A microscopic model of signal transduction mechanisms: olfaction

Jennifer C. Brookes

November 2008

Department of Physics and Astronomy

University College London

Thesis submitted for the degree of Doctor of Philosophy

I, Jennifer C. Brookes, confirm that the work presented in this thesis is my own. Any material that has been derived from other sources has been identified as such and referenced appropriately.

Abstract

This thesis recognizes that: in many systems the initial small molecule - receptor recognition processes, and thus signal transduction, is not fully understood to the highest level of scientific explanation and prediction. One such example of this is olfaction. Molecules cannot necessarily be predicted from a smell, and similarly a smell from molecules. Better understanding of these initial steps, would have important repercussions, especially in the field of rational drug design. So in general the thesis proves the physical feasibility and potential of a novel and generic signaling model, and in particular looks at those processes in olfaction.

The conjecture 'Could humans recognize odours through phonon assisted tunneling?' is tested. This is based on the idea (Turin, 1996) that the nose recognizes an odorant's vibrations (phonons) via inelastic electron tunneling (IETS). The nose thus acts as a 'meat spectroscope'. First the background biology of the olfactory system is evaluated, then the conjecture is posed as a soluble problem. Traditional physics ideas are reconciled with the biological environment. It is proven that no physics based objections hold against the working of this new mechanism, thus a predictive and explanatory theory is now introduced to the field. The parameters of odorant discrimination are explored. In particular the 'Huang-Rhys factor' is modeled as a measure of the electron-phonon coupling integral to signal transduction. Several approaches are considered, 'odorant spectra' is created. Objections to the conjecture are considered, in particular the apparent paradox of enantiomer discrimination. The apparent paradox is shown to be obsolete. A correlation between a certain type of flexibility and whether enantiomer pairs smell the same is found. A rule is established: *The members of an enantiomer pair will smell alike (type 1) when the molecules are rigid, and will smell different (type 2) when they are flexible.* This flexibility refers to a particular property of six-membered rings. A consequence of this finding leads to the investigation of certain steroids in correlation to their bio-effects, and it is found that similar features are apparent, thus the mechanism of biological IETS is applied to other systems.

To Mini-ha-ha.

"These are the times that try men’s souls"

One, of the many, reassuring quotes Gaussian ’03 automates upon completion of a successful job (original source: Thomas Paine).

Contents

Acknowledgements

First and foremost I would like to thank my supervisors Marshall Stoneham and Andrew Horsfield, for the patience, guidance, enthusiasm, and support that I have received from both, in abundance. To Luca Turin also, much credit is due, as really he has been a third supervisor, and I have been very greedy and grateful of his time. I am sincerely thankful to have had known three very different scientists, yet all dynamic and brilliant.

Many thanks to Luca's company, Flexitral, for the CASE award support. I am further grateful to the EPSRC for financial support through the Interdisciplinary Research Collaboration (IRC) in Nanotechnology, without which I would not have had the opportunity to pursue this research.

I would also like to thank Andrew Gormanly, Ché Gannerelli and Lorenzo Stella, for helping me everytime I thought I had broken my computer, and for not letting my computer break me. I owe a lot to many colleagues who are at (and have passed through) the London Centre of Nanotechnology: Thornton Greenland, Dorothy Duffy, Tony Harker, Peter Shusko, Dave Bowler, Simon Gane, Mike Horton, Rachel Mckendry, Kyle Rogers, Denise Ottley, Alex Rutherford, Martina Avellino, Dan Wheatley, Matt Watkins, Flemming Ehlers, Andrew Kerridge, Anna Kimmel, Rathin Choudry, Wisdom Beyhum, Sascha Khakshouri, and Sukina Na. From beyond the LCN and UCL I would like to thank Gavin Vinson at Queen Mary for his interest in this work, and our resulting collaboration.

To the people in my life outside the nano-world, who have so graciously and humorously shared at times my strange obsessions; Lucy-Lee Allen, John Alderton, Karl Collins and Rebecca Mcclane in particular (who have felt the brunt of it) and who I would like to thank for their support.

To my family; my mother, father and brother; for how they have helped me in everything, ceaselessly, and all their advice and tolerance. Last but not least, to Chris Howard, who is not officially family, though he is to me now, thank you for amazing me.

Table of abbreviations

Abbreviation	Meaning
DFT	Density Functional Theory
FRET	Fluoresence Resonance Energy Transfer
FWHM	Full Width Half Maximum
GGA	Generalized Gradient Approximation
Glu	Glutamine
GPCR	G-Protein Coupled Receptor
His	Histidine
HOMO	Highest Occupied Molecular Orbital
HR	Huang-Rhys
IET	Inelastic Electron Tunneling
IR	Infra-Red
LBD	Ligand Binding Domain
LDA	Local Density Approximation
LF	Leap Frog
LSDA	Local Spin Density Approximation
LUMO	Lowest Unoccupied Molecular Orbital
MD	Molecular Dynamics
MK	Merz-Kollman
NADPH	Nicotinamide Adenine Dinuleotide Phosphates
NMR	Nuclear Magnetic Resonance
NVE	fixed atoms N, a fixed volume V, and a fixed internal energy E
OBP	Odorant Binding Protein
OR	Olfactory Receptor
ORN	Olfactory Receptor Neuron
OSN	Olfactory Sensory Neuron
PCR	Polymerase Chain Reaction
PES	Potential Energy Surface
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SARS	Structure Activity RelationS
SCF	Self Consistent Field
SHO	Simple Harmonic Oscillator
STM	Scanning Tunneling Microscopy
TMD	TransMembrane Domain
Tryp	Tryptophan
VV	Velocity Verlet

Table 1: A table of the important abbreviations used throughout this thesis.

Chapter 1

Introduction

What *is* that smell? the answer to this question is surprisingly non-obvious. Smell is caused by small, neutral, volatile molecules (odorants) but what makes one odorant quite different from the others? What *defines* an odorant molecule? There is currently no generally accepted explanation. This is quite a contrast to the scientific explanations we have for other stimuli-responses. In seeing, hearing or tasting, for example, a scientist can predict (and imagine) the colour, pitch, flavour, of the stimuli by analyzing the wavelength, frequency or chemical groups, respectively. A scientist cannot, from analyzing the odorant, predict a smell. More often than not, even very subtle differences to any two odorants can result in very different perceptions. The 'smell' properties of odorants, whilst perceived with astonishing selectivity to the receptors 'eye', are almost invisible to the scientists 'eye'. Design and manufacture of perfumes in industry thus rely on intuition and trial and error, which is an inefficient enterprise, with expensive results.

This problem is exasperated by the elusive nature of the smell receptors. They are notoriously hard to stabilize and examine for their structural data, and so far efforts have been unsuccessful. A lot of imagination then, is required in modelling reasonable interim steps between odorant inhalation and turning the receptor 'on'. An advantage to this however, is that imagination can be used and may be more successful, where relying on old prejudices (see 'The lock and key model' in Chapter 'Current thinking') have failed. Further, the ultimate aim is to predict effect from cause, *not* using the receptors (but using hypothesis) and so it is reasonable that we work without them. That said 60% of pharmaceutical drugs target the G-protein

coupled receptors (the class of protein to which the olfactory receptors belong) and so there is an obvious and pressing need to experimentally probe and theoretically establish the workings of these receptor-proteins.

Signalling transduction is not only mysterious in olfaction, but in *many* ligand (drug)-protein interactions. Steroids for example, exhibit surprising peculiarities similar to those exhibited by odorants. For example: very subtle differences across stereo-configuration (such as the judicial positioning of just one hydrogen atom) can wildly alter hormone effects. This problem is thus not only an interesting nano-scale mystery, but within the answer (potentially) lies many biologically invaluable solutions. There is clearly an under-appreciated requirement that the first steps in ligand-protein interactions are better understood. Whilst drug developments are successful and should not be under-rated, like in perfume design, they require time (years) and very rarely do they satisfy the initial results desired, and very often they come with undesired side effects. A new drug is required *yesterday* and so often hypothesis driven work is skipped over. Reconciling traditional physics with what we know of biological systems, as we do here, could therefore be invaluable in determining working hypotheses for an overlooked problem.

This thesis seeks a possible solution to the problem of signal transduction in olfaction. We conjecture that signal transduction occurs via a phonon excitation in the odorant: Turin's idea in 1996 [1]. The first three chapters give the necessary background; a short history of hypotheses so far, a biological background, and a physical picture to set up the scene. The next two chapters describe the theoretical methods I used for the investigations I made. The following three chapters: the investigations, begin with assessing the feasibility of the conjecture. We develop and establish a feasible new model for signal transduction. The second investigation explores the necessary parameters involved in the previous chapter, and develops a possible identification algorithm: odorant spectra, that could be used predictively. The third and final investigation analyzes a common objection to Turin's proposal, and in doing so exposes a quite independent observation, that may shed light on receptor sensitivities. The last chapter summarizes the findings.

Chapter 2

Current thinking

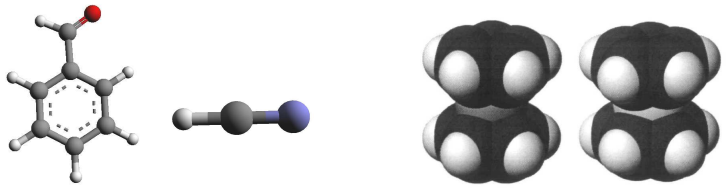
2.1 The "lock and key" model

In 1967 Moncrieff conjectured that the response to scent is initiated by a sort of *lock and key* mechanism [2]: to produce a particular scent a particular fit is required between the odorant and receptor. Emphasis falls on the receptor "feeling" the molecule's shape as opposed to interacting in any physically meaningful way with its functional groups. The idea is borrowed from the types of action exhibited in enzyme behaviour. Amoore wrote that a stereo-chemical mechanism of scent explains why some people exhibit specific *anosmia*- where some people cannot smell certain things due to the lack of or damage to a particular receptor[3]. The anosmia would be due to some blockage or a defective receptor that neglected to accommodate a particular odorant. A more flexible modification of this is the *hand and glove* idea, *i.e.*, that the whole system distorts for a more acceptable fit. There is plenty of evidence for cases where structure does seem important to an odorant's detection. For example, Yoshii, Hirono, and Moriguchi show that ethyl citronellyl oxalate (at a certain conformation) and cyclopentadecanolide are virtually structurally the same, and this explains their similar scents [4]. Further, studies by Araneda *et al* , in 2000 show that octanal and aldehydes with greater than 7 carbon atoms will activate the same cloned receptor, OR17, but shorter chain aldehydes will not [5]. These examples show it is clear that there are some cases where molecules with the same shape produce the same scent, but is this coincidence or rule? And after fit, what stimulates the signal? In addition there is even evidence *against* a structure-fit theory,

for example:

1. Odorants that smell the same whilst being structurally very different, see figure 2.1 a) of benzaldehyde versus hydrogen cyanide.

2. Odorants that smell different whilst being structurally the same, see figure 2.1 b) of ferrocene versus nickelocene.



(a) Smell the same (b) Smell different

Figure 2.1: a) Odorants that smell the same whilst being structurally different, benzaldehyde (left) and hydrogen cyanide (right), both smell of bitter almond. Bitter almond is a striking example of this kind- the scent is recreated in 75 structurally different molecules [6]. b) Odorants that smell different whilst being structurally the same, space filling images taken from Turin (1996) [1]. Two cyclopentadienyl ion complexes; ferrocene (left) and nickelocene (right) are almost exactly the same size and shape save for the differing metal ions tucked inside their structure. This would imply (according to shape theory) that functional groups cannot be smelt when they are hidden, which is not the case, ferrocene and nickelocene smell different; spicy versus oily/chemical, respectively.

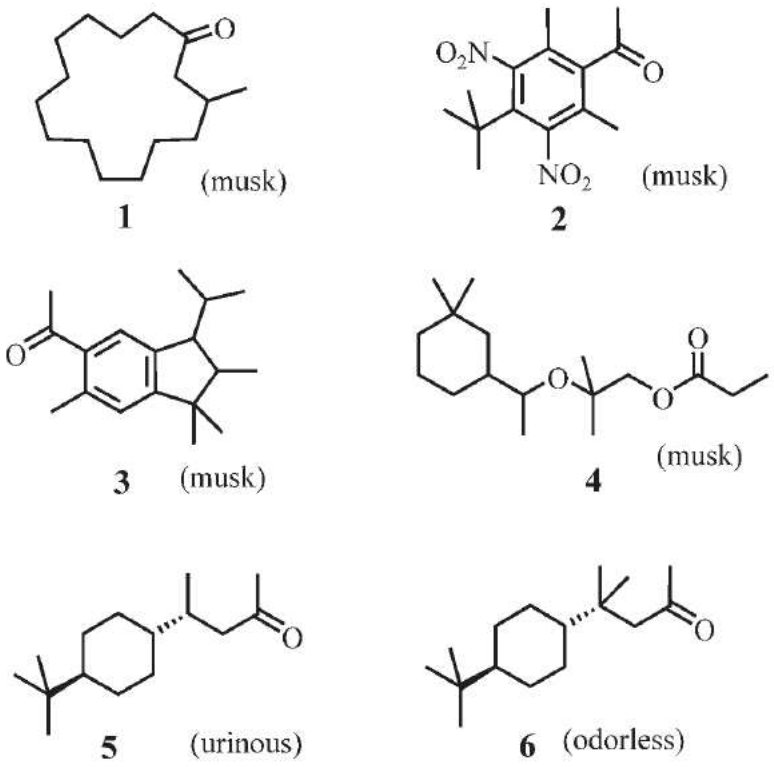
There are many more examples, see figure 2.2, that exemplify the 'unpredictability of scent, as discussed in Sell's 2006 paper [7].

There are some intuitive problems with lock and key also:

1. Mechanical manipulations of soft solids are unlikely, biological systems simply do not work like a lock and key.

2. The molecule's shape will alter due to thermal fluctuations, so precise conformation at the site may be compromised.

Further; if the system is mechanical, what inspires the molecule to leave the receptor? Lock and key does not explain the action that follows fit, unless we attribute



RT

Figure 2.2: The 'unpredictability of scent'; more examples of odorants that smell the same but are structurally different and *vice versa*, taken from Sell (2006), [7].

these particular interactions to random thermal fluctuations. There also appear to be far more different scents than there are receptors, which is in contradiction to the theory which presumes one special receptor per odorant. If this were true a combinatorial explosion would occur, and one that couldn't be accommodated by the anatomy in the nose.

2.2 The odotope theory

This theory crystallizes from the assertion that there must be *some* sort of molecular shape recognition. The odotope theory (see Mori and Shepard 1994) requires that the shape of a molecule is important for the initial reception, though the receptor is not particular to *exact* fit [8]. It may be one structural feature, such as that of the functional group as opposed to the main body of the structure, that a particular receptor responds to as opposed to a general shape, for example a carbonyl or ether group. Therefore the binding with the receptor is dependent on the electronic features of particular groups and how they bind with a certain affinity to a certain receptor. It is intuitive because our noses act like detectors, identifying certain atoms, placing the importance on the atoms present rather than the position of the atoms. However, again this theory gives too many combinatorial probabilities, and is subject to similar insecurities as those above. If the receptors are not particular there will be too many signals, and noise.

2.3 The vibrational theory

First proposed by Dyson (1938) and then Wright (1977) was that the signature of scent is due to a molecule's unique vibrational spectrum as opposed to its structure, [9], [10]. Unique scents are attributed to a unique spectrum in the same way a colour is associated to its frequency of light. According to Turin and Yoshii (2003) no two odorants have ever been found to have exactly the same odour [6]. Characteristic spectra for each molecule would explain this statement. Besides this, there is an intuitive appeal that scent may be a spectral sense, similar to hearing and seeing. Evidence against this theory is given to be the differentiable smells of mirror image related odorants: because they have identical mass and identical spectra when held in an achiral solution, how should they be discriminated in the nose? This is

a *prima facie* evaluation, however, and with a second look we show this evidence is obsolete: see the Chapter 'Chirality and conformations in odorants'. The real detriment to the vibrational theory, in its infancy, was that Dyson was unable to hypothesise a suitable biological mechanism that would act spectroscopically. In the lab we detect a molecules vibration by probing the molecule at the right frequency with an infrared (IR) source. Scientific scepticism arose from taking this spectroscope idea very literally; scientists were unable to envisage a translation of a metal and glass spectroscope into "meat" materials. It wouldn't be possible for our noses to direct IR radiation from a molecule held in some sort of receptor-like-cryostat. So problems arose such as; what would it actually be made of? How could the information be scanned? Optical spectroscopy would be impossible. Wright spoke of thermal vibrations and vibration sensors, being detected mechanically. However, such a mechanism would only occur at below 500cm^{-1}, conflicting with the evidence of strong smelling substances such as hydrogen cyanide -which vibrate above this range [1]. It seemed the mechanism in its bear bones would not be facilitated by the nose's materials.

2.4 Turin's model

In 'A spectroscopic mechanism for primary olfactory reception', published in Chemical Senses 1996, Turin introduces the idea of a working "meat machine". He proposes that the mechanism of detection is biological inelastic electron tunneling (IET), and points to various receptor features that would constitute a spectroscope. Turin postulates that signal transduction is reliant on the success of an electron tunneling from a donor (D) state to an acceptor (A) state, facilitated by the odorant because the energy difference between these states matches a mode of vibration in the odorant. Once at A a chemical reaction is initiated, likely because of a reduction, due to the flow of electrons and then signal activation. This requires that olfactory receptors (ORs) are surrounded by suitable reservoirs for sinks/sources of electrons, and the receptor acts similarly to metal electrodes in a junction. Across a junction current flows through an insulating barrier, from metal 1 to metal 2 (see figure 4.6 in Chapter 'A physical picture') due to the phenomenon of tunneling. Drawing on this analogy to conventional IETS, the electrons may tunnel *elastically* across the gap and give a constant voltage signal- but only if there are degenerate D and

A states and a decent overlap between the filled and empty electronic states [1]. Such elastic electron tunneling spectrosopy (ETS) is observed in semi-conductors, between the valence and conduction bands. If there is a single molecule between the electrodes, however, partial charges on the atoms of the molecule scatter the electron as the electron tunnels, corresponding to the molecule's vibration. As energy is lost the electron tunnels *inelastically*, but, only if the splitting between D and A match the corresponding vibrational excitation of the odorant, again assuming a decent overlap of states. Again, such inelastic electron tunneling spectroscopy (IETS) is observed in semi-conductors. The inelastic contribution can be extracted from differentiating the current-voltage relationship twice. Thus an elastic signal corresponds to negative identification and an inelastic signal corresponds to positive identification of a molecule. Turin proposes that olfaction signalling works in similar ways. Though of course there must be a number of adaptations and to name a few :[1]

1. The theory assumes there is no continuous 'scanning' but discrete energies are picked up where they are activated, and there is a system of certain 'tuned' receptors [6]. This may mean that the nose works in a similar way to colour vision and hearing where receptors are broadly, not specifically tuned, and they overlap to cover the whole spectrum, thus giving more complex grades of colour, harmony, scent.

2. There may not be an electron flow. In conventional ETS a current will flow, but the transfer of one electron may be enough in olfaction to initiate events.

3. The energy states will be different in the receptor as opposed to the Fermi levels in metal junctions.

4. The temperatures will be different. The advantage of conventional IETS (an advantage not enjoyed in the nose) is the junction can be cooled to liquid helium temperatures, thus improving resolution by reducing temperature broadening and simplifying transitions.

5. Usually spectra analysis involves extraction of a small inelastic component from a dominating elastic component. Conversely, in olfaction, it must be that

[1] and there will be more adaptations to this first model!

ETS is suppressed relative to IETS, or there would be signal saturation to the brain.

Probably the strongest test of Turin's theory is provided by isotopes. In isotopic substitutions the structure is controlled, leaving it the same (same bond lengths and angles, and roughly the same volume) whilst crucially altering the overall mass of the compound and thus the IR spectra. Just replacing hydrogen with deuterium is an effective way to actually shift the C-H stretch frequency notably from $3000cm^{-1}$ (for hydrogen) to $2200cm^{-1}$ (for deuterium) see Turin (1996) [1]. A couple of psychophysical studies have been made to refute/support the predictions made from vibration-related theory, see Keller and Voshall (2004) and Roberts (2006), for the conflicting negative/positive results respectively, [11], [12]. For further *support* on isotopic substitution in a range of benzaldehydes see Haffenden, Yaylayan and Fortin (2001) where they show by sensory analysis and IR spectra that certain shifts in the vibrational modes correspond to an alteration in scent [13].

Other interesting phenomena that support Turin's theory include:

1. Physiology studies of rodent olfactory receptor neurons (ORN) have shown that olfactory cells can respond to many odorants that are *not* structurally similar (see Malnic *et al* 1999; Tareilus *et al* 1995, and see Rawson and Gomez 2002)[14] , [15], [16].

2. The carbon chain aldehydes differ in scents according to the number of carbons present. For even carbon atoms the smell is orange, for odd, floral/waxy (see Turin 1996) [1]. This cannot be explained by structure. You might expect the smells to differ incrementally, not alternately. But it can be explained by the vibrations because the evens vibrate differently to odds in alternate patterns.

3. In sila compounds (see Wannagat *et al* 1993), if carbon atoms are swapped with silicon; significantly altering the partial charge and mass, whilst keeping the types of bond the same, then scent is changed [17]. For example dimethylcyclohexane and dimethylsilacyclohexane smell camphoraceous and bleachy respectively [18].

Turin's theory has better explanatory power in contrast to just-structure theories. For example, it explains why some odorants appear to be odourless; odour free

molecules are those that are a) small molecules such as hydrogen as they have no overall partial charge[2], and/or b) molecules that vibrate out of the detectable range or do but weakly and/or c) molecules that are simply too big to get into the receptor site. Also Turin's theory proposes a *mechanism* and introduces a way to quantify scent signals. Despite all this it is simply interesting to test and develop. For these reasons this proposal is the focus of this thesis and, as we shall see, yields interesting results and observations, not to mention the development of a new paradigm briefly introduced below.

2.5 The "swipe card" model

Any link or rule between the structure of an odorant and its odor are at best tenuous. An obvious problem with the structure theory is that it is unreliable, because evidence refutes it. Also, there is no predictive or explanatory power. We wish to seek a more scientifically robust model for signal transduction. It is obvious that shape is important, at least in order to enter and fit into a receptor site. However, it does not seem to be enough. Enter here the *swipe card* metaphor. Like the lock and key, a good enough fit is required to get the swipe card into a lock, but it is not just fit that opens the door: some internal information must be read, such as an electromagnetic strip. In the case of the odorant-receptor interaction this information may be the quantum mechanical vibrations of the odorant. This paradigm was first purported in 1993 by Stoneham *et al* for the neurotransmitter serotonin and its receptor [19]. This thesis describes in more detail the swipe card paradigm for signal transduction in olfaction: see the Chapter 'The olfaction model' for more.

[2]Exceptions to this rule are halogen diatomics such as chlorine- though it is supposed by Turin that these react with biomolecules upon inhalation such that it is the product that we may be smelling [1].

Chapter 3

A biological background

This chapter aims to describe *what* constitutes the system where smell happens, so at least there are parts to put in the puzzle towards explaining *how* smell happens. It provides context for the molecular-level studies in later chapters. We take a top down approach and look at the anatomy from the whole, the parts and then down to the very small: the olfactory receptors (ORs). I will concentrate mainly on the odorant-receptor (ligand-protein) interactions, as this is the main interest of the thesis. Thus this chapter provides the background information on which we base the problem: how does signal transduction occur given the materials it has? Is it possible to reconcile a conventional physics idea with a soft biological environment? Is the anatomy sufficient to support or at least not inhibit the possibility of IETS in the nose?

3.1 The whole: olfaction

Smell is sometimes considered the most intimate sense, because once the odorant is inhaled, our central nervous system (CNS) is then almost directly interacting with the external world. For this to happen the odorant takes a journey. The first part of the odorant's journey, once inhaled, is to meet the *olfactory mucus*. The mucus covers the *olfactory epithelium* and is about 10-40μm thick [20]. The purpose of the olfactory mucus is not obvious, though most likely it simply moderates the concentration of odorants reaching the epithelium; it is said to act as a separation column [21]. A recent discovery is that the mucus layer contains *odorant binding proteins*

(OBP's), *i.e.*, small lipocalin carrier proteins, that seem to play a contentious role [22]. It has been found that these OBP's have a high affinity for aldehydes and large fatty acids [23], it thus seems likely their role is to transport the odorants (largely hydrophobic) across this wet mucus layer to the epithelium. However, it is said also, respiratory cilia (non sensory) embedded in the nasal mucus help to move the odorant molecule to the required site. Beyond the epithelium lie the *olfactory sensory neurons* (OSN), that traverse the epithelium and extend into the mucus terminating in cilia (sensory) that projects from the *dendritic knob*. The cilia are typically 10-50 for each neuron and about $0.3\mu m$ in diameter [24]. The number and type of OR's on the cilia depend upon the species. Each OSN type has cilia covered in a corresponding OR type, of which there are an estimated $\sim$347 [25]. The axon of the OSN's are wrapped in an ensheathing layer, which is famous for its regeneration ability. This regeneration occurs on a timescale of every 30-45 days. This property has been exploited in repair of certain parts of the CNS [26]. An OSN is a bipolar neural cell, depolarization of this cell is discussed in more detail below. The meeting of the odorant with the OSN represents the crucial point where atomic scale detection is converted into an electrical signal. The axons of the OSN's project through the cribriform plate to the *olfactory bulb* (OB). In the bulb, neural axons route to one of ~2,000 globular structures called *glomeruli* (see Buck and Axel, 1991, [27]). Glomeruli are discrete loci on the olfactory bulb. The location they extend to on the brain, are the same in all subjects, for each type of OR. There are four zones of expression in the epithelium that match four zones in the olfactory bulb. It has been found by *polymerase chain reaction* (PCR) -that there is one receptor encoding gene per olfactory cell, [27], [28], [14], [29]. The cell that projects a signal then depends on the response of that receptor, or genetic differences in the neurons and the receptor protein [29]. Each glomerulus receives input from OSN's expressing just one OR, there is thus convergence of information from 10's of 1,000's of OR's to ~2,000 structures on the OB. The direct non branching route from receptor-type to OSN type, and OSN to glomeruli type is referred to as 'zone-to-zone' mapping [30]. Examining the spatial pattern of activity in the olfactory bulb indicates which receptors have been activated. Individual odorants could be identified by their spatial projections onto the OB. What then, happens to the odorant? Perhaps the OBP's clear them away after use. Or the odorants can be degraded or modified at the epithelium by enzymes so that the smell is eventually cleared. It has been pro-

posed that *biotransformation enzymes* act by making the odorant lipid insoluble[1] (all lipids are hydrophobic) and therefore unable to remain bound to the receptor [22].

3.2 The parts: the olfactory epithelium

The olfactory epithelium or *neuroepithelium*, is 30-200μm thick with an area of 2-20 cm^2, depending on the species of vertebrate [30]. See figure 3.2 to picture the epithelium. The body of the OSN is supported by the basal (stem cells) and sustentacular (supporting) cell and are accompanied by the microvillar cells (those that resemble OSN's, not shown in figure 3.2). There are 50 -100 million OSN's typically (see Reed 2004 [29] and Rawson and Gomez 2002 [16]) with an estimated 45,000 OR's that reside per cell, corresponding to a surface density of about 100 receptors per μm^2, at least for the human OR17-40, as estimated by Jacquier, Pick and Vogel [31]. Let's examine more closely the OSN's. The cilia contain olfactory receptors (OR's) which are bound to olfactory G-proteins (G_{olf}), hence they are G-protein coupled receptors, a discovery made recently in Nobel prize winning work for Axel and Buck who found the genes that encode these receptors in 2004. When the odorant binds to the G-protein receptor a G_{olf} causes a subunit G_{alpha} to be released. The G_{alpha} activates the formation of adenyl cyclase III (AC), an enzyme which in turn activates the formation of second messenger cyclic adenosine monophosphate (cAMP). Then cAMP binds cAMP-activated cationic channels and cyclic nuleotide gated (CNG) signalling is released resulting in an ion channel opening and a Ca^{2+} and Na^{2+} influx[2]. This results in depolarization of the OSN, see figure 3.1 for a summary of the events. Calcium channels initiated by CNG activate calcium dependent chloride channels (CAC). These depolarize the cell further such that a signal may be amplified by 85%. For a more detailed review see Rawson and Gomez (2002) [16]. Parallel to this sequence, there is also a secondary transduction mechanism where G-proteins couple to phospholopase C (PLC) as opposed to AC. PLC triggers the production of inositol trisphosphate (IP_3) in a similar way to the signalling mechanism used in hormone detection [32]. The IP_3 then activates IP_3-gated Ca^{2+} channels, and again there is a

[1]Though note, not all odorants are lipids.

[2]Note action potentials of the receptors can also be generated by other ion influxes such as potassium.

depolarization of the cell membrane and production of a receptor potential, though the IP_3 channel is considered unlikely to play a primary role [33] . Thus a change in receptor potential allows a neuron to fire an action potential to the brain.

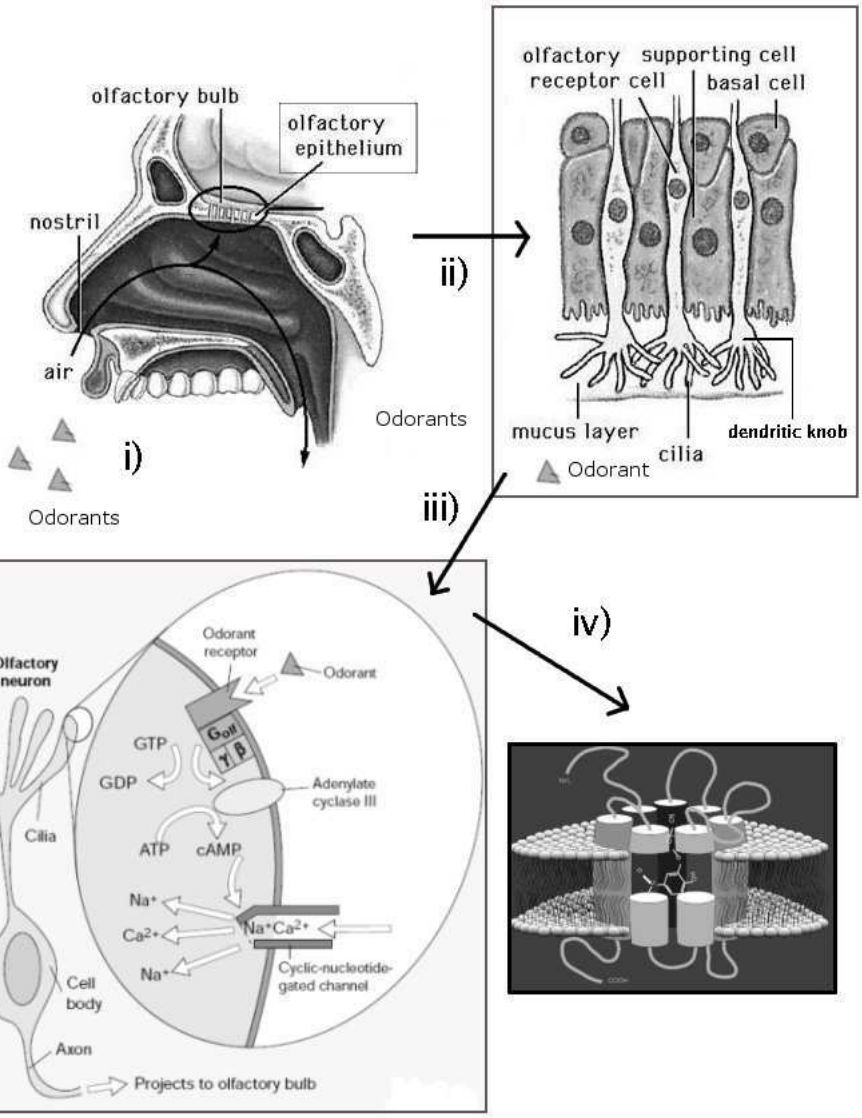


Figure 3.2: The odorant's journey, from i) the air to ii) the epithelium to iii) the OSN and finally iv) it meets the OR.

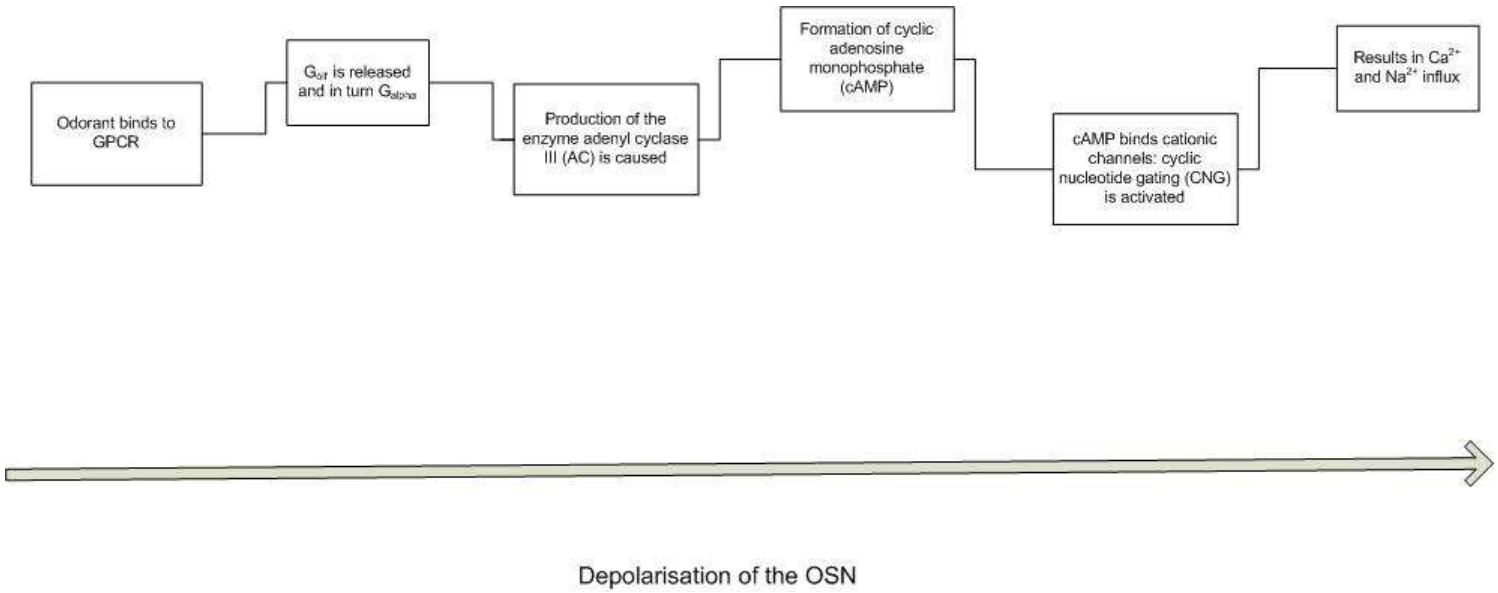


Figure 3.1: A flow chart to show the sequence of events that cause the depolarisation of an OSN.

3.3 The very small: the olfactory receptors

The opening of ion channels described above is known to propagate the odorant signal to the brain. What we are really interested in, however, are the *initiating* events, *i.e.*, what happens between a small odorant-ligand and the GPCR to cause the release of a G-protein for certain receptor types. As put by Colquhoun "How does the the G-protein on the inside of the membrane 'know' about the binding of an agonist on the outside of the membrane?" [34]. The release of G_{olf} is a necessary initial event, as demonstrated by Belluscio *et al* who show G_{olf} deficient mice are anosmic [35]. So we now examine more closely how this release is incurred. Olfactory receptors belong to a class of proteins in cells called the *membrane proteins*. It is estimated that the genes coding for membrane proteins represent 30% of *all* proteins in sequenced genomes [36], [37]. However, only 85 high-resolution structures of membrane protein have been determined so far in sharp contrast to more than 25,000 protein structures currently in the Protein Data Bank (PDB) [3]. None of these 85 are olfactory receptor structures, though they are ubiquitous, OR's are the largest multigene family in multicellular organisms [38]. Determination of these structures is made difficult by their heterological nature; the G-protein-coupled receptors (GPCR's) are notoriously hard to crystallize and get X-ray for their structure information. There is no nuclear magnetic resonance (NMR) or X-ray data for crystallized human olfactory receptors, yet. However, the discovery of the genomic sequences for OR's allow evaluation by *sequence homology*: the amino acid sequencing of the GPCR types can be compared against sequences of related proteins. This is most sensibly done against proteins which have known function and are related structurally and evolutionally to the test sequence. Bovine rhodopsin is then the obvious candidate, given it it is the only GPCR structure known very accurately, see below for more on rhodopsin. [38]. Where residue positions match well, the amino acid units are said to be conserved, where there is considerable deviation, they are said to be variable. From these comparisons, at least, models can be made, such as figure 3.3.

[3] see http://www.rcsb.org/pdb/home/home.do

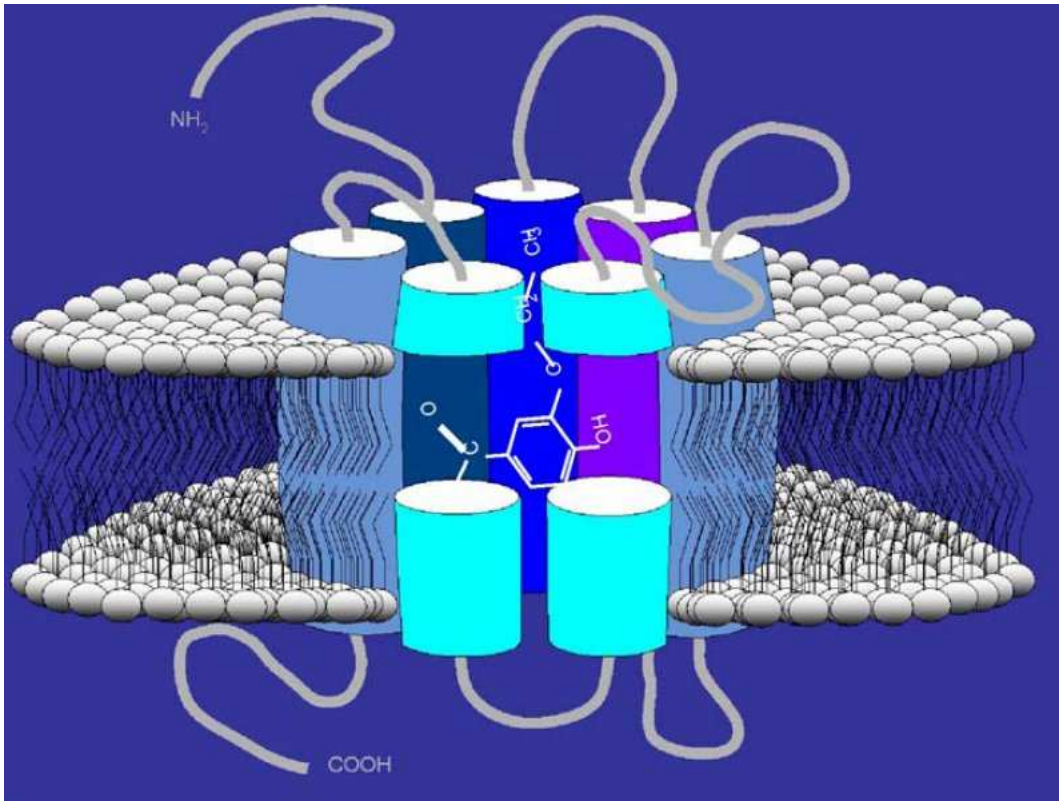


Figure 3.3: Breer's (2003) proposed membrane topology for olfactory receptors [22]. The seven lobes are transmembrane domains (TMDs) . Such a structure is supported by the evidence of the activation of G-proteins, the production of a cyclic second messenger, and the resulting Ca^{2+} produced [16].

G-protein coupled receptors are characterized by seven trans-membrane helices that cross through a hydrophobic region stabilized by a lipid layer. These helices are connected by loops, that terminate with a carboxyl group that extends into the extracellular layer and a nitryl group that extends into the intracellular layer. The loops extend into hydrophilic layers, but the main body of the protein resides within the lipid hydrophobic layer. These loops are flexible and affect the conformation of the helices. Presumably, well within the layer exists the *ligand binding domain* (LBD), though it has been conjectured that the ligand may dock at places on the loops, some odorants (perhaps the large ones) may bind to this part of the protein without entering a pocket. At the LBD, it is usual (as in other ligand-protein systems) that the specificity of the binding site is determined by the partic-

ular amino acid side chains attached to the TMD, and these determine the *ligand complementary-determining* region. As put by Breer in 2003, "it has been proposed that about 20 variable amino acid residues in TM3, TM4 and TM5 may constitute the binding pocket, the ligand complementary determining region. However,... the TM-regions may not be the only determinants for the binding specificity of olfactory receptors" [22].

At ambient temperatures the receptors are of course dynamic, and thus take on certain *conformations*. It is typical that the $\alpha-$helix (as seen here the 7 helices) confer mobility, in contrast to a $\beta-$sheet that confers rigidity. So the olfactory receptors are likely oscillating from active 'on' states and inactive 'off' states and various states in between. Active ligands thus are said to stabilize, for a certain time, an active conformation. Here some definitions for the ligand may be required [34]:

Agonist (an active ligand) - a ligand which stabilizes the *active* conformation of the receptor.

Partial agonist - a ligand which occupies the LBD just as easily as the agonist, so acts competitively, but does not stabilize a completely active conformation. Partial agonism was first observed in GPCR's [34].

Inverse agonist (an inactive ligand or active antagonist) - a ligand which stabilizes an *inactive* conformation of the receptor. They suppress the basal response level.

Thus, there are different ways in which similar ligands interact with the same receptor, for example see ferrocene versus nickelocene in the Chapter 'Current thinking'. There are two popular descriptions of these ligand-protein interactions. One is *induced fit*- the presence of the ligand induces a conformational change of the protein due to the coming together of the two parts into a united preferential geometry, induced by the presence of one another. Before this meeting the receptor resides in an inactive state. This induced conformation change is assumed to be small and local to the agonist [34]. The second idea is *conformational selection*- the presence of the ligand selects a higher energy, perhaps weakly populated, state and the equilibrium shifts in favour of this selection. Before this meeting the protein exists in several states: an ensemble of global (as opposed to the small, local) conformations in dynamic equilibrium. These two descriptions do not wildly differ and as such it

is usually assumed both are at work in some way. The second view works well in modelling single ion channel mechanisms for systems such as the *muscle nicotinc acetylcoline receptor*, for which analysis is the most reliable of all the receptors as this is such a well established system [39]; here there is direct deionization caused by ligand binding. However, for the G-protein-ligand interaction the G_{olf} interim step complicates matters, and further, conformation changes are harder to detect. There are other complications. In single ion channels there are spontaneous openings that correspond to the 'on' state of the receptor. This seems unlikely in olfaction as we do not smell anything without a source. A reasonable way of extracting this random background noise when there are spontaneous combinations (350!) receptors firing, seems unlikely. Something else that stands out for olfaction: in biology there are few pure competitive agonists, the most ubiquitous being inverse agonists [34], but what is an inverse smell? This too indicates that the ORs work via something like induced fit, where there are no inverse agonists, just partial agonists that bind at the same site but do not induce the active receptor state. Though it is conjectured, by Turin in discussion, that odorless examples simply do not get into the LBD, it should be considered that some odorants may act as competitive agonists.

3.3.1 Binding site & interactions

There is a certain paradox surrounding the ligand-receptor selectivity, as the literature refers to the interaction as at once promiscuous and selective [5]. For example Breer records that *broad* ligand specificity is exhibited; "individual olfactory sensory neurons typically actuate a variety of different odorants and each cell shows a unique order of agonist (odorant) potency, indicating that olfactory neurons are highly diverse and broadly tuned" [22] this is also supported by Malnic *et al* 1999 [14]. Conversely individual odorants may respond to a selection of receptors [22]. Due to this broadness in specificity, molecules of nearly identical structures are recognized by different but overlapping sets of receptors" [22]. Thus ORs may act like a set or 'alphabet' of different frequencies and there are certain combinatorial codes for certain scents [13]. For more complex scents there are more letters in the code. As there are hundreds of receptor types [27], and 10,000 types of odours [1] it is very unlikely there is a unique OR for each odorant molecule. So there is no reason to presume that several odorants are not able to fit any one receptor, one at a time. It has been found that, *in vivo* and *in vitro*, receptors do indeed respond to more

than one odorant [6]. However, as some receptors accommodate many structurally related odorants they discriminate against many, in ways that do not just correlate with shape. This evidence is also supported by Sicard and Holley, (1984) [40], and see figure 3.4.

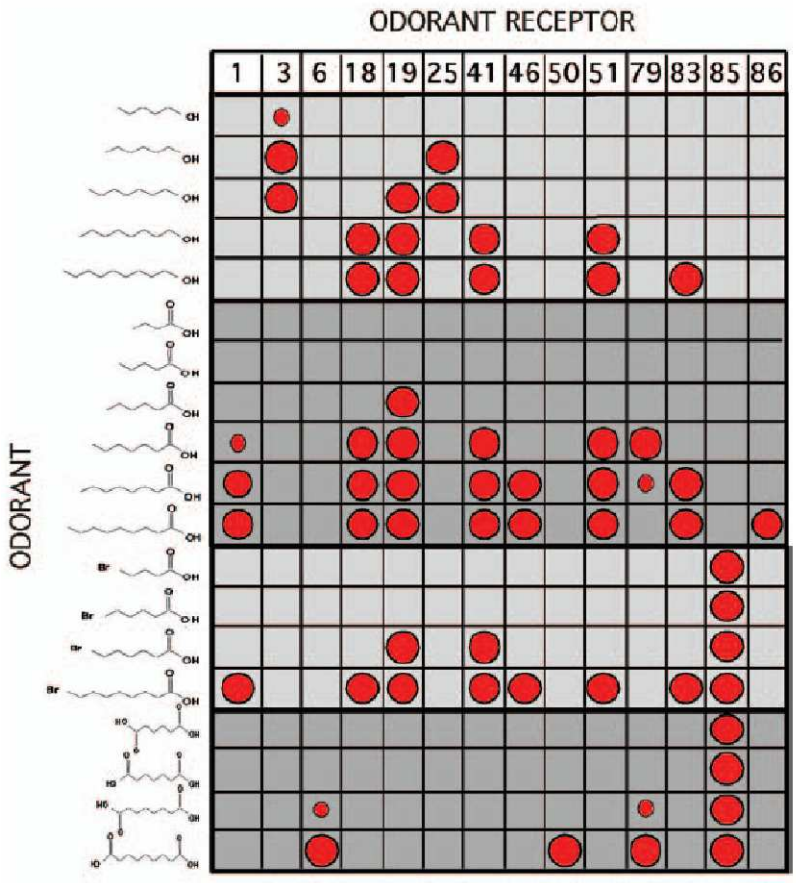
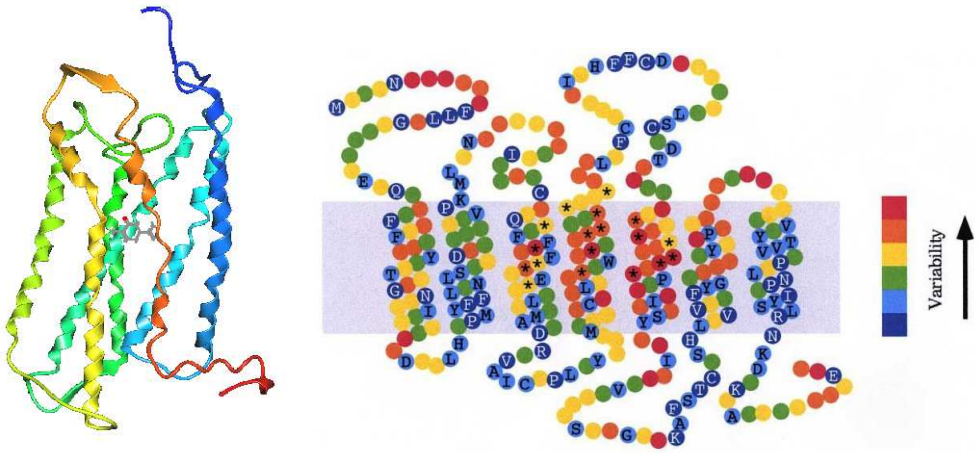


Figure 3.4: Odorant receptors used combinatorially to identify odorants and determine them. The diameter of the circles represent intensity of response [25]. Responses where determined from calcium imaging and single cell *reverse transcription polymerase chain reaction* (RT-PCR) [14].

Presumably for the ~347 receptor types, the binding sites differ, though there may be constant key elements. Determination of even just one binding site has not been achieved[4]. However, estimates can be made (for example see the paper by

[4]For this crystallization is required (with the odorant-ligand).

Arenada *et al* [5]). Usually this uses homology modelling: the amino-acid sequence is aligned against the bovine rhodopsin crystal structure template and this gives a basis for molecular modelling. Arenada *et al* suggest a putative binding site for the mouse receptor that responds to eugenol 'MOR-EG', derived from constructing a model based on the rhodopsin crystal structure template from Palczewski *et al*'s work [41]. A cavity search was performed on this structure to narrow down to 26 important amino acids. Docking analysis then identified 10 amino acids that make hydrogen bonds and hydrophobic contacts and so are likely important in recognition [5]. While it is assumed there is one binding site, however, we don't even know how many binding sites there are: the GPCR'S may be allosteric and work via 2 binding sites, and possibly more. Examination of Leffingwell's putative olfactory receptor structures [42] show helices ~30Å long within the lipid region and distances between the TM helices can be up to 20Å; see figure 3.5 for an appreciation of dimensions. Binding site evaluation is thus a possibly exhausting job, considering; i) the possible 350 types of OR, ii) the 100's of residues that may be involved and iii) the possible multiple interactions. Thus X-ray snap shots of where the odorant is most likely to reside would be invaluable to complement the modelling of the binding-site interactions.



(a) Putative OR

(b) Homology model of OR

Figure 3.5: a) Putative olfactory receptor from Leffingwell [42], with carvone odorant docked within the hydrophobic domain and b) homology model from Fuchs *et al,* shaded box indicates the membrane. Amino acid variability found for multiple alignment of 224 ORs, colour indicates blue for conserved, red for variable. Asterisks indicate hypervariable residues and single letters indicate the most conserved residues [38].

So odorants are typically small (about 4-10Å, long/wide), light (the molecular weight is typically less than 300 Daltons), hydrophobic, neutral and stable. So what do ligands (10's of atoms) do to the receptor (1-10,000's of atoms) to produce such impressive results? What *can* it do? I briefly discuss the possible importance of hydrogen bonding, complexation, and 'orbital steering'. At the LBD it is assumed that the molecule binds to the receptor via hydrogen bonds and van der Waals interactions (about 0.01eV in strength) though exact binding interactions are not known. Hanada *et al* showed in 1994 [1] that the difference between the smell of the enantiomers of carvone tended to disappear at higher temperatures, suggesting that van der Waals forces hold the molecules in certain positions in the receptor, thus supporting the importance of hydrogen bonds. Correlated mutational analyses have showed that, for 'OR5', Lys164 and Asp 204 are essential in ligand binding [43], [44], [45], most likely because these make useful hydrogen bonds to the odorants. To those authors' knowledge, the amino-acids are only considered important in the

context of stabilizing a ligand at the site with hydrogen bonds, possibly limiting translation and rotation of the odorant. Other things to consider for discrimination by frequency, in the light of the above sections, should be noted. One is the effect of hydrogen bonds. They may adversely affect the frequencies in three respects; they may *lower*/*raise* the fundamental frequencies, *e.g.*, the frequency of OH is lowered by electron withdrawing groups and the *intensities* and *widths* of bands may be altered. Nakamoto *et al* report that bent (as opposed to planar) hydrogen bonds, can quite drastically shift the frequencies higher. Further, there is an effect on charge and coordination number to consider. Transition metals often play an important role in ligand stabilization. These metals are observed coordinated in *metalloproteins* at specific environments determined by the side chains of certain amino acids. For example Tyrosinate (Y) is a common 'hard' donor to the complexation of Fe, and Cysteine (C) is a 'soft' donor ligand to Fe. Heme, for example is an Fe-porphyrin complex, it is the active site in *cytochrome c oxidase* which reduces oxygen to water via electron transfer. The most likely role for metal ions is to control a protein conformation as observed in rhodopsin [41], but perhaps the coordination chemistry of ligands to these metals may have an important effect on the ligands performance at the site. The coordination of a ligand about a metal ion could induce a conformation that is determined by *orbital steering*. Orbital steering is a concept first postulated by Koshland and Dafforn [46] to explain the facilitation of enzymatic reactions. Reactions of this sort are kick-started by aligning (steering) the orbitals of substrate and catalytic groups towards the transition state geometry, encouraging the progress of the reactants into the products; this preferential binding lowers the transition state free energy. This is a possibly important interaction for molecular recognition of odorants.

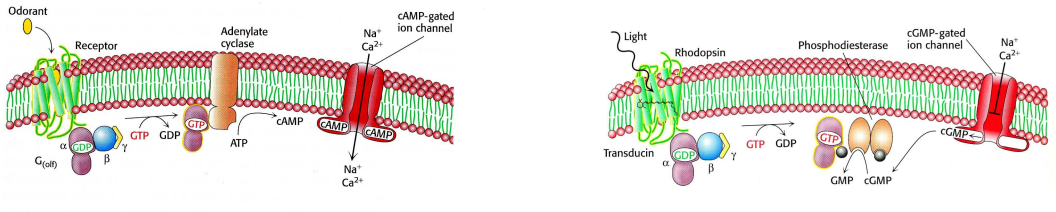
3.3.2 Affinity versus efficacy

One of the problems with the current structure-activity relationships (SARS) is that binding (affinity or selectivity) and activation/gating (efficacy or productivity) are assumed to be independent when they are not necessarily. So are differences in activity due to affinity or efficacy? Do the variations in OR structure affect binding or actuation? Different odorants may be faster at diffusing through the mucus and reaching the OR sites, some may bind more strongly to the odorant site, and some may bind for longer. All of these things may impart pronounced affects on

overall bioeffect, but it is not clear that any of them can be defined as just affinity or efficacy. Efficient binding does not necessarily follow efficient actuation and *vice versa*. There are common quandaries. How for example would a receptor discriminate against an odorant much smaller than its cognate ligand, when there is no excuse of steric hindrance? The smaller molecule should be able to bind (if it has the right functional groups), though it does not necessarily evoke a strong response [47]. See figure 3.4 for examples of small odorants with the right structure, but they do not evoke a response, notice the discrimination receptor 1 has for the shorter chain carboxylic acids, but is this because they lack affinity, efficacy or simply the right partition coefficient (which is in turn part of affinity)? There is also an intensity/concentration conflict. It is usual (in single ion channel mechanisms) to observe a near linear relationship between the concentration of ligand and the intensity of the bio-signal, up to a point of saturation. However, in olfaction, a change in odorant concentration can change the receptor code and so its perceived odour. As Buck says "at higher concentrations, additional ORs were invariably recruited into the odor response" [25]. Another issue is *inverse* affinity, *i.e.*, the odorant must eventually leave the LBD. GPCRs exist in two affinity states towards agonists [48]: high affinity when coupled to the G-protein but low affinity for release, what explains this?

3.4 Parallels: olfaction and vision

We compare the olfactory receptors to the light detecting receptor rhodopsin because it is also of the GPCR class, and for the reasons above, further, the initiating step in vision is very well understood.



(a) Olfactory signal transduction cascade

(b) Visual signal transduction cascade

Figure 3.6: Compare the signalling mechanism in visual processes with scent recognition, a) the binding of the odorant results in the *opening* of cAMP-gated ion channels, b) light induced activation of rhodopsin also leads to the *closing* of ion channels [49].

Compare the two systems shown in figure 3.6. They have, at least, these things in common [50]:

1. Seven TM's.
2. Conservation of certain amino acids.
3. Conservation of loops 1 and 2 where the G-protein may bind.
4. Phosphorylation as a "turn off mechanism".
5. Conservation of amino acids at the N-terminus- structure conservation.
6. Possible binding of Ca^{2+}.
7. Attachment of sugar groups at the N-terminus.

In vision light falls on the optic nerve and the 2 (contrast the ~350) types of sensory neurons, rods and cones, are activated. These are sensitive to low levels of light and colour, respectively [32]. Rhodopsin is present, in thousands, on the outer surface of the rods and cones. There is a chromophore within the domain called 11-*cis*-retinal which binds to the rhodopsin (opsin) via a Schiff base and Lysine (lys) [32]. Upon photoexcitation the 11-*cis*-retinal, initially along the plane of the molecule, makes a transition to all-*trans* retinal, which is a conformational change that affects the whole system and the resulting closure of ion channels. The changes in

ion channels determine the membrane's potential; for example absorption of one photon leads to 1,000 ion channels closing and a change in membrane potential of 1 mV [32]. Re-converting to the *trans*-retinal from *cis* takes milliseconds, whereas the actual photoionization takes 200 femtoseconds[5] . For colour there are three primary cones for blue, green and red, which absorb a region of wavelengths, as determined by three different kinds of opsins. We see different hues because the brain mixes up the three outputs. Different amino acid sequences define the different opsins and different wavelengths absorbed. The interaction of the retinal chromophore with these groups define the maximum wavelength. Colour blindness can result from a mutation in the amino acid sequences, such that detection of different wavelengths are interpreted contrary to the norm. A similar thing in olfaction is specific anosmia. Two points to consider from this comparison may be borne in mind. Firstly, it is possible that like rods and cones, olfactory receptor cells respond over different time scales. Cones have a faster electrical response, they discern colour and detail whereas rods take account of dim background light. Receptors in the nose may have sub-types such as the rods and cones. Secondly, in rhodopsin, it is the variability of certain amino-acids of the protein structure that tune to the absorption of various photons.

3.5 Biological IETS?

Are the appropriate materials available to make Turin's "meat machine?". See figure 3.7. We require at least:

1. An electron source and sink.
2. The right energy levels.
3. A possible donor and acceptor.

In 1996 Turin suggested that *nicotinamide adenine dinucleotide phosphates* (NADPH) could act as the source of electrons. NADPH binds to a receptor via a motif of amino acids that are in fact present in the olfactory system as searched for by Turin and described in his 1996 paper [1]. At the acceptor site Turin asserts the possible importance of the presence of zinc -this is inferred from the anosmia attributed to zinc

[5]see http://www.ks.uiuc.edu/Research/rhodopsin

deficiency in the diet, which is reversible upon supplementation. Turin proposes that zinc docks the G_{olf} forming a disulphide bridge between the receptor and G-protein. Oxidation of the disulphide bridge upon electron transfer results in the release of the G-protein. Interestingly Man *et al*, as noted above, examine mice and human olfactory binding sites, and find that for human OR5U1 it is likely (as in rhodopsin) that a cysteine residue at the N-terminus of the helix TM3 participates in a disulphide bond with a second cysteine in the 2nd extracellular loop. The 2nd extracellular loop would be constrained by the disulphide bond and cover the putative binding pocket. They find 22 sequence positions are highly conserved on the 2nd extracellular loop among the putative orthologs and highly variable among the putative paralogs that they examine, and they project into the hydrophobic domain. This supports the role of a possible bridge reduction. However, this all occurs in the *extra*cellular region, not the *intra*cellular region where the G-protein is actually released, so again the problem of "how does the G-protein know?"

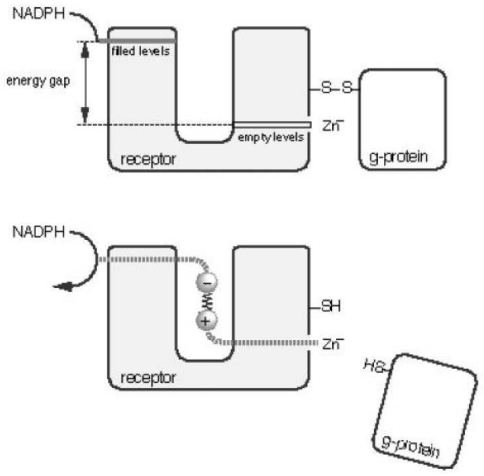


Figure 3.7: Turin's proposed schematic for biological IETS in 1996. The receptor donor takes electrons from NADPH. When there is no odorant, no electrons tunnel because there is no appropriate unfilled energy level. When an odorant (represented by the dipole) docks the electron can tunnel by exciting the vibrational mode of this odorant that matches the energy gap between the donor and acceptor. At the acceptor of the protein, electrons flow and reduce the disulphide bridge via a zinc ion thus releasing the G-protein [1].

