

Supporting information for

A broadly applicable quantitative relative reactivity model for nucleophilic aromatic substitution (S_NAr) using simple descriptors

Jingru Lu, Irina Paci,* and David C. Leitch*

Department of Chemistry, University of Victoria, 3800 Finnerty Rd. Victoria BC, CANADA, V8P 5C2.

*ipaci@uvic.ca; dcleitch@uvic.ca.

Table of Contents

General Considerations	2
Experimental Details	3
Hammett Analyses of <i>para</i> - and <i>meta</i> -Substituted Substrates	28
Preparative Scale Synthesis of S _N Ar products	32
Computational Determination of Molecular Descriptors	44
Construction of the Multivariate Linear Regression Model	52
Cross Validation and Out-of-Sample Prediction	57
<i>LUMO</i> energy as the Electron Affinity alternatives	61
Transition State Calculations of S _N Ar by DFT	64
Experimental details for assessing outliers from external case studies (Fig. 7)	75
References	76

General Considerations

Materials

All solvents, reagents, and organic substrates were used as purchased from commercial suppliers without further purification with the following exceptions. 4-(6-chloropyrimidin-4-yl)morpholine¹, 4-chloro-6-(pyrrolidin-1-yl)pyrimidine¹, 2-chloro-5-(1,3-dioxolan-2-yl)pyridine², 2-chloro-6-isopropoxypyridine³ and 2-chloro-6-(tert-butoxy)-pyridine^{Error! Reference source not found.} were prepared using published procedures; 2-(benzyloxy)-4-chloropyridine and 4-(benzyloxy)-2-chloropyridine were prepared using the general procedures in the *Preparative Scale Synthesis of S_NAr product* section. 4-Chloro-2-methylpyrimidine was purified by dissolving the commercial material in chloroform, followed by filtration to remove insoluble impurities; purity of this material was confirmed by NMR spectroscopy. Potassium *tert*-butoxide was purified by dissolving the commercial material in anhydrous THF, followed by filtration to remove insoluble impurities. The solvent was then evaporated *in vacuo*. This purification was performed inside an MBraun glovebox under an N₂ atmosphere.

Analysis and Spectroscopy

All ¹H and ¹³C NMR spectra were recorded on either a Bruker AVANCE 300 MHz spectrometer or a Bruker AVANCE NEO 500 MHz spectrometer.

High-resolution electrospray ionization mass spectrometric analysis was performed using a Thermo Scientific Ultimate 3000 ESI-Orbitrap Exactive Plus.

LC analysis was performed on either a Shimadzu UPLC or a Waters LCMS. The Shimadzu Nexera X2 UPLC system is equipped with an autosampler, binary pump system, degassing unit, column oven and a diode-array UV/Vis detector. The chromatogram was recorded using a Raptor ARC-18 column (2.1 × 100 mm, particle size 1.8 μm) with 0.05% trifluoroacetic acid (TFA) in water as the mobile phase A and 0.05% TFA in acetonitrile as the mobile phase B. The data was processed using LabSolutions. The Waters Acquity class H UPLC system is equipped with an autosampler, quaternary pump system, column oven, a photodiode array detector and a QDa Mass Spectrometer. The chromatogram was recorded using an ACQUITY UPLC® BEH C18 column (2.1 × 50 mm, particle size 1.7 μm) with 0.1% formic acid (FA) in water as the mobile phase A and 0.1% FA in acetonitrile as the mobile phase B. The data was processed using Masslynx.

Flash column purification of all the S_NAr products in *Preparative Scale Synthesis of S_NAr products* section was performed using a Biotage Selekt system. The system is equipped with a built-in QR reader for Biotage Sfar columns, RFID reader for Selekt collection racks, UV detector, pump, fraction collector, and touch screen. The column information is as follows: Biotage Sfar Silica 60 μm, 25 g, part No. FSRs-0445-0025. All the products were eluted at 70/30 hexanes/EtOAc except for **S4**, which was eluted at 100% methanol.

Experimental Details

Experimental Details of Competition experimentation approach

All S_NAr reactions were performed inside an MBraun glovebox under an N₂ atmosphere. Benzyl alcohol was used as the nucleophile, and potassium *tert*-butoxide (*t*BuOK) was used as the base. Naphthalene or 1,3,5-trimethoxybenzene was used as the internal standard. A library of 74 (hetero)aryl fluorides, chlorides and bromides was used as the substrates. DMSO was used as the reaction solvent.

Stock solutions in DMSO were prepared for each component: the concentration of the substrate stock solutions was 0.5 M, the concentration of the base stock solution was 0.3 M, the concentration of the benzyl alcohol stock solution was 0.2 M, the concentration of the internal standard naphthalene stock solution was 0.03 M and 1,3,5-trimethoxybenzene was 0.2 M. Molecular sieves (4 Å, pellets) were added into the stock solutions of substrates and internal standards to remove possible water content from the commercial materials, which were allowed to sit at least overnight before use.

UPLC or LCMS was used as the analytical technique. To assess each individual (hetero)aryl halide for S_NAr reactivity with benzyl alcohol and determine the retention times of both the substrate and its S_NAr product, individual S_NAr reactions were carried out at room temperature on 1 mL reaction volume scale. In a 4 mL vial containing a stirbar, an aliquot of the substrate stock solution (200 μL, 0.1 mmol) was diluted with additional reaction solvent (DMSO, 600 μL), followed by addition of an aliquot of benzyl alcohol stock solution (100 μL, 0.02 mmol) and an aliquot of *t*BuOK stock solution (100 μL, 0.03 mmol). The resulting solution was mixed for 1-5 hours, then a 200 μL sample was transferred into a 1.5 mL UPLC vial, followed by dilution with 800 μL acetonitrile. The reaction mixture was analyzed by UPLC to confirm the S_NAr reaction occurred and went to completion (from disappearance of the benzyl alcohol peak), and to identify the substrate and its S_NAr product peaks by their retention times.

Competition experiments were performed to obtain the relative activation energies ($\Delta\Delta G^\ddagger_{\text{S}_{\text{N}}\text{Ar}}$) between two substrates in S_NAr. All reactions were conducted at room temperature on 1 mL reaction volume scale under *pseudo* first-order conditions by adding two substrates in excess but equal amount (0.1 mmol, 1.0 equivalent) to compete with benzyl alcohol (0.02 mmol, 0.2 equivalents). The competition reactions were prepared with 200 μL of each substrate stock solution, 200 μL of an internal standard stock solution (the choice of the internal standard is according to the retention times that the peak of the internal standard will not overlap with either the substrate peaks or the product peaks), 200 μL of the reaction solvent DMSO, 100 μL of benzyl alcohol stock solution and 100 μL of *t*BuOK stock solution. The initial concentration of each substrate was 0.1 M, the initial concentration of *t*BuOK was 0.03 M and the initial concentration of benzyl alcohol was 0.02 M. The resulting solution was mixed for 1-5 hours, then a LC sample was prepared by transferring a 200 μL sample and 800 μL acetonitrile into a 1.5 mL UPLC vial (t_{end}). A solution with the initial concentration of the substrates was also prepared by adding 20 μL of each substrate stock solution and 20 μL of the internal standard stock solution into a 1.5 mL UPLC vial then diluting with 940 μL acetonitrile (t_{start}). For each competition experiment, LC chromatograms were recorded for both the t_{start} and t_{end} reaction solutions. The ratio between the two S_NAr rates were obtained from the relative concentrations of the two remaining substrates at t_{end} . This method of quantification avoids the need to obtain relative response factors between all 74 new S_NAr products and the internal standards.

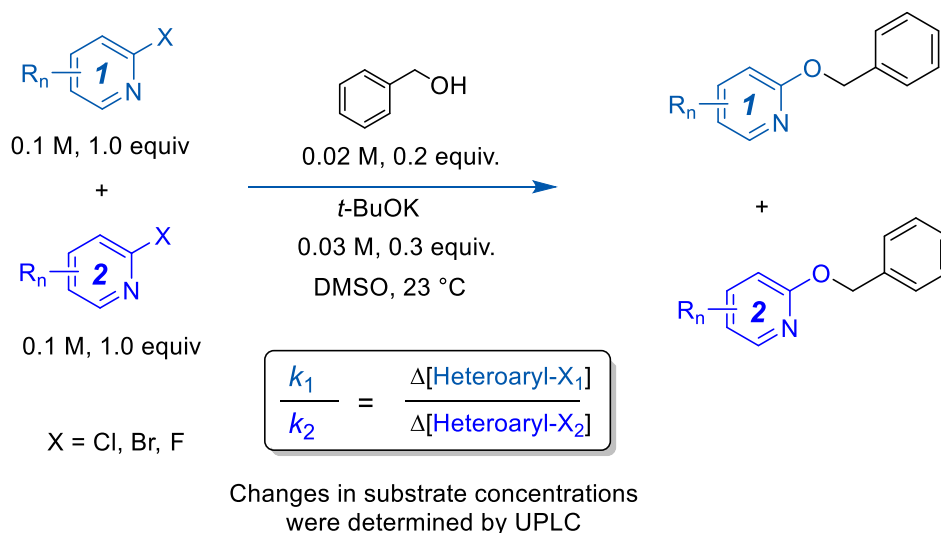


Fig. S1. General competition experiment design. Two electrophiles **1** and **2** in excess but equimolar amounts were reacted with benzyl alcohol as the limiting reagent, using potassium *tert*-butoxide as the base. Outcomes were quantified by changes to electrophile concentration, giving a ratio of *pseudo* first order rate constants k_1/k_2 .

Three typical sets of LC method parameters are shown in Tables S1 to S3 (the 3 min and 10 min ones are on the UPLC and the 6 min one is on the LCMS). The LC inlet methods were adjusted as needed for different pairings of substrates, including the mobile phase gradient, type of columns, runtime and flowrate. To ensure reliable quantitative analysis for every single competition reaction, the inlet method was set up to satisfy the following criteria: the resolution between adjacent peaks should be no less than 1.5, the tailing factor of the peaks of interest should be within 0.8-2.0, and the peak height of the peaks of interest should be above the LOQ (limit of quantification) and below the risk of overloading (usually below 1800 mAu). The UV data was collected at all wavelength range (190 – 800 nm), then the spectra at either 220 nm or 254 nm were extracted to be processed for peak area integration. A representative set of LC chromatograms for a competition experiment is shown in Fig. S2.

Table S1. A typical set of LC method parameters (3 min method) on the Shimadzu Nexera X2 UPLC (equipped with a PDA detector).

LC method parameters

Instrument: Shimadzu UPLC (equipped with a PDA detector)

Instrument Runtime: 3.00 min			
Binary Pump			
Solvent A Name: 0.05% Trifluoroacetic acid in water			
Solvent B Name: 0.05% Trifluoroacetic acid in Acetonitrile			
Gradient Table:			
Time (min)	Flowrate (mL/min)	%A	%B
0.0	0.7	95	5
0.2	0.7	95	5
2.2	0.7	5	95
2.6	0.7	5	95
2.61	0.7	95	5
3.0	0.7	95	5
Column Manager			
Column: Waters Acquity UPLC BEH C18 (2.1 × 30 mm, particle size 1.7 μm)			
Target Column Temperature: 40.0 °C			
AutoSampler			
Needle Wash Solvent: 90:10 Water: Acetonitrile			
Target Sample Temperature: 15.0 °C			
Injection Volume (uL): 1.0			
PDA detector			
Wavelength range: 190 - 800 nm			

Table S2. A typical set of LC method parameters (6 min method) on the Waters Acquity class H UPLC (equipped with a PDA detector a QDa Mass Spectrometer).

LC method parameters

Instrument: Waters Acquity class H UPLC (equipped with a PDA detector a QDa Mass Spectrometer)

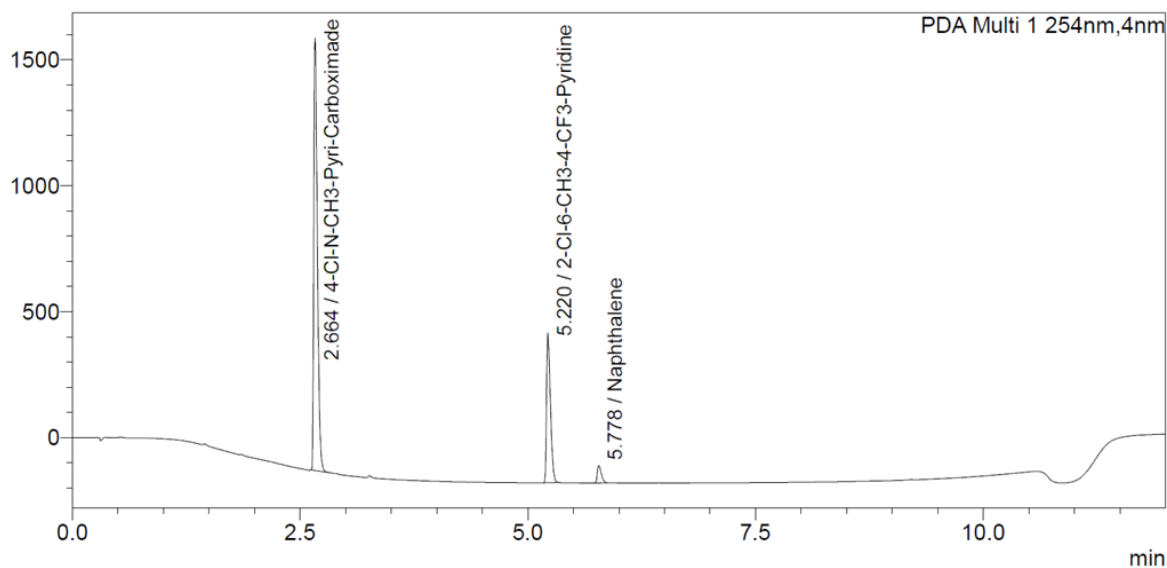
Instrument Runtime: 6.00 min					
Quaternary Pump					
Solvent A Name: 0.1% Formic acid in water					
Solvent B Name: 0.1% Formic acid in Acetonitrile					
Gradient Table:					
Time (min)	Flowrate (mL/min)	%A	%B	%C	%D
0.00	0.5	90	10	0	0
4.00	0.5	10	90	0	0
5.00	0.5	10	90	0	0
5.10	0.5	90	10	0	0
6.00	0.5	90	10	0	0
Waters Acquity Column Manager					
Column: ACQUITY UPLC® BEH C18 column (2.1 × 50 mm, particle size 1.7 μm)					
Target Column Temperature: 40.0 °C					
Waters ACQUITY FTN AutoSampler					
Needle Wash Solvent: 90:10 Water: Acetonitrile					
Target Sample Temperature: 10.0 °C					
Injection Volume (uL) - 1.00					
Waters Acquity PDA detector					
PDA Detector Type: UPLC eLambda 800 nm					
Wavelength range: 190 - 800 nm					
QDa Mass Spectrometer					
Mass Scanning range: 30 Da to 1250 Da					

Table S3. A typical set of LC method parameters (10 min method) on the Shimadzu Nexera X2 UPLC (equipped with a PDA detector).

LC method parameters

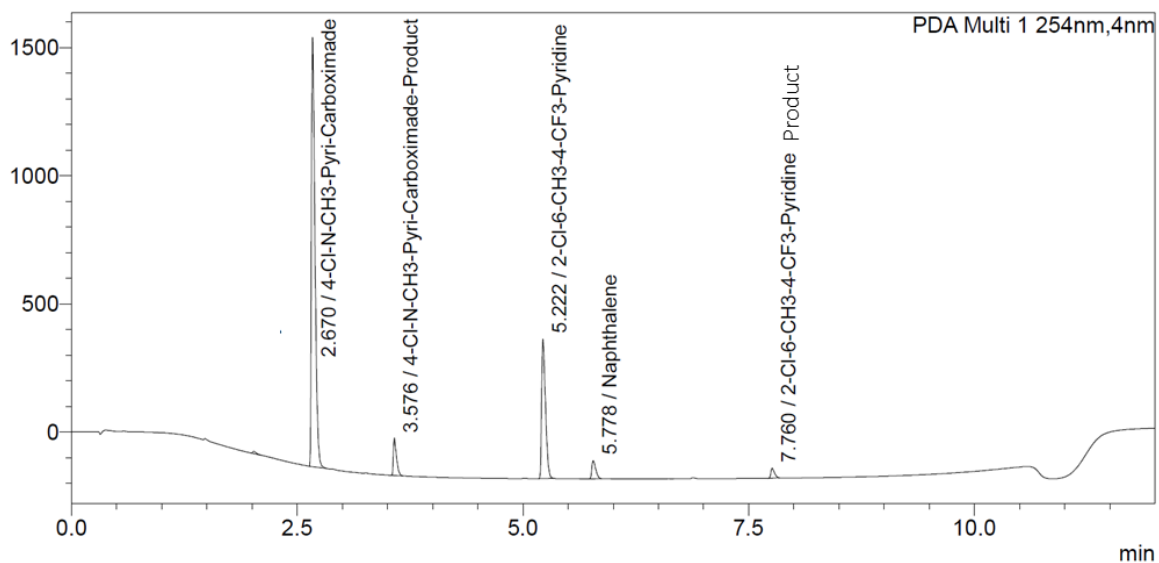
Instrument: Shimadzu UPLC (equipped with a PDA detector)

Instrument Runtime: 10.00 min			
Binary Pump			
Solvent A Name: 0.05% Trifluoroacetic acid in water			
Solvent B Name: 0.05% Trifluoroacetic acid in Acetonitrile			
Gradient Table:			
Time (min)	Flowrate (mL/min)	%A	%B
0.0	0.6	90	10
7.5	0.6	5	95
8.5	0.6	5	95
8.51	0.6	90	10
10	0.6	90	10
Column Manager			
Column: Raptor ARC-18 (2.1 × 100 mm, particle size 1.8 μm)			
Target Column Temperature: 40.0 °C			
AutoSampler			
Needle Wash Solvent: 90:10 Water: Acetonitrile			
Target Sample Temperature: 15.0 °C			
Injection Volume (uL): 1.0			
PDA detector			
Wavelength range: 190 - 800 nm			

A

PDA Ch1 254 nm

Peak#	Name	Ret. Time	Height	Area	Tailing factor	Resolution
1	4-Cl-N-CH3-Pyri-Carboximade	2.664	1718310	4767682	1.674	--
2	2-Cl-6-CH3-4-CF3-Pyridine	5.220	593995	1751829	1.634	31.146
3	Naphthalene	5.778	69512	207583	1.589	6.509

B

PDA Ch1 254 nm

Peak#	Name	Ret. Time	Height	Area	Tailing factor	Resolution
1	4-Cl-N-CH3-Pyri-Carboximade	2.670	1675108	4615358	1.828	--
2	4-Cl-N-CH3-Pyri-Carboximade-Product	3.576	146336	346262	1.830	12.200
3	2-Cl-6-CH3-4-CF3-Pyridine	5.222	543570	1600488	1.668	21.293
4	Naphthalene	5.778	70360	207849	1.603	6.523
5	2-Cl-6-CH3-4-CF3-Pyridine Product	7.760	38942	106810	1.631	23.780

Fig. S2. UPLC chromatograms recorded by Shimadzu Nexera X2 UPLC system at 254 nm for competition reaction between 4-chloro-N-methylpyridine-2-carboxamide and 2-chloro-6-methyl-4 (trifluoromethyl)pyridine: A. LC chromatogram at t_{start} ; B. LC chromatograms at t_{end} .

Experimental Details of kinetic analysis for the touchstone reactions

Three of the S_NAr reactions were chosen as the touchstone reactions and their absolute reaction rates were measured by kinetic analysis. The ΔG^\ddagger_{SNAr} (obtained from the *pseudo* first order rate constant) determined for one of those touchstone reactions (2-chloropyridine and benzyl alcohol in DMSO) was used to calibrate all the relative rates determined by competition reactions, giving the absolute ΔG^\ddagger_{SNAr} values for the entire array of substrates. The other two touchstone reactions were used to validate the accuracy of the rate constants determined from this competition experimentation approach, by comparing the reaction rates determined from kinetic analysis to those obtained from competition experiments.

All of the following procedures were performed inside an MBraun glovebox under an N_2 atmosphere. The reactions were conducted with the same concentrations of each species as for the competition reactions (0.1 M of substrate, 0.02 M of benzyl alcohol, 0.03 M of *t*BuOK and 0.003 M of naphthalene or 0.02 M 1,3,5-trimethoxybenzene as internal standard in 10 mL of DMSO). All the components except for *t*BuOK were added into a 20 mL glass vial charged with a stir bar and mixed well by stirring, then a small amount (200 μ L) of this solution was transferred into an UPLC vial followed by dilution with 800 μ L of acetonitrile as the time zero (t_0) point. A stopwatch was used to keep track of the reaction time immediately after the *t*BuOK stock solution was added. After a certain time period, a portion of the reaction mixture (500 μ L) was transferred into a 2 mL UPLC vial containing 500 μ L of 0.01 M hydrochloric acid in acetonitrile; the excess amount of acid was added to quench the reaction by consuming the base in the reaction solution. Reaction mixture aliquots were withdrawn at 10 time points to construct the reaction progress curves.

These reaction mixture aliquots were analyzed by UPLC and the peak areas of benzyl alcohol and the internal standard were determined. A [BnOH] versus time plot was constructed, and an exponential fit applied. Using a *pseudo* first-order rate law assumption, the rate constant (k) was calculated using Eq (S1). This k value was substituted into the Eyring equation to obtain the Gibbs free energy of this reaction (ΔG^\ddagger_{SNAr}) according to Eq (S2). The results of the three reaction rate measurement experiments are summarized in Fig. S3.

According to the *pseudo* first-order rate law assumption, the rate constant (k) can be determined by Eq (S1):

$$[\text{benzyl alc.}]_t = [\text{benzyl alc.}]_{t=0} e^{-k't}, \text{ where } k' = k[\text{substrate}] \quad \text{Eq(S1)}$$

Then the Gibbs free energy (ΔG^\ddagger_{SNAr}) is determined by the Eyring equation Eq (S2) (transmission coefficient κ assumed to be 1):

$$k = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^\ddagger}{RT}\right) \Rightarrow \Delta G^\ddagger = -RT \ln \frac{kh}{k_B T} \quad \text{Eq(S2)}$$

where R is the ideal gas constant: $8.314 \text{ J} \times (\text{K} \times \text{mol})^{-1}$;

T is the reaction temperature: 299.55 K;

k_B is the Boltzmann constant: $1.38 \times 10^{-23} \text{ J} \times \text{K}^{-1}$;

h is Planck constant: $6.63 \times 10^{-34} \text{ J} \times \text{s}$

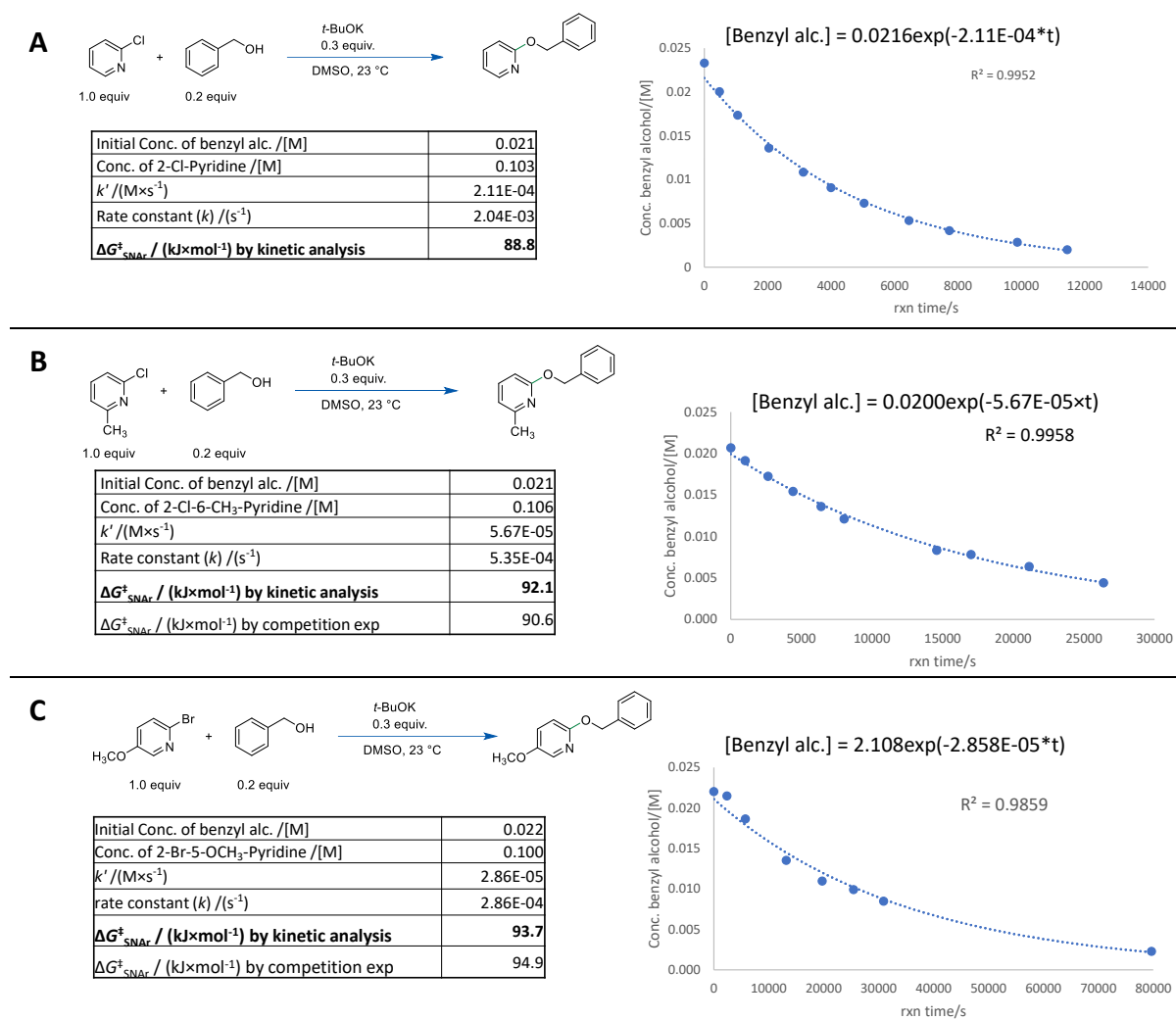


Fig. S3.

Determination of the rate constants for touchstone reactions: **A.** 2-chloropyridine and benzyl alcohol in DMSO; **B.** 2-chloro-6-methylpyridine and benzyl alcohol in DMSO; **C.** 2-bromo-5-methoxypyridine and benzyl alcohol in DMSO.

