 (s, 2H), 3.00-1.60 (m, 10H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 167.3, 134.6, 129.1, 128.1, 128.8, 67.9, 58.7, 41.7; ^{11}B NMR (160 MHz, $CDCl_3$): δ -2.05, -4.98, -9.23, -10.5, -12.0, -12.8; IR (NaCl disc): 2592, 1739, 1178 cm^{-1}; HRMS (ESI, positive) for $C_{11}H_{20}B_{10}O_2$ (m/z): calculated 317.2305 $(M+Na)^+$, found 317.2307.

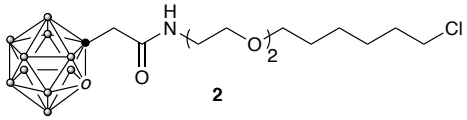
Synthesis of 1-(carboxymethyl)-1,2-dicarba-*closo*-dodecaborane (5)



5

To a solution of **4** (207 mg, 0.71 mmol) in MeOH (10 mL), was added Pd/C wetted with water (34 mg). The resulting mixture was stirred at room temperature until the full conversion of a starting material was observed. The reaction mixture was filtered on celite pad and concentrated. The resulting mixture was dissolved with aq. $NaHCO_3$ and washed with hexane three times. After the water layer was acidified with aq. HCl, extracted with EtOAc, dried over Na_2SO_4, filtered, and concentrated afforede **5** (123 mg, 0.61 mmol, 86%). ^{1}H NMR (400 MHz, CD_3CN): δ 8.93 (br, 1H), 4.64 (s, 1H), 3.28 (s, 2H), 3.10-1.40 (m, 10H). This spectrum is consistent with the reported values.[4]

Synthesis of *N*-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-1,7-dicarba-*closo*-dodecaborane-1-acetamide (2)

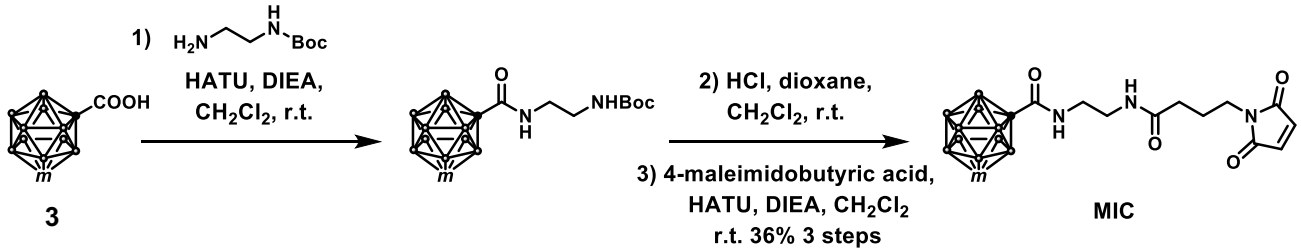


2

To a solution of **5** (40 mg, 0.20 mmol) in CH_2Cl_2 (1 mL), was slowly added excess amount of oxalyl chloride. After the resulting mixture was stirred at 50 ℃ under argon atmosphere for 12 h, the reaction mixture was concentrated under pressure. The crude materials were used to the next reaction without further purification.

To a solution of HaloTag linker **6** (44 mg, 0.20 mmol) and triethylamine (400 μL) in CH_2Cl_2 (1 mL) at 0 ℃ was slowly added the crude material in CH_2Cl_2 (500 μL). After the resulting mixture was stirred at 0 ℃ for 30 min, it was warmed up to room temperature and stirred for 1 h. The reaction mixture was quenched with 1M aq. HCl. The mixture was extracted three times with CH_2Cl_2 and the combined extracts were dried over Na_2SO_4, filtered, and concentrated. Then, the crude materials were purified by column chromatography on silica gel (70% EtOAc in Hexane) afforded **2** as a white solid (50 mg, 0.12 mmol, 62%). m.p. 38-40 ℃; ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.36 (bs, 1H), 4.60 (s, 1H), 3.64-3.62 (m, 2H), 3.59-3.52 (m, 6H), 3.48 (t, J = 6.7 Hz, 2H), 3.43 (q, J = 5.2 Hz, 2H), 3.04 (s, 2H), 2.80-1.60 (m, 10H), 1.81-1.75 (m, 2H), 1.64-1.59 (m, 2H), 1.50-1.44 (m, 2H), 1.40-1.36 (m, 2H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 166.3, 71.4, 70.5, 70.1, 69.3, 69.2, 58.8, 45.1, 43.5, 39.7, 32.6, 29.6, 26.8, 25.5. ^{11}B NMR (160 MHz, $CDCl_3$): δ -2.15, -5.17, -9.53, -10.7, -11.8, -12.7; IR (NaCl disc): 2937, 2865, 2590, 1659, 1556, 1098 cm^{-1}; HRMS (ESI, positive) for $C_{14}H_{34}B_{10}ClNO_3$ (m/z): calculated 432.3084 $(M+Na)^+$, found 432.3073.

Synthesis of 1-((2-(4-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)butanamido)ethyl)carbamoyl)-1,7-dicarba-*closo*-dodecaborane (MIC)



To a solution of **3** (270 mg, 1.43 mmol), HATU (815 mg, 214 mmol), and DIEA (369 mg, 2.86 mmol) in CH_2Cl_2, was added *N*-Boc-ethylenediamine (400 mg, 2.50 mmol) and stirred for 3 h at room temperature. The mixture was washed with 1 M HCl aq., saturated $NaHCO_3$ aq., and Brine. The organic layer was dried over $MgSO_4$, and dried up under the reduced pressure. To a solution of the resulting crude materials in CH_2Cl_2 (1.0 mL), was added 4 M HCl in dioxane (4.0 mL) at 0 ℃ and the mixture was stirred at room temperature for 10 h. After removal of the solvent under the reduced pressure, the residue was suspended in EtOAc and washed with saturated $NaHCO_3$ aq. and brine. The organic layer was dried over $MgSO_4$, the solvent was removed. The crude materials were used to the next reaction without further purification.