Turin suggests that there are about 10 receptors which are broadly tuned and the signal produced is a sum of various parts of overlapping messages from the cells. Restrepo *et al* (1993) and Thüraüf *et al* (1996) have found via patch-clamp techniques that the olfactory cells have a resting potential of around -50- -65 mV and capacitance of 4pF [51], [52]. This corresponds with Turin's estimate for the required emf of the NAPDH at the receptor site. It also corresponds to the vibrational range up to 4000cm^{-1}, and so gives corresponding energy magnitudes (within the IR range for vibrational excitation of a molecule). Measuring the energy transfer for each receptor type would allow reasonable classification of types, and data to categorize odorants with. Turin proposes that each receptor can be tuned to a particular band of frequencies. Within Turin's original theory the donor and acceptor energy levels may be vulnerable to thermal broadening of the range 2kT (of about 400cm^{-1}, thus the nose spectrometer may have poor resolution of approximately this amount) as estimated by Turin and Yoshii (2003). Thus 10 kinds of receptors could be used to monitor the range 0-4000cm^{-1}, if they overlap in similar ways to the three receptors in vision [6].

There can be no electron transfer at all if the appropriate donor *filled* energy levels and acceptor *empty* energy levels are not present. Sequence homology studies such as those described above strongly suggest that a donor/acceptor is an amino acid residue. By Occam's razor, the likely explanation could be that those residues highly conserved (and possibly the same unit) provide donor and acceptor states, and the hypervariability of surrounding residues tune the D-A energy gap splitting across types, much like the receptor frequency modulation in visual processes.

Chapter 4

A physical picture

4.1 Some foundations

4.1.1 A molecule is...

In a molecule there is a collection of interacting electrons and nuclei that can be represented by the Hamiltonian:

$$H = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B<A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{4.1}$$

where i and j denote the N electrons, A and B denote the M nuclei and M_A is the ratio of the mass of the nucleus A to the electron [53]. The Laplacian operators over i and A, respectively, indicate differentiation with respect to the i^{th} position of the electron or A^{th} nucleus. Z_A is the atomic number of the nucleus A. $R_{AB} = |R_{AB}| = |R_A - R_B|$, is the distance between the A^{th} nucleus and the B^{th} nucleus, r_{iA} is the distance between the i^{th} electron and A^{th} nucleus and r_{ij} between the i^{th} electron and j^{th} electron respectively. The first term is the kinetic energy due to the electrons, the second the kinetic energy due to the nuclei, the third is the Coulomb attraction between electrons and nuclei, the fourth is the Coulomb repulsion between electrons and the fifth the Coulomb repulsion between nuclei [53]. We

need this description so we can solve the time-independent Shrödinger equation;

$$H\left|\psi\right\rangle = E\left|\psi\right\rangle, \tag{4.2}$$

and in doing so we can obtain lots of interesting physical properties for the molecule.

4.1.2 The Born-Oppenheimer approximation

The Born-Oppenheimer (BO) approximation is very important because it reduces the level of complexity by allowing us to separate nuclear and electronic motion. This can be expressed mathematically by the wavefunction [53];

$$\psi\left(r;R\right) = \psi_{elec}\left(r;R\right)\psi_{nuc}\left(R\right), \tag{4.3}$$

where r denotes all electronic coordinates and R denotes all nuclear coordinates. The electronic part depends parametrically on nuclear coordinates, whereas the nuclear part depends directly on the nuclear coordinates. This is a reasonable approximation because compared to the much heavier nuclei (by a factor of at least 1836) the electron motion is much faster such that the nuclei seem to be stationary. Thus we need only consider electronic motion, the kinetic energy term for the nuclei can be neglected and the repulsion between them considered constant, then:

$$H_{elec} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}. \tag{4.4}$$

Since, by this description, the electrons are effectively moving in a fixed field of nuclei this method is called the "fixed nucleus" approximation. Solving the electronic Shrödinger equation;

$$H_{elec}(r;R)\psi_{elec}\left(r;R\right) = E_{elec}(R)\psi_{elec}\left(r;R\right), \tag{4.5}$$

gives the electronic eigenfunctions for fixed nuclei R. Equation 4.5 indicates that the electrons follow the nuclei *adiabatically*, *i.e.*, they quickly adjust to the nuclear positions. These electronic solutions are orthogonal and normalized for integration over r at any given value of R . Then, the electronic solution can be substituted into the Shrödinger equation for nuclear motion (we reintroduce the nuclear kinetic

energy, and the constant repulsion term) corresponding to that electronic state. The motion of the electrons relative to the nuclei can be represented as an average, so that the nuclear Hamiltonian is given by:

$$H_{nuc} = -\sum_{A=1}^{M} \frac{1}{2M_A} \nabla_A^2 + \left\langle -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} \right\rangle + \sum_{A=1}^{N} \sum_{B<A}^{M} \frac{Z_A Z_B}{R_{AB}}. \quad (4.6)$$

This gives nuclei motion in a time-average field of electrons. Simplified:

$$H_{nuc} = T_N + E_{total}(R). \quad (4.7)$$

The terms after the nuclei kinetic energy (T_N) in the above equation are equivalent to the *total nuclear potential energy*; a sum of the electronic contributions plus the fixed nucleus repulsion [53]:

$$E_{total}(R) = E_{elec}(R) + \sum_{A=1}^{N} \sum_{B<A}^{M} \frac{Z_A Z_B}{R_{AB}}. \quad (4.8)$$

We can now write, similarly to equation 4.5:

$$[T_N + E_{total}]\psi_{nuc}(R) = E_{nuc}(R)\psi_{nuc}(R). \quad (4.9)$$

Solutions to equation 4.9, describe the vibration, rotation and translation of a molecule; these are quite separate from the electronic transitions. For any given set of nuclear configuration we see there are different sets of energy eigenvalues for motion. A plot of the electronic energy eigenvalue $E_{total}(R)$ against R gives the Born-Oppenheimer potential energy surface (PES) or the *adiabatic surface*. The surface is only "potential energy" by name, in spite of containing *electronic* kinetic energy, because from the *nuclear* point of view there is no kinetic energy involved in $E_{total}(R)$. See section 4.4.1, where we invoke the PES surface to describe donor and acceptor states.

4.1.3 Harmonic oscillators

There are classical harmonic oscillators everywhere: swings, pendulums and in the legs of kangaroos. Nuclei, given they are comparably massive to many other quantum particles, have motion that can be described harmonically and to a reasonable

degree classically. However, in the context of electron transfer we shall see that the quantization of energy levels in oscillators can facilitate interesting phenomena at the nanoscale.

Classic oscillators

The potential energy of a particle displaced from its minimum energy position can be described by an energy diagram as in figure 4.1. The particle is bound to the origin by a restoring force, Hooke's law, given by $F = -kx$, thus the potential energy is $V(x) = \frac{1}{2}kx^2$, where k is the *spring constant* and x is the displacement from equilibrium. This behaviour of linear simple harmonic oscillators (SHO) is characterized by the parabolic energy surface.

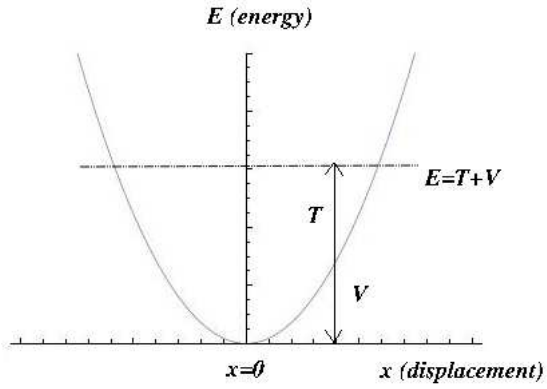


Figure 4.1: An energy diagram of the classic SHO [54].

Hooke's Law: The (elastic) force is as the extension (the restoring force is directly proportional to the displacement from equilibrium).

Analysing figure 4.1, we see that at $x = 0$, the potential energy is minimum corresponding to $F = -\frac{dV}{dx} = 0$: at any stationary point there is by definition no force acting on the particle. When oscillating, as the particle passes through $x = 0$ the energy is mainly kinetic. Far away from the origin (larger displacement), and the energy is mainly potential. At the extrema of the parabola the particle is at rest and all energy is potential. At these end points we get an amplitude of the vibration that

depends on the energy given to the oscillation [54]. The total energy is then just the sum of the kinetic and potential:

$$E(x,\, \dot{x}) = \left(\frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2\right). \tag{4.10}$$

Quantum oscillators

Quantum mechanically we can represent the contributing energies of the oscillator, kinetic and potential, with first and second terms, respectively, in the time-independent Shrödinger equation:

$$-\frac{\hbar}{2\mu}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E(x)\psi(x). \tag{4.11}$$

We remain in one dimension here for simplicity, and to map to the classical picture. The potential energy is $V(x) = \frac{1}{2}\mu\omega^2x^2$ where the angular frequency of the sinusoidal motion is given by $\omega = \sqrt{\frac{k}{\mu}}$ and μ is a *reduced mass*. This potential is found by performing a Taylor expansion of the potential energy about the displacement $x = a$:

$$V(x) = V(a) + V'(a)(x-a) + \frac{1}{2}V''(a)(x-a)^2. \tag{4.12}$$

If the parabola rests at the origin as in figure 4.1, then the term $V(a)$ is at zero, the stationary point by definition is $V'(a) = 0$ and $V''(a) > 0$ and so $V(x) = \frac{1}{2}kx^2$, the spring constant is the second derivative of the potential energy with respect to displacements. For a simple oscillator you can choose $V(a) = 0$.

To get interesting results we seek a solution $\psi(x)$ to equation 4.11. The Schrödinger equation can be simplified to:

$$\frac{d^2\psi}{d\xi^2} + \left(\frac{2E}{\hbar\omega} - \xi^2\right)\psi = 0 \tag{4.13}$$

where $\xi = \sqrt{\frac{\mu\omega}{\hbar}}x$ with a substitution and a change of variables. In the classic picture $|x|$ cannot be greater than the amplitude A; the maximum displacement. Quantum mechanical events are a bit different because there can be tunneling into classically forbidden regions. However, the more the penetration the less likely the event, so the wavefunction falls off ($\psi \to 0$) as $|x| \to \infty$. These requirements are

met by the *Hermite functions*. The Hermite functions are exponential functions with polynomial terms in them. For the first lowest energy state a solution is:

$$\psi = e^{-(\xi^2/2)} v(\xi). \tag{4.14}$$

Differentiation of this wavefunction and insertion into equation 4.13, gives the equation:

$$\frac{d^2v}{d\xi^2} - 2\xi\frac{dv}{d\xi} + 2nv = 0, \tag{4.15}$$

where $E = \hbar\omega(n + \frac{1}{2})$. This gives discrete allowed energy values where n denotes the quantum number. Here, for the lowest level, the *zero point energy,* we have $E = \frac{1}{2}\hbar\omega$. The eigenvalues are equidistant; energy levels are separated by $\hbar\omega$, a distinctive quality of the parabolic potential. We see, with the use of Hermite functions, that the stipulated boundary regions determine possible energy states, which is also true for the classic case [54]. The eigenfunctions are given by the general formula:

$$\psi_n(x) = C_n H_n(\xi) e^{-(\xi^2/2)}, \tag{4.16}$$

where H_n is the Hermite polynomial of degree n, and C_n is the normalization constant, chosen so that we have $\int_{-\infty}^{\infty} |\psi_n(\xi)|^2 \, d\xi$. The Hermite polynomials are given by:

$$H_n(\xi) = (-1)^{n/2} \frac{n!}{(n/2)!} v(\xi). \tag{4.17}$$

Figure 4.2 gives the first eigenfunctions for $n = 0 - 2$, these are illustrated in figure 4.2[55].

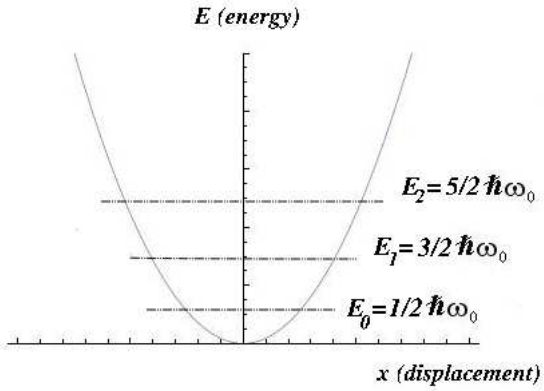


Figure 4.2: An energy diagram of the quantum oscillator.

Here we refer back to section 4.1.2 and the notion of potential energy surfaces for molecules. Plotting a PES for a diatomic molecule, for the energy as a function of atomic displacement, against these displacements, we see something like figure 4.3. This surface shows the repulsive (at short distances) and the attractive (at longer distances) forces between two atoms, and these forces are zero at their equilibrium separation, R_0, when they form a chemical bond and a molecule. If the atoms are displaced from this equilibrium they oscillate, and this can be described harmonically as we can see by a parabolic approximation to the curve. Thus we see how the harmonic representation is useful, at small displacements around the minimum. Indeed for any potential in the neighbourhood of a stable equilibrium position, we can approximate harmonic behaviour for a molecule's potential energy surface. Whilst the Lennard-Jones potentials are more accurate, the parabolic approximation is simple and accurate and so will suffice as a good approximation used throughout this thesis. The real potential energy surface of many systems may not possess symmetry, but, as long as deformations from equilibrium are small, this approximation works.

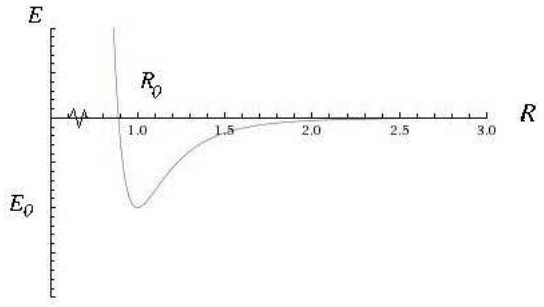


Figure 4.3: Potential energy surface (PES) for a diatomic molecule. $E_{total}(R)$ is just a function of internuclear distance.

4.1.4 Fermi's Golden Rule

Fermi's *golden rule* gives a transition rate: a probability per unit time Γ_{fi} that a transition between two *zero order states* occurs under the presence of a small perturbation $\hat{v}$. It is expressed mathematically by:

$$\Gamma_{fi} = \frac{2\pi}{\hbar} |\langle f | \hat{v} | i \rangle|^2 \rho_f. \tag{4.18}$$

The reciprocal $1/\Gamma_{fi}$ is a measure of time for the transition between the initial eigenstate $|i\rangle$ and the final eigenstate $|f\rangle$. $|\langle f | \hat{v} | i \rangle|$ is known as the transition matrix element, and ρ_f is the density of final states for electron transfer. The rate can be given by:

$$\Gamma_{D \to A} = \frac{2\pi}{\hbar} \sum_{DA} P_D \left| \langle D | \hat{v} | A \rangle \right|^2 \delta(\varepsilon_D - \varepsilon_A), \tag{4.19}$$

where the initial and final states have been replaced with the notation D for *donor* state and A for *acceptor* state respectively, with corresponding energies ε_D and ε_A for $|D\rangle$ and $|A\rangle$ respectively, this notation shall be used constantly hereafter. At thermal equilibrium the population of the initial states of the system is $P_D = exp(-\frac{\varepsilon_D}{k_B T})/Z$ where $Z = \sum_D exp(-\frac{\varepsilon_D}{k_B T})$, and is a partition function, $P_D$averages over all initial possibilities. The sum of P_D is over initial and final states and gives a den-

sity of states. A fast transition requires strong coupling between initial and final states. This coupling is represented by the transition matrix, which can, with certain approximations, be factorized into electronic and nuclear terms. Often strong nuclear wavefunction overlaps are expressed by a *Franck-Condon* factor, a feature used in describing optical absorption and radiation, and discussed further in following sections. The other important contributors to Fermi's golden rule if we examine the above equations are: the weighted sum over independent initial states and the sum over final states which combined with the transition matrix elements and the energy conserving delta function determine which states the system can evolve into. Fermi's golden rule can be used to apply to nuclear decay, radiative atomic transitions and the scattering of particles. We use it to model charge transition in an olfactory receptor, see Chapter "The olfaction model".

4.2 Vibrations in small molecules

We have seen that the nuclei and electrons in a molecule, equation 4.1, work in tandem to keep it always vibrating. Nuclei, the heavy masses of a molecule, displaced at small distances from equilibrium tend back to the starting point, under the forces of the electrons. These nuclei and their geometrical arrangements can be modelled crudely but effectively as balls on springs. We can find the nature of vibrational motion in five "back of the envelope" steps. For a general polyatomic case we have many more atoms and degrees of freedom to consider, and so the back of an envelope is not big enough, but using the simple harmonic approximation we can construct a Hamiltonian which is exactly soluble using some computer package[1]. We start with the simplest approximations before illustrating how accurate vibrations of small molecules, like odorants, can be found by computationally solving the Schrödinger equation in 3-dimensions and for many atoms.

4.2.1 Linear homodiatomics

The simplest case of vibrations in small molecules is illustrated in figure 4.4: there is only one degree of freedom, one bond length and we shall assume identical masses.

[1]Details of using Gaussian '03 to obtain vibrational frequencies is given in the investigation 'Huang-Rhys factors'.

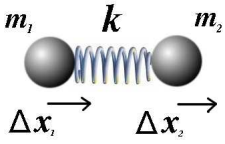


Figure 4.4: A 1 dimensional linear simple harmonic oscillator.

The natural vibrations can be calculated in five steps 1) find the equations of motion; 2) approximate solutions; 3) construct a determinant matrix; 4) find the ratios of displacement amplitude; and finally 5) normalization.

The kinetic (T) and potential energies (V) respectively are given by:

$$T = \frac{1}{2}m_1(\Delta\dot{x}_1)^2 + \frac{1}{2}m_2(\Delta\dot{x}_2)^2 \tag{4.20}$$

$$V = \frac{1}{2}k(\Delta x_2 - \Delta x_1)^2. \tag{4.21}$$

Where $\Delta x_2 - \Delta x_1$ is the extension of the spring with force constant k. We construct a Lagrange equation $L = T - V$, and insert into the Euler-Lagrange to obtain laws of motion for $i = 1, 2.....3N$ atoms, where u is the displacement:

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{u}_i} + \frac{\partial V}{\partial u_i} = 0. \tag{4.22}$$

The general equation 4.22 is of course simplified for us, as there are only 2 atoms and one direction of displacement. Thus we find two simultaneous 2nd-order linear differential equations that describe motion.

$$\begin{aligned} m_1\Delta\ddot{x}_1 + k\Delta x_1 - k\Delta x_2 = 0, \\ m_2\Delta\ddot{x}_2 + k\Delta x_2 - k\Delta x_1 = 0. \end{aligned} \tag{4.23}$$

Now Fourier transforming our equations of motion, we use a solution:$\Delta x = A_1 e^{i\omega t}$, where ω is the frequency of the oscillation, A_1 is the amplitude of the displacement, and t is the time. Inserting this solution into the equations above, we obtain the

Figure 4.5: Normal mode displacement of H_2

secular equations 4.24.

$$\begin{aligned} -\omega^2 m_1 A_1 + kA_1 - kA_2 = 0, \\ -\omega^2 m_2 A_2 + kA_2 - kA_1 = 0. \end{aligned} \tag{4.24}$$

These secular equations have a non-trivial solution when the following determinant is zero:

$$\begin{vmatrix} -\omega^2 m_1 + k & -k \\ -k & -\omega^2 m_2 + k \end{vmatrix} = 0. \tag{4.25}$$

Solving this determinant equation gives the solution $\omega^2 = \left(\frac{m_1+m_2}{m_1 m_2}\right) k = \frac{k}{\mu}$, where μ is the reduced mass as above, and so $\omega = \sqrt{\frac{k}{\mu}}$. This is the *normal mode frequency*, and we see for this simple case that there is only one, according to $3N - 5$ for a free linear diatomic. For linear homodiatomic molecules, for example, H_2, we have $m_1 = m_2 = m$ which gives the solution $\omega = \sqrt{\frac{2k}{m}}$, as before. Inserting the solution into the secular equations 4.24, yields the amplitudes $A_1 = -A_2$, these values are the *normal mode displacement*s, and we see they are equal (in frequency and amplitude) and opposite (in phase) corresponding to figure 4.5. This normal mode is symmetric and so there is no overall change in dipole moment, therefore the molecule does not absorb infrared (IR) radiation: see section 4.3.1 for selection rules. Thus hydrogen is not a very interesting spectroscopic case! The last step is normalization; if u is the displacement of the centre of mass and $u = A_1 + A_2$ (in this case there is no *overall* displacement) then it must be so that $\sum_i A_i^2 = 1$, for normalization.

Note we can obtain the same result if we use a quantum mechanical approach solving the time independent Schrödinger equation 4.11. This is formally equivalent to the classical treatment above, μ being the reduced mass as described. The

solution is thus $\omega = \sqrt{\frac{k}{\mu}}$, referring to section 4.1.3.

4.2.2 Polyatomics

If we consider a perfect lattice, the atoms at each atom site v will be oscillating about equilibrium such that $R_v = R_{0v} + u_v$, where R_{0v} is the equilibrium vector, u_v denotes the displacement vector (the extension) and R_v gives the position vector for atoms at one lattice unit cell l. v indices over 3N for the Cartesian direction vector of each atom site, $u_1,\ u_2,\ u_3....u_{3N} \rightarrow \Delta x_1,\ \Delta y_2,\ \Delta z_3....\Delta z_N$. The kinetic (T) and potential energy (V) terms are now given by:

$$T = \frac{1}{2}\sum_{v=1}^{N} m_v \left[(\Delta \dot{x}_v)^2 + (\Delta \dot{y}_v)^2 + (\Delta \dot{z}_v)^2\right] \tag{4.26}$$

$$V = \frac{1}{2}\sum_{lvl'v'} u_{lv} K_{lvl'v'} u_{l'v'} + O(u^3), \tag{4.27}$$

for small displacements of atoms. Compare equation 4.27 to the Taylor expansion before; equation 4.12. We only consider the term quadratic in displacement, the first term is taken as zero, the term linear in displacement also vanishes as the forces on the atoms at equilibrium are zero, and any higher order terms are neglected within the limit of the harmonic approximation. The remaining term shows how the different motion of nuclei are coupled. We describe the force constant matrix:

$$K_{lvl'v'} = \left(\frac{\partial^2 V}{\partial u_{lv} \partial u_{l'v'}}\right)_0, \tag{4.28}$$

this is a *Hessian* matrix. Molecular mechanics or molecular orbital calculations can be used to obtain the Hessian matrix, by first calculating the potential energy as a function of the position of each atom and then calculating the second derivative. Finding the Hessian is an essential step in some energy minimization techniques, and also essential for frequency analysis which must be done at the relaxed molecular geometry. We can exploit useful properties of the spring constant matrix if we use the periodicity of the lattice, such as $K_{vv'}^{l-l'} = K_{lvl'v'}$. Following similar steps to

the previous diatomic solution we write the equations of motion:

$$m_v \ddot{u}_{lv} = -\sum_{l'v'} K_{lvl'v'} u_{l'v'}. \tag{4.29}$$

We assume a time dependence for the displacements, and construct a solution to equation 4.29 that relates the displacements to the normal modes via a unitary transformation:

$$u_{lv} = \sum_{\alpha} m_v^{-1/2} e_v^{\alpha}(k)\, exp(i[k.l - \omega_\alpha t])\, Q^{\alpha}(k) \tag{4.30}$$

k is a three dimensional wave vector and α labels the phonon branches. As the atom sites have 3 Cartesian directions, we are dealing with 3N vectors, thus 3N modes. $e_v^{\alpha}(k)$ introduces a normalization constraint [56]. We shall call the transformation matrix:

$$\chi_{lv}^{\alpha}(k) = m_v^{-1/2} e_v^{\alpha}(k)\, exp(i[k.l - \omega_\alpha t]), \tag{4.31}$$

which includes the mass factor so that the normal modes are given by:

$$Q(k) = \sum_{lv} \chi_{lv}^{\alpha *}(k) m_v u_{lv}. \tag{4.32}$$

Changing to the normal mode coordinates is a linear transformation and has the effect of eliminating cross terms in the potential energy [56]. Expressing the displacement with a mass-weighting term via the transformation matrix eases the calculation, as we shall see. Substituting 4.30 into 4.29 we obtain:

$$-m_v^{1/2} \omega_\alpha e_v^{\alpha}(k) + \sum_{l'v'} \frac{1}{\sqrt{m_{v'}}} K_{vv'}^{l-l'} exp(-ik[l - l'])e_v^{\alpha}(k) = 0. \tag{4.33}$$

If we sum over lattice vectors $L = l - l'$, so L is the vector between the reference unit cell and a neighbouring unit cell, then:

$$K_{vv'}(k) = \sum_{L} K_{vv'}^{L}(k) exp(-ik.L), \tag{4.34}$$

and equation 4.33 can be written:

$$\sum_{v'} D_{vv'}(k) e^{\alpha}_{v'}(k) = \omega^2_{\alpha} e^{\alpha}_{v}(k) \tag{4.35}$$

where D is a *dynamical matrix* given by:

$$D_{vv'} = (m_{lv} m_{l'v'})^{-1/2} \sum_{L} K^{L}_{vv'}(k) exp(-ik.L). \tag{4.36}$$

Dropping subscripts and generalizing to matrix notation:

$$\mathbf{D}\,\mathbf{e} = \omega^2 \mathbf{e}, \tag{4.37}$$

which has non-trivial solutions when the following secular determinant is zero:

$$\left|\mathbf{D} - \omega^{\mathbf{2}}\mathbf{1}\right| = 0. \tag{4.38}$$

The eigenvalues are real and the eigenvectors form an orthonormal set as **D** is a *Hermitian* matrix. An algorithm that can be used to solve this eigenvalue equation is described in Chapter "Density Functional Theory". The 3N eigenvalues of this determinant ω^2_{α}, are the squares of the *normal mode frequencies*. The $3N \times 3N$ eigenvectors **e** give the *normal mode coordinates*. We have [57]:

$$\sum_{v} e^{*\alpha}_{v}(k) e^{\alpha'}_{v}(k) = \delta_{\alpha\alpha'} \tag{4.39}$$

$$\sum_{\alpha} e^{*\alpha}_{v}(k) e^{\alpha}_{v'}(k) = \delta_{vv'}. \tag{4.40}$$

We see for each mode of motion there is a normal mode coordinate and vice versa. There are 3N sets of **e**, but only 3N-6 normal modes for a polyatomic molecule. There are six roots of the secular determinant for which $\omega = 0$. These correspond to the translational and rotational modes of motion. There is no movement of the centre of mass in a normal mode of vibration. In the case above we see only 3N-5, as there are three coordinates for the centre of mass but only two rotational angles for the linear molecule. A normal mode describes the motion of many atoms with the same frequency and phase. Whilst their amplitudes of motion are different each atom reaches its maximum displacement and passes through its equilibrium at the

same point in time. Each normal mode acts like a simple harmonic oscillator, and one which is independent, *i.e.*, different normal modes do not exchange energy- the energy for excitation of one normal mode is fidelious to that mode.

Application to the diatomic

Referring to the matrix 4.25, we can construct a dynamical matrix:

$$\begin{vmatrix} \frac{k}{m_1} - \omega_\alpha^2 & -\frac{k}{(m_1 m_2)^{\frac{1}{2}}} \\ -\frac{k}{(m_1 m_2)^{\frac{1}{2}}} & \frac{k}{m_2} - \omega_\alpha^2 \end{vmatrix} = 0, \tag{4.41}$$

with this, as well as knowing the solution for ω and the normalization condition $\sum_v |e_v^\alpha|^2 = 1$, we can find $e_1 = -e_2$ given $m_1 = m_2$, and the fact there is just one mode. Solving the secular equation allows you to find ratios of e_v^α, the absolute values are fixed by the normalization condition:

$$e_1^2 + e_2^2 = 1, \tag{4.42}$$

so that the normal coordinates for the mode in H_2 is just:

$$Q_1 = \frac{1}{\sqrt{2}}(q_1 - q_2) \tag{4.43}$$

where q are the mass-weighted coordinates related to displacements by $q_v = \sqrt{m_v} u_v$, where $q_1 = \sqrt{m_1}\Delta x_1$, $q_2 = \sqrt{m_1}\Delta y_1$, $q_3 = \sqrt{m_1}\Delta z_1$, $q_4 = \sqrt{m_2}\Delta x_1 ... etc.$

4.3 Charge distribution and spectroscopy

4.3.1 Infrared spectra

As seen in section 4.1.3, the spacings between vibrational energy levels are to a simple approximation equidistant, and dependent on the potential energy associated with bond stretching. This is exactly so for a single harmonic oscillator. Also, that at absolute zero, the molecule has at least one-half quantum $\frac{1}{2}\hbar\omega$ and is said to have zero point motion, an quanta of excitation are extra increments in energy. This is determined by the uncertainty principle; if there was no vibration the position and momentum of the atoms would be *known* to be zero. The energy spac-

ings determined by small molecules fall into the IR frequency range 200-3,500cm^{-1}, so absorption of these frequencies correspond to vibrationally excited states. The first excited state corresponds to one quantum of vibrational energy absorption and higher energy levels correspond to multiple quanta excitation. The frequency of the quantum transition from one quantum state to the next is called the *fundamental frequency,* on the other hand *overtone frequencies* correspond to transitions between non-adjacent levels. There arises a series of frequencies that are nearly (not harmonic for a real molecule) multiples of the fundamental frequencies.

Generalization of Fermi's golden rule for harmonic perturbations

We can find the probability of a transition between vibrational modes in a molecule by using Fermi's golden rule, from this we realize certain selection criterion and features from which we can relate the probability of a transition to the measurable quantity of IR absorption. To use the golden rule we make certain assumptions; that i) there is a density of states $\rho(\varepsilon_f)$, due to a continuum of closely spaced final energy levels in the molecule, ii) we consider the long time limit from the uncertainty principle, $\tau >> \frac{2\pi\hbar}{\Delta\varepsilon}$, iii) that the perturbation interaction and the density of final states have a weak dependence on the final state, and iv) that first-order perturbation theory is appropriate here. Refer to equation 4.18, if we apply this to the treatment of periodic interactions, $\hat{v}(t) = e^{\pm i\omega t}$, then to derive the rule we write:

$$c_f(t) = -\frac{i}{\hbar}\int_0^\tau dt' \hat{v}_{fi} e^{i(\omega_{fi}\pm\omega)t'} \tag{4.44}$$

$$= -\hat{v}_{fi}\frac{e^{i(\omega_{fi}\pm\omega)\tau} - 1}{\hbar(\omega_{fi}\pm\omega)}, \tag{4.45}$$

in the limit of large τ, this gives:

$$\Gamma_{fi} = \frac{2\pi}{\hbar}|\langle f\,|\hat{v}|\,i\rangle|^2\,\rho(\varepsilon_i \mp \hbar\omega), \tag{4.46}$$

or,

$$\Gamma_{fi} = \frac{2\pi}{\hbar}\,|\langle f\,|\hat{v}|\,i\rangle|^2\,\delta(\varepsilon_f - \varepsilon_i \pm \hbar\omega). \tag{4.47}$$

$e^{i\omega t}$causes transitions where $\varepsilon_f = \varepsilon_i - \hbar\omega$, there is *emission* of light, and $e^{-i\omega t}$; where there is *absorption* of light. An electromagnetic field can cause transitions in both directions. The perturbative potential v, arises from the interaction of the electric field $\vec{E}$ with the charges on the molecule, thus:

$$v(r,t) = e\vec{E}cos\omega t.\left(\sum_i \vec{r}_i + \sum_N z_N \vec{R}_N\right) \tag{4.48}$$

$$= cos\omega t\vec{E}.\vec{\mu} \tag{4.49}$$

where i indexes over electrons, r is the electron coordinate, e is the charge of the electron, N indexes over nuclei, Z is the nuclear charge and $\vec{R}$ the nuclear coordinate. Thus $\vec{\mu}$ is the dipole moment due to the electron and nuclei. Then:

$$v_{fi} = \vec{E}\left\langle f\left|\vec{\mu}\right|i\right\rangle \tag{4.50}$$

$$= \vec{E}.\mu_{\mathbf{fi}}. \tag{4.51}$$

The transition matrix element terms are driven by the electric and nuclear dipole operator, we generalize the $\mu_{\mathbf{fi}}$ dipole moment operator for many particles.

Einstein coefficients

We see that IR absorption or emission will only occur if the transition frequency corresponds to the normal mode vibration that changes the electric moment of the molecule. *Selection rules* determine whether or not this is the case and if a transition is allowed. This is shown in Einstein's coefficient (B_{fi}/B_{if}) of absorption/emission given by [57]:

$$B_{fi} = B_{if} = \frac{8\pi^3}{3h^2}\left[\left|(\mu_x)_{fi}\right|^2 + \left|(\mu_y)_{fi}\right|^2 + \left|(\mu_z)_{fi}\right|^2\right]. \tag{4.52}$$

The frequency of light (v_{if}) stimulating the transition from $i \to f$ is given by:

$$v_{if} = \frac{W_i - W_f}{h}. \tag{4.53}$$

h is Planck's constant, W is the energy of the state involved in the transition $W_i \rightarrow W_f$, μ_x is the x component of the electric dipole moment and $(\mu_x)_{if}$ is an integral over all of space for the complex conjugate of the complete wavefunction of state i and the complete wavefunction of state f; $(\mu_x)_{if} = \int \psi_i^* \mu_x \psi_f d\tau$, and similarly for y and z [57].

IR absorbance

The electric dipole moment of the system is given by: $\mu_x = \sum\limits_a e_a x_a$ and again similarly for y and z, where e is the charge and x the Cartesian coordinate for the a^{th} atom, the sum being over all atoms. These charges e_a are effective charges in the sense the electron distribution changes as the atoms move. This is a new assumption treating the charge as a sum of monopoles centered on atoms. The electric moment can be expanded and expressed using normal modes Q_α over α modes:

$$\mu_x = \mu_x^0 + \sum_{\alpha=1}^{3N-6} \mu_x^{(\alpha)} Q_\alpha, \tag{4.54}$$

where μ_x^0 is the electric moment at equilibrium (the permanent dipole) and $\mu_x^{(\alpha)} = \left(\frac{\partial \mu_x}{\partial Q_\alpha}\right)_0$, the first differential, neglecting higher order terms. The integral $(\mu_x)_{n'n''}$ can be written:

$$\int \psi_{V'}^* \mu_x \psi_{V''} d\tau = \mu_x^0 \int \psi_{V'}^* \psi_{V''} d\tau_V + \sum_{\Omega=1}^{3N-6} \mu_x^{(\alpha)} \int \psi_{V'}^* Q_\alpha \psi_{V''} d\tau_V, \tag{4.55}$$

where we assume that the n' and n'' state can be described by vibrational wavefunctions V' and V'' respectively (separating electronic, rotational and translation wavefunctions). The first term disappears due to orthogonality. The second term survives and is assumed to be a product of harmonic oscillator functions corresponding to normal mode excitations [57]:

$$\int \psi_{V'}^* Q_\alpha \psi_{V''} d\tau_V = \int \psi_{v_1'}^* \psi_{v_1''} d\tau_V \int \psi_{v_2'}^* \psi_{v_2''} d\tau_V .... \int \psi_{v_\alpha'}^* Q_\alpha \psi_{v_\alpha''} d\tau_V. \tag{4.56}$$

For the several states of harmonic oscillator it must be the case that $v_\alpha'' = v_\alpha' + 1$ or $v_\alpha'' = v_\alpha' - 1$ for the total integral to be non-zero. The consequence of this is that only

one quantum number must change, for the rest $v_1'' = v_1'$ and $v_2'' = v_2'$, and this change must be of unit 1, and the quantum number v_α that changes must correspond to a non-zero $\mu_x^{(\alpha)}$, $\mu_y^{(\alpha)}$ or $\mu_z^{(\alpha)}$ term.

The Einstein coefficient mentioned above can be used to derive a measurable quantity: the integrated absorption coefficient $\int k(v)\,dv$. This can be found from;

$$k = \frac{1}{l} ln \frac{I_0}{I}, \tag{4.57}$$

and the differential form:

$$-dI = kI\,dl, \tag{4.58}$$

where I_0 is the initial and I is the final intensity of a beam of radiation crossing a path of length l. The Einstein coefficients as above can be used to express the decrease in intensity of the beam, and the following can be derived;

$$I_\alpha^{IR} = \int_{band} k(v)\,dv = \frac{N\pi}{3c} \sum_v \left[\left(\mu_x^{(\alpha v)}\right)^2 + \left(\mu_y^{(\alpha v)}\right)^2 + \left(\mu_z^{(\alpha v)}\right)^2 \right], \tag{4.59}$$

which describes the IR absorption intensity [57]. For a fuller explanation see Wilson, Decius and Cross [57]. N is the particle density and c is the speed of light. The sum is over atoms v, for α^{th} mode, $\mu_x^{(\alpha v)}$is the dipole derivative for Cartesian direction with respect to the normal mode coordinate for the αth mode (Q_α). The section in parenthesis is the "absolute" IR as calculated in Gaussian and is given in units of km/mol, see the investigation 'Huang-Rhys factors' for more.

4.3.2 Inelastic electron tunneling spectroscopy

Inelastic tunnelling spectroscopy (IETS) is observed in scanning tunnelling microscopy (STM). The system usually consists of two metal electrodes bridged by a small molecule; a metal-insulator-metal junction. A voltage V is applied to the junction, so that *elastic* transitions of the electrons occur and to a first approximation, the current-voltage characteristic is linear. However, at and above a certain threshold V_0 the conductance may increase, where the electrons take another route via *inelastic* transitions. At this point $\hbar\omega_0 \approx eV_0$, energy is conserved by excitation of a molecule, and the corresponding absorption can be seen from spectra produced by plotting $\frac{d^2I}{dV^2}$ against the bias voltage V, see Galperin, Ratner and Nitzan (2004),

[58]. Elastic and inelastic tunnelling due to phonon coupling can then be examined by analysis of the line-widths of peaks. Simonsen and Coleman give an electron tunnelling rate of:

$$W_{l\to r} = \frac{2\pi}{\hbar} |M_{rl}|^2 \, \delta(E_l - E_r - \hbar\omega_0). \tag{4.60}$$

Which is of course, Fermi's golden rule, equation 4.18.This rate describes an electron tunnelling from the right state at (E_r) to a lower energy state on the left (E_l), $\hbar\omega_0$ is the energy of the molecular excitation from the ground to excited state and M_{rl} is the tunneling matrix element. When the energy gap corresponding to $\hbar\omega_0$ is satisfied, IET will occur, and a peak will be seen at this corresponding frequency. IETS has been performed on a variety of organic molecules (aromatic ring compounds included) by Simonsen and Coleman (1973), however, the analysis of odorants is tricky because they are difficult to adsorb on the junction as they readily evaporate during vacuum deposition [1], [59].

4.4 Charge transfer

At the heart of the model of charge transfer is the BO approximation and the corresponding surfaces which describe a transition from a reacting state to a product state. All the important information is captured by two parabolic curves that represent the initial and final configuration; called the *configuration coordinate diagram*. It is grossly simple, a real representation of states would be multidimensional, but it is tremendously effective. The configuration coordinate diagram embodies ideas from two differently authored, but essentially the same, theories: Marcus theory and Huang-Rhys theory. The only real difference is the type of charge transfer they typically characterize. Marcus theory is usually used to describe chemical reaction rates, *i.e* transition state theory charge transfer across *two* mediums, and is often seen in biology textbooks such as Bendall's "Protein Electron transfer" for example [61]. Huang-Rhys theory is typically used to describe Franck-Condon-like optical excitations, where the charge transfer is within *one* medium, and is often found in physics textbooks, see Stoneham's "Theory of Defects in Solids" for example [56]. Both texts will be referred to here. There are many types of charge transfer across homo- (*intra*-protein, solids, polymers, colloids) or hetero- systems (*inter*-protein,

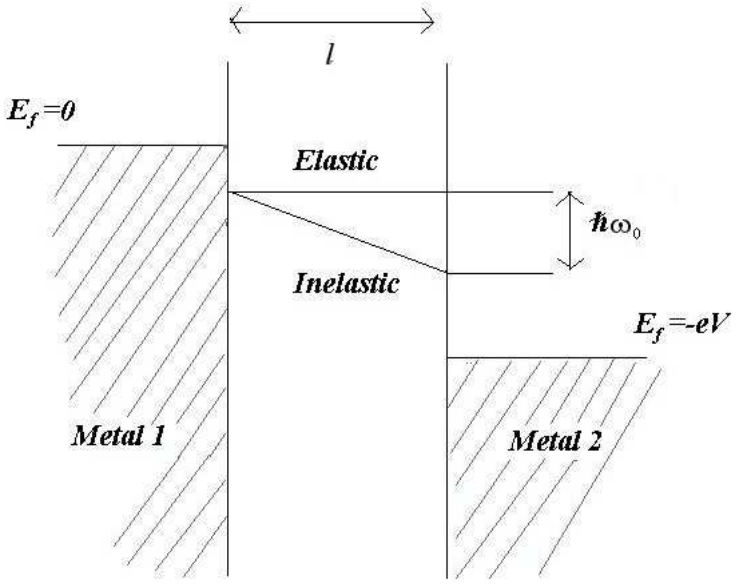


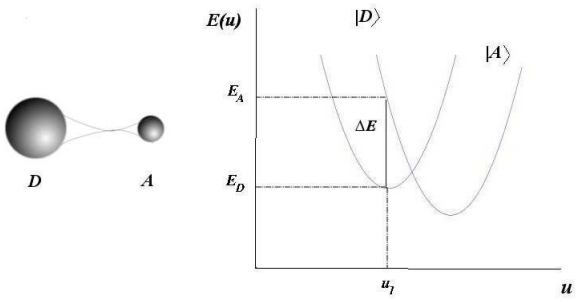
Figure 4.6: Tunneling metal-insulator-metal junction. There are two possible channels that conserve energy; i) *elastic* tunneling- a direct transition across the insulating barrier or ii) *inelastic* tunneling- a transition across the barrier when the electron loses energy $\hbar\omega_0$ to the excitation of an oscillator in the insulating barrier region. This only occurs at $eV \geq eV_0 = \hbar\omega_0$, [60].

polymer-liquid), that may be surrounded by vacuum (semi-conductor electrodes) or in a solvent (ion exchange in a solution). Further there are many types of mobile charge: protons (in hydrolytic enzymes or visual pigment photochemical interactions), holes (aqueous glasses) and excitons. For simplicity, in this thesis the mobile charge described will usually be the electron, and transitions are referred to as electron transfer (ET), however, other types of charge transfer will be considered in the Chapter "The olfaction model". It is a testament to the elegance of the configuration coordinate diagram that it is able to model in similar ways so many different situations. For the most part we use the Huang-Rhys methods to model charge transfer in olfaction although bridges will be made to Marcus theory and we will see that this is a nice opportunity to show that the important principles map, as shown in figures 4.8 and 4.9 and the sections below[2].

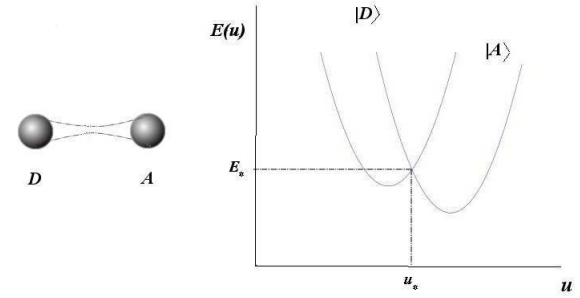
4.4.1 The configuration coordinate diagram

The two states $|D\rangle$ and $|A\rangle$ (see section 4.1.4) are shown as energy surfaces in figure 4.7. Following this diagram and using the ideas described in the following sections, we can visualize the transitions between these states by the following. If a ball (representing an electron) is rolling with kinetic energy on a surface $|D\rangle$ the nuclear coordinate describes its potential energy. Classically, the ball has turning points, where the motion is reversed, however, quantum mechanically there is a certain possibility that it can be found on either states $|D\rangle$ and $|A\rangle$. This is reflected by the probability density of the wavefunction $|\psi(r)|^2$. The electron, however, is formally transferred at the "crossing point" where the energy and coordinates are degenerate for both the donor and acceptor, $\Delta E = E_A(u) - E_D(u) = 0$. This is when the probability density for the electron on $|D\rangle$ or $|A\rangle$ is equal. When the ball is rolling on state $|A\rangle$, again the probability density shows an asymmetry, and the electron is most likely localized here.

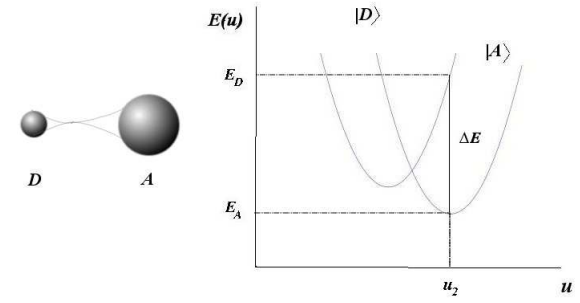
[2] For first hand Marcus theory see R.A. Marcus, *Ann. Rev. Phys. Chem*, 1964 [62] and for first hand Huang-Rhys theory see *Huang and Rhys Proc. Roy. Soc. London*, 1950 [63].



(a) Electron localized at D



(b) Electron crossing



(c) Electron localized at A

Figure 4.7: A schematic of electron transfer from donor (D) to acceptor (A), based on the document 'Introduction to Electron Transfer' at http://home.uva.nl/r.m.williams/Introduction%20to%20ET-30.htm.

Examination of figure 4.7, allows us to define a *reorganization energy*, $\Delta E = E_A(u_2) - E_D(u_2)$. The energies defined by these parabola intersections have important roles in the definition of charge transfer as we shall see below.

Reorganization energy: is the energy required to move from the reactant (donor) equilibrium to the product (acceptor) equilibrium, but remaining on the reactant (donor) surface.

Contrast and compare

The potential energy parabolas in these figures describe the system's state along a reaction coordinate (u); this could be a ferrous-ferric reaction in water [64], ET in *intra*-proteins [61], or in olfaction ET across a D-A past a small molecule in a hydrophobic environment [65]. The differences really lie in deciding the appropriate reaction coordinates, and the appropriate model of the rate, see the section 4.4.2. All the nuclear motion of the whole system is approximated in these parabolas. In a system of many atoms and bonds a single parabola can represent the variation of free energy of many oscillating bonds acting as simple harmonic oscillators and obeying Hooke's law. Configuration coordinates, depending on the system and ease of use, can embrace bond lengths, bond angles, molecular orientation, positions of the reactants, and positions of the solvents. More abstractly, it is the displacement along the vector joining the initial and final relaxed configurations. Thus positions of all atoms in the environment can be accounted for. It can be related to the "spectral density" of the system, which relates to the dielectric response of the system and accounts for the surrounding environment. See section 'Relaxation (reorganization) energy' in Chapter "The olfaction model" and Appendix B for more detail on this point.

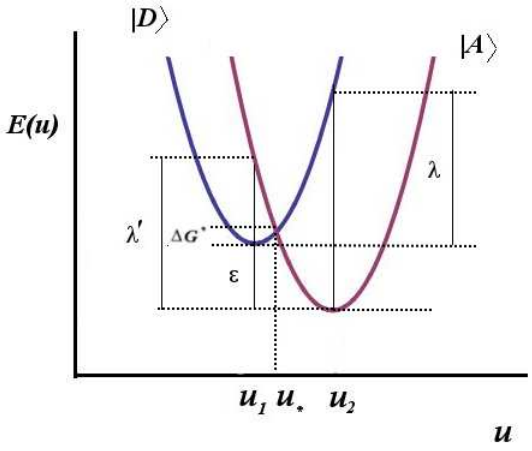


Figure 4.8: A configuration coordinate diagram, "Marcus parabolas" . This figure shows the total potential energy as a function of u for a reactant state and a product state. This type of diagram typically refers to a reaction rate. Showing the "reorganization" energies λ and λ'.

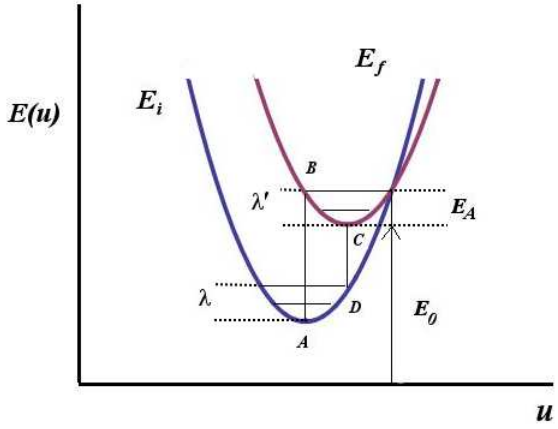


Figure 4.9: A configuration coordinate diagram, "Huang and Rhys". This figure shows the total potential energy as a function of u for two different electronic states. This type of diagram typically refers to energy transfer in optical absorption. Showing the "relaxation" energies λ and λ'. In this diagram the 'cross-over energy' is equal to the reorganization energies (which are also equal); note this may not always be the case.

The ideas of Marcus and Huang and Rhys are integral to the type of charge transfer described and we shall often refer back to these types of diagrams. The equilibrium geometry of the initial ($|D\rangle$) and final ($|A\rangle$) states are the minimum of each potential well respectively. Note these states are described by *equal* parabolas and a change in coordinate u is linear and shifts the parabola by $(u_D - u_A)$. Further, it must be noted that u denotes a *reaction coordinate* and not necessarily a *normal mode* coordinate, see the section 4.2 above, in most cases motions are generalized and so we cannot assume the displacement is the motion of one independent oscillator. Important features, derived from the figures, are discussed in the following sections. Common notation and descriptors for these features are given, for Huang-Rhys theory and Marcus theory, and referenced accordingly.

4.4.1.1 The cross-over energy

$E_A = \Delta G^* = \Delta G^\ddagger$ known as the *cross-over* [66] or *classical activation barrier* [61] or *activation energy* [61], respectively. This is the energy difference between the point

of the upper parabola minimum to the point where the upper parabola crosses onto the lower parabola. It is given by:

$$E_A = \frac{(\varepsilon + \lambda)^2}{4\lambda}, \tag{4.61}$$

when $\lambda = \lambda'$. It is most usually known as the *activation energy* for a traditional adiabatic chemical reaction where it is the energy between the point of cross over and the minimum of the reactant, in a non-adiabatic case. On examination of equation 4.61 we see that when $-\varepsilon = \lambda$, there is no activation energy. When $-\varepsilon < \lambda$, there is "normal" transition and at $-\varepsilon > \lambda$ the activation energy can be large but the reaction rate is actually slower, as shown by the Gaussian dependence on free energy shown in:

$$FC = \frac{1}{\sqrt{4\pi\lambda k_B T}} exp \frac{-(\varepsilon + \lambda)^2}{4\lambda k_B T}, \tag{4.62}$$

where FC denotes the Franck-Condon weighted density of states [61]

4.4.1.2 Relaxation (reorganization) energy

λ is known as the *relaxation energy* or *reorganization energy* [66], [61]. This is defined as the energy that must be added to the reactants (initial state) to move from the equilibrium reactant (initial) geometry to the equilibrium product (final) geometry, but by remaining on the reactant (initial) surface, so by not transferring an electron [61]. Note that this depends on the direction of the reaction, and that there can be two reorganization energies λ and λ', as seen in both figures depending on whether we are discussing endothermic/exothermic reactions or absorption/emission energies respectively, depending on the model used. This is simplified however, because we assume $\lambda = \lambda'$; the parabolas have the same spring constant. Thus the reorganization is always taken as described before: λ, in the direction $D \longrightarrow A$. Note also that twice the reorganization energy is the energy loss in the *Stokes shift* between absorption (A-C) and emission (C-D) as shown in figure 4.9. This is given by:

$$\lambda = \frac{1}{2} M\Omega^2 (u_D - u_A)^2 \tag{4.63}$$

where M is the collective mass, and Ω the collective frequency. This is found explicitly in the investigation "Huang-Rhys factors", and see section 4.4.3, and see section 4.4.3.

4.4.1.3 Driving force

$\varepsilon = E_0 = \Delta G$ known as the *driving force* or the *zero-phonon line* or the *free energy* of the reaction, respectively [61], [66], [61]. It is the energy difference between the two minimums. $E_0 + \lambda$ may refer to a Franck-Condon optical excitation.

4.4.2 Adiabacity versus non-adiabacity

In order to describe charge transfer processes we use an *adiabatic* approximation: that is the electronic and nuclear motions are separated as in the BO approximation above. The nuclei positions u are the *adiabatic parameters*; with gradual changes in u, the electronic Hamiltonian is solved and each corresponding eigenstate is the *adiabatic* state (the BO state). So the PES depicted in figure 4.3 is an *adiabatic surface*, as are the parabolas in the above figures. However, for interesting things to happen states must change, and this corresponds to *non-adiabaticity*. A pioneer of quantum transitions, DeVault, explains as follows: "Briefly, since nuclear motion is generally much slower than electronic motion, one can approximate the electronic part of the wave-function of a molecular system by solving for it with nuclei fixed in position. The electronic energy eigenvalues obtained this way, when plotted as a function of the nuclear positions, form adiabatic surfaces which become potential-energy surfaces for nuclear motion. However, when the nuclei are allowed to move, the wave-functions arrived at by this approximation are no longer exactly eigenfunctions and they can change spontaneously from one to another. The matrix elements causing the changes are made from the terms neglected in the approximation and are called the 'non-adiabaticity operator'. This operator involves derivatives of both the electronic and the nuclear wave-functions with respect to nuclear coordinates"[67]. Some definitions may be useful:

Adiabacity: External parameters of a system change but this does not induce a transition with evolution of the system [56]. The system adapts with gradual changes so that the probability density of wavefunction that describes the final state is modified in the process. A system in an eigenstate of the initial

Hamiltonian will end up in a corresponding eigenstate of the final Hamiltonian.

Non-adiabacity (diabaticity): External parameters of a system change quickly so that the system cannot adopt its adiabatic configuration, the probability density of the wavefunction that describes the final state remains the same. There is no corresponding eigenstate of the final state Hamiltonian with the initial state Hamiltonian. A linear combination of states sum to produce the final probability density.

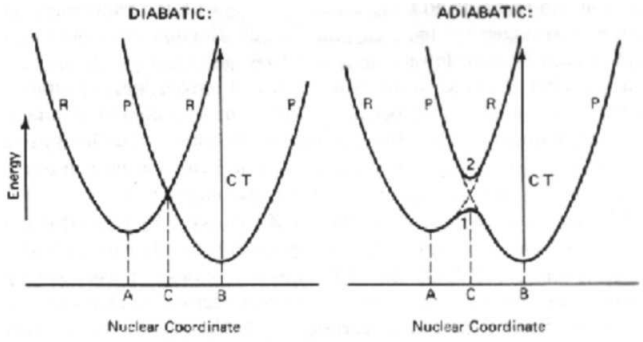


Figure 4.10: Non-adiabatic and adiabatic Marcus energy surfaces [67]. Electron transfer is from R to P. For the non-adiabatic energy surface (left) the coupling between surfaces is *weak* and the reaction rate constant is $k < 1$. For the adiabatic energy surface (right) coupling between surfaces is *strong* and the reaction rate constant is $k = 1$, k is as defined in the text.

Equation 4.64 is given by Marcus theory for an Arrhenius dependent electron transfer rate where $\Delta G^{\ddagger}$, $\Delta S^{\ddagger}$, $\Delta H^{\ddagger}$ are the free energy, the entropy and the enthalpy respectively [67]. Examination of this equation shows us how adiabacity/non-adiabacity effects the rate [67]:

$$k_{ET} = k\,\frac{k_B T}{h} exp(-\Delta G^{\ddagger}/k_B T) = k\,\frac{k_B T}{h} exp(\Delta S^{\ddagger}/k_B T)\, exp(-\Delta H^{\ddagger}/k_B T). \qquad (4.64)$$

The value of the reaction rate constant (k) indicates whether the process is adiabatic ($k = 1$) or non-adiabatic ($k < 1$) and so sets a limiting value. In figure 4.10 for the non-adiabatic transition electron transfer, transfer is less probable, and there

is a quantum 'jump' from one curve to the other where the curves cross. The reactant and product are *weakly* coupled. In contrast the adiabatic process is one in which no quantum jump occurs. The transition is more probable, so that electron transfer back and forth may occur many times. There is a quantum mechanical resonance between the two states, as the electron lingers at the barrier so that electronic orbitals interact and the curves representing the two states smooth to form a continuum, with a quasi-state at the top of the activation barrier. In this case the reactant and product are *strongly* coupled [62]. It is these time scales and strength of coupling that determine the adiabacity and thus the rate of reaction.

For non-adiabatic transfer we can use Fermi's golden rule, see section 4.1.4, to calculate an electron tunneling rate which can be given by:

$$k_{ET} = \frac{2\pi}{\hbar} |H_{AD}|^2 \, FC \tag{4.65}$$

where H_{AD} is the transition matrix element that contains the small non-adiabatic operator, and FC denotes the *Franck-Condon* factor; a large FC corresponds to a favourable overlap between initial and final nuclear wavefunctions and this is expressed as the *lineshape function* in Chapter "The olfaction model". This rate is a good estimation when the donor and acceptor are well separated and distinct, [61]. If they were too close they would interact strongly and the transfer would be adiabatic. According to Marcus theory this type of rate equation is perfectly appropriate and accurate for the types of single electron transitions between redox centres found in bioenergetic membranes, [61].

Franck-Condon principle: that electronic transitions occur with no change in nuclear configuration [56].

4.4.3 Huang-Rhys theory

The Huang-Rhys model precedes and is a special case of Marcus theory, [63]. It describes a full quantitative quantum-mechanical treatment of electronic transitions in solids coupled to the vibrational excitations in the system and surroundings. Typically Huang-Rhys theory is applied to Franck-Condon type optical transitions as depicted in figure 4.9 above. Huang and Rhys derived an equation to describe the optical absorption bands of F centres in ionic crystals (see figure 4.12). The absorption of a phonon can result in electronic transitions coupled with the crystal

lattice vibration. The Huang-Rhys formula regards the single frequency case. With reference to this figure 4.9; a typical transition involves absorption of light to an electronically excited state (A-B), a non-radiative relaxation to C, emission of light (C-D) and a final non-radiative relaxation to A. The *Huang-Rhys factor* (S) measures the relaxation energy (B-C) according to the phonons released when there is an electronic transition. This type of *non-radiative* transition is well described for semi-conductor devices and accounts for the loss of efficiency [66]. These transitions can be appropriately described using the configuration coordinate diagram as above, and simple harmonic approximations. Consider a perfect host lattice in the electronic ground state, the motion of the nuclei oscillating about equilibrium, a lower parabola corresponds to [66];

$$E_i = \sum_\alpha \frac{1}{2}\tilde{m}_\alpha \dot{Q}_\alpha^2 + \sum_\alpha \frac{1}{2} k_\alpha Q_\alpha^2 + E, \tag{4.66}$$

where Q is the normal mode as usual. Higher order terms of E are ignored. Introducing a defect into the lattice imparts a restorative force such that a second parabola can be described:

$$\begin{aligned} E_f &= \sum_\alpha \frac{1}{2}\tilde{m}_\alpha \dot{Q}_\alpha^2 + \sum_\alpha \frac{1}{2} k_\alpha Q_\alpha^2 - \sum_\alpha f_\alpha Q_\alpha \\ &= \sum_\alpha \frac{1}{2}\tilde{m}_\alpha (\dot{Q}_\alpha + \dot{\delta})^2 + \sum_\alpha \frac{1}{2} k_\alpha (Q_\alpha + \delta) - \sum_\alpha \frac{1}{2} f_\alpha^2 / k_\alpha. \end{aligned} \tag{4.67}$$

E_i and E_f describe the two curves in figure 4.9. This linear term due to the disruptive force displaces the normal modes from equilibrium, though the eigenvalues and eigenvectors remain the same. $\delta = \frac{f_\alpha^2}{k}$ corresponds in the linear shift of mean displacement between them, for a given mode α, and the energy $\lambda_\alpha = \frac{^2}{2k}$ is the relaxation energy (or Marcus reorganization energy) as defined above. This shift in parabolas can give us a measure of the electron-phonon coupling, which we define by the ratio of $S_\alpha = \frac{\lambda\alpha}{\hbar\omega_\alpha} = \frac{\frac{1}{2}M\Omega^2(u_{D\alpha}-u_{A\alpha})^2}{\hbar\omega_\alpha}$, where $u_{D\alpha} - u_{A\alpha} = Q_{DA,\alpha}$ is a change in mode α, thus we obtain a dimensionless Huang-Rhys factor (S).