Calculating $\Delta G^{\ddagger}_{SNAr}$ for the entire array of Substrates

All of the competition reactions were conducted under *pseudo* first-order conditions, with the two electrophiles in excess but equal amounts ([substrate]:[nucleophile] = 5:1). The ratio of the reaction rates can be expressed as Eq (S3):

$$\frac{r_1}{r_2} = \frac{k_1 [Nucleophile] [Substrate_1]}{k_2 [Nucleophile] [Substrate_2]} = \frac{k_1 [Substrate_1]}{k_2 [Substrate_2]} \quad Eq(S3)$$

The concentrations of the two substrates are assumed to stay constant throughout the reaction; then, Eq (S1) can be simplified to Eq (S4):

$$\frac{r_1}{r_2} = \frac{k_1}{k_2} = \frac{\frac{d[Substrate_1]}{dt}}{\frac{d[Substrate_2]}{dt}} = \frac{\Delta[Substrate_1]}{\Delta[Substrate_2]} \quad Eq(S4)$$

From UPLC chromatography analysis, $\frac{\Delta[\text{Substrate}_1]}{\Delta[\text{Substrate}_2]}$ is equal to the percent change of the peak area of the two substrates from the reaction mixture at t_{start} to that at t_{end} . By substituting the Eyring equation (Eq (S5)):

$$k = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^\ddagger}{RT}\right) \quad \text{Eq(S5)}$$

into Eq (S4), the relative activation energy ($\Delta\Delta G_{\text{SNAr}}^\ddagger$) of the two S_{NAr} reactions from the competition can be calculated by Eq (S6):

$$\frac{k_1}{k_2} = \exp\left(\frac{\Delta G_{\text{SNAr}}^\ddagger_2 - \Delta G_{\text{SNAr}}^\ddagger_1}{RT}\right) \Rightarrow \Delta\Delta G_{\text{SNAr}}^\ddagger = \ln\left(\frac{k_1}{k_2}\right) RT \Rightarrow$$

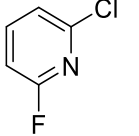
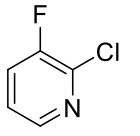
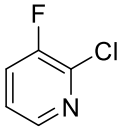
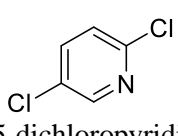
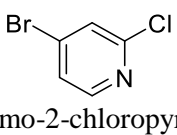
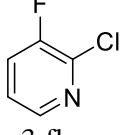
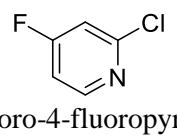
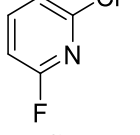
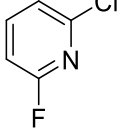
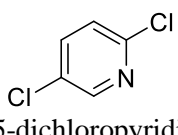
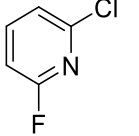
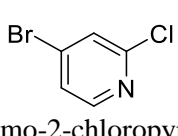
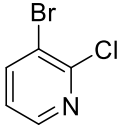
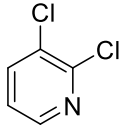
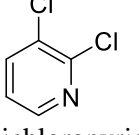
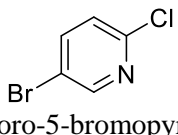
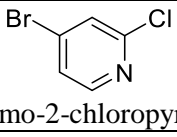
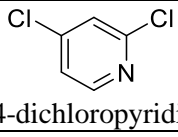
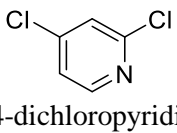
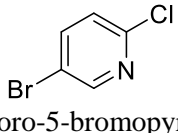
$$\Delta\Delta G_{\text{SNAr}}^\ddagger = \ln\left(\frac{\% \Delta \text{Peak area}_{\text{Substrate1}}}{\% \Delta \text{Peak area}_{\text{Substrate2}}}\right) RT \quad \text{Eq(S6)}$$

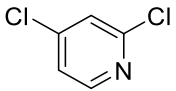
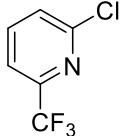
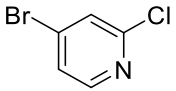
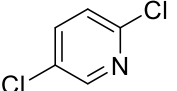
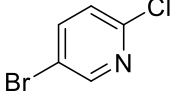
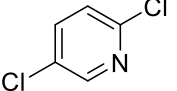
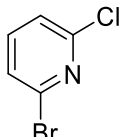
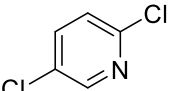
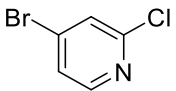

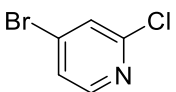
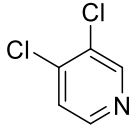
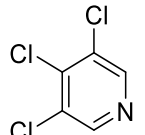
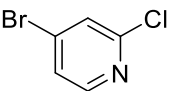
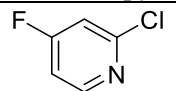
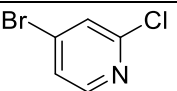
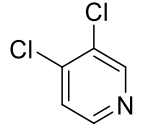
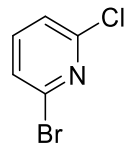
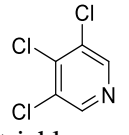
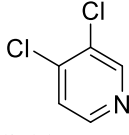
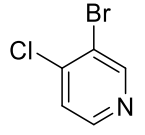
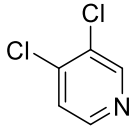
$$\text{where } \Delta\Delta G_{\text{SNAr}}^\ddagger = \Delta G_{\text{SNAr}}^\ddagger_2 - \Delta G_{\text{SNAr}}^\ddagger_1$$

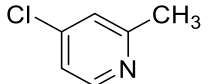
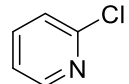
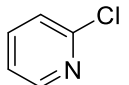
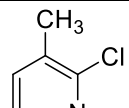
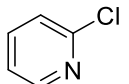
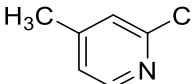
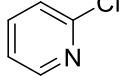
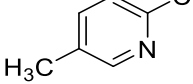
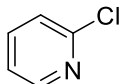
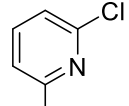
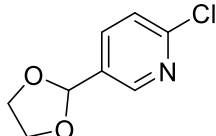
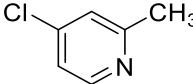
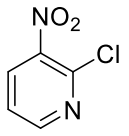
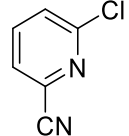
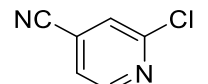
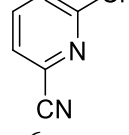
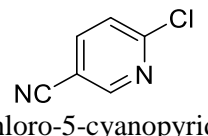
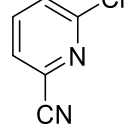
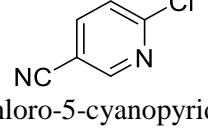
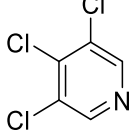
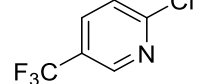
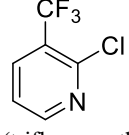
The $\Delta\Delta G_{\text{SNAr}}^\ddagger$ values determined from the competition experiments were related to the absolute $\Delta G_{\text{SNAr}}^\ddagger$ determined from kinetic analysis of the touchstone reaction between 2-chloropyridine and benzyl alcohol in DMSO; therefore, all other $\Delta G_{\text{SNAr}}^\ddagger$ values are given relative to 2-chloropyridine.

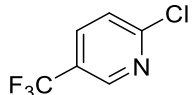
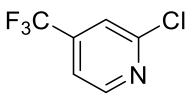
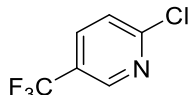
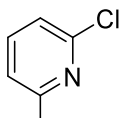
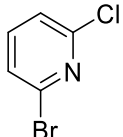
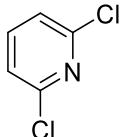
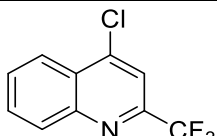
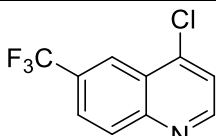
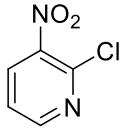
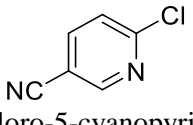
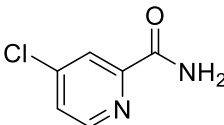
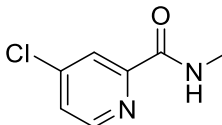
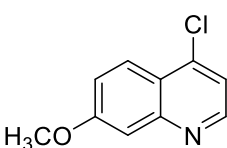
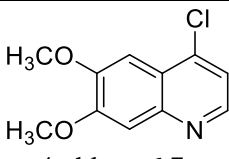
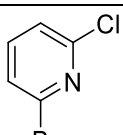
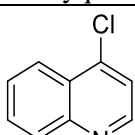
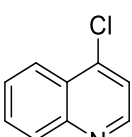
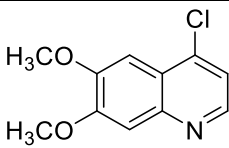
The S_{NAr} rate ratio of 94 competition reactions is shown in Table S4, and the experimental $\Delta G_{\text{SNAr}}^\ddagger$ for the 74 substrates determined from the 94 competition reactions in DMSO is summarized in Table S5. Substrates highlighted are those with $\Delta G_{\text{SNAr}}^\ddagger$ determined by multiple competition experiments with different substrate pairings. The given $\Delta G_{\text{SNAr}}^\ddagger$ values for these substrates are averages of those determined by at least two different competition experiments, and the relative standard deviation (RSD) as well as the $\Delta G_{\text{SNAr}}^\ddagger$ from each individual measurement for these examples are also given.

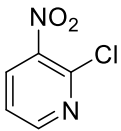
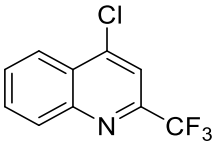
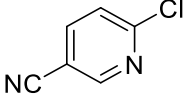
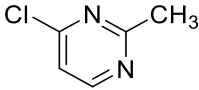
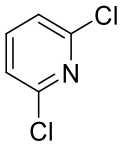
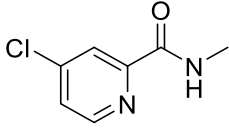
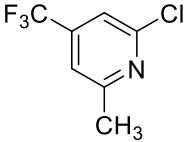
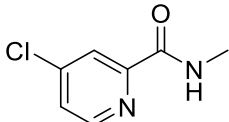
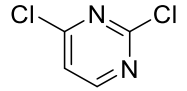
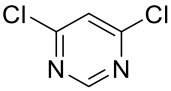
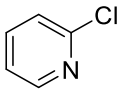
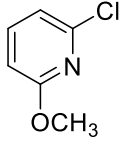
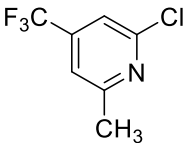
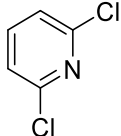
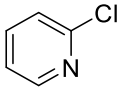
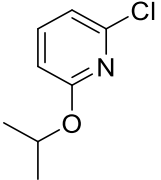
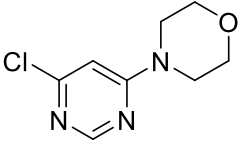
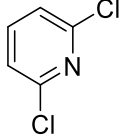
Table S4. S_NAr rate ratios (r_1/r_2) calculated using Eq (S3) for 94 competition reactions in DMSO.

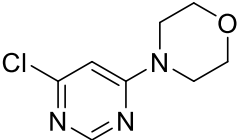
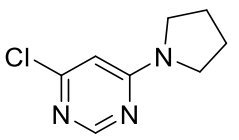
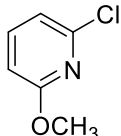
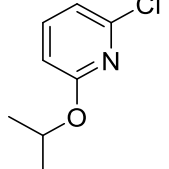
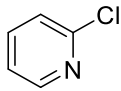
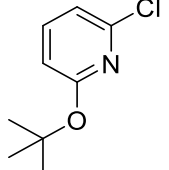
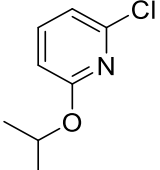
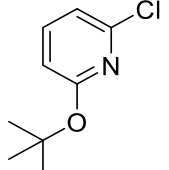
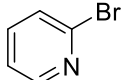
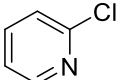
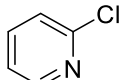
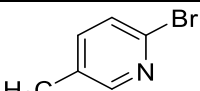
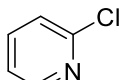
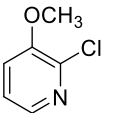
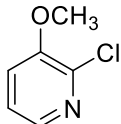
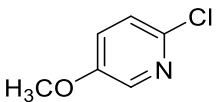
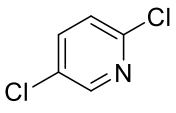
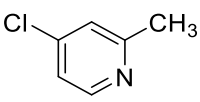
Exp No.	Substrate A	Substrate B	S_NAr rate ratio ^[a]
1	 2-chloro-6-fluoropyridine	 2-chloro-3-fluoropyridine	5.88
2	 2-chloro-3-fluoropyridine	 2,5-dichloropyridine	4.41
3	 4-bromo-2-chloropyridine	 2-chloro-3-fluoropyridine	3.70
4	 2-chloro-4-fluoropyridine	 2-chloro-6-fluoropyridine	1.61
5	 2-chloro-6-fluoropyridine	 2,5-dichloropyridine	27.05
6	 2-chloro-6-fluoropyridine	 4-bromo-2-chloropyridine	4.58
7	 2-chloro-3-bromopyridine	 2,3-dichloropyridine	1.09
8	 2,3-dichloropyridine	 2-chloro-5-bromopyridine	1.61
9	 4-bromo-2-chloropyridine	 2,4-dichloropyridine	1.09
10	 2,4-dichloropyridine	 2-chloro-5-bromopyridine	7.78

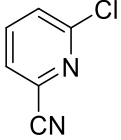
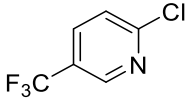
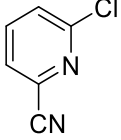
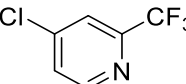
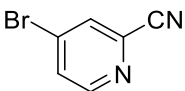
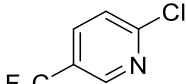
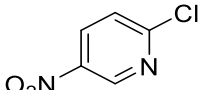
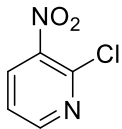
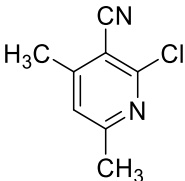
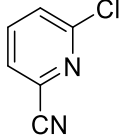
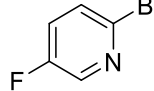
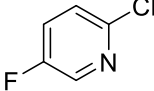
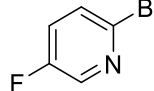
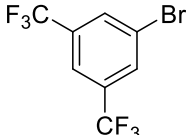
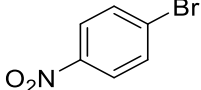
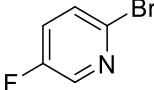
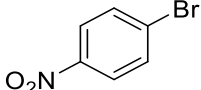
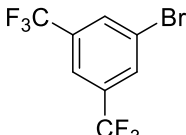
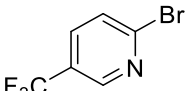
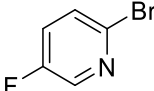
11	 2,4-dichloropyridine	 2-chloro-6-(trifluoromethyl)pyridine	1.87
12	 4-bromo-2-chloropyridine	 2,5-dichloropyridine	10.42
13	 2-chloro-5-bromopyridine	 2,5-dichloropyridine	1.45
14	 2-chloro-6-bromopyridine	 2,5-dichloropyridine	13.16
15	 4-bromo-2-chloropyridine	 2-chloro-3-bromopyridine	5.56
16	 4-bromo-2-chloropyridine	 3,4-dichloropyridine	1.30
17	 3,4,5-trichloropyridine	 4-bromo-2-chloropyridine	4.76
18	 2-chloro-4-fluoropyridine	 4-bromo-2-chloropyridine	9.49
19	 3,4-dichloropyridine	 2-chloro-6-bromopyridine	1.35
20	 3,4,5-trichloropyridine	 3,4-dichloropyridine	4.53
21	 4-chloro-3-bromopyridine	 3,4-dichloropyridine	1.45

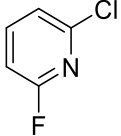
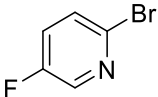
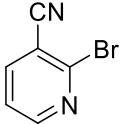
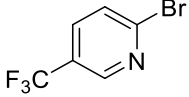
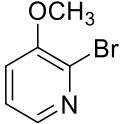
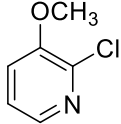
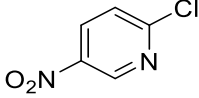
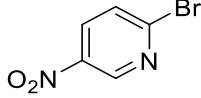
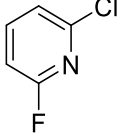
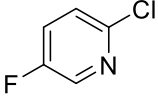
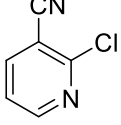
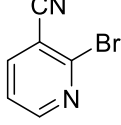
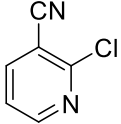
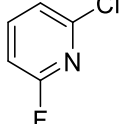
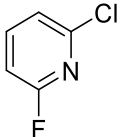
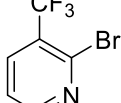
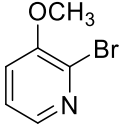
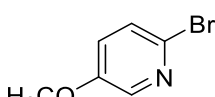
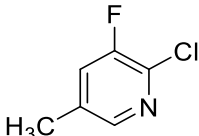
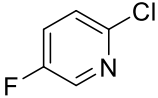
22	 4-chloro-2-methylpyridine	 2-chloropyridine	1.51
23	 2-chloropyridine	 2-chloro-3-methylpyridine	2.72
24	 2-chloropyridine	 2-chloro-4-methylpyridine	3.23
25	 2-chloropyridine	 2-chloro-5-methylpyridine	10.77
26	 2-chloropyridine	 2-chloro-6-methylpyridine	2.49
27	 2-chloro-5-(1,3-dioxolan-2-yl)pyridine	 4-chloro-2-methylpyridine	3.33
28	 2-chloro-3-cyanopyridine	 2-chloro-6-cyanopyridine	1.84
29	 2-chloro-4-cyanopyridine	 2-chloro-6-cyanopyridine	1.37
30	 2-chloro-5-cyanopyridine	 2-chloro-6-cyanopyridine	3.82
31	 2-chloro-5-cyanopyridine	 3,4,5-trichloropyridine	1.88
32	 2-chloro-5-(trifluoromethyl)pyridine	 2-chloro-3-(trifluoromethyl)pyridine	2.44

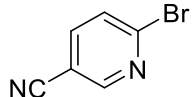
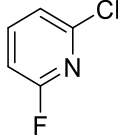
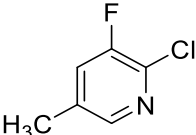
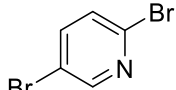
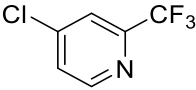
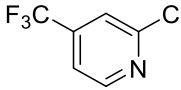
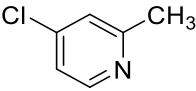
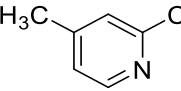
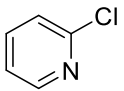
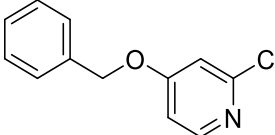
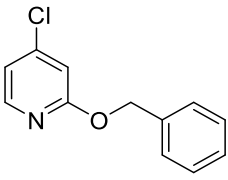
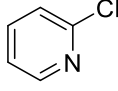
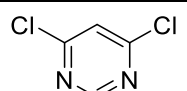
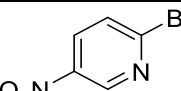
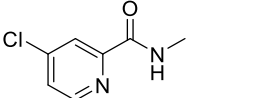
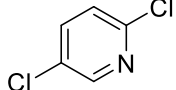
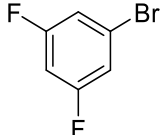
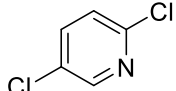
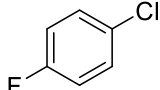
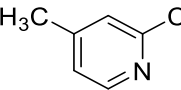
33	 2-chloro-5-(trifluoromethyl)pyridine	 2-chloro-4-(trifluoromethyl)pyridine	2.15
34	 2-chloro-5-(trifluoromethyl)pyridine	 2-chloro-6-(trifluoromethyl)pyridine	5.76
35	 2-bromo-6-chloropyridine	 2,6-dichloropyridine	1.25
36	 4-chloro-2-(trifluoromethyl)quinoline	 4-chloro-6-(trifluoromethyl)quinoline	1.35
37	 2-chloro-3-nitropyridine	 2-chloro-5-cyanopyridine	3.18
38	 4-chloropyridine-2-carboxamide	 4-chloro-N-methylpyridine-2-carboxamide	1.08
39	 4-chloro-7-methoxyquinoline	 4-chloro-6,7-dimethoxyquinoline	7.30
40	 2-bromo-6-chloropyridine	 4-chloroquinoline	3.70
41	 4-chloroquinoline	 4-chloro-6,7-dimethoxyquinoline	22.00

42	 <p>2-chloro-3-nitropyridine</p>	 <p>4-chloro-2-(trifluoromethyl)quinoline</p>	1.83
43	 <p>2-chloro-5-cyanopyridine</p>	 <p>4-chloro-2-methylpyrimidine</p>	2.91
44	 <p>2,6-dichloropyridine</p>	 <p>4-chloro-N-methylpyridine-2-carboxamide</p>	3.07
45	 <p>2-chloro-6-methyl-4-(trifluoromethyl)pyridine</p>	 <p>4-chloro-N-methylpyridine-2-carboxamide</p>	2.64
46	 <p>2,4-dichloropyrimidine</p>	 <p>4,6-dichloropyrimidine</p>	1.30
47	 <p>2-chloropyridine</p>	 <p>2-chloro-6-methoxypyridine</p>	1.10
48	 <p>2-chloro-6-methyl-4-(trifluoromethyl)pyridine</p>	 <p>2,6-dichloropyridine</p>	1.25
49	 <p>2-chloropyridine</p>	 <p>2-chloro-6-isopropoxypyridine</p>	1.39
50	 <p>4-(6-chloropyrimidin-4-yl)morpholine</p>	 <p>2,6-dichloropyridine</p>	1.12

51	 4-(6-chloropyrimidin-4-yl)morpholine	 4-chloro-6-(pyrrolidin-1-yl)pyrimidine	9.07
52	 2-chloro-6-methoxypyridine	 2-chloro-6-isopropoxypyridine	1.54
53	 2-chloropyridine	 2-chloro-6-(tert-butoxy)pyridine	4.42
54	 2-chloro-6-isopropoxypyridine	 2-chloro-6-(tert-butoxy)pyridine	2.00
55	 2-bromopyridine	 2-chloropyridine	2.34
56	 2-chloropyridine	 2-bromo-5-methylpyridine	1.77
57	 2-chloropyridine	 2-chloro-3-methoxypyridine	11.20
58	 2-chloro-3-methoxypyridine	 2-chloro-5-methoxypyridine	1.66
59	 2,5-dichloropyridine	 4-chloro-2-methylpyridine	5.55

60	 2-chloro-6-cyanopyridine	 2-chloro-5-(trifluoromethyl)pyridine	1.35
61	 2-chloro-6-cyanopyridine	 4-chloro-2-(trifluoromethyl)pyridine	1.42
62	 4-bromo-2-cyanopyridine	 2-chloro-5-(trifluoromethyl)pyridine	4.70
63	 2-chloro-5-nitropyridine	 2-chloro-3-nitropyridine	4.42
64	 2-chloro-4,6-dimethyl-3-cyanopyridine	 2-chloro-6-cyanopyridine	2.68
65	 2-bromo-5-fluoropyridine	 2-chloro-5-fluoropyridine	1.38
66	 2-bromo-5-fluoropyridine	 1-bromo-3,5-bis(trifluoromethyl)benzene	7.18
67	 1-bromo-4-nitrobenzene	 2-bromo-5-fluoropyridine	3.01
68	 1-bromo-4-nitrobenzene	 1-bromo-3,5-bis(trifluoromethyl)benzene	18.20
69	 2-bromo-5-(trifluoromethyl)pyridine	 2-bromo-5-fluoropyridine	5.82

70	 2-chloro-6-fluoropyridine	 2-bromo-5-fluoropyridine	11.43
71	 2-bromo-3-cyanopyridine	 2-bromo-5-(trifluoromethyl)pyridine	4.60
72	 2-bromo-3-methoxypyridine	 2-chloro-3-methoxypyridine	2.18
73	 2-chloro-5-nitropyridine	 2-bromo-5-nitropyridine	1.14
74	 2-chloro-6-fluoropyridine	 2-chloro-5-fluoropyridine	17.83
75	 2-chloro-3-cyanopyridine	 2-bromo-3-cyanopyridine	1.18
76	 2-chloro-3-cyanopyridine	 2-chloro-6-fluoropyridine	2.79
77	 2-chloro-6-fluoropyridine	 2-bromo-3-(trifluoromethyl)pyridine	2.96
78	 2-bromo-3-methoxypyridine	 2-bromo-5-methoxypyridine	2.73
79	 2-chloro-3-fluoro-5-methylpyridine	 2-chloro-5-fluoropyridine	3.05

80	 2-bromo-5-cyanopyridine	 2-chloro-6-fluoropyridine	2.96
81	 2-chloro-3-fluoro-5-methylpyridine	 2,5-dibromopyridine	2.84
82	 4-chloro-2-(trifluoromethyl)pyridine	 2-chloro-4-(trifluoromethyl)pyridine	2.94
83	 4-chloro-2-methylpyridine	 2-chloro-4-methylpyridine	3.36
84	 2-chloropyridine	 4-(benzyloxy)-2-chloropyridine	2.48
85	 2-(benzyloxy)-4-chloropyridine	 2-chloropyridine	2.33
86	 4,6-dichloropyrimidine	 2-bromo-5-nitropyridine	2.54
87	 4-chloro-N-methylpyridine-2-carboxamide	 2,5-dichloropyridine	1.71
88	 1-bromo-3,5-difluorobenzene	 2,5-dichloropyridine	2.97
89	 1-chloro-4-fluorobenzene	 2-chloro-4-methylpyridine	1.01

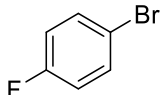
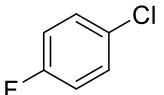
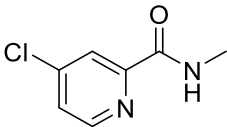
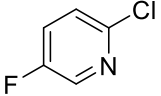
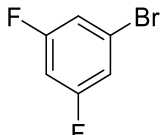
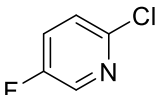
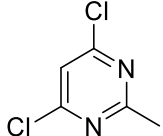
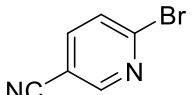
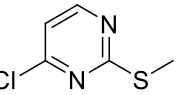
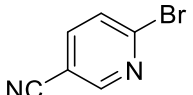
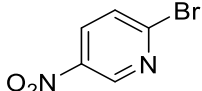
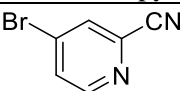
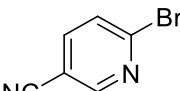
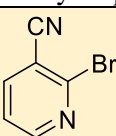
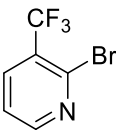
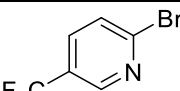
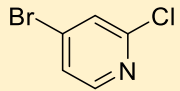
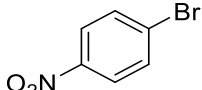
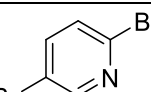
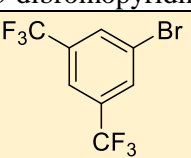
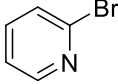
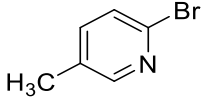
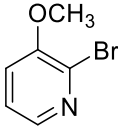
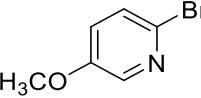
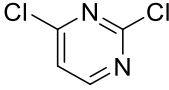
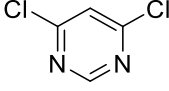
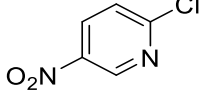
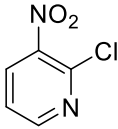
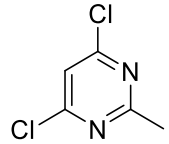
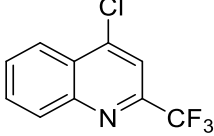
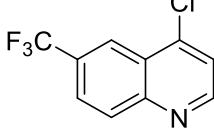
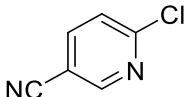
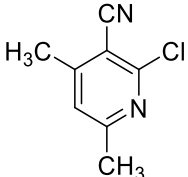
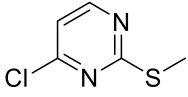
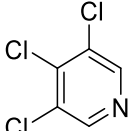
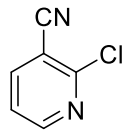
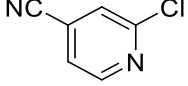
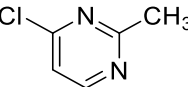
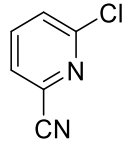
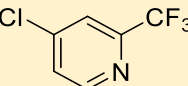
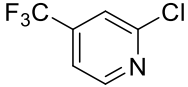
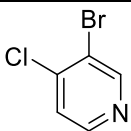
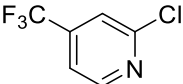
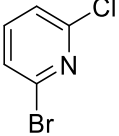
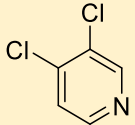
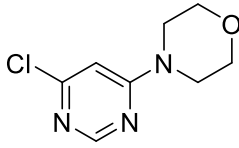
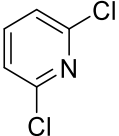
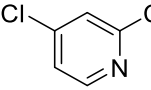
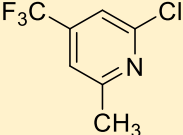
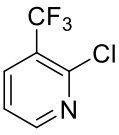
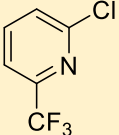
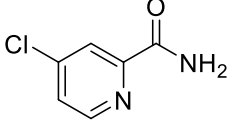
90	 1-bromo-4-fluorobenzene	 1-chloro-4-fluorobenzene	1.42
91	 4-chloro-N-Methylpyridine-2-carboxamide	 2-chloro-5-fluoropyridine	2.54
92	 1-bromo-3,5-difluorobenzene	 2-chloro-5-fluoropyridine	1.37
93	 4,6-dichloro-2-methylpyrimidine	 2-bromo-5-cyanopyridine	3.37
94	 4-chloro-2-(methylthio)pyrimidine	 2-bromo-5-cyanopyridine	1.27

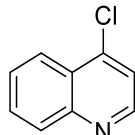
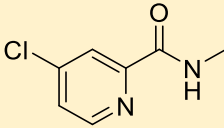
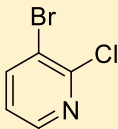
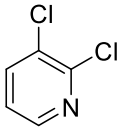
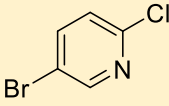
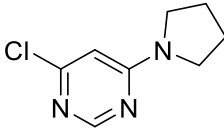
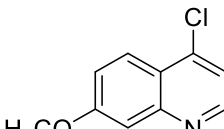
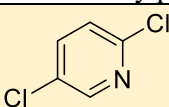
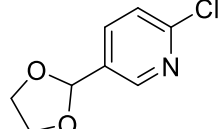
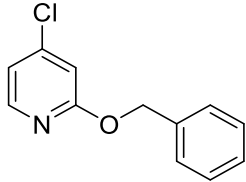
Table S5.Observed $\Delta G^\ddagger_{\text{SNAr}}$ for 74 substrates determined from competition reactions in DMSO.

