Ana Patrícia Imaginário Traguedo
Licenciada em Bioquímica

Theoretical and experimental approaches towards the non-invasive and selective detection of microbial pathogens

Dissertação para obtenção do Grau de Mestre em Bioquímica

Orientadora: Prof.ª Doutora Ana Cecília Afonso Roque, FCT-UNL

Co – orientadora: Doutora Susana Isabel Conde Jesus Palma, FCT-UNL

Júri:
Presidente: Prof. Doutor Pedro António de Brito Tavares, FCT-UNL
Arguente: Doutor Abid Hussain, Biosurfit, S.A.
Vogal: Prof.ª Doutora Ana Cecília Afonso Roque, FCT-UNL

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Outubro 2016

Theoretical and experimental approaches towards the non-invasive and selective detection of microbial pathogens

Agradecimentos

De uma forma geral, gostaria de agradecer a todos aqueles que de alguma forma contribuíram para o desenvolvimento da minha tese de mestrado durante este ano.

Em primeiro lugar gostaria de agradecer à minha orientadora Professora Ana Cecília Roque por me ter dado a oportunidade de fazer parte deste grupo de investigação. Por todo o apoio, orientação e dedicação demonstrados, que fizeram com que tantas vezes superasse as minhas expectativas perante os desafios lançados. Obrigada por toda a paciência, compreensão e entusiasmo que fizeram de mim uma pessoa mais autónoma e confiante.
Também não poderia deixar de agradecer à minha co-orientadora Susana Palma que, pela segunda vez ao longo do meu percurso académico, me acompanhou num projeto com a maior dedicação e carinho possíveis! Por todos os sorrisos, conselhos, motivação e ajuda quando mais precisava um gigantesco obrigada!

Por sua vez, ao Professor Hugo Gamboa quero agradecer por toda a ajuda prestada na análise computacional de dados, por todas as reuniões, e por estar sempre disponível para explicar de forma diferente até eu conseguir perceber uma área que para mim era, até então, praticamente desconhecida. Sem a sua ajuda, este trabalho não teria sido possível.
Quero também expressar os meus agradecimentos a todo o grupo de investigação, pelo bom ambiente de trabalho, conselhos e ajuda disponibilizados.

Gostaria de agradecer à minha família por todo o apoio e carinho demonstrados ao longo de toda a minha vida. Em particular, agradeço aos meus pais e irmã que sempre se mostraram interessados em saber em que consistia este trabalho, apesar de pouco ou nada entenderem acerca desta área, por sempre terem uma palavra positiva e de força, mesmo nos piores momentos, e por me ajudarem sempre a atingir os meus objetivos. Obrigada por me ajudarem a concretizar esta nova etapa!

A todos os meus amigos que estiveram presentes e que me ajudaram de alguma forma a ultrapassar os maus momentos, descontrair e encarar o trabalho que tinha pela frente de forma mais otimista, um muito obrigada!

Finalmente, gostaria de dedicar este trabalho aos meus queridos avós maternos e paternos, em especial aos avós Zé e Jacinta que, apesar de já não estarem presentes, sempre acompanharam a minha vida pessoal e académica com um carinho que nunca irei esquecer.

Abstract

Non-invasive diagnostics of microbial infections based on the volatome is an area of increasing interest. This work set the ground for the development of an artificial nose for the diagnosis of bacterial infections. In a first part, the aim was to find a group of volatile organic compounds (VOCs) able to distinguish between 8 clinically relevant pathogens. The second part aimed to engineer the selectivity of VOC-responding materials (biogels) through the incorporation of VOC-specific peptides.

A systematic review and analysis of available literature data relating the detection of VOCs in human samples with the presence of specific pathogen infections was performed. Statistical classification methods were employed to make a metasearch for potential pathogen VOC biomarkers using those data. A minimal set of VOCs that allows the distinction between *Pseudomonas aeruginosa, Staphylococcus aureus, Aspergillus fumigatus, Escherichia coli, Helicobacter pylori, Proteus mirabilis, Klebsiella pneumoniae* and *Mycobacterium tuberculosis* was suggested.

A comparison between the set of VOCs found by the analysis and biomarkers previously reported for the same pathogens was made. As a result, a set of potential biomarkers for pathogen infection is suggested: indole for *E.coli*; 2-pentylfuran for *A.fumigatus*; isobutene for *H.pylori*; cymol for *M.tuberculosis*; hydrogen cyanide and methyl thiocyanate for *P.aeruginosa*; and 3-methylbutanoic acid for *S.aureus*.

The feasibility of engineering biogels VOC-selectivity was assessed by incorporating in the materials a benzene-sensitive peptide previously reported (P1) and two modified versions containing norleucine (P2) or biphenylalanine (P3) at the C-terminal. The optical response of the as-produced materials to several VOCs was tested on an in-house developed electronic-nose (e-nose). The biogels without any peptide responded more sharply to benzene and acetone. The addition of P1 amplifies the response to benzene and toluene. The addition of P2 and P3 amplified the response signal to both acetone and benzene.

Keywords: volatile organic compounds; biomarkers; pathogen infections; liquid crystal; ionic liquid; electronic-nose.

Resumo

O diagnóstico não invasivo de infeções microbianas baseado no volatoma é uma área de crescente interesse. Este trabalho definiu o caminho para o desenvolvimento de um nariz artificial para o diagnóstico de infeções bacterianas. Numa primeira parte, o objetivo era encontrar um grupo de compostos orgânicos voláteis (VOCs) capaz de distinguir entre 8 patogénios clinicamente relevantes. A segunda parte teve como objetivo desenvolver a seletividade de materiais que respondem à presença de VOCs (biogéis) através da incorporação de três péptidos diferentes específicos para VOCs.

Realizou-se uma revisão sistemática e análise dos resultados disponíveis na literatura relativos à deteção de VOCs em amostras humanas em casos de infeções causadas por agentes patogénicos específicos. Utilizaram-se métodos de classificação estatística para realizar uma metapesquisa para identificar potenciais VOC biomarcadores de patogénios, usando esses dados. Foi sugerido um conjunto mínimo de VOCs que permite distinguir entre *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Aspergillus fumigatus*, *Escherichia coli*, *Helicobacter pylori*, *Proteus mirabilis*, *Klebsiella pneumoniae* e *Mycobacterium tuberculosis*.

Realizou-se depois uma comparação entre o conjunto de biomarcadores voláteis encontrados na nossa análise e os biomarcadores reportados anteriormente, para os mesmos patogénios. Como resultado, foi sugerido um conjunto de potenciais biomarcadores para infeções patogénicas: indole para *E.coli*; 2-pentilfurano para *A.fumigatus*; isobuteno para *H.pylori*, cimeno para *M.tuberculosis*; cianeto de hidrogénio e metil-tiocianato para *P.aeruginosa*; e ácido 3-metilbutanóico para *S.aureus*.

A viabilidade de desenvolver a seletividade para VOCs em biogéis foi avaliada pela incorporação de um péptido, anteriormente reportado, sensível ao benzeno (P1) e duas versões modificadas do mesmo, contendo norleucina (P2) ou bifenilalanina (P3) no C-terminal, nos materiais. A resposta ótica dos materiais produzidos a vários VOCs foi testada no nariz eletrónico (e-nose) desenvolvido *in-house*. Os biogéis sem qualquer péptido responderam de forma mais acentuada ao benzeno e acetona. A adição de P1 amplificou a resposta para o benzeno e o tolueno. A adição de P2 e P3, amplificou o sinal de resposta tanto para acetona como para o benzeno.

Termos-chave: compostos orgânicos voláteis; biomarcadores; infeções por agentes patogénicos; cristal líquido; líquido iónico; nariz eletrónico.

Table of Contents

Agradecimentos v
Abstract vii
Resumo ix
Index of Figures xiii
Index of Tables xvii
Abbreviations xix
1. Introduction 1
1.1 Infectious diseases 1
1.2 Volatile organic compounds (VOC) analysis in clinical samples towards diagnosis 2
1.3 VOCs study experimental outline 3
1.3.1 Sample collection 4
1.3.2 Storage 4
1.3.3 Pre-concentration 5
1.3.4 Analysis methods 5
1.3.4.1. Off-line analysis methods 6
1.3.4.1.1 GC-MS (Gas-chromatography mass spectrometry) 6
1.3.4.1.2 GC-FID (Gas-chromatography coupled to flame ionization detector) . 6
1.3.4.2 Real-time analysis 7
1.3.4.2.1 PTR (Proton Transfer Reaction Mass Spectrometry) 7
1.3.4.2.2 SIFT (Selected Ion Flow Tube Mass Spectrometry) 7
1.3.4.2.3 IMS (Ion Mobility Spectrometry) 8
1.3.4.2.4 SESI-MS (Secondary Electrospray Ionization Mass Spectrometry) 8
1.3.4.2.5 IMR (Ion Molecule Reaction Mass Spectrometry) 8
1.4 Data processing 10
1.5 Electronic noses (e-noses) 10
1.6 Challenges and future directions 12
2. Aim of the work 13
3. Looking for infectious disease VOC biomarkers 15
3.1 Introduction 15
3.2 Materials and methods 16
3.2.1 Literature search and data collection 16
3.2.2 Data handling 19
3.2.3 Statistical analysis 20
3.3 Results and discussion 24
3.3.1 Pathogen- VOC interactions in the bibliography 24
3.3.2 VOCs distinguishing pathogens 35

3.4 Conclusions 48
4. Tailoring biogels selectivity 49
4.1 Introduction 49
4.2 Materials and methods 52
4.2.1 Materials 52
4.2.2 Methods 53
4.2.2.1 Production of gelatin biogels 53
4.2.2.2 Production of agarose biogels 54
4.2.2.3 Production of polyacrylamide biogels 54
4.2.2.4 Coomassie staining of the biogels 55
4.2.2.5 Labeling of peptide P1 with FITC 55
4.2.2.6 Determination of P1 and FITC concentrations in the P1-FITC sample 55
4.2.2.7 Incorporation of P1-FITC in a gelatin biogel 56
4.2.2.8 Incorporation of peptides P1, P2 and P3 in gelatin biogels 56
4.2.2.9 Evaluation of the optical response of gelatin biogels in the presence of different VOCs 57
4.2.2.10 Optical characterization of the biogels by Microscopy 58
4.3 Results and discussion 59
4.3.1 Doping gelatin biogels with P1, P2 and P3 peptides 59
4.3.2 Tracking peptides location within biogels 64
4.3.2.1 Staining with Coomassie blue 64
4.3.2.2 Tracking peptides by Fluorescence 68
4.3.3 Effect of P1, P2 and P3 in the response of gelatin biogels to different VOCs 72
4.3.4 Biogels stability after VOCs exposure 75
4.4 Conclusions 77
5. Concluding remarks and future perspectives 79
6. Bibliography 81

Index of Figures

Figure 1.1- Schematic representation of VOCs study experimental outline. 3

Figure 1.2 – Schematic representation of volatile compounds recognition by an electronic nose device. 11

Figure 2.1 - Schematic representation of the thesis project. 13

Figure 3.1- Strategy followed for the selection of papers. 18

Figure 3.2- Classification steps. 21

Figure 3.3- Support Vector Machine optimal separation plan and separation margin. 21

Figure 3.4- Sequential forward feature selection. 23

Figure 3.5- Representation of the number of articles per year. 25

Figure 3.6- Representation of the number of experiments reported in each article. 25

Figure 3.7- Representation of the number of experiments associated with the 23 pathogens.... 30

Figure 3.8- Samples used to obtain VOC biomarkers. 32

Figure 3.9- Analysis methods used and the corresponding number of experiments. 32

Figure 3.10- Piechart with the relative number of hits (% relative abundance) of compounds in each class that have been detected in all the experiments. 33

Figure 3.11 - Representation of the number of hits found for each individual VOC, in all experiments, for the alcohols and the halogen containing classes. 34

Figure 3.12- Representation of all the 23 pathogens and 418 associated VOCs.. 35

Figure 3.13- Circular layout of the 8 pathogens with more than 3 associated experiments, and the 269 associated VOCs..... 36

Figure 4.1- Micelles with LC molecules inside, IL at the surface and simulation of the peptides proposed interaction with the micelles. 51

Figure 4.2- Structure of peptides P1 (A), P2 (B) and P3 (C). 51

Figure 4.3- Structures of 4-cyano-4-pentylbiphenyl (5CB), 1-butyl-3-methylimidazolium dicyanamide ([BMIM][DCA]) and Fluorescein isothiocyanate (FITC) Isomer I. 52

Figure 4.4- Representation of the glass plate, after spreading the gel. 53

Figure 4.5- In-house developed e-nose assembly, when exposing a biogel to a solvent (A). E-nose sensor array, seen from above (B) and from the side (C).. 58

Figure 4.6- POM images of standard biogel (G_3- Gelatin + [BMIM][DCA] + 5CB + milliQ water) and respective controls.. 60

Figure 4.7- Visualization of a [BMIM][DCA]/5CB/gelatin biogel, with crossed polarizers. A: visualization of the whole biogel. B- detail of a biogel region. 61

Figure 4.8- Standard gelatin biogel and respective controls, observed by POM, with crossed polarizers, in 2 different fields of view. The scale bar corresponds to 50 μm..... 62

Figure 4.9- Gelatin biogels with P1 added, visualized by POM, with crossed polarizers in 2 different fields of view.The scale bar corresponds to 50 μm. 62

Figure 4.10- Gelatin biogels with P2 added, visualized by POM, with crossed polarizers in 2 different fields of view.The scale bar corresponds to 50 μm. 63

Figure 4.11- Gelatin biogels with P3 added, visualized by POM, with crossed polarizers in 2 different fields of view.The scale bar corresponds to 50 μm. 63

Figure 4.12- Appearance of biogels produced using agarose (A) and gelatin (B), as the micelles support.. 65

Figure 4.13- Agarose gels visualization by POM, with crossed polarizers, in 2 different fields of view, before (a to h) and after (a' to h') coomasie staining.. 66

Figure 4.14- Polyacrylamide gels visualization by POM, with crossed polarizers, in 2 different fields of view, before coomasie staining (a to h) and after (a' to h'). 67

Figure 4.15- Chromatograms obtained for peptide P1, FITC and P1-FITC conjugate samples, overlayed (A) and identification of each fraction (B) 68

Figure 4.16- POM images of gelatin biogels without any peptide added (a to c) and P1-FITC (d to f) with uncrossed (a and d), semi-crossed (b and e) and crossed polarizers (c and f). All images have the same scale. The scale bar corresponds to 50 μm. 70

Figure 4.17- Gelatin biogels with (b and f) and without P1-FITC (a and e) visualized by POM, with crossed polarizers, and by fluorescence microscopy (c, g, and d, h) in 2 different fields of view. a, c, e and g: control biogel composed by gelatin, [BMIM][DCA], 5CB, P1 and milliQ water. b, d, f and h: biogel composed by gelatin, [BMIM][DCA], 5CB, P1-FITC and milliQ water.. 71

Figure 4.18- Overlayed signals of the biogels containing [BMIM][DCA]/5CB (A), [BMIM][DCA]/5CB/P1 (C), [BMIM][DCA]/5CB/P2 (B) and [BMIM][DCA]/5CB/P3 (D) when exposed to acetone, benzene, hexane, toluene, xylene and ethanol..... 73

Figure 4.19- Optical signal intensity of each biogel to the different solvents. 74

Figure 4.20- Response fold increase of each biogel, after P1, P2 and P3 incorporation in the standard biogel. 74

Figure 4.21- POM images of standard gelatin biogels (gelatin + [BMIM][DCA] + 5CB), after VOCs exposure, in 2 different fields of view, with uncrossed (a and d), semi-crossed (b and e) and crossed (c and f) polarizers. All images have the same scale. The scale bar corresponds to 50 μm. 75

Figure 4.22- POM images of P1 biogels (gelatin + [BMIM][DCA] + 5CB + P1), after VOCs exposure, in 2 different fields of view, with uncrossed (a and d), semi-crossed (b and e) and crossed (c and f) polarizers. All images have the same scale. The scale bar corresponds to 50 μm. 75

Figure 4.23- POM images of P2 biogels (gelatin + [BMIM][DCA] + 5CB + P2), after VOCs exposure, in 2 different fields of view, with uncrossed (a and d), semi-crossed (b and e) and crossed (c and f) polarizers. All images have the same scale. The scale bar corresponds to 50 μm. 76

Figure 4.24- POM images of P3 biogels (gelatin + [BMIM][DCA] + 5CB + P3), after VOCs exposure, in 2 different fields of view, with uncrossed (a and d), semi-crossed (b and e) and crossed (c and f) polarizers. All images have the same scale. The scale bar corresponds to 50 μm. 76

Index of Tables

Table 1.1- Comparison of the mode of operation, sensibility, advantages and disadvantages of the different analytical methods. 9

Table 1.2- Summary of the different e-noses formats and their mode of operation 11

Table 3.1- Pathogens referred in the selected papers, respective classification, analysis methods used to detect VOCs and number of experiments associated with each pathogen. Pathogens marked (*) were studied in more than 3 experiments. 26

Table 3.2- Set of 26 VOCs that better separates the 8 analyzed pathogens, selected iteractively, by an identification approach. The VOCs are listed from the highest to the lowest score..... 37

Table 3.3- Accuracy rates obtained by classification with and without cross-validation..... 38

Table 3.4- Confusion matrix for the classification without cross-validation, where the column labels represent the predicted pathogen classification and the line labels represent the actual pathogen identity.. 39

Table 3.5 – Associated diseases/infections to each of the 8 studied pathogens. 41

Table 3.6- Representation of the results obtained by the verification approach..... 41

Table 3.7- Comparison between the set of VOCs that better separates the group of 8 pathogens, obtained by identification approach, and the VOCs reported in the literature as present (X) or absent ( ) in healthy bodily fluids.. 43

Table 3.8- Comparison between the VOCs that better separates the 8 pathogens, obtained by a verification approach, and the VOCs reported in the literature as (X) or absent ( ) in healthy bodily fluids. 44

Table 3.9- Comparison between the set of VOCs obtained in our study, the mVOC database and two existing reviews, where "+" means positive biomarker and "-" negative biomarker..... 46

Table 4.1- Reagents and corresponding quantities used to produce biogels (G_3) and the respective negative controls (G_0, G_1, G_2). 53

Table 4.2- Reagents and corresponding quantities used to produce agarose biogels (A_3), and the respective negative controls (A_0, A_1, A_2). 54

Table 4.3- Reagents and corresponding quantities used to produce biogels with polyacrylamide (P_3), and the respective negative controls (P_0, P_1, P_2). 55

Table 4.4- Reagents and corresponding quantities used to incorporate the peptides P1 labeled with FITC in the liquid crystal-ionic liquid micelles, and the respective negative control..... 56

Table 4.5- Reagents and corresponding quantities used to incorporate the peptides P1 (G_3^1), P2 (G_3^2) and P3 (G_3^3) in the biogels, and to produce the respective negative controls (G_0^1, G_1^1, G_2^1), (G_0^2, G_1^2, G_2^2) and (G_0^3, G_1^3, G_2^3). 57

Table 4.6- Biogels with distinct composition and corresponding number of micelles formed and corresponding mean area (μm^2) 64

Table 4.7- Values of absorbance (280 nm) and fluorescence (485-535 nm) for P1 and FITC and the respective concentrations of each compound, in the P1-FITC sample before and after using the rotary evaporator. 69

List of Abbreviations

[BMIM][DCA]- 1-Ethyl-3-methylimidazolium dicyanamide

5CB- 4- Cyano- 4'- pentylbiphenyl

AF- *Aspergillus fumigatus*

AR- Allergic reactions

BI- Bone infections

BLI- Blood infections

CA- *Candida albicans*

CAP- Community-acquired pneumonia

CAR- Carboxen

CD- *Clostridium difficile*

CF- Cystic fibrosis

CJ- *Campylobacter jejuni*

CO_2- Carbon dioxide

DVB- Divinylbenzene

EC- *Escherichia coli*

EF- *Enterococcus faecalis*

EI-MS- Electron ionization- Mass Spectrometry

e-nose- Eletronic-nose

ESI- Electrospray ionization

FID- Flame ionization detector

FITC- Fluorescein isothiocyanate

GC- Gas chromatography

GC-FID- Gas chromatography coupled with flame ionization detector

GC-MS- Gas chromatography Mass Spectrometry

GC-SAW- Gas chromatography surface acoustic wave

GD- *Giardia duodenalis*

GI- Gastric infections

GTI- Gastrointestinal infections

H_2O - Water

H_3O^+ - Hydronium ion

HI- *Haemophilus influenzae*

HP- *Helicobacter pylori*

HS-SPME- Headspace solid-phase microextraction

ICU- Intensive care unit

IL- Ionic liquid

IMR-MS- Ion molecule reaction Mass Spectrometry

IMS- Ion mobility spectrometry

KI- Kidney infections

KP- *Klebsiella pnemoniae*

LC- Liquid crystal

LDR- Light Dependent Resistor

LED- Light Emitting Diode

LP- *Legionella pneumophila*

m/z- Mass to charge ratio

MC- *Moraxella catarrahalis*

MCC- Multi- capillary column

MCC-IMS- Multi- capillary column coupled to ion mobility spectrometry

MM- *Morganella morganii*

MRSA- *Methicillin-resistant Staphylococcus aureus*

MS- Mass Spectrometry

MT- *Mycobacterium tuberculosis*

N_2 - Nitrogen

NM- *Neisseria meningitidis*

NO^+- Nitrosonium ion

NTD- Needle Trap Device

O_2^+ - Oxygen ion

PA- *Pseudomonas aeruginosa*

PCA- Principal component analysis

PCR- Polymerase chain reaction

PDMS- Polydimethylsiloxane

PET- Polyethylene Terephthalate

PF- *Plasmodium falciparum*

PLSDA- Partial least squares Discriminant Analysis

PM- *Proteus mirabilis*

PN- Pneumonia

POM- Polarized Optical Microscopy

ppb- parts per billion

ppm- parts per million

ppt- parts per trillion

PTFE- Polytetrafluoroethylene

PTR-MS- Proton Transfer Reaction Mass Spectrometry

PTR-TOF-MS- Proton Transfer Reaction- time of flight- Mass Spectrometry

PV- *Proteus vulgaris*

SA- *Staphylococcus aureus*

SE- *Staphylococcus epidermidis*

SESI-MS- Secondary Electrospray Ionization Mass Spectrometry

SI- Sinusitis

SIFT-MS- Selected Ion Flow Tube Mass Spectrometry

SKI- Skin infections

SP- *Streptococcus pneumoniae*

SPME- Solid- Phase Microextraction

TB- Tuberculosis

TD-GC-MS- Thermal desorption Gas Chromatography Mass Spectrometry

UTI- Urinary tract infections

VAP- Ventilor- associated pneumonia

VOC- Volatile organic compound

WHO- World health organization

1. Introduction

1.1 Infectious diseases

Diseases caused by pathogenic microorganisms (bacteria, viruses, parasites or fungi) are called infectious diseases [1]. Worldwide, infectious diseases are the leading cause of death of children and one of the leading causes in adults [2], and an early diagnosis is essencial to initiate appropriate antimicrobial therapy for efficient patient management [3].

Tuberculosis, pneumonia and malaria are examples of infectious diseases that affect the population worldwide. Tuberculosis (TB) is caused by the bacterium *Mycobacterium tuberculosis* that usually affect the lungs and occurs in every part of the world [4]. Although it is curable and preventable, an earlier diagnosis is essential. In 2014, 9.6 million people fell ill with TB and 1.5 million died from the disease. Also, globally, an estimated 480 000 people developed multidrug-resistant TB [5]. However, Africa carried the most severe burden, with 281 cases per 100 000 population in 2014 (compared with a global average of 133) [5].

Pneumonia is a form of acute respiratory infection that affects the lungs and it is the largest infectious cause of death in children worldwide, accounting for 15% of all deaths in children under 5 years old, in 2015 [6]. This pulmonary disease can be caused by a number of infectious agents, including viruses, bacteria or fungi. The most common are *Streptococcus pneumoniae* and *Haemophilus influenzae. Pseudomonas aeruginosa* is an uncommon cause of community-acquired pneumonia (CAP), but a common cause of hospital-acquired pneumonia [7]. Ventilator-associated pneumonia (VAP) is a common hospital-acquired infection ocurring in the intensive care unit (ICU) and it is often caused by *Pseudomonas aeruginosa*. It is a complication of mechanical ventilation with an attributable mortality risk of 13% [8] even among patients receiving appropriate antimicrobial therapy. To date, the diagnosis is based on clinical criteria in combination with bacterial culture results.

Malaria is caused by infection with protozoan parasites belonging to the genus *Plasmodium* [9]. The parasites are transmitted to people through bites of infected female *Anopheles* mosquitoes, called "malaria vectors". There are five parasite species that cause malaria in humans [10], and two of these species – *P. falciparum* (most prevalent on the African continent) and *P. vivax* (predominates in many countries outside Africa) – pose the greatest threat [10]. About 3.2 billion people are at risk of malaria [10]. According to World Health Organizarion (WHO) estimates, released in December 2015, there were 214 million cases of malaria in 2015 and 438 000 deaths. Accurate diagnosis of malaria is important to provide adequate treatment and help prevent the emergence of resistant strains of malaria parasites [11]. However, diagnosis continues to present challenges. Currently, the majority of diagnoses rely on a combination of clinical presentation and the old approach of visualizing parasites on a stained blood film. There

remains a need for a simple, inexpensive, and reliable diagnostic test for malaria that can be performed *in situ* or in other primary healthcare settings in remote areas [12].

There are two key contributing factors for the highly negative prognosis in infectious diseases [13]. The first is the late diagnosis, usually performed using invasive and expensive procedures. The second is the lack of medical/laboratorial infrastructures in developing countries. Current methods for detecting microorganisms from clinical samples (culturing, polymerase chain reaction (PCR) and immunological methods) have some limitations regarding time, cost and complexity [14][15][16]. Therefore, there is an urgent need to develop fast, cheap, and accurate tests for the diagnosis of infectious diseases, so that it is possible to initiate early pathogen detection and subsequent specific treatment.

1.2 Volatile organic compounds (VOC) analysis in clinical samples towards diagnosis

Humans emit, normally, a broad range of VOCs, which can be both odorous and non-odorous [17]. VOCs can be emitted from different secretions of the human body [18]. Emission varies with many factors such as age, diet, sex, physiological condition and possibly genetic background [19]. Therefore, body odours can be considered as individual 'odour-fingerprints' [20]. Pathological processes, such as infection and endogenous disorders, can influence those odour fingerprints by producing new VOCs or by changing the ratio of VOCs that are normally produced [21]. The correlation between VOCs and health is well known since the old clinical practices. For instance, Hippocrates recognized the diagnostic value of body odours and reported several disease-specific odours emanated from two different samples: urine and sputum [22].

There are some advantages associated with identifying specific combinations of VOCs (VOCs profiling) associated with human diseases [22]. The composition of clinical samples headspace gives valuable information about both endogenous and exogenous compounds. The first ones will reflect biochemical processes in the body while the second will be originated by exogenous microorganisms, offering new possibilities for non-invasive clinical diagnostics [23]. Headspace sampling can be used to collect VOCs from liquid (urine, blood) and solid (skin, stool) samples [24].

Clinical sample VOC analysis represents a convenient and simple alternative to the time consuming and expensive traditional methods used in clinical laboratory diagnosis. For that, analysis and identification of compounds that are found to be characteristic of a certain infection in clinical samples (infection VOC biomarkers) has been target of substantial research and is emerging as a promising diagnosis tool in modern analytical chemistry [19].

VOC biomarkers for infections are important clinical tools not only due to their disease-detecting potential [24] but also because there are other conceivable applications of VOCs in the field of infectious diseases [25]. Namely:

- Monitoring disease severity and control;
- Predicting prognosis of a disease;
- Evaluating treatment;
- Screening/predicting risk for different diseases in population studies.

VOC analysis from clinical samples has been developing into an attractive proposition because it is non-invasive and the available techniques to measure VOCs (such as Gas Chromatography Mass Spectrometry) are very sensitive (ppt_v-ppb_v) to detect compounds [26], and procedures (such as Selected Ion Flow Tube Mass Spectrometry and Ion Mobility Spectrometry) [26] [27] allow real-time measurement of compounds in the body.

1.3 VOCs study experimental outline

VOCs study can involve different steps, depending on the analytical method employed [28]. Usually, the main steps involved are: sample collection, VOCs handling and storing, analysis, processing of the data obtained and, finally, result output [29] (Figure 1.1). Depending on the chosen analytical method, VOCs may need to be captured, pre-concentrated and then stored [17].

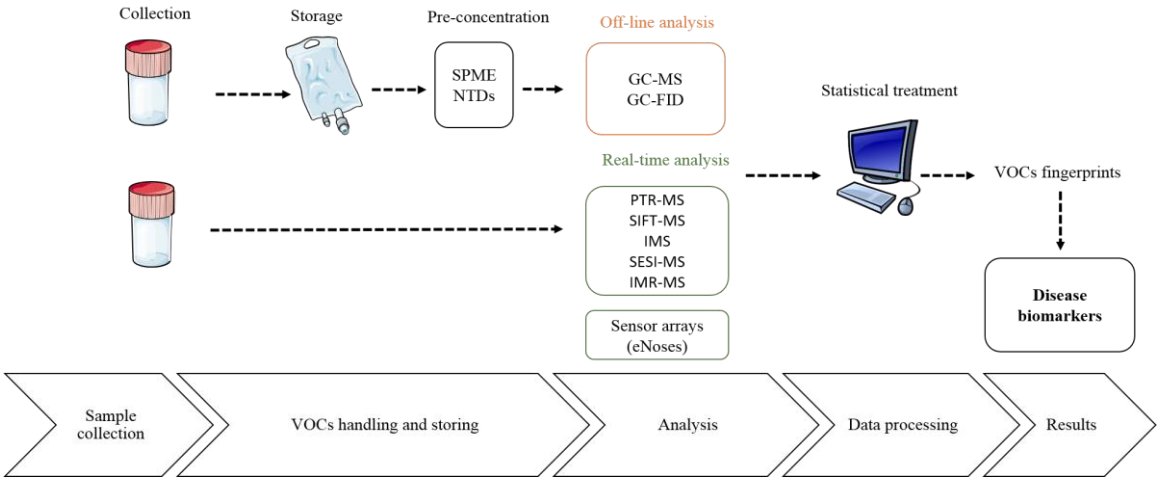

Figure 1.1- Schematic representation of VOCs study experimental outline. SPME- Solid-Phase Microextraction; NTDs- needle trap devices; GC-MS- Gas-Chromatography Mass Spectroscopy; GC-FID- Gas Chromatography coupled to Flame Ionization Detector; PTR-MS- Proton Transfer Mass Spectroscopy; SIFT-MS- Selected Ion Flow Tube Mass Spectroscopy; IMS- Ion Mobility Spectrometry; SESI-MS- Secondary Electrospray Ionization Mass Spectrometry; IMR-MS- Ion Molecule Reaction Mass Spectrometry.

1.3.1 Sample collection

Sampling is one of the most relevant steps. Collection of blood and urine samples, for example, has standardized procedures [30]. Briefly, the sample is introduced in a sealed vial and the volatile components will diffuse into the gas phase until an equilibrium is reached and a sample is then taken from the headspace [31]. In the case of breath analysis, sampling is not so trivial. Breath collection is done by exhaling directly into sampling bags [19]. It can be done in two different ways: through a single breath or multiple breaths. Although single breath collection is less time-consuming, multiple breath analysis is more reproducible in terms of sample composition. So, for screening of potential VOCs associated with a given disease and determination of a specific set of biomarkers, multiple breath analysis is required. The risk of contamination with exogenous compounds from the oral cavity and the surrounding environment is always high and may compromise the analysis and the results [26][32]. These problems result in the variation of the number of compounds and their concentration, which may impair the analytical reproducibility and data reliability [32].

Regarding samples handling, there are some parameters that should be carefully considered to avoid wrong conclusions about the origin of the identified VOCs to be taken [29] [27]. Some of these parameters are sample storage and the interference of environmental VOCs.

1.3.2 Storage

When real-time analysis is not possible, samples need to be stored. Storage should be at very low temperature to reduce VOCs loss, and the samples should be stored as soon after being taken as possible. In the case of liquid or solid samples, they should immediately be placed in an appropriate container and frozen to -80ºC or lower [30]. The container should be clean, produce no VOCs and should not change its characteristics with temperature variation and storage.

Breath sampling can be performed directly or indirectly according to the most suitable analysis to be performed. Direct sampling is preferable because there is no need to store for later analysis, so the decomposition of samples or loss of compounds by diffusion is avoided [19]. However, when direct analysis is not possible, the storage is an important factor to consider [26]. There are several ways to store breath and samples headspace [26]. The most typical examples are:

- Tedlar® bags (PTFE-polytetrafluoroethylene);
- Nalophan bags (PET-polyethylene terephthalate);
- Glass vials (for SPME);
- Thermal desorption tubes (different adsorbents, used in TD-GC-MS).

Currently, Tedlar® bags are the most common materials for VOC storage. However, Nalophan bags are also popular due to its low price, inertness, and relatively good durability [33].

1.3.3 Pre-concentration

During these procedures, interfering compounds could affect the analytical results. To minimize the interference, sometimes an intermediate step between sampling and analysis is required to increase the concentration of the target analytes over the interfering compounds [34]. There are several pre-concentration techniques available [35], such as Solid-Phase Microextraction (SPME) and Needle Trap Device (NTD) [36],[37].

Among the VOC sample pre-concentration methods, SPME is the most used one [24][38] [39][40][41]. This technique involves the use of a fiber coated with an extracting phase which can extract different kinds of analytes, depending on the chosen fiber. The quantity of extracted analyte is proportional to its concentration in the sample. Its headspace variant (HS-SPME), in which analytes belonging to solid or liquid samples are extracted from the headspace, has gained major importance regarding VOC sampling [41].

The Needle Trap Device is an emergent alternative, consisting on a syringe that allows the combination of both the sampling and the pre-concentration steps in a single device [42]. This device is composed by a needle containing a sorbent material packed inside. The sorbent constitution is variable and includes Carboxen (CAR), Divinylbenzene (DVB), Polydimethylsiloxane (PDMS) [43]. In this method, the sample can be actively drawn in and out by diffusion, gas-tight syringe or automated devices, such as vacuum pumps [44]. Unlike SPME, NTD is an exhaustive methodology, allowing an increase in the concentration of several compounds by using more sample volume [38]. Moreover, sample storage, prior to analysis, is also possible with NTD and has been shown to deliver reproducible results for several days of storage, depending on the target analytes [45].

1.3.4 Analysis methods

To identify the different substances within a clinical sample, such as breath or headspace of liquid or solid samples, analytical methods are needed. Since the first reports about exhaled breath composition [46], several methodological improvements and alternatives have been implemented in VOC analysis. Nowadays, VOC analysis is no longer limited to the off-line laboratory approach, as there are many real-time methods available [47]. Table 1.1 summarizes the features of the most commonly used analytical methods for the characterization of gaseous samples towards VOC identification. The real-time analysis alternatives include analytical methods such as proton transfer reaction mass spectrometry (PTR-MS) and its variations

(proton transfer reaction-time-of flight-mass spectrometry (PTR-TOF-MS)), IMS and IMS coupled with multi-capillary columns (MCC-IMS), selected ion flow tube mass spectrometry (SIFT-MS) and secondary electrospray-ionization mass spectrometry (SESI-MS). All these real-time options reduce several experimental steps related with sampling, storage and pre-concentration of the samples, allowing a faster analysis and reducing the loss of information during these steps. More recently electronic noses (e-noses) [48] have been developed and applied to breath analysis with promising results [49]. Many e-nose approaches rely on pattern recognition and perform a qualitative characterization of volatiles unlike the analytical devices that give absolute quantification of volatiles. However, wether the characterization is quantitative or qualitative, the target VOCs always have to be identified by expensive comprehensive methodologies, usually involving mass spectrometry detection [29][38]. Therefore, the search for a reliable tool for VOC analysis assessment is still in progress [32].

1.3.4.1 Off-line analysis methods

Gas chromatography (GC) was the analytical method used in the initial studies in VOC analysis [46] and until today it is the gold standard method when coupled to mass spectrometry (MS). Most exhaled breath VOCs reported so far have been identified and quantified using MS-based methods [38].

1.3.4.1.1 GC-MS (Gas-chromatography mass spectrometry)

GC-MS [20] [26] [32] allows the analysis of compounds in the concentration range from ppb to ppt (Table 1.1). In GC-MS, analysis occurs when volatilized samples are separated in a chromatographic column based on parameters, such as the polarity of the GC column. This system ionizes the target ions, separates them by mass to charge (m/z) ratio and then uses the fragmentation patterns to quantify the amount of each VOC present in the analyzed sample [32].

1.3.4.1.2 GC-FID (Gas-chromatography coupled to flame ionization detector)

In GC-FID [50], VOCs are burned in the FID, producing ions and electrons that can conduct the electric current and this information is used for detection and eventually quantification. GC-FID usually exhibits high sensitivity, large linear response range, and low noise. The FID detector is mass sensitive and its response is not altered significantly by changes in mobile-phase flow rate [32] [38].

1.3.4.2 Real-time analysis

Real-time analysis [13] obtains immediate results and does not require collection and storage of samples, eliminating a major source of experimental errors [26]. It has some advantages when compared with off-line analysis. However, real-time analysis also has some disadvantages, such as the expensive maintenance of the equipment used, the high cost of data acquisition and the fact that detection limits cannot be improved by pre-concentrating the samples. Therefore, vestigial VOCs will not be detected by this approach [26] [32].

1.3.4.2.1 PTR (Proton Transfer Reaction Mass Spectrometry)

PTR-MS [26][38][51][52] application in VOC biomarkers research is increasing, mainly because it can deliver results in a real-time analysis, with high sensitivities for VOCs detection (Table 1.1) and quantification (up to the ppt_v range) [51]. In this analysis H_3O^+ ions are used for proton-transfer reactions with many common VOCs, having almost no reaction with the abundant atmospheric gases (N_2, CO_2 and H_2O) [51][52]. However, PTR-MS has some limitations. This methodology does not allow the identification of compounds with the same molecular weight, because the detection relies on the atomic mass of compounds and the resolution of MS instruments is limited [24]. Therefore, a time of flight mass spectrometer can be linked to the PTR (PTR-TOF-MS) [53][54] to overcome this issue. In this technique, the ions are accelerated to a regular energy by an electric field. Then, the ions travel a defined distance without acceleration and the m/z will determine the time of flight of the compound. This methodological improvement makes possible the separation between distinct chemical compounds with the same molecular weight [38] [53]. However, as mentioned above, since pre-concentration is not possible, trace VOCs can hardly be detected using this approach, and this procedure is a much more expensive technique than GC-MS [53].

1.3.4.2.2 SIFT (Selected Ion Flow Tube Mass Spectrometry)

SIFT-MS [26][38][55] is a technique that allows the measurement of trace concentrations of VOCs in humid air, including breath samples. In a general way, VOCs are collected into the flow tube and ionized with precursor ions (usually H_3O^+, NO^+, or O_2^+), forming the product ions, which are then quantified by MS [56]. This technique, just as PTR-MS, has some disadvantages. Due to the chemical ionization process, not all compounds are detectable (e.g. small hydrocarbons cannot be detected due to their low proton affinity) [26]. The issue of the proper identification of compounds is adressed in SIFT-MS by using different reactant ions (H_3O^+, NO^+, or O_2^+), which exhibit different ion-molecule reactions. Due to the different precursor ion generation, sensivity of SIFT-MS detection and quantification (Table 1.1) is lower (ppb_v range) than PTR-MS (ppt_v range) [26][38].

1.3.4.2.3 IMS (Ion Mobility Spectrometry)

IMS [57] was initially developed for the high sensitive detection of ilegal drugs and explosives, but then it was adapted to industrial and environmental applications, particularly for process control in food quality analysis and air quality control [58]. In the IMS analysis, ions are separated based on their mobility as they travel through a purified gas, in an electric field at the atmospheric pressure [58][59]. This can be achieved using commercially available IMS, without and with different gas chromatographic columns, as MCC (multi-capillary column) /IMS [60],[61]. The sensitivities that can be accomplished with IMS have made it suitable for breath analysis (ppb_v - ppt_v range) [38]. There are already some sucessfull examples of some IMS strategies applied in the diagnosis of pulmonary diseases (lung cancer, lung infections and asthma) as well as other bacterial infections [60]–[63].

1.3.4.2.4 SESI-MS (Secondary Electrospray Ionization Mass Spectrometry)

SESI [64] ionization occurs by proton transfer reactions between the electrospray solution and the volatile analyte, and is therefore suitable for the analysis of hetero-organic molecules, just as in traditional electrospray ionization (ESI). However, unlike the standard procedure, the proton transfer process of SESI occurs in the vapor phase rather than in solution [64]. The distinctive advantage that SESI provides over other ionization methods is that it is possible to fragment specific peaks (provided the appropriate type of mass spectrometer has been applied for SESI), which is an important tool for compound identification. SESI-MS has a sensitive detection limit (ppt_V range) [65] and it has been applied to the detection of explosive gaseous samples, human breath vapor, as well as in the identification of clinically relevant pathogens [66][67].

1.3.4.2.5 IMR (Ion Molecule Reaction Mass Spectrometry)

Electron impact ionization used in conventional mass spectrometry (EI-MS) leads to dissociative ionization of neutral gaseous compounds, thus creating complex fragmentation patterns. This fact limits the identification and quantification of gas mixtures containing different compounds of the same chemical group [68]. To overcome this limitation the ionization of small molecules via ion-molecule reactions (IMR), can be applied, which allows the reduction of fragmentation caused by high energy electron impact ionization. The IMR-MS technique was initially used to measure absolute gas concentrations of cars emissions, caloric plants (fermentation and catalytic processes) and to medical applications [69]. Nowadays, it is also used to the analysis of microbial headspace VOC composition for bacterial species differentiation [70]. This method uses soft chemical ionization for sample molecule ionization and displays no or only minimal fragmentation. The IMR sensitivity varies (ppm_v-ppb_v range) depending on the components measured, system setup and settings. Also, it has the capability of measuring compounds within milliseconds [71]. In this technique, positively charged atomic

ions interact with neutral sample gas molecules [70]. The two-body collision processes result in the formation of product ions whenever the ionization potential of the sample molecule is lower than the potential energy of the incoming primary ion. Differences in ionization potential between primary and product ions may result in a bond rupture and hence a lower molecular weight fragment ion. However, fragmentation is typically avoided due to the soft ionization process [69].

Table 1.1- Comparison of the mode of operation, sensibility, advantages and disadvantages of the different analytical methods.

Analytical method	Mode of operation	Sensibility	Advantages	Disadvantages	Refs
GC-MS	Off-line	ppt_v-ppb_v	Reproducible Identification of unknown VOCs and profiling possible	Pre-concentration needed; Slow; Quantification requires known compounds; Real-time measurements not possible; Expensive; Not suitable for clinical use	[26][38]
PTR-MS	Real-time	ppt_v	No pre-concentration needed Potential for on-line testing	VOC chemical identification and complete profiling not possible	[38][26] [51][52]
SIFT-MS	Real-time	ppb_v	No pre-concentration needed; measures in real-time; fast; Potential for on-line testing; Measures in headspace possible	VOC chemical identification and profiling not possible	[26][38][5 5][56]
IMS	Real-time	ppb_v-ppt_v	No pre-concentration needed; Low cost; Suitable for clinical use	Identification of unknown compounds is not possible	[38] [57]–[59]
SESI-MS	Real-time	ppt_v	No pre-concentration needed; Fast	Complex VOC mixtures can cause unreliable results	[64][65]

Analytical method	Mode of operation	Sensibility	Advantages	Disadvantages	Refs
IMR-MS	Real-time	ppm_v-ppb_v	No pre-concentration needed; Fast	Sometimes formation of secondary ions that may have the same weight as the primary ions	[68]–[70]

1.4 Data processing

The statistical data treatment that follows the analysis step can be particularly inconvenient [26]. Although there is a full range of tools available to handle data complexity, until now there is no agreement regarding the selection and usage of those tools to discover volatile biomarkers that work with acceptable sensitivity and specificity for clinical applications [38]. In the majority of cases there are complex relationships between the group of compounds found in a clinical sample. For these reasons, volatiles identification and profiling using bioinformatics is a promising approach [72]. Specially, when adopting a strategy of identifying patterns instead of individual VOCs a more elaborate method of data analysis is required, such as Principal Component Analysis (PCA) [73] and Partial Least Squares Discriminant Analysis (PLSDA) [27], which allow a reduction of the dimensionality of data. Di Natale et al [74], for instance, found a set of putative biomarkers for lung cancer, in a group of 42 patients, using e-nose and GC-MS methodologies and recurring to PLSDA to analyze the data. On the other hand, Montuschi *et al.* [75] used PCA analysis to obtain an ashtma VOCs pattern based on 27 patients. Data can then be plotted and a visualization of similarities and diferences between data sets is possible. Also, it is possible to identify individual components, instead of patterns, that will be responsible for the diferences observed between data sets and, finally, identify if some of those compounds are biomarkers [72].

1.5 Electronic noses (e-noses)

In order to measure different VOCs, many applications have combined various sensors and materials into a single array, leading to the development of a device able to detect and distinguish odorous compounds- an e-nose [76]. Electronic noses follow an approach which closely resembles mammalian olfaction, by measuring the whole spectrum of VOCs without identification of the individual components [77]. Although individual VOCs cannot be identified, the output of e-noses represents a signature of the VOC pattern (fingerprint) (Figure 1.2), which can be analyzed by pattern recognition algorithms to discriminate VOC mixtures and potentially to detect diseases [19]. There are several formats for e-nose sensors, [77] [78], which are summarized in Table 1.2. Unlike GC-MS and other analytical techniques, e-noses do not contribute to the discovery of biomarkers that are specific for a disease.

Nonetheless, they can be used to compare samples to see if they have similar VOC profiles. Also, due to the low-cost and implementable nature of this technology, it has great potential for clinical use [79]. Pavlou *et al.* [80] sucessfully discriminated some bacterial cultures, associated with tuberculosis, such as *M.tuberculosis* and *M.avium*, by using the volatile patterns resulting from an electronic nose based on a 14 sensor conducting- polymer sensor. Wang *et al.* [81] used a colorimetric sensor to analyze 14 breath biomarkers, such as ammonia, acetone and ethane, in a breath analysis study.

Currently, e-nose research is focusing on finding materials with high sensitivity and good selectivity for VOCs detection to improve the sensitivity and specific discrimination between the pathogens producing them [82][83].

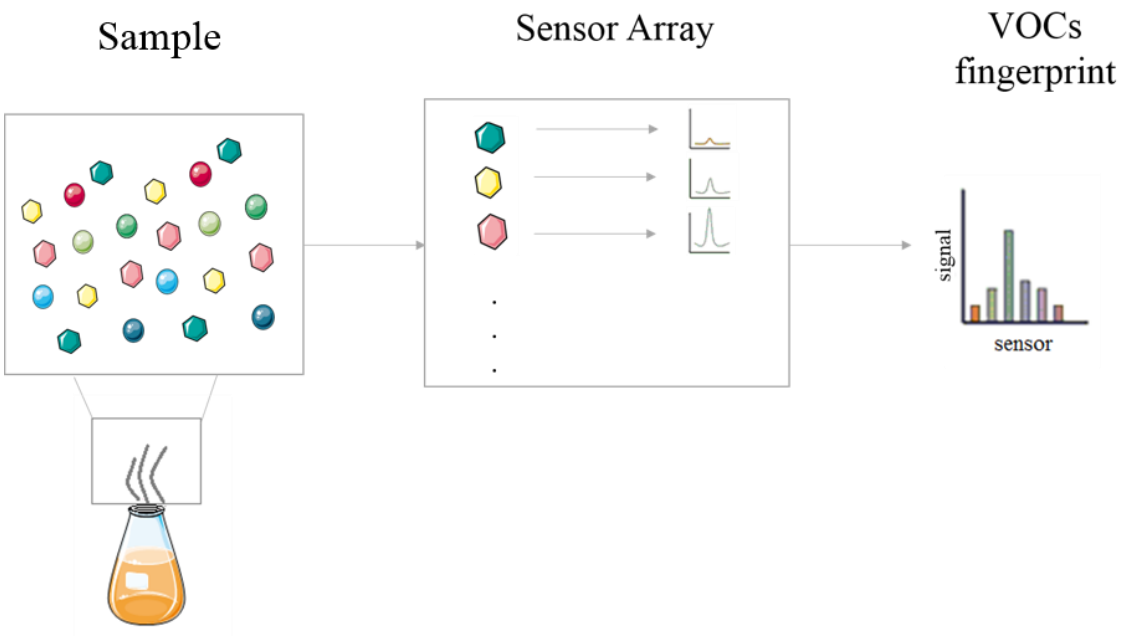

Figure 1.2 – Schematic representation of volatile compounds recognition by an electronic nose device.

Table 1.2- Summary of the different e-noses formats and their mode of operation [78][84].

Sensor Format	Mode of operation
Conducting-polymer sensor	VOCs interact and attach to the polymer surface changing the resistance which results in changes in the signal.
Metal oxide sensor	Oxide materials contain chemically adsorbed oxygen species, which can interact with the VOCs, altering the conductivity of the oxide.
Metal oxide silicon field-effect sensor	Related to metal oxide sensors but the output signal is originated from a change in potential when the VOCs react at a catalytic surface.

Sensor Format	Mode of operation
Piezoelectric crystal	Adsorption of VOCs onto the membrane results on a change in the magnitude of the resonance frequency that is related to the mass of the volatile analyte.
Surface acoustic-waves device	Based on waves emmited along the surface of a crystal by the electric field of surface-deposited aluminium electrodes.
Optical sensor	Based on a light source that interacts with the volatile analyte. The signal is measured in absorbance, fluorescence, reflectance or chemiluminescence.
Electrochemical sensor	Responses are dependent on the electrochemical characteristics of the VOCs that are oxidized or reduced at the working electrode and at the counter electrode. Generated voltage of the reactions between the electrodes is measured.

1.6 Challenges and future directions

Although VOCs profiling is a potential clinical tool, the technique is still not part of routine analysis. Before it is implementated in clinical analysis there are some steps that need to be validated. The first one is an extensive validation of the current available VOCs profiles. Also, further development of the sample-collection devices and the sampling mechanisms is required in order to facilitate taking reliable, reproducible samples [85]. Numerous factors can influence emitted VOCs [86]. More research is required to further identify microbial specific VOCs and this situation is aggravated by the fact that the specific VOCs produced by a given microbe in its natural environment can be different from what is observed *in vitro* due to the use of different growth medium, incubation conditions and possible presence of other microorganisms [40]. Also, the analytical method should be carefully studied. For instance, when direct analysis is not possible, the samples need to be stored, which can affect the original composition of the collected sample [19] [26]. The physiological meaning and biochemichal origin of endogenous VOCs so far are not clear. Although more insight is needed this is not an easy study, because the origin of VOCs can be the result of widely different biochemical pathways [87]. Finally, further refinement of sampling techniques and the development of new tools that combine the strengths of the e-nose (cheap, time efficient) [19][77] IMS (real-time) [57][60]–[63] and GC-MS (sensitive, compound identification) [20] [26] [32] [46] will favour the introduction of VOCs analysis into clinical practice.

2. Aim of the work

Non-invasive diagnostics of microbial infections based on the volatome is an area of increasing interest. E-noses have emerged as a low-cost technology with great potential for clinical use. Also, LC based-devices have been reported as reliable, low-cost and high-sensitive gas sensors for biological and chemical sensing [88][89]. However, there is a need to find sensitive materials with good selectivity for VOCs detection to improve the specific discrimination between the pathogens.

In the first part of this work a search for infectious diseases VOC biomarkers was conducted (Figure 2.1). A literature search was performed and data of interest was collected and organized in a database. A statistical analysis was performed and machine learning algorithms were used to design and develop a model to distinguish pathogens based on the detected VOCs and to identify a possible set of VOC biomarkers for those pathogens.

The second part of this work consisted of a proof-of-concept study in which a proprietary e-nose and a gel-like VOC-sensitive material with optical properties were employed. The work aimed to explore a method for tailoring the selectivity of the material towards a particular set of VOCs (Figure 2.1). As a case-study, three different VOC-specific peptides were incorporated into the materials and the response of the modified material was characterized by Polarized optical microscopy (POM) and further tested on the e-nose.

The long-term goal of this study is to use both the classifier and the combination of responses of the VOC-sensitive biogels to the different solvents in aid to the identification and distinction of volatiles emitted by clinically relevant pathogens.

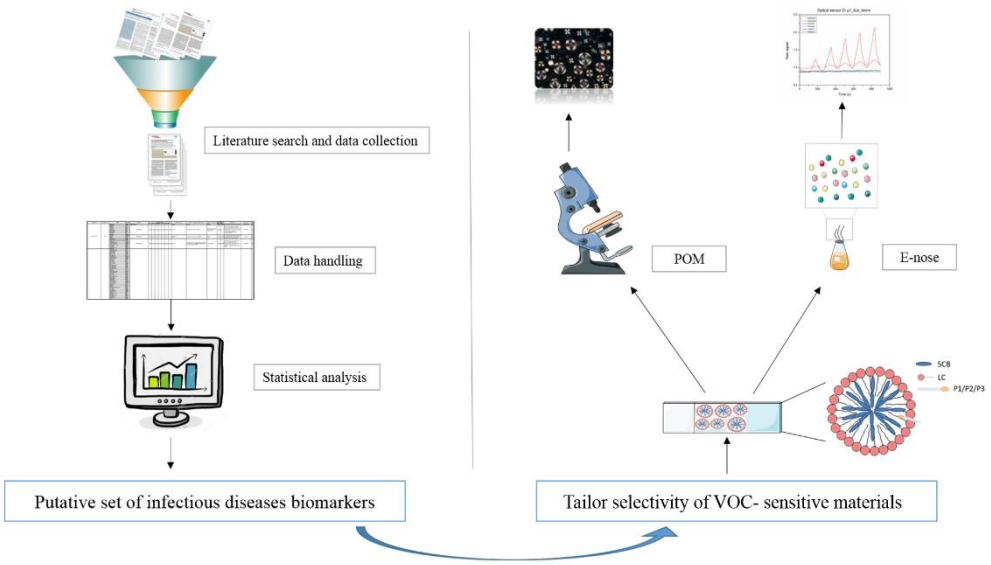

Figure 2.1- Schematic representation of the thesis project.

3. Looking for infectious disease VOC biomarkers

3.1 Introduction

Infectious diseases are one of the major causes of death worldwide. The early detection and identification of the causative microorganisms allows a prompt initiation of appropriate antimicrobial therapy essential for an efficient patient management [3]. Simultaneously, the global spread of antimicrobial resistance to antibiotics is a predominant reason why infectious diseases continue to be target of attention [90]. The misuse of antibiotics is one of the factors that contributes to the selection of drug-resistant pathogens, increasing the need for a correct identification of the microorganisms responsible for infections [91]. The traditional methods for bacterial detection and identification rely on culture and colony counting methods [14], polymerase chain reaction (PCR) [14][16] and immunology-based methods [14][15]. However, these suffer from some major drawbacks. First, the majority can only be used for organisms that can be cultivated *in vitro*. Second, they are still time-consuming and require technical expertise and equipment [4][5][7].

Disease-related biomarkers are chemicals which presence/absence in the body differs according to the health status of an individual [24], indicating the presence or severity of a particular disease state. These biomarkers may have endogenous (produced within the body) or exogenous (introduced into an organism) origin.