Strong and weak electron-phonon coupling

$Q_{DA,\alpha}$ is the linear displacement along mode α, but again, may not necessarily be the normal mode. Within this regime a strong coupling limit is defined by S>>1, and a weak coupling limit is defined by S<<1, see figure 4.11.

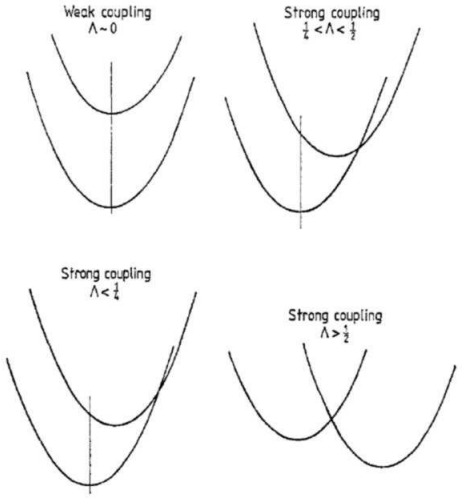


Figure 4.11: Strong and weak coupling in electronic transitions [66].

We use the model for *weak* electron-phonon coupling in non-radiative transitions, to estimate one of the important parameters (S) for more detail see the investigation "Huang-Rhys factors".

4.4.4 Accepting modes and promoting modes

The important modes we have described so far are *accepting* modes. They play a part in the lineshape function which will be seen more explicitly in Chapter "The olfaction model". They appear in the configuration coordinate, and they have an S value. More subtly, but still importantly, are *promoting* modes which also effect the rate equation 4.65. The transition matrix element in equation 4.65 is geometry dependent. Following certain reaction coordinates can result in more favourable mixing of these elements. Thus a transition may be encouraged by a mode that facilitates a favourable path way (promoting) or by providing an appropriate energy/driving force (accepting). It is possible the transition involves both, neither,

or one or the other [66]. Promoting modes have no S value. They are affected by symmetry and distribution in space; a strong promoting mode depends on a strong overlap between molecular orbitals. Thus looking for a strong admixture of initial and final states between transitions as opposed to the Huang-Rhys couplings may be more informative, although there are difficulties here (see Chapter "The olfaction model").

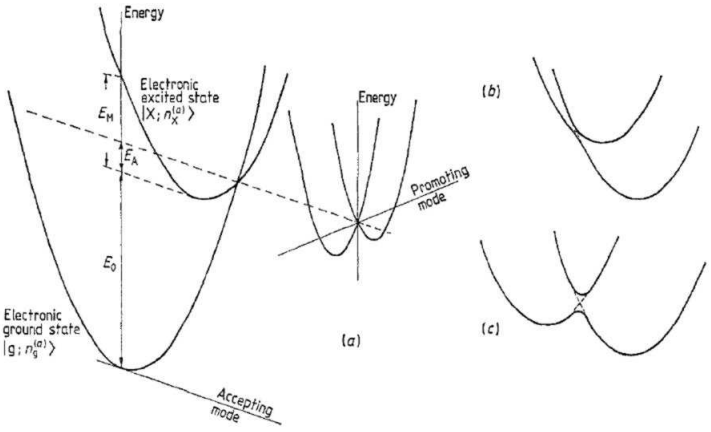


Figure 4.12: Promoting and accepting modes, a) is both a promoting and an accepting mode, b) the accepting mode is also promoting, and c) adiabatic. These configuration coordinates characterize the absorption of photons by F centres in ionic crystals, the electronic excitations are coupled to crystal vibrations, [66].

Chapter 5

Molecular Dynamics

We can learn many things about a solid, liquid or gas, an inorganic, organic or metal material, just from solving Newton's law of motion $F = ma$ for the motions of the atoms. If we have an expression for the forces, then the force gives an acceleration, integrating the acceleration gives a velocity and so from the velocity we can find the new positions of the atoms. These forces are found by deciding the appropriate potentials for your system and defining them as a *force field*. This can be reiterated over time producing a trajectory. Microscopic properties can be converted into macroscopic observables, *e.g.,* atomic positions and velocities are converted into pressure (P). Many thermodynamic properties can be found from molecular dynamics (MD) simulations such as; the density, the chemical potential, the specific heat capacity, diffusion constants and structure factors. These can be derived by differentiation of equations of state and from the laws of thermodynamics. The trajectory also contains dynamics information, which as we shall see, is of most interest in this thesis. The application here is to observe fluctuations in molecular geometry as a function of time corresponding to the thermal motion of a molecule in a low density gas. For brevity then, I only describe in any detail the MD techniques that can be used to explore the flexibility of a molecule over time at a certain temperature. Further I describe methods for DL_POLY (version 2.16) users.

5.1 Phase space, trajectories and ensembles

The positions (r) and momenta (p) of many coordinates make up a configuration of the system and this can be represented as a vector Γ in *phase space*. Space being 3 dimensional there are 6N of these points in phase space. This vector defines a single point in phase space, so each configuration has a corresponding point. The dynamical system explores a trajectory in this phase space and an ensemble consists of a collection of points within it, often defined by certain fixed macroscopic parameters. The enumeration of the conditions defined by the fixed parameters make the *ensemble*. The types of ensembles commonly used are described below [68]:

The canonical or constant -NVT ensemble

For a given thermodynamic macrostate with a fixed number of atoms N, a fixed volume V and a fixed temperature T, this ensemble describes all possible microstates. Thermostats are added to the simulation to moderate the temperature by adding and removing energy.

The isothermal-isobaric or constant -NPT ensemble

For a given thermodynamics macrostate with a fixed number of atoms N, a fixed internal pressure P and a fixed temperature T, this ensemble describes all possible microstates. Barostats are added to the simulation to moderate the pressure.

The grand canonical or constant-μVT ensemble

For a given thermodynamics macrostate with a fixed chemical potential μ, a fixed volume V, and a fixed volume T, this ensemble describes all the possible microstates. The chemical potential is fixed by introducing an external reservoir that allows the number of particles and energy to fluctuate accordingly.

The microcanonical or constant-NVE ensemble

For a given thermodynamics macrostate with a fixed number of atoms N, a fixed volume V, and a fixed internal energy E, this ensemble describes all possible microstates. As there is no exchange of heat into or out of the system, so this is an *adiabatic* process.

5.1.1 The microcanonical ensemble

This latter ensemble is the one I used for those calculations in Chapter 'Chirality and conformations in odorants', and so it will be described as the example here. Whilst the other ensembles are chosen, and were developed, to compare better with experimental results, *i.e.,* temperature is easier to fix than energy, the constant (NVE) is the most suitable here. This is because the internal dynamics of odorants can be well represented by this ensemble, an isolated mechanical system, without the added complications of thermostats and barostats[1]. The points in phase space make a distribution described by a probability $\rho(\Gamma)$ [68]. The probability density for the microcanonical ensemble is:

$$\rho(\Gamma) \propto \delta(E(\Gamma) - E) \tag{5.1}$$

where $E(\Gamma)$ is the Hamiltonian. The specification of the microstate can be represented in phase space Γ as the set of particles defined by their positions and momenta. The delta function selects microstates for which, with a constant N and a constant V, the right energy E is maintained. The quasi-classical partition function can be written [68]:

$$Z(NVE) = \frac{1}{N!}\frac{1}{h^{3N}}\int drdp\delta(E(r,p) - E) \tag{5.2}$$

N! accounts for indistinguishable particles, and the h^{3N} constant allows convergence to the ideal gas approximation, where potential energy is zero so that the energy surface determined by phase space is a 3N-dimensional sphere, the integral is over all 6N phase space coordinates $\int drdp$, and gives the surface area of this energy surface. This is partitioned according to the states. A key point is that the probability of any microstate Z(NVE) is equal for an isolated system in equilibrium. Note equation 5.2 can be used in a quantum development, with the inclusion of Planck's constant and the use of quantum numbers for p and r. For this thesis, however, we take the classical limits and the NVE corresponds to classical Newtonian dynamics of an isolated system, as used in the package DL_POLY. The thermody-

[1]We note, however, that when X-ray or neutron-scattering is available for the olfactory receptor (OR), and hopefully the OR with its ligand, then that may be a good time to modify the model to go with experimental constraints.

namic potential (the negative of entropy) is given by:

$$-S(NVE) = -k_B lnZ(NVE). \tag{5.3}$$

We only assume forces in-between the particles, here are no external forces, and so the total energy (kinetic + potential) remains constant. The trajectory is confined by this, and over time the dynamical system maps out a trajectory on phase space with constant energy. After a long time we assume an average over the trajectory is equal to an average over the ensemble, at equilibrium:

$$A(NVE) = \langle A \rangle_{ens} = \frac{1}{\tau_{obs}} \sum_{\tau=1}^{\tau_{obs}} A(\Gamma(\tau)) \tag{5.4}$$

where A is the macroscopic observable and τ_{obs} is a large finite number of time steps. This is an implementation of the *ergodic hypothesis*, which says that if the states densely cover the constant energy surface, in the limit of long time simulations, then the time averaged values for an observable will match the ensemble averaged values for the observable. This requires that the system does not become locked into a particular region of phase space and that the ensemble averages are not dependent on the initial conditions.

5.2 Equations of motion

Forces acting on the atoms are calculated at each time step and then used to solve Newton's laws of motion and move the nuclear positions. Thus an MD simulation, like the type described here (see below for the integration method) is a deterministic and time-reversible process. Consider a collection of atoms i, with Cartesian coordinates, r_i then:

$$F_i = m_i a_i \tag{5.5}$$

where F is the force, m the mass, a the acceleration, v the velocity, p the momentum (related to velocity $p = mv$), and V is the potential, which is related to the force by:

$$F_i = m_i \ddot{r}_i = -\nabla_{r_i} V. \tag{5.6}$$

Equation 6, shows us an expression that relates the change in potential energy with the particle position with the change in particle position as a function of time. Thus, from initial positions and velocities, a potential is applied, a force found and so an acceleration. From this the integration of acceleration and then velocity:

$$a_i = \frac{dv_i}{dt} \tag{5.7}$$

and:

$$v_i = \frac{dr_i}{dt} \tag{5.8}$$

yields new positions. Thus we have a propagation of momentum and position respectively [69]. First order Taylor expansions give (Euler's approximation):

$$p(t + \delta t) = p(t) + ma(t)\delta t \tag{5.9}$$

and

$$r(t + \delta t) = r(t) + \frac{p(t)}{m}\delta t. \tag{5.10}$$

The initial positions of atoms can be found from a geometry optimization which removes large artificial forces. In a typical MD simulation there are initial conditions, an equilibrating period and then a production period. Initial velocities can be chosen with a distribution corresponding to the required temperature and satisfying zero total linear momentum [68]:

$$P = \sum_{i=1}^{N} m_i v_i = 0. \tag{5.11}$$

During the equilibration period velocities can be scaled according to the kinetic energy:

$$\left\langle \sum_{i=1}^{N} \frac{p_i^2}{2m_i} \right\rangle = \frac{3N-3}{2} k_B \left\langle T \right\rangle \tag{5.12}$$

where i goes over the number of atoms N, and the brackets indicate a time average and there is 3N-3 to account for the number of degrees of freedom. After the equilibration period, the production period generates simulations yielding information to

calculate macroscopic properties at the systems equilibrium point.

A problem with Euler's approximation (equations 5.9 and 5.10) for finite time steps δt, is that we assume constant acceleration at each increment and so we get tangents to the trajectory curve at each point. These may deviate from a physically meaningful path in phase space, though gives good phase space trajectories in the limit $\Delta t \to 0$. So for smaller increments of t the closer the integration to a true trajectory curve. However, this comes at a heavy computational cost. We now consider a better algorithm that led us to use a longer time step.

5.3 Integration algorithms

Equation 5.5 is iteratively solved using the finite difference method of solving ordinary differential equations at a chosen δt. We make a Taylor expansion, and switching from scalar notation to a more general vector notation giving [68]:

$$\mathbf{r}(t+\delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2}\delta t^2 \mathbf{a}(t) + \frac{1}{6}\delta t^3 \mathbf{b}(t) + O(\delta t^4). \tag{5.13}$$

One of the most popular methods of propagation is that developed by Verlet in 1967. Consider the equation 5.13 in forward and backwards time, and add the two variants. The odd power terms are of opposite sign and so cancel, and we obtain, to second order:

$$\mathbf{r}(t+\delta t) \longleftarrow 2\mathbf{r}(t) - \mathbf{r}(t-\delta t) + \mathbf{a}(t)(\delta t)^2 + O(\delta t^4), \tag{5.14}$$

note the velocities have been eliminated. We only require a position, a previous position and an acceleration to drive to the next step. For the first position (where there is no history) equations 5.9 and 5.10 can be used. This is the *Verlet* scheme. Within this scheme is a velocity Verlet (VV) and a Verlet 'leap frog' (LF) algorithm, but I will only describe the latter as it is the default integrator for DL_POLY and the method used here. The starting points are different to the above, notably because the velocities are used in propagation, and they are in fact where you begin. We require the initial positions and integrate the forces to get the velocity half a time step ahead:

$$\mathbf{v}(t+\frac{1}{2}\delta t) \longleftarrow \mathbf{v}(t-\frac{1}{2}\delta t) + \frac{f(t)}{m}(\delta t) + O(\delta t^3) \tag{5.15}$$

where m is the mass of the site. The positions can then be advanced using the new velocities:

$$\mathbf{r}(t+\delta t) \longleftarrow \mathbf{r}(t) + \delta t \mathbf{v}(t+\frac{1}{2}\delta t) + O(\delta t^3) \tag{5.16}$$

and so on. More steps are required in propagation than in Verlet's original method, but force field calculations only take place at an integer time step, not a half integer increment. These are just used to get to the next position (leaped over) where the computer expensive calculations are then performed. So the positions and velocities are projected forward in time in a coupled manner, and this is suitable for the type of calculations I wish to present because the velocities at half-integer time steps can be scaled according to temperature, such that it can be controlled. At each integer time step the velocity can be found from the average of the velocities at half a time step either side of the time t:

$$v(t) = \frac{1}{2}(\mathbf{v}((t+\frac{1}{2}\delta t) + \mathbf{v}(t-\frac{1}{2}\delta t)) + O(\delta t^3) \tag{5.17}$$

which is useful in cases where it is desirable to know quantities depending on the positions and velocities at a given instant [68].

5.4 Force fields

All particle-particle interactions require a potential in order to define the forces in the system and this is what is referred to as the *force field*. Equation 5.18 shows the potentials described that constitute the total configurational, *steric*, energy of a system, this is the definition from the DL_POLY 2.16 user manual. The first 4 potentials are *intra*-molecular interactions; the chemical bond potential, the valence angle potential, the dihedral angle potential and the inversion angle. The next 4 terms are known as the *inter*-molecular interactions; the pair-body, the three body, the Tersoff and the four body interactions. The last two terms are the potential for a metal U_{metal} which relies on a density component and takes account of many body electron-electron interactions, and the external potential U_{extern} that accounts for an external field. In this equation i indexes over the relevant interaction and the vectors $\vec{r}$ denote site positions. N denotes total numbers of the interactions, and

$\vec{R}^N$indicates a many body dependence.

$$
\begin{aligned}
U(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N) = & \sum_{i_{bond}=1}^{N_{bond}} U_{bond}(i_{bond}, \vec{r}_a, \vec{r}_b) \\
& + \sum_{i_{angle}=1}^{N_{angle}} U_{angle}(i_{angle}, \vec{r}_a, \vec{r}_b, \vec{r}_c) \\
& + \sum_{i_{dihed}=1}^{N_{dihed}} U_{dihed}(i_{dihed}, \vec{r}_a, \vec{r}_b, \vec{r}_c, \vec{r}_d) \\
& + \sum_{i_{inv}=1}^{N_{inv}} U_{inv}(i_{inv}, \vec{r}_a, \vec{r}_b, \vec{r}_c, \vec{r}_d) \\
& + \sum_{i=1}^{N-1} \sum_{j>i}^{N} U_{pair}(i, j, |\vec{r}_i - \vec{r}_j|) \\
& + \sum_{i=1}^{N-2N-1} \sum_{j>i}^{N} \sum_{k>j}^{N} U_{3body}(i, j, k, \vec{r}_i, \vec{r}_j, \vec{r}_k) \\
& + \sum_{i=1}^{N-1} \sum_{j>i}^{N} U_{Tersoff}(i, j, k, \vec{r}_i, \vec{r}_j, \vec{R}^N) \\
& + \sum_{i=1}^{N-3N-2N-1} \sum_{j>i}^{N} \sum_{k>j}^{N} \sum_{n>k}^{N} U_{4body}(i, j, k, n, \vec{r}_i, \vec{r}_j, \vec{r}_k, \vec{r}_n) \\
& + \sum_{i=1}^{N} U_{metal}(i, \vec{r}_i, \vec{R}^N) \\
& + \sum_{i=1}^{N} U_{extern}(i, \vec{r}_i, \vec{v}_i)
\end{aligned}
\tag{5.18}
$$

Particle-particle interactions can be extensive if we include all interactions such as the above and in this way we might more accurately define the environment the particles experience. A simulation then, can really only be as good as the potentials that describe the system. In consideration of this quantum mechanical methods are often used in combination with an MD force field in order to refine the potentials, by gaining a greater accuracy to a potentials parameters. For my purposes, as I would like to compare the dynamics of different odorants, modelling isolated gas

phase molecules is enough, thus the force field is greatly simplified. In simulations of larger systems, *e.g.*, liquids and crystals, the interactions of molecules are appropriately accounted for by careful consideration of the boundary conditions. In the MD simulations I describe, *intra*-molecular forces are calculated from energy gradients with respect to the nuclear coordinates at each time step, this is the simple type of potential landscape that is represented pictorially in Chapter 'Chirality and conformations of odorants' figure 9.11.

5.4.1 The Dreiding model

I use a DREIDING force field, based on the success of its accuracy when applied to similar small molecules as odorants (modelled results are compared to crystal structures of organic compounds, rotational barriers and relative conformational energies and barriers) [70]. This force field is simple and proves accurate for the structure and conformational analysis of organic, biological and main group inorganic molecules (see Mayo *et al* [70]). Only nuclear centres are used in generating a force field core model. Figures 5.1 and 5.2 list all the parameters needed to construct a DREIDING force field, as defined by [70]. So the following sections list the potentials required for simulations done in Chapter 'Chirality and conformations of odorants' and discuss the parameters that need to be defined.

Chemical bond potential

The bond stretch interaction is described harmonically:

$$U(r_{ij}) = \frac{1}{2}k(r_{ij} - r_0)^2 \tag{5.19}$$

where $r_{ij} = |\vec{r}_j - \vec{r}_i|$, the *inter*-atomic distances explicitly define the potential. If anharmonic terms are important then a Morse function should be used, but this potential is adequate for small displacements about equilibrium [69]. The spring constant k of bond stretching is set to $k_{ij} = 700$ (kcal/mol)/Å, for single bonds, this is simple in its approximation of k to all bonds, but it proves effective [70]. Other methods may base spring constants on empirical parameters. The equilibrium bond distance is found by $r_0 = R_I^0 + R_J^0$-δ, as given by Mayo *et al* [70], with the addition of bond radii of $-\delta = 0.01$Å.

TABLE I: Geometric Valence Parameters for DREIDING

atom	bond radius R_1^0, Å	bond angle, deg	atom	bond radius R_1^0, Å	bond angle, deg
H_	0.330	180.0	Si3	0.937	109.471
H__HB	0.330	180.0	P_3	0.890	93.3
H_b	0.510	90.0	S_3	1.040	92.1
B_3	0.880	109.471	Cl	0.997	180.0
B_2	0.790	120.0	Ga3	1.210	109.471
C_3	0.770	109.471	Ge3	1.210	109.471
C_R	0.700	120.0	As3	1.210	92.1
C_2	0.670	120.0	Se3	1.210	90.6
C_1	0.602	180.0	Br	1.167	180.0
N_3	0.702	106.7	In3	1.390	109.471
N_R	0.650	120.0	Sn3	1.373	109.471
N_2	0.615	120.0	Sb3	1.432	91.6
N_1	0.556	180.0	Te3	1.280	90.3
O_3	0.660	104.51	I_	1.360	180.0
O_R	0.660	120.0	Na	1.860	90.0
O_2	0.560	120.0	Ca	1.940	90.0
O_1	0.528	180.0	Fe	1.285	90.0
F_	0.611	180.0	Zn	1.330	109.471
Al3	1.047	109.471			

TABLE XI: Rotational Barriers (kcal/mol) about Single Bonds

molecule[a]	experiment: periodicity[b]	experiment: barrier[c]	calculated barrier
CH_3–CH_3	V3	2.882 (0.010)[d]	2.896
CH_3–CH_2CH_3	V3	3.4[e]	3.376
CH_3–$CH_2CH_2CH_3$	V3	3.4[e]	3.410
CH_3CH_2–CH_2CH_3	V3	3.8[e]	3.822
CH_3–$CH(CH_3)_2$	V3	3.9[e]	3.995
CH_3–$C(CH_3)_3$	V3	4.7[e]	5.071
CH_3–CH_2F	V3	3.287 (0.03)[d]	3.172
CH_3–CH_2Cl	V3	3.68[e]	3.487
CH_3–CH_2Br	V3	3.68[e]	3.345
CH_3–CH_2I	V3	3.623 (0.15)[d]	3.336
CH_3–CF_3	V3	3.16 (0.11)[d]	3.768
CH_3–CCl_3	V3	5.10 (0.3)[d]	4.851
F_3C–CF_3	V3	3.92[e]	5.562
CH_3–SiH_3	V3	1.7[e]	2.296
CH_3–GeH_3	V3	1.24[e]	2.037
CH_3–CH_2SiH_3	V3	2.625 (0.01)[d]	3.805
CH_3CH_2–SiH_3	V3	1.979 (0.007)[d]	2.517
CH_3–$Si(CH_3)_3$	V3	1.4[e]	3.191
CH_3–NH_2	V3	1.98[e]	2.085
CH_3–$NHCH_3$	V3	3.62[e]	2.916
CH_3–$N(CH_3)_2$	V3	4.4[d]	3.534
CH_3–PH_2	V3	1.96[e]	1.957
CH_3–OH	V3	0.373 (0.003)[d]	2.117
CH_3–SH	V3	0.445 (0.000)[d]	2.376
CH_3–SeH	V3	0.957 (0.05)[d]	2.183
CF_3–OF	V3	3.900[d]	3.608
CH_3–OCH_3	V3	2.630 (0.007)[d]	3.034
CH_3–SCH_3	V3	2.099 (0.003)[d]	2.902
CH_3–$SeCH_3$	V3	1.498 (0.001)[d]	2.601
CH_3—CH=CH_2	V3	1.995[e]	0.753
CH_3—CH=O	V3	1.143[d]	0.948
CH_3—C(OH)=O	V3	0.481[d]	1.026
CH_3—$C(OCH_3)$=O	V3	0.284[d]	1.030
NH_2—CH=O	V2	18 (3)[e]	24.506
$N(CH_3)_2$—CH=O	V2	19.6 (1.5)[e]	21.037

[a] The dihedral pair is indicated by the single line. [b] Symmetry assumed in experimental analysis. [c] The value in parentheses indicates estimated experimental uncertainty. [d] Reference 25. [e] Reference 26.

Figure 5.1: DREIDING parameters as collected by Mayo *et al* [70].

TABLE IV: DREIDING Torsion Parameters for Equivalent Central Atoms

atom	n	V_{JJ}, kcal/mol	φ, deg
H_		0	
B_3	3	2.0	180
C_3	3	2.0	180
C_R	2	25.0	180
C_2	2	45.0	180
C_1		0	
N_3	3	2.0	180
N_R	2	25.0	180
N_2	2	45.0	180
N_1		0	
O_3	2	2.0	90
O_R	2	25.0	180
O_2		45.0	180
O_1		0	
F_		0	
Al3	3	2.0	180
Si	3	2.0	180
P_3	3	2.0	180
S_3	2	2.0	90
Cl		0	
Ga3	3	2.0	180
Ge	3	2.0	180
As3	3	2.0	180
Se3	2	2.0	90
Br		0	
In3	3	2.0	180
Sn3	3	2.0	180
Sb3	3	2.0	180
Te3	2	2.0	90
I_		0	
Na		0	
Ca		0	
Fe		0	
Zn		0	

TABLE II: The van der Waals Parameters for DREIDING

atom	R_0 Å	D_0, kcal/mol	ζ	source
H	3.195	0.0152	12.382	*n*-hexane crystal[11a]
H__b	3.195	0.0152	12.382	interpolation
H_HB	3.195	0.0001	12.0	H_2O dimer
B	4.02	0.095	14.23	interpolation
C	3.8983	0.0951	14.034	Williams[11a]
N	3.6621	0.0774	13.843	Williams[11a]
O	3.4046	0.0957	13.483	Williams[11d]
F	3.4720	0.0725	14.444	Williams[11b]
Al	4.39	0.31	12.0	interpolation
Si	4.27	0.31	12.0	interpolation
P	4.1500	0.3200	12.0	P_4 crystal
S	4.0300	0.3440	12.0	S_8 crystal
Cl	3.9503	0.2833	13.861	Williams[11c]
Ga	4.39	0.40	12.0	interpolation
Ge	4.27	0.40	12.0	interpolation
As	4.15	0.41	12.0	interpolation
Se	4.03	0.43	12.0	interpolation
Br	3.95	0.37	12.0	interpolation
In	4.59	0.55	12.0	interpolation
Sn	4.47	0.55	12.0	interpolation
Sb	4.35	0.55	12.0	interpolation
Te	4.23	0.57	12.0	interpolation
I	4.15	0.51	12.0	interpolation
Na^+	3.144	0.5	12.0	DREIDING/A
Ca^{2+}	3.472	0.05	12.0	DREIDING/A
Fe^{2+}	4.54	0.055	12.0	DREIDING/A
Zn^{2+}	4.54	0.055	12.0	DREIDING/A
		Implicit Hydrogens		
C_R1	4.23	0.1356	14.034	benzene crystal
C_34	4.2370	0.3016	12.0	CH_4 crystal
C_33	4.1524	0.2500	12.0	interpolation
C_32	4.0677	0.1984	12.0	interpolation
C_31	3.9830	0.1467	12.0	interpolation

Figure 5.2: DREIDING parameters as collected by Mayo *et al [70].*

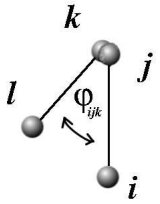


Figure 5.4: The measurement of the dihedral bond.

Valence bond potential

The angle bend interaction is described by the harmonic cosine form:

$$U(\theta_{ijk}) = \frac{k}{2}(cos(\theta_{ijk}) - cos(\theta_0))^2 \tag{5.20}$$

where θ is the angle between atoms ijk, drawn out by the bond vectors:

$$\theta_{ijk} = cos^{-1}\left\{\frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji} r_{jk}}\right\}. \tag{5.21}$$

The force constant k of angle bending is given by $k = 112.49963$(kcal/mol)/ rad^2 as given by Mayo *et al,* where $k = \frac{k'}{(sin\theta_{0j})^2}$ and $k' = 100$ (kcal/mol)/ rad^2. θ_0 is also given by Mayo *et al*, see figure 5.1. Note the potential must be amended to $U(\theta_{ijk}) = k'(1 + cos(\theta_{ijk}))^2$, for linear equilibrium geometries.

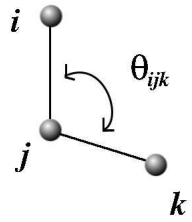


Figure 5.3: The measurement of the valence bond.

Dihedral angle potential

The dihedral angle (or torsional angle) is the angle made by 4 atoms: $ijkl$, it is the angle between the planes made by ijk with jkl.

A torsional angle is periodic by nature, and traditionally is measured from $-\pi < \omega > \pi$, where a clockwise angle is positive and the anticlockwise angle is negative. The potential due to dihedral fluctuations is given by:

$$U(\varphi_{ijkl}) = \frac{1}{2} V_{jk} \left\{1 - cos[n_{jk}(\varphi_{ijkl} - \varphi_{0jk})]\right\} \tag{5.22}$$

$V_{jk}$is the barrier to rotation, $n_{jk}$is an integer for periodicity, $\varphi_{0jk}$is the equilibrium dihedral and $\varphi$is the dihedral found from:

$$\varphi_{ijkl} = cos^{-1} \left\{ \frac{(\vec{r}_{ij} \times \vec{r}_{jk}) \cdot (\vec{r}_{jk} \times \vec{r}_{kl})}{|\vec{r}_{ij} \times \vec{r}_{jk}| \, |\vec{r}_{jk} \times \vec{r}_{kl}|} \right\}. \tag{5.23}$$

Note V_{jk} is the *total* barrier to rotation, so for a total barrier like $V_{jk} = 2.0$ kcal/mol, where there are many dihedrals V_{ijkl} with a common j and k, and different possibilities of i and l, then the energy is distributed over the number of these events. For example, for cyclohexane we define 54 dihedrals for the potential: there are 6 single bond dihedrals between the 6 C atoms, for each of which there are 9 possible ways of measuring this angle. For a case such as this the parameters we require are $V_{jk} = 2.0$ kcal/mol, $n_{jk} = 3$ and $\varphi_{0jk} = 180^{o}$, corresponding to a dihedral single bond with two sp^3atoms, as in cyclohexane. For a full description of dihedral parameters see Mayo *et al*, and figure 5.2.

As mentioned above, a simulation is only as accurate as the force field. In this thesis MD is used predominantly for conformational analysis of an isolated gas phase molecule, so we must be sure to carefully describe the *intra*-molecular potential, whilst on the other hand the *inter*-molecular potentials (we do not yet know what they are) are less crucial, because we are looking for the flexibility properties of a *single* molecule. Conformational analysis involves the examination of all possible stationary points, for a flexible molecule the potential then must allow multiple equilibrium angles, which equation 5.22 does. A common example of conformational analysis is the contrast of two conformers of ethane: staggered, and eclipsed. The staggered conformation is the most stable form because the overall van der Waals (steric) hindrance is minimized; the hydrogen atoms are the most far apart. So for the dihedral potential in DL_POLY the three parameters V_{jk}, n_{jk} and φ_{0jk}, are defined for each atom1-atom2-atom3-atom4, but also the last two columns of data are required (the van der Waals scaling factor and the electrostatic scaling factor) because these forces influence dihedral angles.

van der Waals interactions

It is assumed that 2 and 3 atom interactions are included properly in the bond stretching and angle bending potentials. However, consideration of the 4 atom interactions (as in the dihedral angle potential above) highlights the importance of including short range van der Waals interactions for 1,4 near neighbour atoms, as these will not be included in the other potentials, and they play an important role in the molecule reaching a sterically favourable configuration; those atoms mustn't get too close. Therefore in the force field we need to indicate for which 4 set of atoms we might need to include this term. The van der Waals potential is given by:

$$U(r_{ij}) = \left(\frac{A}{r_{ij}^{12}}\right) - \left(\frac{B}{r_{ij}^{6}}\right) \tag{5.24}$$

where now the ij specify over atom *types* not their indices. The coefficients are given by: $A = D_0 R_0^{12}$ and $B = D_0 R_0^6$ for homo-nuclear atoms, and can be found from figure 5.2. In DREIDING the same van der Waals parameters are used independent of the hybridization. For the heteronuclear atoms usually the default is a geometric mean for $D_{0ij} = [D_{0ii} D_{0jj}]^{\frac{1}{2}}$and an arithmetric mean for $R_{0ij} = \frac{1}{2}(R_{0ii} + R_{0jj})$ to find A and B.

Similarly to above, electrostatic interactions will be accounted for by 1,2 and 1,3 interactions, but also in the force field the electrostatic interactions for 1,4 atom interaction can be defined.

5.5 Calculations using DL_POLY

5.5.1 Construction of input files

For a DL_POLY simulation 3 files are required; the CONTROL, the CONFIG and the FIELD file.

CONTROL file

The CONTROL file in DL_POLY is the file that stipulates the important commands for the type of job to be done. System variables such as the temperature and pressure can be set, and the type of ensemble is stipulated. Each record holds commands necessary for the calculation and is of the form:

Keyword [*options*] {data}

The file must be terminated with a *finish* directive. Definitive instructions for the construction of a CONTROL file can be found in the DL_POLY manual. This is also where instructions can be made to the collection of data to the OUTPUT file. For example you can define the frequency with which data is printed (such as the total energy, volume, total temperature) thus the progress of the simulation can be appropriately monitored. The important constraints required for the calculations presented here are: the number of steps, the equilibration, and the time step (in ps). The other important consideration is the definition of output to the HISTORY file. This key word is *traj* with options [i j k], where i denotes the step where coordinates are first recorded to the HISTORY file, j denotes the interval of steps that is recorded thereafter and k denotes the information it records (here just the coordinates, there are no requirements for the velocities and forces). The final important component is the potential 'cut off'. As remarked above, we are really only concerned with the short-range van der Waals. Usually the long range potential cut off can be approximated as half the unit cell, if there are periodic boundary conditions. In this simplified model, a cut off for the short range potential of about 15% more than the longest length of the molecule (~6 Å) should suffice. Other considerations are greatly simplified. I choose the electrostatic interactions to be ignored in the CONTROL file, following the DREIDING philosophy that the simple case gives the most accurate result (see Mayo *et al*) where they find that including electrostatic interactions for isolated or crystal molecules, actually gives a worse result.

CONFIG file

The CONFIG file contains important starting information about your system. The first record allows a short description of the system. The second gives two keys for the type of information you are supplying. The first key denotes 0 = coordinates, 1= coordinates and velocities, or 2 = coordinates and velocities and forces. The second key determines the type of periodic boundary that may be required. Depending on the above two keys the rest of the molecule must be defined. Again it is greatly simplified for isolated gas phase molecules, just the x, y, z components are required for the atoms involved.

FIELD file

The FIELD file describes the potentials for your system. The first few records give general information: the name of the system, the units, the types of molecules, the number of molecules and the number of atoms. Then the field is described, referring to the potentials outlined above depending on the system requirements, under the headings ATOMS, BONDS, ANGLES, DIHEDRALS, with the corresponding parameters. This information is terminated with a *finish* directive. Then the intermolecular potential is described under the heading VDW for each interacting atom type. The whole potential is terminated with the *close* directive.

5.5.2 Checking calculations

The output files from DL_POLY are OUTPUT, HISTORY, STATIS, REVCON, REVIVE, RDFDAT and ZDNDAT. The first three are the most important for the purposes in this thesis: the HISTORY file for examination of conformations and the STATIS and OUTPUT files are used for simulation checking. Important checks may be:

1. The total energy (potential + kinetic) should be conserved. Plotting the total energy against time should be done to verify there is no systematic drift over time.

2. In NVE MD the linear momentum is conserved, thus the initial and final momentum should remain zero.

3. Though the temperature is stipulated, it will oscillate about this value, defined by the kinetic energy as in equation 5.12. A plot of temperature T should be made against time to check for any wild spikes in the data.

A final other check is to visualize the data to check against expectation, *i.e.*, what makes sense physically and chemically, which is an important part of the analysis in the Chapter 'Chirality and conformations of odorants'.

Chapter 6

Density Functional Theory

Molecular dynamics, though incredibly helpful, only examines closely the nuclear behaviour, and the macroscopic properties that can be derived from which. For some material properties however, such as the stiffness of a metal, the band gap of a semi-conductor, the diffusion of a defect or the colour of a pigment, we need to look more closely at the electron 'glue' that holds these solids together. This involves understanding the behaviour of the interacting electrons in response to the presence of atomic nuclei or ionic cores. The usual way to do this is via solution of the *many-body* time-independent Schrödinger equation:

$$\hat{H}\mathbf{\Psi} = E\mathbf{\Psi}. \tag{6.1}$$

As mentioned previously, $\hat{H}$ is the Hamiltonian operator, E the energy of the system and Ψ is the many-body wave function. Accounting for every interaction of each particle in this way is supposed to sum up to and account for the overall total behaviour. Difficulty lies in the definition of Ψ, and we shall see we have to make certain approximations to find solutions to the equation 6.1. This area of science is referred to as *ab initio* electronic structure theory, and there are many sources that definitively describe methods involved, see Szabo and Ostlund for example [53]. I shall not reiterate these methods here, as the onus of this chapter is a bit more particular; it is to describe a method used in calculations throughout this thesis; *density functional theor*y (DFT), because it accurately describes the total ground-state energy of a system of interacting electrons as a position of the nuclei. The types of calculations where DFT can be appropriate are for molecular:

1. Geometries (see Chapter 'Chirality and conformations in odorants' section 9.2.1).

2. Conformational energies (see Chapter 'Chirality and conformations in odorants' section 9.2.1).

3. Dipole moments (see Chapter 'Huang-Rhys factors' section 8.8).

4. Electrostatic potential fitted partial charges (see Chapter 'Chirality and conformations in odorants' section 9.3.2.2).

5. Huang-Rhys factors (see Chapter 'Huang-Rhys factors').

6. Vibrational frequencies (see Chapter 'Huang-Rhys factors' section 8.4).

To name a few. Further, I describe methods for Gaussian '03 users.

6.1 The problem

Refer back to equation 1 in the Chapter 'A physical picture'. This is a many-body wavefunction as it describes all the interactions in one molecule due to all the electrons and nuclei. In a polyatomic molecule the description of all these interactions is a formidable task, even computationally. To even make the Schrödinger equation machine solvable the first step to simplify the task is to use the previously mentioned Born-Oppenheimer (BO) approximation. This allows us to say that the nuclei in a molecule are fixed and the electrons are moving about them in this fixed potential. Thus solving for just the stationary *electronic* Hamiltonian:

$$\begin{aligned} \hat{H}\mathbf{\Psi}(\mathbf{r}_1, ...\mathbf{r}_\mathrm{N}) &= E\mathbf{\Psi}(\mathbf{r}_1, ...\mathbf{r}_\mathrm{N}) \\ &= \left[\hat{T} + \hat{V} + \hat{U}\right] \Psi(\mathbf{r}_1, ...\mathbf{r}_N), \end{aligned} \tag{6.2}$$

the complexity of the whole system is already reduced. N is the number of electrons, $\hat{T}$ is the electron kinetic energy, $\hat{V}$ is the electron potential energy (due to the fixed field of the nuclei) and $\hat{U}$ is the electron-electron interaction energy. Often $\hat{T}$ and $\hat{U}$ are teamed together as $\hat{H}_0$, because this is the Hamiltonian for the electron system

itself, without interaction with the external field due to the nuclei. Written more explicitly:

$$\hat{T} = -\sum_{i=1}^{N} \hbar^2 \frac{\nabla_i^2}{2m_e}, \tag{6.3}$$

$$\hat{V} = -\sum_{i=1}^{N}\sum_{A=1}^{M} \frac{e^2 Z_A}{4\pi\varepsilon_0 \mathbf{r}_{iA}} = \sum_{i}^{N} v(\mathbf{r}_i), \tag{6.4}$$

$$\hat{U} = \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{e^2}{4\pi\varepsilon_0 \mathbf{r}_{ij}} = \sum_{i\neq j}^{N} u(\mathbf{r}_i, \mathbf{r}_j), \tag{6.5}$$

where $\hat{T}$, $\hat{V}$ and $\hat{U}$ are just the first, third and fourth terms from equation 1 in the Chapter 'A physical picture', i labels the N electrons and A labels the M nuclei, but instead of $\hbar = e = 4\pi\varepsilon_0 = m_e = 1$, they have been included here to show definitively the interactions. The terms have also been written in a simplified hand for brevity (for $\hat{V}$ and $\hat{U}$) because they will be referred to later, especially $v(\mathbf{r})$; as the *external potential*. The second and fourth terms from equation 1 in chapter 'A physical picture' can be neglected because of the BO approximation; the first because of the significantly larger mass term and the fourth because the nuclear-nuclear repulsion energy is a constant term for a given set of fixed nuclear coordinates. Wavefunctions are invariant to constant terms in the Hamiltonian, so this term can be included on the eigenvalue for the electronic energy, but is not required for solving for these eigenvalues. The number of interactions are now reduced, however, the wavefunction Ψ is dependent on $(\mathbf{r}_1 s_1; ....; \mathbf{r}_N, s_N)$; one spin and three spatial coordinates for every electron (at fixed nuclear positions), so despite these simplifications, solving the electronic Hamiltonian in this form still involves too much data.

6.2 Thomas-Fermi density functional theory

The Hamiltonian in equation 6.2 now only requires the position and atomic numbers Z_A of the nuclear coordinates and the number N of electrons. In order to obtain these parameters and avoid some quantum mechanical complexity, Thomas

and Fermi suggested to look at a physical observable; the electron density:

$$\rho(\mathbf{r}) = \int d^3\mathbf{r}_1 \int d^3\mathbf{r}_2 .... \int d^3\mathbf{r}_N \left|\mathbf{\Psi}(\mathbf{r}_1,\, \mathbf{r}_2 ... \mathbf{r}_\mathrm{N})\right|^2 . \tag{6.6}$$

Integration of the electron density, over all space, gives us N, the total number of electrons. Further, where there is a maximal electron density, the nuclear positions can be deduced, and the assignment of nuclear atomic number Z_A. Thus we can write the above components of energy in terms of this electron density:

$$V\left(\rho(\mathbf{r})\right) = -\sum_A^M \int \frac{Z_A}{r_{iA}} \rho(\mathbf{r}) d\mathbf{r}, \tag{6.7}$$

for the interaction of the electrons and nuclei, and

$$U\left(\rho(\mathbf{r})\right) = \frac{1}{2} \int \int \frac{\rho(\mathbf{r}_i)\rho(\mathbf{r}_j)}{\mathbf{r}_{ij}} d\mathbf{r}_i d\mathbf{r}_j \tag{6.8}$$

for the interaction of electrons and electrons i and j, all over space[69]. We return to $\hbar = e = 4\pi\varepsilon_0 = m_e = 1$ for brevity. Note these energy contributions are functions that depend on functions and are therefore called *functionals*. The most difficult part that remains, is to account for the continuous charge distribution in the kinetic energy term. For this Thomas and Fermi derived:

$$T\left(\rho(\mathbf{r})\right) = \frac{3}{10}(3\pi^2)^{2/3} \int \rho^{5/3}(\mathbf{r}) d\mathbf{r} \tag{6.9}$$

which assumes a particular electronic distribution; in fact a *uniform electron gas* that has a constant non-zero density. Note these energy components can be calculated, without considering the wave function and we can approximate a ground state energy of a molecule, just by looking at its particle density.

6.3 The Hohenberg-Kohn theorems

The principle above, of avoiding the wavefunction and its complications, was popular for a while, but was not wholly accurate [69], because there is no *exact* solution. Enter Hohenberg and Kohn with solutions for the quantum mechanical formalism in the form of two important theorems.

Theorem 1: It is impossible that two different external potentials $v(\mathbf{r})$

give rise to the same ground-state density distribution $\rho(\mathbf{r})$.

This first proposition tells us a one-to-one mapping exists between the external potential and the ground state density:

$$v(\mathbf{r}) \leftrightarrow \rho(\mathbf{r}) \tag{6.10}$$

One dictates the other, and *vice versa*. The proof of this theorem can be made by assuming the contradiction of this statement, *i.e.*, that it is possible for 2 different potentials $v(\mathbf{r})$ and $v'(\mathbf{r})$ to give us the same ground state, but we end up with the contradiction $E_0 + E_0 < E'_0 + E_0$: for a complete proof see Parr and Yang [71]. Implementing this theorem the many body wave function can now be written as

$$\mathbf{\Psi}(\mathrm{r}_1, \mathrm{r}_2 ... \mathrm{r}_\mathrm{N}) = \Psi[\rho(\mathbf{r})]. \tag{6.11}$$

The consequence of this is that *all* observables are now functionals of the density. The ground state (lowest stable state) is given by

$$E_g[\rho(\mathbf{r})] = \left\langle \Psi[\rho(\mathrm{r})] \left| \hat{H} \right| \Psi[\rho(\mathrm{r})] \right\rangle \tag{6.12}$$

and the excitation energies are $E_{ex} = E_{ex}[\rho(\mathbf{r})]$. Thus the density contains a wealth of information. We can write, for the ground state energy functional:

$$\begin{aligned} E_g[\rho] &= \left\langle \Psi[\rho] \left| \hat{T} + \hat{V} + \hat{U} \right| \Psi[\rho] \right\rangle \\ &= \left\langle \Psi[\rho] \left| \hat{T} + \hat{U} \right| \Psi[\rho] \right\rangle + \int dr^3 v(\mathbf{r}) \rho(\mathbf{r}), \end{aligned} \tag{6.13}$$

where we call the first term $F[\rho]$, and we have separated out the terms not dependent on the external potential. Now we come to the second theorem.

> Theorem 2: The ground-state energy for a given $v(\mathbf{r})$ is obtained by minimising $E_g[\rho]$ with respect to $\rho(\mathbf{r})$ for fixed $v(\mathbf{r})$. The $\rho(\mathbf{r})$ that gives the minimum is the ground state.

This theorem implements the *variational principle* for the ground state, where the average measurement for the energy observable is [71]:

$$E\left[\Psi\right] = \frac{\langle\Psi|\,\hat{H}|\mathbf{\Psi}\rangle}{\langle\Psi\,|\Psi\rangle},$$

then $E\left[\Psi\right] \geq E_g$, which is tantamount to saying the energy found from any initial guessed wavefunction is an upper bound to the true ground state energy. This second theorem works by analogy to the variational principle, we have $E_g[\rho] \geq E_g$. Again for proof of this theorem see Parr and Yang [71].

6.4 The Kohn-Sham scheme

So the ground state energy of a many electron system can be obtained, as above, from the minimum of the energy functional

$$E_g[\rho] = F[\rho] + \int dr^3 v(\mathbf{r})\rho(\mathbf{r}), \tag{6.14}$$

where $F[\rho] = T[\rho]+U[\rho]$ [71]. Refer back to equation 6.13. The term $\hat{U}$, the electron-electron interaction energy, contains the classical repulsion given in equation 6.8, which we now call $J[\rho]$, but for completeness we need to account for a non -classical term:

$$U[\rho] = J[\rho] + non - classical\, term.$$

By analogy to the Hohenberg-Kohn method, the Kohn-Sham technique is to write the functionals:

$$F[\rho] = T[\rho] + J[\rho] + E_{xc}[\rho],$$

where

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] + U[\rho] - J[\rho].$$

$E_{xc}[\rho]$ is the *exchange-correlation energy*, $T_s[\rho]$ is the kinetic energy of the *non-interacting* electrons and $T[\rho]$ is the kinetic energy for *interacting* electrons (and the difference between these two energies is taken to be small). $T_s[\rho]$ only holds for densities that are ground state densities for some well-defined non-interacting system. What is required now is an exact value for the exchange-correlation energy

(see below) and the kinetic energy part. The kinetic energy part is determined by artificial Kohn-Sham one-electron orbitals φ_i^{KS}:

$$\rho(\mathbf{r}) = \sum_i^N \sum_s \left|\varphi_i^{KS}(\mathbf{r}, s)\right|^2, \tag{6.15}$$

which are orthonormal wavefunctions, they are called *artificial* because they are constructed as solutions to the Hamiltonian for a non-interacting system of electrons moving in an *effective potential* v_{eff}:

$$-\frac{\hbar^2}{2m}\nabla^2\varphi_i^{KS}(\mathbf{r}, s) + v_{eff}(\mathbf{r})\varphi_i^{KS}(\mathbf{r}, s) = \varepsilon_i\varphi_i^{KS}(\mathbf{r}, s). \tag{6.16}$$

The problem is thus reduced to a one-particle system; the Hamiltonian is a sum of one-electron operators, the eigenfunctions are *Slater* determinants of individual one-electron operators and the eigenvalues are the sum of one-electron eigenvalues ε_i, which are meaningless unless they are summed to give the total kinetic energy we are seeking. The effective potential is [71]:

$$v_{eff}(\mathbf{r}) = v(\mathbf{r}) + \frac{\delta J[\rho]}{\delta\rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})} \tag{6.17}$$

These *Kohn-Sham equation*s are used thus; a $\rho(\mathbf{r})$ is guessed for the initial $v_{eff}(\mathbf{r})$ in equation 6.17. Then equation 6.16 is solved to obtain the Kohn-Sham orbitals φ_i^{KS}, and hence a new $\rho'(\mathbf{r})$ is found. This process is reiterated until the difference $\rho(\mathbf{r}) - \rho'(\mathbf{r})$ is minimal, within a certain criterion. Thus the Kohn-Sham process is referred to as an iterative *self-consistent field* (SCF) procedure.

6.5 Exchange -correlation functionals

The effective potential is calculated from the exchange-correlation functional:

$$E_{xc}[\rho(\mathbf{r})] = \int \rho(\mathbf{r})\varepsilon_{xc}[\rho(\mathbf{r})]d\mathbf{r}, \tag{6.18}$$

which is the exchange-correlation energy per particle with density $\rho(\mathbf{r})$ [69]. The dependence of E_{xc} is written as an electron density, and an energy density ε_{xc}, which is a sum of the individual exchange and correlation contributions. The *exchange* corresponds to the energy difference when the spatial coordinates of two electrons are

interchanged. The *correlation* corresponds to the energy difference from the electrons being represented by the density, not as point charges; when we use densities the interaction that occurs when electrons move, and their effect on each other, is lost. Values of ε_{xc} for a homogeneous electron gas are usually found from *Monté Carlo* calculations of the energy of a homogeneous electron gas at various densities, which we will not explore, but a brief discussion of methods for ε_{xc} is given below.

The local density approximation (LDA)

In the local density approximation we treat $\varepsilon_{xc}[\rho]$ as a function of $\rho(\varepsilon_{xc}[\rho(\mathbf{r})])$. The value of ε_x for the *exchange*, is found from ρ at a certain position r, and hence determines a *local* value of ρ:

$$\varepsilon_x[\rho(\mathbf{r})] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\rho^{1/3}(\mathbf{r}). \tag{6.19}$$

The value of ε_c for the *correlation*, is determined by many workers, especially well for metals, from the assumption of a local uniform-electron gas [72]. The combination for a total function of density $E_{xc}[\rho(\mathbf{r})]$ is tabulated and parameterized. This method is very accurate ($\sim 0.1\%$) [73], especially if $\rho(\mathbf{r})$ varies over a large length scale.

The local spin density approximation (LSDA)

The LDA method above can be extended to include spin-polarization using [69]:

$$\varepsilon_x[\rho(\mathbf{r}),\zeta] = \varepsilon^0_{xc}[\rho(\mathbf{r})] + \left\{\varepsilon^1_{xc}[\rho(\mathbf{r})] - \varepsilon^0_{xc}[\rho(\mathbf{r})]\right\}\left[\frac{(1+\zeta)^{4/3}+(1-\zeta)^{4/3}-2}{2\left(2^{1/3}-1\right)}\right], \tag{6.20}$$

The first term is equation 6.19, the second term is decided from the consideration of a uniform electron gas composed only of electrons with the same spin. This converges to the unpolarized case in LDA when $\zeta = 0$ and the second term becomes zero. ε_{xc} is a function of the total electronic density and the net spin density.

The generalized gradient approximation (GGA)

This method defines the exchange-correlation energy in terms of spin polarized electron densities ($\rho_\uparrow$ and $\rho_\downarrow$) and their gradients ($\nabla\rho_\uparrow$ and $\nabla\rho_\downarrow$)[74]:

$$E_{xc}[\rho_\uparrow, \rho_\downarrow] = \int d\mathbf{r} f\left(\rho_\uparrow(\mathbf{r}), \rho_\downarrow(\mathbf{r}), \nabla\rho_\uparrow(\mathbf{r}), \nabla\rho_\downarrow(\mathbf{r})\right). \tag{6.21}$$

This is known as the generalized gradient approximation (GGA), as this method is sensitive to a varying electron density. There are many types of correlation functionals for GGA; B86, P86, PW91, [69] , [74]. A popular choice is 'LYP' which computes the correlation energy from four empirical parameters fit to the helium atom. It includes an exact cancellation of the self-interaction error in one-electron systems [69].

The Becke Hybrid functionals

There are various density functional theory methods that describe the exchange-correlation functionals. The method used predominantly for this thesis, is called 'B3LYP' after the Becke Three Parameter Hybrid Functionals, developed by Becke in 1993 [75]. Hybrid functionals include some exact exchange. B3LYP is a combination of Becke's parameters for exchange, and the correlation LYP as discussed above. Becke describes a functional based on experimental atomization energies that give better performance than the gradient (GGA) corrections only. The exchange-correlation energy E_{XC} is given by

$$E_{XC} = \int_0^1 U_{XC}^{\zeta} d\zeta, \tag{6.22}$$

where ζ is the interelectronic coupling-strength parameter that turns on the Coulomb repulsion between electrons, and U_{XC} is the potential energy of exchange -correlation. Though this is a potential energy, the kinetic part is generated from the ζ integration, which goes from 0 (fully interacting Kohn-Sham reference system) to 1 (non interacting Kohn-Sham reference system). At limit $\zeta = 0$, there is *exact* exchange, E_X^{exact}, and Becke argues that this should be included in the functional, moderated by a constant a_0 that reflects the rate of the onset of correlation. For the full

exchange-correlation approximation:

$$E_{XC} = E_{XC}^{LSDA} + a_0 \left(E_X^{exact} - E_X^{LSDA}\right) + a_X \Delta E_X^{B88} + a_C \Delta E_C^{PW91}, \tag{6.23}$$

where a_0, a_X and $a_C$are the semi-empirical coefficients, ΔE_X^{B88} is an LSDA gradient approximation and E_C^{PW91} is the gradient correction for correlation by Perdew and Wang [76] . For greater detail see this and papers I and II by Becke [75].

6.6 Basis sets

The idea of the basis set is to define a set of wavefunctions that adequately account for the electron configuration and describe the molecular orbitals. This is usually done via the orbitals of the individual atoms in a *linear combination of atomic orbitals* (LCAO) [54]:

$$\varphi_i = \sum_{\mu=1}^{K} C_{\mu i} \phi_i \tag{6.24}$$

Thus we can approximate the molecular orbitals φ_i, with atomic orbitals ϕ_i ; the coefficients $C_{\mu i}$ are adjusted to minimize the energy. An accurate expression for an orbital is given by a Slater type function [53]:

$$\phi^{SF}(\zeta, \mathbf{r} - \mathbf{R}_A) = \left(\frac{\zeta^3}{\pi}\right)^{1/2} e^{-\zeta|\mathbf{r}-\mathbf{R}_A|} \tag{6.25}$$

where $|\mathbf{r} - \mathbf{R}_A|$ is the radius in Angstroms from the centre of a function $\mathbf{R}_A$, and ζ is the *Slater orbital exponent*. However, this is an expensive calculation, and it is faster (though the pay off is accuracy) to combine Gaussian functions of the form [53]:

$$\phi^{GF}(\alpha, \mathbf{r} - \mathbf{R}_A) = \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\zeta|\mathbf{r}-\mathbf{R}_A|^2} \tag{6.26}$$

where α is the *Gaussian orbital exponent*.

Minimal basis set STO-3G

This basis is *minimal* because it uses the minimum number of functions for each atom; the *core* orbitals, and the occupied *valence* orbitals of the atom [53]. Thus $1s$ for hydrogen and, $1s$, $2s$, $2p_x$, $2p_y$, $2p_z$, 5 for the second row atoms. Since there are few functions, it pays to make them the best form describable for the shape of the atomic orbitals; so Slater functionals are a good candidate for this purpose. The acronym STO-3G comes from *S*later-*T*ype *O*rbitals, created by adding 3 Gaussian functions together (a fixed linear combination otherwise known as contracted). The Gaussians are adjusted to fit as well as possible the Slater orbitals. These orbitals are contracted and constrained.

4-31G and 6-31G

4-31G and 6-31G are refered to as *split-valence* basis sets, because the atomic orbitals are split into a compact inner orbital and an outer valence orbital, thus a molecule can be better represented in its neutral *and* charged state. The notation here refers to 4 (or 6) contracted Gaussian functions for the core orbitals, 3 contracted Gaussians for the inner valence orbitals and 1 (not contracted) Gaussian for the outer valence orbital.

6-31G* and 6-31G**

The basis sets 6-31G* and 6-31G** (also known as 6-31G(d) and 6-31G(d,p) and double zeta polarization and double zeta 2 polarization respectively) account for the polarization of orbitals due to their interactions with neighbouring other orbitals. For example, a hydrogen atom in isolation is represented by the $1s$ orbital, but when in proximity to another atom it is disturbed by the introduction of a nonuniform electron field, such that an introduction of a p-type function (a hybridization) better describes the overall effect [53]. For atoms with occupied p orbitals, $d-$type functions can be included to account for polarization. The 6-31G* and 6-31G** are in this sense 6-31G with this polarization effect added. So the notation is as above, except one * indicates that d-type uncontracted Gaussian functions are added (to the heavier atoms), and ** indicates the addition of p-type uncontracted Gaussian functions (to the hydrogen atoms).

cc-pVTZ

This acronym stands for a *c*orrelation-*c*onsistent *p*olarized *v*alence (only) basis set with *t*riple *z*eta polarization. This gives even larger shells of polarizing (and correlating) functions than the above.

6.7 Choice of method and basis set

Throughout this thesis a density functional B3LYP is used with a basis set of 6-31G**. This method is known to reproduce fairly robust geometries for organic compounds [69], and B3LYP is supposed to give energies that are correct with a very small absolute deviation of 2.4 kcal/mol [75], and very good estimates for population analysis [77]. The problem with LDA/LSDA is that most molecules have varying non-uniform electron densities, which are not well represented by these methods. LSDA gives moderately good molecular structures, vibrational frequencies and charge densities and GGA gives better thermo-chemistry, energetics and structure of H-bond networks. Neither LSDA nor GGA are good for barrier heights [73]. Further the LSDA functional makes systematic overestimation of bond strengths and is inaccurate for biomolecular modelling that involves hydrogen bonds [78]. The LSDA methods would not be used to its best capabilities in this work, which is analysis of neutral molecules (closed shell). The correlation effects introduced by Becke are generally perceived to be more expensive computationally, but very accurate for small organic compounds (such as odorants). In conjunction with this method, an extended basis set is usually used. The higher orbital extended basis sets are considered superior for more accurate representation of the molecular charge density, and double polarized basis sets are more accurate for the IR absorbency calculations which are important considerations we come to in the Chapter 'Huang-Rhys factors' [79]. 6-31G* and 6-31G** basis sets are considered superior to STO-3G and 4-31G levels in this sense [53]. STO-3G is good for geometries but not for energies, especially small rings , which again is something we come to in later sections.

6.8 Calculations using Gaussian '03

Gaussian '03 is an *ab initio* quantum chemistry package, and has been the tool used for the electronic structure calculations described in this thesis. Here I describe the types of calculations involved, with some instructions for anyone wishing to recreate similar calculations using these density functional methods.

6.8.1 Single point energies

The simplest (default) calculation using Gaussian is a *single point energy* calculation to full *self-consistent field* (SCF) consistency. All that is required is your method and basis set of choice and a starting molecular geometry. From this point all one-electron and two-electron integrals are calculated. Wavefunctions are estimated from this starting point, and improved upon by choosing an electron and finding the average field on the electron, given all the others. A new one-electron density matrix is then constructed. This is done repeatedly until the density matrix fails to show any appreciable difference from the previous calculation. The positions of the atoms are unaltered.

Input file

For a template example of an input file in Gaussian see figure 6.1.

```
%chk=HCl_SCF

#B3LYP/6-31** SCF=Tight

A single point energy calculation for hydrogen chloride

0 1
H
Cl 1 R1

R1=1.286132
```

Figure 6.1: A template for Gaussian input. This particular job describes a single point energy calculation for HCl at a stipulated bond length, with tight SCF convergence criterion.

In the first record the user defines a name for the checkpoint (.chk) file that

will be printed in the output (likely based on your test molecule). The .chk file records all useful information that will be necessary for a continuing calculation. Also Gaussian has utilities that can be used on the machine readable .chk file, such as *freqcheck* which extracts useful information in analysis, such as the vibrational frequencies. This record is blank line terminated. The next record is the *route card*, this describes the method/basis set you wish to use followed by any keywords needed to define the type of calculation you wish to do. In this example *SCF=tight* is used. Any DFT single point energy calculation using a basis sets with diffuse functions should use this keyword to request the tight SCF convergence criteria [80]. This record too is blank line terminated. The next record is used to describe your job, for future reference (this information is not read by the program). The last part of the input describes the system; the charge (0 here for a neutral molecule) and the spin multiplicity (1 for a singlet), and the starting geometry of the system. Here the geometry is defined in z-matrix notation, but often Cartesian xyz coordinates will be used, depending on convenience in the calculation. Thus the input in Gaussian is simply two important parts, the rest is non-mandatory; the route card that describes the calculation, and the system specification.