To a solution of 4-maleimidobutyric acid (256 mg, 1.40 mmol), HATU (570 mg, 1.50 mmol), and DIEA (387 mg, 3.00 mmol) in CH_2Cl_2 was added the crude material in CH_2Cl_2, and stirred at room temperature for 11 h. The resulting mixture was washed with 1 M HCl aq., saturated $NaHCO_3$ aq., and Brine. The organic layer was dried over $MgSO_4$, and dried up under the reduced pressure. The crude materials were purified by column chromatography on silica gel (20% EtOAc in Hexane) afforded **MIC** as a white solid (204 mg, 0.516 mmol, 36% 3 steps). m.p. 48-50 ℃;^{1}H NMR (500 MHz; $CDCl_3$) δ 7.09 (s, 1H), 6.73 (s, 2H), 6.39 (s, 1H), 3.58 (t, *J* = 6.5 Hz, 2H), 3.33-3.38 (m, 4H), 3.05 (s, 1H,), 2.17 (t, *J* = 7.0 Hz, 2H), 1.94 (quin, *J* = 6.5 Hz, 2H). ^{13}C NMR (125 MHz; $CDCl_3$) δ 173.2, 171.0, 161.18, 134.2, 75.45, 42.10, 38.82, 36.93, 33.08, 24.60. ^{11}B NMR (160 MHz; $CDCl_3$) δ -5.63, -7.46, -10.87, -11.57, -13.27, -15.73; IR (NaCl disc): 3343, 3062, 2939, 2604, 1706. HRMS (ESI, positive) for $C_{13}H_{25}B_{10}N_3O_4$ (m/z): calculated 420.2690 $(M+Na)^+$; found 420.2688.

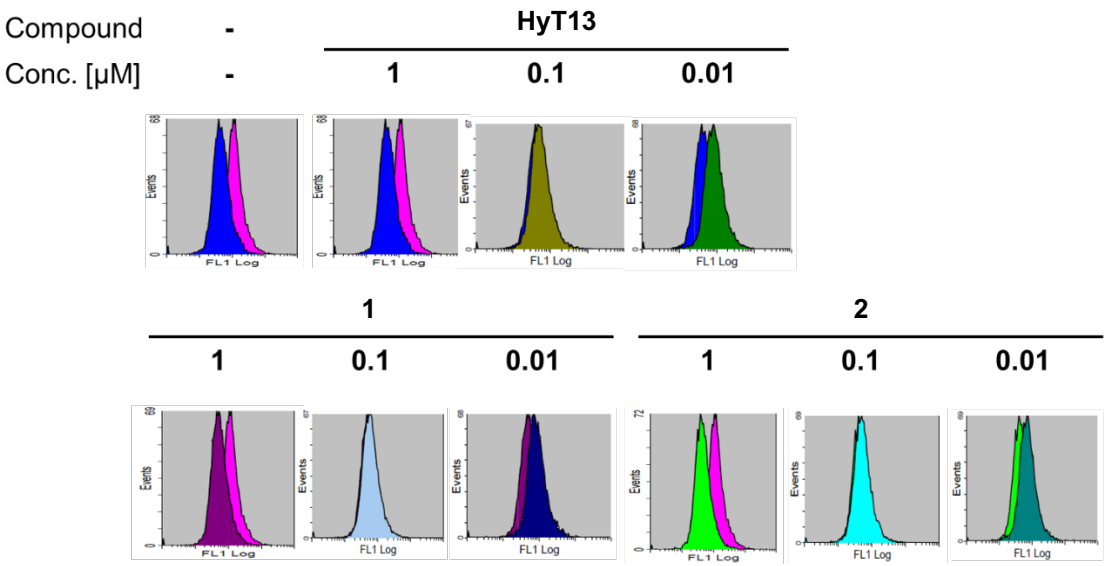


Figure S1. Flow cytometric analysis of HA–HaloTag2-EGFP levels in Flp-In 293 cells treated with HyT13 and compounds **1** and **2**.

Table S1 Condition of electroporation for FITC-BSA

Condition[a]	Voltage [V]	Pulse length [ms]	Interval [ms]	Number	Decay rate [%]	Polarity
1	150	10	50	5	20	+
2	150	20	50	5	20	+
3	150	10	50	1	20	+
4	150	30	50	5	20	+
5	150	10	50	5	10	+
6	150	20	50	5	10	+

[a] Condition for poring pulse. Condition for transfer pulse was constant as follows: transfer pulse; pulse voltage, 20 V; pulse length, 50 ms; pulse interval, 50 ms; pulse number, 20; decay rate, 30; polarity, +/-.

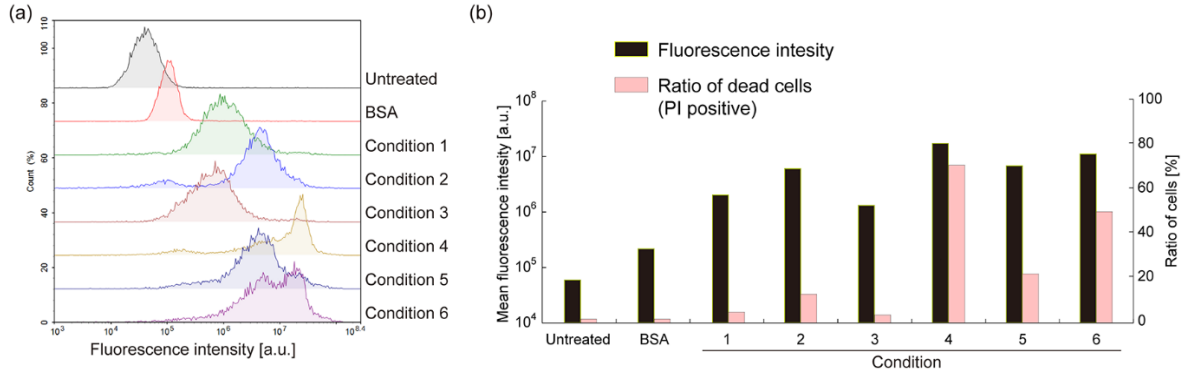


Figure S2 (a) Histograms of LNCaP cells in which FITC-BSA (1 mg/mL) was treated for 24 h and electroporated under various condition (Table S1). (b) Mean fluorescence intensity of FITC-BSA internalized into LNCaP, the ratio of dead cells (PI positive).

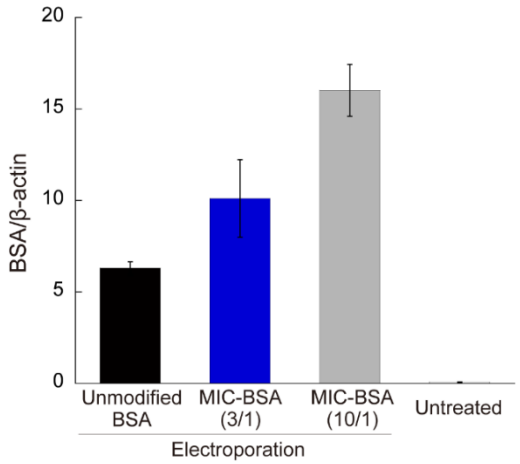


Figure S3 BSA/β-actin values in LNCaP cells electroporated with unmodified BSA, MIC-BSA(3/1), and MIC-BSA(10/1).

