# Supporting Information for:

## Mining Natural Products for Macrocycles to Drug Difficult Targets

Fabio Begnini,[1,#] Vasanthanathan Poongavanam,[1,#] Björn Over,[2,#] Marie Castaldo,[3] Stefan Geschwindner,[4] Patrik Johansson,[4] Mohit Tyagi,[1] Christian Tyrchan,[5] Lisa Wissler,[4] Peter Sjö,[6] Stefan Schiesser[5,*] and Jan Kihlberg[1,*]

[1] Department of Chemistry - BMC, Uppsala University, Box 576, SE-751 23 Uppsala, Sweden

[2] Department of Medicinal Chemistry, Research and Early Development, Early Cardiovascular, Renal and Metabolism, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

[3] Discovery Biology, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden

[4] Structure, Biophysics & Fragment-Based Lead Generation, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden

[5] Department of Medicinal Chemistry, Research and Early Development, Early Respiratory and Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

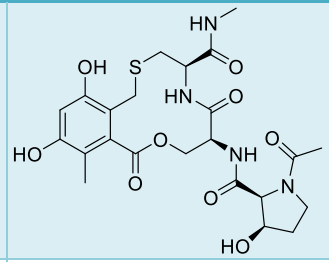
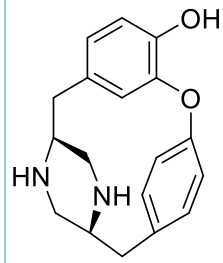
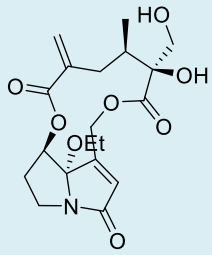
[6] Drugs for Neglected Diseases initiative (DNDi), 15 Chemin Louis Dunant, 1202 Geneva, Switzerland

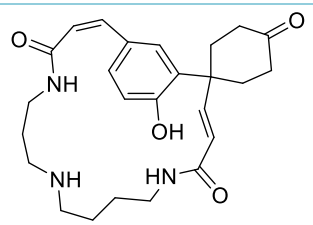
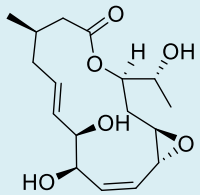
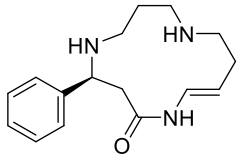
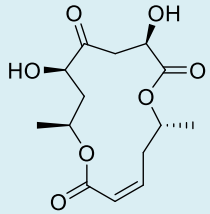
[#] Equally contributing authors

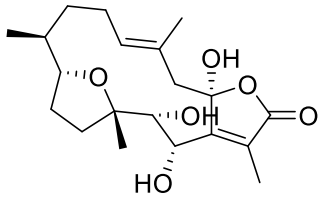
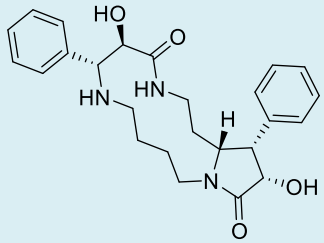
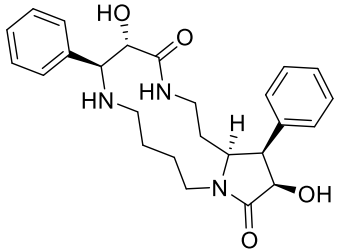
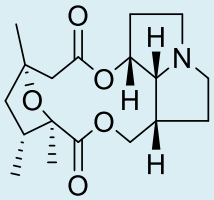
* Corresponding authors: stefan.schiesser@astrazeneca.com , jan.kihlberg@kemi.uu.se

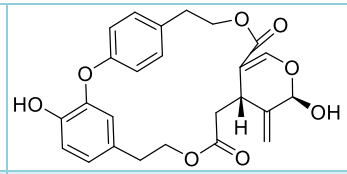
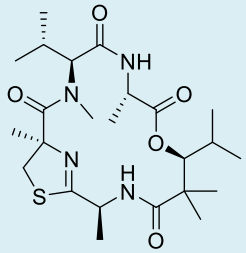
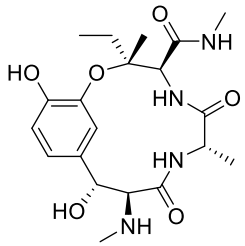
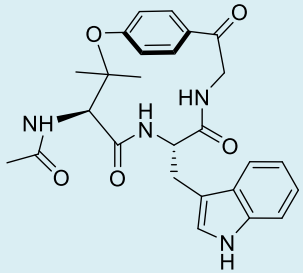
# Supporting Tables

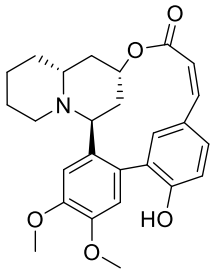
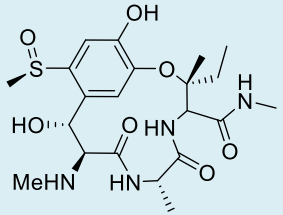
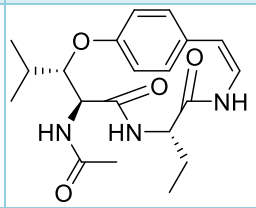
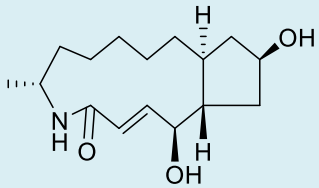
**Table S1**. Structures of the cores in the lead-like set and their molecular descriptors, calculated using MOE [a,b]

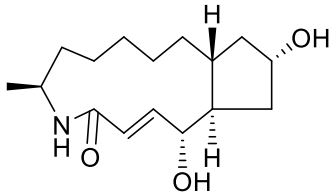
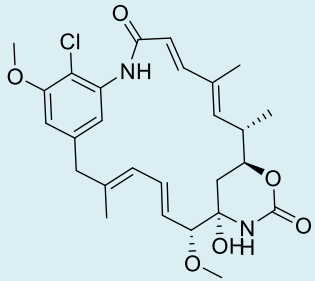
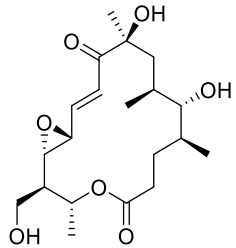
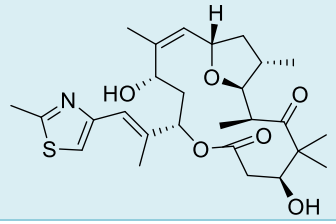
| Core No. | Natural Product | Structure | MW | HBA | HBD | NRotB | cLogP | TPSA |
|---|---|---|---|---|---|---|---|---|
| **1** | Cyclothialidine |  | 538.6 | 8 | 6 | 6 | -1.4 | 194.6 |
| **2** | Piperazinomycin |  | 296.4 | 3 | 3 | 0 | 2.5 | 53.5 |
| **3** | 8-Ethoxy-3-oxo-1,2-dehydroretrorsine |  | 395.4 | 6 | 2 | 3 | 0.5 | 122.6 |

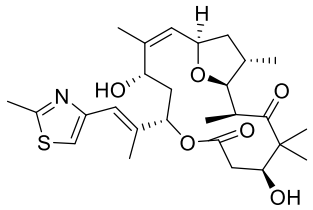
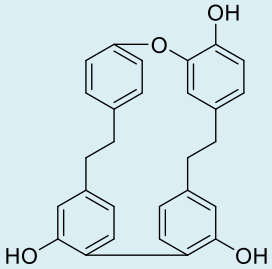
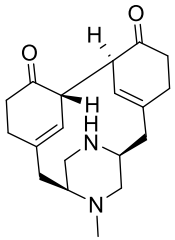
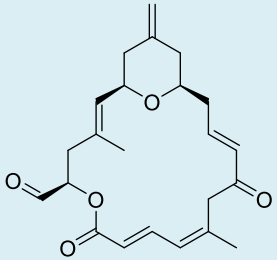
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **4** | Numismine |  | 439.6 | 5 | 4 | 0 | 2.8 | 107.5 |
| **5** | Iriomoteolide 3a |  | 326.4 | 5 | 3 | 1 | 1.0 | 99.5 |
| **6** | 4-Phenyl-1,5,9-triazacyclotridec-12-en-2-one |  | 273.4 | 3 | 3 | 1 | 1.0 | 53.2 |
| **7** | Grahamimycin B |  | 300.3 | 5 | 2 | 0 | 0.1 | 110.1 |

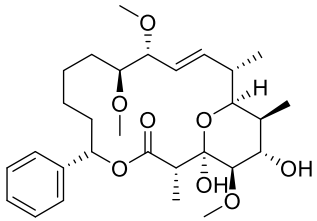
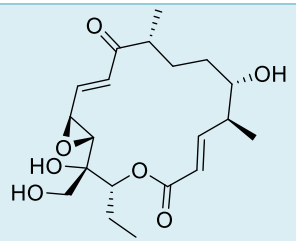
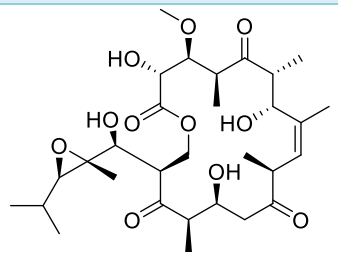
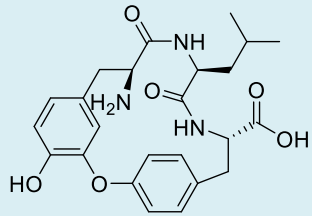
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **8** | 4,7-Epoxy-2,3,14-trihydroxy-1(15),11-cembradien-16,14-olide |  | 366.5 | 5 | 3 | 0 | 2.3 | 96.2 |
| **9A** | 7-Hydroxypleurocorine Enantiomer A |  | 437.5 | 5 | 4 | 2 | 1.6 | 101.9 |
| **9B** | 7-Hydroxypleurocorine Enantiomer B |  | 437.5 | 5 | 4 | 2 | 1.6 | 101.9 |
| **10** | Nemorensine |  | 337.4 | 4 | 0 | 0 | 2.0 | 65.1 |

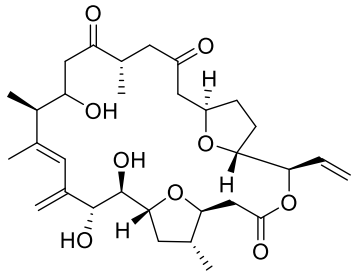
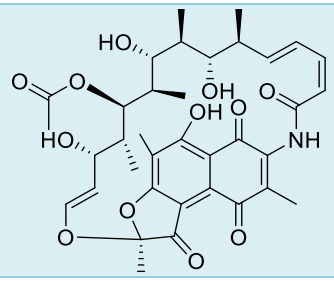
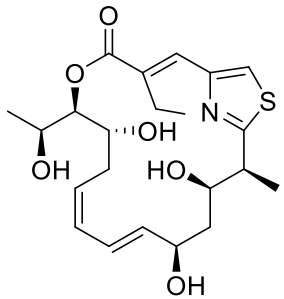
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **11** | Uhdoside A |  | 452.5 | 5 | 2 | 0 | 2.8 | 111.5 |
| **12** | Halipeptin A |  | 496.7 | 5 | 2 | 2 | 2.5 | 117.2 |
| **13** | Ustiloxin F |  | 422.5 | 7 | 6 | 4 | 0.0 | 149.0 |
| **14** | Hymenocardine |  | 476.5 | 5 | 4 | 4 | 2.1 | 129.4 |

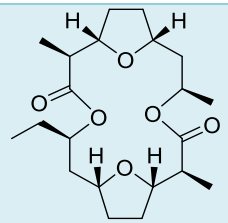
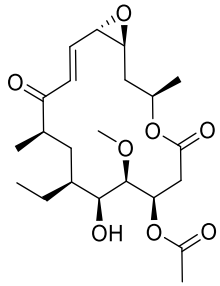
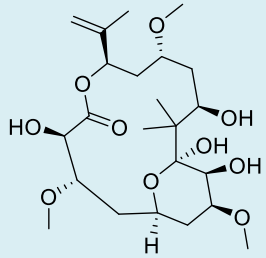
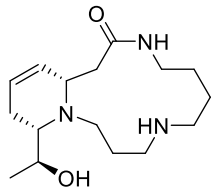
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **15** | Vertine |  | 435.5 | 5 | 1 | 2 | 5.5 | 68.2 |
| **16** | Ustiloxin B |  | 484.6 | 8 | 6 | 5 | -1.3 | 185.3 |
| **17** | Discarine M |  | 373.5 | 4 | 3 | 4 | 2.0 | 96.5 |
| **18A** | IFB Lactam 1 Enantiomer A |  | 281.4 | 3 | 3 | 0 | 2.3 | 69.6 |

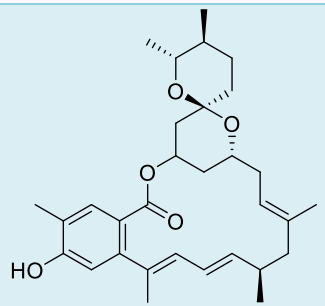
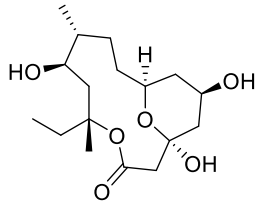
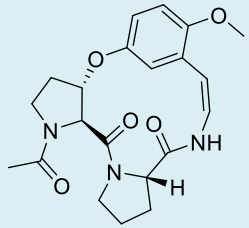
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **18B** | IFB Lactam 1 Enantiomer B |  | 281.4 | 3 | 3 | 0 | 2.3 | 69.6 |
| **19** | Maysenine |  | 517.0 | 5 | 3 | 2 | 4.6 | 106.1 |
| **20** | Aldgamycin C |  | 384.5 | 6 | 3 | 1 | 1.6 | 116.6 |
| **21A** | Epothilone tetrahydrofuran Enantiomer A |  | 505.7 | 6 | 2 | 2 | 3.2 | 106.0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **21B** | Epothilone tetrahydrofuran Enantiomer B |  | 505.7 | 6 | 2 | 2 | 3.2 | 106.0 |
| **22** | Polymorphatin A |  | 424.5 | 3 | 3 | 0 | 7.0 | 69.9 |
| **23** | Herquline b |  | 314.4 | 4 | 1 | 0 | -0.7 | 49.4 |
| **24** | Dactylolide |  | 384.5 | 4 | 0 | 1 | 2.6 | 69.7 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **25** | Soraphen A |  | 520.7 | 7 | 2 | 4 | 4.8 | 103.7 |
| **26** | Mycinamycin XII |  | 382.5 | 6 | 3 | 2 | 1.5 | 116.6 |
| **27** | Tedanolide |  | 584.7 | 10 | 4 | 4 | 0.2 | 180.2 |
| **28** | Renieramide |  | 427.5 | 6 | 6 | 2 | 1.6 | 151.0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **29** | Amphidinolide C |  | 562.7 | 8 | 3 | 1 | 2.2 | 139.6 |
| **30** | Antibiotic A 39079S-1 |  | 681.7 | 11 | 5 | 2 | 3.7 | 206.0 |
| **31** | Thuggacin B |  | 437.6 | 6 | 4 | 2 | 1.9 | 120.1 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **32** | Homononactic nonactic dilactone |  | 382.5 | 4 | 0 | 1 | 4.2 | 71.1 |
| **33** | Deltamycin X |  | 426.5 | 6 | 1 | 4 | 2.1 | 111.7 |
| **34** | Peloruside A |  | 476.6 | 9 | 4 | 4 | 1.7 | 144.1 |
| **35** | Palustrine |  | 295.4 | 4 | 3 | 1 | 0.4 | 64.6 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **36** | Milbemycin □3 |  | 494.7 | 4 | 1 | 0 | 7.6 | 65.0 |
| **37** | Lyngbouilloside |  | 330.4 | 5 | 3 | 1 | 2.9 | 96.2 |
| **38** | Oxyphylline D |  | 399.4 | 5 | 1 | 2 | 0.8 | 88.2 |

[a]The name of the original natural product is given in column 2.· MW, molecular weight; HBA, hydrogen bond acceptor; HBD, hydrogen bond donor; NRotB, number of rotatable bonds; cLogP, calculated lipophilicity; cLogS, TPSA, topological polar surface area.

[b]SMILES codes for this set of 41 cores are given in Supporting Excel Sheet, together with the SMILES codes for the set of 217 cores.

**Table S2.** Summary of Glide scores for docking of the 41 lead-like macrocyclic cores into the four selected Keap1 crystal structures (PDB ID: 4IQK, 3VNG, 4IFJ and 1ZGK).[a,b]

| Core No. | Natural Product | Charge | 4IQK | 3VNG | 4IFJ | 1ZGK |
|---|---|---|---|---|---|---|
| **1** | Cyclothialidine | 0 | -5.27 | -6.13 | -5.06 | -4.49 |
| **2** | Piperazinomycin | 1 | -4.81 | -5.70 | -6.21 | -7.31 |
| **2** | Piperazinomycin | 1 | -4.95 | -3.98 | -3.96 | -6.10 |
| **2** | Piperazinomycin | 0 | -4.17 | -5.43 | -4.14 | -3.38 |
| **2** | Piperazinomycin | 0 | -3.99 | -3.44 | -3.75 | - |
| **3** | 8-Ethoxy-3-oxo-1,2-dehydroretrorsine | 0 | -4.99 | -5.23 | -4.97 | -3.25 |
| **4** | Numismine | 1 | -4.38 | -4.71 | -4.46 | -4.67 |
| **4** | Numismine | 0 | -3.94 | -5.35 | -4.41 | -5.23 |
| **5** | Iriomoteolide 3a | 0 | -5.88 | -5.05 | -5.02 | -5.21 |
| **6** | 4-Phenyl-1,5,9-triazacyclotridec-12-en-2-one | 0 | -4.60 | -3.72 | -4.84 | -4.80 |
| **6** | 4-Phenyl-1,5,9-triazacyclotridec-12-en-2-one | -1 | -4.20 | -4.16 | -4.80 | -5.00 |
| **7** | Grahamimycin B | 0 | -5.19 | -5.24 | -4.16 | -4.15 |
| **8** | 4,7-Epoxy-2,3,14-trihydroxy-1(15),11-cembradien-16,14-olide | 2 | -4.32 | -3.39 | -3.36 | -3.60 |
| **8** | 4,7-Epoxy-2,3,14-trihydroxy-1(15),11-cembradien-16,14-olide | 1 | -4.76 | -3.97 | -4.04 | -4.01 |
| **9A** | 7-Hydroxypleurocorine Enantiomer A | 1 | -4.90 | -2.76 | -3.87 | -3.96 |
| **9A** | 7-Hydroxypleurocorine Enantiomer A | 0 | -5.34 | -3.26 | -3.56 | -3.59 |
| **9B** | 7-Hydroxypleurocorine Enantiomer B | 1 | -6.52 | -3.80 | -4.08 | -3.86 |
| **9B** | 7-Hydroxypleurocorine Enantiomer B | 0 | -5.28 | -3.21 | -4.75 | -3.58 |
| **10** | Nemorensine | 1 | -3.69 | -3.99 | -3.79 | -3.93 |
| **11** | Uhdoside A | 0 | -5.89 | -4.52 | -3.27 | -3.95 |
| **11** | Uhdoside A | -1 | -3.18 | -4.06 | -3.12 | -2.95 |
| **12** | Halipeptin A | 0 | -3.38 | -4.02 | -3.33 | -3.66 |
| **13** | Ustiloxin F | 1 | -4.25 | -3.99 | -4.19 | -3.47 |
| **13** | Ustiloxin F | 0 | -4.19 | -4.59 | -2.76 | -2.81 |
| **14** | Hymenocardine | 0 | -4.63 | -4.20 | -4.15 | -4.80 |
| **15** | Vertine | 1 | -4.72 | -3.68 | -3.58 | -3.56 |
| **15** | Vertine | 0 | -3.70 | -3.72 | -3.83 | -4.06 |
| **16** | Ustiloxin B | 1 | -4.76 | -6.68 | -3.92 | -4.59 |
| **16** | Ustiloxin B | 0 | -4.62 | -7.10 | -4.33 | -3.90 |
| **16** | Ustiloxin B | (+/-) | -4.82 | -6.38 | -3.99 | -4.18 |
| **17** | Discarine M | 0 | -5.00 | -3.48 | -3.78 | -4.05 |
| **18A** | IFB Lactam 1 Enantiomer A | 0 | -3.67 | -4.93 | -3.99 | - |

| | | | | | | |
|---|---|---|---|---|---|---|
| **18B** | IFB Lactam 1 Enantiomer B | 0 | -4.69 | -4.50 | -4.76 | - |
| **19** | Maysenine | 0 | -4.56 | -4.45 | -4.32 | -4.50 |
| **20** | Aldgamycin C | 0 | -3.66 | -4.20 | -3.48 | -3.83 |
| **21A** | Epothilone tetrahydrofuran Enantiomer A | 0 | -4.96 | -4.07 | -3.05 | -3.59 |
| **21B** | Epothilone tetrahydrofuran Enantiomer B | 0 | -4.96 | -4.07 | -3.05 | -3.59 |
| **22** | Polymorphatin A | 0 | -4.30 | -3.96 | -5.44 | -3.96 |
| **22** | Polymorphatin A | -1 | -3.94 | -4.06 | -5.20 | -3.72 |
| **22** | Polymorphatin A | -1 | -3.88 | -3.59 | -3.47 | -4.12 |
| **22** | Polymorphatin A | -1 | -3.45 | -3.25 | -3.58 | -3.03 |
| **23** | Herquline b | 1 | -4.52 | -3.12 | -4.13 | -2.81 |
| **23** | Herquline b | 1 | -4.52 | -3.44 | -4.16 | -4.12 |
| **23** | Herquline b | 0 | -4.51 | -4.42 | -5.03 | -3.76 |
| **24** | Dactylolide | 0 | -4.97 | -3.89 | -3.93 | -3.58 |
| **24** | Dactylolide | -1 | -3.98 | -2.80 | -2.83 | -3.22 |
| **25** | Soraphen A | 0 | -2.60 | -2.58 | -3.98 | -3.87 |
| **26** | Mycinamycin XII | 0 | -4.98 | -4.55 | -3.88 | -3.58 |
| **27** | Tedanolide | 0 | -3.19 | -5.89 | -3.64 | -3.63 |
| **28** | Renieramide | 0 | -4.22 | -4.19 | -3.21 | -4.13 |
| **28** | Renieramide | -1 | -4.81 | -4.97 | -3.17 | -3.70 |
| **28** | Renieramide | -1 | -4.23 | -4.50 | -3.21 | -3.36 |
| **29** | Amphidinolide C | 0 | -2.39 | -3.11 | -2.73 | -2.97 |
| **30** | Antibiotic A 39079S-1 | -1 | -5.67 | -4.80 | -3.93 | -3.92 |
| **30** | Antibiotic A 39079S-1 | 0 | -4.43 | -4.03 | -4.00 | -4.39 |
| **30** | Antibiotic A 39079S-1 | 0 | -2.93 | -3.68 | -3.45 | -3.01 |
| **31** | Thuggacin B | 0 | -3.64 | -4.33 | -4.46 | -4.48 |
| **32** | Homononactic nonactic dilactone | 0 | -3.48 | -4.14 | -4.13 | -3.94 |
| **33** | Deltamycin X | 0 | -4.61 | -4.53 | -3.56 | -3.48 |
| **34** | Peloruside A | 0 | -4.13 | -3.62 | -3.39 | -2.81 |
| **35** | Palustrine | 1 | -5.07 | -4.14 | -4.46 | -5.53 |
| **35** | Palustrine | 2 | -3.65 | -4.79 | -4.24 | -4.58 |
| **36** | Milbemycin β3 | 0 | -2.72 | -2.63 | -2.57 | -2.54 |
| **37** | Lyngbouilloside | 0 | -4.25 | -4.12 | -4.55 | -4.72 |
| **38** | Oxyphylline D | 0 | -4.48 | -3.61 | -3.32 | -3.64 |

***a*** The Epik tool from LigPrep[1] was used to predict the protonation state at pH 7.0 of ionizable residues in the cores. For some cores two protonation states were obtained, then both states were used in the docking.

***b***Glide scores were determined from the standard precision mode in Glide. For each core the docked pose that had the lowest Glide score for binding to each of the Keap1 structures has been identified, and the score has been listed.

**Table S3.** Combined list of the top-ten cores that had the lowest Glide score for docking into each of the four Keap1 crystal structures (PDB ID: 4IQK, 3VNG, 4IFJ and 1ZGK).[a]

| Core No. | Scaffolds | 4IQK | 3VNG | 4IFJ | 1ZGK | Sum |
|---|---|---|---|---|---|---|
| **1** | ***Cyclothialidine*** | + | + | + | + | ++++ |
| **2** | ***Piperazinomycin*** | - | + | + | + | +++ |
| **3** | ***8-Ethoxy-3-oxo-1,2-dehydroretrorsine*** | + | + | + | - | +++ |
| **4** | ***Numismine*** | - | + | + | + | +++ |
| **5** | ***Iriomoteolide 3a*** | + | + | - | + | +++ |
| **6** | 4-Phenyl-1,5,9-triazacyclotridec-12-en-2-one | - | - | + | + | ++ |
| **7** | Grahamimycin B | + | + | - | - | ++ |
| **9B** | 7-Hydroxypleurocorine Enantiomer B | + | - | - | - | + |
| **11** | Uhdoside A | + | - | - | - | + |
| **14** | Hymenocardine | - | - | - | + | + |
| **16** | Ustiloxin B | - | + | - | + | ++ |
| **17** | Discarine M | + | - | - | - | + |
| **18A** | IFB Lactam 1 Enantiomer A | - | + | + | - | ++ |
| **18B** | IFB Lactam 1 Enantiomer B | - | - | + | - | + |
| **19** | Maysenine | - | - | - | + | + |
| **21A** | Epothilone tetrahydrofuran Enantiomer A | + | - | - | - | + |
| **22** | Polymorphatin A | - | - | + | - | + |
| **26** | Mycinamycin XII | + | - | - | - | + |
| **27** | Tedanolide | - | + | - | - | + |
| **30** | Antibiotic A 39079S-1 | + | - | - | - | + |
| **31** | Thuggacin B | - | - | + | - | + |
| **35** | Palustrine | - | - | - | + | + |
| **37** | Lyngbouilloside | - | - | + | + | ++ |

[a]A + sign indicates that the core was found in the top-ten list of cores that had the lowest Glide score for docking into the crystal structure given at the top of the column. The names of the fives cores that were found on the top-ten list for three or four of the crystal structures are in boldface italic.

**Table S4.** Reported biological activities and synthetic routes for the natural product parents of the top-five cores from docking into Keap1[a]

| **Natural product** | **Biological activity** | **Synthetic route** |
|---|---|---|
| Cyclothialidine | Inhibition of DNA gyrase B ATPase activity. Brvar, M., et al., *Bioorg. Med. Chem. Lett.* **2010**, 20, 3, 958–962 & Boehm, H. J., et al., *J. Med. Chem.* **2000**, 43, 14, 2664-2674. | Goetschi, E., et al., *Helvetica Chimica Acta* **1996**, 79, 8 2219-2234. |
| Piperazinomycin | Inhibition of *Trichophyton mentagrophytes* 833 and *Trichophyton asteroides* 429. Tamai, S., et al., *J. Antibiot*., **1982**, 35, 1130- 1136. | Nishiyama, N., et al., *Tetrahedron Lett.* **1986**, 27, 37, 4484-4484 & Boger, L. D., et al., *J. Am. Chem. Soc.* **1993**, 115, 24, 11426-11433. |
| 8-Ethoxy-3-oxo-1,2-dehydroretrorsineine | ----[b] | ----[b] |
| Numismine | ----[b] | ----[b] |
| Iriomoteolide 3a | Antiproliferative activity *vs* DAUDI and HL-60 cell lines. Cribiu, R., et al., *Angew. Chem. Int. Ed*. **2009**, 48, 8780 –8783. | Cribiu, R., et al., *Angew. Chem. Int. Ed*. **2009**, 48, 8780 –8783. |

[a]Data retrieved from ChEMBL, PubChem and PubMed

[b]No data reported in the literature

**Table S5.** Data collection and refinement statistics for the complex between Keap1 and macrocycle **14**. Values within parenthesis refer to the highest resolution shell.

| **Data collection** | |
|---|---|
| Space group | $P2_12_12_1$ |
| Cell dimensions (Å) | 75.7 75.8 203.1 |
| Resolution (Å) | 47.37 – 2.37 (2.45) |
| Measured reflections | 312117 |
| Unique reflections | 48272 |
| $R_{merge}$ | 0.073 (1.4) |
| $\langle I / \sigma I \rangle$ | 15.4 (1.3) |
| Completeness (%) | 99.9 (99.9) |
| Redundancy | 6.5 (6.2) |
| **Refinement** | |
| Resolution (Å) | 47.37 – 2.37 (2.43) |
| $R_{work}$ / $R_{free}$ | 0.192 / 0.210 |
| No. atoms | |
| Protein | 4732 |
| Water | 182 |
| Ligand | 34 |
| Average *B*-factors | |
| Protein (Å$^2$) | 70.1 |
| Water (Å$^2$) | 68.0 |
| Ligand (Å$^2$) | 78.0 |
| Ramachandran outliers (%) | 0.53 |
| R.m.s deviations | |
| Bond lengths (Å) | 0.010 |
| Bond angles (°) | 1.16 |

**Table S6.** Binding free energies calculated for compound **9**, **12-15** with Prime MM-GBSA. Energies are provided in kcal/mol.[a]

| **Energy terms** | **(9) OMe*** | **(12)NH2** | **(13) NMe*** | **(14) $NMe_2$** | **(15) $NEt_2$*** |
|---|---|---|---|---|---|
| $\mathbf{E_{Coul}}$ | -15.97 | -27.02 | -15.46 | -16.53 | -6.83 |
| $\mathbf{E_{vdW}}$ | -49.58 | -47.05 | -50.37 | -52.47 | -41.82 |
| $\mathbf{E_{Lipo}}$ | -14.78 | -12.76 | -14.66 | -17.29 | -15.79 |
| $\mathbf{E_{Covalent}}$ | 2.99 | 2.98 | 1.69 | 0.94 | 1.65 |
| $\mathbf{E_{HB}}$ | -1.04 | -1.55 | -1.07 | -1.08 | -0.41 |
| $\mathbf{E_{Packing}}$ | -1.48 | -1.49 | -1.37 | -1.42 | -1.72 |
| $\mathbf{E_{SelfCont}}$ | 0.30 | 0.53 | 0.25 | 0.25 | 0.29 |
| $\mathbf{E_{Solv\ GB}}$ | 21.36 | 31.34 | 23.44 | 23.95 | 13.34 |
| $\mathbf{MM\text{-}GBSA_{Bind}}$ | -58.20 | -55.01 | -57.55 | -63.64 | -51.29 |

[a] Energies for compounds **9**, **13** and **15** are averages of different bound conformations. Calculated energies for the individual conformations are given in Table S7.

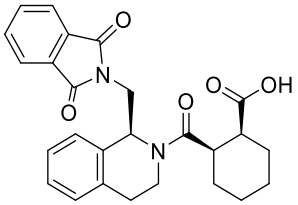
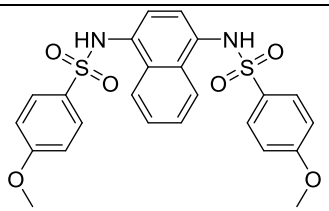
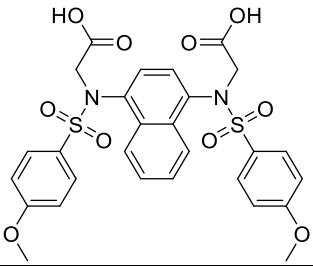
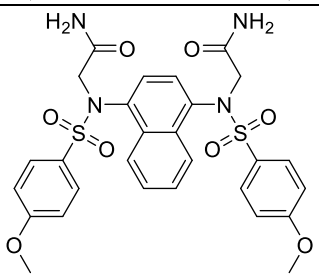
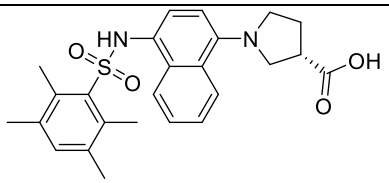
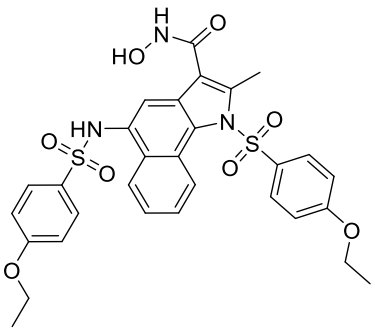
**Table S7.** Binding free energies calculated for the different conformations in which compounds **9**, **13** and **15** may bind to Keap1. Energies were calculated with Prime MM-GBSA and are provided in kcal/mol.

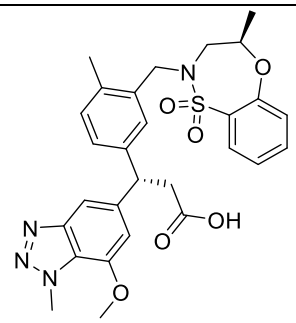
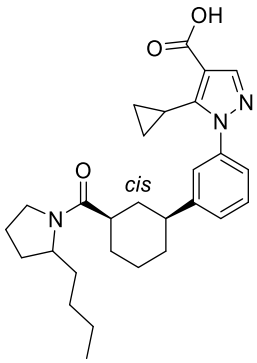
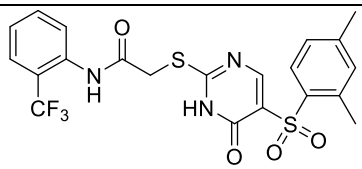
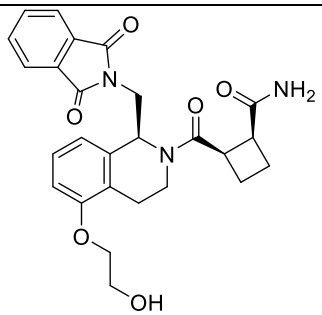
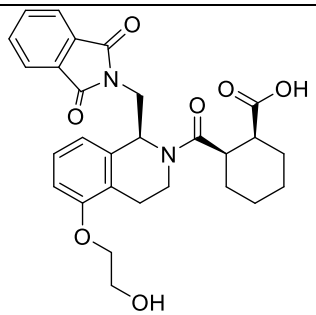
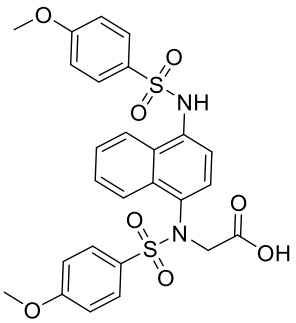
| **Contribution** | **(9) OMe** | | **(13) NMe** | | **(15) NEt2** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *trans* | *cis* | *trans* | *cis* | 1 | 2 | 3 | 4 | 5 | 6 |
| **$E_{Coul}$** | -17.53 | -14.41 | -14.69 | -16.23 | -3.13 | -1.91 | -10.93 | -4.87 | -11.37 | -8.79 |
| **$E_{vdW}$** | -49.61 | -49.54 | -50.81 | -49.94 | -41.44 | -41.66 | -44.04 | -41.46 | -41.05 | -41.27 |
| **$E_{Lipo}$** | -14.66 | -14.90 | -14.36 | -14.96 | -16.56 | -15.04 | -16.52 | -15.30 | -15.77 | -15.55 |
| **$E_{Covalent}$** | 2.99 | 2.99 | 2.70 | 0.68 | -0.41 | 1.07 | 3.63 | 0.94 | 3.26 | 1.41 |
| **$E_{HB}$** | -1.00 | -1.08 | -1.05 | -1.09 | -0.42 | -0.38 | -0.41 | -0.38 | -0.46 | -0.42 |
| **$E_{Packing}$** | -1.47 | -1.50 | -1.31 | -1.42 | -1.58 | -2.04 | -1.58 | -1.96 | -1.55 | -1.59 |
| **$E_{SelfCont}$** | 0.33 | 0.27 | 0.27 | 0.23 | 0.47 | 0.43 | 0.33 | 0.18 | -0.01 | 0.35 |
| **$E_{Solv\ GB}$** | 23.52 | 19.21 | 22.90 | 23.99 | 10.14 | 9.56 | 18.48 | 10.86 | 16.08 | 14.92 |
| **$MMPBSA_{Bind}$** | -57.44 | -58.97 | -56.36 | -58.74 | -52.92 | -49.97 | -51.04 | -51.98 | -50.87 | -50.94 |

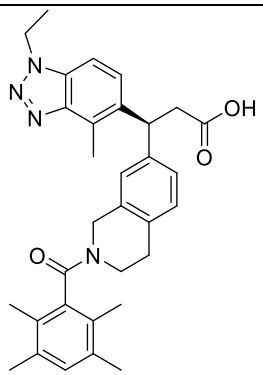
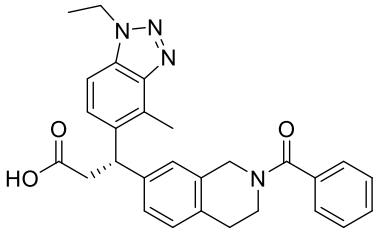
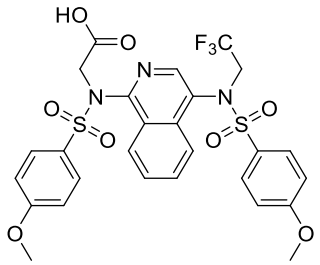
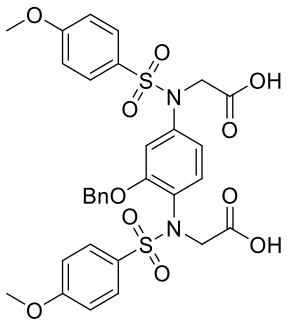
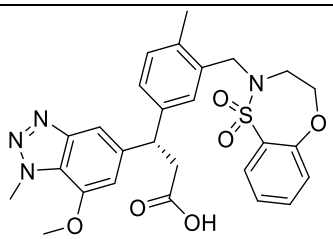
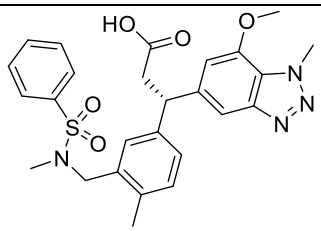
**Table S8.** PubChem IDs of the compounds included in the ¨PubChem¨ set of inhibitors of the Keap1-Nrf2 PPI.

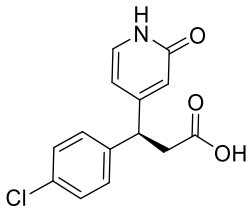
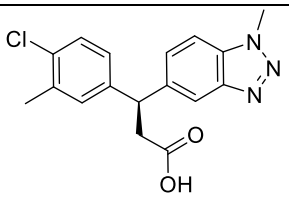
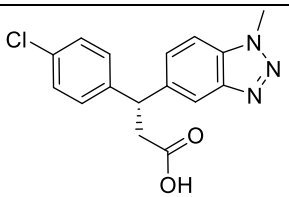
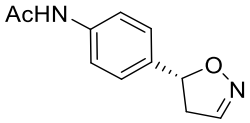
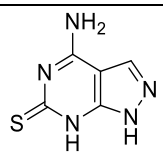
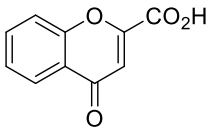
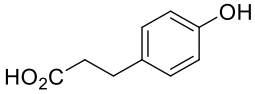
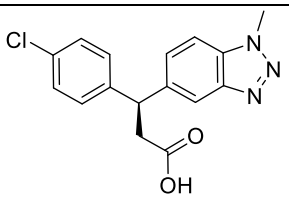
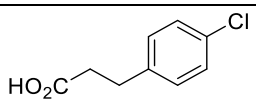
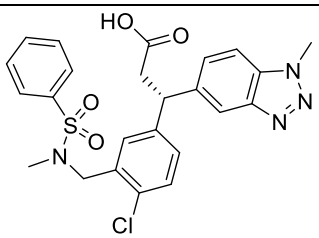
| Pubchem ID | Pubchem ID | Pubchem ID | Pubchem ID | Pubchem ID | Pubchem ID | Pubchem ID | Pubchem ID |
|---|---|---|---|---|---|---|---|
| 51089829 | 26728148 | 7974131 | 56422303 | 49669481 | 24779150 | 24825722 | 137276122 |
| 89854938 | 24829566 | 57260573 | 14742838 | 24825143 | 92764796 | 49671788 | 137276148 |
| 24811348 | 7972928 | 26732045 | 49642153 | 4248390 | 17509878 | 85199487 | 24798269 |
| 17507037 | 14733521 | 7966617 | 22404448 | 49645395 | 22403398 | 89853346 | 137276123 |
| 862895 | 26662707 | 49680279 | 17508591 | 857242 | 26731456 | 57267239 | 49665518 |
| 24811166 | 14728463 | 57264490 | 49821344 | 24830045 | 49667306 | 26665871 | 137276118 |
| 26730097 | 47199028 | 57266210 | 14731549 | 26731697 | 57256794 | 57264421 | 17514406 |
| 24832765 | 49648331 | 26659839 | 14731051 | 57256179 | 17515518 | 47200946 | 137276135 |
| 26728852 | 56463520 | 22414565 | 24830685 | 26662442 | 24817037 | 137276110 | 49646564 |
| 26728495 | 57258364 | 81066397 | 57266997 | 4263656 | 24779880 | 137276101 | 24802470 |
| 49665230 | 17507017 | 49679995 | 7965518 | 7973906 | 865524 | 7977252 | 14739396 |
| 49640558 | 56423153 | 22412996 | 49717937 | 89852938 | 26726782 | 49665326 | 22404075 |
| 85201922 | 24837832 | 26663318 | 856072 | 47196791 | 47195322 | 49680120 | 137276146 |
| 22414251 | 93577218 | 26726754 | 7966286 | 57258271 | 57258297 | 24829662 | 26671199 |
| 49671082 | 47200693 | 26728484 | 26727713 | 24823674 | 26666864 | 24810226 | 137276129 |
| 47199711 | 17514836 | 81066003 | 49680398 | 87349223 | 22402458 | 85199342 | 137276151 |
| 26726720 | 7968999 | 26728720 | 26664273 | 26724325 | 49640994 | 863985 | 85267952 |
| 17503854 | 57264483 | 47197290 | 17506362 | 85271528 | 859771 | 26665310 | 17513336 |
| 24836548 | 85273725 | 17511470 | 24793026 | 14739295 | 49680993 | 26661686 | 7975141 |
| 26665209 | 57257957 | 17506046 | 85269638 | 24794776 | 57256138 | 50086765 | 14735592 |
| 56463290 | 26728500 | 87349331 | 24780759 | 14737603 | 24793940 | 57267658 | 26663704 |
| 7967765 | 49826718 | 49679961 | 26662852 | 861951 | 85147438 | 137276096 | 56423121 |
| 26662278 | 49668975 | 47197717 | 26730133 | 24794011 | 7977444 | 137276137 | 125082067 |
| 47195933 | 24822559 | 49674600 | 47200971 | 49671604 | 57264186 | 49665687 | 137276130 |
| 14719363 | 49681567 | 26662177 | 57257934 | 24819706 | 3716071 | 137276160 | 137276140 |
| 17512528 | 4255867 | 137276167 | 26659873 | 57257807 | 57264743 | 22414615 | 137276175 |
| 24830410 | 24790693 | 14723775 | 26662506 | 49731689 | 24784098 | 17401506 | 137276181 |
| 26664968 | 47199273 | 49646670 | 7971429 | 26727807 | 57258869 | 17508621 | 864554 |
| 49680443 | 85146382 | 26664933 | 8139864 | 49668827 | 7973765 | 137276112 | 137276128 |
| 17401681 | 24840439 | 22405155 | 24807802 | 57257737 | 26664392 | 92764778 | 17515884 |
| 56321175 | 49666263 | 47196922 | 85269124 | 137276174 | 49736663 | 57265921 | 24815048 |
| 26731548 | 85147439 | 57260004 | 26725751 | 57262988 | 85148992 | 85270327 | 56463649 |
| 24819784 | 17401788 | 137276165 | 47201832 | 17514464 | 137276107 | 85271088 | |
| 49726002 | 47194273 | 49646460 | 47198276 | 14739794 | 57264930 | 137276178 | |
| 57267546 | 845397 | 14729511 | 49665355 | 14742382 | 24780599 | 14736033 | |
| 24785705 | 57262489 | 49645652 | 26667935 | 47199147 | 49675230 | 89853209 | |
| 47196106 | 89852629 | 7976131 | 49668562 | 11533021 | 26729871 | 137276095 | |
| 92764363 | 87336285 | 856223 | 49821301 | 49674604 | 57257054 | 47197329 | |
| 26729204 | 17386900 | 14718786 | 57265681 | 137276157 | 22401801 | 137276116 | |
| 22411822 | 26727606 | 57260741 | 14738815 | 57261299 | 57257404 | 49678901 | |
| 49641105 | 4258633 | 49666621 | 24823253 | 24806768 | 24840600 | 49824847 | |
| 92764290 | 85149207 | 57264495 | 49645097 | 7965179 | 85304455 | 24785386 | |
| 49645020 | 87338532 | 24802960 | 57255915 | 137276121 | 87345622 | 57255884 | |
| 47196075 | 7975468 | 49647358 | 859482 | 26668969 | 49646559 | 137276139 | |
| 49668408 | 24782090 | 24822771 | 87345778 | 17432625 | 57267231 | 137276127 | |
| 85199467 | 14731284 | 24840333 | 47202582 | 49819194 | 24791506 | 17513335 | |
| 24823505 | 22404715 | 849500 | 49735399 | 49665488 | 85271114 | 47203201 | |
| 85198922 | 24787398 | 56423686 | 57268294 | 24823842 | 85201603 | 87335628 | |
| 2678521 | 24820754 | 57265022 | 57260179 | 49643845 | 137276153 | 22416313 | |

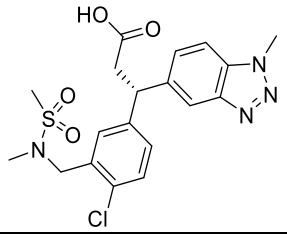
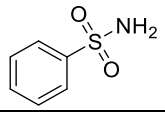
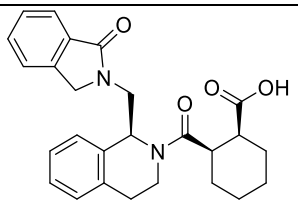
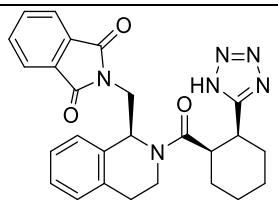
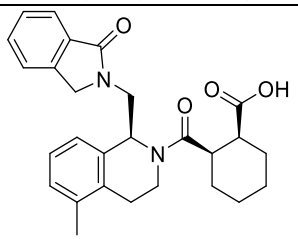
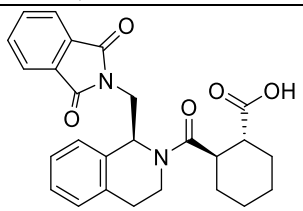
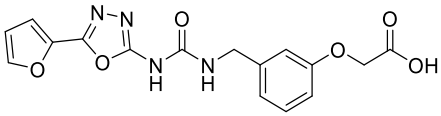
**Table S9.** Structures of the compounds in the ¨validated¨ set of inhibitors of the Keap1-Nrf2 PPI, with their enumeration in the article by Tran et al.[2] and the PDB ID of those that have been co-crystallized with Keap1.

| Identifier | Structure | Compound number in Tran et al.[2] | PDB ID |
|---|---|---|---|
| **V1** |  | **3** | 4L7B |
| **V2** |  | **4** | 4IQK |
| **V3** |  | **5** | --- |
| **V4** |  | **6** | 4XMB |
| **V5** |  | **7** | 5CGJ |
| **V6** |  | **8** | --- |

| | | | |
|---|---|---|---|
| **V7** |  | **9** | 5FNU |
| **V8** |  | **10** | --- |
| **V9** |  | **11** | 4IN4 |
| **V10** |  | --- | 6SP4 |
| **V11** |  | --- | 6SP1 |
| **V12** |  | --- | 6V6Z |

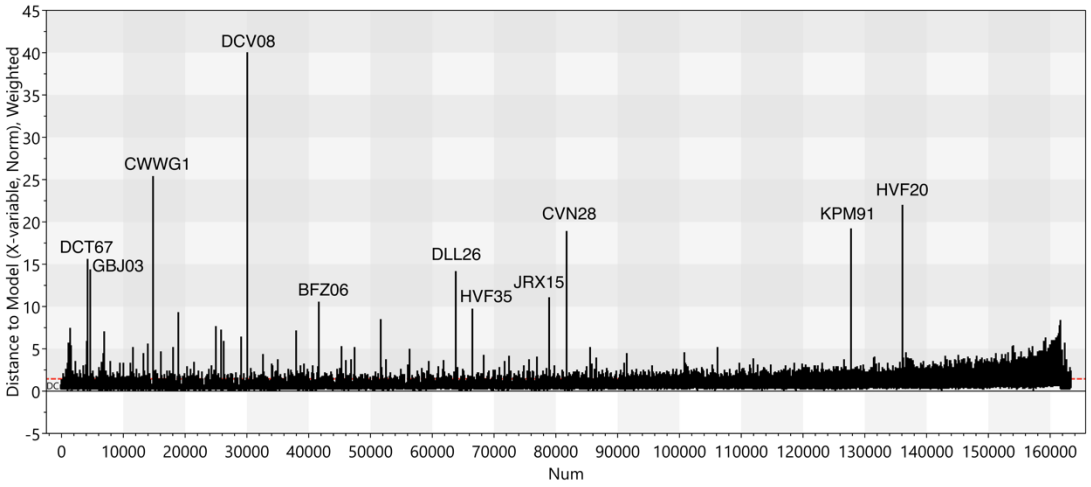
| | | | |
|---|---|---|---|
| **V13** |  | --- | 6TYM |
| **V14** |  | --- | 6TYP |
| **V15** |  | --- | 6UF0 |
| **V16** |  | --- | 6HWS |
| **V17** |  | --- | 6QMK |
| **V18** |  | --- | 6QMJ |

| | | | |
|---|---|---|---|
| **V19** |  | --- | 6QMC |
| **V20** |  | --- | 6QME |
| **V21** |  | --- | 6QMD |
| **V22** |  | --- | 6FFM |
| **V23** |  | --- | 5WIY |
| **V24** |  | --- | 5WHO |
| **V25** |  | --- | 5WHL |
| **V26** |  | --- | 5FNR |
| **V27** |  | --- | 5FNQ |
| **V28** |  | --- | 5FNT |

| | | | |
|---|---|---|---|
| **V29** |  | --- | 5FNS |
| **V30** |  | --- | 5FZN |
| **V31** |  | --- | 4N1B |
| **V32** |  | --- | 4L7C |
| **V33** |  | --- | 4L7D |
| **V34** |  | --- | 4IFN |
| **V35** |  | --- | 3VNG |

# Supporting Figures

**a)**




**b)**

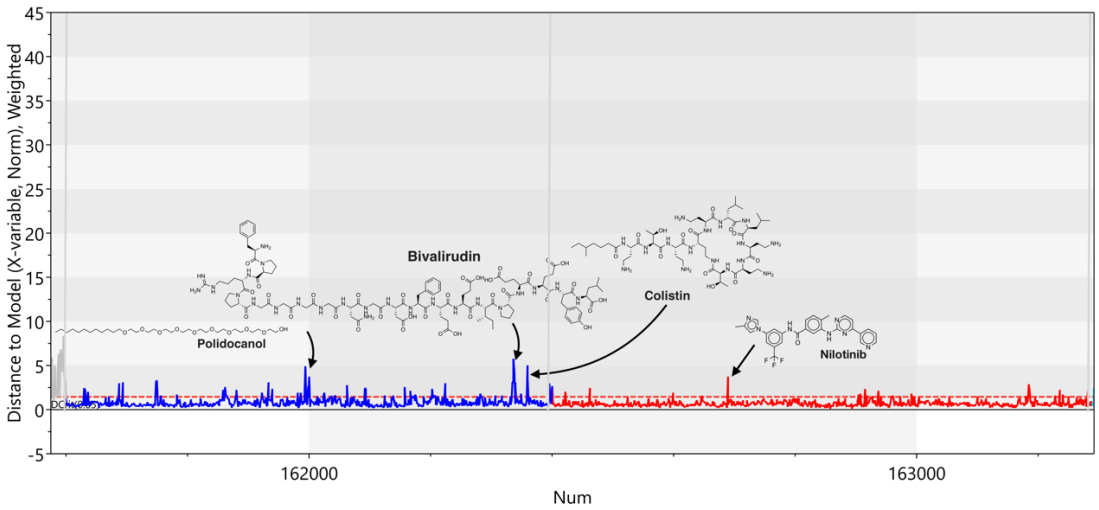



**Figure S1.** Distance to model (DModX) plots from the first two principle components in the PCA analysis. DModX shows the distance of a given observation to the model plane. Distance to model for **a)** all compounds in the Dictionary of Natural Products and **b)** all drugs in DrugBank (non-oral are in blue, and oral in red). DModX variables (normalized and weighted, Y-axis) is plotted against compound number (X-axis).

