**pKa measurements for the SAMPL6 prediction challenge for a set of kinase inhibitor-like fragments**

**Mehtap Işık1,2 , Dorothy Levorse3 , Ariën S. Rustenburg1,4 , Ikenna E. Ndukwe5 , Heather Wang6 , Xiao Wang5 , Mikhail Reibarkh5 , Gary E. Martin5 , David L. Mobley7 , Timothy Rhodes3 , John D. Chodera1\*,6**

1 Computational and Systems Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065, United States

2 Tri-Institutional PhD Program in Chemical Biology, Weill Cornell Graduate School of Medical Sciences, Cornell University, New York, NY 10065, United States

3 Pharmaceutical Sciences, MRL, Merck & Co., Inc., 126 East Lincoln Avenue, Rahway, New Jersey 07065, United States

4 Graduate Program in Physiology, Biophysics, and Systems Biology, Weill Cornell Medical College, New York, NY 10065, United States

5 Process and Analytical Research and Development, Merck & Co., Inc., Rahway, NJ 07065, United States

6 Process and Analytical Research and Development, Merck & Co., Inc., Rahway, NJ 07065, United States

7 Department of Pharmaceutical Sciences and Department of Chemistry, University of California, Irvine, Irvine, California 92697, United States

\*For correspondence:

john.chodera@choderalab.org (JDC)

SUPPLEMENTARY INFORMATION

**Table SI 1.** Procurement details of SAMPL6 pKa challenge compounds

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SAMPL6 Molecule ID** | **Group** | **Supplier** | **LOT** | **CAT** | **Supplier reported purity** | **CAS** | **eMolecules ID** | **canonical isomeric SMILES** | **Experimental Molecule ID** |
| SM01 | fragment-like | AchemBlock | 11549 | 10222 | 95% | 521937-07-05 | 6679830 | c1cc2c(cc1O)c3c(o2)C(=O)NCCC3 | M01 |
| SM02 | fragment-like | ChemDiv | CM02432403 | 3232-0333 |  |  | 1327907 | c1ccc2c(c1)c(ncn2)Nc3cccc(c3)C(F)(F)F | M02 |
| SM03 | fragment-like | ChemDiv |  | Z27474679 |  |  | 1228629 | c1ccc(cc1)Cc2nnc(s2)NC(=O)c3cccs3 | M03 |
| SM04 | fragment-like | ChemDiv |  | Z126957826 |  |  | 30719859 | c1ccc2c(c1)c(ncn2)NCc3ccc(cc3)Cl | M04 |
| SM05 | fragment-like | ChemDiv |  | Z119335440 |  |  | 18908671 | c1ccc(c(c1)NC(=O)c2ccc(o2)Cl)N3CCCCC3 | M05 |
| SM06 | fragment-like | ChemDiv |  | Z28487401 |  |  | 18893169 | c1cc2cccnc2c(c1)NC(=O)c3cc(cnc3)Br | M06 |
| SM07 | fragment-like | Enamine | 2017-0168841 | Z57161635 | 95% | 100818-54-0 | 1327878 | c1ccc(cc1)CNc2c3ccccc3ncn2 | M07 |
| SM08 | fragment-like | Enamine | 2017-0168838 | Z57157353 | 95% | 65418-08-8 | 1367649 | Cc1ccc2c(c1)c(c(c(=O)[nH]2)CC(=O)O)c3ccccc3 | M08 |
| SM09 | fragment-like | Enamine | 2017-0168839 | Z220564816 | 95% |  | 1865544 | COc1cccc(c1)Nc2c3ccccc3ncn2.Cl | M09 |
| SM10 | fragment-like | Enamine | 2017-0168843 | Z69130143 | 95% | 35056-22-5 | 23354217 | c1ccc(cc1)C(=O)NCC(=O)Nc2nc3ccccc3s2 | M10 |
| SM11 | fragment-like | Maybridge | 142989 | RJC00689SC | 90% | 5334-30-5 | 719540 | c1ccc(cc1)n2c3c(cn2)c(ncn3)N | M11 |
| SM12 | fragment-like | Maybridge | 265423 | DP00818SC |  |  | 1859493 | c1ccc2c(c1)c(ncn2)Nc3cccc(c3)Cl.Cl | M12 |
| SM13 | fragment-like | Maybridge | 248841 | GK03474SC |  |  | 5828805 | Cc1cccc(c1)Nc2c3cc(c(cc3ncn2)OC)OC | M13 |
| SM14 | fragment-like | Enamine |  | Z57290870 |  |  | 31653344 | c1ccc(cc1)n2cnc3c2ccc(c3)N | M15 |
| SM15 | fragment-like | Enamine |  | Z1318268952 |  |  | 37095168 | c1ccc2c(c1)ncn2c3ccc(cc3)O | M16 |
| SM16 | fragment-like | VitaScreen |  | STK098832 |  |  | 1284691 | c1cc(c(c(c1)Cl)C(=O)Nc2ccncc2)Cl | M18 |
| SM17 | fragment-like | VitaScreen |  | STK032731 |  |  | 1444229 | c1ccc(cc1)CSc2nnc(o2)c3ccncc3 | M19 |
| SM18 | drug-like | Enamine |  | Z278071350 |  |  | 18897105 | c1ccc2c(c1)c(=O)[nH]c(n2)CCC(=O)Nc3ncc(s3)Cc4ccc(c(c4)F)F | D01 |
| SM19 | drug-like | Enamine |  | Z30206127 |  |  | 3365457 | CCOc1ccc2c(c1)sc(n2)NC(=O)Cc3ccc(c(c3)Cl)Cl | D02 |
| SM20 | drug-like | VitaScreen |  | STL282831 |  |  | 46568819 | c1cc(cc(c1)OCc2ccc(cc2Cl)Cl)/C=C/3\C(=O)NC(=O)S3 | D05 |
| SM21 | drug-like | VitaScreen |  | STL368658 |  |  | 1574612 | c1cc(cc(c1)Br)Nc2c(cnc(n2)Nc3cccc(c3)Br)F | D06 |
| SM22 | drug-like | VitaScreen |  | STK070581 |  |  | 536848 | c1cc2c(cc(c(c2nc1)O)I)I | D07 |
| SM23 | drug-like | VitaScreen |  | STK097966 |  |  | 4375254 | CCOC(=O)c1ccc(cc1)Nc2cc(nc(n2)Nc3ccc(cc3)C(=O)OCC)C | D08 |
| SM24 | drug-like | VitaScreen |  | STK090644 |  |  | 1415746 | COc1ccc(cc1)c2c3c(ncnc3oc2c4ccc(cc4)OC)NCCO | D09 |

**Table SI 2.** Calculated properties and descriptors of compounds of 28 compounds selected and procured for pKa challenge. pKa measurement experiments were successful 24 molecules labeled SM01-SM24.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SAMPL6 Molecule ID** | **Experimental Molecule ID** | **group** | **Epik pKas in [3,11] range** | **OpenEye XlogP** | **Molecular Weight (g/mol)** | **eMolecules reported availability (mg)** | **Number of rotatable bonds** | **Number of UV-chr. units** | **eMolecules ID** | **canonical isomeric SMILES** |
| SM01 | M01 | fragment-like | [9.119] | 3.27 | 289.26 | 184 | 0 | 27 | 6679830 | c1cc2c(cc1O)c3c(o2)C(=O)NCCC3 |
| SM02 | M02 | fragment-like | [4.05] | 3.61 | 301.39 | 101 | 3 | 36 | 1327907 | c1ccc2c(c1)c(ncn2)Nc3cccc(c3)C(F)(F)F |
| SM03 | M03 | fragment-like | [7.12] | 3.52 | 269.73 | 379 | 5 | 14 | 1228629 | c1ccc(cc1)Cc2nnc(s2)NC(=O)c3cccs3 |
| SM04 | M04 | fragment-like | [5.564] | 3.79 | 304.77 | 415.5 | 3 | 36 | 30719859 | c1ccc2c(c1)c(ncn2)NCc3ccc(cc3)Cl |
| SM05 | M05 | fragment-like | [5.346] | 2.92 | 328.16 | 424.3 | 4 | 18 | 18908671 | c1ccc(c(c1)NC(=O)c2ccc(o2)Cl)N3CCCCC3 |
| SM06 | M06 | fragment-like | [4.001, 10.328] | 2.91 | 235.28 | 406 | 3 | 37 | 18893169 | c1cc2cccnc2c(c1)NC(=O)c3cc(cnc3)Br |
| SM07 | M07 | fragment-like | [5.564] | 2.80 | 293.32 | 208.7 | 3 | 36 | 1327878 | c1ccc(cc1)CNc2c3ccccc3ncn2 |
| SM08 | M08 | fragment-like | [4.109] | 3.14 | 287.74 | 232.1 | 3 | 59 | 1367649 | Cc1ccc2c(c1)c(c(c(=O)[nH]2)CC(=O)O)c3ccccc3 |
| SM09 | M09 | fragment-like | [4.05] | 2.72 | 311.36 | 119.7 | 3 | 36 | 1865544 | COc1cccc(c1)Nc2c3ccccc3ncn2.Cl |
| SM10 | M10 | fragment-like | [8.672] | 1.50 | 211.22 | 149.1 | 6 | 28 | 23354217 | c1ccc(cc1)C(=O)NCC(=O)Nc2nc3ccccc3s2 |
| SM11 | M11 | fragment-like | [3.869] | 3.90 | 292.16 | 3430 | 1 | 31 | 719540 | c1ccc(cc1)n2c3c(cn2)c(ncn3)N |
| SM12 | M12 | fragment-like | [4.05] | 2.60 | 295.34 | 7366 | 2 | 36 | 1859493 | c1ccc2c(c1)c(ncn2)Nc3cccc(c3)Cl.Cl |
| SM13 | M13 | fragment-like | [4.267] | 2.91 | 235.28 | 1864 | 4 | 36 | 5828805 | Cc1cccc(c1)Nc2c3cc(c(cc3ncn2)OC)OC |
| - | M14 | fragment-like | [5.564] | 2.33 | 209.25 | 208.7 | 3 | 36 | 1327878 | c1ccc(cc1)CNc2c3ccccc3ncn2 |
| SM14 | M15 | fragment-like | [6.348] | 2.22 | 210.23 | 50213 | 1 | 40 | 31653344 | c1ccc(cc1)n2cnc3c2ccc(c3)N |
| SM15 | M16 | fragment-like | [5.82, 8.709] | 4.23 | 263.33 | 21650.2 | 1 | 40 | 37095168 | c1ccc2c(c1)ncn2c3ccc(cc3)O |
| - | M17 | fragment-like | [3.158] | 2.93 | 267.11 | 283.7 | 2 | 24 | 45809595 | CC(C)c1ccc(cc1)/C=C\2/c3ccccc3NC2=O |
| SM16 | M18 | fragment-like | [4.714, 9.847] | 3.31 | 269.32 | 385 | 3 | 20 | 1284691 | c1cc(c(c(c1)Cl)C(=O)Nc2ccncc2)Cl |
| SM17 | M19 | fragment-like | [4.902] | 3.34 | 426.44 | 170 | 4 | 30 | 1444229 | c1ccc(cc1)CSc2nnc(o2)c3ccncc3 |
| SM18 | D01 | drug-like | [9.381, 10.773] | 5.17 | 381.28 | 247.7 | 7 | 37 | 18897105 | c1ccc2c(c1)c(=O)[nH]c(n2)CCC(=O)Nc3ncc(s3)Cc4ccc(c(c4)F)F |
| SM19 | D02 | drug-like | [9.167] | 5.78 | 403.31 | 489.9 | 6 | 28 | 3365457 | CCOc1ccc2c(c1)sc(n2)NC(=O)Cc3ccc(c(c3)Cl)Cl |
| - | D03 | drug-like | [4.113] | 4.72 | 401.48 | 324.5 | 6 | 35 | 10794751 | CC(C)(C)c1cc(n(n1)c2ccccc2)NC(=O)Nc3cccc(c3Cl)Cl |
| - | D04 | drug-like | [3.199] | 5.24 | 380.25 | 636.9 | 5 | 34 | 3064762 | c1ccc(cc1)C(=O)Nc2ccc(cc2)Oc3c4c5c(sc4ncn3)CCCC5 |
| SM20 | D05 | drug-like | [8.05] | 4.14 | 438.09 | 154 | 4 | 24 | 46568819 | c1cc(cc(c1)OCc2ccc(cc2Cl)Cl)/C=C/3\C(=O)NC(=O)S3 |
| SM21 | D06 | drug-like | [3.892] | 3.37 | 396.95 | 222 | 4 | 28 | 1574612 | c1cc(cc(c1)Br)Nc2c(cnc(n2)Nc3cccc(c3)Br)F |
| SM22 | D07 | drug-like | [3.511, 6.794] | 2.94 | 420.46 | 239 | 0 | 29 | 536848 | c1cc2c(cc(c(c2nc1)O)I)I |
| SM23 | D08 | drug-like | [6.336] | 2.79 | 391.42 | 319 | 10 | 28 | 4375254 | CCOC(=O)c1ccc(cc1)Nc2cc(nc(n2)Nc3ccc(cc3)C(=O)OCC)C |
| SM24 | D09 | drug-like | [4.829] | 3.27 | 289.26 | 398 | 7 | 71 | 1415746 | COc1ccc(cc1)c2c3c(ncnc3oc2c4ccc(cc4)OC)NCCO |