Checking calculations

In a calculation such as this, search in the output for the final SCF value:

SCF Done: E(RB+HF-LYP) = -460.818518782 A.U. after 10 cycles

The letters in the parenthesis note the method used, here B3LYP. In this example, the job took 10 iterations to reach convergence. If the SCF energy is not getting more negative, or is oscillating, then there is a problem with the calculation.

6.8.2 Geometry optimization

In a geometry optimization, after a single point energy has been found from the starting point, then the positions of the atoms are moved according to the optimization algorithm chosen (in Gaussian this is a Berny algorithm using redundant internal coordinates [80]) and then the SCF procedure is repeated. This is done again iteratively until the optimization criteria are met; when the forces on atoms are

close enough to zero given precision. So at each new position analytic first derivatives of the energy are found and this continues until a stationary point is reached. An efficient way to find a close starting structure is to build the molecule using a package like ArgusLab [81] and optimize the geometry using a fast semi-empirical method. Usually then, unless the system is complicated, the calculation at a higher level of theory such as B3LYP/6-31G** will not require too many iterations to reach the geometry desired.

Input file

For a geometry optimization the key words:

#B3LYP/6-31** SCF=Tight Integral(Grid=UltraFine) Opt=Tight

can be used in the route card. This specifies not just tight convergence for the SCF but also for the optimization, over an ultra-fine integration grid (DFT methods use numerical integration). This combination of an accurate grid for integration with tight optimization criterion is important for exact conformations of molecules and for molecular systems with small force constants (low frequency vibrational modes), and so is recommended for reliable frequencies [80]. Though the small molecules tested in this thesis may have relatively high frequencies (compared to proteins) given the onus in this thesis on the frequency calculation, and there is high commendation of this method by other users for similar ligands [82], this combination was implemented in calculations throughout this thesis.

Checking calculations

Sometimes the geometry optimization terminates in error, and it is usually because the calculation does not satisfy the convergence criterion within the default cycles. Monitoring the SCF energy levels at each output is important; if the values are oscillating and not converging it is probably because the molecule is 'stuck' at an unfruitful geometry. Restarting a calculation at the geometry before this point may avoid this, and/or it may be best to tweak the geometry from near this position, and increase the SCF cycles in the next calculation with the keyword *MaxCycle=N*, *i.e.*, you can increase N the number of optimization steps (to N=100 for example). If a job is terminating because of a bad geometry, than including the keyword *iop(5/13=1)*

will force the calculation to continue optimizing despite convergence failure, but it should be used with care; it is important to check the resulting geometry makes sense and is actually a minimum geometry.

```
         Item               Value     Threshold  Converged?
Maximum Force            0.000001     0.000015     YES
RMS     Force            0.000000     0.000010     YES
Maximum Displacement     0.000027     0.000060     YES
RMS     Displacement     0.000005     0.000040     YES
Predicted change in Energy=-2.829118D-10
Optimization completed.
   -- Stationary point found.
```

Figure 6.2: An example of Gaussian output for a successful geometry optimization; convergence criterion are met. This is for decaborane, a tight convergence in both SCF and optimization using B3LYP/6-31**.

Sometimes a job only converges on the *basis of negligible forces*, in other words, the displacements criterion are not met. In this case it is wise to check the geometry and the frequencies (as below) to investigate the true nature of the stationary point.

6.8.3 Vibrational analysis and InfraRed absorbancy

As we have seen in the theory described (Chapter 'A physical picture') a frequency calculation computes the second derivative of the energy (the force constants) and the resulting vibrational frequencies by solving the Hessian matrix.

Input file

The success of the geometry optimization above is integral; at this point (the relaxed geometry) *only* must the frequencies be calculated, and using the same method and basis set. For this reason the two jobs are usually concatenated and called for in the route card as in the above section but with the keyword:

#P B3LYP/6-31** SCF=Tight Integral(Grid=UltraFine) Opt=Tight Freq

Freq added at the end. Often also *P* is included right at the beginning because it requests additional information that maybe useful, *i.e.,* bar-graph spectra for small molecules and polarizability tensors.

Checking calculations

After checking the calculation begins at the end of a relaxed geometry, it is important to check the resulting low frequencies, see figure 6.3, for example. These correspond to the translations and rotations that are projected out of the Hessian. In the example below, figure 6.3, the translations are very close to zero and the rotations are below 50 cm^{-1}, which is good for a DFT method [83]. Another thing to verify is that the second line of low frequencies, match the frequencies printed out below; the normal modes of vibration. If they do not this indicates that the translation and rotation modes have not been removed properly from the Hessian containing the vibrational modes [83]. If the nature of the stationary point is in contention, it is also best to check if these eigenvalues are real; if so a true minimum has been found. If there are any negative (imaginary frequencies) then you can identify possible transitions states, one negative frequency suggests a a *first order transition state* and two imaginary frequencies a *second order transition state*.

```
Full mass-weighted force constant matrix:
Low frequencies ---    0.0004    0.0005    0.0008   15.1507   15.7820   18.4012
Low frequencies ---  231.2139  339.0967  371.6392
Diagonal vibrational polarizability:
       4.1792024       6.2058199       2.9934527
Harmonic frequencies (cm**-1), IR intensities (KM/Mole), Raman scattering
activities (A**4/AMU), depolarization ratios for plane and unpolarized
incident light, reduced masses (AMU), force constants (mDyne/A),
and normal coordinates:
                     1                      2                      3
                     A                      A                      A
Frequencies --   231.2131               339.0967               371.6362
Red. masses --     2.4727                 3.5618                 4.8485
Frc consts  --     0.0779                 0.2413                 0.3945
IR Inten    --     0.0000                 1.2950                 4.5540
Atom AN      X      Y      Z        X      Y      Z        X      Y      Z
  1   1     0.09  -0.08  -0.30     0.02  -0.01  -0.01    -0.27   0.00   0.10
  2   5     0.00   0.00   0.00     0.00  -0.02   0.11     0.00   0.16  -0.10
  3   1    -0.15   0.26  -0.07    -0.23  -0.05   0.02    -0.13   0.08  -0.04
  4   5     0.00   0.04   0.00     0.00   0.00  -0.06     0.00  -0.08   0.00
```

Figure 6.3: An example of a Gaussian output for a successful frequency calculation. This is for decaborane as above again a tight convergence in both SCF and optimization using B3LYP/6-31** .

6.8.4 Potential energy surfaces

Optimized geometries are most successful when you have a good starting geometry, so in a sense you must already find (roughly) what it is (exactly) you are looking for. Potential energy surfaces (PES) are easy for diatomics (see Chapter 'A physical picture') but far more complicated for even small molecules like odorants. A potential energy search can be made in Gaussian using a search algorithm, to find the pathway from the starting point (usually a global minimum) a final point (usually a local minimum) and the transition state (you give a guess for this structure also), and the calculation finds the points in between. Though more often than not, the search is unsuccessful. It is more useful to make structures on a program like ArgusLab, or even using a Dreiding ball and stick model and to find various stationary points, based on intuition. In this work a lot of the odorants have 6-membered rings, and so exhibit potential energy pathways involving similar geometries to the well established system of cyclohexane. Once various important geometries along the pathway have been established, there is usually a clear deviation of some bond angle that links and transforms one conformer state of the molecule to another. In the case of cyclohexane, the incrementation of one dihedral can be used to define the potential energy surface and conformational sampling of the molecule. Once the initial identification of a pathway have been done, verification of the states can then be made by the searching algorithm.

Input file

To find a potential energy surface using Gaussian a constrained optimization is appropriate, and this can be called by using the keywords:

#B3LYP/6-31G** Opt=(modredun) NoSymm

in the route card. *NoSymm* is useful because it stops any reorientation of the molecule, which would cause difficulties if the symmetry of the molecule changes. The *modredun* requests modification of the redundant internal coordinate definitions and requires that after the geometry specification a line is added, to define the modifications you require, for example :

N1 N2 N3 N4 **S** nsteps stepsize [[min]max]

N describes the atom number, *S* calls a relaxed potential energy surface scan [80]. If you define 4 atoms, the angle between them (the dihedral) can be incremented a size of *stepsize* degrees for *nsteps* steps. At each increment the geometry is *partially* optimized (partially because the angle of the dihedral must be preserved) and the energy for this geometry is found.

A plot from the above results allows identification of the important stationary points, then the search algorithm mentioned above can be used for further verification: the *Synchronous-Transit-Guided-Quasi-Newton (STQN)* method [84], [85]. Such a job can be called by the route card

#P B3LYP/6-31* Opt=(QST3,calcfc,eigentest) iop(5/13=1) Freq

The keyword *QST3* calls the STQN algorithm, *calcfc* requests the force constants at the starting point and *eigentest* tests the curvature of the Berny optimization which is useful in determination of a transition state. Keyword *iop(5/13=1)* can also be used here, because complete convergence is not required (we are looking for intermediate states not optimum geometries) and a frequency analysis is useful for the states found, because then the nature of the stationary points can be analyzed, as above. The system specification must include the initial, final and the intermediate geometries.

Checking calculations

At the end of a successful PES scan a search for the energy eigenvectors can be made and extracted, as in figure 6.4.

Summary of Optimized Potential Surface Scan

	1	2	3	4	5
EIGENVALUES --	-235.88035	-235.87971	-235.87857	-235.87704	-235.87524
R1	1.53750	1.53838	1.53934	1.54059	1.54212
R2	1.53821	1.54030	1.54272	1.54538	1.54826
R3	1.09756	1.09729	1.09692	1.09657	1.09638
R4	1.10070	1.10095	1.10122	1.10107	1.10089

Figure 6.4: An example of the output from a potential energy scan for cyclohexane, using the method described.

6.8.5 Population analysis, charge, dipole moments and ESP

The issue of charge is contentious, and the methods of determination varied. A common way of attributing charge is to assign fractions to each individual atom in a molecule. These *partial* charges however, ignore the wave-like nature of electron distribution. One method of determination of a partial charge (that used by Gaussian unless the route card stipulates otherwise) is *Mulliken population analysis*, where in DFT, the molecular wavefunctions are partitioned into atomic contributions depending on the Kohn-Sham orbitals φ_i^{KS} [86]. A second method uses the electrostatic potential (ESP) to derive charge assignments. The molecular electrostatic potential (MEP) can be calculated; the degree to which a molecule is repelled/attracted by a point charge, and mapped onto a van de Waals surface of the molecule. The degree of negative to positive potential is often indicated by gradation in colour. This typically gives a good idea of a molecules overall charge distribution. Matching the ESP to a monopole expansion gives us assignment for partial charges, see Besler *et al* (1990) for details on the algorithm used to this end [87].

Input files

By default Gaussian uses the density matrix for population analysis. The population is analyzed at the end of a single point calculation, or at the first and last points of a geometry optimization. For detailed information based on the DFT density matrix, then a route card such as:

```
#P B3LYP/6-31** SCF=Tight Pop=Full Density=Current
```

can be used. The keyword *Pop=Full* requests the printing of all orbitals (occupied and virtual) together with the atomic charge distribution (Mulliken) and multipole moments. Thus the dipole moment (in Debyes), and higher order moments, is easily extracted from the output. The *cubeman* utility can be used on this output to extract cubes of electron density and electrostatic potential from the calculation, which is useful for analysis using software such as GopenMol, where such surfaces can be plotted and visualized.

For a partial charge calculation, then:

```
#B3LYP/6-31G** Pop=MK
```

requests a population analysis based on the atomic charge assignment via the Merz-Kollman (MK) algorithm [87]. This technique is described in Chapter 'Huang-Rhys factors', where partial charges on atoms are required.

Checking calculations

The dipole moment can be checked against experimental values. The atomic centered charges, however, derived from fitting molecular electrostatic potential to outside a van der Waals radius, cannot be compared directly with experimental evidence. They can only be tested in comparing with observables such as the free energy of solvation [78].

Chapter 7

The olfaction model

The previous chapters have given a bit of background and all that is necessary for us to describe biological inelastic electron tunneling in olfaction as a signaling transduction mechanism. I begin with a qualitative description of events to set the scene, then the mathematical model and the evaluation of rates and finally the parameters that are involved in order to arrive at these rates. The rate evaluations are required in order to find an average time τ_{T0} for non-discriminating ("elastic") tunneling or τ_{T1} for discriminating ("inelastic") tunneling. Only the inelastic contribution is sensitive to the odorant (M) oscillator frequency ω_0, and so needs to dominate the elastic contribution ($\tau_{T0} \gg \tau_{T1}$): this is contrary to typical IETS, where usually the inelastic channel is overwhelmed by the elastic channel. The crucial result is we find for biological IETS in olfaction that this situation is actually reversed.

7.1 The physical mechanism

7.1.1 The sequence of events

A smellable molecule 'M' upon inhalation is carried to a cavity behind the nose and in between the eyes and mouth. At this cavity M travels through a layer of mucus. This most likely involves collisions with water and proteins, as M progresses randomly and eventually meets the cell membrane of the olfactory neuron. Depending on its solubility M may spend some time in the lipid membrane. It is possible that odorant binding proteins facilitate the transition across the hydrophillic-hydrophobic layer. After some time M reaches the hydrophobic cavity and somehow

interacts with parts on some of the 7 transmembrane (TM) helice walls. In particular, when M reaches the ligand binding domain (LBD), something happens. The orientation of M is most likely to be changeable, however, at the LBD weak bonds will be made involving van der Waals and electrostatic interactions, possibly stabilizing an overall configuration. This induces an 'on' configuration of the receptor 'R' and M, that we call R+M. The assembly of this configuration involves two special points 'D' and 'A' -the donor and acceptor linked to a source and sink of electrons respectively. D and A interact with M, but not strongly. A time later, M experiences a sudden change in force caused by the electron jumping from the D to A site. After this kick, M then oscillates at a vibrational frequency corresponding to the D-A splitting, and the electron has crossed. The transfer to A triggers the release of the α-subunit of a neighboring G-protein (G), which initiates a large influx of Ca^{2+} ions into the cell; a signal that can be communicated by consequent firing of neurons to the brain. Hence the signalling machinery is initiated. See the cartoon 7.1 for a schematic of these initial events at the LBD.

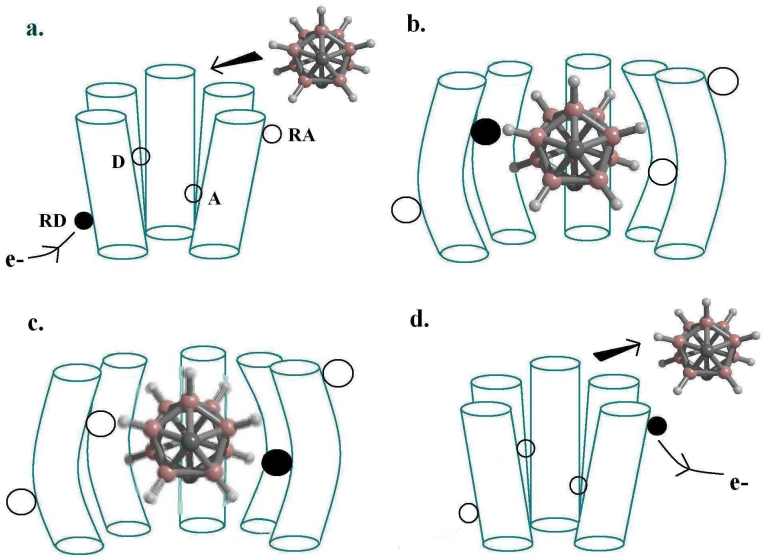


Figure 7.1: A cartoon for the proposal of biological IETS in the nose, only 5 TM helices for the OR are shown (cylinders) here for clarity. a) M approaches R, meanwhile an electron diffuses to position RD on a TM helix. b) M docks at the LBD, the overall configuration of R+M changes, meanwhile the electron *intra*-protein tunnels to D and it spends some time there. c) The electron jumps from D to A causing M to vibrate vigorously. d) M is expelled from the LBD and the electron *intra*-protein tunnels to RA. Signal transduction is initiated with the G-protein release.

7.1.2 A qualitative discussion

We will now examine more closely, the 'on' state of R+M and the important states involved. Another cartoon of the scene is described below, see figure 7.2, in a little bit more detail the crucial components and timescales.

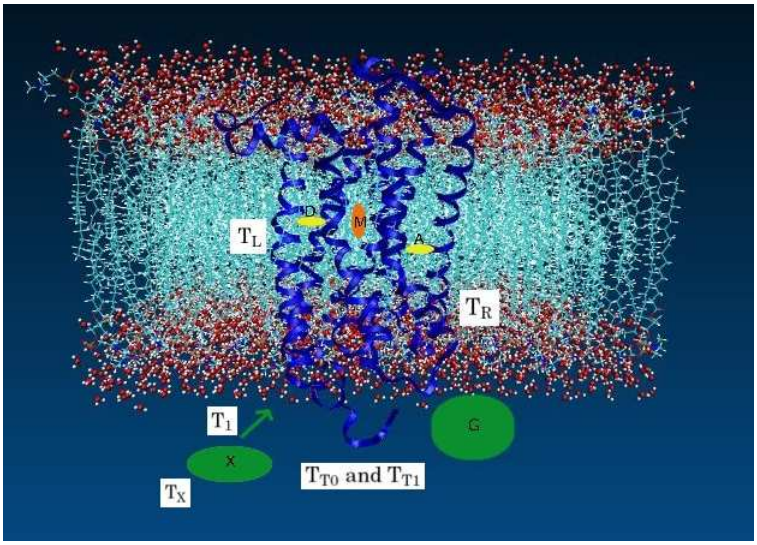


Figure 7.2: The olfactory receptor is a G-protein coupled receptor with seven hydrophobic transmembrane helices (TMH). This figure shows a cartoon for the OR protein (blue ribbon) embedded into a lipid membrane layer. The light blue chains show the hydrophobic lipid region and the red (oxygen) and white (hydrogen) atoms at the *intra* and *extra* cellular layers show water molecules. A charged particle is localized on the donor D in one TMH from where it hops to the acceptor A, these are indicated by the yellow ellipses. The odorant molecule, represented by the orange ellipse, docks somewhere within the hydrophobic domain. The approach of the reducing/oxidizing species is indicated by X in the green ellipse. The G-protein is indicated by the green circle. For more details on the T timescales featured in this figure see the text below, the most important times are: τ_{T0} for non-discriminating ("elastic") tunneling or τ_{T1} for discriminating ("inelastic") tunneling. This figure was produced using VMD software (see http://www.ks.uiuc.edu/Research/vmd/) to embed a conjectural olfactory receptor structure into a lipid layer.

Figure 7.2 shows the mucus layer, the cell membrane of the olfactory neuron, and one olfactory receptor containing an odorant. The odorant M with natural frequency[1] ω_0 is situated between D and A, and exchanges energy with the mobile particle enabling it to move between sites that differ in energy by $\hbar\omega_0$.

Further simplifying this picture, see figure 7.3 to represent the important en-

[1]In general an odorant will have multiple vibrational modes, but we consider only one here for simplicity. The particle is also able to exchange energy with other molecules in the surrounding environment. For now we will treat these as having a weak effect, and will reintroduce these interactions at a later stage.

ergy levels involved: D, M and A, and the electron transition. Examining these states, we assume that an electron on D and A is localized and distinct due to the distance between them, which we approximate as an average of ~8Å from the distances between helices. This is a good estimate if we believe the likelihood that these electron source/sinks are amino acids, see figure 3.5 in the Chapter 'A biological background', and also if we compare to distances for rhodopsin [41]. The M represents a state on the odorant the electron can pass through, and so is shown as the highest energy state. Its energy relative to D and A and can be estimated from the HOMO-LUMO gap for a typical odorant, which is of order 10's eV in contrast to the small D-A splitting which must correspond to a typical odorant fundamental mode of around 0.2 eV.

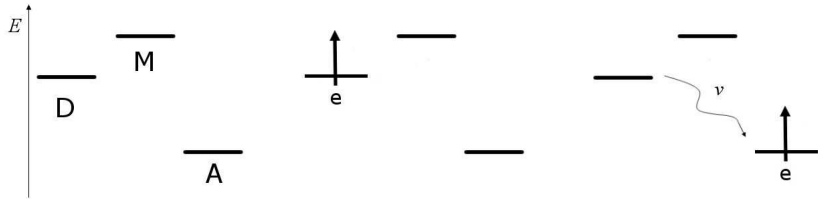


Figure 7.3: The important energy states D, M and A. Showing the transition of the electron from the donor and the acceptor, hopping with v between them, without spending any time on M. M facilitates the transition but does not host it for any length of time.

Absorption or emission spectroscopy?

We assume that the electron tunnels from a higher D state to a lower A state, thus exciting a phonon in M and giving the odorant energy, as opposed to the reverse process. This is determined because a calculation estimating the probability P_{ex} that M arrives in an excited state and loses energy, is given by:

$$P_{ex} = exp\left(-\frac{\hbar\omega_0}{k_B T}\right). \tag{7.1}$$

Taking a typical frequency $\omega_0 = 0.2eV$ and $k_B T = 0.027eV$ (normal human body temperature) the result is about 0.0006, and so very few odorants entering the receptor could produce a signal. It is unlikely at room temperature that odorants will be in an excited vibrational state, so it is more likely the case that the receptor works by giving energy to the odorant. Hence the positioning of energy states in

figure 7.3. A population inversion is unlikely. Hence electron transfer in olfaction involves phonon absorption (temperature independent) and not emission (temperature dependent).

Overall recognition time scale

We assume the important processes are as shown in figure 7.2, and we give values:

1. T_X is defined by the diffusion of a reducing (oxidising if hole transport) species. This is obtained from the Stokes-Einstein relation and gives $10\mu s - 1ms$.
2. T_1 can be determined by Marcus theory as this involves the charge from species X tunneling into the GPCR and this is estimated from typical rates to be $1ms - 1\mu s$, l for proteins [88].
3. T_L is defined by hopping transport. We assume charge injection on a timescale $1ms - 1\mu s$.
4. T_{T0} and T_{T1} are ns timescales.
5. T_R is described similarly to T_L.

Since the sense of smell involves a fairly modest timescale (tens of milliseconds overall), the actual probability per unit time for the inelastic channel can be very small and still consistent with experiment, the crucial result is, that it should be greater than the elastic channel. Thus, the inelastic transition must be faster then the elastic. More detail on these timescales is given below.

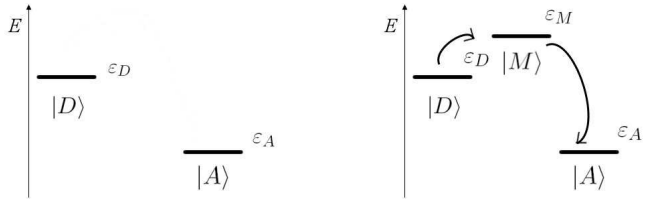
7.2 A mathematical model

Here I describe the states in our system and show how we arrive at inelastic and elastic rates that define the olfaction model using Fermi's golden rule.

7.2.1 The electron

The important electronic states are those where the electron is on the donor, represented by the notation $|D\rangle$, and when the electron is on the acceptor, this is represented by the notation $|A\rangle$. We make the assumption that the electronic wavefunctions evolve adiabatically, as electron motion is very rapid compared to the nuclear

motion during transitions. This is why we can use the configuration coordinate to represent the overall states of the system. However, the matrix element for the transition between them is non-adiabatic and this is the reason we can use Fermi's Golden rule to construct a rate.



(a) (b)

Figure 7.4: A figure to show a) the electronic states when there is a donor (D) and acceptor (A) but no odorant and b) when an odorant (M) is present. When an odorant is present, the electron is facilitated by travelling though an electronic state on M.

See figure 7.4 a). The Hamiltonian to describe these energetic states is:

$$H_e = \begin{pmatrix} \varepsilon_D & t \\ t & \varepsilon_A \end{pmatrix}, \tag{7.2}$$

where t is very small; the donor and acceptor are *weakly* coupled. If we introduce an odorant M into the equation, as in figure 7.4 b), then the electron can go from D to A via the molecule with state $|M\rangle$. The Hamiltonian for this scenario is:

$$H'_e = \begin{pmatrix} \varepsilon_D & v & 0 \\ v & \varepsilon_M & v \\ 0 & v & \varepsilon_A \end{pmatrix}. \tag{7.3}$$

In order to generalize to an effective two state Hamiltonian, we can use a technique called *downfolding*:

$$\begin{pmatrix} \varepsilon_D & v & 0 \\ v & \varepsilon_M & v \\ 0 & v & \varepsilon_A \end{pmatrix} \begin{pmatrix} c_D \\ c_M \\ c_A \end{pmatrix} = \varepsilon \begin{pmatrix} c_D \\ c_M \\ c_A \end{pmatrix}. \tag{7.4}$$

This yields the secular equations:

$$\begin{aligned} \varepsilon_D c_D + v c_M &= \varepsilon c_D \\ v c_D + \varepsilon_M c_M + v c_A &= \varepsilon c_M \\ v c_M + \varepsilon_A c_A &= \varepsilon c_A. \end{aligned} \tag{7.5}$$

So $c_M = (\varepsilon - \varepsilon_M) = v\,(c_A + c_D)$, and $c_M = \frac{v}{\varepsilon - \varepsilon_M}\,(c_A + c_D)$ which means that we can rewrite the 3×3 matrix as 2×2:

$$\begin{pmatrix} \varepsilon_D + \frac{v^2}{\varepsilon - \varepsilon_M} & \frac{v^2}{\varepsilon - \varepsilon_M} \\ \frac{v^2}{\varepsilon - \varepsilon_M} & \varepsilon_A + \frac{v^2}{\varepsilon - \varepsilon_M} \end{pmatrix} \begin{pmatrix} c_D \\ c_A \end{pmatrix} = \varepsilon \begin{pmatrix} c_D \\ c_A \end{pmatrix}. \tag{7.6}$$

This indicates a coupling $\frac{v^2}{\varepsilon - \varepsilon_M} = t \approx \frac{v^2}{\varepsilon_D - \varepsilon_M}$, from the off diagonal effective matrix elements. We approximate $\varepsilon = \varepsilon_D$ as we do not know the energy eigenstate ε, but assume $\varepsilon = \varepsilon_D$ or ε_A, since ε_D and ε_A differ by very little (meV), as compared to the difference between ε_D and ε_M (10's eV). Thus the initial electronic state will involve an admixture of D and M due to the presence of the odorant, and the final electronic state is similarly an admixture of A and M. This interaction drives the transition as we shall see below. This implies the presence of an odorant M is integral to an electron transfer process, so even if the protein R is sampling conformational space in microseconds it is unlikely that without the introduction of something like M that it will have the right mix of matrix elements to initiate a signal.

7.2.2 The odorant harmonic oscillator

We can use a configuration coordinate diagram and the harmonic approximation to represent the odorant[2], see figure 7.5. The Hamiltonian for the odorant harmonic oscillator is:

$$H_o = \frac{p_o^2}{2m_o} + \frac{1}{2} M_o \omega_o^2 u_o^2, \tag{7.7}$$

where p_o is the odorant momentum, M_o the odorant mass, u_o the odorant displacement and ω_o the odorant vibration. The last term is E_o, as in figure 7.5.

[2]These methods are described in Chapter 'A physical picture'

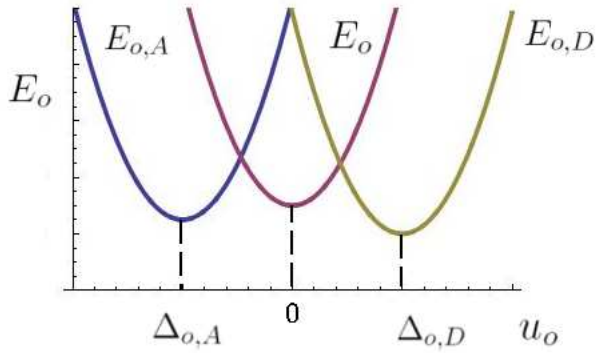


Figure 7.5: Odorant harmonic oscillator, showing three parabolas. The leftmost (blue) parabola indicates the forces on the odorant when the electron is on the donor, the middle (pink) parabola indicates the forces on the odorant when there is no electron, and the rightmost (yellow) parabola indicates the forces on the odorant when there is an electron on the acceptor.

The electron exerts a force on the odorant, $E_{o,D}$, when it is on the donor and a force, $E_{o,A}$, when it is on the acceptor. The Hamiltonian when the electron is on the donor is:

$$H_{o,D} = \frac{p_o^2}{2m_o} + \frac{1}{2}M_o\omega_o^2 u_o^2 - F_{o,D}u_o \tag{7.8}$$

$$= \frac{p_o^2}{2m_o} + \frac{1}{2}M_o\omega_o^2 \left(u_o - \Delta_{o,D}\right)^2 - \frac{1}{2}M_o\omega_o^2\Delta_{o,D}^2. \tag{7.9}$$

The last two terms make $E_{o,D}$. The Hamiltonian when the electron is on the acceptor is:

$$H_{o,A} = \frac{p_o^2}{2m_o} + \frac{1}{2}M_o\omega_o^2 u_o^2 - F_{o,A}u_o \tag{7.10}$$

$$= \frac{p_o^2}{2m_o} + \frac{1}{2}M_o\omega_o^2 \left(u_o - \Delta_{o,A}\right)^2 - \frac{1}{2}M_o\omega_o^2\Delta_{o,A}^2. \tag{7.11}$$

The last two terms make $E_{o,A}$. The forces are $F_{o,D} = M_o\omega_o^2\Delta_{o,D}$ and $F_{o,A} = M_o\omega_o^2\Delta_{o,A}$. Δ is the shift of the parabolas, see figure 7.5; the force on the odorant forces the parabola down in energy and along u_o, displacing the atoms.

7.2.3 Environment oscillators

There is not just the *odorant* reponse to the moving charge; but we must also consider all other oscillations in the environment. We treat all the environment oscillators, again harmonically as above, but generalized as one. [3] The Hamiltonian for oscillations in the environment (E) is:

$$H_{E,D} = \frac{p_E^2}{2m_E} + \frac{1}{2}M_E\omega_E^2u_E^2 - F_{E,D}u_E \tag{7.12}$$

$$= \frac{p_E^2}{2m_E} + \frac{1}{2}M_E\omega_E^2\left(u_E - \Delta_{E,D}\right)^2 - \frac{1}{2}M_E\omega_E^2\Delta_{E,D}^2 \tag{7.13}$$

and similarly for A.

Here we make a key assumption that $\omega_E << \omega_o$, *i.e.*, we assume low frequencies of oscillation in the environment, such as the 'soft' floppy protein backbone fluctuations observed in most proteins. It must be that the odorant frequency is significantly higher than all the environmental oscillating frequencies. For discrimination in this model it is necessary that these other modes are weakly coupled to the electron transfer and there is negligible probability of their excitation. The scenario can be pictured as a 2 dimensional configuration coordinate diagram, with the reaction coordinate as *all* displacements in the environment. A shallower 2d-surface parabola would represent the whole environment and the odorant frequency would stand out as a steeper parabola on this surface. The use of configuration coordinate diagrams in this way, allow us to see the emergence of a relaxation energy as in general charge transfer. The tunneling event from D to A imparts a change in force on the modes of the odorant molecule, and this is exactly like the change in force due to an electronic state change in Huang-Rhys theory, see Chapter 'Huang-Rhys factors'. The ratio of $S = \frac{\lambda}{\hbar\omega_0}$ corresponds to the dimensionless 'Huang-Rhys' factor which has been outlined before and will be explored in detail in the Chapter 'Huang-Rhys factors'.

[3] Though the principle can be extended to many oscillators.

7.2.4 For the complete interacting system

We construct the Hamiltonian for everything:

$$H = \begin{pmatrix} H_D & H_{DA} \\ H_{DA} & H_A \end{pmatrix}, \tag{7.14}$$

where the Hamiltonians we have found are:

$$\begin{aligned} H_D = \varepsilon_D - \frac{1}{2}M_o\omega_o^2\Delta_{o,D}^2 - \frac{1}{2}M_E\omega_E^2\Delta_{E,D}^2 - t \\ + \frac{p_o^2}{2m_o} + \frac{1}{2}M_o\omega_o^2\left(u_o - \Delta_{o,D}\right)^2 \\ + \frac{p_E^2}{2m_E} + \frac{1}{2}M_E\omega_E^2\left(u_E - \Delta_{E,D}\right),^2 \end{aligned} \tag{7.15}$$

which is:

$$H_D = \varepsilon_D + H'_{o,D} + H'_{E,D}, \tag{7.16}$$

for the donor, and similarly for A. $H_{DA} = t$.

7.2.5 Unperturbed eigenstates

Now we can determine the eigenstates required for solving Fermi's golden rule. Examining the Hamiltonians:

$$H'_{o,D}\left|\chi_{o,D,n}\right\rangle = \left(n + \frac{1}{2}\right)\hbar\omega_o\left|\chi_{o,D,n}\right\rangle, \tag{7.17}$$

$$H'_{o,A}\left|\chi_{o,A,n}\right\rangle = \left(n + \frac{1}{2}\right)\hbar\omega_o\left|\chi_{o,A,n}\right\rangle, \tag{7.18}$$

$$H'_{E,D}\left|\chi_{E,D,N}\right\rangle = \left(N + \frac{1}{2}\right)\hbar\omega_E\left|\chi_{E,D,N}\right\rangle, \tag{7.19}$$

$$H'_{E,A}\left|\chi_{E,A,N}\right\rangle = \left(N + \frac{1}{2}\right)\hbar\omega_E\left|\chi_{E,A,N}\right\rangle, \tag{7.20}$$

where $|\chi_{o,D,n}\rangle$, for example, is a total wavefunction given by a product of a set of vibrationally independent modes:

$$\chi_{o,D,\{n\}}(Q) = \prod_{\alpha} \zeta(n_\alpha; Q_{D\alpha}), \tag{7.21}$$

[56], where α denotes modes; $Q_{X\alpha} = Q - \bar{Q}_{X\alpha}$, n_α is a mode and the set of occupation numbers is given by $\{n\}$. All four states are described in similar ways; $\{n'\}$ being the set of final oscillating states on the odorant and $N = \{n_q\}$ and $N' = \{n'_q\}$ being the set of oscillating states in the environment, initially and finally, respectively. Then total wavefunctions and eigenstates can be written:

$$\begin{aligned} H_D \left|D, \chi_{o,D,n}, \chi_{E,D,N}\right\rangle &= \left(\varepsilon_D + \left(n + \frac{1}{2}\right)\hbar\omega_o + \left(N + \frac{1}{2}\right)\hbar\omega_E\right)\left|D, \chi_{o,D,n}, \chi_{E,D,N}\right\rangle \\ &= U_{DnN}\left|D, \chi_{o,D,n}, \chi_{E,D,N}\right\rangle \end{aligned} \tag{7.22}$$

$$\begin{aligned} H_A \left|A, \chi_{o,A,n'}, \chi_{E,A,N'}\right\rangle &= \left(\varepsilon_A + \left(n' + \frac{1}{2}\right)\hbar\omega_o + \left(N' + \frac{1}{2}\right)\hbar\omega_E\right)\left|A, \chi_{o,A,n'}, \chi_{E,A,N'}\right\rangle \\ &= U_{An'N'}\left|D, \chi_{o,A,n'}, \chi_{E,A,N'}\right\rangle. \end{aligned} \tag{7.23}$$

The total wavefunctions are the products of the electronic wavefunction, the odorant vibrational wavefunction and the environmental vibrational wavefunction for D or A.

7.2.6 Fermi's golden rule

If the starting point involves the electron in state $|D\rangle$, and the odorant in state $|\chi_{o,D,0}\rangle$, $n = 0$, and the finishing point is with the electron in state $|A\rangle$ and the odorant in state $|\chi_{o,A,n}\rangle$, we construct Fermi's golden rule for a transition between these two states, for detection in olfaction.

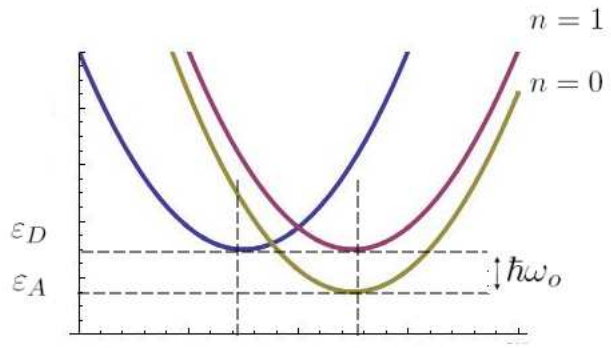


Figure 7.6: A configuration coordinate diagram to show the inelastic $(n = 1)$ versus the elastic $(n = 0)$ route.

The key oscillation in the odorant is characterized by n and corresponds to a mode of the odorant Q_0. We conjecture the important transitions correspond to i) *elastic* tunneling where $n = 0$ and no phonon is excited and there is no occupation of mode Q_0 and ii) *inelastic* tunnelling where $n = 1$ and one phonon is excited and there is an occupation of mode Q_0 in the odorant. So in the construction of these two rates we have in the initial state $n = 0$ and in the final state $n = 0$ or 1, for the odorant vibrational state, depending on the rate we evaluate, see figure 7.6. Thus we must treat odorant oscillations explicitly in this way although all other oscillations can include others of the odorant and of the receptor.

$$\frac{1}{\tau_{D,0\to A,n}} = \frac{2\pi}{\hbar}\sum_{NN'} P_N \left|\langle D\chi_{oD0}\chi_{EDN}\left|H\right|A\chi_{oAn}\chi_{EAN'}\rangle\right|^2 \delta\left(U_{D0N} - U_{AnN'}\right). \quad (7.24)$$

The kind of quantum mechanical analysis made by Huang and Rhys for the absorption of light in F-centres of crystals [63] can be applied to this problem in olfaction. The coupling of a single vibrational frequency to the electronic excitation is caused by the absorption of a photon, and can be compared to the electron transfer coupled

with a single vibrational frequency in an odorant, except we have a contribution from many more states in the environment.

In using the Born-Oppenheimer approximation we can examine the *environments response (*refered to as *lattice* response in Huang-Rhys theory*)* separately to the electronic part, as the vibrational energy is independent of the electronic state, thus:

$$= \frac{2\pi}{\hbar}\sum_{NN'} P_N \left|H_{DA}.\left\langle \chi_{oD0}\left|\chi_{oAn}\right.\right\rangle.\left\langle \chi_{EDN}\left|\chi_{EAN'}\right.\right\rangle\right|^2 \delta\left(\varepsilon_D - \varepsilon_A - n\hbar\omega_o + \left(N-N'\right)\hbar\omega_E\right), \tag{7.25}$$

we can separate these parts. Averaging over the initial states of the environment and a sum over the final states of the environment we obtain:

$$= \frac{2\pi}{\hbar}\left|t\right|^2\left|\left\langle \chi_{oD0}\left|\chi_{oAn}\right.\right\rangle\right|^2 \sum_{NN'} \frac{e^{-N\hbar\omega_E/k_BT}}{Z_E}\left|\left\langle \chi_{EDN}\left|\chi_{EAN'}\right.\right\rangle\right|^2 \delta\left(\varepsilon_D - \varepsilon_A - n\hbar\omega_o + \left(N-N'\right)\hbar\omega_E\right), \tag{7.26}$$

where we use $Z_E = \sum\limits_N e^{-N\hbar\omega_E/k_BT}$. We use a technique that makes use of the Franck-Condon factor mentioned previously; which describes the overlap of initial and final vibrational states. For a complete rate of reaction, we must appreciate the likelihood the system occupies the *right* initial state and that there is an appropriate final state to get to. This is done by multiplying the 'Franck-Condon' factor by the fraction of instances when the system is in the initial state, at the given temperature, and then summing these products over the final states. The Franck-Condon factors are thus corrected for by Boltzmann probabilities. A way to refer to this part of the rate is as a *response component*, is traditionally, as a $line-shapefunction\ G(\varepsilon)$ [56]:

$$G(\varepsilon) = \underset{\{n\}}{Av}\sum_{\{n'\}}\left|\int dQ\chi^*_{o,D\{n\}}(Q)\chi_{o,A\{n'\}}(Q)\right|^2 \delta\left(\varepsilon - \varepsilon_{DA,nn'}\right) \tag{7.27}$$

We see the stronger the overlap of the nuclear wavefunction before and after the electron transfers then the faster the transition. This line-shape function can be

written in terms of the Huang-Rhys factor:

$$G(\varepsilon) = \sum_{n=0}^{\infty} \left(\frac{e^{-S} S^n}{n!} \right) \delta \left(\varepsilon_D - \varepsilon_A - n\hbar\omega_0 \right), \tag{7.28}$$

for the low temperature limit; for the odorant it is the case that $\hbar\omega_0 >> k_B T$. This is derived by Huang and Rhys for the single frequency and strong coupling case [56], however, the same result can be found for the case here: many frequencies and the weak coupling case. For thorough references to the algebra involved in this result see Chapter 10 in 'Defects in solids' [56]. Note the emergence of the aforementioned Huang-Rhys factor $S = \frac{\lambda}{\hbar\omega_0}$ for the single frequency of the odorant. We assume S<<1. This weak coupling limit allows an immediate indication of what proportion of zero phonon, one phonon, two phonon, *etc.,* transitions will arise, though we assume two-phonon excitation or beyond does not occur. The equation 7.28 gives a measure of how readily the lattice vibrations take up the energy and that the delta function selects instances where $\varepsilon_D - \varepsilon_A = n\hbar\omega_0$. So electron transfer only occurs when there is conservation of energy and this is true when there is a *driving force* (see Chapter 'A physical picture') given by $\varepsilon_n = \varepsilon_D - \varepsilon_A - n\hbar\omega_0$ for two instances. These two possibilities of electron transfer at ε_0 and ε_1 are our elastic and inelastic pathways, for $n = 0$ or $n = 1$, respectively.

We now have a standard Marcus problem, with $\sigma_n = \frac{e^{-S} S^n}{n!}$; we have a rate factorized into an electronic matrix component t, and an environment response element σ_n (which relates to the Huang-Rhys factor in an odorant). The last step is to fully account for the rest of the rest of the world:

$$\begin{aligned} \frac{1}{\tau_{D,0\to A,n}} = \frac{2\pi}{\hbar} \sigma_n \left| t \right|^2 \sum_{NN'} \frac{e^{-N\hbar\omega_E / k_B T}}{Z_E} \left| \langle \chi_{EDN} \left| \chi_{EAN'} \right\rangle \right|^2 \\ \delta \left(\varepsilon_D - \varepsilon_A - n\hbar\omega_o + \left(N - N' \right) \hbar\omega_E \right) \end{aligned} \tag{7.29}$$

We do this by making another key assumption: that we can use the harmonic approximation to couple the system to a *bath of modes* [89]. Referring to configuration coordinate diagrams, the two states D and A entail a reaction coordinate that corresponds to the reaction pathway. However, there are nuclear degrees of freedom that do not belong to the reaction coordinate. We account for this by coupling the system to a bath of modes that represent the environment: the q oscillators.

We extend from one oscillator to *many* in the environment. The real full quantum expression for the rate then gives:

$$\frac{1}{\tau_{D,0\rightarrow A,n}} = \frac{\beta}{\hbar} |t|^2 \sigma e^{-\beta\varepsilon_n/2} \int_{-\infty}^{\infty} d\xi \, exp(-i\beta\varepsilon_n\xi - f(\xi)), \tag{7.30}$$

where:

$$f(\xi) = \frac{2}{\pi\hbar} \int_0^{\infty} \frac{J(\omega)}{\omega} \frac{cosh(\beta\hbar\omega/2) - cosh(i\beta\hbar\omega\xi)}{sinh(\beta\hbar\omega/2)} \, d\omega. \tag{7.31}$$

We use the $\sigma_n = \frac{e^{-S}S^n}{n!}$, $\beta = 1/k_BT$, the condition $\varepsilon = \varepsilon_D - \varepsilon_A - \hbar\omega_0$, the Huang-Rhys factors S and S_q and the electronic matrix element is t as above. This factor is important, indeed without coupling to this thermal bath the rate would not exist. $J(\omega)$ is called the *spectral density* of this environment and it is related to the reorganization energy of the environment by:

$$J(\omega) = \sum_q \pi\delta(\omega - \omega_q) S_q \frac{\hbar\omega_q^2}{2}, \tag{7.32}$$

$$\lambda = \frac{2}{\pi} \int_0^{\infty} \left(\frac{J(\omega)}{\omega} \right) d\omega. \tag{7.33}$$

where q denotes oscillations in the environment. This part of the rate accounts for energy broadening due to fluctuations in the environment, where $\lambda = \Sigma_q S_q \hbar\omega_q$ is the reorganization energy, as appeared in the above discussion. To minimize this broadening and for the receptor to be discerning we require that the reorganization energy be small, which indicates a small coupling between oscillations in the environment and the itinerant electron, see section below. Since $f(\xi)$, in equation 7.30, has a deep minimum about $\xi = 0$, values near the minimum dominate in the exponential and so the harmonic approximation remains valid. As noted, for the odorant it is the case that $\hbar\omega_0 >> k_BT$, however, for the other oscillations in the environment we assume low background oscillations within the classical limit, $\hbar\omega_q << k_BT$. Given this, we approximate to $\beta\hbar\omega \ll 1$, and thus approach the classical limit, so we recover an expression from traditional Marcus theory;

$$environment\, response\, component = \frac{\sigma_n}{\sqrt{4\pi k_B T\lambda}} exp\left(-\frac{(\varepsilon_n - \lambda)^2}{4k_BT\lambda} \right), \tag{7.34}$$

and so obtain the crucial result below, from using Huang and Rhys's methods.

7.3 The crucial result

Thus, we find the rate equation for olfaction, using Fermi's golden rule, to be:

$$\frac{1}{\tau_{D,0\rightarrow A,n}} = \frac{2\pi}{\hbar}\left|t\right|^2 \frac{\sigma_n}{\sqrt{4\pi k_B T\lambda}} exp\left(-\frac{\left(\varepsilon_n-\lambda\right)^2}{4k_B T\lambda}\right). \tag{7.35}$$

The important parameters inserted are described in the table 7.1. Using these values we get $\tau_{T0} \sim 87ns$ and $\tau_{T1} \sim 0.15ns$ which satisfies the condition that $\tau_{T1} << \tau_{T0}$. The inelastic tunneling rate dominates the elastic rate, which is contrary to the opposite (almost universal) case.

Quantity	$\hbar\omega_0$	S	λ	$\left\|t\right\|$
Value	200 meV	0.1	30 meV	1 meV

Table 7.1: A table to show estimated values for the parameters needed to compute τ_{T0} and τ_{T1} [65].

Let us briefly analyze some features of this rate, using the table of parameters. First, figure 7.7, is simply a plot of the two channels: inelastic versus elastic for time versus reorganization energy λ to find the point at which the relationship is inverted. This happens at $\sim 62meV$, showing that the rate is very sensitive to this parameter, below $62meV$ discrimination is possible via inelastic tunneling, above this amount, the elastic rate rules.

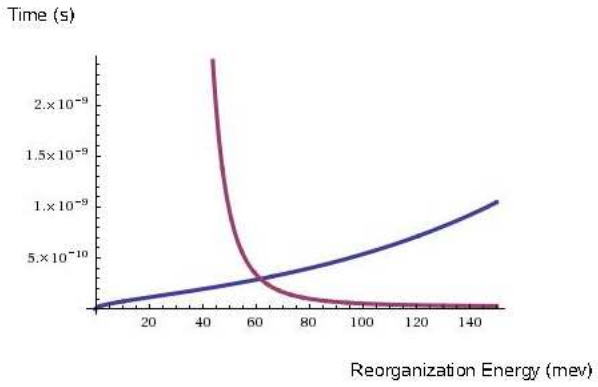


Figure 7.7: A plot to show the time (s) for an inelastic transmission (blue) versus the elastic transmission (red) all parameters given in table are kept the same, the only variable is the reorganization energy.

Secondly we examine figure 7.8. This plots the rate for each channel keeping the other parameters the same but varying the energy splitting $\delta = \varepsilon_D - \varepsilon_A$. Note maximum rates are obtained when the $\varepsilon_n = \lambda$, because of the Gaussian function, this is then an 'activationless' reaction [61] . This indicates that at very small splittings that match the reorganization energy, the receptor would be saturated with an elastic signal, and the criterion for discrimination is inverted ($\tau_{T1} >> \tau_{T0}$). The splittings must not get this small, which is unlikely anyway as the range of molecular vibrations is more like 70-400meV (~500-3300cm^{-1}), where below 70meV is lost to thermal noise. In fact the evaluation of this rate indicates energy splittings below 110meV are likely to be out of the discriminating region. For the elastic channel to be at its maximum, such an event however, would require that the reorganization energy, being small, is matched by $\hbar\omega_0$, and this is unlikely for the odorant vibrations of interest. For low reorganization energies and comparably high vibrational modes, we are safely within the discriminating region ($\tau_{T0} >> \tau_{T1}$). This plot highlights, as does figure 7.8, the sensitivity on reorganization energy: we really require λ much closer to 0meV than 200meV.

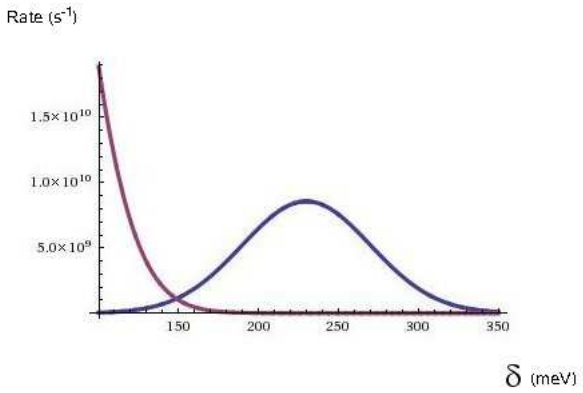


Figure 7.8: A plot to the show the rate (s^{-1}) for an inelastic transmission (blue) versus the elastic transmission (red) all parameters given in table are kept the same, the variable is the energy splitting.

Finally, see plot 7.9, we show the same as above, except just for the inelastic channel, and we see how temperature can effect the rate, for 4K (green), 77K (yellow), 295K (red) and 310K (blue). We see an expected decrease in the full width half maximum (FWHM) with decrease in temperature, improving resolution. In olfaction, at temperatures in the body of ~310K, we see a FWHM of approximately 100meV ($2\sqrt{2ln2}.\sqrt{2\lambda k_B T}$), thus above the range 110meV, an odorant can activate a given receptor type within the range $\pm 50meV$, of the types $\varepsilon_D - \varepsilon_A$ tuning. Thus, both the average amount of energy transferred and the mean square variance of the energy transferred (which gives a broadening) can be related to the corresponding reorganisation energy (we can separate this into contributions from individual modes in all simple cases).

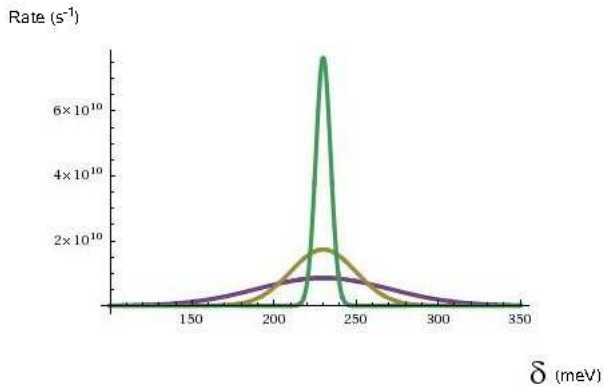


Figure 7.9: A plot to the show a rate (s^{-1}), for the inelastic channel, all parameters given in table are kept the same, the variable is the energy splitting. We have for four temperatures, green is 4K, yellow is 77K, red is 295K and blue is 310K.

Of the parameters in the table $\hbar\omega_0$ and S can be calculated directly and accurately for odorants, and these parameters are discussed in greater detail elsewhere and a whole investigation is devoted to the Huang-Rhys factor S. The other facets of the system that play important roles in this crucial result are discussed below.

7.4 The parameters

By now it is obvious that realizing the atomistic models for a donor and acceptor would be very helpful, however, positions of atoms to 0.1-0.2Å are needed, which is well beyond the best current data which can resolve at best to 2Å, and anyway this is an order of magnitude too inaccurate for the determination of tunneling rates because at this distance there are no interactions. So we proceed by using a model receptor in which the key parameters are estimated from the ranges of known biomolecular systems.

7.4.1 The relaxation (reorganization energy)

This model is extremely sensitive to the reorganization energy of the environment λ. At and below 30 meV the elastic channel is suppressed relative to the inelastic, but an increase of this energy to just 62meV inverts this relationship. Here we must evaluate what the reorganization energy is, and consider whether we can assume that it is so small. In our model when an electron makes a transition, some of the system, like the other electrons, respond essentially almost instantaneously, and some other parts of the system, like the changes in positions of atoms, respond more slowly. This slower part of the system relaxes to equilibrium under the new forces, and there's a relaxation energy associated with this which includes the work done by the new forces and the energy needed to stretch or compress bonds. When the electron transition occurs, some energy is given to every degree of freedom that couples to the transition so the reorganisation energy here refers to all those other modes, not just the key vibration of the odorant. If the reorganization energy were too large, there would be broadening, see above, and so less discrimination and perhaps confusing sidebands to the transition. So simply; the reorganization energy represents the changes to the nuclear positions corresponding to the charge movement.

We assume there is something about the domain environment which may promote electron transfer in the presence of an odorant. In other words the reorganization energy takes into account the energy required when the environment dipoles electrostatically prefer the electron on the donor to where the environment dipoles electrostatically prefer it on the acceptor. So we may take into account the polarizing property of a traveling electron such that the dielectric is included. Altogether the reorganization energy corresponds to a contribution from what Marcus calls *inner shell atoms* (λ_i) and the *outer atoms* (λ_o), respectively. The first relates to the inner shell vibrational modes approximated by:

$$\lambda_i = \frac{1}{2}\Sigma_\alpha k Q_\alpha^2 \tag{7.36}$$

where α sums over modes. This equation is represented by the parabolas, and has of course arisen in previous sections, where we assume a classical limit, and small characteristic frequencies compared to $k_B T$, that are weakly coupled. The outer shell element λ_o refers to atoms further out, part of the 'solvent' and is estimated

by the polarizability of the environment:

$$\lambda_o = \frac{(\Delta e)}{4\pi\varepsilon_o}\left(\frac{1}{2r_1}+\frac{1}{2r_2}+\frac{1}{r_{12}}\right)\left(\frac{1}{D_{op}}-\frac{1}{D_S}\right) \tag{7.37}$$

where r_1 and r_2 are the radii of the two reactants when they are in contact and D_{op} is the square of the refractive index (the fast response dielectric constant) and D_S the static dielectric constant (the slow response dielectic constant) and Δe is the charge that is transferred. Very unpolar environments, such as the hydrophobic LBD, are little effected by charge transfer such that there is small rearrangement of nuclei- this leads us to believe the reorganization energy for the olfactory system will be small. Further, electron transfer over small distances compared to the redox centres correspond to small rearrangements in the environment. We compare the olfactory environment to the photosynthetic bacteria *Rhodobacter capsulatus* which has values of λ below 0.03 eV at room temperature. If the receptor had been hydrophilic, and the system behaved as in water, we might have had a reorganisation energy of say 1 eV: these energies are estimated from the redox potentials of ions in water [90]. Calculations for the relaxation energy in the aqueous case are well described, see for example, Sit *et al* [64]. For the electron transfer in a ferrous-ferric system in water; the introduction gives a good description of Marcus theory, but water is highly polarisable, which is why this energy is large. So we *desire* a small reorganization energy for discrimination against noise, but we can also estimate one reasonably in consideration of equations 7.36 and 7.37 in terms of parallels with similar systems.

7.4.2 The donor and acceptor

For the model to work we have important requirements of the D and A:

1. D and A must be sharp levels and interact only weakly with the oscillator via the perturbation $\hat{v}$, and weakly to each other by hopping integral t.

2. D and A must be able to return back to their original (charged) states repeatedly and elastically.

3. To detect odorants within ms, though tunneling via an odorant can be much faster, the replenishment of D and A could be within ms but not longer.

4. There must be a small energy splitting $\varepsilon_D - \varepsilon_A$ that corresponds to the small $\hbar\omega_0$.

Let us now consider the above.

Transport issues

The electron may spend a comparatively long time at, or getting to (exiting from) the states $|D\rangle$ and $|A\rangle$, respectively. This can be conceived if the D and A states are reached via a staircase of states differing in energy. We assume that these states are well defined for each receptor type, and they are reached by slowish incoherent hopping of the electron. We find that overall charge injection at site D and extraction at site A can work overall between $ms - \mu s$, described by the components in figure 7.2. This indicates that no matter how fast a tunneling event may be between D and A, the re-population of D and re-emptying of A puts a bound on how many tunneling events can occur in one receptor. We must assume that the reverse process of charge transfer is frustrated, otherwise the use of transition rate theory is not justified. Further, only in a simple potential barrier system is it likely that the charge will go back and forth resonantly; this depends on the distance and relative energies of the walls. It is more likely that there is a range of donor/acceptor level barriers and wells that break up this resonant process. Plus the relaxation of the excited odorant would block this reverse process.

For timescales on transport, if we estimate a simple 3D diffusion rate, we produce a value within the range $1\mu s - 1ms$, if we assume simply, $\tau \sim \frac{\delta^2}{D}$, no bias, where D is the diffusion constant for 3D diffusion. This is based on a value given for water, $D \sim 10^{-4} cm^2/s$, which corresponds to a jump of 1 molecular radius each lattice vibration period. If we estimate δ as the distance the charge must move to the donor; we estimate $100nm = 0.1\mu m$. Increasing δ to $1\mu m$ and the diffusion rate still falls within a reasonable region ($< ms$). Even if we consider a slightly more detailed model: where charge motion is still diffusion, but it is now biased by a field $E = V/d$, the rate still falls in the right region. This is perhaps more realistic, when we consider the local environment; a voltage V driven by some chemical reaction, that push the electrons in the right direction. d is the distance the charge must move from where the excess potential is established to the donor. The electron mobility can be defined by μ as it moves along a type of 1-dimensional wire, this is the

drift velocity per unit applied field and it is related to a diffusion constant D via the Nernst-Einstein relationship:

$$\mu = \left(\frac{q}{k_B T}\right) D, \tag{7.38}$$

showing a temperature dependence, and q is the charge [91]. The relation for the mobility to the diffusion constant is $\frac{\mu(cm^2/V\,sec)}{D(cm^2/sec)} = 1.14 \times 10^4 T^{\text{-}1} q$ [91]. The mobility is:

$$\mu = \frac{u}{(V/d)} \tag{7.39}$$

and u is the drift velocity. From this we can estimate a rate (in s^{-1}) with: $k = 1.14 \times 10^4 qD/T.V/d^2$, if we take $q = 1$ and $T = 310K$, then simply $k = 37DV/d^2$. The potential V can be estimated as $0.5V$ (bigger than the odorant quanta of vibration) and $D \sim 10^{\text{-}4} cm^2/s$, which is true for most liquids for 3D diffusion. We take d from before as $d \sim \delta \sim 100nm \sim 1 \times 10^{-5} cm$, so that we would expect a rate $\sim 18.5 \times 10^{-6} s^{-1}$, a time of $\sim 0.5\mu s$. Here we adopt typical convention, *i.e.*, that where $\mu > 1$ and the mean free path exceeds nearest neighbour distance, then there is conduction and when $\mu < 1$, if the jumps do not exceed one lattice vibration period, then there is electron jumping. This means, if the biological 'wire' for electron transfer is a conductor and $\mu > 1$, then electron transport to the donor is fast. If there is hopping even if d is increased to $1\mu m$ then the time is still less than ms, and decreasing the distance d, if there is a very small mobility, the rate is still easily within this limit. We conclude then that diffusion of the charge to the GPCR receptor, occurs within the range $10\mu s - 1ms$ for the value T_X.