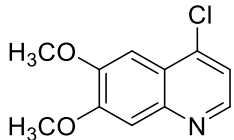
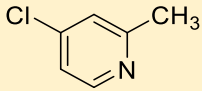
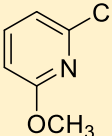
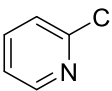
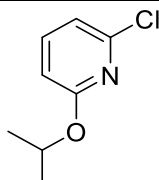
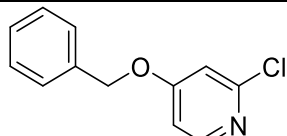
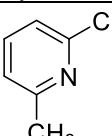
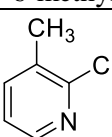
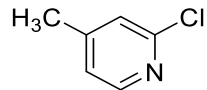
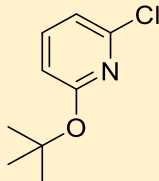
Substrate	Leaving group	$\Delta G^\ddagger_{\text{SNAr}}$ (kJ/mol)	RSD ^[a]
			$\Delta G^\ddagger_{\text{SNAr}}$ from independent measurements
 2-bromo-5-nitropyridine	Br	64.36	
 4-bromo-2-cyanopyridine	Br	70.84	
 2-bromo-5-cyanopyridine	Br	72.28	
 2-bromo-3-cyanopyridine	Br	73.01	0.32%
			72.84 73.17
 2-bromo-3-(trifluoromethyl)pyridine	Br	76.92	
 2-bromo-5-(trifluoromethyl)pyridine	Br	76.97	
 4-bromo-2-chloropyridine	Br	76.99	2.00%
			76.07 76.14 78.78
 1-bromo-4-nitrobenzene	Br	78.61	
 2,5-dibromopyridine	Br	81.98	
 1-bromo-3,5-bis(trifluoromethyl)benzene	Br	85.98	0.23%
			86.12 85.84

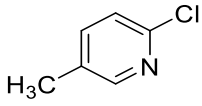
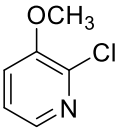
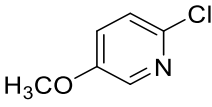
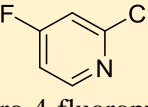
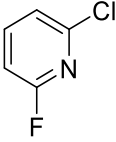
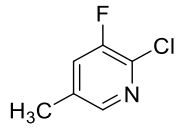
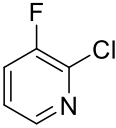
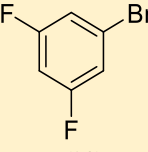
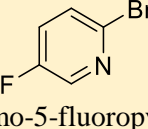
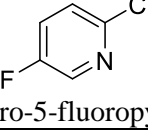
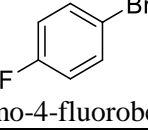
 2-bromopyridine	Br	86.22	
 2-bromo-5-methylpyridine	Br	90.22	
 2-bromo-3-methoxypyridine	Br	92.41	
 2-bromo-5-methoxypyridine	Br	94.91	
 2,4-dichloropyrimidine	Cl	61.38	
 4,6-dichloropyrimidine	Cl	62.03	
 2-chloro-5-nitropyridine	Cl	64.03	
 2-chloro-3-nitropyridine	Cl	67.73	
 4,6-dichloro-2-methylpyrimidine	Cl	67.82	
 4-chloro-2-(trifluoromethyl)quinoline	Cl	69.24	
 4-chloro-6-(trifluoromethyl)quinoline	Cl	69.98	

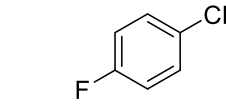
 2-chloro-5-cyanopyridine	Cl	70.61	
 2-chloro-4,6-dimethyl-3-cyanopyridine	Cl	71.49	
 4-chloro-2-(methylthio)pyrimidine	Cl	71.69	
 3,4,5-trichloropyridine	Cl	72.18	
 2-chloro-3-cyanopyridine	Cl	72.43	
 2-chloro-4-cyanopyridine	Cl	73.17	
 4-chloro-2-methylpyrimidine	Cl	73.27	
 2-chloro-6-cyanopyridine	Cl	73.95	
 4-chloro-2-(trifluoromethyl)pyridine	Cl	74.37	0.86%
 2-chloro-5-(trifluoromethyl)pyridine	Cl	74.70	74.82 73.92
 4-chloro-3-bromopyridine	Cl	75.75	

 2-chloro-4-(trifluoromethyl)pyridine	Cl	76.60	
 2-chloro-6-bromopyridine	Cl	76.62	
 3,4-dichloropyridine	Cl	76.76	1.31%
			75.95 77.65 75.88
 4-(6-chloropyrimidin-4-yl)morpholine	Cl	76.90	
 2,6-dichloropyridine	Cl	77.18	
 2,4-dichloropyridine	Cl	77.20	
 2-chloro-6-methyl-4-(trifluoromethyl)pyridine	Cl	77.36	1.34%
			78.09 76.62
 2-chloro-3-(trifluoromethyl)pyridine	Cl	77.69	
 2-chloro-6-(trifluoromethyl)pyridine	Cl	78.91	0.27%
			78.76 79.06
 4-chloropyridine-2-carboxamide	Cl	79.78	

 4-chloro-quinoline	Cl	79.88	
 4-chloro-N-Methylpyridine-2-carboxamide	Cl	80.51	1.29%
			79.97 81.71 79.84
 2-chloro-3-bromopyridine	Cl	81.09	0.30%
			80.92 81.27
 2,3-dichloropyridine	Cl	81.13	
 2-chloro-5-bromopyridine	Cl	82.21	0.17%
			82.31 82.12
 4-chloro-6-(pyrrolidin-1-yl)pyrimidine	Cl	82.39	
 4-chloro-7-methoxyquinoline	Cl	82.63	
 2,5-dichloropyridine	Cl	83.04	0.23%
			83.09 83.20 82.83
 2-chloro-5-(1,3-dioxolan-2-yl)pyridine	Cl	84.78	
 2-(benzyloxy)-4-chloropyridine	Cl	86.23	

 4-chloro-6,7-dimethoxyquinoline	Cl	87.58	
 4-chloro-2-methylpyridine	Cl	87.77	0.75%
			87.31 88.24
 2-chloro-6-methoxypyridine	Cl	88.56	0.76%
			89.04 88.08
 2-chloropyridine	Cl	88.80 touchstone reaction	
 2-chloro-6-isopropoxypyridine	Cl	89.16	
 4-(benzyloxy)-2-chloropyridine	Cl	90.60	
 2-chloro-6-methylpyridine	Cl	90.61	
 2-chloro-3-methylpyridine	Cl	90.83	
 2-chloro-4-methylpyridine	Cl	91.26	
 2-chloro-6-(tert-butoxy)-pyridine	Cl	91.69	1.25%
			92.50 90.88

 2-chloro-5-methylpyridine	Cl	94.26	
 2-chloro-3-methoxypyridine	Cl	94.35	
 2-chloro-5-methoxypyridine	Cl	95.61	
 2-chloro-4-fluoropyridine	F	72.60	2.35%
			73.80 71.39
 2-chloro-6-fluoropyridine	F	74.99	
 2-chloro-3-fluoro-5-methylpyridine	F	79.38	
 2-chloro-3-fluoropyridine	F	79.40	
 1-bromo-3,5-difluorobenzene	F	80.85	0.92%
			80.33 81.38
 2-bromo-5-fluoropyridine	F	81.21	0.27%
			81.36 81.05
 2-chloro-5-fluoropyridine	F	82.16	
 1-bromo-4-fluorobenzene	F	90.36	

 1-chloro-4-fluorobenzene	F	91.23	
---	---	-------	--

^[a]RSD values determined from the $\Delta G^\ddagger_{\text{S}_{\text{N}}\text{Ar}}$ values obtained by independent competition experiments with different substrates (see Table S4).

Hammett Analyses of *para*- and *meta*-Substituted Substrates

To further validate the kinetic parameters obtained by competition experiments, we have obtained reaction constants (ρ) through construction of Hammett plots $-\log(k_Z/k_H)$ versus substituent σ values – for two sets of *para*-substituted substrates and two set of *meta*-substituted substrates undergoing $\text{S}_{\text{N}}\text{Ar}$. These include: 5-substituted-2-chloropyridines and 5-substituted-2-bromopyridines (σ_{para}); 4-substituted-2-chloropyridines and 6-substituted-2-chloropyridines (σ_{meta}). Substituent σ values were obtained from published tables.⁵

We obtain reaction constants of $\rho = 5.1$ in DMSO for $\text{S}_{\text{N}}\text{Ar}$ of the 5-Z-2-Cl-pyridines (Fig. S4), $\rho = 4.4$ in DMSO for $\text{S}_{\text{N}}\text{Ar}$ of the 5-Z-2-Br-pyridines (Fig. S5), $\rho = 5.1$ in DMSO for $\text{S}_{\text{N}}\text{Ar}$ of the 4-Z-2-Cl-pyridines (Fig. S6), and $\rho = 4.8$ in DMSO for $\text{S}_{\text{N}}\text{Ar}$ of the 6-Z-2-Cl-pyridines (Fig. S7).

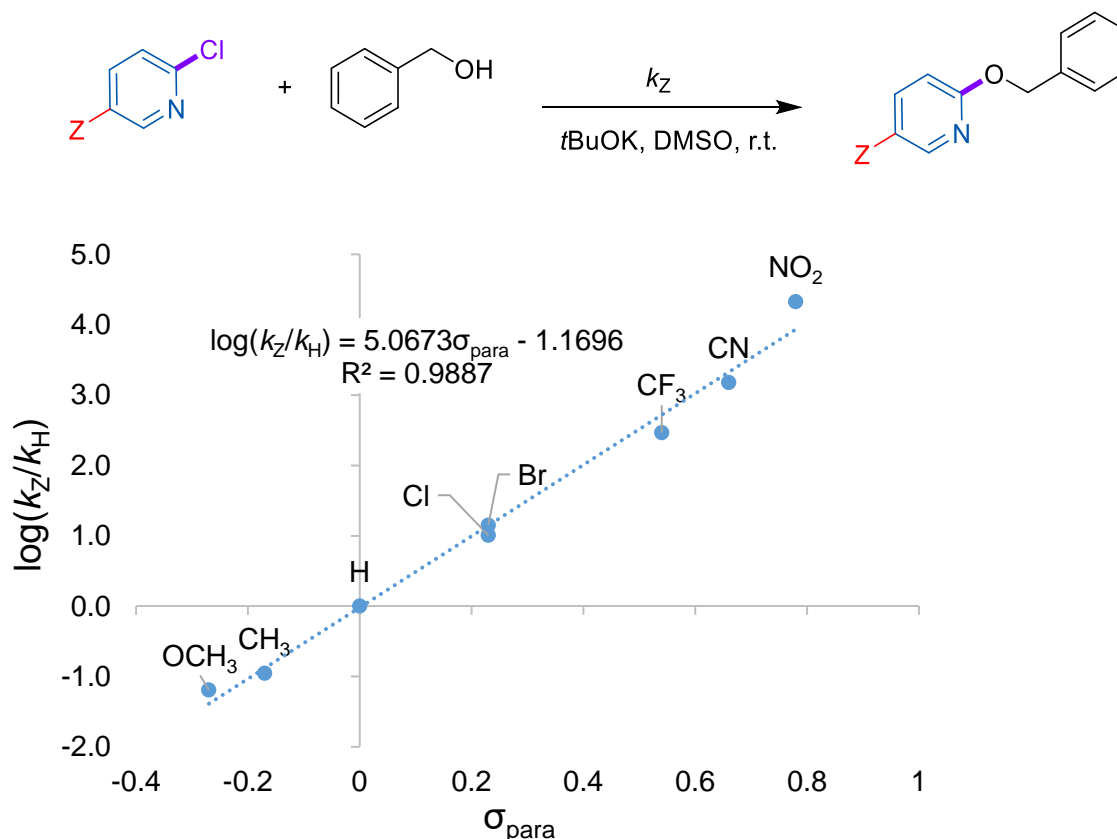


Fig. S4.

Hammett plot of $\log(k_Z/k_H)$ versus σ_{para} for $\text{S}_{\text{N}}\text{Ar}$ with benzyl alcohol of a group of 5-Z-2-chloropyridines in DMSO.

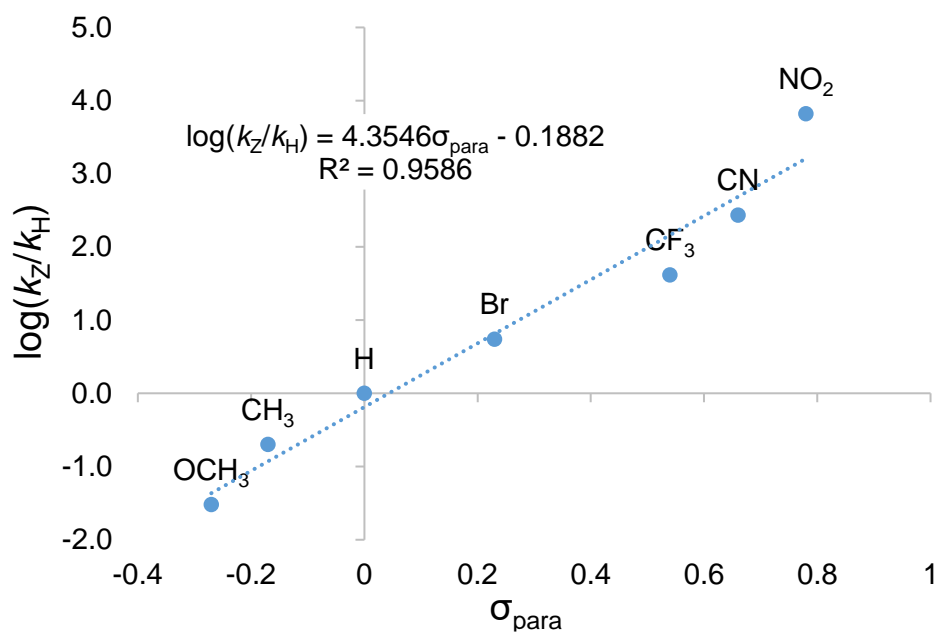
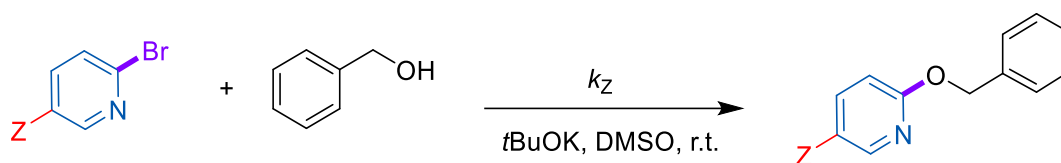


Fig. S5.

Hammett plot of $\log(k_Z/k_H)$ versus σ_{para} for S_NAr with benzyl alcohol of a group of 5-Z-2-bromopyridines in DMSO.

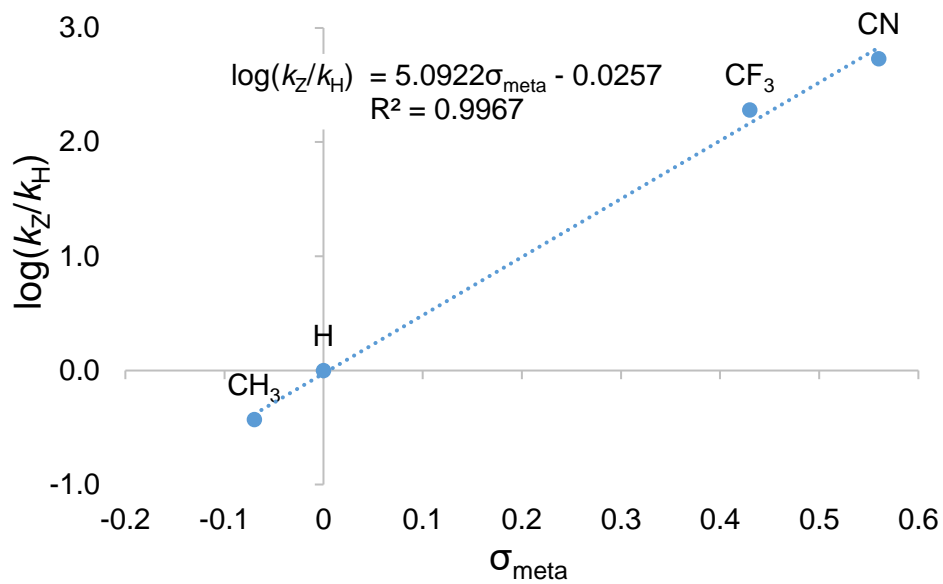
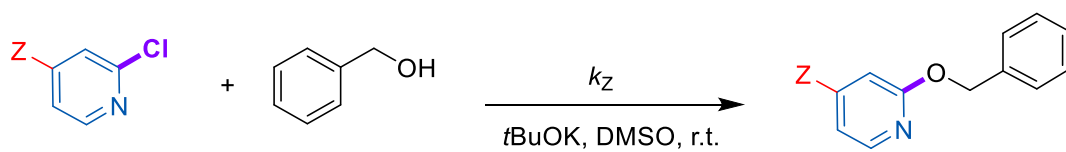


Fig. S6.

Hammett plot of $\log(k_Z/k_H)$ versus σ_{para} for S_NAr with benzyl alcohol of a group of 4-Z-2-chloropyridines in DMSO.

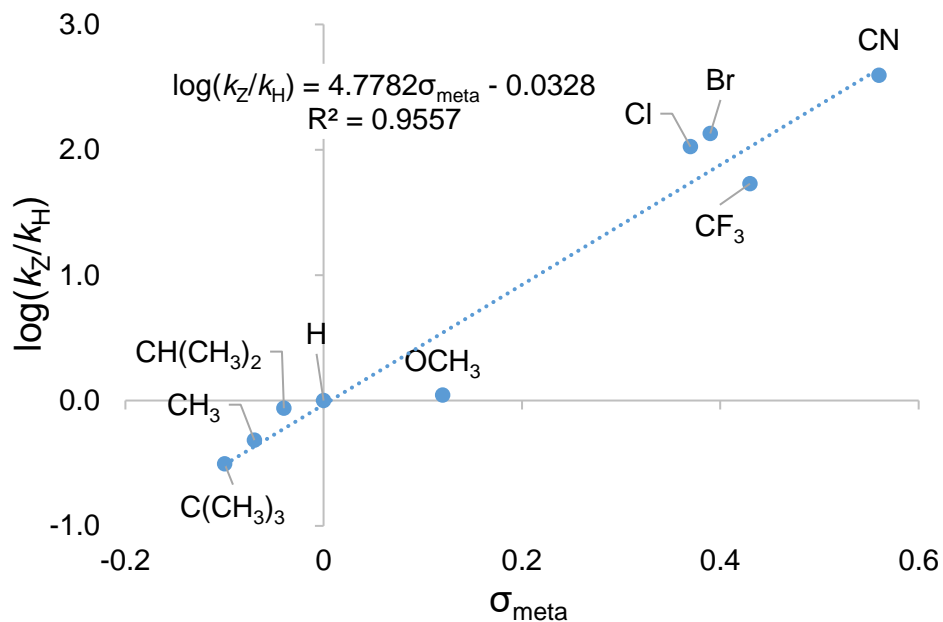
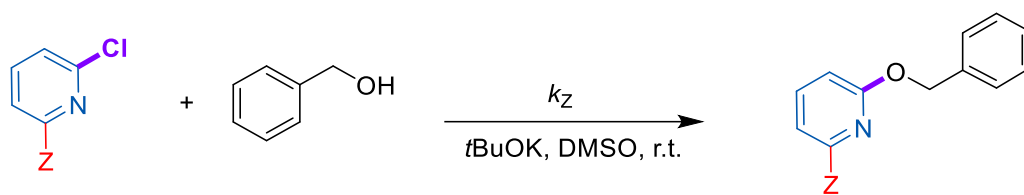


Fig. S7.

Hammett plot of $\log(k_Z/k_H)$ versus σ_{para} for S_NAr with benzyl alcohol of a group of 6-Z-2-chloropyridines in DMSO .

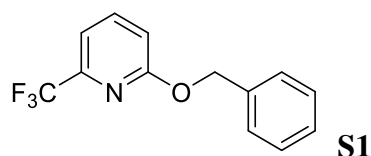
Preparative Scale Synthesis of S_NAr products

Five representative products from S_NAr reactions were isolated, purified, and characterized using the following general procedure.

In the glovebox, a 50 mL round bottom flask containing a stir bar was charged with 3 mmol of the substrate (1.0 equiv.), 3.9 mmol of benzyl alcohol (1.3 mmol) and 15 mL DMSO. 4.5 mmol *t*BuOK was dissolved into 5 mL DMSO and the solution was added dropwise into the round bottom flask. The reaction mixture was stirred at room temperature overnight. On the next day, the reaction mixture was transferred outside the glovebox for workup and isolation.

The reaction mixture was poured into 30 mL water, then extracted with ethyl acetate (3×20 mL). The combined ethyl acetate phase was washed using saturated NaCl solution (30 mL) then dried over Mg₂SO₄, filtered and evaporated under vacuum. Then the crude product was purified by automated flash chromatography (Biotage Selekt) on silica gel using hexanes/ethyl acetate to obtain the purified product.

Characterization of Isolated S_NAr products



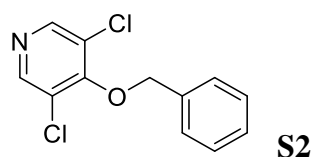
The S_NAr product **S1** was prepared by the general procedure using 2-chloro-6-(trifluoromethyl)pyridine. 307.9 mg of colorless liquid was obtained (31% yield)

¹H NMR: (500 MHz, CDCl₃): δ 5.3 (s, 2H, 2 x CH₂), 6.8 (d, 1H, 1 x Py-H), 7.1 (d, 1H, 1 x Py-H), 7.2-7.4 (m, 5H, 5 x Ph-H), 7.5-7.6 (t, 1H, 1 x Py-H).

¹³C{¹H} NMR: (126 MHz, CDCl₃): δ 68.3 (1 x CH₂), 113.4 (1 x Py), 114.9 (1 x Py), 118.2-124.8 (q, J=273.6 Hz, 1 x CF₃), 128.2-128.6 (5 x Ph), 136.7 (1 x Ph), 139.5 (1 x Py), 145.0-145.9 (q, J=34.8 Hz, 1 x Py), 163.6 (1 x Py).

¹⁹F{¹H} NMR: (471 MHz, CDCl₃): δ -68.4.

HRMS (ESI): [C₁₃H₁₀F₃NO·H]⁺ (major isotopomer, protonated): 254.07873 (calc'd), 254.07871 (found).

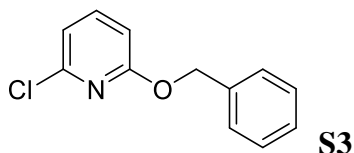


The S_NAr product **S2** was prepared by the general procedure using 3,4,5-trichloropyridine. 172.5 mg of white powder was obtained (23% yield)

¹H NMR: (500 MHz, CDCl₃): δ 5.1 (s, 2H, 2 x CH₂), 7.5-7.3 (m, 5H, 5 x Ph-H), 8.4 (s, 2H, 2 x Py-H).

¹³C{¹H} NMR: (126 MHz, CDCl₃): δ 75.4 (1 x CH₂), 126.7 (2 x Py), 128.6-128.8 (5 x Ph), 135.3 (1 x Ph), 149.2 (2 x Py), 157.5 (1 x Py).

HRMS (ESI): [C₁₂H₉Cl₂NO·H]⁺ (major isotopomer, protonated): 254.01340 (calc'd), 254.01339 (found).

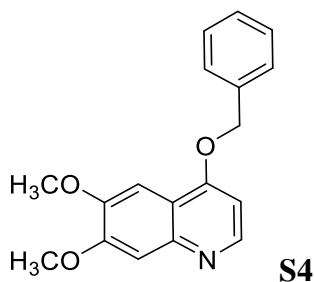


The S_NAr product **S3** was prepared by the general procedure using 2,6-dichloropyridine. 320.4 mg of colorless liquid was obtained (49% yield)

1H NMR: (500 MHz, $CDCl_3$): δ 5.3 (s, 2H, 2 x CH_2), 6.6 (dd, 1H, 1 x Py-H), 6.8 (dd, 1H, 1 x Py-H), 7.2-7.4 (m, 5H, 5 x Ph-H), 7.4-7.5 (1H 1 x Py-H).

$^{13}C\{^1H\}$ NMR: (126 MHz, $CDCl_3$): δ 68.3 (1 x CH_2), 109.4 (1 x Py), 116.5 (1 x Py), 128.1-128.5 (5 x Ph), 136.6 (1 x Ph), 140.7 (1 x Py), 148.3 (1 x Py), 163.2 (1 x Py).

HRMS (ESI): $[C_{12}H_{10}ClNO \cdot H]^+$ (major isotopomer, protonated): 220.05237 (calc'd), 220.05239 (found).

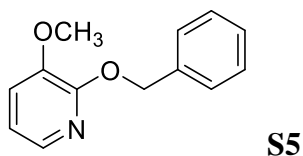


The S_NAr product **S4** was prepared by the general procedure using 4-chloro-6,7-dimethoxyquinoline. During the flash chromatography purification, the targeted product could not be eluted with hexanes/ethyl acetate gradient flow, and the flash column was flushed with 100% methanol to collect the targeted product. 212.3 mg of bright yellow crystal was obtained (24% yield).

1H NMR: (500 MHz, $CDCl_3$): δ 3.9 (s, 3H, 3 x OCH_3), 3.9 (s, 3H, 3 x OCH_3), 5.2 (s, 2H, 2 x CH_2), 6.6 (d, 1H, 1 x Qu-H), 7.3-7.4 (m, 7H, 5 x Ph-H and 2 x Qu-H), 8.5 (d, 1H, 1 x Qu-H).

