CAMBRIDGE

An introduction to chemical terms for non-chemists

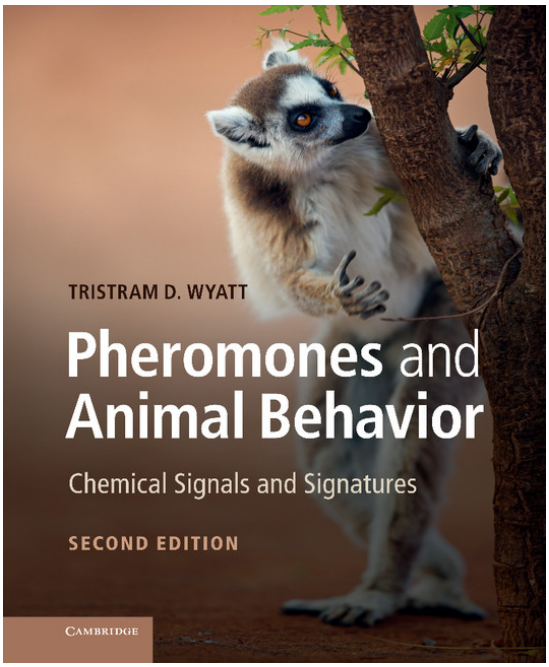
A **free** download available from www.cambridge.org/pheromones

Reprinted with permission from

Pheromones and Animal Behavior (2nd edition)

(Cambridge University Press, 2014)

by Tristram D. Wyatt



Winner, Best Postgraduate Textbook Prize, Society of Biology 2014

The judges said: “Revised and extended since the first edition, this splendid, comprehensive resource covers both classic ideas in the field of chemical communication as well as recent advances. ... All in all, an outstanding textbook and a worthy winner…”

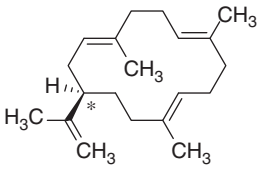
www.cambridge.org/pheromones

CAMBRIDGE UNIVERSITY PRESS

APPENDIX

An introduction to some chemical terms for non-chemists

The chemical names for semiochemicals can seem complicated to a biologist. One of my favorites is the termite sex and trail pheromone (1*E*,5*E*,9*E*,12*R*)-1,5,9-trimethyl-12-(1-methylethenyl)-1,5,9-cyclotetradecatriene (not surprisingly better known by its common name, neocembrene) (Chapter 7) (Bordereau & Pasteels 2011).

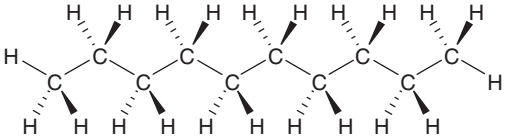


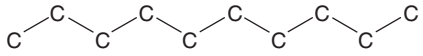
Neocembrene.

However, getting the molecule right in every detail can be essential for getting a response and the systematic naming protocols developed by chemists allow us to describe a particular molecule unambiguously. The aim of this short appendix is to explain the basics of chemical names. Some of the ideas and examples for this appendix come from chapters by Stevens in Howse *et al.* (1998).

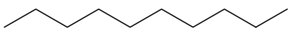
We need precise specification of molecules because chemosensory receptor proteins are activated by the particular characteristics of ligand molecules. These characteristics include a ligand's functional groups, charge, and size as well as its shape in three dimensions, hydrophobicity and flexibility (Chapter 9) (Kato & Touhara 2009; Leal 2013; Reisert & Restrepo 2009). This potential range of molecular characteristics, in biosynthesis and reception, has been acted on by natural selection so these variations are vitally important for responses to semiochemicals. Each of the possible variations between molecules described below can be important biologically, as in each case the molecule is different and it may stimulate a different range of olfactory receptors (Chapter 9). The same considerations of molecular characteristics apply to the selectivity of enzymes catalyzing the synthesis of biological signal molecules.