We now consider the limits on getting the charge from the reducing species to the localized donor state, via the protein for T_1 and T_L. These values can be estimated using known values for protein charge transfer and using what we know about conduction in DNA [88]. The process is likely hopping transport given by [92]:

$$T = (J^2/\hbar)(\pi/4\varepsilon k_B T)^{1/2} exp(-2\alpha R)\, exp\left(-\frac{\varepsilon}{k_B T}\right) \tag{7.40}$$

We have assumed there are similar states, to and from, between which the charge hops, and that they are separated by a distance R. J is the transfer matrix between these two states, *i.e*, between bases in DNA. We take J to be $\sim 0.1eV$- a

half-bandwidth as determined by:

$$2B = 2NJ = 2\hbar^2/m^* a^2, \tag{7.41}$$

where $m^*$is the effective mass, N is nearest neighbours at distance a and B is the bandwidth [91]. For DNA we estimate an activation energy ε of 0.5eV, and α is the decay constant ~ 1.1Å, typical for proteins [88]. The distance between bases in DNA is $R \sim 3.4$Å. Using these parameters, and $T = 310K$ we obtain a time $\sim 175\mu s$. We conclude then within slight variations of the hopping distance and activation energy that the hopping process within the protein occurs within $1ms - 1\mu s$. Thus transport to D is not rate limiting. It seems acceptable to assume that the electron resides at D (and similarly A) for a period of μs which corresponds to many atomic vibrations.

The small energy difference

For detection a very small splitting of $\varepsilon_D - \varepsilon_A$ corresponding to an odorant's vibration is required at resonance. However, it is common that most band gaps in biomolecular units (HOMO-LUMO splittings) are 10's of eV not meV. We imagine that likely donor/acceptor candidates are amino acid residues, perhaps of the same unit such as His, and so electron transfer across one unit to the other from LUMO to HOMO is ruled out because of the typically large energy difference. What seems more likely in order to achieve this necessary criterion; is transfer across HOMO-HOMO or LUMO-LUMO, in which case ionization potentials and electron affinities (respectively) must be carefully found. In such cases we can imagine a small splitting thus. If D and A are largely the same, for example 2 tryptophan residues, and can be represented by a state $|x\rangle$ at site D or A, one may differ by something small: such as a charge $q\,|e|$, this will be asymmetrically placed with regard to D or A. For example if the charge is 5Å from D and 4Å from A the resulting energy separation$\frac{e^2}{\varepsilon R_A} - \frac{e^2}{\varepsilon R_D}$ would be $\sim 0.72/\varepsilon$ eV for $\varepsilon = 3$ which corresponds to 1935cm^{-1}. It is intuitive to think there may be staple units for the D and A that define the receptor types but that the energy states are subtly modified by surrounding residues to get a fine tuning according to different phonons. These ideas however, must be verified with calculation and experiment to test for the likely stability and fidelity one electron would have to a particular D/A unit.

Possible candidates: zinc and tryptophan

There are two important observations about zinc in olfaction, made by Turin in 1996, i) that anosmia occurs with lack of zinc, and is reversed by dietary supplements and ii) those odorants that bind strongly to zinc, also smell very intensely. This suggests immediately that zinc is important certainly to the overall healthy configuration and effective folding of the OR protein, but possibly and maybe more interestingly, that it has an effect on the binding and so perception of the ligand. Figure 7.10 shows a a putative OR structure from Leffingwell's data [42], highlighted in yellow haloes are the amino acids glutamine (Glu) and histidine (His). These residues coordinate zinc. Typical ligands for zinc are water, cysteine, histidine, and glutamic aspartic acids [93]. Well within the domain there is a histidine residue in the middle of the 6^{th} helix and a glutamine is opposite on the 3^{rd}. For all of Leffingwell's receptors His exists in the middle of the 6th helix and on several residues on the loops. Thus, in light of the above it is possible that the small splitting is achieved by charges on residues or a transition metal near D and/or A such as zinc (or even copper). Determination of any transition metal near or at the binding site, would be incredibly useful in modelling because not only may the position of charge define various D-A splittings, but also to consider:

1. In the presence of a transition metal there are complexation energies that would influence the odorant. For $Zn^{2+}$and Cu^{2+}there can be planar or octahedrally coordinated ligands, for Cu^{+}, tetrahedrally coordinated ligands. The D-A splitting may then be *induced* by the tethering of the odorant to the donor, so binding occurs at the same site as the donor and may even be essential to turn on the receptor. In fact this sort of 'tethering' would affect the hopping integral (it would vary linearly with nuclear displacement) such as to increase IETS rate by 10% and leave the ETS rate unaltered.

2. The presence of a transition metal may also affect the electron-phonon coupling. This may occur via a pseudo-Jahn Teller like effect that couples with fluxional molecules.

3. The inclusion of a transition metal would alter the reorganization energy. Usually cofactors in the environment contribute to small reorganization energies [61]. The presence of a stabilizing metal ion such as zinc or copper, may by

enforcing a geometry, reduce the reorganization energy to the small amount required, such as in 'Blue Cu' centres in plastosyanin (a protein that transfers electrons in plant chloroplasts) where very fast electron transfer occurs with a very small driving force. It has been found that embedding a redox center inside a low dielectric cavity can lower the outer sphere λ by 50% [88]. Further, constraining the coordination environment around zinc or copper could reduce the inner sphere λ.

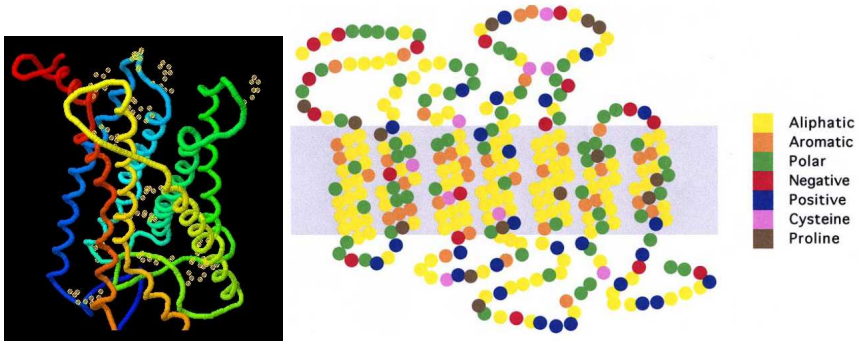


Figure 7.10: On the left, a putative olfactory receptor [42], Glu and His amino acids are indicated by the yellow haloes. Next to it on the right, is a sequence analysis from Fuchs *et al* [38]. Membrane is the shaded box. The amino acid groups are coloured and grouped according to their type. Note the cysteine residues highlighted in pink.

Tryptophan, deserves a mention here also, because of its constant appearance in the olfactory receptor repertoire, see figure 3.5 in section 'A biological background'. The sequence homology data by Fuchs *et al* show a clearly conserved tryptophan on TM4. Further this region is surrounded by highly variable neighbouring amino acids, and right in the middle of the domain[4] . Note also that tryptophan is 'the only side chain capable of participating as a donor in charge-transfer complexes with pyridinium compounds and other electrophiles' [94]. Pyridinium complexes include DNA bases Thymine(T) and Cytosine(C). Thus the donor (and also possible acceptor) may be tryptophan. Tryptophan absorbs light strongly at $280nm$ and emits light in a way that is very sensitive to its surrounding environment, *i.e.,* it is solvatochromatic, and this depends on the polarity surrounding the residue. Thus different

[4]see http://www.blc.arizona.edu/MolecularGraphics/DNA_Structure/DNA_tutorial.HTML#Pyrimidine

conformations of a protein containing tryptophan will quench or enhance its natural fluorescence. Though these energy scales are larger than the small splitings we require, $280nm$ corresponds to a HOMO-LUMO gap of ~4.4 eV, a determation of wavelength categories with receptor types could be useful. Perhaps fluorescence studies using *fluoresence resonance energy transfer* (FRET) for the ORs would be useful in determining D-A splittings.

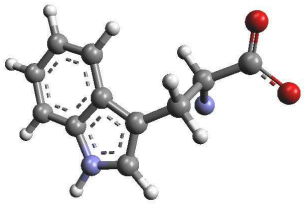


Figure 7.11: Tryptophan (Tryp)

7.4.3 The vibrations

C-H stretch vibrations are abundant in the environment, and occur at around 0.36eV (2911.3 cm^{-1}) [95]. We can estimate a root mean square displacement (rms) thermal atomic displacements of these stretches using:

$$\bar{x}^2 = \frac{1}{2}\frac{\hbar}{M\omega} coth\left(\frac{\hbar\omega}{2k_BT}\right), \tag{7.42}$$

which for our instance, has $coth \to 1$, as $\hbar\omega >> k_BT$ [91], as in the case C-H $360meV$ compared to $k_BT \sim 25meV$. Using $\frac{\hbar^2}{M} = 4.14meV$Å with M=1 for hydrogen in atomic units, and $\hbar\omega = 360meV$, we get a rms of 0.076Å. The smallness of these amplitudes of vibration, add to the assumption that the nuclear and electronic parts above are separable, because nuclear vibrations of the order 0.1Å would have little influence on the probability of electron tunneling over several Å [67]. However, if these oscillations are abundant, they would likely dominate; we must assume then that the frequencies around common CH stretches are 'blind-spots' where the receptor type does not exist [1].

7.4.4 The Huang-Rhys factors

The Huang-Rhys factor S, has important value in optical transitions, and also in olfaction, as we see here it can be used as a measure of the electron-phonon coupling strength. We take it as a calculable parameter, given we know about odorants, and have accurate methods to calculate vibrational modes, and hopefully it can allow us to make predictions for different odorants, based on the strength of this factor. Reasonable measures are 0.1-0.3, anything above this value does not significantly alter the rate and may lead to 2, 3 phonon processes. For more detail, see the investigation.

7.4.5 The electronic matrix element

A very important parameter is t- the electronic matrix element, and it is not unambiguously defined. t will be different in the absence of an odorant to the presence of the odorant, as there are shifts in geometry and energy caused. The localized energy states of D and A determine their coupling. Here (not knowing exactly what the D and A may be) we assume, as above, that they are single molecular orbital energies where the hopping integral between them is determined by: $t = v^2/(\varepsilon_M - \varepsilon_A)$, where ε_M is the energy level of the molecule and ε_A of the acceptor, if the former is a LUMO and the latter a HOMO. This can be an energy difference of up to $10eV$. The hopping integral v can be given an upper bound of $0.1eV$, determined from the strength of hydrogen bonds between the donor, acceptor and molecule. Knowledge of these structures would aid in obtaining a more specific value: the energy of tunneling should be determined with respect to the odorant, for the electron on the donor in the geometry of the activated complex R+M. For example Newton *et al* calculate the matrix element in Fe^{2+}-Fe^{3+}, from the overlap of the orbitals from these two iron atoms [67]. Nonetheless, the ratio of elastic to inelastic transmission is comparably insensitive to this value, so this is not essential for a decent analysis of the rate. We can compare to the experimental data for $C.vinosum$, from which, when the experimental parameters are inserted into a similar equation to 7.35 (with the same assumptions of non-adiabacity and low temperatures), then the matrix element obtained is 2.4meV, which indicates the right magnitude [67].

An interesting consideration of 'promoting' and 'accepting' modes in non-radiative transitions arises here. Accepting modes are those typically seen in configuration

coordinate diagrams. Promoting modes allow the right matrix elements for a transfer to occur, whereas the accepting modes provide the energy/driving force for such a transition, as mentioned in Chapter 'A physical picture'. Whilst the promoting modes may not account for smell (they can have Huang-Rhys factor 0) it is possible they are responsible for the intensity, if promoting and accepting modes occur simultaneously.

7.5 Discussion

Examining the important parts of the rate equation 7.35: the environment component (internal information; the vibrations) and electronic matrix elements (conformation, deformation, binding) could point to measurements of efficacy versus affinity respectively. Though the elements are similarly non-straightforward because the two concepts are both non-separable and interdependent.

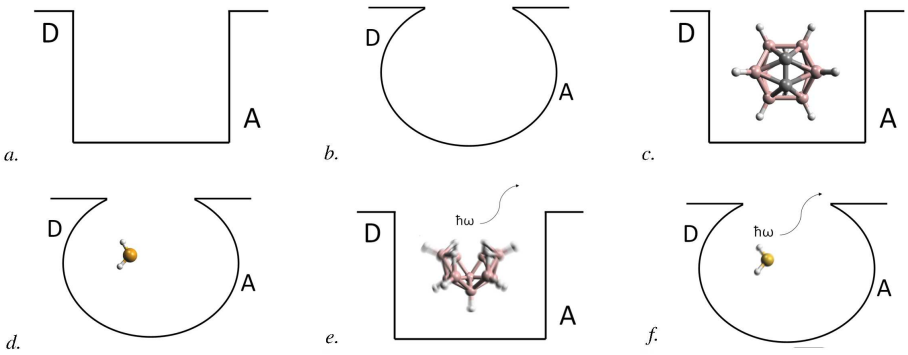


Figure 7.12: Possible Scenarios

For the sake of the soluble, we have kept the system under study very simple. We assume only:

1. One receptor type.
2. One odorant that binds.
3. One itinerant electron.
4. One odorant phonon that is probed.

Considering only these factors, we can imagine possible scenarios and examine what effect there would be on the rate and the model for scent discrimination. See figure 7.12, where there are 6 cartoons that represent possible scenarios. The receptor is depicted as a bucket with energy states D and A either side. Binding is indicated by deformation of the well, and actuation is indicated by the excitation of the right mode. Assume the receptor type is tuned to 'sulphur' which detects the S-H vibration at around 2600cm^{-1}. So we assume the D-A splitting corresponds to this, and that there is a good probability of electron transfer, that goes with, and is not necessarily separable from, a favourable lattice response. In other words the important mixing of orbitals are those particular to a molecule with the right vibration. Consider the various M's in the following scenarios:

1. Scenario a: R-M. There is no odorant, so S=0, unless there are CH vibrations in the environment (which is very likely). If we assume the efficient hopping integral of 1 meV is weaker when the 'right' configuration is not induced (*i.e.*, there is a wrong mixing of orbitals), if t=0.01 meV, then the inelastic time is still faster than the elastic time, though perhaps too slow (now μs) for a good signal. An inelastic signal is unlikely, anyway, as the CH vibrations are $\sim 300cm^{-1}$ higher in frequency than the D-A splitting, and the elastic signal is certainly too slow, approaching ms. We assume under these conditions the receptor is 'off'.

2. Scenario b: R-M. There is a probability, due to receptor dynamics, that the system makes an efficient configuration that allows the mixing of appropriate states for a decent electronic matrix element. The receptor could turn itself 'on'. Given the environment oscillations, there is then a good enough probability of elastic tunnelling, though estimations based on *inter*-protein tunneling should be compared, see below. We assume under these conditions the receptor has a likelihood of turning itself 'on'. If the matrix elements are not induced by an odorant however, this scenario is likely quite random, and so it seems unlikely the brain can extract such signals. This implies there must be some way the receptors are detuned to CH vibrations, as Turin suggested in 1996 [1].

3. Scenario c: R+M. There is an odorant, but it *does not* allow favourable mixing of the transition matrix elements, it has an S factor but not for the right vi-

bration. The receptor is 'off', as it is unlikely there is either IETS or ETS, in the same way as scenario a. We assume under these conditions the receptor is 'off'. An example would be M = carborane which smells camphoraceous, and not at all sulphuraceous.

4. Scenario d: R+M. There is an odorant, that *does* allow favourable mixing of the transition matrix elements, it has an S value, but it does not have the right vibration. The receptor is 'off', as it is unlikely there is IETS, though there could be ETS. An example would be M = selenium hydride, which has an 'irritating smell of decayed horseradish'. This is a similar scenario to 2, except presence of the odorant makes it much more likely, to induce the right state. Under these conditions, we have to assume either, the elastic rate under these conditions is too slow ($\tau_{T0} \sim 87ns$, and so the receptors, are extremely sensitive), or that there simply isn't a degenerate A state for the electron to get to, or that the elastic channel is suppressed in some other way, to assume this receptor is 'off'.

5. Scenario e: R+M. There is an odorant, but it does not allow the favourable mixing of the transition matrix elements, it has an S value, but it does have the right vibrations. The receptor is 'on' , as there can be IETS and ETS, though IETS dominates. An example would be M = decaborane which smells of sulphur. However, if for example the IETS rate for discrimination was given by the parameters in table 7.1, then we require t > 0.1 meV, as a t value much higher would make τ_{T1} of the scale $\tau_{T0} \sim 10's\, ns$ which we assume above is comparably too slow for this receptor R.

6. Scenario f: R+M. There is an odorant that *does* allow favourable mixing of the transition matrix elements, it has an S value, and it has the 'right' vibration. The receptor is 'on', as there can be IETS and ETS. However, as we see above the inelastic route is favourable. An example is M = sulphur, which smells sulphuraceous!

We see in scenarios b and d that there could be a real chance of ETS which we can't ignore. However, implicit is the assumption that receptor types are tuned by their D-A splitting. Thus where cases of ETS are possible, this goes against the assumption of the receptor 'type'. The electron has to have a final state that is

appropriate, A, where the electron can go to cause signal initiation. If the electron doesn't have a state to reside it may resonantly go back and forth, and not drive a forward non-reversible reaction, which is required for signal transduction. The D-A states cannot be represented by the continuous levels you get in metal junctions, but we are limited by HOMO and LUMO states, most likely on the amino acids of the OR. In conventional IETS, usually the elastic channel will always dominate; because there is a Fermi level with empty states above and filled below, whereas we require a discrete level A. What we would really desire, for absolute determination, is clarification of possible donor and acceptor states.

7.5.1 Timescales

Above we have a list of the important time-scales for overall recognition and we assume that this suffices for G-protein release, under the simplifications given in section 7.3. To summarize:

1. Overall odorant recognition is estimated in ms, see Bhandawat *et al*, for their estimates on this timescale for odorant transduction times in frogs of $\sim 50ms$ [96].

2. Diffusion of M through the mucus layer, of if we assume the maximum thickness of the epithelium $\delta = 200\mu m$, [30] then diffusion would take $\sim \delta^2/D \sim 20ms$.

3. Peak production of cAMP occurs at $\sim 25ms$ according to the results of Boekhoff *et al* [97], given overall transduction, this suggests a dwell time of M at the binding site of ms.

4. In the meantime, it looks as though the D and A are replenished in $1ms - 1\mu s$. So with a dwell time of ms, there is a possibility 1000's of electrons may transfer if they arrive and leave in μs. We assume that the electron leaves site A and the G-protein is replenished on a similar timescale.

5. Electron transfer is slower than the vibrational lifetime of the odorant. The normal vibration quanta are not stable, they impart energy to phonons at a decay rate of $\tau \sim 10^{-12} - 10^{-13}$secs [98]. A lifetime estimated from our FWHM at ~100cm^{-1}, indicates a lifetime of 0.33ps. So the energy transferred

to vibrational degrees of freedom dissipates, relatively quickly, compared to electron transfer.

6. What are the dynamics of the odorant, indeed the odorant and the OR, or the OR isolated? Indeed, as noted above, thermal motion may cause the OR protein to spontaneously transform conformational states. Protein fluctutation barriers, for the heme pocket of myoglobin, are around 10kcal/mol which correspond to 1-10μs [99]. For the docked system conformational changes are likely induced by electronic interactions, this manipulation of electronic orbitals is referred to as 'orbital steering' [100], into certain favourable geometries for the ligand and the protein. This may happen on the faster scale of several ps [100]. We may assume this change goes with a favourable mixing of TM elements, in which case the tunneling of the electron occurs at an equilibrium configuration. However if it is the case that the lifetime of the complex formation is longer than the electron transition time we must account for it passing through non-equilibrium conformations.

7.5.2 Intensity

The intensity is not directly involved in the rate equation but is obviously important in detection nonetheless. A simple linear correlation between the amount of odorant emitted from a source and the perceived intensity of the source does not exist [25]. Increasing the concentration of a gas does not necessarily result in a correspondingly more *powerful* sensation, it actually results in a *different* sensation, as more of the receptor types are activated, which make sense from a saturation point of view, and indicates ORs are not considerably selective. Also, see Keller and Voshall (2007), who show in psychophysical methods that for 85% of *aldehydes* tested, humans perceived the higher the vapour pressure (VP) of the source the stronger the response [101]. The VP corresponds to how many odorants enter the nose. Conversely for 72% of *alcohols* the higher the VP the lesser the intensity. It seems intensity does not simply go with bombardment of odorants to receptors. So, other factors that may influence the intensity of a signal may be:

1. How *long* the odorant may be in the receptor; the time scale for the receptors "on" state may reflect the intensity of a signal.

2. We do not know how *many* electrons are fired at the odorant; we assume one for simplicity, many more may result in a larger signal corresponding to a current. Note this may not necessarily go with the above point 1.

3. How *many* odorants may be at the binding site. It may be detrimental or beneficial if a receptor is saturated by an odorant. If many are within the binding site, we still assume one odorant per binding site, the odorant probed, will be that with the best S factor which again will depend on the fit and the vibration.

4. The alignment of the odorant within the receptor; this in turn may affect the efficiency of the tunneling route, *i.e.*, the faster the rate. We can test this by estimating a rate for the most potent chemicals, based on the S factor, which is examined in the next chapter.

5. Peripheral mechanisms may sort different intensities according to different types; for example, dangerous smells are always amplified. Intensity may be entirely a post recognition issue, where amplification/deamplification occurs at the synapse level.

7.5.3 Charge transfer in general

We only consider the quantum mechanics for electron transfer, but of course there are other possibilities including other mobile charged particles:

1. *A triplet excitation.* Its motion corresponds to the exchange of two electrons between D and A. The origin of the triplet excitation of D might be singlet oxygen, though the low concentration in the atmosphere is probably insufficient for this.

2. *A proton.* The origin of this would be acidic species. The short distance over which protons can tunnel makes this relatively unlikely because of the need to exert a substantial change in force on the odorant.

All of the above mechanisms can be described by the same quantum mechanical formalism, with appropriately modified definitions of the key terms. Proton release is considered elsewhere [19].

7.6 Conclusions

7.6.1 Inelastic versus elastic tunnelling

In virtually all cases since its earliest observations, tunneling with phonon emission is a weak adjunct to the elastic component. Here we show for olfaction, that the inelastic channel can overwhelm the elastic. As long as the inelastic tunneling time is shorter than microseconds (overall recognition in milliseconds) then this should be fast enough. The crucial point is that the elastic tunneling time is longer. There is no problem getting fast enough inelastic tunneling, the problem is ensuring that elastic tunneling is far less rapid. We can estimate the possibility of electron transfer occurring when there is no odorant present by referring back to equation 7.40. If we assume the other parameters are relatively unchanged, but now the electron has to travel a considerably greater distance (contrast within the protein helix to across protein helix, see figure 7.10) then we can see how the time will increase. Increasing the hopping distance to just 8Å (as above see section 7.1.2) increases the time to greater than ms (27ms), *i.e.*, too slow for recognition. So if D and A are simply too far away from each other, then ETS is unlikely. It may even be that prior to odorant docking $\varepsilon_D < \varepsilon_A$, thus inhibiting tunneling further in the absence of the odorant. It may also be that docking of the odorant significantly lowers the reorganization energy (via the effects noted above) and the brain detects a *variance* in the current from elastic to inelastic.

7.6.2 The 'swipe card' paradigm

Indeed the protein fold may play an essential role in lowering the reorganization energy [88], if the protein folds with the ligand to induce a favourable transition state for electron transfer. Strongly smelling odorants that bind well to zinc may have the effect of lowering the reorganization energy to give the most efficient electron transfer. This implies then, that the faster the transfer of an electron then the better the ligands 'efficacy'. A proper evaluation of protein folding and the likely interactions of protein and ligand and redox centre with the appropriate geometries are thus likely important. The important point to note, really, is that it is not just the vibrations of the odorant that are important to detection, but in turn, unavoidably is the importance of the shape of the odorant, because it plays a part in many

other important parameters in the rate, to get an optimal value.

The culmination of this result, and similar appearances elsewhere [19], has led to the introduction of the 'swipe card' paradigm to rival the 'lock and key' idea. In the 'lock and key' complementary fit is required to 'fit the lock' but there is no explanation how as to 'open the door'. In the 'swipe card' model, just as in the modern key card used in hotel rooms: the right kind of shape is *necessary* but not *sufficient* to open the door. What opens the door is a combination of fit *and* the right internal information, in the case of this analogy, an electromagnetic strip. In olfaction a complementary fit is required (for optimization of the parameters above, such as λ, t, and as we shall see next in more detail S) but it is the right vibration which facilitates the electrons journey and causes a forwards signalling transduction event.

Chapter 8

Huang-Rhys factors

As we have seen in the previous chapter, an important parameter in the rate equation, contained in $\sigma(0, n)$, is the 'Huang-Rhys' factor, which for brevity hereafter we shall refer to as the 'HR factor'. It measures how readily energy is taken up by an oscillator in an electronic Hamiltonian and is measured by the dimensionless factor S [56]:

$$S_\alpha = \frac{1}{2\hbar\omega_\alpha \tilde{m}_\alpha} \left(\frac{\Delta f_\alpha}{\omega_\alpha} \right)^2 = \frac{1}{2\hbar\omega_\alpha} \left| \Delta Q_\alpha \right|^2 \omega_\alpha^2. \tag{8.1}$$

For an in detail derivation of these equations, see the Appendix. ΔQ_α is the change in mode, ω_α and $\tilde{m}_\alpha$ are the odorant vibrational frequency and reduced mass, respectively, for mode α and is easily obtained as described in a coming section. More ambiguous is the change in force, Δf_α, caused by the electron transition, and this is explored in this chapter in later sections.

Since S measures the strength of electron-phonon coupling, we conjecture that it is an important indication of an odorant signal's strength. Thus, we seek to model 'odorant spectra' , *i.e.*, a Huang-Rhys factor versus the mode of vibration. This allows us to see if Turin's theory agrees with observation and if similar smelling odorants look similar spectrally, in a predictive and informative manner. The concept of the Huang-Rhys factors allow us, despite knowing so little about the receptors, to evaluate the strength possible activating capacity of an odorant. There are plenty of interesting test cases in nature to examine, for example:

1. Odorants with similar structures that smell different.

2. Odorants with different structures that smell the same.

3. Enantiomers (mirror image) molecules.

4. Isotope molecules.

5. Agonists versus antagonists.

Of this list we test examples of point 2 above. This allows us to examine if the swipe card paradigm is able to predict and explain the results of different odorants, which is the ultimate aim for olfaction and all related drug protein interactions.

To estimate a Huang-Rhys factor for a given odorant vibrational mode, we need a good representation of the electron- odorant (oscillator) system. Herein lies the difficulty. We present three main models, the first calculates total energies and avoids analysis of component forces, the second and third models approximate, in different ways, as we shall see, explicitly the long range electrostatic interactions between the itinerant electron and the odorant atomic partial charges. We shall call these models the 'parabola model', the 'IR absorbancy model and the 'partial charge' model respectively.

8.1 The parabola model

As we have seen previously, from the harmonic approximation, we can find the reorganization energy and hence a Huang-Rhys factor, via a configuration coordinate diagram like figure 8.1. Examining this figure, we can obtain a value for S using simple algebra, as long as we can calculate the total energies, and this avoids any complicated arguments for forces and definitions for charge.

8.1.1 The model

We have:

$$E_D = \varepsilon + Cq^2, \tag{8.2}$$

for the donor surface D and;

$$E_A = C(q - q_0)^2 = Cq^2 + 2Cqq_0 + Cq_0^2, \tag{8.3}$$

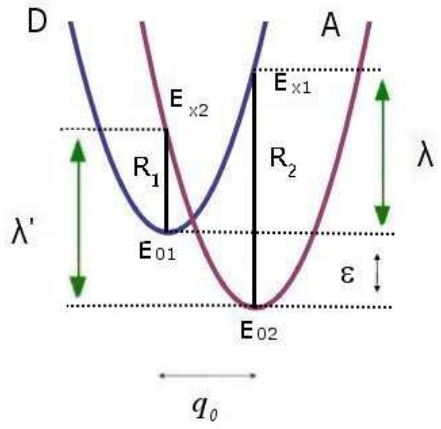


Figure 8.1: Configuration coordinate for electron transfer from D to A. The two reorganization energies are depicted, see Chapter 'A physical picture' for definitions, and the vertical transition energies R_1 and R_2.

for the acceptor surface A where (from Hooke's law) $C = 1/2k = 1/2M\omega^2$. Then the vertical transitions in figure 8.1 are:

$$R_1 = Cq_0^2 - \varepsilon, \tag{8.4}$$

and

$$R_2 = \varepsilon + Cq_0^2. \tag{8.5}$$

where ε is the *driving force* (see figure 8.1). These equations give $\varepsilon = (R_2 - R_1)/2$ and $Cq_0^2 = (R_1 + R_2)/2$. Examination of the change in force as we displace by q_0, in either direction $D \rightarrow A$ or $A \rightarrow D$; we get equal forces:

$$\left.\frac{dE_D}{dq}\right|_{q=q_0} = 2Cq_0, \tag{8.6}$$

$$\left.\frac{dE_A}{dq}\right|_{q=0} = 2Cq_0. \tag{8.7}$$

We require a value for the relaxation energy due to this change in force. This is given by $\Delta f^2/2k = (2Cq_0)^2/2k = 1/2kq_0^2 = Cq_0^2 = (R_1 + R_2)/2$, so a Huang-Rhys factor is simply:

$$S = \frac{(R_1 + R_2)}{2\hbar\omega}. \tag{8.8}$$

8.1.2 The method

We use the simplest system: a diatomic molecule (HCl) is positioned with its bond parallel and between the two points; site 1 and site 2. Site 1 (the donor position) is where the electron is first placed and site 2 (the acceptor position) is where the electron moves to. In this calculation the electron is placed at the donor 3Å away from the chlorine atom and then on the acceptor 3Å away from the chlorine atom. These calculations were done using Gaussian methods described elsewhere, using the method B3LYP and basis set 6-31G**. For $(R_1 + R_2)$, we must calculate the total energies E_{01}, E_{02}, E_{X1} and E_{X2}, as seen above in figure 8.1. The computations made were:

E_{01} This is the energy with the electron at site 1. A full self consistent relaxation and geometry optimization (see Chapter 'Density Functional Theory' for details on these) is performed.

E_{X2} This is the energy with the electron at site 2. This time freezing the geometry at the previous geometry for E_{01} and calculating to full self-consistency but with no geometry relaxation.

E_{02} This is the energy with the electron at site 2. A full self consistent relaxation and geometry optimization is performed.

E_{X1} This is the energy with the electron at site 1. Now freezing the geometry at the previous geometry for E_{02} and calculating to full self-consistency but with no geometry relaxation.

The results are given in the table 8.1 below.

Electron at site...	Relaxed (a.u)	Not relaxed (a.u)	Reorganization Energy (a.u)
1	E_{01}=-460.801628148	E_{X1}=-460.801566537	$\lambda = E_{X1} - E_{01} = 0.000062$ (0.001687 eV)
2	E_{02}=-460.818529377	E_{X2}=-460.818518782	$\lambda^{'} = E_{X2} - E_{02} = 0.000011$ (0.000299 eV)

Table 8.1: A table to show total energy calculations for HCl between two sites for an electron charge, The energy levels correspond to those annotated in figure 8.1.

Considering a simple diatomic molecule, translations and rotations are not considered, we take $\hbar\omega = 0.36eV$, which corresponds to the only mode (3N-5 for linear diatomics). We get, converting a.u to eV ($\times 27.212$), then $S = 0.0027$, as we need S >> 0.0001, from the rate equation in the previous chapter, we easily satisfy the criterion for a significant signal.

8.1.3 Analysis

Examining energy levels in table 8.1, the picture 8.1, most likely looks something like figure 8.2, which agrees for the small coupling assumptions we have assumed here (small couplings still give us efficient rates). Of course a full calculation of the potential energy surface would be required to verify this, and indeed maybe a more accurate method: no optimizations would be required; just energy calculations at various configurations. However, this model even in the simple state described here, is quite inefficient and costly in both computational and real time. Another point to raise is the degree of precision: we have values of energy to 6 d.p of a.u which is too small to attach too much significance to. Further, we have used a point charge at the location 3Å from the atomic charge, a better representation may be a uniform electric field, for reasons we describe below. Referring back to the generic equation 8.1, where E_R is the relaxation energy $S = \frac{E_R}{\hbar\omega}$; there are ways to write this equation for the energy in terms of forces and this may be more convenient. The change in force arises from an oscillating molecule interacting with a charge distribution. So we examine two other possible models in depth.

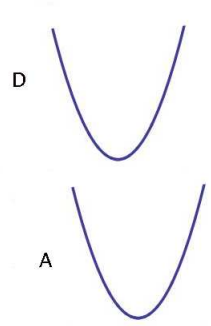


Figure 8.2: Configuration coordinate for a phonon excitation of the mode HCl.

8.2 The diatomic

Consider, the simplest model for monopoles on each atom, as in figure 8.3.

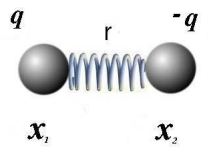


Figure 8.3: A simple diatomic dipole.

The dipole moment for this system is

$$\mu = q\left(x_2 - x_1\right), \tag{8.9}$$

whilst for a continuous charge distribution (still in 1 dimension) the dipole moment is

$$\mu = \int \rho(x)\, x\, dx, \tag{8.10}$$

From equation 8.9, we can define an effective atomic charge, for a bond length r; $q = \mu/r$, and thus a rate of change, $q = \frac{\partial \mu}{\partial r}$, this is assuming q is independent of

bond length. Consider two positions for the two atoms:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{8.11}$$

$$\vec{x}_0 = \begin{pmatrix} 0 \\ r \end{pmatrix}. \tag{8.12}$$

Then the displacements are given by:

$$\Delta\vec{x} = \vec{x} - \vec{x}_0 = \begin{pmatrix} x_1 \\ x_2 - r \end{pmatrix}. \tag{8.13}$$

We require the displacements in terms of normal modes. For a dimer the equations of motion from equilibrium are:

$$\begin{aligned} m_1\ddot{u}_1 &= -k\left(u_1 - u_2\right) \\ m_2\ddot{u}_2 &= -k\left(u_2 - u_1\right). \end{aligned} \tag{8.14}$$

Which we use because Gaussian works in units of displacement, u_1 and u_2 from equilibrium. Solving for eigenvalues and eigenvectors in the usual way (see Chapter 'A physical picture' and below) we get the usual expression for ω and the normal mode eigenvector is:

$$\vec{u} = \left(-\frac{m_2}{\sqrt{m_1^2 + m_2^2}}, \frac{m_1}{\sqrt{m_1^2 + m_2^2}}\right). \tag{8.15}$$

Now, if the center of mass does not move, then the displacement of atoms can be described in terms of modes. The displacement $\Delta\vec{x}$ is related to the Gaussian normalized displacements by $\Delta\vec{x} = Q.\vec{u}$, where Q is the size of the displacement in Å, because the length of $\Delta\vec{x}$ is given by $\sqrt{\Delta\vec{x}.\Delta\vec{x}} = |Q|\sqrt{\vec{u}.\vec{u}} = |Q|$. The atomic positions are:

$$\vec{x}(Q) = \begin{pmatrix} Qu_1 \\ r + Qu_2 \end{pmatrix}, \tag{8.16}$$

such that the distance between atoms 1 and 2 is then given by:

$$x_2 - x_1 = r + Q(u_2 - u_1) = r + Q\frac{(m_1 + m_2)}{\sqrt{m_1^2 + m_2^2}}, \tag{8.17}$$

if we insert the eigenvectors given by equation 8.15. We see explicitly that the change in dipole moment changes with the normal mode, the dipole moment depends on bond length. Inserting equation 8.17 into equation 8.9, we get:

$$\mu(Q) = q.\left(r + Q\frac{(m_1 + m_2)}{\sqrt{m_1^2 + m_2^2}}\right), \tag{8.18}$$

and we have a dipole moment that depends on the size of the displacement along a mode. Thus:

$$\frac{\partial \mu}{\partial Q} = q\frac{(m_1 + m_2)}{\sqrt{m_1^2 + m_2^2}}, \tag{8.19}$$

which (for a diatomic) should agree with what Gaussian computes in an IR absorbance calculation. The Gaussian program obtains the dipole moment derivative by finding the second derivative of energy with respect to nuclear coordinates and the electric field perturbations. This works if we assume the donor and acceptor produce uniform electric fields when charged, as above. To check the analytical values with some numerical values, we define a change in bond length:

$$\Delta r = x_2 - x_1 - r = Q\frac{(m_1 + m_2)}{\sqrt{m_1^2 + m_2^2}}, \tag{8.20}$$

these are values that are important later in determining the partial charges for a simple diatomic.

8.3 Generic math model

Repeating the generic model for a Huang-Rhys factor:

$$S_\alpha = \frac{E_R}{\hbar\omega} = \frac{1}{2\hbar\omega_\alpha \tilde{m}_\alpha}\left(\frac{\Delta f_\alpha}{\omega_\alpha}\right)^2 \tag{8.21}$$

where E_R is the relaxation energy. The electron moving from donor to acceptor exerts a force on the partial charges which is a response to an electric field. Though

the two models we describe next are different, they should converge at certain approximations. Can we approximate the point charge calculation to IR absorbance? The HR factor from the point charge model can be written:

$$S_\alpha = \frac{1}{2\hbar\omega_\alpha}\frac{1}{\omega_\alpha^2}\left(\Sigma_I \frac{\overrightarrow{u}_{\alpha,I}\cdot\Delta\overrightarrow{F}_I}{\sqrt{M}_I}\right)^2 \tag{8.22}$$

α is over the 3N vibrational modes, and I denotes atoms as usual. $\Delta\overrightarrow{F}_I$ is given by:

$$\Delta\overrightarrow{F_I} = \frac{-z_I e}{4\pi\varepsilon_{\ 0}}\int(\rho_D(\overrightarrow{r}) - \rho_A(\overrightarrow{r}))\frac{\left(\overrightarrow{r} - \overrightarrow{R}_I^{(0)}\right)}{\left|\overrightarrow{r} - \overrightarrow{R}_I^{(0)}\right|^3}d\overrightarrow{r} \tag{8.23}$$

Which is the change in force on atom I. $\rho_D(\overrightarrow{r})$ and $\rho_A(\overrightarrow{r})$ are the charge densities at each site respectively, and $\overrightarrow{R}_I^{(0)}$ is the relaxed positions of the atoms. If $\Delta\overrightarrow{F}_I$ from equation 8.23 is substituted into equation 8.22 and we use the relationship between displacements and modes:

$$S_\alpha = \frac{1}{2\hbar\omega_\alpha}\frac{1}{\omega_\alpha^2}\left(\frac{\partial}{\partial Q_\alpha}\sum_I z_I e\int\frac{\rho_D(\overrightarrow{r}) - \rho_A(\overrightarrow{r})}{4\pi\varepsilon_0}\left|\overrightarrow{r} - \overrightarrow{R}_I^{(0)}\right|d\vec{r}\right)^2 \tag{8.24}$$

Which, if we approximate the donor and acceptor as infinite parallel sheets (rather than point charge) we get similar results, considering the different charge distributions, where :

$$S = \frac{1}{2}\frac{2mE^2}{\hbar\omega_\alpha^3}\left(\frac{\partial\mu}{\partial Q_\alpha}\right)^2. \tag{8.25}$$

Where we assume the electric field $E = E_D = E_A$, and accounted for the mass factor (for a diatomic) which is $m = \frac{m_1^2+m_2^2}{m_1 m_2(m_1+m_2)}$, and μ is the dipole moment. So we see directly the dependency of the Huang-Rhys factor on the derivative of the dipole moment with respect to the normal modes of vibration. Essentially the Huang-Rhys factor is proportional to IR absorbance. Equation 8.25 however, requires the definition of a uniform electric field, the models below assume point charges for D and A as opposed to the infinite parallel sheets of charge as shown in figure 8.4.

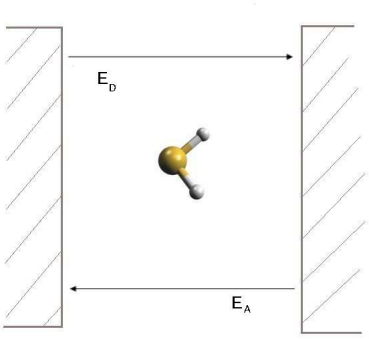


Figure 8.4: Figure to show an odorant between the receptor acting as a parallel plate capacitor.

8.4 Vibrational analysis for Gaussian '03

A lot of the information required in calculating S, using the methods described in the following, comes from the information output from the computer program Gaussian '03. Here, we briefly describe what the important information is, and how we extract and use it. For the modes of vibration, the method of solution described here will be that used by the program Gaussian '03 and as described by J. Ochterski in 'Vibrational Analysis in Gaussian' and also by Wilson, Decius and Cross (1955) , [83], [57]. Starting from the calculation of the Hessian matrix:

$$f_{cartij} = \left(\frac{\partial^2 V}{\partial\varepsilon_i \partial\varepsilon_j}\right)_0 \tag{8.26}$$

Equation 8.26 holds the second partial derivatives with respect to Cartesian displacement coordinates; *i.e.*, the force or 'spring' constants, of course at equilibrium position. $\varepsilon_1, \varepsilon_2, \varepsilon_3 .... \varepsilon_{3N} \rightarrow \Delta x_1 \Delta y_1 \Delta z_1 ... \Delta z_N$. The Hessian is then transformed to mass-weighted coordinates;

$$f_{mwcij} = \left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_0 = \frac{f_{cartij}}{\sqrt{m_i m_j}}, \tag{8.27}$$

where $q_1 = \sqrt{m_1}\varepsilon_1 = \sqrt{m_1}\Delta x_1$, and so on, as we have seen in the Chapter 'A physical picture', see equation 28. This matrix can then be diagonalized for 3N eigenvalues and 3N eigenvectors (the normal modes) but the 6 translations and rotations have to be separated out: see Ochterski for detail on how this is done. Put briefly, a *Schmidt* orthogonalization is used to generate the 3N-6 vectors which are orthogonal to the rotation and translation vectors and the result is a transformation where $S = Dq$, S are the *internal coordinates* and D transforms the mass-weighted coordinates q, into this new frame. So the Hessian f_{mwc} is then converted to these new coordinates by:

$$f_{int} = D^{\dagger} f_{mwc} D. \tag{8.28}$$

Then diagonalized:

$$L^{\dagger} f_{int} L = \wedge, \tag{8.29}$$

where $\wedge$ is a diagonal matrix consisting of $\lambda_\alpha = 4\pi^2 v_\alpha^2$, α denotes the 3N-6 modes, as usual. These eigenvalues are the *normal mode frequencies* and are converted into such (in wavenumbers $\tilde{v}_\alpha$, in cm^{-1}) by:

$$\tilde{v}_\alpha = \sqrt{\frac{\lambda_\alpha}{4\pi^2 c^2}}, \tag{8.30}$$

let $\omega_\alpha = 2\pi v_\alpha = 2\pi c \tilde{v}_\alpha$ for α =3N-6 modes. So we can pick ω_0 for whichever mode of vibration we think might be interesting. The wavenumber of the fundamental frequencies are printed by Gaussian, along with, for each mode, the displacements, the reduced mass, the force constants. $Q_{cart} = MDL$ is calculated where $M_I = \frac{1}{\sqrt{m_I}}$ is a diagonal matrix, where I $\rightarrow$x,y,z for each atom, as usual. This is:

$$\chi_{cartI,\alpha} = \sum_{j}^{3N} \frac{D_{i,j} L_{j,\alpha}}{\sqrt{m_j}}, \tag{8.31}$$

which contains the normal mode column vectors, and a mass factor, which ensures the trace of the transformation matrix is unity. These *normal modes* in Cartesian coordinates are normalized in Gaussian to become the *displacements* in Cartesian

coordinates. The normalizing factor is:

$$N_\alpha = \sqrt{\left(\sum_I^{3N} \chi^2_{cart\,I,\alpha}\right)^{-1}}, \tag{8.32}$$

and each element of $\mathrm{I_{cart}}$ is scaled by N_α for each mode. Then the reduced mass is given by:

$$\tilde{m}_\alpha = \left(\sum_I^{3N} \chi^2_{cart\,I,\alpha}\right)^{-1} = \left(\sum_I^{3N} \left(\frac{\chi^2_{mwc\,I,\alpha,}}{\sqrt{m_I}}\right)^2\right)^{-1} = N^2_\alpha, \tag{8.33}$$

which is obviously slightly different for the polyatomic case, than compared with a simple diatomic. Gaussian prints the normalized Cartesian displacement $\sum_I u^2_{\alpha,I} = 1$.

$$u_{cart\alpha,I} = \frac{\chi_{cart\,\alpha,I}}{\sqrt{\left(\sum_I^{3N} \chi^2_{cart\,I,\alpha,}\right)}} = \frac{\chi_{cart\,\alpha,I}}{|\vec{\chi}_\alpha|} \tag{8.34}$$

Where $\chi_{\alpha,I} = \frac{Q_{\alpha,I}}{\sqrt{m_I}}$, from above equation 8.31. u are the normalized displacements in Cartesian displacements and these are the values printed in the output. Linear combinations transforming the displacements to normal modes eliminates any cross terms, as usual. The force constant for each mode is given by:

$$k_\alpha = 4\pi^2 \tilde{\nu}^2 \tilde{m}_\alpha, \tag{8.35}$$

where $\tilde{\nu} = \frac{1}{2\pi}\sqrt{\frac{k_\alpha}{m_\alpha}}$. So for each mode of vibration $\omega_0$of an odorant (ω_α) a reduced mass, a force constant and the displacements of the atoms are found. Also, if requested, the normal modes are used to calculate the Infrared (IR) intensity of the modes, which is a quantity we see to be useful. This value for each mode (Q_α) is found using equation 4.59 in Chapter 'A physical picture', the section in parenthesis is the "absolute" IR as calculated in Gaussian and is given in units of km/mol. Unit conversion is given by: 1 $(D/\text{Å})^2\text{amu}^{-1}$=42.255 km/mol = 171.65$\text{cm}^{-2}\text{atm}^{-1}$ at 0^oC and 1 atm. All the molecular properties $\omega_\alpha, \tilde{m}_\alpha$ and $\left|\frac{d\mu}{dQ_\alpha}\right|^2$ (unless stated otherwise) are calculated using Gaussian'03 with DFT method B3LYP and the basis set 6-31**, for consistency.

8.5 The IR absorbance model

We conjecture above that the HR factors depend on the rate of change of electric dipole moment with nuclear displacement, which can be derived from the IR absorbance.

8.5.1 The model

We treat the odorant (oscillating dipole) as a dipole with charges $\pm qe$, and calculate the change in force upon the electron transition from D to A. The electron is treated as a point charge on either D or A. We assume the odorant, upon excitation of the right mode, has an overall dipole moment vector in a direction parallel to the donor and acceptor. This mimics a receptor-odorant system that would maximize the change in force, and so the HR factor. For simplicity, we also assume the donor and acceptor are equidistant from the centre of the dipole. The system configuration is represented pictorially in figure 8.5.

$$S = 4q^2 \frac{m_e}{\tilde{m}_\alpha} \left(\frac{Ry}{\hbar \omega_\alpha} \right)^3 \left(\frac{\hat{R}_D \cdot \hat{p}}{\left(\frac{R_D}{a_o} \right)^2} - \frac{\hat{R}_A \cdot \hat{p}}{\left(\frac{R_A}{a_o} \right)^2} \right)^2 \tag{8.36}$$

The terms involved are; m_e the electron mass, $\tilde{m}_\alpha$ the mass of the odorant (here the reduced mass as calculated by Gaussian for each mode), ω_α as usual, $\hat{R}_D$ is the vector position of the donor from the centre of the dipole (and similarly for A, see figure 8.5), and $\hat{p}$ is the direction of the dipole. Consequentially the Huang-Rhys factor has a simple dependence in the position from the D/A. Here the last term is a geometrical term, where we will assume a 'perfect' alignment of the dipole, *i.e.*, one that maximizes the HR factor.

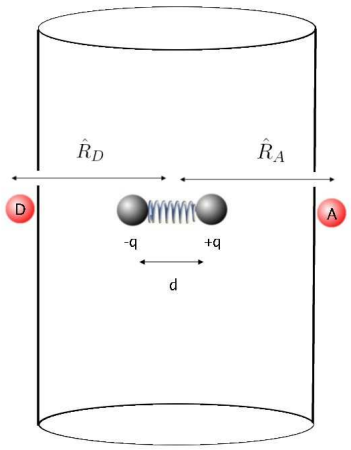


Figure 8.5: The model of odorant (as a simple dipole of $\pm qe$) within the TM domain (a simple cylinder) with an electron at the donor (or acceptor position). $\hat{R}_D = \hat{R}_A$ measures distance from D/A to centre of the dipole, d is the length of the dipole.

8.5.2 The method: determination of partial charge 'q'

The parameter q in equation 8.36 is found from the IR absorbance as calculated from Gaussian for each mode. q is a partial charge (a fraction of electron charge e) such that the odorant is simplified as a dipole. Optical transitions are governed by a change in this dipole moment and a large IR absorbance indicates that there has been a significant change in dipole moment as a certain mode is excited. The IR absorption is defined as [57]:

$$I_{\alpha}^{IR} = \frac{N\pi}{3c} \sum_{v} \left[\left(\mu_x^{(av)} \right)^2 + \left(\mu_y^{(av)} \right)^2 + \left(\mu_z^{(av)} \right)^2 \right] \tag{8.37}$$

as mentioned previously in Chapter 'A physical picture'. The sum is over atoms v. For the α^{th} mode, $\mu_x^{(ia)}$is the dipole derivative of the electric moment for the Cartesian direction with respect to the normal mode coordinate (Q_α), from this the partial charge is extracted. Gaussian prints the dipole moment derivatives (the section in parenthesis) as the absolute IR intensity in km/mol [79]. Unit conversion is given by: 1 $(D/\text{Å})^2\text{amu}^{-1}$=42.255 km/mol = 171.65$\text{cm}^{-2}\text{atm}^{-1}$ at 0^oC and 1 atm

[1]. All the molecular properties ω_α, $\tilde{m}_\alpha$ and $\left|\frac{\partial\mu}{\partial Q_\alpha}\right|^2$ used in calculating the HR factor via equation 8.36 (unless stated otherwise) were calculated using Gaussian '03 with the DFT functional B3LYP and the basis set 6-31** consistently.

8.5.3 The assumptions and accuracy

The important assumptions used here are:

1. The acceptor to centre of dipole distance is equidistant to the donor to centre of dipole distance.

2. There is a 'perfect' alignment of the dipole parallel to donor and acceptor.

Considering the sensitivity to $\hat{R}_D/\hat{R}_A$, the first assumption may be crucial, further it may not be the right assumption for convergence to other models. The second assumption may also be crucial: the model is limited in finding the largest Huang-Rhys factor, but we lose sensitivity to geometrical factors by using an average dipole moment. This is important, as we have noted before; the relative directions of the electron and the dipole (thus the direction of scattering) is important, see for example the investigation chapter examining chiral molecules. It may be more important in 'odorant spectra', for cross comparisons: to fix the odorant at one overall geometry (as it may be fixed in binding) and then to see the impact on the couplings for all the other modes.

As regards accuracy in general, the electronic structure method of calculation for IR may not be totally reliable. Theoretically determined intensities may differ from the experimental by up to 10-50%. The accuracy of IR calculations from first principles is disputable [102]. Calculations for hydrogen sulphide in particular seem to be very sensitive to the basis set used, see table 8.2 for comparisons.

[1]The km/mol given for the IR intensities in Gaussian is *kilometers per mol,* which is strange but true. The absorption of IR is a function $1/cL$ times the integral lnI_0/I over the vibrational modes, where c is the concentration and L the path length, I is absorbance and I_0 absorbance under initial conditions. The units then work out as $cm^{-1}.cm^3mol^{-1}.cm^{-1}.km/10^5cm$.

Mode	B3LYP/6-31**	B3LYP/cc-pVTZ	$I^{IR}/\left(D/\text{Å}\right)^2 amu^{-1}$
$1A_1$	4.9 km/mol	0.49 km/mol	=0.0116
$2A_1$	6.7 km/mol	0.74 km/mol	=0.0175
$1B_2$	8.6 km/mol	0.80 km/mol	=0.0189

Table 8.2: IR intensities for the vibrational modes of hydrogen sulphide. For more detail on the IR spectrum of hydrogen sulphide see the below section on the test molecules. See also, Chapter 'Density Functional Theory' for a discussion of methods used here. The latter column is converted into $\left(D/\text{Å}\right)^2 amu^{-1}$ for comparisons with IR absorption intensities of water done by other researchers in table 8.3.

Though the order is preserved, there is a factor of 10 difference in absorbancies after a change in basis set; the numbers become even smaller. They are very small, especially compared to the results of structurally and electronically similar water, see table 8.3.

Mode	LDA	B3LYP	EXPERIMENT
$1A_1$	1.841	2.303	1.16-1.59
$2A_1$	0.094	0.142	0.059
$1B_2$	1.742	1.252	1.00-1.42

Table 8.3: IR intensities for the vibrational modes of water. Units of IR intensity are in $\left(D/\text{Å}\right)^2 amu^{-1}$

In table 8.3 experimental values and theoretical (LDA and B3LYP) values are compared, as measured by Porezag and Pederson (1996) [79]. Double polarized basis sets are more accurate for IR, but it is noted that absolute values should be treated with trepidation, because of this wide variability with basis set. Information about the *relative* absorbance, however, can be discerned and compared.

8.6 The point charge model

Considering the above, we seek a second model. The above results do not include positions of all the relative partial charges, so with the ease of calculation comes the expense of analysing for particular configurations; the change in dipole moment is averaged out over all directions. This second model considers the directions of each

component force.

8.6.1 The model

We treat the odorant (oscillator) as a collection of dipoles for each bond. We calculate the change in force upon electron transition from D to A, by adding vectorally all the component Coulombic forces. Thus we explicitly account for every charge. For the interaction of a point charge at D and then A with all the partial charges on the odorant, we get an expression for the change in force:

$$\Delta f_\alpha = -\frac{e}{4\pi\varepsilon_0}\Sigma_I \left(\frac{\overrightarrow{u}_{\alpha,I} \cdot q_I(\overrightarrow{r}_A - \overrightarrow{R}_I^{(0)})}{\left|\overrightarrow{r}_A - \overrightarrow{R}_I^{(0)}\right|^3} - \frac{\overrightarrow{u}_{\alpha,I} \cdot q_I(\overrightarrow{r}_D - \overrightarrow{R}_I^{(0)})}{\left|\overrightarrow{r}_D - \overrightarrow{R}_I^{(0)}\right|^3} \right). \tag{8.38}$$

Each atom of the oscillating molecule is assigned a partial charge q_I , see figure 8.6, which can be inserted into equation 8.1, for a Huang-Rhys factor.

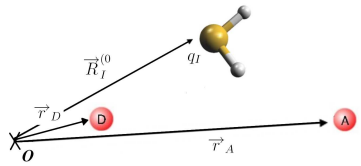


Figure 8.6: Figure to show how the forces are calculated, summing the Coulombic contributions.

Most of the information for the above calculation is then obtained using Gaussian; the displacements $\overrightarrow{u}_{\alpha,I}$ and eigenvalues ω_α are found as described above. The position of the itinerant electron from the relaxed individual atomic positions that make up the odorant are found vectorially, thus this approach reflects an electron transfer mechanism sensitive to relative positions of charge in the odorant. The change in force can be calculated for various positions of the donor and acceptor relative to the odorant. We do not just calculate the 'maximal' for one mode as in the above. Although for convergence of these two models, we approximate to the same degree, *i.e.*, Δf_α is calculated for D and A equidistant from the odorant, though this time the centre of mass, is not necessarily the centre of dipole. Again we model within a cylinder, as cylindrical coordinates prove convenient in modeling, and also

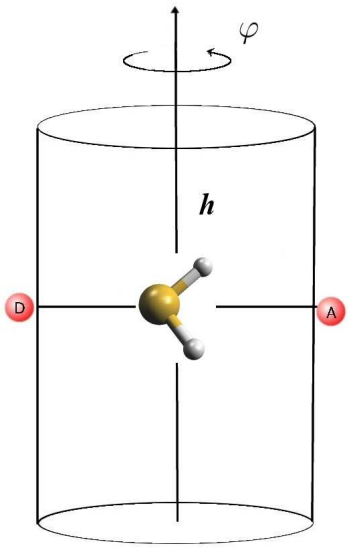


Figure 8.7: The model of odorant within the TM domain (a simple cylinder) with an electron at the donor (or acceptor position).

again the cylinder represents the area the 7 TMs encircle. Calculations of the HR factor can then be done for the odorant rotating about and up the height axis (the z-axes), and rotation is given by $\varphi = theta \times \pi/12$, where we refer to $theta$ as the integer increment we rotate the odorant by and the height goes up to 5Å.

8.6.2 The method: determination of atomic partial charges "q"

The partial charge "q_I" can be found from electrostatic potential fitting (ESP) [69]:

$$V_{MEP}(r) = \sum_{I}^{nuclei} \frac{Z_I}{|r - r_I|} - \sum_{r,s} P_{rs} \int \varphi_r(r') \frac{1}{|r - r'|} \varphi_s(r') dr'. \tag{8.39}$$

Here r and s run over the atomic orbitals, φ orbitals consists of the basis set and P is the one-electron density matrix. This measures the response of the molecule to a positive or negative test charge. This method is very good at determining surface charges, so should be very useful in assigning charge to planar odorants, but perhaps less so for (less common) odorants where atoms may be disguised within the molecule and thus will be less well represented by this mapping to a surface. The

V_{MEP} is matched and minimized with respect to a monopole expansion [69]:

$$V_{ESP}(r) = \sum_{I}^{nuclei} \frac{q_I}{|r - r_I|}, \tag{8.40}$$

when there is minimal difference between $V_{MEP}(r)$ and $V_{ESP}(r)$ the q_I is determined. We use the 'Merz-Kollman' algorithm in a Gaussian calculation to match the ESP at a high density of points along a cubic grid determined by the van der Waals surface of the molecule, for discussion of this method see Chapter 'Density Functional Theory' [87]. Using Gaussian '03 the partial charges were found using the MK algorithm and the same method and basis set are used for the relaxed atoms. All the other parameters in equation 8.38, can be calculated as usual, the atomic positions are easily found. Thus we obtain our q_I, which is, as before, a partial charge (a fraction of electron charge e) though found through different means.

8.6.3 The assumptions and accuracy

The important assumptions used here are:

1. Both equation 8.38 and 8.40 do not take into account the polarization in response to the moving charge.

2. There is nothing intelligent about the way the odorant is placed and rotated between the donor and acceptor- it is placed at the origin where its centre of mass falls as determined by the calculation. Knowledge of a 'tether' or way in which odorants may bind would eliminate the many possible degrees of freedom the odorant may take within the binding domain. To compare the two models however, we must seek, using this model, configurations that appear to 'maximize' the HR factor, in the way the above model does.

Further, as mentioned above, the method of ESP fitting is very good for surfaces, but may not be quite right for application here. There are two main pitfalls to this method in determining accurate partial charge q_I; they are 1) the problem of conformation and 2) of atoms not near the van der Waals edge, for the reasons of approximating to a surface as mentioned in the above section. As partial charges depend on geometrical spacing, freely rotating groups may not be well represented,

so we have here a very static picture, sensitive to the overall orientation of charge, when we are not sure what the important orientations may be.

8.7 Application of HR factor models to small molecule detection.

8.7.1 Boranes and sulphur

Decaborane and hydrogen sulphide (H_2S) both smell sulphuraceous, even though they are 10's of atoms different in structure and decaborane contains no sulphur. Structurally related to decaborane with similar icosahedra structure are the *meta*, *ortha* and *para* carborane isomers which smell camphoraceous[2]. If decaborane and hydrogen sulphide share similar spectra, and the carboranes do not, then we have a good test of Turin's proposal.

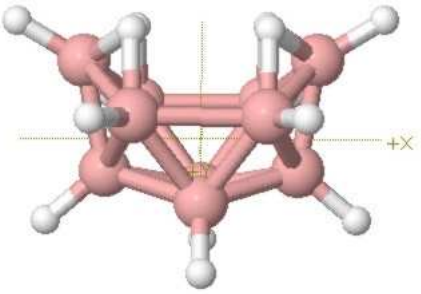


Figure 8.8: Decaborane $(B_{10}H_{14})$

[2]In 'Structure Odour Relations' Turin (2003) comments that the *o*-carborane is *slightly* sulphuraceous, compared to the other isomers [6].