$^{13}C\{^1H\}$ NMR: (126 MHz, $CDCl_3$): δ 54.1 (2 x OCH_3), 68.3 (1 x CH_2), 97.9 (1 x Qu), 98.5 (1 x Qu), 105.8 (1 x Qu), 114.0 (1 x Qu), 125.4-126.8 (5 x Ph), 134.0 (1 x Ph), 144.2 (1 x Qu), 147.1 (1 x Qu), 147.2 (1 x Qu), 150.6 (1 x Qu), 158.2 (1 x Qu).

HRMS (ESI): $[C_{18}H_{17}NO_3 \cdot H]^+$ (major isotopomer, protonated): 296.12812 (calc'd), 296.12813 (found).



The S_NAr product **S5** was prepared by the general procedure using 2-chloro-3-methoxypyridine. 168.5 mg of colorless liquid was obtained (26% yield).

1H NMR: (500 MHz, $CDCl_3$): δ 3.8 (s, 3H, 3 x OCH_3), 5.4 (2, 2H, 2 x CH_2), 6.7-6.8 (dd, 1H, 1 x Py-H), 7.0 (d, 1H, 1 x Py-H), 7.2-7.4 (m, 5H, 5 x Ph-H), 7.6-7.7 (d, 1H, 1 x Py-H).

$^{13}C\{^1H\}$ NMR: (126 MHz, $CDCl_3$): δ 55.7 (1 x OCH_3), 67.5 (1 x CH_2), 117.0 (1 x Py), 117.5 (1 x Py), 127.7-128.4 (5 x Ph), 137.0 (1 x Ph), 137.4 (1 x Py), 144.2 (1 x Py), 153.9 (1 x Py).

HRMS (ESI): $[C_{13}H_{13}NO_2 \cdot H]^+$ (major isotopomer, protonated): 216.10191 (calc'd), 216.10193 (found).

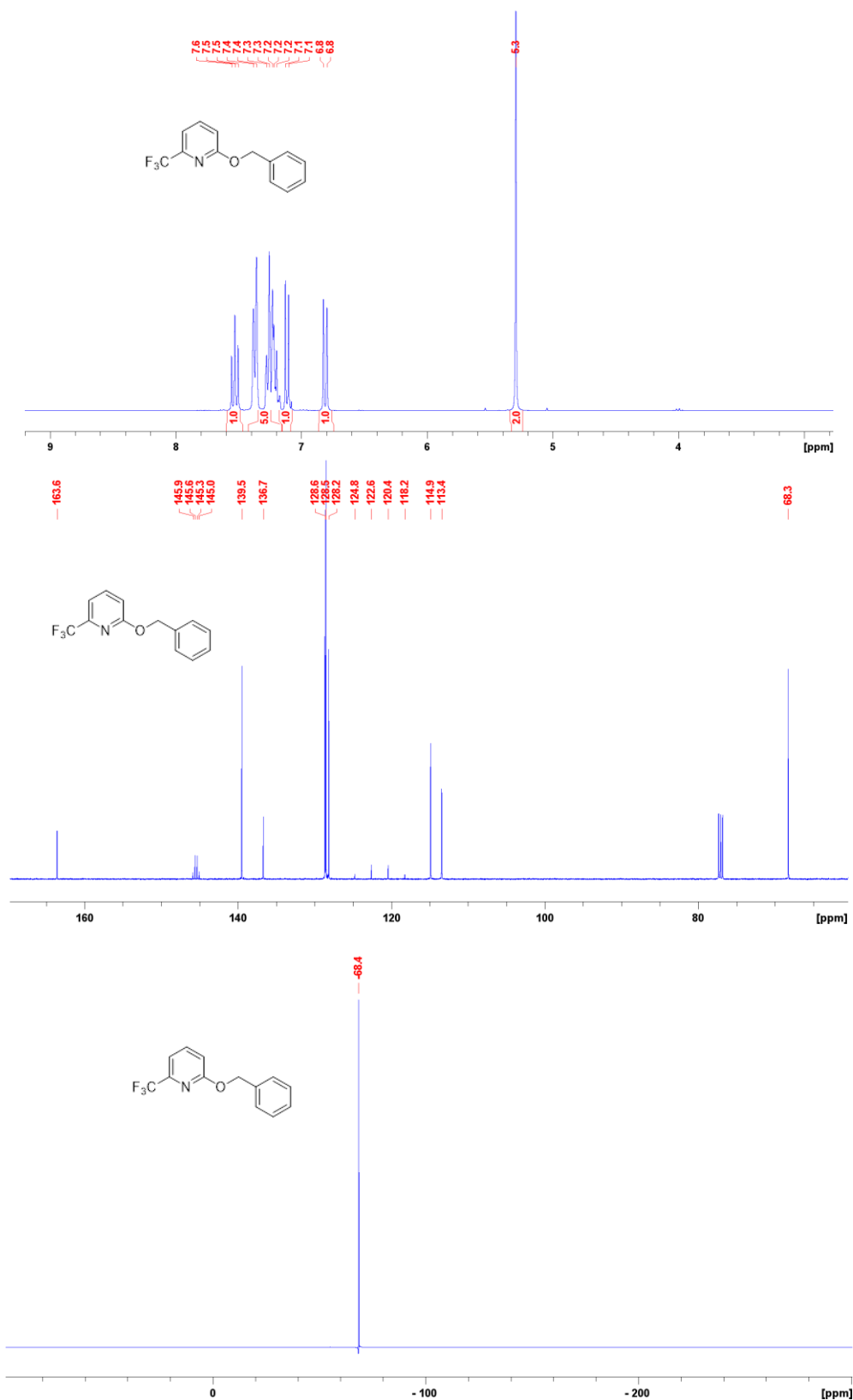


Fig. S8. ^1H (500 MHz, CDCl_3), $^{13}\text{C}\{^1\text{H}\}$ (126 MHz, CDCl_3) and $^{19}\text{F}\{^1\text{H}\}$ NMR: (471 MHz, CDCl_3) NMR spectra of **S1**.

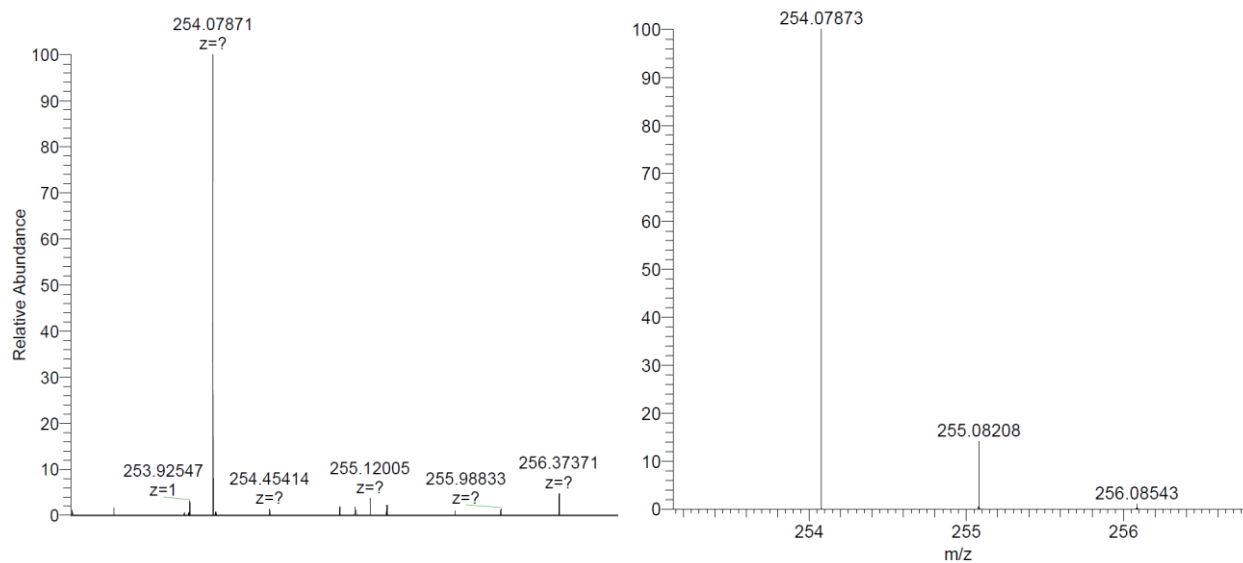


Fig. S9.

Left: Experimental HRMS-ESI spectrum of $[S1+H]^+$. Right: Calculated HRMS isotope pattern for $[S1+H]^+$.

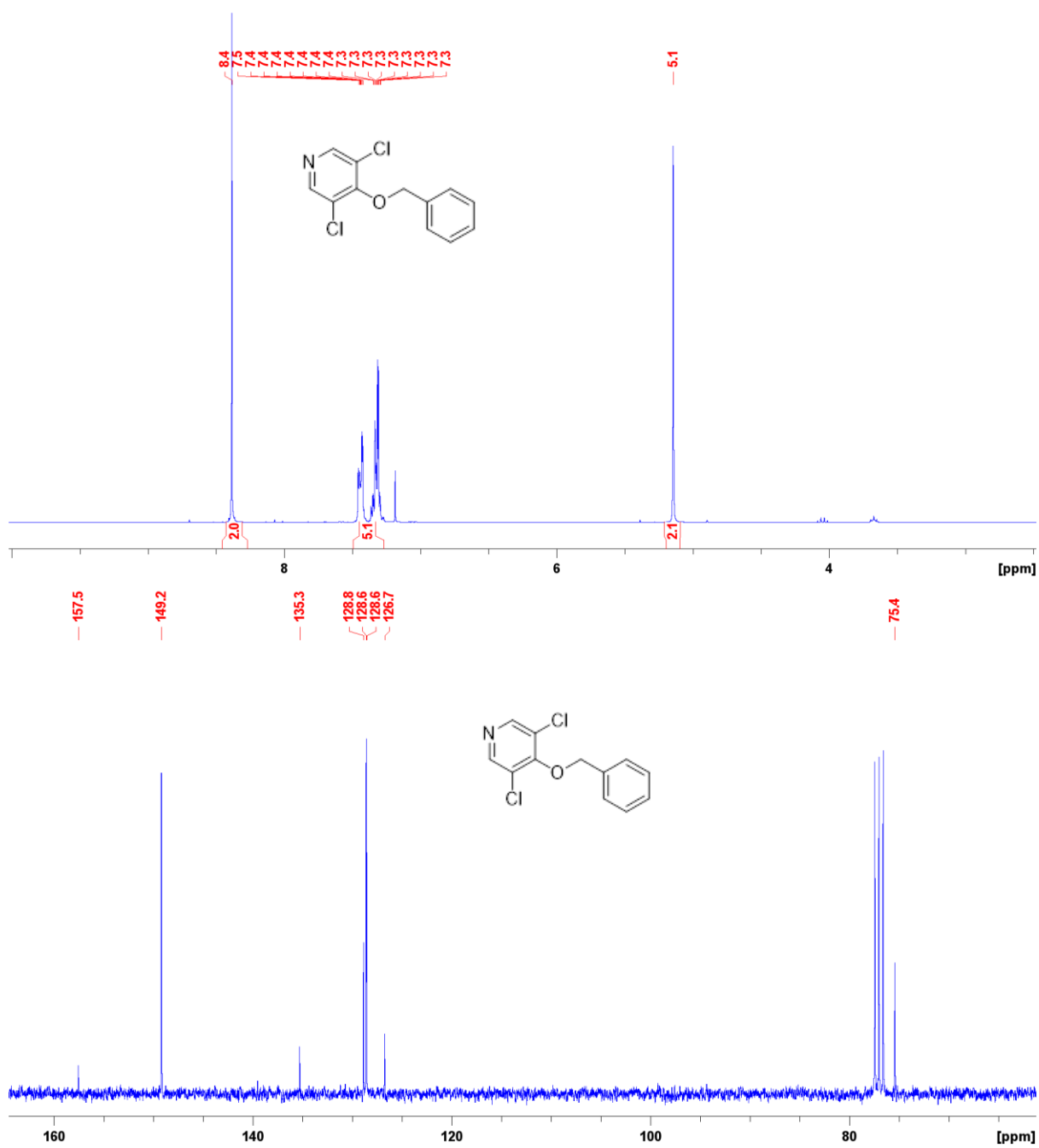


Fig. S10.
¹H (500 MHz, CDCl₃), and ¹³C{¹H} (126 MHz, CDCl₃) NMR spectra of S2.

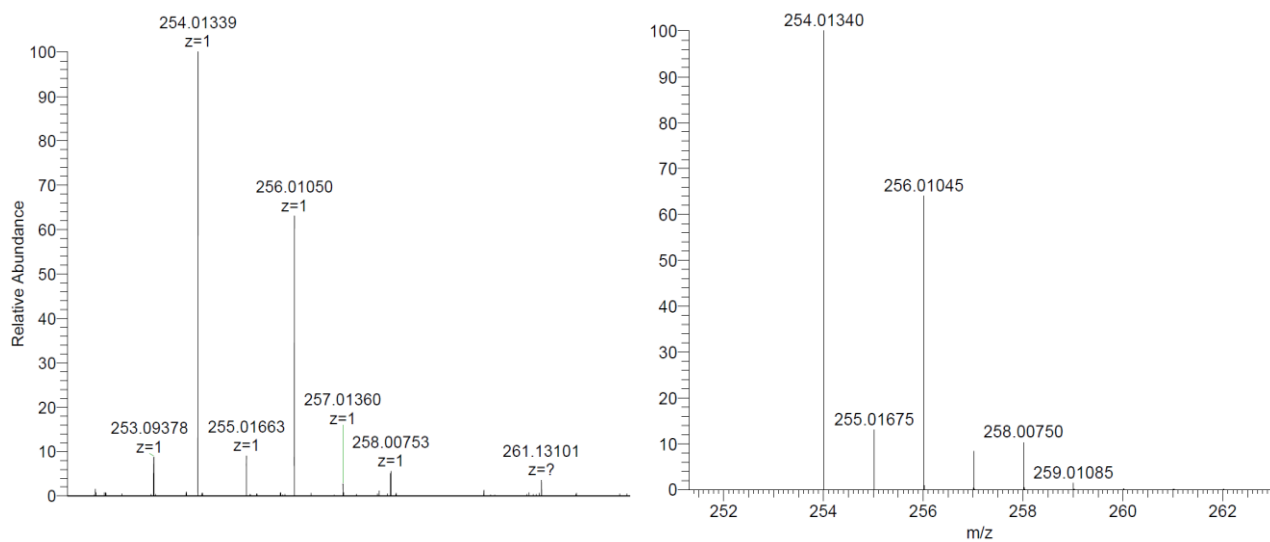


Fig. S11.

Left: Experimental HRMS-ESI spectrum of $[S2+H]^+$. Right: Calculated HRMS isotope pattern for $[S2+H]^+$.

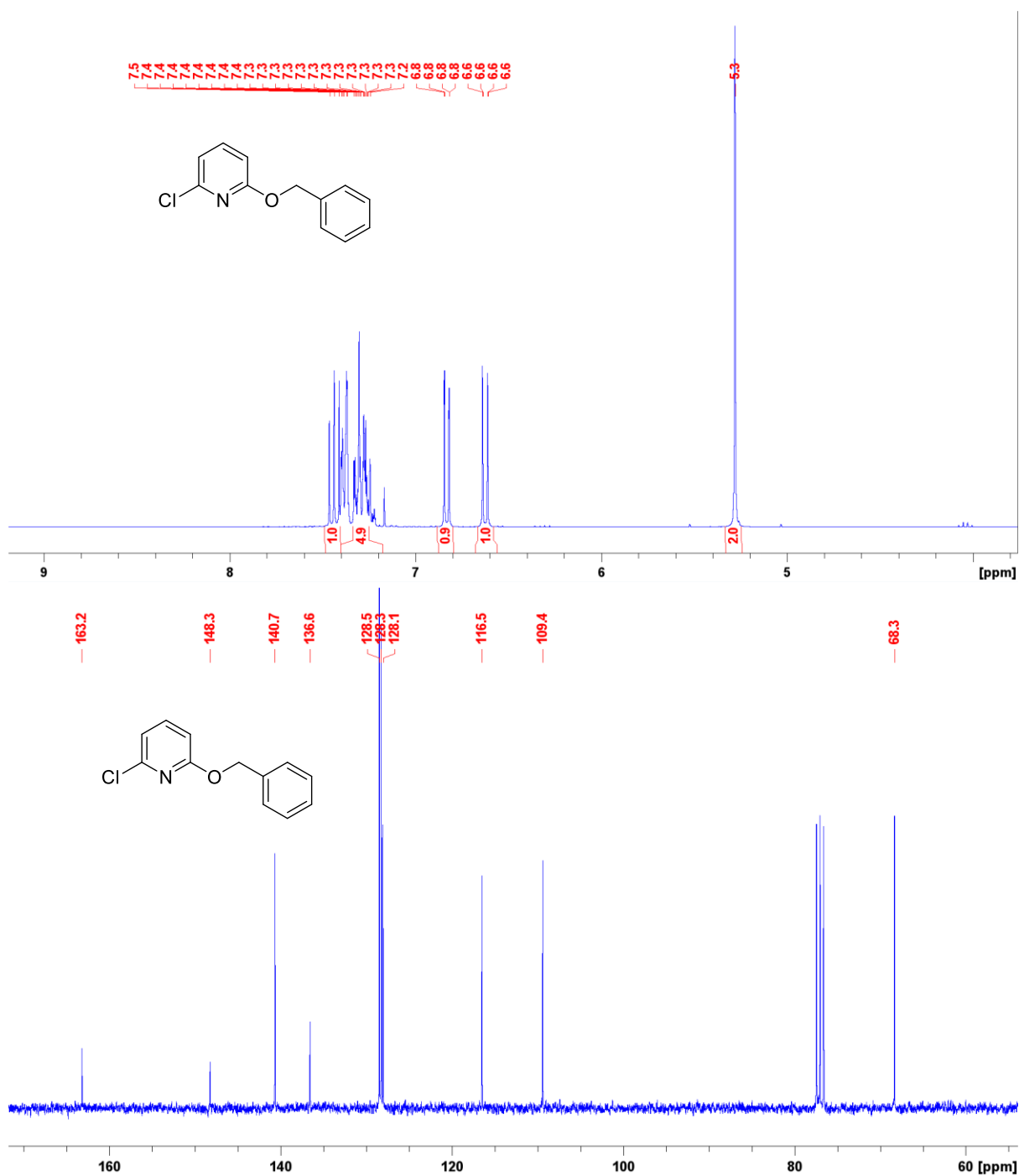


Fig. S12. ¹H (500 MHz, CDCl₃), and ¹³C{¹H} (126 MHz, CDCl₃) NMR spectra of S3.

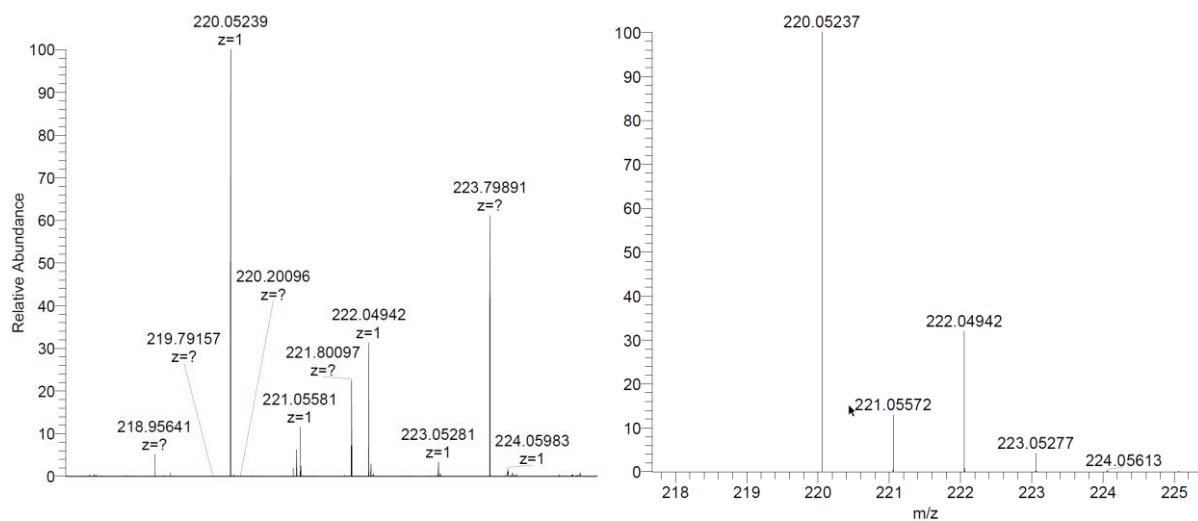


Fig. S13.
 Left: Experimental HRMS-ESI spectrum of $[S3+H]^+$. Right: Calculated HRMS isotope pattern for $[S3+H]^+$.

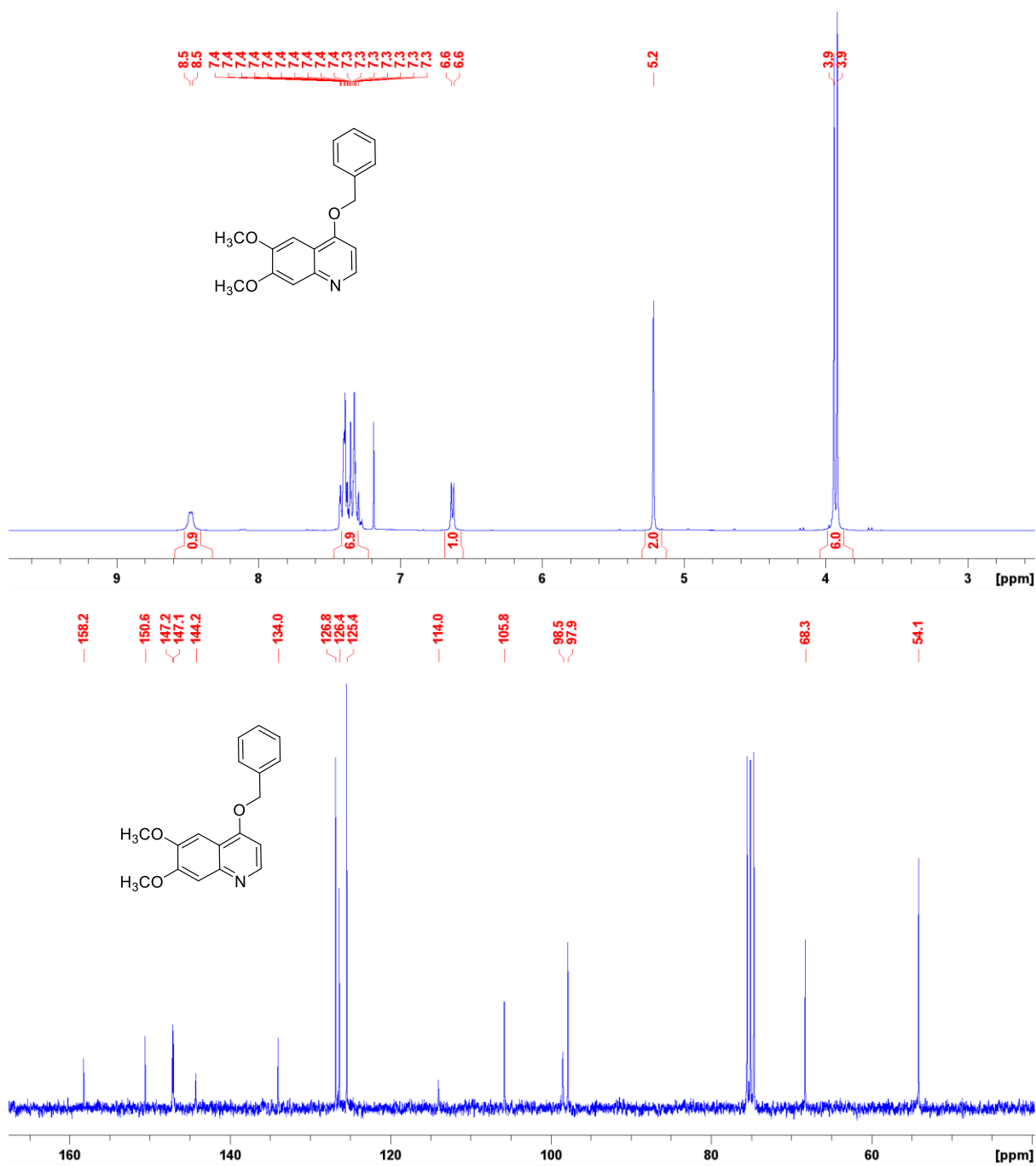


Fig. S14. ^1H (500 MHz, CDCl_3), and $^{13}\text{C}\{^1\text{H}\}$ (126 MHz, CDCl_3) NMR spectra of S4.

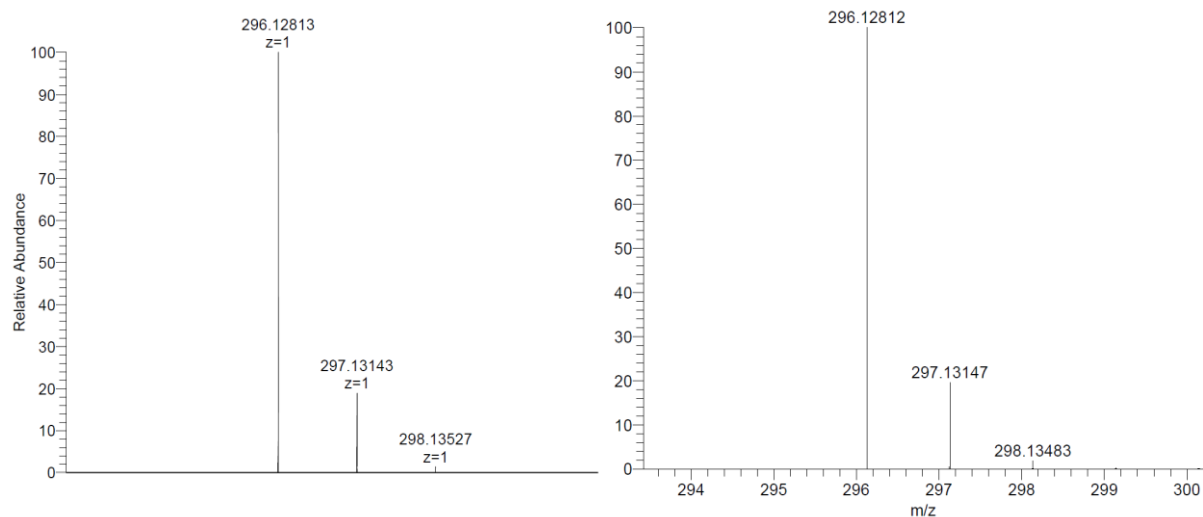


Fig. S15.

Left: Experimental HRMS-ESI spectrum of $[S4+H]^+$. Right: Calculated HRMS isotope pattern for $[S4+H]^+$.

