Sustainable Solvent Development: Exploring Bio-Based Alternatives Assisted By Computational Modelling

Tabitha Holly Marguerite Petchey

PhD

University of York

Chemistry

October 2018

Abstract

Solvents are the mainstay of many chemical reactions and formulations in the chemical industry. In recent years, concern regarding the potential long-term health and environmental risks of many existing solvents has vastly increased, leading to stricter legislation and pressure to discontinue their use. As such, there has been more focus on alternatives, from solvent selection guides informing users about solvent safety rankings, to research into new and existing compounds, usually from bio-available sources, which show potential as replacements.

In this thesis, the development of new bio-based solvents from a pilot-scale production platform molecule is explored. The synthesis and purification of lactones derived from levoglucosenone were studied extensively. Literature syntheses were critiqued and improved upon to scale-up the production of these compounds, previously disregarded as potential solvents. Also, p-cymene was investigated as an alternative solvent for direct amidation, with promising results, and thoroughly compared with its analogue, petrochemically derived toluene. In both approaches, computational modelling was employed to guide the decision-making process, and the limitations of the techniques used as predictive tools were assessed.

p-Cymene was found to be a suitable alternative to toluene for most crystalline amides, with the yield appreciably advantaged by the higher reflux temperature. The Yalkowsky approximation was shown to compare well with experimental solubilities of the amides and a flow system was put into practice. HSP of 5-membered ring lactones were predicted using HSPiP showing promise for some as NMP replacements. Furthermore, (S)-γ-hydroxymethyl-α,β-butyrolactone was made with 98% selectivity in a one-step method using an L-lysine catalyst and subsequently scaled-up to 175 g using a simplified method, then O-methylated. In initial tests, both lactones displayed good resistance to organic bases, faring better than Cyrene.

Contents

List of Figures

List of Schemes

List of Tables

Acknowledgements

Author's Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. Where experimental work has been done in collaboration with colleagues, this has been stated in the text. All sources are acknowledged as References.

Yalkowsky calculations, presented in Chapter 2, were originally run by Dr. James Sherwood and have since been modified by the author. MSc student, Martyna Kundrotaite, under the supervision of the author, conducted a study on the L-lysine promoted Baeyer-Villiger oxidation of levoglucosenone. *t*-Butylation of a lactone was done in collaboration with Dr. Fergal Byrne. Both of these have been included and acknowledged in Chapter 3.

The author of this thesis was lead author of the following publication; part of the work disclosed in Chapter 2 was published in this article:

- "Optimization of Amidation Reactions Using Predictive Tools for the Replacement of Regulated Solvents with Safer Biobased Alternatives" -
Tabitha H. M. Petchey, James W. Comerford, Thomas J. Farmer*, Duncan J. Macquarrie, James Sherwood, and James H. Clark, *ACS Sustainable Chem. Eng.*, 2018, **6**, 1550-1554

The author also collaborated on the following Review, which bears relevance to this thesis:

- "Tools and techniques for solvent selection: green solvent selection guides" - Fergal P. Byrne, Saimeng Jin, Giulia Paggiola, Tabitha H. M. Petchey, James H. Clark*, Thomas J. Farmer, Andrew J. Hunt, C. Robert McElroy and James Sherwood, *Sustain. Chem. Process*, 2016, **4**, 7-30

Chapter 1

Introduction

Global pollution and excessive resource use are driven by human ambition to create a more convenient, sensational, and luxurious existence. It is an uphill battle to reduce our burden on the planet, whilst trying to keep humanity in the lifestyle to which it has become accustomed. Add to this the ethical need to reduce socioeconomic inequality and it is almost impossible to reverse the tide of commercialism and industrial development. Thus, now more than ever, we have a responsibility to expand research into finding the most efficient, low impact methods to meet our collective global needs [Royal 2012].

1.1 Solvents

1.1.1 Overview

Solvents are essential facilitators across the board in the chemical industry [Cseri 2018]. Where so many other chemical components are specific to their application, solvents have always had a more generic, one-size-fits-all, arrangement. While reactants, catalysts, additives, etc. are scrutinised to the utmost extent, solvents are often ignored and chosen on the basis of convenience [Constable 2007]. This is largely because they do a comparatively simple job and, once they are known to work, there are few reasons to probe further. Now though, as so many existing solvents are derived from petrochemicals and harbour malignant properties that threaten the safety of workers and the environment, there are compelling grounds and increasing momentum to find alternatives.

What is a solvent?

A solvent is defined as the largest component in a solution, such as water in a beverage; it contains the many sugars and flavour compounds found therein, which are said to be dissolved, and each one is defined as a solute [IUPAC 2019a]. Importantly, solutes dissolved in a solution do not sediment as long as they remain under the same conditions, differentiating them from particles in a suspension. Although many solvents are used in liquid form, they may also include solids and supercritical fluids [IUPAC 2019b]. In this text, "solvent" will be used to describe a liquid which can be used to dissolve other materials, or is under investigation as such a substance.

General applications

In many chemical reactions, the solvent is by far the largest component in the reaction, even though it typically does not contribute towards the formula of the product. In general, it performs the function of dissolving the reactants to maximise the contact between them, and transferring heat to or away from the reaction as necessary [Kerton 2013]. However, use as a reaction medium is not the whole story since huge quantities of solvent are often used for purification, particularly in the pharmaceutical industry, and many have roles in consumer product formulations as well, such as paints and coatings, cleaning chemicals, and cosmetics [Constable 2007].

In reactions and formulations alike, the solvent is used to create homogeneity and so appropriate properties are key to good performance. Although solvents may be considered a benign, non-reactive component of a reaction, they can have a significant effect on the reaction equilibrium and kinetics. This is because the properties of a given solvent influence the stability of reactants and products as well as the transition state in ways including conformation and tautomeric equilibria [Reichardt 2011].

Negative health and environmental impact

There are some fairly intuitive reasons to look at the replacement of certain traditional solvents. For example, many are derived from petrochemicals, leaving a clear and escalating problem in long-term supply of those chemicals. In addition, hydro-

carbon solvents from oil are hydrophobic, which can be beneficial in applications such as purification and degreasing, but if released into the environment they are liable to bioaccumulate [Gissi 2015]. Regrettably, the negative qualities of a solvent, such as toxicity and flammability, may be linked to the same structural features that give the solvent advantageous properties for its purpose.

Another property of significance is volatility since, in many scenarios, it is necessary to remove the solvent at some juncture and a high volatility allows this to be done easily with minimal energy input by distillation. However, highly volatile solvents pose a greater risk of exposure to workers, are more likely to escape into the atmosphere, and are a more appreciable fire and explosion hazard if also flammable [Byrne 2016]. Chlorinated solvents have also long been known to harbour dangerous properties, for example chloroform is classified as toxic if inhaled and causes long-term organ damage through repeat exposure, including possible reproductive toxicity. The obvious replacement for chloroform is dichloromethane (DCM), but this in turn is a suspected carcinogen [ECHA 2018b]. Unfortunately, it is hard to replace such solvents because of their individual properties. They have a small amount of polarity due to asymmetry of the C-Cl bond geometry, but are insoluble with water, making for good separations. They are also non-flammable and not generally reactive, but they vapourise very easily, making the health implications that much more significant. This particular issue was addressed back in 1999 by the VOC Solvents Emissions Directive [EC 1999] and other legislation that is discussed further in the next section.

Legislation

The introduction of legislation designed to reduce the environmental damage caused by the chemical industry has led to a need to replace certain undesirable solvents to avoid legal repercussions. The European Union regulation concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) became law in the UK in 2007 and affects all chemicals manufactured or imported into the EU in >1 t a^{-1}. Naturally, this affects all solvents, which are bulk chemicals. Since then, companies have been required to register the chemicals that they manufacture or import and provide "a package of technical information on the chemical and its

hazards", which has happened over a long phase-in process that ended in May 2018 [EC 2006]. Some of the solvent-related restrictions imposed include halogenated solvents chloroform ($CHCl_3$), which:

> "Shall not be placed on the market, or used, as substances, as constituents of other substances, or in mixtures in concentrations equal to or greater than 0.1% by weight, where the substance or mixture is intended for supply to the general public and/or is intended for diffusive applications such as in surface cleaning and cleaning of fabric" [EC 2006]

and dichloromethane (DCM), which:

> "Paint strippers containing dichloromethane in a concentration equal to or greater than 0.1% by weight shall not be placed on the market...[or] used by professionals." [EC 2006]

As previously mentioned, both of these substances are now classified as suspected carcinogens. Other solvent restrictions include toluene and dipolar aprotics such as N-methyl-2-pyrrolidone (NMP), which will be facing restriction in 2020 [ECHA 2018c]. A solvent of a similar bearing, N,N-dimethylformamide (DMF), has not yet been restricted but appears on the 'Candidate list of substances of very high concern' due to its reproductive toxicity [ECHA 2018a]. Those chemicals not in compliance with the regulations will face removal from the market by the 'rapid alert system for dangerous non-food products' (RAPEX) [RAPEX 2018]. Certain chemicals classified as persistent organic pollutants have been banned altogether, according to the Stockholm Convention on Persistent Organic Pollutants (POPs) [Stockholm 2018].

Further legislation refers specifically to the atmospheric implications of solvent release. For example, the VOC Solvents Emissions Directive was passed in 1999 by the European Commision (EC Directive 1999/13/EC) and required member states of the European Union to reduce volatile organic compounds (VOCs) from a wide range of applications, including coatings, surface cleaners and dry cleaning fluid [EC 1999]. Prior to this, the USA passed the Clean Air Act in 1970, which faced ammendments in 1990 including a final ruling in 1998 to limit solvents in consumer

products (National Volatile Organic Compound Emission Standards for Consumer Products) to reduce the build up of unsafe ozone levels at ground level [EPA 1998].

Whilst many of the restrictions are imposed on consumer products rather than those for industrial use, there has been some leniency to account for commercial considerations affected by complete bans and it is likely that these restrictions will become more strict in the future, especially as more health-related information comes to light. History shows that, although the upheaval may be significant, sizeable legislative constraints can be passed when there is enough evidence of harm. For example, the Montreal protocol of 1987, which restricted the use of ozone-depleting substances (such as carbon tetrachloride) due to increasing concern over rates of skin cancer, was ratified globally and resulted in vital repair [UNEP 2018].

Pharmaceutical perspective

Solvent usage is of particular significance in the pharmaceutical industry where syntheses are complex and there are very high standards of purity, resulting in many steps and sensitive purification methods such as recrystallisation and chromatography. In 2007, the American Chemical Society Green Chemistry Institute Pharmaceutical Roundtable (ACS GCIPR) conducted a study to highlight the types of materials used in the manufacture of an active pharmaceutical ingredient (API) [Henderson 2011]. Figure 1.1 shows the relative amounts used in a typical process.

With solvents making up 56% of the inputs (far more if water is considered a solvent) into a standard process, it is unsurprising that the pharmaceutical industry produces the highest amount of waste per kg of product than any other sector of the chemical industry. Back in 2000, the figure stood at anywhere between 25 and 100 kg kg^{-1} of product, compared to 1-5 kg kg^{-1} for bulk chemicals [Sheldon 2000]. It is unlikely that this figure has changed significantly since publication due to the extensive purification processes previously mentioned, although strides have been taken to improve solvent utilisation choices (see Section 1.1.3). This scale of use and high turnover of research are likely the reasons that the pharmaceutical industry has taken the lead on tackling our dependence on toxic and unsustainable solvents, which is discussed further in section 1.1.3 [Constable 2007, Henderson 2011, Prat 2016].
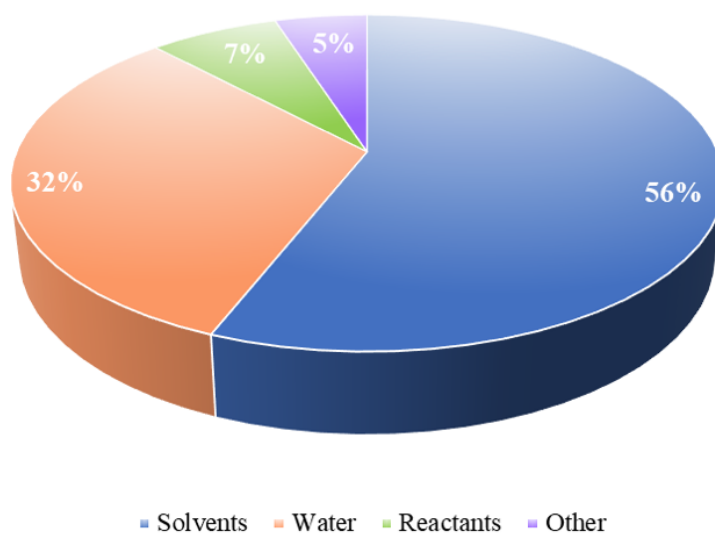
Figure 1.1: Summary of materials used (by mass) in the pharmaceutical industry. From ACS GCIPR. Data from [Henderson 2011].

1.1.2 Green Solvents

The study of green solvents, much as any other area of green chemistry, carries the ambition to reduce the environmental and human health impact brought about by chemical processes and products of the chemical industry, including the energy associated with production and use of a chemical [Byrne 2016]. To this end, there has been a lot of interest surrounding the development of safer, more sustainable solvents than many currently in use among practitioners of green chemistry [Anastas 2013, Clark 2017].

Bio-Based alternatives

As highlighted by Byrne et al. in 2016, the renewability of solvents has been far less extensively addressed than the hazards and environmental fate of those solvents [Byrne 2016]. While so many solvents currently in use are derived from dwindling petrochemical sources, it is imperative to put some emphasis on the origin of the solvent. Nonetheless, significant strides are being made in this area of research [Gu 2013].

The initial steps towards sustainable solvents have been finding alternative routes to known solvents, in particular bioethanol and, via that, ethyl acetate [Carotenuto

2013]. However, in more recent years, there has been much more research into previously unused solvent candidate molecules synthesised from renewable sources [Clarke 2018]. A few notable examples are shown in figure 1.2.

(a) GVL   **(b)** TMTHF   **(c)** PC
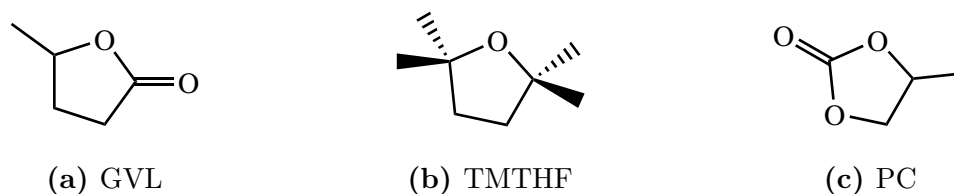
Figure 1.2: Structures of γ-valerolactone (GVL), 2,2,5,5-tetramethyltetrahydrofuran (TMTHF) and propylene carbonate (PC).

Lactones, ethers and cyclic carbonates have attracted a fair amount of interest in green chemistry research as they can be synthesised from renewable sources. γ-Valerolactone (GVL) can be made from the hydrogenation of bio-based platform molecule, levulinic acid, a process usually catalysed by supported ruthenium [Piskun 2016]. It has already found use as a dipolar aprotic solvent in mild biomass degradation [Shuai 2015], as well as in palladium catalysed cross-coupling reactions [Ismalaj 2014].

2,2,5,5-Tetramethyltetrahydrofuran (TMTHF) has recently been synthesised in a 3-step process from acetone and acetylene (not necessarily but potentially bio-derived) using hydrogenation with a palladium on alumina catalyst, followed by cyclisation using an H-Beta zeolite catalyst. It has been shown to perform well in Fischer esterifications and amidation reactions, and has the advantage of not forming peroxides, which is a hazard of other ethers [Byrne 2017].

Propylene carbonate can also be synthesised catalytically from bio-derived propylene glycol and CO_2, so in addition to producing a useful solvent, it also represents a form of carbon capture [Castro-Osma 2015]. PC has been reported as a successful solvent for use in the Heck reaction [Parker 2014].

This is by far not the limit of bio-based solvent research and recent reviews shine more light on the subject [CalvoF. 2018 Clarke 2018, Pena-P. 2015, Gu 2013]. Other platform molecules and derivatives specific to this work (D-limonene and levoglucosenone) are discussed in section 1.3.

Supercritical fluids, ionic liquids, switchable solvents and aqueous systems

Aside from direct replacements for traditional solvents, as bio-based alternatives tend to be, there are a number of other techniques that have been developed to tackle the health and environmental problems associated with many conventional solvents. It is beyond the scope of this work to explore these in great depth, but is important to be cognisant of the types of alternative research in this field.

The first to be mentioned is the use of supercritical fluids. A substance becomes supercritical when it is raised above its critical temperature and pressure [Eckert 1996], which causes alterations to the physical properties of the substance. As the temperature tends towards the critical point, it causes thermal expansion of the liquid fraction so that it becomes less dense, while the increasing pressure causes the gas fraction to become more dense so that they eventually become indistinguishable beyond the critical point [Hunt 2018]. The most prevalent of supercritical fluids is supercritical carbon dioxide ($scCO_2$), which has long had applications including separation and extraction, which can be of value in pharmaceutical science [Subram. 1997]. It is considered a green alternative because CO_2 is environmentally benign (when not being produced to excess), non-toxic and has a low critical temperature (31.1 °C) leading to a low energy cost.

The next type of solvent to mention is the ionic liquid, so called because it is made up entirely of ions. It is a salt with a low melting point. Well known examples of this include imidazolium-based ionic liquids due to their low viscosity and high stability to reaction conditions [Ghandi 2014]. This type of solvent has been promoted as green on the basis of low volatility and flammability, hence they pose a low risk of exposure to workers and will not escape into the atmosphere. However, this idea has remained controversial due to the high toxicity of some of these liquids and that it can be difficult to remove other chemicals from the solution [Pham 2010]. They remain a large arm of the green solvents research area and more is being done to produce them from sustainable sources [Hulsbosch 2016].

Switchable solvents aim to address the problems with multistep chemical processes, which may require a different solvent at each step to account for the different solvation parameters required. The idea is to create a solvent which can be induced

to alter its properties accordingly, reducing waste significantly [Herrero 2017]. The first examples of solvents like this were the switchable-polarity solvents designed by Jessop et al. [Jessop 2005], which were "switched" to higher polarity by the introduction of CO_2 at 1 bar, and back to lower polarity by heating or with the introduction of N_2 [Jessop 2012]. While less well established than all other solvent types mentioned, switchable solvents are certainly an innovative way of perceiving the potential for solvent use.

One of the more obvious solvents of choice for the green chemist is, of course, water, due to its neutral impact on the environment and complete lack of health implications. It is also non-flammable. The two main drawbacks of using water as a solvent are the amount of energy it takes to evaporate water due to its strong hydrogen-bonding ability, and the poor solubility of many organic molecules in water. However, this can be advantageous in promoting two hydrophobic species to bind together [Breslow 2010].

Solvent-free reactions

As green chemists it is also important to consider the possibility of using no solvent at all. Instinctively, it is clear that the fewer inputs required in a reaction, the less wasteful it will be, and less dependent on resources and high energy requirements to produce the inputs. There are examples of reaction techniques requiring no solvent, for instance, mechanochemistry, which involves the use of mechanical energy by the means of ball-milling or other grinding action [James 2012]. Successful solvent-free reactions include certain reduction, oxidation, elimination, rearrangement and hydrolysis reactions, among others, which have been comprehensively described by Tanaka, up to 2003 [Tanaka 2003].

While it is relevant to keep the possibility in mind, working without a solvent is not feasible for all reactions due to the safety of rapid heat transfer, and chemical formulations rely on solvents for consistency and performance. Solventless processing cannot usually be applied to pharmaceuticals which, as has been described, rely heavily on solvent use for purification purposes [Byrne 2016].

1.1.3 Solvent Selection

Selection guides

If a chemist has the intention to pursue a more environmentally benign chemical reaction by substituting in a greener solvent, they will have to ask themself how they should choose. There are many aspects to consider, some of which have been mentioned already, for example, high volatility implies that a solvent is likely to cause worker exposure and be released into the atmosphere. Low water solubility indicates that a solvent is likely to bioaccumulate. Various pieces of toxicity data will show that two different solvents are hazardous to health, one in the short term and one in the long term [Rama 2014]. These being just a few of the relevant details, they may have a hard time rating the green credentials of one solvent against another and it is not really practical to perform a life cycle analysis for each option they might pick.

It was with this in mind that some major pharmaceutical companies began the development of their "solvent selection guides", largely to help advise employees in their process development, but the data has been published and is potentially useful to other chemists. The three main contenders in this were SmithKline Beecham (now GSK) who originally published the first solvent selection guide in 1999, using the now familiar "traffic-light system" to give a clear visual distinction between good and bad solvents [Curzons 2007]. The guide broke solvents down into four categories of waste, impact, health and safety, and gave each solvent a score out of 10 in each category. Pfizer and Sanofi later followed suit [Alfonsi 2008, Prat 2013], with a specific medicinal chemistry slant. Sanofi's guide expanded the range of solvents contained and incorporated legislation into the guidance which added to the applicability. Pfizer also provided a "solvent substitution guide", which administers the user with relevant alternatives to adopt in preference to undesirable solvents.

GSK have since updated their guide to incorporate legislative "flags" and a life cycle score, which makes the set more comprehensive [Henderson 2011], if a little unwieldy. In their publication, GSK also included a "quick reference guide" which is somewhat more user-friendly, and is summarised in figure 1.3.
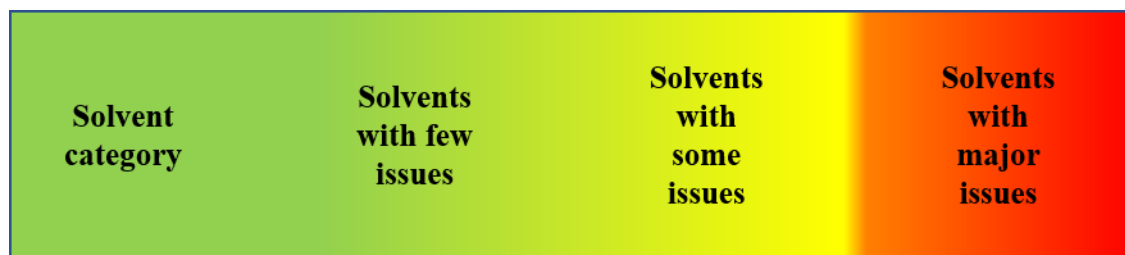
Figure 1.3: A representation of GSK's quick reference guide for solvent selection [Henderson 2011].

Ashcroft et al. reported that GSK had reduced dichloromethane from the third most used solvent at the time of the first selection guide, to eighth most common in 2005, demonstrating a sharp drop [Ashcroft 2015]. Similarly, Pfizer had reportedly reduced its use of chloroform by 99% between 2007 and 2015. This is suggestive of the positive effect of raising awareness among employees.

Substitution towards modern solvents

The story thus far shows just how much care should be taken when making solvent selections. The selection guides provided by industry demonstrate that there is concern over the negative impacts of solvent use. However, there are not many listed in the way of green solvents, other than ethanol, meaning that substitution options are limited, especially in the dipolar aprotic region, where dimethyl sulfoxide (DMSO) is the only listed alternative. Although not toxic, DMSO can carry other compounds through the skin on exposure, and can decompose violently at high temperatures [Fierz 1994]. This should serve to inform the prevailing research trends to (1) keep up the search for new alternative solvents, and (2) probe the properties of these solvents to the largest extent in order to engender confidence in industry that will get new solvents onto the substitution lists.

In 2016, CHEM21, a collaboration of the above pharmaceutical companies, alongside Charnwood consultants and the University of York, published another solvent selection guide, which has a strong basis in the Globally Harmonized System (GHS) on classification and labelling and was developed by review of the pre-existing selection guides [Prat 2016, 2014]. The ranking system is based on the worst attribute of each solvent, differentiating it from the prior publications. The highlight of this collaboration is the inclusion of "less classical solvents", which

marks a step forward in acceptance of these new molecules.

1.2 Predictive Modelling

1.2.1 Hansen Solubility Parameters in Practice (HSPiP)

Development

In 1936, the founder of solubility theory, Joel Hildebrand, defined a parameter to describe the behaviour of a solvent as the square root of the cohesive energy density, which later became known as the Hildebrand solubility parameter, δ [Hildebrand 1936]. This parameter relies on the observation that the solubilising power of a liquid is correlated with its cohesive energy density, since the same attractive intermolecular forces which must be broken to gasify a liquid, must also be broken to dissolve another substance in the liquid. The simplicity of the Hildebrand solubility parameter is that it is a measure of the total intermolecular forces of a solvent and is quite reliable for non-polar, non-hydrogen bonding species. However, to define solubility with a broader scope, it is necessary to understand that the type and contribution of intermolecular forces also play a part and must be defined individually to really follow the "like dissolves like" adage.