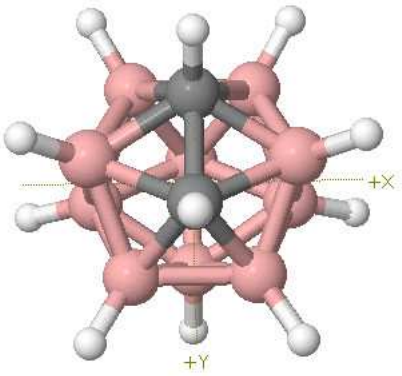


Figure 8.10: *o*-carborane ($o - C_2B_{10}H_{12}$)

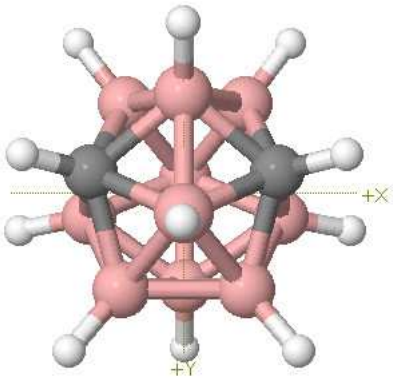


Figure 8.9: *m*-carborane ($m - C_2B_{10}H_{12}$)

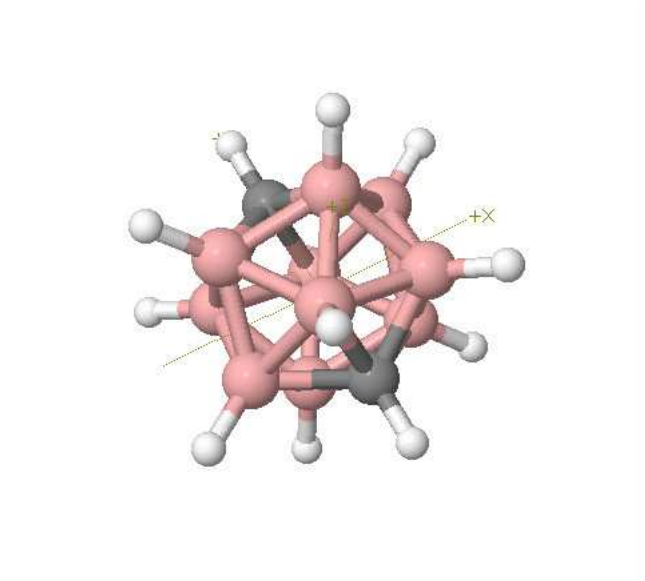


Figure 8.11: *p*-carborane ($p - C_2B_{10}H_{12}$)

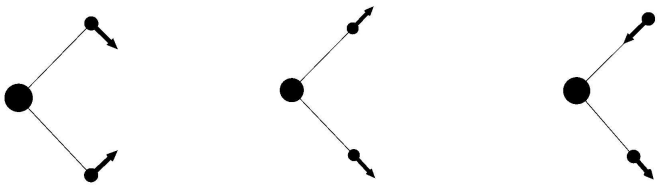


Figure 8.12: Normal modes of hydrogen sulphide; $1A_1$ the 'bending' mode, $2A_1$ the symmetric stretch' and $1B_2$ the 'anti-symmetric stretch', these are labelled to represent their C_{2v} point group symmetry (the lower two modes have A_1 symmetry and the highest frequency mode has B_2 symmetry). Note the first two result in dipole moment vectors parallel to the bisector of the apex angle, whereas the third, highest frequency mode, results in a dipole moment vector perpendicular to the bisector of the apex angle.

	Theory*		Theory$	
Mode	Waveno.	Intensity	Waveno.	Intensity
$1A_1$	1218	1.7	1223	4.9
$2A_1$	2711	0.8	2689	6.7
$1B_2$	2731	2.0	2709	8.6

Figure 8.13: The InfraRed spectrum of hydrogen sulphide as found by theoretical methods. *denotes results taken from Li and Hamilton [103], who calculated these values using the Becke three-parameter hybrid exchange functional with the Perdew gradient corrected correlation functional combined with a double zeta basis set (B3P86/SDD**). $ denotes calculations I performed for hydrogen sulphide with the usual method of B3LYP/6-31G**. Units are vibrational mode in cm^{-1} and the IR intensity of absorption in km/mol.

The IR spectrum has been measured experimentally, see Nielsen and Barker [104], Sprague and Nielson [105], for detail. They find the bands of H_2S are observed, in order of intensity at 1260cm^{-1}($1A_1$), 2685cm^{-1}($1B_2$) and 2615cm^{-1} ($2A_1$). Note the latter is only observed in Raman spectroscopy (vibrations are symmetric about the centre of symmetry). Note in the theory we have intensities ordered $I(2A_1) < I(1A_1) < I(1B_2)$ versus $I(1A_1) < I(2A_1) < I(1B_2)$. Considering the simplicity of hydrogen sulphide (see figure 8.12), this is a good example for testing, as there are only 3 modes, and we conjecture that the vibration at 2615cm^{-1}($2A_1$) is the one responsible for sulphuraceousness, given it is the higher frequency mode that is experimentally detectable. Hereafter we say ~2600cm^{-1}, as this is the right region and an exact value (see table 8.13) is contestable. The mode investigated for H_2S will always be the symmetric stretch in this analysis. Thus we shall use hydrogen sulphide, decaborane, *m*-, *o*- and *p*-carborane as test molecules for the Huang-Rhys factor determination of 'odorant spectra', assuming a ~2600cm^{-1} tuned receptor. We will test both models that examine forces.

IR absorbance model

We use the method outlined above to investigate the 'odorant spectra' similarity for the aforementioned molecules. Results are shown in figure 8.14 which examines the whole range of vibrations and shows that HR couplings are strikingly stronger overall for decaborane as compared to the carboranes, but stronger even compared to hydrogen sulphide. Further, it has interim modes within 1500-2100cm^{-1}, where

the others do not. This suggests the possibility that decaborane may activate a ~2600cm^{-1} receptor additively, note the strong couplings at ~1600cm^{-1}and ~1000cm^{-1}.

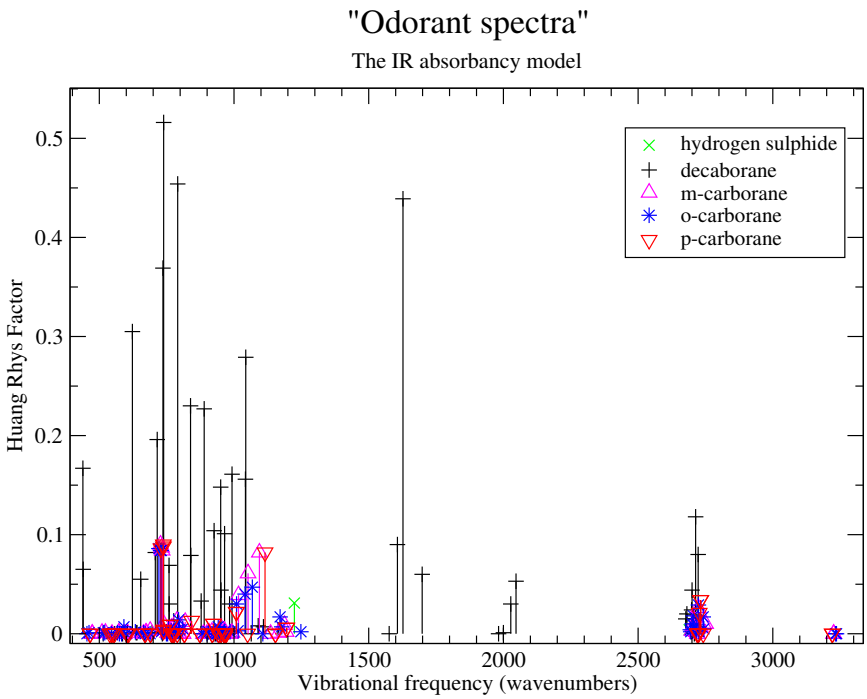


Figure 8.14: IR HR factors for the whole detectable range (>500cm^{-1}), for the test molecules. Decaborane, notably, has stronger couplings for most modes compared to the others, even compared to hydrogen sulphide.

Looking more closely at the "sulphur" region of around 2500cm^{-1}-2700cm^{-1}, see figure 8.15, we can see more clearly the hydrogen sulphide couplings and the magnitudes of S are given for hydrogen sulphide and decaborane. The couplings for hydrogen sulphide are relatively small and considering the proximity and strength of the carborane modes, these spectra would suggest *all* the molecules have a fair chance at smelling sulphuraceous. The carborane modes are shifted more towards the higher frequency end, but only by a couple of wavenumbers (*o*-carborane is closest to the SH stretch in sulphur, which is interesting as it is the one averred to be slightly sulphuraceous and less camphoraceous [6]). The dipole moment vector of the modes with respect to D and A *at one orientation* for the boranes should really be considered given, and some symmetric vibrations are unlikely to be observed, as for the 2615cm^{-1} ($2A_1$) vibration in H_2S . Perhaps some of these modes are redun-

dant. If this model is reliable then the results suggest that the receptor is more sensitive to the position of the vibrational mode in frequency than to the strength of the coupling.

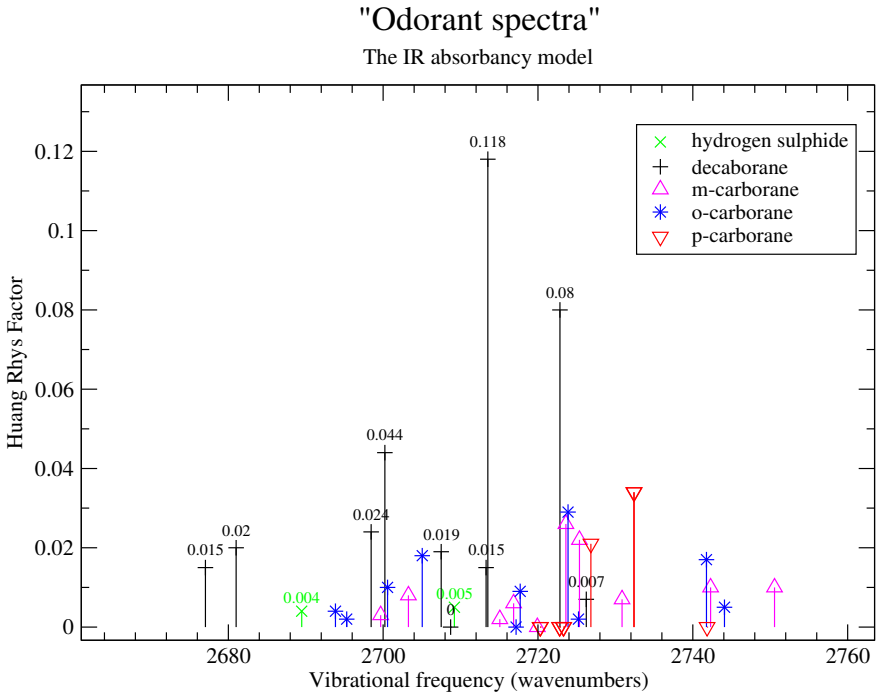


Figure 8.15: IR HR factors for the "sulphur" range (~2600cm^{-1}), for the test molecules. A closer examination of this range, from the full spectrum allows us the see the hydrogen sulphide couplings, which are very small compared to the decaborane couplings.

Point charge model

The point charge model, allows for orientation variation, at a putative 'binding site', thus we can allow for fluctuations of a molecule. As a simple example see figure 8.16 which shows how a HR factor can vary with orientation within the 'binding site' for hydrogen sulphide. By symmetry the HR factor values for H_2S repeat every $6\pi/12$, (theta=6). An initial value of $\vec{r}_D = \vec{r}_A = 3$Å was chosen as the closest fit for hydrogen sulphide because this distance would not involve any covalent bonding (around 0.96Å) with the odorant but isn't too far from the hydrogen bonding limit (around 1.97Å). It is found that within 3-6 Angstrom hydrogen sulphide has

sufficient coupling (for at least theta 0-3) from between 0.01 to 0.10. It might be expected then, that decaborane would show coupling within these limits at about 2600cm^{-1} frequency and *not* the carboranes to explain the sulphuraceous character of the former and the absence in the later.

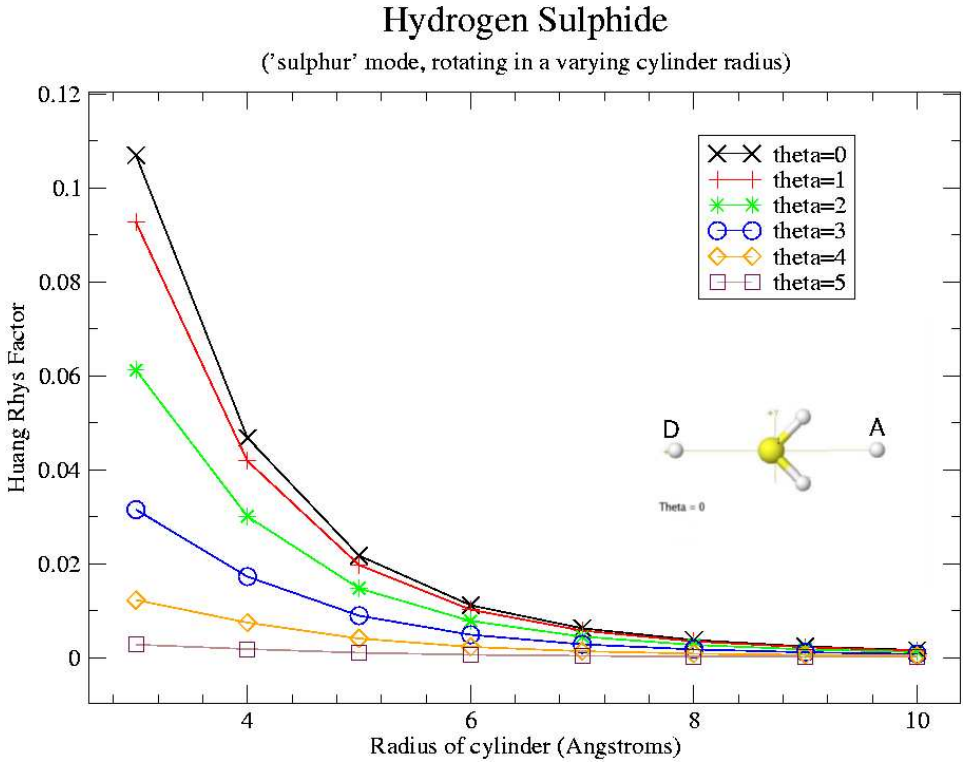


Figure 8.16: HR factors for hydrogen sulphide rotating equidistant from D and A at various radius'. These are the couplings for mode $\sim 2600cm^{-1}$, the 'sulphur' mode. Hydrogen sulphide is shown, inset in the graph, to show the starting position where theta=0 and the radius is 3Å.

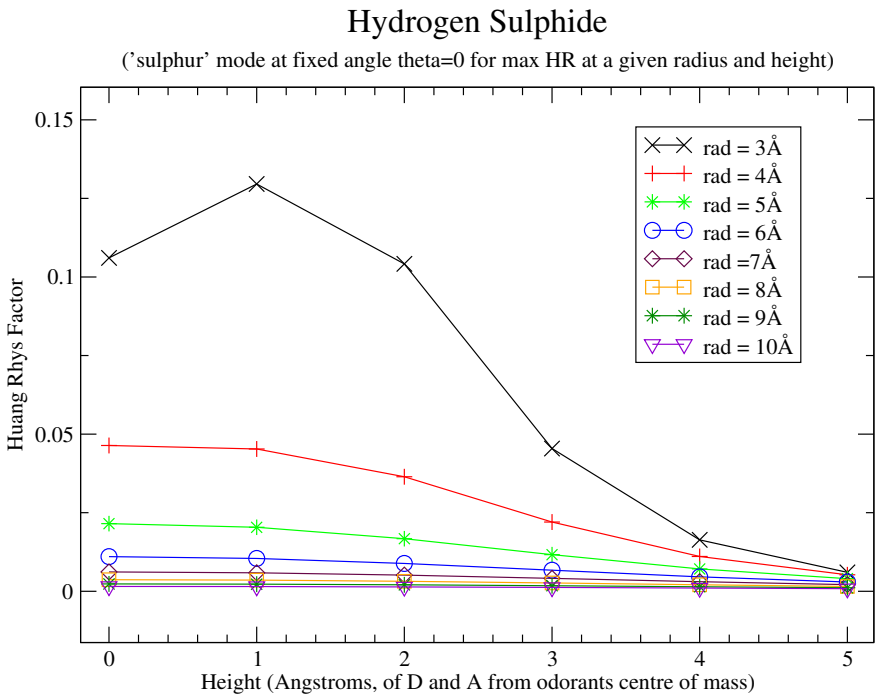


Figure 8.17: the HR factors for Hydrogen Sulphide at optimal theta, equidistant from D and A and at various heights and radius.These are the couplings for mode $\sim 2600cm^{-1}$, the 'sulphur' mode.

Figure 8.17 shows an optimal position for a maximum HR factor at height = 1Å and radius= 3Å. Using this model, if we assume that the "sulphur" detecting receptor is designed for hydrogen sulphide and so is designed to maximize this HR factor, we fit the other molecules into this accordingly. Problem is, given the boranes are much larger, there will be fewer configurations where the molecules can fit within this criterion without the donor and acceptor unit being too close to the rest of the molecule. We must assume then, that at the binding site, different configurations occur, but that they do not upset the D-A splittings, *i.e.*, the D-A splitting are separate units not too sensitive to binding site interactions. We only look at the 'sulphur' region for differences in spectra, since this is the important region.

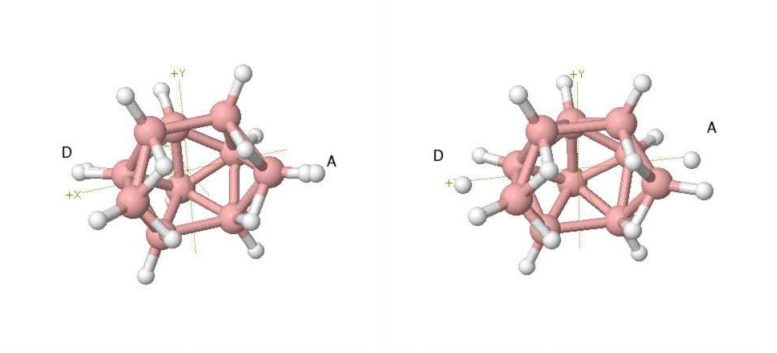


Figure 8.18: Decaboranes both point charge D and A at 3Å, and height from origin 0Å. On the left the angle of rotation φ= $4\pi/12$ (theta-4), on right the angle of rotation $\varphi = 6\pi/12$ (theta=6).

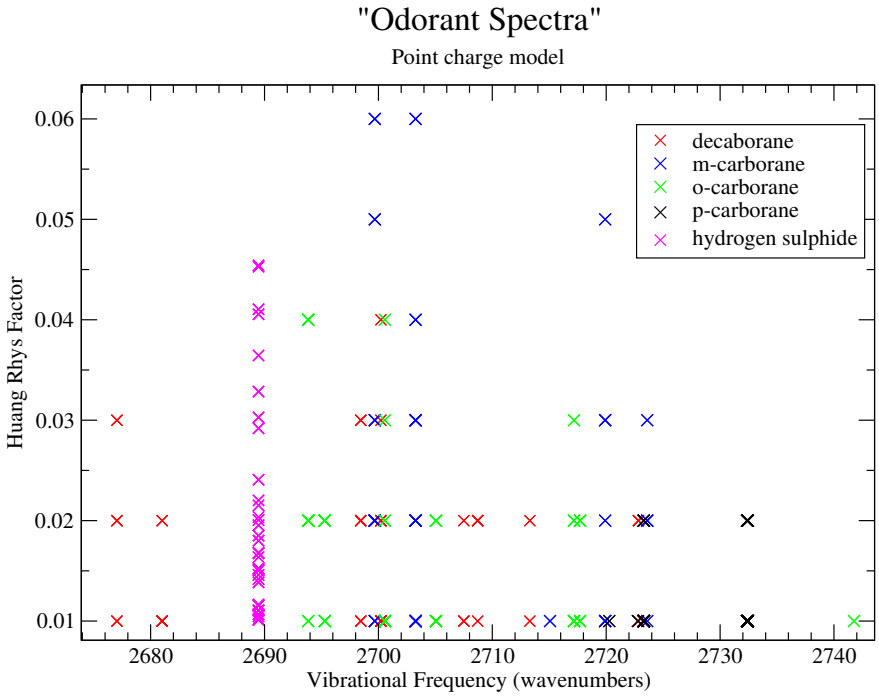


Figure 8.19: Point charge model HR factors for the “sulphur” range (~2600cm^{-1}) for the test molecules. Coupling data is shown for instances where HR factor > 0.01, discounting values for combinations where the radius = 3Å and the height up the z-axes is < 3Å, and where the radius is 4Å and the height is 0Å, as this creates a ’binding site’ too small.

See the results in figure 8.19. This spectrum excludes sites which are probably too small (where rad =3Å, height < 3Å and rad=4Å, height = 0Å) and show results for the "sulphur" region for all the molecules at various orientations against D and A. All modes shown have couplings greater than 0.01 H_2S is shown also for values within the same radius and height constraints as the boranes, assuming they make similar interactions at the binding site. Analyzing this spectrum it might be predicted that *m*-carborane smells sulphuraceous, because of its stronger coupling, unless the tuning of the H_2S receptor has a threshold at about 2693 cm^{-1} (again, *o*-carborane is slightly sulphuraceous). A test of deuterated hydrogen sulphide would be useful: if it doesn't smell of sulphur then we could discern a "cut-off" value for the sulphur receptor. Note these results differ quite a lot from the IR model: the decaborane does not stand out with stronger couplings, and the hydrogen sulphide is not comparatively very weak. What figure 8.19 indicates is that what might be more important than strength, is the *frequency* of strong couplings, of which hydrogen sulphide has ubiquitous couplings compared to the others. However, decaborane does not follow suit, for these conditions.

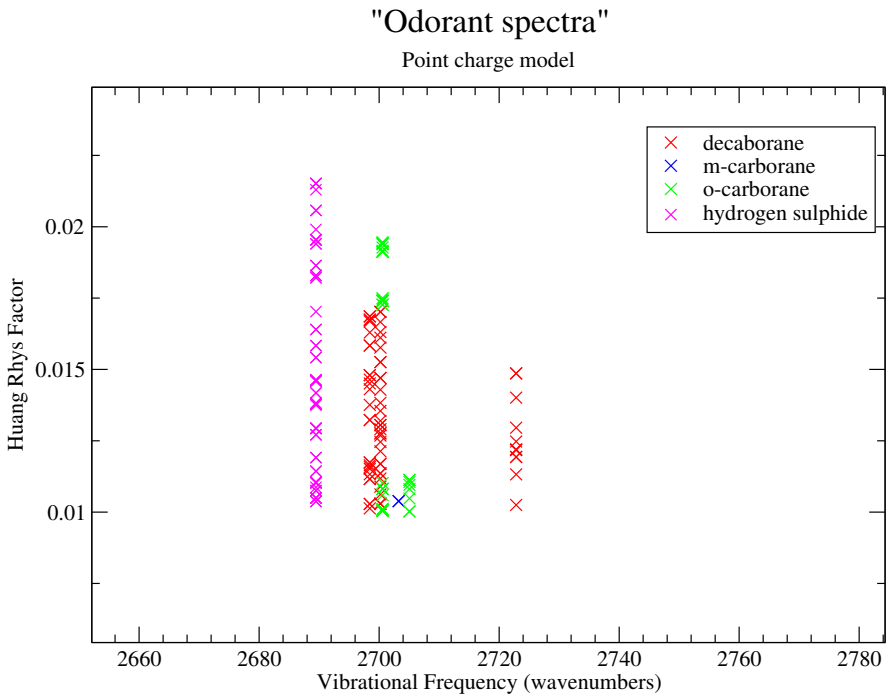


Figure 8.20: Point charge model HR factors for the "sulphur" range (~2600cm^{-1}) for the test molecules. Coupling data is shown for instances where HR factor > 0.01 the radius is 5Å and there are varying heights of D and A.

Figure 8.20, however, shows results for when the donor and acceptor are equidistant from the origin of odorant and equal heights. Figure 8.19 show instances where the radius is 5Å (4Å is probably too small) and the HR factors are calculated for combinations of donor and acceptor heights (D =0-5 and A =0-5 in increments of 1Å), only HR factors of >0.01 are plotted. In this particular scenario of set radius but differing heights *p*-carborane doesn't feature at all, and the results become more reminiscent of those based on IR, compare to figure 8.15. Refer back to figure 8.20, *o*-carborane has the strongest coupling but decaborane appears a lot more frequently. Further, there isn't a particular donor and acceptor position (for example D= 0 Å and A = 4 Å) in which *o*-carborane *does not* have a higher HR value than decaborane. However, if we stipulate a binding site cavity at this radius decaborane and H_2S stand out because they both have strong and frequent couplings compared to the carboranes.

Discussion

Neither model shows a striking spectral similarity for decaborane and hydogen sulphide. Probably the most convincing 'odorant spectra' is seen in figure 8.20, where we have chosen an overall configuration for the D-odorant-A unit that maximizes couplings for decaborane and hydrogen sulphide but not for the others. This model entails that D and A are separated by 10Å and that the heights of D and A may vary (without upsetting the D-A splitting). This suggests that if the odorant molecule M fluctuates around the binding site, then there are plenty more *opportunities* for decaborane and hydrogen sulphide to provide enough electron-phonon coupling as compared to the carborane isomers. The overall geometry dependence of these models (in particular the point charge model) has illustrated how useful determination of D and A sites, with relation to the binding site, would be. We also point out, that the frequency of a good S factor, considering various orientations at the site, may in fact be more important than less frequent, but stronger signals. This in turn points to the relevance of symmetry in the molecules. The very symmetric odorants, may advantageously have more orientations that provide a significant S factor. For the isomers, symmetry operations are 20 for the *para* position and 4 for the *meta* and *ortha*, see figures 8.8, 8.9, 8.10 and 8.11. Decaborane, *p*-carborane and hydrogen sulphide are highly symmetric compared to the remaining two molecules. The others have S values around the right frequency (*p*-carborane vibrates at the higher end of the spectrum). This may mean that over an average time it is important to consider that the receptor will "see" similar orientations, and it may be important to weight the results according to how frequently the odorant exists in a position conducive to tunneling. The more often the odorant hits a reasonable HR value at a certain orientation, the more likely to evoke a 'sulphur' signal.

One of Turin's observations is not just the correlation between odour and frequency, but that there is a change in intensity with a change in the magnitude of the charges on the odorant. Now one of the reasons the boranes are interesting is because they are "electron deficient", that is there are more atomic bonds (H=1, B=5) than valence electrons in the constituent atoms. Decaborane has B-H-B bridges that are asymmetric: for $B_{10}H_{14}$ the B-H (bridge)'s are 1.30 Å and 1.35 Å and the B-B bonds are long at 1.97Å. As a consequence decaborane has much more pronounced partial charges, thus bigger HR factors and better IET efficiency, which seems to indicate IETS efficiency is reflected in an intensity effect. It would be an

interesting test to see if inhaling more of the carboranes in parts per billion (ppb) result in the gradual introduction of sulphur character.

Which model?

Note that we can either approximate D or A as infinite sheets of plane charge or as point charges at a site D or A, see section 8.3, but we expect the results for one charge distribution to converge. However, the resulting spectra above for the test molecules, look quite different.

	IR absorbance model	point charge model
3Å	0.004	0.107
4Å	0.001	0.046
5Å	0.001	0.022

Table 8.4: A table to show the HR factors for the symmetric stretch mode in Hydrogen sulphide- the IR absorbance model versus point charge model. They are compared for various cylinder radii.

To examine the relation of the 2 models see table 8.4 , which gives, for hydrogen sulphide, the HR factors at given radii of the cavites for each model. This is just for symmetric stretch. At the orientation that maximizes the change in force for the 'sulphur' / symmetric stretch mode (as in figure 8.16, relative to D and A) a consequence is that the antisymmetric stretch is undetected at this alignment. This point highlights a big difference across models: the average change in dipole for the HR factor assumes a radius for D and A, but also an alignment that is able to detect the change in dipole, which can be variable for each mode. In which case we are not comparing HR factors at the same orientation relative to D and A because the IR absorbance model assumes the best alignment, whatever the mode. However, even when we compare for the same orientation as in table 8.4, the results do not agree. The Huang-Rhys factors from the IR absorbancy approximation are a couple orders of magnitude smaller than the Huang-Rhys factors derived from the point charge model. The table indicates the strong dependence on the radius of the cavity, but it does not suggest good agreement between models.

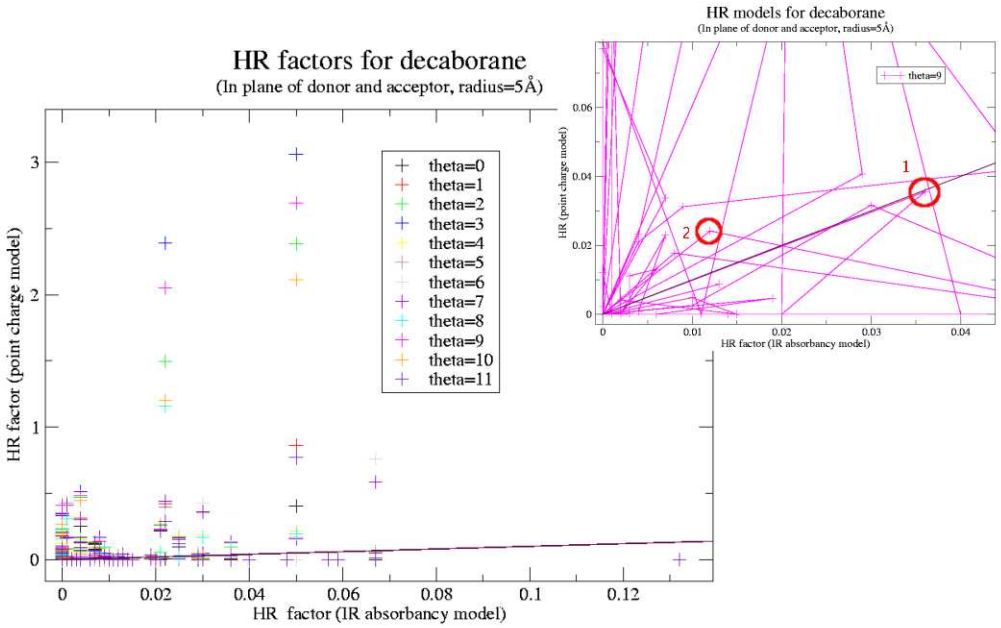


Figure 8.21: Decaborane IR absorbance model HR factors versus point charge model HR factors. Inset graph zooms into $x = y$ region near the origin. Theta =9 is an orientation at which a couple of HR factors agree reasonably, they are circled and noted, see body of text. Order of modes are connected by lines.

As 2 points (the first two modes of H_2S) do not a straight line make, a plot was forgone for the hydrogen sulphide example. However, in figure 8.21 we can plot many values for decaborane, especially as we do not know what the 'sulphur' mode is and we sample a region. We plot HR factor versus HR factor for all the modes in decaborane, for various orientations for the values up to theta=11 (after this by symmetry the values repeat themselves). Compare to $x = y$; again there

is little agreement between the two models, the inset graph zooms into the $x = y$, the order of modes are connected by lines, and we can see the best agreement for the factors occurs at the orientation theta=9. Interestingly, if we examine point 1, this corresponds to a mode at 1043cm^{-1}, where $S_{IR} = 0.03 \approx S_q = 0.036$, and point 2, corresponds to a point relatively close to $x = y$, that corresponds to a mode at 1606cm^{-1} that has $S_{IR} = 0.012$ and $S_q = 0.024$, at the orientation theta=9.

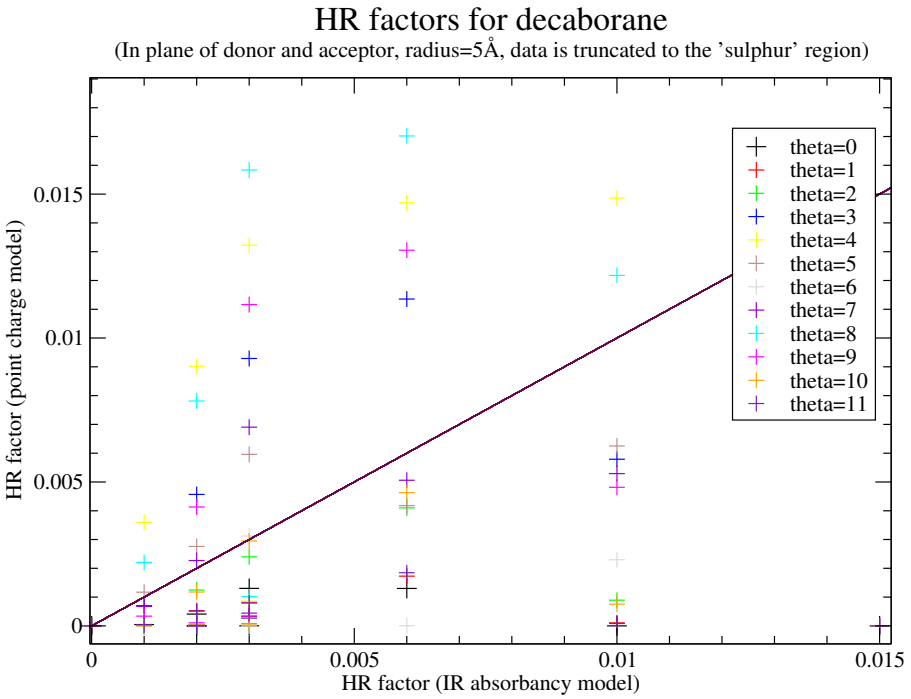


Figure 8.22: Decaborane IR HR factor versus point charge HR factor for 'sulphur' range. The straight line is $x = y$ for comparison.

Plots8.21 and 8.22 are very busy because we consider many orientations with respect to the point charge model. Considering all orientations, it is more likely more important to seek a theta which has the most points, on a straight line, that corresponds to a *strong* coupling, rather than $x = y$. Presumably this indicates a favourable orientation for decaborane in the binding site, at least for the mode we are interested in. Looking in particular at those with strong point charge HR factor (>0.01) for the range around ~2600cm^{-1} for the modes, see figure 8.22, it looks as though for $\varphi = 8\pi/12$ (theta = 8) in Figure 8.22, there is maximum HR factor and a

possible straight line of best fit, this orientation is shown in figure 8.23

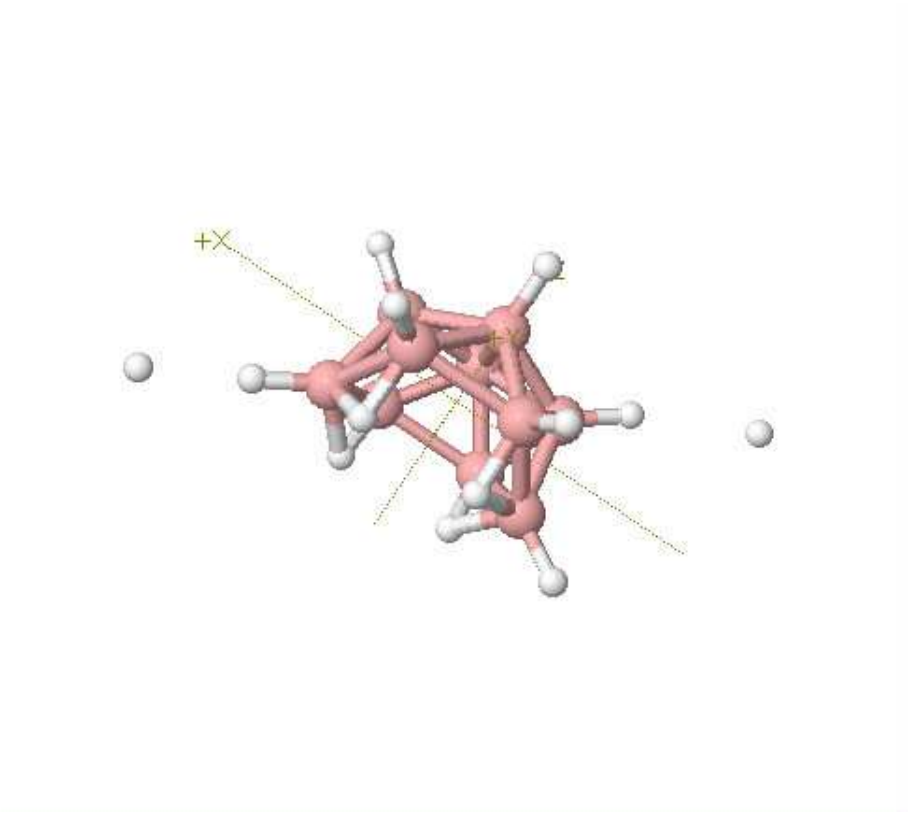


Figure 8.23: Decaborane, with point charges at D and A.

For the borane molecules and sulphur we do not have very good agreement across models. If there was more reassuring evidence, the carboranes could be positioned similarly to decaborane, in figure 8.23 and equivalent graphs plotted for smaller increments around theta=8. There may be orientations that fit a straight line better, and the results may be more striking compared to decaborane and hydrogen sulphide at a quite fixed orientation. The next couple of sections then, look to see if the models agree for simpler systems, and to see where the divergence is coming from, hopefully then we will better be able to evaluate which, if either model, is accurate enough to describe an 'odorant spectrum'.

Diatomic	$\left(\frac{\partial\mu}{\partial Q}\right)_{IR}$	$\left(\frac{\partial\mu}{\partial Q}\right)_{q}$
Hydrogen chloride	0.1359	0.1461
Lithium hydride	0.3503	0.3718
Carbon monoxide	0.2638	-0.9642

Table 8.5: A table to show the change in dipole moment with normal mode, calculated in one instance analytically (leftmost column) and the other numerically (rightmost column). The units are eÅ/Å$\sqrt{amu}$. Calculations both use method and basis set B3LYP/6-31G**.

8.8 Comparisons: determining partial charges

In the point charge model we rely heavily on an accurate value for the partial charge on the atoms, and this may be where problems arise. It is interesting to see, for the simplest system, whether the analytical result from Gaussian $\frac{\partial\mu}{\partial Q_\alpha}$ converges to a numerical calculation that is based on charges to calculate the change in dipole moment. In the IR absorbance model, we approximate the odorant to a simple dipole, extracting a partial charge from the change in dipole moment that we get from IR absorbance. What we want to determine, is whether we can derive similar values from effective atomic charges q:

$$\left(\frac{\partial\mu}{\partial Q_\alpha}\right)_{IR} = \left(\frac{\partial\mu}{\partial Q_\alpha}\right)_{q}.$$

So for three test molecules we calculate the dipole $\mu(\Delta r)$ at various displacements and calculate Q from rearrangement of equation 8.20 , plots of these values (see section 8.2) allow us to obtain a gradient and so $\left(\frac{\partial\mu}{\partial Q}\right)_{q}$. Calculations for population analysis are described in the chapter 'Density Functional Theory' (I use B3LYP/6-31G**). So table 8.5 compares the results below. $\left(\frac{\partial\mu}{\partial Q}\right)_{IR}$ is simply extracted from the Gaussian calculation of IR absorbance (also calculated using B3LYP/6-31G**), noting for comparisons, it is easiest to convert to eÅand Gaussian uses Debyes, $1D = 0.208e$Å.

Table 8.5 shows a decent agreement for the first two test molecules, but the results wildly deviate for carbon monoxide; the gradient for the numerical value is negative. The strange result likely comes from problems in determining the Mulliken atomic charge. There is quite a drastic difference in bonding for these di-

Diatomic	q_{MK}	q_{dipole}
Hydrogen chloride	0.254	0.232
Lithium hydride	0.710	0.723
Carbon monoxide	0.001	0.011

Table 8.6: A table to show partial charges derived from the MK algorithm (leftmost column) and from the dipole moment calculated from Mulliken charges. The units are fractions of electronic charge e. Calculations both use method and basis set B3LYP/6-31G**.

atomics; contrast the triple covalent bonds in CO with ionic LiH and HCl. Mulliken charges overestimate the covalent character of a bond, and it is unlikely this one method is equally reliable for the different bonding in this selection of molecules.

The numerical derivation of $\left(\frac{\partial \mu}{\partial Q}\right)_q$ can be compared to the partial charges used in the point charge model. It is interesting to see if a partial charge derived from the dipole moment, from $q = \mu/r$, as above, would agree with the MK algorithm for partial charge, described in section 8.6.2.

Again, see table 8.6, there is an interesting divergence for carbon monoxide, the other two diatomics have partial charges that agree well, whereas the last example has a difference of a factor of ten across the two models. Table 8.6 exemplifies a previously made issue (see Chapter 'Density Functional Theory') that charge is not unambiguously defined. We especially have problems in the agreement of q_{MK} and q_{dipole} for CO; but this is not exceptional. The description of charge is very much dependent on your choice of algorithm. Different algorithms are used for charge definition; some based on electronic orbitals (such as MK, Mulliken and ESP derived from the one electron density matrix), others based on charge density (Bader, Hirshfield). See Haiduke and Bruns, for example, who compare results for AIM (atom-in-molecule) charges with experimentally derived fluxes, they get 0.240, 0.909, and 1.101 for HCl, LiH and CO respectively, where their assignment of charge sign agree with mine (negative partial charge is on Cl, Li and O for HCl, LiH and CO respectively) [106]. Though the agreement of sign in the charge is there, again we see a big difference in magnitude for CO. Haiduke and Bruns, also use high level theory and polarized basis sets so this exemplifies how different algorithms (but same method/basis set) for charge fitting can give various results. Keeping the algorithm the same, however, and contrasting with different *methods and basis sets*

according to the chemistry of the molecule can also be inappropriate, see Larin *et al* for example [107]. Larin *et al*'s results show (see figure 8.24) that for CO the dipole moment can change in sign: contrast BH and HLYP with B3P86 for example, and that the Mulliken atomic charge can differ by *several* orders of magnitude across methods.

TABLE I
Properties of the XO molecule, X = C or N, optimized with different basis sets: X–O equilibrium distance ρ_e (Å), dipole moment μ (1 a.u. = 1 ea_0), Mulliken atomic charge $q(X) = -q(O)$ (1 a.u. = 1 e), and harmonic vibrational frequency ω_e (cm^{-1}).

Basis set	Method	ρ_e	μ	$-q(X)$	ω_e[a]
CO					
cc-pVTZ	BHandHLYP	1.1134	0.0098	0.003	2258
	B3PW91	1.1260	0.0582	0.019	2156
	B3P86	1.1253	0.0540	0.019	2160
	MP2	1.1380	0.1154	0.076	2125
[6s4p4d2f][b]	BHandHLYP	1.1183	−0.0113	0.163	2259
	B3PW91	1.1312	0.0379	0.254	2154
	B3P86	1.1306	0.0346	0.257	2159
	MP2	1.1387	0.1108	1.193	2124
Experiment		1.13[c]	0.0481[d]	—	2170[c]
NO					
cc-pVTZ	B3PW91	1.1435	0.0495	0.0014	1945
	B3P86	1.1429	0.0488	0.0005	1948
Experiment		1.149[c]	0.0622[e]	—	1904[c]

[a] Harmonic DFT frequencies are scaled by 0.9716 as in Ref. [13]; no scaling for MP2.
[b] Basis set from Ref. [14].
[c] Ref. [15].
[d] Ref. [16].
[e] Ref. [17].

Figure 8.24: A table to show the differences method and basis sets make to the determination in charge and dipole moment for CO, from Larin *et al* [107].

8.9 Conclusions

The objective of this chapter was to generate 'odorant spectra' that test Turin's ideas, and hopefully provide a predictive algorithm. However, the two models based on forces do not agree very well and it appears the main problem is the definition of type of charge/field caused by the presence of the electron at site D and A, and in turn the proper calculation of a force change. This tells us it's really a large electronic structure calculation problem. In the point charge model we approximate a continuous distribution by point charges, which shouldn't be too bad because of i) the averaging effect of Coulomb interactions, which vary slowly over $1/r$, where r is distance from charge and ii) the tendency of electrons to follow the nuclei. Although aside from this there are other issues to do with the models independently, which may account for the overall sources of error and disagreement. The IR absorbance model is reliant very much on calculated values which do not seem convincing. Refer back to table 8.2. In the theory we have $I(2A_1) < I(1A_1) < I(1B_2)$ and $I(1A_1) <$

$I(2A_1) < I(1B_2)$. We see varying orders of magnitude for the intensity according to basis set but also even the relative values of IR intensity disagree for different models, which is a problem that the usual scaling factor across from experiment to theory will not fix. Thus absolute values are unreliable, and so too even relative values, which is a problem since we are so reliant on such a factor. The point charge model, also, is reliant on unconvincing calculations, where the definition of charge is notoriously contentious.

Analysis for diatomics, show indeed that for the simplest picture, and the same charge distribution, the two models (point charge and IR absorbance) should agree. Simple algebra gives a force F_D at D (and similarly for A):

$$F_D = \frac{qe^2}{4\pi\varepsilon_0\left(R_D^2 - h^2\right)} \frac{R_D}{\left(R_D^2 - h^2\right)^{1/2}}, \tag{8.41}$$

Which we simplify even further for $h = 0$, as in the IR absorbance model, and so we eliminate the cos theta component and just have $\Delta f = 2\frac{qe^2}{4\pi\varepsilon_0}\frac{1}{|R|^2}cos\theta = 2\frac{qe^2}{4\pi\varepsilon_0 R_D^2}$, where $R = R_D = -R_A$ see figure 8.25.

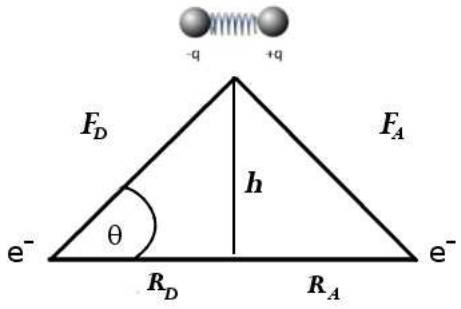


Figure 8.25: Figure to show the change in force on an oscillating dipole as the electron moves from point D to A.

Results 8.16 and 8.17 approximate at least qualitatively to this approximation. The fall of F as $1/R^3$ at constant θ is seen in figure 8.17 whereas the fall off as cosθ at constant R is seen in figure 8.16. The couplings are sensitive to R, and this is apparent in all models, but it is first seen in the parabola model. Calculations are for distances 3Å away from the point charge (so Cl, or H, depending on which

	HF/STO-3G		B3LYP/6-31G**	
Mode	Frequencies (cm^{-1})	IR absorbance (km/mol)	Frequencies (cm^{-1})	IR absorbance (km/mol)
$1A_1$	1609	23.1	1223	4.9
$2A_1$	3274	17.2	2689	6.7
$1B_2$	3322	10.5	2709	8.6

Table 8.7: A table to contrast the results for HF/STO-3G with B3LYP/6-31G** for H_2S.

site the electron resides) of the dipole/diatomic. This was because both the IR absorbance model and the point charge model converge to the same S value using this assumption. However, there are various ways for better convergence to the other 2 models, in particular for polyatomics, where it might be better to approximate the electron 3Å away from the *centre of mass*, or the *centre of dipole* where we cannot assume the odorant is a point dipole anymore. So it may be useful to go back to the, relatively successful, first model to resolve this issue. The parabola model has the added advantage of overcoming the biggest issue which is that; even with all the simplifications for a diatomic, the definition of charge (and so then the forces) is ill-defined, and so will vary too much across models.

The problem with partial charge assignment is that there are no observables, and so, no quantum mechanical formalism. Probably the best way to improve and test further the idea of an 'odorant spectrum', is to investigate the relation for $S \propto$ to the IR absorbance and further investigate for simpler systems, with the theoretical absorbancies evaluated against experimental absorption spectrum (so at least relative intensities must agree, and we have some reassurance from observables). Another important step is to decide the most realistic (based on determination of the D and A) distribution of charge at these units. Further a different choice of test molecules may be better; hydrogen sulphide, a seemingly simple odorant, perhaps was not the best choice. Sulphur is a third row element and so the molecules IR absorbance may be better represented by a minimal basis set such as STO-3G which has an advantage in that it describes cheaply third row atoms with $1s$, $2s$, $2p_x$, $2p_y$, $2p_z$, $3s$, $3p_x$, $3p_y$, $3p_z$. For example compare HF/STO-3G with B3LYP/6-31G** for H_2S; though the frequencies are more accurate for the latter choice the IR intensities may be more reasonable in the former.

See table 8.7, the IR intensities can differ in methods by orders of magnitude (sometimes several), so we can see easily were the loss in agreement occurs, depending on what value is closer to a true observable.

Chapter 9

Chirality and conformations of odorants

We see in the Chapter 'A physical picture' that the vibration frequencies are identical for mirror-image odorant molecules. However, it is famously known that many mirror-image, "chiral" (enantiomer) molecules smell different in their right and left handed forms. Chiral discrimination is hardly unique to olfaction. For example take the effects of cocaine (*1R, 2R, 3S, 5S* is psychoactive while *1S, 2S, 3S, 5R* is inactive), the hormone thyroxine (*R* is inactive, while *S* is active) and the gypsy moth pheromone disparlure (*7R, 8S* activates a response, whereas *7S, 8R* inhibits a response). A distinction is also made in taste: *D*-form amino acids are sweet whereas the *L*-forms are usually tasteless. It is therefore not surprising that the chiral helices of the olfactory G-protein coupled receptors (GPCR's) are also enantioselective.

But how could a "meat" spectroscope tell a difference? This is an alleged refutation of Turin's theory. In actual fact, the discrimination of chiral molecules does not refute either the "lock and key" or the "swipe card" model. It should be emphasized here that in lock and key, the 'right' shape is considered necessary *and* sufficient for signal transduction whereas in the swipe card paradigm shape is necessary *but not* sufficient. Either way the odorant has to adequately make the right contacts with a binding site, but it must also *do* something else. If the "something else" is biological IETS, as we have seen in Chapter 'The olfaction model', the signal is indeed dependent on the relative positions of oscillating atoms and so is enantioselective. Thus the case of chiral odorants makes an interesting study, especially

as there happens to be a wealth of data on this subject; see www.leffingwell.com. In May 2006 this website contained over 400 pairs[1] (>800 odorants) of example enantiomer odorants and their associated smells. By documenting certain features of these examples, such as, the position and number of chiral centres, the presence of 'osmophoric' groups (see section 9.3.2.2) and the type and number of substituent group, a rule emerged for odorants with a 6-membered ring. Analysis shows this rule correlates a conformational flexibility of an odorant with whether it has a certain bioeffect - a difference in perception for its mirror images. We investigate this feature and we see it has interesting implications for the view point of "receptor eye" signal transduction.

9.1 Data analysis

9.1.1 Classifying chiral odorants and a note on subjectivity

It is a must to treat all experimental data on smell with caution. First, there is some subjectivity in how people describe smells, though there is usually broad agreement. Secondly, the character of a smell can depend on the strength of the source, sometimes dramatically. Thirdly, purity is very important: even small traces of some strong-smelling impurity can transform an odour. Even with 98% enantiomeric purity, incorrect results occur if one enantiomer is significantly stronger (e.g. with a 0.0002 ppb threshold): the smell character of the weaker can be easily dominated by a very small fraction of the stronger component. This is especially an issue for enantiomers, since impurities may arise in synthesis, or during storage before use, as a result of reactions that are chirally-sensitive. Such reactions could misleadingly make enantiomers smell different. For these reasons data produced using gas chromatography should be used. There should also be checks that the sample indeed contains just one enantiomer (say at least 99% pure). For trademarked odorants, the in-house quality procedures may suffice.

Throughout this chapter I refer to data published by Leffingwell [42], so this study is based on data, from this website, accurate as of May 2006. All odour descriptors are taken from Leffingwell's site and all odorants are categorized into the

[1]Note sometimes there are *several* pairs of hands, this depends on the number of chiral atoms in the odorant. There are 2^n stereoisomers for 'n' chiral centres. For example if there are two chiral centres there are (*R, R*), (*S, S*), (*R, S*) and (*S, R*) enantiomeric forms.

same groups Leffingwell uses. Like all large databases, the information may be uneven in quality. If we analyse the Leffingwell data, without being too critical about what could well be trace contaminants (that is we ignore small differences in odour character or intensity), then 59% of enantiomer pairs smell the same and 41% different. This predominance of pairs that smell largely the same should be borne in mind when we discuss differences in detail. However, if we demand odour character and intensity be essentially identical then only 5% of enantiomer pairs smell the same. In the interest of scientific rigor 4 classes of enantiomer odorants are defined thus:

- Type 1. The left and right-handed versions smell the same in character, with similar thresholds. Presumably both enantiomers have very similar affinity and efficacy for the same receptors.
- Type 1c (character). The left and right-handed forms smell much the same, but have small, sometimes subtle, differences in character. Clearly, one possibility is that these type 1c odorants should be type 1 but impurity species are involved somehow. Many odorants fall into this category, and it is hard to give an unambigous example.
- Type 1i (intensity). The left and right-handed versions smell the same, but have very different thresholds. Since presumably several different receptors are actuated, but the only difference is in intensity, there is an implication that both affinity and efficacy simply scale on going from one enantiomer to another.
- Type 2. The left and right-handed forms are clearly distinct. Type 2 enantiomer pairs activate different receptors, through different affinity or efficacy or both.

Type 1 or type 2 are thus described scrupulously. For example, if one enantiomer is described as “amber, woody, cedarwood, animal and strong” and the other member of the pair as “woody, camphoraceous, amber, spicy”, there are very strong similarities (woody, amber), but the scents strictly are different, so the pair is classed as type 2. I shall be equally scrupulous in distinguishing type 2 from type 1i or type 1c, since they have their own characteristic features. So, as above, we distinguish

between *strong* similarity (5% type 1, 95% other) and *weak* similarity (types 1, 1c, 1i the same, 59%, and type 2 , 41%). I emphasize this because of the subjectivity objection which is often voiced in olfactory research. This study uses *objective* information, i.e., whilst the descriptions of differences may be subjective, the fact of whether they are the same or not is less so. The schematic representation of these odorant classifications can be seen in figure 9.1.

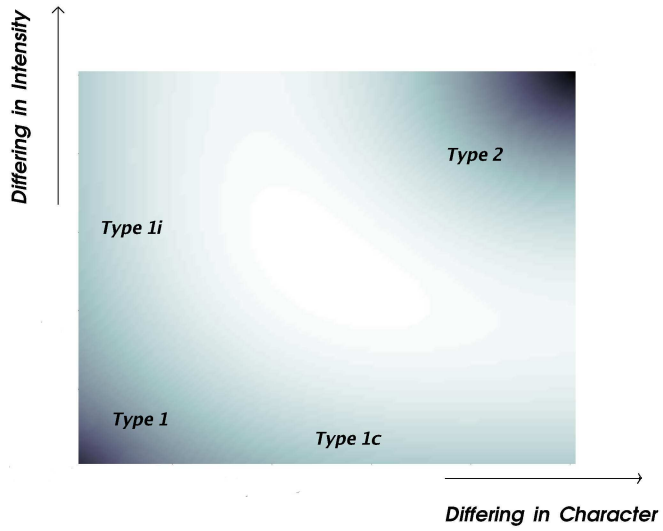


Figure 9.1: An 'odorant square' representing in shades odorant classifications, within the strong regime, where type 1 is 5%. The bottom-left corner represents type 1, the top-right corner type 2. Along the x-axis represents type 1c and along the y-axis represents type 1i. Graduations in shade are meant to indicate areas of ambiguity and show overlap between the four categorizations (type 1, 2, 1i and 1c). Whilst graduated areas are contentious, we note all examples examined in this paper are taken from odorants that lie within a triangle. In other words, where it may be disputable *how* enantiomer pairs may smell different from one another, it is not disputable *whether* they smell the *same* or *whether* they smell *different*.

9.1.2 Receptor combinatorics

9.1.2.1 Pure enantiomers

The brain appears to identify smells through a combinatorial analysis of the intensities of the signals it receives and knowledge of the receptors from which the signals come. Probably the brain can interpret even incomplete signals from the receptors, e.g. if a fraction of the receptors cannot initiate a signal for some reason, just as the eye fills in parts of images that may be missing [108]. We illustrate some possible cases in figures 9.2 to 9.6, to show an apparently random code of receptor actuations in response to enantiomers R and S. Each box in the figures represents a receptor type, and the size of the circle is proportional to the strength of the direct signal to the brain from that receptor.

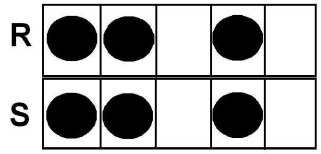


Figure 9.2: Type 1, receptor sees enantiomers as essentially the same. This pair would exist in the black corner of figure 9.1.

Type 1 enantiomers would lead to the situation shown in Fig 9.2. Both R and S give the same signals to the brain. As an example (table 9.5 in this chapter, no. 3), the fenchones smell camphoraceous. (1R, 4S)-Fenchone has an odour threshold of 510 ppb, and that of its antipode is not very different, 350 ppb. The fenchones give the closest recorded example of such a type 1. The natural interpretation is that there is no relevant structure change when these molecules enter the chiral receptor, and likewise no other critical change, e.g., in vibrational frequencies.

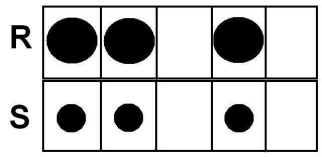


Figure 9.3: Type 1i, receptor sees enantiomers as essentially the same, except in one hand the intensity of the signal across all receptors is greatly reduced. This pair would exist somewhere along the y-axis of figure 9.1.

Figure 9.3 shows what could be either type 1 or type 1i behaviour, in that the signals from one of the enantiomers are weaker. As an example (see Table 9.5, no. 4), the camphor enantiomers both possess the same camphoraceous character, but are readily distinguished by their relative intensities [109].

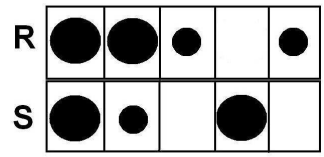


Figure 9.4: Type 1c, receptor sees enantiomers completely differently by virtue of different types of receptors being activated with different affinities. This pair would exist somewhere within the shady area of ambiguity in figure 9.1.

Qualitatively distinct behaviour is shown in figure 9.4, where the enantiomers initiate different signals to the brain. The behaviour is typical of type 2 pairs. Thus, for example (table 9.6 in this chapter, no 3) (4*R*)-carvone smells of spearmint (43 ppb threshold) yet (4*S*)-carvone smells of caraway (600ppb threshold). Nootkatones [110] are another example of 'true' type 2's, which differ both in character and power (see figure 9.30).

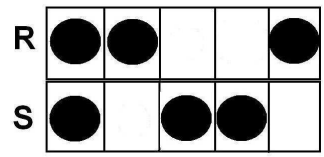


Figure 9.5: Type 2, receptor sees enantiomers differently by virtue of different types of receptors being activated. This pair would exist somewhere along the x-axis of figure 9.1.

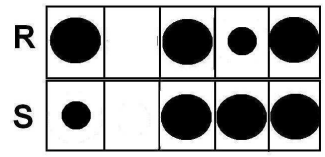


Figure 9.6: Type 1c, receptor sees enantiomers differently by virtue of the same types of receptors being activated but with different affinities. This pair would exist somewhere within the shady area of ambiguity in figure 9.1.

In figure 9.5, members of the hypothetical enantiomer pair activate different receptor types, even though their affinities for their respective receptors are much the same. The brain receives different signals for the members of the pair. For the brain, the situation shown in figure 9.6 may appear similar, but the differences do not have the same origin. Here, the enantiomers activate the same receptor types but with differing affinities. This might correspond to type 1c pairs, with largely similar scent, but hints of differences. The figures show there is a distinct type 1 case (figure 9.2) and a distinct type 2 case (figure 9.5), which are the ones we consider in the rest of the study, but the others (figures 9.3, 9.4 and 9.6) indicate areas of ambiguity.

9.1.2.2 Comparisons of pure enantiomers and racemic mixtures

Can one learn more from racemic mixtures, in which the proportions of the two enantiomers can be varied?

Racemic mixtures of type 1's

For type 1 enantiomers, and also for type 1i, the scent of the racemic mixture should be exactly the same as the individual scents (see figure 9.7).

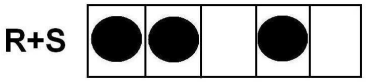


Figure 9.7: A mixture of figure 9.2.

