Petra Bombicz[a*], Nóra V. May[a], Dániel Fegyverneki[b],
Avirmed Saranchimeg[a], Laura Bereczki[a*]

Methods for easy recognition of isostructurality -
LabJack-like crystal structures of halogenated 2-phenylbenzimidazoles

[a]*Research Laboratory of Chemical Crystallography, Research Centre for Natural Sciences*
[b]*Institute of Organic Chemistry, Research Centre for Natural Sciences,*
1117 Budapest, Magyar Tudósok körútja 2., Hungary
e-mail: bombicz.petra@ttk.hu and nagyne.bereczki.laura@ttk.hu

Supplementary information

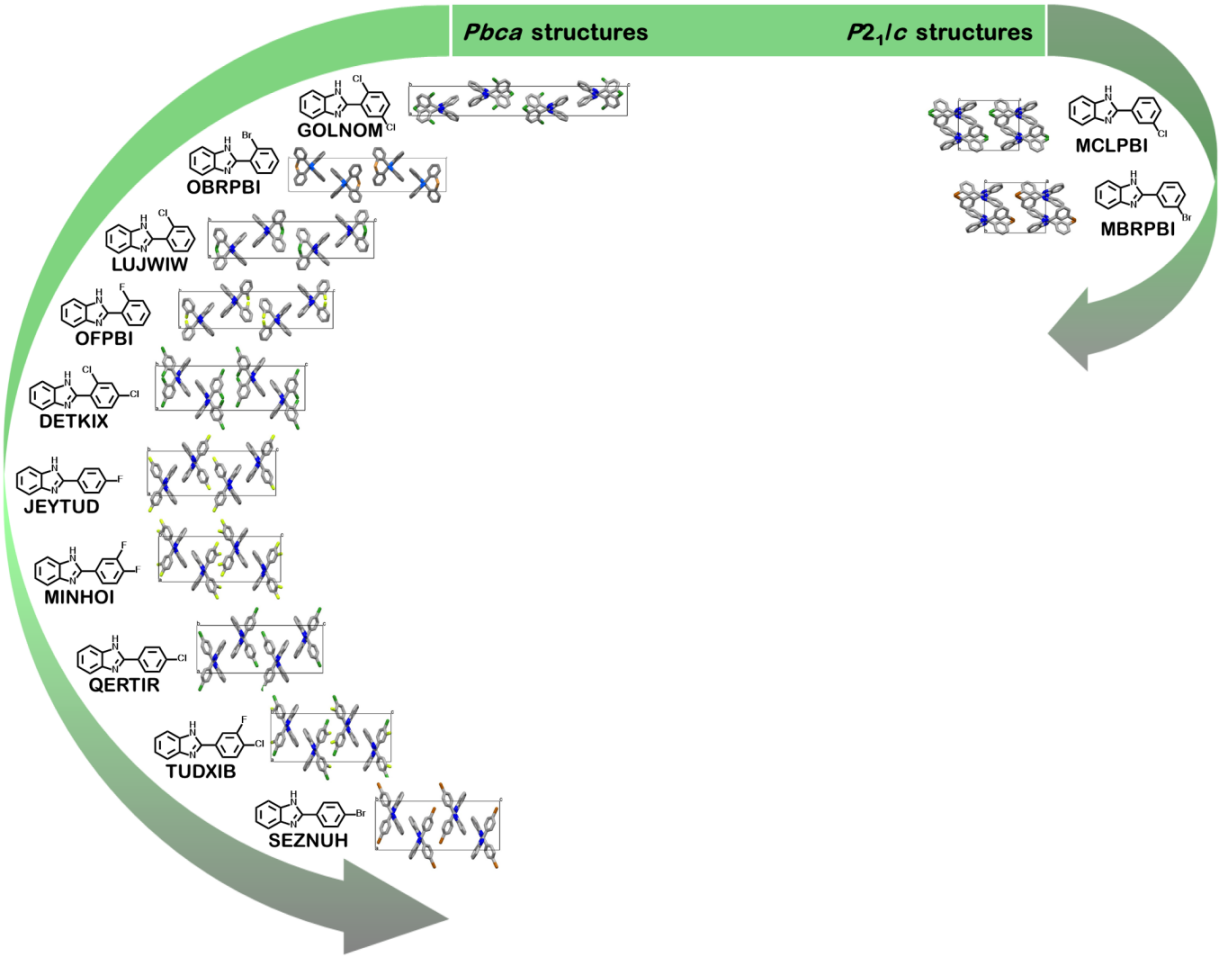

Figure S1 The fine-tuned series of the studied halogenated 2-phenylbenzimidazoles listed by the gradual altering unit cells within the isostructural groups.

