A Model for Odour Thresholds

Michael H. Abraham, Joëlle M.R. Gola, J. Enrique Cometto-Muniz[1] and William S. Cain[1]

Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK and [1]Chemosensory Perception Laboratory, Department of Surgery (Otolaryngology), University of California, San Diego, Mail Code 0957, La Jolla, CA 92093-0957, USA

Correspondence to be sent to: Michael H. Abraham, Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK. e-mail: m.h.abraham@ucl.ac.uk
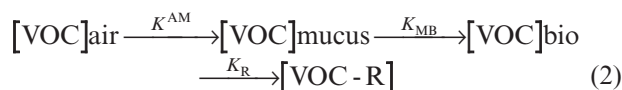
Abstract

Odour detection thresholds, that we have previously obtained, have been analysed by a general equation for selective transport. It is shown that such selective transport can account for some 77% of the total effect. The remainder is due to a specific size effect, that might involve odour-binding proteins, and a specific effect for aldehydes and carboxylic acids. Our analysis raises the question of whether selective transport is physically separable from the specific effects of receptor activation. The model predicts a chemical cut-off in odour detection along any homologous series.

Introduction

There have been a number of correlations of odour detection thresholds (ODT) with various properties of odorants, the study by Laffort and Patte (Laffort and Patte, 1987) being one of the first to employ a physicochemical analysis. Chastrette (Chastrette, 1997) has reviewed work up to 1996; many studies involved sets of odorants of similar structure, and none led to any conclusions of mechanistic significance. Two subsequent studies related ODT to properties of homologous series of odorants. Yamanaka (Yamanaka, 1995) showed that odour thresholds of Devos *et al.* (Devos *et al.*, 1990) for several homologous series could be correlated with the odorant activity coefficient in water, γ_W, through a set of equations of the type:

$$\log(1/\text{ODT}) = a \log \gamma_W + b \qquad (1)$$

where a and b differ for each homologous series. A much more detailed analysis was carried out by Hau and Connell (Hau and Connell, 1998) who used the following representation of a possible mechanism:

$$[\text{VOC}]\text{air} \xrightarrow{K^{AM}} [\text{VOC}]\text{mucus} \xrightarrow{K_{MB}} [\text{VOC}]\text{bio}$$
$$\xrightarrow{K_R} [\text{VOC-R}] \qquad (2)$$

where [VOC]air is the concentration of an odorant, or volatile organic compound (VOC) in air, [VOC]mucus is the concentration in the mucus, [VOC]bio is the VOC concentration in the biophase that contains the olfactory receptor, and VOC-R is the concentration of the VOC–receptor complex. The equilibrium constants for the three stages are denoted here as K^{AM}, K_{MB} and K_R. From equation (2), a connection between ODT values and VOC partition coefficients was deduced as:

$$\log [\{\text{ODT}\}K^W] = -a \log \text{Poct} + b \qquad (3)$$

where K^W is the air–water partition coefficient, also known as the Ostwald solubility coefficient, and Poct is the water–octanol partition coefficient. The ODT values were from the AIHA compilation (AIHA, 1989); the coefficients a and b vary from one homologous series to another. The interpretation of equation (3) was that K^W is an approximation for K^{AM}, as both refer to the equilibrium between the gas phase and an aqueous condensed phase, and that K_{MB} and K_R are both functions of Poct.

Both equations (1) and (3) suffer from shortcomings as predictive equations, in that only homologous series can be considered. This excludes numerous types of important VOCs such as inhalation anaesthetics and terpenes that do not fall into any homologous series. The model of Hau and Connell (Hau and Connell, 1998) is significant, however, because it is the only real attempt to correlate ODT values on any mechanistic basis.

There have been studies using sets of varied structural types of VOCs, rather than restriction to homologous series. Dravnieks (Dravnieks, 1974) correlated four sets of threshold data of vapours, using various structural features as the independent variables, but results were not very good, with r^2 ranging from 0.42 to 0.58 with four independent variables. Such methods may be useful as empirical correlations, but yield little mechanistic information.

In order to investigate odour thresholds in more detail, it is important to understand the way in which olfactory perception is processed, via the relationship between odour stimuli and the receptive surface (Pearce *et al.*, 1998). Once in the airspace above the olfactory mucosa, the molecules must diffuse through a layer of mucus (10–30 μm thick) to gain final access to the receptors themselves (Hornung and Mozell, 1981; Snyder *et al.*, 1988). Such diffusion, or transport, may involve (at least in part) odorant binding proteins (OBPs) that can act as carriers (Bianchet *et al.*, 1996; Brownlow and Sawyer, 1996; Tegoni *et al.*, 1996; Löbel *et al.*, 2001). The central pocket in the OBP has dimensions of $11 \times 10 \times 7$ Å (i.e. 770 Å^3) with an opening size of 6×7 Å (Tegoni *et al.*, 1996), although a much larger cavity of 1100–1300 Å^3 has been suggested (Bianchet *et al.*, 1996). Once transported across the mucosal layer to a receptor area or biophase, the VOC (or the VOC/OBP complex) can then interact with odour receptors at the surface of the cilia membrane of the olfactory neuron. The actual binding pocket in the rat OR5 receptor, however, is no less than 12 Å from the extracellular surface of the receptor (Singer and Shepherd, 1994). A general model that we suggest is shown in Figure 1. It is useful to consider two types of interaction. Simple transport processes are selective, in that different VOCs will have different equilibrium constants, depending on their structure. However, small changes in structure or small positional changes of functional groups have rather small effects on such processes. On the other hand, in processes such as ligand/receptor interactions, small changes in structure can have very large effects; we refer to these processes as having specific effects. In Figure 1 we indicate which processes may be selective and those that may be specific in nature.

Whether the VOC/OBP interactions and the VOC/R interactions are general interactions that can be modelled by a physicochemical transport process, or whether they are more specific interactions, is a crucial point. The analysis of Hau and Connell (Hau and Connell, 1998) certainly supposes that the VOC/R interaction is a general interaction that can be modelled by a simple physicochemical descriptor, such as log Poct.