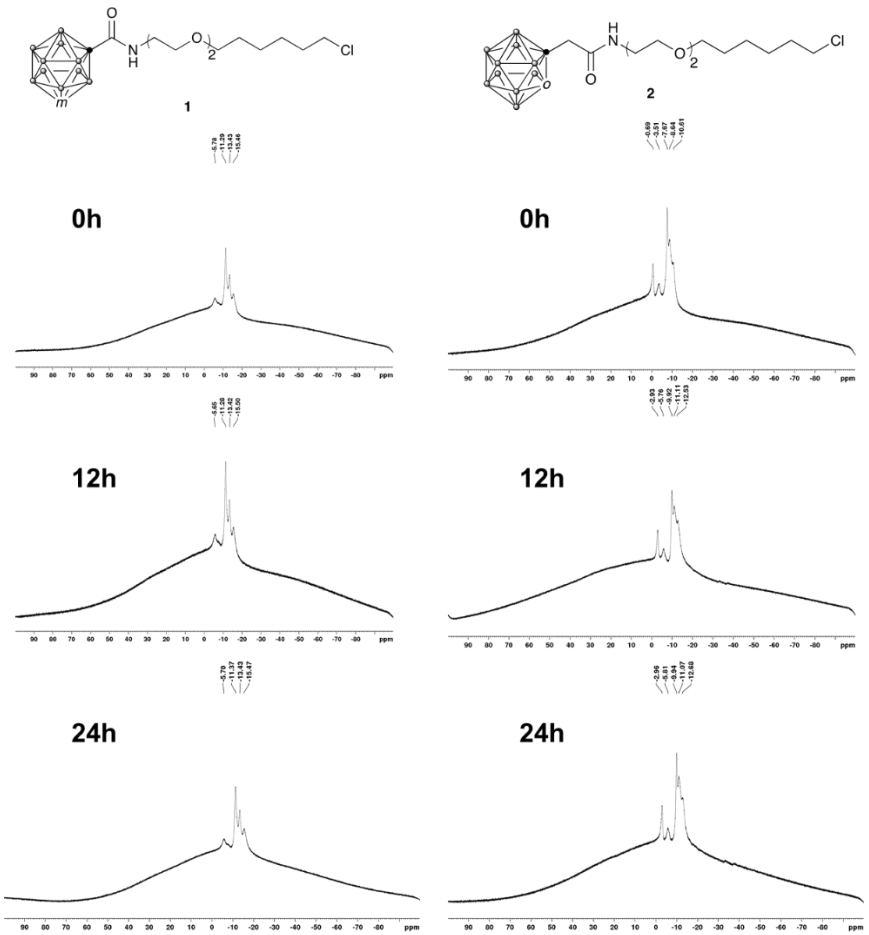


Figure S4 Stability experiments of compounds **1** and **2** in DMSO/H_2O (3/1) monitored by ^{11}B NMR spectroscopy (160 MHz, DMSO-d_6) for 24 h.[5] The *m*-carborane derivative **1** was very stable, whereas the *o*-carborane derivative **2** was slightly deboronized (-33 and -37 ppm) in a time-dependent manner (up to 24h).

3. Reference

(1) Neklesa, T. K.; Tae, H. S.; Schneekloth, A. R.; Stulberg, M. J.; Corson, T. W.; Sundberg, T. B.; Raina, K.; Holley, S. A.; Crews, C. M., Small-molecule hydrophobic tagging-induced degradation of HaloTag fusion proteins. *Nat. Chem. Biol.* **2011**, *7* (8), 538–543.

(2) Tae, H. S.; Sundberg, T. B.; Neklesa, T. K.; Noblin, D. J.; Gustafson, J. L.; Roth, A. G.; Raina, K.; Crews, C. M., Identification of hydrophobic tags for the degradation of stabilized proteins. *ChemBioChem* **2012**, *13* (4), 538–541.

(3) Scholz, M.; Blobaum, A. L.; Marnett, L. J.; Hey-Hawkins, E., *ortho*-Carbaborane Derivatives of Indomethacin as Cyclooxygenase (COX)-2 Selective Inhibitors. *Bioorganic Med. Chem.* **2012**, *20* (15), 4830–4837.

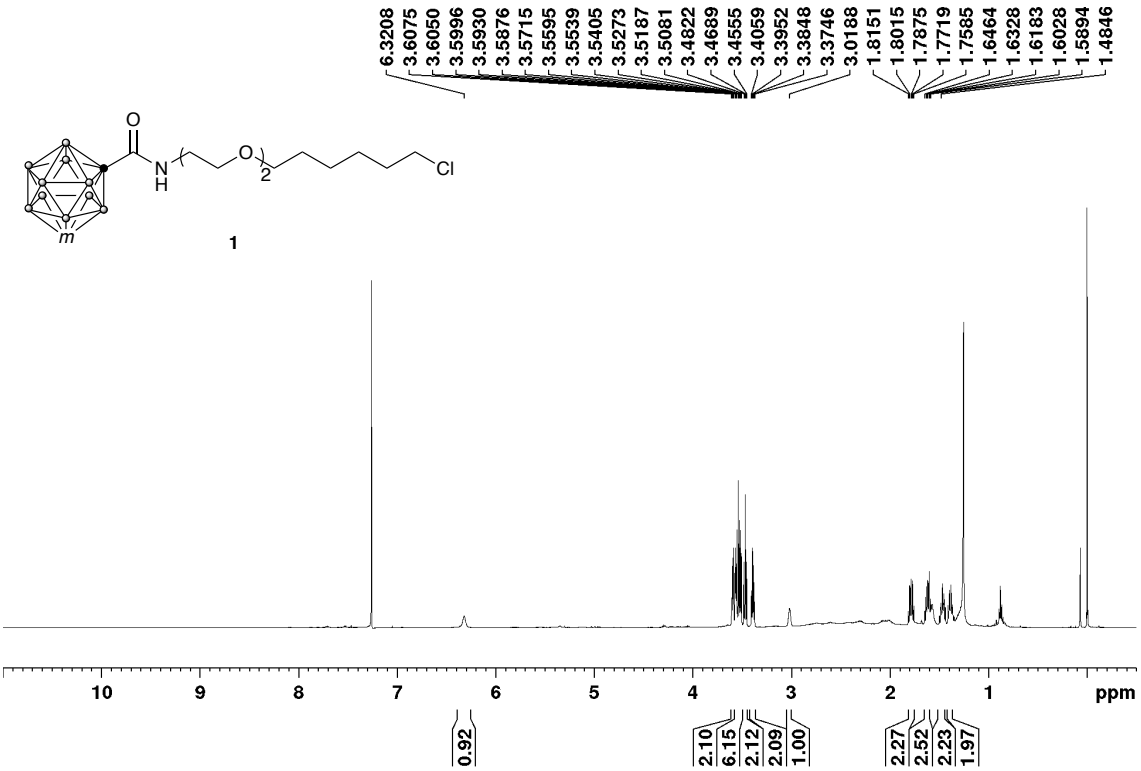
(4) Brozek, E. M.; Mollard, A. H.; Zharov, I., Silica nanoparticles carrying boron-containing polymer brushes. *J. Nanoparticle Res.* **2014**, *16* (4), 2407.

