1 General Introduction

1.1 Anthropogenic CO_2 causes climate change

The anthropogenic caused climate change is nowadays seen as the biggest threat to humanity[1] and is publicly discussed at least since the mid- to late 1980s[2], although scientifically published a decade earlier[3]. To put this into perspective: the internet[4] and the polymerase chain reaction (PCR)[5] are as old as it is public knowledge that human greenhouse gas emissions cause a dangerous rise in global temperatures. To gather information and communicate causes, effects, and proposals to politicians and the public, the United Nations (UN) founded the Intergovernmental Panel on Climate Change (IPCC) in 1988[6]. The latest IPCC report[7] and its summary for policymakers[8] finds clear and drastic words for the climate change that has already happened: mankind caused the increase in greenhouse gas concentrations in the atmosphere, which increased the surface temperature by 1.09 °C in comparison to 1850-1900. This increase cannot be explained by other phenomena than anthropogenic greenhouse gas emissions. Global see levels have already risen 0.2 m since. These changes will worsen and are irreversible for centuries.

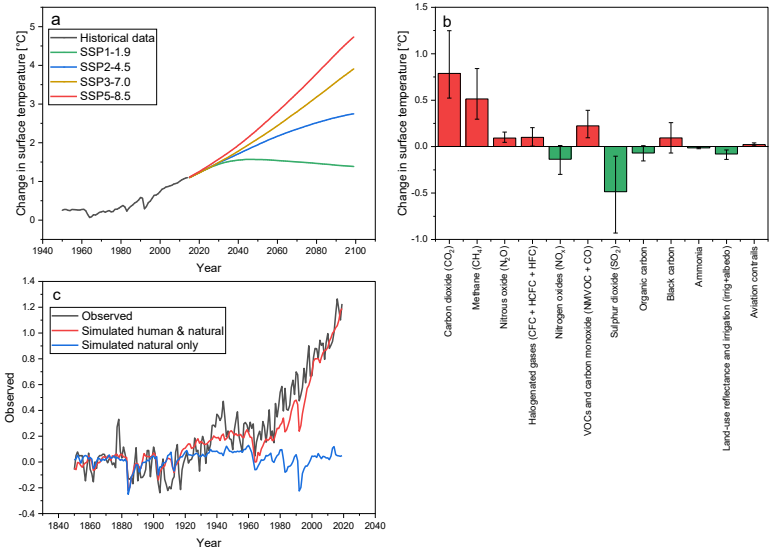

Figure 1: CO_2 is the main contributor to human-caused climate change. Change in surface temperature relative to 1850-1900 (black), and simulations regarding the change due to anthropogenic (red) and natural causes (blue) (a), simulations of surface temperature change in the next 80 years based on greenhouse gas emission scenarios (b), and temperature changes caused by individual molecules (c). Data from [7] with individual data sets [8–10].

1.2 Biotechnology has a key role in minimising CO_2 emissions

With anthropogenic global greenhouse gas (GHG) emissions identified as the sole cause of climate change, the reduction of primarily CO_2 up to climate neutrality is an ongoing political goal[12]. The focus of discussions is the de-fossilisation of the electricity production[13] and the transport sector, where in the latter battery electric vehicles and biofuels are prominent research foci[14]. Often overlooked is the chemical industry, which produces around 7 % of GHG[15] and uses around 8 % of total fossil fuel production as feedstocks[16], which accounts for over 500 Mt/a in form of coal (55 Mt/a), natural gas (199 Mt/a) and liquid oil products (260 Mt/a)[15]. The carbon from these fossil feedstocks will sooner or later be released into the atmosphere as CO_2 and thus will contribute to global warming. One aim of biotechnology is to use renewable feedstocks, allowing carbon-neutral product cycles.

As an example, the most important biotechnological process is the *Saccharomyces cerevisiae*-based production of bioethanol with 88 Mt/a[17]. This process is efficient with > 90 % theoretical yield, fast with > 1.0 g/L/h, and the end product can accumulate > 100 g/L, facilitating purification[18]. Bioethanol can be mixed with gasoline and thus be implemented directly into the transport sector energy chain, which alone reduced CO_2 emissions by 43.5 Mt/a[19]. Albeit, this becomes minuscule comparing it to the transport sector CO_2 emissions of over 7,000-8,000 Mt/a[20]. First-generation biofuels use C_6 sugars mostly derived from corn as biomass feedstocks, which puts them, even at this low overall contribution to CO_2 emission reduction, into direct competition in using these plants and cropland for food. The emerging food-vs-fuel debate gave rise to the development of second- and third-generation biofuels[21]. Second-generation biofuels use a wide range of otherwise unused feedstock such as municipal waste or lignocellulosic biomass, while third-generation fuels aim to use CO_2 as a carbon source, for example by using algal biomass[22].

To alleviate the problems arising in using land- or water-grown biomass as feedstocks, the use of industrially producible C1-compounds as feedstocks gained attention[23,24]. This group consist of methane (CH_4), methanol (CH_4O), formaldehyde (CH_2O), formate (CH_2O_2), and carbon dioxide (CO_2) or monoxide (CO), which both need H_2 as an electron donor. These can be produced industrially, cheaply, and easily from CO_2 and green H_2. But established production strains such as *S. cerevisiae* or *E. coli* are not able to use gaseous substrates (methane, CO_2/H_2)[25]. Formate contains only very low energy and causes problems with its acidity. The use of methanol is extensively explored[26], but its toxicity and volatility cause problems in lab-scale experiments[27].

