Chem. Senses 27: 399–405, 2002

The Creation of a Database of Odorous Compounds Focused on Molecular Rigidity and Analysis of the Molecular Features of the Compounds in the Database

Fumiko Yoshii, Yoshiaki Yamada[1], Takashi Hoshi[2] and Hisahiro Hagiwara

Graduate School of Science and Technology, Niigata University, 8050 Ikarashi 2 Nocho, Niigata 950-2181, [1]Graduate School of Medical and Dental Sciences, Niigata University, 2–5274 Gakkocho-dori, Niigata 951-8514 and [2]Faculty of Engineering, Niigata University, 8050 Ikarashi 2 Nocho, Niigata 950-2181, Japan

Correspondence to be sent to: Fumiko Yoshii, Graduate School of Science and Technology, Niigata University, 8050 Ikarashi 2 Nocho, Niigata 950-2181, Japan. e-mail: fuyoshii@mb.infoweb.ne.jp

Abstract

It is important to select odorous molecules for experiments on olfaction and for the development of an electronic nose. Odorous molecules having a small number of conformers, namely structurally rigid molecules, are assumed to interact with a small number of types of olfactory receptor proteins or to interact with the proteins in a simpler manner than that of fairly flexible molecules. Focusing on the rigidity of molecular structures, we collected 287 odorous molecules from data sources, which included 1205 chemicals in total and a database of the 287 odorous molecules (*DB_odMOL*) was created using CS ChemFinder Pro (version 5.0). The logarithmic value of the octanol/water partition coefficient (log *P*) and melting point, boiling point and vapour pressure of the molecules were estimated using CS Chem3D Pro. The database *DB_odMOL* accumulates these estimated data in addition to literature values for odour quality, odour detection thresholds and the safety of molecules. The rigidity of the 287 molecules was further analysed by conformational analysis performed by molecular mechanics using Conformer® in CS Chem3D Pro (version 5) and 72 rigid odorous molecules were selected. The 287 molecules were also analysed based on atomic composition, substructure and molecular size. Sixty-two odorous molecules among the 72 rigid odorous molecules were further selected based on their atomic composition. The 62 rigid molecules with simple atomic composition that were finally selected should be useful for researchers in choosing odorous molecules for the study of olfaction, including the fields of molecular biology, physiology, structure–odour relationships and other fields of the study concerning odour.

Introduction

Olfactory sensation is one of the chemical senses and is triggered by odorous chemical molecules. However, contrary to the importance of the molecular features of odorants, in most cases sufficient attention has not been paid to selecting odorants for research into olfaction. Furthermore, how many researchers in the study of olfaction have noticed that the shape of an odorous molecule itself can sometimes be changed to different shapes due to molecular flexibility? The conformation of odorants has been considered in the study of quantitative or qualitative structure–odour relationships (Buchbauer *et al.*, 1994; Rognan and Chastrette, 1994; Yoshii and Hirono, 1996; Gorbachov and Rossiter, 1999) by assuming interaction between odorous molecules and olfactory receptor proteins. The interaction of fairly flexible molecules and receptor proteins is considered to be more complicated than that between rather rigid molecules and the proteins. There is no yet proven direct experimental evidence that the flexibility of conformers of odorous compounds influences actual odour sensation in human subjects.

In this study rather rigid odorous molecules were collected and a database of the odorous molecules was constructed with accumulation of other molecular parameters. The aim of this study was that the creation of the database would help researchers in the study of olfaction in selecting odorous molecules by consulting molecular properties. The creation of this database is also the first step in the selection of standard odorous molecules that can be used for a wide range of experiments on olfaction. When standard odorous molecules are employed in various fields of study of olfaction, including electrophysiology, molecular biology, olfactory evaluation and the development of sensors, comparison of the data from different studies should become much more general and easier and bring fruitful

results towards an understanding of the mechanism of olfaction.

Materials and methods

A flow chart of the procedure used in the creation of the database is described in Figure 1.

Selection of compounds

The data sources for the odorous compounds are shown in Figure 1. Standard odorous materials recommended by the International Organization for Standardization (ISO 5496:1992), T&T olfactometer (clinical test materials for olfaction that are widely used in Japan), compilation and standardized data of detection threshold values of odorous molecules in the gas phase (Devos *et al.*, 1990), literature relating to perfumery materials (Nakajima, 1995) and a World Wide Web site of flavour materials (Acree and Arn, 1997) were selected for choosing a wide range of odorous compounds. The data sources were considered as covering odorous compounds for practical use with intense odour and perfume and flavour materials. The odorous molecules for the database were selected from these data sources by the criteria described below.

