Supporting Information for

Predicting Fluorescence to Singlet Oxygen Generation Quantum Yield Ratio for BODIPY Dyes Using QSPR and Machine Learning

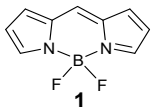
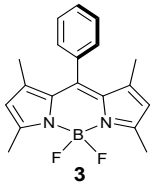
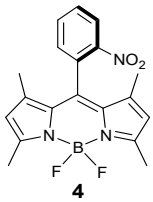
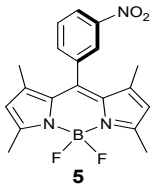
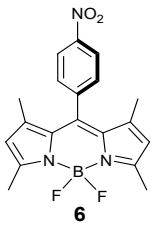
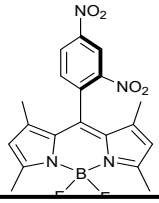
Platon P. Chebotaev,[a] Andrey A. Buglak,[a,b*] Aimee Sheehan,[c] Mikhail A. Filatov[c*]

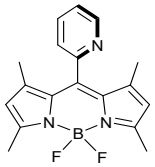
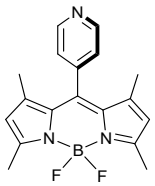
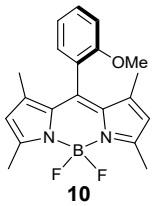
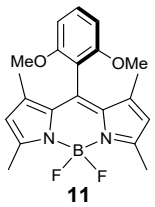
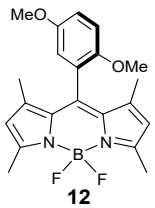
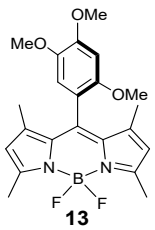
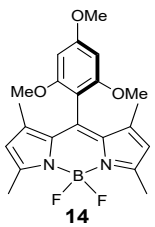
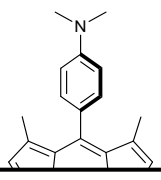
[a] Faculty of Physics, Saint-Petersburg State University, Universiteteskaya Emb. 7-9, 199034 St. Petersburg, Russia. E-mail: andreybuglak@gmail.com

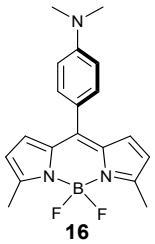
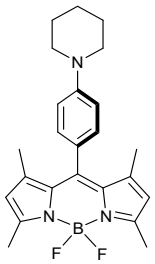
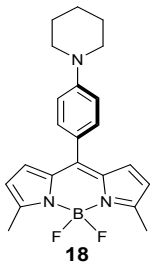
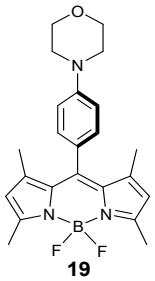
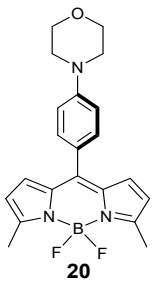
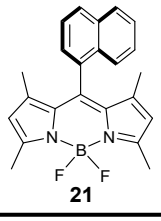
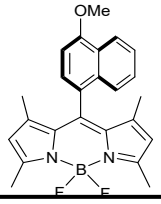
[b] Institute of Physics, Kazan Federal University, 18 Kremlyovskaya street, 420008, Kazan, Russia

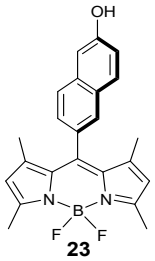
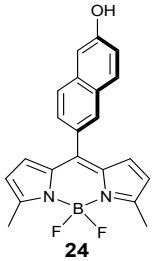
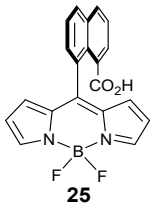
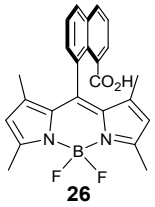
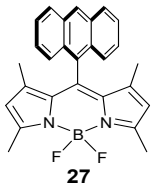
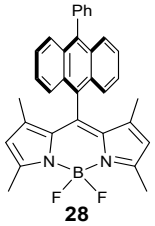
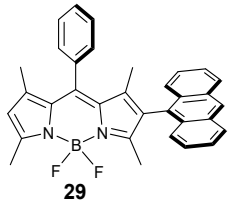
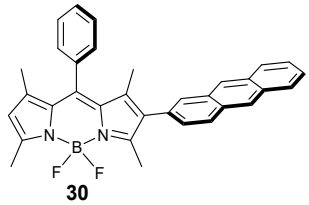
[c] School of Chemical and Biopharmaceutical Sciences, Technological University Dublin, City Campus, Grangegorman, Dublin 7, Ireland. E-mail: mikhail.filatov@tudublin.ie

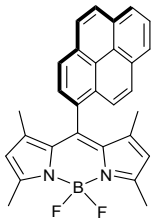
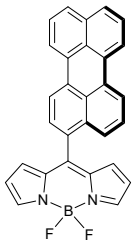
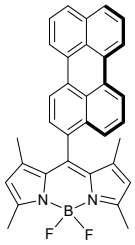
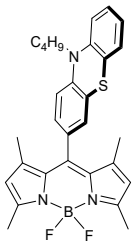
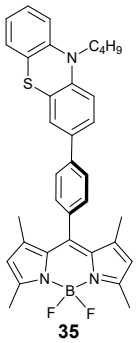
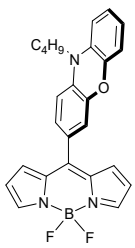
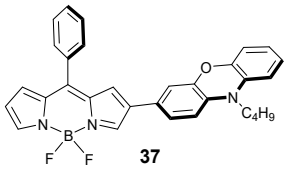
Table S1. BODIPY dataset.

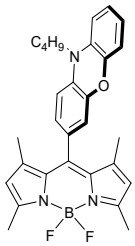
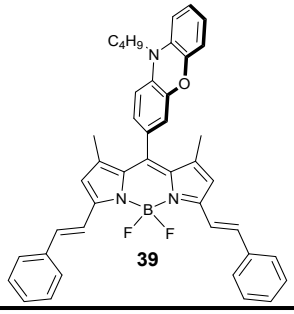
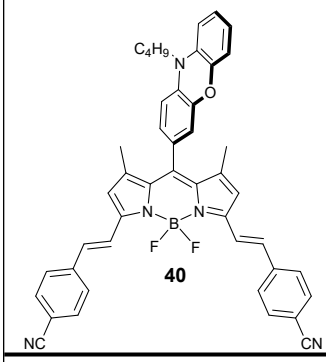
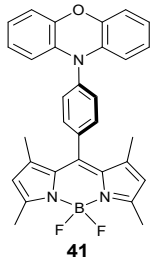
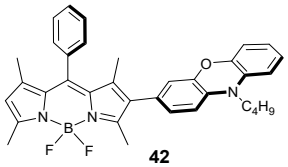
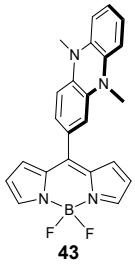
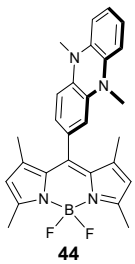
Structure	Solvent	Φ_{Δ}	Φ_{Fl}	Comment	Reference
 1	hexane	0.066	0.92		1
	toluene	0.12	0.79		1
	THF	0.071	0.86		1
	MeOH	0.083	0.96		1
 2	hexane	0.03	0.78		2
	toluene	0.061	0.74		2
	CCl_4	0.100	0.81		2
	CH_2Cl_2	0.062	0.74		2
	THF	0.091	-	Not reported	2
	EtOH	0.058	0.8		2
	CH_3CN	0.069	0.81		2
 3	hexane	0.038	0.56		3
	toluene	0.023			4
	EtOAc	0.052	0.58		3
	THF	0.13	0.56		4
	pinacolone	0.11	0.55		3
	acetone	0.050	0.46		3
	EtOH	0.030	-	Not reported	4
	MeOH	0.031	0.58		3
	CH_3CN	0.017	0.52		4
 4	hexane	0.018	0.023		3
	EtOAc	0.027	0.016		3
	THF	0.026	0.05		3
	pinacolone	0.079	0.0064		3
	acetone	0.051	0.0038		3
	MeOH	0.0083	0.027		3
	CH_3CN	0.020	0.0032		3
 5	hexane	0.01	0.42		3
	EtOAc	0.031	0.23		3
	THF	0.028	0.25		3
	pinacolone	0.07	0.074		3
	acetone	0.029	0.07		3
	MeOH	0.0062	0.14		3
	CH_3CN	0.0044	0.019		3
 6	hexane	0.0067	0.19		3
	EtOAc	0.021	0.14		3
	THF	0.019	0.11		3
	pinacolone	0.047	0.047		3
	acetone	0.0093	0.027		3
	MeOH	0.0036	0.03		3
	CH_3CN	0.0043	0.015		3
	hexane	0.021	0.033		3
	EtOAc	0.026	0.025		3
	THF	0.026	0.0087		3
	pinacolone	0.073	0.012		3
	acetone	0.012	0.0091		3
	MeOH	0.0055	0.045		3