1.3 Molecular basis of volatility[28–30]

Volatility, *in general*, is often described as a compound's likelihood of vapourising or existing in the gas phase. From this ill-defined term, one is often redirected to the

vapour pressure for measurements. The vapour pressure is defined as the pressure of a vapour over its liquid or solid phase in a closed environment if both are in thermodynamic equilibrium. The vapour pressure of a compound can in simplest terms be dependent on temperature and pressure, but following Gibb's Phase rule, the number of dependencies increases with the number of compounds in the system:

$$F = C - P + 2 \qquad \textbf{Equation 1}$$

$$p = p(T, x) \qquad \textbf{Equation 2}$$

With F being the number of degrees of freedom (i.e. number of dependencies), C being the number of components, and P being the number of phases. This means in a system with two components (water and a metabolite) and two phases (liquid and gas phase), two degrees of freedom exist, which means the vapour pressure (p) is dependent on temperature (T) and concentration of the metabolite (x).

For the calculations in solutions, the vapour pressure is derived from the molar concentration (x) and the saturated vapour pressure (p^{sat}, see Equation 3), which can be determined experimentally in one compound systems. The p^{sat} is an exponential function and can thus be described by Equation 4. Except for the temperature (T), p^{sat} is dependent on 3 parameters, the Antoine parameters. For numerous compounds, these are already determined and can be accessed for example in the NIST database[31].

$$y \cdot \Phi \cdot p = x \cdot \gamma \cdot p^{sat} \qquad \textbf{Equation 3}$$

$$p^{sat} = 10^{A - \frac{B}{C+T}} \qquad \textbf{Equation 4}$$

$$\textbf{With } x \to 1 \quad p = x \cdot p^{sat} \qquad \textbf{Equation 5}$$

$$\textbf{With } x \to 0 \quad p = x \cdot H \qquad \textbf{Equation 6}$$

In an idealised system, the pointing factor Φ and y can be disregarded. Having a solution with a high concentration of analyte (x→1), the activity coefficient becomes 1, giving Raoult's law (Equation 5), with an infinite dilution (x→0), the activity coefficient approaches a constant value, which can be expressed in combination with the saturated vapour pressure as the Henry coefficient (Equation 6). Although interdependencies in multi-component mixtures exist, for infinite dilutions these become negligible.

This leads to the vapour pressure being an inherent property of a compound. The higher the intermolecular forces in the system, the lower is the vapour pressure, *in general*. With hydrogen bonding, dipole forces, and van der Waals forces, *in general*, larger molecules with more polar functional groups are less volatile. But with counteracting effects this is hard to predict and should be measured whenever possible[32]. But as different fields of natural science need to deal with a large number of mole-

cules, structure-based prediction of vapour pressure and the related parameter is inevitable[32–34]. Most of these calculations are based on group contribution methods (GCMs), where properties of a base molecule are expanded with properties of a functional group, whose effect was experimentally determined with a different molecule[33]. An example could be the prediction of properties for 1,4-heptanediole, which could be calculated based on the properties of 1-heptanol (base group), as well as 1-hexanol, and 1,4-hexanediol (group contribution).

1.4 Volatile organic compounds and the smell of yeast

The variety of slightly different definitions for the term volatile organic compounds (VOCs) reflects the range of scientific fields where these are of interest. Commonly, VOCs are described to contain 15 or fewer carbon atoms and having a vapour pressure greater than 10 Pa at 25 °C[35]. This would include carbon dioxide and methane, which are often viewed separately based on their importance and relatively high abundance (see chapter 1.1).

Besides microbial VOCs, volatile organic compounds are of interest in diverse fields of industry including those producing materials emitting VOCs (building material[36] or plastics[37,38]), those measuring the VOCs impact on human health[39,40], and those providing air filtering systems[41]. Natural and anthropogenic hydrocarbon emissions are of interest to atmospheric scientists[35,42] and VOCs measured in human exhalations are used as a biomarker for, e.g., lung cancer[43]. Even narrowing down the field to microbial produced VOCs leaves hundreds of organisms, metabolic pathways, and diverse classes of organic compounds. Some VOCs are purposefully produced as semiochemicals because they can act more easily over prolonged distances and are not limited by the body of liquid the microbes reside in. But arguably most of the VOCs are produced involuntarily as metabolites and intermediates just happen to have sufficiently high vapour pressure but distinguishing between correlation and causation is nearly impossible after millions of years of evolution and symbiosis[44].

1.4.1 VOCs derived from *S. cerevisiae*'s central carbon metabolism

In *S. cerevisiae*'s central carbon metabolism alone, five volatile organic substances are found as shown in Figure 2: pyruvate, acetaldehyde, acetate, ethanol, and carbon dioxide. As discussed here, their volatility seems to be accidental, as the yeast does not seem to use these as semiochemicals, although insects and animals use ethanol and its precursors as an indicator for high-energy ripe fruits[45,46]. Monitoring the levels of pyruvate, ethanol, and acetate can give insights into the cell's energy production, redox status, and pH stress, which can be beneficial for industrial processes.

Pyruvate is the first component of central carbon metabolism, which has a vapour pressure above 10 Pa. Because pyruvate plays a central role in the energy metabolism, it seems unlikely that it is purposefully produced as a volatile. For the same reason, to the best of my knowledge, it does not seem to act as an infochemical. But

pyruvate is described to have an acetate-like smell and is, therefore, a contributor to food smell and taste[47].