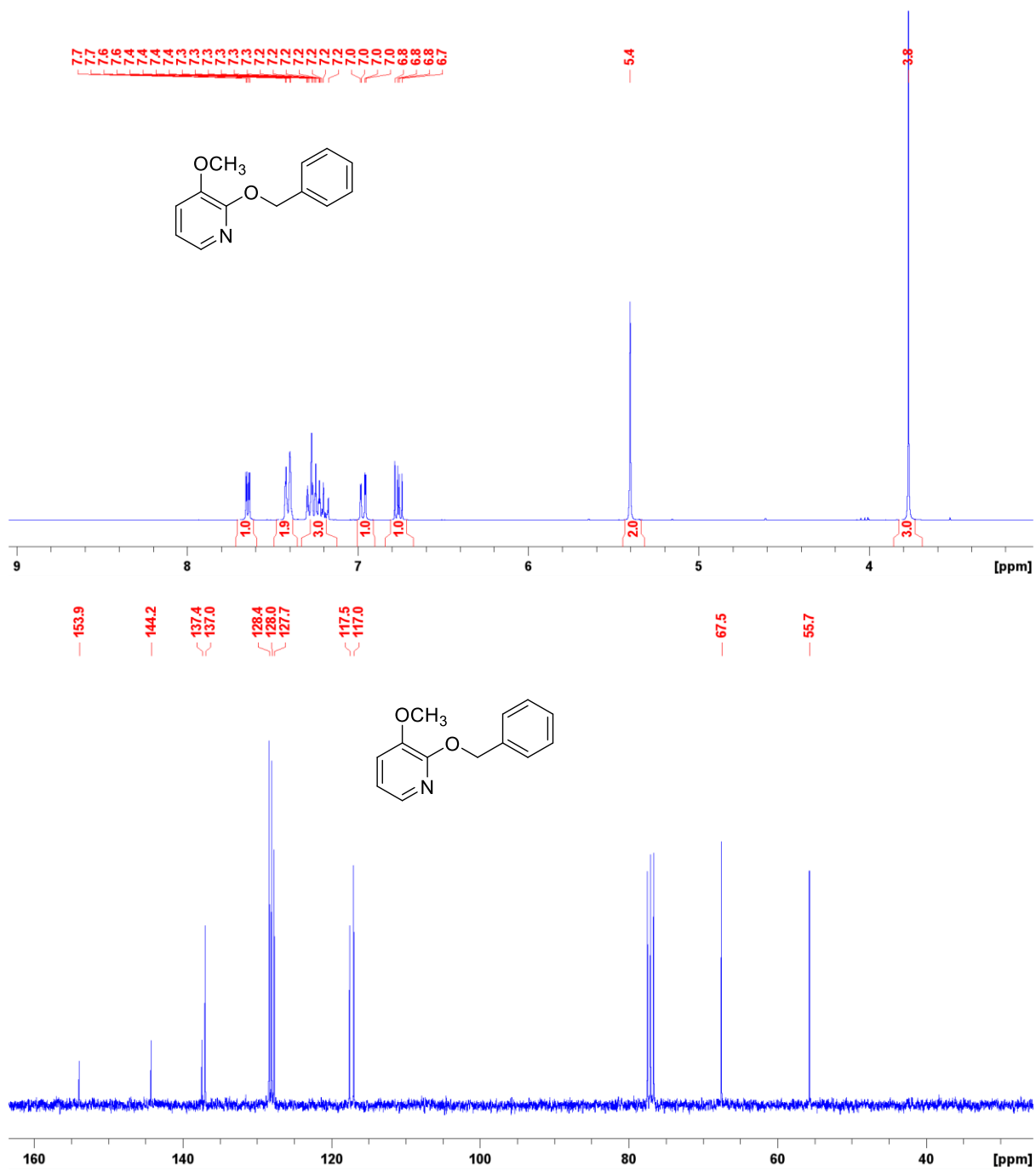


Fig. S16.
 ^1H (500 MHz, CDCl_3), and $^{13}\text{C}\{^1\text{H}\}$ (126 MHz, CDCl_3) NMR spectra of S5.

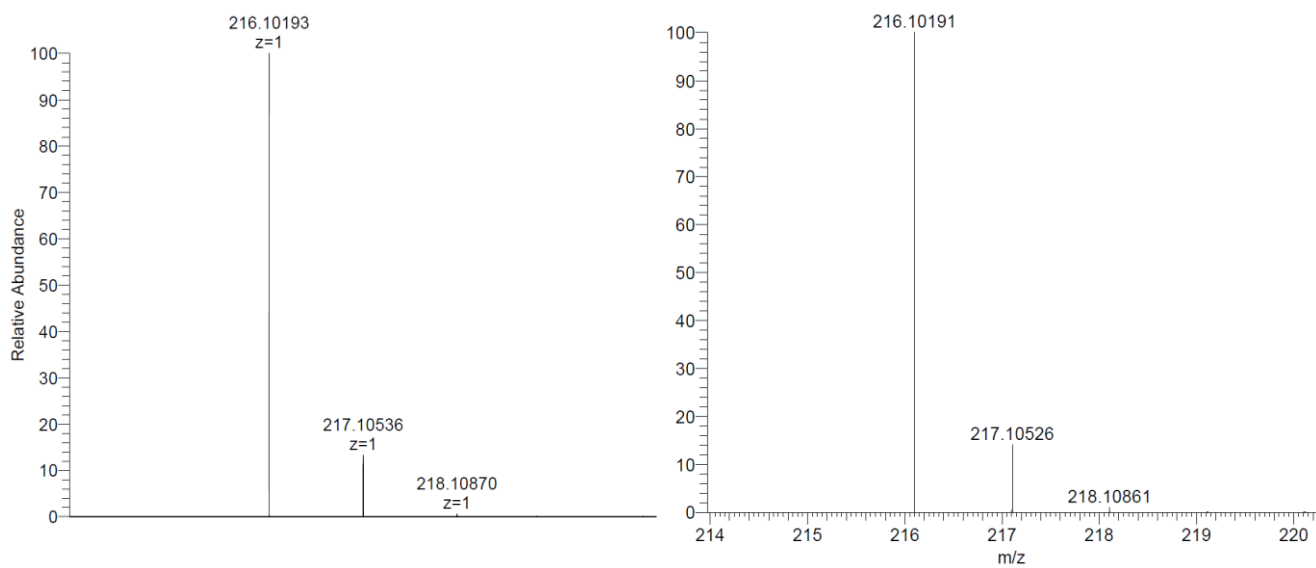


Fig. S17. Left: Experimental HRMS-ESI spectrum of $[\mathbf{S5}+\mathbf{H}]^+$. Right: Calculated HRMS isotope pattern for $[\mathbf{S5}+\mathbf{H}]^+$.

Computational Determination of Molecular Descriptors

All molecular geometry optimizations were performed using ORCA version 4.0.1.2⁶ and all electrostatic potential (*ESP*) calculations were performed using Multiwfn version 3.7^{7,8}.

Initial substrate structures were either downloaded from the ChemSpider database⁹ or generated using Avogadro version 4.1.¹⁰ Geometry optimizations were performed for the uncharged species (N-electron system) using a B3LYP/def2-TVZPD approach. The energy of the anionic species (N+1-electron system) was calculated at the optimized geometry of the uncharged species.

Using the conceptual link between *LUMO* energy and electron affinity (*EA*) related to Koopmans's theorem¹¹, we calculated *EA* for each substrate according to:

$$EA = E(N) - E(N + 1) \sim -\epsilon(LUMO) \quad Eq(S7)$$

The corrections to the *EA* values for substrates in the main text Fig. 5 that proceed through overlap of the *LUMO*+1 orbitals were conducted by adding the orbital energy difference between the *LUMO*+1 and *LUMO*, as output by the DFT calculations, to the *EA* values.

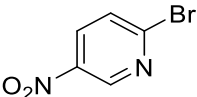
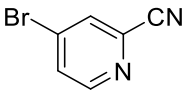
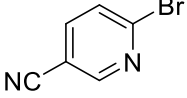
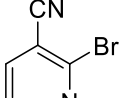
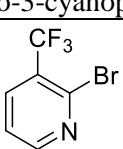
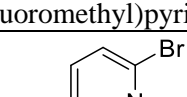
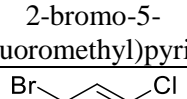
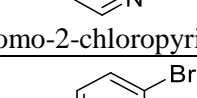
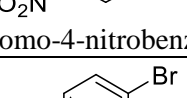
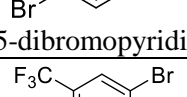
A *.gbw* file that contains a binary summary of the calculation was generated automatically during the execution of ORCA, then transformed to a *.wfn* file by the utility program ORCA_2aim. The *.wfn* file contains electron density information and can be used for molecular surface analysis by the Multiwfn program. The utility program ORCA_2mkl was used to generate a *.fch* file, containing information on the full range of calculated MOs from the *.gbw* file.

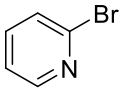
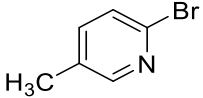
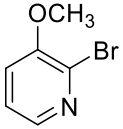
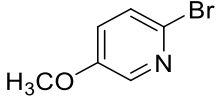
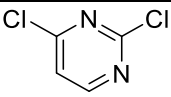
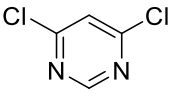
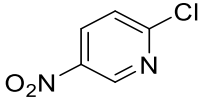
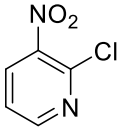
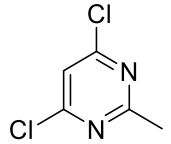
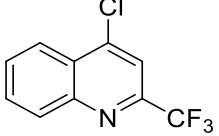
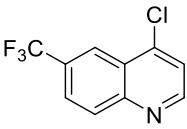
The *.wfn* file was imported into Multiwfn for *ESP* calculations. The molecular *ESP* was calculated at a 0.004 au isosurface of electron density. By selecting *Output surface properties of each atom*, the surface area occupied by each atom, as well as the maximal, minimal and average *ESP* values at that surface were displayed. The average *ESP* values at the reactive center (*ESP*₁) and sum of the *ESP* values at the ortho and para atoms in relation to the reactive center (*ESP*₂) were used as the electronic descriptors in construction of the S_NAr predictive model.

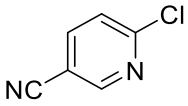
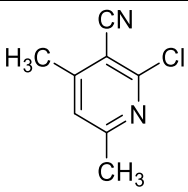
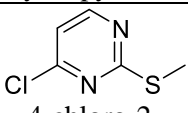
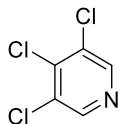
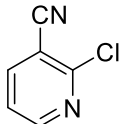
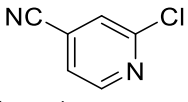
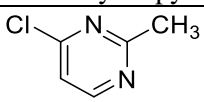
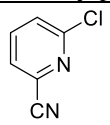
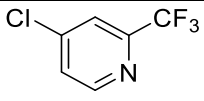
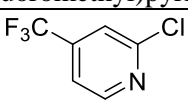
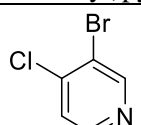
The electrostatic potential maps and molecular orbitals were plotted and rendered in VMD, a molecular visualization program.¹² To do so, the electron density cube file and the *ESP* cube file were generated in Multiwfn from the *.wfn* file. The two cube files can be read by VMD and the *ESP* map can be plotted, by projecting the *ESP* values at each grid point to a defined electron density isosurface. To obtain molecular orbital pictures, orbital cube files were generated in Multiwfn using *.fch* file as input, then plotted in VMD at a defined electron density isosurface.

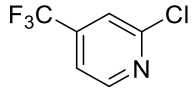
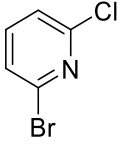
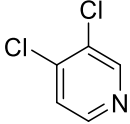
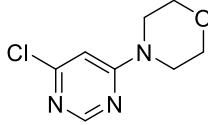
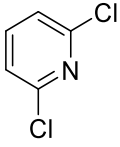
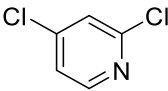
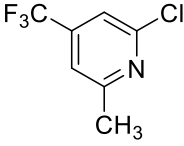
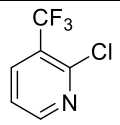
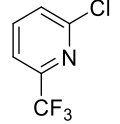
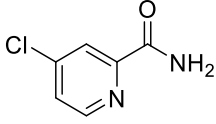
Table S6.

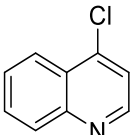
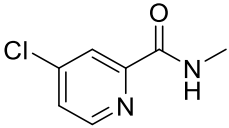

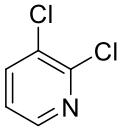
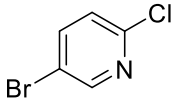
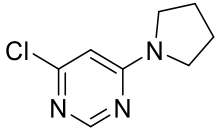
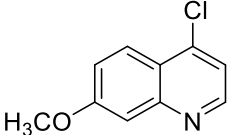
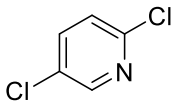
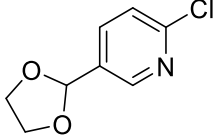
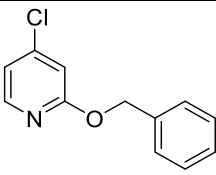
Calculated molecular descriptors for the 74 (hetero)aryl halide substrates used in S_NAr experiments.

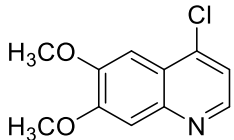
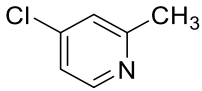
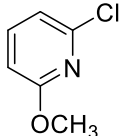
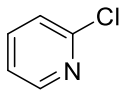
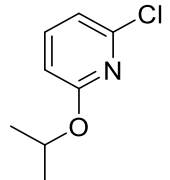
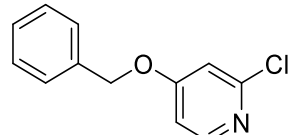
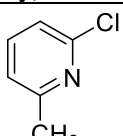
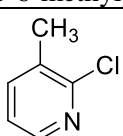
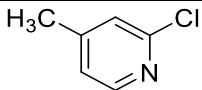
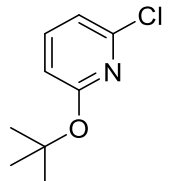
Substrate	Leaving group	Electron Affinity (EA) (kJ/mol)	Average ESP_1 (kJ/mol)	Average ESP_2 (kJ/mol)
 2-bromo-5-nitropyridine	Br	127.8	98.66	200.30
 4-bromo-2-cyanopyridine	Br	54.4	98.14	141.97
 2-bromo-5-cyanopyridine	Br	52.8	90.68	167.00
 2-bromo-3-cyanopyridine	Br	52.5	88.83	150.75
 2-bromo-3-(trifluoromethyl)pyridine	Br	6.5	74.41	110.57
 2-bromo-5-(trifluoromethyl)pyridine	Br	55.1	74.39	123.51
 4-bromo-2-chloropyridine	Br	-16.4	68.43	48.24
 1-bromo-4-nitrobenzene	Br	101.6	72.67	212.67
 2,5-dibromopyridine	Br	-11.4	57.10	68.66
 1-bromo-3,5-bis(trifluoromethyl)benzene	Br	17.2	81.76	239.30

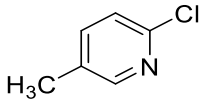
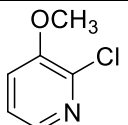
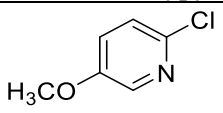
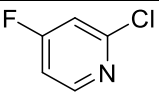
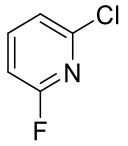
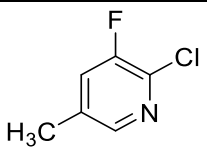

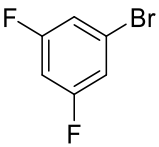
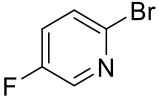
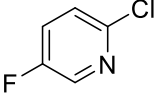
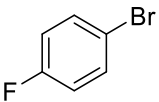
 2-bromopyridine	Br	-54.0	35.53	-11.39
 2-bromo-5-methylpyridine	Br	-56.9	24.37	-17.20
 2-bromo-3-methoxypyridine	Br	-50.1	34.40	-2.06
 2-bromo-5-methoxypyridine	Br	-53.9	19.06	-8.79
 2,4-dichloropyrimidine	Cl	12.6	101.17	17.30
 4,6-dichloropyrimidine	Cl	4.6	99.15	24.34
 2-chloro-5-nitropyridine	Cl	123.2	105.30	196.46
 2-chloro-3-nitropyridine	Cl	113.9	101.82	176.89
 4,6-dichloro-2-methylpyrimidine	Cl	-2.3	86.08	-0.84
 4-chloro-2-(trifluoromethyl)quinoline	Cl	69.5	75.27	69.50
 4-chloro-6-(trifluoromethyl)quinoline	Cl	74.3	74.17	59.38

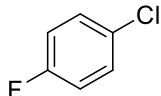
 2-chloro-5-cyanopyridine	Cl	46.3	96.68	163.10
 2-chloro-4,6-dimethyl-3-cyanopyridine	Cl	24.6	74.40	96.11
 4-chloro-2-(methylthio)pyrimidine	Cl	-16.8	67.66	-38.07
 3,4,5-trichloropyridine	Cl	-2.9	90.18	108.80
 2-chloro-3-cyanopyridine	Cl	49.0	95.86	150.24
 2-chloro-4-cyanopyridine	Cl	69.9	91.93	149.60
 4-chloro-2-methylpyrimidine	Cl	-31.8	68.02	-60.39
 2-chloro-6-cyanopyridine	Cl	54.5	91.91	147.58
 4-chloro-2-(trifluoromethyl)pyridine	Cl	3.1	88.14	96.13
 2-chloro-5-(trifluoromethyl)pyridine	Cl	0.6	80.54	117.45
 4-chloro-3-bromopyridine	Cl	15.0	76.32	102.23

 2-chloro-4-(trifluoromethyl)pyridine	Cl	-21.2	64.00	54.86
 2-chloro-6-bromopyridine	Cl	-25.1	70.99	42.09
 3,4-dichloropyridine	Cl	-28.9	71.48	46.38
 4-(6-chloropyrimidin-4-yl)morpholine	Cl	-46.3	39.22	-98.45
 2,6-dichloropyridine	Cl	-4.7	66.41	76.70
 2,4-dichloropyridine	Cl	-28.0	62.58	49.83
 2-chloro-6-methyl-4-(trifluoromethyl)pyridine	Cl	-23.2	71.98	46.53
 2-chloro-3-(trifluoromethyl)pyridine	Cl	-5.0	77.18	105.19
 2-chloro-6-(trifluoromethyl)pyridine	Cl	4.7	77.82	100.65
 4-chloropyridine-2-carboxamide	Cl	-5.7	70.82	25.19

 4-chloro-quinoline	Cl	27.3	44.28	-33.90
 4-chloro-N-Methylpyridine-2-carboxamide	Cl	8.1	68.47	1.89
 2-chloro-3-bromopyridine	Cl	-23.2	59.81	45.92
 2,3-dichloropyridine	Cl	-28.0	60.07	49.43
 2-chloro-5-bromopyridine	Cl	-17.4	61.37	63.41
 4-chloro-6-(pyrrolidin-1-yl)pyrimidine	Cl	-56.9	22.65	-161.21
 4-chloro-7-methoxyquinoline	Cl	13.8	32.95	-64.56
 2,5-dichloropyridine	Cl	-22.2	60.40	65.10
 2-chloro-5-(1,3-dioxolan-2-yl)pyridine	Cl	-45.3	29.87	-32.64
 2-(benzyloxy)-4-chloropyridine	Cl	-28.2	42.62	-47.42

 4-chloro-6,7-dimethoxyquinoline	Cl	3.7	21.63	-97.34
 4-chloro-2-methylpyridine	Cl	-58.9	39.32	-45.53
 2-chloro-6-methoxypyridine	Cl	-73.0	29.09	-30.54
 2-chloropyridine	Cl	-60.8	38.82	-18.34
 2-chloro-6-isopropoxypyridine	Cl	-68.5	25.20	-38.78
 4-(benzyloxy)-2-chloropyridine	Cl	-40.8	30.25	-45.79
 2-chloro-6-methylpyridine	Cl	-62.7	29.53	-41.13
 2-chloro-3-methylpyridine	Cl	-68.5	34.39	-28.53
 2-chloro-4-methylpyridine	Cl	-67.5	28.98	-42.60
 2-chloro-6-(<i>tert</i> -butoxy)pyridine	Cl	-75.3	23.28	-26.10

 2-chloro-5-methylpyridine	Cl	-63.4	27.60	-24.94
 2-chloro-3-methoxypyridine	Cl	-71.4	25.55	-28.55
 2-chloro-5-methoxypyridine	Cl	-58.3	24.22	-15.54
 2-chloro-4-fluoropyridine	F	-56.9	84.42	44.08
 2-chloro-6-fluoropyridine	F	-39.7	80.25	44.12
 2-chloro-3-fluoro-5-methylpyridine	F	-38.2	70.42	142.55
 2-chloro-3-fluoropyridine	F	-40.2	79.83	166.13
 1-bromo-3,5-difluorobenzene	F	-61.0	64.00	119.66
 2-bromo-5-fluoropyridine	F	-25.6	81.20	177.91
 2-chloro-5-fluoropyridine	F	-32.0	78.99	175.28
 1-bromo-4-fluorobenzene	F	-64.2	43.98	70.00

 1-chloro-4-fluorobenzene	F	-82.9	42.12	68.86
---	---	-------	-------	-------

Construction of the Multivariate Linear Regression Model

The selection of the molecular descriptors used to correlate $\Delta G^{\ddagger}_{\text{S}_{\text{N}}\text{Ar}}$ with substrate structures was guided by the mechanistic features of nucleophilic aromatic substitution,^{13,14} DFT calculations on the transition states, and iterative refinement of the included descriptors based on our experimental observations. As summarized in Fig. S18, there are three descriptors that lead to accurate predictions of $\text{S}_{\text{N}}\text{Ar}$ reactivity: *Electron affinity* (*EA*) of the substrate (a measurable molecular property to approximate the *LUMO* energy), average electrostatic potential (*ESP*₁) at the reactive carbon, sum of the average electrostatic potential at the *ortho* and *para* atoms in relation to the reactive carbon (*ESP*₂). Linear regression analysis with the normalized descriptors has been performed to obtain the contribution of the individual descriptor to the reactivity of the substrate, and the results are also summarized in Fig. S19.

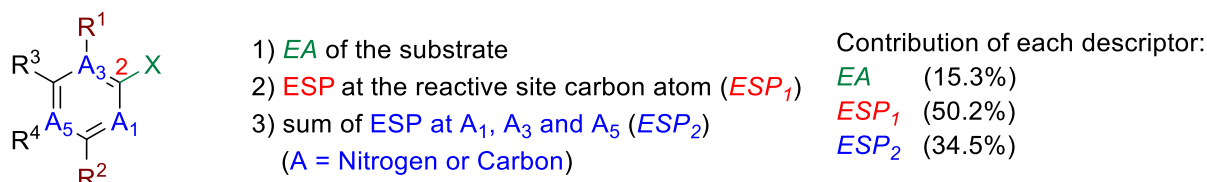


Fig. S18.

Substrate molecular descriptors to construct the nucleophilic aromatic substitution ($\text{S}_{\text{N}}\text{Ar}$) predictive model.

The initial multivariate linear regression model constructed from these descriptors and the $\Delta G^{\ddagger}_{\text{S}_{\text{N}}\text{Ar}}$ of 74 (hetero)aryl halides in DMSO achieved excellent linear correlation with a squared correlation coefficient (R^2) of 0.92, and a mean absolute error (MAE) of 1.87 kJ/mol, and has an expected random distribution of residuals as summarized in Fig. S19.

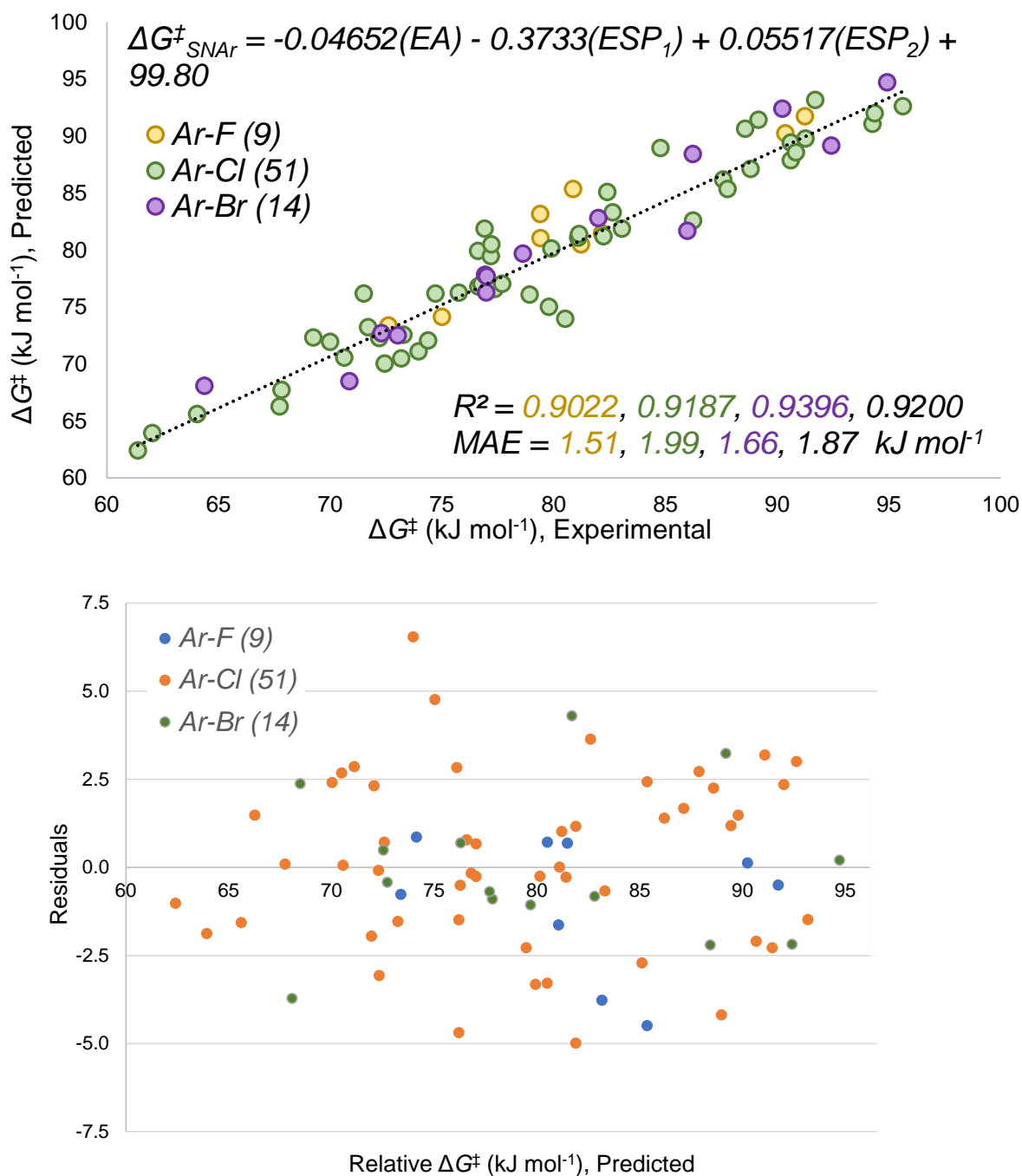


Fig. S19.

Multivariate linear regression model of $\Delta G^{\ddagger}_{SNAr}$ for 74 substrates, including experimental versus predicted plot (top) and predicted versus residuals plot (bottom).

To validate the inclusion of these descriptors and evaluate simpler potential models, we have compared the three-descriptor model to a series of alternatives (Table S7).

First, we evaluated three univariate linear regression models constructed by each individual descriptor. The linear regression with ESP_1 gives fairly good performance with an R^2 of 0.77 and an MAE of 3.0 kJ/mol, as shown in entry 2. In contrast, the univariate models with the other two descriptors only give a weak linear relationship (EA , entry 4) and almost no linear relationship (ESP_2 , entry 3).

ESP_1 is the descriptor that contributes the most to predicting S_NAr reactivity among the three, as indicated by the performance of the univariate models. Then we evaluated two bivariate linear regression models as shown in entries 5 and 6. Compared to the univariate model in entry 2, the inclusion of one more descriptor, either EA or ESP_2 improves the prediction accuracy, but still in a slightly lower level compared to the prediction accuracy of the multivariate model with all the three descriptors included.