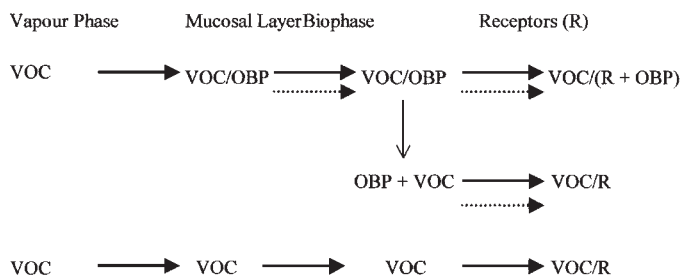Our approach is first to use a model that simply reflects a passive physicochemical transport property. Comparison with physicochemical transport to various solvents or to various biophases will then indicate whether or not such passive transport can model all or part of the odour detection process.

Methodology

We have devised a very general equation for the correlation of a variety of processes in which VOCs are transferred from the gas phase to some condensed phase (Abraham *et al.*, 1991; Abraham, 1993):

$$\log \text{SP} = c + e\mathbf{E} + s\mathbf{S} + a\mathbf{A} + b\mathbf{B} + l\mathbf{L} \qquad (4)$$

The dependent variable, log SP, is some property of a series of VOCs in a given system. The independent variables in equation (4) are (Abraham, 1993) properties of the VOCs. We use a simplified nomenclature, with the original nomenclature in parentheses: $\mathbf{E}$ (R_2) is an excess molar refraction, $\mathbf{S}$ (π_2^H) is the dipolarity/polarizability, $\mathbf{A}$ ($\Sigma\alpha_2^H$) and $\mathbf{B}$ ($\Sigma\beta_2^H$) are the overall or effective hydrogen-bond acidity and basicity, and $\mathbf{L}$ (log L^{16}) is defined through L^{16}, the solute Ostwald solubility coefficient on hexadecane at 298K. The $\mathbf{L}$-descriptor is itself a combination of two solute properties: (i) a general measure of solute size, and (ii) the ability of a solute to interact with a solvent phase through dispersion forces. The units of $\mathbf{E}$ are cm^3/10; the other descriptors have no units because they are all derived from the logarithm of an equilibrium constant. The coefficients c, e, s, a, b and l are found by multiple linear regression analysis. They reflect the complementary properties of the receptor phase. The e-coefficient gives the tendency of the phase to interact with VOCs through polarizability-type interactions, mostly via electron pairs. The s-coefficient is a measure of the phase dipolarity/polarizability. The a-coefficient represents the complementary property to VOC hydrogen-bond acidity and so is a measure of the phase hydrogen-bond basicity. Likewise, the b-coefficient is a measure of the phase hydrogen-bond acidity. Finally, the l-coefficient is a measure of the hydrophobicity of the phase. Equation (4) has been applied to numerous gas–solvent partitions (Abraham *et al.*, 1994b, 1998b, 1999a,b), to gas–biophase partitions (Abraham and Weathersby, 1994), and to a very large number of gas chromatographic systems (Abraham *et al.*, 1999c), so it is a well tried and tested equation.

We have previously used equation (4) to correlate nasal pungency threshold values (NPT, in p.p.m.) for 43 varied compounds (Abraham *et al.*, 1998a), resulting in equation (5):

$$\log(1/\text{NPT}) = -8.519 + 2.154\mathbf{S} + 3.522\mathbf{A} + 1.397\mathbf{B} + 0.860\mathbf{L} \qquad (5)$$

$$n = 43, \; r^2 = 0.955, \; \text{SD} = 0.27, \; F = 201$$

Figure 1   A possible model for odour thresholds. Selective processes: ➤; specific processes: ┄➤.

Here and elsewhere, n is the number of data points (i.e. the number of VOCs), r is the correlation coefficient, SD is the standard deviation in the dependent variable, and F is the F-statistic. The e-coefficient of the independent variable, **E**, was statistically not significant. The reciprocal of NPT values were used, so that the more potent the VOC the larger is the value of log(1/NPT).

The coefficients in equation (5) can be compared with those for various gas-condensed phase partitions that take place by simple transfer mechanisms, as shown in Table 1 (Abraham et al., 1994b, 1998b, 1999a,b). There is considerable similarity between the NPT equation and equations for the solubility of gaseous VOCs in solvents such as wet 1-octanol and methanol. There is also some similarity with equations for the solubility of gaseous VOCs in a number of biophases (Abraham and Weathersby, 1994). There is therefore nothing extraordinary about equation (5), which can be regarded as an equation for simple transfer of VOCs from the gas phase to a biophase. It is noteworthy that equation (5) encompasses a wide variety of VOCs, including carboxylic acids, aldehydes, ketones, alcohols, etc., with but one outlier—acetic acid. Following the analysis of Abraham et al. (Abraham et al., 1994a), equation (5) could be interpreted as arising from transport of VOCs to a biophase, followed by activation of a receptor through an 'on–off' mechanism that was independent of the structure of the VOC.

Our strategy is to apply the general equation (4) to ODT values, in the hope that we might deduce whether or not the resulting equation is consistent with simple transfer of VOCs from the gas phase to a biophase.

Results and discussion

General analysis

ODTs for a series of 64 compounds, including esters, aldehydes, ketones, alcohols, carboxylic acids, aromatic

hydrocarbons, terpenes and a number of other VOCs, have been determined previously (Cometto-Muniz and Cain, 1990, 1991, 1993, 1994, 1995) and by Cometto-Muniz et al. (Cometto-Muniz et al., 1998a,b), using a standardized protocol. This protocol entails direct measurement of vapour phase concentration of the VOCs for as many steps on the dilution series of each VOC as the sensitivity of an FID gas chromatographic detector (or sometimes a PID detector) allows; this is a rarity in olfactory research. The average standard deviation for all odour thresholds, expressed as log(1/ODT) is 0.63 log unit. The VOCs used in these studies are listed in Table 2, together with log(1/ODT) values, where ODT is in p.p.m. The corresponding VOC descriptors are given in Table 3. As a first step we applied equation (4) to all the VOCs except the carboxylic acids and aliphatic aldehydes that were clearly out of line. The VOCs, propanone, 1-octanol, methylacetate and t-butylacetate were then also revealed to be outliers, and were removed to yield the correlation equation,

