Investigation of new actinobacteria for the biodesulphurisation of diesel fuel

Selva Manikandan Athi Narayanan

A thesis submitted in partial fulfilment of the requirements of Edinburgh Napier University, for the award of Doctor of Philosophy

May 2020

Abstract

Biodesulphurisation (BDS) is an emerging technology that utilises microorganisms for the removal of sulphur from fossil fuels. Commercial-scale BDS needs the development of highly active bacterial strains which allow easier downstream processing. In this research, a collection of actinobacteria that originated from oil-contaminated soils in Russia were investigated to establish their phylogenetic positions and biodesulphurisation capabilities.

The eleven test strains were confirmed as members of the genus *Rhodococcus* based on 16S rRNA and *gyrB* gene sequence analysis. Two organisms namely strain F and IEGM 248, confirmed as members of the species *R. qingshengii* and *R. opacus*, respectively based on the whole-genome sequence based OrthoANIu values, exhibited robust biodesulphurisation of dibenzothiophene (DBT) and benzothiophene (BT), respectively. *R. qingshengii* strain F was found to convert DBT to hydroxybiphenyl (2-HBP) with DBTO and $DBTO_2$ as intermediates. The DBT desulphurisation genes of strain F occur as a cluster and share high sequence similarity with the *dsz* operon of *R. erythropolis* IGTS8. *Rhodococcus opacus* IEGM 248 could convert BT into benzofuran. The BDS reaction of both strains follows the well-known 4S pathway of desulphurisation of DBT and BT.

When cultured directly in a biphasic growth medium containing 10% (v/v) oil (*n*-hexadecane or diesel) containing 300 ppm sulphur, strain F formed a stable oil-liquid emulsion, making it unsuitable for direct industrial application despite the strong desulphurisation activity. Whereas the strain 248 formed distinct oil, biomass and aqueous phases which enabled easy extraction of the desulphurised oil with more than 80% reduction in total sulphur content, making it a potential candidate strain for the development of a robust BDS biocatalyst to upgrade crude oils and refinery streams at industrial scale.

Whole-genome analysis of these strains also revealed the presence of a high copy number of various monooxygenases and sulphur metabolism related genes that occurred in clusters. These genes offer potential target sites for future mutation strategies to forestall sulphate induced repression of the desulphurisation genes which warrants future investigation.

Declaration

I declare that this thesis is a result of my own independent work and all the work presented in it was undertaken and written by me. I also declare that no part of this work has been submitted for any other degree or professional qualification.

Selva Manikandan Athi Narayanan Date 25/05/2020

Dedication

To my beloved parents and my dear brother who funded my research and have been supportive of all my experiments with life, seeing me through successes and failures

Acknowledgement

I thank my Director of Studies, Dr Fiona M Stainsby for her comments, criticisms and guidance for the entirety of the research work and for steering me through university regulatory aspects regarding the course. Her profound feedback on my thesis chapters was hugely helpful in structuring the thesis to its form.

I would like to thank my co-supervisor Prof Nick Christofi, Recircle Ltd, who has a key role in conceiving the idea for this project, providing the bacteria and persistent guidance throughout the project in terms of the research activities and great moral support.

My sincere thanks to Ms Patricia Gonzalez-Iglesias and her team at the Microbiology Laboratory, School of Applied Science for their daily support and timely services during my laboratory-ridden days. I thank Ms Sadie Kemp for providing consultation regarding PCR techniques and the use of the thermal cycler. I thank Dr Sonja Rueckert for demystifying the procedures related to gene sequencing. I thank Mr Chris Edmans for putting up with my standard and at times strange purchase requests and especially for his patience and cooperation while dealing with sequencing related results that were sometimes sent to his inbox. The little succours from these helpful colleagues made me feel positive and reassured while trying out those techniques for the first time.

I am immensely thankful to Ms Tatyana Peshkur and Ms Mara Knapp at the University of Strathclyde, for their role in providing consultation and demonstration of the analytical techniques used in this research. I thank Mr Chris Gould and Ms Louise MacNicol at Petroineos Refining, Grangemouth, UK for providing with HDS stream gas oil used in this research.

I am very grateful to my lovely wife, Mrs Divyapradeepa, for the sacrifices she made, the love she showered, the care and support she offered during my long student life.

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1 Literature Review

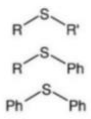
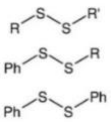
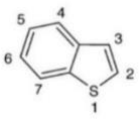
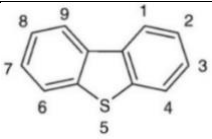
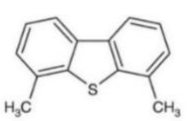
Fossil fuels are used on a vast scale throughout the world as an energy source for domestic and industrial use, and for transportation. The dependency on these finite fossil fuels is so fundamental to the growth and development of nations that humankind continues to rely on this consumption despite knowing the seriously deleterious environmental and health impacts associated with their use. However, growing environmental awareness has resulted in increased regulation of fuel quality, particularly regarding the harmful components inherently present in them. Sulphur, one of the key elements present in fossil fuels, is released during combustion as sulphoxide gases that pose hazards to human health and the environment (WHO, 2007). With increasingly stringent legislation that limits the permitted levels of sulphur in fossil fuels, considerable effort has been made to research and develop methods to minimise or eliminate sulphur from fossil fuels (El-Gendy and Nassar, 2018b). Current desulphurisation technologies rely on energy-intensive catalytic chemical processes such as hydrodesulphurisation (HDS), which is costly and ineffective against the desulphurisation of polycyclic aromatic compounds (PACs). This has necessitated exploration into alternative desulphurisation technologies. Biodesulphurisation (BDS) technology utilises microbial enzymatic processes to remove sulphur from fossil fuels, with the advantage of relatively lower pollution and operational costs. However, this is not a full-scale technology for commercial processing even though several microorganisms capable of BDS have been discovered.

In this literature review, the compelling reasons for fossil fuel desulphurisation are presented along with a critical review of the technologies that are being used or developed to achieve the same. BDS technology and the mechanisms of BDS reaction are described, followed by a review of previously reported microorganisms capable of BDS activity. The existing body of research is critically examined to understand the significant challenges associated with the development of BDS technology on a commercial scale. Finally, potential avenues to address some of the crucial issues around the development of robust biocatalysts capable of efficient desulphurisation of fossil fuels are explored. With a clear understanding of the current research trends and

methodology used for the development of BDS catalysts, the priorities and specific aims of this research are defined at the end of the review.

1.1 Overview of sulphur compounds present in crude oil

Petroleum is a broad term that includes crude oil and petroleum products. Crude oil is a complex mixture of hydrocarbons, organic and inorganic compounds, with trace amounts of metals such as vanadium and nickel that exists as a liquid in underground geologic formations and remains a liquid when brought to the surface. Petroleum products can be produced by the processing of crude oil or other forms of fossil fuels such as natural gas or synthesised by blending different finished petroleum products. Liquids produced at natural gas processing plants are not included within the term crude oil. The chemical composition and physical properties of crude oil vary significantly based on the location, origin and the type of crude. After carbon and hydrogen elements, sulphur is the third most abundant element in crude oils and can account for 0.05% to 14% (w/w) (Hamme *et al.*, 2003). Sulphur in crude oil exists as elemental sulphur, hydrogen sulphide, pyrites and organic forms like thiols, sulphides, disulphides, thiolanes and thiophenes such as benzothiophene (BT), dibenzothiophene (DBT) and their alkylated derivatives (Kilbane and Le Borgne, 2004; Kropp and Fedorak, 1998; Soleimani *et al.*, 2007) as shown in Figure 1.1. Crude oil containing high levels of sulphur – more than 0.5% are classified as ‘sour’, whereas those with relatively low levels of sulphur are termed ‘sweet'. Sweet crude is the preferred form of crude because it is safer to extract and transport, and easier to refine than sour crude, and yields a higher proportion of the more valuable final petroleum products (Fattouh, 2010). Estimates of the sulphur content in crude oil vary from 1000 to 30,000 ppm. Typical sulphur concentrations in ‘straight run’ diesel (produced by simple distillation of crude oil with reasonably low aromatic hydrocarbons) often exceed 5000 ppm (Monticello, 2000). Due to the corrosive nature of sulphur, sour crude causes more damage to machinery in refineries and thus results in higher maintenance costs over time. The Carbon-Sulphur (C-S) bonds contribute to the high molecular weight and the inherent viscosity which makes crude oil processing both energy-intensive and costly for transportation through pipelines (Alomair *et al.*, 2015; Kirkwood *et al.*, 2005).

			Thiophenes		
Thiols	**Sulphides**	**Disulphides**	**Benzothiophene (BT)**	**Dibenzothiophene (DBT)**	**Alkylated DBT (alkyl DBT)**
					

R, R' – Alkyl groups; Ph – Phenolic groups

Figure 1.1 Various forms of organic sulphur compounds in crude oil

After extraction from the oil wells, crude oil is transported to oil refineries for fractional distillation into petroleum products based on the range of boiling points, with each fraction containing different levels of sulphur compounds as shown in Table 1.1. The distillates are then subjected to 'hydroprocessing' to meet environmental specifications, followed by 'catalytic cracking' (breaking up of large hydrocarbon molecules into smaller molecules using a combination of heat and catalytic action) and the final step of 'blending' of the various hydrocarbons components as per customer requirements (Speight, 2014). Gas oil is the technical name for the same petroleum product, which is called diesel by consumers, but the difference is the government tax associated with them. Over the years of usage, diesel engines are known for high efficiency, durability, and reliability together with their low-operating cost, which makes them the most preferred engines especially for heavy-duty vehicles (Reşitoğlu *et al.*, 2015). The popularity of the diesel engines results in ever increasing demand for middle distillate (diesel) fuels, with a share of 36.1% of the total global oil consumption in 2018 (BP, 2019). Among the petroleum products obtained from crude oil, the middle distillate fractions contain a significant amount of sulphur (Table 1.1). Sulphur present in fuels inhibits the proper functioning of after-treatment systems such as diesel particulate filters, lean NOx traps, and selective catalytic reduction that are designed to reduce tailpipe emissions in vehicles.

Table 1.1. Percentage distribution of sulphur compounds in crude oil with 1.2% total sulphur.

Distillation range (°C)	Sulphur content of the fraction (%)	Sulphur compound distribution (%)			
		Thiols	Sulphides	Thiophenes	BT, DBT & their alkyl derivatives
70–180 (light distillate, naphtha)	0.02	50	50	Trace	-
160–240 (kerosene)	0.2	25	25	35	15
230–350 (middle distillate, gas oil, diesel)	0.9	15	15	35	35
350–550 (vacuum gas oil)	1.8	5	5	30	60
> 550 (vacuum residue)	2.9	Trace	Trace	10	90

The table, adapted from Javadli & Klerk (2012), shows the representative sulphur levels present in the various distillates (fractions) obtained from a sample of crude oil with 1.2% total sulphur content. Within each fraction, light distillates comprising of aviation and motor gasoline and naphtha have a low sulphur content occurring as mostly thiols and sulphides, and only trace levels of thiophenes. The popular middle distillate fractions have higher 0.9% sulphur content, the majority of which is contributed by thiophenes, BT, DBT and alkylated derivatives.

1.2 Impact of sulphoxide emission

Sulphur is released into the atmosphere by natural processes such as volcanic eruptions; however, anthropogenic activities are the primary source of atmospheric sulphur pollution. When naturally existing fossil fuels are burned, sulphoxide gases (SO_x), primarily sulphur dioxide (SO_2) and sulphur trioxide (SO_3), are released. Sulphur dioxide is a colourless gas with a pungent, irritating odour and taste. Among the sulphoxides (SO, SO_2, SO_3), SO_2 is the major component and draws the maximum concern, and so SO_2 is used as the indicator in the literature to refer to all the sulphoxides. In the UK, fossil fuel combustion in energy and transformation industries accounted for 35.2% of total SO_2 emissions in 2017, while that from manufacturing industries and transportation-related sources was 22.5% and 9.2%, respectively (NAEI, 2018).

Plants are very susceptible to SO_2, with levels of 1-2 ppm causing immediate damages, possibly by inhibiting photosynthesis (Schmidt *et al.*, 1973). In animals, the respiratory system and the eyes are the primary organs affected by air pollution. SO_2 inhalation may produce an urgent irritant effect on the

respiratory mucosa, and overexposure can lead to inflammation of the respiratory epithelium, bronchoconstriction, dyspnoea and other respiratory tract infections which lead to poor health and ultimately increased healthcare costs (Balmes *et al.*, 1987; Ghozikali *et al.*, 2015; Ko *et al.*, 2007; Medley *et al.*, 2002). SO_2 deposition also occurs as acid rain (wet deposition) and as gas and particles suspended in the atmosphere (dry deposition; Griffith, 2009). Acid rain accelerates the erosion of historical buildings, monuments and other artefacts exposed to the atmosphere and agricultural crops that have a direct impact on the economy (Bender and Weigel, 2011; Larssen *et al.*, 2006). The acid is transferred to soil, where it damages foliage and agricultural crops, with a direct impact on the economy. Acid depresses the pH of lakes with low buffering capacity and endangers marine life (Breeze, 2017). Sulphuric acid deposition in areas where soils have been degraded due to deforestation depletes the levels of free Ca^{2+}, Mg^{2+}, and K^{+} cations which are essential nutrients for plants. Insoluble Aluminium compounds are leached and converted into soluble Al^{3+} and $AlOH^{2+}$ ions under acidic conditions. These soluble ions are toxic to terrestrial plants and animals. As the forests and ponds are increasingly acidified by acid rain, the lifecycle of animals such as anurans dependent on them are also severely affected (Antal and Puttonen, 2006). In mining areas, inadvertent cracks are on the aquiclude, through which acid rain percolates and may reduce the strength of the weakest rocks and consequently lead to landslides (Zhang and McSaveney, 2018). SO_2 combines with dust to form sulphate aerosols that caused haze problems in China and India (Li *et al.*, 2017). Indeed, sulphate aerosols have been correlated to the reduction in sunlight reaching earth's surface causing a ‘global dimming’ effect which influences the global climate and life on our planet (Stern, 2006; Wild, 2014).

1.3 Standard sulphur limits in fuel

When the legislation limiting sulphur levels was initially introduced, not all countries engaged readily with this change. It took multinational efforts and formation of international agencies to set the limits which have been made increasingly stringent over the period since their introduction. In Europe, a timeline for achieving lower sulphur levels in diesel fuels was established in

the 1999 Gothenburg Protocol that set emission targets to abate acidification. The Directive 1999/32/EC required the sulphur content in diesel to be decreased gradually from 2000 ppm in the year 1994 to 350 ppm in 2000 (EURO III), and subsequently to <50 ppm in 2005 (EURO IV) and to <10 ppm in the year 2010 (EURO V). Because of its historical importance since the industrial revolution, monitoring of SO_2 has been extensive in Europe, and a long-term database of 24-hour SO_2 measurements is available. Most countries use the European (EURO) standards to measure sulphur in fuels but use different names to describe them. The current Ultra-Low Sulphur (ULS) fuel specification around the world include Bharat Stage 6 (BS-6) in India, China VI in China, Tier III in the US, AFRI 4 in Africa, among others. The latest environmental regulations, such as the TIER III and EURO V demand a maximum of 10 ppm total sulphur in diesel oils. In Japan sulphur limits were drastically reduced from 2,000 ppm in 1992 to 500 ppm in 1997, followed by a further 10 fold reduction in 2004, and by 2007 it was reduced to 10 ppm. As of 2017, India required on-road diesel and gasoline nationwide to meet BS-4 specifications (<50 ppm sulphur), which will tighten to BS-6 (<10 ppm sulphur) by 2020. The maximum of 10 ppm sulphur limit is also included in the Chinese norms GB 19147-2016 for diesel fuels deriving from the CHINA V environmental program which was fully implemented at the end of 2017. As of 2017, 11 African countries Morocco, Mauritius, Kenya, Uganda, Tanzania, Rwanda, Burundi, Ghana, Mozambique, Malawi and Zimbabwe moved to low-sulphur fuels (AFRI 4 - equivalent to EURO 4), with Morocco and Tunisia having implemented ULS fuels already. The push for lower-sulphur transportation fuels has also impacted the global shipping sector, as well. After years of unwillingness to implement stringent sulphur limits on marine fuels, the International Maritime Organization (IMO) made a landmark decision to limit sulphur from the current level of 3.5% to 0.5% in marine fuels globally from the year 2020 (IMO, 2016).

1.4 Technologies for desulphurisation of crude oil-derived fuels

Adhering to low-sulphur and ULS standards is capital-intensive for refiners, as it requires upgrades and new units to the existing refineries to meet Euro IV and Euro V standards. This has driven the need for improvements in

hydrodesulphurisation technology and the development of alternative desulphurisation technologies.

1.4.1 Hydrodesulphurisation (HDS)

Currently, Hydrodesulphurisation (HDS) in combination with carbon rejection technologies, such as coking and fluid catalytic cracking (FCC), is the most common technology in the petroleum industry to reduce the sulphur content of crude oil. In the HDS process, the fossil fuel and hydrogen gas are fed together to a fixed-bed reactor containing an HDS catalyst such as Nickel–Molybdenum (NiMo) or Cobalt–Molybdenum (CoMo) packed on an Al_2O_3 matrix, which under high temperature (200 - 455 °C) and pressure (150 - 3000 psi), remove sulphur from organosulphur compounds as H_2S by hydrogenation or hydrogenolysis, respectively (Boniek *et al.*, 2015). The hydrogenation and hydrogenolysis reactions depend on the size and structure of the individual S compound. In the case of DBT, HDS has been reported to follow two pathways. One is a direct hydrogenolysis pathway (DDS) where sulphur is removed without interfering with the aromatic ring, whereas in the hydrogenation desulphurisation (HYD) pathway, aromatic rings of DBT compounds are preferentially hydrogenated to 4H-or 6H-DBT intermediates that are subsequently desulphurised as shown in Figure 1.2. The DDS proceeds faster than the HYD pathway (Srivastava, 2012).

As the number of rings and methyl substituents is increased, the HDS reactivity of the sulphur compounds is notably reduced as the substituents make it difficult for hydrogen to reach the sulphur atom. This is termed steric hindrance, which slows down the reaction rate. HDS of inorganic and aliphatic sulphur compounds such as thiols and sulphides, aliphatic acyclic sulphides (thioethers) and cyclic sulphides (thiolanes) is consequently more straightforward than desulphurisation of compounds that contain aromatic sulphur, such as thiophenics and its benzologs (e.g. benzothiophene (BT), dibenzothiophene (DBT), benzonaphthothiophene) (Gray *et al.*, 1995). HDS becomes increasingly expensive, and the reaction becomes less efficient in handling sulphur removal as lower sulphur levels are reached in the fuel feed (Linguist and Pacheco, 1999), as illustrated in Figure 1.3

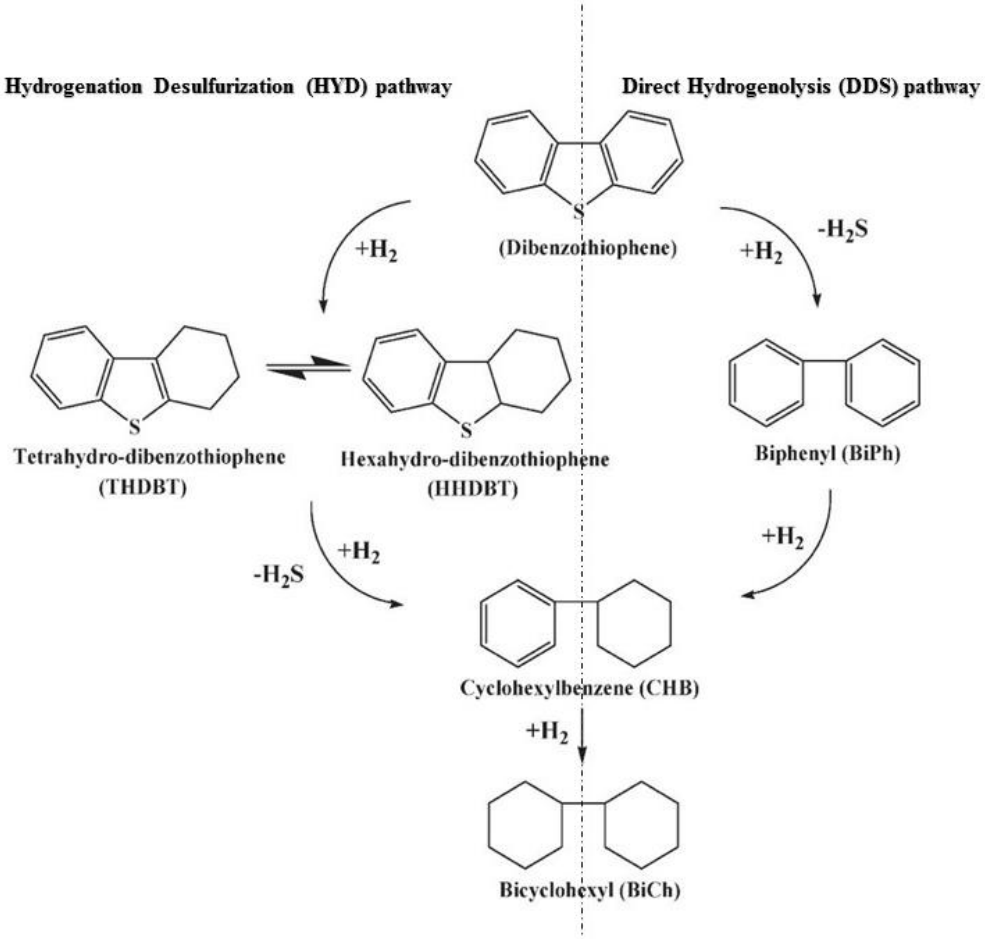


Figure 1.2: Pathways for the HDS of DBT at 300°C and 1500 psi catalysed by CoMo/Al_2O_3.
Adapted from Houalla *et al.* (1980)

The bulk of inorganic sulphurs and simple organic sulphurs can be removed by existing HDS technology, however up to 70% of the sulphur in crude oil is found as dibenzothiophene (DBT) and substituted DBTs which are particularly recalcitrant to removal by HDS (Borgne and Quintero, 2003). Under deep desulphurisation conditions in HDS, the reactivity of the sterically hindered alkyl DBTs is inhibited by different contaminants such as H_2S, nitrogen-containing compounds and aromatic molecules. Organic nitrogen compounds have stronger adsorptive strength for the catalyst's active sites than that of sulphur-containing compounds. The resulting degree of inhibition depends on the type and concentration of the organic nitrogen compounds (Stanislaus *et al.*, 2010). The average nitrogen content in straight-run light gas oil feeds is in the range of 100–300 ppm, whereas cracked distillate feeds (e.g. light cycle oil, coker gas oil), usually contain higher nitrogen levels (>500ppm). Reducing

the nitrogen content of diesel feeds increased HDS reactivity significantly (Yang *et al.*, 2004).

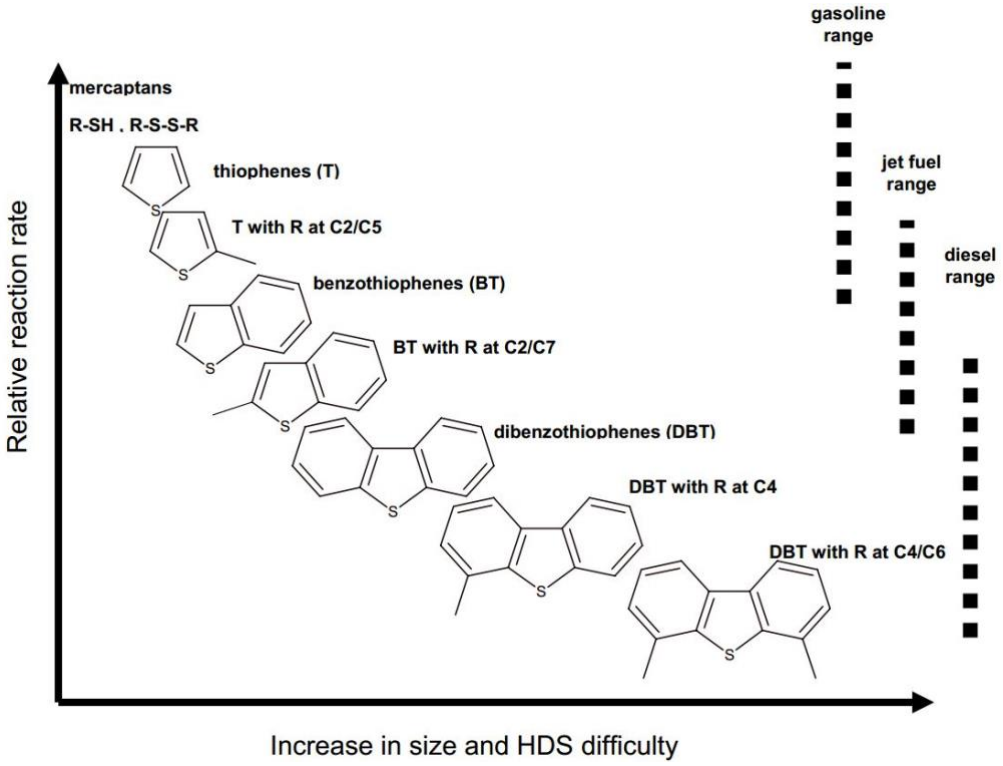


Figure 1.3 Relationship between HDS reactivity and the size of the model organosulphur compounds.
Adapted from Song (2003)

Thiophenic sulphur is resonance stabilised and resistant to removal by cracking. Hence extreme hydrogenation (deep desulphurisation) using a $NiMo/Al_2O_3$ catalyst under high temperature and pressure conditions is required for the HDS of refractory compounds such as DBT and 4,6-dimethyl dibenzothiophene (4,6-DMDBT) in order to achieve legally required levels of desulphurisation (Bataille *et al.*, 2000). This increases the cost of HDS further, and it is not particularly effective against DMDBTs (Ma *et al.*, 1994; Shafi and Hutchings, 2000). During deep desulphurisation, the HDS process releases metals (Ni, V) found in the fuels and large amounts of organic nitrogen compounds present in the diesel. These are deposited over time on the catalyst surface and lead to catalyst deactivation. Aromatic compounds in diesel burn out at a high temperature and are reduced to elemental carbon (coking) - this also leads to catalyst deactivation (Corma *et al.*, 2001).

1.4.2 Other desulphurisation techniques

HDS is an established method of industrial-scale desulphurisation, hence any feasible alternative desulphurisation method must prove more efficient and/or cost-effective than HDS. Oxidative desulphurisation (ODS), oxidation extraction desulphurisation (OEDS), and adsorptive desulphurisation (AD) are alternative desulphurisation technologies being explored to produce ultra-clean fuels. However, these technologies also have potential drawbacks (Srivastava, 2012).

The phenomenon that organosulphur compounds are more soluble than hydrocarbons in certain solvents has been exploited to develop solvent-extraction based desulphurisation technology. This method does not require hydrogen gas and operates at ambient temperature (Babich & Moulijn 2003). Many research studies employing conventional solvents (e.g. ethanol, acetone, polyethylene glycol, acetonitrile, γ-butyrolactone, dimethylformamide), nitrogen-containing solvents (e.g. amines, pyrrolidones), or sulphur containing solvents (e.g. DMSO and sulpholane) for extracting the sulphur from fuels have been reported (Stanislaus *et al.*, 2010). However, extraction using these solvents does not reduce the sulphur content to ultra-low levels (<10 ppm).

Extractive desulphurisation becomes less selective as the heaviness of the oil increases. Also, aromatic hydrocarbons may be co-extracted with the sulphur, which affects the calorific value of the fuel. Extractive desulphurisation methodologies are currently being developed using ionic liquids, salt-like materials that exist in a liquid state at room temperature up to 100°C. These are non-volatile, and their hydrophobic or hydrophilic nature can be modulated by modifying both cation and anion. Ionic liquids such as chloroaluminate, hexafluorophosphate and tetrafluoroborate have been shown to possess good selectivity for organosulphur compounds in crude oil (Bösmann *et al.*, 2001). This technology is still in the development stage as there are several issues to be overcome, such as regeneration of ionic liquids for repeated cycles of extraction, reducing the number of cycles for efficient separation and the cost-effectiveness of the process.

Oxidation based desulphurisation methods use oxidants (usually H_2O_2 or H_2SO_4) to convert sulphur compounds to sulphones with a high polarity that can subsequently be readily extracted using polar solvents. However, the oxidant is stoichiometrically consumed during the process, and it would require large quantities of the oxidant to treat heavy oils on an industrial scale, which would make the process very expensive (Javadli and Klerk, 2012). Desulphurisation of gas oil using an ultrasound-assisted catalytic oxidative process is reported to benefit from fast and mild operating conditions with high efficiency (Ja'fari *et al.*, 2018). However, the operational factors such as time, temperature, the oxidizing agent, frequency, power, and catalyst must be studied in more detail before commercial application.

Adsorption desulphurisation using biologically derived sorbent materials (e.g. carbonised material from date palm kernel) has been proposed as an economically attractive alternative technology (Zubaidy *et al.*, 2013). The process is operated at ambient temperature and pressure, without consumption of hydrogen or oxygen. Building on this, several different kinds of adsorbents based on activated carbon, and agro-waste such as pomegranate leaf powder (Sadare and Daramola, 2018) and bamboo-derived porous biochar (Yang *et al.*, 2018) based adsorbents have been explored by researchers for the desulphurisation of crude oil distillates, especially diesel. The operating cost of the process differs with the choice of the adsorbent used. It has been proposed that the activity of the adsorbents could be enhanced by molecular imprinting technology to create specific molecular recognition sites in polymers to identify sulphur-bearing template molecules (Yang *et al.*, 2014). This approach is only in the early stages of the investigation.

In addition to achieving the necessary ultra-low sulphur level fuels on a commercial scale, an ideal desulphurisation process should require minimal energy (operational costs) and involve straightforward means for the removal/extraction of the final oxidized sulphur product. Achieving efficient deep desulphurisation of fuels is the major challenge with current technologies. Hence, alternative methods that can solve this problem would be of enormous value and commercial importance. Biodesulphurisation, using microorganisms (bacteria or fungi), has been shown to solve the problem of

removing sulphur from the refractory organic compounds and therefore is a highly active area of research. The challenges outlined above for HDS and alternative chemical-based technologies mean that alternative, environment-friendly and cheaper technologies for desulphurisation of refractory organic compounds found in fuels have been explored.

1.5 Biodesulphurisation (BDS)

Sulphur is an essential nutrient for growth and metabolic activity in all living organisms. It is a vital component of proteins through the amino acids cysteine and methionine, and the disulphide bonds that render the proteins their 3D structure. It is an active constituent of numerous coenzymes and prosthetic groups (e.g. iron sulphur centres, coenzyme-A, thiamine, lipoic acid, S-adenosylmethionine, glutathione). In the case of bacteria, sulphur represents about 0.5–1% of their dry weight (Guobin *et al.*, 2006). Inorganic sulphate is the preferred sulphur source for the growth of most microorganisms, including bacteria. However, inorganic sulphates are not prevalent in nature, whereas alternative sulphur sources such as sulphonates (R-C-SO_3) and sulphonate esters are found as naturally occurring products or as xenobiotic compounds, making up over 95% of the sulphur content of most aerobic soils (Autry and Fitzgerald, 1990; Kertesz, 2000; Scherer, 2009). Bacteria have evolved ways to acquire sulphur from alternative sources available in the environment under sulphur-limiting conditions, by producing additional proteins that enable metabolism of sulphur sources available in the environment (Kertesz and Wietek, 2001; Kirkwood *et al.*, 2005). These sulphate starvation-induced proteins could either be enzymes with catalytic functions or transport systems involved in scavenging and metabolising the alternative sulphur sources (Kertesz and Wietek, 2001; Kertesz, 2000). The ability of certain microorganisms to utilise the refractory organic compounds as their sulphur source could potentially be exploited for desulphurisation of sulphur-containing fuels (Kilbane II, 2006; Kilbane, 1989).

In the higher boiling fractions of crude oil, more than 60% of the sulphur occurs as DBT and substituted DBT derivatives (Kropp and Fedorak, 1998). BT predominates in gasoline, and DBT and its alkylated derivatives predominate in the diesel fraction (Gupta *et al.*, 2005; Ma, 2010; McFarland *et al.*, 1998).

BT and DBT have the fundamental ring structure typical for all polycyclic aromatic compounds (PACs) found in fuel, and therefore, they have been historically used as ideal model compounds to test for BDS capability in the development stages of BDS research.

Several bacteria, fungi and yeast have been shown to metabolise DBT. However, only the naturally existing microorganisms capable of extracting sulphur by non-destructive degradation of polycyclic aromatic organosulphur compounds are of value for BDS. Fungi such as *Paecilomyces sp*. TLi (Faison *et al.*, 1991), *Cunninghamella elegans* (Crawford and Gupta, 1990), *Aspergillus*-like fungus (Acharya *et al.*, 2005), *Stachybotrys* sp. WS4 (Torkamani *et al.*, 2008), white-rot fungus *Trametes versicolor* ATCC 200801 and *Phanerochaete chrysosporium* ME 446 (Aytar *et al.*, 2008), were previously reported to be capable of DBT oxidation without the formation of biphenyl. Yeast such as *Rhodosporidium toruloides*, *Candida parapsilosis* S1-Y1 and *Cryptococcus humicolus* S1-Y2 (Bayoumi *et al.*, 2009) and filamentous ascomycetes *Exophiala spinifera* (Elmi *et al.*, 2015) and *Stachybotrys bisbyi* (Gherbawy *et al.*, 2016) have been proposed for BDS recently, where the metabolism of DBT proceeds further beyond 2HBP to less toxic metabolites. In general, fungi only e DBT effectively into DBT-sulphones. The metabolic pathway further downstream varies across fungal species and is not yet fully understood (Linder, 2018). On the other hand, metabolic pathways of DBT in bacterial systems are well established, and are more efficient than fungi, making bacteria the preferred microorganisms for the development of BDS technology. Collated and curated information about bacterial BDS pathways, enzymes and genes are available in specialised knowledge bases for microbial biocatalytic reactions and biodegradation pathways. The University of Minnesota Biocatalysis/Biodegradation Database (Eawag-BBD; Gao *et al.*, 2010) includes the DBT degradation (Kodama) pathway (http://eawag-bbd.ethz.ch/dbt2/dbt2_map.html); DBT desulphurisation (4S) pathway (http://eawag-bbd.ethz.ch/dbt/dbt_map.html); and the BT desulphurisation pathway (http://eawag-bbd.ethz.ch/btp/btp_map.html). MetaCyc (SRI International, Canada) is another comprehensive database that provides a curated collection of metabolic pathways and enzymes from all domains of life; the MetaCyc Pathway code for DBT desulphurisation is PWY-681. The 4S

pathway is described up to the formation of 2-hydroxybiphenyl (2HBP). Interestingly, MetaCyc provides links to another metabolic pathway (PWY-7008), by which 2HBP is metabolised into benzoate by *Pseudomonas nitroreducens*.

1.6 Bacterial BDS Pathways

Bacteria can metabolise organosulphur compounds for different nutritional requirements, including both carbon and sulphur source. When used as a carbon source, the hydrocarbon backbone of the organic compound is broken or destroyed. While early research on BDS technology (pre-1990) focussed on bacterial strains that were only capable of complete metabolism (degradation) of the organosulphur compounds in fuels, later investigations focussed on strains (both anaerobic and aerobic bacteria) with a non-destructive sulphur-specific metabolic pathway. The efficiency of anaerobic BDS process was less than that of aerobic BDS process, and there is little evidence for the commercial potential of anaerobic desulphurisation owing to the high costs associated with maintaining culture conditions and supplying the hydrogen required for the anaerobic BDS process. Therefore, aerobic desulphurisation has been more widely considered (Gupta *et al.*, 2005; Nazari *et al.*, 2017).

1.6.1 The Kodama pathway

In the early stages of BDS research, bacteria that could metabolise DBT as their energy source were explored. A Japanese research team showed that microorganisms could attack thiophenic compounds in a series of oxidation steps, converting them into water-soluble compounds and termed this the ‘Kodama pathway’ (Kodama *et al.*, 1973). In this pathway, the peripheral aromatic ring of DBT is cleaved following 3 main steps - hydroxylation, ring cleavage and hydrolysis (Gupta *et al.*, 2005; Mohebali and Ball, 2008; Soleimani *et al.*, 2007) as shown in Figure 1.4. Members of several bacterial genera have been reported to follow this pathway for DBT utilisation, the predominant one being *Pseudomonas* (for example, *P. alcaligenes* and *P. putida*) (Hartdegen *et al.*, 1984; Monticello and Finnerty, 1985).

The Kodama pathway is a destructive BDS pathway in which C-C bonds in the DBT molecule are broken, and sulphur is not selectively removed from the organic molecule, as shown in Figure 1.4. The end product, water-soluble 3-hydroxy-2-formylbenzothiophene, is still a complex organic sulphur-containing compound but has a lower carbon content than DBT, and removal of this compound from the fuel undesirably reduces its overall hydrocarbon carbon content. Hence, due to its destructive nature, the Kodama pathway reduces the calorific value of the fuel that is being desulphurised.

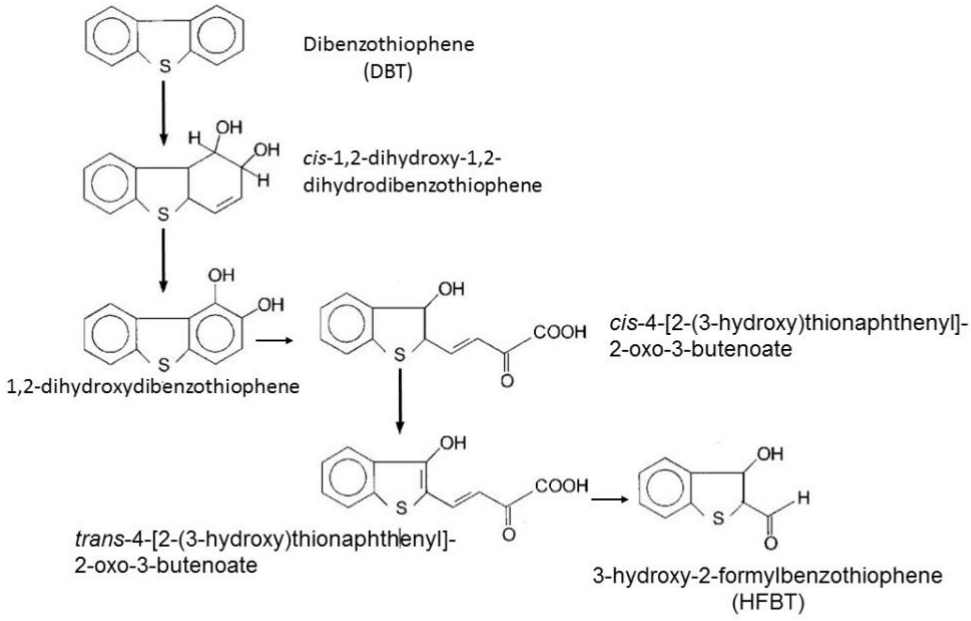


Figure 1.4: Kodama pathway of DBT metabolism
Adapted from (Gupta *et al.*, 2005; Soleimani *et al.*, 2007) The pathway beyond HFBT are not precisely known.

Ideally, bacterial strain considered suitable for application in an industrial desulphurisation process, should cleave only the C-S bonds, and leave the C-C bonds to remain intact. This selectivity is necessary because it means that fuel can be desulphurised without affecting the calorific value (Mohebali and Ball, 2008). The majority of research in the past two decades has therefore focussed on strains that follow the '4S pathway', an oxidative desulphurisation pathway that cleaves only the carbon-sulphur bond in DBT and leaves the hydrocarbon skeleton intact (Abin-Fuentes *et al.*, 2013).

1.6.2 The 4S pathway

The 4-step pathway of biodesulphurisation or the '4S pathway' involves an oxidative cleavage of the C-S bond in DBT. This pathway was first described in 1988 by pioneers at Gas Technology Institute, USA (formerly, Institute of Gas Technology (IGT)) for the DBT desulphurisation mechanism of *Rhodococcus erythropolis* strain IGTS8 (Gallagher *et al.*, 1993; Kilbane, 1990; Mohebali and Ball, 2008). The '4S pathway' is a targeted desulphurisation pathway in which DBT is desulphurised and converted to 2-hydroxybiphenyl (2HBP). Through this pathway, shown in Figure 1.5, only the C-S bond in the DBT is specifically cleaved, and the hydrocarbon skeleton of DBT is released intact with no loss in the calorific value of the fuel. Hence the use of this pathway is proposed for the commercial desulphurisation of petroleum products in production fields and also refineries (McFarland, 1999; Monticello, 2000).

The 4S pathway is a multi-enzymic process involving a sequential activity of two cytoplasmic flavin-dependent monooxygenases (DszC and DszA) and a desulfinase (DszB) that catalyse 4-step sequential reactions as illustrated in Figure 1.5. The reduced flavin ($FMNH_2$) required for the activity of DszC and DszA is supplied by the concurrent activity of an NADH-dependent FMN oxidoreductase (DszD) (Gallagher *et al.*, 1993; Gray *et al.*, 2003, 1996). The enzyme DszC converts DBT into DBT-sulphoxide (step 1) and DBT-sulphone (step 2). The enzyme DszA converts DBT into 2′-hydroxybiphenyl-2-sulfinic acid (step 3), and as the final step 4, the enzyme DszB hydrolyzes 2′-hydroxybiphenyl-2-sulfinic acid into 2-hydroxybiphenyl (2HBP) and sulphite (SO_3^{2-}). These enzymes have been associated with the BDS activity of *R. rhodochrous* IGTS8 (Gray *et al.*, 1996), *R. erythropolis* D-1 (Ohshiro *et al.*, 1997), and *Paenibacillus* sp. strain A11-2 (Konishi *et al.*, 2000a). The structural and functional aspects of purified forms of these enzymes have been studied in detail (Duan *et al.*, 2013; Hino *et al.*, 2017; Lee *et al.*, 2006a; Liu *et al.*, 2014b; Okai *et al.*, 2017).

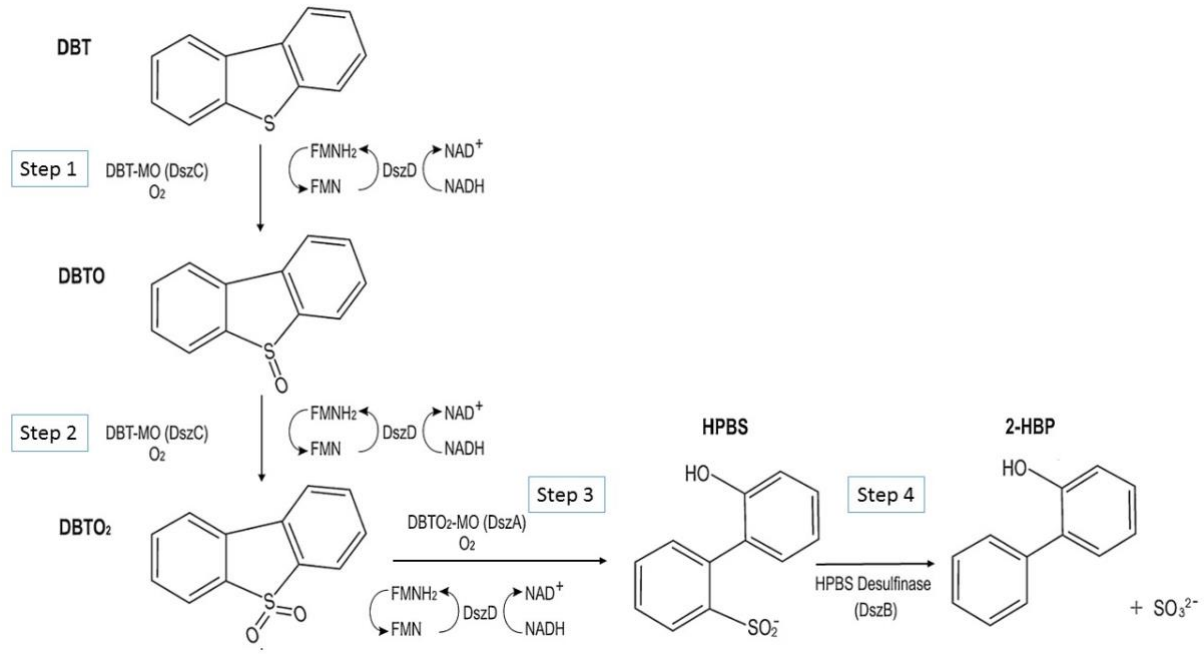


Figure 1.5: The 4S pathway of DBT desulphurisation
Adapted from (Boniek *et al.*, 2015; Soleimani *et al.*, 2007; Gallagher *et al.*, 1993; Javadli and Klerk, 2012)

DszC or DBT monooxygenase - catalyses the conversion of DBT to DBT-sulphone ($DBTO_2$) in a two-step process with DBT sulfoxide (DBTO) being the intermediate compound. Isotopic labelling studies indicate that the two oxygen atoms needed for the reactions are molecular oxygen obtained from the air during the aerobic growth of the microorganism. The first step (DBT to DBTO) with a rate constant of 0.06 min^{-1} is 10 times slower than the second step (DBTO to $DBTO_2$) with a rate constant of 0.5 min^{-1}. DszC enzyme of *R. rhodochrous* IGTS8 shows homology to acyl coenzyme-A and is a homotetramer with a subunit molecular weight of 50 kDa. In the case of *R. erythropolis* D-1, this enzyme is a homohexamer with a subunit molecular weight of 45 kDa. This enzyme can act on the DBT and its derivatives such as 4,6-dimethyl DBT, 2,8-dimethyl DBT, and 3,4-benzo-DBT but it does not show any activity on other polycyclic aromatic compounds with a structure similar to that of DBT, such as carbazole, dibenzofuran or fluorene. In these latter compounds, the sulphur atom is replaced by nitrogen, oxygen and hydrogen, respectively.

DszA or DBT-sulphone monooxygenase - oxidizes $DBTO_2$ to 2-(2′-hydroxyphenyl) benzene sulfinate (HPBS). It is a homodimer with a subunit molecular weight of 50 kDa, and the optimum pH and temperature for its function are pH 7.5 and 35°C, respectively. However, the DszA from thermophilic *Paenibacillus* sp. strain A11-2 is optimal at 45°C and stable up to 60°C. Monooxygenase enzymes, in general, show a 1.7-fold higher activity towards 4,6-dimethyl-DBT-sulphone as compared to DBT-sulphone. The activity of the enzyme from *R. erythropolis* D-1 is inhibited by 50% in the presence of 1 mM EDTA or any other chelating agents but no inhibition is observed in the case of *R. rhodochrous* IGTS8 even in the presence of 10 mM EDTA (Gray *et al.*, 1996; Ohshiro *et al.*, 1997).

Flavin dependant monooxygenases contain flavin as a prosthetic group and require NADP or NADPH as a coenzyme. Based on the chemical reactions catalysed by them, DszC and DszA are classified as

- Oxidoreductases (EC 1)
- Acting on paired donors, with incorporation or reduction of molecular oxygen (EC 1.14)
- With reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen into the other donor given enzyme commission numbers (EC 1.14.14)

Unlike other monooxygenases which transfer a single oxygen atom to a carbon, DszC supplies two oxygen atoms sequentially to the sulphur atom of DBT. This makes DszC a unique member of the flavin monooxygenase family. The DszC and DszA have been assigned Enzyme Commission numbers EC 1.14.14.21 and EC 1.14.14.22, respectively. These enzymes have been recorded to form a two-component system with NADH-dependent FMN reductase DszD with EC number 1.5.1.42, based on its reaction as an oxidoreductase (EC 1); acting on the CH-NH group of donors (EC 1.5), and with NAD+ or NADP+ as acceptor (EC 1.5.1).

DszB or HPBS desulfinase - catalyses the conversion of 2’-hydroxybiphenyl-2-sulfinate (HBPS) to 2-hydroxybiphenyl (2HBP) and sulphite. It is produced in relatively lower quantities than DszA and DszC, and therefore it is the rate-

limiting enzyme of the 4S pathway. It is a monomer with a subunit molecular weight of 39 kDa (Lee *et al.*, 2006b), and it is functional over a broad temperature range (25–50°C), the optimum being 35°C between a pH range of 6.0 – 7.5 in the case of *Rhodococcus* IGTS8 (Watkins *et al.*, 2003). Interestingly, the enzyme has been shown to lose activity irreversibly when the only cysteine residue in its amino acid sequence was modified. The catalytic activity of DszB is inhibited by one of the reaction products 2HBP, but not by sulphite. Moreover, the DszB enzyme has a narrow substrate specificity: only HBPS and its derivates (Nakayama *et al.*, 2002), and it was proposed to belong to a unique class of desulfinase enzymes based on its 3D structure (Ohshiro *et al.*, 2007). DszB is not assisted by pyridoxal 5'-phosphate or any other cofactor, which further distinguishes it from the other enzymes that catalyse desulfination, (e.g. cysteine sulfinate desulfinase, L-aspartate β-decarboxylase) (Geronimo *et al.*, 2017). Based on the chemical reaction catalysed by DszB, it was classified as a

- Hydrolase (EC 3)
- Acting on carbon-sulphur bonds (EC 3.13)
- Acting on carbon-sulphur bonds (only sub-subclass identified to date) (EC 3.13.1),

and it remains as the only member of its class EC 3.13.1.

The end product of the 4S pathway, 2-hydroxybiphenyl (2HBP) and its derivatives are often used as the indicator for a strong BDS reaction. It can be detected qualitatively by Gibbs assay (Kayser *et al.*, 1993; Mohamed *et al.*, 2015; Mohebali *et al.*, 2008; Wang *et al.*, 2013a) or measured quantitatively using GC-MS (Khedkar and Shanker, 2015; Labana *et al.*, 2005; Li *et al.*, 2005b) or HPLC (Calzada *et al.*, 2012; Caro *et al.*, 2007; Piddington *et al.*, 1995). Gibbs assay is a colourimetric method that is widely used to confirm the BDS of DBT. The test relies on the reaction between phenol formed as a result of DBT desulphurisation (2HBP) and the Gibbs reagent (2,6 dichloroquinone-4-chloroimide) to form indophenol that is visualised as a deep blue colour (Gibbs, 1927; Pallagi *et al.*, 1994; Svobodová *et al.*, 1978). Gibbs assay is described in detail in section 3.1.2. The inhibitory effects of 2HBP on the growth and BDS activity of bacteria have been recognised since the early

days of BDS research (Akhtar *et al.*, 2009) and it remains a significant problem (Kilbane II, 2017). The minimum amount of cytoplasmic 2HBP concentration to inhibit the activity of the enzymes DszA, DszB, and DszC by 50%, is 60, 110 and 50 µM, respectively (Abin-Fuentes *et al.*, 2013).

1.6.3 Pathways of BT desulphurisation

In contrast to DBT-desulphurising bacteria, little is known about the organisms that can desulphurise BT exclusively, both in terms of the enzymes and genes associated with the BT-desulphurisation pathway (Ma, 2010). Owing to the mutagenic and carcinogenic nature of benzothiophene derivatives, only those bacterial species that have developed tolerance to these compounds survive in oil-contaminated environments (Kropp and Fedorak, 1998). It has been proposed and widely understood that sulphur-specific degradation (the cleavage of carbon-sulphur bonds) in BT occurs by a mechanism similar to the 4S pathway of dibenzothiophene desulphurisation (Gilbert *et al.*, 1998; Kirimura *et al.*, 2002).

Figure 1.6 shows two divergent pathways (a & b) for BT desulphurisation, with the two different end products proposed to date (Kirkwood *et al.*, 2007a). In both pathways, benzothiophene (BT) is first oxidized to benzothiophene-S,S-dioxide in a two-step process with benzothiophene-S-oxide as intermediate. In the case of *Gordonia desulfuricans* 213E which exhibits pathway b, the sulfinate group is removed with oxygenation of the molecule, giving 2-(2'-hydroxyphenyl)ethan-1-al (HPEal; Gilbert *et al.*, 1998). In the case of *Sinorhizobium* sp. KT55 (pathway a), the final product is o-hydroxystyrene, produced through desulfination of the molecule, without the oxygenation of the carbon atom (Konishi *et al.*, 2000b). Both end products were reported for BT desulphurisation by *Rhodococcus* sp. strain WU-K2R, where the HPEal was further converted to Benzofuran (Kirimura *et al.*, 2002). Benzofuran was also reported as the end product in the case of *Gordonia rubropertinctus* T08 (Matsui *et al.*, 2001b) and *Mycobacterium phlei* WU-0103 (Ishii *et al.*, 2005).

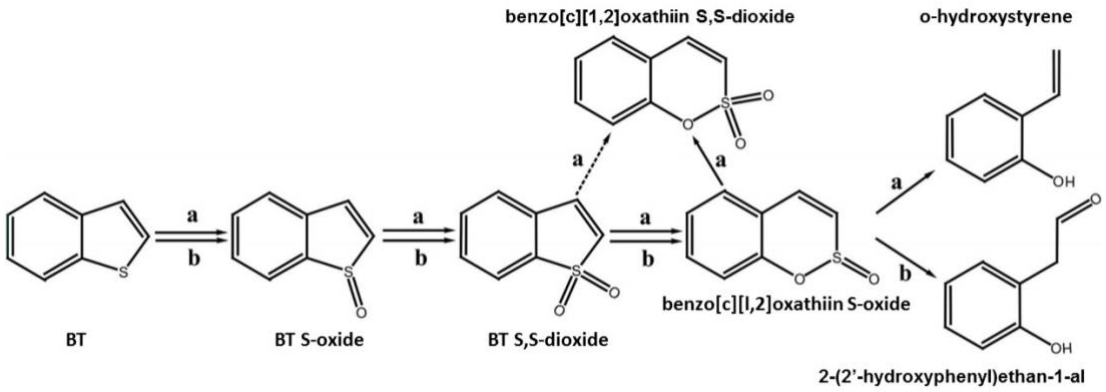


Figure 1.6: Proposed BT desulphurisation pathways
a - BT desulphurisation pathway reported for *Sinorhizobium* sp. KT55
b - BT desulphurisation pathway proposed for *Gordonia* sp. 213E

1.7 Avenues for the improvement of BDS technology

The '4S Pathway' is a multi-enzyme process, involving the enzymes DszC and DszB that require a constant supply of $FMNH_2$ as a co-substrate, and therefore it is desirable to have growing cells to sustain the BDS functionality and to develop continuous bioprocess (Lin and Tao, 2017; Setti *et al.*, 1997). BDS research ultimately aims to develop a commercial process for petroleum desulphurisation. Mass production of biocatalysts with a broad substrate specificity and high BDS rates has been suggested as an effective way to decrease the bioprocessing costs. However, there are other inherent critical issues in terms of the bacterial biocatalysts used for the process. Despite being a very critical element for growth, sulphur is not required in high quantities by bacteria naturally (Kilbane, 1990). Therefore, natural *dsz* gene regulation in BDS strains, which is often based on the sulphur needs of the bacteria, results in lower levels of *dsz* gene expression than would be required for industrial application. Monticello (2000) estimated that consistent desulphurisation rate of over 1.2 mmol of DBT/g (dry cell weight)/hour of the catalyst was needed to achieve <15 ppm in fuels. To achieve current sulphur specification (<10 ppm), biocatalysts with a desulphurisation activity of 3 mmol of DBT/g (dry cell weight)/hour are required (Kilbane II, 2006; Mohebali and Ball, 2008). Furthermore, the BDS activity of various naturally occurring bacteria can be completely repressed by sulphate or other readily bioavailable sulphur sources, including methionine, cysteine, taurine, methanesulfonic acid

and Casamino acids. The sulphate induced repression of BDS activity is an ongoing challenge (El-Gendy and Nassar, 2018; Mohebali and Ball, 2008).

The main strategies to date for BDS technology improvement include process engineering optimisation and development of recombinant BDS strains with enhanced BDS activity, notwithstanding the search for new naturally occurring bacterial isolates with desired BDS characteristics naturally (Martínez *et al.*, 2016). In terms of strain enhancement, research has focussed on increasing *dsz* gene expression and elimination of natural gene repression. The *dszABC* operon of *R. erythropolis* IGTS8 (NCBI accession: L37363) described by Piddington and co-workers (1995) serves as the reference gene sequence for many genetic engineering and mutation studies in this field. The *dsz* operon of IGTS8 is a single plasmid-borne operon containing three open reading frames *dsz*A, *dsz*B and *dsz*C, that encode three proteins, DszA, DszB, and DszC, respectively that are involved in the conversion of DBT to 2HBP. The *dszD* gene for DszD is not located near this cluster but occurs in the chromosomal DNA (Gray *et al.*, 1996) (Oldfield *et al.*, 1997). These *dsz* genes have been isolated, cloned, mutated and overexpressed by various genetic engineering strategies with the aim of generating recombinant BDS strains or enhancing the BDS capability of the native strains (Aliebrahimi *et al.*, 2015; Gallardo *et al.*, 1997; Li *et al.*, 2008a; Li *et al.*, 1996; Martínez *et al.*, 2016; Meesala *et al.*, 2008; Shavandi *et al.*, 2009). The genetic engineering strategies (summarised in Table 1.2) included the rearrangement of the *dsz* gene cluster, increasing the copy number of the *dsz* genes, and the expression of heterogeneous equivalents of the FMN reductase (DszD). Cloning and expression of genes such as *hpaC* (encodes 4-Hydroxyphenylacetate 3-monooxygenase for the oxidation of 4-hydroxyphenylacetate in *E. coli*; Galán *et al.*, 2000), *vbg* (encodes *Vitreoscella* haemoglobin VHb; Mu *et al.*, 2017; Xiong *et al.*, 2007), FMN:NADPH oxidoreductase genes from *Vibrio harveyi* (Xi *et al.*, 1997) and *Photobacterium fischeri* (Lei and Tu, 1996) have been shown to improve the activity of DszC and DszA. As an alternative approach, the *dszABC* genes cloned downstream of different promoters, e.g. 16S rRNA rrn promoter from *Rhodococcus sp.* strain T09 (Matsui *et al.*, 2002); heat shock protein Hsp60 promoter from *Mycobacterium sp.* strain SP3 (Takada *et al.*, 2005); *lac* promoter from *E. coli* (Alves *et al.*, 2007; Shavandi *et al.*, 2009) have

been used to achieve constitutive expression of *dsz* genes even in the presence of sulphates.

Table 1.2 Genetic engineering strategies for BDS strain enhancement

Genetic Engineering Strategy	Expression Host Organism	Reference
Rearranging the *dsz* gene cluster	*Rhodococcus* sp. T09	Matsui *et al.*, 2001a
	R. erythropolis KA2-5-1	Hirasawa *et al.*, 2001
	R. erythropolis KA2-5-1	Kilbane II, 2006
	R. erythropolis DS-3	Li *et al.*, 2008a
	E. coli DH10B	Reichmuth *et al.*, 2004
Expression of heterogeneous or modified FMN reductase *dszD* gene	*E. coli* and *P. putida*	Galán *et al.*, 2000; Reichmuth *et al.*, 2000
	Pseudomonas sp.	Galán *et al.*, 2000
	E. coli BL21	Lei and Tu, 1996
	E. coli DH10B	Reichmuth *et al.*, 2004
	R. erythropolis LSSE8-1	Xiong *et al.*, 2007
	R. erythropolis IGTS8	Kamali *et al.*, 2010
Altering the promoter sequence of *dszABC* operon	*R. erythropolis IGTS8*	Rambosek *et al.*, 1999
	Rhodococcus sp. T09	Matsui *et al.*, 2002
	R. erythropolis KA2-5-1	Noda *et al.*, 2003
	Mycobacterium sp. G3	Takada *et al.*, 2005
	E. coli DH5α	Alves *et al.*, 2007
	Gordonia alkanivorans RIPI90A	Shavandi *et al.*, 2009
	E. coli BL21	Khosravinia *et al.*, 2018

The table shows the popular gene engineering strategies used to improve BDS activity. Some of the recombinant plasmids were expressed in *E. coli* which is not a suitable organism for BDS activity. This reflects the difficulty in genetic manipulation of rhodococci. Also, the success of the genetic manipulation does not necessarily translate into beneficial effects in real BDS application.

In the case of heterologous hosts, the level of desulphurisation activity achieved by genetic manipulation is often limited by factors that are not clearly understood. A better understanding of the host factors that contribute to the functioning of the pathway is necessary in order to successfully develop genetically engineered strains with very high levels of expression of the *dsz* genes (Kilbane II, 2006). For technical (e.g. gene expression and codon usage preferences) and regulatory reasons, the use of original strains has been recommended (Monticello, 1998), and self-cloning is generally regarded as

more effective than heterologous recombination, as gene expression and DBT permeation are readily achieved (Hirasawa *et al.*, 2001; Li *et al.*, 1996; Matsui *et al.*, 2001a). Directed evolution methods, e.g. DNA shuffling, chemostat enrichment have been applied on the wild-type BDS *Rhodococcus* strains to develop highly recombined genes and evolved enzymes with extended substrate range (gain of function) and improved (20 fold increased) activity of DszC (Arensdorf *et al.*, 2002; Coco *et al.*, 2001).

The complexity of the 4S pathway enzyme system and its cofactors requirement prohibit the use of purified enzyme systems, and whole-cell biocatalysis is recommended for an efficient BDS process (Mohebali and Ball, 2016; Setti *et al.*, 1997). Fuel oils are complex mixtures of organic solvents which could be very toxic to microorganisms even at low concentrations. Therefore, solvent tolerance of the bacteria is the main factor influencing the suitability of the bacteria for the purpose and other process engineering parameters, e.g. fuel - culture medium ratio in the biphasic BDS process. The oil/water ratio is a critical parameter that influences operational costs associated with water handling, separation and disposal (Foght, 2004). Culturing oil-tolerant bacteria in biphasic systems (oil-aqueous), where the growing cells are presented along with the fuel to be desulphurised, has been shown as a more efficient approach than using resting cells, where non-growing mature desulphurising cells are suspended in oil-phosphate buffer system to carry out the BDS reaction in their resting state (Adlakha *et al.*, 2016; Tao *et al.*, 2006).

Many naturally occurring BDS strains do not have high solvent tolerance, the latter being a highly desirable feature for BDS on a commercial scale. The solvent tolerance of Gram-positive bacterial strains of the genera *Bacillus*, *Rhodococcus* and *Arthrobacter* is lower than that of the Gram-negative *Pseudomonas* strains (Sardessai and Bhosle, 2002). Therefore, recombinant *Pseudomonas* strains have been developed to exploit their high solvent tolerance and relatively higher growth rates.For example, by introducing the BDS gene cluster *dsz*ABCD from *R. erythropolis* XP into the solvent-tolerant strain *P. putida* Idaho, Tao and co-workers (2006) constructed *Pseudomonas putida* A4 strain which retained the BDS capability and a high solvent tolerance

(10% (v/v) *p*-xylene) for a biphasic reaction mixture. However, Caro *et al*. (2007) showed that a genetically modified strain *P. putida* CECT 5279 was more sensitive to DBT mass transfer limitation than *R. erythropolis* IGTS8 in biphasic systems with DBT and n-hexadecane, and therefore recommended the latter strain for application in commercial BDS process. Indeed, rhodococci have a high capability to uptake organic compounds from the oil interface because of their naturally highly hydrophobic cell membranes and the production of surface-active modules (Monticello, 2000). Kawaguchi *et al.* (2012) constructed a recombinant strain by introducing desulphurising genes from *Rhodococcus erythropolis* IGTS8 into an organic solvent-tolerant *Rhodococcus opacus* strain B-4, to avoid transcriptional inhibition by the sulphate end product. DBT consumption rates increased by 80% in oil (n-hexadecane)/water biphasic reaction mixtures and resting cells were predominantly localized in the emulsion layer. This demonstrates that despite the solvent tolerant nature of *Pseudomonas* strains, rhodococci are the preferred type of organisms for BDS development; hence a selection of rhodococci were examined for their BDS capability in the current research.

In terms of the bioprocess, the separation of the product (desulphurised fuel) is a challenge, particularly in the case of biphasic systems. Schilling and team (2002) proposed a continuous bioreactor system equipped with a settler device for the separation of the desulphurised oil from the aqueous phase. When BDS strains with high hydrophobicity such as *Gordonia* sp. strain CYKS1 were used, the bacteria adhered to the oil phase. In such cases, phase separation in BDS of diesel oil was optimised by the addition of ethanol as a de-emulsifier to produce a three-phase system (oil-biocatalyst-aqueous phase) in order to recover the biocatalyst (Choi *et al.*, 2003). A microchannel reactor system for biodesulphurisation was developed by Noda *et al.,* (2008), who reported that the rate of reaction in the oil/water phase of the microchannel reaction was more than nine-fold that in a batch (control) reactor. Alternatively, to avoid the problems associated with biphasic systems regarding the optimisation of the volumetric ratio between the organic and aqueous phases, triphasic systems (oil-bacteria-aqueous) such as immobilisation have been explored (Amin, 2011; Derikvand and Etemadifar, 2014; Dinamarca *et al.*, 2014a, 2014b, 2010; Li *et al.*, 2005a; Naito *et al.*, 2001;

Tang *et al.*, 2012). The immobilised systems are proposed to be more straightforward downstream processing when compared to a continuously stirred tank bioreactor systems and more cost-effective than biphasic systems (Lee *et al.*, 2008; Setti *et al.*, 1997). Amin and team (2013) applied their expertise to a vertical rotating immobilised cell reactor design for BDS by developing a two-stage system that incorporated a surfactant producing *Bacillus subtilis* stage, followed by BDS active *R. erythropolis.* Nevertheless, strains that offered easy separation of the desulphurised oil phase from the aqueous growth phase naturally remain ideal for commercialisation, and therefore, this aspect was identified as an essential criterion for selecting candidate BDS strains in the current research study.

However, after nearly three decades of research, the technology is yet to be made applicable for commercial oil biodesulphurisation. Kilbane II (2017) recognised that failure to develop superior biocatalysts whose BDS activity remains stable and uninhibited by the end product (2HBP) or sulphates in the culture system was a primary reason for this. As described in the literature review above, a variety of microorganisms have been shown to be capable of BDS, and there have been various reports of genetically engineered strains with enhanced BDS capabilities and bioprocess strategies that aim to improve the practical applicability of the technology for fossil fuel desulphurisation. The sulphate dependent repression of *dsz* gene expression has been addressed through genetic engineering. However, there are other major obstacles facing the commercialisation of BDS technology, viz, (i) biocatalyst activities are significantly lower than that required for BDS rates to match HDS rates (Kilbane II, 2006) and (ii) biocatalysts cannot maintain activity for an extended period of time. Genetic engineering strategies hitherto have focussed on the manipulation and monitoring of previously reported *dsz* genes, and have failed to consider other potential genes indirectly influencing the growth of the strains in a fossil fuel-rich medium and sulphur assimilation in general. Maass *et al.* (2015a) observed that the main goal in most BDS related research has been to determine the DBT degradation and 2HBP formation specific rates and to propose an empiric kinetic model that does not necessarily accurately describe the BDS process under different growth and operational conditions. Consequently, scaling up the BDS process with genetically modified BDS

organisms using real-world fuels has proven difficult. Therefore, rather than making several genetic modifications in an existing BDS strain to impart desired characteristics, research seeking new bacteria that exhibit many of such characteristics inherently remains a promising way forward. Recognising the challenges with genetic engineering-based strain enhancement, naturally occurring BDS bacteria are still considered as a viable means by which a commercial scale BDS technology may be achieved. Isolates from diverse genera of bacteria such as *Achromobacter, Chelatococcus, Bacillus, Gordonia, Paenibacillus, Microbacterium, Mycobacterium* and *Rhodococcus* have been reported to be capable of BDS activity, with *Rhodococcus* being the predominant group. Members of this genus are also found naturally in various oil-contaminated environments and are capable of oil degradation. Further, they are hydrophobic in nature and capable of the production of surface-active compounds to aid the breakdown of oil.

In addition to the biocatalyst, an additional challenge to scale-up is the model growth system used for strain selection. The successful BDS outcomes achieved in laboratory-scale research have not been translated to commercial scale because most strain development-related work employs model organosulphur compounds (such as BT, DBT) and model oil (such as hexadecane), whereas fossil fuels, which require treatment in a full-scale commercial technology, are complex mixtures of several components. Hence, the findings from the resting cells systems in a laboratory setting using these model chemicals become less relevant for real-world scenarios where the feedstock composition is not strictly controlled, and the mass transfer dynamics are different.

1.8 Scope of the thesis

Biodesulphurisation technology offers several advantages over the current hydrodesulphurisation approach as a method to produce cleaner low-sulphur fuels. Indeed, it offers the possibility of a more environmentally sustainable process, with no reduction in the calorific value of fuel and the capability to reach ultra-low sulphur concentrations. However, the successful commercialisation of BDS technology depends on the application of superior biocatalysts either through the continued improvements in known BDS

organisms and/or through the discovery and identification of new strains with robust BDS activity and high tolerance to the solvent mixture in fuels (Nuhu, 2013). Furthermore, laboratory-based studies must evaluate biocatalysts not only using model organosulphur compounds and model oils but fossil fuels directly. In this research, the latter route was taken for the development of BDS technology by looking for new BDS capable strains and testing their capabilities under real-world conditions.

In the current study, research will focus on the investigation of naturally occurring strains for BDS capabilities using both model organosulphur compounds and diesel. A selection of putative rhodococci strains, previously isolated from oil-polluted environments, will be identified using phylogenetic analyses and screened for biodesulphurisation activity. The BDS activity spectrum of positive strains, (i.e. ability to desulphurise BT and/or DBT) will be determined, initially in a wholly aqueous growth medium. The strains with BDS activity will then be tested for their ability to grow and express BDS activity in a biphasic medium containing *n*-hexadecane (model oil) and with gas oil (diesel) obtained downstream of an HDS process and reconstituted with added thiophenes (BT and DBT), to simulate real-world fuel conditions. The strains that exhibit the most desirable culture characteristics for potential commercial application will be identified. A known benzothiophene desulphurising bacterium, *G. desulfuricans* 213E$_T$ will be used a reference strain and positive control for experiments wherever suitable. This strain has been used for a patented bio-devulcanisation process by Recircle Ltd., UK for the removal of sulphur from rubber tyres. Its 16S rRNA gene sequence (AF101416) and whole-genome sequence (NZ_BCNF00000000.1) are known. In addition, whole-genome sequence analysis will be used to identify putative genes involved in the desulphurisation pathway which may be the target of future strain modifications. The overall aim of the research study is to identify one or more naturally occurring bacterium with suitable BDS activity and growth characteristics for further development in a full-scale commercial BDS process.

Chapter 2 Phylogenetic study of putative rhodococci isolated from hydrocarbon-contaminated environments

2.1 Introduction

Taxonomy is the study of identification, naming and classification of organisms. In this chapter, the complications in the taxonomy of the genus *Rhodococcus* is explained, which emphasizes the importance of establishing the taxonomic identity of the test strains used in the study given that they were isolated during the period when taxonomic lines were unclear. A brief review of the current standards followed in prokaryotic classification is described with a particular focus on genomic sequence based typing methods.

2.1.1 Biology of *Rhodococcus*

Rhodococci are Gram-positive (sometimes Gram-variable), obligate aerobes, catalase-positive, partially lysozyme sensitive, non-motile, non-endospore or non-conidia forming bacteria. They are chemoorganotrophic and have an oxidative type of metabolism. The cells appear as cocci during the early stages of growth, and as short rods as they mature, forming side projections or simple branching. These rod filaments usually undergo fragmentation to form the subsequent generations. Though few strains like *R. hoagii* require thiamine, most rhodococci can grow on standard cultivation media at 30°C, and visible colonies are formed in solid media between 1–3 days (Finnerty, 1992; Goodfellow et al., 1998). Rhodococci are known for diverse physiology, but in general, the cells have a thicker cell membrane dominated by the presence of an arabinogalactan cell wall polysaccharide and large 2-alkyl,3-hydroxy-branched-chain fatty acids (mycolic acids), which are covalently assembled into a peptidoglycan-arabinogalactan-mycolic acid matrix that renders them highly hydrophobic, which enables the bacteria to attach to oil/water interface while growing in aqueous–hydrocarbon environment (Borole *et al.*, 2002; Sutcliffe, 1998). Mycolic acids consist of peptidoglycolipids, glycolipids and lipids (Nishiuchi *et al.*, 2000). *R. opacus* has the unique ability to accumulate more than 50% of its dry cell mass as triacylglycerols (Fei *et al.*, 2015; Holder *et al.*, 2011; Kalscheuer *et al.*, 2000).

Although a few species are pathogenic, most of them are benign and thrive in a wide range of environmental niches, including soil, water, graves, plants and animals. The type strains of *R. rhodochrous*, *R. ruber*, *R. opacus*, *R.*

aetherivorans, *R. percolatus*, *R. coprophilus*, *R. degradans*, *R. erythropolis*, and *R. lactis* were isolated from activated sludge obtained from waste treatment or contaminated areas (Briglia *et al.*, 1996; Goodfellow *et al.*, 2004; Kumari *et al.*, 2015; Rowbotham and Cross, 1977). *R. jostii* was isolated from a medieval grave (Takeuchi *et al.*, 2002). *R. kyotonensis*, *R. kunmingensis* and *R. pyschrotolerans* were isolated from rhizosphere soil (Li *et al.*, 2007a; Silva *et al.*, 2017; Wang *et al.*, 2008). The type strains of *R. koreensis*, *R. pyridinivorans*, *R. phenolicus*, and *R. defluvii*, were isolated from industrial wastewater (Kampfer *et al.*, 2014; Rehfuss and Urban, 2005; Yoon *et al.*, 2000a, 2000b). Among the pathogenic rhodococci, *R. cercidiphylli, R. gannanensis,* and *R. trifolli* were isolated from plants (Kampfer *et al.*, 2013; Li *et al.*, 2008b; Ma *et al.*, 2017), and *R. hoagii*, *R. rhodnii*, *R. gordoniae, R. kronopolitis* and *R. triatomae* were isolated from live animals (Goodfellow and Alderson, 1977; Jones *et al.*, 2004; Liu *et al.*, 2014a; Yassin, 2005). Whilst the moderately halophilic type strains of *R. enclensis* and *R. marinonascens* were isolated from marine sediments (Dastager *et al.*, 2014; Helmke and Weyland, 1984), the type strains of *R. psychrotolerans*, *R. kroppenstedii*, and *R. baikonurensis* isolated from cold regions occur as psychrotrophic organisms (Mayilraj *et al.*, 2006; Silva *et al.*, 2017; Yoon *et al.*, 2010).

Rhodococci isolated from anthropogenically contaminated environments that contain complex aromatic compounds have been studied for a wide range of metabolic capabilities, including degradation of hydrocarbons and desulphurisation of fossil fuels. (Ceniceros *et al.*, 2017; McLeod *et al.*, 2006). They have a variety of plasmids ranging from small cryptic, closed circular plasmids to large linear ones (Larkin *et al.*, 2010; Matsui *et al.*, 2007; Van Der Geize and Dijkhuizen, 2004), and hence they have evolved with a number of genes encoding a wide range of metabolic capabilities. The 3 large linear plasmids of *R. jostii* RHA1 contain genes responsible for the degradation of polychlorinated biphenyls (McLeod *et al.*, 2006). The genes involved in the BDS activity of *R. erythropolis* IGTS8 occur in a large (150 kb) plasmid (Denis-Larose *et al.*, 1997). More about rhodococcal whole-genome sequence is described in section 4.1.1.

2.1.2 Role of polyphasic approaches in resolving rhodococcal systematics

Presently, to describe a new bacterial species, the authors must describe the isolate in terms of multiple parameters involving traditional microbiology and molecular biology methods, analytical methods and bioinformatics. This is a polyphasic approach integrating phenotypic, chemotaxonomic (identification of specific metabolites such as fatty acids, polar lipids, cell wall composition and exopolysaccharides) and genotypic (comparison of a selection of marker genes sequences, DNA–DNA hybridisation, G+C content variation) information about an organism and delineates microbial species based on a consensus of available data (Konstantinidis and Stackebrandt, 2013). These properties of the new isolate are used to distinguish its uniqueness from other previously validly described strains by comparison with well-curated databases for protein and nucleic acid. Taxonomy is richly informed by phylogenetics – the study of the evolutionary history of species (Choudhuri and Choudhuri, 2014). In the context of prokaryotic systematics, the term phylogenetic analysis refers to the identification of the tree best representing the evolutionary distances between selected semantides. A phylogenetic tree is a diagram that illustrates the lines of evolutionary descent of various species, organisms, or genes from a common ancestor.

However, it is not a surprise when the outcomes of the various methods in the polyphasic taxonomy toolbox do not entirely agree with the suggested description of the same isolate. The individual strains (isolates) can show variations in their phenotype and metabolic versatility despite tight genotypic clustering of strains, and *vice versa*. Organisms that pass the polyphasic approach are recognised as a validly published new species after their complete characterisation is described, their unique properties identified, and the etymology explained by publication in a recognised journal such as International Journal of Systematic and Evolutionary Biology (IJSEM).

2.1.3 Use of genomic sequences in resolving rhodococcal taxonomy

The idea of using sequence-based quantitative assessment of phylogenetic (evolutionary) relationships between representatives of the major known kinds

of organisms became popular in the late 1970s (Pace *et al.*, 2012). The macromolecules that are common to all cells and that change very slowly in time, called sementides, are used for understanding the phylogeny. The taxonomic assignments were done by comparing the nucleotide (nt) sequence to previously submitted reference sequences in the database (Huse *et al.*, 2008; Liu *et al.*, 2008; Wang *et al.*, 2007b) or clustering-based methods that identify taxon-independent operational taxonomic units (OTUs) using a sequence similarity threshold (Huse *et al.*, 2010; Schloss *et al.*, 2009; Schloss and Handelsman, 2005). OTU refers to a terminal leaf in a phylogenetic tree and is defined by a nucleic acid sequence. As the OTUs are frequently defined by the comparison of the respective sequences between organisms, sequence identity <95% are not considered to form part of the same OTU (Achtman and Wagner, 2008; Konstantinidis *et al.*, 2006).

Ribosomal RNA (rRNA) genes have been used as standard phylogenetic markers in molecular taxonomic studies since the pioneering studies on the tree of life by Woese and Fox (1977). There are a variety of reasons why rRNA genes have been selected as standard genes for molecular taxonomy. First, rRNA is an essential constituent in all living organisms. Second, the existence of many conserved regions in the rRNA genes allows the alignment of their sequences derived from distantly related organisms, while their variable regions are useful for the distinction of closely related organisms. Furthermore, the horizontal transfer of rRNA genes is believed to be rare.

With the 16S rRNA gene sequence becoming a standard molecular marker for the description of new species, NCBI maintains a curated bacterial and archaeal 16S rRNA dataset and allows a customised BLAST analysis over the '16S ribosomal RNA sequences (Bacteria and Archaea)' database (O'Leary *et al.*, 2016), which, as of May 2018, included 19,213 reference sequences for the bacterial 16S ribosomal RNA, of which over 95% were from type strains. A homology of ≥97% across the entire 16S rRNA gene or some variable region of the gene is considered the same OTU (Claesson *et al.*, 2010; Konstantinidis *et al.*, 2006). There are several curated 16S rRNA gene databases such as SILVA rRNA database (https://www.arb-silva.de/); the Ribosomal Database Project, RDP (https://rdp.cme.msu.edu/); EzBioCloud

(http://www.ezbiocloud.net/) that help steer clear of ambiguous unreviewed sequences and enable a faster approach to assign a new isolate to already known taxa, at least at the genus level (Oren and Garrity, 2014; Woo *et al.*, 2008).

Until 1996, a valid publication of new *Rhodococcus* species did not involve the 16S rRNA sequence based phylogenetic analysis, and the classification was based on numerical taxonomy involving their phenotypic characteristics, mycolic acid composition and the metabolic capabilities. The type strains of *R. coprophilus*, *R. erythropolis*, *R. fascians*, *R. globerulus*, *R. hoagii, R. marinonascens, R. rhodochrous*, *R. rhodnii*, and *R. ruber* were described before the year 1994, without their 16S rRNA gene information.

Revision of classification is not uncommon in bacterial systematics. Even before 16S rRNA gene analysis was the norm, *R. chlorophenolicus* was reclassified as *Mycobacterium chlorophenolicum* based on mycolic acid analyses (Hagglblom *et al.*, 1994). However, significant breakthroughs in the rhodococcal taxonomy happened as a result of 16S rRNA sequencing, and it started with the reclassification of the polychlorinated biphenyl-degraders *Acinetobacter* sp. P6 and *Corynebacterium* sp. MB 1 as *Rhodococcus globerulus* (Asturias *et al.*, 1994). The trend to include 16S rRNA gene sequence data began with the valid publication of the nutritionally versatile species *R. opacus* (Klatte *et al.*, 1994), and the type strain of *R. percolatus* MBS1^T described soon after included the phylogenetic tree showing how the 12 validly described rhodococci of that time clustered (Briglia *et al.*, 1996). Interestingly, their report revealed high levels of 16S rRNA gene sequence similarity (95.1-99.3%) between the 13 rhodococci species.

When the 16S rRNA gene sequence similarity is >98.7%, the species delineation was based by measuring the ability of the DNA fractions to form heteroduplexes under optimal conditions, usually 25°C below the melting temperature of the homoduplexes (with a maximum thermostability difference of 5°C). This method called DNA:DNA hybridisation (DDH) (Christensen *et al.*, 2000; De-Ley *et al.*, 1970; Ezaki *et al.*, 1989, 1988) was recognised as a way to resolve conflict in phylogenetic analysis (Murray *et al.*, 1990; Vandamme *et al.*, 1996). It has since been undertaken as a part of the polyphasic approach,

and DDH is the current gold standard, and a DDH value of <70% is the prevailing norm for the delineation of new species (Klenk *et al.*, 2014; Richter and Rossello-Mora, 2009). The type strain of *R. zopfii* T1^T with toxicant degrading capabilities was described based on DDH but did not include 16S rRNA gene analysis (Stoecker *et al.*, 1994). However, both approaches were included as a part of polyphasic taxonomy from then on, which led to several reclassifications of type strains across several genera and the genus *Rhodococcus* as follows,

- *Arthrobacter picolinophilus* DSM 20665^T was transferred to *R. erythropolis* (Koch *et al.*, 1995)
- *R. maris* was reclassified into a new genus, *Dietzia* (Rainey *et al.*, 1995b)
- *Tsukamurella wratislaviensis* N805^T was reclassified as *Rhodococcus wratislaviensis* (Goodfellow *et al.*, 2002)
- Despite being sourced from forest topsoil in the same region, the type strains of *R. humicola* and *R. pedocola* were distinguished by 33.05-35.60% DDH value, based on which they were recognised as two novel species (Nguyen and Kim, 2016)

The phylogenetic analysis of 10 *Nocardia* species and 22 strains of 16 *Rhodococcus* species conducted by Rainey *et al.* (1995a) resolved the close relationship of *Rhodococcus* with *Nocardia*. With a significant bootstrap value of 71% determined for the branching point that separated *Nocardia* and *Rhodococcus*, they concluded that *Nocardia* does not appear to be a sister taxon of *Rhodococcus* but branches off from within the radiation of *Rhodococcus*, and hence its species can be considered to be derived from a *Rhodococcus* ancestor. Based on 16S rRNA sequence analysis, Oberreuter *et al.* (2002) assessed that the *R. erythropolis* strains have a low intraspecific diversity, which indicated that having fewer strains included in the database could represent this species adequately for reliable identification.

When the number of sequences increases, evaluating the best fitting tree that reflects the phylogenetic relationship of the group becomes difficult, as the evaluation of each tree is a computationally-intensive process (Felsenstein, 1978). Bootstrapping is a resampling technique that is often used to evaluate

the tree topology and increase the confidence that the inferred tree is correct, and it is usually performed with 1000 resamplings for bacterial phylogenetic analyses (Efron, 1979; Felsenstein, 1985). There are many programs that use heuristic algorithms to infer the potentially best tree, but they come with their own advantages and drawbacks (Zhou *et al.*, 2018). The fastest and the most popularly cited phylogenetic method in systematics is the Neighbour-Joining (NJ) method, which uses a distance-based clustering program (Saitou and Nei, 1987; Van Noorden *et al.*, 2014). NJ is a heuristic approach that does not guarantee to find the perfect result, but under normal conditions has a very high probability of doing so. It has excellent computational efficiency, making it well suited for large datasets. NJ method is still considered a good starting point for more sophisticated methods (Nguyen *et al.*, 2015). However, statistical approaches, such as maximum likelihood (ML), produce more reliable results than distance and parsimony methods (Zhou *et al.*, 2018). IQ-TREE program is a fast ML-based phylogenetic program that offers ultrafast bootstrap approximation (UFBoot) to assess branch supports (Minh *et al.*, 2013). IQ-TREE explores the tree space efficiently and often achieves higher likelihoods than other programs like RAxML and PhyML (Nguyen *et al.*, 2015; Trifinopoulos *et al.*, 2016; Zhou *et al.*, 2018). Until 2010, only NJ tree was presented on the species description papers to show the relationship between the new *Rhodococcus* species with other type strains of *Rhodococcus*. For the valid description of *R. nanhaiensis* SCSIO 10187^{T}, phylogenetic trees were constructed using a combination of NJ and ML methods (Li *et al.*, 2012), and it has become a norm ever since.

16S rRNA gene sequences are now available for all rhodococcal type strains, and these are used as reference sequences for comparative analysis to determine the relatedness between strains. The lengths of these reference sequences range from 1,306 nt (for *R. hoagii*) to 1,537 nt (for *R. soli*). Actinobacteria constitute one of the most abundant and ancient taxonomic phyla within the domain bacteria and are well known for their secondary metabolites. Considerable variation in the metabolic properties, genome size and GC content of the members of this phylum have been observed. Therefore, the placement of new or existing species based on 16S rRNA gene sometimes becomes problematic due to the low congruence level (Barka *et*

al., 2016; Vandamme *et al.*, 1996; Verma *et al.*, 2013). Also, rRNA sequences do not discriminate well between closely related species or even genera, where only small numbers of substitutions occur between the compared rRNA sequences.