In 1967, Charles Hansen first defined a method for describing solubility, based on 3 parameters, namely hydrogen-bonding (δH), polarity (δP) and dispersive (δD) or van der Waals forces [Hansen 1967], building on the previous work of Hildebrand. With the help of Stephen Abbott and Hiroshi Yamamoto, this set of parameters was used as the basis to develop a solvent modelling program known as Hansen Solubility Parameters in Practice (HSPiP), which allows the user to define the parameters of a solvent along three axes and, thus, get a clear visual idea of how it compares to other solvents (along with a numerical output) [Hansen 2018]. In this work, "HSP space" will be the term used to describe the 3-dimensional graphical output of the HSPiP software.

Theory

HSPiP is designed to be usable by the practical chemist or chemical engineer, so it is possible to put to use without delving deeply into the mathematics behind it. However, it is useful to have a basic understanding of the principles to help interpret the results and to identify any unexpected anomalies. The theory presented was developed by Hansen in his thesis [Hansen 1967], but the information is publicly available on the HSPiP accompanying website [Hansen 2018]. Hansen Solubility Parameters (HSP) are based on thermodynamics and are used entirely to predict whether or not something can dissolve. They do not consider kinetics and so do not describe how quickly a substance will dissolve. HSPiP can be used to model solvent and solute molecules. Where the solvent is described in the theory below, it can be assumed that the same applies to the solute.

HSP use the same basic theory of Hildebrand, in that the square of the total parameter is equal to the cohesive energy density, c (i.e. cohesive energy per molar volume). However, the total parameter is divided into the three separate parameters described above, (δH), (δP) and (δD). Thus, the square of the total parameter is the sum of the squares of each individual parameter.

$$c = \delta^2 = \delta D^2 + \delta P^2 + \delta H^2 \tag{1.1}$$

While hydrogen bond forces can be considered a type of polar force, it is useful to separate these two parameters since there are many solvents which have a high polarity but exhibit little hydrogen-bonding ability. In real-world applications, this can have a big effect on how the solvent performs. The result of plotting these three parameters in a 3-dimensional graph is that the distance between two points gives an idea of how similar the two solvents are and whether they are likely to mix. The closer they are in HSP-space, the more thermodynamically favourable mixing will be. A calculation to determine the distance is given in equation 1.2, although the software can provide this information.

$$\text{Distance}^2 = 4(\delta D_A - \delta D_B)^2 + (\delta P_A - \delta P_B)^2 + (\delta H_A - \delta H_B)^2 \tag{1.2}$$

The factor of four in equation 1.2 relating to dispersion forces is an approximation, allowing for the fact that the differential heat of mixing term is considered larger for dispersive forces than polar and hydrogen bonding forces. HSPiP has continued to use this factor as it produces spherical plots and suits much of the practical testing that has been carried out.

It is worth noting that HSP do not work for ionic solids or primarily aqueous solutions. This is in part due to the fact that adding a fourth dimension to cover ionic interactions would be impractical to visualise graphically, and aqueous solutions require their own interrogation due to the unusual properties of water [Hansen 2018]. However, ionic and aqueous systems are not modelled in this work.

Practical Uses

If the aim is to find a replacement solvent (or combination of solvents) for use with a known solute, HSPiP makes it possible to accurately predict what solvents could be used for this purpose. To achieve this, the solubility parameters of the solute are found experimentally by dissolving the compound in a range of solvents covering a wide area of HSP space. The researcher can then enter their results into the software using a scale of 1 to 6 to indicate how well the compound dissolved in each solvent. The software can then calculate a "sphere of solubility" within which the solvents sit, and outside which the non-solvents sit. The software produces a 3-dimensional graph (figure 1.4a), which is easy to use in a virtual setting. For the purposes of a report, it can be more useful to show the graph from three different 2D perspectives (figure 1.4b).

The green dot in figure 1.4 represents the solute and is surrounded by a green sphere, which represents an area of solubility. Once the HSP of the solute are know, this 3D image can be used to predict whether untested solvents could be used to dissolve the known compound, since the solvent HSP should also sit within the sphere. Another benefit of the software is that HSP of a mixture of solvents can be calculated, as they are simply the average of the vectors modified by the percentage contribution of each solvent. This means that, although there may be no "desirable solvents" within the sphere of solubility, two outliers may combine to form a mixture for which the HSP do fall within the sphere. It is easy to see how this fairly simple
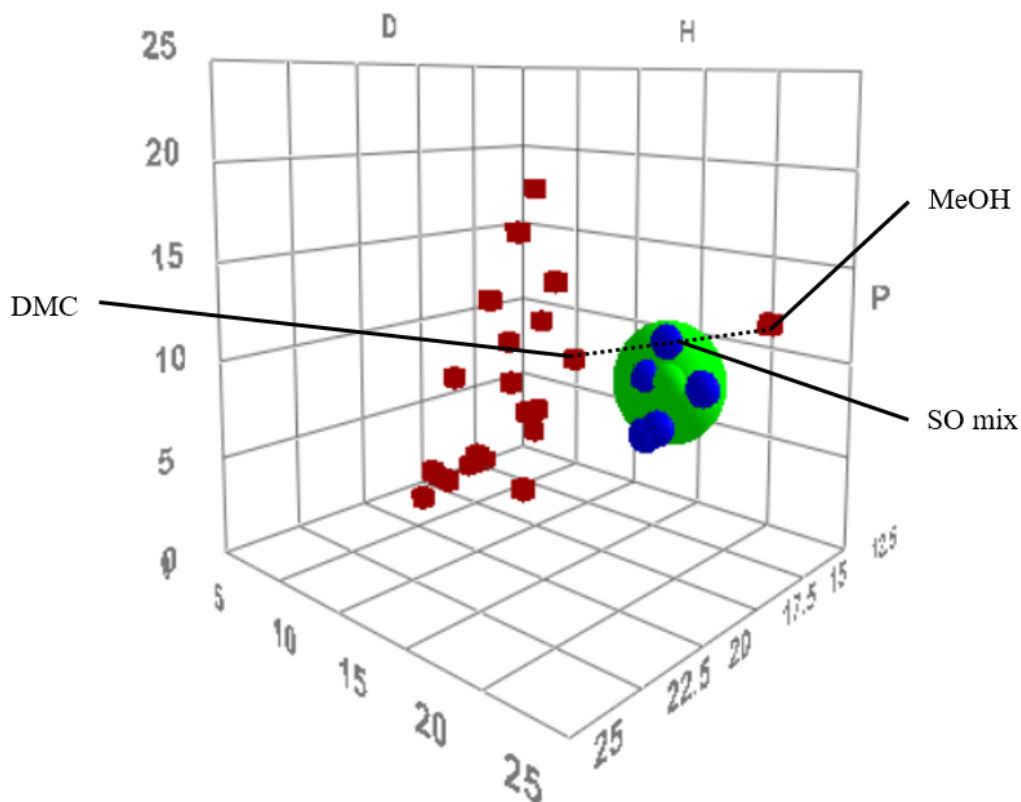
(a) 3-Dimensional graphical output from HSPiP software. A representative selection of commonly used solvents is shown by red and blue dots. The solvent mix (SO mix) is a 50:50 mixture of dimethyl carbonate (DMC) and methanol (MeOH), and falls within the sphere of solubility for this fictional compound.
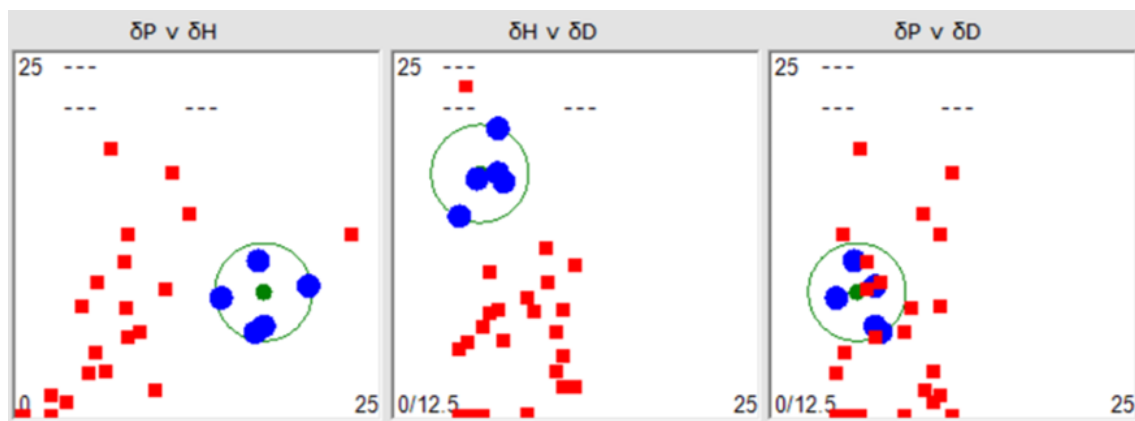
(b) 2-Dimensional representation of figure 1.4a.

Figure 1.4: An example of HSPiP use and visual output of the software: (a) 3D perspective and (b) 2D perspective.

software could save a lot of time in the endeavour to replace undesirable solvents.

Group contribution

A group contribution method divides the target molecule (or mixture) up into individual functional groups and uses mathematical assumptions about those individual groups to build a prediction about the character of the whole molecule [Su 2017]. The extent of specificity about the groups can theoretically be altered according to the users choice, with an effect on the accuracy. For example, if a C=O carbonyl group is the group of choice, it would be possible to make predictions based on the assumption that all C=O groups have identical behaviour, and likewise for other similarly sized groups. However, it is also possible to define several different types of C=O groups according to what atom is adjacent, and even two bonds away, and whether it is a cyclic structure, etc. The more specificity that is used for the predictive method, the more accurate the outcome is likely to be, but this is also at the loss of statistical data on which to base the assumptions. Additionally, the calculation could become extremely complex, which in practice would increase the waiting time significantly.

Group contribution methods include the Jankowski method [Jankowski 2008], designed primarily for use in biochemical systems, and the Stefanis-Panayiotou method [Stefanis 2008], designed for pure organic compounds. For the purposes of this work, the predictive method designed for HSPiP by one of its developers, Dr. Hiroshi Yamamoto, has been used to make predictions about the HSP of various compounds. Therefore, it seems appropriate to give a short summary of how it works.

As group contribution does not generally account for the geometry of a molecule, only the number of groups present, this can lead to inaccurate predictions of δP since the polarity of a molecule as a whole can be affected by the relative positions of two polar groups. Yamamoto's so called "Neural Network" (NN) methodology accounts for interactions between different groups. He devised a "Molecular Breaking" programme as part of the Y-MB method, which divides a given molecule into subgroups [Pirika 2018]. This is done automatically by the software after inputting the SMILES code. The behavioural contributions of each group is modified by its NN interactions and summed to give the final HSP prediction [Yamamoto 2018].

1.2.2   COSMO-RS

Background

In 1993, Andreas Klamt and Gerrit Schüürmann first published work detailing a new approach to solubility modelling, akin to pre-existing conductor screening models [Klamt 1993]. He named the algorithm COSMO or "COnductor-like Screening MOdel". In this approach, the molecule is assumed to be encapsulated in a perfect virtual conductor described by Klamt as a "dielectric continuum of permittivity". The cavity occupied by the molecule is edged by the so-called "σ-surface", an interface between the molecule and dielectric continuum, which is also the "solvent-accessible surface". This essentially describes the charge density at the surface of the molecule.

Klamt noted some limitations of the dielectric continuum approach and published an improved model, known as COSMO for Realistic Solvation (COSMO-RS), which is a statistical thermodynamics theory based on the original COSMO [Klamt 1995]. The dielectric continuum approach has difficulty accounting for solvents with dielectric constants very close to one another, but where the solvent properties differ. Klamt claims that COSMO is effective because of its "algorithmic simplicity" and "numerical stability" [Klamt 2018].

COSMOconf and COSMOtherm

COSMO-RS can be used in the following way to make solubility predictions. First, a molecular structure is drawn by the user in a molecular modelling program (ArgusLab, which can perform a geometry optimisation, was used for this work). The molecular data file is imported into the COSMOconf program, which can be used to calculate the most likely conformation of the molecule. This calculation can take hours or days depending on the complexity of the molecule. Once this is done, the program generates a set of conformer files, listed from lowest to highest energy. In this research, the ground state conformer has always been used for COSMOtherm modelling.

Once the conformational analysis is complete, the data files can be imported into the COSMOtherm program, which calculates the σ-surface. An example of a con-

former calculated using COSMOconf, then the σ-surface added using COSMOtherm, is given in figure 1.5.
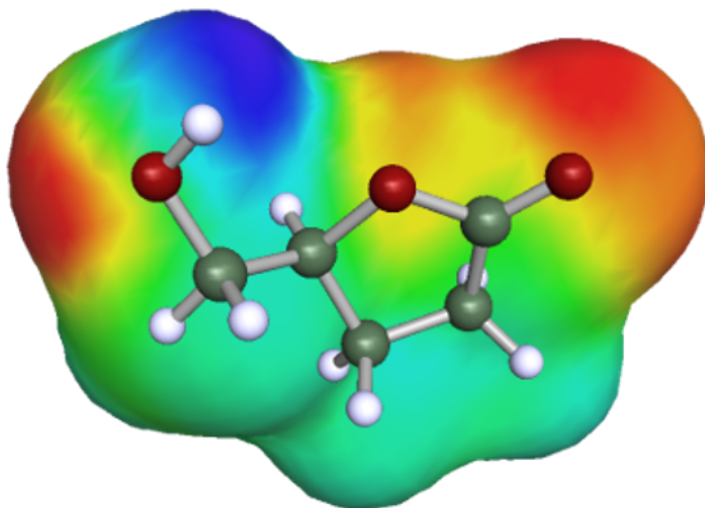
Figure 1.5: Conformational analysis of (S)-γ-hydroxymethyl-α,β-butyrolactone using COSMOconf showing the σ-surface, calculated using COSMOtherm. Blue represents areas of positive charge density. Red represents areas of negative charge density.

Bright blue represents areas of positive charge density while bright red corresponds to negative. Areas between yellow and bright blue are only weakly polar. Once the σ-surface has been calculated, this provides all the information needed for COSMOtherm to calculate the chemical potential, which is visualised in the form of two graphs exemplified in figure 1.6.

(a) σ-Profile
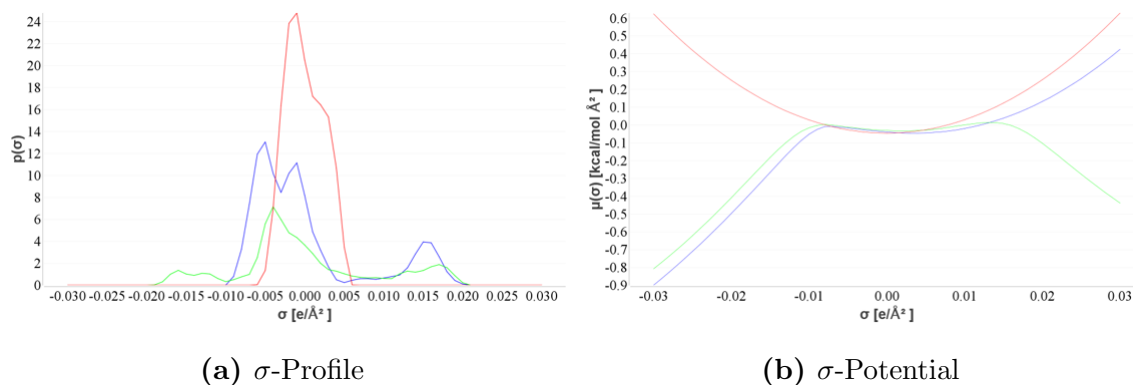
(b) σ-Potential

Figure 1.6: An example of the σ-profile and σ-potential output of COSMOtherm. Red = hexane; green = methanol; blue = dimethylformamide.

The σ-profile in figure 1.6a shows how much of the surface is covered by what charge, with positive polarity being shown towards the left of the x-axis, and negative polarity represented towards the right. A nonpolar, hydrocarbon solvent will tend to have a large peak around 0 e $\mathring{A}^{-2}$ because of its low polarity, whereas an alcohol, such as methanol in figure 1.6, manifests as a slight peak in the negative, hydrogen-bond donor, region and another peak in the positive, hydrogen-bond acceptor, region. COSMOtherm also uses the σ-surface to calculate the σ-potential of the molecule, as shown in figure 1.6b. It illustrates the chemical potential $\mu_s(\sigma)$. To use the example of a non-polar hydrocarbon again, this would exhibit a U-shape, because it would be energetically unfavourable to make contact with a hydrogen-bond donor or acceptor. In the case of methanol, the opposite is observed, since it has both hydrogen-bond donating and accepting characteristics [Durand 2011].

1.2.3 The Yalkowsky Approximation

Arguably, the simplest method for estimating solubility is to use the relationship between the melting point (°C) of the solute and its crystal-liquid fugacity ratio (CLFR) as described by Yalkowsky and Wu [Yalkowsky 2010].

$$\text{Log CLFR} = -0.01(mp - 25) \tag{1.3}$$

This approximation uses the melting point of the solute to predict its solubility at a given temperature. The CLFR describes how the crystallinity of a solute affects its solubility in an ideal solvent, since the CLFR is equal to the solubility, and is given by mole fraction. Due to generally weaker intermolecular interactions in substances of lower melting points, it stands to reason that less energy is generally required to break them, which in this case, translates to higher ideal solubilities. The CLFR is not specific to the solvent, only the solute. This approximation assumes that solubility decreases linearly with increase in the melting point of a substance and other factors are excluded. It is, therefore, used as a guide with the expectation that some variance should be expected due to moderating influences such as the amount of interaction between solvent and solute molecules.

1.3   Bio-Based Solvents

1.3.1   *p*-Cymene vs. Toluene

Comparison

Toluene is a widely used aromatic hydrocarbon solvent with multiple uses including as a reaction medium and as an industrial cleaning agent [Paggiola 2016]. Toluene is already listed by the European Chemicals Agency (ECHA) with the following restrictions for use:

> "Shall not be placed on the market, or used, as a substance or in mixtures in a concentration equal to or greater than 0.1% by weight where the substance or mixture is used in adhesives or spray paints intended for supply to the general public." [EC 2006]

As toluene is derived from petrochemicals, one of the crucial aspects to consider is its sustainability. In looking for a suitable bio-based alternative, there is one structure which looks both promising and structurally similar, *p*-cymene, shown in figure 1.7c.

(a) Toluene (111 °C)   **(b)** D-Limonene (176 °C)   **(c)** *p*-Cymene (177 °C)
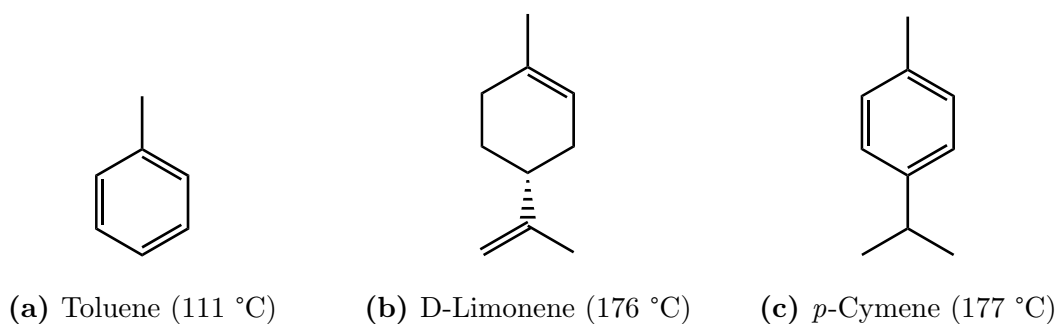
Figure 1.7: Structures of toluene, D-limonene and 4-methylisopropylbenzene (*p*-cymene) with boiling points.

p-Cymene, is a known aromatic compound related to the terpene family, which can be produced by the alkylation of toluene [Vora 2003]. However, there are also production routes to *p*-cymene from dipentenes, particularly D-limonene (figure 1.7b). D-Limonene is a naturally occurring terpene, prevalent in citrus peel, and thus a by-product of the citrus processing industry [Paggiola 2016]. Besides the sustainability perspective, there are also health benefits to making this substitution.

Table 1.1 provides a summary of the hazards associated with each of these three substances.

Table 1.1: Known hazard information for toluene, D-limonene (dipentene) and *p*-cymene. Hazard statements gathered from the European Chemicals Agency(ECHA) website [ECHA 2018b].

Solvent	Hazard statements
Toluene	Highly flammable liquid and vapour, is suspected of damaging the unborn child, may be fatal if swallowed and enters airways, may cause damage to organs through prolonged or repeated exposure, causes skin irritation and serious eye irritation, may cause drowsiness or dizziness, suspected of damaging fertility or the unborn child, is harmful to aquatic life with long lasting effects.
Dipentene	Flammable liquid and vapour, causes skin irritation and may cause an allergic skin reaction, very toxic to aquatic life with long lasting effects.
p-Cymene	Flammable liquid and vapour, may be fatal if swallowed and enters airways, causes serious eye irritation, skin irritation and may cause respiratory irritation.

It is clear from table 1.1 that there are considerable benefits to using *p*-cymene over toluene, as it is less flammable, and does not have the same known long-term health effects. It is also less likely to cause exposure by inhalation, since it is less volatile.

Production of *p*-cymene from biomass

Whilst *p*-cymene is still produced by alkylation of toluene [Vora 2003], there have been promising studies demonstrating its potential as a bio-based compound. Martin-Luengo et al. published a study reporting 100% conversion of limonene, by using a sepiolite catalyst which had been modified by either nickel, iron or manganese, forming *p*-cymene in yields of 88, 82 and 77%, respectively [Martin-Luengo 2010]. The reaction was done fast and efficiently by using microwave heating for 20 minutes

(165 °C). The drawback to this reaction was that large amounts of solid (up to 500 mg per 5 mL limonene) were required.

Nguyen et al. subsequently reported a similar technique using Vietnamese montmorillonite, which naturally contains various metal oxides. They acidified the catalyst using 2M H_2SO_4 and achieved a selectivity of 84% and yield of 75% of p-cymene by heating the solventless mixture to 163 °C, also by microwave irradiation. Similarly, the reaction lasted only 19 minutes and used 10 wt% of catalyst to limonene [Nguyen 2012]. At present, research is still needed to improve the efficiency of this conversion, but it is necessary to pursue green sources of solvents simultaneously with finding suitable solvent candidate molecules.

Industry practicality

In 2016, Paggiola et al. produced a publication regarding the use of toluene as a cleaning agent in industry, and its potential for replacement by limonene extracted from citrus waste [Paggiola 2016]. The authors surveyed an extensive selection of publicly available market data to estimate the types and volume of use of toluene, as well as the production volumes and location of citrus and citrus waste. As the leading use for citrus, only juicing was considered, while other processes such as the production of preserves, which may not generate peel waste, were disregarded.

The study estimated that 65 million MT of citrus waste are produced globally each year. Based on the consumption of 14.8 million MT of toluene globally in 2014 (about 65% of which went into synthesis of other chemicals, with 13.5% going into paints and coatings, and 5.5% being used for cleaning), the authors concluded that total global substitution of toluene by limonene (and by implication p-cymene, if obtained from limonene) would not be feasible. However, they were positive about the possibility of replacing toluene in industrial cleaning for citrus-producing countries, such as Brazil, where the D-limonene potential from juicing alone is more than twice as high as the industrial use of toluene for cleaning.

In relation to this work, the implication is that limonene and p-cymene from limonene cannot be universal replacements for toluene, but the potential for limonene recovery from waste sources is sufficiently high as to have value in targeted uses. It also highlights that there is not going to be a "quick fix" to solvent substitution and

chemists must be prepared to use intelligent and varied strategies to cope with this.

1.3.2   Levoglucosenone and Cyrene

Development and production

The identification of levoglucosenone (LGO), **1** (scheme 1.2), from the acidic pyrolysis of cellulosic biomass was first made in 1973 [Halpern 1973], following various reports of the reduction in yield of levoglucosan (figure 1.8) during pyrolysis when an acid catalyst was used, and simultaneous production of an unidentified compound [Tsuchiya 1970]. Prior to this, acid and base catalysts in this process were primarily used as fire retardants, since it was assumed that decreased levoglucosan yields from pyrolysis were correlated with lower tar yields and, therefore, reduced flammability of the resultant product [Tamaru 1951].
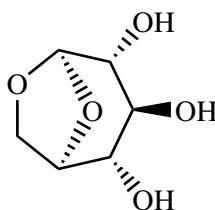
Figure 1.8: Structure of levoglucosan, a precursor in the formation of levoglucosenone (LGO), first named in 1918 [Pictet 1918], and formally identified in 1929 [Josephson 1929].

The structure of levogucosenone had previously been misinterpreted on first discovery but was later redefined by Halpern et al. [Fung 1972, Wodley 1971]. Despite this misidentification, the theory that lower levoglucosan yields reduced flammability were debunked. However, the yield of levoglucosan was correctly disassociated with flammability . Once the structure of LGO was known, the potential of this enantiomerically pure compound began to be explored, especially with respect to pharmaceuticals, as has been reviewed extensively in the literature [Miftakhov 1994, Comba 2018].