**Table SI 3**. pKa results of replicate UV-metric pKa measurements. "UV-metric pKa" assay indicates spectrophotometric pKa measurements done with Sirius T3 in ISA water. "UV-metric pKa with cosolvent" assay refers to pKa determination by Yasuda-Shedlovsky extrapolation from psKa measurements in various ratios of ISA methanol:water mixtures. Triplicate measurements were performed at 25 ± 0.5°C and in the presence of approximately 150 mM KCl to adjust ionic strength.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Molecule ID | pKa1 | pKa2 | pKa3 | Assay Type | Experiment ID | Experiment Report |
| SM01 | 9.54 |  |  | UV-metric pKa | 17I-15024 | SM01\_17I-15024\_M01\_UV-metric pKa\_report.pdf |
| SM01 | 9.53 |  |  | UV-metric pKa | 17I-15025 | SM01\_17I-15025\_M01\_UV-metric pKa\_report.pdf |
| SM01 | 9.53 |  |  | UV-metric pKa | 17I-16001 | SM01\_17I-16001\_M01\_UV-metric pKa\_report.pdf |
| SM02 | 5.04 |  |  | UV-metric pKa with cosolvent | 17I-22022 | SM02\_17I-22022\_M02\_UV-metric psKa\_report.pdf |
| SM02 | 5.04 |  |  | UV-metric pKa with cosolvent | 17I-22023 | SM02\_17I-22023\_M02\_UV-metric psKa\_report.pdf |
| SM02 | 5.02 |  |  | UV-metric pKa with cosolvent | 17I-22024 | SM02\_17I-22024\_M02\_UV-metric psKa\_report.pdf |
| SM03 | 7.01 |  |  | UV-metric pKa with cosolvent | 17I-19004 | SM03\_17I-19004\_M03\_UV-metric psKa\_report.pdf |
| SM03 | 7.01 |  |  | UV-metric pKa with cosolvent | 17I-19005 | SM03\_17I-19005\_M03\_UV-metric psKa\_report.pdf |
| SM03 | 7.03 |  |  | UV-metric pKa with cosolvent | 17I-19006 | SM03\_17I-19006\_M03\_UV-metric psKa\_report.pdf |
| SM04 | 6.03 |  |  | UV-metric pKa | 17I-18018 | SM04\_17I-18018\_M04\_UV-metric pKa\_report.pdf |
| SM04 | 6.02 |  |  | UV-metric pKa | 17I-18019 | SM04\_17I-18019\_M04\_UV-metric pKa\_report.pdf |
| SM04 | 6.02 |  |  | UV-metric pKa | 17I-18020 | SM04\_17I-18020\_M04\_UV-metric pKa\_report.pdf |
| SM05 | 4.57 |  |  | UV-metric pKa with cosolvent | 17I-19007 | SM05\_17I-19007\_M05\_UV-metric psKa\_report.pdf |
| SM05 | 4.6 |  |  | UV-metric pKa with cosolvent | 17I-19008 | SM05\_17I-19008\_M05\_UV-metric psKa\_report.pdf |
| SM05 | 4.6 |  |  | UV-metric pKa with cosolvent | 17I-19009 | SM05\_17I-19009\_M05\_UV-metric psKa\_report.pdf |
| SM06 | 2.96 | 11.76 |  | UV-metric pKa | 17I-18021 | SM06\_17I-18021\_M06\_UV-metric pKa\_report.pdf |
| SM06 | 3.11 | 11.74 |  | UV-metric pKa | 17I-18022 | SM06\_17I-18022\_M06\_UV-metric pKa\_report.pdf |
| SM06 | 3.02 | 11.71 |  | UV-metric pKa | 17I-18023 | SM06\_17I-18023\_M06\_UV-metric pKa\_report.pdf |
| SM07 | 6.07 |  |  | UV-metric pKa | 17I-16002 | SM07\_17I-16002\_M07\_UV-metric pKa\_report.pdf |
| SM07 | 6.07 |  |  | UV-metric pKa | 17I-16004 | SM07\_17I-16004\_M07\_UV-metric pKa\_report.pdf |
| SM07 | 6.09 |  |  | UV-metric pKa | 17I-20001 | SM07\_17I-20001\_M07\_UV-metric pKa\_report.pdf |
| SM08 | 4.23 |  |  | UV-metric pKa | 17I-19001 | SM08\_17I-19001\_M08\_UV-metric pKa\_report.pdf |
| SM08 | 4.2 |  |  | UV-metric pKa | 17I-19002 | SM08\_17I-19002\_M08\_UV-metric pKa\_report.pdf |
| SM08 | 4.22 |  |  | UV-metric pKa | 17I-19003 | SM08\_17I-19003\_M08\_UV-metric pKa\_report.pdf |
| SM09 | 5.37 |  |  | UV-metric pKa with cosolvent | 17I-16014 | SM09\_17I-16014\_M09\_UV-metric psKa\_report.pdf |
| SM09 | 5.35 |  |  | UV-metric pKa with cosolvent | 17I-16015 | SM09\_17I-16015\_M09\_UV-metric psKa\_report.pdf |
| SM09 | 5.4 |  |  | UV-metric pKa with cosolvent | 17I-16016 | SM09\_17I-16016\_M09\_UV-metric psKa\_report.pdf |
| SM10 | 9.01 |  |  | UV-metric pKa with cosolvent | 17I-20020 | SM10\_17I-20020\_M10\_UV-metric psKa\_report.pdf |
| SM10 | 9.02 |  |  | UV-metric pKa with cosolvent | 17I-20021 | SM10\_17I-20021\_M10\_UV-metric psKa\_report.pdf |
| SM10 | 9.02 |  |  | UV-metric pKa with cosolvent | 17I-20022 | SM10\_17I-20022\_M10\_UV-metric psKa\_report.pdf |
| SM11 | 3.89 |  |  | UV-metric pKa | 17I-16005 | SM11\_17I-16005\_M11\_UV-metric pKa\_report.pdf |
| SM11 | 3.89 |  |  | UV-metric pKa | 17I-16006 | SM11\_17I-16006\_M11\_UV-metric pKa\_report.pdf |
| SM11 | 3.89 |  |  | UV-metric pKa | 17I-16007 | SM11\_17I-16007\_M11\_UV-metric pKa\_report.pdf |
| SM12 | 5.28 |  |  | UV-metric pKa | 17I-21002 | SM12\_17I-21002\_M12\_UV-metric pKa\_report.pdf |
| SM12 | 5.28 |  |  | UV-metric pKa | 17I-21003 | SM12\_17I-21003\_M12\_UV-metric pKa\_report.pdf |
| SM12 | 5.27 |  |  | UV-metric pKa | 17I-21004 | SM12\_17I-21004\_M12\_UV-metric pKa\_report.pdf |
| SM13 | 5.79 |  |  | UV-metric pKa | 17I-16011 | SM13\_17I-16011\_M13\_UV-metric pKa\_report.pdf |
| SM13 | 5.76 |  |  | UV-metric pKa | 17I-16012 | SM13\_17I-16012\_M13\_UV-metric pKa\_report.pdf |
| SM13 | 5.76 |  |  | UV-metric pKa | 17I-16013 | SM13\_17I-16013\_M13\_UV-metric pKa\_report.pdf |