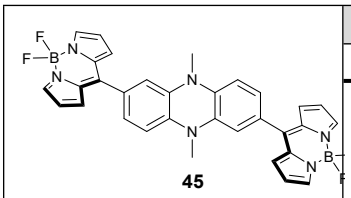
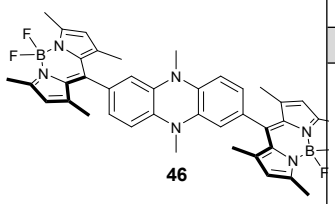
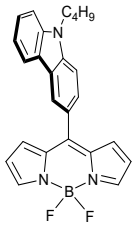
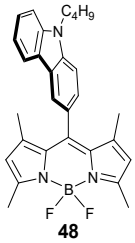
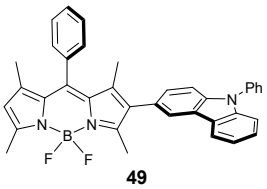
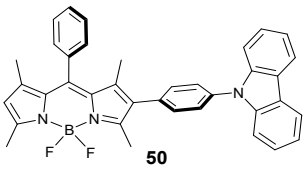
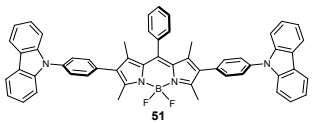
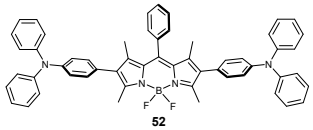
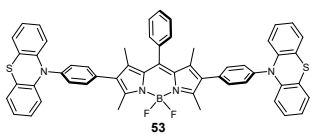
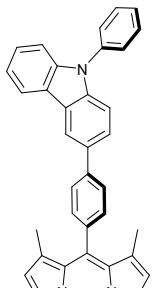
	CH_3CN	0.0049	0.0033		3
 **8**	hexane	0.0052	0.059		3
	EtOAc	0.0039	0.062		3
	THF	0.012	0.061		3
	pinacolone	0.012	0.04		3
	acetone	0.012	0.049		3
	MeOH	0.013	0.071		3
	CH_3CN	0.024	0.038		3
 **9**	hexane	0.0091	0.25		3
	EtOAc	0.032	0.38		3
	THF	0.024	0.33		3
	pinacolone	0.020	0.17		3
	acetone	0.031	0.2		3
	MeOH	0.012	0.12		3
	CH_3CN	0.037	0.13		3
 **10**	hexane	0.029	0.98		5
	EtOAc	0.057	0.76		5
	THF	0.061	0.87		5
	pinacolone	0.078	0.74		5
	acetone	0.17	0.78		5
	MeOH	0.021	0.69		5
	CH_3CN	0.18	0.57		5
 **11**	hexane	0.040	0.71		5
	EtOAc	0.073	0.47		5
	THF	0.051	0.64		5
	pinacolone	0.081	0.64		5
	acetone	0.082	0.72		5
	MeOH	0.036	0.58		5
	CH_3CN	0.18	0.54		5
 **12**	hexane	0.026	0.971		6
	EtOAc	0.178	0.499		6
	THF	0.462	0.457		6
	pinacolone	0.680	0.481		6
	acetone	0.250	0.059		6
	MeOH	0.023	0.037		6
	CH_3CN	0.125	0.01		6
 **13**	hexane	0.114	0.863		6
	EtOAc	0.291	0.004		6
	THF	0.357	0.005		6
	pinacolone	0.392	0.004		6
	acetone	0.068	0.003		6
	MeOH	0.004	0.003		6
	CH_3CN	0.033	0.001		6
 **14**	hexane	0.024	0.95		5
	EtOAc	0.063	0.64		5
	THF	0.059	0.78		5
	pinacolone	0.16	0.69		5
	acetone	0.11	0.8		5
	MeOH	0.074	0.68		5
	CH_3CN	0.31	0.55		5
	hexane	0.102	0.441		7
	EtOAc	0.412	0.011		7
	THF	0.623	0.014		7
	pinacolone	0.490	0.015		7
	acetone	0.114	0.009		7

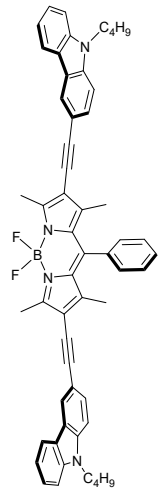
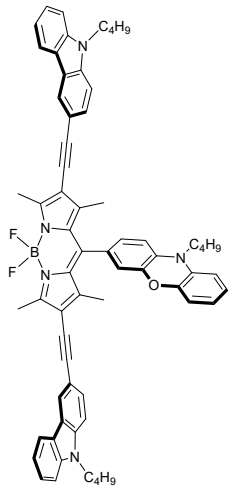
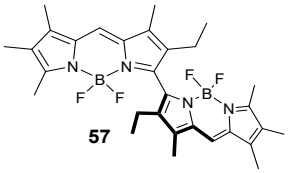
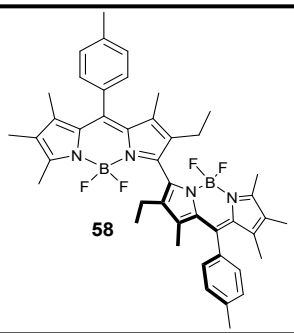
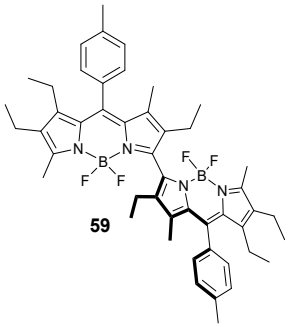
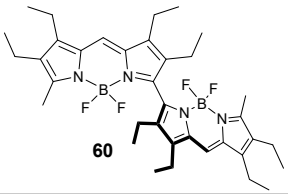
	MeOH	0.073	0.003		7
	CH_3CN	0.062	0.005		7
 **16**	hexane	0.23	0.376		7
	EtOAc	0.171	0.013		7
	THF	0.321	0.011		7
	pinacolone	0.439	0.024		7
	acetone	0.087	0.006		7
	MeOH	0.019	0.03		7
	CH_3CN	0.052	0.008		7
 **17**	hexane	0.058	0.384		7
	EtOAc	0.511	0.006		7
	THF	0.612	0.008		7
	pinacolone	0.644	0.024		7
	acetone	0.145	0.002		7
	MeOH	0.037	0.01		7
	CH_3CN	0.083	0.005		7
 **18**	hexane	0.258	0.341		7
	EtOAc	0.220	0.005		7
	THF	0.401	0.008		7
	pinacolone	0.457	0.004		7
	acetone	0.099	0.002		7
	MeOH	0.046	0.003		7
	CH_3CN	0.057	0.002		7
 **19**	hexane	0.120	0.45		7
	EtOAc	0.676	0.029		7
	THF	0.535	0.011		7
	pinacolone	0.588	0.063		7
	acetone	0.192	0.001		7
	MeOH	0.038	0.002		7
	CH_3CN	0.083	0.001		7
 **20**	hexane	0.225	0.388		7
	EtOAc	0.290	0.008		7
	THF	0.338	0.004		7
	pinacolone	0.281	0.002		7
	acetone	0.100	0.002		7
	MeOH	0.008	0.005		7
	CH_3CN	0.033	0.02		7
 **21**	hexane	0.05	0.87		4
	toluene	0.043	0.77		4
	THF	0.13	0.85		4
	EtOH	0.041	0.9		4
	CH_3CN	0.057	0.83		4
	hexane	0.011	0.906		6
	EtOAc	0.165	0.411		6
	THF	0.232	0.438		6
	pinacolone	0.460	0.479		6
	acetone	0.471	0.389		6
	MeOH	0.274	0.648		6

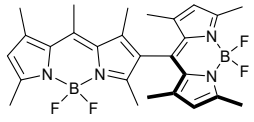
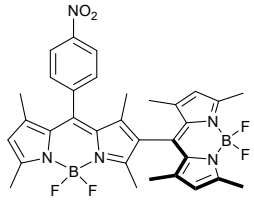
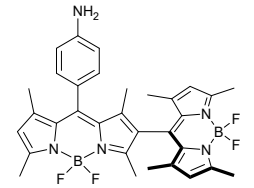
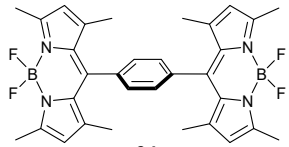
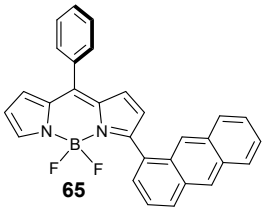
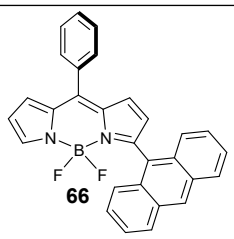
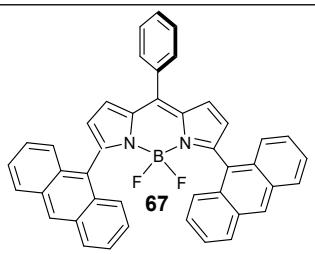
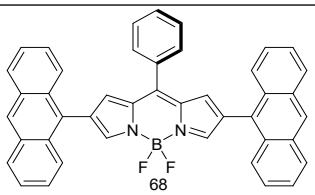
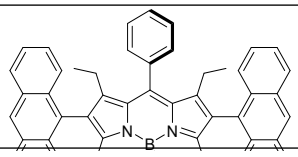
	CH_3CN	0.872	0.123		6
 **23**	hexane	0.047	0.637		6
	EtOAc	0.104	0.695		6
	THF	0.442	0.581		6
	pinacolone	0.382	0.621		6
	acetone	0.111	0.483		6
	MeOH	0.131	0.11		6
	CH_3CN	0.081	0.118		6
 **24**	hexane	0.116	0.269		6
	EtOAc	0.106	0.129		6
	THF	0.19	0.167		6
	pinacolone	0.317	0.11		6
	acetone	0.070	0.101		6
	MeOH	0.046	0.03		6
	CH_3CN	0.011	0.092		6
 **25**	hexane	n.d.	0.92		2
	toluene	0.066	-	Not Reported	2
	CCl_4	0.15	-	Not Reported	2
	CH_2Cl_2	0.20	0.3		2
	THF	0.15	0.7		2
	EtOH	0.30	0.42		2
	CH_3CN	0.084	0.45		2
 **26**	hexane	0.066	0.92		2
	toluene	0.038	0.86		2
	CCl_4	0.061	-	Not Reported	2
	CH_2Cl_2	0.068	0.89		2
	THF	0.066	0.79		2
	EtOH	0.18	0.81		2
	CH_3CN	0.092	0.57		2
 **27**	hexane	0.01	0.99		8
	toluene	0.045	0.92		4,8
	THF	0.21	0.36		4
	EtOH	0.53	0.038		8
	CH_3CN	0.22	0.009		4
 **28**	hexane	0.04	0.9		8
	toluene	0.10	0.81		9
	EtOH	0.59	0.02		8
	CH_3CN	0.84	0.01		9,8
 **29**	toluene	0.20	0.39		9
	CH_2Cl_2	0.24	0.1		9
	CH_3CN	0.11	0.01		9
 **30**	toluene	0.11	0.42		9
	CH_2Cl_2	0.13	0.2		9
	CH_3CN	0.05	0.04		9

