

# Relationships between Chemical Structures and Mutagenicity: A Preliminary Survey for a Database of Mutagenicity Test Results of New Work Place Chemicals

Katsuhiko SAWATARI<sup>1\*</sup>, Yoshifumi NAKANISHI<sup>1</sup> and Taijiro MATSUSHIMA<sup>2</sup>

<sup>1</sup> National Institute of Industrial Health, 21–1, Nagao 6-chome, Tama-ku, Kawasaki 214-8585, Japan

<sup>2</sup> Japan Bioassay Research Center, 2445 Hirasawa, Hadano, Kanagawa 257-0015, Japan

*Received March 23, 2001 and accepted June 7, 2001*

**Abstract:** A database of mutagenicity test results of new chemicals has been developed. Based on the amendment of the Industrial Safety and Health Law (ISHL) in 1979, manufacturers and importers in Japan are required to register any new work place chemicals with bacterial mutagenicity test results. At present more than ten thousand substances have been examined. We have surveyed correlations between 44 substructures and mutagenicity in 2,857 ISHL data as well as in 1,207 National Toxicology Program data as a preliminary analysis. The percentages of the mutagenic compounds were calculated. High percentages were found for electrophilic reagents such as epoxides(63%), aromatic nitro compounds(49%) and primary alkyl monohalides(46%). 71% of peroxides was found to be mutagenic. The results suggest that several types of reactions such as nucleophilic substitution reaction, nitrenium cation reaction and radical reaction are included in the process of the mutagenic alterations of DNA.

**Key words:** Database of bacterial mutagenicity tests, Molecular substructure, QSAR

In Japan based on the Industrial Safety and Health Law (ISHL) amended in 1979, manufacturers and importers, who introduce any new work place chemicals, are required to conduct bacterial mutagenicity tests and to submit the reports to the Ministry of Labour of Japan. By the beginning of the year 2001, the number of tested for mutagenicity and registered new chemicals has exceeded ten thousand. The authors have started to develop a database using the ISHL data for a preliminary analysis of the chemical substructure and mutagenicity relationships: our future objective is to extend the use of the database for the quantitative structure-activity relationships (QSAR) analyses of these chemicals. For the preliminary analysis we also developed a database using published U.S. National Toxicology Program data (NTP data)<sup>1–5</sup>.

Compounds with one or more electrophilic substructure(s) such as halogenated aliphatic substructure and aromatic nitro

substructure, have been known to be mutagenic. Their mutagenic potentials were accounted for by the fact that DNA molecules with many lone pair electrons are prone to electrophilical attack<sup>6</sup>. High correlations between the electrophilicity of compounds and their mutagenicity and/or carcinogenicity were also reported<sup>7–9</sup>. Ashby *et al.* proposed a concept 'structural alert' using a few hundred data of NTP, and reported that a high correlation (more than 90% overall) existed between some substructures of molecules and mutagenicity<sup>10, 11</sup>. However, Ashby *et al.* did not present numerically individual correlations between individual substructures and mutagenicity since their database was not so large enough. ISHL data would form a large database, and a preliminary analysis is presented. Ashby's main interest was in carcinogenicity, therefore they surveyed only 301 mutagenicity data though more than 1200 data were reported by NTP. A survey of whole NTP mutagenicity data would also be informational.

The databases were developed with ChemFinder Pro 5.0

\*To whom correspondence should be addressed.