Rigidity

Odorous molecules were selected when flexible single carbon–carbon bonds in the main or side chain structure were shorter than the four-membered carbon bond (C-C-C-C). When the compounds had one or more ring structures, the compounds that had smaller rings than a seven-membered ring were selected. When more than one double bond existed in a seven-membered ring, the compounds which contained the seven-membered ring were selected as exceptions.

Odour strength

The detection threshold value in the gas phase was treated as the strength of an odour. The detection threshold value is the lowest concentration of odorous compounds at which human subjects can detect the existence of the odour. The *p*. ol (volume) weighted and compiled in the literature (Devos *et al.*, 1990) was used for selection of the molecules. The *p*. ol (volume) is the negative log of the concentration expressed in volumetric fractions, where *p*. ol = 6 equals the threshold of 1 p.p.m. and *p*. ol = 9 equals the threshold of 1 p.p.b. The *p*. ol (volume) weighted is the standardized threshold by weighing threshold values from different sources. The value of *p*. ol (volume) weighted was rounded off down to one decimal point for selection of the compounds. Compounds were selected when their rounded off values were larger than 5.

Atomic composition

Compounds consisting of only carbon atoms (C), hydrogen atoms (H), nitrogen atoms (N), oxygen atoms (O) and sulphur atoms (S) were selected. The majority of organic

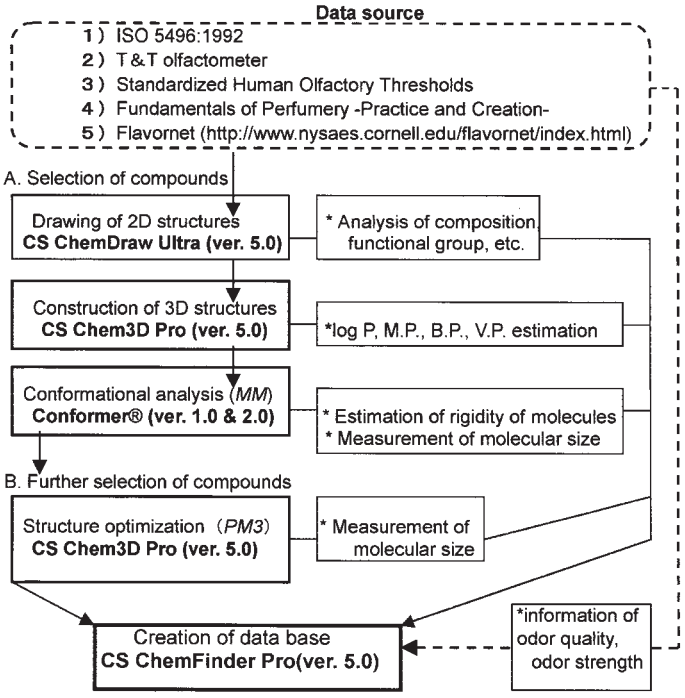

Figure 1 Procedure for the analysis of odorous molecules and creation of the database.

odorous compounds consist of C plus H and N plus O plus S, though there are some compounds in the data sources involving selenium atoms or halogen atoms, such as chlorine and bromine.

Other criteria

Molecules having one or more asymmetric carbons in their structures were not selected because remarkable odour quality differences were sometimes recognized between enantiomers (Boelens *et al.*, 1993; Laska and Teubner, 1999). Gaseous compounds at room temperature were also excluded because the difficulty of handling them does not meet the criteria for future application in experiments on olfaction.

