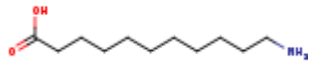
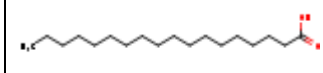
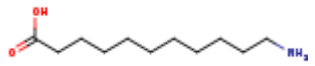
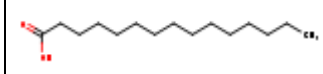
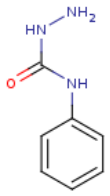
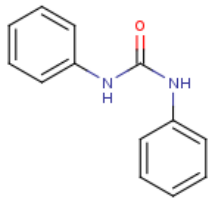
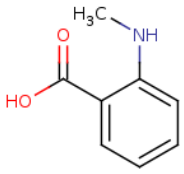
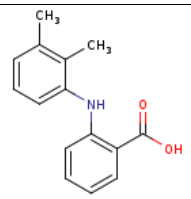
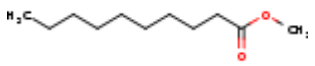
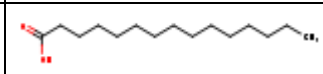
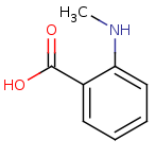
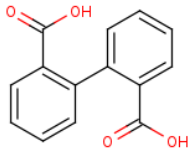
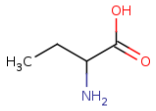
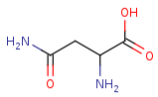
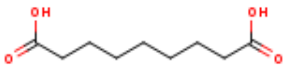
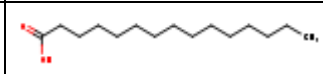
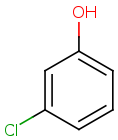
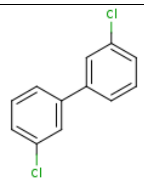
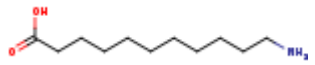
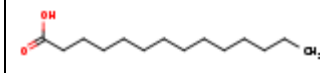
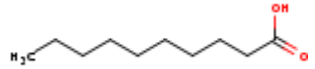
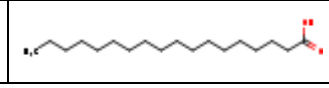
Supplemental Information

Comprehensive Results

Table S1 shows the compounds with the highest GSE errors in the solubility technique comparison, with corresponding ALOGPS logS errors. Table S2 and Table S3 show the comprehensive list of all functional group substitutions and endpoints that were found in the MMP list. These are being reported as they may be of use in ADMET optimisation work.

Table S1 — The Compounds with Absolute Error ≥ 2 log units in the GSE ΔlogS Prediction of the ALOGPS training set

From	To	GSE ΔlogS Error	ALOGPS ΔlogS Error
		–3.98	–0.28
		–3.81	–0.88
		–3.26	–1.07
		–3.01	–1.89
		–2.66	–2.18
		–2.63	–2.42
		2.6	0.61
		–2.36	–1.52
		2.34	0.42
		–2.26	0.87
		–2.23	–1.13

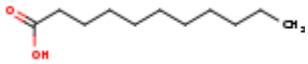
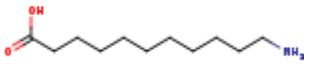
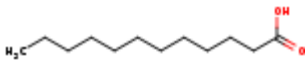
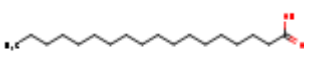
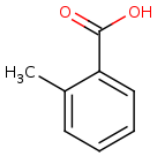
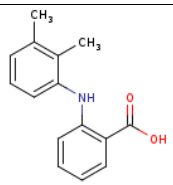
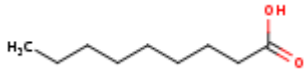
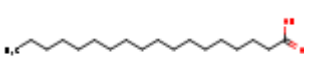
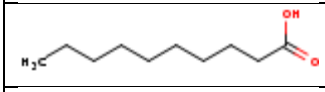
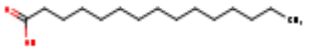
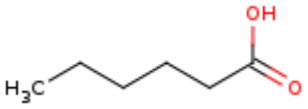
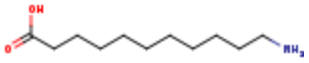
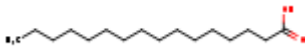
		2.17	–0.31
		–2.09	–1.27
		–2.09	–0.88
		–2.06	–1.73
		–2.06	–0.95
		2.02	–1.96
		–2.00	–1.65

Table S2 — The results of functional group substitution analysis for all compounds