 **31**	hexane	0.01	0.967		10
	toluene	0.086	0.83		4
	THF	0.20	0.94		4
	EtOH	0.34	0.651		10
	CH_3CN	0.34	0.59		4
 **32**	hexane	0.1	0.381		11
	toluene	0.31	0.067		12
 **33**	toluene	0.18	0.619		11
	THF	0.21	0.082		11
	CH_2Cl_2	0.42	0.037		11
	CH_3CN	0.11	0.017		11
 **34**	hexane	0.349	-	Not reported	13
	toluene	0.673	-	Not reported	13
	CH_2Cl_2	0.013	-	Not reported	13
 **35**	hexane	0.018	-	Not reported	13
	toluene	0.246	-	Not reported	13
 **36**	hexane	0.28	-	Not reliable	14
	toluene	0.08	-	Not reliable	14
 **37**	toluene	0.02	0.0004		14

 **38**	hexane	0.11	0.155		14
	toluene	0.42	-	Not reliable	14
 **39**	toluene	0.23	0.393		15
 **40**	hexane	0.05	0.354		15
	toluene	0.18	0.018		15
 **41**	hexane	0.01	0.303		14
	toluene	0.39	0.019		14
 **42**	hexane	0.02	0.16		14
	toluene	0.04	0.033		14
 **43**	toluene	0.01	0.0012		16
	THF	0.003	-	Not reported	16
 **44**	cyclohexane	0.04	-	Not reported	16
	toluene	0.014	0.001		16
	THF	0.006	-	Not reported	16

 45	toluene	0.001	0.005		16
	THF	0.008	-	Not reported	16
	CH_3CN	0.009	-	Not reported	16
 46	cyclohexane	0.30	-	Not reported	16
	toluene	0.05	0.0013		16
	THF	0.004	-	Not reported	16
 47	toluene	0.033	0.2		17
	CH_2Cl_2	0.58	0.04		17
	CH_3CN	0.024	-	Not reported	17
 48	toluene	0.023	0.83		17
	CH_2Cl_2	0.082	0.72		17
	CH_3CN	0.54	0.18		17
 49	toluene	0.083	0.38		17
	CH_2Cl_2	0.026	0.15		17
 50	toluene	0.09	0.82		18
 51	toluene	0.11	0.81		18
 52	toluene	0.19	0.31		18
 53	toluene	0.32	0.09		18
	toluene	0.02	0.59		17
	CH_2Cl_2	0.043	0.52		17

 **55**	hexane	0.03	0.462		15
	toluene	0.04	0.36		15
 **56**	Hexane	0.04	0.484		15
	toluene	0.12	0.299		15
 **57**	toluene	0.4	0.71		19
	CH_2Cl_2	0.5	0.56		19
 **58**	toluene	0.4	0.67		19
	CH_2Cl_2	0.5	0.6		19
 **59**	toluene	0.3	0.69		19
	CH_2Cl_2	0.5	0.62		19
 **60**	toluene	0.3	0.76		19

	CH_2Cl_2	0.5	0.63		19
 61	toluene	0.24	0.804		20
	$CHCl_3$	0.75	0.22		21
	THF	0.86	0.17		21
	CH_2Cl_2	0.64	-	Not reported	20
	acetone	0.5	0.84		21
	CH_3CN	0.25	0.01		21
 62	toluene	0.44	0.033		20
	CH_2Cl_2	0.20	-	Not reported	20
 63	toluene	0.144	0.628		20
	CH_2Cl_2	0.68	-	Not reported	20
	CH_3CN	0.112	0.001		20
 64	toluene	0.01	0.573		20
	CH_2Cl_2	0.09	-	Not reported	20
	CH_3CN	0.03	0.007		20
 65	Hexane	-	0.123	Not reported	22
	toluene	0.071	0.265		22
	CH_2Cl_2	0.12	0.205		22
	THF	-	0.187	Not reported	22
	CH_3CN	-	0.013	Not reported	22
 66	Hexane	0.065	0.096		22
	toluene	0.33	0.121		22
	CH_2Cl_2	0.12	0.027		22
	THF	0.061	0.021		22
	CH_3CN	-	0.003	Not reported	22
 67	Hexane	0.12	0.173		22
	toluene	0.37	0.166		22
	CH_2Cl_2	0.25	0.062		22
	THF	0.089	0.05		22
	CH_3CN	0.027	0.008		22
 68	Hexane	0.023	0.077		22
	toluene	0.011	0.025		22
	CH_2Cl_2	-	0.004	Not reported	22
	THF	-	0.003	Not reported	22
	CH_3CN	0.005	0.002		22
	Hexane	0.22	0.617		23
	toluene	0.289	0.513		23
	THF	0.307	0.291		23
	CH_2Cl_2	-	0.207	Not reported	23

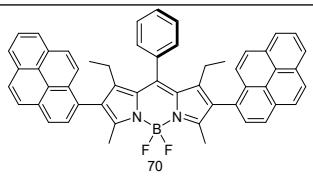
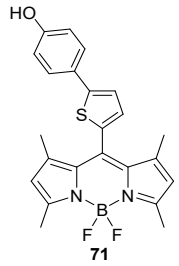
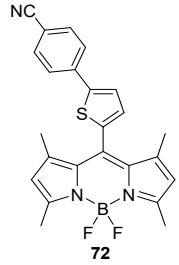
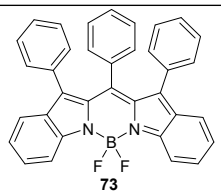
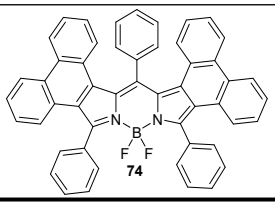
	EtOH	0.299	0.094		23
	CH_3CN	0.14	0.025		23
 70	Hexane	0.18	0.676		23
	toluene	0.203	0.634		23
	THF	0.185	0.445		23
	CH_2Cl_2	-	0.352	Not reported	23
	EtOH	0.19	0.191		23
	CH_3CN	0.09	0.078		23
 71	Hexane	0.007	0.028		24
	toluene	0.046	0.039		24
	CH_2Cl_2	0.441	0.019		24
	CH_3CN	0.175	0.0001		24
 72	Hexane	-	0.023	Not reported	24
	toluene	0.026	0.028		24
	CH_2Cl_2	0.012	0.031		24
	CH_3CN	0.026	0.020		24
 73	Hexane	0.026	0.001		25
	toluene	0.058	0.0009		25
	CH_2Cl_2	0.028	0.0002		25
	CH_3CN	0.023	-	Not Reported	25
 74	Hexane	-	0.696	Not Reported	25
	toluene	-	0.74	Not Reported	25
	CH_2Cl_2	-	0.723	Not Reported	25
	CH_3CN	0.055	0.644	Not Reported	25

Table S2. Statistical parameters of the best performing QSPR models for BODIPYs in toluene obtained with MLR, SVR and RFR method.

Parameter	MLR	SVR	RFR
R^2_{train}	0.863	0.845	0.839
$RMSE_{train}$	0.280	0.281	0.304
$R^2_{adjusted}$	0.829	0.806	0.798
$\mathbf{q^2}$	0.771	0.793	0.556

Table S3. Values of the descriptor used by Model 2 (SVR, toluene).

Compound	VE1_RG	$\frac{Lg\ \Phi_{Fl}}{Lg\ \Phi_{\Delta}}$
BDP-1	3.739	0.108
BDP-2 [test]	4.153	0.228
BDP-21	3.961	0.111
BDP-26	3.915	0.153
BDP-27	3.816	0.374
BDP-28	3.700	1.618
BDP-29	4.205	0.121
BDP-30	3.635	1.460
BDP-31 [test]	3.925	0.049
BDP-32	5.25	2.218
BDP-33	3.867	0.046
BDP-37	3.008	1.060
BDP-39 [test]	2.943	0.083
BDP-40	3.016	0.308
BDP-42 [test]	3.110	0.240
BDP-43	3.771	0.082
BDP-44	3.945	0.705
BDP-45	3.774	0.437
BDP-46	5.043	0.076
BDP-47 [test]	3.039	0.135
BDP-48	2.964	0.472

BDP-49	3.845	0.095
BDP-50	3.790	0.393
BDP-51	3.736	0.767
BDP-52	2.543	0.389
BDP-53 [test]	3.521	0.569
BDP-54	3.708	1.054
BDP-55	3.116	0.280
BDP-56	3.231	0.585
BDP-57 [test]	4.230	0.317
BDP-58	4.273	2.113
BDP-59	3.700	0.027
BDP-60	5.169	0.635
BDP-61	4.056	0.980
BDP-63	2.512	2.000
BDP-64	3.952	2.343
BDP-65	3.435	0.286
BDP-66 [test]	3.462	0.092
BDP-67	3.520	0.538
BDP-68	3.439	0.502
BDP-69	3.302	2.308
BDP-70 [test]	2.766	1.905
BDP-71	3.733	0.818
BDP-72	3.769	1.806
BDP-73	5.754	2.463

Table S4. Most significant descriptors of Model 3 (toluene, RFR).