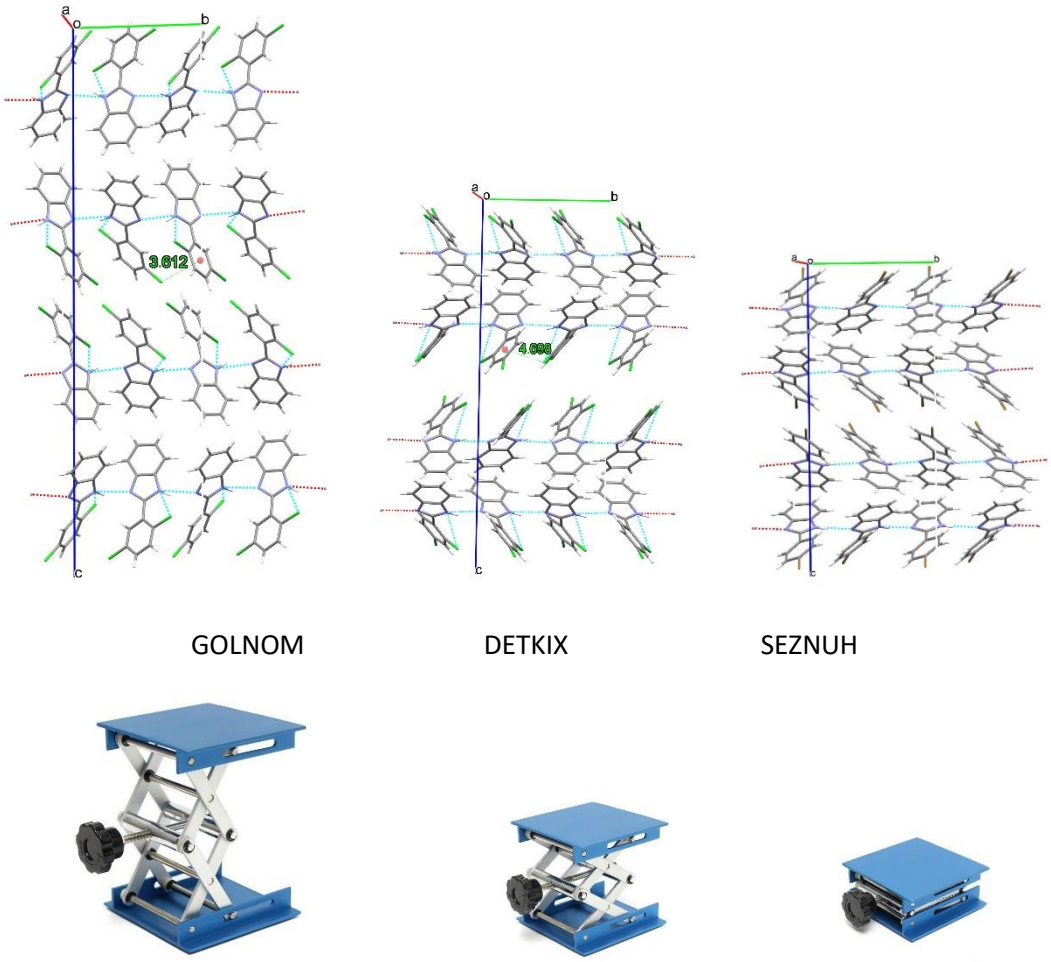

Figure S2 The lab Jack in pulled (GOLNOM), medium (DETKIX) and pushed (SEZNUH) position as a result of the placement of the halogen substituents. Determining supramolecular interactions are highlighted. Crystallographic c axes are approximately proportional to the real values. DETRIX is the "swich" structure between the two isostructural subgroups.