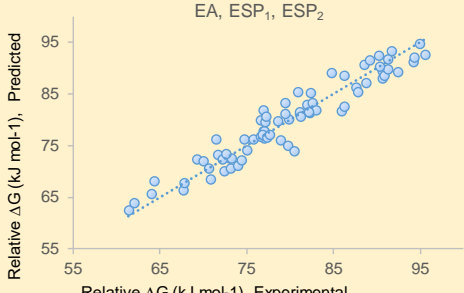
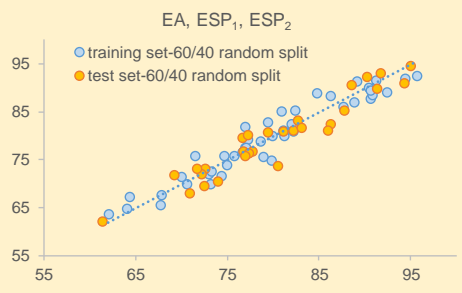
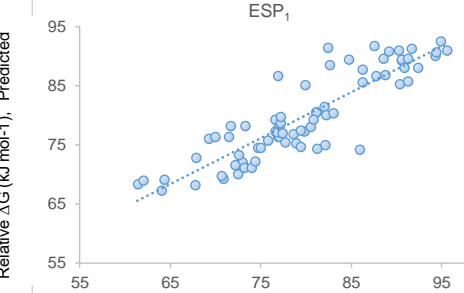
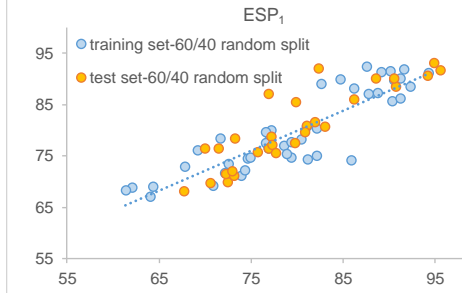
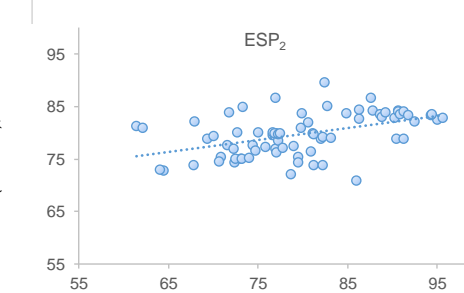
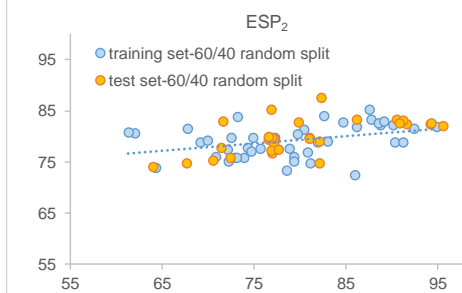
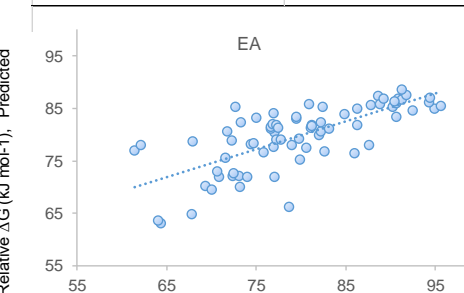
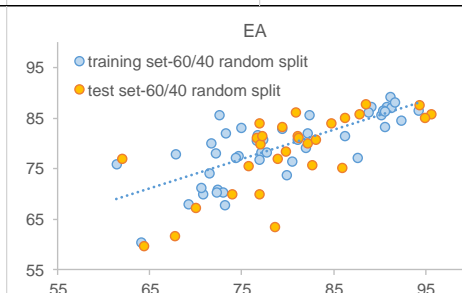
The inclusion of ESP_2 (sum of the ESP at the *ortho* and *para* atoms) in the model was explored based on our DFT transition state calculations and the experimental observations. From our iterative analysis of the linear regression model, we discovered that the two ESP values – ESP at the *ortho* atoms and the ESP at the *para* atom – have approximately equal contributions when linear regression fitting is done with ESP_{ortho} and ESP_{para} as separate descriptors. This alternate, four-descriptor model is shown in entry 7. An R^2 of 0.92 and a MAE of 1.9 kJ/mol were obtained from the four-descriptor linear regression model, which are almost identical to those obtained from the original three-variable regression model (entry 1). The two models give the same level of prediction accuracy, and thus the ESP_{ortho} and ESP_{para} are treated as one summed value for simplification.

The model with steric effect as an additional descriptor (steric A value¹⁵) has also been evaluated. The statistic measures in the entry 8 show that the prediction accuracy is almost identical compared to the initial model shown in Fig. S19. Linear regression analysis with normalized descriptors shows that the contribution of the steric A value is only 2%, which is much lower than any other descriptor. As a result, the steric descriptor is not included because of its insignificant impact to the reactivity of the substrates in our training dataset.

Overall, the alternatives presented in Table S7 demonstrate that all of the three descriptors are necessary and sufficient to predict $\Delta G^\ddagger_{S_NAr}$ for this diverse set of substrates.

Table S7.

Comparison of model performance for different combinations of molecular descriptors.

Comparison of Model Performance				
Variables	Multivariate linear regression model performance		60/40 random split performance for test set	
	R^2	MAE (kJ/mol)	Q^2	MAE (kJ/mol)
EA, ESP ₁ , ESP ₂	0.92	1.90	0.91	2.0
EA, ESP ₁ , ESP ₂				
ESP ₁	0.77	3.0	0.82	2.6
ESP ₁				
ESP ₂	0.23	5.8	0.36	5.50
ESP ₂				
EA	0.53	4.4	0.45	4.8
EA				

EA, ESP ₁	R^2	MAE (kJ/mol)	Q^2	MAE (kJ/mol)
	0.80	2.9	0.82	2.7
ESP ₁ , ESP ₂	R^2	MAE (kJ/mol)	Q^2	MAE (kJ/mol)
	0.88	2.4	0.92	2.0
EA, ESP ₁ , ESP _{2-ortho} , ESP _{2-para}	R^2	MAE (kJ/mol)	Q^2	MAE (kJ/mol)
	0.92	1.9	0.91	1.9
EA, Steric A value ESP ₁ , ESP ₂	R^2	MAE (kJ/mol)	Q^2	MAE (kJ/mol)
	0.92	1.9	0.88	2.2

Cross Validation and Out-of-Sample Prediction

To further evaluate the linear regression model, we performed cross-validation by doing five random 60/40 training/test data splits (Figs S20-S24). Excellent linear correlation was achieved between the observed and predicted $\Delta G^{\ddagger}_{\text{SNAr}}$, as indicated by the range of R^2 from 0.89 to 0.93 for training set, Q^2 from 0.86 to 0.93, and MAE from 1.63 kJ/mol to 2.30 kJ/mol for test set. The good agreement between the observed and predicted $\Delta G^{\ddagger}_{\text{SNAr}}$ obtained from this random split cross-validation has indicated that our multivariate linear regression model is appropriately fitted, with no overfitting issues.

We also calculated the 95% prediction intervals of the predicted $\Delta G^{\ddagger}_{\text{SNAr}}$ for the test set for one of the 60/40 random split model (Fig. S20). The narrow prediction intervals (± 5.1 kJ/mol to ± 5.5 kJ/mol) indicate that the model makes reliable prediction of the S_{NAr} rates.

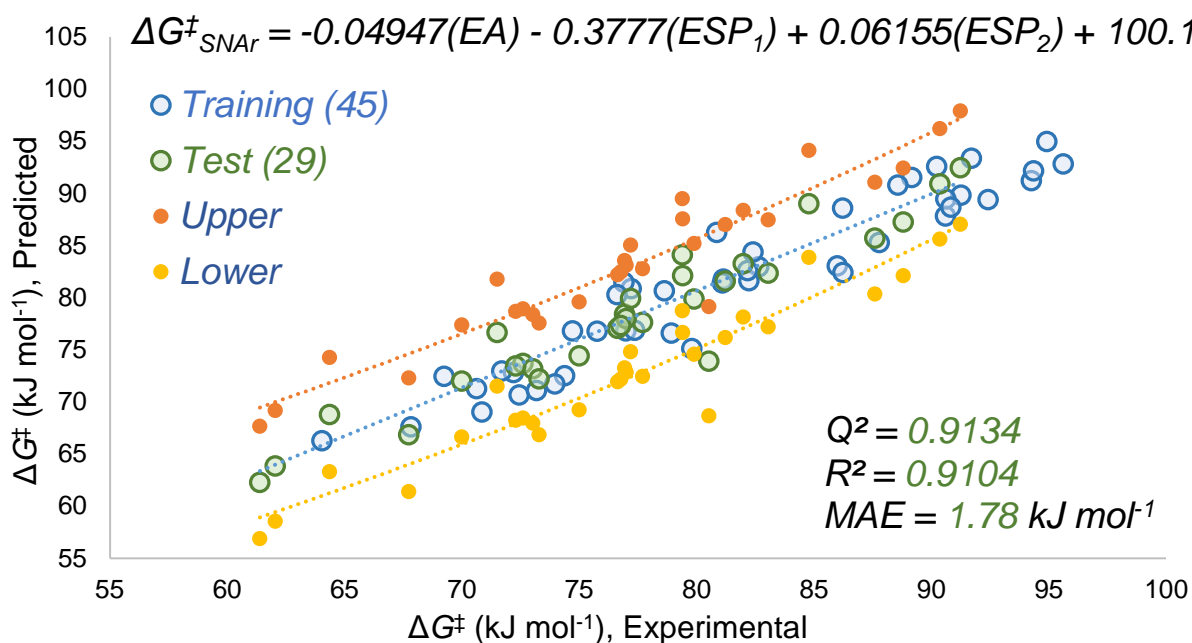


Fig. S20.

Multivariate linear regression model from one of the five 60/40 random split divisions (1/5), and the 95% prediction intervals of the 29 predicted for the test set. The prediction intervals are in the range of ± 5.1 kJ/mol to ± 5.5 kJ/mol.

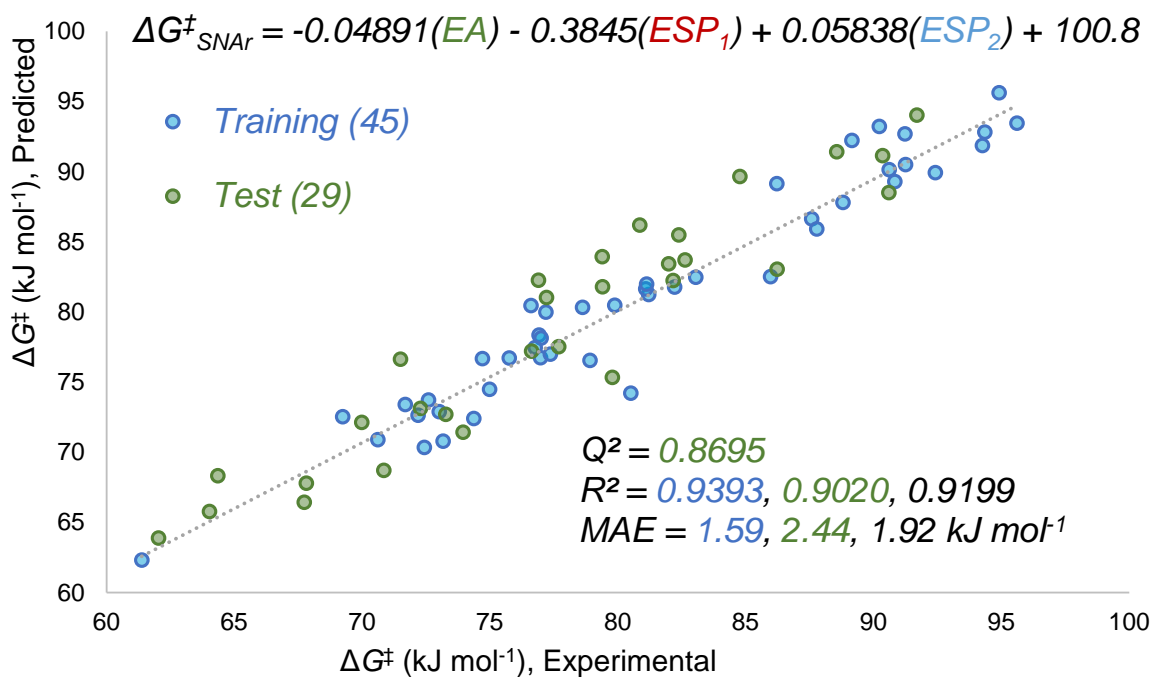


Fig. S21.

Multivariate linear regression model from one of the five 60/40 random split divisions (2/5).

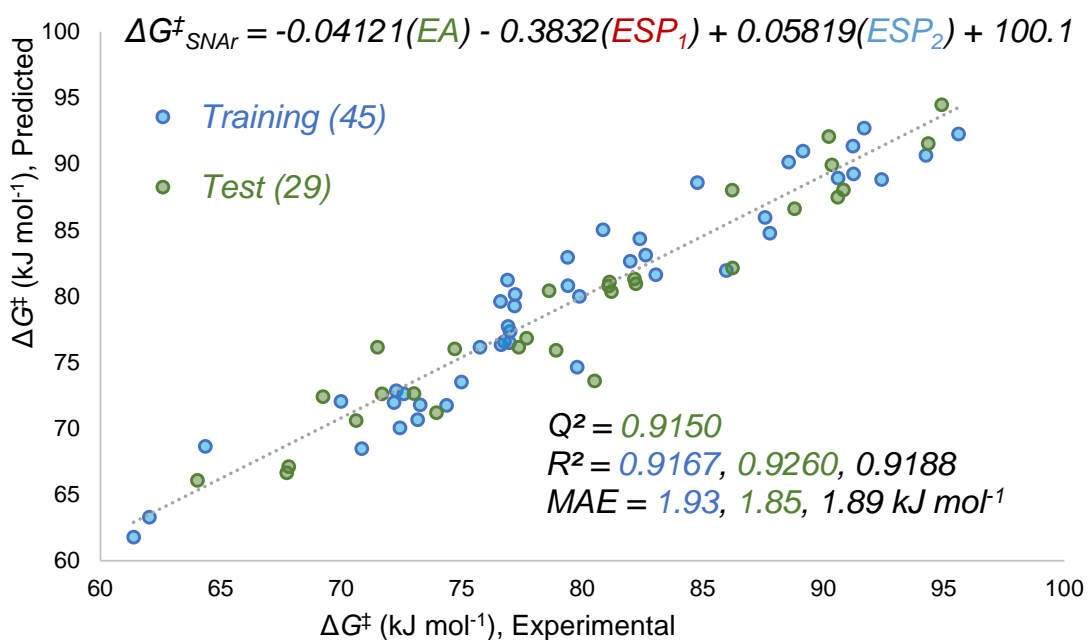


Fig. S22.

Multivariate linear regression model from one of the five 60/40 random split divisions (3/5).

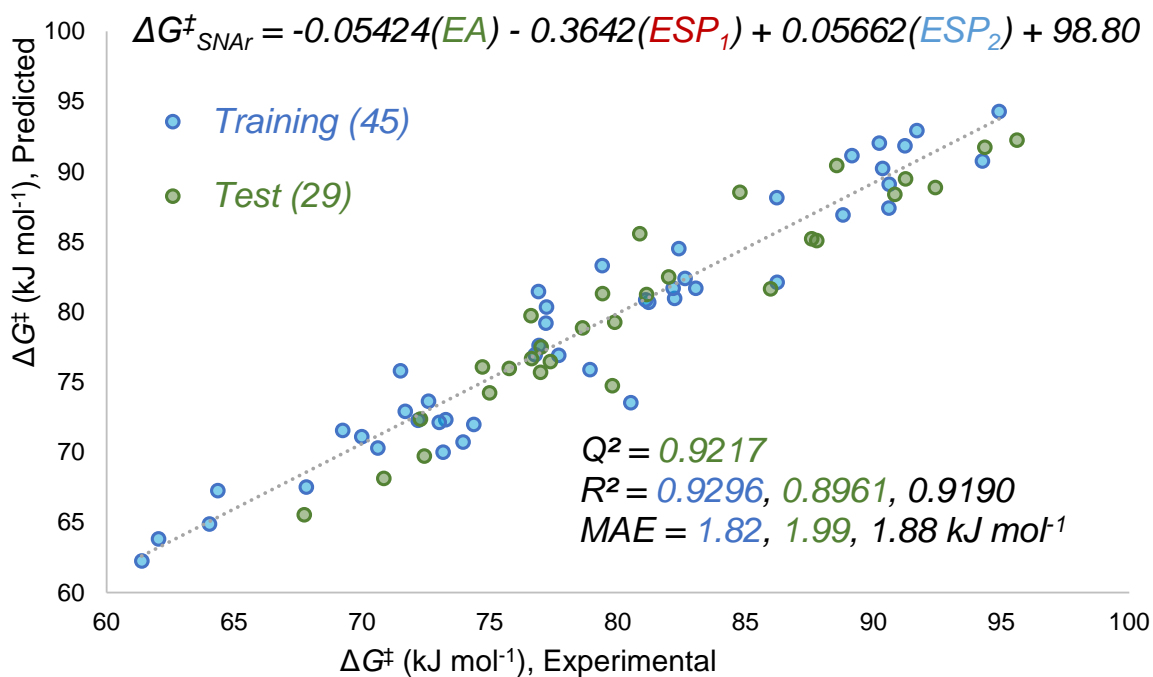


Fig. S23.

Multivariate linear regression model from one of the five 60/40 random split divisions (4/5).

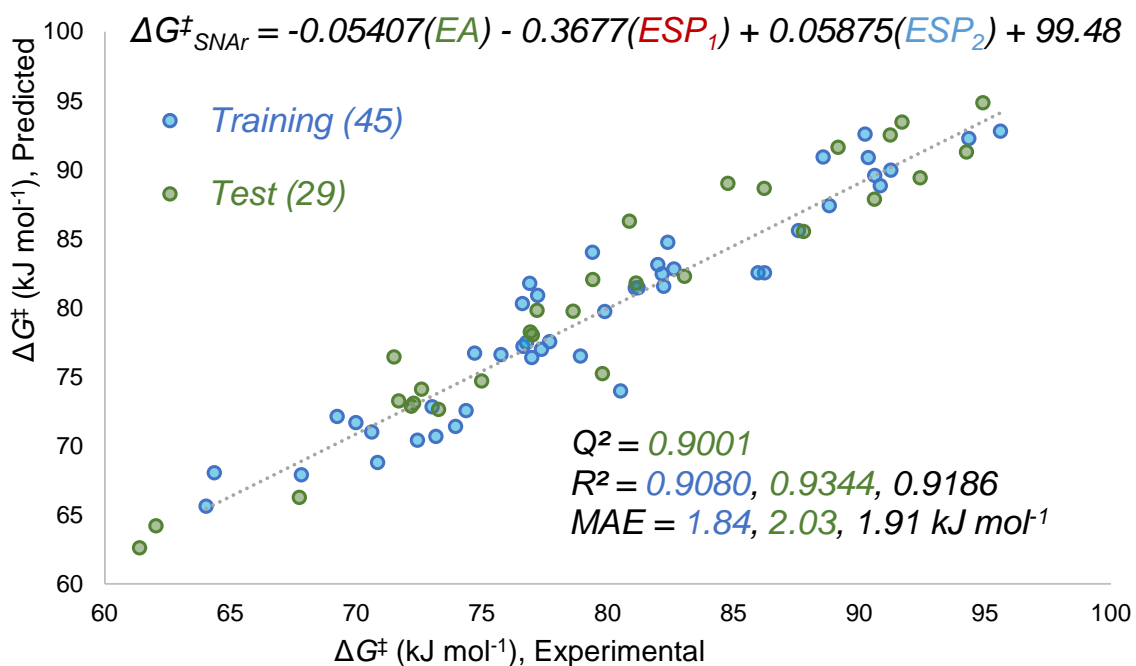


Fig. S24.

Multivariate linear regression model from one of the five 60/40 random split divisions (5/5).

The model performance was further evaluated by an out-of-sample prediction. To test if the model can give reliable predictions for molecules with a variety of structural features, we split the data set into a training set containing only the chloropyridine substrates, and a test set containing all other substrates. The model has achieved an excellent performance with a R^2 of 0.92 for the training set, and a MAE of 1.81 kJ/mol and a Q^2 of 0.93 for the test set (Fig. S25).

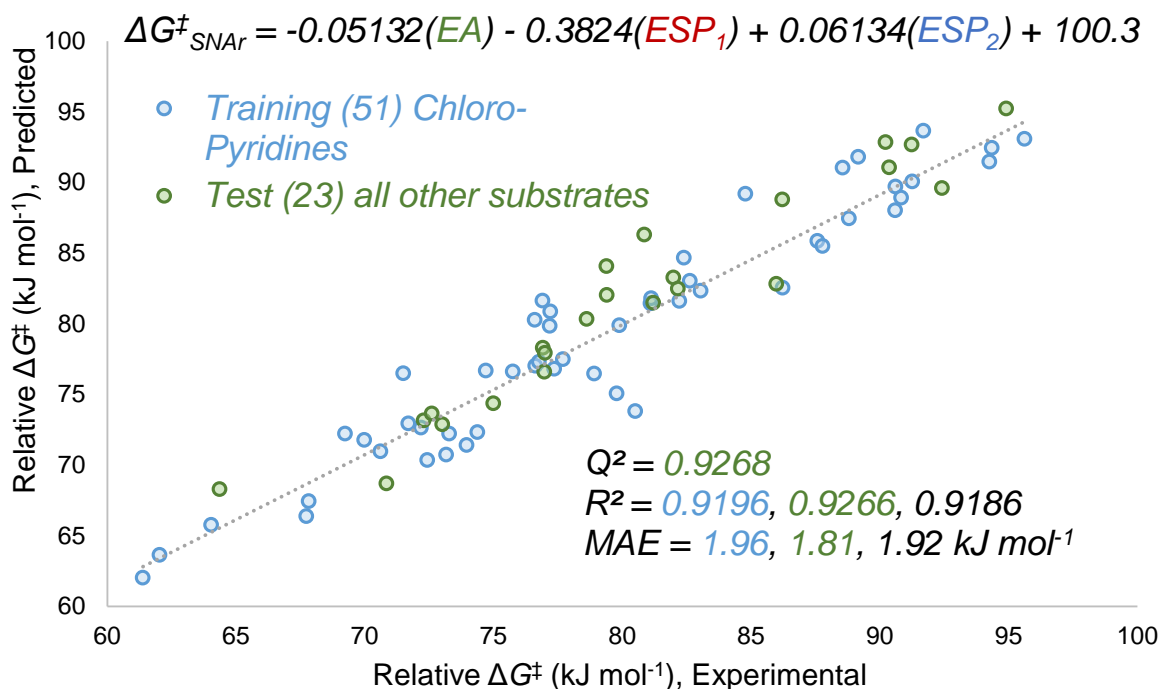


Fig. S25.

Multivariate linear regression model obtained using a training set containing all chloropyridines, and a test set containing all other substrates.

LUMO energy as the Electron Affinity alternatives

LUMO energy at B3LYP/def2-TVZPD level by Orca

We use the electron affinity of a substrate as an approximation to its LUMO energy. Electron affinity is a measurable physical property, it is defined as the amount of energy released after an electron is attached to a neutral molecule. Electron affinity is calculated as the energy difference between a neutral molecule and its anionic radical by DFT. The orbital energy calculation requires less computational effort since only the DFT calculation on the neutral molecule is needed. While LUMO is a virtual orbital, and the calculated virtual orbital energy can introduce uncertainties since it is sensitive to the selection of the basis set.

We compared the LUMO energies to the electron affinity calculated at the same level of theory and basis set (B3LYP/def2-TVZPD), obtaining a linear correlation as shown in Fig. S26. The strong correlation ($R^2=0.94$) indicates that the LUMO energies may be used as an alternative to the electron affinity to reduce computational cost. Then we constructed a multivariate linear model using LUMO energies and ESP values as shown in Fig. S27, and the model performance ($R^2=0.92$ and MAE=1.91 kJ/mol) is equivalent to the model using electron affinity (Fig. S19).

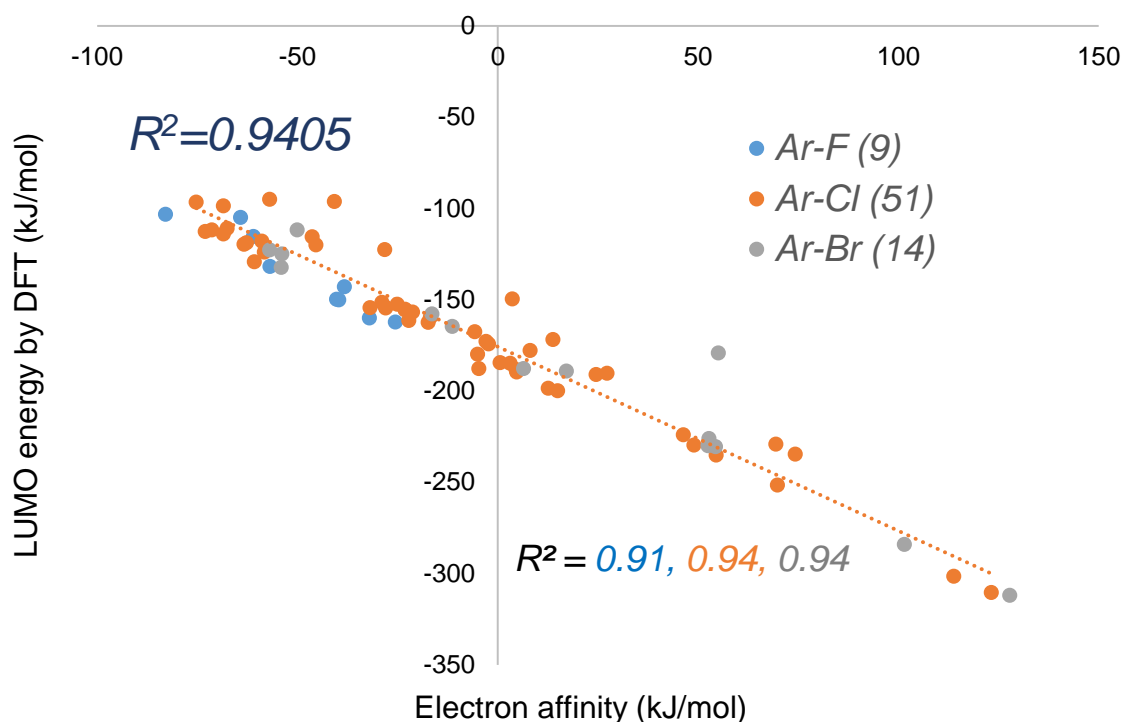


Fig. S26.

Linear correlation of electron affinity (*EA*) vs. DFT-calculated *LUMO* energies at B3LYP/def2-TVZPD of the 74 substrates in the training dataset.

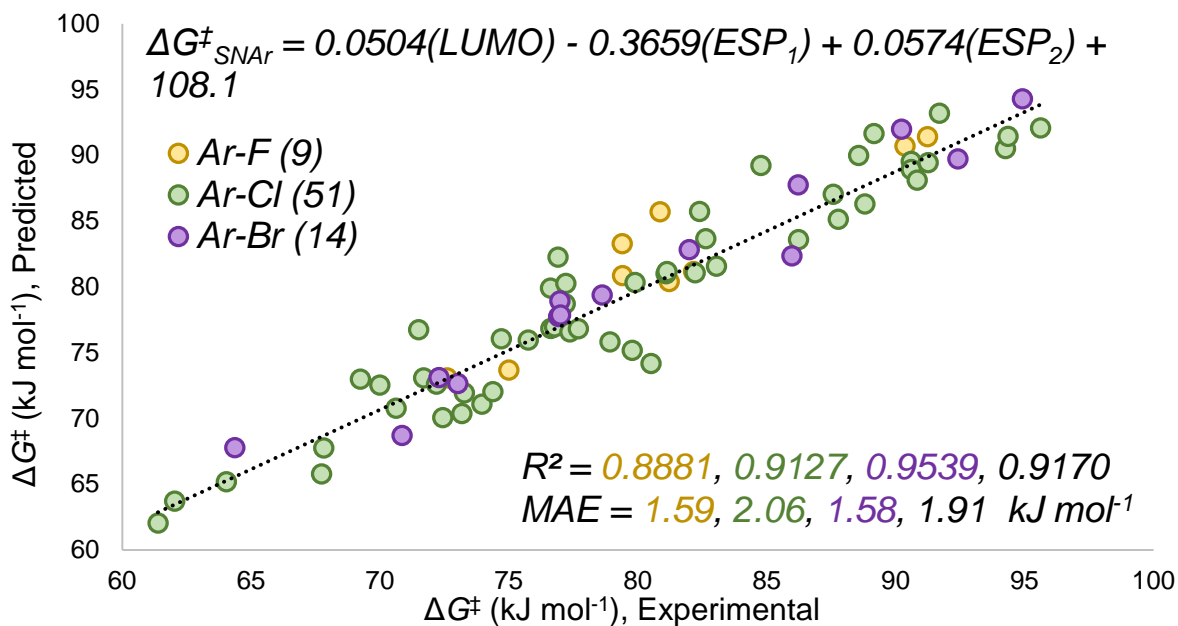


Fig. S27.