Pathogenic microbes release unique combinations of metabolites in the body, which represent potential infectious disease biomarkers [13]. Analyses of volatile organic chemicals from different bodily fluids (blood, saliva, urine, faeces, milk, breath and skin) have the potential for bacterial identification and differentiation [92][24][93], since some pathogenic metabolites are VOCs, and therefore offer the possibility of fast diagnosis and disease monitoring, when compared to traditional methods. This approach has only began to receive attention recently, mainly because VOCs are present in the body in low concentrations (ppt_v-ppm_v), making it essential to use analytical methods with high sensitivity ranges [23]. The advances in analytical techniques increased the potential for VOCs detection and GC-MS has become the gold standard instrument for headspace VOC analysis as it offers high sensitivity (ppt_v-ppb_v) and extensive compound libraries are available, making compound identification easier [19]. Before VOC analysis can be implemented in clinical diagnosis, possible volatile biomarkers must be known. Currently, the search for VOCs as disease biomarkers has been the focus of many studies that, in some cases achieved different results. Kunze *et al.* [61], for instance, identified 2-ethyl-1-hexanol, acetone, 2-phenylacetaldehyde, ammonia (dimer), dodecane, nonanal and ammonia in the headspace of a *P.aeruginosa* isolate obtained from a blood sample. On the other

hand, Savelev *et al*. [39] found 2-nonanone, 2,4-dimethyl-1-heptene, 1-heptene, isopentanol and limonene in a *P.aeruginosa* isolate obtained from a breath sample. The use of distinct samples (breath, blood, urine, skin, faeces), sampling methods and analytical techniques contributes to a variety of results, making difficult to have firm conclusions about the relevance of each VOC as infection biomarker.

In this chapter, we performed a systematic review of existing literature, and analysis of published results relating the detection of VOCs in human samples with the presence of specific pathogen infections. Hence, we suggest a minimal set of VOCs that allow the distinction between 8 clinically relevant pathogens - *Pseudomonas aeruginosa, Staphylococcus aureus, Aspergillus fumigatus, Escherichia coli, Helicobacter pylori, Proteus mirabilis, Klebsiella pneumoniae* and *Mycobacterium tuberculosis*. The set of volatile biomarkers found in our analysis was compared with the biomarkers found in other works, for the same pathogens. As a result, we identified a set of potential biomarkers for pathogen infection: indole for *E.coli*; 2-pentylfuran for *A.fumigatus*; isobutane for *H.pylori*; cymol for *M.tuberculosis*; hydrogen cyanide and methyl thiocyanate for *P.aeruginosa*; and 3-methylbutanoic acid for *S.aureus*.

3.2 Materials and methods

3.2.1 Literature search and data collection

Data was collected from bibliography obtained through a systematic search in the PubMed and the Web of Science on-line libraries, performed in the period between October 12th, 2015 and February 10th, 2016. The search terms in PubMed were: “volatile organic compounds”; “volatile organic compounds + health/breath/pathogen/disease”, “volatile biomarkers + disease” and “exhaled volatile organic compounds”. For the Web of Science search, the terms were the same except for “volatile organic compounds”, that was not used alone, because the PubMed search with that same term included many articles that were not associated with health (pathogen).

The retrieved articles were selected for examination if the title and/or abstract suggested the investigation of microbial pathogens and the measurement of VOCs regarding human clinical perspective.

Further screening of the selected articles was performed according to the flowchart depicted in Figure 3.1. Relevant articles were selected for the study according to a set of inclusion criteria:

i. The article’s subject should concern the human clinical research field (plants, soils and animals research fields were not included);
ii. The article should indicate the name of the disease-associated pathogen (the disease name alone was not informative enough);

iii. The article should report quantitative or qualitative information regarding the individual VOCs instead of reporting only VOC patterns;

iv. The analytical method used to detect and quantify the VOCs should be described;

v. The article should provide a summary of the detected VOCs per pathogen.

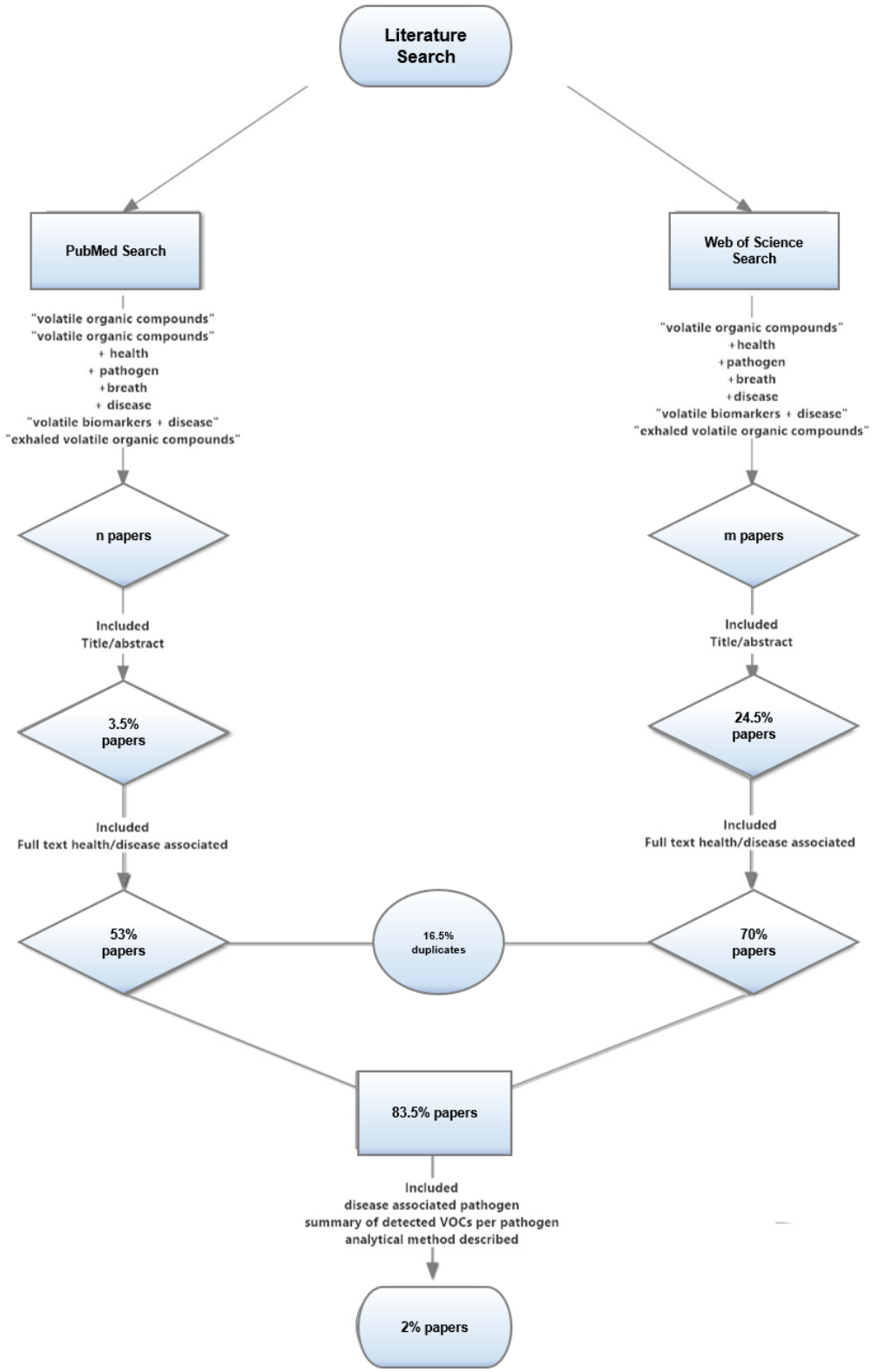

Figure 3.1- Strategy followed for the selection of papers.

3.2.2 Data handling

Data of interest was collected after reading the fulltext of the selected articles and organized in a database. A major table was compiled with information retrieved from the articles, organized in the following columns: pathogen identification (name and classification), bacterial strain (when applicable), VOCs associated with each pathogen and corresponding PubChem ID (for univocal VOC identification), the type of sample (saliva, blood, breath, skin, urine, faeces and milk) where each VOC was identified (when applicable), the analytical method used to detect it, culture conditions/growth medium, incubation time before analysis, detected VOC concentration range (value and unit) and the respective bibliographical reference (Appendix 2).

Since data collection and respective organization in the database was performed by a single subject, table filling errors might occur. To quantify those errors, data validation was performed by a second independent subject. 3 articles were chosen randomly from the set of 44, and the database was filled by the second subject with the information collected from the articles. It was found that in 100 VOCs present in those articles there were 3 table filling errors, resulting in an associated error of 3%.

A new database was created by a transformation on the structure of the previous described database to facilitate further data processing. Data was re-organized and a new parameter was added: the number of experiments. Some articles included results from more than one experimental condition: for example, results obtained with distinct growth media, with different analytical methods, or even with multiple bacterial strains. To account for these situations, for the same article, each dataset obtained in a specific experimental condition was considered as a distinct experiment. Hence, an article may describe several experiments and one pathogen may have a higher number of associated experiments than the corresponding number of papers. The database organization was one entry per experiment related to one pathogen, and the Boolean (true or false) indication of the detection of each of the VOC.

The number of hits was considered as the number of times that each VOC was detected in all of the experiments. In some cases data was normalized to percentage of the total number of experiments to facilitate data visualization and interpretation.

Cytoscape software was used as a visualization tool, to generate pathogen-VOC interaction graphs that allow an easy identification of the evolution of data processing during this work.

3.2.3 Statistical analysis

A new filtering criteria was applied to the full dataset (major table modified), in order to reduce the unbalance of data between pathogens. Therefore, a sub-dataset including only pathogens with more than 3 associated experiments was used. The result was 8 pathogens: 2 $Gram^+$ bacteria, 1 fungus and 5 $Gram^-$ bacteria.

The main goals of this study were to devise a model to distinguish pathogens based on the detected VOCs and to identify a putative set of VOC biomarkers for those pathogens. For that, machine learning methods based on statistical classification were used to design and develop the algorithms for pathogen classification from the transformed sub-data matrix with 8 pathogens, 269 VOCs and 174 experiments. Each line of the matrix is a features vector, which refers to an experiment where a pathogen was present and contains a binary vector reporting the identification (or not) of a VOC. Each VOC presence is considered a feature in the features vector.

A set of computing steps were executed in order to generate classifiers and estimate the classification error rate (detailed in Figure 3.2). This computational work was performed by Prof. Hugo Filipe Silveira Gamboa (Faculdade de Ciências e Tecnologia- Universidade Nova de Lisboa, Departamento de Física).

The first computational step consisted in transforming the database with the collected information on the papers to generate the binary vector of features. We separated the data into training and test datasets for validation purposes. Then, we executed a feature selection mechanism to identify a good subset of features that generated low classification error. The classifier was trained in the process to search for the best subset of features. The process ended with a validation step where we computed the final error and classifier behavior by computing the confusion matrix when a dataset not used in the training phase was used.

In the classification process we used statistical based classifiers to be able to separate the pathogens based on the binary VOC input data.

We tested several standard classifiers [94]: decision trees, naive Bayes classifier, nearest neighbour classifier and support vector machines (SVM). In our preliminar tests the results generated by the SVM always outperformed the other classifiers. We selected the SVM with linear kernels as the classifier to execute the feature selection process.

The support vector machine [95] classification method operates a transformation on the data projecting the data to a higher dimension space than the original data structure and applies an optimization technique to find an optimal separation plan in the new transformed space. In

figure 3.3 we show an example of a separation plan selection that maximizes the separation margin between the two classes.

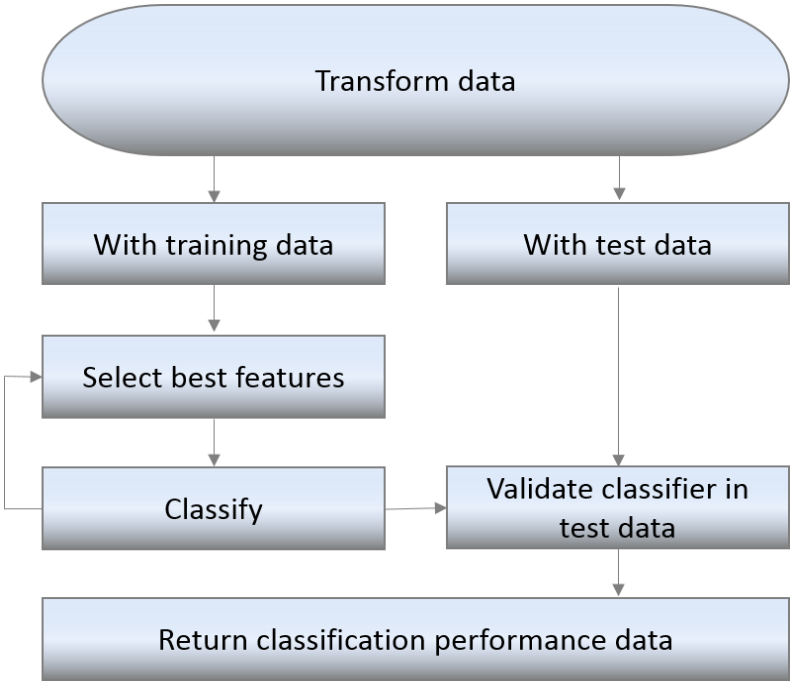

Figure 3.2- Classification steps.

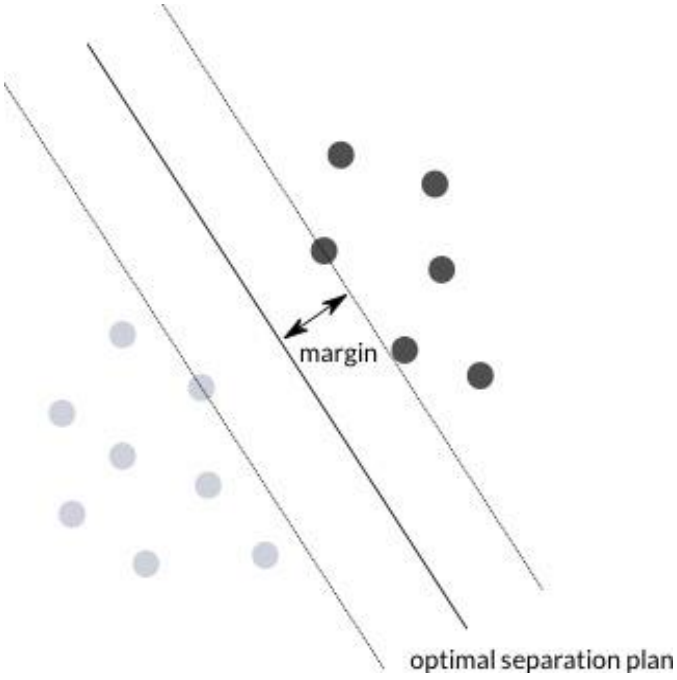

Figure 3.3- Support Vector Machine optimal separation plan and separation margin.

The base learning process of the SVM optimizes the margin distance by selecting a separation plan of a two class problem. This process is replicated to each pair of pathogens.

The classification task was performed in two modes [96].

i. Multi- class: we defined a multi-class problem where we considered each pathogen as a separated class and a set of SVM were adjusted to each pathogen. This mode is also called identification mode, when we have several classes and we want to identify to which class the data belongs (multi-class problem). The question this classifier will address is: "Based on these VOC what is the most probable pathogen (from the set of selected pathogens)".

ii. Dual-class: for each pathogen we were interested in verifying if the sample VOCs corresponded to the pathogen or to any other one. This is also called verification, where we are interested in verifying if our assumption of the data belonging to a specific class is true. We are answering to this type of questions: "Does the new data belong to, e.g. *P. aeruginosa*?"

A selection of the best VOCs subset was executed by standard feature selection mechanisms implemented in both modes of classification.

Two non-optimal mechanisms are normally applied: the sequential forward feature selection and the sequential backward feature selection. In the first case we start with an empty vector of features, adding one feature at a time and growing the vector until the classification error stops decreasing. In the backward mechanism we start with the full feature vector and remove one feature at a time, reducing the dimension of the feature vector. We used the later one because it requires a lower computational complexity and the results tend to be similar [97]. The steps executed in the sequential forward selection algorithm are depicted in figure 3.4.

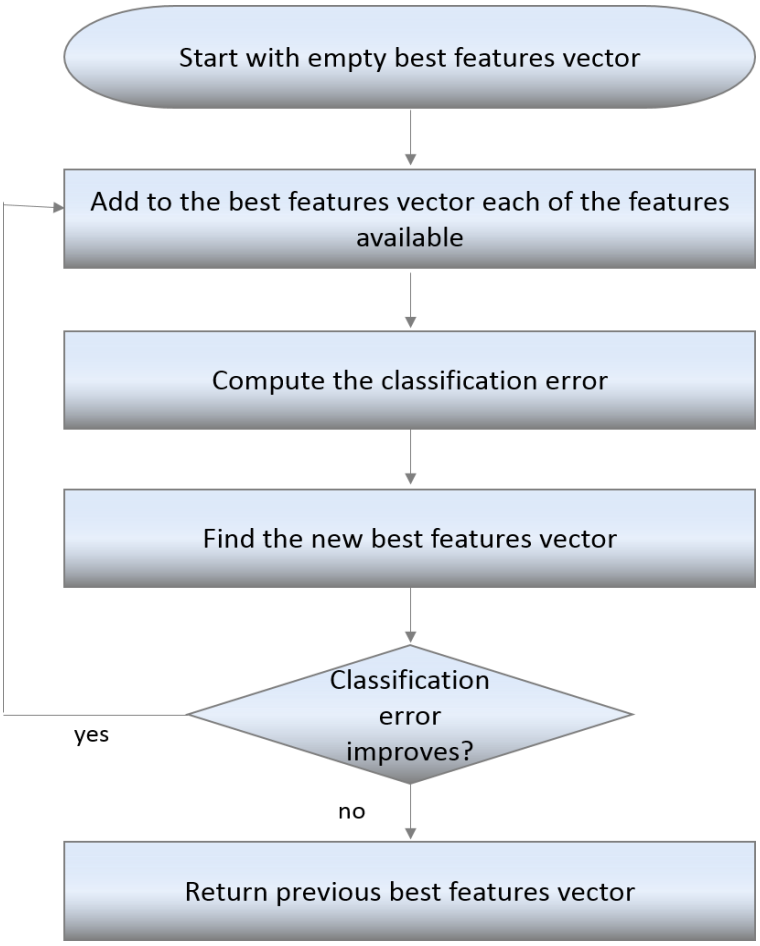

Figure 3.4- Sequential forward feature selection.

In the case of multi-class, we executed the feature selection for all the classes, returning a vector of the best features for separating the pathogens.

In the dual class problem, where we verified the possibility of a pathogen against all the other pathogens, we executed the feature selection for each case returning for each pathogen a set of features that better separated the pathogen from all the others.

All the results are reported based on a cross validation mechanism where we use a training dataset to find the best features and train the classifier, and a testing dataset where we report the classification error rate and the confusion matrix. We used the leave-one-out [98] cross-validation method, that removes only one pathogen example from the training data set and tests the classification in this sample that has never been presented to the classifier. The results are the average values of running this process for each pathogen example.

3.3 Results and discussion

3.3.1 Pathogen-VOC interactions in the bibliography

The PubMed search resulted in 9738 articles of which 341 articles were selected based on title and/or abstract. Full text was read and tested for inclusion criteria, and this step resulted in the inclusion of 180 articles in this study (Figure 3.1). The Web of Science search resulted in 474 articles. 116 articles were selected based on title and/or abstract. After reading the full text, 81 additional publications fitted the criteria. The results obtained using the two online databases were compared and duplicated hits were removed, finally resulting in the inclusion of 44 articles in the study, based on their full text content.

The included articles were published between 1977 and 2016 (Figure 3.5), with an accentuated increase in the number of publications between 2011 and 2012. Most of the publications concerns the last 11 years (2005-2016), corresponding to 88.2% of the total number of collected articles, while the articles concerning the previous years correponds to 11.8% of the total articles, showing that VOCs have been increasingly studied as potential disease biomarkers over the last 10 years.

In total, the 44 articles report 23 pathogens, 418 VOCs and 199 experiments. The number of experiments present in each articles varies. Papers 11 and 7 are the ones with more associated experiments (36 and 26, respectively) while there are many papers with only 1 reported experiment, such as articles 8, 9 and 10 (Figure 3.6).

The selected articles refer to 23 distinct disease associated pathogens (Table 3.1 and Figure 3.7): *Pseudomonas aeruginosa (PA), Staphylococcus aureus (SA), Methicillin-resistant Staphylococcus aureus (MRSA), Streptococcus pneumoniae (SP), Klebsiella pneumoniae (KB), Haemophilus influenzae (HI), Aspergillus fumigatus (AF), Morganella morganii (MM), Proteus mirabilis (PM), Proteus vulgaris (PV), Staphylococcus epidermidis (SE), Enterococcus faecalis (EF), Candida albicans (CA), Escherichia coli (EC), Helicobacter pylori (HP), Legionella pneumophila (LP), Clostridium difficile (CD), Campylobacter jejuni (CJ), Mycobacterium tuberculosis (MT), Giardia duodenalis (GD), Plasmodium falciparum (PF), Neisseria meningitidis (NM)* and *Moraxella catarrhalis (MC)*.

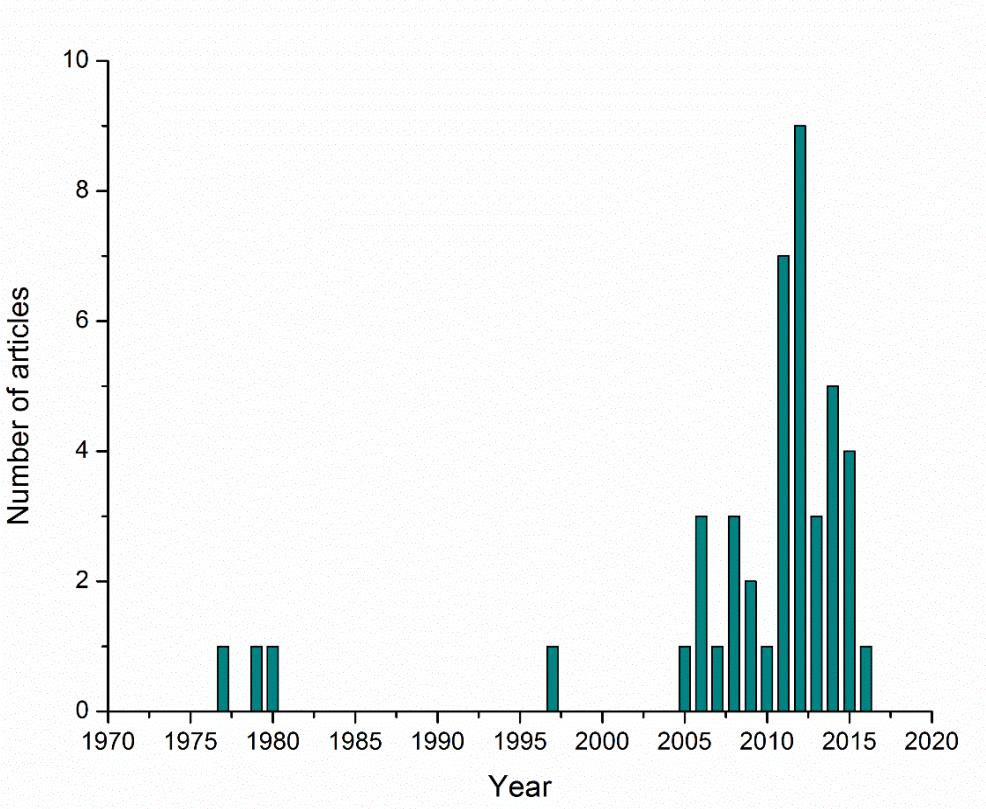

Figure 3.5- Representation of the number of articles per year.

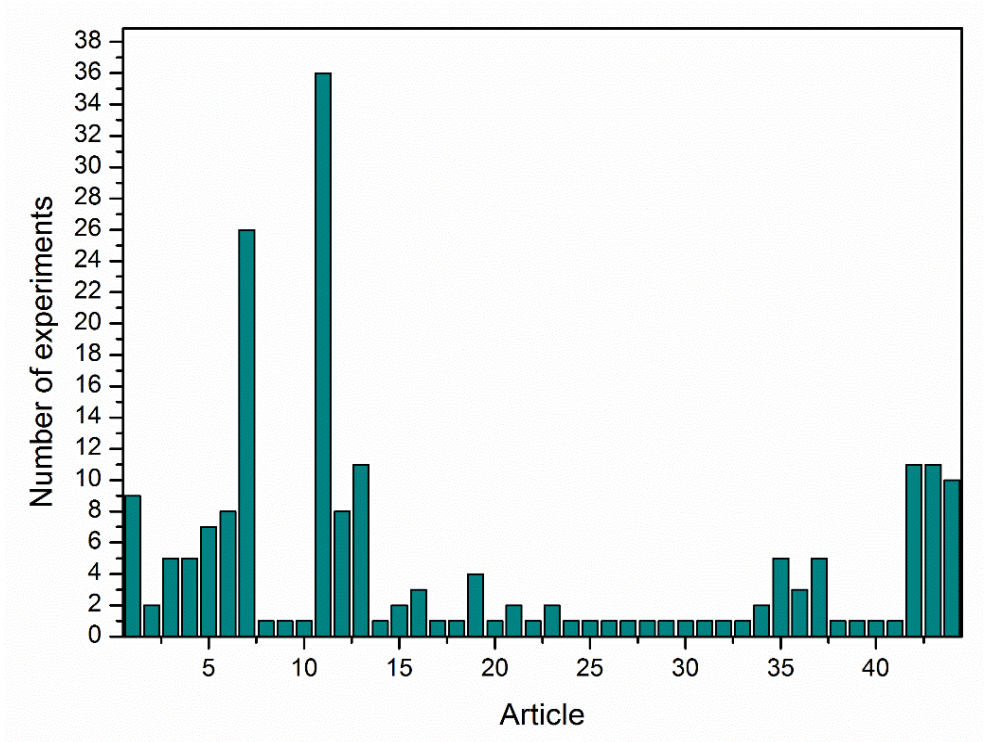

Figure 3.6- Representation of the number of experiments reported in each article.

Table 3.1- Pathogens referred in the selected papers, respective classification, analysis methods used to detect VOCs and number of experiments associated with each pathogen. Pathogens marked (*) were studied in more than 3 experiments. *Pseudomonas aeruginosa (PA), Staphylococcus aureus (SA), Methicillin-resistant Staphylococcus aureus (MRSA), Streptococcus pneumoniae (SP), Klebsiella pneumoniae (KB), Haemophilus influenzae (HI), Aspergillus fumigatus (AF), Morganella morganii (MM), Proteus mirabilis (PM), Proteus vulgaris (PV), Staphylococcus epidermidis (SE), Enterococcus faecalis (EF), Candida albicans (CA), Escherichia coli (EC), Helicobacter pylori (HP), Legionella pneumophila (LP), Clostridium difficile (CD), Campylobacter jejuni (CJ), Mycobacterium tuberculosis (MT), Giardia duodenalis (GD), Plasmodium falciparum (PF), Neisseria meningitidis (NM)* and *Moraxella catarrhalis (MC)*.

Pathogen	Classification	Analytical Method	Number of experiments	Type of sample	Refs.	Example
Aspergillus fumigatus*	Fungus	GC-MS; IMS; SIFT-MS	**4**	clinical isolate reporter-labeled strains	[24][37], [55], [99]	*AF* was detected in a breath sample headspace, associated with sinusitis [24]
Campylobacter jejuni	Gram$^-$ bacterium	GC-MS	1	clinical isolate	[100]	*CJ* was detected in faecal samples headspace, associated with gastroenteritis [100]
Candida albicans	Fungus	SIFT-MS	1	reporter-labeled strains	[101]	*CA* was inoculated into healthy males urine samples [101]
Clostridium difficile	Gram$^+$ bacterium	GC-MS	2	clinical isolate	[100] [102]	*CD* was detected in faecal samples associated with gastroenteritis [100]
Enterococcus faecalis	Gram$^+$ bacterium	SIFT-MS	2	clinical isolate reporter-labeled strains	[101] [103]	*EF* was inoculated into healthy males urine samples [101]
Escherichia coli*	Gram$^-$ bacterium	GC-MS; IMS; IMR-MS;PTR-MS; SESI-MS; SIFT-MS	**36**	clinical isolate reporter-labeled strains	[101] [61], [62], [70], [92], [103]–[110]	*EC* was detected in blood, urine and skin samples headspaces, associated with urinary tract infection [13]

Pathogen	Classification	Analytical Method	Number of experiments	Type of sample	Refs.	Example
Giardia duodenalis	Protozoa	GC-MS	1	clinical isolate	[111]	*GD* was detected in faecal samples headspace, associated with Giardasis [111]
Haemophilus influenzae	Gram$^-$ bacterium	GC-MS	3	clinical isolate	[40], [112], [113]	*HI* was found in clinical isolates (origin not reported) from patients with sinusitis [113]
Helicobacter pylori*	Gram$^-$ bacterium	GC-MS; PTR-MS	**4**	clinical isolate reporter-labeled strains	[24][114]	*HP* was detected in breath sample headspace, associated with gastric infections [24][114]
Klebsiella pneumoniae*	Gram$^-$ bacterium	GC-MS; SIFT-MS	**8**	clinical isolate reporter-labeled strains	[101] [104] [62] [105] [110]	*KP* was found in clinical isolates (origin not reported) from patients with pneumonia [105] and from patients with urinary tract infections [110].
Legionella pneumophila	Gram$^-$ bacterium	GC-MS	1	unknown reporter-labeled strain	[113]	*LP* was inoculated in blood cultures, associated with Legionellosis [113]
Methicillin-resistant Staphylococcus aureus	Gram$^+$ bacterium	GC-MS	2	clinical isolate reporter-labeled strains	[104] [115]	*MRSA* was grown in Mueller Hinton broth and trypticase soy agar at 37ºC [115]
Moraxella catarrhalis	Gram$^-$ bacterium	GC-MS	2	clinical isolate	[113] [40]	*MC* was found in clinical isolates (origin not reported) from patients with sinusitis [113]

Pathogen	Classification	Analytical Method	Number of experiments	Type of sample	Refs.	Example
Morganella morganii	Gram$^-$ bacterium	GC-MS	1	reporter-labeled strains	[106]	*MM* was inoculated into brain-heart-infusion broth, associated with urinary tract infections [106]
Mycobacterium tuberculosis*	Gram$^+$ bacterium	GC-MS; GC-SAW	**8**	clinical isolate reporter-labeled strains	[24][116]–[121]	*MT* was detected in breath [117], sputum [116] [119] and urine [24] [121] samples headspaces, associated with tuberculosis
Neisseria meningitidis	Gram$^-$ bacterium	GC-MS; SIFT-MS	2	clinical isolate reporter-labeled strains	[108] [122]	*NM* was detected in blood samples headspaces [108]
Plasmodium falciparum	Protozoa	GC-MS	1	clinical isolate	[12]	*PF* was detected in breath samples associated with malaria [12]
Proteus mirabilis*	Gram$^-$ bacterium	GC-MS; SIFT-MS	**4**	clinical isolate reporter-labeled strains	[103] [106] [110]	*PM* was detected in urine sample headspace, associated with urinary tract infection [103]
Proteus vulgaris	Gram$^-$ bacterium	IMR-MS; SIFT-MS	2	reporter-labeled strains	[101] [70]	*PV* was inoculated into healthy males urine samples [101]
Pseudomonas aeruginosa*	Gram$^-$ bacterium	GC-MS; IMS; IMR-MS; SESI-MS; SIFT-MS	**94**	clinical isolate reporter-labeled strains	[24][37][101] [61], [62], [92], [103], [104] [108] [70] [110] [40] [39], [123]–[130]	*PA* was detected in breath [24], blood, urine and skin [61] samples headspaces, associated with pneumonia

Pathogen	Classification	Analytical Method	Number of experiments	Type of sample	Refs.	Example
Staphylococcus aureus*	$Gram^+$ bacterium	GC-MS; IMS; SESI-MS; SIFT-MS	**16**	clinical isolate reporter-labeled strains	[101][62], [92], [103], [104] [105] [108] [40] [123][131]	*SA* was detected in breath samples headspaces, associated with pneumonia [123][131]
Staphylococcus epidermidis	$Gram^+$ bacterium	SIFT-MS	1	reporter-labeled strains	[101]	*SE* was inoculated into healthy males urine samples [101]
Streptococcus pneumoniae	$Gram^+$ bacterium	GC-MS; SIFT-MS	3	clinical isolate reporter-labeled strains	[108] [112] [40]	*SP* was detected in breath samples, from patients with pneumonia [112]

During this work, it was evident that there are pathogens more studied than others. The pathogen with more associated experiments is the bacterium *Pseudomonas aeruginosa* (94), corresponding to 47.24% of the total number of experiments. This pathogen is followed by *Escherichia coli* (36 associated experiments), *Staphylococcus aureus* (16 associated experiments), *Klebsiella pneumoniae* (8 associated experiments), *Mycobacterium tuberculosis* (8 associated experiments), *Aspergillus fumigatus* (4 associated experiments), *Proteus mirabilis* (4 associated experiments) and *Helicobacter pylori* (4 associated experiments) (Figure 3.7). The rest of the microorganisms has less than 4 associated experiments (Table 3.1).

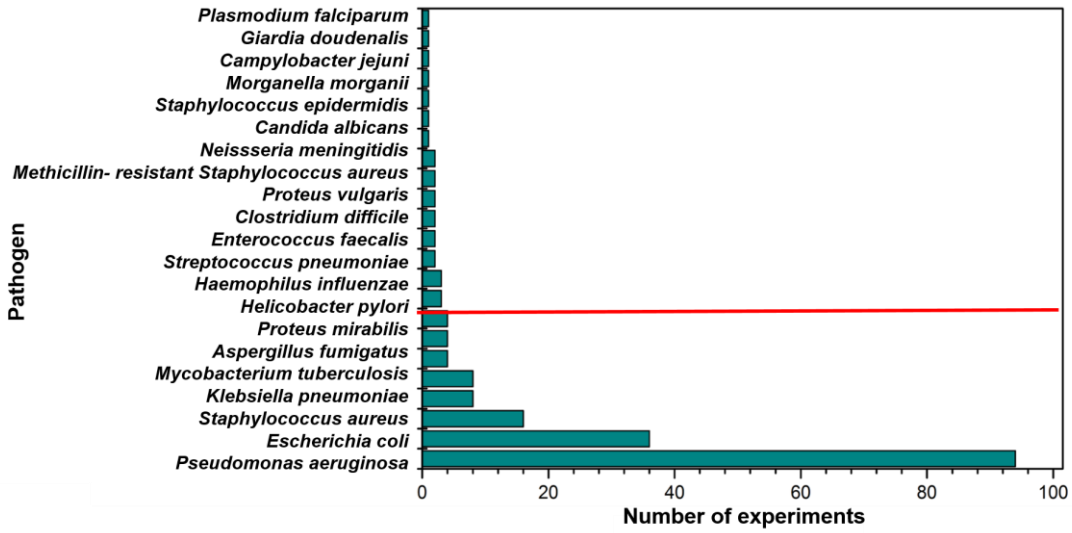

Figure 3.7- Representation of the number of experiments associated with the 23 pathogens. The red line separates the pathogens with more than 3 associated experiments (below the line) from the pathogens with 3 or less associated experiments (above the line).

The experimental analysis of VOCs involves several steps that vary with the chosen analytical method [14][24][84]. The general steps include sampling (collection and storage, if needed, of the sample), headspace analysis [61][86][104], VOC quantification/identification [103][125], data processing [72], and, finally, as an output, putative VOC biomarkers should be obtained.

The experiments associated with each pathogen vary in the type of sample used (Figure 3.8). It was found that most of the experiments used commercial acquired microbial strains, known as reporter-labeled strains (50.2%), 16.6% of the experiments did not indicate the strain used in the study, referring only the microbe species group. The remaining 33.2% of the experiments used human clinical isolates collected from urine (11.1%), breath (8.5%), blood (6.1%), skin (5.5%) and faeces (2.0%). The high percent of experiments using urine and breath may be due to the growing interest in non-invasive diagnostic tools [58][59][60].

Different methods of analysis were used in the studied publications and sometimes more than one analysis method was employed in the same article. Clear differences were observed, regarding the most oftenly used methods (Figure 3.9). The 3 mostly used analytical methods to detect and quantify VOCs were IMS, SIFT-MS and GC-MS, mainly due to their high sensitivity (ppt_v-ppb_v) [16][58][61]–[65]. Interestingly, IMS and SIFT-MS allow real-time analysis while GC-MS only offers an off-line analysis, despite remaining the gold standard method [77][87]. Together these three analytical methods account for 96% of the total number of experiments, while the remaining 4 methods (GC-SAW, PTR-MS, IMR-MS and SESI-MS) only account for 4%.

The reported VOCs were grouped according to their chemical class to allow an easier overall comparison. Some compounds could be fitted in more than one class. In these situations, one of those chemical classes was randomly chosen. For instance, methyl thiocyanate contains both nitrogen and sulfur and it was classified as a sulfur containing compound (See Appendix 1).

The relative abundance of each class was calculated by dividing the number of hits of each class by the total number of hits concerning all the classes (eq.1). The order of abundance of each compound class in all of the experiments was the following: alcohols > hydrocarbons > nitrogen-containing > ketones > sulfur-containing > aldehydes > esters > acids > furans and ethers > others > halogen-containing (Figure 3.10). The most abundant class is the class of the alcohols. However, the abundance values for the 5 most abundant classes are in the same order of magnitude, being almost equally distributed.

$$Relative\ abundance\ (\%) = \frac{Number\ of\ hits_{class}}{Total\ number\ of\ hits} \times 100$$

(eq.1)

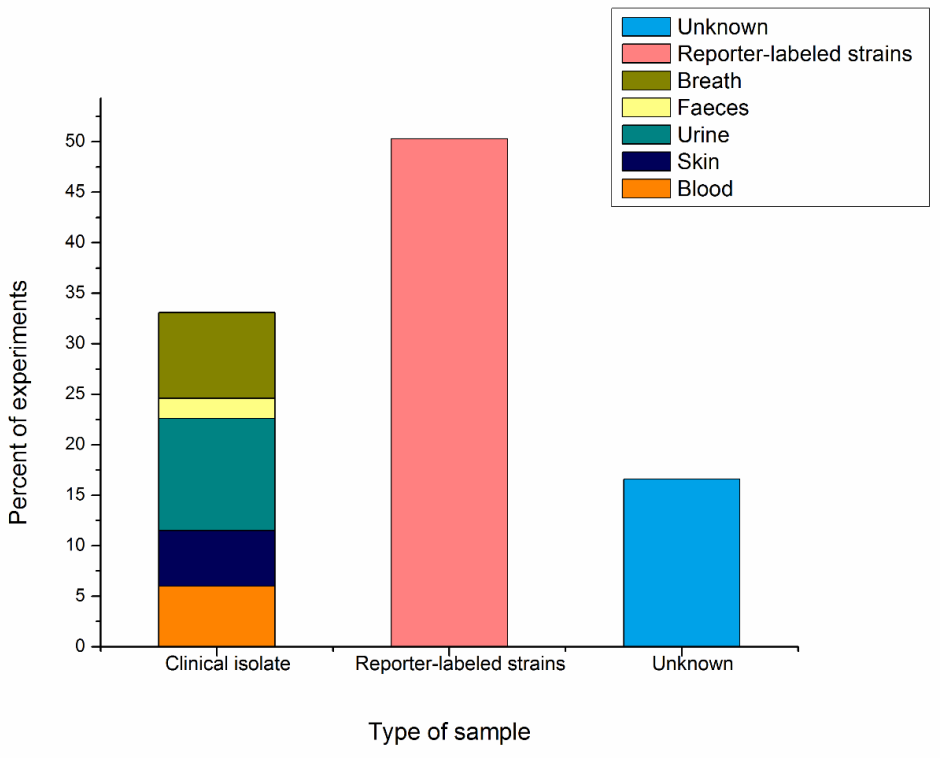

Figure 3.8- Samples used to obtain VOC biomarkers.

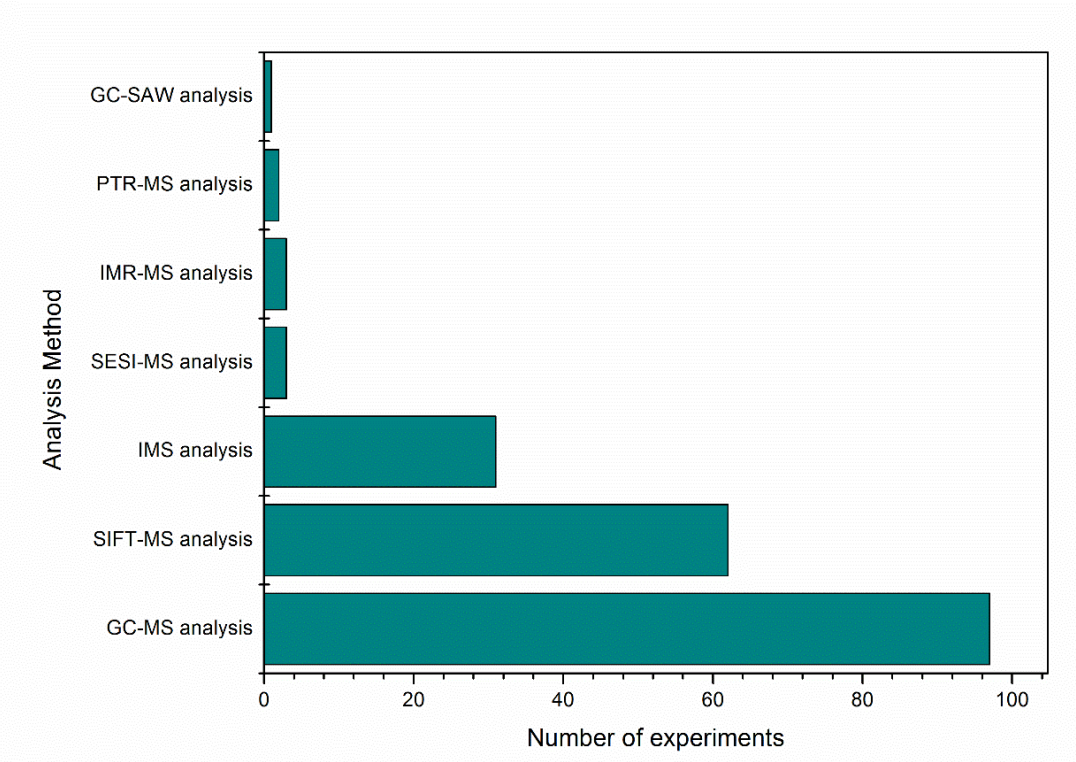

Figure 3.9- Analysis methods used and the corresponding number of experiments.

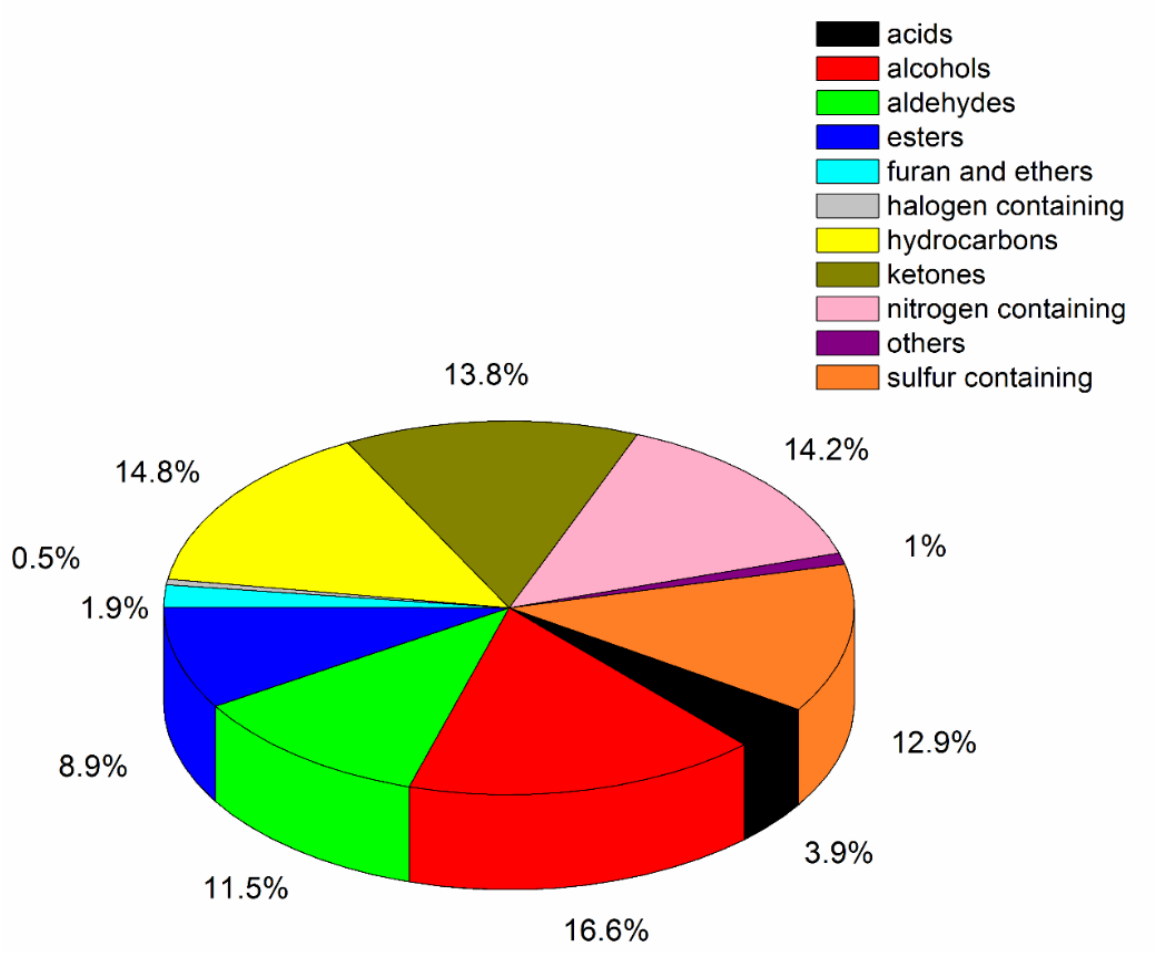

Figure 3.10- Piechart with the relative number of hits (% relative abundance) of compounds in each class that have been detected in all the experiments.

Ethanol is the most frequently detected VOC in pathogen culture headspaces. This volatile is produced by 14 of the 23 studied pathogens (AF, CA, CD, EF, EC, GD, HP, KP, MRSA, NM, PM, PA, SA, SP) (Figure 3.11). Regarding hydrocarbons, 1-undecene is the most reported, but is associated to only 2 microorganisms (PA, KB), (See Appendix 1) and the most referred nitrogen containing compound is hydrogen cyanide, associated with PA and HP (See Appendix 1). The class of hydrocarbons contains 106 distinct VOCs, being the most diverse class (See Appendix 1), while the less diversified are the halogen containing, and acids classes (See Appendix 1).

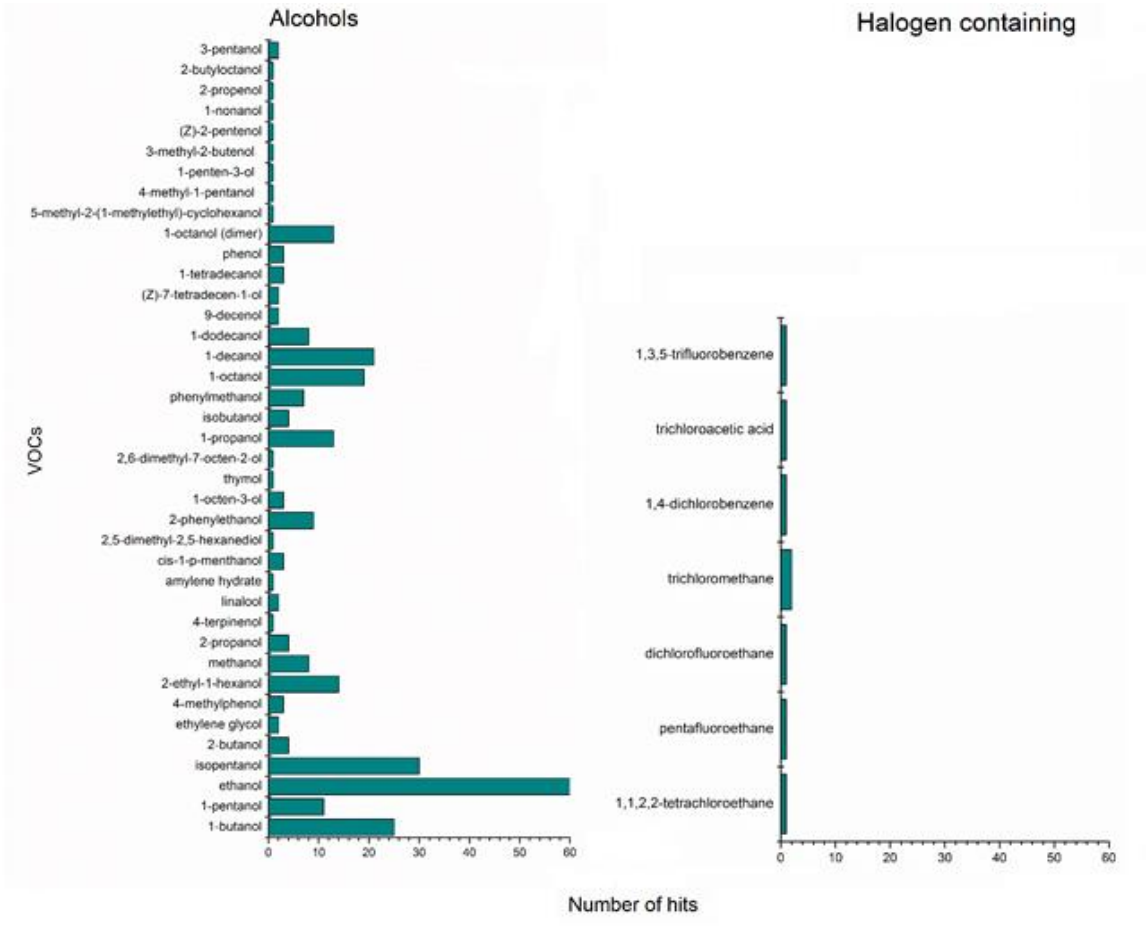

Figure 3.11 - Representation of the number of hits found for each individual VOC, in all experiments, for the alcohols and the halogen containing classes.

The literature search resulted in 44 articles reporting 23 associated pathogens, each of them associated to a VOC pattern. In the circular graph representation shown in figure 3.12 the complexity of the pathogen -VOC relations is notable. While some VOCs are associated with more than one pathogen (the ones represented in the circumference), others were only reported to be associated with one pathogen (the ones outside the circle). For instance, for pathogen 3 (*Escherichia coli*) there is a group of VOCs that is exclusively associated with this bacterium, but these relations were referred only one time (one hit). Those are represented as nodes outside the circle with a straight line connecting them to the *Escherichia coli* node. It is also notable that one of the *Escherichia coli* exclusive VOCs (outside the circle) has been mentioned in several experiments (more than one hit), represented by the oval shaped lines connecting it to *Escherichia coli* and evidenced by a red arrow (inset of Figure 3.12).

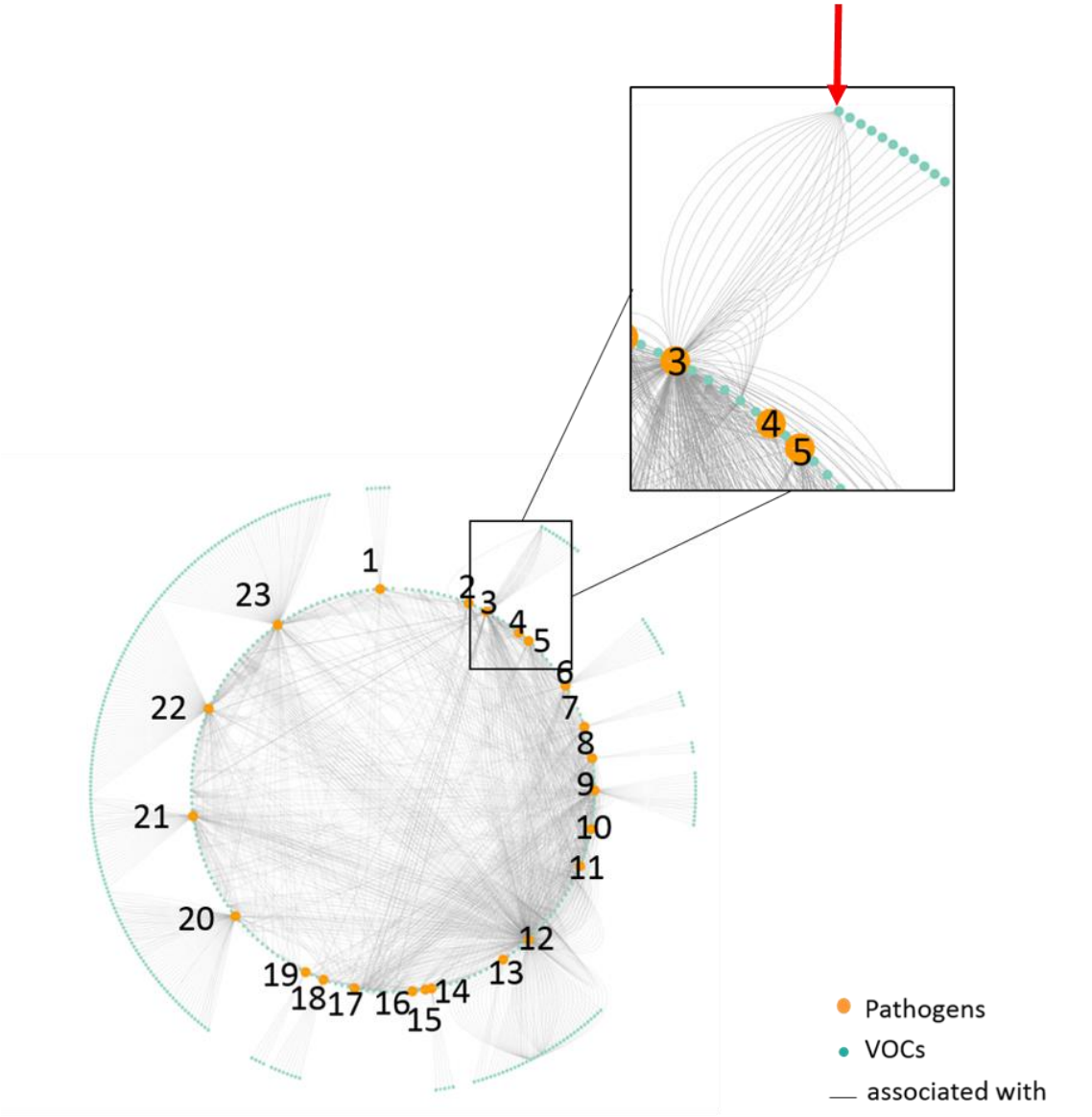

Figure 3.12- Representation of all the 23 pathogens and 418 associated VOCs. *1- Aspergillus fumigatus, 2- Enterococcus faecalis, 3- Escherichia coli, 4-Morganella morganii, 5- Proteus mirabilis, 6-Haemophilus influenzae, 7- Streptococcus pneumoniae, 8- Klebsiella pneumoniae, 9- Staphylococcus aureus, 10- Proteus vulgaris, 11- Staphylococcus epidermidis, 12- Pseudomonas aeruginosa, 13-Plasmodium falciparum, 14- Neisseria meningitidis, 15- Legionella pneumophila, 16-Moraxella catarrhalis, 17- Candida albicans, 18- Giardia duodenalis, 19- Methicilin-resistant staphylococcus aureus, 20- Mycobacterium tuberculosis, 21- Helicobacter pylori, 22- Campylobacter jejuni, 23-Clostridium difficile.* The arrow in the inset highlights a VOC that is associated with one pathogen and that was reported in more than one experiment.

3.3.2 VOCs distinguishing pathogens

It was evident from the circular layout representing the full dataset (Figure 3.12), that a distinction between the pathogens based on the reported associated VOCs could be possible. The next step was to identify the set of VOCs that play the major roles in this distinction. A supervised machine-learning algorithm was used to devise an automated VOC-based pathogen classifier. The VOC-pathogen interaction subset data, including pathogens with more than 3

experiments (8 pathogens, 269 VOCs) was used as input for the algorithm to ensure a better equilibrium between the information in the model. The circular graph for the 8 pathogens (Figure 3.13) shows the significant reduction of complexity, compared with the representation of the full dataset (Figure 3.12). For this set of microorganisms there are still shared VOCs (those in the circumference), however, it is visually remarkable the existence of a group of VOCs exclusive to almost each pathogen (outside the circle), as it is the case of the *Helicobacter pylori* (inset of Figure 3.13).

