Supporting Information

A framework for constructing linear free energy relationships to design molecular transition metal catalysts

Zhenzhuo Lan,[a] and Shaama Mallikarjun Sharada*[a,b]

1 Asymmetric Systems

Table 1 shows descriptors, LFER, and DFT barriers for **A1-A38** asymmetrically substituted dicopper systems.

Table 1: Sum of four Hammett parameters σ_p, bite angles θ (°) and Sterimol parameters B1 (Å), BS barriers $\Delta E^{\ddagger}_{DFT}$ calculated by DFT (ωB97X-D/srsc/6-311+G*) and predicted by LFER model $\Delta E^{\ddagger}_{LFER}$ for 38 asymmetrically substituted catalysts (barriers in kJ/mol)

Index	$\sum \sigma_p$	θ_1	θ_2	$B1_1$	$B1_2$	$\Delta E^{\ddagger}_{DFT}$	$\Delta E^{\ddagger}_{LFER}$
A1	-0.54	96.904	95.999	1.59	1.68	83.3	87.3
A2	-0.34	95.159	95.142	1.72	1.68	84.6	85.5
A3	1.10	96.230	96.264	1.88	1.68	73.8	81.2
A4	1.56	97.011	97.396	1.59	1.69	71.3	75.2
A5	-0.88	98.156	98.476	1.72	1.63	94.6	94.9
A6	0.56	99.521	97.618	1.76	1.73	73.1	88.5
A7	1.02	98.153	96.147	1.72	1.63	74.4	79.9
A8	0.76	93.548	97.384	1.72	1.91	79.0	83.0
A9	1.22	95.367	96.050	1.73	1.64	70.5	75.9
A10	2.66	97.254	97.866	1.91	1.89	60.3	78.0
A11	0.55	95.913	96.567	1.68	1.64	83.8	80.5
A12	-0.17	95.041	96.555	1.69	1.66	98.3	84.8
A13	0.78	96.617	96.773	1.69	1.61	70.0	79.6
A14	-0.27	96.455	96.574	1.66	1.61	92.4	85.6
A15	1.10	96.797	96.302	1.67	1.68	84.0	78.1
A16	-0.34	94.998	95.488	1.64	1.67	94.9	84.1
A17	1.56	95.890	97.397	1.64	1.68	63.2	74.7
A18	-0.54	96.501	96.587	1.66	1.67	95.1	88.5
A19	1.10	96.466	96.829	1.64	1.89	82.4	81.5
A20	-0.34	95.610	96.562	1.71	1.64	89.2	86.6
A21	1.56	96.229	96.658	1.61	1.65	68.0	73.2
A22	-0.54	98.769	97.079	1.65	1.62	89.3	90.4
A23	1.65	97.383	94.966	1.79	1.66	78.7	75.4
A24	-0.51	95.664	96.212	1.72	1.68	94.5	88.3
A25	2.33	96.637	97.239	1.69	1.59	69.7	69.5
A26	-0.81	96.648	98.809	1.69	1.66	93.0	93.2
A27	2.66	97.611	96.867	1.64	1.91	71.7	72.9
A28	0.76	93.188	97.381	1.71	1.91	75.9	82.4
A29	1.22	93.739	97.388	1.65	1.59	71.6	73.2
A30	0.56	98.961	97.223	1.64	1.87	80.2	87.8
A31	-0.88	98.788	96.999	1.64	1.72	90.8	94.2
A32	1.02	98.827	96.645	1.64	1.61	75.7	79.4
A33	-0.44	96.578	96.007	1.65	1.62	90.8	86.2
A34	0.28	96.537	96.460	1.66	1.89	87.3	86.9
A35	0.51	96.153	97.078	1.66	1.60	85.5	80.5
A36	0.38	95.836	96.162	1.63	1.76	90.8	82.3
A37	0.61	95.652	95.677	1.63	1.66	93.1	78.4
A38	1.33	96.687	95.739	1.66	1.67	82.2	75.5

[a] Mork Family Department of Chemical Engineering and Materials Science,
[b] Department of Chemistry, University of Southern California, Los Angeles CA 90089, USA.
E-mail: ssharada@usc.edu

2 Transferability Tests

Table 2 lists LFER descriptors calculated using **IS** geometries at various levels of theory.