Multivariate linear regression model constructed using DFT-calculated *LUMO* energies and *ESP* values at B3LYP/def2-TVZPD, and the linear regression analysis for experimental versus predicted $\Delta G^{\ddagger}_{SNAr}$.

LUMO energy at GFN1-xTB level by Entos Envision

To assess whether a rapid method at a lower level of theory could be suitable for generating *LUMO* energies that give accurate predictions, we evaluated Entos Envision¹⁶, an interactive web-based platform for molecular simulation and visualization developed by Entos, Inc. It requires the molecular structure as input to run computational calculations using GFN1-xTB,¹⁷ a semi-empirical tight-binding method, and outputs a series of properties including molecular orbital energies. Calculation at this theory level only takes a few seconds for the size of a typical substrate in S_NAr reaction, much faster than hybrid functional DFT calculations, which may run for hours. We compared the *LUMO* energies computed in Entos to electron affinity at B3LYP/def2-TVZPD level, obtaining a linear correlation as shown in Fig. S28. This good correlation, with an $R^2 = 0.88$, indicates that these semi-empirical calculations give reasonable *LUMO* energy values. We used the *LUMO* energies obtained from Entos in our multivariate regression model, and compared its performance with the model built using DFT-calculated electron affinity (Fig. S19). The two models are almost identical in their performance, in terms of R^2 and MAE values (Fig. S29).

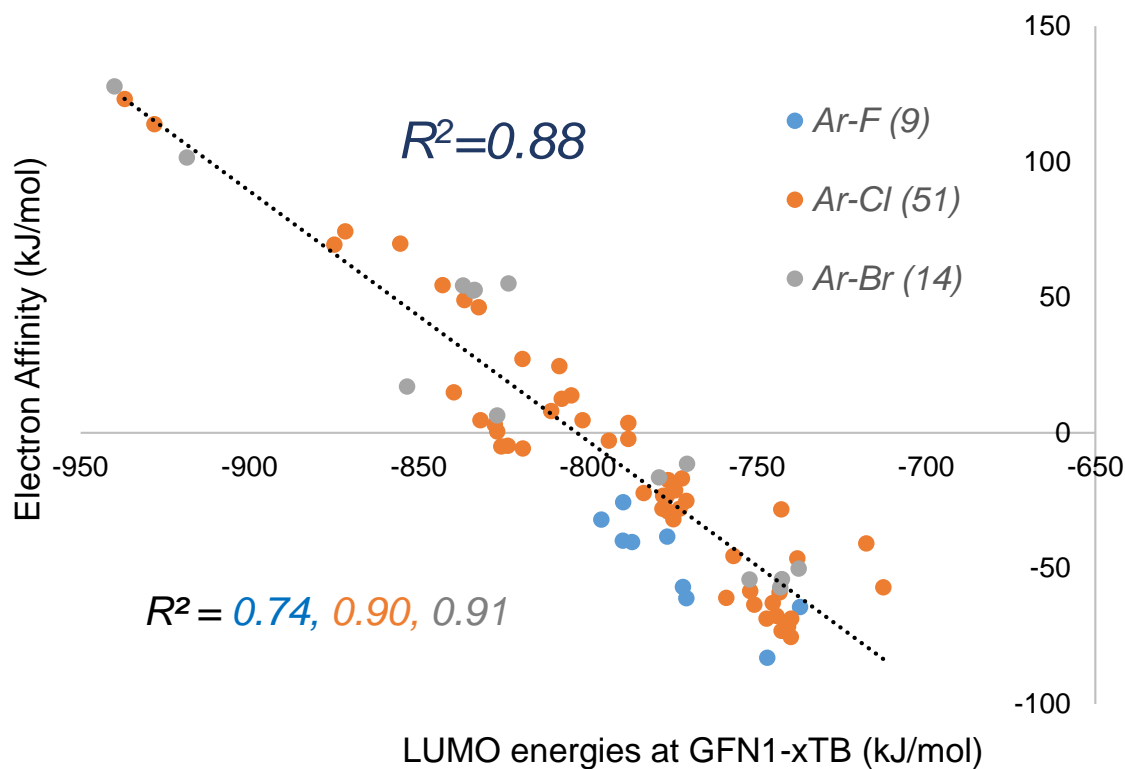


Fig. S28.

Linear correlation of the electron affinity at B3LYP/def2-TVZPD vs. the LUMO energies at GFN1-xTB of the 74 substrates in the training dataset.

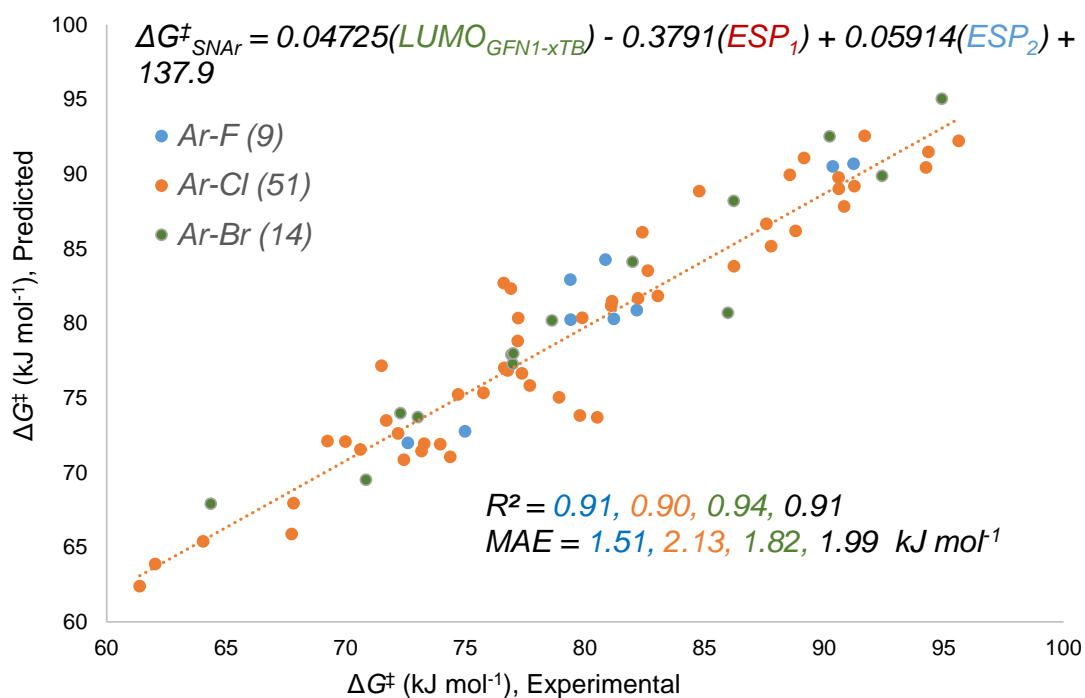


Fig. S29.

Multivariate linear regression model constructed using *LUMO* energies at GFN1-xTB by Entos and *ESP* values at B3LYP/def2-TVZPD by ORCA/Multiwfn, and the linear regression analysis for experimental versus predicted $\Delta G_{\text{SNAr}}^\ddagger$.

Transition State Calculations of S_NAr by DFT

All transition state calculations were performed in ORCA 4.0.1.2. Geometry optimization and frequency calculations for starting materials, intermediates, transition states and products were performed at B3LYP/def2-SVPD level with D3BJ dispersion for all atoms and implicit solvation model CPCM when solvent effect is included. Additional single point calculations were performed at B3LYP/def2-TVZPD level with D3BJ dispersion and CPCM when needed.

Transition state geometries were sought through relaxed scans along a simplified reaction coordinate (the breaking or forming bond length), at a B3LYP/def2-SVPD/CPCM (when needed) level with D3BJ dispersion. The geometry at the maximum energy point of the reaction coordinate scan was used as input geometry for continuing the transition state search via saddle-point optimization.

The optimization to a saddle point was performed using a B3LYP/def2-SVPD/CPCM (when needed) approach with D3BJ dispersion. The transition state structure was confirmed by vibrational analysis. The vibrational mode corresponding to the single imaginary frequency was visualized to ensure that it lied along the reaction coordinate bond stretch.

The free energies of the starting materials, intermediates and transition states were calculated as the sum of the electronic energies (single point energy at B3LYP/def2-TVZPD level) and the Gibbs free energy correction (obtained from frequency calculations at B3LYP/def2-SVPD level). Gibbs free energies of activation ($\Delta G^\ddagger_{\text{S}_{\text{N}}\text{Ar}}$) for the reactions were calculated as the difference in free energies between the transition state and the starting materials.

These labor-intensive transition state calculations were pursued for 6 di-halogenated heterocycles at both reactive sites except for 2,4-dichloro-1,3,5-triazine, for which the two sites are identical. In implicit solvation model CPCM(DMSO), concerted pathways were found at all reactive centers, and no sigma-complex intermediates could be located; the starting materials transformed into the products *via* only one intrinsic reaction coordinate (the bond length between the reactive carbon at the substrate and the negatively charged oxygen at deprotonated benzyl alcohol). Transition state structures and energy diagrams are shown in Figs. S30 – S39. All molecular images were generated in Vesta.¹⁸

In gas phase, stepwise mechanisms were found for the carbon-fluorine sites at 2-chloro-3-fluoro-pyridine and 2-chloro-4-fluoro-pyridine (Figs. S38 – S39). The stable sigma-complex was located from geometry optimization and confirmed by frequency calculations, that no negative frequencies were found. The first step, addition of the nucleophile, is an exothermic reaction that involves a very early transition state. In these cases, the transition state could not be found as no maximum point was accessible from the potential energy surface scan. This computational challenge has been reported by Jacobsen and co-workers' in their S_NAr mechanistic study, where transition state searches failed in similar exothermic steps.¹³ Transition states for the second step (the departure of the fluorine) were located and the structures were optimized and confirmed by vibrational analysis for both of the carbon-fluorine sites.

HOMO diagrams for the transition state were calculated at both the major and minor reactive sites for the following 3 substrates: 2,4-dichloropyridine (Fig. 6, main text), 2-chloro-4-bromopyridine (Fig. S32) and 2,4-dichloropyrimidine (Fig. S34). One motivation to conduct transition state calculations is to look for the mechanistic factors that determine the intramolecular selectivity, and our orbital analysis shows that the frontier orbital interaction is certainly among those important factors. Transition state frontier orbital analysis shows that different sites of a substrate use different antibonding orbitals to interact with the nucleophile, where those orbitals are accessible through thermal fluctuations. In these cases, the major site undergoes S_NAr corresponding to the *LUMO* orbital, while the minor site corresponds to *LUMO*+1.

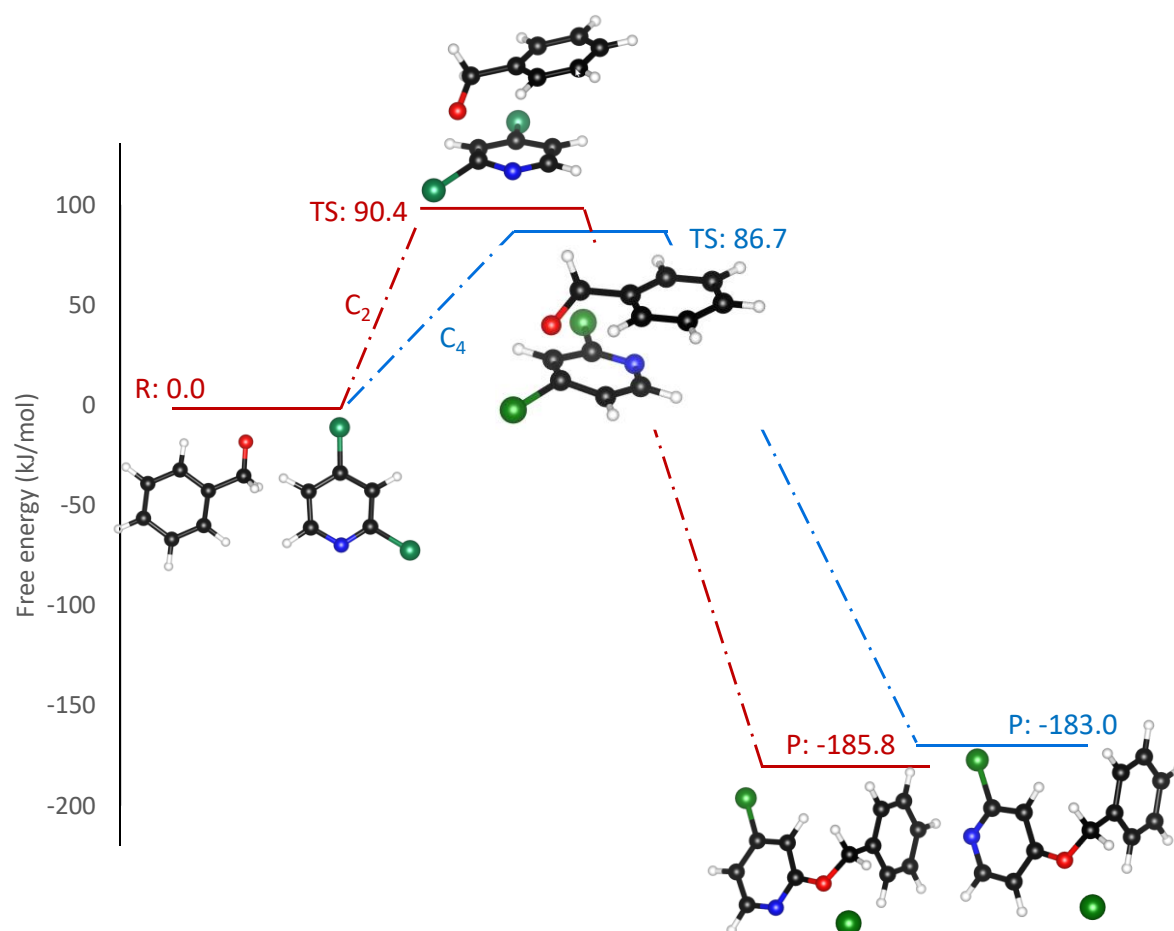


Fig. S30.

Calculated reaction coordinates for S_NAr of 2,4-dichloropyridine and the anion of benzyl alcohol in DMSO. The blue pathway is the S_NAr coordinate at C₄, the major reactive site; and the red pathway is the S_NAr coordinate at C₂, the minor reactive site.

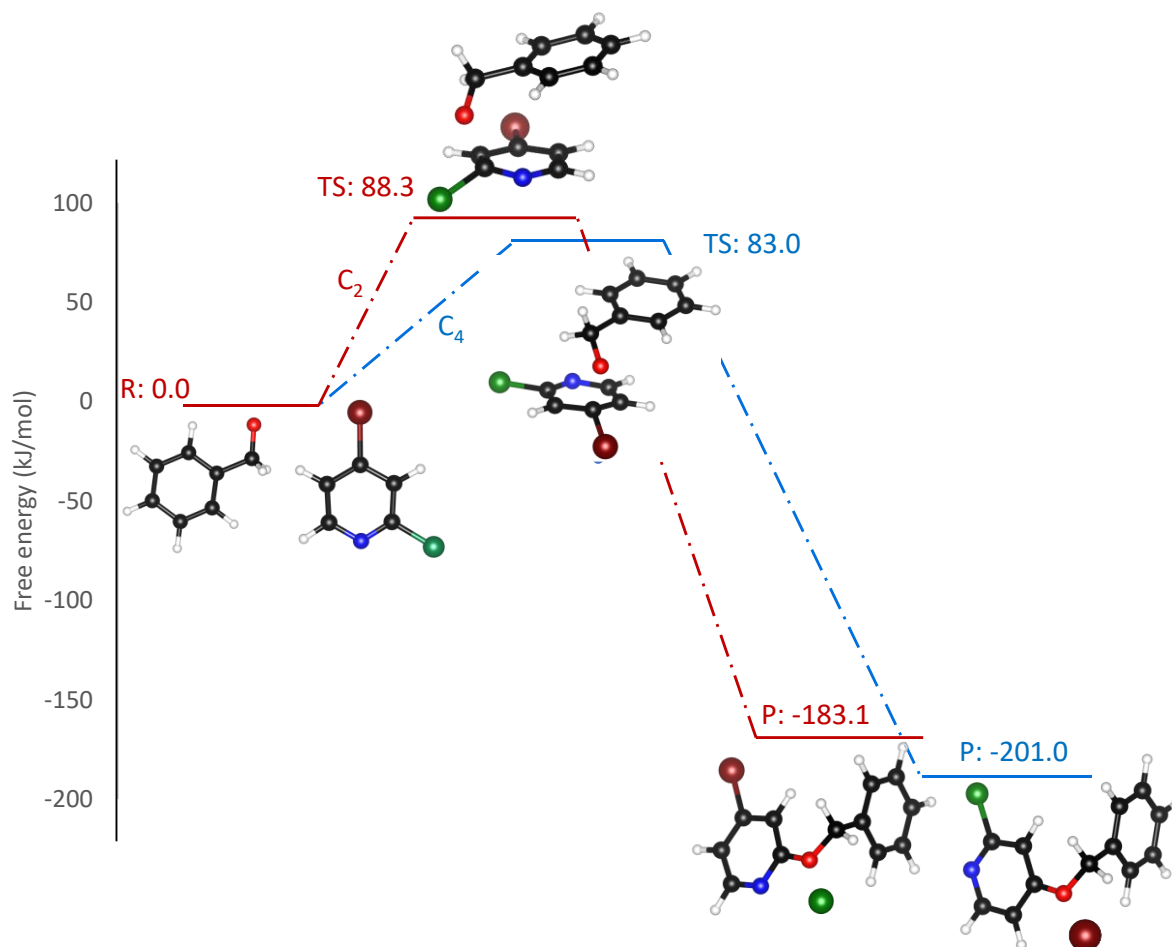


Fig. S31.

Calculated reaction coordinates for S_NAr of 2-chloro-4-bromopyridine and the anion of benzyl alcohol in DMSO. The blue pathway is the S_NAr coordinate at C₄, the major reactive site; and the red pathway is the S_NAr coordinate at C₂, the minor reactive site.

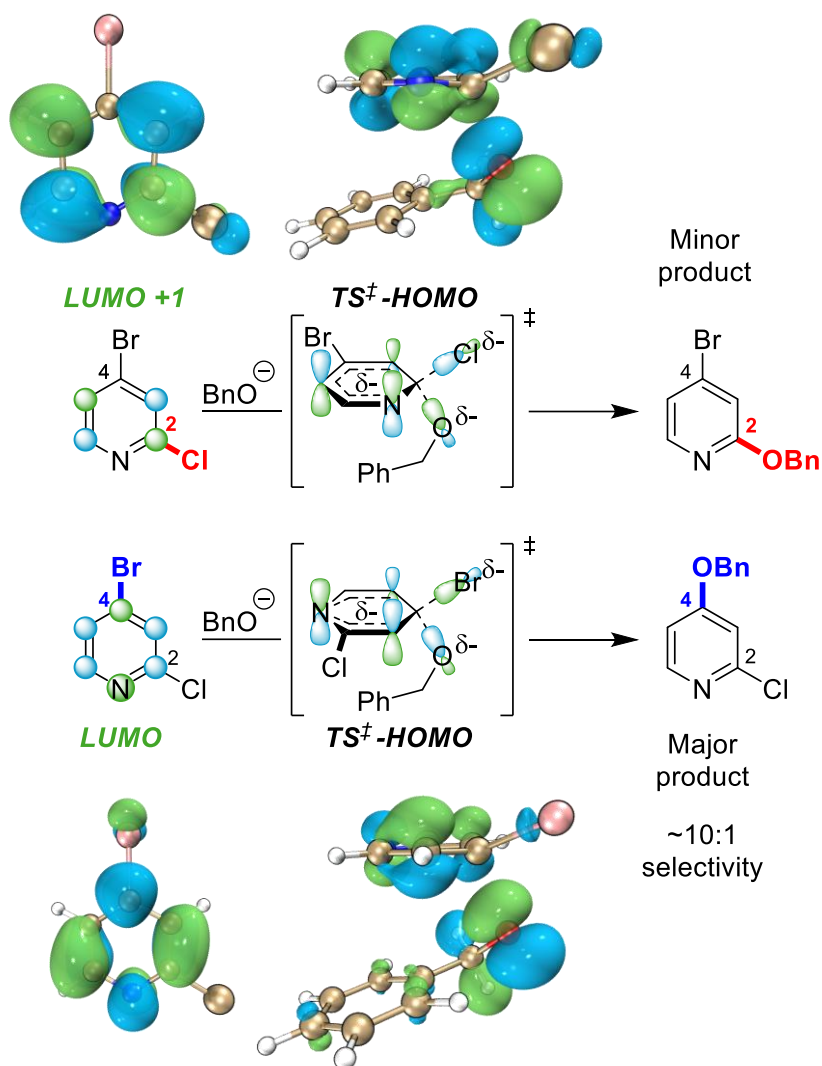


Fig. S32.

Frontier orbitals of 2-chloro-4-bromopyridine and the anion of benzyl alcohol involved in S_NAr , and $HOMO$ orbitals of the transition states at C_2 and C_4 . The major reactive site C_4 undergoes S_NAr with the $LUMO$ orbital, as shown as the blue reaction pathway in Fig. S31. The minor reactive site C_2 undergoes S_NAr with the $LUMO+1$ orbital, as shown as the red reaction pathway in Fig. S31.

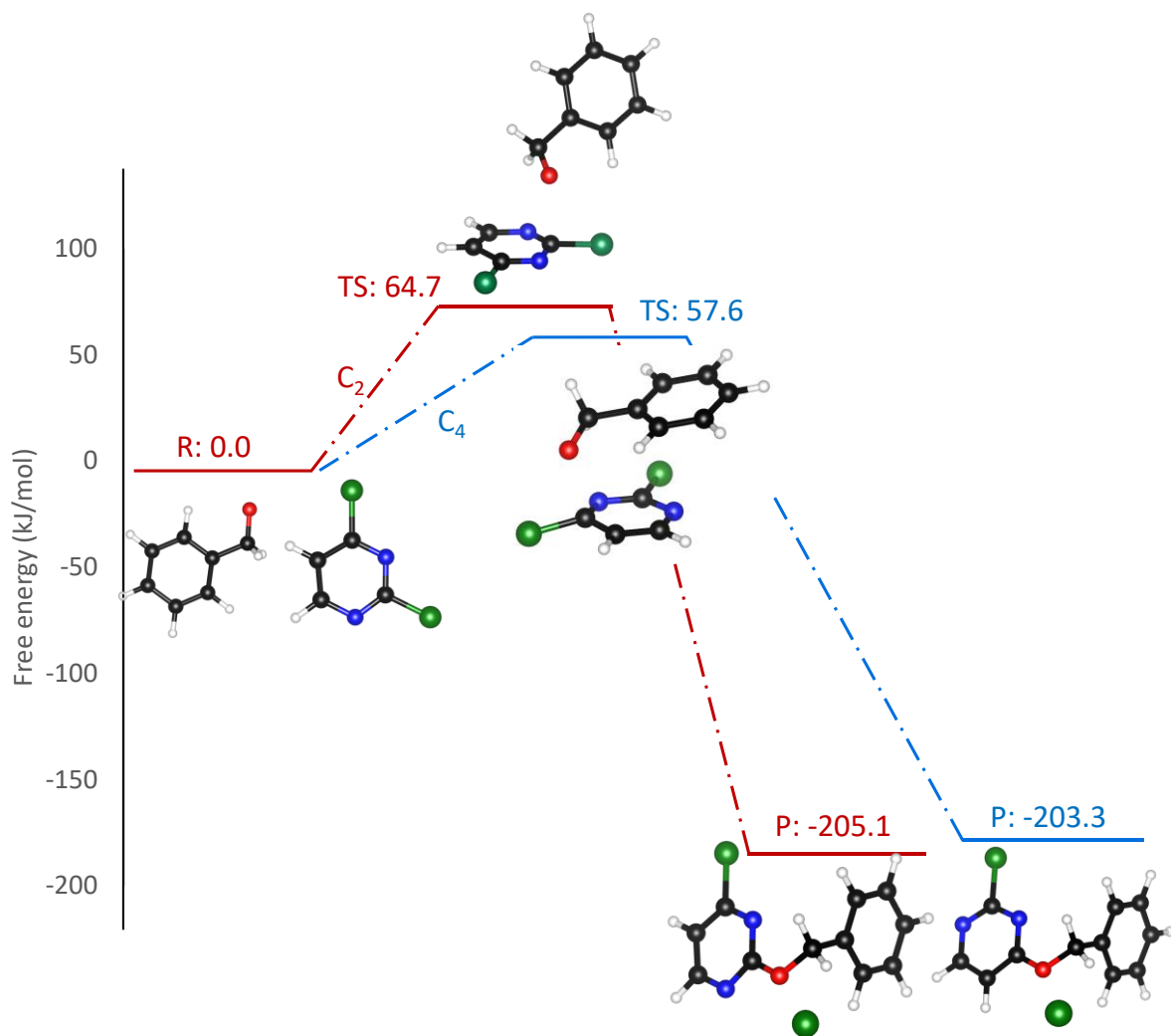


Fig. S33.

Calculated reaction coordinates for S_NAr of 2,4-dichloro-pyrimidine and the anion of benzyl alcohol in DMSO. The blue pathway is the S_NAr coordinate at C₄, the major reactive site; and the red pathway is the S_NAr coordinate at C₂, the minor reactive site.

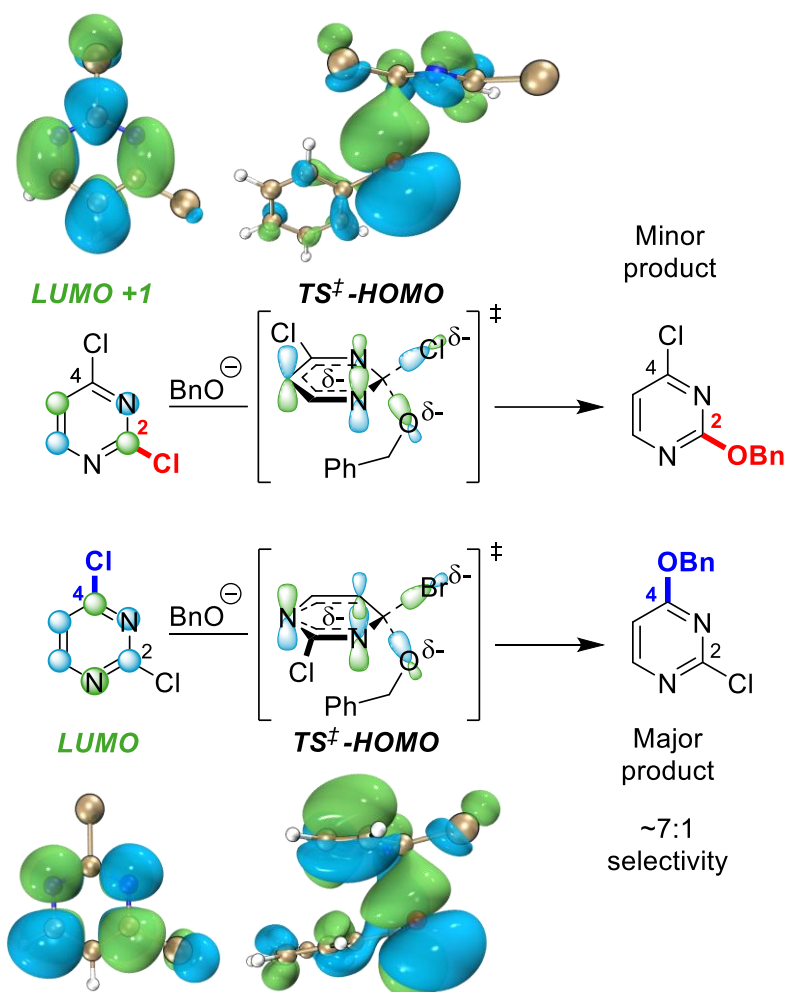


Fig. S34.