Organic molecules are based on a chain of carbon and attached hydrogen atoms. The carbon backbone forms a zigzag because of the tetrahedral arrangement of the four carbon bonds (these angles are important for other characteristics of the molecular shape, see Section A.2). The hydrogen atoms attached to the carbon backbone lie in two planes, above and below the paper (represented here by bonds as solid wedges (the plane above) and as dashed wedges (the plane below)):





The structure is often simplified to show just the carbon backbone:



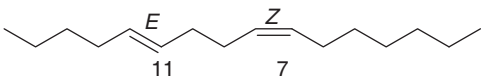
When other atoms such as oxygen or nitrogen, or other functional groups, are added to the chain or substituted for hydrogens or carbons, the chemical nature of the molecule changes (see Table A.1 for common functional groups).

The naming of compounds tells the reader the length of the carbon chain and what and where the important functional groups occur. The name also indicates the number and position of carbon double bonds (C=C). For example, the formal name for one component of the

Table A.1 **Prefixes and suffixes for common functional groups. Adapted from Howse *et al.* (1998).**

Functional group	Formula	Prefix	Suffix
Alcohol	-OH	Hydroxy-	-ol
Aldehyde	-CH=O	Formyl-	-al
Amine	$-NH_2$	Amino-	-amine
Carboxylic acid	-COOH	Carboxy-	-oic acid
Ester	-COOR	*R*-oxycarbonyl-	-*R*-oate
Ketone	>C=O	Oxo-	-one

pheromone of the pink bollworm, *Pectinophora gossypiella*, is (*Z*,*E*)-7,11-hexadecandienyl acetate (the other component is the *Z*,*Z* isomer; see Section A.2.2.1). The 'dien' tells us that there are two double bonds, the '7,11' that these occur at carbons 7 and 11:



(*Z*,*E*)-7,11-hexadecandienyl acetate.

Isomers

Having the formula for a molecule, e.g., C_2H_6O, is just the beginning. This molecular formula could describe many molecules. Isomers are groups of compounds that have the same atoms, in the same molecular formula, but have different structures and shapes. They may have different physical and chemical properties too. The word "isomer" comes from the Greek (*isos* = "equal" (same), méros = "part").

There are two main different kinds of isomers: **structural** (constitutional) isomers have the same atoms connected in different ways, whereas **stereoisomers** have atoms connected in the same way but they differ in the arrangement of atoms in space. Isomers exclude any temporary arrangements simply due to the molecule rotating as a whole or around particular bonds.

A.1 Structural (constitutional) isomers

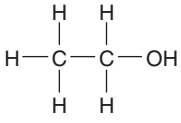
Structural (constitutional) isomers have the atoms connected in different ways, creating either different functional groups and/or different shaped molecules.

A.1.1 Chain isomers

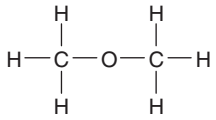
These arise because carbon chains can branch. For example there are two isomers of butane, C_4H_{10}. One has the carbons in a "straight chain," the other is branched.

A.1.2 Functional group isomers

Atoms with the same molecular formula can be connected in ways that produce different functional groups, giving quite different chemical properties to the molecule. For example, a molecular formula C_2H_6O could be ethanol (an alcohol) or methoxymethane (an ether):



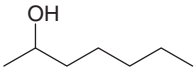
Ethanol



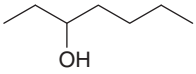
Methoxymethane ("dimethyl ether").

A.1.3 Positional isomers

Positional isomers share the same molecular formula but differ, for example, in the position of a functional group. Heptan-2-ol and heptan-3-ol, with the –OH group on the second or third carbon respectively:

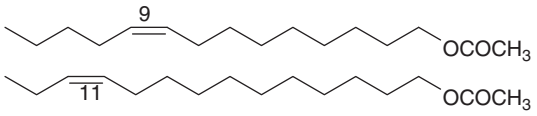


Heptan-2-ol.