Compound	ATSC4e	TDB01e	$\frac{Lg\ \Phi_{Fl}}{Lg\ \Phi_{\Delta}}$
BDP-1	0.4766	1.352	0.108
BDP-2 [test]	0.2636	1.303	0.228
BDP-21	0.2131	1.296	0.111
BDP-26	0.3993	1.32	0.153
BDP-27	0.2332	1.296	0.374
BDP-28	0.2635	1.292	1.618
BDP-29	0.2519	1.292	0.121
BDP-30	0.2564	1.292	1.460
BDP-31 [test]	0.2284	1.301	0.049
BDP-32	0.2012	1.32	2.218
BDP-33	0.2421	1.301	0.046
BDP-37	0.3406	1.319	1.060
BDP-39 [test]	0.4071	1.298	0.083
BDP-40	0.3874	1.308	0.308
BDP-42 [test]	0.3682	1.303	0.240
BDP-43	0.3268	1.332	0.082
BDP-44	0.2794	1.31	0.705
BDP-45	0.7844	1.349	0.437
BDP-46	0.4533	1.315	0.076
BDP-47 [test]	0.2759	1.306	0.135
BDP-48	0.3333	1.29	0.472
BDP-49	0.3052	1.296	0.095
BDP-50	0.3033	1.296	0.393
BDP-51	0.4015	1.297	0.767
BDP-52	0.4386	1.29	0.389
BDP-53 [test]	0.4141	1.321	0.569
BDP-54	0.3203	1.296	1.054
BDP-55	0.4717	1.285	0.280
BDP-56	0.612	1.292	0.585

BDP-57 test	0.3469	1.298	0.317
BDP-58	0.5153	1.289	2.113
BDP-59	0.5681	1.281	0.027
BDP-60	0.4435	1.282	0.635
BDP-61	0.5342	1.311	0.980
BDP-63	0.5007	1.312	2.000
BDP-64	0.4591	1.308	2.343
BDP-65	0.1555	1.308	0.286
BDP-66 test	0.1463	1.309	0.092
BDP-67	0.1769	1.302	0.538
BDP-68	0.2564	1.301	0.502
BDP-69	0.3185	1.286	2.308
BDP-70 test	0.314	1.292	1.905
BDP-71	0.336	1.328	0.818
BDP-72	0.2738	1.322	1.806
BDP-73	0.1584	1.301	2.463

Table S5. Statistical parameters of the models for BODIPYs in acetonitrile.

Parameter	MLR	SVR	RFR
R^2_{train}	0.753	0.732	0.784
$RMSE_{train}$	0.466	0.482	0.436
$R^2_{adjusted}$	0.677	0.651	0.718
q^2	0.563	0.483	0.528

Table S6. Most significant molecular descriptors of Model 4 (acetonitrile, MLR).

Compound	VE1sign_G/D	VE2sign_G/D	F01[C-N]	$\frac{Lg\ \Phi_{Fl}}{Lg\ \Phi_{\Delta}}$
BDP-2	0.005	0.00015	4	0.079
BDP-3	0.216	0.00501	4	0.160
BDP-4	0.173	0.00385	5	1.468

BDP-5 [test]	0.221	0.00490	5	0.730
BDP-6	0.229	0.00510	5	0.771
BDP-7	0.185	0.00395	6	1.074
BDP-8 [test]	0.214	0.00510	6	0.877
BDP-9	0.200	0.00476	6	0.619
BDP-10	0.225	0.00480	4	0.328
BDP-11	0.227	0.00446	4	0.359
BDP-12	0.272	0.00533	4	2.215
BDP-13 [test]	0.281	0.00511	4	2.025
BDP-14	0.282	0.00514	4	0.510
BDP-15	0.315	0.00618	7	1.905
BDP-16 [test]	0.075	0.00167	7	1.633
BDP-17	0.351	0.00604	7	2.129
BDP-18	0.109	0.00209	7	2.169
BDP-19	0.333	0.00595	7	2.775
BDP-20 [test]	0.101	0.00201	7	1.147
BDP-21	0.181	0.00370	4	0.065
BDP-23	0.256	0.00511	4	0.850
BDP-24	0.012	0.00026	4	0.529
BDP-25 [test]	0.188	0.00408	4	0.322
BDP-26	0.036	0.00070	4	0.236
BDP-27	0.144	0.00262	4	3.111
BDP-29	0.185	0.00285	4	2.086
BDP-30	0.226	0.00348	4	1.074
BDP-31	0.188	0.00331	4	0.489
BDP-33	0.208	0.00331	4	1.846
BDP-48 [test]	0.335	0.00516	7	2.783
BDP-61	0.241	0.00360	8	3.322
BDP-63	0.135	0.00178	9	3.155
BDP-64	0.000	0.00000	8	1.415
BDP-67	0.055	0.00073	4	1.337
BDP-68	0.010	0.00013	4	1.173

BDP-69	0.199	0.00214	4	1.876
BDP-70	0.136	0.00140	4	1.059
BDP-72	0.240	0.00471	5	1.072
BDP-74 [test]	0.042	0.00048	4	0.152

Table S7. Values of the most significant descriptors used by Model 5.

Compound	F06[N-B]	B05[O-O]	$\frac{Lg\ \Phi_{Fl}}{Lg\ \Phi_{\Delta}}$
BDP-2	0	0	0.079
BDP-3	0	0	0.160
BDP-4	1	0	1.468
BDP-5 [test]	0	0	0.730
BDP-6	0	0	0.771
BDP-7	1	0	1.074
BDP-8 [test]	0	0	0.877
BDP-9	0	0	0.619
BDP-10	0	0	0.328
BDP-11	0	0	0.359
BDP-12	0	1	2.215
BDP-13 [test]	0	1	2.025
BDP-14	0	0	0.510
BDP-15	0	0	1.905
BDP-16 [test]	0	0	1.633
BDP-17	0	0	2.129
BDP-18	0	0	2.169
BDP-19	0	0	2.775
BDP-20 [test]	0	0	1.147
BDP-21	0	0	0.065
BDP-23	0	0	0.850
BDP-24	0	0	0.529
BDP-25 [test]	0	0	0.322
BDP-26	0	0	0.236
BDP-27	0	0	3.111
BDP-29	0	0	2.086
BDP-30	0	0	1.074
BDP-31	0	0	0.489
BDP-33	0	0	1.846

BDP-48 [test]	0	0	2.783
BDP-61	3	0	3.322
BDP-63	3	0	3.155
BDP-64	0	0	1.415
BDP-67	0	0	1.337
BDP-68	0	0	1.173
BDP-69	0	0	1.876
BDP-70	0	0	1.059
BDP-72	0	0	1.072
BDP-74 [test]	0	0	0.152

Table S8. Most significant descriptors of Model 6.

Compound	H2u	P_VSA_LogP_5	$\frac{Lg\ \Phi_{Fl}}{Lg\ \Phi_{\Delta}}$
BDP-2	2.052	6.371	0.079
BDP-3	2.763	12.740	0.160
BDP-4	3.021	15.500	1.468
BDP-5 [test]	2.998	15.500	0.730
BDP-6	2.736	15.500	0.771
BDP-7	2.994	18.260	1.074
BDP-8 [test]	2.674	21.360	0.877
BDP-9	2.732	30.840	0.619
BDP-10	3.480	28.990	0.328
BDP-11	3.712	45.230	0.359
BDP-12	3.734	45.230	2.215
BDP-13 [test]	3.829	61.470	2.025
BDP-14	3.788	61.470	0.510
BDP-15	3.186	18.620	1.905
BDP-16 [test]	3.020	12.250	1.633
BDP-17	3.343	18.620	2.129
BDP-18	3.063	12.250	2.169
BDP-19	3.535	18.620	2.775
BDP-20 [test]	3.435	12.250	1.147
BDP-21	3.506	17.890	0.065
BDP-23	3.153	17.890	0.850
BDP-24	3.056	11.520	0.529
BDP-25 [test]	3.426	14.700	0.322
BDP-26	3.524	21.070	0.236
BDP-27	3.862	23.030	3.111
BDP-29	3.855	29.410	2.086
BDP-30	3.366	29.410	1.074
BDP-31	3.810	28.180	0.489
BDP-33	3.872	33.330	1.846

BDP-48 [test]	4.057	19.790	2.783
BDP-61	3.990	22.300	3.322
BDP-63	4.118	28.240	3.155
BDP-64	3.173	25.490	1.415
BDP-67	4.173	33.330	1.337
BDP-68	4.026	39.700	1.173
BDP-69	4.055	46.070	1.876
BDP-70	3.983	56.360	1.059
BDP-72	3.220	46.990	1.072
BDP-74 [test]	3.818	44.840	0.152

ALE toluene MLR

We applied the accumulated local effects (ALE) methodology to analyze the performance of our ML models in toluene. Figure 1 presents the results for the MLR model (Model 1). The ALE graphs illustrate how local variations in the descriptor values influence the dependent variable yyy. The slope coefficients of the graphs match those from the MLR method, confirming that the linear model accurately captures linear dependencies. The interpretation of the ALE plot is that the ALE value for a given feature represents the deviation from the average effect of that feature. In other words, the ALE value reflects the relative impact of that feature on the prediction at a specific feature value.

