ORCA – Online Research @ Cardiff

Characterising human odorant signals - insights from insect semiochemistry and *in silico* modeling

Authors: Ashish Radadiya, John A. Pickett

Institution: School of Chemistry, Cardiff University, CF10 3AT, Wales, UK

Abstract: Interactions relating to human chemical signalling, although widely acknowledged, are relatively poorly characterised, chemically, except for human axillary odour. However, the extensive chemical ecology of insects, involving countless pheromone and other semiochemical identifications, may offer insights into overcoming problems of characterising human-derived semiochemicals more widely. Current techniques for acquiring insect semiochemicals are discussed, particularly in relation to the need for samples to relate, as closely as possible, to the ecological situation in which they are naturally deployed. Analysis is facilitated by chromatography coupled to electrophysiological preparations from the olfactory organs of insects *in vivo*. This is not feasible with human olfaction, but there are now potential approaches using molecular genetically reconstructed olfactory preparations already in use with insect systems.

There are specific insights of value for characterising human semiochemicals from advanced studies on semiochemicals of haematophagous insects, which include those involving human hosts, in addition to wider studies on farm and companion animals. The characterisation of the precise molecular properties recognised in olfaction could lead to new advances in analogue design and a range of novel semiochemicals for human benefit. There are insights from successful synthetic biology studies on insect semiochemicals using novel biosynthetic precursors. Already, wider opportunities in olfaction emerging from *in silico* studies, involving a range of theoretical and computational approaches to molecular design and understanding olfactory systems at the molecular level, are showing promise for studying human semiochemistry.

Authors for correspondence:

Ashish Radadiya,
E-mail: radadiyaa@cardiff.ac.uk

John A. Pickett,
E-mail: pickettj4@Cardiff.ac.uk

1. Introduction

Studies on insect chemical ecology receive more attention than human semiochemistry because of their impact as pests and also, more recently, with the growing concern for raising the value of insects in ecosystem services such as pollination. However, vertebrate and particularly mammal semiochemistry is more sophisticated in terms of the behaviours evoked, presumably on account of the highly evolved interactions with the central nervous system of these animals, especially true for human interactions. Here, some insights are offered from studies on insects that may be of value in chemically characterising the cues underlying olfactory communication in human beings and other vertebrates. The first step in characterising semiochemicals involves capturing samples for confirmation as being biologically active. For insects, there is usually an emphasis on such samples being derived directly from the insect itself and its own exocrine secretory system. Although bacteria living superficially on the cuticle of insects or at sites of excretion have provided some insights, for human studies there is more value in studying bacterial and other parasite-associated semiochemistry for non-invasive disease detection and relating directly to human interactions. The paper in this edition by Andreas Natsch and Roger Emter [1] describes the production of highly impacting human semiochemicals from conversion of essential precursors by human commensal bacteria. Semiochemicals from insects are preferably captured during a situation ecologically relevant to the signalling process under study, e.g. mating or host location. There is often a specific chemical composition of the pheromones for individual insect species, although there can be some intraspecific variation. However, human beings and other mammals often produce particular pheromonal signatures that differentiate individuals.

For insects, we can access the olfactory signal recognition system by insertion of microelectrodes into the olfactory organs, *i.e.* antennal electrophysiology. This allows direct monitoring of chromatographically separated components of complex mixtures comprising semiochemicals, and obviates the need to identify individual components for behavioural testing by allowing early pinpointing of biologically active compounds. It is also possible genetically to modify insect olfactory systems, for example recent advances in the neurophysiology of the fruit fly *Drosophila melanogaster* to investigate more sophisticated aspects of insect semiochemistry, in addition to chemical characterisation. This raises new prospects for studying human semiochemistry, with the potential for using overexpressed human molecular recognition genetics in non-sentient animal systems. However, whilst we await advances towards this approach, we can still exploit the insect system, particularly where insects have an evolutionary dependence on human semiochemistry, and we already have evidence for biting flies detecting not only human attractants, but semiochemicals that might relate directly to human interactions.

The rational design of better-performing insect semiochemical analogues is more difficult than for drug and pesticide discovery because of the substantially more precise nature of semiochemical recognition, but some opportunities are now being realised by synthetic biological techniques. We may also have new theoretical approaches to molecular design which could be applied to human semiochemical development.

2. Capturing semiochemicals from insects.

An early dispute over the natural composition of a honeybee pheromone involved in foraging and swarming behaviour was resolved using samples taken directly form the site of pheromone generation. This was achieved by injecting, and then retracting, minute volumes of inert solvent (Figure 1) from the closed secretory Nasonov gland known to produce the pheromone [2]. This approach established that aerial oxidation did not produce artefactual components, as had been

claimed, but that these were generated biosynthetically during release [3]. This pheromonal composition was then successfully mimicked in the field by a polymeric slow release formulation [4] and is currently a commercially valuable tool in apiculture, e.g. Swarm Commander. This approach may not appear to offer direct application in human semiochemical elucidation, but sampling of regions of human skin, and particularly organs within the epidermis, would comprise an analogous approach (Figure 2). For studies on human and other vertebrate pheromones, many solvents cause irritation whereas, for insects, there is much higher tolerance of solvent contact. Even inter-tergite regions of insects can accept hydrocarbon solvents without trauma. Instead of using solvents, polymeric absorbents, e.g. those for reverse-phase liquid chromatography [5], can be used to extract
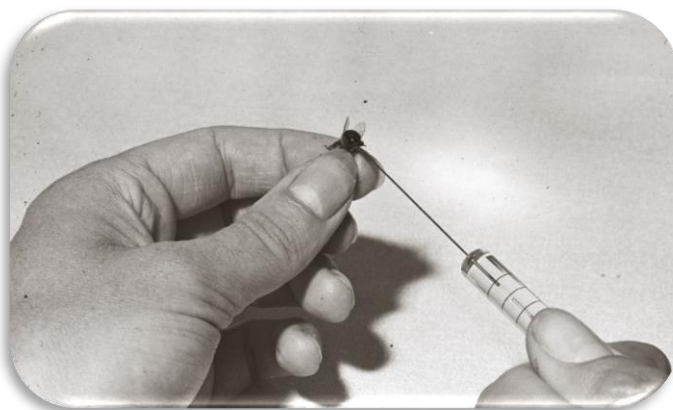
Figure 1 The Nasonov pheromone of the honey bee is sampled by placing the tip of a blunted syringe needle into the closed Nasonov gland minimising aerial oxidation or other perturbation of composition. (Courtesy Ingrid Williams, Rothamsted Research)