**Table SI 3**. Continued.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Molecule ID | pKa1 | pKa2 | pKa3 | Assay Type | Experiment ID | Experiment Report |
| SM14 | 2.59 | 5.31 |  | UV-metric pKa | 17I-29002 | SM14\_17I-29002\_M15\_UV-metric pKa\_report.pdf |
| SM14 | 2.57 | 5.29 |  | UV-metric pKa | 17I-29003 | SM14\_17I-29003\_M15\_UV-metric pKa\_report.pdf |
| SM14 | 2.57 | 5.29 |  | UV-metric pKa | 17I-29004 | SM14\_17I-29004\_M15\_UV-metric pKa\_report.pdf |
| SM15 | 4.71 | 8.96 |  | UV-metric pKa | 17K-10009 | SM15\_17K-10009\_M16\_UV-metric pKa\_report.pdf |
| SM15 | 4.7 | 8.94 |  | UV-metric pKa | 17K-10010 | SM15\_17K-10010\_M16\_UV-metric pKa\_report.pdf |
| SM15 | 4.69 | 8.92 |  | UV-metric pKa | 17K-10011 | SM15\_17K-10011\_M16\_UV-metric pKa\_report.pdf |
| SM16 | 5.37 | 10.64 |  | UV-metric pKa | 17J-03025 | SM16\_17J-03025\_M18\_UV-metric pKa\_report.pdf |
| SM16 | 5.37 | 10.65 |  | UV-metric pKa | 17J-03026 | SM16\_17J-03026\_M18\_UV-metric pKa\_report.pdf |
| SM16 | 5.38 | 10.65 |  | UV-metric pKa | 17J-03027 | SM16\_17J-03027\_M18\_UV-metric pKa\_report.pdf |
| SM17 | 3.15 |  |  | UV-metric pKa | 17J-02024 | SM17\_17J-02024\_M19\_UV-metric pKa\_report.pdf |
| SM17 | 3.16 |  |  | UV-metric pKa | 17J-02025 | SM17\_17J-02025\_M19\_UV-metric pKa\_report.pdf |
| SM17 | 3.17 |  |  | UV-metric pKa | 17J-02026 | SM17\_17J-02026\_M19\_UV-metric pKa\_report.pdf |
| SM18 | 2.1 | 9.51 | 10.93 | UV-metric pKa with cosolvent | 17I-28001 | SM18\_17I-28001\_D01\_UV-metric psKa\_report.pdf |
| SM18 | 2.15 | 9.59 | 11.03 | UV-metric pKa with cosolvent | 17I-28002 | SM18\_17I-28002\_D01\_UV-metric psKa\_report.pdf |
| SM18 | 2.19 | 9.63 | 11.11 | UV-metric pKa with cosolvent | 17I-28003 | SM18\_17I-28003\_D01\_UV-metric psKa\_report.pdf |
| SM19 | 9.59 |  |  | UV-metric pKa with cosolvent | 17I-28004 | SM19\_17I-28004\_D02\_UV-metric psKa\_report.pdf |
| SM19 | 9.58 |  |  | UV-metric pKa with cosolvent | 17I-28005 | SM19\_17I-28005\_D02\_UV-metric psKa\_report.pdf |
| SM19 | 9.52 |  |  | UV-metric pKa with cosolvent | 17I-28006 | SM19\_17I-28006\_D02\_UV-metric psKa\_report.pdf |
| SM20 | 5.72 |  |  | UV-metric pKa with cosolvent | 17J-12002 | SM20\_17J-12002\_D05\_UV-metric psKa\_report.pdf |
| SM20 | 5.75 |  |  | UV-metric pKa with cosolvent | 17J-12003 | SM20\_17J-12003\_D05\_UV-metric psKa\_report.pdf |
| SM20 | 5.64 |  |  | UV-metric pKa with cosolvent | 17K-16013 | SM20\_17K-16013\_D05\_UV-metric psKa\_report.pdf |
| SM21 | 4.11 |  |  | UV-metric pKa with cosolvent | 17J-06002 | SM21\_17J-06002\_D06\_UV-metric psKa\_report.pdf |
| SM21 | 4.09 |  |  | UV-metric pKa with cosolvent | 17J-06003 | SM21\_17J-06003\_D06\_UV-metric psKa\_report.pdf |
| SM21 | 4.09 |  |  | UV-metric pKa with cosolvent | 17J-06004 | SM21\_17J-06004\_D06\_UV-metric psKa\_report.pdf |
| SM22 | 2.45 | 7.42 |  | UV-metric pKa with cosolvent | 17J-07003 | SM22\_17J-07003\_D07\_UV-metric psKa\_report.pdf |
| SM22 | 2.37 | 7.45 |  | UV-metric pKa with cosolvent | 17J-07004 | SM22\_17J-07004\_D07\_UV-metric psKa\_report.pdf |
| SM22 | 2.37 | 7.41 |  | UV-metric pKa with cosolvent | 17J-07005 | SM22\_17J-07005\_D07\_UV-metric psKa\_report.pdf |
| SM23 | 5.45 |  |  | UV-metric pKa with cosolvent | 17J-07006 | SM23\_17J-07006\_D08\_UV-metric psKa\_report.pdf |
| SM23 | 5.44 |  |  | UV-metric pKa with cosolvent | 17J-07007 | SM23\_17J-07007\_D08\_UV-metric psKa\_report.pdf |
| SM23 | 5.45 |  |  | UV-metric pKa with cosolvent | 17J-07008 | SM23\_17J-07008\_D08\_UV-metric psKa\_report.pdf |
| SM24 | 2.61 |  |  | UV-metric pKa with cosolvent | 17J-06007 | SM24\_17J-06007\_D09\_UV-metric psKa\_report.pdf |
| SM24 | 2.58 |  |  | UV-metric pKa with cosolvent | 17J-06008 | SM24\_17J-06008\_D09\_UV-metric psKa\_report.pdf |
| SM24 | 2.6 |  |  | UV-metric pKa with cosolvent | 17J-06009 | SM24\_17J-06009\_D09\_UV-metric psKa\_report.pdf |

**Table SI 4**. UV-metric pKa measurments with and without cosolvent for 12 pKa challenge compounds with higher aqueous solubility and pyridoxine HCl. pKa values were measured with both methods to test if cosolvent method introduces bias.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Molecule ID** | **pKa1** | **pKa2** | **pKa3** | **Assay Type** | **Experiment ID** | **Experiment Report** |
| SM01 | 9.54 |  |  | UV-metric pKa | 17I-15024 | SM01\_17I-15024\_M01\_UV-metric pKa\_report.pdf |
| SM01 | 9.53 |  |  | UV-metric pKa | 17I-15025 | SM01\_17I-15025\_M01\_UV-metric pKa\_report.pdf |
| SM01 | 9.53 |  |  | UV-metric pKa | 17I-16001 | SM01\_17I-16001\_M01\_UV-metric pKa\_report.pdf |
| SM01\* | 9.71 |  |  | UV-metric pKa with cosolvent | 17J-06011 | SM01\_17J-06011\_M01\_UV-metric psKa\_report.pdf |
| SM04 | 6.03 |  |  | UV-metric pKa | 17I-18018 | SM04\_17I-18018\_M04\_UV-metric pKa\_report.pdf |
| SM04 | 6.02 |  |  | UV-metric pKa | 17I-18019 | SM04\_17I-18019\_M04\_UV-metric pKa\_report.pdf |
| SM04 | 6.02 |  |  | UV-metric pKa | 17I-18020 | SM04\_17I-18020\_M04\_UV-metric pKa\_report.pdf |
| SM04\* | 5.97 |  |  | UV-metric pKa with cosolvent | 17K-09020 | SM04\_17K-09020\_M04\_UV-metric psKa\_report.pdf |
| SM06 | 2.96 | 11.76 |  | UV-metric pKa | 17I-18021 | SM06\_17I-18021\_M06\_UV-metric pKa\_report.pdf |
| SM06 | 3.11 | 11.74 |  | UV-metric pKa | 17I-18022 | SM06\_17I-18022\_M06\_UV-metric pKa\_report.pdf |
| SM06 | 3.02 | 11.71 |  | UV-metric pKa | 17I-18023 | SM06\_17I-18023\_M06\_UV-metric pKa\_report.pdf |
| SM06\* | 3.46 | >12 |  | UV-metric pKa with cosolvent | 17K-09021 | SM06\_17K-09021\_M06\_UV-metric psKa\_report.pdf |
| SM06\* | 3.4 | >12 |  | UV-metric pKa with cosolvent | 17K-10008 | SM06\_17K-10008\_M06\_UV-metric psKa\_report.pdf |
| SM06\* | 3.45 | >12 |  | UV-metric pKa with cosolvent | 17K-16011 | SM06\_17K-16011\_M06\_UV-metric psKa\_report.pdf |
| SM07 | 6.07 |  |  | UV-metric pKa | 17I-16002 | SM07\_17I-16002\_M07\_UV-metric pKa\_report.pdf |
| SM07 | 6.07 |  |  | UV-metric pKa | 17I-16004 | SM07\_17I-16004\_M07\_UV-metric pKa\_report.pdf |
| SM07 | 6.09 |  |  | UV-metric pKa | 17I-20001 | SM07\_17I-20001\_M07\_UV-metric pKa\_report.pdf |
| SM07\* | 5.96 |  |  | UV-metric pKa with cosolvent | 17J-12006 | SM07\_17J-12006\_M07\_UV-metric psKa\_report.pdf |
| SM08 | 4.23 |  |  | UV-metric pKa | 17I-19001 | SM08\_17I-19001\_M08\_UV-metric pKa\_report.pdf |
| SM08 | 4.2 |  |  | UV-metric pKa | 17I-19002 | SM08\_17I-19002\_M08\_UV-metric pKa\_report.pdf |
| SM08 | 4.22 |  |  | UV-metric pKa | 17I-19003 | SM08\_17I-19003\_M08\_UV-metric pKa\_report.pdf |
| SM08\* | 4.47 |  |  | UV-metric pKa with cosolvent | 17J-12007 | SM08\_17J-12007\_M08\_UV-metric psKa\_report.pdf |
| SM11 | 3.89 |  |  | UV-metric pKa | 17I-16005 | SM11\_17I-16005\_M11\_UV-metric pKa\_report.pdf |
| SM11 | 3.89 |  |  | UV-metric pKa | 17I-16006 | SM11\_17I-16006\_M11\_UV-metric pKa\_report.pdf |
| SM11 | 3.89 |  |  | UV-metric pKa | 17I-16007 | SM11\_17I-16007\_M11\_UV-metric pKa\_report.pdf |
| SM11\* | 3.74 |  |  | UV-metric pKa with cosolvent | 17J-04003 | SM11\_17J-04003\_M11\_UV-metric psKa\_report.pdf |
| SM12 | 5.28 |  |  | UV-metric pKa | 17I-21002 | SM12\_17I-21002\_M12\_UV-metric pKa\_report.pdf |
| SM12 | 5.28 |  |  | UV-metric pKa | 17I-21003 | SM12\_17I-21003\_M12\_UV-metric pKa\_report.pdf |
| SM12 | 5.27 |  |  | UV-metric pKa | 17I-21004 | SM12\_17I-21004\_M12\_UV-metric pKa\_report.pdf |
| SM12\* | 5.16 |  |  | UV-metric pKa with cosolvent | 17J-11014 | SM12\_17J-11014\_M12\_UV-metric psKa\_report.pdf |
| SM13 | 5.79 |  |  | UV-metric pKa | 17I-16011 | SM13\_17I-16011\_M13\_UV-metric pKa\_report.pdf |
| SM13 | 5.76 |  |  | UV-metric pKa | 17I-16012 | SM13\_17I-16012\_M13\_UV-metric pKa\_report.pdf |
| SM13 | 5.76 |  |  | UV-metric pKa | 17I-16013 | SM13\_17I-16013\_M13\_UV-metric pKa\_report.pdf |
| SM13\* | 5.68 |  |  | UV-metric pKa with cosolvent | 17J-04005 | SM13\_17J-04005\_M13\_UV-metric psKa\_report.pdf |
| SM14 | 2.59 | 5.31 |  | UV-metric pKa | 17I-29002 | SM14\_17I-29002\_M15\_UV-metric pKa\_report.pdf |
| SM14 | 2.57 | 5.29 |  | UV-metric pKa | 17I-29003 | SM14\_17I-29003\_M15\_UV-metric pKa\_report.pdf |
| SM14 | 2.57 | 5.29 |  | UV-metric pKa | 17I-29004 | SM14\_17I-29004\_M15\_UV-metric pKa\_report.pdf |
| SM14\* | 2.56 | 5.29 |  | UV-metric pKa with cosolvent | 17J-12008 | SM14\_17J-12008\_M15\_UV-metric psKa\_report.pdf |

\* These UV-metric pKa measurement with cosolvent were not used in the calculation of pKa mean and SEM in the experimental data reported for SAMPL6 pKa challenge because replicate water based experiments also existed for these molecules.