Table 2: Bite angles θ (°), Sterimol parameters B1 (Å), BS DFT barriers, $\Delta E^{\ddagger}_{DFT}$, using various levels of theory and LFER model predictions $\Delta E^{\ddagger}_{LFER}$ for 11 symmetric substituted catalysts (kJ/mol). Hammett parameters are based on experiment and therefore invariant with level of theory

	X_{im}	θ_1	θ_2	$B1_1$	$B1_2$	$\Delta E^{\ddagger}_{DFT}$	$\Delta E^{\ddagger}_{LFER}$
PBE0/	$N(CH_3)_2$	97.871	94.480	1.96	1.84	82.0	104.1
6-311+G*	$NHCH_3$	96.969	97.150	1.50	1.50	82.1	88.4
(Cu: srsc)	NH_2	96.979	97.235	1.57	1.55	73.6	89.6
	OH	98.888	98.876	1.66	1.65	71.5	89.1
	OCH_3	96.866	98.978	1.65	1.64	86.6	84.0
	CH_3	96.752	92.646	1.6	1.71	71.9	74.8
	H	96.102	96.111	1.66	1.65	73.4	73.4
	CHO	95.333	95.248	1.62	1.61	67.2	59.1
	COOH	97.618	97.561	1.62	1.62	62.6	63.5
	CF_3	95.958	98.341	1.91	1.77	54.4	67.7
	NO_2	97.336	97.750	1.62	1.61	45.2	54.5
SCAN0/	$N(CH_3)_2$	94.580	98.370	1.95	1.92	62.0	85.6
6-311+G*	$NHCH_3$	96.874	96.742	1.59	1.57	65.0	70.2
(Cu: srsc)	NH_2	97.006	97.099	1.57	1.59	56.5	69.7
	OH	98.724	98.938	1.66	1.68	57.8	69.1
	OCH_3	98.754	98.896	1.67	1.70	75.6	66.9
	CH_3	92.672	92.937	1.72	1.71	50.6	52.4
	H	95.678	95.803	1.65	1.62	51.4	51.4
	CHO	95.278	95.264	1.62	1.63	51.2	39.0
	COOH	97.469	97.316	1.62	1.61	47.7	42.4
	CF_3	95.536	97.808	1.79	1.88	41.6	46.0
	NO_2	97.179	96.610	1.65	1.59	30.4	32.9
ωB97X-V/	$N(CH_3)_2$	95.003	97.866	1.95	1.78	109.4	131.5
6-311+G*	$NHCH_3$	96.982	96.908	1.58	1.57	114.3	118.9
(Cu: srsc)	NH_2	97.133	97.376	1.60	1.59	105.9	119.2
	OH	97.614	97.185	1.67	1.66	106.3	114.3
	OCH_3	98.704	98.828	1.66	1.66	127.1	114.4
	CH_3	93.002	96.016	1.75	1.63	102.3	103.7
	H	96.014	95.926	1.61	1.65	100.3	100.3
	CHO	95.535	95.406	1.60	1.60	100.4	87.0
	COOH	97.633	97.550	1.62	1.61	95.2	91.3
	CF_3	97.948	95.579	1.91	1.77	79.6	94.9
	NO_2	97.185	97.614	1.59	1.64	77.3	82.3
ωB97X-D/	$N(CH_3)_2$	94.638	97.840	1.96	1.79	159.8	185.2
6-31G*	$NHCH_3$	96.649	96.556	1.63	1.58	164.8	173.0
	NH_2	96.665	96.743	1.60	1.58	160.0	171.6
	OH	98.090	98.219	1.67	1.67	159.7	169.9
	OCH_3	97.904	98.066	1.69	1.70	177.0	167.7
	CH_3	92.881	96.125	1.72	1.71	158.2	158.4
	H	95.706	95.534	1.63	1.64	153.4	153.4
	CHO	95.272	95.261	1.61	1.63	155.4	141.1
	COOH	97.442	96.944	1.61	1.62	145.0	144.2
	CF_3	97.961	95.365	1.89	1.77	144.4	148.1
	NO_2	97.279	97.168	1.60	1.60	128.8	135.1
ωB97X-D/	$N(CH_3)_2$	98.896	94.725	1.90	1.78	71.4	99.0
def2-SVP	$NHCH_3$	97.294	97.262	1.55	1.56	85.6	86.5
	NH_2	97.576	97.459	1.57	1.57	79.6	86.4