Figure 2 Sampling air from over a particular region of human skin for isolation of volatile semiochemicals. (Courtesy Jetske de Boer, Netherlands Institute of Technology)

human skin volatiles directly, and then either be eluted with solvent or the volatiles desorbed by a flow of inert gas, for analysis. For identifying insect pheromones as they are released into the air above an insect, entrainment onto a porous polymer is most effective and can accurately measure the naturally-released pheromone composition [6]. A very simple system employs solid phase micro-

extraction (SPME) such as by a retractable absorbent polymer fibre, which is exposed to the headspace above the pheromone-releasing insect and combined as a syringe device for direct insertion for gas chromatographic (GC) analysis [7]. This approach has the disadvantage of a relatively small part of the headspace being sampled and the sample is directly eluted into the GC apparatus, precluding multiple analysis, for example, of the chemical composition and biological activity [7]. The most successful approach is by the entrainment of air from above the biological sample (Figure 3) and by drawing the air through a porous polymer and eluting with a purified solvent, to obtain a liquid sample for multiple types of analysis [6]. For long-term storage, the sample must be sealed in a glass ampoule under an inert gas, preferably nitrogen. Insects can be held in a glass vessel through which a stream of clean air is passed. If the system is not sealed prior to the airflow being analysed, a positive pressure should be established to ensure no contamination enters the system. This can be readily applied to the study of mammal semiochemicals, and has been used for studies on human beings [8, 9] and farm [10] and companion animals [11]. Here, polyester or other non-contaminating polymeric bags are used to enclose parts of the body to be sampled selectively, and regions producing extraneous volatiles to be excluded. Thus, volatiles from the head and thereby breath can be obviated [8] (Figure 2). Such a procedure led to the identification of a group of simple compounds from children, comprising short chain oxo-substituted alkanes (Figure 4), which in the context of other human volatile compounds increased attraction of mosquitoes vectoring the causative pathogens of malaria [9].
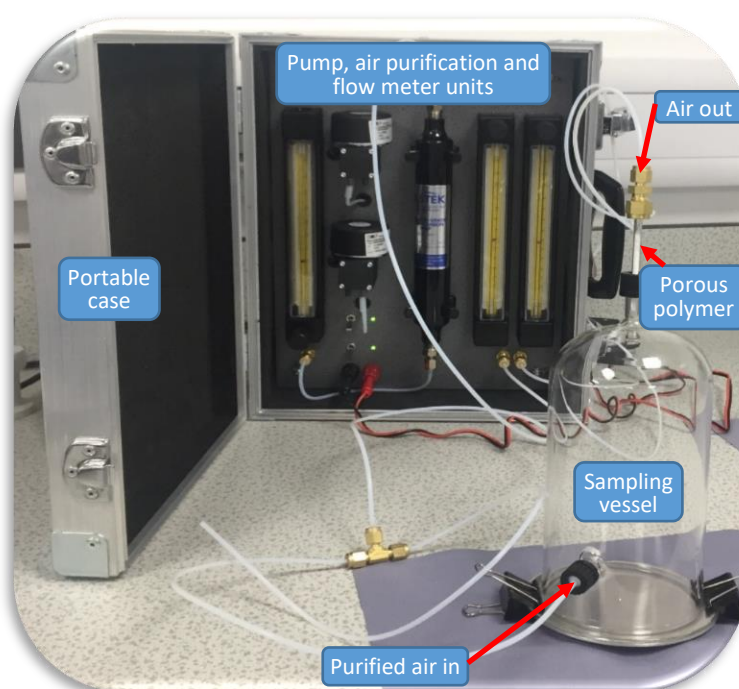
Figure 3 Portable apparatus for entrainment of volatiles from over biological samples, under positive pressure of purified air, by absorption on porous polymer. (Courtesy Irene Castellan, Cardiff University)
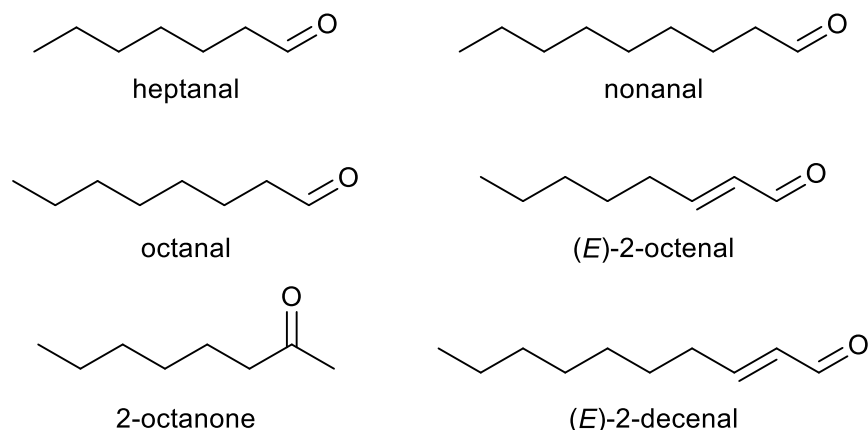
Figure 4 Human volatile compounds, associated with infection by the causative pathogen of malaria and increased attractiveness to the mosquito vector, isolated from children by air entrainment and identified by GC-MS following GC-EAG studies.

For insects, most pheromones are emitted away from waste excretory sites. However, as with human semiochemistry, there are exceptions. Thus, an aggregation pheromone, and faecal or detritus-related semiochemicals, can be entrained from air above cellulosic substrates acting as refuges for the common bed bug, *Cimex lectularius,* and show maximal aggregation activity [12]. Indeed, socks worn by human subjects function maximally to capture, directly, human-related semiochemicals for behavioural assays with haematophagous insects [13]. However, for GC-coupled electrophysiological and chemical analysis, entrainment onto porous polymers such Porapak Q or Tenax TA (Figure 3), from air having passed over the feet of individuals, is employed [9]. In more sophisticated investigations, for example on the subtle inherited differences of human semiochemical profiles, volatile samples are isolated directly by entrainment of air from over human hands [8, 14].

3. Capturing semiochemicals from the insect environment, including on the host

Insects can alter their pheromone composition or response when releasing on or near their host, e.g. sex pheromones released when the insect is on its host and thus indicating a food supply for ensuing larvae, particularly by synergising the pheromone response [15, 16]. Entrainment of the semiochemicals via porous polymers, with desorption by inert gas stripping, can allow analysis of compounds normally emerging with the solvent from the chromatograph, again using positive air pressure in glass vessels or polyester bags surrounding the host or part of the host (Figure 3) [7]. When insects feed on their hosts, this can reduce attractiveness of the host to subsequent attack, particularly by that insect species, thus avoiding competition. For plant hosts, this has been recently reviewed [17] and also discussed for vertebrate hosts [18, 19], particularly in terms of the contextual relationship with feeding stress and non-host related repellents. The analogy with human and other vertebrate semiochemical studies is that we can study signalling under stress to identify semiochemical markers of pathogen or parasite development [9], which can be useful in early diagnosis and may, perhaps, offer remedies for psychological problems.