**Table SI 4**. Continued.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Molecule ID** | **pKa1** | **pKa2** | **pKa3** | **Assay Type** | **Experiment ID** | **Experiment Report** |
| SM15 | 4.71 | 8.96 |  | UV-metric pKa | 17K-10009 | SM15\_17K-10009\_M16\_UV-metric pKa\_report.pdf |
| SM15 | 4.7 | 8.94 |  | UV-metric pKa | 17K-10010 | SM15\_17K-10010\_M16\_UV-metric pKa\_report.pdf |
| SM15 | 4.69 | 8.92 |  | UV-metric pKa | 17K-10011 | SM15\_17K-10011\_M16\_UV-metric pKa\_report.pdf |
| SM15\* | 4.67 | 9.01 |  | UV-metric pKa with cosolvent | 17J-12009 | SM15\_17J-12009\_M16\_UV-metric psKa\_report.pdf |
| SM16 | 5.37 | 10.64 |  | UV-metric pKa | 17J-03025 | SM16\_17J-03025\_M18\_UV-metric pKa\_report.pdf |
| SM16 | 5.37 | 10.65 |  | UV-metric pKa | 17J-03026 | SM16\_17J-03026\_M18\_UV-metric pKa\_report.pdf |
| SM16 | 5.38 | 10.65 |  | UV-metric pKa | 17J-03027 | SM16\_17J-03027\_M18\_UV-metric pKa\_report.pdf |
| SM16\* | 5.35 | 11.01 |  | UV-metric pKa with cosolvent | 17J-12010 | SM16\_17J-12010\_M18\_UV-metric psKa\_report.pdf |
| SM17 | 3.15 |  |  | UV-metric pKa | 17J-02024 | SM17\_17J-02024\_M19\_UV-metric pKa\_report.pdf |
| SM17 | 3.16 |  |  | UV-metric pKa | 17J-02025 | SM17\_17J-02025\_M19\_UV-metric pKa\_report.pdf |
| SM17 | 3.17 |  |  | UV-metric pKa | 17J-02026 | SM17\_17J-02026\_M19\_UV-metric pKa\_report.pdf |
| SM17\* | 3.2 |  |  | UV-metric pKa with cosolvent | 17J-12011 | SM17\_17J-12011\_M19\_UV-metric psKa\_report.pdf |
| Pyridoxine HCl | 4.81 | 8.85 |  | UV-metric pKa | 18E-22003 | 18E-22003\_Pyridoxine HCl\_UV-metric pKa\_0417936-0002.pdf |
| Pyridoxine HCl | 4.83 | 8.87 |  | UV-metric pKa | 18E-22004 | 18E-22004\_Pyridoxine HCl\_UV-metric pKa\_0417936-0002.pdf |
| Pyridoxine HCl | 4.84 | 8.86 |  | UV-metric pKa | 18E-22010 | 18E-22010\_Pyridoxine HCl\_UV-metric pKa\_0417936-0002.pdf |
| Pyridoxine HCl | 4.85 | 8.79 |  | UV-metric pKa with cosolvent | 18E-22011 | 18E-22011\_Pyridoxine HCl\_UV-metric psKa\_0417936-0002.pdf |
| Pyridoxine HCl | 4.85 | 8.79 |  | UV-metric pKa with cosolvent | 18E-22012 | 18E-22012\_Pyridoxine HCl\_UV-metric psKa\_0417936-0002.pdf |
| Pyridoxine HCl | 4.86 | 8.77 |  | UV-metric pKa with cosolvent | 18E-22013 | 18E-22013\_Pyridoxine HCl\_UV-metric psKa\_0417936-0002.pdf |

\* These UV-metric pKa measurement with cosolvent were not used in the calculation of pKa mean and SEM in the experimental data reported for SAMPL6 pKa challenge because replicate water based experiments also existed for these molecules.

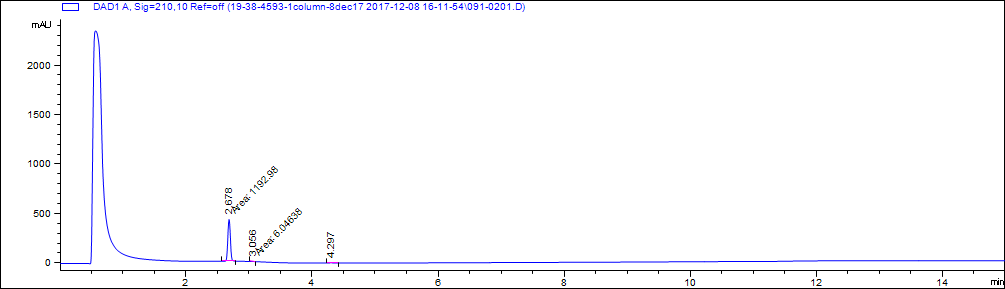
**Table SI 5.** pKa mean and SEM results of with (cosolvent) and without cosolvent (water) replicate experiment. pKa values without SEM reported were measured with 1 replicate. pKa values reported with SEM were measured in triplicates.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Molecule ID** | **pKa ID** | **pKa mean water** | **pKa SEM water** | **pKa mean cosolvent** | **pKa SEM cosolvent** |
| SM01 | SM01\_pKa1 | 9.53 | 0.01 | 9.71 |  |
| SM04 | SM04\_pKa1 | 6.02 | 0.01 | 5.97 |  |
| SM06 | SM06\_pKa1 | 3.03 | 0.04 | 3.44 | 0.02 |
| SM07 | SM07\_pKa1 | 6.08 | 0.01 | 5.96 |  |
| SM08 | SM08\_pKa1 | 4.22 | 0.01 | 4.47 |  |
| SM11 | SM11\_pKa1 | 3.89 | 0.01 | 3.74 |  |
| SM12 | SM12\_pKa1 | 5.28 | 0.01 | 5.16 |  |
| SM13 | SM13\_pKa1 | 5.77 | 0.01 | 5.68 |  |
| SM14 | SM14\_pKa1 | 2.58 | 0.01 | 2.56 |  |
| SM14 | SM14\_pKa2 | 5.3 | 0.01 | 5.29 |  |
| SM15 | SM15\_pKa1 | 4.7 | 0.01 | 4.67 |  |
| SM15 | SM15\_pKa2 | 8.94 | 0.01 | 9.01 |  |
| SM16 | SM16\_pKa1 | 5.37 | 0.01 | 5.35 |  |
| SM16 | SM16\_pKa2 | 10.65 | 0.01 | 11.01 |  |
| SM17 | SM17\_pKa1 | 3.16 | 0.01 | 3.2 |  |
| Pyridoxine HCl | Pyridoxine HCl\_pKa1 | 4.83 | 0.01 | 4.85 | 0.01 |
| Pyridoxine HCl | Pyridoxine HCl\_pKa2 | 8.86 | 0.01 | 8.78 | 0.01 |

**Table SI 6.** Summary of LC-MS purity results of pKa challenge compounds. LC-MS purity was calculated as percentage of area under the curve relative to all the peaks observed.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **SAMPL6 molecule ID** | **Supplier reported purity (%)** | **LC-MS purity (Area%)** | **Supplier reported molecular weight (g/mol)** | **eMolecules reported molecular weight (g/mol)** | **LC/MS molecular weight (g/mol)** | **eMolecules ID** |
| SM01 | 95 | 98.516 | 217.23 | 217.221 | 217.221 | 6679830 |
| SM02 |  | 97.62 |  | 289.255 | 289.255 | 1327907 |
| SM03 |  | 90.798 |  | 301.387 | 301.387 | 1228629 |
| SM04 |  | 99.884 |  | 269.729 | 269.729 | 30719859 |
| SM05 |  | 98.749 |  | 304.771 | 304.771 | 18908671 |
| SM06 |  | 98.913 |  | 328.163 | 328.163 | 18893169 |
| SM07 | 95 | 99.384 | 235.284 | 235.284 | 235.284 | 1327878 |
| SM08 | 95 | 99.242 | 293.317 | 293.317 | 293.317 | 1367649 |
| SM09 | 95 | 98.969 | 287.744 | 287.744 | 287.744 | 1865544 |
| SM10 | 95 | 96.98 | 311.357 | 311.358 | 311.358 | 23354217 |
| SM11 | 90 | 98.022 |  | 211.223 | 211.233 | 719540 |
| SM12 |  | 98.134 |  | 292.163 | 292.163 | 1859493 |
| SM13 |  | 97.512 |  | 295.336 | 295.336 | 5828805 |
| SM14 |  | 96.484 |  | 209.247 | 209.247 | 31653344 |
| SM15 |  | 98.698 |  | 210.231 | 210.231 | 37095168 |
| SM16 |  | 97.733 | 267.114 | 267.111 | 267.111 | 1284691 |
| SM17 |  | 96.994 |  | 269.322 | 269.322 | 1444229 |
| SM18 |  | 92.601 |  | 426.439 | 426.439 | 18897105 |
| SM19 |  | 94.464 |  | 381.276 | 381.276 | 3365457 |
| SM20 |  | 97.176 | 380.25 | 380.245 | 380.245 | 46568819 |
| SM21 |  | 96.791 | 438.1 | 438.092 | 438.092 | 1574612 |
| SM22 |  | 86.975 | 396.95 | 396.951 | 396.951 | 536848 |
| SM23 |  | 96.011 | 420.47 | 420.461 | 420.461 | 4375254 |
| SM24 |  | 99.331 | 391.43 | 391.42 | 391.42 | 1415746 |

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | area | Area% |
| 1 | 2.7 | 1193 | 98.5 |
| 2 | 3.1 | 6 | 0.5 |
| 3 | 4.3 | 11.9 | 1.0 |

**B**

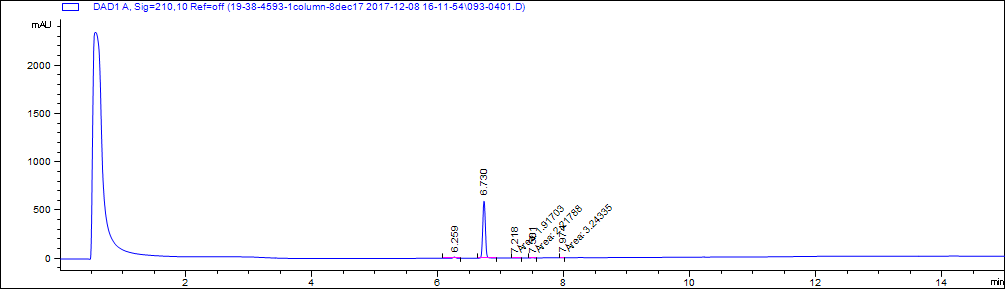


M+1

2M+Na

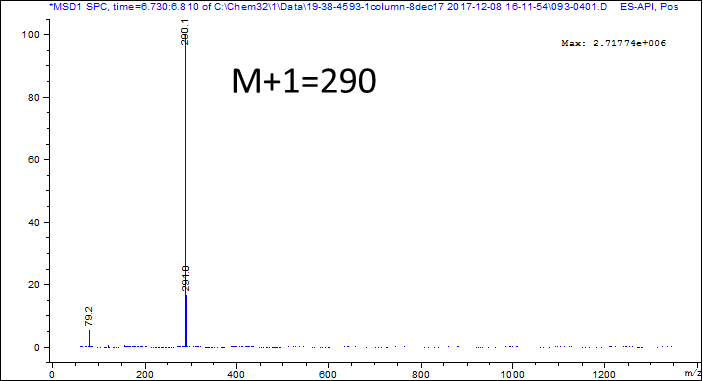
**Figure SI 1.** LC-MS chromatograms for SM01. Molecular weight estimated from mass spectrometry is 217.221 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