Heptan-3-ol.

or the position of a double bond, for example (*Z*)-9-tetradecen-1-ol acetate and (*Z*)-11-tetradecen-1-ol acetate, the pheromone components of the summer fruit tortrix moth, *Adoxophyes orana*:



The highly variable cuticular hydrocarbons of ants, used in colony odors for nestmate recognition, include many different isomers using various positions of double bonds and/or the position and number of methyl ($-CH_3$) groups (Chapters 1 and 6) (Martin *et al.* 2008b).

A.2 Stereoisomers (spatial isomers)

Stereoisomers are compounds with the same atoms (the same molecular formula) and the same order of connecting the atoms together (connectivity), but with different spatial orientations of the atoms (configurations), changing the shape of the molecule. There are several types of stereoisomer. These are described in the following sections and the nomenclature used is shown in Table A.2. Stereoisomers are usually treated by receptors as different molecules (so proper chemical identification of semiochemical molecules must include stereochemistry, Mori 2007).

If the stereoisomers are mirror images of each other, they are called enantiomers. If they are not mirror images

Table A.2 **Naming of isomers: the meanings of the letters and symbols. The development of chemical nomenclature to precisely describe molecules has left us with a number of different naming schemes, including some that are synonymous and others that appear similar but are based on different principles.**

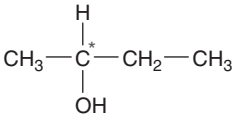
Symbols	Synonym	Type of isomer
E, Z		Geometrical spatial arrangement around double bond
l, d	–, +	Optical isomers (*laevo, dextro* rotation of polarized light)
L, D		Configurations of a sugar or amino acid, based on the absolute configuration of glyceraldehyde (Fischer nomenclature)
S, R		Different absolute configurations of stereocenters (asymmetric centers)

of each other then they are called diastereoisomers. Geometric isomers are a subclass of diastereoisomers.

A.2.1 Enantiomers (mirror image molecules)

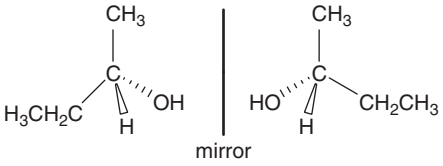
Chiral molecules are mirror images of each other and, just like your hands, cannot be superimposed on each other (chiral comes from the Greek word *cheir*, meaning "hand"). The two mirror images are called **enantiomers**. The ability to distinguish between a pair of enantiomers requires a chiral agent; a glove is a chiral agent that distinguishes right and left hands. Shoes do the same for feet. Enzymes and receptors are just such systems in biology. Nature is inherently chiral (Mori 2007).

The chirality is a consequence of the tetrahedral geometry of carbon atoms: if different groups are attached to each of the four bonds of one of the carbon atoms, the molecule can be made in two different ways (or as two enantiomers, see diagram below). A central carbon atom with four *different* groups attached is called a stereogenic or asymmetric center (sometimes called a chiral center). An asymmetric center can be in one of the two forms, *R* or *S* (see below) (Table A.2). It is called asymmetric because there is no way to split the molecule in half to give two equal halves. For example, butan-2-ol has four different groups attached to the central carbon:



The asymmetric carbon (chiral center) is marked with a star.

And represented spatially:



In this and other diagrams, solid wedges ▶ represent bonds coming out of the page toward you and dashed wedges [ıııı] (or dotted lines) represent bonds going away from you behind the plane of the paper. Lines – are bonds in the plane of the paper.