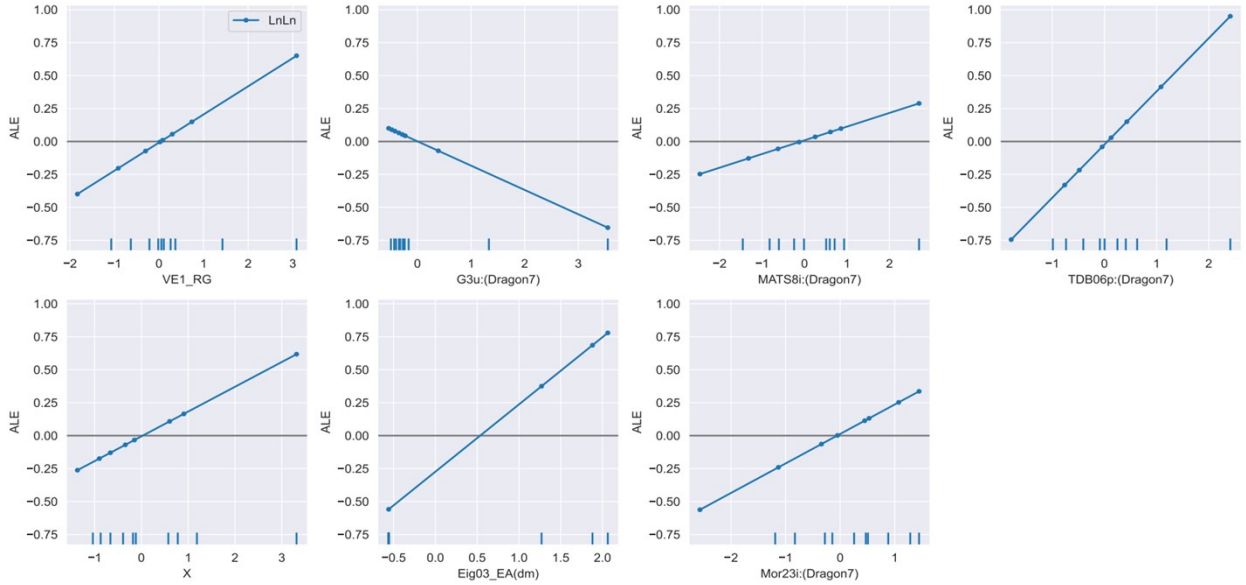


Figure S1. ALE plots for the MLR method for molecules studied in toluene.

ALE graphs for Model 1 are linear and the slope coefficient of each graph coincides with the coefficient obtained in the MLR method. This indicates that Model 1 works correctly and accurately describes all linear dependencies.

RFR

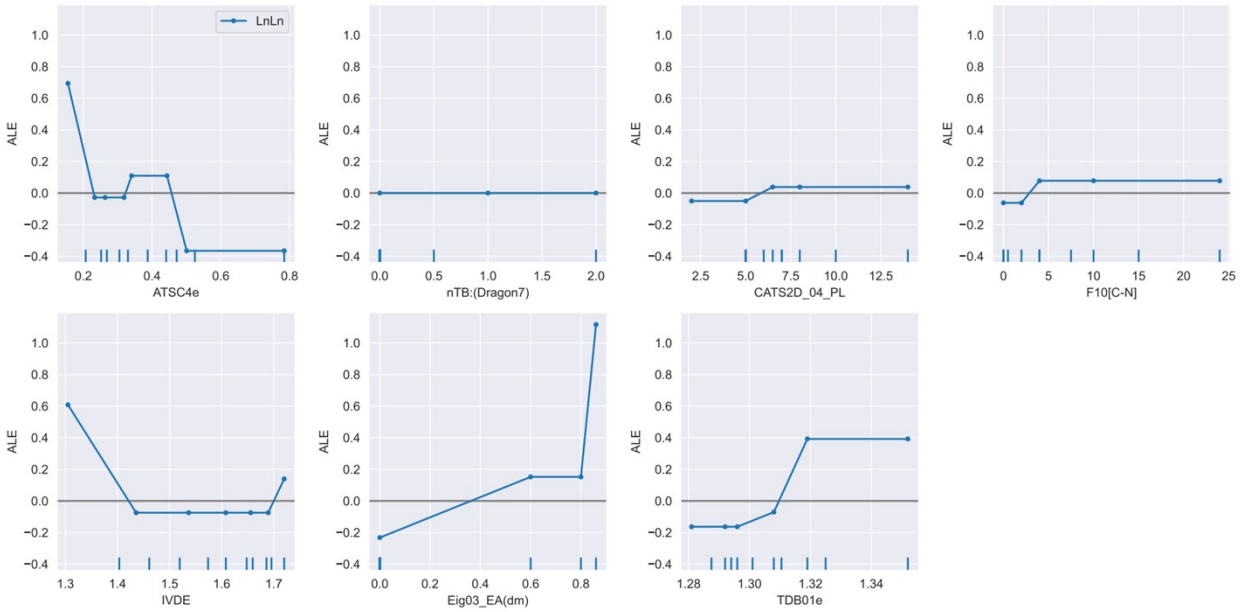


Figure S2. ALE plots for the RFR model (Model 2) in toluene.

The results of the ALE methodology for the RFR model (Model 2) built for toluene are presented in Figure S2. The graphs indicate that the descriptors ATSC4e, Eig03_EA(dm), and TDB01e have the strongest impact on the model. As the value of the ATSC4e descriptor increases, the dependent variable decreases. The Eig03_EA(dm) descriptor causes a monotonic increase in the dependent variable yyy, with a sharp jump occurring between values 0.6 and 0.8. Similarly, the TDB01e descriptor also has a monotonic effect on yyy up to a value of 1.32, after which it no longer influences yyy. The remaining descriptors have minimal impact, as indicated by their relative contributions.

SVR

The results of the ALE methodology for the SVR model (Model 3) built for toluene are presented in Figure S3. This model uses a linear kernel, and all ALE dependencies are linear. The slope coefficients of the graphs correspond to the relative significance of the descriptors. Only the G3u descriptor is inversely proportional to the dependent variable yyy, while the other descriptors are directly proportional to y.

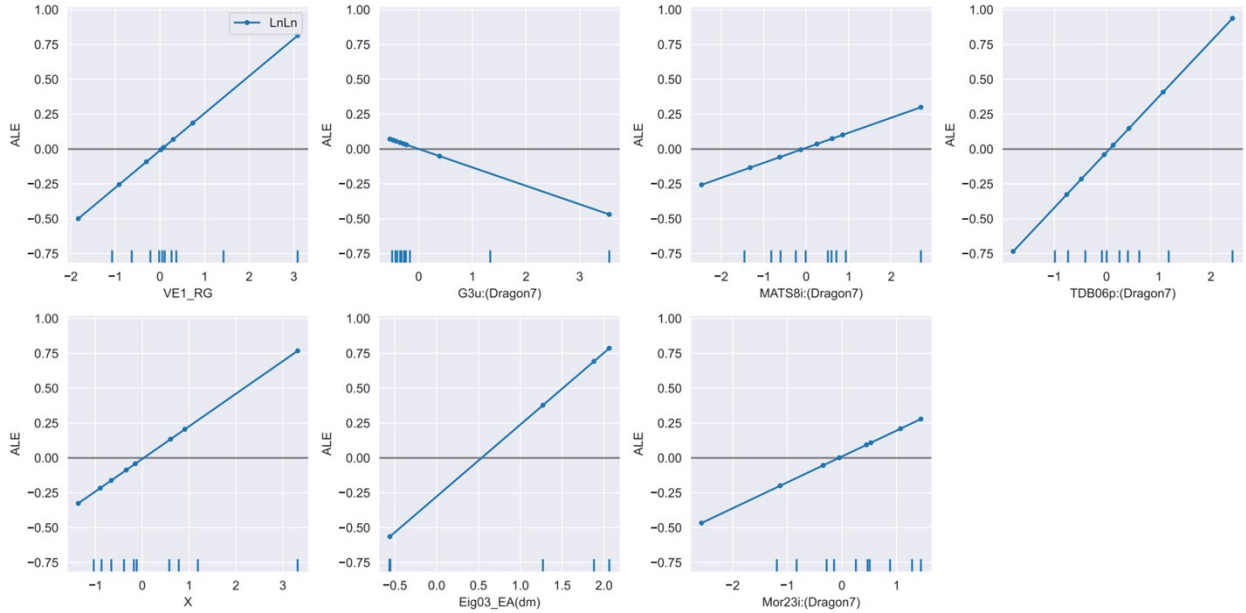


Figure S3. ALE plots for the SVR Model 6 in toluene.

Acetonitrile MLR

We also applied the ALE methodology to evaluate the significance of descriptors in our acetonitrile ML models. Similar to the other linear regression model (Model 1), in the case of Model 4, the ALE graphs are linear, and the slope coefficients of each graph align with the coefficients obtained from the MLR method. The results of the ALE methodology for the MLR model in acetonitrile (Model 4) are presented in Figure S4.

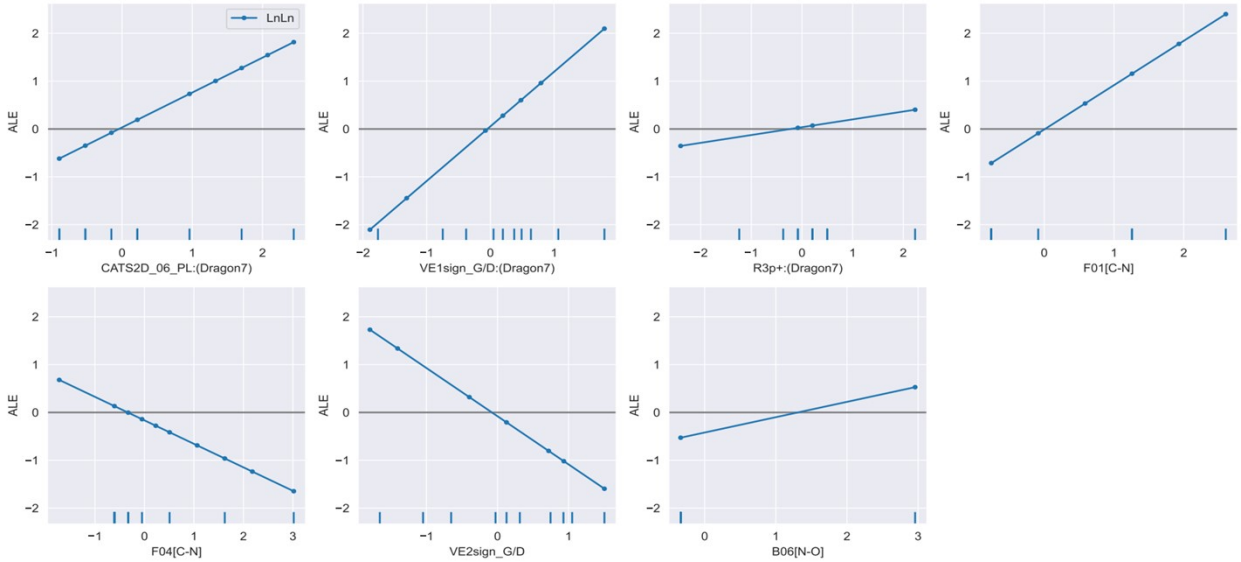


Figure S4. ALE plots for the MLR method for molecules studied in acetonitrile.