$$\log(1/\text{ODT}) = -5.145 + 0.533\textbf{E} + 1.912\textbf{S} + 1.276\textbf{A} + 1.559\textbf{B} + 0.699\textbf{L}$$
$$\phantom{\log(1/\text{ODT}) = -5.145 +} 0.410 \quad 0.455 \quad\;\; 0.623 \quad\;\; 0.755 \quad\;\; 0.732 \quad\;\; 0.072 \tag{6}$$

$$n = 50,\ r^2 = 0.773,\ q^2 = 0.603,\ \text{SD} = 0.579,\ F = 28.7$$

Here, and elsewhere, q^2 is the coefficient of cross-validation, a useful measure of internal self-consistency. The SD values of the coefficients themselves are given below the coefficients. It is of considerable interest to compare the coefficients in the above equation with the coefficients for the other processes shown in Table 1. It can be seen that the coefficients in equation (6) are of the same sign and similar order of magnitude as those for transfer from the gas phase to organic solvents. For example, equation (6) compares well with the equations for gas/methanol or gas/wet 1-octanol, as well as with the NPT equation (5). Results of our analysis are therefore compatible with the possibility that simple transfer from the gas phase to a biophase might play a substantial part in the relationship of odour thresholds to the structure of VOCs, of the order of 77% of the total effect.

The closeness of equation (6) to equation (5), shows also that values of log(1/NPT) and log(1/ODT), except for the aldehydes and carboxylic acids, will be reasonably well correlated, which is indeed the case. Inspection of Table 1 also leads to the conclusion that the aqueous mucus layer that covers the olfactory epithelium does not influence the transport process, because the equation for gas/water transfer (Abraham et al., 1994b) is completely different to equation (6). The latter equation is also in agreement with the finding that the odour receptor binding pocket, at least for the OR5 receptor, is a considerable distance away from the extracellular surface of the receptor (Singer and Shepherd, 1994).

Table 1 Regression coefficients in equation (4) for gas–solvent (phase) partitions at 298K

Phase	e	s	a	b	l
Wet 1-octanol	0.002	0.709	3.519	1.429	0.858
Dry methanol	−0.215	1.173	3.701	1.432	0.769
Chloroform	−0.467	1.203	0.138	1.432	0.994
Acetone	−0.277	1.522	3.258	0.078	0.863
Dimethylformamide	−0.189	2.327	4.756	0.000	0.808
Water	0.822	2.743	3.904	4.814	−0.213
Brain[a]	0.427	0.286	2.781	2.787	0.609
Muscle[a]	0.544	0.216	3.471	2.924	0.578
Fat[a]	−0.172	0.729	1.747	0.219	0.895
Nasal pungency[a]	0.000	2.154	3.522	1.397	0.860
ODT, equation (6)[a]	0.533	1.912	1.276	1.559	0.699

[a]At 310K.

Table 2 Values of log(1/ODT) with ODT in p.p.m.

VOC	Log(1/ODT)	VOC	Log(1/ODT)
Methanol	−3.180	Hexylbenzene	0.190
Ethanol	−1.850	Heptylbenzene	0.250
1-Propanol	−1.150	Octylbenzene	0.430
2-Propanol	−2.700	Oct-1-ene	−2.310
1-Butanol	−0.300	Oct-1-yne	−2.130
2-Butanol	−1.980	Chlorobenzene	−1.110
2-Methy1-1-propanol	−2.780	2-Phenylethanol	2.190
1-Pentanol	−0.110	s-Butylacetate	−0.670
1-Hexanol	0.050	t-Butylacetate	−0.110
1-Heptanol	1.000	Butyraldehyde	−0.477
4-Heptanol	−0.910	Pentanal	−0.699
1-Octanol	2.150	Hexanal	1.097
Pyridine	−0.110	Heptanal	1.523
Methylacetate	−3.460	Octanal	2.398
Ethylacetate	−2.240	Formic acid	−0.886
Propylacetate	−1.390	Acetic acid	2.000
Butylacetate	−0.380	Butanoic acid	2.444
Pentylacetate	−0.070	Hexanoic acid	2.585
Hexylacetate	0.200	Octanoic acid	4.959
Heptylacetate	0.010	Menthol	1.660
Octylacetate	0.410	Cumene	−0.033
Decylacetate	0.500	*p*-Cymene	−0.121
Dodecylacetate	1.360	Δ-3-Carene	−0.223
Propanone	−4.070	Linalool	0.022
2-Pentanone	−0.930	1,8-Cineole	0.495
2-Heptanone	0.150	Geraniol	1.070
2-Nonanone	0.030	α-Terpinene	−0.152
Toluene	−2.190	γ-Terpinene	−0.992
Ethylbenzene	−1.260	α-Pinene	−1.277
Propylbenzene	−0.470	β-Pinene	−1.070
Butylbenzene	−0.630	(*R*)-(+)-limonene	−0.994
Pentylbenzene	−0.004	(*S*)-(+)-limonene	−0.659

Table 3 VOC parameters used in the present work

Solute	E	S	A	B	L	D (A)
Methanol	0.278	0.440	0.430	0.470	0.970	5.150
Ethanol	0.246	0.420	0.370	0.480	1.485	6.378
1-Propanol	0.236	0.420	0.370	0.480	2.031	7.649
2-Propanol	0.212	0.360	0.330	0.560	1.764	6.634
1-Butanol	0.224	0.420	0.370	0.480	2.601	8.882
2-Butanol	0.217	0.360	0.330	0.560	2.338	7.890
2-Methyl-1-propanol	0.180	0.300	0.310	0.600	1.963	6.638
1-Pentanol	0.219	0.420	0.370	0.480	3.106	10.146
1-Hexanol	0.210	0.420	0.370	0.480	3.610	11.396
1-Heptanol	0.211	0.420	0.370	0.480	4.115	12.654
4-Heptanol	0.180	0.360	0.330	0.560	3.850	11.650
1-Octanol	0.199	0.420	0.370	0.480	4.619	13.910
Pyridine	0.631	0.840	0.000	0.520	3.022	6.814
Methylacetate	0.142	0.640	0.000	0.450	1.911	7.650
Ethylacetate	0.106	0.620	0.000	0.450	2.314	8.870
Propylacetate	0.092	0.600	0.000	0.450	2.819	10.154
Butylacetate	0.071	0.600	0.000	0.450	3.353	11.340
Pentylacetate	0.067	0.600	0.000	0.450	3.844	12.760
Hexylacetate	0.056	0.600	0.000	0.450	4.351	13.880