Construction of the structures of the odorous molecules, conformational analysis of the molecules and further selection of the molecules

Two-dimensional molecular structures of all the selected odorous molecules were drawn by CS ChemDraw Ultra (version 5.0, CambridgeSoft Corporation, Cambridge, MA). These structures were converted to three-dimensional structures using CS Chem3D Pro (version 5.0, CambridgeSoft Corporation). Minimization of energy of the molecules was performed by molecular mechanics parameters in CS Chem3D Pro. Conformer® (versions 1 and 2, Princeton Simulations, Princeton, NJ) performed conformational analysis of all the selected odorous molecules using molecular mechanics parameters in CS Chem3D Pro. The

criterion for selection of the conformers by conformational analysis was the calculated energy of the conformers. Conformers were selected when their energies were between the energy of the lowest energy conformer and 5 kcal/mol higher than the lowest energy. The total number of conformers searched according to the energy criteria is considered as the index of flexibility, though the conformers searched are still dependent on the level of calculation. The molecule was considered to be rigid when only one conformer was selected by conformational analysis. The rigid molecules were further optimized by the semi-empirical molecular orbital method PM3 in CS Chem3D Pro. The hardware used for this work was Celebris GL 6200 (Compaq Computer Kabushiki Kaisha, previously Digital Equipment Corporation Japan, Tokyo, Japan) and Dynabook DS50C/1CA model (Toshiba Corporation, Tokyo, Japan). The operating system was Windows 98 (Microsoft Corporation, Redmond, WA).

Calculation and analysis of the molecular features

The two-dimensional structures of the selected odorous molecules were analysed in terms of atomic composition and functional groups (ring structure, carbonyl group, hydroxyl group, ether substructure, etc.). The logarithmic values of the 1-octanol/water partition coefficient (log *P*) and the melting points, boiling points and vapour pressures of the molecules were estimated using CS Chem3D Pro, as long as the estimation was possible for the structure. The units of the melting and boiling points are degrees Celsius, whereas the unit of the vapour pressure is pascals.

The most stable conformer, which has the lowest energy, was selected for each molecule by conformational analysis. The size (Å) of the most stable conformer for each molecule was measured using the distance between the two atoms that were the furthest from each other within a molecule. The sizes of the molecules that were adopted for further optimization by PM3 were also measured in the same way.

Creation of the database

All the data obtained by calculation and by analysis of the molecules was accumulated and the database was constructed by CS ChemFinder Pro (version 5.0, CambridgeSoft Corporation). Their molecular formulae and molecular weights (mol. wts) were automatically calculated using CS ChemFinder Pro. The expression of the database is shown in Figure 2. The description 'estim' in the database represents estimation and 'multi' represents multiple bonds. Descriptions of odour quality and the detection threshold values were inserted into the database when this information was available in the data source. Descriptions of odour quality were translated into English when they were written in Japanese. The safety of the molecules in the database was also searched from other sources (Indo, 1996; Kanto Chemical Co. Inc., 2000) and included in the database, but their sources are not referred to in the database.

Results

The number of odorous molecules selected from the data sources totalled 287 compounds.

The features of the 287 molecules in the database

The atomic compositions of the 287 molecules are shown in Figure 3. More than half of the compounds in the database were made of C, H and O atoms. The atomic compositions of 46 molecules classified as 'other' in Figure 3 were either C plus H plus N plus O, C plus H plus N plus S, C plus H plus O plus S or C plus H plus O plus N plus S. Substructures of the 287 molecules were analysed. More than half of the compounds had one ring structure, as shown in Table 1. More than one-third of the molecules had a carbonyl group in their structures, as shown in Table 2, although some molecules had a multiple number of functional groups. The molecules' molecular weights, log *P* values, melting and boiling points and vapour pressures are listed in Table 3. The lightest molecule was ethanol (mol. wt = 46) and the heaviest one was 1,1,3,3,5-pentamethyl-4,6-dinitroindane, musk moskene (mol. wt = 278). Calculation of log *P* was not possible for the two isocyanides in the database using CS Chem3D Pro. The smallest value of log *P* was –1.6 for 3-hydroxy-2-methyl-4H-pyran-4-one (maltol) and the largest one was 6.07 for 4-acetyl-6-*t*-butyl-1,1-dimethylindane (celestolide). Only 16 molecules showed negative values of log *P*. The estimation of vapour pressure using CS Chem3D Pro was only possible for 74 molecules in the database. Supplementation by actual data or by another estimation method was particularly necessary in order to complete the vapour pressure values. The number of conformers found for each molecule in the database is shown in Figure 4. Two conformers were manually generated in the case of allylcarbylamine, since conformational analysis of the amine was not possible using Conformer®. Seventy-two of the 287 molecules had only one conformer for one molecule by conformational analysis. The 72 molecules were considered to be rather rigid. The range of the size of the most stable conformers for each of the 287 molecules was between 3.93 and 12.0 Å. The number of most stable conformers of the 287 molecules, which fall into a specific size, is shown in Figure 5. It reveals that more than 240 most stable conformers out of the 287 were within 5–9 Å length.