DNA-DNA hybridisation method inherently suffers from non-reproducibility and cannot be used to build incremental databases (Achtman and Wagner, 2008; Gevers *et al.*, 2005). It is widely criticised as an expensive, laborious and time-consuming technique (Li *et al.*, 2015; Whitman, 2015), that is not well standardised, and hence error-prone (Meier-Kolthoff *et al.*, 2013; Schweizer, 2008). Hence, research into describing alternative marker genes for bacterial identification are being explored.Where 16S rRNA gene sequence analysis is not sufficient, either the core genes that control functions such as cell division and metabolic activity (Gil *et al.*, 2004) or the house-keeping genes that code for proteins with conserved functions maintain cellular function (Martens *et al.*, 2008), have been used as OTU to assess phylogenetic diversity. These protein-coding genes are ubiquitous among bacterial species, evolve at a higher rate than the rRNA gene, but are not frequently transmitted horizontally. Apart from the 16S rRNA gene, genes coding for beta-subunit of DNA gyrase (*gyrB*), catechol 1,2-dioxygenase (*catA*), alkane 1-monooxygenase (*alkB*), protein export pathway (*secA*), the beta-subunit of RNA polymerase (*rpoB*), the sigma 70 (sigma D) factor of RNA polymerase (*rpoD*), recombinase A (*recA*), the beta-subunit of ATP synthase F0F1 (*atpD*), translation initiation factor IF-2 (*infB*), tRNA modification GTPase ThdF or TrmE (*thdF*), the chaperonin GroEL (*groEL*), sporulation-specific cell division protein (*ssgB*), heat shock protein (*grpE*), preprotein translocase subunit (*secY*) have been used to establish phylogeny of certain groups like bifidobacteria, mycobacteria and *Streptomyces* (Girard *et al.*, 2013; Táncsics *et al.*, 2015, 2014; Verma *et al.*, 2013). The phylogenies produced using single gene are unstable. Therefore, an integrative approach must be adopted where the phylogenetic information from individual markers are combined, and decisions are made from the consensus data.

The average base-substitution rate of 16S rRNA genes was 1% per 50 million years, while that of *gyrB* at the synonymous sites was estimated to be 0.7 - 0.8% per one million years (Yamamoto and Harayama, 1996). Therefore, some species with entirely identical 16S rRNA gene sequences can be differentiated by using their *gyrB* gene sequences (Kasai *et al.*, 1998; Peeters and Willems, 2011; Takeda *et al.*, 2010; Yamamoto *et al.*, 1999; Yamamoto and Harayama, 1995; Yáñez *et al.*, 2003). The divergence of 16S rRNA gene sequences among the members of *Nocardia, Gordonia* and *Rhodococcus* species is too low to satisfy taxonomic analyses. The usefulness and limits of the *gyrB* gene as an alternative marker for phylogenetic studies of these genera is discussed below.

After studying a number of protein-coding genes, Kasai *et al.* (1998) PCR-amplified the 1.2 kb-long *gyrB* segments from about 1,000 bacterial species using degenerate primers and determined their nucleotide sequences to establish a database. They proposed that the *gyrB* gene could be used as a potential taxonomic marker. Since then, many scientists have used it for further resolution of organisms to the species level as listed in Table 2.1. Following the suggestion of previous researchers, Shen *et al.* (2006b) developed new primer sets for the *gyrB* gene so as to identify, compare and assess diversity among members of the genus *Gordonia* based on the *gyrB* gene as a taxonomic marker. They observed that *gyrB* gene sequence greatly resolved to distinguish between the strains within *G. amicalis* and *G. terrae,* where the 16S rRNA gene sequences shared a high (>99.5%) sequence similarity with their respective type strains. However, in the case of 4 strains of *G. alkanivorans* isolated from oil-contaminated soil but from different geographic regions, i.e., DSM 44369^T, DSM 44187, DSM 44499 and CC-JG39, and 2 clinical strains of *G. rubripertincta*, i.e., DSM 43248 and JCM 3199, there was almost 100% similarities in both *gyrB* gene and 16S rRNA gene sequences, indicating the limitations in the use of *gyrB* as a taxonomic marker. They attribute the higher conservation (homogeneity) of *gyrB* genes in *G. alkanivorans* to environmental stress or natural selection pressure, and broad strain diversity among *G. terrae* strains to the natural ecosystems with minimum anthropogenic disturbance or from minimally contaminated environments from where they were isolated. They observed 16S rRNA gene

sequence similarities between 29 *Gordonia* species ranged from 93.1% to 99.8%, *gyrB* gene sequence similarities between 23 *Gordonia* species ranged from 77.5% to 97.3%. They later recommended using alkane hydroxylase gene (*alkB*) as an additional molecular marker for improved resolution (Shen *et al.*, 2010). However, *gyrB* sequence analysis is regularly undertaken for new *Godonia* species (*G. iterans*, *G. hongkongensis*, *G. phthalatica*) described from 2014 onwards (Chan *et al.*, 2016; de Menezes *et al.*, 2016; Jin *et al.*, 2017; Kang *et al.*, 2014).

As in the case of the polluted environment derived *Gordonia* strains, there was a high (>98.5%) *gyrB* gene homology between the type strains of *R. qingshengii* and *R. erythropolis* that were also isolated from polluted environments (Táncsics *et al.*, 2014). *Rhodococcus* species are well-known for their ability to degrade monoaromatic pollutants, and commonly possess catechol 1,2-dioxygenase (*catA*) genes in their chromosomes, which Táncsics *et al.* (2008) suggested as potential biomarkers.

Kang *et al.* (2009) *gyrB* gene showed the highest interspecies variability for *Gordonia*, and the next best separation was obtained using *secA1* gene analysis (whose similarity range is 81.9–98.0%). Takeda *et al.* (2010) analysed the phylogenetic relationship of 56 type species of *Nocardia* based on *gyrB* gene sequences and reported an interspecies similarity range of 82.4–99.9 % for *Nocardia gyrB* gene sequences, corresponding to nucleotide differences of 270–2 nt, against an interspecies similarity range of 94.4–100.0 % for the 16S rRNA sequences, corresponding to nucleotide differences of 75–0 nt. They reported that *gyrB* sequence clearly distinguished between the type strains of *N. paucivorans*, *N. brevicatena*, *N. carnea* and *N. flavorosea* that otherwise showed a high >99% similarity between their 16S rRNA gene sequences. However, discrepancies were observed in the phylogenetic positions of *N. exalbida, N. miyunensis* and *N. vaccinii* based on *gyrB* or 16S rRNA gene sequences.

Kirby *et al.* (2010) observed that the 390-nt sequence of the *gyrB* gene of a *Kribbella* isolate was enough to assess whether a strain is likely to represent a new species before time and effort is invested in the polyphasic taxonomic characterisation of the isolate. The genus *Flavobacterium* contained 33

extreme psychrophilic species occurring in Antarctica which show an interspecies 16S rRNA gene sequence similarity of 97.2–98.7%, whereas the *gyrB* similarity range within the genus was 79.1-94.9%, which showed its higher discriminatory power than the 16S rRNA gene (Peeters and Willems, 2011). Recently, it was reported that *gyrB* could be used successfully alongside the 16S rRNA gene to determine the species composition of food microbiota with improved resolution (Poirier *et al.*, 2018).

Table 2.1 Use of *gyrB* gene as a taxonomic marker for bacterial identification

Organism genus / group	Reference
***Pseudomonas putida* strains**	Yamamoto and Harayama, 1995
Acinetobacter	Yamamoto *et al.*, 1999
Aeromonas	Yáñez *et al.*, 2003
Salmonella, Shigella, E. coli	Fukushima *et al.*, 2002
***Bacillus cereus* strains**	Bavykin *et al.*, 2004
Gordonia	Chan *et al.*, 2016; Kang *et al.*, 2014; le Roes *et al.*, 2008; Li *et al.*, 2014; Liu *et al.*, 2011; Shen *et al.*, 2006a, 2006b
***Bacillus subtilis* strains**	Wang *et al.*, 2007a
Kribbella	Kirby *et al.*, 2010
Nocardia	Takeda *et al.*, 2010
***Rhodococcus* sp.**	Táncsics *et al.*, 2014

DDH involves only pairwise comparisons of two prokaryotic genomes, and the individual specimens cannot be comparatively analysed with a database using a defined set of standards. As the speed and costs of whole-genome sequencing have become more favourable, *in silico* methods based on the comparison of wholly sequenced genomes, have been suggested as an alternative to DDH for the identification of microorganisms (Auch *et al.*, 2010; Deloger *et al.*, 2009; Konstantinidis and Tiedje, 2005). This topic is described in detail in sections 4.1.4 and 4.2.3. The concept of obtaining average nucleotide identity (ANI) values between sequenced genome of bacterial species is actively endorsed as the future of bacterial taxonomy (Borriss *et al.*, 2011; De Vos *et al.*, 2017; Kim *et al.*, 2014; Konstantinidis *et al.*, 2006; Li *et al.*, 2015; Rodriguez-R and Konstantinidis, 2014).

2.1.4 Current phylogenetic structure of the genus *Rhodococcus*

The genus name *Rhodococcus* was first proposed by Zopf in the year 1891 and presently positioned under the phylum and class of Actinobacteria, the order of Actinomycetales, and family of *Nocardiaceae* (Goodfellow, 2014). In their review article, Goodfellow *et al.* (1998) clarify the twists and turns in the taxonomic history of the genus *Rhodococcus* and its nomenclature until the end of the twentieth century. In the past, 16S rRNA gene sequence based phylogenetic analysis confirmed the close relationship between the genera *Corynebacterium*, *Mycobacterium*, *Nocardia* and *Rhodococcus* and showed that these taxa formed a suprageneric group within the evolutionary radiation encompassed by actinomycetes. It supported the view that *Rhodococcus* species could be assigned to two aggregate groups each of which merited generic status and led to the revival of the genus *Gordonia* Tsukamura 1971 (Stackebrandt *et al.*, 1988) and proposal of the genus *Tsukamurella* (Collins *et al.*, 1988) containing members reclassified from genus *Rhodococcus*. At the end of the year 1996, there were 13 species described under *Rhodococcus* namely *rhodochrous* (type species), *erythropolis, equi, fascians, coprophilus, rhodnii, ruber, marinonascens, globerulus, opacus, zopfii, wratislaviensis, and percolatus*. Ever since the taxonomy of *Rhodococcus* has witnessed dramatic changes and developments thanks to the application of 16S rRNA and additional house-keeping genes to establish the phylogeny of these organisms and to identify and classify members of this and related groups. There are 40 new *Rhodococcus* species validly published since the year 2000, totalling 53 validly described species to date. The list of all validly described *Rhodococcus* species names is presented in Appendix 7.3.

Based on the phylogeny of 16S rRNA gene sequences, the members have been categorised into three subclades *R. erythropolis*, *R. equi (R. hoagii)* and *R. rhodochrous.* The members of the *R. erythropolis* subclade are used for many biotransformation and remediation applications. *R. erythropolis* was the first characterised bacteria in its subclade, while *R. jostii* was the first to have its genome sequenced (McLeod *et al.*, 2006). DNA-DNA hybridisation values revealed that type strains of *R. equi* and *Corynebacterium hoagii* were closely related. Because the oldest epithet, *hoagii* (Morse, 1912), had priority these

species were renamed in combination as *Rhodococcus hoagii* (Kampfer *et al.*, 2014; Sangal *et al.*, 2015; Tindall, 2014). It should be noted that *"Rhodococcus australis"* (Ferreira and Tracey 1984) was not effectively published.

2.1.5 Background information about the identity of the test strains

The strains used in this study were received from the Laboratory of Alkanotrophic Microorganisms at the Institute of Ecology and Genetics of Microorganisms (IEGM) that hosts a vast collection of non-pathogenic bacterial strains isolated from extremely diverse ecological regions in Russia. The IEGM is included in the World Federation for Culture Collections (acronym IEGM, WFCC#768; http://www.iegm.ru/iegmcol). The IEGM collection has a particular interest in maintaining microorganisms that oxidize natural and anthropogenic hydrocarbons and thus participate in the biogeochemical processes of the biosphere (Ivshina and Kuyukina, 2013). *Rhodococcus* strains that are predominant among the hydrocarbon-oxidizing microorganisms have been the core of the collection (Ivshina *et al.*, 1994; Ivshina and Kuyukina, 2018). *Rhodococcus* species held by the IEGM collection are represented not by single strains (often only type strains) but by numerous natural isolates from various habitats. The selected non-pathogenic strains of rhodococci with active oxygenase complexes are excellent targets for screening of new producers of valuable substances, degraders, and transformers of complex organic compounds and designing new effective technologies, production of fodder from unconventional sources (propane, n-butane), enzymatic transformation of carbon compounds, oil and gas prospecting, optimisation of secondary oil recovery, control and clean-up of hydrocarbon contamination of air and water, and bioremediation of oil-contaminated soils. The first strain of bacteria with a successful BDS activity was *R. erythropolis* strain IGTS8 developed by Institute of Gas Technology, USA (Kilbane and Jackowski, 1992). Rhodococci isolated from oil-polluted regions are preferred over other BDS strains as they have a broad substrate range and exhibit deep desulphurising activity (Ma, 2010). Therefore, 11 IEGM rhodococci strains (listed in Table 2.2) that were previously isolated from hydrocarbon-contaminated environments were chosen to study their potential

for biodesulphurisation (BDS) activity. The rationale for selecting rhodococci strains is discussed in detail in section 3.1.4.

The identity of the strains as received from IEGM was inferred from three sources – IEGM catalogue (http://www.iegm.ru/iegmcol/strains/index.html) available online, a printed version of the catalogue authored by Prof. Nick Christofi (1998), and a published research article by Bell *et al.* (1999). Nevertheless, between these 3 sources of information, the strain identity details concurred for strains 20, 60, 87, 213, 369, and 505; and differed for strains 208, 248, 488, and 508. Among the 11 test strains, 4 strains (20, 208, 213, 505) were classified as *R. erythropolis*, 2 strains (87, 369) were classified as *R. ruber*, and 1 strain (248) was classified as *R. opacus*. However, the species identity of 3 other strains (60, 488, 508) was ambiguous, and one strain was unknown (strain F).

At IEGM, for the purpose of cataloguing, *Rhodococcus* strains isolated from natural samples at sites with high anthropogenic loads are characterised by emulsifying and biodegrading abilities towards individual hydrocarbons and petroleum products, increased resistance to heavy metals (Cd, Cr, Cu, Mo, Ni, Pb, V, Zn), and maintenance of their activity under extreme acidity (pH 2.0–6.0) and salinity (2–6% NaCl). Even in the PCR based classification of IEGM strains by Bell *et al.* (1999) only a partial region of the 16S rRNA gene was targeted for amplification, and the species were assigned based on the strength of the amplification achieved by the species-specific primers developed by them. They had developed primers specific for *R. erythropolis* (Re), *R. ruber* (Ru), *R. globerulus* (Rg), and *R. opacus* (Ro), and used them primarily to verify the classification of strains determined by IEGM. In some cases, only weak amplification was observed leading to inconclusive results, e.g. primer Ro and strain 60. They reported slight cross-reactivity between primer Rr and *R. ruber* strains and primer Ro and *R. erythropolis* strains. The taxonomic identification of these test strains has not been updated to current standards which could be attributed to the operational difficulties faced by IEGM (Ivshina, 2012; Ivshina and Kuyukina, 2018, 2013). Clearly, the IEGM did not use the standard polyphasic approach for the classification of these strains, whereas prokaryotic systematics have progressed very much since.

The freeze-dried vials of the test strains used in the study were prepared in the early 1990s. The phylogenetic structure of the genus has changed since. Owing to the addition of several new closely related species with high 16S rRNA gene similarity, and because of the non-standard grounds on which the test strains were previously classified, there is a need to revisit their taxonomic classification and confirm the identity of the test strains based on currently recommended practices in phylogenetic analysis. It is essential to establish the taxonomic positions of the test strains prior to any investigation and future potential commercial application. This will help in choosing the right growth medium, precautions to be taken in terms of safety, choice of reference strains for comparative analysis, and their biotechnological capability could be attributed accurately.

2.1.6 Aim

To identify the test strains isolated from various hydrocarbon-contaminated environments to species level.

Objectives

1. To establish that the morphological and phenotypical properties of the reactivated strains that have been in long-term storage are in accordance with the previous description.
2. To determine whether the test strains contain mycolic acid in the cell membrane, and therefore belong to the suborder *Corynebacterineae*.
3. To determine the species identity of the test strains by phylogenetic analysis based on 16S rRNA and *gyrB* gene sequences.

2.2 Materials and Methods

2.2.1 Test strains

In this study, 11 naturally occurring bacterial isolates that were sourced from oil-polluted environments and deposited as various members of the genus *Rhodococcus* in the Regional Specialized Alkanotrophic Microorganism Collection of the Institute of Ecology and Genetics of Microorganisms (IEGM), Russia, were examined to confirm their taxonomic identity (Table 2.2). The validly described species *Gordonia desulfuricans* 213E$_T$, a benzothiophene biodesulphurising organism with a known 16S rRNA gene sequence (NR_028734) was used as a positive control. The 11 test strains were received as lyophilised biomass from Recyclatech Group Ltd., UK. Additional type strains of the genera *Corynebacterium, Dietzia, Gordonia, Mycobacterium, Nocardia, Rhodococcus* and *Tsukamurella* were obtained as -80°C cryopreserved cultures from the mycolic acid-containing actinomycete culture collection held at Edinburgh Napier University.

2.2.2 Cultivation, maintenance and preservation of test strains

Glass vials containing freeze-dried bacterial biomass were cut-open using a diamond cutter. About 0.5 - 1 ml of nutrient broth was added under sterile conditions and each pellet allowed to rehydrate for 30 seconds at room temperature. Each bacterial suspension was gently mixed and transferred to 10 ml of nutrient broth (NB) (Oxoid, UK) or Glucose Yeast Extract (GYE; Gordon & Mihm, 1962) medium (NCIMB growth media catalogue: 486) and incubated at 30°C with orbital shaking at 180 rpm (New Brunswick™ Innova® 44/44R, Eppendorf) for 3-4 days. A small aliquot of each bacterial suspension was streaked using sterile disposable loops (10µl) on the surface of Nutrient Agar (NA) (Oxoid, UK) plates and incubated, inverted at 30°C for 3-4 days to obtain single colonies. Liquid cultures were subsequently streaked for single colonies on NA and GYEA plate and incubated as above to check for purity by visual examination and Gram-stained smears (Hucker and Conn, 1923). All pure test strains were as plates or slopes at room temperature for no more than 2 weeks.

Preparation of glycerol stocks of the cultures

The glycerol suspensions were also prepared from 1 or 2 loopful of biomass scraped, using sterile disposable loops, from GYEA plates streaked for single colonies and incubated at 30°C for 5 days. Inocula were transferred to 2.5 ml cryo-vials and emulsified in approximately 1.5 ml of sterile glycerol solution (50% v/v) in order to obtain an even suspension of cells. Working frozen glycerol stocks provided a source of inocula and replica stocks served as a means of long-term preservation. Working inocula were obtained by thawing the glycerol suspensions at room temperature and transferring a small sample of the suspended biomass to GYEA plates before promptly returning the vials to the freezer.

Table 2.2. Isolation and classification details of the 11 test strains obtained from IEGM

Lab Code	IEGM strain code	Species	Geographical origin and isolation substrate	Classified based on	Known biotechnologically relevant traits
20	IEGM 20	*Rhodococcus erythropolis*	oil-polluted soil, Ukraine	Immuno PCR	Biosurfactants[1]
60	IEGM 60	*Rhodococcus sp.* – could be *R. opacus*	oil-polluted soil, oilfield, Ukraine	PCR[2]	Biosurfactants[1]; transforms thioanisole
87	IEGM 87	*Rhodococcus ruber*	core sample 39 m deep, Uljanovsk region, Russia	PCR	
208	IEGM 208	*Rhodococcus erythropolis*	oil-polluted soil, Ukraine	Immuno PCR[3]	
213	IEGM 213	*Rhodococcus erythropolis*	sewage, Kharbin, China	PCR	drotaverine hydrochloride-resistant
248	IEGM 248	*Rhodococcus opacus*	soil from a lavsan (polyether fibre) production facility, Belarus	PCR	
369	IEGM 369	*Rhodococcus ruber*	oil-polluted water, Sverdlovsk oilfield region, Russia	PCR 16S rRNA gene[4]	Biosurfactants[5]
Source: List of Species and Strains of IEGM Collection (http://www.iegm.ru/iegmcol/strains/index.html; accessed on 01/12/2017).					

[1] Produced biosurfactants when grown on *n*-alkanes (C10-C16)
[2] PCR using R. opacus specific primers Ro1 and Ro2, performed at the lower annealing temperature of 62 °C, gave weak positive signal
[3] gives a positive PCR with *R. opacus* specific primers even though it is still mentioned as *R. erythropolis*
[4] 16S rRNA gene sequence data is not available because it was published in Russian (Novoselova et al., 2011), and not deposited in public databases
[5] Produced biosurfactants when grown on n-alkanes (C12-C17)

Lab Code	IEGM strain code	Species	Geographical origin and isolation substrate	Classified based on	Known biotechnologically relevant traits
488	IEGM 488	*Rhodococcus erythropolis* – could be *R. opacus*	oil-polluted water, Unva river, oil-extracting enterprise area, Perm region, Russia	PCR[4]	
505	IEGM 505	*Rhodococcus erythropolis*	oil-polluted water, Nozhovskoe oilfield, Perm region, Russia	PCR	
508	IEGM 508	*Rhodococcus opacus*	N/A		
F	N/A	*Rhodococcus sp.*	N/A		
213E	N/A	*Gordonia desulfuricans 213T*	the soil in the vicinity of an oil shale spoil heap, West Calder, West Lothian, Scotland, UK	16S rRNA gene (AF101416)	Benzothiophene desulphurisation
Gordonia desulfuricans 213E$_T$ is not listed in the IEGM catalogue. It was obtained from the culture collection at Recyclatech Group Ltd., UK.					

Immuno: The strain demonstrated the reaction of complete identity with a polyclonal antiserum against *Rhodococcus erythropolis* IEGM 7$_T$ – original article Ivshina *et al.*, 1986
PCR: species-specific primers suggested by (Bell *et al.*, 1999)
Strain 488 – not found in IEGM online index (http://www.iegm.ru/iegmcol/strains/index.html)
Strain 505 – not mentioned in the printed 'IEGM CATALOGUE OF STRAINS' published by Prof. Nick Christofi (1998)
Strain 508 – mentioned only in (Bell *et al.*, 1999) and not listed in the web catalogue
Strain F – the label of the vial containing the lyophilised bacteria read "*Rhodococcus* sp." and did not include the species name

2.2.3 Colony and micromorphology characteristics

NA and GYEA plates were streaked for single colonies and incubated at 30°C for 3-4 days. The colony features of the test strains were examined after 4 and 6 days by both eye and using a binocular zoom plate microscope (Leica Stereo Microscope MDG41 using PLANAPO 1.0X objective). The colonies were examined for size and shape, elevation, margin, pigmentation, surface appearance, and for the development of substrate mycelium and aerial hyphae.

Air-dried smears prepared from 3, 4 and 5-day old cultures were Gram-stained (modified from Hucker & Conn, 1923) and their acid-fastness was studied (modified of the Zeihl-Neelson method). The stained preparations were viewed by bright-field microscopy under oil immersion using an x100 objective (Olympus BX51). Cell shape and the presence of hyphae and primary and secondary branching were noted.

2.2.4 Chemotaxonomy

2.2.4.1 Extraction and detection of mycolic acids

In this study, the validly described species *Corynebacterium amycolatum* S160^T (negative control); *Dietzia maris* N1015^T, *Gordonia bronchialis* DSM 43247^T, *Mycobacterium peregrinum* M6^T, *Nocardia brasiliensis* N318^T, *Rhodococcus rhodochrous* DSM 43241^T and *Tsukamurella paurometabola* DSM 20162^T were used as reference strains to study the mycolic acid profiles by Thin Layer Chromatography (TLC). Single colonies grown on GYEA plates incubated at 30ºC for 3 days were transferred to screw-capped conical flasks containing 150 ml of GYE broth. *Dietzia maris* was cultured in NB as it did not grow well in GYE. The caps of flasks were not tightly screwed, and the flasks were incubated with shaking at 180 rpm in an orbital incubator at 30ºC for 5 days. After incubation, Gram-stained smears were examined for purity, and the cultures harvested by centrifugation at 3200 *x g* for 20 minutes and washed three times with sterile distilled water. Harvested biomass was esterified by acid or alkaline methanolysis to derive extracts with mycolic acid and fatty acid methyl esters.

2.2.4.2 Acid methanolysis

The bacterial lipids were extracted and the mycolic acids and fatty acids derivatised to methyl esters using a modification of the whole-organism methods previously described (Hamid *et al.*, 1993; Minnikin *et al.*, 1980). Dried biomass (ca. 50 mg) of each test strain was transferred to 8.5 ml Corning glass test tubes fitted with PTFE lined screw caps (Corning, UK) containing 3 ml of methanolic sulphuric acid solution (methanol-toluene-concentrated sulphuric acid; 30:15:1, v/v/v). The test tubes were tightly closed, sealed with parafilm to prevent evaporation and methanolysis performed overnight in a covered water bath held at 75°C. After cooling to room temperature, 2 ml of n-hexane was added to each preparation, and the contents were shaken vigorously for 5 minutes to extract the methyl esters, prior to centrifugation at low speed (2000 rpm) for 5 minutes to separate the two layers. The upper organic layers, which contained the hexane extract, were transferred to clean tubes using a Pasteur pipette. A further 2 ml of n-hexane was added to each of the initial preparations, and the extraction process repeated. The upper organic layers were removed and combined with the first extracts. The pooled extracts were then transferred to small glass vials and evaporated under a stream of nitrogen gas at room temperature for approximately 20 minutes to concentrate the samples. Once dried, the crude methyl ester residues were redissolved in approximately 50 µl of petroleum ether and stored at 4ºC until further separation by thin-layer chromatography.

2.2.4.3 Analytical thin-layer-chromatography

The mixtures of fatty acid methyl esters (FAMES) and mycolic acid methyl esters (MAMES) were separated by one-dimensional ascending thin-layer-chromatography (TLC), as described by O'Donnell *et al.* (1982). Aluminium backed silica gel TLC sheets (Merck) (20 cm x 20 cm) were cut down to 10 cm x 20 cm in size. The samples were redissolved in 50 µl *n*-hexane then applied in small volumes (ca. 5µl), using a capillary pipette, to a line marked 1 cm from the base of each sheet. Spots were immediately dried using a hand-held hair dryer to prevent excessive spreading. Each TLC plate was loaded with 7 of the methanolysate samples from the reference strains and those prepared from the test strains. The aluminium plates loaded with samples were

developed in glass tanks containing light petroleum ether and acetone (92:8, v/v). The plates were removed when the solvent front approached the top, and the chromatogram air-dried then developed a second time in the same direction. The chromatograms were sprayed with a 5% solution of ethanolic molybdophosphoric acid (molybdophosphoric acid [10%, w/v] in ethanol [95%. v/v] (Sigma-Alrich, UK)) and immediately dried in an oven at 180°C for 5 minutes. The positions of the separated FAMES and MAMES were revealed as dark blue/black spots on a pale green background. The MAMES of the test strains separate according to chain length and were identified according to their relative mobilities described as Retention factor (R*f*) values by comparison against those derived from the reference strains.

Majidzadeh & Fatahi-Bafghi (2018) reported that TLC had been used as a tool to assign strains to one of 3 genus groupings based on the retention factor (R*f*): group 1 with R*f* ≥ 0.6 which consists of *Tsukamurella* and *Mycobacterium*; group 2 with 0.2 < Rf < 0.6 which consists of *Nocardia*, *Gordonia*, and *Rhodococcus*; and group 3 with Rf ≤ 0.2 which consists of *Corynebacterium*.

2.2.5 Genomic DNA extraction and purification

Various methods for the isolation of bacterial genomic DNA were applied and assessed using all the test strains. A standard DNA extraction procedure (Sambrook *et al.*, 1989) involving phenol:chloroform extraction and ethanol precipitation yielded poor-quality DNA from all the test strains. Subsequently, the ISOLATE II Genomic DNA Kit (Bioline, UK), which involves a similar chemical lysis method was tested, but this did not improve the quality or yield of the genomic DNA significantly. Further optimisation involving the addition of lysozyme (20 $mg.ml^{-1}$) to the kit lysis solution did not improve results.

To improve cell lysis, a kit incorporating physical lysis (bead-beating) was chosen. The PowerLyzer® PowerSoil® DNA Isolation Kit (MO BIO Laboratories Inc, Canada) designed to extract DNA from soil samples resulted in higher yields of gDNA from most strains, but the harsh steps in the protocol meant the DNA samples were prone to shearing.

Finally, the desired chemically enhanced bead beating process was achieved by using UltraClean® Microbial DNA Isolation Kit (MO BIO Laboratories Inc.,

Canada), which involves steps that lyse the bacterial cells using a combination of heat, detergent, and mechanical force with specialised beads. The test strains were grown in nutrient broth for 48 hours, and the genomic DNA was isolated following the manufacturer's protocol. To facilitate lysis without shearing, the preps were heated at 65°C for 10 minutes with occasional bump vortexing for a few seconds every 2-3 minutes. The isolated genomic DNA was suspended in 50 µl of 10 mM Tris, pH 8, and the concentration and purity were determined by Nanovue (NanoVue™ UV/Visible Spectrophotometers, GE Healthcare, UK) before storing at 4°C for further use.

2.2.6 PCR amplification of the 16S rRNA gene and *gyrB* gene

The genomic DNA extracts from the 11 putative rhodococci, and *Gordonia desulfuricans* 213E$_T$ were used as templates for the PCR amplification of the 16S rRNA gene. Primers for PCR amplification were purchased at the 200 nmol scales from Eurofins Genomics, UK, as dry, partially purified, precipitated DNA, and resuspended in PCR grade water (Bioline, UK) to provide primer at a 100 pmol.µl^{-1} stock concentration. The primer stocks were stored at -20°C, and they were used to prepare working stocks solutions with a concentration of 10 pmol.µl^{-1}.

The melting temperature (Tm) of each primer used for PCR amplification was estimated from the base content of each oligonucleotide sequence using a formula for oligonucleotides up to 25 bases in length (Thein & Wallace, 1986):

$$T_m\ (^oC) = 2(A + T) + 4(G + C)$$

The optimum annealing temperature was obtained by subtracting 5ºC from the estimated Tm value. The lowest temperature value of the two primers used in the PCR was selected as a starting annealing temperature from which PCR conditions were then optimised, where necessary.

The universal primers 27F (forward) and 1525R (reverse), used widely in the literature for the amplification of the 16S rRNA gene (Weisburg *et al.*, 1991) were selected. For the PCR amplification of the *gyrB* gene, TancF (forward) and TancR (reverse) primers were used. These were originally designed by Táncsics and co-workers (2014), by comparing the conserved domains in the amino acid sequence of *gyrB* protein found in *Rhodococcus jostii* RHA1

(CP000431), *R. opacus* B4 (AP011115), *R. equi* 103S (FN563149) and *R. erythropolis* PR4 (AP008957) and reverse transcribing them to matching nucleotide sequences. The forward primer TancF was a 20-mer oligonucleotide corresponding to GGKFDSD (*R. erythropolis* PR4 *gyrB* amino acid position 109–115). The reverse primer TancR was a 23-mer oligonucleotide targeting the amino acid sequence KIINVEKA (*R. erythropolis* PR4 *gyrB* amino acid position 486–493). The expected length of the amplified *gyrB* gene fragment using this primer set is 1,154 bp. The primers used in the study are listed in Table 2.3 below.

Table 2.3 Primers for the amplification of the 16S rRNA gene and *gyrB* gene

Primer	Sequence 5' – 3'	Length (bp)	Binding site[b] 5' – 3'	T_m (°C)	GC%
27F	AGAGTTTGATCATGGCTCA	19	8 - 27	52.4	42.1
1525R[a]	AAGGAGGTG**W**TCCARCC	17	1544 - 1525	54.0	55.9
TancF	GGCGGCAAGTTCGACTTCGA	20	325-375	61.4	60
TancR	GCCTTCTCGACGTTGATGATC	21	1457-1478	59.8	52.4

The primers 27F and 1525R were used to amplify the 16S rRNA gene, and the primers TancF and TancR were used for *gyrB* gene amplification.

[a]Degeneracy in the 1525R primer sequence according to Weisburg *et al.*, (1991): W = A:T;
[b]Binding site on the 16S rRNA molecule is the numbered according to the *Escherichia coli* numbering system (Brosius *et al.*, 1978) and the binding site on the *gyrB* gene sequence of *Rhodococcus erythropolis* PR4 known from its whole genome sequence available in the NCBI database (AP008957.1: 10128-12164)

PCR based amplification was carried out in 200 µl thinly walled PCR microfuge tubes containing BioMix™ Red (Bioline) a 2x reaction mix containing ultra-stable *Taq* DNA polymerase with *Taq* polymerase, deoxyribonucleotides (dNTPs), buffer and $MgCl_2$ (5 mM). Approximately 10ng of purified genomic DNA template was added to the PCR tube. In each PCR run, multiple samples were amplified, including negative controls lacking gDNA. PCR reagents were stored separately from nucleic acid samples at -20ºC.

The PCR mixture composition for both 16S rRNA gene and *gyrB* amplification was prepared as follows:

Component	Volume	Final concentration
BioMix™ Red (Bioline) x2	25 µl	
Forward Primer (10 pmol.µl^{-1})	5 µl	1 pmol.µl^{-1}
Reverse Primer (10 pmol.µl^{-1})	5 µl	1 pmol.µl^{-1}
Bacterial DNA template	Varies per sample	10 ng
PCR grade water	Varies per sample	
Total volume	50 µl	

PCR amplification was performed in a peqSTAR 96X Universal Gradient thermocycler (PEQLAB Biotechnologie GmbH, Germany) under the following conditions.

PCR Condition	16 rRNA gene	*gyrB* gene
Heat Lid	110°C	
Initial denaturation	95°C for 2 minutes	
Amplification cycles	30	
-dsDNA Denaturation	95°C for 1 minute	
-Annealing	45°C for 30 sec	52°C for 30 sec
-Extension	72°C for 1.5 minutes	
Final extension	72°C for 5 minutes, 1 cycle	
Store indefinitely	4°C, hold	

For each test strain, the 16S rRNA and *gyrB* genes were amplified from the DNA template in four separate reactions and subjected to gel electrophoresis. After confirming the amplification and integrity, the amplicons from the 4 repetitions were pooled and subjected to gel electrophoresis once again. The successful amplification of the 16S rRNA gene was visualised by the appearance of an apparent single band in line with the 1,500 bp marker band. The *gyrB* gene was visualised as a bright single band between the 1,000 bp and 1,500 bp marker bands, but closer to the 1,000 bp marker band. This approach reduced the effects of base insert error that occurs in PCR and

provided amplified product for purification and sequencing purposes. PCR products were stored at -20°C, ready for purification and sequencing.

2.2.7 Detection of DNA and integrity check

The presence and integrity of the extracted genomic DNA were checked by agarose gel electrophoresis to ensure that DNA samples were not excessively sheared or fragmented during the extraction procedure. Agarose gel (1%, w/v) was prepared by dissolving 1 g of agarose in 100 ml of 1x TAE buffer in a 250 ml conical flask. The mixture was dissolved by heating in a microwave oven in a 20-sec burst with gentle mixing until the solution became clear. Care was taken to avoid boiling the solution. The molten agarose was cooled to ~50°C in by placing the flask on a rotary shaker (150 rpm), and SYBR Safe™ DNA gel stain (Invitrogen) added to a final concentration of 0.5 $\mu g.ml^{-1}$ before casting the gel in acrylic gel tray (15 x 15 cm) fitted with a comb to provide ca. 10 µl volume capacity wells and allowed to set.

The gel was immersed into the electrophoresis tank containing 1x Tris-acetate-EDTA (TAE) buffer, and the comb was removed. The 1x TAE buffer contained 40 mM Tris, 20 mM acetic acid, and 1 mM EDTA at pH 8.0. Genomic DNA (2 µl) from each test strain was transferred to a 200 µl microcentrifuge tube containing 6 µl of PCR grade water and 2 µl of 5X gel loading buffer (Bioline) and mixed to give 10 µl sample, of which 5 µl was used to load into each submersed well. Five µl of DNA molecular size marker (1 kb Hyperladder; Bioline) was added to the first well. The DNA samples were then electrophoresed at a constant voltage (50 mV) for 90 minutes or until the marker dye approached the end of the gel. The gel was briefly exposed to UV light from a UV transilluminator (ChemiDoc™ XRS+ System, Bio-Rad) and visualised in the computer using Image Lab™ Software (Bio-Rad). The genomic DNA samples appeared as slowly migrated samples close to the loading wells. Sheared DNA appeared as a smear along the track of the sample lane, in which case the sample was rejected and necessitated modification of the DNA extraction procedure.

The purity and quantity of genomic DNA were also determined by spectrophotometry (NanoVue Plus Spectrophotometer, GE Healthcare). The

machine was set to measure DNA, and about 2 µl of the DNA sample was pipetted directly onto the sample plate. After the measurement, the sample was discarded by wiping the sample plate clean for the next measurement. Absorbance values were determined at 230, 260 and 280 nm, using the final elution buffer from the kit as blank. Examination of sample spectra was used to confirm the purity of the samples. NanoVue Plus displayed both individual absorbance values and absorbance ratios (260/280 and 260/230) on the screen, along with the sample concentration value.

The corresponding reading at 230 nm measured any contamination by small molecules such as polysaccharides and a ratio between the readings at 260nm and 230nm (A260/A230) in the range of 2.0–2.2 was indicative of a carbohydrate-free sample. The 260/230 values for “pure” nucleic acid are often higher than the respective 260/280 values. The reading at 280 nm provided an indication of protein contamination, and a ratio of A260/A280 of about 1.8 indicated a protein-free sample (Sambrook *et al.*, 1989). DNA samples with low ratio values were subjected to a further clean-up step so as to avoid any interference from residual phenol from nucleic acid extraction or residual guanidine (often used in column-based kits).

PCR products were obtained using Biomix-Red PCR mix (Bioline) and already contained an additional inert red dye that permitted easy visualisation and direct loading to 1% w/v agarose gels. Five µl of PCR product was loaded to each well directly, and the remaining steps in the electrophoresis performed as described above. The PCR amplicons appear as migrated bands in line with the correspondingly sized marker band Appendix 7.2. Any additional or unexpected bands along the track of the sample lane indicated the presence of non-target amplicons and necessitated either PCR optimisation or carefully avoiding the unwanted bands while proceeding with PCR product clean-up by gel slicing.

2.2.8 Sequencing of purified 16S rRNA and *gyrB* gene

The pooled amplicons were cleaned up using Wizard® SV Gel and PCR Clean-Up System (Promega) following the manufacturer procedure. In the case of the 16S rRNA gene amplicon, the PCR product was subjected to clean

up directly using the kit. In the case of *gyrB* gene amplicon, the pooled amplicon was run on 1% agarose gel, and the bands corresponding to the 1.2 kb marker were sliced out using an x-tracta gel extraction tool (Sigma-Aldrich, UK), and the amplicon was recovered from gel slice using the kit. The clean PCR products were run on 1% agarose gel to check for purity.

The clean PCR products were prepared for sequencing by Eurofins Genomics, UK following the procedure described for their PlateSeq Service. The samples were prepared by mixing 2 µl of the primer (10 pmol.µl^{-1} primer concentration) with 15 µl of purified PCR product (10 ng.µl^{-1} DNA template) in 1.5 ml microcentrifuge tubes or into wells of a custom-designed microtitre plate. The concentration of primers with wobble bases was calculated according to the following formula: n^X * 10 pmol.µl^{-1}, where n = number of bases within a wobble according to IUPAC code; X = number of wobbles within the primer sequence. In the case of the 16S rRNA gene, the internal primers listed in Table 2.4 were included for sequencing purposes. Internal sequencing primers were not used in the case of *gyrB* sequencing owing to the lack of data on conserved regions of the gene sequence across a broad group of bacteria.

Table 2.4 Internal primers used for 16S rRNA gene sequencing

Primer	Sequence 5' – 3'	Region covered on the nucleotide sequence of 16S rRNA gene	Concentration used in the sequencing sample (pmol.µl^{-1})
27F	AGAGTTTGATC**M**TGGCTCAG	1-27	20
MG4f*	AATTCCTGGTGTAGCGGT	600-610	10
782r*	ACCAGGGTATCTAATCCTGT	710-720*	10
MG5f*	AAACTCAAAGGAATTGACGG	840-850	10
1525R	AAGGAGGTG**W**TCCA**R**CC	14	40

*internal primers

The nucleotide sequences were obtained in both FASTA format and as a tabulated form in spreadsheets. The quality of the sequencing results was cross-checked with the corresponding chromatograms. The quality report is the trace file in *.pdf format in which the quality value of every single base is shown in colour code below every single peak. Different colours represent the four specified quality ranges.

2.2.9 Consolidation of sequencing data

The reliability of the nucleotide sequence data for the 16S rRNA gene and *gyrB* gene, obtained from Eurofins Genomics, UK was checked. The sequences obtained were assembled using the CAP3 Assembler programme (Huang, 1999) (http://doua.prabi.fr/software/cap3) to check for overlapping regions and to assemble the sequence fragments into a single full gene sequence. The fully assembled sequences were analysed.

The 16S rRNA gene sequences were subjected to BLAST® analysis using the 'blastn' suite over the subset of all 16S ribosomal RNA gene sequences in the NCBI database, to look at the degree of alignment with previously reported bacterial 16S rRNA sequences. The *gyrB* gene sequences were subjected to 'blastn' analysis over the entire non-redundant nucleotide sequences available in the NCBI database to look for highly similar sequences.

The near-complete 16S rRNA gene sequences for the 11 strains and corresponding sequences of the type strains of 51 validly described *Rhodococcus* species and 13 type species representing the other genera (*Corynebacterium, Dietzia, Gordonia, Millisia, Skermania, Hoyosella, Mycobacterium, Nocardia, Rhodococcus, Smaragdicoccus, Segniliparus, Tsukamurella, Williamsia*) of the Order Corynebacteriales retrieved from the EMBL-GenBank-DDBJ (Benson *et al.*, 2000) and the RDP (Ribosomal Database Project; Maidak *et al.*, 1997) databases were aligned using the algorithms MUSCLE (Edgar, 2004) and MAFFT (Katoh *et al.*, 2002; Katoh and Standley, 2013) and pairwise evolutionary distance matrices calculated. The MUSCLE alignment of the sequences was further used to construct the phylogenetic trees. The almost complete *gyrB* gene sequences (around 1,020 bp) for the 11 strains were aligned with corresponding sequences of the type strains of 22 validly described *Rhodococcus* species.

2.2.10 Construction of phylogenetic trees

After aligning the sequences using MUSCLE algorithm, the phylogenetic trees were reconstructed based on the neighbour-joining (Saitou and Nei, 1987) and maximum-likelihood (Felsenstein, 1981) tree-mapping algorithms using Mega 7 suite (Caspermeyer, 2016) and the Species Delimitation plugin (Masters *et*

al., 2011) of the Geneious software suite (Biomatters Ltd http://www.biomatters.com). Tree topologies were evaluated based on bootstrap analysis of 1000 datasets.

The MUSCLE aligned sequences were subjected to phylogenetic analysis using IQ-Tree (Nguyen *et al.*, 2015) which is a fast and effective stochastic algorithm to infer phylogenetic trees by maximum likelihood with similar computing time as RAxML and PhyML packages. The relationships between the test strains and the reference strains were obtained in terms of percentage similarity and number of different nucleotides.

2.3 Results

The 11 bacterial test strains used in this research were initially isolated from oil-polluted soils in Russia in the early 1990s. These alkanotrophic organisms are biotechnologically important bacterial cultures involved in the degradation of oil hydrocarbons (including gaseous *n*-alkanes) and other natural and synthetic organic compounds. They have been preserved as members of the genus *Rhodococcus* at the IEGM, Russia. The basis of their taxonomical classification has remained vague, and the only verifiable source of information regarding their classification is the PCR based detection method developed by Bell *et al.* (1999) who had designed a set of species-specific primers targeting the 16S rRNA gene to obtain partially amplified sequences. Since 1999, several closely related new species have been described within the genus *Rhodococcus*. Therefore, it was essential to ensure the validity of the imprecise classification of the test strains using the 16S rRNA gene and *gyrB* phylogeny and other micromorphological and chemotaxonomical characteristics.

2.3.1 Mycolic acid profile of the test strains

Mycolic acids were extracted by acid methanolysis and analysed according to the protocol of Minnikin *et al.* (1980). The patterns obtained following one-dimensional thin-layer chromatography of whole organism methanolysates of the 11 strains revealed the general mycolic acid composition of these organisms as multi-spot patterns that correspond to the different types of mycolates found in them, as shown in Figure 2.1. The components with Rf values of 0.1-0.5 usually correspond to mycolic acid methyl esters (MAMES) methyl esters of mycolic acids, whereas those with higher Rf values (0.8-1.0) are attributable to non-hydroxylated fatty acid methyl esters (FAMES). Longer carbon chains have low polarity, resulting in faster migration on the plates. The expected order of migration distance of the MAMES bands for the samples used in this research is *Mycobacterium* > *Tsukamurella* > *Gordonia* > *Nocardia* > *Rhodococcus* > *Dietzia* > *Corynebacterium*.

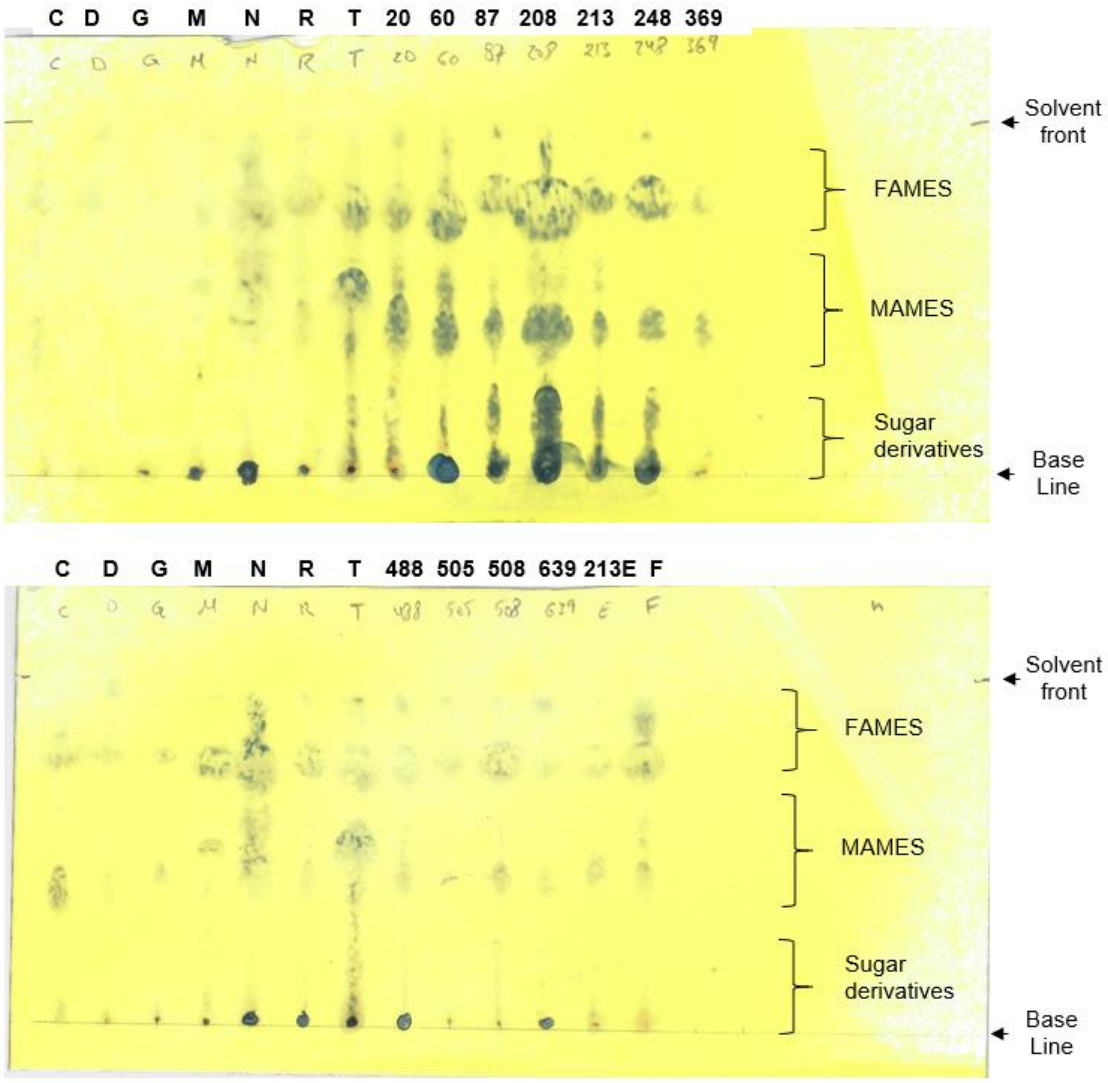


Figure 2.1 TLC of whole organism methanolysates of the 11 test strains and the reference strains from other genera under the order Corynebacteriales

Whole-cells of the test strains were lyophilised, and the mycolic acids were extracted using the acid-methanolysis method. The mycolic acid methyl esters (MAMES) and fatty acid methyl esters (FAMES) in the methanolysate extracts of **(C)** *Corynebacterium amycolatum* S160^T (negative control), **(D)** *Dietzia maris* N1015^T, **(G)** *Gordonia bronchialis* DSM 43247^T, **(M)** *Mycobacterium peregrinum* M6^T, **(N)** *Nocardia brasiliensis* N318^T, **(R)** *Rhodococcus rhodochrous* DSM 43241^T and **(T)** *Tsukamurella paurometabola* DSM 20162^T and test strains **20, 60, 87, 208, 213, 248, 369, 488, 505, 508,** *G. desulfuricans* **213E**^{T} and strain **F** were detected by one-dimensional thin layer chromatography using solvent petroleum ether and acetone (92:8, v/v) and after development with spray solvent 5% solution of ethanolic molybdophosphoric acid and developed at 180 °C for 5 mins. The strain **639** shown on the chromatogram is not a part of this research.

The chromatogram also revealed the non-standardisation of the quantity of sample loaded on the baseline, especially in the case of samples such as *Nocardia*, *Tsukamurella*, 60 and 208 that were clearly overloaded, which subsequently reduced the resolution and caused band overlaps. The absence of well-defined bands in the case of samples *Gordonia*, *Dietzia*, *Rhodococcus*, 369, 488, 505, 639 and 213E indicates that the samples were too dilute. Although the appearance of FAMES band was expected in the case of the negative control (*Corynebacterium amycolatum*), the chromatogram shows a weak MAMES band which advocates revisiting the authenticity of the strain received. The relative front (R*f*) value of each mycolic acid spot was compared with those of the reference strains included on each chromatogram. The methanolysates of all the test strains produced MAME spots with R*f* values ~0.5 which corresponded most closely with the MAMES of the reference strains *Gordonia bronchialis* DSM 43247^T, *Nocardia brasiliensis* N318^T and *Rhodococcus rhodochrous* DSM43241^T. Although the migration of the bands is not clearly visualised precisely in the TLC plate images (Figure 2.1), the observations confirmed the presence of mycolic acid methyl esters in all the test strains, indicating that they are mycolic acid-containing actinomycetes and hence belong to the suborder *Corynebacterineae*.

2.3.2 Phylogeny of the strains based on 16S rRNA gene and *gyrB* gene sequence analyses

BLASTn analysis

The 16S rRNA gene nucleotide sequences obtained for the 11 strains (coded 20, 60, 87, 208, 213, 248, 369, 488, 505, 508 and F) were at least 1,495 bp long. These sequences were compared with the entire non-redundant nucleotide sequence database (45,538,299 sequences available as on 07/12/2017), and with the curated database of all 16S ribosomal RNA sequences from bacteria and archaea available at the NCBI (19,573 sequences available as on the date 07/12/2017), using the Basic Local Alignment Search Tool programme, optimised for Highly Similar Sequences (megablast). Additionally, the BLAST settings were adjusted to include only reference 16S rRNA gene sequences of type strains. BLAST finds regions of

similarity between nucleotide sequences and calculates the statistical significance of these. The *gyrB* nucleotide sequences for the 11 strains were at least 1,066 bp long. These sequences were compared with the entire non-redundant nucleotide sequence database, using BLAST analyses. The top hits from the BLAST analysis for the 16S rRNA gene and *gyrB* sequences of the test strains are shown in Table 2.5, and results obtained for each strain is presented in Appendix 7.4. The 16S rRNA gene sequence of *R. degradans* CCM 4446^{T} was not included in the curated 16S rRNA database for type strains available at the NCBI. However, it was available in the general non-redundant nucleotide database (accession number: JQ776649) which was used in the detailed phylogenetic analysis conducted later in the study. Therefore, the blast results presented in Table 2.5 do not include the *R. degradans* strain, and instead, *Nocardia coeliaca* DSM 44595 is presented as the top hit. Based on the revised 16S rRNA gene sequences of the type strains of *N. coeliaca* and *N. globerula*, Kampfer *et al.* (2014) suggested the reclassification of these species into the *Rhodococcus* genus.

Table 2.5 Top hits from the NCBI nucleotide database obtained for the BLASTn analysis of 16S rRNA gene and *gyrB* sequences of the test strains

Strain	Test strain 16S rRNA gene length (bp)	Top BLAST hit from the curated 16S rRNA sequence database for the 16S rRNA gene sequence	Accession code of the 16S rRNA gene of the top hit	% similarity		Test strain *gyrB* gene length (bp)	BLAST hit from the entire non-redundant nucleotide collection for *gyrB* gene sequence	Accession code of *gyrB* gene of the top hit	% similarity
20	1494	+*Nocardia coeliaca* DSM 44595	NR_104776	99		1066	*R. erythropolis* ATCC 47072	AB014272	99
60	1499	*R. rhodochrous* 372	NR_037023	99		1076	**R. rhodochrous* NCTC 10210	LT906450	99
87	1502	*R. ruber* DSM 43338	NR_026185	100		1073	**R. ruber* IFO 15591 (DSM 4338)	AB014174	100
208	1492	*R. jostii* IFO 16295	NR_024765	99		1079	**R. erythropolis* DSM 1069	AB014111	90
213	1493	*N. coeliaca* DSM 44595	NR_104776	99		1070	*R. globerulus* ATCC 21506	AB014251	99
248	1495	*R. wratislaviensis* NCIMB 13082	NR_026524	99		1080	*R. erythropolis* DSM 1069	AB014111	99
369	1501	*R. ruber* DSM 43338	NR_026185	100		1077	**R. ruber* IFO 15591 (DSM 4338)	AB014174	99
488	1460	*R. koreensis* DNP505	NR_114500	99		1076	**R. erythropolis* DSM 1069	AB014111	90
505	1495	*N. coeliaca* DSM 44595	NR_104776	99		1070	*R. erythropolis* JAM 1484	AB014172	99
508	1495	*R. jostii* IFO 16295	NR_024765	99		1078	**R. erythropolis* DSM 1069	AB014111	90
F	1495	*N. coeliaca* DSM 44595	NR_104776	99		1066	*R. erythropolis* JCM 2892	AB355724	99

The blastn output obtained for the 16S rRNA and *gyrb* gene sequences of the test strains showed multiple sequences in the database with high pairwise similarity (>98%) with the query sequence. The top hit for each gene's sequence was selected on the basis of high max score, % query cover values, i.e. how much of the query sequence is covered by the target sequence, and high % identity values. In the case of *gyrB* gene sequence, the hits from whole-genome sequencing projects and unclassified *Rhodococcus* species were excluded.

*There were other sequences with equally high similarity values. However, the top hit was selected based on concurrency with 16S rRNA gene BLAST results.
+Reclassification of *Nocardia coeliaca* into the genus *Rhodococcus* has been proposed (Kampfer *et al.*, 2014). The 16S rRNA gene sequence of *N. coeliaca* shares that of *R. degradans* CCM 4446 (NR_145886.1) that was included in the curated 16S rRNA gene database for type material on a later date.

Phylogenetic trees

Unrooted phylogenetic trees were constructed using the Neighbour-Joining (Figure 2.2) and IQ-TREE (Figure 2.3) methods. It is evident from both phylogenetic trees that the 11 strains examined in this study are located within the evolutionary radiation encompassed by validly described members of the genus *Rhodococcus,* which confirms the original genus assignment of these strains. Further, the 11 strains formed independent lines of descent representing several distinct species. Strains 208, 488 and 508 formed a tight cluster in both the NJ and IQ-TREE based methods trees, supported by bootstrap values of 99% in both cases. Strains 20, 213, 505 and F also formed a tight cluster that included *R. degradans* in both trees (bootstrap values of 64% and 99%, respectively). Strains 87 and 369 are most closely related and cluster with *R. ruber* in both trees (bootstrap values of 97% and 100%, respectively). Strains 60 and 248 formed single lines of descent in distinct parts of the *Rhodococcus* phylogenetic tree.

Unrooted phylogenetic trees based on the *gyrB* gene were constructed using the NJ (Figure 2.4) and IQ-TREE (Figure 2.5) methods. The 11 strains were positioned/located within the evolutionary radiation encompassed by the validly described *Rhodococcus* species for which the *gyrB* gene sequence data were published, in agreement with the results of the 16S rDNA analysis. The clustering of the *gyrB* gene sequences followed similar trends as observed in the case of 16S rRNA gene sequences such that the strains 208, 488 and 508 formed an independent cluster, strains 87 and 369 formed clustered with *R. ruber,* strains 60 and 248 formed single lines of descent in distinct parts of the *Rhodococcus* phylogenetic tree that included *R. rhodochrous* and *R. wratislaviensis,* respectively. Strains 20, 213, 505 and F also formed a tight cluster that included *R. degradans, R. erythropolis and R. qingshengii.*

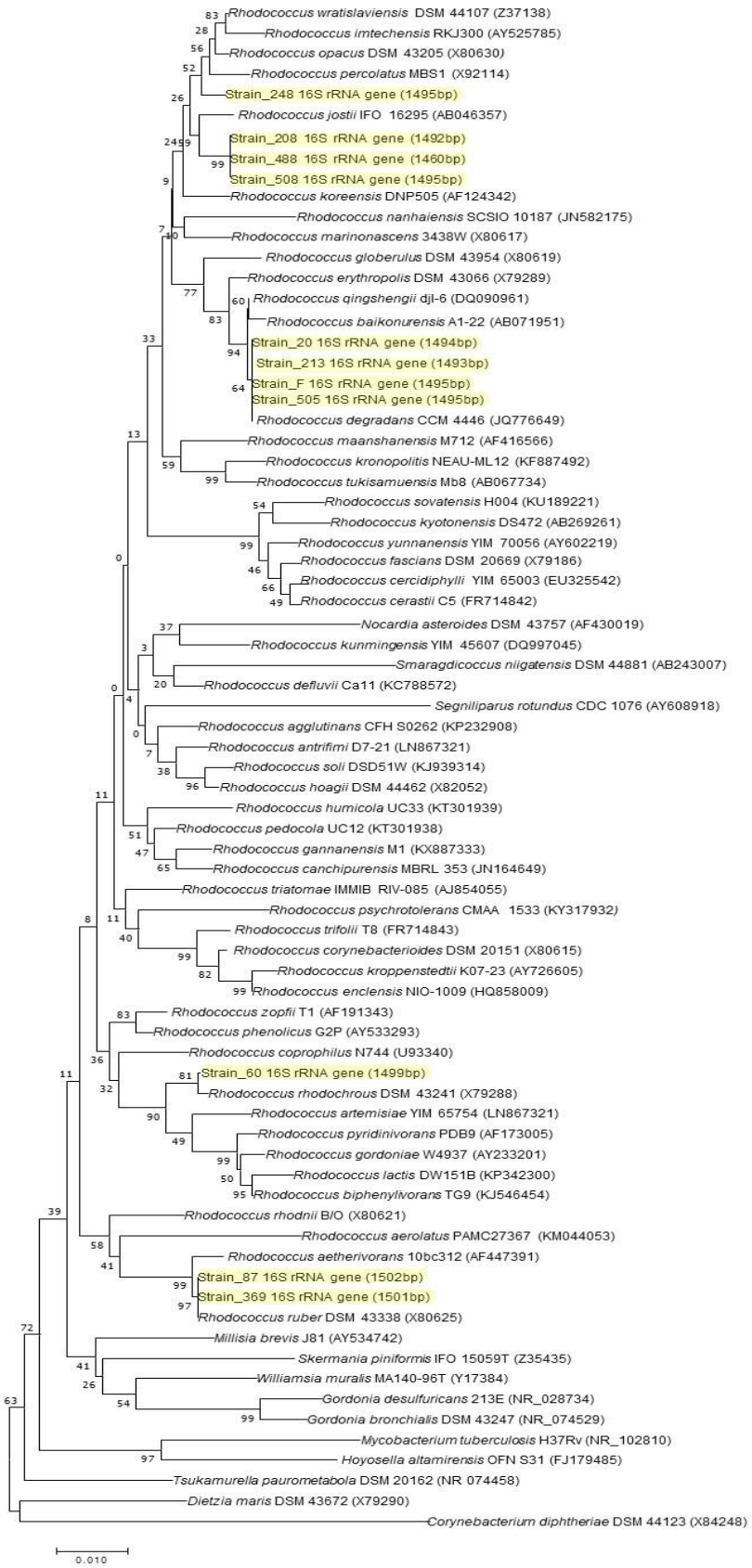


Figure 2.2 NJ-TREE based on the 16S rRNA gene sequences

Neighbour-joining tree showing the 16S rRNA gene sequence based phylogenetic relationship of the 11 test strains with 51 type strains of genus *Rhodococcus* and type species from 13 genera belonging to Order Corynebacteriales. Numbers at nodes indicate bootstrap support values based on 1000 resampled dataset. Bar, 1 nt substitution per 100 nt. The test strains are highlighted in yellow.

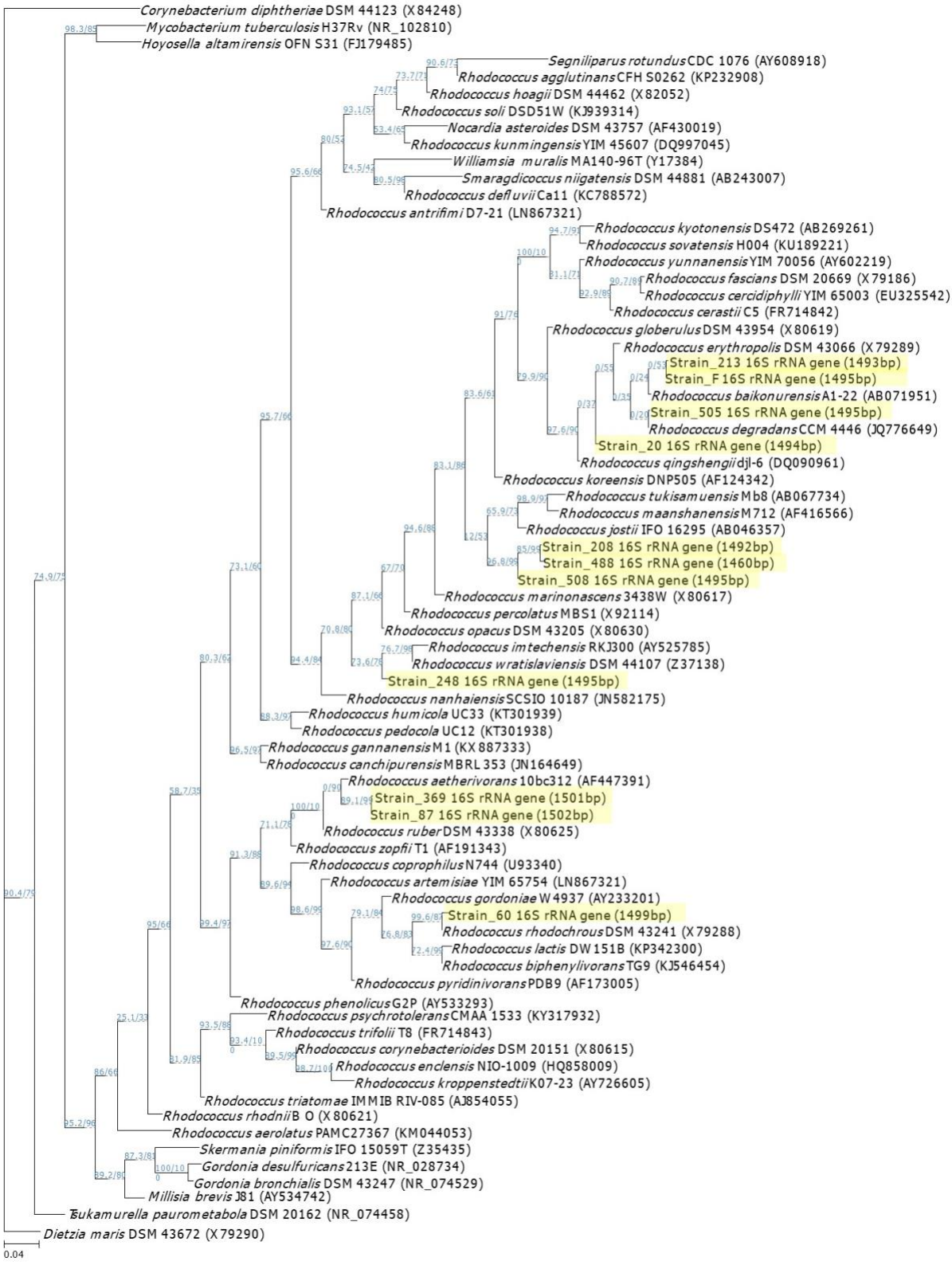


Figure 2.3 IQ-TREE based on 16S rRNA gene sequences

The output from the web-based IQ-TREE program showing the 16S rRNA gene sequence based phylogenetic relationship of the 11 test strains with 51 type strains of genus *Rhodococcus* and type species from 13 genera belonging to Order Corynebacteriales. Numbers at nodes indicate bootstrap support values based on 1000 resampled dataset. Bar, 5 nt substitutions per 100 nt. The tree is UNROOTED although outgroup taxon *Corynebacterium diphtheriae* DSM 44123 (X84248) is drawn at the root. Numbers at the nodes are SH-aLRT support (%) / ultrafast bootstrap support (%). Bar, 5 nt substitutions per 100 nt. The test strains are highlighted in yellow.

Examination of the NJ and IQ Tree topologies reveals that members of the genus *Rhodococcus* form a distinct clade within the suborder *Corynebacterineae,* which is made up of several deep-rooted subclades but is distinct from the type species of the other genera and supported by high bootstrap values. Phylogenetic examination of the 51 validly described species of the genus *Rhodococcus* reveals that the percentage similarity of 16S rRNA gene sequences amongst the type strains range from a maximum of 99.78% between *R. baikonurensis* A1-22 (AB071951) and *R. degradans* CCM 4446 (JQ776649) and a minimum of 87.77% between *R. tukisamuensis* Mb8 (AB067734) and *R. lactis* DW151B (KP342300), with a mean value of 95.63%, which corresponds to nucleotide (nt) differences of 3 to 170, with a mean difference of 62 nt. Across all the 14 type species under the order Corynebacteriales, the type strain of the type species of the genus *Millisia* (*M. brevis* J81^{T}) was observed as the closest relative of the type strain of the type species of the genus *Rhodococcus* (*R. rhodochrous* DSM 43241^{T}) with 94.79% similarity (Table 2.6). Members of the genus *Rhodococcus* show lower percentage similarities with the valid representative species belonging to the other genera under the order Corynebactyeriales. Therefore, on the basis of the 16S rRNA gene sequences the eleven test strains used in this study could be reliably classified under the genus *Rhodococcus*. For further resolution in their identification at the species level, *gyrB* gene-based phylogenetic analyses were undertaken.

Table 2.6 Percentage similarity between the 16S rRNA gene sequence of *Rhodococcus rhodochrous* DSM 43241T and other type species belonging to various genera under the Order Corynebacteriales

Type species belonging to various genera under Order Corynebacteriales (16S rRNA accession code)	% similarity with *R. rhodochrous* DSM 43241T (X79288)
Corynebacterium diphtheriae **DSM 44123T (X84248)**	91.47
Dietzia maris **DSM 43672T (X79290)**	94.49
Tsukamurella paurometabola **DSM 20162T (NR 074458)**	94.44
Millisia brevis **J81T (AY534742)**	94.79
Nocardia asteroides **DSM 43757T (AF430019)**	93.58
Gordonia bronchialis **DSM 43247T (NR 074529)**	94.44
Williamsia muralis **MA140-96TT (Y17384)**	94.52
Smaragdicoccus niigatensis **DSM 44881T (AB243007)**	94.37
Skermania piniformis **IFO 15059T (Z35435)**	93.88
Segniliparus rotundus **CDC 1076T (AY608918)**	93.06
Hoyosella altamirensis **OFN S31T (FJ179485)**	93.05
Mycobacterium tuberculosis **H37RvT (NR 102810)**	91.61
Turicella otitidis **DSM 8821T (X73976)**	85.92
Parvopolyspora pallida **DSM 43888T (AB006157)**	89.03

The table shows the percentage similarity between the 16S rRNA gene sequence of type species of genus *Rhodococcus* (*R. rhodochrous* DSM 43241T) and the type species of various genera under order Corynebacteriales. Based on the multiple sequence alignment of the 16S rRNA using MUSCLE algorithm, *Millisia brevis* J81T was found as the closest relative of the type species of genus *Rhodococcus*.

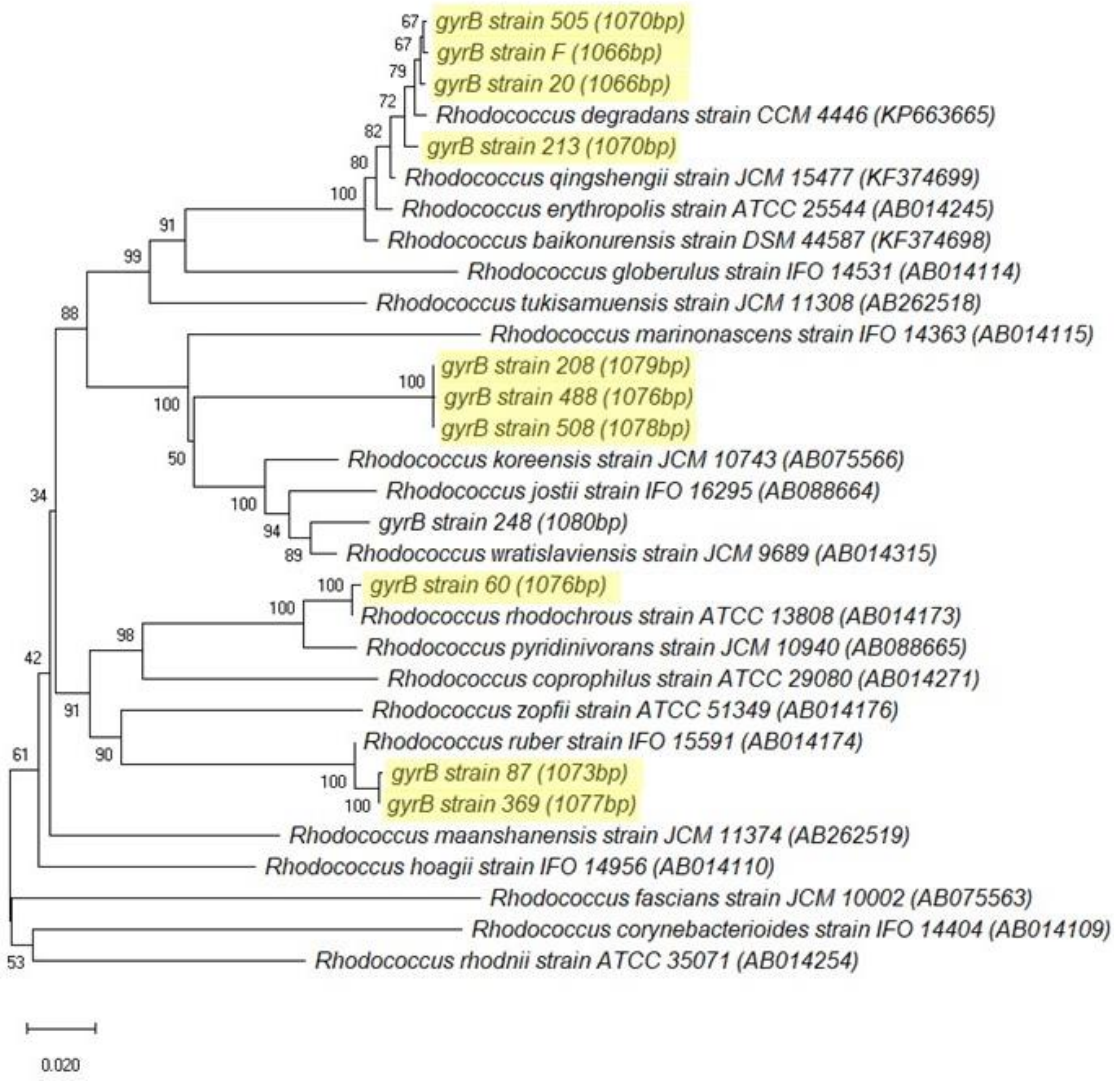


Figure 2.4 NJ-TREE based on *gyrB* gene sequences

The neighbour-joining tree is showing the *gyrB* gene sequence based phylogenetic relationship of the 11 test strains with 20 type strains of genus *Rhodococcus*. Numbers at nodes indicate bootstrap support values based on 1000 resampled dataset. Bar, 2 nt substitutions per 100 nt.

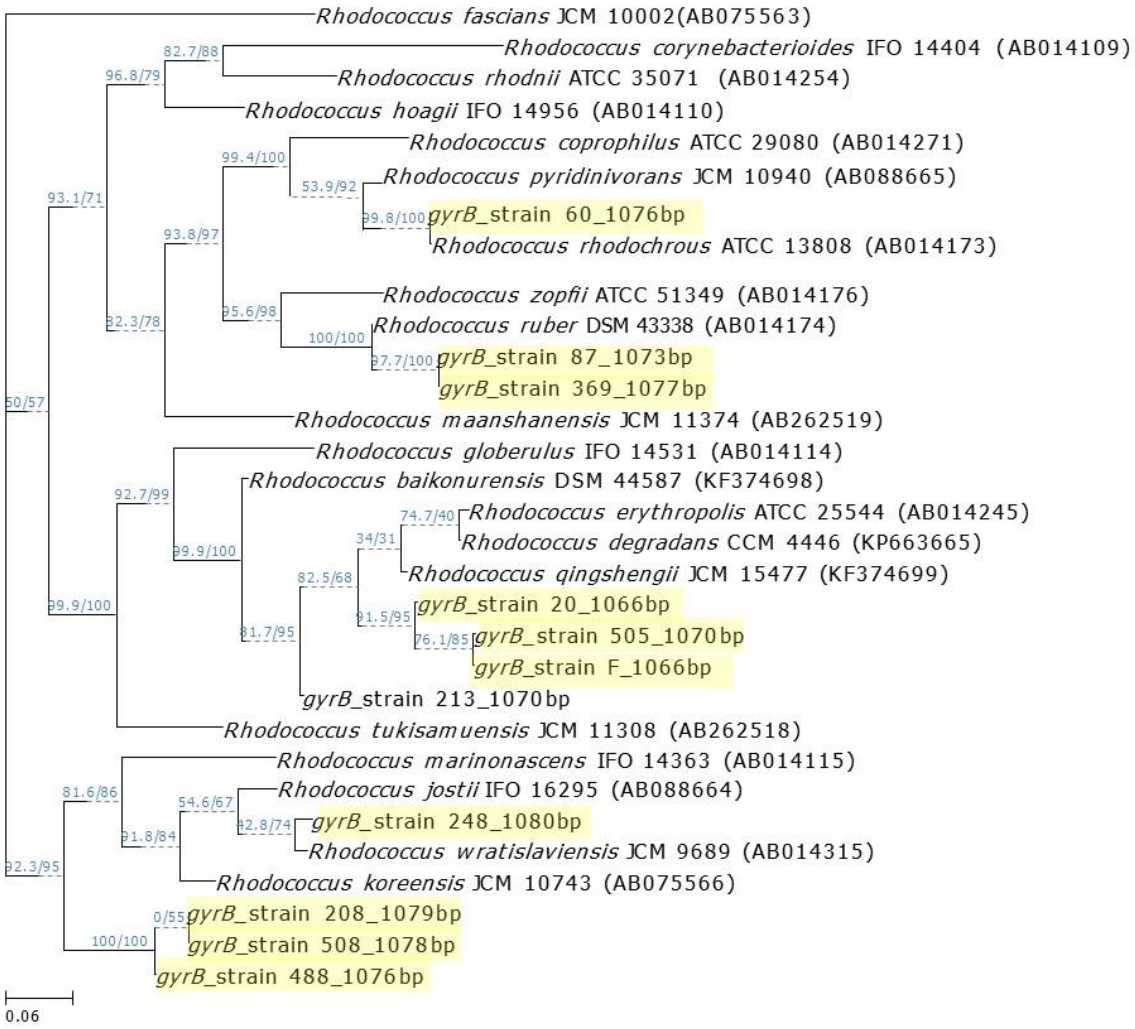


Figure 2.5 IQ-TREE based on *gyrB* gene sequences

The output from the web-based IQ-TREE program showing the *gyrB* gene sequence based phylogenetic relationship of the 11 test strains with 20 type strains of genus *Rhodococcus*. The tree is UNROOTED although outgroup taxon *Rhodococcus fascians* JCM 10002 (AB075563) is drawn at the root. Numbers at the nodes are SH-aLRT support (%) / ultrafast bootstrap support (%). Bar, 6 nt substitutions per 100 nt. The test strains are highlighted in yellow, and the BDS strains among them are highlighted in green.

Phylogenetic analysis of the 20 published *gyrB* gene sequences of the validly described *Rhodococcus* species reveal that the percentage similarity of 16S rRNA gene sequences amongst the type strains ranges from a maximum 99.14% (9/1275 nt differences) shared between *R. qingshengii* JCM 15477 (KF374699) and *R. degradans* CCM 4446 (KP663665). The lowest percentage similarity of 73.15% (290/1275 nt differences) is observed between *R. degradans* CCM 4446 (KP663665) and *R. corynebacterioides* IFO 14404 (AB014109).

There were other general trends observed in the relatedness among the *Rhodococcus* type strains. It was observed that *R. fascians* JCM 10002 (AB075563) and *R. corynebacterioides* IFO 14404 (AB014109) were the most distantly related to other rhodococci type strains in terms of *gyrB* gene sequences. The closest type strains were observed to be *R. erythropolis* ATCC 25544 (AB014245), *R. degradans* CCM 4446 (KP663665), *R. qingshengii* JCM 15477 (KF374699) with at least 98.34% similarity (a maximum of 20/1275 nt differences) among them. Second set of closely related species were *R. rhodochrous* ATCC 13808 (AB014173) and *R. pyridinivorans* JCM 10940 (AB088665) with 97.07% similarity (37/1275 nt differences) between them, while *R. jostii* IFO 16295 (AB088664), *R. koreensis* JCM 10743 (AB075566), *R. wratislaviensis* JCM 9689 (AB014315) formed the third set of closely related cluster with at least 94.3% similarity (a maximum of 72/1275 nt differences) among them.

The closest species assignment for each of the strains based on both the 16S rDNA and *gyrB* gene sequence analyses generated in this study is provided (Table 2.7) and the data supporting each assignment described in more detail in the following sections.

Table 2.7 Confirmed identity of each strain and closest species identity based on phylogenetic analyses of 16S rDNA and *gyrB* gene sequences undertaken in this study.

Strain code	Species assignment based on IEGM catalogue (printed or online versions).	Species assignment based on Bell *et al.* (1999)	Closest identities based on phylogenetic analyses in the current study*	
			16S rRNA gene-based	*gyrB* gene-based
Strain group A**				
208	*Rhodococcus erythropolis*	*Rhodococcus opacus*	*Rhodococcus* sp.	*Rhodococcus* sp.
488	*Rhodococcus erythropolis*	*Rhodococcus opacus*	*Rhodococcus* sp.	*Rhodococcus* sp.
508	*NA*	*Rhodococcus opacus*	*Rhodococcus* sp.	*Rhodococcus* sp.
Strain group B				
20	*Rhodococcus erythropolis*	*Rhodococcus erythropolis*	*Rhodococcus degradans*	*Rhodococcus degradans*
213	*Rhodococcus erythropolis*	*Rhodococcus erythropolis*	*Rhodococcus degradans*	*Rhodococcus degradans*
505	*Rhodococcus erythropolis*	*Rhodococcus erythropolis*	*Rhodococcus degradans*	*Rhodococcus degradans*
F	*Rhodococcus* sp.	*Rhodococcus* sp.	*Rhodococcus degradans*	*Rhodococcus degradans*
Strain group C				
87	*Rhodococcus ruber*	*Rhodococcus ruber*	*Rhodococcus ruber*	*Rhodococcus ruber*
369	*Rhodococcus ruber*	*Rhodococcus ruber*	*Rhodococcus ruber*	*Rhodococcus ruber*
Single strains				
60	*NA*	*Rhodococcus opacus**	*Rhodococcus rhodochrous*	*Rhodococcus rhodochrous*
248	*Rhodococcus erythropolis*	*Rhodococcus opacus*	*Rhodococcus wratislaviensis*	*Rhodococcus wratislaviensis*

NA, Information not available.

*, Species identification based on similarity matrices and phylogenetic position in evolutionary trees generated using MUSCLE alignment distance matrix, and Neighbour-Joining and IQ-TREE tree algorithms.

**Putatively novel based on % nt difference; strains with nt differences greater than 2% from closest validly described species may be novel.

2.3.3 Identification of strains 208, 488 and 508

At the time of receipt, strains 208 and 488 were listed as *R. erythropolis* in the IEGM online catalogue; no information was available for strain 508. These strains were classified as *R. opacus* based on PCR by Bell (1999). These strains formed a distinct clade within the genus *Rhodococcus* in the present study which is closely related to neither *R. erythropolis* nor *R. opacus*. The following section provides details of the 16S rDNA and *gyrB* phylogenetic data upon which the closest species assignments for these strains are based, along with descriptions of the cell and colony characteristics.

Colony, micromorphology and staining properties of strains 208, 488 and 508.

The cell morphology of strains 208, 488 and 508 was observed to be consistent with their assignment to the genus *Rhodococcus*. The cells were aerobic, Gram-positive, weakly acid-fast, non-spore-forming and non-motile. Observation by light microscopy at x1000 magnification under oil immersion revealed the cells were thin and filamentous during the early growth phase (up to 48 h) and most cells appeared as fragmented filaments rod or coccoid shaped when sampled at later stages.

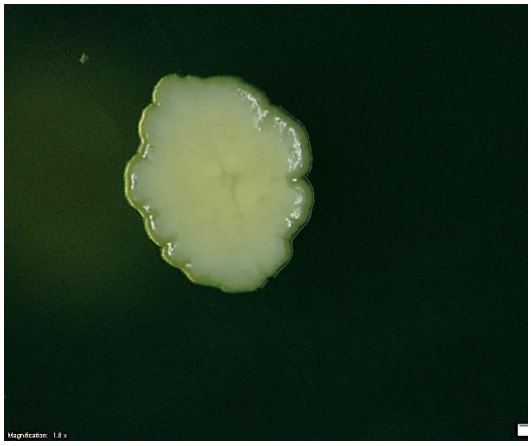
Colonies of strain 208 appeared smooth, whitish-pale grey in colour, irregular in form and were flat with lobate edges when grown on NA plates and umbonate in elevation with irregular edges when grown on GYEA, after 3 days of incubation at 28°C (Figure 2.6 a). Strains 488 formed very similar-looking colonies to strain 208 when grown on NA plates. They appeared dull, with a raised elevation and some folding in the centre of the colony after 3 days at 28°C incubation. But when grown on GYEA, they were no longer similar as the colonies of 488 were often cream to peach in pigmentation with undulated edges and more irregular contoured surface when grown on GYEA (Figure 2.6 d). Strain 508 colonies appeared similar as that of strains 208 and 488 when grown on NA plates with colonies that were whitish-pale grey, irregular in form with lobate edges, raised elevation, smooth surface and dull appearance. These colonies were also whitish-grey in pigment on GYEA plates but were round in form, flat and with irregular edges (Figure 2.6 f).

a) Strain 208 on NA

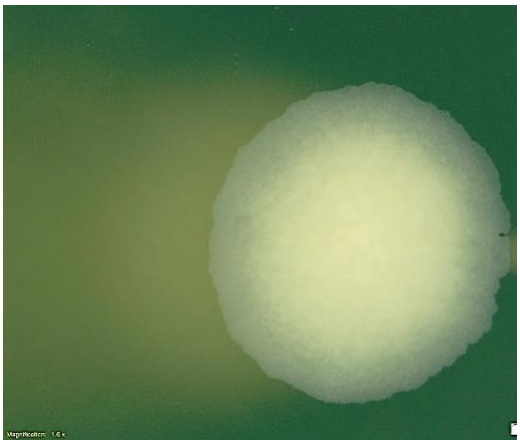
b) Strain 208 on GYEA

c) Strain 488 on NA

d) Strain 488 on GYEA



e) Strain 508 on NA



f) Strain 508 on GYEA

Figure 2.6 Colonies of strains 208, 488 and 508

Colonies of strains 208 (a & b), 488 (c & d) and 508 (e & f) growing as macroscopically visible colonies on nutrient agar (NA) and glucose yeast-extract agar (GYEA) plates, respectively after incubation for 3 days at 28ºC. Bar [500 µm].

Phylogenetic identification of strains 208, 488 and 508 based on the 16S rRNA gene and gyrB trees.

In the current study, the 16S rRNA gene sequences of these strains clustered together in both the NJ (Figure 2.2) and IQ-TREE/maximum likelihood (IQ/ML) phylogenetic trees (Figure 2.3). In both trees, these strains formed a distinct clade that did not include any valid *Rhodococcus* species with 99% bootstrap values. The 16S rRNA gene nucleotide similarity values and the number of nucleotide differences between the three strains and the valid *Rhodococcus* species recovered in these clades (NJ and IQ trees) are shown in Table 2.8. The three strains assigned to this clade share a high 16S rRNA gene sequence similarity of 99.66% which corresponds to 5/1520 nucleotide (nt) difference. The most closely related valid species to this clade, *R. koreensis* DNP505 (AF124342), shared the highest similarity (98.91%;16/1520 nt differences) with strain 208 and 508, and (98.76%;18/1520 nt differences) with 488. *R. jostii* IFO 16295 (AB046357), the next most similar valid species to members of this clade, shares the highest similarity of 98.37% with strain 508 in this clade.

The *gyrB* gene sequences of these strains also formed a distinct cluster with 100% bootstrap values, as seen in Figure 2.4 and Figure 2.5. From Table 2.9, it could be noted that the closest validly described species was *R. wratislaviensis* JCM 9689 (AB014315) with more than 89.24% (116/1078 nt differences) similarity with these strains, closely followed by *R. koreensis* JCM 10743 (AB075566) with more than 89.06% (118/1078 nt differences), although *R. jostii* IFO 16295 (AB088664) showed a high similarity of more than 88.68% corresponding to 120/1078 nt differences.

These results show that strains 208, 488 are neither *R. erythropolis* nor *R. opacus* as previously assigned. The identity of the strains 208, 488 and 508 based on 16S rDNA and *gyrB* gene sequence could be confirmed only to the genus level and could potentially be novel.