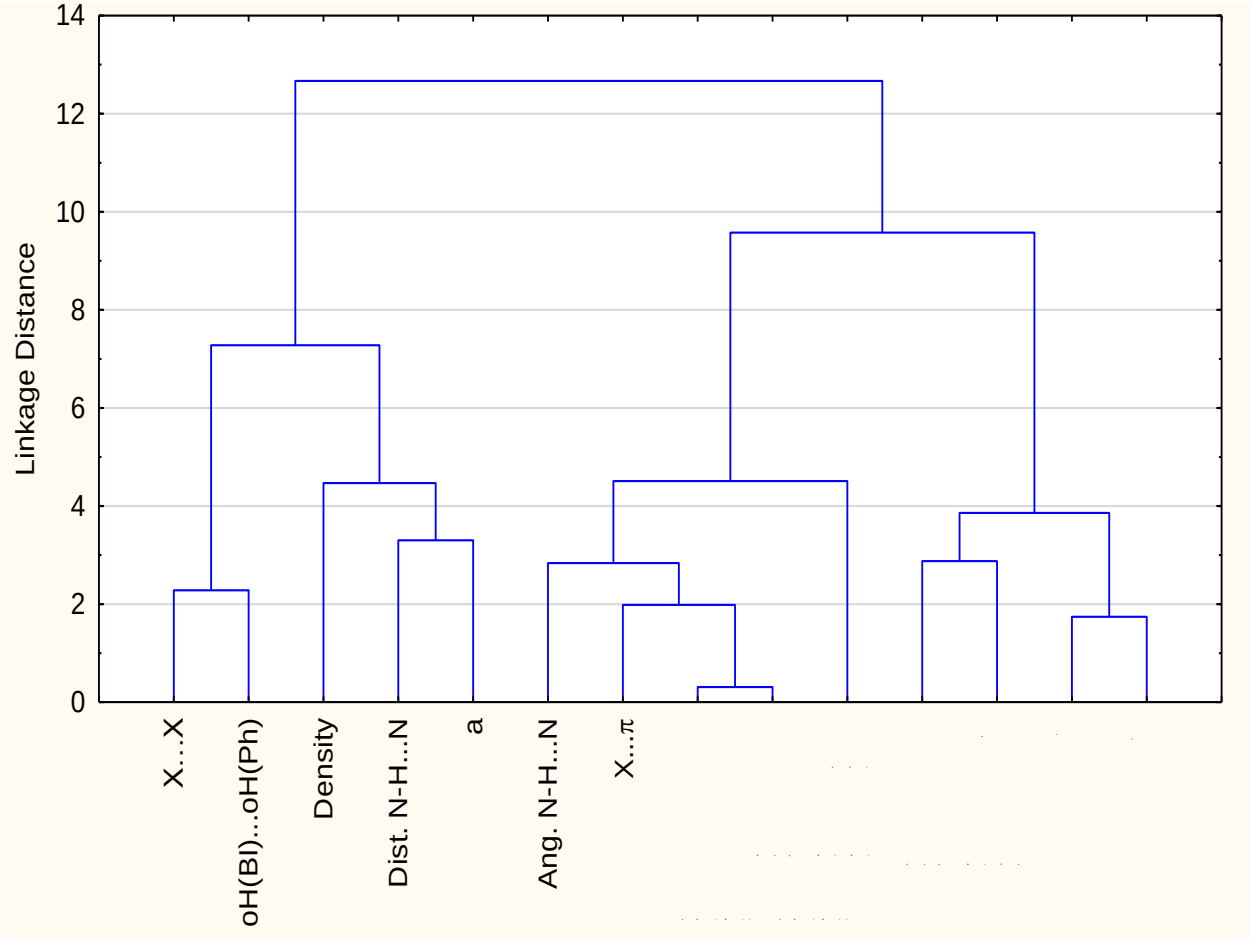

Figure S3 Tree diagram calculated for 14 variables (after standardization) shown in Table S2 by Ward's method using Euclidean distances. The parameters that are close in linkage distance are correlated with each other.