Frontier orbitals of 2,4-dichloropyrimidine and the anion of benzyl alcohol involved in S_NAr , and $HOMO$ orbitals of the transition states at C_2 and C_4 . The major reactive site C_4 undergoes S_NAr with the $LUMO$ orbital, as shown as the blue reaction pathway in Fig. S33. The minor reactive site C_2 undergoes S_NAr with the $LUMO+1$ orbital, as shown as the red reaction pathway in Fig. S33.

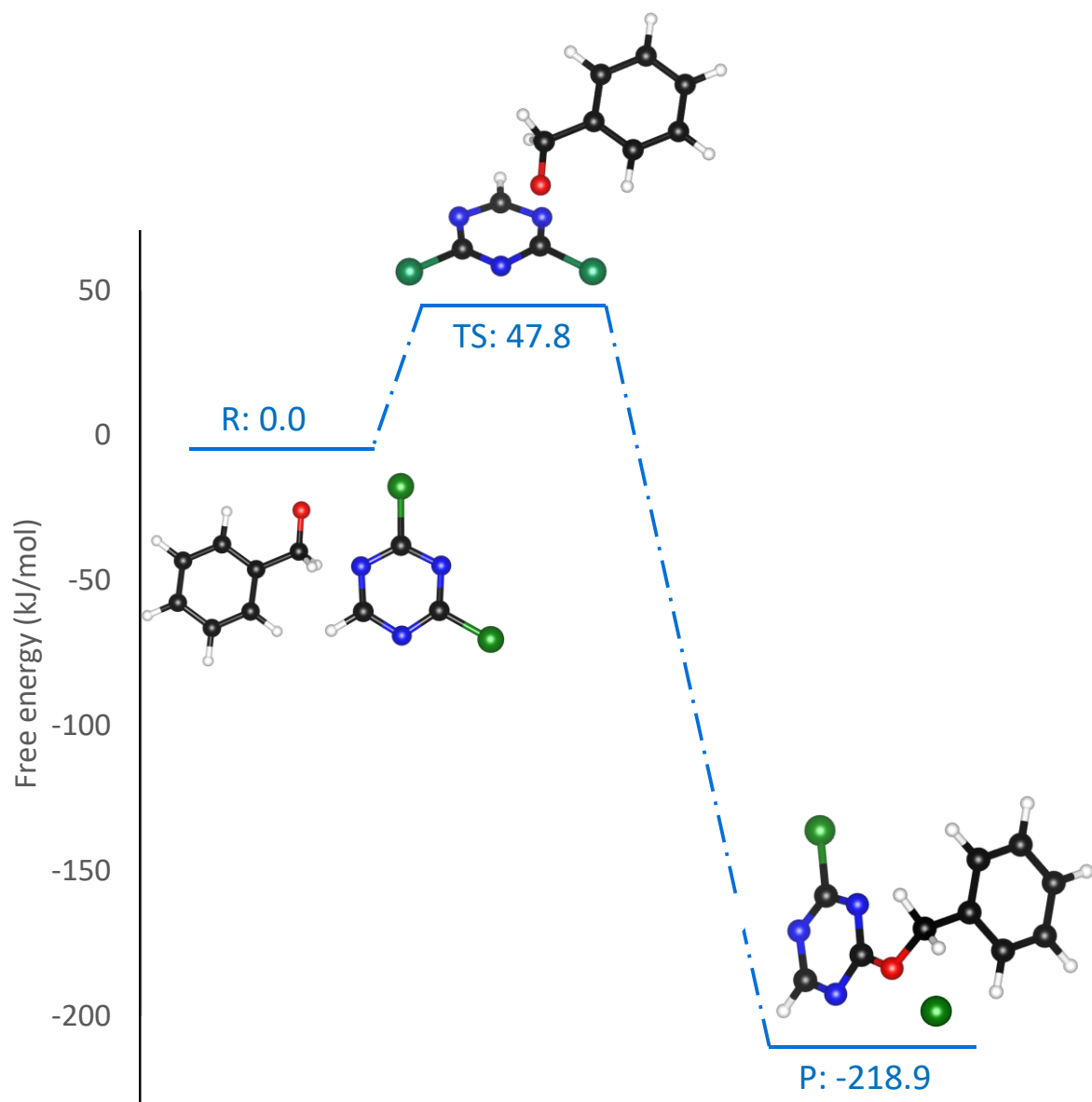


Fig. S35. Calculated reaction coordinates for S_NAr of 2,4-dichloro-1,3,5-triazine and the anion of benzyl alcohol in DMSO.

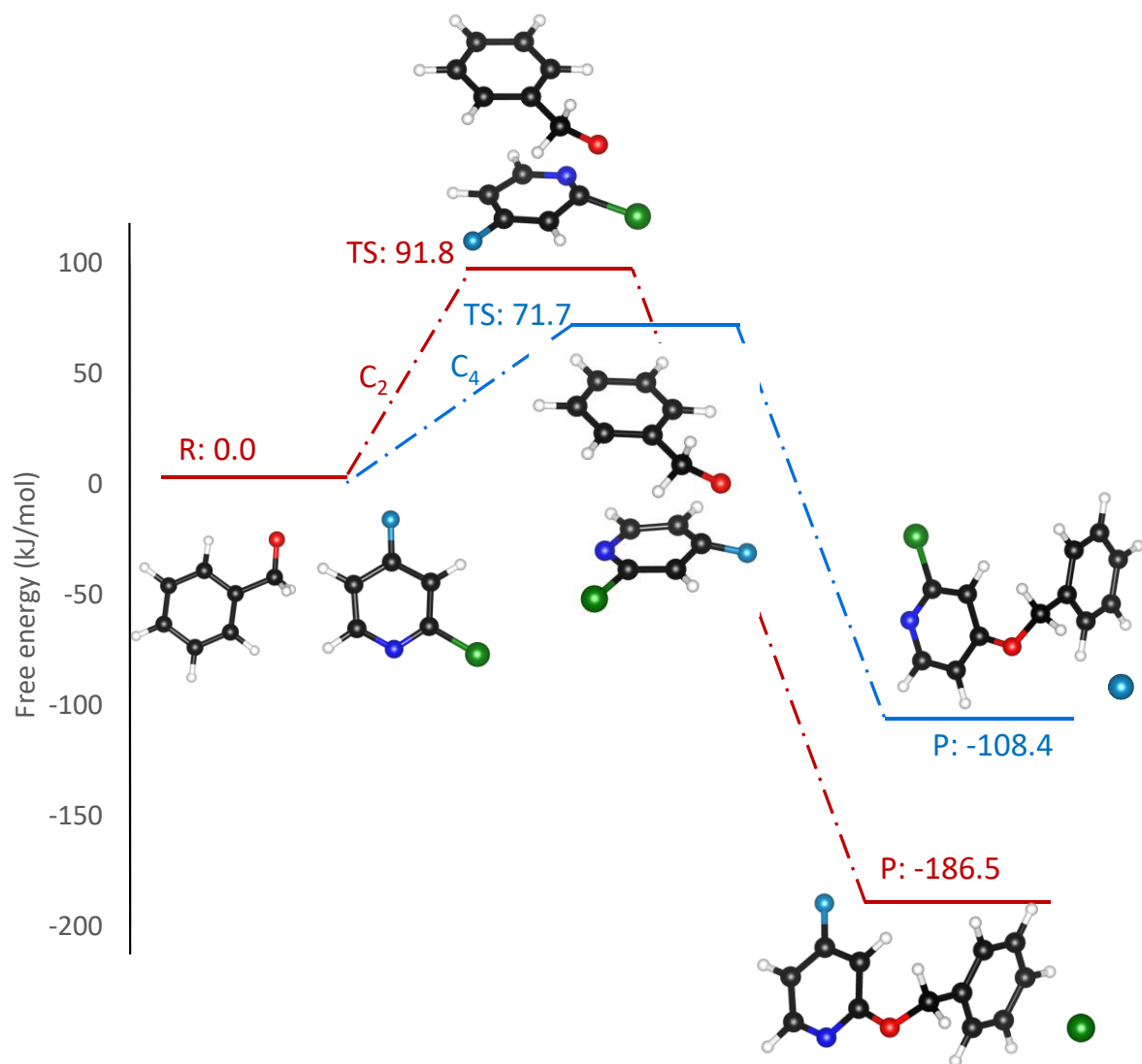


Fig. S36. Calculated reaction coordinates for S_NAr of 2-chloro-4-fluoropyridine and the anion of benzyl alcohol in DMSO. The blue pathway is the S_NAr coordinate at C₄, the major reactive site; and the red pathway is the S_NAr coordinate at C₂, the minor reactive site.

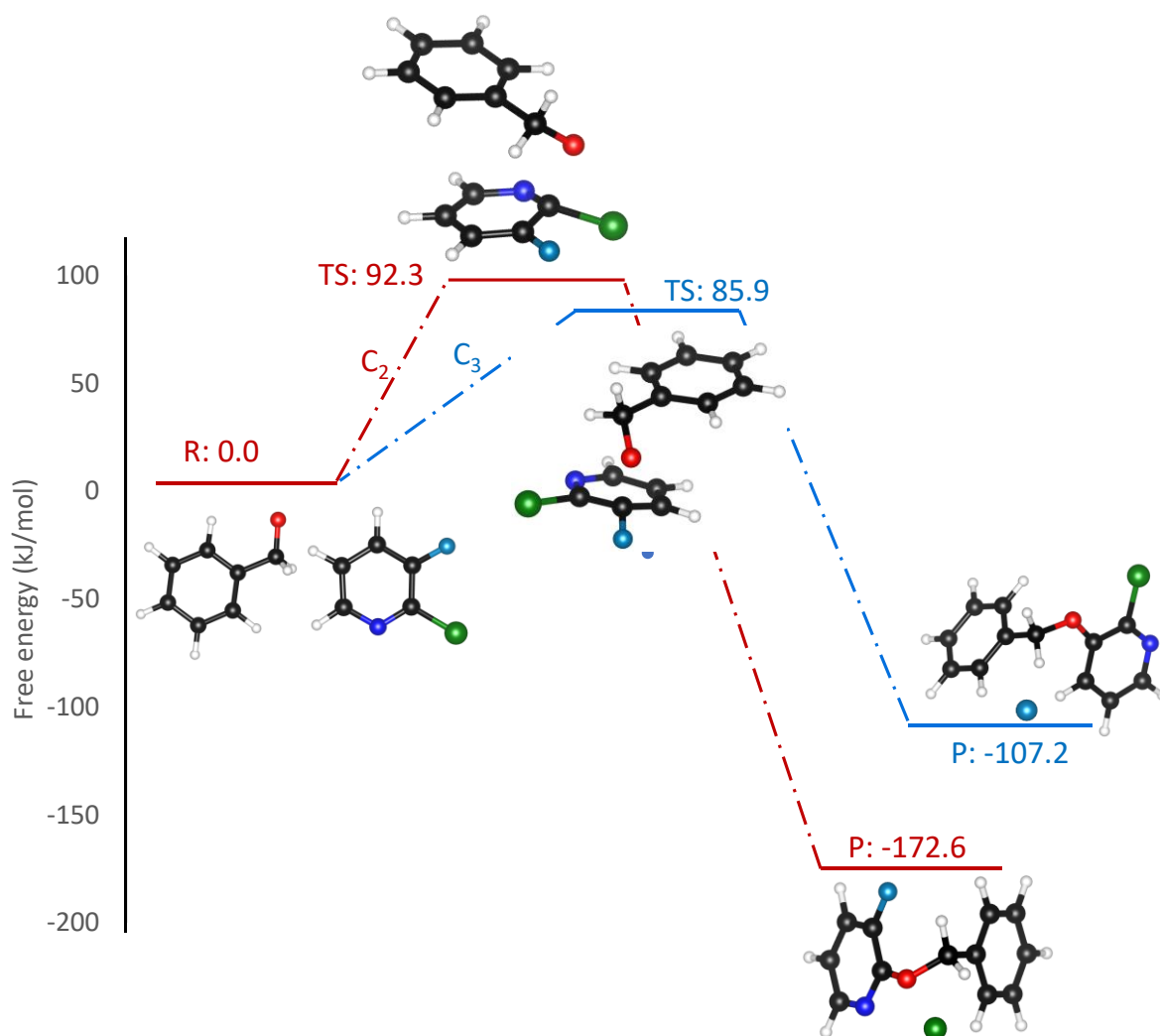


Fig. S37. Calculated reaction coordinates for S_NAr of 2-chloro-3-fluoropyridine and the anion of benzyl alcohol in DMSO. The blue pathway is the S_NAr coordinate at C₃, the major reactive site; and the red pathway is the S_NAr coordinate at C₂, the minor reactive site.

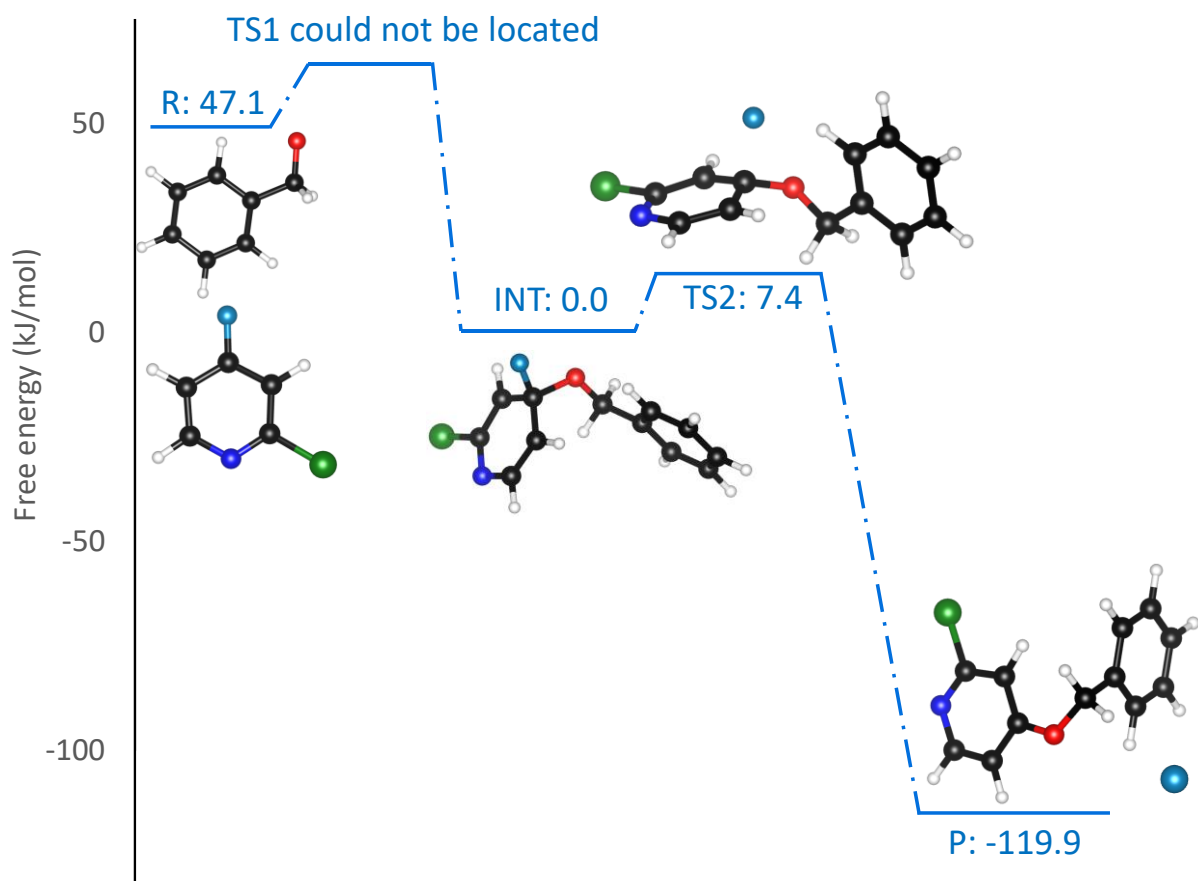


Fig. S38. Calculated reaction coordinate for S_NAr at the carbon-fluorine site (C_4) of 2-chloro-4-fluoropyridine and the anion of benzyl alcohol in the gas phase.

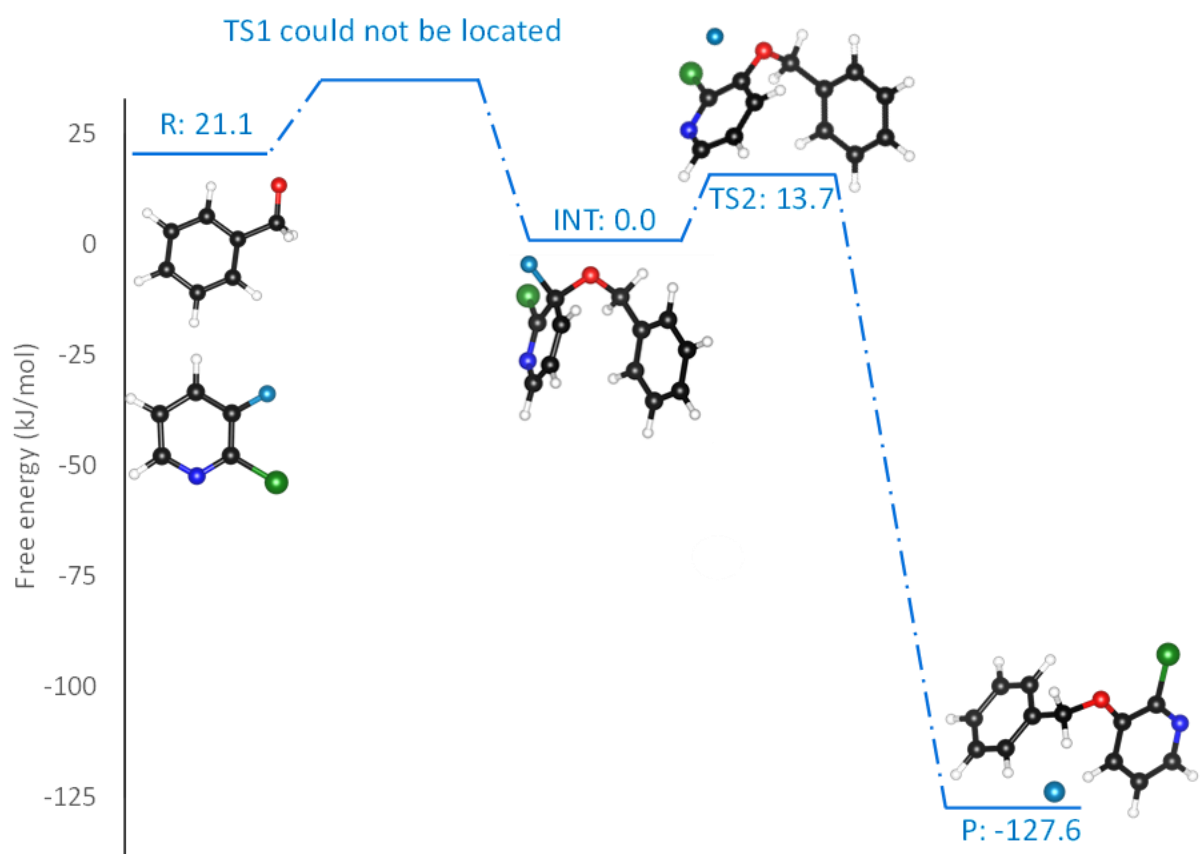


Fig. S39. Calculated reaction coordinates for S_NAr at the carbon-fluorine site (C_3) of 2-chloro-3-fluoropyridine and the anion of benzyl alcohol in the gas phase.

Experimental details for assessing outliers from external case studies (Fig. 7C)

To assess the model's applicability on new S_NAr reactions beyond our training dataset, we conducted a series of external validation on both S_NAr rate correlation (main text, Fig. 7) and site-selectivity prediction (main text, Fig. 8-12). We also performed experimental checks on the one significant outlier that we observed from the S_NAr rate correlation test (Substrate 6, Fig. 7C).

The substrate 6 from Fig. 7C (1-bromo-4-fluoro-2-nitrobenzene) has two potential S_NAr sites (Ar-F and Ar-Br) and is expected to form multiple products from reaction with piperidine.

The substrate was purchased from a commercial supplier and used as is. The reaction was conducted using the same reaction conditions from literature²² (which is also shown in Fig. S40A) for two days, then the reaction solution was analyzed by LCMS. The molecular ion peaks from the LCMS analysis gives experimental evidence that the reaction of substrate 6C and piperidine leads to two S_NAr products. The product ratio of 1.5:1 was obtained from the peak area ratio from the UV spectrum at 254 nm, with the Ar-Br site slightly favoured.

Excellent linear correlation ($R^2 = 0.90$) was obtained for the dataset in Fig. 7C with all the substrates included, and exclusion of the outlier 6C gives rise to a stronger correlation, with an R^2 of 0.94.

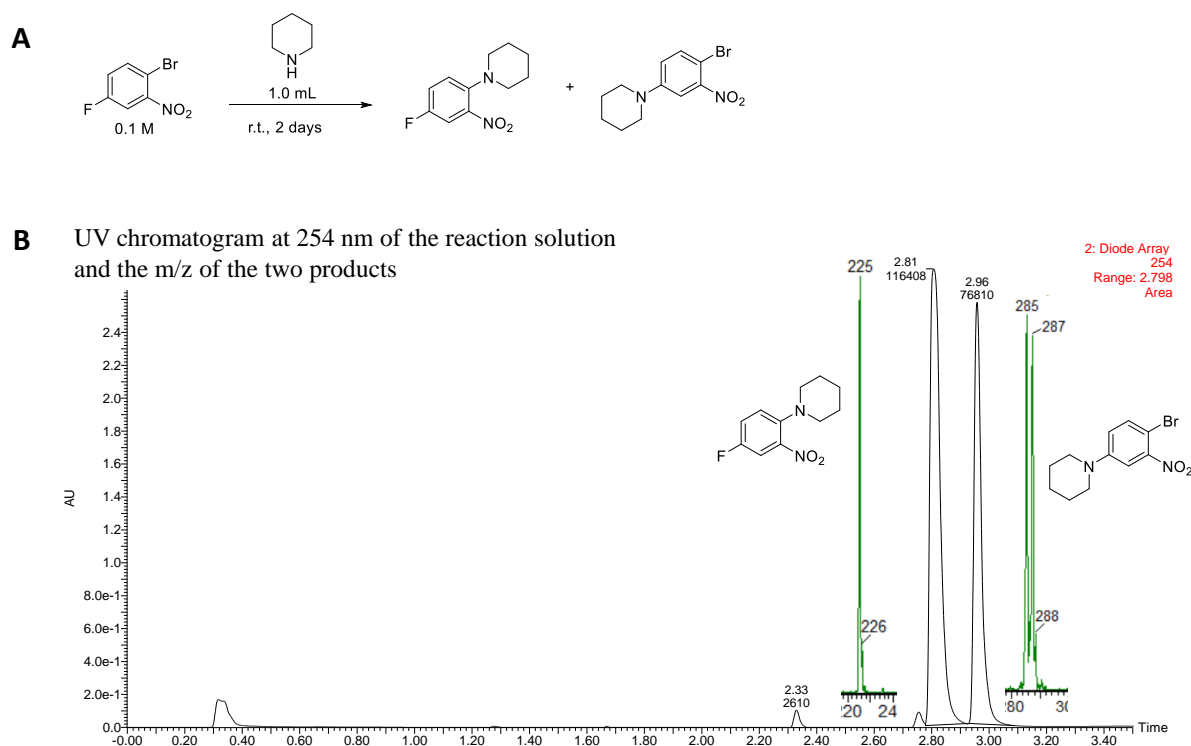


Fig. S40.

A). Reaction condition of substrate 6 in Fig.7C reacting with piperidine. B). UV chromatogram at 254 nm extracted from LCMS analysis of the reaction solution, and the m/z of the two S_NAr products.

References

1. S. Sengmany, J. Lebre, E. Le Gall, E. Leonel, Selective mono-amination of dichlorodiazines. *Tetrahedron* 2015, 71, 4859-4867.
2. S. Pöller, W. Schuhmann, A miniaturized voltammetric pH sensor based on optimized redox polymers. *Electrochimica Acta* 2014, 140, 101–107.
3. B. Cash, N. Prevost, F. Wagner, D. Comins, Studies toward the Total Synthesis of Dihydrolycolucine. Preparation of AB and CEF Ring Fragments. *J. Org. Chem.* 2014, 79, 5740–5745.
4. H. Gellrich, J. Huang, M. Seiche, M. Keller, M. Meuwly, B. Breit, Ligand Self-Assembling through Complementary Hydrogen-Bonding in the Coordination Sphere of a Transition Metal Center: The 6-Diphenylphosphanylpyridin-2(1H)-one System. *J. Am. Chem. Soc.* 2011, 133, 964–975.
5. C. Hansch, A. Leo, R. W. Taft, A survey of Hammett substituent constants and resonance and field parameters. *Chem. Rev.* 1991, 91, 165-195.
6. F. Neese, F. Wennmohs, U. Becker, C. Riplinger, The ORCA quantum chemistry program package. *J. Chem. Phys.* 2020, 152, 224108.
7. T. Lu, F. Chen, Multiwfn: A Multifunctional Wavefunction Analyzer. *J. Comput. Chem.* 2012, 33, 580-592.
8. T. Lu, F. Chen, Quantitative analysis of molecular surface based on improved Marching Tetrahedra algorithm. *J. Mol. Graph. Model.* 2012, 38, 314-323.
9. H. E. Pence, A. Williams, ChemSpider: An Online Chemical Information Resource. *J. Chem. Educ.* 2010, 87, 1123-1124.
10. M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* 2012, 4, 17.
11. T. Koopmans, Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. *Physica* 1934, 1, 104–113.
12. W. Humphrey, A. Dalke, K. Schulten, VMD - Visual Molecular Dynamics. *J. Molec. Graph.* 1996, 14, 33-38.

13. E. Kwan, Y. Zeng, H. Besser, E. Jacobsen, Concerted nucleophilic aromatic substitutions. *Nature Chem.* 2018, 10, 917–923.
14. F. Terrier, Rate and Equilibrium Studies in Jackson-Meisenheimer Complexes. *Chem. Rev.* 1982, 82, 77-152.
15. J. A. Hirsch Topics in Stereochemistry, John Wiley & Sons, Ltd, 1967, pp. 199–222.
16. Entos Envision. <http://www.entos.ai/envision> (accessed 15 July 2022).
17. S. Grimme, C. Bannwarth, P. Shushkov, A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All *spd*-Block Elements ($Z = 1-86$). *J. Chem. Theory Comput.* 2017, 13, 1989–2009.
18. K. Momma, F. Izumi, VESTA: a three-dimensional visualization system for electronic and structural analysis. *J. Appl. Cryst.* 2008, 41, 653-658.
19. K. Fukui, T. Yonezawa, H. Shingu, A Molecular Orbital Theory of Reactivity in Aromatic Hydrocarbons. *J. Chem. Phys.* 1952, 20, 722-725.
20. K. Houk, Frontier molecular orbital theory of cycloaddition reactions. *Acc. Chem. Res.* 1975, 8, 361–369.
21. S. Giri, R. Gour, and K. Kartha, Diazepinium perchlorate: a neutral catalyst for mild, solvent-free acetylation of carbohydrates and other substances. *RSC Adv.* 2017, 7, 13653-13667.
22. E. Berliner, L. Monack, Nucleophilic Displacement in the Benzene Series. *J. Am. Chem. Soc.* 1952, 74, 1574–1579.