Electronic supplementary materials *Mendeleev Commun.*, 2024, **34**, 351–353

Comparison of assessments of overall toxicity of chemical reactions upon using cytotoxicity and acute toxicity data

Ksenia S. Egorova, Alexandra V. Posvyatenko, Alexey S. Galushko and Andrey E. Kolesnikov

Contents

Building of bio-Strips and calculation of BFs and TPs

A bio-Strip is a visual representation of a chemical reaction in accordance with the amounts (*n*, mmol) and half-maximal cytotoxicity concentrations (CC_{50}, $mmol \cdot L^{-1}$, mM) or median lethal doses (here - LD_{50}, $mmol \cdot kg^{-1}$ b.w.) of all its participant substances. In the bio-Strip, the length of each section corresponds to the "normalized (cyto)toxicity" (in the case of CC_{50} – NC; see equation 1) of the corresponding substance (starting material (SM), catalyst (CT), additional reagent (R), solvent (S), product (P), or byproduct (BP).

$$\mathrm{NT} = \frac{n}{\mathrm{CC_{50}}} \quad (1)$$

The color of the sections shows the CC_{50} or LD_{50} of the corresponding compounds measured in a particular biological object (here – the cell lines CaCo-2 (colorectal adenocarcinoma), HEK293T (human embryonic kidney), and FRSN (mesenchymal stem cells from foreskin) or rats upon oral administration). The lowest value is red, while the highest value is green.

For each reaction, the following metrics are calculated: the bio-Factor (BF), which shows the change in the "overall (cyto)toxicity" during the reaction (see equation 2), and three (cyto)toxicity potentials (TPs; in the case of cytotoxicity - CPs): the initial TP (TP_i), which characterizes the compounds entering the reaction (SM, CT, S, R; see equation 3); the final TP (TP_f), which characterizes the compounds leaving the reaction (P, BP and compounds that can be recycled or regenerated, such as CT and S; see equation 4), and the relative final TP (TP_{f_rel}), which characterizes the compounds leaving the reaction except for the target product (BP and compounds that can be recycled or regenerated; see equation 5).

$$\mathrm{BF} = \frac{\sum \mathrm{NT}_{out}}{\sum \mathrm{NT}_{in}} = \frac{\sum \frac{n}{\mathrm{CC_{50}}}(out)}{\sum \frac{n}{\mathrm{CC_{50}}}(in)} \quad (2),$$

$$\mathrm{TP}_i = \sum \mathrm{NT}_{in} \quad (3),$$

$$\mathrm{TP}_f = \sum \mathrm{NT}_{out} \quad (4),$$

$$\mathrm{TP}_{f\_rel} = \sum \mathrm{NT}_{out} - \mathrm{NT}_{product} \quad (5),$$

where *in* and *out* mark the compounds entering or leaving the reaction, respectively. According to the above-shown equations, lower TP values correspond to reactions with lower "overall toxicity".

Table S1. Experimental data used in this work.