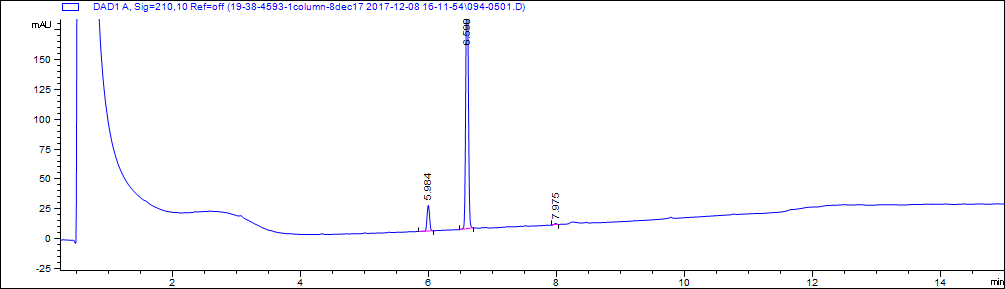
|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 6.3 | 34.9 | 2.0 |
| 2 | 6.7 | 1733.8 | 97.6 |
| 3 | 7.2 | 1.9 | 0.1 |
| 4 | 7.5 | 2.2 | 0.1 |
| 5 | 8.0 | 3.2 | 0.2 |

**B**



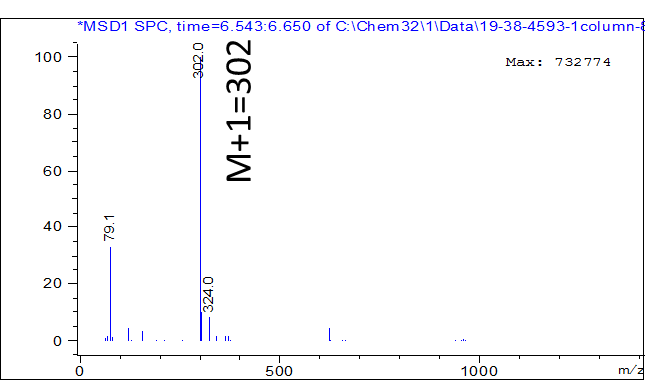
**Figure SI 2.** LC-MS chromatograms for SM02. Molecular weight estimated from mass spectrometry is 289.255 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



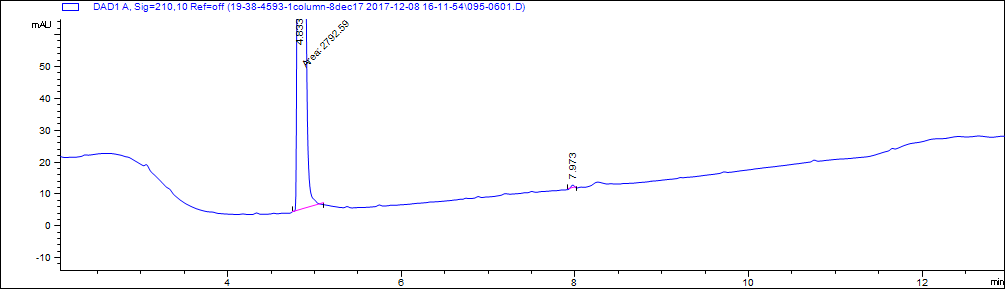
|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 6.0 | 58.2 | 8.7 |
| 2 | 6.6 | 606.7 | 90.8 |
| 3 | 8.0 | 3.3 | 0.5 |

**B**



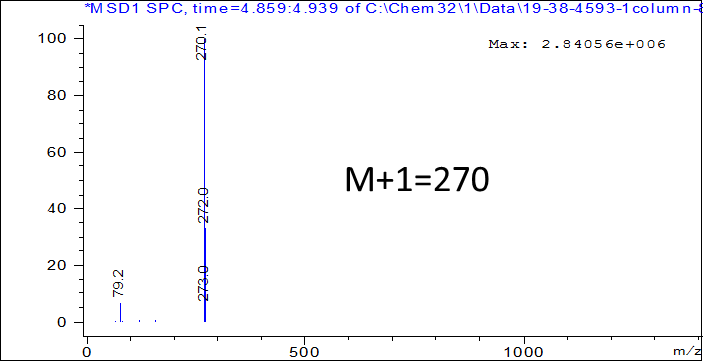
**Figure SI 3.** LC-MS chromatograms for SM03. Molecular weight estimated from mass spectrometry is 301.387 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



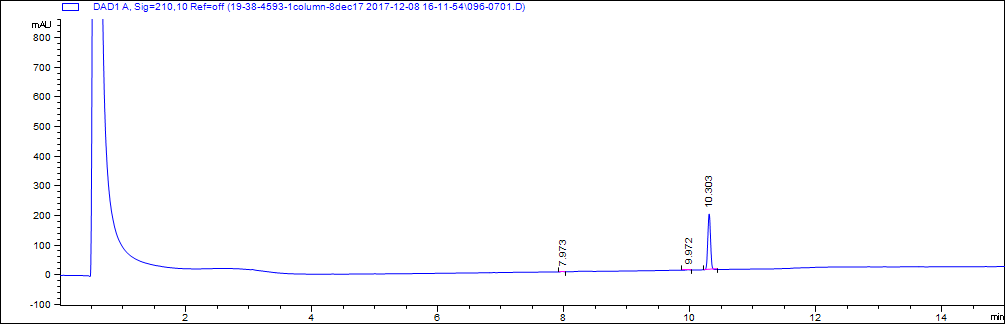
|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 4.8 | 2792.6 | 99.9 |
| 2 | 8.0 | 3.2 | 0.1 |

**B**



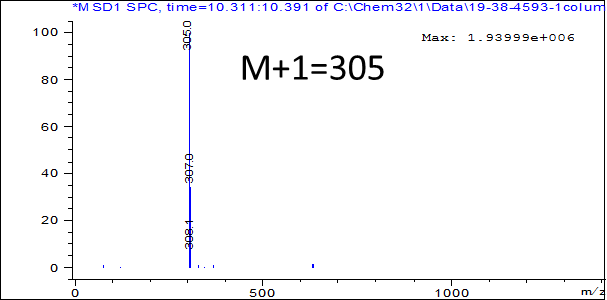
**Figure SI 4.** LC-MS chromatograms for SM04. Molecular weight estimated from mass spectrometry is 269.729 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



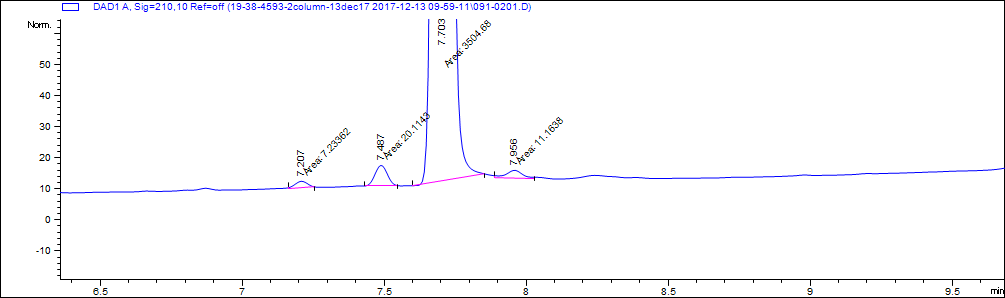
|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 8.0 | 3.3 | 0.6 |
| 2 | 10.0 | 4 | 0.7 |
| 3 | 10.3 | 569.7 | 98.7 |

**B**



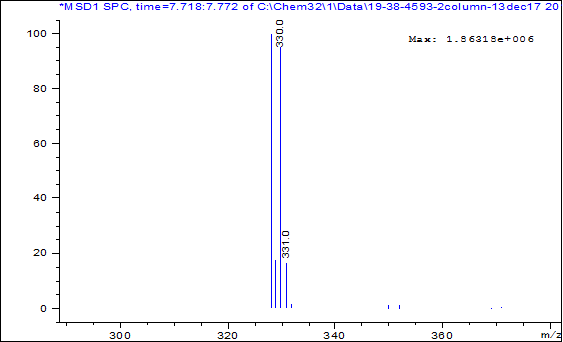
**Figure SI 5.** LC-MS chromatograms for SM05. Molecular weight estimated from mass spectrometry is 304.771 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 7.2 | 7.2 | 0.2 |
| 2 | 7.5 | 20.1 | 0.6 |
| 3 | 7.7 | 3504.7 | 98.9 |
| 4 | 8.0 | 11.2 | 0.3 |

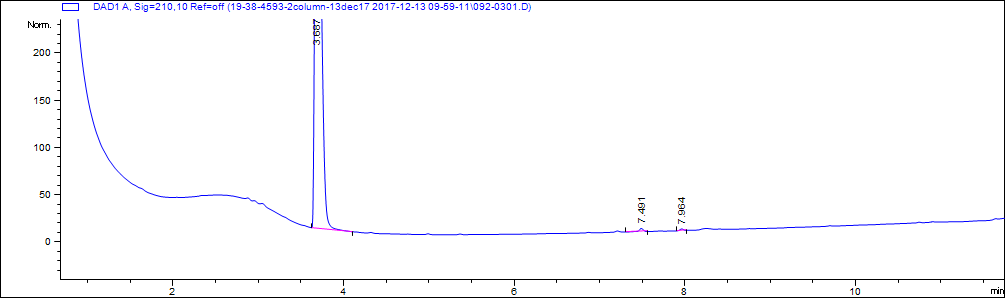
**B**



M/Z=328 and 330

**Figure SI 6.** LC-MS chromatograms for SM06. Molecular weight estimated from mass spectrometry is 328.163 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 3.7 | 2822.4 | 99.4 |
| 2 | 7.5 | 11.3 | 0.4 |
| 3 | 8.0 | 6.2 | 0.2 |