Table 3   Continued

Solute	E	S	A	B	L	D (A)
Heptylacetate	0.050	0.600	0.000	0.450	4.865	15.150
Octylacetate	0.029	0.600	0.000	0.450	5.364	16.395
Decylacetate	0.033	0.600	0.000	0.450	6.373	18.940
Dodecylacetate	0.012	0.600	0.000	0.450	7.381	21.380
Propanone	0.179	0.700	0.040	0.490	1.696	6.612
2-Pentanone	0.143	0.680	0.000	0.510	2.755	9.110
2-Heptanone	0.123	0.680	0.000	0.510	3.760	11.610
2-Nonanone	0.119	0.680	0.000	0.510	4.735	14.120
Toluene	0.601	0.520	0.000	0.140	3.325	8.080
Ethylbenzene	0.613	0.510	0.000	0.150	3.778	9.303
Propylbenzene	0.604	0.500	0.000	0.150	4.230	10.124
Butylbenzene	0.600	0.510	0.000	0.150	4.730	11.650
Pentylbenzene	0.594	0.510	0.000	0.150	5.230	12.778
Hexylbenzene	0.591	0.500	0.000	0.150	5.720	14.080
Heptylbenzene	0.577	0.480	0.000	0.150	6.219	15.231
Octylbenzene	0.579	0.480	0.000	0.150	6.714	16.466
Oct-1-ene	0.094	0.080	0.000	0.070	3.568	12.808
Oct-1-yne	0.155	0.220	0.090	0.100	3.521	12.771
Chlorobenzene	0.718	0.650	0.000	0.070	3.657	8.360
2-Phenylethanol	0.811	0.910	0.300	0.640	4.628	10.090
s-Butylacetate	0.044	0.570	0.000	0.470	3.054	10.149
t-Butylacetate	0.025	0.540	0.000	0.470	2.802	8.943
Butanal	0.187	0.650	0.000	0.450	2.270	8.44
Pentanal	0.163	0.650	0.000	0.450	2.851	9.690
Hexanal	0.146	0.650	0.000	0.450	3.357	10.950
Heptanal	0.140	0.650	0.000	0.450	3.865	12.200
Octanal	0.160	0.650	0.000	0.450	4.361	13.460
Formic acid	0.300	0.790	0.720	0.340	1.400	5.260
Acetic acid	0.265	0.650	0.610	0.440	1.750	6.298
Butanoic acid	0.210	0.620	0.600	0.450	2.830	8.790
Hexanoic acid	0.174	0.600	0.600	0.450	3.920	10.290
Octanoic acid	0.150	0.600	0.600	0.450	5.000	13.800
Menthol	0.400	0.500	0.230	0.580	5.177	10.590
Cumene	0.602	0.490	0.000	0.160	4.084	9.300
p-Cymene	0.607	0.490	0.000	0.190	4.590	10.476
Δ-3-Carene	0.511	0.220	0.000	0.100	4.649	6.930
Linalool	0.398	0.550	0.200	0.670	4.794	12.749
1,8-Cineole	0.383	0.330	0.000	0.760	4.688	8.788
Geraniol	0.513	0.632	0.390	0.660	5.479	13.749
α-Terpinene	0.526	0.250	0.000	0.150	4.715	10.477
γ-Terpinene	0.497	0.320	0.000	0.200	4.815	10.499
α-Pinene	0.446	0.140	0.000	0.120	4.308	9.000
β-Pinene	0.530	0.240	0.000	0.190	4.394	8.828
(R)-(+)-Limonene	0.488	0.280	0.000	0.450	4.725	9.550
(S)-(+)-Limonene	0.488	0.280	0.000	0.450	4.725	9.550

In order to ascertain what other factors, as well as simple transport, influence the ODT values, it is instructive to plot the residuals in equation (6), i.e. [log(1/ODT)obs – log(1/ODT)calc] against the 'size' parameter, **L**. The residuals are not random, and both small VOCs and large VOCs are less potent than expected. An even more informative plot, shown in Figure 2, is of the residuals versus the maximum length, **D**, of the VOC. The latter was obtained by means of a computer-assisted molecular-modelling program (Molecular Modeling Pro, 1992). The maximum value for **D** in a VOC was obtained after geometry optimization. In Figure 2, only the residuals for some homologous series are given, for clarity. It can then be seen that the residuals follow a 'parabolic-like' curve: as molecular size increases, the residual value increases to a maximum value and then decreases.
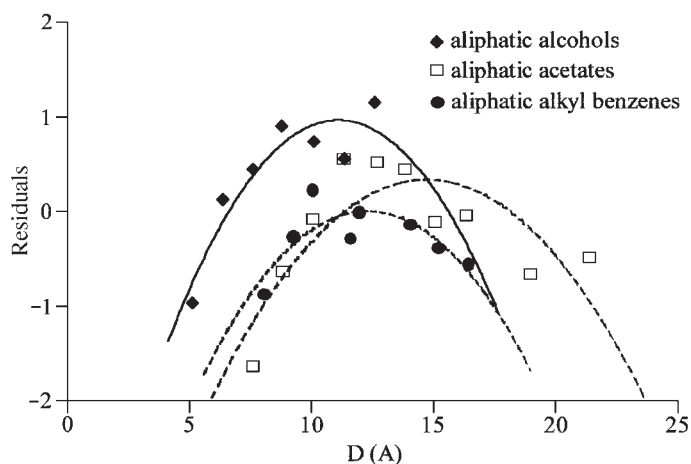
Figure 2   Residuals (observed – calculated values in equation 6) against the VOC maximum length.