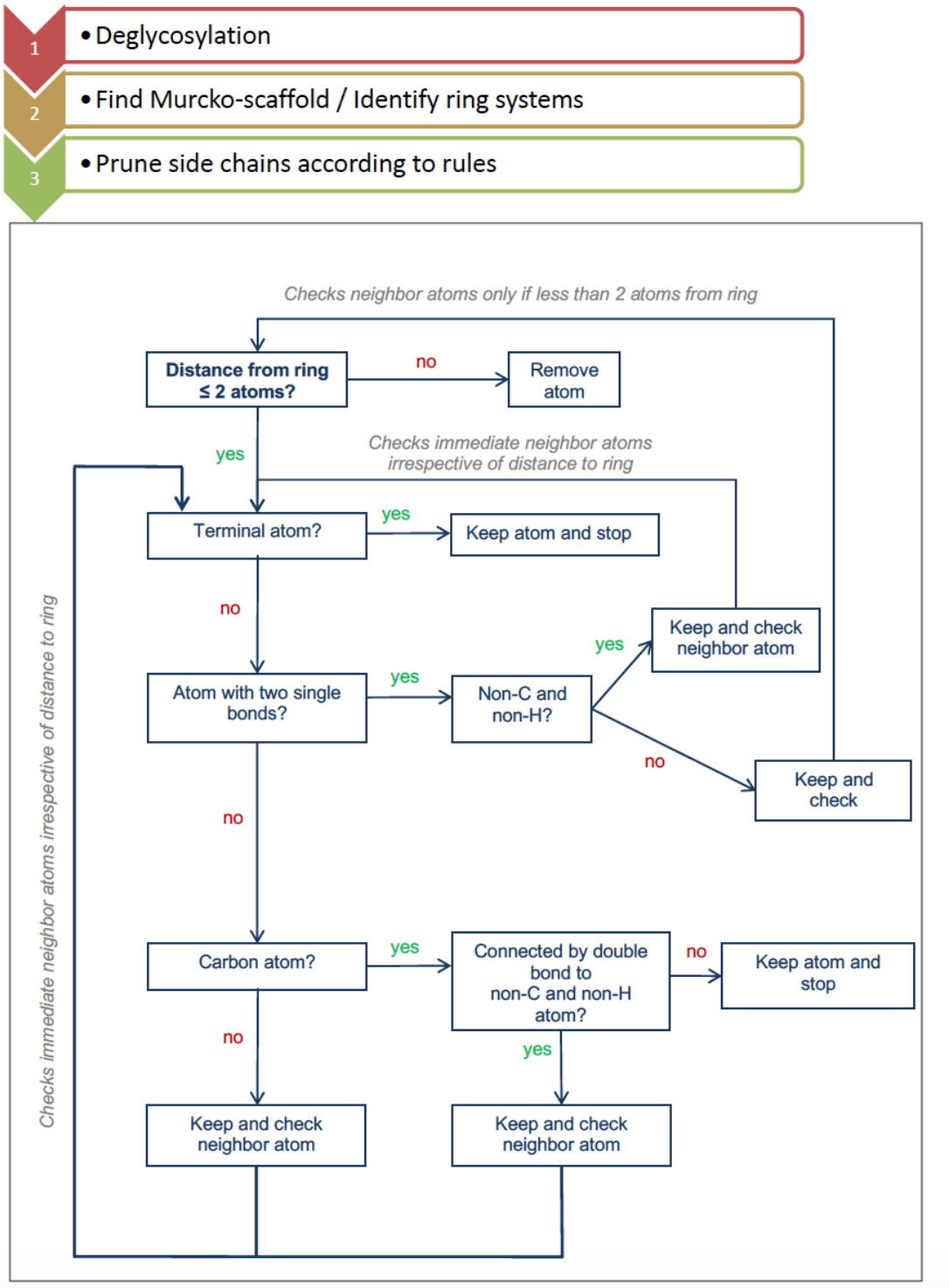



**Figure S2.** Workflow of the algorithm used for pruning of side chains from the macrocyclic natural products to give macrocyclic cores

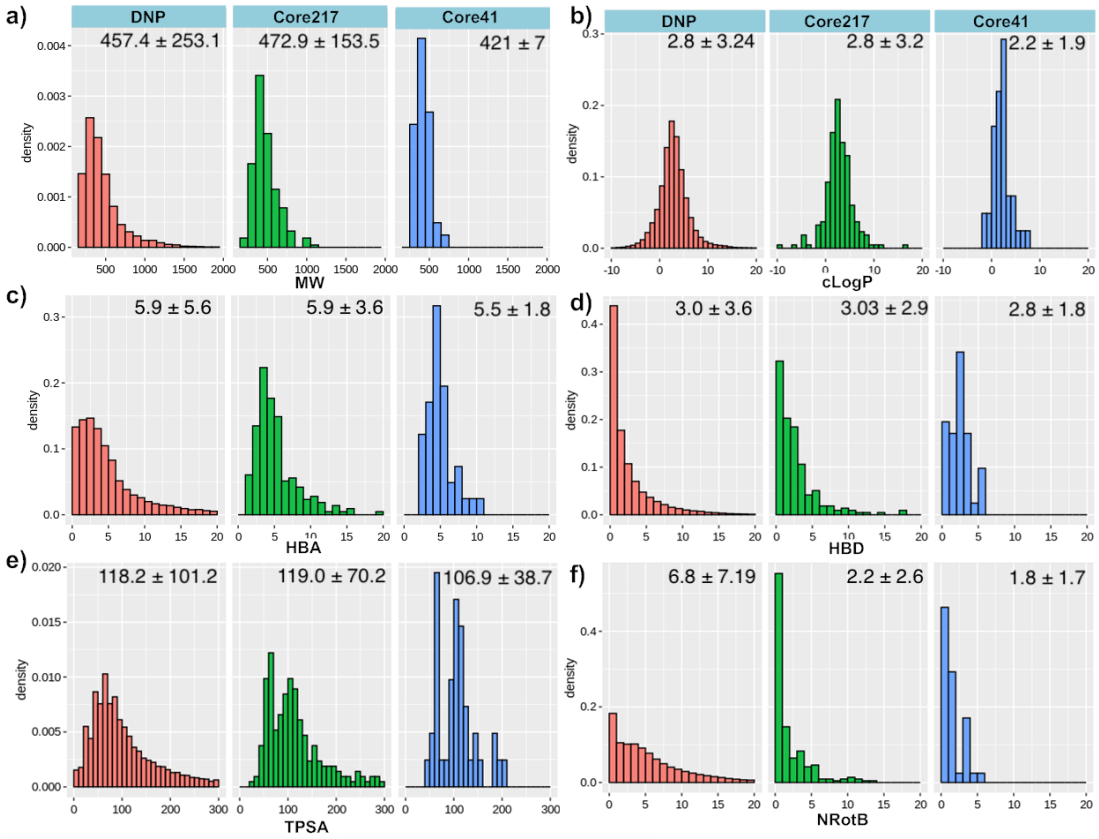



**Figure S3.** Molecular descriptor distribution for all of the natural products in the Dictionary of Natural Products (DNP) and the sets of 217 and 41 macrocyclic cores extracted from the DNP. **a)** MW, molecular weight; **b)** cLogP, calculated lipophilicity; **c)** HBA, hydrogen bond acceptor; **d)** HBD, hydrogen bond donor; **e)** TPSA, topological polar surface area; **f)** NRotB, number of rotatable bonds. The mean value and the standard deviation for the descriptor is given in each panel. Descriptors were calculated using the MOE descriptor module.[3]

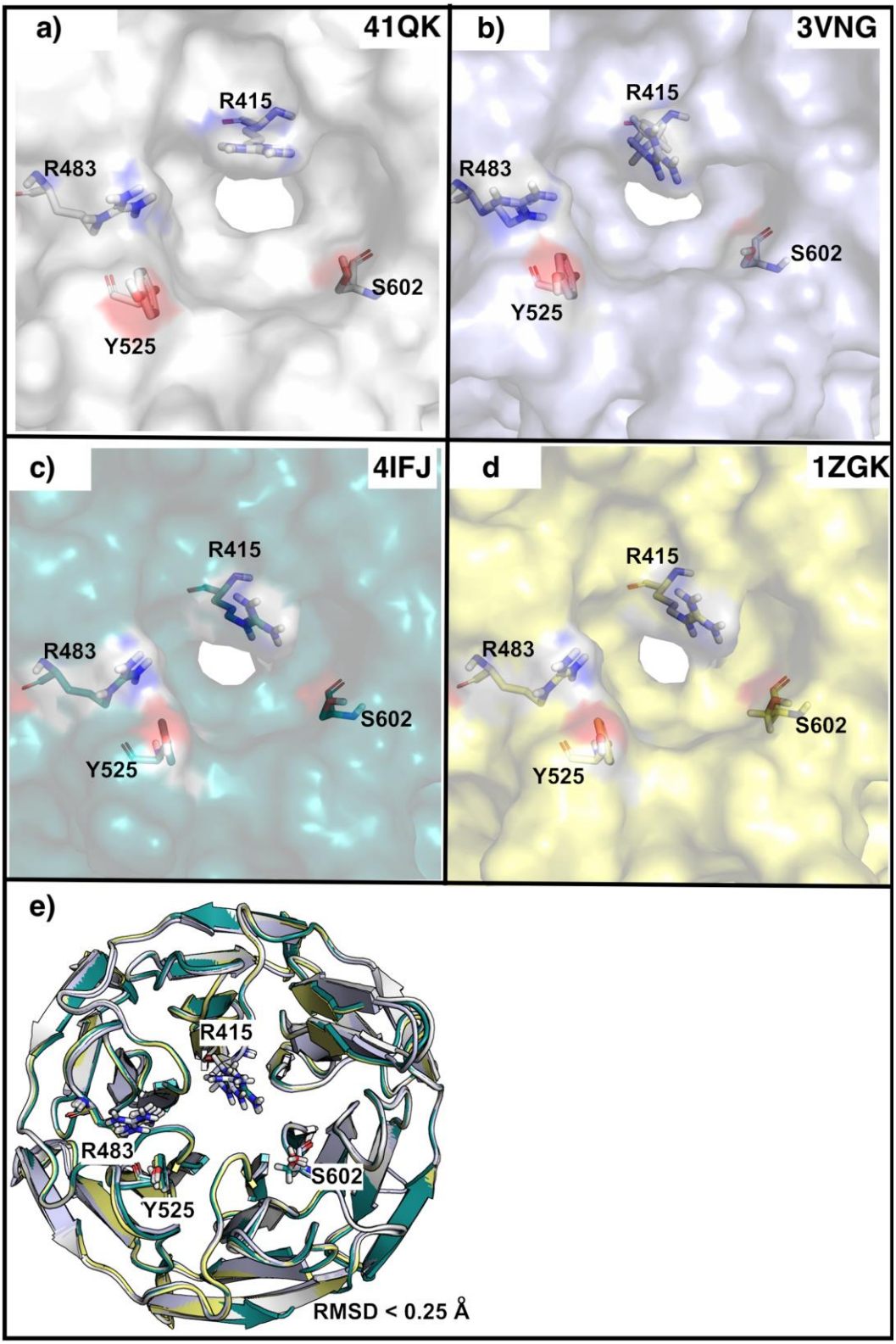
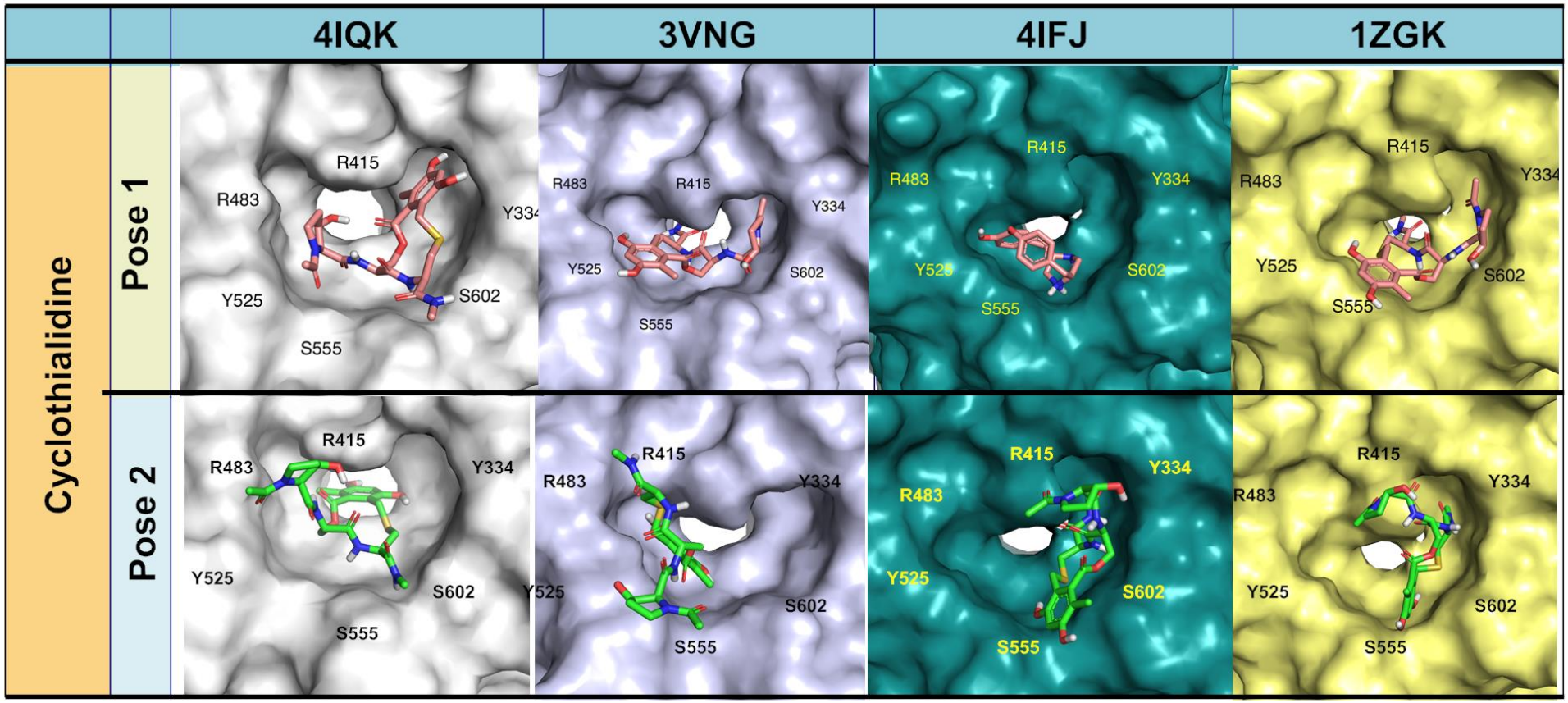
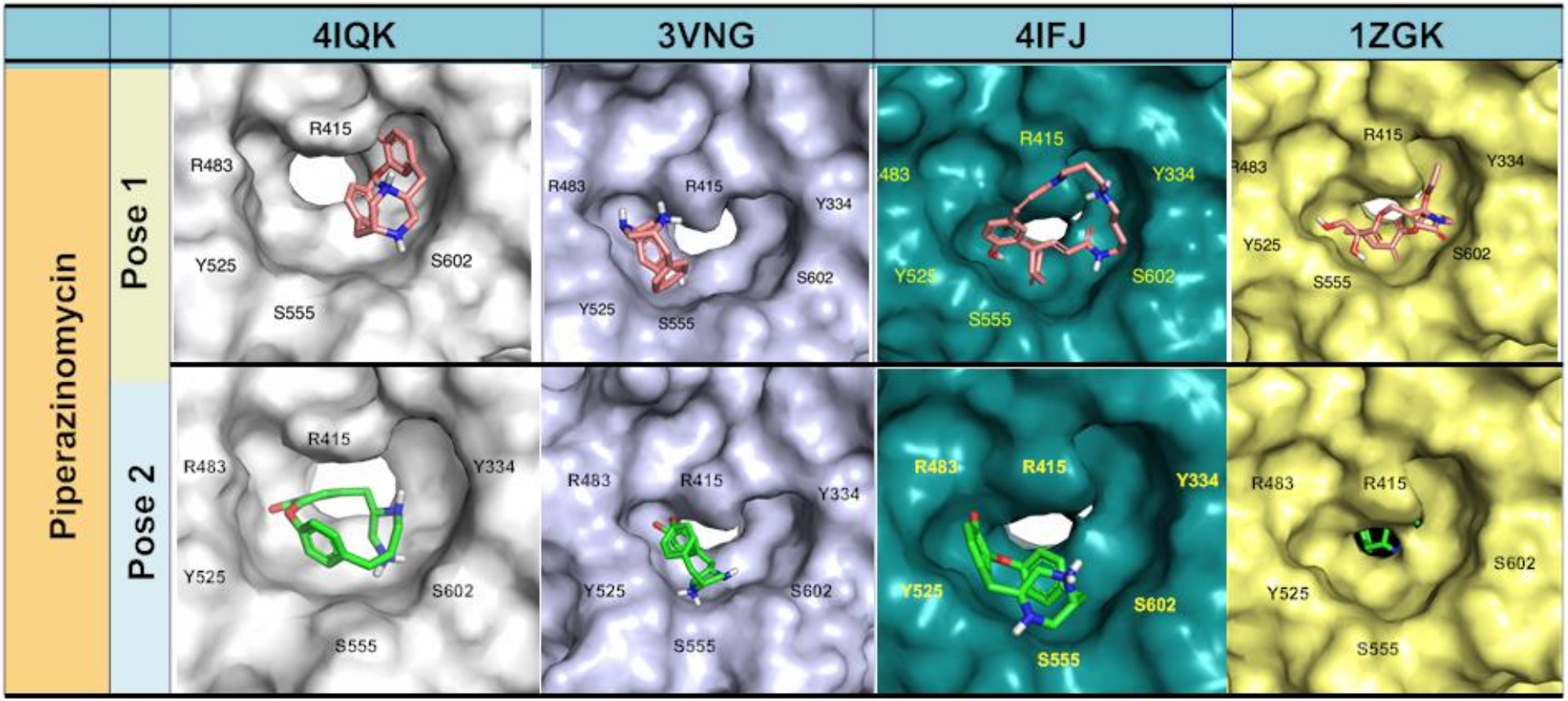
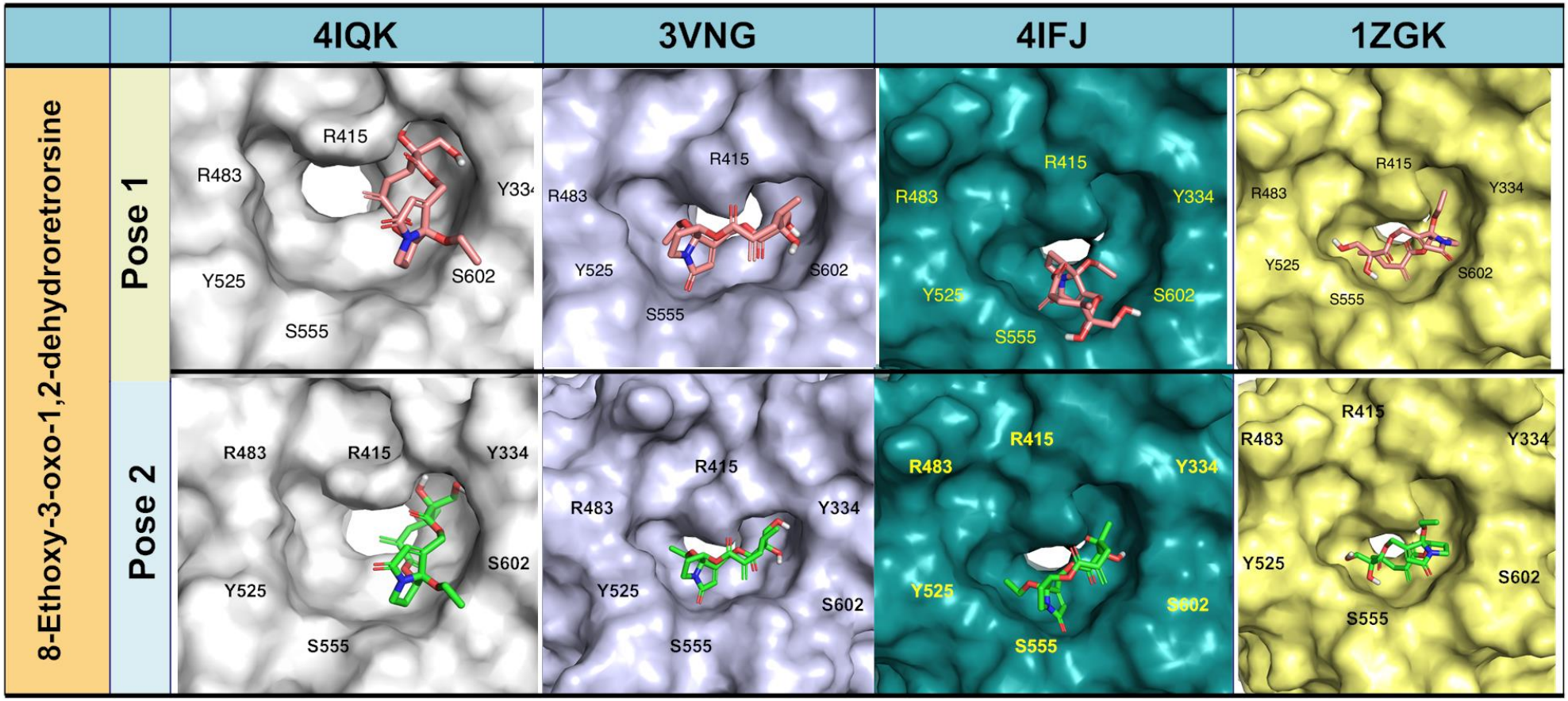


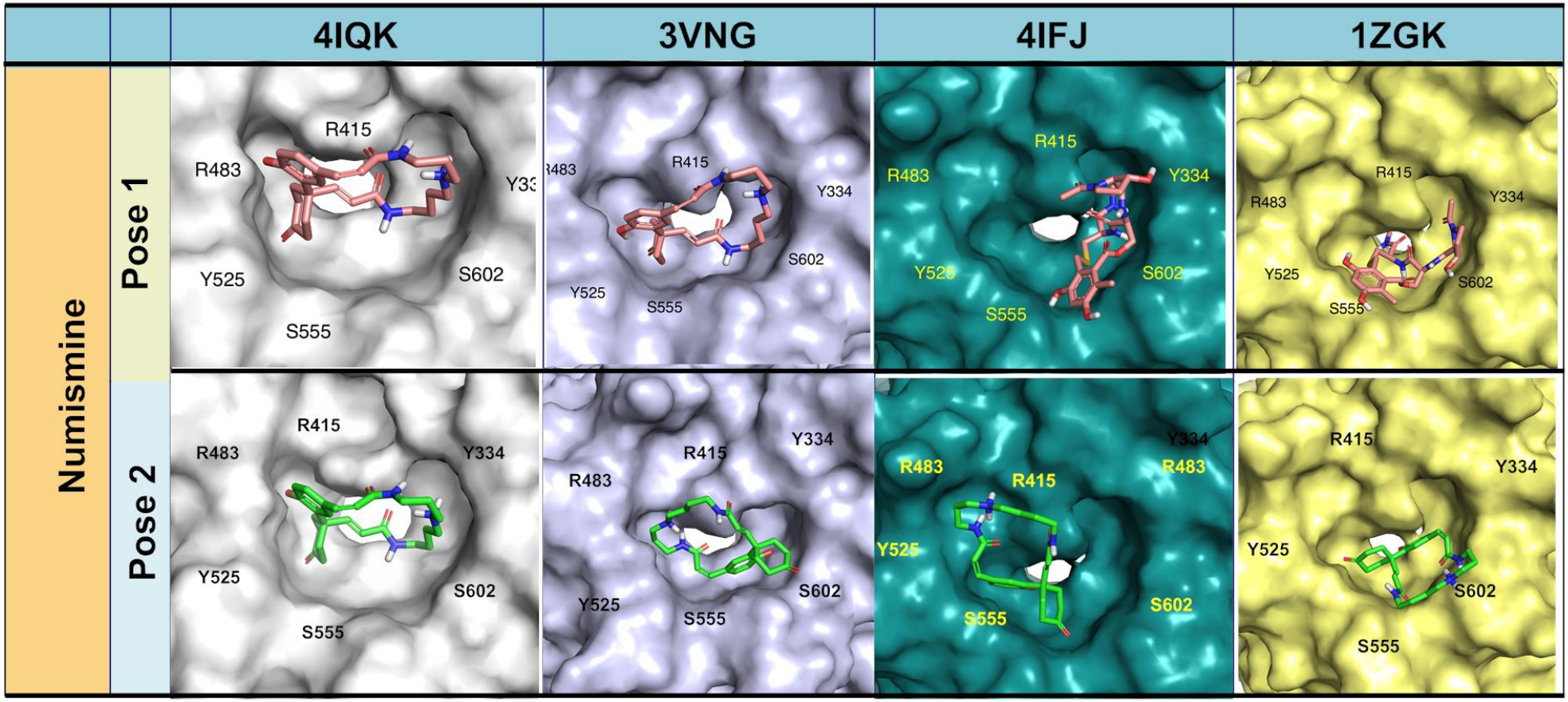

**Figure S4.** Conformational mobility of the binding site on Keap1 for Nrf2. **a)-d)** Arg415 adopts different positions in the complexes with small molecule inhibitors (4IQK[4] and 3VNG[5]) as compared to in the apo form (4IFJ and 1ZGK). Some other amino acids, e.g. Arg483, also undergo minor conformational changes between the different crystal structures. **e)** Superimposition of the atoms in the backbones of Keap1 in the four crystal structures. Four key amino acids in the binding site of Keap1 are indicated in the five panels.

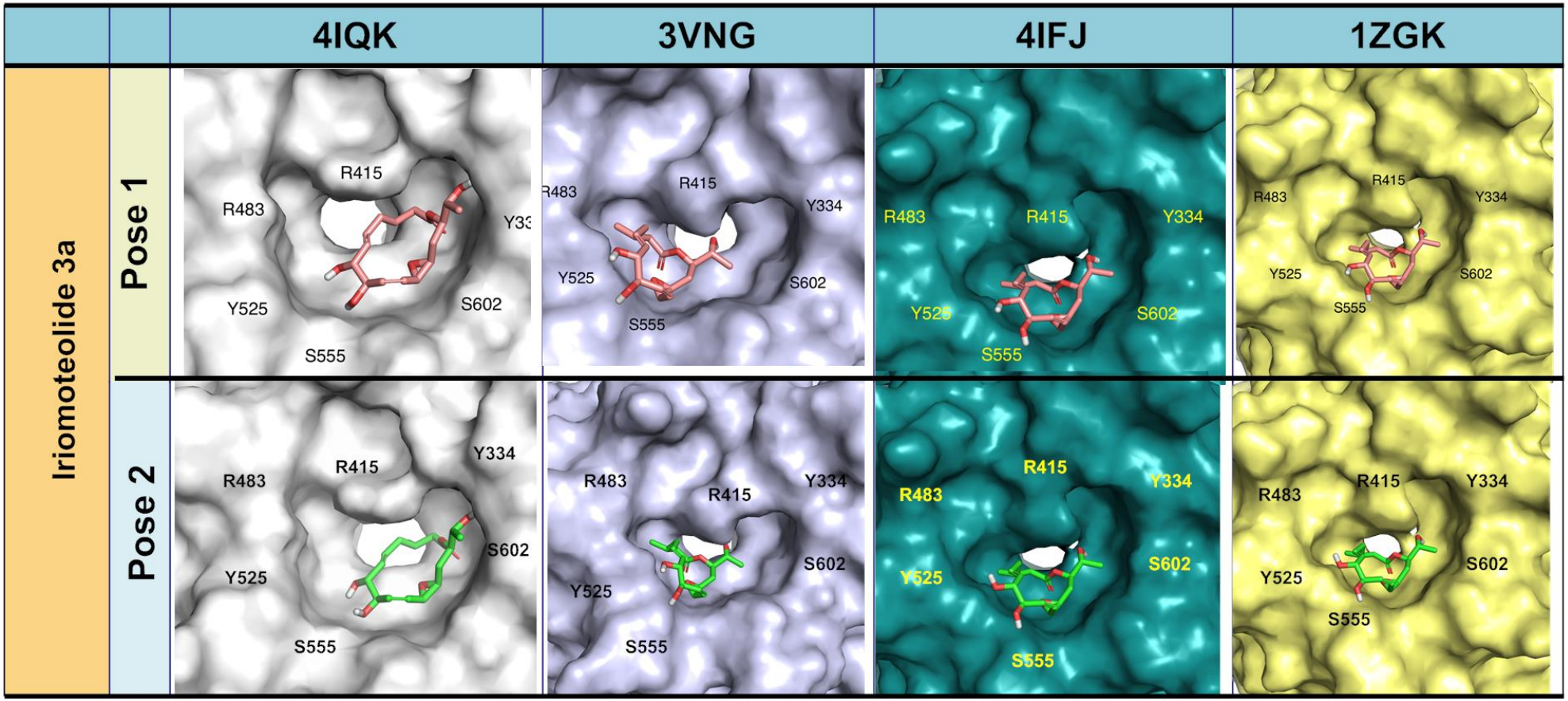
**Figure S5.** Binding poses of the top five cores from the docking in the binding site of Keap1 in the four selected crystal structures (PDB ID: 4IQK, 3VNG, 4IFJ and 1ZGK). Cores. **1**: Cyclothialidine; **2**: Piperazinomycin; **3**: 8-Ethoxy-3-oxo-1,2-dehydroretrorsine; **4**: Numismine; **5**: Iriomoteolide 3a.




4IQK
3VNG
4IFJ
1ZGK
Piperazinomycin
Pose 1
Pose 2
R415
R483
Y525
S602
S555
Y334


4IQK
3VNG
4IFJ
1ZGK
8-Ethoxy-3-oxo-1,2-dehydroretrorsine
Pose 1
Pose 2
R415
R483
Y334
Y525
S602
S555


4IQK
3VNG
4IFJ
1ZGK
Numismine
Pose 1
Pose 2
R415
R483
Y334
Y525
S602
S555


4IQK
3VNG
4IFJ
1ZGK
Iriomoteolide 3a
Pose 1
Pose 2
R415
R483
Y525
S602
S555
Y334

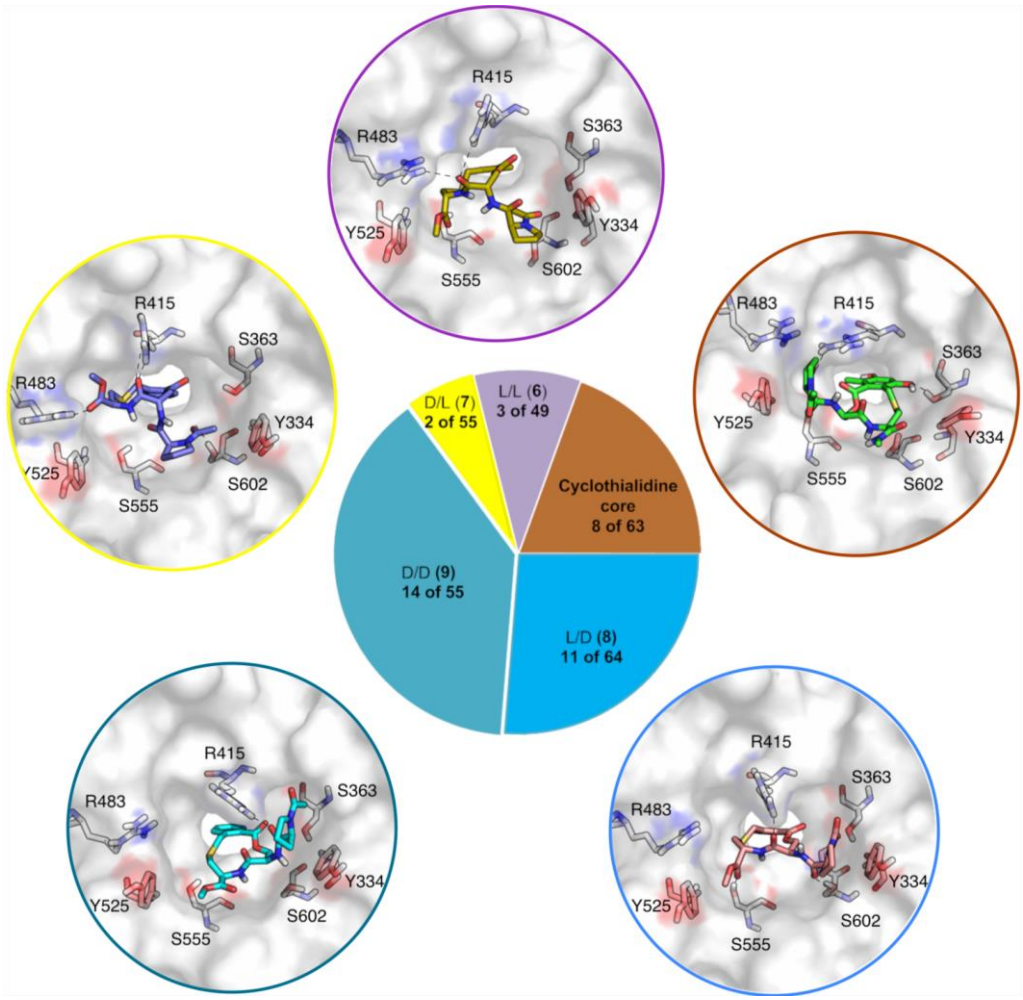



**Figure S6.** "Probabilities" of macrocycles **6**-**9** and cyclothialidine core **1** to bind with their phenyl group reaching deep into the Kelch channel of Keap1, as estimated with induced-fit docking (IFD). The probability was calculated as the fraction of poses which displayed this "preferable" binding mode, as compared to the total number of poses obtained from IFD (central pie chart). Selected poses that illustrates how each of **6**-**9** and **1** adopt this "preferable" binding mode when docked into the binding site of Keap1 are shown in the circles located around the central pie chart. IFD was performed using a crystal structure of Keap1 co-crystallized with a small molecule ligand (PDB ID 4IQK[4]).

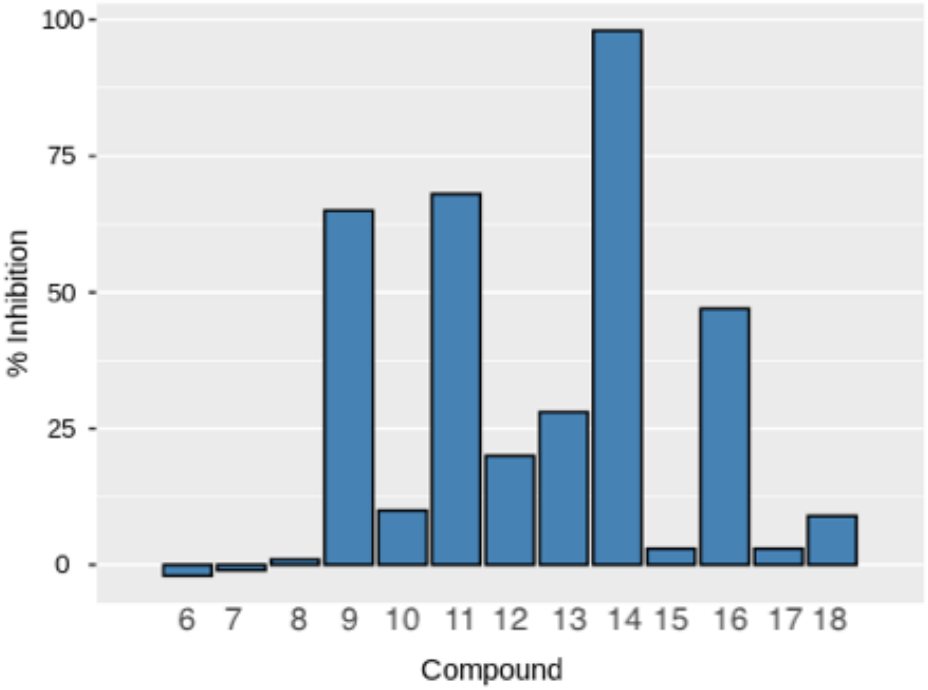



**Figure S7.** Potency (% Inhibition) of compounds **6**-**18** as inhibitors of the Keap1-Nrf2 PPI determined at 500 μM in the surface plasmon resonance-based inhibition in solution assay (ISA)

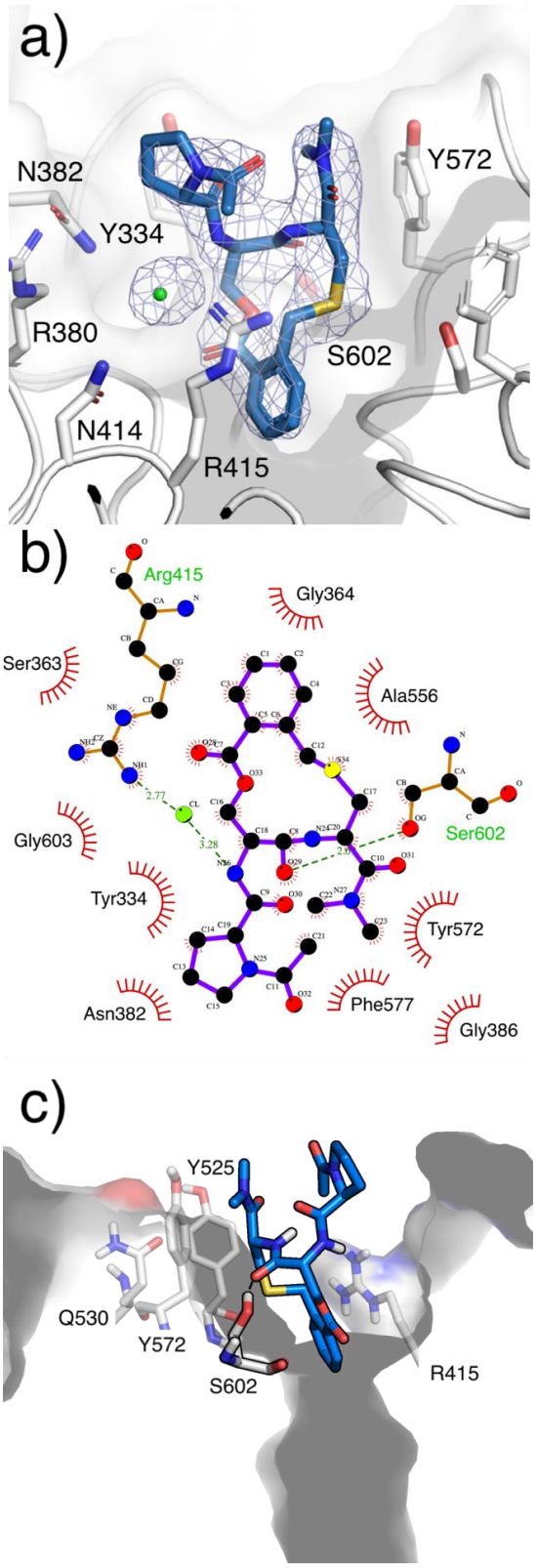



**Figure S8**. **a)** 2Fo-Fc electron density of compound **14** in the crystalline complex with Keap1 contoured at 1.3 sigma, with the chloride ion in green. **b)** Overview of the ligand-protein interactions formed in the complex of **14** and Keap1, drawing generated with LigPlot v.2.1[6] c) View of the structure of the complex of **14** and Keap1 in which the hydrogen bond between S602 and the carbonyl group of Ser in **14** is highlighted.