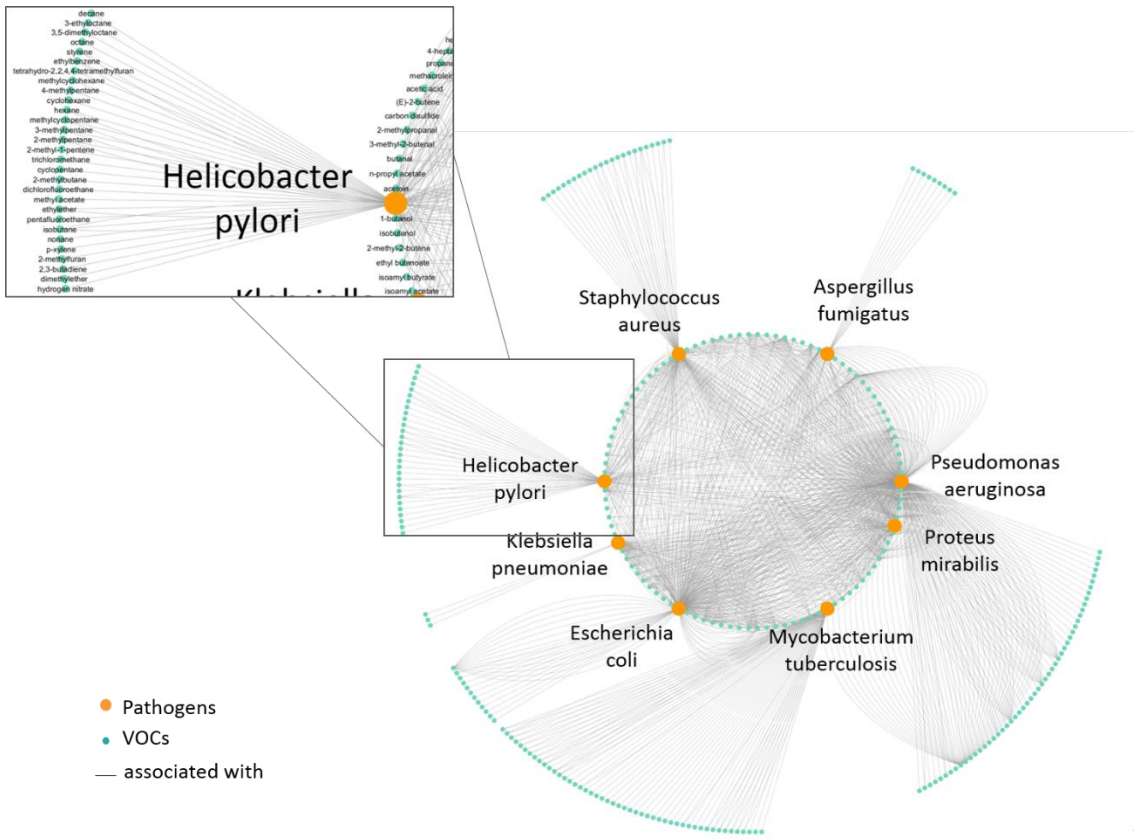

Figure 3.13- Circular layout of the 8 pathogens with more than 3 associated experiments, and the 269 associated VOCs.

From the 269 VOCs in the dataset, a profile with 26 VOCs was identified by the algorithm as the set of VOCs that allows to better separate the 8 pathogens. This selection was based on the VOCs score contributions for the separation model (Table 3.2).

Table 3.2- Set of 26 VOCs that better separates the 8 analyzed pathogens, selected iteractively, by an identification approach. The VOCs are listed from the highest to the lowest score.

Set of VOCs that better separates the 8 analyzed pathogens	
1.	indole
2.	3-methylbutanoic acid
3.	1,3,5-trimethylbenzene
4.	cymol
5.	isobutane
6.	2-phenylanisol
7.	1,1,2,2-tetrachloroethane
8.	1-decanol
9.	1-undecene
10.	isopentanol
11.	acetoin
12.	1-heptene
13.	1-hydroxy-2-butanone
14.	2-methylbutanal
15.	2-pentanone
16.	2-phenylethanol
17.	dimethyl disulfide
18.	ammonia
19.	1-pentanol
20.	3-octanone
21.	acetic acid
22.	hydrogen nitrate
23.	1-dodecanol
24.	1-propanol
25.	isoprene
26.	1-hydroxy-2-propanone

The classification model was validated using the "leave-one-out" cross-validation method. The limitation of evaluating a model without cross-validation is that we do not know how well the classifier will do when it is asked to make new predictions for data that it has never seen before. In our case, the obtained accucary rate was of 81%, slightly less than when no cross-validation was used (94%) (Table 3.3).

Despite having high accuracy rates, in both cases, is important to refer that the classes (pathogens) are not equally represented in the model, since the number of associated experiments varies. So, the reason to the high accuracy may be due to the existance of imbalanced data, because the model looks at the data and decides that the best thing to do is to always predict *Pseudomonas aeruginosa*, since it is the class with more instances.

Table 3.3- Accuracy rates obtained by classification with and without cross-validation.

	Accuracy rate
With cross-validation	0.810
Without cross-validation	0.936

The confusion matrix for the classification with highest accuracy rate (without cross-validation) value shows the correct predictions (diagonal) and the types of incorrect predictions made by the model (which pathogens are being confused by another) (Table 3.4). It can be seen that the classifier never confuses *H.pylori* or *M.tuberculosis* with other pathogens. There are 3 microorganisms that are only mislabeled once, namely: *A. fumigatus* as *P.aeruginosa*, both associated with sinusitis (Table 3.5), *P.mirabilis* as *E.coli*, both associated with urinary tract infections, and *P.aeruginosa* as *E.coli*, both found in urinary tract infections. In addition, *E.coli* was mislabeled twice as *P. aeruginosa*. Finally, *S.aureus* and *K.pneumoniae* were mislabeled 3 times each. The first was confused 2 times with *E.coli* (both found in urinary tract infections) and 1 with *P.aeruginosa* (both found in urinary tract infections, pneumonia and cystic fibrosis). The second was mislabeled 1 time with *E.coli* (both associated with urinary tract infections) and 2 with *P.aeruginosa* (both found in urinary tract infections, pneumonia, blood and bone infections), respectively). So, in total there were 163 true positives (predicted pathogen was the same as the actual pathogen) and 11 false positives, corresponding to a classification accuracy rate of 93.6%. This confusion can be due to the fact that the mislabeled pathogens emit similar compounds. For instance, *P.aeruginosa* and *S. aureus* both share 36 VOCs with *E.coli*, being the pathogens with higher number of shared compounds. Also, regarding these 3 pathogens, within the shared compounds, the most relevant classes are the aldehydes, ketones, alcohols, nitrogen containing and sulfur containing. *P.aeruginosa* and *E.coli*, share 8 aldehydes and 7 ketones (which contain both a carbonyl group), 5 alcohols, 5 sulfur containing and 6 nitrogen containing compounds. Similarly, *E.coli* and *S. aureus*, have 7 aldehydes, 6 ketones, 6 sulfur containing, 5 nitrogen containing compounds and 5 alcohols in common.

Given the class unbalance, in cases where there are reduced samples of the pathogen, the trivial classification as not the pathogen (uninformed accuracy) has already a high accuracy level. To better report the improvements of the classifier learning, we present the base accuracy and compare with the final classification accuracy after selecting the feature subset and train the classifier (Table 3.6). As a result, we obtained a putative set of VOCs that allows a separation between the pathogens with high final classification accuracy values.

Table 3.4- Confusion matrix for the classification without cross-validation, where the column labels represent the predicted pathogen classification and the line labels represent the actual pathogen identity. Red represents the incorrect predictions made by the model and green the correct prredictions (diagonal).

			Predicted							
			A.fumigatus	***E.coli***	***H.pylori***	***K.pneumoniae***	***M.tuberculosis***	***P.mirabilis***	***P.aeruginosa***	***S.aureus***
Actual	Fungus	***A.fumigatus***	3	0	0	0	0	0	1	0
	$Gram^-$ bacterium	***E.coli***	0	34	0	0	0	0	2	0
	$Gram^-$ bacterium	***H.pylori***	0	0	4	0	0	0	0	0
	$Gram^-$ bacterium	***K.pneumoniae***	0	1	0	5	0	0	1	1
	$Gram^+$ bacterium	***M.tuberculosis***	0	0	0	0	8	0	0	0
	$Gram^-$ bacterium	***P.mirabilis***	0	1	0	0	0	3	0	0
	$Gram^-$ bacterium	***P.aeruginosa***	0	1	0	0	0	0	93	0
	$Gram^+$ bacterium	***S.aureus***	0	2	0	0	0	0	1	13

These pathogens are mainly associated with respiratory infections (tuberculosis (MT), pneumonia (SA, KP, PM, PA), cystic fibrosis (AF, SA, PA) and sinusitis (AF, SA, PA)) urinary tract infections (SA, EC, KP, PM, PA), gastric (SA, HP) and gastrointestinal infections (EC, PA) (Table 3.5). Tuberculosis is one of the clinically most relevant disease, since, although it is curable and preventable, affects every part of the world and an earlier diagnosis is essential to a positive outcome. Also, pneumonia is the largest infectious cause of death in children and elderly, worldwide and an early detection of the disease and identification of the envolved microorganism have vital importance. Therefore, the determination of pathogen specific biomarkers is crucial.

Biomarkers have been emerging as a dynamic and powerful approach to screen and detect a disease. In this study we have successfully determinate a set of VOCs that allow pathogen separation, based on VOCs released by microbial pathogens by consolidating results obtained by several authors, respecting pathogens causing infections.

A recent review descriminates 1840 VOC compounds identified from healthy humans, and the respective bodily fluids [133]. Although the typical VOC concentrations ranges in the normal and disease states are not described, it is important to see whether the sets of VOCs that better separates the 8 pathogens, obtained by identification and verification approaches are present in the healthy body or not, and if so in which bodily fluid they can be found. By comparing the healthy body data with the set of VOCs that better separates the studied pathogens, obtained by an identification approach, 3 VOCs were found that are not reported, so far, in the healthy human body (Table 3.7). These compounds are 2-phenylanisol, 1-hydroxy-2-butanone and hydrogen nitrate. Interestingly, each of these 3 VOCs was reported as being associated with only one pathogen and is part of the set of VOCs that better separates determined in this work for that pathogen: 2-phenylanisol was associated twice with *M.tuberculosis* (reporter-labeled strain) [119][120], 1-hydroxy-2-butanone was associated once with *K.pneumoniae* (unknown type of sample) [105] and hydrogen nitrate was associated once with *H.pylori* (reported-labeled strain) [24]. By comparing the list of VOCs that better separates the 8 pathogens, obtained by a verification approach, with the VOCs found in the healthy body, 4 more VOCs were never found in the headspace of any healthy bodily fluid: 1-methylethenyl-pyrazine associated with *A.fumigatus* [37], 3-methylcyclohexene, associated with *K.pneumoniae* [104], 2,3,4,5-tetrahydropyridazine and 2-methylbutanoate, associated with *S.aureus* [105] (Table 3.8). Therefore, the detection of any of these 7 VOCs in some human sample's headspace may facilitate the distinction of pathogens present in an infection.

Knowing the human bodily fluids where a VOC is normally found may facilitate the choice of the most suitable fluid to be analyzed. We verified that when testing for the presence of *E.coli* in

urinary tract infections, the type of collected sample varies between blood, skin and urine [61]. Curiously, indole is reported to be present in healthy skin and urine samples but not in breath, milk and blood samples. Therefore, when investigating the presence of *E.coli*, the presence of indole in blood may be an indicator of an infection caused by that pathogen.

Table 3.5 – Associated diseases/infections to each of the 8 studied pathogens. AR- allergic reactions. BLI- blood infections. BI-bone infections. CF- cystic fibrosis. GI- gastric infections. GTI- gastrointestinal infections. HI- heart infections. KI- kidney infections. PN-pneumonia. SI- sinusitis. SKI- skin infections. TB- tuberculosis. UTI- urinary tract infections.

Pathogen		Associated diseases/infections												
		AR	**BLI**	**BI**	**CF**	**GI**	**GTI**	**HI**	**KI**	**PN**	**SI**	**SKI**	**TB**	**UTI**
A. fumigatus	Fungus	**x**			**x**						**x**			
S. aureus	$Gram^+$ bacterium		**x**	**x**	**x**	**x**		**x**		**x**	**x**	**x**		**x**
E. coli	$Gram^-$ bacterium						**x**							**x**
H. pylori	$Gram^-$ bacterium					**x**								
K. pneumoniae	$Gram^-$ bacterium		**x**	**x**						**x**				**x**
M. tuberculosis	$Gram^+$ bacterium												**x**	
P. mirabilis	$Gram^-$ bacterium								**x**	**x**				**x**
P. aeruginosa	$Gram^-$ bacterium		**x**	**x**	**x**		**x**	**x**		**x**	**x**			**x**

Table 3.6- Representation of the results obtained by the verification approach.

Pathogen	Number of experiments	Set of VOCs that better separates	Uninformed accuracy	Final classification accuracy
A. fumigatus	4	1-methylethenyl-pyrazine 2-pentylfuran cyclohexanone pentanal	0.98	1.0
E. coli	36	indole	0.79	0.92
H. pylori	4	isobutane hydrogen nitrate	0.98	1.0
K. pneumoniae	8	1-hydroxy-2-butanone 3-methylcyclohexene	0.95	0.97
M. tuberculosis	8	1,3,5-trimethylbenzene cymol 2-phenylanisol	0.95	1.0
P. mirabilis	4	(E)-2-butene	0.98	0.98

Pathogen	Number of experiments	Set of VOCs that better separates	Uninformed accuracy	Final classification accuracy
P. aeruginosa	94	ethanol 1-decanol 3-methylbutanal 2-phenylethanol 1,3,5-trimethylbenzene cymol hydrogen cyanide 2-pentene ethane	0.54	0.86
S. aureus	16	3-methylbutanoic acid 1,1,2,2-tetrachloroethane 2,3,4,5-tetrahydropyridazine 2-methylbutanoate pyrimidine	0.91	0.95

Table 3.7- Comparison between the set of VOCs that better separates the group of 8 pathogens, obtained by identification approach, and the VOCs reported in the literature as present (X) or absent ( ) in healthy bodily fluids. NR-not reported.

VOCs	PubChem ID	Healthy body						
		Faeces	Urine	Breath	Skin	Milk	Blood	Saliva
indole	798	x	x		x			x
3-methylbutanoic acid	10430	x		x	x			
1,3,5-trimethylbenzene	7947			x				
cymol	7463		x	x	x			x
isobutane	6360			x				
2-phenylanisol	6835	NR	NR	NR	NR	NR	NR	NR
1,1,2,2-tetrachloroethane	6591		x				x	
1-decanol	8174	x			x			
1-undecene	13190			x				
isopentanol	31260	x	x				x	x
acetoin	179	x		x				x
1-heptene	11610			x				
1-hydroxy-2-butanone	521300	NR	NR	NR	NR	NR	NR	NR
2-methylbutanal	7284	x	x	x	x			x
2-pentanone	7895	x	x	x	x	x		x
2-phenylethanol	6054	x			x	x		x
dimethyl disulfide	12232	x	x			x		x
ammonia	222		x	x			x	
1-pentanol	6276	x	x	x	x	x	x	x
3-octanone	246728	x	x	x		x		x
acetic acid	176	x	x	x	x	x		x
hydrogen nitrate	944	NR	NR	NR	NR	NR	NR	NR
1-dodecanol	8193	x			x			x
1-propanol	1031	x	x	x		x	x	x
isoprene	6557			x	x	x	x	
1-hydroxy-2-propanone	8299	x		x				

Table 3.8- Comparison between the VOCs that better separates the 8 pathogens, obtained by a verification approach, and the VOCs reported in the literature as (X) or absent ( ) in healthy bodily fluids. NR-not reported. (+) -positive biomarker. (-) – negative biomarker.

Pathogen	VOCs	PubChem ID	Healthy body						
			Faeces	Urine	Breath	Skin	Milk	Blood	Saliva
AF	1-methylethenyl-pyrazine (+)	62897	NR	NR	NR	NR	NR	NR	NR
	2-pentylfuran (+)	19602	x	x	x		x		x
	cyclohexanone (+)	7967		x	x			x	
	pentanal (+)	8063	x	x	x		x		x
EC	indole (+)	798	x	x		x			x
HP	isobutane (+)	6360			x				
	hydrogen nitrate (+)	944	NR	NR	NR	NR	NR	NR	NR
KP	1-hydroxy-2-butanone (+)	521300	NR	NR	NR	NR	NR	NR	NR
	3-methylcyclohexene (+)	11573	NR	NR	NR	NR	NR	NR	NR
MT	1,3,5-trimethylbenzene (+)	7947			x				
	cymol (+)	7463		x	x	x			x
	2-phenylanisol (+)	6835	NR	NR	NR	NR	NR	NR	NR
PM	(E)-2-butene (-)	62695			x				
PA	ethanol (+)	702	x	x	x			x	x
	1-decanol (-)	8174	x			x			
	3-methylbutanal (+)	11552	x	x	x				x
	2-phenylethanol (+)	6054	x			x	x		x
	1,3,5-trimethylbenzene (-)	7947			x				
	cymol (-)	7463		x	x	x			x
	hydrogen cyanide (+)	768			x				
	2-pentene (+)	12585			x				
	ethane (+)	6324			x				
SA	3-methylbutanoic acid (+)	10430	x		x	x			
	1,1,2,2- tetrachloroethane (+)	6591		x				x	
	2,3,4,5- tetrahydropyridazine (+)	not available	NR	NR	NR	NR	NR	NR	NR
	2-methylbutanoate (+)	22253297	NR	NR	NR	NR	NR	NR	NR
	pyrimidine (+)	9260				x			

Other authors also focused on the determination of VOC biomarkers [24][134][135]. The published databases acquired data through extensive literature searches. Our results are compared with the VOC fingerprints suggested in those studies in Table 3.9. However, different methods were employed to determine the VOC biomarkers. In our study, as described above, we used a verification method to obtain a set of VOCs that allow a separation between a specific pathogen and the other 7 studied pathogens, based on the information collected from published research. This method selected both negative and positive biomarkers (identified as – or + in Table 3.9). In the mVOC database, available in http://bioinformatics.charite.de/mvoc/ [135], a similarity determination between the compound of interest and the compounds of the mVOC database, is used based on the Tanimoto coefficient.

In the Bos et al. review [134] the compounds referred in the selected articles were organized in tables per functional group. Then, all the compounds produced by at least one of the bacteria were included in an interaction graph that connected the compounds with all bacteria known to produce them. This step allowed a visual representation of VOCs that were produced by only one pathogen. Those volatiles were suggested as possible biomarkers for the pathogen. In the Sethi review [24], the VOC biomarkers/profiles referred in each article were summarized. In both studies, there was no further processing of the information contained in those articles.

Although the methods to obtain the VOC fingerprints were distinct between the compared databases, some of the compounds identified as putative biomarkers for a certain pathogen were identified in more than one database. For instance, hydrogen cyanide was identified in 3 out of 4 databases as *P.aeruginosa* biomarker, while 2-pentylfuran, indole, isobutane, cymol, methyl thiocyanate and 3-methylbutanoic acid were identified in 2 of the compared databases as *A. Fumigatus, E.coli, H.pylori, M.tuberculosis, P.aeruginosa and S.aureus* biomarkers, respectively (Table 3.9). The consistent finding of the same putative biomarkers in distinct databases using different data pre-processing and processing methods empowers the possibility of these compounds actually being biomarkers of the pathogens in question.

Table 3.9- Comparison between the set of VOCs obtained in our study, the mVOC database and two existing reviews, where "+" means positive biomarker and "-" negative biomarker.

Pathogen		VOC fingerprint			
		this work	**mVOC database** [135]	**Bos et al. review** [134]	**Sethi et al. review** [24]
Fungus	***A. fumigatus***	1-methylethenyl-pyrazine + 2-pentylfuran + cyclohexanone + pentanal + - -	- - - - 2,4-pentadione 3-methyl-1,3-pentadione	- - - - - -	- 2-pentylfuran - - - -
Gram⁻ bacterium	***E.coli***	indole + - - - -	- - - - pentyl cyclopropane	indole methanol 1-pentanol ethyl acetate -	- - - - -
	H.pylori	isobutane + hydrogen nitrate + - -	- - - -	- - - -	isobutane - 2-butanone ethyl acetate
	P. aeruginosa	ethanol + 1-decanol - 3-methylbutanal + 2-phenylethanol + 1,3,5-trimethylbenzene - cymol - hydrogen cyanide + 2-pentene + ethane + - - - - - -	- - - - - - - - - - - - - - -	- - - - - - hydrogen cyanide - - 1-undecene 2-butanone 2,4-dimethylheptane methyl thiocyanate 4-methyl-quinazoline -	- - - - - - hydrogen cyanide - - - - - methyl thiocyanate - 2-aminoacetophenone

Pathogen		VOC fingerprint				
		this work	**mVOC database** [135]	**Bos et al. review** [134]	**Sethi et al. review** [24]	
	K.pneumoniae	1-hydroxy-2-butanone + 3-methylcyclohexene +	- -	- -	- -	
	P.mirabilis	(E)-2-butene -	-	-	-	
Gram⁺ bacterium	***M. tuberculosis***	1,3,5-trimethylbenzene + cymol + 2-phenylanisol + - - - - - - - -	- - - - - - - - methyl nicotinate methyl phenylacetate -	- - - - - - - - - - -	- - - 1-methyl-naphthalene 1,4-dimethylcyclohexane - - - - - -	- cymol - - - o-xylene isopropyl acetate 3-pentanol - - dimethylstyrene
	S. aureus	3-methylbutanoic acid + 1,1,2,2-tetrachloroethane + 2,3,4,5-tetrahydropyridazine + 2-methylbutanoate + pyrimidine +	- - - - -	3-methylbutanoic acid - - - -	- - - - -	

3.4 Conclusions

Advances in analytical technologies for detecting and measuring VOCs in clinical matrices have generated an increasing interest in the exploitation of VOCs as biomarkers of different diseases. In this work we were able to conclude that the volatolome of *Pseudomonas aeruginosa* is the most studied, which is mostly associated with urinary tract infections and pulmonary infections, such as pneumonia and cystic fibrosis. The collected articles reported mostly reporter-labeled strains. However, when they used clinical isolates, the most used human sample was urine, probably due to the existing standarized procedures for handling and analysis. The GC-MS analysis was the gold standard method to identify the compounds, followed by the SIFT-MS. After grouping the reported VOCs according to their chemical class, it was found that the alcohols class was the most abundant one, while halogen containing compounds were the less abundant. Our systematic literature search resulted in 44 articles, published between 1977 and 2016, reporting 23 distinct pathogens and 418 associated VOCs. We have employed a machine learning method to classify pathogens with more than 3 associated experiments based on the emitted VOCs. VOC-pathogen interaction data was used to build an input data matrix with 8 clinically relevant pathogens (*Pseudomonas aeruginosa, Staphylococcus aureus, Aspergillus fumigatus, Escherichia coli, Helicobacter pylori, Proteus mirabilis, Klebsiella pneumoniae* and *Mycobacterium tuberculosis*) and 269 VOCs. The classifier was able to distinguish the 8 pathogens based on 26 of the 269 VOCs. That set of VOCs was compared with the reported VOCs emitted from a healthy human body [133] and it was found that 3 VOCs (2-phenylanisol, 1-hydroxy-2-butanone and hydrogen cyanide) are not reported, so far, in the healthy human body. We have identified a minimal set of VOCs that allowed the separation of a specific pathogen from the others and compared those VOC lists with the ones found in other studies, for the same pathogens. It was found that indole for *E.coli*; 2-pentylfuran for *A.fumigatus*; isobutane for *H.pylori*; cymol for *M.tuberculosis*; hydrogen cyanide and methyl thiocyanate for *P.aeruginosa*; and 3-methylbutanoic acid for *S.aureus*, were referred in ours and other of the compared databases that used distinct data processing methods. Therefore, these compounds have strong probability of being biomarkers. Nonetheless, more work is required to define the range of normality/disease state in VOCs from humans in terms of concentration ranges in all bodily fluids. This data could then be used to interpretate the constitution of each collected sample obtained from patients, and to monitor their health state or infer about possible pathogen invasions.

4. Tailoring biogels selectivity

4.1 Introduction

Over the last years, GC-MS and associated methodologies became the most widely used analytical gas detectors for clinical diagnosis. However, these instruments are large, expensive and require trained operators, which represent significant limitations to their massive application for diagnosis purposes [26]. E-noses have become target of research, because of their potential ability of being non-invasive, simple and fast tools for detecting VOCs from the human body [84]. An e-nose is a device that comprises of an array of chemical sensors with different selectivity, a signal-preprocessing unit and a pattern recognition system [79]. The interaction of VOCs with an array of sensors generates a characteristic fingerprint which can then be recognized by comparing it with previously recorded patterns in the recognition system [78][79][84]. Several diseases have already been detected using electronic noses, such as urinary tract infections [136][137], tuberculosis [117] and ashtma [49]. However, despite the advances in e-nose research areas, sensors selectivity to detect VOCs remains a major challenge.

Liquid crystals (LCs) are intermediate phases between solid and liquid states (mesophase) [138], in which the matter has fluid properties like liquids and anisotropic properties like crystals [139]. Nowadays, LCs are well known mainly due to their application in electronic display devices. Due to their optical properties, materials that present LC phases are also attractive for other applications, such as chemical sensors [89][139][140]. One of their most advantageous characteristics is that LC molecules are able to rotate the polarization of light that passes through them [141]. Nowadays, the detection principle of LC-based chemical sensors relies on the disruption of the orientations of thermotropic LC molecules at LC/solid or LC/aqueous interfaces upon interaction with analytes [142]. This orientational change of the LC molecules can be observed in a microscope using crossed polarizers due to the LCs birefringence [142]. LCs are promising materials for VOCs detection due to high sensitivity to changes in their molecular ordering under external influences. Several LC based sensors have been sucessfully tested. For example, Ding *et al.* [142] tested a LC based optical sensor to detect butylamine in the air, Winterbottom and collegues [140] have also used cholesteric LCs to detect ethanol,water vapour, and vaporous analytes such as amines in air and Sen *et al*. [143] have reported a LC based sensor to selectively detect nitrogen dioxide.

Ionic liquids (ILs) are organic salts with a stable liquid phase over a wide temperature range around room temperature, whose cations and anions can be varied at will to change their

chemical and physical properties. ILs have drawn attention to polymer chemistry, as well as extraction processes of volatile organic solvents worldwide, since they offer a potentially clean method to carry out chemical processes [144] [145]. ILs can act as surfactants [146] [147]. By taking advantage of this property, a proprietary composite material was developed at the Biomolecular Engineering Lab. It is possible to obtain micelles with LC inside and IL at the surface (Figure 4.1). By using a biopolymer and water the micelles are imobilized in the biopolimeric matrix, forming a thin film [148] [149]. When exposed to VOCs, the LC molecules change their organization [150] [151] that can be detected by Polarizing Optical Microscopy (POM) [152].

In this work we used the proprietary composite gel-like material composed by the LC 5CB, the IL [BMIM][DCA] and the biopolymer gelatin. This material is able to form thin films responsive to VOCs. The films, here referred as biogels, produce an optical response in presence of VOC molecules due to the change of conformation of 5CB molecules within 5CB-[BMIM][DCA] micelle structures. The optical response is observable by POM and quantified using an e-nose developed in-house. The e-nose consists of an array of sensing elements (each composed by a LED, a biogel thin film placed between two crossed polarizers and a LDR), and a signal processing module that quantifies optical response of the biogels.

Some VOCs are recognized as disease biomarkers. Therefore, tailoring the selectivity of the e-nose response towards certain VOC biomarkers would benefit its usability in disease detection. Ju *et al.* [153] demonstrated that tailor-made small peptides can be promising specific receptors for VOCs detection. In their study they identified a specific peptide, with the aminoacid sequence DSWAADIP (Figure 4.2 A), that showed selectivity towards benzene over toluene, xylene, hexane, acetone and ethanol. The behavior of this peptide could provide a very useful foundation for qualitative and quantitative sensing of VOCs for future applications, such as non-invasive testing of health conditions or environmental risk monitoring.

In this work, we have accessed the feasibility of adding VOC-selectivity to biogels thin films by incorporating in the standard biogel the benzene-sensitive peptide identified by Ju *et al.* [153] and two modified versions of it, that contained norleucine (P2) or biphenylalanine (P3) added to their C-terminal (Figure 4.2 B and C), to facilitate its entry into the micelles. The biogels doped with these modified peptides were observed by POM, the location of the added peptides was verified by FITC labeling and the respective optical response to several VOCs was tested on the in-house developed e-nose.

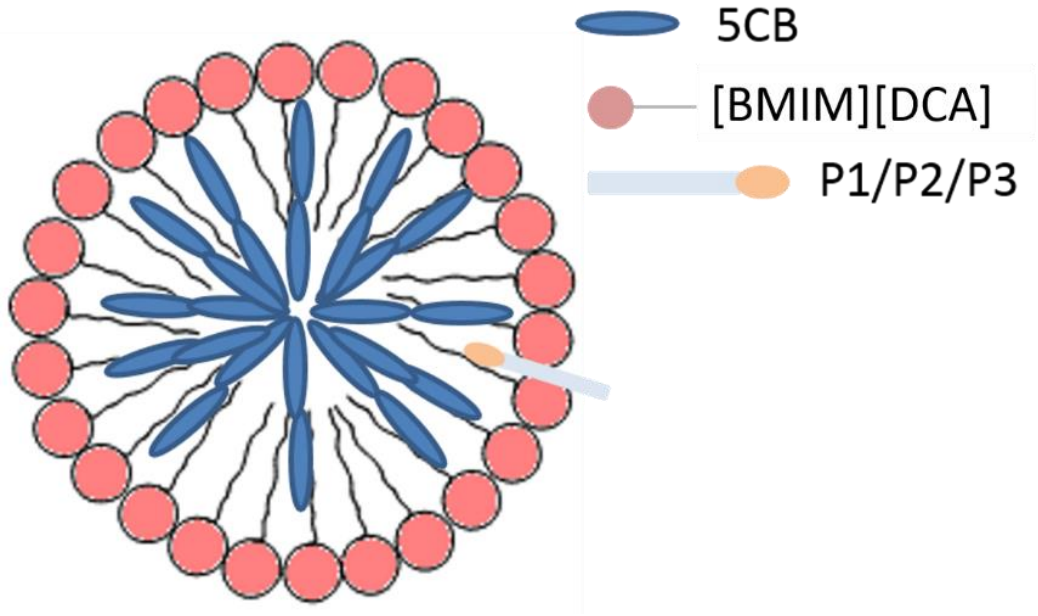

Figure 4.1- Micelles with LC molecules inside, IL at the surface and simulation of the peptides proposed interaction with the micelles.

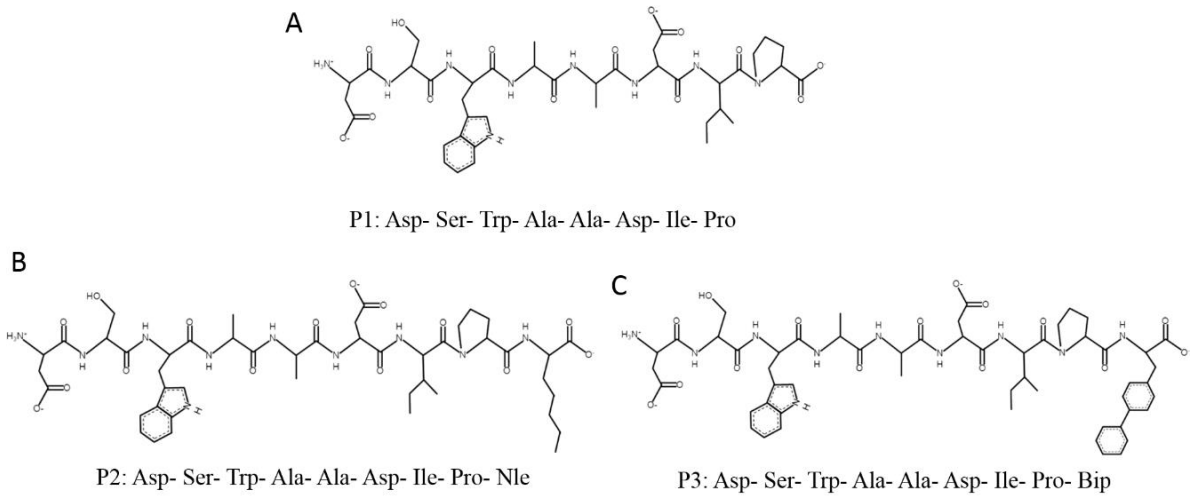

Figure 4.2- Structure of peptides P1 (A), P2 (B) and P3 (C).

4.2 Materials and methods

4.2.1 Materials

Materials: 4-cyano-4-pentylbiphenyl (5CB, >98% purity, purchased from TCI) (Figure 4.3), 1-butyl-3-methylimidazolium dicyanamide ([BMIM][DCA], >98% purity, purchased from io-li-tec) (Figure 4.3), Ammoniumperoxodisulfate (APS, ≥ 98% purity, purshased from Roth), Tetramethylethylenediamine (TEMED, ≥ 98%, purchased from nzytech), acrylamide/bis-acrylamide solution, Fluorescein isothiocyanate (FITC) Isomer I (99% purity, purchased from Sigma) (Figure 4.3), agarose (ultrapure grade, from nzytech), gelatin from bovine skin (purchased from Sigma), acetone (>99.5% purity, purshased from Roth), toluene (>99.5% purity, purchased from Panreac), benzene (>98% purtity, purchased from Quimilabo), xylene (>98% purity, purchased from Riedel de-Haen), hexane (95% purity, purchased from PA Fisher), ethanol (96% purity, purchased from Panreac), peptide P1 (Asp- Ser- Trp- Ala- Ala- Asp- Ile- Pro, 98.9% purity, purchased from GeneCust), peptide P2 (Asp- Ser- Trp- Ala- Ala- Asp- Ile- Pro- Nle, 98.2% purity, purchased from GeneCust), peptide P3 (Asp- Ser- Trp- Ala- Ala- Asp- Ile- Pro- Bip, 99.3% purity, purchased from GeneCust).

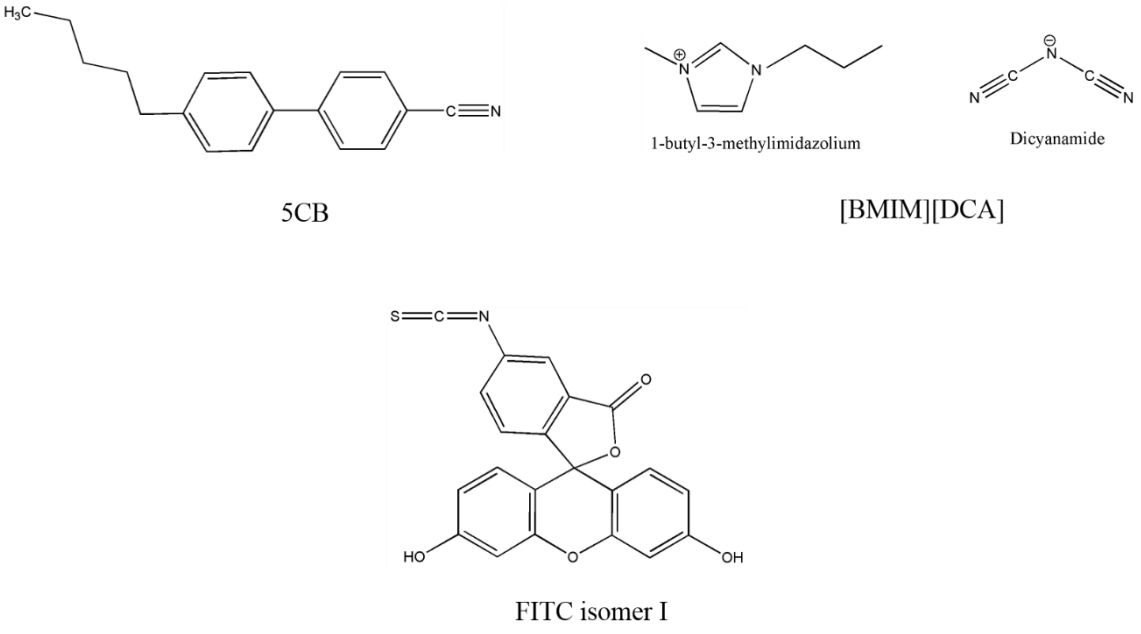

Figure 4.3 - Structures of 4-cyano-4-pentylbiphenyl (5CB), 1-butyl-3-methylimidazolium dicyanamide ([BMIM][DCA]) and Fluorescein isothiocyanate (FITC) Isomer I.

4.2.2 Methods

4.2.2.1 Production of gelatin biogels

For the production of biogels with liquid crystal-ionic liquid micelles immobilized in gelatin a procedure developed at FCT/UNL by the research group of Prof. Dr. Ana Cecília Roque was followed. A glass vial containing a small magnetic stirrer was pre-heated to 37ºC in a stirring hotplate (Echotherm ™ HS40, Torrey Pines Scientific®) 75 µl of [BMIM][DCA] were introduced in the vial and stirred at 340 rpm for 5 min, after which, 5 µl of 5CB were added and kept stirring for 10 min. Then, 25 mg of gelatin were slowly added and kept stirring at 600 rpm for 10 min to obtain a homogeneous mixture. Finally, 25 µl of milliQ water at 37ºC were added and the mixture was kept stirring for 10 min. Negative controls (G_0, G_1, G_2) were also produced (Table 4.1).

A drop (45 µl) of the mixtures was pipetted into warm (37˚C) glass slides with adhesive tape on the sides (Figure 4.4), spread with a glass rod and the film was left to dry.

Table 4.1- Reagents and corresponding quantities used to produce biogels (G_3) and the respective negative controls (G_0, G_1, G_2). G_0- gelatin + milliQ water; G_1- [BMIM][DCA] + gelatin + milliQ water; G_2- 5CB + gelatin + milliQ water; G_3- [BMIM][DCA]+ 5CB + gelatin + milliQ water.

	[BMIM][DCA]	5CB	gelatin	milliQ water
G_0	0 µl	0 µl	25 mg	105 µl
G_1	75 µl	0 µl	25 mg	30 µl
G_2	0 µl	5 µl	25 mg	75 µl + 25 µl
G_3	75 µl	5 µl	25 mg	25 µl

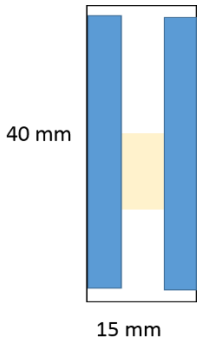

Figure 4.4- Representation of the glass plate, after spreading the gel.

4.2.2.2 Production of agarose biogels

To test agarose as a support matrix to the 5CB/[BMIM[DCA] micelles, a similar protocol was followed. Namely, a hotplate was heated to 45˚C and a small magnetic stirrer was placed in a glass vial. Then, 75 µl of [BMIM][DCA] was introduced in the vial and stirred at 340 rpm for 5 min, 5 µl of 5CB was added to the mixture and was kept stirring for 10 min. After that, 84.5 µl of 5% agarose were added and kept stirring at 600 rpm for 10 min. Negative controls were also produced (A_0, A_1, A_3) (Table 4.2). A drop (45 µl) of the mixtures was pipetted into warm (45˚C) glass slides with adhesive tape on the sides (Figure 4.4), spread with a glass rod and the film was left to dry.

Table 4.2- Reagents and corresponding quantities used to produce agarose biogels (A_3), and the respective negative controls (A_0, A_1, A_2). A_0- agarose + milliQ water; A_1- [BMIM][DCA] + agarose + milliQ water; A_2- 5CB + agarose + milliQ water; A_3- [BMIM][DCA]+ 5CB + agarose + milliQ water.

	[BMIM][DCA]	5CB	agarose	milliQ water
A_0	0 µl	0 µl	84.5 µl	80 µl
A_1	75 µl	0 µl	84.5 µl	5 µl
A_2	0 µl	5 µl	84.5 µl	75 µl
A_3	75 µl	5 µl	84.5 µl	0 µl

4.2.2.3 Production of polyacrylamide biogels

An acrylamide/bisacrylamide solution was prepared by adding 1.67 ml of acrylamide/bisacrylmide solution to 19 µl of PSA and 1.25 µl TEMED. A glass vial containing a small magnetic stirrer was pre-heated to 45ºC in a stirring hotplate (Echotherm ™ HS40, Torrey Pines Scientific®). Then, 75 µl of [BMIM][DCA] was introduced in the vial and stirred at 340 rpm for 5 min, after which, 5 µl of 5CB were added and kept stirring for 10 min. Then, 84.5 µl of the previously prepared polyacrylamide were added and kept stirring at 600 rpm for 10 min. Negative controls were also produced (P_0, P_1, P_2) (Table 4.3). A drop (45 µl) of the mixtures was pipetted into warm (45˚C) glass slides with adhesive tape on the sides (Figure 4.4), spread with a glass rod and the resulting film was left to dry.

Table 4.3- Reagents and corresponding quantities used to produce biogels with polyacrylamide (P_3), and the respective negative controls (P_0, P_1, P_2). P_0- polyacrylamide + milliQ water; P_1- [BMIM][DCA] + polyacrylamide + milliQ water; P_2- 5CB + polyacrylamide + milliQ water; P_3- [BMIM][DCA]+ 5CB + polyacrylamide + milliQ water.

	[BMIM][DCA]	5CB	polyacrylamide	milliQ water
P_0	0 µl	0 µl	84.5 µl	80 µl
P_1	75 µl	0 µl	84.5 µl	5 µl
P_2	0 µl	5 µl	84.5 µl	75 µl
P_3	75 µl	5 µl	84.5 µl	0 µl

4.2.2.4 Coomassie staining of the biogels

The biogels produced using agarose and polyacrylamide as support matrices were stained with coomassie blue R-250 solution. For that, the glass slides were immersed in coomassie blue R-250 solution for 30 min, under gentle agitation. Then, the staining solution was discarded and the biogels were washed several times with ddH_2O. Finally, the glass slides were immersed in destaining solution for 30 min, slowly agitating. After this step the destaining solution was removed and the gels were left to dry in the hotte.

4.2.2.5 Labeling of peptide P1 with FITC

For FITC labeling, 30 mg/ml stock solution of peptide P1 in carbonate-bicarbonate buffer (0.1 M pH = 9.0) was prepared and 0.44 mg of FITC were diluted in 1 ml of DMSO. Then, 300 µl of the FITC solution were added, in aliquots of 10 µl, slowly while agitating, to 100 µl of P1 solution. The final mixture was left to incubate in the dark overnight. P1-FITC conjugate was recovered by HPLC. Since the resulting sample was very diluted (0.80 mg/ml P1 to 0.30 mg/ml FITC) it was then concentrated (3.46 mg/ml P1 to 1.40 mg/ml FITC) by using a rotary evaporator.

4.2.2.6 Determination of P1 and FITC concentrations in the P1-FITC sample

To determine P1 and FITC concentrations in the P1-FITC sample, two aliquots were taken before and after using the rotary evaporator, respectively, and the absorvance values (280 nm) and fluorescence intensity (485-535 nm) were measured in a microplate reader (infinite 200, Tecan i-control). P1 in carbonate-bicarbonate buffer solutions were prepared, in 5 different concentrations (6 mg/ml, 3 mg/ml, 1.5 mg/ml, 0.75 mg/ml and 0.375 mg/ml) and the respective absorvance values were measured at 280 nm (infinite 200, Tecan i-control). FITC in DMSO

solutions were prepared in 6 distinct concentrations (13 mg/ml, 2.6 mg/ml, 1.3 mg/ml, 0.65 mg/ml, 0.325 mg/ml and 0.1625 mg/ml) and the fluorescence intensity was measured at 485-535 nm in a microplate reader (infinite 200, Tecan i-control). Two calibration lines were constructed and the concentrations of P1 and FITC, for the two P1-FITC aliquots, were calculated.

4.2.2.7 Incorporation of P1-FITC in a gelatin biogel

The gelatin biogel was produced following the protocol detailed in 4.2.2.1 with slight modifications: P1-FITC solution (as obtained after the evaporation of excess solvent) was added to the mixture before the addition of the gelatin, and no water was added (Table 4.4).

Table 4.4- Reagents and corresponding quantities used to incorporate the peptides P1 labeled with FITC in the liquid crystal-ionic liquid micelles, and the respective negative control. F_0- [BMIM][DCA] + 5CB + gelatina + milliQ water; F_1- [BMIM][DCA] + 5CB + P1-FITC + gelatin.

	[BMIM][DCA]	5CB	gelatin	P1-FITC	milliQ water
F_0	75 µl	5 µl	25 mg	0 µl	25 µl
F_1	75 µl	5 µl	25 mg	25 µl	0 µl

4.2.2.8 Incorporation of peptides P1, P2 and P3 in gelatin biogels

Different concentrations of P1 in water were tested to optimize the protocol. The most successful one is detailed hereafter. First, 75 µl of [BMIM][DCA] were added to a glass vial containing a small magnetic stirrer and pre-heated to 37˚C for 5 min with stirring at 340 rpm. Then,5 µl of 5CB were added and kept stirring for 10 min. After that, 10 µl of a 10 mg/ml P1 solution or 11.3 µl of a 10 mg/ml P2 solution or 12.5 mg/ml of a 10 mg/ml P3 solution were added to the mixture and kept stirring for 10 min. After that, 25 mg of gelatin were slowly added and kept stirring at 600 rpm for 10 min to obtain an homogeneous mixture. Finally, 25 µl of warm milliQ water were added and the mixture was kept stirring for 10 min. Negative controls (G_0^1, G_1^1, G_2^1, G_0^2, G_1^2, G_2^2, G_0^3, G_1^3, G_2^3) were also produced, for each peptide (Table 4.5).

Table 4.5- Reagents and corresponding quantities used to incorporate the peptides P1 (G_3^1), P2 (G_3^2) and P3 (G_3^3) in the biogels, and to produce the respective negative controls (G_0^1, G_1^1, G_2^1), (G_0^2, G_1^2, G_2^2) and (G_0^3, G_1^3, G_2^3). G_0^1- gelatin + P1 + milliQ water; G_1^1- [BMIM][DCA] + P1 + gelatin + milliQ water; G_2^1- 5CB + P1 + gelatin + milliQ water; G_3^1- [BMIM][DCA]+ 5CB + P1 + gelatin + milliQ water. G_0^2- gelatin + P2 + milliQ water; G_1^2- [BMIM][DCA] + P2 + gelatin + milliQ water; G_2^2- 5CB + P2 + gelatin + milliQ water; G_3^2- [BMIM][DCA]+ 5CB + P2 + gelatin + milliQ water. G_0^3- gelatin + P3 + milliQ water; G_1^3- [BMIM][DCA] + P3 + gelatin + milliQ water; G_2^3- 5CB + P3 + gelatin + milliQ water; G_3^3- [BMIM][DCA]+ 5CB + P3 + gelatin + milliQ water.

		[BMIM][DCA]	5CB	gelatin	Peptide	milliQ water
Biogels with P1	**G_0^1**	0 µl	0 µl	25 mg	10 µl	75+15 µl
	G_1^1	75 µl	0 µl	25 mg	10 µl	15 µl
	G_2^1	0 µl	5 µl	25 mg	10 µl	75+15 µl
	G_3^1	75 µl	5 µl	25 mg	10 µl	15 µl
Biogels with P2	**G_0^2**	0 µl	0 µl	25 mg	11.3 µl	75+13.7 µl
	G_1^2	75 µl	0 µl	25 mg	11.3 µl	13.7 µl
	G_2^2	0 µl	5 µl	25 mg	11.3 µl	75+13.7 µl
	G_3^2	75 µl	5 µl	25 mg	11.3 µl	13.7 µl
Biogels with P3	**G_0^3**	0 µl	0 µl	25 mg	12.5 µl	75+12.5 µl
	G_1^3	75 µl	0 µl	25 mg	12.5 µl	12.5 µl
	G_2^3	0 µl	5 µl	25 mg	12.5 µl	75+12.5 µl
	G_3^3	75 µl	5 µl	25 mg	12.5 µl	12.5 µl

4.2.2.9 Evaluation of the optical response of gelatin biogels in the presence of different VOCs

The different gelatin biogel films ([BMIM][DCA] + 5CB + milliQ water + gelatin, [BMIM][DCA] + 5CB + P1+ milliQ water + gelatin, [BMIM][DCA] + 5CB + P2 + milliQ water + gelatin, [BMIM][DCA] + 5CB + P3 + milliQ water + gelatin, and the correspondent controls) were positioned in the sensor array of the in-house developed e-nose (Figure 4.5) and tested regarding their responses to acetone, hexane, benzene, xylene, toluene and ethanol vapours. For all the solvents the biogels were subjected to 5 cycles of exposure/recovery (60 s of exposure followed by 100 s of recovery, totalizing 15 min per test). The solvents were kept in a bath thermostatized at 37ºC and the respective vapours were pumped to the sensors array chamber (during the exposure time) alterned with clean air (during the recovery time).

The optical signal intensities were calculated by determining the signal amplitude corresponding to each biogel after being exposed to the solvents. The response fold-increase was calculated by the ratio of the biogels with P1, P2 or P3 and the standard biogel signal intensities.

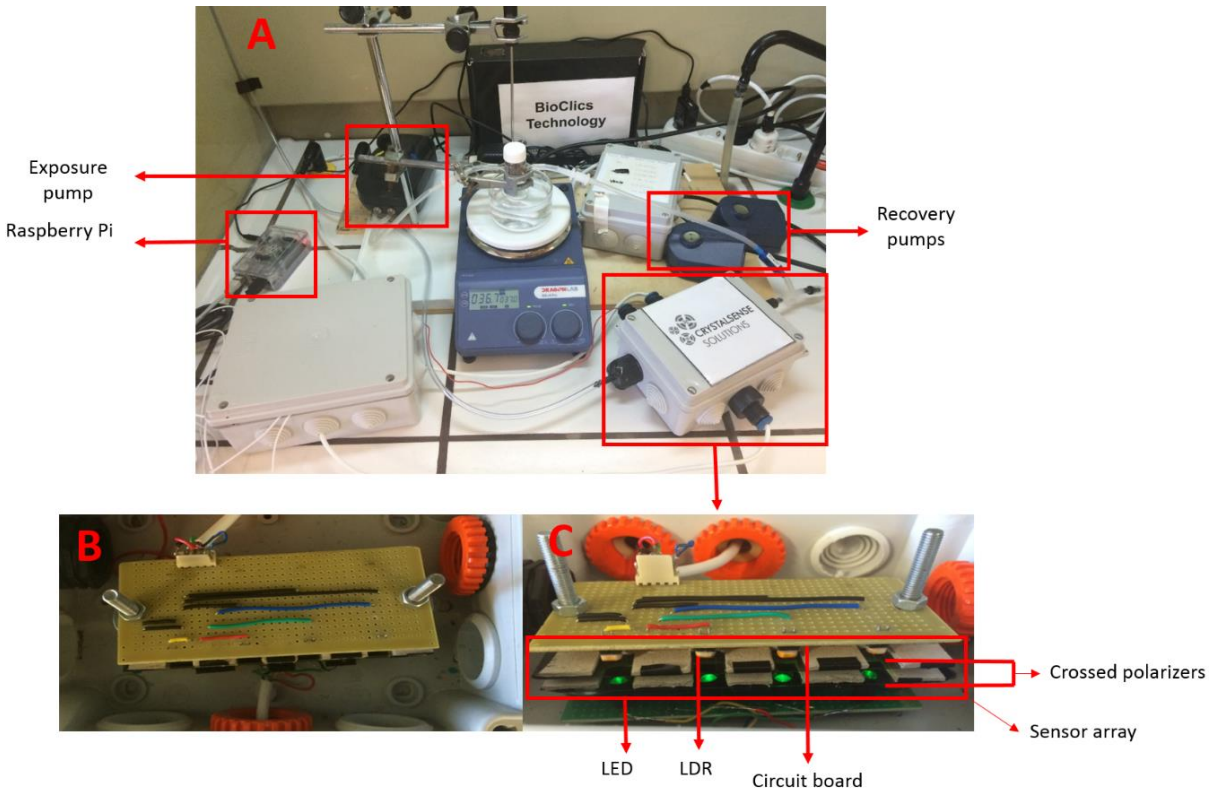

Figure 4.5- In-house developed e-nose assembly, when exposing a biogel to a solvent (A). E-nose sensor array, seen from above (B) and from the side (C). LED- Light Emitting Diode. LDR- Light Dependent Resistor.

4.2.2.10 Optical characterization of the biogels by Microscopy

To characterize the biogels regarding the formation of LC/IL micelles and the respective morphology, the glass slides containing the biogels were observed under polarized light using a ZEISS, Observer.Z1 optical microscope equipped with a ZEISS, Axiocam 503 color camera. To observe the biogels containing P1-FITC, the same microscope was employed, using a source of UV light and a green fluorescent filter. The imaging software for microcopy (ZEN 2.3 (blue edition), ZEISS) was used to process and analyze the images.

4.3 Results and discussion

4.3.1 Doping gelatin biogels with P1, P2 and P3 peptides

As shown by Ju *et al*. the peptide Asp- Ser- Trp- Ala- Ala- Asp- Ile- Pro (P1) has the ability to recognize benzene over other volatile molecules [153]. As a case-study for tailoring the selectivity of the biogels, new peptide-containing biogels were produced by incorporating this peptide or its analogs P2 and P3 in the composition of the biogels. The resulting biogels were then evaluated regarding their optical response towards different volatile solvents to access the effect of adding a selective component in the material.

The proprietary composite gel-like material composed by the LC 5CB, the IL [BMIM][DCA] and te biopolymer gelatin allow the formation of opticaly active micelles with LC inside and IL at the surface, observable by POM (Figure 4.6).

Controls without the addition of the peptides were produced (Figure 4.7 and Appendix 1) and we verified that micelles were only formed in a mixture containing [BMIM][DCA], 5CB and water (Figure 4.8). The IL acts as a surfactant and, since LC molecules are hydrophobic, occurs the self-assembly of 5CB/[BMIM][DCA] in micellar structures with radial configuration [154]. This configuration has one point of defect at the micelle center, causing the micelles to appear as crosses, when observed by POM with crossed-polarizers [154] [155].

We found that by adding P1 to liquid crystal-ionic liquid mixture, micelles were also formed. However, the negative controls with gelatin, P1, milliQ water and gelatin, [BMIM][DCA], P1, milliQ water presented birefrigent structures, which could indicate that the peptide may have LC properties [156] [157] and self-assemble in some structure with birefrigent properties (Figure 4.9).

After finding that the incorporation of the peptide P1 in the biogels was possible and did not affect the formation of micelles (See Appendix 1), we proceeded to the incorporation of two similar peptides, containing an artificial aminoacid in the C-terminal (norleucine in P2 and bifenilalanine in P3) that could act as hydrophobic tails in the peptides termination, to favour their entry into the micelles.

The addition of either P2 or P3 to the liquid crystal-ionic liquid mixture also resulted in the formation of micelles (See Appendix 1). Regarding P2 controls, we verified that the 5CB/P2 gels presented droplets and micelles, and that the [BMIM][DCA]/P2 gels contained rod-shaped birefrigent structures (Figure 4.10), also possibly due to the molecular rigidity [156], while the P3 controls with only [BMIM][DCA], P3 and water, presented micelles (Figure 4.11). P3 contains biphenylalanine which has a similar structure to 5CB (Figure 4.3), which could lead to these peptide behaving similar to 5CB, under these conditions.