In the diagram below, a symmetrical molecule A with two different groups attached to the carbon is contrasted with an asymmetric molecule C with four different groups:

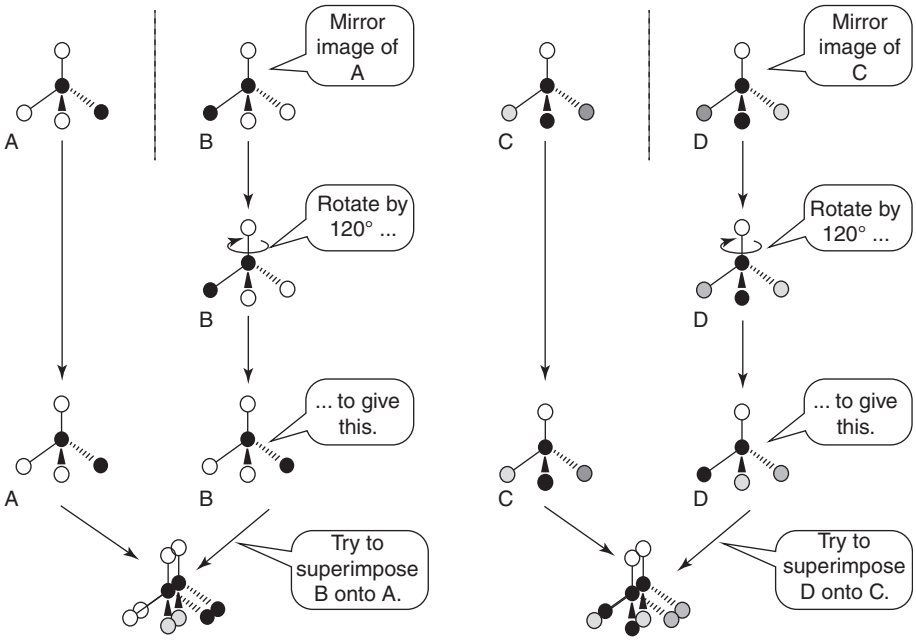
One way to see this in a molecule is to imagine the page of this book as a plane of symmetry and look along it:

A special property of chiral compounds, which led to the discovery of the phenomenon, is optical activity. Solutions of pure enantiomers rotate the plane of polarization of plane-polarized light passing through them: opposite enantiomers rotate it in opposite directions. The plane of polarization is rotated to the left by l-molecules, and to the right by d-molecules. The abbreviations come from Latin: l (*laevo*, for "left") and d (*dextro*, "right"). The letters l and d have been replaced by (–) and (+) respectively (Table A.2). A different naming system uses D and L.[1]

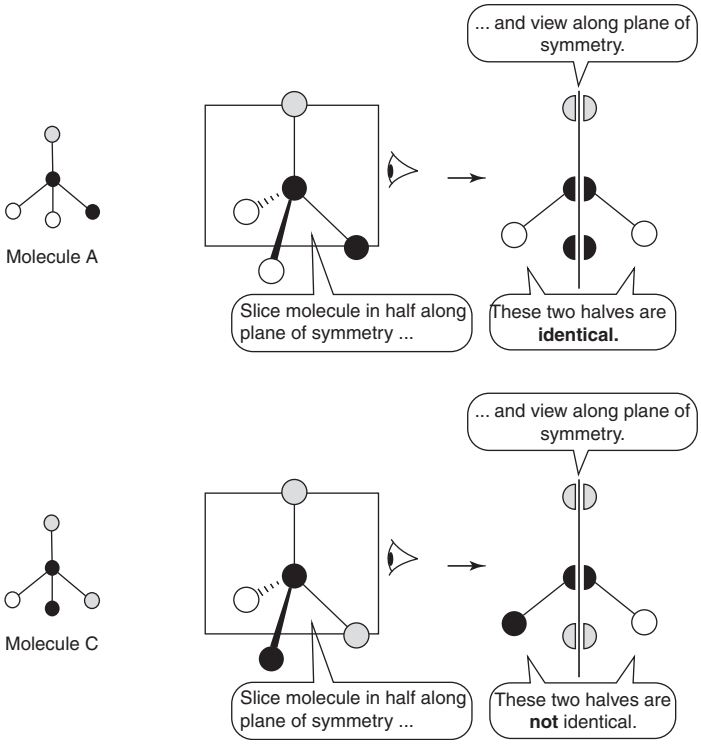
A **racemic mixture** or **racemate** is an equal mix of the two enantiomers. As they cancel each other out optically, such solutions do not rotate the plane of polarization of plane-polarized light, so they are optically inactive. Most chemically synthesized compounds, unless special steps are used in synthesis or purification, are racemic mixtures. Molecules produced by enzymes tend to be one enantiomer or the other.