4. Exploitation of haematophagous insects

Insects that have vertebrate hosts, particularly for the consumption of blood, not only determine suitability of potential hosts by taxonomic recognition via semiochemicals but can also avoid closely related animals which are less suitable hosts. This involves the insect detecting the presence, or heightened levels, of certain compounds which, in the context of normal host semiochemicals, cause avoidance or even repulsion. This phenomenon is an example of contextual repellency [18, 19]. Thus,

the tsetse flies, *Glossina* species, comprising the vectors of trypanosomes causing animal trypanosomiasis or nagana in sub-Saharan Africa, are known to avoid the non-host waterbuck, *Kobus ellipsiprymnus defassa,* although the livestock cattle hosts are in the same family (Bovidae). The semiochemicals identified as effecting this avoidance*, i.e.* pentanoic acid, 2-methoxyphenol, 6-propyltetrahydro-2*H*-2-one, (*E*)-6,10-dimethylundeca-5,9-dien-2-one, are used to protect cattle in the field against infection by trypanosomes [20]. Even within a breed of cattle [9] and within the human species, variation in the semiochemistry of individuals causes contextual inhibition of potential host characteristics [8]. This provides repellents for experimental treatments on human arms to repel mosquitoes, including *Anopheles gambiae*, the vector of the causative pathogens (*Plasmodium* species) of malaria [21], in favourable comparison with DEET, the "gold standard" of mosquito repellents.

From these studies, it is evident that haematophagous insects can detect semiochemicals underpinning host choice relating to human physiology, in addition to the manipulation by *Plasmodium* species of human semiochemistry causing attraction of the vector *An. gambiae*. There is evidence that even changes in human physiology, e.g. throughout the menstrual cycle, can be detected behaviourally by the yellow fever mosquito, *Aedes aegypti* (JG Logan and CM Woodcock 2000, unpublished). This means that, whilst development of non-invasive molecular tools directly for neurophysiological detection of human semiochemicals is awaited, live recordings comprising electrophysiological preparations from the antennae of haematophagous insects can be used to detect human semiochemicals unrelated, directly, to insect behaviour. Not only electroantennogram preparations of entire antennae, but also single neuron preparations, can be coupled to high resolution gas chromatography (GC) for human semiochemical identification studies and, as a technique for studying mosquito olfaction, has been available for over two decades [22].

The close relationship between human volatile production and olfactory recognition by haematophagous insects suggests that human pheromones may be detected by these insects. Thus, insect electrophysiology could be used in the identification of human pheromone components. Furthermore, olfactory organs of insects, specifically the "empty" neuron of the fruit fly, *Drosophila melanogaster*, can be modified to express functional olfactory proteins heterologously from other insects [23] to understand the mechanism, at the molecular level, that helps to discover largely unrecognised olfactory proteins and receptors. More speculatively, the opportunity may eventually arise for expressing human olfactory receptor proteins in these electrophysiologically more compliant organisms for direct identification studies of human semiochemistry. The fundamental differences between insect and mammalian olfactory systems are well understood (see §5a) and so, for this speculative approach to be realised, the amino acid sequences for molecular recognition by the proteins involved in human olfaction would need to be genetically expressed within the functional background of the insect olfactory system.

5. Rational design of better-performing semiochemical analogues

(a) Biosynthetic approaches

In learning from insect semiochemistry, there are well defined differences in the olfactory processes of insects compared with mammals. Insects have a family of odorant-gated ion channels comprising a highly conserved co-receptor (Orco), specifically functioning as the odorant-gated ion channel, and a divergent odorant receptor (OR) that confer chemical specificity and chemical signal amplification. In mammals, this is delivered by G-protein coupled ORs [24, 25]. Both are bathed in an aqueous lumen, with transport to the OR from the air mediated by odorant binding proteins (OBPs) or lipocalin-like proteins respectively. However, the molecular recognition systems in both, although insufficiently well

understood for predictable design of analogues, appear to be essentially similar. The rational design of semiochemical analogues, *i.e.* synthetic chemicals with biological activities related to those of known insect semiochemicals but with improved properties, e.g. resistance to oxidation or isomerisation under sunlight, would be an obvious way forward. However, the molecular recognition of semiochemicals, as for all olfactory recognition systems, as well as being poorly understood is extremely selective because of the need to discriminate semiochemicals from the vast range of other volatile compounds in the environment [24].

For drugs and pesticides, the recognition system is protected by the blood/brain barrier and other physiological filters, making the task of designing biologically active analogues a much a more rational process than for semiochemicals. However, some opportunities are now being realised for insect semiochemicals by a synthetic biological approach, which produces novel analogues by biosynthesis. In this approach, novel non-natural precursors of semiochemical ligands, from the late stages of the biosynthesis route, are fed to the final synthase enzymes in the biosynthesis of the semiochemical. This strategy has tested the hypothesis that, if the enzyme accepts the novel precursor and a semiochemical-like product ensues, the product will have a sufficiently similar chemical "space" to the natural semiochemical to show related biological activity. This hypothesis has tested true for the phytophagous insects, aphids, by which analogues of the natural plant stress-related sesquiterpene (*S*)-germacrene D have been generated that are highly active [26]. This work is currently being further developed for pest control in a push-pull system, by which the natural repellent (*S*)-germacrene D (push), and novel analogues, e.g. (*S*)-15-methylgermacrene, together with a novel attractant, (*S*)-14,15-dimethylgermacrene D (pull), are tested in the field. New semiochemicals have been targeted for related studies, for example 7-epizigiberene [27], a powerful whitefly repellent from wild tomato plants which is produced by a related biosynthetic pathway with similar potential late-stage biosynthetic precursors. This approach could be applied to human semiochemistry once new work has identified the specific enzymology by which some human semiochemicals are biosythesised, including terpenoids related to the plant semiochemical above, in which homozygous (identical) human twins are analysed for mosquito attraction and the correlation investigated for associated semiochemicals and genetics already known to exist [14].

(b) In silico approaches

Currently, we do not have sufficient understanding of the interaction between olfactory ligands and olfactory recognition proteins for the predictable design of novel synthetic semiochemicals. However, for drug and pesticide design, considerable knowledge is available and the synthesis of highly active analogues of natural and synthetic lead structures for molecular design is routine. In this process, computer-based (*in silico*) molecular design is categorized as either ligand-based or structure-based. Ligand-based (*in silico*) strategies are used from the existing knowledge of ligand structures, which for this example would be known semiochemicals. This technique is particularly useful when the experimental three-dimensional structure of the recognition protein is not available, *i.e.* the situation with olfaction. Indeed, historically, ligand-based methods have been used in semiochemical studies (see §5b(i)). However, without information on a direct protein interaction, we can potentially use an approach called homology modeling if coordinates from a related system are available (see §5b(ii)). This would then make available structure-based studies for olfaction by a virtual approach.