Despite the fact that micelles formed in the presence of the three peptides, the biogel that contained the higher number of micelles per biogel area was the one containing [BMIM][DCA], 5CB and P3 with 23497 micelles and a mean micelle area of approximately 312 μm^2 (19.9 μm diameter). However, despite the standard gelatin biogel (without peptide) being the one with the smaller number of micelles formed (992) it was the gel that contained micelles with the highest mean area value (1098 μm^2 and 37.9 μm diameter) (Table 4.6). This finding suggests that the addition of small peptides to the gelatin biogel composition tends to reduce the size and increase the number of micelles.

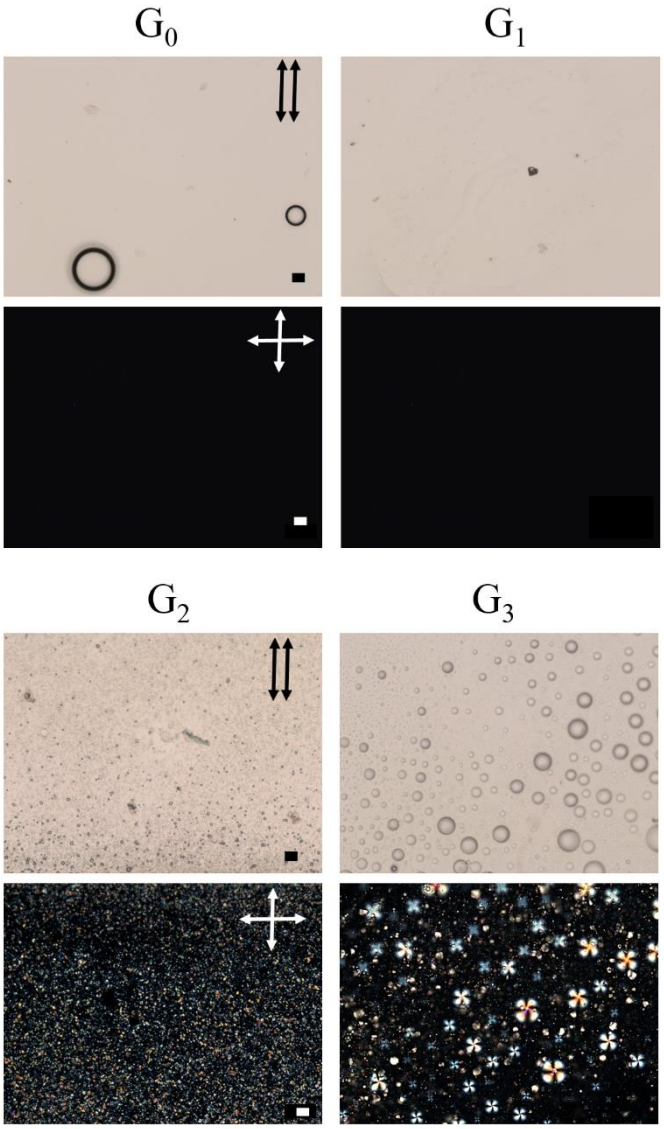

Figure 4.6- POM images of standard biogel (G_3- Gelatin + [BMIM][DCA] + 5CB + milliQ water) and respective controls. G_0- Gelatin + milliQ water. G_1- Gelatin + [BMIM][DCA] + milliQ water. G_2- Gelatin + 5CB + milliQ water. All the images have the same scale. The scale bar corresponds to 50 μm.

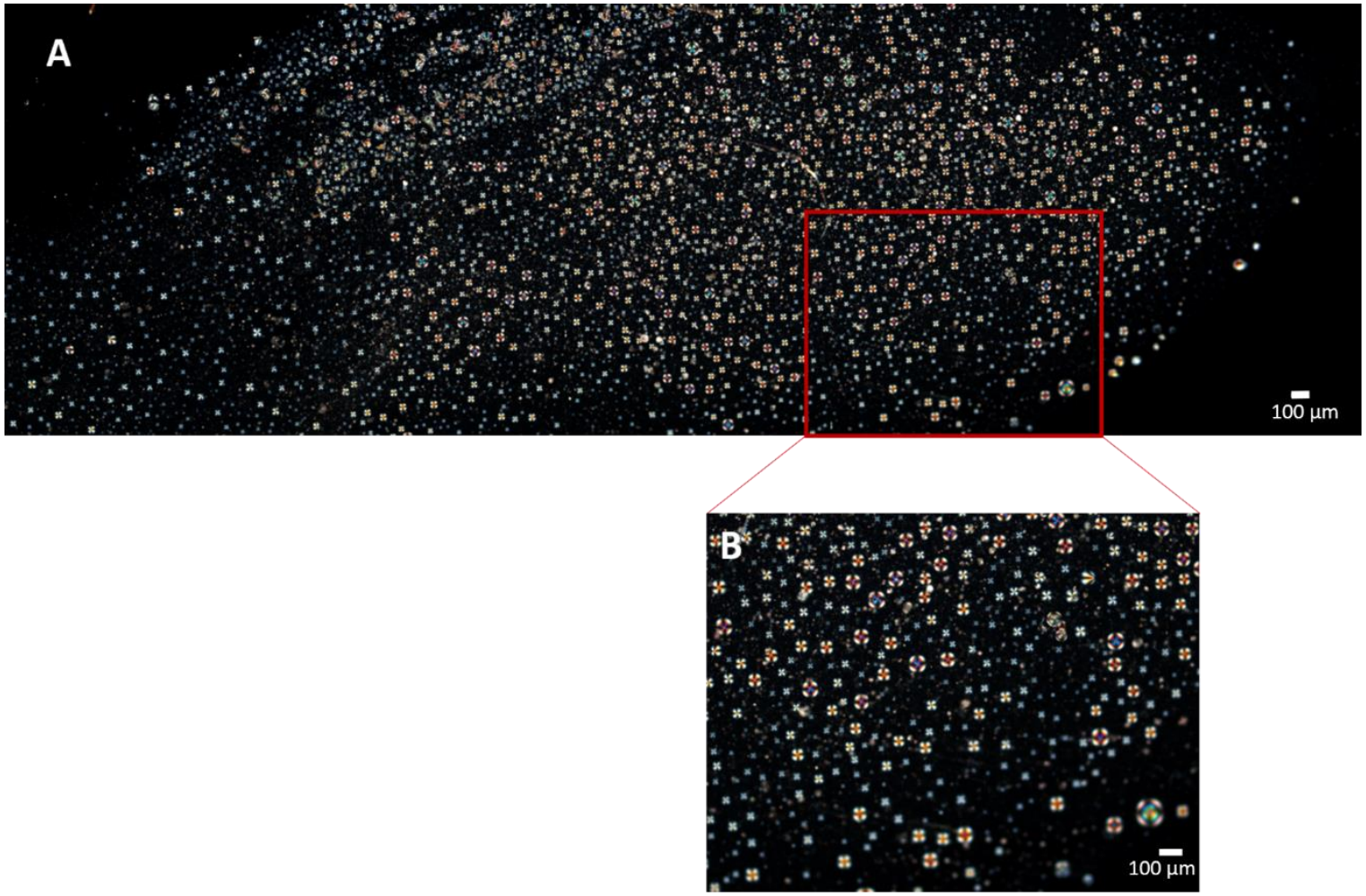

Figure 4.7- Visualization of a [BMIM][DCA]/5CB/gelatin biogel, with crossed polarizers. A: visualization of the whole biogel. B- detail of a biogel region.

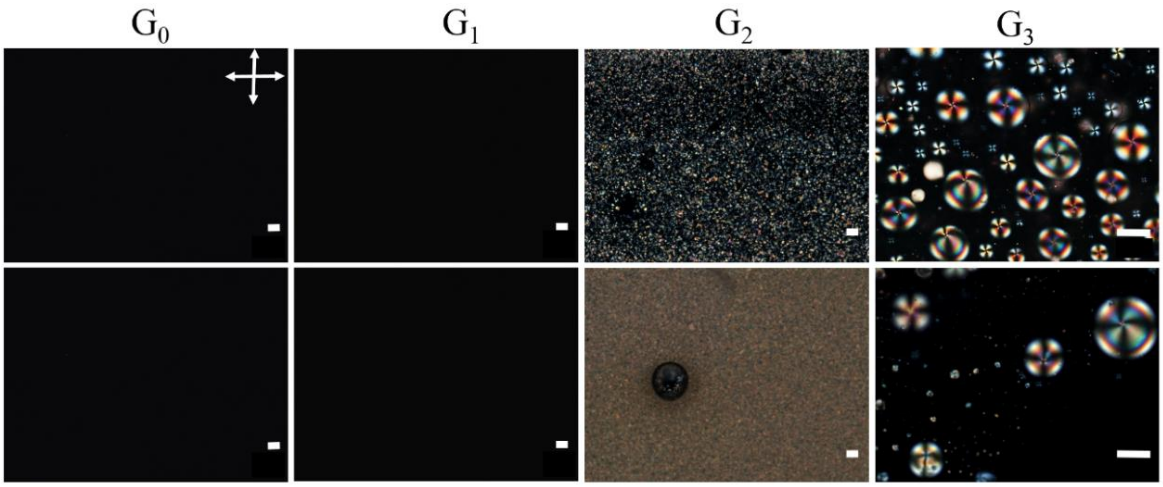

Figure 4.8- Standard gelatin biogel and respective controls, observed by POM, with crossed polarizers, in 2 different fields of view. G_0: Gelatin + milliQ water. G_1: Gelatin + [BMIM][DCA] + milliQ water. G_2: Gelatin + 5CB+ milliQ water. G_3: Gelatin + [BMIM][DCA] + 5CB + milliQ water. The scale bar corresponds to 50 µm.

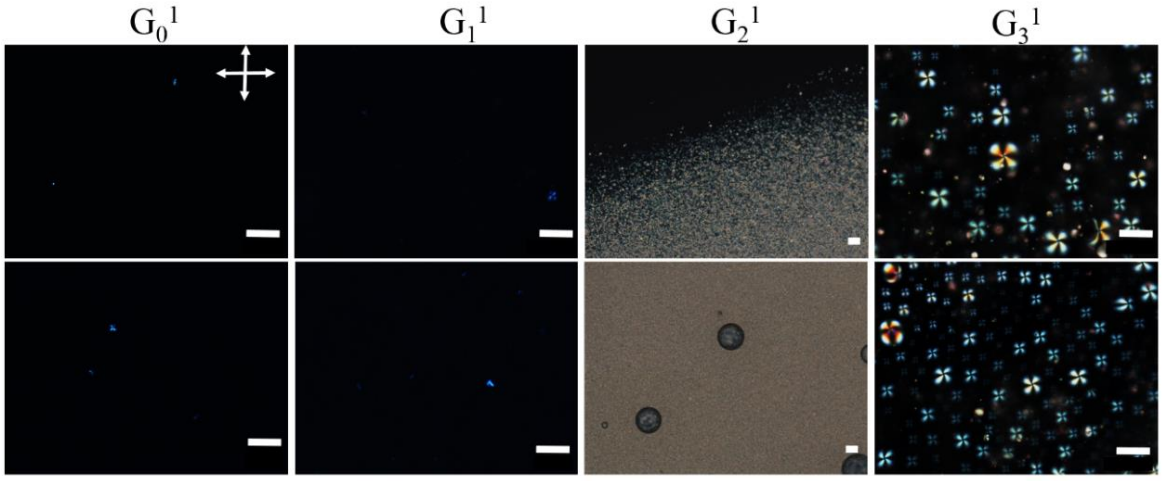

Figure 4.9- Gelatin biogels with P1 added, visualized by POM, with crossed polarizers in 2 different fields of view. G_0^1: Gelatin + P1 + milliQ water. G_1^1: Gelatin + [BMIM][DCA] + P1+ milliQ water. G_2^1: Gelatin + 5CB + P1 + milliQ water. G_3^1: Gelatin + [BMIM][DCA] + 5CB + P1 + milliQ water. The scale bar corresponds to 50 µm.

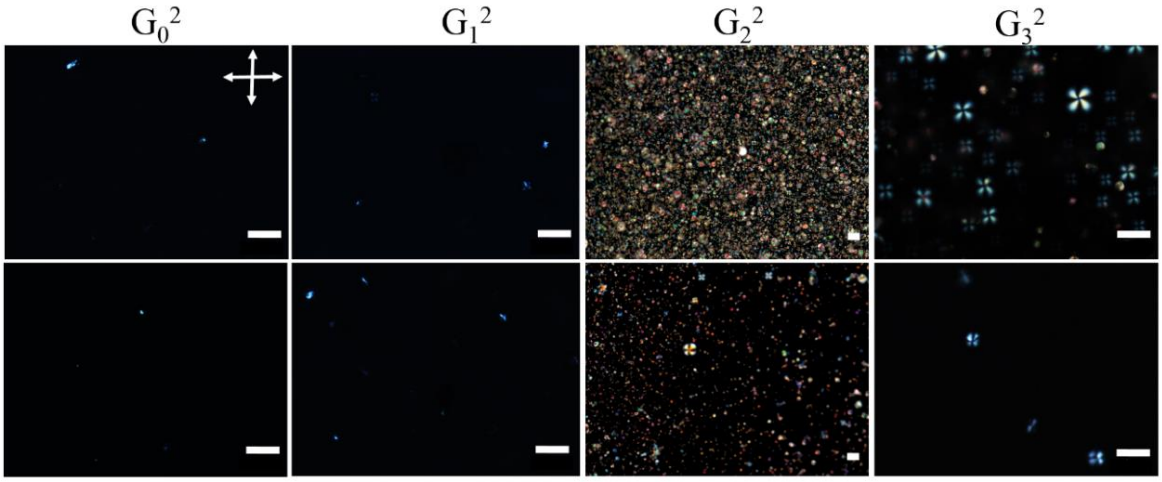

Figure 4.10- Gelatin biogels with P2 added, visualized by POM, with crossed polarizers in 2 different fields of view. G_0^2: Gelatin + P2 + milliQ water. G_1^2: Gelatin + [BMIM][DCA] + P2+ milliQ water. G_2^2: Gelatin + 5CB + P2 + milliQ water. G_3^2: Gelatin + [BMIM][DCA] + 5CB + P2 + milliQ water. The scale bar corresponds to 50 μm.

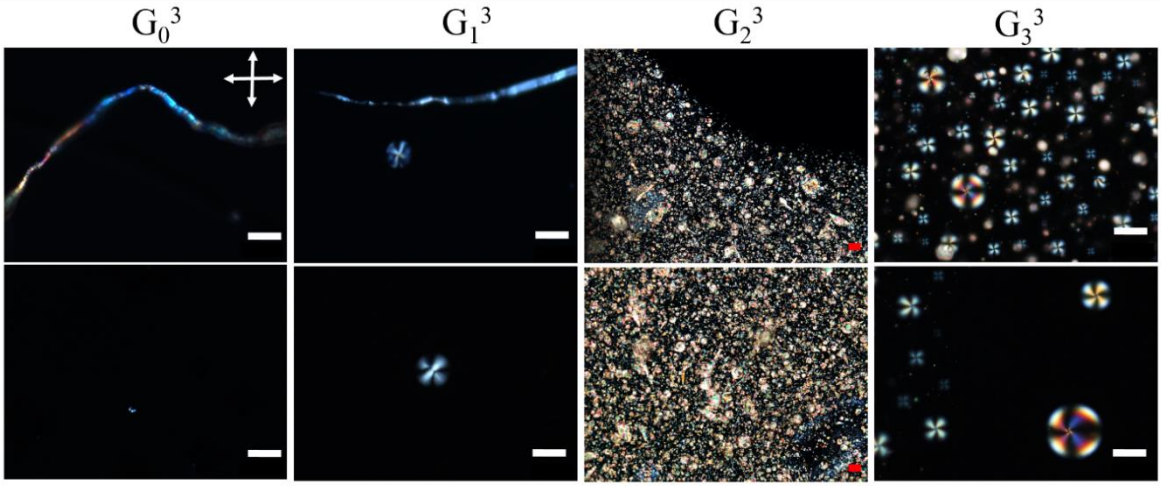

Figure 4.11 - Gelatin biogels with P3 added, visualized by POM, with crossed polarizers in 2 different fields of view. G_0^3: Gelatin + P3 + milliQ water. G_1^3: Gelatin + [BMIM][DCA] + P3+ milliQ water. G_2^3: Gelatin + 5CB + P3 + milliQ water. G_3^3: Gelatin + [BMIM][DCA] + 5CB + P3 + milliQ water. The scale bar corresponds to 50 μm.

Table 4.6 - Biogels with distinct composition and corresponding number of micelles formed and corresponding mean area (μm^2).

Biogel composition	Number of miceles	Mean micelle area (μm^2)	Area with micelles (%)	Biogel area (μm^2)	Number of micelles per biogel area (micelles/μm^2)
[BMIM][DCA] + 5CB + milliQ water + gelatin	992	1098.14	2.6	1089350.06	0.0009
[BMIM][DCA] + 5CB + P1 + milliQ water + gelatin	3905	849.50	4.9	3317295.01	0.0012
[BMIM][DCA] + 5CB + P2 + milliQ water + gelatin	1744	833.22	2.3	1453143.60	0.0012
[BMIM][DCA] + 5CB + P3 + milliQ water + gelatin	23497	312.67	8.1	7346790.87	0.0032

4.3.2 Tracking peptides location within biogels

4.3.2.1 Staining with Coomassie blue

To study the morphology of the new peptide-containing biogels we needed a strategy to verify if those peptides were located in the micelles or dispersed in the biogel matrix. Since the aim was to locate the peptides, by specifically staining them, the gelatin component of the biogels was removed, because it is a proteic component and the coomassie blue would also stain it. The gelatin component was substituted by agarose or polyacrylamide, for further coomassie staining of those biogels.

We verified that the biogels obtained with agarose and polyacrylamide in their composition did not have the same consistency of the ones produced with gelatin (Figure 4.12). The gelatin biogel formed more resistant, peelable films while the agarose and polyacrylamide ones resulted in sticky films. By POM we verified that, when using agarose, the liquid crystal-ionic liquid micelles were sucessfully produced (Figure 4.13 a to h). On the other hand, the polyacrylamide gels did not allow the formation of micellar structures, producing other birefrigent structures instead (Figure 4.14 a to h). Curiously, the negative control composed by 5CB, polyacrylamide

and milliQ water by itself produced liquid crystal droplets and some micelles (Figure 4.14 c and g).

The agarose and the polyacrylamide gels were both stained with Coomassie Blue R-250 and observed by POM. We verified that the staining procedure damaged both biogels and destroyed the micelles in the agarose support (Figures 4.13 and 4.14). In the polyacrylamide gels we did not see any birefrigent structures after the staining (Figure 4.14). Micelles formation was only observed in the agarose biogels, however, those gels were also damaged by the staining procedure. Therefore, peptides incorporation was not carried out in agarose and polyacrylamide biogels.

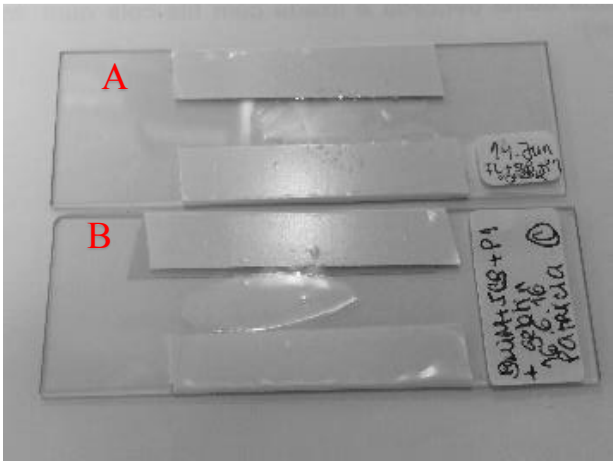

Figure 4.12 - Appearance of biogels produced using agarose (A) and gelatin (B), as the micelles support.

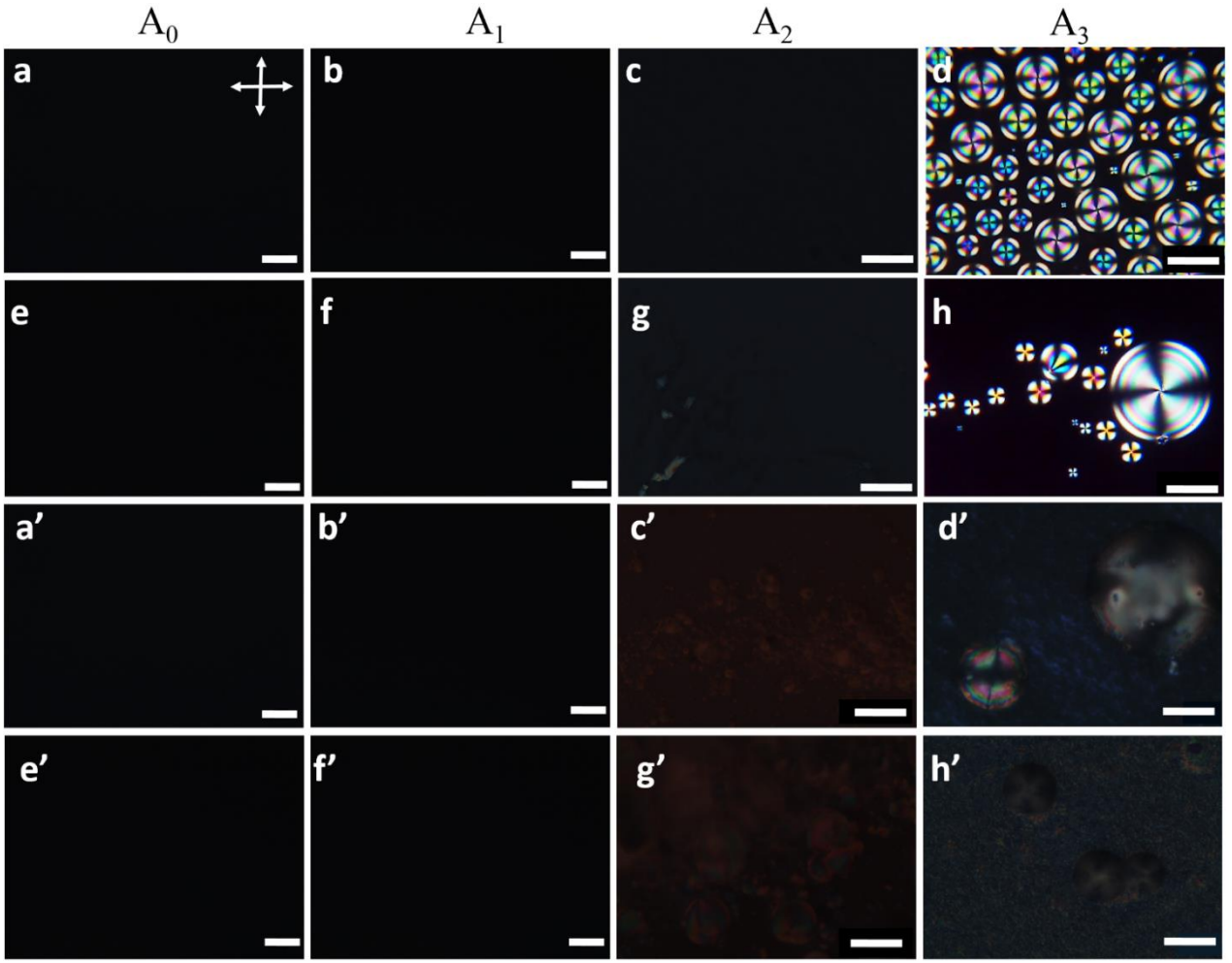

Figure 4.13 - Agarose gels visualization by POM, with crossed polarizers, in 2 different fields of view, before (a to h) and after (a' to h') coomasie staining. A_0- Agarose + milliQ water. A_1- Agarose + [BMIM][DCA] + milliQ water. A_2- Agarose + 5CB + milliQ water. A_3- Agarose + [BMIM][DCA] + 5CB + milliQ water. The scale bar corresponds to 50 μm.

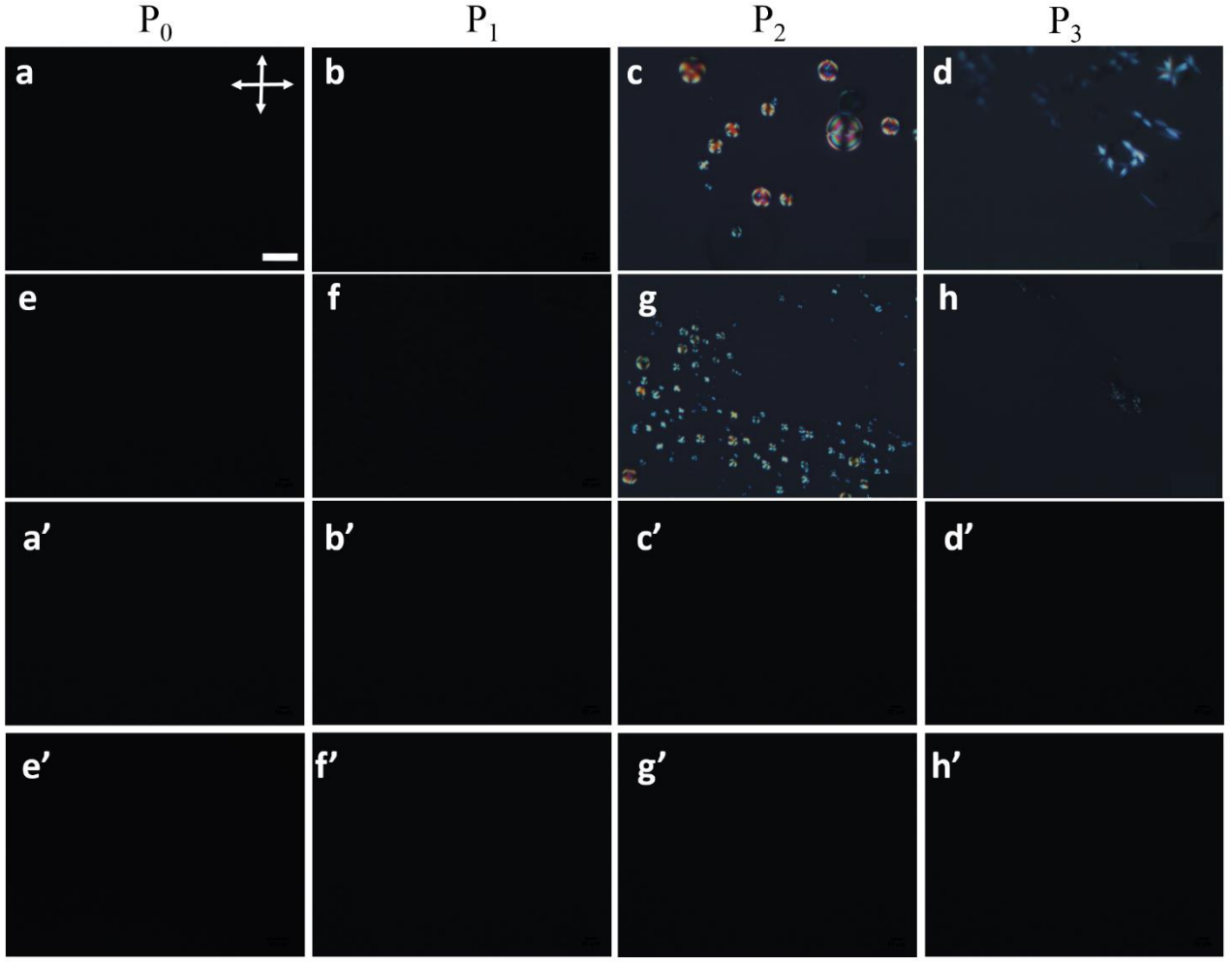

Figure 4.14 - Polyacrylamide gels visualization by POM, with crossed polarizers, in 2 different fields of view, before coomasie staining (a to h) and after (a' to h'). P_0- Polyacrylamide + milliQ water. P_1- Polyacrylamide + [BMIM][DCA] + milliQ water. P_2- Polyacrylamide + 5CB + milliQ water. P_3- Polyacrylamide + [BMIM][DCA] + 5CB + milliQ water. All the images have the same scale. The scale bar corresponds to 50 µm.

4.3.2.2 Tracking peptides by fluorescence

Since the staining with coomasie blue R-250 was destructive for the biogels, a different strategy has to be devised for tracking the peptides in the biogels. We applied the method of labeling the peptide P1 with FITC so that it could be detected by observing the biogel by fluorescence microscopy.

After the P1-FITC conjugation reaction, HPLC was performed in order to identify each component and separate P1-FITC conjugate from unlabeled peptide P1 and eventual remains of FITC present in the reaction mixture (Figure 4.15 A and Appendix 1). After identifying each component, the fraction eluted at 11.3-12.5 min, corresponding to the P1-FITC conjugate was collected (Figure 4.15 B).

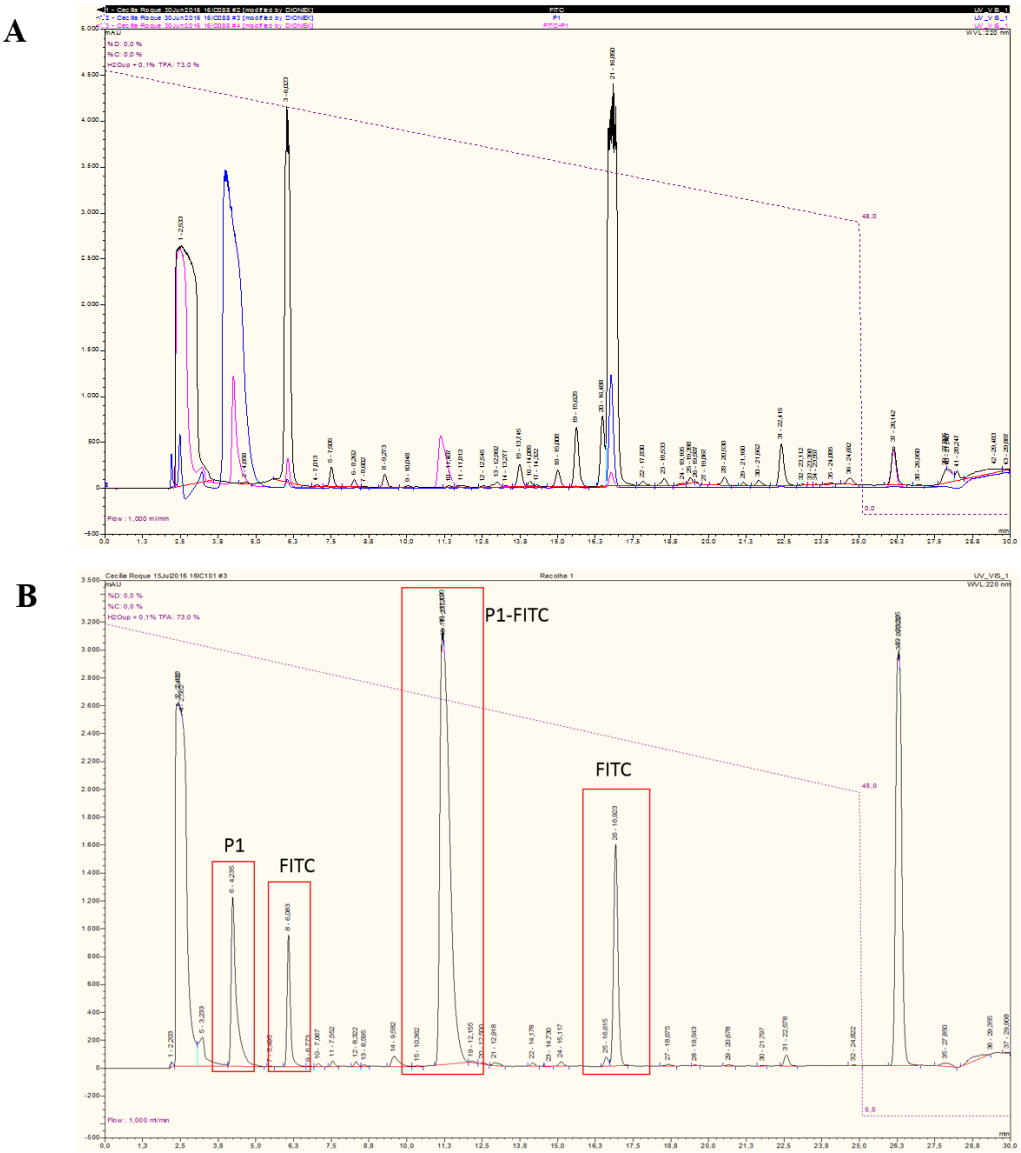

Figure 4.15 - Chromatograms obtained for peptide P1, FITC and P1-FITC conjugate samples, overlayed (A) and identification of each fraction (B)

The collected P1-FITC fraction was very diluted, therefore, the rotary evaporator was used to concentrate the sample. After that, P1 and FITC concentrations in the collected fraction were determined. We verified that with the rotary evaporator we were able to concentrate the sample, approximately 4 times. Also, although the FITC and the P1 were added in a 1:1 ratio (1.14×10^{-7} mol), the final conjugate contained those components in a 1:2.5 ratio, respectively (Table 4.7), which may indicate that some amount of the added FITC did not bind to P1.

Table 4.7 - Values of absorbance (280 nm) and fluorescence (485-535 nm) for P1 and FITC and the respective concentrations of each compound, in the P1-FITC sample before and after using the rotary evaporator.

	Abs $_{280\ nm}$	**Fluorescence** $_{485\text{-}535\ nm}$	**Concentration (mg/ml)**
P1-FITC sample before rotary evaporator	0.129	3562	P1 = 0.80 FITC = 0.30
P1-FITC sample after rotary evaporator	0.346	6976	P1= 3.46 FITC = 1.40

The concentrated P1-FITC sample (3.46 to 1.40 mg/ml) was incorporated in the biogels with gelatin during their production and the resulting films were observed by POM and fluorescence microscopy. POM revealed that in the films containing [BMIM][DCA], 5CB, and gelatin the formation of micelles was not affected by the addition of P1-FITC (Figure 4.16). By observing the biogel with FITC labeled P1 with fluorescence microscopy we concluded that the matrix and the micelles were both green (Figure 4.17, d and h), compared to the control biogel, which did not present any fluorescence (Figure 4.17, c and g), some bubbles also seem to have encapsulated P1-FITC, since they present fluorescence (red rectangle in Figure 4.17). However the inside of the micelles presented a brighter colour than the surrounding matrix, indicating that P1 was successfully incorporated within the micelles and probably a much smaller amount was also dispersed in the gel (red highlights in Figure 4.17). It is assumed that, due to the similarity in the aminoacid sequence, peptides P2 and P3 behave like P1 and will also be partially incorporated into the micelles.

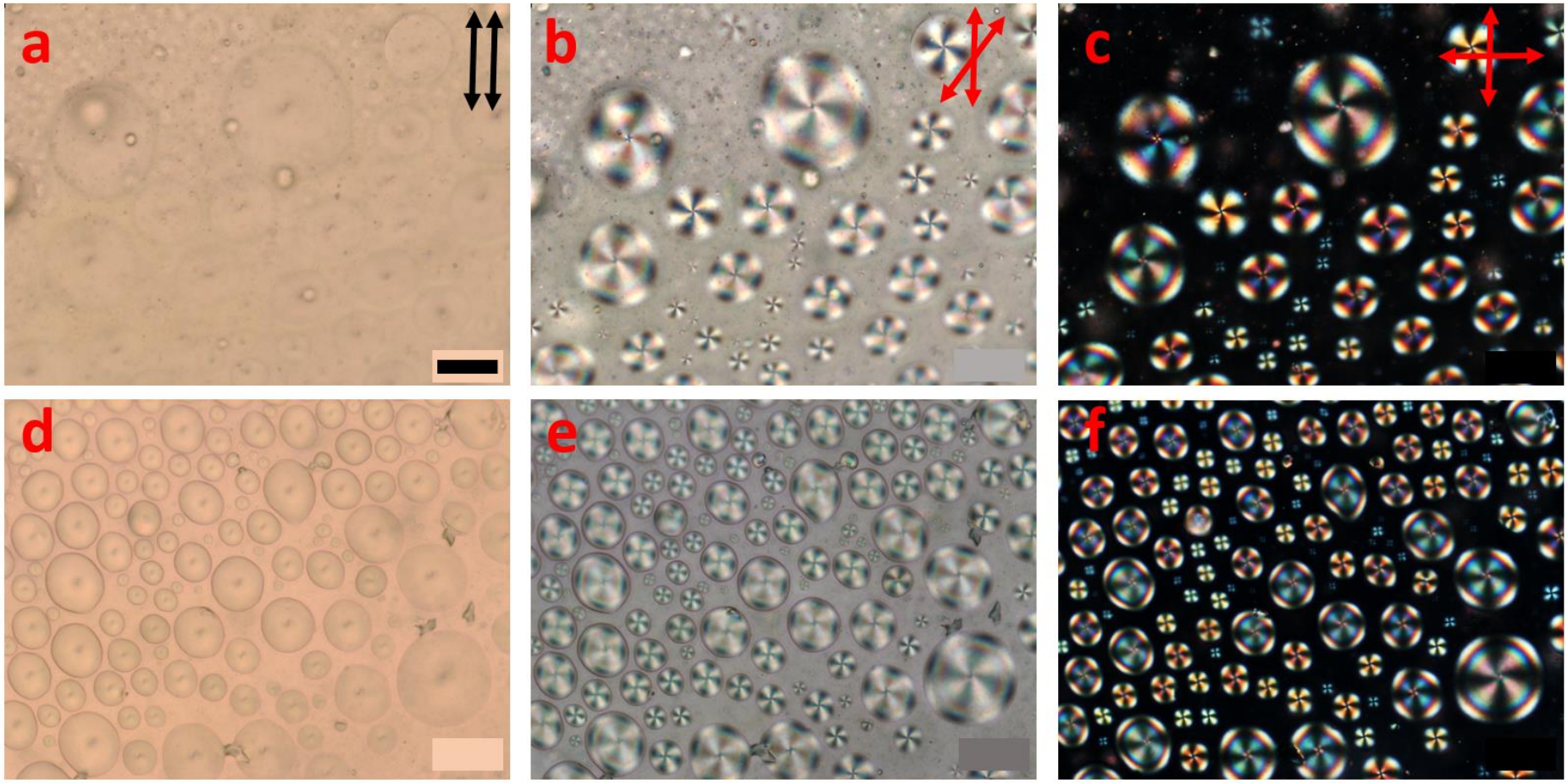

Figure 4.16 - POM images of gelatin biogels without any peptide added (a to c) and P1-FITC (d to f) with uncrossed (a and d), semi-crossed (b and e) and crossed polarizers (c and f). All images have the same scale. The scale bar corresponds to 50 μm.

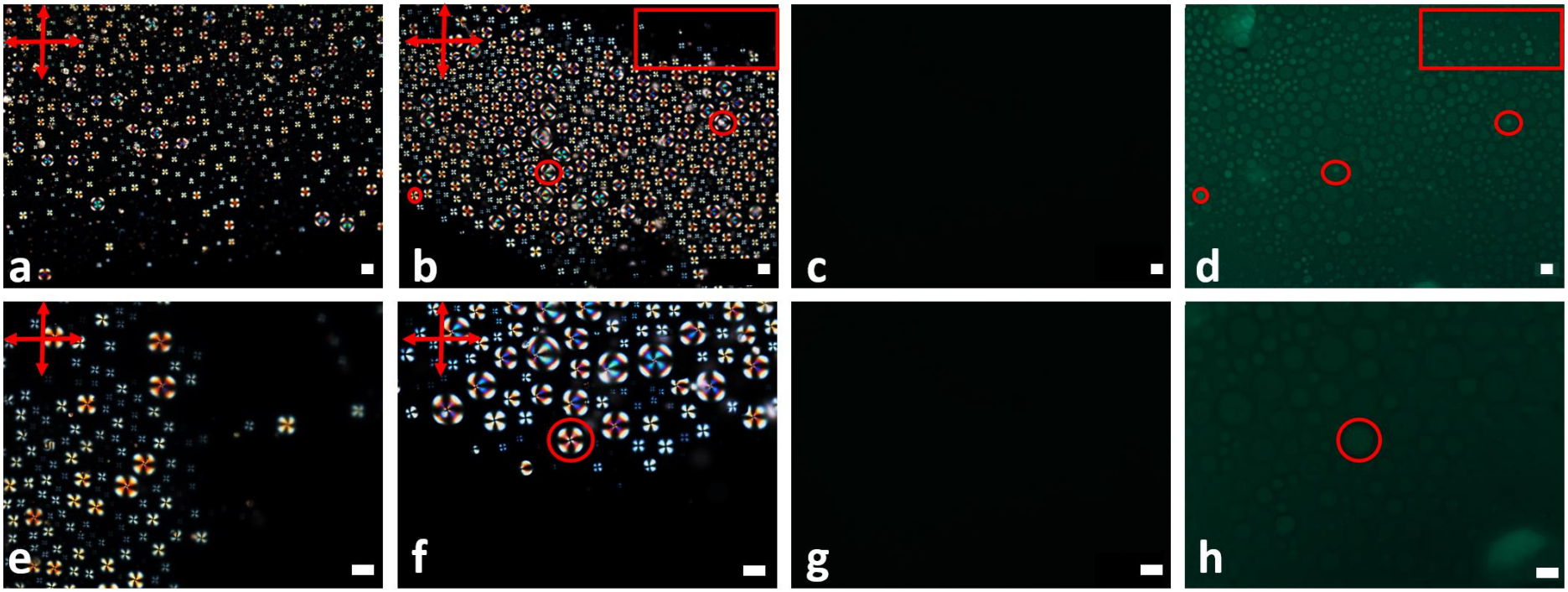

Figure 4.17 - Gelatin biogels with (b and f) and without P1-FITC (a and e) visualized by POM, with crossed polarizers, and by fluorescence microscopy (d, h, and c, g) in 2 different fields of view. a, c, e and g: control biogel composed by gelatin, [BMIM][DCA], 5CB, P1 and milliQ water. b, d, f and h: biogel composed by gelatin, [BMIM][DCA], 5CB, P1-FITC and milliQ water. The scale bar corresponds to 50 μm. The red marks highlight fluorescent bubles (rectangle) and Bright spots inside the micelles (circules).

4.3.3 Effect of P1, P2 and P3 in the response of gelatin biogels to different VOCs

After incorporating the peptides into the biogels, micelles were still produced, therefore, we exposed those materials to a group of solvent vapours in order to evaluate the effect of the peptides in the VOC response of the new materials compared to the standard one. The responses to VOCs were evaluated using a proprietary custom built e-nose. Since we verified that, in some cases, micelles or rod-shaped birefrigent structures were formed without the addition of liquid crystal, we also exposed those control films to the same solvents (See Appendix 1).

The solvents to which the biogels were exposed were the same tested in [153] (acetone, benzene, xylene, toluene, hexane and ethanol). These compounds belong to three of the most abundant chemical classes (ketones, hydrocarbons and alcohols) found in Chapter 3 and ethanol was found to be the most frequently detected VOC in pathogen culture headspaces in that same chapter.

According to Ju *et al.* [153], the peptide P1 presented great selectivity towards benzene, but also presented a minimal response to toluene.

We observed that, although the standard biogel already responded significantly to benzene, compared to other VOCs (Figures 4.18 A and 4.19), the biogel containing P1 presented a much more intense response (Figures 4.18 C and 4.19).

The incorporation of P1 also increased the biogel response to acetone, hexane and toluene, when compared to the material without P1 (Figures 4.18 A and C and Figure 4.19). However, the highest increase occurs for toluene and hexane (Figure 4.20). Biogels containing the modified versions of P1 (P2 and P3) were also tested to see if the addition of norleucine and biphenylalanine, respectively, changed the signal response observed for P1. The incorporation of P2 in the biogels modified the sensors response (Figure 4.19). The sensor responded with a higher signal intensity to benzene, acetone and toluene when P2 was present in the biogel (Figures 4.18 B and 4.19). The greatest response increase ocurred in presence of acetone and toluene (Figure 4.20). Regarding the incorporation of P3 we noted that both acetone and benzene produced a pronounced response of the biogel (Figures 4.18 D) and the higher response increase occurs for acetone and hexane (Figure 4.20). In fact, the biogel containing P3 was the one that responded the most when exposed to benzene and the biogel containing P2 responded the most to acetone, xylene, toluene and (Figure 4.19).

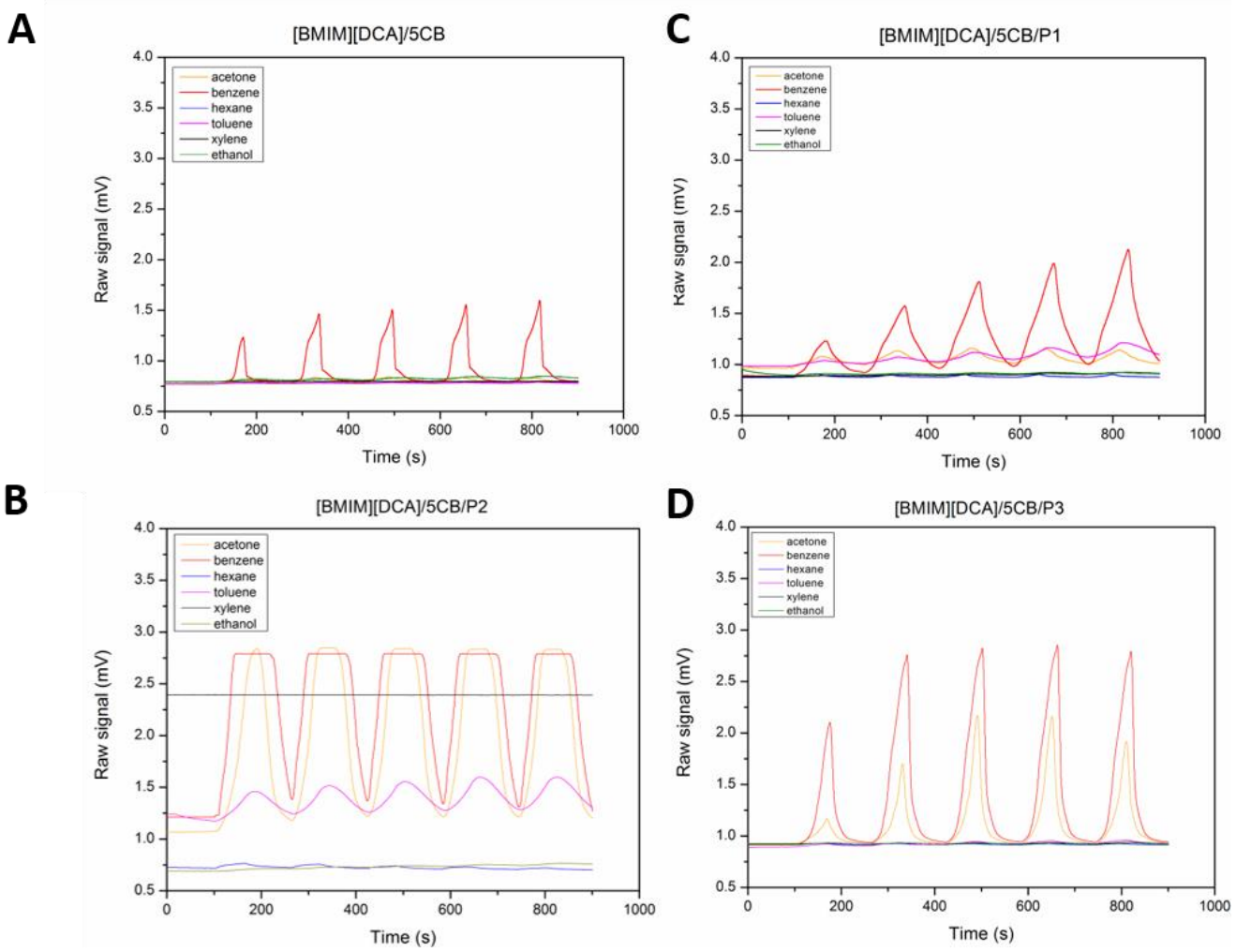

Figure 4.18 -Overlayed signals of the biogels containing [BMIM][DCA]/5CB (A), [BMIM][DCA]/5CB/P1 (C), [BMIM][DCA]/5CB/P2 (B) and [BMIM][DCA]/5CB/P3 (D) when exposed to acetone, benzene, hexane, toluene, xylene and ethanol.

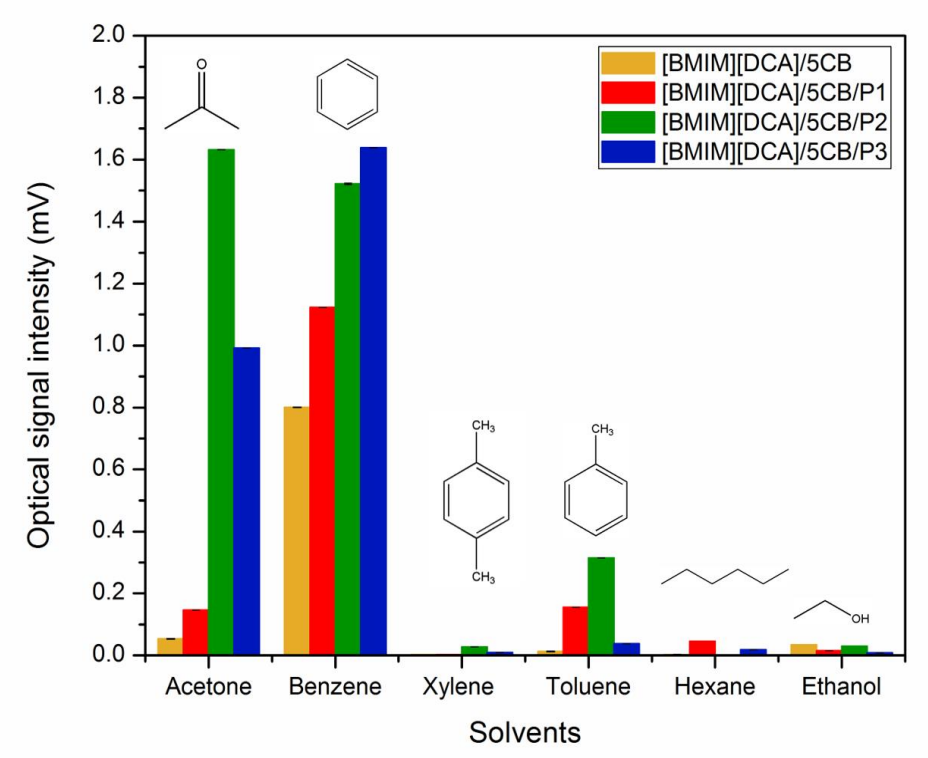

Figure 4.19 - Optical signal intensity of each biogel to the different solvents.

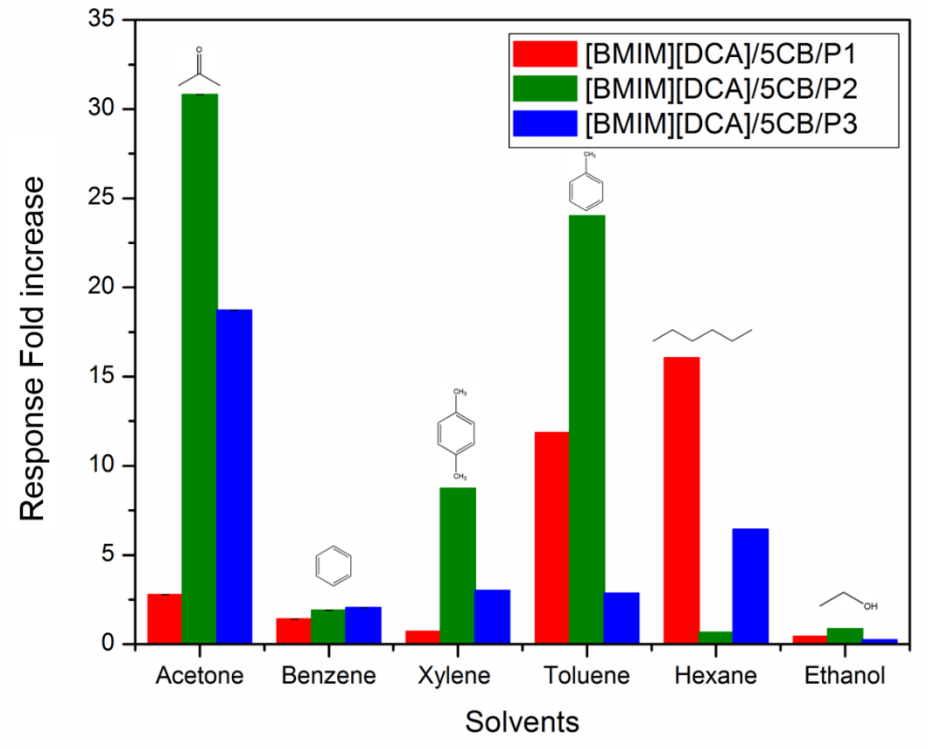

Figure 4.20 - Response fold increase of each biogel, after P1, P2 and P3 incorporation in the standard biogel.

4.3.4 Biogels stability after VOCs exposure

After VOCs exposure the biogels were observed again by POM. Altough some micelles appear to remain stable, all the gels showed that bubbles and some micelles were destroyed (Figures 4.21-4.24). The material containing P3 is the one that remains more stable (Figure 4.24). However, one of the VOCs was an alcohol and the gels stability could be affected after being exposed to it, since it may interfer with the ionic liquid at the micelles surface [157] [156], altering their structure and destabilize the gelatin component [158] [159] .

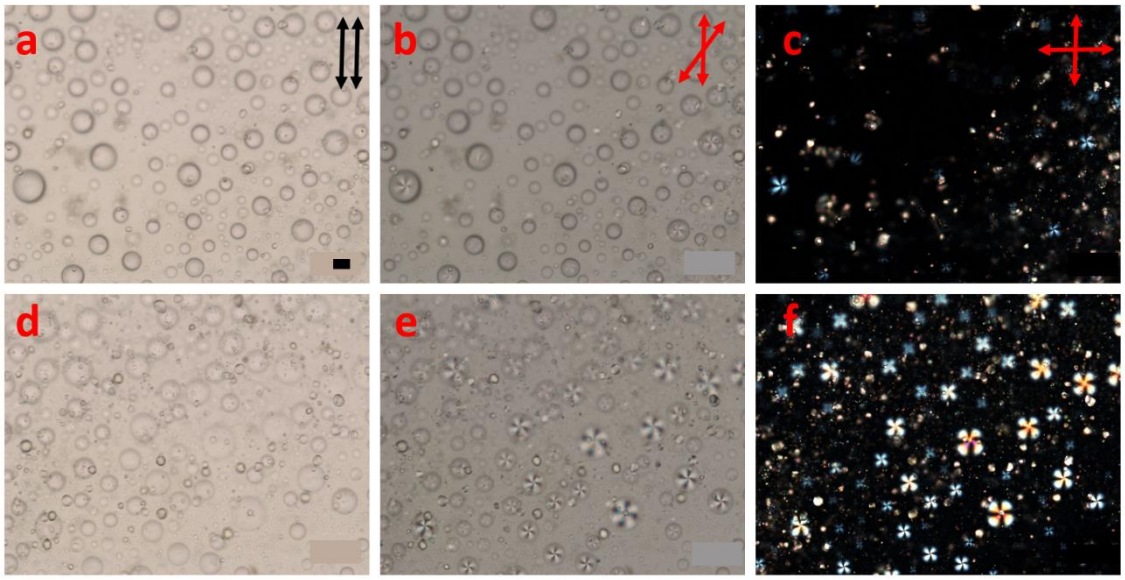

Figure 4.21 - POM images of standard gelatin biogels (gelatin + [BMIM][DCA] + 5CB), after VOCs exposure, in 2 different fields of view, with uncrossed (a and d), semi-crossed (b and e) and crossed (c and f) polarizers. All images have the same scale. The scale bar corresponds to 50 µm.