From	To	Number of Samples	Mean ΔT (°C)	Standard Error of Mean (°C)	P-value
Sulfonamides	Sulfonic acids	39	90	±17	<0.0001
Phosphonic acids	Phosphonic acid esters	37	–85	±8.6	<0.0001
Thiocarboxylic acid amides	Thiocarboxylic acid esters	22	–73	±7.4	<0.0001
Dialkylethers	Carboxylic acid secondary amides	20	72	±9.9	<0.0001
Carboxylic acid primary amides	Carboxylic acid esters	176	–68	±3.8	<0.0001
Carboxylic acids	Carboxylic acid esters	7,056	–65	±0.6	<0.0001
Alkynes	Carboxylic acid primary amides	26	64	±7.1	<0.0001
Diarylethers	Carboxylic acid secondary amides	36	57	±12	<0.0001
Aryl iodides	Carboxylic acid esters	35	–57	±11	<0.0001
Sulfones	Dialkylethers	92	–57	±5.1	<0.0001
Sulfonyl halides	Sulfonamides	108	57	±5.1	<0.0001
Enols	Enolethers	24	–53	±9.4	<0.0001
Carboxylic acid primary amides	Nitro compounds	20	–53	±6.2	<0.0001
Carboxylic acid primary amides	Carboxylic acid tertiary amides	347	–52	±2.8	<0.0001
Carboxylic acid tertiary amides	Carboxylic acids	30	52	±8.8	<0.0001
Oximes	Carboxylic acid esters	21	–49	±9.6	<0.0001
Alkynes	Carboxylic acids	43	47	±8.5	<0.0001
Carboxylic acids	Sulfones	28	–46	±7.1	<0.0001
Nitriles	Dialkylethers	67	–46	±6.0	<0.0001
Nitriles	Carboxylic acid primary amides	124	46	±4.3	<0.0001
Oximes	Oxime ethers	85	–44	±6.3	<0.0001
Sulfones	Carboxylic acid derivatives	21	–42	±5.7	<0.0001
Nitriles	Thiocarboxylic acid amides	41	42	±7.5	<0.0001
Carboxylic acids	Nitriles	209	–42	±3.7	<0.0001
Sulfones	Oxime ethers	36	–42	±6.1	<0.0001
Carboxylic acids	Hemiaminals	32	–41	±10	<0.0001
Sulfonamides	Sulfonic acid esters	62	–40	±5.1	<0.0001

Carboxylic acid unsubstituted imides	Carboxylic acid substituted imides	84	–40	±5.6	<0.0001
Carboxylic acid derivatives	Oximes	31	39	±8.7	<0.0001
Carboxylic acids	Nitro compounds	137	–39	±4.8	<0.0001
Alkynes	Dialkylethers	56	–39	±7.3	<0.0001
Alkynes	Aryl chlorides	33	38	±13	<0.005
Ureas	Carbamic acid esters (urethanes)	701	–37	±2.0	<0.0001
Carboxylic acid esters	Sulfones	75	36	±4.4	<0.0001
Carboxylic acids	Carboxylic acid secondary amides	80	–35	±7.2	<0.0001
Thioethers	Sulfones	105	34	±3.7	<0.0001
Phenols	Oxohetarenes	22	–34	±9.5	<0.0001
Hydroxamic acids	Carboxylic acid esters	93	–33	±6.1	<0.0001
Pyrazines (HS)	Pyridazines (HS)	37	32	±7.9	<0.0001
1,2-Aminoalcohols	Arenes	30	–32	±8.4	<0.0001
Nitro compounds	Carboxylic acid secondary amides	25	32	±13	<0.005
Nitriles	Oximes	63	31	±6.3	<0.0001
Carboxylic acid secondary amides	Carboxylic acid tertiary amides	1,570	–31	±1.4	<0.0001
Sulfoxides	Carboxylic acid esters	32	–31	±7.2	<0.0001
Carboxylic acid primary amides	Carboxylic acid secondary amides	1,046	–30	±1.4	<0.0001
Nitriles	Azides	22	–30	±7.5	<0.0001
Nitriles	Carboxylic acid esters	276	–30	±3.1	<0.0001
Carboxylic acid esters	Acetals / Ketals	23	–30	±8.3	<0.0001
Alkynes	Diarylethers	27	–30	±10	<0.005
Aryl fluorides	Sulfones	29	29	±11	<0.005
Nitriles	Acetals / Ketals	37	–29	±8.7	<0.0005
Aryl chlorides	Sulfones	32	28	±11	<0.01
Aryl fluorides	Nitriles	85	28	±4.2	<0.0001
Indoles	Unsaturated five-membered heterocycles with one heteroatom (LS)	33	–28	±4.9	<0.0001
Dialkylethers	Alkylarylethers	507	27	±2.1	<0.0001