We suggest that the pattern of residuals in Figure 2 is due to an extra effect, in addition to simple transfer. The effect can be quantified and incorporated into an equation for log(1/ODT) through addition of a parabolic term in $(D – D^2)$:

$$\log(1/ODT) = -6.757 + 0.533E + 1.912S + 1.276A + 1.559B + 0.699L + 0.297D – 0.013D^2 \quad (7)$$

$$n = 50, r^2 = 0.82, SD = 0.511$$

The statistics of equation (7) are quite good, bearing in mind the experimental error in the ODT values. We can include the carboxylic acids and aliphatic aldehydes into the regression equation by means of an indicator variable, **H**, chosen as 2.0 for the carboxylic acids and aldehydes and zero for all other VOCs,

$$\log(1/ODT) = -7.445 + 0.304E + 1.652S + 2.104A + 1.500B + 0.822L + 0.369D – 0.016D^2 + 1.000H \quad (8)$$

$$n = 60, r^2 = 0.84, SD = 0.601$$

Equation (8) is a general equation for log(1/ODT) values, and could be used to predict further values to ~0.6 log units (see later), of the order of experimental error. Four compounds are again outliers to equation (8), namely, propanone, methylacetate, t-butylacetate and 1-octanol. We shall use equation (8) as the basis of our model of odour detection, but suggest that an alternative predictive equation can be constructed by using a parabolic term in **L**, rather than in **D**:

$$\log(1/ODT) = -7.720 – 0.060E + 2.080S + 2.829A + 1.139B + 2.028L – 0.148L^2 + 1.000H \quad (9)$$

$$n = 60, r^2 = 0.85, q^2 = 0.536, SD = 0.598, F = 44$$

In equation (9) the indicator variable for aldehydes and carboxylic acids takes the value $H = 1.6$. The advantage of equation (9) over equation (8) is that it is not necessary to obtain the maximum length, **D**, in order to predict further values of log(1/ODT).

The necessity for the use of an indicator variable for aldehydes and carboxylic acids arises because these two sets of compounds are more potent than predicted by equation (7). There is precedent for the extra potency of aldehydes and carboxylic acids. Alarie *et al.* (Alarie *et al.*, 1998) have shown that these compounds are more potent than expected in sensory irritation in mice, and suggest that they undergo some actual chemical reaction. However, aldehydes and carboxylic acids (except acetic acid) fit our equation for nasal pungency thresholds (Abraham *et al.* 1998a) without use of any indicator variable, see equation (5). There is also the problem of the four outliers, propanone, methylacetate, t-butylacetate and 1-octanol. There may be extra experimental error with the first three compounds. Loss of propanone and methylacetate due to their high volatility would result in the compounds appearing to be of lower potency. In the case of t-butylacetate, the compound seemed to form an emulsion in some experiments, and this would result in an erroneous estimation of the ODT value. However, we have no explanation for the increased potency of 1-octanol.

Very recently, the EVA spectral descriptor has been applied to a selection of ODT values (Turner and Willett, 2000). No details were given other than for 52 ODT values, q^2 was 0.57 and for 44 log ODT values q^2 was 0.71; unfortunately EVA results cannot be interpreted in any chemical way and so cannot lead to any mechanistic conclusions.

Predictive capability

The statistics given for the various equations in log(1/ODT) do not lead to any assessment of their predictive capability, but only of their correlative ability. One method of estimating the predictive power of an equation is to divide total set of data into a training set and a test set. The training set is used to develop a correlation equation that in turn is used to predict the values for the test set. As the latter values have not been used to set up the correlation equation, a comparison of predicted and observed values for the test set is a very useful guide to the predictive power of the training equation. Equations (7), (8) and (9) cannot be studied in this way, because the parabolic terms have been imposed and are not the result of a straightforward correlation. However, equation (6) is an example of a multiple correlation, and so we have used this equation as an example. In order to have sufficient data points to construct a correlation equation for the training set, we used 38 points for the training set and 12 for a test set. It is important

Table 4   A summary of the training set correlations

Set no.	c	e	s	a	b	l	r^2	SD
1(ODT)	−5.331	0.299	2.144	2.135	1.018	0.753	0.772	0.615
2(E)	−5.247	0.659	1.402	1.435	2.284	0.702	0.816	0.553
3(S)	−5.123	0.785	2.010	1.000	1.611	0.646	0.804	0.529
4(A)	−5.284	0.709	1.659	0.722	2.122	0.716	0.793	0.580
5(B)	−5.482	0.452	1.881	1.762	1.493	0.800	0.761	0.622
6(L)	−5.311	0.533	1.656	1.346	1.858	0.728	0.813	0.562
Av	−5.296	0.572	1.792	1.400	1.731	0.724	0.793	0.577

that the test set is a representative sample of the entire set. We listed the 50 compounds in order of increasing values of log(1/ODT) and then selected every fourth compound as a member of the test set, leaving 38 compounds as the training set; we refer to this training/test set as 1(ODT). We then listed the 50 compounds in order of the dependent variable **E**, and chose every fourth compound as a member of a new test set, again leaving 38 compounds as a training set; the new training/test set is denoted as 2(**E**). A similar process was used to obtain training/test sets by ordering compounds by the other independent variables. This gave six different training/test sets.

A summary of the statistics for the six 38-compound training sets is in Table 4, and a comparison of the predicted (pred) and observed (obs) values of log(1/ODT) for the 12-compound test sets is given in Table 5. We give the usual standard deviation as

$$\sqrt{\{\Sigma[(obs) - (pred)]/(n - 1)\}}$$

where $n = 12$. The average deviation is

$$AD = \Sigma[(obs) - (pred)/12]$$

and the average absolute deviation is

$$AAD = |\Sigma[(obs) - (pred)/12]|$$

Also in Tables 4 and 5 are the average values of the various coefficients and statistics for the training and test sets. The six training sets have somewhat different coefficients and statistics to the correlation equation (6), but the average values are within any statistical error the same as those for equation (6). This can be seen from the SD values for the coefficients given in equation (6). The various training sets predict values of log(1/ODT) with an average SD value of 0.608 log units, as compared with the correlation SD of 0.579 log units, that is only 0.029 units higher. We can therefore take the value of 0.608 as a measure of the predictive capability of equation (6). In Table 5 are listed also values of AD and AAD. The former is negligible, at 0.006 log units, and shows that there is no bias towards too

Table 5   Comparison of predicted and observed log(1/ODT) values for test sets

Set no.	SD	AD	AAD
1(ODT)	0.537	0.166	0.455
2(E)	0.577	−0.190	0.577
3(S)	0.754	0.252	0.630
4(A)	0.628	−0.109	0.512
5(B)	0.487	−0.265	0.542
6(L)	0.670	0.183	0.555
Av.	0.608	0.006	0.545

positive or too negative predicted values. The AAD values are always less than the SD values, and simply provide another estimate of predictive capability.