Descriptions of odour quality for 168 of the compounds in the database were available in the data sources. More than 100 different words were used for describing the odour qualities of the molecules, even though overlapped words were excluded. Odour detection threshold values were available in the data sources for 176 molecules (Devos *et al.*, 1990). The lowest detection threshold value was 4.54 for ethanol and the highest was 10.5 for 2,3-benzopyrrole (indole), although accumulation of odour detection threshold values is not complete in the database.

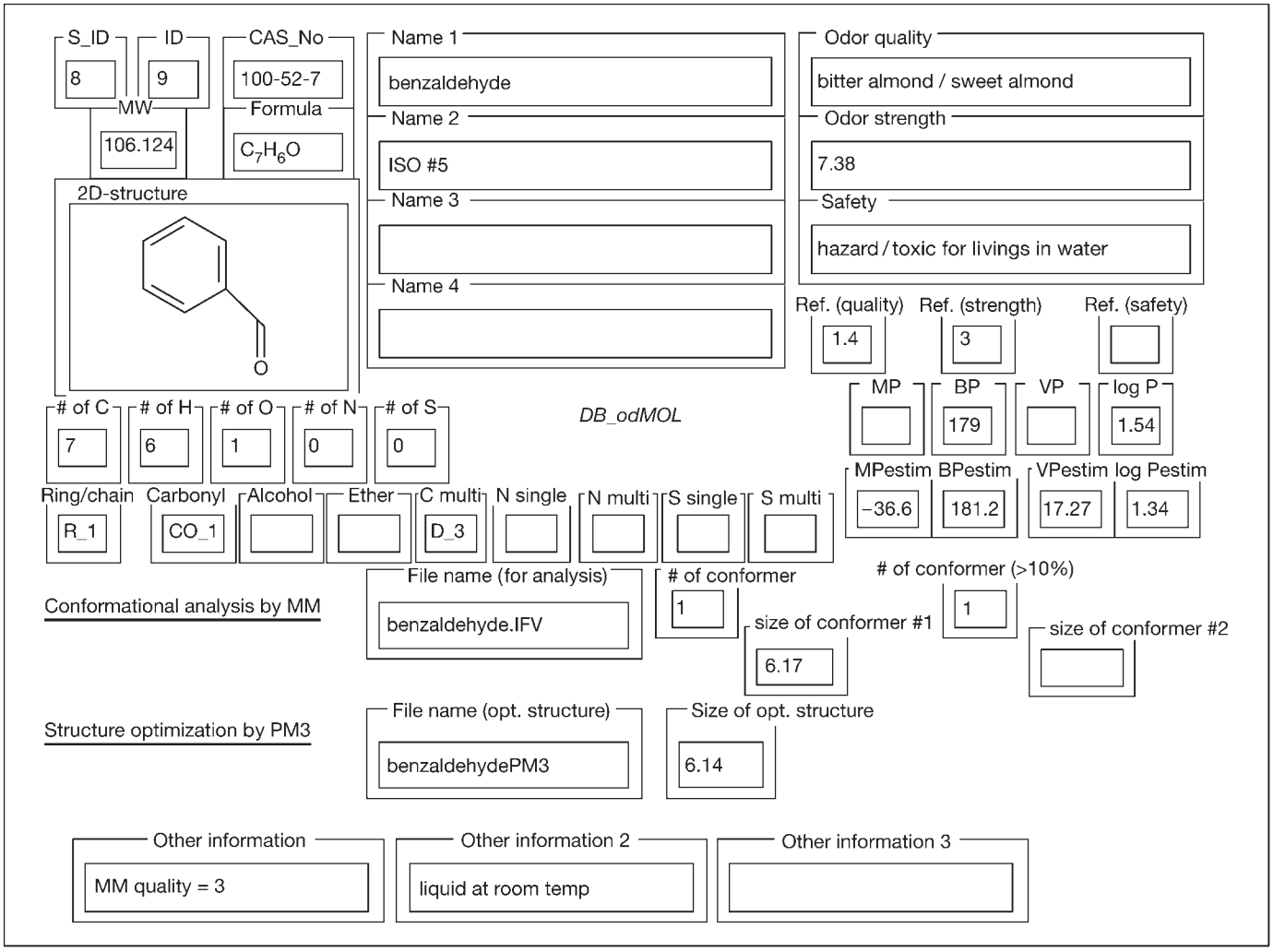

Figure 2 Expression of a molecule in the database: example of benzaldehyde.