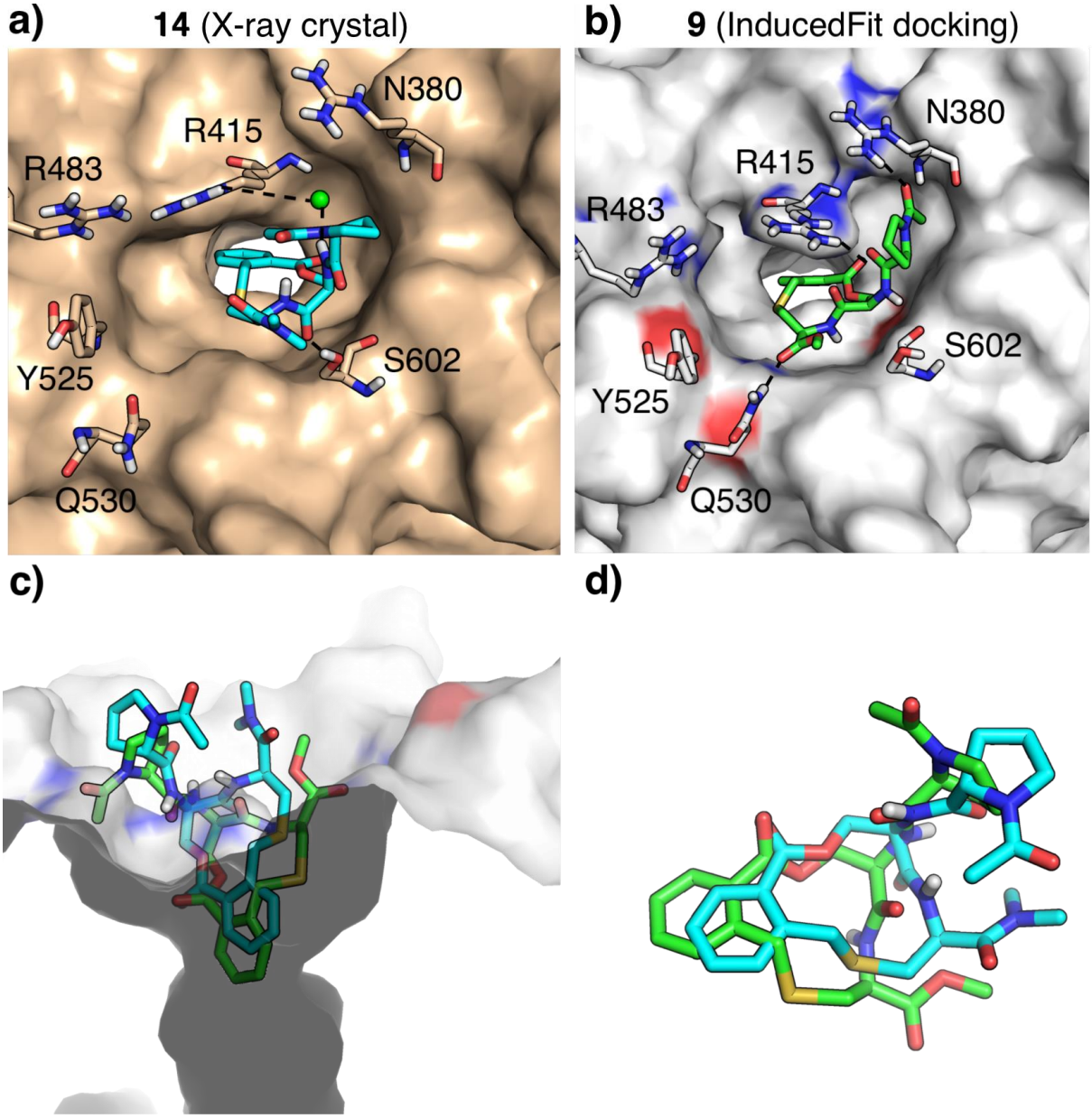
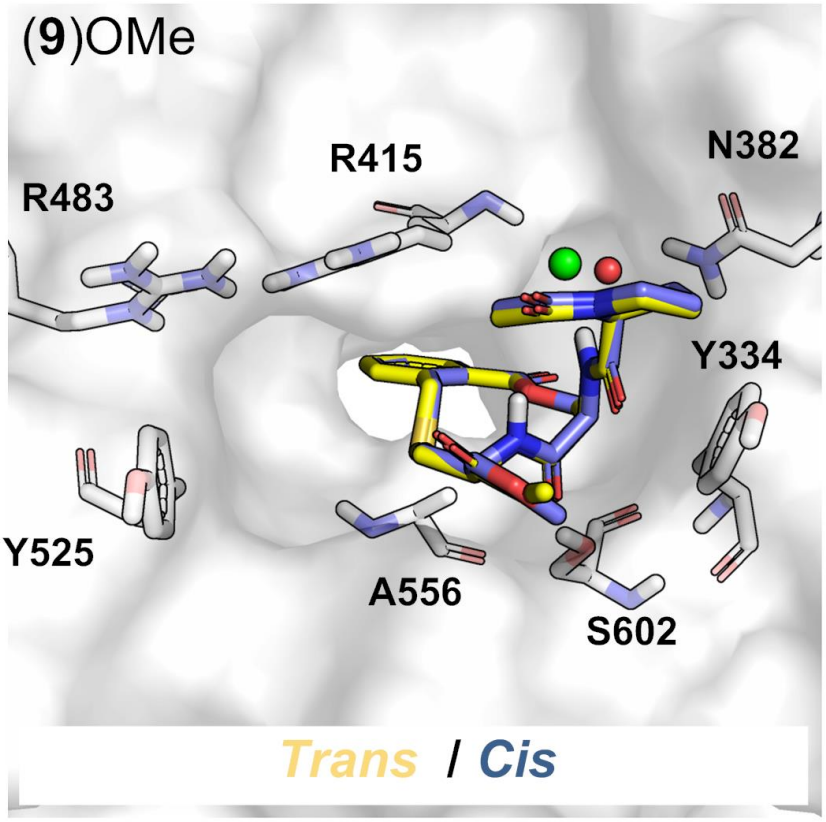
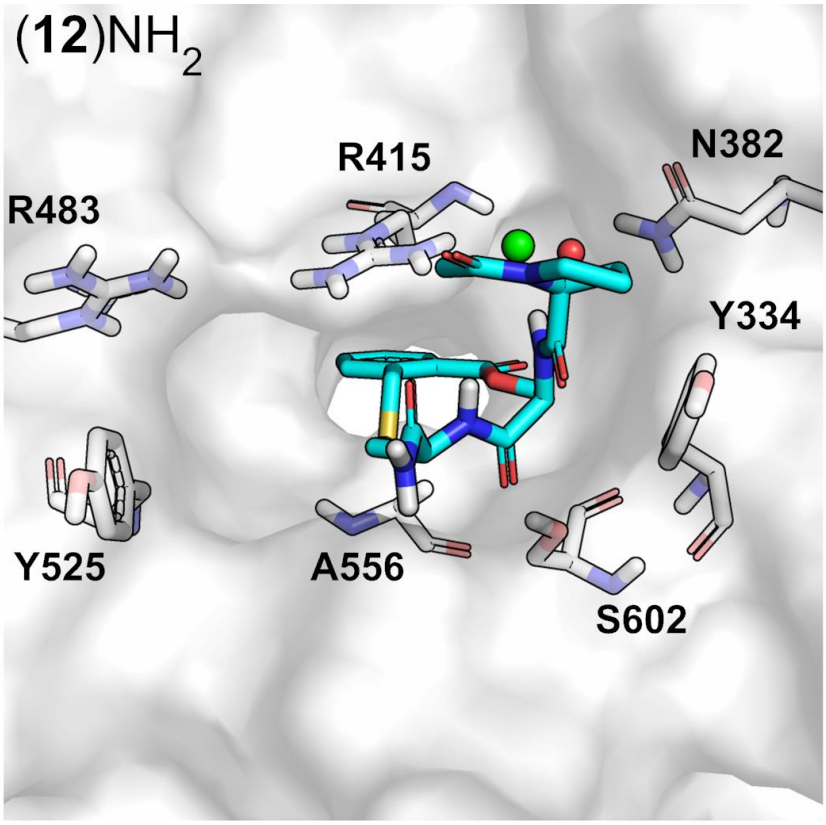
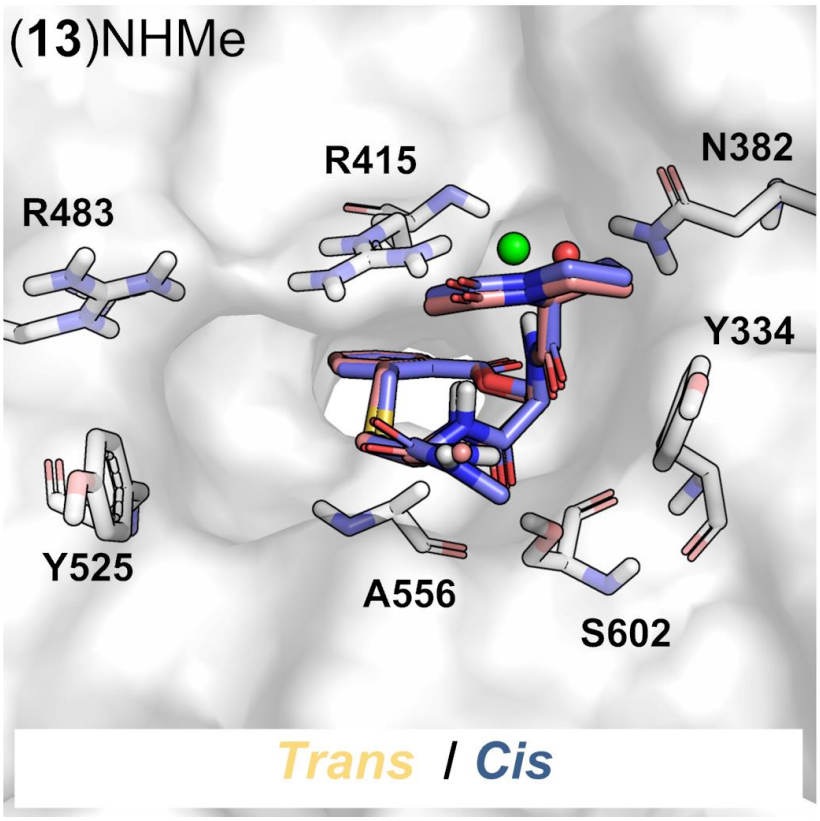
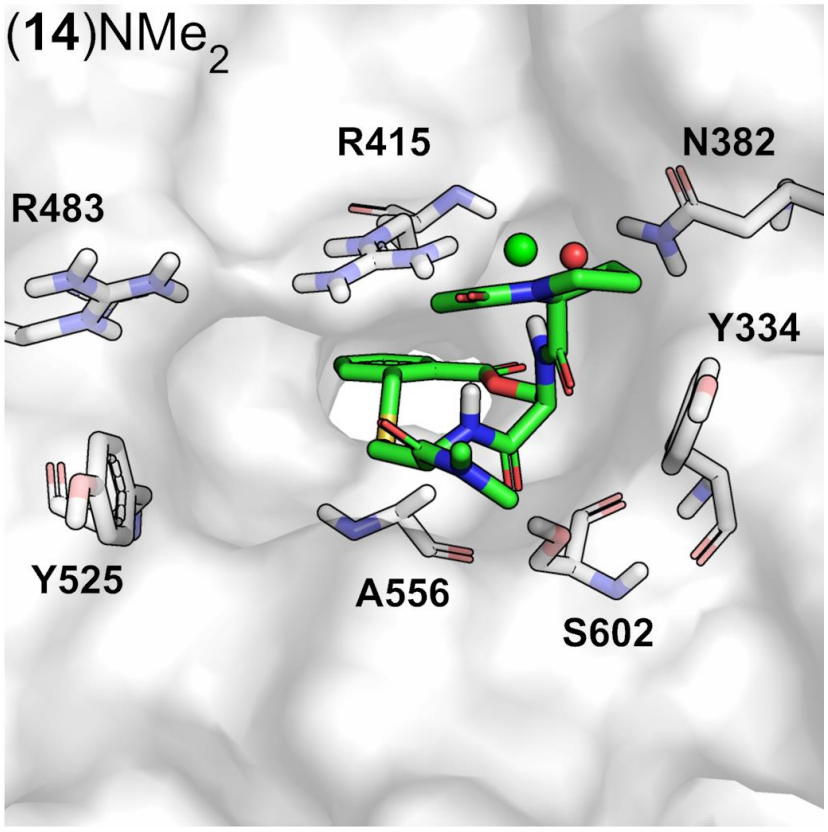



**Figure S9.** Comparison of one of the docked poses of macrocycle **9** in Keap1 and the crystal structure of the complex between Keap1 and **14**. **a)** View of macrocycle **9** docked into the binding site of crystalline Keap1 (PDB ID: 4IQK[4]). **b)** View of **14** in the crystalline complex with Keap1. **c)** and **d)** Superimposition of the complexes and the bound ligands in panels a) and b), respectively.


(9)OMe
R415
N382
R483
Y334
Y525
A556
S602
Trans / Cis


(12)NH2
R415
N382
R483
Y334
Y525
A556
S602


(13)NHMe
R415
N382
R483
Y334
Y525
A556
S602
Trans / Cis


(14)NMe2
R415
N382
R483
Y334
Y525
A556
S602

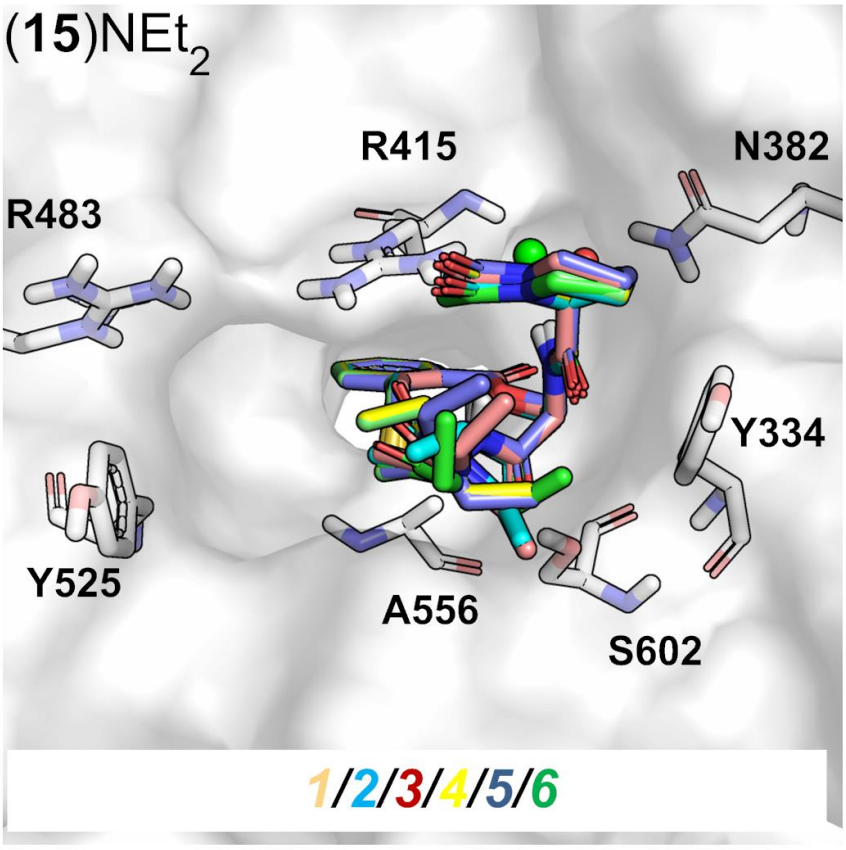



**Figure S10.** Comparison of the Keap1 bound poses of compounds **9, 12-15** used in the binding affinity calculations. The ligand binding pocket of Keap1 is shown in surface representation.

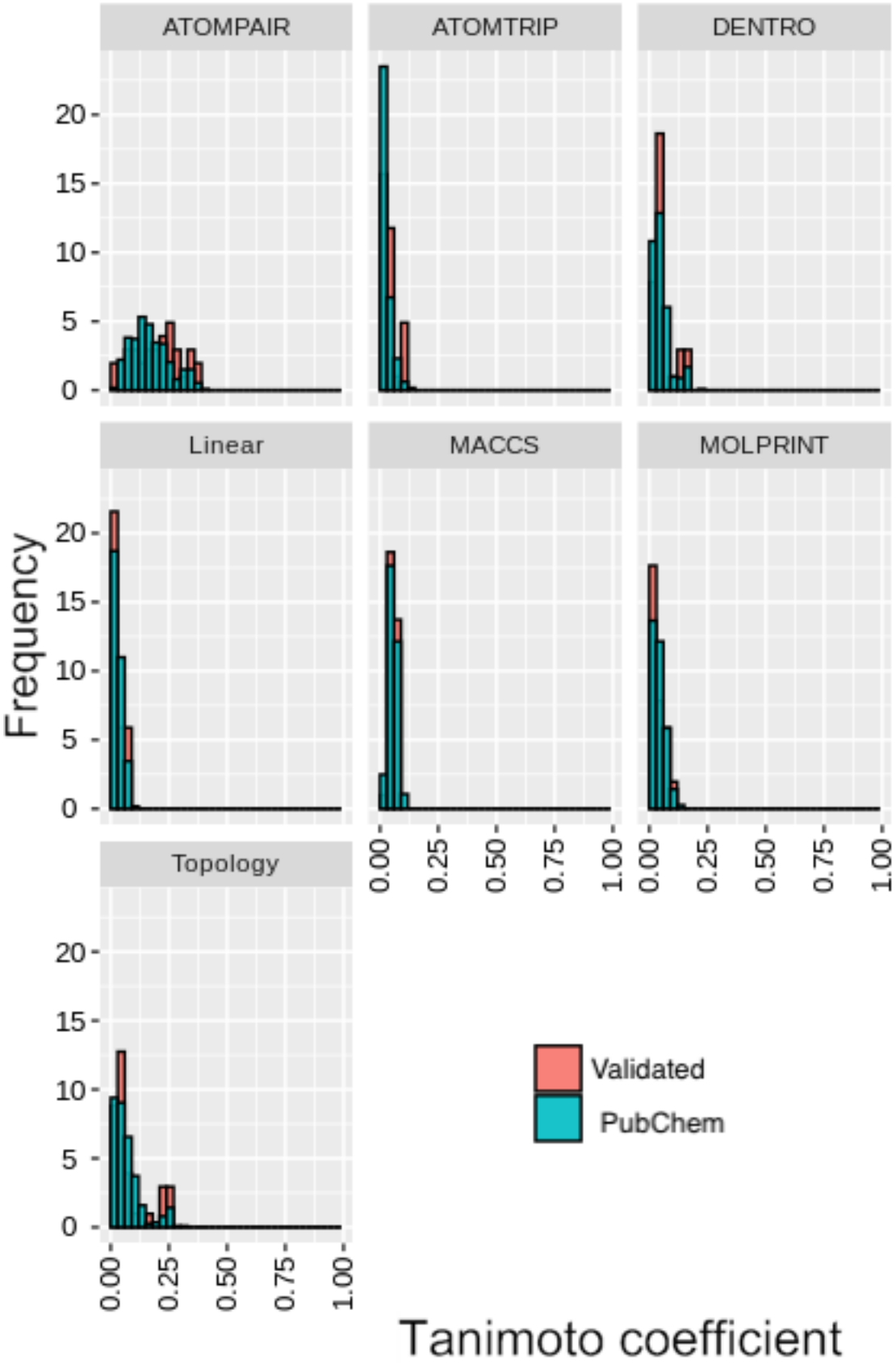



**Figure S11.** Tanimoto coefficients comparing the structure of macrocycle **14** to the Validated and PubChem sets of inhibitors of the Keap1-Nrf2 PPI using seven different fingerprints, i.e. *ATOMPAIR*

(Pairs of atoms, differentiated by type, and the distance separating them), *ATOMTRIP* (Triplets of atoms differentiated by type, and the three distances separating them. Topological torsions: Linear paths of length three), *DENTRO* (Linear and branched fragments), *Linear* (Linear fragments + ring closures), *MACCS* (Structural fingerprints), *MOLPRINT* (A radial-like fingerprint that encodes atom environments using lists of atom types located at different topological distances) and *Topology* (Topological torsions: Linear paths of length three).

# Computational Procedures

## Computational Procedure 1: Validation of the protocol for docking into Keap1

Prior to the docking of 41 macrocyclic cores into the crystal structures of Keap1, the docking protocol was validated by comparing the conformation (the pose) of the bound ligands as obtained from docking with the one determined experimentally for two structures [PDB ID: 3VNG (2.1 Å)[5] and 4IQK (1.97 Å)[4]] co-crystalized with small molecule ligands [N'N-naphthalene-1,4-diylbis(4-methoxybenzenesulfonamide and 2-(3-((3-(5-(furan-2-yl)-1,3,4-oxadiazol-2-yl)ureido)methyl)phenoxy)acetic acid), respectively]. Both crystal structures were imported and refined with Protein Preparation Wizard[7] using the Schrödinger Suite.[8] Structure refinement included adding hydrogen atoms, assigning bond orders, building disulfide bonds and removal of water molecules beyond 5 Å from the ligand atoms. The PROPKA[9] tool from Protein Preparation Wizard was used to predict the protonation states of the ionizable residues at pH 7.0. Subsequently, the positions of the hydrogen atoms in the Keap1 structure were energy minimized using the OPLS3 force field.[10] Ligands were prepared using the LigPrep[11] module from the Schrödinger Suite.

The receptor grid generation module of Glide[12] was used to define the active site for the docking experiments. The active site of Keap-1 was defined using the bound ligand in the two chosen crystal structures as the centroid of the grid box with a radius of 15 Å from centroid. Subsequently, ligands were docked into Keap-1 structures using the standard precision mode, where the ligand sampling was set to be flexible and the top 20 best poses per ligand were extracted. Other docking parameters were set to default.

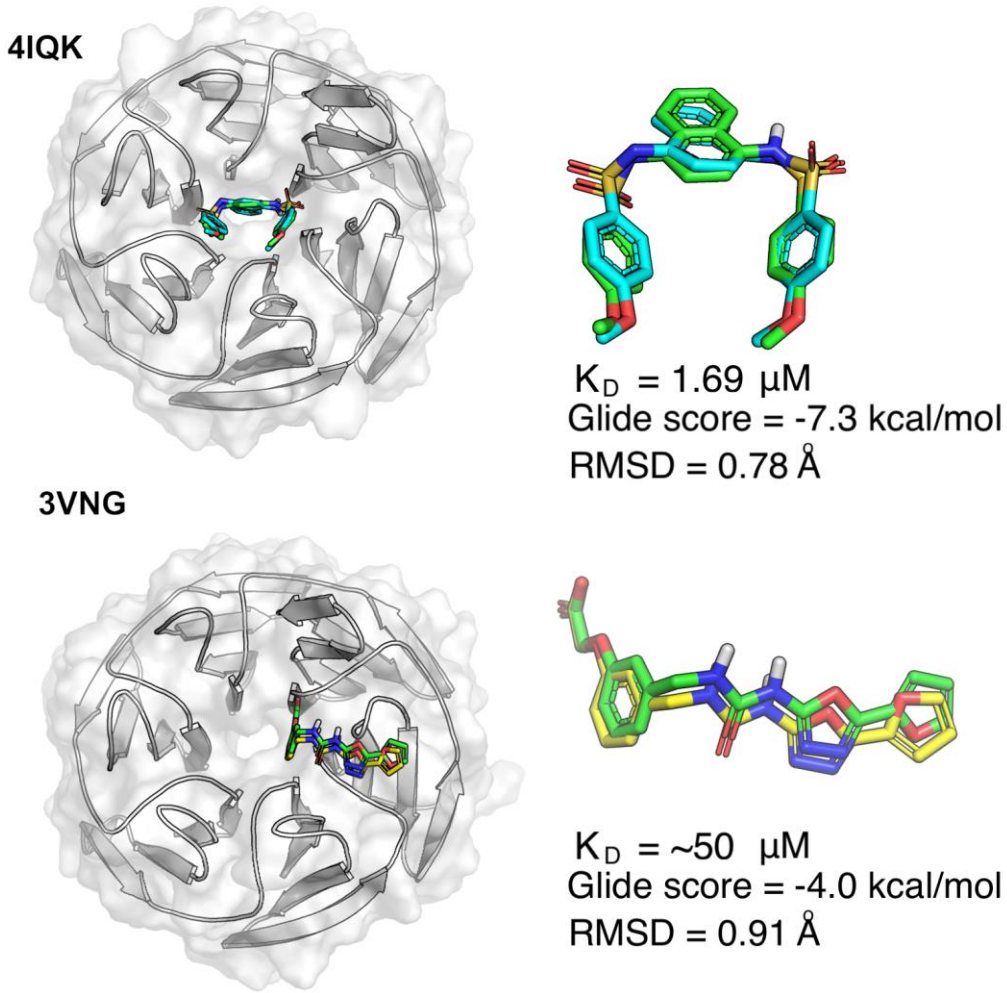
Docking poses were compared to the bound ligand conformation by root-mean-square deviation (RMSD). The docking protocol used in this validation study was able to reproduce the bound conformation of the ligand in each of the two crystal structures with RMSD < 1 Å within the top 20 ranked poses (Computational Procedure 3, Table 1 and Figure 1).

**Computational Procedure 1, Table 1.** Summary of ligand RMSD and Glide score values for the top 20 poses of ligands re-docked into crystal structures 4IQK and 3VNG[a]

| Rank | 4IQK | | | 3VNG | | |
|---|---|---|---|---|---|---|
| | Pose No | GlideScore (kcal/mol) | RMSD (Å) | Pose No | GlideScore (kcal/mol) | RMSD (Å) |
| **1** | 1 | -7.73 | 1.55 | 1 | -5.38 | 2.38 |
| **2** | **2** | **-7.31** | **0.78** | 2 | -5.27 | 3.16 |
| **3** | 3 | -6.75 | 1.92 | 3 | -5.03 | 7.94 |
| **4** | 4 | -5.45 | 1.30 | 4 | -4.95 | 2.35 |
| **5** | 5 | -5.32 | 1.43 | 5 | -4.68 | 2.25 |
| **6** | 6 | -5.31 | 1.49 | 6 | -4.67 | 2.29 |
| **7** | 7 | -5.12 | 1.17 | 7 | -4.56 | 6.58 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **8** | 8 | -4.97 | 1.51 | 8 | -4.52 | 9.29 |
| **9** | 9 | -4.95 | 1.69 | 9 | -4.51 | 2.60 |
| **10** | 10 | -4.88 | 2.33 | 10 | -4.51 | 8.85 |
| **11** | 11 | -4.65 | 1.50 | 11 | -4.44 | 2.48 |
| **12** | 12 | -4.56 | 1.56 | 12 | -4.44 | 7.03 |
| **13** | 13 | -4.49 | 1.95 | 13 | -4.42 | 3.28 |
| **14** | 14 | -4.47 | 1.28 | 14 | -4.27 | 2.70 |
| **15** | 15 | -4.26 | 0.99 | 15 | -4.26 | 3.46 |
| **16** | 16 | -4.21 | 1.97 | 16 | -4.19 | 7.90 |
| **17** | 17 | -4.12 | 2.70 | 17 | -4.10 | 3.48 |
| **18** | 18 | -3.89 | 2.15 | **18** | **-4.04** | **0.91** |
| **19** | 19 | -3.76 | 1.59 | 19 | -4.01 | 1.28 |
| **20** | No pose | - | - | 20 | -3.80 | 3.56 |

[a]Docking poses in which the bound ligand is most similar to that in the crystal structure are marked in yellow.

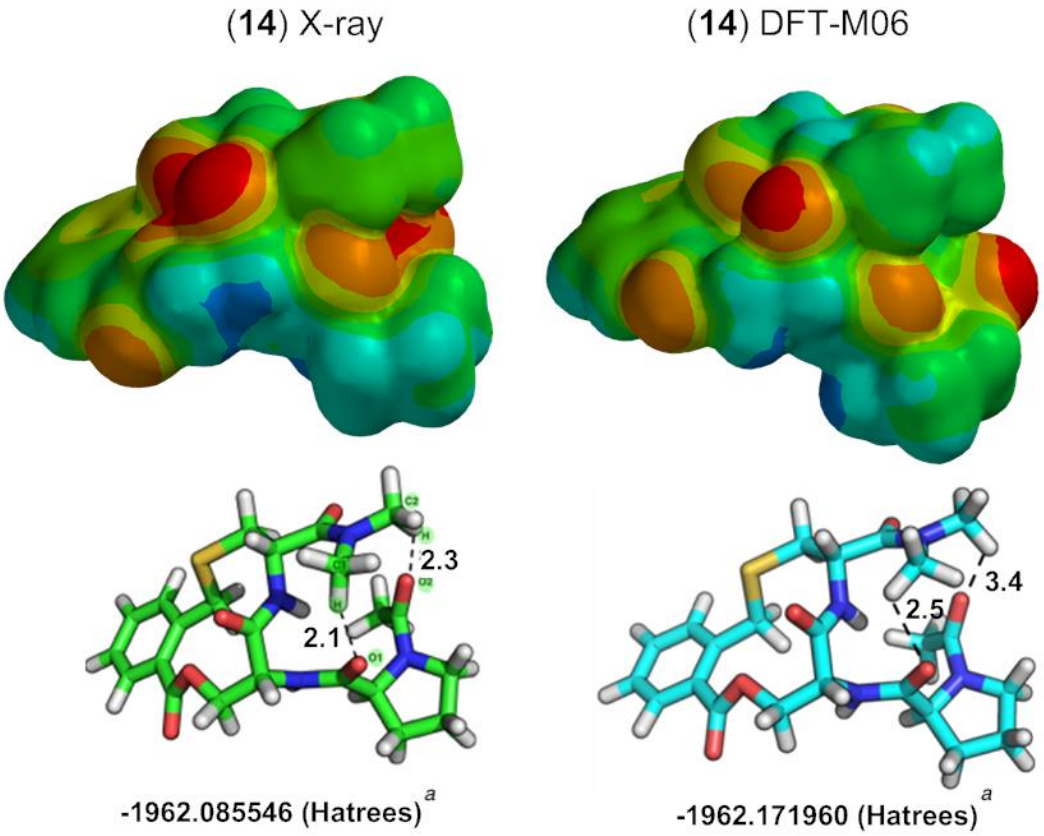



**Computational Procedure 1, Figure 1.** Validation of ligand binding mode. The pose of the docked ligands is shown in cyan (PDB: 4IQK) and yellow (3VNG), while the bound conformation determined

by X-ray crystallography of each ligand is shown in green. The $IC_{50}$ values for the co-crystalized ligands were obtained from the literature (4IQK[4] and 3VNG[13]).

## Computational Procedure 2: Quantum mechanical calculations of intramolecular non-classical hydrogen bonds

Quantum mechanical calculations were carried out using the Spartan 14 software[14] to study the potential non-classical intramolecular hydrogen bonds of compound **14**. The conformation of compound **14** in the crystalline complex with Keap1 was geometry optimized using the DFT combination functional RM06 with the 6-31G** basis set in vacuum, while other parameters were set to default as implemented in Spartan. The RM06 functional has been shown to accurately identify the global minimum energy conformer and correctly rank low energy conformers as compared to other DFT functionals (e.g., DFT B3LYP).[15] Electrostatic potential map (EPS) were constructed for crystalline **14** and its geometry optimized conformation at an isovalue of 0.002 electrons/au$^3$ for the electron density. High to low electron density regions are shown in red to blue colors, respectively (Computational Procedure 7, Figure 1).

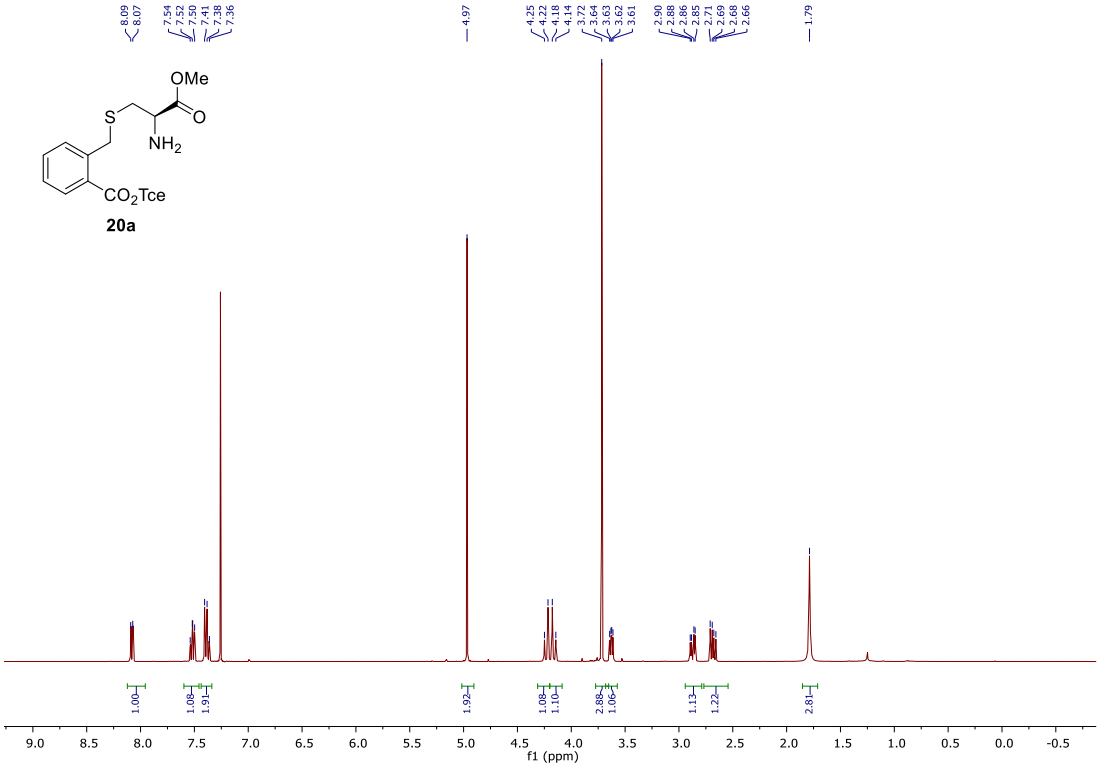



**Computational Procedure 2, Figure 1.** Conformations of **14** in the crystalline complex with Keap1 (left) and after geometry optimization (right) using DFT M06 6-13G**. Electrostatic potential map (top), and structures (bottom) with distances between CH and carbonyl oxygen atoms (Å) in the indicated potential non-classical intramolecular hydrogen bonds.

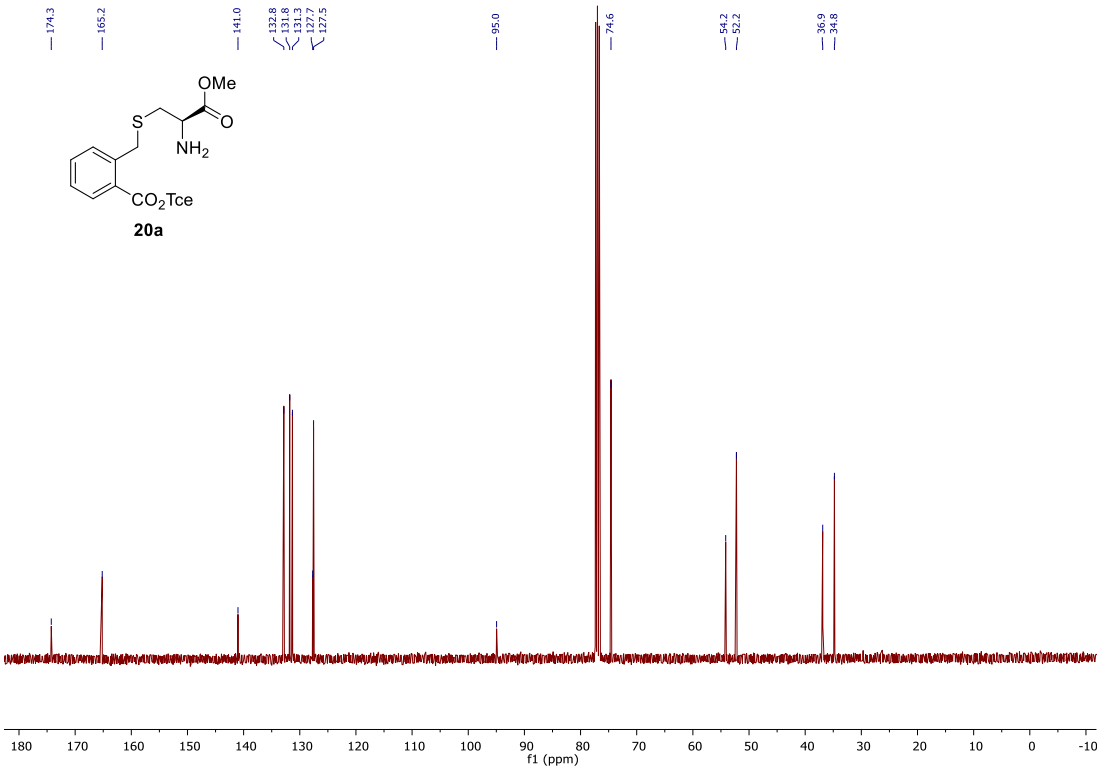
[a] Total energy of each of the two conformations in vacuum.

# NMR Spectra

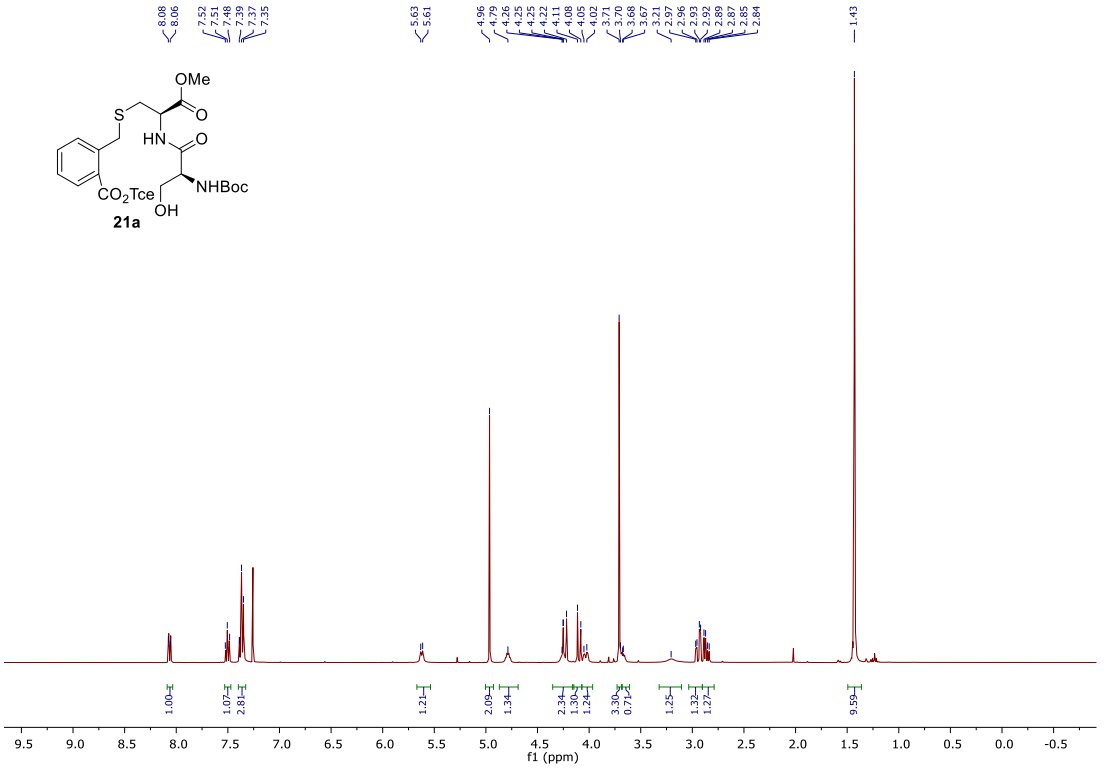
**20a** $^{1}$H NMR (400 MHz, $CDCl_3$)



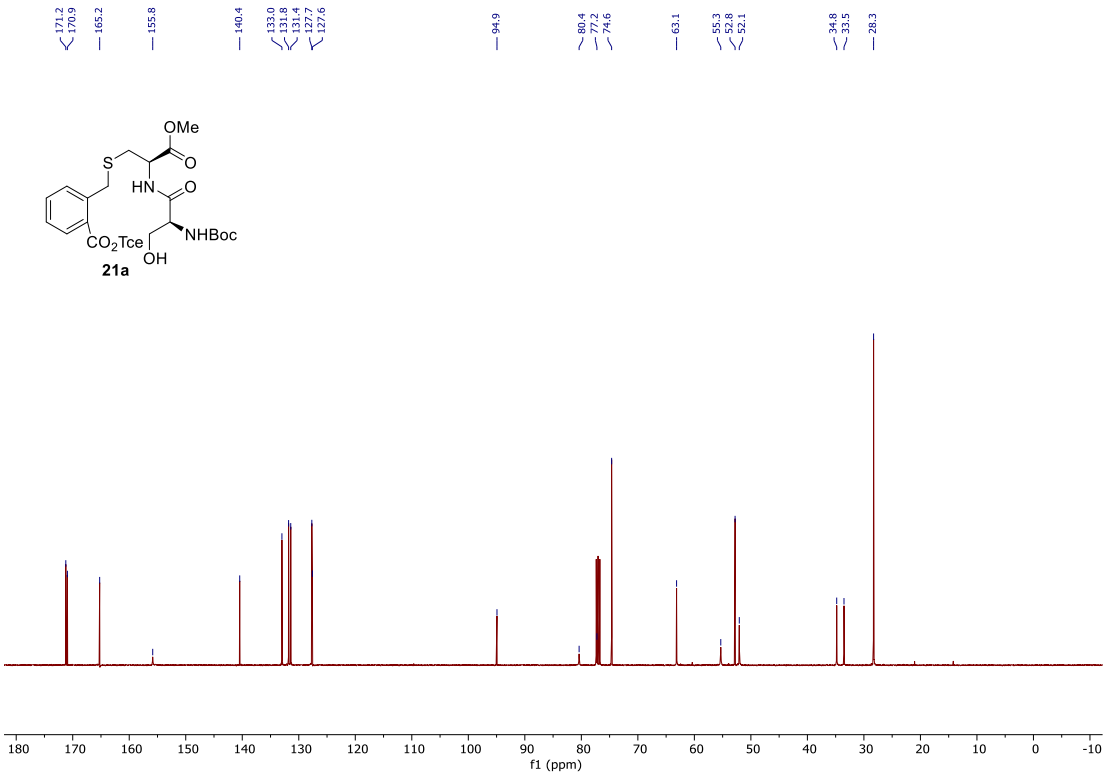

**20a** $^{13}$C NMR (101 MHz, $CDCl_3$)



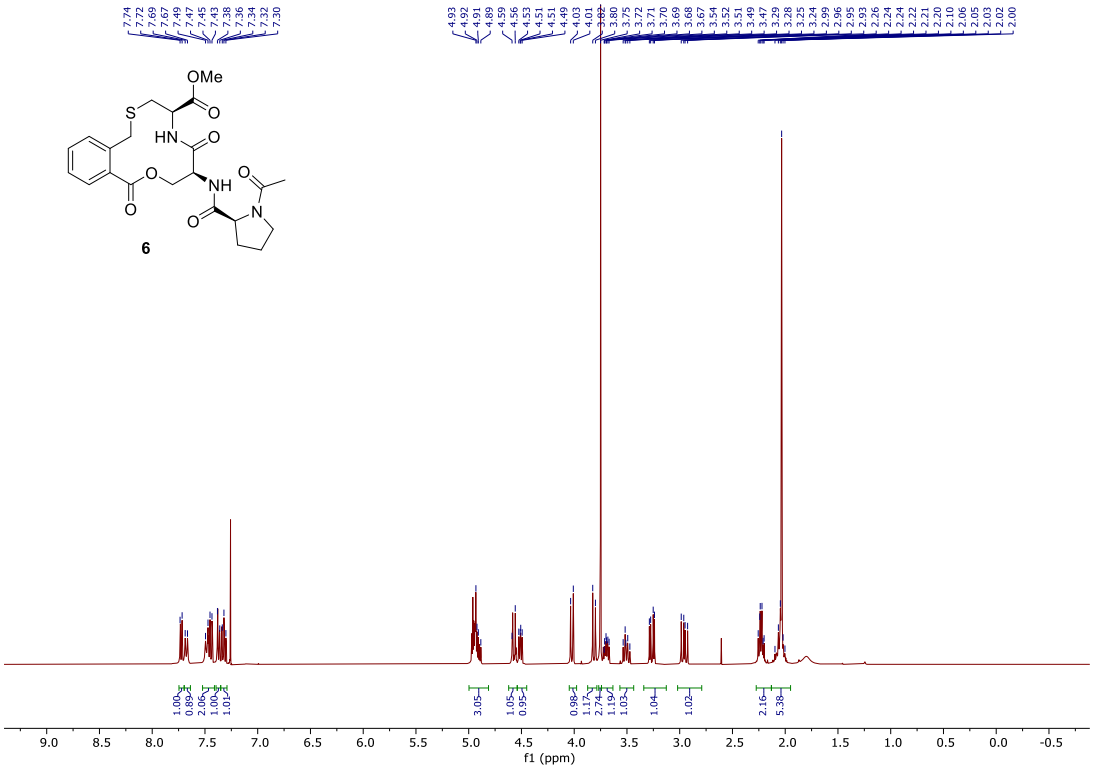
**21a** $^{1}$H NMR (400 MHz, $CDCl_3$)