The increment of descriptors F04[C-N] and VE2sign_G/D leads to a decrease of the dependent parameter y, whereas the increment of other descriptors leads to an increase of the dependent variable y.

RFR

The results of the ALE method for the RFR model in acetonitrile are presented in Figure S5. The graphs confirm that the descriptors H2u, P_VSA_LogP_5, and F01[C-N] are the most influential. Analyzing these ALE plots reveals how different attributes nonlinearly affect the predictions of the random forest model. The increase in the F01[C-N] descriptor up to a value of 7 leads to a rise in the dependent variable yyy, after which it ceases to have an effect. The increase in the most significant descriptor, H2u, up to a value of 3.2 also causes a rise in the dependent variable, followed by a sharp decline at 3.5 and then an increase up to 4. The P_VSA_LogP_5 descriptor, on the other hand, monotonically decreases the dependent variable.

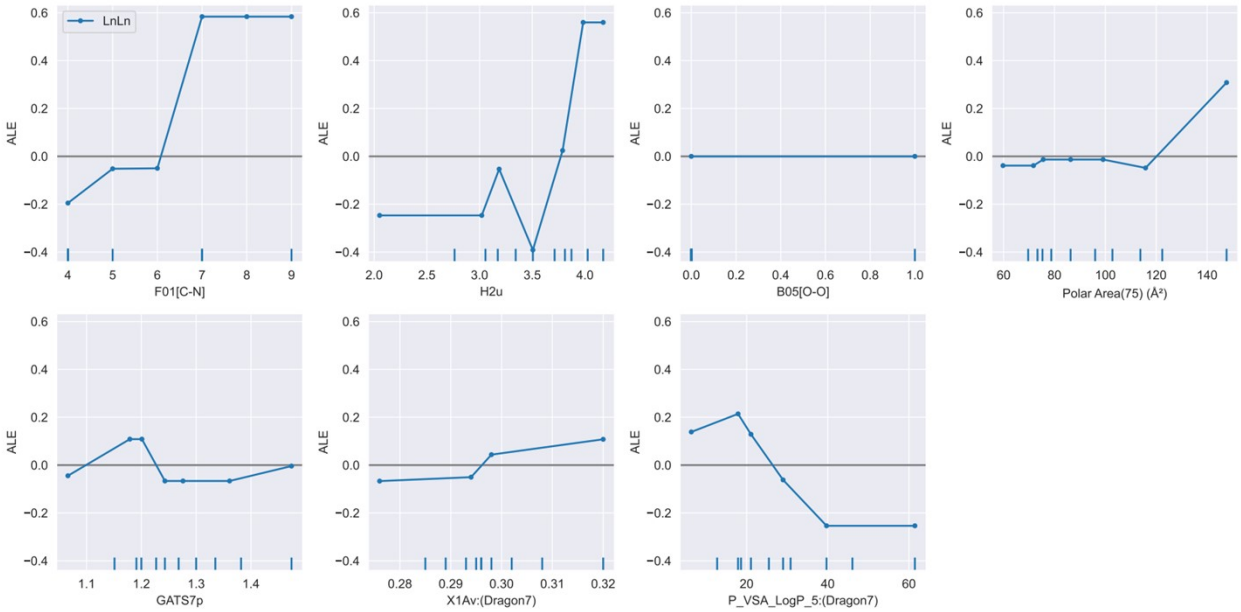


Figure S5. ALE plots for the RFR method for molecules dissolved in acetonitrile.

SVR

The results of the ALE method for the SVR model developed for acetonitrile are presented in Figure S6. The graphs show that all dependencies are linear. Since the model uses a linear kernel, Model 6 does not account for non-linear dependencies. As the values of the descriptors LDI, F04[C-B], and LLS_02 increase, the dependent variable yyy decreases. The remaining descriptors cause yyy to increase as their values increase.

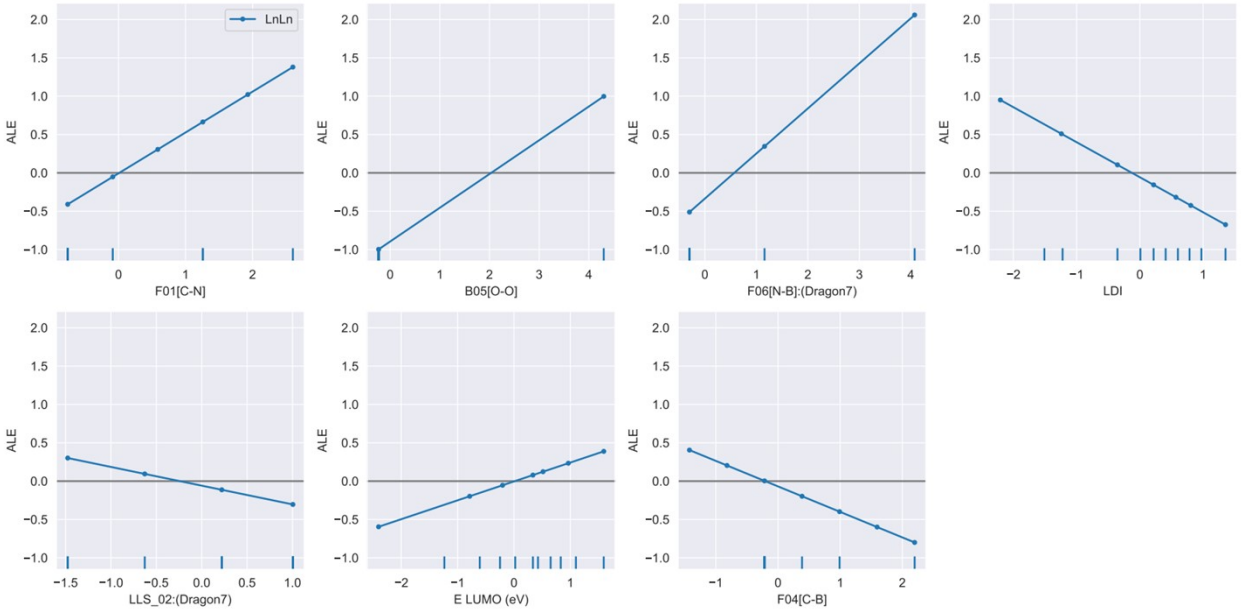


Figure S6. ALE plots for the SVR method for molecules studied in acetonitrile.

SHAP toluene MLR

We used the SHAP methodology to evaluate the contribution of descriptors to our QSPR models. SHAP (Shapley additive explanations) is a theoretical approach for interpreting the results of any ML model. SHAP combines optimal credit allocation with local explanations by using traditional Shapley values from game theory and their related applications. Specifically, we applied SHAP to analyze the performance of our QSPR models for molecules dissolved in toluene. Figure S7 presents the SHAP results for the MLR Model 1 in toluene. The distribution of SHAP values and their relationship with actual feature values are depicted using the color bar on the right. According to SHAP, the Eig03_EA(dm) descriptor is the main contributor with the highest impact on the model output.

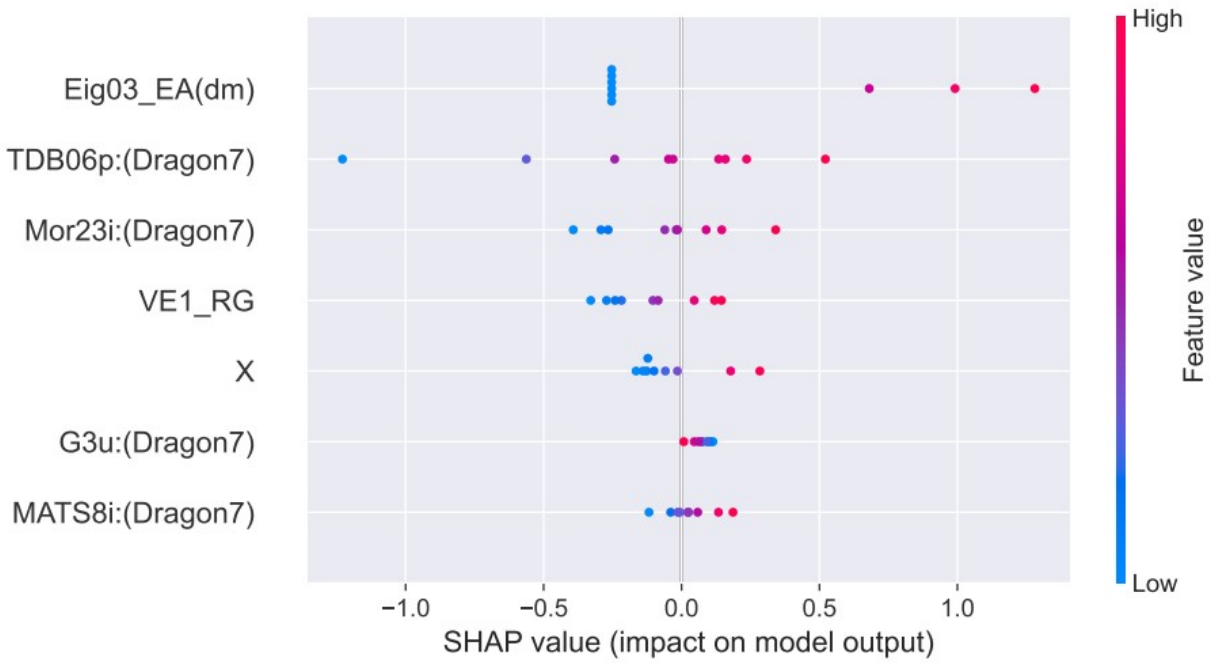


Figure S7. Beeswarm plot of SHAP values generated using the MLR Model 1 developed for toluene.