Carboxylic acid esters	Carboxylic acid secondary amides	130	26	±4.1	<0.0001
Nitriles	Hydrazine derivatives	23	26	±9.3	<0.005
Carboxylic acids	Pnictogens (nitrogen group)	34	–25	±11	<0.05
Benzyl fluorides	Benzyl chlorides	43	25	±4.2	<0.0001
Dialkylethers	Arenes	26	25	±6.3	<0.0001
Aryl bromides	Nitriles	126	25	±4.3	<0.0001
Alkylarylethers	Sulfones	26	24	±9.9	<0.01
Alkynes	Carboxylic acid esters	148	–24	±4.7	<0.0001
Aryl iodides	Aryl fluorides	464	–22	±1.9	<0.0001
Aryl bromides	Sulfones	29	22	±9.3	<0.05
Alkynes	Oxime ethers	20	–22	±10	<0.05
1,3-Aminophenols	1,4-Aminophenols	48	21	±6.2	<0.0005
Nitro compounds	Carboxylic acid esters	108	–20	±4.9	<0.0001
Benzyl chlorides	Benzyl bromides	74	20	±5.2	<0.0001
Carboxylic acids	Hydroxamic acids	203	–19	±3.7	<0.0001
Aryl fluorides	Nitro compounds	94	19	±3.9	<0.0001
Dialkylethers	Acetals / Ketals	34	19	±7.2	<0.01
Secondary aromatic amines	Secondary mixed amines (aryl alkyl)	434	–19	±2.1	<0.0001
Hydrazones	Carboxylic acid derivatives	30	18	±9.8	<0.05
Nitriles	Aryl chlorides	117	–18	±3.7	<0.0001
Sulfoxides	Sulfones	560	18	±1.7	<0.0001
Dialkylthioethers	Alkylarylthioethers	164	17	±3.3	<0.0001
Nitriles	Alkylarylethers	42	–17	±6.2	<0.01
Carboxylic acids	Oximes	20	–17	±8.1	<0.05
1,2,4-Oxadiazoles (HS)	1,3,4-Oxadiazoles (HS)	168	17	±3.1	<0.0001
Aryl iodides	Nitriles	50	16	±5.6	<0.005
Pyrimidines (HS)	Pyrazines (HS)	147	–16	±3.3	<0.0001
Aryl bromides	Alkynes	42	–16	±5.8	<0.005
Alkynes	Acetals / Ketals	24	–16	±8.1	<0.01
Aryl fluorides	Alkynes	24	16	±4.9	<0.001

Alkynes	Nitriles	152	15	±4.7	<0.001
Tertiary mixed amines (aryl alkyl)	Tertiary aliphatic amines	225	–15	±3.3	<0.0001
Aryl fluorides	Carboxylic acid esters	22	–15	±7.7	<0.05
Alkyl bromides	Alkyl iodides	56	14	±4.4	<0.0001
Aryl fluorides	Aryl bromides	1,883	13	±0.8	<0.0001
Primary aromatic amines	Primary aliphatic amines	274	–13	±4.6	<0.005
Aryl chlorides	Aryl iodides	626	13	±1.5	<0.0001
Alkynes	Arenes	44	13	±6.0	<0.05
Aryl bromides	Nitro compounds	87	13	±4.8	<0.005
Diarylethers	Thioethers	66	12	±4.9	<0.005
Oxime ethers	Carboxylic acid esters	34	–12	±6.4	<0.05
Guanidines	Thioureas	96	11	±4.6	<0.01
Alkylarylethers	Diarylethers	1,012	10	±1.8	<0.0001
Nitriles	Sulfoxides	68	–9.7	±5.3	<0.05
Imidazoles (HS)	Pyrazoles (HS)	285	–8.9	±3.0	<0.005
Alkyl chlorides	Alkyl fluorides	90	–7.8	±3.8	<0.05
Tertiary alcohols	Primary alcohols	201	7.6	±3.2	<0.01
Aryl iodides	Aryl bromides	678	–7.1	±1.2	<0.0001
Aryl fluorides	Aryl chlorides	6,039	7.0	±0.5	<0.0001
Tertiary mixed amines (aryl alkyl)	Alkylarylethers	113	6.9	±3.2	<0.05
Carboxylic acid primary amides	Carboxylic acids	153	–6.9	±3.8	<0.05
Furans (HS)	Thiophenes (HS)	825	5.1	±1.4	<0.0001
Aryl chlorides	Aryl bromides	3,322	5.1	±0.6	<0.0001
1,4-Aminophenols	1,2-Aminophenols	40	1.1	±6.6	<0.0001