(5) Neumann, W.; Xu, S.; Sárosi, M. B.; Scholz, M. S.; Crews, B. C.; Ghebreselasie, K.; Banerjee, S.; Marnett, L. J.; Hey-Hawkins, E. Nido-Dicarbaborate Induces Potent and Selective Inhibition of Cyclooxygenase-2. *ChemMedChem* **2016**, *11* (2), 175–178.

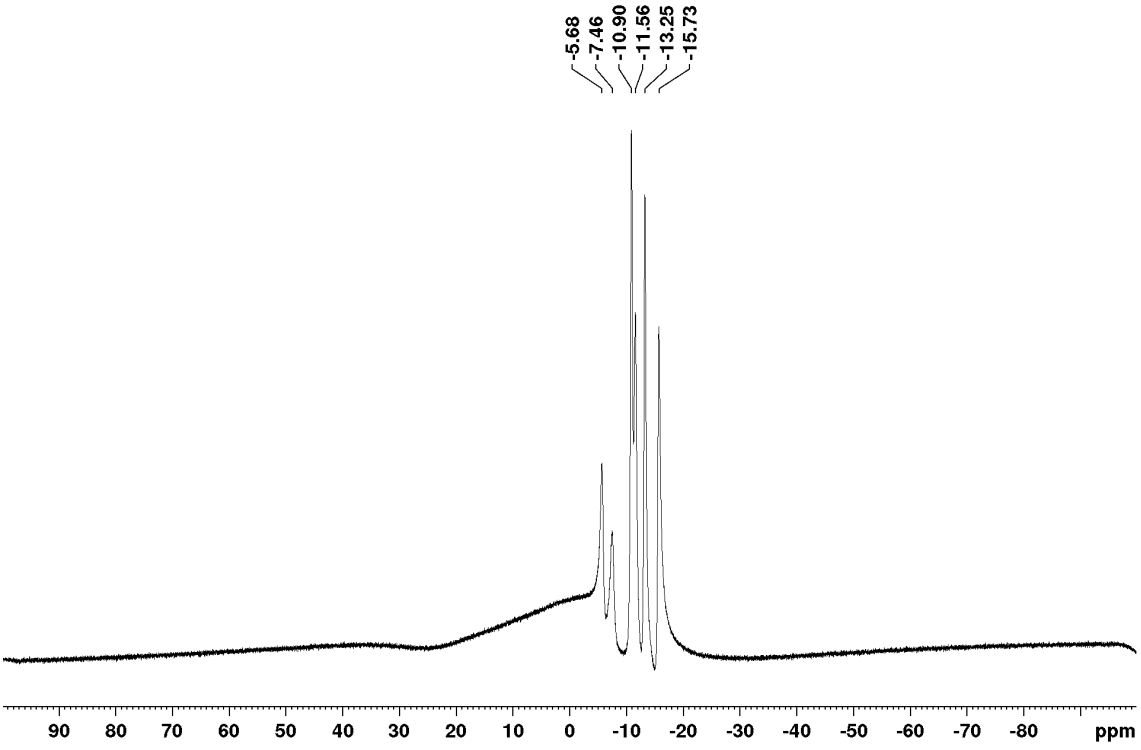
4. Spectra data

***N*-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-1,7-dicarba-*closo*-dodecaborane-1-carboxamide (3)**

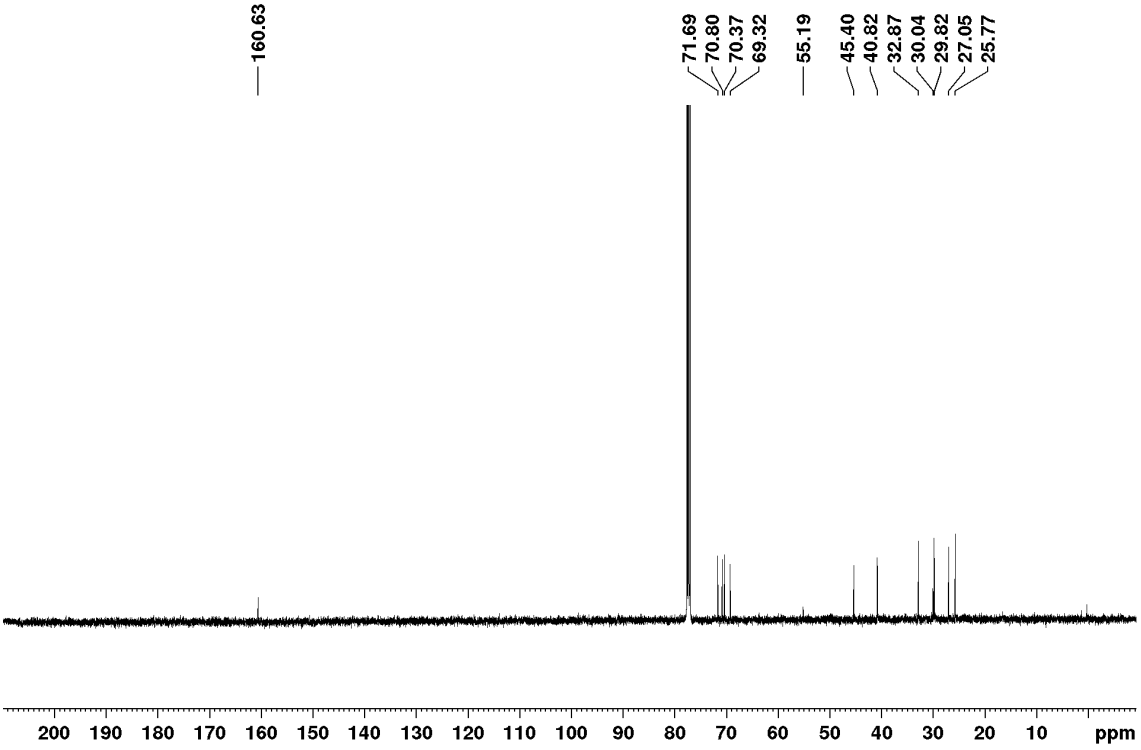
^{1}H NMR (400 MHz, $CDCl_3$) (**1**)



^{11}B NMR (160 MHz, $CDCl_3$)