There may be further useful information in the way the threshold varies with composition. If so, one could use racemic mixtures to identify 'true' type 1's. An in-

teresting pair are the aforementioned camphor enantiomers. The thresholds listed on Leffingwell's site give 1000-1290ppb for (+) camphor; but there are no data for (-) camphor. Does this mean a big difference in intensity, *i.e.*, type 1i? For a racemic mixture, the threshold is 4600 ppb for equal amounts of (+) and (-). If the two enantiomers contributed independently, one would expect somewhere between 1000-1290 ppb for (+) and (-) having equal odours, and 2000-2580 ppb if (-) had no odour at all. Why is a higher concentration needed when the (-) camphor is added? Individually the enantiomers are type 1's. Presumably both activate the same receptors with similar affinities. The higher threshold shows that the weaker (-) somehow obstructs detection of (+). The (-) form might antagonize some of the receptors through competitive docking, though there seem to be two problems with this view. First, if (-) antagonizes as a mixture you would expect it to antagonize receptors in its pure form, and then (+) and (-) would not smell the same individually. Secondly, one would expect (+) and (-) each to have a 50:50 chance of exciting the same receptors, yet the competition is asymmetric. Yet there is evidently competition. There are at least two explanations here that will be discussed in turn: pairing and competition. Firstly, in the racemic mixture the opposite hands may dimerize with each other so more individuals are required for adequate perception, a dimer of camphor may be too big for detection and so doesn't smell. This could be tested by a study of the correlation of proportions of the racemic and the threshold for detection. Secondly, there may be discrimination during diffusion into the mucus layer or simply according to the affinity olfactory binding proteins may have to each handed molecule so there is competition between hands to get at the site. However, this requires that what is chirally discerning before reaching the receptor site becomes chirally ignorant (camphor is type 1) and this seems unlikely. It is more likely the case that (+) and (-) both occupy two separate sites of the olfactory receptor but in doing so decrease the efficiency of actuation in some way.

Racemic mixtures of type 2's

Figure 9.8 shows a mix of the components shown separately in figure 9.4.

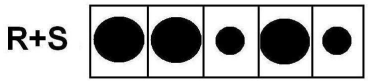


Figure 9.8: A mixture of figure 9.4.

Presumably the racemic mix should smell different from both individual components, whether a composite elemental (smells like sum of parts) or mixed (a gestalt). As a crude example, in figure 9.4, the right handed isomer smells 'citrusy' and the left smells 'flowery'. It is possible that the mixture shown in figure 9.8 results in two obvious components of citrus and lemon. It is equally possible the mixture smells like something similar to the parts, but not exactly so, such as 'sweet pea'. The results of gestalt mixing will depend on the types of receptor activated, neuronal mapping, and perhaps other factors. This may be tested, for example, by exposing the receptor repertoire to the R and L isomers and then to real sweet pea, and seeing if the same peripheral parts are excited. The different types of mixing in perception may be reflected by the spatial mixing pattern on the olfactory bulb. The character of a mixture may be further distinguished by temporal patterning. For example the excitation of 'citrus' perception may be separated from 'flowery' via time constraints of the receptors involved. Whatever these mixtures are, perception of such mixtures indicates different receptors are activated in R and L, with the consequence that receptors are (quite often) chirally selective.

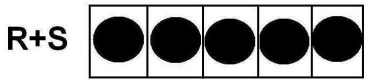


Figure 9.9: A mixture of figure 9.5.

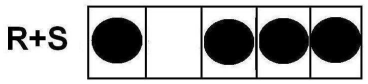


Figure 9.10: A mixture of 9.6.

Figures 9.9 and 9.10 show a mix of the components shown separately in figures 9.5 and 9.6 respectively. If the racemic mixture smells different from the compo-

nents, it is still not determined whether this is due to the binding or activation or both. Given the possibility already mentioned about competitive docking, we note that here it is assumed the strength of the signal from a certain receptor goes hand in hand with the odorant's affinity for it, *i.e*, a strong signal means dominant binding. This may not always be the case, as shown above for camphor. The intensity of the signal may depend on the geometrical configuration at the site, but also be complicated by unknowns governed by binding. The odorant-receptor interaction is vulnerable to the familiar problem of affinity versus efficacy, see Chapter 'A biological background'. So the information on type 2's is ambiguous.

Unfortunately it seems unlikely that any racemic mixtures, as exemplified above, would expose the true nature of type 2 enantiomers whether of the type shown in figure 9.4, 9.5, or 9.6. This is because for every scenario detailed and depicted in figures 9.8, 9.9, 9.10 we have the smell of $R \neq S \neq (R+S)$. The sum does not tell us what the parts are. Further, if a racemic smells exactly alike to one of its constituent hands, we are still no wiser as to the origin of type 2, all that can be inferred is that one hand has a greater affinity or rate at getting to the receptors it activates compared with its rival. Once at site an odorant may not activate as enthusiastically, and this may be due to an inadequate feature of the molecule or alignment, which we cannot know.

There is one outcome that may be useful: if there exists a type 1c where R and S differ in character, but one is significantly more powerful, we expect the racemate to smell of the dominant enantiomer, not different. Crudely, this might indicate that the more powerful enantiomer has a greater actuation capability for more receptor types. Thus, potentially, examples of such type 1c could yield information about typical receptor binding sites.

9.1.3 An important rule

Examination of the Leffingwell data for May 2006 [42] shows that a majority (52%) of enantiomer odorants contain 6-membered rings. In particular there is a novel systematic feature. The type 2 enantiomer pairs are distinctive. I have identified what seems a universal result that flexible enantiomers (at least those with the *6-membered ring flexibility*) are type 2, *i.e*. produce two different scents. We now examine the conjecture that 6-membered ring flexibility implies type 2, whilst rigidity implies type 1. To do so, the next section examines what we mean by 'flexible'.

The rule I propose is this: The members of an enantiomer pair will smell alike (type 1) when the molecules are rigid, and will smell different (type 2) when they are flexible.

9.2 Conformational analysis

If the nose discerns seemingly subtle enantiomeric conformations, then it makes sense to consider all other stereoisomeric conformations as it is unclear which conformation the nose is actually smelling. What is the receptor response to any configuration of the molecule it smells?

9.2.1 Cyclohexane

This study begins with cyclohexane because i) there are well established experimental results for the geometries of cyclohexane and ii) because it represents the nature of six-membered ring flexibility that we are interested in. So we start with some verified results.

Cyclohexane samples a 'conformational globe' - there are at least 6 possible stationary points that the molecule may visit along a potential energy surface (PES). The first three are: the global minimum 'chair', the high energy minimum 'twist-boat', and the second transition state 'boat'. More ambiguously the second three are the transition states 'envelope' and 'inverted half-chair ' (C_1 and C_2 respectively and in figure 9.11) and a higher energy stationary point- the 'pseudo-rotational form' [111] (C_S). It is considered [112] that the first saddle point is quite flat, and the two possible transition states are connected via a pseudo-rotational pathway. Similarly there is a relatively small energy separation between the twist-boat and boat structures- this area is quite flat. The boat is the transition state for the pseudo-rotation of the twist, so we expect rapid interconversion between these states. So from chair-twist there is a highly fluxional area connected by these three forms. The boat, chair and twist-boat are thought to be the basis conformations that constitute in various linear combinations the other three forms [112], i.e. chair + boat = envelope, inverted chair + twist-boat = inverted half-chair, and boat + twist-boat = pseudo-rotational form. Figure 9.11 shows the potential energy surface for cyclohexane that describes a chair-chair' interconversion, where now and hereafter the prime symbol ' refers to the mirror image version of a conformer. One chair form

converts to its mirror image and this proceeds via the twist state. During this process axial and equatorial positions swap and the molecule is fundamentally flipped inside out. This has been experimentally verified via NMR spectroscopy and it is estimated that this conformational flipping occurs at a rate of 10^{11} times a second at room temperature according to results descibed by Eliel *et al* [113]. So at room temperature there are conformational changes identified by the inversion. However, because the chair form is dominantly more favourable, identifying boats and twist-boats at lower temperatures via NMR is difficult because there is a low equilibrium concentration of these forms[112]. Table 9.1 gives a summary of what has been found by various methods to be the energy barriers of these conformational changes.

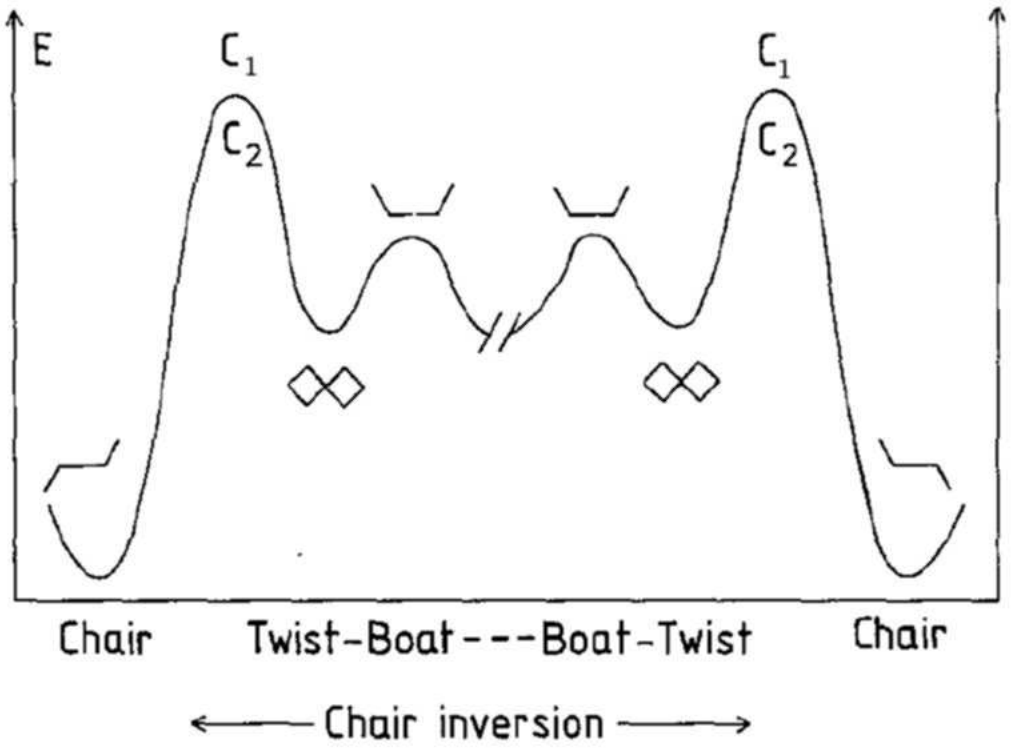


Figure 9.11: Potential energy surface for cyclohexane showing the chair-chair interconversion [114].

Looking at these values, we can see that, depending on the force field used, the barrier for twist pseudo-rotation is 1-2 kcal/mol. Inserting 1.5 kcal/mol into

Geometry	$\triangle E$ (MM)[1]	$\triangle E$ (*ab initio* SCF)[2]	$\triangle E$ (*ab initio* MP-2)[2]	$\triangle E$ (*ab initio*)[3]
Chair	0	0	0	
Twist-boat	5.7	6.79	6.95	
Boat	8.4	7.92	8.44	6.46
C_1	14.3	12.25	12.49	
C_2	13.3	12.21	12.42	
C_S	—	12.44	12.92	11.53

Table 9.1: Conformal energetics in kcal/mol for cyclohexane, a comparison of methods. [1]References data from Juaristi [112], these are molecular mechanics computations performed by Hendrickson and [2]is data from *ab initio* molecular orbital theory calculations by Dixon *et al* [111].The last set of calculations [3] are those I found using Gaussian '03 and B3LYP/6-31G*. Using the conformer QST3 search (see section 9.3.1.1) the C_S was found: a second order transition state, characterized by two imaginary frequencies.

the rate equation of interconversion for stereoisomers [115], we get a rate of $k = 4.92 \times 10^{11} s^{-1}$ at 298K, with which cyclohexane pseudo-rotates between twist and boat.

$$k = 2.084 \times 10^{10} T e^{-\triangle G^{\ddagger}/1.986T} \tag{9.1}$$

Equation 9.1 uses temperature (T) in Kelvin and the barrier height $G^{\ddagger}$ in calories per mole. Judging by this we can expect at room temperature at least one event in a picosecond timescale. So, simulating cyclohexane dynamics in picoseconds and at temperatures higher that 298K should definitely yield conformational exploration. Note, however, these results are based on experiment for liquids, the models I use, see Chapter 'Molecular Dynamics', are for isolated gas phase so these values must be treated as guide-lines only.

9.2.2 MD simulations method

All calculations start with an energy minimum for the initial geometry. This is found using the DFT method B3LYP and a 6-31G* basis set as in previous *ab initio* calculations unless otherwise stated. Though, we are looking at dynamics, so an absolute minimum is not necessary, this method should yield a good starting point for the initial positions in the MD simulation. The equilibrium geometry in Cartesian

.xyz form is converted into a dl_poly compatible CONFIG file, for an isolated gas phase molecule. The CONFIG file can be viewed using the graphical user interface (gui) and visually checked. A FIELD file is then generated from the gui implementing a Dreiding force field. This describes appropriate values for the known dihedral angles, and all other degrees of freedom. The CONTROL file is the final necessary file which describes particulars of the desired calculation. We use an NVE ensemble which stipulates a constant number of particles (N) volume (V) and energy (E). A constant energy simulation isolates the system so as it is unable to exchange heat or energy with its surroundings. A target temperature is set for the system to obtain, about which it averages and randomly orientated initial velocities are chosen accordingly. A good way to test the reliability of a calculation and the accuracy of a potential is to check afterwards the oscillations of temperature over the simulation. This again can be done on the statistics section of the gui: any large deviations from the set temperature indicate a problem. The time step is set to 0.2 fs, so at least 100000 steps are required to simulate 20 ps. As mentioned before the choice of a suitable time step is important to recreate a suitable trajectory. I chose this value, expecting the fastest bond vibration to correspond to fs, so this time period is shorter. Throughout the simulation configurations are recorded to a HISTORY file in .xyz format once every 100 steps, which corresponds to 0.02 ps. This is a time interval which seems appropriate- it is unlikely that within this short time events will be missed, thereby degrading the resolution. No electrostatics are calculated, as it is not trivial, but note this may help to better define the intermolecular potential and so better represent a barrier. The only two things that will vary for a simulation in the CONTROL file will be i) the temperature and ii) the number of steps. Simulations will be run for temperatures of 600K, 300K, and 10K. This range allows us to visualize motions at approximately room temperature, with two contrasting higher and lower values. The number of steps stipulated will usually be 100,000 or 500,000 or 600,000 to correspond to 20 ps, 100 ps, 120 ps long simulations respectively. According to the rate given by equation 9.1, we expect with higher temperature simulations and longer simulations to see more events corresponding to conformer changes. Based on liquid cyclohexane at room temperature the temperature of 300K with 20 ps should be sufficient for event resolution, however, contrast with the other different simulations provide a fuller picture. With these appropriate commands added to the CONTROL file, the simulation is ready

Dihedral	Boat	Chair	Twist
C1-C2-C3-C4	50.4	55.8	61.7
C2-C3-C4-C5	0	-55.8	-29.8
C3-C4-C5-C6	-50.4	55.8	-29.8
C4-C5-C6-C1	50.4	-55.8	61.7
C5-C6-C1-C2	0	55.8	-29.8
C6-C1-C2-C3	-50.4	-55.8	-29.8

Table 9.2: A table to describe the dihedral angles (in degrees) that define various geometries. These are optimized geometry parameters from *ab initio* molecular orbital theory calculated by Dixon *et al* [111]. Only cyclohexane chair has been verified experimentally. This is because the chair form is the most 'rigid', whereas the twist and boat forms are highly fluxional and 'flexible'.

to run. Afterwards, checks are made to the OUTPUT file to ensure a successful simulation, for example no wild deviations to the rolling averages. Apart from this the HISTORY file contains the most valuable information, because it shows snapshots (every 0.02 ps) of the molecule's configuration at a corresponding step in the simulation. The HISTORY file is converted back to .xyz format retaining the MD step information, ready for analysis. One of the first things that can be done is to view the .xyz information, animating the molecule throughout the steps. This can be done with an application like jmol, and should in essence indicate the molecules motion as if you were looking at it under a very powerful microscope, though not in real time.

9.2.3 Conformational Analysis

What characterizes the chair, boat and twist form of cyclohexane is the sequence and value of the 6 dihedral bonds made by the 6 carbon atoms. The dihedral measures the angle between two planes outlined by 4 consecutive atoms. A positive dihedral is where the angle follows a clockwise direction about the middle bond of the 4 atoms, negative is the anticlockwise direction. The dihedral angles used are those described in table 9.2.

These definitions in table 9.2 go into a PARAM (parameters file). In this analysis the 6 carbon atoms that constitute the ring have to be identified. This can be done by visualizing the output in .xyz form with jmol, which shows the labelling of the atoms, the clockwise order (and they *must* be in the right order) of the 6 car-

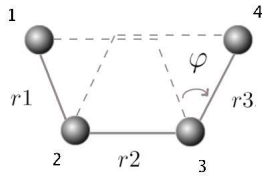


Figure 9.12: Dihedral angle measurement.

bon atoms also goes into the PARAM file. Then for each configuration (step) the 6 dihedral angles are calculated given by:

$$\varphi = atan\left\{\frac{sin\varphi}{cos\varphi}\right\} \tag{9.2}$$

$$sin\varphi = \frac{r2.\left(p2 \times p1\right)}{|p1|\,|p2|\,|r2|} \tag{9.3}$$

$$cos\varphi = \frac{p1.p2}{|p1|\,|p2|} \tag{9.4}$$

Where r1, r2 and r3 are the 3 vectors between the 4 atoms, and p1 and p2 are the two planes, described by $p1 = r1 \times r2$ and $p2 = r2 \times r3$. Equation 9.2 preserves the sign, necessary in this analysis. This gives a sequence of 6 dihedral values $\varphi_1 - \varphi_6$, these values are permutated six times ($\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6$ then $\varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6, \varphi_1$ etc) and these dihedrals are checked against the stationary point dihedrals shown in the sequence described in the PARAM file. So for each stationary point we have the difference from that target $d = \sqrt{\Sigma_i \left(\varphi_i - \varphi_{sp,i}\right)^2}$, where i denotes 1-6 and sp denotes the dihedral of the stationary point. The permutating ensures the *minimum* difference from each stationary point is found, and so accounts for inverted chairs (chair') for example. This gives us the difference from boat, chair and twist that can be plotted in a 3D scatter graph. Given the time stipulations above, for a 20 ps simulation there are 1000 (every 100 of 100,000) geometries recorded to the HISTORY file for a 0.02 ps period between each geometry. For 100 ps 5000 geometries are recorded to the HISTORY file for a 0.02 ps period between each diagram. A scatter graph will contain at least 1000 points therefore and at these geometries using the above method I calculate the nearness the geometry has to 'chair', 'boat' and 'twist'.

The difference is given in radians, 0 indicates most-like, 3 indicates least-like.

9.2.4 Results

9.2.4.1 3D scatter plots

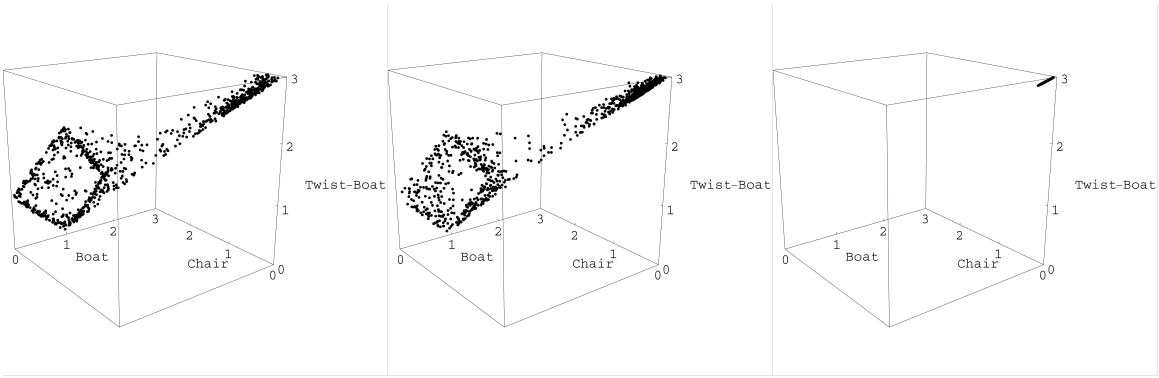
The results show scatter plots as described above for 600K, 300K and 10K. Figure 9.13 shows results for a first 20 ps equilibrating period and figure 9.14 shows results for a 100 ps period taking the initial starting point of this simulation as the end point of the previous 20 ps simulation. There are three main points to be derived from these results.

1. They show that isolated cyclohexane explores a wide conformational space, *i.e.*, there is a great distribution of points on the scatter graph at 600K and even at around room temperature. In contrast at 10K very little is happening and the molecule remains chair-like.

2. For the higher temperatures, see figure 9.13 a) and b) and figure 9.14 a) and b), we see that during a simulation a distinct square forms on the boat-twist plane. This looks to indicate cyclohexane's preference to pseudo-rotate about the boat and twist forms. To verify this, see figure 9.15, I have extracted points from the vertices of the square of 9.14, for 300K and examined the corresponding .xyz coordinates, see figure 9.16. The four vertices correspond to boat, twist, boat' and twist'. Where boat' and twist' correspond to the mirror reflection (the negative of the dihedral values as stipulated in 9.2). These four points are interconnected by the pseudo-rotating states and it looks, for cyclohexanol, that there is no especially *preferential* state along this square, and the molecule easily fluctuates from one state to the next in either direction.

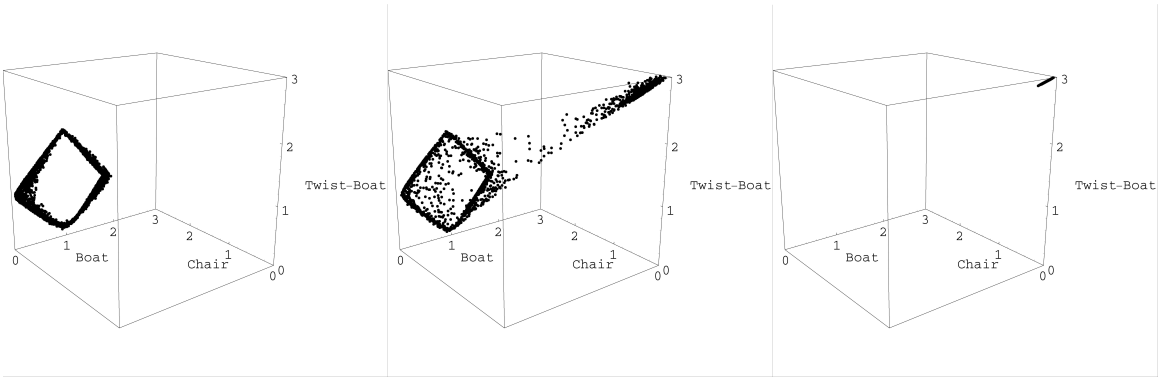
3. The longer simulation at 600K, see figure 9.14 a), stands out. At the higher temperature the energy barrier required to approach this pseudo-rotating square is easily surmounted. However, it appears there is no interconversion, only population of the pseudo-rotating square. In fact we can't be sure from any of these diagrams if there is full interconversion from chair-chair'. Throughout the first 20 ps at 600K, there is a graduation of states from chair-like to the pseudo-rotating square. However, after this 20 ps 'equilibrating' period the

100 ps simulation shows no graduation and it appears cyclohexane remains around the middle flat region of the PES as described in figure 9.14.



(a) (b) (c)

Figure 9.13: Cyclohexane simulation of 20 ps, showing the conformational space of an isolated gas phase molecule at a) 600K, b) 300K and c)10K. Scatter graphs show 3-axes scaled 0-3 radians for chair-like, boat-like or twist-like, 0 indicating most like (with the least deviation from the relevant geometry). A large distribution of points indicate large conformational sampling, and a high density of points indicate a popular conformer. We see with increasing temperature that cyclohexane explores a much wider conformational space: at 600 K more heavily populating pseudo-rotating states between boat and twist.



(a) (b) (c)

Figure 9.14: Cyclohexane simulation of 100 ps, showing from left to right the conformational space of an isolated gas phase molecule at a) 600K, b) 300K and c) 10K. Scatter graphs show 3-axes scaled 0-3 radians for chair-like, boat-like or twist-like, 0 indicating most like (with the least deviation from the relevant geometry). A large distribution of points indicate large conformational sampling, and a high density of points indicate a popular conformer. We see after the equilibrating 20 ps, that at higher temperatures cyclohexane will easily surmount the energy barrier and heavily populate pseudo-rotating states between boat and twist, throughout a longer simulation.

Because of the symmetry of the cyclohexane it is the case that information is lost as to which conformation is a chair or a mirror image chair, this is because the dihedral stipulations for the chair (see table 9.2) define both, so in permutating the 6 dihedrals a match would be found whether the cyclohexane is actually 'chair' 1 or 'chair' 2, see table 3.

Examining table 9.3, we see that two chair geometries are related by reflection, however, between these states there are 6 boat and 6 twist geometries (for both-three of which are related by reflection) that can be sampled in-between chair-chair' inversion. Figure 9.16 shows a possible route for chair-chair' interconversion that was extracted from the data in figure 9.14, for 300 K. The conformations are labelled with the position they are found on the pseudo-rotating square of figure 9.15. Note however, this is just an example, it is becoming apparent that there are many possible routes between chair-chair' via boat-twist interconversion that can take place.

Dihedral (in degrees)	'Chair' 1	'Chair' 2
55.8	C2-C6-C1-C5	C4-C2-C6-C1
-55.8	C6-C1-C5-C3	C2-C6-C1-C5
55.8	C1-C5-C3-C4	C6-C1-C5-C3
-55.8	C5-C3-C4-C2	C1-C5-C3-C4
55.8	C3-C4-C2-C6	C5-C3-C4-C2
-55.8	C4-C2-C6-C1	C3-C4-C2-C6

Table 9.3: Table to show loss of data, 'Chair' 1 corresponds to the 'chair' form in figure 9.16 and 'Chair' 2 corresponds to the mirror 'chair' form in figure 9.16. The method of detection however, fails to differentiate, they can only be told apart by examination of a trajectory to see which particular twist or boat state they pass through.

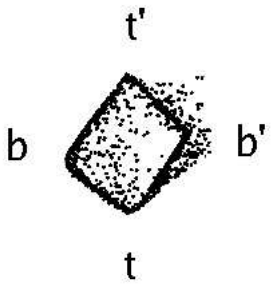


Figure 9.15: A closer examination of some points in simulation b) from figure 9.14, for 300 K. The 4 vertices of the square correspond to boat, twist, boat' and twist'. Where boat' and twist' correspond to the mirror reflection (the negative of the dihedral values as stipulated in 9.2). I extract some of these points, in figure 9.16.

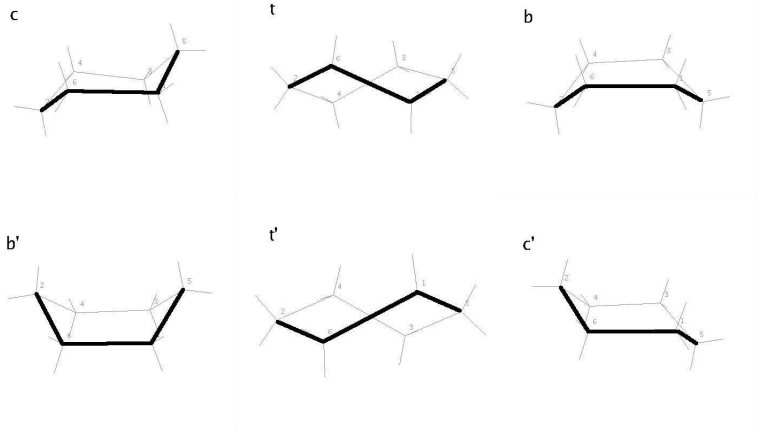


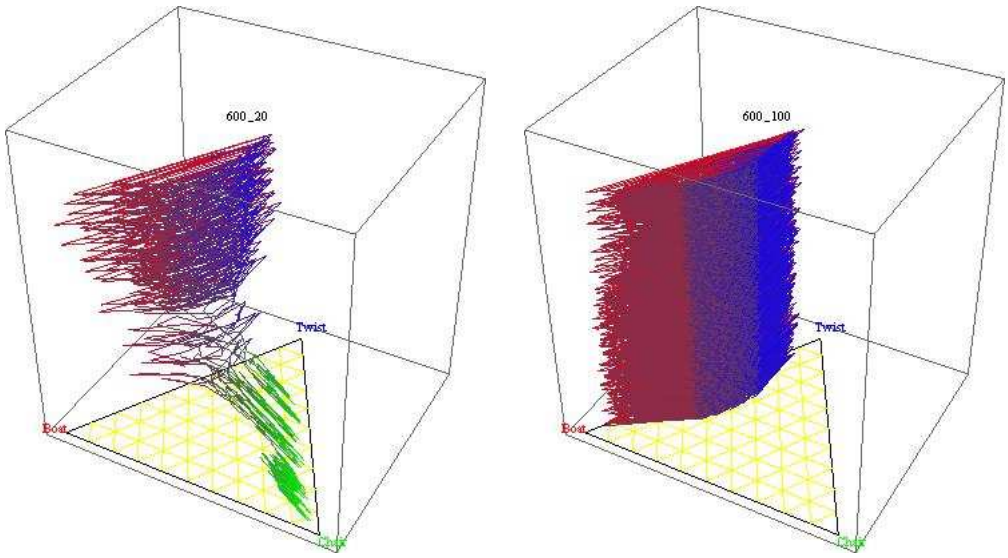
Figure 9.16: An example of interconversion chair to mirror image chair and the various states. Dark lines show bonds at the front of the page. This is an example of the chair-chair' interconversion defined by the swapping of dihedrals shown in table 9.3. There are other ways of converting chair-chair' due to symmetry, however, these are examples extracted from the data of figure 9.15, and are shown to be interlinked.

Points 1) and 2) are consistent with expectation, and so is perhaps 3), however a closer examination of the *trajectories* between states may be clearer and is desired, as a lot of the plots are overlaid and so no history can be seen.

9.2.4.2 3D ternary diagrams of trajectories

Since we require 4 dimensions (chair-like, boat-like, twist-like and time) there is a difficulty in presenting this data. However, *phase-diagram* -like plots can show clearly this type of information by making use of ternary axes. So onto a plane made by 3 corners of an equilateral triangle we can plot the normalized data. A plot in the middle of the ternary graph would indicate a configuration equally like the chair, the boat and the twist. Plots nearest the chair corner indicate chair, and so on for the others. A line is drawn between each point in chronological order throughout the simulation steps. The history is thus shown by time increasing out of the plane. Colour assignments go: green for chair-like, red for boat-like, and blue for twist-like.

Since the plots above for cyclohexane at 600K for 20 ps and 100 ps are intriguing they are investigated using this method (see figure 9.17).



(a) 20 ps at 600K

(b) 100 ps at 600K

Figure 9.17: A 3D ternary diagram trajectory for a) the 20 ps cyclohexane simulation at 600K and b) for the 100 ps cyclohexane simulation at 600K.

Figure 9.17 a) of the first 20 ps shows transitions through green-red-blue-green, at the beginning of the simulation that indicates interconversion chair-chair' can happen, though it is exhibited here as a relatively rare event. There are about 6 of these transitions in a 20 ps time-frame. However, this does not conflict with the experimental findings, as mentioned in section 9.2.1. Figure 9.17 b) shows the simulation after this 20 ps for another 100 ps, and that once an initial barrier is surmounted, the preference is to stay at the relatively flat pseudo-rotating region, this implies greater entropy contribution and complicates the simple one dimensional picture of a potential energy surface (figure 9.11).

9.2.4.3 Potential energy surfaces

To complement the molecular dynamics simulations, and in consideration of the above, *ab initio* calculations were run to investigate the potential energy surface of cyclohexane, to see if the transition chair-chair' can be seen. Simulations begin with a starting point geometry (indicated in the figures) and these were optimized

using B3LYP/6-31G* on Gaussian '03 and the single point energy found. A potential energy scan (PES) can be performed by Gaussian by defining a C-C-C-C dihedral angle, in the molecule, to be incremented by 5 degrees over a range. The other atoms are held fixed and the rest of the molecule is optimized partially, for each incremental step. The total energy is extracted for each nuclear geometry and plotted against the dihedral increment. In figures 9.18 to 9.22, the potential energy surfaces are found for different paths, and extracted geometries for various stationary points are pictured.

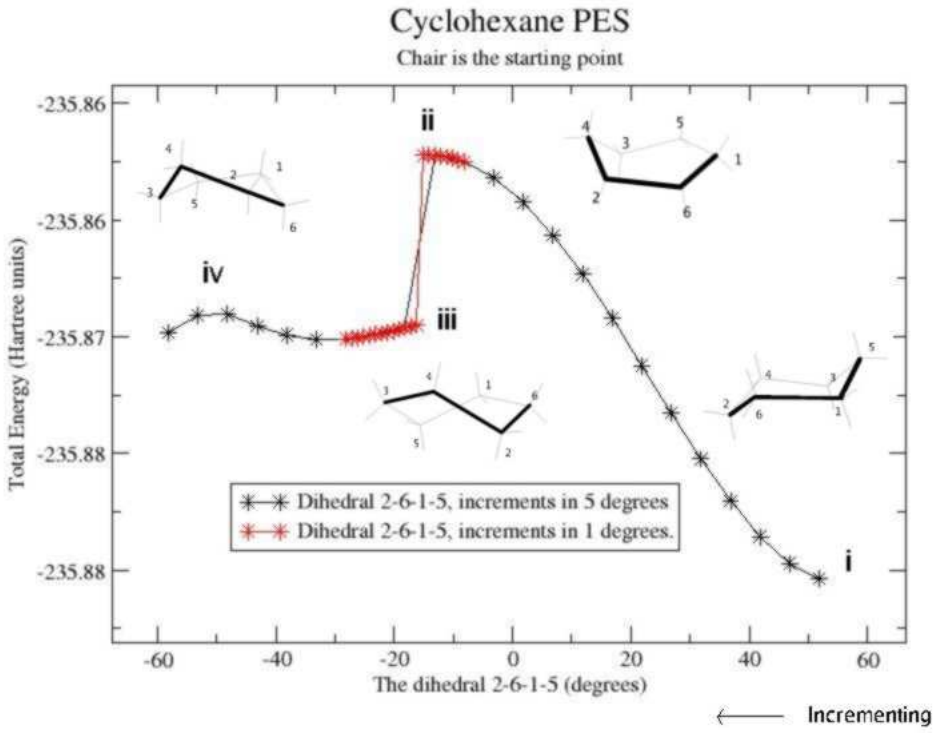


Figure 9.18: Potential energy surface for cyclohexane. Starting the increments of the dihedral C2-C6-C1-C5 at the optimized chair (extracted from figure 9.14 at 300K) the conformation evolves into a pseudo-rotating state.

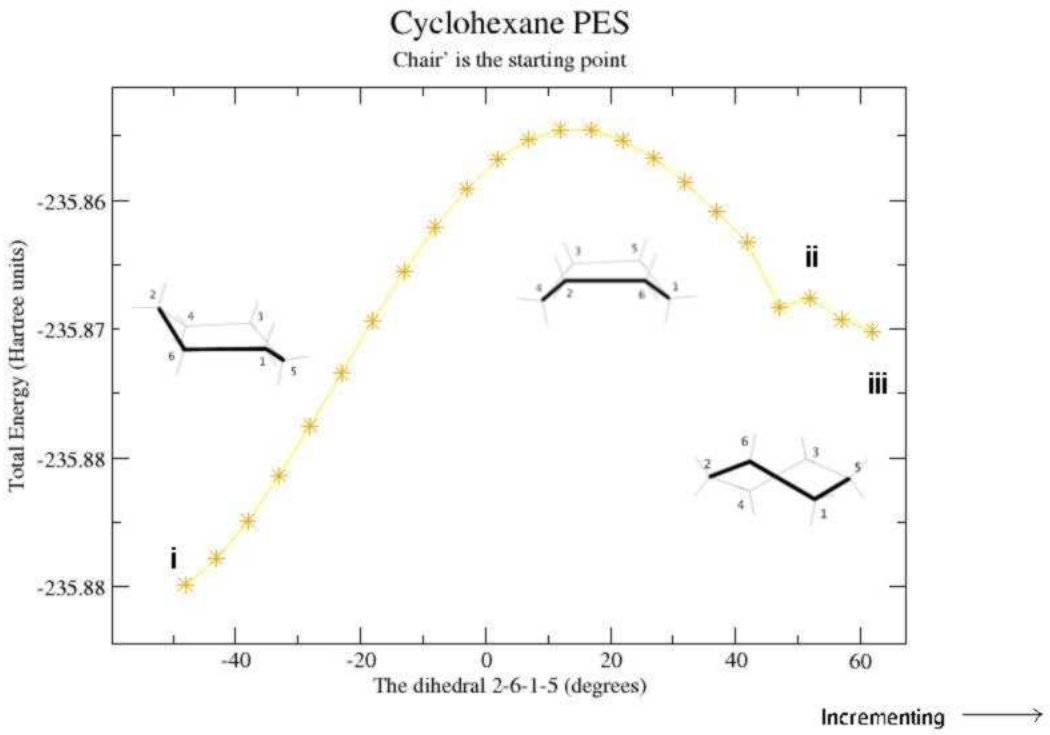


Figure 9.19: Potential energy surface for cyclohexane. Starting the increments of the dihedral C2-C6-C1-C5 at the optimized chair' (extracted from figure 9.14 at 300K) the conformation evolves into a boat state.

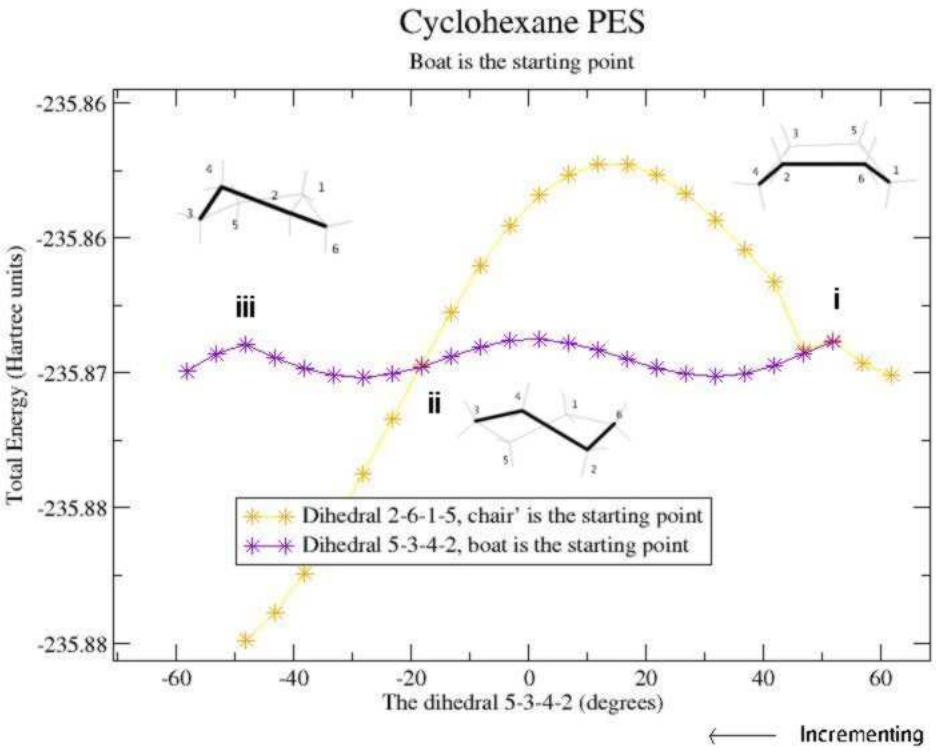


Figure 9.20: Potential energy surface for cyclohexane. Starting increments of the dihedral C5-C3-C4-C2 at the boat state (from the end point of figure 9.19) the conformation evolves into a pseudo-rotating state.

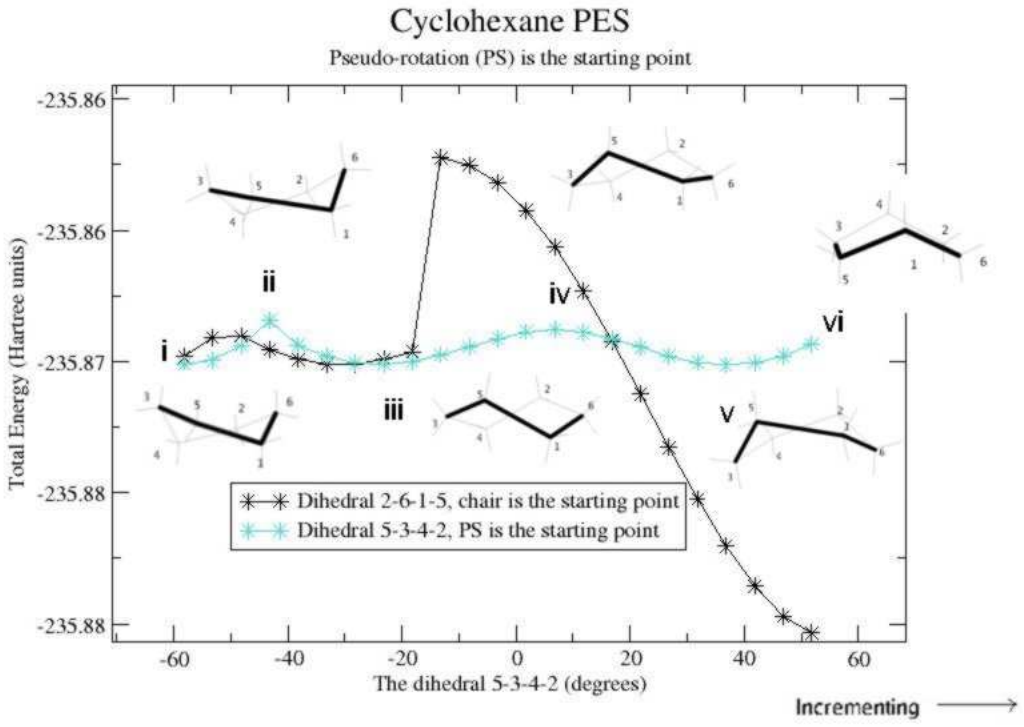


Figure 9.21: Potential energy surface for cyclohexane. Starting increments of the dihedral C5-C3-C4-C2 at the pseudo-rotating state (the end point of figure 9.18) the conformation evolves into the same boat state as figure 9.19.

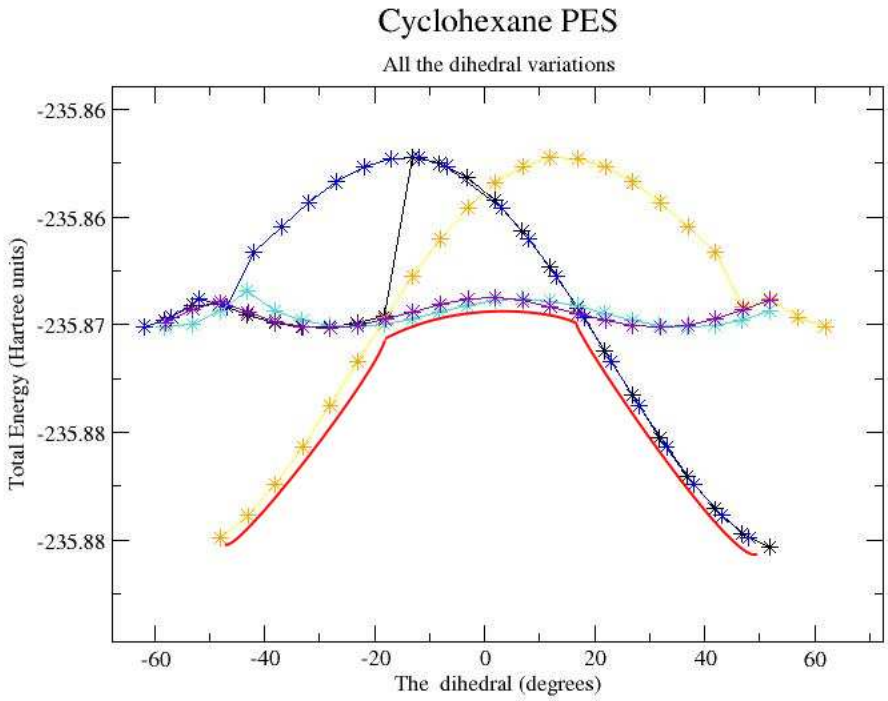


Figure 9.22: A potential energy surface for cyclohexane showing all previous plots, with an extra data set of 9.19 with the dihedral direction reversed to examine overlay with set 9.18. The red line drawn in shows a lowest energy path way.

Two chair forms related by reflection are extracted from the data in section 9.2.4.1, from the 300K simulation shown in figure 9.14. The first surface, figure 9.18 seeks to increment the C2-C6-C1-C5 angle until a boat is found, as in figure 9.11. However, the i) optimized chair conformation evolves into ii) a maximum point which is the envelope, then there is a discontinuity at iii) which corresponds to a twist geometry and the end point is iv) a pseudo-rotating state. The red line shows narrower increments through the discontinuity region to verify its existence. Interestingly in contrast to the second surface, figure 9.19, the mirror surface does not show this discontinuity. The i) optimized chair' conformation evolves into a maximum point which is the envelope and then ii) a boat form and the end point is iii) a twist form. Since a boat form was successfully reached in figure 9.19, the next scan, figure 9.20 starts at this point and increments the dihedral C5-C3-C4-C2 in

order to obtain another chair form. However the scan runs from the i) optimized boat to ii) the twist and iii) pseudo-rotating state seen in 9.18. So going backwards the boat does not tend to take the higher energy path through a transition to a chair, but favours the flatter pseudo-rotating transition. Similarly the end point of figure 9.18 is taken as the starting point of scan 9.21 and the dihedral C5-C3-C4-C2 is incremented in order to obtain another chair form. Again the pseudo-rotating path is taken. The scan runs from i) the optimized pseudo-rotation form to ii) something mid-way to iii) a twist to iv) something midway to v) a pseudo-rotation in the opposite direction of i) to vi) the boat state reached from the chair' in figure 9.19. The final scan 9.22 superimposes them all, with an extra plot of 9.19 but with the direction of the dihedral reversed. This combination of plots indicates a lowest energy path that is possibly preferable- the lines that connect chair-chair' via one twist-boat-twist'- the path of least resistance.

9.2.5 Analysis

Consideration of the 3D scatter plots, the ternary trajectories and potential energy surfaces in combination give a fuller picture of cyclohexane dynamics. Points 1) and 2) are consistent with expectation and point 3) makes more sense after careful examination. Although for chair-chair' interconversion there is a minimum energy path, see figure 9.22, there are many more possibilities for the geometry along the flat pseudo-rotating paths in figures 9.20 and 9.21. The strong implication is that an enthalpy-entropy compensation factor is at work. We note also that this analysis, for the PES surfaces, is one dimensional only, and there can be many flat paths between the chair-chair'. It is likely that entropic factors favour cyclohexane motion along these directions.

The important point is that the MD method and conformational analysis highlights which type of transitions a 6-membered ring is *able* to make. How these transitions are induced or weighted is something that cannot be reasonably modelled without the olfactory receptor structure. The emergence of the rule, however, in section 9.1.3, strongly indicates this flexibility is important in perception, so now we illustrate how these methods can be applied to analyze the dynamics of odorants, referring back to the rule.

9.3 The rule

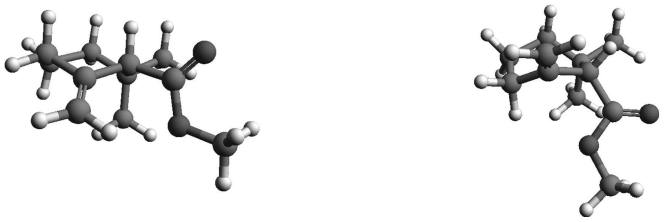
Reiterating the rule: *the members of an enantiomer pair will smell alike (type 1) when the molecules are rigid, and will smell different (type 2) when they are flexible.* The validity of this rule is explored below.

9.3.1 6-membered ring flexibility

There are various types of 'flexibility'; for 6-membered rings alone there are at least 3 groups. The first group are those having *cis-trans* transitions about the 6-membered ring, see figure 9.23. These correspond to molecules where two functional groups lie across the plane of the ring, and where it is feasible that these groups can isomerize by switching between sides of the plane. Secondly, there are molecules whose flexibility stems from cyclohex*ane* ring twist transitions from chair to boat conformations, as shown in figure 9.24. Thirdly, there are those molecules for which cyclohex*ene* ring twists are possible, similar to cyclohexane chair-boat, but more strained, as shown in figure 9.25. Figures 9.23, 9.24 and 9.25 show the minimum energy geometries of chiral odorants. We shall it is these types of 'twist' transitions that appear to be important to perception.

(a) 1S-4R *Cis* (b) 1R-4R *Trans*

Figure 9.23: Type 2 *cis* and *trans-p*-menthan-8-thiol-3-ones, (1*S*,4*R*) *cis* smells 'blackcurrant leaf, tropical note of passion fruit, intensive fruit note'. (1*R*,4*S*) *cis* smells 'rubber, mercaptan-note, isopulegone note, sulfurous, disagreeable'. 1R-4R *trans* smells 'onion-like, weak fruity, tropical, dirty'. the (1*S*,4*S*) *trans* smells 'stronger than (1*R*,4*R*) isomer, tropical, sulfurous, pronounced buchu leaf oil notes'.



(a) Chair (b) Boat

Figure 9.24: Type 2 chair and boat (1*R*)-γ-methyl-cyclogeranate, smells 'camphoraceous, corky, cellar'. (1*S)* smells 'aromatic, damascone-like, thujone, fruity'.

(a) Equatorial (b) Axial

Figure 9.25: Type 2 equatorial and axial (4*R*)-carvone, smells 'sweet spearmint, fresh herbal' and (4*S*)- smells 'caraway, fresh herbal'.

Many other types of flexibility are possible, though we shall not consider them here. They include the entegen-zusammen (*E*-*Z*) transitions for which *cis*-*trans* cannot be unambiguously defined. An (*E*-*Z*) transition across a double bond in a carbon chain can make a vast difference to odor perception [7]. One may contrast the "creamy, butter, odour" of *Z*-4-heptanal with the "aggressive, green and putty-like" odour of its *E* isomer. There may also be transitions across 5-membered rings or odorant rings with more than 6 atoms (for example see no 14 in the table 9.5) that are not unlike *cis*-*trans* or chair-boat flexibilities observed in 6-membered rings. Such transitions may be of interest in another study.

Secondary to their relative ubiquity another reason for studying odorants containing 6-membered rings is that they are similar in form to 11-*cis*-retinal. Receptors, such as the ORs, evolved from rhodopsin could be sensitive to the structural

changes for these types of transitions. The conformational change in 11-*cis*-retinal to all-*trans*-retinal upon photoexcitation is the initiating step in visual signal transduction, with photons of several eV being needed to trigger the protein-ligand conformation change. Olfaction involves rather different energies (probably of order 0.1 eV as illustrated below) from the optically-induced processes. However, it is possible that in olfaction the molecular environment might reduce the energy difference between alternative structures, for example by stabilizing a higher energy form through interactions with molecular components of the receptor. This could be achieved through favourable binding and folding. It is common that G-protein coupled receptors (GPCRs) change conformation upon binding. For example, the intercellular loop 3 (IL3) is known to be important in G-protein activation [116] and Day *et al* observe that IL3 in the GPCR $\beta_2 AR$ connect and can move the transmembrane (TM) helices TM5 and TM6 such that they realign with TM3, producing relatively different responses depending on the agonist that induced the changes. If such properties of GPCR's are conserved in the olfactory receptor repertoire, similar activation should not be ruled out. The Leffingwell database contains many odorants that are structurally similar to 11-*cis*-retinal, such as the strong smelling irones, ionones, and damascones. It is reported that the β-ionone ring of 11-*cis*-retinal exists 8Å from a zinc coordination site within the binding site in rhodopsin [117]. With zinc deficiency the receptor rhodopsin malfunctions, similarly a lack of zinc causes anosmia. It seems the six-membered ring is an important feature in olfaction, and so its conformational properties may explain certain observations. Thus *cis*- and *trans* 1,2-difluoroethenes have differing dipole moments as one conformer exists out of the plane (is non-planar) [118]. This affects their observed IR spectra, and could be relevant in any model that relates selectivity to odorant vibrations. It is interesting that in figures 9.23 to 9.25 these types of geometry changes produce the biggest changes in positions of various dipole moments within the molecule. Further these regions that change position probably define how the molecule binds to the receptor, see section 9.3.2. So there are several reasons to look at the flexibilities of 6-membered ring molecules. The data analyzed [42] included 428 enantiomer pairs (so 856 molecules). Of these, 450 molecules were flexible, following the above definition; 404 molecules were rigid, and 2 were ones that are ambiguous.

9.3.1.1 Relative energy calculations

It is usual that energy differences between the alternative structural forms of the isolated molecule are larger than thermal energies. For an idea of these magnitudes table 9.4 shows the calculated isolated gas phase energies for initial and final state geometries and the transition states. We make comparisons with cyclohexane. Cyclohexane is not chiral but it is useful as a good control and measure of flexibility, as in the section 9.2.1 above. Chair and boat *like* geometries were optimized using B3LYP/6-31G* on Gaussian '03 and the single point energy found. Transition states were initially estimated from a potential energy scan (PES) implemented again using B3LYP/6-31G*. For each molecule a C-C-C-C dihedral angle is incremented by 5 degrees over a range of around 100 degrees with these atoms held fixed and the rest of the molecule is optimized partially, as above, section 9.2.4.3. From these results an initial guess was taken from the potential energy surface for the transition state which is then implemented in a Gaussian 'QST3' calculation which uses the *Synchronous-Transit-Guided-Quasi-Newton* method. Thus the transition state is found and verified by analysis of the second derivatives to find the zero gradient and one imaginary frequency. We analyze cyclohexane, which is well established, as a test of the method, and compare our results with Peng *et al* who find $E_1 = 0.2eV$, $E_2 = 0.5eV$, and so $E_3 = 0.3eV$ using Hartree-Fock methods [84], where E_1, E_2 and E_3 are the energies defined in table 9.4, and so find a satisfactory agreement.

Molecule	T-B (E_1)	T-C (E_2)	C-B (E_3)
Cyclohexane	0.22	0.50	0.28
R-γ-methyl-cyclogeranates	0.17	0.33	0.16
R-carvone	0.27	0.34	0.07

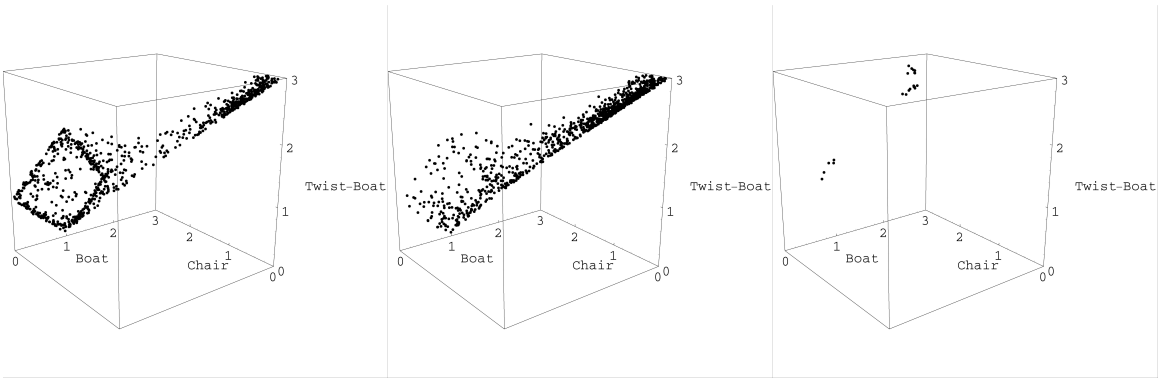
Table 9.4: Comparison of the energy gaps exhibited by chair-boat flexible molecules. Energies are in eV. C is the energy of the 'chair' configuration, B that of the 'boat', and T that of the 'transition state'. Calculated using B3LYP/6-31G* on Gaussian '03. E_1 is the difference from transition state to boat, E_2 is the difference from chair to transition state and so E_3 is the difference from chair to boat.

The energy gaps for odorants are comparable to cyclohexane and small, certainly small enough for the molecule to explore a wide range of geometries. So we expect to see chair, boat and transition-state *like* internal motion in odorants, rem-

iniscent of the type of conformational sampling explored by cyclohexane at room temperature[112].

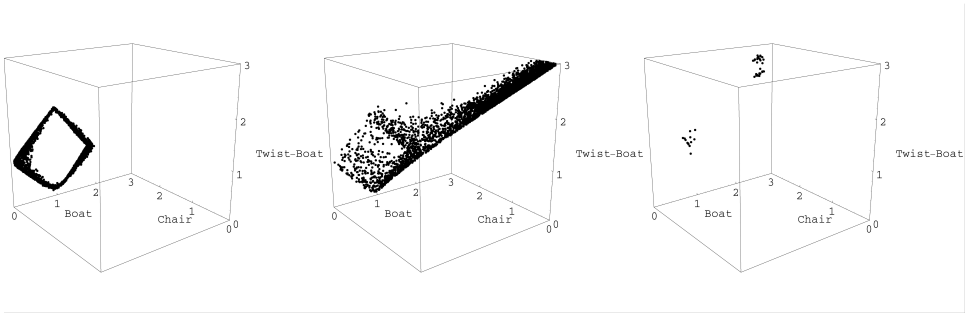
9.3.1.2 Results for small chiral odorants

We compare type 1 (1R,4S)- fenchone (see the table 9.5, number 3), and type 2 (4R)-carvone (see the table 9.6, number 3) of Leffingwell's data. These two are chosen because they are indisputably of definite type. The same simulations and conformational analysis were performed for these odorants that is outlined in the sections above 9.2.2 and 9.2.3 for cyclohexane. Figures 9.26 and 9.27 contrast the behaviour of (4R)-carvone and (1R,4S)-fenchone with cyclohexane, and we see very clearly that (4R)-carvone explores much more conformational space than (1R,4S)-fenchone, as is expected because fenchone is kept rigid in the twist-boat-like geometry by a carbon bridge. This is quite a drastic difference as is found in their perception.



(a) (b) (c)

Figure 9.26: From left to right, a) cyclohexane, b)(4R)-carvone (type 2) and c) (1R,4S)-fenchone (type 1) at 600K. We see clearly that cyclohexane reaches the boat-twist pseudo-rotating barrier, (4R)-carvone approaches it, but (1R,4S)-fenchone does not.



(a) (b) (c)

Figure 9.27: From left to right, a) cyclohexane, b) (4*R*)-carvone (type 2) and c) (1*R*,4*S*)-fenchone (type 1) at 600K. The simulations ran for 100 ps. We see clearly that cyclohexane easily reaches the boat-twist pseudo-rotating barrier, as does (4*R*)-carvone, but (1*R*,4*S*)-fenchone even after a longer simulation, does not.

So referring back to the rule, see section 9.1.3, 'flexible' is the case where an odorant has a 6-membered ring that is able to make these types of 'twisting' conformational changes that is seen in cyclohexane.

9.3.2 Are all type 1's rigid?

We avoid cases that are type 1i or type 1c, since the data raise other questions. Given within the *strong* similarity regime there are so many pairs that fall into the *smell different* category, the approach was not just to show support of the rule, see table 9.6, but to be rigorous and seek a *counter example* to the rule, see table 9.5. So in table 9.6 we seek an enantiomer that has a 6-membered ring but is type 1.

The conjecture implies that there should be few, if any, flexible enantiomers that smell the same. The data agree. From Table 9.5, it certainly seems that most of the rather few clear type 1's are rigid by our definition. Of the 11 definite type 1's (see the notes for why 1, 5, 6, 7, 12, 13 are exempt) 3 are rigid through constraint of the carbon ring at the 1 and 4 carbon position restricting twist (see 2-methylborneol, fenchone and camphor), and the rest have more exotic geometries (see no. 14). We consider the three possible counter examples (10, 11 and 16 in the table). Do these molecules contradict our conjecture?