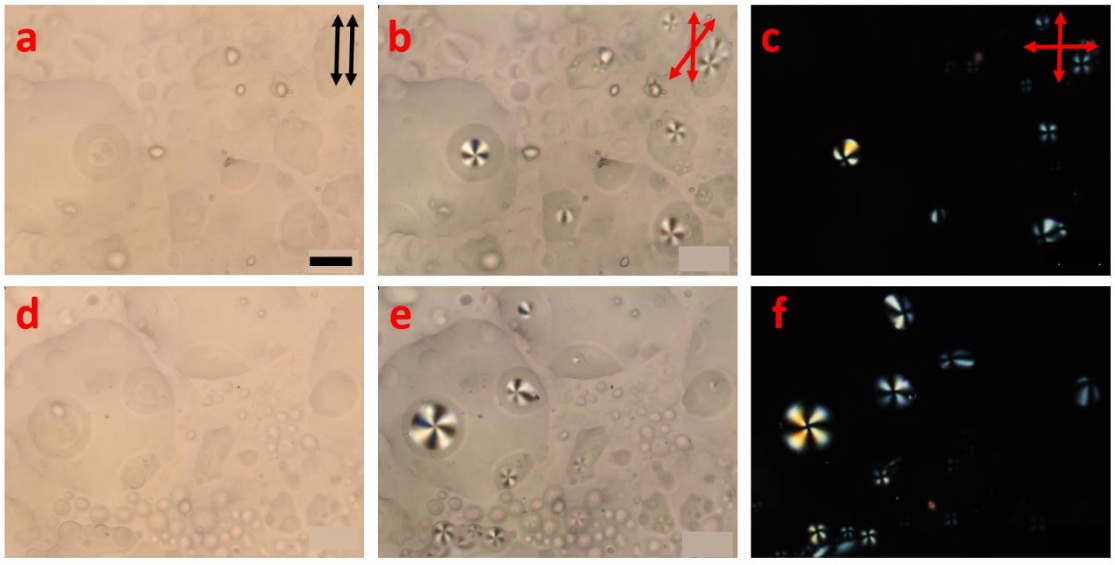

Figure 4.22 - POM images of P1 biogels (gelatin + [BMIM][DCA] + 5CB + P1), after VOCs exposure, in 2 different fields of view, with uncrossed (a and d), semi-crossed (b and e) and crossed (c and f) polarizers. All images have the same scale. The scale bar corresponds to 50 µm.

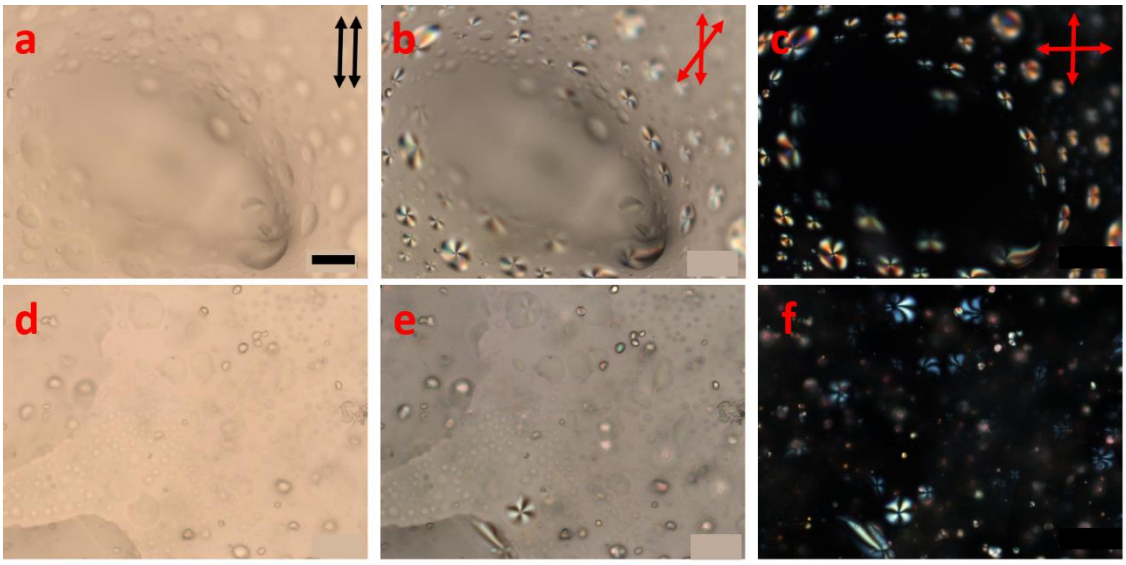

Figure 4.23 - POM images of P2 biogels (gelatin + [BMIM][DCA] + 5CB + P2), after VOCs exposure, in 2 different fields of view, with uncrossed (a and d), semi-crossed (b and e) and crossed (c and f) polarizers. All images have the same scale. The scale bar corresponds to 50 μm.

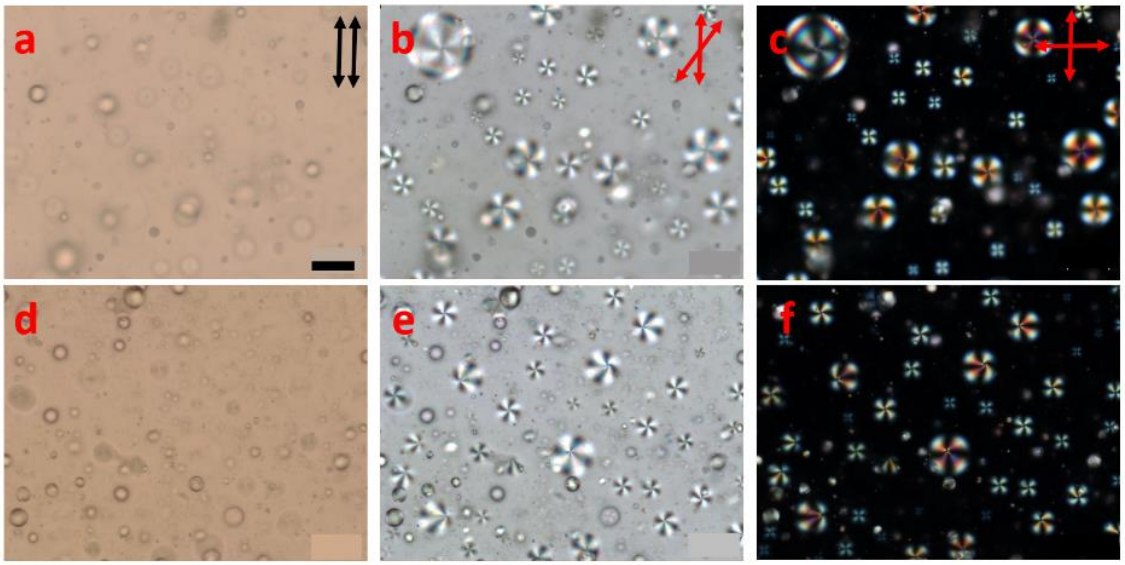

Figure 4.24 - POM images of P3 biogels (gelatin + [BMIM][DCA] + 5CB + P3), after VOCs exposure, in 2 different fields of view, with uncrossed (a and d), semi-crossed (b and e) and crossed (c and f) polarizers. All images have the same scale. The scale bar corresponds to 50 μm.

4.4 Conclusions

Tailoring selectivity of VOC-responding materials was conducted by adding three different peptides (P1, P2 and P3) to a standard biogel. Selectivity was assessed using a proprietary electronic nose based on the optical response of the biogels upon their interaction with VOCs.

In this work we used a biogel, a proprietary composite gel-like sensor material composed by the LC 5CB and the IL [BMIM][DCA] self-assembled in micelle structures immobilized within the biopolymer gelatin. This material is able to form thin films responsive to VOCs. The biogel produce an optical response in presence of VOC molecules. The optical response is observable by POM and quantified using an e-nose developed in-house.

In this work, we have accessed the feasibility of adding VOC-selectivity to biogels thin films by incorporating in the standard biogel the benzene-sensitive peptide identified by Ju *et al.* (P1) and two modified versions of it, that contained norleucine (P2) or biphenylalanine (P3) added to their C-terminal to facilitate self-assembly and LC encapsulation within the micelles.

By labeling P1 with FITC we were able to verify that it was successfully incorporated in the micelles, although some P1 was also dispersed in the gel. P1, P2 and P3 were incorporated in the biogels and micelles were always produced.

We were also able to verify by POM that the peptides self-assemble in some structures with birefrigent properties. P3, for instance, forms micelles with only the addition of [BMIM][DCA] and water. By testing a set of VOCs in a proprietary custom built e-nose we verified that the biogels without any peptide added responded more sharply to benzene and acetone. The addition of P1 seemed to amplify the response to benzene and toluene. The addition of P2 and P3 amplified the response signal to both acetone and benzene. Addition of P2 also increased the response intensity to toluene when compared to the addition of P1 to the standard biogel. Since the standard biogel already responded significantly towards benzene it would be of interest to incorporate P1 in biogels that do not respond when exposed to that solvent and test if a selectivity improvement towards benzene occurs.

5. Concluding remarks and future perspectives

In this work the distinction between 8 clinically relevant pathogens based on the emitted VOCs reported in literature was investigated. Data of interest was collected and machine learning methods were employed to classify the pathogens, based on the emitted VOCs. Data regarding VOC-pathogen interaction found in research articles between 1977 and 2016 was used to build an input data matrix with *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Aspergillus fumigatus*, *Escherichia coli*, *Helicobacter pylori*, *Proteus mirabilis*, *Klebsiella pneumoniae* and *Mycobacterium tuberculosis* and 269 VOCs. That set of VOCs was compared with the reported VOCs emitted from a healthy human body [133] to assess whether those compounds were also present in the healthy body. It was found that 3 VOCs (2-phenylanisol, 1-hydroxy-2-butanone and hydrogen cyanide) are not reported, so far, in the healthy human body. We have identified a minimal set of VOCs that allowed the separation of a specific pathogen from the others and compared those VOC lists with the ones found in other studies, for the same pathogens. It was found that indole for *E.coli*; 2-pentylfuran for *A.fumigatus*; isobutane for *H.pylori*; cymol for *M.tuberculosis*; hydrogen cyanide and methyl thiocyanate for *P.aeruginosa*; and 3-methylbutanoic acid for *S.aureus*, were referred in ours and other databases. Therefore, these compounds have strong probability of being biomarkers. Nonetheless, more work is required to define the range of normality/disease state in VOCs from humans in terms of concentration ranges in all bodily fluids. This data could then be used to interpret the constitution of each collected sample obtained from patients, to monitor their health state or infer about possible pathogen invasions. Also, since there are pathogens much more studied than others, from a statistical point of view, it would be important to balance the classifier input data, choosing, for example, 4 experiments for each pathogen, randomly, and repeat this classification procedure several times to confirm the reliability of the classifier results.

In the second part of the work, the selectivity of VOC-responding materials (biogels) was also tailored by adding three different peptides (P1, P2 and P3) to a standard biogel. Selectivity was assessed using a proprietary electronic nose based on the optical response of the biogels upon their interaction with VOCs.

E-noses allow the development of non-invasive, simple and fast tools for detecting VOCs from the human body [84]. However, despite the advances in e-nose research areas, sensors selectivity to detect VOCs remains a major challenge. In this work we used a biogel, a proprietary composite gel-like sensor material composed by the LC 5CB and the IL [BMIM][DCA] self-assembled in micelle structures immobilized within the biopolymer gelatin. This material is able to form thin films responsive to VOCs. The biogel produce an optical response in presence of VOC molecules. The optical response is observable by POM and

quantified using an e-nose developed in-house. Since some VOCs are recognized as disease biomarkers, tailoring the selectivity of biogel response towards certain VOC biomarkers would benefit its usability in future applications, namely for non-invasive testing of health conditions or environmental risk monitoring. In this work, we have accessed the feasibility of adding VOC-selectivity to biogels thin films by incorporating in the standard biogel the benzene-sensitive peptide identified by Ju *et al.* (P1) and two modified versions of it, that contained norleucine (P2) or biphenylalanine (P3) added to their C-terminal to facilitate self-assembly and LC encapsulation within the micelles. By labeling P1 with FITC we were able to verify that it was successfully incorporated in the micelles, although some P1 was also dispersed in the gel. P1, P2 and P3 were incorporated in the biogels and micelles were always produced. We were also able to verify by POM that the peptides self-assemble in some structures with birefrigent properties. P3, for instance, forms micelles with only the addition of [BMIM][DCA] and water. By testing a set of VOCs in a proprietary custom built e-nose we verified that the biogels without any peptide added responded more sharply to benzene and acetone. The addition of P1 seemed to amplify the response to benzene and toluene. The addition of P2 and P3 amplified the response signal to both acetone and benzene. Since the standard biogel already responded significantly towards benzene it would be of interest to incorporate P1 in biogels that do not respond when exposed to that solvent and test if a selectivity improvement towards benzene occurs.

6. Bibliography

[1] "WHO | Infectious diseases," *WHO*, 2016.

[2] Director of Strategy and Communications, "Current issue of Microbiology Today | Microbiology Society," 2014. [Online]. Available: http://www.microbiologysociety.org/all-microsite-sections/microbiology-today/index.cfm/article/E57E1D0F-F935-422E-89EC82A82944199B. [Accessed: 14-Apr-2016].

[3] A. Boutayeb and S. Boutayeb, "International Journal for Equity in The burden of non communicable diseases in developing countries," vol. 8, pp. 1–8, 2005.

[4] K. Dheda, M. Ruhwald, G. Theron, J. Peter, and W. C. Yam, "Point-of-care diagnosis of tuberculosis: Past, present and future," *Respirology*, vol. 18, no. 2, pp. 217–232, 2013.

[5] "WHO | Tuberculosis," *WHO*, 2016.

[6] "WHO | Pneumonia," *WHO*, 2016.

[7] S. Fujitani, H. Sun, and V. L. Yu, "Recent Advances in Chest Medicine Pneumonia Due to Pseudomonas aeruginosa," pp. 909–919, 2016.

[8] R. Schnabel, R. Fijten, A. Smolinska, J. Dallinga, and F. J. Van Schooten, "Analysis of volatile organic compounds in exhaled breath to diagnose ventilator-associated pneumonia," *Nat. Publ. Gr.*, October, pp. 1–10, 2015.

[9] L. Penny, P. Neuwald, M. Urdea, H. Associates, and H. Street, "Biomarkers for Infectious Disease Diagnostics in the Developing World : Rapid , Home-based Diagnosis of Malaria in Symptomatic Individuals and Screening of Asymptomatic Pregnant Women," pp. 1–41, 2009.

[10] "WHO | Malaria," *WHO*, 2016.

[11] "WHO | Antimicrobial resistance," *WHO*, 2016.

[12] A. Z. Berna, J. S. McCarthy, R. X. Wang, K. J. Saliba, F. G. Bravo, J. Cassells, B. Padovan, and S. C. Trowell, "Analysis of Breath Specimens for Biomarkers of *Plasmodium falciparum* Infection," *J. Infect. Dis.*, vol. 212, no. 7, pp. 1120–1128, 2015.

[13] A. Mazzatenta, C. Di Giulio, and M. Pokorski, "Pathologies currently identified by exhaled biomarkers," *Respir. Physiol. Neurobiol.*, vol. 187, no. 1, pp. 128–134, Jun. 2013.

[14] O. Lazcka, F. J. Del Campo, and F. X. Muñoz, "Pathogen detection: A perspective of traditional methods and biosensors," *Biosens. Bioelectron.*, vol. 22, no. 7, pp. 1205–1217, 2007.

[15] A. H. Peruski and L. F. Peruski, "Immunological methods for detection and identification of infectious disease and biological warfare agents.," *Clin. Diagn. Lab. Immunol.*, vol. 10, no. 4, pp. 506–513, 2003.

[16] D. Klein, "Quantification using real-time PCR technology: Applications and limitations," *Trends Mol. Med.*, vol. 8, no. 6, pp. 257–260, 2002.

[17] C. Turner, "Techniques and issues in breath and clinical sample headspace analysis for disease diagnosis," 2016.

[18] R. P. Arasaradnam, E. Westenbrink, M. J. Mcfarlane, R. Harbord, S. Chambers, N. O'Connell, C. Bailey, C. U. Nwokolo, K. D. Bardhan, R. Savage, J. A. Covington, N. O. Connell, C. Bailey, C.

U. Nwokolo, K. D. Bardhan, R. Savage, and J. A. Covington, “Differentiating Coeliac Disease from Irritable Bowel Syndrome by Urinary Volatile Organic Compound Analysis - A Pilot Study,” *PLoS One*, vol. 9, no. 10, p. e107312, 2014.

[19] C. Lourenço and C. Turner, “Breath Analysis in Disease Diagnosis: Methodological Considerations and Applications,” pp. 465–498, 2014.

[20] M. Shirasu and K. Touhara, “JB Review The scent of disease : volatile organic compounds of the human body related to disease and disorder,” vol. 150, no. 3, pp. 257–266, 2011.

[21] H. Kataoka, K. Saito, H. Kato, and K. Masuda, “Noninvasive analysis of volatile biomarkers in human emanations for health and early disease diagnosis,” *Bioanalysis*, vol. 5, no. 11, pp. 1443–1459, Jun. 2013.

[22] A. W. Boots, J. J. B. N. van Berkel, J. W. Dallinga, A. Smolinska, E. F. Wouters, and F. J. van Schooten, “The versatile use of exhaled volatile organic compounds in human health and disease.,” *J. Breath Res.*, vol. 6, no. 2, p. 27108, Jun. 2012.

[23] D. Smith and S. Patrik, “The challenge of breath analysis for clinical diagnosis and therapeutic monitoring,” pp. 390–396, 2007.

[24] S. Sethi, R. Nanda, and T. Chakraborty, “Clinical Application of Volatile Organic Compound Analysis for Detecting Infectious Diseases,” *Clin. Microbiol. Rev.*, vol. 26, no. 3, pp. 462–475, Jul. 2013.

[25] K. D. van de Kant, L. J. van der Sande, Q. Jobsis, O. C. van Schayck, and E. Dompeling, “Clinical use of exhaled volatile organic compounds in pulmonary diseases: a systematic review,” *Respir. Res.*, vol. 13, no. 1, p. 117, 2012.

[26] S. J. Amann, A., Miekisch W, Pleil J, Risby T, “Methodological issues of sample collection and analysis of exhaled breath,” 2010.

[27] J. Pereira, P. Porto-Figueira, C. Cavaco, K. Taunk, S. Rapole, R. Dhakne, H. Nagarajaram, and J. Câmara, “Breath Analysis as a Potential and Non-Invasive Frontier in Disease Diagnosis: An Overview,” *Metabolites*, vol. 5, no. 1, pp. 3–55, 2015.

[28] N. J. W. Rattray, Z. Hamrang, D. K. Trivedi, R. Goodacre, and S. J. Fowler, “Taking your breath away : metabolomics breathes life in to personalized medicine,” *Trends Biotechnol.*, vol. 32, no. 10, pp. 538–548, 2014.

[29] K.-H. Kim, S. A. Jahan, and E. Kabir, “A review of breath analysis for diagnosis of human health,” *TrAC Trends Anal. Chem.*, vol. 33, pp. 1–8, Mar. 2012.

[30] M. Bishop, E. Fody, and L. Schoeff, *Clinical Chemistry - Techniques, Principles, Correlations*, 6th ed. 2010.

[31] A. Tipler, “An Introduction to Headspace Sampling in Gas Chromatography,” *PerkinElmer, Inc.*, pp. 3–33, 2013.

[32] W. Cao, Y. Duan, W. Cao, and Y. Duan, “Current Status of Methods and Techniques for Breath Analysis,” in *Critical Reviews in Analytical Chemistry*, p. 8347, February, 2007.

[33] T. H. E. Analyst, “Stability of selected volatile breath constituents in Tedlar , Kynar and Flexfilm sampling bags,” 2013.

[34] S. M. Cho, Y. Jun, G. Suk, and S. Shin, “Two-step preconcentration for analysis of exhaled gas of human breath with electronic nose,” vol. 117, pp. 50–57, 2006.

[35] M. Li, S. Biswas, M. H. Nantz, R. M. Higashi, and X. Fu, "Sensors and Actuators B : Chemical A microfabricated preconcentration device for breath analysis," *Sensors Actuators B. Chem.*, vol. 180, pp. 130–136, 2013.

[36] J. C. Reynolds, M. A. Jimoh, and C. Guallar-hoyas, "Analysis of human breath samples using a modified thermal desorption : gas chromatography electrospray ionization interface," vol.8, pp 1-10, 2014.

[37] A. H. Neerincx, B. P. Geurts, M. F. J. Habets, J. A. Booij, J. van Loon, J. J. Jansen, L. M. C. Buydens, J. van Ingen, J. W. Mouton, F. J. M. Harren, R. A. Wevers, P. J. F. M. Merkus, S. M. Cristescu, and L. A. J. Kluijtmans, "Identification of Pseudomonas aeruginosa and Aspergillus fumigatus mono- and co-cultures based on volatile biomarker combinations," *J. Breath Res.*, vol. 10, 2016.

[38] J. Pereira, P. Porto-figueira, C. Cavaco, K. Taunk, S. Rapole, R. Dhakne, H. Nagarajaram, and J. S. Câmara, "Breath Analysis as a Potential and Non-Invasive Frontier in Disease Diagnosis: An Overview," pp. 3–55, 2014.

[39] S. U. Savelev, J. D. Perry, S. J. Bourke, H. Jary, R. Taylor, a. J. Fisher, P. a. Corris, M. Petrie, and a. De Soyza, "Volatile biomarkers of Pseudomonas aeruginosa in cystic fibrosis and noncystic fibrosis bronchiectasis," *Lett. Appl. Microbiol.*, vol. 52, pp. 610–613, 2011.

[40] G. Preti, E. Thaler, C. W. Hanson, M. Troy, J. Eades, and A. Gelperin, "Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus : Analysis by solid-phase microextraction and gas chromatography – mass spectrometry," *J. Chromatogr. B*, vol. 877, pp. 2011–2018, 2009.

[41] E. Tait, J. D. Perry, S. P. Stanforth, and J. R. Dean, "Identification of Volatile Organic Compounds Produced by Bacteria Using HS-SPME-GC – MS," pp. 363–373, 2014.

[42] C. Zscheppank, H. L. Wiegand, and C. Lenzen, "Investigation of volatile metabolites during growth of Escherichia coli and Pseudomonas aeruginosa by needle trap-GC-MS," pp. 6617–6628, 2014.

[43] H. L. Lord, W. Zhan, and J. Pawliszyn, "Fundamentals and applications of needle trap devices. A critical review," *Anal. Chim. Acta*, vol. 677, pp. 3–18, 2010.

[44] A. Wang, F. Fang, and J. Pawliszyn, "Sampling and determination of volatile organic compounds with needle trap devices," vol. 1072, pp. 127–135, 2005.

[45] M. Mieth, S. Kischkel, J. K. Schubert, D. Hein, and W. Miekisch, "Multibed Needle Trap Devices for on Site Sampling and Preconcentration of Volatile Breath Biomarkers," vol. 81, no. 14, pp. 5851–5857, 2009.

[46] L. Pauling, A. B. Robinson, R. O. Y. Teranishit, and P. Cary, "Quantitative Analysis of Urine Vapor and Breath by Gas-Liquid Partition Chromatography," vol. 68, no. 10, pp. 2374–2376, 1971.

[47] J. Herbig, J. Beauchamp, D. Smith, and P. Span, "Mass spectrometry for real-time quantitative breath analysis," vol.8, 2014.

[48] F. Rock, N. Barsan, and U. Weimar, "Electronic Nose : Current Status and Future Trends," *Am. Chem. Soc.*, vol. 108, pp. 705–725, 2008.

[49] S. Dragonieri, R. Schot, B. J. A. Mertens, S. Le Cessie, S. A. Gauw, A. Spanevello, O. Resta, N. P. Willard, T. J. Vink, K. F. Rabe, E. H. Bel, and P. J. Sterk, "An electronic nose in the discrimination of patients with asthma and controls," *J. Allergy Clin. Immunol.*, vol. 120, no. 4, pp. 856–862, 2007.

[50] T. Holm, “Aspects of the mechanism of the flame ionization detector,” vol. 842, pp. 221–227, 1999.

[51] X. Zhan, J. Duan, Y. Duan, and I. Introduction, “Recent developments of proton-transfer reaction mass spectrometry (PTR-MS) and its applications and its application in medical research,” pp. 143–165, 2013.

[52] A. A. Schwarz, K., W. Filipiak, “Determining concentration patterns of volatile compounds in exhaled breath by PTR-MS,” *J. Breath Res.*, 2009.

[53] A. Jordan, S. Haidacher, G. Hanel, E. Hartungen, L. Märk, H. Seehauser, R. Schottkowsky, P. Sulzer, and T. D. Märk, “A high resolution and high sensitivity proton-transfer-reaction time-of-flight mass spectrometer (PTR-TOF-MS),” *Int. J. Mass Spectrom.*, vol. 286, pp. 122–128, 2009.

[54] J. Herbig, M. Müller, S. Schallhart, T. Titzmann, M. Graus, and A. Hansel, “On-line breath analysis with PTR-TOF,” *J. Breath Res.*, vol. 3, no. 2, p. 27004, 2009.

[55] T. W. E. Chippendale, F. J. Gilchrist, P. Španěl, A. Alcock, W. Lenney, D. Smith, P. Spanel, A. Alcock, W. Lenney, and D. Smith, “Quantification by SIFT-MS of volatile compounds emitted by Aspergillus fumigatus cultures and in co-culture with Pseudomonas aeruginosa, Staphylococcus aureus and Streptococcus pneumoniae,” *Anal. Methods*, vol. 6, no. 20, pp. 8154–8164, Jul. 2014.

[56] D. Smith and S. Patrik, “Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace gas analysis,” *Mass Spectrom. Rev.*, no. 24.5, pp. 661–700, 2005.

[57] H. H. Hili, F. William, R. H. S. Louis, and D. Mcminn, “Ion Mobility Spectrometry,” vol. 62, no. 23, pp. 1201–1209, 1990.

[58] H. Borsdorf and G. A. Eiceman, “Ion Mobility Spectrometry : Principles and Applications,” pp. 323–375, February 2006.

[59] A. B. Kanu and P. H. H. Hill, “Ion mobility spectrometry : recent developments and novel applications,” no. May, 2004.

[60] T. Perl, M. Ju, W. Vautz, M. Kuhns, M. B. Zepelin, and M. Quintel, “Detection of characteristic metabolites of Aspergillus fumigatus and Candida species using ion mobility spectrometry-metabolic profiling by volatile organic compounds,” *Mycoses*, no. 54.6, pp. 828–837, 2011.

[61] N. Kunze, J. Göpel, M. Kuhns, M. Jünger, M. Quintel, and T. Perl, “Detection and validation of volatile metabolic patterns over different strains of two human pathogenic bacteria during their growth in a complex medium using multi-capillary column-ion mobility spectrometry (MCC-IMS).,” *Appl. Microbiol. Biotechnol.*, vol. 97, no. 8, pp. 3665–3676, 2013.

[62] M. Juenger, W. Vautz, M. Kuhns, L. Hofmann, S. S. Ulbricht, J. I. J. I. Baumbach, M. Quintel, T. Perl, “Ion mobility spectrometry for microbial volatile organic compounds: a new identification tool for human pathogenic bacteria,” *Appl. Environ. Microbiol.*, vol. 93, no. 6, pp. 2603–2614, Mar. 2012.

[63] B. J. . Ruzsanyi, V., “Analysis of human breath using IMS,” *Int J Ion Mobil. Spectrom*, vol. 8, pp. 5–7, February 2005.

[64] H. D. Bean, J. Zhu, and J. E. Hill, “Characterizing Bacterial Volatiles using Secondary Electrospray Ionization Mass Spectrometry (SESI-MS),” *J.Vis.Exp*, no. 52, 2011.

[65] J. Zhu and J. E. Hill, “Detection of Escherichia coli via VOC profiling using secondary electrospray ionization-mass spectrometry (SESI-MS),” *Food Microbiol.*, vol. 34, no. 2, pp. 412–417, 2013.

[66] J. Zhu, H. D. Bean, J. Jiménez-Díaz, J. E. Hill, J. Jimenez-Diaz, and J. E. Hill, "Secondary electrospray ionization-mass spectrometry (SESI-MS) breathprinting of multiple bacterial lung pathogens, a mouse model study," *J. Appl. Physiol.*, vol. 114, no. 11, pp. 1544–1549, 2013.

[67] H. D. Bean, J. Zhu, J. C. Sengle, and J. E. Hill, "Identifying methicillin-resistant *Staphylococcus aureus* (MRSA) lung infections in mice via breath analysis using secondary electrospray ionization-mass spectrometry (SESI-MS)," *J. Breath Res.*, vol. 8, no. 4, p. 041001, 2014.

[68] U. Tegtmeyer, H. P. Weiss, and R. Schlogl, "Gas analysis by IMR-MS : a comparison to conventional mass spectrometry," *J. Anal. Chem.*, no. 347, pp. 263–268, 1993.

[69] K.Leiter, J.Hirber, M.Andriollo, and W.Lindinger, "IMR-MS (ION-MOLECULE-REACTION MASS-SPECTROMETRY)- A DIAGNOSTIC TOOL FOR ANALYSIS OF GASEOUS COMPONENTS." 1991.

[70] M. E. Dolch, C. Hornuss, C. Klocke, S. Praun, J. Villinger, W. Denzer, G. Schelling, and S. Schubert, "Volatile compound profiling for the identification of Gram-negative bacteria by ion-molecule reaction-mass spectrometry," *J. Appl. Microbiol.*, vol. 113, no. 5, pp. 1097–1105, 2012.

[71] H. S. U. Guntow and K. H. R. Schlögl, "Performance and application potential of ion-molecule reaction mass spectrometry (IMR-MS) in the analysis of complex gas mixtures," *J. Anal. Chem.*, no. 356, pp. 127–137, 1996.

[72] C. Lourenço and C. Turner, "Breath Analysis in Disease Diagnosis: Methodological Considerations and Applications," *Metabolites*, vol. 4, no. 2, pp. 465–498, 2014.

[73] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 2, no. 4, pp. 433–459, 2010.

[74] C. Di Natale, A. Macagnano, E. Martinelli, R. Paolesse, G. D'Arcangelo, C. Roscioni, A. Finazzi-Agrô, and A. D'Amico, "Lung cancer identification by the analysis of breath by means of an array of non-selective gas sensors," *Biosens. Bioelectron.*, vol. 18, no. 10, pp. 1209–1218, 2003.

[75] P. Montuschi, M. Santonico, C. Mondino, G. Pennazza, G. Maritini, E. Martinelli, R. Capuano, G. Ciabattoni, R. Paolesse, C. Di Natale, P. J. Barnes, and A. D'Arnico, "Diagnostic performance of an electronic nose, fractional exhaled nitric oxide, and lung function testing in asthma," *Chest*, vol. 137, no. 4, pp. 790–796, 2010.

[76] M. A. Craven and J. . Gardner, "Electronic noses - development and future prospects," *Trends Anal. Chem.*, vol. 15, no. 9, pp. 486–493, 1996.

[77] N. Fens, M. P. Van Der Schee, P. Brinkman, and P. J. Sterk, "Exhaled breath analysis by electronic nose in airways disease . Established issues and key questions," *Clin. Exp. Allergy*, no. 43, pp. 705–715, 2012.

[78] A. P. F. Turner and N. Magan, "Electronic noses and disease diagnostics," *Nat. Rev. Microbiol.*, vol. 2, no. 2, pp. 161–166, 2004.

[79] A. D. Wilson, "Advances in Electronic-Nose Technologies for the Detection of Volatile Biomarker Metabolites in the Human Breath," *Metabolites*, no. 5, pp. 140–163, 2015.

[80] A. K. Pavlou, N. Magan, J. Meecham, J. Brown, P. Klatser, and A. P. F. Turner, "Detection of Mycobacterium tuberculosis (TB) in vitro and in situ using an electronic nose in combination with a neural network system," vol. 20, pp. 538–544, 2004.

[81] C. Wang and P. Sahay, "Breath analysis using laser spectroscopic techniques: Breath biomarkers, spectral fingerprints, and detection limits," *Sensors*, vol. 9, no. 10, pp. 8230–8262, 2009.

[82] L. Aba, "Selectivity Improvement of Gas Sensor Based on Poly(3,4-ethylenedioxythiophene):Poly(styrenesulfonate) Thin Film by Using Imprinting Method," *J. Mod. Phys.*, vol. 03, pp. 529–533, 2012.

[83] J.-C. Lei, C. Hou, D.-Q. Huo, X. Luo, M. Bao, X. Li, M. Yang, and H. Fa, "A novel device based on a fluorescent cross-responsive sensor array for detecting lung cancer related volatile organic compounds," *Rev. Sci. Instrum.*, vol. 86, no. 2, p. 25106, 2015.

[84] S. Chen, Y. Wang, and S. Choi, "Applications and Technology of Electronic Nose for Clinical Diagnosis," *Open J. Appl. Biosens.*, vol. 02, pp. 39–50, 2013.

[85] J. M. Sanchez and M. Alonso, "Analytical challenges in breath analysis and its application to exposure monitoring," *Trends Anal. Chem.*, vol. 44, 2013.

[86] K. Schmidt and I. Podmore, "Current Challenges in Volatile Organic Compounds Analysis as Potential Biomarkers of Cancer," 2015.

[87] R. M. S. Thorn and J. Greenman, "Microbial volatile compounds in health and disease conditions," *J. Breath Res.*, vol. 6, no. 2, Jun. 2012.

[88] D. S. Miller, X. Wang, and N. L. Abbott, "Design of functional materials based on liquid crystalline droplets," *Chem. Mater.*, vol. 26, no. 1, pp. 496–506, 2014.

[89] Y.-T. Lai, J.-C. Kuo, and Y.-J. Yang, "A novel gas sensor using polymer-dispersed liquid crystal doped with carbon nanotubes," *Sensors Actuators A Phys.*, vol. 215, pp. 83–88, 2014.

[90] R. Wise, T. Hart, O. Cars, M. Streulens, R. Helmuth, P. Huovinen, and M. Sprenger, "Antimicrobial resistance. Is a major threat to public health.," *BMJ*, vol. 317, no. 7159, pp. 609–610, 1998.

[91] T. M. Barbosa and S. B. Levy, "The impact of antibiotic use on resistance development and persistence," *Drug Resist.*, vol. 3,pp. 303–311, 2000.

[92] M. Sohrabi, L. Zhang, K. Zhang, A. Ahmetagic, and M. Q. Wei, "Volatile Organic Compounds as Novel Markers for the Detection of Bacterial Infections," *Clin. Microbiol.*, vol. 3, no. 3, pp. 1–6, 2014.

[93] F. Buljubasic and G. Buchbauer, "The scent of human diseases: a review on specific volatile organic compounds as diagnostic biomarkers," *Flavour Fragr. J.*, vol. 30, no. 1, pp. 5–25, 2015.

[94] C. M. Bishop, *Pattern Recognition and Machine Learning*, vol. 53, no. 9. 2013.

[95] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, vol. 22, no. 2. 2000.

[96] K. P. Murphy, "Machine Learning: A Probabilistic Perspective," *MIT Press*, p. 414, 1997.

[97] K.Z.Mao, "Orthogonal Forward Selection and Backward Elimination Algorithms for Feature Subset Selection," *IEEE Trans. Syst. man, Cybern. B Cybern.*, vol. 34, no. 1, pp. 629–634, 2004.

[98] T.-T. Wong, "Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation," *Pattern Recognit.*, vol. 48, no. 9, pp. 2839–2846, 2015.

[99] T. Perl, M. Juenger, W. Vautz, J. Nolte, M. Kuhns, M. B. Zepelin, and M. Quintel, "Detection of characteristic metabolites of Aspergillus fumigatus and Candida species using ion mobility spectrometry - metabolic profiling by volatile organic compounds," *Mycoses*, vol. 54, no. 6, pp. E828–E837, 2011.

[100] C. E. Garner, S. Smith, B. de Lacy Costello, P. White, R. Spencer, C. S. J. Probert, N. M. Ratcliffe, B. de L. Costello, P. White, R. Spencer, C. S. J. Probert, N. M. Ratcliffe, V. O. C. Garner, D. Lacy, and N. M. Volatile, "Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease," *Faseb J.*, vol. 21, no. 8, pp. 1675–1688, 2007.

[101] M. K. Storer, K. Hibbard-Melles, B. Davis, and J. Scotter, "Detection of volatile compounds produced by microbial growth in urine by selected ion flow tube mass spectrometry (SIFT-MS)," *J. Microbiol. Methods*, vol. 87, no. 1, pp. 111–113, 2011.

[102] I. Ahmed, R. Greenwood, B. D. L. Costello, N. M. Ratcliffe, and C. S. Probert, "An Investigation of Fecal Volatile Organic Metabolites in Irritable Bowel Syndrome," *PLoS One*, vol. 8, no. 3, p. e58204, 2013.

[103] R. Michael, S. Thorn, D. Michael, and J. Greenman, "Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro," *J. Microbiol. Methods*, vol. 84, no. 2, pp. 258–264, 2011.

[104] A. W. Boots, A. Smolinska, J. J. B. N. van Berkel, R. R. R. Fijten, E. E. Stobberingh, M. L. L. Boumans, E. J. Moonen, E. F. M. Wouters, J. W. Dallinga, and F. J. Van Schooten, "Identification of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography-mass spectrometry.," *J. Breath Res.*, vol. 8, no. 2, p. 27106, Jun. 2014.

[105] E. Tait, J. D. Perry, S. P. Stanforth, and J. R. Dean, "Identification of Volatile Organic Compounds Produced by Bacteria Using HS-SPME-GC – MS," *J. Chromatogr. Sci.*, no. 52, pp. 363–373, 2014.

[106] E. Tait, J. D. Perry, S. P. Stanforth, and J. R. Dean, "Use of volatile compounds as a diagnostic tool for the detection of pathogenic bacteria," *Trends Anal. Chem.*, vol. 53, pp. 117–125, 2014.

[107] S. Maddula and L. M. Blank, "Detection of volatile metabolites of Escherichia coli by multi capillary column coupled ion mobility spectrometry," pp. 791–800, 2009.

[108] R. A. Allardyce, V. S. Langford, A. L. Hill, and D. R. Murdoch, "Detection of volatile metabolites produced by bacterial growth in blood culture media by selected ion flow tube mass spectrometry (SIFT-MS)," vol. 65, pp. 361–365, 2006.

[109] M. Bunge, N. Araghipour, T. Mikoviny, R. Schnitzhofer, A. Hansel, F. Schinner, A. Wisthaler, R. Margesin, and T. D. Ma, "On-Line Monitoring of Microbial Volatile Metabolites by Proton Transfer Reaction-Mass Spectrometry," *Appl. Environ. Microbiol.*, vol. 74, no. 7, pp. 2179–2186, 2008.

[110] N. J. Hayward, T. H. Jeavons, A. J. C. Nicholson, and A. G. Thornton, "Development of Specific Tests for Rapid Detection of Escherichia coli and all Species of Proteus in Urine," *J. Clin. Microbiol.*, vol. 6, no. 3, pp. 195–201, 1977.

[111] A. Bond, A. Vernon, S. Reade, A. Mayor, C. Minetti, J. Wastling, K. Lamden, and C. Probert, "Investigation of Volatile Organic Compounds Emitted from Faeces for the Diagnosis of Giardiasis," *J. Gastrointest. Liver Dis.*, vol. 24, no. 3, pp. 281–286, 2015.

[112] W. Filipiak, a. Sponring, M. M. Baur, C. Ager, a. Filipiak, H. Wiesenhofer, M. Nagl, J. Troppmair, and a. Amann, "Characterization of volatile metabolites taken up by or released from Streptococcus pneumoniae and Haemophilus influenzae by using GC-MS," *Microbiology*, vol. 158, no. 12, pp. 3044–3053, 2012.

[113] A. Abd, E. Qader, D. Lieberman, and Y. Shemer, "Volatile organic compounds generated by cultures of bacteria and viruses associated with respiratory infections," *Biomed. Chromatogr.* April, 2015.

[114] A. Ulanowska, T. Kowalkowski, K. Hrynkiewicz, M. Jackowski, and B. Buszewski, "Determination of volatile organic compounds in human breath for Helicobacter pylori detection by SPME-GC/MS," *Biomed. Chromatogr.*, vol. 25, no. 3, pp. 391–397, 2011.

[115] B. Jia, B. Sohnlein, K. Mortelmans, M. Coggiola, and H. Oser, "Distinguishing Methicillin-Resistant and Sensitive Staphylococcus aureus Using Volatile Headspace Metabolites," *IEEE Sens. J.*, vol. 10, no. 1, pp. 71–75, 2010.

[116] A. Cheepsattayakorn and R. Cheepsattayakorn, "Breath Tests in Diagnosis of Pulmonary Tuberculosis," *Recent Pat. Biotechnol.*, vol. 8, no. No.2, pp. 172–175, 2015.

[117] M. Phillips, V. Basa-Dalay, G. Bothamley, R. N. Cataneo, P. Kim, M. Piedad, R. Natividad, P. Schmitt, J. Wai, P. K. Lam, M. P. R. Natividad, P. Schmitt, and J. Wai, "Breath biomarkers of active pulmonary tuberculosis," *Tuberculosis*, vol. 90, no. 2, pp. 145–151, Mar. 2010.

[118] M. Phillips, R. N. Cataneo, R. Condos, G. A. Ring, J. Greenberg, V. La, M. I. Munawar, and O. Tietje, "Volatile biomarkers of pulmonary tuberculosis in the breath," pp. 44–52, 2007.

[119] G. F. Mgode, B. J. Weetjens, T. Nawrath, D. Lazar, C. Cox, M. Jubitana, A. Mahoney, D. Kuipers, R. S. Machang, J. Weiner, S. Schulz, and S. H. E. Kaufmann, "Mycobacterium tuberculosis volatiles for diagnosis of tuberculosis by Cricetomys rats," *Tuberculosis*, vol. 92, no. 6, pp. 535–542, 2012.

[120] M. Syhre and S. T. Chambers, "The scent of Mycobacterium tuberculosis," *Tuberculosis*, vol. 88, no. 4, pp. 317–323, 2008.

[121] K. M. Banday, K. K. Pasikanti, E. C. Y. Chan, R. Singla, K. V. S. Rao, V. S. Chauhan, and R. K. Nanda, "Use of Urine Volatile Organic Compounds To Discriminate Tuberculosis Patients from Healthy Subjects," *Anal. Chem.*, vol. 83, no. 14, pp. 5526–5534, 2011.

[122] W. L. Wood, D. J. Higbee, M. Gooldy, S. Glogowski, R. Fitzpatrick, R. J. Karalus, T. D. Wood, and D. J. Mangino, "Analysis of Volatile Bacterial Metabolites by Gas Chromatography – Mass Spectrometry," *Spectroscopy*, vol. 21, pp. 21–28, 2006.

[123] W. Filipiak, A. Sponring, M. M. Baur, A. Filipiak, C. Ager, H. Wiesenhofer, M. Nagl, J. Troppmair, and A. Amann, "Molecular analysis of volatile metabolites released specifically by staphylococcus aureus and pseudomonas aeruginosa," *Bmc Microbiol.*, vol. 12, no. 1, p. 113, 2012.

[124] M. Barker, M. Hengst, J. Schmid, H.-J. Buers, B. Mittermaier, D. Klemp, and R. Koppmann, "Volatile organic compounds in the exhaled breath of young patients with cystic fibrosis.," *Eur. Respir. J.*, vol. 27, no. 5, pp. 929–36, 2006.

[125] P. C. Goeminne, T. Vandendriessche, J. Van Eldere, B. M. Nicolai, and M. L. Hertog, "Detection of Pseudomonas aeruginosa in sputum headspace through volatile organic compound analysis," *Respir. Res.*, vol. 13, no. 1, p. 1, 2012.

[126] V. Shestivska, A. Nemec, P. Dřevínek, K. Sovová, K. Dryahina, and P. Spaněl, "Quantification of methyl thiocyanate in the headspace of Pseudomonas aeruginosa cultures and in the breath of cystic fibrosis patients by selected ion flow tube mass spectrometry.," *Rapid Commun. Mass Spectrom.*, vol. 25, no. 17, pp. 2459–2467, 2011.

[127] C. Scholler, M. Sorer, and K. Wilkins, "Volatile metabolites from some gram-negative bacteria," *Chemosphere*, vol. 35, no. 7, pp. 1487–1495, 1997.

[128] W. Carroll, A. Alcock, and D. Smith, "Detection of Volatile Compounds Emitted by Pseudomonas aeruginosa Using Selected Ion Flow Tube Mass Spectrometry," vol. 456, no. September 2004, pp. 452–456, 2005.

[129] C. D. Cox and J. Parker, "Use of 2-Aminoacetophenone Production in Identification of Pseudomonas aeruginosa," vol. 9, no. 4, pp. 479–484, 1979.

[130] J. N. Labows, K. J. Mcginley, G. U. Y. F. Webster, and J. J. Leyden, "Headspace Analysis of Volatile Metabolites of Pseudomonas aeruginosa and Related Species by Gas Chromatography-Mass Spectrometry," vol. 12, no. 4, pp. 521–526, 1980.

[131] W. Filipiak, R. Beer, A. Sponring, A. Filipiak, C. Ager, and A. Schiefecker, "Breath analysis for in vivo detection of pathogens related to ventilator-associated pneumonia in intensive care patients : A prospective pilot study Breath analysis for in vivo detection of pathogens related to ventilator-associated pneumonia in intensive," *J. Breath Res.*, no. October, p. 16004, 2015.

[132] E. Brodrick, A. Davies, P. Neill, L. Hanna, and E. M. Williams, "Breath analysis: translation into clinical practice.," *J. Breath Res.*, vol. 9, no. 2, p. 27109, Jun. 2015.

[133] B. de L. Costello, A. Amann, H. Al-Kateb, C. Flynn, W. Filipiak, T. Khalid, D. Osborne, N. M. Ratcliffe, B. de Lacy Costello, A. Amann, H. Al-Kateb, C. Flynn, W. Filipiak, T. Khalid, D. Osborne, and N. M. Ratcliffe, "A review of the volatiles from the healthy human body," *J. Breath Res.*, vol. 8, no. 1, p. 14001, 2014.

[134] L. D. J. Bos, P. J. Sterk, and M. J. Schultz, "Volatile Metabolites of Pathogens: A Systematic Review," *PLoS Pathog.*, vol. 9, no. 5, p. e1003311, 2013.

[135] M. C. Lemfack, J. Nickel, M. Dunkel, R. Preissner, and B. Piechulla, "mVOC: a database of microbial volatiles.," *Nucleic Acids Res.*, vol. 42, pp. D744–D748, Jan. 2014.

[136] A. K. Pavlou, N. Magan, C. McNulty, J. Jones, D. Sharp, J. Brown, and A. P. F. Turner, "Use of an electronic nose system for diagnoses of urinary tract infections.," *Biosens. Bioelectron.*, vol. 17, no. 10, pp. 893–899, 2002.

[137] N. Guernion, N. M. Ratcliffe, P. T. N. Spencer-Phillips, and R. a. Howe, "Identifying bacteria in human urine: Current practice and the potential for rapid, near-patient diagnosis by sensing volatile organic compounds," *Clin. Chem. Lab. Med.*, vol. 39, no. 10, pp. 893–906, 2001.

[138] M. J. Stephen and J. P. Straley, *Physics of liquid crystals Liquid Crystals in General*, vol. 46, no. 4. 1974.

[139] P. V Shibaev, M. Wenzlick, J. Murray, and A. Tantillo, "Rebirth of Liquid Crystals for Sensoric Applications : Environmental and Gas Sensors," vol. 2015, 2015.

[140] D. a. Winterbottom, R. Narayanaswamy, and I. M. Raimundo, "Cholesteric liquid crystals for detection of organic vapours," *Sensors Actuators, B Chem.*, vol. 90, no. 1–3, pp. 52–57, 2003.

[141] S. Mohanty, "Liquid crystals — The 'fourth' phase of matter," *Resonance*, vol. 8, pp. 52–70, 2003.

[142] X. Ding and K.-L. Yang, "Liquid crystal based optical sensor for detection of vaporous butylamine in air," *Sensors Actuators B Chem.*, vol. 173, pp. 607–613, 2012.

[143] A. Sen, K. a. Kupcho, B. a. Grinwald, H. J. Vantreeck, and B. R. Acharya, "Liquid crystal-based sensors for selective and quantitative detection of nitrogen dioxide," *Sensors Actuators, B Chem.*, vol. 178, pp. 222–227, 2013.

[144] A. Biswas, R. L. Shogren, D. G. Stevenson, J. L. Willett, and P. K. Bhowmik, "Ionic liquids as solvents for biopolymers: Acylation of starch and zein protein," *Carbohydr. Polym.*, vol. 66, no. 4, pp. 546–550, 2006.

[145] D. Mecerreyes, *Applications of Ionic Liquids in Polymer Science and Technology*. .

[146] K. Goossens, K. Lava, C. W. Bielawski, and K. Binnemans, “Ionic Liquid Crystals: Versatile Materials.,” *Chem. Rev.*, vol. 116, no. 8, pp. 4643–4807, 2016.

[147] N. a. Smirnova and E. a. Safonova, “Ionic liquids as surfactants,” *Russ. J. Phys. Chem. A*, vol. 84, no. 10, pp. 1695–1704, 2010.

[148] L. Marin, M. C. Popescu, A. Zabulica, H. Uji-I, and E. Fron, “Chitosan as matrix for bio-polymer dispersed liquid crystal systems,” *Carbohydr. Polym.*, vol. 95, no. 1, pp. 16–24, 2013.

[149] A. Isabel, K. Petrova, M. Teresa, and J. Sotomayor, “New Polymer Networks for PDLC Films Application,” *New Polym. Spec. Appl.*, 2012.

[150] Q.-Z. Hu and C.-H. Jang, “Spontaneous formation of micrometer-scale liquid crystal droplet patterns on solid surfaces and their sensing applications,” *Soft Matter*, vol. 9, no. 24, pp. 5779–5784, 2013.

[151] P. V Shibaev, M. Wenzlick, J. Murray, and a Tantillo, “Liquid Crystalline Compositions as Gas Sensors,” *Mol. Cryst. Liq. Cryst.*, vol. 611, no. 1, pp. 94–99, 2015.

[152] R. J. Carlton, J. T. Hunter, D. S. Miller, R. Abbasi, P. C. Mushenheim, L. N. Tan, and N. L. Abbott, “Chemical and biological sensing using liquid crystals,” vol. 1, no. 1, pp. 29–51, 2014.

[153] S. Ju, K.-Y. Lee, S.-J. Min, Y. K. Yoo, K. S. Hwang, S. K. Kim, and H. Yi, “Single-carbon discrimination by selected peptides for individual detection of volatile organic compounds,” *Sci. Rep.*, vol. 5, p. 9196, 2015.

[154] D. Wang, S.-Y. Park, and I.-K. Kang, “Liquid crystals: emerging materials for use in real-time detection applications,” *J. Mater. Chem. C*, vol. 3, no. 35, pp. 9038–9047, 2015.

[155] T. Bera and J. Fang, “Interaction of Surfactants and Polyelectrolyte-Coated Liquid Crystal Droplets,” pp. 1–7, 2014.

[156] R. R. Rola, G. N. Bhola, and U. C. Bhoya, “Dependence of Mesomorphism on Molecular Rigidity,” *Int. Lett. Chem. Phys. Astron.*, vol. 47, pp. 67–76, 2015.

[157] M. L. Chauhan, B. K. Patel, S. N. Pandya, and a. V. Doshi, “Mesomorphism dependence on molecular rigidity,” *Mol. Cryst. Liq. Cryst.*, vol. 625, no. 1, pp. 47–54, 2016.

[158] C. N. Pace, S. Trevino, E. Prabhakaran, and J. M. Scholtz, “Protein structure, stability and solubility in water and other solvents,” *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 359, p. 1225, 2004.

[159] D. S. Mathew and R. S. Juang, “Role of alcohols in the formation of inverse microemulsions and back extraction of proteins/enzymes in a reverse micellar system,” *Sep. Purif. Technol.*, vol. 53, no. 3, pp. 199–215, 2007.

Appendix

Appendix 1

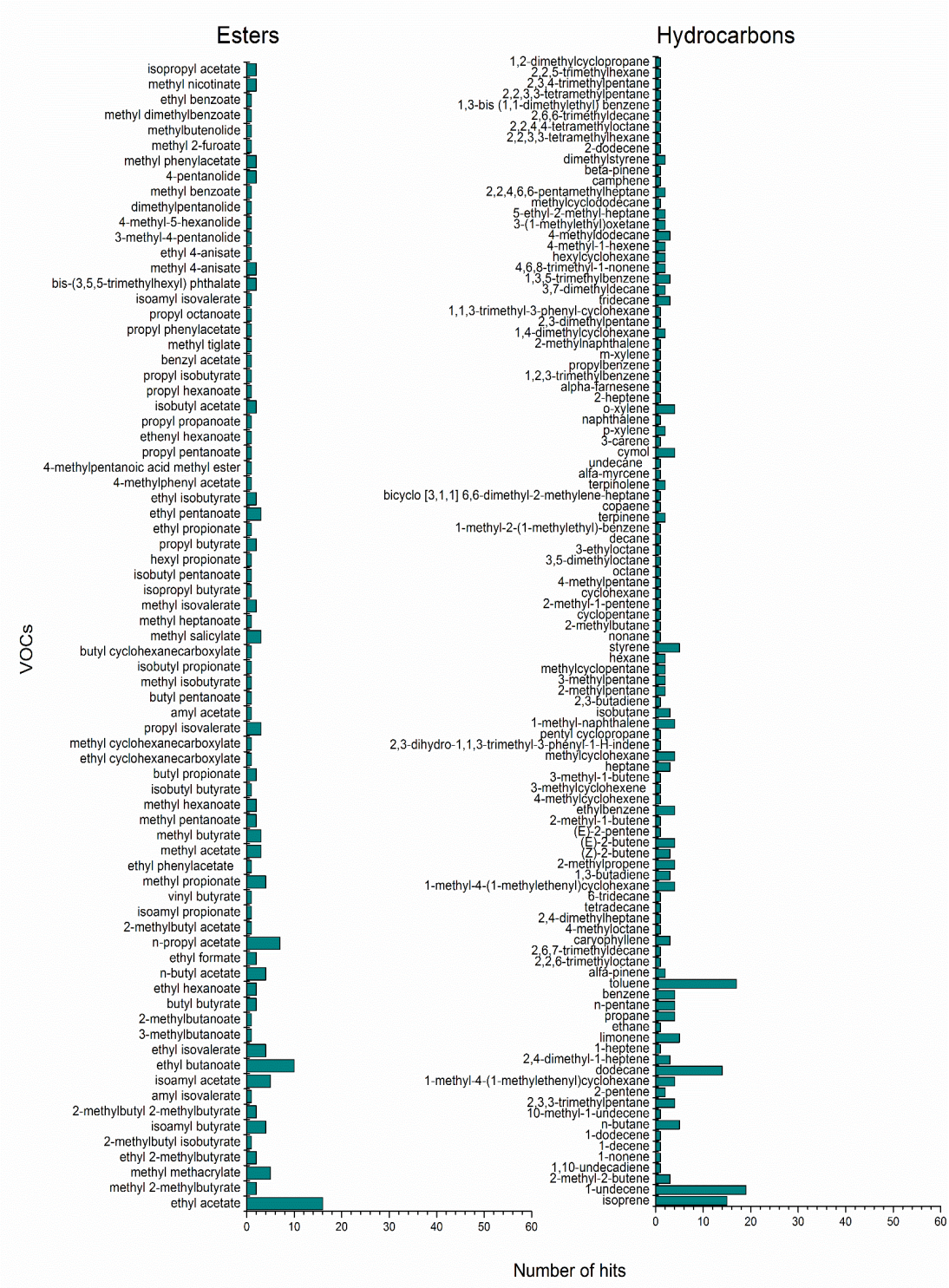

Figure A1- Representation of the number of hits for each individual VOC, in all the experiments, for the esters and hydrocarbons.