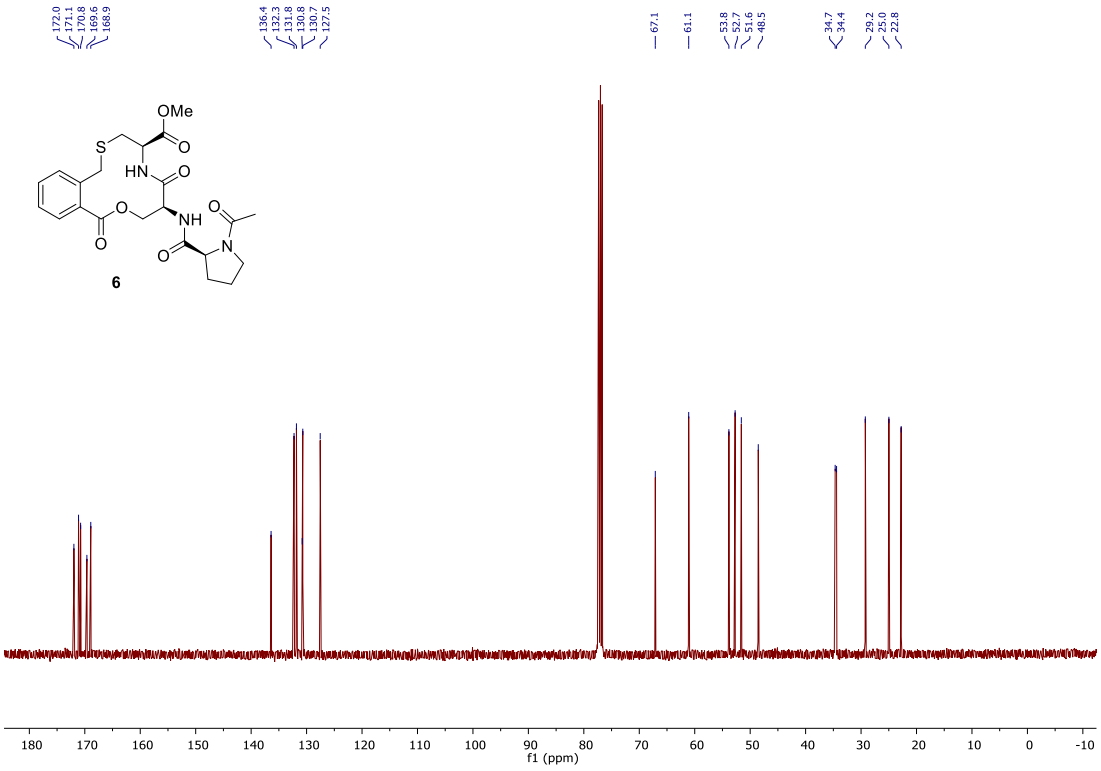

**21a** $^{13}$C NMR (101 MHz, $CDCl_3$)



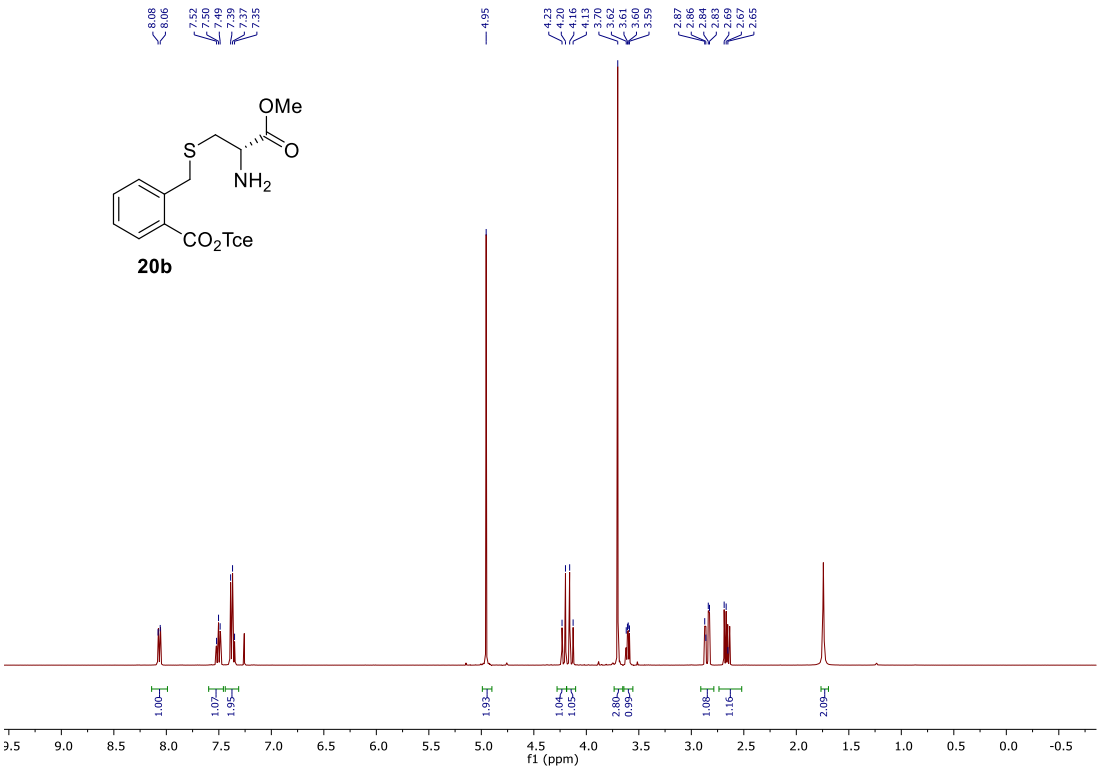
**6** $^{1}$H NMR (400 MHz, $CDCl_3$)



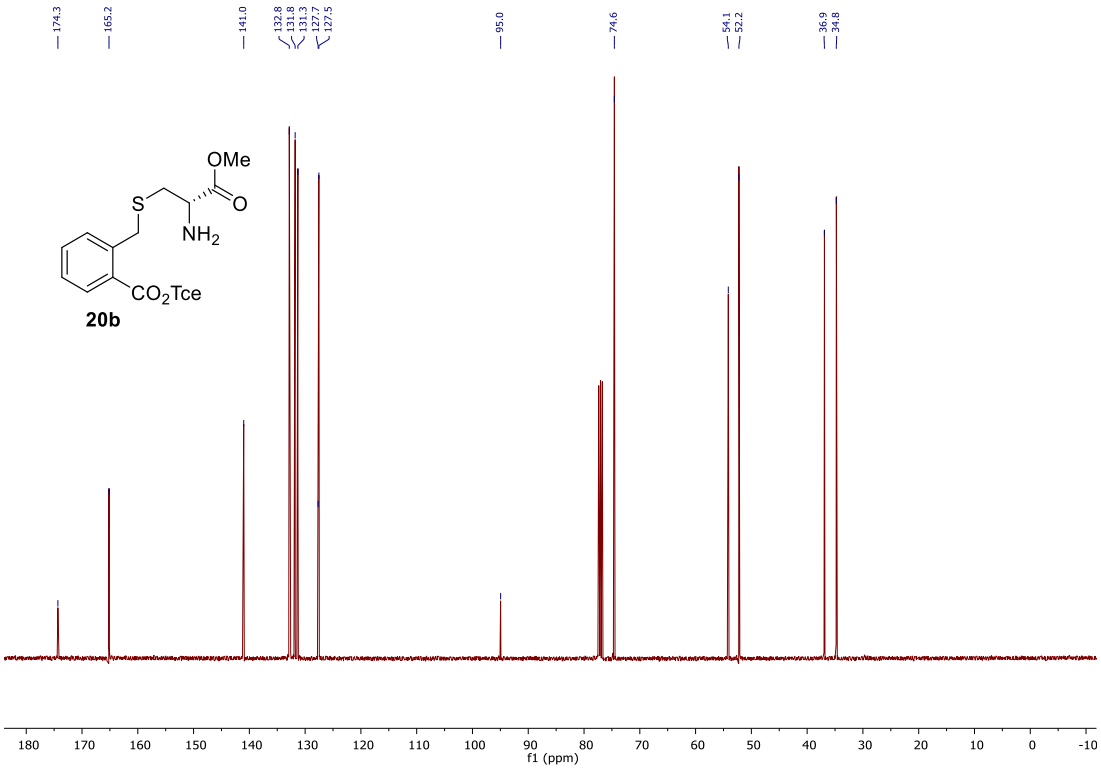

**6** $^{13}$C NMR (101 MHz, $CDCl_3$)



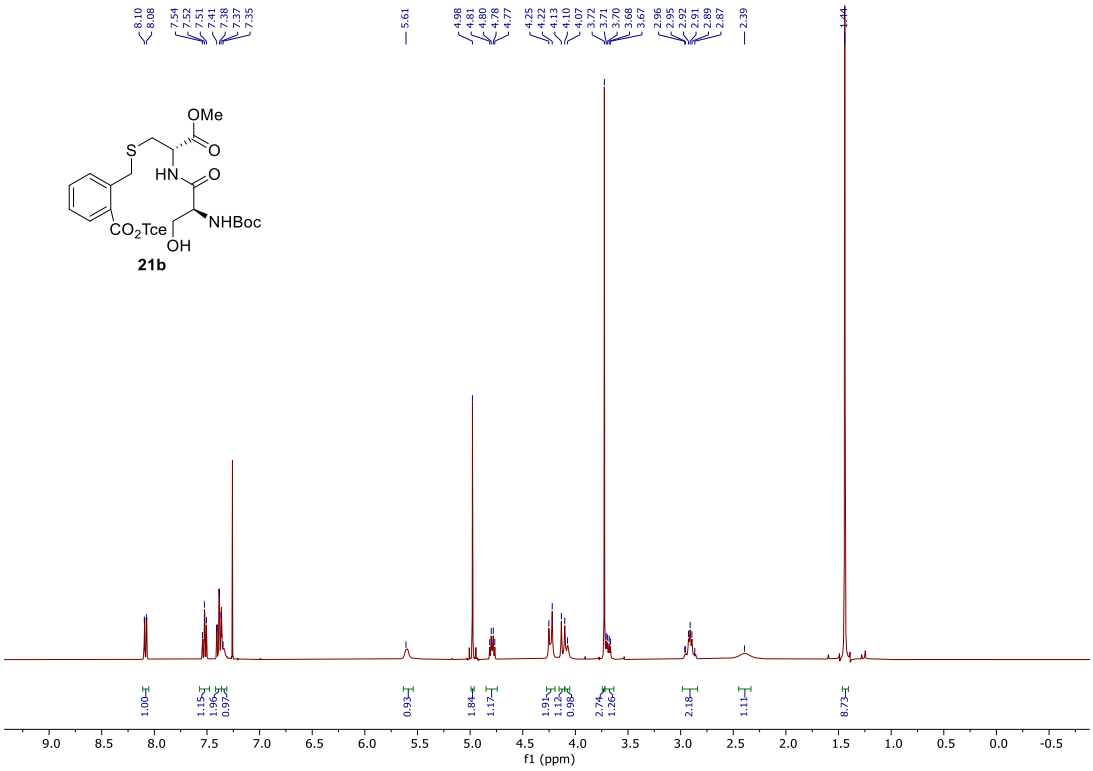
**20b** $^{1}$H NMR (400 MHz, $CDCl_3$)



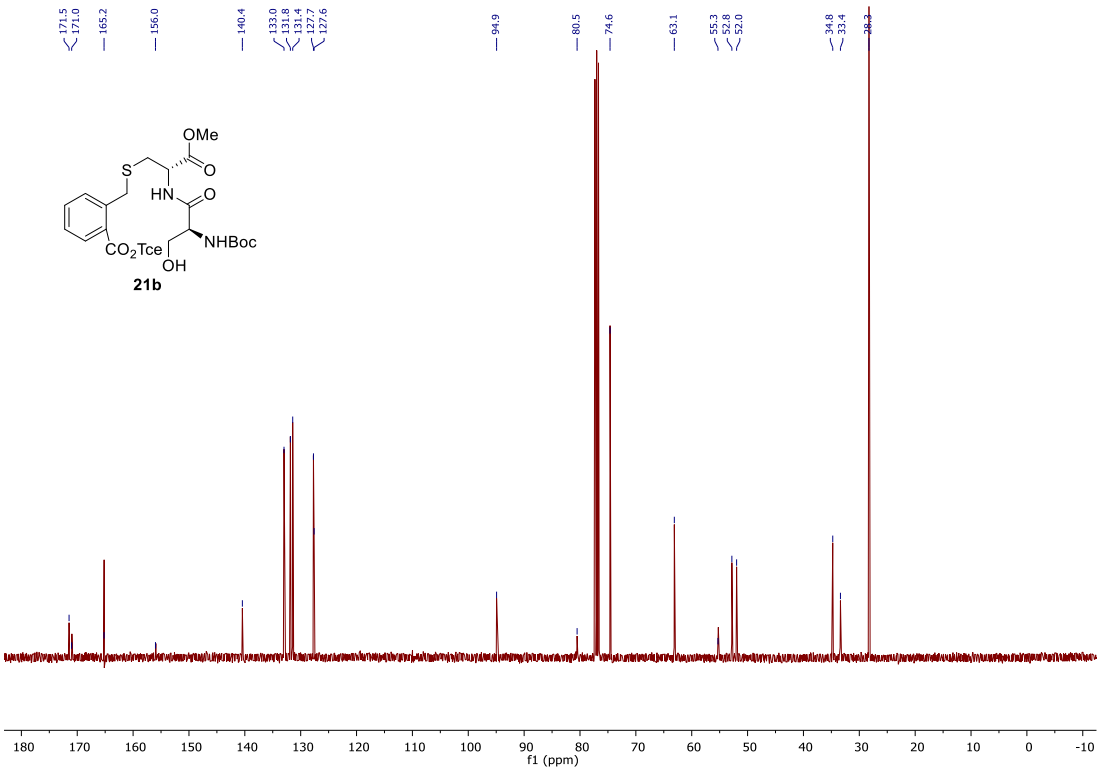

**20b** $^{13}$C NMR (101 MHz, $CDCl_3$)



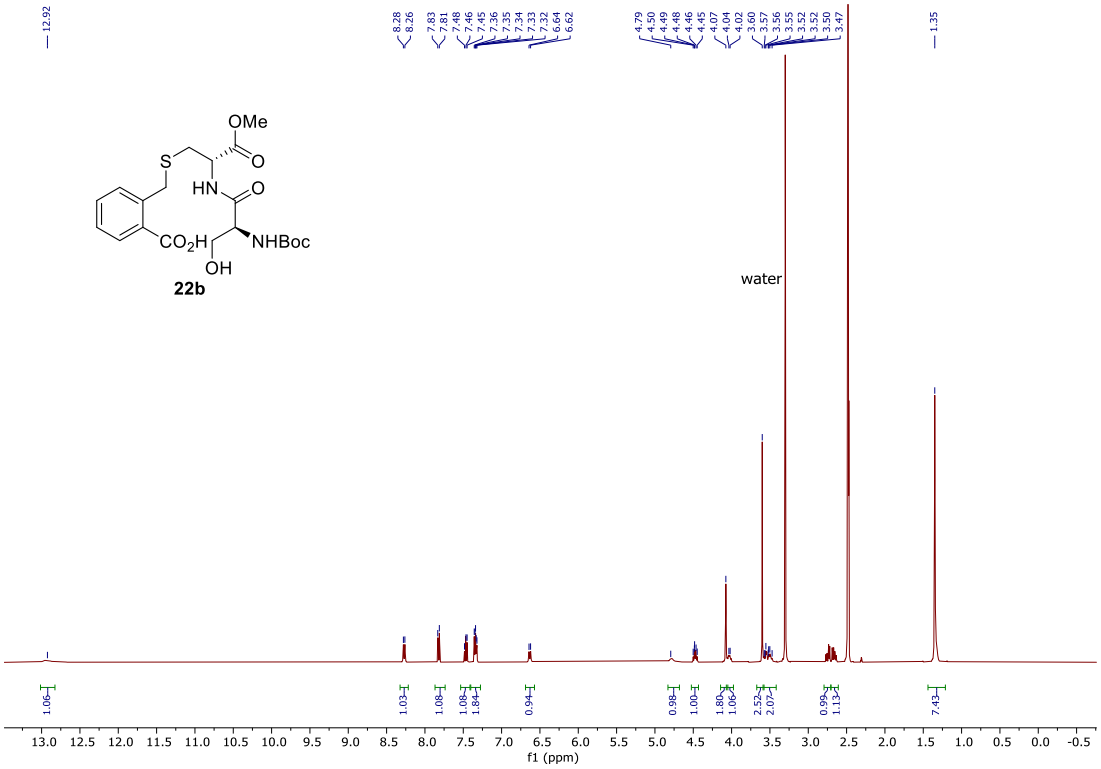
**21b** $^{1}$H NMR (400 MHz, $CDCl_3$)



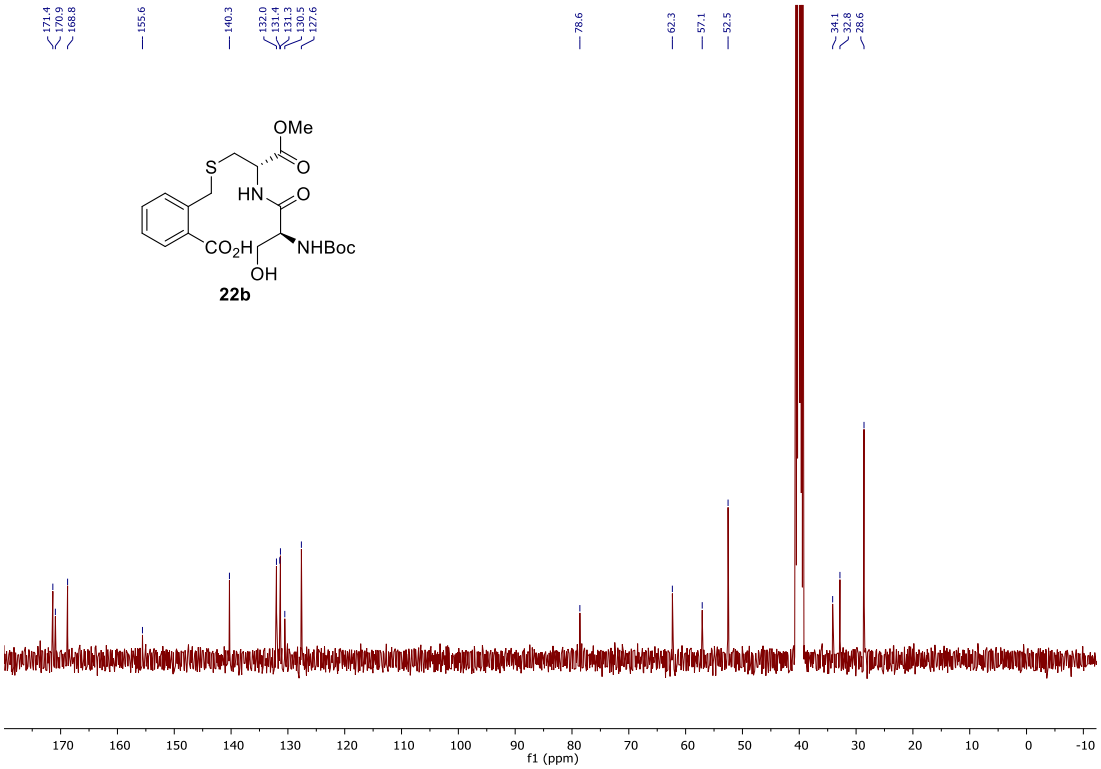

**21b** $^{13}$C NMR (101 MHz, $CDCl_3$)



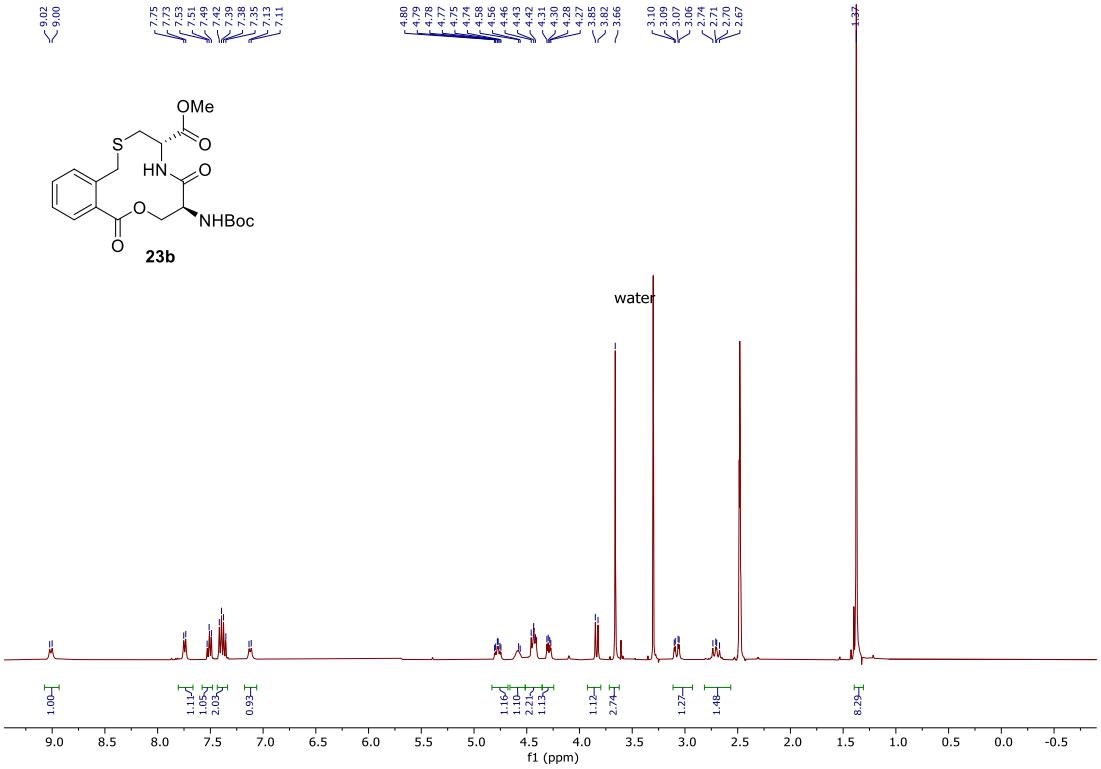
**22b** $^{1}$H NMR (400 MHz, DMSO-$d_6$)



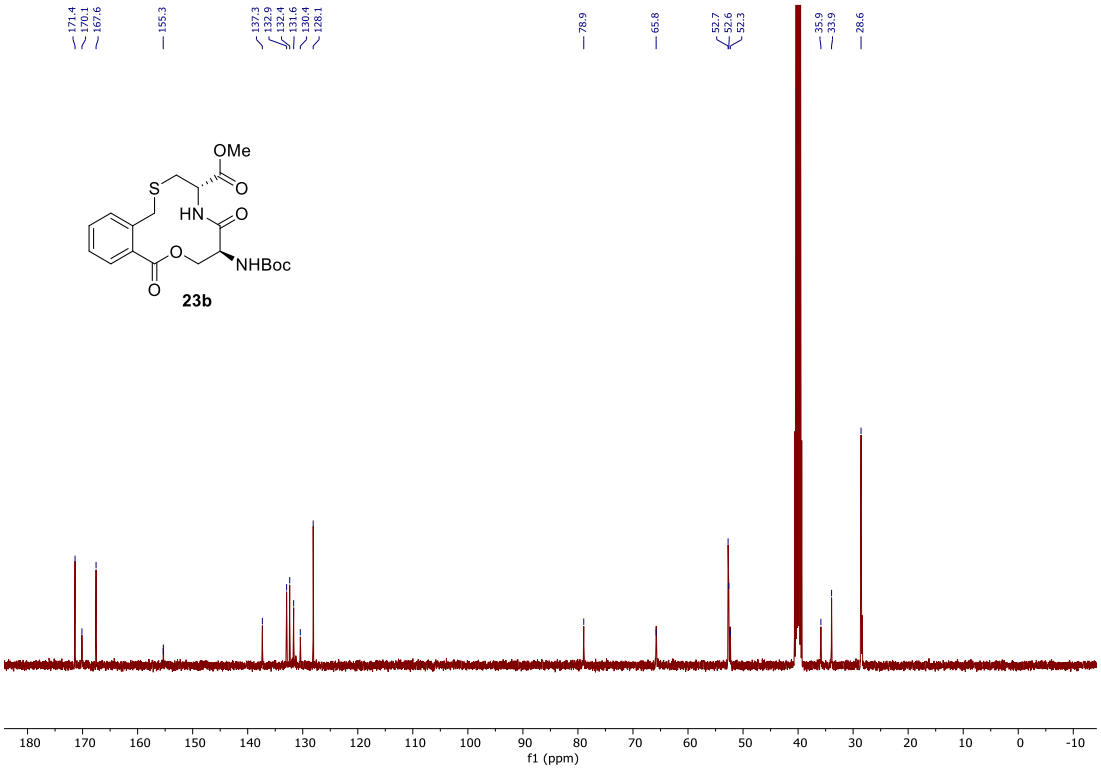

**22b** $^{13}$C NMR (101 MHz, DMSO- $d_6$)



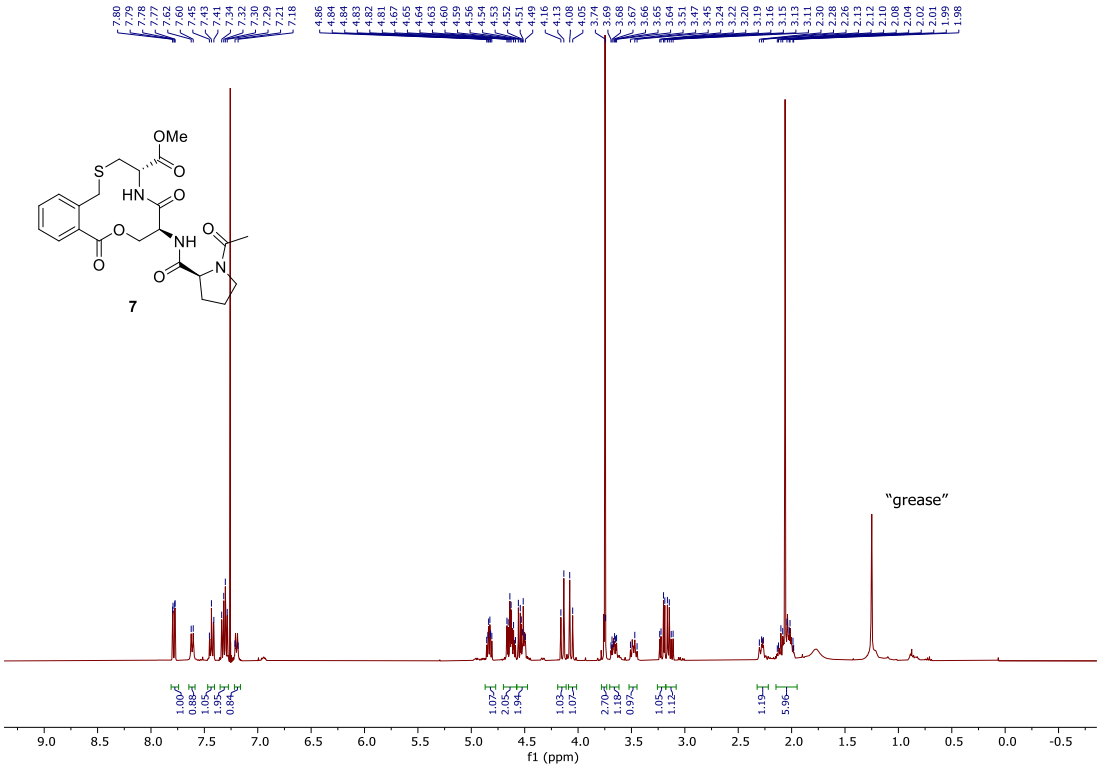
**23b** $^{1}$H NMR (400 MHz, DMSO-$d_6$)



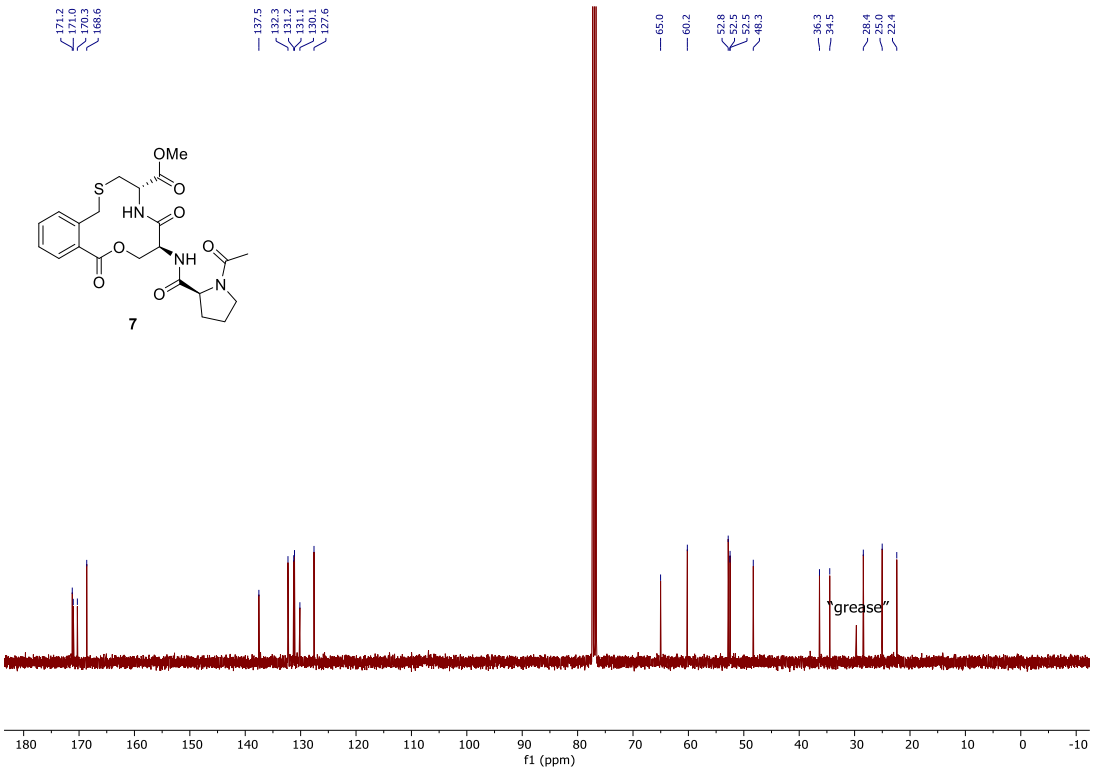

**23b** $^{13}$C NMR (101 MHz, DMSO-$d_6$)



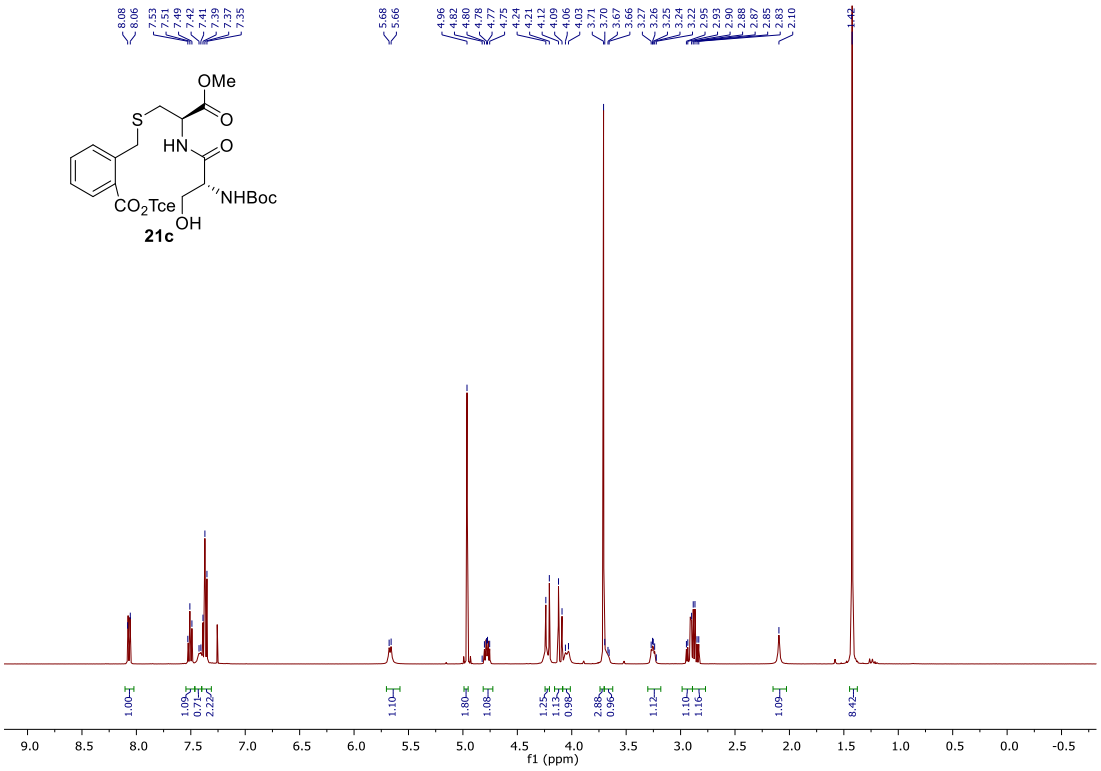
**7** $^{1}$H NMR (400 MHz, $CDCl_3$)



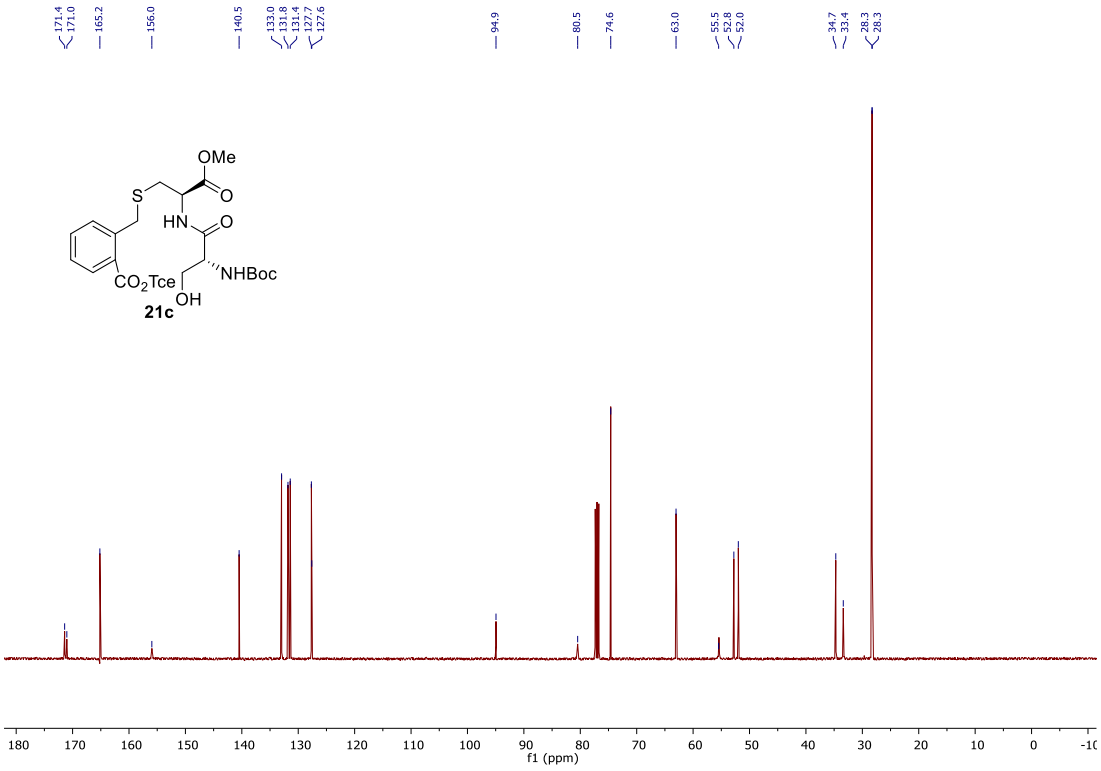

**7** $^{13}$C NMR (101 MHz, $CDCl_3$)



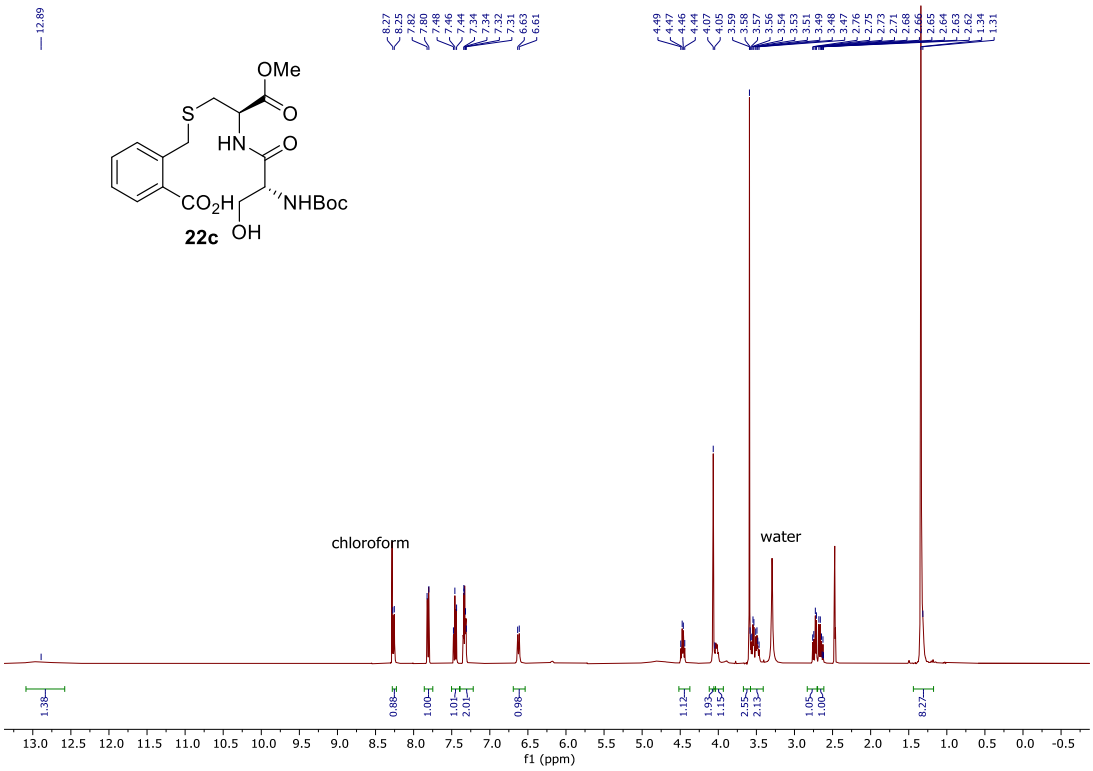
**21c** $^{1}$H NMR (400 MHz, $CDCl_3$)



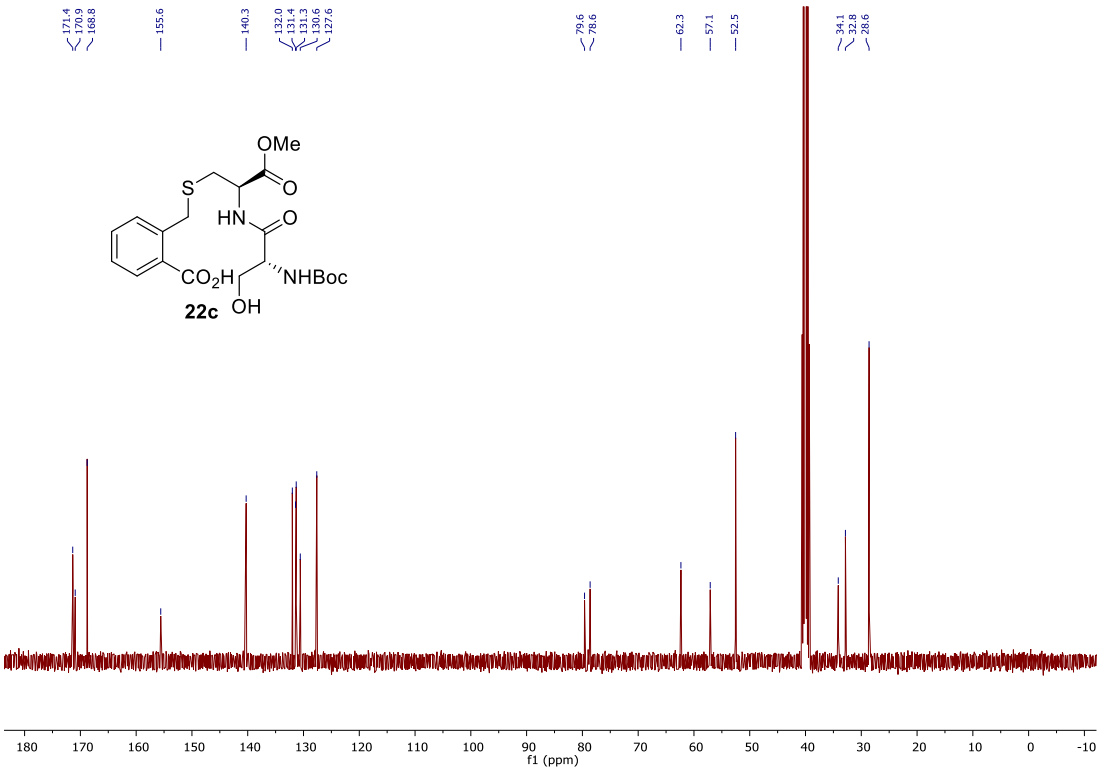

**21c** $^{13}$C NMR (101 MHz, $CDCl_3$)



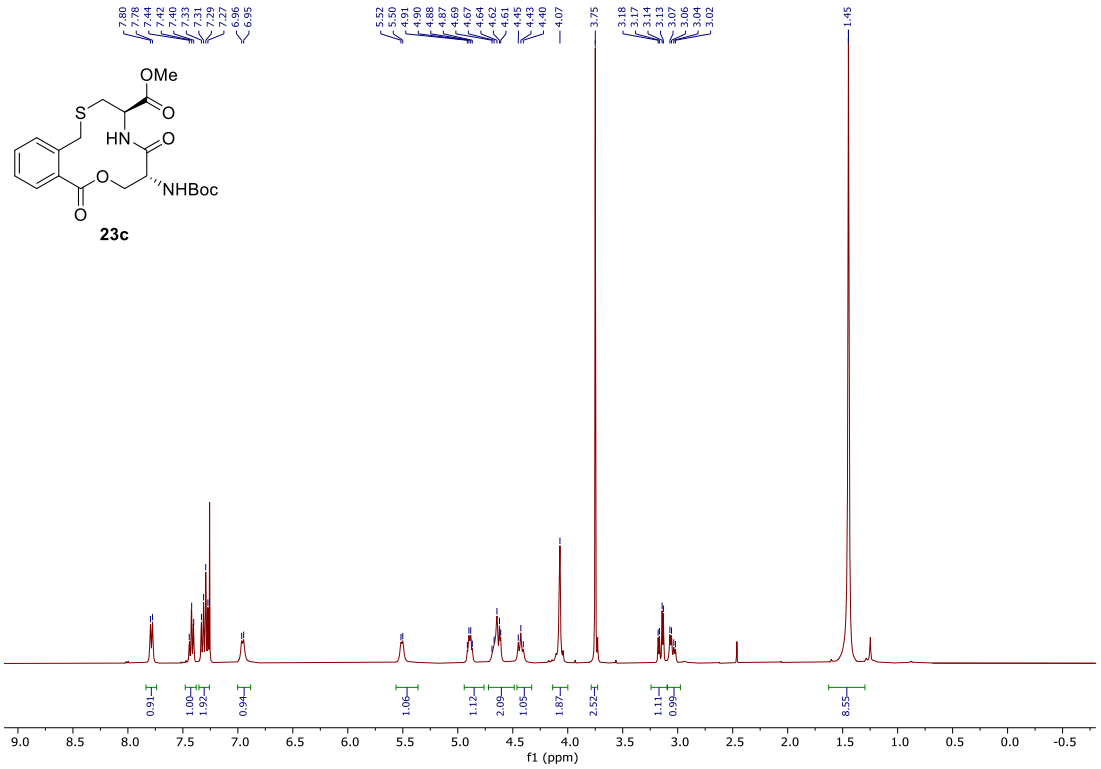
**22c** $^{1}$H NMR (400 MHz, DMSO-$d_6$)