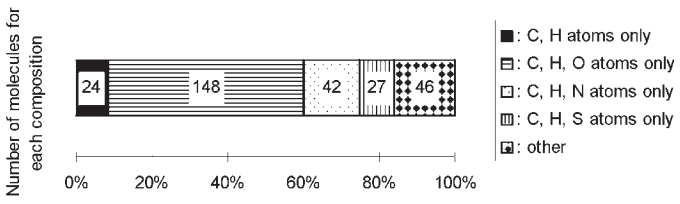

Figure 3 Atomic compositions of the molecules in the database.

Table 2 Functional groups concerning oxygen atoms of the odorous molecules in the database

Functional group	No. of molecules
CO	118
COC	80
OH	52

CO, carbonyl group; COC, ether group; OH, hydroxyl group.

Table 1 Number of ring structures of the odorous molecules in the database

No. of ring structures	No. of molecules
0	90
1	165
2	31
3	1
Total	287

Table 3 Molecular weights, log *P* values, melting and boiling points and vapour pressures of compounds in the database

	Lowest value	Highest value
Range of molecular weight	46.1	278.3
Range of log *P* value	−1.6	6.07
Range of melting point(°C)	−225	178
Range of boiling point (°C)	38	484
Range of vapour pressure (Pa)	0.00152	6974.33

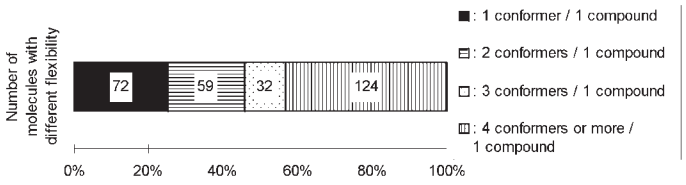

Figure 4 Rigidity of the molecules by conformational analysis: the number of conformers found for each molecule.

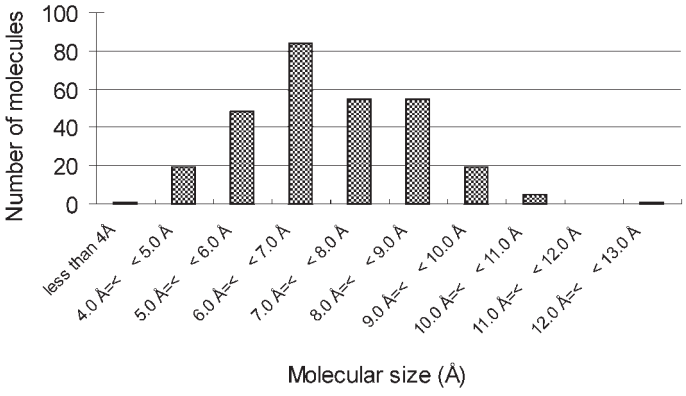

Figure 5 Size variations of the most stable conformers of the 287 molecules in the database.

Selection of odorous molecules by atomic composition and rigidity

Sixty-two rigid molecules were further selected from the 72 rigid molecules according to their atomic composition. The structures of the 62 molecules listed in Figure 6 have simple atomic compositions (C plus H only, C plus H plus O only, C plus H plus N only or C plus H plus S only). The names, molecular weights, sizes, odour qualities, odour strengths, boiling points and vapour pressures of the molecules in Figure 6 are listed in Table 4 from *DB_odMOL*. These 62 compounds cover from 4 to 10 Å in size and cover from 54 to 178 in molecular weight.