Carbon dioxide, as mentioned above is mostly dismissed because it is odourless to humans and is present in high concentrations in air. But microbial CO_2 is a factor in the production of many foods, e.g., sparkling wine, bread, cake, and cheese as well as for special packaging[48]. In process control, CO_2 rates are closely monitored because carbon dioxide accumulation can lead to acidification of the medium due to bicarbonates (see Equation 7). CO_2 can also negatively influence biotechnological production due to end-product inhibition, it negatively affects membrane biosynthesis, and was shown to reduce the performance of aerobic cultivations by up to 50 %[49]. But CO_2 in anaplerotic reactions is vital, as *S. cerevisiae* carbonic anhydrase knock-out mutants proved to be lethal[50]. It can even function as a signal molecule as for example *Drosophila* is attracted by higher CO_2 concentrations[51].

$$CO_{2_{gas}} \leftrightarrow CO_{2_{liquid}} \overset{H_2O}{\longleftrightarrow} H_2CO_3 \leftrightarrow H^+ + HCO_3^- \leftrightarrow H^+ + CO_3^{2-} \quad [49] \qquad \textbf{Equation 7}$$

Acetaldehyde is the intermediate in both ethanol and acetate production and acts respirotoxic[52]. It is at the same time classified as cancerogenic but also an often-used food additive due to its fruity flavours[53,54]. Because of its industrial relevance, microbial overproduction of this compound has been researched since the 1980s and is also part of this thesis and discussed later (see chapter 3.2, page 49). Furthermore,

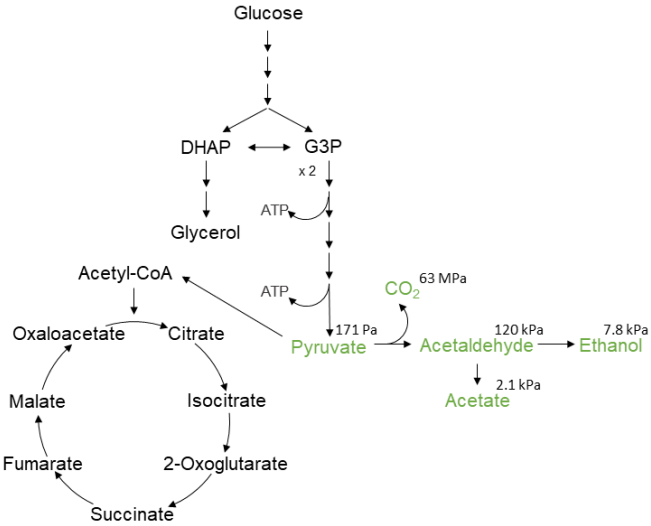

Figure 2: Yeast central carbon metabolism with vapour pressures. Marked in green are those which fit the volatile organic compound (VOC) definition and will contribute to the volatilome. Vapour pressures are just shown if < 10 Pa and were taken from [59].

acetaldehyde acts as an attractant at least for insects, speculatively because it indicates the presence of glucose-rich fruits[55,56].

Acetate is produced either from cleaving acetyl groups with the special case of acetyl-coenzyme A (acetyl-CoA) or directly from the enzymatic conversion of acetaldehyde. While the effect of acetate and the microbial usage of this C2-compound is thoroughly studied, its production remains to be fully elucidated[57]. It is often regarded as an overflow metabolite with a similar function as ethanol in Crabtree positive yeasts, but it is also known that acetate production levels increase multiple fold upon an increased pH[58]. Verduyn et al. account this to an increased acetate efflux[58]. But the enzymes ALD4-6, which catalyse acetaldehyde to acetate conversion, are also upregulated under alkaline stress[59]. This could hint that acetate is also produced for pH regulation. While there is much literature about acetate esters[60], acetate in itself seems not to be discussed as a semiochemical, maybe for the same reason as pyruvate.

Ethanol is arguably the oldest[61] and today the most important product of not only *S. cerevisiae*, but also the whole food and biotech industry[18]. Metabolically, ethanol is produced from acetaldehyde to recycle NADH if the glycolysis and the tricarboxylic acid (TCA) cycle outruns the respiration capacity, which occurs under O_2 deficiency or high glucose concentrations in Crabtree-positive yeasts[62]. Globally, there are three structurally independent groups of ADHs which are responsible for ethanol breakdown, and they are found in all phyla of life, hinting at the dietary importance of ethanol-containing food[63]. There are many behavioural studies regarding this compound ranging from *Drosophila*[45] and beetles[46] to monkeys[64] and humans[65].

1.4.2 VOCs derived from *S. cerevisiae*'s amino acid metabolism

The amino acid degradation pathways are interconnected and often at least partly redundant. A simplified representation is given in Figure 3. The first step is predominantly oxidative deamination via the central α-ketoglutarate/glutamate cycle, yielding α-keto acids. In humans, for example, all amino acids can further be degraded into intermediates in the central carbon metabolism[66]. In *S. cerevisiae* on the other hand seven amino acids, namely leucine, isoleucine, valine, methionine, tryptophane, tyrosine, and phenylalanine are degraded via the Ehrlich pathway[67]. In general, much about the biological functions and mechanisms of the Ehrlich pathway are not completely resolved. Here, after the initial transamination, an irreversible decarboxylation to a fusel aldehyde takes place. The aldehyde is subsequently either oxidised or reduced to its corresponding fusel acid or alcohol, whereby the ratio strongly depends on the cultivation conditions. While in aerobic, glucose-limited cultures with multiple N-sources, mostly fusel acids are produced, in fermentative conditions, the fusel alcohols dominate[67].