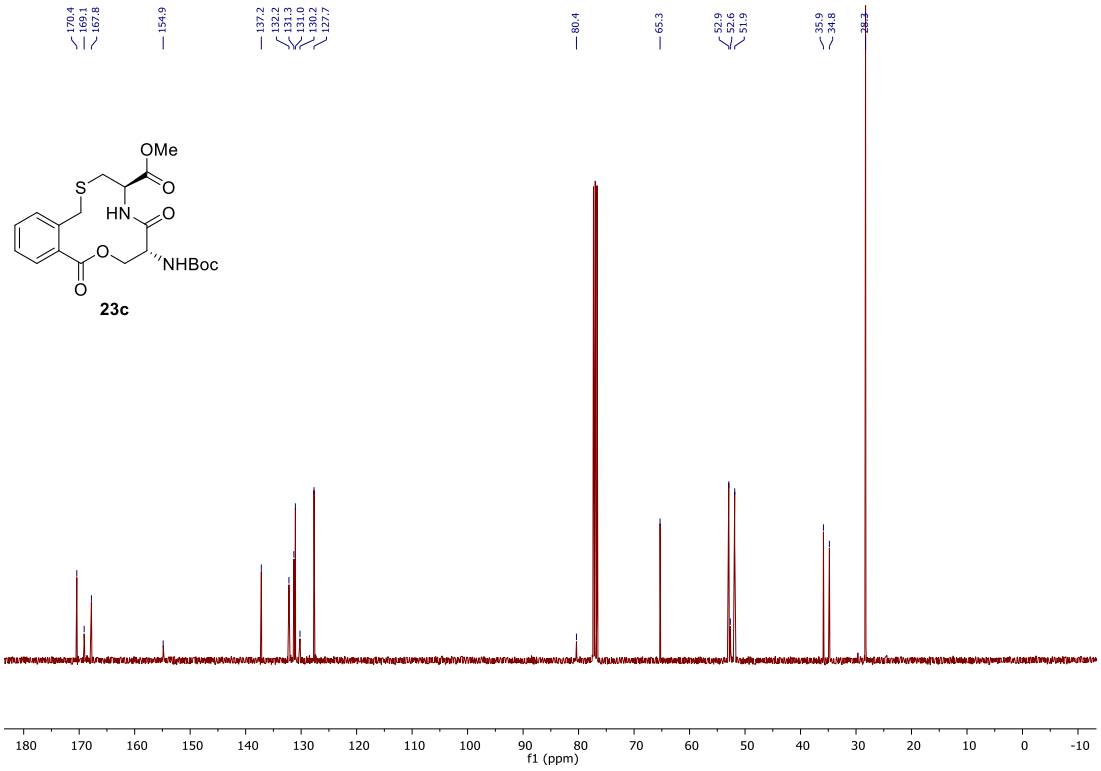

**22c** $^{13}$C NMR (101 MHz, DMSO-$d_6$)



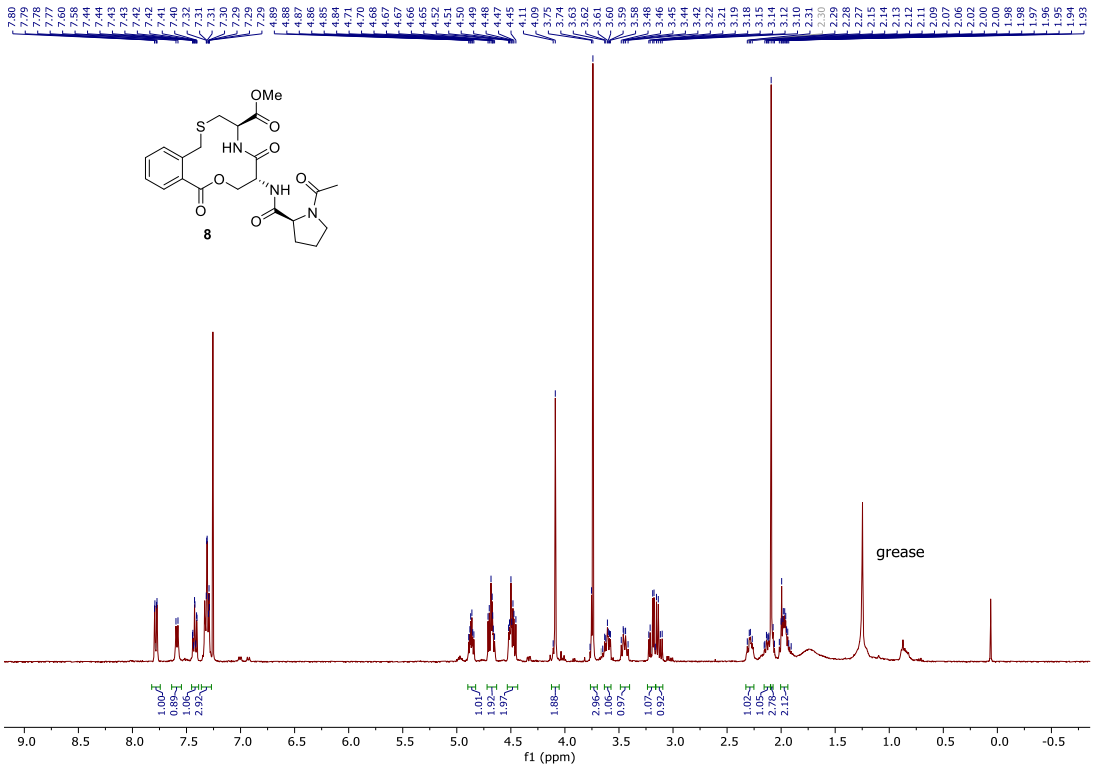
**23c** $^{1}$H NMR (400 MHz, $CDCl_3$)



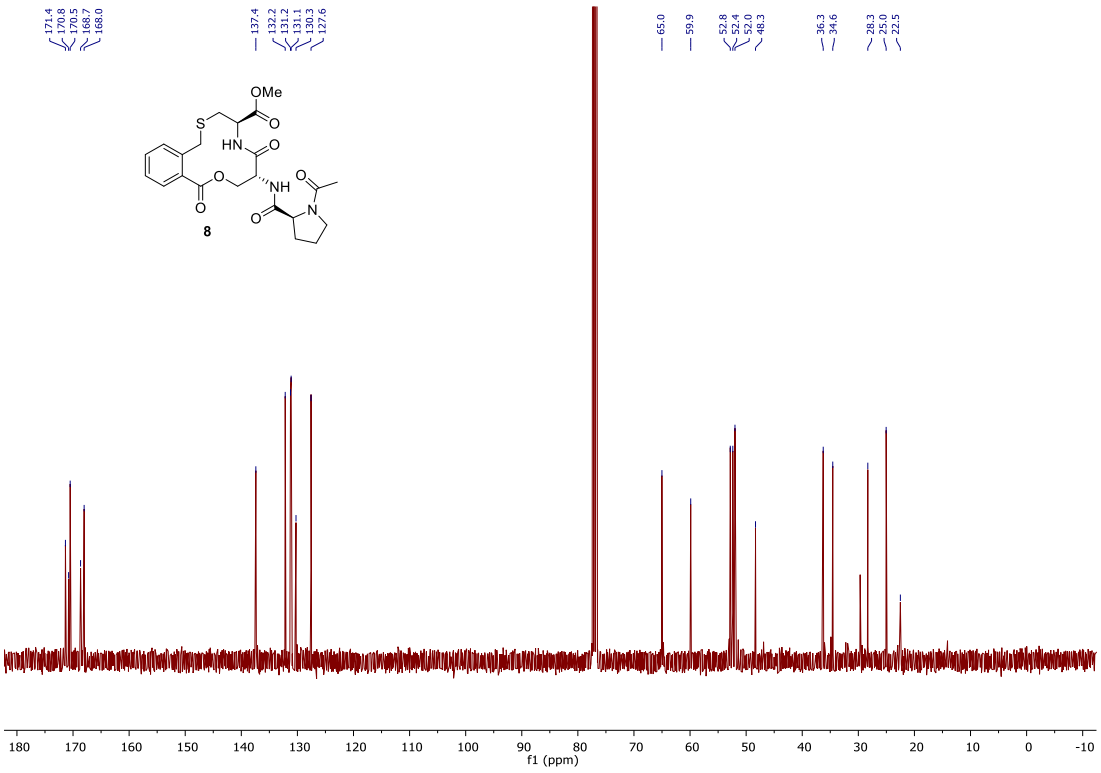

**23c** $^{13}$C NMR (101 MHz, $CDCl_3$)



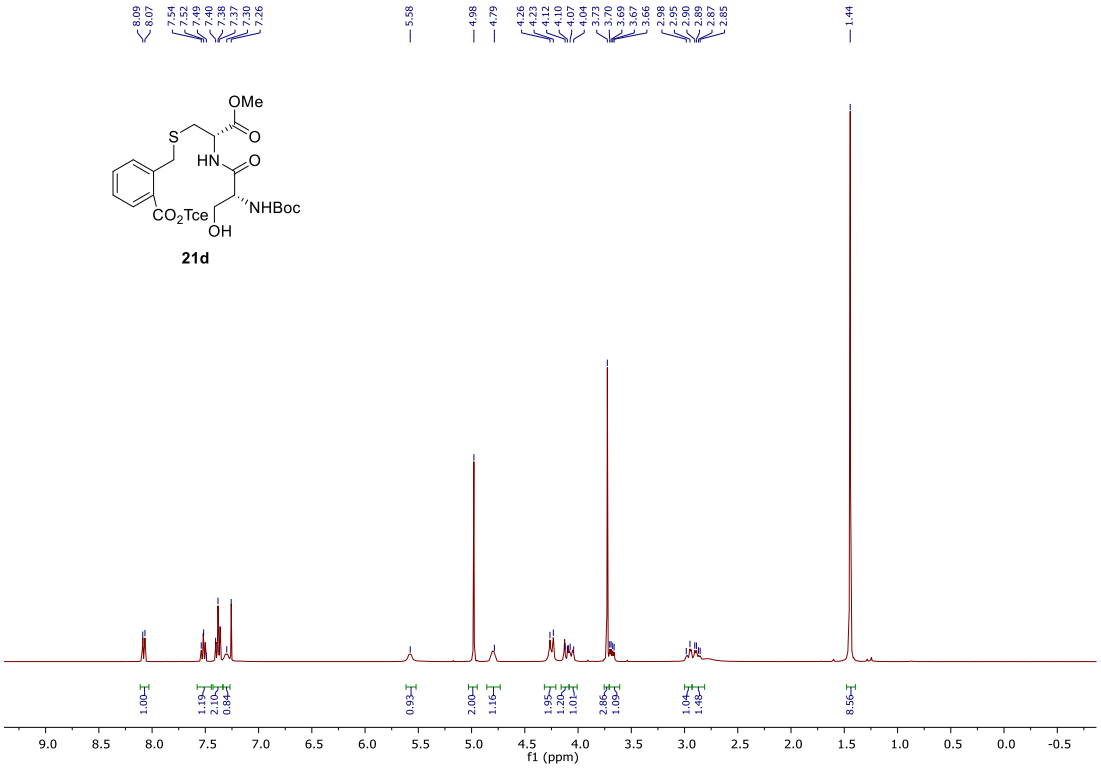
**8** $^{1}$H NMR (400 MHz, $CDCl_3$)



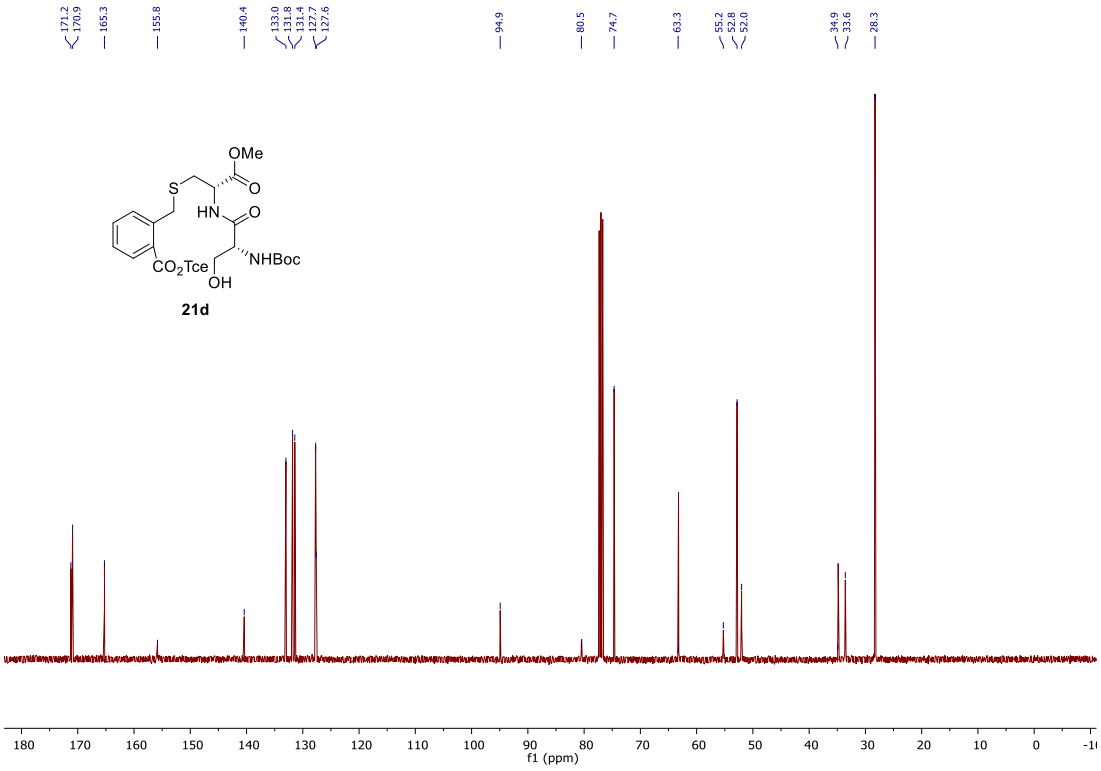

**8** $^{13}$C NMR (101 MHz, $CDCl_3$)



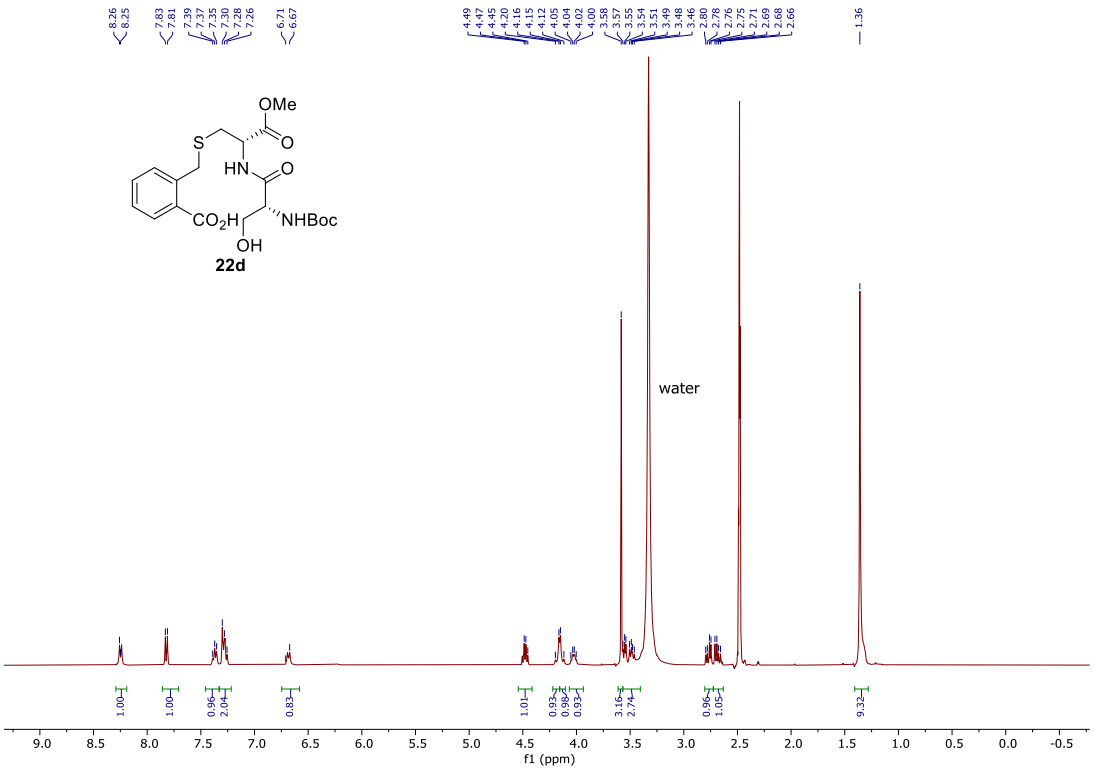
**21d** $^{1}$H NMR (400 MHz, $CDCl_3$)



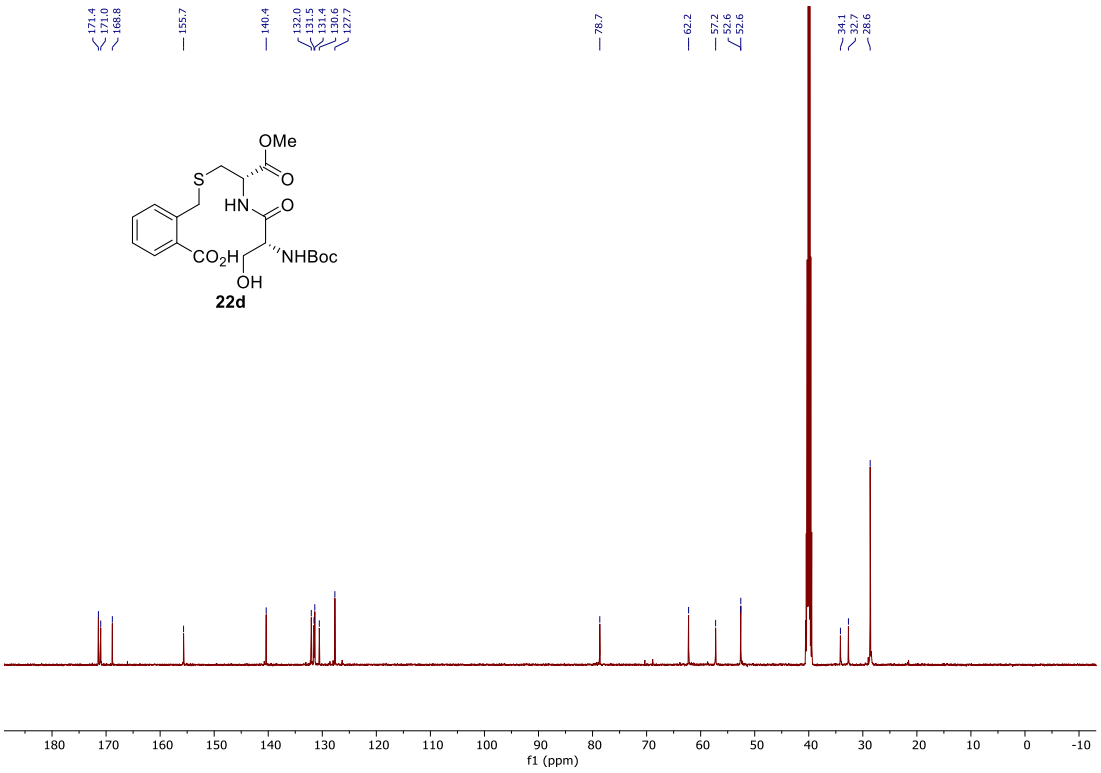

**21d** $^{13}$C NMR (101 MHz, $CDCl_3$)



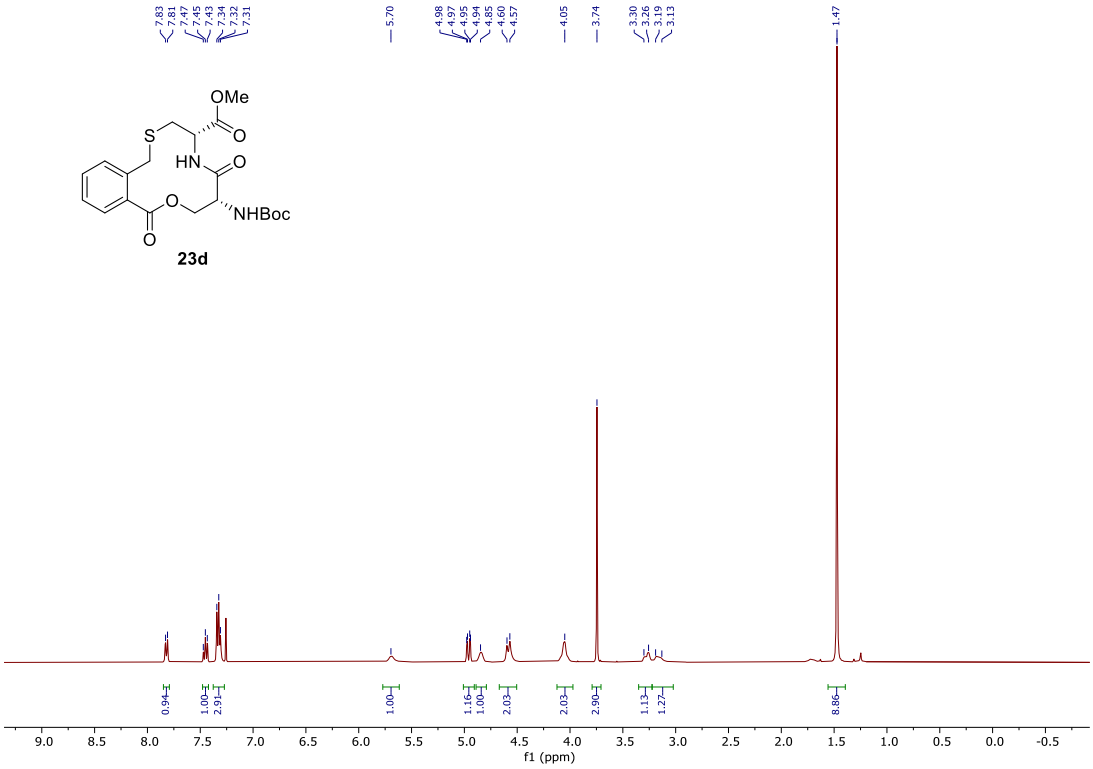
**22d** $^{1}$H NMR (400 MHz, DMSO-$d_6$)



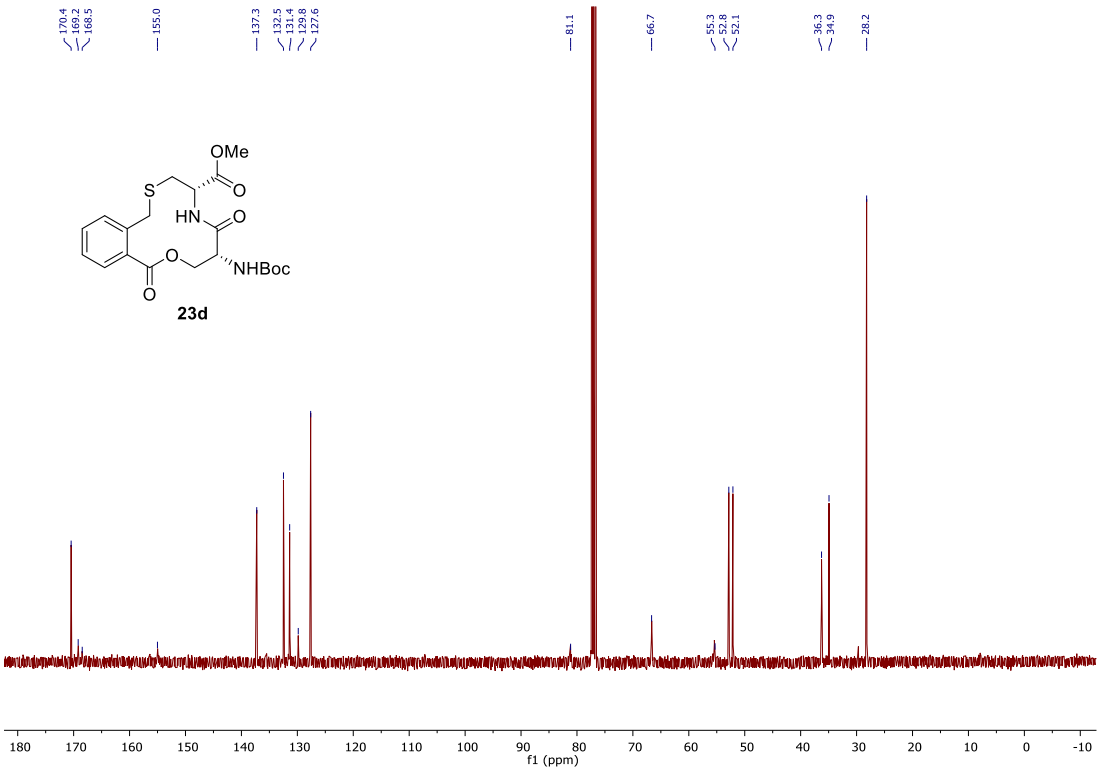

**22d** $^{13}$C NMR (101 MHz, DMSO-$d_6$)



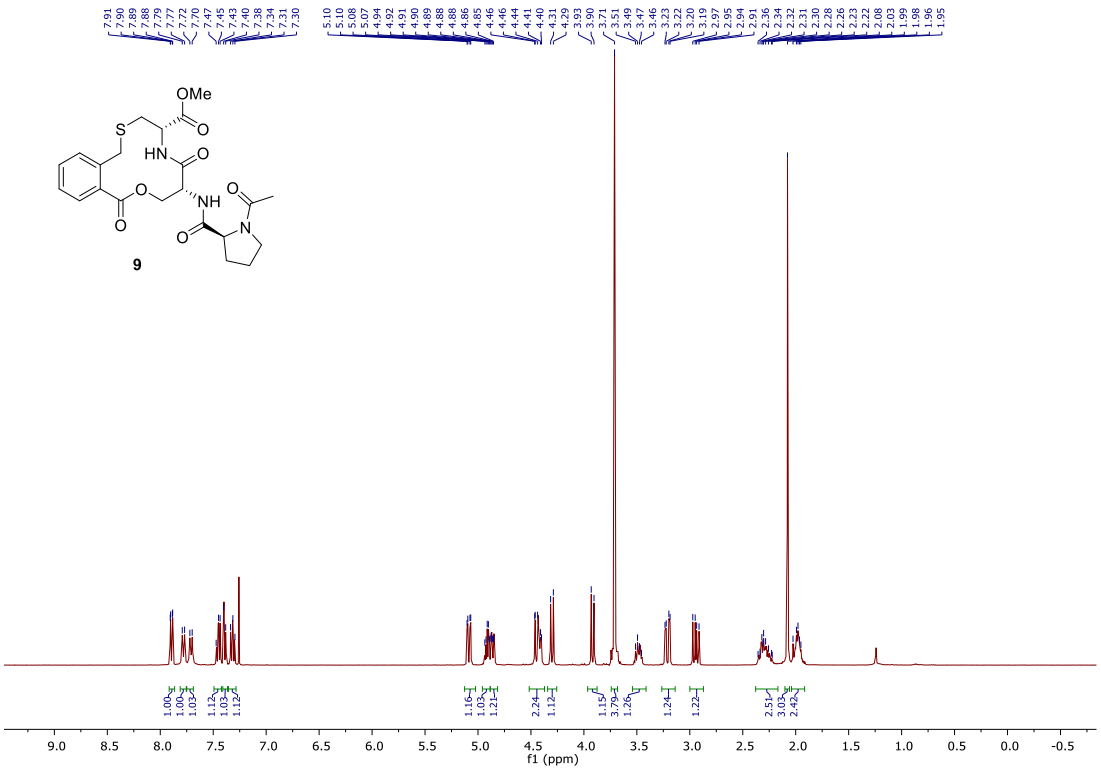
**23d** $^{1}$H NMR (400 MHz, $CDCl_3$)



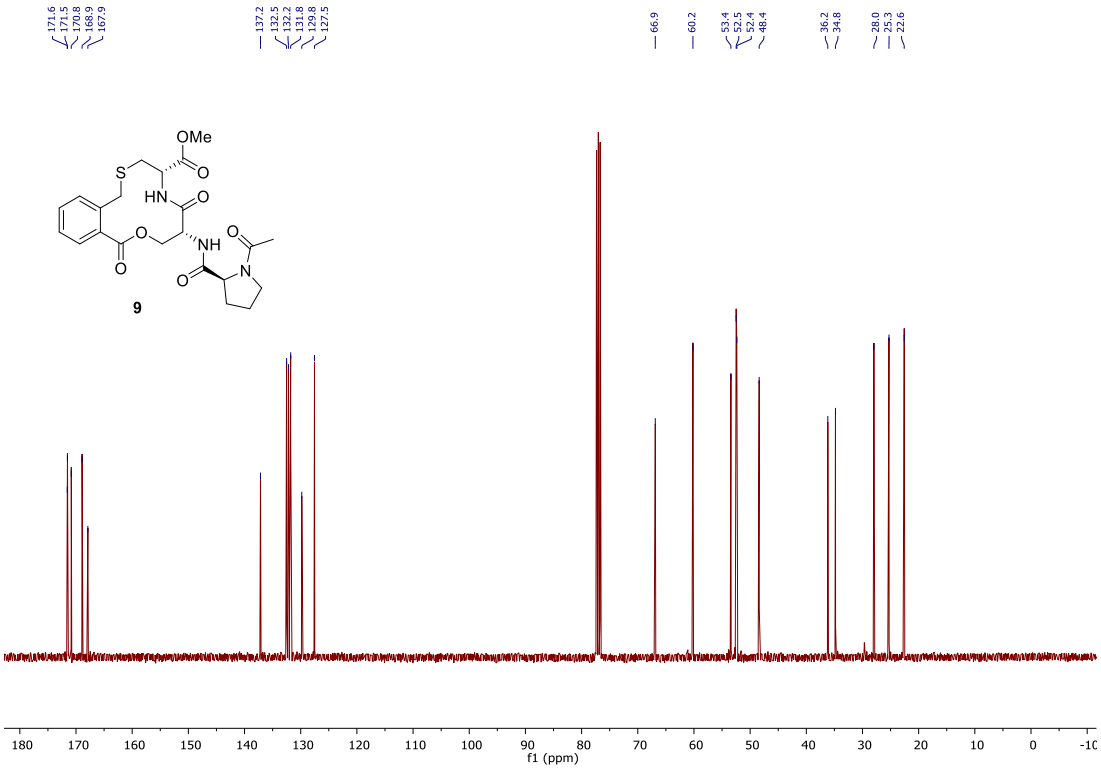

**23d** $^{13}$C NMR (101 MHz, $CDCl_3$)



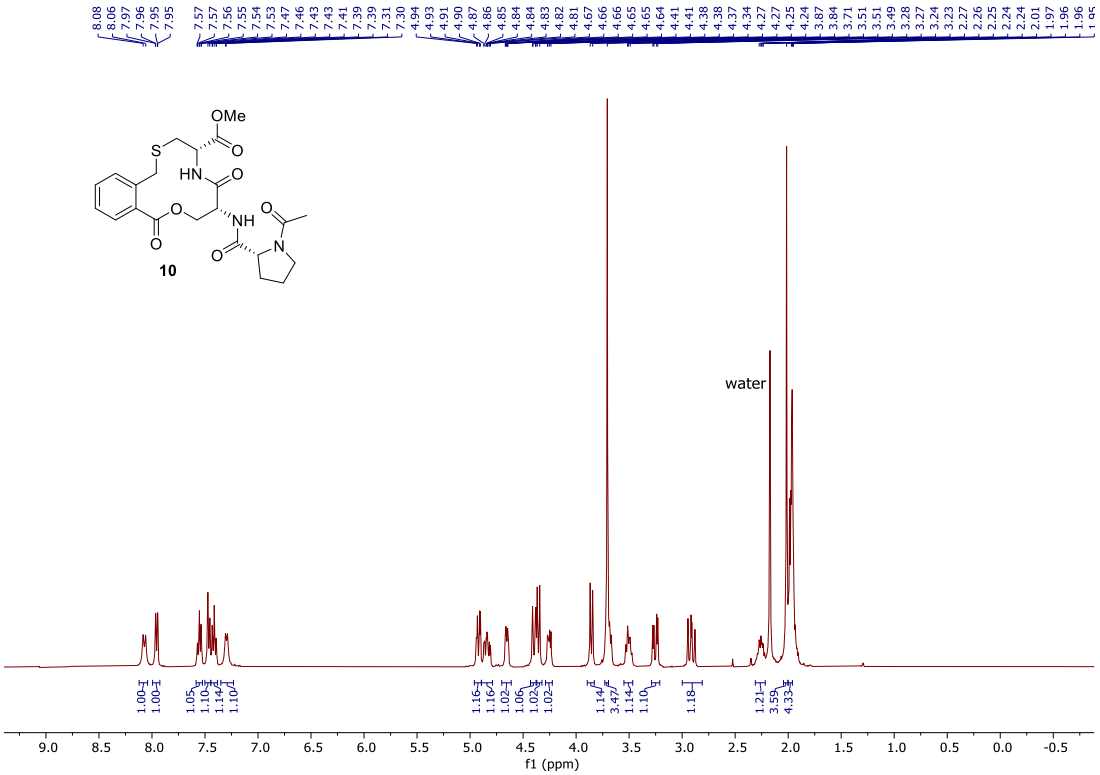
**9** $^{1}$H NMR (400 MHz, $CDCl_3$)



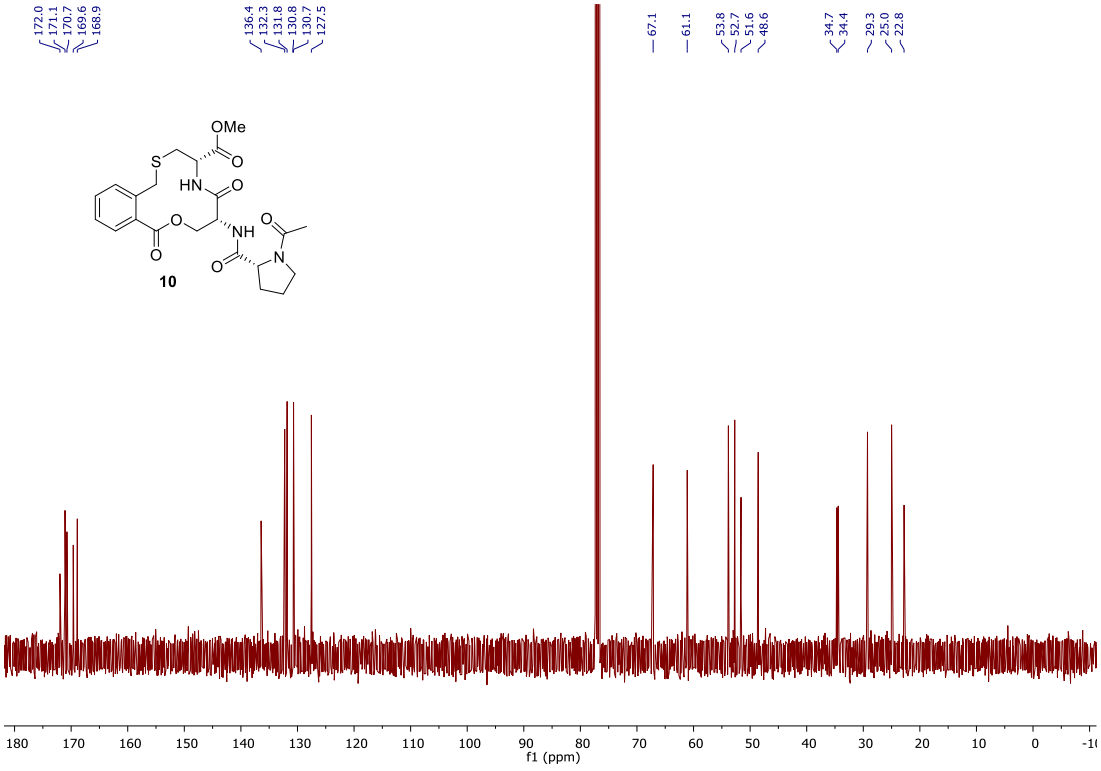

**9** $^{13}$C NMR (101 MHz, $CDCl_3$)



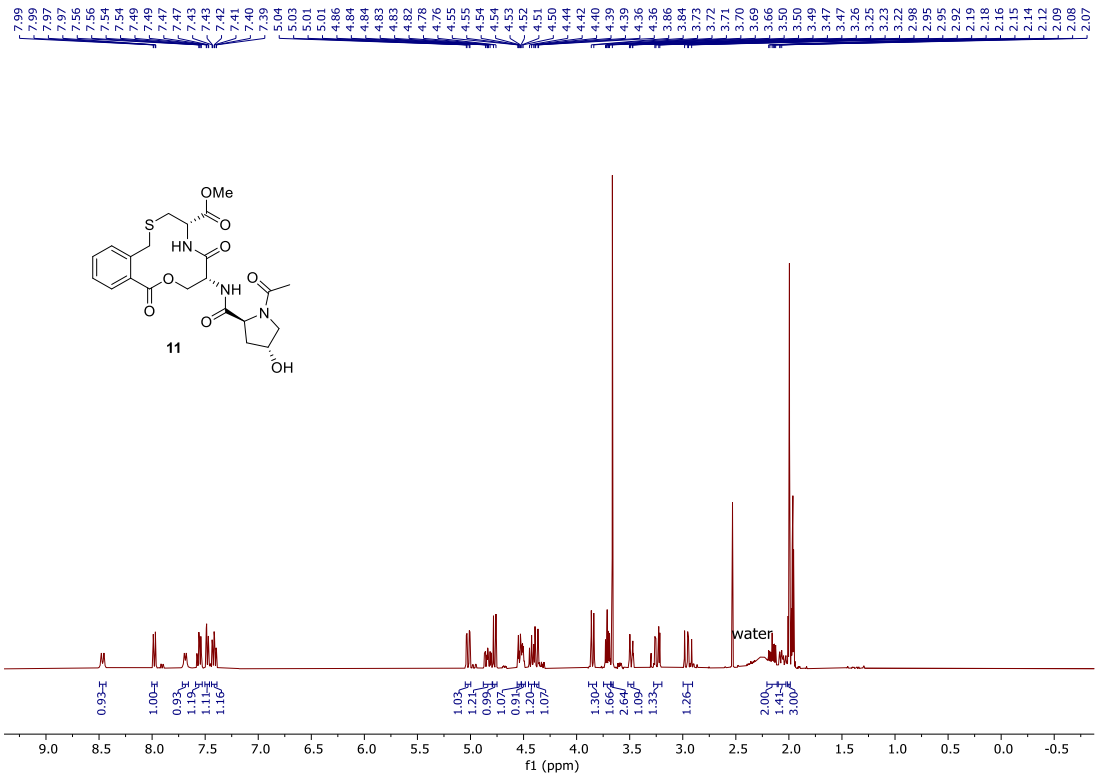
**10** $^{1}$H NMR (400 MHz, $CD_3CN$)



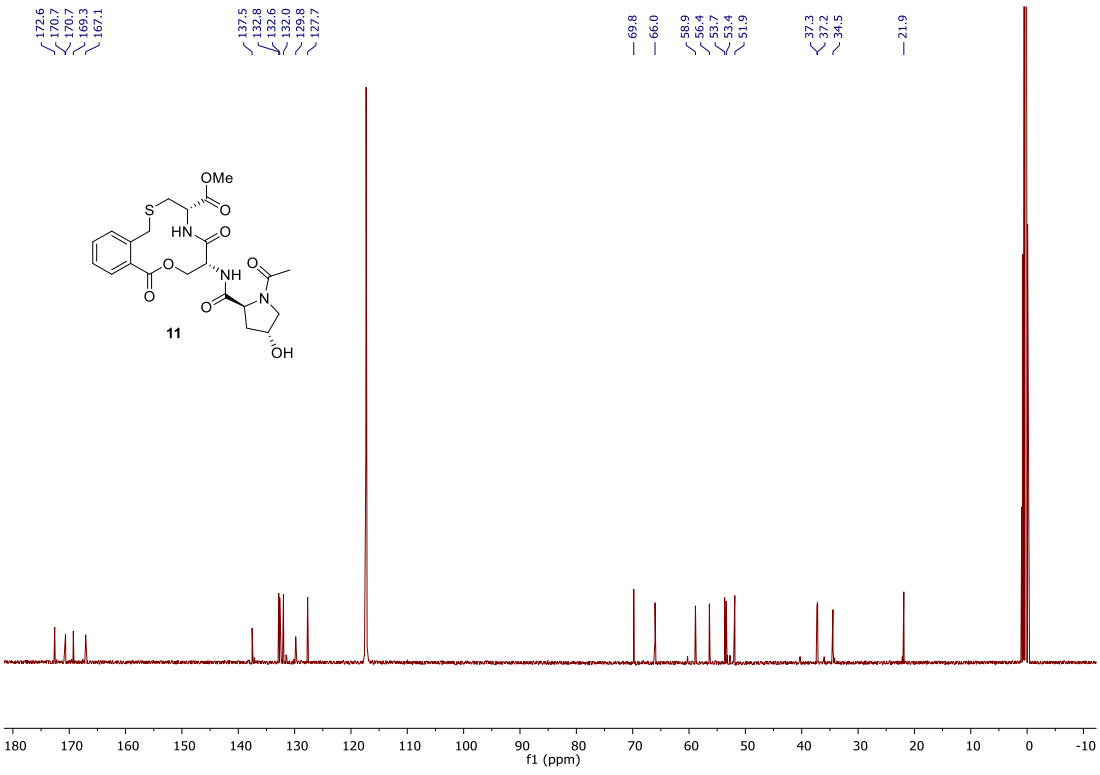

**10** $^{13}$C NMR (101 MHz, $CDCl_3$)



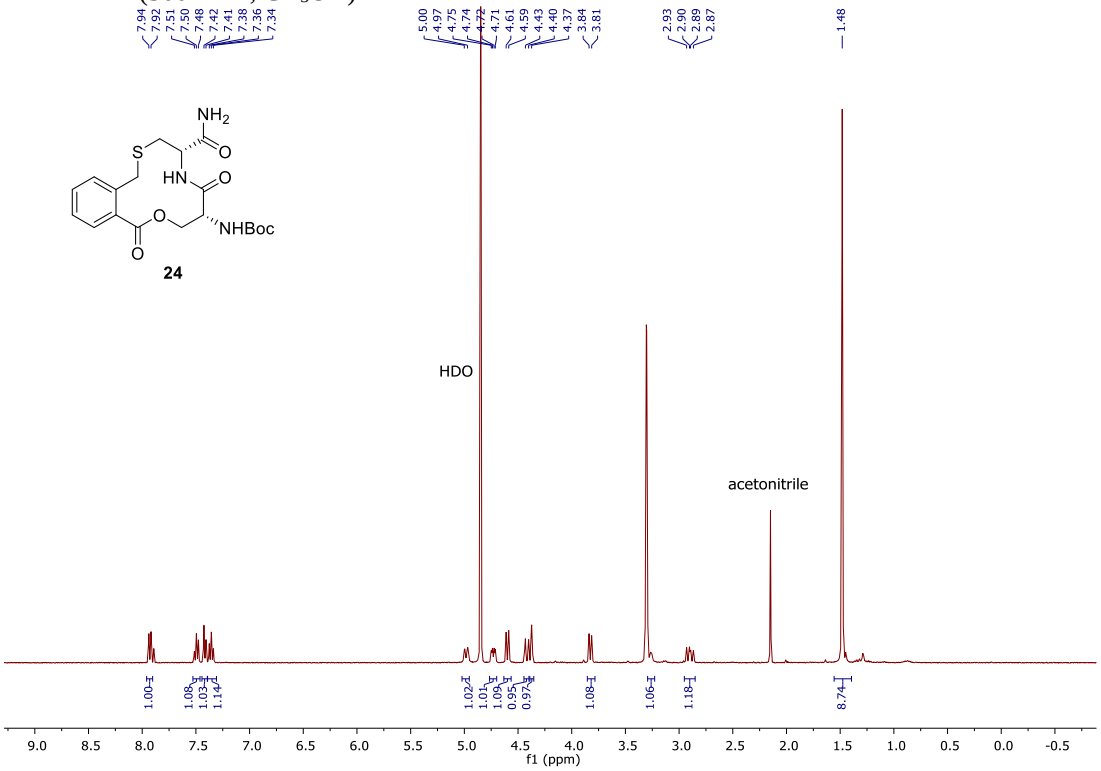
**11** $^{1}$H NMR (400 MHz, $CD_3CN$)



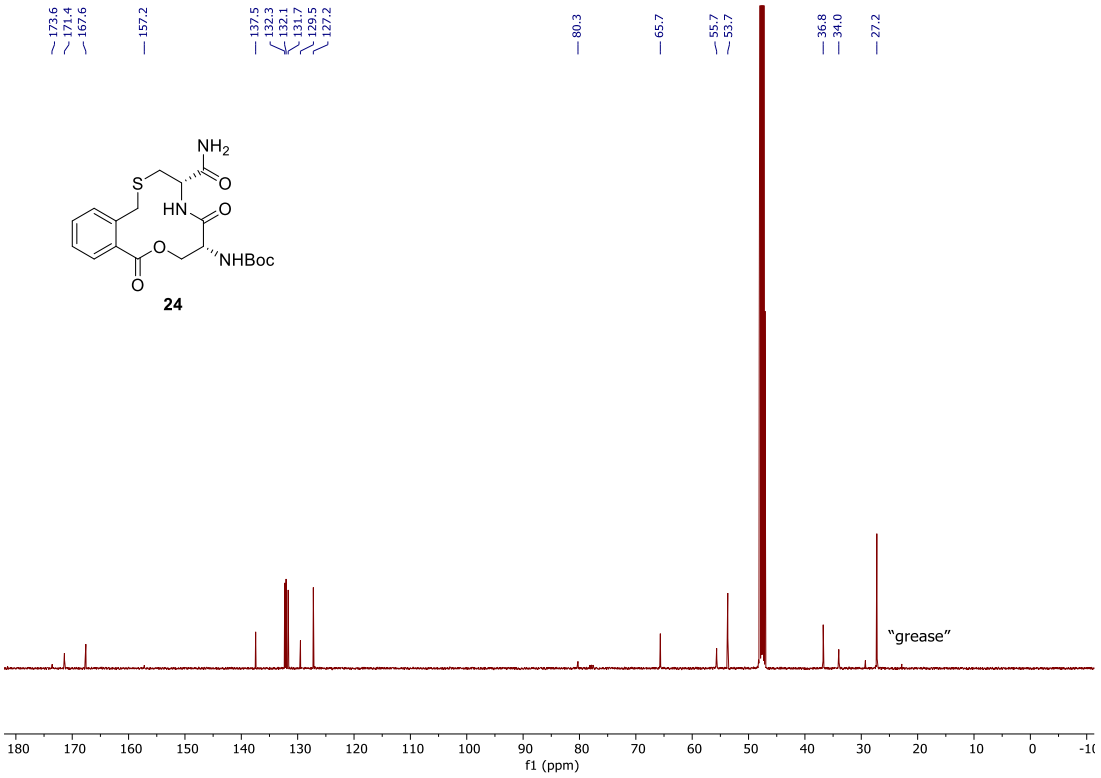

**11** $^{13}$C NMR (101 MHz, $CD_3CN$)



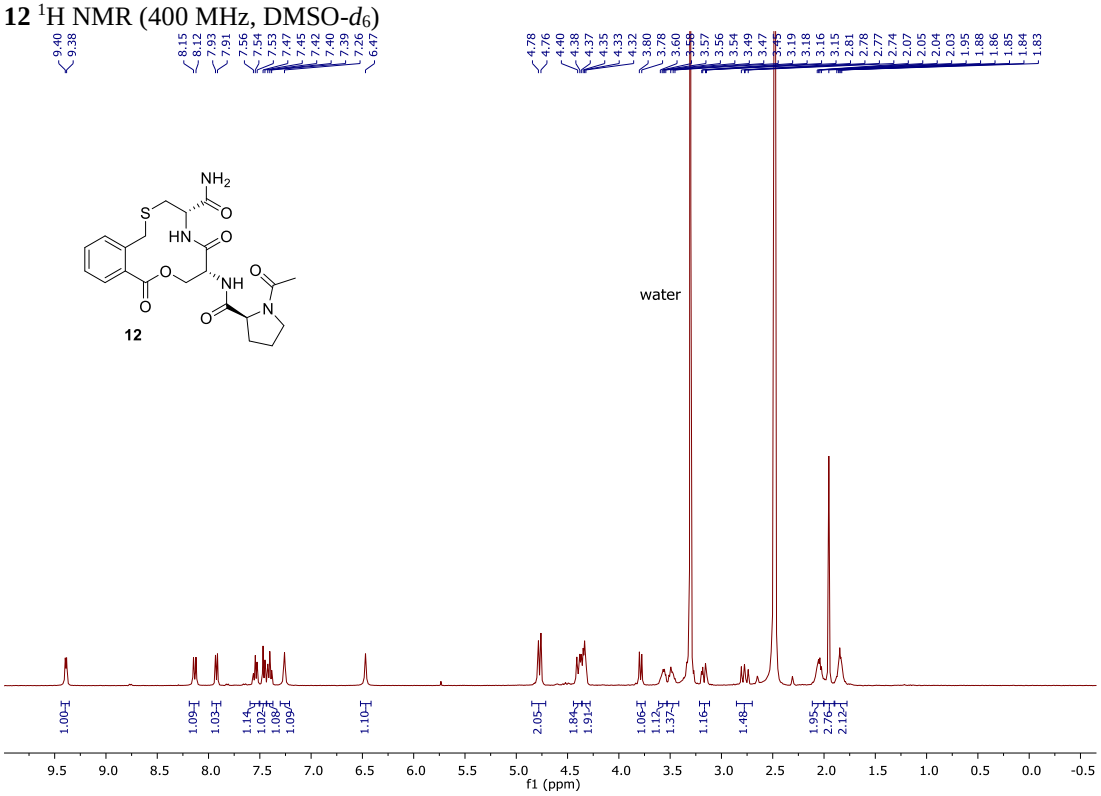
**24** $^{1}$H NMR (500 MHz, $CD_3OD$)