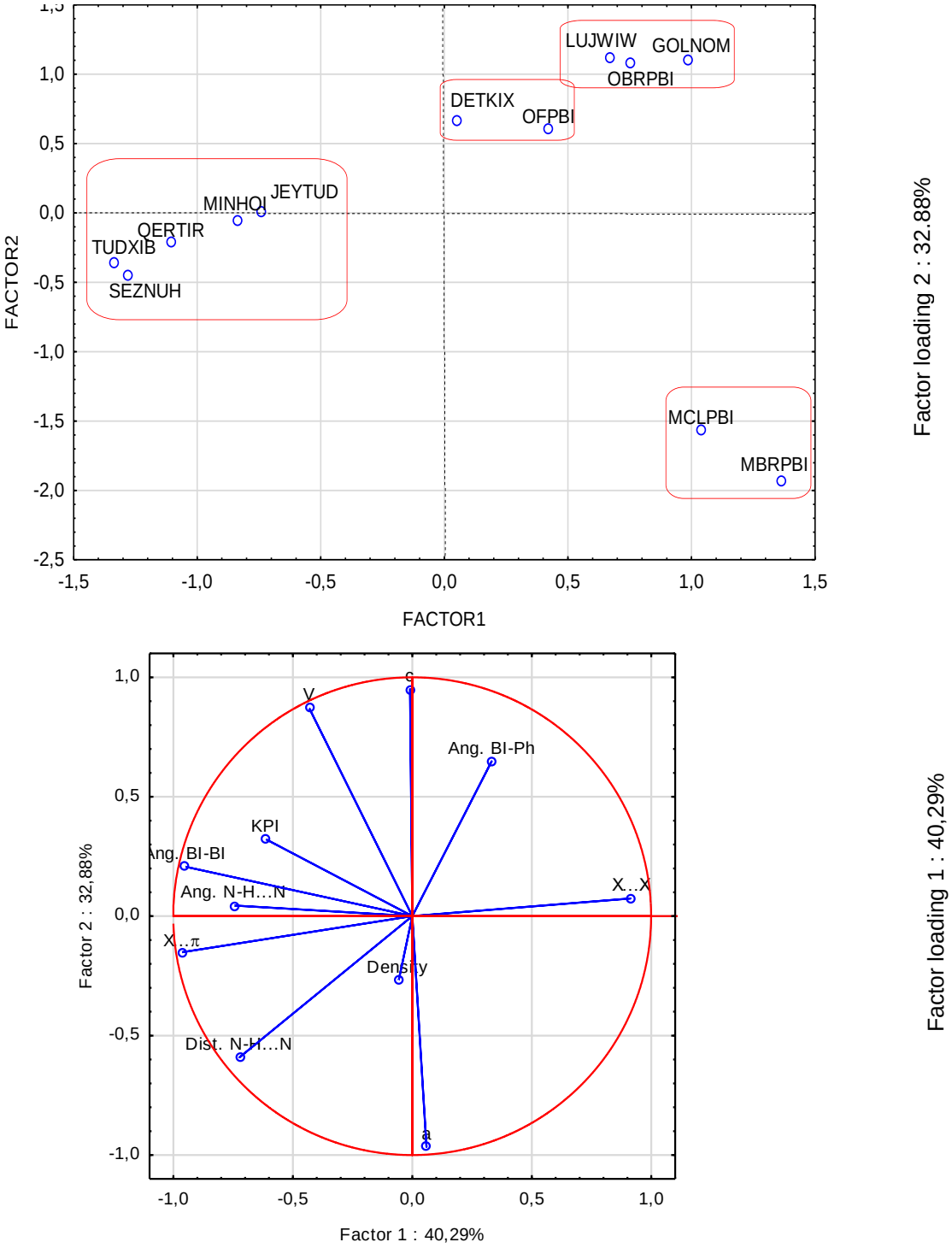

Figure S4. Scatterplot of Factor 2 against Factor 1 obtained from the PCA, indicating the groups obtained from the 12 investigated PBI cases (upper figure). The projection of the variables on the factor-plane showing the directionality of the values (lower figure).