Not using the amino acids' carbon backbones in the central carbon metabolism seems costly, but two main reasons are often given for the existence of this pathway. On the one hand, the transamination reaction is close to thermodynamic equilibrium,

thus a high flux for maximum nitrogen acquisitions is made possible by removing reactions products, i.e. α-keto acids, by a thermodynamically favoured decarboxylation. On the other hand, the variable oxidation/reduction step in the Ehrlich pathways is argued to help maintaining redox equilibrium[67,68]. Further, as fusel alcohols are produced under amino acid starvation they are also involved in the translation regulation[69].

Fusel alcohols generally have a pleasant odour and at least some have sufficiently high vapour pressures to be classified as VOCs[70]. Many occur widely in nature and have therefore diverse functions in inter-and intraspecies communication. Biotechnological applications include the use in food and fragrance industries but also as biofuels and feedstocks as in the case of isoamyl alcohol[71]. Fusel acids on the other hand are easily converted with acetyl-CoA into ethyl esters[72]. The biological function of ester formation is not yet well understood, but there are four main hypotheses for this pathway in yeast, which not only includes Ehrlich pathway derived fusel acids, but also medium-chain fatty acids as described later (see chapter 3.4). This process could be used for detoxification to ensure intracellular pH homeostasis. Because *S. cerevisiae* cannot produce unsaturated fatty acids (UFAs) under anaerobic conditions, the esters could be used as UFA analogous to maintain membrane fluidity or simply to free CoA stuck as acyl-CoA in UFA synthesis. Lastly, these ethyl esters have been proven to be a strong attractant for *D. melanogaster*, which help dissipate yeasts from fermenting food[60].

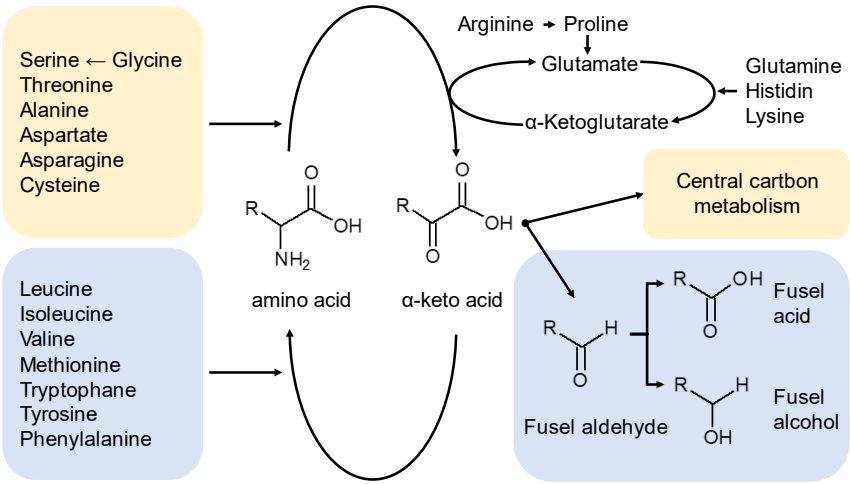

Figure 3: Schematic amino acid degradation pathways in *S. cerevisiae*. While in higher animals all amino acids are degraded to intermediates in the central carbon metabolism, in *S. cerevisiae* seven amino acids are degraded via the Ehrlich pathway to fusel alcohols or fusel acids.

1.4.3 VOCs derived from other *S. cerevisiae* pathways

Fatty acids (FA) in *S. cerevisiae* are mainly monounsaturated C_{16} and C_{18} fatty acids, but there is a considerable spectrum of diverse fatty acids in lower amounts[73,74]. The fatty acid synthesis requires acetyl-CoA, ATP, and NADH, while the β-oxidation again requires ATP but produces FADH, NADH, and acetyl-CoA. Further, O_2 is required for the introduction of a double bond to produce unsaturated fatty acids[75]. In theory, the complete synthesis or degradation of FA does not yield any volatile compounds, but in reality, these processes are often incomplete and free FA are released into the cell. Fatty acids are again easily converted to esters for the same reasons as fusel acids are (see 1.4.2, page 8). But instead of free carboxylic acids and acetyl-CoA, free ethanol and CoA-bound fatty acids are used, but compared to the acetate ester formation, the ethyl ester formation is only sparsely researched. Key enzymes are the ethanol hexanoyl transferase (Eht1) and ethyl ester biosynthesis gene (Eeb1)[60,72]. Interestingly, although their deletion lowers the ethyl ester content in yeast, their overexpression, even in combination with increased substrate availability, does not increase ethyl ester production[72]. Mainly ethyl hexanoate, -octanoate and -decanoate are produced, which smell fruity or apple-like[72]. Therefore, they are important contributors to the smell and taste of *S. cerevisiae* fermentation products.

Terpene derived compounds are the most abundant biogenic VOCs in the atmosphere with the prime example isoprene[76]. But their spectrum in yeast is quite limited and key enzymes for the biosynthetic pathways could not yet be experimentally proven[68,77]. Terpenes are important flavour molecules in beer and wine, but their *de novo* synthesis by yeast contributes only minor amounts, in contrast to the grapes used, which can contain free, but mostly glycosidically bound terpenes. During wine ageing, these terpenes can be released either through acidic or catalysed hydrolysis by *S. cerevisiae*. Furthermore, yeasts are known to catalyse conversions of, e.g., geraniol to *β*-citronellol (reduction) or geranial (reduction) as well as acetate ester formation of terpenoids[78].