Table 2.8 Distance matrix for the phylogenetic analysis of the 16S rRNA sequence of strains 208, 488 and 508

	R. tukisamuensis Mb8 (AB067734)	*R. maanshanensis* M712 (AF416566)	*R. nanhaiensis* SCSIO 10187 (JN582175)	*R. globerulus* DSM 43954 (X80619)	*R. erythropolis* DSM 43066 (X79289)	*R. degradans* CCM 4446 (JQ776649)	*R. qingshengii* djl-6 (DQ090961)	*R. baikonurensis* A1-22 (AB071951)	*R. imtechensis* RKJ300 (AY525785)	*R. opacus* DSM 43205 (X80630)	*R. wratislaviensis* DSM 44107 (Z37138)	*R. marinonascens* 3438W (X80617)	*R. jostii* IFO 16295 (AB046357)	*R. koreensis* DNP505 (AF124342)	**208 16S rRNA gene (1492bp)**	**508 16S rRNA gene (1495bp)**	**488 16S rRNA gene (1460bp)**
R. tukisamuensis Mb8 (AB067734)		69	88	84	48	86	86	88	85	77	76	76	83	77	75	76	76
R. maanshanensis M712 (AF416566)	94.85		40	46	34	40	40	40	41	37	37	31	37	33	29	29	31
R. nanhaiensis SCSIO 10187 (JN582175)	93.34	97.13		45	44	38	38	40	36	30	26	31	48	39	37	37	37
R. globerulus DSM 43954 (X80619)	93.83	96.75	96.79		27	16	21	17	41	39	37	37	38	26	27	27	29
R. erythropolis DSM 43066 (X79289)	96.42	97.46	96.65	98.02		11	11	14	41	42	38	30	43	34	30	30	32
R. degradans CCM 4446 (JQ776649)	93.65	97.17	97.29	98.91	99.24		5	3	43	43	39	33	44	32	31	31	33
R. qingshengii djl-6 (DQ090961)	93.73	97.17	97.29	98.57	99.24	99.66		3	48	48	44	38	50	37	37	38	38
R. baikonurensis A1-22 (AB071951)	93.47	97.01	96.97	98.74	99.01	99.78	99.78		45	45	41	34	45	34	33	33	35
R. imtechensis RKJ300 (AY525785)	93.8	97.1	97.43	97.22	96.96	97.08	96.75	96.66		17	10	34	39	27	30	30	29
R. opacus DSM 43205 (X80630)	94.38	97.4	97.88	97.37	96.9	97.09	96.76	96.68	98.87		7	24	31	25	23	24	24
R. wratislaviensis DSM 44107 (Z37138)	94.39	97.38	98.14	97.48	97.19	97.35	96.99	96.96	99.32	99.54		24	31	23	20	20	22
R. marinonascens 3438W (X80617)	94.45	97.82	97.8	97.5	97.81	97.76	97.43	97.5	97.71	98.41	98.38		32	23	19	20	20
R. jostii IFO 16295 (AB046357)	93.98	97.41	96.6	97.44	96.83	97.04	96.65	96.69	97.44	97.96	97.92	97.89		26	26	25	27
R. koreensis DNP505 (AF124342)	94.31	97.66	97.21	98.23	97.49	97.83	97.47	97.48	98.17	98.32	98.44	98.45	98.26		16	16	18
208 16S rRNA gene (1492bp)	94.49	97.95	97.36	98.17	97.79	97.9	97.48	97.55	97.99	98.46	98.64	98.73	98.3	98.91		5	5
508 16S rRNA gene (1495bp)	94.43	97.95	97.36	98.17	97.79	97.9	97.42	97.55	97.99	98.4	98.64	98.66	98.37	98.91	99.66		5
488 16S rRNA gene (1460bp)	94.29	97.79	97.36	97.99	97.6	97.72	97.36	97.36	98.02	98.36	98.48	98.63	98.17	98.76	99.66	99.66	

Percentage similarity

Nucleotide difference

The table shows the pairwise similarity values (%) and the number of nucleotide differences found between the 16S rRNA gene sequences of strains 208, 488 and 508, and closely related type strains of validly described *Rhodococcus* species recovered in the same clade on the NJ-TREE and IQ-TREE. Darker shades indicate high similarity and low number of nucleotide differences between the sequence pair.

Table 2.9 Distance matrix for the phylogenetic analysis of the *gyrB* sequence of strains 208, 488, 508

	R. tukisamuensis JCM 11308 (AB262518)	*R. marinonascens* IFO 14363 (AB014115)	***gyrB*_208 (1079bp)**	***gyrB*_508 (1078bp)**	***gyrB*_488 (1076bp)**	*R. jostii* IFO 16295 (AB088664)	*R. koreensis* JCM 10743 (AB075566)	*R. wratislaviensis* JCM 9689 (AB014315)
R. tukisamuensis JCM 11308 (AB262518)		249	211	211	212	222	206	207
R. marinonascens IFO 14363 (AB014115)	80.52		156	155	156	161	155	142
***gyrB*_208 (1079bp)**	80.73	85.56		0	2	121	119	115
***gyrB*_508 (1078bp)**	80.71	85.63	100		2	120	118	114
***gyrB*_488 (1076bp)**	80.6	85.53	99.81	99.81		122	120	116
R. jostii IFO 16295 (AB088664)	82.63	87.25	88.8	88.88	88.68		72	54
R. koreensis JCM 10743 (AB075566)	83.88	87.73	88.98	89.06	88.87	94.3		52
R. wratislaviensis JCM 9689 (AB014315)	83.8	88.76	89.35	89.43	89.24	95.72	95.88	

Percentage similarity

Nucleotide difference

The table shows the pairwise similarity values (%) and the number of nucleotide differences found between the *gyrb* gene sequences of strains 208, 488 and 508, and closely related type strains of validly described *Rhodococcus* species recovered in the same clade on the NJ-TREE and IQ-TREE. Darker shades indicate high similarity and low number of nucleotide differences between the sequence pair.

2.3.4 Identification of strains 20, 213, 505 and F

The tests strains 20, 213, 505 and F, that were previously classified as members of the validly described species *Rhodococcus erythropolis* (Gray and Thornton 1928) in the IEGM catalogue (online/printed) and confirmed as the same by Bell *et al.* (1999), were found in the present study to be most highly related to another species *R. degradans*. The following section provides details of the 16S rRNA gene and *gyrB* gene sequence based phylogenetic data upon which the closest species assignment for these strains is based, along with descriptions of the cell and colony characteristics of each strain.

Colony, micromorphology and staining properties of strains 20, 213, 505 and F.

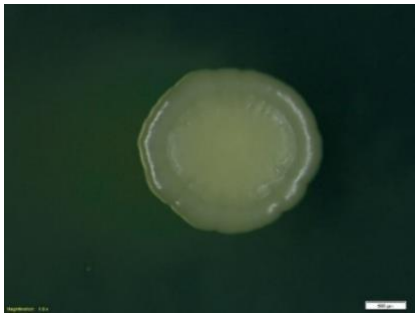
The cell morphologies of strains 20, 213, 505 and F were observed to be consistent with their assignment to the genus *Rhodococcus*. The cells were aerobic, Gram-positive, weak to non-acid fast, non-spore-forming and non-motile. Observation by light microscopy revealed a typical rod–coccus developmental cycle. Cells were rods which then fragmented into short rods or cocci (48–60 h) into the stationary phase (72 h) which is a typical rod–coccus cycle of the members of the genus *Rhodococcus.*

Strain 20 formed smooth, pale-white, convex, and opaque colonies with undulated edges on NA plate, while similar-looking colonies but with a pale-orange colouration were observed on GYEA plates after 3 days of incubation at 28°C (Figure 2.7 a & b).

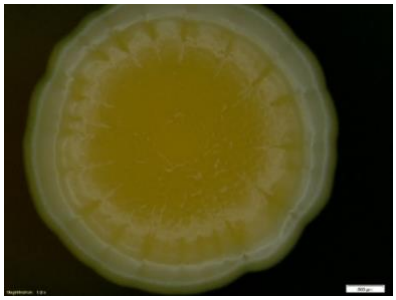
Strain 213 formed smooth and shiny, raised irregularly shaped colonies with undulated margins on an NA plate, and on GYEA plates strain 213 formed smooth but dull surfaced, round-shaped colonies with umbonate elevation and undulated margin (Figure 2.7 c & d).

Strain 505 formed smooth, circular, raised, pale white colonies on NA plate (Figure 2.7 e). When grown on GYEA, strain 505 formed pale-beige coloured colonies with a smooth surface, a round-entire margin and crateriform elevation after 3 days of incubation at 28°C (Figure 2.7 f). Strain F grew as smooth and shiny colonies with round-entire margins, convex, sticky and

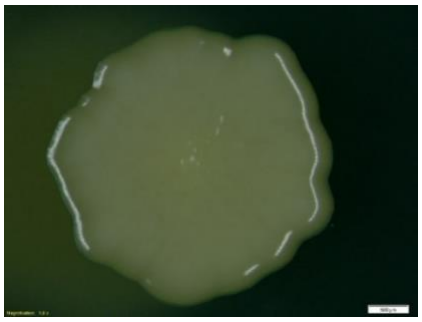
opaque colonies with regular edges when grown on NA and GYEA plates. The colouration was pale white on NA plates, while cream-coloured on GYEA plates. Among all the test strains that were used in this study, strain F exhibited uniquely mucoidal sticky colonies (Figure 2.7 g & h).



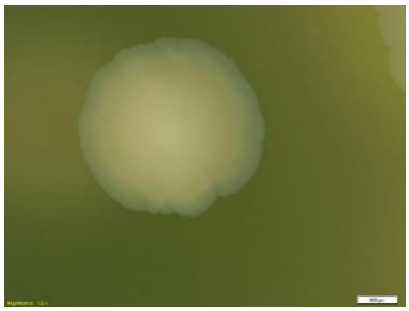
a) Strain 20 growing in NA



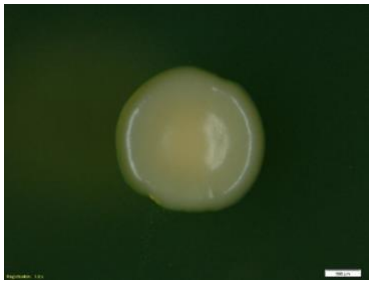
b) Strain 20 on GYEA



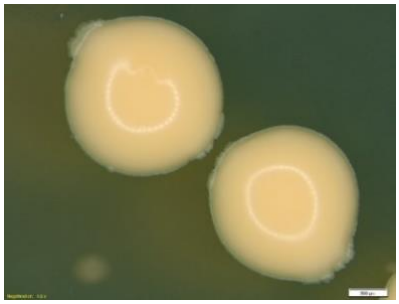
c) Strain 213 on NA



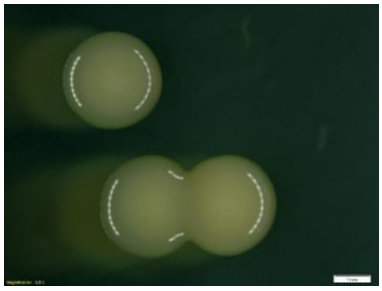
d) Strain 213 on GYEA



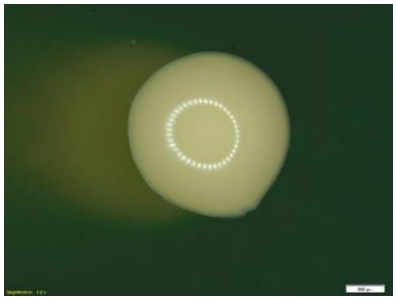
e) Strain 505 on NA



f) Strain 505 on GYEA



g) Strain F on NA



h) Strain F on GYEA

Figure 2.7 Colonies of strains 20, 213, 505 and F

Strains 20 (a & b), 213 (c & d), 505 (e & f) and F (g,h) growing as macroscopically visible colonies on nutrient agar (NA) and glucose yeast-extract agar (GYEA) plates, respectively after incubation for 3 days at 28ºC.

Phylogenetic identification of strains 20, 213, 505 and F based on the 16S rDNA and gyrB trees.

The 16S rRNA gene sequences of these four strains formed a tight cluster together with four validly described *Rhodococcus* species in both the NJ-TREE (Figure 2.2) and IQ-TREE (Figure 2.3). The 16S rRNA gene sequences of the four strains shared between 99.53% (strains 20 and F) and 99.87% (strains 505 and F) similarity and average similarity of 99.7% with each other, and differed by between 2 to 7 nucleotides (Table 2.10). All four strains shared identical 16S rRNA gene sequences (100% sequence homology) with the type strain of *R. degradans* CCM 4446^T (1,473 bp). *R. baikonurensis* A1-22^T shared the next highest similarity of 99.78% (3/1348 nucleotide differences) with the four strains in this clade (Table 2.10).

The NJ and IQ TREE (Figure 2.4 and Figure 2.5) show tight clustering among 20, 505, and F with more than 99.44% similarity among each other. Strain 213 formed a loose member of this cluster with 98.78% similarity with strain 20, 505 and F. The *gyrB* of *R. degradans* CCM 4446^T was the closest (Table 2.11). Based on the similarities observed in the 16S rRNA gene and *gyrB* gene sequence analysis, it could be concluded that the strains 20, 213, and 505 should be classified as members of *R. degradans* rather than their original classification as *R. erythropolis,* along with the unknown rhodococcal strain F. During this analysis, it was also observed that the *gyrB* sequence of *R. erythropolis* ATCC 25544^T (AB014245), *R. qingshengii* JCM 15477^T (KF374699), and *R. degradans* CCM 4446^T (KP663665) shared an average of 98.4% similarity among themselves, which indicated the close relationship of the type strains.

Table 2.10 Distance matrix for the phylogenetic analysis of the 16S rRNA sequence of strains 20, 213, 505 and F

	R. koreensis DNP505 (AF124342)	*R. globerulus* DSM 43954 (X80619)	*R. erythropolis* DSM 43066 (X79289)	*R. qingshengii* djl-6 (DQ090961)	**20 16S rRNA gene (1494bp)**	**213 16S rRNA gene (1493bp)**	**505 16S rRNA gene (1495bp)**	**F 16S rRNA gene (1495bp)**	*R. degradans* CCM 4446 (JQ776649)	*R. baikonurensis* A1-22 (AB071951)
R. koreensis DNP505 (AF124342)		26	34	38	32	32	32	32	32	34
R. globerulus DSM 43954 (X80619)	98.23		27	22	16	16	16	16	16	17
R. erythropolis DSM 43066 (X79289)	97.49	98.02		11	11	11	11	11	11	14
R. qingshengii djl-6 (DQ090961)	97.41	98.51	99.24		10	10	9	7	6	3
20 16S rRNA gene (1494bp)	97.83	98.91	99.24	99.32		5	6	7	0	3
213 16S rRNA gene (1493bp)	97.83	98.91	99.24	99.32	99.67		4	3	0	3
505 16S rRNA gene (1495bp)	97.83	98.91	99.24	99.39	99.6	99.73		2	0	3
F 16S rRNA gene (1495bp)	97.83	98.91	99.24	99.53	99.53	99.8	99.87		0	3
R. degradans CCM 4446 (JQ776649)	97.83	98.91	99.24	99.59	100	100	100	100		3
R. baikonurensis A1-22 (AB071951)	97.48	98.74	99.01	99.78	99.78	99.78	99.78	99.78	99.78	

Percentage similarity

Nucleotide difference

The table shows the pairwise similarity values (%) and the number of nucleotide differences found between the 16S rRNA gene sequences of strains 20, 213, 505 and F, and closely related type strains of validly described *Rhodococcus* species recovered in the same clade on the NJ-TREE and IQ-TREE. Darker shades indicate high similarity and low number of nucleotide differences between the sequence pair.

Table 2.11 Distance matrix for the phylogenetic analysis of the *gyrB* sequence of strains 20, 213, 505 and F

	R. globerulus IFO 14531 (AB014114)	*R. tukisamuensis* JCM 11308 (AB262518)	*R. baikonurensis* DSM 44587 (KF374698)	*R. erythropolis* ATCC 25544 (AB014245)	***gyrB*_213 (1070bp)**	*R. qingshengii* JCM 15477 (KF374699)	*R. degradans* CCM 4446 (KP663665)	***gyrB*_20 (1066bp)**	***gyrB*_505 (1070bp)**	***gyrB_F (1066bp)***
R. globerulus IFO 14531 (AB014114)		198	160	169	140	162	141	145	144	145
R. tukisamuensis JCM 11308 (AB262518)	84.05		142	157	137	149	138	139	140	139
R. baikonurensis DSM 44587 (KF374698)	86.8	88.27		18	14	20	13	18	20	18
R. erythropolis ATCC 25544 (AB014245)	86.33	87.53	98.51		20	17	12	19	19	21
***gyrB*_213 (1070bp)**	86.99	87.27	98.67	98.13		15	11	13	13	13
R. qingshengii JCM 15477 (KF374699)	86.63	87.7	98.34	98.59	98.57		9	10	13	12
R. degradans CCM 4446 (KP663665)	86.76	87.03	98.75	98.87	98.96	99.14		10	8	7
***gyrB*_20 (1066bp)**	86.5	87.05	98.29	98.22	98.78	99.05	99.06		6	4
***gyrB*_505 (1070bp)**	86.62	86.99	98.1	98.23	98.79	98.76	99.24	99.44		4
gyrB_F (1066bp)	86.5	87.05	98.29	98.03	98.78	98.86	99.34	99.62	99.63	

Percentage similarity

Nucleotide difference

The table shows the pairwise similarity values (%) and the number of nucleotide differences found between the *gyrb* gene sequences of strains 20, 213, 505 and F, and closely related type strains of validly described *Rhodococcus* species recovered in the same clade on the NJ-TREE and IQ-TREE. Darker shades indicate high similarity and low number of nucleotide differences between the sequence pair.

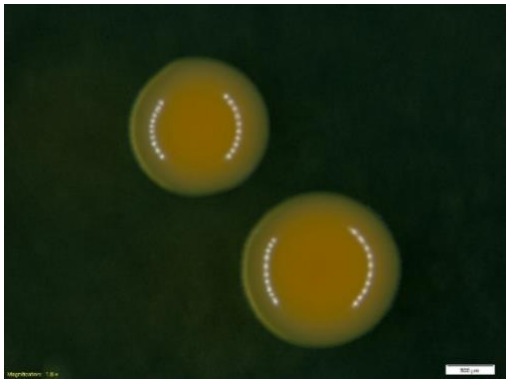
2.3.5 Identification of strains 87 and 369

Strains 87 and 369 were classified as members of the species *Rhodococcus ruber* in the IEGM catalogue. The following section provides details of the 16S rDNA and *gyrB* phylogenetic data upon which the closest species assignments for these strains are based, along with descriptions of the cell and colony characteristics.

Colony, micromorphology and staining properties of strains 87 and 369

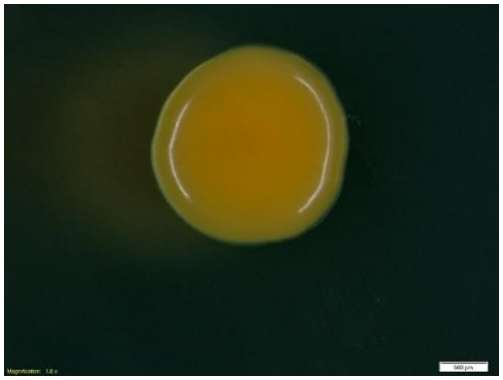
The cell morphologies of strains 87 and 369 were observed to be consistent with their assignment to the genus *Rhodococcus*. The cells were aerobic, Gram-positive, weak to non-acid fast, non-spore-forming and non-motile. Observation by light microscopy revealed a typical short rod–coccus developmental cycle after 3 days of growth.

Both strains 87 and 369 exhibited highly similar colony characteristics when grown on NA and GYEA plates. On NA plates, they appeared as round colonies with a smooth and shiny surface, a convex elevation and reddish-orange in colour. On GYEA plates, they grew as irregular shaped colonies with undulated margin and a crateriform surface. The colouration ranged from a yellowish central region to pale-yellowish towards the edge, as shown in Figure 2.8.



(a) Strain 87 on NA

(b) Strain 87 on GYEA



(b) Strain 369 on NA

(d) Strain 369 on GYEA

Figure 2.8 Colonies of strains 87 and 369
Strains 87 (a & b) and 369 (c & d) growing as macroscopically visible colonies on nutrient agar (NA) and glucose yeast-extract agar (GYEA) plates, respectively after incubation for 3 days at 28ºC.

Phylogenetic identification of strains 87 and 369 based on the 16S rDNA and gyrB trees

The 16S rRNA gene sequence of strains 87 and 369 clustered together with a bootstrap value of 97/100 in NJ and IQ/ML trees, respectively and formed a tight cluster with two validly described *Rhodococcus* species in NJ (Figure 2.2) and IQ-TREE (Figure 2.3). The 16S rRNA gene sequences of strains 87 and 369 shared 99.87% similarity (2/1501 nucleotide differences), and both of them share 99.73% similarity with *R. ruber* DSM 43338^T (X80625) which corresponds to 4/1481 nucleotide differences also occurring at the ends of the type strain. They cluster together with *R. aetherivorans,* which shared high 99.42% similarity (8/1409 nucleotide difference between 1096 – 1103 positions) with *R. ruber,* as shown in Table 2.12. A similar trend was observed in the case of *gyrB* gene sequences (Table 2.13), where the strains 87 and 369 shared 99.91% similarity (1/1073 nucleotide difference) which occur at the

872nd position on the gene sequence and clustered together with *R. ruber* in the *gyrB* phylogenetic tree (Figure 2.4 and Figure 2.5). Again, *R. ruber* DSM 43338 (X80625) shares 99.44% and 99.54% similarity with strains 87 and 369, respectively, which corresponds to 6/1266 and 5/1266 nucleotide differences, respectively. The *gyrB* gene sequence of the type strain of *R. aetherivorans* was not available to further support the findings of the 16S rRNA gene sequence analysis. Therefore, *R. zopfii* ATCC 51349 (AB014176) was observed as the second closest relative, which showed an average of 86.8% similarity with strains 87 and 369. The findings prove that the original classification of strains 87 and 369 as *R. ruber* is correct and it has been confirmed through 16S rRNA and *gyrB* phylogeny.

Table 2.12 Distance matrix for the phylogenetic analysis of the 16SrRNA sequence of strains 87 and 369

	R. zopfii T1 (AF191343)	*R. aetherivorans* 10bc312 (AF447391)	**369 16S rRNA gene (1501bp)**	**87 16S rRNA gene (1502bp)**	*R. ruber* DSM 43338 (X80625)
R. zopfii T1 (AF191343)		34	30	30	30
R. aetherivorans 10bc312 (AF447391)	97.56		11	10	8
369 16S rRNA gene (1501bp)	97.97	99.22		2	4
87 16S rRNA gene (1502bp)	97.97	99.29	99.87		4
R. ruber DSM 43338 (X80625)	97.96	99.42	99.73	99.73	

Nucleotide difference

Percentage similarity

The table shows the pairwise similarity values (%) and the number of nucleotide differences found between the 16S rRNA gene sequences of strains 87 and 369, and closely related type strains of validly described *Rhodococcus* species recovered in the same clade on the NJ-TREE and IQ-TREE. Darker shades indicate high similarity and low number of nucleotide differences between the sequence pair.

Table 2.13 Distance matrix for the phylogenetic analysis of the *gyrB* sequence of strains 87 and 369

	R. coprophilus ATCC 29080 (AB014271)	*R. pyridinivorans* JCM 10940 (AB088665)	*R. rhodochrous* ATCC 13808 (AB014173)	*R. hoagii* IFO 14956 (AB014110)	*R. maanshanensis* JCM 11374 (AB262519)	*R. zopfii* ATCC 51349 (AB014176)	*R. ruber* IFO 15591 (AB014174)	***gyrB*_87 (1073bp)**	***gyrB*_369 (1077bp)**
R. coprophilus ATCC 29080 (AB014271)		152	154	182	217	191	187	165	164
R. pyridinivorans JCM 10940 (AB088665)	87.97		37	190	204	170	179	164	163
R. rhodochrous ATCC 13808 (AB014173)	87.81	97.07		185	199	169	173	156	156
R. hoagii IFO 14956 (AB014110)	85.59	84.96	85.35		182	181	165	156	155
R. maanshanensis JCM 11374 (AB262519)	82.82	83.85	84.24	85.56		206	200	189	188
R. zopfii ATCC 51349 (AB014176)	84.88	86.54	86.62	85.67	83.69		156	143	142
R. ruber DSM 43338 (AB014174)	85.19	85.83	86.3	86.9	84.13	87.65		6	5
***gyrB*_87 (1073bp)**	84.67	84.76	85.5	85.5	82.39	86.71	99.44		1
***gyrB*_369 (1077bp)**	84.81	84.91	85.56	85.65	82.54	86.85	99.54	99.91	

Percentage similarity — **Nucleotide difference**

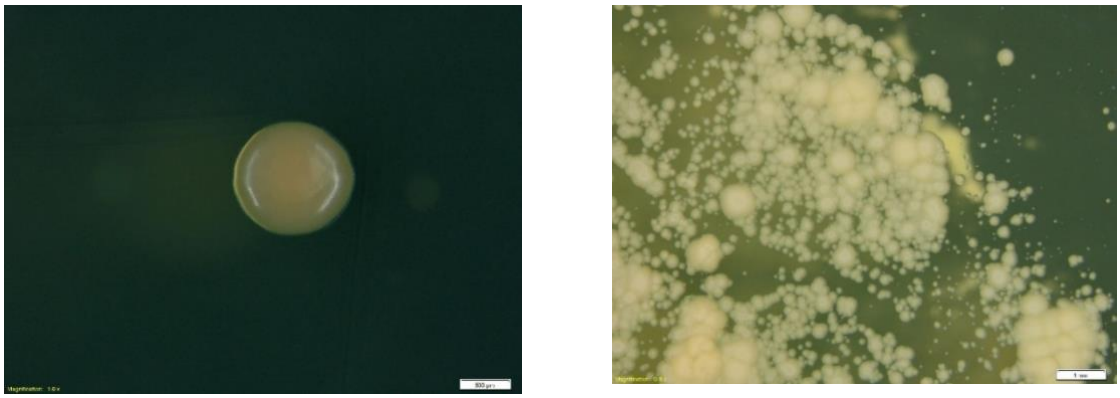
The table shows the pairwise similarity values (%) and the number of nucleotide differences found between the *gyrb* gene sequences of strains 87 and 369, and closely related type strains of validly described *Rhodococcus* species recovered in the same clade on the NJ-TREE and IQ-TREE. Darker shades indicate high similarity and low number of nucleotide differences between the sequence pair.

2.3.6 Species identification of strain 60

The identity of the test strain 60 was not precisely known. Whilst it is listed as a member of the genus *Rhodococcus* in the IEGM catalogue, its species classification had not been delineated. Moreover, Bell *et al.* (1999) had described that their species-specific primers for *R. opacus* gave a weak positive signal for this strain. In the present study, strain 60 was shown to be only distantly related to this species and instead fell within a distinct clade containing several other *Rhodococcus* species. The following section provides details of the 16S rDNA and *gyrB* phylogenetic data upon which the closest species assignment(s) for this strain is/are based, along with descriptions of the cell and colony characteristics.

Colony, micromorphology and staining properties of strain 60

The cell morphology of strain 60 was consistent with its assignment to the genus *Rhodococcus*. The cells were aerobic, Gram-positive, weakly acid-fast, non-spore-forming and non-motile. Under the microscope, the cells appeared as cocci occurring as clusters and occasionally arranged as filaments. On NA plate, the colonies of strain 60 appeared as circular colonies with round-entire margin, having a raised elevation and pale-orange coloured after 3 days of incubation at 28°C. The colonies were smaller and convex when cultured in GYEA Figure 2.9.



(a) Strain 60 growing on NA **(b) Strain 60 growing on GYEA**

Figure 2.9 Colonies of strain 60
Strain 60 growing as macroscopically visible colonies on **(a)** nutrient agar (NA) and **(b)** glucose yeast-extract agar (GYEA) plates, respectively after incubation for 3 days at 28ºC.

Phylogenetic identification of strain 60 based on the 16S rDNA and gyrB trees.

In this research, the 16S rRNA gene of this strain exhibited a 99.59% similarity to *Rhodococcus rhodochrous* DSM 43241 (X79288). This corresponded to 6/1474 nucleotide differences – 4 of which occurred at the ends. They occurred as a tight cluster with a bootstrap value of 81/91 in NJ (Figure 2.2) and IQ-TREE (Figure 2.3), respectively. The IQ tree grouped the above strains under *R. gordoniae* W4937 (AY233201) subclade that included *R. lactis* DW151B (KP342300), *R. biphenylivorans* TG9 (KJ546454) and *R. pyridinivorans* PDB9 (AF173005) which had 97.97%, 98.6%, 98.85% similarity with strain 60, respectively. It is clear from the phylogenetic trees and by the nucleotide similarity values (Table 2.14) that strain 60 can be distinguished from representatives of all the validly described *Rhodococcus* species.

The NJ and IQ/ML trees obtained using the *gyrB* gene sequence of strain 60 (Figure 2.4 and Figure 2.5) showed that it clustered with that of *R. rhodochrous* ATCC 13808 (AB014173) with 100/100 bootstrap values. The latter strain shares 97.07% similarity with *R. pyridinivorans* JCM 10940 (AB088665), and it could be noted that *R. pyridinivorans* is the second closest to strain 60 with 85.06% similarity Table 2.15. From these results, it could be concluded that strain 60 should be classified as a member of *R. rhodochrous,* thus resolving its unclear taxonomic identification.

Table 2.14 Distance matrix for the phylogenetic analysis of the 16S rRNA sequence of strain 60

	R. artemisiae YIM 65754 (LN867321)	*R. gordoniae* W4937 (AY233201)	*R. rhodochrous* DSM 43241 (X79288)	**60 16S rRNA gene (1499bp)**	*R. lactis* DW151B (KP342300)	*R. biphenylivorans* TG9 (KJ546454)	*R. pyridinivorans* PDB9 (AF173005)
R. artemisiae YIM 65754 (LN867321)		32	36	43	45	30	26
R. gordoniae W4937 (AY233201)	97.72		21	23	23	11	13
R. rhodochrous DSM 43241 (X79288)	97.56	98.49		6	27	11	15
60 16S rRNA gene (1499bp)	97.13	98.36	99.59		29	21	17
R. lactis DW151B (KP342300)	96.86	98.31	98.11	97.97		17	24
R. biphenylivorans TG9 (KJ546454)	98.02	99.22	99.25	98.6	98.81		8
R. pyridinivorans PDB9 (AF173005)	98.24	99.07	98.98	98.85	98.32	99.46	

Nucleotide difference

Percentage similarity

The table shows the pairwise similarity values (%) and the number of nucleotide differences found between the 16S rRNA gene sequences of strain 60 and closely related type strains of validly described *Rhodococcus* species recovered in the same clade on the NJ-TREE and IQ-TREE. Darker shades indicate high similarity and low number of nucleotide differences between the sequence pair.

Table 2.15 Distance matrix for the phylogenetic analysis of the *gyrB* sequence of strain 60

	R. ruber IFO 15591 (AB014174)	*R. zopfii* ATCC 51349 (AB014176)	*R. coprophilus* ATCC 29080 (AB014271)	*gyrB*_639 (1079bp)	*R. pyridinivorans* JCM 10940 (AB088665)	***gyrB*_60 (1076bp)**	*R. rhodochrous* ATCC 13808 (AB014173)
R. ruber IFO 15591 (AB014174)		159	197	315	177	301	173
R. zopfii ATCC 51349 (AB014176)	87.43		192	312	169	298	166
R. coprophilus ATCC 29080 (AB014271)	84.44	84.83		282	152	283	154
*gyrB*_639 (1079bp)	74.39	74.63	77.02		213	61	208
R. pyridinivorans JCM 10940 (AB088665)	86.02	86.65	87.97	82.64		183	37
***gyrB*_60 (1076bp)**	75.49	75.73	76.9	94.34	85.06		150
R. rhodochrous ATCC 13808 (AB014173)	86.33	86.89	87.81	83.05	97.07	87.76	

Percentage similarity

Nucleotide difference

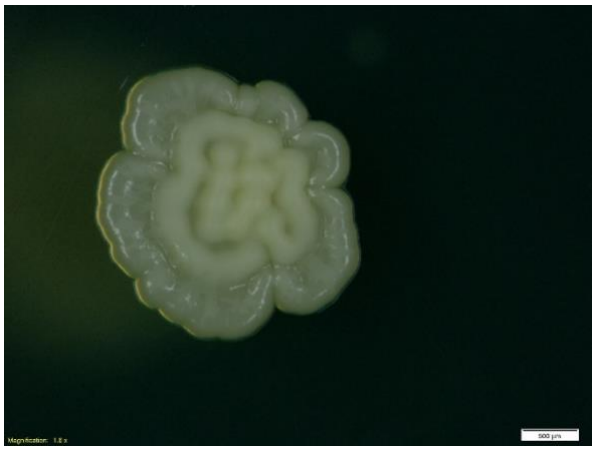
The table shows the pairwise similarity values (%) and the number of nucleotide differences found between the *gyrb* gene sequences of strain 60 and closely related type strains of validly described *Rhodococcus* species recovered in the same clade on the NJ-TREE and IQ-TREE. Darker shades indicate high similarity and low number of nucleotide differences between the sequence pair.

2.3.7 Identification of strain 248

Strain 248, previously classified as *R. erythropolis* in the printed version of IEGM catalogue, was identified as a member of the validly described species *Rhodococcus opacus* by Bell *et al.* (1999). It is currently listed as *R. opacus* in the IEGM online catalogue, but in the present study, it was shown to be highly related to both *R. opacus* and *R. wratislaviensis* and more distantly related to *R. erythropolis*. The following section provides details of the 16S rRNA gene and *gyrB* phylogenetic data upon which the closest species assignment for this strain is based, along with descriptions of the cell and colony characteristics.

Colony, micromorphology and staining properties of strain 248.

The cells of strain 248 were aerobic, weakly Gram-positive, non-acid-fast, non-spore-forming and non-motile. Observation by light microscopy revealed a typical rod–coccus like arrangement of the cells. After 3 days of incubation at 28°C, the GYEA and NA plates gave rise to irregularly shaped colonies with highly undulated margins. The colonies on NA plates were pale-white coloured with contoured surface elevation, whereas on GYEA plates the colonies were pale-beige in colour with a characteristic crateriform elevation in the centre, and contoured towards the edge as shown in Figure 2.10.



(a) Strain 248 on NA

(b) Strain 248 on GYEA

Figure 2.10 Colonies of strain 248
The colony of strain 248 growing as macroscopically visible colonies on **(a)** nutrient agar (NA) and **(b)** glucose yeast-extract agar (GYEA) plates, respectively after incubation for 3 days at 28ºC

Phylogenetic identification of strain 248 based on the 16S rDNA and gyrB trees.

The NJ Tree (Figure 2.2) and IQ-TREE (Figure 2.3) obtained using the 16S rRNA gene sequences show that strain 248 forms a distinct position that included *R. opacus* DSM 43205 (X80630), *R. percolatus* MBS1 (X92114), *R. imtechensis* RKJ300 (AY525785) and *R. wratislaviensis* DSM 44107 (Z37138) under its subclade (bootstrap value 52). The type strain of *R. wratislaviensis* N805^T (DSM 44107) also shared high similarity with the type strain of *R. imtechensis* that was validly published later (Ghosh *et al.*, 2006).

From the distance matrix of the 16S rRNA gene sequence (Table 2.16), it could be noted that strain 248 shared 99.73% with *R. wratislaviensis* DSM 44107 (Z37138) which is the closest validly described strain, followed by *R. opacus* DSM 43205 (X80630) with 99.14% similarity and *R. imtechensis* RKJ300 (AY525785) with 98.53% similarity in order. These percentage similarity values correspond to 4/1474, 13/1481 and 22/1495 nucleotide differences, respectively. The close relationship of strain 248 with *R. wratislaviensis* JCM 9689 (AB014315) is also supported by the high 97.41% similarity between their *gyrB* gene sequences (Table 2.17) and the tight clustering observed between 248 and *R. wratislaviensis* in the NJ and IQ-TREE (Figure 2.4 and Figure 2.5). Based on the similarities observed in the 16S rRNA gene and *gyrB* sequence analysis, it could be concluded that strain 248 should be classified as a member of *R. wratislaviensis* rather than an *R. opacus* strain.

Table 2.16 Distance matrix for the phylogenetic analysis of the 16S rRNA sequence of strain 248

	R. nanhaiensis SCSIO 10187 (JN582175)	*R. marinonascens* 3438W (X80617)	*R. percolatus* MBS1 (X92114)	*R. imtechensis* RKJ300 (AY525785)	*R. opacus* DSM 43205 (X80630)	**248 16S rRNA gene (1495bp)**	*R. wratislaviensis* DSM 44107 (Z37138)
R. nanhaiensis SCSIO 10187 (JN582175)		31	50	36	30	27	26
R. marinonascens 3438W (X80617)	97.8		39	34	24	26	24
R. percolatus MBS1 (X92114)	96.61	97.54		40	26	35	28
R. imtechensis RKJ300 (AY525785)	97.43	97.71	97.5		17	22	10
R. opacus DSM 43205 (X80630)	97.88	98.41	98.43	98.87		13	7
248 16S rRNA gene (1495bp)	98.07	98.25	97.83	98.53	99.14		4
R. wratislaviensis DSM 44107 (Z37138)	98.14	98.38	98.27	99.32	99.54	99.73	

Percentage similarity

Nucleotide difference

The table shows the pairwise similarity values (%) and the number of nucleotide differences found between the 16S rRNA gene sequences of strain 248 and closely related type strains of validly described *Rhodococcus* species recovered in the same clade on the NJ-TREE and IQ-TREE. Darker shades indicate high similarity and low number of nucleotide differences between the sequence pair.

Table 2.17 Distance matrix for the phylogenetic analysis of the *gyrB* sequence of strain 248

	R. marinonascens IFO 14363 (AB014115)	*R. jostii* IFO 16295 (AB088664)	*R. koreensis* JCM 10743 (AB075566)	***gyrB_248 (1080bp)***	*R. wratislaviensis* JCM 9689 (AB014315)
R. marinonascens IFO 14363 (AB014115)		161	155	137	142
R. jostii IFO 16295 (AB088664)	87.25		72	49	54
R. koreensis JCM 10743 (AB075566)	87.73	94.3		48	52
gyrB_248 (1080bp)	87.33	95.47	95.56		28
R. wratislaviensis JCM 9689 (AB014315)	88.76	95.72	95.88	97.41	

Percentage similarity

Nucleotide difference

The table shows the pairwise similarity values (%) and the number of nucleotide differences found between the *gyrb* gene sequences of strain 248 and closely related type strains of validly described *Rhodococcus* species recovered in the same clade on the NJ-TREE and IQ-TREE. Darker shades indicate high similarity and low number of nucleotide differences between the sequence pair.

2.4 Discussion

Most microbiologists concur with the idea that a microbial species is an independently evolving population, whose evolution is sculpted within the group by a variety of events leading to gaps between species that allow their differentiation (De Vos *et al.*, 2017). The ultimate goal of taxonomy is to construct a classification that is of operative and predictive use for any discipline in microbiology, and that is also inherently stable (Richter and Rossello-Mora, 2009). New technologies allowing high-throughput genomic data acquisition have improved bacterial classifications significantly and have led to a database-based taxonomy centred on portable and interactive data (Rosselló-Móra *et al.*, 2017). Among currently used methods available for the classification and identification of prokaryotes, the analysis of 16S rRNA gene sequences offers a reproducible and technically easy procedure that is also scalable.

Taxonomic identity of bacterial strains, especially ones with high commercial value, is important to ensure proper communication and to avoid confusion in legal matters such as patents. Several species of *Rhodococcus* are well known for their metabolic versatility. With the test strains used in this study being rhodococci, and the research focussed on exploiting one of their metabolic pathways (biodesulphurisation of organosulphur compounds) for a potential commercial application, it is essential to clear out any confusion in their taxonomic identity. The work presented in this chapter was purposed towards confirming the identity of the test strains based on their morphological and staining characteristics, together with the use of recently recommended phylogenetic approaches. Given the complicated taxonomic history of the genus *Rhodococcus*, this aspect gains even more interest.

It is not uncommon for taxonomic classification of bacterial strains to be revisited and updated. In the past, several naturally occurring BDS strains have witnessed revisions in their taxonomy. The BDS isolates that were reported as *Pseudomonas delafieldii*, *Mycobacterium phlei*, *Desulfobacterium anilini*, and *Gordonia nitida* are currently reclassified as *Acidovorax delafieldii* (Willems *et al.*, 1990), *Mycolicibacterium phlei* (Gupta *et al.*, 2018), *Desulfatiglans anilini* (Suzuki *et al.*, 2014), and *Gordonia alkanivorans*

(Arenskotter *et al.*, 2005), respectively. Even the original desulphurising strain IGTS8, initially described as *Rhodococcus rhodochrous*, was later reclassified as *R. erythropolis*, and it is available in the culture collections as ATCC 53968, BCRC 16363, CCRC 16363, and CECT 5044.

Until 1996, which is the latest year of isolation for the test strains used in this study, only 13 *Rhodococcus* species were described. Since then, there has been a considerable increase in the number of new species described under this genus, with current count at 55 (as of December 2019). When the DDH method was used for taxonomical classification, the minimum recommended % homology values to be classified as one species was 70% (Wayne *et al.*, 1987). However, Briglia *et al.* (1996) recognised that the type strains of the species *R. opacus* and *R. percolatus* showed distinguishing phenotypic characteristics despite sharing a DNA–DNA similarity value above the 70% cut-off point and suggested that classification should not be based solely on the genomic similarities. Later, a DDH homology value above 80% was recommended for species-level kinship in the genus *Rhodococcus* (Sukhoom, 1999). On the basis of 16S rRNA, *gyrB* and *catA* gene sequence comparisons and reaffirmed DNA–DNA hybridisation, *R. jialingiae* was confirmed as a later synonym of *R. qingshengii* (Táncsics *et al.*, 2014).

The original classification of the test strains used in this study was based on morphological features, their metabolic ability to utilise alkanes as carbon source and partial amplification of 16S rRNA gene achieved through species-specific primers developed by Bell *et al.* (1999). The partial 16S rRNA sequence analysis is no longer the reliable means to distinguish the identity of closely related species that tend to share more than 99.5% similarity (Fox *et al.*, 1992; Novoselova *et al.*, 2011; Woo *et al.*, 2008). In the context of the complexity existing in the taxonomy of genus *Rhodococcus* and the need for refinement of the taxonomy of the organisms within the IEGM culture collection (as highlighted by Ivshina, 2012), it becomes imperative to resolve the species identity of the test strains.

The mycolic acid profiles, as well as the BLASTn analyses using 16S rRNA gene and *gyrB* sequences, show unambiguously that the test strains used in this study are affiliated to the genus *Rhodococcus*. It should be noted that the

16S rRNA sequences for *R. qingshengii* and *R. jialingiae* were shown to have 0.2% difference, which was later not found when 1525R primer was used (Táncsics *et al.*, 2014). This was because the variable sequence stretch upstream of the 1492R primer site did not exist in the original 16S rRNA gene sequence obtained by Wang *et al.* (2010). Therefore, to avoid any misleading outcomes, primers 27F and 1525R were used in this study to reveal the sequence stretch upstream of the 1492R primer site. For each strain, the top 20 highest sequence matches, percentage homology, sequence length compared, and corresponding E-values revealed that the 16S rRNA gene sequences of the test strains matched to strains identified by the database as members of the genus *Rhodococcus*. In the case of strains 20, 213, 505 and strain F, there was a high similarity with the type strain of *Nocardia globerula* DSM 44596 amongst other rhodococci in the top 10 list, indicating a closer link between the two genera.

The choice of alternative marker depends on the genus under investigation and the richness of the sequence data in the database. In the case of *Rhodococcus*, at the time of writing there were *gyrB* and *alkB* sequences for 20 and 30 type strains, respectively with the similarity ranges of (73.15-99.14%) and (59.23-99.16%), respectively. In the analysis of both genes, the type strains of *R. baikonurensis*, *R. globerulus*, *R. qingshengii*, *R erythropolis* formed a close cluster, and the type strains of *R. pyridinivorans* and *R. rhodochrous* formed another distinct closely related cluster. In the *alkB* gene-based phylogenetic analysis, *R. opacus*, *R. percolatus*, *R. imtechensis* and *R. wratislaviensis*, *R. jostii* and *R. koreensis* formed a loose cluster. For the type strain of *R. degradans,* the *alkB* sequence was not available, and only the *gyrB* sequence was available, which shared a high similarity (99.14%) with that of *R. qingshengii* which corresponded to 9/1059 nucleotide differences. The 16S rRNA gene analysis suggested *R. degradans* as the closely related species to strain 20, 213, 488 and strain F. Based on these observations, the *gyrB* gene was chosen to provide further resolution in the phylogeny.

The gene sequences were aligned by MUSCLE and MAFFT algorithms, both of which produced similar similarity matrices. Therefore, MUSCLE aligned sequences were used to obtain the phylogenetic trees through the popular NJ-

Treeing algorithm and the recently developed IQ-Tree method, with the latter providing a higher resolution than the former.

In the 16S rRNA gene-based phylogenetic analysis, strain 87 and 369 clustered along with *R. ruber.* This observation indicated that they are indeed *R. ruber* strains as previously classified by Bell *et al.* (1999). In the case of strains 20, 213, 505 and F, determination of pairwise 16S rRNA gene sequence similarity showed that their previous classification as *R. erythropolis* could be correct. The phylogenetic trees based shown in Figure 2.2 and 2.3, show that the strains 20, 213, 505 and F fell within the *R. erythropolis* clade encompassing *R. degradans* and *R. qingshengii*. Strain F (initially received as *Rhodococcus* sp.) and strains 20, 213 and 505 (received initially as *R. erythropolis*) grouped closely with *R. degradans,* whose type strain CCM 4446^T, was recently reclassified from the *R. erythropolis* group based on *catA* and *gyrB* gene analysis (Švec *et al.*, 2015).

The identity of strains 208, 488 and 508 differed in the IEGM web-catalogue, the printed catalogue, and the findings by Bell *et al.* (1999). In the 16S rRNA and *gyrB* phylogenetic trees obtained in this study, they occur together as a distinct cluster in the phylogenetic analysis indicating that they are highly likely to be potentially new species. Strain 248 was received as *R. opacus, and* it was found closely related to *R. wratislaviensis* (99.73% similarity) as per 16S rRNA gene sequence-based phylogenetic analysis and the BLASTn results obtained in this research. As per the taxonomic history, the type strain *R. wratislaviensis* N805^T was reclassified from the genus *Tsukamurella* which was earlier described as closely related to *R. opacus* and *R. percolatus* (Goodfellow *et al.*, 2002). The high pairwise similarity of 16S rRNA gene sequence of strain 248 with *R. opacus* (99.54%) could not be resolved using gyrB gene sequence analysis owing to the lack of a validated *gyrB* gene sequence for *R. opacus* type strain.

This study supports original genus classification of all strains but highlights the inaccuracies in the IEGM classification and the limitations of the study by Bell and team, based on diagnostic PCR using species-specific primers. The present study reveals 1 or 2 new lines of descent representing one or two possible previously undescribed species found in hydrocarbon-contaminated

soil. Besides reinforcing the value of 16S rRNA gene sequence analysis to establish the phylogenetic position of unknown strains, the study showed the *gyrB* gene-based phylogenetic analyses offered only a limited level of resolution, affected mainly by the lack of reference sequences for all the type strains. However, the results of the phylogenetic analysis of *gyrB* were broadly congruent with 16S rRNA.

The micro- and macromorphology features of the test strains observed in this research were compared with the species description available in the Bergey's Manual Vol.5 (Goodfellow *et al.*, 2012). Among the Corynebacteriales, the main distinguishing morphological features for rhodococci is the ability of most strains to form hyphae that fragment into rods and cocci, although they do show considerable heterogeneity. Elementary branching prior to fragmentation was observed in the case of strains 20, 213, 505 and strain F, which is also a characteristic feature among *R. erythropolis*. The growth of these colonies on NA plates, (flat, circular colonies with whole margins and pale salmon-pink pigment concurs with the description of *R. degradans* colonies growing on the same medium and therefore, they could be reclassified as *R. degradans*. In this study, the strains 208, 488 and 508 also showed elementary branching prior to fragmentation, as observed in the case of *R. erythropolis* strains. But unlike the smooth edges of *R. erythropolis* colonies*,* these strains formed cream-coloured, opaque, convex colonies with irregular edges. This observation, together with the clustering observed in the phylogenetic trees, indicates that they could potentially be a different new species altogether. Strains 87 and 369 formed rough, pink to red colonies on GYEA which matched the morphological description of *R. ruber* species, also supporting the observations made from the phylogenetic trees. The macromorphology of colonies of strain 248 (cream-coloured colonies with an entire margin and a rough surface; a depressed centre after several days of incubation as seen in Figure 2.10) is more concordant with the species description for *R. opacus* than the other strains.

Apart from the 11 test strains, it could be noted from the IQ-TREE (Figure 2.3) that the type species of four other mycolic acid-containing genera also appear to fall within the evolutionary span encompassed by members of the genus

Rhodococcus; the type strain of *Williamsia muralis* MA140-96T, *Nocardia asteroides* DSM43757T, *Segniliparus rotundus* CDC1076T and *Smaragdicoccus niigatensis* DSM44881T. These type species of different genera fall within a distinct, deep-rooted subclade containing 6 validly described *Rhodococcus* species. *Segniliparus* species were reported to fall within the phylogenetic radiation encompassed by *Rhodococcus* species, and most closely with the type strain of *R. equi*, in the original description paper by Butler *et al.* (2005). However, these were distinguished from this species and the genus *Rhodococcus* based on DDH and other chemotaxonomic data. It should be noted that the 6 valid *Rhodococcus* species in this subclade were described subsequent to the description of *Segniliparus* (Butler *et al.*, 2005) and *Smaragdicoccus* (Adachi *et al.*, 2007). The present study also reveals the high degree of homology of the 16S rRNA gene sequence among rhodococci despite the diverse polluted soil environments from where *R. erythropolis, R. rhodochrous* and *R. ruber* (all isolated from soil from activated sludge), *R. opacus* (soil around a defective town gas pipe), *R. degradans* (a soil contaminated by organic pollutants) and *R. qingshengii* (a carbendazim-contaminated soil) were originally isolated.

Chapter 3 Biodesulphurisation of benzothiophene and dibenzothiophene present in aqueous medium and diesel oil by new rhodococci strains

3.1 Introduction

The metabolic abilities of several species under the genus *Rhodococcus* have found a wide range of biotechnological applications. They are applied for the degradation of a variety of environmental pollutants and transformation or synthesis of compounds with possible useful applications. Biodesulphurisation (BDS) is one of their mainstream applications. In this chapter, the importance of natural bacterial isolates for the development of the BDS technology and the methodology followed for their identification are described. Based on the current challenges in BDS technology, the ideal characteristics of a commercially applicable strain are recognised. The high suitability of rhodococci for biodesulphurisation is then established, followed by examination of the test strains for BDS activity in a wholly aqueous medium, and subsequently in an oil-water biphasic medium in order to select the candidate BDS strains for future scaling up.

3.1.1 Development of bacterial catalysts for BDS

In the early period of development, most of the environmental isolates and enrichment cultures of mixed bacterial populations that were tested for BDS activity exhibited metabolisation of organosulphur compounds by initiating biodegradation at the C-C bond. However, there was a mixed culture growing in a medium enriched with thiophene as its sole source of sulphur, which exhibited C-S bond cleavage for about 20% of its products (Yamada *et al.*, 1968). The most successful microorganism for sulphur utilisation from organosulphur compounds was *Pseudomonas* isolated from enrichment cultures employing DBT as the sole source of sulphur. The organism, however, used DBT as a carbon substrate as well (Stoner *et al.*, 1990). Thus, an unnatural, selective mutation process was utilised to develop a microorganism having selective sulphur metabolism. A mixture of bacterial isolates sourced from oil-contaminated soils was subjected to directed evolution by exposing the isolates to the mutagen NTG -1 methyl-3-nitro-1-nitrosoguanidine in a continuous flow bioreactor at the Institute of Gas Technology (IGT), USA. As a result, a mixed culture of 7 mutant strains with capabilities for sulphur specific C-S bond cleavage was developed (Kilbane, 1990, 1989). This group of microorganisms was code-named IGTS7. Among

these strains, *Rhodococcus rhodochrous* IGTS8 and *Bacillus sphaericus* IGTS9 were found to be highly capable of BDS (Kayser *et al.*, 1993; Kilbane and Jackowski, 1992). The development of IGTS8 led to significant improvements in the field of biodesulphurisation such as the description of the metabolic pathway (the 4S pathway) of the sequential enzymatic desulphurisation of DBTand the genes encoding the enzymes (Denome *et al.*, 1994; Oldfield *et al.*, 1997; Piddington *et al.*, 1995). The important criterion for a candidate BDS strain is the ability to assimilate sulphur from the organic compounds without destroying the carbon skeleton of the compound (Boniek *et al.*, 2015; Izumi *et al.*, 1994; Labana *et al.*, 2005; Mohebali and Ball, 2016, 2008; Stanislaus *et al.*, 2010; Xu *et al.*, 2009). This is essential to preserve the calorific value of the fuel. This is a primary differentiating factor between BDS strains and others that are capable of biodegradation of oil, in which case CO_2 is the desired product. In general, the anaerobic processes were slower and poorly efficient, and hence microorganisms that follow an aerobic process were preferred (Debabov, 2010; Setti *et al.*, 1997). Although successful BDS organisms have been identified from both Gram-positive and Gram-negative types, there have been only a few reported cases of the latter (Gunam, 2013).

The naturally occurring desulphurising strains were often sourced from oil-polluted environments and were isolated by selective enrichment culturing techniques (Ahmad *et al.*, 2014; Bhatia and Sharma, 2010; Chauhan *et al.*, 2015; Chen *et al.*, 2008; Davoodi-Dehaghani *et al.*, 2010; Gilbert *et al.*, 1998; Gunam *et al.*, 2006; Izumi *et al.*, 1994; Khedkar and Shanker, 2015; Kilbane II, 2006; Kirkwood *et al.*, 2005; Labana *et al.*, 2005; Li *et al.*, 2005b; Papizadeh *et al.*, 2010; Rhee *et al.*, 1998; Song and Ma, 2003; Tanaka *et al.*, 2001; Wu *et al.*, 2011; Yang and Marison, 2005). Since the description of the first BDS strain IGTS8, many researchers have isolated new naturally occurring bacteria from diverse taxonomic groups capable of BDS of BT and/or DBTas listed in Appendix 7.5. Screening for bacterial strains with BDS activity is widely done by enrichment culturing in a chemically defined Mineral Salts Medium (MSM) devoid of any easily assimilable sulphur sources (like sulphates). The MSM is supplemented with the polycyclic aromatic sulphur compound of interest as the sole sulphur source. Benzothiophene (BT) and Dibenzothiophene (DBT) are the widely used model polycyclic aromatic sulphur compounds to study

bacterial desulphurisation, whilst some authors include alkyl derivatives of DBT (alkyl-DBT) and asymmetric structural isomers of DBT such as naphthothiophene in their studies to show the broader substrate range of the BDS strains (Table 3.1).

Table 3.1 Naturally occurring bacterial strains capable of BDS of alkyl derivatives of BT and DBT

Bacterium	alky-DBT substrate	Reference
***Achromobacter* sp.**	4-methyl DBT	Bordoloi *et al.*, 2014
***Chelatococcus* sp.**	4,6-dimethyl DBT	Bordoloi *et al.*, 2016
***Bacillus subtilis* WU-S2B**	2,8-dimethyl DBT 4,6-dimethyl DBT 3,4-benzo DBT	Kirimura *et al.*, 2001
***Gordonia sp.* HS126-4N**	2,8-dimethyl DBT	Akhtar *et al.*, 2018
***Paenibacillus sp*. A11-2**	Methyl, ethyl, dimethyl, trimethyl and propyl DBTs	Onaka *et al.*, 2001b
***Microbacterium sp.* ZD-M2**	4,6-dimethyl DBT	Li *et al.*, 2005b
***Mycobacterium sp.* G3**	4,6-dibutyl DBT 4,6-dipentyl DBT	Okada *et al.*, 2002
	4,6-dimethyl DBT 4,6-diethyl DBT	Nomura *et al.*, 2005
***Mycobacterium sp.* ZD-19**	4,6-dimethyl DBT	Chen *et al.*, 2008
***Rhodococcus sp*. ECRD-1**	4,6-diethyl DBT	Lee *et al.*, 1995
***Rhodococcus sp.* SA11**	1-methyl DBT, 4-methyl DBT, 2,3-dimethyl DBT, and 4,6-dimethyl DBT	Mohamed *et al.*, 2015
R. erythropolis* IGTS8	C1 and C2 substituted DBTs	Kaufman *et al.*, 1999
***R. erythropolis* H-2**	2,8-dimethyl DBT 4,6-dimethyl DBT 3,4-benzo DBT	Ohshiro, 1996
***R. erythropolis* KA2-5-1**	Methyl, ethyl, dimethyl, trimethyl DBTs	Kobayashi *et al.*, 2000
***R. erythropolis* XP**	4-methyl DBT 4,6-dimethyl DBT Benzonaphthothiophene	Yu *et al.*, 2006a
Strain RIPI-S81	4-methyl DBT 4,6-dimethyl DBT	Rashidi *et al.*, 2006

Although it is known that DBT desulphurising strains can implicitly exhibit the BDS of alkyl-derivatives of DBT also, some authors studied the latter activity explicitly. All the strains listed in the table can desulphurise DBT also.

A majority of BDS strains can desulphurise either DBTs or BTs, while only a few species are capable of desulphurising both (Bachmann *et al.*, 2014; Khedkar and Shanker, 2015). The majority of reported BDS bacteria that are capable of desulphurising the symmetric heterocyclic sulphur in DBT are not reported to desulphurise the asymmetric heterocyclic sulphur in BT and its derivatives (Mohamed *et al.*, 2015; Soleimani *et al.*, 2007; Srivastava, 2012). Table 3.2 shows examples of naturally occurring bacterial isolates with different range of substrates for desulphurisation exhibiting different levels of BDS activity. It should be noted that there is a lack of a standardised way of reporting the BDS activity in the literature to make meaningful comparisons. *Gordonia desulfuricans* 213E (NCIMB 40816) (Gilbert *et al.*, 1998), *Rhodococcus* sp. T09 (Matsui *et al.*, 2000) and *Sinorhizobium sp*. KT55 (Tanaka *et al.*, 2001), are some of the strains reported to desulphurise BT only. A few bacteria, such as *Paenibacillus sp*. A11-2 (Ishii *et al.*, 2000; Konishi *et al.*, 2000b), *Mycobacterium goodii* X7B (Li *et al.*, 2005a) and *Gordonia alkanivorans* RIPI90A (Shavandi *et al.*, 2009) harbouring DBT desulphurisation genes or their homologs are capable of the biodesulphurisation of both DBT and BT. *Gordonia* sp. IITR100 desulphurised thianthrene, which is a non-thiophenic sulphur compound, alongside DBT (Ahmad *et al.*, 2014). Some bacteria such as *Bacillus subtilis* WU-S2B (Kirimura *et al.*, 2001), *Paenibacillus* sp. A11-2 (Konishi *et al.*, 2000b), and *Mycobacterium phlei* WU-F1 (Furuya *et al.*, 2001), exhibited thermophilic BDS of DBT and its alkyl derivates, at 50 °C. Interestingly, desulphurising bacteria sourced from the same location may not have the same desulphurisation spectrum. It was observed that *Rhodococcus* sp. SA11, *Stenotrophomonas* sp. SA21 and *Rhodococcus* sp. SA31 that were isolated from the same area had similar capabilities for desulphurising DBT and its alkylated homologs, but only the SA11 strain could use BT as a sulphur source (Mohamed *et al.*, 2015).

Table 3.2 Naturally occurring bacterial isolates and their BDS activity

Desulphurisation organism	Substrate	BDS activity	Reference
***Corynebacterium* sp. SY1**	DBT	17%	Omori *et al.*, 1992
***Gordonia alkanivorans* 1B**	DBT, methyl-DBT	77% (168 h)	Alves *et al.*, 2008
***G. alkanivorans* RIPI90A**	DBT	182 µM h^{-1}	Mohebali *et al.*, 2007
***G. desulfuricans* 213E**	BT, DBT	N/A	Gilbert *et al.*, 1998
***Gordonia* sp. ZD-7**	DBT	93% (48 h)	Li *et al.*, 2006
***Lysinibacillus sphaericus* DMT-7**	DBT	60% (360 h)	Bahuguna *et al.*, 2011
***Microbacterium* sp. NISOC-06**	DBT	94.8% (336 h)	Papizadeh *et al.*, 2010
***Microbacterium* ZD-M2**	DPS, DBT, dimethyl-DBT, TH, BT	70–100% (70 h)	Li *et al.*, 2005b
***Mycobacterium goodie* X7B**	DBT, Lianing crude oil	59% (24 h), 99% (72 h)	Li *et al.*, 2007b
***Mycobacterium phlei* WU-0103**	BT Naphtho [2, 1-b] TH	52% (72 h)	Ishii *et al.*, 2005
***Mycobacterium* sp. ZD-19**	TH or BT, DBT or 4,6-DMDBT	100% (10h or 42 h) 100% (50h or 56 h)	Chen *et al.*, 2008
***Pantoea agglomerans* D23W3**	DBT, alkyl-DBT	93% (24h)	Bhatia and Sharma, 2010
***Pseudomonas stutzeri* TCE3**	DBT	N/A	Dinamarca *et al.*, 2010
***R. erythropolis* FSD-2**	Diesel	97% (12 h)	Zhang *et al.*, 2007
***R. erythropolis* LSSE8-1**	DBT	79.4% (24 h)	Li *et al.*, 2009
***R. erythropolis* SHT87**	DBT, DBT-sulphone, TH, alkyl-TH	100% (10 h)	Davoodi-Dehaghani *et al.*, 2010
***R. erythropolis* XP**	Jilian FCC gasoline, SR gasoline	30–85% (12 h)	Yu *et al.*, 2006a
***Rhodococcus* sp. JVH1**	BT, methyl-BT	Varied significantly	Kirkwood *et al.*, 2007a
***Rhodococcus* sp. 1awq**	DBT	0.26 µmol g^{-1} min^{-1}	Ma *et al.*, 2006
***Shewanella putrefaciens* NCIMB8768**	DBT	43.5 mmol/L HBP (72 h)	Ansari *et al.*, 2007
***Sphingomonas subarctica* T7b**	BT, alkyl-DBT,	41% (36 h)	Gunam *et al.*, 2006

TH – thiophenes; BT - benzothiophene; DBT – dibenzothiophene; DPS - diphenylsulphide

The BDS activity of the strains is presented in the table, as reported in the original research article. In some publications, the activity BDS is represented in terms of % reduction in sulphur levels observed for a certain time of reaction/incubation. In other articles, the BDS activity is reported in terms of the amount of substrate utilised or product formed. There is no standardised way of reporting BDS activity in the literature. The different ways of presenting BDS activity makes it difficult to make meaningful comparisons between the performances of different BDS strains.

As a result of desulphurisation, thiophenes are converted into corresponding monohydroxy biphenyls (phenolic end products) that are later detected using the Gibbs assay (described in section 3.1.2). Following the successful BDS activity exhibited by the growing cells in an oil-free aqueous growth medium, they are prepared as resting cell systems (which are non-growing live cells that retain most of the enzyme activities of growing cells, collected at the late exponential phase and resuspended in phosphate buffer at a high cell density). The resting cells are then mixed with fuels to assess their oil desulphurisation capability. There are also reports where the BDS strains were grown directly in an oil-water biphasic medium to achieve desulphurisation of the oil. The oil phase could be laboratory reagent grade solvents (model oil) or naturally sourced crude oil or fossil fuels obtained downstream of the HDS process. The antimicrobial action of a solvent is correlated to its hydrophobicity and can be measured as the logarithm of the octanol-water partition coefficient (log P values). Hydrophobic solvents, with a log P value > 4, accumulate in the membrane, but will not reach a high membrane concentration and are not toxic because of their low water solubility, whereas the highly water-soluble solvents with log P value < 4 are toxic to most organisms (de Bont, 1998). Solvents such as *n*-hexadecane, *n*-tetradecane, *n*-dodecane, and *n*-octane which have log P values of 8.8, 7.2, 7.0, and 4.5, respectively have been frequently used as model oils, and the optimal oil-water ratios for BDS activity are studied. Some components of fossil fuels, such as cycloalkanes and naphthalenes, are toxic to microorganisms (Sikkema *et al.*, 1995; Tao *et al.*, 2006). Therefore, only the microorganisms that are inherently or adaptively tolerant to petroleum-based solvents could be used for the BDS of fossil fuels (Table 3.3). Higher oil content causes mass transfer limitations and also limits the oxygen supply, and therefore, the ratio of oil/water in the biphasic medium must be optimised (Adlakha *et al.*, 2016).

Table 3.3 Desulphurisation of real-world petroleum by naturally occurring bacterial isolates

Desulphurising bacteria	Desulphurisation yield	References
***Rhodococcus erythropolis* I-19**	67% desulphurisation of petroleum	Folsom *et al.*, 1999
***Rhodococcus* sp. ECRD-1**	669 p.p.m desulphurisation middle distillate oil	Grossman *et al.*, 2001
Rhodococcus* sp. and *Athrobacter sulphurcus	50% desulphurisation of diesel oil	Labana *et al.*, 2005
***R. erythropolis* XP**	94.5% desulphurisation of diesel oil	Yu *et al.*, 2006a
***R. globerulus* DAQ3**	1580 ppm desulphurisation of diesel oil	Yang and Marison, 2005
***Gordonia* sp. CYKS1**	Reduction of sulphur from 0.15% (wt/wt) to 0.06% (wt/wt) of middle distillate unit feed (MDUF)	Rhee *et al.*, 1998
***P. delafieldii* R-8**	313 mg.L^{-1} desulphurisation from 591 mg.L^{-1} of diesel oil	Guobin *et al.*, 2006
***Nocardi*a sp. CYKS2**	0.3 - 0.24 wt.% desulphurisation of diesel oil	Chang *et al.*, 2000
***Mycobacterium phlei* WU-0103**	52% desulphurisation gas oil fraction from 1000 to 475 ppm	Ishii *et al.*, 2005
***Mycobacterium goodii* X7B**	86% desulphurisation diesel oil	Li *et al.*, 2003
***Mycobacterium goodii* X7B**	59% desulphurisation of Liaoning crude oil from 3600 to 1478 ppm	Li *et al.*, 2007b
Mycobacterium goodii X7B	immobilised cells of strain X7B the total sulphur content significantly decreased, from 227 to 71 ppm at 40 °C.	Li *et al.*, 2005a

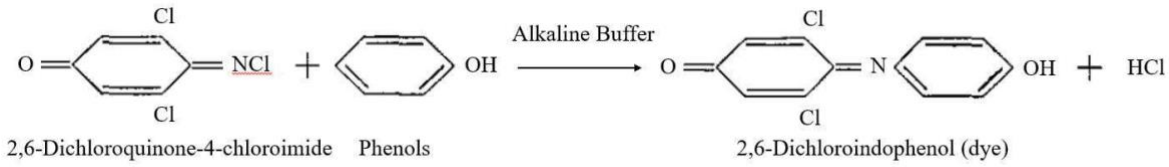
The table shows examples of research where BDS strains were used for the desulphurisation of real-world fuels. The ability of the organisms to survive the solvent (fuel) environment is an important trait that determines the suitability for industrial process. There is no standard across BDS research in terms of the feedstock specification to use for validating the fuel desulphurisation activity.

Additionally, when biphasic (oil-water) or tri-phasic (oil-water-microorganisms) systems are used, it is critical to use strains that offer efficient separation and uncomplicated downstream processing because the many oil tolerant BDS strains form emulsions which make product recovery difficult (Derikvand and Etemadifar, 2014; Naito *et al.*, 2001; Todescato *et al.*, 2017). The biotransformation of organic compounds using growing cells is usually conducted in complex growth media. The growth of the bacteria results in a complicated mixture of metabolites produced in the culture system. Konishi *et al.,* (1997) factored this into account and mentioned the potential interference that these components could pose while measuring the desulphurising activity, and therefore for the purpose of quantification, they resorted to using whole resting cells. This method is popular ever since and widely reported in the literature. Nevertheless, they also indicated that using growing cells for studying biodesulphurisation has the advantage that the growth of bacterial cells can be a reliable indicator of their utilisation of the chemicals as their essential nutritional components.

Following successful desulphurisation, the sulphur levels in the sample are quantified and presented as the percentage reduction in the total sulphur content of the feed. There are current ASTM methods specifying Induction Coupled Plasma – Optical Emission Spectroscopy (ICP-OES) for the quantification of a variety of elements in similar matrices (given in parentheses), such as D7151-15 (insulating oils), D7111-16 (middle distillate fuels), D5185-13e1 (lubricating oils and base oils), and D4951-14 (lubricating oils). Measuring the sulphur levels in aqueous solutions using ICP-OES is a straightforward procedure, whereas careful considerations are needed in the sample preparation step when analysing unstable systems like emulsions and oils (Amais *et al.*, 2014; Young *et al.*, 2011). The authors, who proposed xylene diluted solutions of petroleum crude oil, also emphasized the importance of sample introduction system and torch maintenance to avoid memory effects and remove carbon deposits in the system (Fabec and Ruschak, 1985). They also used a base oil in the calibration solutions to minimise viscosity differences.

3.1.2 Gibbs assay

Gibbs (1927) reported that dihalogen substituted quinone-chloroimides such as 2,6-dichloroquinonechloroimide and 2,6-dibromoquinonechloroimide formed stable, measurable intense blue coloured indophenols when they react with phenols. This reaction came to be known as the Gibbs assay for the detection of phenols, and it works for phenols without para-substitution. In the case of para-substituted phenols, the specific product depends on the substituent and affects the colour development and hence the absorbance values.



2,6-Dichloroquinone-4-chloroimide Phenols 2,6-Dichloroindophenol (dye)

This reaction is very sensitive – can detect phenol 1 part in 20,000,000, and colour development is favoured between pH 8-10 (pH 9.4 preferred), and the readings are measured at 610 nm, the peak of the absorption band for indophenol. However, complete colour development requires up to 24 hours. Additionally, the wavelength of maximum extinction varies depending on the type of phenol studied (Ettinger and Ruchhoft, 1948).

Gibbs assay has been used to monitor the desulphurisation activity of bacteria since the first desulphurising strain *R. rhodochrous* IGTS8 was reported (Kayser *et al.*, 1993). Ever since, Gibbs assay has become one of the routine techniques in the field of biodesulphurisation research that uses BT or DBT and their derivatives as the target model organosulphur compounds (Alkhalili *et al.*, 2017; Ansari *et al.*, 2007; Castorena *et al.*, 2002; Chauhan *et al.*, 2015; Gherbawy *et al.*, 2016; Gilbert *et al.*, 1998; Gunam *et al.*, 2006; Kayser *et al.*, 1993; Kirkwood *et al.*, 2005; Mohamed *et al.*, 2015; Peng and Zhou, 2016; Rashidi *et al.*, 2006; Yu *et al.*, 2006a, 2006b).

3.1.3 Adaptation of Gibbs reaction for the development of newer methods

The Gibbs reaction mechanism of detecting phenolics has been exploited to develop biosensors and for other extended applications. In most cases, Gibbs reaction was used for quicker qualitative and quantitative measurements to eschew time consuming and expensive chromatographic methods like HPLC. Petcu and team (2004) developed a new fast and reliable test based on molecular imprinting (MIP) techniques for the rapid quantification of propofol (anaesthetic) in blood based on Gibbs reaction. Bhuiya and Liu (2009) developed a convenient and cost-effective assay method for measuring the catalytic properties of enzymatic transmethylation of phenolics using Gibbs reagent where the catalytic activity of the enzymes was measured as colour changes that happened during the regiospecific methylation reaction catalysed by enzymes. Bashir and Liu (2009) developed an α-cyclodextrin/Gibbs reagent biosensor on Platinum-sputtered glass slide, in which the Gibbs reagent was held trapped in the α-cyclodextrin. When the phenol molecule is in close proximity to the Gibbs reagent, charge transfer happens between the two molecules that make up the 'detecting element' entrapped within a sol-derived type matrix. The charge transfer led to a colour change that was detected colourimetrically or cyclic voltammetrically. This modification had advantages such as high substrate selectivity, lower cost of analysis per sample, miniaturization and simplicity. Arip and team (2013) developed an optical chemical sensor for rapid detection of permethrin (a protectant for wood against termite attack) based on the reaction between permethrin and Gibbs reagent. In this research, deriving inspiration from above-described exemplar studies where Gibbs reaction principle was used to develop high-throughput sensors, a modified Gibbs test methodology was attempted aimed at direct detection of Gibbs positive bacterial colonies growing on solidified medium. If successful, this methodology can be adopted to detect and isolate a single colony of BDS bacteria from a mixed population.

3.1.4 Rhodococci used for BDS of fossil fuels

Members of *Rhodococcus* are well known for their metabolic versatility and with proven tolerance and degrading ability against organic solvents (Kim et

al., 2018). For example, tolerance to high levels of benzene was exhibited by *Rhodococcus* strain 33 (Paje *et al.*, 1997). *Rhodococcus opacus* has been shown to be tolerant of toluene, styrene, xylene, octane, decane (Na *et al.*, 2005) and *Rhodococcus* was the dominant genus of polychlorinated biphenyls degrading bacteria (Leigh *et al.*, 2006). *R. erythropolis* strains have exhibited tolerance to acetonitrile (Langdahl *et al.*, 1996). The rhodococcal extracellular polysaccharides act as a natural barrier protecting *Rhodococcus* species from toxic aromatics like *n*-hexadecane (Iwabuchi *et al.*, 2000) and have been reported as the cause for emulsification, moisture retention and adsorption, and thickening (Urai *et al.*, 2007). The high tolerance to solvents and toxic chemicals indicate that rhodococci are appropriate microorganisms for use in the BDS process (Stancu, 2014; Todescato *et al.*, 2017). The highly hydrophobic nature of the rhodococci cell membrane is an added advantage as it increases the contact between the cell and the sulphur compounds in the fuels (Monticello, 2000). In a biphasic medium, rhodococci preferentially exist at the oil-water interface reaction which naturally facilitated the uptake of DBT occurring in the oil phase. The strain SY1 has been shown to desulphurise dibenzyl sulphoxide to benzyl alcohol and toluene (Omori *et al.*, 1995), and the strain IGTS8 can utilise 1-chloroethyl sulphide as the sole sulphur source, indicating a potential application in the detoxification of warfare agent mustard (2,2′ dichlorodiethyl sulphide) (Kilbane II and Jackowski, 1996).

The recovery of the desulphurised oils and the short life of biocatalysts were major hindrances in developing free whole-cell systems. Pacheco *et al.* (1999) proposed that the reaction time and longevity of cells should be of 1 and 400 hours, respectively, for a cost-effective BDS process. Immobilisation of the cells improved longevity and oil recovery. When *R. erythropolis* KA2-5-1 was immobilised by entrapping them with calcium alginate, agar, photo-cross-linkable resin prepolymers, the BDS activity was sustained repeatedly in *n*-tetradecane containing DBT for more than 900 h with reactivation. However, the average desulphurisation rate was lower than that in an *n*-tetradecane/water/cells triphasic system with the same bacterial strain (Naito *et al.* 2001). Calcium alginate immobilisation in combination with coating the surface of cells with superparamagnetic Fe_3O_4 nanoparticles, resulted in the easy separation and reuse of *Rhodococcus erythropolis LSSE8-1* by growing

in an *n*-dodecane/water biphasic medium (Li *et al.*, 2009). This strategy was adopted by Dai *et al.* (2014) who reported that desulphurising activities of immobilised and free cells of *Brevibacterium lutescens* CCZU12-1 exhibited the similar time courses (8 hours), and the combination of magnetic Fe_3O_4 nanoparticles enabled easy recovery and reuse of the immobilised biocatalyst up to 4 times, but with subsequent batches taking a longer time to achieve complete desulphurisation. Encapsulation of *R. erythropolis* R1 in calcium alginate beads with a combination of nano γ-Al_2O_3 resulted in a very effective BDS characterised by increased DBT consumption. The nanoparticles were adsorbed on the cells creating additional surface area and pores for enhanced DBT capture. The viability of the immobilised cells was 86%, whereas that of the free growing cells was 98%. Although the 12% reduction in the viability of the cells was acceptable, due to nanoparticles adsorption, some of the endproduct (2HBP) remained at the cell surface and was not released entirely into the medium (Derikvand and Etemadifar, 2014). Thus it is clear that immobilisation of biocatalysts helps in easy separation and reusability. However, the mass transfer resistance is enhanced compared to that in free-cell systems, mainly due to internal mass transfer limitations (León *et al.*, 1998). Therefore, biphasic reaction systems are the ones to be applied in the real industry, enhancing the solubility and availability of hydrophobic substrates to the biocatalyst and limiting biocatalyst inhibition by hampering the accumulation of 2-HBP, which is recovered into the oil phase from the aqueous phase.

Based on this understanding, Yu *et al.* (2006a) studied the BDS of methylated benzothiophene (3-M-BT) present in gasoline by *Rhodococcus erythropolis* XP grown free cells in a biphasic medium and reported its ability to take up high concentrations of 3-M-BT from n-octane. Yang *et al.* (2007) reported that the BDS efficiency of whole growing cells of *Rhodococcus globerulus* DAQ3 was higher in a fed-batch system (200 ppm sulphur removal) than in a batch system (120 ppm sulphur removal). Maass *et al.* (2015a) reported DBT-BDS by the *R. erythropolis* ATCC 4277 in a batch reactor using a biphasic system; DBT dissolved in *n*-dodecane/water of different ratios 20, 80, and 100 % (v/v), recording BDS efficiencies of 93.3, 98.0, and 95.5 %. Free growing cells of the

strain ATCC 4277 also helps to achieve other desirable effects such as denitrogenation, besides the desulphurisation (Maass *et al.*, 2015b).

There are more than 90 naturally occurring bacterial BDS strains belonging to diverse genera reported until December 2017 (Appendix 7.5), and *Rhodococcus* predominates the list, especially *R. erythropolis* strains (El-Gendy and Nassar, 2018; Xu *et al.*, 2009). Since the development of the first BDS strain **IGTS8**, other strains of *R. erythropolis* with BDS capability have been isolated, such as **SY1** (identified initially as a *Corynebacterium* sp.) (Omori *et al.*, 1995, 1992); **D-1** (Izumi *et al.*, 1994); **N1-36**, **N1-43**, **Q1a-22** (Wang and Krawiec, 1996); **I-19** (Folsom *et al.*, 1999); **T09** (Matsui *et al.*, 2000); **KA2-5-1** (Onaka *et al.*, 2001a); **KT462** (Y. Tanaka *et al.*, 2002); **XP** (Yu *et al.*, 2006b, 2006a); **ATCC 4277** (Maass *et al.*, 2015b); **PD1** (Derikvand *et al.*, 2015); **AF21875** (Parravicini *et al.*, 2016). These naturally occurring strains were all isolated from polluted soil environments and are very closely related based on their 16S rRNA gene sequence similarity but differed in terms of their BDS activity. Strain T09 exhibited BDS of BT only, whereas original strains KA2-5-1 exhibited BDS of DBT only, and the strains KT462 could desulphurise both model compounds. Only a few other naturally occurring BDS strain belonging to other species under genus the *Rhodococcus* have been reported. *R. fascians* Eu-32 isolated from the roots of a eucalyptus tree and *R. ruber* 9C isolated from coal contaminated soil are the only other naturally occurring non-erythropolis *Rhodococcus* strain with BDS activity (Akhtar *et al.*, 2015; Mishra *et al.*, 2017). There are many other naturally occurring rhodococci isolates whose species identity remains unresolved to date, and several recombinant rhodococci strains with enhanced BDS activity have been developed. Nevertheless, genetically engineered *R. opacus* ROD2-8 (modified from a non-BDS strain B-4) and strain ATCC 17039; and *Rhodococcus ruber* G3 were developed to take advantage of their relatively higher solvent tolerance nature compared to *R. erythropolis* species (Franchi *et al.*, 2003; Kawaguchi *et al.*, 2012; Pan *et al.*, 2013). Therefore, it would be a sensible approach to look for new naturally occurring BDS capable strains among metabolically versatile members of genus *Rhodococcus* such as *R. erythropolis*, *R. opacus, and R. ruber,* that were primarily isolated from oil-contaminated regions and study their BDS activity as growing cells in a biphasic medium.

3.1.5 Aim

The aim of this study is to investigate the biodesulphurisation capabilities amongst the rhodococci isolates obtained from oil-polluted environments and identified to species level in Chapter 2. The aim extends further into identification of one or more candidate BDS strains that can be applied for commercial biodesulphurisation of diesel fuels. Such strains(s) must possess stable desulphurisation activity, grow in a low-cost culture medium, and exhibit culture characteristics that will readily allow downstream separation when grown in a biphasic growth medium containing diesel oil.

Objectives

1) To identify strains, from the bank of 11 rhodococci isolates, that demonstrate BDS capabilities when grown on mineral salts medium (MSM) containing benzothiophene (BT) or dibenzothiophene (DBT) as the sole source of sulphur, using Gibbs detection test.

2) To identify and assess the BDS activity of strains when grown in an MSM containing minimal ingredients, for the purpose of ensuring its robustness and also minimising the cost of the growth medium for future commercial development.

3) To investigate the growth properties and desulphurisation activity of candidate rhodococcal strains when grown as resting cell suspensions and as growing cells in a biphasic medium containing BT and/or DBT dissolved in either a model oil (*n*-hexadecane) or in diesel.

4) To develop a modified Gibbs test methodology for the rapid detection of BDS capable bacterial colonies from a mixed population of primary isolates and/or mutant libraries grown on a solid medium.

3.2 Materials and Methods

3.2.1 Preparation of the test strains for BDS study

The bacteria were revived from their respective glycerol stocks for studying their BDS capability. About 100 µl of glycerol stock was inoculated into 250 ml shake flasks containing 100 ml of nutrient broth (NB) medium and incubated at 28ºC over a period of 4 days, with orbital shaking at 180 rpm (Infors HT Ecotron, Bottmingen, Germany). The increase in biomass was confirmed by turbidity measurements (Jenway 6320D, Techne, UK). The cells were harvested by centrifugation at 3250 x *g* for 5 minutes and resuspended in 5 ml of sterilised quarter-strength Ringer's solution, which is an isotonic diluent for bacterial cells.

It was imperative to avoid the carryover of any nutrients, especially sulphur components from the richer NB medium into the mineral salts medium (MSM) used for studying the BDS activity. Therefore, the cells were washed thrice using Ringer's solution before inoculation into the enrichment medium (Ansari *et al.*, 2007). The cell washing cycle involved centrifugation of the cell suspension at 3250 *x g* for 5 minutes (Hettich Mikro 200, Andreas Hettich GmbH & Co. KG, Germany) and resuspension of the pellets in fresh Ringer's solution (5 ml) at every cycle. After the third wash cycle, the washed cell pellets were resuspended in 1 ml Mineral Salts Solution (MSS) and used to inoculate the MSM medium with an initial cell concentration of 0.1 optical density at 600 nm ($O.D_{600}$) to study the desulphurisation activity.

3.2.2 Identification of the test strains with BDS capability

The sequence of experiments conducted to identify the strains with BDS activity from the 11 test strains, and the approach used to develop a rapid screening method is illustrated in Figure 3.1.

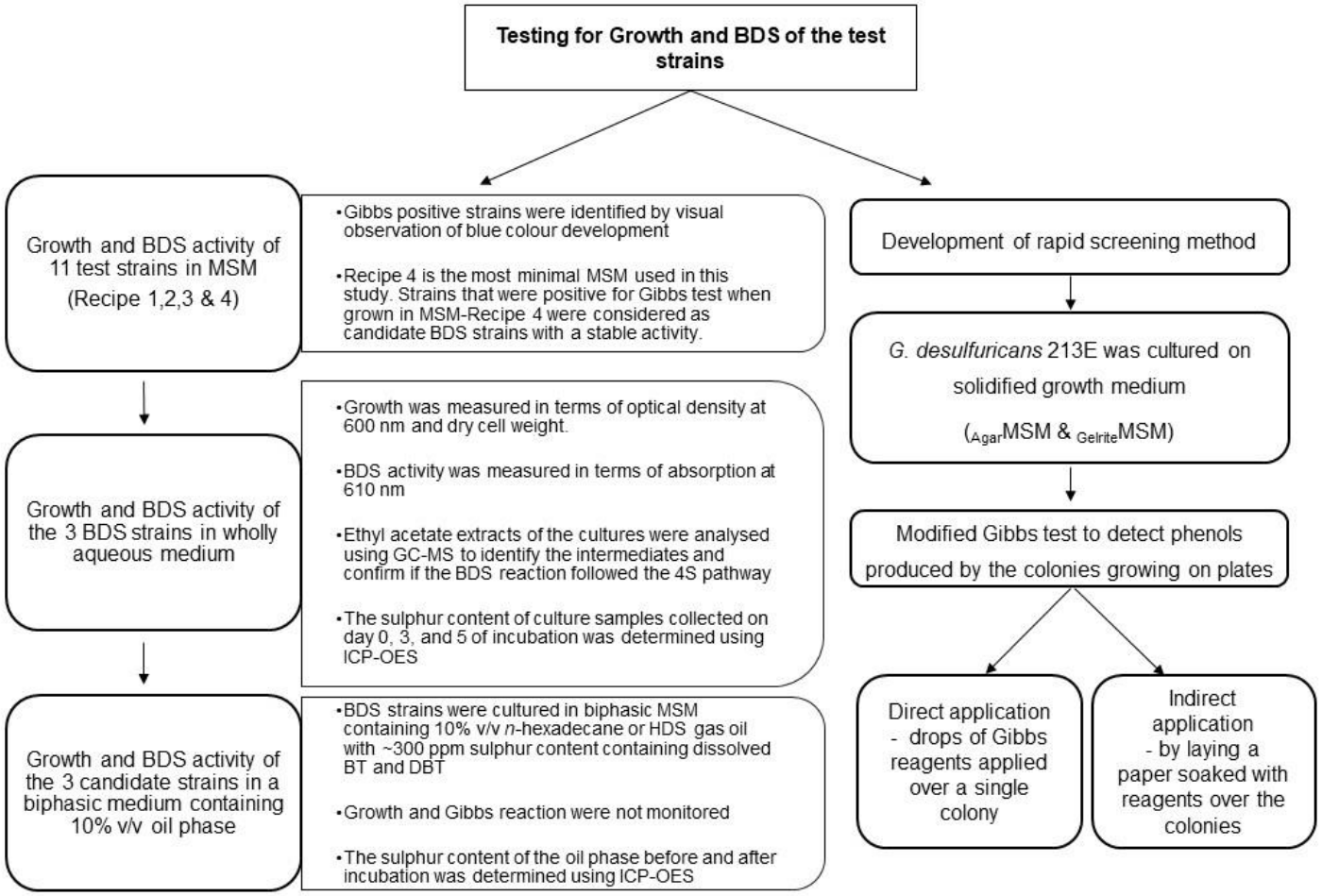


Figure 3.1 Workflow followed in this research to study the growth and BDS of test strains

3.2.2.1 Media for testing the BDS activity

The enrichment of the test strains was performed in mineral salts medium (MSM) whose composition was broadly based on Hutner's mineral medium which was originally developed for organisms like *Euglena* but subsequently used for algae such as *Chlamydomonas* and recommended for the cultivation of Gram-negative bacteria (e.g. *Pseudomonas*, *Halomonas, Ancalomicrobium*, *Ancyclobacter*, *Prosthecomicrobium*) (Hutner *et al.*, 1950, Cohen-Bazire *et al.*, 1957). The metal-sulphate salts in the Hutner's basal salts solution and "Metals 44" solution (Atlas, 2010) were replaced by respective metal-chlorides salts to make up the sulphur-free mineral salts solution (MSS). The modified MSS consisted of Na_2HPO_4, KH_2PO_4, NH_4Cl, as the primary mineral nutrients and trace levels of other mineral salts such as $MgCl_2.6\ H_2O$, $CaCl_2.2\ H_2O$, $(NH_4)_6Mo_7O_{24}.4\ H_2O$, $FeCl_3.6\ H_2O$, EDTA, $ZnCl_2$, $FeCl_3.6\ H_2O$, $MnCl_2.4\ H_2O$, $CuCl_2.2\ H_2O$, $Co(NO_3)_2.6\ H_2O$, $Na_2B_4O_7.10\ H_2O$ as additional nutrients. The sources of the chemicals used in the study are listed in Appendix 7.1 and the composition of the different solution components of the MSS is presented in Table 3.4.