A major problem in earlier attempts to design novel semiochemicals in olfaction research is the perceived need to have relatively low molecular weights, and thereby boiling points, to provide sufficient volatility to reach the olfactory apparatus of insects. However, these essential physicochemical restrictions do allow for some higher molecular weights, provided the compounds are sufficiently lipophilic for a positive vapour pressure. For example, although (5*R*,6*S*)-6-acetoxy-5-
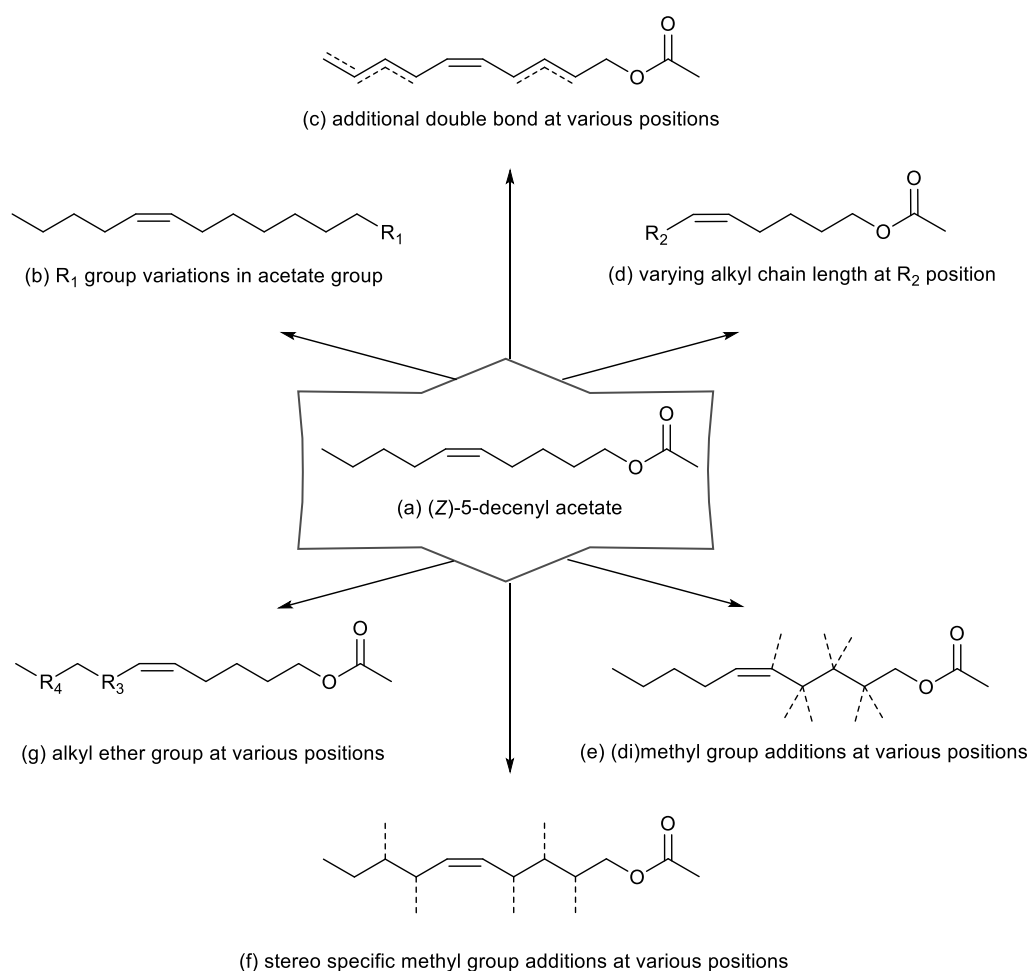
hexadecanolide contains 4 oxygen atoms with the molecular weight of 312.45 Da, it is the oviposition pheromone for *Culex* spp. Mosquitos, where the response is via olfaction.

As described in §5a, semiochemicals are perceived by insects in association with OBPs which, when bound together, are recognised by ORs [24, 25]. Furthermore, when studied *in vitro*, OBPs are apparently highly promiscuous and many ORs are not apparently selective, although certain ORs do selectively recognise pheromones. However, when this arrangement is complete *in vivo*, molecular recognition is spectacularly selective and sensitive. Additionally, insect ORs are only active when forming an heteromeric complex with the co-expressed Orco (also known as "tuning" receptor subunits [28-31]) and OR neurons also play a mediating role in signal transduction [32, 33].This complex molecular recognition system poses difficulties, of course, even in ligand-based modeling.

(i) Ligand-based approaches

The first ligand-based studies in olfaction involved Structure Activity Relationships (SARs), *i.e.* the relationships between the chemical structure of pheromones and other semiochemicals and their olfactory activities in insects. For further advances, these relationships were studied quantitatively by measuring biological activity against different but related chemical structures, termed Quantitative Structure Activity Relationships (QSARs). In this, QSAR provides a mathematical model derived from a series of similar molecules possessing known physiological activities, correlated with their physicochemical properties. For insects, physiological activity can readily be derived from electrophysiological activity (see §4). However, for human physiological activity, this would be more difficult to measure unless, in the future, it proves possible to express human molecular olfactory recognition in insect systems as proposed speculatively earlier (see §4). For physicochemical properties in QSAR of semiochemicals, predicted volatility from molecular parameters [34] would need to be added to other components of physicochemical properties, normally used in the study of QSARs for drugs and pesticides [35, 36]. In practical terms, a series of semiochemicals and analogues possessing experimental olfactory activity is divided into training and test sets. The training set is used to derive a QSAR model and the test set is used to determine the quality of the model. For semiochemicals, the advanced QSAR modeling technique, 3D-QSAR, developed using the three-dimensional (3D) properties of chemical structures, including molecular shape, electrostatic and hydrophobic properties calculated in 3D space, could be useful in developing new semiochemicals.

For insect olfaction, SAR studies were initiated with the identification of (*Z*)-7-dodecenyl acetate and other minor components, including (*Z*)-5-decenyl acetate, as the sex pheromone produced by the female turnip moth, *Agrotis segetum* [37]. Synthetic analogues of (*Z*)-7-dodecenyl acetate led to SARs emphasizing the importance of the acetate group, along with strict requirements for the polar functional group for pheromonal activity in *A. segetum* [38]. The SAR study was extended to using (*Z,E*)-dienic analogues of the minor pheromonal component (*Z*)-5-decenyl acetate, which identified a new role for *E* double bonds in conferring pheromonal activity [39]. Varying the alkyl chain length among the analogues of (*Z*)-5-decenyl acetate led to an extended SAR, suggesting that the terminal alkyl chain requires a specific chain length for interaction with a hydrophobic part of the protein target [40]. Subsequently, the introduction of methyl groups into the carbon chain of (*Z*)-5-decenyl acetate [41] showed discrimination amongst the ensuing optical isomers [42]. This investigation provided the first SAR example (Figure 5) of a systematic variation in chemical structures and impact on pheromonal activity. Although this was an interesting initiation to the SAR study of semiochemicals, the lack of improved activity suggested further refinements for the approach, and so the data from 49 analogues were used for 3D-QSAR modeling [43]. Although this 3D-QSAR provided 3D contour maps of the potential binding pocket of the target protein, it again did not lead to the production of new semiochemicals and pointed to the need for further experiments.

