

Supporting Information

Design and Evaluation of Bonded Atom Pair Descriptors

Hany E. A. Ahmed, Martin Vogt, and Jürgen Bajorath^{*}

Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical
Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn,
Dahmannstr. 2, D-53113 Bonn, Germany.

Supplementary Tables S1 and S2

Table S1. Bonded atom pair dictionary.

Atom Pair	Index	Atom Pair	Index	Atom Pair	Index
C-!@C	1	O=!@P	51	n-!@C	101
C-@C	2	N-!@N	52	n-!@N	102
C=!@C	3	N-@N	53	n-!@O	103
C=@C	4	N=!@N	54	n-!@S	104
C#!@C	5	N=@N	55	n-@P	105
C#@C	6	[#7]:[#7\$(a1aaaa1)]	56	n-@C	106
[#6]:[#6\$(a1aaaa1)]	7	[#7]:[#7\$(a1aaaaa1)]	57	n-@N	107
[#6]:[#6\$(a1aaaaa1)]	8	N-!@[#G7]	58	n-@O	108
C-!@N	9	N-@[#G7]	59	n-@S	109
C-@N	10	n-[#G7]	60	C!@[C\$(C@*)]	110
C=!@N	11	N-!@P	61	C!@[N\$(N@*)]	111
C=@N	12	N-@P	62	N!@[C\$(C@*)]	112
C#!@N	13	N=!@P	63	N!@[N\$(N@*)]	113
[#6]:[#7\$(a1aaaa1)]	14	[#7]:[#15\$(a1aaaa1)]	64	O!@[C\$(C@*)]	114
[#6]:[#7\$(a1aaaaa1)]	15	[#7]:[#15\$(a1aaaaa1)]	65	C!@[N\$(C@*)]	115
C-!@O	16	N-!@S	66	C!@[N\$(O@*)]	116
C-@O	17	N-@S	67	C!@[N\$(S@*)]	117
C=!@O	18	N=!@S	68		
[#6]:[#8\$(a1aaaa1)]	19	[#7]:[#16\$(a1aaaa1)]	69		
[#6]:[#8\$(a1aaaaa1)]	20	[#7]:[#16\$(a1aaaaa1)]	70		
C-!@S	21	S-!@S	71		
C-@S	22	S-@S	72		
C=!@S	23	S=!@S	73		
[#6]:[#16\$(a1aaaa1)]	24	S=@S	74		
[#6]:[#16\$(a1aaaaa1)]	25	[#16]:[#16\$(a1aaaa1)]	75		
C-!@P	26	[#16]:[#16\$(a1aaaaa1)]	76		
C-@P	27	S-!@[#G7]	77		
C=!@P	28	S-!@P	78		
[#6]:[#15\$(a1aaaa1)]	29	S-@P	79		
[#6]:[#15\$(a1aaaaa1)]	30	S=!@P	80		
C-!@[#G7]	31	[#16]:[#15\$(a1aaaa1)]	81		
C-C[#G7]	32	[#16]:[#15\$(a1aaaaa1)]	82		
C=C[#G7]	33	P-!@P	83		
C=@A[#G7]	34	P-@P	84		
a-[#G7]	35	P=!@P	85		
C-@A[#G7]	36	P=@P	86		
N-!@O	37	[#15]:[#15\$(a1aaaa1)]	87		
N-@O	38	[#15]:[#15\$(a1aaaaa1)]	88		
N=!@O	39	P-!@[#G7]	89		
n@a[#8]	40	P-@[#G7]	90		
[#7]:[#8\$(a1aaaa1)]	41	p-[#G7]	91		
[#7]:[#8\$(a1aaaaa1)]	42	c-!@C	92		
O-!@O	43	c-!@N	93		
O-@O	44	c-!@O	94		
O-!@S	45	c-!@S	95		
O-@S	46	c-@P	96		
O=!@S	47	c-@C	97		
O-!@[#G7]	48	c-@N	98		
P-!@O	49	c-@O	99		
P-@O	50	c-@S	100		

The complete set of bonded atom pairs using generalized halogen atoms is provided

in SMARTS representation.

Table S2. Bonded atom pair dictionary with specific halogen atom types.