As mentioned above, we cannot apply the training/test set method to estimate the predictive capability of equations (7), (8) and (9), but we think it reasonable to assign estimates as ~0.03 log units higher than the correlation SD values.

A model of odour detection

Equation (8) is not only a predictive equation, but can be considered to be compatible with the model shown in Figure 1. A large part of the variation in log(1/ODT) values with the structure of the VOCs is due to simple transport of the VOC from the gas phase to a biophase. In addition, there is an effect that we suggest is due to the size of the VOC, specifically to the maximum length [two recent papers have stressed the importance of molecular length as a factor in odour recognition (Araneda *et al.*, 2000; Johnson and Leon, 2000)]. The potency of VOCs in an homologous series has a maximum deviation from the simple transport equation (6) when the VOC has a maximum length of ~11–12 Å. Now this length is almost the same as the maximum dimension of the central pocket in OBPs, namely, 11 Å (Tegoni *et al.*, 1996); the alternative volume of Bianchet *et al.* (Bianchet *et al.*, 1996) suggests a maximum length of the central pocket of 12–13 Å. Thus one possible mechanism includes simple transfer from the gas phase to a biophase mediated by transport by OBPs. The exceptions are aldehydes and carboxylic acids that are more potent than calculated by

about a factor of 100. We do not suggest that there is only one OBP or even one type of OBP; there may be several types with maximum dimensions ~10–15 Å.

Of course, the above is not the only mechanism that fits our data analysis. It is possible that the OBPs have no discrimination at all, and that the 'maximum length' effect takes place on activation of the receptor. In any event, we do suggest that at least two types of interaction contribute to the overall threshold effect.

We can obtain some information as to the role of OBPs from recent work (Vincent *et al.*, 2000) in which complexation constants for a number of VOCs with porcine OBP were obtained. Details are in Table 6, with the complexation constants given as $\log(1/IC_{50})$. Over the seven VOCs studied, values of $\log(1/IC_{50})$ vary by 0.75 log unit, whereas $\log(1/ODT)$ varies by no less than 3.99 log units. It is therefore possible that the effect of OBPs is not the prime reason

Table 6 Comparison of complexation of VOCs with porcine OBPs, and odour thresholds

VOC	Log(1/ODT)[a]	Log(1/IC_{50})[b]
Benzylbenzoate	4.58	−0.59
Benzophenone	4.25	−0.56
Thymol	2.40	−0.40
2-Isobutyl-3-methoxy pyrazine	1.27	0.05
Undecanal	0.73	0.16
Dihydromyrcenol	0.59	0.10

[a]Equation (8). [b]From Vincent *et al.* (Vincent *et al.*, 2000).

for the variation of log(1/OTD), but that complexation to OBPs (or possibly the rate of complexation to OBPs) just mediates the effect of transport to, and interactions with, the receptor.

Equation (8) has other consequences, including the effect of homologues. Descriptors for the higher homologues are given in Table 7. The linear dependence of log(1/ODT) on **L**, as in equation (6), would lead to a regular increase in log(1/ODT) along an homologous series, as shown in Figure 3. However, the parabolic dependence on $(\mathbf{D} - \mathbf{D}^2)$ considerably modifies the linear increase and results in the prediction shown in Figure 3. The values of log(1/ODT) gradually become smaller than expected from the linear relationship, and eventually even begin to decrease, see Figure 3. This corresponds to a chemical cut-off in potency, a prediction that is completely outside the scope of previous analyses (Yamanaka, 1995; Hau and Connell, 1998). This predicted cut-off effect has a very important consequence. Hau *et al.* (Hau *et al.*, 2000) have used their partition model (Hau and Connell, 1998) to predict odour thresholds for VOCs found in the indoor environment. As pointed out above, these partition models do not include any cut-off effect at all, and hence higher homologues will be predicted to be more potent than on our model.

Another, very important, consequence follows from the initial equation (6). The dependent variable, log(1/ODT), conceptually takes the place of the dependent variable, log K, where K is a gas/biophase equilibrium constant given by:

$$K = \frac{[\text{no. of molecules of VOC in the biophase}]}{[\text{no. of molecules of VOC in the gas phase}]} \quad (10)$$
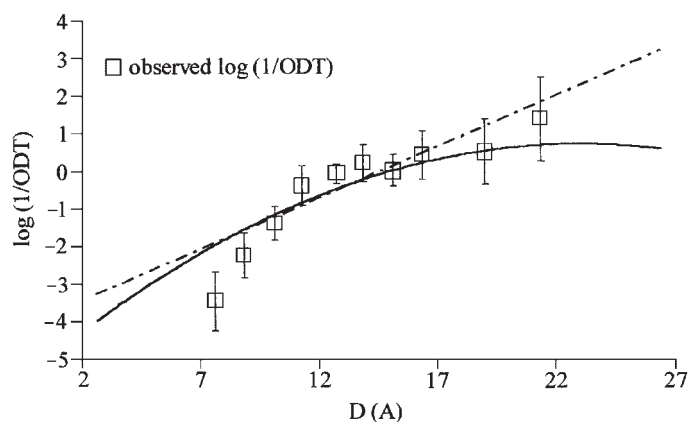
Table 7 Descriptors for higher homologous