Conclusion

Structurally rigid molecules are assumed to interact with a small number of types of olfactory receptor protein or to interact with the proteins in a simpler manner than that of fairly flexible molecules. Two hundred and eighty-seven rather rigid odorous molecules were compiled in the database. The selection of odorants was mainly based on the rigidity of their molecular structures. The atomic compositions, substructures (ring structure, functional group, etc.), rigidities and sizes of the molecules were analysed and compiled into the database. Other information relating to their physicochemical parameters, such as log *P* values, melting and boiling points and vapour pressures, along with odour quality, odour intensity and safety, were also included

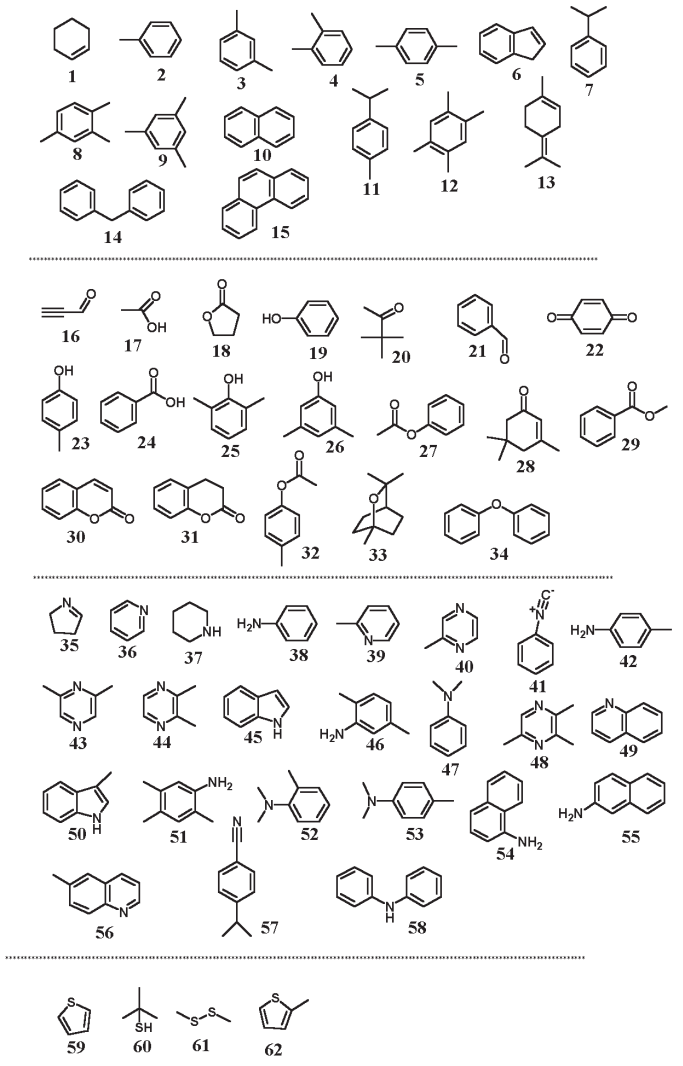

Figure 6 Two-dimensional structures of the 62 rigid molecules of simple atomic composition.

in the database by estimation using computer calculation or the data sources used for this work. This database, namely *DB_odMOL*, should be useful for researchers in choosing odorants for the study of olfaction based on the substructures, sizes and other parameters of the odorous molecules. *DB_odMOL* also helps other researchers in the study of the relation between odour qualities and molecular properties. We are planning to provide *DB_odMOL* in the form of a compact disk or by the Internet. Further accumulation of information about odorous molecules, some correction by experimental data or prediction data and supplementation to the odour qualities, intensity, safety and some other physicochemical properties are in progress.

The 62 rigid molecules (Figure 6) thus selected by conformational analysis and atomic composition were proposed as the first choice of odorous molecules for the study of olfaction, although further selection is necessary according to the odour qualities and safety of the molecules. The 62

Table 4 Names, molecular weights, sizes, log *P* values, odour qualities, odour strengths, boiling points and vapour pressures of the 62 molecules in Figure 6