Table 3.4 Composition of various components of Mineral Salts Medium

Mineral Salts Solution (MSS)

Chemical	Mol. mass (g/mol)	Quantity (per litre)	Final Molarity in the medium (mM)
Na_2HPO_4	141.96	7.2 g	50.36
KH_2PO_4	136.09	6.4 g	47.02
NH_4Cl	53.49	1.88 g	35.14
Modified Hutner's basal salts solution		10 ml	
After the addition of carbon source (Sucrose) to MSS, it was referred to as Mineral Salts Medium (MSM). Other than for culturing purposes, the MSS was used for resuspending the cell pellets and making dilutions, where needed. At this composition, the MSM had a pH 6.8-7.0 at 20 °C.			

Modified Hutner's basal salts solution

Chemical	Mol. mass (g/mol)	Quantity (per litre)	Molarity in the solution (mM)
Nitrilotriacetic acid*	191.14	10 g	52.31
$MgCl_2.6H_2O$	203.30	24.4 g	120.02
$CaCl_2.2H_2O$	147.01	3.34 g	22.71
$(NH_4)_6Mo_7O_{24}.4H_2O$	1235.85	0.009 g	0.008
$FeCl_3.6H_2O$	270.30	0.1088 g	0.4
Modified "Metals 44" solution		50 ml	

*Nitriloacetic acid was first dissolved in approximately 600ml of distilled water by heat stirring the water. After cooling down to obtain a clear solution, 7.3g KOH was added to neutralise the pH. Hutner's basal salts solution was filter sterilised (Sterile Syringe Filter 0.2 µm, VWR Europe) rather than autoclaving to avoid precipitation of the salts.

Modified "Metals 44" solution

Chemical	Mol. mass (g/mol)	Quantity (per litre)	Molarity in the solution (mM)
EDTA	292.24	2.50 g	8.55
$ZnCl_2$	136.28	5.20 g	38.15
$FeCl_3.6H_2O$	270.30	4.89 g	18.09
$MnCl_2.4H_2O$	197.91	2.00 g	10.1
$CuCl_2.2H_2O$	170.48	0.30 g	1.75
$Co(NO_3)_2.6HO$	291.03	0.25 g	0.85
$Na_2B_4O_7.10H_2O$	381.3	0.20 g	0.52

Throughout this research, sucrose was used as the carbon source. The MSM was sterilised by autoclaving at 121 °C for 20 minutes, and after cooling down to room temperature, the organosulphur compounds (BT or DBT) were added to the MSM prior to inoculation. Owing to the low solubility of BT (0.13 $g.L^{-1}$) and DBT (0.0015 $g.L^{-1}$), the stock solutions of BT and DBT were prepared by dissolving in ethanol. The BT or DBT stock solution was added to the sterile MSM to achieve final concentrations as required by the experiment, usually 0.2 mM final concentration which equated to 26.8 $mg.L^{-1}$ of BT or 36.8 $mg.L^{-1}$ of DBT. This composition of MSM used in the preliminary stages of the research is referred to as Recipe 1 in this thesis.

3.2.2.2 Growth conditions for testing the BDS capability

For BT DBT desulphurisation experiments, shake-flask cultures consisting of 50 ml of MSM containing 0.2 mM final concentration of the sulphur source in the medium was used. In this thesis, the MSM containing BT as the sole sulphur source is reported as $MSM_{BT,}$ and when DBT is the sole sulphur source, it is written as MSM_{DBT}. When the MSM contains both BT and DBT as the sulphur source, it is denoted as MSM_{BD}.

The composition of the complete MSM-Recipe 1 was altered by reduction or elimination of some ingredients to obtain recipes 2, 3 & 4. This was based on the MSM-Recipe 1 components that are not prevalently reported in then literature describing the growth media used to study BDS activity, thereby making it cheaper.

Recipe 1 – MSM based on Hutner's mineral medium

Recipe 2 – Recipe 1 without Nitriloacetic acid

Recipe 3 – Recipe 1 without Nitriloacetic acid and "Metals 44 solution"

Recipe 4 – Recipe 3 containing reduced levels of sucrose (2 $g.l^{-1}$ instead of 6 $g.l^{-1}$)

Recipe 1 is the richer medium among 4 compositions, while Recipe 4 was the most minimal medium used in the study.

The growth and desulphurisation of the test strains in each of the modified MSMs were assessed with a view to determining the strain which can exhibit growth and sustain a stable desulphurisation activity in a minimal most MSM, and to study the growth phase-dependent nature of BDS activity as observed in the case of several previously reported BDS strains.

Abiotic controls contained sterile MSM supplemented with the same amounts of the sulphur compounds. A culture of *Gordonia desulfuricans* 213E was used as a positive control.

3.2.2.3 Measurement of growth by optical density

The bacteria were incubated for a period of 7 days with periodical sampling, during which the strains with BDS capability converted the BT or DBT into phenolic compounds. About 1.5 ml aliquots of the culture samples were collected in 1.5 ml microcentrifuge tubes at periodic intervals to measure the growth and desulphurisation activity. From every sample, 200 µl of the sample was transferred to a disposable spectrophotometer cuvette with 1800µl of distilled water (10x diluted), to measure the cell growth through spectrophotometry at 600 nm wavelength (JENWAY 6300). The remaining 1.3 ml of the samples were stored tubes at -20ºC for carrying out Gibbs test on them at the end of the incubation period to measure the phenols formed at each sampling point. The growth of the bacteria was inferred from the increase in the turbidity of the culture medium ($O.D_{600nm}$). Samples were collected every 12 hours during the seven days of incubation to determine the growth and BDS activity.

In this research, growth of the 12 strains (i.e. the 11 test strains plus the positive control) during the screening stages (grown in MSM-Recipe 1 & 2) was estimated by visual observation, and not every passage was consistently quantified by spectrophotometry, because the primary interest was to identify the strains with BDS activity, and moreover, other than the 3 BDS strains, they exhibited a varying degree of non-homogenous culture characteristics after day 3. However, the BDS strains exhibited a homogenous growth in all the recipes, and their growth in MSM-Recipe 3 & 4 was measured as described above.

3.2.2.4 Measurement of growth by dry cell weight

The growth of bacteria was measured by dry cell weight (DCW) measurements using the traditional filter paper method with a view to understanding the correlation between the DCW and the turbidity measurements. Three 5 ml aliquots of the culture were taken at periodic time points during growth and filtered through pre-weighed (Fisherbrand PS-200) 0.45 µm nanopore filter membrane discs (Sartorius Stedium cellulose nitrate filter) using a manual vacuum apparatus. The filter papers loaded with bacteria were dried at 60ºC, and the decrease in weight was periodically monitored until the value was constant. The difference between the weight of the dried filter discs and its original weight was taken as the dry cell weight per 5 ml, from which the dry cell weight in grams per litre (DCW.g.L^{-1}) was calculated.

3.2.3 Gibbs test for detection of phenolic compounds in the culture

Gibbs reagent reacts with aromatic hydroxyl groups, such as phenol (and its ortho or meta-substituted derivatives), at a pH 8.0 to form a deep blue coloured indophenol complex that can be monitored spectrophotometrically at 610 nm (Kayser *et al.*, 1993). A freshly prepared solution of Gibbs reagent (1% w/v) dissolved in ethanol was used for the test. Aliquots (approx. 1.5 ml) of current MSM culture or the thawed MSM culture aliquot were centrifuged at 9400 g for 5 minutes to obtain cell-free culture supernatants (CFCS) which contain the phenolic compounds produced by BT or DBT desulphurisation. As reported previously in the literature, the phenolic compound formed is 2-hydroxybiphenyl (2HBP) in the case of DBT and 2-(2'-hydroxyphenyl)ethan-1-al or *o*-hydroxystyrene in the case of BT.

Approximately 1 ml of the CFCS was transferred to a clean cuvette, and the pH was adjusted to 8.0 - 9.0, by adding 200 µl of the 1M $NaHCO_3$ solution. Finally, 20 µl of the freshly prepared Gibbs reagent was added and mixed thoroughly by gentle vortexing. After 30 minutes of incubation in darkness at room temperature, a positive reaction between the Gibbs reagent and the phenolic compound in the transparent CFCS was viewed as the development of deep blue colour. This colour change was measured for absorbance at 610 nm, where required. Maximum accuracy was achieved in the Gibbs assay by

precisely controlled conditions of pH and time of incubation/colour development (30 minutes) (Kayser *et al.*, 1993).

In this study, 0.1 mM solution of 2-hydroxybiphenyl (2HBP) dissolved in ethanol and diluted in ultra-pure water was used as a positive control for Gibbs reaction, and uninoculated MSM that was incubated alongside the bacterial cultures was used as a negative control.

3.2.4 Preparation of resting cells for desulphurisation *n*-hexadecane

The *Rhodococcus* strain 248 was cultivated in 2L of MSM_{BT} in a 5L flask. After culturing for 5 days, the culture was concentrated by centrifugation for 5 minutes at 3026 *x g* (Eppendorf Centrifuge 5800R). A potassium phosphate buffer (0.1 M, pH 7) prepared by mixing KH_2PO_4 (39 ml of 0.2 M solution) and K_2HPO_4 (61 ml of 0.2 M solution) was used to wash and resuspend the cell pellets to a high cell density of 15 ($O.D_{600}$). Hexadecane containing BT (2 mM) was added to the cell suspension in the oil/water ratio of 1:1, 1:2 and 1:9 in triplicates.

3.2.5 Preliminary BDS studies in a biphasic medium

In order to identify the strains capable of accessing the BT or DBT dissolved in diesel oils and selectively desulphurising them, they were cultured in a biphasic MSM containing *n*-hexadecane (model oil). The MSM-Recipe 4 was prepared without the addition of ethanolic solutions of BT or DBT. It was supplemented with *n*-hexadecane containing dissolved BT (1258.1 $mg.Kg^{-1}$) or DBT (1726.8 $mg.Kg^{-1}$) which equate to 300 $mg.Kg^{-1}$ (ppm by weight) of final sulphur content in the oil phase. In this thesis, the MSM containing *n*-hexadecane is denoted as MSM_{Hex}. The strains that were previously grown in MSM_{BD} were used for inoculating the fresh MSM_{Hex}. The growth of the strains in MSM_{HexBT} and MSM_{HexDBT} was monitored. Owing to the non-homogenous nature of growth exhibited by the bacteria, their growth was not measured by turbidity measurements. However, their BDS activity was tested by Gibbs assay. The strains that exhibited BDS activity against the BT or DBT dissolved in oil were considered as potential candidate strains for the BDS of fuel oil.

3.2.6 Biodesulphurisation of BT and DBT in gas oil by growing cells

3.2.6.1 Testing the oil tolerance of the test strains

The successful candidate BDS strains identified through the preliminary biphasic medium based BDS study described above were further grown in biphasic medium with various oil/water ratios of 1%, 2%, 5%,10%, 20% and 50% (v/v), to determine the optimal ratio for the strain. The growth of the strains was negatively affected at 20% and 50% (v/v) levels of *n*-hexadecane, while at 1%, 2% and 5%, it was difficult to distinguish between the oil, biomass and aqueous phases visually. Therefore, further growth experiments using biphasic medium were carried out at 10% (v/v) oil in the aqueous MSM. The model oil was replaced with actual gas oil (diesel) obtained downstream of the HDS process at the Petroineos refinery, Grangemouth, UK, as the oil phase to make MSM_{Diesel}. The HDS gas oil which contained less than 10 ppm sulphur content was supplemented with equimolar amounts of BT and DBT to obtain ~300 ppm sulphur content in the fuel, before being used for the experiments (MSM_{D300}).

3.2.6.2 Growth conditions for BDS activity

The BDS strains were cultured in 100 ml of MSM_{BD} in 250 ml Erlenmeyer flasks at 28 °C with shaking at 180 rpm. After 3 days of growth, their BDS activity was confirmed by Gibbs assay. The cells were pelleted by centrifugation and washed once by resuspending them in 5 ml of sterile MSS to remove any carryover of nutrients or metabolites. The washed cells were used to inoculate 1 litre of MSM_{Hex} containing 10% (v/v) *n*-hexadecane with 300 ppm Sulphur content. In this medium, the only source of sulphur is the dissolved BT or DBT present in the oil phase. The cultures were maintained in triplicates alongside another triplicate set of entirely aqueous MSM supplemented with 0.2 mM BT or DBT (ethanolic solutions) for comparative studies between the oil-free aqueous and biphasic MSMs. In the case of strain 248 and *Gordonia desulfuricans* 213E, MSM_{BT} and MSM_{HexBT} were used for the growth and BDS studies. In the case of strain F, MSM_{DBT} and MSM_{HexDBT} were used. The growth of the BDS strains was measured by turbidimetry and dry cell weight methods in the case of Aqueous MSM. Owing to the non-homogenous nature

of growth in the presence of oil, the growth of the strains in the case of MSM_{Hex} was measured by weight methods only.

3.2.7 Determination of the metabolic intermediates by GC-MS

Gas chromatography-mass spectrometry (GC-MS) was conducted to elucidate the molecular identity of metabolites of BT and DBT desulphurisation occurring in the MSM_{BT} and MSM_{DBT} cultures of the BDS strains. The richer MSM (Recipe 2) was used in this study to allow for the fullest expression of BDS activity.

3.2.7.1 Sample preparation for GC-MS

The BDS strains 248, F and *G. desulfuricans* 213E were cultured in MSM_{BT}, MSM_{DBT} and MSM_{BD}. The cell-free culture supernatant (CFCS) of the 5-day old culture was tested to confirm the presence (formation) of phenol by the Gibbs test. Gibbs positive cultures were selected for further GC-MS analysis to identify the phenol and other metabolic intermediates formed as a result of biodesulphurisation of BT and DBT by the bacteria. Day 0 cultures were used as controls to confirm the presence of only BT or DBT and the absence of phenols at the start of the cultures and to identify the retention time of BT and DBT.

About 25 ml of the bacterial culture was taken in 50 ml falcon tubes and acidified by gradually adding 6 *N* HCl, up to pH 2.0 at 20 °C. To the acidified culture, an equal volume of HPLC grade ethyl acetate (Sigma-Aldrich, USA) was added and shaken on a Vortex-Genie (Scientific Industries Inc., USA) for thorough mixing. After allowing to fractionate into layers for 5 minutes, the ethyl acetate layer was extracted and used for further GC-MS analysis. Ethyl acetate extract of MSM without BT or DBT was used to check for background peaks in the chromatogram caused by the solvents used in the procedure, and the ethyl acetate solution of 1 mM 2HBP was used to generate reference chromatogram for identification of its retention time.

3.2.7.2 Analytical conditions for qualitative GC-MS analysis of the ethyl acetate extracts

The qualitative analysis of the ethyl acetate extracts was performed using GC-MS scan mode. The analysis was performed by Thermo Scientific Trace Ultra™ GC (equipped with a split/splitless injector connected with a Thermo Scientific DSQ quadruple mass spectrometer). The metabolites were separated on a 30 m long x 0.25 mm I.D. ZB-SemiVolatiles column (Phenomenex, USA) with 0.25 μm film thickness. Each sample was injected at 1 μl using splitless mode. The temperature conditions were set by taking into consideration the physical properties of ethyl acetate solvent. The inlet temperature was set at 280 °C. The carrier gas (helium) was set at 2 $ml.min^{-1}$ throughout the run. The temperature was 60 °C with 0.2 minutes isothermal, then increased at the rate of 10 °C to 325 °C hold for 5 minutes. The total run time was 31.5 minutes. The mass spectrometer was set for the full scan from mass 50 to 550 amu at the scan rate of 500 amu s^{-1}. Transfer line and ion source temperatures were set at 325 and 230 °C, respectively.

The chromatograms obtained were scanned manually for the peaks of interest and the corresponding mass spectra corresponding to the peaks were searched against the NIST MS Search database (version 2.3) to identify the intermediates of 4S pathway such as Dibenzothiophene-5-oxide, Dibenzothiophene-5,5-dioxide (Dibenzothiophene sulphone) and the final phenolic end product 2-hydroxybiphenyl (2HBP), whose mass spectral data are presented in Appendix 7.6. The peaks conserved in all the chromatograms obtained for the day 0 samples and the day 5 samples were eliminated.

3.2.8 Determination of sulphur content in aqueous samples by ICP-OES

Since the BT and DBT have a very low solubility in water, as the bacteria grew in the MSM, they metabolised the BT and DBT into more soluble and hence assimilable forms of sulphur. The amount of dissolved sulphur present in the culture system at various stages of bacterial growth was measured with a view to understanding the quantity of sulphur remaining in the culture medium after assimilation by the bacteria. The BDS strains were grown in MSM_{BT} (for strain 248) and MSM_{DBT} (for strain F) as the sole sulphur source. The CFCS was

collected on days 0, 3, 5 and 7 of incubation, and filtered through 0.45 µm syringe filters (Millex) before being used for the determination of the dissolved concentrations of sulphur by ICP-OES (Thermo Scientific iCAP 6200 Duo View ICP Spectrometer, Thermo Fisher Scientific, Cambridge, UK). The aqueous samples were analysed directly using standard sample injection system consisted of Mira Mist nebuliser and glass cyclonic mixing chamber. Instrument operation conditions and emission intensity used for the measurement are given in Appendix 7.7. Thermo iTEVA Analyst software (version 2.4.0.81) was used for instrument operation and for data handling and processing. The wavelengths for minimum interference were 180.73 and 182.03 nm (Amais *et al.*, 2014; Kirkbright *et al.*, 1972; Yang and Marison, 2005). Sulphur calibration standards of 0.01, 0.1, 1 and 10 ppm were prepared from a Sulphur standard solution of 1000 ppm concentration (ThermoFisher Scientific, Cambridge, UK), and the dilutions were prepared in ultra-pure de-ionised water (resistivity of 18.2 MΩ cm, Direct Q3 Millipore Water Purifier, Millipore, USA). A solution of 5 ppm Yittrium in 0.5% nitric acid was used as the internal standard. The 0.2 and 2 ppm analytical quality control (AQC) solutions were used to check for elemental recovery and drift in the measurement during analysis. The check using AQCs was done at the beginning of the sequence run, every 10 samples thereafter and at the end of the experiment. Analytical precision (RSD) was typically 1-5% (n=3). The instrument was calibrated, and the spiked blank was analysed, and the recovery calculated. A detection limit study was carried out by analysing the calibration blank with ten replicates and multiplying the standard deviation of this analysis by three. This was repeated three times, and the average values for sulphur detection limits were calculated as 0.06 ppm.

3.2.9 Determination of sulphur content in oil samples by ICP-OES

3.2.9.1 Preparation of oil samples for the ICP-OES

For the analysis of diesel oil and hexadecane samples, the oil phase of the bi-phasic culture system was extracted and subjected to ICP-OES. As there was no ASTM published for the measurement of sulphur content of oil by ICP-OES, the methodology recommended by the manufacturer (Beauvir, 2016) for lubricating oil elemental analysis was followed with modification. All solvents

and standards were purchased from Conostan® (SCP SCIENCE, Baie-D'Urfé, Canada). Silicon was used as the internal standard, and PremiSolv™ was used as the base solvent. Unlike the aqueous samples, the internal standard was not introduced into the system through a separate line in the case of oil sample analysis. The internal standard (Si) was added to the base solvent to obtain a 10 mg.L^{-1} solution used for all further dilutions.

The organic samples to be analysed included *n*-hexadecane and diesel oils in their pure form and as recovered from the bacterial cultures. Therefore, to minimise the differences in viscosity, the final solution was ensured to always contain 10% oil, by adding CONOSTAN 75cSt Blank Oil (which has a kinematic viscosity of 75 x 10^{-6} m_2 s^{-1}) to the sample and standards prior to the addition of the diluent. The oil samples were always mixed with the diluent containing the internal standard (Silicon) and blank oil.

3.2.9.2 Measurement of sulphur content in oil samples by ICP-OES

The sample introduction system in the ICP-OES equipment was fitted with Duo Volatile Organic Sample Introduction Kit (Thermo Fisher Scientific, UK) and the operational settings were changed for oil analysis as described in Appendix 7.7. Sulphur calibration standards of 3, 5 and 50 ppm were prepared from a stock sulphur standard solution of 5000 ppm concentration. A 10 ppm analytical quality control (AQC) solution was used to check for elemental recovery and drift in the measurement during analysis. The samples were introduced manually without an autosampler. During the analysis, the system was rinsed using the diluent between every sample to flush out any remnant sulphur from the system. The check using AQCs was done at the beginning of the sequence run, at the end of the experiment samples thereafter and at the end of the experiment. Analytical precision (RSD) was typically 1-5% (n=3). The instrument was calibrated, and the spiked blank was analysed, the recovery calculated. A detection limit study was carried out by analysing the calibration blank with ten replicates and multiplying the standard deviation of this analysis by three. This was repeated three times, and the average detection limit was calculated as 0.13 ppm of sulphur.

3.2.10 Modified Gibbs test on solidified medium

In order to develop a method of detecting phenolic compounds produced by colonies, the MSM (Recipe 2) was solidified using agar (1%, 1.5%, 2% w/v) as the gelling agent. Although solidification was attempted initially using agars of different quality grades, Noble Agar was chosen to ensure most minimal sulphur contaminants. After autoclaving, the molten $_{Agar}$MSM was poured aseptically into 90 mm sterile disposable Petri dishes and allowed to cool and solidify for about 1 hour. The surface of the solidified MSM was coated with 200 µl of BT or DBT (100 mM stock solutions BT and DBT) using a sterile L-rod. After allowing to sit for 30 minutes, a loopful of washed cells were streaked onto the $_{Agar}$MSM plates. Alternatively, MSM was solidified using Gelrite™ (0.75%), which is a gellan gum-based solidifying agent to obtain clear gels rather than the brownish coloured gels obtained using Agar. The above procedure was repeated using the $_{Gelrite}$MSM plates.

After 4 days of incubation, the plates were sprayed with sodium bicarbonate solution and Gibbs reagent was introduced to the colonies of *G. desulfuricans* 213E on the plates either directly (2 µl drops on the colonies). Alternatively, filter paper (6 cm diameter) soaked with 200 µl 1M sodium bicarbonate solution and 20 µl Gibbs reagent was laid over the colonies. Any colour development occurring around the colonies over a period of 12 hours was noted.

3.3 Results

3.3.1 Identification of the stable BDS activity of strain 248 and strain F

The 11 rhodococci test strains and the control strain *G. desulfuricans* 213E were grown in MSM-Recipe 1 containing either BT or DBT as the sole sulphur source. Increase in biomass was observed in all the cultures, with the strains growing as a homogenous cell suspension with various degrees of growth. As the primary objective of the study was to identify strains that exhibited BDS activity, the growth of the test strains was not monitored periodically, and only the BDS activity of the test strains was checked on days 3 and 7 of incubation by Gibbs assay. The strains with BDS capability (BDS strains) converted the BT and DBT into a phenolic product that was detected using the Gibbs reagent, where the presence of phenol (positive reaction) was indicated by the

development of a blue colour. The preliminary screening study using MSM-Recipe 1 showed that the CFCS obtained from the MSM_{BT} culture of strain 248, and the MSM_{DBT} cultures of strains 248, 488, 508 and strain F exhibited the characteristic blue colour development in the Gibbs assay, like that of the known BDS strain *G. desulfuricans* 213E, which indicated their BDS capability and the intensity of the blue colour was assessed visually. The intensity of the blue colouration was interpreted as higher amounts of phenols produced, indicating an intense BDS activity.

Exclusion of the nitriloacetic acid from MSM-Recipe 2 did not change the growth and BDS characteristics of the test strains. However, when the 'Metals 44' solution, was removed from the medium composition, the growth of strains 20, 60, 87, 213, 369, 505 and 639 was affected, whereas the growth of strains 208, 248, 488, 508, strain F and *G. desulfuricans* 213E was not affected, indicating that 5 of the test strains could thrive in the minimal MSM-Recipe 3 (data not shown). However, changes were observed in their BDS characteristics when repeatedly subcultured in MSM-Recipe 3, such that the DBT desulphurisation activities of strains 488 and 508 were unstable, as indicated by the less intense blue colour development in subsequent passages and three passages thereafter, the BDS activity (blue colour development) was not observed.

In the case of strain 248, repeated subculturing in minimal MSM resulted in the loss of BDS activity against DBT, whereas in the case of strain F, it resulted in the loss of BDS activity against BT. However, the strain 248 and strain F exhibited stable desulphurisation activity against BT and DBT, respectively, even in the minimal MSM, indicating they are predominantly BT and DBT desulphurising strains, respectively. The control strain *G. desulfuricans* 213E, which is a known BT desulphurising organism, followed the same BDS trend as strain 248 when grown in minimal MSM-Recipe 3.

Based on the sustainment of BDS activity in a minimal MSM, strains 248, F and *G. desulfuricans* 213E were selected as candidate BDS strains. In the case of MSM_{BD} cultures, a positive Gibbs reaction was obtained for all the 3 BDS strains. But as the medium contained both BT and DBT, it could not be ascertained at this stage as to what extent the respective phenols produced

from the desulphurisation of BT and DBT contributed to the blue colour formation. However, this indicated that the presence of a less preferred organosulphur compound in the medium did not affect the growth of the cells.

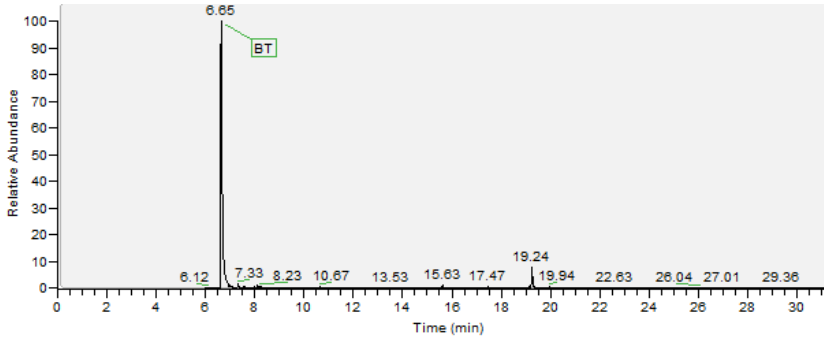
3.3.2 Elucidation of the biodesulphurisation pathway of the test strains

The ethyl acetate extracts of the CFCS obtained from the 5-day old cultures of strains 248, F and *G. desulfuricans* 213E grown in MSM_{BT}, MSM_{DBT} and MSM_{BD} (Recipe 1) were subjected to GC-MS analysis to study the metabolic intermediates and the phenolic end products formed as the results of BT DBT desulphurisation. The 5 days of incubation allows significant utilisation of BT or DBT, and hence substantial metabolic intermediates and end product (phenol) formation. According to the established 4S pathway, the phenolic end product of DBT desulphurisation is 2-hydroxybiphenyl (2HBP) with Dibenzothiophene-5,5-dioxide (Dibenzothiophene sulphone) as one of the intermediates.

In the gas chromatography carried out in this study, BT and DBT were observed at the retention times of 6.65 and 13.62 minutes, respectively as observed in (Figure 3.2 A), (Figure 3.3 A), and (Figure 3.4 A). The significant conserved peaks were observed in all the MSM_{BT} samples at retention times 7.33, 15.63, and 19.24 minutes referred to non-relevant compounds. The peak corresponding to BT was observed as the predominant peak only in the case of uninoculated MSM_{BT} (Figure 3.2 A), whereas in all other chromatograms (Figure 3.2 B, C &D), this peak was very much diminished indicating BT metabolisation by the bacteria. Interestingly, in the case of strain 248 (Figure 3.2 B) and *G. desulfuricans* 213E (Figure 3.2 D), a peak was observed at 11.98 minutes, whose mass spectra corresponded to benzofuran which is a metabolic by-product of BT desulphurisation by the 4S pathway (Gilbert *et al.*, 1998; Kirimura *et al.*, 2002). It has been suggested that 2-Hydroxyphenylacetaldehyde was recovered as benzofuran due to the extreme dehydration caused by the addition of acid while preparing the cultures for ethyl acetate extraction (Kirkwood *et al.*, 2007a). However, the peak corresponding to benzofuran was not observed in the case of strain F (Figure 3.2B). This concurred with the previous observation that strain F was predominantly a DBT desulphurising strain.

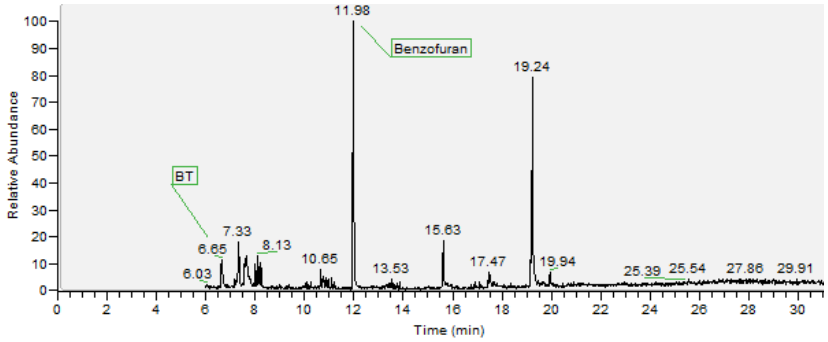
A)

Uninoculated MSM$_{BT}$



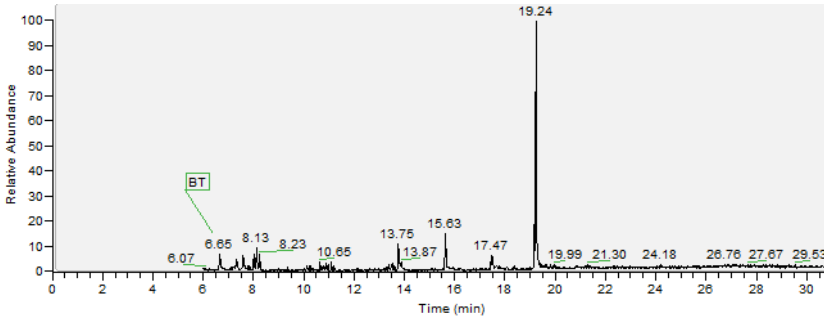
B)

Strain 248 in MSM$_{BT}$



C)

Strain F in MSM$_{BT}$



D)

G. desulfuricans 213E in MSM$_{BT}$

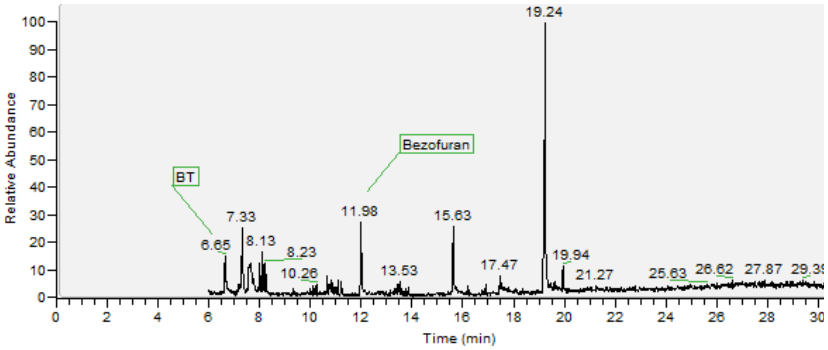


Figure 3.2 GC-MS profile of MSM$_{BT}$ cultures after 5 days of incubation

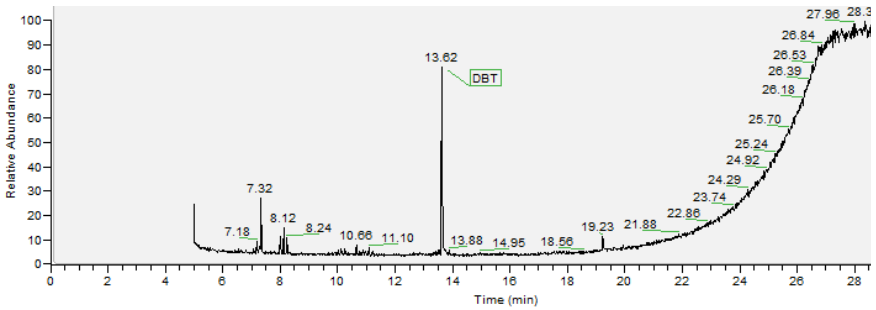
The chromatograms obtained by the gas chromatography of the ethyl acetate extracts of MSM$_{BT}$ cultures of strains 248, F, *G. desulfuricans* 213E and uninoculated (control) after 5 days of incubation is presented where the peak observed at the retention time of 6.65 minutes corresponds to benzothiophene (BT). The uninoculated sample (Fig 3.2 A) has a significant BT peak. In the case of strain 248 (Fig 3.2B) and 213E (Fig 3.2D), the end product benzofuran was detected at 11.98 minutes. The GC profile of strain F (Fig 3.2C) does not reveal any known end product of BT BDS pathway by scanned chromatograms, indicating that it is a DBT desulphurising strain.

In the case of MSM_{DBT} cultures, the peaks corresponding to DBT, the intermediate DBT-sulphone and the final phenolic compound 2HBP were obtained at a retention time of 13.62, 16.19 and 10.82 minutes, respectively for all culture samples as presented in Figure 3.3 A, B, C & D. These results confirmed that the BDS activity of strains 248, F and 213E followed the established 4S pathway of DBT desulphurisation.

Gordonia desulfuricans 213E was initially published as a BT desulphurising bacteria. In this research, the DBT desulphurising activity of strain 213E was evidenced through Gibbs assay and the GC-MS chromatogram which shows the peak for 2HBP (Figure 3.3 D). However, along with 2HBP, significant levels of DBT were also detected (Figure 3.3 D). This meant that the conversion was not so significant as strain 248 or strain F, where relatively lower levels of DBT remained as indicated by the insignificant peak at RT 13.62 minutes (Figure 3.3 B & C). This result, together with the benzofuran detected in the MSM_{BT} cultures of strain 213E (Figure 3.2 D) concurs with the original description of *G. desulfuricans* 213E by Gilbert *et al.* (1998) as a BT desulphurising strain.

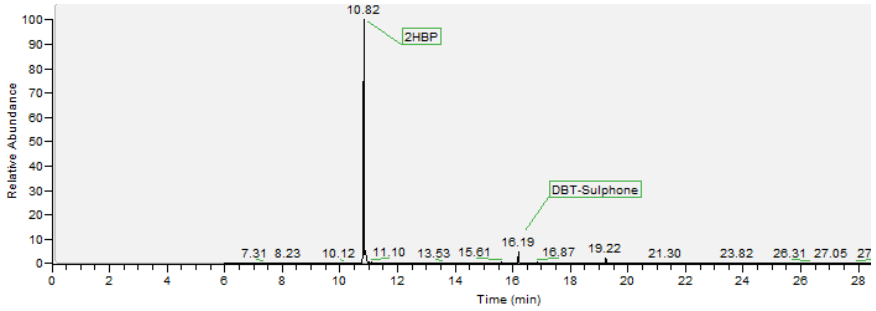
A)

Uninoculated MSM$_{DBT}$



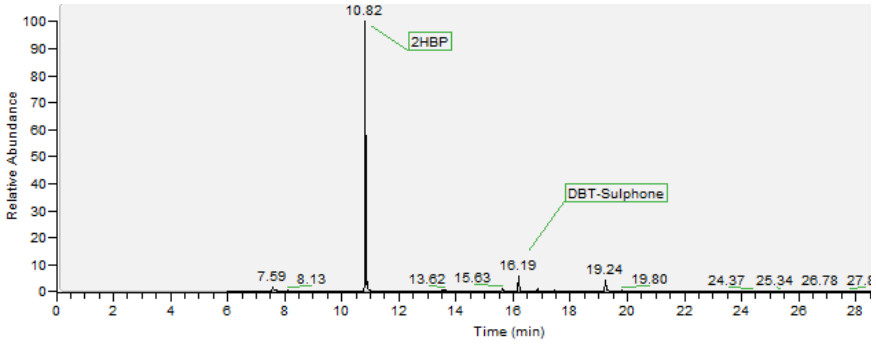
B)

Strain 248 in MSM$_{DBT}$



C)

Strain F in MSM$_{DBT}$



D)

G. desulfuricans 213E in MSM$_{DBT}$

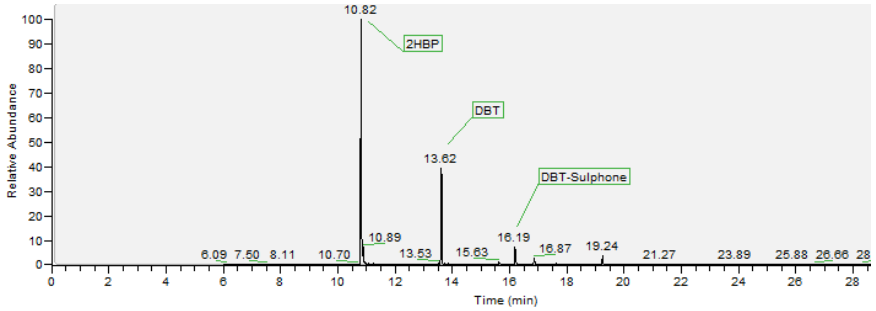


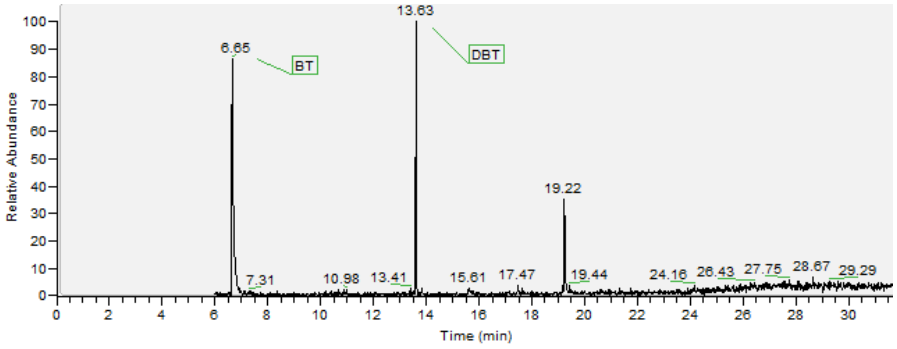
Figure 3.3 GC-MS profile of MSM$_{DBT}$ cultures after 5 days of incubation

The chromatograms obtained by the gas chromatography of the ethyl acetate extracts of MSM$_{DBT}$ cultures of strains 248, F, *G. desulfuricans* 213E and uninoculated (control) after 5 days of incubation is presented where the peak observed at the retention time of 13.62 minutes corresponds to dibenzothiophene (DBT). The uninoculated sample (Fig 3.3A) has a significant DBT peak. In the case of strain 248 (Fig 3.3B), strain F (Fig 3.3C) and 213E (Fig 3.3D), the end product 2-hydroxybiphenyl (2HBP) and the intermediate DBT-sulphone were detected at 10.82 and 16.19 minutes, respectively, indicating that these organisms followed the 4S pathway of DBT desulphurisation.

Using Gibbs assay, it was known that the 5-day old MSM_{BD} cultures of strains 248, F and 213E were all Gibbs positive, but it was unknown as to what phenolic end product in the culture caused the positive reaction (blue colour formation), i.e. if it was the 2HBP formed as result of DBT desulphurisation or another phenolic end product formed as a result of BT desulphurisation or a combination of both. The MSM_{BD} cultures were started with same levels of BT and DBT, and the GC chromatogram of uninoculated MSM_{BD} collected at the end of the 5 days of incubation, confirmed the presence of both thiophenes as shown in Figure 3.4 A. The strain F exhibited efficient metabolism of DBT into 2HBP, which is indicated by the diminished DBT peak at RT 13.62 minutes and presence of a significant 2HBP peak at RT 10.82 minutes (Figure 3.4 C). In the case of strain F MSM_{BD} cultures, the BT peak at RT 6.65 minutes was also diminished, but the corresponding phenolic compound could not be detected. Both BT and DBT were utilised by strain 248 (Figure 3.4 B) and *G. desulfuricans* 213E (Figure 3.4 D), but not entirely. The phenolic end product of DBT desulphurisation (2HBP) was detected at RT 13.63 minutes, but the phenolic end product of BT desulphurisation (benzofuran) which was detected in the MSM_{BT} cultures of these strains (Figure 3.2 B & D) was not found in the MSM_{BD} cultures. However, the diminished BT peak indicated their preference for BT over DBT when the thiophenes were presented together. Interestingly, MSM_{BD} cultures of strain 248 and 213E showed a small peak at RT 29.77 minutes (Figure 3.4 B & D) which corresponds to dibenzyl ketoxime, which is a phenolic compound (Smith, 2003). Also, the strain 213E presented a peak at RT 16.85 minutes (Figure 3.4 D) that corresponded to dibenzofuran possibly produced by the activity of the enzymes that produced the benzofuran when MSM_{BT} was used. The differences in the GC-MS profile of strain 248 and 213E reveal the differences in their metabolism.

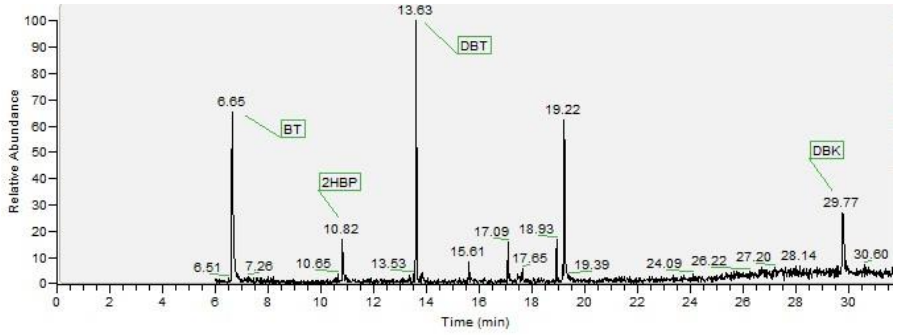
A)

Uninoculated MSM$_{BD}$



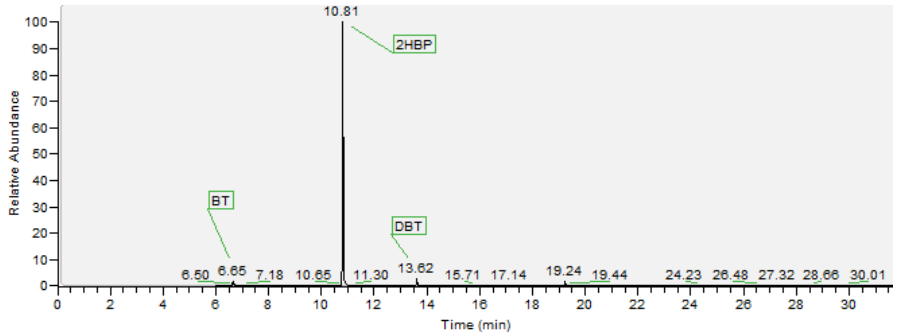
B)

Strain 248 in MSM$_{BD}$



C)

Strain F in MSM$_{BD}$



D)

*G. desulfurican*s 213E in MSM$_{BD}$

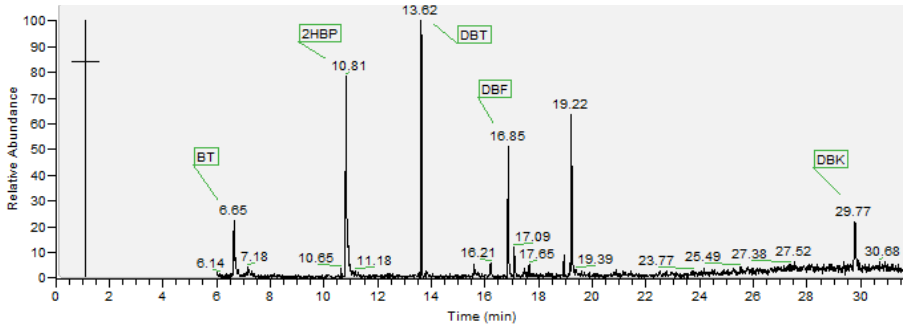


Figure 3.4 GC-MS profile of MSM$_{BD}$ cultures after 5 days of incubation

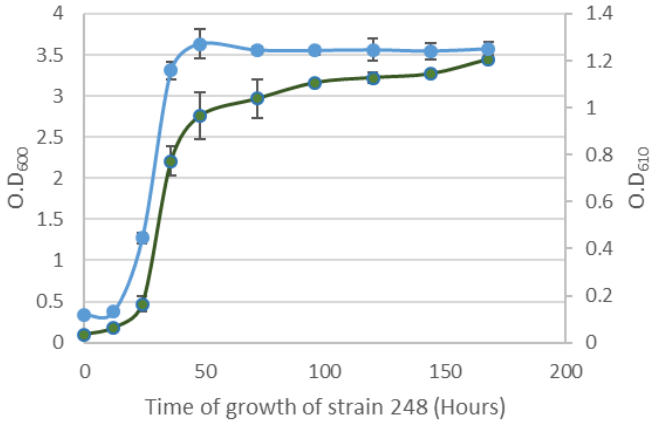
The chromatograms obtained by the gas chromatography of the ethyl acetate extracts of MSM$_{BD}$ cultures of strains 248, F, *G. desulfuricans* 213E and uninoculated (control) after 5 days of incubation is presented where the peak observed at the retention time (RT) of 6.65 and 13.62 minutes corresponds to benzothiophene (BT) and dibenzothiophene (DBT), respectively. The uninoculated sample (Fig 3.4A) has a significant BT & DBT peaks. In the case of strain 248 (Fig 3.4B), strain F (Fig 3.4D) and 213E (Fig 3.4D), the end product 2-hydroxybiphenyl (2HBP) was detected at 10.82, indicating that these organisms followed the 4S pathway of DBT desulphurisation. A peak corresponding to dibenzyl ketoxime (DBK) was observed at RT 29.77 minutes in the MSM$_{BD}$ cultures of 248 and 213E, while the latter strain also showed a peak at RT 16.85 minutes corresponding to dibenzofuran (DBF)

The GC-MS results confirmed the presence of 2HBP in the cultures when DBT was used as a sulphur source in the medium (MSM_{DBT} and MSM_{BD}) which indicates the strains 248, F and 213E exhibited the 4S pathway of DBT desulphurisation. The end product of BT desulphurisation is not definitely known, and it has been reported to vary between BDS organisms. When grown in MSM_{BT}, benzofuran was detected in the case of strain 248 and 213E, but no peaks corresponding to any known phenolic end product were detected in the case of strain F.

3.3.3 Growth and biodesulphurisation of BT by strains 248 and F growing in aqueous MSM_{BT}

When cultured in MSM_{BT} the strain 248 exhibited growth as presented in Figure 3.5 A, where after a short initial lag phase for 24 hours, the bacteria exhibited exponential growth, reaching a peak of 2.7 $O.D_{600nm}$ just at the end of the second day of incubation. The growth continued for the next 4 days of incubation, although at a slower rate. The turbidity of the culture was homogeneous throughout the period of incubation, as shown in Figure 3.6. The sulphur level of pure MSM_{BT} medium was 0.63 ppm. Upon incubation, the levels of dissolved sulphur initially raised to 2.52 ppm on day 2 and recorded 3.37 ppm at the end of 7 days of incubation, as shown in Figure 3.8. When cultured in MSM_{BT}, *G. desulfuricans* 213E exhibited a growth trend as shown in Figure 3.5 B, where the bacteria exhibited an exponential growth after a short initial lag phase for 24 hours, which reached a peak 3.5 $O.D_{600nm}$ just after the second day of incubation, 56 hours approximately. The growth plateaued beyond that point and remained at the stationary phase for the next 2 days of incubation, after which the turbidity declined. In both strains, the curve for the Gibbs reaction attained the peak levels just after the late exponential phase. The gradual increase in the sulphur levels of the aqueous medium with the time of incubation, as shown in Figure 3.8, is because of the accumulation of the sulphates released into the medium as a result of desulphurisation of BT by strains 248 and 213E. The poorly soluble BT is converted into soluble sulphates that are released into the aqueous medium.

A)



B)

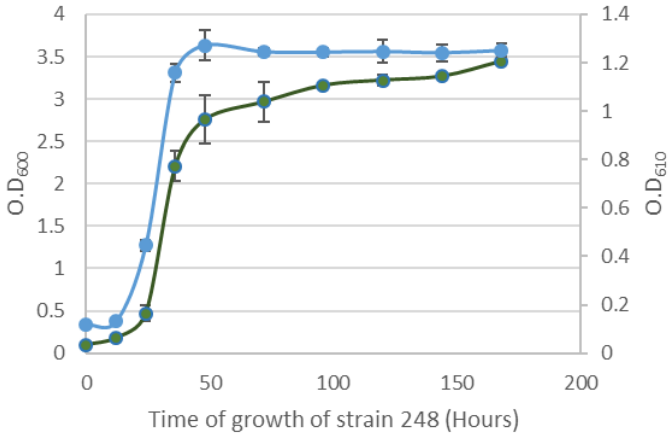


Figure 3.5 Growth and BDS activity of strain 248 and *Gordonia desulfuricans* 213E grown in MSM$_{BT}$

The figure shows the growth (green line) and desulphurisation activity (blue line) of strain 248 **(A)** and *Gordonia desulfuricans* 213E **(B)** grown in MSM with BT as a sole sulphur source. The MSM was supplied with an ethanolic solution of BT as the sole sulphur source (0.2 mM final concentration) and inoculated with 0.1 O.D$_{600nm}$ of bacteria at the beginning of the culture. Samples were collected every 12 hours, and the turbidity of the culture owing to biomass increase was measured by spectrophotometry. The BDS activity was measured by Gibbs reaction, and the intensity of the blue colouration was measured by spectrophotometry at O.D$_{610nm}$. Results were reported as mean calculated from triplicates experiments with ± standard deviation from the mean indicated as errors bars.

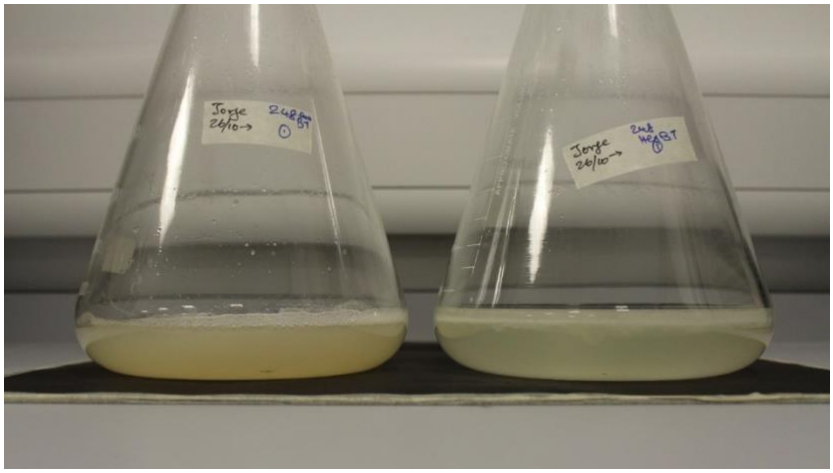


Figure 3.6 Comparison of culture characteristics of strain 248 growing in aqueous MSM and biphasic growth medium
The image shows the homogenous nature of the culture of 248 when grown in a wholly aqueous MSM observed as a turbid culture (left flask), and the non-homogenous dispersion of cells in the presence of oil (*n*-hexadecane) observed as a biomass rich layer and less dense aqueous layer (right flask). The clumped nature of the growth in the biphasic medium resulted in inconsistent sampling, and hence, the growth could not be measured using optical density or dry weight measurements.

3.3.4 Growth and biodesulphurisation of DBT by strain F growing in aqueous MSM_{DBT}

When cultured in MSM_{DBT}, strain F exhibited a growth as presented in Figure 3.7, where the strain showed an exponential growth after a short initial lag phase for 24 hours, which reached a peak 3.5 $O.D600_{nm}$ just at the end of the second day of incubation, followed by a short stationary phase for 12 hours. The strain F exhibited a further exponential phase of growth for the next 4 days of incubation. Whilst the primary rapid growth phase was accompanied by an increase in the accumulation of phenolic product (in this case, the phenolic product is 2HBP), the secondary growth phase saw a decrease in the concentration of the phenolic product.

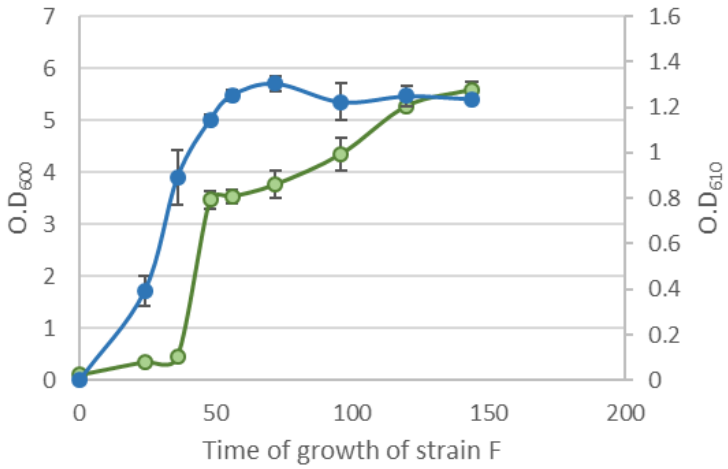


Figure 3.7 Growth and BDS activity of strain F grown in wholly aqueous MSM_{DBT}

The figure shows the growth (green line) and desulphurisation activity (blue line) of strain F grown in MSM with DBT as a sole sulphur source. The MSM was supplied with an ethanolic solution of DBT as the sole sulphur source (0.2 mM final concentration) and inoculated with 0.1 $O.D_{600nm}$ of bacteria at the beginning of the culture. The values represented the mean of three replicates, and the standard deviations were shown as error bars. The BDS activity was measured by Gibbs reaction, and the intensity of the blue colouration was measured by spectrophotometry at $O.D_{610nm}$. Results were reported as mean calculated from triplicates experiments with ± standard deviation from the mean indicated as errors bars.

At the start of the culture, the sulphur content of the MSM_{BT} and MSM_{DBT} were measured as 0.63 ppm and 4.53 ppm, respectively. These values are lower than the expected value of 26 ppm (corresponding to 0.2 mM). A potential reason for the low initial reading could be because of the poor solubility and non-homogeneous dispersion of BT and DBT in the aqueous medium. However, as the biodesulphurisation occurred during incubation, the sulphates were produced, which dissolved and dispersed homogeneously in the aqueous environment, and the accumulation of soluble sulphates was observed, as shown in Figure 3.8.

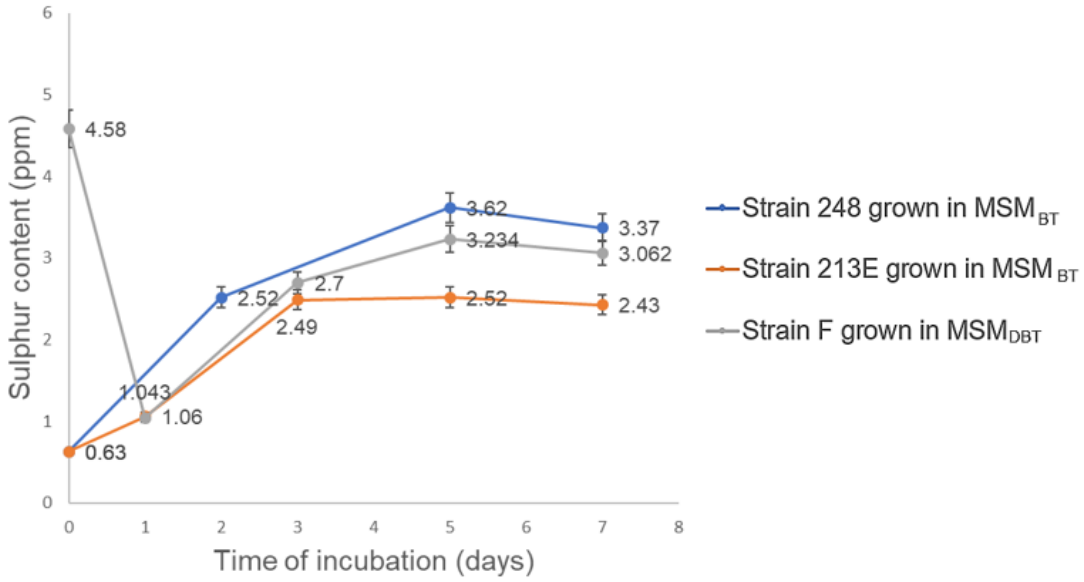


Figure 3.8 Sulphur content of the aqueous MSM cultures of the BDS strains

The graph shows the sulphur levels measured in the CFCS of MSM cultures sampled at various stage of incubation, by ICP-OES technique. BT and DBT were supplied as ethanolic solutions at a starting concentration of 0.2mM (26 ppm) of BT or DBT, which corresponds to 6 ppm sulphur content. The variation in the levels on Day 0 (start of culture) is because of the difference in the level of solubility, volatility and density of BT and DBT. In the case of MSM_{BT} cultures, immediately after addition, the BT moved to the surface level of the aqueous MSM, which resulted in little BT getting picked during sampling and hence a lower initial value. In the case of MSM_{DBT} culture, immediately after addition, the DBT was dispersed throughout the aqueous MSM, which resulted in the realistic sampling and hence an appropriate sulphur content (4.58 ppm) value was measured. The BT and DBT were adsorbed to the cell surface on Day 1, leading to a decrease in the measured sulphur levels. The sulphates produced as a result of BDS reaction are released into the aqueous MSM subsequent days of incubation, as reflected in the increase in measured sulphur content.

3.3.5 BDS activity of growing cells in biphasic medium containing *n*-hexadecane & diesel

When grown in a biphasic medium, all the strains exhibited a non-homogenous culture, as shown in Figure 3.9, which made the sampling inconsistent. So, their growth could not be tracked by optical density measurements or dry cell weight measures. The HDS gas oil (diesel) was reconstituted with added BT and DBT to make MSM_{Diesel}. When dealing with solvents like *n*-hexadecane and gas oil, volatile substances like BT and DBT, inaccuracies are prone to occur. The MSM_{D300} was prepared to contain 300 ppm of sulphur, but its sulphur content measured as 292 ppm by ICP-OES. The shaking effect during incubation caused a natural loss of these volatile compounds, as indicated by the sulphur content of uninoculated systems which measured a slightly lower

273 ppm after 7 days of incubation. Most of this decrease was caused by the loss of BT (melting point 32°C), rather than DBT (melting point 100°C). The reduction in the sulphur content of the oil phase was measured using ICP-OES as presented in Table 3.5, where it is seen that all the strains offered a significant (> 80%) reduction in the sulphur levels over 7 days of incubation. However, they differed in the amount of diesel recoverable after incubation owing to the different levels of emulsification observed in the biphasic culture as presented in Figure 3.9, where it can be seen that 213E culture is completely emulsified (Figure 3.9 D). Upon allowing to stand and fractionate into phases naturally, the MSM_{D300} culture of strain 248 formed distinct aqueous, biomass and oil phases as shown in Figure 3.10, from which about 85% of the original volume of oil phase (desulphurised diesel) was recovered from the culture. The MSM_{D300} culture of strain F formed an emulsified biomass-oil layer and a relatively clear aqueous layer (Figure 3.10). By centrifugation of the emulsified layer, about 50% of the original volume of oil-phase was recovered. The sulphur content of the recovered desulphurised diesel phases was measured by ICP-OES. When grown in MSM_{D300}, strains 248 and F achieved 84.5% and 87.2% reduction in sulphur levels, respectively, as presented in Table 3.5.

Table 3.5 Biodesulphurisation of oil by growing cells of strain 248 and F

Sample	Total sulphur content (ppm)		Reduction (%)
	Day 0	Day 7	
Uninoculated MSM_{D300}	292	273	6.5
248 grown in MSM_{D300}	296	45.86	84.5
248 grown in MSM_{HexBT}	355.4	38.8	89
Strain F grown in MSM_{D300}	293	37.4	87.2

The table shows the decrease in the sulphur content of the oil phase achieved by strains 248 and F over 7 days of incubation in the respective biphasic medium. The slight decrease observed in the uninoculated sample shows the natural loss due to volatility of the thiophenes. The BDS activity of *G. desulfuricans* 213E was not quantified because the oil phase was not recoverable. The 89% BDS activity of 248 grown in MSM_{HexBT} indicates that it is a robust BT desulphurising bacteria.

a) MSMD300 incubated with No Bacteria (Control)

b) Strain 248 growing in MSM_{D300}

c) Strain F growing in MSM_{D300}

d) Strain 213E growing in MSM_{D300}

Figure 3.9 Strains 248, 213E and F growing in a biphasic medium containing Diesel

The image shows cultures of strain 248 **(b)**, F **(c)** and 213E **(d)** growing in a biphasic medium containing 10% (v/v) HDS gas oil with ~300ppm sulphur content (MSM_{D300}). The uninoculated flask **(a)** was used as control.

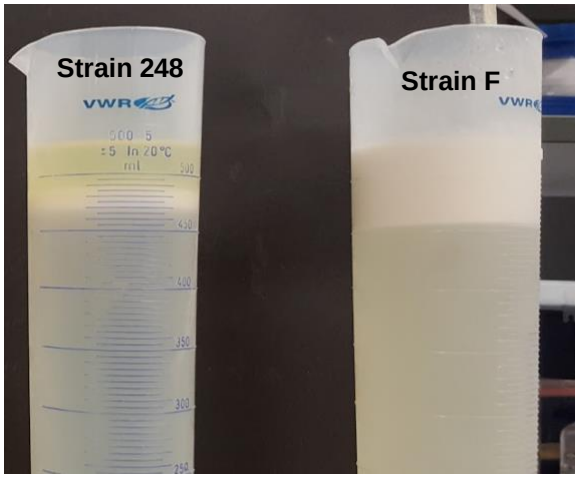


Figure 3.10 Fractionation of biphasic cultures of strain 248 and strain F

When allowed to stand for 12 hours, oil-water phases of the biphasic cultures of strains 248 and F clearly fractionated. In the case of strain 248 (left), a distinct biomass layer was formed, which enabled easy extraction of the desulphurised fuel. The oil phase fuel was significantly emulsified in the case of strain F (right), and therefore only a limited volume was recoverable.

3.3.6 BDS activity of the colonies growing on solidified medium

When grown on solidified MSM ($_{Agar}$MSM or $_{Gelrite}$MSM), *Gordonia desulfuricans* 213E produced visible colonies after 48 hours. The Gibbs test procedure was modified to study the *in situ* BDS activity of the 4-day old colonies. When Gibbs reagent was added directly over the colonies, no blue colour development was observed. However, when the filter paper pre-wetted with Gibbs reagent was laid over the colonies grown on solidified $_{Agar}MSM_{BD}$,

a blue colouration was observed over regions with the clustered colonies, but only sporadically over single colonies as shown in Figure 3.11. Colour development appeared patchy and smudged around clustered colonies, i.e. not confined to a single colony.

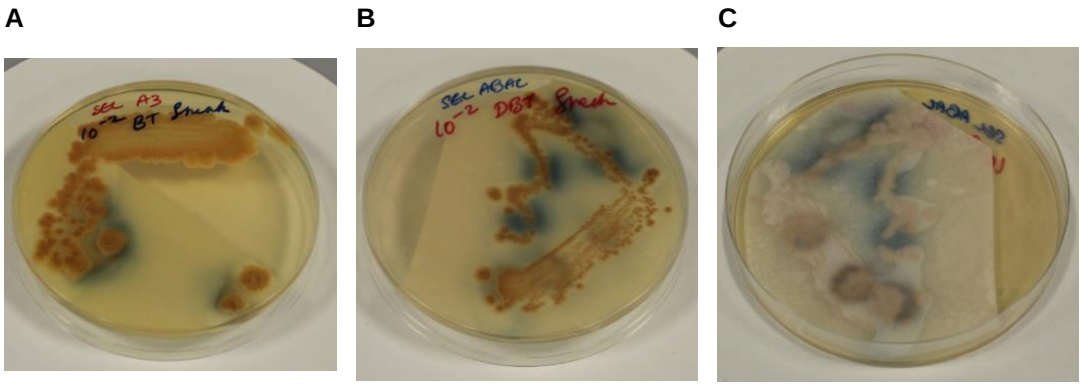


Figure 3.11 Gibbs Test conducted on *G. desulfuricans* 213E colonies growing on solidified MSM

The image shows the colour development observed 4 hours after the Gibbs reagent was introduced to the colonies of *G. desulfuricans* 213E growing on $_{Agar}$MSM$_{BT}$ **(A)**, $_{Agar}$MSM$_{DBT}$ **(B & C).** A filter paper pre-soaked in Gibbs reagent and applied over the 4-day old 213E colonies grown on solidified MSM. Plates without any bacteria and incubated for the same number of days were used as the negative control, where no colour development was observed (image not is shown).

The reaction mixture dispersed through the solidified medium and resulted in a transient and poorly visualised blue colour. The diffusion was not prevented even when the porosity of the gels was decreased by increasing the concentration of agar (1-5%) or Gelrite (0.75-2%). Moreover, increasing the agar concentration led to an even darker coloured gel which hindered visualisation of the colour development, where the Gibbs reaction mixture that appeared deep blue on clear plastic is seen pale blue-greenish on an $_{Agar}$MSM plate as shown in Figure 3.12. Therefore this approach was considered unreliable and not explored further in this research.

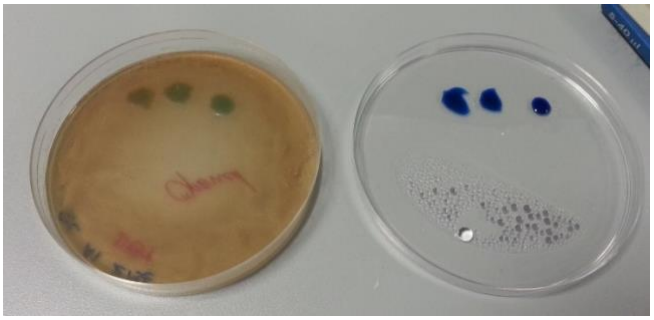


Figure 3.12 Gibbs reaction of 2-hydroxybiphenyl taken on $_{Agar}$MSM

The figure shows the comparative visualisation of the blue colour developed by the Gibbs reaction of 2-hydroxybiphenyl on an agar gel matrix (left) and over a clear plastic (right).

3.4 Discussion

This study presents the screening of soil actinomycete bacteria previously isolated from oil-contaminated soil and classified as members of genus *Rhodococcus*, for biodesulphurisation capabilities and assessment of their activity.

3.4.1 Stable BDS activity in the minimal most MSM

The primary aim of the study was to identify strains with BDS activity, and so the test strains were cultured in mineral salts medium containing BT or DBT as the sole sulphur source. Whole living bacterial cells were used as biocatalysts in this study rather than resting cell suspensions, and so the growth of the organism was crucial to the biodesulphurisation (BDS) activity. In general, the growth of bacteria is affected by the nutrient composition of the growth medium, due to cofactor requirements of many enzymes involved in the BDS biochemical pathway. However, the *Gordonia* and *Rhodococcus* strains used for BDS have not been reported to require special additives for their growth, and they could synthesize all the essential precursor metabolites from pure elemental sources in a medium (Arenskotter *et al.*, 2004; Finnerty, 1992). Only in a few cases, trace amounts of thiamine were added to the medium (Denome *et al.*, 1994; Gallardo *et al.*, 1997; Omori *et al.*, 1992).

A defined enrichment mineral salts medium (MSM) containing the vital inorganic nutrients that were commonly used in the mineral medium compositions widely used as in the literature was designed. The composition of the mineral salts medium (MSM) initially used in the research was based on the Hutner's mineral medium (Cohen-Bazire *et al.*, 1957; Hutner *et al.*, 1950). The same composition was used in the original description paper of the type strain *G. desulfuricans* 213E and subsequently used by Recylatech Group Ltd, for growing the bacteria for their bio-devulcanisation process. In its composition, it included Hutner's Basal Salts (HBS) solution, which contains several metal salts and involves tedious preparatory steps. To maintain consistency across several batches, growth medium (MSM) for commercial-scale application is prepared concentrated, and it is diluted as required for the batch. Precipitation of mineral salts occurs when mineral-rich medium contains

calcium and iron salts. Therefore, the recipe was changed to identify the minimal medium that supported the growth and BDS activity whilst being cheaper to make and easy to handle.

In the fourth edition of the Handbook of Microbiological Media (Atlas, 2010), there are several undefined and defined growth media recommended for the culturing of actinomycetes bacteria. Among the recommended defined media, especially for *Rhodococcus* strains, are Medium K (Kievskaya Broth) and Raymond's medium which contained salts of K, P, N, Na, Mg, Ca and Fe ions as the necessary components. Based on this observation, the MSM composition was altered gradually to contain only the essential elements.

Fluctuations in the pH affect the growth of the *R. erythropolis* and hence their BDS activity, and therefore, the medium was formulated to have a neutral pH at the beginning and during the course of incubation. This was a crucial determinant in the choice of carbon source because it was reported that citrate and glutamic acid as carbon sources led to a gradual increase in the pH of the culture and yielded lower biomass than glucose (del Olmo *et al.*, 2005; Martin *et al.*, 2005). To achieve higher biomass, and hence more biocatalyst that would be ideal for BDS, sucrose was used as the carbon source in the final MSM, as it was previously reported to yield higher biomass (Borgne and Quintero, 2003). Acetic acid and glycerol are other widely used carbon sources in BDS research, but they were not used in this study to keep the cost of the medium minimal. As the growth of the BDS strains was conserved in all the MSM recipes, the MSM-Recipe 4, which has the advantage of being a cheaper medium for commercial application, was used routinely in this study. Moreover, this version of MSM contained all minerals commonly found in the biodesulphurisation medium widely used in the literature.

For a commercially successful BDS technology, it is necessary to have bacteria with stable activity. The reproducibility of the results has been a problem since the beginning of BDS research (Kilbane, 1989). The use of repeated subculturing in the BT or DBT containing enrichment medium ensured the selection of organisms with stable BDS functionality (Kobayashi *et al.*, 2000; Wang *et al.*, 2017; Yoshikawa *et al.*, 2002). Based on this observation, the strains tested in this study were subjected to repeated

subculturing in the MSM_{BT} and MSM_{DBT} to allow for adaptive selection of the test strain with stable BDS activity.

It has been reported that growth on *n*-alkanes ranging in length from C20 to C36 could be challenging to determine due to poor solubility of *n*-alkanes in water. Growth in MSM containing a medium-chain alkane like *n*-hexadecane is a crucial indicator for the ability to grow in medium containing fossil fuels (Zampolli *et al.*, 2014). An increase of microbial DBT desulphurisation activity was observed when *R. erythropolis* strain H-2 was cultured in biphasic media containing tetradecane (70%) was also observed by Ohshiro *et al.* (1995). Patel *et al*. (1997) reported similar trends when *Rhodococcus* strain IGTS8 was cultured with 50% (v/v) hexadecane. It was suggested that this positive effect might be due to the extraction of the growth-inhibiting product 2HBP into the organic phase (Monticello, 2000), coupled with improved DBT availability and oxygen transfer in the presence of hexadecane (Abbad-Andaloussi *et al.*, 2003). In this study, the ability of the strains to grow in a biphasic (oil-water) medium was tested using *n*-hexadecane, a 16 carbon (C16) alkane and the primary hydrocarbon in diesel (Caro *et al.*, 2008; Rhee *et al.*, 1998). All the test strains survived growth in biphasic medium containing 1% *n*-hexadecane (data not shown). When grown in a biphasic medium, the test strains assembled themselves into flocs concurring with previous observations by Dorobantu *et al.*, (2004). Mohebali *et al.* (2007) reported that increasing cell concentrations led to a decrease in dibenzothiophene degradation by the formation of cellular flocs due to the hydrophobic nature of desulphurising bacteria. This phenomenon was observed when the strains were cultured in the biphasic medium. It was also reported that the level of desulphurisation activity varies with the kind of fossil fuel derivatives used. The wild-type strain *Rhodococcus* sp. strain IGTS8 did not result in a significant decrease in the total sulphur content of crude oil, whereas the BDS activity was better when refined products such as gasoline and diesel oil were used (Kaufman *et al.*, 1999). Therefore, two different biphasic systems containing 10% *n*-hexadecane and HDS-gas oil were used to study the BDS activity of strains 248, F and 213E.

3.4.2 Growth phase-dependent biodesulphurisation activity of the growing cells in a wholly aqueous medium

The BDS activity of all the strains was a growth-dependent process. In the case of strain 248, the BDS activity improved with an increase in biomass, with a peak level of accumulation occurring during the late exponential phase. The concentration of phenolic stopped as the growth reached the stationary phase. In the case of *G. desulfuricans* 213E, after continuing in the stationary phase, the biomass started to decline, indicating a death phase after 6 days of incubation. Interestingly, strain F showed a renewed growth phase after a short stationary phase, which was accompanied by a decrease in the concentration of the phenolic compound. A possible explanation for this could be that strain F utilised the sucrose as a carbon source during the primary growth phase and then adapted to using the phenol as a substrate leading to a second growth phase. The periodic dry cell weight measurements (DCW) of the culture samples followed a similar trend as the optical density values until day 4 (data not shown). In this research, DCW values were conducted to confirm if the increase in O.D. measurements were contributed by the actual increase in biomass and not because of turbidity caused by any emulsifying metabolites produced by the bacteria during growth.

3.4.3 Desulphurisation pathway of strains 248, 213E and F

The detection of phenols in the cultures using Gibbs reagent indicated that the 3 BDS strains might follow the 4S pathway. A limitation of the Gibbs reaction-based approach for detection of phenols is that the absorption peak is not very sensitive to the substitution of the phenol, which makes it ideal for determining the total phenol content in a sample, but not in distinguishing between specific substituted phenols (Mistry and Wenthold, 2018). Moreover, Ettinger and Ruchhoft (1948) indicated that standard phenolic curves are not repeatable but show a linear relationship up to 100 ppb, with the temperature and time being critical factors influencing reproducibility. Gibbs assay was mainly used in this research to study the formation of phenolic end products qualitatively, and the spectrometric measurements obtained to track the accumulation were relied upon only to understand the trend of BDS activity along with growth measurements. Therefore, GC-MS was performed to confirm the identity of

the phenolic end products produced by strains 248, 213E and F. As the DBT desulphurisation pathway and its intermediates are well known, and their spectral data were readily available, it was easy to detect them by GC-MS. The reaction intermediates of BT desulphurisation are not unambiguously established and their spectral data are not available for analysis. Therefore, only the end product of BT metabolism (benzofuran) was detected.

When presented with a mixture of aromatic sulphur compounds, even broad-spectrum BDS bacteria tend to be preferential for either BTs or DBTs, even though they can use any of the aromatic compounds as a sulphur source. *Rhodococcus* sp. WU-K2R preferred BT when presented with a BT/naphthothiophene mixture (Kirimura *et al.*, 2002). *Rhodococcus* sp. K1bD preferred DBT over 1,4-dithiane when presented together (Kirkwood *et al.*, 2005), and a consortia of BDS strains used for bunker oil desulphurisation showed a preference for DBT over benzo[*b*]naphtho[1,2-*d*]thiophene (Jiang *et al.*, 2014). A similar trend is reflected in the BDS activity of strain 248, which preferred BT over DBT, and for strain F, which preferred DBT over BT when grown in minimal MSM. The growth pattern of each strain was similar in all the MSM recipes, but when presented with a richer MSM-Recipe 1 or 2, the BDS activity spectrum of the strains was improved, as the strains 248 and F were able to desulphurise both BT and DBT.

3.4.4 Desulphurisation of *n*-hexadecane and diesel by growing cells

Aromatic hydrophobic compounds are toxic to bacteria due to their high partition into the membrane, and only microorganisms that are resistant to the highly toxic nature of the fuel grow to produce a thick pellicle in the bi-phasic system. Growth inhibition by toxic compounds (either end product or feedstock) is a major hampering factor in terms of the commercial prospects of biochemical processes (Nicolaou *et al.*, 2010). In this study, the basis of enrichment procedure was the ability of the test strains to assimilate sulphur derived from the polycyclic aromatic hydrocarbons dissolved in the hydrophobic phase. In the case of biocatalytic transformations of hydrophobic water-immiscible chemicals, oil-water biphasic reaction systems are considered suitable (Quijano *et al.*, 2009; Watanabe *et al.*, 2008). The practical importance of bacterial predilection for the organic phase should be carefully

considered for the development of efficient whole-cell biocatalyst acting in oil-water biphasic systems. Such two-liquid-phase culture systems resulted in better accession of the oil phase by hydrophobic bacteria like *R. opacus* B-4 and *R. erythropolis* PR4 (Hamada *et al.*, 2009, 2008).

The BDS associated advantages of growing *R. erythropolis* in a biphasic medium were previously reported (Ohshiro *et al.*, 1996, 1995). Strain 248 formed a distinct biomass layer when grown in biphasic MSM containing *n*-hexadecane or diesel, as shown in Figure 3.10. To avoid consumption of the fuel as the carbon source, sucrose, an easily assimilable sugar, was included in the growth medium. However, the richer lipid membrane brought the *R. erythropolis* strain F into close contact with the fuel and the natural tendency for hydrocarbon degradation and biosurfactant production by the *R. erythropolis* spp. could have led to the formation of stable emulsions in the case of strain F, as shown in Figure 3.10.

In a biphasic growth medium, the cultures for strains 248, 213E and F were of non-homogeneous nature, as shown in Figure 3.9, and the sample aliquots varied in biomass load. Therefore the cell growth could not be measured reliably using turbidity or dry cell weight methods. For all the 3 strains, the biphasic growth medium initially consisted of *n*-hexadecane as the oil phase. The ratio of 10% (v/v) oil phase was selected based on recommendations by authors who used a biphasic medium to study BDS activity by *Rhodococcus* strains previously (Davoodi-Dehaghani *et al.*, 2010; Yu *et al.*, 2006a). It has been suggested that high oil content inhibits oxygen supply and causes mass transfer limitations, while low oil-water ratio leads to dilution of the enzyme activity. One of the gaps in the literature is that the justification for the choice of oil-water ratios is not explicitly stated in many of the studies (Adlakha *et al.*, 2016). When resting cell systems were used, all the BDS strains formed thick foamy flocs with the oil phase which made the oil recovery impossible. Therefore that approach was not followed through in this research.

3.4.5 Problems in the development of a rapid screening method

The regular Gibbs assay procedure involved the use of liquid culture aliquots taken in a cuvette to which the Gibbs reagent was added. This procedure was

adapted in this study to work on a solidified medium so as to develop a rapid screening method to detect single bacterial colonies that exhibited BDS activity. Noble Agar was used as the solidifying agent to make up the solidified MSM ($_{Agar}$MSM), in order to avoid introducing trace sulphates to the bacteria that could be present in other inferior quality. Colonies were formed on the AgarMSM and Gibbs test was conducted over them. The greening brown background colour of the $_{Agar}$MSM matrix made the visualisation of the blue colour formation difficult, as seen in Figure 3.12. Even when Gelrite based clear solidified MSM was used, this approach was not successful because the colouration development was either very short-lived or unspecific to a colony.

The idea to develop a rapid screening method was inspired by the reports of Gibbs reaction-based biosensors described in section 3.1.3, which involved molecular imprinting that allowed for molecular assembly of desired chemicals structure on a solid polymer substrate. The principle of Gibbs reaction was adapted to develop fast, responsive sensors based on the colour-forming reaction between the reagent and phenols in the sample (Arip *et al.*, 2013; Bashir and Liu, 2009). While those references adopted microfluidic approaches, the methodology adopted in this research involved the use of relatively larger volumes of reagents that diffused substantially through the solid matrix (agar or gelrite) which resulted in a non-specific (diffused) and short-lived colour development, unlike the persistent deep blue colour observed in the conventional method using cuvettes. Therefore, this approach did not serve the purpose of screening for specific single colonies with BDS activity on a culture plate containing a mixed population of isolates or random mutants.

Chapter 4 Whole-genome sequence based analysis of sulphur metabolism-related genes in the desulphurising strains and establishment of their species identity

4.1 Introduction

Biodesulphurisation (BDS) is a multi-enzymes process, and the genes that encode this capability vary between different BDS strains. The enzymes involved in DBT desulphurisation are well established, but there is no explicit knowledge of the enzymes and genes involved in BT desulphurisation. Sulphate induced repression of the desulphurisation enzymes expression and inadequate levels of desulphurisation achieved by microorganisms have been critical problems affecting the commercial prospects of biodesulphurisation technology. Therefore, solving these problems by genetic manipulation requires a thorough understanding of the genes involved in the process in the specific organisms. In this chapter, the whole-genome sequence of the BDS strains will be investigated to identify the arrangement of genes encoding BDS activity and other sulphur metabolism related genes, for the purpose of identifying potential target genes that could be manipulated for strain enhancement in the future. In addition, the whole genome sequences of the BDS strains will be applied as a taxonomic tool to confirm the species identities proposed in Chapter 2.

4.1.1 Availability and quality of rhodococcal genome sequences

At the time of this research, 314 curated rhodococcal genomic assemblies covering 31 valid *Rhodococcus* species were available in the NCBI database. Among them, complete genomes were available for 9 species and others at various levels of assembly (contigs, scaffolds, chromosome level) which lack segments of the genome, leading to unreliable bioinformatic analysis outcomes. The strains with complete genome sequence data are not always the type strain of the species; some of them are published in journals, and others are direct submissions to the database. There is no curated ‘reference RefSeq’ (manually identified high-quality whole-genome sequences) genome assembly recognised for any species under this genus; nevertheless there are 6 ‘representative RefSeq’ (computationally determined high-quality genome) genome assemblies. From the whole genome data available on rhodococci investigated so far, it could be understood that they have large and complex genomes that probably acquired many genes by recombination in the distant past (Larkin *et al.*, 2010). There is increasing evidence that multiple pathways

and gene homologs are present that further increase *Rhodococcus* catabolic versatility. Comparative and functional genomic studies have been carried out to identify the genetic basis of metabolic capabilities of rhodococci (Pathak *et al.*, 2016), which show the existence of multiple pathways and gene homologs in their genomes. All three species of the "erythropolis" subclade (*R. erythropolis, R. opacus, R. jostii*), catabolise a wide range of oligosaccharides and organic compounds. In terms of biotechnology, *R. erythropolis* has two technical advantages, viz. tolerating a broad range of temperatures and possessing a relatively smaller genome, subsequently a faster growth rate. However, the larger genomes and plasmids of *R. jostii* and *R. opacus* contain genes encoding the ability to metabolise a wide range of organic substances.

4.1.2 Desulphurisation associated genes

The nucleotide sequence of the 9.7 kb DNA fragment containing the desulphurisation operon of *R. erythropolis* IGTS8, containing three genes (*dszA*, *dszB*, and *dszC*) that encode DBT desulphurisation enzymes DszA, DszB and DszC, respectively is available in the NCBI GenBank under the accession number U08850, and the 5.5 kb complete CDS of the *dszABC* is available under accession code L37363. The desulphurisation associated genes are often denoted as *dsz*, *tds, sox* or *bds* genes in the literature and often occur as operons. In their attempts to understand the *dsz* operon regulation, Li and team (1996) suggested that there could be a Dsz repressor which is induced directly by sulphates, Cys, Met and sulphur-rich Casamino acids. They also suggested a possible overlap between the operator and promoter region. There is only limited knowledge about the repressor or activator proteins associated with the operator site of the *dsz* operon. The 5' and 3' termini of the *dsz/bds/tds* genes are not conserved regions, and hence PCR primers based on the 5' and 3' ends of *R. erythropolis* IGTS8 *dszABC* genes failed to amplify relevant genes from *G. amicalis* F.5.25.8 (Kilbane II and Robbins, 2007). Rather than developing bespoke primers sets to suit for each type of BDS strain, obtaining the whole-genome sequence of the BDS strains and utilising the existing knowledge about the nucleotide sequence of *dsz* genes and amino acid sequence of Dsz enzymes obtained from various BDS strains, may be a more useful means by which BDS associated genes of

new BDS strains may be identified. However, despite a large number of BDS capable rhodococci and gordonae strains reported in the literature, validly published whole-genome sequence is available for only a some of them as shown in Table 4.1.

Table 4.1 List of BDS capable actinobacteria with validly published genome sequence containing BDS genes

Organism	Metabolic capability	Bioproject Accession	Authors
***Gordonia alkanivorans* CGMCC 6845**	hydrocarbon-degradation bacterium isolated from petroleum-contaminated saline soil	PRJNA227487	Wang *et al.*, 2014
***Gordonia amicalis* CCMA 559**	biosurfactant production	PRJNA215264	Domingos *et al.*, 2013
***Gordonia* sp. IITR100**	DBT desulphurisation	PRJNA376065	Jaishankar *et al.*, 2017
***Gordonia terrae* C-6**	BT desulphurisation	PRJNA196464	Wang *et al.*, 2013b
***Rhodococcus jostii* RHA1**	polychlorinated-biphenyls degradation	PRJNA13693	McLeod *et al.*, 2006
***Rhodococcus* sp. JVH1**	fluorinated organosulphur compounds desulphurisation	PRJNA46601	Brooks and Van Hamme, 2012
***Rhodococcus erythropolis* VSD3**	diesel fuel degradation	PRJNA348829	Stevens *et al.*, 2017
***Rhodococcus qingshengii* TUHH-12**	piezotolerant bacterium growing on crude oil and tetracosane as sole carbon sources	PRJNA246036	Lincoln *et al.*, 2015

*These strains were not published originally as BDS strains, but their genome contained genes homologous to *dszABC* genes which indicated their potential for DBT desulphurisation. While the genome sequence of *Rhodococcus erythropolis* IGTS8 is not available, the genome sequence of *Gordonia desulfuricans* 213E$_T$ is available (PRJDB465; GenBank: BCNF) but not validly published.