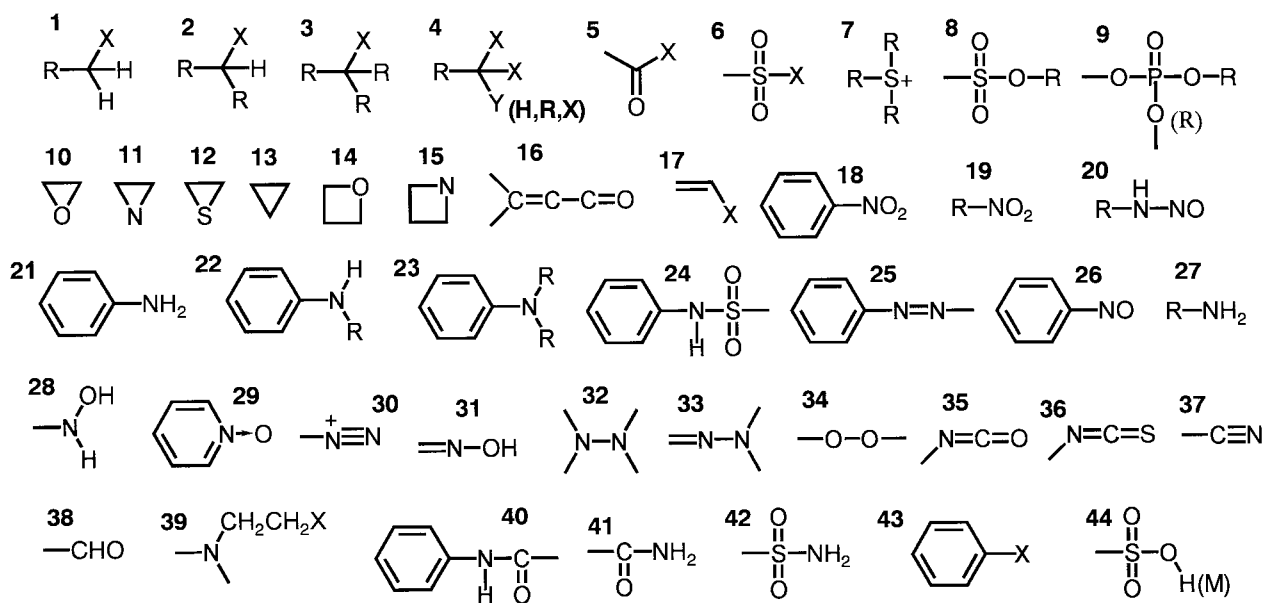


Fig. 1. Substructures examined. X: halogen.

(CambridgeSoft Corp., Massachusetts). At the present time, mutagenicity test results of 4,224 chemicals of the ISHL data have been compiled in the database. Those chemicals include 678 mixtures, 662 polymers and 27 inorganic substances. Therefore the data of 2,857 organic compounds were surveyed in the present work. NTP data consists of 1,207 organic compounds. Relationships between mutagenicity and 44 substructures, listed in Table 1 and shown in Fig. 1, were examined about ISHL data as well as NTP data. The 44 substructures included most of the structural alerts presented by Ashby *et al.* Each of the 2,857 (ISHL) and 1,207 (NTP) compounds was inquired whether it had one or more of the 44 substructures and whether it was reported as mutagenicity positive or negative. And the numbers of positive compounds and negative compounds of each substructure were summed up. Compounds with two or more substructures are counted about each substructure. Percentages of positive compounds in compounds of each substructure were calculated (Table 1). ChemFinder can extract compounds with a certain substructure from the whole data, but the program cannot possibly recognize some substructures, therefore all data were ascertained by repeated inquiries with human eyes.

13% of 3,863 new work place chemicals registered in Japan by 1989 were reported as positive<sup>12</sup>, and 13% of 776 new chemicals surveyed in Germany were reported as positive<sup>13</sup>. The percentage of positive chemicals did not change when 10,512 chemicals were surveyed at the

beginning of the year of 2001 in Japan.

As for organic compounds surveyed in the present work, 15.8% of ISHL organic compounds were positive. On the other hand, 34.2% of NTP data were positive. NTP is supposed to have examined chemicals which were suspected as mutagenic. In order to compare ISHL data with NTP data, both data were normalized. The values indicating relative mutagenic potential were calculated by dividing the percentage of positive compounds of each substructure by the percentage of the whole positive compounds: 15.8% for the ISHL data and 34.2% for the NTP data (Table 1). The results of both ISHL data and NTP data are similar to each other basically. However there are some differences which could not be neglected. This suggests that the number of NTP data is not enough large. The number of ISHL data was greater than NTP data, and ISHL data were not biased in selection of tested chemicals. Therefore the values of ISHL data are discussed below. However, ISHL data of the present work might not possibly be enough for fine discussions. General discussions could be held.

In substructures about which more than 10 test results were reported, epoxide structure was most highly correlated with mutagenicity, and the percentage of positive compounds was 63.0%; followed by aromatic nitro structure (48.8%), primary alkyl monohalide structure (45.8%) and ester of sulphonic acid structure (42.2%). The percentages of positive compounds of secondary alkyl monohalide structure and tertiary alkyl monohalide structure were 31.8% and 40.0%,