Substance	Mw, g·mol^{-1}	Amount in reaction, g [a]	24-h CC_{50}, mM [b]			LD_{50}, mmol·kg^{-1} (rat, oral) [c]
			CaCo-2	FRSN	HEK293T	
Starting materials						
Phenylboronic acid	121.93	0.122	14.158 (8.824-19.491)	14.378 (9.347-19.408)	3.072 (1.824-4.320)	6.069
Iodobenzene	204.01	0.204	3.349 (1.595-5.104)	3.333 (1.915-4.750)	9.543 (6.517-12.568)	8.573
Bromobenzene	157.01	0.157	5.546 (4.449-6.643)	4.803 (1.874-7.732)	17.497 (12.613-22.380)	15.177
Catalysts						
$Pd(OAc)_2$	224.50	0.002	1.035 (0.421-1.650)	0.709 (0.510-0.908)	1.096 (0.640-1.553)	22.762
$PdCl_2$	177.33	0.002	0.698 (0.522-0.875)	0.385 (0.305-0.465)	0.724 (0.347-1.102)	15.248
$Pd(acac)_2$	304.64	0.003	0.037 (0.026-0.049)	0.007 (0.005-0.009)	0.014 (0.009-0.019)	6.565
Reagents						
Na_2CO_3	105.988	0.106	45.423 (31.753-59.094)	46.313 (39.130-53.497)	32.063 (24.104-40.022)	38.589
K_2CO_3	138.205	0.138	29.997 (23.581-36.413)	49.953 (24.034-75.873)	32.360 (25.446-39.274)	13.531
Solvents						
Ethanol	46.07	1.263	598.875 (456.787-740.963)	1266.750 (895.474-1638.026)	607.700 (504.107-711.293)	153.245
NMP	99.13	1.648	144.300 (108.218-180.382)	118.167 (86.989-149.345)	192.500 (161.918-223.082)	39.484
Products						
1,1'-biphenyl	154.21	0.154	32.307 (30.673-33.940)	16.593 (7.768-25.419)	>20 [d]	21.270
Byproducts						
NaI	149.8940	<0.15	184.733 (167.782-201.684)	127.267 (89.611-164.922)	80.103 (77.624-82.583)	28.954
NaBr	102.8900	<0.103	185.880 (107.519-264.241)	163.100 (137.261-188.939)	80.040 (64.682-95.398)	34.017
KI	166.0030	<0.166	112.300 (103.109-121.491)	64.477 (57.074-71.880)	55.627 (49.419-61.835)	16.741
KBr	119.0100	<0.119	314.733 (256.230-373.237)	89.727 (55.567-123.886)	67.760 (42.527-92.993)	25.796
$Na_2B_4O_7$ [e]	381.37	<0.076	37.453 (23.232-51.674)	43.623 (26.873-60.374)	46.313 (39.594-53.033)	6.975
$K_2B_4O_7$ [e]	305.5	<0.06	13.213 (6.274-20.152)	32.343 (20.201-44.486)	24.583 (22.180-26.985)	>8.183
H_3BO_3	61.84	<0.06	246.075 (165.622-326.528)	120.067 (76.847-163.287)	115.788 (75.313-156.262)	43.014

[a] 100% conversion is considered. For clarity, maximal possible amounts of byproducts were used in calculations. [b] 95% confidence intervals are shown in parentheses. [c] LD_{50} values for rats upon oral administration are obtained from the NCBI PubChem database or from Sigma-Aldrich safety data sheets. [d] The exact value could not be measured due to technical difficulties (insufficient solubility in the cultural media). [e] $K_2B_4O_7{\cdot}4H_2O$ and $Na_2B_4O_7{\cdot}10H_2O$ were tested.

Table S2. bio-Factors and toxicity potentials for synthesis of 1,1'-biphenyl.