(c) additional double bond at various positions

(b) R₁ group variations in acetate group

(d) varying alkyl chain length at R₂ position

(a) (Z)-5-decenyl acetate

(g) alkyl ether group at various positions

(e) (di)methyl group additions at various positions

(f) stereo specific methyl group additions at various positions

Figure 5 Development of analogues of the insect sex pheromone from the female moth *Agrotis segetum* for (a) minor component (*Z*)-5-decenyl acetate analogues ; (b) major component (*Z*)-7-dodecenyl acetate analogues by varying group R₁; (c) analogues introduced with one or more *E* double bonds; (d) analogues with varying carbon chain length at R₂; (e) methyl analogues with increased branching; (f) methyl analogues with defined chirality; (g) analogues incorporating ether functionality at R₃ or R₄ and having varying alkyl chain lengths.

Human olfaction. An early attempt to study human semiochemistry by QSAR targeted the olfactorily active components of ambergris. Although ambergris is a naturally produced lipoidal secretion known for its fixative properties in perfumery [44], the principal component, a triterpene alcohol, ambrein, oxidized into a range of compounds with human odorant properties. A QSAR study on 98 of these compounds was made, in which physiological activities were measured using a panel of human volunteers trained on a series of 10 semantic odour descriptors (*i.e.* different types of smells including earthy, woody, camphor, fruity, rosy, marine, sandalwood, musky and cedarwood) [45]. This was correlated with a set of physicochemical properties expressing descriptors of molecular characteristics, including chirality and other topological elements and atomic connectivity etc., calculated with the help of a range of computer software tools [45]. This type of modeling could be used to derive synthetic ambergris and other fragrances, depending upon the availability of a wide range of natural analogues.

An artificial intelligence approach has been made to correlate physicochemical parameters of human semiochemicals with their olfactory perception. This involved establishing a crowdsourcing non-profit

community working through contributors across the research spectrum, including researchers from universities, technology companies, not-for-profit agencies, and biotechnology and pharmaceutical companies, entitled the DREAM Olfaction Prediction Consortium. The objective was to develop machine-learning algorithms that could differentiate 476 structurally diverse odorant molecules [46]. For this, the physiological parameters, determined according to 21 semantic odorant descriptors, were derived initially from the individual perceptions of a panel of 49 volunteers. An additional 19 semantic odorant descriptors were derived by averaging the population perceptions, also from the 49 volunteers. The artificial intelligence approach allowed a greater number of physicochemical properties than normally available to QSAR to be correlated with the olfactory properties of the odorant molecules [47]. Among the dataset from the 476 molecular structures, 338 were given to 19 participating teams at the international level, while the omitted 69 were used in qualifying tests. Of the participating 19 teams, one team was found to be the most effective in developing machine-learning *in silico* models, specialized in predicting individual perceptions. A second team was found to be most effective in developing *in silico* predictive models, specialized in population perception [47, 48]. These two separate artificial machine-learning models were then used for differentiating structurally diverse molecules from the 376, based on odour intensity and pleasantness, and could also correctly predict 8 among 19 semantic odour descriptors, specifically garlic, fish, sweet, fruit, burnt, spices, flower and sour.

Future ligand-based directions. Use of simple QSAR and 3D-QSAR based approaches to semiochemical analogue production could be more valuable if vastly more molecular structures were analysed, also employing the apparent advantages of a machine-learning approach [47, 48]. Because QSAR studies usually do not use entirely random libraries, it is essential to include structural diversity so as not to miss unexpected structure-activity relationships, particularly important for semiochemical invention where there are structurally unrelated compounds with similar biological activity. *In silico* searches of vast theoretical libraries could provide virtual hits to be followed by synthesis, but this objective would not be served by use of combinatorial libraries which, by definition, lack diversity. Indeed, for semiochemical analogue discovery, the creation of specifically diverse semiochemical-related libraries would be valuable. Commercially available virtual libraries, such as Enamine-REAL [49] and ZINC [50] containing millions of molecular structures, could be assessed for the availability of semiochemical-like molecules. This could be achieved by searching natural semiochemical structures, eliminating non-volatile compounds and those likely to interact pharmacologically with the insect recognition processes.

(ii) Structure-based approaches

Many aspects of the mechanisms and physiology of olfactory systems, including the key ORs, remain unknown. There are nearly 400 human ORs identified but no X-ray, NMR or even Cryo-EM based structures have been reported [51, 52]. Since these ORs are relatively high molecular weight transmembrane proteins and difficult to crystalize, this limits olfactory research and the identification of human pheromones. A Cryo-EM structure for an insect Orco (a co-expressed co-receptor), which is essential for the functioning of insect ORs (see §5a) has recently been reported, but only at low resolution [24].

Essentially, for advancing human olfactory research, it is important to identify the functionality and molecular mechanisms of G-protein coupled ORs. Therefore, in the absence of experimental structural data, we can use homology modeling, involving structure-based computational biology in conjunction with chemical biology. Homology models of ORs can be built using computational approaches in multiple steps, sequentially starting from protein sequence alignment search tools [53, 54], for identifying similar protein or protein domain structures from the protein data banks [55, 56]. The best

matches can then be used as templates for construction of the homology model and there are well developed methods available for this general purpose [57, 58]. However, it is crucial to assess the quality of the newly constructed homology model by means of various, equally well established, techniques [57-59]. This homology model can be tested *in silico* using relevant semiochemical structures and further refinements made to the homology model. The model is then tested experimentally to access discrimination between semiochemical structures.

The putative human odorant receptors OR5AN1 and OR1A1, for musk-related odorants used in perfumery, have been modeled to derive virtual 3D homology models [60]. In this process, an X-ray crystal structure of a specific muscarinic acetylcholine receptor was used as a template. This receptor comprises the closest homologue protein for which there is a high-grade crystal structure, and which is defined by interaction with the alkaloid muscarine. The homology models of OR5AN1 and OR1A1 were used to identify the binding specificity of 35 musk-related odorants, including all three classes of aromatic nitro-, polycyclic- and macrocyclic- musk odorants. Computer-based studies of molecular interactions showed how these musk odorants differentially interacted with the homology models of OR5AN1 and OR1A1. This was experimentally confirmed using a human cell line (Hana3A) to express, heterologously, the OR5AN1 and OR1A1 proteins. In addition, the amino acid residues identified as essential for molecular recognition in the homology model were confirmed experimentally by directed mutation studies, Tyr260 to Phe in OR5AN1, and Tyr251 to Phe and Tyr258 to Phe in OR1A1 proteins [60].