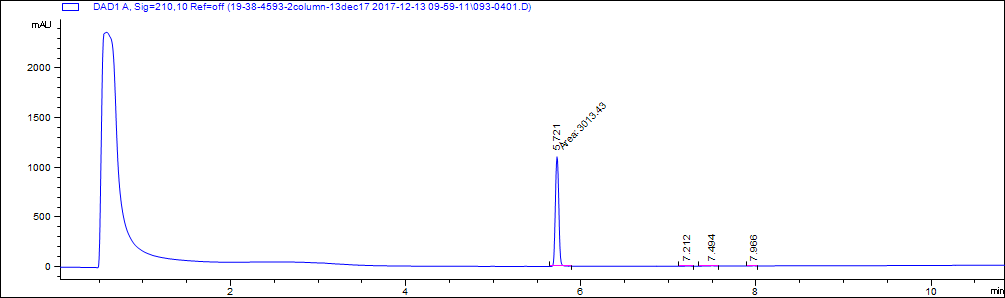
**B**



M+1=236

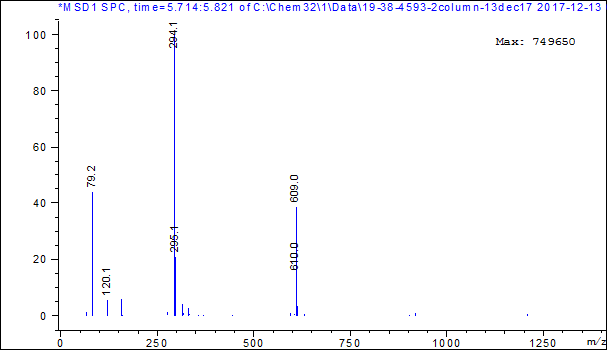
**Figure SI 7.** LC-MS chromatograms for SM07. Molecular weight estimated from mass spectrometry is 235.284 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 5.7 | 3013.4 | 99.2 |
| 2 | 7.2 | 5 | 0.2 |
| 3 | 7.5 | 11.1 | 0.4 |
| 4 | 8.0 | 6.9 | 0.2 |

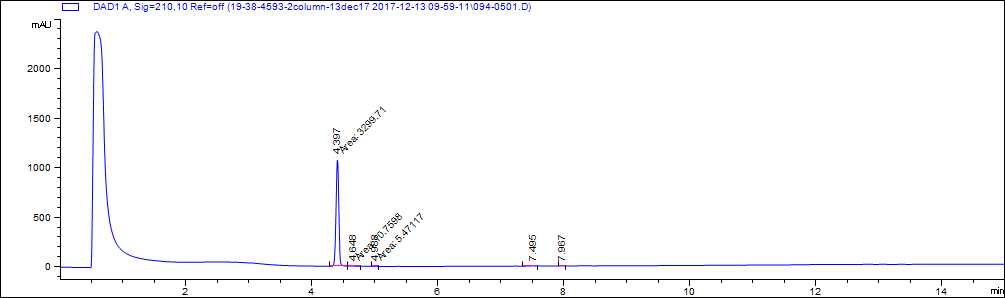
**B**



M+1=294

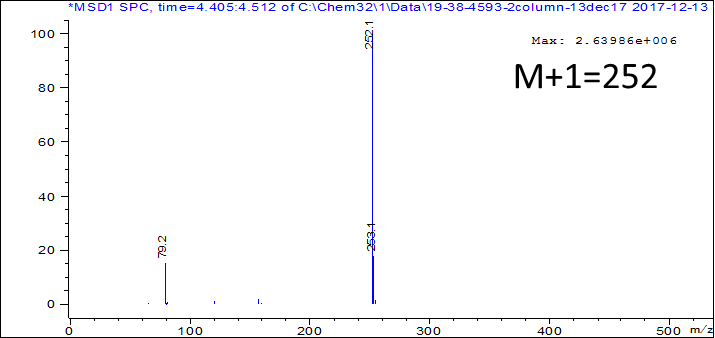
**Figure SI 8.** LC-MS chromatograms for SM08. Molecular weight estimated from mass spectrometry is 293.317 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



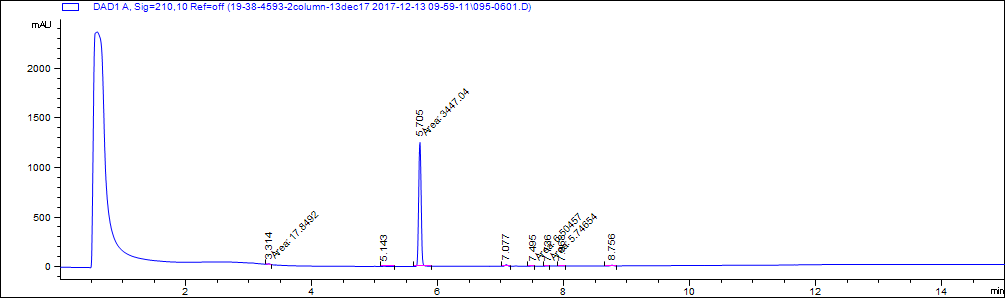
|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 4.4 | 3299.7 | 99.0 |
| 2 | 4.6 | 10.8 | 0.3 |
| 3 | 5.0 | 5.5 | 0.2 |
| 4 | 7.5 | 12.6 | 0.4 |
| 5 | 8.0 | 5.7 | 0.2 |

**B**



**Figure SI 9.** LC-MS chromatograms for SM09. Molecular weight estimated from mass spectrometry is 287.744 g/mol (m/z = 251.244). (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 2 | 5.1 | 5.2 | 0.1 |
| 3 | 5.7 | 3447 | 97.0 |
| 4 | 7.1 | 43.7 | 1.2 |
| 5 | 7.5 | 6.5 | 0.2 |
| 6 | 7.7 | 5.7 | 0.2 |
| 7 | 8.0 | 6.8 | 0.2 |
| 8 | 8.8 | 21.5 | 0.6 |

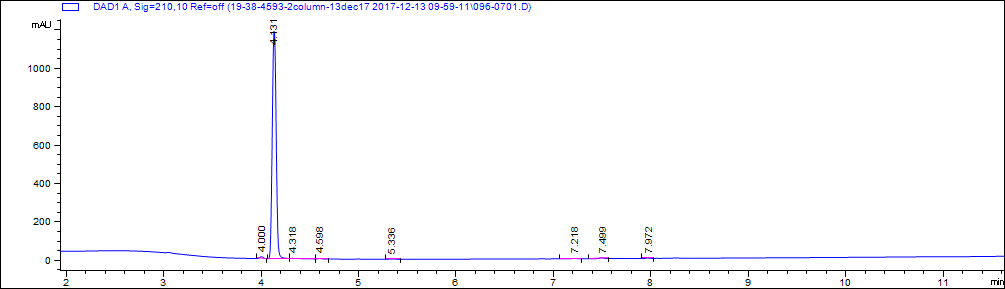
**B**



M+1=312

**Figure SI 10.** LC-MS chromatograms for SM10. Molecular weight estimated from mass spectrometry is 311.358 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 4 | 25.1 | 0.8 |
| 2 | 4.1 | 3281.8 | 98.0 |
| 3 | 4.6 | 5.6 | 0.2 |
| 4 | 5.3 | 6.5 | 0.2 |
| 5 | 7.2 | 7.1 | 0.2 |
| 6 | 7.5 | 15.2 | 0.5 |
| 7 | 8.0 | 6.8 | 0.2 |

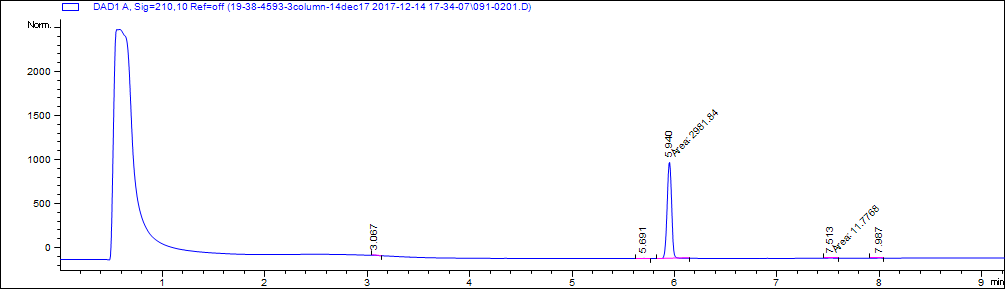
**B**



M+1=212

**Figure SI 11.** LC-MS chromatograms for SM11. Molecular weight estimated from mass spectrometry is 211.233 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 3.1 | 5.1 | 0.2 |
| 2 | 5.7 | 7.3 | 0.2 |
| 3 | 5.9 | 2981.8 | 98.1 |
| 4 | 7.5 | 11.8 | 0.4 |
| 5 | 8.0 | 6.7 | 0.2 |
| 6 | 12.3 | 25.7 | 0.8 |

**B**

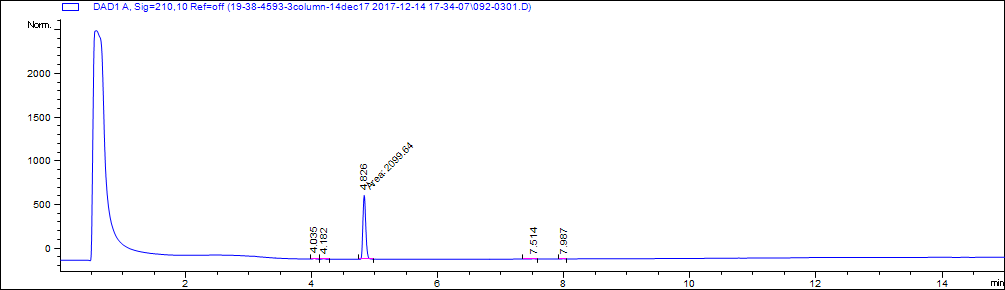


M-1=254

M+1=256

**Figure SI 12.** LC-MS chromatograms for SM12. Molecular weight estimated from mass spectrometry is 292.163 g/mol (m/z = 255). (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 4.0 | 18.1 | 0.8 |
| 2 | 4.2 | 15.4 | 0.7 |
| 3 | 4.8 | 2099.6 | 97.5 |
| 4 | 7.5 | 13.5 | 0.6 |
| 5 | 8.0 | 6.5 | 0.3 |

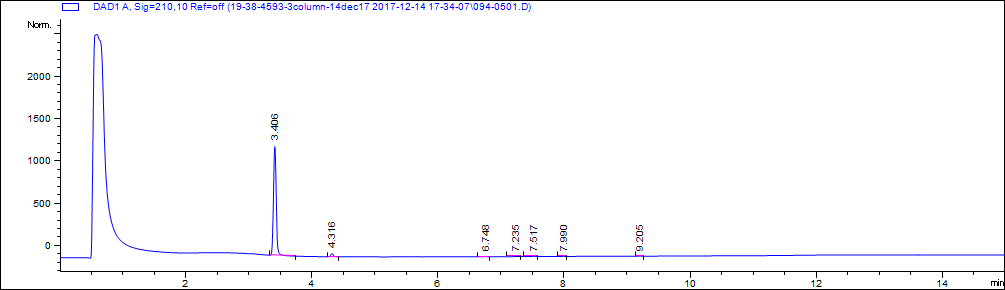
**B**



M+1=296

**Figure SI 13.** LC-MS chromatograms for SM13. Molecular weight estimated from mass spectrometry is 295.336 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 3.4 | 3561.5 | 96.5 |
| 2 | 4.3 | 90.1 | 2.4 |
| 3 | 6.7 | 7.5 | 0.2 |
| 4 | 7.2 | 5.9 | 0.2 |
| 5 | 7.5 | 14.3 | 0.4 |
| 6 | 8.0 | 6.6 | 0.2 |
| 7 | 9.2 | 5.3 | 0.1 |