It is only in recent years, however, that a process has been developed which allowed large-scale manufacture. In 2010, Circa Group Pty Ltd filed a patent for their Furacell™ process and developed four pilot plants, working towards producing LGO on a 50 tonne a^{-1} scale [Court 2012].
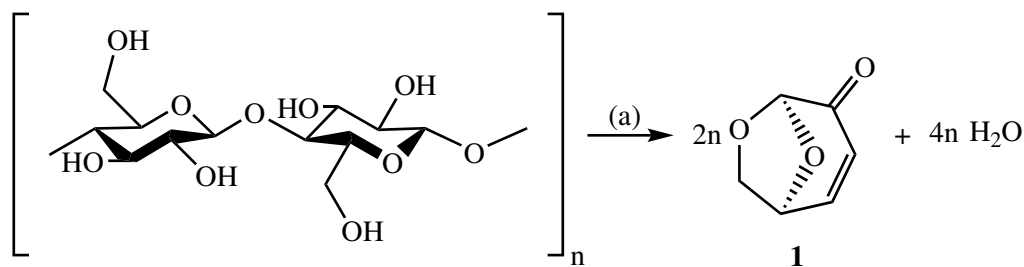
Scheme 1.1: Schematic representation of the Furacell™ process, converting cellulose into levoglucosenone, char and water. (a) H_3PO_4, 430 °C [Court 2012].

The C=C double bond in LGO allowed for hydrogenation to dihydrolevoglucosenone, now produced by Circa Group under the trade name of Cyrene (**2**), as shown in scheme 1.2. This removed its Michael accepting ability and significantly lowered its reactivity, which is instinctively more useful for a solvent.
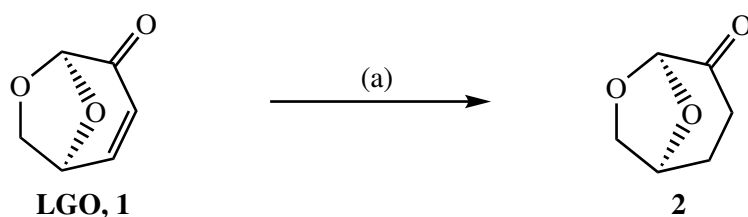
Scheme 1.2: The solvent-free conversion of levoglucosenone (LGO, **1**) to Cyrene (**2**): (a) Pd/C (10%), H_2, 80 bar, RT, 2 h [Sherwood 2014].

The Furacell process is capable of utilising lignocellulosic feedstock, reducing the need for extensive pre-treatment, reducing cost and increasing applicability to genuine waste streams. With the promise of large-scale production of these versatile substances, and the threat of dwindling fossil resources, a new perspective evolved that brought these molecules into the realm of solvents, since which time a whole new body of research has developed. Cyrene itself has since been established as a useful dipolar aprotic solvent for a number of purposes, including graphite exfoliation and graphene dispersion production [Salavagione 2017], as a medium for the Sonagashira reaction [Wilson 2016], Suzuki-Miyaura couplings [Wilson 2018a], for amide coupling using HATU coupling reagent [Wilson 2018b], in metal-organic framework (MOF) synthesis [Zhang 2016] and in the synthesis of ureas [Mistry 2017].

Levoglucosenone derivatives

Due to the increased availability of LGO and Cyrene, there was sudden scope for many more bio-based derivatives, some of which are highlighted in figure 1.9.
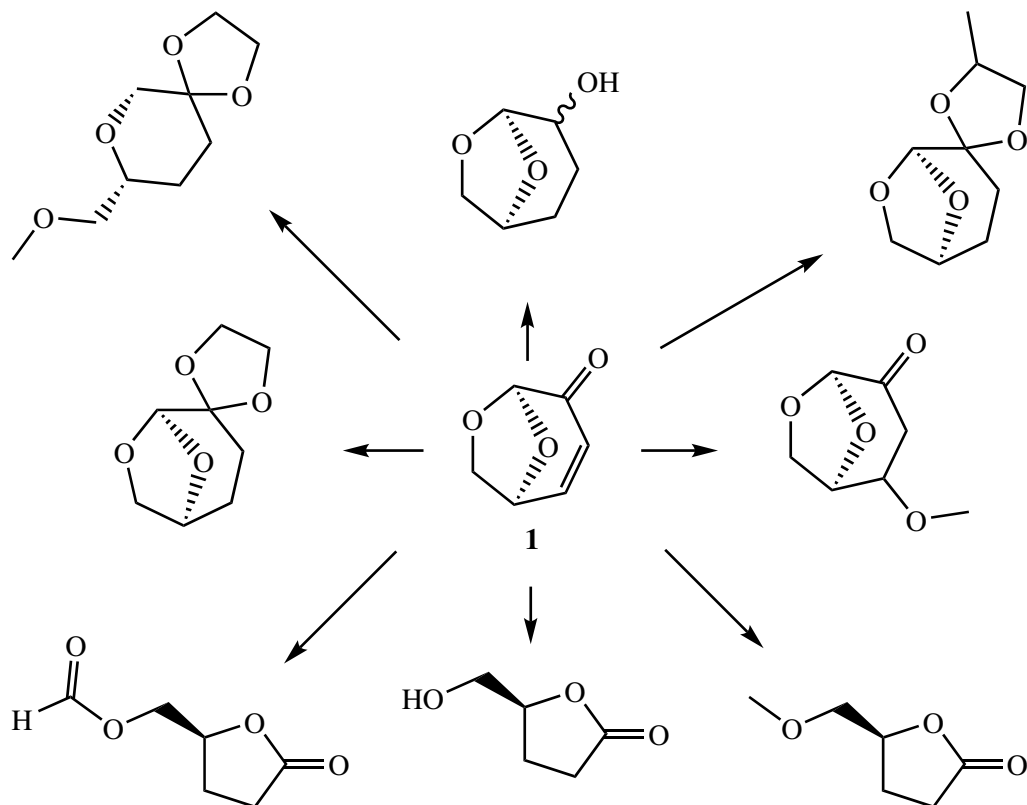
Figure 1.9: A selection of LGO derivatives.

The scope of this platform molecule is much greater than can be elaborated on here, so only relatively simple conversions have been include. A wide range of derivative structures has previously been reported by Pacheco et al., with a focus on those close to NMP in HSP space [Pacheco 2016], and by Camp in a comprehensive review of syntheses [Camp 2018]. Lactones (v)-(vii) are the primary focus of this work. Structure (iv), also known as Cygnet 0.0, was the first modification of Cyrene identified by the Clark research group [Sherwood 2014]. The cyclic ketal serves to protect the carbonyl from forming a geminal diol or dimerising (discussed further in section 3.1.2. However, the effect of this is to raise the melting point of the substance above room temperature, (noted by Pacheco et al. as just under 60 °C), which is not always convenient for a solvent, hence the search is still on for dipolar aprotic solvent replacement candidates [Pacheco 2016].

Chapter 2

p-Cymene as a Medium for Direct Amidation

2.1 Introduction to Amidation

2.1.1 Background

Prevalence in the pharmaceutical industry

Amides are important functional groups in organic chemistry, especially in pharmaceuticals, due to their prevalence in nature. The appearance of amides in a biological context provides a compatibility with the human body, which is already evolved to be receptive towards and metabolise this type of molecule [Rajput 2018]. Additionally, amides are more stable to hydrolysis than esters, another biologically available functional group, and tend to remain active for longer in the body [Snape 2010].

The importance of amide formation to the production of pharmaceutical products was highlighted in two industrial surveys: firstly, by Carey et al. in 2006 [Carey 2006] and again by Roughley et al. in 2011 [Roughley 2011]. Carey et al. focused on 128 drug candidate molecules in the early stages of development at the time, while Roughley et al. used data from publications in the Journal of Medicinal Chemistry, Bioorganic and Medicinal Chemistry, and Medicinal Chemistry Letters from 2008 onward. Each perspective looked at the pharmaceuticals under development at AstraZeneca, GlazoSmithKline and Pfizer and categorised the transformations into reaction classes. Although this type of review is limited by confidentiality and scope

(biopharmaceuticals were not included, nor were chemicals already in production), they provide an indication of the types of transformations most relevant to chemists in pharmaceutical development. This is of pertinence here because the pharmaceutical industry has such a high incidence of solvent use, as alluded to in section 1.1.1. From the given data in these two publications, it is indicated that 12 - 22% of reactions done in the pursuit of drug candidate molecules was an acylation, and 66 - 71% of those were amide bond formations. Roughley et al. denote N-acylation to amide as the most prevalent at 16% of all reactions. The focus on amide functional groups in pharmaceuticals is to be expected because of their prevalence in living organisms, thus they are germane to issues of biological compatibility.

Coupling reagents

In 44% of the amide forming reactions carried out in the pharmaceutical sector reported by Carey et al., the coupling started with an amine and an acyl chloride, resulting in the formation of HCl [Carey 2006]. This has the benefit of being a tried and tested method, but it does mean that the reactions always produce acidic waste, adding to the atom economy, and acyl chlorides are typically corrosive substances, making them more dangerous to use. Additionally, acyl chlorides are generally made by activation of the carboxylic acid using reagents such as thionyl chloride ($SOCl_2$), oxalyl chloride (($COCl)_2$) and phosphorous oxychloride ($POCl_2$), which adds an extra step, and are typically catalysed by DMF [Dunetz 2015]. All of these reagents cause burns and are classified as toxic or fatal if inhaled [Sigma 2018]. Also, $SOCl_2$ may form known carcinogen, dimethylcarbamoyl chloride when used in combination with DMF, and ($COCl)_2$ produces CO_2 and CO gases, which must be safely removed [Dunetz 2015].

Amides can also be formed directly from an amine and a carboxylic acid, however, this can result in the formation of the corresponding ammonium-carboxylate salt instead of the product. For this reason, amidations are commonly executed using one of a number of coupling agents. Besides the acyl chloride method, the two most used coupling reagents, reported by Dunetz et al., were 1,1′-carbonyldiimidazole (CDI) and 1-ethyl-3-(3′-dimethylaminopropyl)-carbodiimide hydrochloride (EDC) based on the number of times reported in the literature at >100 mmol scale. The

structures of each are shown in figure 2.1.

(a) EDC   **(b)** (CDI)
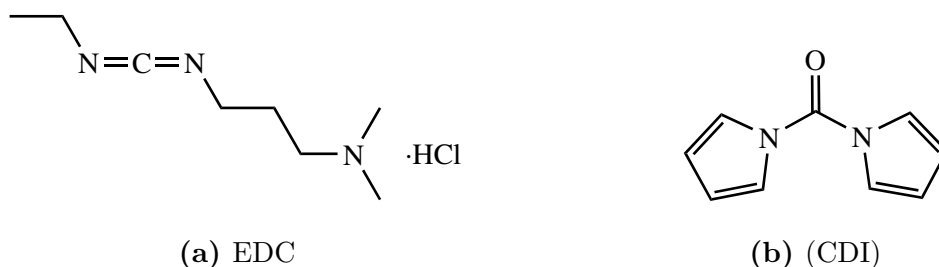
Figure 2.1: Structures of (a) 1-ethyl-3-(3′-dimethylaminopropyl)-carbodiimide hydrochloride (EDC) and (b) 1,1′-carbonyldiimidazole (CDI).

EDC is a carbodiimide used to form the activated ester of a carboxylic acid. Whilst effective, there is a high waste associated with the use of such compounds since a stoichiometric amount of byproduct is produced. It is also quite common to make use of an auxiliary to prevent an unwanted intramolecular rearrangement in which the acyl group transfers onto the nitrogen. One such nucleophile is the frequently used, shock sensitive 1-hydroxybenzotriazole (HOBt). CDI, on the other hand, is safe in comparison and more economical, however, it produces two equivalents of imidazole and one of CO_2 for each use [Dunetz 2015]. The full scope of coupling agents for amide synthesis has been explored thoroughly in the literature [Al-Warhi 2012, El-Faham 2011, Valeur 2009, Montalbetti 2005, Han 2004]. Most of these produce significant waste and call for the betterment of catalytic methods.

2.1.2 Developments

Catalyst for direct amidation

A catalytic method previously shown to have promise was reported by Comerford et al. and made use of a simple activated silica catalyst to produce amides directly from amines and carboxylic acids via a condensation reaction [Comerford 2009]. Since the catalyst was heterogeneous, it was reportedly easy to remove from the reaction by filtration, and reusable with a decrease in activity of about 10%, but which could be regained on reactivation. Conveniently, the thermally activated K60 silica gave the best results and it was noted that supported Lewis acids only served to decrease the activity. A general scheme for the method applied is given in scheme 2.1.
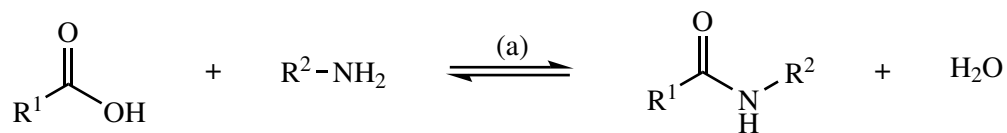
Scheme 2.1: Direct amide bond formation from a carboxylic acid and primary amine. (a) 12 mmol of each reactant, 20 %w/w activated K60 silica, 20 mL toluene, reflux, 24 h.

The silica catalyst also fared well against zeolites, EPZ 'enviro' catalysts, and free Lewis acids $FeCl_3$ and $ZnCl_2$. The study also showed that the changes caused by activation of the silica up to 600 °C were reversible, but above about 650 ° C, the changes became irreversible, a subject that has been explored previously [Vansant 1995]. Physisorbed water can be removed at relatively low temperatures (100-200 °C dependent on porosity) but dehydroxylation occurs only above 200 °C and is not complete until 1200 °C.
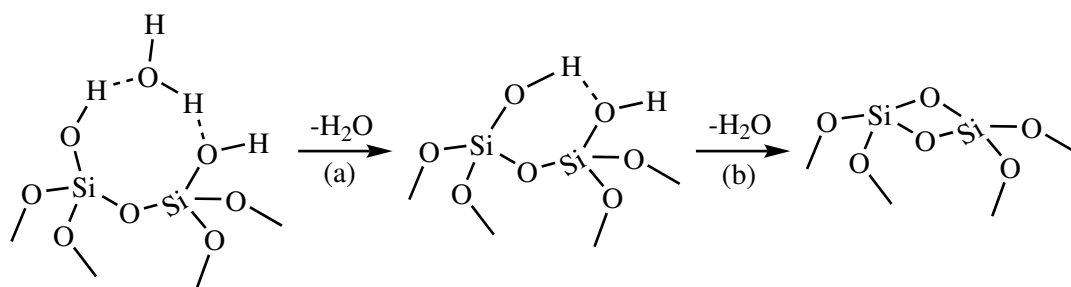
Scheme 2.2: Activation of silica: (a) dehydration, 200-400 °C; (b) dehydroxylation, >400 °C but incomplete until 1200 °C.

The sparse silanol groups present after heating to 700 °C cause a significant retardation of the rehydroxylation process as the bond strength of the siloxane bridge is enforced, and because hydroxylation occurs via pre-existing silanol groups [Vansant 1995]. Comerford et al. proposed that the sparse dehydrated silanol groups were sufficiently acidic to catalyse the amidation, yet not acidic enough to cause protonation of the amine.

It is necessary to acknowledge that there have been investigations into the successful use of arylboronic acids as catalysts for direct amidation [Ishihara, 1996, Arkhipenko 2018] and it has even been reported that catalytic amounts of boric acid can be used to assist the reaction [Tang 2005], however, these require aqueous workup to remove the catalyst and do not follow the reasoning behind the current

method.

Solvents

Solvents typically used for amide formation, given that the majority use some kind
of coupling reagent, are often polar aprotic. Reactions promoted by a carbodiimide
reagent are typically run in DCM and with additional DMF to aid solvation if
necessary [El-faham 2011]. These regularly contain halogens or nitrogen heteroatoms
that are generally associated with unwanted hazards including reproductive toxicity,
which are discussed further in section 3.1.1. However, direct amidation can be
run in low polarity solvents, such as toluene, as demonstrated in the method of
Comerford et al. In combination with the activated silica catalyst, this becomes
quite important, because it is undesirable for the solvent to interact strongly with
the catalyst, lowering kinetic activity by blocking active sites [Comerford 2009].
This, however, is impaired by the known toxic and non-renewable nature of toluene,
previously described in section 1.3.1. An additional benefit was that the reaction
in toluene showed a high tolerance to water and so it was not necessary to use
Dean-Stark apparatus or other water-removal methods [Comerford 2009].

2.1.3   Aims of this Chapter

The intention of this project was to find a greener alternative solvent for use in direct
amidations with the chosen activated silica catalyst. The goal for this particular
procedure was to find which amides would be soluble at high temperatures, allowing
for filtration of the heterogeneous catalyst, but precipitate out of solution at cold
temperatures, allowing recovery of the product without distillation of the solvent.
In doing so, the aim was to find out whether HSPiP could be used to predict this
outcome. In addition, it was hoped to discover the benefits and limitations of relying
on HSPiP in a predictive capacity to discern which reactions would be successful
without resorting to excessive man-hours of practical solubility testing.

2.2   Predictive Modelling

2.2.1   HSPiP

Choice of amides

A range of amides were chosen to test the performance of an alternative solvent against toluene. Structures are shown in figure 2.2.

These amides were chosen to represent a range of solvent space to test the usefulness of HSP predictions in such a study. The amides chosen for testing covered a range of couplings of carboxylic acids to amines, including aromatic-aromatic, aromatic-aliphatic, and aliphatic-aliphatic, with the incorporation of chlorine and nitro groups in some compounds to give variance to the polarity, as well as inductive and mesomeric effects during coupling. Aniline and its derivatives were used several times since this was known to be a difficult coupling that would see appreciable improvements with the chosen activated silica catalyst, thus giving a good comparison between toluene and its alternative [Comerford 2009]. The compounds chosen to synthesise in this study were: *N*-phenyl-4-phenylbutanamide (**3**), 2,*N*-diphenylacetamide (**4**), *N*-phenylbenzamide (**5**), 2-(4-chlorophenoxy)-*N*-phenylacetamide (**6**), *N*-(2-chlorophenyl)-2-phenylacetamide (**7**), *N*-(2,6-dimethylphenyl)-2-phenylacetamide (**8**), phenylacetylpyrrolidine (**9**), *N*-butylbenzamide (**10**), *N*-hexyl-4-nitrobenzamide (**11**), 4-nitro-*N*-phenylbenzamide (**12**), *N*-hexylcyclohexanecarboxamide (**13**), *N*-phenylcyclohexanamide (**14**), and furan-2-carboxylic acid phenylamide (**15**). The structures of each amide is shown in figure 2.2.

Using HSPiP to identify an alternative

In this case, the structure of *p*-cymene was known and could reasonably be expected to behave similarly to toluene due to structural similarities between the two. However, HSPiP helped to visualise the solubility parameters in relation to the amides and other known alternative solvents, which gave assurance that *p*-cymene was the best choice to work with. The SMILES code for each amide was entered into HSPiP, which was used to find a prediction of the HSP. Ordinarily, a single solute would be plotted amongst several solvents and practical solubility testing used to determine the radius of the sphere. Two examples of this are shown in figure 2.3.
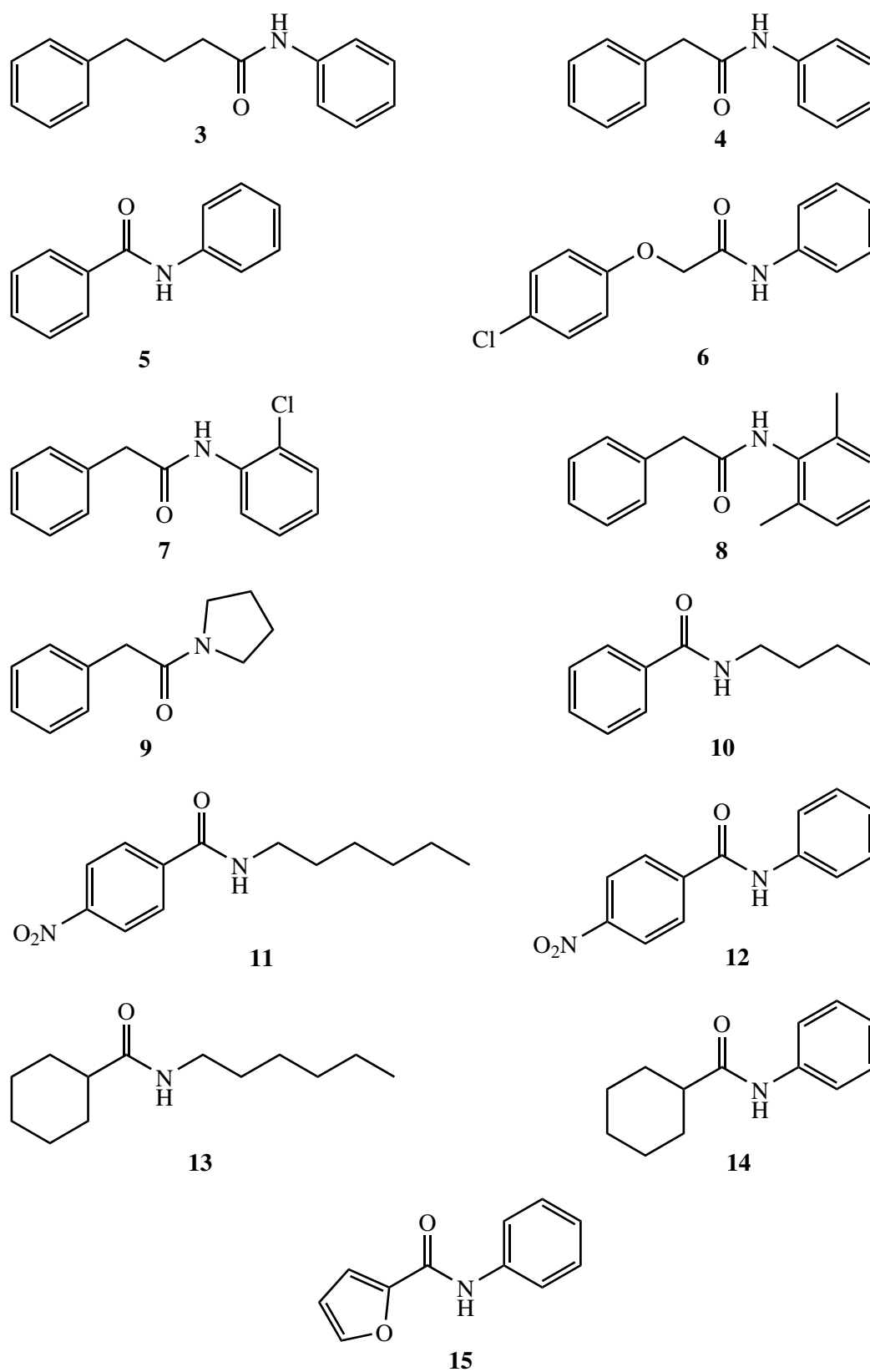
Figure 2.2: Structures of the thirteen amides chosen for this study.

In this case, the radius of the green solvent sphere is arbitrary and chosen to show, graphically, the output of the software. This is because the aim was to use the software predictively, and preparing excess amides for hundreds of solubility tests was outside the scope of this project. Predictive HSP plots for all other amides are given in appendix I.
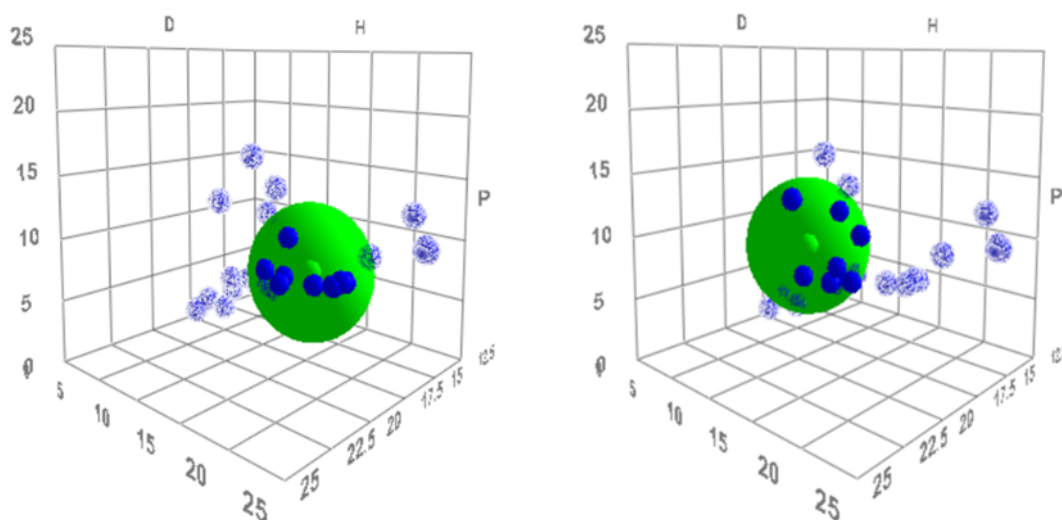
(a) *N*-Phenyl-4-phenylbutanamide (**3**)   **(b)** *N*-Hexylcyclohexanecarboxamide (**13**)

Figure 2.3: Two of the chosen amides (green dots) plotted in HSPiP alongside a selection of solvents designated "recommended" by the CHEM21 selection guide [Prat 2016]. DMF and DMSO have also been included to represent dipolar aprotic space, which green solvents have not really covered.