Hydrogen sulfide can be formed by yeasts from inorganic sulfur compounds such as sulfate. After internalisation, SO_4^{2-} is reduced under ATP consumption to HSO_3 to be sequestered with serine or homoserine to methionine and cysteine. Under nitrogen limitation, these compounds are not present in sufficient quantities and a pH-driven reduction of HSO_3 yields free H_2S, which can diffuse out of the cell. Hydrogen sulfide can in turn react with other *S. cerevisiae* products to form volatile organic thiols such as ethanethiol[79,80].

1.5 Analytical methods for exploring volatile metabolites of *S. cerevisiae*

Possible methods for analysing volatile metabolites can be divided into two groups: measuring the liquid phase or measuring the gas phase. Using the first has

the advantage that liquids, under normal temperature and pressure, contain around 1,200 times more molecules per volume than an ideal gas (55.5 mol/L for water and 0.04 mol/L for an ideal gas). This means, that gas-phase analytics must be multiple orders of magnitude more sensitive to obtain the same results. Further, handling and storage of liquid samples is much simpler than with gas-phase samples. But using liquid chromatography (e.g., HPLC) for volatilome analysis, has two major problems: 1) the non-volatile matrix can mask the dilute volatiles, and 2) the gas-phase concentrations of compounds with sufficiently high *theoretical* vapour pressures remain uncertain. Therefore, the true volatilome can only be measured by gas-phase methods. In general, every analysis step as for example sampling or chromatography facilitates compound identification but distorts volatile concentrations.

1.5.1 Gas chromatography-based methods

Gas chromatography (GC) was invented in 1952 and is still one of the standard analytical techniques today[81]. Most often liquid injection GC is used, where the liquid sample is evaporated using high temperatures in the injection port, while this can be skipped during the injection of gaseous samples. Recently, the use of solid-phase microextraction fibres (SPME) gained attention as a third alternative[82]. The fibres are loaded with volatile molecules in the headspace of the sample and then either stored or directly moved to the GC, where the molecules are desorbed using high temperatures[83].

The sample is then led through a capillary column by a mobile phase of either helium, hydrogen, or nitrogen and the analytes are separated by their affinity to the stationary phase. The higher the affinity the longer the retention time. There are a plethora of different techniques modifying this easy principle beginning by using different columns[84] over two-dimensional GC (GC×GC)[85] up to chemical derivatisation[86]. Usually, either flame ionisation detector (FID) or mass spectrometry (MS) are used as detectors. The first is comparatively easy, cheap, and completely sufficient for routine analysis[87], but MS offers the additional information of the analyte's mass and in MS/MS setups can even identify the compound without the need for external standards.

1.5.2 Ion mobility spectrometry (IMS)

The IMS has also been in use for decades and its physical principles, use, and evolution are already reviewed conclusively[62,88–90]. IMS relies on the principle of sample injection, analyte ionisation and separation, followed by detection. While radioactive sources are commonly used, other ionisation techniques, which are more or less congruent with those used in other MS techniques, can also be applied (see 1.5.3 for more details). The ions are led into the drift tube, where they move due to an electric field and are hindered by the drift gas, for example, nitrogen, moving in the opposite direction. Each molecule has a specific ion mobility (K), which is on the one hand influenced by instrument settings like the electric field strength and on the

other hand on molecule-specific properties like reduced mass and collision cross section[88]. Depending on the specific ionisation method used, the drift gas and residual water can also be ionised, forming so-called reactant ion peaks (RIPs). The presence of molecules with a higher proton affinity than water quenches the RIP, and the RIP could potentially overlay analytes with the same drift time at the detector. As a detector, often a simple Faraday's plate or cup is used, where the ions transfer their charge. The metal is then discharged to measure a current proportional to the number of ions[91]. This detector type is simple and inexpensive but does not provide qualitative information beyond the number of ions per second. Therefore, other electrodes that do not stop the ion beam[88] and MS have also been explored[92].

1.5.3 Mass spectrometry-based methods

MS-based methods are now the standard of measuring not only the volatilome but in most explorative methods including liquid phase metabolomics. MS can be used as detectors in already established methods like in GC-MS, or as a standalone technique like the SESI-Orbitrap MS[93,94]. Most techniques ionise the sample, after optional chromatographic separation, directly prior to MS analysis. Notably, exceptions to this are SIFT MS and PTR-ToF-MS, which are explained later.

Important variables for the assessment of MS techniques are dimensionless mass to charge ratio *m/z*, and the also dimensionless resolution or resolving power of the MS: *R*. This is often defined as the *m/z* of a peak divided by its full-width half-height maximum (FWHM)[95].

1.5.3.1 Ion sources

Giving a conclusive review of all possible ionisation techniques is far out of scope for this work. A good overview of different techniques is given here [91,96,97]. A general concern for ionisation methods used in gas-phase MS is the ionisability of the matrix, mostly consisting of ambient air N_2, O_2, CO_2, or H_2O. The non-ionisation of these compounds has the same effect as a major concentration gain for the analytes, although this could mean losing some information on the composition of the gas phase.

Spray ionisation techniques: Most of the methods are based on electrospray ionisation (ESI), where the liquid sample is nebulized by forcing it through a capillary on which high voltages, positive or negative, are applied[98]. Sheath gas or liquid can be applied to aid in directing the forming droplets from the tip into the gas phase and subsequently into the detector. Low flow rates, surface tensions and electrolyte concentrations are advantageous, but increasing voltages can compensate for unfavourable conditions, up to a certain point. It should be mentioned, that in positive mode, negative charges are removed from the droplets rather than positives charges are introduced. After the droplet leaves the tip of the electrospray capillary, it begins to shrink in size due to the evaporation of the solvent. Two models have been postulated about the exact mechanism going from charged cluster to single ions[99]. In the ion evaporation model, the droplets shrink until the point where the coulombic repulsions

inside the cluster overcome the surface tension, causing the release of solvated ions. The charge residue model on the other side assumes that each droplet only contains one analyte ion, which is released only due to solvent evaporation.