Reaction	SM2 (ArX)	CT (PdA_2)	R (M_2CO_3)	S	BP1 (MX)	BP2 ($M_2B_4O_7$)	CaCo-2				FRSN				HEK293T				Rat (oral)			
							BF	TP_i	TP_f	TP_{f_rel}	BF	TP_i	TP_f	TP_{f_rel}	BF	TP_i	TP_f	TP_{f_rel}	BF	TP_i	TP_f	TP_{f_rel}
A-A-A-A	I	OAc	Na	EtOH	NaI	Na	0.23	0.45	0.1	0.07	0.27	0.43	0.12	0.06	0.25	0.52	0.13	0.08	0.64	0.49	0.31	0.27
A-A-A-B	I	OAc	Na	NMP	NaI	Na	0.33	0.52	0.17	0.14	0.43	0.55	0.24	0.18	0.31	0.56	0.17	0.12	0.76	0.73	0.55	0.51
A-A-B-A	I	OAc	K	EtOH	KI	K	0.25	0.46	0.11	0.08	0.3	0.43	0.13	0.07	0.27	0.52	0.14	0.09	0.62	0.53	0.33	0.29
A-A-B-B	I	OAc	K	NMP	KI	K	0.35	0.53	0.18	0.15	0.45	0.54	0.24	0.18	0.32	0.56	0.18	0.13	0.74	0.78	0.57	0.53
A-B-A-A	I	Cl	Na	EtOH	NaI	Na	0.23	0.45	0.11	0.07	0.29	0.44	0.13	0.07	0.26	0.52	0.13	0.08	0.64	0.49	0.31	0.27
A-B-A-B	I	Cl	Na	NMP	NaI	Na	0.34	0.52	0.18	0.14	0.44	0.56	0.25	0.19	0.31	0.56	0.18	0.13	0.76	0.73	0.55	0.51
A-B-B-A	I	Cl	K	EtOH	KI	K	0.26	0.46	0.12	0.09	0.31	0.44	0.14	0.08	0.28	0.52	0.14	0.09	0.62	0.53	0.33	0.29
A-B-B-B	I	Cl	K	NMP	KI	K	0.35	0.53	0.19	0.16	0.46	0.56	0.26	0.2	0.33	0.56	0.18	0.13	0.74	0.78	0.57	0.53
A-C-A-A	I	acac	Na	EtOH	NaI	Na	0.51	0.71	0.36	0.33	0.83	1.84	1.53	1.47	0.68	1.22	0.83	0.78	0.64	0.49	0.31	0.27
A-C-A-B	I	acac	Na	NMP	NaI	Na	0.55	0.78	0.43	0.4	0.84	1.96	1.65	1.59	0.69	1.26	0.88	0.83	0.76	0.73	0.56	0.51
A-C-B-A	I	acac	K	EtOH	KI	K	0.52	0.72	0.37	0.34	0.84	1.84	1.54	1.48	0.69	1.22	0.84	0.79	0.62	0.54	0.33	0.29