The +/– naming system is based on the observation of the direction in which polarized light is shifted but

[1] An example is L–kynurenine, an amino acid used by masu salmon, *Oncorhynchus masou*, as a female sex pheromone. D and L are used to define the absolute configuration of a sugar or of an amino acid, based on the absolute configuration of glyceraldehyde (Fischer nomenclature). D, L (absolute configuration) should not be confused with d, l (rotation of the plane of polarization of plane-polarized light). The rules for naming are not important here, but biologically the differences between the different naming systems can be important (see Howse *et al.* 1998, p. 152).



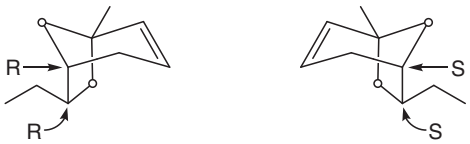
Molecule A and its mirror image (B) can be superimposed on each other. Molecule C, with four different groups around the central carbon, cannot be superimposed on its mirror image D. Molecules C and D are enantiomers. Figure from Crowe and Bradshaw (2010).



Molecule A is symmetrical: it can be divided in half along the plane of the page to generate two half molecules, which are identical. However, the enantiomer, Molecule C, is asymmetric. It cannot be divided in half along the plane of the page, or along any other plane to generate two identical half molecules. Caption and figure from Crowe and Bradshaw (2010).

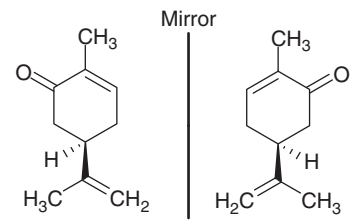
does not tell us the actual position of all atoms in space around a molecule. This is done by the **absolute configuration**, which uses a rather complicated set of rules that allow the chemist to describe the absolute configuration at each of the asymmetric centers as either *R* (from *rectus* = "right") and *S* (from *sinister* = "left"). As the *R* and *S* are defined according to a set of nomenclatural rules, these do not necessarily predict which way the whole molecule will shift polarized light (+/–). The full names of the carvone enantiomers are (*R*)-(–)-carvone and its mirror image, (*S*)-(+)-carvone.

An example of stereoselectivity in a mammalian pheromone is dehydro-*exo*-brevicomin, with two chiral centers, which comes in *R*, *R* and *S*, *S* forms; only the *R*, *R* form is active biologically and only this form is produced by male mice (Novotny *et al.* 1995):



(*R*,*R*)-3,4-dehydro-*exo*-brevicomin (*S*,*S*)-3,4-dehydro-*exo*-brevicomin.

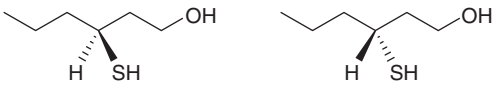
We can smell a difference between the enantiomers of many but not all molecules that have them (Laska 2004). An example of our sensitivity to enantiomers is the way we can distinguish (–)-carvone, which gives us the smell of spearmint, from its mirror image, (+)-carvone, perceived by us as the spicy smell of caraway seed:



(–)-carvone
smell: spearmint

(+)-carvone
caraway

Another pair of enantiomers that we can distinguish are molecules produced by bacteria from odorless precursors secreted in the armpits of 80% of the world's population: one enantiomer, (*S*)-3-methyl-3-sulfanylhexan-1-ol, is responsible for the oniony armpit smell of sweat:



(*S*)-3SH (*R*)-3SH

by contrast, the other enantiomer (*R*)-3-methyl-3-sulfanylhexan-1-ol, smells fruity, of grapefruit (Chapter 13) (Troccaz *et al.* 2009).