RFR

The results of the SHAP methodology applied to the RFR Model 2 for toluene are presented in Figure S8. The highest contribution is observed for the Eig03_EA(dm) descriptor, which is similar to Model 1 (toluene, MLR).

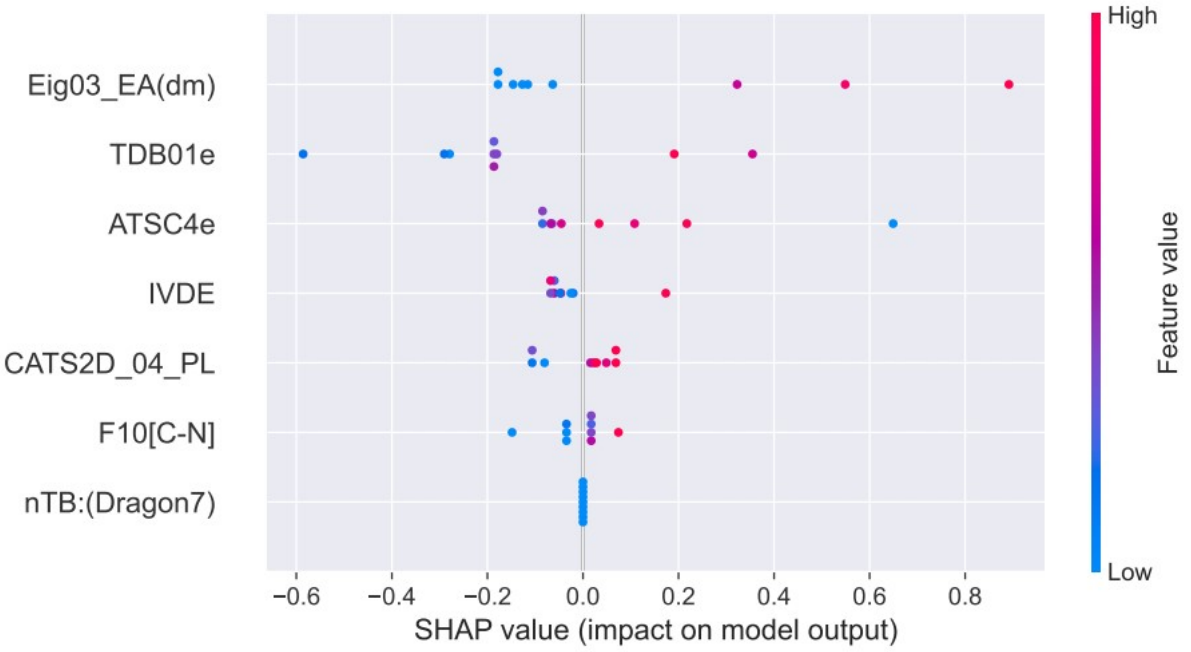


Figure S8. Summary plot of SHAP values generated from the test dataset using the RFR model developed for toluene-extracted molecules.

SVR

Application of SHAP methodology to SVR Model 3 in toluene is presented on Figure S9. The predominant descriptor is Eig03_EA(dm).

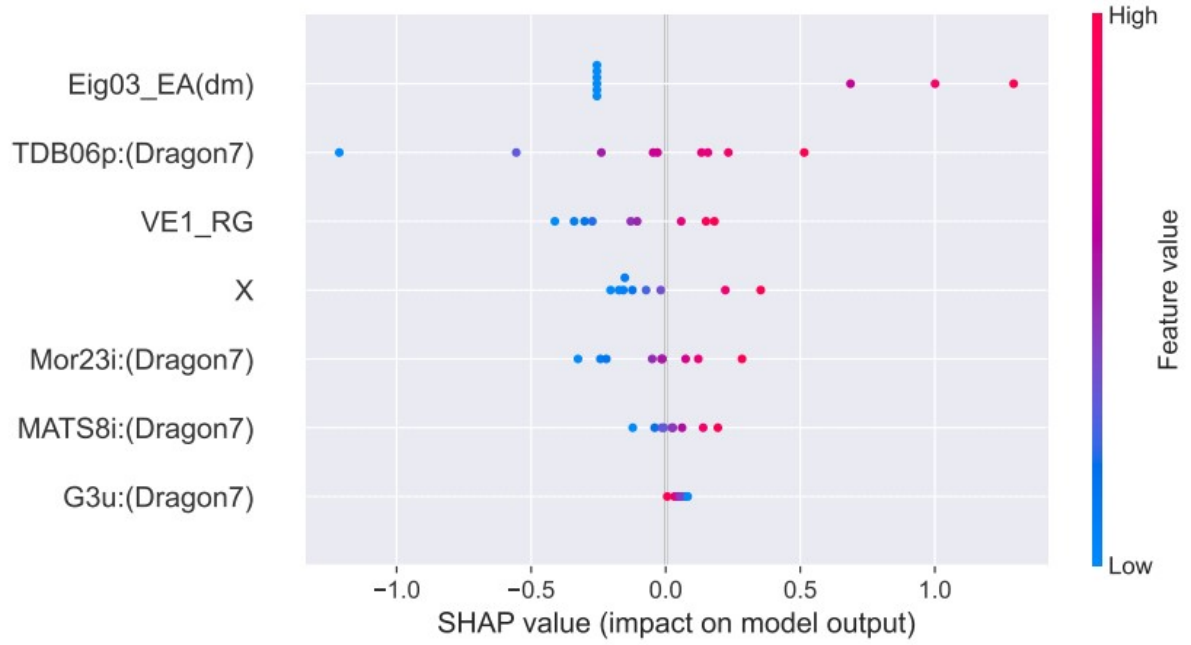


Figure S9. Beeswarm plot of SHAP values generated using the SVR method for toluene.

Acetonitrile MLR

The results of the SHAP method for the MLR Model 4 in acetonitrile are presented in Figure S10. The main contributors of Model 4 are VE1sign_G/D (directly proportional to the

dependent variable y), VE2sing_G/D (inversely correlated with y) and F01[C-N] (directly proportional to y).

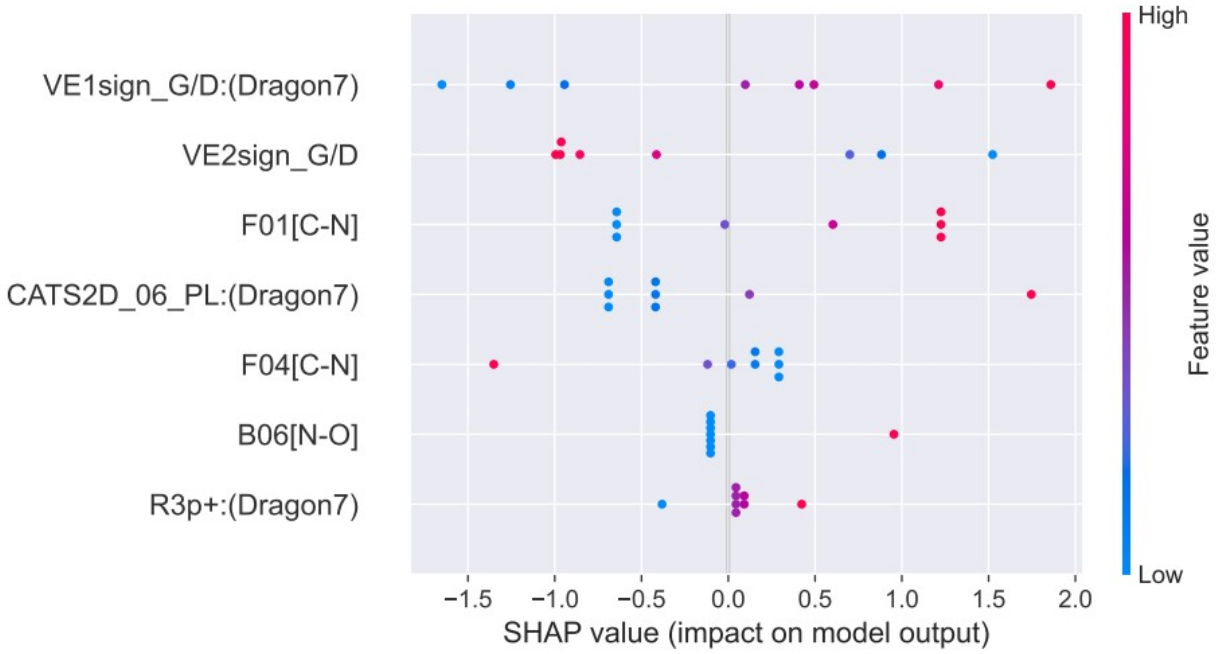


Figure S10. Summary plot of SHAP values generated from the test dataset using the MLR model developed for acetonitrile-extracted molecules.

RFR

The results of the SHAP method for the RFR Model 5 in acetonitrile are presented in Figure S11. According to SHAP, the major contributions are made by H2u and F01[C-N] descriptors, both of which are directly proportional to the dependent variable y.