Secondary electrospray ionisation (SESI) was refined from ESI sources to aid gas-phase MS. In the development, the liquid sample stream was replaced with a simple electrolyte solution, e.g., formic acid or silver nitrate, in a commercial ESI source, into which human breath was blown[100]. In modern SESI units, the sample gas stream is led through the electrospray. While the mechanisms of the water cluster formation are still debated, Tejero Rioseras et al. could show, that the analyte ionisation happens in the gas phase through ion-molecule interactions[101].

Plasma desorption techniques: Atmospheric pressure chemical ionisation (APCI) was developed in the 1970s[102]. As with ESI, the liquid is nebulized, but without charging the droplets. With N_2 as carrier gas and atmospheric water, the corona discharge electrode produces mostly H_3O^+ ions, which then collide with the analytes transferring their charge. APCI outperforms ESI at high flow rates and has a better tolerance against salts and buffers. Similar to the electrospray-based methods, APCI can be used to measure the headspace of samples online, as reported, e.g., by Aznar et al.[103].

Direct analysis in real time (DART) is also a plasma-based technique related to APCI. First, electronically excited neutral helium atoms are produced and then shot onto the liquid, solid, or gaseous sample. The ionisation of analyte molecules is not achieved through the primarily produced exited neutral atoms, but reactant ions, typically protonated water clusters formed from laboratory air. The beauty of this method is, that solid objects can be tested without previous sample preparation[104]. But also gas-phase measurements are possible, as shown by, e.g., Li[105] or Busman et al.[106].

Photoionization techniques: Atmospheric pressure photoionisation (APPI) is the youngest of the three API methods and is derived from APCI. Again, the analyte stream is nebulized, but the ionisation takes place by photoionisation directly instead of collision with previously generated ions. Commonly, krypton lamps are used, which have two discrete photon energy lines at 10.0 and 10.6 eV[96]. This can be used to selectively ionise analytes and leave out air and solvent molecules, which makes this method less vulnerable to matrix effects[107]. APPI was used for volatilome measurements in the same way as the methods already described above, e.g. by Zhou et al.[108]

1.5.3.2 Detectors

As for the ionisation sources, presenting all MS detector types is out of scope for this work and again a good overview is given by Gross[91].

Time of flight[91,109] (ToF) detectors are arguably the oldest widespread used detectors, constructed in the mid-1940s and commercially available in the mid-1950s. An ion package is accelerated in an electric field and depending on their mass-to-charge ratio (*m/z*), the ions reach different velocities. Lower *m/z* ions will traverse the

subsequent field-free drift region faster. While earlier instruments only had a linear flight path, modern instruments can have a single or multiple reflectors, which increases the flight time difference of the ions drastically. Because faraday cups as used e.g. with IMS are not sensitive enough, secondary electron multipliers (SEMs), are used. Here, upon impact of an ion, multiple electrons are emitted from a metal or semiconductor and focused by an electrode to the next multiplier element. These SEMs can be arranged as discrete dynodes or as continuous (micro) channels.

Linear quadrupoles[91,110] (Q) are comprised of two pairs of cylindrical electrodes arranged in a square. The pairs of electrodes are held at the same voltage comprised of a DC and a radio voltage (RF) part, the opposite of the other pair. Ions that were previously accelerated are disturbed in their flight path. With the frequency of the RF component, the *m/z* ratio, which can stably cross the quadrupole, is controlled. Thus, the quadrupole acts as a filter and the same SEM-type detectors as described above can be used. To gain information from MS/MS experiments, triple quadrupoles (QqQ, TQ) can be used. While the first and the last quadrupole work as described above, the one in the middle acts as a collision cell, applying only the RF voltage part purposefully inducing unstable ion flight paths. Instead of a third quadrupole, also a ToF detector (Q-ToF) can be used[111], a common example of hybrid MS techniques.

Orbitrap[112,113] detectors although based on an idea from the 1920s are only commercially available since 2005. In contrast to ToF- or TQ-MS, the ions do not hit a SEM detector but are analysed non-destructively in a trap. Orbitraps have this in common with Fourier transform ion cyclotron resonance (FT-ICR) MS systems but do not need large cryogenically cooled magnets. In most modern commercial machines, upon entering the MS, non-ionised molecules are sorted out by a bent flatpole. Next, a quadrupole is used as a mass filter to limit the number of ions progressing to the curved linear trap (c-trap), which can hold the ions. Now the ions are either sent to a TQ-like collision cell and back or directly to the orbitrap mass analyser. The analyser consists of a spindle-shaped centre electrode and an outer electrode. The complex electrostatic field generated by the shape of the inner electrode does not only force the ions into an orbit around it but also induces back-and-forth oscillations, whose frequencies are solely dependent on the *m/z* ratio of the ions. This induces an image current in the outer electrodes, where each *m/z* ratio produces a sine wave, which can be resolved by Fourier transformation.

While the ion sources each have individual advantages and disadvantages based on their physical way of ion production, the targets for MS detectors are quite simple: high resolution and large *m/z* range. While ToF, Q, and Q-ToF instruments offer a resolution of around 10,000, Orbitraps can easily reach 100,000. But they only provide an *m/z* range of up to 2,000[95]. An exceptional FT-ICR MS was constructed to reach a resolution of over 2,100,000[114], but the costs generally scale with the resoltution[95].