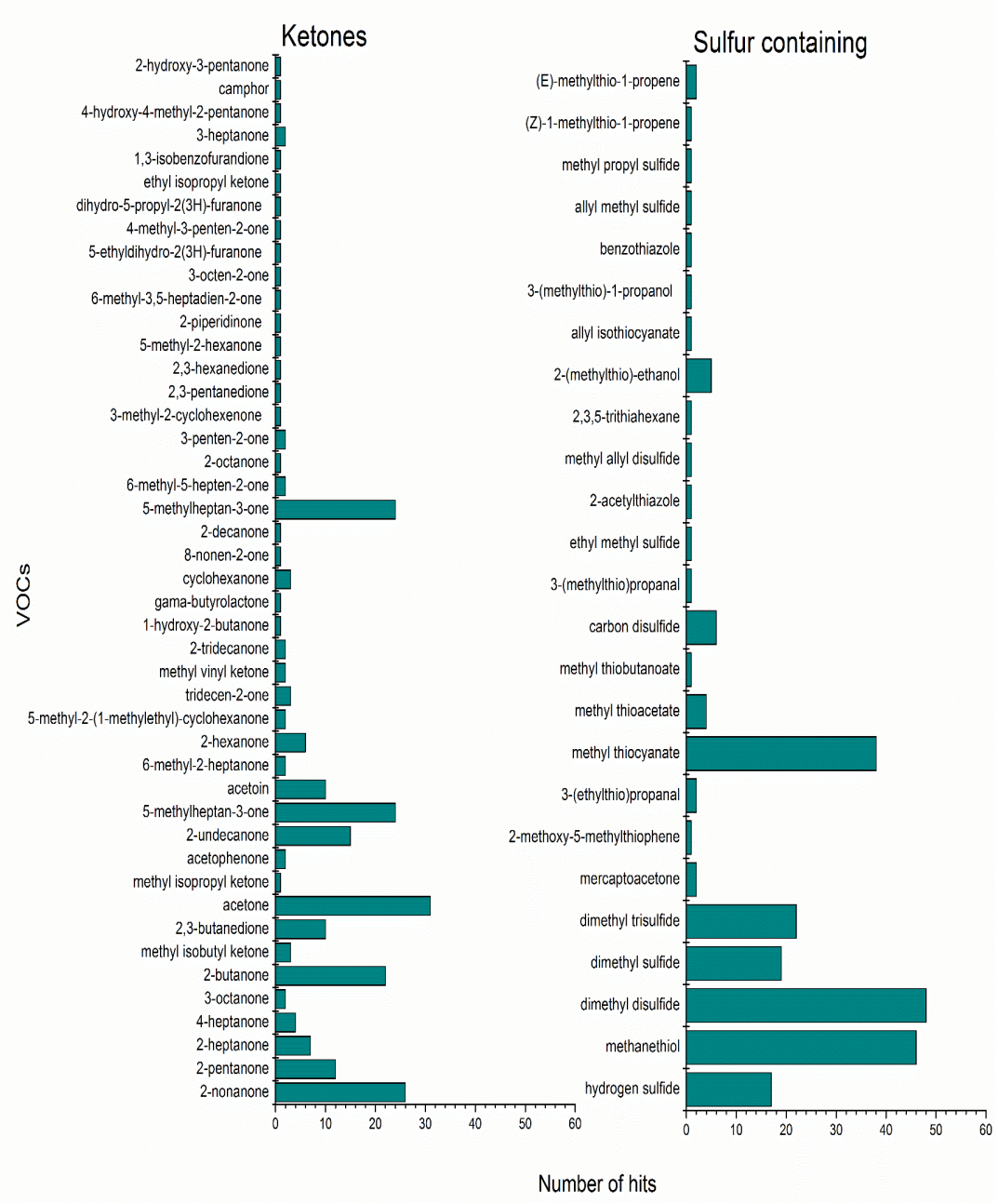

Figure A2- Representation of the number of hits for each individual VOC, in all the experiments, for the ketones and sulfur containing.

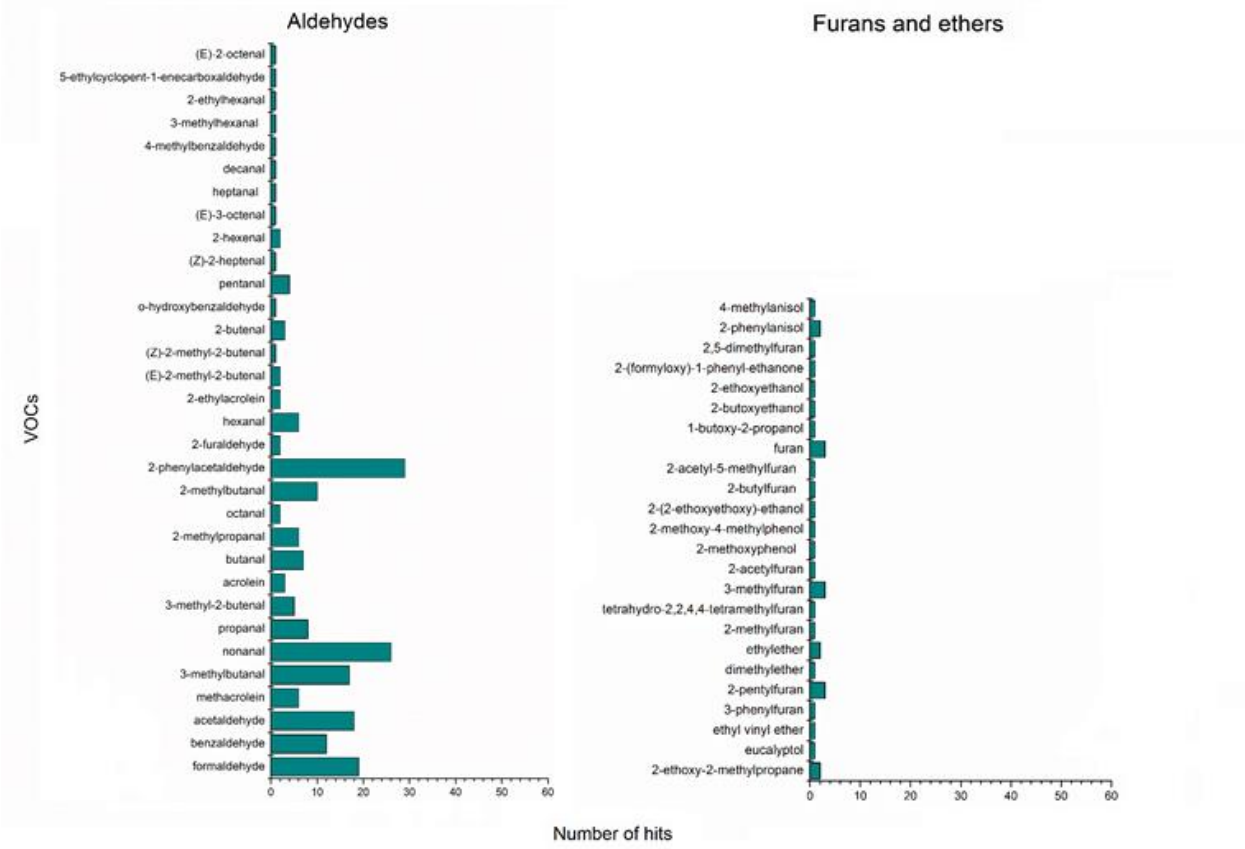

Figure A3 - Representation of the number of hits for each individual VOC, in all the experiments, for aldehydes and furans and ethers.

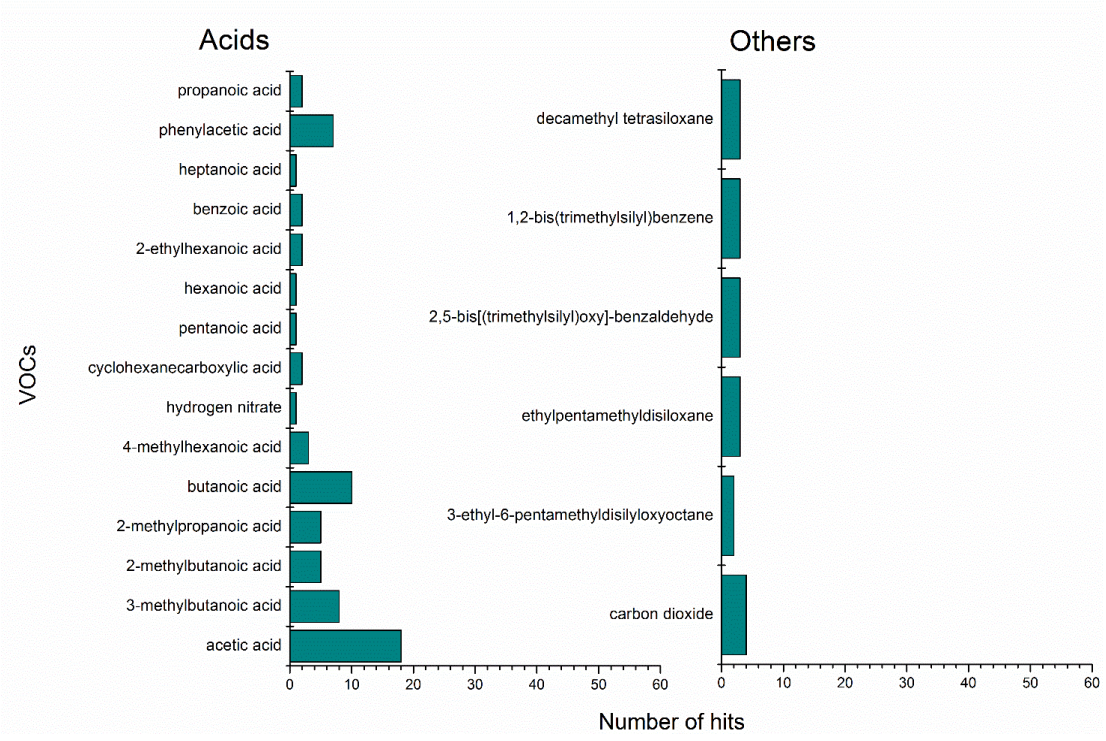

Figure A4- Representation of the number of hits for each individual VOC, in all the experiments, for acids and others.

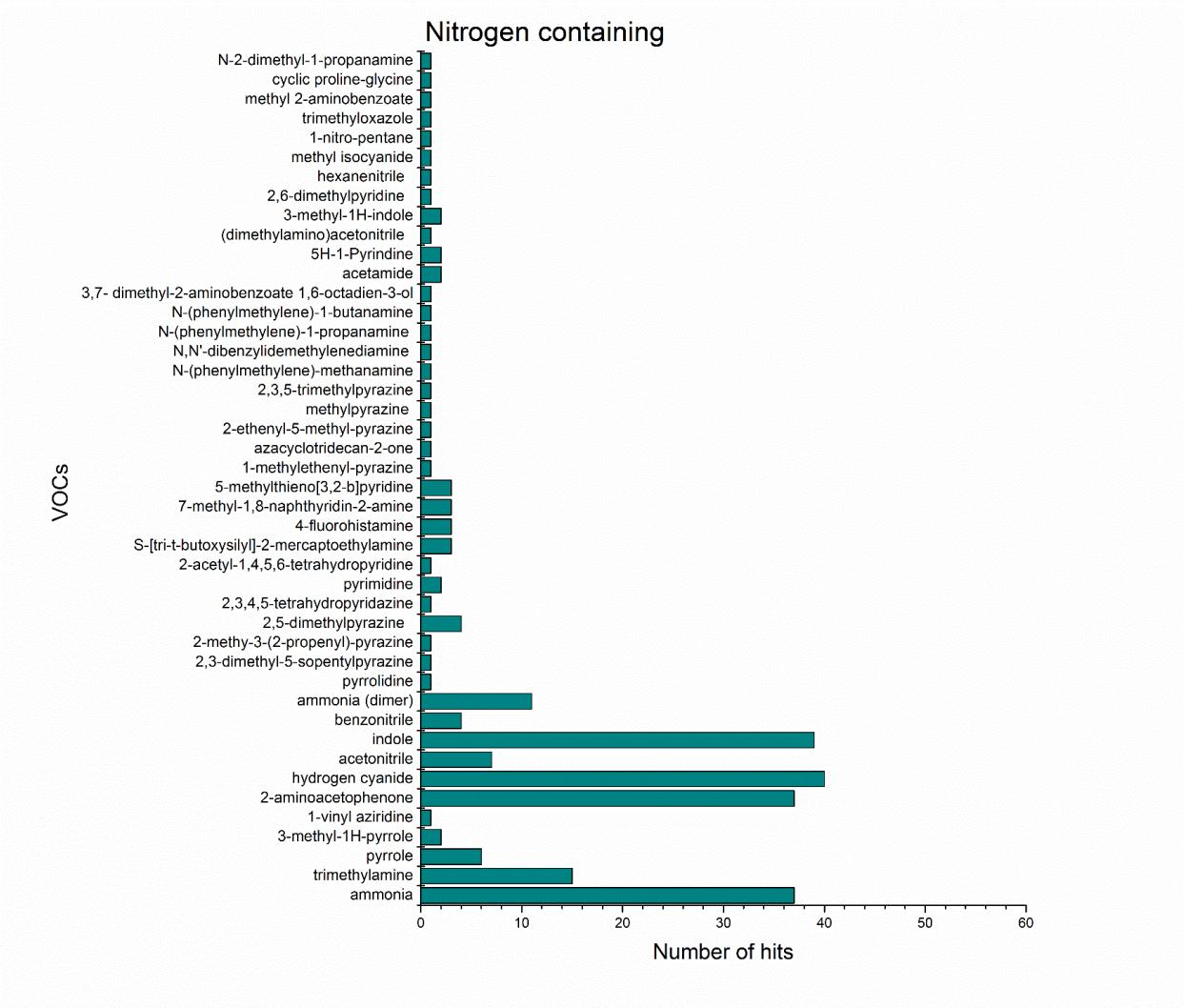

Figure A5- Representation of the number of hits for each individual VOC, in all the experiments, for the Nitrogen containing.

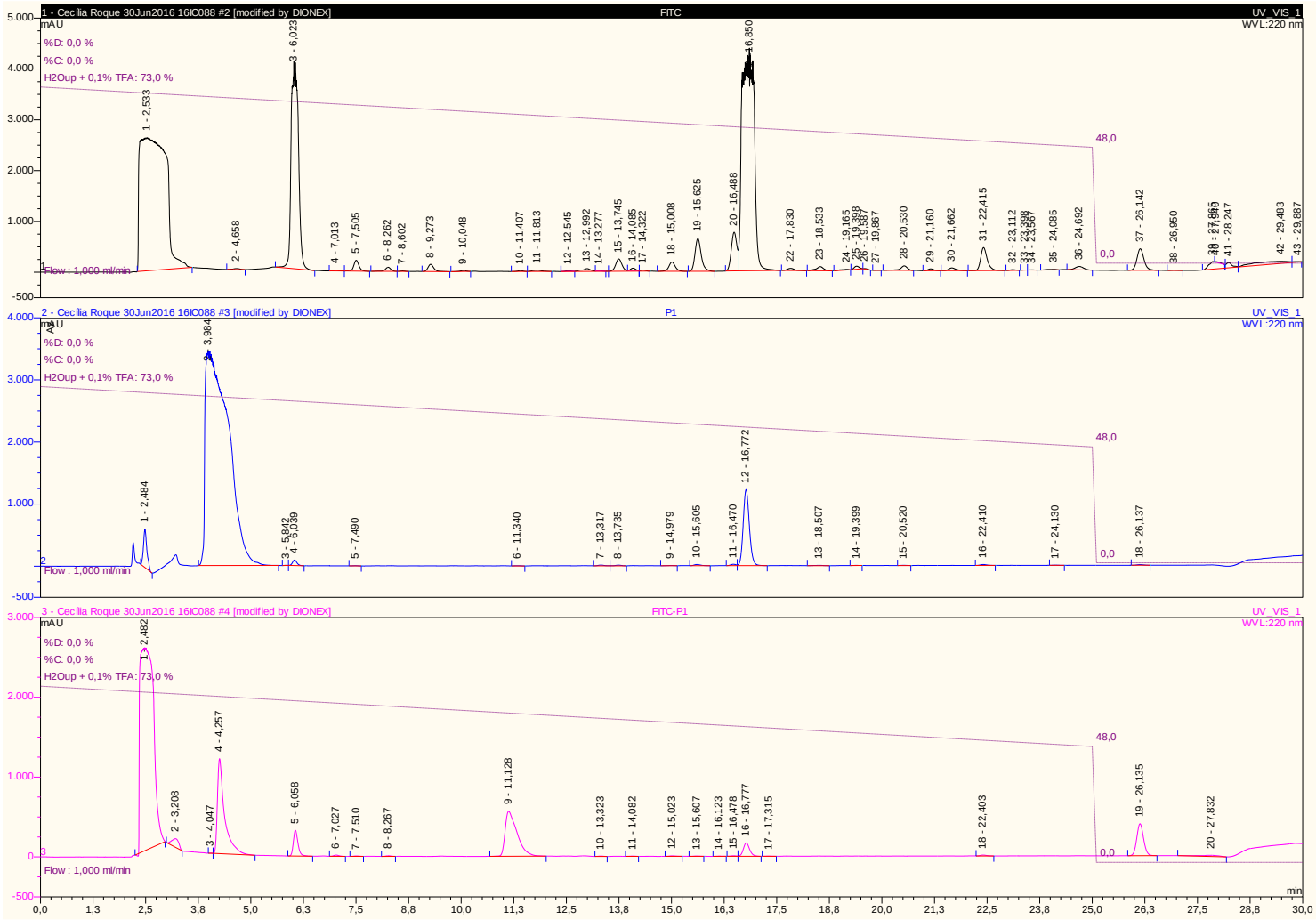

Figure A6- Chromatograms obtained for peptide P (blue), FITC (black) and P1-FITC conjugate (pink) samples.

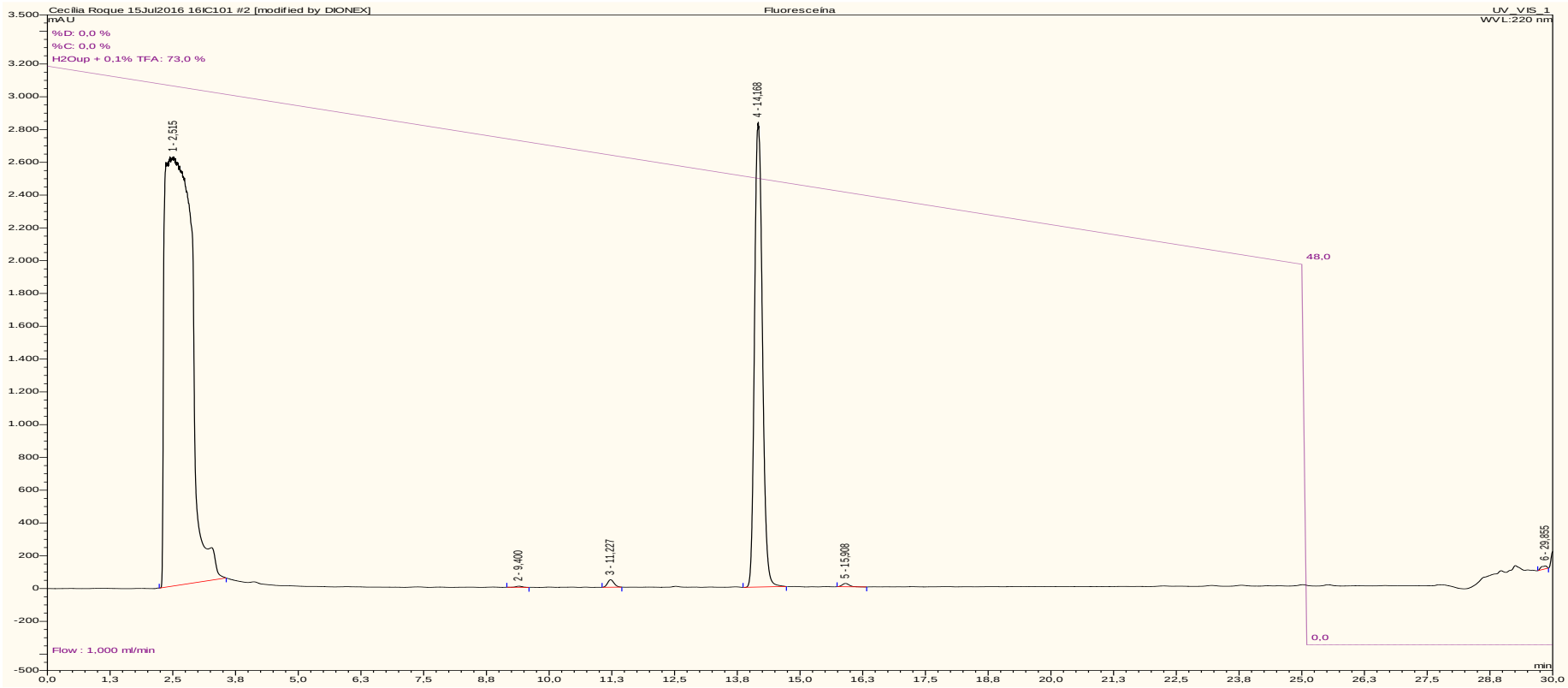

Figure A7- Chromatograms obtained for a fluorescein sample.

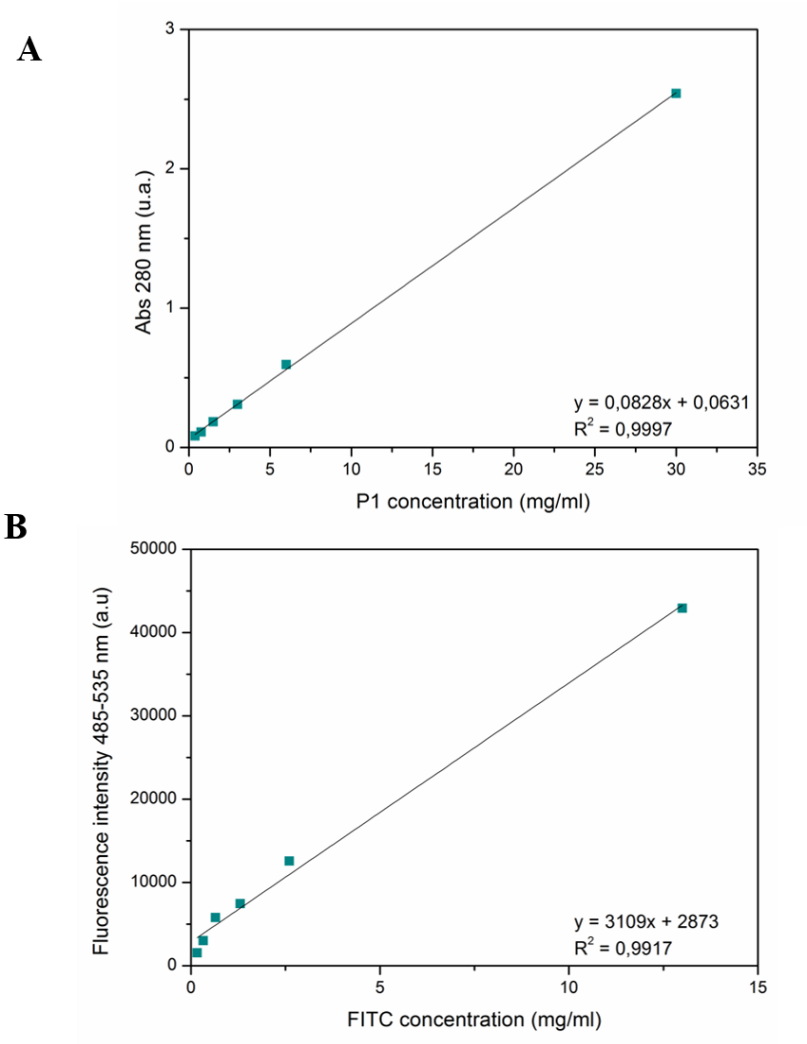

Figure A8- Calibration lines to quantify the amount of P1 (A) and FITC (B) present in the collected fraction.

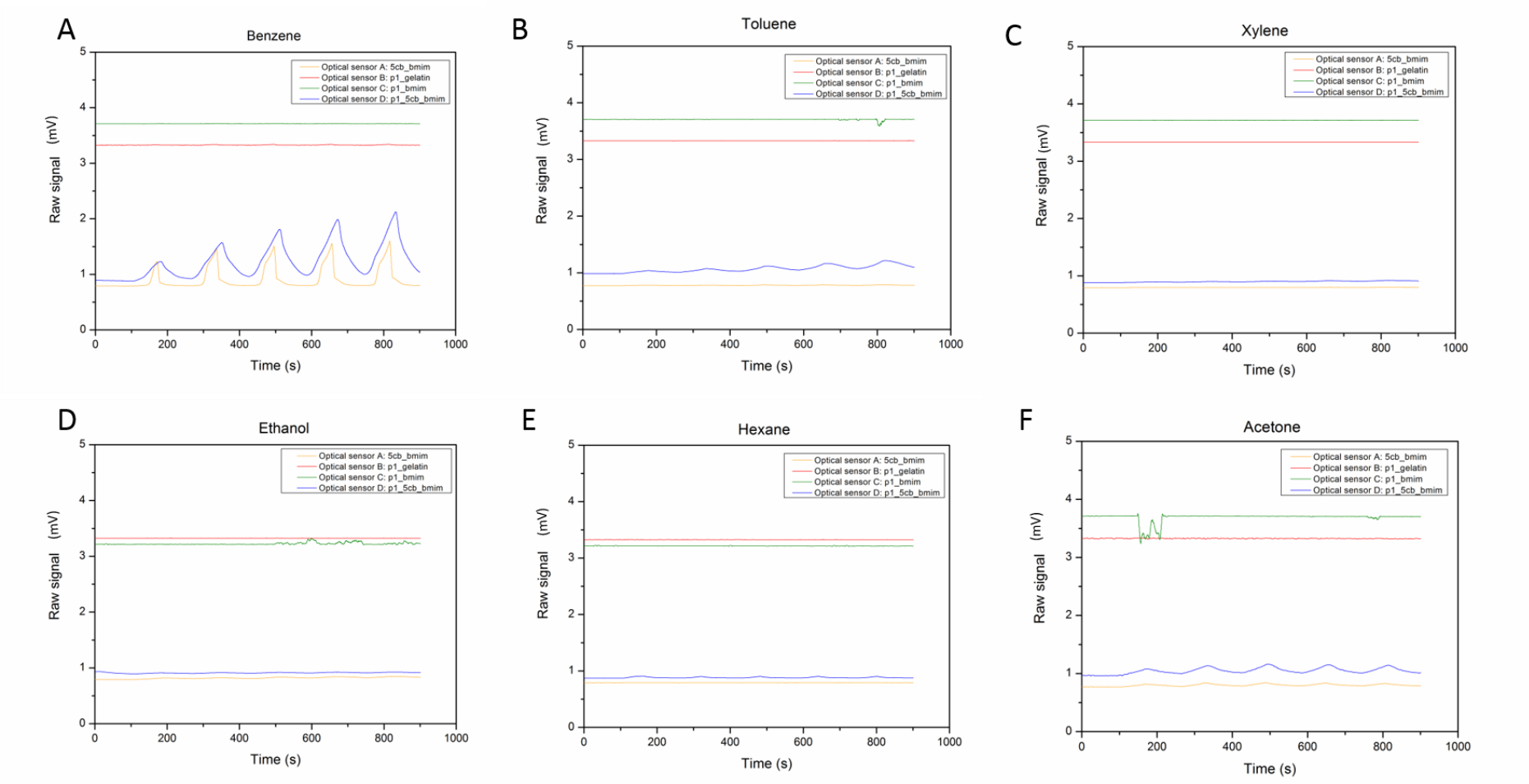

Figure A9- Responses of the biogels containing [BMIM][DCA]/5CB (yellow), gelatin/P1 (red), [BMIM][DCA]/P1 (green) and [BMIM][DCA]/5CB/P1 (blue) when exposed to benzene (A), toluene (B), xylene (C), ethanol (D), hexane (E) and acetone (F).

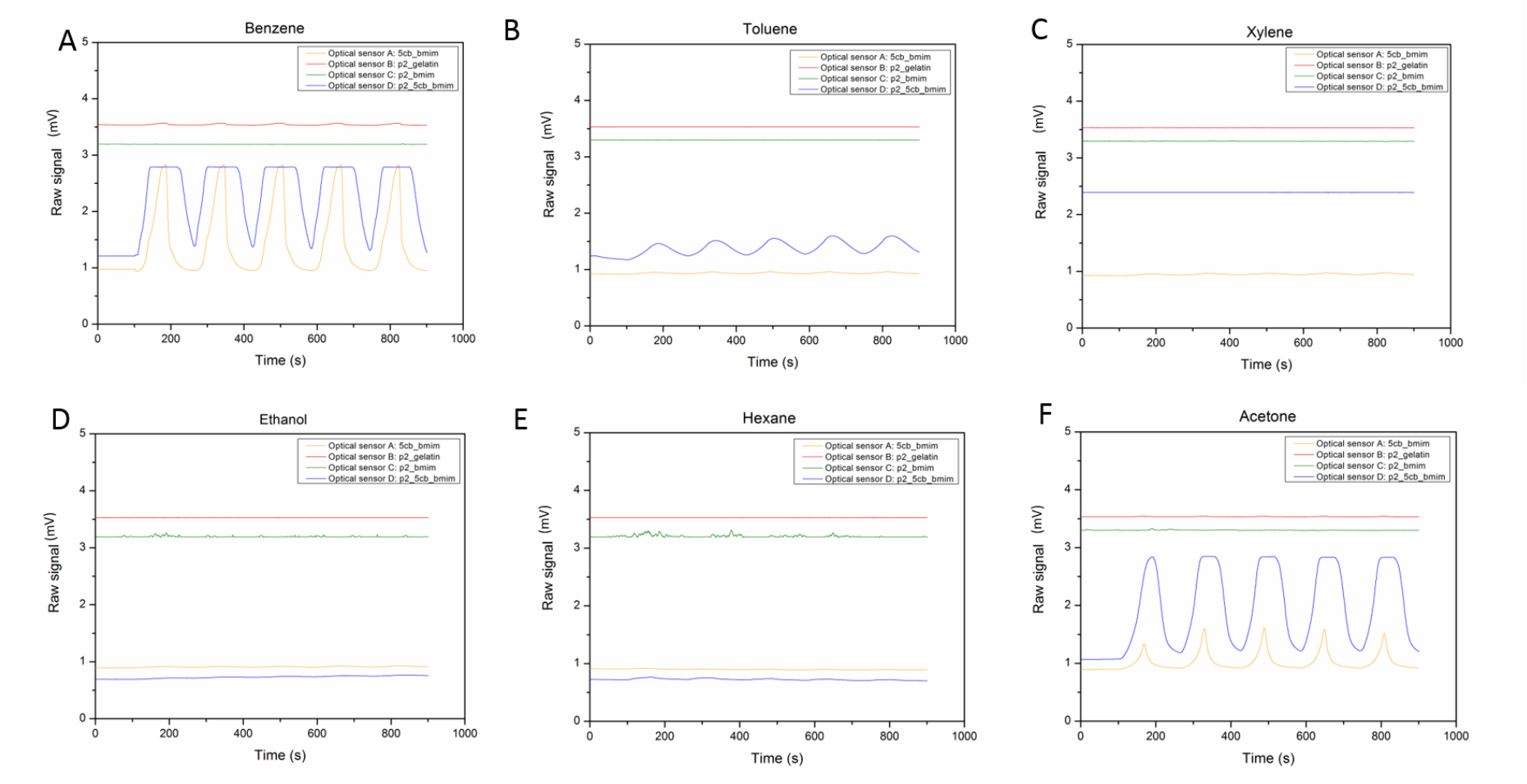

Figure A10- Responses of the biogels containing [BMIM][DCA]/5CB (yellow), gelatin/P2 (red), [BMIM][DCA]/P2 (green) and [BMIM][DCA]/5CB/P2 (blue) when exposed to benzene (A), toluene (B), xylene (C), ethanol (D), hexane (E) and acetone (F).

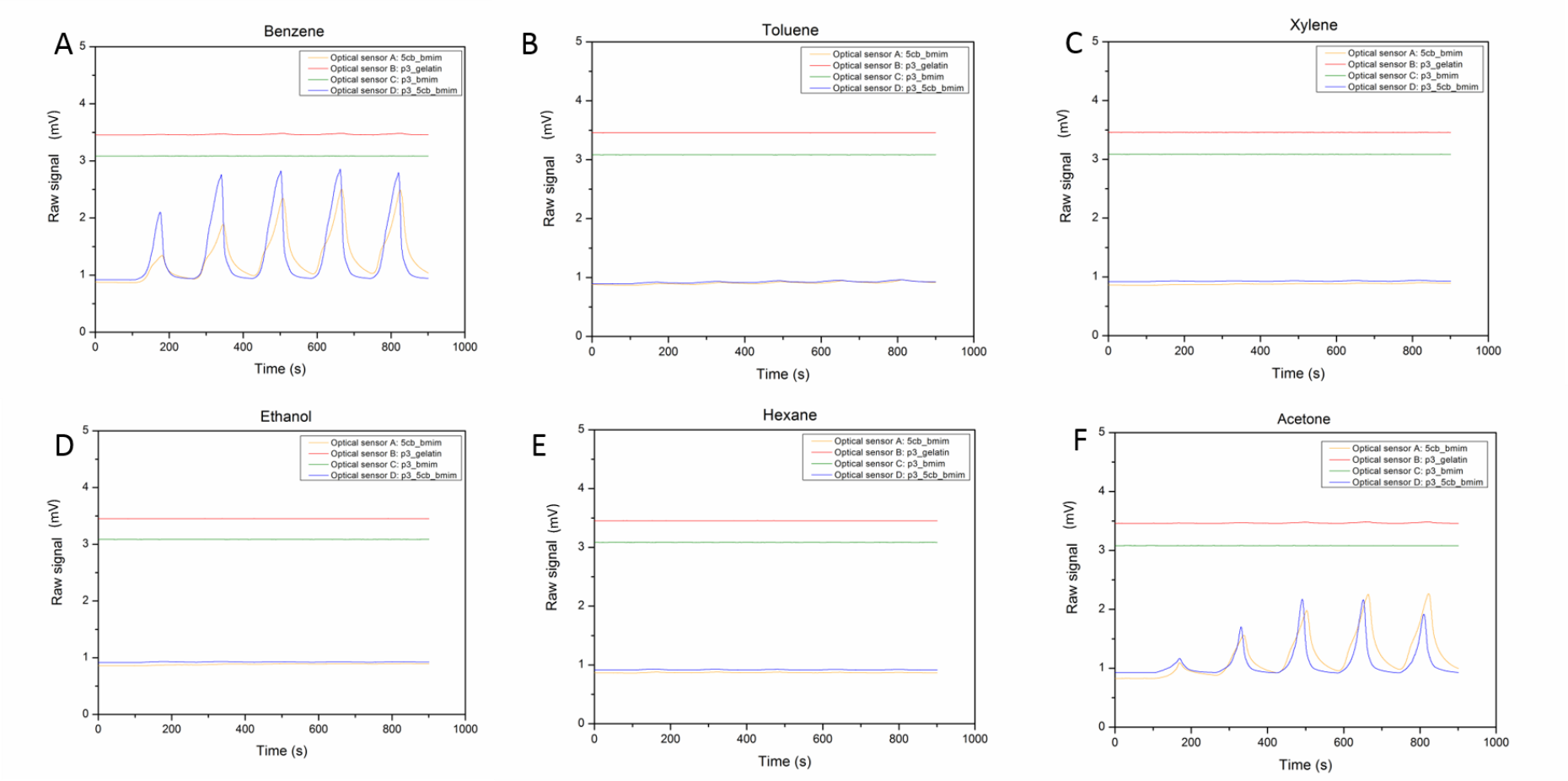

Figure A11- Responses of the biogels containing [BMIM][DCA]/5CB (yellow), gelatin/P3 (red), [BMIM][DCA]/P3 (green) and [BMIM][DCA]/5CB/P3 (blue) when exposed to benzene (A), toluene (B), xylene (C), ethanol (D), hexane (E) and acetone (F).

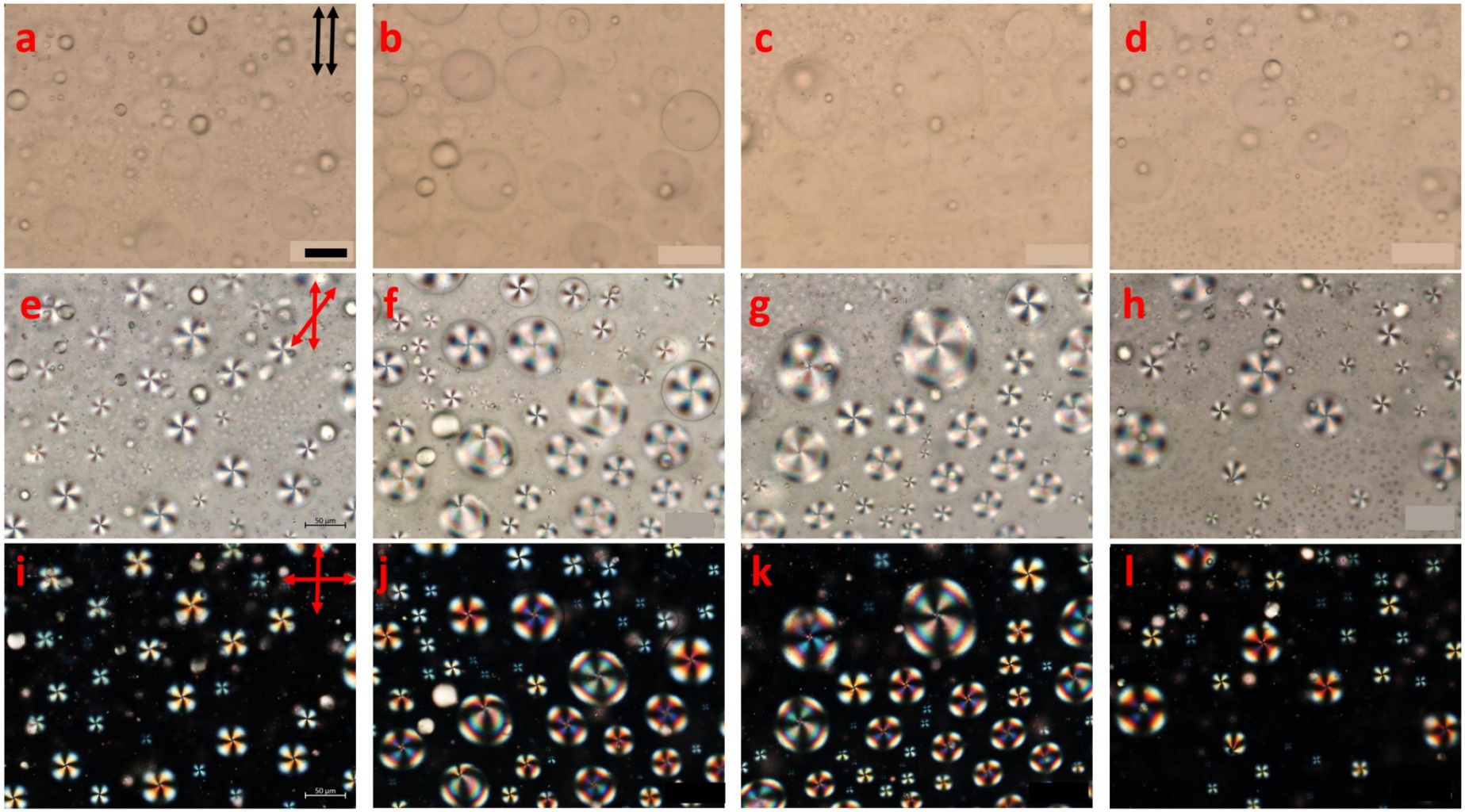

Figure A12- POM images of standard gelatin biogels ([BMIM][DCA]/5CB), in 5 different fields of view. a to d: uncrossed polarizers. e to h: semi-crossed polarizers. i to l: crossed polarizers. All images have the same scale. The scale bar corresponds to 50 µm.

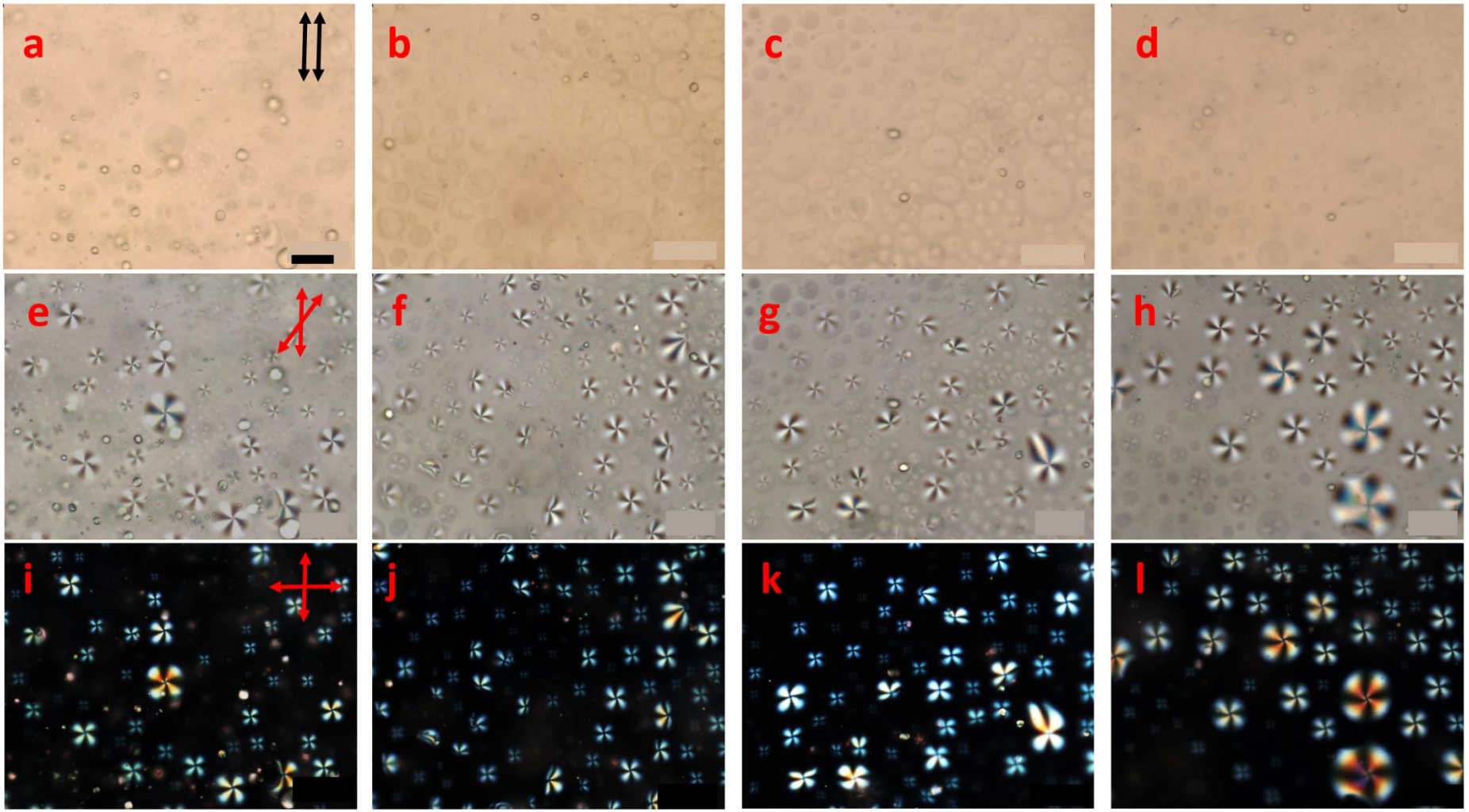

Figure A13- POM images of P1 gelatin biogels ([BMIM][DCA]/5CB/P1), in 5 different fields of view. a to d: uncrossed polarizers. e to h: semi-crossed polarizers. i to l: crossed polarizers. All images have the same scale. The scale bar corresponds to 50 μm.

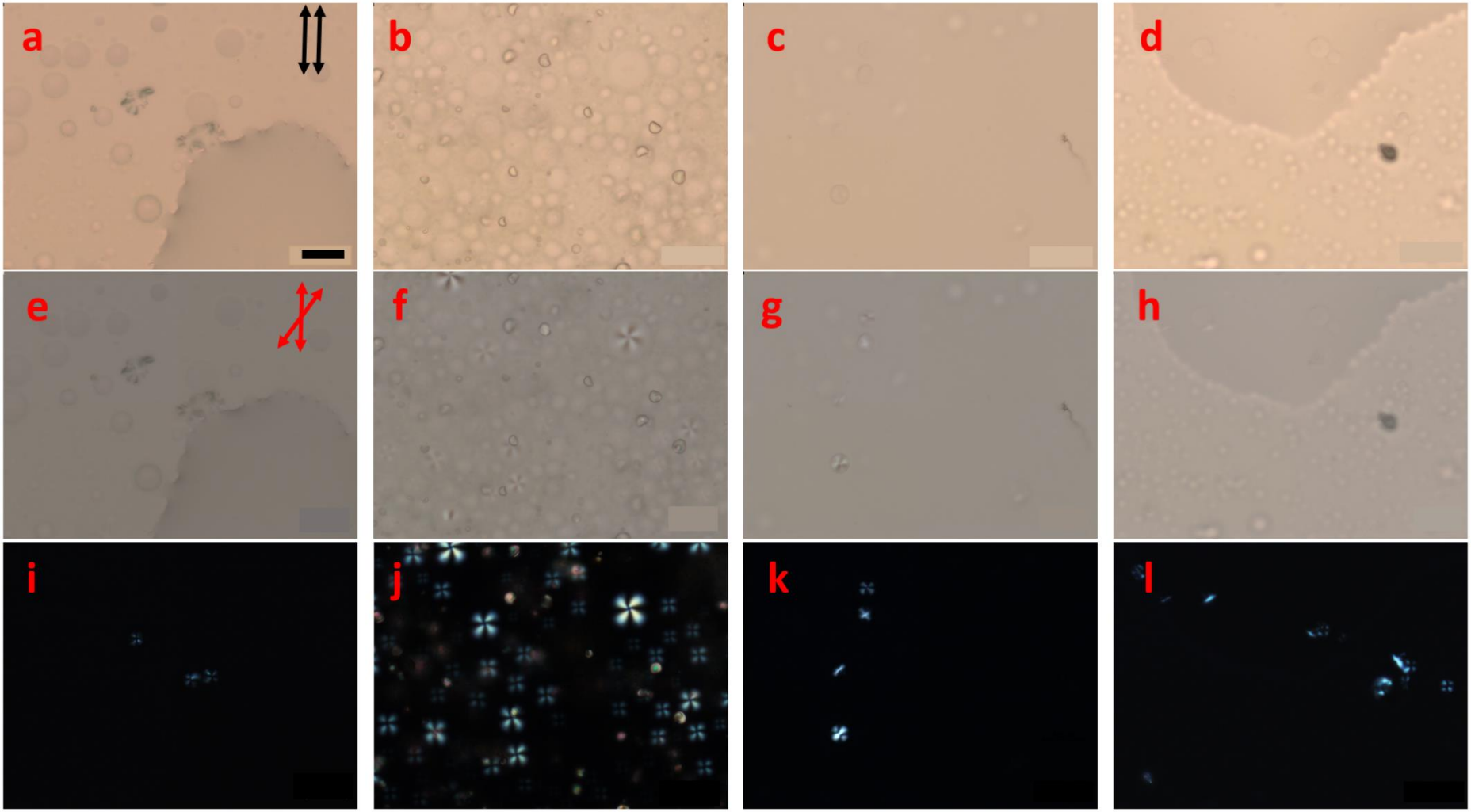

Figure A14- POM images of P2 gelatin biogels ([BMIM][DCA]/5CB/P2), in 5 different fields of view. a to d: uncrossed polarizers. e to h: semi-crossed polarizers. i to l: crossed polarizers. All images have the same scale. The scale bar corresponds to 50 µm.

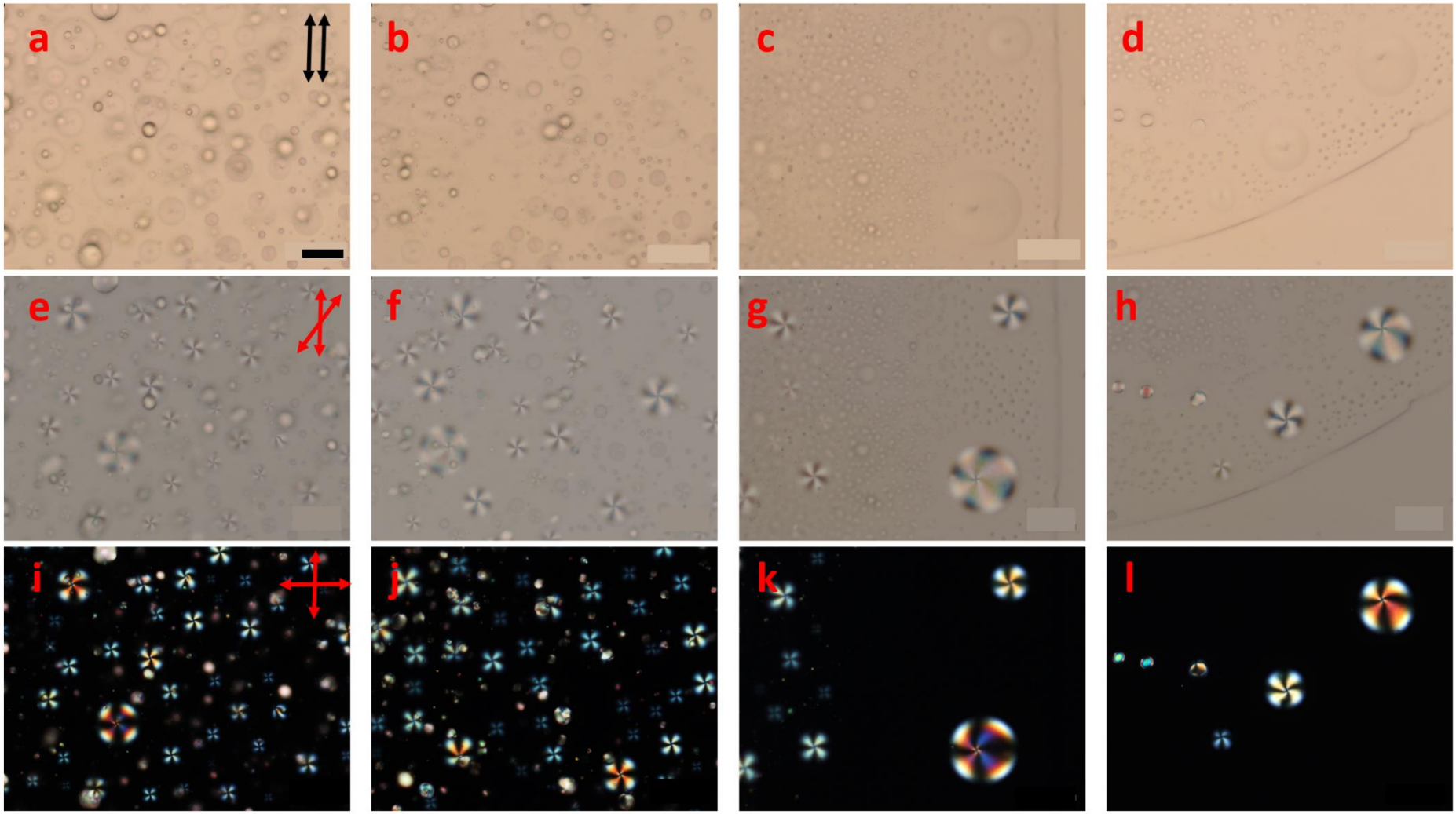

Figure A15- POM images of P3 gelatin biogels ([BMIM][DCA]/5CB/P3), in 5 different fields of view. a to d: uncrossed polarizers. e to h: semi-crossed polarizers. i to l: crossed polarizers. All images have the same scale. The scale bar corresponds to 50 µm.