It is usual for the solvent sphere of the solute to be plotted with respect to several solvents, as the radius of the sphere may be different for each solute. The aim, in this context, was to use the most simple method to predict the likelihood of a good result during the following amidation study (section 2.3). Rather than finding the specific radius for each amide sphere, which would involve dissolving each amide in about 20 different solvents, all of the amides were plotted alongside a selection of green solvents. It is apparent from figure 2.3 that the 3D output from HSPiP is not very clear when printed. Therefore, the results have been plotted 2-dimensionally, as shown in figure 2.4.

The two other perspectives, δH vs. δD and δP vs. δD, are given in appendix II. The third axis, showing dispersion forces, has been omitted since there was the least

variance for this parameter (δD = 15.5-20.6 MPa$^{1/2}$; δP = 1.4-13.7 MPa$^{1/2}$; δH = 2.0-11.3 MPa$^{1/2}$). The scale of each axis can be altered in the software and has been chosen to provide the best spread of data points. It is clear from the graph that the selection of amides covered a range of polarities and H-bonding abilities. Those with greater δP and δH were further in HSP space from the traditional solvent, toluene.
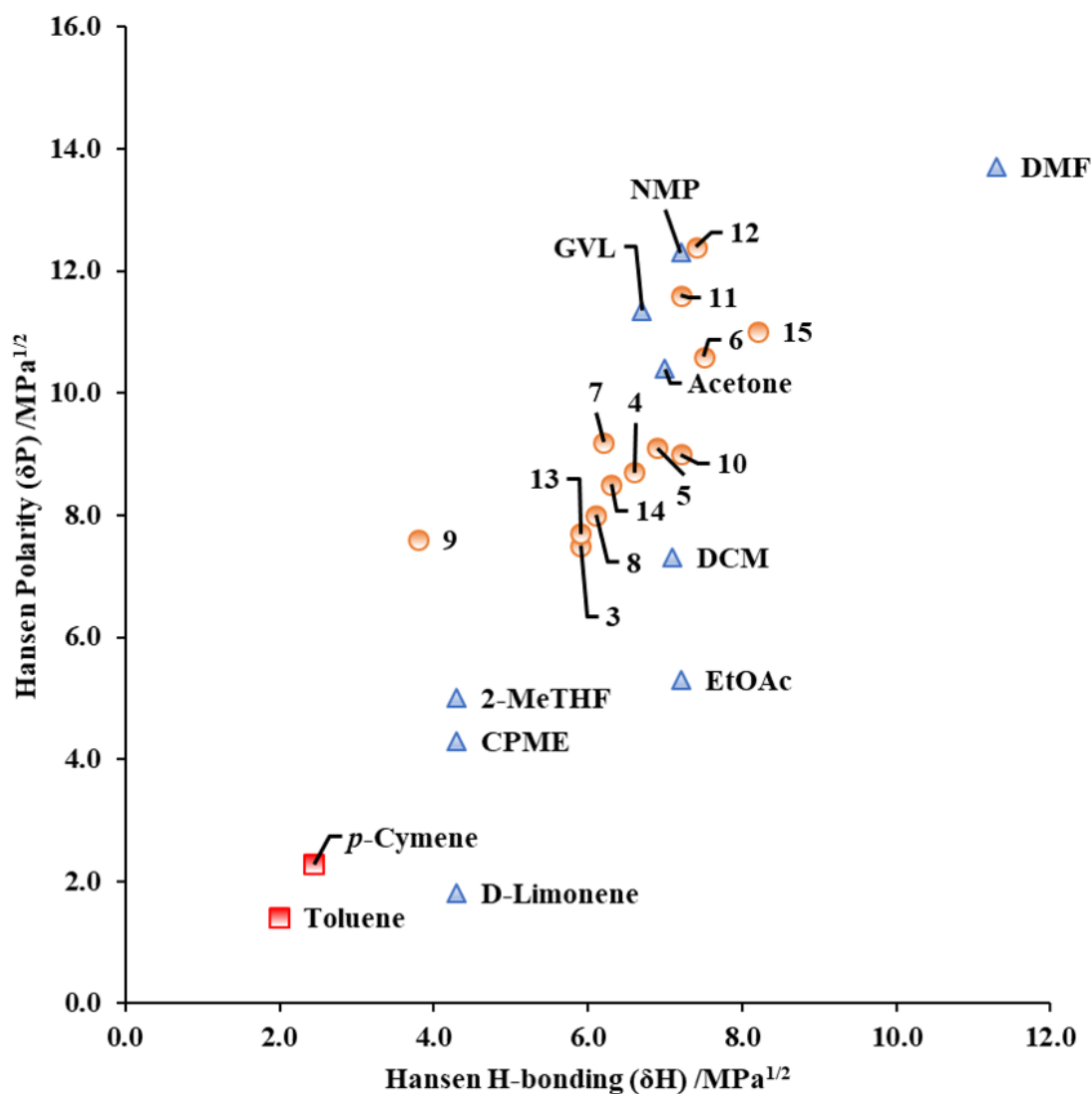
Figure 2.4: HSP predictions for all 13 amides, calculated in HSPiP. Parameter δP against δH. A few green solvents are shown as well as DMF, NMP and DCM for comparison.

This factor was important for the methodology of the amidations because two solvency observations were required: firstly, that the amide would not be soluble at room temperature and could be recovered by filtration, and secondly, that the amide would be soluble at high temperatures so that the silica catalyst could be

filtered off prior to precipitation of the product.

From the graph, *p*-cymene and toluene sit quite close in solvent space, with *p*-cymene being a little closer to the amides on the δP and δH axes. They are similarly close in both other 2D perspectives of solvent space, although toluene is closer to most of the amides on the δD axis. It was, therefore, not immediately clear which would have greater solvency of the amides, but validated the expectation that both solvents would behave similarly. To clarify the modelling output of HSPiP, equation 1.2 (section 1.2.1) was used to determine a numerical value for the distance between each of the amides and both solvents. The results of this calculation are shown in table 2.1 along with the difference in each parameter.

Table 2.1: Distance in HSP space between each amide and solvents, toluene and *p*-cymene.

Amide	Distance from toluene /MPa$^{1/2}$				Distance from *p*-cymene /MPa$^{1/2}$			
	δD	δP	δH	Total	δD	δP	δH	Total
3	1.0	6.1	3.9	7.5	1.6	5.2	3.5	7.0
4	1.6	7.3	4.6	9.2	2.2	6.4	4.2	8.9
5	1.9	7.7	4.9	9.9	2.5	6.8	4.5	9.6
6	2.2	9.2	5.5	11.6	2.8	8.3	5.1	11.3
7	2.0	7.8	4.2	9.7	2.6	6.9	3.8	9.5
8	1.4	6.6	4.1	8.3	2.0	5.7	3.7	7.9
9	1.0	6.2	1.8	6.8	1.6	5.3	1.4	6.4
10	0.3	7.6	5.2	9.2	0.9	6.7	4.8	8.5
11	0.7	10.2	5.2	11.5	1.3	9.3	4.8	10.8
12	2.6	11.0	5.4	13.3	3.2	10.1	5.0	13.0
13	0.8	6.3	3.9	7.6	0.2	5.4	3.5	6.5
14	0.8	7.1	4.3	8.5	1.4	6.2	3.9	7.9
15	1.9	9.6	6.2	12.0	2.5	8.7	5.8	11.6

The difference between individual parameters was calculated by simply subtracting δX$_{solvent}$ from δX$_{amide}$ and finding the modulus, which indicates the contribution and variance of each parameter to the overall distance. The total distance of each amide from toluene was greater than the same amides from *p*-cymene, which

suggested that they would have greater solubility in p-cymene than toluene. Additionally, there are noticeable trends in structural effects, for instance, amides **9** and **13** are nearest, both having aliphatic components contributing to a lower polarity. Amides **12** and **15** are furthest, each having rigid aromatic groups on either side of the amide functional group.The actual solubilities of each amide were found experimentally and are discussed further in section 2.4.

2.2.2 Yalkowsky Approximation

The Yalkowsky approximation was used to predict the ideal solubility of each amide. As described in section 1.2.3, due to generally weaker intermolecular interactions in substances of lower melting points, it stands to reason that less energy is generally required to break them, which in this case, translates to higher ideal solubilities. For the purposes of this study, equation 1.3 was used to estimate the ideal mole fractional solubility (X^{ideal}) of each amide, as rearranged below.

$$X^{ideal} = 10^{-0.01(mp-25)} \tag{2.1}$$

The calculation in equation 2.1 was applied to all 13 amides and the results have been visualised in figure 2.5. The ideal solubility was predicted with respect to a system at room temperature, which has been defined as 25 °C.

Since most of the amides are highly crystalline, many were predicted to have limited solubility at room temperature. This was considered the ideal result for the proposed amidation method since, as previously explained, it depended upon the precipitation of the amides on cooling to room temperature. By contrast, amides **9**, **10**, **13** and, to some extent, **11** were more likely to remain in solution, rendering filtration impossible.

2.3 Amidation Study

2.3.1 Methodology

The effectiveness of p-cymene as a solvent for direct amidation was investigated in comparison to petroleum-based solvent toluene. Thirteen amides were chosen
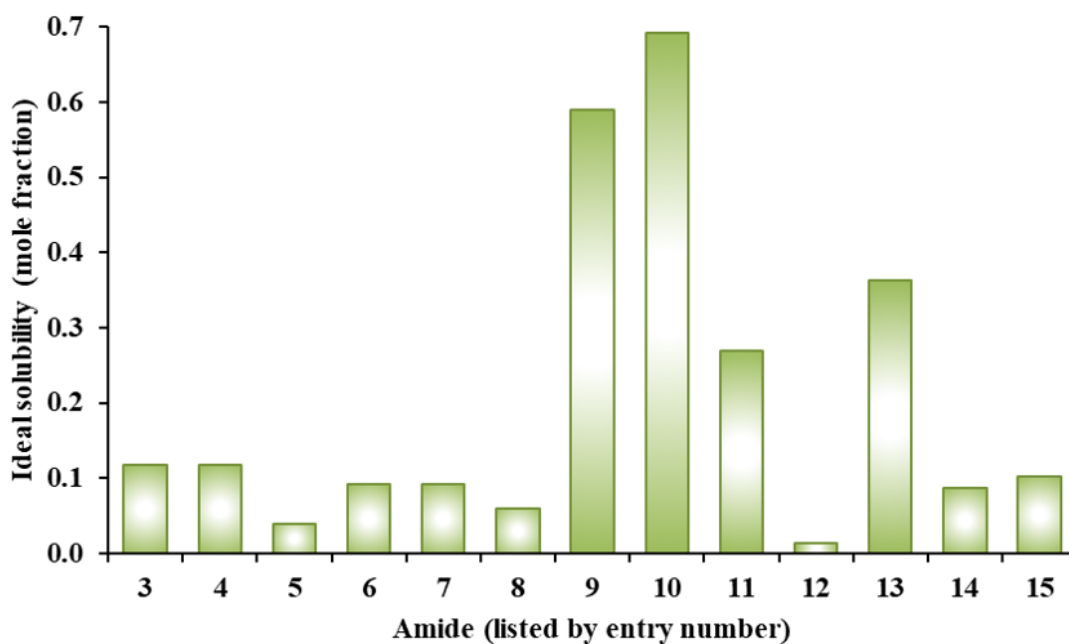
Figure 2.5: Yalkowsky approximation applied to each amide.

for this purpose, which covered a range of polarities, functional groups and aromatic/aliphatic molecules as described above (section 2.2.1). The thirteen chosen amides were synthesised in both refluxing toluene and *p*-cymene, as well as *p*-cymene at 111 °C to be comparable with the temperature conditions of refluxing toluene. K60 silica activated at 700 °C was used as a catalyst for reasons explained in section 2.1.2. Amide yields were obtained after 24 h for each reaction by GC analysis of the reaction mixture with reference to an internal standard of hexadecane or tetradecane, and also by isolation of the amide, to highlight recovery issues caused by varying solubility in the solvent. Wherever possible, recovery was done by filtration of the catalyst at high temperature, followed by filtration of precipitated amide at room temperature, and if not, this has been incorporated into the discussion.

2.3.2   Results

Amide yields in both solvents

The yields of all reactions are given in table 2.2, 2.3 and 2.3. "Blank" yields indicate that no catalyst was used. Percentage yields both from product isolation and determined by gas chromatography are shown to indicate which products were readily

recovered by filtration.

The yields given in tables 2.2, 2.3 and 2.4 (results of all amide formations) show a common trend with respect to the effect of the solvent, which is that p-cymene did not generally perform as well as toluene when the reactions were run at 111 °C. There are several possible reasons for this and the cause has not been confirmed. Speculatively, since the p-cymene was not refluxing, there was less turbulence in the solution, which could imply that there was less mixing. Another possible cause is that the toluene reactions were more concentrated since the toluene was boiling and left a greater proportion in the vapour phase. Additionally, although considerable care was taken to stop the reactions losing volume over the 24 h period (see section 2.3.3), vapourised toluene was more likely to escape the system than p-cymene and the condensers were unlikely to be 100% efficient.

It is clear that the higher temperature available to p-cymene as a solvent had a significant effect on driving some of the more difficult reactions forward within the 24 h time frame. In some cases, the temperature was a much more significant factor than the catalyst. To make the effect of temperature and catalyst on the yield clearer, the results have been plotted as a bar chart in figure 2.6.
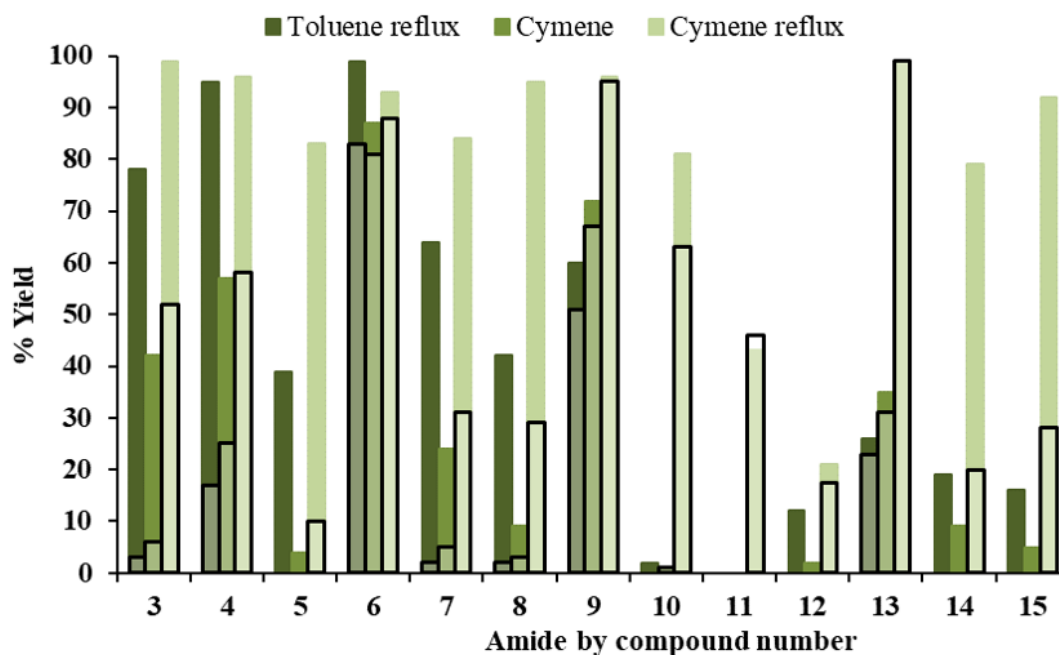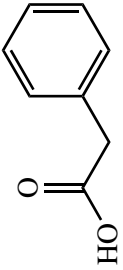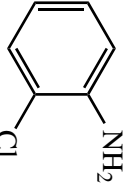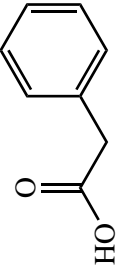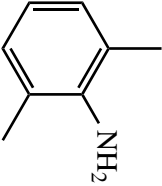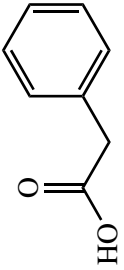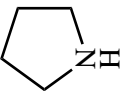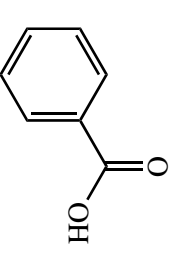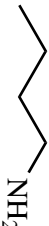
Figure 2.6: Catalysed vs. catalyst-free reactions. Black outlines indicate catalyst-free reaction yields. Yields obtained by GC analysis.

Table 2.2: Average yields resulting from each amide formation - Part 1

Entry #	Acid	Amine	Solvent	GC % Conversion[a]	Blank % Conversion[b]	Isolated % Yield
3			Toluene	78	3	61
			p-Cymene*[c]	42	6	37
			p-Cymene	>99	52	92
4			Toluene	95	17	80
			p-Cymene*	57	25	57
			p-Cymene	96	58	88
5			Toluene	39	0	57
			p-Cymene*	4	0	10
			p-Cymene	83	10	69
6			Toluene	>99	83	74
			p-Cymene*	87	81	75
			p-Cymene	93	88	77

[a]Calculated by calibration of relative GC peak intensity with internal standard.

[b]Reaction run without catalyst. Calculated from GC.

[c]p-Cymene* indicates reaction was run at 111 °C in accordance with refluxing toluene.

Table 2.3: Average yields resulting from each amide formation - Part 2

Entry #	Acid	Amine	Solvent	GC % Conversion[a]	Blank % Conversion[b]	Isolated % Yield
7	phenylacetic acid	o-chloroaniline	Toluene	64	2	18
			p-Cymene*	24	5	6
			p-Cymene	84	31	55
8	phenylacetic acid	2,6-dimethylaniline	Toluene	42	2	20
			p-Cymene*	9	3	2
			p-Cymene	95	29	69
9	phenylacetic acid	pyrrolidine	Toluene	60	51	52
			p-Cymene*	72	67	28
			p-Cymene	96	95	22
10	benzoic acid	butylamine	Toluene	2	<1	-[c]
			p-Cymene*	1	1	-
			p-Cymene	81	63	55
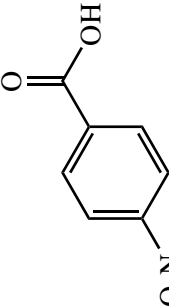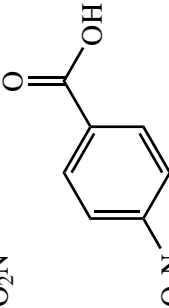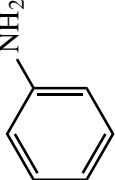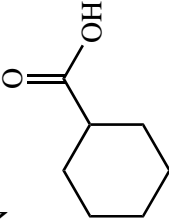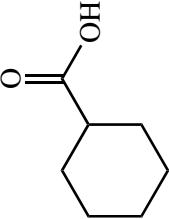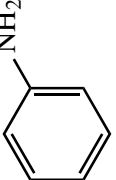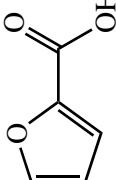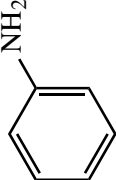
[a]Calculated by calibration of relative GC peak intensity with internal standard.
[b]Reaction run without catalyst. Calculated from GC.
[c]Yield too low to isolate amide.

Table 2.4: Average yields resulting from each amide formation - Part 3

Entry #	Acid	Amine	Solvent	GC % Conversion[a]	Blank % Conversion[b]	Isolated % Yield
11			Toluene	<1	<1	-
			p-Cymene*	<1	<1	-
			p-Cymene	43	46	34
12			Toluene	12	<1	8
			p-Cymene*	2	<1	<1
			p-Cymene	21	17	11
13			Toluene	26	23	9
			p-Cymene*	35	31	16
			p-Cymene	>99	>99	>99
14			Toluene	19	<1	11
			p-Cymene*	9	<1	5
			p-Cymene	79	20	70
15			Toluene	16	0	-
			p-Cymene*	5	0	-
			p-Cymene	92	28	78

[a]Calculated by calibration of relative GC peak intensity with internal standard.
[b]Reaction run without catalyst. Calculated from GC.

The activated silica catalyst did not make a significant improvement in the yields of amides **6** and **9**, which were fairly successful to begin with in the blank reactions (88% and 95% respectively when run in refluxinf *p*-cymene). Similarly, the catalyst had little effect on the formation of **10**, **11** and **13**. With the exception of **6**, these amides were not formed from couplings of aniline or aniline derivatives, which were known to benefit from the presence of activated silica [Comerford 2009]. It is likely that the electron-withdrawing nature of the acid structure left the carbonyl more susceptible to attack and so **6** formed very easily.

The yields of nitro-containing amides **11** and **12** were particularly low. It was observed during testing that 4-nitrobenzoic acid was poorly soluble in both solvents and so mixing was not ideal. This was helped somewhat by the higher temperature of refluxing *p*-cymene, but there was significant precipitation above the line of the solvent, where the glass became cooler. This effect could be reduced by changing the experimental set up but, for consistency, all reactions were run the same way. Another difficulty with these reactions, and in particular amide **12**, was that the product was so insoluble, even at higher temperatures, that it was prone to precipitate prior to the work-up. Specifically, excess precipitate had to be melted back into solution with a heat gun in order to retrieve a GC sample. Further to this, the product was prone to precipitation during filtration even though the funnel was pre-heated, which made it difficult to isolate.

This very low solubility seems to mark the limits of this methodology at the high polarity end of the scale. Amide **12** was the furthest from either solvent in HSP space, a prediction of zero solubility using the Yalkowsky calculation, and showed no sign of dissolving during solubility testing. Since this method relied upon the amide dissolving at high temperatures and precipitating at room temperature, this compound fell outside the bounds of our target. At the opposite end of the scale, amide **9**, **10** and **13** were completely soluble in both solvents and had to be recovered by distillation to remove the solvent followed by column chromatography to remove the internal standard. Losses incurred during these procedures meant that a significant drop in yield was observed between that obtained by GC analysis and the isolated product, particularly in the case of **9**, which achieved a yield of 96%, but only 22% was recovered. Due to the high boiling point of the solvents, especially

p-cymene at 177 °C, distillation is not a desirable method of isolation, and not the aim of this methodology, so these compounds mark the lower limits in this dataset. A positive aspect of this is that an upper and lower limit for this modelling strategy when applied to silica catalysed direct coupling of carboxylic acids and amines in bio-derivable *p*-cymene was approximated.

2.3.3 Consistency

A large number of amide formations were run in the early stages of this study, which have not been mentioned in the previous results. This is due to a number of inconsistencies leading to a drawn out process of preparatory work to ensure reliable results were produced. The problems encountered and solutions found are summarised in this section.

Internal standard

Initially, reactions were run with no internal standard. It was expected that the amide yield could be ascertained by calibrating against the carboxylic acid starting material, and working out the conversion. An internal standard was initially avoided with the understanding that extraneous inputs could have an effect on the solubility, which was a factor in the amide recovery. However, it became clear that relying on the GC response to acid would not be effective. For example, early reactions were run with smaller acids, such as propionic acid. With a boiling point of only 141 °C, there was a possibility that the acid could escape the system, particularly in *p*-cymene at reflux, skewing the results. This could also lead to partial obscurance of the GC signal by the solvent, due to low acid retention times. A further issue of note was that the acid peaks often suffered from tailing in the chromatogram, due to the non-polar column used, leading to unreliable peak areas. If this problem was addressed with a polar column, this would have resulted in a similar situation for the amide. Ultimately, the decision was made to use an internal standard throughout the reaction. Either tetradecane or hexadecane were used, according to visibility on the chromatogram. This did not appear to cause any problems for the study, but it is necessary to understand that the presence of these compounds is likely to have slightly decreased the overall polarity of the solvent medium.

Temperature agreement and solvent loss

On some occasions, there was a notable loss of solvent in the toluene reactions, which also resulted in an increase in yield, due to the higher concentration. Additionally, a temperature difference was measured between heating points on the apparatus. It was necessary to replace a few pieces of equipment before the full selection of amides could be reliably tested. To ensure that the multipoint hotplate/condenser did not have a temperature bias between flask positions, p-cymene was heated to 111 °C in all 6 positions and the temperature tested for consistency. With a full set of undamaged fittings, there was no more than 1-2 °C difference between positions (possibly due to the draft extraction in the fume hood). Toluene was refluxed overnight to check that solvent was not being lost.