A-C-B-B	I	acac	K	NMP	KI	K	0.56	0.79	0.44	0.41	0.85	1.96	1.66	1.6	0.7	1.26	0.88	0.83	0.74	0.78	0.58	0.53
B-A-A-A	Br	OAc	Na	EtOH	NaBr	Na	0.31	0.33	0.1	0.07	0.34	0.34	0.11	0.05	0.28	0.47	0.13	0.08	0.7	0.44	0.31	0.26
B-A-A-B	Br	OAc	Na	NMP	NaBr	Na	0.43	0.4	0.17	0.14	0.51	0.45	0.23	0.17	0.33	0.51	0.17	0.12	0.81	0.68	0.55	0.5
B-A-B-A	Br	OAc	K	EtOH	KBr	K	0.32	0.34	0.11	0.08	0.36	0.33	0.12	0.06	0.29	0.47	0.14	0.09	0.64	0.48	0.31	0.26
B-A-B-B	Br	OAc	K	NMP	KBr	K	0.43	0.41	0.18	0.15	0.53	0.45	0.24	0.18	0.35	0.51	0.18	0.13	0.76	0.73	0.55	0.51
B-B-A-A	Br	Cl	Na	EtOH	NaBr	Na	0.32	0.33	0.11	0.07	0.36	0.35	0.13	0.07	0.28	0.47	0.13	0.08	0.7	0.44	0.31	0.26
B-B-A-B	Br	Cl	Na	NMP	NaBr	Na	0.43	0.4	0.18	0.14	0.53	0.47	0.25	0.19	0.34	0.51	0.18	0.13	0.81	0.68	0.55	0.5
B-B-B-A	Br	Cl	K	EtOH	KBr	K	0.33	0.34	0.11	0.08	0.39	0.35	0.13	0.07	0.3	0.47	0.14	0.09	0.64	0.48	0.31	0.26
B-B-B-B	Br	Cl	K	NMP	KBr	K	0.44	0.41	0.18	0.15	0.54	0.46	0.25	0.19	0.35	0.51	0.18	0.13	0.76	0.73	0.55	0.51
B-C-A-A	Br	acac	Na	EtOH	NaBr	Na	0.61	0.59	0.36	0.33	0.87	1.75	1.53	1.47	0.71	1.17	0.83	0.78	0.7	0.44	0.31	0.26
B-C-A-B	Br	acac	Na	NMP	NaBr	Na	0.65	0.66	0.43	0.4	0.88	1.87	1.65	1.59	0.72	1.21	0.88	0.83	0.81	0.68	0.55	0.5
B-C-B-A	Br	acac	K	EtOH	KBr	K	0.61	0.6	0.37	0.34	0.88	1.75	1.54	1.48	0.72	1.17	0.84	0.79	0.64	0.49	0.31	0.27
B-C-B-B	Br	acac	K	NMP	KBr	K	0.65	0.67	0.44	0.41	0.89	1.87	1.65	1.59	0.73	1.21	0.88	0.83	0.76	0.73	0.55	0.51