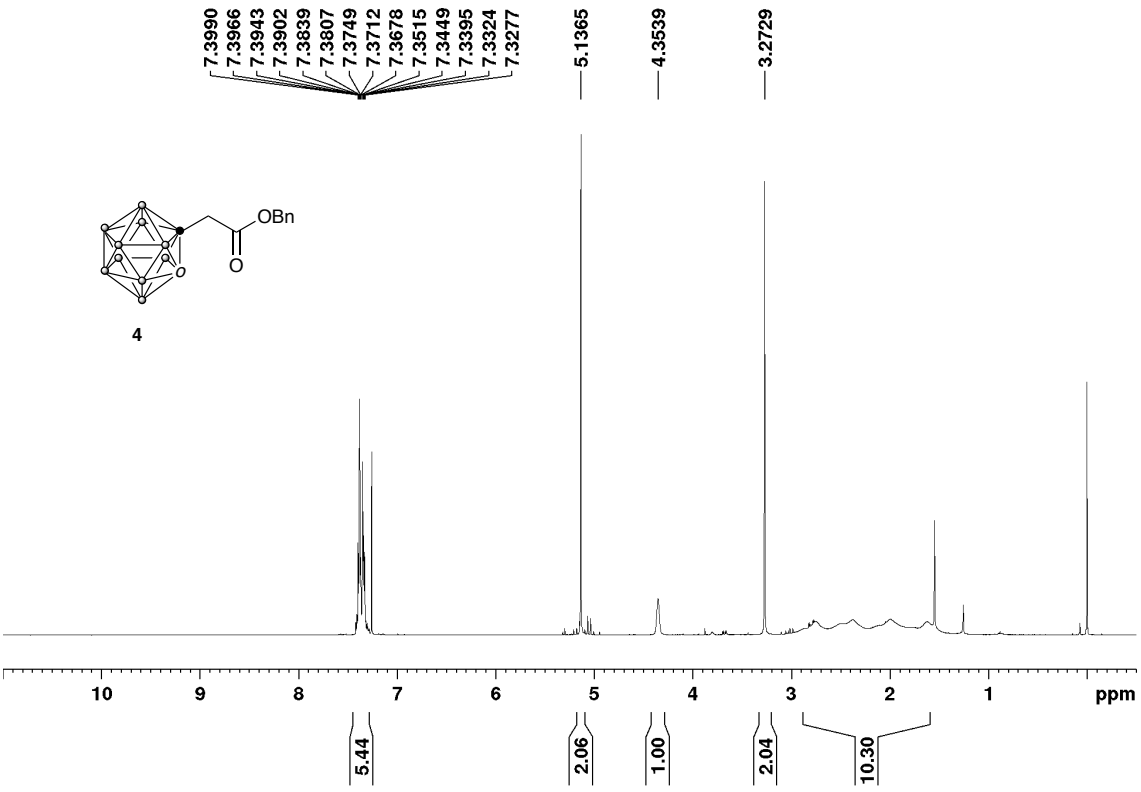


^{13}C NMR (125 MHz, $CDCl_3$) (**1**)

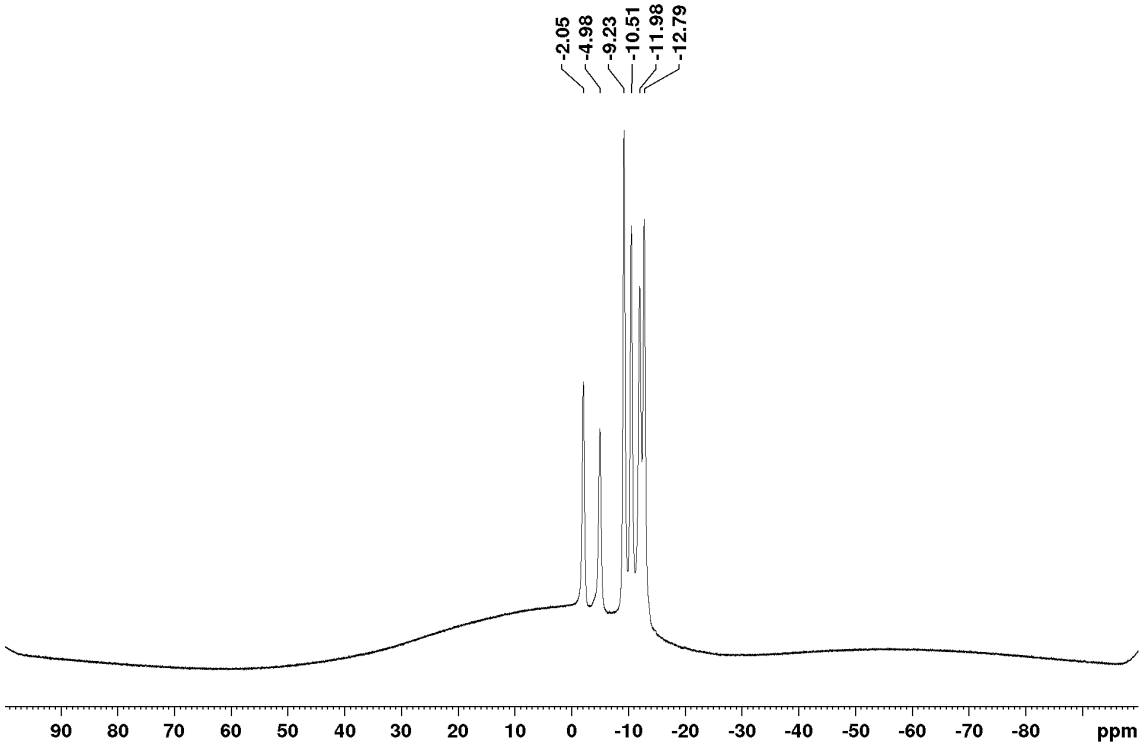


1-(2-(benzyloxy)-2-oxoethyl)-1,2-dicarba-*closo*-dodecaborane (4)