Santos *et al.* (2006) reported that the *dsz* operon sequence (4,149 bp) of *Gordonia amicalis* F.5.25.8 showed 84% homology with that of *G. alkanivorans* 1B (AY678116). Interestingly, the *dszB* gene sequence for the strain F.5.25.8 was only 95% homologous to the previously obtained partial sequence for the same (DQ174770.1). The individual genes *dszA*, *dszB* and *dszC* shared 85, 84 and 83% homology with those of *R. erythropolis* IGTS8. However, the similarities were low with other distantly related BDS strains such as *B. subtilis* (*bdsABC*) and *Paenibacillis* sp. A11 (*tdsABC*). In another study, comparative molecular analysis of 6 desulphurising rhodococci strains isolated from different geographic locations revealed a conserved nature of the *dsz* genes (Denis-Larose *et al.*, 1997).

In addition to the known desulphurisation genes and enzymes, alkane sulphonate monooxygenases (SsuDs), sulphonate ABC transporters (SsuABC) and sulphate permeases (CysQ) and sulphatases have been hypothesized to be associated with organosulfur compounds degradation (Ellis, 2011; Erwin *et al.*, 2005; Hatzios and Bertozzi, 2011; van der Ploeg *et al.*, 2001; Van Hamme *et al.*, 2013). In soil-dwelling bacteria, all the proteins that are highly homologous to the desulfinase enzyme DszB contain Cys, His, and Arg amino acid residues in their N-terminal regions. The genes encoding these enzymes occur in the genomic sequences of various bacterial species, not necessarily as a part of a defined operon system (Lee *et al.*, 2006a). These findings indicate the possibilities for novel sulphur metabolic pathways.

Gordonia sp. NB4-1Y, a very closely related species to *Gordonia desulfuricans* 213E, is capable of metabolising sulfonated poly-fluoroalkyl compounds. A neighbour-joining tree showed that the *ssuD* sequences of strain NB4-1Y tended to cluster with putative *ntaA* genes rather than with previously characterised *ssuD* sequences (Van Hamme *et al.*, 2013). The initial degradation of nitrilotriacetate, a widely applied chelator, is also catalysed by a class C flavoprotein monooxygenase NtaA (Knobel *et al.*, 1996; Uetz *et al.*, 1992; Xu *et al.*, 1997) which is encoded by *ntaA* gene. It has been reported that the monooxygenases such as NtaA/SnaA/SoxA/DszA could oxidize nitriloacetate using reduced flavin mononucleotide ($FMNH_2$) and O_2, as well as catalyse the 3rd step of the 4S pathway, which is the conversion of $DBTO_2$ (Dibenzothiophene-5,5-dioxide) to HBPS (2'-Hydroxybiphenyl-2-sulfinate). Alkane sulfonate monooxygenases (SsuD; EC 1.14.14.5) enables bacteria to metabolise a wide range of alkanesulfonates as sulphur source by using a reduced flavin (supplied by a SsuE enzyme) and dioxygen, SsuD generates a C4a-(hydro)peroxyflavin intermediate that is directly responsible for catalysing the conversion of alkanesulfonate to sulphite and the corresponding aldehyde (Armacost *et al.*, 2014; Ellis, 2011). SsuD was not known to desulfonate aromatic sulfonates, but Wang *et al.* (2013a) reported that SsuD was one of the upregulated genes when *G. terrae* C-6 was grown in the presence of BT as a sole sulphur source, indicating a possible role of SsuD as a BT desulphurising enzyme. Recently, genes encoding for three putative alkanesulfonate monooxygenases, seven putative sulphonate ABC

transporters, and two putative sulphate permeases were associated with the thermophilic biodesulphurisation activity of *Geobacillus thermoglucosidasius* W-2 (Zhu *et al.*, 2016).

In terms of the amino acid sequence, the enzyme DszC (EC 1.14.14.21) shared sequence homology with several acyl-coenzyme A dehyrdogenases, and the enzyme DszA (EC 1.14.14.22) shared significant homology to oxygenases such as the SnaA subunit of pristinamycin II_A synthase in *Streptomyces pristinaespiralis* and to component B of nitriloacetate monooxygenase in *Chelobacter* ATCC 29600. Knowledge of the degree of homology of the Dsz enzymes helps in the identification of putative genes encoding BT desulphurisation, which are presently not clearly established.

4.1.3 Sulphate metabolism genes

Biodesulphurisation activity is a natural starvation-induced process that the bacteria have evolved to meet their sulphur requirements, which is relatively less than their other nutritional requirements such as carbon, hydrogen or oxygen. The desulphurisation levels are naturally sufficient to meet the organisms' need, but very low for application in a commercial scale BDS technology. Knowledge of the sulphur uptake gene could be useful to develop a gene silencing strategy for BDS strain improvement. By devoiding bacteria of the ability to uptake sulphates, the bacteria will be forced to depend only on alternate sulphur sources in the fuels, thereby avoiding sulphate induced repression of *dsz* genes and increased BDS rates.

In *E. coli*, the majority of sulphate assimilation genes belong to the cysteine (*cys*) regulon that is positively regulated by the transcription factor CysB (Kredich, 1992). In the absence of sulphate and cysteine, *E. coli* can use aliphatic sulfonates as a source of sulphur for growth by the expression of the taurine transport enzymes (encoded by *tauABCD*) and alkanesulfonate transport enzymes (encoded by *ssuEADCB*) (Eichhorn *et al.*, 1999; van der Ploeg *et al.*, 2001). Currently, there is no study that exclusively focuses on the sulphate metabolism of rhodococci. Therefore results obtained for a closely related genus *Mycobacterium* could be used for comparative studies to understand the sulphate metabolism genes in the *Rhodococcus* test strains

used in this study. Wooff *et al.*, (2002) obtained the genome sequence of *Mycobacterium tuberculosis* BCG and found that in the case of the Gram-positive mycobacterial strains, sulphur assimilation was mediated by Cys·T·W·A· SubI ABC transporter complex (encoded by the *subI, cysTWA* cluster). The periplasmic sulphate must be activated to provide the necessary energy for either reduction or transfer. The enzyme ATP sulfurylase adenylates the inorganic sulphate to generate high energy phosphoric-sulphuric acid anhydride bond of adenosine-5'-phosphosulphate (APS). The energy used for this reaction is compensated by the GTP hydrolysis which results in a two-subunit enzyme complex comprising of a catalytic sulfurylase subunit (CysD) and a G protein-like GTPase (CysN). The APS then undergoes either reduction or phosphorylation. In the reduction pathway, it gets sequentially reduced by APS reductase (CysH) and sulphite reductase (CysJ) towards the biosynthesis of reduced sulphur metabolites. Alternatively, an APS kinase (CysC) completes 3'-phosphoadenosine-5'-phosphosulphate (PAPS) synthesis with phosphorylation of the 3'-hydroxyl of APS, using the energy derived by an ATP hydrolysis. In mycobacteria, APS kinase (CysC) is fused to the GTPase domain (CysN) of ATP sulfurylase (Hatzios and Bertozzi, 2011; Schelle and Bertozzi, 2006; Wooff *et al.*, 2002). In this research, a rational approach will be followed to identify the sulphur uptake associated genes and a potential target gene among them that can be silenced in the future as a way for BDS improvement.

4.1.4 Phylogenetic study using the whole genome sequence

Until very recently, bacterial taxonomy has relied on a polyphasic approach based on the combination of phenotypic, chemotaxonomic and genotypic characteristics. Phylogenetic analyses based on 16S rRNA gene sequence nucleotide similarity has informed and shaped bacterial taxonomy significantly, and various additional housekeeping genes have also been utilised for this purpose. Yet, there are limitations with this approach and indeed with DNA–DNA hybridisation (DDH) which is used as the 'gold standard' to delineate bacterial species. Whole-genome sequences contain the entirety of genetic information of bacterial strains and can help to resolve the taxonomic identity of bacteria. Goris and team (2007) cut the whole genome sequences into

1,000 bp long fragments and applied the BLAST algorithm to identify high-scoring segment pairs between whole-genome sequences to mimic the DDH method. These measures have been termed as the overall genome-related index (OGRI) values (Chun and Rainey, 2014), and the minimum standards and workflow (Figure 4.1) have been proposed for the use of genomic sequences for prokaryotic taxonomy (Chun *et al.*, 2018). The taxonomic resolution of OGRI is limited to differentiate only closely related species and not suitable for phylogenetic inference at the suprageneric rank level. Average Nucleotide Identity (ANI) is a similarity-type index, which is calculated by fragmentation of genome sequences, followed by nucleotide sequence search (usually BLAST), alignment and obtaining the scores for identity calculation. More accurate genomic sequences result in reliable ANI similarity values, whose coherence has been validated by comparison with corresponding 16S rRNA gene similarity and DDH values (Kim *et al.*, 2014; Klappenbach *et al.*, 2007).

Although DDH is still the gold-standard in species delineation, new proposals and standardisations are being put forward to enable taxonomists to avoid tedious determinations of DDH values in wet-lab experiments and to resolve previously dubious classifications based on whole-genome sequences homology (Chun *et al.*, 2018; De Vos *et al.*, 2017). ANIb (ANI algorithm using BLAST) has been used most widely for classification and identification of bacteria and archaea (Camelo-Castillo *et al.*, 2014; Chan *et al.*, 2012; Haley *et al.*, 2010; Hoffmann *et al.*, 2012; Jiménez *et al.*, 2013; Lee *et al.*, 2013; Loffler *et al.*, 2013; Lucena *et al.*, 2012; Ruvira *et al.*, 2013; Tran *et al.*, 2017; Yi *et al.*, 2012). As the 16S rRNA gene sequence of several *Rhodococcus* species share >99% similarity, this is actively promoted as a reliable and cost-effective approach to resolve the phylogeny of *Rhodococcus* species (Creason *et al.*, 2014; Sangal *et al.*, 2016, 2015). ANIb causes variations, usually minor when reciprocal calculations are compared, meaning, the ANI values for the same pair of genomic sequences differ depending on which one of them was assigned as a query or as a subject. This led to the development of an improved ANI algorithm, OrthoANI (Lee *et al.*, 2016) which accounts for the concept of orthology – homology among sequences descended from the same ancestral sequence for which function of the gene or sequence has been

conserved across evolutionary time. OrthoANIu (OrthoANI using USEARCH program; Edgar, 2010) shows good correlation with the popular ANIb, and it works significantly even faster than ANIb, making OrthoANIu as suitable for large-scale comparative studies (Yoon *et al.*, 2017).

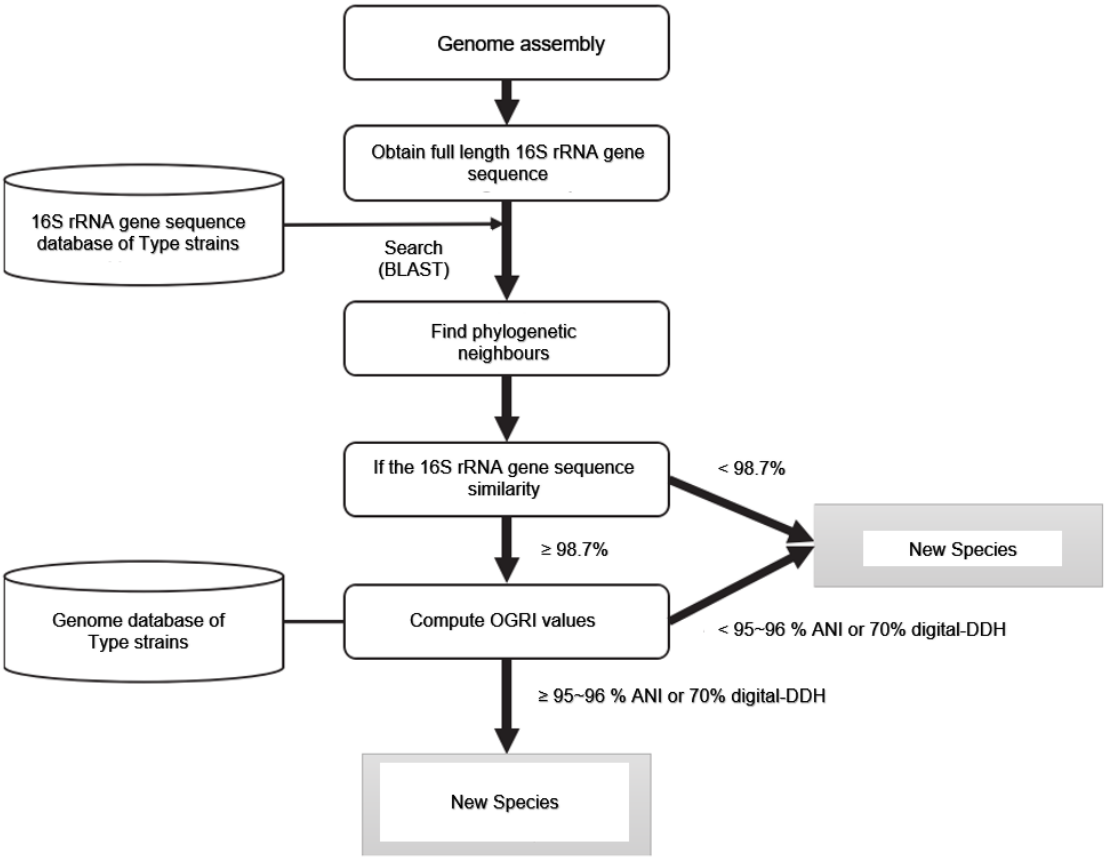


Figure 4.1 Workflow of whole-genome sequence based classification of bacteria at the species level

Alternatively, there are other bioinformatic tools for the rapid identification of bacteria based on raw genomic or metagenomic (mixed genomes) sequence reads. The Reads2Type web service (Saputra *et al.*, 2015) compares the similarity of each sequence in the input file with their 50-mer marker sequence database and rapidly identify any bacterial isolate to its species level. A more recent program StrainSeeker (Roosaare *et al.*, 2017) is a strain identification tool that offers the advantage of not having to upload the protected sequence reads to external servers for the analysis.

In this research, the species identity of the test strains was initially determined using their 16S rRNA and *gyrB* gene sequences (in Chapter 2) which confirmed their assignment as members of genus *Rhodococcus*. The BDS strains 248 and F were determined as *R. wratislaviensis* and *R. degradans,* respectively, although very high percentage similarity was observed with other closely related valid species and hence DDH analysis would be required for confirmation. As stated earlier, whole-genome sequence similarity offers a valid alternative to DDH, and therefore, the draft whole-genome sequence of these BDS strains will be used to confirm their species identity through ANI values obtained by comparison with the genome sequences of closely related *Rhodococcus* species available in the database. Although a draft whole genome sequence of the *G. desulfuricans* 213E / NBRC 100010 is already available (Genbank: BCNF01000001), a new sequence of the 213E strain will be obtained in this research so as to support the validity of the approach in terms of the DNA extraction and also ANI analysis.

4.1.5 Aim

This research study involves *in silico* analysis of whole-genome sequence data from the rhodococci test strains IEGM 248 and strain F, and the desulphurising species *Gordonia desulfuricans* 213E$_T$ to discover the gene systems involved in sulphur metabolism, with a particular focus on genes encoding desulphurising activity. It is expected that this information will enable the future development of targeted mutagenesis strategies for improving biodesulphurisation capabilities, by way of avoiding sulphate induced repression, in these and other strains.

Additionally, the whole genome sequence data from these strains will be used for taxonomic purposes, to identify them to species level using OGRI, and to compare the findings with those of the phylogenetic studies based on single gene (16S rRNA and *gyrB* gene) analysis conducted in Chapter 2, in order to evaluate these approaches.

Objectives

The main objectives of the study are as follows:

1) To identify the genes involved in sulphur metabolism and biodesulphurisation of BT and DBT in each of the three BDS strains, and study the genetic loci, occurrence patterns and arrangement of the genes, by using comparative genomic approaches, in order to inform future strain improvement strategies.

2) To determine the whole-genome similarities between strains IEGM 248, F and closely related valid *Rhodococcus* species in order to confirm species-level identifications.

4.2 Methods

4.2.1 Whole-genome sequencing of test strains

The genomic DNA of the test strains were obtained, as described in section 2.2.5. EDTA has been known to inhibit NGS library preparation, and so when preparing the DNA extractions for whole-genome sequencing, they were eluted in a buffer (10 mM Tris-HCl pH 8.5) with no EDTA. The purity and integrity were checked as described in section 2.2.7. The DNA was sent for whole genome sequencing by MicrobesNG, which is a BBSRC-funded *(grant number BB/L024209/1)* collaboration between the University of Birmingham and the University of Sheffield.

At MicrobesNG, the DNA in the samples were fragmented to comply with Illumina library preparation method. The DNA fragments were converted into the library by ligation to sequencing adapters containing specific sequences designed to interact with the surface of the flow cell used in Illumina platform, followed by clonal amplification of the library by cluster generation. Finally, sequencing was done on the Illumina HiSeq 2500 platform using 2x250bp paired-end reads. The reads were trimmed using Trimmomatic (Bolger *et al.*, 2014). The nucleotide reads data were subjected to bioinformatics analyses such as the identification of the closest available reference genome using Kraken (Wood and Salzberg, 2014), and mapping of the reads to the reference genome using BWA-MEM (Burrows-Wheeler Aligner) (Li, 2013) to assess the quality of the data. The reads were also subjected to *de novo* assembly using SPAdes (Bankevich *et al.*, 2012), and mapped back to the resultant contigs, again using BWA-MEM to obtain further metrics on the quality. The reference contigs were reordered and reoriented relative to the reference genome based on a MUMmer (Kurtz *et al.*, 2004) whole-genome alignment, and an automated annotation was performed using Prokka (Seemann, 2014).

These analyses were performed at MicrobesNG, and the resulting annotated draft genome assembly was available via a user-friendly web interface. Repeat sequences larger than ~1000bp (e.g. IS elements and rRNA operons) cannot be resolved using the sequencing methodology used by MicrobesNG, and which cause a break in the assembly. Therefore, a closed genome would not

be obtained *per se,* and the sequences were obtained as several contigs. The output was provided in standard bioinformatics file formats such as *fasta*, *gbk* (GenBank file format), *gff* (Generic Feature Format). Contaminating DNA sequences could occur during both culturing and DNA sequencing steps and even in minor amounts could get incorporated into the NGS sequence assembly. The quality of the draft sequence produced by MicrobesNG was tested using ContEst16S program (Lee *et al.*, 2017).

4.2.2 Importing the whole genome sequence into Geneious

Geneious™ (Biomatters Ltd., Auckland, New Zealand), a recognised bioinformatics suite for phylogenetic analyses (Czech *et al.*, 2017; Masters *et al.*, 2011; Smith, 2015) and genomics (Abbasian *et al.*, 2016; Butler III *et al.*, 2016; Wang *et al.*, 2017) was used to conduct bioinformatics analyses. Geneious package offers tools for NGS analyses for storing, organising and analysing (*de novo* assembly or mapping to a reference sequence) NGS data; tools for evolutionary analyses such as MSA, phylogenetic trees construction, repeat identification; inbuilt database searching and importing data into the software environment; extensions or plugins to include further operations.

4.2.3 Phylogenetic analysis using whole-genome sequences

The phylogenetic identity of strains 248 and F was studied using OrthoANIu (Yoon *et al.*, 2017) downloaded from http://www.ezbiocloud.net/tools/ani. The whole-genome sequence of the test strains was selected in fasta format. The closely related rhodococci species were chosen based on the similarity values obtained by the comparative analysis of their 16S rRNA gene and *gyrB* gene (described in chapter 2), and their representative genome sequences were obtained by searching the NCBI Prokaryotic Genome database (Table 4.2). Wherever a closely related representative genome was not available, the next high-quality genome of the closely related type species was chosen. The genome of strain 248 was compared with the highest quality genome available for *R. wratislaviensis, R. opacus, R. imtechensis* and *R. jostii*. As the whole-genome sequence of *R. degradans* was not available in the database at the time of the experiment, a comparative analysis was not possible between strain 248 and *R. degradans*. The genome of strain F was compared with *R. erythropolis* and *R. qingshengii*. The draft-genome sequence of *Gordonia*

desulfuricans 213E$_T$ obtained in this research was compared with the previously reported whole-genome sequence for the same strain available in the database as control. Additionally, the web-based StrainSeeker and Read2Type programs were used to identify the strains 248 and F.

Table 4.2 Representative rhodococcal whole-genome sequences used for OrthoANIu analysis in this study

Organism	Genbank Code	Contigs	Total length (kbp)	GC (%)
Strain 248		339	9,375	66.82
***R. wratislaviensis* NBRC 100605**	NZ_BAWF	151	10,403	66.78
***R. opacus* DSM 43205**	NZ_LRRG	382	8,534	67.28
***R. imtechensis* RKJ300**	NZ_AJJH	178	8,231	67.22
***R. jostii* RHA1**	NC_FNTL	1	7,804	67.52
Strain F		71	6,421	62.51
***R. erythropolis* NBRC 15567**	BCRM	67	6,588	62.4
***R. qingshengii* djl-6-2**	CP025959	1	6,518	62.44

The table shows the genomes of closely related organisms selected for pairwise comparison with the genomes of strains 248 and F using OrthoANIu analysis to obtain the genome-wide similarity values. Strain 248 was compared with the highest quality genome available for *R. wratislaviensis, R. opacus, R. imtechensis* and *R. jostii*. The strain F genome was compared with *R. erythropolis* and *R. qingshengii*. The genome of *R. degradans* was not available for comparison.

4.2.4 Identification of desulphurisation and sulphur assimilation genes

The *dsz* operon sequence of *Rhodococcus erythropolis IGTS8* (Accession: L37363) was used as the reference sequence. A short sequence between nucleotides (901 - 960) (5' aagtactaccaacacatcgcccgtactctggagcgcggcaagttcgatctgttgtttctg 3') of *dsz* operon was used as a query to identify other similar desulphurisation operons in the NCBI Nucleotide database using BLAST and from the database hits that had a higher score, expect value, identity (more than 85%) and query coverage (more than 65%), a list of 22 organisms which contained *dsz* genes were obtained as shown in Table 4.3. In the case of strain F, the genes related

to desulphurisation were known directly from the annotation of the whole genome sequence made by Prokka and by matching the above query with the genome using 'Map to Reference' functionality in Geneious program. In the case of the strains 248 and 213E, the genes related to desulphurisation were known by searching the annotations for all monooxygenases class of enzymes (EC 1.14) such as alkanesulfonate monooxygenases, flavin-dependent monooxygenases, acyl-coA dehydrogenases on the genome sequence using gene names, product name and EC numbers.

Table 4.3 DBT desulphurisation operon sequences available in the NCBI database

Organism	NCBI Nucleotide Accession No.
Acidovorax delafieldii	DQ062154.1
***Agrobacterium tumefaciens* FD-3**	AY960127.1
Bacillus subtilis	AB076745.1
Brevibacillus brevis	DQ062161.1
***Gordonia alkanivorans* 1B**	AY678116
***Gordonia alkanivorans* RIPI90A**	EU364831.1
***Gordonia amicalis* F.5.25.8**	EF026089
***Gordonia nitida* (reclassified as *G. alkanivorans*)**	AY714057.1
***Gordonia* sp. CYKS2 (putative operon)**	AY396519.1
***Gordonia* sp. IITR100**	KC693733
***Mycobacterium goodii* X7B**	JF740062.1
***Mycobacterium phlei* SM120-1**	KP202690.1
Nocardia globerula	AY714059.1
Rhodococcus erythropolis	AY714058.1
***Rhodococcus erythropolis* HN2**	KJ021035.1
***Rhodococcus* sp. IGTS8**	L37363.1; U08850.1; RERDSZA
***Rhodococcus* sp. DS-3**	DQ444325.1
***Rhodococcus* sp. SDUZAWQ**	AY789136.1
***Rhodococcus erythropolis* XP**	AY278323.1
synthetic construct (*Rhodococcus* sp. LY822)	EF570781.1

4.3 Results

The complete genome sequences data for the new BDS strains 248 and F generated by MicrobesNG, UK, were analysed using various bioinformatics tools with an aim to establish their species identity and to study the genes encoding the BDS activity.

4.3.1 Whole-genome sequences

The draft whole-genome sequences of the strains 248, F and *G. desulfuricans* 213E obtained from MicrobesNG have been deposited at NCBI/GenBank under BioProject numbers PRJNA603520, PRJNA603528 and PRJNA603534, respectively. The NCBI Prokaryotic Genome Annotation Pipeline (PGAP) annotated whole-genome sequences of *R. opacus* IEGM 248, *R. qingshengii* strain F and *G. desulfuricans* 213E are available from GenBank using accession numbers JAAECF000000000, JAADZT000000000 and JAADZU000000000, respectively.

As a part of MicrobesNG's standard analysis pipeline, the nucleotide sequence reads were subjected to *de novo* assembly using SPAdes to obtain the contigs and eventually a draft genome was constructed. Additionally, a nearest related reference genome was also identified using Kraken. The nucleotide sequence reads were mapped to both the newly obtained draft genome and also the reference genome using BWA-MEM to assess the quality of data and generate quality metrics for the QUality ASsesment Tool (QUAST) report. The assembly statistics of contigs in the draft genome sequence are provided in Table 4.4.

Table 4.4 QUAST report on the whole genome sequence output

Statistics without reference	Strain 248	Strain F	Strain 213E
# contigs	237	49	204
# contigs (≥ 0 bp)	339	71	266
# contigs (≥ 1000 bp)	209	43	184
Largest contig length (bp)	466,750	729,836	142,683
Total number of bases in the assembly	9,332,585	6,414,301	5,534,161
N50	87,833	307,870	53,400
N75	55,695	202,740	28,386
L50	27	7	33
L75	59	14	69
GC (%)	66.93	62.51	68
Mismatches			
# N's	0	0	0
# N's per 100 kbp	0	0	0

All statistics are based on contigs of size ≥ 500 bp, unless otherwise noted (e.g., "# contigs (≥ 0 bp)" and "Total length (≥ 0 bp)" include all contigs.)

The N50 is the contig length such that using longer or equal length contigs produces half (50%) of the bases of the assembly. N75 is the contig length such that using longer or equal length contigs produces 75% of the bases of the assembly. Usually, there is no exact value that produces the 50% or 75%, and so the technical definition is that they represent the maximum length x such that using contigs of length at least x accounts for at least 50 or 70% of the total assembly length. L50 and L75 represent the number of contigs that have lengths equal to N50 and N75, respectively.

4.3.2 Desulphurisation genes of strain F

Strain F is a DBT desulphurising bacteria and therefore the known *R. erythropolis* IGTS8 *dsz* operon sequence (NCBI Nucleotide: RERDSZA or L37363) (Piddington *et al.*, 1995) was used as the reference sequence to locate the desulphurisation genes in the genome of strain F using the "Map to reference" option in Geneious. The desulphurising operon *dszABC* of strain F was located in the contig 20 of the draft genome sequence obtained in this study, as shown in Figure 4.2. Three gene sequences occurred in succession as in the case of *dszA*, *dszB* and *dszC* and they were annotated as the genes encoding Nitrilotriacetate monooxygenase component A (*ntaA*), 2'-hydroxybiphenyl-2-sulfinate desulfinase (*soxB*) and Dibenzothiophene desulphurisation enzyme C (*soxC*), respectively. Multiple sequence alignment of the *dszABC* sequence of strain F and other known DBT desulphurisation gene cluster sequences reported from different DBT desulphurising strains showed a 100% match with *dszABC* sequence of the IGTS8 strain and also showed a high similarity with that of *Acidovorax delafieldii* (NCBI Nucleotide: DQ062154), *Brevibacillus brevis* (NCBI Nucleotide: DQ062161) and *Rhodococcus* sp. strain IGTS8 (NCBI Nucleotide: U08850) (Denome *et al.*, 1994) as shown in Table 4.5.

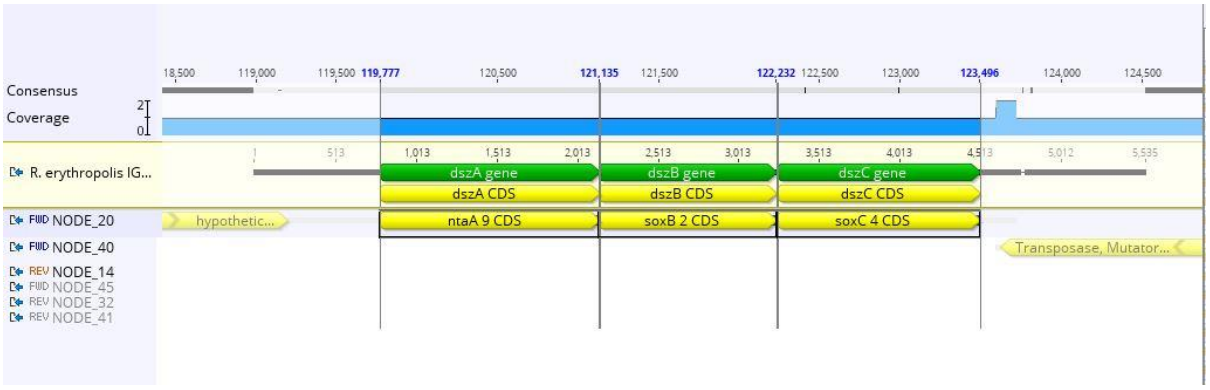


Figure 4.2 Detection of the *dsz* genes of strain F using comparative genomics

As the strain F was a dibenzothiophene (DBT) desulphurising bacteria, the 5.5 kb *dszABC* operon sequence of *R. erythropolis* IGTS8 (Genbank: L37363) was used to search the draft whole-genome sequence of strain F to detect the locus and the nucleotide sequences flanking the DBT desulphurisation operon. The 100% matching sequence containing all the 3 genes was located on the contig 20 (Node 20) of the draft sequence. The blue coloured numbers indicate the nucleotide count on the contig 20.

Table 4.5 Comparison of *dsz* operon sequences of different DBT desulphurising bacteria

	M. goodii X7B (JF740062)	*Bacillus subtilis* (AB076745)	*M. phlei* SM120-1 (KP202690)	G. amicalis F.5.25.8 (EF026089)	*G. alkanivorans* 1B (AY678116)	*G. alkanivorans* RIPI90A (EU364831)	*Gordonia* sp. CYKS2 (AY396519)	*Gordonia sp.* IITR100 (KC693733)	*R. erythropolis* HN2 (KJ021035)	*G. nitida* (AY714057)	*Rhodococcus* sp. SDUZAWQ (AY789136)	*Nocardia globerula* (AY714059)	*R. erythropolis* (AY714058)	*R. erythropolis* IGTS8 dszABC (RERDSZA)	*Agrobacterium tumefaciens* (AY960127)	*R. erythropolis* XP (AY278323)	*Acidovorax delafieldii* (DQ062154)	*Brevibacillus brevis* (DQ062161)	*Rhodococcus* sp. (RSU08850)	Strain_F dszABC
M. goodii X7B (JF740062)		636	1843	1580	2224	2207	2465	3971	1208	1716	1161	1709	1711	2588	1259	1157	1158	1158	3832	1395
Bacillus subtilis (AB076745)	85		342	1516	1482	1466	1584	1576	1183	1573	1134	1569	1568	1563	1236	1130	1131	1131	1564	1381
M. phlei SM120-1 (KP202690)	67.56	92.33		2041	2705	2689	3773	6384	1183	2389	1134	2385	2384	3198	1377	1130	1131	1131	5424	1614
G. amicalis F.5.25.8 (EF026089)	61.91	63.09	56.45		895	885	1499	1493	671	1119	598	1109	1109	1107	676	591	592	592	1102	778
G. alkanivorans 1B (AY678116)	56.53	64.86	51.88	77.74		2	282	2752	504	669	417	661	659	1487	482	407	410	410	3308	574
G. alkanivorans RIPI90A (EU364831)	56.65	65.03	51.95	77.89	99.96		280	2731	504	658	417	650	648	1476	484	407	410	410	3272	576
Gordonia sp. CYKS2 (AY396519)	53.46	63.63	46.12	68.27	93.79	93.8		2818	500	1240	413	1231	1229	2157	480	403	406	406	4341	572
Gordonia sp. IITR100 (KC693733)	43.36	63.81	37.87	68.4	61.85	61.83	66.85		503	1243	416	1234	1232	1937	483	406	409	409	5046	575
R. erythropolis HN2 (KJ021035)	67.64	68.31	68.31	81.48	86.45	86.45	86.56	86.48		120	111	111	109	108	106	105	104	104	104	104
G. nitida (AY714057)	61.92	63.71	54.69	74.35	84.3	84.49	75.15	75.09	96.77		23	31	35	25	21	17	16	16	37	16
Rhodococcus sp. SDUZAWQ (AY789136)	68.9	69.62	69.62	83.5	88.79	88.79	88.9	88.82	97.02	99.38		14	12	11	9	8	7	7	7	7
Nocardia globerula (AY714059)	62.07	63.81	54.76	74.58	84.49	84.68	75.33	75.27	97.02	99.32	99.62		26	16	12	8	7	7	31	10
R. erythropolis (AY714058)	62.05	63.85	54.8	74.59	84.53	84.72	75.37	75.31	97.07	99.23	99.68	99.43		20	10	6	5	5	34	6
R. erythropolis IGTS8 dszABC (RERDSZA)	53.01	63.94	49.13	74.63	71.13	71.23	63.79	67.72	97.1	99.45	99.7	99.65	99.56		9	5	4	4	32	4
Agrobacterium tumefaciens (AY960127)	67.63	68.22	65.83	82.02	87.67	87.62	87.72	87.64	97.15	99.46	99.76	99.69	99.74	99.77		3	2	2	5	5
R. erythropolis XP (AY278323)	68.98	69.71	69.71	83.69	89.05	89.05	89.16	89.08	97.18	99.54	99.78	99.78	99.84	99.87	99.92		1	1	1	1
Acidovorax delafieldii (DQ062154)	68.98	69.7	69.7	83.66	88.98	88.98	89.09	89.01	97.2	99.57	99.81	99.81	99.87	99.89	99.95	99.97		0	0	0
Brevibacillus brevis (DQ062161)	68.98	69.7	69.7	83.66	88.98	88.98	89.09	89.01	97.2	99.57	99.81	99.81	99.87	99.89	99.95	99.97	100		0	0
Rhodococcus sp. IGTS8 (RSU08850)	43.98	63.92	38.75	74.74	53.08	53.21	47.94	51.43	97.2	99.19	99.81	99.32	99.26	99.42	99.87	99.97	100	100		0
Strain_F dszABC	65.87	66.21	62.85	80.1	86.02	85.98	86.07	86	97.2	99.61	99.81	99.75	99.85	99.9	99.87	99.97	100	100	100	

Nucleotide difference

Percentage similarity

Molecular phylogenetic analysis based on the complete *dszABC* sequence of strain F and other desulphurising bacterial species by Maximum Likelihood method. The NCBI accession code for the *dsz* operon is provided in parenthesis. The matrix shows variation among the gene sets that encode the same functionality. Darker shades indicate high similarity.

The *dszABC* gene cluster was not found in the draft whole-genome sequence of strains 248 and *G. desulfuricans* 213E, which are predominantly BT desulphurising organisms. As noted earlier, genes encoding the desulphurisation of benzothiophene (BT) are not clearly established in the literature, and there was no reference sequence that could be used to probe the BT desulphurisation genes in the whole-genome sequence of strains 248 and 213E. Therefore, the whole-genome sequences of strain 248 and 213E were searched to identify genes encoding various monooxygenases and identify putative genes for BT desulphurisation.

4.3.3 Monooxygenases genes occurring in the genome of strains 248, F and *Gordonia desulfuricans* 213E

Strain F

There were 91 instances of monooxygenase related genes (EC 1.14.x.x) (EC 1.13.x.x) occurring at various nodes in the draft genome sequences. There was an abundant occurrence of genes encoding flavin-dependent monooxygenase (*hsaA, hsaB*), alkane monooxygenase (*alkB*), alkane sulfonate monooxygenase (*ssuD*), limonene 1,2-monooxygenase (*limB*), nitrilotriacetate monooxygenase (*ntaA*) and nitronate monooxygenase at various regions of the genome (Appendix 7.8).

A single instance of NAD(P)H-dependent FAD/FMN reductase occurred on Node_14. At Node_12, genes related to the metabolism of organic sulphur compounds occur in proximity. The alkanesulfonate monooxygenase genes (*ssuD_3, ssuD_4*) occur consecutively and immediately followed by *soxB_1* (gene for the HBPS desulfinase) and other sulphur metabolism related genes such as aliphatic sulfonates import ATP-binding protein (*ssuB_4*), putative aliphatic sulfonates transport permease protein (*ssuC_5*), and putative aliphatic sulfonates-binding protein precursor (*ssuA_3*). As described in section 4.3.2, the full *dszABC* operon was found in the whole-genome sequence of strain F .

Strain 213E

In the case of strain 213E, the draft genome sequence revealed 79 monooxygenase related genes are occurring at various nodes of the genome sequence **(**Appendix 7.9**)**. Repeated occurrence of the genes encoding 2,4-dichlorophenol 6-monooxygenase (*tfdB*), 4-hydroxy acetophenone monooxygenase (*hapE*), 4-nitrophenol 4-monooxygenase (*npcB*), alkanal monooxygenase alpha chain (*luxA*), alkanesulfonate monooxygenase (*ssuD*), dimethyl-sulfide monooxygenase (*dmoA*), flavin-dependent monooxygenase (*hsaA*), limonene 1,2-monooxygenase, nitrilotriacetate monooxygenase component A (*ntaA*), nitronate monooxygenase and pyrimidine monooxygenase (*rutA*) were observed.

At Node_27, the genes encoding FMN reductases (*ntaB* and *ssuE*) occur consecutively and closely followed by a gene for alkanesulfonate monooxygenase (*ssuD_2*). The FMN reductase NtaB (EC 1.5.1.42), unlike SsuE (EC 1.5.1.38), has a strong preference for NADH over NADPH.

The draft genome shows 6 instances of the gene encoding DszC enzyme occurring at various nodes. At Node_12, Node_38, and Node_52, the gene *dszC* occur successively with other monooxygenases encoding genes such as *ntaA*, *msuD* and *dmoA_3,* respectively.

At Node_76, the genes related to aliphatic sulphur metabolism such as *ssuA* (binding), *ssuB* (import) and *ssuC* (transport permease) occur clustered and in proximity to *ssuE_3* (FMN reductase).

Strain 248

About 159 instances of monooxygenase related genes were observed in the draft genome of strain 248 which included genes encoding 2,4-dichlorophenol 6-monooxygenase (*tfdB*), 3-ketosteroid-9-alpha-monooxygenase oxygenase subunit (*kshA*), 4-hydroxy acetophenone monooxygenase (*hapE*), alkanal monooxygenase alpha chain (*luxA*), alkanesulfonate monooxygenase (*ssuD*), antibiotic biosynthesis monooxygenase, flavin-dependent monooxygenase oxygenase subunit (*hsaA*), flavin-dependent monooxygenase reductase subunit (*hsaB*), limonene 1,2-monooxygenase (*limB*), nitrilotriacetate

monooxygenase component A (*ntaA*), nitronate monooxygenase, pentachlorophenol 4-monooxygenase (*pcpB*) and pyrimidine monooxygenase (*rutA*), as shown in Appendix 7.10.

Node_40 consists of two clusters of monooxygenase genes that are made up of genes *ntaA_7*, *ntaA_8* and *ntaA_9* occurring as a tight cluster and genes *ntaA_10*, *moxC_2* and *limB_5* occurring as another tight cluster.

The *dszC* gene occurs at Node_5 clustered with *msuD_2* (gene for methanesulfonate monooxygenase), and at Node_83 as a cluster with *ntaA_12* and *ntaA_13*.

4.3.4 Sulphur metabolism related genes occurring in the genome of strains 248, F and 213E

The Cys family of proteins encoded by *cys* genes play vital roles in the sulphate metabolism of bacteria, with CysB being the master controller. The genes encoding sulphate binding proteins *sbpA,* and the *cysTWA* genes occur as conserved clusters in the whole-genome sequence of the three BDS strains analysed in this research, as shown in Figure 4.3, where it could also be seen that the gene associated with cysteine biosynthesis *cysD* and *cysH* also occur clustered. The bifunctional protein CysNC plays a role in the synthesis of sulphite from sulphate. The gene for *cysNC* always occur adjacent to *cysD* genes and translated in the same direction, with more (at least 3) copies spread across the genome than the *cysTWA* genes which occur only once in all three BDS strains. In the BT desulphurising strains 248 and 213E, one copy of the *cysNC-cysD* pair occur in the same contig containing the *cysTWA* genes, distinguishing them from the DBT desulphurising strain F where all copies of the *cysNC-cysD* pair occurs at different loci. This observation concurred with the findings of Woof *et al.* (2002) who reported the clustered arrangement of these sulphur uptake genes in *M. tuberculosis* BCG.

The knowledge of sulphate assimilation genes is important because regulation of sulfur metabolism relies on the transcriptional response of sulphate assimilation enzymes to diverse environmental cues and regulatory proteins that influence flux through the sulfate assimilation pathway (Hatzios and Bertozzi, 2011). The knowledge of their location in the genonme and the

sequences flanking them will be helpful to develop gene silencing approaches in the future to generate genetically modified strains. These mutant strains will have their sulphate genes silenced, and therefore depend on the constitutive expression of *dsz* genes to meet their sulphur requirements even in the presence of sulphates and exhibit improved BDS rates (proposed as future work in section 5.1.6).

A) Sulphate uptake genes of *G. desulfuricans* 213E (contig 4)

B) Sulphate uptake genes of strain 248 (contig 11)

C) Sulphate uptake genes of strain F (contig 3)

Figure 4.3 Sulphate uptake-related genes occurring on the genomes of strains 213E, 248 and F

The segment of whole-genome sequence showing a clustered arrangement of genes associated with uptake of sulphates found in strains *G. desulfuricans 213E* (A), *Rhodococcus opacus* IEGM 248 (B) and *R. qingshengii* F (C) are shown. The contig number of the draft whole-genome sequence where these genes occur is given in parentheses. The degree of conservation of the arrangement is such that gene for sulphate binding precursor protein *sbpA* and genes for sulphate transportation proteins *cysTWA* occur clustered, and are translated in forward direction. The genes encoding cysteine biosynthesis *cysD* and *cysH* occur clustered, and are translated in reverse direction. This observation further supports the evolutionary idea that genes encoding proteins associated with a single process tend to occur closer to each other. The three BDS strains differ in terms of the other genes occurring in between the *cysTWA* and *cysDH* clusters. The arrangement of the sulphur uptake genes of strains 248, F and 213E is homologous to *Mycobacterium tuberculosis* BCG,which is an actinobacterial strain.

4.3.5 Identification of Strain 248

The Read2Type rapid bacterial identification program, after 18 minutes of operation, divided the submitted genome sequence of strain 248 into 38,839 50-mer sequences and predicted 4 different possible identities for the strain 248 as *Nocardia farcinica, Haloarcula hispanica, Rhodococcus jostii,* and *Rhodococcus opacus*. The StrainSeeker program compared the genome sequence of strain 248 against a database of more than 4,300 bacterial strains obtained from NCBI RefSeq repository and detected that *R. jostii* RHA1 and *R. opacus PD630* genome sequences occurring at a relative frequency of 57.24% and 42.76% on the whole genome sequence of strain 248.

The OrthoANIu based comparison of strain 248 with genome sequences of *R. wratislaviensis* NBRC 100605, *R. opacus* DSM 43205 and *R. imtechensis* RKJ300 revealed a percentage match of 94.59%, 96.66% and 96.18%, respectively. Interestingly, a high similarity was obtained when the genomes of the type species were subjected to the pairwise comparison as given in Table 4.6. Based on the OrthoANIu values, the species identity of strain 248 could be reliably confirmed as *R. opacus*.

Table 4.6 OrthANIu values for the pairwise comparison of strain 248 and the genomes of its closely related type strains

Metric	*R. wratislaviensis* vs 248	*R. opacus* vs 248	*R. imtechensis* vs 248	*R. jostii* vs 248	*R. wratislaviensis* vs *R. imtechensis*	*R. opacus* vs *R. imtechensis*	*R. wratislaviensis* vs *R. opacus*
OrthoANIu value (%)	**94.59**	**96.66**	**96.18**	**94.99**	**94.34**	**97.88**	**94.51**
Genome A length (bp)	10,329,540	8,336,460	8,139,600	7,804,020	10,329,540	8,336,460	10,329,540
Genome B length (bp)	9,208,560	9,208,560	9,208,560	9,208,560	8,139,600	8,139,600	8,336,460
Average aligned length (bp)	4,907,944	4,796,859	4,450,152	4,827,779	4,242,131	4,441,082	4,512,155
Genome A coverage (%)	47.51	57.54	54.67	61.86	41.07	53.27	43.68
Genome B coverage (%)	53.3	52.09	48.33	52.43	52.12	54.56	54.13

The draft genome sequence of strain 248 was compared with closely related representative genomes of type strains of *R. wratislaviensis* NBRC 100605, *R. opacus* DSM 43205, *R. jostii* RHA1and *R. imtechensis* RKJ300. Genome A refers to the reference genome obtained from the NCBI database, and Genome B refers to the WGS of strain F obtained in this research.

4.3.6 Identification of Strain F

The StrainSeeker program detected that genomes of *R. erythropolis* strains BG43, R138, CCM2595 and PR4 occur at a relative frequency of 31.35%, 26.56%, 21.10% and 21.00% on the whole genome sequence of strain F. The OrthoANIu based comparison of strain F with the genomes of *R. erythropolis* and *R. qingshengii* showed a percentage similarity value of 95.48% and 98.93%, respectively (Table 4.7). Based on the high OrthoANIu values, strain F could be reliably confirmed as a member of the species *R. qingshengii*. It was also observed that strain F shared comparable percentage similarity values of 95.48% and 95.51% with *R. erythropolis* strains NBRC 15567 and XP, respectively. The former is the type strain of the species, and the latter is a widely studied BDS capable strain with representative grade whole genome sequence. Interestingly, a high similarity of 95.48% was obtained for the pairwise comparison of the genomes of the type strains of *R. erythropolis* and *R. qingshengii,* as given in Table 4.7.

Table 4.7 OrthANIu values for pairwise comparison of strain F and the genomes of its closely related type strains

Metric	***R. erythropolis*** **NBRC 15567 vs Strain F**	***R. qingshengii*** **vs Strain F**	***R. erythropolis*** **XP vs Strain F**	***R. qingshengi*** **djl-6-2*i*** ***vs*** ***R. erythropolis*** **NBRC 15567**
OrthoANIu value (%)	**95.48**	**98.93**	**95.51**	**95.48**
Genome A length (bp)	6,557,580	6,517,800	7,225,680	6,517,800
Genome B length (bp)	6,386,220	6,386,220	6,386,220	6,557,580
Average aligned length (bp)	4,236,638	4,259,426	4,287,519	4,063,974
Genome A coverage (%)	64.61	65.35	59.34	62.35
Genome B coverage (%)	53.3	52.09	67.14	48.33

The draft genome sequence of strain F was compared with closely related representative genomes of type strains of *R. erythropolis* strains NBRC 15567 and XP, and *R. qingshengii* djl-6-2. The strain XP was included because it was a well studied BDS strain with the known whole genome sequence. Genome A refers to the reference genome obtained from the NCBI database, and Genome B refers to the WGS of strain F obtained in this research.

4.3.7 Confirmation of Strain 213E identity by OrthoANIu values

The whole-genome sequence of the *Gordonia desulfuricans* strain 213E obtained through this research was compared with that of the representative genome sequence of *Gordonia desulfuricans* NBRC 100010 already available in the NCBI Genome (RefSeq: NZ_BCNF00000000.1) which confirmed a very high similarity value of 99.94% as shown in Table 4.8.

Table 4.8 OrthoANIu values of pairwise comparison of whole-genome sequences of *G. desulfuricans* strain 213E

Metric	***G. desulfuricans* 213E NBRC Vs Strain 213E**
OrthoANIu value (%)	**99.94**
Genome A length (bp)	5,308,080
Genome B length (bp)	5,430,480
Average aligned length (bp)	3,909,026
Genome A coverage (%)	73.64
Genome B coverage (%)	71.98

Genome A refers to the reference genome available in the NCBI database, and Genome B refers to the WGS obtained for strain 213E in this research.

4.4 Discussion

Bacteria have efficient cellular mechanisms to adapt to the change in culture environment by the activity of global regulatory proteins, that generally act at the transcriptional level. The regulation often carried out by two-component systems (Hoch, 2000), that involve detection of extracellular signals and transduction of the signals into the cytosol. There might also be a cross-talk between the regulators and unknown proteins that influence the regulation (Yamamoto *et al.*, 2005). Additionally, σ factors also play essential roles in the transcription, where they allow RNA polymerase to be recruited at specific DNA sequences in the promoter regions. The BDS enzymes are produced by the bacteria as a stress-response (sulphate limitation in the growth medium). The *dsz* genes involved in DBT metabolism occur in most DBT degrading bacteria and with approximately 70% homology (Monticello, 2000). Despite the genetic homology, they exhibit differences in Dsz phenotypes in terms of specificity and substrate preferences (Kilbane II, 2006). Therefore, comprehensive knowledge of all the genes present in the organism could provide insights into its genetic preparedness for the activity. Mutation strategies could be developed based on this knowledge. Recently, the bioinformatics-based approach has been used to predict potential BDS strains using their genome sequences data without the need to grow them in the lab (Bhanjadeo *et al.*, 2018).

The whole-genome sequence obtained for the strains 248, F and *G. desulfuricans* 213E from MicrobesNG in this research are draft sequences at the contig level of assembly. Using ContEst16S program (Lee *et al.*, 2017), at least 7 of the publicly available rhodococci genome assemblies were found to be contaminated with sequences from other species. The quality of genome sequence is a critical factor for bioinformatics analyses, and therefore, contamination-free high-quality RefSeq or Representative sequences were used in this study. The genome sequences obtained in this research were not contaminated. The QUAST report on the quality of the genomes as described in Table 4.4 show that the sequences could be reliably used for bioinformatics analyses based on the N50 and N75 scores. It also shows the difference in the size of the genomes, with strain 248 being the largest (9.3 Mb), followed

by strain F (6.4 Mb) and 213E (5.5 Mb). It could be noted from Table 4.2 that the GC% of the strains 248 (66.82%) and strain F (62.51%) is similar to that of the respective reference genome sequence of selected for comparison from the Genbank.

4.4.1 Phylogenetic analysis of DBT *dsz* operon in strain F

All known bacteria with the ability to desulphurise DBT into 2HBP through the 4S pathway possess three desulphurisation enzymes. From Table 4.5, it could be noted that the *dsz* operon occurs predominantly in the members of phylum Actinobacteria and that the *dsz* genes appear more conserved among *Rhodococcus* species, in concordance with previously reported observations (Akhtar *et al.*, 2015; Bhanjadeo *et al.*, 2018; Duarte *et al.*, 2001). The infrequent distribution of the *dsz* operon is observed in Proteobacteria and Firmicutes phyla, which Bhanjadeo *et al.,* (2018) described as a scattered pattern of conservation of *dsz* genes owing to the horizontal gene transfer mode of origin of the operon.

4.4.2 Genetic arrangement of the BDS genes

The genes encoding the enzymes involved in the 4S pathway of DBT desulphurisation and their clustered arrangement in the genome of bacteria belonging to diverse genera are so well established, that Chauhan *et al.* (2014) grouped the *dsz* operons into 6 types. Quite the reverse, the genes encoding BT desulphurisation enzymes are not precisely known. It was suggested that in the case of organisms exhibiting BDS of both BT and DBT, the genes coding for the desulphurisation enzymes share sequence homology (Kirimura *et al.*, 2004). The only research done towards establishing the BT desulphurisation genes to date is the genetic analysis published by Wang *et al.* (2013a), who investigated the BT desulphurisation genes, by obtaining the draft whole-genome sequence of the BT desulphurising strain *Gordonia terrae* C-6 (Genbank: AQPW00000000.1) and comparing the transcriptomic profiles when cultured in the presence of BT or Na_2SO_4 as the sole sulphur source. Among the 135 upregulated genes which were mostly alkane sulfonate monooxygenases, they identified a gene cluster consisting of a desulfinase gene, a flavin-dependent monooxygenase gene and an alkanesulfonate monooxygenase gene whose products were functionally analogous to DszB,

DszC and DszA, and designated it as *bdsABC* operon (NCBI accession: KC831580). This highlighted the possible role of alkanesulfonate monooxygenases (*ssuD*) in desulphurisation, and therefore, in this study, the genomes of strains 248, 213E and F were searched for the occurrence of *ssuD* genes and genes encoding flavin-dependent monooxygenases.

The genome of strain 248 was more abundant than that of *G. desulfuricans* 213E and strain F in terms of the number of genes encoding monooxygenase enzymes. However, a clustered occurrence of monooxygenase and desulfinase genes was not detected in the draft genome sequence of strains 248 and *G. desulfuricans* 213E obtained in this study. Nevertheless, both strains contained the *soxC* gene encoding DszC and a gene encoding NADPH-dependent FMN reductase that underlies the modest ability of these strains to desulphurise DBT. Whereas the significant DBT desulphurisation activity of strain F, as shown in Figure 3.8 and Table 3.5, could be attributed to the several instances of the *dszABC* genes occurring in its genome along with several other flavin-dependent monooxygenases (Appendix 7.8).

One of the hypotheses of this research was that the location, arrangement and reoccurrence of the BDS related genes would be a determinant for the level of BDS activity expressed by them. This could be established by selecting the BDS strains with known genome sequence, studying the location and arrangement of *dsz* genes and other monooxygenases and correlate with the specific activity reported for them. Searching through the NCBI Genbank shows that BDS genes are often seen as individual annotations on WGS projects and occur scattered (not as a cluster). The number of annotations of *dszC* > *dszA* > *dszB*. At the time of writing, validly published whole-genome sequence data were available for only 8 other BDS capable actinobacteria, as shown in Table 4.1. The complete *dsz* operon was available for only 20 BDS strains listed in Table 4.5, among which only 10 included the sequences flanking the *dszABC* genes. The lack of diversity in the BDS strain WGS availability and complete *dszABC* operon limited the possibility of the investigation.

4.4.3 Genome sequence-based identification of strains 248 and F

The high speed and low cost of draft genome sequencing opens the door to *in silico* comparisons that are reminiscent of DDH. The power of genome sequences to resolve the taxonomy was utilised to confirm the identity of strains 248, F and 213E. It was previously found that digitally derived genome-to-genome distances showed a better correlation with 16S rRNA gene sequence distances than DDH values (Auch *et al.*, 2010). The pairwise average nucleotide identity (ANI) values were obtained in this research by the comparison of the genome sequences of the test strains with a representative genome of the closely related species. The closely related species were determined by 16S rRNA and *gyrB* gene sequence analysis, as described in Chapter 2, where the corresponding reference sequences were taken from the type strains for comparison. Stackebrandt (2011) suggested that the level of genome sequence identity (ANI value) among two strains must be >96% to be considered as an equivalent of a DDH similarity value of higher than 70%, which is the current standard for species delineation. It should be noted carefully that the 4% variation allowed between the genome sequences could encode for any phenotypic differences between them.

Strain 213E

The WGS of *G. desulfuricans* 213E obtained in this research (5104_213E) and the representative genome for the same strain existing in the NCBI Genome database (NZ_BCNF00000000.1) were subjected to ANI value calculation. This was done mainly for the purpose of evaluating the accuracy of the OrthoANIu algorithm for species delineation, and to confirm the reliability of the workflow followed in this research. A high genome coverage >70% used in the pairwise comparison and very high ANI value of 99.94% as shown in Table 4.8 is a promising result that happened as expected, as the two genome sequences were from the same strain. Besides serving as a further confirmation of the identity of the strain, the results show that the OrthoANIu values are reliable.

Strain 248

Based on 16S rRNA gene and *gyrB* sequence similarity analysis, the identity of strain 248 was determined (section 2.3.7) as *R. wratislaviensis,* as against its original classification as *R. opacus* which was the second closest relative (99.14% 16S rRNA similarity) and closely followed by *R. imtechensis* (98.53% 16S rRNA similarity). The highest quality reference whole-genome sequences available for these closely related strains was selected for evaluating the pairwise relatedness. Based on the ANI values shown in Table 4.6, it could be seen that the whole-genome sequence of strain 248 shares the highest relatedness (96.66%) to *R. opacus* rather than *R. wratislaviensis* (94.59%).

Moreover, it should be noted that when the reference genomes of *R. opacus*, *R. wratislaviensis* and *R. imtechensis* were compared to each other, a high percentage similarity (97.88%) was observed between *R. opacus* and *R. imtechensis*, which is higher than the threshold needed for being classified as a single species. Sangal *et al.*, (2016) reported the high similarity between these two strains and went on to suggest that *R. imtechensis* RKJ300^{T} represents a later heterotypic synonym of *R. opacus* DSM 43205^{T}. In the description of *R. imtechensis* type strain, Gosh and team (2006) acknowledge the close relationship with *R. opacus* and *R. wratislaviensis* but used the differences exhibited by them in the hydrolysis of tween 80 and utilisation of specific substrates as carbon sources (determined using Biolog GP2 Microplate) to present it as a novel species. As the genes associated with this phenotype are not clearly defined, it could only be assumed that they may have occurred in the unmatched regions of their genome.

Although programs like Read2Type and Strainseeker indicated the close relationship between strain 248 and *R. jostii*, the OrthoANIu value between them was less than 95% threshold. Based on higher OrthoANIu value of 96.66%, the strain 248 could be concluded as *R. opacus IEGM 248.*

Strain F

Based on the 16S rRNA and *gyrB* sequence similarity analysis, the identity of the strain F was determined as *R. degradans* (section 2.3.4). However, the whole-genome sequence data of *R. degradans* was not available at the time

of research and hence the genomes of other closely related species *R. qingshengii* was included in the analysis. It could be seen from Table 4.7, that the strain F shares ANI value of 95.48% with *R. erythropolis* and an even higher value of 98.93% with *R. qingshengii*. Interestingly the two reference genomes of *R. erythropolis* and *R. qingshengii* also shared a high 95.48% similarity, which hints that the differing 5.5% genome could contain the genes that led to the phenotypic differences based on which the latter strain was distinguished from the former and *R. baikonurensis* DSM 44587^T in order to be described as a new species (Xu *et al.*, 2007). The higher OrthoANIu value (98.93%) shared by strain F with *R. qingshengii* than with *R. erythropolis* strains (~95%) confirmed that strain F belonged to *R. qingshengii*.

The relationship between *R. degradans* and *R. qingshengii* could not be resolved through WGS based approaches because the *R. degradans* whole-genome sequence was not available. However, the conclusion that strain F is a *R. qingshengii* strain is further supported by other reports of *R. qingshengii* strains with BDS related characteristics. Licoln *et al.* (2015), reported *R. qingshengii* strain TUHH-2 harbouring BDS genes. Also, a desulphurisation-negative *R. qingshengii* CW25 has been genetically transformed into more enhanced BDS strain (Wang *et al.*, 2017), which indicated the natural adaptability of the species for BDS purposes. There is no known BDS capable *R. degradans* strain. This circumstantial evidence also favours the conclusion that strain F could be a *R. qingshengii* strain.

5.1.1 The scope of BDS technology

The race for commercial exploitation of BDS technology was on even before the development of the IGTS8 strain in 1990. Maliyantz (1935) reported the first known bacterial desulphurisation of petroleum oil with the accumulation of hydrogen sulphide. Since then, there have been significant investments in BDS research and development, and a series of US patents 2521761 (1950), 2574070 (1951), 2641564 (1953) and 2975103 (1961) filed for commercial application of BDS technology. The ENCHIRA Biotechnology Corporation (ENBC), USA (formerly, Energy Biosystems Corporation) pioneered a commercially applicable BDS technology, and interest in this was also shared by various other organisations including the Japanese Petroleum Energy Centre, Institute of Gas Technology, the Korean Advanced Institute of Science and Technology and Exxon Research & Engineering Company (Bachmann *et al.*, 2014; Borgne and Quintero, 2003).

BDS technology has some apparent advantages over current hydrodesulphurisation (HDS) processes. The energy requirements and associated CO_2 (greenhouse gas) emissions of BDS based processes are much less compared to the energy-intensive and expensive HDS process. This is because BDS processes operate at room temperature and normal pressure, much milder (and safer) process conditions than for HDS (Alves *et al.*, 2015; Linguist and Pacheco, 1999). Atlas *et al.* (1999) estimated that in the HDS process, the cost of reducing sulphur content from 200 to 50 ppm would be 4–5 times or higher than the cost of lowering the sulphur content from 500 to 200 ppm. Although there will be logistical costs for BDS associated with maintaining a contamination-free environment, sanitation handling, shipment and storage of bacterial cultures, this would be relatively cheaper than the operational costs of HDS (Kirkwood *et al.*, 2007b). In addition, the capital costs to set up a BDS process would reportedly be 50% of that for HDS. The advantages of BDS technology also include low operational costs (approximately 15% less than HDS) and also the ability to target thiophenes (Nuhu, 2013).

Despite these advantages and interest shown by multinational establishments, the BDS technology is not currently applied on an industrial scale because of practical and technical problems yet to be solved (these are summarised in Table 5.1). Indeed, the most favourable bioprocess method for a commercial-scale operation has not yet been established. Petro Star Incorporated, a fuel refinery in Alaska, obtained several rounds of funding from EBS and the US Department of Energy to develop enhanced enzymes and genetically modified host strains for BDS; however, scale-up of the process beyond pilot-scale was beset with issues.

Lab-scale research is often being reported in shake-flask level batch bioprocess experiments and continuous tank reactors and immobilised systems. Lately, advancements have been made in order to increase BDS efficiency by genetic engineering techniques (recombinant strains for overexpression of *dsz* genes) and classic microbiology techniques such as removal of the inhibiting end product from the culture, co-substrate addition (Martínez *et al.*, 2015), and to seamlessly integrate BDS technology as a complementing process to the existing processes (HDS technology). By these approaches, it has been possible to generate ultra-low sulphur diesel (Alves *et al.*, 2015; Nazari *et al.*, 2017). The lower capital and operating costs, minimal pollution and the possibilities to produce high valuable by-products, such as biosurfactants, mean that the BDS technology still a promising avenue of research (Alves *et al.*, 2015; Javadli and Klerk, 2012). As the desulphurisation enzymes require constant regeneration of cofactors for catalysing the oxidation-reduction reactions, whole cells were recommended for the BDS process to avoid problems of cofactor recycling and regeneration (Lin and Tao, 2017; Setti *et al.*, 1997). Strains with broader substrate ranges are also needed to address better the complicated mixture of chemicals present in petroleum. Characteristics such as reaction rate, emulsion formation and breakage, biocatalyst recovery, fuel recovery and both gas and liquid mass transport characteristics have a direct impact on the commercial prospects of the BDS technology in terms of setting up and operational costs. Hence there is a constant quest for new naturally occurring strains that would ideally require minimal if not no reformations for application in commercial scale.

Table 5.1 Critical factors influencing the large-scale application of BDS technology

Reasons favouring BDS technology	Factors of concern
Operation releases a very minimal amount of acid rain gases (Izumi *et al.*, 1994)	Operational costs associated with maintaining sterile conditions, and storage and use of viable microbial cells within the refinery environment to avoid contamination (McFarland, 1999)
High targeted enzymatic activity for the DBT and its derivatives and other desired compounds in the fuel are unaffected (Konishi *et al.*, 2000b)	Sulphate dependent repression of *dsz* genes and toxic inhibition by 2HBP produced as the end product of DBT desulphurisation (Alves and Paixão, 2011)
Lower capital and operating costs than HDS (Guobin *et al.*, 2006)	Conversion of sulphite to extracellular sulphates using sulfite oxidoreductase would affect *dsz* gene expression (Aggarwal *et al.*, 2012)
Yields ultra-low sulphur fuels (Soleimani *et al.*, 2007)	Cost of substrate in the bioprocess (Silva *et al.*, 2013)
More environmentally sustainable process as refractory organosulphur compounds are desulphurised under mild pressures and temperatures (Caro *et al.*, 2007)	Challenges with downstream oil/microbial biomass separation (Li *et al.*, 2009)
Yields high valuable by-products such as biosurfactants (Alves and Paixão, 2014a)	Multi-enzyme process requires whole-cell catalyst for full functionality (Alves and Paixão, 2014b)

With this in mind, the current study sought to identify new naturally occurring bacterial strains with BDS capabilities that have the potential to be future commercial strains in a full-scale BDS process. The focus of the research in this study was influenced by the knowledge that natural actinobacterial strains, notably rhodococci, are known to be metabolic powerhouses, capable of an array of transformation and degradation processes including biodesulphurisation. Therefore, the overarching aim of the research was to identify new naturally occurring, rhodococci and/or closely related actinobacteria, capable of biodesulphurisation of benzothiophene (BT) and/or dibenzothiophene (DBT). A total of 11 strains, isolated from various hydrocarbon-contaminated environments, and putatively identified as species belonging to the genus *Rhodococcus*, were selected from the IEGM culture collection in Russia. These aerobic, mesophilic organisms were tested for their ability to grow in mineral media, at neutral pH, with BT or DBT as the sole sulphur source. A routine procedure in the workflow to select BDS strains was adopted in this study: the strains were grown in defined selective media, MSM_{BT} or MSM_{DBT} with n-hexadecane or diesel, and the formation of phenol measured as an indirect indicator of BDS activity. During the study, the formulation for MSM was successfully modified to decrease the elemental composition and thereby reduce media associated costs.

To ensure confidence in the results of BDS screening, the type strain of *Gordonia desulfuricans*, strain 213E, was utilised as a positive control as this validly described species is well characterised and its benzothiophene desulphurising capabilities understood. However, it should be noted that its suitability for the BDS of diesel has not been explored previously. *Gordonia desulfuricans* 213E was also included in the phylogenetic studies, including single gene (16S rRNA and *gyr*B) and whole-genome sequence analyses, for the same purpose.

Prior to the current investigation, the taxonomic identities of the test strains were based on a limited number of phenotypic tests undertaken more than two decades ago, later followed by diagnostic PCR-based identifications using species-specific primer sets targeting variable regions of the 16S rRNA gene (Bell *et al.*, 1999). Confirmation of the taxonomic status of these organisms

was an important starting point to the current study to ensure that any strains found to be capable of BDS through this study are correctly classified. The basis of the previous identifications, however, was tenuous. The taxonomy of the genus *Rhodococcus* has transformed almost beyond recognition since these isolates were first characterised. Species membership has grown from around 11 to over 50, with many other species removed and reassigned to other genera or indeed combined with existing species. Phylogenetic analyses have played a significant role in shaping rhodococcal systematics and the approach to species identification more broadly. The significance of these developments in the context of the current study is that the previous identifications have no formal grounding; the criteria for classification as a member of this genus have evolved and there is now also a plethora of sequence data upon which to make phylogenetic identifications. Hence, the identifications made by the IEGM and by Bell *et al.* (1999) represent a low level of resolution by current taxonomic practices. Indeed, the species-specific primers designed by Bell and co-workers are no longer considered specific to the target species and based on *in silico* analyses would target multiple species, including those described at the time but for which 16S rRNA gene sequences were unavailable and more recently reported novel species.

Therefore, in this study, a phylogenetic study using high-quality 16S rRNA and *gyrB* gene sequences was successfully conducted in order to establish the correct genus and species assignments of all the test strains, and this enabled confirmation of the original genus assignment to *Rhodococcus* and enabled the putative species-level identifications made by the previous workers to be confirmed or amended. However, the findings of this study not only highlighted the comparative resolution of 16S rRNA versus *gyr*B gene analyses but also the limitations of single gene based phylogeny. Whole-genome sequencing taxonomy was therefore employed later in the study to firmly establish the species identities of those strains that were found capable of BDS.

The growth and BDS activities of the rhodococci isolates were investigated at bench-scale in shake flasks grown in batch mode. The overarching aim of this research study was successfully achieved; 2 of the 11 strains from the IEGM collection, namely strain 248 and strain F, were discovered to be robust BDS strains. Furthermore, these two strains were capable of BDS in MSM with a modified formulation with reduced concentrations and removal of some elements. The significance of this finding is that these strains demonstrate the potential for use in full-scale processes in which media costs will by necessity need to be kept to a minimum.

5.1.2 Identity and biodesulphurisation activity of strain 248

In the course of this study, strain 248 was found to be a benzothiophene (BT) desulphurising actinomycete belonging to genus *Rhodococcus*. A detailed phylogenetic analysis conducted using the high-quality 16S rRNA (1,495 bp) and *gyrB* (1,080 bp) gene sequences of strain 248, indicated that the strain was most closely related to the species *R. wratislaviensis* and most likely a member of this species. This finding disagreed with the original putative identification from the IEGM as R. *erythropolis* and that of Bell and co-workers, *R. opacus.* However, based on colony morphology and the WGS based OrthoANIu values from the WGS study, strain 248 was indeed confirmed as a member of the species *R. opacus*. *Rhodococcus opacus* is highly metabolically versatile and usually contain large plasmids, which explains the larger size of the draft genome (9.3 Mb) obtained in this research. When grown in a sulphate free mineral medium, *R. opacus* IEGM 248 exhibited desulphurisation of BT and DBT, with BT being the preferred substrate. This activity is similar to that of *Gordonia desulfuricans* 213E which was reported as a mainly BT desulphurising organism by Kim *et al.*, (1999) but later reported to be capable of DBT desulphurisation also and used for biological devulcanisation of ground rubber tyres (Tatangelo *et al.*, 2016). The occurrence of *dszC* in 6 different contigs in the draft genome of *G. desulfuricans* 213E obtained in this research further supports its broad range desulphurisation capability and hence its potential for future application.

5.1.3 Identity and biodesulphurisation activity of strain F

The current research study also revealed that the strain F (received as *Rhodococcus* sp.) is a DBT desulphurising bacterium. A detailed phylogenetic analysis conducted using the high-quality 16S rRNA (1,495 bp) and *gyrB* (1,066 bp) gene sequences of strain F, indicated that the strain was most closely related to the species *R. degradans* (100%) and closely followed by *R. qingshengii*. The16S rRNA and *gyrB* gene sequences of the type strains of these two validly described species also shared high similarity with each other, of 99.59% and 99.14%, respectively. However, based on OrthoANIu values obtained in the WGS analyses, the strain F was found to be a member of the species *R. qingshengii*.

In a biphasic medium containing 1% (v/v) *n*-hexadecane, the culture of strain F formed an emulsion, whereas the other isolates assigned to the same '*R. erythropolis'* group, also analysed in this study, namely strains 20, 213 and 508, grew as flocs suspended in a relatively less turbid medium (data not shown). This type of growth indicates the production of emulsifying surface-active agents by strain F for the purpose of degrading the oil substrate. However, this type of growth is unsuitable for a full-scale BDS process as it would lead to significant challenges for downstream separation of the oil and biomass. Hence, strain F is not the leading candidate.

When grown in a sulphate free mineral medium, *R. qingshengii* F exhibited the desulphurisation of BT and DBT, with DBT being the preferred substrate. The whole-genome sequence of the strain F revealed the presence of *dszABC* operon (type 1) which supported its DBT desulphurising phenotype.

The genomes of strains 248, F and 213E were searched for the occurrence of sulphur metabolism associated genes in order to identify their location on the genome and the nucleotide sequences flanking the genes. This would then enable identification of the protospacer adjacent motif (PAM) sequence upon which suitable guide RNAs could be designed for CRISPRi based gene silencing approach in the future. It should be pointed out that gene repression methodologies had not been developed for rhodococci at the outset of this study. During the course of the research, the CRISPR based genome

engineering technology witnessed rapid developments in terms of modified endonuclease enzymes requiring different PAM requirements (Choi and Lee, 2016; Selle and Barrangou, 2015; Wenyan Jiang, David Bikard, David Cox, Feng Zhang, 2013). Also, a preliminary molecular tool kit for rhodococci based on CRISPRi approach has been developed recently (DeLorenzo *et al.*, 2018; DeLorenzo and Moon, 2018). Therefore, genetic manipulation was not conducted, and this research was restricted to *in silico* investigation and identification of suitable target genes for suppression in the future.

5.1.4 Conclusion

In this research, a selection of putative rhodococci bacteria isolated from hydrocarbon-contaminated environments were screened for the capability of biodesulphurisation (BDS) of BT and/or DBT with a view to establishing strains with promising characteristics for potential commercial application of gas oil (diesel) desulphurisation. Two strains (strain 248 and strain F) with stable BDS capabilities were discovered and their taxonomic identities were confirmed as *Rhodococcus opacus* IEGM 248 and *Rhodococcus qingshengii* strain F. Interestingly, members of the latter species have not previously been reported capable of BDS, although a strain has been genetically modified to achieve DBT activity. The identity of the strains arrived at through the single gene (16S rRNA, *gyrB*) and whole-genome approaches (OrthoANI) only agree at the genus level, which reflects the complexity in the taxonomy of the genus *Rhodococcus* and the limitations of the taxonomic tools currently in use. There is discussion amongst the bacterial taxonomy community regarding current standards for species description and the value of including whole-genome sequence based approaches as standard within the polyphasic framework. Although both strains 248 and F exhibited deep desulphurisation of thiophenes present in diesel, based on the ease of recovery of the desulphurised oil observed in the case of strain 248, it was identified as the preferred potential candidate strain for commercial application. Through comprehensive analysis of the gene sequence data available about desulphurisation related genes, this research work highlights the limitations in the data currently available that hinder meaningful comparative analysis to understand the correlation between the organisation of *dsz* genes and the resulting BDS activity. The other strains

analysed in this study, although not suitable for application to BDS, were able to grow in MSM containing oil, which is an indicator of other noteworthy metabolic activities for biotechnological investigation but beyond the scope of this research.

5.1.5 Future Works

The test strains, especially 208, 488 and 508, can benefit from a thorough polyphasic taxonomic study combining biochemical tests and chemotaxonomic approaches (mycolic acid profiling), DNA-DNA Hybridisation and whole-genome sequencing, leading to full taxonomic description and notification of new species in the International Journal of Systematic and Evolutionary Microbiology (IJSEM) – recognised publication.

The qualitative data obtained in this research work regarding the BDS capability of the strains 248, F and 213E are supported by enough replicates and are consistent throughout. However, the extent of the quantitative data was limited by the access to analytical equipment such as ICP-OES or GC-MS. The HDS gas oil (diesel) used in this study was from a single batch obtained from PetroIneos refinery. The composition of diesel fuel is prone to geographic variations, and therefore, it would be interesting to study how the organisms fare with gas oil from different geo-locations. As the organism interacts to some extent with other naturally occurring organic compounds in the oil, a 2D-GC-MS analysis on the oil would reveal the changes caused in the fuel by the action of the bacteria and the subsequent effect on its calorific value. Moreover, in order to compare and assess the performance of the organisms accurately, BDS research can benefit from standardisation of the feedstock (gas oil) specification that the researchers use in biphasic systems. In terms of bioprocess, immobilising the BDS bacteria on 3D printed multichannel scaffolds that have distinct channels to supply the nutrients for growth and fuel feedstock for desulphurisation could also be a promising way forward.

The bioinformatic approach followed in this research relies on the sequence information available in the databases. The *dsz* genes show both a high degree of conservation and variation between bacterial species. However,

there are only a few complete *dsz* operon sequences available in the database and the ones covering the regulatory regions are further limited. A comparative analysis of *dsz* genes, regulatory regions and the respective enzymes of different strains can offer insights regarding bacterial desulphurisation at the molecular level. The quality of the whole-genome sequence of the strains 248 and strain F is currently at contig level of assembly. These should be updated to the complete genome sequence with a clear distinction of the chromosomal and plasmid sequences. As the *dsz* genes could be plasmid-borne in *R. opacus* strains, such detailed information would be helpful to develop refined mutation strategies. The CRISPRi based gene silencing technique is emerging for *Rhodococcus* sp. recently. When fully developed these techniques could be used for silencing of various sulphur metabolism genes to arrive at constitutive mutants suitable for commercial application.

5.1.6 Suppression of sulphate assimilation genes as the future approach for the development of BDS strains

The two main obstacles for the commercialisation of BDS technology are, firstly, sulphate induced repression of *dsz* gene expression, and secondly, reduced BDS rates of the strains. Several genetic engineering strategies have been attempted to augment the BDS activity as described in detail in section 1.7. However, none of the approaches has led to the commercial application of the technology yet. Many of the members of the order Actinomycetales, and particularly rhodococci are recalcitrant to conventional plasmid-based genetic manipulation methods and modification of the genes owing to the high GC content in the DNA and the diverse genomic contents, which hinder the progress in systematic metabolic engineering (Tong *et al.*, 2015). However, gene repression has not been reported as a strategy for the development of BDS strains so far.There are two gene silencing methods available for prokaryotic gene repression, antisense RNA and CRISPRi, neither of which has been validated and optimised for *Rhodococcus* spp until recently. DeLorenzo *et al.* (2018) developed a CRISPRi based prototype molecular tool kit for tunable and targeted gene repression in *Rhodococcus opacus* based on the CRISPRi principles developed for mycobacteria which utilised a different ortholog of the deactivated Cas9 called dCas$_{Sth1}$. The approach requires the

knowledge of the whole-genome sequence of the organism. The authors, however, have declared the need for further refinement and optimisation of the method. To date, CRISPR-Cas based genetic engineering has not been reported with reference to BDS genes or to manipulate any other metabolic aspect of BDS bacteria. In the future, these approaches should be adopted, and their success rates for the development of improved BDS strains determined.

The *dsz* operon is a coordinated unit of genetic expression containing a set of genes, namely a regulator, an operator, and structural genes (Dejaloud *et al.*, 2017). Therefore, repression of a repressor of the *dsz* operon could be a promising step forward to obtain constitutive expression of genes. Aggarwal and team (2011) performed several *in silico* experiments and metabolic flux balance analyses using the available genomic, biochemical and experimental data of the BDS strain *R. erythropolis* PR4 and found that components of the 4S pathway occur among these essential genes and reactions for the survival on medium containing DBT. In the absence of DBT, the components of the 4S pathway, *dsz* genes, are not essential. However, when DBT is the sole sulphur source, any malfunction in the components of the 4S pathway would lead to cell death. Moreover, the *dsz* operon is not strictly conserved among the BDS organisms which would demand the development of bespoke genetic manipulation for each type rather than a standardised kit. Therefore, the repression of the components of *dsz* operon poses more difficulty and risk.

It has been known that bacteria such as *E. coli*, *B. subtilis*, *P. aeruginosa* and *P. putida* can also utilise sulphonates such as taurine or isethionate as a sulphur source as do rhodococci. This capability is endowed by the gene cluster (*ssuEADCB* operon) that shares a high degree of sequence similarity between the bacteria, and generally, their expression in all the bacteria is repressed by sulphates (van Der Ploeg *et al.*, 1999). Therefore, a more rational genetic engineering approach to improve the BDS capability of the strains would be to repress the genes associated with sulphate uptake through CRISPRi approaches, forcing the organism to express *dsz* genes in order to utilise the thiophenes, irrespective of the presence of sulphates in the medium. As the BDS reaction occurs in the cytoplasm, the organism may not be

affected by the inability to obtain sulphates from the environment. The BDS activity of bacteria is a manifestation of the organisms' efforts to meet their sulphur requirements for survival. Suppressing other forms of sulphur intake guarantees that all sulphur requirements are met through the BDS reaction, which would result in improved BDS levels, thereby tackling the second problem in the BDS technology.

Genes encoding the primary sulphate transport complex of *M. tuberculosis* BCG are induced following treatment with hydrogen peroxide (*cysT*) and nutrient starvation (*cysA1, cysT, cysW, subI*), conditions that also induce the ATP sulfurylase genes *cysD* and *cysN*, as well as the bifunctional *cysNC* gene. It was observed that *cysA-* mutants phenotype was similar to that of the wild-type BCG strains and hence it was proposed that although CysA is essential for sulphate transport in rich media, there could be secondary sulphate transporter involved in sulphate uptake to compensate for its role. It was also observed that *cysH-* mutants, could not survive *in vitro* containing only sulphates and were dependent on the acquisition of reduced sulphur (cysteine or methionine) from the environment for survival (Schelle and Bertozzi, 2006). This highlights the importance of the *cysH* gene for sulphur assimilation and therefore, a potential target gene for repression for avoiding sulphate induced repression of the desulphurisation genes. As the *cysH-* mutant organisms could not utilise the inorganic sulphates, the *dsz* genes would be induced. Silencing of the *cysH* gene encoding the enzyme phosphoadenosine phosphosulfate reductase (EC 1.8.4.8) could be achieved using CRISPRi.

Methionine and cysteine are other sulphur sources that were reported to impact *dsz* gene expression. In the case of *R. erythropolis* KA2-5-1 and its mutants described by Tanaka *et al.* (2002), no DBT desulphurisation activity was observed for the wild type in the presence of either methionine or cysteine. However, a disruption in the *cbs* gene, that encodes for cystathionine β-synthase, an enzyme responsible for converting methionine to cysteine, enabled the *cbs-* mutants with desulphurising activity in the presence of methionine, albeit with slower growth, showing that methionine does not repress *dsz* genes. No desulphurisation activity or growth was expressed by the mutants in the presence of cysteine, and therefore they concluded that

cysteine repressed the *dsz* genes. However, Aggarwal and team (2011) proposed that by designing a strain that is unable to convert cysteine to any sulphur-containing metabolite that is essential for growth, then such a mutant would take DBT even in the presence of cysteine. By trial and error *in silico* mutation studies, they found that mutants with a simultaneous knockout of *cbs* gene and the gene for cysteine desulfhydrase forced the strain to use DBT, even in the presence of cysteine, showing that cysteine may not repress the *dsz* genes. *In silico* mutants with the simultaneous lack of genes for cysteine desulfhydrase, cystathionine γ-lyase, and cystathionine β-synthase could exhibit DBT desulphurisation in the presence of both methionine and cysteine in the medium. The *cbs* gene was found in genomes of strain 248 (contig_2 and contig_81), F (contig_2) and 213E (contig_33). The *cysH* gene occurred in the genomes of strain 248 (contig_11 and contig_74), F (contig_3) and 213E (contig_15). However, the implications of silencing the *cysH* or *cbs* genes remain unknown, and therefore the development of temporary gene silencing systems would be an ideal way forward.

Chapter 6 References

Abbad-Andaloussi, S., Lagnel, C., Warzywoda, M., and Monot, F. (2003). Multi-criteria comparison of resting cell activities of bacterial strains selected for biodesulfurization of petroleum compounds. *Enzyme and Microbial Technology*. **32**(3–4): 446–454.

Abbasian, F., Lockington, R., Megharaj, M., and Naidu, R. (2016). Identification of a new operon involved in desulfurization of dibenzothiophenes using a metagenomic study and cloning and functional analysis of the genes. *Enzyme and Microbial Technology*. **87**–**88:** 24–28.

Abin-Fuentes, A., Mohamed, M.E.-S.S., Wang, D.I.C., and Prather, K.L.J. (2013). Exploring the mechanism of biocatalyst inhibition in microbial desulfurization. *Applied and Environmental Microbiology*. **79**(24): 7807–7817.

Acharya, C., Sukla, L., and Misra, V. (2005). Biological elimination of sulphur from high sulphur coal by *Aspergillus*-like fungi. *Fuel*. **84**(12): 1597–1600.

Achtman, M. and Wagner, M. (2008). Microbial diversity and the genetic nature of microbial species. *Nature Reviews Microbiology*. **6**(6): 431.

Adachi, K., Katsuta, A., Matsuda, S., Peng, X., Misawa, N., Shizuri, Y., Kroppenstedt, R.M., Yokota, A., and Kasai, H. (2007). *Smaragdicoccus niigatensis* gen. nov., sp. nov., a novel member of the suborder *Corynebacterineae*. *International Journal of Systematic and Evolutionary Microbiology*. **57**(2): 297–301.

Adlakha, J., Singh, P., Ram, S.K., Kumar, M., Singh, M.P., Singh, D., Sahai, V., and Srivastava, P. (2016). Optimization of conditions for deep desulfurization of heavy crude oil and hydrodesulfurized diesel by *Gordonia* sp. IITR100. *Fuel*. **184:** 761–769.

Aggarwal, S., Karimi, I.A., Kilbane Ii, J.J., and Lee, D.Y. (2012). Roles of sulfite oxidoreductase and sulfite reductase in improving desulfurization by *Rhodococcus erythropolis*. *Molecular Biosystems*. **8**(10): 2724–2732.

Aggarwal, S., Karimi, I.A., and Lee, D.Y. (2011). Reconstruction of a genome-scale metabolic network of *Rhodococcus erythropolis* for desulfurization studies. *Molecular BioSystems*. **7**(11): 3122.

Ahmad, A., Chauhan, A.K., Javed, S., and Kumar, A. (2014). Desulfurization of thianthrene by a *Gordonia* sp. IITR100. *Biotechnology Letters*. **36**(11): 2209–2214.

Akhtar, N., Akhtar, K., and Ghauri, M.A. (2018). Biodesulfurization of thiophenic compounds by a 2-hydroxybiphenyl-resistant *Gordonia* sp. hs126-4n carrying *dszABC* genes. *Current Microbiology*. **75**(5): 597–603.

Akhtar, N., Ghauri, M.A., Anwar, M.A., and Akhtar, K. (2009). Analysis of the dibenzothiophene metabolic pathway in a newly isolated *Rhodococcus* spp. *FEMS Microbiology Letters*. **301**(1): 95–102.

Akhtar, N., Ghauri, M.A., Anwar, M.A., and Heaphy, S. (2015). Phylogenetic characterization and novelty of organic sulphur metabolizing genes of *Rhodococcus* spp. (Eu-32). *Biotechnology Letters*. **37**(4): 837–847.

Aliebrahimi, S., Raheb, J., Ebrahimipour, G., Bardania, H., Nurollah, M., and Aghajani, Z. (2015). Designing a new recombinant indigenous *Klebsiella oxytoca* ISA4 by cloning of *dsz* Genes. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*. **37**(19): 2056–2063.

Alkhalili, B.E., Yahya, A., Abrahim, N., and Ganapathy, B. (2017). Biodesulfurization of Sour Crude Oil. *Modern Applied Science*. **11**(9): 104.

Alomair, O., Jumaa, M., Alkoriem, A., and Hamed, M. (2015). Heavy oil viscosity and density prediction at normal and elevated temperatures. *Journal of Petroleum Exploration and Production Technology*. **6**(2):253-263

Alves, L., Marques, S., Matos, J., Tenreiro, R., and Gírio, F.M. (2008). Dibenzothiophene desulfurization by *Gordonia alkanivorans* strain 1B using recycled paper sludge hydrolyzate. *Chemosphere*. **70**(6): 967–973.

Alves, L., Melo, M., Mendonça, D., Simões, F., Matos, J., Tenreiro, R., and Gírio, F.M. (2007). Sequencing, cloning and expression of the *dsz* genes required for dibenzothiophene sulfone desulfurization from *Gordonia alkanivorans* strain 1B. *Enzyme and Microbial Technology*. **40**(6): 1598–1603.