1.5.3.3 Selected-ion flow-tube (SIFT) and Proton transfer reaction (PTR) MS

SIFT-MS[115,116], in contrast to previous methods, which were developed for liquid-phase analytes and were later adapted for VOCs, was developed for trace-gas analysis. First, a microwave discharge ion source produces reactant ions (H_3O^+, NO^+, $O_2^{+\bullet}$) from laboratory air and water, which are then filtered by a first quadrupole to only select one of these at a time. In a helium-filled flow tube, the sample gas is introduced and ionised by gas-phase collisions with the selected reactant ion. The analysis is then performed by a second quadrupole. The ionisation chemistry of the three reactant ions is very different and can produce different spectra, widening the range of analysable molecules and at the same time, this can be used to strengthen the identification of measured ions. **PTR-MS**[117] is a further evolution of SIFT-MS. Instead of filtering reactant ions, H_3O^+ ions are produced in high purity by a hollow discharge cathode from water. Further, the flow tube is replaced by a shorter drift tube, where the ions are guided by an electrical field instead of a carrier gas, which minimises dilution effects. Although linear quadrupoles are predominantly used, ToF and even ion traps can be used as mass analysers, again closing the gap between the different techniques.

1.6 The exploitation of *S. cerevisiae* volatiles

With such a broad spectrum of volatiles, there are numerous industrial applications, including the production of bulk chemicals (chapter 1.6.1), fine chemicals (chapter 1.6.2), the removal of VOCs, or monitoring yeast growth.

1.6.1 Biofuel production

Bioethanol production by *S. cerevisiae*, as mentioned above, for blending or replacing gasoline in transportation is the volumetrically biggest biotechnological process[17]. The ethanol pathway in *S. cerevisiae* has a very high flux and efficiency and is additionally linked to the cells' energy production (see Figure 2). Glycolysis and the subsequent TCA cycle produce large amounts of NADH. If the cells' respiratory system cannot regenerate the cofactors, for example under oxygen-limited or high glucose conditions, pyruvate is decarboxylated to acetaldehyde and then reduced to ethanol, recycling one NADH per molecule. In theory, one molecule of glucose is converted into two molecules of ethanol and two molecules of carbon dioxide. In reality, however, mainly due to the formation of biomass, additional reduced compounds, e.g., glycerol, need to be formed. With > 90 % theoretical yield, > 1.0 g/L/h, and the end product concentrations > 100 g/L, this process is fast and efficient[18]. Furthermore, the carbon dioxide produced during ethanol fermentation is very pure and is commonly used in the beverage industry[118].

Bioethanol production can roughly be divided into three parts: obtaining fermentable sugars, fermenting sugars into ethanol, product separation and purification. The pre-treatment is heavily dependent on the type of biomass used. First-generation biofuels used starchy or sucrose-rich plants, like corn, sugarcane, or sugar beet, which need only little pre-treatment but could also directly be used as food[119]. This gave

rise to the “food-vs-fuel” debate[21]. Second-generation biofuels utilise non-food sources such as waste or lignin-biomass. Although this technology is commercialised, the share of second-generation biofuels is neglectable[120]. Bioethanol production is mostly run as a batch process, and after cell filtration, the ethanol has to be separated from the fermentation broth containing salts, proteins, and other molecules. This is done by distillation or rectification up to a purity of 92-94 %, due to the ethanol-water azeotrope. Multiple different techniques can be used to obtain higher purities like pressure-swing distillation[121], but both process steps are energy-intensive totalling up to 15 % of the overall process costs[122].

Biodiesel[123,124], mostly medium-chain fatty acid methyl (FAME) or ethyl (FAEE) esters, is being developed to overcome the weaknesses of bioethanol such as hygroscopicity and comparatively low energy density[123]. As discussed above (chapter 1.4.3, page 10), the production of FAEEs is native to *S. cerevisiae*, but extensive metabolic studies were performed to optimise fluxes and yields. As an example of the complex challenges in optimising cell-based production, the increase of acetyl-CoA as FA precursor availability might serve. One reaction draining the acetyl-CoA pool is the production of ethanol which could be prevented by the deletion of either alcohol dehydrogenases, leading to the accumulation of toxic acetaldehyde, or pyruvate decarboxylases, leading to severe growth defects, as the concentration of cytosolic acetyl-CoA is decreased. Many other factors have to be balanced to maximise yeast-based biodiesel production, such as the concentration of malonyl-CoA, redox equilibrium, glucose flux, and enzyme levels for the subsequent steps[123]. Also relying on the *de novo* synthesis of fatty acids is the production of fatty alcohols and alkanes, both shown to be possible in *S. cerevisiae*. For both compounds, extensive industrial interest exists[123].

1.6.2 Flavour volatiles

The smell and taste of yeast fermented products like beer, bread, and wine are heavily influenced by VOCs. In the last decades, ongoing advances have been made in identifying new molecules contributing to the bouquet, and nowadays it is generally accepted that rather than the raw material, the yeasts are responsible for most of them[125]. With the help of modern genetic tools, key enzymes for the production of flavour-active compounds, for example, fusel alcohols, acids, and esters are easily identified[80]. But because genetically engineered organisms (GMOs) for food applications are only seldom approved, extensive research on the influence of process parameters like fermentation temperature on the VOC formation has been performed[126]. Taking this further, grape must contains a plethora of different non-*Saccharomyces* species, which were previously seen as contamination, but now are studied, showing their contribution to the overall VOC composition[127].