In another study, a homology model of OR7D4 was developed from the template protein structure of a β2-adrenergic receptor [61]. This model was then used to investigate the differential perception of androstenone as being either foul-smelling or pleasant. This work related to the natural mutant variants in the human OR, OR7D4-WM, confirming *in silico* that the mutations Arg88 to Typ and Thr133 to Met should change the perception of androstenone from foul to pleasant. As with the previous study, the amino acids identified as crucial for olfactory discrimination were confirmed by experimental mutation studies, with a success rate of 75% [61].

6. Conclusions

As the growing need for exploiting insect semiochemistry becomes even more apparent, there will perhaps be more insights gained in the quest to overcome the many challenges associated with further understanding human chemical signalling networks, and associated molecular modulators, at the physiological level. The identification of insect semiochemical structures is rendered simple, in principle, by use of electrophysiological preparations from the insect olfactory recognition apparatus. However, with the eventual possibility of transferring, by heterologous gene expression into the insect recording systems, crucial gene products from human and other vertebrate olfactory recognition systems, the analytical convenience of insect electrophysiological analysis could be made available. As has been done for many insects, the specific semiochemicals with which each OR interacts, at physiologically relevant stimulus concentrations, need to be identified to advance this aspect of human semiochemistry research. We have the task of exploring new opportunities beyond those from insect semiochemistry research and should try, as some are doing in other sections of this issue [1], to research ways for advancing human wellbeing with mood enhancers and better therapies. As the exploitation of insect semiochemistry impacts more on food sustainability, and also on farm animal and human health, we can point to the social and economic value of olfaction research and development in this field. Although it would perhaps be specious to believe that we could manipulate human behaviour in quite the same dramatic and predictable way as can be achieved with insects, the value of insights from insect semiochemistry may provide opportunities to expand the understanding of human semiochemistry and help move towards innovation in human health and wellbeing.

References

[1] Natsch, A. & Emter, R. 2020 The specific biochemistry of human axilla odour formation viewed in an evolutionary context. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **375**, 20190269. (DOI:10.1098/rstb.2019.0269).

[2] Pickett, J. A., Williams, I. H., Martin, A. P. & Smith, M. C. 1980 Nasonov Pheromone of the Honey Bee, Apis-Mellifera L (Hymenoptera, Apidae) .1. Chemical Characterization. *Journal of Chemical Ecology* **6**, 425-434. (DOI:Doi 10.1007/Bf01402919).

[3] Pickett, J. A., Williams, I. H., Smith, M. C. & Martin, A. P. 1981 Nasonov Pheromone of the Honey Bee, Apis-Mellifera L (Hymenoptera, Apidae) .3. Regulation of Pheromone Composition and Production. *Journal of Chemical Ecology* **7**, 543-554. (DOI:Doi 10.1007/Bf00987702).

[4] Kigatiira, I. K., Beament, J. W. L., Free, J. B. & Pickett, J. A. 1986 Using Synthetic Pheromone Lures to Attract Honeybee Colonies in Kenya. *J Apicult Res* **25**, 85-86. (DOI:10.1080/00218839.1986.11100698).

[5] Gikonyo, N. K., Hassanali, A., Njagi, P. G. N., Gitu, P. M. & Midiwo, J. O. 2002 Odor composition of preferred (buffalo and ox) and nonpreferred (waterbuck) hosts of some savanna tsetse flies. *Journal of Chemical Ecology* **28**, 969-981. (DOI:10.1023/A:1015205716921).

[6] Harvey, D. J., Vuts, J., Hooper, A., Finch, P., Woodcock, C. M., Caulfield, J. C., Kadej, M., Smolis, A., Withall, D. M., Henshall, S., et al. 2018 Environmentally vulnerable noble chafers exhibit unusual pheromone-mediated behaviour. *Plos One* **13**, e0206526. (DOI:10.1371/journal.pone.0206526).

[7] Agelopoulos, N. G. & Pickett, J. A. 1998 Headspace analysis in chemical ecology: Effects of different sampling methods on ratios of volatile compounds present in headspace samples. *Journal of Chemical Ecology* **24**, 1161-1172. (DOI:Doi 10.1023/A:1022442818196).

[8] Logan, J. G., Birkett, M. A., Clark, S. J., Powers, S., Seal, N. J., Wadhams, L. J., Mordue, A. J. & Pickett, J. A. 2008 Identification of human-derived volatile chemicals that interfere with attraction of Aedes aegypti mosquitoes. *Journal of Chemical Ecology* **34**, 308-322. (DOI:10.1007/s10886-008-9436-0).

[9] Robinson, A., Busula, A. O., Voets, M. A., Beshir, K. B., Caulfield, J. C., Powers, S. J., Verhulst, N. O., Winskill, P., Muwanguzi, J., Birkett, M. A., et al. 2018 Plasmodium-associated changes in human odor attract mosquitoes. *P Natl Acad Sci USA* **115**, E4209-E4218. (DOI:10.1073/pnas.1721610115).

[10] Birkett, M. A., Agelopoulos, N., Jensen, K. M. V., Jespersen, J. B., Pickett, J. A., Prijs, H. J., Thomas, G., Trapman, J. J., Wadhams, L. J. & Woodcock, C. M. 2004 The role of volatile semiochemicals in mediating host location and selection by nuisance and disease-transmitting cattle flies. *Med Vet Entomol* **18**, 313-322. (DOI:DOI 10.1111/j.0269-283X.2004.00528.x).

[11] Borges, L. M. F., de Oliveira, J. G., Ferreira, L. L., Louly, C. C. B., Pickett, J. A. & Birkett, M. A. 2015 Identification of non-host semiochemicals for the brown dog tick, Rhipicephalus sanguineus sensu lato (Acari: Ixodidae), from tick-resistant beagles, Canis lupus familiaris. *Ticks Tick-Borne Dis* **6**, 676-682. (DOI:10.1016/j.ttbdis.2015.05.014).

[12] Weeks, E. N. I., Logan, J. G., Birkett, M. A., Pickett, J. A. & Cameron, M. M. 2013 Tracking bed bugs (Cimex lectularius): a study of the effect of physiological and extrinsic factors on the response to bed bug-derived volatiles. *J Exp Biol* **216**, 460-469. (DOI:10.1242/jeb.074930).

[13] Verhulst, N. O., Andriessen, R., Groenhagen, U., Kiss, G. B., Schulz, S., Takken, W., van Loon, J. J. A., Schraa, G. & Smallegange, R. C. 2010 Differential Attraction of Malaria Mosquitoes to Volatile Blends Produced by Human Skin Bacteria. *Plos One* **5**, e15829. (DOI:10.1371/journal.pone.0015829).