Alves, L. and Paixão, Susana M (2014). Enhancement of dibenzothiophene desulfurization by *Gordonia alkanivorans* strain 1B using sugar beet molasses as alternative carbon source. *Applied Biochemistry and Biotechnology*. **172**(6): 3297–305.

Alves, L. and Paixão, Susana M. (2014). Fructophilic behaviour of *Gordonia alkanivorans* strain 1B during dibenzothiophene desulfurization process. *New Biotechnology*. **31**(1): 73–79.

Alves, L. and Paixão, S.M. (2011). Toxicity evaluation of 2-hydroxybiphenyl and other compounds involved in studies of fossil fuels biodesulphurisation. *Bioresource Technology*. **102**(19): 9162–9166.

Alves, L., Paixão, S.M., Pacheco, R., Ferreira, A.F., and Silva, C.M. (2015). Biodesulphurization of fossil fuels: energy, emissions and cost analysis. *RSC Advances*. **5**(43): 34047–34057.

Amais, R.S., Donati, G.L., and Nobrega, J.A. (2014). Sulfur determination in fuels by ICP-OES and ICP-MS to meet increasingly stricter legislation requirements. *Spectroscopy*. **29**(3): 24–33.

Amin, G.A. (2011). Integrated two-stage process for biodesulfurization of model Oil by vertical rotating immobilized cell reactor with the bacterium *Rhodococcus erythropolis*. *Journal of Petroleum & Environmental Biotechnology*. **2**:107.

Amin, G.A., Bazaid, S.A., and Abd El-Halim, M. (2013). A two-stage immobilized cell bioreactor with *Bacillus subtilis* and *Rhodococcus erythropolis* for the simultaneous production of biosurfactant and biodesulfurization of model oil. *Petroleum Science and Technology*. **31**(21): 2250–2257.

Ansari, F., Prayuenyong, P., and Tothill, I.E. (2007). Biodesulfurization of dibenzothiophene by *Shewanella putrefaciens* NCIMB 8768. *Journal of Biological Physics and Chemistry*. **7:** 75–78.

Antal, Z. and Puttonen, P. (2006). The effects of acid rain on anurans and the possibilities of protection in Hungary, focusing on the challenges of forest management. *North-Western Journal of Zoology*. **2**(1): 47–58.

Arensdorf, J.J., Loomis, A.K., DiGrazia, P.M., Monticello, D.J., and Pienkos, P.T. (2002). Chemostat approach for the directed evolution of biodesulfurization gain-of-function mutants. *Applied and Environmental Microbiology*. **68**(2): 691–698.

Arenskotter, M., Broker, D., and Steinbuchel, A. (2004). Biology of the metabolically diverse genus *Gordonia*. *Applied and Environmental Microbiology*. **70**(6): 3195–3204.

Arenskotter, M., Linos, A., Schumann, P., Kroppenstedt, R.M., and Steinbüchel, A. (2005). *Gordonia nitida* Yoon *et al.* 2000 is a later synonym of *Gordonia alkanivorans* Kummer *et al.* 1999. *International Journal of Systematic and Evolutionary Microbiology*. **55**(2): 695–697.

Arip, M.N.M., Heng, L.Y., Ahmad, M., and Hasbullah, S.A. (2013). Reaction of 2,6-dichloroquinone-4-chloroimide (Gibbs reagent) with permethrin - an optical sensor for rapid detection of permethrin in treated wood. *Chemistry Central Journal*. **7**(1): 122.

Armacost, K., Musila, J., Gathiaka, S., Ellis, H.R., and Acevedo, O. (2014). Exploring the catalytic mechanism of alkanesulfonate monooxygenase using molecular dynamics. *Biochemistry*. **53**(20): 3308–3317.

Asturias, J.A., Moore, E., Yakimov, M.M., Klatte, S., and Timmis, K.N. (1994). Reclassification of the polychlorinated biphenyl-degraders *Acinetobacter* sp. Strain P6 and *Corynebacterium* sp. Strain MB1 as *Rhodococcus globerulus*. *Systematic and Applied Microbiology*. **17**(2): 226–231.

Atlas, R.M. (2010). *Handbook of Microbiological Media, Fourth Edition*. 4th ed. Boca Raton: CRC Press.

Atlas, R.M., Boron, D.J., Deever, W.R., Johnson, A.R., McFarland, B.L., and Meyer, J.A. (1999). Method for removing organic sulfur from heterocyclic sulfur-containing organic compounds.

Auch, A.F., von Jan, M., Klenk, H.-P., and Göker, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Standards in genomic sciences*. **2**(1): 117–134.

Autry, A.R. and Fitzgerald, J.W. (1990). Sulfonate S: A major form of forest soil organic sulfur. *Biology and Fertility of Soils*. **10**(1): 50–56.

Aytar, P., Şam, M., and Çabuk, A. (2008). Microbial desulphurization of Turkish lignites by white rot fungi. *Energy & Fuels*. **22**(2): 1196–1199.

Babich, I. and Moulijn, J.A. (2003). Science and technology of novel processes for deep desulfurization of oil refinery streams: a review. *Fuel*. **82**(6): 607–631.

Bachmann, R.T., Johnson, A.C., and Edyvean, R.G.J. (2014). Biotechnology in the petroleum industry: An overview. *International Biodeterioration and Biodegradation*. **86:** 225–237.

Bahuguna, A., Lily, M.K., Munjal, A., Singh, R.N., and Dangwal, K. (2011). Desulfurization of dibenzothiophene (DBT) by a novel strain *Lysinibacillus sphaericus* DMT-7 isolated from diesel contaminated soil. *Journal of Environmental Sciences*. **23**(6): 975–982.

Balmes, J.R., Fine, J.M., and Sheppard, D. (1987). Symptomatic bronchoconstriction after short-term inhalation of sulfur dioxide. *The American Review of Respiratory Disease*. **136**(5): 1117–1121.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M.A., and Pevzner, P.A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*. **19**(5): 455–477.

Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-vaillant, N., Jacquard, C., Klenk, H.-P., Clément, C., Ouhdouch, Y., and van Wezel, G.P. (2016). Taxonomy, physiology, and natural products of actinobacteria. *Microbiology and Molecular Biology Reviews*. **80**(1): 1–43.

Bashir, S. and Liu, J.L. (2009). Construction and characterization of phenol-based sensor derived from colloidal chemistry. *Sensors and Actuators B*. **139:** 584–591.

Bataille, F., Lemberton, J.-L., Michaud, P., Pérot, G., Vrinat, M., Lemaire, M., Schulz, E., Breysse, M., and Kasztelan, S. (2000). Alkyldibenzothiophenes hydrodesulfurization-promoter effect, reactivity, and reaction mechanism. *Journal of Catalysis*. **191**(2): 409–422.

Bavykin, S.G., Lysov, Y.P., Zakhariev, V., Kelly, J.J., Jackman, J., Stahl, D.A., and Cherni, A. (2004). Use of 16S rRNA, 23S rRNA, and *gyrB* gene sequence analysis to determine phylogenetic relationships of *Bacillus cereus* group microorganisms. *Journal of clinical microbiology*. **42**(8): 3711–30.

Bayoumi, R.A., Mohmed, E.E.-D.M., El-Louboudy, S.S., and Hendawey, A.H. (2009). Microbial biodesulfurization of dibenzo-thiophene in Egyptian petroleum crude oil. *The Egyptian Journal of Experimental Biology (Botany)*. **5:** 157–167.

Beauvir, M. (2016). Sprint analysis of lubricating oils using the Thermo Scientific iCAP 7600 ICP-OES - Application Note 43161.

Bell, K.S., Kuyukina, M.S., Heidbrink, S., Philp, J.C., Aw, D.W.J., Ivshina, I.B., Christofi, A., and Christofi, N. (1999). Identification and environmental detection of *Rhodococcus* species by 16S rDNA-targeted PCR. *Journal of Applied Microbiology*. **87**(4): 472–480.

Bender, J. and Weigel, H.-J. (2011). Changes in atmospheric chemistry and crop health: A review. *Agronomy for Sustainable Development*. **31**(1): 81–89.

Bhanjadeo, M.M., Rath, K., Gupta, D., Pradhan, N., Biswal, S.K., Mishra, B.K., and Subudhi, U. (2018). Differential desulfurization of dibenzothiophene by newly identified MTCC strains: Influence of Operon Array. *PloS one*. **13**(3): e0192536.

Bhatia, S. and Sharma, D.K.K. (2010). Biodesulfurization of dibenzothiophene, its alkylated derivatives and crude oil by a newly isolated strain *Pantoea agglomerans* D23W3. *Biochemical Engineering Journal*. **50**(3): 104–109.

Bhuiya, M.-W. and Liu, C.-J. (2009). A cost-effective colorimetric assay for phenolic O-methyltransferases and characterization of caffeate 3-O-methyltransferases from *Populus trichocarpa*. *Analytical Biochemistry*. **384**(1): 151–158.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. **30**(15): 2114–2120.

Boniek, D., Figueiredo, D., dos Santos, A.F.B., and de Resende Stoianoff, M.A. (2015). Biodesulfurization: a mini review about the immediate search for the future technology. *Clean Technologies and Environmental Policy*. **17**(1): 29–37.

de Bont, J.A.. (1998). Solvent-tolerant bacteria in biocatalysis. *Trends in Biotechnology*. **16**(12): 493–499.

Bordoloi, N.K., Bhagowati, P., Chaudhuri, M.K., and Mukherjee, A.K. (2016). Proteomics and metabolomics analyses to elucidate the desulfurization pathway of *Chelatococcus* sp. V. Achal, ed. *PLOS ONE*. **11**(4): e0153547.

Bordoloi, N.K., Rai, S.K., Chaudhuri, M.K., and Mukherjee, A.K. (2014). Deep-desulfurization of dibenzothiophene and its derivatives present in diesel oil by a newly isolated bacterium *Achromobacter* sp. to reduce the environmental pollution from fossil fuel combustion. *Fuel Processing Technology*. **119:** 236–244.

Borgne, S. Le and Quintero, R. (2003). Biotechnological processes for the refining of petroleum. *Fuel Processing Technology*. **81**(2): 155–169.

Borole, A.P., Kaufman, E.N., Grossman, M.J., Minak-Bernero, V., Bare, R., and Lee, M.K. (2002). Comparison of the emulsion characteristics of *Rhodococcus erythropolis* and *Ecsherichia coli* soxc-5 cells expressing biodesulfurization genes. *biotechnology progress*. **18**(1): 88–93.

Borriss, R., Rueckert, C., Blom, J., Bezuidt, O., Reva, O., and Klenk, H.-P. (2011). Whole genome sequence comparisons in taxonomy. *Methods in Microbiology*. **38:** 409–436.

Bösmann, A., Datsevich, L., Jess, A., Lauter, A., Schmitz, C., and Wasserscheid, P. (2001). Deep desulfurization of diesel fuel by extraction with ionic liquids. *Chemical Communications*. (23): 2494–2495.

BP (2019). *BP Statistical Review of World Energy*.

Breeze, P. (2017). Combustion plant emissions: sulfur dioxide, nitrogen oxides, and acid rain, in *Electricity Generation and the Environment*. Elsevier, pp. 33–47.

Briglia, M., Rainey, F.A., Stackebrandt, E., Schraa, G., and Salkinoja-Salonen, M.S. (1996). *Rhodococcus percolatus* sp. nov., a bacterium degrading 2,4,6-Trichlorophenol. *International Journal of Systematic Bacteriology*. **46**(1): 23–30.

Brooks, S.L. and Van Hamme, J.D. (2012). Whole-genome shotgun sequence of *Rhodococcus* species strain JVH1. *Journal of Bacteriology*. **194**(19): 5492–5493.

Brosius, J., Palmer, M.L., Kennedy, P.J., and Noller, H.F. (1978). Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. **75**(10): 4801–5.

Butler III, R.R., Wang, J., Stark, B.C., and Pombert, J.-F. (2016). Complete genome sequences of two interactive moderate thermophiles, *Paenibacillus napthalenovorans* 32O-Y and *Paenibacillus* sp. 32O-W. *Genome Announcements*. **4**(1): e01717-15.

Butler, W.R., Floyd, M.M., Brown, J.M., Toney, S.R., Daneshvar, M.I., Cooksey, R.C., Carr, J., Steigerwalt, A.G., and Charles, N. (2005). Novel mycolic acid-containing bacteria in the family *Segniliparaceae* fam. nov., including the genus *Segniliparus* gen. nov., with descriptions of *Segniliparus rotundus* sp. nov. and *Segniliparus rugosus* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*. **55**(4): 1615–1624.

Calzada, J., Alcon, A., Santos, V.E., and Garcia-Ochoa, F. (2012). Extended kinetic model for DBT desulfurization using *Pseudomonas putida* CECT5279 in resting cells. *Biochemical Engineering Journal*. **66**(0): 52–60.

Camelo-Castillo, A., Benitez-Paez, A., Belda-Ferre, P., Cabrera-Rubio, R., and Mira, A. (2014). *Streptococcus dentisani* sp. nov., a novel member of the mitis group. *International Journal of Systematic and Evolutionary Microbiology*. **64**(Pt 1): 60–65.

Caro, A., Boltes, K., Leton, P., and Garcia-Calvo, E. (2008). Biodesulfurization of dibenzothiophene by growing cells of *Pseudomonas putida* CECT 5279 in biphasic media. *Chemosphere*. **73**(5): 663–669.

Caro, A., Boltes, K., Letón, P., and García-Calvo, E. (2007). Dibenzothiophene biodesulfurization in resting cell conditions by aerobic bacteria. *Biochemical Engineering Journal*. **35**(2): 191–197.

Caspermeyer, J. (2016). MEGA evolutionary software re-engineered to handle today's big data demands. *Molecular Biology and Evolution*. **33**(7): 1887.

Castorena, G., SuÃ¡rez, C., Valdez, I., Amador, G., FernÃ¡ndez, L., and Borgne, S. (2002). Sulfur-selective desulfurization of dibenzothiophene and diesel oil by newly isolated *Rhodococcus* sp. strains. *FEMS Microbiology Letters*. **215**(1): 157–161.

Ceniceros, A., Dijkhuizen, L., Petrusma, M., and Medema, M.H. (2017). Genome-based exploration of the specialized metabolic capacities of the genus *Rhodococcus*. *BMC Genomics*. **18**(1): 593.

Chan, J.F.W., Lam, J.Y.W., Woo, P.C.Y., Chen, J.H.K., Wu, A.K.L., Xiong, L., Leung, K.-W., Poon, R.W.S., Tsang, C.-C., and Lau, S.K.P. (2016). *Gordonia hongkongensis* sp. nov., isolated from blood culture and peritoneal dialysis effluent of patients in Hong Kong. *International Journal of Systematic and Evolutionary Microbiology*. **66**(10): 3942–3950.

Chan, J.Z.M.Z.-M., Halachev, M.R., Loman, N.J., Constantinidou, C., and Pallen, M.J. (2012). Defining bacterial species in the genomic era: Insights from the genus *Acinetobacter*. *BMC Microbiology*. **12**(1): 302.

Chang, J.H., Chang, Y.K., Ryu, H.W., and Chang, H.N. (2000). Desulfurization of light gas oil in immobilized-cell systems of *Gordona* sp. CYKS1 and *Nocardia* sp. CYKS2. *FEMS Microbiology Letters*. **182**(2): 309–312.

Chauhan, A.K., Ahmad, A., Singh, S.P., and Kumar, A. (2015). Biodesulfurization of benzonaphthothiophene by an isolated *Gordonia* sp. IITR100. *International Biodeterioration and Biodegradation*. **104:** 105–111.

Chauhan, A.K., Singh, S.P., and Kumar, A. (2014). Cloning, sequencing and characterization of the 'desulfurization gene cluster' from an isolated strain *Gordonia* sp. IITR100. *Journal of Applied Biotechnology*. **2**(2): 65.

Chen, H., Zhang, W.-J., Chen, J.-M., Cai, Y.-B., and Li, W. (2008). Desulfurization of various organic sulfur compounds and the mixture of DBT + 4,6-DMDBT by *Mycobacterium* sp. ZD-19. *Bioresource Technology*. **99**(9): 3630–3634.

Choi, K.R. and Lee, S.Y. (2016). CRISPR technologies for bacterial systems: Current achievements and future directions. *Biotechnology Advances*.

Choi, O.K., Cho, K.S., Ryu, H.W., and Chang, Y.K. (2003). Enhancement of phase separation by the addition of de-emulsifiers to three-phase (diesel oil/biocatalyst/aqueous phase) emulsion in diesel biodesulfurization. *Biotechnology Letters*. **25**(1): 73–77.

Choudhuri, S. and Choudhuri, S. (2014). Phylogenetic Analysis. In *Bioinformatics for Beginners*. Elsevier, pp. 209–218.

Christensen, H., Angen, Oø., Mutters, R., Olsen, J.E., and Bisgaard, M. (2000). DNA--DNA hybridization determined in micro-wells using covalent attachment of DNA. *International Journal of Systematic and Evolutionary Microbiology*. **50**(3): 1095–1102.

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D.R., da Costa, M.S., Rooney, A.P., Yi, H., Xu, X.-W., De Meyer, S., and Trujillo, M.E. (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*. **68**(1): 461–466.

Chun, J. and Rainey, F.A. (2014). Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. *International Journal of Systematic and Evolutionary Microbiology*. **64**(2): 316–324.

Claesson, M.J., Wang, Q., O'Sullivan, O., Greene-Diniz, R., Cole, J.R., Ross, R.P., and O'Toole, P.W. (2010). Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic Acids Research*. **38**(22): e200.

Coco, W.M., Levinson, W.E., Crist, M.J., Hektor, H.J., Darzins, A., Pienkos, P.T., Squires, C.H., and Monticello, D.J. (2001). DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nature Biotechnology*. **19**(4): 354–359.

Cohen-Bazire, G., Sistrom, W.R., and Stanier, R.Y. (1957). Kinetic studies of pigment synthesis by non-sulfur purple bacteria. *Journal of Cellular and Comparative Physiology*. **49**(1): 25–68.

Collins, M.D., Smida, J., Dorsch, M., and Stackebrandt, E. (1988). *Tsukamurella* gen. nov. Harboring *Corynebacterium paurometabolum* and *Rhodococcus aurantiacus*. *International Journal of Systematic Bacteriology*. **38**(4): 385–391.

Corma, A., Martínez, C., Ketley, G., and Blair, G. (2001). On the mechanism of sulfur removal during catalytic cracking. *Applied Catalysis A: General*. **208**(1–2): 135–152.

Crawford, D.L. and Gupta, R.K. (1990). Oxidation of dibenzothiophene by *Cunninghamella elegans*. *Current Microbiology*. **21**(4): 229–231.

Creason, A.L., Davis, E.W., Putnam, M.L., Vandeputte, O.M., and Chang, J.H. (2014). Use of whole genome sequences to develop a molecular phylogenetic framework for *Rhodococcus fascians* and the *Rhodococcus* genus. *Frontiers in Plant Science*. **5:** 406.

Czech, L., Huerta-Cepas, J., and Stamatakis, A. (2017). A critical review on the use of support values in tree viewers and bioinformatics toolkits. *Molecular Biology and Evolution*. **34**(6): 1535–1542.

Dai, Y., Shao, R., Qi, G., and Ding, B.-B. (2014). Enhanced dibenzothiophene biodesulfurization by immobilized cells of *Brevibacterium lutescens* in *n*-octane–water biphasic system. *Applied Biochemistry and Biotechnology*. **174**(6): 2236–2244.

Dastager, S.G., Mawlankar, R., Tang, S.-K., Krishnamurthi, S., Ramana, V. V., Joseph, N., and Shouche, Y.S. (2014). *Rhodococcus enclensis* sp. nov., a novel member of the genus *Rhodococcus*. *International Journal of Systematic and Evolutionary Microbiology*. **64**(Pt 8): 2693–2699.

Davoodi-Dehaghani, F., Vosoughi, M., and Ziaee, A.A. (2010). Biodesulfurization of dibenzothiophene by a newly isolated *Rhodococcus erythropolis* strain. *Bioresource Technology*. **101**(3): 1102–1105.

De-Ley, J., Cattoir, H., and Reynaerts, A. (1970). The quantitative measurement of DNA hybridization from renaturation rates. *European Journal of Biochemistry*. **12**(1): 133–142.

Debabov, V.G. (2010). Microbial desulfurization of motor fuel. *Applied Biochemistry and Microbiology*. **46**(8): 733–738.

Dejaloud, A., Vahabzadeh, F., and Habibi, A. (2017). *Ralstonia eutropha* as a biocatalyst for desulfurization of dibenzothiophene. *Bioprocess and Biosystems Engineering*. **40**(7): 969–980.

Deloger, M., El Karoui, M., and Petit, M.-A. (2009). A genomic distance based on MUM indicates discontinuity between most bacterial species and genera. *Journal of bacteriology*. **191**(1): 91–99.

DeLorenzo, D.M. and Moon, T.S. (2018). Selection of stable reference genes for RT-qPCR in *Rhodococcus opacus* PD630. *Scientific Reports*. **8**(1): 6019.

DeLorenzo, D.M., Rottinghaus, A.G., Henson, W.R., and Moon, T.S. (2018). Molecular toolkit for gene expression control and genome modification in *Rhodococcus opacus* PD630. *ACS Synthetic Biology*. **7**(2): 727–738.

Denis-Larose, C., Labbé, D., Bergeron, H., Jones, A.M., Greer, C.W., Al-Hawari, J., Grossman, M.J., Sankey, B.M., and Lau, P.C. (1997). Conservation of plasmid-encoded dibenzothiophene desulfurization genes in several rhodococci. *Applied and Environmental Microbiology*. **63**(7): 2915–9.

Denome, S.A., Oldfield, C., Nash, L.J., and Young, K.D. (1994). Characterization of the desulfurization genes from *Rhodococcus* sp. strain IGTS8. *Journal of Bacteriology*. **176**(21): 6707–6716.

Derikvand, P. and Etemadifar, Z. (2014). Improvement of biodesulfurization rate of alginate immobilized *Rhodococcus erythropolis* R1. *Jundishapur Journal of Microbiology*. **7**(3): 1–7.

Derikvand, P., Etemadifar, Z., and Biria, D. (2015). RSM optimization of dibenzothiophene biodesulfurization by newly isolated strain of *Rhodococcus erythropolis* PD1 in aqueous and biphasic systems. *Microbiology*. **84**(1): 65–72.

Dinamarca, M.A., Ibacache-Quiroga, C., Baeza, P., Galvez, S., Villarroel, M., Olivero, P., and Ojeda, J. (2010). Biodesulfurization of gas oil using inorganic supports biomodified with metabolically active cells immobilized by adsorption. *Bioresource Technology*. **101**(7): 2375–2378.

Dinamarca, M.A., Orellana, L., Aguirre, J., Baeza, P., Espinoza, G., Canales, C., and Ojeda, J. (2014a). Biodesulfurization of dibenzothiophene and gas oil using a bioreactor containing a catalytic bed with *Rhodococcus rhodochrous* immobilized on silica. *Biotechnology Letters*. **36**(8): 1649–1652.

Dinamarca, M.A., Rojas, A., Baeza, P., Espinoza, G., Ibacache-Quiroga, C., and Ojeda, J. (2014b). Optimizing the biodesulfurization of gas oil by adding surfactants to immobilized cell systems. *Fuel*. **116:** 237–241.

Domingos, D.F., Dellagnezze, B.M., Greenfield, P., Reyes, L.R., Melo, I.S., Midgley, D.J., and Oliveira, V.M. (2013). Draft genome sequence of the biosurfactant-producing bacterium *Gordonia amicalis s*train CCMA-559, isolated from petroleum-impacted sediment. *Genome Announcements*. **1**(6): e00894-13.

Dorobantu, L.S., Yeung, A.K.C., Foght, J.M., and Gray, M.R. (2004). Stabilization of oil-water emulsions by hydrophobic bacteria. *Applied and Environmental Microbiology*. **70**(10): 6333–6336.

Duan, X., Zhang, L., Zhou, D., Ji, K., Ma, T., Shui, W., Li, G., and Li, X. (2013). Crystallization and preliminary structural analysis of dibenzothiophene monooxygenase (DszC) from *Rhodococcus erythropolis*. *Acta Crystallographica Section F Structural Biology and Crystallization Communications*. **69**(6): 597–601.

Duarte, G.F., Rosado, A.S., Seldin, L., de Araujo, W., and van Elsas, J.D. (2001). Analysis of bacterial community structure in sulfurous-oil-containing soils and detection of species carrying dibenzothiophene desulfurization (*dsz*) genes. *Applied and Environmental Microbiology*. **67**(3): 1052–62.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*. **32**(5): 1792–7.

Edgar, R.C. (2010). Search and clutering orders of magnitude faster than BLAST. *Bioinformatics*. **26**(19): 2460–2461.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*. **7**(1): 1–26.

Eichhorn, E., van der Ploeg, J.R., and Leisinger, T. (1999). Characterization of a two-component alkanesulfonate monooxygenase from *Escherichia coli*. *Journal of Biological Chemistry*. **274**(38): 26639–26646.

El-Gendy, N.S. and Nassar, H.N. (2018a). *Biodesulfurization in Petroleum Refining*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

El-Gendy, N.S. and Nassar, H.N. (2018b). Challenges and opportunities. In *Biodesulfurization in Petroleum Refining*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 973–1117.

Ellis, H.R. (2011). Mechanism for sulfur acquisition by the alkanesulfonate monooxygenase system. *Bioorganic Chemistry*. **39**(5–6): 178–184.

Elmi, F., Etemadifar, Z., and Emtiazi, G. (2015). A novel metabolite (1,3-benzenediol, 5-hexyl) production by *Exophiala spinifera* strain FM through dibenzothiophene desulfurization. *World Journal of Microbiology and Biotechnology*. **31**(5): 813–821.

Erwin, K.N., Nakano, S., and Zuber, P. (2005). Sulfate-dependent repression of genes that function in organosulfur metabolism in *Bacillus subtilis* requires Spx. *Journal of Bacteriology*. **187**(12): 4042–4049.

Ettinger, M.B. and Ruchhoft, C.C. (1948). Determination of phenol and structurally related compounds by gibbs method. *Analytical Chemistry*. **20**(12): 1191–1196.

Ezaki, T., Hashimoto, Y., Takeuchi, N., Yamamoto, H., Liu, S.L., Miura, H., Matsui, K., and Yabuuchi, E. (1988). Simple genetic method to identify viridans group streptococci by colorimetric dot hybridization and fluorometric hybridization in microdilution wells. *Journal of Clinical Microbiology*. **26**(9): 1708–13.

Ezaki, T., Hashimoto, Y., and Yabuuchi, E. (1989). Fluorometric deoxyribonucleic acid-deoxyribonucleic acid hybridization in microdilution wells as an alternative to membrane filter hybridization in which radioisotopes are used to determine genetic relatedness among bacterial strains. *International Journal of Systematic Bacteriology*. **39**(3): 224–229.

Fabec, J.L. and Ruschak, M.L. (1985). Determination of nickel, vanadium, and sulfur in crudes and heavy crude fractions by inductively coupled argon plasma atomic emission spectrometry and flame atomic absorption spectrometry. *Analytical Chemistry*. **57**(9): 1853–1863.

Faison, B.D., Clark, T.M., Lewis, S.N., Ma, C.Y., Sharkey, D.M., and Woodward, C.A. (1991). Degradation of organic sulfur compounds by a coal-solubilizing fungus. *Applied Biochemistry and Biotechnology*. **28**–**29**(1): 237–251.

Fattouh, B. (2010). The dynamics of crude oil price differentials. *Energy Economics*. **32**(2): 334–342.

Fei, Q., Wewetzer, S.J., Kurosawa, K., Rha, C., and Sinskey, A.J. (2015). High-cell-density cultivation of an engineered *Rhodococcus opacus* strain for lipid production via co-fermentation of glucose and xylose. *Process Biochemistry*. **50**(4): 500–506.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. **39**(4): 783–791.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*. **17**(6): 368–376.

Felsenstein, J. (1978). The Number of Evolutionary Trees. *Systematic Zoology*. **27**(1): 27.

Ferreira, N. P. and Tracey, R. P. (1984) Numerical taxonomy of cholesterol degrading soil bacteria. *Journal of Applied Bacteriology*. **57**(3): 429–446.

Finnerty, W.R. (1992). The Biology and Genetics of the Genus *Rhodococcus*. *Annual Review of Microbiology*. **46**(1): 193–218.

Foght, J.M. (2004). Whole-cell bio-processing of aromatic compounds in crude oil and fuels. In R. Vazquez-Duhalt & R. Quintero-Ramirez, eds. *Petroleum Biotechnology: Developments and Perspectives*. Elsevier, pp. 145–175.

Folsom, B.R., Schieche, D.R., DiGrazia, P.M., Werner, J., and Palmer, S. (1999). Microbial desulfurization of alkylated dibenzothiophenes from a hydrodesulfurized middle distillate by *Rhodococcus* erythropolis I-19. *Applied and Environmental Microbiology*. **65**(11): 4967–72.

Fox, G.E., Wisotzkey, J.D., and Jurtshuk, P. (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *International Journal of Systematic Bacteriology*. **42**(1): 166–170.

Franchi, E., Rodriguez, F., Serbolisca, L., and De Ferra, F. (2003). Vector development, isolation of new promoters and enhancement of the catalytic activity of the dsz enzyme complex in *Rhodococcus* Sp. Strains. *Oil & Gas Science and Technology*. **58**(4): 515–520.

Fukushima, M., Kakinuma, K., and Kawaguchi, R. (2002). Phylogenetic analysis of *Salmonella*, *Shigella*, and *Escherichia coli* strains on the basis of the *gyrB* gene sequence. *Journal of Clinical Microbiology*. **40**(8): 2779–2785.

Furuya, T., Kirimura, K., Kino, K., and Usami, S. (2001). Thermophilic biodesulfurization of dibenzothiophene and its derivatives by *Mycobacterium phlei* WU-F1. *FEMS Microbiology Letters*. **204**(1): 129–133.

Galán, B., Díaz, E., and García, J.L. (2000). Enhancing desulphurization by engineering a flavin reductase-encoding gene cassette in recombinant biocatalysts. *Environmental Microbiology*. **2**(6): 687–694.

Gallagher, J.R., Olson, E.S., and Stanley, D.C. (1993). Microbial desulfurization of dibenzothiophene: a sulfur-specific pathway. *FEMS Microbiology Letters*. **107**(1): 31–35.

Gallardo, M., Ferrandez, A., De Lorenzo, V., Garcia, J., and Diaz, E. (1997). Designing recombinant *Pseudomonas* strains to enhance biodesulfurization. *Journal of Bacteriology*. **179**(22): 7156–7160.

Gao, J., Ellis, L.B.M., and Wackett, L.P. (2010). The University of Minnesota Biocatalysis/Biodegradation Database: improving public access. *Nucleic Acids Research*. **38**(suppl_1): d488–D491.

Van Der Geize, R. and Dijkhuizen, L. (2004). Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. *Current Opinion in Microbiology*. **7**(3): 255–261.

Geronimo, I., Nigam, S.R., and Payne, C.M. (2017). Desulfination by 2′-hydroxybiphenyl-2-sulfinate desulfinase proceeds via electrophilic aromatic substitution by the cysteine-27 proton. *Chemical Science*. **8**(7): 5078–5086.

Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J., Stackebrandt, E., de Peer, Y. Van, Vandamme, P., Thompson, F.L., and Swings, J. (2005). Opinion: Re-evaluating prokaryotic species. *Nature Reviews Microbiology*. **3**(9): 733–739.

Gherbawy, Y., El-Deeb, B., Hazzani, A. Al, Maher, A., and Shehata, A. (2016). Mycobiota of oil-contaminated soil samples and their abilities for dibenzothiophene desulfurization. *Geomicrobiology Journal*. **33**(7): 618–624.

Ghosh, A., Paul, D., Prakash, D., Mayilraj, S., and Jain, R.K. (2006). *Rhodococcus imtechensis* sp. nov., a nitrophenol-degrading actinomycete. *International Journal of Systematic and Evolutionary Microbiology*. **56**(8): 1965–1969.

Ghozikali, M.G., Mosaferi, M., Safari, G.H., and Jaafari, J. (2015). Effect of exposure to O_3 , NO_2, and SO_2 on chronic obstructive pulmonary disease hospitalizations in Tabriz, Iran. *Environmental Science and Pollution Research International*. **22**(4): 2817–2823.

Gibbs, H.D. (1927). Phenol tests: III. The indophenol test. *Journal of Biological Chemistry*. **72**(2): 649–664.

Gil, R., Silva, F.J., Peretó, J., and Moya, A. (2004). Determination of the core of a minimal bacterial gene set. *Microbiology and Molecular Biology Reviews*. **68**(3): 518–537.

Gilbert, S.C., Morton, J., Buchanan, S., Oldfield, C., and McRoberts, A. (1998). Isolation of a unique benzothiophene-desulphurizing bacterium, *Gordona* sp. strain 213E (NCIMB 40816): and characterization of the desulphurization pathway. *Microbiology*. **144**(Pt 9): 2545–2553.

Girard, G., Traag, B.A., Sangal, V., Mascini, N., Hoskisson, P.A., Goodfellow, M., and van Wezel, G.P. (2013). A novel taxonomic marker that discriminates between morphologically complex actinomycetes. *Open Biology*. **3**(10).

Goodfellow, M. (2014). The Family Nocardiaceae. In E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, & F. Thompson, eds. *The Prokaryotes: Actinobacteria*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 595–650.

Goodfellow, M. and Alderson, G. (1977). The actinomycete-genus *Rhodococcus*: a home for the 'rhodochrous' complex. *Journal of General Microbiology*. **100**(1977): 99–122.

Goodfellow, M., Alderson, G., and Chun, J. (1998). Rhodococcal systematics: problems and developments. *Antonie van Leeuwenhoek*. **74**(1/3): 3–20.

Goodfellow, M., Chun, J., Stackebrandt, E., and Kroppenstedt, R.M. (2002). Transfer of *Tsukamurella wratislaviensis* Goodfellow et a. 1995 to the genus *Rhodococcus* as *Rhodococcus wratislaviensis* comb. nov.. *International Journal of Systematic and Evolutionary Microbiology*. **52**(3): 749–755.

Goodfellow, M., Jones, A.L., Maldonado, L.A., and Salanitro, J. (2004). *Rhodococcus aetherivorans* sp. nov., a new species that contains methyl t-butyl ether-degrading actinomycetes. *Systematic and Applied Microbiology*. **27**(1): 61–65.

Goodfellow, M., Kämpfer, P., Busse, H.-J., Trujillo, M.E., Suzuki, K., Ludwig, W., and Whitman, W.B. eds. (2012). *Bergey's Manual® of Systematic Bacteriology*. New York, NY: Springer New York.

Gordon, R. E. and Mihm, J. M. (1962). Identification of *Nocardia caviae* (Erikson) nov. comb.*, *Annals of the New York Academy of Sciences*. **98**(3): 628–636.

Gray, K.A., Mrachko, G.T., and Squires, C.H. (2003). Biodesulfurization of fossil fuels. *Current Opinion in Microbiology*. **6**(3): 229–235.

Gray, K.A., Pogrebinsky, O.S., Mrachko, G.T., Xi, L., Monticello, D.J., and Squires, C.H. (1996). Molecular mechanisms of biocatalytic desulfurization of fossil fuels. *Nature Biotechnology*. **14**(13): 1705–1709.

Gray, M.R., Ayasse, A.R., Chan, E.W., and Veljkovic, M. (1995). Kinetics of hydrodesulfurization of thiophenic and sulfide sulfur in Athabasca bitumen. *Energy and Fuels*. **9**(3): 500–506.

Gray, P. H. H. and Thornton, H. G. (1928). Soil bacteria that decompose certain aromatic compounds, *Zentralblatt Fur Bakteriologie, Parasitenkunde, Infektionskrankheiten Und Hygiene.* **73**: 74–96.

Griffith, N. (2009). Acid rain in China. *New Scientist*. **203**(2726): 49.

Grossman, M.J., Lee, M.K., Prince, R.C., Minak-Bernero, V., George, G.N., and Pickering, I.J. (2001). Deep desulfurization of extensively hydrodesulfurized middle distillate oil by *Rhodococcus* sp. strain ECRD-1. *Applied and Environmental Microbiology*. **67**(4): 1949–1952.

Gunam, I.B.W. (2013). Biodesulfurization of dibenzothiophene and its derivatives using resting and immobilized cells of *Sphingomonas subarctica* T7b. *Journal of Microbiology and Biotechnology*. **23**(4): 473–482.

Gunam, I.B.W., Yaku, Y., Hirano, M., Yamamura, K., Tomita, F., Sone, T., and Asano, K. (2006). Biodesulfurization of alkylated forms of dibenzothiophene and benzothiophene by *Sphingomonas subarctica* T7b. *Journal of Bioscience and Bioengineering*. **101**(4): 322–327.

Guobin, S., Huaiying, Z., Jianmin, X., Guo, C., Wangliang, L., and Huizhou, L. (2006). Biodesulfurization of hydrodesulfurized diesel oil with *Pseudomonas delafieldii* R-8 from high density culture. *Biochemical Engineering Journal*. **27**(3): 305–309.

Gupta, N., Roychoudhury, P.K., and Deb, J.K. (2005). Biotechnology of desulfurization of diesel: prospects and challenges. *Applied Microbiology and Biotechnology*. **66**(4): 356–366.

Gupta, R.S., Lo, B., and Son, J. (2018). Phylogenomics and comparative genomic studies robustly support division of the genus *Mycobacterium* into an emended genus *Mycobacterium* and four novel genera. *Frontiers in microbiology*. **9:** 67.

Hagglblom, M.M., Nohynek, L.J., Palleroni, N.J., Kronqvist, K., Nurmiaho-Lassila, E.-L., Salkinoja-Salonen, M.S., Klatte, S., and Kroppenstedt, R.M. (1994). Transfer of polychlorophenol-degrading *Rhodococcus chlorophenolicus* (Apajalahti *et al.* 1986) to the genus *Mycobacterium* as *Mycobacterium chlorophenolicum* comb. nov. *International Journal of Systematic Bacteriology*. **44**(3): 485–493.

Haley, B., Grim, C., Hasan, N., Choi, S.Y., Chun, J., Brettin, T., Bruce, D., Challacombe, J., Detter, J., Han, C., Huq, A., and Colwell, R. (2010). Comparative genomic analysis reveals evidence of two novel *Vibrio* species closely related to *V. cholerae*. *BMC Microbiology*. **10**(1): 154.

Hamada, T., Maeda, Y., Matsuda, H., Sameshima, Y., Honda, K., Omasa, T., Kato, J., and Ohtake, H. (2009). Effect of cell-surface hydrophobicity on bacterial conversion of water-immiscible chemicals in two-liquid-phase culture systems. *Journal of Bioscience and Bioengineering*. **108**(2): 116–120.

Hamada, T., Sameshima, Y., Honda, K., Omasa, T., Kato, J., and Ohtake, H. (2008). A comparison of various methods to predict bacterial predilection for organic solvents used as reaction media. *Journal of Bioscience and Bioengineering*. **106**(4): 357–362.

Hamid, M.E., Minnikin, D.E., and Goodfellow, M. (1993). A simple chemical test to distinguish mycobacteria from other mycolic-acid-containing actinomycetes. *Journal of General Microbiology*. **139**(9): 2203–2213.

Van Hamme, J.D., Bottos, E.M., Bilbey, N.J., and Brewer, S.E. (2013). Genomic and proteomic characterization of *Gordonia* sp. NB4-1Y in relation to 6 : 2 fluorotelomer sulfonate biodegradation. *Microbiology*. **159**(Pt_8): 1618–1628.

Hamme, J.D. Van, Singh, A., and Ward, O.P. (2003). Recent advances in petroleum microbiology. *Microbiology and Molecular Biology Reviews*. **67**(4): 503–549.

Hartdegen, F.J. el, Coburn, J.M., and Roberts, R.L. (1984). Microbial desulfurization of petroleum. *Chemical Engineeríng Progress*. **80**(5): 63–67.

Hatzios, S.K. and Bertozzi, C.R. (2011). The regulation of sulfur metabolism in *Mycobacterium tuberculosis* M. Manchester, ed. *PLoS Pathogens*. **7**(7): e1002036.

Helmke, E. and Weyland, H. (1984). *Rhodococcus marinonascens* sp. nov., an Actinomycete from the Sea. *International Journal of Systematic Bacteriology*. **34**(2): 127–138.

Hino, T., Hamamoto, H., Suzuki, H., Yagi, H., Ohshiro, T., and Nagano, S. (2017). Crystal structures of TdsC, a dibenzothiophene monooxygenase from the thermophile *Paenibacillus* sp. A11-2, reveal potential for expanding its substrate selectivity. *Journal of Biological Chemistry*. **292**(38): 15804–15813.

Hirasawa, K., Ishii, Y., Kobayashi, M., Koizumi, K., and Maruhashi, K. (2001). Improvement of desulfurization activity in *Rhodococcus erythropolis* KA2-5-1 by genetic engineering. *Bioscience, Biotechnology, and Biochemistry*. **65**(2): 239–246.

Hoch, J.A. (2000). Two-component and phosphorelay signal transduction. *Current Opinion in Microbiology*. **3**(2): 165–170.

Hoffmann, M., Monday, S.R., Allard, M.W., Strain, E.A., Whittaker, P., Naum, M., Mccarthy, P.J., Lopez, J. V., Fischer, M., and Brown, E.W. (2012). *Vibrio caribbeanicus* sp. nov., isolated from the marine sponge *Scleritoderma cyanea*. *International Journal of Systematic and Evolutionary Microbiology*. **62**(8): 1736–1743.

Holder, J.W., Ulrich, J.C., DeBono, A.C., Godfrey, P.A., Desjardins, C.A., Zucker, J., Zeng, Q., Leach, A.L.B., Ghiviriga, I., Dancel, C., Abeel, T., Gevers, D., Kodira, C.D., Desany, B., Affourtit, J.P., Birren, B.W., and Sinskey, A.J. (2011). Comparative and functional genomics of *Rhodococcus opacus* PD630 for biofuels development. P. M. Richardson, ed. *PLoS Genetics*. **7**(9): e1002219.

Houalla, M., Broderick, D.H., Sapre, A.V., Nag, N.K., Beer, V.H.J. de, Gates, B.C., and Kwart, H. (1980). Hydrodesulfurization of methyl-substituted dibenzothiophenes catalyzed by sulfided Co-Mo/gamma-Al_2O_3. *Journal of Catalysis*. **61**(2): 523–527.

Huang, X. (1999). CAP3: A DNA sequence assembly program. *Genome Research*. **9**(9): 868–877.

Hucker, G.J. and Conn, H.J. (1923). Methods of gram staining. *Technical Bulletin of the New York State Agricultural Experiment Station*. **23**(March): 1–37.

Huse, S.M., Dethlefsen, L., Huber, J.A., Welch, D.M., Relman, D.A., and Sogin, M.L. (2008). Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genetics*. **4**(11).

Huse, S.M., Welch, D.M., Morrison, H.G., and Sogin, M.L. (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology*. **12**(7): 1889–1898.

Hutner, S.H., Provasoli, L., Schatz, A., and Haskins, C.P. (1950). Some approaches to the study of the role of metals in the metabolism of microorganisms. *Proceedings of the American Philosophical Society*. **94**(2): 152–170.

IMO (2016). IMO sets 2020 date for ships to comply with low sulphur fuel oil requirement. [online]. Available from: http://www.imo.org/en/MediaCentre/PressBriefings/Pages/MEPC-70-2020sulphur.aspx?platform=hootsuite [Accessed April 20, 2018].

Ishii, Y., Konishi, J., Okada, H., Hirasawa, K., Onaka, T., and Suzuki, M. (2000). Operon structure and functional analysis of the genes encoding thermophilic desulfurizing enzymes of *Paenibacillus* sp. A11-2. *Biochemical and Biophysical Research Communications*. **270**(1): 81–88.

Ishii, Y., Kozaki, S., Furuya, T., Kino, K., and Kirimura, K. (2005). Thermophilic biodesulfurization of various heterocyclic sulfur compounds and crude straight-run light gas oil fraction by a newly isolated strain *Mycobacterium phlei* WU-0103. *Current Microbiology*. **50**(2): 63–70.

Ivshina, I.B. (2012). Current situation and challenges of specialized microbial resource centres in Russia. *Microbiology*. **81**(5): 509–516.

Ivshina, I.B., Kamenskikh, T.N., and Liapunov, Y.E. (1994). *IEGM Catalogue of Strains of Regional Specialised Collection of Alkanotrophic Microorganisms*. Moscow: Russian Academy of Sciences.

Ivshina, I.B., Koblova, I. V, Bezmaternykh, G.I., Nesterenko, O.A., Kvasnikov, E.I., and Shkaruba, V. V (1986). [The identification of bacteria of the genus *Rhodococcus* by immunodiffusion]. *Mikrobiologicheskiĭ zhurnal*. **48**(2): 3–8.

Ivshina, I.B. and Kuyukina, M.S. (2018). Specialized microbial resource centers: a driving force of the growing bioeconomy. In S. K. Sharma & A. Varma, eds. *Microbial Resource Conservation: Conventional to Modern Approaches*. Cham: Springer International Publishing, pp. 111–139.

Ivshina, I.B. and Kuyukina, M.S. (2013). Turning Russian specialized microbial culture collections into resource centers for biotechnology. *Trends in Biotechnology*. **31**(11): 609–611.

Iwabuchi, N., Sunairi, M., Anzai, H., Nakajima, M., and Harayama, S. (2000). Relationships between colony morphotypes and oil tolerance in *Rhodococcus* rhodochrous. *Applied and Environmental Microbiology*. **66**(11): 5073–5077.

Izumi, Y., Ohshiro, T., Ogino, H., Hine, Y., and Shimao, M. (1994). Selective desulfurization of dibenzothiophene by *Rhodococcus erythropolis* D-1. *Applied and Environmental Microbiology*. **60**(1): 223–226.

Ja'fari, M., Ebrahimi, S.L., and Khosravi-Nikou, M.R. (2018). Ultrasound-assisted oxidative desulfurization and denitrogenation of liquid hydrocarbon fuels: A critical review. *Ultrasonics Sonochemistry*. **40**(Part A): 955–968.

Jaishankar, J., Singh, P., and Srivastava, P. (2017). Draft genome sequence of a biodesulfurizing bacterium, *Gordonia* sp. strain IITR100. *Genome announcements*. **5**(17): e00230-17.

Javadli, R. and Klerk, A. (2012). Desulfurization of heavy oil. *Applied Petrochemical Research*. **1**(1–4): 3–19.

Jiang, X., Yang, S., and Li, W. (2014). Biodesulfurization of model compounds and de-asphalted bunker oil by mixed culture. *Applied Biochemistry and Biotechnology*. **172**(1): 62–72.

Jiménez, G., Urdiain, M., Cifuentes, A., López-López, A., Blanch, A.R., Tamames, J., Kämpfer, P., Kolstø, A.B., Ramón, D., Martínez, J.F., Codoñer, F.M., and Rosselló-Móra, R. (2013). Description of *Bacillus toyonensis* sp. nov., a novel species of the *Bacillus cereus* group, and pairwise genome comparisons of the species of the group by means of ANI calculations. *Systematic and Applied Microbiology*. **36**(6): 383–391.

Jin, D., Kong, X., Jia, M., Yu, X., Wang, X., Zhuang, X., Deng, Y., and Bai, Z. (2017). *Gordonia phthalatica* sp. nov., a di-n-butyl phthalate-degrading bacterium isolated from activated sludge. *International Journal of Systematic and Evolutionary Microbiology*.

Jones, A.L., Brown, J.M., Mishra, V., Perry, J.D., Steigerwalt, A.G., and Goodfellow, M. (2004). *Rhodococcus gordoniae* sp. nov., an actinomycete isolated from clinical material and phenol-contaminated soil. *International Journal of Systematic and Evolutionary Microbiology*. **54**(2): 407–411.

Kalscheuer, R., Luftmann, H., Steinbüchel, A., Baumeister, D., and Wältermann, M. (2000). *Rhodococcus opacus* strain PD630 as a new source of high-value single-cell oil? Isolation and characterization of triacylglycerols and other storage lipids. *Microbiology*. **146**(5): 1143–1149.

Kamali, N., Tavallaie, M., Bambai, B., Karkhane, A.A., and Miri, M. (2010). Site-directed mutagenesis enhances the activity of NADH-FMN oxidoreductase (DszD) activity of *Rhodococcus erythropolis*. *Biotechnology Letters*. **32**(7): 921–927.

Kampfer, P., Dott, W., Martin, K., and Glaeser, S.P. (2014). *Rhodococcus defluvii* sp. nov., isolated from wastewater of a bioreactor and formal proposal to reclassify [*Corynebacterium hoagii*] and *Rhodococcus equi* as *Rhodococcus hoagii* comb. nov. *International Journal of Systematic and Evolutionary Microbiology*. **64**(Pt 3): 755–761.

Kampfer, P., Wellner, S., Lohse, K., Lodders, N., and Martin, K. (2013). *Rhodococcus cerastii* sp. nov. and *Rhodococcus trifolii* sp. nov., two novel species isolated from leaf surfaces. *International Journal of Systematic and Evolutionary Microbiology*. **63**(Pt 3): 1024–1029.

Kang, Y.-Q., Ming, H., Gonoi, T., Chen, Y., Cao, Y., Wang, Y.-Y., Cheng, J., Koga, T., Mikami, Y., and Li, W.-J. (2014). *Gordonia iterans* sp. nov., isolated from a patient with pneumonia. *International Journal of Systematic and Evolutionary Microbiology*. **64**(Pt 10): 3520–3525.

Kang, Y., Takeda, K., Yazawa, K., and Mikami, Y. (2009). Phylogenetic studies of *Gordonia* species based on *gyrB* and *secA1* gene analyses. *Mycopathologia*. **167**(2): 95–105.

Kasai, H., Watanabe, K., Gasteiger, E., Bairoch, A., Isono, K., Yamamoto, S., and Harayama, S. (1998). Construction of the *gyrB* database for the identification and classification of bacteria. *Genome informatics*. **9:** 13–21.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*. **30**(14): 3059–3066.

Katoh, K. and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*. **30**(4): 772–780.

Kaufman, E.N., Borole, A.P., Shong, R., Sides, J.L., and Juengst, C. (1999). Sulfur specificity in the bench-scale biological desulfurization of crude oil by *Rhodococcus* IGTS8. *Journal of Chemical Technology & Biotechnology*. **74**(10): 1000–1004.

Kawaguchi, H., Kobayashi, H., and Sato, K. (2012). Metabolic engineering of hydrophobic *Rhodococcus opacus* for biodesulfurization in oil-water biphasic reaction mixtures. *Journal of Bioscience and Bioengineering*. **113**(3): 360–366.

Kayser, K.J., Bielaga-Jones, B.A., Jackowski, K., Odusan, O., and Kilbane, J.J. (1993). Utilization of organosulphur compounds by axenic and mixed cultures of *Rhodococcus rhodochrous* IGTS8. *Journal of General Microbiology*. **139**(12): 3123–3129.

Kertesz, M.. A. and Wietek, C. (2001). Desulfurization and desulfonation: applications of sulfur-controlled gene expression in bacteria. *Applied Microbiology and Biotechnology*. **57**(4): 460–6.

Kertesz, M.A. (2000). Riding the sulfur cycle - metabolism of sulfonates and sulfate esters in Gram-negative bacteria. *FEMS Microbiology Reviews*. **24**(2): 135–175.

Khedkar, S. and Shanker, R. (2015). Isolation and classification of a soil actinomycete capable of sulphur-specific biotransformation of dibenzothiophene, benzothiophene and thianthrene. *Journal of Applied Microbiology*. **118**(1): 62–74.

Khosravinia, S., Mahdavi, M.A., Gheshlaghi, R., Dehghani, H., and Rasekh, B. (2018). Construction and characterization of a new recombinant vector to remove sulfate repression of *dsz* promoter transcription in biodesulfurization of dibenzothiophene. *Frontiers in Microbiology*. **9:** 1578.

Kilbane II, J.J. (2017). Biodesulfurization: how to make it work? *Arabian Journal for Science and Engineering*. **42**(1): 1–9.

Kilbane II, J.J. (2006). Microbial biocatalyst developments to upgrade fossil fuels. *Current Opinion in Biotechnology*. **17**(3): 305–314.

Kilbane II, J.J. and Jackowski, K. (1996). Biocatalytic detoxification of 2-chloroethyl ethyl sulfide. *Journal of Chemical Technology & Biotechnology*. **65**(4): 370–374.

Kilbane II, J.J. and Robbins, J. (2007). Characterization of the *dszABC* genes of *Gordonia amicalis* F.5.25.8 and identification of conserved protein and DNA sequences. *Applied Microbiology and Biotechnology*. **75**(4): 843–851.

Kilbane, J.J. (1989). Desulfurization of coal: the microbial solution. *Trends in Biotechnology*. **7**(4): 97–101.

Kilbane, J.J. (1990). Sulfur-specific microbial metabolism of organic compounds. *Resources, Conservation and Recycling*. **3**(2–3): 69–79.

Kilbane, J.J. and Le Borgne, S. (2004). Petroleum biorefining: the selective removal of sulfur, nitrogen, and metals. R. Vazquez-Duhalt & R. Quintero-Ramirez, eds. *Petroleum Biotechnology: Developments and Perspectives*. Elsevier, pp. 29–65.

Kilbane, J.J. and Jackowski, K. (1992). Biodesulfurization of water-soluble coal-derived material by *Rhodococcus rhodochrous* IGTS8. *Biotechnology and Bioengineering*. **40**(9): 1107–1114.

Kim, D., Choi, K.Y., Yoo, M., Zylstra, G.J., and Kim, E. (2018). Biotechnological potential of *Rhodococcus* biodegradative pathways. *Journal of Microbiology and Biotechnology*. **28**(7): 1037–1051.

Kim, M., Oh, H.-S., Park, S.-C., and Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*. **64**(Pt 2): 346–351.

Kim, S.B., Brown, R., Oldfield, C., Gilbert, S.C., and Goodfellow, M. (1999). *Gordonia desulfuricans* sp. nov., a benzothiophene-desulphurizing actinomycete. *International Journal of Systematic Bacteriology*. **49**(4): 1845–1851.

Kirby, B.M., Everest, G.J., and Meyers, P.R. (2010). Phylogenetic analysis of the genus *Kribbella* based on the *gryB* gene: proposal of a *gryB*-sequence threshold for species delineation in the genus Kribbella. *Antonie van Leeuwenhoek*. **97**(2): 131–142.

Kirimura, K., Furuya, T., Nishii, Y., Ishii, Y., Kino, K., and Usami, S. (2001). Biodesulfurization of dibenzothiophene and its derivatives through the selective cleavage of carbon-sulfur bonds by a moderately thermophilic bacterim *Bacillus subtilis* WU-S2B. *Journal of Bioscience and Bioengineering*. **91**(3): 262–266.

Kirimura, K., Furuya, T., Sato, R., Ishii, Y., Kino, K., and Usami, S. (2002). Biodesulfurization of naphthothiophene and benzothiophene through selective cleavage of carbon-sulfur bonds by *Rhodococcus* sp. strain WU-K2R. *Applied and Environmental Microbiology*. **68**(8): 3867–72.

Kirimura, K., Harada, K., Iwasawa, H., Tanaka, T., Iwasaki, Y., Furuya, T., Ishii, Y., and Kino, K. (2004). Identification and functional analysis of the genes encoding dibenzothiophene-desulfurizing enzymes from thermophilic bacteria. *Applied Microbiology and Biotechnology*. **65**(6): 703–713.

Kirkbright, G.F., Ward, A.F., and West, T.S. (1972). The determination of sulphur and phosphorus by atomic emission spectrometry with an induction-coupled high-frequency plasma source. *Analytica Chimica Acta*. **62**(2): 241–251.

Kirkwood, K.M., Andersson, J.T., Fedorak, P.M., Foght, J.M., and Gray, M.R. (2007a). Sulfur from benzothiophene and alkylbenzothiophenes supports growth of *Rhodococcus* sp. strain JVH1. *Biodegradation*. **18**(5): 541–549.

Kirkwood, K.M., Ebert, S., Foght, J.M., Fedorak, P.M., and Gray, M.R. (2005). Bacterial biodegradation of aliphatic sulfides under aerobic carbon- or sulfur-limited growth conditions. *Journal of Applied Microbiology*. **99**(6): 1444–1454.

Kirkwood, K.M., Foght, J.M., and Gray, M.R. (2007b). Selectivity among organic sulfur compounds in one- and two-liquid-phase cultures of *Rhodococcus* sp. strain JVH1. *Biodegradation*. **18**(4): 473–480.

Klappenbach, J.A., Goris, J., Vandamme, P., Coenye, T., Konstantinidis, K.T., and Tiedje, J.M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology*. **57**(1): 81–91.

Klatte, S., Kroppenstedt, R.M., and Rainey, F.A. (1994). *Rhodococcus opacus* sp.nov., an unusual nutritionally versatile *Rhodococcus*-species. *Systematic and Applied Microbiology*. **17**(3): 355–360.

Klenk, H.-P., Meier-Kolthoff, J.P., and Göker, M. (2014). Taxonomic use of DNA G+C content and DNA–DNA hybridization in the genomic age. *International Journal of Systematic and Evolutionary Microbiology*. **64**(2): 352–356.

Knobel, H.R., Egli, T., and Van Der Meer, J.R. (1996). Cloning and characterization of the genes encoding nitrilotriacetate monooxygenase of *Chelatobacter heintzii* ATCC 29600. *Journal of Bacteriology*. **178**(21): 6123–6132.

Ko, F.W.S., Tam, W., Wong, T.W., Chan, D.P.S., Tung, A.H., Lai, C.K.W., and Hui, D.S.C. (2007). Temporal relationship between air pollutants and hospital admissions for chronic obstructive pulmonary disease in Hong Kong. *Thorax*. **62**(9): 780–785.

Kobayashi, M., Onaka, T., Ishii, Y., Konishi, J., Takaki, M., Okada, H., Ohta, Y., Koizumi, K., and Suzuki, M. (2000). Desulfurization of alkylated forms of both dibenzothiophene and benzothiophene by a single bacterial strain. *FEMS Microbiology Letters*. **187**(2): 123–126.

Koch, C., Klatte, S., Schumann, P., Burghardt, J., Kroppenstedt, R.M., and Stackebrandt, E. (1995). Transfer of *Arthrobacter picolinophilus* Tate and Ensign 1974 to *Rhodococcus erythropolis*. *International Journal of Systematic Bacteriology*. **45**(3): 576–577.

Kodama, K., Umehara, K., Shimizu, K., Nakatani, S., Minoda, Y., and Yamada, K. (1973). Identification of microbial products from dibenzothiophene and its proposed oxidation pathway. *Agricultural and Biological Chemistry*. **37**(1): 45–50.

Konishi, J., Ishii, Y., Onaka, T., Ohta, Y., Suzuki, M., and Maruhashi, K. (2000a). Purification and characterization of dibenzothiophene sulfone monooxygenase and FMN-dependent NADH oxidoreductase from the thermophilic bacterium *Paenibacillus* sp. strain A11-2. *Journal of Bioscience and Bioengineering*. **90**(6): 607–613.

Konishi, J., Ishii, Y., Onaka, T., Okumura, K., and Suzuki, M. (1997). Thermophilic carbon-sulfur-bond-targeted biodesulfurization. *Applied and Environmental Microbiology*. **63**(8): 3164–9.

Konishi, J., Onaka, T., Ishii, Y., and Suzuki, M. (2000b). Demonstration of the carbon-sulfur bond targeted desulfurization of benzothiophene by thermophilic *Paenibacillus* sp. strain A11-2 capable of desulfurizing dibenzothiophene. *FEMS Microbiology Letters*. **187**(2): 151–154.

Konstantinidis, K. and Stackebrandt, E. (2013). Defining taxonomic ranks. In E. Rosenberg, E. DeLong, S. Lory, E. Stackebrandt, & F. Thompson, eds. *The Prokaryotes: Prokaryotic Biology and Symbiotic Associations*. New York: Springer, p. 2013:229.

Konstantinidis, K.T., Ramette, A., and Tiedje, J.M. (2006). The bacterial species definition in the genomic era. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. **361**(1475): 1929–40.

Konstantinidis, K.T. and Tiedje, J.M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences*. **102**(7): 2567–257.

Kredich, N.M. (1992). The molecular basis for positive regulation of cys promoters in *Salmonella typhimurium* and *Escherichia coli*. *Molecular Microbiology*. **6**(19): 2747–2753.

Kropp, K.G. and Fedorak, P.M. (1998). A review of the occurrence, toxicity, and biodegradation of condensed thiophenes found in petroleum. *Canadian Journal of Microbiology*. **44**(7): 605–22.

Kumari, A., Pinnaka, A.K., Singh, P.K., Chawla, N., and Korpole, S. (2015). *Rhodococcus lactis* sp. nov., an actinobacterium isolated from sludge of a dairy waste treatment plant. *International Journal of Systematic and Evolutionary Microbiology*. **65**(11): 4215–4220.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004). Versatile and open software for comparing large genomes. *Genome Biology*. **5**(2): r12.

Labana, S., Pandey, G., and Jain, R.K. (2005). Desulphurization of dibenzothiophene and diesel oils by bacteria. *Letters in Applied Microbiology*. **40**(3): 159–163.

Langdahl, B.R., Bisp, P., and Ingvorsen, K. (1996). Nitrile hydrolysis by *Rhodococcus erythropolis* BL1, an acetonitrile-tolerant strain isolated from a marine sediment. *Microbiology*. **142**(1): 145–154.

Larkin, M.J., Kulakov, L.A., and Allen, C.C.R. (2010). Genomes and Plasmids in *Rhodococcus*. In H. M. Alvarez, ed. *Biology of Rhodococcus*. Microbiology Monographs. Springer Berlin Heidelberg, pp. 73–90.

Larssen, T., Lydersen, E., Tang, D., He, Y., Gao, J., Liu, H., Duan, L., Seip, H.M., Vogt, R.D., Mulder, J., Shao, M., Wang, Y., Shang, H., Zhang, X., Solberg, S., Aas, W., Okland, T., Eilertsen, O., Angell, V., Li, Q., Zhao, D., Xiang, R., Xiao, J., and Luo, J. (2006). Acid rain in china. *Environmental Science and Technology*. **40**(2): 418–425.

Lee, I., Chalita, M., Ha, S.M., Na, S.I., Yoon, S.H., and Chun, J. (2017). ContEst16S: An algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. *International Journal of Systematic and Evolutionary Microbiology*. **67**(6): 2053–2057.

Lee, I., Ouk Kim, Y., Park, S.-C., and Chun, J. (2016). OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *International Journal of Systematic and Evolutionary Microbiology*. **66**(2): 1100–1103.

Lee, I.S., Bae, H.-S., Ryu, H.W., Cho, K.-S., and Chang, Y.K. (2008). Biocatalytic desulfurization of diesel oil in an air-lift reactor with immobilized *Gordonia nitida* CYKS1 Cells. *Biotechnology Progress*. **21**(3): 781–785.

Lee, K., Park, S.C., Yi, H., and Chun, J. (2013). *Flavobacterium limnosediminis* sp. nov., isolated from sediment of a freshwater lake. *International Journal of Systematic and Evolutionary Microbiology*. **63**(PART 12): 4784–4789.

Lee, M.K., Senius, J.D., and Grossman, M.J. (1995). Sulfur-specific microbial desulfurization of sterically hindered analogs of dibenzothiophene. *Applied and Environmental Microbiology*. **61**(12): 4362–4366.

Lee, W.C., Ohshiro, T., Matsubara, T., Izumi, Y., and Tanokura, M. (2006a). Crystal structure and desulfurization mechanism of 2'-hydroxybiphenyl-2-sulfinic acid desulfinase. *Journal of Biological Chemistry*. **281**(43): 32534–32539.

Lee, W.C., Ohshiro, T., Matsubara, T., Izumi, Y., and Tanokura, M. (2006b). Crystal structure and desulfurization mechanism of 2′-hydroxybiphenyl-2-sulfinic acid desulfinase. *Journal of Biological Chemistry*. **281**(43): 32534–32539.

Lei, B. and Tu, S.C. (1996). Gene overexpression, purification, and identification of a desulfurization enzyme from *Rhodococcus* sp. strain IGTS8 as a sulfide/sulfoxide monooxygenase. *Journal of Bacteriology*. **178**(19): 5699–5705.

Leigh, M.B., Prouzova, P., Mackova, M., Macek, T., Nagle, D.P., and Fletcher, J.S. (2006). Polychlorinated biphenyl (PCB)-degrading bacteria associated with trees in a PCB-contaminated site. *Applied and Environmental Microbiology*. **72**(4): 2331–2342.

León, R., Fernandes, P., Pinheiro, H.M., and Cabral, J.M.S. (1998). Whole-cell biocatalysis in organic media. *Enzyme and Microbial Technology*. **23**(7–8): 483–500.

Li, B., Furihata, K., Ding, L.-X., and Yokota, A. (2007a). *Rhodococcus kyotonensis* sp. nov., a novel actinomycete isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*. **57**(9): 1956–1959.

Li, C., McLinden, C., Fioletov, V., Krotkov, N., Carn, S., Joiner, J., Streets, D., He, H., Ren, X., Li, Z., and Dickerson, R.R. (2017). India is overtaking china as the world's largest emitter of anthropogenic sulfur dioxide. *Scientific Reports*. **7**(1): 14304.

Li, F, Xu, P., Feng, J., Meng, L., Zheng, Y., Luo, L., and Ma, C. (2005a). Microbial desulfurization of gasoline in a *Mycobacterium goodii* X7B immobilized-cell system. *Applied and Environmental Microbiology*. **71**(1): 276–281.

Li, F., Xu, P., Ma, C.Q., Luo, L.L., and Wang, X.S. (2003). Deep desulfurization of hydrodesulfurization-treated diesel oil by a facultative thermophilic bacterium *Mycobacterium* sp. X7B. *FEMS Microbiology Letters*. **223**(2): 301–307.

Li, F, Zhang, Z., Feng, J., Cai, X., and Xu, P. (2007b). Biodesulfurization of DBT in tetradecane and crude oil by a facultative thermophilic bacterium *Mycobacterium goodii* X7B. *Journal of Biotechnology*. **127**(2): 222–228.

Li, G. Q, Li, S., Zhang, M., Wang, J., Zhu, L., Liang, F., Liu, R., and Ma, T. (2008a). Genetic rearrangement strategy for optimizing the dibenzothiophene biodesulfurization pathway in *Rhodococcus erythropolis*. *Applied and Environmental Microbiology*. **74**(4): 971–976.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN]

Li, J, Zhao, G.-Z., Chen, H.-H., Qin, S., Xu, L.-H., Jiang, C.-L., and Li, W.-J. (2008b). *Rhodococcus cercidiphylli* sp. nov., a new endophytic actinobacterium isolated from a *Cercidiphyllum japonicum* leaf. *Systematic and Applied Microbiology*. **31**(2): 108–113.

Li, J., Zhao, G.-Z., Long, L.-J., Wang, F.-Z., Tian, X.-P., Zhang, S., and Li, W.-J. (2012). *Rhodococcus nanhaiensis* sp. nov., an actinobacterium isolated from marine sediment. *International Journal of Systematic and Evolutionary Microbiology*. **62**(Pt 10): 2517–2521.

Li, M., Squires, C., Monticello, D., and Childs, J. (1996). Genetic analysis of the *dsz* promoter and associated regulatory regions of *Rhodococcus erythropolis* IGTS8. *Journal of Bacteriology*. **178**(22): 6409–6418.

Li, S.-H., Jin, Y., Cheng, J., Park, D.-J., Kim, C.-J., Hozzein, W.N., Wadaan, M.A.M., Shu, W.-S., Ding, L.-X., and Li, W.-J. (2014). *Gordonia jinhuaensis* sp. nov., a novel actinobacterium, isolated from a VBNC (viable but non-culturable) state in pharmaceutical wastewater. *Antonie van Leeuwenhoek*. **106**(2): 347–356.

Li, W., Wang, M.-D., Chen, H., Chen, J.-M., and Shi, Y. (2006). Biodesulfurization of dibenzothiophene by growing cells of *Gordonia* sp. in batch cultures. *Biotechnology Letters*. **28**(15): 1175–1179.

Li, W, Zhang, Y., Wang, M.D., and Shi, Y. (2005b). Biodesulfurization of dibenzothiophene and other organic sulfur compounds by a newly isolated *Microbacterium* strain ZD-M2. *FEMS Microbiology Letters*. **247**(1): 45–50.

Li, X., Huang, Y., and Whitman, W.B. (2015). The relationship of the whole genome sequence identity to DNA hybridization varies between genera of prokaryotes. *Antonie van Leeuwenhoek*. **107**(1): 241–249.

Li, Y.-G., Gao, H.-S., Li, W.-L., Xing, J.-M., and Liu, H.-Z. (2009). In situ magnetic separation and immobilization of dibenzothiophene-desulfurizing bacteria. *Bioresource Technology*. **100**(21): 5092–5096.

Lin, B. and Tao, Y. (2017). Whole-cell biocatalysts by design. *Microbial Cell Factories*. **16**(1): 106.

Lincoln, S.A., Hamilton, T.L., Valladares Juárez, A.G., Schedler, M., Macalady, J.L., Müller, R., and Freeman, K.H. (2015). Draft genome sequence of the piezotolerant and crude oil-degrading bacterium *Rhodococcus qingshengii* strain TUHH-12. *Genome announcements*. **3**(2): e00268-15.

Linder, T. (2018). Assimilation of alternative sulfur sources in fungi. *World Journal of Microbiology and Biotechnology*. **34**(4): 51.

Linguist, L. and Pacheco, M. (1999). Enzyme-based diesel desulfurization process offers energy, CO_2 advantages. *Oil Gas Journal*. **97**(8): 45–48.

Liu, H, Zhang, Y., Liu, C., Fang, B., Li, C., Guan, X., Li, L., Wang, X., and Xiang, W. (2014a). *Rhodococcus kronopolitis* sp. nov., a novel actinobacterium isolated from a millipede (*Kronopolites svenhedind* Verhoeff). *Antonie van Leeuwenhoek*. **106**(6): 1207–1214.

Liu, S, Zhang, C., Su, T., Wei, T., Zhu, D., Wang, K., Huang, Y., Dong, Y., Yin, K., Xu, S., Xu, P., and Gu, L. (2014b). Crystal structure of DszC from *Rhodococcus* sp. XP at 1.79 Å. *Proteins*. **82**(9): 1708–1720.

Liu, Y., Ge, F., Chen, G., Li, W., Ma, P., Zhang, G., and Zeng, L. (2011). *Gordonia neofelifaecis* sp. nov., a cholesterol side-chain-cleaving actinomycete isolated from the faeces of *Neofelis nebulosa*. *International Journal of Systematic and Evolutionary Microbiology*. **61**(1): 165–169.

Liu, Z., Desantis, T.Z., Andersen, G.L., and Knight, R. (2008). Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucleic Acids Research*. **36**(18).

Loffler, F.E., Yan, J., Ritalahti, K.M., Adrian, L., Edwards, E.A., Konstantinidis, K.T., Muller, J.A., Fullerton, H., Zinder, S.H., and Spormann, A.M. (2013). *Dehalococcoides mccartyi* gen. nov., sp. nov., obligately organohalide-respiring anaerobic bacteria relevant to halogen cycling and bioremediation, belong to a novel bacterial class, Dehalococcoidia classis nov., order Dehalococcoidales ord. nov. and famil. *International Journal Of Systematic and Evolutionary Microbiology*. **63**(Pt 2): 625–635.

Lucena, T., Ruvira, M.A., Arahal, D.R., Macián, M.C., and Pujalte, M.J. (2012). *Vibrio aestivus* sp. nov. and *Vibrio quintilis* sp. nov., related to Marisflavi and Gazogenes clades, respectively. *Systematic and Applied Microbiology*. **35**(7): 427–431.

Ma, C.-Q.Q., Feng, J.-H.H., Zeng, Y.-Y.Y., Cai, X.-F.F., Sun, B.-P.P., Zhang, Z.-B. Bin, Blankespoor, H.D., and Xu, P. (2006). Methods for the preparation of a biodesulfurization biocatalyst using *Rhodococcus* sp. *Chemosphere*. **65**(1): 165–169.

Ma, J., Zhang, L., Wang, G., Zhang, S., Zhang, X., Wang, Y., Shi, C., Si, L., Zhao, H., Liu, F., Zhang, B., and Huang, X. (2017). *Rhodococcus gannanensis* sp. nov., a novel endophytic actinobacterium isolated from root of sunflower (Helianthus annuus L.). *Antonie van Leeuwenhoek*. **110**(9): 1113–1120.

Ma, T. (2010). The desulfurization pathway in *Rhodococcus*. In H. M. Alvarez, ed. *Biology of Rhodococcus*. Microbiology Monographs. Springer Berlin Heidelberg, pp. 207–230.

Ma, X., Sakanishi, K., and Mochida, I. (1994). Hydrodesulfurization reactivities of various sulfur compounds in diesel fuel. *Industrial and Engineering Chemistry Research*. **33**(2): 218–222.

Maass, D., Mayer, D.A., Moritz, D.E., Oliveira, D., de Souza, A.A.U., and Souza, S.M.A.G. (2015a). An evaluation of kinetic models in the biodesulfurization of synthetic oil by *Rhodococcus erythropolis* ATCC 4277. *Applied Biochemistry and Biotechnology*. **177**(3): 759–770.

Maass, D., Todescato, D., Moritz, D.E., Oliveira, J.V., Oliveira, D., Ulson de Souza, A.A., and Guelli Souza, S.M.A. (2015b). Desulfurization and denitrogenation of heavy gas oil by *Rhodococcus erythropolis* ATCC 4277. *Bioprocess and Biosystems Engineering*. **38**(8): 1447–1453.

Majidzadeh, M. and Fatahi-Bafghi, M. (2018). Current taxonomy of *Rhodococcus* species and their role in infections. *European Journal of Clinical Microbiology & Infectious Diseases*. **37**(11): 2045–2062.

Maliyantz, A. (1935). Bakteriologische methoden der entschwefelung von erdöl. *Azerbaidzhanskoe Neftyanoe Khoz*. **6:** 89–93.

Martens, M., Dawyndt, P., Coopman, R., Gillis, M., De Vos, P., and Willems, A. (2008). Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *International Journal of Systematic and Evolutionary Microbiology*. **58**(1): 200–214.

Martin, A.B., Alcon, A., Santos, V.E., and Garcia-Ochoa, F. (2005). Production of a biocatalyst of *Pseudomonas putida* CECT5279 for DBT biodesulfurization: influence of the operational conditions. *Energy and Fuels*. **19**(3): 775–782.

Martínez, I., García, J.L., and Díaz, E. (2016). Genetic engineering for removal of sulfur from fuel aromatic heterocycles. In *Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids: Production of Fuels and Chemicals*. Cham: Springer International Publishing, pp. 1–23.

Martínez, I., Santos, V.E., Alcon, A., and Garcia-ochoa, F. (2015). Enhancement of the biodesulfurization capacity of *Pseudomonas putida* CECT5279 by co-substrate addition. *Process Biochemistry*. **50**(1): 119–124.

Masters, B.C., Fan, V., and Ross, H.A. (2011). Species delimitation - a geneious plugin for the exploration of species boundaries. *Molecular Ecology Resources*. **11**(1): 154–157.

Matsui, T., Hirasawa, K., Koizumi, K., Maruhashi, K., and Kurane, R. (2001a). Optimization of the copy number of dibenzothiophene desulfurizing genes to increase the desulfurization activity of recombinant *Rhodococcus* sp. *Biotechnology Letters*. **23**(20): 1715–1718.

Matsui, T., Noda, K., Tanaka, Y., Maruhashi, K., and Kurane, R. (2002). Recombinant *Rhodococcus* sp. strain T09 can desulfurize DBT in the presence of inorganic sulfate. *Current Microbiology*. **45**(4): 240–244.

Matsui, T., Onaka, T., Maruhashi, K., and Kurane, R. (2001b). Benzo[b]thiophene desulfurization by *Gordonia rubropertinctus* strain T08. *Applied Microbiology and Biotechnology*. **57**(1–2): 212–215.

Matsui, T., Onaka, T., Tanaka, Y., Tezuka, T., Suzuki, M., and Kurane, R. (2000). Alkylated benzothiophene desulfurization by *Rhodococcus* sp. strain T09. *Bioscience, Biotechnology, and Biochemistry*. **64**(3): 596–599.

Matsui, T., Saeki, H., Shinzato, N., and Matsuda, H. (2007). Analysis of the 7.6-kb cryptic plasmid pNC500 from *Rhodococcus rhodochrous* B-276 and construction of *Rhodococcus-E. coli* shuttle vector. *Applied Microbiology and Biotechnology*. **74**(1): 169–175.

Mayilraj, S., Krishnamurthi, S., Saha, P., and Saini, H.S. (2006). *Rhodococcus kroppenstedtii* sp. nov., a novel actinobacterium isolated from a cold desert of the Himalayas, India. *International Journal of Systematic and Evolutionary Microbiology*. **56**(5): 979–982.

McFarland, B.L. (1999). Biodesulfurization. *Current Opinion in Microbiology*. **2**(3): 257–264.

McFarland, B.L., Boron, D.J., Deever, W., Meyer, J.A., Johnson, A.R., and Atlas, R.M. (1998). Biocatalytic sulfur removal from fuels: applicability for producing low sulfur gasoline. *Critical Reviews in Microbiology*. **24**(2): 99–147.

McLeod, M.P., Warren, R.L., Hsiao, W.W.L., Araki, N., Myhre, M., Fernandes, C., Miyazawa, D., Wong, W., Lillquist, A.L., Wang, D., Dosanjh, M., Hara, H., Petrescu, A., Morin, R.D., Yang, G., Stott, J.M., Schein, J.E., Shin, H., Smailus, D., Siddiqui, A.S., Marra, M.A., Jones, S.J.M., Holt, R., Brinkman, F.S.L., Miyauchi, K., Fukuda, M., Davies, J.E., Mohn, W.W., and Eltis, L.D. (2006). The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proceedings of the National Academy of Sciences*. **103**(42): 15582–15587.

Medley, A.J., Wong, C.-M., Thach, T.Q., Ma, S., Lam, T.-H., and Anderson, H.R. (2002). Cardiorespiratory and all-cause mortality after restrictions on sulphur content of fuel in Hong Kong: an intervention study. *The Lancet*. **360**(9346): 1646–1652.

Meesala, L., Balomajumder, C., Chatterjee, S., and Roy, P. (2008). Biodesulfurization of dibenzothiophene using recombinant *Pseudomonas* strain. *Journal of Chemical Technology and Biotechnology*. **83**(3): 294–298.

Meier-Kolthoff, J.P., Göker, M., Spröer, C., and Klenk, H.-P. (2013). When should a DDH experiment be mandatory in microbial taxonomy? *Archives of Microbiology*. **195**(6): 413–418.

de Menezes, C.B.A., Afonso, R.S., de Souza, W.R., Parma, M., de Melo, I.S., Zucchi, T.D., and Fantinatti-Garboggini, F. (2016). *Gordonia didemni* sp. nov. an actinomycete isolated from the marine ascidium *Didemnum* sp. *Antonie van Leeuwenhoek*. **109**(2): 297–303.

Minh, B.Q., Nguyen, M.A.T., and von Haeseler, A. (2013). Ultrafast Approximation for Phylogenetic Bootstrap. *Molecular Biology and Evolution*. **30**(5): 1188–1195.

Minnikin, D.E., Hutchinson, I.G., Caldicott, A.B., and Goodfellow, M. (1980). Thin-layer chromatography of methanolysates of mycolic acid-containing bacteria. *Journal of Chromatography A*. **188**(1): 221–233.

Mishra, S., Panda, S., Pradhan, N., Satapathy, D., Biswal, S.K., and Mishra, B.K. (2017). Insights into DBT biodegradation by a native *Rhodococcus* strain and its sulphur removal efficacy for two Indian coals and calcined pet coke. *International Biodeterioration & Biodegradation*. **120**, 124–134.

Mistry, S. and Wenthold, P.G. (2018). Mass spectrometric detection of the Gibbs reaction for phenol analysis. *Journal of Mass Spectrometry*. **53**(10): 947–953.

Mohamed, M.E.-S., Al-Yacoub, Z.H., and Vedakumar, J. V. (2015). Biocatalytic desulfurization of thiophenic compounds and crude oil by newly isolated bacteria. *Frontiers in Microbiology*. **6**.

Mohebali, G. and Ball, A.S. (2008). Biocatalytic desulfurization (BDS) of petrodiesel fuels. *Microbiology*. **154**(Pt 8): 2169–2183.

Mohebali, G. and Ball, A.S. (2016). Biodesulfurization of diesel fuels – Past, present and future perspectives. *International Biodeterioration and Biodegradation*. **110**, 163–180.

Mohebali, G., Ball, A.S., Kaytash, A., and Rasekh, B. (2008). Dimethyl sulfoxide (DMSO) as the sulfur source for the production of desulfurizing resting cells of *Gordonia alkanivorans* RIPI90A. *Microbiology*. **154**(Pt 3): 878–885.

Mohebali, G., Ball, A.S., Rasekh, B., and Kaytash, A. (2007). Biodesulfurization potential of a newly isolated bacterium, *Gordonia alkanivorans* RIPI90A. *Enzyme and Microbial Technology*. **40**(4): 578–584.

Monticello, D.J. (2000). Biodesulfurization and the upgrading of petroleum distillates. *Current Opinion in Biotechnology*. **11**(6): 540–546.

Monticello, D.J. (1998). Riding the fossil fuel biodesulfurization wave. *ChemTech*. **28**(7): 38–45.

Monticello, D.J. and Finnerty, W.R. (1985). Microbial desulfurization of fossil fuels. *Annual Review of Microbiology*. **39**(1): 371–389.

Morse, M.E. (1912). A study of the diphtheria group of organisms by the biometrical method. *Journal of Infectious Diseases*. **11:** 253–285.

Mu, T., Yang, M., Zhao, J., Sharshar, M.M., Tian, J., and Xing, J. (2017). Improvement of desulfurizing activity of haloalkaliphilic *Thialkalivibrio versutus* SOB306 with the expression of *Vitreoscilla* hemoglobin gene. *Biotechnology Letters*. **39**(3): 447–452.

Murray, R.G.E., Brenner, D.J., Colwell, R.R., De Vos, P., Goodfellow, M., Grimont, P.A.D., Pfennig, N., Stackebrandt, E., and Zavarzin, G.A. (1990). Report of the ad hoc committee on approaches to taxonomy within the Proteobacteria. *International Journal of Systematic Bacteriology*. **40**(2): 213–215.

Na, K., Kuroda, A., Takiguchi, N., Ikeda, T., Ohtake, H., and Kato, J. (2005). Isolation and characterization of benzene-tolerant *Rhodococcus opacus* strains. *Journal of Bioscience and Bioengineering*. **99**(4): 378–382.

NAEI (National Atmospheric Emmissions Inverntory) report titled ‘Air Pollution in the UK 2017’, dated September 2018, accessed online from the website of UK Department for Environment, Food & Rural Affairs https://uk-air.defra.gov.uk/assets/documents/annualreport/air_pollution_uk_2017_issue_1.pdf

Naito, M., Kawamoto, T., Fujino, K., Kobayashi, M., Maruhashi, K., and Tanaka, A. (2001). Long-term repeated biodesulfurization by immobilized *Rhodococcus erythropolis* KA2-5-1 cells. *Applied Microbiology and Biotechnology*. **55**(3): 374–378.

Nakayama, N., Matsubara, T., Ohshiro, T., Moroto, Y., Kawata, Y., Koizumi, K., Hirakawa, Y., Suzuki, M., Maruhashi, K., Izumi, Y., and Kurane, R. (2002). A novel enzyme, 2′-hydroxybiphenyl-2-sulfinate desulfinase (DszB), from a dibenzothiophene-desulfurizing bacterium *Rhodococcus erythropolis* KA2-5-1: gene overexpression and enzyme characterization. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. **1598**(1–2): 122–130.

Nazari, F., Kefayati, M., and Raheb, J. (2017). The study of biological technologies for the removal of sulfur compounds. *Journal of Sciences*. **28**(3): 205–219.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*. **32**(1): 268–274.

Nguyen, T.M. and Kim, J. (2016). *Rhodococcus pedocola* sp. nov. and *Rhodococcus humicola* sp. nov., two antibiotic-producing actinomycetes isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*. **66**(6): 2362–2369.

Nicolaou, S.A., Gaida, S.M., and Papoutsakis, E.T. (2010). A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation. *Metabolic Engineering*. **12**(4): 307–331.

Nishiuchi, Y., Baba, T., and Yano, I. (2000). Mycolic acids from *Rhodococcus*, *Gordonia*, and *Dietzia*. *Journal of Microbiological Methods*. **40**(1): 1–9.

Noda, K.-I., Kogure, T., Irisa, S., Murakami, Y., Sakata, M., and Kuroda, A. (2008). Enhanced dibenzothiophene biodesulfurization in a microchannel reactor. *Biotechnology Letters*. **30**(3): 451–454.

Noda, K., Watanabe, K., and Maruhashi, K. (2003). Cloning of a rhodococcal promoter using a transposon for dibenzothiophene biodesulfurization. *Biotechnology letters*. **25**(3): 1875–82.

Nomura, N., Takada, M., Okada, H., Shinohara, Y., Nakajima-Kambe, T., Nakahara, T., and Uchiyama, H. (2005). Identification and functional analysis of genes required for desulfurization of alkyl dibenzothiophenes of *Mycobacterium* sp. G3. *Journal of Bioscience and Bioengineering*. **100**(4): 398–402.

Van Noorden, R., Maher, B., and Nuzzo, R. (2014). The top 100 papers. *Nature*. **514**(7524): 550–553.

Novoselova, N.., Ananina, L.., Kuyukina, M.., and Ivshina, I.B. (2011). Taxonomic revision using 16s rRNA gene analysis of actinobacterial cultures maintained in the regional specialized collection of alkanotrophic microorganisms. *Herald of Ural Medical Academic Science (In Russian)*. **4/1**(38): 43–44.

Nuhu, A.A. (2013). Bio-catalytic desulfurization of fossil fuels: A mini review. *Reviews in Environmental Science and Biotechnology*. **12**(1): 9–23.

O'Donnell, A.G., Goodfellow, M., and Minnikin, D.E. (1982). Lipids in the classification of *Nocardioides*: Reclassification of *Arthrobacter simplex* (Jensen) lochhead in the genus *Nocardioides* (Prauser) emend. O'Donnell *et al.* as *Nocardioides simplex* comb. nov. *Archives of Microbiology*. **133**(4): 323–329.