Continued on next page

Table 2: Bite angles θ (°), Sterimol parameters B1 (Å), BS DFT barriers, $\Delta E^{\ddagger}_{DFT}$, using various levels of theory and LFER model predictions $\Delta E^{\ddagger}_{LFER}$ for 11 symmetric substituted catalysts (kJ/mol). Hammett parameters are based on experiment and therefore invariant with level of theory

	X_{im}	θ_1	θ_2	$B1_1$	$B1_2$	$\Delta E^{\ddagger}_{DFT}$	$\Delta E^{\ddagger}_{LFER}$
	OH	99.589	99.445	1.66	1.67	74.3	86.5
	OCH_3	99.488	99.385	1.68	1.7	91.6	84.5
	CH_3	96.593	93.017	1.71	1.73	70.1	73.0
	H	96.362	96.217	1.64	1.62	68.5	68.5
	CHO	95.477	95.455	1.62	1.62	63.5	55.3
	COOH	96.883	97.575	1.63	1.63	62.2	58.7
	CF_3	98.317	96.102	1.90	1.77	53.8	63.3
	NO_2	96.992	97.136	1.64	1.60	44.4	49.3
ωB97X-D/	$N(CH_3)_2$	94.882	99.079	1.9	1.76	82.0	101.8
def2-SVPD	$NHCH_3$	96.976	96.910	1.59	1.59	88.5	89.7
	NH_2	97.112	97.103	1.60	1.59	80.9	89.2
	OH	98.977	99.266	1.66	1.67	74.6	88.4
	OCH_3	98.894	99.134	1.68	1.71	95.0	86.5
	CH_3	92.874	96.615	1.76	1.73	75.4	76.5
	H	95.866	96.255	1.65	1.63	71.2	71.2
	CHO	95.165	95.127	1.63	1.63	68.8	57.7
	COOH	97.499	97.088	1.62	1.62	64.8	61.2
	CF_3	95.474	98.349	1.89	1.77	61.7	65.2
	NO_2	97.399	97.191	1.64	1.60	47.5	52.6

3 Sensitivity Analysis

Figure 1 shows the results of SOBOL sensitivity analysis, indicating the dependence of model performance on the three sets of descriptors chosen to construct the LFER.

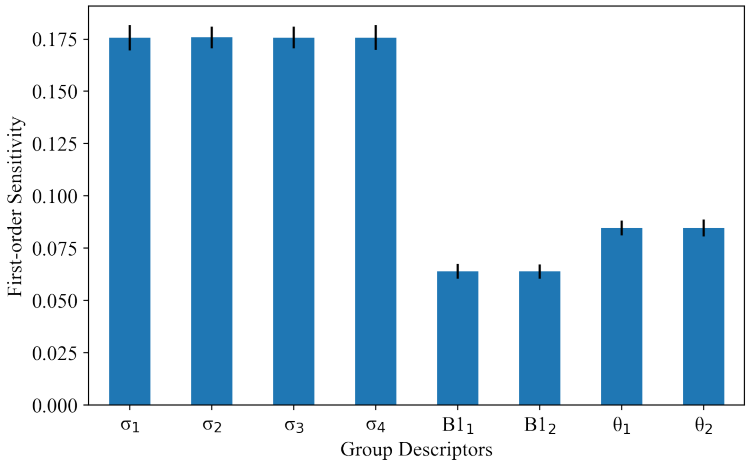

Figure 1: LFER model sensitivity analysis to interaction (σ_p) and strain ($B1, \theta$) descriptors.