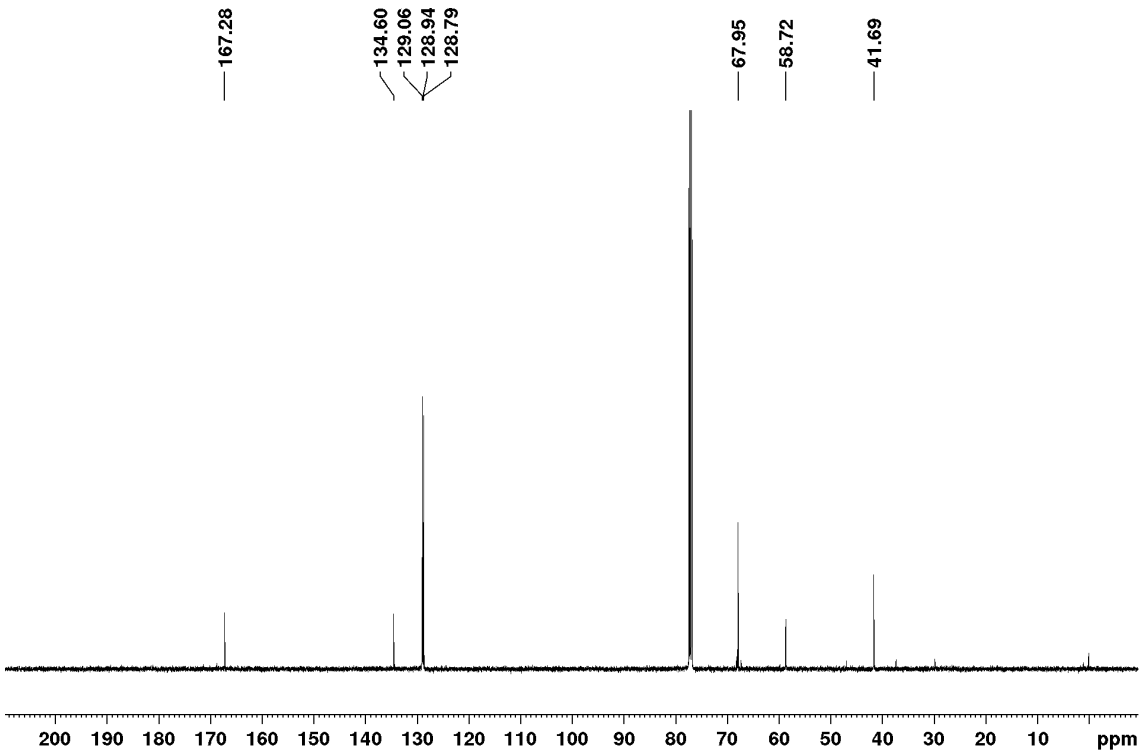
^{1}H NMR (400 MHz, $CDCl_3$) (**4**)



^{11}B NMR (160 MHz, $CDCl_3$) (**4**)

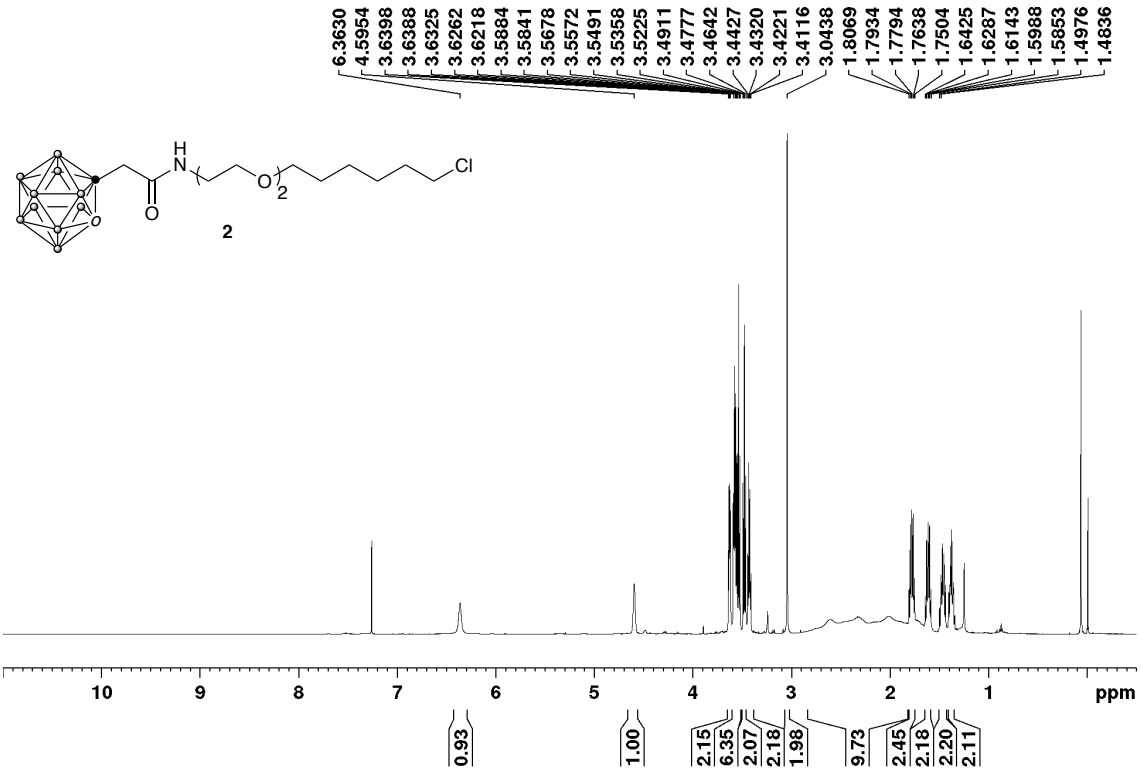


^{13}C NMR (125 MHz, $CDCl_3$) (**4**)