Target product P (1,1'-biphenyl) and starting material SM1 (phenylboronic acid) are not shown. ArI, iodobenzene; ArBr, bromobenzene; EtOH, ethanol; NMP, *N*-methylpyrrolidone; SM, starting material; CT, catalyst; R, reagent; S, solvent; BP, byproduct; BF, bio-Factor; TP_i, initial toxicity potential; TP_f, final toxicity potential; TP_{f_rel}, relative final toxicity potential. The lowest TPs in each biological object are highlighted with green, whereas the names of the reactions with the lowest TPs in all the biological objects are highlighted with blue.

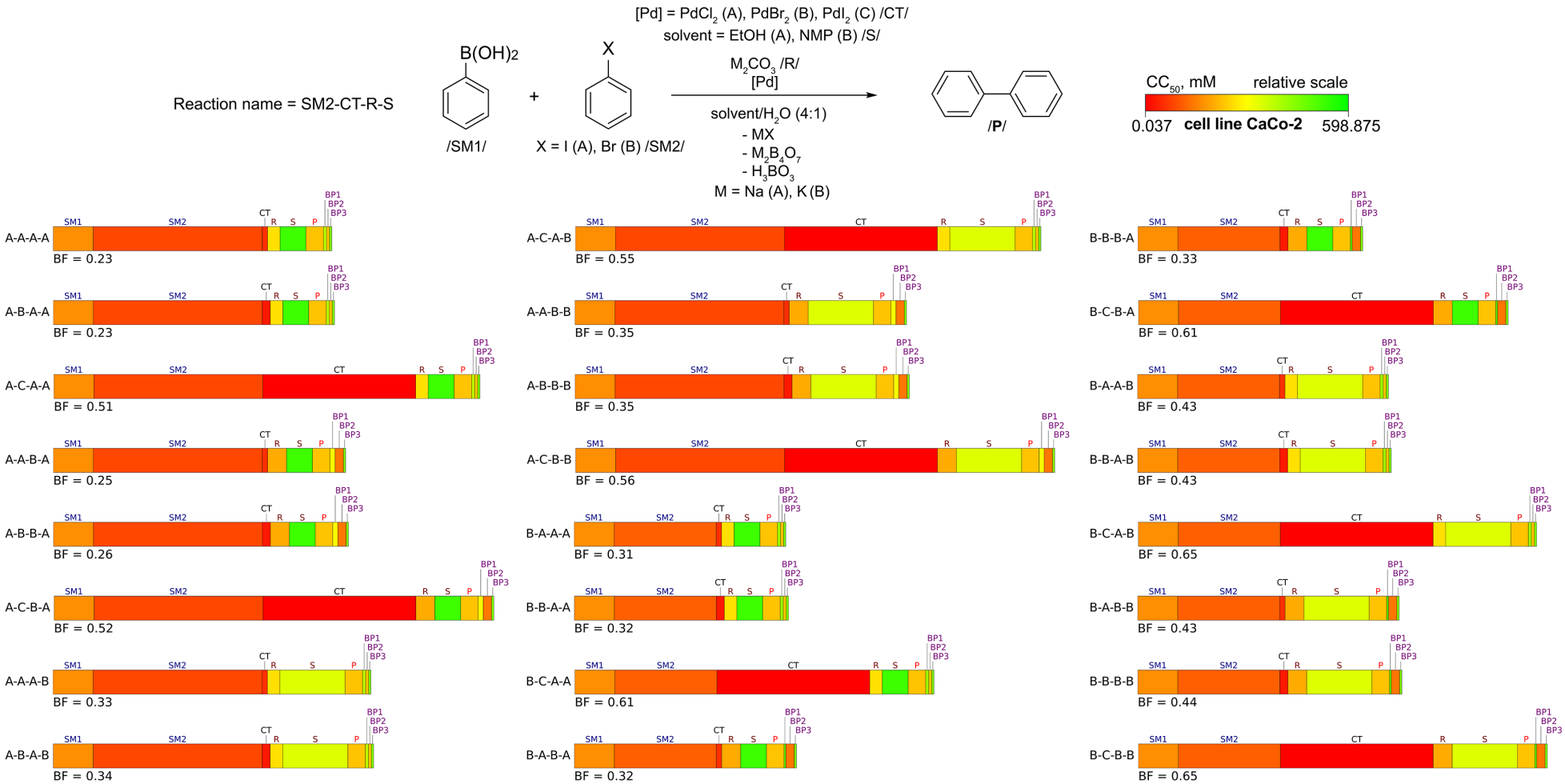

Figure S1. bio-Strips of 24 synthetic routes for 1,1'-biphenyl based on 24-h CC_{50} values measured in CaCo-2 cells. The 1st, 2nd, 3rd, and 4th letters in the reaction names correspond to the types of starting material 2 (SM2: iodobenzene (A) or bromobenzene (B)), catalyst (CT: $Pd(OAc)_2$ (A), $PdCl_2$ (B), or $Pd(acac)_2$ (C)), reagent (R: Na_2CO_3 (A) or K_2CO_3 (B)), and solvent (S: ethanol (A) or NMP (B)), respectively. The color of the strip sections corresponds to the CC_{50} of a particular substance in CaCo-2 cells (see the cytotoxicity scale above the bio-Strips). bio-Factors are shown below the strips.