Appendix 2- Pathogen- VOCs Database

Pathogen	Classification	VOCs	PubChem ID	Methods	*In vivo* Sample							Bacterial strain	Culture conditions/Growth medium	Incubation time before analysis	Concentration range		Reference	Journal Code	Year
					Saliva	Blood	Breath	Skin	Urine	Faeces	Milk				Value	Unit			
		2-pentylfuran	19602	SPME-GC-MS analysis			×					unknown	unknown	unknown	unknown	unknow	Sethi, S., Nanda, R., & Chakraborty, T. (2013).	Clin. Microbiol. Rev.	2013
		ammonia (↑)	222																
		methanol (↑)	887																
		ethanol (↑)	702																
		1-propanol (↑)	1031																
		acetone (↑)	180	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	cultured overnight on agar plates and identified	The cultures were incubated at 37 C for 72 h prior to the headspace analysis	unknown	ppb	Chippendale, T. W., Gilchrist, F. J., Spanêl, P., Alcock, A., Lenney, W., & Smith, D. (2014). Quantification by SIFT-MS of volatile compounds emitted by Aspergillus fumigatus cultures and in co-culture with Pseudomonas aeruginosa, Staphylococcus aureus and Streptococcus pneumoniae. *Analytical Methods ,6* (20), 8154-8164. DOI:	Anal. Methods	2014
		methanethiol (↑)	878																
		dimethyl sulfide (↑)	1068																
		acetaldehyde (↓)	177																
		butanal (↓)	261																
		pentanal (↓)	8063																
Aspergillus fumigatus	*fungus*	3-octanone	246728	MCC-IMS analysis	not used	not used	not used	not used	not used	not used	not used	DSM 21023	grown on columbia sheep blood agar for 24h 37ºC	unknown	unknown	unknow	Perl, T., Jünger, M., Vautz, W., Nolte, J., Kuhns, M., Borg-von Zepelin, M., & Quintel, M. (2011). Detection of characteristic metabolites of Aspergillus fumigatus and Candida species	Mycoses	2011
		isopentanol	31260																
		ethanol	702																
		cyclohexanone	7967																
		8-nonen-2-one	21108	TD-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	stored in 10% glycerol broth at −80 °C, and revived by subculturing on Sabouraud dextrose agar (SAD) supplemented with 0.02% chloramphenicol, for 2 × 7 d at 37 °C.	16h, 24h and 48h	unknown	unknow	Neerincx, A. H., Geurts, B. P., Habets, M. F. J., Booij, J. A., van Loon, J., Jansen, J. J., ... & Wevers, R. A. (2016). Identification of Pseudomonas aeruginosa and Aspergillus fumigatus mono-and co-cultures based on volatile biomarker combinations. *Journal of breath research , 10* (1), 016002.	J. Breath Res.	2016
		2,3-dihydro-1,1,3-trimethyl-3-phenyl-1-H-ind	19793																
		2-tridecanone	11622																
		1-methylethenyl-pyrazine	62897																
		azacyclotridecan-2-one	13690																
		2-ethenyl-5-methyl-pyrazine	26335																
		2-undecanone	8163																
		2-nonanone	13187																
		1-hydroxy-2-propanone	8299																
		2-acetylthiazole	520108																
		butanoic acid (↓)	264	SPME-GC-MS analysis						×		unknown	unknown	unknown	unknown	unknow	Garner, C. E., Smith, S., de Lacy Costello, B., White, P., Spencer, R., Probert, C. S., & Ratcliffe, N. M. (2007). Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. The FASEB Journal, 21(8), 1675-1688.	FASEB J.	2007
		2-methylbutanoic acid (↓)	8314																
		carbon disulfide (↓)	6348																
		dimethyl sulfide (↓)	1068																
		dimethyl disulfide (↓)	12232																
		indole (↓)	798																
		4-methylphenol (↓)	2879																
		3-methylfuran (↓)	13587																
		3-methylbutanoic acid (↓)	10430																
		dimethyl trisulfide (↓)	19310																
		limonene (↓)	22311																
		methanethiol (↓)	878																
		propanal (↓)	527																
		methyl acetate (↓)	6584																
		2-hexanone (↓)	11583																
		cymol (↓)	7463																
		2-phenylacetaldehyde (↓)	998																
		methyl butyrate (↓)	12180																
		methyl propionate (↓)	11124																
		2,3-pentanedione (↑)	11747																
		phenol (↑)	996																
		2-propanol (↓)	3776																
		2-methylpropanoic acid (↓)	6590																
		3-methyl-1H-indole (↓)	6736																
		isobutanol (↓)	6560																
		3-carene (↓)	26049																
		heptanoic acid (↑)	8094																
		methyl isobutyl ketone (↓)	7909																
		isopentanol (↓)	31260																
		2-phenylethanol (↑)	6054																
		1-propanol (↑)	1031																
		acetophenone (↑)	7410																
		2,3-hexanedione (↓)	19707																
		caryophyllene (↓)	5322111																
		n-propyl acetate (↑)	7997																
		styrene (↑)	7501																
		toluene (↑)	1140																
		pentanal (↑)	8063																
		2-(2-ethoxyethoxy)-ethanol (↑)	8146																
		2-butanol (↓)	6568																
		n-pentane (↓)	8003																
		2-pentylfuran (↑)	19602																
		ethyl pentanoate (↓)	10882																
		1-octanol (↑)	957																
		2-ethylhexanoic acid (↑)	8697																
		methanol (↑)	887																
		3-penten-2-one (↑)	12248																
		benzoic acid (↑)	243																
		ethylbenzene (↑)	7500																
		linalool (↑)	6549																
		heptanal (↑)	8130																
		phenylmethanol (↑)	244																
		octanal (↑)	454																
		3-heptanone (↓)	7802																
		methyl salicylate (↓)	4133																
		ethenyl hexanoate (↑)	76451																
		2-ethyl-1-hexanol (↑)	7720																
		1-nonanol (↓)	8914																
		2-butenal (↑)	447466																
		5-methyl-2-hexanone (↓)	8034																
		2-piperidinone (↑)	12665																
		6-methyl-3,5-heptadien-2-one (↑)	5370101																
		allyl isothiocyanate (↓)	5971																
Campylobacter jejuni	Gram negative bacterium	dodecane (↑)	8182																
		nonanal (↑)	31289																
		propyl propanoate (↑)	7803																
		p-xylene (↑)	7809																
		acetamide (↓)	178																
		3-methyl-2-butenal (↑)	61020																
		2-butylfuran (↑)	20534																
		decanal (↑)	8175																
		ethyl hexanoate (↑)	31265																
		1-octen-3-ol (↑)	18827																
		2-acetyl-5-methylfuran (↑)	14514																
		3-octen-2-one (↑)	15475																
		cyclohexanecarboxylic acid (↑)	7413																
		5-ethyldihydro-2(3H)-furanone (↑)	12756																
		isobutyl acetate (↓)	8038																
		acetonitrile (↑)	6342																
		4-methylbenzaldehyde (↑)	7725																
		isoamyl butyrate (↑)	7795																
		propyl hexanoate (↑)	12293																
		methacrolein (↑)	6562																
		naphthalene (↑)	931																
		2,6-dimethylpyridine (↑)	7937																
		trichloromethane (↑)	6212																
		benzonitrile (↑)	7505																
		3-methylhexanal (↑)	140511																
		hexanenitrile (↑)	12352																
		propyl isobutyrate (↓)	12571																
		4-methyl-3-penten-2-one (↑)	8858																
		3-(methylthio)-1-propanol (↓)	10448																
		pyrimidine (↑)	9260																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1						*In vivo* Sample										Concentration range				
2	Pathogen	Classification	VOCs	PubChem ID	Methods	Saliva	Blood	Breath	Skin	Urine	Faeces	Milk	Bacterial strain	Culture conditions/Growth medium	ıcubation time before analy	Value	Unit	Reference	Journal Code	Year
151			o-xylene (↑)	7237																
152			dihydro-5-propyl-2(3H)-furanone (↑)	7742																
153			furan (↑)	8029																
154			2-heptene (↑)	11611																
155			2-hexenal (↑)	5281168																
156			1-butoxy-2-propanol (↑)	21210																
157			2-propenol (↑)	7858																
158			acrolein (↑)	7847																
159			5H-1-Pyrindine (↑)	575987																
160			benzyl acetate (↑)	8785																
161			alpha-farnesene (↑)	5281516																
162			1,2,3-trimethylbenzene (↑)	10686																
163			propylbenzene (↑)	7668																
164			ethyl isopropyl ketone (↑)	11265																
165			m-xylene (↑)	7929																
166			ethyl isovalerate (↑)	7945																
167			propyl isovalerate (↑)	11176																
168			cis-1-p-menthanol (↑)	89437																
169			2-butoxyethanol (↑)	8133																
170			2-ethoxyethanol (↑)	8076																
171			2-(formyloxy)-1-phenyl-ethanone (↑)	569595																
172			2,5-dimethylfuran (↑)	12266																
173			2-ethylhexanal (↑)	31241																
174			methyl isocyanide (↑)	11646																
175			methyl tiglate (↑)	5323652																
176			methyl vinyl ketone (↑)	6570																
177			2-methylnaphthalene (↑)	7055																
178			1-nitro-pentane (↑)	220639																
179			ethyl pentanoate (↑)	10882																
180			propyl phenylacetate (↑)	221641																
181			propyl octanoate (↑)	69351																
182			2,5-dimethylpyrazine (↑)	31252																
183			isoamyl isovalerate (↑)	12613																
184			trimethylamine (↑)	1146																
185			phenylacetic acid (↑)	999																
186			ethanol	702																
187	*Candida albicans*	Fungus	formaldehyde	712	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	CDC S-24	sterile urine (20 mL) from healthy males inoculated to a concentration of between 10^7 and 10^9 cfu/mL	37˚C for 6h	unknown	ppb	Storer, M. K., Hibbard-Melles, K., Davis, B., & Scotter, J. (2011). Detection of volatile compounds produced by microbial growth in	J. Microbiol. Methods	2011
188			methanethiol	878																
189			methyl butyrate	12180																
190			methanol	887																
191			methyl propionate	11124																
192			methyl pentanoate	12206																
193			caryophyllene	5322111																
194			1-methyl-2-(1-methylethyl)-benzene	10812																
195			propyl butyrate	7770																
196			butyl butyrate	7983																
197			methyl hexanoate	7824																
198			terpinene	7462																
199			copaene	19725																
200			n- butyl acetate	31272																
201			isoamyl butyrate	7795																
202			isobutyl butyrate	10885																
203			butyl propionate	11529																
204			ethyl cyclohexanecarboxylate	18686																
205			bicyclo [3,1,1] 6,6-dimethyl-2-methylene-hep	440967																
206			terpinolene	11463																
207			methyl cyclohexanecarboxylate	20748																
208			propyl isovalerate	11176																
209			amyl acetate	12348																
210			butyl pentanoate	61137	GC-MS analysis						×		unknown	unknown	From each sample, a 2 gm aliquot was placed into a 18 ml glass vial (Sigma Aldridge), sealed with a silicone/polytetraflouroeth ylene septum, within 6 hours of sample production and were frozen at -20˚C until analyzed.	unknown	unknow	Ahmed, I., Greenwood, R., Costello Bde, L., Ratcliffe, N. M., & Probert, C. S. (2013). An investigation of fecal volatile organic metabolites in irritable bowel syndrome. *PloS one*, *8* (3), e58204. doi:10.1371/journal.pone.0058204.	PLoS One	2013
211			2-hexanone	11583																
212			alfa-pinene	6654																
213			cyclohexanecarboxylic acid	7413																
214			6-methyl-5-hepten-2-one	9862																
215			methyl 2-methylbutyrate	13357																
216			methyl isobutyrate	11039																
217			2-butanol	6568																
218			isobutyl propionate	10895																
219			3,7- dimethyl-2-aminobenzoate 1,6-octadien	23535																
220			5-methyl-2-(1-methylethyl)-cyclohexanol	1254																
221			alfa-myrcene	519324																
222			butyl cyclohexanecarboxylate	81040																
223			terpinene	7462																
224			methyl salicylate	4133																
225			methyl heptanoate	7826																
226			methyl isovalerate	11160																
227			2-methylbutyl 2-methylbutyrate	17129																
228			5-methyl-2-(1-methylethyl)-cyclohexanone	6986																
229			isopropyl butyrate	61184																
230			isobutyl pentanoate	66356																
231			methyl allyl disulfide	62434																
232			hexyl propionate	88454																
233			6-methyl-5-hepten-2-one (↓)	9862																
234			acetaldehyde (↓)	177																
235			acetic acid (↓)	176																
236			butanoic acid (↓)	264																
237			2-methylbutanoic acid (↓)	8314																
238			carbon disulfide (↓)	6348																
239			dimethyl sulfide (↓)	1068																
240			dimethyl disulfide (↓)	12232																
241			pentanoic acid (↓)	7991																
242			3-methylfuran (↓)	13587																
243			propanoic acid (↓)	1032																
244			3-methylbutanoic acid (↓)	10430																
245			dimethyl trisulfide (↓)	19310																
246			hexanoic acid (↓)	8892																
247			limonene (↓)	22311																
248			methanethiol (↓)	878																
249			propanal (↓)	527																
250			methyl acetate (↓)	6584																
251			butanal (↓)	261																
252			ethanol (↓)	702																
253			1-pentanol (↓)	6276																
254			2-methylbutanal (↓)	7284																
255			2-hexanone (↓)	11583																
256	Clostridium difficile	Gram positive bacterium	2-phenylacetaldehyde (↓)	998																
257			methyl butyrate (↓)	12180																
258			ethyl acetate (↓)	8857																
259			methyl propionate (↓)	11124																
260			2-methylpropanoic acid (↓)	6590																
261			3-methyl-1H-indole (↓)	6736																
262			ethyl butanoate (↓)	7762																
263			methyl isobutyl ketone (↓)	7909																

Pathogen	Classification	VOCs	PubChem ID	Methods	*In vivo* Sample							Bacterial strain	Culture conditions/Growth medium	Incubation time before analysis	Concentration range		Reference	Journal Code	Year
					Saliva	Blood	Breath	Skin	Urine	Faeces	Milk				Value	Unit			
		1-propanol (↓)	1031																
		2-acetylfuran (↓)	14505																
		styrene (↑)	7501																
		toluene (↓)	1140																
		pentanal (↓)	8063																
		n- butyl acetate (↓)	31272																
		propyl butyrate (↓)	7770																
		methyl pentanoate (↓)	12206																
		2-butanol (↓)	6568																
		ethyl propionate (↓)	7749																
		2-octanone (↓)	8093																
		methyl thioacetate (↓)	73750																
		2-pentylfuran (↓)	19602																
		methyl hexanoate (↓)	7824																
		ethyl pentanoate (↓)	10882	SPME-GC-MS analysis						×		unknown	unknown	fresh or 7 days	unknown	unknown	Garner, C. E., Smith, S., de Lacy Costello, B., White, P., Spencer, R., Probert, C. S., & Ratcliffe, N. M. (2007). Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. The FASEB Journal, 21(8), 1675-1688.	FASEB J.	2007
		2-ethylhexanoic acid (↓)	8697																
		methanol (↓)	887																
		butyl propionate (↓)	11529																
		3-penten-2-one (↑)	12248																
		benzoic acid (↑)	243																
		phenylmethanol (↑)	244																
		undecane (↓)	14257																
		2,3,5-trithiahexane (↓)	93236																
		2-ethyl-1-hexanol (↑)	7720																
		dodecane (↓)	8182																
		2-butenal (↓)	447466																
		acetamide (↓)	178																
		2-furaldehyde (↓)	7362																
		2-nonanone (↓)	13187																
		3-methyl-2-butenal (↑)	61020																
		1-octen-3-ol (↑)	18827																
		4-heptanone (↑)	31246																
		methacrolein (↓)	6562																
		ethyl isobutyrate (↑)	7342																
		benzonitrile (↑)	7505																
		isoamyl acetate (↑)	31276																
		4-methyl-1-pentanol (↑)	12296																
		1-penten-3-ol (↑)	12020																
		3-methyl-2-butenol (↑)	11173																
		3-methyl-2-cyclohexenone (↑)	14511																
		6-methyl-2-heptanone (↑)	13572																
		(Z)-2-heptenal (↑)	5362616																
		2-hexenal (↑)	5281168																
		(E)-3-octenal (↑)	5283325																
		(Z)-2-pentenol (↑)	5364919																
		acrolein (↑)	7847																
		5H-1-Pyrindine (↑)	575987																
		4-methylphenyl acetate (↑)	8797																
		(dimethylamino)acetonitrile (↑)	61237																
		methyl isovalerate (↑)	11160																
		propyl isovalerate (↑)	11176																
		cis-1-p-menthanol (↑)	89437																
		cyclohexanone (↑)	7967																
		cyclopentane (↑)	9253																
		2-(methylthio)-ethanol (↑)	78925																
		4-methylpentanoic acid methyl ester (↑)	17008																
		propyl pentanoate (↑)	67328																
		2-methoxyphenol (↑)	460																
		2-methoxy-4-methylphenol (↑)	7144																
		2,5-dimethylpyrazine (↑)	31252																
		acetone	180																
		2-butanone	6569																
		2-pentanone	7895																
		formaldehyde	712																
		2-methylbutanal	7284																
		ethyl butanoate	7762	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	NTCC 775	sterile urine (20 mL) from healthy males inoculated to a concentration of between 10^7 and 10^9 cfu/mL	37˚C for 6h	unknown	ppb	Storer, M. K., Hibbard-Melles, K., Davis, B., & Scotter, J. (2011). Detection of volatile compounds produced by microbial growth in urine by selected ion flow tube mass spectrometry (SIFT-MS). Journal of microbiological methods, 87(1), 111-113. doi: 10.1016/j.mimet.2011.06.012.	J. Microbiol. Methods	2011
		n-propyl acetate	7997																
		hydrogen sulfide	402																
		dimethyl sulfide	1068																
		dimethyl disulfide	12232																
		methanethiol	878																
		ammonia	222																
		propene	8252																
		1-butanol	263																
		1-propanol	1031																
		1-pentanol	6276																
		phenylacetic acid	999																
		formaldehyde	712																
		2-butanone	6569																
		2-pentanone	7895	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	Bos, L. D., Sterk, P. J., & Schultz, M. J. (2013). Volatile metabolites of pathogens: a systematic review. doi:10.1371/journal.ppat.1003311	PLOS	2013
Enterococcus faecalis	Gram positive bacterium	acetone	180																
		ethyl butanoate	7762																
		n-propyl acetate	7997																
		dimethyl sulfide	1068																
		hydrogen sulfide	402																
		methanethiol	878																
		pyrrole	8027																
		1-butanol	263												146.13				
		1-pentanol	6276												14.06				
		2-aminoacetophenone	11952												1.34				
		acetone	180												55.07				
		dimethyl disulfide	12232												17.72				
		ethanol	702												365.59				
		ethyl butanoate	7762	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	blood agar	48h	4.13	ppb	Thorn, R., Reynolds, D. M. and Greenman, J. (2011) Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro.Journal of Microbiological Methods, 84 (2). pp. 258-264. ISSN 0167-7012 DOI: 10.1016/j.mimet.2010.12.001	J. Microbiol. Methods	2011
		formaldehyde	712												335.86				
		hydrogen sulfide	402												148.46				
		isoprene	6557												6.58				
		methanethiol	878												334.37				
		phenylacetic acid	999												41.93				
		pyrrole	8027												0.8				
		trimethylamine	1146												44.28				
		acetone	180																
		acetic acid	176																
		methanol	887																
		ethanol	702																
		formaldehyde	712																
		ethyl acetate	8857																
		ethyl butanoate	7762																
		n-propyl acetate	7997	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	W310	sterile urine (20 mL) from healthy males inoculated to a concentration of between 10^7 and 10^9 cfu/mL	37˚C for 6h	unknown	ppb	Storer, M. K., Hibbard-Melles, K., Davis, B., & Scotter, J. (2011). Detection of volatile compounds produced by microbial growth in urine by selected ion flow tube mass spectrometry (SIFT-MS). *Journal of microbiological methods* , *87* (1), 111-113. doi: 10.1016/j.mimet.2011.06.012.	J. Microbiol. Methods	2011
		hydrogen sulfide	402																
		dimethyl disulfide	12232																
		dimethyl sulfide	1068																
		methanethiol	878																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
						In vivo Sample										Concentration range				
1	Pathogen	Classification	VOCs	PubChem ID	Methods	Saliva	Blood	Breath	Skin	Urine	Faeces	Milk	Bacterial strain	Culture conditions/Growth medium	ncubation time before analys	Value	Unit	Reference	Journal Code	Year
347			trimethylamine	1146																
348			2-aminoacetophenone	11952																
349			indole	798																
350			ethanol	702																
351			1-pentanol	6276																
352			formaldehyde	712																
353			acetaldehyde	177																
354			acetic acid	176																
355			hydrogen sulfide	402																
356			methanethiol	878																
357			dimethyl sulfide	1068	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	ATCC 25922	blood culture bottles	24 h	unknown	unknown	Allardyce, Randall A., et al. "Detection of volatile metabolites produced by bacterial growth in blood culture media by selected ion flow tube mass spectrometry (SIFT-MS)." *Journal of microbiological methods* 65.2 (2006): 361-365.	J. Microbiol. Methods	2006
358			dimethyl disulfide	12232																
359			trimethylamine	1146																
360			indole	798																
361			1-propanol	1031																
362			2-aminoacetophenone	11952																
363			hexanal	6184																
364			carbon dioxide	280																
365			ammonia	222	IMR-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	DH5 5678	blood agar plates	24 h	unknown	unknown	Dolch, M. E., et al. "Volatile compound profiling for the identification of Gram-negative bacteria by ion-molecule reaction–mass spectrometry." *Journal of*	J. Appl. Microbiol.	2012
366			methanethiol	878																
367			indole	798																
368			1-butanol	263												173.38				
369			1-pentanol	6276												47.69				
370			acetoin	179												27.26				
371			butanoic acid	264												40.86				
372			ethanol	702												212.95				
373			ethyl acetate	8857												120.97				
374			ethyl butanoate	7762									NCTC 10418			18.35				
375			formaldehyde	712												555.63				
376			hydrogen sulfide	402												318.21				
377			indole	798												551.51				
378			isoprene	6557												101.01				
379			methanethiol	878												566.71				
380			phenylacetic acid	999												7.28				
381			trimethylamine	1146	SIFT-MS analysis	not used	not used	not used	not used	not used	not used	not used		nutrient agar; incubated at 37˚C aerobically when required	24 h	131.92	ppb	Thorn, R., Reynolds, D. M. and Greenman, J. (2011) Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro.Journal of Microbiological Methods, 84 (2). pp. 258-264. ISSN 0167-7012 DOI: 10.1016/j.mimet.2010.12.001.	J. Microbiol. Methods	2011
382			1-butanol	263												123.04				
383			1-pentanol	6276												26.66				
384			2-aminoacetophenone	11952												1.95				
385			acetoin	179												20.2				
386			butanoic acid	264												29.65				
387			ethanol	702												216.13				
388			ethyl acetate	8857									NCTC 12900			91.71				
389			formaldehyde	712												530.19				
390			hydrogen sulfide	402												362.26				
391			indole	798												419.61				
392			methanethiol	878												572.05				
393			phenylacetic acid	999												46.68				
394			trimethylamine	1146												133.58				
395			isoprene	6557																
396			1-propanol	1031																
397			3-methylbutanal	11552																
398			2-methylbutanal	7284																
399			2,3,3-trimethylpentane	11215																
400			benzaldehyde	240	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 25922	blood agar plates;incubated overnight at 37˚C; transfer to sterile Brain Heart Infusion broth, growth for 4h with constant agitation at 37˚C	unknown	unknown	unknown	Boots AW, Smolinska A, van Berkel JJ, Fijten RR, Stobberingh EE, et al. (2014) Identification of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography-mass spectrometry. J Breath Res 8: 027106. doi:10.1088/1752-7155/8/2/027106.	J. Breath Res.	2014
401			acetic acid	176																
402			2,3-butanedione	650																
403			n-propyl acetate	7997																
404			indole	798																
405			3-methyl-4-(1-methylethenyl)cyclohexane	14299																
406			acetonitrile	6342	SESI-MS analysis	not used	not used	not used	not used	not used	not used	not used	unknown	unknown	unknown	unknown	unknown	Sohrabi M, Zhang L, Zhang K, Ahmetagic A, Wei MQ (2014) Volatile Organic Compounds as Novel Markers for the Detection of Bacterial	J. Clin. Microbiol.	2014
407			ethanol	702																
408			indole	798																
409			ethanol	702										The strain carried the plasmid pLB4 encoding the acetone synthesis pathway of Clostridium acetobutylicum ATCC 824; Neidhardt minimal salt medium was used with 3-(N-morpholino)propanesulfonic acid (MOPS) as a buffering						
410			acetone	180	MCC-IMS analysis	not used	not used	not used	not used	not used	not used	not used	BL21 pLB4		3, 4 and 6h	unknown	unknown	Maddula, S., Blank, L. M., Schmid, A., & Baumbach, J. I. (2009). Detection of volatile metabolites of Escherichia coli by multi capillary column coupled ion mobility	Anal. Bioanal. Chem.	2009
411			2-nonanone	13187																
412			2-heptanone	8051																
413			1-octanol	957																
414			1-decanol	8174																
415			1-dodecanol	8193	HS-SPME-GC-FID analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	50 mL TS broth in 125 mL vial. Samples incubated with sha	unknown	unknown	unknown			
416			2-undecanone	8163																
417			tridecen-2-one	53427438																
418			ethanol	702																
419			1-propanol	1031																
420			isopentanol	31260																
421			1-octanol	957																
422			9-decenol	25612	HS-SPME-GC-FID/GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	10 mL TS broth. Samples incubated without shaking for 18	unknown	unknown	unknown	Tait, E., Perry, J. D., Stanforth, S. P., & Dean, J. R. (2014). Identification of volatile organic compounds produced by bacteria using HS-SPME-GC–MS. Journal of chromatographic science, 52(4), 363-373. doi:10.1093/chromsci/bmt042	J. Chromatogr. Sci.	2014
423			1-decanol	8174																
424			indole	798																
425			1-dodecanol	8193																
426			(Z)-7-tetradecen-1-ol	5362795																
427			1-tetradecanol	8209																
428			dimethyl disulfide	12232	HS-SPME-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	Super broth (tryptone, yeast, NaCl and NaOH); 5 mL culture medium in 20 mL vial; samples incubated without shaking for 18h at 35ºC.	unknown	unknown	unknown			
429			ethanol	702																
430			2-nonanone	13187																
431			2-heptanone	8051																
432			pentyl cyclopropane	75640																
433			indole	798																
434			ethanol (↑)	702	MCC-IMS analysis	not used	not used	not used	not used	not used	not used	not used								
435			indole (↑)	798																
436	*Escherichia coli*	Gram negative bacterium	2-(methylthio)-ethanol (↑)	78925																
437			3-methylbutanal (↑)	11552																
438			dimethyl disulfide (↑)	12232																
439			methylpyrazine (↑)	7976																
440			2-(methylthio)-ethanol (↑)	78925																
441			phenol (↑)	996																
442			dimethyl trisulfide (↑)	19310																
443			benzonitrile (↑)	7505																
444			2,3,5-trimethylpyrazine (↑)	26808	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	DSM 1103	Columbia Sheep blood agar	24 h at 37 °C	unknown	unknown	Jünger, M., Vautz, W., Kuhns, M., Hofmann, L., Ulbricht, S., Baumbach, J. I., ... & Perl, T. (2012). Ion mobility spectrometry for microbial volatile organic compounds: a new identification tool for human pathogenic bacteria. Applied microbiology and biotechnology, 93(6), 2603-2614. doi: 10.1007/s00253-012-3924-4.	Appl. Microbiol. Biotechnol.	2012
445			N-(phenylmethylene)-methanamine (↑)	73954																
446			2-nonanone (↑)	13187																
447			N,N'-dibenzylidemethylenediamine (↑)	66033																
448			2-decanone (↑)	12741																
449			N-(phenylmethylene)-1-propanamine (↑)	250250																
450			ethyl phenylacetate (↑)	7590																
451			N-(phenylmethylene)-1-butanamine (↑)	296031																
452			indole (↑)	798																
453			1-methyl-naphthalene (↑)	7002																
454			1-decanol	8174										brain-heart-infusion broth	unknown	unknown	unknown			
455			indole	798																
456			1-dodecanol	8193																
457			acetic acid	176	SPME-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	NCTC 10418					Tait, E., Perry, J. D., Stanforth, S. P., & Dean, J. R. (2014). Use of volatile compounds as a diagnostic tool for the detection of pathogenic	TrAC, Trends Anal. Chem.	2014
458			1-decanol	8174																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Pathogen	Classification	VOCs	PubChem ID	Methods	*In vivo* Sample							Bacterial strain	Culture conditions/Growth medium	Incubation time before analysis	Concentration range		Reference	Journal Code	Year
2						Saliva	Blood	Breath	Skin	Urine	Faeces	Milk				Value	Unit			
459			indole	798										1% glucose enteric fermentation broth	unknown	unknown	unknown	bacteria. TRA: Trends in Analytical Chemistry, 53, 117-125.		
460			1-dodecanol	8193																
461			1-tetradecanol	8209																
462			1-decanol	8174	MCC-IMS analysis		×		×	×			DSM 25944 and 12 clinical isolates	Lysogeny Broth (LB) fluid medium	72 h incubation	unknown	unknown	Kunze, N., Göpel, J., Kuhns, M., Jünger, M., Quintel, M., & Perl, T. (2013). Detection and validation of volatile metabolic patterns over different strains of two human pathogenic bacteria during their growth in a complex medium using multi-capillary column-ion mobility spectrometry (MCC-IMS). *Applied microbiology and biotechnology* , *97* (8), 3665-3676. DOI 10.1007/s00253-013-4762-8.	Appl. Microbiol. Biotechnol.	2013
463			5-methylheptan-3-one	7822																
464			2-phenylacetaldehyde	998																
465			ethanol	702																
466			nonanal	31289																
467			ammonia	222																
468			indole	798																
469			1-octanol	957																
470			1-octanol (dimer)	not available																
471			methanol	887	PTR-MS analysis	not used	not used	not used	not used	not used	not used	not used	DSMZ 30083	complex medium	24h	unknown	ppt-ppm	Bunge, Michael, et al. "On-line monitoring of microbial volatile metabolites by proton transfer reaction-mass spectrometry." *Applied and environmental microbiology* 74.7 (2008): 2179-2186.	Appl. Environ. Microbiol.	2008
472			acetaldehyde	177																
473			ethanol	702																
474			methanethiol	878																
475			acetone	180																
476			acetic acid	176																
477			indole	798																
478			dimethyl disulfide	12232	HS-GLC analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate 1	YEB medium	18h	unknown	unknown	Hayward, N. J., et al. "Development of specific tests for rapid detection of Escherichia coli and all species of Proteus in urine." *Journal of clinical microbiology* 6.3 (1977): 195-201.	J. Clin. Microbiol.	1977
479			dimethyl disulfide	12232									clinical isolate 2							
480			dimethyl disulfide	12232									clinical isolate 3							
481			dimethyl disulfide	12232									clinical isolate 4							
482			dimethyl disulfide	12232									clinical isolate 5							
483			dimethyl disulfide	12232									clinical isolate 6							
484			1-propanol	1031		unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	urine + buffered lactose PW						
485			ethanol	702																
486			ethanol	702	unknown		×			×			unknown	unknown	unknown	unknown	unknown	Broza, Y. Y., & Haick, H. (2013). Nanomaterial-based sensors for detection of disease by volatile organic compounds. Nanomedicine, *8* (5), 785-806. doi: 10.2217/nnm.13.64.	Nanomedicine	2013
487			dimethyl disulfide	12232																
488			methanethiol	878																
489			trimethylamine	1146																
490			ammonia	222																
491			hexane	8058	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	Bos, L. D., Sterk, P. J., & Schultz, M. J. (2013). Volatile metabolites of pathogens: a systematic review. doi:10.1371/journal.ppat.1003311	PLOS	2013
492			2-methyl-1-butanol	8723																
493			1-butanol	263																
494			1-decanol	8174																
495			1-dodecanol	8193																
496			ethanol	702																
497			methanol	887																
498			1-propanol	1031																
499			octanol	957																
500			1-pentanol	6276																
501			phenylacetic acid	999																
502			propanoic acid	1032																
503			3-methylbutanal	11552																
504			acetaldehyde	177																
505			benzaldehyde	240																
506			formaldehyde	712																
507			hexanal	6184																
508			dodecane	8182																
509			2-heptanone	8051																
510			acetoin	179																
511			acetone	180																
512			1-methyl-naphthalene	7002																
513			2-methylnaphthalene	7055																
514			2-methylphenol	335																
515			phenol	996																
516			ethyl acetate	8857																
517			ethyl butanoate	7762																
518			ethyl phenylacetate	7590																
519			n-propyl acetate	7997																
520			propyl phenylacetate	221641																
521			2-(methylthio)-ethanol	78925																
522			dimethyl disulfide	12232																
523			dimethyl trisulfide	19310																
524			hydrogen sulfide	402																
525			methanethiol	878																
526			2,3,5-trimethylpyrazine	26808																
527			2-aminoacetophenone	11952																
528			3-methyl-1H-indole	6736																
529			4-chloro-1H-indole	91345																
530			acetonitrile	6342																
531			benzonitrile	7505																
532			indole	798																
533			methylpyrazine	7976																
534			N,N'-dibenzylideneethylenediamine	66033																
535			N-butyl-1-phenylmethanimine	296031																
536			N-phenylmethylene-1-propanamine	250250																
537			N-phenylmethylene-methanamine	73954																
538			trimethylamine	1146																
539	*Giardia duodenalis*	Protozoa	2,2,4,4-tetramethyloctane (↑)	182333	GC-MS analysis						×		unknown	unknown	unknown	unknown	unknown	Bond, A., Vernon, A., Reade, S., Mayor, A., Wastling, J., Minetti, C., ... & Probert, C. (2015). PWE-173 Investigation of volatile organic compounds emitted from faeces for the diagnosis of giardiasis. *Gut* , *64* (Suppl 1), A288-A288. DOI: http://dx.doi.org/10.15403/jgld.2014.1121.243.abo	Gut	2015
540			acetic acid (↑)	176																
541			2,2,4,6,6-pentamethylheptane (↑)	26058																
542			cyclopentane (↑)	9253																
543			2-pentanone (↑)	7895																
544			2,6,6-trimethyldecane (↑)	545605																
545			1-propanol (↓)	1031																
546			3-methylfuran (↓)	13587																
547			1,3-bis (1,1-dimethylethyl) benzene (↑)	136810																
548			ethanol (↓)	702																
549			2,5-dimethylpyrazine (↑)	31252																
550			propanoic acid (↑)	1032																
551			pentanal (↑)	8063																
552			4-pentanolide (↑)	7921																
553			2-hydroxy-3-pentanone (↑)	521790																
554			2,2,3,3-tetramethylpentane (↑)	92723																
555			5-ethylcyclopent-1-enecarboxaldehyde (↑)	580057																
556			(E)-2-octenal (↑)	5283324																
557			ethyl isobutyrate (↓)	7342																
558			o-xylene (↓)	7237																
559			terpinolene (↓)	11463																
560			indole	798	SPME-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	blood agar or chocolate blood agar	48h	unknown	unknown	Preti, George, et al. "Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: analysis by solid-phase microextraction and gas	Clin. Microbiol. Rev.	2009
561			benzaldehyde	240																
562			acetic acid	176																
563			phenylmethanol	244																
564			acetaldehyde (↑)	177																
565			butanal (↑)	261																
566			propanal (↑)	527																
567			1-butanol (↑)	263																
568			methanol (↑)	887																
569			2,3-butanedione (↑)	650																
570			2-pentanone (↑)	7895																

Pathogen	Classification	VOCs	PubChem ID	Methods	*In vivo* Sample							Bacterial strain	Culture conditions/Growth medium	Incubation time before analysis	Concentration range		Reference	Journal Code	Year
					Saliva	Blood	Breath	Skin	Urine	Faeces	Milk				Value	Unit			
Haemophilus influenzae	Gram negative bacterium	4-heptanone (↑)	31246	GC-MS analysis			×					clinical isolate	chocolate agar plates; liquid cultures: tryptic soy broth	, 4.5, 6 and 7.5 h after inocu	unknown	ppt-ppm	W. Filipiak, A. Sponring, M. M. Baur, C. Ager, A. Filipiak, H. Wiesenhofer, M. Nagl, J. Troppmair, A. Amann. Characterization of volatile metabolites taken up by or released from Streptococcus pneumoniae and Haemophilus influenzae by using GC-MS. Microbiology 2012, 158, 3044. doi: 10.1099/mic.0.062687-0.	Microbiol.	2012
		acetic acid (↑)	176																
		ethyl acetate (↑)	8857																
		methyl methacrylate (↑)	6658																
		vinyl butyrate (↑)	31247																
		methyl propionate (↑)	11124																
		3-(ethylthio)propanal (↑)	229467																
		dimethyl sulfide (↑)	1068																
		methanethiol (↑)	878																
		dimethyl disulfide (↑)	12232																
		carbon disulfide (↑)	6348																
		methyl thioacetate (↑)	519840																
		ethyl methyl sulfide (↑)	12230																
		dimethyl trisulfide (↑)	19310																
		2-methyl-2-butene (↑)	10553																
		isoprene (↑)	6557																
		3-methyl-1-butene (↑)	11239																
		o-hydroxybenzaldehyde (↑)	6998																
		furan (↑)	8029																
		2-acetyl-1,4,5,6-tetrahydropyridine (↑)	520194																
		gama-butyrolactone (↑)	7302																
		3-ethyl-6-pentamethyldisilyloxyoctane	590048	GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	blood culture test tube over 48 h	one aliquot was first cultivated in SP4 broth (dilution of 1:10 for 24 h at 37°C);further diluted 1:50 (24 h), when the log phase was observed	unknown	unknown	Abd El Qader, A., Lieberman, D., Shemer Avni, Y., Svobodin, N., Lazarovitch, T., Sagi, O., & Zeiri, Y. (2015). Volatile organic compounds generated by cultures of bacteria and viruses associated with respiratory infections.*Biomedical Chromatography* . DOI: 10.1002/bmc.3494.	Biomed. Chromatogr.	2015
		heptane	8900																
		S-[tri-t-butoxysilyl]-2-mercaptoethylamine	6058																
		methylcyclohexane	7962																
		4-fluorohistamine	541569																
		isopentanol	31260																
		7-methyl-1,8-naphthyridin-2-amine	594420																
		ethylpentamethyldisiloxane	20667832																
		5-methylthieno[3,2-b]pyridine	591057																
		2,5-bis[(trimethylsilyl)oxy]-benzaldehyde	622536																
		1,2-bis(trimethylsilyl)benzene	519794																
		decamethyl tetrasiloxane	8852																
		isobutane	6360	SPME-GC-MS analysis			×					unknown	unknown	unknown	unknown	unknown	Sethi, S., Nanda, R., & Chakraborty, T. (2013). Clinical application of volatile organic compound analysis for detecting infectious diseases. *Clinical microbiology reviews* , *26* (3), 462-475. doi: 10.1128/CMR.00020-13.	Clin. Microbiol. Rev.	2013
		2-butanone	6569																
		ethyl acetate	8857																
		hydrogen cyanide	768	PTR-MS analysis			×					NCTC 11637	unknown	unknown	unknown	unknown			
		hydrogen nitrate	944																
		dimethylether	8254	SPME-GC-MS analysis			×					clinical isolate	air sample transferred from a Tedlar bag to a glass vial; SPME fiber was inserted into the vial and exposed to the gaseous mixture	15 min	unknown	unknown	Ulanowska, A., Kowalkowski, T., Hrynkiewicz, K., Jackowski, M., & Buszewski, B. (2011). Determination of volatile organic compounds in human breath for Helicobacter pylori detection by SPME-GC/MS. *Biomedical Chromatography* , *25* (3), 391-397. DOI 10.1002/bmc.1460.	Biomed. Chromatogr.	2011
		2,3-butadiene	not available																
		acetaldehyde	177																
		ethanol	702																
		isobutane	6360																
		acetonitrile	6342																
		n-butane	7843																
		acetone	180																
		2-propanol	3776																
		ethylether	3283																
		isoprene	6557																
		n-pentane	8003																
		2-methylfuran	10797																
		2-butanone	6569																
		ethyl acetate	8857																
		2-methylpentane	7892																
		3-methylpentane	7282																
		benzene	241																
		methylcyclopentane	7296																
		hexane	8058																
		toluene	1140																
		2-hexanone	11583																
		ethylbenzene	7500																
		p-xylene	7809																
		styrene	7501																
		benzaldehyde	240																
		nonane	8141																
		propane	6334	SPME-GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	unknown	isolation from patient stomach mucous membrane biopsies; selective medium BD BBLTM Stacker Plates; culture at 37ºC in microaerofilic conditions 5-6 days; suspension of isolated baceria in sterile water for analysis	1 h	unknown	unknown			
		acetaldehyde	177																
		ethanol	702																
		methanethiol	878																
		(E)-2-butene	62695																
		isobutane	6360																
		2-methylpropene	8255																
		acetonitrile	6342																
		n-butane	7843																
		pentafluoroethane	9633																
		acetone	180																
		carbon disulfide	6348																
		2-propanol	3776																
		ethylether	3283																
		methyl acetate	6584																
		dichlorofluoroethane	15586																
		2-methylbutane	6556																
		2-pentene	12585																
		n-pentane	8003																
		cyclopentane	9253																
		2-methylpropanal	6561																
		trichloromethane	6212																
		2-butanone	6569																
		ethyl acetate	8857																
		2-methyl-1-pentene	12986																
		isobutanol	6560																
		2-methylpentane	7892																
		3-methylpentane	7282																
		benzene	241																
		methylcyclopentane	7296																
		hexane	8058																
		cyclohexane	8078																
		toluene	1140																
		mercaptoacetone	520144																
		3-methylbutanal	11552																
		2-ethoxy-2-methylpropane	12512																
		dimethyl disulfide	12232																
		1-pentanol	6276																
		4-methylpentane	not available																
		methylcyclohexane	7962																
		tetrahydro-2,2,4,4-tetramethylfuran	27010																
		ethylbenzene	7500																
		styrene	7501																
		2,4-dimethyl-1-heptene	123385																
		octane	356																
		3,5-dimethyloctane	139989																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
						In vivo Sample										Concentration range				
	Pathogen	Classification	VOCs	PubChem ID	Methods	Saliva	Blood	Breath	Skin	Urine	Faeces	Milk	Bacterial strain	Culture conditions/Growth medium	ncubation time before analys	Value	Unit	Reference	Journal Code	Year
682			3-ethyloctane	79985																
683			decane	15600																
684			ethanol	702																
685			formaldehyde	712																
686			hydrogen sulfide	402																
687			dimethyl sulfide	1068	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	NCTC 9633	sterile urine (20 mL) from healthy males inoculated to a concentration of between 10^7 and 10^9 cfu/mL	37˚C for 6h	unknown	ppb	Storer, M. K., Hibbard-Melles, K., Davis, B., & Scotter, J. (2011). Detection of volatile compounds produced by microbial growth in urine by selected ion flow tube mass spectrometry (SIFT-MS). Journal of microbiological methods, 87(1), 111-113. doi: 10.1016/j.mimet.2011.06.012.	J. Microbiol. Methods	2011
688			dimethyl disulfide	12232																
689			methanethiol	878																
690			ammonia	222																
691			trimethylamine	1146																
692			isoprene (↑)	6557																
693			1-propanol (↑)	1031																
694			4-methylcyclohexene (↑)	11572																
695			2-butanone (↑)	6569																
696			3-methylcyclohexene (↑)	11573	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 700683	blood agar plates; incubated overnight at 37˚C; transfer to sterile Brain Heart Infusion broth, growth for 4h with constant agitation at the same temperature	unknown	unknown	unknown	Boots AW, Smolinska A, van Berkel JJ, Fijten RR, Stobberingh EE, et al. (2014) Identification of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography-mass spectrometry. J Breath Res 8: 027106. doi:10.1088/1752-7155/8/2/027106.	J. Breath Res.	2014
697			3-methylbutanal (↓)	11552																
698			2-methylbutanal (↓)	7284																
699			2,3,3-trimethylpentane (↓)	11215																
700			benzaldehyde (↓)	240																
701			2,3-butanedione (↓)	650																
702			1-methyl-4-(1-methylethenyl)cyclohexane (↓	14299																
703			1-undecene	13190																
704			1-octanol	957																
705			1-decanol	8174																
706			1-dodecanol	8193	HS-SPME-GC-FID analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	50 mL TS broth in 125 mL vial	Samples incubated with shaking for 22-26 h at 32º C	unknown	unknown			
707			2-undecanone	8163																
708			tridecen-2-one	53427438																
709			2-nonanone	13187																
710			2-tridecanone	11622																
711			ethanol	702																
712			1-propanol	1031																
713			isopentanol	31260																
714			1-octanol	957																
715			9-decenol	25612	HS-SPME-GC-FID/GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	10 mL Mueller Hinton broth in 25 mL vials	Samples incubated for up to 10h at 37ºC without shaking	unknown	unknown	Tait, E., Perry, J. D., Stanforth, S. P., & Dean, J. R. (2014). Identification of volatile organic compounds produced by bacteria using HS-SPME-GC–MS. Journal of chromatographic science, 52(4), 363-373. doi:10.1093/chromsci/bmt042	J. Chromatogr. Sci.	2014
716			1-decanol	8174																
717	*Klebsiella pneumoniae*	Gram negative bacteriu	1-dodecanol	8193																
718			(Z)-7-tetradecen-1-ol	5362795																
719			1-tetradecanol	8209																
720			3-methylbutanoic acid	10430																
721			2-methylbutanoic acid	8314																
722			2-methylpropanoic acid	6590	SPME- GC–MS/GC–flame photometric detector ana	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	Blood agar and chocolate blood agar	Samples incubated for at lea	unknown	unknown			
723			1-hydroxy-2-butanone	521300																
724			butanoic acid	264																
725			4-methylhexanoic acid	15271																
726			2-phenylethanol	6054																
727			ethanol (↑)	702	MCC-IMS analysis	not used	not used	not used	not used	not used	not used	not used	DSM 2026	Columbia Sheep blood agar	24 h at 37 °C	unknown	unknown	Jünger, M., Vautz, W., Kuhns, M., Hofmann, L., Ulbricht, S., Baumbach, J. I., ... & Perl, T. (2012). Ion mobility spectrometry for microbial volatile organic compounds: a new identification tool for human pathogenic bacteria. Applied microbiology and	Appl. Microbiol. Biotechnol.	2012
728			2-phenylethanol (↑)	6054																
729			3-methylbutanal (↑)	11552	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used								
730			isopentanol (↑)	31260																
731			2-(methylthio)-ethanol (↑)	78925																
732			2-nonanone (↑)	13187																
733			undecane	14257	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	Bos, L. D., Sterk, P. J., & Schultz, M. J. (2013). Volatile metabolites of pathogens: a systematic review. doi:10.1371/journal.ppat.1003311	PLOS	2013
734			2-butanol	6568																
735			isopentanol	31260																
736			ethanol	702																
737			isobutanol	6560																
738			2-methyl-1-butanol	8723																
739			3-methylbutanal	11552																
740			formaldehyde	712																
741			2-nonanone	13187																
742			2-heptanone	8051																
743			toluene	1140																
744			isoamyl acetate	31276																
745			2-(methylthio)-ethanol	78925																
746			dimethyl disulfide	12232																
747			hydrogen sulfide	402																
748			methanethiol	878																
749			ammonia	222																
750			trimethylamine	1146																
751			3-ethyl-6-pentamethyldisilyloxyoctane	590048																
752			heptane	8900																
753			S-[tri-t-butoxysilyl]-2-mercaptoethylamine	6058																
754			methylcyclohexane	7962																
755			4-fluorohistamine	541569																
756			isopentanol	31260	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	unknown	blood culture test tube over 48 h	one aliquot was first cultivated in SP4 broth (dilution of 1:10 for 24 h at 37°C);further diluted 1:50 for another 24 h, when the log phase was observed	unknown	unknown	Abd El Qader, A., Lieberman, D., Shemer Avni, Y., Svobodin, N., Lazarovitch, T., Sagi, O., & Zeiri, Y. (2015). Volatile organic compounds generated by cultures of bacteria and viruses associated with respiratory infections.*Biomedical Chromatography* . DOI: 10.1002/bmc.3494.	Biomed. Chromatogr.	2015
757	*Legionella pneumophila*	Gram negative bacteriu	7-methyl-1,8-naphthyridin-2-amine	594420																
758			ethylpentamethyldisiloxane	20667832																
759			5-methylthieno[3,2-b]pyridine	591057																
760			2,5-bis[(trimethylsilyl)oxy]-benzaldehyde	622536																
761			1,2-bis(trimethylsilyl)benzene	519794																
762			decamethyl tetrasiloxane	8852																
763			ethanol (↑)	702																
764			3-methylbutanal (↑)	11552	GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	NRS 382	Mueller Hinton broth (MHB) and trypticase soy agar (TSA) ; incubation at 37˚C	10 min and 8 h exposure	unknown	unknown	Jia, B., Sohnlein, B., Mortelmans, K., Coggiola, M., & Oser, H. (2010). Distinguishing methicillin-resistant and sensitive Staphylococcus aureus using volatile	IEEE Sens. J.	2010
765			1-propanol (↑)	1031																
766			benzaldehyde (↓)	240																
767	ethicillin-resistant *Staphylococcus aure*	Gram positive bacterium	1,4-dichlorobenzene	4685																
768			2-heptanone	8051																
769			2-pentanone	7895	GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	blood agar plates; incubated overnight at 37˚C; transfer to sterile Brain Heart Infusion broth, growth for 4h with constant agitation at the same temperature	unknown	unknown	unknown	Boots AW, Smolinska A, van Berkel JJ, Fijten RR, Stobberingh EE, et al. (2014) Identification of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography-mass spectrometry. J Breath Res 8: 027106. doi:10.1088/1752-	J. Breath Res.	2014
770			2-dodecene	522440																
771			trichloroacetic acid	6421																
772			2,2,3,3-tetramethylhexane	26057																
773			(E)-methylthio-1-propene	637915																
774			heptane	8900																
775			S-[tri-t-butoxysilyl]-2-mercaptoethylamine	6058																
776			methylcyclohexane	7962																
777			4-fluorohistamine	541569																
778			isopentanol	31260	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	unknown	blood culture test tube over 48 h	one aliquot wasfirst cultivated in SP4 broth at a dilution of 1:10 for 24 h at 37°C; further diluted 1:50 for another 24 h, when the log phase was observed	unknown	unknown	Abd El Qader, A., Lieberman, D., Shemer Avni, Y., Svobodin, N., Lazarovitch, T., Sagi, O., & Zeiri, Y. (2015). Volatile organic compounds generated by cultures of bacteria and viruses associated with respiratory infections.*Biomedical Chromatography* . DOI: 10.1002/bmc.3494.	Biomed. Chromatogr.	2015
779			7-methyl-1,8-naphthyridin-2-amine	594420																
780	*Moraxella catarrhalis*	Gram negative bacteriu	ethylpentamethyldisiloxane	20667832																
781			5-methylthieno[3,2-b]pyridine	591057																
782			2,5-bis[(trimethylsilyl)oxy]-benzaldehyde	622536																
783			1,2-bis(trimethylsilyl)benzene	519794																
784			decamethyl tetrasiloxane	8852																
785			benzaldehyde	240	SPME-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	blood agar or chocolate blood agar	48h	unknown	unknown	Preti, George, et al. "Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: analysis by solid-phase	Clin. Microbiol. Rev.	2009
786			phenylmethanol	244																
787			2-phenylethanol	6054																
788			dimethyl disulfide	12232														Tait, E., Perry, J. D., Stanforth, S. P., & Dean, J.		
789			isopentanol	31260																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1						*In vivo* Sample										Concentration range				
2	Pathogen	Classification	VOCs	PubChem ID	Methods	Saliva	Blood	Breath	Skin	Urine	Faeces	Milk	Bacterial strain	Culture conditions/Growth medium	ncubation time before analys	Value	Unit	Reference	Journal Code	Year
790	*Morganella morganii*	Gram negative bacteriu	dimethyl trisulfide	19310	SPME-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	WILD 10257	brain-heart-infusion broth	overnight incubation 37˚C	unknown	unknown	R. (2014). Use of volatile compounds as a diagnostic tool for the detection of pathogenic bacteria. TrAc Trends in Analytical Chemistry, 53, 117-125.	TrAC, Trends Anal. Chem.	2014
791			1-decanol	8174																
792			1-dodecanol	8193																
793			phenol	996																
794			1-methyl-naphthalene	7002	ATD/GC-MS analysis			×					unknown	unknown	unknown	unknown	unknown	Cheepsattayakorn, A., & Cheepsattayakorn, R. (2014). Breath Tests in Diagnosis of Pulmonary Tuberculosis. Recent patents on biotechnology, 8(2), 172-175. DOI: 74/1872208309666140904115813	Recent Pat. Biotechnol.	2014
795			1,4-dimethylcyclohexane	11523																
796			1,3-isobenzofurandione	6811																
797			2,3-dimethylpentane	11260																
798			acetaldehyde	177																
799			phenylmethanol	244																
800			1,1,3-trimethyl-3-phenyl-cyclohexane	not available																
801			tridecane	12388																
802			3,7-dimethyldecane	28468																
803			5-ethyl-2-methyl-heptane	26056																
804			1,3,5-trimethylbenzene	7947																
805			4,6,8-trimethyl-1-nonene	41077																
806			hexylcyclohexane	20283																
807			4-methyl-1-hexene	19589																
808			bis-(3,5,5-trimethylhexyl) phthalate	34277																
809			4-methyldodecane	521958																
810			3-(1-methylethyl)oxetane	543882																
811			3-(1-methylethyl)oxetane	543882	GC-MS analysis			×					unknown	unknown	unknown	unknown	unknown	Phillips, M., Basa-Dalay, V., Bothamley, G., Cataneo, R. N., Lam, P. K., Natividad, M. P. R., ... & Wai, J. (2010). Breath biomarkers of active pulmonary tuberculosis. *Tuberculosis* , *90* (2), 145-151. doi: 10.1016/j.tube.2010.01.003.	Tuberculosis	2010
812			4-methyldodecane	521958																
813			hexylcyclohexane	20283																
814			bis-(3,5,5-trimethylhexyl) phthalate	34277																
815			1,3,5-trimethylbenzene	7947																
816			3,7-dimethyldecane	28468																
817			tridecane	12388																
818			4,6,8-trimethyl-1-nonene	41077																
819			5-ethyl-2-methyl-heptane	26056																
820			4-methyl-1-hexene	19589																
821			1-methyl-naphthalene	7002	GC-MS analysis			×					H37RV	Myco bottles containing 1.0 ml of Growth Supplement were inoculated with 0.5 ml of a 1.0 McFarland suspension in sterile saline prepared from isolates grown on Lowenstein Jensen medium	Samples incubated an additional 2 days after the Myco bottle yielded a positive signal	unknown	unknown	Phillips, M., Cataneo, R. N., Condos, R., Erickson, G. A. R., Greenberg, J., La Bombardi, V., ... & Tietje, O. (2007). Volatile biomarkers of pulmonary tuberculosis in the breath. Tuberculosis, 87(1), 44-52. doi: 10.1016/j.tube.2010.01.003.	Tuberculosis	2007
822			3-heptanone	7802																
823			methylcyclododecane	524446																
824			2,2,4,6,6-pentamethylheptane	26058																
825			cymol	7463																
826			1,4-dimethylcyclohexane	11523																
827			methyl nicotinate	7151	GC-MS analysis			×					H37RV [1] H37RA [1] clinical isolates [1 2 3]	[1] Middlebrook 7H11 agar [2]Middlebrook 7H9 broth [3] Sauton liquid medium	18-24 h	unknown	unknown	Mgode, G. F., Weetjens, B. J., Nawrath, T., Lazar, D., Cox, C., Jubitana, M., ... & Kaufmann, S. H. (2012). Mycobacterium tuberculosis volatiles for diagnosis of tuberculosis by Cricetomys rats.*Tuberculosis* , *92* (6), 535-542. doi: 10.1016/j.tube.2012.07.006	Tuberculosis	2012
828			methyl 4-anisate	8499																
829			2-phenylanisol	6835																
830			4-methylanisol	7731																
831			ethyl 4-anisate	60979																
832			trimethyloxazole	30215																
833			methyl 2-aminobenzoate	8635																
834			benzothiazole	7222																
835			4-hydroxy-4-methyl-2-pentanone	31256																
836			3-methyl-4-pentanolide	248934																
837			4-methyl-5-hexanolide	544628																
838			dimethylpentanolide	not available																
839			cyclic proline-glycine	456653																
840			2-phenylethanol	6054																
841			methyl benzoate	7150																
842			4-pentanolide	7921																
843			methyl phenylacetate	7559																
844			methyl 2-furoate	11902																
845			methyl salicylate	4133																
846			camphor	2537																
847			methylbutenolide	not available																
848			methyl dimethylbenzoate	32786																
849			phenylmethanol	244																
850			ethyl benzoate	7165																
851			methyl phenylacetate	7559	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	H37RA	Lowenstein–Jensen/Glycerol, sheep blood agar and BacT/Alert® MP	incubated at 37 ˚C;sampling was carried out in a semicontinuous mode every 24 h for 3	unknown	unknown	Syhre, M., & Chambers, S. T. (2008). The scent of Mycobacterium tuberculosis. *Tuberculosis* , *88* (4), 317-323. doi: 10.1016/j.tube.2008.01.002.	Tuberculosis	2008
852			methyl 4-anisate	8499																
853			methyl nicotinate	7151																
854			2-phenylanisol	6835																
855			o-xylene (↑)	7237	GC-MS analysis					×			clinical isolates	unknown	samples were stored and transported at 4˚C; analyzed preferentially on the same day of sample collection, or stored with	unknown	unknown	Banday, K. M., Pasikanti, K. K., Chan, E. C. Y., Singla, R., Rao, K. V. S., Chauhan, V. S., & Nanda, R. K. (2011). Use of urine volatile organic compounds to discriminate tuberculosis patients from healthy subjects.	Anal. Chem.	2011
856			isopropyl acetate (↑)	7915																
857			3-pentanol (↓)	11428																
858			dimethylstyrene (↓)	62385																
859			cymol (↓)	7463																
860			o-xylene	7237	GC-MS analysis					×			clinical isolates	unknown	unknown	unknown	unknown	Sethi, S., Nanda, R., & Chakraborty, T. (2013). Clinical application of volatile organic compound analysis for detecting infectious diseases. *Clinical microbiology reviews* , *26* (3), 462-475. doi: 10.1128/CMR.00020-13.	Clin. Microbiol. Rev.	2013
861			isopropyl acetate	7915																
862			3-pentanol	11428																
863			dimethylstyrene	62385																
864			cymol	7463																
865			camphene	6616	ATD-GC-SAW analysis			×					unknown	unknown	unknown	unknown	unknown	Phillips, M., Basa-Dalay, V., Blais, J., Bothamley, G., Chaturvedi, A., Modi, K. D., ... & Udwadia, Z. F. (2012). Point-of-care breath test for biomarkers of active pulmonary tuberculosis. *Tuberculosis* , *92* (4), 314-320. doi: 10.1016/j.tube.2012.04.002.	Tuberculosis	2012
866			beta-pinene	14896																
867			1,3,5-trimethylbenzene	7947																
868			1-methyl-naphthalene	7002																
869			tridecane	12388																
870			2-butyloctanol	19800																
871			4-methyldodecane	521958																
872			methyl nicotinate	7151	unknown			×					unknown	unknown	unknown	unkno wn		Broza, Y. Y., & Haick, H. (2013). Nanomaterial-based sensors for detection of disease by volatile organic compounds. Nanomedicine, *8* (5), 785-806. doi: 10.2217/nnm.13.64.	Nanomedicine	2013
873			methyl phenylacetate	7559																
874			methyl 4-anisate	8499																
875			2-phenylanisol	6835																
876			3-(1-methylethyl)oxetane	543882																
877			4-methyldodecane	521958																
878			hexylcyclohexane	20283																
879			bis-(3,5,5-trimethylhexyl) phthalate	34277																
880			1,3,5-trimethylbenzene	7947																
881			3,7-dimethyldecane	28468																
882			tridecane	12388																
883			4,6,8-trimethyl-1-nonene	41077																
884			5-ethyl-2-methyl-heptane	26056																
885			4-methyl-1-hexene	19589																
886			ethanol	702	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	NZESR 1033	blood culture bottles	24 h	unknown	unknown	Allardyce, Randall A., et al. "Detection of volatile metabolites produced by bacterial growth in blood culture media by selected ion flow tube mass spectrometry (SIFT-MS)." *Journal of microbiological*	J. Microbiol. Methods	2006
887			formaldehyde	712																
888			methanethiol	878																
889			indole	798																
890	Neisseria meningitidis	Gram negative bacteriu	ethanol	702	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 13077	Roswell Park growth media with murine macrophages	amounts of time (1, 3, 6, 24,	unknown	unknown	Wood, William L., et al. "Analysis of volatile bacterial metabolites by gas chromatography-mass spectrometry." *Spectroscopy* 21.6 (2006).	Spectroscopy	2006
891			dimethyl sulfide	1068																
892			1,3,5-trifluorobenzene	9745																
893			2,3,4-trimethylpentane	11269																
894			2,2,5-trimethylhexane	19041																
895			styrene	7501																
896			1,2-dimethylcyclopropane	102832																
897			2-methylpropanal	6561																
898			methacrolein	6562																
899			N-2-dimethyl-1-propanamine	12249																
900			3-methylbutanal	11552																
901			carbon dioxide	280																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Pathogen	Classification	VOCs	PubChem ID	Methods	*In vivo* Sample							Bacterial strain	Culture conditions/Growth medium	ncubation time before analys	Concentration range		Reference	Journal Code	Year
2						Saliva	Blood	Breath	Skin	Urine	Faeces	Milk				Value	Unit			
903			isoprene	6557																
904			acetone	180																
905			benzene	241																
906	*Plasmodium falciparum*	Protozoa	cyclohexanone	7967	GC-MS analysis			×					3D7	1640 RPMI medium, 37˚C in 270mL polystyrene flasks, with a culture volume of 50mL, and a low-O2 gaseous environment (1% O2, 3% CO2, and 96% N2); medium changed daily	samples collected before and during early-stage malaria and after antimalarial drug administration	unknown	unknown	Berna, A. Z., McCarthy, J. S., Wang, R. X., Saliba, K. J., Bravo, F. G., Cassells, J., ... & Trowell, S. C. (2015). Analysis of Breath Specimens for Biomarkers of Plasmodium falciparum Infection. *The Journal of infectious diseases* . doi: 10.1093/infdis/jiv176.	J. Infect. Dis.	2015
907			allyl methyl sulfide	66282																
908			methyl propyl sulfide	19754																
909			(Z)-1-methylthio-1-propene	5364225																
910			(E)-methylthio-1-propene	637915																
911			dimethyl disulfide	12232																
912			isopentanol	31260																
913			1-octanol	957	SPME-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	NCTC 10975	brain-heart-infusion broth	overnight incubation 37˚C	unknown	unknown	Tait, E., Perry, J. D., Stanforth, S. P., & Dean, J. R. (2014). Use of volatile compounds as a diagnostic tool for the detection of pathogenic bacteria. TrAc Trends in Analytical Chemistry, 53, 117-125.	TrAC, Trends Anal. Chem.	2014
914			isoamyl acetate	31276																
915			1-decanol	8174																
916			1-dodecanol	8193																
917			methanethiol	878	HS-GC analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	YEB or LAS medium	18h	unknown	unknown	Hayward, N. J., et al. "Development of specific tests for rapid detection of Escherichia coli and	J. Clin. Microbiol.	1977
918			dimethyl disulfide	12232																
919			1-butanol	263	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate NP1	nutrient agar; incubated at 37˚C aerobically when required	24 h	121.36	ppb	Thorn, R., Reynolds, D. M. and Greenman, J. (2011) Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro.Journal of Microbiological Methods, 84 (2). pp. 258-264. ISSN 0167-7012 DOI: 10.1016/j.mimet.2010.12.001.	J. Microbiol. Methods	2011
920			1-pentanol	6276												1962.27				
921			2-aminoacetophenone	11952												3.08				
922			acetoin	179												14.09				
923			butanoic acid	264												37.07				
924			ethanol	702												335.75				
925			ethyl acetate	8857												18.64				
926			ethyl butanoate	7762												17.12				
927			formaldehyde	712												1911.06				
928			hydrogen sulfide	402												4010.85				
929			isoprene	6557												111.91				
930	*Proteus mirabilis*	Gram negative bacterium	methanethiol	878												2024				
931			phenylacetic acid	999												8.09				
932			pyrrole	8027												0.88				
933			trimethylamine	1146												54.49				
934			1-butanol	263									clinical isolate NP2			157.1				
935			1-pentanol	6276												1850.04				
936			2-aminoacetophenone	11952												2.48				
937			acetoin	179												14.59				
938			butanoic acid	264												38.39				
939			dimethyl disulfide	12232												1759.54				
940			ethanol	702												170.2				
941			ethyl acetate	8857												19.3				
942			ethyl butanoate	7762												20.88				
943			formaldehyde	712												7618.38				
944			hydrogen sulfide	402												4223.55				
945			isoprene	6557												80.44				
946			methanethiol	878												7995.08				
947			phenylacetic acid	999												10.77				
948			pyrrole	8027												3.84				
949			trimethylamine	1146												81.32				
950			carbon dioxide	280																
951			ammonia	222																
952			methanethiol	878	IMR-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	ATCC 13315	blood agar plates	24 h	unknown	unknown	Dolch, M. E., et al. "Volatile compound profiling for the identification of Gram-negative bacteria by ion-molecule reaction–mass spectrometry." *Journal of*	J. Appl. Microbiol.	2012
953			indole	798																
954			acetone	180																
955			acetaldehyde	177																
956			formaldehyde	712																
957	*Proteus vulgaris*	Gram negative bacterium	n-propyl acetate	7997																
958			hydrogen sulfide	402																
959			dimethyl disulfide	12232	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	NCTC 4175	sterile urine (20 mL) from healthy males inoculated to a concentration of between 10^7 and 10^9 cfu/mL	37˚C for 6h	unknown	ppb	Storer, M. K., Hibbard-Melles, K., Davis, B., & Scotter, J. (2011). Detection of volatile compounds produced by microbial growth in urine by selected ion flow tube mass spectrometry (SIFT-MS). Journal of microbiological methods, 87(1), 111-113. doi: 10.1016/j.mimet.2011.06.012.	J. Microbiol. Methods	2011
960			dimethyl sulfide	1068																
961			methanethiol	878																
962			ammonia	222																
963			trimethylamine	1146																
964			2-aminoacetophenone	11952																
965			indole	798																
966			1-butanol	263	SIFT-MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 15692	Frozen stocks (-80˚C); Resuscitated onto nutrient agar; incubated at 37˚C aerobically when required	5h and 24h	25.38	ppb	Thorn, R., Reynolds, D. M. and Greenman, J. (2011) Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro.Journal of Microbiological Methods, 84 (2). pp. 258-264. ISSN 0167-7012 DOI: 10.1016/j.mimet.2010.12.001	J. Microbiol. Methods	2011
967			1-pentanol	6276												35.05				
968			ammonia	222												337.74				
969			ethanol	702												271.46				
970			formaldehyde	712												526.03				
971			hydrogen sulfide	402												218.15				
972			isoprene	6557												87.15				
973			methanethiol	878												565.73				
974			trimethylamine	1146												102.08				
975			1-butanol	263									ATCC 9027			19.22				
976			ammonia	222												494.3				
977			dimethyl disulfide	12232												990.88				
978			ethanol	702												306.98				
979			formaldehyde	712												734.81				
980			hydrogen sulfide	402												152.01				
981			methanethiol	878												780.09				
982			pyrrole	8027												0.64				
983			trimethylamine	1146												61.56				
984			isopentanol	31260	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 27853	inoculated in a 4 ml liquid preculture and grown over night at 37°C without agitation; cultivated in tryptic soy broth medium	1,5 h- 28,0 h	unknown	ppt-ppm	Filipiak, Wojciech, et al. "Molecular analysis of volatile metabolites released specifically by Staphylococcus aureus and Pseudomonas aeruginosa." BMC microbiology DOI: 10.1186/1471-2180-12-113.	BMC Microbiol.	2012
985			ethanol	702																
986			2-butanol	6568																
987			2-nonanone	13187																
988			2-pentanone	7895																
989			2-heptanone	8051																
990			4-heptanone	31246																
991			3-octanone	246728																
992			2-butanone	6569																
993			methyl isobutyl ketone	7909																
994			ethyl acetate	8857																
995			methyl 2-methylbutyrate	13357																
996			methyl methacrylate	6658																
997			ethyl 2-methylbutyrate	24020																
998			2-methylbutyl isobutyrate	97883																
999			isoamyl butyrate	7795																
1000			2-methylbutyl 2-methylbutyrate	17129																
1001			amyl isovalerate	95978																
1002			dimethyl sulfide	1068																
1003			dimethyl trisulfide	19310																
1004			methanethiol	878																
1005			mercaptoacetone	520144																
1006			2-methoxy-5-methylthiophene	141615																
1007			3-(ethylthio)propanal	229467																
1008			1-undecene	13190																
1009			2-methyl-2-butene	10553																
1010			1,10-undecadiene	139543																
1011			1-nonene	31285																
1012			1-decene	13381																
1013			1-dodecene	8183																
1014			n-butane	7843																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
						In vivo Sample										Concentration range				
	Pathogen	Classification	VOCs	PubChem ID	Methods	Saliva	Blood	Breath	Skin	Urine	Faeces	Milk	Bacterial strain	Culture conditions/Growth medium	ncubation time before analys	Value	Unit	Reference	Journal Code	Year
1014			isoprene	6557																
1015			10-methyl-1-undecene	519941																
1016			pyrrole	8027																
1017			3-methyl-1H-pyrrole	12023																
1018			1-vinyl aziridine	21843																
1019			2,3-butanedione (↓)	650																
1020			benzaldehyde (↓)	240																
1021			acetaldehyde (↓)	177																
1022			methacrolein (↓)	6562																
1023			3-methylbutanal (↓)	11552																
1024			nonanal (↓)	31289																
1025			propanal (↓)	527																
1026			3-methyl-2-butenal (↓)	61020																
1027			acrolein (↓)	7847																
1028			butanal (↓)	261																
1029			2-methylpropanal (↓)	6561																
1030			octanal (↓)	454																
1031			isoprene	6557	GC-*tof* -MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 27853	growth on blood agar plates and incubated overnight at 37˚C; transfer to sterile Brain Heart Infusion broth, growth for 4h with constant agitation at the same temperature	unknown	unknown	unknown	Boots AW, Smolinska A, van Berkel JJ, Fijten RR, Stobberingh EE, et al. (2014) Identification of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography-mass spectrometry. J Breath Res 8: 027106. doi:10.1088/1752-7155/8/2/027106	J. Breath Res.	2014
1032			3-methylbutanal (↓)	11552																
1033			2-methylbutanal (↓)	7284																
1034			2,3,3-trimethylpentane (↓)	11215																
1035			benzaldehyde (↓)	240																
1036			1-undecene	13190																
1037			2-pentene	12585																
1038			2,3-butanedione (↓)	650																
1039			2-butanone	6569																
1040			2-heptanone	8051																
1041			2-nonanone	13187																
1042			1-methyl-4-(1-methylethenyl)cyclohexane (↓	14299																
1043			ethanol	702	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	unknown	unknown	unknown	unknown	unknown	Sohrabi M, Zhang L, Zhang K, Ahmetagic A, Wei MQ (2014) Volatile Organic Compounds as Novel Markers for the Detection of Bacterial Infections. Clin Microbial 3: 151. DOI: 10.13140/2.1.5009.0887 doi:10.4172/2327-5073.1000151	J. Clin. Microbiol.	2014
1044			acetone	180																
1045			2-butanone	6569																
1046			2-pentanone	7895																
1047			isoprene	6557																
1048			2-aminoacetophenone	11952																
1049			dimethyl sulfide	1068																
1050			dimethyl disulfide	12232																
1051			dimethyl trisulfide	19310																
1052			methyl thiocyanate	11168																
1053			methyl isopropyl ketone	11251																
1054			acetophenone	7410																
1055			methyl thioacetate	519840																
1056			methyl thiobutanoate	62444																
1057			hydrogen cyanide	768																
1058			acetonitrile	6342	SIFT-MS analysis															
1059			ethanol	702																
1060			acetone	180																
1061			acetic acid	176																
1062			ethylene glycol	174																
1063			2-pentanone	7895																
1064			4-methylphenol	2879																
1065			indole	798																
1066			2-aminoacetophenone	11952																
1067			2-nonanone	13187	SESI-MS analysis															
1068			2-undecanone	8163																
1069			2-aminoacetophenone	11952																
1070			acetic acid	176	unknown															
1071			acetone	180																
1072			acetonitrile	6342																
1073			ammonia	222																
1074			2-butanone	6569																
1075			dimethyl sulfide	1068																
1076			dimethyl disulfide	12232																
1077			ethanol	702																
1078			hydrogen cyanide	768																
1079			isoprene	6557																
1080			methanol	887																
1081			methanethiol	878																
1082			hydrogen cyanide (↑)	768	SIFT-MS analysis															
1083			methyl thiocyanate	11168																
1084			2-aminoacetophenone	11952	GC-MS analysis			×					unknown	unknown	unknown	unknown	unknown	Sethi, S., Nanda, R., & Chakraborty, T. (2013). Clinical application of volatile organic compound analysis for detecting infectious diseases. *Clinical microbiology reviews* , *26* (3),	Clin. Microbiol. Rev.	2013
1085					SPME-GC-MS analysis															
1086			ethanol (↑)	702	MCC-IMS analysis	not used	not used	not used	not used	not used	not used	not used								
1087			3-methylbutanal (↑)	11552	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	DSM 4638	Columbia Sheep blood agar	24 h at 37 °C	unknown	unknown	Jünger, M., Vautz, W., Kuhns, M., Hofmann, L., Ulbricht, S., Baumbach, J. I., ... & Perl, T. (2012). Ion mobility spectrometry for microbial volatile organic compounds: a new identification tool for human pathogenic bacteria. Applied microbiology and biotechnology, 93(6), 2603-2614. doi:	Appl. Microbiol. Biotechnol.	2012
1088			dimethyl disulfide (↑)	12232																
1089			isopentanol (↑)	31260																
1090			benzonitrile (↑)	7505																
1091			1-undecene (↑)	13190																
1092			2-nonanone (↑)	13187																
1093			acetone	180	MCC-IMS analysis	not used	not used	not used	not used	not used	not used	not used	DSM 1117	Lysogeny Broth (LB) fluid medium	72h	unknown	unknown	Kunze, N., Göpel, J., Kuhns, M., Jünger, M., Quintel, M., & Perl, T. (2013). Detection and validation of volatile metabolic patterns over different strains of two human pathogenic bacteria during their growth in a complex medium using multi-capillary column-ion mobility spectrometry (MCC-IMS). *Applied microbiology and biotechnology* , *97* (8), 3665-3676. DOI 10.1007/s00253-013-4762-8.	Appl. Microbiol. Biotechnol.	2013
1094			2-phenylacetaldehyde	998																
1095			ammonia	222																
1096			5-methylheptan-3-one	7822																
1097			nonanal	31289																
1098			ammonia (dimer)	not available																
1099			dodecane	8182																
1100			2-ethyl-1-hexanol	7720																
1101			acetone	180																
1102			2-phenylacetaldehyde	998			×		×	×			12 clinical isolates			unknown	unknown			
1103			ammonia	222																
1104			5-methylheptan-3-one	7822																
1105			nonanal	31289																
1106			ammonia (dimer)	not available																
1107			2-nonanone	13187	SPME-GC-MS analysis			×					unknown	unknown	unknown	unknown	unknown	Savelev, S. U., Perry, J. D., Bourke, S. J., Jary, H., Taylor, R., Fisher, A. J., ... & De Soyza, A. (2011). Volatile biomarkers of Pseudomonas aeruginosa in cystic fibrosis and noncystic fibrosis bronchiectasis. *Letters in applied*	Lett. Appl. Microbiol.	2011
1108			2,4-dimethyl-1-heptene	123385																
1109			1-heptene	11610																
1110			isopentanol	31260																
1111			limonene	22311																
1112			ethane (↑)	6324	GC-MS analysis			×					unknown	unknown	unknown	unknown	ppb	Barker, M., Hengst, M., Schmid, J., Buers, H. J., Mittermaier, B., Klemp, D., & Koppmann, R. (2006). Volatile organic compounds in the exhaled breath of young patients with cystic fibrosis. *European respiratory journal* , *27* (5), 929-936. DOI: 10.1183/09031936.06.00085105.	Eur. Respir. J.	2006
1113			propane	6334																
1114			n-pentane (↑)	8003																
1115			methanol (↓)	887																
1116			ethanol (↓)	702																
1117			2-propanol (↓)	3776																
1118			acetone (↓)	180																
1119			isoprene (↓)	6557																
1120			benzene (↑)	241																
1121			toluene (↑)	1140																
1122			dimethyl sulfide (↓)	1068																
1123			limonene (↑)	22311																
1124			alfa-pinene	6654																
1125			2,2,6-trimethyloctane	522006																