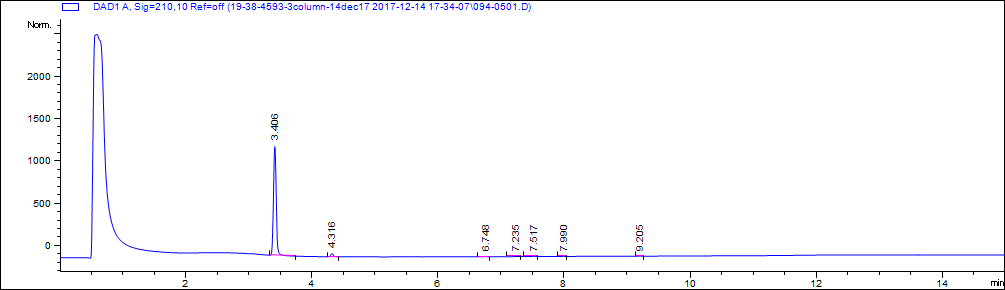
**B**



M+1=210

**Figure SI 14.** LC-MS chromatograms for SM14. Molecular weight estimated from mass spectrometry is 209.247 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



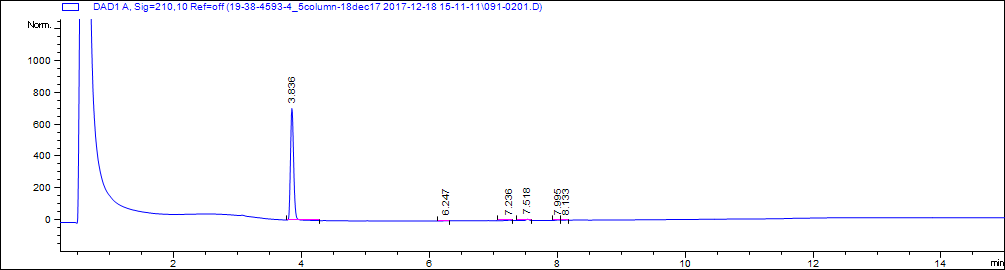
|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 4.4 | 2822.7 | 98.7 |
| 2 | 5.5 | 15.6 | 0.5 |
| 3 | 5.8 | 5.5 | 0.9 |
| 4 | 7.5 | 9.4 | 1.2 |
| 5 | 8.0 | 6.7 | 1.0 |

**B**



**Figure SI 15.** LC-MS chromatograms for SM15. Molecular weight estimated from mass spectrometry is 210.231 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 3.8 | 2375.8 | 97.7 |
| 2 | 6.2 | 5.3 | 0.2 |
| 3 | 7.2 | 9.6 | 0.4 |
| 4 | 7.5 | 23.9 | 1.0 |
| 5 | 8.0 | 10.2 | 0.4 |
| 6 | 8.1 | 6.1 | 0.3 |

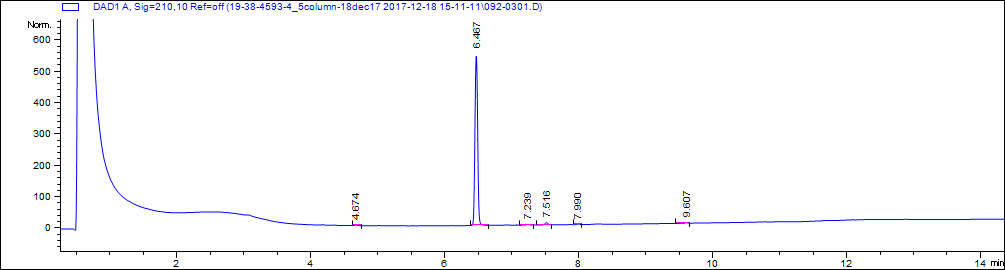
**B**



M/Z=267,269

**Figure SI 16.** LC-MS chromatograms for SM16. Molecular weight estimated from mass spectrometry is 267.111 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 4.7 | 5.3 | 0.3 |
| 2 | 6.5 | 1541.6 | 97.0 |
| 3 | 7.2 | 7.3 | 0.5 |
| 4 | 7.5 | 22.5 | 1.4 |
| 5 | 8.0 | 6.4 | 0.4 |
| 6 | 9.6 | 6.3 | 0.4 |

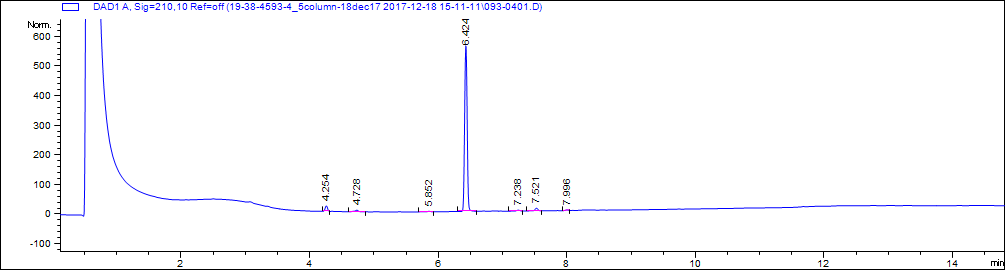
**B**



M+1=270

**Figure SI 17.** LC-MS chromatograms for SM17. Molecular weight estimated from mass spectrometry is 269.322 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 4.3 | 47.8 | 2.9 |
| 2 | 4.7 | 22.5 | 1.4 |
| 3 | 5.9 | 5.7 | 0.3 |
| 4 | 6.4 | 1538.3 | 92.6 |
| 5 | 7.2 | 10.7 | 0.6 |
| 6 | 7.5 | 28.5 | 1.7 |
| 7 | 8.0 | 7.8 | 0.5 |

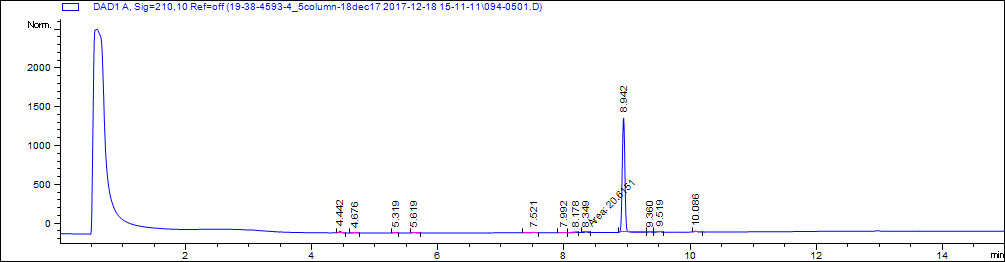
**B**



M+1=427

**Figure SI 18.** LC-MS chromatograms for SM18. Molecular weight estimated from mass spectrometry is 426.439 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 4.4 | 51.6 | 1.3 |
| 2 | 4.7 | 5.7 | 0.1 |
| 3 | 5.3 | 14.3 | 0.3 |
| 4 | 5.6 | 17.8 | 0.4 |
| 5 | 7.5 | 17.4 | 0.4 |
| 6 | 8.0 | 8.2 | 0.2 |
| 7 | 8.2 | 14.1 | 0.3 |
| 8 | 8.3 | 20.6 | 0.5 |
| 9 | 8.9 | 3893.8 | 94.5 |
| 10 | 9.4 | 15.4 | 0.4 |
| 11 | 9.5 | 26.1 | 0.6 |
| 12 | 10.1 | 37.1 | 0.9 |

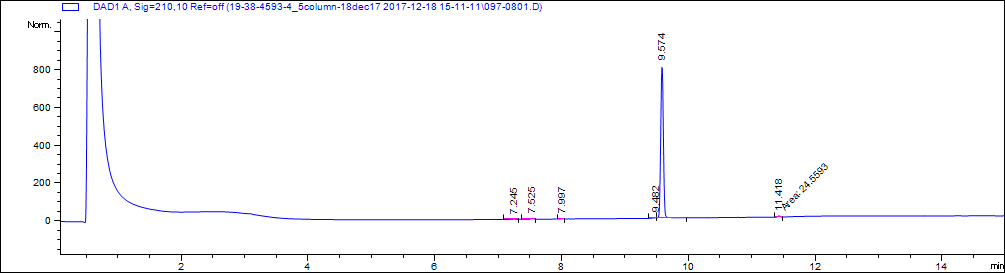
**B**



M/Z=381 and 383

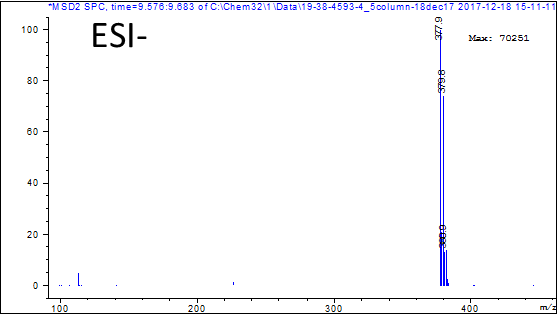
**Figure SI 19.** LC-MS chromatograms for SM19. Molecular weight estimated from mass spectrometry is 381.276 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 7.2 | 5.6 | 0.2 |
| 2 | 7.5 | 16.4 | 0.7 |
| 3 | 8.0 | 7.0 | 0.3 |
| 4 | 9.5 | 15.4 | 0.6 |
| 5 | 9.6 | 2376.4 | 97.2 |
| 6 | 11.4 | 24.6 | 1.0 |

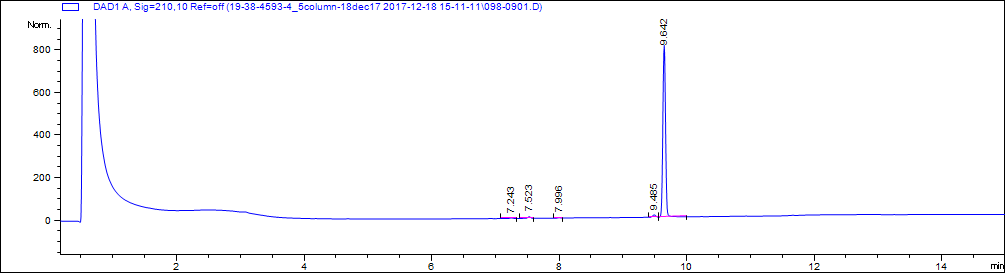
**B**



M/Z=378 and 380

**Figure SI 20.** LC-MS chromatograms for SM20. Molecular weight estimated from mass spectrometry is 380.245 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 7.2 | 9.1 | 0.4 |
| 2 | 7.5 | 23.6 | 1.0 |
| 3 | 8.0 | 6.4 | 0.2 |
| 4 | 9.5 | 39.9 | 1.6 |
| 5 | 9.6 | 2381.2 | 96.8 |