4 LFER Model Predictions: Interaction & Strain Energies

Table 3 shows differences between LFER and DFT values of ASM components for A1-A38. Catalyst and substrate strain terms are listed in Table 4 for the analysis of sources of model error.

Table 3: Differences between LFER and DFT-based (ωB97X-D/srsc, 6-311+G*) ΔE_{INT}, ΔE_{STR}, and $\Delta E^{\ddagger}$ (kJ/mol) for **A1-A38**. ASM components for the outliers (Main text, Figure 4) for which the magnitudes of LFER deviations exceed 10 kJ/mol are highlighted in boldface.

Index	$\Delta E_{INT,LFER}$-$\Delta E_{INT,DFT}$	$\Delta E_{STR,LFER}$-$\Delta E_{STR,DFT}$	$\Delta E^{\ddagger}_{LFER}$-$\Delta E^{\ddagger}_{DFT}$
A1	4.7	-0.7	4.0
A2	6.3	-5.5	0.8
A3	-3.1	10.4	7.4
A4	-17.9	21.8	3.9
A5	-0.6	0.9	0.3
A6	**-7.0**	**22.4**	**15.4**
A7	-12.4	17.9	5.5
A8	-1.7	5.6	3.9
A9	-12.8	18.2	5.4
A10	**-17.1**	**34.8**	**17.7**
A11	-2.0	-1.3	-3.3
A12	**1.1**	**-14.5**	**-13.4**
A13	-4.2	13.8	9.6
A14	-1.6	-5.2	-6.8
A15	-5.8	-0.1	-5.9
A16	**6.9**	**-17.7**	**-10.8**
A17	**-9.6**	**21.1**	**11.5**
A18	0.6	-7.2	-6.5
A19	-4.2	3.2	-0.9
A20	7.6	-10.2	-2.6
A21	-13.2	18.4	5.2
A22	-0.6	1.7	1.1
A23	-12.2	9.0	-3.3
A24	3.7	-9.9	-6.3
A25	-19.6	19.5	-0.2
A26	1.2	-0.9	0.3
A27	-17.9	19.0	1.1
A28	-1.5	8.0	6.5
A29	-9.0	10.6	1.6
A30	-6.9	14.5	7.6
A31	1.3	2.0	3.3
A32	-13.8	17.4	3.7
A33	-2.8	-1.8	-4.6
A34	-9.2	8.8	-0.4
A35	-10.9	5.9	-5.0
A36	-7.4	-1.0	-8.4
A37	**-7.7**	**-7.1**	**-14.7**
A38	-8.4	1.7	-6.7

Table 4: Catalyst ($\Delta E_{STR,CAT}$) and substrate strain energies ($\Delta E_{STR,SUB}$) in kJ/mol for **S1-S11** and **A1-A38** calculated using EDA fragment energy differences (ωB97X-D/srsc, 6-311+G*).

Index	$\Delta E_{STR,SUB}$	$\Delta E_{STR,CAT}$
S1	121.3	61.2
S2	131.3	67.2
S3	130.1	63.9
S4	128.1	63.2
S5	132.0	74.2
S6	139.1	57.0
S7	126.8	62.0
S8	126.5	74.0
S9	124.2	52.1
S10	114.0	68.0
S11	111.1	51.1
A1	133.0	61.2
A2	136.1	62.5
A3	123.9	63.0
A4	116.1	57.1
A5	127.9	67.2
A6	128.7	48.1
A7	124.3	51.5
A8	135.8	57.1
A9	117.5	55.2
A10	114.5	52.6
A11	130.5	65.6
A12	134.2	72.9
A13	123.0	58.4
A14	129.5	64.1
A15	126.4	67.9
A16	137.7	71.7
A17	115.1	56.5
A18	131.4	66.8
A19	123.5	70.6
A20	138.5	69.3
A21	120.6	55.2
A22	131.2	61.2
A23	116.3	66.6
A24	136.8	68.5
A25	114.9	57.7
A26	133.0	63.8
A27	113.2	62.0
A28	126.7	61.7
A29	125.5	55.2
A30	126.0	59.1
A31	132.6	61.3
A32	121.1	55.3
A33	127.9	62.9
A34	120.0	64.1
A35	121.8	62.5
A36	120.0	68.5
A37	126.9	68.9
A38	127.0	65.3.