Solute	E	S	A	B	L	D (A)
1-Nonanol	0.193	0.420	0.370	0.480	5.124	15.160
1-Decanol	0.191	0.420	0.370	0.480	5.628	16.400
1-Undecanol	0.181	0.420	0.370	0.480	6.139	17.670
1-Dodecanol	0.175	0.420	0.370	0.480	6.640	18.910
1-Tridecanol	0.169	0.420	0.370	0.480	7.149	20.110
1-Tetradecanol	0.163	0.420	0.370	0.480	7.656	21.430
Tridecylacetate	0.000	0.600	0.000	0.450	7.878	22.670
Tetradecylacetate	0.000	0.600	0.000	0.450	8.380	23.910
Pentadecylacetate	0.000	0.600	0.000	0.450	8.883	25.180
Hexadecylacetate	0.000	0.600	0.000	0.450	9.386	26.430
2-Decanone	0.108	0.680	0.000	0.510	5.245	15.390
2-Undecanone	0.101	0.680	0.000	0.510	5.732	16.630
2-Dodecanone	0.103	0.680	0.000	0.510	6.167	17.890
2-Tridecanone	0.100	0.680	0.000	0.510	6.672	19.140
2-Nonadecanone	0.100	0.680	0.000	0.510	9.554	26.650
Nonylbenzene	0.578	0.480	0.000	0.150	7.212	17.640
Decylbenzene	0.579	0.470	0.000	0.150	7.708	18.980
Undecylbenzene	0.579	0.470	0.000	0.150	8.159	20.180
Dodecylbenzene	0.571	0.470	0.000	0.150	8.600	21.390
Tridecylbenzene	0.570	0.470	0.000	0.150	9.132	22.590
Tetradecylbenzene	0.570	0.470	0.000	0.150	9.619	23.950

Figure 3 Plot of observed values of log(1/ODT) against the VOC maximum length D, for the homologous series of acetates. ---- Values calculated from equation 6; — values calculated from equation 7.

The ODT value itself represents the number of molecules in the gas phase, so that the only way that 1/ODT can take the place of an equilibrium constant, K, is if the number of molecules of a VOC in the biophase in equilibrium with the gas phase threshold value of the VOC, is the same for each VOC. This is a more general conclusion than the supposition of Hau and Connell (Hau and Connell, 1998) that the minimum proportion of available receptors necessary for the detection of odours is the same for all members of a homologous series, but differs from series to series.

The odour perception of enantiomers is well known, but invariably in terms of odour quality (Rossiter, 1996; Pybus and Sell, 1999). Rossiter (Rossiter, 1996) and Laska *et al*. (Laska *et al*., 1999) list pairs of enantiomers that elicit different sensations of odour quality. The latter workers tested odour discrimination of 10 pairs of enantiomers and concluded that within their experimental procedure, differences in odour intensity played little or no part in discrimination of the two enantiomeric forms. Other workers have shown that ODTs for $R(+)$- and $S(-)$-nicotine are essentially the same (Thuerauf *et al*., 1999). This again suggests selective, rather than specific, transport of VOCs to the biophase.

Regarding the potential implications of our results for the interpretation of olfactory receptor expression studies, we have shown that an equation set for selective transport of VOCs to the olfactory biophase is able to account for 77% of the total effect, measured as ODT. In order to account for the remaining effect, 'specific processes' need to be considered. The addition of a parabolic term in D (a maximum length parameter) or in L (a size parameter) raises the explained effect to ~85%. Thus, our data indicate that additional specific parameters, for example those derived from receptor-ligand studies, might be needed to account completely for the ODT measured. The question of whether selective transport is physically separable from the effects of receptor activation remains to be explored: If transport is not an intrinsic part of the stimulation of the receptors, but merely a filter, then research on receptors may well need to look at the residual after the transport aspects are subtracted.

Acknowledgements

References

Abraham, M.H. (1993) *Scales of solute hydrogen-bonding: their construction and application to physicochemical and biochemical processes.* Chem. Soc. Rev., 22, 73–83.

Abraham, M.H. and **Weathersby, P.K.** (1994) *Hydrogen bonding. 30. Solubility of gases and vapors in biological liquids and tissues.* J. Pharm. Sci., 83, 1450–1456.

Abraham, M.H., Whiting, G.S., Doherty, R.M. and **Shuely, W.J.** (1991) *Hydrogen bonding. 16. A new solute solvation parameter, π_2^H, from gas chromatographic data.* J. Chromatogr., 587, 213–228.

Abraham, M.H., Nielsen, G.D. and **Alarie, Y.** (1994a) *The Ferguson principle and an analysis of biological activity of gases and vapors.* J. Pharm. Sci., 83, 680–688.

Abraham, M.H., Andonian-Haftvan, J., Whiting, G.S., Leo, A. and **Taft, R.W.** (1994b) *Hydrogen bonding. Part 34. The factors that influence the solubility of gases and vapours in water at 298K, and a new method for its determination.* J. Chem. Soc., Perkin Trans., 2, 1777–1791.

Abraham, M.H., Kumarsingh, R., Cometto-Muniz, J.E. and **Cain, W.S.** (1998a) *An algorithm for nasal pungency thresholds in man.* Arch. Toxicol., 72, 227–232.

Abraham, M.H., Whiting, G.S., Carr, P.W. and **Ouyang, H.** (1998b) *Hydrogen bonding. Part 45. The solubility of gases and vapours in methanol at 298K: an LFER analysis.* J. Chem. Soc. Perkin Trans., 2, 1385–1390.

Abraham, M.H., Platts, J.A., Hersey, A., Leo, A.J. and **Taft, R.W.** (1999a) *Correlation and estimation of gas–chloroform and water–chloroform partition coefficients by a linear free energy relationship method.* J. Pharm. Sci., 88, 670–679.

Abraham, M.H., Le, J. and **Acree, W.E. Jr** (1999b) *The solvation properties of the aliphatic alcohols.* Collect. Czech. Chem. Commun., 64, 1748–1760.

Abraham, M.H., Poole, C.F and **Poole, S.K.** (1999c) *Classification of stationary phases and other materials by gas chromatography.* J. Chromatogr., A842, 79–114.

Alarie, Y., Schaper, M., Nielsen, G.D. and **Abraham, M.H.** (1998) *Structure–activity relationships of volatile organic chemicals as sensory irritants.* Arch. Toxicol., 72, 125–140.