No.	Name	Mol. wt	Size (Å)	Log *P* value	Odour quality	Odour strength	Boiling point (°C)	Vapour pressure (Pa)
1	Cyclohexene	82.1	5.03	2.02		6.44	93.0	1208.6
2	Methylbenzene (toluene)	92.1	5.94	2.28	paint		125.7	386.1
3	1,3-Dimethylbenzene	106.2	6.54	2.70	geranium	6.49	148.3	89.9
4	1,2-Dimethylbenzene	106.2	6.04	2.70		6.07	148.3	89.9
5	1,4-Dimethylbenzene	106.2	6.90	2.70		6.31	148.3	120.2
6	Indene	116.2	6.75	2.36		8.06	182.4	
7	Isopropylbenzene (cumene)	120.2	7.25	2.94		7.62	157.1	61.2
8	1,2,4-Trimethylbenzene	120.2	6.99	3.11		6.81	170.0	28.6
9	1,3,5-Trimethylbenzene	120.2	6.82	3.11		6.64	170.0	34.0
10	Naphthalene	128.2	7.23	2.99	tar	7.83	***218***	1.2
11	1-Isopropyl-4-methylbenzene	134.2	8.23	3.36	solvent/weak citrus	8.67	178.4	
12	1,2,4,5-Tetramethylbenzene	134.2	6.91	3.53		7.58	***197***	
13	Terpinolene	136.2	8.17	1.72	piney		178.2	
14	Diphenylmethane	168.2	8.83	***4.20***	green, geranium-like	7.63	***262***	0.2
15	Phenanthrene	178.2	9.33	4.11		8.11	327.3	0.0
16	Propynal	54.0	4.71	–0.19		7.01	68.7	
17	Acetic acid	60.1	4.02	–0.07	sour	6.84	122.3	213.9
18	γ-Butyrolactone	86.1	4.49	0.34	sweet weak/caramel		177.0	
19	Phenol	94.1	5.65	1.48		6.96	170.1	4.8
20	3,3-Dimethylbutan-2-one	100.2	5.66	1.49		6.73	95.1	
21	Benzaldehyde	106.1	6.17	***1.54***	bitter almond/sweet almond	7.38	***179***	17.3
22	1,4-Benzoquinone	108.1	5.34	-0.41		7.97	218.1	
23	4-Methylphenol	108.1	6.61	1.89	medicinal	8.73	190.8	1.5
24	Benzoic acid	122.1	7.09	1.22	sweet balsamic/urine		***249***	0.0
25	2,6-Dimethylphenol	122.2	6.81	2.31		9.12	210.7	1.4
26	3,5-Dimethylphenol	122.2	6.53	2.31		9.81	***483.9***	
27	Phenyl acetate	136.2	7.66	1.60		6.98	196.0	
28	3,5,5-Trimethyl-2-cyclohexen-1-one	138.2	6.89	1.87		6.2	201.5	33.5
29	Methyl benzoate	136.2	7.24	1.69	ylang ylang-like/dry-fruity/prune	6.97	***200***	5.3
30	2H-1-Benzopyran-2-one (coumarin)	146.1	7.30	***1.95***	powdery/sweet balsamic	9.14	290.8	
31	3,4-Dihydro-2H-benzopyrane (dihydrocoumarin)	148.2	7.10	***1.73***	sweet/sweet herbal		288.8	
32	4-Methylphenyl acetate	150.2	8.71	2.02	narcis-like as diluted	9.11	***212***	
33	1,8-Epoxy-*p*-menthane (1,8-cineole)	154.3	6.49	***2.31***	camphor-like/peppermint	7.79	***176***	
34	Diphenyl oxide	170.2	8.87	***4.28***	jasmine-like floral/geranium like, floral	8.15	***216***	0.3
35	3,4-Dihydro-2H-pyrrole (1-pyrroline)	69.1	4.31	0.03		7.61	81.9	
36	Pyridine (azabenzene)	79.1	4.33	0.75		7.07	113.4	282.2
37	Piperidine (hexahydropyridine)	85.1	4.97	0.76		6.43	127.9	409.2
38	Aminobenzene (aniline)	93.1	5.87	1.05		6.17	184.0	6.7
39	2-Methylpyridine	93.1	5.91	1.40		7.37	136.4	152.3
40	2-Methylpyrazine	94.1	5.80	0.29	popcorn		146.9	
41	Phenyl isocyanide	103.1	6.52			9.12		
42	4-Methylaniline	107.2	6.93	1.47		6.48	204.2	
43	2,6-Dimethylpyrazine	108.1	6.63	0.75	sweet roast fried potato/nutty		168.7	
44	2,3-Dimethylpyrazine	108.1	5.84	0.94	roasted hazelnuts like		***156***	
45	2,3-Benzopyrrole (indole)	117.2	6.69	***2.07***	animal/floral as diluted/mothballs	10.5	***254***	
46	2,5-Dimethylaniline	121.2	7.06	1.88		7.66	223.4	
47	*N*,*N*-Dimethylaniline	121.2	7.27	2.70		7.66	169.4	9.5
48	Trimethylpyrazine	122.2	6.81	1.40	cocoa, coffee-like/roasted cocoa		***171***	
49	Quinoline (1-azanaphthalene)	129.2	7.19	2.12		7.83	240.2	0.8
50	Skatole (3-methyl-1H-indole)	131.2	7.40	***2.59***	animal-like/floral as diluted/mothballs	9.27	***226***	
51	2,4,5-Trimethylaniline	135.2	6.90	2.30		7.34	241.8	