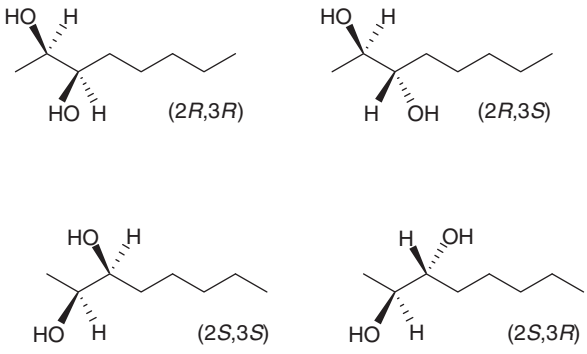
We and other organisms can distinguish different enantiomers of molecules because of the chiral characteristics of chemosensory receptors. In those cases where we can distinguish different enantiomers, they are stimulating different combinations of receptors, giving the signal to the brain that they are different smells (Chapter 9).

A.2.2 Diastereoisomers

When a molecule has more than one asymmetric center, a second form of stereoisomerism is possible, diastereoisomerism (stereoisomers that are not mirror images, so are not enantiomers). Each of the asymmetric centers can again be in one of the two forms, *R* and *S*. If there are two asymmetric centers there will be $2^2 = 4$ stereoisomers. Some may be mirror image molecules, others will not be (diastereoisomers).

For example, one component of the male sex pheromone of the cerambycid grape borer beetle, *Xylotrechus pyrrhoderus*, 2,3-octandiol occurs in four forms: (2*R*,3*R*), (2*R*,3*S*), (2*S*,3*S*), and (2*S*,3*R*):

Some of these molecules represent mirror images of each other, for example, (2*S*,3*S*) and (2*R*,3*R*) and are enantiomers. Other pairs of isomers, such as (2*R*,3*R*) and (2*S*,3*R*) are not mirror images and these are called diastereoisomers, that is, they are stereoisomers that are not enantiomers. These are common



The four optical isomers of 2, 3-octandiol.

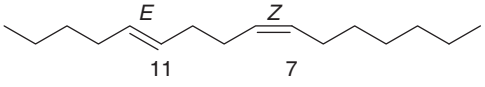
The four isomers of 2,3-octandiol. Only one, (2*S*,3*S*), is active as a pheromone component for the grape borer beetle, *Xylotrechus pyrrhoderus* (Sakai *et al.* 1984). Insects that use chiral pheromones typically produce and respond to either a single stereoisomer or to a species-specific blend of only some of the possible stereoisomers (Mori 2007).

in branched compounds. If there are three asymmetric centers in the molecule there will be eight isomers (2^3), as the number of optical isomers is 2^N where N is the number of asymmetric centers. This complicates the work of both the synthetic chemist and biologist as only some isomers (or combinations of them) may be active as pheromones.

A.2.2.1 Geometrical isomers

Geometrical isomers, also known as *cis–trans* or *E–Z* isomers, are a particular form of diastereoisomer. During biosynthesis of many moth pheromones, precursors have particular hydrogens removed, catalyzed by specific desaturase enzymes, to leave double bonds (see e.g., Figure 3.15). The double bond makes the carbon chain rigid at that point. Because of the angles of the chain, there are two versions of the molecule depending on the way molecules are arranged around the double bond, either on the same side (*Z*) or opposite sides (*E*) (from the German, *Zusammen* = "together" and *Entgegen* = "opposite"). Like other isomers, these two forms are different chemical compounds, often with slightly different chemical and physical properties. In simple molecules, the older terms *cis* and *trans* are roughly equivalent to *Z* and *E* respectively (though not strictly so by modern, more systematic naming rules better suited to more complicated structures, see www.chemguide.co.uk/basicorg/isomerism/ez.html).