With these problems being fixed, the yields of various reactions (isolated or by GC if isolation was difficult) were run twice to ensure that results were comparable. A selection of the repeats is shown in table 2.5.

Table 2.5: A selection or repeat amide reactions compared for result consistency.

Amide	Solvent	Temp. °C	% Yield 1	% Yield 2
3[a]	Toluene	111 °C	62.3	58.9
	p-Cymene	111 °C	35.6	37.7
	p-Cymene	177 °C	91.1	93.6
4[a]	Toluene	111 °C	79.8	81.6
	p-Cymene	111 °C	57.5	55.6
	p-Cymene	177 °C	87.7	87.5
8[a]	Toluene	111 °C	19.4	19.6
9[b]	p-Cymene	111 °C	70.6	73.5
10[b]	Toluene	111 °C	2.1	2.3
	p-Cymene	111 °C	1.4	1.2

[a]Isolated yields.
[b]Yields obtained by GC analysis.

For context, amide **4** had achieved isolated yields of 38%, 75% and 53% in three preceding reactions run in toluene. Similar disparaties were noticed in p-cymene, and for amide **3**. Once the yields were within 4% of one another, the error was

deemed reasonable for the purposes of this study.

2.4   Solubility Tests

2.4.1   Concepts and Development

Optimising the method

The solubility testing method was developed using amide **3**, as the yield had been high and so a reasonable quantity of pure amide was available. Initial tests were run on a scale of 20 mL solvent, using 0.5 g amide. This was subsequently scaled down due to the lower quantities of other amides available. During the first attempt, the solutions were raised to 60 °C in order to fully dissolve the amide. The intention was to ensure that the solutions were super-saturated on cooling to 25 °C, whereupon they were allowed to equilibrate for 23 h, causing excess amide to precipitate out of solution. A typical setup is shown in figure 2.7, picture (A).
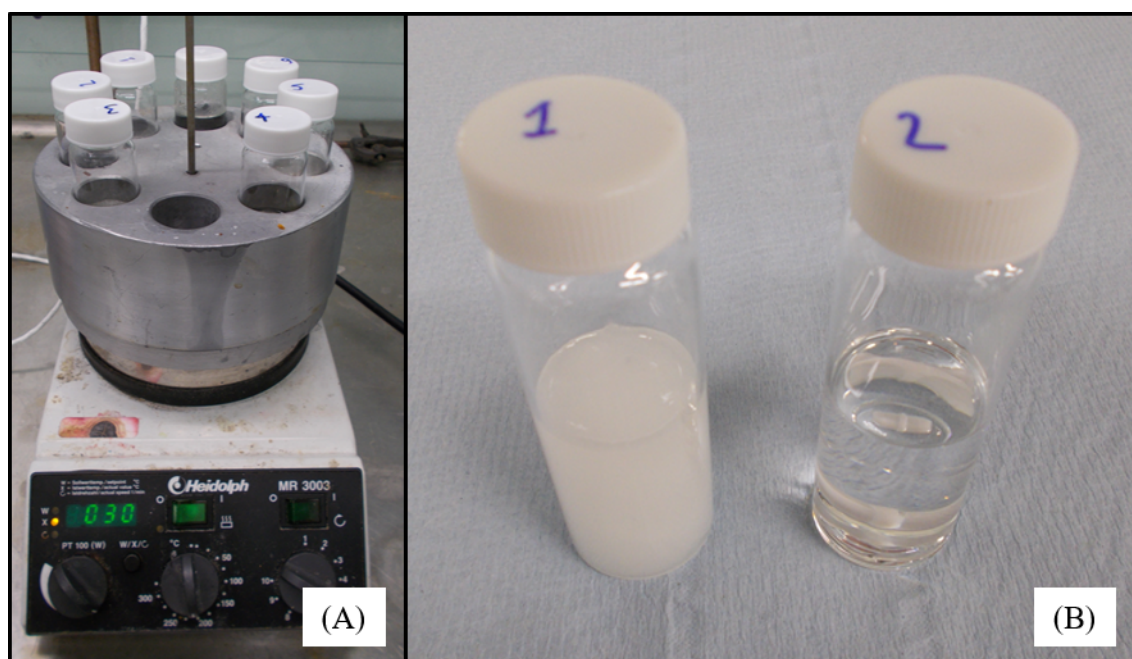
Figure 2.7: (A) Setup of solubility tests; (B) extensive precipitation of amide **3** on cooling in *p*-cymene (1) and none in toluene (2).

Unfortunately, as the amide precipitated, it also expanded to fill the volume of the solvent, as shown in the second image (B - left). In order to find the concentration of amide in solution, it was necessary to be able to take a sample of the clear

liquid, which was not possible. On the right, the image shows that the amide had not precipitated out of toluene upon cooling, perhaps requiring seeding, hence the method had to be modified.

Ultimately, the solubility testing was run over two days at a constant temperature of 25 °C, the first day with stirring and the second without to allow any undissolved compound to settle. It was hoped that this would be sufficient time to allow a saturated solution to be reached. As the solubility of a solute in a given solvent increases with temperature, care was taken to ensure the temperature remained constant throughout the test. At the time of sampling, a small aliquot of the clear liquid was taken using a pre-heated pipette so that precipitation would not occur on the glass. This was immediately weighed and mixed with a prepared internal standard solution of either hexadecane or tetradecane. As this method relied on the accuracy of weight measurements, the results have been provided in moles of amide per kilogram of solution.

Repeatability

As solubility testing was reliant upon the production of pure amide from the amidation study itself (section 2.3), and a portion of amide was also required for GC calibration, repeat tests could not be done for all 13 amides. To test the repeatability of the solubility tests, amide **3** was tested three times in each solvent to find the mean concentration of a saturated solution, to be quantified by GC analysis. The individual results are provided in table 2.6.

Table 2.6: Solubility of amide **3** at 25 °C: repeat tests. Quantities determined by GC.

Solvent	Solubility of **3** /mol kg^{-1}			Mean	SD[a]	RSD[b]
	Repeat 1	Repeat 2	Repeat 3	/mol kg^{-1}	/mol kg^{-1}	
Toluene	0.074	0.078	0.078	0.077	0.0022	2.86%
p-Cymene	0.017	0.017	0.017	0.017	0.00027	1.61%

[a]Standard deviation
[b]Relative Standard Deviation

It was evident that the spread of data points was not far from the mean giving reasonably low relative standard deviations (<3%) so this method was considered

reliable enough to continue use for all 13 amides.

2.4.2   Results and Discussion

Solubility of amides in toluene and *p*-cymene

Each of amides **3** - **15** were tested in both toluene and *p*-cymene to determine the solubilities in each solvent. A comparison of the results is shown in figure 2.8.
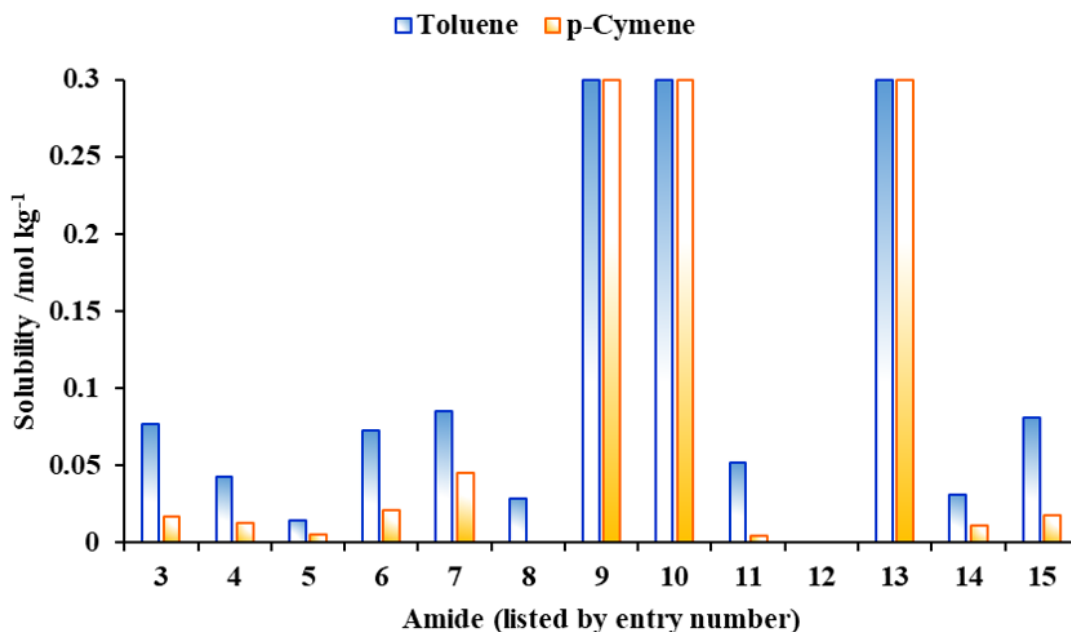
Figure 2.8: Solubility of each amide in toluene and *p*-cymene at 25 °C. Amides shown reaching 0.3 mol kg^{-1} were fully soluble and a saturated solution could not be reached with the available quantities.

The full numerical data for these solubilities have been included in appendix IV. As presented in the chart, the two solvents, *p*-cymene and toluene, behaved similarly in terms of their relative ability to solubilise each amide. In general, the more polar amides did not dissolve well, whereas the more aliphatic amides, **9**, **10** and **13**, were completely soluble in both solvents. One surprising result was that the amides exhibited greater solubility in toluene than in *p*-cymene, despite the greater distance in HSP space. Figure 2.9 compares the distance in HSP space with experimental solubility of the amides to visualise the trends observed.

For the most part, a greater distance in solvent space corresponded with a lower observed solubility, which was expected. It is not clear why *p*-cymene was a worse
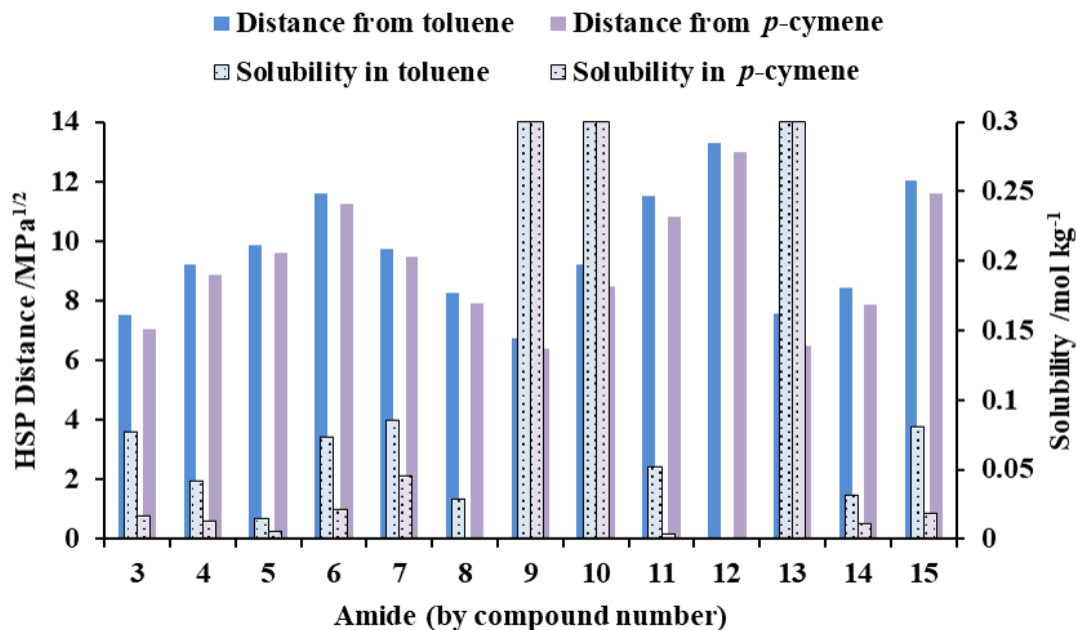
Figure 2.9: A comparison of the solubilities of each amide with the HSP distances calculated in table 2.1.

solvent for the amides despite having closer HSP. This may be because the HSP values of *p*-cymene are not yet as refined as they will become in HSPiP, with further data aggregation, since the group contribution method is heavily reliant on statistical datasets. Amides **9**, **10** and **13** appeared to be excessively soluble, but this was likely effected by their low melting points, which are better accounted for by the Yalkowsky approach, as demonstrated in figure 2.10.

It is possible that kinetics, which is not accounted for by HSPiP, played a part in amides **3** and **8** dissolving poorly in spite of their shorter distances, although a reasonable effort was made to allow the solutions to equilibrate. Kinetics may also be responsible for the unexpectedly lower solubility of all amides in *p*-cymene, as its kinematic viscosity is higher at 0.85 mm^2 s^{-1} compared to 0.65 mm^2 s^{-1} for toluene [Parr 2018].

Referring to its examination in figure 2.5, the Yalkowsky approximation seems to have been a good fit for this dataset. There is a clear correlation between the approximation and the solubility data presented in figure 2.8. For example, amides **5**, **8** and **12** were all expected to be virtually insoluble, which proved to be true experimentally. The outliers, **9**, **10** and **13**, all dissolved easily in both solvents, which
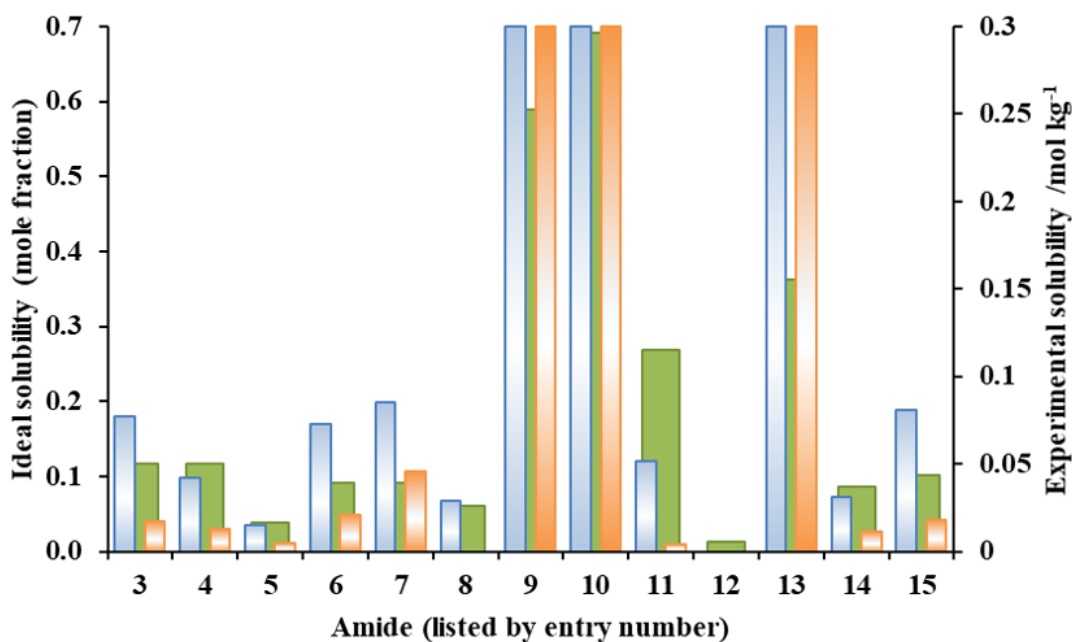
Figure 2.10: A comparison of the Yalkowsky solubility predictions with amide solubilities determined experimentally. Green = Yalkowsky prediction; blue = solubility in toluene; orange = solubility in *p*-cymene. A solubility of 0.3 mol kg^{-1} indicates that the amide was too soluble to gain a true solubility limit.

also correlates with the Yalkowsky approximation. The only significant diversion from the prediction was amide **11**, which was barely soluble despite its higher solubility prediction than most of the other amides. It appears that the polarity of this compound was significant enough to interfere with its dissolution in toluene and *p*-cymene, for which the melting point-based Yalkowsky approximation did not account. In this case, a melting point-based approach was a reasonably successful predictor for solubility. However, since we know that not all low melting substances are miscible, it can reasonably be assumed that this method should not be solely relied upon.

2.4.3 Identification of Rotamer

An interesting observation was made with regard to the characterisation of amide **8** (*N*-(2,6-Dimethylphenyl)-2-phenylacetamide). At first glance, the proton NMR spectrum appeared to show contaminants in the crystalline amide as shown in figure 2.11.
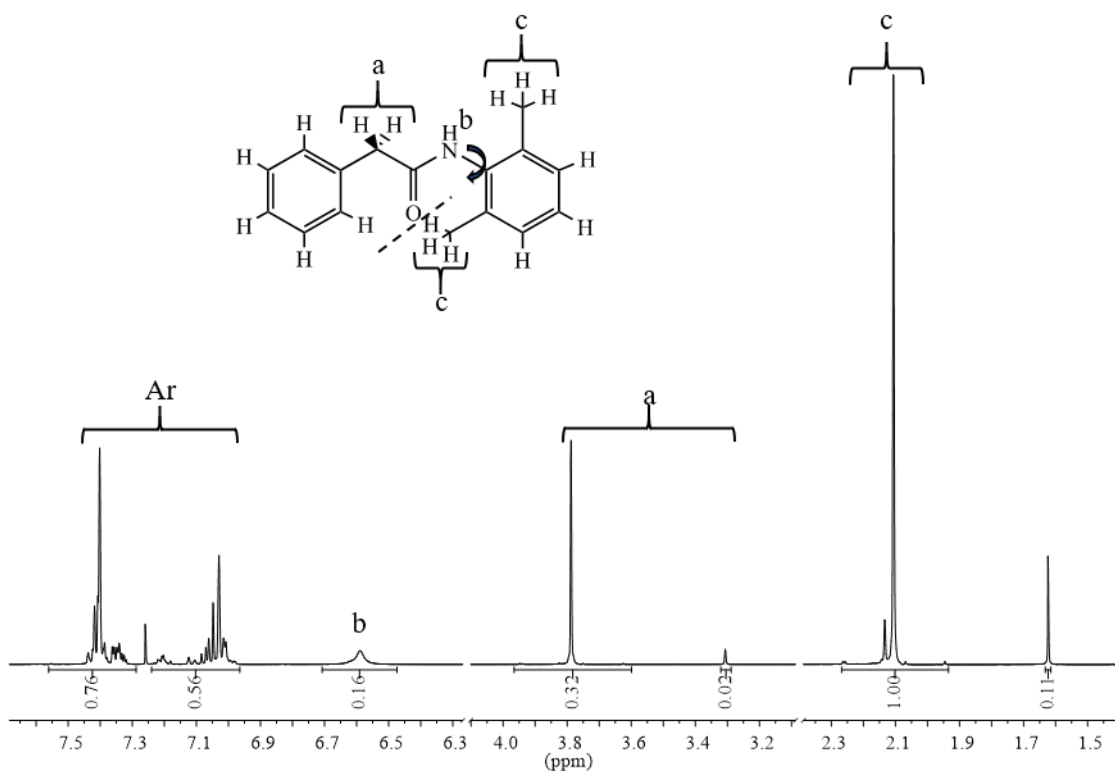
Figure 2.11: Proton NMR (400 MHz) of *N*-(2,6-dimethylphenyl)-2-phenylacetamide (**8**) run in CDCl$_3$. Water signal at δ 0.11 ppm. Full spectrum provided in appendix X.

As can be seen in the spectrum, there are unexpected singlets at δ 3.3 and 2.1 ppm as well as unexpectedly complex multiplets in the aryl region between δ 7.0-7.5 ppm. Multiple steps were taken to try to elucidate the identity of this compound, including TLC in the following solvents:

- 80:20 petroleum ether:ethyl acetate

- 50:50 pet ether:ethyl acetate

- DCM with 2 drops ethyl acetate

- 95:5 DCM:ethyl acetate

No second spot was observed and neither by chromatography column run in a solvent system of 95:5 DCM:ethyl acetate. Re-runs of the NMR spectra after separation techniques returned the same observations. The GC chromatogram also returned clean (see appendix VI).

The speculation is that these extra signals may be caused by a rotamer of the amide, as identified in figure 2.11. It is proposed that the close proximity of the

methyl groups to the carbonyl oxygen, caused steric hindrance and induced-dipole interactions, represented by the dotted line, to the extent that the rotation of the C-N bond was slowed down enough to observe two distinct species. This was evidenced by the fact that the sum of the integration, of what were defined as analogous protons, was the same as expected for a single species of this molecule. Carbon NMR run of the same sample gave an equivalent number of extra signals (see appendix XI). The signals did not match up with the starting material amide or acid. Variable temperature proton NMR was run at 25, 30 and 55 °C to determine if the duplicate signals could be aligned by overcoming the thermodynamic barrier to rotation. However, the results were inconclusive as there was negligible change in chemical shift.

Other methods were used to probe this theory, including filtering the NMR sample through Amberlyst 15 hydrogen type, and using $CDCl_3$ that had been pre-dried with $MgSO_4$ and neutralised with basic alumina, for the elimination of water from the samples. Neither of these methods cause a shift in the signals. A further test was carried out by running proton NMR spectroscopy with a drop of conc. HCl. This did not remove the secondary peaks, but the analagous signals were caused to shift closer together. However, a proton NMR spectrum of the same sample run in DMSO saw the disappearance of the secondary peaks.

2.5   Flow System

Following successful application of p-cymene as a solvent for direct amide formation, it was hypothesised that the process could be made greener by the use of an alternative technological setup. In light of previous work done by Comerford et al. using a flow reactor to form amides in toluene [Comerford 2012], p-cymene was tested a solvent in a flow system, using the reaction of 4-phenylbutyric acid and aniline to give amide **3** as a model. An image of the reactor used is shown in figure 2.12.

The reaction was set up such that the reaction mixture could circulate through the system. The reactants were initially mixed with p-cymene and put into a reservoir, which could be connected up to the system via two connecting tubes. The catalyst chamber was filled with activated K60 silica catalyst (3.5 g) and heated
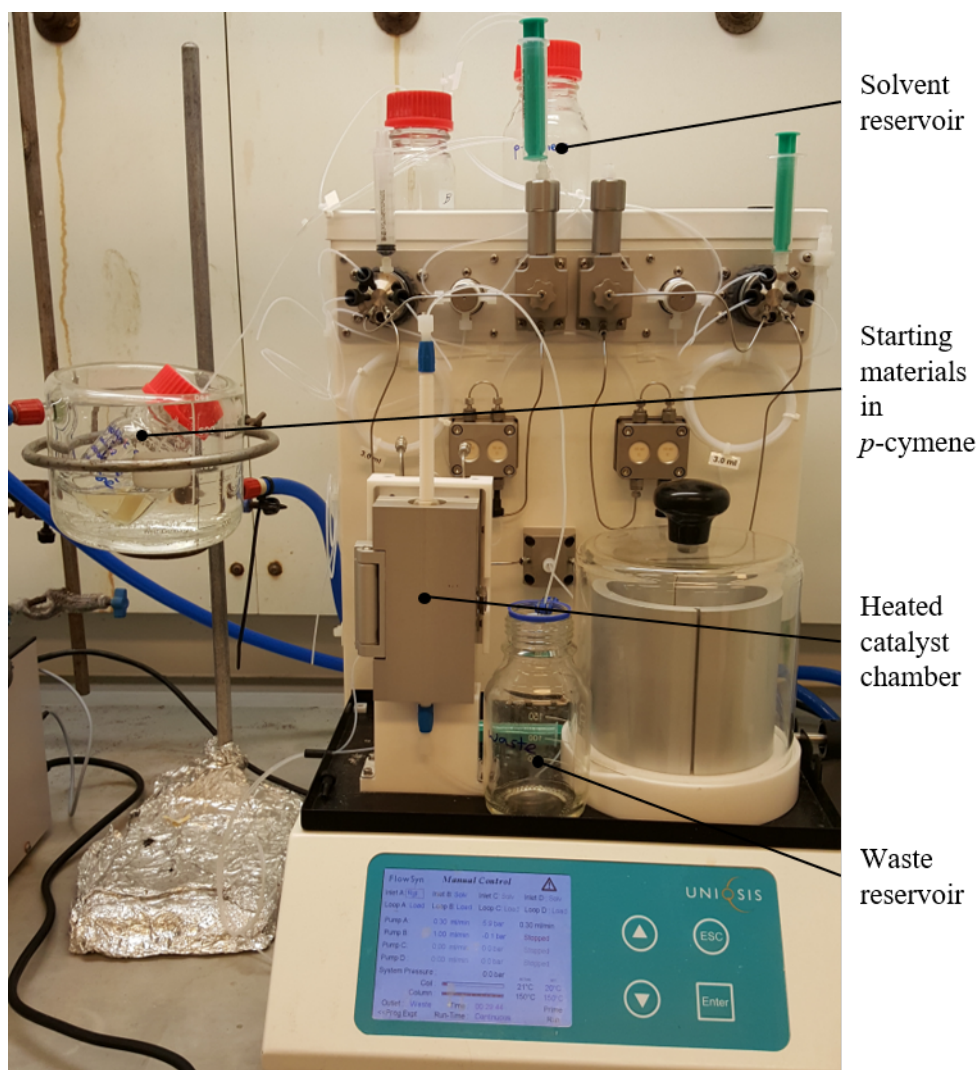
Figure 2.12: Flow system used for the production of amide **3**.

to 150 °C. *p*-Cymene was flushed through the system from the solvent reservoir, then the input line was switched, via the display screen, to come from the reactant vessel. The solvent exiting the system was initially captured in the waste reservoir, but after a few mL had been collected, the exit tube was connected up to the reactant reservoir so that the mixture could recirculate. It was expected that the amide would precipitate in the reservoir, which was cooled by circulating water, thereby driving the equilibrium of the reaction forward. The results of flow reactions are shown in table 2.7.