[14] Fernandez-Grandon, G. M., Gezan, S. A., Armour, J. A. L., Pickett, J. A. & Logan, J. G. 2015 Heritability of Attractiveness to Mosquitoes. *Plos One* **10**, e0122716. (DOI:10.1371/journal.pone.0122716).

[15] Dawson, G. W., Griffiths, D. C., Pickett, J. A., Wadhams, L. J. & Woodcock, C. M. 1987 Plant-Derived Synergists of Alarm Pheromone from Turnip Aphid, Lipaphis (Hyadaphis) Erysimi (Homoptera, Aphididae). *Journal of Chemical Ecology* **13**, 1663-1671. (DOI:Doi 10.1007/Bf00980207).

[16] Dickens, J. C., Smith, J. W. & Light, D. M. 1993 Green leaf volatiles enhance sex attractant pheromone of the tobacco budworm,Heliothis virescens (Lep.: Noctuidae). *CHEMOECOLOGY* **4**, 175-177. (DOI:10.1007/bf01256553).

[17] Pickett, J. A. & Khan, Z. R. 2016 Plant volatile-mediated signalling and its application in agriculture: successes and challenges. *New Phytol* **212**, 856-870. (DOI:10.1111/nph.14274).

[18] Pickett, J. A., Birkett, M. A., Dewhirst, S. Y., Logan, J. G., Omolo, M. O., Torto, B., Pelletier, J., Syed, Z. & Leal, W. S. 2010 Chemical Ecology of Animal and Human Pathogen Vectors in a Changing Global Climate. *Journal of Chemical Ecology* **36**, 113-121. (DOI:10.1007/s10886-010-9739-9).

[19] Pickett, J. A., Barasa, S. & Birkett, M. A. 2014 Vertebrate pheromones and other semiochemicals: the potential for accommodating complexity in signalling by volatile compounds for vertebrate management. *Biochem Soc T* **42**, 846-850. (DOI:10.1042/Bst20140134).

[20] Saini, R. K., Orindi, B. O., Mbahin, N., Andoke, J. A., Muasa, P. N., Mbuvi, D. M., Muya, C. M., Pickett, J. A. & Borgemeister, C. W. 2017 Protecting cows in small holder farms in East Africa from tsetse flies by mimicking the odor profile of a non-host bovid. *Plos Neglect Trop D* **11**, e0005977. (DOI:10.1371/journal.pntd.0005977).

[21] Logan, J. G., Stanczyk, N. M., Hassanali, A., Kemei, J., Santana, A. E. G., Ribeiro, K. A. L., Pickett, J. A. & Mordue, A. J. 2010 Arm-in-cage testing of natural human-derived mosquito repellents. *Malaria J* **9**, 239. (DOI:10.1186/1475-2875-9-239).

[22] Pickett, J. A. & Woodcock, C. M. 1996 The role of mosquito olfaction in oviposition site location and in the avoidance of unsuitable hosts. *Ciba F Symp* **200**, 109-123. (DOI:10.1002/9780470514948.ch9).

[23] Dobritsa, A. A., van der Goes van Naters, W., Warr, C. G., Steinbrecht, R. A. & Carlson, J. R. 2003 Integrating the molecular and cellular basis of odor coding in the Drosophila antenna. *Neuron* **37**, 827-841. (DOI:Doi 10.1016/S0896-6273(03)00094-1).

[24] Butterwick, J. A., Del Marmol, J., Kim, K. H., Kahlson, M. A., Rogow, J. A., Walz, T. & Ruta, V. 2018 Cryo-EM structure of the insect olfactory receptor Orco. *Nature* **560**, 447-452. (DOI:10.1038/s41586-018-0420-8).

[25] Silbering, A. F. & Benton, R. 2010 Ionotropic and metabotropic mechanisms in chemoreception: 'chance or design'? *Embo Rep* **11**, 173-179. (DOI:10.1038/embor.2010.8).

[26] Touchet, S., Chamberlain, K., Woodcock, C. M., Miller, D. J., Birkett, M. A., Pickett, J. A. & Allemann, R. K. 2015 Novel olfactory ligands via terpene synthases. *Chem Commun* **51**, 7550-7553. (DOI:10.1039/c5cc01814e).

[27] Avila, C. A., Marconi, T. G., Viloria, Z., Kurpis, J. & Del Rio, S. Y. 2019 Bactericera cockerelli resistance in the wild tomato Solanum habrochaites is polygenic and influenced by the presence of Candidatus Liberibacter solanacearum. *Sci Rep-Uk* **9**. (DOI:s41598-019-50379-7).

[28] Larsson, M. C., Domingos, A. I., Jones, W. D., Chiappe, M. E., Amrein, H. & Vosshall, L. B. 2004 Or83b encodes a broadly expressed odorant receptor essential for Drosophila olfaction. *Neuron* **43**, 703-714. (DOI:10.1016/j.neuron.2004.08.019).

[29] Neuhaus, E. M., Gisselmann, G., Zhang, W., Dooley, R., Stortkuhl, K. & Hatt, H. 2005 Odorant receptor heterodimerization in the olfactory system of Drosophila melanogaster. *Nat Neurosci* **8**, 15-17. (DOI:10.1038/nn1371).

[30] Benton, R., Sachse, S., Michnick, S. W. & Vosshall, L. B. 2006 Atypical membrane topology and heteromeric function of Drosophila odorant receptors in vivo. *PLoS Biol* **4**, e20. (DOI:10.1371/journal.pbio.0040020).

[31] Sato, K., Pellegrino, M., Nakagawa, T., Nakagawa, T., Vosshall, L. B. & Touhara, K. 2008 Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**, 1002-1006. (DOI:10.1038/nature06850).

[32] Bhandawat, V., Olsen, S. R., Gouwens, N. W., Schlief, M. L. & Wilson, R. I. 2007 Sensory processing in the Drosophila antennal lobe increases reliability and separability of ensemble odor representations. *Nat Neurosci* **10**, 1474-1482. (DOI:10.1038/nn1976).

[33] Luo, S. X., Axel, R. & Abbott, L. F. 2010 Generating sparse and selective third-order responses in the olfactory system of the fly. *Proc Natl Acad Sci U S A* **107**, 10713-10718. (DOI:10.1073/pnas.1005635107).

[34] Olsen, E. & Nielsen, F. 2001 Predicting vapour pressures of organic compounds from their chemical structure for classification according to the VOC-directive and risk assessment in general. *Molecules* **6**, 370-389. (DOI:Doi 10.3390/60400370).

[35] Gramatica, P. 2007 Principles of QSAR models validation: internal and external. *Qsar Comb Sci* **26**, 694-701. (DOI:10.1002/qsar.200610151).

[36] Tropsha, A. 2010 Best Practices for QSAR Model Development, Validation, and Exploitation. *Molecular Informatics* **29**, 476-488. (DOI:10.1002/minf.201000061).