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C.M., Goldfarb, T., Gupta, T., Haft, D., Hatcher, E., Hlavina, W., Joardar, V.S., Kodali, V.K., Li, W., Maglott, D., Masterson, P., McGarvey, K.M., Murphy, M.R., O'Neill, K., Pujar, S., Rangwala, S.H., Rausch, D., Riddick, L.D., Schoch, C., Shkeda, A., Storz, S.S., Sun, H., Thibaud-Nissen, F., Tolstoy, I., Tully, R.E., Vatsan, A.R., Wallin, C., Webb, D., Wu, W., Landrum, M.J., Kimchi, A., Tatusova, T., DiCuccio, M., Kitts, P., Murphy, T.D., and Pruitt, K.D. (2016). Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*. **44**(D1): d733–D745.

Oberreuter, H., Charzinski, J., and Scherer, S. (2002). Intraspecific diversity of *Brevibacterium linens*, *Corynebacterium glutamicum* and *Rhodococcus erythropolis* based on partial 16S rDNA sequence analysis and Fourier-transform infrared (FT-IR) spectroscopy. *Microbiology*. **148**(5): 1523–1532.

Ohshiro, T. (1996). Desulfurization of dibenzothiophene derivatives by whole cells of *Rhodococcus erythropolis* H-2. *FEMS Microbiology Letters*. **142**(1): 65–70.

Ohshiro, T., Hirata, T., Hashimoto, I., and Izumi, Y. (1996). Characterization of dibenzothiophene desulfurization reaction by whole cells of *Rhodococcus erythropolis* H-2 in the presence of hydrocarbon. *Journal of Fermentation and Bioengineering*. **82**(6): 610–612.

Ohshiro, T., Hirata, T., and Izumi, Y. (1995). Microbial desulfurization of dibenzothiophene in the presence of hydrocarbon. *Applied Microbiology and Biotechnology*. **44**(1–2): 249–252.

Ohshiro, T., Ohkita, R., Takikawa, T., Manabe, M., Lee, W.C., Tanokura, M., and IZUMI, Y. (2007). Improvement of 2'-hydroxybiphenyl-2-sulfinate desulfinase, an enzyme involved in the dibenzothiophene desulfurization pathway, from *Rhodococcus erythropolis* KA2-5-1 by site-directed mutagenesis. *Bioscience, Biotechnology, and Biochemistry*. **71**(11): 2815–2821.

Ohshiro, T., Suzuki, K., and Izumi, Y. (1997). Dibenzothiophene (DBT) degrading enzyme responsible for the first step of DBT desulfurization by *Rhodococcus erythropolis* D-1: purification and characterization. *Journal of Fermentation and Bioengineering*. **83**(3): 233–237.

Okada, H., Nomura, N., Nakahara, T., and Maruhashi, K. (2002). Analysis of dibenzothiophene metabolic pathway in *Mycobacterium* strain G3. *Journal of Bioscience and Bioengineering*. **93**(5): 491–497.

Okai, M., Lee, W.C., Guan, L.-J., Ohshiro, T., Izumi, Y., and Tanokura, M. (2017). Crystal structure of dibenzothiophene sulfone monooxygenase BdsA from *Bacillus subtilis* WU-S2B. *Proteins: Structure, Function, and Bioinformatics*. **85**(6): 1171–1177.

Oldfield, C., Pogrebinsky, O., Simmonds, J., Olson, E.S., and Kulpa, C.F. (1997). Elucidation of the metabolic pathway for dibenzothiophene desulphurization by *Rhodococcus* sp. strain IGTS8 (ATCC 53968). *Microbiology*. **143**(1): 2961–2973.

del Olmo, C.H., Alcon, A., Santos, V.E., and Garcia-Ochoa, F. (2005). Modeling the production of a *Rhodococcus erythropolis* IGTS8 biocatalyst for DBT biodesulfurization: influence of media composition. *Enzyme and Microbial Technology*. **37**(2): 157–166.

Omori, T., Monna, L., Saiki, Y., and Kodama, T. (1992). Desulfurization of dibenzothiophene by *Corynebacterium* sp. strain SY1. *Applied and Environmental Microbiology*. **58**(3): 911–915.

Omori, T., Saiki, Y., Kasuga, K., and Kodama, T. (1995). Desulfurization of alkyl and aromatic sulfides and sulfonates by dibenzothiophene-desulfurizing *Rhodococcus* sp. strain SY1. *Bioscience, Biotechnology, and Biochemistry*. **59**(7): 1195–1198.

Onaka, T., Kobayashi, M., Ishii, Y., Konishi, J., and Maruhashi, K. (2001a). Selective cleavage of the two C-S bonds in asymmetrically alkylated dibenzothiophenes by *Rhodococcus erythropolis* KA2-5-1. *Journal of Bioscience and Bioengineering*. **92**(1): 80–82.

Onaka, T., Konishi, J., Ishii, Y., and Maruhashi, K. (2001b). Desulfurization characteristics of thermophilic *Paenibacillus* sp. strain A11-2 against asymmetrically alkylated dibenzothiophenes. *Journal of Bioscience and Bioengineering*. **92**(2): 193–196.

Oren, A. and Garrity, G.M. (2014). Then and now: a systematic review of the systematics of prokaryotes in the last 80 years. *Antonie van Leeuwenhoek*. **106**(1): 43–56.

Pace, N.R., Sapp, J., and Goldenfeld, N. (2012). Phylogeny and beyond: Scientific, historical, and conceptual significance of the first tree of life. *Proceedings of the National Academy of Sciences of the United States of America*. **109**(4): 1011–1018.

Pacheco, M.A., Lange, E.A., Pienkos, P.T., Yu, L.Q., Rouse, M.P., Lin, Q., and Linguist, L.K. (1999). Recent advances in desulfurization of diesel fuel. In *National Petrochemical and Refiners Association*. San Antonio, Texas, USA, p. AM-99-27.

Paje, M.L.F., Neilan, B.A., and Couperwhite, L. (1997). A *Rhodococcus* species that thrives on medium saturated with liquid benzene. *Microbiology*. **143**(9): 2975–2981.

Pallagi, I., Toro, A., and Farkas, O. (1994). Mechanism of the gibbs reaction . 3. indophenol formation via radical electrophilic aromatic-substitution ($S_{RE}Ar$) on phenols. *Journal of Organic Chemistry*. **59**(17): 6543–6557.

Pan, J., Wu, F., Wang, J., Yu, L., Khayyat, N.H., Stark, B.C., and Kilbane, J.J. (2013). Enhancement of desulfurization activity by enzymes of the *Rhodococcus dsz* operon through coexpression of a high sulfur peptide and directed evolution. *Fuel*. **112:** 385–390.

Papizadeh, M., Ardakani, M.R., Ebrahimipour, G., and Motamedi, H. (2010). Utilization of dibenzothiophene as sulfur source by *Microbacterium* sp. NISOC-06. *World Journal of Microbiology and Biotechnology*. **26**(7): 1195–1200.

Parravicini, F., Brocca, S., and Lotti, M. (2016). Evaluation of the conformational stability of recombinant desulfurizing enzymes from a newly isolated *Rhodococcus* sp. *Molecular Biotechnology*. **58**(1): 1–11.

Patel, S.B., Kilbane, J.J., and Webster, D.A. (1997). Biodesulphurisation of dibenzothiophene in hydrophobic media by *Rhodococcus* sp. strain IGTS8. *Journal of Chemical Technology and Biotechnology*. **69**(1): 100–106.

Pathak, A., Chauhan, A., Blom, J., Indest, K.J., Jung, C.M., Stothard, P., Bera, G., Green, S.J., and Ogram, A. (2016). Comparative genomics and metabolic analysis reveals peculiar characteristics of *Rhodococcus opacus* strain M213 particularly for naphthalene degradation M.-J. Virolle, ed. *PLOS ONE*. **11**(8): e0161032.

Peeters, K. and Willems, A. (2011). The *gyrB* gene is a useful phylogenetic marker for exploring the diversity of *Flavobacterium* strains isolated from terrestrial and aquatic habitats in Antarctica. *FEMS Microbiology Letters*. **321**(2): 130–140.

Peng, B. and Zhou, Z. (2016). Study on growth characteristic and microbial desulfurization activity of the bacterial stain MP12. *Biochemical Engineering Journal*. **112:** 202–207.

Petcu, M., Schaare, P.N., and Cook, C.J. (2004). Propofol-imprinted membranes with potential applications in biosensors. *Analytica Chimica Acta*. **504**(1): 73–79.

Piddington, C., Kovacevich, B., and Rambosek, J. (1995). Sequence and molecular characterization of a DNA region encoding the dibenzothiophene desulfurization operon of *Rhodococcus* sp. strain IGTS8. *Applied and Environmental Microbiology*. **61**(2): 468–475.

van der Ploeg, J., Eichhorn, E., and Leisinger, T. (2001). Sulfonate-sulfur metabolism and its regulation in *Escherichia coli*. *Archives of Microbiology*. **176**(1–2): 1–8.

van Der Ploeg, J.R., Iwanicka-Nowicka, R., Bykowski, T., Hryniewicz, M.M., and Leisinger, T. (1999). The *Escherichia coli* ssuEADCB gene cluster is required for the utilization of sulfur from aliphatic sulfonates and is regulated by the transcriptional activator Cbl. *The Journal of biological chemistry*. **274**(41): 29358–65.

Poirier, S., Rué, O., Peguilhan, R., Coeuret, G., Zagorec, M., Champomier-Vergès, M.-C., Loux, V., and Chaillou, S. (2018). Deciphering intra-species bacterial diversity of meat and seafood spoilage microbiota using *gyrB* amplicon sequencing: a comparative analysis with 16S rDNA V3-V4 amplicon sequencing G.-J. Nychas, ed. *PLOS ONE*. **13**(9): e0204629.

Quijano, G., Hernandez, M., Thalasso, F., Muñoz, R., and Villaverde, S. (2009). Two-phase partitioning bioreactors in environmental biotechnology. *Applied Microbiology and Biotechnology*. **84**(5): 829–846.

Rainey, F.A., Burghardt, J., Kroppenstedt, R.M., Klatte, S., and Stackebrandt, E. (1995a). Phylogenetic analysis of the genera *Rhodococcus* and *Nocardia* and evidence for the evolutionary origin of the genus *Nocardia* from within the radiation of *Rhodococcus* species. *Microbiology*. **141**(2): 523–528.

Rainey, F.A., Klatte, S., Kroppenstedt, R.M., and Stackebrandt, E. (1995b). *Dietzia*, New Genus Including *Dietzia maris* comb. nov., Formerly *Rhodococcus maris*. *International Journal of Systematic Bacteriology*. **45**(1): 32–36.

Rambosek, J., Piddington, C.S., Kovacevich, B.R., Young, K.D., and Denome, S.A. (1999). Recombinant DNA encoding a desulfurization biocatalyst.

Rashidi, L., Mohebali, G., and Rasekh, B. (2006). Biodesulfurization of dibenzothiophene and its alkylated derivatives through the sulfur-specific pathway by the bacterium RIPI-S81. *African Journal of Biotechnology*. **5**(4): 351–356.

Rehfuss, M. and Urban, J. (2005). *Rhodococcus phenolicus* sp. nov., a novel bioprocessor isolated actinomycete with the ability to degrade chlorobenzene, dichlorobenzene and phenol as sole carbon sources. *Systematic and Applied Microbiology*. **28**(8): 695–701.

Reichmuth, D.S., Blanch, H.W., and Keasling, J.D. (2004). Dibenzothiophene biodesulfurization pathway improvement using diagnostic GFP fusions. *Biotechnology and Bioengineering*. **88**(1): 94–99.

Reichmuth, D.S., Hittle, J.L., Blanch, H.W., and Keasling, J.D. (2000). Biodesulfurization of dibenzothiophene in *Escherichia coli* is enhanced by expression of a *Vibrio harveyi* oxidoreductase gene. *Biotechnology and Bioengineering*. **67**(1): 72–79.

Reşitoğlu, İ.A., Altinişik, K., and Keskin, A. (2015). The pollutant emissions from diesel-engine vehicles and exhaust aftertreatment systems. *Clean Technologies and Environmental Policy*. **17**(1): 15–27.

Rhee, S.-K., Chang, J.H., Chang, Y.K., and Chang, H.N. (1998). Desulfurization of dibenzothiophene and diesel oils by a newly isolated *Gordona* strain, CYKS1. *Applied and Environmental Microbiology*. **64**(6): 2327–2331.

Richter, M. and Rossello-Mora, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*. **106**(45): 19126–19131.

Rodriguez-R, L.M. and Konstantinidis, K.T. (2014). Bypassing cultivation to identify bacterial species. *Microbe Magazine*. **9**(3): 111–118.

le Roes, M., Goodwin, C.M., and Meyers, P.R. (2008). *Gordonia lacunae* sp. nov., isolated from an estuary. *Systematic and Applied Microbiology*. **31**(1): 17–23.

Roosaare, M., Vaher, M., Kaplinski, L., Möls, M., Andreson, R., Lepamets, M., Kõressaar, T., Naaber, P., Kõljalg, S., and Remm, M. (2017). StrainSeeker: fast identification of bacterial strains from raw sequencing reads using user-provided guide trees. *PeerJ*. **5:** e3353.

Rosselló-Móra, R., Trujillo, M.E., and Sutcliffe, I.C. (2017). Introducing a digital protologue: a timely move towards a database-driven systematics of Archaea and Bacteria. *Systematic and Applied Microbiology*. **40**(3): 121–122.

Rowbotham, T.J. and Cross, T. (1977). *Rhodococcus coprophilus* sp. nov.: an aerobic nocardioform actinomycete belonging to the 'rhodochrous' complex. *Journal of General Microbiology*. **100**(1): 123–138.

Ruvira, M.A., Lucena, T., Pujalte, M.J., Arahal, D.R., and Macián, M.C. (2013). *Marinifilum flexuosum* sp. nov., a new *Bacteroidetes* isolated from coastal Mediterranean Sea water and emended description of the genus *Marinifilum* Na *et al.*, 2009. *Systematic and Applied Microbiology*. **36**(3): 155–159.

Sadare, O.O. and Daramola, M.O. (2018). Adsorptive removal of dibenzothiophene from petroleum distillates using pomegranate leaf (*Punica granatum*) powder as a greener adsorbent. *Chemical Engineering Communications*, 1–13.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*. **4**(4): 406–425.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. 2nd ed. Plainview, N.Y: Cold Spring Harbor Laboratory Press.

Sangal, V., Goodfellow, M., Jones, A.L., Schwalbe, E.C., Blom, J., Hoskisson, P.A., and Sutcliffe, I.C. (2016). Next-generation systematics: An innovative approach to resolve the structure of complex prokaryotic taxa. *Scientific Reports*. **6**(1): 38392.

Sangal, V., Jones, A.L., Goodfellow, M., Hoskisson, P.A., Kämpfer, P., and Sutcliffe, I.C. (2015). Genomic analyses confirm close relatedness between *Rhodococcus defluvii* and *Rhodococcus equi (Rhodococcus hoagii)*. *Archives of Microbiology*. **197**(1): 113–116.

Santos, S.C.C., Alviano, D.S., Alviano, C.S., Pádula, M., Leitão, A.C., Martins, O.B., Ribeiro, C.M.S., Sassaki, M.Y.M., Matta, C.P.S., Bevilaqua, J., Sebastián, G. V., and Seldin, L. (2006). Characterization of *Gordonia* sp. strain F.5.25.8 capable of dibenzothiophene desulfurization and carbazole utilization. *Applied Microbiology and Biotechnology*. **71**(3): 355–362.

Saputra, D., Rasmussen, S., Larsen, M. V, Haddad, N., Sperotto, M.M., Aarestrup, F.M., Lund, O., and Sicheritz-Pontén, T. (2015). Reads2Type: a web application for rapid microbial taxonomy identification. *BMC Bioinformatics*. **16**(1): 398.

Sardessai, Y. and Bhosle, S. (2002). Tolerance of bacteria to organic solvents. *Research in Microbiology*. **153**(5): 263–268.

Schelle, M.W. and Bertozzi, C.R. (2006). Sulfate metabolism in mycobacteria. *ChemBioChem*. **7**(10): 1516–1524.

Scherer, H.W. (2009). Sulfur in soils. *Journal of Plant Nutrition and Soil Science*. **172**(3): 326–335.

Schilling, B.M., Alvarez, L.M., Wang, D.I.C., and Cooney, C.L. (2002). Continuous desulfurization of dibenzothiophene with *Rhodococcus rhodochrous* IGTS8 (ATCC 53968). *Biotechnology Progress*. **18**(6): 1207–1213.

Schloss, P.D. and Handelsman, J. (2005). Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Applied and Environmental Microbiology*. **71**(3): 1501–1506.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., and Weber, C.F. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*. **75**(23): 7537–7541.

Schmidt, M., Siebert, W., and Bagnall, K. (1973). The chemistry of sulfur, selenium, tellurium and polonium. In *Pergamon Texts in Inorganic Chemistry*. Oxford: Pergamon Press, p. 825.

Schweizer, H. (2008). Bacterial genetics: past achievements, present state of the field, and future challenges. *Biotechniques*. **44**(5): 633–634, 636–641.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. **30**(14): 2068–2069.

Selle, K. and Barrangou, R. (2015). Harnessing CRISPR–Cas systems for bacterial genome editing. *Trends in Microbiology*. **23**(4): 225–32.

Setti, L., Lanzarini, G., and Pifferi, P.G. (1997). Whole cell biocatalysis for an oil desulfurization process. *Fuel Processing Technology*. **52**(1–3): 145–153.

Shafi, R. and Hutchings, G.J. (2000). Hydrodesulfurization of hindered dibenzothiophenes: an overview. *Catalysis Today*. **59**(3–4): 423–442.

Shavandi, M., Sadeghizadeh, M., Zomorodipour, A., and Khajeh, K. (2009). Biodesulfurization of dibenzothiophene by recombinant *Gordonia alkanivorans* RIPI90A. *Bioresource Technology*. **100**(1): 475–479.

Shen, F. T., Goodfellow, M., Jones, A.L., Chen, Y.P., Arun, A.B., Lai, W.A., Rekha, P.D., and Young, C.-C. (2006a). *Gordonia soli* sp. nov., a novel actinomycete isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*. **56**(11): 2597–2601.

Shen, F. T, Lu, H.L., Lin, J.L., Huang, W.S., Arun, A.B., and Young, C.C. (2006b). Phylogenetic analysis of members of the metabolically diverse genus *Gordonia* based on proteins encoding the *gyrB* gene. *Research in Microbiology*. **157**(4): 367–75.

Shen, F.T., Young, L. Sen, Hsieh, M.F., Lin, S.Y., and Young, C.C. (2010). Molecular detection and phylogenetic analysis of the alkane 1-monooxygenase gene from *Gordonia* spp. *Systematic and Applied Microbiology*. **33**(2): 53–59.

Sikkema, J., de Bont, J.A., and Poolman, B. (1995). Mechanisms of membrane toxicity of hydrocarbons. *Microbiology and Molecular Biology Reviews*. **59**(2): 201–222.

Silva, L.J., Souza, D.T., Genuario, D.B., Hoyos, H.A.V., Santos, S.N., Rosa, L.H., Zucchi, T.D., and Melo, I.S. (2017). *Rhodococcus psychrotolerans* sp. nov., isolated from rhizosphere of Deschampsia antarctica. *Antonie van Leeuwenhoek*, 1–8.

Silva, T.P., Paixão, S.M., Teixeira, A. V., Roseiro, J.C., and Alves, L. (2013). Optimization of low sulfur carob pulp liquor as carbon source for fossil fuels biodesulfurization. *Journal of Chemical Technology and Biotechnology*. **88**(5): 919–923.

Smith, A. (2003). The structures of phenolic oximes and their complexes. *Coordination Chemistry Reviews*. **241**(1–2): 61–85.

Smith, D.R. (2015). Buying in to bioinformatics: an introduction to commercial sequence analysis software. *Briefings in Bioinformatics*. **16**(4): 700–9.

Soleimani, M., Bassi, A., and Margaritis, A. (2007). Biodesulfurization of refractory organic sulfur compounds in fossil fuels. *Biotechnology Advances*. **25**(6): 570–596.

Song, C. (2003). An overview of new approaches to deep desulfurization for ultra-clean gasoline, diesel fuel and jet fuel. *Catalysis Today*. **86**(1–4): 211–263.

Song, C. and Ma, X. (2003). New design approaches to ultra-clean diesel fuels by deep desulfurization and deep dearomatization. *Applied Catalysis B: Environmental*. **41**(1–2): 207–238.

Speight, J.G.J.G. (2014). *The Chemistry and Technology of Petroleum*. 5th ed. Boca Raton: CRC Press, Taylor & Francis Publishers.

Srivastava, V.C. (2012). An evaluation of desulfurization technologies for sulfur removal from liquid fuels. *RSC Advances*. **2**(3): 759–783.

Stackebrandt, E. (2011). Molecular taxonomic parameters. *Microbiology Australia*. **32:** 59–61.

Stackebrandt, E., Smida, J., and Collins, M.D. (1988). Evidence of phylogenetic heterogeneity within the genus *Rhodococcus*: Revival of the genus *Gordona* (Tsukamura). *Journal of General and Applied Microbiology*. **34**(4): 341–348.

Stancu, M.M. (2014). Physiological cellular responses and adaptations of *Rhodococcus erythropolis* IBBPo1 to toxic organic solvents. *Journal of Environmental Sciences*. **26**(10): 2065–2075.

Stanislaus, A., Marafi, A., and Rana, M.S. (2010). Recent advances in the science and technology of ultra low sulfur diesel (ULSD) production. *Catalysis Today*. **153**(1–2): 1–68.

Stern, D.I. (2006). Reversal of the trend in global anthropogenic sulfur emissions. *Global Environmental Change*. **16**(2): 207–220.

Stevens, V., Thijs, S., McAmmond, B., Langill, T., Van Hamme, J., Weyens, N., and Vangronsveld, J. (2017). Draft genome sequence of *Rhodococcus erythropolis* VSD3, a diesel fuel-degrading and plant growth-promoting bacterium isolated from *Hedera helix* leaves. *Genome Announcements*. **5**(8): e01680-16.

Stoecker, M.A., Herwig, R.P., and Staley, J.T. (1994). *Rhodococcus zopfii* sp. nov., a toxicant-degrading bacterium. *International Journal of Systematic Bacteriology*. **44**(1): 106–110.

Stoner, D.L., Wey, J.E., Barrett, K.B., Jolley, J.G., Wright, R.B., and Dugan, P.R. (1990). Modification of water-soluble coal-derived products by dibenzothiophene-degrading microorganisms. *Applied and Environmental Microbiology*.

Sukhoom, A. (1999). *Selective isolation and characterisation of rhodococci for use in bioremediation*. University of Newcastle, Newcastle upon Tyne, UK.

Sutcliffe, I.C. (1998). Cell envelope composition and organisation in the genus *Rhodococcus*. *Antonie van Leeuwenhoek*. **74**(1/3): 49–58.

Suzuki, D., Li, Z., Cui, X., Zhang, C., and Katayama, A. (2014). Reclassification of *Desulfobacterium anilini* as *Desulfatiglans anilini* comb. nov. within *Desulfatiglans* gen. nov., and description of a 4-chlorophenol-degrading sulfate-reducing bacterium, *Desulfatiglans parachlorophenolica* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*. **64**(Pt 9): 3081–3086.

Švec, P., Pantu˚ček, R., Vojtková, H., Sedláček, I., Vandamme, P., Busse, H.-J., Mašlaňová, I., Králová, S., Černohlávková, J., and Cnockaert, M. (2015). Classification of strain CCM 4446T as *Rhodococcus degradans* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*. **65**(12): 4381–4387.

Svobodová, D., Křenek, P., Fraenkl, M., and Gasparič, J. (1978). The colour reaction of phenols with the Gibbs reagent. *Mikrochimica Acta*. **70**(3–4): 197–211.

Takada, M., Nomura, N., Okada, H., Nakajima-Kambe, T., Nakahara, T., and Uchiyama, H. (2005). De-repression and comparison of oil-water separation activity of the dibenzothiophene desulfurizing bacterium, *Mycobacterium* sp. G3. *Biotechnology Letters*. **27**(12): 871–874.

Takeda, K., Kang, Y., Yazawa, K., Gonoi, T., and Mikami, Y. (2010). Phylogenetic studies of Nocardia species based on *gyrB* gene analyses. *Journal of Medical Microbiology*. **59**(2): 165–171.

Takeuchi, M., Hatano, K., Sedlácek, I., Pácová, Z., and Takeuchi, M. (2002). *Rhodococcus jostii* sp. nov., isolated from a medieval grave. *International Journal of Systematic and Evolutionary Microbiology*. **52**(2): 409–413.

Tanaka, Y., Matsui, T., Konishi, J., Maruhashi, K., and Kurane, R. (2002). Biodesulfurization of benzothiophene and dibenzothiophene by a newly isolated *Rhodococcus* strain. *Applied Microbiology and Biotechnology*. **59**(2–3): 325–328.

Tanaka, Y., Onaka, T., Matsui, T., Maruhashi, K., and Kurane, R. (2001). Desulfurization of benzothiophene by the Gram-negative bacterium, *Sinorhizobium* sp. KT55. *Current Microbiology*. **43**(3): 187–191.

Tanaka, Yasuhiro, Yoshikawa, O., Maruhashi, K., and Kurane, R. (2002). The cbs mutant strain of *Rhodococcus erythropolis* KA2-5-1 expresses high levels of Dsz enzymes in the presence of sulfate. *Archives of Microbiology*. **178**(5): 351–357.

Táncsics, A., Benedek, T., Farkas, M., Máthé, I., Márialigeti, K., Szoboszlay, S., Kukolya, J., and Kriszt, B. (2014). Sequence analysis of 16S rRNA, *gyrB* and *catA* genes and DNA-DNA hybridization reveal that *Rhodococcus jialingiae* is a later synonym of *Rhodococcus qingshengii*. *International Journal of Systematic and Evolutionary Microbiology*. **64**(Pt 1): 298–301.

Táncsics, A., Benedek, T., Szoboszlay, S., Veres, P.G., Farkas, M., Máthé, I., Márialigeti, K., Kukolya, J., Lányi, S., and Kriszt, B. (2015). The detection and phylogenetic analysis of the alkane 1-monooxygenase gene of members of the genus *Rhodococcus*. *Systematic and Applied Microbiology*. **38**(1): 1–7.

Táncsics, A., Szoboszlay, S., Kriszt, B., Kukolya, J., Baka, E., Márialigeti, K., and Révész, S. (2008). Applicability of the functional gene catechol 1,2-dioxygenase as a biomarker in the detection of BTEX-degrading *Rhodococcus* species. *Journal of Applied Microbiology*. **105**(4): 1026–1033.

Tang, H., Li, Q., Wang, Z., Yan, D., and Xing, J. (2012). Simultaneous removal of thiophene and dibenzothiophene by immobilized *Pseudomonas delafieldii* R-8 cells. *Chinese Journal of Chemical Engineering*. **20**(1): 47–51.

Tao, F., Yu, B., Xu, P., and Ma, C.Q. (2006). Biodesulfurization in biphasic systems containing organic solvents. *Applied and Environmental Microbiology*. **72**(7): 4604–4609.

Tatangelo, V., Mangili, I., Caracino, P., Anzano, M., Najmi, Z., Bestetti, G., Collina, E., Franzetti, A., and Lasagni, M. (2016). Biological devulcanization of ground natural rubber by *Gordonia desulfuricans* DSM 44462T strain. *Applied Microbiology and Biotechnology*, 1–12.

Tindall, B.J. (2014). The correct name of the taxon that contains the type strain of *Rhodococcus equi*. *International Journal of Systematic and Evolutionary Microbiology*. **64**(Pt 1): 302–308.

Thein, S. T. and Wallace, R. R. (1986). The use of synthetic oligonucleotides as specific hybridization probes in the diagonis of genetic disorders., in Davis, K. (ed.) *Human genetic diseases: A practical approach*. Oxford IRL, pp. 33–50.

Todescato, D., Maass, D., Mayer, D.A., Vladimir Oliveira, J., de Oliveira, D., Ulson de Souza, S.M.A.G., and Ulson de Souza, A.A. (2017). Optimal production of a *Rhodococcus erythropolis* ATCC 4277 biocatalyst for biodesulfurization and biodenitrogenation applications. *Applied Biochemistry and Biotechnology*.

Tong, Y., Charusanti, P., Zhang, L., Weber, T., and Lee, S.Y. (2015). CRISPR-Cas9 based engineering of actinomycetal genomes. *ACS Synthetic Biology*. **4**(9): 1020–9.

Torkamani, S., Shayegan, J., Yaghmaei, S., and Alemzadeh, I. (2008). Study of the first isolated fungus capable of heavy crude oil biodesulfurization. *Industrial and Engineering Chemistry Research*. **47**(19): 7476–7482.

Tran, P.N., Savka, M.A., and Gan, H.M. (2017). *In-silico* taxonomic classification of 373 genomes reveals species misidentification and new genospecies within the genus *Pseudomonas*. *Frontiers in Microbiology*. **8**(JUL): 1296.

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., and Minh, B.Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*. **44**(W1): w232–W235.

Uetz, T., Schneider, R., Snozzi, M., and Egli, T. (1992). Purification and characterization of a two-component monooxygenase that hydroxylates nitrilotriacetate from '*Chelatobacter*' strain ATCC 29600. *Journal of Bacteriology*. **174**(4): 1179–1188.

Urai, M., Yoshizaki, H., Anzai, H., Ogihara, J., Iwabuchi, N., Harayama, S., Sunairi, M., and Nakajima, M. (2007). Structural analysis of mucoidan, an acidic extracellular polysaccharide produced by a pristane-assimilating marine bacterium, *Rhodococcus erythropolis* PR4. *Carbohydrate Research*. **342**(7): 927–932.

Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K., and Swings, J. (1996). Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiological Reviews*. **60**(2): 407–38.

Verma, M., Lal, D., Kaur, J.J., Saxena, A., Kaur, J.J., Anand, S., and Lal, R. (2013). Phylogenetic analyses of phylum Actinobacteria based on whole genome sequences. *Research in Microbiology*. **164**(7): 718–728.

De Vos, P., Thompson, F., Thompson, C., and Swings, J. (2017). A flavor of prokaryotic taxonomy: systematics revisited. In *Microbial Resources*. pp. 29–44.

Wang, J., Butler, R.R., Wu, F., Pombert, J.-F., Kilbane, J.J., and Stark, B.C. (2017). Enhancement of microbial biodesulfurization via genetic engineering and adaptive evolution. *PLOS ONE*. **12**(1): e0168833.

Wang, L.T., Lee, F.L., Tai, C.J., and Kasai, H. (2007a). Comparison of *gyrB* gene sequences, 16S rRNA gene sequences and DNA-DNA hybridization in the *Bacillus subtilis* group. *International Journal of Systematic and Evolutionary Microbiology*. **57**(8): 1846–1850.

Wang, P. and Krawiec, S. (1996). Kinetic analyses of desulfurization of dibenzothiophene by *Rhodococcus erythropolis* in batch and fed-batch cultures. *Applied and Environmental Microbiology*. **62**(5): 1670–5.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007b). Naive Bayesian classifier for rapid assignment of rRNA Sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*. **73**(16): 5261–5267.

Wang, W., Ma, T., Lian, K., Zhang, Y., Tian, H., Ji, K., and Li, G. (2013a). Genetic analysis of benzothiophene biodesulfurization pathway of *Gordonia terrae* strain C-6. *PloS one*. **8**(12): e84386.

Wang, W., Ma, T., Ren, Y., and Li, G. (2013b). Draft genome sequence of a benzothiophene-desulfurizing bacterium, *Gordona terrae* strain C-6. *Genome Announcements*. **1**(3): e00381-13.

Wang, X., Jin, D., Zhou, L., Wu, L., An, W., and Zhao, L. (2014). Draft genome sequence of *Gordonia alkanivorans* Strain CGMCC6845, a halotolerant hydrocarbon-degrading bacterium. *Genome Announcements*. **2**(1): e01274-13.

Wang, Y.-X., Wang, H.-B., Zhang, Y.-Q., Xu, L.-H., Jiang, C.-L., and Li, W.-J. (2008). *Rhodococcus kunmingensis* sp. nov., an actinobacterium isolated from a rhizosphere soil. *International Journal of Systematic and Evolutionary Microbiology*. **58**(6): 1467–1471.

Wang, Z., Xu, J., Li, Y., Wang, K., Wang, Y., Hong, Q., Li, W.-J., and Li, S.-P. (2010). *Rhodococcus jialingiae* sp. nov., an actinobacterium isolated from sludge of a carbendazim wastewater treatment facility. *International Journal of Systematic and Evolutionary Microbiology*. **60**(2): 378–381.

Watanabe, H., Tanji, Y., Unno, H., and Hori, K. (2008). Rapid conversion of toluene by an *Acinetobacter* sp. Tol 5 mutant showing monolayer adsorption to water-oil interface. *Journal of Bioscience and Bioengineering*. **106**(3): 226–230.

Watkins, L.., Rodriguez, R., Schneider, D., Broderick, R., Cruz, M., Chambers, R., Ruckman, E., Cody, M., and Mrachko, G.. (2003). Purification and characterization of the aromatic desulfinase, 2-(2′-hydroxyphenyl) benzenesulfinate desulfinase. *Archives of Biochemistry and Biophysics*. **415**(1): 14–23.

Wayne, L.G., Brenner, D.J., Colwell, R.R., Grimont, P.A.D., Kandler, O., Krichevsky, M.I., Moore, L.H., Moore, W.E.C., Murray, R.G.E., Stackebrandt, E., Starr, M.P., and Truper, H.G. (1987). Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *International Journal of Systematic and Evolutionary Microbiology*. **37**(4): 463–464.

Weisburg, W.G., Barns, S.M., Pelletier, D.A., and Lane, D.J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *Journal of bacteriology*. **173**(2): 697–703.

Wenyan Jiang, David Bikard, David Cox, Feng Zhang, and L.A.M. (2013). CRISPR-assisted editing of bacterial genomes. *Nature Biotechnology*. **31**(3): 233–239.

Whitman, W.B. (2015). Genome sequences as the type material for taxonomic descriptions of prokaryotes. *Systematic and Applied Microbiology*. **38**(4): 217–222.

WHO (2007). Sulfur dioxide. In *Air Quality Guidelines for Europe*. Copenhangen: WHO Regional Publications, European Series No. 91, pp. 194–198.

Wild, M. (2014). Global Dimming and Brightening. *Global Environmental Change*, 39–47.

Willems, A., Falsen, E., Pot, B., Jantzen, E., Hoste, B., Vandamme, P., Gillis, M., Kersters, K., and De Ley, J. (1990). Acidovorax, a new genus for *Pseudomonas facilis*, *Pseudomonas delafieldii*, E. Falsen (EF) Group 13, EF Group 16, and several clinical isolates, with the species *Acidovorax facilis* comb. nov., *Acidovorax delafieldii* comb. nov., a. *International Journal of Systematic Bacteriology*. **40**(4): 384–398.

Woese, C.R. and Fox, G.E. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences*. **74**(11): 5088–5090.

Woo, P.C.Y., Lau, S.K.P., Teng, J.L.L., Tse, H., and Yuen, K.-Y. (2008). Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clinical Microbiology and Infection : The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*. **14**(10): 908–934.

Wood, D.E. and Salzberg, S.L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*. **15**(3): r46.

Wooff, E., Michell, S.L., Gordon, S. V., Chambers, M.A., Bardarov, S., Jacobs, W.R., Hewinson, R.G., and Wheeler, P.R. (2002). Functional genomics reveals the sole sulphate transporter of the *Mycobacterium tuberculosis* complex and its relevance to the acquisition of sulphur in vivo. *Molecular Microbiology*. **43**(3): 653–663.

Wu, X., Wang, Y., Dai, Q., Liang, R., and Jin, D. (2011). Isolation and characterization of four di-n-butyl phthalate (DBP)-degrading *Gordonia* sp. strains and cloning the 3,4-phthalate dioxygenase gene. *World Journal of Microbiology and Biotechnology*. **27**(11): 2611–2617.

Xi, L., Squires, C.H., Monticello, D.J., and Childs, J.D. (1997). A Flavin Reductase Stimulates DszA and DszC Proteins of *Rhodococcus erythropolis* IGTS8 *in vitro*. *Biochemical and Biophysical Research Communications*. **230**(1): 73–75.

Xiong, X., Xing, J., Li, X., Bai, X., Li, W., Li, Y., and Liu, H. (2007). Enhancement of biodesulfurization in two-liquid systems by heterogeneous expression of *Vitreoscilla* hemoglobin. *Applied and Environmental Microbiology*. **73**(7): 2394–2397.

Xu, J.-L., He, J., Wang, Z.-C., Wang, K., Li, W.-J., Tang, S.-K., and Li, S.-P. (2007). *Rhodococcus qingshengii* sp. nov., a carbendazim-degrading bacterium. *International Journal of Systematic and Evolutionary Microbiology*. **57**(12): 2754–2757.

Xu, P., Feng, J., Yu, B., Li, F., and Ma, C. (2009). Recent developments in biodesulfurization of fossil fuels. *Advances in Biochemical Engineering/Biotechnology*. **113:** 255–274.

Xu, Y., Mortimer, M.W., Fisher, T.S., Kahn, M.L., Brockman, F.J., and Xun, L. (1997). Cloning, sequencing, and analysis of a gene cluster from *Chelatobacter heintzii* ATCC 29600 encoding nitrilotriacetate monooxygenase and NADH:flavin mononucleotide oxidoreductase. *Journal of Bacteriology*. **179**(4): 1112–1116.

Yamada, K., Minoda, Y., Kodama, K., Nakatani, S., and Akasaki, T. (1968). Microbial conversion of petro-sulfur compounds. *Agricultural and Biological Chemistry*. **32**(7): 840–845.

Yamamoto, K., Hirao, K., Oshima, T., Aiba, H., Utsumi, R., and Ishihama, A. (2005). Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*. *Journal of Biological Chemistry*. **280**(2): 1448–56.

Yamamoto, S., Bouvet, P.J.M., and Harayama, S. (1999). Phylogenetic structures of the genus Acinetobacter based on *gyrB* sequences: comparison with the grouping by DNA-DNA hybridization. *International Journal of Systematic Bacteriology*. **49**(1): 87–95.

Yamamoto, S. and Harayama, S. (1995). PCR amplification and direct sequencing of *gyrB* genes with universal primers and their application to the detection and taxonomic analysis of *Pseudomonas putida* strains. *Applied and environmental microbiology*. **61**(3): 1104–9.

Yamamoto, S. and Harayama, S. (1996). Phylogenetic analysis of *Acinetobacter* strains based on the nucleotide sequences of *gyrB* genes and on the amino acid sequences of their products. *International Journal of Systematic Bacteriology*. **46**(2): 506–511.

Yáñez, M.A., Catalán, V., Apráiz, D., Figueras, M.J., and Martínez-Murcia, A.J. (2003). Phylogenetic analysis of members of the genus *Aeromonas* based on *gyrB* gene sequences. *International Journal of Systematic and Evolutionary Microbiology*. **53**(Pt 3): 875–83.

Yang, E., Yao, C., Liu, Y., Zhang, C., Jia, L., Li, D., Fu, Z., Sun, D., Robert Kirk, S., and Yin, D. (2018). Bamboo-derived porous biochar for efficient adsorption removal of dibenzothiophene from model fuel. *Fuel*. **211**(Supplement C): 121–129.

Yang, H., Chen, J., Fairbridge, C., Briker, Y., Zhu, Y.J., and Ring, Z. (2004). Inhibition of nitrogen compounds on the hydrodesulfurization of substituted dibenzothiophenes in light cycle oil. *Fuel Processing Technology*. **85**(12): 1415–1429.

Yang, J., Hu, Y., Zhao, D., Wang, S., Lau, P.C.K., and Marison, I.W. (2007). Two-layer continuous-process design for the biodesulfurization of diesel oils under bacterial growth conditions. *Biochemical Engineering Journal*. **37**(2): 212–218.

Yang, J. and Marison, I.W. (2005). Two-stage process design for the biodesulphurisation of a model diesel by a newly isolated *Rhodococcus globerulus* DAQ3. *Biochemical Engineering Journal*. **27**(1): 77–82.

Yang, Y., Liu, X., and Xu, B. (2014). Recent advances in molecular imprinting technology for the deep desulfurization of fuel oils. *New Carbon Materials*. **29**(1): 1–14.

Yassin, A.F. (2005). *Rhodococcus triatomae* sp. nov., isolated from a blood-sucking bug. *International Journal of Systematic and Evolutionary Microbiology*. **55**(4): 1575–1579.

Yi, H., Cho, Y.J., Yoon, S.H., Park, S.C., and Chun, J. (2012). Comparative genomics of Neisseria weaveri clarifies the taxonomy of this species and identifies genetic determinants that may be associated with virulence. *FEMS Microbiology Letters*. **328**(2): 100–105.

Yoon, J., Miwa, H., Ahmed, I., Yokota, A., and Fujiwara, T. (2010). *Rhodococcus baikonurensis* BTM4c, a boron-tolerant actinobacterial strain isolated from soil. *Bioscience, Biotechnology, and Biochemistry*. **74**(1): 178–181.

Yoon, J.H., Cho, Y.G., Kang, S.S., Kim, S.B., Lee, S.T., and Park, Y.H. (2000a). *Rhodococcus koreensis* sp. nov., a 2,4-dinitrophenol-degrading bacterium. *International Journal of Systematic and Evolutionary Microbiology*. **50**(3): 1193–1201.

Yoon, J.H., Kang, S.S., Cho, Y.G., Lee, S.T., Kho, Y.H., Kim, C.J., and Park, Y.H. (2000b). *Rhodococcus pyridinivorans* sp. nov., a pyridine-degrading bacterium. *International Journal of Systematic and Evolutionary Microbiology*. **50**(6): 2173–2180.

Yoon, S.H., Ha, S. min, Lim, J., Kwon, S., and Chun, J. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*. **110**(10): 1281–1286.

Yoshikawa, O., Ishii, Y., Koizumi, K.-I., Ohshiro, T., Izumi, Y., and Maruhashi, K. (2002). Enhancement and stabilization of desulfurization activity of *Rhodococcus erythropolis* KA2-5-1 by feeding ethanol and sulfur components. *Journal of Bioscience and Bioengineering*. **94**(5): 447–452.

Young, C.G., Amais, R.S., Schiavo, D., Garcia, E.E., Nóbrega, J.A., and Jones, B.T. (2011). Determination of sulfur in biodiesel microemulsions using the summation of the intensities of multiple emission lines. *Talanta*. **84**(3): 995–999.

Yu, B., Ma, C., Zhou, W., Wang, Y., Cai, X., Tao, F., Zhang, Q., Tong, M., Qu, J., and Xu, P. (2006a). Microbial desulfurization of gasoline by free whole-cells of *Rhodococcus erythropolis* XP. *FEMS Microbiology Letters*. **258**(2): 284–289.

Yu, B., Xu, P., Shi, Q., and Ma, C. (2006b). Deep desulfurization of diesel oil and crude oils by a newly isolated *Rhodococcus erythropolis* strain. *Applied and Environmental Microbiology*. **72**(1): 54–58.

Zampolli, J., Collina, E., Lasagni, M., and Di Gennaro, P. (2014). Biodegradation of variable-chain-length n-alkanes in *Rhodococcus opacus* R7 and the involvement of an alkane hydroxylase system in the metabolism. *AMB Express*. **4**(1): 73–82.

Zhang, H., Shan, G., Liu, H., and Xing, J. (2007). Surface modification of γ-Al_2O_3 nano-particles with gum arabic and its applications in adsorption and biodesulfurization. *Surface and Coatings Technology*. **201**(16–17): 6917–6921.

Zhang, M. and McSaveney, M.J. (2018). Is air pollution causing landslides in China? *Earth and Planetary Science Letters*. **481**: 284–289.

Zhou, X., Shen, X.-X., Hittinger, C.T., and Rokas, A. (2018). Evaluating fast maximum likelihood-based phylogenetic programs using empirical phylogenomic data sets. *Molecular Biology and Evolution*. **35**(2): 486–503.

Zhu, L., Li, M., Guo, S., and Wang, W. (2016). Draft genome sequence of a thermophilic desulfurization bacterium, *Geobacillus thermoglucosidasius* strain W-2. *Genome announcements*. **4**(4): e00793-16.

Zubaidy, I.A.H. Al, Tarsh, F. Bin, Darwish, N.N., Majeed, B.S.S.A., Sharafi, A. Al, and Chacra, L.A. (2013). Adsorption process of sulfur removal from diesel oil using sorbent materials. *Journal of Clean Energy Technologies*. **1**(1): 66–68.

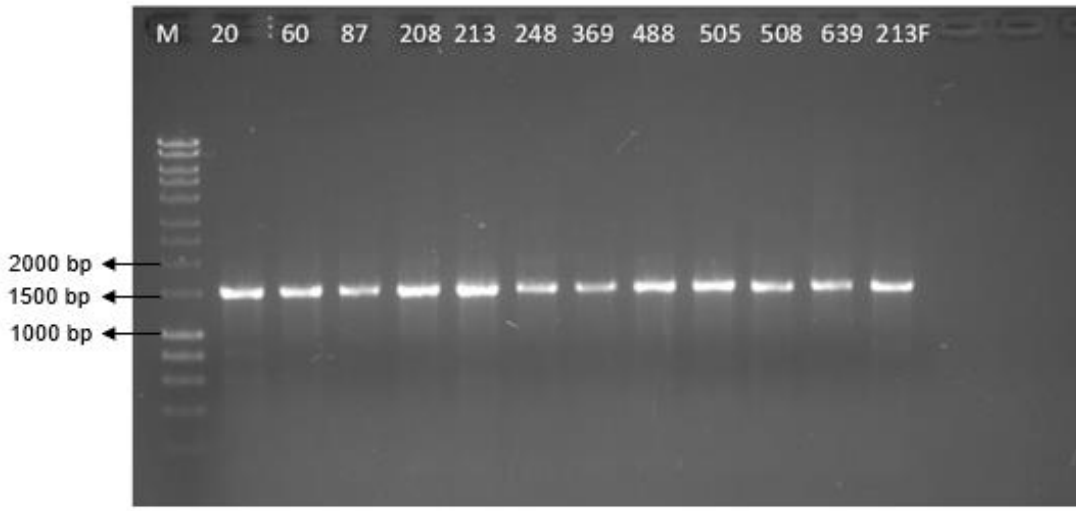
Chapter 7 Appendices

Appendix 7.1 List of chemicals and manufacturer used to study the growth and BDS activity of the test strains

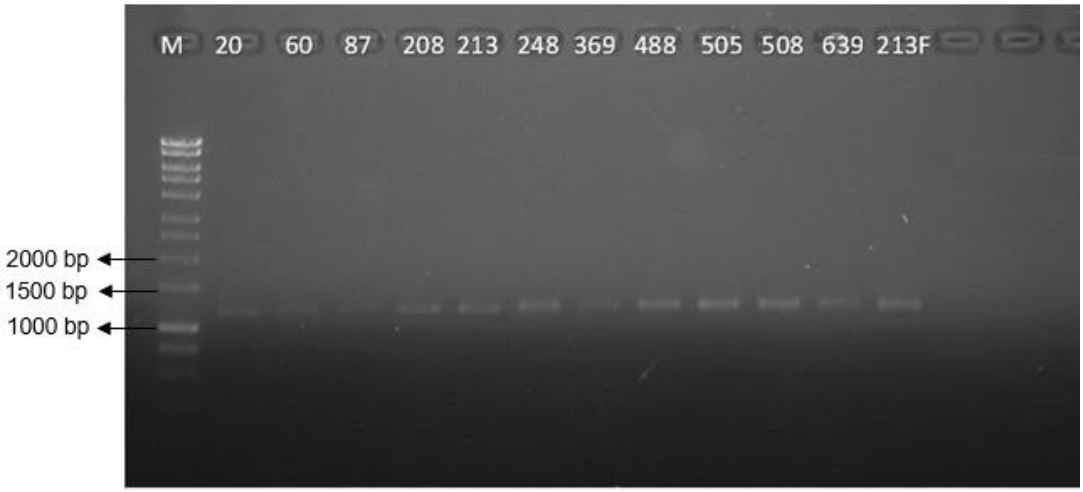
Chemical	Chemical Formula	Manufacturer (code)
Distilled water	H_2O	Elga Purelab Water Purifier
Nutrient Agar		Oxoid (CM0003)
Nutrient Broth		Oxoid (CM0001)
Agar No.1		Oxoid (LP0011)
Agar No.3		Oxoid (LP0013)
Agar Noble		Sigma (A5431)
Gibbs reagent (95%)	$C_6H_2Cl_3NO$	Sigma, UK
Benzothiophene (98%)	C_8H_6S	Alfa Aesar, UK
Dibenzothiophene (98%)	$C_{12}H_8S$	Alfa Aesar, UK
Sucrose	$C_{12}H_{22}O_{11}$	VWR Chemicals (GPR RECTAPUR®)
Ringer's solution		Oxoid (10239632)
Disodium phosphate	Na_2HPO_4	VWR Chemicals (GPR RECTAPUR®)
Dipotassium phosphate	KH_2PO_4	Fisher Scientific, USA
Ammonium chloride	NH_4Cl	Aldrich, USA
Sodium bicarbonate	$NaHCO_3$	Fisher Scientific, UK
EDTA		VWR Chemicals (AnalaR NORMAPUR®)
Salts used to make Trace Minerals	$MgCl_2.6H_2O$ $CaCl_2.2H_2O$ $(NH_4)_6Mo_7O_{24}.4\ H_2O$ $FeCl_3.6\ H_2O$ $ZnCl_2$, $FeCl_3.6\ H_2O$ $MnCl_2.4\ H_2O$ $CuCl_2.2\ H_2O$ $Co(NO_3)_2.6\ H_2O$ $Na_2B_4O_7.10\ H_2O$	VWR Chemicals (GPR RECTAPUR®)

Appendix 7.2 Gel electrophoresis of purified 16S rRNA and *gyrB* gene amplicons of test strains

A



B



The gel electrophoresis images of the **(A)** 16S rRNA gene amplicons and **(B)** *gyrB* gene amplicons obtained from the test strains is shown. The numbers printed over the wells indicate the respective strains from which the amplicons were obtained. By comparison with lanes formed from the 1Kb Hyperladder™ that was run on the wells marked (M), it could be seen that the bands of 16S rRNA gene amplicon and *gyrB* amplicon have migrated along with their own lanes without any distortion and parallel to the 1500 and 1000 bp marker bands, respectively. No obscure patches are observed that indicated the integrity (non-sheared DNA) and purity of the samples. The clear lane on the Blank control (B) showed the high quality of sample preparation for the electrophoresis.

Appendix 7.3 List of validly described species under genus *Rhodococcus*

Species	Nomenclature Revision History	Environment of isolation of the type strain	Type strain 16S rRNA gene sequence Accession Number
Rhodococcus rhodochrous	*Nocardia rubra*; *Nocardia salmonicolor*; *Rhodococcus roseus* Tsukamura *et al.* 1991; *Mycobacterium rhodochrous*; *Rhodococcus rhodochrous*	soil/activated sludge foam	X79288
Rhodococcus erythropolis	*Arthrobacter picolinophilus* Tate and Ensign 1974 *"Corynebacterium hydrocarboclastum" "Nocardia canicruria" Nocardia calcarea* Metcalf and Brown 1957 *Nocardia erythropolis* (Gray)	soil	X79289
Rhodococcus fascians	*Corynebacterium fascians* (Tilford 1936) Dowson 1942; *"Mycobacterium luteum"*; *Rhodococcus luteus* Nesterenko *et al.* 1982	leaves of nfected plant *Chrysanthemum morifolium*	X79186
Rhodococcus coprophilus	*Mycobacterium rhodochrous*	lake mud	U93340
Rhodococcus rhodnii		gut of Reduviid bug (Rhodnius prolixus)	X80621
Rhodococcus ruber	*Nocardia pellegrino*; *Nocardia rubra*		X80625
Rhodococcus marinonascens		the uppermost layer of marine sediments from various sea sites in the northeastern Atlantic Ocean	X80617
Rhodococcus globerulus		soil	X80619
Rhodococcus opacus		soil from the surroundings of a defective town gas pipe (Siebert, 1969)	X80630

Rhodococcus zopfii		toluene-phenol bioreactor operated by the Department of Civil Engineering, University of Washington	AF191343
Rhodococcus wratislaviensis	*Tsukamurella wratislaviensis* Goodfellow *et al.* 1995	soil	Z37138
Rhodococcus percolatus		percolator that was seeded with contaminated sludge and sediment samples and was continuously fed with 2,4,6-tricholorophenol, Finland	X92114
Rhodococcus koreensis		industrial wastewater in Cheong-Ju, Korea	AF124342
Rhodococcus pyridinivorans		industrial wastewater in Korea	AF173005
Rhodococcus jostii		femur of the remains of Jost Lucembursky, margrave in Moravia, Brno (Czech Republic)	AB046357
Rhodococcus maanshanensis		soil sample that had been collected from Maanshan Mountain in Anhui Province, China	AF416566
Rhodococcus tukisamuensis		soil in Sapporo City, Hokkaido, Japan	AB067734

Rhodococcus aetherivorans		activated sludge	AF447391
Rhodococcus baikonurensis		air in the Russian space laboratory Mir	AB071951
Rhodococcus gordoniae		blood culture of an immunocompetent patient with fatal pneumonia associated with adult respiratory disease syndrome	AY233201
Rhodococcus corynebacterioides	*Nocardia corynebacterioides* (Serrano et al 1972)	air-contaminated culture medium	X80615
Rhodococcus phenolicus		Johnson Space Center graywater bioprocessor	AY533293
Rhodococcus triatomae		blood-sucking bug of the genus *Triatoma*	AJ854055
Rhodococcus yunnanensis		forest soil sample in Yunnan Province, China	AY602219
Rhodococcus imtechensis		pesticide-contaminated site in Punjab State, India	AY525785.2
Rhodococcus kroppenstedtii		Lahaul-Spiti Valley, a cold desert of the Himalayas, India	AY726605
Rhodococcus kyotonensis		soil sample in Kyoto city, Japan	AB269261

Rhodococcus qingshengii	*Rhodococcus jialingiae* (Wang *et al.* 2010)	carbendazim-contaminated soil sample from Jiangsu province, China	DQ090961
Rhodococcus cercidiphylli		surface sterilized leaf sample of Cercidiphyllum japonicum collected from Yunnan province, south-west China	EU325542
Rhodococcus kunmingensis		soil sample collected from the rhizosphere of Taxus chinensis in Kunming, SW China	DQ997045
***Rhodococcus* jialingiae**	now regrouped under *Rhodococcus qingshengii* (Xu *et al.*, 2007)	the sludge of a carbendazim wastewater treatment facility in Jiangsu province, China	DQ185597.2
Rhodococcus artemisiae		pharmaceutical plant Artemisia annua L	GU367155
Rhodococcus nanhaiensis		sediment sample collected from the South China Sea (Nanhai sea area) at a depth of 84.5 m	JN582175
Rhodococcus canchipurensis		limestone quarry at Hundung, Manipur, India	JN164649
Rhodococcus cerastii		phyllosphere of Cerastium holosteoides in the Hainich-Dun region, Germany	FR714842

Rhodococcus trifolii		leaf surface of Trifolium repens, Hainich-Dün region, Thuringia, Germany	FR714843
Rhodococcus defluvii		wastewater treatment bioreactor in Aachen, Germany which showed extensive phosphorus removal	KC788572
Rhodococcus enclensis		marine sediment sample collected from Chorao Island, Goa, India	HQ858009
Rhodococcus hoagii	*Corynebacterium equi* Magnusson 1923 *Corynebacterium hoagii* (Morse 1912) Eberson 1918 *Nocardia restricta* (Turfitt 1944) McClung 1974 ''*Prescotella equi*'' *''Prescottia equi*" *Rhodococcus equi* (Magnusson 1923) Goodfellow and Alderson 1977	lung abscess of the foal	X80614; X82052*
Rhodococcus kronopolitis		millipede (Kronopolites svenhedind Verhoeff), which was collected from Fenghuang Mountain in Wuchang, Heilongjiang Province, north China	KF887492
Rhodococcus aerolatus		subarctic rainwater	KM044053
Rhodococcus agglutinans		soil	KP232908

Rhodococcus antrifimi		dried bat dung inside a natural cave on Jeju Island	LN867321
Rhodococcus biphenylivorans		river sediment in Taizhou city, Zhejiang province, eastern China - e-waste disassembly site for nearly 30 years	KJ546454
Rhodococcus degradans	*Arthrobacter* sp. HA1T (Scholtz *et al.*, 1987b).		JQ776649
Rhodococcus lactis		sludge sample of a dairy industry effluent treatment plant	KP342300
Rhodococcus soli		soil, Kyoto Park, Japan	KJ939314
Rhodococcus humicola		forest topsoil in Suwon, Gyeonggi-Do, South Korea	KT301939
Rhodococcus pedocola		forest topsoil in Suwon, Gyeonggi-Do, South Korea	KT301938
Rhodococcus gannanensis		sunflower root (Helianthus annuus L.) in Gannan, China	KX887333
Rhodococcus sovatensis		water sample of the hypersaline and heliothermal Lake Ursu, Sovata, Romania	KU189221
Rhodococcus olei		oil-contaminated soil of Biratnagar, Morang, Nepal	MF405107

Rhodococcus pyschrotolerans		rhizosphere of Deschampsia antarctica collected at King George Island, Antarctic Peninsula	KY317932
Rhodococcus electrodiphilus		a marine coral reef collected from the Shivrajpur coast, Kachhighadi, Gujarat, India	LT630357
Rhodococcus daqingensis		petroleum-contaminated soil collected in Daqing, Heilongjiang province, China	MH205096
Rhodococcus subtropicus		soil sampled in a natural cave on Jeju Island, Republic of Korea	MK605285

The table presents a chronological list of the validly described species under genus *Rhodococcus* as of December 2019. Their taxonomical revision history and details about the environment of isolation where available are included along with the NCBI accession code for the 16S rRNA of their respective type strains. *In this research, *R. olei, R. pyschrotolerans, R. electrodiphilus*, *R. daqingensis* and *R. subtropicus* were not included in any of the phylogenetic analyses because they were described after December 2017.

Appendix 7.4 Results of the BLASTn analysis using the 16S rRNA and *gyrB* gene sequence

In the case of 16S rRNA gene, the curated subset of the NCBI nucleotide database containing 16S ribosomal RNA sequences of type strains (Bacteria and Archaea) was used. In the case of *gyrB* gene, the entire non-redundant nucleotide database was selected. The output were obtained using *blastn* version 2.5.0 for the analysis done in December 2016.

Strain 20:

- received as *Rhodococccus erythropolis* IEGM 20 – classified based on the species specific PCR primers designed by Bell *et al.* (1999)

BLAST analysis output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2702	2702	99%	0.0	99%	NR_104776.1
Rhodococcus jialingiae strain djl-6-2 16S ribosomal RNA gene, partial sequence	2700	2700	97%	0.0	100%	NR_115708.1
Rhodococcus qingshengii strain djl-6 16S ribosomal RNA gene, partial sequence	2700	2700	97%	0.0	100%	NR_043535.1
Nocardia globerula strain DSM 44596 16S ribosomal RNA gene, partial sequence	2686	2686	99%	0.0	99%	NR_104795.1
Rhodococcus erythropolis strain N11 16S ribosomal RNA gene, partial sequence	2671	2671	98%	0.0	99%	NR_037024.1
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, complete sequence	2641	2641	98%	0.0	99%	NR_026184.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2555	2555	98%	0.0	98%	NR_114500.1
Rhodococcus marinonascens strain DSM 43752 16S ribosomal RNA gene, partial sequence	2551	2551	98%	0.0	98%	NR_026183.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2521	2521	98%	0.0	98%	NR_024973.1
Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2519	2519	98%	0.0	97%	NR_042352.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2516	2516	99%	0.0	97%	NR_024765.1
Rhodococcus imtechensis strain RKJ300 16S ribosomal RNA gene, complete sequence	2512	2512	99%	0.0	97%	NR_042946.1
Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2510	2510	98%	0.0	97%	NR_026524.1
Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2501	2501	98%	0.0	97%	NR_026186.1
Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2488	2488	99%	0.0	97%	NR_136860.1
Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2486	2486	99%	0.0	97%	NR_109454.1
Rhodococcus yunnanensis strain YIM 70056 16S ribosomal RNA gene, partial sequence	2486	2486	98%	0.0	97%	NR_043009.1
Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2481	2481	99%	0.0	97%	NR_028654.1
Nocardia shimofusensis strain YZ-96 16S ribosomal RNA gene, partial sequence	2475	2475	99%	0.0	97%	NR_028650.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2475	2475	99%	0.0	97%	NR_041873.1
Rhodococcus percolatus strain MBS1 16S ribosomal RNA gene, partial sequence	2473	2473	99%	0.0	96%	NR_044878.2
Rhodococcus baikonurensis strain A1-22 16S ribosomal RNA gene, partial sequence	2473	2473	90%	0.0	99%	NR_024784.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2466	2466	98%	0.0	97%	NR_119107.1
Nocardia kroppenstedtii strain N1286 16S ribosomal RNA, partial sequence	2464	2464	99%	0.0	97%	NR_133794.1
Rhodococcus cercidiphylli strain YIM 65003 16S ribosomal RNA gene, partial sequence	2464	2464	98%	0.0	97%	NR_116275.1
Nocardia araoensis strain NBRC 100135 16S ribosomal RNA gene, partial sequence	2459	2459	99%	0.0	97%	NR_028652.1
Nocardia farcinica strain DSM 43665 16S ribosomal RNA gene, complete sequence	2459	2459	99%	0.0	97%	NR_114643.1

BLAST analysis output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 47072	1932	1932	100%	0.0	99%	AB014272.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 12674	1927	1927	100%	0.0	99%	AB014116.1
Rhodococcus globerulus gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21506	1927	1927	100%	0.0	99%	AB014251.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: IAM 1414	1921	1921	100%	0.0	99%	AB355726.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 6821	1921	1921	100%	0.0	99%	AB355707.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10797.icb	1921	1921	100%	0.0	99%	AB018771.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10804.icb	1921	1921	100%	0.0	99%	AB018745.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21220	1916	1916	100%	0.0	99%	AB014095.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 2892	1916	1916	100%	0.0	99%	AB355724.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain IAM 1484	1910	1910	100%	0.0	99%	AB014172.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 53968	1910	1910	100%	0.0	99%	AB014262.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 2895	1910	1910	100%	0.0	99%	AB355715.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10789.icb	1910	1910	100%	0.0	99%	AB018770.1
Rhodococcus globerulus gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 15076	1905	1905	100%	0.0	99%	AB014247.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain TA 422	1905	1905	100%	0.0	99%	AB014277.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21785	1905	1905	100%	0.0	99%	AB014258.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 6825	1905	1905	100%	0.0	99%	AB355725.1
Rhodococcus degradans strain CCM 4446 DNA gyrase subunit B (gyrB) gene, partial cds	1899	1899	99%	0.0	99%	KP663665.1
Rhodococcus qingshengii strain RGN4 DNA gyrase subunit B (gyrB) gene, partial cds	1897	1897	98%	0.0	99%	KF374695.1
Rhodococcus qingshengii strain K5 DNA gyrase subunit B (gyrB) gene, partial cds	1897	1897	98%	0.0	99%	KF374694.1
Rhodococcus qingshengii strain Ba49 DNA gyrase subunit B (gyrB) gene, partial cds	1897	1897	98%	0.0	99%	KF374690.1
Rhodococcus sp. 008, complete genome	1893	1893	100%	0.0	99%	CP012749.1
Rhodococcus qingshengii strain DSM 45222 DNA gyrase subunit B (gyrB) gene, partial cds	1886	1886	98%	0.0	99%	KF374699.1
Rhodococcus qingshengii strain BBG1 DNA gyrase subunit B (gyrB) gene, partial cds	1886	1886	98%	0.0	99%	KF374693.1
Rhodococcus qingshengii strain PT3-14 DNA gyrase subunit B (gyrB) gene, partial cds	1886	1886	98%	0.0	99%	KF374691.1
Rhodococcus sp. gyrB gene for DNA gyrase B subunit, partial cds, strain MBIC 1337	1882	1882	100%	0.0	99%	AB014175.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: PR4 (= NBRC	1882	1882	100%	0.0	99%	AB355727.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: DSM 11397	1882	1882	100%	0.0	99%	AB355712.1

Strain 60:

- received as *Rhodococcus sp.* – suggested as *R. opacus* strain based on a weak positive PCR signal obtained using the species specific PCR primers designed by Bell *et al.* (1999)

BLAST analysis output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

	Description	Max score	Total score	Query cover	E value	Ident	Accession
☐	Rhodococcus rhodochrous strain 372 16S ribosomal RNA gene, partial sequence	2704	2704	98%	0.0	99%	NR_037023.1
☐	Rhodococcus biphenylivorans strain TG9 16S ribosomal RNA, partial sequence	2678	2678	99%	0.0	99%	NR_134798.1
☐	Rhodococcus rhodochrous strain DSM 43241 16S ribosomal RNA gene, partial sequence	2658	2658	96%	0.0	99%	NR_116689.1
☐	Rhodococcus pyridinivorans strain PDB9 16S ribosomal RNA gene, partial sequence	2634	2634	98%	0.0	99%	NR_025033.1
☐	Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2595	2595	99%	0.0	98%	NR_115082.1
☐	Rhodococcus zopfii strain DSM 44108 16S ribosomal RNA gene, partial sequence	2573	2573	98%	0.0	98%	NR_041775.1
☐	Rhodococcus artemisiae strain YIM 65754 16S ribosomal RNA gene, partial sequence	2553	2553	99%	0.0	98%	NR_108785.1
☐	Rhodococcus coprophilus strain CUB 687 16S ribosomal RNA gene, partial sequence	2536	2536	97%	0.0	98%	NR_029206.1
☐	Rhodococcus rhodochrous strain DSM 43241 16S ribosomal RNA gene, partial sequence	2534	2534	91%	0.0	99%	NR_118621.1
☐	Rhodococcus ruber strain DSM 43338 16S ribosomal RNA gene, complete sequence	2516	2516	98%	0.0	97%	NR_026185.1
☐	Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2477	2477	99%	0.0	97%	NR_136860.1
☐	Rhodococcus gordoniae strain W4937 16S ribosomal RNA gene, partial sequence	2470	2470	93%	0.0	98%	NR_025730.1
☐	Rhodococcus rhodnii strain B/O 16S ribosomal RNA gene, partial sequence	2444	2444	98%	0.0	97%	NR_037029.1
☐	Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2442	2442	98%	0.0	97%	NR_042352.1
☐	Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2436	2436	99%	0.0	96%	NR_134799.1
☐	Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2436	2436	99%	0.0	96%	NR_109454.1
☐	Rhodococcus equi strain DSM 20307 16S ribosomal RNA gene, complete sequence	2435	2435	98%	0.0	96%	NR_041910.1
☐	Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2425	2425	99%	0.0	96%	NR_028654.1
☐	Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2425	2425	98%	0.0	96%	NR_026524.1
☐	Rhodococcus pyridinivorans strain DSM 44555 16S ribosomal RNA gene, partial sequence	2423	2423	90%	0.0	99%	NR_118620.1
☐	Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete seque	2414	2414	98%	0.0	96%	NR_041873.1
☐	Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2410	2410	98%	0.0	96%	NR_026186.1
☐	Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2403	2403	99%	0.0	96%	NR_104776.1
☐	Nocardia araoensis strain NBRC 100135 16S ribosomal RNA gene, partial sequence	2403	2403	98%	0.0	96%	NR_028652.1
☐	Nocardia amamiensis strain TT 00-78 16S ribosomal RNA gene, partial sequence	2398	2398	98%	0.0	96%	NR_041531.1
☐	Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete seque	2398	2398	98%	0.0	96%	NR_119107.1

BLAST analysis output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

	Description	Max score	Total score	Query cover	E value	Ident	Accession
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10809.icb	1978	1978	100%	0.0	99%	AB018780.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10805.icb	1973	1973	100%	0.0	99%	AB018776.1
☐	Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 13808	1973	1973	100%	0.0	99%	AB014173.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10803.icb	1962	1962	100%	0.0	99%	AB018775.1
☐	Rhodococcus rhodochrous strain BX2 DNA gyrase subunit B (gyrB) gene, partial cds	1929	1929	100%	0.0	99%	JF506741.1
☐	Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 15906	1801	1801	100%	0.0	97%	AB014260.1
☐	Rhodococcus sp. p52, complete genome	1796	1796	100%	0.0	97%	CP016819.1
☐	Rhodococcus pyridinivorans SB3094, complete genome	1796	1796	100%	0.0	97%	CP006996.1
☐	Rhodococcus sp. HS-D2 DNA gyrase subunit B gene, partial cds	1790	1790	100%	0.0	97%	KF742503.1
☐	Rhodococcus pyridinivorans gyrB gene for DNA gyrase subunit B, partial cds	1790	1790	100%	0.0	97%	AB088665.1
☐	Rhodococcus pyridinivorans strain AK37 DNA gyrase subunit B (gyrB) gene, partial cds	1786	1786	98%	0.0	97%	KF374703.1
☐	Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 4004	1779	1779	100%	0.0	96%	AB014252.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10806.icb	1652	1652	100%	0.0	94%	AB018777.1
☐	Rhodococcus sp. BX2 DNA gyrase subunit B (gyrB) gene, partial cds	1513	1513	93%	0.0	94%	GQ457455.1
☐	Rhodococcus coprophilus gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 29080	1253	1253	100%	0.0	88%	AB014271.1
☐	Rhodococcus sp. WB1, complete genome	1238	1238	100%	0.0	88%	CP015529.1
☐	Rhodococcus aetherivorans strain IcdP1, complete genome	1238	1238	100%	0.0	88%	CP011341.1
☐	Rhodococcus zopfii gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 51349	1171	1171	99%	0.0	86%	AB014176.1
☐	Rhodococcus ruber gyrB gene for DNA gyrase B subunit, partial cds, strain IFO 15591	1168	1168	99%	0.0	86%	AB014174.1
☐	Rhodococcus sp. YYL DNA gyrase subunit B (gyrB) gene, partial cds	1157	1157	99%	0.0	86%	KJ469760.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10807.icb	1157	1157	99%	0.0	86%	AB018778.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10258.icb	1142	1142	100%	0.0	86%	AB018782.1
☐	Rhodococcus sp. TCH14 gyrB gene for DNA gyrase subunit B, partial cds	1114	1114	100%	0.0	85%	AB183445.1
☐	Nocardia caishijiensis culture-collection CDC<USA-GA>:W9706 DNA gyrase subunit B (gyrB) ge	1109	1109	100%	0.0	85%	GQ496124.1
☐	Rhodococcus opacus strain 1CP, complete genome	1103	1103	100%	0.0	85%	CP009111.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10604.icb	1103	1103	100%	0.0	85%	AB018774.1
☐	Rhodococcus sp. TKN46 gyrB gene for DNA gyrase subunit B, partial cds	1099	1099	99%	0.0	85%	AB183443.1
☐	Rhodococcus sp. TKN45 gyrB gene for DNA gyrase subunit B, partial cds	1099	1099	99%	0.0	85%	AB183442.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10794.icb	1098	1098	100%	0.0	85%	AB018773.1

Strain 87:

- received as *Rhodococcus ruber* IEGM 87 – classified based on the species specific PCR primers designed by Bell *et al.* (1999)

BLAST analysis output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus ruber strain DSM 43338 16S ribosomal RNA gene, complete sequence	2728	2728	98%	0.0	100%	NR_026185.1
Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2591	2591	99%	0.0	98%	NR_115082.1
Rhodococcus zopfii strain DSM 44108 16S ribosomal RNA gene, partial sequence	2564	2564	98%	0.0	98%	NR_041775.1
Rhodococcus aetherivorans strain 10bc312 16S ribosomal RNA gene, partial sequence	2538	2538	93%	0.0	99%	NR_025208.1
Rhodococcus ruber strain DSM 43338 16S ribosomal RNA gene, partial sequence	2536	2536	91%	0.0	100%	NR_118602.1
Rhodococcus pyridinivorans strain PDB9 16S ribosomal RNA gene, partial sequence	2534	2534	98%	0.0	98%	NR_025033.1
Rhodococcus biphenylivorans strain TG9 16S ribosomal RNA, partial sequence	2518	2518	99%	0.0	97%	NR_134798.1
Rhodococcus rhodochrous strain 372 16S ribosomal RNA gene, partial sequence	2505	2505	97%	0.0	97%	NR_037023.1
Rhodococcus artemisiae strain YIM 65754 16S ribosomal RNA gene, partial sequence	2494	2494	99%	0.0	97%	NR_108785.1
Rhodococcus aetherivorans strain DSM 44752 16S ribosomal RNA gene, partial sequence	2492	2492	91%	0.0	99%	NR_118619.1
Rhodococcus coprophilus strain CUB 687 16S ribosomal RNA gene, partial sequence	2466	2466	97%	0.0	97%	NR_029206.1
Rhodococcus rhodnii strain B/O 16S ribosomal RNA gene, partial sequence	2455	2455	98%	0.0	97%	NR_037029.1
Rhodococcus rhodochrous strain DSM 43241 16S ribosomal RNA gene, partial sequence	2453	2453	96%	0.0	97%	NR_116689.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2449	2449	98%	0.0	97%	NR_041873.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2433	2433	98%	0.0	96%	NR_119107.1
Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2429	2429	99%	0.0	96%	NR_136860.1
Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2425	2425	98%	0.0	96%	NR_042352.1
Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2422	2422	99%	0.0	96%	NR_109454.1
Nocardia amamiensis strain TT 00-78 16S ribosomal RNA gene, partial sequence	2416	2416	98%	0.0	96%	NR_041531.1
Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2399	2399	98%	0.0	96%	NR_026524.1
Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2394	2394	99%	0.0	96%	NR_028654.1
Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2386	2386	91%	0.0	98%	NR_118611.1
Rhodococcus equi strain DSM 20307 16S ribosomal RNA gene, complete sequence	2386	2386	98%	0.0	96%	NR_041910.1
Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2385	2385	98%	0.0	96%	NR_026186.1
Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2383	2383	99%	0.0	96%	NR_134799.1
Nocardia araoensis strain NBRC 100135 16S ribosomal RNA gene, partial sequence	2377	2377	98%	0.0	96%	NR_028652.1
Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2375	2375	99%	0.0	96%	NR_104776.1
Nocardia higoensis strain NBRC 100133 16S ribosomal RNA gene, partial sequence	2375	2375	99%	0.0	96%	NR_028651.1

BLAST analysis output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

	Description	Max score	Total score	Query cover	E value	Ident	Accession
☐	Rhodococcus ruber strain DSM 43338 16S ribosomal RNA gene, complete sequence	2728	2728	98%	0.0	100%	NR_026185.1
☐	Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2591	2591	99%	0.0	98%	NR_115082.1
☐	Rhodococcus zopfii strain DSM 44108 16S ribosomal RNA gene, partial sequence	2564	2564	98%	0.0	98%	NR_041775.1
☐	Rhodococcus aetherivorans strain 10bc312 16S ribosomal RNA gene, partial sequence	2538	2538	93%	0.0	99%	NR_025208.1
☐	Rhodococcus ruber strain DSM 43338 16S ribosomal RNA gene, partial sequence	2536	2536	91%	0.0	100%	NR_118602.1
☐	Rhodococcus pyridinivorans strain PDB9 16S ribosomal RNA gene, partial sequence	2534	2534	98%	0.0	98%	NR_025033.1
☐	Rhodococcus biphenylivorans strain TG9 16S ribosomal RNA, partial sequence	2518	2518	99%	0.0	97%	NR_134798.1
☐	Rhodococcus rhodochrous strain 372 16S ribosomal RNA gene, partial sequence	2505	2505	97%	0.0	97%	NR_037023.1
☐	Rhodococcus artemisiae strain YIM 65754 16S ribosomal RNA gene, partial sequence	2494	2494	99%	0.0	97%	NR_108785.1
☐	Rhodococcus aetherivorans strain DSM 44752 16S ribosomal RNA gene, partial sequence	2492	2492	91%	0.0	99%	NR_118619.1
☐	Rhodococcus coprophilus strain CUB 687 16S ribosomal RNA gene, partial sequence	2466	2466	97%	0.0	97%	NR_029206.1
☐	Rhodococcus rhodnii strain B/O 16S ribosomal RNA gene, partial sequence	2455	2455	98%	0.0	97%	NR_037029.1
☐	Rhodococcus rhodochrous strain DSM 43241 16S ribosomal RNA gene, partial sequence	2453	2453	96%	0.0	97%	NR_116689.1
☐	Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2449	2449	98%	0.0	97%	NR_041873.1
☐	Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2433	2433	98%	0.0	96%	NR_119107.1
☐	Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2429	2429	99%	0.0	96%	NR_136860.1
☐	Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2425	2425	98%	0.0	96%	NR_042352.1
☐	Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2422	2422	99%	0.0	96%	NR_109454.1
☐	Nocardia amamiensis strain TT 00-78 16S ribosomal RNA gene, partial sequence	2416	2416	98%	0.0	96%	NR_041531.1
☐	Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2399	2399	98%	0.0	96%	NR_026524.1
☐	Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2394	2394	99%	0.0	96%	NR_028654.1
☐	Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2386	2386	91%	0.0	98%	NR_118611.1
☐	Rhodococcus equi strain DSM 20307 16S ribosomal RNA gene, complete sequence	2386	2386	98%	0.0	96%	NR_041910.1
☐	Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2385	2385	98%	0.0	96%	NR_026186.1
☐	Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2383	2383	99%	0.0	96%	NR_134799.1
☐	Nocardia araoensis strain NBRC 100135 16S ribosomal RNA gene, partial sequence	2377	2377	98%	0.0	96%	NR_028652.1
☐	Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2375	2375	99%	0.0	96%	NR_104776.1
☐	Nocardia higoensis strain NBRC 100133 16S ribosomal RNA gene, partial sequence	2375	2375	99%	0.0	96%	NR_028651.1

Strain 208:

- received as *Rhodococcus erythropolis* IEGM 208 – classified based on the species specific PCR primers designed by Bell *et al.* (1999)

BLAST analysis output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2632	2632	98%	0.0	99%	NR_114500.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2623	2623	99%	0.0	99%	NR_024765.1
Rhodococcus marinonascens strain DSM 43752 16S ribosomal RNA gene, partial sequence	2617	2617	98%	0.0	99%	NR_026183.1
Nocardia globerula strain DSM 44596 16S ribosomal RNA gene, partial sequence	2614	2614	99%	0.0	98%	NR_104795.1
Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2610	2610	98%	0.0	99%	NR_026524.1
Rhodococcus imtechensis strain RKJ300 16S ribosomal RNA gene, complete sequence	2608	2608	99%	0.0	98%	NR_042946.1
Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2606	2606	98%	0.0	99%	NR_026186.1
Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2597	2597	99%	0.0	98%	NR_104776.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2588	2588	98%	0.0	98%	NR_024973.1
Rhodococcus percolatus strain MBS1 16S ribosomal RNA gene, partial sequence	2575	2575	99%	0.0	98%	NR_044878.2
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, complete sequence	2575	2575	98%	0.0	98%	NR_026184.1
Rhodococcus erythropolis strain N11 16S ribosomal RNA gene, partial sequence	2571	2571	98%	0.0	98%	NR_037024.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2560	2560	95%	0.0	99%	NR_118421.1
Rhodococcus jialingiae strain djl-6-2 16S ribosomal RNA gene, partial sequence	2545	2545	97%	0.0	98%	NR_115708.1
Rhodococcus qingshengii strain djl-6 16S ribosomal RNA gene, partial sequence	2545	2545	97%	0.0	98%	NR_043535.1
Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2512	2512	99%	0.0	97%	NR_136860.1
Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2503	2503	98%	0.0	97%	NR_042352.1
Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2499	2499	99%	0.0	97%	NR_109454.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete seque	2483	2483	99%	0.0	97%	NR_041873.1
Rhodococcus yunnanensis strain YIM 70056 16S ribosomal RNA gene, partial sequence	2481	2481	98%	0.0	97%	NR_043009.1
Rhodococcus equi strain DSM 20307 16S ribosomal RNA gene, complete sequence	2475	2475	99%	0.0	97%	NR_041910.1
Nocardia amamiensis strain TT 00-78 16S ribosomal RNA gene, partial sequence	2466	2466	98%	0.0	97%	NR_041531.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete seque	2466	2466	98%	0.0	97%	NR_119107.1
Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2464	2464	99%	0.0	97%	NR_115082.1
Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2460	2460	99%	0.0	97%	NR_134799.1
Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2460	2460	99%	0.0	97%	NR_028654.1
Rhodococcus cercidiphylli strain YIM 65003 16S ribosomal RNA gene, partial sequence	2453	2453	98%	0.0	97%	NR_116275.1
Rhodococcus maanshanensis strain M712 16S ribosomal RNA gene, partial sequence	2447	2447	94%	0.0	98%	NR_025190.1
Rhodococcus wratislaviensis strain DSM 44107 16S ribosomal RNA gene, partial sequence	2442	2442	92%	0.0	99%	NR_118605.1

BLAST analysis output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus opacus B4 DNA, complete genome	1382	1382	99%	0.0	90%	AP011115.1
Rhodococcus opacus PD630, complete genome	1373	1373	99%	0.0	90%	CP003949.1
Rhodococcus opacus strain 1CP, complete genome	1371	1371	99%	0.0	90%	CP009111.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain DSM 1069	1371	1371	99%	0.0	90%	AB014111.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10604.icb	1371	1371	99%	0.0	90%	AB018774.1
Rhodococcus sp. TCH14 gyrB gene for DNA gyrase subunit B, partial cds	1371	1371	99%	0.0	90%	AB183445.1
Rhodococcus sp. TCH4 gyrB gene for DNA gyrase subunit B, partial cds	1371	1371	99%	0.0	90%	AB183444.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10794.icb	1365	1365	99%	0.0	90%	AB018773.1
Rhodococcus sp. TKN46 gyrB gene for DNA gyrase subunit B, partial cds	1365	1365	99%	0.0	90%	AB183443.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10793.icb	1360	1360	99%	0.0	89%	AB018772.1
Rhodococcus wratislaviensis gyrB gene for DNA gyrase B subunit, partial cds, strain JCM 9689	1360	1360	99%	0.0	89%	AB014315.1
Rhodococcus sp. TKN14 gyrB gene for DNA gyrase subunit B, partial cds	1360	1360	99%	0.0	90%	AB183441.1
Rhodococcus sp. TKN45 gyrB gene for DNA gyrase subunit B, partial cds	1354	1354	99%	0.0	89%	AB183442.1
Rhodococcus opacus strain R7 sequence	1349	1349	99%	0.0	89%	CP008947.1
Rhodococcus sp. gyrB gene for DNA gyrase B subunit, partial cds, strain RHA1	1338	1338	99%	0.0	89%	AB014313.1
Rhodococcus jostii RHA1, complete genome	1338	1338	99%	0.0	89%	CP000431.1
Rhodococcus koreensis gyrB gene for DNA gyrase subunit B, partial cds	1338	1338	99%	0.0	89%	AB075566.1
Rhodococcus jostii gyrB gene for DNA gyrase subunit B, partial cds	1327	1327	99%	0.0	89%	AB088664.1
Rhodococcus koreensis gyrB gene for DNA gyrase subunit B, partial cds	1293	1293	97%	0.0	89%	AB450821.1
Rhodococcus marinonascens gyrB gene for DNA gyrase B subunit, partial cds, strain IFO 14363	1144	1144	99%	0.0	86%	AB014115.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10258.icb	1088	1088	99%	0.0	85%	AB018782.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 15906	1061	1061	99%	0.0	84%	AB014260.1
Rhodococcus sp. p52, complete genome	1055	1055	99%	0.0	84%	CP016819.1
Rhodococcus pyridinivorans SB3094, complete genome	1055	1055	99%	0.0	84%	CP006996.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 4004	1055	1055	99%	0.0	84%	AB014252.1
Rhodococcus pyridinivorans gyrB gene for DNA gyrase subunit B, partial cds	1055	1055	99%	0.0	84%	AB088665.1
Rhodococcus sp. HS-D2 DNA gyrase subunit B gene, partial cds	1050	1050	99%	0.0	84%	KF742503.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10809.icb	1050	1050	99%	0.0	84%	AB018780.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10803.icb	1050	1050	99%	0.0	84%	AB018775.1

Strain 213:

- received as *Rhodococcus erythropolis* IEGM 213 – classified based on the species specific PCR primers designed by Bell *et al.* (1999)

BLAST analysis output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus jialingiae strain djl-6-2 16S ribosomal RNA gene, partial sequence	2700	2700	97%	0.0	100%	NR_115708.1
Rhodococcus qingshengii strain djl-6 16S ribosomal RNA gene, partial sequence	2700	2700	97%	0.0	100%	NR_043535.1
Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2699	2699	99%	0.0	99%	NR_104776.1
Nocardia globerula strain DSM 44596 16S ribosomal RNA gene, partial sequence	2682	2682	99%	0.0	99%	NR_104795.1
Rhodococcus erythropolis strain N11 16S ribosomal RNA gene, partial sequence	2671	2671	98%	0.0	99%	NR_037024.1
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, complete sequence	2641	2641	98%	0.0	99%	NR_026184.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2555	2555	98%	0.0	98%	NR_114500.1
Rhodococcus marinonascens strain DSM 43752 16S ribosomal RNA gene, partial sequence	2551	2551	98%	0.0	98%	NR_026183.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2521	2521	98%	0.0	98%	NR_024973.1
Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2516	2516	98%	0.0	98%	NR_042352.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2514	2514	99%	0.0	97%	NR_024765.1
Rhodococcus imtechensis strain RKJ300 16S ribosomal RNA gene, complete sequence	2510	2510	99%	0.0	97%	NR_042946.1
Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2510	2510	98%	0.0	97%	NR_026524.1
Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2501	2501	98%	0.0	97%	NR_026186.1
Rhodococcus yunnanensis strain YIM 70056 16S ribosomal RNA gene, partial sequence	2486	2486	98%	0.0	97%	NR_043009.1
Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2484	2484	99%	0.0	97%	NR_136860.1
Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2483	2483	99%	0.0	97%	NR_109454.1
Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2477	2477	99%	0.0	97%	NR_028654.1
Rhodococcus baikonurensis strain A1-22 16S ribosomal RNA gene, partial sequence	2473	2473	90%	0.0	99%	NR_024784.1
Nocardia shimofusensis strain YZ-96 16S ribosomal RNA gene, partial sequence	2471	2471	99%	0.0	97%	NR_028650.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2471	2471	99%	0.0	97%	NR_041873.1
Rhodococcus percolatus strain MBS1 16S ribosomal RNA gene, partial sequence	2470	2470	99%	0.0	97%	NR_044878.2
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2466	2466	98%	0.0	97%	NR_119107.1
Nocardia kroppenstedtii strain N1286 16S ribosomal RNA, partial sequence	2460	2460	99%	0.0	97%	NR_133794.1
Rhodococcus cercidiphylli strain YIM 65003 16S ribosomal RNA gene, partial sequence	2460	2460	98%	0.0	97%	NR_116275.1
Nocardia araoensis strain NBRC 100135 16S ribosomal RNA gene, partial sequence	2455	2455	98%	0.0	97%	NR_028652.1
Nocardia farcinica strain DSM 43665 16S ribosomal RNA gene, complete sequence	2455	2455	99%	0.0	97%	NR_114643.1
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, partial sequence	2440	2440	90%	0.0	99%	NR_118617.1
Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2438	2438	99%	0.0	96%	NR_134799.1