5 Intramolecular Interactions & Catalyst Strain

Analysis of ligand-ligand interaction differences between **TSoxo** and **IS** is carried out by separating N-donors coordinating to different Cu atoms and carrying out EDA for each pair of N-donors.

Table 5: Interaction energy differences between **TSoxo** and **IS**, ΔE_{INT} (ωB97X-D/srsc/6-311+G*, kJ/mol), for **S1-S11**. Subscript 1 and 2 denote fragment 1 and 2, each composed of two N-donors connecting to the same Cu. $\Delta E_{INT,total}$ is calculated via $\Delta E_{INT,1}$ + $\Delta E_{INT,2}$. ΔE_{INT} for the outliers for which the magnitudes of LFER deviations exceed 10 kJ/mol are highlighted in boldface.

Index	X_{im}	$\Delta E_{INT,1}$	$\Delta E_{INT,2}$	$\Delta E_{INT,total}$
S1	**$N(CH_3)_2$**	**-27.1**	**10.6**	**-16.5**
S2	$NHCH_3$	-5.8	-4.9	-10.7
S3	**NH_2**	**-5.0**	**-6.1**	**-11.1**
S4	**OH**	**-8.1**	**-6.8**	**-14.9**
S5	OCH_3	-12.3	-8.6	-20.9
S6	CH_3	1.0	-15.5	-14.5
S7	H	-5.6	-4.8	-10.4
S8	**CHO**	**-3.9**	**-12.8**	**-16.7**
S9	COOH	-4.1	-4.7	-8.8
S10	CF_3	-0.2	-5.1	-5.3
S11	NO_2	-3.5	-3.6	-7.1

6 Strain LFER: Compensation

Table 6 shows descriptors, interaction energy, and DFT barriers for one reference and seven bidentate N-donor **SI1-SI7** substituted dicopper systems. All descriptors are calculated for optimized bare catalysts while ΔE_{INT} and ΔE_{STR} are calculated by the difference between **TSoxo** and **IS**. R^2 for the linear fit of $\Delta E_{STR,LFER}$-$\Delta E_{STR,DFT}$ versus $\Delta E^{\ddagger}_{INT,DFT}$ - $\Delta E^{\ddagger}_{INT,DFT,ref}$ is 0.93 and the slope is unity. Geometries of these bidendate N-donor substituents are shown in Figure 2.

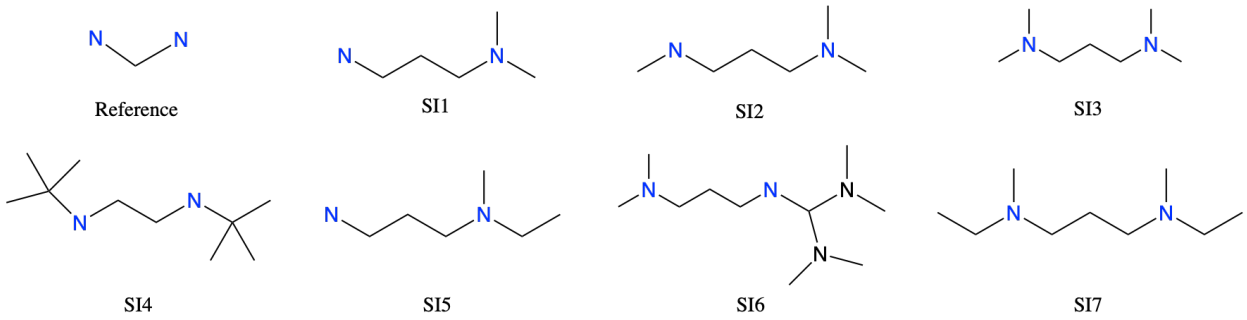

Figure 2: Geometries of eight bidentate N-donor substituent groups.