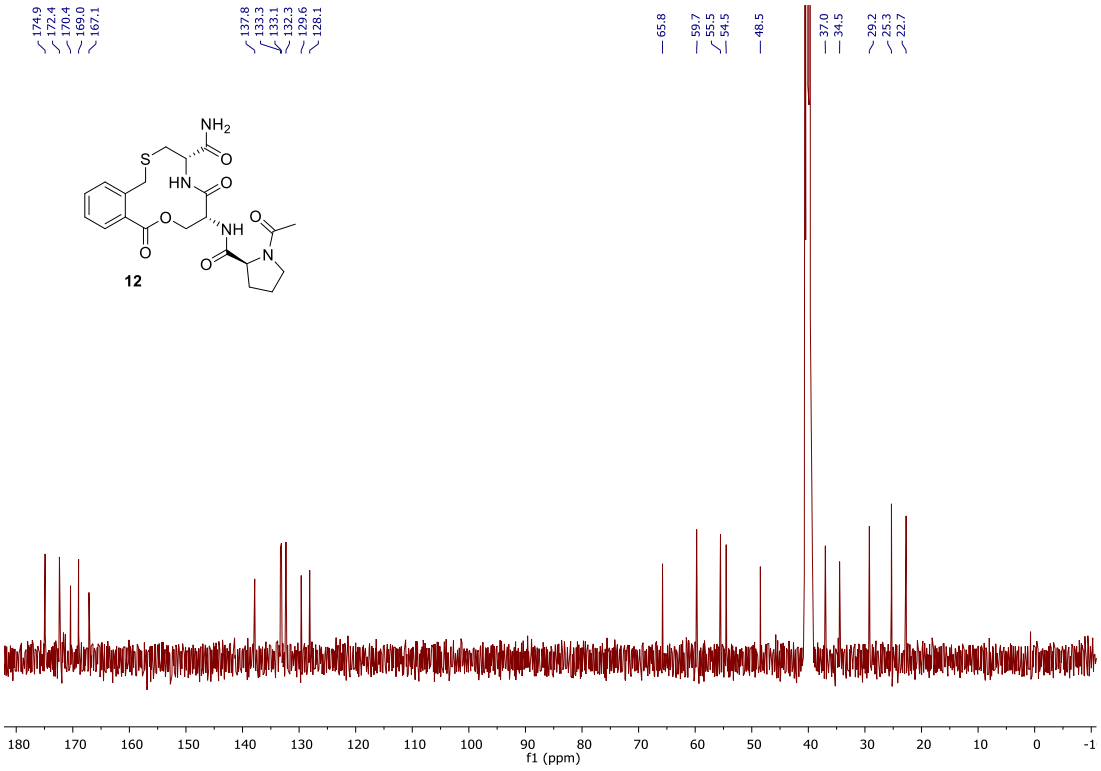

**24** $^{1}$H NMR (126 MHz, $CD_3OD$)



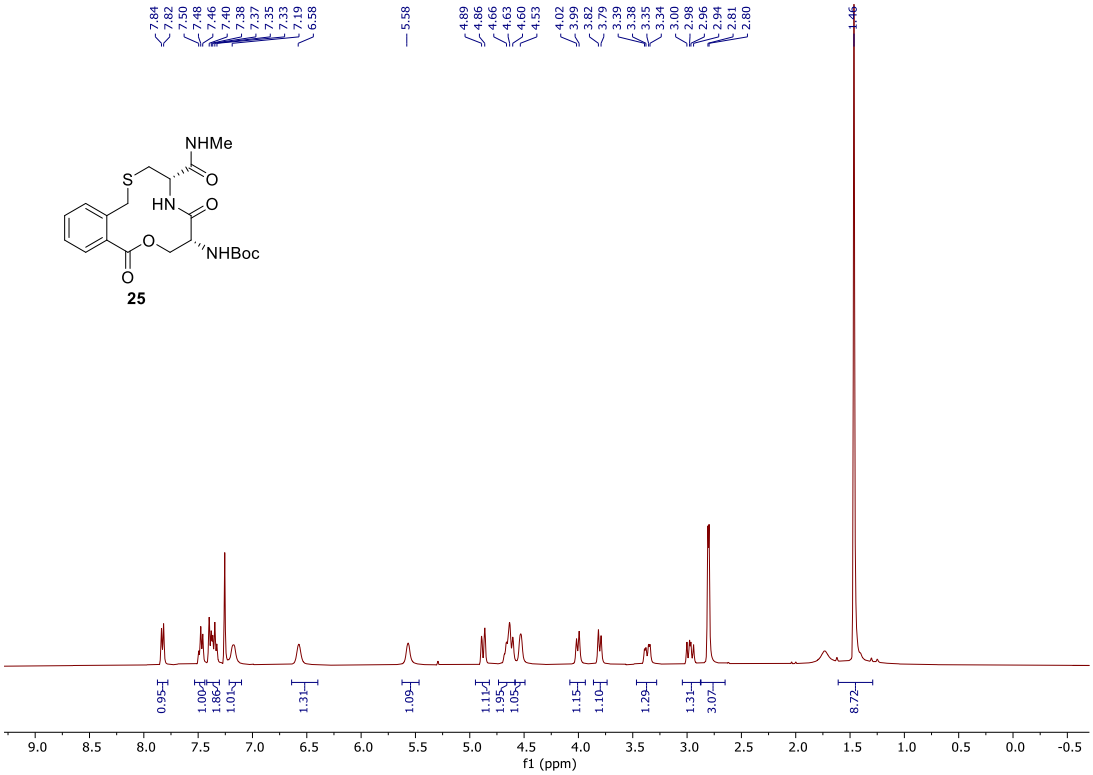
**12** $^{1}$H NMR (400 MHz, DMSO-$d_6$)



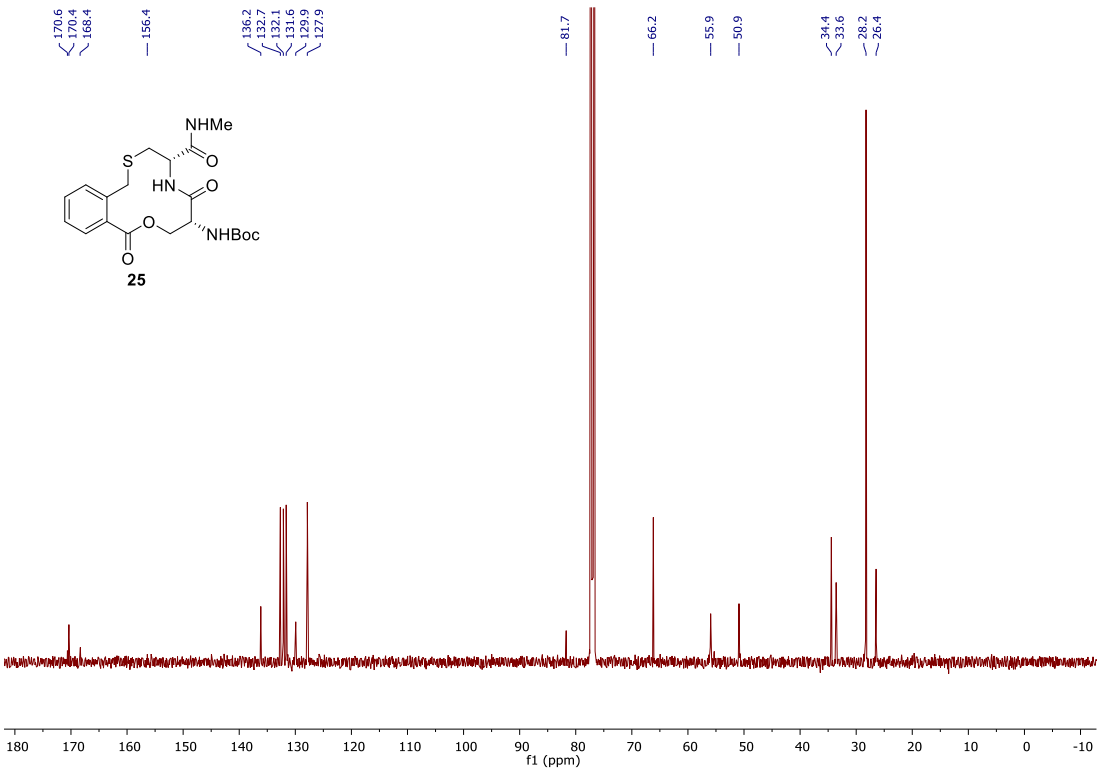

**12** $^{13}$C NMR (101 MHz, DMSO-$d_6$)



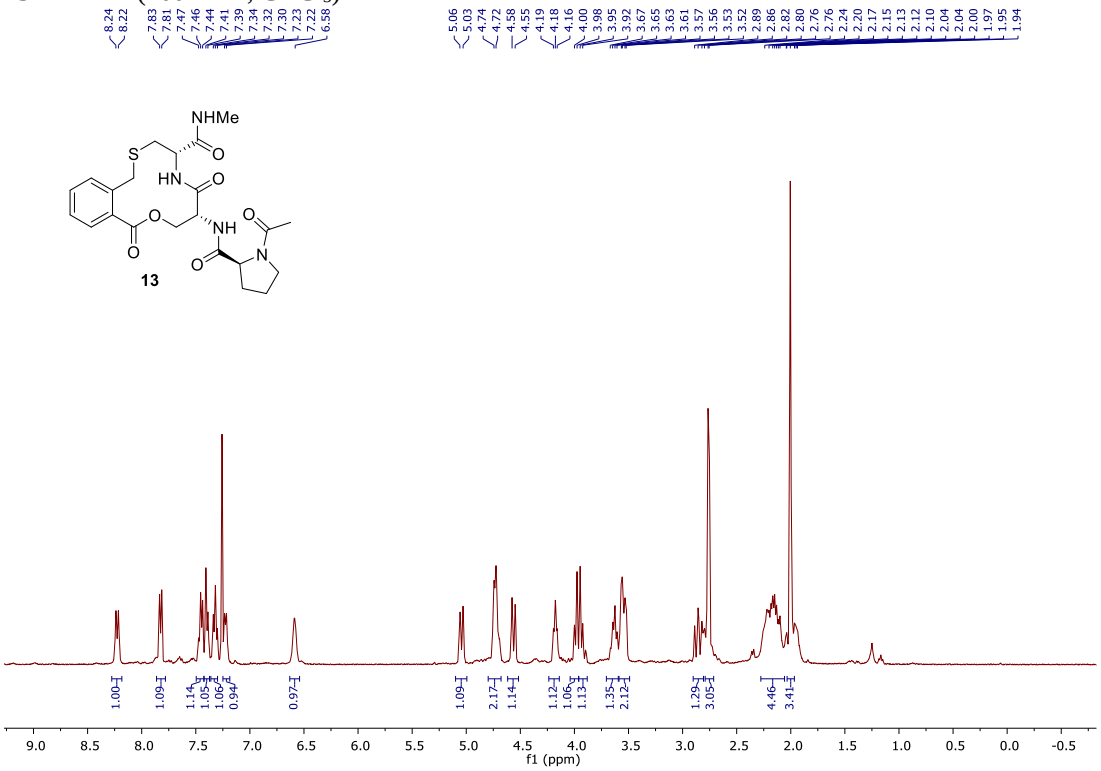
**25** $^{1}$H NMR (400 MHz, $CDCl_3$)



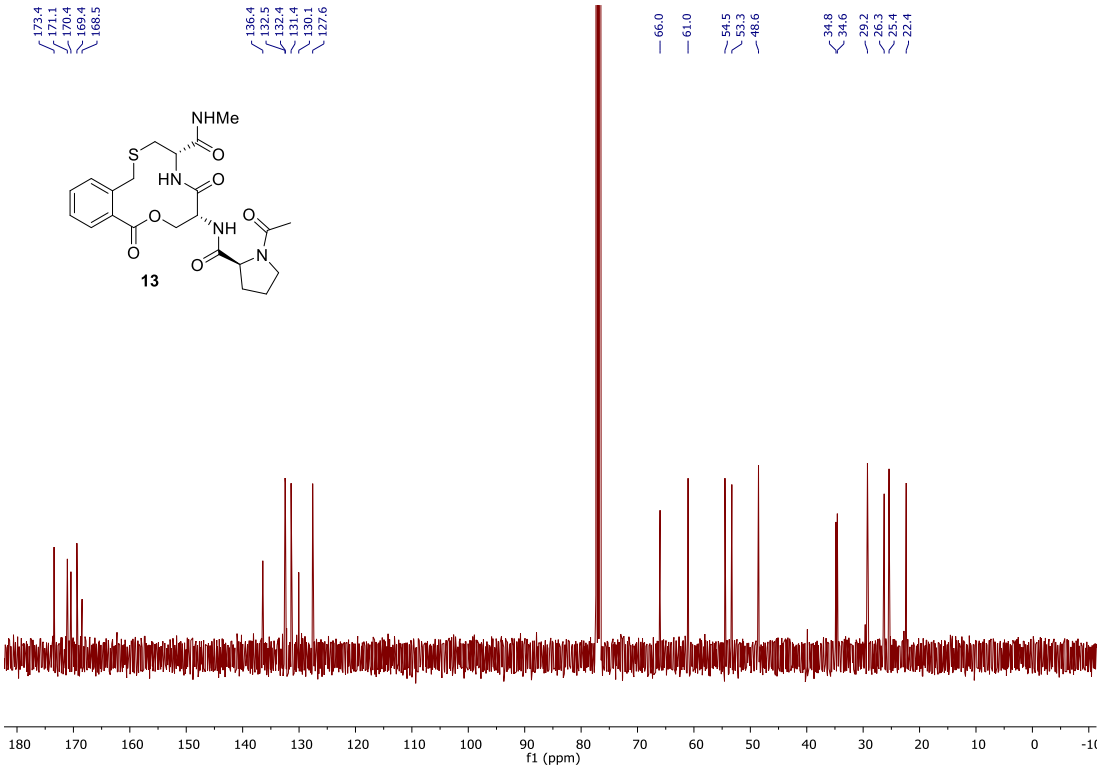

**25** $^{13}$C NMR (126 MHz, $CDCl_3$)



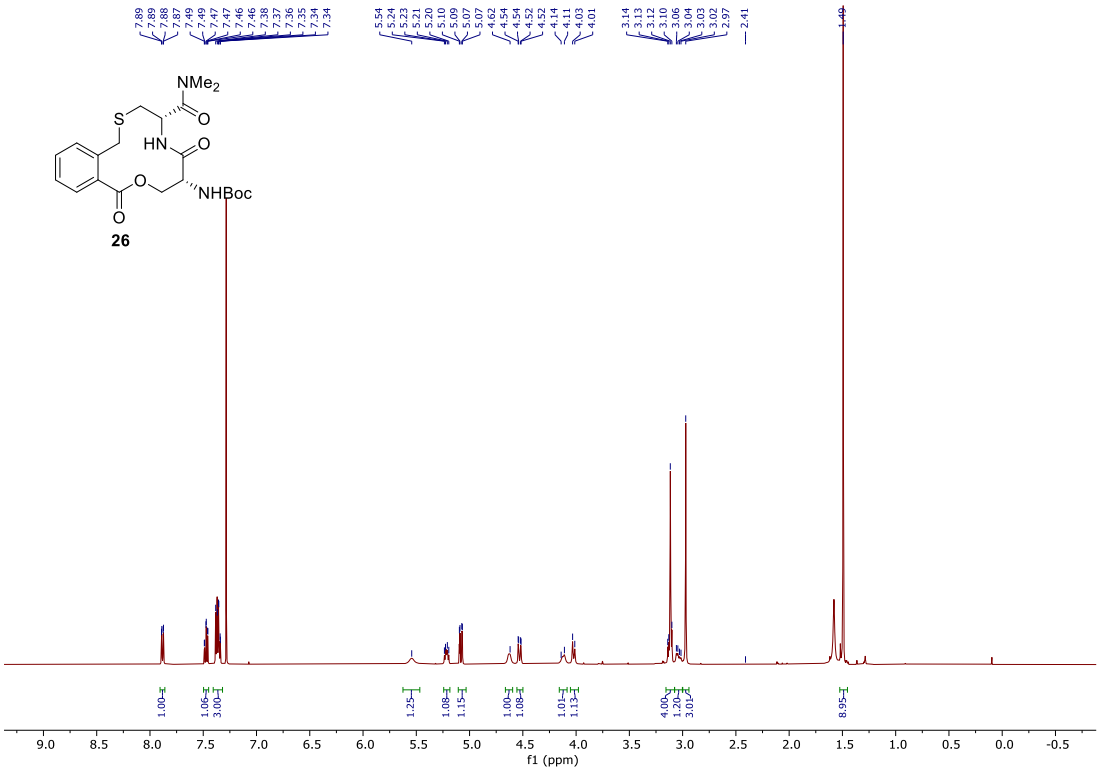
**13** $^{1}$H NMR (400 MHz, $CDCl_3$)



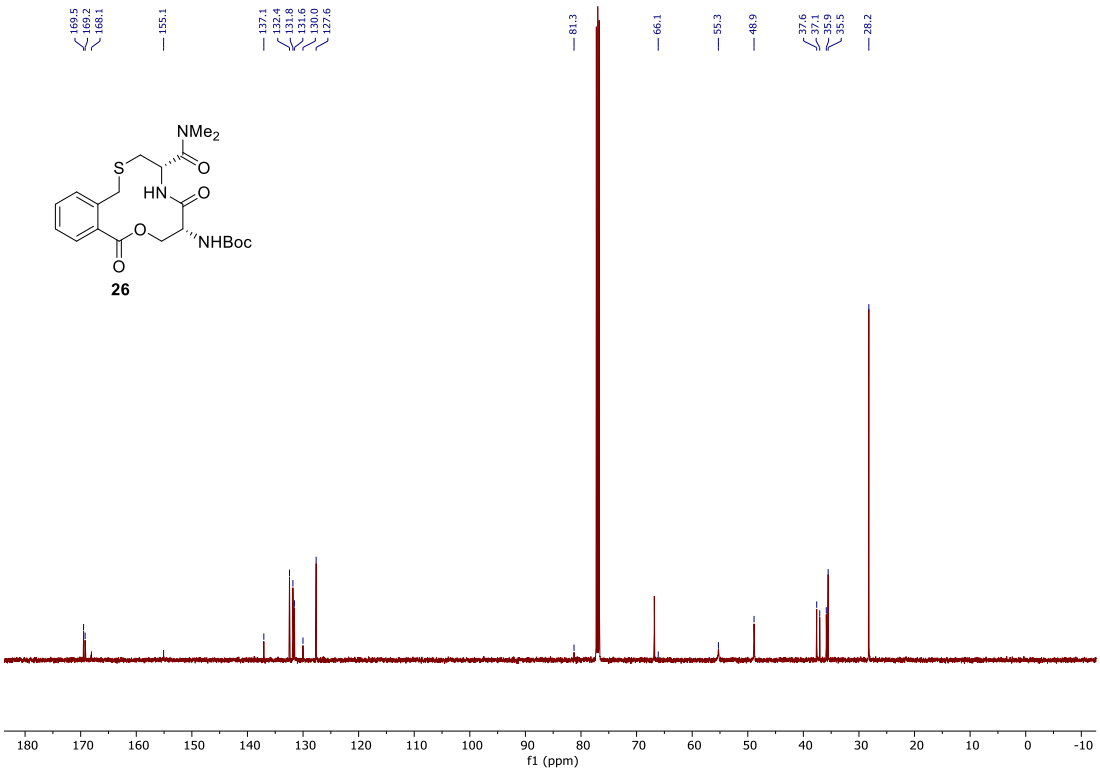

**13** $^{13}$C NMR (101 MHz, $CDCl_3$)



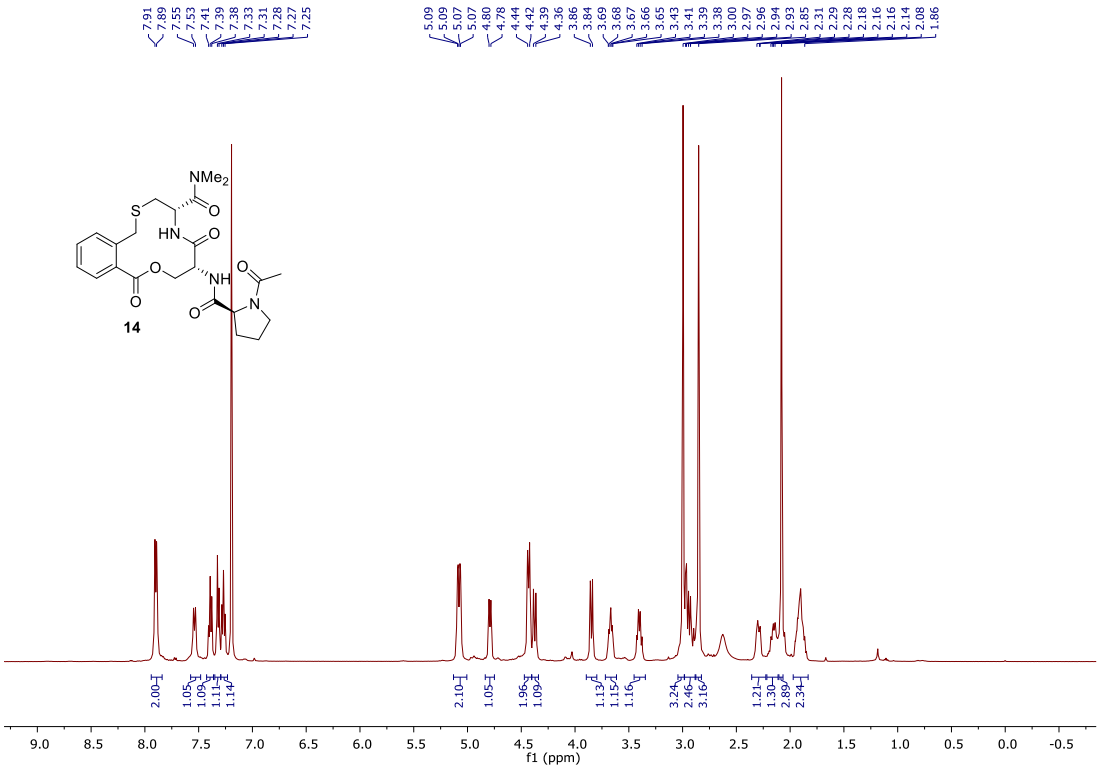
**26** $^{1}$H NMR (500 MHz, $CDCl_3$)



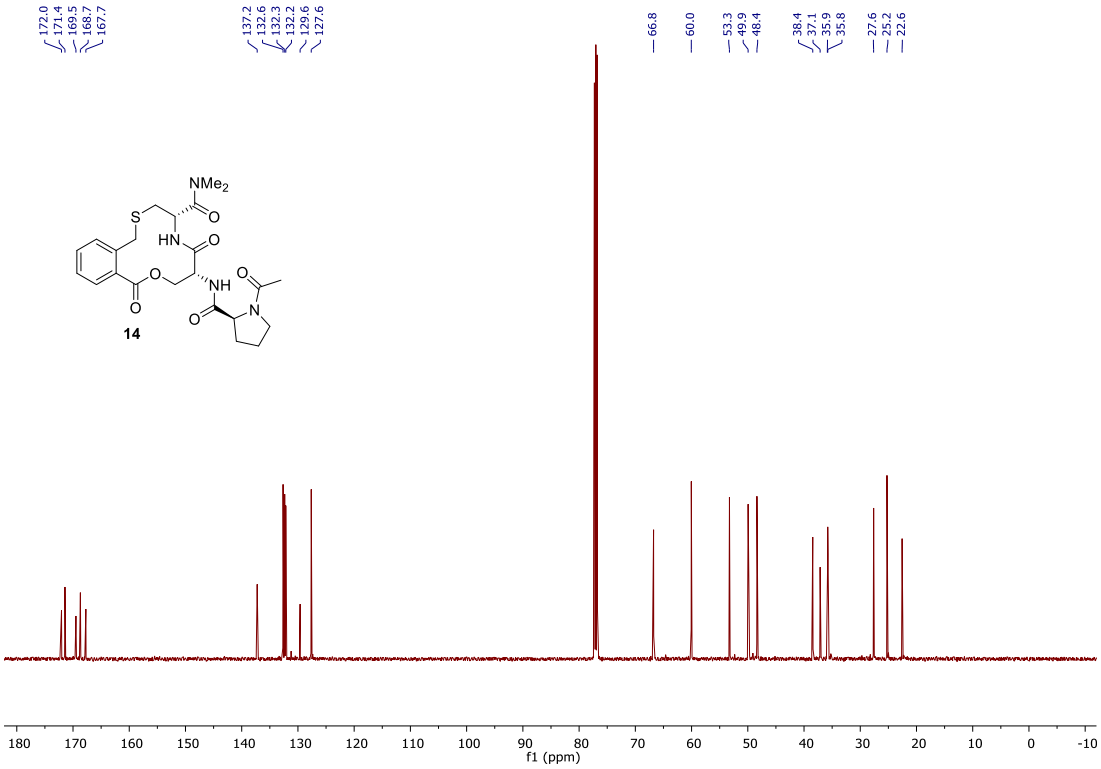

**26** $^{13}$C NMR (126 MHz, $CDCl_3$)



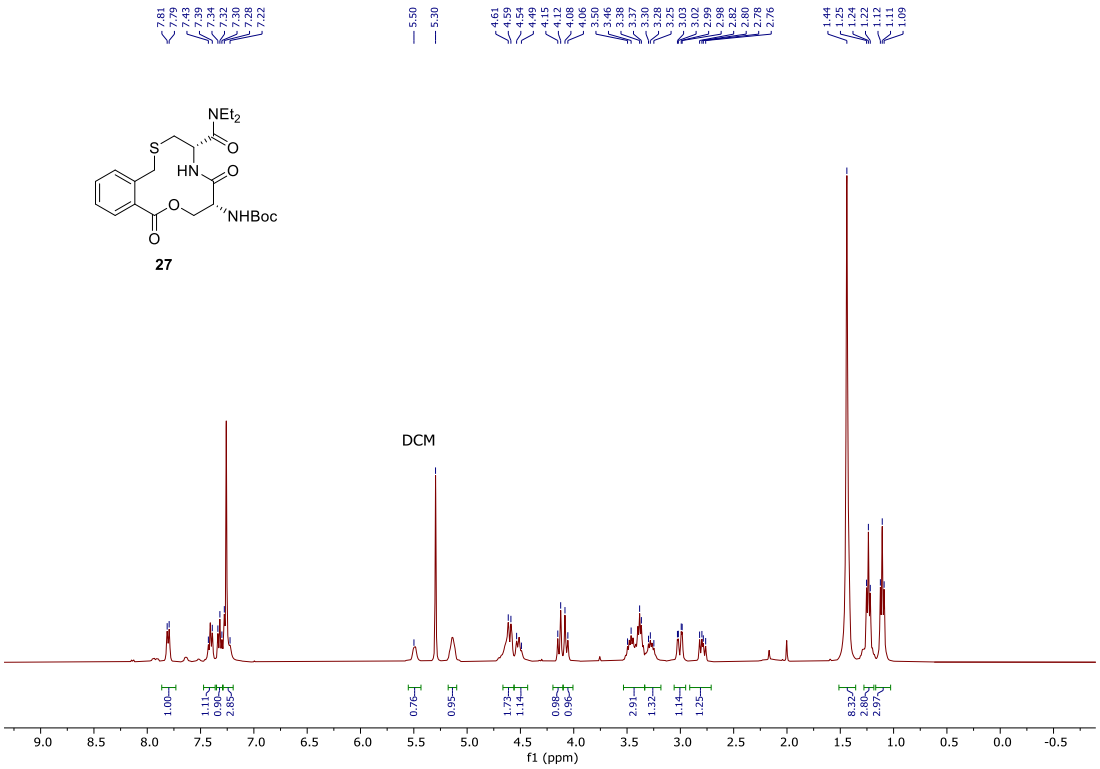
**14** $^{1}$H NMR (500 MHz, $CDCl_3$)



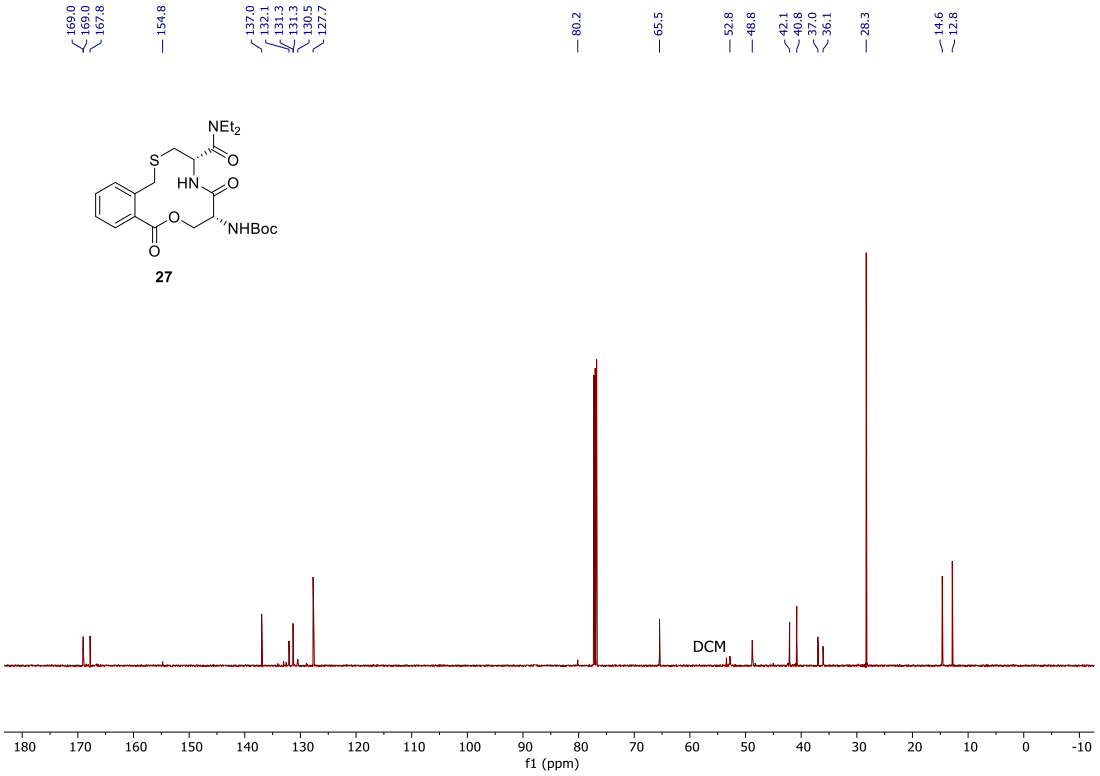

**14** $^{13}$C NMR (126 MHz, $CDCl_3$)



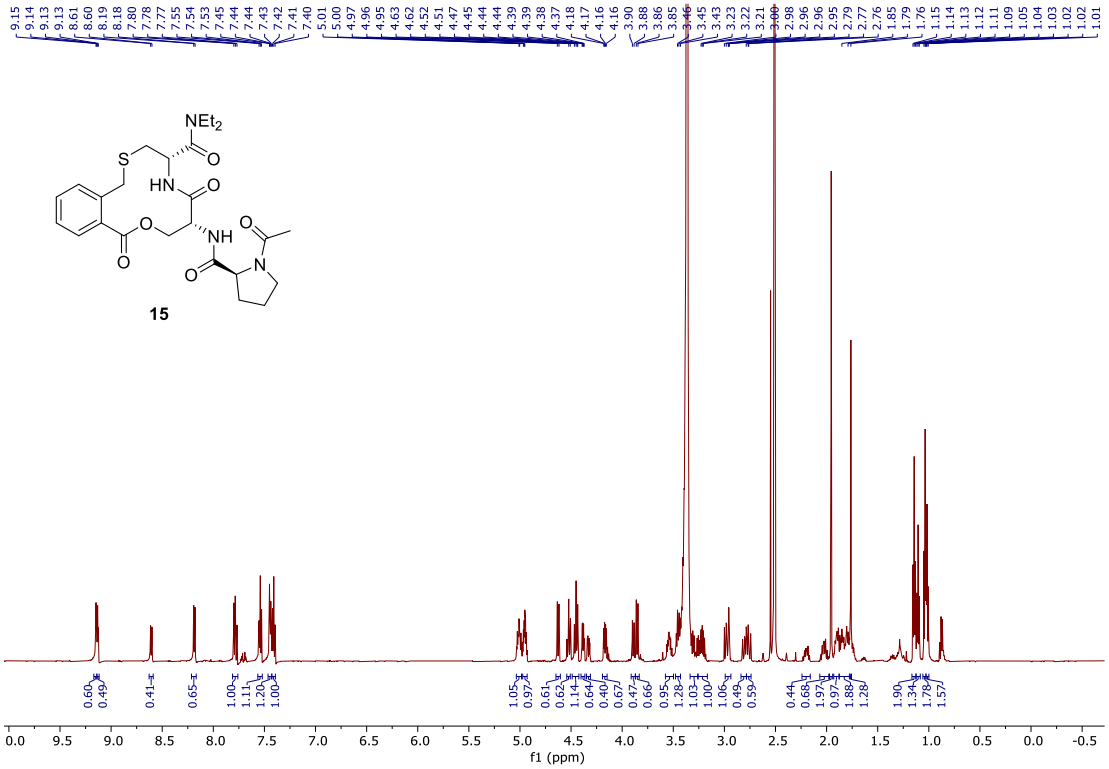
**27** $^{1}$H NMR (400 MHz, $CDCl_3$)



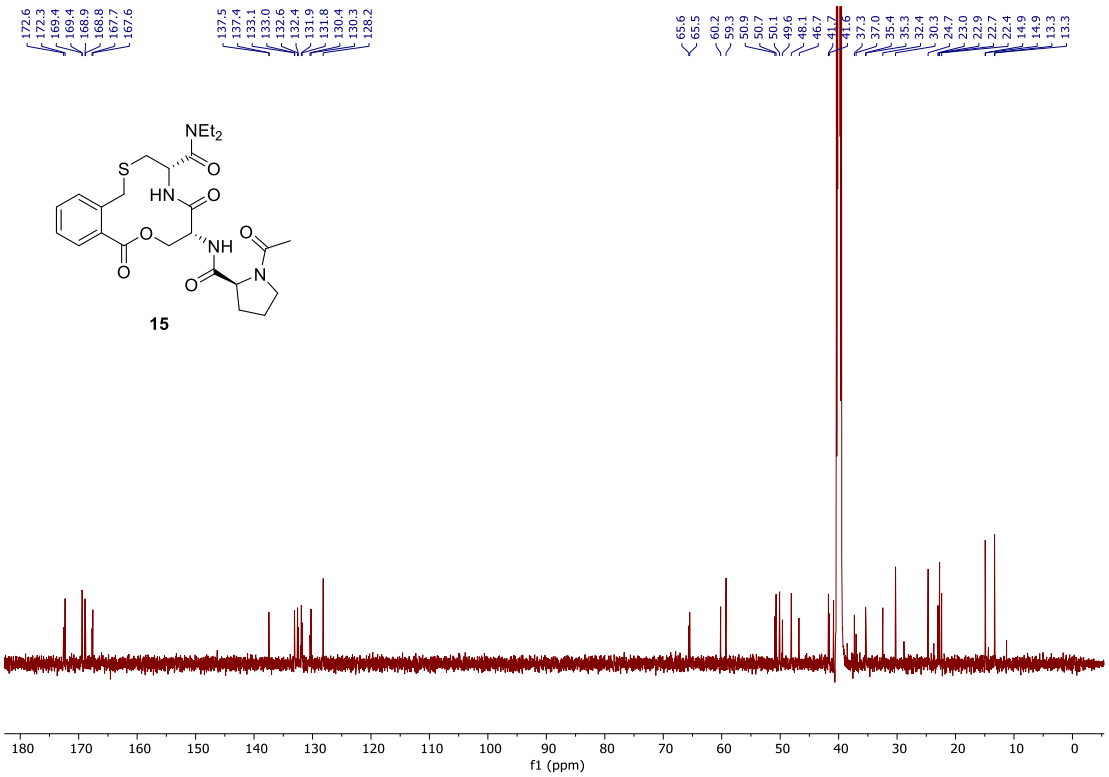

**27** $^{13}$C NMR (126 MHz, $CDCl_3$)



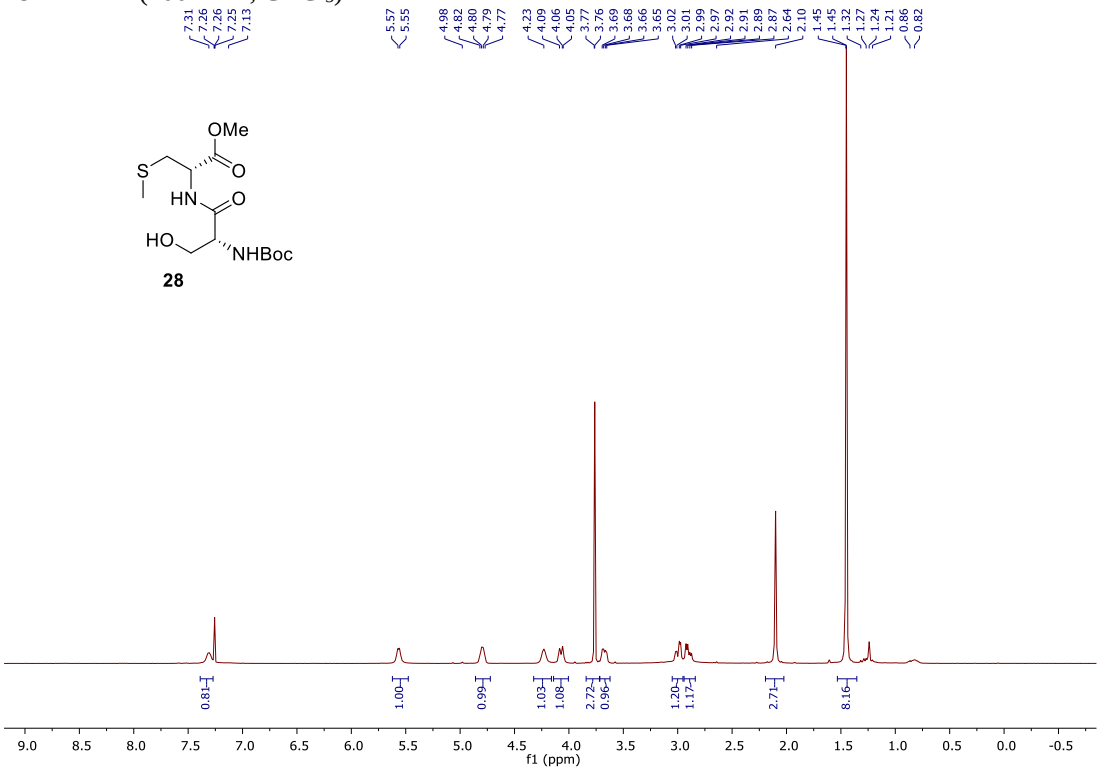
**15** $^{1}$H NMR (600 MHz, DMSO-$d_6$, mixture of 2 rotamers in ratio 6:4)



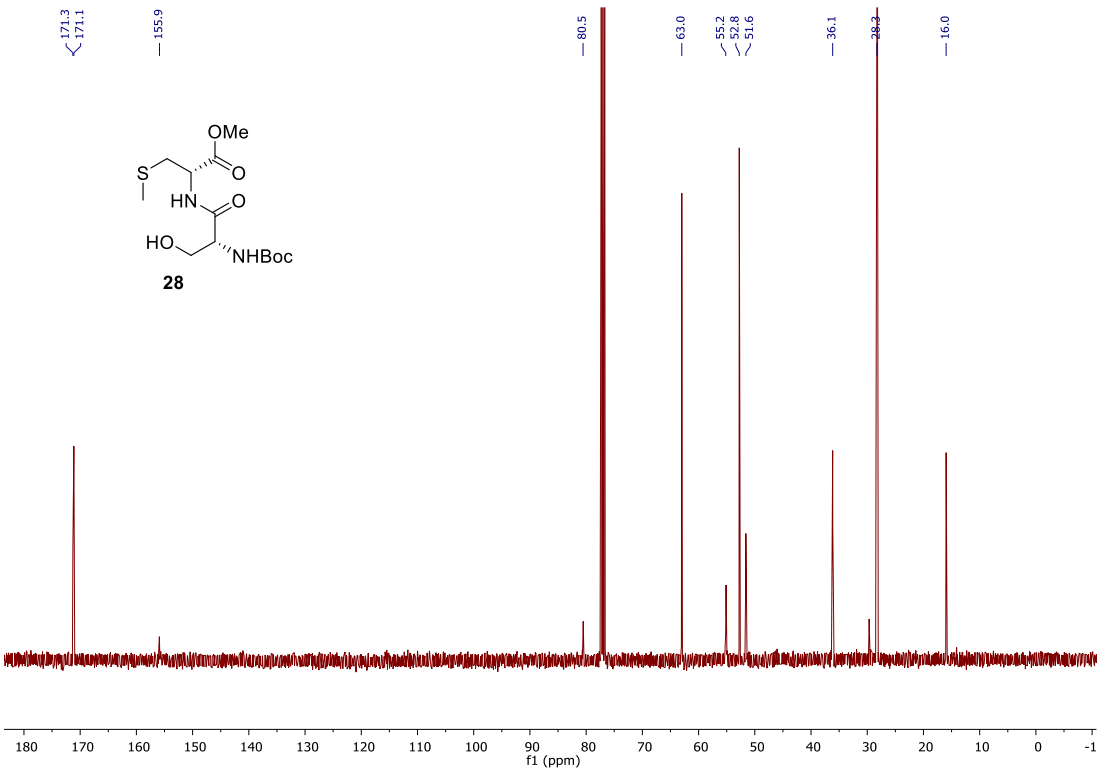

**15** $^{13}$C NMR (151 MHz, DMSO-$d_6$, mixture of 2 rotamers)



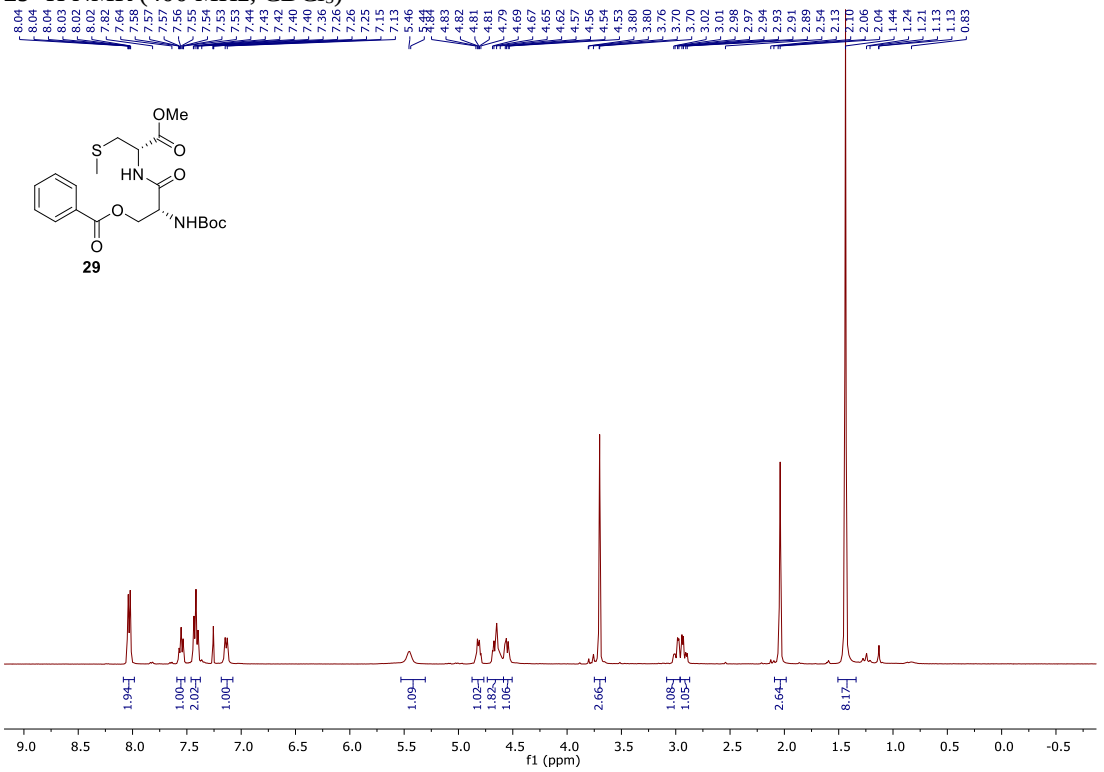
**28** $^{1}$H NMR (400 MHz, $CDCl_3$)