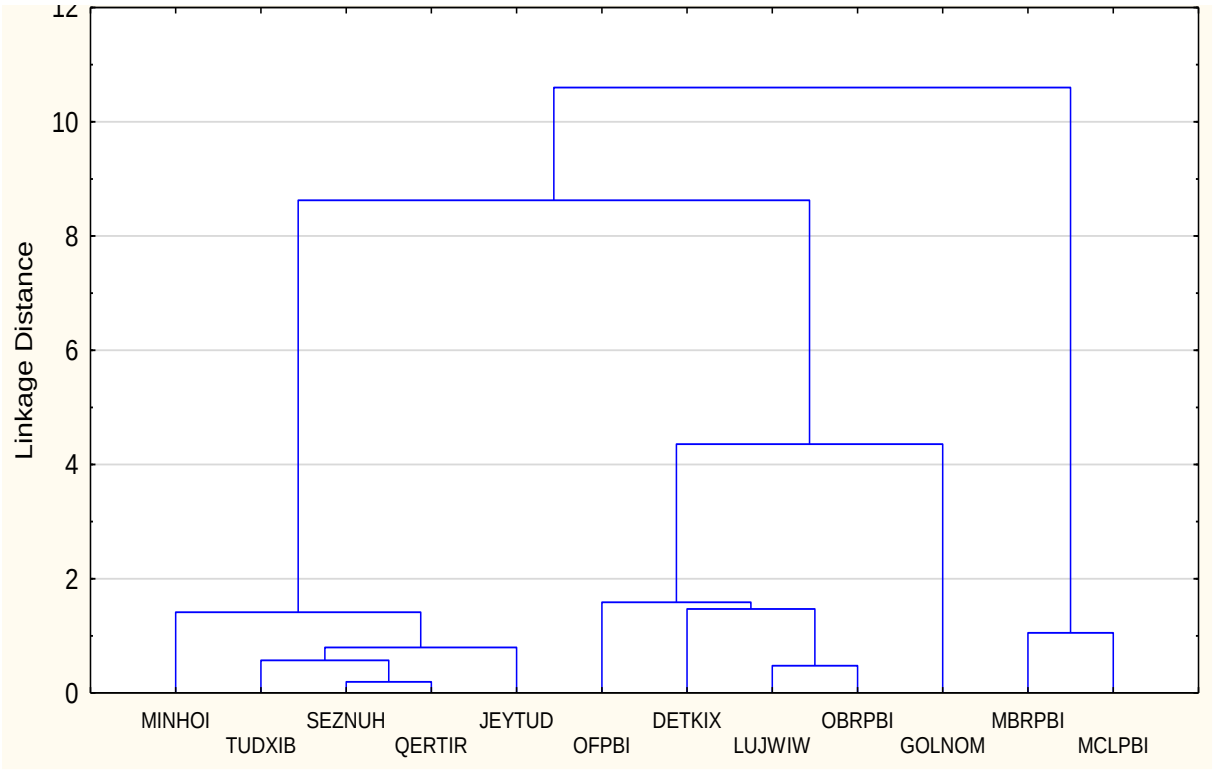

Figure S5 Tree diagram calculated for 5 variables (a, b, c cell lengths, Ang BI-BI and Ang BI-Ph) by Ward's method using Euclidean distances. The same groups could be detected as with the calculation using 14 variables.