Table 6: Bite angles θ (°) and Sterimol parameters B1 (Å), ΔE_{INT}, difference between LFER and DFT-based (ωB97X-D/srsc/6-311+G*) ΔE_{STR} for eight bidentate N-donor substituted catalysts including a reference system (barriers in kJ/mol)

Index	θ_1	θ_2	$B1_1$	$B1_2$	$\Delta E^{\ddagger}_{INT,DFT}$ - $\Delta E^{\ddagger}_{INT,DFT,ref}$	$\Delta E_{STR,LFER}$-$\Delta E_{STR,DFT}$
Reference	71.805	71.809	1.55	1.55	0.0	0
SI1	96.299	96.257	2.83	1.64	20.0	15.0
SI2	95.448	95.449	2.84	2.43	24.4	17.2
SI3	100.864	100.757	2.83	2.84	30.1	29.3
SI4	86.649	85.547	3.21	2.02	18.6	13.5
SI5	96.936	97.782	1.65	2.85	24.1	16.0
SI6	94.381	94.381	3.25	2.82	30.5	25.7
SI7	101.145	102.677	2.85	2.83	39.6	34.6

7 Reference Analysis

LFER model predictions, in principle, must be independent of the choice of reference system. This is demonstrated in Table 7, where LFER 1 and 2 for amine groups (X_{am} = OCH_3, CH_3, H, CF_3, NO_2), constructed with different references, are two identical LFERs with the same Hammett slope. When we utilize these to predict the reactivity for the systems examined in this study, the different intercepts and different references cancel out,resulting in identical predictions. Similarly, LFER 3 and 4 are the same with different references for imidazole groups (X_{im} = OCH_3, CH_3, H, CF_3, NO_2). To match the form of the Taft equation where there is no intercept, we set intercept to 0 in the linear regression model and obtain LFER 5 and 6 with smaller MAE but a cost of slightly decreasing R^2. We note however that the choice of reference is critical to model construction. Our earlier study predicts poor model fit for an amine reference (LFERs 1, 2, or 5) reflected in the low R^2 value. This leads to very large MAEs when applied to the systems examined here. We utilize LFER 6 as the interaction component of the LFER, described by Equation 7 in the main text.

Table 7: Analysis of linear regression model for two different references — $X_{am} = H$ and $X_{im} = H$. The 'group' column refers to the N-donor variations employed in construction of the interaction term. The electronic LFERs for frozen monodentate complexes are established for amine and imidazole groups with mean absolute error (MAE) (kJ/mol) and coefficient determination (R^2) calculated. LFER 1-4 are constructed without a fixed intercept in the linear regression model and LFER 5-6 are constructed with intercept set to zero.

LFER index	Reference	Group	$-log(k/k_{ref})$	MAE	R^2
1	$X_{am} = H$	amine	-3.37$\sum_{i=1}^{4} \sigma_{pi} + 1.46$	15.02	0.635
2	$X_{im} = H$	amine	-3.37$\sum_{i=1}^{4} \sigma_{pi} - 4.68$	15.02	0.635
3	$X_{am} = H$	imidazole	-2.28$\sum_{i=1}^{4} \sigma_{pi} + 7.78$	3.70	0.873
4	$X_{im} = H$	imidazole	-2.28$\sum_{i=1}^{4} \sigma_{pi} - 0.23$	3.70	0.873
5	$X_{am} = H$	amine	-3.05$\sum_{i=1}^{4} \sigma_{pi}$	14.07	0.589
6	$X_{im} = H$	imidazole	-2.33$\sum_{i=1}^{4} \sigma_{pi}$	3.55	0.871