	Pathogen	Classification	VOCs	PubChem ID	Methods	*In vivo* Sample							Bacterial strain	Culture conditions/Growth medium	Incubation time before analysis	Concentration range		Reference	Journal Code	Year
						Saliva	Blood	Breath	Skin	Urine	Faeces	Milk				Value	Unit			
1126			dodecane	8182																
1127			4-terpinenol	11230																
1128			1-undecene	13190																
1129			linalool	6549																
1130			2,6,7-trimethyldecane	43924																
1131			indole	798																
1132			toluene	1140																
1133			ethanol	702																
1134			acetoin	179																
1135			acetic acid	176																
1136			amylene hydrate	6405																
1137			caryophyllene	5322111																
1138			cis-1-p-menthanol	89437																
1139			2,5-dimethyl-2,5-hexanediol	8031																
1140			2-nonanone	13187																
1141			acetone	180	GC-MS analysis			×					unknown	Blood agar; Mannitol Salt agar; Mac-Conkey agar	24h at 37°C	unknown	unknown	Goeminne, P. C., Vandendriessche, T., Van Eldere, J., Nicolai, B. M., Hertog, M. L., & Dupont, L. J. (2012). Detection of Pseudomonas aeruginosa in sputum headspace through volatile organic compound analysis. *Respir Res , 13* , 87. doi: 10.1186/1465-9921-13-87.	Respir. Res.	2012
1142			2-ethyl-1-hexanol	7720																
1143			2-heptanone	8051																
1144			2-ethoxy-2-methylpropane	12512																
1145			2-phenylethanol	6054																
1146			1-octen-3-ol	18827																
1147			4-methyloctane	16665																
1148			isoamyl acetate	31276																
1149			limonene	22311																
1150			eucalyptol	2758																
1151			6-methyl-2-heptanone	13572																
1152			thymol	6989																
1153			2-phenylacetaldehyde	998																
1154			2-hexanone	11583																
1155			2,4-dimethyl-1-heptene	123385																
1156			5-methyl-2-(1-methylethyl)-cyclohexanone	6986																
1157			2,4-dimethylheptane	16656																
1158			pyrrolidine	31268																
1159			2,6-dimethyl-7-octen-2-ol	29096																
1160			methyl thiocyanate	11168	SIFT-MS analysis	not used	not used	not used	not used	not used	not used	not used	NIPH 2414	Shaken submersion culture of each strain in 7mL of Mueller-Hinton Broth (MHB) liquid medium (Oxoid) in a 100mL hermetically sealed flask	unknown	31	ppb	Shestivska, V., Nemec, A., Dřevínek, P., Sovová, K., Dryahina, K., & Španěl, P. (2011). Quantification of methyl thiocyanate in the headspace of Pseudomonas aeruginosa cultures and in the breath of cystic fibrosis patients by selected ion flow tube mass spectrometry. Rapid Communications in Mass Spectrometry, 25(17), 2459-2467. doi: 10.1002/rcm.5146.	Rapid Commun. Mass Spectrom.	2011
1161			hydrogen cyanide	768												654				
1162			methyl thiocyanate	11168									NIPH 2415			8				
1163			hydrogen cyanide	768												17				
1164			methyl thiocyanate	11168									NIPH 2418			16				
1165			hydrogen cyanide	768												317				
1166			methyl thiocyanate	11168									NIPH 2421			20				
1167			hydrogen cyanide	768												3849				
1168			methyl thiocyanate	11168									NIPH 2423			58				
1169			hydrogen cyanide	768												7604				
1170			methyl thiocyanate	11168									NIPH 2425			182				
1171			hydrogen cyanide	768												2367				
1172			methyl thiocyanate	11168									NIPH 2427			2				
1173			hydrogen cyanide	768												6925				
1174			methyl thiocyanate	11168									NIPH 2430			887				
1175			hydrogen cyanide	768												5912				
1176			methyl thiocyanate	11168									NIPH 2433			6				
1177			hydrogen cyanide	768												1460				
1178			methyl thiocyanate	11168									NIPH 2440			17				
1179			hydrogen cyanide	768												5110				
1180			methyl thiocyanate	11168									NIPH 2443			32				
1181			hydrogen cyanide	768												5506				
1182			methyl thiocyanate	11168									NIPH 2449			15				
1183			hydrogen cyanide	768												5521				
1184			methyl thiocyanate	11168									NIPH 2451			67				
1185			hydrogen cyanide	768												2257				
1186			methyl thiocyanate	11168									NIPH 2452			11				
1187			hydrogen cyanide	768												1282				
1188			methyl thiocyanate	11168									NIPH 2454			2				
1189			hydrogen cyanide	768												12				
1190			methyl thiocyanate	11168									NIPH 2455			12				
1191			hydrogen cyanide	768												48				
1192			methyl thiocyanate	11168									NIPH 2457			9				
1193			hydrogen cyanide	768												5998				
1194			methyl thiocyanate	11168									NIPH 2462			14				
1195			hydrogen cyanide	768												46				
1196			methyl thiocyanate	11168									NIPH 2463			12				
1197			hydrogen cyanide	768												14				
1198			methyl thiocyanate	11168									NIPH 2471			1784				
1199			hydrogen cyanide	768												2763				
1200			methyl thiocyanate	11168									NIPH 2472			5				
1201			hydrogen cyanide	768												49				
1202			methyl thiocyanate	11168									NIPH 2473			6				
1203			hydrogen cyanide	768												125				
1204			methyl thiocyanate	11168									NIPH 2480			260				
1205			hydrogen cyanide	768												7611				
1206			methyl thiocyanate	11168									NIPH 2491			306				
1207			hydrogen cyanide	768												4783				
1208			methyl thiocyanate	11168									NIPH 2492			12				
1209			hydrogen cyanide	768												2606				
1210			methyl thiocyanate	11168									NIPH 2495			41				
1211			hydrogen cyanide	768												202				
1212			methyl thiocyanate	11168									NIPH 2496			2				
1213			hydrogen cyanide	768												2532				
1214			methyl thiocyanate	11168									NIPH 2498			2				
1215			hydrogen cyanide	768												6099				
1216			methyl thiocyanate	11168									NIPH 2502			367				
1217	*Pseudomonas aeruginosa*	Gram negative bacterium	hydrogen cyanide	768												1104				
1218			methyl thiocyanate	11168									NIPH 2507			30				
1219			hydrogen cyanide	768												488				
1220			methyl thiocyanate	11168									NIPH 2511			61				
1221			hydrogen cyanide	768												3502				
1222			methyl thiocyanate	11168									NIPH 2512			256				
1223			hydrogen cyanide	768												1658				
1224			methyl thiocyanate	11168									ANC 3157			10				
1225			hydrogen cyanide	768												16				
1226			methyl thiocyanate	11168									ANC 3209			3				
1227			hydrogen cyanide	768												64				
1228			methyl thiocyanate	11168									ANC 3488			313				
1229			hydrogen cyanide	768												4752				
1230			methyl thiocyanate	11168									PAO13888			83				
1231			hydrogen cyanide	768												5339				
1232			acetone	180	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	UMRL 1203	sterile urine (20 mL) from healthy males inoculated to a concentration of between 10^7 and 10^9 cfu/mL	37˚C for 6h	unknown	ppb	Storer, M. K., Hibbard-Melles, K., Davis, B., & Scotter, J. (2011). Detection of volatile compounds produced by microbial growth in urine by selected ion flow tube mass	J. Microbiol. Methods	2011
1233			ethanol	702																
1234			formaldehyde	712																
1235			methanethiol	878																
1236			2,3-dimethyl-5-sopentylpyrazine	not available																
1237			2-methy-3-(2-propenyl)-pyrazine	583833																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
						In vivo Sample										Concentration range				
	Pathogen	Classification	VOCs	PubChem ID	Methods	Saliva	Blood	Breath	Skin	Urine	Faeces	Milk	Bacterial strain	Culture conditions/Growth medium	Incubation time before analysis	Value	Unit	Reference	Journal Code	Year
1238			methyl thioacetate	519840																
1239			2-furaldehyde	7362																
1240			dimethyl trisulfide	19310																
1241			tetradecane	12389																
1242			1-undecene	13190	TD-GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 27853	Brain Heart Infusion (BHI)	16h, 24h and 48h	unknown	unknown	Neerincx, A. H., Geurts, B. P., Habets, M. F. J., Booij, J. A., van Loon, J., Jansen, J. J., ... & Wevers, R. A. (2016). Identification of Pseudomonas aeruginosa and Aspergillus fumigatus mono-and co-cultures based on volatile biomarker combinations. *Journal of breath research* , *10* (1), 016002.	J. Breath Res.	2016
1243			hexanal	6184																
1244			6-tridecane	142600																
1245			dimethyl disulfide	12232																
1246			butanal	261																
1247			3-methyl-1H-pyrrole	12023																
1248			2-methylbutanal	7284																
1249			hydrogen sulfide	402																
1250			methanethiol	878	SIFT-MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 27853	blood culture bottles	24 h	unknown	unknown	Allardyce, Randall A., et al. "Detection of volatile metabolites produced by bacterial growth in blood culture media by selected ion	J. Microbiol. Methods	2006
1251			dimethyl sulfide	1068																
1252			carbon dioxide	280																
1253			ammonia	222																
1254			methanethiol	878	IMR-MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 27853	blood agar plates	24 h	unknown	unknown	Dolch, M. E., et al. "Volatile compound profiling for the identification of Gram-negative bacteria by ion-molecule reaction–mass spectrometry." *Journal of*	J. Appl. Microbiol.	2012
1255			indole	798																
1256			2-aminoacetophenone	11952																
1257			dimethyl disulfide	12232																
1258			1-undecene	13190	SPME-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	blood agar or chocolate blood agar	48h	unknown	unknown	Preti, George, et al. "Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: analysis by solid-phase microextraction and gas chromatography–mass spectrometry." *Journal*	Clin. Microbiol. Rev.	2009
1259			2,5-dimethylpyrazine	31252																
1260			dimethyl sulfide	1068																
1261			isoprene	6557																
1262			dimethyl disulfide	12232	CG-FID analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 10145	minimal salt AB medium+ 1% citrate	overnight	unknown	unknown	Schöller, Charlotte, Søren Molin, and Ken Wilkins. "Volatile metabolites from some gram-negative bacteria." *Chemosphere* 35.7 (1997): 1487-1495.	Chemosphere	1997
1263			dimethyl trisulfide	19310																
1264			1-undecene	13190																
1265			hydrogen cyanide	768																
1266			ammonia	222																
1267			acetonitrile	6342	SIFT-MS analysis			×					clinical isolate	blood agar (BA) and Pseudomonas selective media (PSM)	48h at 37ºC	unknown	ppb	Carroll, Will, et al. "Detection of volatile compounds emitted by Pseudomonas aeruginosa using selected ion flow tube mass spectrometry." *Pediatric pulmonology* 39.5 (2005): 452-456.	Pediatr. Pulmonol.	2005
1268			dimethyl disulfide	12232																
1269			ethanol	702																
1270			2-aminoacetophenone	11952	GC-MS/ Colorimetric analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 15692	blood agar plates	20h	unknown	unknown	Cox, Charles D., and J. Parker. "Use of 2-aminoacetophenone production in	J. Clin. Microbiol.	1979
1271						unknow	unknow	unknow	unknow	unknow	unknow	unknow	10 clinical isolates							
1272			dimethyl disulfide	12232																
1273			dimethyl trisulfide	19310																
1274			2-nonanone	13187																
1275			2-undecanone	8163																
1276			methanethiol	878									ATCC 19660							
1277			2-aminoacetophenone	11952																
1278			1-butanol	263																
1279			toluene	1140																
1280			2-butanone	6569																
1281			1-undecene	13190																
1282			isopentanol	31260																
1283			dimethyl disulfide	12232																
1284			dimethyl trisulfide	19310																
1285			2-nonanone	13187																
1286			2-undecanone	8163																
1287			methanethiol	878									ATCC 27313							
1288			2-aminoacetophenone	11952																
1289			1-butanol	263																
1290			toluene	1140																
1291			2-butanone	6569																
1292			1-undecene	13190																
1293			isopentanol	31260																
1294			dimethyl disulfide	12232																
1295			dimethyl trisulfide	19310																
1296			2-nonanone	13187																
1297			2-undecanone	8163																
1298			methanethiol	878									ATCC 7700							
1299			2-aminoacetophenone	11952																
1300			1-butanol	263																
1301			toluene	1140																
1302			2-butanone	6569																
1303			1-undecene	13190																
1304			isopentanol	31260																
1305			dimethyl disulfide	12232																
1306			dimethyl trisulfide	19310																
1307			2-nonanone	13187																
1308			2-undecanone	8163																
1309			methanethiol	878									ATCC 17423							
1310			2-aminoacetophenone	11952																
1311			1-butanol	263																
1312			toluene	1140																
1313			2-butanone	6569																
1314			1-undecene	13190																
1315			isopentanol	31260																
1316			dimethyl disulfide	12232																
1317			dimethyl trisulfide	19310																
1318			2-nonanone	13187																
1319			2-undecanone	8163																
1320			methanethiol	878									ATCC 27312							
1321			2-aminoacetophenone	11952																
1322			1-butanol	263																
1323			toluene	1140																
1324			2-butanone	6569																
1325			1-undecene	13190																
1326			isopentanol	31260																
1327			dimethyl disulfide	12232																
1328			dimethyl trisulfide	19310																
1329			2-nonanone	13187																
1330			2-undecanone	8163																
1331			methanethiol	878																
1332			2-aminoacetophenone	11952	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 27316	trypticase soy agar	24h at 37ºC	unknown	unknown	Labows, JOHN N., et al. "Headspace analysis of volatile metabolites of Pseudomonas aeruginosa and related species by gas chromatography-mass spectrometry." *Journal of Clinical Microbiology* 12.4 (1980): 521-526.	J. Clin. Microbiol.	1980
1333			1-butanol	263																
1334			toluene	1140																
1335			2-butanone	6569																
1336			1-undecene	13190																
1337			isopentanol	31260																
1338			dimethyl disulfide	12232																
1339			dimethyl trisulfide	19310																
1340			2-nonanone	13187																
1341			2-undecanone	8163																
1342			methanethiol	878									ATCC 17429							
1343			2-aminoacetophenone	11952																
1344			1-butanol	263																
1345			toluene	1140																
1346			2-butanone	6569																
1347			1-undecene	13190																
1348			isopentanol	31260																
1349			dimethyl disulfide	12232																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
						In vivo Sample										Concentration range				
	Pathogen	Classification	VOCs	PubChem ID	Methods	Saliva	Blood	Breath	Skin	Urine	Faeces	Milk	Bacterial strain	Culture conditions/Growth medium	ncubation time before analysis	Value	Unit	Reference	Journal Code	Year
1350			dimethyl trisulfide	19310									ATCC 17423							
1351			2-nonanone	13187																
1352			2-undecanone	8163																
1353			methanethiol	878																
1354			2-aminoacetophenone	11952																
1355			1-butanol	263																
1356			toluene	1140																
1357			2-butanone	6569																
1358			1-undecene	13190																
1359			isopentanol	31260																
1360			dimethyl disulfide	12232									ATCC 7701							
1361			dimethyl trisulfide	19310																
1362			2-nonanone	13187																
1363			2-undecanone	8163																
1364			methanethiol	878																
1365			2-aminoacetophenone	11952																
1366			1-butanol	263																
1367			toluene	1140																
1368			2-butanone	6569																
1369			1-undecene	13190																
1370			isopentanol	31260																
1371			dimethyl disulfide	12232									CDC 9104							
1372			dimethyl trisulfide	19310																
1373			2-nonanone	13187																
1374			2-undecanone	8163																
1375			methanethiol	878																
1376			2-aminoacetophenone	11952																
1377			1-butanol	263																
1378			toluene	1140																
1379			2-butanone	6569																
1380			1-undecene	13190																
1381			isopentanol	31260																
1382			dimethyl disulfide	12232									CDC 9171							
1383			dimethyl trisulfide	19310																
1384			2-nonanone	13187																
1385			2-undecanone	8163																
1386			methanethiol	878																
1387			2-aminoacetophenone	11952																
1388			1-butanol	263																
1389			toluene	1140																
1390			2-butanone	6569																
1391			1-undecene	13190																
1392			isopentanol	31260																
1393			dimethyl disulfide	12232	HS-GLC analysis	unknow	unknow	unknow	unknow	unknow	unknow	unknow	clinical isolate	YEB medium	18h	unknown	unknow	Hayward, N. J., et al. "Development of specific tests for rapid detection of Escherichia coli and all species of Proteus in urine." *Journal of clinical microbiology* 6.3 (1977): 195-201.	J. Clin. Microbiol.	1977
1394			1,10-undecadiene	139543																
1395			1-decene	13381																
1396			1-dodecene	8183																
1397			1-nonene	31285																
1398			1-undecene	13190																
1399			10-methyl-1-undecene	519941																
1400			2,4-dimethylheptane	16656																
1401			2-methyl-2-butene	10553																
1402			isoprene	6557																
1403			n-butane	7843																
1404			undecane	14257																
1405			1-undecene	13190																
1406			ethylene glycol	174																
1407			2-butanol	6568																
1408			2-methyl-2-propanol	6386																
1409			2-pentanol	22386																
1410			isopentanol	31260																
1411			isobutanol	6560																
1412			2-methyl-1-butanol	8723																
1413			(E)-2-octenal	5283324																
1414			3-methylbutanal	11552																
1415			1-phenyl-1-butanone	10315																
1416			2,3-butanedione	650																
1417			2-nonanone	13187																
1418			2-butanone	6569																
1419			2-heptanone	8051																
1420			2-tridecanone	11622																
1421			3-decanone	13576																
1422			methyl isopropyl ketone	11251																
1423			3-methyl-2-pentanone	11262																
1424			3-methyl-3-penten-2-one	5364579																
1425			3-octanone	246728																
1426			4-heptanone	31246																
1427			methyl isobutyl ketone	7909																
1428			4-methyl-4-penten-2-one	19543																
1429			2-methylphenol	335	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	Bos, L. D., Sterk, P. J., & Schultz, M. J. (2013). Volatile metabolites of pathogens: a systematic review. doi:10.1371/journal.ppat.1003311	PLOS	2013
1430			acetophenone	7410																
1431			limonene	22311																
1432			1-phenylethanol	7409																
1433			toluene	1140																
1434			2-methylbutyl isobutyrate	97883																
1435			2-methylbutyl 2-methylbutyrate	17129																
1436			amyl isovalerate	95978																
1437			ethyl 2-methylbutyrate	24020																
1438			isoamyl butyrate	7795																
1439			isoamyl acetate	31276																
1440			methyl methacrylate	6658																
1441			methyl 2-methylbutyrate	13357																
1442			2-methoxy-5-methylthiophene	141615																
1443			3-(ethylthio)propanal	229467																
1444			dimethyl disulfide	12232																
1445			dimethyl sulfide	1068																
1446			dimethyl trisulfide	19310																
1447			mercaptoacetone	520144																
1448			methanethiol	878																
1449			methyl thioacetate	73750																
1450			1-vinyl aziridine	21843																
1451			2,4-dimethyl-quinazoline	not available																
1452			2,5-dimethylpyrazine	31252																
1453			2-(3-methylbutyl)-3,5-dimethyl-pyrazine	578812																
1454			2-aminoacetophenone	11952																
1455			2-benzoxazole	not available																
1456			3-ethyl-3-(methylthio)-pyrazine	175366																
1457			2-isopropyl-3-methylpyrazine	519203																
1458			3-methyl-1H-pyrrole	12023																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
						In vivo Sample										Concentration range				
1	Pathogen	Classification	VOCs	PubChem ID	Methods	Saliva	Blood	Breath	Skin	Urine	Faeces	Milk	Bacterial strain	Culture conditions/Growth medium	ncubation time before analys	Value	Unit	Reference	Journal Code	Year
1464			4-methylquinazoline	241520																
1465			acetonitrile	6342																
1466			ammonia	222																
1467			benzonitrile	7505																
1468			benzoxazole	9228																
1469			hydrogen cyanide	768																
1470			methyl thiocyanate	11168																
1471			3-methylbutanoic acid	10430	SPME-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	blood agar or chocolate blood agar	48h	unknown	unknown	Preti, George, et al. "Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: analysis by solid-phase microextraction and gas chromatography–mass spectrometry." *Journal of Chromatography B* 877.22 (2009): 2011-2018.	Clin. Microbiol. Rev.	2009
1472			2-methylbutanoic acid	8314																
1473			2-methylpropanoic acid	6590																
1474			1-hydroxy-2-propanone	8299																
1475			acetoin	179																
1476			butanoic acid	264																
1477			4-methylhexanoic acid	15271																
1478			2-phenylethanol	6054																
1479			ethanol	702	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	ATCC 25923	blood culture bottles	24 h	unknown	unknown	Allardyce, Randall A., et al. "Detection of volatile metabolites produced by bacterial growth in blood culture media by selected ion flow tube mass spectrometry (SIFT-MS)." *Journal of microbiological methods* 65.2 (2006): 361-365.	J. Microbiol. Methods	2006
1480			acetaldehyde	177																
1481			hydrogen sulfide	402																
1482			methanethiol	878																
1483			trimethylamine	1146																
1484			indole	798																
1485			2-aminoacetophenone	11952																
1486			hexanal	6184																
1487			ethanol (↑)	702	MCC-IMS analysis	not used	not used	not used	not used	not used	not used	not used								
1488			3-methylbutanal (↑)	11552	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	DSM 13661	Columbia Sheep blood agar	24 h at 37 °C	unknown	unknown	Jünger, M., Vautz, W., Kuhns, M., Hofmann, L., Ulbricht, S., Baumbach, J. I., ... & Perl, T. (2012). Ion mobility spectrometry for microbial volatile organic compounds: a new	Appl. Microbiol. Biotechnol.	2012
1489			isopentanol (↑)	31260																
1490			2-(methylthio)-ethanol (↑)	78925																
1491			tridecen-2-one	53427438	HS-SPME-GC-FID analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	50 mL TS broth in 125 mL vial.	ubated with shaking for 22-2	unknown	unknown			
1492			dimethyl disulfide	12232																
1493			acetic acid	176		unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	VF (peptone, NaCl) and VL broth (casein hydrolysate, yeast extract, beef extract, cysteine, glucose, NaCl). 1 mL	48 h incubations	unknown	unknown			
1494			3-methylbutanoic acid	10430																
1495			ethanol	702	HS-SPME-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	10 mL Mueller Hinton broth in 25 mL vials	up to 10 h at 37ºC without shaking	unknown	unknown			
1496			1-propanol	1031																
1497			isopentanol	31260																
1498			acetaldehyde	177																
1499			propanal	527																
1500			3-methylbutanal	11552																
1501			2-phenylacetaldehyde	998																
1502			2,3,4,5-tetrahydropyridazine	not available																
1503			acetaldehyde	177		unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	samples infected with bacteria; 5 mL sample in 20 mL vial	14 h at 37ºC	unknown	unknown	Tait, E., Perry, J. D., Stanforth, S. P., & Dean, J. R. (2014). Identification of volatile organic compounds produced by bacteria using HS-SPME-GC–MS. Journal of chromatographic science, 52(4), 363-373. doi:10.1093/chromsci/bmt042	J. Chromatogr. Sci.	2014
1504			2,3-butanedione	650																
1505			2-butanone	6569																
1506			ethyl acetate	8857																
1507			3-methylbutanal	11552																
1508			2-methylbutanal	7284																
1509			2-pentanone	7895																
1510			acetic acid	176																
1511			isopentanol	31260																
1512			acetoin	179																
1513			ethyl butanoate	7762																
1514			ethyl isovalerate	7945																
1515			ethyl 2-methylbutyrate	24020																
1516			butanoic acid	264																
1517			3-methylbutanoate	3587356																
1518			2-methylbutanoate	22253297																
1519			2-heptanone	8051																
1520			butyl butyrate	7983																
1521			ethyl hexanoate	31265																
1522			3-methylbutanoic acid	10430	SPME- GC–MS/GC–flame photometric detector anal	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	Blood agar and chocolate blood agar.	ples incubated for at least 4	unknown	unknown			
1523			2-methylbutanoic acid	8314																
1524			2-methylpropanoic acid	6590																
1525			1-hydroxy-2-propanone	8299																
1526			acetoin	179																
1527			butanoic acid	264																
1528			4-methylhexanoic acid	15271																
1529			2-phenylethanol	6054																
1530			1-butanol	263	SIFT-MS analysis	not used	not used	not used	not used	not used	not used	not used	NCTC 6571	Frozen stocks (-80˚C), Resuscitated onto nutrient agar; incubated at 37˚C aerobically when required	5h and 24h	93.49	ppb	Thorn, R., Reynolds, D. M. and Greenman, J. (2011) Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro.Journal of Microbiological Methods, 84 (2). pp. 258-264. ISSN 0167-7012 DOI: 10.1016/j.mimet.2010.12.001	J. Microbiol. Methods	2010
1531			ethanol	702												110.9				
1532			ethyl butanoate	7762												271.77				
1533			formaldehyde	712												17.26				
1534			hydrogen sulfide	402												1493.36				
1535			indole	798												2362.62				
1536			methanethiol	878												2.97				
1537			phenylacetic acid	999												1566.82				
1538			pyrrole	8027												12.74				
1539			trimethylamine	1146												0.72				
1540			1-butanol	263									ATCC 8325			114.92				
1541			1-pentanol	6276												26.1				
1542			ethanol	702												33.9				
1543			formaldehyde	712												87.11				
1544			hydrogen sulfide	402												1080.61				
1545			methanethiol	878												332.46				
1546			acetone	180	GC-*tof* -MS analysis	not used	not used	not used	not used	not used	not used	not used	clinical isolate	blood agar plates and incubated overnight at 37˚C; transfer to sterile Brain Heart Infusion broth, growth for 4h with constant agitation at the same temperature	unknown	unknown	unknown	Boots AW, Smolinska A, van Berkel JJ, Fijten RR, Stobberingh EE, et al. (2014) Identification of microorganisms based on headspace analysis of volatile organic compounds by gas chromatography-mass spectrometry. J Breath Res 8: 027106. doi:10.1088/1752-7155/8/2/027106	J. Breath Res.	2014
1547			3-methylbutanal (↓)	11552												1131.22				
1548			2-methylbutanal (↓)	7284																
1549			dimethyl disulfide	12232																
1550			2,3,3-trimethylpentane (↓)	11215																
1551			benzaldehyde (↓)	240																
1552			2,3-butanedione (↓)	650																
1553			1,1,2,2-tetrachloroethane	6591																
1554			dimethyl trisulfide	19310																
1555			1-methyl-4-(1-methylethenyl)cyclohexane (↓	14299																
1556			propanal	527	GC-MS analysis	not used	not used	not used	not used	not used	not used	not used	ATCC 25923	inoculated in a 4 ml liquid preculture and grown	1,5h-28,0 h	unknown	ppt-ppm	Filipiak, Wojciech, et al. "Molecular analysis of volatile metabolites released specifically by Staphylococcus aureus and Pseudomonas aeruginosa." *BMC microbiology* 12.1 (2012): 113.	BMC Microbiol.	2012
1557			2-ethylacrolein	70203																
1558			(E)-2-methyl-2-butenal	5321950																
1559			benzaldehyde (↓)	240																
1560			methacrolein	6562																
1561			acetaldehyde	177																
1562			3-methylbutanal	11552																
1563			2-methylpropanal	6561																
1564			1-butanol	263																
1565			isobutanol	6560																
1566			isopentanol	31260																
1567			ethanol	702																
1568			acetoin	179																
1569			1-hydroxy-2-propanone	8299																
1570			2,3-butanedione	650																
1571			acetic acid	176																
1572			3-methylbutanoic acid	10430																
1573			ethyl acetate	8857																
1574			n- butyl acetate	31272																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Pathogen	Classification	VOCs	PubChem ID	Methods	*In vivo* Sample							Bacterial strain	Culture conditions/Growth medium	Incubation time before analysis	Concentration range		Reference	Journal Code	Year
2						Saliva	Blood	Breath	Skin	Urine	Faeces	Milk				Value	Unit			
1571			ethyl isovalerate	7945																
1572			isoamyl acetate	31276																
1573			ethyl formate	8025																
1574	*Staphylococcus aureus*	Gram positive bacterium	methyl methacrylate	6658																
1575			methanethiol	878																
1576			dimethyl disulfide	12232																
1577			1,3-butadiene	7845																
1578			2-methylpropene	8255																
1579			n-butane	7843																
1580			(Z)-2-butene	5287573																
1581			(E)-2-butene	62695																
1582			propane	6334																
1583			acetonitrile	6342	SESI-MS analysis	not used	not used	not used	not used	not used	not used	not used	unknown	unknown	unknown	unknown	unknown	Sohrabi M, Zhang L, Zhang K, Ahmetagic A, Wei MQ (2014) Volatile Organic Compounds as Novel Markers for the Detection of Bacterial Infections. Clin Microbial 3: 151. doi:10.4172/2327-5073.1000151	J. Clin. Microbiol.	2014
1584			ethanol	702																
1585			1-butanol	263																
1586			acetone	180																
1587			acetic acid	176																
1588			ethylene glycol	174																
1589			isopentanol	31260																
1590			pyrimidine	9260																
1591			2-pentanone	7895																
1592			4-methylphenol	2879																
1593			2-nonanone	13187																
1594			acetaldehyde (↑)	177	TD-GC-MS analysis			×					clinical sample	blood agar, chocolate agar and macConkey agar plates; overnight incubation	unknown	unknown	ppb	Filipiak, W., Beer, R., Sponring, A., Filipiak, A., Ager, C., Schiefecker, A., ... & Amann, A. (2015). Breath analysis for in vivo detection of pathogens related to ventilator-associated pneumonia in intensive care patients: a prospective pilot study. *Journal of breath research* , *9* (1), 016004. DOI: 10.1088/1752-7155/9/1/016004	J. Breath Res.	2015
1595			2-methylpropene (↑)	8255																
1596			n-butane (↑)	7843																
1597			2-pentanone	7895																
1598			propanal (↑)	527																
1599			ethyl acetate (↑)	8857																
1600			methyl vinyl ketone	6570																
1601			hexanal	6184																
1602			1,3-butadiene (↑)	7845																
1603			benzaldehyde (↓)	240																
1604			4-heptanone	31246																
1605			dimethyl sulfide	1068																
1606			ethanol (↑)	702																
1607			propane (↑)	6334																
1608			3-methylbutanal (↑)	11552																
1609			methacrolein (↑)	6562																
1610			(Z)-2-butene (↑)	5287573																
1611			acetic acid (↑)	176																
1612			(E)-2-butene (↑)	62695																
1613			2,3-butanedione (↑)	650																
1614			carbon disulfide	6348																
1615			2-methylpropanal (↑)	6561																
1616			3-methyl-2-butenal (↑)	61020																
1617			butanal	261																
1618			n-propyl acetate	7997																
1619			(E)-2-pentene	5326161																
1620			acetoin (↑)	179																
1621			(E)-2-methyl-2-butenal (↑)	5321950																
1622			(Z)-2-methyl-2-butenal (↑)	10336																
1623			1-butanol (↑)	263																
1624			1-propanol	1031																
1625			2-ethylacrolein (↑)	70203																
1626			2-methylbutyl acetate	12209																
1627			isobutanol (↑)	6560																
1628			2-methyl-2-butene	10553																
1629			isopentanol (↑)	31260																
1630			dimethyl disulfide	12232																
1631			ethyl formate (↑)	8025																
1632			ethyl isovalerate (↑)	7945																
1633			ethyl butanoate	7762																
1634			1-hydroxy-2-propanone (↑)	8299																
1635			isoamyl butyrate	7795																
1636			isoamyl propionate	7772																
1637			isobutyl acetate	8038																
1638			isoamyl acetate (↑)	31276																
1639			3-methylbutanoic acid (↑)	10430																
1640			methanethiol (↑)	878																
1641			methyl methacrylate (↑)	6658																
1642			n- butyl acetate (↑)	31272																
1643			ethyl vinyl ether	8023																
1644			formaldehyde	712	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	NCTC 7447	sterile urine (20 mL) from healthy males inoculated to a concentration of between 10^7 and 10^9 cfu/mL	37˚C for 6h	unknown	ppb	Storer, M. K., Hibbard-Melles, K., Davis, B., & Scotter, J. (2011). Detection of volatile compounds produced by microbial growth in urine by selected ion flow tube mass	J. Microbiol. Methods	2011
1645			2-methylbutanal	7284																
1646			methanethiol	878																
1647			ammonia	222																
1648			(E)-2-butene	62695	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	Bos, L. D., Sterk, P. J., & Schultz, M. J. (2013). Volatile metabolites of pathogens: a systematic review. doi:10.1371/journal.ppat.1003311	PLOS	2013
1649			2-methylpropene	8255																
1650			n-butane	7843																
1651			propane	6334																
1652			ethylene glycol	174																
1653			2-butanol	6568																
1654			isobutane	6360																
1655			isopentanol	31260																
1656			ethanol	702																
1657			isobutanol	6560																
1658			2-methyl-1-butanol	8723																
1659			acetic acid	176																
1660			3-methylbutanoic acid	10430																
1661			phenylacetic acid	999																
1662			2-ethylacrolein	70203																
1663			(E)-2-methyl-2-butenal	5321950																
1664			2-methylbutanal	7284																
1665			methacrolein	6562																
1666			3-methylbutanal	11552																
1667			acetaldehyde	177																
1668			benzaldehyde	240																
1669			hexanal	6184																
1670			2,3-butanedione	650																
1671			2-nonanone	13187																
1672			2-heptanone	8051																
1673			1-hydroxy-2-propanone	8299																
1674			toluene	1140																
1675			n- butyl acetate	31272																
1676			ethyl formate	8025																
1677			ethyl isovalerate	7945																
1678			methyl methacrylate	6658																
1679			2-(methylthio)-ethanol	78925																
1680			acetonitrile	6342																
1681			ammonia	222																
1682			pyrimidine	9260																
1683			acetone	180																

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1						*In vivo* Sample										Concentration range				
2	Pathogen	Classification	VOCs	PubChem ID	Methods	Saliva	Blood	Breath	Skin	Urine	Faeces	Milk	Bacterial strain	Culture conditions/Growth medium	ncubation time before analys	Value	Unit	Reference	Journal Code	Year
1683			2-butanone	6569																
1684			2-pentanone	7895																
1685			2-hexanone	11583																
1686			acetaldehyde	177																
1687			2-methylbutanal	7284																
1688	*Staphylococcus epidermidis*	Gram positive bacterium	ethyl acetate	8857	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	ATCC 14990	sterile urine (20 mL) from healthy males inoculated to a concentration of between 10^7 and 10^9 cfu/mL	37˚C for 6h	unknown	ppb	Storer, M. K., Hibbard-Melles, K., Davis, B., & Scotter, J. (2011). Detection of volatile compounds produced by microbial growth in urine by selected ion flow tube mass spectrometry (SIFT-MS). Journal of microbiological methods, 87(1), 111-113. doi: 10.1016/j.mimet.2011.06.012.	J. Microbiol. Methods	2011
1689			n-propyl acetate	7997																
1690			hydrogen sulfide	402																
1691			dimethyl disulfide	12232																
1692			dimethyl sulfide	1068																
1693			ammonia	222																
1694			indole	798																
1695			benzaldehyde	240																
1696			phenylmethanol	244																
1697			2-phenylethanol	6054	SPME-GC-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	clinical isolate	blood agar or chocolate blood agar	48h	unknown	unknown	Preti, George, et al. "Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: analysis by solid-phase microextraction and gas chromatography–mass spectrometry." *Journal*	Clin. Microbiol. Rev.	2009
1698			acetic acid	176																
1699			methanethiol	878																
1700			acetaldehyde (↑)	177																
1701			2-butenal (↑)	447466																
1702			propanal (↑)	527																
1703			2-methylpropanal (↑)	6561																
1704			3-methyl-2-butenal (↑)	61020																
1705			butanal (↑)	261																
1706			1-butanol (↑)	263																
1707			ethanol (↑)	702																
1708			2,3-butanedione (↑)	650																
1709			2-pentanone (↑)	7895																
1710			2-nonanone (↑)	13187																
1711			2-butanone (↑)	6569																
1712			acetone (↑)	180																
1713			acetic acid (↑)	176																
1714			methyl methacrylate (↑)	6658																
1715			ethyl acetate (↑)	8857	GC-MS analysis			×					nical isolate (later identified as ATCC 4961	blood agar plates and tryptic soy broth	3, 3.75, 4.5, 6 and 7.5 h after inoculation	unknown	ppt-ppm	W. Filipiak, A. Sponring, M. M. Baur, C. Ager, A. Filipiak, H. Wiesenhofer, M. Nagl, J. Troppmair, A. Amann. Characterization of volatile metabolites taken up by or released from Streptococcus pneumoniae and Haemophilus influenzae by using GC-MS. Microbiology 2012, 158, 3044. doi: 10.1099/mic.0.062687-0.	Microbiol.	2012
1716			methanethiol (↑)	878																
1717			dimethyl disulfide (↑)	12232																
1718			carbon disulfide (↑)	6348																
1719			dimethyl trisulfide (↑)	19310																
1720			dimethyl sulfide (↑)	1068																
1721			1,3-butadiene (↑)	7845																
1722	*Streptococcus pneumoniae*	Gram positive bacterium	2-methylpropene (↑)	8255																
1723			(E)-2-butene (↑)	62695																
1724			(Z)-2-butene (↑)	5287573																
1725			2-methyl-1-butene (↑)	11240																
1726			furan (↑)	8029																
1727			3-(methylthio)propanal (↑)	18635																
1728			ethylbenzene (↑)	7500																
1729			3-phenylfuran (↑)	518802																
1730			3-methylbutanal (↓)	11552																
1731			hexanal (↓)	6184																
1732			ethanol	702																
1733			formaldehyde	712																
1734			acetaldehyde	177																
1735			dimethyl sulfide	1068	SIFT-MS analysis	unknown	unknown	unknown	unknown	unknown	unknown	unknown	ATCC 49619	blood culture bottles	24 h	unknown	unknown	Allardyce, Randall A., et al. "Detection of volatile metabolites produced by bacterial growth in blood culture media by selected ion flow tube mass spectrometry (SIFT-MS)." *Journal of microbiological methods* 65.2 (2006): 361-365.	J. Microbiol. Methods	2006
1736			trimethylamine	1146																
1737			indole	798																
1738			2-aminoacetophenone	11952																
1739			hexanal	6184																
1740			ethanol	702																
1741			3-methylbutanal	11552																
1742			acetaldehyde	177																
1743			formaldehyde	712																
1744			hexanal	6184																
1745			acetone	180	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown	Bos, L. D., Sterk, P. J., & Schultz, M. J. (2013). Volatile metabolites of pathogens: a systematic review. doi:10.1371/journal.ppat.1003311	PLOS	2013
1746			2-pentylfuran	19602																
1747			dimethyl sulfide	1068																
1748			2-aminoacetophenone	11952																
1749			benzonitrile	7505																
1750			trimethylamine	1146																