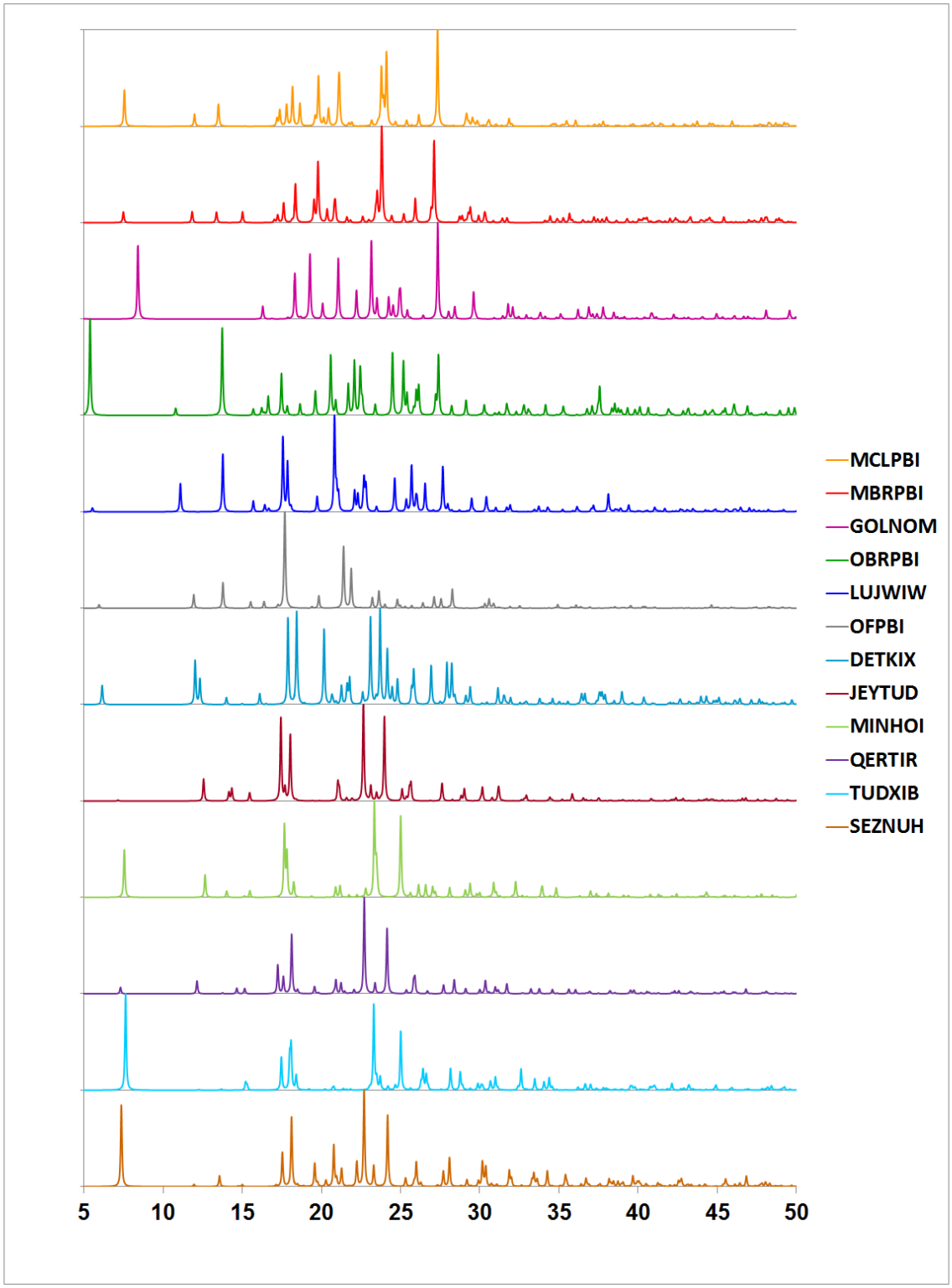

Figure S6 Simulated powder diffraction patterns of all 2-phenylbenzimidazoles derivatives in the fine-tuned series.

Table S1 Crystal data and details of the structure determination and refinement of the newly synthesized 2-phenylbenzimidazoles derivatives.

Sample code	MCLPBI	MBRPBI	OBRPBI	OFPBI
Empirical formula	C_{13} H_9 C_l N_2	C_{13} H_9 Br N_2	C_{13} H_9 Br N_2	C_{13} H_9 F N_2
Formula weight	228.67	273.13	273.13	212.22
Temperature	293(2) K	293(2) K	293(2) K	293(2) K
Radiation and wavelength	Mo-Kα λ =0.71075Å	Mo-Kα λ =0.71073Å	Cu-Kα λ =1.54178Å	Cu-Kα λ =1.54187Å
Crystal system	monoclinic	monoclinic	orthorhombic	orthorhombic
Space group	$P2_1/c$	$P2_1/c$	*Pbca*	*Pbca*
Unit cell dimensions	*a* =12.5894(10)Å	*a* =12.6755(16)Å	*a* =7.007(3)Å	*a* =7.1246(3)Å
	b =9.5051(8)Å	*b* =9.6559(18)Å	*b* =9.936(3)Å	*b* =10.0327(4)Å
	c =9.7495(7)Å	*c* =9.7645(15)Å	*c* =32.745(3)Å	*c* =29.6291(14)Å
	α =90°	α =90°	α =90°	α =90°
	β =111.933(8)°	β =111.638(9)°	β =90°	β =90°
	γ =90°	γ =90°	γ =90°	γ =90°
Volume	1082.22(16)Å^3	1110.9(3)Å^3	2279.7(16)Å^3	2117.86(16)Å^3
Z	4	4	8	8
Density (calculated)	1.403 Mg/m^3	1.633 Mg/m^3	1.592 Mg/m^3	1.331 Mg/m^3
Absorption coefficient, μ	0.322 mm^{-1}	3.671 mm^{-1}	4.675 mm^{-1}	0.763 mm^{-1}
F(000)	472	544	1088	880
Crystal colour	colourless	colourless	colourless	colourless
Crystal description	block	prism	block	needle
Crystal size	0.83 x 0.31 x 0.13 mm	0.5 x 0.1 x 0.05 mm	0.41 x 0.35 x 0.26 mm	0.56 x 0.14 x 0.06 mm
Absorption correction	numerical	numerical	numerical	numerical
Max. and min. transmission	0.986 1.000	0.3300 1.0000	0.839 0.916	0.997 0.972
θ–range for data collection	5.238 ≤ θ ≤ 25.345°	3.080 ≤ θ ≤ 25.338°	5.403 ≤ θ ≤ 68.165°	2.983 ≤ θ ≤ 68.219°
Index ranges	-15 ≤ *h* ≤15;-11≤ *k* ≤11;	-15 ≤ *h* ≤15;-11≤ *k* ≤11;	-8 ≤ *h* ≤8;-11≤ *k* ≤11;	-8 ≤ *h* ≤8;-11≤ *k* ≤11;
	-11 ≤ *l* ≤11	-11 ≤ *l* ≤11	-39 ≤ *l* ≤39	-35 ≤ *l* ≤35
Reflections collected	19656	28474	23169	21923
Completeness to 2θ	0.991	0.999	0.999	0.996
Independent reflections	1959 [*R*(int) =0.0906]	2033 [*R*(int) =0.1219]	2078 [*R*(int) =0.0779]	1924 [*R*(int) =0.0553]
Reflections *I*>2σ(*I*)	1509	1444	1916	1343
Refinement method	full-matrix least-squares on F^2	full-matrix least-squares on F^2	full-matrix least-squares on F^2	full-matrix least-squares on F^2
Data / restraints / parameters	1959 /0 /149	2033 /0 /145	2078 /0 /149	1924 /0 /149
Goodness-of-fit on *F*2	1.184	1.063	1.169	1.261
Final *R* indices [*I*>2σ(*I*)]	*R*1 =0.0706,	*R*1 =0.0832,	*R*1 =0.0629,	*R*1 =0.0944,
	*wR*2 =0.1056	*wR*2 =0.1344	*wR*2 =0.1377	*wR*2 =0.1417
R indices (all data)	*R*1 =0.1044,	*R*1 =0.1142,	*R*1 =0.0692,	*R*1 =0.1422,
	*wR*2 =0.1155	*wR*2 =0.1543	*wR*2 =0.1418	*wR*2 =0.1583
Max. and mean shift/esd	0.000;0.000	0.000;0.000	0.000;0.000	0.000;0.000
Largest diff. peak and hole	0.146;-0.186 e.Å^{-3}	0.470;-0.562 e.Å^{-3}	0.589;-0.602 e.Å^{-3}	0.149;-0.139 e.Å^{-3}