[37] Lofstedt, C., Vanderpers, J. N. C., Lofqvist, J., Lanne, B. S., Appelgren, M., Bergstrom, G. & Thelin, B. 1982 Sex-Pheromone Components of the Turnip Moth, Agrotis-Segetum Schiff, Lepidoptera, Noctuidae - Chemical-Identification, Electro-Physiological Evaluation and Behavioral Activity. *Journal of Chemical Ecology* **8**, 1305-1321. (DOI:Doi 10.1007/Bf00987764).

[38] Liljefors, T., Thelin, B. & Van Der Pers, J. N. 1984 Structure-activity relationships between stimulus molecule and response of a pheromone receptor cell in turnip moth,Agrotis segetum : Modifications of the acetate group. *J Chem Ecol* **10**, 1661-1675. (DOI:10.1007/BF00987353).

[39] Bengtsson, M., Liljefors, T. & Hansson, B. S. 1987 Dienic analogs of (Z)-5-decenyl acetate, a pheromone component of the turnip moth, Agrotis segetum: Synthesis, conformational analysis and structure-activity relationships. *Bioorganic Chemistry* **15**, 409-422. (DOI:10.1016/0045-2068(87)90036-8).

[40] Bengtsson, M., Liljefors, T., Hansson, B. S., Lofstedt, C. & Copaja, S. V. 1990 Structure-activity relationships for chain-shortened analogs of (Z)-5-decenyl acetate, a pheromone component of the turnip moth,Agrotis segetum. *J Chem Ecol* **16**, 667-684. (DOI:10.1007/BF01016478).

[41] Jonsson, S., Liljefors, T. & Hansson, B. S. 1992 Introduction of methyl groups to acetate substituted chain of (Z)-5-decenyl acetate, a pheromone component of turnip moth,agrotis segetum: synthesis, single-sensillum recordings, and structure-activity relationships. *J Chem Ecol* **18**, 637-657. (DOI:10.1007/BF00987825).

[42] Jonsson, S., Malmstrom, T., Liljefors, T. & Hansson, B. S. 1993 Enantiomers of methyl substituted analogs of (Z)-5-decenyl acetate as probes for the chirality and complementarity of its receptor inAgrotis segetum 1: Synthesis and structure-activity relationships. *J Chem Ecol* **19**, 459-484. (DOI:10.1007/BF00994319).

[43] Norinder, U., Gustavsson, A.-L. & Liljefors, T. 1997 A 3D-QSAR Study of Analogs of (Z)-5-Decenyl Acetate, A Pheromone Component of the Turnip Moth, Agrotis segetum. *Journal of Chemical Ecology* **23**, 2917-2934. (DOI:10.1023/A:1022531531072).

[44] Rowland, S. J., Sutton, P. A. & Knowles, T. D. J. 2019 The age of ambergris. *Nat Prod Res* **33**, 3134-3142. (DOI:10.1080/14786419.2018.1523163).

[45] Kovatcheva, A., Golbraikh, A., Oloff, S., Xiao, Y. D., Zheng, W., Wolschann, P., Buchbauer, G. & Tropsha, A. 2004 Combinatorial QSAR of ambergris fragrance compounds. *J Chem Inf Comput Sci* **44**, 582-595. (DOI:10.1021/ci034203t).

[46] Keller, A. & Vosshall, L. B. 2016 Olfactory perception of chemically diverse molecules. *BMC Neurosci* **17**, 55. (DOI:10.1186/s12868-016-0287-2).

[47] Keller, A., Gerkin, R. C., Guan, Y., Dhurandhar, A., Turu, G., Szalai, B., Mainland, J. D., Ihara, Y., Yu, C. W., Wolfinger, R., et al. 2017 Predicting human olfactory perception from chemical features of odor molecules. *Science* **355**, 820-826. (DOI:10.1126/science.aal2014).

[48] Li, H., Panwar, B., Omenn, G. S. & Guan, Y. 2018 Accurate prediction of personalized olfactory perception from large-scale chemoinformatic features. *Gigascience* **7**, gix127. (DOI:10.1093/gigascience/gix127).

[49] Enamine. https://enamine.net/library-synthesis/real-compounds/real-space-navigator.

[50] Sterling, T. & Irwin, J. J. 2015 ZINC 15--Ligand Discovery for Everyone. *J Chem Inf Model* **55**, 2324-2337. (DOI:10.1021/acs.jcim.5b00559).

[51] Block, E. 2018 Molecular Basis of Mammalian Odor Discrimination: A Status Report. *J Agric Food Chem* **66**, 13346-13366. (DOI:10.1021/acs.jafc.8b04471).

[52] Massberg, D. & Hatt, H. 2018 Human Olfactory Receptors: Novel Cellular Functions Outside of the Nose. *Physiol Rev* **98**, 1739-1763. (DOI:10.1152/physrev.00013.2017).

[53] Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402. (DOI:10.1093/nar/25.17.3389).

[54] Boratyn, G. M., Schaffer, A. A., Agarwala, R., Altschul, S. F., Lipman, D. J. & Madden, T. L. 2012 Domain enhanced lookup time accelerated BLAST. *Biol Direct* **7**, 12. (DOI:10.1186/1745-6150-7-12).

[55] Berman, H., Henrick, K. & Nakamura, H. 2003 Announcing the worldwide Protein Data Bank. *Nat Struct Biol* **10**, 980-980. (DOI:10.1038/nsb1203-980).

[56] Burley, S. K., Berman, H. M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., Christie, C., Dalenberg, K., Duarte, J. M., Dutta, S., et al. 2019 RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res* **47**, D464-D474. (DOI:10.1093/nar/gky1004).

[57] Muhammed, M. T. & Aki-Yalcin, E. 2019 Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chem Biol Drug Des* **93**, 12-20. (DOI:10.1111/cbdd.13388).

[58] Orry, A. J. W. & Abagyan, R. 2012 *Homology Modeling*. Totowa, NJ, Humana Press.

[59] Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. 1963 Stereochemistry of polypeptide chain configurations. *J Mol Biol* **7**, 95-99. (DOI:10.1016/s0022-2836(63)80023-6).

[60] Ahmed, L., Zhang, Y., Block, E., Buehl, M., Corr, M. J., Cormanich, R. A., Gundala, S., Matsunami, H., O'Hagan, D., Ozbil, M., et al. 2018 Molecular mechanism of activation of human musk receptors OR5AN1 and OR1A1 by (R)-muscone and diverse other musk-smelling compounds. *Proc Natl Acad Sci U S A* **115**, E3950-E3958. (DOI:10.1073/pnas.1713026115).

[61] de March, C. A., Topin, J., Bruguera, E., Novikov, G., Ikegami, K., Matsunami, H. & Golebiowski, J. 2018 Odorant Receptor 7D4 Activation Dynamics. *Angew Chem Int Ed Engl* **57**, 4554-4558. (DOI:10.1002/anie.201713065).