Atom Pair	Index	Atom Pair	Index	Atom Pair	Index
C-!@C	1	[#7]:[#16\$(a1aaaaa1)]	60	C=@A[I]	119
C-@C	2	S-!@S	61	a-[Cl]	120
C=!@C	3	S-@S	62	a-[F]	121
C=@C	4	S=!@S	63	a-[Br]	122
C#!@C	5	S=@S	64	a-[I]	123
C#@C	6	[#16]:[#16\$(a1aaaaa1)]	65	C-@A[Cl]	124
[#6]:[#6\$(a1aaaaa1)]	7	[#16]:[#16\$(a1aaaaa1)]	66	C-@A[F]	125
[#6]:[#6\$(a1aaaaa1)]	8	S-!@P	67	C-@A[Br]	126
C-!@N	9	S-@P	68	C-@A[I]	127
C-@N	10	S=!@P	69	O-!@[Cl]	128
C=!@N	11	[#16]:[#15\$(a1aaaaa1)]	70	O-!@[F]	129
C=@N	12	[#16]:[#15\$(a1aaaaa1)]	71	O-!@[Br]	130
C#!@N	13	P-!@P	72	O-!@[I]	131
[#6]:[#7\$(a1aaaaa1)]	14	P-@P	73	N-!@[Cl]	132
[#6]:[#7\$(a1aaaaa1)]	15	P=!@P	74	N-!@[F]	133
C-!@O	16	P=@P	75	N-!@[Br]	134
C-@O	17	[#15]:[#15\$(a1aaaaa1)]	76	N-!@[I]	135
C=!@O	18	[#15]:[#15\$(a1aaaaa1)]	77	N-@[Cl]	136
[#6]:[#8\$(a1aaaaa1)]	19	c-!@C	78	N-@[F]	137
[#6]:[#8\$(a1aaaaa1)]	20	c-!@N	79	N-@[Br]	138
C-!@S	21	c-!@O	80	N-@[I]	139
C-@S	22	c-!@S	81	n-[Cl]	140
C=!@S	23	c-@P	82	n-[F]	141
[#6]:[#16\$(a1aaaaa1)]	24	c-@C	83	n-[Br]	142
[#6]:[#16\$(a1aaaaa1)]	25	c-@N	84	n-[I]	143
C-!@P	26	c-@O	85	S-!@[Cl]	144
C-@P	27	c-@S	86	S-!@[F]	145
C=!@P	28	n-!@C	87	S-!@[Br]	146
[#6]:[#15\$(a1aaaaa1)]	29	n-!@N	88	S-!@[I]	147
[#6]:[#15\$(a1aaaaa1)]	30	n-!@O	89	P-!@[Cl]	148
N-!@O	31	n-!@S	90	P-!@[F]	149
N-@O	32	n-@P	91	P-!@[Br]	150
N=!@O	33	n-@C	92	P-!@[I]	151
n@a[#8]	34	n-@N	93	P=P@[Cl]	152
[#7]:[#8\$(a1aaaaa1)]	35	n-@O	94	P=P@[F]	153
[#7]:[#8\$(a1aaaaa1)]	36	n-@S	95	P=P@[Br]	154
O-!@O	37	C!@[C\$(C@*)]	96	P=P@[I]	155
O-@O	38	C!@[N\$(N@*)]	97	p-[Cl]	156
O-!@S	39	N!@[C\$(C@*)]	98	p-[F]	157
O-@S	40	N!@[N\$(N@*)]	99	p-[Br]	158
O=!@S	41	O!@[C\$(C@*)]	100	p-[I]	159
P-!@O	42	C!@[N\$(C@*)]	101		
P-@O	43	C!@[N\$(O@*)]	102		
O=!@P	44	C!@[N\$(S@*)]	103		
N-!@N	45	C-!@[Cl]	104		
N-@N	46	C-!@[F]	105		
N=!@N	47	C-!@[Br]	106		
N=@N	48	C-!@[I]	107		
[#7]:[#7\$(a1aaaaa1)]	49	C-C[Cl]	108		
[#7]:[#7\$(a1aaaaa1)]	50	C-C[F]	109		
N-!@P	51	C-C[Br]	110		
N-@P	52	C-C[I]	111		
N=!@P	53	C=C[Cl]	112		
[#7]:[#15\$(a1aaaaa1)]	54	C=C[F]	113		

<chem>[#7]:[#15\$(a1aaaaa1)]</chem>	55	<chem>C=C[Br]</chem>	114
<chem>N-!@S</chem>	56	<chem>C=C[I]</chem>	115
<chem>N-@S</chem>	57	<chem>C=@A[Cl]</chem>	116
<chem>N=!@S</chem>	58	<chem>C=@A[F]</chem>	117
<chem>[#7]:[#16\$(a1aaaaa1)]</chem>	59	<chem>C=@A[Br]</chem>	118

The complete set of bonded atom pairs using specific halogen atom types is provided in SMARTS representation.