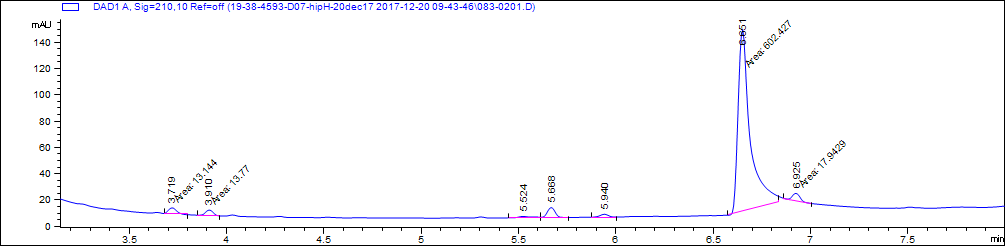
**B**



M+1=439

**Figure SI 21.** LC-MS chromatograms for SM21. Molecular weight estimated from mass spectrometry is 438.092 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 3.7 | 13.1 | 1.9 |
| 2 | 3.9 | 13.8 | 2.0 |
| 3 | 5.5 | 8.0 | 1.2 |
| 4 | 5.7 | 24.9 | 3.6 |
| 5 | 5.9 | 12.4 | 1.8 |
| 6 | 6.7 | 602.4 | 87.0 |
| 7 | 6.9 | 17.9 | 2.6 |

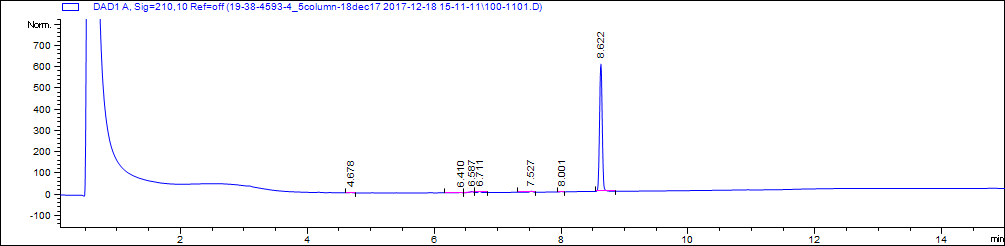
**B**



M+1=398

**Figure SI 22.** LC-MS chromatograms for SM22. Molecular weight estimated from mass spectrometry is 396.951 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 4.7 | 5.4 | 0.3 |
| 2 | 6.4 | 5.6 | 0.3 |
| 3 | 6.6 | 22.7 | 1.2 |
| 4 | 6.7 | 18.1 | 1.0 |
| 5 | 7.5 | 16.1 | 0.9 |
| 6 | 8.0 | 5.6 | 0.3 |
| 7 | 8.6 | 1770.7 | 96.0 |

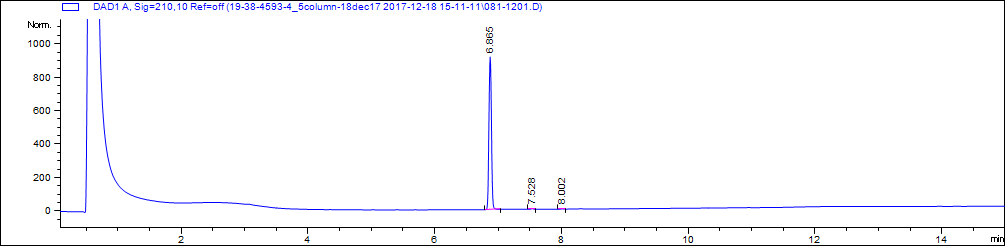
**B**



M+1=421

**Figure SI 23.** LC-MS chromatograms for SM23. Molecular weight estimated from mass spectrometry is 420.461 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.

**A**



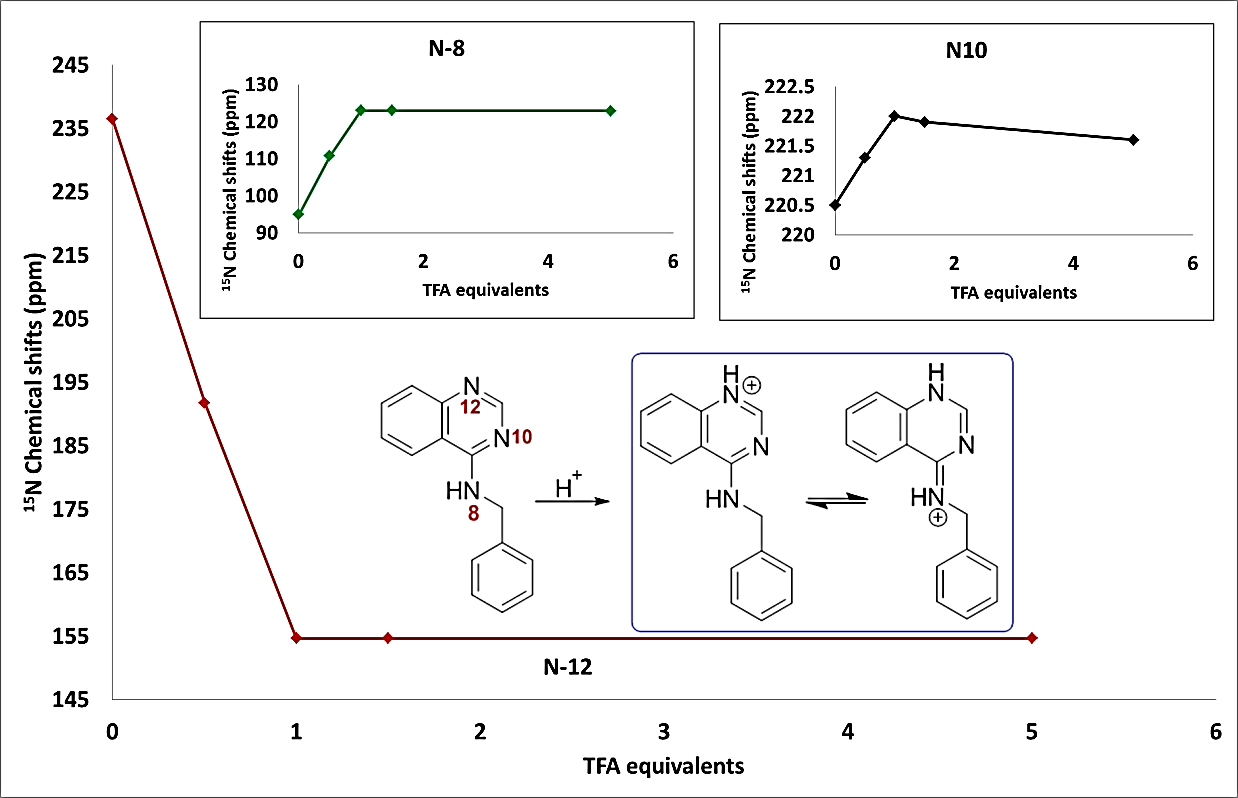
|  |  |  |  |
| --- | --- | --- | --- |
| # | Retention time (minutes) | Area | Area% |
| 1 | 6.9 | 2511.1 | 99.3 |
| 2 | 7.5 | 11.0 | 0.4 |
| 3 | 8.0 | 5.9 | 0.2 |

**B**

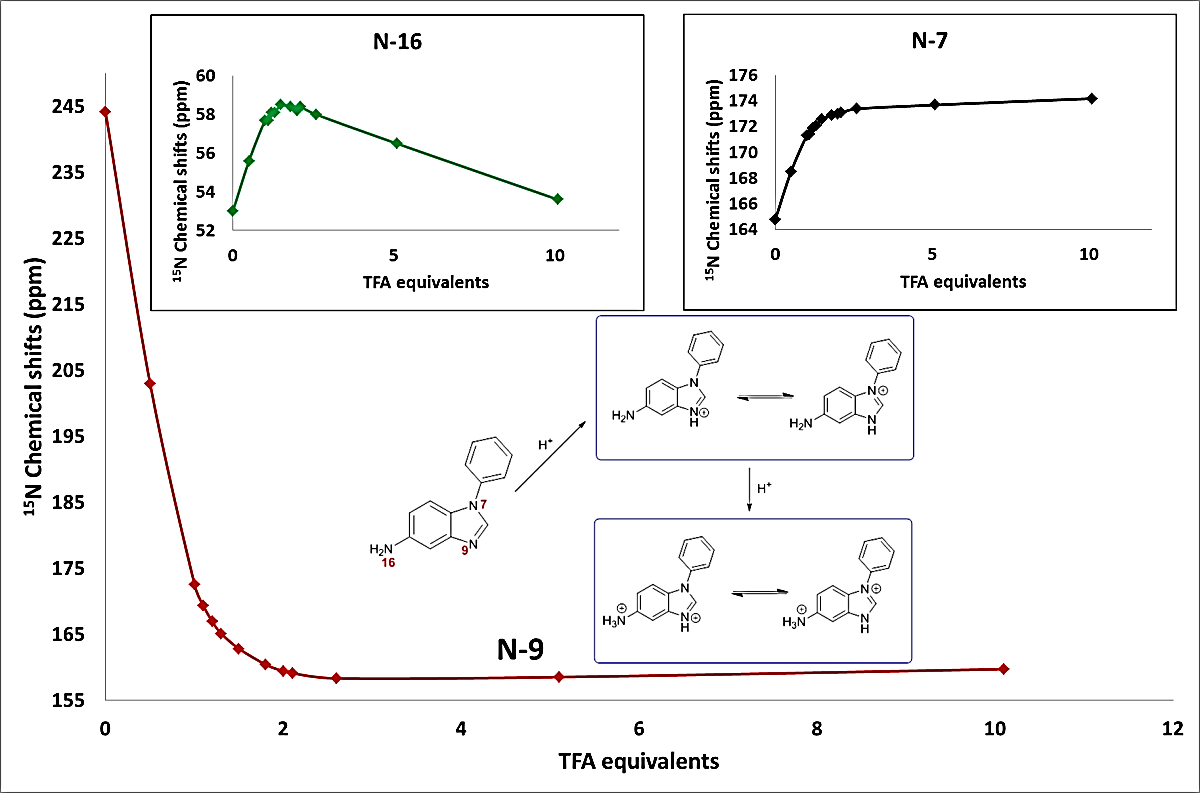


M+1=392

**Figure SI 24.** LC-MS chromatograms for SM24. Molecular weight estimated from mass spectrometry is 391.42 g/mol. (A) LC chromatogram and table of peak properties. %Area describes relative area under the curve of selected peaks and interpreted as an estimate of purity. (B) MS detection.



**Figure SI 25.** The plot of 15N chemical shifts of N-12 of SM07 vs titrated TFA-*d* equivalents, showing the mono-protonation of N-12 as evidenced by its large upfield chemical shifts change. Insets show the plots for the 15N chemical shifts of N-8 and N-10 of SM07 *vs* TFA-*d* equivalents (showing electronic effects due to protonation of N-12) and the protonation pathway and resonance structures for the protonated SM07.



**Figure SI 26.** Plot of 15N chemical shifts for the N-9 nuclide of SM14 vs TFA-*d* additions of 0.0, 0.5, 1.0, 1.1, 1.2, 1.3, 1.5, 1.8, 2.0, 2.1, 2.6, 5.1 and 10.1 equivalents, showing the first protonation at N-9 – with a large upfield chemical shifts change of 71.6 ppm between 0 and 1 equivalents of TFA-*d*. Insets show plots of N-7 and N-16 showing the downfield chemical shift, between 0 and 2 equivalents of TFA-*d*, due to electronic effect from the protonation of N-9. N-16 also exhibited a small upfield chemical shift change of 4.4 ppm between 2.5 and 10 equivalents of TFA-*d*, which indicated this to be the second site of protonation. The protonation pathway and subsequent resonance structures for SM14 is shown inset.