AIHA (1989) Odor thresholds for chemicals with established occupational health standards. American Industrial Hygiene Association, Ohio.

Araneda, R.C., Kini, A.D. and **Firestein, S.** (2000) *The molecular receptive range of an odorant receptor.* Nat. Neurosci. 3, 1248–1255.

Bianchet, M.A., Bains, G., Pelosi, P., Pevsner, J., Snyder, S.H., Monaco, H.L. and **Amzel, L.M.** (1996) *The three-dimensional structure*

of bovine odorant binding protein and its mechanism of odor recognition. Nat. Struct. Biol., 3, 934–939.

Brownlow, S. and **Sawyer, L.** (1996) *Waiting to inhale.* Nat. Struct. Biol., 3, 902–906.

Chastrette, M. (1997) *Trends in structure–odor relationship.* SAR QSAR Environ. Res., 6, 215–254.

Cometto-Muniz, J.E. and **Cain, W.S.** (1990) *Thresholds for odor and nasal pungency*. Physiol. Behav., 48, 719–725.

Cometto-Muniz, J.E. and **Cain, W.S.** (1991) *Nasal pungency, odor and eye irritation thresholds for homologous acetates.* Pharmacol. Biochem. Behav., 39, 983–989.

Cometto-Muniz, J.E. and **Cain, W.S.** (1993) *Efficacy of volatile organic compounds in evoking nasal pungency and odor*. Arch. Environ. Health, 48, 309–314.

Cometto-Muniz, J.E. and **Cain, W.S.** (1994) *Sensory reactions of nasal pungency and odor to volatile organic compounds—the alkylbenzenes*. Am. Ind. Hyg. Assoc. J., 55, 811–817.

Cometto-Muniz, J.E. and **Cain, W.S.** (1995) *Relative sensitivity of the ocular trigeminal, nasal trigeminal, and olfactory systems to airborne chemicals*. Chem. Senses, 20, 191–198.

Cometto-Muniz, J.E., Cain, W.S and **Abraham, M.H.** (1998a) *Nasal pungency and odor of homologous aldehydes and carboxylic acids.* Exp. Brain Res., 118, 180–188.

Cometto-Muniz, J.E., Cain, W.S., Abraham, M.H. and **Kumarsingh, R.** (1998b) *Trigeminal and olfactory chemosensory impact of selected terpenes.* Pharmacol. Biochem. Behav., 60, 765–779.

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

Dravnieks, A. (1974) *A building block model for the characterization of odorant molecules and their odors.* Ann. N.Y. Acad. Sci., 237, 144–163.

Hau, K.M. and **Connell, D.W.** (1998) *Quantitative structure–activity relationships (QSARs) for odor thresholds of volatile organic compounds (VOCs).* Indoor Air, 8, 23–33.

Hau, K.M., Connell, D.W. and **Richardson, B.J.** (2000) *Use of partition models in setting health guidelines for volatile organic compounds.* Reg. Toxicol. Pharmacol., 31, 22–29.

Hornung, D.E. and **Mozell, M.M.** (1981) *Accessibility of odorant molecules to the receptors.* In Cagan, R.H. and Kare, M.R. (eds), Biochemistry of Taste and Olfaction. Academic Press, New York, pp. 33–45.

Johnson, B.A. and **Leon, M.** (2000) *Odorant molecular length: one aspect of the olfactory code.* J. Comp. Neurol., 426, 330–338.

Laffort, P. and **Patte, F.** (1987) *Solubility factors established by gas-liquid chromatography.* J. Chromatogr., 406, 51–74.

Laska, M., Liesen, A. and **Teubner, P.** (1999) *Enantioselectivity of odor perception squirrel monkeys and humans.* Am. J. Physiol., 277 (Regulatory Integrative Comp. Physiol. 46), R1098–R1103.

Löbel, D., Strotman, J., Jacob, M. and **Breer, H.** (2001) *Identification of a third rat odorant-binding protein (OBP3).* Chem. Senses, 26, 673–680.

Molecular Modeling Pro (TM) Revision 3.1 (1992). ChemSW Software Inc., 420F Executive Ct, N. Fairfield, CA 94585-4019, USA.

Pearce, T.C., Gardner, J.W. and **Gopel, W.** (1998) *Strategies for mimicking olfaction*. In Baltes, H., Gopel, W. and Hesse, J. (eds), Sensors Update, Vol. 3. Wiley-UCH, p. 73.

Pybus, D.H. and **Sell, C.S.** (1999) The Chemistry of Fragrances. Royal Society of Chemistry, London.

Rossiter, K.J. (1996) *Structure–odor relationships.* Chem. Rev., 96, 3201–3240.

Singer, M.S. and **Shepherd, G.M.** (1994) *Molecular modelling of ligand–receptor interactions on the OR5 olfactory receptor.* NeuroReport, 5, 1297–1300

Snyder, S.H., Sklar, P.B. and **Pevsner, J.** (1988) *Molecular mechanism of olfaction.* J. Biol. Chem., 263, 13971–13974.

Tegoni, M., Ramoni, R., Bignetti, E., Spinelli, S. and **Cambillau, C.** (1996) *Domain swapping creates a third putative combining site in bovine odorant binding protein dimer.* Nat. Struct. Biol., 3, 863–867.

Thuerauf, N., Kaegler, M., Dietz, R., Barocka, A. and **Kobal, G.** (1999) *Dose-dependent stereoselective activation of the trigeminal sensory system by nicotine in man.* Psychopharmacology, 142, 236–243.

Turner, D.B. and **Willett, P.** (2000) *The EVA spectral descriptor.* Eur. J. Med. Chem., 35, 367–375.

Vincent, F., Spinelli, S., Raamolni, R., Pelosi, P., Cambillo, C. and **Tegoni, M.** (2000) *Complexes of porcine odorant binding proteins with odorant molecules belonging to different chemical classes.* J. Mol. Biol., 300, 127–139.

Yamanaka, T. (1995) *Contribution of activity coefficients of methylene homologues to odor and flavor detection thresholds.* Chem. Senses, 29, 471–475.