**Table 1. Substructures examined and the results of countings and calculations**

#	Substructure	ISHL data				NTP data			
		Positive <sup>a</sup>	Negative <sup>b</sup>	% <sup>c</sup>	Relative Mutagen. Potential <sup>d</sup>	Positive <sup>e</sup>	Negative <sup>f</sup>	% <sup>g</sup>	Relative Mutagen Potential <sup>h</sup>
	All	450	2407	15.8	—	413	794	34.2	—
1	Primary alkyl monohalide	44	52	45.8	2.91	41	15	73.7	2.15
2	Secondary alkyl monohalide	14	30	31.8	2.02	11	4	73.3	2.14
3	Tertiary alkyl monohalide	4	6	40.0	2.54	3	7	30.0	0.88
4	Alkyl multihalide	11	110	9.1	0.58	18	43	29.5	0.86
5	Acyl halide	21	30	41.2	2.61	9	2	81.8	2.39
6	Sulfonyl halide	9	11	45.0	2.86	0	0	—	—
7	Sulfonium salt	4	3	57.1	3.63	0	0	—	—
8	Esters of sulfonic acid	19	26	42.2	2.68	4	1	80.0	2.34
9	Esters of phosphoric acid	1	21	4.5	0.29	13	19	40.6	1.19
10	Epoxides	17	10	63.0	4.00	9	4	69.2	2.02
11	Ethyleneimine	3	0	100.0	6.35	2	0	100.0	2.92
12	Ethylen sulfide	2	1	66.7	4.23	0	0	—	—
13	Cyclopropane	2	58	3.3	0.21	0	0	—	—
14	Oxetane	0	3	0.0	0.00	3	0	100.0	2.92
15	Trimethyleneimine	6	53	10.2	0.65	0	2	0.0	0.00
16	$\alpha$ , $\beta$ -unsaturated carbonyl	22	82	21.2	1.34	17	46	27.0	0.79
17	Monohaloalkene	3	10	23.1	1.47	7	4	63.6	1.86
18	Aromatic nitro	101	106	48.8	3.10	81	24	77.1	2.25
19	Aliphatic nitro	4	2	66.7	4.23	10	6	62.5	1.82
20	N-Nitroso alkyl	0	0	—	—	7	1	87.5	2.56
21	Aromatic primary amine	38	142	21.1	1.34	97	61	61.4	1.79
22	Aromatic secondary amine	20	142	12.3	0.78	19	24	44.2	1.29
23	Aromatic tertiary amine	11	41	21.2	1.34	15	5	75.0	2.19
24	Aromatic sulfonamide	9	24	27.3	1.73	0	2	0.0	0.00
25	Aromatic azo	41	79	34.2	2.17	22	18	55.0	1.61
26	Aromatic nitroso	0	1	0.0	0.00	3	1	75.0	2.19
27	Aliphatic primary amine	16	90	15.1	0.96	9	40	18.4	0.54
28	Hydroxylamine	1	0	100.0	6.35	3	0	100.0	2.92
29	Pyridine-N-oxide	4	5	44.4	2.82	0	0	—	—
30	Diazonium salt	3	7	30.0	1.90	0	0	—	—
31	Oxime	5	18	21.7	1.38	2	2	50.0	1.46
32	Hydrazine	18	30	37.5	2.38	7	10	41.2	1.20
33	Hydrazone	6	12	33.3	2.12	5	2	71.4	2.09
34	Peroxide and peracid	5	2	71.4	4.53	4	5	44.4	1.30
35	Isocyanate	0	1	0.0	0.00	3	2	60.0	1.75
36	Isothiocyanate	0	4	0.0	0.00	1	0	100.0	2.92
37	Nitryl	39	129	19.4	1.23	7	31	18.4	0.54
38	Aldehyde	12	46	20.7	1.31	4	29	12.1	0.35
39	Nitrogen mustard	0	0	—	—	5	2	71.4	2.09
40	Anilide	32	187	14.6	0.93	16	17	48.5	1.42
41	Carboxylic acid amide	0	13	0.0	0.00	10	25	28.6	0.84
42	Sulfamide	0	4	0.0	0.00	0	4	0.0	0.00
43	Halogenated aromatic ring	73	464	13.6	0.86	37	115	24.3	0.71
44	Sulfonic acid	20	129	13.4	0.85	11	24	31.4	0.92

a, e: numbers of compounds which have each substructure and were reported as mutagenicity positive; b, f: those of mutagenicity negative; c= $a/(a + b) \times 100$ ; g= $e/(e + f) \times 100$ ; d= $c/15.8$  (the percentage of the positive compounds of all of ISHL data); h= $g/34.2$  (the percentage of the positive compounds of all of NTP data).

respectively. The percentages of positive compounds of acyl halide structure and sulfonyl halide structure were 41.2% and 45.0%, respectively. Alkyl monohalides, acyl halides, sulfonyl halides and esters of sulphonic acid are known to undergo nucleophilic substitution reaction. Open-ring reactions of epoxides are similar to nucleophilic substitution reaction basically. These substances are supposed to readily react with nucleophilic substances such as DNA. Sulfonium salts are known as strong alkylating reagents and readily undergo nucleophilic substitution reaction, and the percentage of positive compounds was high (57.1%). On the other hand, alkyl multihalides are not ready to undergo nucleophilic substitution reactions, and the percentage of positive compounds was low (9.1%). Aromatic nitro compounds are known to be reduced to hydroxylamines and to become nitrenium cations which also readily react with nucleophilic substances. On the other hand, the percentage of positive compounds of aromatic primary amines was not so high (21.1%) though aromatic primary amines are also oxidized to hydroxylamines and become nitrenium cations. Aromatic azo compounds are decomposed oxidatively to hydroxylamines and become nitrenium cations, and the percentage of positive compounds was 34.2%. The percentages of positive compounds of monohaloalkene, aldehyde and  $\alpha$ ,  $\beta$ -unsaturated carbonyl structure, which were regarded as structural alerts by Ashby and are known to undergo nucleophilic substitution reaction, were 23.1%, 20.7% and 21.2%, respectively. The values were not so high as expected. The percentage of positive compounds of ester of phosphoric acids was very low (0.0%), though it is one of the Ashby's structural alerts and known to undergo nucleophilic substitution reactions. Hydrazines and hydrazones are very unstable and reactive compounds and the percentages of positive compounds were high (37.5% and 33.3%, respectively). Peroxides, which generate free radicals, are moreover reactive and the percentage of positive compounds was very high (71.4%) though the number of the test data was small.