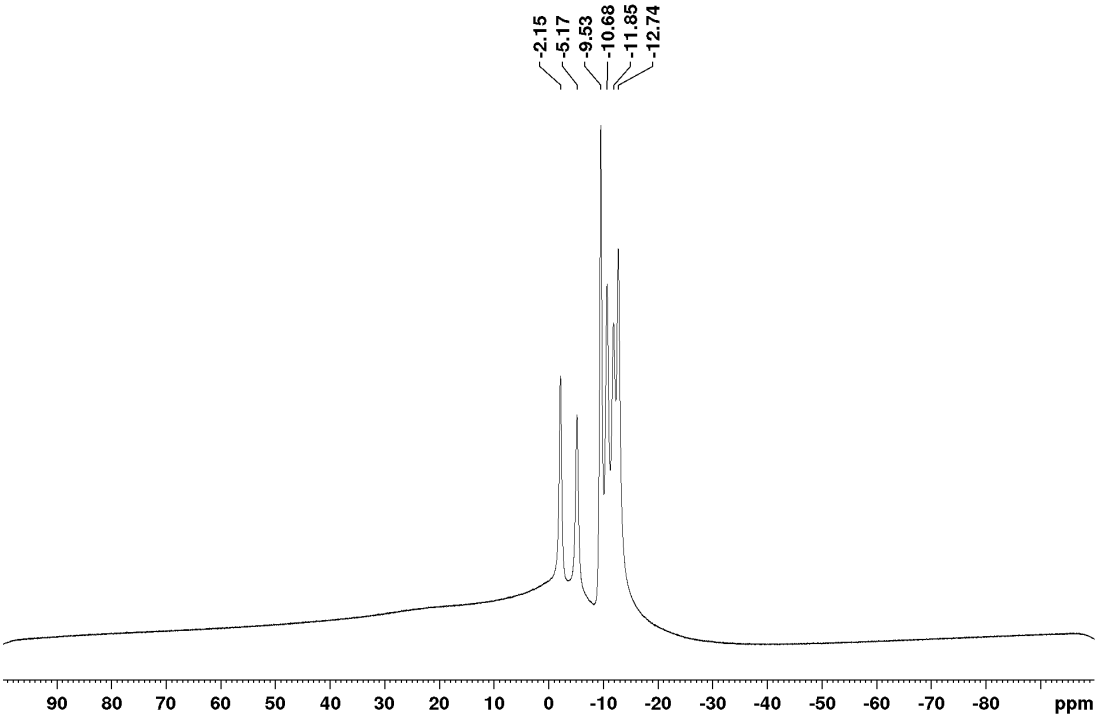


***N*-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-1,7-dicarba-*closo*-dodecaborane-1- acetamide (2)**

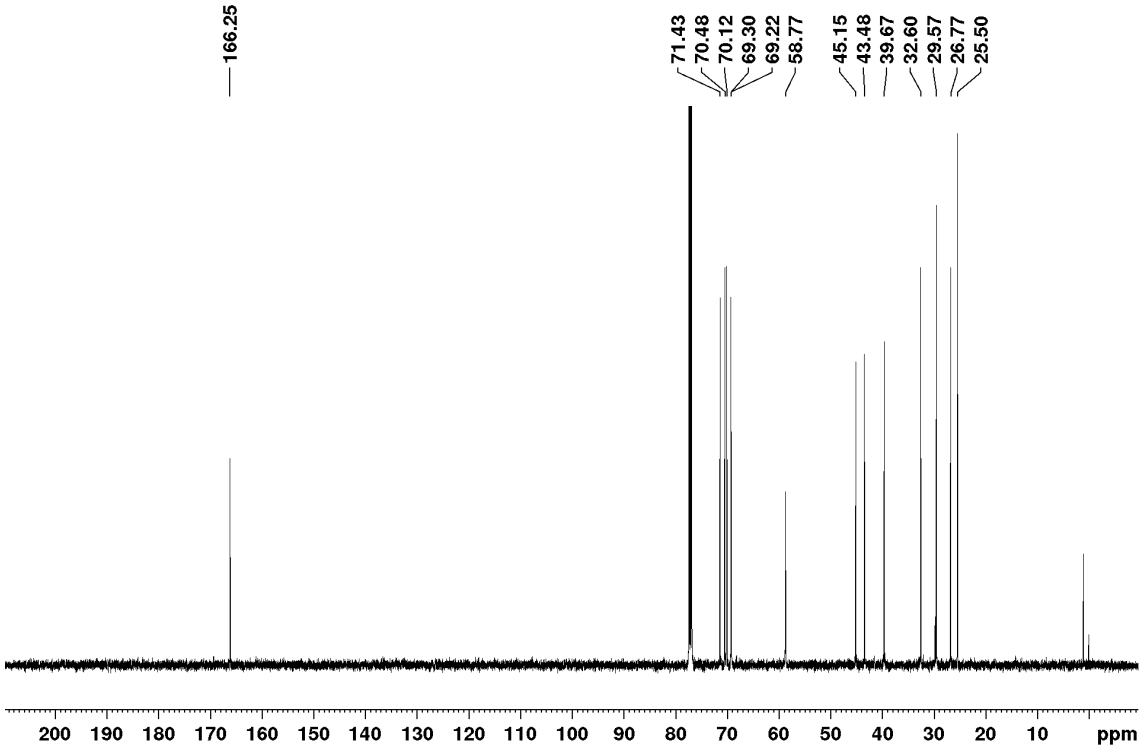
^{1}H NMR (400 MHz, $CDCl_3$) (**2**)



^{11}B NMR (160 MHz, $CDCl_3$) (**2**)

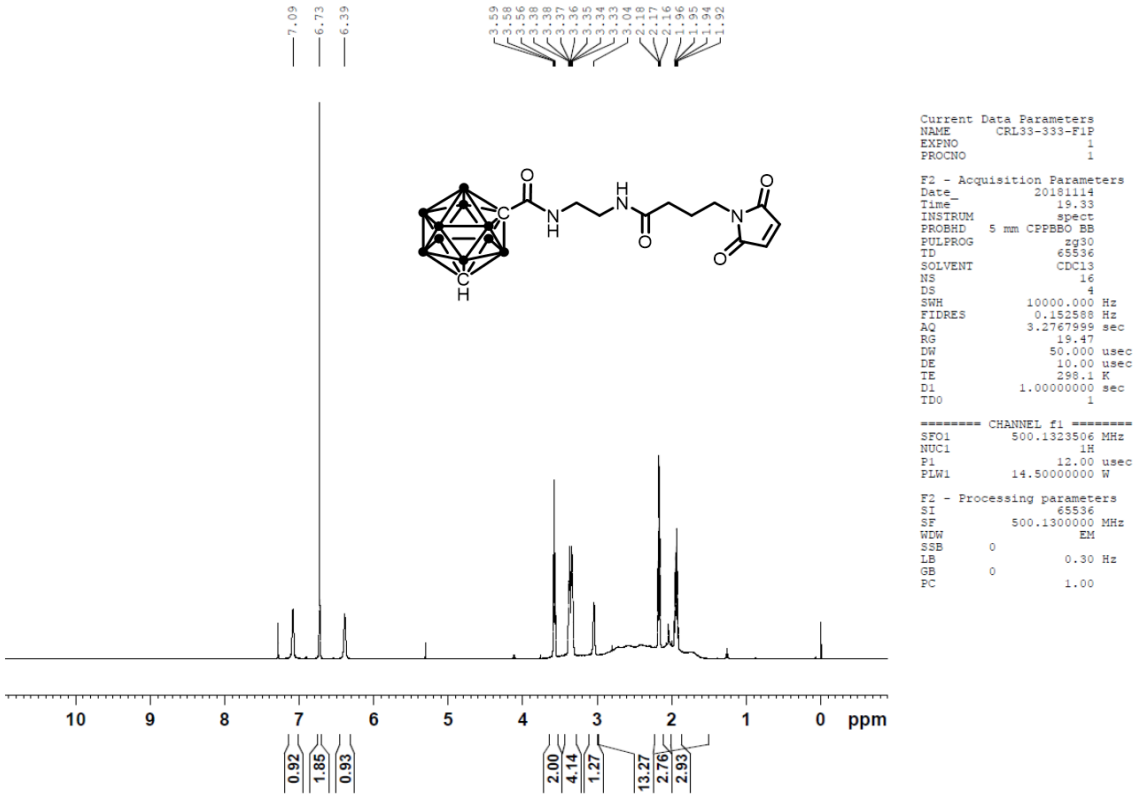


^{13}C NMR (125 MHz, $CDCl_3$) (**2**)