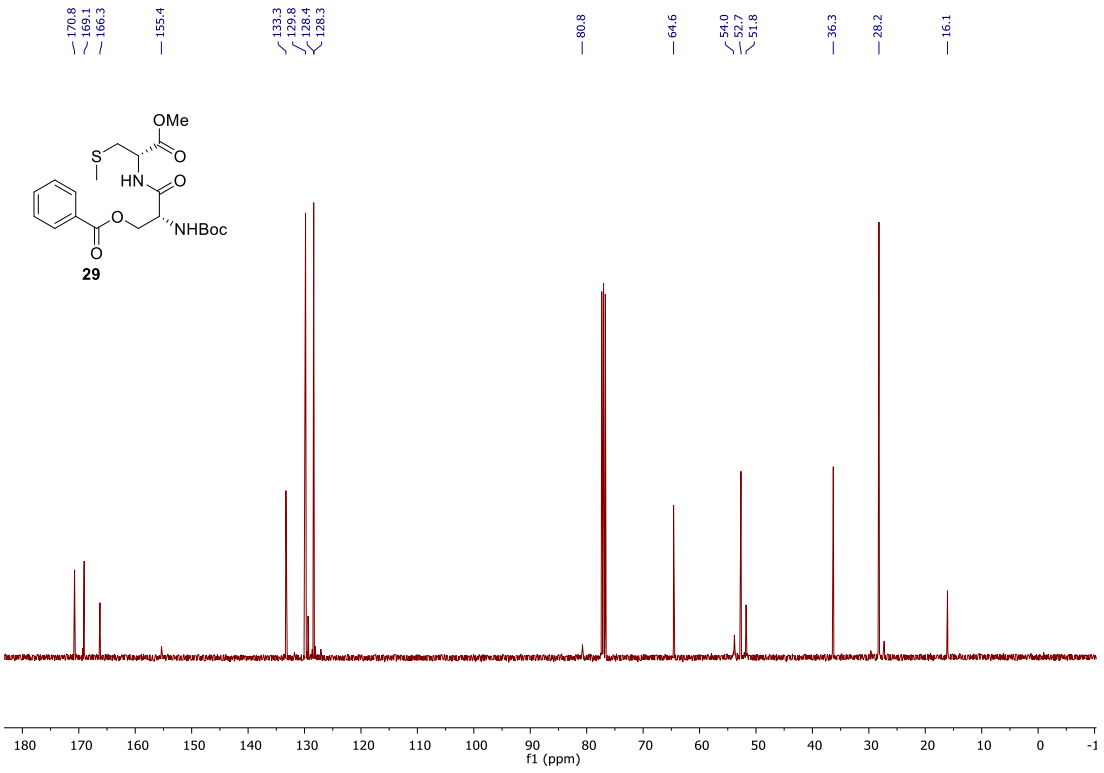

**28** $^{13}$C NMR (101 MHz, $CDCl_3$)



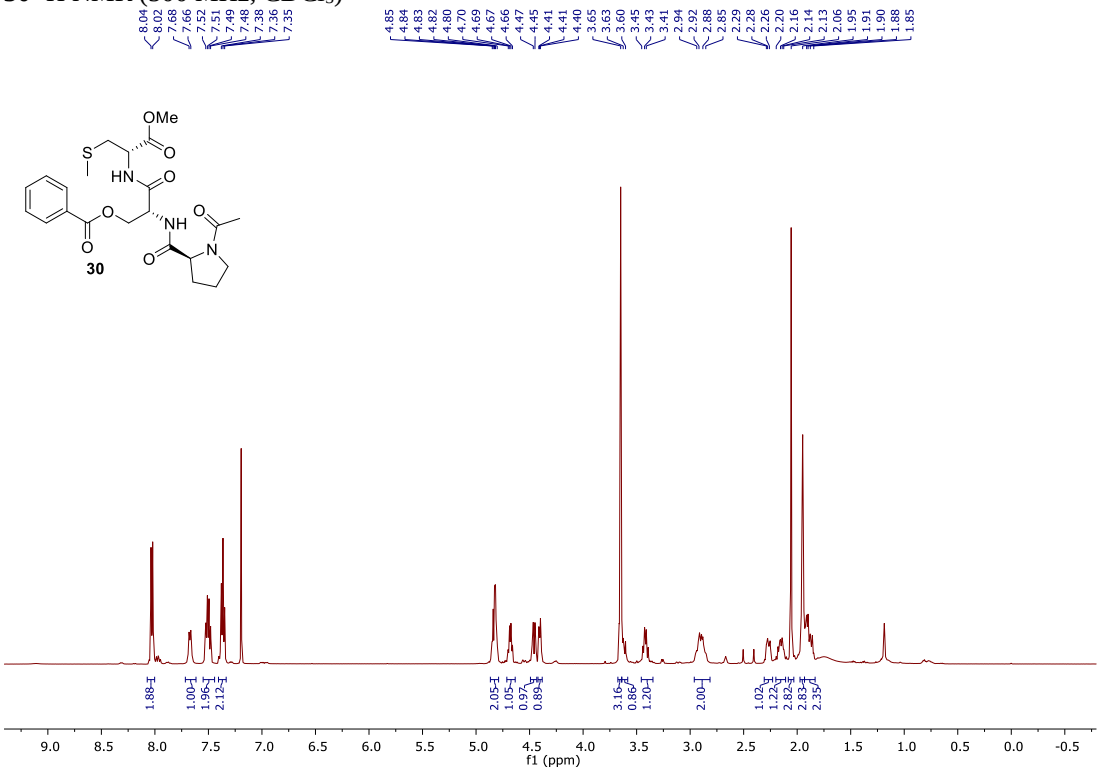
**29** $^{1}$H NMR (400 MHz, $CDCl_3$)



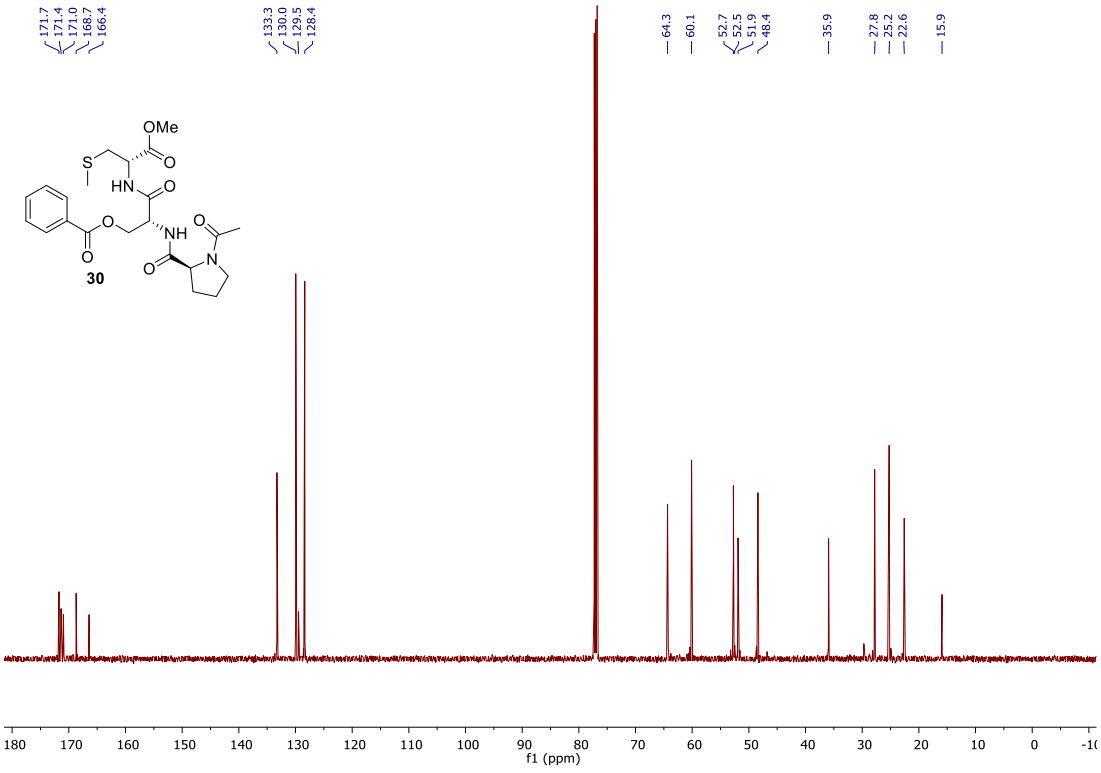

**29** $^{13}$C NMR (101 MHz, $CDCl_3$)



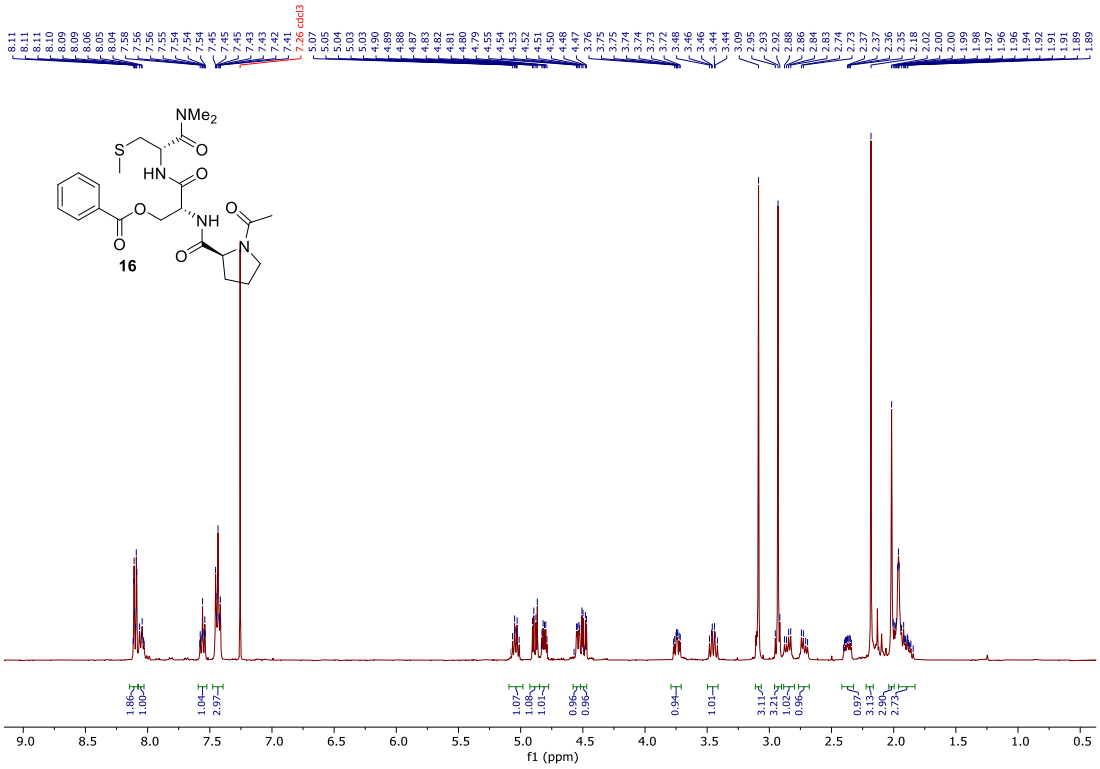
**30** $^{1}$H NMR (500 MHz, $CDCl_3$)



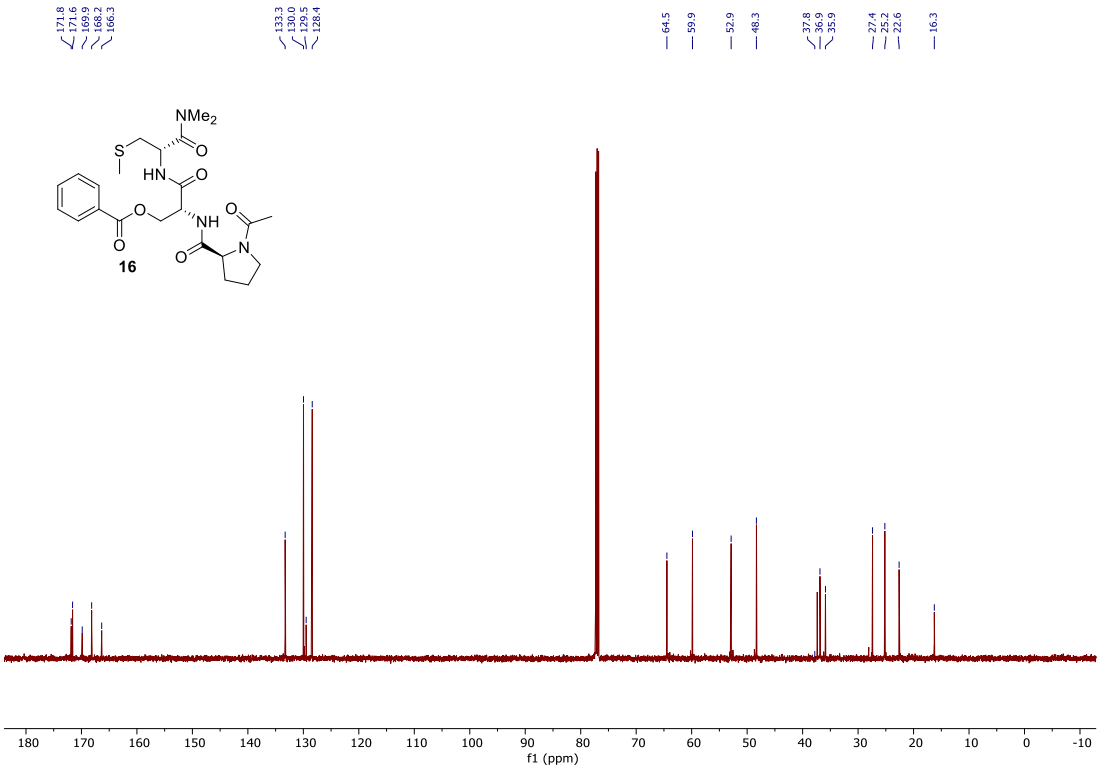

**30** $^{13}$C NMR (126 MHz, $CDCl_3$)



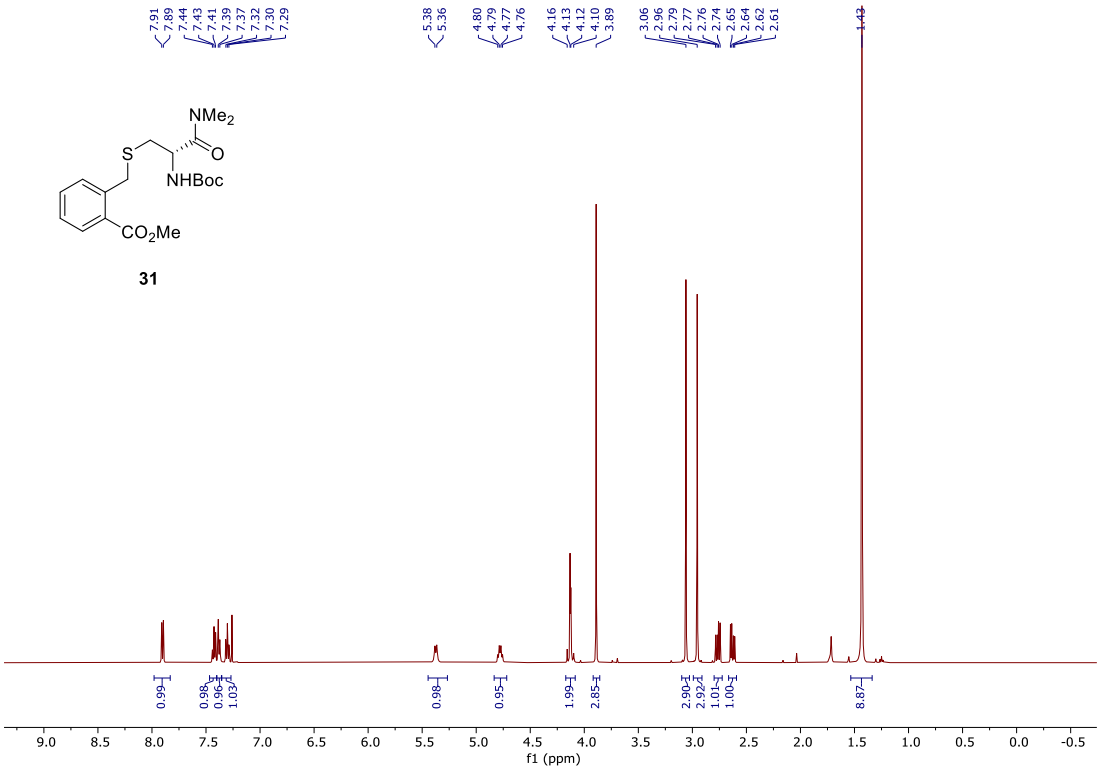
**16** $^{1}$H NMR (400 MHz, $CDCl_3$)



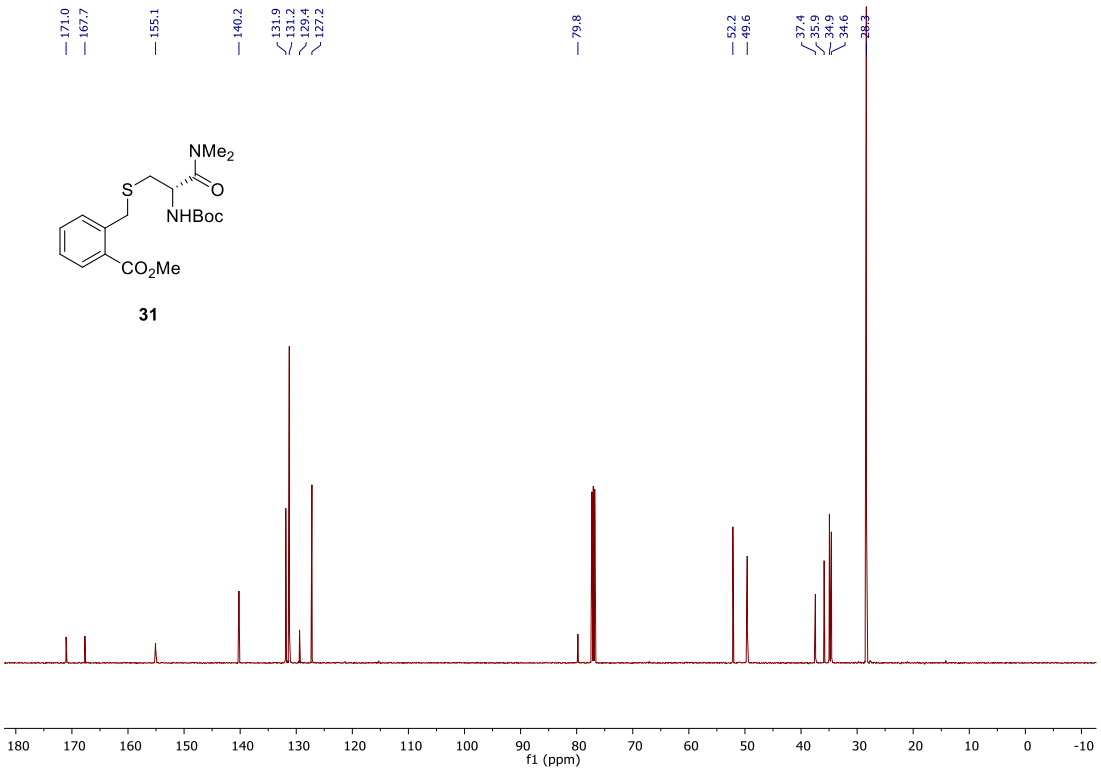

**16** $^{13}$C NMR (126 MHz, $CDCl_3$))



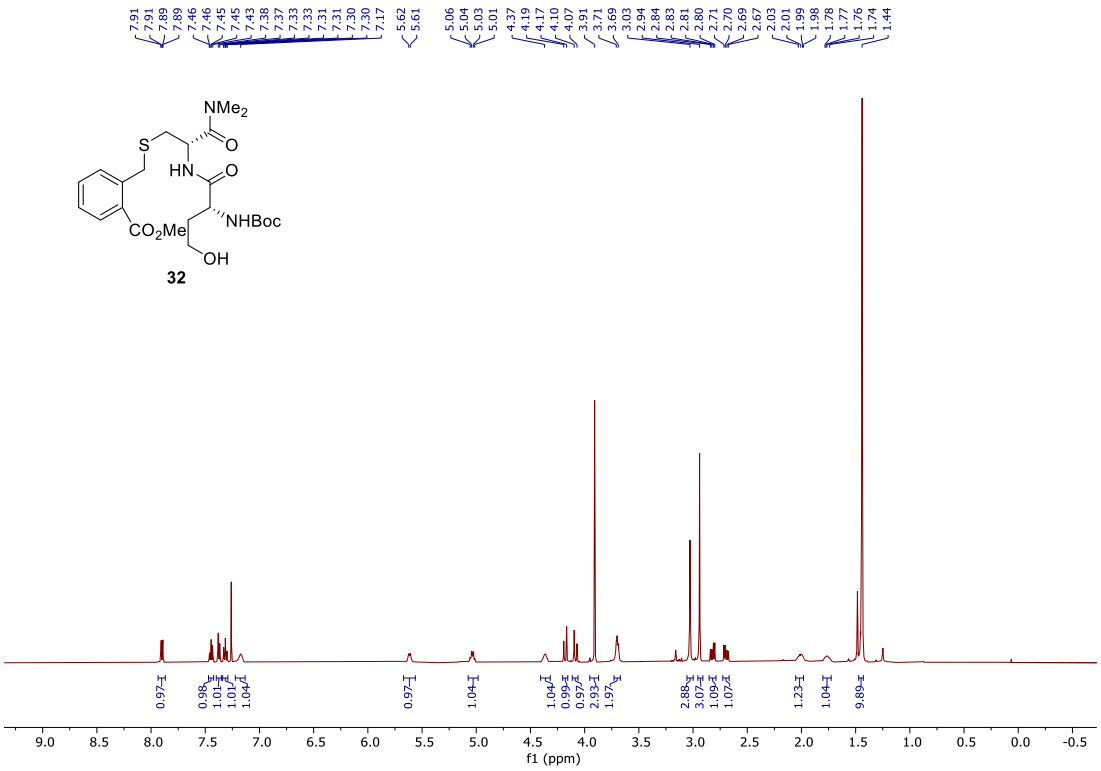
**31** $^{1}$H NMR (500 MHz, $CDCl_3$)



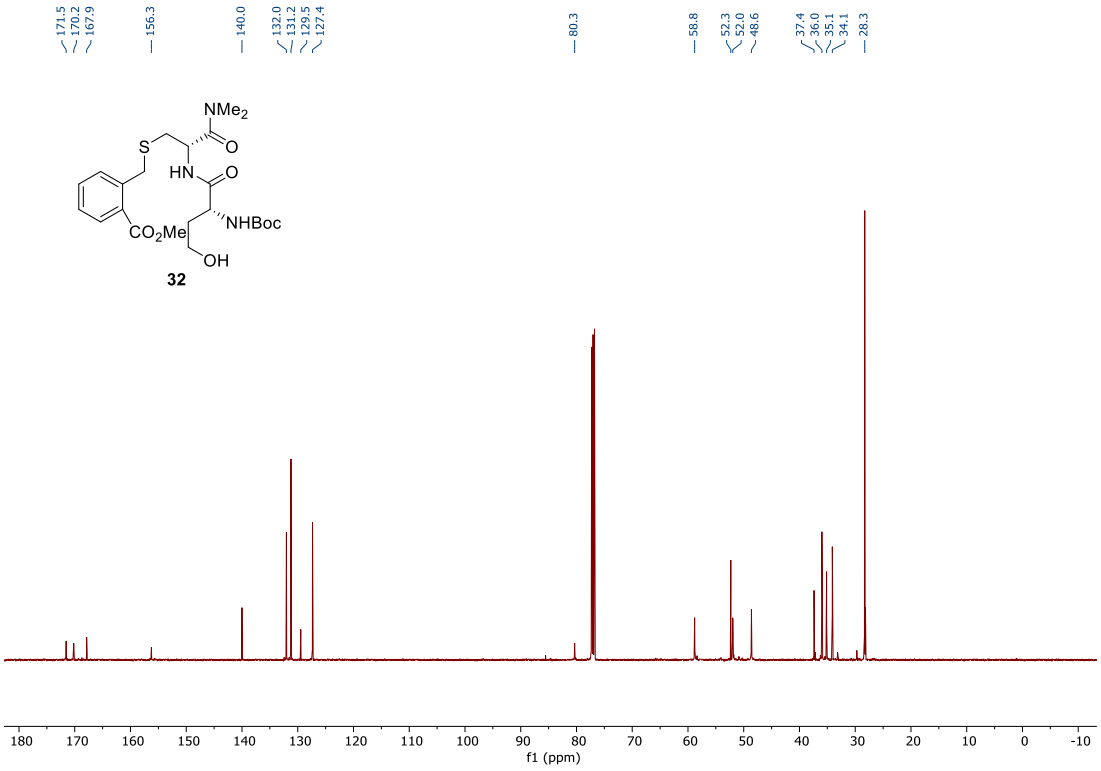

**31** $^{13}$C NMR (126 MHz, $CDCl_3$)



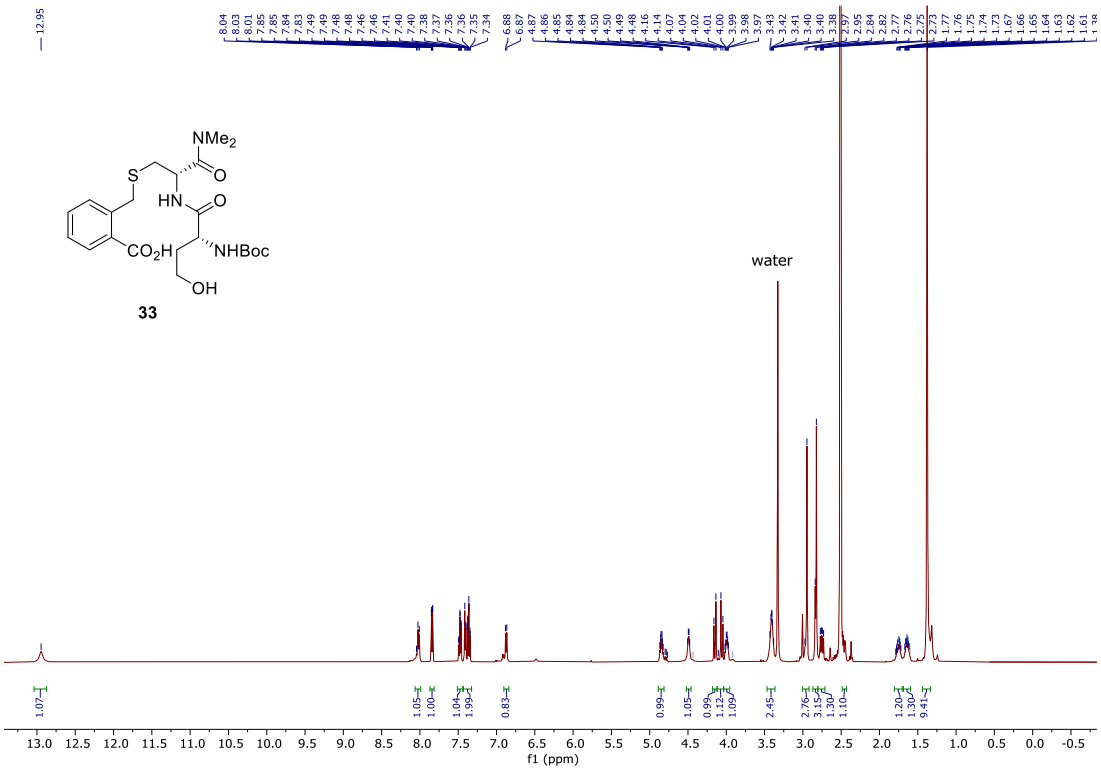
**32** $^{1}$H NMR (500 MHz, $CDCl_3$)



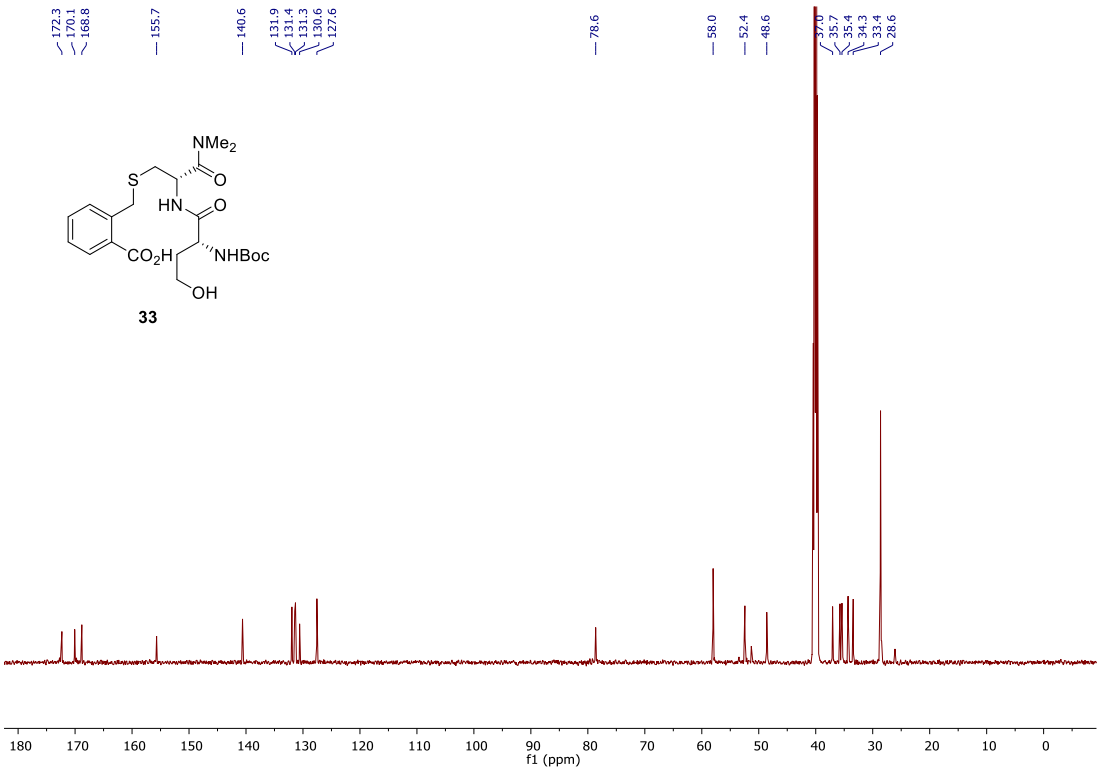

**32** $^{13}$C NMR (126 MHz, $CDCl_3$)



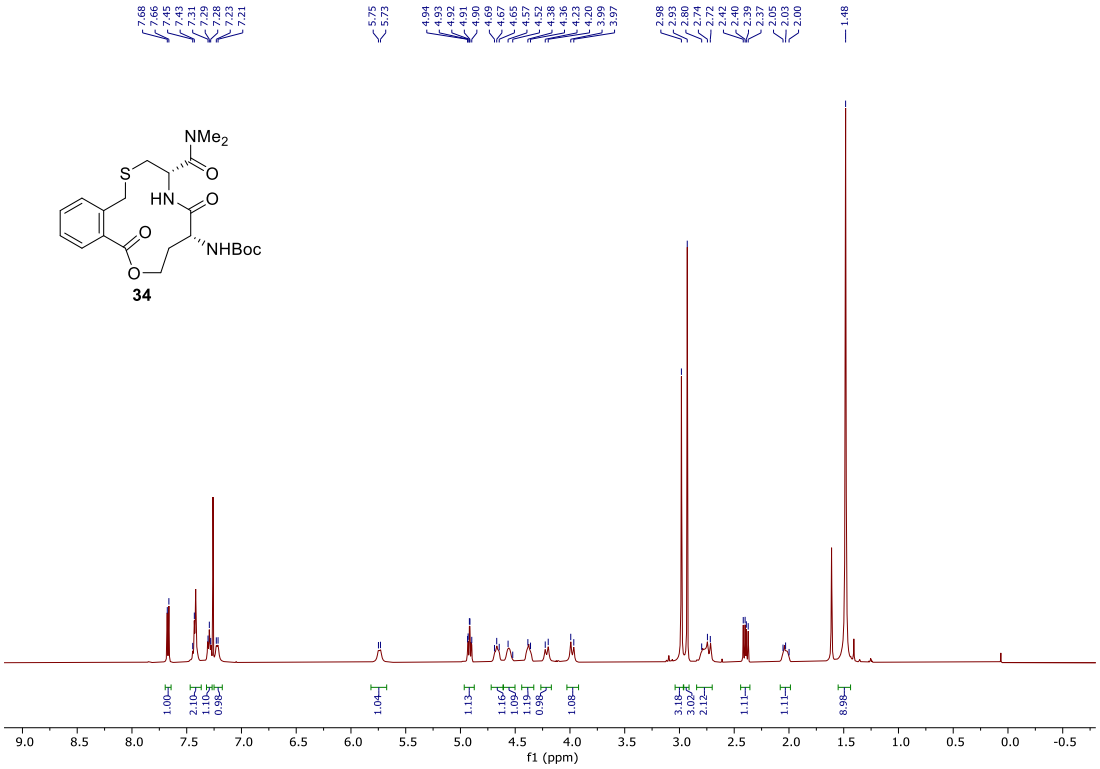
**33** $^{1}$H NMR (600 MHz, DMSO-$d_6$)



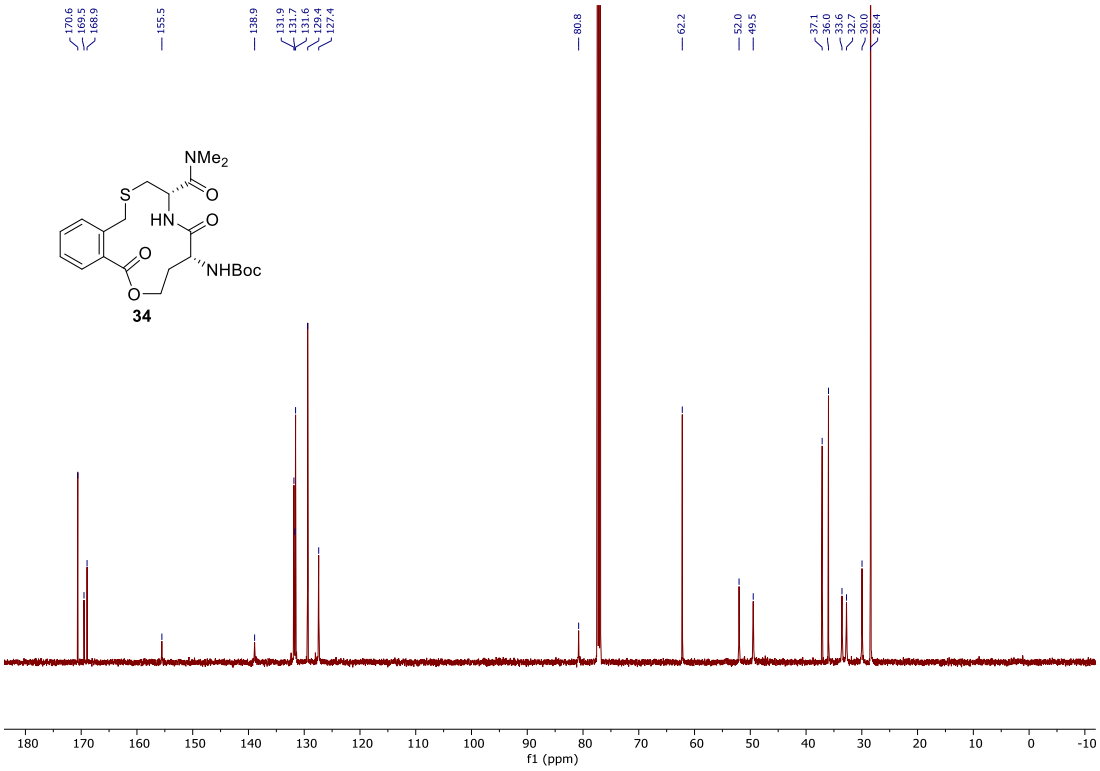

**33** $^{13}$C NMR (151 MHz, DMSO-$d_6$)



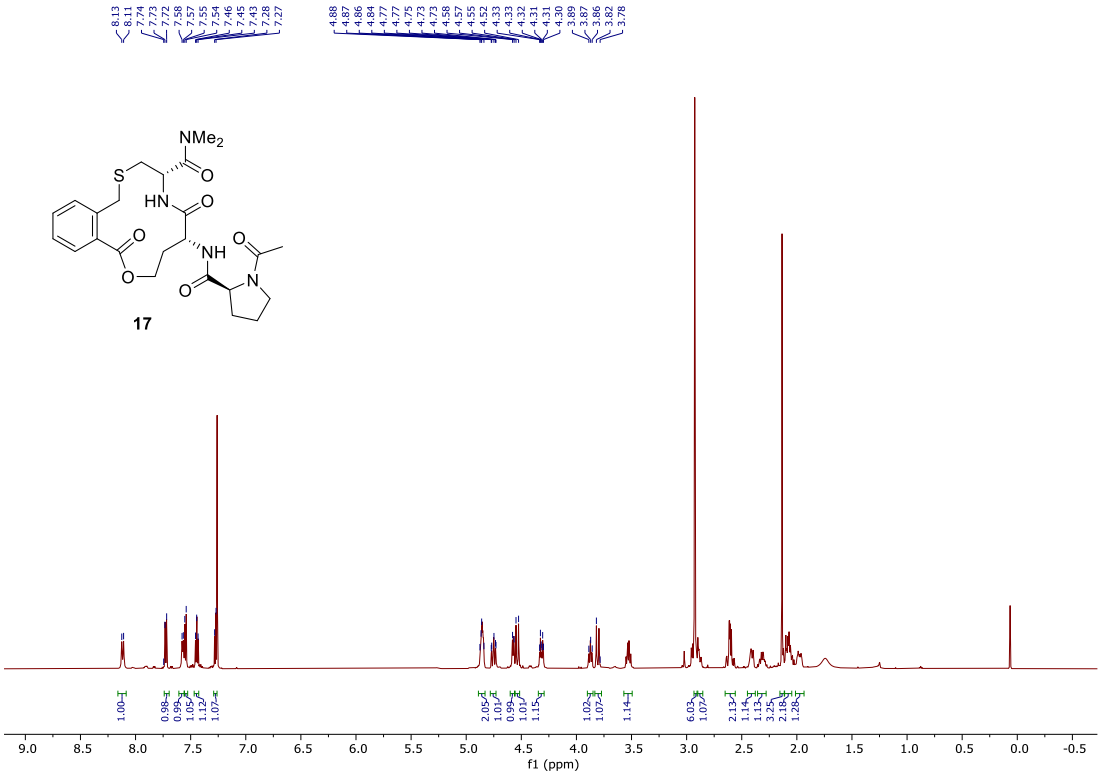
**34** $^{1}$H NMR (500 MHz, $CDCl_3$)



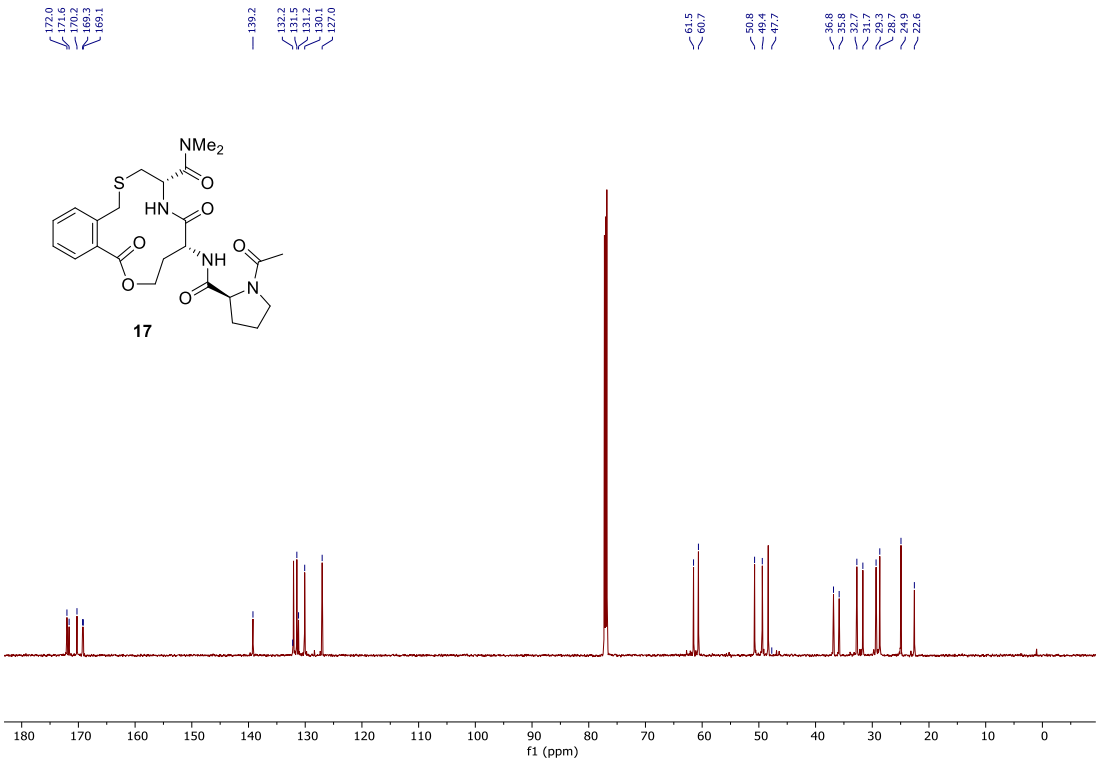

**34** $^{13}$C NMR (126 MHz, $CDCl_3$)



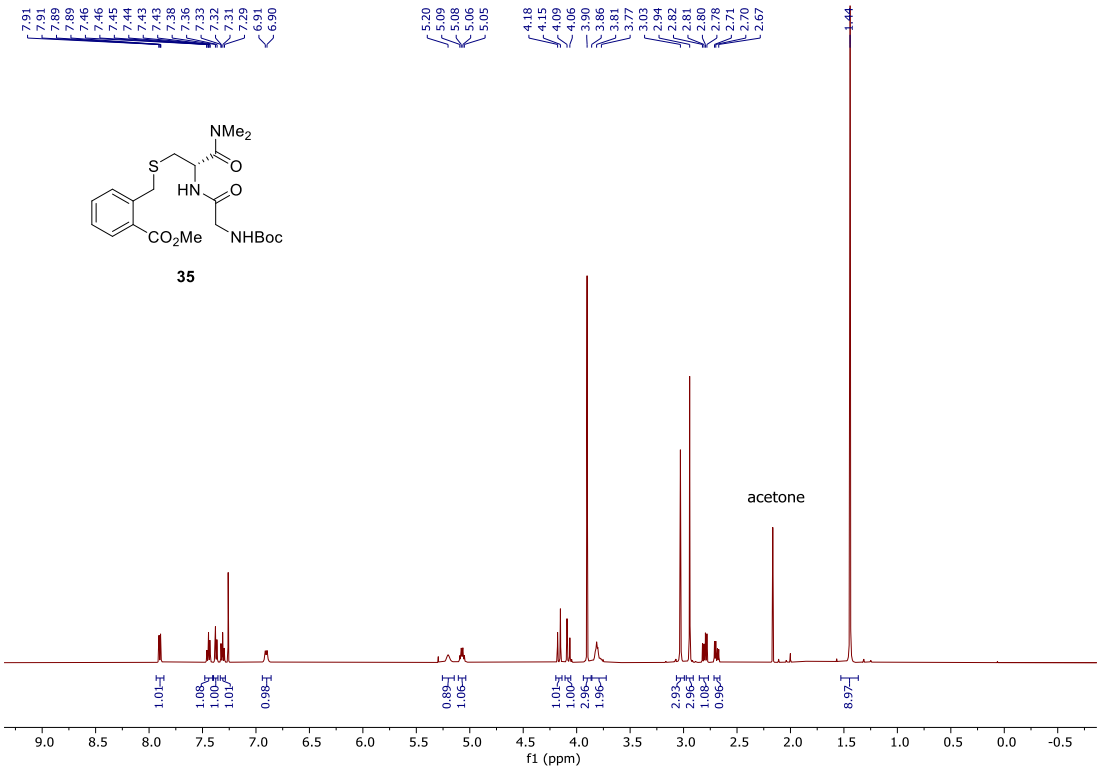
**17** $^{1}$H NMR (600 MHz, $CDCl_3$)