Table S2 The data of crystal parameters, molecular conformations and secondary interactions involved in the statistical analysis.

	V	a	b	c	KPI	Density	Ang. BI-BI	Dist. N-H...N	Ang. N-H...N	oH(BI)...oH(BI)	oH(BI)...oH(Ph)	Ang. BI-Ph	X...π[a]	X…X[a]
MCLPBI	1084.55	12.5984	9.5126	9.7559	68	1.4	67.83	2.869	153.55	2.854	3.921	26.81	4.058	-
MBRPBI	1112.48	12.682	9.6591	9.7701	58	1.625	68.53	2.875	153.64	2.900	3.931	28.2	4.073	-
GOLNOM	2301.73	5.6259	9.7396	42.007	69.7	1.519	74.49	2.824	147.71	2.934	3.710	27.7	3.612	-
OBRPBI	2279.76	7.007	9.936	32.745	66.1	1.592	92.88	2.801	161.98	4.033	2.641	42.43	4.186	-
LUJWIW	2220.55	7.0197	9.9261	31.8686	66.9	1.368	94.72	2.801	162.87	4.023	2.696	39.6	4.302	-
OFPBI	2117.86	7.1246	10.0327	29.6291	66.1	1.331	99.21	2.863	159.1	4.048	2.868	35.22	4.738	-
DETKIX	2434.95	8.563	9.910	28.694	65.6	1.436	112.33	2.866	154.96	4.911	2.429	42	4.698	3.506
JEYTUD	2080.09	8.574	9.830	24.68	67.4	1.355	117.76	2.875	163.31	5.066	2.391	29.23	-	3.456
MINHOI	2028.27	8.7195	9.9454	23.389	71.9	1.508	119.17	2.874	157.87	5.030	2.308	30	-	2.849
QERTIR	2157.21	9.1284	9.783	24.156	68.3	1.408	123.57	2.905	166.97	5.365	2.322	27	-	3.346
TUDXIB	2103.35	9.2302	9.8500	23.1247	72.5	1.558	124.4	2.924	166.4	5.416	2.286	26.9	-	3.37
SEZNUH	2205.7	9.3969	9.7876	23.982	68.5	1.645	126.66	2.925	168.05	5.476	2.307	26.77	-	3.433

[a]Not measurable distances were treated with value of 7 in the statistical analysis

Table S3 Multivariate data analysis: correlations between the variables of Table 2. Correlations marked with red are significant at $p < 0.05$