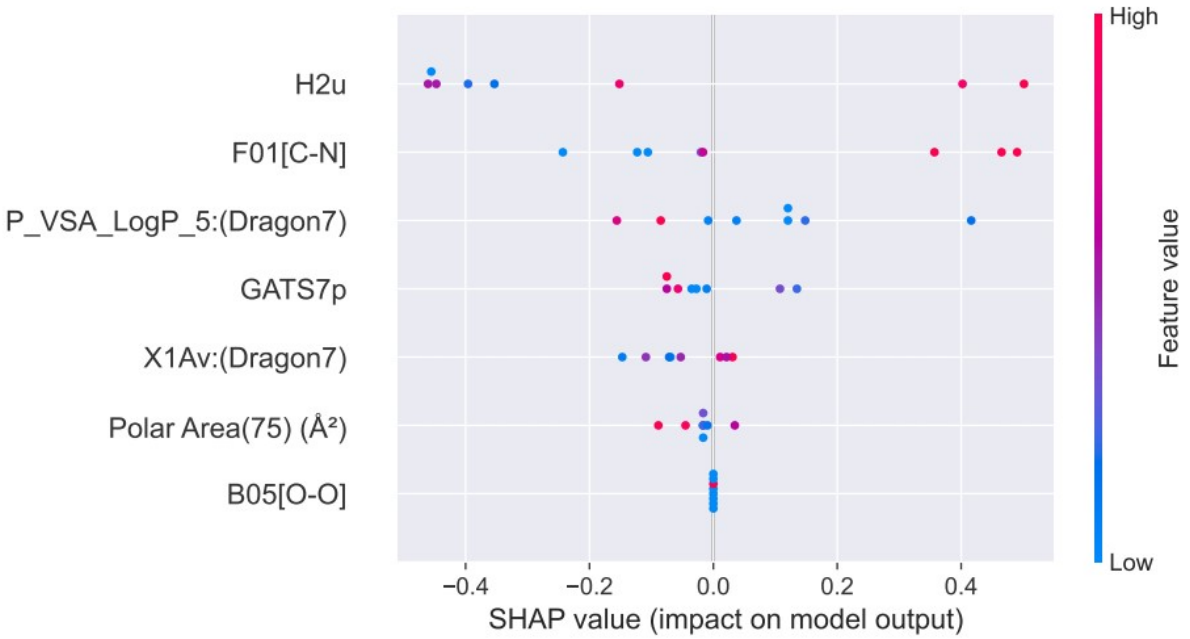


Figure S11. Summary plot of SHAP values generated using the RFR model in acetonitrile.

SVR

The results of the SHAP method for the SVR Model 6 developed for acetonitrile are presented in Figure S12. According to SHAP, the major contributor is the LDI descriptor, which is inversely correlated with the dependent variable y.

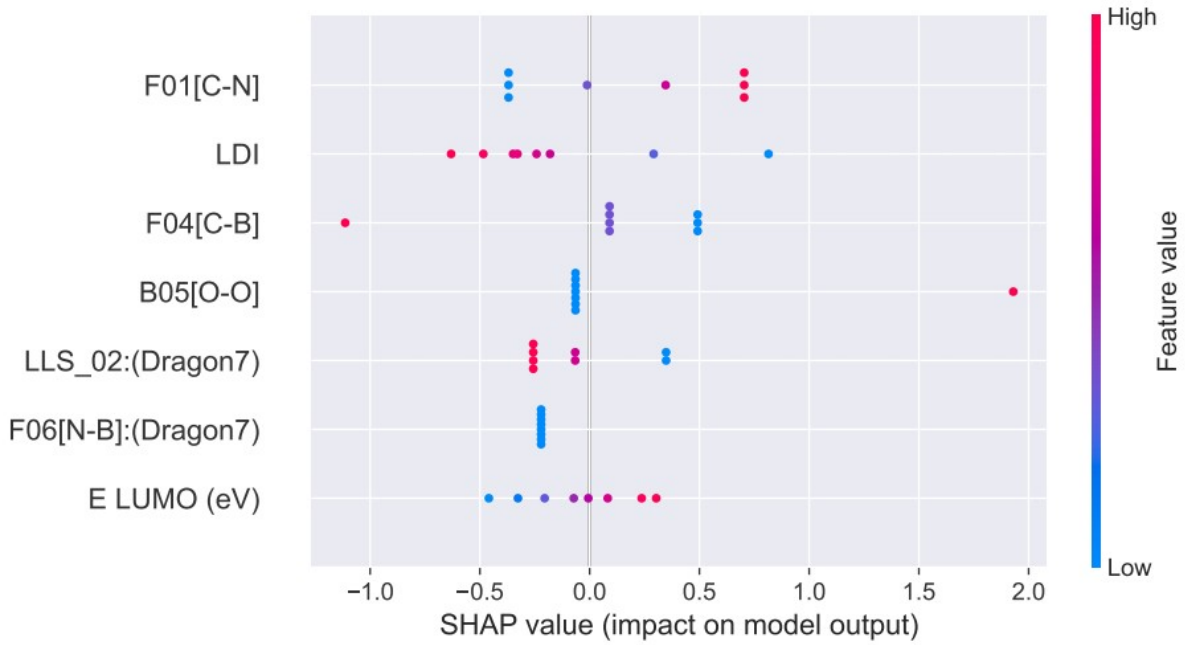


Figure S12. Beeswarm visualization of SHAP generated using the SVR Model 6 developed for acetonitrile.

References

[1] X.-F. Zhang, J. Zhu, *J. Luminescence*, 2019, **205**, 148-157.
[2] X.-F. Zhang, N. Feng, *Spectrochim. Acta A*, 2018, **189**, 13-21.
[3] W. Hu, M. Liu, X.-F. Zhang, Y. Wang, Y. Wang, H. Lan, H. Zhao, *J. Phys. Chem. C*, 2019, **123**, 15944-15944.

[4] X.-F. Zhang, N. Feng, *Chem. Asian. J.*, 2017, **12**, 2447-2456.
[5] W. Hu, Y. Lin, X.-F. Zhang, M. Feng, S. Zhao, J. Zhang, *Dyes Pigm.*, 2019, **164**, 139-147.
[6] W. Hu, X.-F. Zhang, X. Lu, S. Lan, D. Tian, T. Li, L. Wang, S. Zhao, M. Feng, J. Zhang, *J. Lumin.*, 2018, **194**, 185-192.
[7] W. Hu, X.-F. Zhang, X. Lu, S. Lan, D. Tian, T. Li, L. Wang, S. Zhao, M. Feng, J. Zhang, *Dyes Pigm.*, 2018, **149**, 306-314.
[8] M. A. Filatov, S. Karuthedath, P. M. Polestshuk, S. Callaghan, K. J. Flanagan, M. Telitchko, T. Wiesner, F. Laquai, M. O. Senge, *Phys. Chem. Chem. Phys.*, 2018, **20**, 8016-8031.
[9] Z. Wang, J. Zhao, *Org. Lett.*, 2017, **19**, 4492-4495.
[10] M. A. Filatov, S. Karuthedath, P. M. Polestshuk, S. Callaghan, K. Flanagan, T. Wiesner, F. Laquai, M. O. Senge, *ChemPhotoChem*, 2018, **2**, 606-615.
[11] Z. Wang, M. Ivanov, Y. Gao, L. Bussotti, P. Foggi, H. Zhang, N. Russo, B. Dick, J. Zhao, M. Di Donato, G. Mazzone, L. Luo, M. Fedin, *Chem. Eur. J.*, 2020, **26**, 1091-1102.
[12] Z. Wang, J. Zhao, M. Di Donato, G. Mazzone, *Chem. Commun.*, 2019, **55**, 1510-1513.
[13] K. Chen, W. Yang, Z. Wang, A. Iagatti, L. Bussotti, P. Foggi, W. Ji, J. Zhao, M. Di Donato, *J. Phys. Chem. A*, 2017, **121**, 7550-7564.
[14] Y. Dong, A. A. Sukhanov, J. Zhao, A. Elmali, X. Li, B. Dick, A. Karatay, V.K. Voronkova, *J. Phys. Chem. C*, 2019, **123**, 22793-22811.
[15] Y. Dong, A. Elmali, J. Zhao, B. Dick, A. Karatay, *ChemPhysChem*, 2020, **21**, 1388-1401.
[16] K. Chen, M. Taddei, L. Bussotti, P. Foggi, J. Zhao, M. Di Donato, *ChemPhotoChem*, 2020, **4**, 487-501.
[17] Y. Hou, I. Kurganskii, A. Elmali, H. Zhang, Y. Gao, L. Lv, J. Zhao, A. Karatay, L. Luo, M. Fedin, *J. Chem. Phys.*, 2020, **152**, 114701.
[18] V.-N. Nguyen, Y. Yim, S. Kim, B. Ryu, K. M. K. Swamy, G. Kim, N. Kwon, C.-Y. Kim, S. Park, J. Yoon, *Angew. Chem. Int. Ed.*, 2020, **59**, 8957-8962.
[19] B. Ventura, G. Marconi, M. Bröring, R. Krüger, L. Flamigni, *New J. Chem.*, 2009, **33**, 428-438.
[20] L. Ya, J. Zhao, A. Iagatti, L. Bussotti, P. Foggi, E. Castellucci, M. Di Donato, K. Han, *J. Phys. Chem. C*, 2018, **122**, 2502-2511.
[21] N. Epelde-Elezcano, E. Palao, H. Manzano, A. Prieto-CastaCeda, A.R. Agarrabeitia, A. Tabero, A. Villanueva, S. de la Moya, C. Ljpez-Arbeloa, V. Martinez-Martinez, M.J. Ortiz, *Chem. Eur. J.*, 2017, **23**, 4837-4848.
[22] Y. Hu, Y. Hou, Z. Wang, Y. Li, J. Zhao, *J. Chem. Phys.*, 2020, **153**, 224304.
[23] T. Mikulchyk, S. Karuthedath, C. S. P. De Castro, A. A. Buglak, A. Sheehan, A. Wieder, F. Laquai, I. Naydenova, M. A. Filatov., *J. Matter. Chem. C*, 2022, **10**, 11588-11597.
[24] Y. Dong, M. Taddei, S. Doria, L. Bussotti, J. Zhao, G. Mazzone, M. Di Donato, *J. Phys. Chem. B*, 2021, **125**, 4779-4793.
[25] Y. Yan, A. A. Sukhanov, M. H. E. Bousquet, Q. Guan, J. Zhao, V. K. Voronkova, D. Escudero, A. Barbon, Y. Xing, G. G. Gurzadyan, D. Jacquemin, *J. Phys. Chem. B*, 2021, **125**, 6280-6295.