Table 2.7: Results of reaction to form amide **3** in a flow reactor, using *p*-cymene as a solvent.

Recirculation	Residence time[a] /min	Total time /min	Space time yield[b] /mg g^{-1} h^{-1}
1	23.3	67	60
2	46.6	134	48
3	69.9	201	39

[a]Calculated from volume of catalyst (3 g) and flow rate (0.3 mL min^{-1}).
[b]The incremental decrease in output is expected as the starting materials were not replenished.

The results show that, as expected, the space time yield decreased with further circulations of the reactants, since the reactant concentration was not replenished. Unfortunately, the reaction did not produce a high yield, which was unexpected as this amide worked reasonably well in batch testing, achieving >99% GC yield in refluxing *p*-cymene and 42% in *p*-cymene at 111 °C. As the flow reaction was run at 150 °C, due to the constraints of the machinery, it was expected that the temperature would be sufficient to promote the coupling. However, the maximum residence time in this study was just under 70 minutes, so it is not directly comparable with the batch reactions.

Operation

There were a number of operational difficulties that hindered a successful outcome, which are summarised here. For the flow system to work, it was reliant upon the reactants being soluble in the solvent at lower temperatures than the amide, which appeared to be the case. However, it was also hoped that the amide would precipitate

out of solution in the reservoir. This was not obviously the case, possibly due to the low concentrations of amide that had been produced in the time period. Unfortunately, the reactor would not continue to run for four or more cycles. It was initially tested overnight in order to achieve several circulations, but the automatic cutoff was triggered when the pressure rose above 7 bar. On occasions when the system stopped and cooled down, the silica catalyst became extremely dense and tough to remove. It is speculated that this was caused by the precipitation of amide in the silica chamber, however, it is not known whether that occurred before or after the build-up of pressure. As the amide is known to be fully soluble at 150 °C in p-cymene, there is no obvious reason for that to happen, and a longer study with an engineering perspective would be needed to solve the problem. Nonetheless, it has been demonstrated that, in principle, a flow system could be applied to certain amide reactions in p-cymene.

2.6 Observations on Solvent Stability

2.6.1 p-Cymene Stability Tests

While testing amidation reactions conducted in p-cymene, it was observed that the solvent became an orange/brown colour. For a solvent to be useful as a reaction medium, its stability towards the reaction conditions must be established, so a test was required to determine the cause of the discolouration. The discolouration was more profound during the high temperature reactions. It was speculated that a radical oxidation akin to the cumene process could be responsible for this [Hock 1944], due to the ability of the arene ring to mesomerically stabilise a radical on the pendant isopropyl group. A possible mechanism for this is given in scheme 2.3.

In order to test the influencing factors in the reaction, p-cymene was heated for 24 h and stirred with activated K60 silica and phenylacetic acid respectively, each at two different temperatures, 111 °C and 177 °C, in line with the amidation reactions carried out in this study. The extent of discolouration is shown in figure 2.13.

It is clear from the colour gradient in the solutions that the carboxylic acid, in this case phenylacetic acid (right), has a much greater impact on the oxidation of p-cymene than the activated K60 silica catalyst (left). This is likely because activating
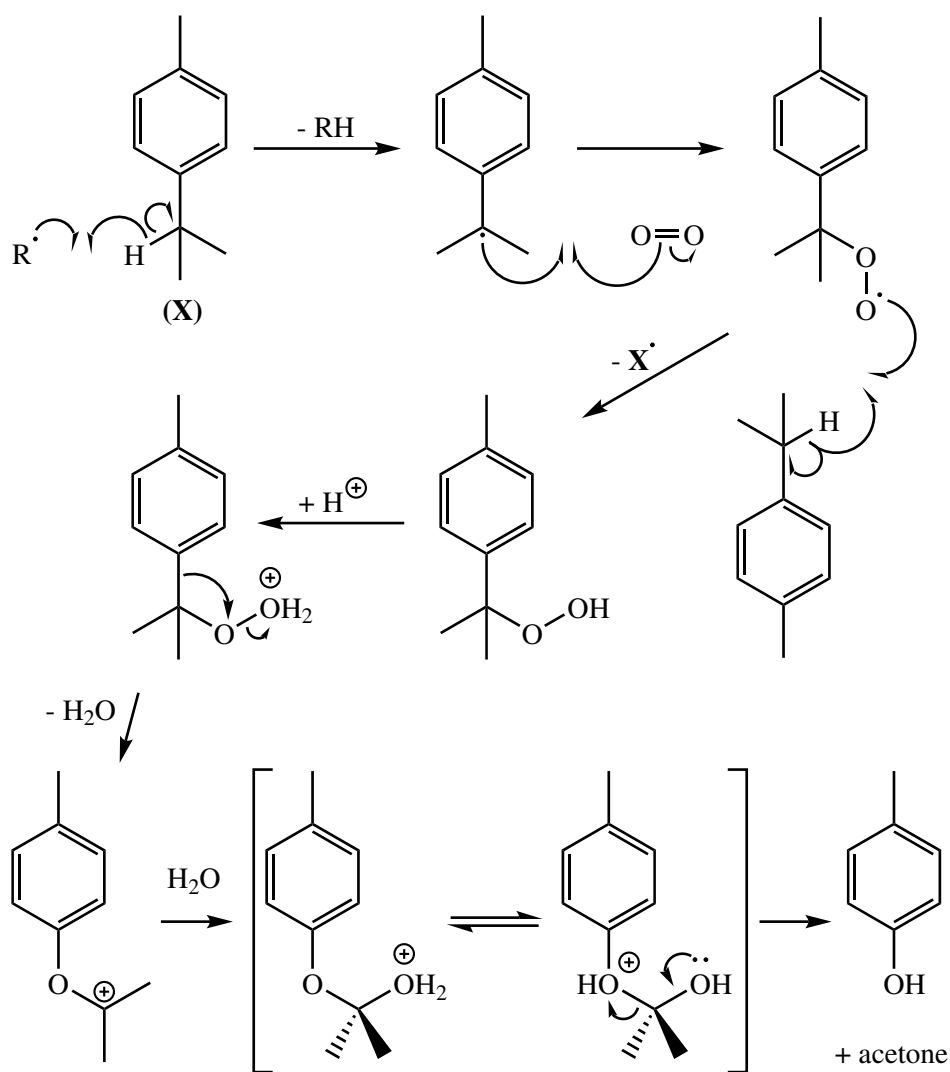
Scheme 2.3: A speculative mechanism for the oxidation of *p*-cymene analagous to the cumene process.

silica at 700 °C removes most of the silanol units, which could otherwise behave similarly to the carboxylic acid. Raising the temperature to 177 °C from 111 °C also intensified the orange colour signifying, predictably, that a higher temperature and increased contact with O_2 increased the extent of oxidation. This coincides with the cumene process, which relies on the presence of a proton donor. Although phenol itself is uncoloured, various oxidised forms could contribute to the colouring.

As the system was under an air atmosphere, it was assumed that exposure to oxygen was instrumental in the oxidation of the *p*-cymene. To assess this theory, a repeat of the test heating phenylacetic acid in *p*-cymene to reflux was run under an inert atmosphere of argon, as shown on the right in 2.13. The *p*-cymene was degassed
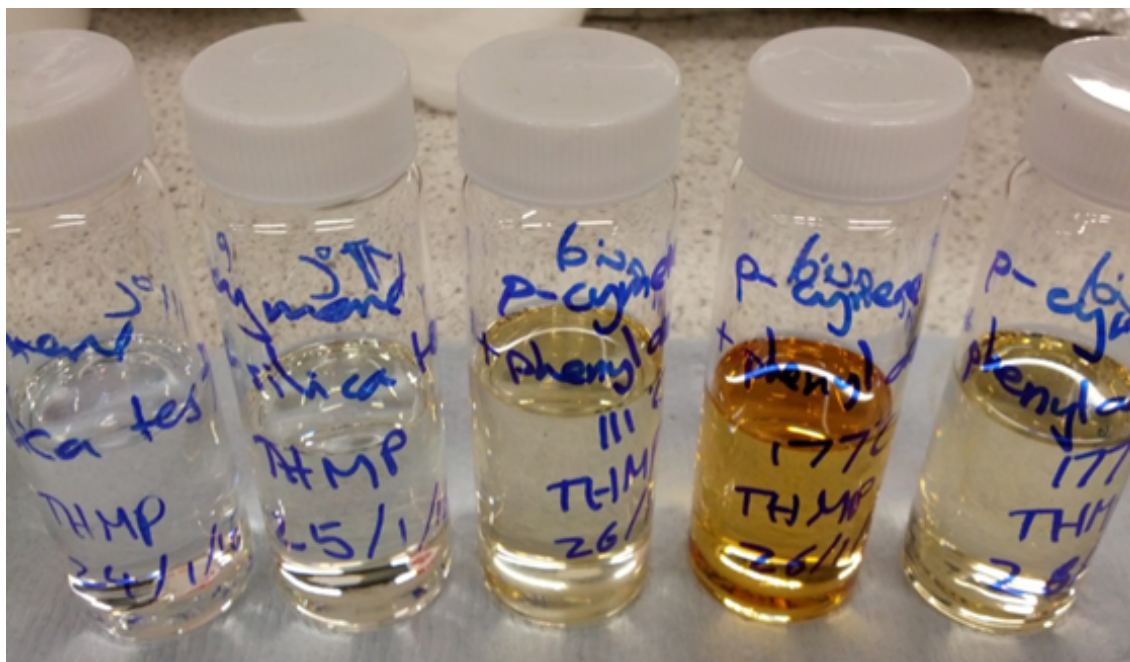
Figure 2.13: Colour of mixtures after heating for 24 h. *p*-Cymene stirred with (from L to R): activated silica at 111 °C, activated silica at 177 °C, phenylacetic acid at 111 °C, phenylacetic acid at 177 °C, phenylacetic acid at 177 °C under an argon atmosphere.

with argon overnight prior to the test to remove residual oxygen. The discolouration was reduced greatly, although not completely, possibly due to an imperfectly sealed system. In this case, the test was purely qualitative, but it could prove useful in the future to generate more detailed characterisation by NMR to identify the contamination. It was not considered necessary to conduct all experiments under argon as the small quantities of contamination products were not crucial to the assessment of the solvent validity and undetectable in the analysis. However, it is common practice in industry to use an inert atmosphere for reasons of safety and quality, so this was not anticipated to be a significant issue.

2.6.2 Superiority of *p*-Cymene over Limonene Precursor

As *p*-cymene has been justified for use in the previous experiments as a potential downstream derivative of limonene, it would be reasonable to question why the conversion needs to take place at all, rather than using limonene itself as a solvent.

In order to ascertain whether D-limonene could be used as a solvent as well as *p*-cymene, an amidation using 4-phenylbutyric acid and aniline was run to produce

amide **3** using D-limonene as the solvent. This amide was chosen because it had a mid-range yield in *p*-cymene at 111 °C of 42% (from GC analysis). This allowed for the reaction in limonene to establish a yield comparable with a *p*-cymene solvent, rather than choosing a reaction that worked extremely well or extremely poorly.

The reaction mixture became very dark during the course of the reaction, which was to be expected as limonene is susceptible to attack from radicals. No crystals precipitated out of solution after cooling to 4 °C and the GC trace was comparatively messy.
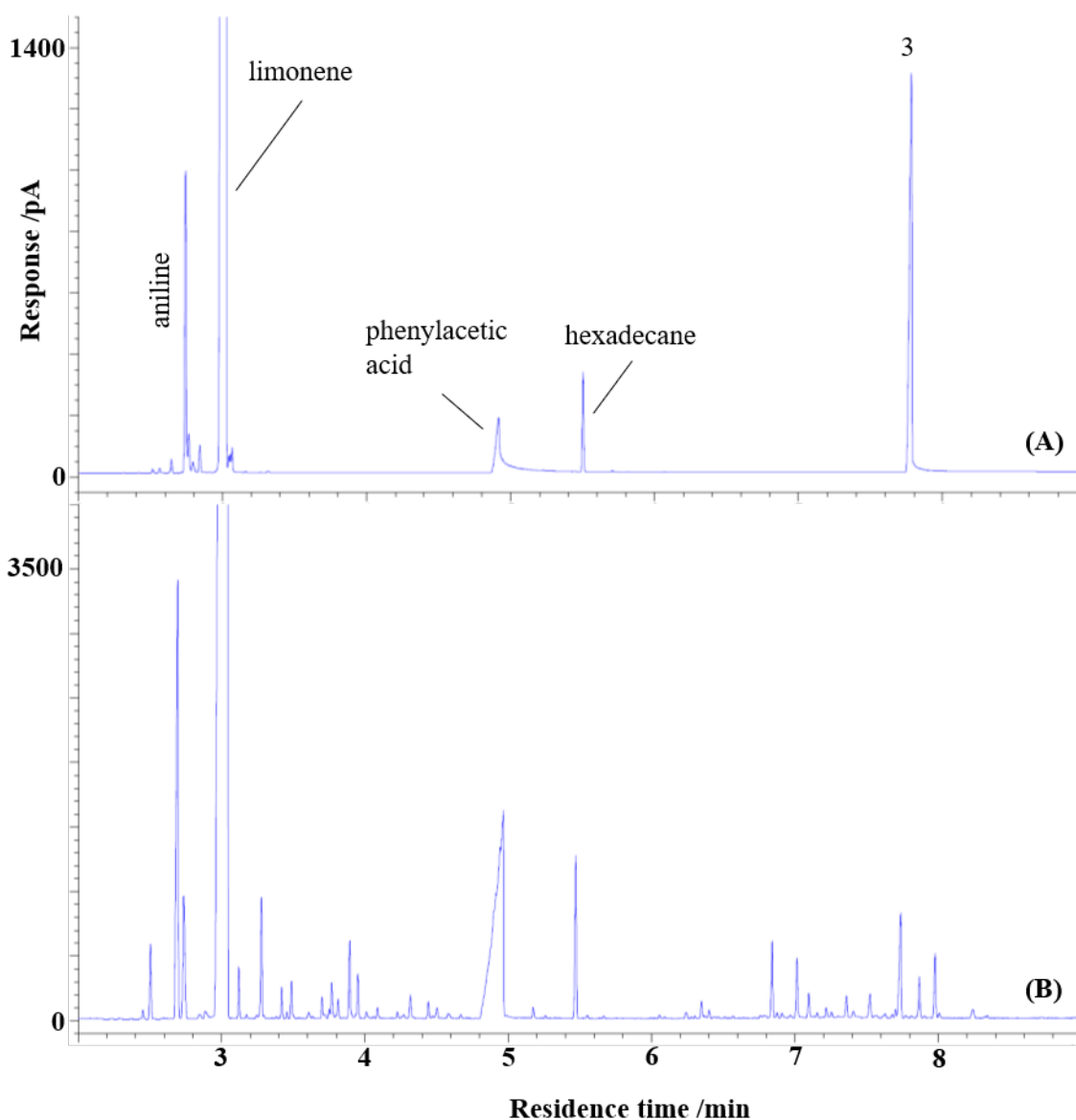
Figure 2.14: GC traces of reaction mixtures from: (A) formation of *N*-Phenyl-4-phenylbutanamide (**3**) in *p*-cymene at 111 °C; (B) formation of **3** in D-limonene at 111 °C.

From the GC in figure 2.14, a yield of <6% was determined, which does not fare well compared to the 42% achieved in *p*-cymene at the same temperature. Furthermore, considering the number of breakdown products apparent at this temperature, it is likely that even more would appear should the reaction be done at reflux, so, even though limonene has the same boiling point as *p*-cymene, it can be ruled out as an alternative solvent for direct amidation. The additional stability of *p*-cymene, therefore, justifies the extra step of converting limonene. However, this does not rule out the use of limonene as a solvent for non-reactive purposes, such as cleaning agents [Paggiola 2016], and the deposition of electronic materials [Lamarche 2017].

2.7   Conclusions

In this study, it has been shown that *p*-cymene is a suitable substitute for toluene as a solvent for direct amide bond formation, in particular, by making use of the higher reflux temperature available to *p*-cymene (177 °C). *p*-Cymene worked well in combination with the activated silica catalyst, by dissolving most of the amides at high temperatures, allowing for catalyst recovery, and then allowing the amides to precipitate for filtration, with the exception of more aliphatic amides.

The amides chosen were all marginally less soluble in *p*-cymene than in toluene despite predictions regarding distance in HSP space by HSPiP suggesting the opposite. The Yalkowsky approximation was used to predict the solubilities of the amides in an ideal solvent and this proved to be a reasonable approximation of real-world observations.

Flow reactions were demonstrative of the possibility of improving the greenness of the reaction in the future, although the methodology requires optimisation as yields were low.

2.8   Future Work

It would be useful to expand the dataset for this research considering the usefulness and variety of amides in industry. It would, therefore, be prudent to include a range of heteroaromatic starting materials, particularly nitrogen containing heterocycles, as well as establishing the relevance to the pharmaceutical industry by attempting

the total synthesis of a known drug compound, which was outside the scope of this work.

Considering the success using p-cymene as a substitute for toluene in amidations, there is a precedent for finding further alternative, safer solvents. It is not the aim of this work to promote p-cymene as the only possible option. Byrne et al. recently published research detailing a catalytic route to 2,2,5,5-tetramethyltetrahydrofuran (TMTHF) by the ring-closure of 2,5-dimethylhexane-2,5-diol [Byrne 2017]. The Hansen solubility parameters of TMTHF are reportedly close to those of toluene and it has exhibited comparable behaviour in model Fischer esterifications and amidation studies. It follows that a more comprehensive study should take place. It would be complementary to complete a study on the use of p-cymene in esterification reactions alongside this, utilising HSPiP as a solubility tool to compare is effectiveness with this study.

An area that requires significant development is the optimisation of the flow reaction. Theoretically, this type of equipment can greatly improve the efficiency of a process because the solvent is recirculated, which is a point of particular importance with regard to sustainability, especially where limonene derived p-cymene is concerned [Paggiola 2016]. Additionally, research into the improvement of the conversion of limonene into p-cymene it necessary to make it a viable green solvent candidate.

Chapter 3

Development of 5-Membered Ring Lactones as Cellulose-Derived Solvents

3.1 Introduction

The recent industrial development in production of levoglucosenone and Cyrene from waste lignocellulosic material, as described in 1.3.2, has spurred a plethora of studies on the potential derivatives of these molecules in the world of green chemistry [Pacheco 2016, Camp 2018]. Among the applications tested for these compounds include replacements for solvents with known toxic effects, Cyrene itself having already been identified as a potential candidate for the substitution of dipolar aprotics [Sherwood 2014].

3.1.1 Replacement of Dipolar Aprotic Solvents

Toxicity concerns

In 2015, Ashcroft et al. published a set of data regarding the use of solvents reported in publications from *Organic Process Research & Development* between 1997-2012 [Ashcroft 2015]. Only papers describing chemistry on a significant scale (>100 g) were considered, with the majority of operations being on the 1-100 kg scale. In a section focusing on dipolar aprotic solvents, by number of papers citing use of each

solvent, the authors showed that the five most prevalent polar aprotic solvents were acetonitrile (MeCN), dimethylformamide (DMF), dimethyl sulfoxide (DMSO), *N*-methyl-2-pyrrolidone (NMP), and dimethylacetamide (DMAc) respectively at the time of publication. Of these, DMF, DMAc and NMP are of particular concern due to reprotoxicity issues and have both been named on the REACh substances of very high concern list (SVHC) [ECHA 2018a]. DMSO does not have the same toxicity concerns but poses a serious thermal decomposition risk [Lam 2006], which warrants finding alternatives. Although NMP was cited only a fraction as often as MeCN and DMF, it was the solvent which displayed the most significant growth in use, doubling between the periods of 2005-2008 to 2009-2012. This can be explained by the fact that the health issues associated with DMF were discovered first and NMP was previously regarded as a suitable substitute. It is important to bear in mind the limitations of this review as it could not cover chemical processes that are protected as trade secrets and the statistics were based on only one journal, but it does show the general areas of use for dipolar aprotic solvents.

Table 3.1 shows the hazards associated with these five solvents as defined by the European Chemicals Agency (ECHA).

The prevalence of teratogenic effects is the biggest concern leading to the search for alternatives to DMF, NMP and DMAc in particular. On top of this, the presence of heteroatoms in each of these solvents leads to the production of NOx and SOx on incineration, which must be filtered from the emissions or they will lead to further atmospheric damage and negative health effects [Le Cloirec 2012]. It is also worth noting that acetonitrile, although lesser in hazard than the previous three solvents, has been classified as toxic and a possible carcinogen by companies registering to REACH [ECHA 2018b], and all classifications are subject to increased severity given time and the acquisition of further data.

Applications of dipolar aprotic solvents

In the aforementioned publication, it was shown that nearly 50% of the reported uses of DMF, DMAc, NMP and DMSO concern their application in nucleophilic substitution reactions [Ashcroft 2015]. A breakdown of these uses is shown in figure 3.1.

Table 3.1: Known toxicity information for the top five most commonly used polar aprotic solvents. Each solvent is manufactured or imported in the European Economic Area on a scale of 10 000 - 100 000 t a^{-1}. Information gathered from European Chemicals Agency (ECHA) website [ECHA 2018b].

Solvent	Hazard statements
MeCN	Highly flammable liquid and vapour, is harmful if swallowed, is harmful in contact with skin, causes serious eye irritation and is harmful if inhaled.
DMF[a]	May damage the unborn child, is harmful in contact with skin and is harmful if inhaled.
DMSO	No hazards have been classified.
NMP	May damage the unborn child, causes serious eye irritation, causes skin irritation and may cause respiratory irritation.
DMAc[a]	May damage the unborn child, is harmful in contact with skin and is harmful if inhaled.

[a]Additionally, DMF and DMAc have been linked to toxic liver disease after occupational exposure [Malag. 2012].

This observation can be explained by the fact that dipolar aprotic solvents promote fast reaction times for nucleophilic subsitution reactions, especially where small, negatively charged nucleophiles are involved [Parker 1961]. This is because the hydrogen bonding ability of protic solvents allows the nucleophile, particularly smaller, more basic nucleophiles, to be shielded. If the nucleophile interacts strongly with a solvent, it follows that it will be slower to undergo a substitution. In addition, nucleophilic aromatic substitutions can be accelerated by up to a factor of 10^5 when compared to systems in a protic solvent [Miller 1961]. Therefore, it is important to find alternative solvents which perform to an equal or higher standard in order for them to be adopted in industry.

Applications of NMP falling under the "other" heading, in other words not related to the manufacture of a chemical, include use as an extraction solvent in the petrochemicals industry, solvent/co-solvent for insecticides and pesticides in the agriculture industry, solvent for polyurethane paints, acrylic and epoxy resins, and
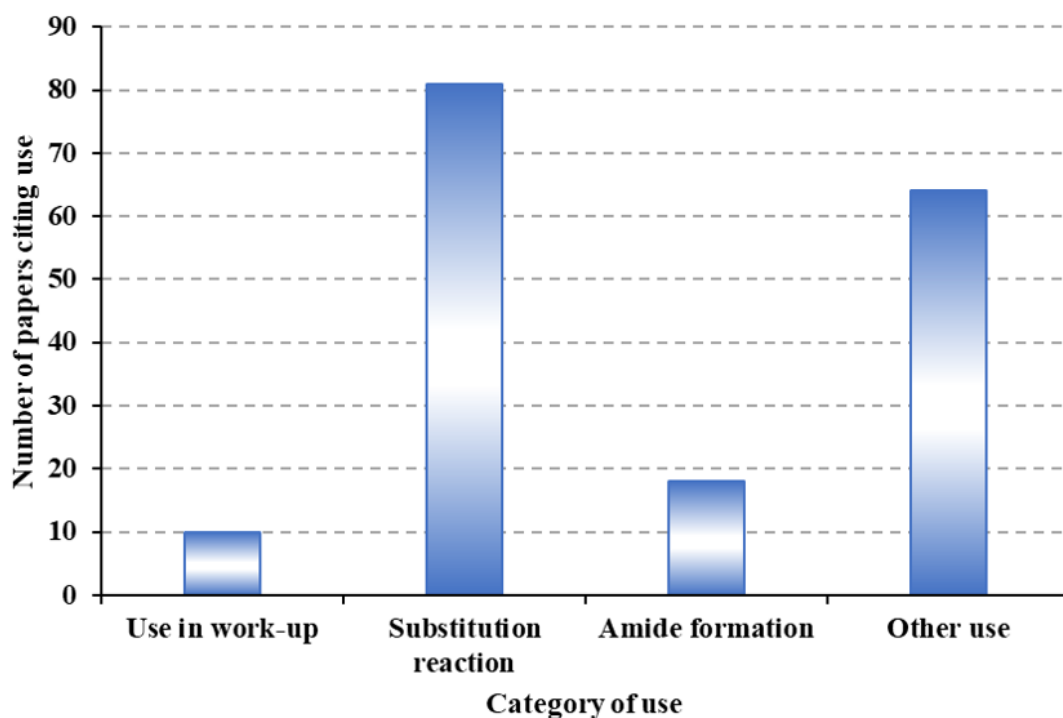
Figure 3.1: Use of dipolar aprotic solvents in work-up, substitution reactions, amide formation and other uses (image from [Ashcroft 2015]). Combined analysis for DMF, DMAc, NMP, and DMSO showing number of uses reported in papers published by *Organic Process Research & Development* between 1997-2012. Only reactions on a scale of 100 g or above were considered.

enamels in the paints and coatings industry, floor cleaning and paint stripping, and significant usage as a photoresist stripper in the electronics industry [FMI 2015].