Variable	V	a	b	c	KPI	Density	Ang. BI-BI	Dist. N-H...N	Ang. N-H...N	Ang. BI-Ph	oH(BI)...oH(BI)	oH(BI)...oH(Ph)	X...π	X…X
V	1.0000	-0.8643	0.7450	0.8618	0.4542	-0.0979	0.6076	-0.1557	0.3391	0.4481	0.5804	-0.7138	0.2594	-0.3618
a	-0.8643	1.0000	-0.7040	-0.9803	-0.3828	0.2058	-0.2209	0.5047	-0.0464	-0.4608	-0.1705	0.3718	0.0742	-0.0623
b	0.7450	-0.7040	1.0000	0.5978	0.1988	-0.1995	0.5176	-0.2258	0.3371	0.6236	0.4849	-0.6810	0.2034	-0.1974
c	0.8618	-0.9803	0.5978	1.0000	0.3494	-0.1217	0.1584	-0.5051	-0.0383	0.4394	0.1135	-0.2940	-0.1478	0.0804
KPI	0.4542	-0.3828	0.1988	0.3494	1.0000	-0.1116	0.5225	0.2227	0.2815	-0.2461	0.4878	-0.4858	0.5166	-0.4841
Density	-0.0979	0.2058	-0.1995	-0.1217	-0.1116	1.0000	-0.0248	0.2086	0.0285	-0.2017	0.0007	0.0765	0.0456	-0.0668
Ang. BI-BI	0.6076	-0.2209	0.5176	0.1584	0.5225	-0.0248	1.0000	0.5651	0.7708	-0.0413	0.9957	-0.9536	0.8912	-0.8851
Dist. N-H...N	-0.1557	0.5047	-0.2258	-0.5051	0.2227	0.2086	0.5651	1.0000	0.4381	-0.6670	0.5720	-0.3137	0.7396	-0.6967
Ang. N-H...N	0.3391	-0.0464	0.3371	-0.0383	0.2815	0.0285	0.7708	0.4381	1.0000	-0.0357	0.7915	-0.7762	0.7305	-0.5172
Ang. BI-Ph	0.4481	-0.4608	0.6236	0.4394	-0.2461	-0.2017	-0.0413	-0.6670	-0.0357	1.0000	-0.0224	-0.2288	-0.4291	0.2631
oH(BI)...oH(BI)	0.5804	-0.1705	0.4849	0.1135	0.4878	0.0007	0.9957	0.5720	0.7915	-0.0224	1.0000	-0.9543	0.8859	-0.8909
oH(BI)...oH(Ph)	-0.7138	0.3718	-0.6810	-0.2940	-0.4858	0.0765	-0.9536	-0.3137	-0.7762	-0.2288	-0.9543	1.0000	-0.7608	0.7631
X...π	0.2594	0.0742	0.2034	-0.1478	0.5166	0.0456	0.8912	0.7396	0.7305	-0.4291	0.8859	-0.7608	1.0000	-0.8936
X…X	-0.3618	-0.0623	-0.1974	0.0804	-0.4841	-0.0668	-0.8851	-0.6967	-0.5172	0.2631	-0.8909	0.7631	-0.8936	1.0000

Table S4 Results of the principal component analysis (PCA)[a] showing the Factor scores

Case	Factor score 1	Factor score 2	Factor score 3	Factor score 4
MCLPBI	1.03990	-1.56353	-0.58671	-1.43546
MBRPBI	1.36559	-1.93514	0.60780	1.19183
GOLNOM	0.98696	1.09918	-2.53075	0.35633
OBRPBI	0.75370	1.08190	0.79443	1.46221
LUJWIW	0.67033	1.11529	0.91591	-0.60185
OFPBI	0.42281	0.60710	0.43639	-1.02666
DETKIX	0.05348	0.66293	1.01117	0.06694
JEYTUD	-0.74008	0.00916	0.43438	-1.08147
MINHOI	-0.83511	-0.05702	-0.73309	-0.12855
QERTIR	-1.10431	-0.20942	0.31867	-0.63109
TUDXIB	-1.33355	-0.35832	-0.60598	0.37607
SEZNUH	-1.27971	-0.45211	-0.06222	1.45170

[a]Principal component analysis (PCA) was applied on the data of Table S2 in order to reduce the data dimensionality. PCA transforms the original measured variables into new uncorrelated variables which are the linear combination of the original measured variables.

Table S5 Results of the principal component analysis (PCA) showing the Factor loadings[a]

Variable	Factor loading 1	Factor loading 2	Factor loading 3	Factor loading 4
V	-0.704710	-0.649947	0.126592	0.136603
a	0.365228	0.867490	-0.286128	-0.049409
b	-0.589857	-0.640201	-0.287120	0.004676
c	-0.299769	-0.857734	0.347636	0.139822
KPI	-0.601941	-0.036673	0.656766	-0.164052
Density	0.046785	0.274915	0.051327	0.950521
Ang. BI-BI	-0.985746	0.133485	-0.076787	-0.006573
Dist. N-H...N	-0.454414	0.810699	0.066736	0.030701
Ang. N-H...N	-0.755930	0.206891	-0.312821	0.020675
oH(BI)...oH(BI)	-0.974610	0.163567	-0.128075	0.012613
oH(BI)...oH(Ph)	0.968562	0.101926	0.202596	0.002830
Ang. BI-Ph	-0.016724	-0.767766	-0.557581	0.020577
X...pa	-0.847383	0.481696	0.057675	-0.055986
X…Xa	0.831036	-0.399668	-0.050351	0.001585
Explained Variance	6.478876	4.133687	1.217574	0.976217
Proportion of the Total variance	0.462777	0.295263	0.086970	0.069730

[a]Factor 1 and 2 explains the 75.8% of the variance