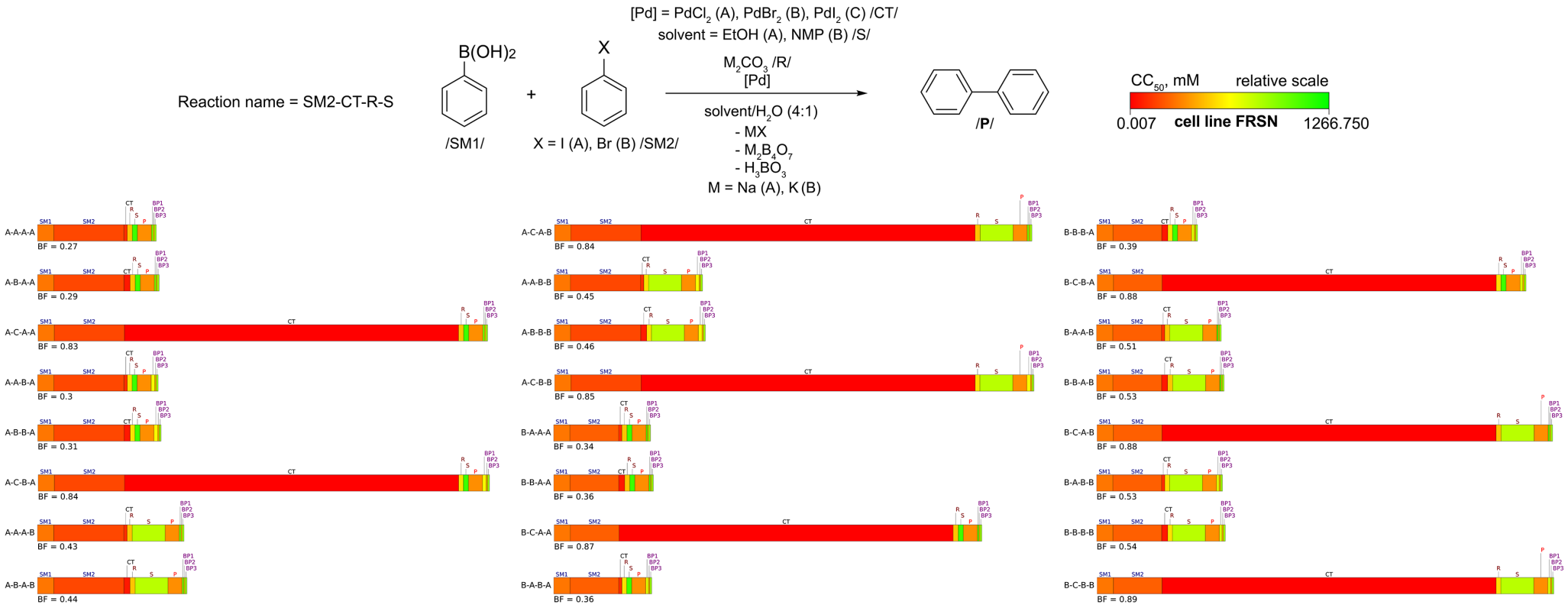

Figure S2. bio-Strips of 24 synthetic routes for 1,1'-biphenyl based on 24-h CC_{50} values measured in FRSN cells. The 1st, 2nd, 3rd, and 4th letters in the reaction names correspond to the types of starting material 2 (SM2: iodobenzene (A) or bromobenzene (B)), catalyst (CT: $Pd(OAc)_2$ (A), $PdCl_2$ (B), or $Pd(acac)_2$ (C)), reagent (R: Na_2CO_3 (A) or K_2CO_3 (B)), and solvent (S: ethanol (A) or NMP (B)), respectively. The color of the strip sections corresponds to the CC_{50} of a particular substance in FRSN cells (see the cytotoxicity scale above the bio-Strips). bio-Factors are shown below the strips.