The Environmental Protection Agency (EPA) in the USA identified the use of paint strippers in both an occupational and consumer setting to be the most likely routes of exposure to NMP [EPA 2015]. NMP is extremely effective at removing paints and coatings and has been widely used as an alternative to dichloromethane-based strippers. However, proposals for both solvents to be banned in paint removers were raised towards the end of the Obama administration, prompting a search for alternatives [EPA 2017]. Although the likelihood of a ban is uncertain under the current administration, these proposals highlight the widespread use and exposure to NMP and support the value of research into replacements. Moreover, there are few suggested green alternatives for dipolar aprotics given in the solvent selection guides discussed in section 1.1.3 [Henderson 2011, Prat 2013].

3.1.2   Known Problems with Cyrene and a Possible Solution

In 2014, this research group published a communication detailing the possibility of using Cyrene, **2**, as an alternative to dipolar aprotic solvents by showing that it has similar dipolarity to the commonly used solvents DMF and NMP, and demonstrating its efficacy in the Menschutkin and fluorination reactions [Sherwood 2014]. Since then Cyrene has also been exhibited as a potential alternative to DMF for metal-organic framework (MOF) synthesis [Zhang 2016], and as a platform molecule for many other interesting structures [Pacheco 2016, Camp 2018].

However, as with all new solvents, new information has arisen following continued testing of the substance, some of which has shown that Cyrene is not a suitable replacement for dipolar aprotics in all circumstances. For example, the ketone moiety exists in equilibrium with the geminal diol in the presence of water [Sigma 2018], a phenomenon which diminished the surface areas observed in the MOFs mentioned above [Zhang 2016], meaning that Cyrene may often need to be used under anhydrous conditions. A second unwanted phenomenon is the formation of a dimer under various basic conditions at elevated temperature [Wilson 2016]. This is due to the acidity of the alpha proton to the ketone. Figure 3.2 gives the structures of the geminal diol formed in water (**16**), and the two aldol products (**17**) and (**18**).

(a) Geminal diol   **(b)** Aldol product 1   **(c)** Aldol product 2
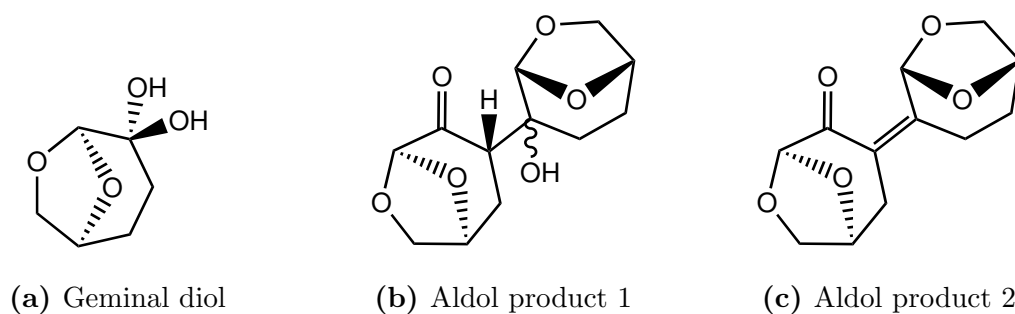
Figure 3.2: Structures of Cyrene's geminal diol (**16**), and two aldol dimers (**17**) and (**18**)

Wilson et al. reported that organic bases were generally tolerated at room temperature. Bases triethylamine and *N,N*-diisopropylethylamine did not cause dimerisation even at 50 °C. However, almost all of the inorganic bases tested caused a reaction at room temperature. The base sensitivity of Cyrene does not preclude it from use as a solvent, but shows that there is still cause to explore other solvents

in the same space. It is possible that an alteration to the structure of Cyrene to a 5-membered ring system could reduce the acidity of the alpha-proton enough to make a difference in its stability towards bases.

3.1.3 Aims of this Chapter

The aims of this chapter were to explore routes from levoglucosenone (**1**) to its lactone derivatives, with the primary focus of creating more diverse options than just Cyrene within the remit of carbonyl containing, cyclic solvents. Considering the base-sensitivity of Cyrene, the aim was to find more stable structures, with the hope of optimising their syntheses to reduce waste and toxicity by trying alternative synthetic conditions. In addition, computer modelling techniques were used to assess the likelihood of a useful outcome of the solvent candidate molecules, by predicting their positions in solvent space relative to known solvents, some of which are still in need of replacement due to their inherent hazards. Finally, the objective was to assess the greenness of the tested reaction methods to give this research standing in the wider context of green chemistry.

3.1.4 Chapter Structure

1. Section 3.2.1 Lactone Concept Development

   - A description of how the research came about and initial conversion of levoglucosenone to a lactone.

2. Section 3.2.2 Application to Cyrene

   - Results of applying the same synthetic methods to Cyrene.

3. Section 3.2.3 Problems with the m-CPBA method

   - An explanation of why a greener synthetic method was sought

4. Section 3.2.4 Synthetic Methods in the Literature

   - A discussion of other synthetic methods that have been used.

5. Section 3.2.5 Use of Amino Acid, L-Lysine, as a Catalyst

- The results of a greener method development.

6. Section 3.3.1 Solvent Space Prediction in HSPiP

   - Predictive modelling done in HSPiP to show what solvent properties are expected from the proposed lactone structures.

7. Section 3.3.2 COSMO-RS

   - Predictive modelling done in COSMO-RS to visualise the chemical potential of each lactone structure.

8. Section 3.4.1 Investigation of L-Lysine Method

   - An exploration of the time-dependence and scaling up of L-lysine catalysed Baeyer-Villiger reaction.

9. Section 3.4.2 Determination of Side-Products of L-Lysine Catalysed Reaction

   - A summary of the assessment of other compounds present, by Martyna Kundrotaite.

10. Section 3.6 Catalyst-Free Baeyer-Villiger Oxidation

    - An exploration of the reaction run without L-lysine.

11. Section 3.4.4 Scale-up

    - Results of an increased reaction size of the catalyst-free Baeyer-Villiger reaction.

12. Section 3.5.1 Hydrogenation of Unsaturated Lactone

    - A description of the hydogenation method applied to LGO derived lactone.

13. Section 3.5.2 Separation

    - A discussion of the reasons for difficult purification.

14. Section 3.5.3 Scale-up

- Results of increasing the size of the lactone hydrogenation.

15. Section 3.5.4 Alternative hydrogenation method

    - Results of an attempt to improve the greenness of the hydrogenation reaction.

16. Section 3.6.1 MeI method

    - Methylating the lactone to create a polar aprotic solvent.

17. Section 3.6.2 Scale-up

    - Results of increasing the volume of the lactone methylation.

18. Section 3.6.3 Greening the Reaction

    - Results of an attempt to improve the greenness of the methylation reaction.

19. Section 3.7.1 Base Sensitivity Tests

    - Testing both hydrolysed and methylated lactones for stability toward bases.

3.2 Oxidation of Levoglucosenone

3.2.1 Lactone Concept Development

Generation of formate ester

It is relevant to the narrative of this work to include an observation previously made by Dr Thomas Farmer regarding the epoxidation of levoglucosenone (LGO, **1**) [Farmer 2013]. In order to develop the functionality of the double bond, an attempt was made to introduce an epoxide ring to LGO following a published method [Berhal 2009] of alkene epoxidation. The referenced work showed that *meta*-chloroperoxy benzoic acid (*m*-CPBA) could be successfully used as an oxidising agent in the epoxidation of an alkene in the α, β- position to an ester, in this case to form
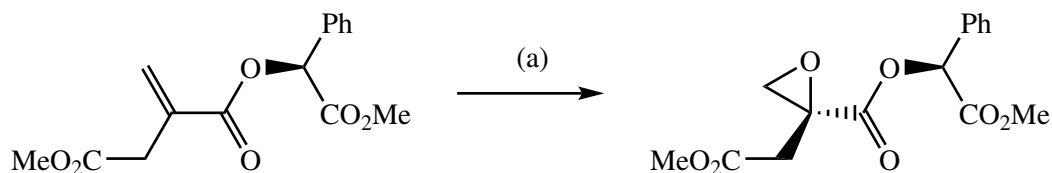
Scheme 3.1: (a) 2 equivalents of enriched *m*-CPBA (95% purity), 2,4,6-tri-tert-butylphenol (2% w/w *m*-CPBA) in 1,2-dichloroethane (DCE), reflux, 16 h [Berhal 2009].

an intermediate in the synthesis of chiral alkaloids, achieving a 72% yield for the conversion shown in scheme 3.1.

It was hoped that these conditions would have the same effect on the double-bond of LGO. Scheme 3.2 shows the anticipated mechanism of LGO (**1**) epoxidation to its epoxidised counterpart (**19**).
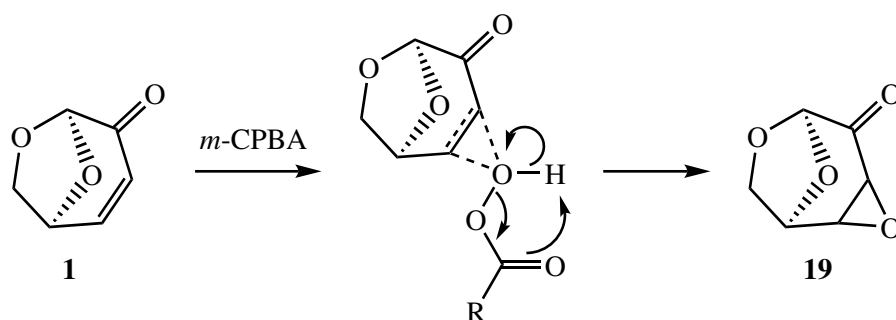
Scheme 3.2: Mechanism of epoxidation initially expected from the reaction between levoglucosenone (**1**) and *m*-CPBA to produce epoxide **19**.

On analysis of the product, however, it became clear that product **19** had not been formed due to the presence of two distinct signals at δ 6.2 and δ 7.4 ppm in the ^{1}H NMR spectrum, indicative of two alkene protons. In the case of the epoxide, the alkene peaks would be expected to disappear and be replaced by new peaks at δ 3-4. In fact, it was established that the oxidation had caused an oxygen atom to insert within the 7-membered ring, next to the carbonyl group, and on the opposite side of the double-bond. This can be explained by the Baeyer-Villiger mechanism, which results in structure **20**, as shown in scheme 3.3. It is worth noting that, during an epoxidation, the mechanism involves a peracid that remains protonated, while the Baeyer-Villiger involves the deprotonation of the peracid by the carbonyl, forming a hard nucleophile.
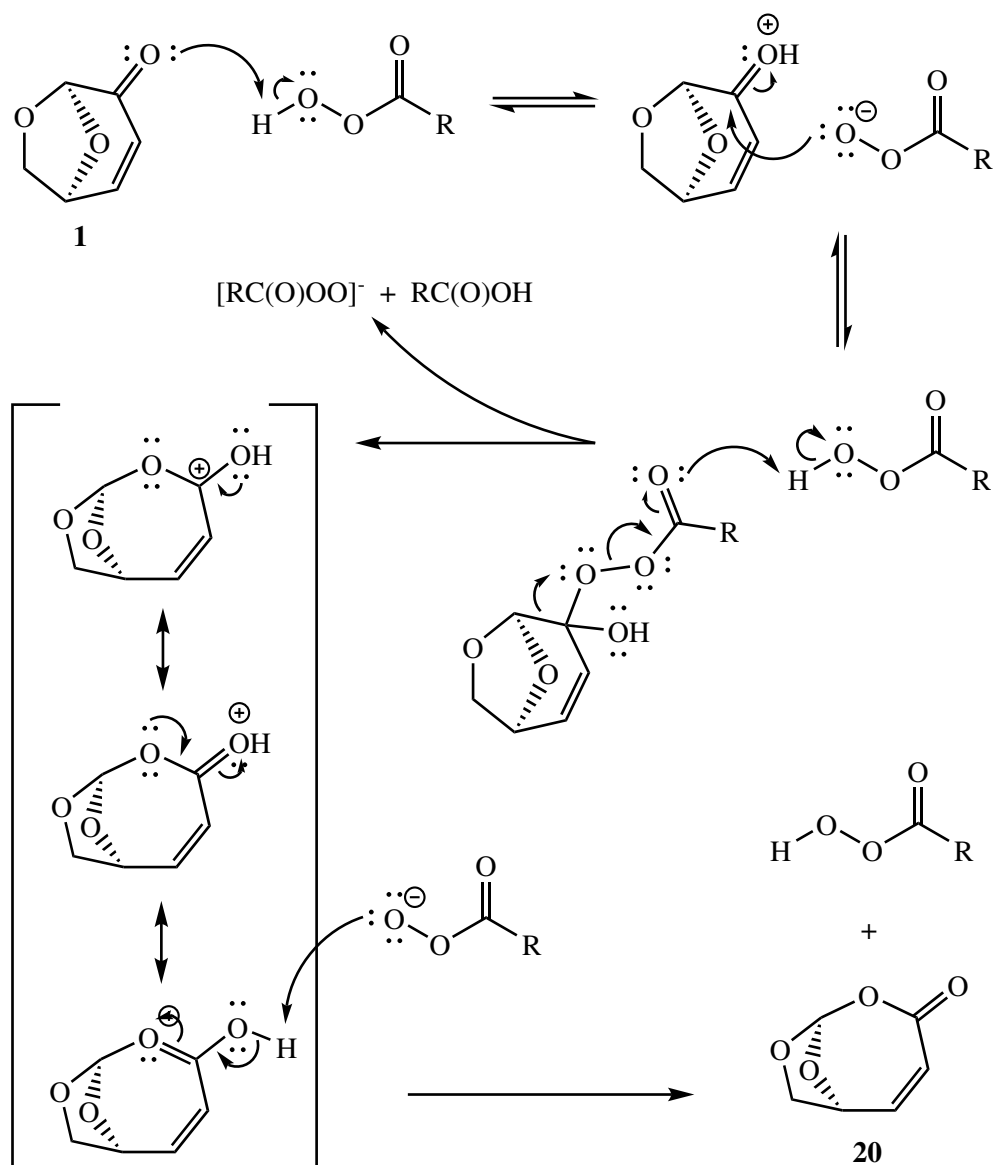
Scheme 3.3: Proposed Baeyer-Villiger reaction mechanism of levoglucosenone (LGO, **1**)
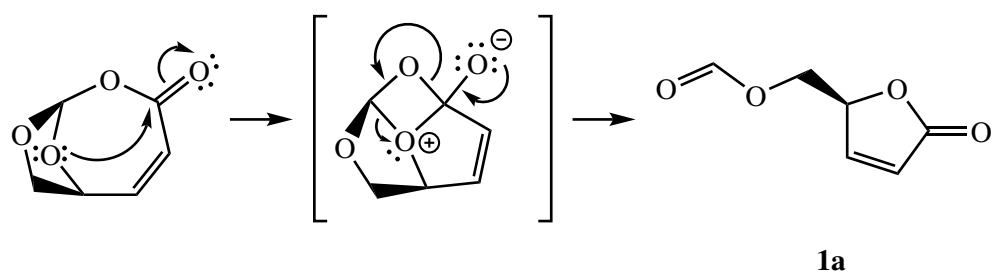
Scheme 3.4: Proposed mechanism of the Baeyer-Villiger product (**20**) rearrangement into lactone (**1a**), giving unracemized lactone.

While this mechanism holds true logically and in accordance with previous reports, the Baeyer-Villiger product (**20**) was not actually observed. Structure (**20**)

was found to have rearranged intramolecularly to form the resulting 5-membered ring lactone with a formate ester functional group (**1a**), as shown in scheme 3.4. The Baeyer-Villiger product **20** was first described by Shafizadeh et al. [Shafizadeh 1979], although the rearrangement of (**20**) was not speculated upon, causing confusion about the "surprisingly deshielded" tertiary proton at δ 8.17 in the ^{1}H NMR spectrum.

The rearrangement was later reported in the literature, also using *m*-CPBA as the oxidising agent [Koseki 1990], clarifying that the formate hydrogen of (**1a**) has a distinctive signal at δ 8 in the ^{1}H NMR spectrum, which differentiates it from the Baeyer-Villiger product (**20**). This rearrangement seems to happen so readily that the Baeyer-Villiger product itself has never been isolated.
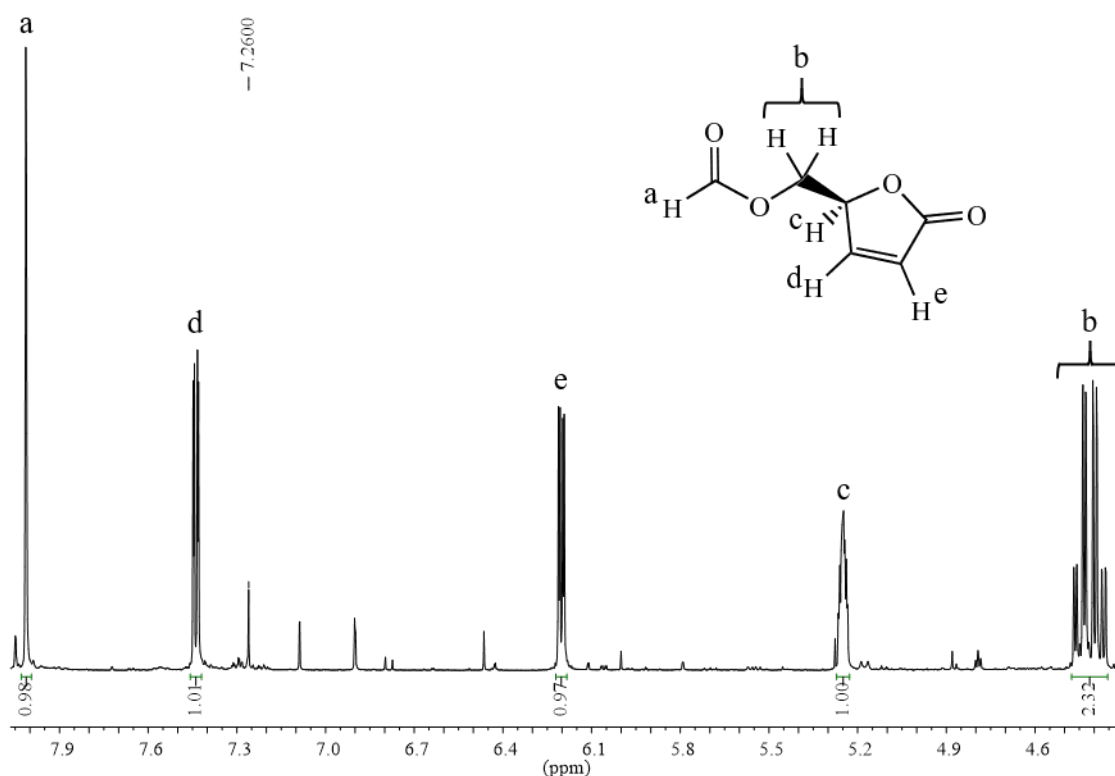
Figure 3.3: Proton (400 MHz) NMR spectrum for crude **1a** in CDCl$_3$, with annotation. For full spectrum, see appendix XII.

The functionality of molecule **1a** led to speculation about further derivatisation, both through hydrogenation of the double bond, either post-lactonisation or by performing the Baeyer-Villiger oxidation on Cyrene to form saturated lactone **2a**, and through hydrolysis of the formate ester to produce **1b**.

(a) 2a   (b) 1b
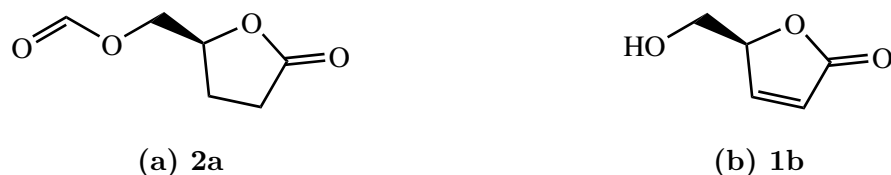
Figure 3.4: Structures of saturated lactone **2a**, and hydrolysed lactone **1b**.

Both of these possibilities had been proven to work in the literature [Koseki 1990] and so an avenue for adjustments in the physical and chemical properties of the solvent molecules was unfurled.

3.2.2 Application to Cyrene

Cyrene, **2**, also known as dihydrolevoglucosenone, was tested using the same reaction conditions as previously described to confirm that the Baeyer-Villiger oxidation/rearrangement could be applied to this molecule. The reaction is shown in scheme 3.5.
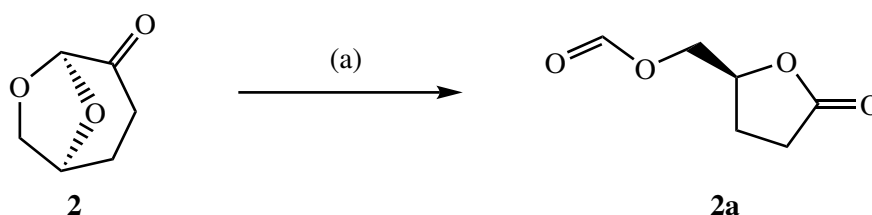
Scheme 3.5: Conditions (a): Cyrene (8 mmol) and tert-tributylphenol (0.4 mmol) dissolved in dichloroethane (32 mL) and mixed with m-CPBA (12 mmol), RT, 24 h.

With the C=C double bond saturated, Cyrene and the resulting lactone, **2a**, were considered to be less susceptible to attack in general, and immune to Michael addition. Due to the lengthy workup procedure, a considerable mass-loss was observed (approximately 50%) after the solvent was removed from the crude product. However, as the reaction was done on a 1 g scale, the loss was expected to improve on scale-up. Radical inhibitor 2,4,6-tri-*tert*-butylphenol, ttbp, also remained as a contaminant owing to solubility in the organic phase, therefore further separation was necessary. Proton NMR showed no clear signal indicating the presence of m-CPBA in the crude product. However, TLC showed spots consistent with m-CPBA alone. This is likely because m-CPBA is a strong absorber of UV radiation and

could be detected in low concentrations. m-CPBA produced four spots on the TLC plate (KMnO$_4$ solution), indicating a number of breakdown products. This is a factor which should be considered with regard to the efficiency of the reaction and the difficulty observed in isolating the product. In the interest of green chemistry, it must be acknowledged that this reaction is not clean with regard to the inputs required and the poor performance discussed hereafter.

The GC trace of the crude product after biphasic separation, see appendix VII), also shows no presence of m-CPBA at 4.6 minutes and minimal amounts of Cyrene at 3.4 minutes. Due to the large number of contaminants visible in the gas chromatogram, column chromatography was deemed the most likely method to achieve effective isolation of the product. Therefore, after testing by TLC, an automated chromatography column was run using a gradient of 2% up to 20% ethyl acetate in cyclohexane as the mobile phase. Detection was by UV absorbance, but due to limited UV absorbance by the product, all fractions were collected and monitored by TLC.

In the first trace, several eluents were seen early on, but the product did not elute, so the column was flushed through with a mobile phase gradient reaching 100% ethyl acetate, at which point the product began to elute. This procedural observation cast light on the significance of the polarity and hydrogen bonding of product **2a**, with regard to its separation.

This initial test produced a yield of only 14%, however, the formation of lactone **2a** was confirmed by proton NMR spectroscopy (see appendix XIII), the main difference from **1a** being the presence of CH$_2$ signals between δ 2-3 instead of the CH signals at δ 6.2 and 7.4. IR was used to prove the existence of two carbonyl groups present in **2a**, which differentiate it from the Baeyer-Villiger product, **21**, as shown in figure 3.5.