If a non-electrophilic substructure such as halogenated aromatic ring substructure was included in a compound together with electrophilic substructures such as aromatic nitro substructure, and if the compound was positive, the number of positive compounds of the non-electrophilic substructure was increased by the present way of counting. But it would be irrational because the mutagenicity of the compound probably resulted from the electrophilic substructure. The non-electrophilic substructure was correlated with mutagenicity to an excessive degree. But any data modification would be arbitrary. Therefore no data

modification was adopted. The values of high percentages in Table 1 rather exactly represent the correlations, but those of low percentages could not so exactly represent the correlations.

Correlations between mutagenicity and individual substructures were presented as numerical values using 2,857 test data. Correlations between mutagenicity and electrophilic substructures were high. This supports the validity of the hypothesis of Miller *et al.* The results further suggest that several types of reactions such as nucleophilic substitution reaction, nitrenium cation reaction and radical reaction are included in the process of the mutagenic alterations of DNA. QSAR analyses based on these ideas would be successful.

## References

- 1) Haworth S, Lawlor T, Mortelmans K, Speck W (1983) Salmonella mutagenicity test results for 250 chemicals. *Environ Mutagen* **5** Suppl 1, 3–142.
- 2) Mortelmans K, Haworth S, Lawlor T, Speck W (1986) Salmonella mutagenicity tests: II Results from the testing of 270. *Environ Mutagen* **8** Suppl 7, 1–119.
- 3) Zeiger E, Anderson B, Haworth S, Lawlor T, Mortelmans K, Speck W (1987) Salmonella mutagenicity tests: III Results from the testing of 255. *Environ Mutagen* **9** Suppl 9, 1–109.
- 4) Zeiger E, Anderson B, Haworth S, Lawlor T, Mortelmans K, Speck W (1987) Salmonella mutagenicity tests: IV Results from the testing of 300. *Environ Mol Mutagen* **11** Suppl 12, 1–157.
- 5) Zeiger E, Anderson B, Haworth S, Lawlor T, Mortelmans K (1992) Salmonella mutagenicity tests: V Results from the testing of 311. *Environ Mol Mutagen* **19** Suppl 21, 2–141.
- 6) Miller JA, Miller EC (1977) Ultimate chemical carcinogens as reactive mutagenic electrophiles. In: *Origins of human cancer*. eds. by Hiatt HH, Watson JD, Wintson JA, 605–28, Cold Spring Harbor Laboratory, New York.
- 7) Bartsch H, Terracini B, Malaveille C, Tomatis L, Wahrendorf J, Brun G, Dodet B (1983) Quantitative comparison of carcinogenicity, mutagenicity and electrophilicity of 10 direct-acting alkylating agents and the initial O<sup>6</sup>; 7-alkylguanine ratio in DNA with carcinogenic potency in rodents. *Mutat Res* **110**, 181–219.
- 8) Shelby MD (1988) The genetic toxicity of human carcinogens and its implication. *Mutat Res* **204**, 3–15.

- 9) Benigni R, Cotta-Ramusino M, Andreoli C, Giuliani A (1992) Electrophilicity as measured by  $K_e$ : molecular determinants, relationship with other physical-chemical and quantum mechanical parameters, and ability to predict carcinogenicity. *Carcinogenesis* **13**, 547–53.
- 10) Ashby J (1985) Fundamental structural alerts to potential carcinogenicity or noncarcinogenicity. *Environ Mutagen* **7**, 919–21.
- 11) Ashby J, Tennant RW (1991) Definitive relationships among chemical structure, carcinogenicity and mutagenicity tested by U.S. NTP. *Mutat Res* **257**, 229–306.
- 12) Matsushima T (1990) Genotoxicity of new Japanese Chemicals. In: *Mutation and the environment*. eds. by Mendelsohn ML, Albertini RJ, Part E Environmental genotoxicity, risk, and modulation, 251–5, Wiley-Liss, New York.
- 13) Broschinski L, Madle S, Hensel C (1998) Genotoxicity tests for new chemicals in Germany: routine in vitro test systems. *Mutat Res* **418**, 121–9.