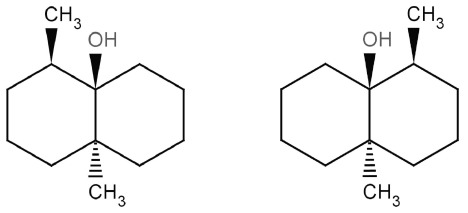
It is interesting to compare the type 1's geosmin and tetrahydronootkatone with

Table to show type 1 odorants, seeking a possible counter example

Section	Enantiomer Pair	No.	Flexible?	Notes-a counter example?
Cyclic Terpenoids	(2S,4S) 2-Acetyl-p-menth-6-ene and (2R,4R)	1	+	No, the (2S,4R) and (2R,4S) isomers smell different.
Bicyclic Terpenoids	(S) and (R) 2-Methylborneol	2	-	
	(1R, 4S) and (1S, 4R) Fenchone	3	-	
	(1R, 4R) and (1S, 4S) Camphor	4	-	
	(1S,4R,5R) and (1R,4S,5S) 5 -Acetoxy-1,8-cineole	5	-	'Virtually too weak to evaluate'.
Ionones, Irones, Damascones & Structurally Related Odorants	(1′S,2S,2′R, 5′S) and (1′R, 2R, 2′S, 5′R) 5-Methyl-ambercore	6	+	No, (1′R, 2S, 2′S, 5′R) and (1′S, 2R, 2′R, 5′S) smell different and comes with the description 'vague'.
Acyclic Terpenoids	(2R,3S), (2S,3R), (2S,3S) and (2R,3R) 3-Mercapto-2-methyl-pentane-1-ol	7	-	Note all 4 are type 1. 'The odour description appears to be for the 4 isomer mixture, but the enantiomers are differentiated by their threshold data.'
Lactones	(4R) and (4S) γ-octathionolactone	8	-	
	(3R) and (3S) Butylphthalide	9	-	
Sesquiterpenoid Related Odorants	(4R,4aS,6R,8aS) and (4S, 4aR, 6S, 8aR) Tetrahydronootkatone	10	+	Possible counter example.
	(5S) and (5R) 10-Demethyl-β-vetivone	11	+	Possible counter example.
	(4R,5R) and (4S,5S) 3,5-Dimethyl-4-(4-methyl-pent-3-enyl)-cyclohex-2-enone	12	+	No, 'Nearly odourless' and their isomers smell different, a partial structure of β -vetivone.
Steroids and Sandalwood type Odorants	(1R, 2′R, 3′R) and (1S,2′S, 3′S) Ebanol	13	-	Not very reliable- only descriptor is 'no sandalwood relevance'.
Musks	(3S) and (3R) 3-Methyl-1,4-dioxacylcopentadecan-2-one	14	-	
Misc	(R) and (S) Lilial	15	-	
	(4S, 4aS, 8aR) and (4R, 4aR, 8aS) Geosmin	16	+	Possible counter example.
	(R) and (S) 2-Methoxy-3-[1-methylpropyl] pyrazine	17	-	

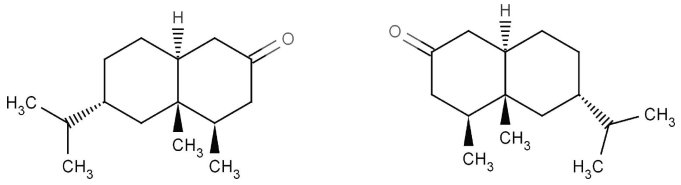
Table 9.5: Type 1's, + denotes a positive response, - a negative response to whether the odorant is flexible.

type 2 nootkatone, as they are structurally similar (see figures 9.28, 9.29 and 9.30).



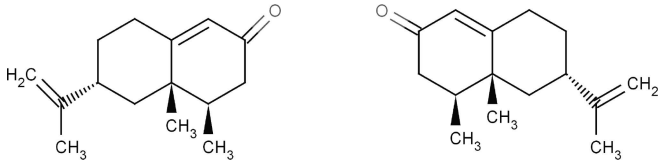
(a) RRS-Geosmin (b) SSR-Geosmin

Figure 9.28: Type 1i, geosmins smell 'earthy, musty'.



(a) RSRS-Tetrahydronootkatone (b) SRSR-Tetrahydronootkatone

Figure 9.29: Type 1, tetrahydronootkatones smell 'dusty-woody, fresh, green, sour, spicy, herbal, slightly fruity, animal, erogenic'.



(a) RSR-Nootkatone (b) SRS-Nootkatone

Figure 9.30: Type 2, nootkatones: the (4R, 4aS, 6R)-(+) enantiomer smells of grapefruit (0.8ppm) and its mirror is 'woody, spicy' (600ppm) [110]. Note also the (+) enantiomer is around 750 x more potent than the (-) [115].

Prima facie, one might expect all these odorant pairs to be all type 1 or all type

2. However, closer inspection shows that, in fact, the first two examples are rigid, whilst the third is not. Although all three exhibit 6-membered rings, the presence of two rings, combined with the transpositions of groups across the join exhibited in the first two, produces a restriction in movement not exhibited by the third. The presence of the two groups, *trans* across the join (H is equatorial and CH_3 is axial) locks both rings into two conformationally rigid chair-chair formations and, in addition, ensures that any substituents about the rings will occupy definite axial or equatorial positions [115]. If the configuration were *cis* (for example if both H and CH_3 where axial) then the rings have the ability to flex from chair-boat, and the substituent groups are mobile about axial-equatorial positions. Further, if there is no group, *cis* or *trans*, but a double bond as in nootkatone, then it seems the rings are also flexible.

9.3.2.1 Examination of the possible counter example

We use the same method as described above to evaluate type 1 (4*R*, 4*S*, 6*R*, 8*S*)-tetrahydronootkatone (figure 9.31) and type 2 (4*R*, 4*S*, 6*R*)-nootkatone (figure 9.32) and see that indeed tetrahydronootkatone is rigidized whereas nootkatone is flexible approaching the same type of pseudo-rotation exhibited by cyclohexane (see figure 9.14).

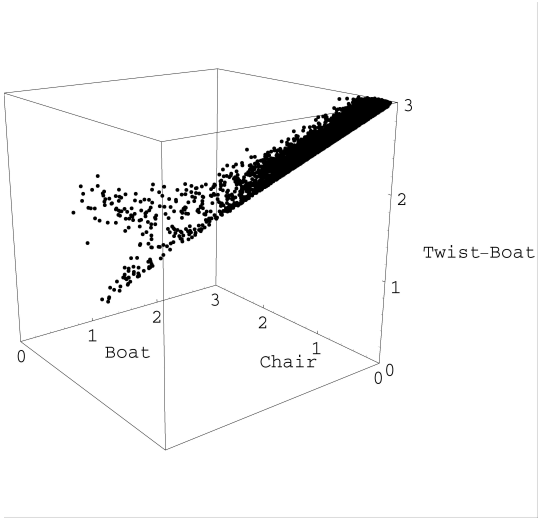


Figure 9.31: Type 1, (4*R*, 4*S*, 6*R*, 8*S*)-tetrahydronootkatone, 600K simulation for 100 ps. This odorant *does not* reach the pseudo-rotating contour shown in cyclohexane, figure 6. It is not flexible.

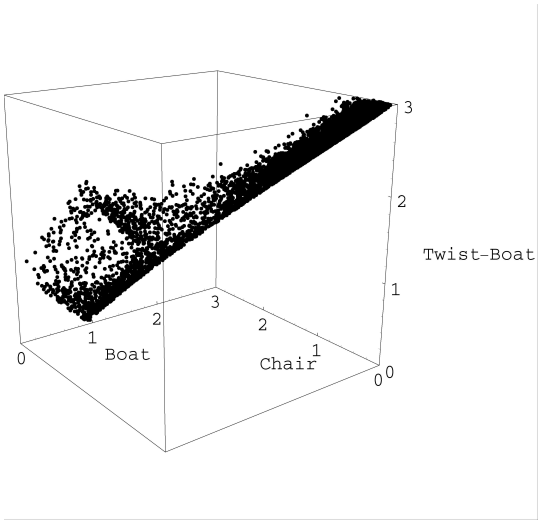
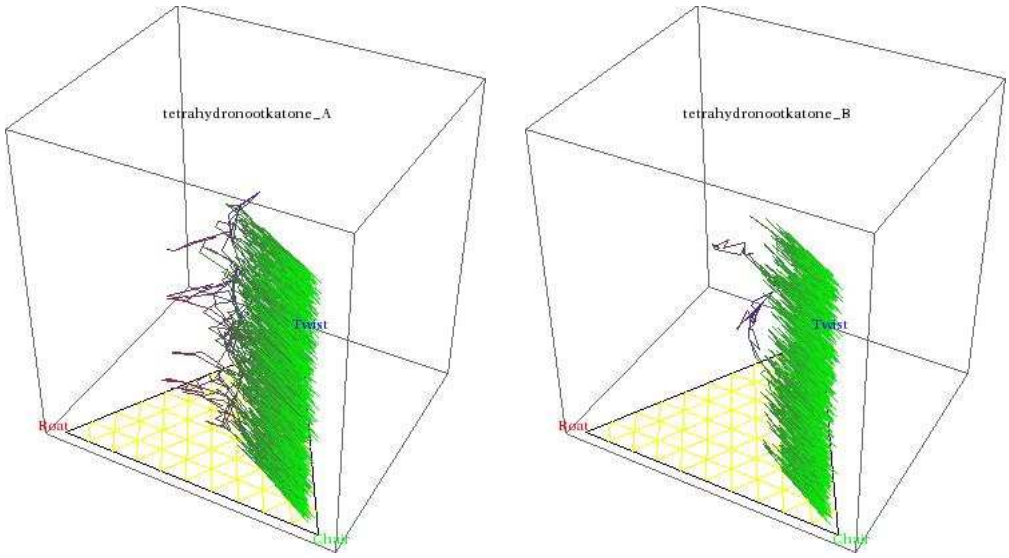
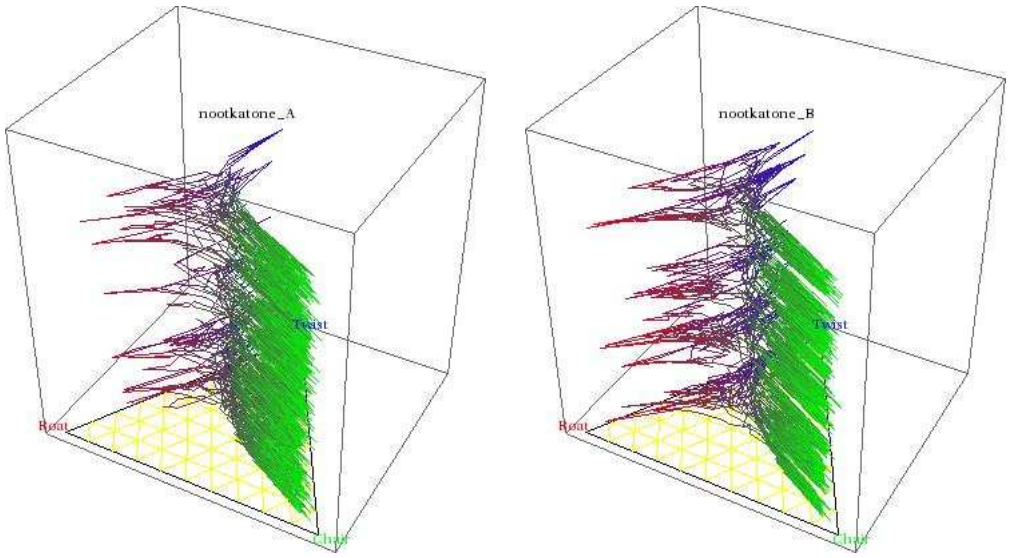


Figure 9.32: Type 2, (4*R*, 4*S*, 6*R*)-nootkatone, 600K simulation for 100 ps. This odorant *does* reach the pseudo-rotating contour shown in cyclohexane, figure 6. It is flexible.



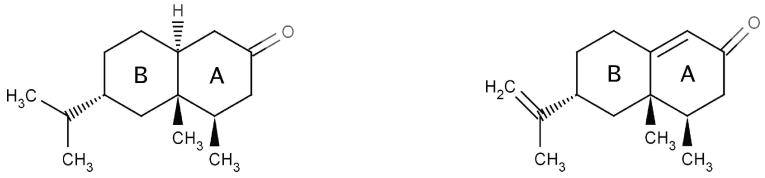
(a) Ring A (b) Ring B

Figure 9.33: A 3D ternary diagram trajectory for the 100 ps (4*R*, 4*S*, 6*R*, 8*S*)-tetrahydronootkatone simulation at 600K.



(a) Ring A (b) Ring B

Figure 9.34: A 3D ternary diagram trajectory for the (4*R*, 4*S*, 6*R*)-nootkatone 100 ps simulation at 600K.



(a) (b)

Figure 9.35: The ring notation for a) (4*R*, 4*S*, 6*R*, 8*S*)-tetrahydronootkatone and b) (4*R*, 4*S*, 6*R*)-nootkatone.

Figures 9.31 and 9.32 indicate some difference but a closer examination using the ternary diagram really highlights a definite distinctness in dynamics. These simulations separate the dynamics of the two rings in each molecule see figure 9.35, for the ring notation. Figure 9.33 shows the trajectories of (4*R*, 4*S*, 6*R*, 8*S*)-tetrahydronootkatone and figure 9.34 shows (4*R*, 4*S*, 6*R*)-nootkatone (both for 100 ps and 600K). We can see in nootkatone that both rings frequently sample much

more conformational space and interconvert. This is not the case at all for the ring B in tetrahydronootkatone which is almost completely restricted to a chair conformation and the ring A comparably is only able to make certain twist conformations. Comparing the saturated ring B in nootkatone with its neighbour ring A, we see it is comparably more flexible as the double bond in A endows some rigidity, as is expected. However, it is interesting that this effect is nowhere near as drastic as the difference a *trans* positioned single H atom makes to the overall behaviour of a molecule.

So, this possible counter-example, examined in detail, actually reinforces a strong correlation between flexibility and discrimination. We find, to the authors knowledge, a new observation that though nootkatone is not *cis* because it does not possess two groups, it exhibits a similar type of conformational freedom as compared to the *trans* version.

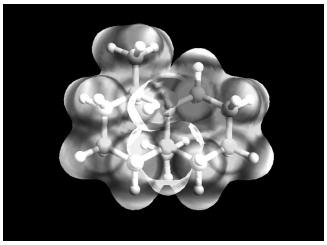
(a) R-10-Demethyl-β-vetivone

(b) S-10-Demethyl-β-vetivone

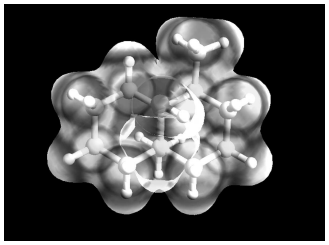
Figure 9.36: Type 1, 10-Demethyl-β-vetivone smells 'intense cresolic'.

The final possible counter-example is pictured in figure 9.36 (no. 11 in table 9.5 of this chapter). We suggest that again this structure is in fact rigid, as the 5-membered ring stabilizes the 6-membered ring in a similar way to that observed in the previous example. We conclude, that all probable *counter*-examples in fact are rigid, and *positive* examples. The significant conclusion from a thorough investigation is thus that there are few true type 1's, and what they all share is the property of rigidity and, further, one osmophore, as is shown below. We have found no contradiction to this proposition. Further we suggest an inverse relationship between rigidity and the number of receptors activated, *i.e.* rigid molecules can actuate fewer receptors.

9.3.2.2 'Osmophoric' groups

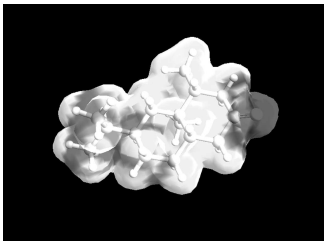


(a) (4R,4R,8S)-Gesomin

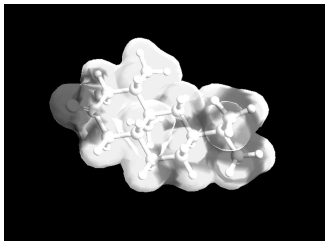


(b) (4S,4S,8R)-Geosmin

Figure 9.37: Type 1i, geosmins: electrostatic potential fit to the electron density isosurface. Note there is *one* main region of electronegativity, corresponding to the 'osmophoric group'.

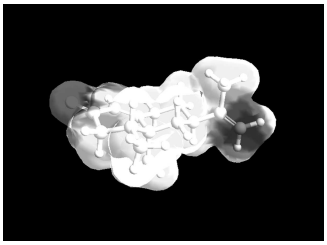


(a) (4R,4S,6R,8S)-Tetrahydronootkatone

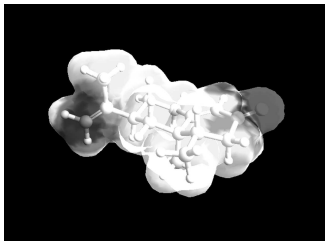


(b) (4S,4R,6S,8R)-Tetrahydronootkatone

Figure 9.38: Type 1, tetrahydronootkatones: electrostatic potential fit to the electron density isosurface. Note there is *one* main region of electronegativity.



(a) (4R,4S,6R)-nootkatone



(b) (4S,4R,6S)-nootkatone

Figure 9.39: Type 2, nootkatones: electrostatic potential fit to the electron density isosurface: Note, there are *two* main regions of electronegativity, one weaker than the other.

So, how does this rigidity, or rather its absence, affect important groups on the odorant? 'Osmophoric groups' have been coined as a phrase to describe polar parts of an odorant molecule. Usually odorants have one strongly polar group and a secondary weaker polar group [7]. Having discovered one correlation between two rare odorant qualities, namely that there are very few true type 1's and that they are rigid, it is interesting to observe that the same molecules seem only to have one osmophoric group. Here an osmophoric group means one that is electronegative, and an osmophore is a region rich in such groups. This was investigated by looking at the electrostatic potential (ESP) which gives a measure of which parts of the molecule can participate in electrostatic non-bonding interactions, crucial to the binding and alignment of the odorant in the receptor. The ESP, calculated using ArgusLab [81], is that felt by a positive point charge, and is mapped onto the electron density using ZINDO for several odorant pairs at their isolated phase equilibrium geometries (see figures 9.37, 9.38 and 9.39). In contrast, it is observed that the type 2 nootkatone exhibits two areas of higher regions of electronegativity about opposite ends of the conjoined rings (at the C=O and C=C) whereas the type 1i geosmin and type 1 tetrahydronootkatone exhibit only one area of electronegativity (about the alcohol and carbonyl group, respectively). Thus we propose that a combination of rigidity and a sole osmophoric group [7] actually renders left and right handed odorants superimposable (in the sense that the same atoms exist in similar space) within a receptor. In this case the receptor is unable to discriminate, thus these unlikely cases we perceive as type 1. The presence of two osmophoric groups, however, in

particular at either side of a flexible ring, introduces a degree of complexity in signal generation that may arise from the alignment of these osmophoric groups with contacts in the receptor. This complexity is reflected in our ability to differentiate enantiomer pairs, as indicated by figure 9.39 and table 9.6, and is discussed in section 9.4.1 .

9.3.3 Are all type 2's flexible?

As noted above, the division into type 1 and type 2 can be tricky, for there are various reasons as to why such a categorization could be unreliable and we have been very cautious in assigning enantiomers to type 1. We should be equally scrupulous in deciding which are of type 2 (Table 9.6). Given all the issues raised in section 9.1.1, when giving examples to the rule we shall only give odorants that lie within strict classifications, *i.e,* odorants that lie within the opposite corners of figure 9.1, so there can be no bone of contention concerning the 'subjectivity' argument

For those pairs with 6-membered rings, the positions of the functional groups have been counted in a clockwise direction for one hand; the prime symbol ' denotes cases where the group is positioned away from the plane of the ring (into the page). Similarly, the asterisk * denotes a chiral centre, and ^ denotes if the functional group is electron rich (*i.e.* there exists a π-bond, or something electronegative like oxygen and sulphur). We anticipate that a combination of ^ and * would make the receptor sensitive to chirality, as superimposability is then unlikely. This suggestion is based on three ideas. First, it seems likely that the electronically interesting part of an odorant to a receptor are the regions to which it may bind. Secondly, binding will be chiral-sensitive. If the above parts are positioned chirally they are opposite in space and thus different in character for each enantiomer unless a receptor is ambidextrous or equally and oppositely designed (chiral receptors exist for every chiral odorant). Thirdly, binding and detection are further sensitive to conformational change. The receptor may be rendered even more sensitive to chirality when conformational changes affect chiral and electronic groups.

Excluding the acyclics category, table 9.6 indicates a typical type 2 odorant recurrent pattern. The position of certain important groups about the ring is highlighted for their role in chiral discrimination. This pattern is consistent with our discussion above. It is also consistent with results from Laska *et al*, who found that wherever an odorant bears an isopropenyl group at the chiral carbon atom, a methyl group

Table to show type 2 odorants, seeking possible examples

Section	Enantiomer Pair	No.	Flexible?	Planar?	Chiral centre?	Electron rich?	Other?
Cyclic	(1R, 4S) and (1S, 4R) *cis* and *trans*-p-Menthan-*8*-thiol-3-ones	1	+	-	1,4	3,4	-
Terpenoids	Linalool oxides (pyranoids)	2	+	-	1,4	1,4	2,3
	R and S carvones	3	+	-	4	2,4	1
	Carvone epoxides	4	+	-	1,2,5	1,2, 3,5	-
	Carveols	5	+	2	2,4	2,4	1
	Trans dihydrocarvones	6	+	1	1,4	2,4	-
	Carvotanacetols	7	+	2	2,4	2	1
	(2S, 4R) and (2R, 4S) 2-Acetyl-p-menth-6-enes	8	+	4	2,4	2	1
Bicyclic Terpenoids	4- Cylohexylmethylene-isopinocamphores	9	+	1	1	2, (E-Z)3	5
Ionones, Irones,	Methyl-α-cyclogeranates	10	+	-	2	2	1,3
Damascones &	Methyl-γ-cyclogeranates	11	+	-	2	2,3	1
Structurally	Trimethyl-cyclohexyl-pentan-3-ones	12	+	2	2,3	2	1
Related	Trimethylcylcohexyl-pent-1-en-3-ones	13	+	2	2,3	2	1
Odorants	Tetrahydroirones	14	+	1,3	1,3	3	2
	Trans γ irones	15	+	1	1,3	3,4	2
	Theaspiranes	16	+	-	2	2	1,3
Acyclics	Methylhexanols	17	No ring	-	-	-	-
(Alcohols,	2-Methyylbutanoic acids	18	No ring	-	-	-	-
Esters, Acids,	2 Pentanols	19	No ring	-	-	-	-
Aldehydes)	3-Hydroxyheptanes	20	No ring	-	-	-	-
	1-Hexen-3-ols	21	No ring	-	-	-	-
	1-Hepten-3-ols	22	No ring	-	-	-	-
	1-Octen-3-ols	23	No ring	-	-	-	-
	1-Nonen-3-ols	24	No ring	-	-	-	-
	1-Decen-3-ols	25	No ring	-	-	-	-
	3-Mercaptohexanal	26	No ring	-	-	-	-
	3-Acetylthiohexanals	27	No ring	-	-	-	-
Misc	Methyl-1,4, dimethylcyclohex-3-ene-1-carboxylates	28	+	-	4	4	1
	2-Methylbutyl-2-methylpyrolidines	29	Pentane ring	-	-	-	1,2
	2-Methyl-N-(3*'*-methylbutylidene)butan-1-amines	30	No ring	-	-	-	1,2

Table 9.6: Type 2's, + denotes a positive response, - denotes a negative response to whether the odorant is flexible.

at the para-position and/or an oxygen containing group at the meta-position, then enantiomer pairs could be discriminated [119]. Further, as Laska *et al* discuss, this feature gives some support for Ohloff's multipoint attachment theory, whereby the receptor is "two toothed", binding- by using at least two points on the odorant.

At this juncture, odorants appear to fall into at least three natural categories: (a) true type 1's, (b) true type 2's, and (c) acyclics. In our quite substantial data set, it is always the case that that the two members of enantiomer pair with a 6-membered ring flexibility produce two different scents (type 2).

9.4 Discussion

For chiral discrimination in the nose we might postulate four possible explanations. Firstly, there may be discrimination *prior* to actuation through metabolism *en route*. One isomer only might break into chiral parts so the nose is smelling different molecules, and thus perceives them differently. Also diffusion processes to the cilia may be chirally biased, much like gas chromotography. Secondly, there may be discrimination *during* actuation of a receptor that may depend in the majority on adaptability of fit. Perhaps rigid molecules only activate receptors where there is a good fit, while flexible molecules can activate even when there is a poor fit: the poorness of the fit then means that the difference between left and right handedness actually matters, whereas for a good fit it might not. Thirdly, there might be allosteric regulation (one binding site regulates the response of a second binding site) of GPCRs. For example, one odorant could change a key geometry or energy probed by the second odorant. Allosteric regulation is known to occur in other GPCRs (*e.g.* muscarinic acetylcholine receptors). Fourthly, there may be discrimination *post* actuation in terms of how the brain might interpret the combinations of signals from a number of receptors. However, given the zone-to-zone evidence, and the lack of evidence for metabolizing enzymes, it seems likely that the discrimination of enantiomer pairs occurs because the receptor somehow differentiates these molecules.

9.4.1 How might flexibility affect olfaction?

In normal thermal diffusion processes occur during rare events, near critical values of reaction coordinates, and we suspect that relatively rare geometries have an

important role in olfaction, just as in diffusion. If so, thermal fluctuations allow relatively soft molecules to explore a wider range of geometries. Such an exploration of molecular conformations permitted by flexibility might be important before docking, or at the receptor, or both. Our discussion will concentrate on situations made possible by the relative ease of larger structural changes, and the effect of this at the OR site.

Flexibility complicates the interpretation of both the lock and key and the swipe card descriptions. Intuitively, one would expect rigid molecules to be sensitive to receptor structure (and so type 2), whereas flexible molecules might be more able to evade chiral constraints (and so type 1). Theimer *et al* [109] suggest that very flexible enantiomers, such as the acyclics, should always smell the same because they are able to explore mutually similar and wide conformational spaces. These expectations seem plausible in almost any model of olfaction. Yet both expectations prove false. This is less surprising when we recognise that the conformational space explored by an odorant is likely to be limited by a chiral receptor. Within the chiral space defined by a receptor, the motion will be subject to restrictions, but a flexible molecule will still be able to explore a wider range of conformations than a rigid one and so produce a more varied response.

We note also that it could only take a single extra receptor activated by just one hand (say, by L) for us to perceive a difference between L and R forms. Flexibility then offers the opportunity for the odorant to be at once promiscuous and selective in activating a receptor. This depends on the receptor type, an observation noted often in the literature [22] and this brings us back to the issue of a 'good' fit. For the type 1 fenchone, 1*R*,4*S* and 1*S*,4*R* are a good fit: they appear superimposable, and so may activate the same receptors and smell the same. However, the type 2 carvone has 4*R*, 4*R'* , 4*S* and 4*S'* forms (figure 9.25), where the prime ' denotes distinct equilibrium geometries. These are non-superimposable, so 4*R*, 4*R'* and 4*S* may adapt to fit and activate receptors equally, but 4*S'* may not. If so, chiral pairs would be discriminated. So, if the shapes, structures and positions of groups matter, which they clearly do, then flexibility must be also be a factor in olfaction.

Consider, for example, the nose's ability to discriminate between 4*R* and 4*S* carvone (see figure 9.25). If a molecule has n chiral centres, then there are always 2^n forms in space. When we introduce flexibility, there are also what we may call pseudoequlibrium states, metastable, but likely to be achieved with a useful probability.

With m such states, there will be $(m+1)2^n$ forms. For carvone, there are at least $2^2 = 4$ geometries where the equatorial form is the equilibrium and the axial the pseudoequilibrium form. Two of these states are shown in figure 4. Statistically, such molecules as these should have more chances of activating an appropriate chiral receptor. Combinatorially, there are not enough types of receptor to accommodate all these structures, so it is not surprising there is an asymmetry favouring one hand to succeed and its antipode to fail. Rigid chiral molecules however, should activate the same receptors with equal success owing to the superimposability of their fewer 2^n forms.

Whether chair/boat-type conformations are preferred or not, however, will depend on the binding site and on the contacts the odorant may make with that site (see figures 9.37, 9.38 and 9.39). Binding hinders motion of the odorant, reducing its effective symmetry and so the number of ways a receptor can "see" the odorant. Most odorants exhibit at least one polar group, the so-called osmophoric group or osmophore, [7] likely to make a hydrogen bond to the receptor. The combination of flexibility and the presence of more than one osmophoric group, as discussed in section 9.3.2, defines the odorant's non-superimposability, and thus its type. Important too, may be the time spent at the site. Lai *et al* [7] showed that when ligands are fitted to a binding pocket then, once natural vibratory motion was included, some odorants stayed within the pocket, whereas others moved out. Their results correlated with experiment, in that those that stayed within the domain activated the OR17 receptor, whereas the others did not.

Conformation change thus complicates simpler 'lock and key' models because it becomes important to consider which is the structure that the receptor actually recognizes.There are also complications for 'swipe card' models. In Turin's model [1], for instance, some odorant orientations favour electron inelastic tunnelling, and some not so. The first requirement for a favourable conformation is alignment of dipoles within the odorant to maximize the change in force as the electron tunnels; the rate and intensity of the signal is sensitive to this. The overall distance the electron tunnels should be small, for tunnelling rates drop exponentially with distance. The odorant must have a mode with the right vibrational frequency for the conformations that satisfy the other criteria. This mode, directly involved in actuation in this model, is probably distinct from the motions associated with flexibility. However, the motions associated with flexibility could have another role. In solid

state physics, there is a well-known role for what are termed promoting modes, which affect the rate of some processes. The promoting mode [66] is not the reaction coordinate, but makes a key step easier. In solid state diffusion, a promoting mode might open a gap, or channel, between atoms so a diffusing atom can migrate more easily. Fluctuations associated with flexibility may have a role in enabling the actuation event, perhaps in making inelastic tunnelling more probable. Of course, analogous enhancements might be possible for other actuation mechanisms.

9.5 Conclusions

Examination of a very large number of enantiomer pairs, with a proper concern for the likely accuracy of the data, shows a general rule: *The members of an enantiomer pair smell will alike (type 1) when the molecules are rigid, and will smell different (type 2) when they are flexible.* The flexibility referred to is not related to the geometric changes needed to deform one enantiomer into the other. This result is counter-intuitive both for shape theories and for those for which selectivity depends on some other factor.

It is possible that the processes leading to the type 2 differences could occur *en route* to the receptor, *e.g.*, by chirally-selective catalysis, or during diffusion steps through a chiral medium. We cannot rule this out, but the large number of type 2 pairs makes it seem unlikely. It is also possible that the critical step is docking at the receptor, where (for example) one member only of a flexible enantiomer pair might be able to wriggle into one or more receptors that the other member cannot access. This would lead to a different ensemble of signals being sent to the brain, and hence different smells. We cannot rule this out, but again the large number of type 2 pairs makes it seem improbable as the whole explanation of the rule.

Flexible molecules will be able to explore a wider range of conformations than their rigid counterparts. Our molecular dynamics calculations for free molecules (and small groups of them) confirm that this is so. What seems probable to us is that actuation of the receptor involves some of the rare conformations of the odorant that become possible for a flexible molecule, *i.e.*, fluctuations from the average shape are important. This would be analogous to other situations. Thus Luchinsky *et al* note that large, though infrequent, fluctuations are often the cause of important changes in many systems, for example 'nucleation at phase transitions, chemical

reactions, mutations in DNA sequences, protein transport in biological cells and failures of electronic devices' [120]. Further, as noted, the fluctuations associated with flexibility may act as promoting modes that enhance the electron tunnelling [66].

It seems likely that constrained fluctuations within the chiral receptor are critical. We have rather little information on the receptor structure at the appropriate level of detail. However, we can ask about the likely way the odorant will be tethered to the receptor, since this will limit some of the motions. If the left and right handed molecules do indeed enter the same receptors then, when there, the type 1 odorants must "look" the same. As discussed above, our particular concern is whether or not the two members of the enantiomer pair are superimposable or not, placing the active sites on the odorants in a precise way relative to the receptor.

If the odorant is only bound to the receptor at a single point, then the molecule will have quite a lot of freedom to adjust its position and orientation. There is a significant possibility that the two members of the enantiomer pair will be superimposable, in the sense just mentioned. Rigidity assists superimposability (for flexible molecules left and right handed forms might be deformed differently), and having only one site makes it insensitive to handedness. However, if the odorant is bound at two or more points, then it becomes more likely that the two members will not be superimposable. Many of the type 2 odorants have electronegative groups placed so as to make two-point binding likely. Thus conformational analysis must be considered in olfactory modelling because it appears the OR is extremely sensitive to the state of the molecule it is detecting.

The analysis supports the general "swipe card" picture of receptor actuation, in that actuation involves processes beyond the odorant having a good enough shape. It is surprising that molecular flexibility proves so crucial. The most probable reasons, we feel, suggest fluctuations of molecular shape, not just the average shape, though this must matter also, are very important.

Chapter 10

The future

To summarize the particular conclusions from the three investigations:

1. The 'swipe card' model as a signalling transduction mechanism in olfaction is entirely possible (possible without violating any laws of physics) and as such is an exciting new paradigm. Verification of the crucial parameters will be provided with establishment of the GPCRs and modelling rates will only become more reliable.

2. One of the parameters crucial to an accurate rate: the Huang-Rhys factor, is unfortunately difficult to non-ambiguously determine. Unless calculations avoiding the contentious issue of charge can be reliably established, some observable must be reasonably chosen: it would be a start to model a Huang-Rhys factor based on experimental Infrared spectrum.

3. One of the subtleties olfactory receptors seem to be sensitive to are the dynamics of molecules: in particular the chair, boat and twist like transitions of 6-membered rings. This could be tested with the synthesis and testing of new odorant molecules with constrained geometries.

In general there are many more questions in olfaction science yet to be answered, and a lot would be complemented with experiments. For example binding and functional assays and surface plasmon resonance to go with odorant-binding simulations. However, as noted at the beginning in the introduction, there is still so much theory can do. Especially since we are interested in what *one* environment (the

receptor type) sees differently about *many* ligands. Notoriously to the scientists perspective, see Charles Sell 'On the unpredictability of scent' for examples [7], we see very little, but the tools such as MD and DFT described in this thesis allow us to begin to see from the receptor perspective: at the nanoscale. Thus, for the sake of solubility and simplicity, we can (and have) made progress by first examining odorant behaviour in the isolated gas phase. This is a start, the ultimate aim of course being to model the dynamics of the system as a whole when such techniques are available. For now, we solve problems that we can!

Appendix

The Huang-Rhys factor is best understood with reference to the configuration coordinate, see figure 10.1.

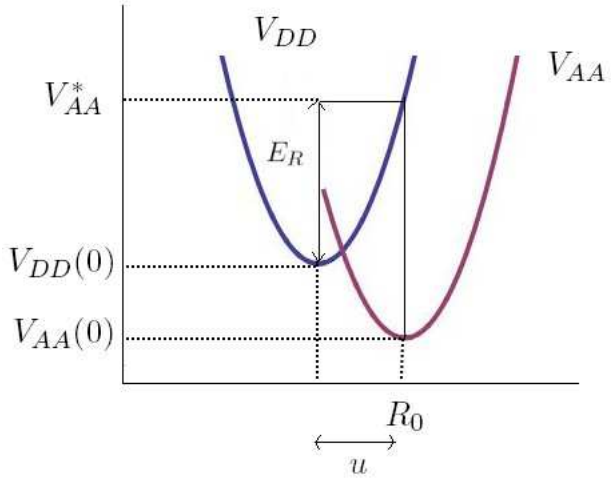


Figure 10.1: Configuration coordinate, V_{DD} and V_{AA}describe the donor and acceptor surfaces, λ is the relaxation energy we seek to calculate a Huang-Rhys factor.

Assuming the two states are the same, except displaced in energy and coordinates, we define $R_v = R_{0,v} + u_v$, where R is the reaction coordinate as usual, v indexes i atoms, with x, y, z Cartesian coordinates. $R_{0,v}$ is the relaxed coordinate, u_v

is the displacement from this equilibrium. Lets describe the lower energy surface:

$$V_{AA} = V_{AA}(0) + \frac{1}{2}\sum_{vv'} u_v K_{vv'} u_{v'}. \tag{10.1}$$

By definition the $\frac{\partial V_{AA}}{\partial R_{0,v}}$ term $\rightarrow$0 as this describes a minimum point. To describe the potential energy surface:

$$V_{DD} = V^*_{AA} - \sum_v F^*_v u_v + \frac{1}{2}\sum_{vv'} u_v K^*_{vv'} u_{v'}, \tag{10.2}$$

we make a key assumption that $K_{vv'} = K^*_{vv'}$, the spring constant matrices are the same for all surfaces. This assumption is OK, because of the weak coupling between the odorant and the receptor. We use Newton's Law:

$$m_v \ddot{u}_v = -\frac{\partial V_{AA}}{\partial u_v}$$

$$= -\sum_v K_{vv'} u_{v'} \tag{10.3}$$

For ease of calculation, we convert from displacements to mass-weighted coordinates:

$$\sqrt{m}_v \ddot{u}_v = -\sum_{v'} \frac{K_{vv'}}{\sqrt{m_v m_{v'}}} u_{v'} \tag{10.4}$$

So make a substitution $q_v = \sqrt{m}_v u_v$for mass-weighed coordinates. Then we introduce the dynamical matrix $D_{vv'}$by:

$$\ddot{q}_v = -\sum_{v'} D_{vv'} q_{v'} \tag{10.5}$$

$$D_{vv'} = \frac{K_{vv'}}{\sqrt{m_v m_{v'}}} \tag{10.6}$$

This is assuming harmonicity so:

$$q_v(t) = \hat{q}_v exp^{i\omega t} \tag{10.7}$$

$$\ddot{q}_v = -\omega^2 \hat{q}_v \tag{10.8}$$

Substituting 10.5into 10.8 yields a matrix eigenequation with eigenvalues $\omega$and eigenfunctions $\hat{q}_{v'}$ that can be solved computationally.

$$\omega^2 \hat{q}_v = \sum_{v'} D_{vv'} \hat{q}_{v'} \tag{10.9}$$

Say $\frac{F_v^*}{\sqrt{m_v}} = f_v$gives a mass-weighted force. Equation 10.2 reduces to:

$$V_{DD} = V_{AA}^* - \sum_v f_v q_v + \frac{1}{2} \sum_{vv'} q_v D_{vv'} q_{v'} \tag{10.10}$$

Now, we introduce *normal modes* which are linear combinations of the displacements, making them independently dynamic thus:

$$q_v = \sum_\alpha a_\alpha q_{\alpha,v}, \tag{10.11}$$

where α denotes the 3N-6 normal modes of vibration for non-linear molecules, a is a coefficient and $q_{\alpha,v}$ is the normal mode vector. The normal modes diagonalize $D_{vv'}$.

$$V_{DD} = V_{AA}^* - \sum_{\substack{v \\ \alpha}} f_v a_\alpha q_{\alpha,v} + \frac{1}{2} \sum_{\substack{vv' \\ \alpha\alpha'}} a_\alpha a_{\alpha'} q_{\alpha,v} D_{vv'} q_{\alpha' v'} \tag{10.12}$$

$$V_{DD} = V_{AA}^* - \sum_\alpha a_\alpha \left(\sum_v f_v q_{\alpha,v} \right) + \frac{1}{2} \sum_{\alpha\alpha'} a_\alpha a_{\alpha'} \left(\sum_{vv'} D_{vv'} q_{\alpha v} q_{\alpha' v'} \right) \tag{10.13}$$

So:

$$V_{DD} = V_{AA}^* - \sum_\alpha \tilde{f}_\alpha a_\alpha + \frac{1}{2} \sum_\alpha a_\alpha^2 \omega_\alpha^2 \tag{10.14}$$

In this way, cross terms are eliminated which is what makes these displacements *modes*. To obtain the Huang-Rhys factor we require the ratio of relaxation energy to the corresponding mode energy $\hbar\omega_\alpha$. The relaxation energy by definition corresponds to $V_{DD} - V_{DD}(0) = \lambda$.

At the vertical transition point from the minimum of $V_{AA}$to a point on the energy

surface V_{DD}, $V_{DD} = V^*_{AA}$and $q = 0$ so $a = 0$.

$$F_{DD,\alpha} = -\frac{\partial V_{DD}}{\partial a_\alpha} = \tilde{f}_\alpha - \omega^2_\alpha a_\alpha = 0 \tag{10.15}$$

Rearranging for: $a_\alpha = \frac{\hat{f}_\alpha}{\omega^2_\alpha}$ and substituting into 10.14 we get:

$$V_{DD} - V_{DD}(0) = -\sum_\alpha \frac{\tilde{f}^2_\alpha}{\omega^2_\alpha} + \frac{1}{2}\sum_\alpha \frac{\tilde{f}^2_\alpha}{\omega^2_\alpha} = -\frac{1}{2}\sum_\alpha \left(\frac{\tilde{f}_\alpha}{\omega_\alpha}\right)^2 \tag{10.16}$$

And so our dimensionless Huang-Rhys factor 'S' is given by this change in force and [118]:

$$S = \frac{\lambda_\alpha}{\hbar\omega_\alpha}. \tag{10.17}$$

Bibliography

[1] L. Turin. A spectroscopic mechanism for primary olfactory reception. *The Chemical Senses*, 21/6:773–791, 1996.

[2] R.W. Moncrieff. *The Chemical Senses*. NY: John Wiley and Sons, 1944.

[3] J.E. Amoore. The stereochemical theory of olfaction. *Nature*, 199:912–913, 1963.

[4] F. Yoshii, S. Hirono, and I. Moriguchi. Relations between the odor of (r) ethyl citronellyl oxalate and its stable conformations. *Quantitative structure-activity relationships*, 13:(2):144–147, 1994.

[5] R. C. Araneda, A. D. Kini, and S. Firestein. The molecular receptive range of an odorant receptor. *Nature Neuroscience*, 3, no 12, 2000.

[6] L. Turin and F. Yoshii. Struture-odor relations: a modern perspective. 2003.

[7] C. S. Sell. On the unpredictability of odor. *Angew. Chem. Int. Ed.*, 45:6254–6261, 2006.

[8] K. Mori and G.M. Shepard. Emerging principles of molecular signal processing by mitral/tufted cells in the olfactory bulb. *Semin. Cell. Biol.*, 5:65–74, 1994.

[9] G.M. Dyson. The scientific basis of odour. *Chem. Ind.*, 57:647–651, 1938.

[10] R.H. Wright. Odor and molecular vibrations: neural coding of olfactory information. *Journal of theoretical biology*, 64:473–502, 1977.

[11] A. Keller and L.B. Vosshall. A psychophysical test of the vibration theory of olfaction. *Nature Neuroscience*, Breif communication:1–2, 2004.

[12] C. Roberts. Psychopysical tests of the vibration theory of olfaction. *Chemical Senses*, 31, 2006.

[13] L.J. Haffenden, V.A. Yaylagan, and J. Fortin. Investigation of vibrational theory of olfaction with variously labelled benzaldehydes. *Food Chemistry*, 73:67–72, 2001.

[14] B. Malnic, J. Hirono, T. Sato, and L.B Buck. Combinatorial receptor codes for odors. *Cell*, 96:713–723, 1999.

[15] E. Tareilus, J. Noe, and H. Breer. Calcium signalling in olfactory neurons. *Biochem. Biophysics Acta*, 1269:129–138, 1995.

[16] N.E. Rawson and M. G. Gomez. Cell and molecular biology of human olfaction. *Microscopy research and technique*, 58:142–151, 2002.

[17] U. Wannagat, R. Munstedt, and U. Harder. *Leibigs. Ann. Chem.*, 950, 1985.

[18] L. Turin. A method for the calculation of odor character from molecular structure. *Journal of theoretical biology*, 7:367, 2002.

[19] D. Wallace, A.M. Testa, A. Harker, and A.H. Ramos. A new approach to the quantum modelling of biochemicals. *Molecular simulation*, 2(6):385–400, 1993.

[20] P.P.C. Graziadei. *Handbook of sensory physiology, chemical senses, part1, Olfaction, vol4.* New York Springer, 1971.

[21] M. M. Mozell. Evidence for a chromatographic model of olfaction. *The Journal of General Physiology*, pages 46–62, 1970.

[22] H. Breer. Olactory receptors: molecular basis for recognition and discrimination of odors. *Anal. Bioanal. Chem*, 377:427–433, 2003.

[23] L.C. Briand, C. Eloit, V. Nespoulos, J.-C. Huet, and C. Henry. Evidence of an odorant-binding protein in the human olfactory mucus: location, structural characterization, and odorant-binding properties. *Biochemistry*, 41:7241–7252, 2002.

[24] Simon Gane. Review article: something fishy in the state of olfaction science. *Pending publication*, 2008.

[25] L. B. Buck. Unravelling the sense of smell. *Nobel Lecture*, 2004.

[26] M. Teresa Moreno-Flores, J. Diaz-Nido, F. Wandosell, and J. Avila. Olfactory ensheathing glia: drivers of axonal regeneration in the central nervous system? *J. Biomed. Biotechnol.*, 2(1):37–43, 2002.

[27] L. Buck and R. Axel. A novel multi-gene family may encode odorant receptors. *Cell*, 65:175–187, 1991.

[28] A. Chess, I. Simon, H. Cedar, and R. Axel. Allelic inactivation regulates olfactory receptor gene expression. *Cell*, 78:823–834, 1994.

[29] R.R. Reed. After the holy grail: Establishing a molecular basis for mammalian olfaction. *Review Cell*, 116:329–336, 2004.

[30] F. M. Simoes de Souza and G. Antunes. Biophysics of olfaction. *Reports on Progress in Physics*, 70:451–491, 2007.

[31] V. Jacquier, H. Pick, and H. Vogel. Characterization of an extended receptive ligand repertoir of the human olfactory receptor or17-40 comprising structurally related compounds. *Journal of Neurochemistry*, 97:537–544, 2006.

[32] D. L. Nelson and M. M. Cox. *Lehniger: Principles of biochemistry*. W.H. Freeman and company, New York., 2005.

[33] F. A. W. Ebrahimi and A. Chess. Olfactory g proteins: Simple and complex signal transduction. *Current Biology*, 8:R431–R433, 1998.

[34] David. Colquhoun. Binding, gating, affinity and efficacy: The interpretation of struture-activity relationships for agonists and of the effects of mutating receptors. *British Journal of Pharmacology*, 125:923–947, 1998.

[35] L. Belluscio, G.H. Gold, A. Nemes, and R. Axel. Mice deficient in g(olf) are anosmic. *Neuron*, 20:69–81, 1998.

[36] E. Wallin and G. von Heijne. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.*, 7(4):1029–38, 1998.

[37] P.J. Loll. Membrane protein crystallization: The high throughput challenge. *J. Struct. Biol.*, 142:144–152, 2003.

[38] T. Fuchs, G. Glusman, S. Horn-Saban, D. Lancet, and Y.Pilpel. The human olfactory subgenome: from sequence to structure and evolution. *Human genetics*, 108:1–13, 2001.

[39] R. Lape, D. Colquhoun, and L.G. Sivilotti. On the nature of partial agonism in the nicotinic receptor subfamily. *Nature Articles*, 454:722–727, 2008.

[40] G. Sicard and A. Holley. Receptor cell responses to odorants: similarities and difference among odorants. *Brain Res.*, 292:283–296, 1984.

[41] K. Palczewski, T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Tong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, and M. Miyano. Crystal structure of rhodopsin: A g protein-coupled receptor. *Science*, 289(5480):739–45, 2000.

[42] Leffingwell and associates. http://www.leffingwell.com.

[43] M. S Singer and G. M. Shepard. Molecular modelling of ligand-receptor interactions in the or5 olfactory receptor. *Neuroreport*, 5:1297–1300, 1994.

[44] Peter C. Lai, Michael. S. SInger, and Chiquito J. Crasto. Structural activation pathways from dynamic olfactory receptor-odorant interactions. *Chemical Senses*, 30(9):781–792, 2005.

[45] B.D. Rubin and L.C. Katz. Spatial coding of enantiomers in the rat olfactory bulb. *Nature Neuroscience*, 4:355–356, 2001.

[46] A. Dafforn and D.E. Koshland JR. Theoretical aspects of orbital steering. *Proc. Nat. Acad. Sci*, 68:2463–2467, 1971.

[47] A. Fersht. *Enzyme structure and mechanisms*. W.H. Freeman and Company, 1985.

[48] Y.B. Chaim, B. Chanda, N. Dascal, F. Bezanilla, I. Parnas, and H. Parnas. Movement of 'gating charge' is coupled to ligand binding in a g-protein coupled receptor. *Nature Letters*, 2006.

[49] J. M. Berg, J.L. Tymoczko, and L. Stryer. *Biochemistry*. W.H. Freeman and company, New York, 2001.

[50] A. Lewis and L.V. Del Prioire. The biophysics of visual photoreception. *Physics Today*, pages 38–46, 1988.

[51] D. Restrepo, Y. Okada, J.H Teeter, L.D. Lowry, and B. Cowart. Human olfactory neurons respond to odor stimuli with an increase in cytoplasmis ca++. *Biophysics*, 64:1961–1966, 1993.

[52] H. Thurauf, M. Gjuric, G. Kobal, and H. Hatt. Cyclic nucleotide-gated channels in identified human olfactory receptor neurons. *European Journal of Neuroscience*, 8:2080–2089, 1996.

[53] A. Szabo and N.S. Ostlund. *Modern Quantum Chemistry: Introduction to advanced electronic structure theory*. McGraw-Hill, Inc, 1989.

[54] C.A. Coulson. *Valence*. Cambridge University Press, 1961.

[55] E. Merzbacher. *Quantum Mechanics 2nd Edition*. John Wiley and Sons, 1970.

[56] A.M. Stoneham. *Theory of defects in solids. Electronic structure of defects in insulators and semiconductors.* Oxford University Press, 1975.

[57] J.C. Decius E. Bright and P.C. Cross. *The theory of Infrared and raman vibrational spectra*. McGraw-Hill publishing company, 1955.

[58] M. Galperin, M.A. Ratner, and A. Nitzan. Inelastic electron tunneling spectroscopy in molecular junctions: peaks and dips. *The Journal of Chemical Physics*, 121:11965–11979, 2004.

[59] M.G. Simonsen and R.V Coleman. Inelastic-tunnelling spectra of organic compounds. *Phys. Rev B.*, 8:5875–5876, 1973.

[60] J. Lambe and R. C. Jaklevic. Molecular vibration spectra by inelastic electron tunneling. *Physical Review*, 165:821–832, 1968.

[61] D.S. Bendall, editor. *Protein electron transfer*. Bios Scientific publisher, 1996.

[62] R.A. Marcus. Chemical and electrochemical electron-transfer theory. *Annual Review of Physical Chemistry*, 15:155–196, 1964.

[63] K. Huang and A. Rhys. Theory of light absorption and non-radiative transitions in f-centres. *Proc. R. Soc. Lond.*, 204:406–423, 1950.

[64] P.H.-L. Sit, M. Cococcioni, and N. Marzari. Realistic quantitative descriptions of electron transfer reactions: Diabatic free-energy surfaces from first-principles molecular dynamics. *Physical review letters*, 97:028303, 2006.

[65] J. C. Brookes, F. Hartoutsiou, A. P. Horsfield, and A. M. Stoneham. Could humans recognize odor by phonon assisted tunneling? *Physical Review Letters*, 98:038101, 2007.

[66] A. M. Stoneham. Non-radiative transitions in semiconductors. *Rep. Prog. Physics*, 44:1251–1295, 1981.

[67] D. DeVault. *Quantum-mechanical tunnelling in biological systems*. Cambridge University Press, 2008.

[68] M. P. Allen and D. J. Tildesley. *Computer Simulation of liquids*. Oxford Science Publications, 1989.

[69] C.Cramer. *Essentials of computational chemistry theories and methods*. John Wiley and sons, 2002.

[70] S. L. Mayo, B. D. Olafson, and W. A. Goddard III. Dreiding: A generic force field for molecular simulations. *J. Phys. Chem.*, 94:8897–8909, 1990.

[71] R. G. Parr and W. Yang. *Density-functional theory of atoms and molecules*. Oxford Science Publications, 1989.

[72] O. Gunnarsson and B.I. Lundqvist. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Phys. Rev. B*, 13:4274, 1976.

[73] W. Kohn, A.D. Becke, and R.G. Parr. Density functional theory of electronic structure. *J. Phys. Chem.*, 100(31):12974–12980, 1996.

[74] Che M.S. Gannarelli. *Properies of the Earth's deep interior studied using ab initio modelling techniques*. PhD thesis, University College London, 2005.

[75] A. D. Becke. Density-functional thermochemistry. iii. the role of exact exchange. *J. Chem. Phys*, 98:5648–5652, 1993.

[76] J.P. Perdew. *Electronic structure of solids*. Akadenmic Verlag, Berlin, 1991.

[77] C.C. Trout and J.D.Kubicki. Deprotonation energies of a model fulvic acid i. carboxylic acid groups. *Geochemica et cosmochimica acta*, 70:44–55, 2006.

[78] A. St.-Amant, W. D. Cornell, and P.A. Kollman. Calcultion of molecular geometries, relative conformational energies, dipole moments, and molecular electrostatic potential fitted charges of small organic molecules of biochemical interest by density functional theory. *Journal of Computational Chemistry*, 16:1483–1506, 1995.

[79] D. Porezag and M.R. Pederson. Infrared intensities and raman-scattering activities within density-functional theory. *Physical Review B*, 54:no. 11, 1996.

[80] http://www.gaussian.com/home.htm.

[81] Mark A. Thompson. http://www.arguslab.com. ArgusLab 4.0.1, Planaria Software, LCC, Seattle, WA.

[82] T. van Mourik and H.A. Fruchtl. The potential energy landscape of noradrenaline: An electronic structure study. *Molecular Physics*, 103:1641–1654, 2005.

[83] J. W. Ochterski. Vibrational analysis in gaussian.

[84] Chunyang Peng, Phillippe Y. Ayala, H. Bernard Schlegel, and Michael J. Frisch. Using redundant internal coordinates to optimize equilibrium geometries and transition states. *Journal of Computational Chemistry*, 17:49–56, 1996.

[85] C. Peng and H. B. Schlegel. Combining synchronous transit and quasi-newton methods to find transition states. *Israel Journal of Chemistry*, 33:449, 1993.

[86] R.S. Mulliken. Electronic population analysis on lcao-mo molecular wave functions (i). *The Journal of Chemical Physics*, 23:1833–1840, 1955.

[87] B.H. Besler, K.M. Merz Jr, and P.A. Kollman. Atomic charges derived from semiempirical methods. *J. Comp. Chem.*, 11:431–439, 1990.

[88] H.B. Gray and J. R. Winkler. Long-range electron transfer. *PNAS*, 102:3534–3539, 2005.

[89] X. Song and R. A. Marcus. Quantum correction for electron transfer rates. comparison of polarizable versus nonpolarizable descriptions of solvent. *Journal of Chemical Physics*, 99(10):7768–7773, 1993.

[90] A. Stoneham and M.M.D. Ramos. Equilibria and electronic processes. *Journal of solid state chemistry*, 106:2–12, 1993.

[91] W. Hayes and A.M. Stoneham. *Defects and defect processes in nonmetallic solids*. Dover Publications Inc., 2004.

[92] G. P. Triberis and L. R. Friedman. A percolation treatment of the conductivity for the high-temperature small-polaron hopping regime in disordered systems. *J. Phys. C: Solid State Phys.*, 14:4631–4639, 1981.

[93] M. Elstner, Q. Cui, P.Munih, E. Kaxiras, T. Frauenheim, and M. Karplus. Journal of computational chemistry. *Modelling zinc in biomolecules with the self consistent charge-density functional tight binding (SCC-DFTB) method: Applications to structural and energetic analysis*, pages 565–581, 2003.

[94] T. Creighton. *Proteins: structures and molecular properties*. W. H. Freeman, 1992.

[95] T.W. Graham Solomons. *Organic Chemistry*. John Wiley & Sones, 1995.

[96] V. Bhandawat, J. Reiset, and Yau. K-W. Elementary response of olfactory receptor neurons to odorants. *Science*, 308:1931–1934, 2005.

[97] I. Boekhoff, K. Touhara, S. Danner, J. Inglese, M. J. Lohse, H. Breer, and R.J. Lefkowitz. Phosducin, potential role in modulation of olfactory signalling. *The journal of biological chemistry*, 272:4606–4612, 1997.

[98] L.A. Blumenfeld and D.S. Chernavskii. Tunneling of electrons in biological processes. *J. theor. Biol.*, 39:1–7, 1973.

[99] A. Blondel, J.P. Renaud, S. Fischer, D. Moras, and M. Karplus. Retinoic acid receptor: A simulation analysis of retinoic acid binding and th resulting conformational changes. *Journal of Molecular Biology*, 291:101–115, 1999.

[100] Jens Ulstrup. *Lecture notes in chemistry*. Springer-Verlag, 1979.

[101] Andreas Keller and Leslie.B. Vosshall. Influence of odorant receptor repertoire on odor perception in humans and fruit flies. *Proceeding of the National Academy of the United States of America*, 104:5614–5619, 2007.

[102] L.A. Curtiss, S.R. Langhoff, and G.D. Carney. Ab initio scf and cl calculations of the dipole moment funciton of ozone. *J. Chem. Phys*, 71(12):5016–5021, 1979.

[103] G.P. Li and I.P. Hamilton. Infrared and raman spectra for complexes of small cationic gold clusters and hydrogen sulphide. *Canadian journal of analytical sciences and spectroscopy.*, 49:no.5, 2004.

[104] H. H. Nielsen and E. F. Barker. Infrared absorption bands in hydrogen sulfide. *Physical Review*, 37:727–732, 1931.

[105] A.D. Sprague and H.H. Nielson. The infrared absorption by hydrogen sulfide. *The Journal of chemical physics*, 5:no 2, 1937.

[106] R. L. A. Haiduke and R.E. Bruns. An atomic charg-charge flux-dipole flux atom-in-molecule decomposition for molecular dipole moment derivatives and infrared fundamental intensities. *Journal of Physicsl Chemistry A.*, 109:2680–2688, 2005.

[107] A. V. Larin, A.N. Trubnikov, and D. P. Vercauteren. Differences between the co and no properties for stability of alkali metal complexes me(xo)n+, x=c or n. *International Journal of Quantum Chemistry*, 90:541–548, 2002.

[108] Francis Crick. *The Astonishing Hypothesis, The Scientific Search for the Soul.* Touchstone, 1995.

[109] E. T. Theimer, T. Yoshida, and E. M. Klaiber. Olfaction and molecular shape. chirality as a requisite for odor. *Journal of Agricultural and Food Chemistry.*, 25:1168–1177, 1977.

[110] R. Bentley. The nose as a stereochemist. enantiomers and odor. *Chemical Reviews*, 106:4099–4112, 2006.

[111] D. A. Dixon and A. Komornicki. Ab initio conformational analysis of cyclohexane. *Journal of Physical Chemistry*, 94:5630–5636, 1990.

[112] Eusebio Juaristi, editor. *Conformational behavior of six-membered rings, analysis, dynamics, and steroelectronic effects*. VCH publishers, 1995.

[113] E. L. Eliel, N.L. Allinger, S. J. Angyal, and G. A. Morrison. *Conformational Analysis*. John Wiley and Sons, 1965.

[114] Anton Beyer and Peter Schuster. Internal dynamics of flexible molecules: Cyclohexane. *Monatshefte fur Chemie Chemical Monthly*, 121:339–349, 1990.

[115] E. L. Eliel and S. H. Wilen. *Stereochemistry of Organic Compounds*. John Wiley and Sons, 1993.

[116] P. W. Day, S. G. F. Rasmussen, C. Parnot, J. J. Fung, A. Masood, T. S. Kobilka, X. J. Yao, H. J. Choi, W. I. Weis, D. K. Rohner, and B. K. Kobilka. A monoclonal antibody for g-protein-coupled-receptor crystallography. *Nature Methods*, 4:927–929, 2007.

[117] A. Stojanovic, J. Stitham, and J. Hwa. Critical role of transmembrane segment zinc binding in the structure and function of rhodopsin. *Journal of Biological Chemistry*, 279:35932–35941, 2004.

[118] M. Allan, A. C. Craig, and L. V. McCarty. Vibrational excitation of cis- and trans-1,2-difluoroethenes by electron impact; effect of dipole moment on the threshold peaks. *Journal of Physics B. Atomic molecular and optical physics*, 35:523–532, 2002.

[119] M. Laska, D. Genzel, and A. Wieser. The number of functional olfactory receptor genes and the relative size of olfactory brain structures are poor predictors of olfactory discrimination performances with enantiomers. *Chemical Senses*, 30:171–175, 2005.

[120] D. G. Luchinsky, P. V. E. McClintock, and M. I. Dykman. Analogue studies of nonlinear systems. *Rep. Prog. Phys.*, 61:889–997, 1998.