Table S3 — The results of functional group endpoint analysis for all compounds

Group	Number of Samples	Mean ΔT (°C)	Standard Error of Mean (°C)	P-value
Pyrazoles (HS)	21	–70	±17	<0.0001
Sulfenic acid derivatives	49	–55	±6.2	<0.0001
Thiocarboxylic acids	25	52	±7.9	<0.0001
1,3-Diphenols	22	51	±8.6	<0.0001
Alkyl iodides	21	48	±13	<0.0005
Carboxylic acids	2,449	45	±1.2	<0.0001
Oxohetarenes	187	44	±4.2	<0.0001
Secondary alcohols	190	43	±3.6	<0.0001
1,2-Diphenols	28	42	±9.9	<0.0001
Carboxylic acid primary amides	562	42	±2.5	<0.0001
Dialkylthioethers	39	–41	±6.0	<0.0001
Azo compounds	65	40	±8.1	<0.0001
Pnictogens (nitrogen group)	88	39	±7.3	<0.0001
Phenols	312	36	±2.9	<0.0001
Ureas	65	–36	±6.9	<0.0001
Hydrazine derivatives	291	35	±3.5	<0.0001
Thioenolethers	65	35	±6.4	<0.0001
a-Aminoacids	94	33	±7.6	<0.0001
1,2-Diols	183	33	±3.5	<0.0001
Unconjugated alkadienes (1,4-alkadienes)	101	–32	±3.9	<0.0001
Aminals	31	30	±11	<0.005
Sulfones	2,045	30	±1.2	<0.0001
Nitrates	49	–30	±6.6	<0.0001
Carboxylic acid amidines	68	30	±8.0	<0.0001
Alkylarylthioethers	287	–27	±2.7	<0.0001
Ketones	43	27	±7.0	<0.0005
Thiocarboxylic acid amides	96	27	±5.9	<0.0001

Oximes	255	26	±3.5	<0.0001
Ketene acetal derivatives	73	26	±4.5	<0.0001
Dialkylethers	2,108	–25	±1.0	<0.0001
Thiohemiaminals	30	–23	±8.7	<0.005
Five-membered heterocycles with one heteroatom (LS)	37	23	±8.3	<0.005
Disulfides	64	–23	±6.1	<0.0001
Carboxylic acid secondary amides	545	22	±2.3	<0.0001
Tertiary aliphatic amines	215	–22	±3.7	<0.0001
Nitro compounds	3,842	22	±0.8	<0.0001
a, ß-Unsaturated carboxylic acids	937	22	±1.8	<0.0001
Carboxylic acid tertiary amides	461	–22	±2.3	<0.0001
Imido esters	30	–20	±9.8	<0.05
Benzothiophenes	86	20	±3.4	<0.0001
Conjugated alkadienes (1,3-alkadienes)	398	19	±2.6	<0.0001
Enamines	166	19	±3.5	<0.0001
Sulfonic acid esters	31	–18	±9.0	<0.05
Nitriles	4,618	18	±0.8	<0.0001
Carboxylic acid esters	3,486	–18	±0.9	<0.0001
Hydroxamic acids	68	17	±5.7	<0.001
Tertiary alcohols	140	15	±3.2	<0.0001
Aryl iodides	626	15	±1.8	<0.0001
Secondary aliphatic amines	45	–14	±6.7	<0.05
Oxime ethers	344	–14	±2.3	<0.0001
1,3-Oxazoles (HS)	23	–13	±7.0	<0.01
Metalloids	342	–13	±2.5	<0.0001
Sulfoxides	748	13	±1.8	<0.0001
Acetals / Ketals	650	–12	±1.9	<0.0001
Aryl bromides	2,318	12	±0.8	<0.0001
Secondary mixed amines (aryl alkyl)	155	–11	±4.3	<0.01
Aldehydes	107	9.6	±5.4	<0.05

CO2 derivative (general)	347	–8.8	±2.8	<0.001
Gem-trihalides	403	–8.5	±2.5	<0.0005
Carboxylic acid derivatives	427	8.5	±2.5	<0.0005
Hemiaminals	199	–8.0	±3.4	<0.01
Primary alcohols	283	7.7	±2.7	<0.005
Arenes	4,278	7.3	±0.7	<0.0001
Primary aliphatic amines	187	–6.7	±3.5	<0.05
Aryl chlorides	3,499	6.2	±0.8	<0.0001
Tertiary mixed amines (aryl alkyl)	244	5.7	±2.9	<0.05
Alkyl bromides	153	5.7	±3.2	<0.05
Alkynes	3,148	5.0	±0.9	<0.0001
Diarylethers	690	3.9	±1.8	<0.05
Alkylarylethers	1,022	–3.5	±1.3	<0.005