BLAST analysis output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10789.icb	1921	1921	100%	0.0	99%	AB018770.1
Rhodococcus globerulus gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21506	1916	1916	100%	0.0	99%	AB014251.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: IAM 1414	1910	1910	100%	0.0	99%	AB355726.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 6821	1910	1910	100%	0.0	99%	AB355707.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10797.icb	1910	1910	100%	0.0	99%	AB018771.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10804.icb	1910	1910	100%	0.0	99%	AB018745.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21220	1905	1905	100%	0.0	99%	AB014095.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain TA 422	1905	1905	100%	0.0	99%	AB014277.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21785	1905	1905	100%	0.0	99%	AB014258.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 2892	1905	1905	100%	0.0	99%	AB355724.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain IAM 1484	1899	1899	100%	0.0	99%	AB014172.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 47072	1899	1899	100%	0.0	99%	AB014272.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 2895	1899	1899	100%	0.0	99%	AB355715.1
Rhodococcus degradans strain CCM 4446 DNA gyrase subunit B (gyrB) gene, partial cds	1895	1895	98%	0.0	99%	KP663665.1
Rhodococcus sp. 008, complete genome	1893	1893	100%	0.0	99%	CP012749.1
Rhodococcus globerulus gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 15076	1893	1893	100%	0.0	99%	AB014247.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 12674	1893	1893	100%	0.0	99%	AB014116.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 6825	1893	1893	100%	0.0	99%	AB355725.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: DSM 11397	1893	1893	100%	0.0	99%	AB355712.1
Rhodococcus erythropolis CCM2595, complete genome	1888	1888	100%	0.0	99%	CP003761.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 19369	1888	1888	100%	0.0	99%	AB014244.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 53968	1888	1888	100%	0.0	99%	AB014262.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 17895	1888	1888	100%	0.0	99%	AB014259.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 3191	1888	1888	100%	0.0	99%	AB355722.1
Rhodococcus sp. gyrB gene for DNA gyrase B subunit, partial cds, strain MBIC 1337	1882	1882	100%	0.0	98%	AB014175.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 17041	1882	1882	100%	0.0	98%	AB014253.1
Nocardia coeliaca strain W9704 DNA gyrase subunit B gene, partial cds	1882	1882	100%	0.0	98%	GQ984365.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: PR4 (= NBRC	1882	1882	100%	0.0	98%	AB355727.1
Rhodococcus erythropolis PR4 DNA, complete genome	1882	1882	100%	0.0	98%	AP008957.1

Strain 248:

- received as *Rhodococcus opacus* IEGM 248 – classified based on the species specific PCR primers designed by Bell *et al.* (1999)

BLAST output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2700	2700	98%	0.0	99%	NR_026524.1
Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2680	2680	98%	0.0	99%	NR_026186.1
Rhodococcus imtechensis strain RKJ300 16S ribosomal RNA gene, complete sequence	2669	2669	99%	0.0	99%	NR_042946.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2627	2627	98%	0.0	99%	NR_024973.1
Rhodococcus percolatus strain MBS1 16S ribosomal RNA gene, partial sequence	2597	2597	99%	0.0	98%	NR_044878.2
Rhodococcus marinonascens strain DSM 43752 16S ribosomal RNA gene, partial sequence	2595	2595	98%	0.0	99%	NR_026183.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2593	2593	98%	0.0	98%	NR_114500.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2593	2593	99%	0.0	98%	NR_024765.1
Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2573	2573	99%	0.0	98%	NR_136860.1
Nocardia globerula strain DSM 44596 16S ribosomal RNA gene, partial sequence	2564	2564	99%	0.0	98%	NR_104795.1
Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2558	2558	99%	0.0	98%	NR_104776.1
Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2555	2555	99%	0.0	98%	NR_109454.1
Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2553	2553	98%	0.0	98%	NR_042352.1
Rhodococcus equi strain DSM 20307 16S ribosomal RNA gene, complete sequence	2553	2553	99%	0.0	98%	NR_041910.1
Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2538	2538	99%	0.0	97%	NR_134799.1
Rhodococcus wratislaviensis strain DSM 44107 16S ribosomal RNA gene, partial sequence	2532	2532	92%	0.0	99%	NR_118605.1
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, complete sequence	2531	2531	98%	0.0	98%	NR_026184.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2529	2529	95%	0.0	99%	NR_118421.1
Rhodococcus erythropolis strain N11 16S ribosomal RNA gene, partial sequence	2527	2527	98%	0.0	98%	NR_037024.1
Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2519	2519	99%	0.0	97%	NR_115082.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete seque	2516	2516	99%	0.0	97%	NR_041873.1
Nocardia shimofusensis strain YZ-96 16S ribosomal RNA gene, partial sequence	2510	2510	99%	0.0	97%	NR_028650.1
Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2505	2505	99%	0.0	97%	NR_028654.1
Rhodococcus jialingiae strain djl-6-2 16S ribosomal RNA gene, partial sequence	2501	2501	97%	0.0	98%	NR_115708.1
Nocardia amamiensis strain TT 00-78 16S ribosomal RNA gene, partial sequence	2501	2501	98%	0.0	97%	NR_041531.1
Rhodococcus qingshengii strain djl-6 16S ribosomal RNA gene, partial sequence	2501	2501	97%	0.0	98%	NR_043535.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete seque	2495	2495	98%	0.0	97%	NR_119107.1
Nocardia higoensis strain NBRC 100133 16S ribosomal RNA gene, partial sequence	2486	2486	99%	0.0	97%	NR_028651.1

BLAST output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain DSM 1069	1973	1973	100%	0.0	99%	AB014111.1
Rhodococcus opacus strain 1CP, complete genome	1923	1923	100%	0.0	99%	CP009111.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10604.icb	1923	1923	100%	0.0	99%	AB018774.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10794.icb	1917	1917	100%	0.0	99%	AB018773.1
Rhodococcus sp. TKN46 gyrB gene for DNA gyrase subunit B, partial cds	1917	1917	99%	0.0	99%	AB183443.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10793.icb	1912	1912	100%	0.0	99%	AB018772.1
Rhodococcus sp. TKN14 gyrB gene for DNA gyrase subunit B, partial cds	1912	1912	99%	0.0	99%	AB183441.1
Rhodococcus opacus PD630, complete genome	1901	1901	100%	0.0	98%	CP003949.1
Rhodococcus wratislaviensis gyrB gene for DNA gyrase B subunit, partial cds, strain JCM 9689	1840	1840	100%	0.0	97%	AB014315.1
Rhodococcus opacus strain R7 sequence	1834	1834	100%	0.0	97%	CP008947.1
Rhodococcus sp. gyrB gene for DNA gyrase B subunit, partial cds, strain RHA1	1812	1812	100%	0.0	97%	AB014313.1
Rhodococcus jostii RHA1, complete genome	1812	1812	100%	0.0	97%	CP000431.1
Rhodococcus opacus B4 DNA, complete genome	1740	1740	100%	0.0	96%	AP011115.1
Rhodococcus sp. TCH14 gyrB gene for DNA gyrase subunit B, partial cds	1735	1735	100%	0.0	96%	AB183445.1
Rhodococcus sp. TKN45 gyrB gene for DNA gyrase subunit B, partial cds	1729	1729	99%	0.0	96%	AB183442.1
Rhodococcus koreensis gyrB gene for DNA gyrase subunit B, partial cds	1729	1729	100%	0.0	96%	AB075566.1
Rhodococcus jostii gyrB gene for DNA gyrase subunit B, partial cds	1724	1724	100%	0.0	95%	AB088664.1
Rhodococcus sp. TCH4 gyrB gene for DNA gyrase subunit B, partial cds	1685	1685	100%	0.0	95%	AB183444.1
Rhodococcus koreensis gyrB gene for DNA gyrase subunit B, partial cds	1679	1679	97%	0.0	96%	AB450821.1
Rhodococcus marinonascens gyrB gene for DNA gyrase B subunit, partial cds, strain IFO 14363	1247	1247	100%	0.0	88%	AB014115.1
Rhodococcus sp. WB1, complete genome	1112	1112	100%	0.0	85%	CP015529.1
Rhodococcus aetherivorans strain IcdP1, complete genome	1112	1112	100%	0.0	85%	CP011341.1
Rhodococcus ruber gyrB gene for DNA gyrase B subunit, partial cds, strain IFO 15591	1107	1107	100%	0.0	85%	AB014174.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10809.icb	1105	1105	99%	0.0	85%	AB018780.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10803.icb	1105	1105	99%	0.0	85%	AB018775.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10805.icb	1103	1103	100%	0.0	85%	AB018776.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 13808	1099	1099	99%	0.0	85%	AB014173.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10258.icb	1092	1092	100%	0.0	85%	AB018782.1
Rhodococcus equi strain ATCC 6939 DNA gyrase B subunit (gyrB) gene, partial cds	1086	1086	100%	0.0	85%	FJ438562.1

Strain 369:

- received as *Rhodococcus ruber* IEGM 369 – classified based on the species specific PCR primers designed by Bell *et al.* (1999) and an inaccessible 16S rRNA gene sequence published in Russian literature (Novoselova *et al.*, 2011)

BLAST analysis output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus ruber strain DSM 43338 16S ribosomal RNA gene, complete sequence	2728	2728	98%	0.0	100%	NR_026185.1
Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2591	2591	99%	0.0	98%	NR_115082.1
Rhodococcus zopfii strain DSM 44108 16S ribosomal RNA gene, partial sequence	2564	2564	98%	0.0	98%	NR_041775.1
Rhodococcus aetherivorans strain 10bc312 16S ribosomal RNA gene, partial sequence	2538	2538	93%	0.0	99%	NR_025208.1
Rhodococcus ruber strain DSM 43338 16S ribosomal RNA gene, partial sequence	2536	2536	91%	0.0	100%	NR_118602.1
Rhodococcus pyridinivorans strain PDB9 16S ribosomal RNA gene, partial sequence	2534	2534	98%	0.0	98%	NR_025033.1
Rhodococcus biphenylivorans strain TG9 16S ribosomal RNA, partial sequence	2518	2518	99%	0.0	97%	NR_134798.1
Rhodococcus rhodochrous strain 372 16S ribosomal RNA gene, partial sequence	2505	2505	98%	0.0	97%	NR_037023.1
Rhodococcus artemisiae strain YIM 65754 16S ribosomal RNA gene, partial sequence	2494	2494	99%	0.0	97%	NR_108785.1
Rhodococcus aetherivorans strain DSM 44752 16S ribosomal RNA gene, partial sequence	2492	2492	91%	0.0	99%	NR_118619.1
Rhodococcus coprophilus strain CUB 687 16S ribosomal RNA gene, partial sequence	2466	2466	97%	0.0	97%	NR_029206.1
Rhodococcus rhodnii strain B/O 16S ribosomal RNA gene, partial sequence	2455	2455	98%	0.0	97%	NR_037029.1
Rhodococcus rhodochrous strain DSM 43241 16S ribosomal RNA gene, partial sequence	2453	2453	96%	0.0	97%	NR_116689.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2449	2449	98%	0.0	97%	NR_041873.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2433	2433	98%	0.0	96%	NR_119107.1
Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2429	2429	99%	0.0	96%	NR_136860.1
Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2425	2425	98%	0.0	96%	NR_042352.1
Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2422	2422	99%	0.0	96%	NR_109454.1
Nocardia amamiensis strain TT 00-78 16S ribosomal RNA gene, partial sequence	2416	2416	98%	0.0	96%	NR_041531.1
Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2399	2399	98%	0.0	96%	NR_026524.1
Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2394	2394	99%	0.0	96%	NR_028654.1
Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2386	2386	91%	0.0	98%	NR_118611.1
Rhodococcus equi strain DSM 20307 16S ribosomal RNA gene, complete sequence	2386	2386	98%	0.0	96%	NR_041910.1
Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2385	2385	98%	0.0	96%	NR_026186.1
Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2383	2383	99%	0.0	96%	NR_134799.1
Nocardia araoensis strain NBRC 100135 16S ribosomal RNA gene, partial sequence	2377	2377	98%	0.0	96%	NR_028652.1
Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2375	2375	99%	0.0	96%	NR_104776.1
Nocardia higoensis strain NBRC 100133 16S ribosomal RNA gene, partial sequence	2375	2375	99%	0.0	96%	NR_028651.1

BLAST analysis output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus sp. YYL DNA gyrase subunit B (gyrB) gene, partial cds	1984	1984	100%	0.0	99%	KJ469760.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10807.icb	1984	1984	100%	0.0	99%	AB018778.1
Rhodococcus ruber gyrB gene for DNA gyrase B subunit, partial cds, strain IFO 15591	1962	1962	100%	0.0	99%	AB014174.1
Rhodococcus sp. WB1, complete genome	1618	1618	100%	0.0	94%	CP015529.1
Rhodococcus aetherivorans strain IcdP1, complete genome	1618	1618	100%	0.0	94%	CP011341.1
Rhodococcus zopfii gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 51349	1243	1243	100%	0.0	88%	AB014176.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10805.icb	1171	1171	100%	0.0	86%	AB018776.1
Nocardia neocaledoniensis culture-collection CDC<USA-GA>:W8372 DNA gyrase subunit B (gyr	1170	1170	100%	0.0	86%	GQ496101.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10809.icb	1164	1164	99%	0.0	86%	AB018780.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10803.icb	1164	1164	99%	0.0	86%	AB018775.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 13808	1164	1164	99%	0.0	86%	AB014173.1
Rhodococcus equi strain ATCC 6939 DNA gyrase B subunit (gyrB) gene, partial cds	1160	1160	100%	0.0	86%	FJ438562.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10258.icb	1160	1160	100%	0.0	86%	AB018782.1
Rhodococcus equi gyrB gene for DNA gyrase subunit B, complete cds	1160	1160	100%	0.0	86%	AB201534.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10782.icb	1155	1155	100%	0.0	86%	AB018747.1
Rhodococcus equi gyrB gene for DNA gyrase B subunit, partial cds, strain IFO 14956	1155	1155	100%	0.0	86%	AB014110.1
Rhodococcus rhodochrous strain BX2 DNA gyrase subunit B (gyrB) gene, partial cds	1147	1147	99%	0.0	86%	JF506741.1
Rhodococcus equi 103S chromosome	1144	1144	100%	0.0	86%	FN563149.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10788.icb	1142	1142	100%	0.0	86%	AB018767.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 15906	1138	1138	100%	0.0	86%	AB014260.1
Rhodococcus sp. TCH4 gyrB gene for DNA gyrase subunit B, partial cds	1138	1138	100%	0.0	86%	AB183444.1
Rhodococcus sp. p52, complete genome	1133	1133	100%	0.0	86%	CP016819.1
Rhodococcus sp. HS-D2 DNA gyrase subunit B gene, partial cds	1133	1133	100%	0.0	86%	KF742503.1
Rhodococcus pyridinivorans SB3094, complete genome	1133	1133	100%	0.0	86%	CP006996.1
Nocardia asteroides culture-collection CDC<USA-GA>:W7948 DNA gyrase subunit B (gyrB) gene	1131	1131	100%	0.0	86%	GQ496120.1
Nocardia thailandica culture-collection CDC<USA-GA>:W9872 DNA gyrase subunit B (gyrB) gen	1131	1131	100%	0.0	86%	GQ496090.1
Nocardia asteroides gyrB gene for DNA gyrase subunit B, partial cds	1131	1131	100%	0.0	86%	AB074926.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 4004	1127	1127	100%	0.0	86%	AB014252.1
Rhodococcus pyridinivorans gyrB gene for DNA gyrase subunit B, partial cds	1127	1127	100%	0.0	86%	AB088665.1

Strain 488:

- received as *Rhodococcus erythropolis* IEGM 488 – suggested as probable *R. opacus* strain based on the species specific PCR primers designed by Bell *et al.* (1999). This strain was not listed in the IEGM online catalogue (December 2017)

BLAST analysis output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2580	2580	99%	0.0	99%	NR_114500.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2562	2562	99%	0.0	98%	NR_024765.1
Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2558	2558	99%	0.0	99%	NR_026524.1
Rhodococcus marinonascens strain DSM 43752 16S ribosomal RNA gene, partial sequence	2555	2555	98%	0.0	99%	NR_026183.1
Nocardia globerula strain DSM 44596 16S ribosomal RNA gene, partial sequence	2553	2553	99%	0.0	98%	NR_104795.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2551	2551	97%	0.0	99%	NR_118421.1
Rhodococcus imtechensis strain RKJ300 16S ribosomal RNA gene, complete sequence	2547	2547	99%	0.0	98%	NR_042946.1
Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2543	2543	98%	0.0	98%	NR_026186.1
Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2536	2536	99%	0.0	98%	NR_104776.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2536	2536	99%	0.0	98%	NR_024973.1
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, complete sequence	2514	2514	98%	0.0	98%	NR_026184.1
Rhodococcus erythropolis strain N11 16S ribosomal RNA gene, partial sequence	2514	2514	98%	0.0	98%	NR_037024.1
Rhodococcus percolatus strain MBS1 16S ribosomal RNA gene, partial sequence	2512	2512	99%	0.0	98%	NR_044878.2
Rhodococcus jialingiae strain djl-6-2 16S ribosomal RNA gene, partial sequence	2483	2483	98%	0.0	98%	NR_115708.1
Rhodococcus qingshengii strain djl-6 16S ribosomal RNA gene, partial sequence	2483	2483	98%	0.0	98%	NR_043535.1
Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2471	2471	99%	0.0	97%	NR_042352.1
Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2449	2449	99%	0.0	97%	NR_136860.1
Rhodococcus wratislaviensis strain DSM 44107 16S ribosomal RNA gene, partial sequence	2442	2442	94%	0.0	99%	NR_118605.1
Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2436	2436	99%	0.0	97%	NR_109454.1
Rhodococcus yunnanensis strain YIM 70056 16S ribosomal RNA gene, partial sequence	2435	2435	98%	0.0	97%	NR_043009.1
Rhodococcus cercidiphylli strain YIM 65003 16S ribosomal RNA gene, partial sequence	2431	2431	99%	0.0	97%	NR_116275.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete seque	2431	2431	99%	0.0	97%	NR_041873.1
Rhodococcus equi strain DSM 20307 16S ribosomal RNA gene, complete sequence	2423	2423	99%	0.0	97%	NR_041910.1
Rhodococcus maanshanensis strain M712 16S ribosomal RNA gene, partial sequence	2422	2422	96%	0.0	98%	NR_025190.1
Nocardia amamiensis strain TT 00-78 16S ribosomal RNA gene, partial sequence	2414	2414	99%	0.0	97%	NR_041531.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete seque	2403	2403	98%	0.0	97%	NR_119107.1
Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2401	2401	99%	0.0	96%	NR_115082.1
Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2398	2398	99%	0.0	96%	NR_134799.1
Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2398	2398	99%	0.0	96%	NR_028654.1

BLAST analysis output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

	Description	Max score	Total score	Query cover	E value	Ident	Accession
☐	Rhodococcus opacus B4 DNA, complete genome	1371	1371	99%	0.0	90%	AP011115.1
☐	Rhodococcus opacus strain 1CP, complete genome	1360	1360	99%	0.0	90%	CP009111.1
☐	Rhodococcus opacus PD630, complete genome	1360	1360	99%	0.0	90%	CP003949.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain DSM 1069	1360	1360	99%	0.0	90%	AB014111.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10604.icb	1360	1360	99%	0.0	90%	AB018774.1
☐	Rhodococcus sp. TCH14 gyrB gene for DNA gyrase subunit B, partial cds	1360	1360	99%	0.0	90%	AB183445.1
☐	Rhodococcus sp. TCH4 gyrB gene for DNA gyrase subunit B, partial cds	1360	1360	99%	0.0	89%	AB183444.1
☐	Rhodococcus sp. TKN46 gyrB gene for DNA gyrase subunit B, partial cds	1360	1360	99%	0.0	90%	AB183443.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10794.icb	1354	1354	99%	0.0	89%	AB018773.1
☐	Rhodococcus sp. TKN14 gyrB gene for DNA gyrase subunit B, partial cds	1354	1354	99%	0.0	89%	AB183441.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10793.icb	1349	1349	99%	0.0	89%	AB018772.1
☐	Rhodococcus wratislaviensis gyrB gene for DNA gyrase B subunit, partial cds, strain JCM 9689	1349	1349	99%	0.0	89%	AB014315.1
☐	Rhodococcus sp. TKN45 gyrB gene for DNA gyrase subunit B, partial cds	1349	1349	99%	0.0	89%	AB183442.1
☐	Rhodococcus opacus strain R7 sequence	1338	1338	99%	0.0	89%	CP008947.1
☐	Rhodococcus sp. gyrB gene for DNA gyrase B subunit, partial cds, strain RHA1	1327	1327	99%	0.0	89%	AB014313.1
☐	Rhodococcus jostii RHA1, complete genome	1327	1327	99%	0.0	89%	CP000431.1
☐	Rhodococcus koreensis gyrB gene for DNA gyrase subunit B, partial cds	1327	1327	99%	0.0	89%	AB075566.1
☐	Rhodococcus jostii gyrB gene for DNA gyrase subunit B, partial cds	1315	1315	99%	0.0	89%	AB088664.1
☐	Rhodococcus koreensis gyrB gene for DNA gyrase subunit B, partial cds	1293	1293	97%	0.0	89%	AB450821.1
☐	Rhodococcus marinonascens gyrB gene for DNA gyrase B subunit, partial cds, strain IFO 14363	1138	1138	99%	0.0	86%	AB014115.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10258.icb	1077	1077	99%	0.0	85%	AB018782.1
☐	Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 15906	1050	1050	99%	0.0	84%	AB014260.1
☐	Rhodococcus pyridinivorans SB3094, complete genome	1044	1044	99%	0.0	84%	CP006996.1
☐	Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 4004	1044	1044	99%	0.0	84%	AB014252.1
☐	Rhodococcus pyridinivorans gyrB gene for DNA gyrase subunit B, partial cds	1044	1044	99%	0.0	84%	AB088665.1
☐	Rhodococcus sp. HS-D2 DNA gyrase subunit B gene, partial cds	1038	1038	99%	0.0	84%	KF742503.1
☐	Nocardia tenerifensis strain DSM 44704 DNA gyrase subunit B (gyrB) gene, partial cds	1038	1038	99%	0.0	84%	FJ765062.1
☐	Rhodococcus sp. WB1, complete genome	1035	1035	99%	0.0	84%	CP015529.1
☐	Rhodococcus aetherivorans strain IcdP1, complete genome	1035	1035	99%	0.0	84%	CP011341.1

Strain 505:

- received as *Rhodococcus erythropolis* IEGM 505 – classified based on the species specific PCR primers designed by Bell *et al.* (1999). This strain was not listed in the printed IEGM catalogue written by Prof. Christofi

BLAST analysis output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus jialingiae strain djl-6-2 16S ribosomal RNA gene, partial sequence	2702	2702	97%	0.0	100%	NR_115708.1
Rhodococcus qingshengii strain djl-6 16S ribosomal RNA gene, partial sequence	2702	2702	97%	0.0	100%	NR_043535.1
Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2700	2700	99%	0.0	99%	NR_104776.1
Nocardia globerula strain DSM 44596 16S ribosomal RNA gene, partial sequence	2684	2684	99%	0.0	99%	NR_104795.1
Rhodococcus erythropolis strain N11 16S ribosomal RNA gene, partial sequence	2673	2673	98%	0.0	99%	NR_037024.1
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, complete sequence	2641	2641	98%	0.0	99%	NR_026184.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2555	2555	98%	0.0	98%	NR_114500.1
Rhodococcus marinonascens strain DSM 43752 16S ribosomal RNA gene, partial sequence	2553	2553	98%	0.0	98%	NR_026183.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2521	2521	98%	0.0	98%	NR_024973.1
Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2516	2516	98%	0.0	98%	NR_042352.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2514	2514	99%	0.0	97%	NR_024765.1
Rhodococcus imtechensis strain RKJ300 16S ribosomal RNA gene, complete sequence	2512	2512	99%	0.0	97%	NR_042946.1
Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2510	2510	98%	0.0	97%	NR_026524.1
Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2503	2503	98%	0.0	97%	NR_026186.1
Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2486	2486	99%	0.0	97%	NR_136860.1
Rhodococcus yunnanensis strain YIM 70056 16S ribosomal RNA gene, partial sequence	2486	2486	98%	0.0	97%	NR_043009.1
Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2484	2484	99%	0.0	97%	NR_109454.1
Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2479	2479	99%	0.0	97%	NR_028654.1
Rhodococcus baikonurensis strain A1-22 16S ribosomal RNA gene, partial sequence	2473	2473	90%	0.0	99%	NR_024784.1
Nocardia shimofusensis strain YZ-96 16S ribosomal RNA gene, partial sequence	2473	2473	99%	0.0	97%	NR_028650.1
Rhodococcus percolatus strain MBS1 16S ribosomal RNA gene, partial sequence	2471	2471	99%	0.0	97%	NR_044878.2
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2471	2471	98%	0.0	97%	NR_041873.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2468	2468	98%	0.0	97%	NR_119107.1
Nocardia kroppenstedtii strain N1286 16S ribosomal RNA, partial sequence	2462	2462	99%	0.0	97%	NR_133794.1
Rhodococcus cercidiphylli strain YIM 65003 16S ribosomal RNA gene, partial sequence	2460	2460	97%	0.0	97%	NR_116275.1
Nocardia araoensis strain NBRC 100135 16S ribosomal RNA gene, partial sequence	2455	2455	98%	0.0	97%	NR_028652.1
Nocardia farcinica strain DSM 43665 16S ribosomal RNA gene, complete sequence	2455	2455	98%	0.0	97%	NR_114643.1
Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2440	2440	99%	0.0	96%	NR_134799.1
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, partial sequence	2440	2440	90%	0.0	99%	NR_118617.1

BLAST analysis output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain IAM 1484	1938	1938	100%	0.0	99%	AB014172.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 2895	1938	1938	100%	0.0	99%	AB355715.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21220	1932	1932	100%	0.0	99%	AB014095.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 2892	1932	1932	100%	0.0	99%	AB355724.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: IAM 1414	1927	1927	100%	0.0	99%	AB355726.1
Rhodococcus globerulus gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 15076	1921	1921	100%	0.0	99%	AB014247.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 47072	1921	1921	100%	0.0	99%	AB014272.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21785	1921	1921	100%	0.0	99%	AB014258.1
Rhodococcus globerulus gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21506	1921	1921	100%	0.0	99%	AB014251.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 6825	1921	1921	100%	0.0	99%	AB355725.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 12674	1916	1916	100%	0.0	99%	AB014116.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 6821	1916	1916	100%	0.0	99%	AB355707.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10797.icb	1916	1916	100%	0.0	99%	AB018771.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10789.icb	1916	1916	100%	0.0	99%	AB018770.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10804.icb	1916	1916	100%	0.0	99%	AB018745.1
Rhodococcus degradans strain CCM 4446 DNA gyrase subunit B (gyrB) gene, partial cds	1912	1912	98%	0.0	99%	KP663665.1
Rhodococcus sp. 008, complete genome	1910	1910	100%	0.0	99%	CP012749.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain TA 422	1910	1910	100%	0.0	99%	AB014277.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 53968	1899	1899	100%	0.0	99%	AB014262.1
Rhodococcus qingshengii strain K5 DNA gyrase subunit B (gyrB) gene, partial cds	1897	1897	98%	0.0	99%	KF374694.1
Rhodococcus sp. gyrB gene for DNA gyrase B subunit, partial cds, strain MBIC 1337	1888	1888	100%	0.0	99%	AB014175.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: PR4 (= NBRC	1888	1888	100%	0.0	99%	AB355727.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: DSM 11397	1888	1888	100%	0.0	99%	AB355712.1
Rhodococcus erythropolis PR4 DNA, complete genome	1888	1888	100%	0.0	99%	AP008957.1
Rhodococcus qingshengii strain RGN4 DNA gyrase subunit B (gyrB) gene, partial cds	1886	1886	98%	0.0	99%	KF374695.1
Rhodococcus qingshengii strain PT3-14 DNA gyrase subunit B (gyrB) gene, partial cds	1886	1886	98%	0.0	99%	KF374691.1
Rhodococcus qingshengii strain Ba49 DNA gyrase subunit B (gyrB) gene, partial cds	1886	1886	98%	0.0	99%	KF374690.1
Rhodococcus erythropolis CCM2595, complete genome	1882	1882	100%	0.0	98%	CP003761.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 19369	1882	1882	100%	0.0	98%	AB014244.1

Strain 508:

- received as *Rhodococcus opacus* IEGM 508 – classified based on the species specific PCR primers designed by Bell *et al.* (1999)

BLAST analysis output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2632	2632	98%	0.0	99%	NR_114500.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2627	2627	99%	0.0	99%	NR_024765.1
Nocardia globerula strain DSM 44596 16S ribosomal RNA gene, partial sequence	2621	2621	99%	0.0	98%	NR_104795.1
Rhodococcus marinonascens strain DSM 43752 16S ribosomal RNA gene, partial sequence	2619	2619	98%	0.0	99%	NR_026183.1
Rhodococcus imtechensis strain RKJ300 16S ribosomal RNA gene, complete sequence	2614	2614	99%	0.0	98%	NR_042946.1
Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2610	2610	98%	0.0	99%	NR_026524.1
Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2608	2608	98%	0.0	99%	NR_026186.1
Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2604	2604	99%	0.0	98%	NR_104776.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2588	2588	98%	0.0	98%	NR_024973.1
Rhodococcus percolatus strain MBS1 16S ribosomal RNA gene, partial sequence	2580	2580	99%	0.0	98%	NR_044878.2
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, complete sequence	2575	2575	98%	0.0	98%	NR_026184.1
Rhodococcus erythropolis strain N11 16S ribosomal RNA gene, partial sequence	2573	2573	98%	0.0	98%	NR_037024.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2560	2560	95%	0.0	99%	NR_118421.1
Rhodococcus jialingiae strain djl-6-2 16S ribosomal RNA gene, partial sequence	2547	2547	97%	0.0	98%	NR_115708.1
Rhodococcus qingshengii strain djl-6 16S ribosomal RNA gene, partial sequence	2547	2547	97%	0.0	98%	NR_043535.1
Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2518	2518	99%	0.0	97%	NR_136860.1
Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2508	2508	98%	0.0	97%	NR_042352.1
Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2505	2505	99%	0.0	97%	NR_109454.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete seque	2486	2486	99%	0.0	97%	NR_041873.1
Rhodococcus yunnanensis strain YIM 70056 16S ribosomal RNA gene, partial sequence	2481	2481	98%	0.0	97%	NR_043009.1
Rhodococcus equi strain DSM 20307 16S ribosomal RNA gene, complete sequence	2479	2479	99%	0.0	97%	NR_041910.1
Rhodococcus phenolicus strain DSM 44812 16S ribosomal RNA gene, partial sequence	2470	2470	99%	0.0	97%	NR_115082.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete seque	2468	2468	98%	0.0	97%	NR_119107.1
Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2466	2466	99%	0.0	97%	NR_134799.1
Nocardia amamiensis strain TT 00-78 16S ribosomal RNA gene, partial sequence	2466	2466	98%	0.0	97%	NR_041531.1
Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2466	2466	99%	0.0	97%	NR_028654.1
Rhodococcus cercidiphylli strain YIM 65003 16S ribosomal RNA gene, partial sequence	2459	2459	98%	0.0	97%	NR_116275.1
Rhodococcus maanshanensis strain M712 16S ribosomal RNA gene, partial sequence	2447	2447	94%	0.0	98%	NR_025190.1
Rhodococcus wratislaviensis strain DSM 44107 16S ribosomal RNA gene, partial sequence	2442	2442	92%	0.0	99%	NR_118605.1

BLAST analysis output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

	Description	Max score	Total score	Query cover	E value	Ident	Accession
☐	Rhodococcus opacus B4 DNA, complete genome	1382	1382	100%	0.0	90%	AP011115.1
☐	Rhodococcus opacus PD630, complete genome	1373	1373	99%	0.0	90%	CP003949.1
☐	Rhodococcus opacus strain 1CP, complete genome	1371	1371	100%	0.0	90%	CP009111.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain DSM 1069	1371	1371	100%	0.0	90%	AB014111.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10604.icb	1371	1371	100%	0.0	90%	AB018774.1
☐	Rhodococcus sp. TCH14 gyrB gene for DNA gyrase subunit B, partial cds	1371	1371	100%	0.0	90%	AB183445.1
☐	Rhodococcus sp. TCH4 gyrB gene for DNA gyrase subunit B, partial cds	1371	1371	100%	0.0	90%	AB183444.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10794.icb	1365	1365	100%	0.0	90%	AB018773.1
☐	Rhodococcus sp. TKN46 gyrB gene for DNA gyrase subunit B, partial cds	1365	1365	99%	0.0	90%	AB183443.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10793.icb	1360	1360	100%	0.0	89%	AB018772.1
☐	Rhodococcus wratislaviensis gyrB gene for DNA gyrase B subunit, partial cds, strain JCM 9689	1360	1360	100%	0.0	89%	AB014315.1
☐	Rhodococcus sp. TKN14 gyrB gene for DNA gyrase subunit B, partial cds	1360	1360	99%	0.0	90%	AB183441.1
☐	Rhodococcus sp. TKN45 gyrB gene for DNA gyrase subunit B, partial cds	1354	1354	99%	0.0	89%	AB183442.1
☐	Rhodococcus opacus strain R7 sequence	1349	1349	100%	0.0	89%	CP008947.1
☐	Rhodococcus sp. gyrB gene for DNA gyrase B subunit, partial cds, strain RHA1	1338	1338	100%	0.0	89%	AB014313.1
☐	Rhodococcus jostii RHA1, complete genome	1338	1338	100%	0.0	89%	CP000431.1
☐	Rhodococcus koreensis gyrB gene for DNA gyrase subunit B, partial cds	1338	1338	100%	0.0	89%	AB075566.1
☐	Rhodococcus jostii gyrB gene for DNA gyrase subunit B, partial cds	1327	1327	100%	0.0	89%	AB088664.1
☐	Rhodococcus koreensis gyrB gene for DNA gyrase subunit B, partial cds	1293	1293	97%	0.0	89%	AB450821.1
☐	Rhodococcus marinonascens gyrB gene for DNA gyrase B subunit, partial cds, strain IFO 14363	1144	1144	100%	0.0	86%	AB014115.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10258.icb	1088	1088	100%	0.0	85%	AB018782.1
☐	Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 15906	1061	1061	100%	0.0	84%	AB014260.1
☐	Rhodococcus sp. p52, complete genome	1055	1055	100%	0.0	84%	CP016819.1
☐	Rhodococcus pyridinivorans SB3094, complete genome	1055	1055	100%	0.0	84%	CP006996.1
☐	Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 4004	1055	1055	100%	0.0	84%	AB014252.1
☐	Rhodococcus pyridinivorans gyrB gene for DNA gyrase subunit B, partial cds	1055	1055	100%	0.0	84%	AB088665.1
☐	Rhodococcus sp. HS-D2 DNA gyrase subunit B gene, partial cds	1050	1050	100%	0.0	84%	KF742503.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10809.icb	1050	1050	100%	0.0	84%	AB018780.1
☐	Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10803.icb	1050	1050	100%	0.0	84%	AB018775.1

Strain F:

- received as *Rhodococcus* sp. – the species identity of strain F was unknown in the beginning of the research.

BLAST analysis output for 16S rRNA gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus jialingiae strain djl-6-2 16S ribosomal RNA gene, partial sequence	2700	2700	97%	0.0	100%	NR_115708.1
Rhodococcus qingshengii strain djl-6 16S ribosomal RNA gene, partial sequence	2700	2700	97%	0.0	100%	NR_043535.1
Nocardia coeliaca strain DSM 44595 16S ribosomal RNA gene, partial sequence	2699	2699	99%	0.0	99%	NR_104776.1
Nocardia globerula strain DSM 44596 16S ribosomal RNA gene, partial sequence	2682	2682	99%	0.0	99%	NR_104795.1
Rhodococcus erythropolis strain N11 16S ribosomal RNA gene, partial sequence	2671	2671	98%	0.0	99%	NR_037024.1
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, complete sequence	2641	2641	98%	0.0	99%	NR_026184.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2555	2555	98%	0.0	98%	NR_114500.1
Rhodococcus marinonascens strain DSM 43752 16S ribosomal RNA gene, partial sequence	2551	2551	98%	0.0	98%	NR_026183.1
Rhodococcus koreensis strain DNP505 16S ribosomal RNA gene, partial sequence	2521	2521	98%	0.0	98%	NR_024973.1
Rhodococcus triatomae strain IMMIB RIV-085 16S ribosomal RNA gene, partial sequence	2516	2516	98%	0.0	98%	NR_042352.1
Rhodococcus jostii strain IFO 16295 16S ribosomal RNA gene, partial sequence	2514	2514	99%	0.0	97%	NR_024765.1
Rhodococcus imtechensis strain RKJ300 16S ribosomal RNA gene, complete sequence	2510	2510	99%	0.0	97%	NR_042946.1
Rhodococcus wratislaviensis strain NCIMB 13082 16S ribosomal RNA gene, partial sequence	2510	2510	98%	0.0	97%	NR_026524.1
Rhodococcus opacus strain DSM 43205 16S ribosomal RNA gene, partial sequence	2501	2501	98%	0.0	97%	NR_026186.1
Rhodococcus yunnanensis strain YIM 70056 16S ribosomal RNA gene, partial sequence	2486	2486	98%	0.0	97%	NR_043009.1
Rhodococcus agglutinans strain CFH S0262 16S ribosomal RNA, partial sequence	2484	2484	99%	0.0	97%	NR_136860.1
Rhodococcus canchipurensis strain MBRL 353 16S ribosomal RNA gene, partial sequence	2483	2483	99%	0.0	97%	NR_109454.1
Nocardia arthritidis strain IFM 10035 16S ribosomal RNA gene, complete sequence	2477	2477	99%	0.0	97%	NR_028654.1
Rhodococcus baikonurensis strain A1-22 16S ribosomal RNA gene, partial sequence	2473	2473	90%	0.0	99%	NR_024784.1
Nocardia shimofusensis strain YZ-96 16S ribosomal RNA gene, partial sequence	2471	2471	99%	0.0	97%	NR_028650.1
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2471	2471	99%	0.0	97%	NR_041873.1
Rhodococcus percolatus strain MBS1 16S ribosomal RNA gene, partial sequence	2470	2470	99%	0.0	97%	NR_044878.2
Rhodococcus corynebacterioides strain DSM 20151 16S ribosomal RNA gene, complete sequ	2466	2466	98%	0.0	97%	NR_119107.1
Nocardia kroppenstedtii strain N1286 16S ribosomal RNA, partial sequence	2460	2460	99%	0.0	97%	NR_133794.1
Rhodococcus cercidiphylli strain YIM 65003 16S ribosomal RNA gene, partial sequence	2460	2460	98%	0.0	97%	NR_116275.1
Nocardia araoensis strain NBRC 100135 16S ribosomal RNA gene, partial sequence	2455	2455	98%	0.0	97%	NR_028652.1
Nocardia farcinica strain DSM 43665 16S ribosomal RNA gene, complete sequence	2455	2455	99%	0.0	97%	NR_114643.1
Rhodococcus globerulus strain DSM 43954 16S ribosomal RNA gene, partial sequence	2440	2440	90%	0.0	99%	NR_118617.1
Rhodococcus soli strain DSD51W 16S ribosomal RNA, partial sequence	2438	2438	99%	0.0	96%	NR_134799.1

BLAST analysis output for *gyrB* gene sequence

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10789.icb	1921	1921	100%	0.0	99%	AB018770.1
Rhodococcus globerulus gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21506	1916	1916	100%	0.0	99%	AB014251.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: IAM 1414	1910	1910	100%	0.0	99%	AB355726.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 6821	1910	1910	100%	0.0	99%	AB355707.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10797.icb	1910	1910	100%	0.0	99%	AB018771.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, clone:gy10804.icb	1910	1910	100%	0.0	99%	AB018745.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21220	1905	1905	100%	0.0	99%	AB014095.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain TA 422	1905	1905	100%	0.0	99%	AB014277.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 21785	1905	1905	100%	0.0	99%	AB014258.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 2892	1905	1905	100%	0.0	99%	AB355724.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain IAM 1484	1899	1899	100%	0.0	99%	AB014172.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 47072	1899	1899	100%	0.0	99%	AB014272.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 2895	1899	1899	100%	0.0	99%	AB355715.1
Rhodococcus degradans strain CCM 4446 DNA gyrase subunit B (gyrB) gene, partial cds	1895	1895	98%	0.0	99%	KP663665.1
Rhodococcus sp. 008, complete genome	1893	1893	100%	0.0	99%	CP012749.1
Rhodococcus globerulus gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 15076	1893	1893	100%	0.0	99%	AB014247.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 12674	1893	1893	100%	0.0	99%	AB014116.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 6825	1893	1893	100%	0.0	99%	AB355725.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: DSM 11397	1893	1893	100%	0.0	99%	AB355712.1
Rhodococcus erythropolis CCM2595, complete genome	1888	1888	100%	0.0	99%	CP003761.1
Rhodococcus erythropolis gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 19369	1888	1888	100%	0.0	99%	AB014244.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 53968	1888	1888	100%	0.0	99%	AB014262.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 17895	1888	1888	100%	0.0	99%	AB014259.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: JCM 3191	1888	1888	100%	0.0	99%	AB355722.1
Rhodococcus sp. gyrB gene for DNA gyrase B subunit, partial cds, strain MBIC 1337	1882	1882	100%	0.0	98%	AB014175.1
Rhodococcus rhodochrous gyrB gene for DNA gyrase B subunit, partial cds, strain ATCC 17041	1882	1882	100%	0.0	98%	AB014253.1
Nocardia coeliaca strain W9704 DNA gyrase subunit B gene, partial cds	1882	1882	100%	0.0	98%	GQ984365.1
Rhodococcus erythropolis gyrB gene for DNA gyrase subunit B, partial cds, strain: PR4 (= NBRC	1882	1882	100%	0.0	98%	AB355727.1
Rhodococcus erythropolis PR4 DNA, complete genome	1882	1882	100%	0.0	98%	AP008957.1

Appendix 7.5 List of naturally occurring bacterial isolates reported to exhibit biodesulphurisation of DBT

Bacteria	**Author & Year**
Agrobacterium sp. MC501	Constanti *et al.,* 1994
Arthrobacter sp. ECRD-1	Lee *et al.,* 1995
Arthrobacter sulfureus	Labana *et al.,* 2005
Bacillus brevis R-6	Jiang *et al.,* 2002
Bacillus cereus HN	Arabian *et al.,* 2014
Bacillus sphaericus R-16	Jiang *et al.,* 2002
Bacillus subtilis Fds-1	Ma *et al.,* 2006c
Bacillus subtilis WU-S2B	Kirimura *et al.,* 2001
Corynebacterium sp. P32C1	Maghsoudi *et al.,* 2000
Corynebacterium sp. SY1 (reclassified as *Rhodococcus* sp. SY1)	Omori *et al.,* 1992
Corynebacterium sp. ZD-1	Wang *et al.,* 2006
Desulfobacterium anilini (reclassified as *Desulfatiglans anilini*)	Aribike *et al.,* 2009
Desulfovibrio desulfuricans	Yamada *et al.,* 1968
Gordonia alkanivorans 1B	Alves & Paixão, 2011
	Alves *et al.,* 2008
Gordonia alkanivorans RIPI90A	Mohebali *et al.,* 2007
	Mohebali *et al.,* 2008
Gordonia alkanivorans 1B	Alves *et al.,* 2005
Gordonia alkanivorans RIPI90A	Mohebali *et al.,* 2007
Gordonia nitida CYKS1 (reclassified as *G. alkanivorans* CYKS1)	Chang *et al.,* 2001
	Rhee *et al.,* 1998
Gordonia sp. F.5.25.8	Duarte *et al.,* 2001
Gordonia sp. IITR100	Ahmad *et al.,* 2014
Gordonia sp. ZD-7	Li *et al.,* 2006

Gordonia sp. WQ-01	Jia *et al.,* 2006
Klebsiella sp.	Dudley & Frost, 1994
Lysinibacillus sphaericus DMT-7	Bahuguna *et al.,* 2011
Microbacterium sp. NISOC-06	Papizadeh *et al.,* 2010
Microbacterium ZD-M2	Li *et al.,* 2005b
Mycobacterium goodii X7B	Li *et al.,* 2005a
	Li *et al.,* 2007b
	Chen *et al.,* 2008
Mycobacterium phlei GTIS10	Kayser *et al.,* 2002
Mycobacterium phlei SM120-1	Srinivasaraghavan *et al.,* 2006
Mycobacterium phlei WU-0103	Ishii *et al.,* 2005
Mycobacterium phlei WU-0104	Ishii *et al.,* 2005
Mycobacterium phlei WU-F1	Furuya *et al.,* 2002
	Furuya *et al.,* 2003
Mycobacterium sp. G3	Nekodzuka *et al.,* 1997
Mycobacterium sp. MR65	Watanabe *et al.,* 2003
Mycobacterium sp. X7B	Nekodzuka *et al.,* 1997
	Li *et al.,* 2003
Mycobacterium sp. ZD-19	Li *et al.,* 2003
	Chen *et al.,* 2008
Nocardia asteroides	Olson, 2000
Nocardia globerula	Wang & Krawiec, 1994
Nocardia globerula R-9	Jiang *et al.,* 2002
	Mingfang *et al.,* 2003
Nocardia sp. CYKS2	Chang *et al.,* 1998
Paenibacillus sp. A11-2	Konishi *et al.,* 1997
	Onaka *et al.,* 2001

Pantoea agglomerans D23W3	Bhatia & Sharma, 2010
Pseudomonas abikonensis DDA109	Yamada *et al.,* 1968
Pseudomonas delafieldii R-8 (reclassified as *Acidovorax delafieldii)*	Jiang *et al.,* 2002
	Mingfang *et al.,* 2003
	Guobin *et al.,* 2005
	Guobin *et al.,* 2006
	Zhang *et al.,* 2007
	Zhang *et al.,* 2008
	Li *et al.,* 2009
	Guobin *et al.,* 2005
	Guobin *et al.,* 2006
Pseudomonas jianii DDC279	Yamada *et al.,* 1968
Pseudomonas jianii DDE 27	Yamada *et al.,* 1968
Pseudomonas putida	Gomez *et al.,* 2006
Pseudomonas putida A4	Tao *et al.,* 2006
Pseudomonas putida CECT5279	Alcon *et al.,* 2005
	Martin *et al.,* 2005
	Alcon *et al.,* 2005
	Caro *et al.,* 2007
	Caro *et al.,* 2008
Pseudomonas sp. ARK	Honda *et al.,* 1998
Pseudomonas stutzeri TCE3	Dinamarca *et al.,* 2010
Pseudomonas stutzeri UP-1	Hou *et al.,* 2005
Ralstonia eutropha	Dejaloud *et al.,* 2017
Rhodococcus erythropolis	Amin, 2011
Rhodococcus erythropolis 1awq	Feng *et al.,* 2006
Rhodococcus erythropolis D-1	Izumi *et al.,* 1994

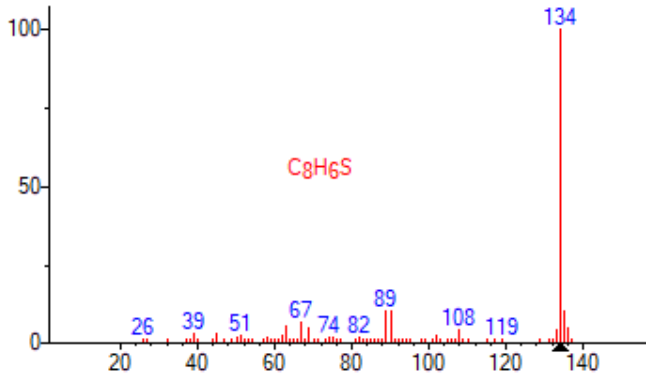
Rhodococcus erythropolis DR-1	Ma *et al.,* 2006
Rhodococcus erythropolis DS-3	Yu *et al.,* 2006
Rhodococcus erythropolis FSD-2	Zhang *et al.,* 2007
Rhodococcus erythropolis H-2	Ohshiro *et al.,* 1996
Rhodococcus erythropolis I-19	Denis-Larose *et al.,* 1997
	Folsom *et al.,* 1999
Rhodococcus erythropolis IGTS8	Kilbane, 1992
	del Olmo *et al.,* 2005
	Caro *et al.,* 2007
	Tangaromsuk *et al.,* 2008
	Ansari *et al.,* 2009
Rhodococcus erythropolis KA2- 5-1	Folsom *et al.,* 1999
Rhodococcus erythropolis LSSE8-1	Li *et al.,* 2007
Rhodococcus erythropolis NCC-1	Li *et al.,* 2007a
Rhodococcus erythropolis SHT87	Davoodi-Dehaghani *et al.,* 2010
Rhodococcus erythropolis USTB-03	Yan *et al.,* 2008
Rhodococcus erythropolis XP	Yu *et al.,* 2006
	Ma *et al.,* 2006a
	Yu *et al.,* 2006
Rhodococcus globerulus DAQ3	Yang *et al.,* 2007
	Yang & Marison, 2005
	Yang *et al.,* 2007
Rhodococcus sp. 1awq	Yu *et al.,* 2006
	Ma *et al.,* 2006
Rhodococcus sp.	Labana *et al.,* 2006
Rhodococcus sp. B1	Denis-Larose *et al.,* 1997
Rhodococcus sp. DS-3	Akbarzadeh *et al.,* 2003

Rhodococcus sp. ECRD-1	Lee *et al.,* 1995
	Grossman *et al.,* 1999
	Grossman *et al.,* 2001
Rhodococcus sp. HN2	El-Gendy *et al.,* 2014
Rhodococcus sp. IMPS02	Castorena *et al.,* 2002
	Matsui *et al.,* 2001
Rhodococcus sp. JUBT1	Guchhait *et al.,* 2005
Rhodococcus sp. JVH1	Kirkwood *et al.,* 2007a
Rhodococcus sp. P32C1	Kobayashi *et al.,* 2000
	Maghsoudi *et al.,* 2001
Rhodococcus sp. SA11	Mohamed *et al.,* 2015
Rhodococcus sp. T09	Maghsoudi *et al.,* 2000
Rhodococcus sp. UM3	Purdy *et al.,* 1993
Rhodococcus sp. UM9	Purdy *et al.,* 1993
Rhodococcus sp. X309	Omori *et al.,* 1995
RIPI-S81 (classification unknown)	Rashidi *et al.,* 2006
Shewanella putrefaciens NCIMB 8768	Ansari *et al.,* 2007
Sphingomonas sp. AD109	Zhang *et al.,* 2007
Sphingomonas subarctica T7b	Darzins & Mrachko 1998
	Gunam *et al.,* 2006
Xanthomonas sp.	Gunam *et al.,* 2006

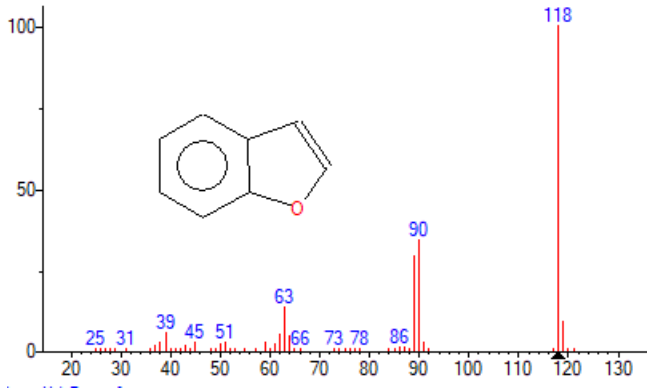
The list was compiled up to December 2017. The information was sourced from original research literature in the PubMed collection and from verified citations used in one of the comprehensive books on the subject, “Biodesulfurization in Petroleum Refining” by El-Gendy & Nassar (2018).

Appendix 7.6 Mass Spectral profile of intermediates of 4S pathway obtained from NIST Mass Spectrometry Data Center Collection with major peaks marked

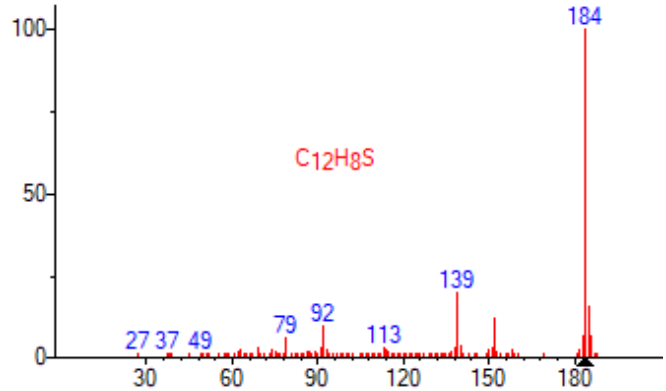
Benzothiophene C_8H_6S
MW: 134 Exact Mass: 134.019021 CAS#: 95-15-8



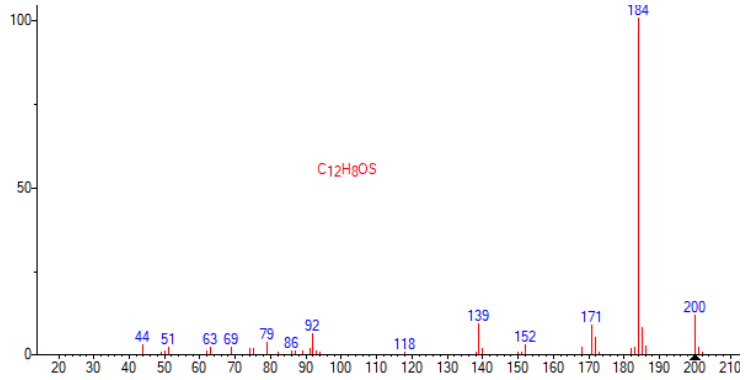
Benzofuran C_8H_6O
MW: 118 Exact Mass: 118.041865 CAS#: 271-89-6



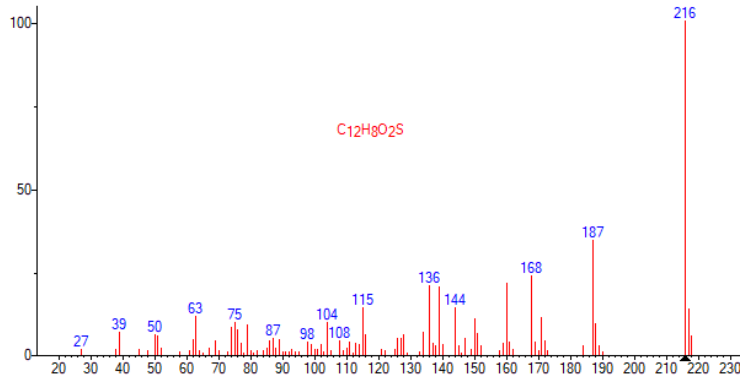
Dibenzothiophene $C_{12}H_8S$
MW: 184 Exact Mass: 184.034671 CAS#: 132-65-0



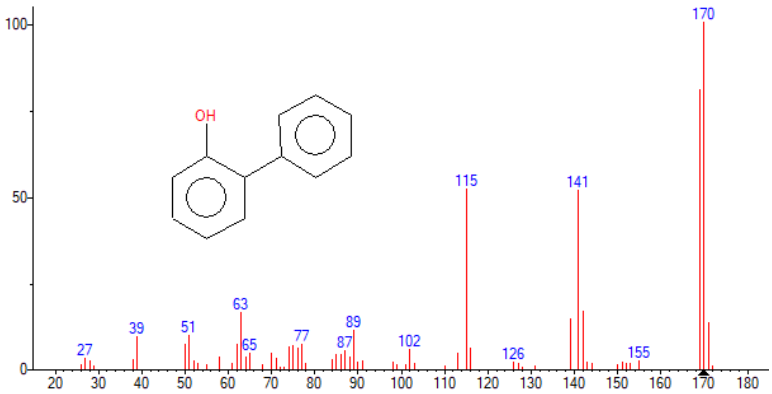
Dibenzothiophene-5-oxide $C_{12}H_8OS$
MW: 200 Exact Mass: 200.029586 CAS#: 1013-23-6



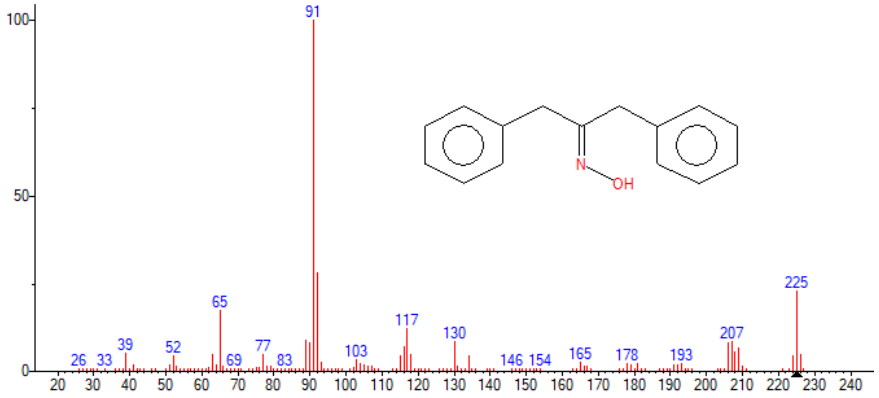
Dibenzothiophene-5,5-dioxide $C_{12}H_8O_2S$
MW: 216 Exact Mass: 216.024501 CAS#: 1016-05-3



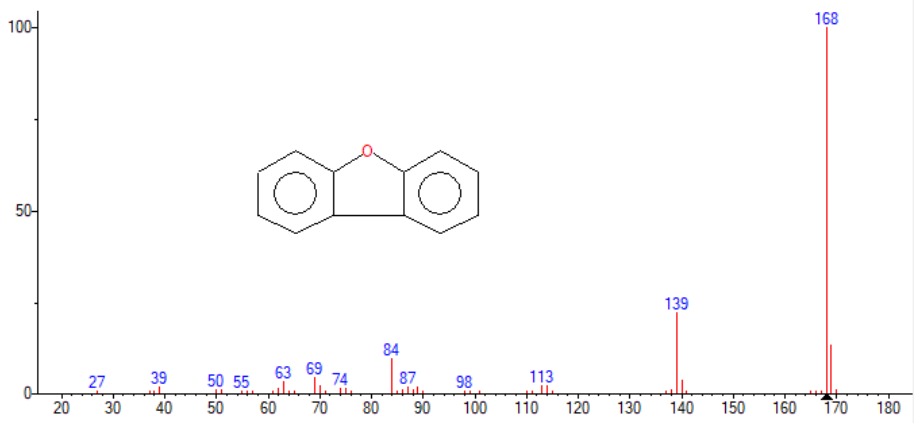
2-Hydroxybiphenyl $C_{12}H_{10}O$
MW: 170 Exact Mass: 170.073165 CAS#: 90-43-7



Dibenzyl ketoxime $C_{15}H_{15}NO$
MW: 225 Exact Mass: 225.115364 CAS#: 1788-31-4 NIST#: 185226



Dibenzofuran $C_{12}H_8O$
MW: 168 Exact Mass: 168.057515 CAS#: 132-64-9 NIST#: 228192



Appendix 7.7 ICP-OES operating conditions for aqueous and oil samples

		Aqueous	Oil
	Analysis Preferences		
Sample Options	# Repeats	3	3
	Delay Time	0.0 seconds	0.0
	Sample Flush Time	30 seconds	80
Source	Sample Introduction	Mira Mist Nebuliser	V-Groove Nebuliser
	Spray Chamber	Glass cyclonic	
	Plasma View	Line Selection	Line Selection
Analysis Maximum Integration Times (sec)	Low WL Range	Axial 15 Radial 15	Axial 15 Radial 15
	High WL Range	Axial 5 Radial 5	Axial 5 Radial 5
Calibration Mode	Concentration		
Trailing Full Frame	Intelli-Frame	Yes	
Options	Max Integration Time (sec)	30	
	WL Range	Low	
	View	Radial	
	Auto-Increment Sample Names	Yes	
	Use Sample Weight Corrections	No	
	Source Settings		
Nebuliser Pump	Flush Pump Rate (rpm)	100	25
	Analysis Pump Rate (rpm)	50	25
	Pump Relaxation Time (sec)	5	0
Pump Tubing Type	Sample Tube	Tygon orange-white	Viton orange-white (0.64 mm ID)
	Internal Standard Tube	Tygon blue-white	-
	Drain Tube	Tygon white-white	Viton white-white (1.02 mm ID)
	Torch		
	RF Power	1150 W	
	Nebuliser Flow	0.50 L/min	
	Auxiliary Gas	0.5 L/min	

Appendix 7.8 Genes encoding monooxygenases and FMN reductase found in the draft whole-genome sequence of strain F

Gene Loci	Gene_ copy number	Product	EC Number	Length (bp)	Direction
Contig_4	*kshA_1*	3-ketosteroid-9-alpha-monooxygenase oxygenase subunit	1.14.13.142	1161	reverse
Contig_4	*kshA_2*			1155	reverse
Contig_11	*kshA_3*			1164	forward
Contig_13	*kshA_4*			1197	forward
Contig_2	*hapE_1*	4-hydroxyacetophenone monooxygenase	1.14.13.84	1557	forward
Contig_2	*hapE_2*			1488	reverse
Contig_5	*hapE_3*			1485	forward
Contig_12	*hapE_4*			2085	reverse
Contig_14	*hapE_5*			1584	reverse
Contig_25	*hapE_6*			1575	forward
Contig_15	*nphA1*	4-nitrophenol 2-monooxygenase, oxygenase component	1.14.13.29	1629	forward
Contig_2	*npcB*	4-nitrophenol 4-monooxygenase/4-nitrocatechol 2-monooxygenase, reductase component	1.14.13.166	450	reverse
Contig_14	*luxA_1*	Alkanal monooxygenase alpha chain	1.14.14.3	1059	reverse
Contig_17	*luxA_2*			1095	forward
Contig_4	*alkB_1*	Alkane 1-monooxygenase	1.14.15.3	1176	forward
Contig_7	*alkB_2*			1227	forward
Contig_21	*alkB_4*			1152	reverse
Contig_6	*alkB1_1*			1167	forward
Contig_9	*alkB1_2*			1161	forward
Contig_3	*ssuD_1*	Alkanesulfonate monooxygenase	1.14.14.5	1110	forward
Contig_5	*ssuD_2*			996	forward
Contig_12	*ssuD_3*			1155	reverse
Contig_12	*ssuD_4*			1110	reverse
Contig_14	*ssuD_5*			846	reverse
Contig_16	*ssuD_6*			876	forward
Contig_3		Antibiotic biosynthesis monooxygenase		297	forward
Contig_3				972	reverse
Contig_19		Cyclohexanone 1,2-monooxygenase	1.14.13.22		forward
Contig_2	*dmoA*	Dimethyl-sulfide monooxygenase	1.14.13.131	1419	reverse
Contig_14	*ethA*	FAD-containing monooxygenase EthA	1.14.13.-	1482	reverse
Contig_3	*hsaA_1*	Flavin-dependent monooxygenase, oxygenase subunit HsaA	1.14.14.12	1179	forward
Contig_4	*hsaA_2*			1176	forward
Contig_11	*hsaA_3*			1167	reverse
Contig_11	*hsaA_4*			1179	reverse
Contig_13	*hsaA_5*			1182	forward
Contig_4	*hsaB_1*	Flavin-dependent monooxygenase, reductase subunit HsaB	1.5.1.36	585	forward
Contig_13	*hsaB_2*			501	forward
Contig_16	*hsaB_3*			654	reverse
Contig_4	*mhuD*	Heme-degrading monooxygenase HmoB	1.14.99.3	309	reverse
Contig_28	*pvdA*	L-ornithine 5-monooxygenase	1.13.12.-	1347	reverse
Contig_5		Lactate 2-monooxygenase	1.13.12.4	1248	forward
Contig_2	*limB_1*	Limonene 1,2-monooxygenase	1.14.13.107	1167	reverse
Contig_2	*limB_2*			1194	forward
Contig_6	*limB_3*			855	forward
Contig_11	*limB_4*			876	forward
Contig_16	*limB_5*			990	reverse

Contig_17	*limB_6*			1008	forward
Contig_22	*limB_7*			1251	forward
Contig_20	*linC_3*	Linalool 8-monooxygenase	1.14.13.151	1389	forward
Contig_15		Luciferase-like monooxygenase		897	forward
Contig_6	*msuD_1*	Methanesulfonate monooxygenase	1.14.14.5	1125	forward
Contig_15	*msuD_2*			1164	reverse
Contig_2	*ntaA_1*	Nitrilotriacetate monooxygenase component A **(DszA)**	1.14.14.10	1347	reverse
Contig_2	*ntaA_2*			1299	reverse
Contig_2	*ntaA_3*			1395	reverse
Contig_3	*ntaA_4*			1281	reverse
Contig_8	*ntaA_5*			1215	reverse
Contig_13	*ntaA_6*			1371	reverse
Contig_13	*ntaA_7*			1119	reverse
Contig_13	*ntaA_8*			1353	reverse
Contig_20	*ntaA_9*			1362	forward
Contig_22	*ntaA_10*			1242	forward
Contig_4	*nmo*	Nitronate monooxygenase	1.13.12.16	990	reverse
Contig_4				984	reverse
Contig_12				984	reverse
Contig_27				1116	forward
Contig_12				1029	forward
Contig_4				1062	reverse
Contig_11				963	forward
Contig_3				1137	reverse
Contig_3				930	forward
Contig_4	*pcpB_1*	Pentachlorophenol 4-monooxygenase	1.14.13.50	1608	reverse
Contig_7	*pcpB_2*			1437	forward
Contig_3	*pamO_1*	Phenylacetone monooxygenase	1.14.13.92	1482	forward
Contig_6	*pamO_2*			1389	reverse
Contig_13	*pamO_3*			1626	reverse
Contig_15	*mymA*	Putative FAD-containing monooxygenase MymA	1.14.13.-	1500	reverse
Contig_2	*moxC_1*	Putative monooxygenase MoxC	1.14.-.-	1080	reverse
Contig_21	*moxC_2*			1305	reverse
Contig_22	*moxC_3*			1368	forward
Contig_3		Putative monooxygenase Rv0793	1.-.-.-	285	forward
Contig_8				291	reverse
Contig_6	*ycnE*	Putative monooxygenase YcnE	1.-.-.-	288	forward
Contig_4	*rutA_1*	Pyrimidine monooxygenase RutA	1.14.99.46	864	reverse
Contig_6	*rutA_2*			870	forward
Contig_10	*rutA_3*			1020	reverse
Contig_4		Steroid C26-monooxygenase	1.14.13.141	1251	forward
Contig_19				1254	forward
Contig_24				1269	reverse
Contig_4	*iaaM*	Tryptophan 2-monooxygenase	1.13.12.3	1689	forward
Contig_13	*xylA*	Xylene monooxygenase electron transfer component		747	forward
Contig_14		NAD(P)H-dependent FAD/FMN reductase	1.5.1.45	567	reverse
		DszB			
Contig_12	*soxB_1*	2'-hydroxybiphenyl-2-sulfinate desulfinase	3.13.1.3	1041	reverse
Contig_20	*soxB_2*	2'-hydroxybiphenyl-2-sulfinate desulfinase	3.13.1.3	1098	forward
		DszC			
Contig_6	*soxC_1*	Dibenzothiophene desulfurization enzyme C	1.14.14.21	1221	reverse
Contig_10	*soxC_3*	Dibenzothiophene desulfurization enzyme C	1.14.14.21	1278	reverse

Contig_10	*soxC_2*	Dibenzothiophene desulfurization enzyme C	1.14.14.21	1224	reverse
Contig_20	*soxC_4*	Dibenzothiophene desulfurization enzyme C	1.14.14.21	1254	forward

The shaded rows indicate the genes associated with desulphurisation of benzothiophene and dibenzothiophene occurring at various contigs of the draft genome sequence of the strain F. The annotated whole-genome sequence was researched for genes homologous to the *dsz* genes and as a result, the 3 desulphurisation genes *soxC, ntaA* and *soxB* corresponding to *dszC, dszA* and *dszB*, respectively occured as cluster *dszABC* operon on contig_10. There were other instances of these genes at different contigs where they did not occur as a cluster. A gene encoding NADPH dependent FMN reductase was found at contig_14.

Appendix 7.9 Genes encoding monooxygenases and FMN reductase found in the draft whole-genome sequence of *Gordonia desulfuricans* 213E

Sequence Name	Gene_ copy number	Product	EC_number	Length (bp)	Direction
Contig_15	*tfdB_2*	2,4-dichlorophenol 6-monooxygenase	1.14.13.20	894	reverse
Contig_15	*tfdB_1*			705	reverse
Contig_97	*tfdB_3*			846	forward
Contig_8	*kshA*	3-ketosteroid-9-alpha-monooxygenase oxygenase subunit	1.14.13.142	1209	forward
Contig_1	*hapE_1*	4-hydroxyacetophenone monooxygenase	1.14.13.84	1521	forward
Contig_57	*hapE_2*			1533	reverse
Contig_112	*hapE_3*			1716	reverse
Contig_58	*hpaC*	4-hydroxyphenylacetate 3-monooxygenase reductase component	1.5.1.36	498	forward
Contig_26	*nphA1_1*	4-nitrophenol 2-monooxygenase, oxygenase component	1.14.13.29	1614	reverse
Contig_45	*nphA1_2*			1572	reverse
Contig_6	*npcB_1*	4-nitrophenol 4-monooxygenase/4-nitrocatechol 2-monooxygenase, reductase component	1.14.13.166	582	forward
Contig_26	*npcB_2*			588	reverse
Contig_127	*npcB_3*			516	forward
Contig_22	*luxA_1*	Alkanal monooxygenase alpha chain	1.14.14.3	1113	forward
Contig_27	*luxA_2*			1191	reverse
Contig_83	*luxA_3*			1023	reverse
Contig_137	*luxA_4*			1041	forward
Contig_51	*alkB*	Alkane 1-monooxygenase	1.14.15.3	1230	forward
Contig_8	*ssuD_1*	Alkanesulfonate monooxygenase	1.14.14.5	1107	forward
Contig_27	*ssuD_2*			1164	reverse
Contig_64	*ssuD_3*			1173	reverse
Contig_15		Antibiotic biosynthesis monooxygenase		318	reverse
Contig_126				321	forward
Contig_24	*dmoA_1*	Dimethyl-sulfide monooxygenase	1.14.13.131	1386	reverse
Contig_38	*dmoA_2*			1458	forward
Contig_52	*dmoA_3*			1377	forward
Contig_52	*dmoA_4*			1368	forward
Contig_23	*ethA*	FAD-containing monooxygenase EthA	1.14.13.-	1512	forward
Contig_6	*tftD_1*	FADH(2)-dependent monooxygenase TftD	1.14.14.-	1512	forward
Contig_127	*tftD_2*			1497	forward
Contig_6	*hsaA_1*	Flavin-dependent monooxygenase, oxygenase subunit HsaA	1.14.14.12	996	reverse
Contig_8	*hsaA_2*			1185	reverse
Contig_12	*hsaA_3*			1269	forward
Contig_29	*hsaA_4*			1137	reverse
Contig_58	*hsaA_5*			1170	forward
Contig_60	*hsaA_6*			1173	reverse

Contig_8	*hsaB_1*	Flavin-dependent monooxygenase, reductase subunit HsaB	1.5.1.36	606	reverse
Contig_23	*hsaB_2*			498	forward
Contig_60	*hsaB_3*			510	reverse
Contig_3	*mhuD*	Heme-degrading monooxygenase HmoB	1.14.99.3	309	reverse
Contig_16	*pvdA*	L-ornithine 5-monooxygenase	1.13.12.-	1311	reverse
Contig_6		Lactate 2-monooxygenase	1.13.12.4	1290	forward
Contig_1	*limB_1*	Limonene 1,2-monooxygenase	1.14.13.107	1026	reverse
Contig_2	*limB_2*			990	forward
Contig_11	*limB_3*			1035	forward
Contig_23	*limB_4*			1167	forward
Contig_58	*limB_5*			1140	reverse
Contig_64	*limB_6*			1170	forward
Contig_86	*limB_7*			876	reverse
Contig_87	*limB_8*			978	reverse
Contig_75	*mmoX*	Methane monooxygenase component A alpha chain	1.14.13.25	1638	forward
Contig_75	*mmoC*	Methane monooxygenase component C	1.14.13.25	1053	forward
Contig_38	*msuD*	Methanesulfonate monooxygenase	1.14.14.5	1110	reverse
Contig_12	*ntaA_1*	Nitrilotriacetate monooxygenase component A	1.14.14.10	1287	reverse
Contig_27	*ntaA_3*			1371	reverse
Contig_27	*ntaA_2*			1347	reverse
Contig_41	*ntaA_5*			1380	forward
Contig_41	*ntaA_4*			1173	forward
Contig_87	*ntaA_6*			1347	reverse
Contig_131	*ntaA_7*			1353	forward
Contig_3		Nitronate monooxygenase	1.13.12.16	1155	forward
Contig_3				1092	reverse
Contig_73				978	forward
Contig_134				1032	reverse
Contig_97	*pcpB*	Pentachlorophenol 4-monooxygenase	1.14.13.50	1638	reverse
Contig_5	*pamO_1*	Phenylacetone monooxygenase	1.14.13.92	1386	forward
Contig_13	*pamO_2*			1569	reverse
Contig_18	*pamO_3*			1482	forward
Contig_23	*pamO_4*			1602	forward
Contig_67		Putative ammonia monooxygenase		1152	reverse
Contig_51		Putative monooxygenase Rv0793	1.-.-.-	318	forward
Contig_30	*rutA_1*	Pyrimidine monooxygenase RutA	1.14.99.46	879	forward
Contig_65	*rutA_2*			987	forward
Contig_94	*rutA_4*			1086	forward
Contig_94	*rutA_3*			942	forward
Contig_3		Steroid C26-monooxygenase	1.14.13.141	1290	forward
Contig_3				1260	reverse

Contig_2	*styA*	Styrene monooxygenase StyA	1.14.14.11	1401	reverse
Contig_5	*iaaM*	Tryptophan 2-monooxygenase	1.13.12.3	1662	Reverse
Contig_27	*ntaB*	FMN reductase (NADH) NtaB	1.5.1.42	501	reverse
Contig_22	*ssuE_1*	FMN reductase (NADPH)	1.5.1.38	681	forward
Contig_27	*ssuE_2*			501	reverse
Contig_76	*ssuE_3*			501	reverse
Contig_94	*nfrA1*			798	Reverse
Contig_1	*soxC_1*	Dibenzothiophene desulfurization enzyme C	1.14.14.21	1200	forward
Contig_12	*soxC_2*			1254	reverse
Contig_28	*soxC_3*			1236	reverse
Contig_38	*soxC_4*			1251	reverse
Contig_52	*soxC_5*			1194	forward
Contig_52	*soxC_6*			651	forward

The shaded rows indicate the genes associated with desulphurisation of benzothiophene and dibenzothiophene occurring at various contigs of the draft genome sequence of the strain 213E. The annotated whole-genome sequence was researched for genes homologous to the *dsz* genes and as a result, genes *soxC*, *ntaA* and *ssuE* that correspond to *dszC*, *dszA* and *dszD*, respectively were identified explicitly, but a gene encoding with functionality of desulfinase (*dszB*) was not found from the annotations.

Appendix 7.10 Genes encoding monooxygenases and FMN reductase in the draft whole-genome sequence of strain 248

Sequence Name	Gene_ copy number	Product	EC_number	Length (bp)	Direction
Contig_97	*cinA_2*	1,8-cineole 2-endo-monooxygenase	1.14.13.156	1224	forward
Contig_19	*tfdB_1*	2,4-dichlorophenol 6-monooxygenase	1.14.13.20	1647	forward
Contig_22	*tfdB_2*			1386	forward
Contig_36	*tfdB_3*			1932	reverse
Contig_57	*tfdB_4*			1707	reverse
Contig_75	*tfdB_5*			1692	forward
Contig_5	*camP_1*	2,5-diketocamphane 1,2-monooxygenase	1.14.13.162	996	reverse
Contig_42	*camP_2*			1167	forward
Contig_42	*36*	3,6-diketocamphane 1,6 monooxygenase	1.14.13.-	1116	reverse
Contig_2	*kshA_1*	3-ketosteroid-9-alpha-monooxygenase oxygenase subunit	1.14.13.142	1161	forward
Contig_3	*kshA_2*			1185	forward
Contig_5	*kshA_3*			1161	forward
Contig_72	*kshA_4*			1206	reverse
Contig_100	*kshA_5*			1167	forward
Contig_12	*hapE_1*	4-hydroxyacetophenone monooxygenase	1.14.13.84	1482	reverse
Contig_13	*hapE_2*			1464	forward
Contig_30	*hapE_3*			1545	reverse
Contig_42	*hapE_4*			1959	reverse
Contig_43	*hapE_5*			1812	reverse
Contig_50	*hapE_6*			1491	forward
Contig_99	*hapE_7*			1602	forward
Contig_108	*hapE_8*			1548	forward
Contig_56	*hpaC*	4-hydroxyphenylacetate 3-monooxygenase reductasecomponent	1.5.1.36	594	reverse
Contig_5	*nphA1_1*	4-nitrophenol 2-monooxygenase, oxygenase component	1.14.13.29	1617	reverse
Contig_5	*nphA1_2*			1617	reverse
Contig_34	*npcA*	4-nitrophenol 4-monooxygenase/4-nitrocatechol 2-monooxygenase, oxygenase component	1.14.13.166	1587	forward
Contig_5	*npcB_1*		1.14.13.166	561	reverse

Contig_34	*npcB_2*	4-nitrophenol 4-monooxygenase/4-nitrocatechol 2-monooxygenase, reductase component		543	forward
Contig_39	*npcB_3*			513	reverse
Contig_16	*hspB*	6-hydroxy-3-succinoylpyridine 3-monooxygenase HspB	1.14.13.163	1209	reverse
Contig_31		6-hydroxynicotinate 3-monooxygenase precursor	1.14.13.114	1131	reverse
Contig_48				1212	forward
Contig_34	*luxA_1*	Alkanal monooxygenase alpha chain	1.14.14.3	993	reverse
Contig_63	*luxA_2*			1173	reverse
Contig_72	*luxA_3*			1173	forward
Contig_72	*luxA_4*			1122	forward
Contig_62	*luxB*	Alkanal monooxygenase beta chain	1.14.14.3	1029	reverse
Contig_5	*alkB1*	Alkane 1-monooxygenase 1	1.14.15.3	1233	reverse
Contig_10	*ssuD_1*	Alkanesulfonate monooxygenase	1.14.14.5	921	reverse
Contig_31	*ssuD_2*			930	forward
Contig_39	*ssuD_3*			1053	reverse
Contig_81	*ssuD_5*			651	reverse
Contig_81	*ssuD_4*			450	reverse
Contig_3	*hpaH_1*	Anthranilate 3-monooxygenase oxygenase component	1.14.14.8	1461	forward
Contig_19	*hpaH_2*			1458	reverse
Contig_1		Antibiotic biosynthesis monooxygenase		984	forward
Contig_6				558	reverse
Contig_7				303	forward
Contig_9				621	reverse
Contig_16				285	reverse
Contig_35				309	reverse
Contig_51				363	reverse
Contig_114				327	forward
Contig_114				276	reverse
Contig_62		Cyclohexanone 1,2-monooxygenase	1.14.13.22	1635	reverse
Contig_15	*cpnB_1*	Cyclopentanone 1,2-monooxygenase	1.14.13.16	1638	reverse
Contig_15	*cpnB_2*			1617	reverse

Contig_23	*cpnB_3*			1647	reverse
Contig_36	*cpnB_4*			1632	forward
Contig_48	*cpnB_5*			1650	reverse
Contig_60	*dmoA*	Dimethyl-sulfide monooxygenase	1.14.13.131	1416	reverse
Contig_23	*ethA_1*	FAD-containing monooxygenase EthA	1.14.13.-	1494	forward
Contig_66	*ethA_2*			1548	reverse
Contig_115	*ethA_3*			591	forward
Contig_2	*hsaA_2*	Flavin-dependent monooxygenase, oxygenase subunit HsaA	1.14.14.12	1179	reverse
Contig_2	*hsaA_1*			1167	reverse
Contig_3	*hsaA_3*			1176	reverse
Contig_5	*hsaA_4*			1182	forward
Contig_16	*hsaA_5*			1200	reverse
Contig_56	*hsaA_6*			1161	reverse
Contig_103	*hsaA_7*			1179	reverse
Contig_2	*hsaB_1*	Flavin-dependent monooxygenase, reductase subunit HsaB	1.5.1.36	537	forward
Contig_3	*hsaB_2*			588	reverse
Contig_3	*hsaB_3*			501	forward
Contig_5	*hsaB_4*			519	forward
Contig_12	*hsaB_5*			555	reverse
Contig_28	*hsaB_6*			522	reverse
Contig_42	*hsaB_7*			582	forward
Contig_111	*hsaB_8*			1044	reverse
Contig_20	*mhuD*	Heme-degrading monooxygenase HmoB	1.14.99.3	315	forward
Contig_1	*pvdA*	L-ornithine 5-monooxygenase	1.13.12.-	1347	forward
Contig_43		Lactate 2-monooxygenase	1.13.12.4	1287	reverse
Contig_60				192	forward
Contig_2	*limB_1*	Limonene 1,2-monooxygenase	1.14.13.107	897	forward
Contig_4	*limB_2*			1011	reverse
Contig_10	*limB_3*			975	forward
Contig_28	*limB_4*			1194	reverse
Contig_40	*limB_5*			1062	reverse
Contig_85	*limB_6*			966	forward
Contig_1		Luciferase-like monooxygenase		828	reverse
Contig_5				855	reverse
Contig_101				855	reverse

Contig_114	*mmoX*	Methane monooxygenase component A alpha chain	1.14.13.25	1635	forward
Contig_114	*mmoY*	Methane monooxygenase component A beta chain	1.14.13.25	1107	forward
Contig_52	*mmoC_1*	Methane monooxygenase component C	1.14.13.25	768	forward
Contig_114	*mmoC_2*			1044	forward
Contig_3	*msuD_1*	Methanesulfonate monooxygenase	1.14.14.5	1071	reverse
Contig_5	*msuD_3*			1170	forward
Contig_5	*msuD_2*			1125	forward
Contig_1	*ntaA_3*	Nitrilotriacetate monooxygenase component A	1.14.14.10	1362	forward
Contig_1	*ntaA_2*			1344	forward
Contig_1	*ntaA_1*			1296	forward
Contig_3	*ntaA_4*			1275	reverse
Contig_10	*ntaA_5*			1359	forward
Contig_38	*ntaA_6*			1299	forward
Contig_40	*ntaA_10*			1395	reverse
Contig_40	*ntaA_8*			1365	reverse
Contig_40	*ntaA_9*			1137	reverse
Contig_40	*ntaA_7*			915	reverse
Contig_55	*ntaA_11*			1365	reverse
Contig_83	*ntaA_13*			1296	forward
Contig_83	*ntaA_12*			900	forward
Contig_108	*ntaA_14*			1356	reverse
Contig_1		Nitronate monooxygenase	1.13.12.16	1116	forward
Contig_1				1062	reverse
Contig_7				990	reverse
Contig_11				939	forward
Contig_15				636	reverse
Contig_22				1029	reverse
Contig_22				990	reverse
Contig_23				981	forward
Contig_31				978	forward
Contig_45				963	forward
Contig_74				975	forward
Contig_74				576	reverse

Contig_74				438	reverse
Contig_140				987	reverse
Contig_1	*pcpB_2*	Pentachlorophenol 4-monooxygenase	1.14.13.50	1617	forward
Contig_1	*pcpB_1*			1500	reverse
Contig_5	*pcpB_3*			1521	forward
Contig_12	*pcpB_4*			1443	forward
Contig_21	*pcpB_5*			1545	reverse
Contig_64	*pcpB_6*			1428	forward
Contig_166	*pcpB_7*			1530	forward
Contig_5	*pamO_1*	Phenylacetone monooxygenase	1.14.13.92	1662	forward
Contig_58	*pamO_2*			1647	forward
Contig_80		Putative ammonia monooxygenase		1119	reverse
Contig_3	*moxC_1*	Putative monooxygenase MoxC	1.14.-.-	1392	forward
Contig_40	*moxC_2*			1059	reverse
Contig_112	*moxC_3*			1299	forward
Contig_25		Putative monooxygenase Rv0793	1.-.-.-	294	reverse
Contig_5	*ycnE*	Putative monooxygenase YcnE	1.-.-.-	324	reverse
Contig_1	*rutA_1*	Pyrimidine monooxygenase RutA	1.14.99.46	864	forward
Contig_1	*rutA_2*			864	reverse
Contig_6	*rutA_3*			873	reverse
Contig_34	*rutA_4*			888	reverse
Contig_36	*rutA_5*			915	forward
Contig_39	*rutA_6*			1131	reverse
Contig_63	*rutA_7*			1080	reverse
Contig_101	*rutA_8*			1167	reverse
Contig_1		Steroid C26-monooxygenase	1.14.13.141	1257	forward
Contig_1				1218	forward
Contig_5				1224	reverse
Contig_60	*styA*	Styrene monooxygenase StyA	1.14.14.11	1359	reverse
Contig_21	*tcmH*	Tetracenomycin-F1 monooxygenase	1.13.12.21	240	forward
Contig_4	*tsaM1*	Toluene-4-sulfonate monooxygenase system iron-sulfur subunit TsaM1	1.14.14.-	1074	reverse
Contig_48	*iaaM_1*		1.13.12.3	1671	forward

Contig_49	*iaaM_2*	Tryptophan 2-monooxygenase		1689	reverse
Contig_32	*xylA*	Xylene monooxygenase electron transfer component		753	Forward
Contig_3	*rutF*	FMN reductase (NADH) for Rut pathway	1.5.1.42	621	forward
Contig_22	*nfrA1*	FMN reductase (NADPH)	1.5.1.38	828	reverse
Contig_3	*ssuE_1*	FMN reductase (NADPH)	1.5.1.38	510	forward
Contig_72	*ssuE_2*			531	forward
Contig_2		NADPH-dependent FMN reductase		732	Forward
Contig_5	*soxC_1*	Dibenzothiophene desulfurization enzyme C		1212	Reverse
Contig_83	*soxC_2*			1230	Forward

The shaded rows indicate the genes associated with desulphurisation of benzothiophene and dibenzothiophene occurring at various contigs of the draft genome sequence of the strain 248. The annotated whole-genome sequence was researched for genes homologous to the *dsz* genes and as a result, genes *soxC*, *ntaA* and *ssuE* that correspond to *dszC*, *dszA* and *dszD*, respectively were identified explicitly, but a gene encoding with functionality of desulfinase (*dszB*) was not found from the annotations.