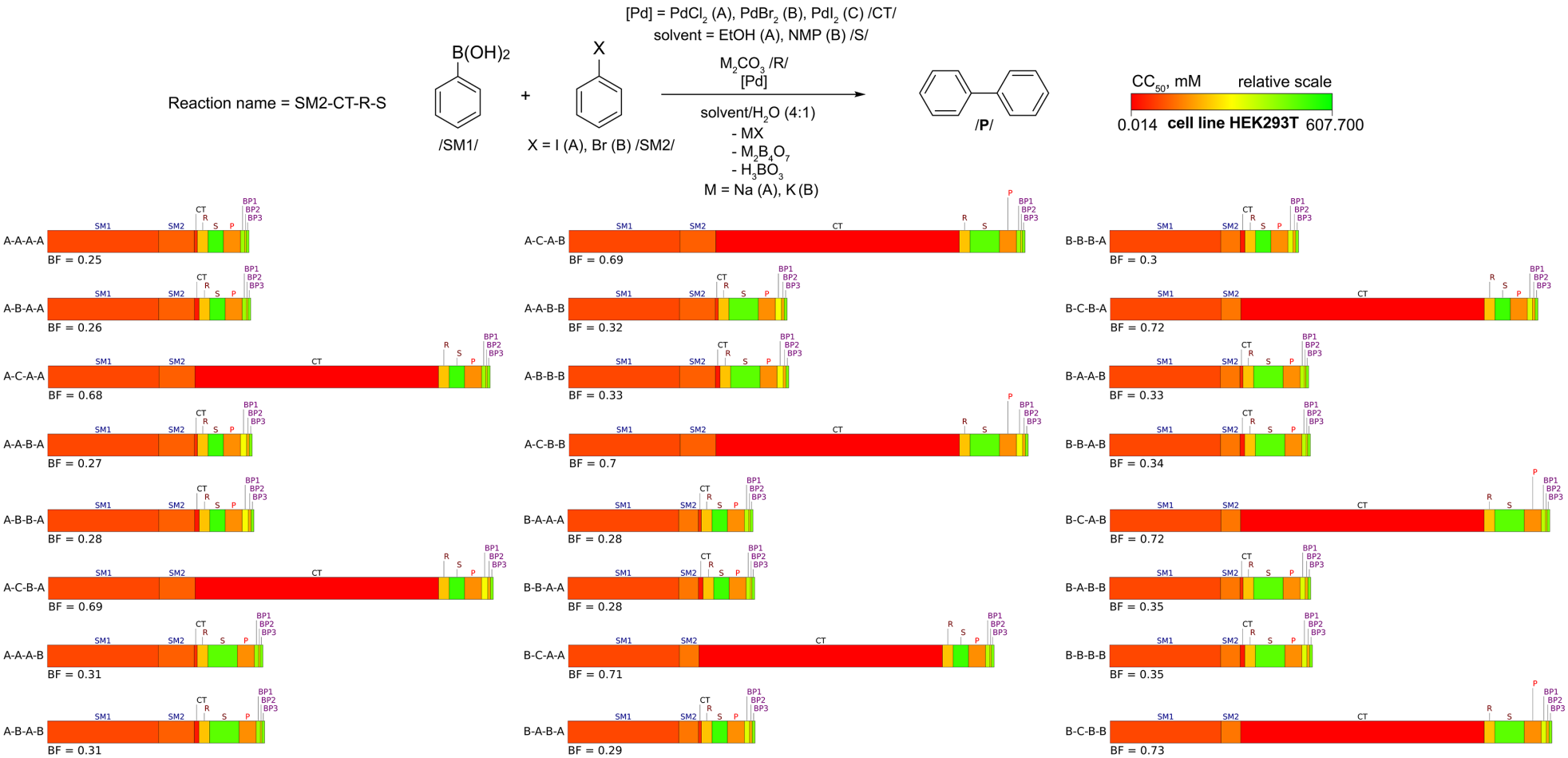

Figure S3. bio-Strips of 24 synthetic routes for 1,1'-biphenyl based on 24-h CC_{50} values measured in HEK293T cells. The 1st, 2nd, 3rd, and 4th letters in the reaction names correspond to the types of starting material 2 (SM2: iodobenzene (A) or bromobenzene (B)), catalyst (CT: $Pd(OAc)_2$ (A), $PdCl_2$ (B), or $Pd(acac)_2$ (C)), reagent (R: Na_2CO_3 (A) or K_2CO_3 (B)), and solvent (S: ethanol (A) or NMP (B)), respectively. The color of the strip sections corresponds to the CC_{50} of a particular substance in HEK293T cells (see the cytotoxicity scale above the bio-Strips). bio-Factors are shown below the strips.