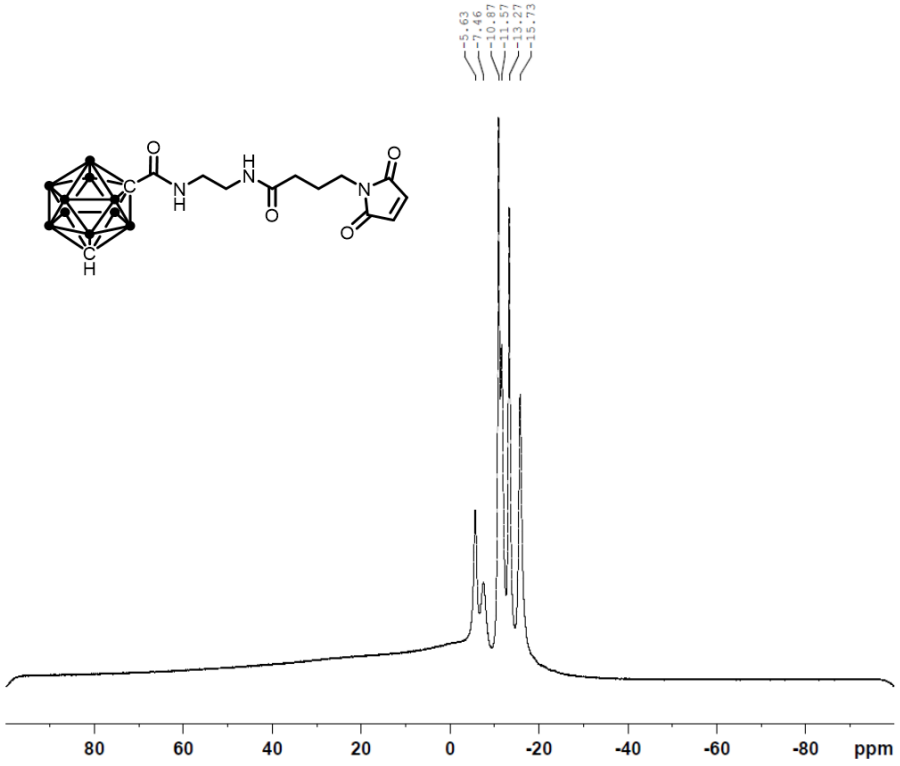


1-((2-(4-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)butanamido)ethyl)carbamoyl)-1,7-dicarba-*closo*-dodecaborane (MIC)

^{1}H NMR (400 MHz, $CDCl_3$)

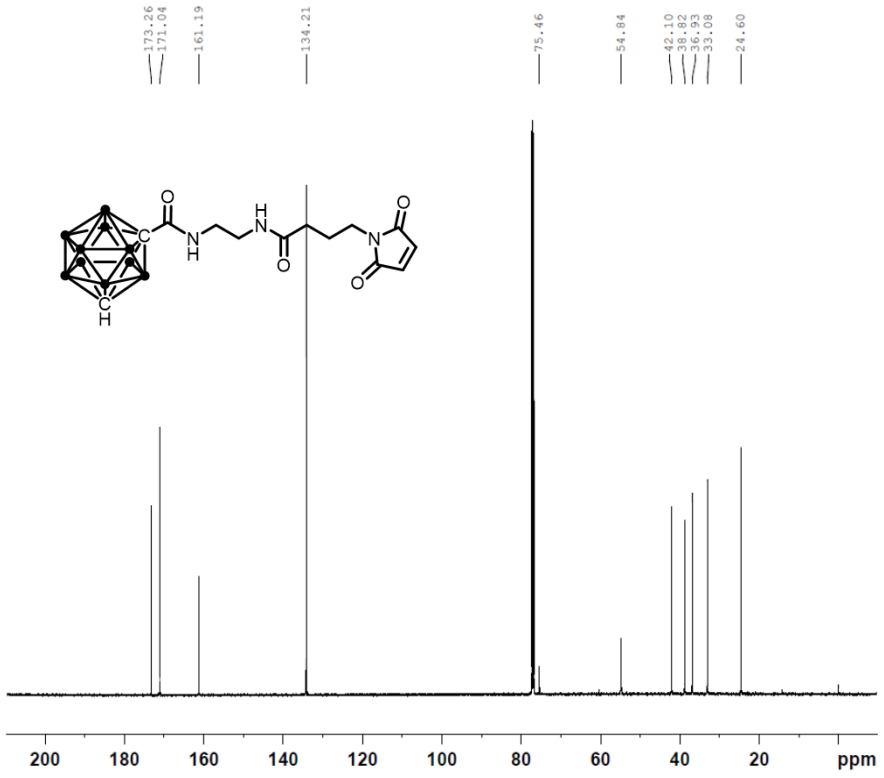


^{11}B NMR (160 MHz, $CDCl_3$) (**MIC**)




```
Current Data Parameters
NAME       CRL33-333-F1B
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20181114
Time_              19.36
INSTRUM            spect
PROBHD   5 mm CPPBBO BB
PULPROG          zgpg30
TD                65536
SOLVENT           CDCl3
NS                   43
DS                    4
SWH           32051.281 Hz
FIDRES        0.489064 Hz
AQ            1.0223616 sec
RG               191.56
DW               15.600 usec
DE                18.00 usec
TE                298.1 K
D1           1.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        160.4615790 MHz
NUC1               11B
P1               11.93 usec
PLW1        67.00000000 W

======== CHANNEL f2 ========
SFO2        500.1323506 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2             80.00 usec
PLW2        14.50000000 W
PLW12        0.32624999 W
PLW13        0.20880000 W

F2 - Processing parameters
SI                32768
SF          160.4615790 MHz
WDW                EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

^{13}C NMR (125 MHz, $CDCl_3$) (**MIC**)




```
Current Data Parameters
NAME       CRL33-333-F1C
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20181114
Time_              19.59
INSTRUM            spect
PROBHD   5 mm CPPBBO BB
PULPROG          zgpg30
TD                65536
SOLVENT           CDCl3
NS                  386
DS                    4
SWH           27573.529 Hz
FIDRES        0.420739 Hz
AQ            1.1883861 sec
RG               191.56
DW               18.133 usec
DE                18.00 usec
TE                298.1 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1

======== CHANNEL f1 ========
SFO1        125.7703648 MHz
NUC1               13C
P1               10.00 usec
PLW1        55.70000076 W

======== CHANNEL f2 ========
SFO2        500.1323506 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2             80.00 usec
PLW2        14.50000000 W
PLW12        0.32624999 W
PLW13        0.20880000 W

F2 - Processing parameters
SI                32768
SF          125.7577890 MHz
WDW                EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```