Overexpression of aroma compounds since the dawn of genetic editing is no longer limited to a smell’s native organism but can be done in any chassis organism. Baker’s yeast is an attractive host based on the variety of established genetic tools,

longstanding experience, and fermentation infrastructure. Vanillin, although it naturally occurs in orchids and has been produced chemically for over 100 years. Until now, no commercial microbial production strategy has been implemented, but considerable progress has been made in recent years. Starting from the intermediate of the shikimate pathway, 3-dehydroshikimate, four heterologous genes from hosts reaching from bacterium to human are needed to produce this high-value compound in *S. cerevisiae*. Since vanillin exhibits antimicrobial activity and inhibits growth in yeast at only 50 mg/L, it is detoxified to vanillin-*b*-$_D$-glucose, which allows yields of up to 500 mg/L[128].

In contrast to vanillin, the yeast-based production of valencene and nootkatone is already commercialised, e.g., by Evolva. Both compounds are bicyclic sesquiterpenes with a low odour threshold and a fruity, citrus-like aroma[129,130]. Valencene is produced from glucose in *S. cerevisiae* after the introduction of a single gene, the valencene synthase[131], but further genetic optimisation was needed to obtain reasonable yields[132]. Nootkatone, on the other hand, can be obtained from valencene through multiple conversions. Nootkatone can be obtained with 98 % purity, and in addition to its uses in food and flavour industries, it is used as an insecticide[133].

1.6.3 Other applications

Biofilters are a longstanding concept to remove VOCs from polluted printing press air due to the ability of different fungi to degrade for example alkanes, phenols, and formaldehyde[134]. *S. cerevisiae* and *Candida guilleriermondii* were used by Granström et al. to detoxify printing press air, including the complete removal of ethanol, ethyl acetate, 1-, and 2-propanol[135].

Biomedical applications of microbial VOCs analysis include human breath analysis for finding infection biomarkers. Using the human breath in comparison to blood samples has the advantage of being non-invasive. This means that patients are more willing to participate in studies and even infants can be tested regularly[136]. For example, SESI-Orbitrap MS was used to identify infection biomarkers from viruses and bacteria[137,138]. This connects to baker's yeast volatilome studies, as, although *S. cerevisiae* is mostly considered a non-pathogen, there are reports of invasive infections[139,140].

Monitoring mould growth is possible using yeast and fungi VOCs. Contrasting wipe tests and colony forming unit determination, this is even possible without finding the mould spot. The applicability was shown for building and agricultural environments. No single VOC could to this day be identified to show and quantify mould growth, but among the important compounds are 1-octen-3-ol or dimethyl sulfide[141].

1.7 Scope and Outline of this thesis

The overall goal of this doctoral thesis was to gain deeper insights into the gas phase above yeast-based fermentations. Novel combinations of chemical and biological catalysis were explored to capture and re-use CO_2 emerging from bioethanol

fermentations. Acetaldehyde as an example for volatile metabolites was produced with *in situ* separation and capturing. And lastly, SESI-Orbitrap MS for the online gas-phase analysis of such compounds was established.

Chapter 1 explores the origin of anthropogenic CO_2 production and the role of biotechnology to minimize the need for fossil fuels and feedstocks. In accordance with the aim of this work, it introduces the reader to the molecular principles of volatility and the current status of volatile analytics, with special attention to mass spectrometry-based methods. Lastly, the chapter is dedicated to the exploration and exploitation of the *S. cerevisiae* volatile space.

Chapter 2 contains the material and methods used in this study. More information on the experimental procedure can be found in the appendix, the previously published articles or through the author: hendrik.mengers@rwth-aachen.de

Chapter 3.1 describes how chemical catalysis and yeast-based fermentations can be combined in a one-step, one-pot reaction. These two are often seen as contradictory based on the diverging demands of metallic and cell catalysts, e.g solvents, temperature, and pressure. Carefully conflating the reaction conditions, this chapter shows that by adding a Ru-catalyst and H_2 pressure, CO_2 from a bioethanol fermentation can be captured and upgraded *in situ*, killing two birds with two stones.

Chapter 3.2 shows the process of establishing yeast-based acetaldehyde production. Acetaldehyde is mainly synthesised from petrochemical feedstocks with a production of over 1 Mt/a. Its volatile nature makes *in situ* separation in an aerated fermenter inevitable but also favourable, as downstream purification costs make up a considerable amount of biotech processes. Fermentations in 200- and 750-mL scale were performed with a genetically modified *S. cerevisiae* and acetaldehyde capture using water traps established and characterised.

Chapter 3.3 established SESI-Orbitrap MS for the gas-phase analytics of metabolites on the example of the heat-labile sulfur-containing allicin, the main flavour ingredient in garlic. The compound was measurable over an open microreaction tube with a pure solution, over freshly crushed garlic and in human breath after garlic consumption. The kinetics with a time-resolution of 0.4 Hz regarding the known thiosulfinate chemistry in garlic, ramsons, and onions could be followed by this machine.

Chapter 3.4 finally explores the possibilities of using SESI-Orbitrap MS for online, real-time analysis of the complete yeast volatilome. On the one hand, the applicability of this system for reaction monitoring was demonstrated by tracking ethanol, with results comparable to HPLC or GC, and acetaldehyde, detected hours earlier. On the other hand, multiple shifts in the volatilome were observable. For example, amino acid degradation products were found at the end of the lag phase and fatty acid ethyl esters were detected during the C-source shift upon glucose depletion.

Chapter 4 concludes the findings of this doctoral thesis and places them in a larger context. It again shows the breadth of yeast volatiles and the potential for future industrial exploitation of the molecules and measurement methods.