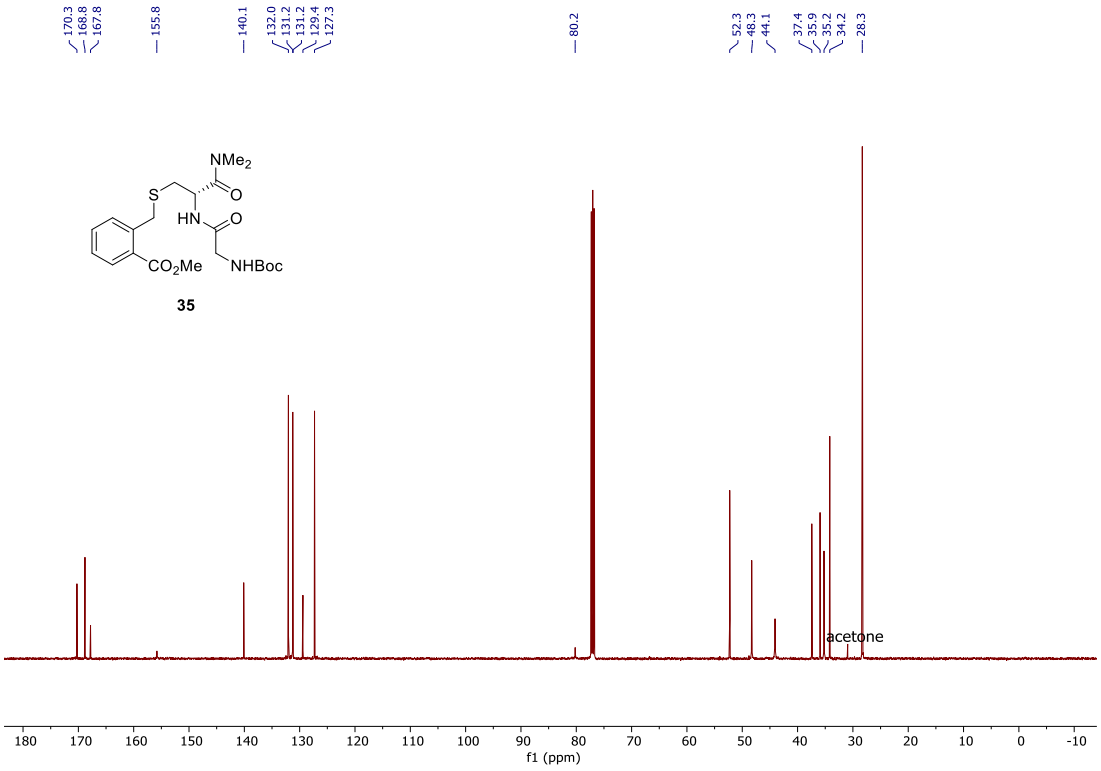

**17** $^{13}$C NMR (151 MHz, $CDCl_3$)



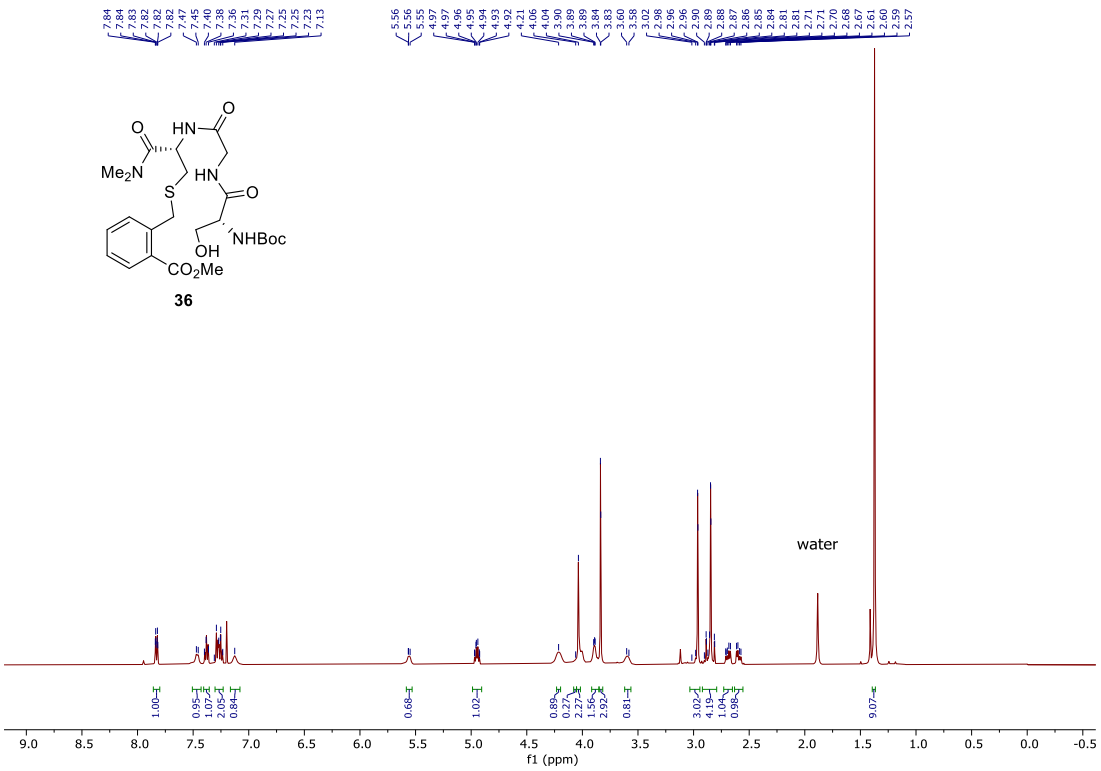
**35** $^{1}$H NMR (400 MHz, $CDCl_3$)




**35** $^{13}$C NMR (126 MHz, $CDCl_3$)

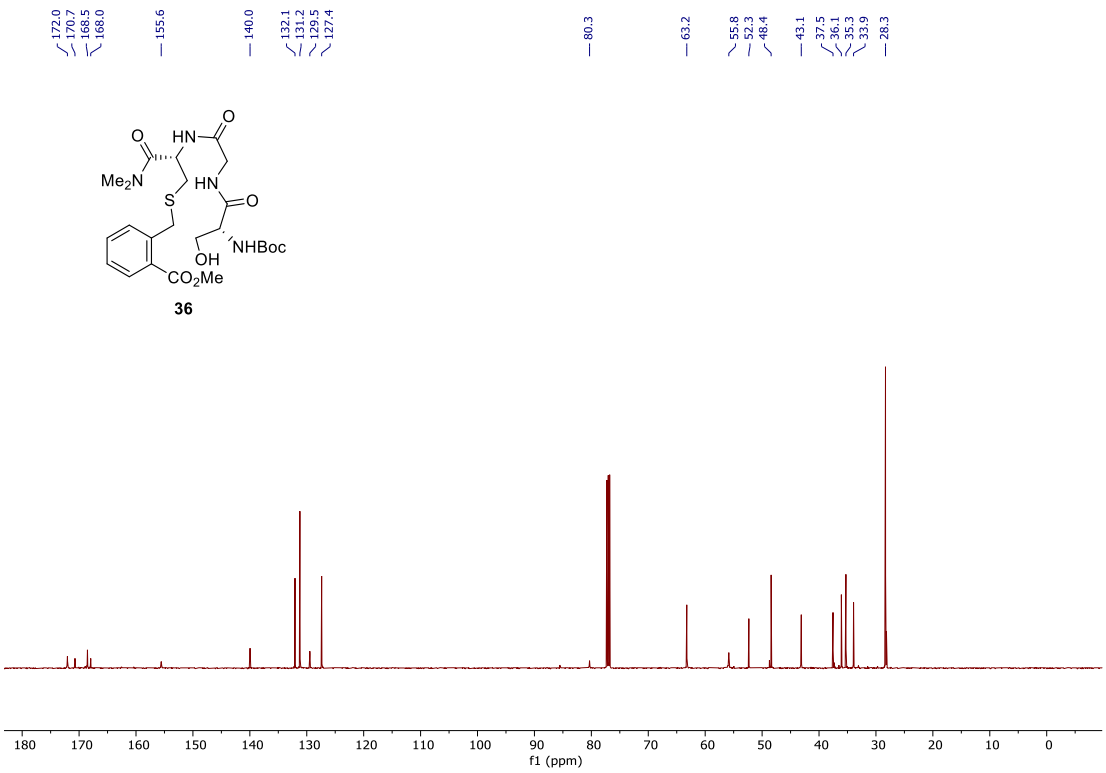


**36** $^{1}$H NMR (500 MHz, $CDCl_3$)



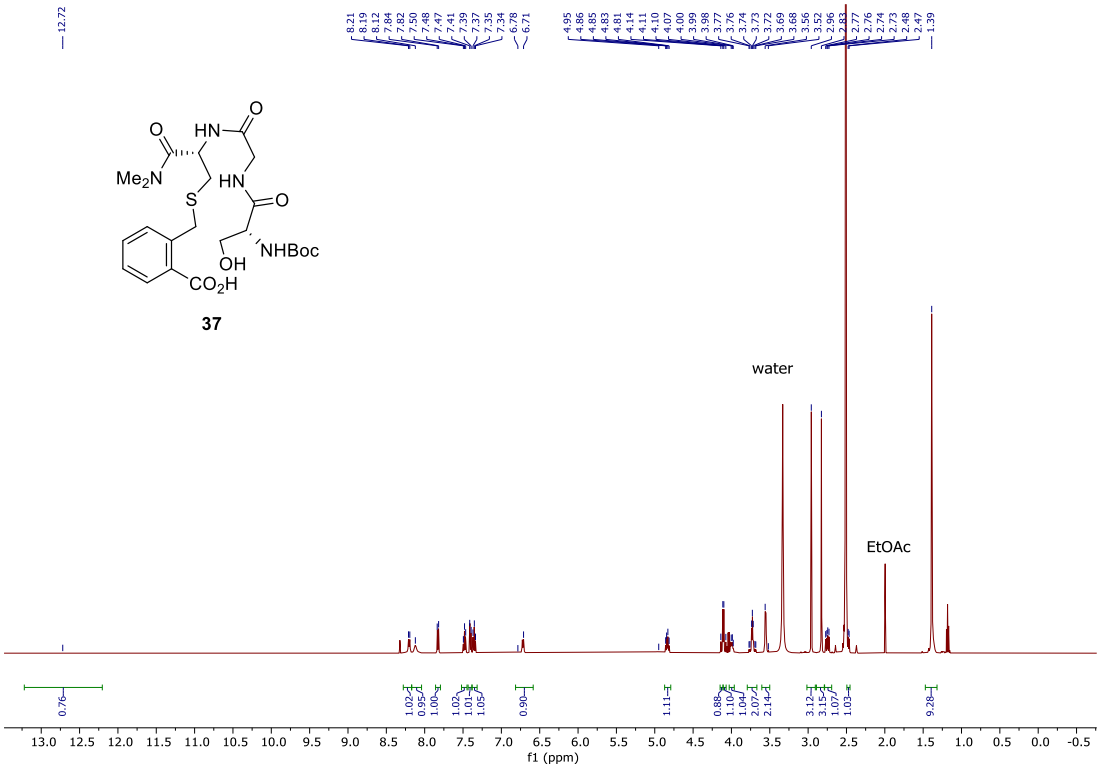

**36**

**36** $^{13}$C NMR (126 MHz, $CDCl_3$)

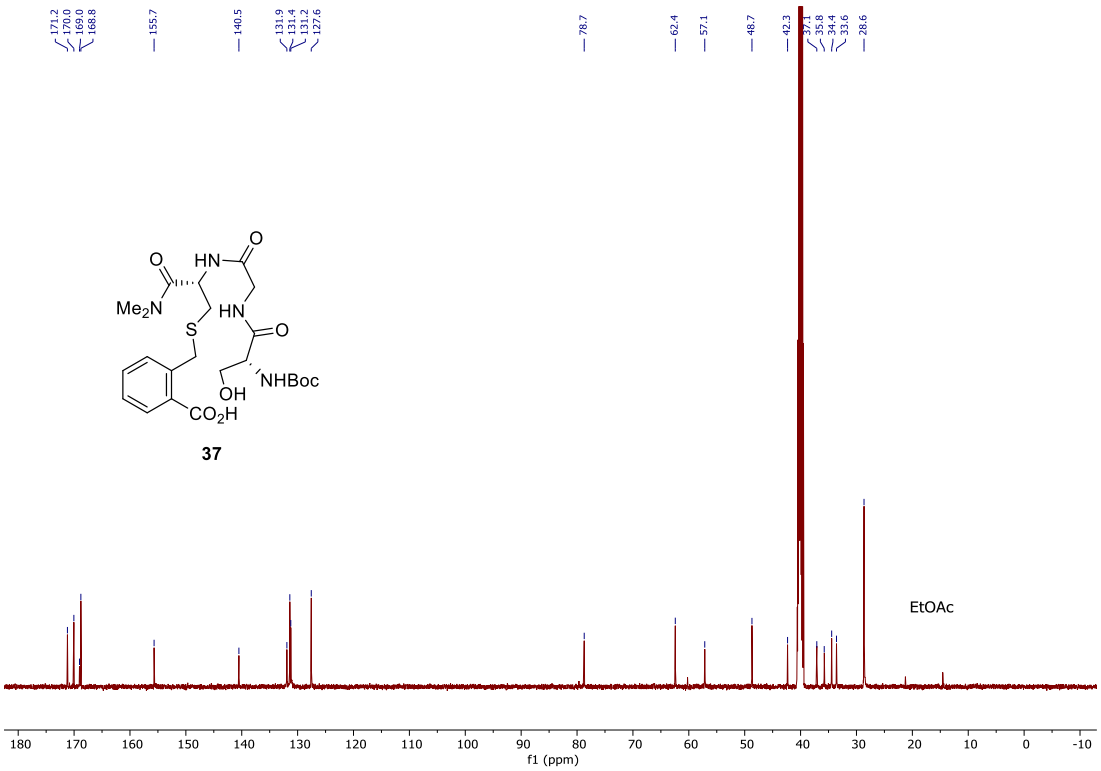



**36**

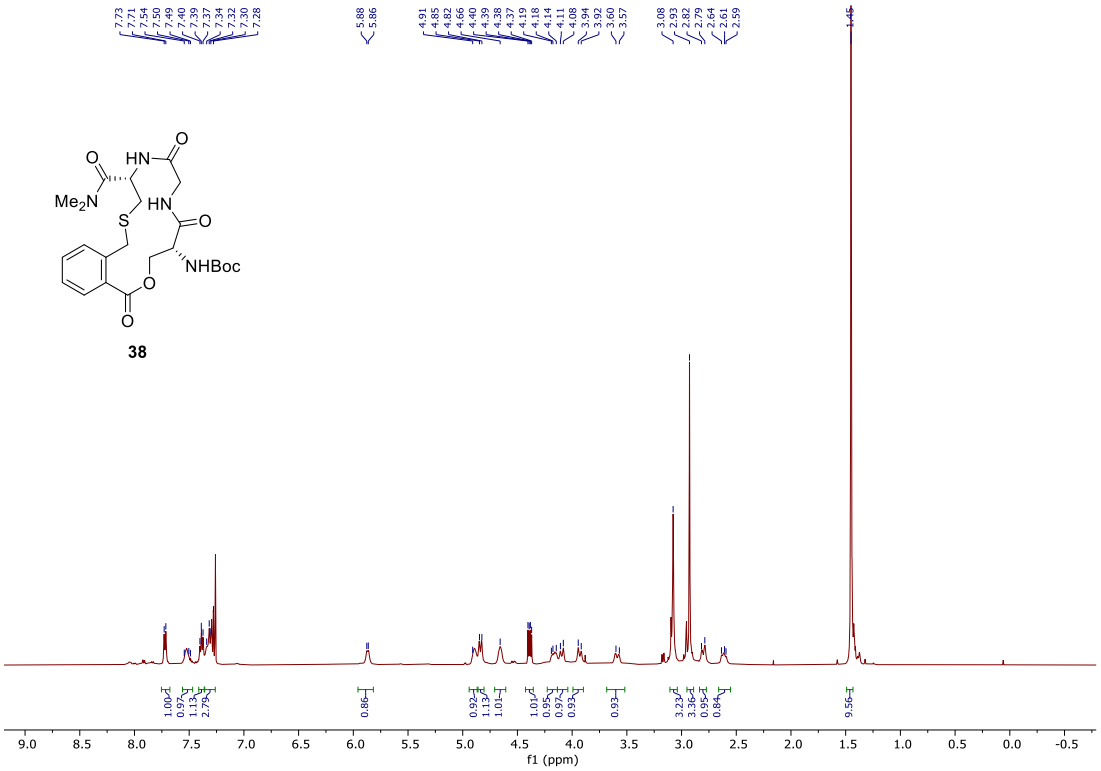
**37** $^{1}$H NMR (500 MHz, DMSO-$d_6$)



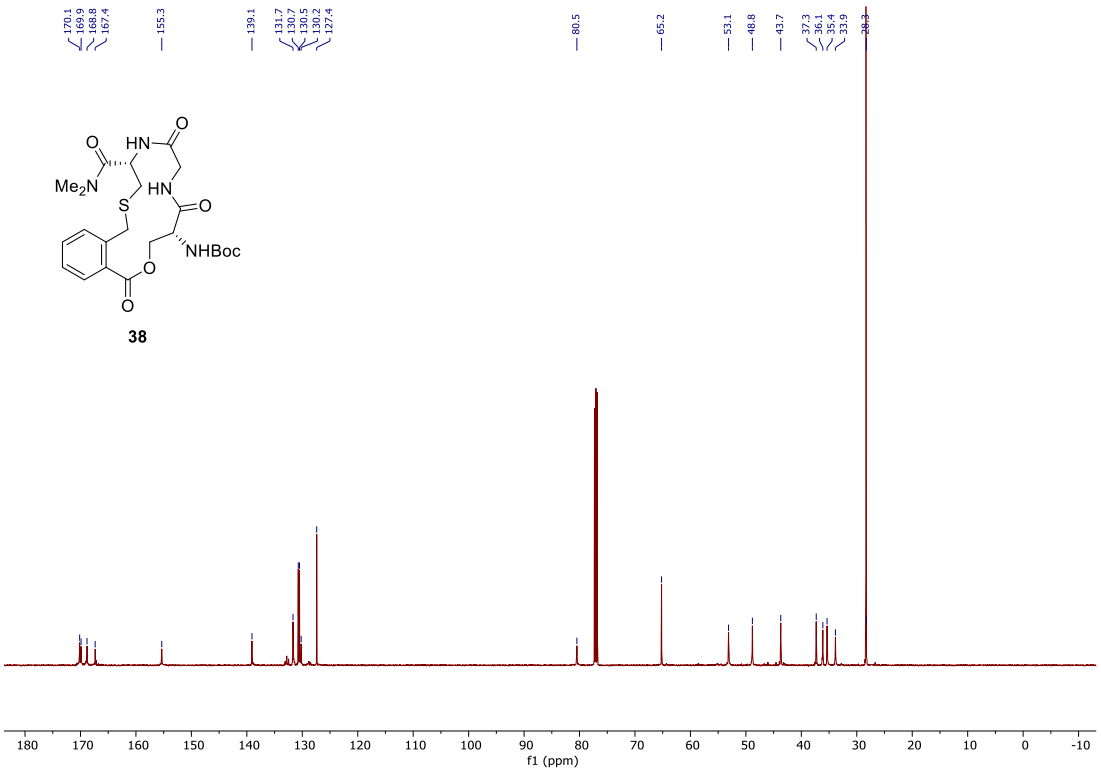

**37** $^{13}$C NMR (126 MHz, DMSO-$d_6$)



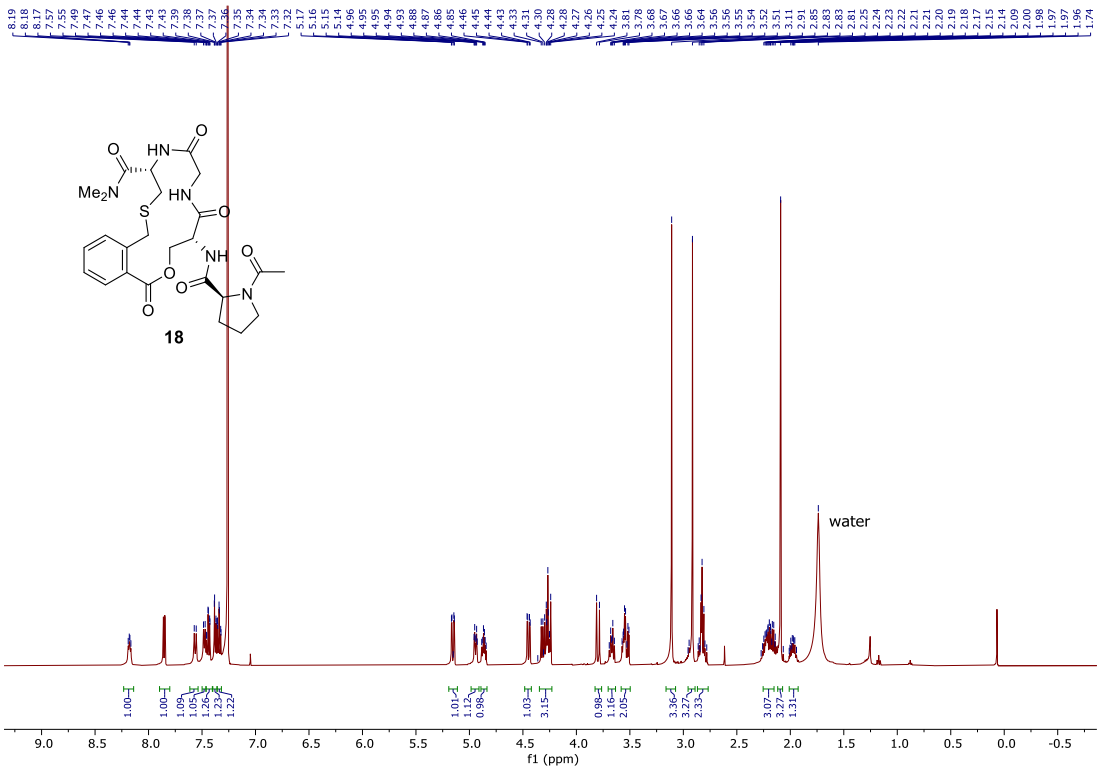
**38** $^{1}$H NMR (500 MHz, $CDCl_3$)



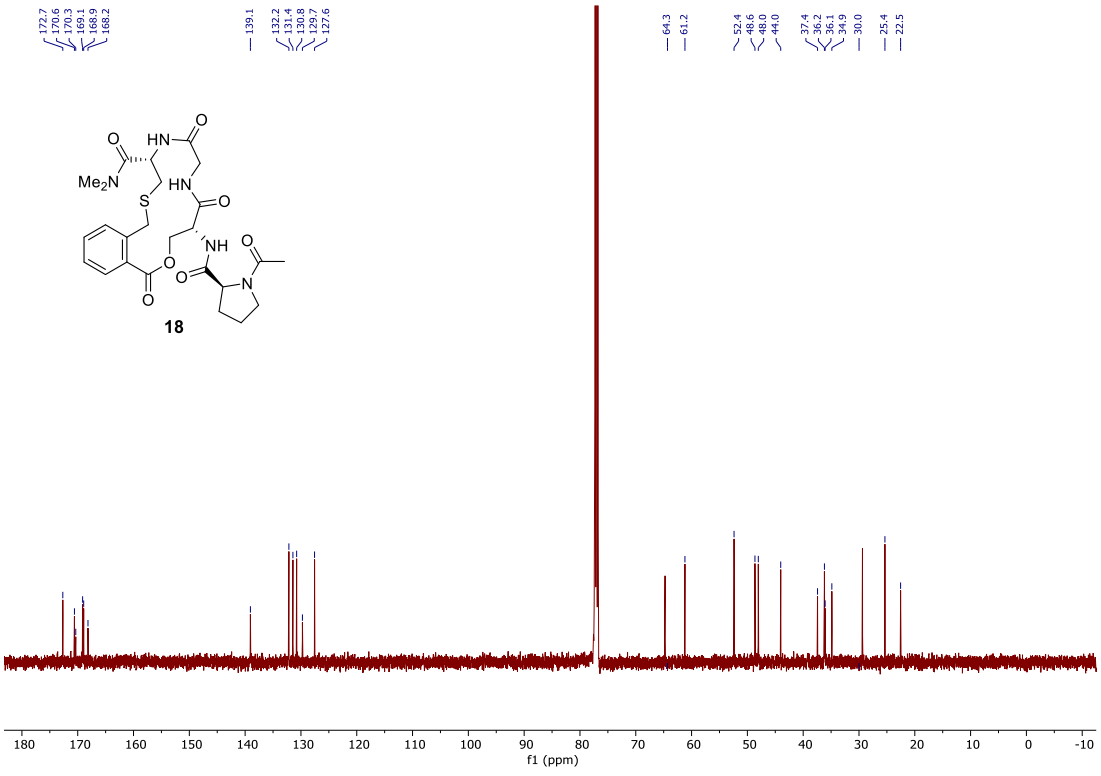

**38** $^{13}$C NMR (126 MHz, $CDCl_3$)



**18** $^{1}$H NMR (500 MHz, $CDCl_3$)



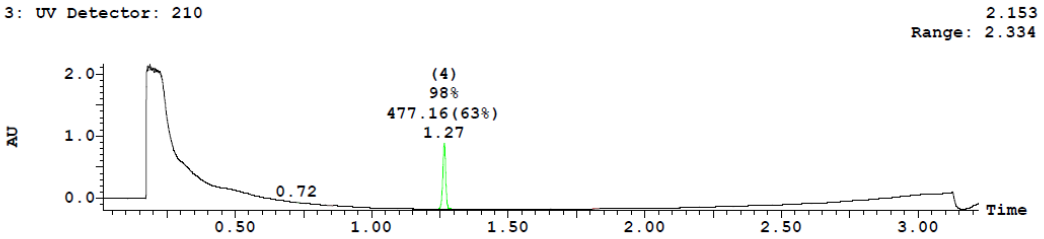

**18** $^{13}$C NMR (126 MHz, $CDCl_3$)



# Purity reports for compounds 6-18

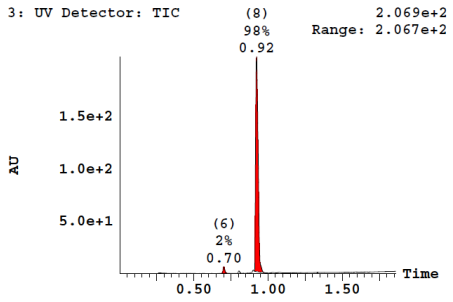
Compound **6**:

**Sample Report:**

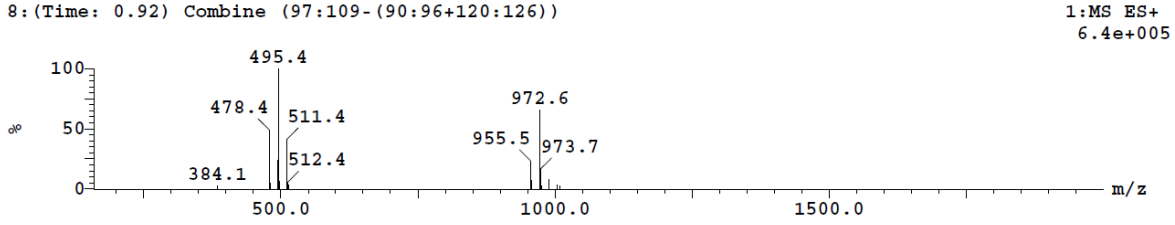



| Peak Number | Compound | Time | AreaAbs | Area %Total | Height | Mass Found |
|---|---|---|---|---|---|---|
| 1 | Found | 0.72 | 2e+001 | 0.11 | 1e+003 | 477.16 |
| 2 | | 0.84 | 7e+001 | 0.49 | 5e+003 | |
| 3 | | 1.03 | 2e+001 | 0.13 | 3e+003 | |
| 4 | Found | 1.27 | 1e+004 | 97.91 | 1e+006 | 477.16 |
| 5 | Found | 1.44 | 1e+001 | 0.11 | 2e+003 | 477.16 |
| 6 | Found | 1.52 | 5e+001 | 0.34 | 4e+003 | 477.16 |
| 7 | Found | 1.62 | 6e+001 | 0.45 | 5e+003 | 477.16 |
| 8 | | 1.82 | 6e+001 | 0.45 | 5e+003 | |

Compound **7**:

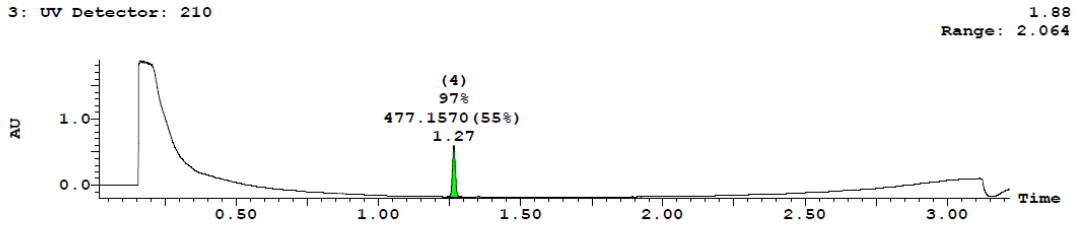



| Peak ID | Time | Mass Found |
|---|---|---|
| 8 | 0.93 | |



Compound **8**:

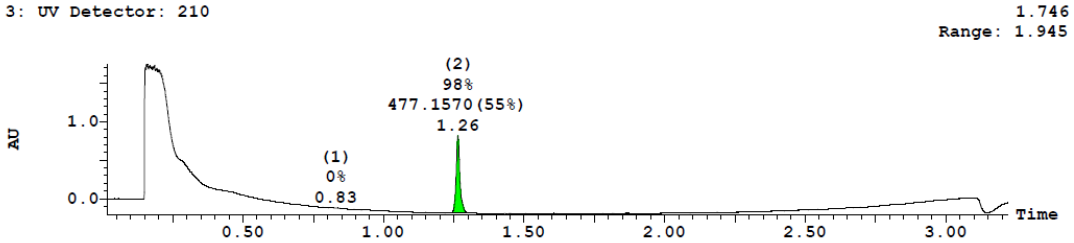
**Sample Report:**




| Peak Number | Compound | Time | AreaAbs | Area %Total | Height | Mass Found |
|---|---|---|---|---|---|---|
| 1 | | 0.87 | 9e+001 | 0.92 | 8e+003 | |
| 2 | Found | 1.21 | 1e+001 | 0.12 | 1e+003 | 477.16 |
| 3 | Found | 1.24 | 9e+000 | 0.09 | 2e+003 | 477.16 |
| 4 | Found | 1.26 | 9e+003 | 96.81 | 8e+005 | 477.16 |
| 5 | Found | 1.35 | 9e+001 | 0.89 | 9e+003 | 477.16 |
| 6 | | 1.42 | 5e+000 | 0.06 | 9e+002 | |
| 7 | | 1.48 | 2e+001 | 0.21 | 2e+003 | |
| 8 | | 1.51 | 8e+000 | 0.08 | 1e+003 | |
| 9 | | 1.56 | 1e+001 | 0.10 | 1e+003 | |
| 10 | | 1.61 | 1e+001 | 0.11 | 1e+003 | |
| 11 | | 1.64 | 5e+001 | 0.47 | 3e+003 | |
| 12 | | 1.89 | 1e+001 | 0.15 | 2e+003 | |

Compound **9**:

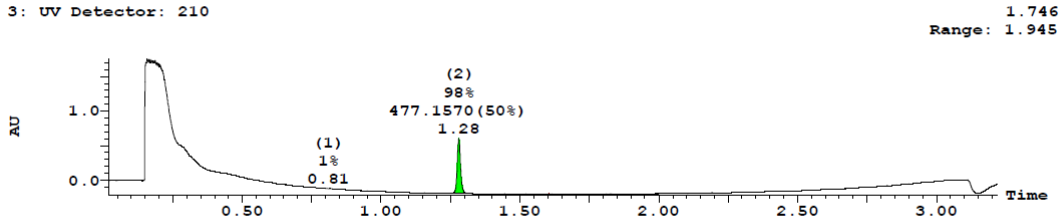
**Sample Report:**




| Peak Number | Compound | Time | AreaAbs | Area %Total | Height | Mass Found |
|---|---|---|---|---|---|---|
| 1 | | 0.83 | 5e+001 | 0.31 | 4e+003 | |
| 2 | Found | 1.26 | 2e+004 | 97.92 | 1e+006 | 477.16 |
| 3 | Found | 1.43 | 2e+001 | 0.11 | 2e+003 | 477.16 |
| 4 | | 1.51 | 5e+001 | 0.30 | 4e+003 | |
| 5 | | 1.60 | 9e+001 | 0.59 | 7e+003 | |
| 6 | | 1.86 | 1e+002 | 0.69 | 9e+003 | |
| 7 | | 2.28 | 2e+001 | 0.10 | 2e+003 | |

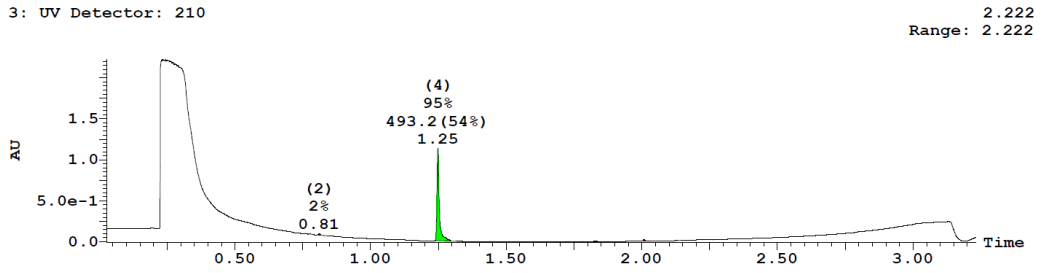
Compound **10**:

**Sample Report:**



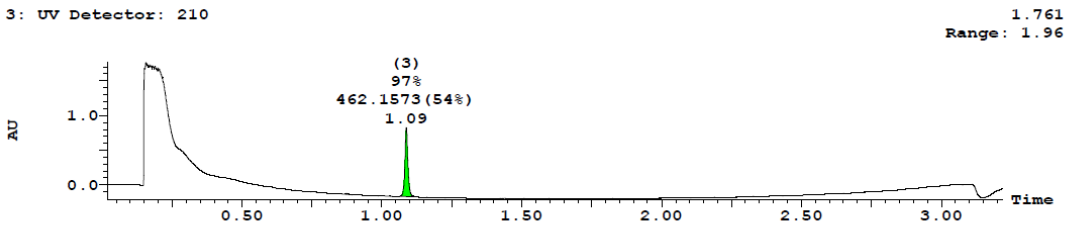

| Peak Number | Compound | Time | AreaAbs | Area %Total | Height | Mass Found |
|---|---|---|---|---|---|---|
| 1 | | 0.81 | 6e+001 | 0.50 | 4e+003 | |
| 2 | Found | 1.28 | 1e+004 | 97.96 | 8e+005 | 477.16 |
| 3 | | 1.51 | 5e+001 | 0.49 | 4e+003 | |
| 4 | | 1.60 | 1e+002 | 1.06 | 8e+003 | |

Compound **11**:



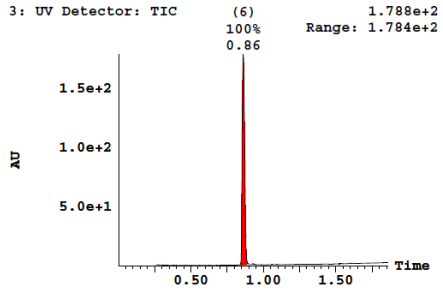

Compound **12**:

**Sample Report:**

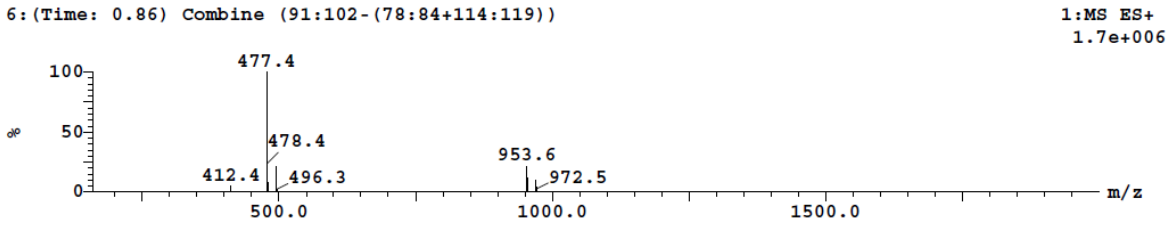



| Peak Number | Compound | Time | AreaAbs | Area %Total | Height | Mass Found |
|---|---|---|---|---|---|---|
| 1 | | 0.87 | 2e+002 | 1.46 | 2e+004 | |
| 2 | Found | 1.02 | 6e+001 | 0.49 | 6e+003 | 462.16 |
| 3 | Found | 1.09 | 1e+004 | 96.81 | 1e+006 | 462.16 |
| 4 | | 1.44 | 2e+001 | 0.13 | 2e+003 | |
| 5 | | 1.60 | 1e+002 | 0.79 | 7e+003 | |
| 6 | | 1.87 | 4e+001 | 0.33 | 4e+003 | |

Compound **13**:

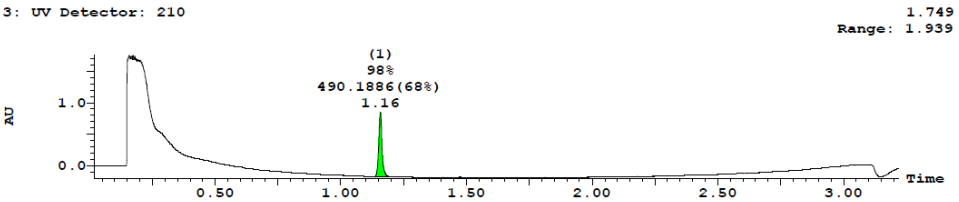



| Peak ID | Time | Mass Found |
|---|---|---|
| 6 | 0.86 | |




Compound **14**:

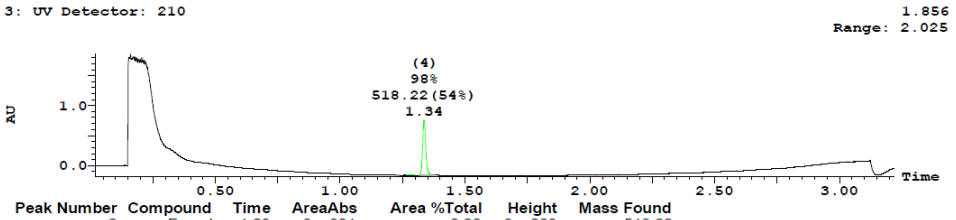
**Sample Report:**




| Peak Number | Compound | Time | AreaAbs | Area %Total | Height | Mass Found |
|---|---|---|---|---|---|---|
| 1 | Found | 1.16 | 2e+004 | 98.18 | 1e+006 | 490.19 |
| 2 | | 1.44 | 3e+001 | 0.21 | 4e+003 | |
| 3 | | 1.47 | 9e+001 | 0.56 | 8e+003 | |
| 4 | | 1.60 | 1e+002 | 0.72 | 8e+003 | |
| 5 | | 1.86 | 5e+001 | 0.29 | 4e+003 | |
| 6 | | 2.60 | 6e+000 | 0.04 | 1e+003 | |

Compound **15**:

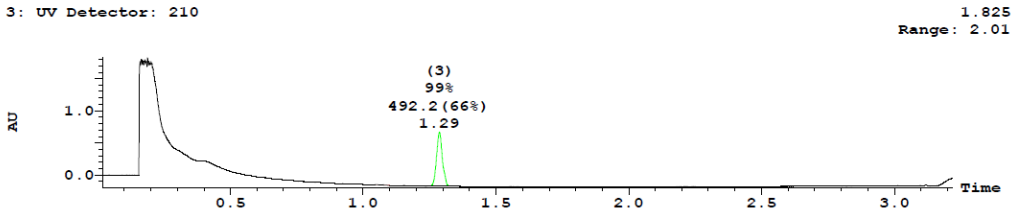
**Sample Report:**




| Peak Number | Compound | Time | AreaAbs | Area %Total | Height | Mass Found |
|---|---|---|---|---|---|---|
| 2 | Found | 1.23 | 3e+001 | 0.20 | 3e+003 | 518.22 |
| 3 | Found | 1.28 | 2e+002 | 1.45 | 1e+004 | 518.22 |
| 4 | Found | 1.34 | 1e+004 | 98.13 | 9e+005 | 518.22 |
| 5 | | 1.48 | 2e+001 | 0.18 | 3e+003 | |
| 7 | | 2.65 | 6e+000 | 0.05 | 2e+003 | |

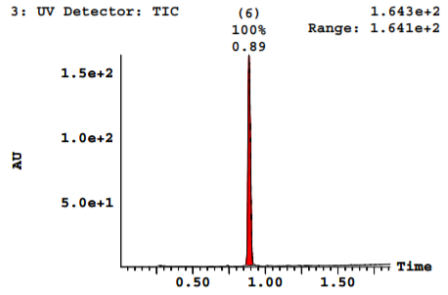
Compound **16**:

Sample Report:

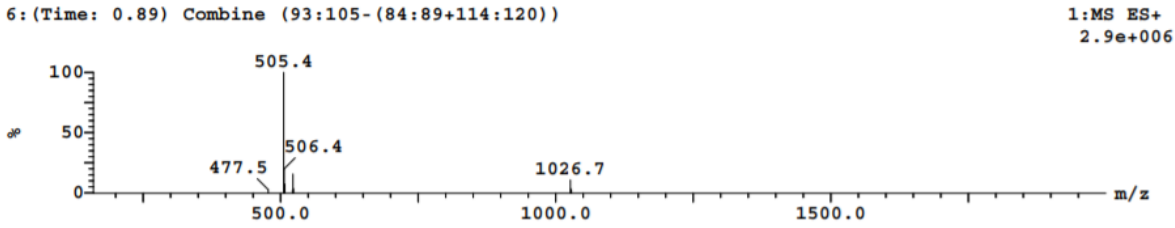



| Peak Number | Compound | Time | AreaAbs | Area %Total | Height | Mass Found |
|---|---|---|---|---|---|---|
| 1 | | 1.06 | 1e+002 | 0.57 | 6e+003 | |
| 2 | | 1.08 | 9e+001 | 0.41 | 6e+003 | |
| 3 | Found | 1.29 | 2e+004 | 98.91 | 8e+005 | 492.20 |
| 5 | Found | 1.49 | 2e+001 | 0.11 | 2e+003 | 492.20 |

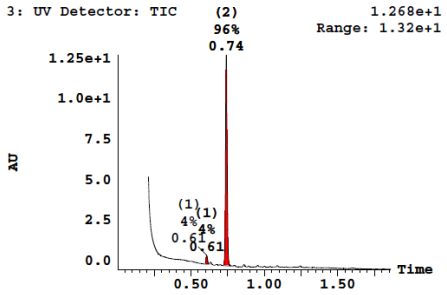
Compound **17**:

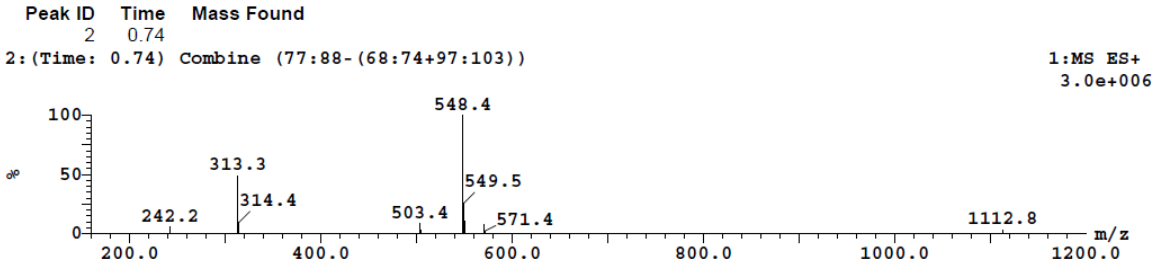



| Peak ID | Time | Mass Found |
|---|---|---|
| 6 | 0.88 | |



Compound **18**:


3: UV Detector: TIC
(2)
96%
0.74
1.268e+1
Range: 1.32e+1
1.25e+1
1.0e+1
7.5
5.0
2.5
0.0
AU
(1)
4%
0.61
Time
0.50
1.00
1.50


Peak ID Time Mass Found
2 0.74
2:(Time: 0.74) Combine (77:88-(68:74+97:103))
1:MS ES+
3.0e+006
100
50
0
%
548.4
313.3
549.5
314.4
242.2
503.4
571.4
1112.8
m/z
200.0
400.0
600.0
800.0
1000.0
1200.0

## References


1. Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M., Epik: a software program for pKaprediction and protonation state generation for drug-like molecules. *J. Comput. Aided. Mol. Des.* **2007,** *21* (12), 681–691.

2. Tran, K. T.; Pallesen, J. S.; Solbak, S. M. Ø.; Narayanan, D.; Baig, A.; Zang, J.; Aguayo-Orozco, A.; Carmona, R. M. C.; Garcia, A. D.; Bach, A., A comparative assessment study of known small-molecule Keap1−Nrf2 protein−protein interaction inhibitors: Chemical synthesis, binding properties, and cellular activity. *J. Med. Chem.* **2019,** *62*, 8028–8052.

3. Labute, P., A widely applicable set of descriptors. *J. Mol. Graph. Model.* **2000,** *18* (4-5), 464–477.

4. Marcotte, D.; Zeng, W.; Hus, J. C.; McKenzie, A.; Hession, C.; Jin, P.; Bergeron, C.; Lugovskoy, A.; Enyedy, I.; Cuervo, H.; Wang, D.; Atmanene, C.; Roecklin, D.; Vecchi, M.; Vivat, V.; Kraemer, J.; Winkler, D.; Hong, V.; Chao, J.; Lukashev, M.; Silvian, L., Small molecules inhibit the interaction of Nrf2 and the Keap1 Kelch domain through a non-covalent mechanism. *Bioorg. Med. Chem.* **2013,** *21* (14), 4011–4019.

5. Kunishima, N., Tanaka, T., Satoh, M., Saburi, H, Crystal structure of Keap1 in complex with synthetic small molecular based on a co-crystallization. **2012** To be Published.

6. Laskowski, R. A.; Swindells, M. B., LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model* **2011,** *51* (10), 2778–2786.

7. Greenidge, P. A.; Kramer, C.; Mozziconacci, J.-C.; Wolf, R. M., MM/GBSA binding energy prediction on the PDBbind data eet: Successes, failures, and directions for further improvement. *J. Chem. Inf. Model.* **2013,** *53*, 201−209.

8. Schrödinger Release 2017-1: Maestro (Version 11.1.011), Schrödinger, LLC, New York, NY, (2017).

9. Li, H.; Robertson, A. D.; Jensen, J. H., Very fast empirical prediction and rationalization of protein pKa values. *Proteins* **2005,** *61* (4), 704–721.

10. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A., OPLS3: A force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* **2016,** *12* (1), 281–296.

11. Schrödinger Release 2019-4: LigPrep, Schrödinger, LLC, New York, NY, (2019).

12. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004,** *47* (7), 1750–1759.
13. Satoh, M.; Saburi, H.; Tanaka, T.; Matsuura, Y.; Naitow, H.; Shimozono, R.; Yamamoto, N.; Inoue, H.; Nakamura, N.; Yoshizawa, Y.; Aoki, T.; Tanimura, R.; Kunishima, N., Multiple binding modes of a small molecule to human Keap1 revealed by X-ray crystallography and molecular dynamics simulation. *FEBS Open Bio* **2015,** *5*, 557–570.
14. Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C. P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L., 3rd; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F., 3rd; Kong, J.; Krylov, A. I.; Gill, P. M.; Head-Gordon, M., Advances in methods and algorithms in a modern quantum chemistry program package. *Phys. Chem. Chem. Phys.* **2006,** *8* (27), 3172–3191.
15. Walker, M.; Harvey, A. J.; Sen, A.; Dessent, C. E., Performance of M06, M06-2X, and M06-HF density functionals for conformationally flexible anionic clusters: M06 functionals perform better than B3LYP for a model system with dispersion and ionic hydrogen-bonding interactions. *J. Phys. Chem. A* **2013,** *117* (47), 12590–12600.