Table 4 Continued

No.	Name	Mol. wt	Size (Å)	Log *P* value	Odour quality	Odour strength	Boiling point (°C)	Vapour pressure (Pa)
52	*N*,*N*-2-Trimethylaniline	135.2	7.21	3.12		7.26	190.2	
53	*N*,*N*-4-Trimethylaniline	135.2	8.28	3.12		7.21	190.2	
54	1-Aminonaphthalene	143.2	7.25	2.17		7.47	293.2	
55	2-Aminonaphthalene	143.2	8.21	2.17		6.56	293.2	
56	6-Methylquinoline	143.2	8.06	***2.50***	tobacco, civet-like, floral as diluted		***259***	
57	4-Isopropylbenzonitrile	145.2	8.75	3.30	cumin, spicy-like		254.8	
58	Diphenylamine	169.2	10.23	***3.34***		7.65	***302***	0.0
59	Thiophene	84.1	4.61	1.96	sulphur/garlic	6.48	114.0	1083.1
60	2-Methylpropane-2-thiol	90.2	4.59	1.43		9.48	***64***	
61	Dimethyl disulfide	94.2	5.72	***1***	onion	7.91	113.7	
62	2-Methylthiophene	98.2	5.37	2.30	sulphur		137.0	

The log *P* and boiling point values in bold and italic are experimental values. Odour strength: standard detection threshold value in the gas phase, *p*. ol (volume), from data compiled by Devos *et al*. (Devos *et al*., 1990).

molecules are candidates for standard odorous molecules in general studies relating to olfaction.

Acknowledgement

This work was supported by the Kayamori Informational Technology Foundation.

References

Acree, T. and **Arn, H.** (1997) *Flavornet*. In Acree, T. and Arn, H. (eds), Gas Chromatography–Olfactometry (GCO) of Natural Products.

Boelens, M.H., Boelens, H. and **Van Gemert, L.J.** (1993) *Sensory properties of optical isomers*. Perfume Flavor., 18, 1–16.

Buchbauer, G., Hillisch, A., Mraz, K. and **Wolschann, P.** (1994) *Conformational parameters of the sandalwood-odor activity: conformational calculations on sandalwood odor*. Helv. Chim. Acta, 77, 2286–2296.

Devos, M., Patte, F., Rouault, J., Laffort, P. and **Van Gemert, L.J.** (eds) (1990) Standardized Human Olfactory Thresholds. IRL Press, Oxford.

Gorbachov, M.Y. and **Rossiter, K.J.** (1999) *A new electronic-topological investigation of the relationship between chemical structure and ambergris odour*. Chem. Senses, 24, 171–178.

Indo, M. (ed.) (1996) Synthetic Perfume and Flavor Materials—Chemistry and Products Information. The Chemical Daily Co., Ltd, Tokyo (in Japanese).

Kanto Chemical Co. Inc. (2000) General Catalog No. 21 (2000). Kanto Chemical Co. Inc., Tokyo.

Laska, M. and **Teubner, P.** (1999) *Olfactory discrimination ability of human subjects for ten pairs of enantiomers*. Chem. Senses, 24, 161–170.

Nakajima, M. (ed.) (1995) Fundamentals of Perfumery—Practice and Creation. Sangyo Tosho, Tokyo (in Japanese).

Rognan, C. and **Chastrette, M.** (1994) *Structure–odor relationships: a highly predictive tridimensional interaction model for the bell-pepper note*. Eur. J. Med. Chem., 29, 595–609.

Yoshii, F. and **Hirono, S.** (1996) *Construction of a quantitative three-dimensional model for odor quality using comparative molecular field analysis (CoMFA)*. Chem. Senses, 21, 201–210.

Accepted January 7, 2002