This pheromone, one component of the pink bollworm moth, *Pectinophora gossypiella*, female pheromonehas two double bonds, one in each orientation:



(*Z*,*E*)-7,11-hexadecandienyl acetate.

Further reading

There is more about naming conventions for organic molecules here www.chemguide.co.uk/basicorg/convmenu.html and isomers here www.chemguide.co.uk/basicorg/isomermenu.html. These are websites written and maintained by Clark (2009).

Crowe and Bradshaw (2010) introduce chemistry for the biosciences.

The Leffingwell website www.leffingwell.com/chirality/chirality2.htm, has molecular structures of odor molecules and additional information on olfactory thresholds for these.

Mori (2007) describes the importance of chirality for pheromones, giving many examples of the ways that chirality determines specificity of pheromones in animals as varied as elephants and beetles.

For the structures of many pheromones see Pherobase www.pherobase.com (El-Sayed, 2013).

For the structure and naming of organic molecules try www.chemspider.com of the Royal Society of Chemistry. It allows you to search by common name and shows synonyms as well as the systematic names.

References

Bordereau, C & Pasteels, J M (2011) Pheromones and chemical ecology of dispersal and foraging in termites. In Bignell, D E, Roisin, Y & Lo, N (eds.) *Biology of Termites: a Modern Synthesis*, 2nd edn. pp. 279–320. Dordrecht: Springer.

Clark, J & Edexcel (2009) *Edexcel International GCSE Chemistry*. Harlow: Pearson Education. Supported by the website www.chemguide.co.uk.

Crowe, J & Bradshaw, T (2010) *Chemistry for the Biosciences: the Essential Concepts*, 2nd edn. Oxford: Oxford University Press.

El-Sayed, A M (2013) The Pherobase: database of pheromones and semiochemicals. [Online]. Available: www.pherobase.com [Accessed 6 June 2013].

Howse, P E, Stevens, I D R & Jones, O T (1998) *Insect Pheromones and their Use in Pest Management*. London: Chapman & Hall.

Kato, A & Touhara, K (2009) Mammalian olfactory receptors: pharmacology, G protein coupling and desensitization. *Cell Mol Life Sci* **66**: 3743–53.

Laska, M (2004) Olfactory discrimination ability of human subjects for enantiomers with an isopropenyl group at the chiral center. *Chem Senses* **29**: 143–52.

Leal, W S (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol* **58**: 373–91.

Martin, S J, Helanterä, H & Drijfhout, F P (2008b) Colony specific hydrocarbons identify nest mates in two species of Formica ant. *J Chem Ecol* **34**: 1072–80.

Mori, K (2007) Significance of chirality in pheromone science. *Biorg Med Chem* **15**: 7505–23.

Novotny, M V, Xie, T M, Harvey, S et al. (1995) Stereoselectivity in mammalian chemical communication – male-mouse pheromones. *Experientia* **51**: 738–43.

Reisert, J & Restrepo, D (2009) Molecular tuning of odorant receptors and its implication for odor signal processing. *Chem Senses* **34**: 535–45.

Sakai, T, Nakagawa, Y, Takahashi, J, Iwabuchi, K & Ishii, K (1984) Isolation and identification of the male sex pheromone of the grape borer *Xylotrechus pyrrhoderus* Bates (Coleoptera: Cerambycidae). *Chem Lett* **1984**: 263–4.

Troccaz, M, Borchard, G, Vuilleumier, C et al. (2009) Gender-specific differences between the concentrations of nonvolatile (*R*)/(*S*)-3-methyl-3-sulfanylhexan-1-ol and (*R*)/(*S*)-3-hydroxy-3-methyl-hexanoic acid odor precursors in axillary secretions. *Chem Senses* **34**: 203–10.