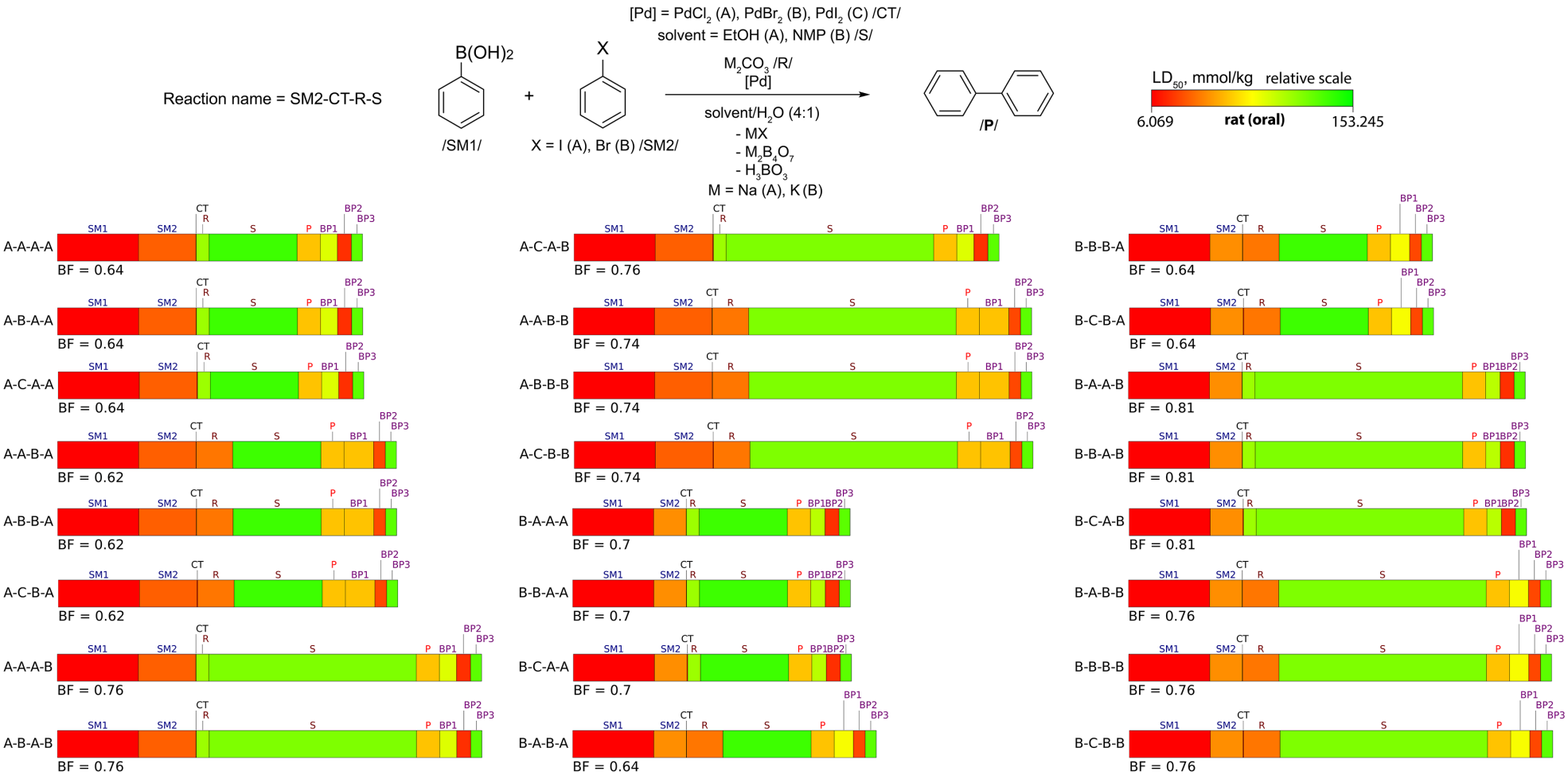

Figure S4. bio-Strips of 24 synthetic routes for 1,1'-biphenyl based on LD_{50} values for rats upon oral administration. The 1st, 2nd, 3rd, and 4th letters in the reaction names correspond to the types of starting material 2 (SM2: iodobenzene (A) or bromobenzene (B)), catalyst (CT: $Pd(OAc)_2$ (A), $PdCl_2$ (B), or $Pd(acac)_2$ (C)), reagent (R: Na_2CO_3 (A) or K_2CO_3 (B)), and solvent (S: ethanol (A) or NMP (B)), respectively. The color of the strip sections corresponds to the LD_{50} of a particular substance in rats (see the toxicity scale above the bio-Strips). bio-Factors are shown below the strips.