Shafizadeh et al. reported a 37% yield of Baeyer-Villiger product **21** using 85% pure m-CPBA at 1.2 molar equivalents. Koseki reported a 59% yield when Shafizadeh's procedure was used, although the purity of the oxidising agent was not stated. The various peaks visible in the GC trace shown in appendix VII show that the reaction does not go cleanly. Speculatively, this could be due to differences in the quality of the m-CPBA used in the reaction. In the current work, m-CPBA
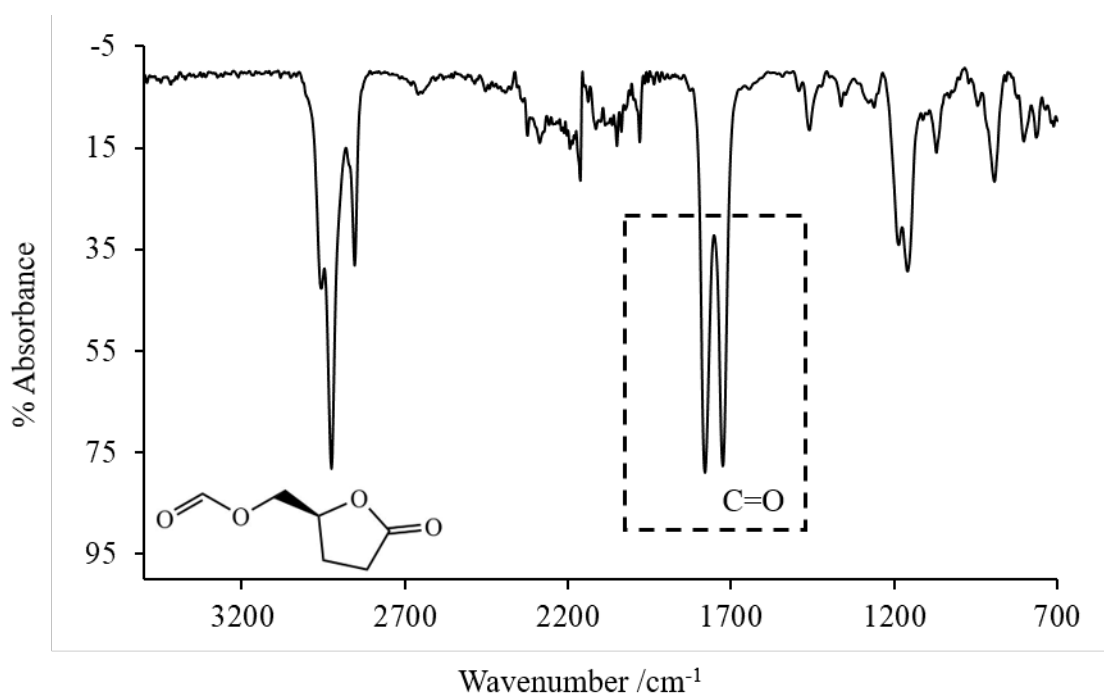
Figure 3.5: IR spectrum of **2a** evidencing the presence of two carbonyl groups (dashed line).

of purity ⩽ 77% was used, and it was clear from the TLC that the chemical had at least three impurities, or was prone to breaking down during chromatography. Unfortunately, a higher purity is not readily available commercially.

Simplification

When considering how to simplify the separation of multiple resulting compounds from this reaction, it was noted that ttbp was originally used by Berhal et al. as a radical inhibitor to avoid polymerisation of the alkene during epoxidation [Berhal 2009]. Using Cyrene as a starting material, this issue was side-stepped and so the reaction was repeated without ttbp to produce a cleaner product. The oxidation was repeated for varying lengths of time: 7 h, 24 h and 3 days. Proton NMR spectra are given for the products of each of these reactions along with Cyrene and *m*-CPBA starting materials in figure 3.6.

What can be gleaned from this comparison is that the conversion of Cyrene was fairly limited at the 7 h mark but appeared to have gone to completion within 24 h, as is illustrated by the disappearance of signals at δ 3.95, 4.05 and 5.10. Accurate conversion and selectivity were not determined due to the difficulty obtaining well
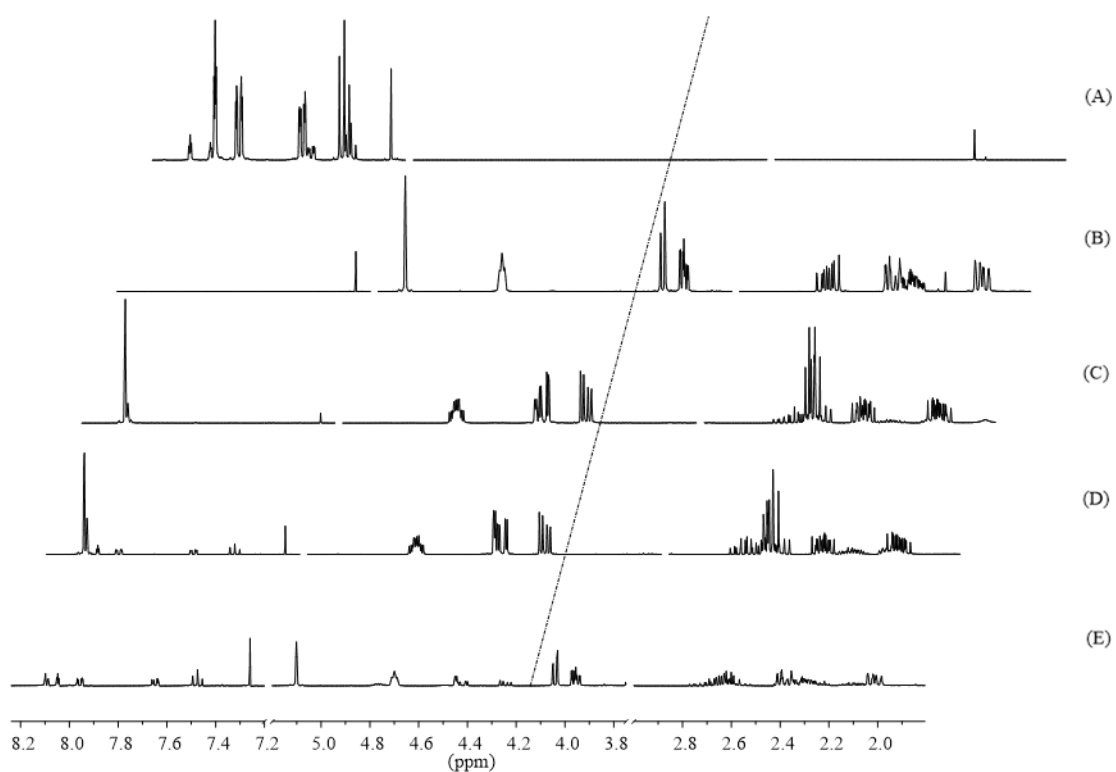
Figure 3.6: Proton (400 MHz) NMR spectrum for *m*-CPBA (A), Cyrene (B), product from Cyrene oxidation by *m*-CPBA after 3 days (C), 24 h (D), and 7 h (E). All run in CDCl$_3$. To the immediate right of the dashed line are the CH$_2$ signals next to the acetal of Cyrene. To the left are the CH$_2$ signals of the pendant formate ester group of **2a**.

resolved GC traces. However, since the overall performance of the reaction was quite poor and has negative connotations in a green context, optimisation of the separation was not pursued.

3.2.3 Problems with the m-CPBA method

As is common for halogenated solvents, dichloroethane may cause cancer after prolonged exposure, but it is also highly flammable [ECHA 2018b], which is a disadvantage in comparison to chloroform and dichloromethane. In addition to the safety concerns, the biphasic workup to the reaction is resource intensive in relation to the quantity of product produced, and m-CPBA is a bulky reactant, which provides only an oxygen atom towards the final product. This gives the reaction an atom economy of only 48%.

The use of unsustainable materials alone was a good reason to look for alternatives to this reaction. Furthermore, the performance of the reaction was poor and, considering the number of breakdown products in the crude material, aiming for adequate isolation of the desired product was getting to be costly and time-consuming.

3.2.4 Synthetic Methods in the Literature

Following the successful application of the m-CPBA promoted Baeyer-Villiger reaction, a literature search was done to find less wasteful synthetic methods. On contemplation of the structure of **1a** and **2a**, it was recognised that the formate ester functional group could be easily hydrolysed to produce an alcohol, a concept that was also explored in the work of Koseki et al. [Koseki 1990]. By modifying the pendant chain in this way, it was expected that a more stable molecule could be formed whilst entering a different area of "solvent-space". A literature search yielded three previously known methods to produce alcohols **1b** and **2b** from levoglucosenone and Cyrene, as described in table 3.2.

A common feature to all of these methods is the need for a hydrolysis step to convert all of the formate ester functional groups to alcohol groups. Koseki and Flourat both used heating in a mixture of methanol hydrochloric acid to achieve this. Paris et al. used acidic Amberlyst 15 to push the hydrolyis to near completion, although their work tested a variety of zeolites as catalysts for the Baeyer-Villiger oxidation,

Table 3.2: A summary of the methods to produce lactones **1a**, **1b**, **2a** and **2b** published prior to the execution of this research.

Author	Reaction conditions	Compounds cited	% Yield
[Koseki 1990][a]	(1) AcOOH (1.2 equiv.), AcOH, RT, 48 h; (2) Me$_2$S, 1 h; (3) MeOH, HCl (catalytic), 45 °C, overnight	**1a**, **1b**, **2a**, **2b**	Not provided (intermediates)[b]
[Paris 2013]	1,4-Dioxane, H$_2$O$_2$ (35%), zeolite, 100 °C, 4 h; Amberlyst 15, RT, 6 h	**1a**, **1b**	72, 89 respectively (Highest of each)
[Flourat 2015][c]	(1) CAPSO[d], acid, salt forms, CAL-B, H$_2$O$_2$ (1.2. equiv.); (2) Methanol, HCl (12 equiv.), 45 °C or Amberlyst-15, 1,4-dioxane, RT	**1b**, **2b**	83, 78 respectively

[a]Furthered in [Kawakami 1990]

[b]The yield of the resulting compounds were 65% via lactone **1b** and 87% via lactone **2b**, so the yields of **1b** and **2b** can be assumed to have been higher than this.

[c]Optimised in [Teixeira 2016]

[d]3-(cyclohexylamino)-2-hydroxy-1-propanesulfonic acid

achieving varying selectivites of **1a** and **1b** prior to the hydrolysis step [Paris 2013]. Clearly, the ideal synthesis would be done in one step from levoglucosenone and Cyrene, which is explored further in this work.

There are other advantages and disadvantages to each of these methods. For example, in Koseki's publication [Koseki 1990], the solvent used was acetic acid (AcOH), which can be bioderived. Although fermentation currently makes up only about 10% of global production [Vidra 2018], it is still a better alternative to the CH$_2$Cl$_2$ used previously. Another advantage is that the peracetic acid used to perform the oxidation is less bulky than m-CPBA (76.0514 g mol^{-1} compared to 172.57 g mol^{-1}), and can potentially be made in-situ using H$_2$O$_2$ and the already present AcOH, which means there is less inherent waste associated with the reaction due

to a higher atom economy. However, AcOH is corrosive and flammable. Further to this, the use of dimethyl sulfide $((CH_3)_2S)$ to quench excess peroxide forms DMSO, which is difficult to remove by vacuum distillation and is less desirable than catalytically decomposing the peroxide. Paris et al. have developed a method with the advantage of a heterogeneous catalyst in that it can be filtered off afterwards [Paris 2013]. This is a significant benefit to the workup of the procedure, however, the catalysts themselves were formed by a convoluted process involving tetraethylammonium hydroxide, tetraethylorthosilicate, a variety of metal complexes and, significantly, hydrofluoric acid (48%), which would be of significant concern on a larger scale due to its extremely corrosive nature, and ability to enter and de-calcify the body. The conditions required 140 °C for 21 days to prepare the catalyts. In addition, 1,4-dioxane, used as a solvent in the actual oxidation, is another highly flammable, suspected carcinogen.

Flourat et al. used a slightly different approach, making use of a yeast lipase enzyme, *Candida antarctica* B (CAL-B), to convert levoglucosenone into the lactone [Flourat 2015]. This method is arguably the greenest, using the mildest conditions and achieving shorter reaction times. However, it was still necessary to perform a hydrolysis afterwards, for which heterogeneous catalyst Amberlyst-15 could be used, but again involved the use of 1,4-dioxane. CAL-B, being immobilised on acrylic resin, could be removed by filtration and re-used after washing with ethyl acetate. However, the authors reported that suspending the CAL-B in ethyl acetate caused approximately 45% of the enzymatic activity to be lost, which meant more had to be used.

While there is a limited number of examples of publications citing production of the formate ester lactones **1a** and **2a**, many more refer to the production of the alcohols **1b** and **2b**, generally as an intermediate within a longer synthesis. Since a full discussion of all of these methods would be bulky and digress significantly from the topic of this thesis, only the two syntheses derived from bio-based compounds have been described, as pictured in figure 3.7b.

The synthesis of lactone **2b** from L-glutamic acid was originally described by Taniguchi et al. in 1974 and most recently reported by Anderl et al. [Taniguchi 1974, Anderl 2018]. The method has seen little improvement in this time, save for

(a) L-Glutamic acid

(b) D-Mannitol

d
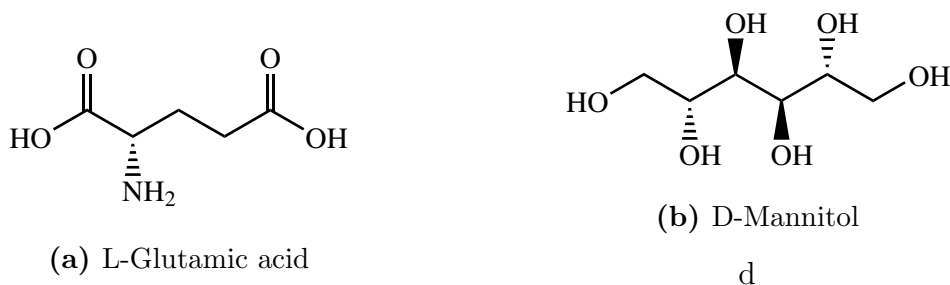
Figure 3.7: Structures of L-glutamic acid and D-mannitol used to synthesise **2b**: (a) i) NaNO$_2$, HCl, H$_2$O, 0 °C-RT, 79%; ii) BH$_3$·SMe$_2$, dry THF, 0 °C-RT, 69% [Anderl 2018]; (b) i) ZnCl$_2$, acetone, 55%; ii) NaIO$_4$, NaHCO$_3$, MeOH, 1 h, 0 °C, NaBH$_4$, 81%; iii) *p*-toluenesulfonyl chloride, pyridine, 96%; iv) sodium iodide in refluxing acetone; v) ethyl malonate, DMF, sodium hydride, 3 h, 100 °C, 63%; vi) MgC$_{12}$·$_6$H$_{20}$, DMF, 20 h, reflux, 95% [Takano 1981a+b].

the removal of a middle step refluxing the intermediate in a combination of ethanol and benzene with *p*-toluenesulfonic acid. An alternative route from D-mannitol was first published by Takano et al. in 1981 [Takano 1981b]. By contrast, this route has seen very little replication in the literature, probably due to the lengthy, 6-step process. It seems reasonable to rule out the second method as a viable competitor on account of the multiple steps and undesirable inputs. The synthetic complexity s reduced when using the L-glutamic acid route, but still requires the use of a toxic borane dimethyl sulfide complex, which must be kept completely dry to avoid the production of H$_2$ gas.

3.2.5 Use of Amino Acid, L-Lysine, as a Catalyst

An alternative method for Baeyer-Villiger oxidation was tested, using a combination of hydrogen peroxide (H$_2$O$_2$ 30% aqueous solution) and L-lysine, shown in figure 3.8.
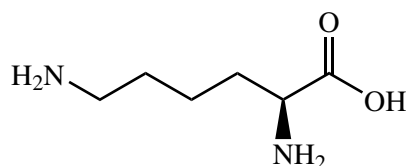
Figure 3.8: Structure of L-lysine

This was based on the publication of research by Kim et al., in which they demonstrated the use of various amino acids in the epoxidation of cyclohex-2-enone
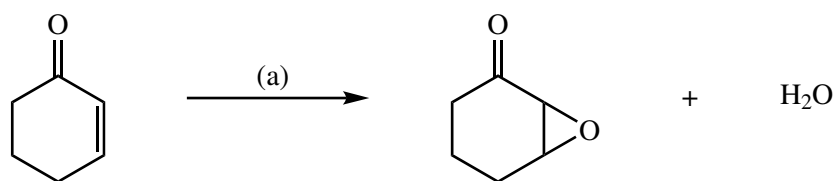
Scheme 3.6: Epoxidation of cyclohex-2-enone as published by Kim et al. with conditions: (a) cyclohexen-2-one, amino acid (0.1 equiv.), H_2O_2 (30%, 3 equiv.), *t*-amyl alcohol (internal standard), water, 25 °C, 2 h [Kim 2011].

They reported that lysine and arginine converted >99% of the starting material, while histidine achieved a 70% conversion, and all other amino acids tested had almost had virtually no effect at all. The publication provided only results obtained via GC analysis and no complementary characterisation and therefore it can be speculated that the products observed were not in fact epoxides. Since the *m*-CPBA catalysed reaction had resulted in Baeyer-Villiger oxidation of levoglucosenone rather than epoxidation, it was hypothesised that lysine would have the same effect, with the added benefit of solubility in water so that the whole reaction could be run in aqueous solution. It was hoped that the amphiphilic amino acid would serve to promote both Baeyer-Villiger oxidation and subsequent hydrolysis of the formate ester (scheme 3.7).
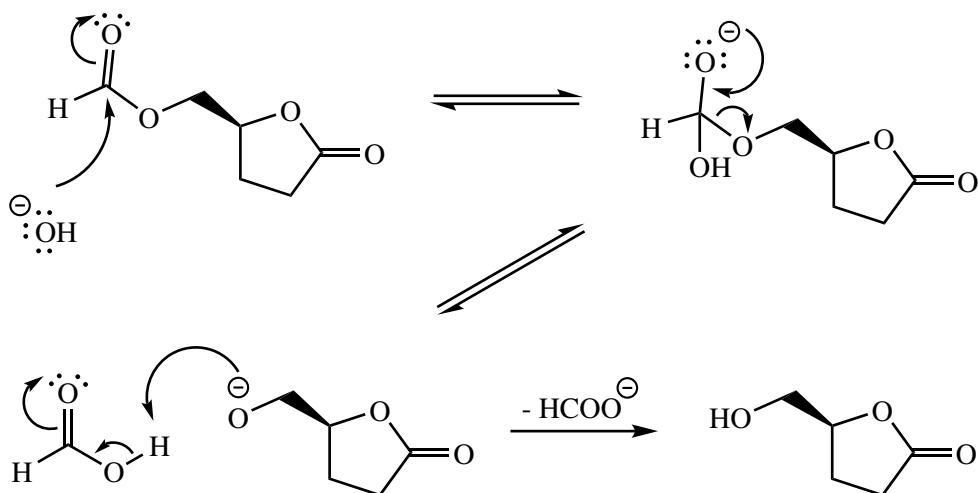
Scheme 3.7: Hydrolysis of lactone **2a** to form **2b**.

The reaction was initially tried on both LGO and Cyrene, using water as a solvent to dissolve all of the components. The reactions were run at room temperature

for 1 day using 3 equivalents of H_2O_2 and 0.1 equivalents of L-lysine. To extract the product from L-lysine, the mixture was subsequently dried and triturated with acetonitrile. The reaction is summarised in scheme 3.8.
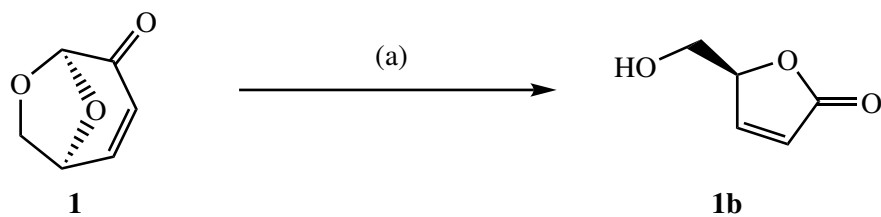
Scheme 3.8: Conditions (a): levoglucosenone (10 mmol) and L-lysine (1 mmol) were dissolved in wate (16 mL). H_2O_2 (30 wt%, 3 equiv.) added dropwise and stirred at RT, 24 h.

When compared to the product from the *m*-CPBA reaction, initial proton NMR results showed that lactones **1a** and **2a** were still present after 24 h, but no longer the major product. Instead, new signals were observed for **2b** at δ 3.87 and 3.63 (both dd), corresponding to the CH_2 protons on the pendant alcohol, indicating that some formate ester hydrolysis had occurred (scheme 3.4). These protons were more shielded than the equivalent protons on **2a** (δ 4.38 and 4.20), which is consistent with the slightly less electron-withdrawing, neighbouring alcohol group. Although full conversion of Cyrene was obtained, selectivity was achieved of about 2:1 of lactones **2b:2a**, with smaller unidentified side-products present. Similar results were obtained when running the reaction on LGO, culminating in a selectivity ratio of 5:3 of lactones **1b:1a**. Again, a number of impurities were observed in the NMR spectrum but it was generally cleaner than the Cyrene analogue and so further testing was primarily based on LGO with the assumption that hydrogenation could be performed in a later step.

3.3 Predictive Modelling

3.3.1 Solvent Space Prediction in HSPiP

HSPiP was previously used to anticipate the relative solubility of a set of amides in toluene and *p*-cymene by predicting their Hansen solubiility parameters (HSP). In the development of previously unexplored solvents, it can be a useful tool for

gauging the position of a particular substance in HSP space. By comparison with traditional solvents with known properties, it is possible to get an idea of how the new solvent will behave in terms of intermolecular interactions, physical properties, and usefulness in certain reaction types. Due to the serendipitous success of synthesising lactone **1a** in section 3.2.1, it was deemed necessary to systematically screen the product and its close analogues by modelling to predict the solvent behaviour of each. HSPiP was, therefore, used to predict the HSP for each of the lactones **1a**, **1b**, **2a** and **2b**, along with LGO and Cyrene. In addition, as polar protic solvents are not in the highest demand with respect to green chemistry, ethanol and water being examples of the "greenest" solvents available, the lactones were also considered with a view to methylating the pendant alcohol and thus the possibility of new aprotic solvents (**1c**) and (**2c**), depicted in figure 3.9.

(a) **1c** (b) **2c**

Figure 3.9: Structures of methylated lactones from LGO and Cyrene.

To find out where each of the lactones sat with respect to other solvents in HSP space, they were modelled in HSPiP amongst a variety of traditional solvents, as shown in figure 3.10. It is important to note here that the modelling of lactones **1a**, **2a**, **1b**, **2b**, **2c**, LGO and Cyrene by HSPiP was published in a paper subsequent to this work being carried out [Pacheco 2016].

Cyrene was already known to be close in proximity to NMP [Sherwood 2014] and, in fact, both LGO, Cyrene and lactones **1a**, **1c**, **2a** and **2c** fall within the green sphere shown surrounding NMP in figure 3.10a. From the 2D plot in figure 3.10b, the two alcohols **1b** and **2b** are shown to be similar in polarity to NMP but much much further away with respect to H-bonding than the other lactones. This H-bonding was the primary reason for methylating the lactones to form **1c** and **2c**, and the prediction suggested this would be effective, although some polarity would be lost in the process. The alternative 2D perspectives given in appendix III show that, in terms of dispersion forces, lactones **1b** and **2b** were the closest to NMP, although the spread was much less broad than H-bonding. It should be noted that the δD
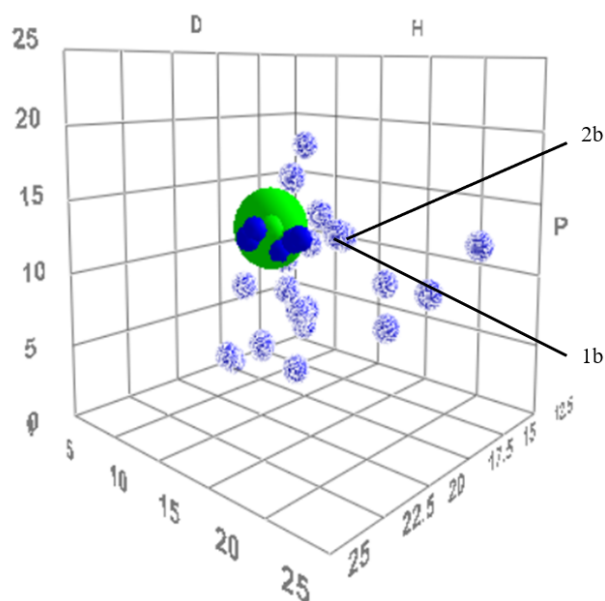
(a) 3D output from HSPiP depicting various traditional solvents, LGO, Cyrene and all 6 lactones. NMP is represented by the green dot. Lactones **1a**, **1c**, **2a** and **2c** all fall within close range of NMP (solid blue dots). Lactones **1b** and **2b** fall further to the right due to H-bonding of the alcohol group.
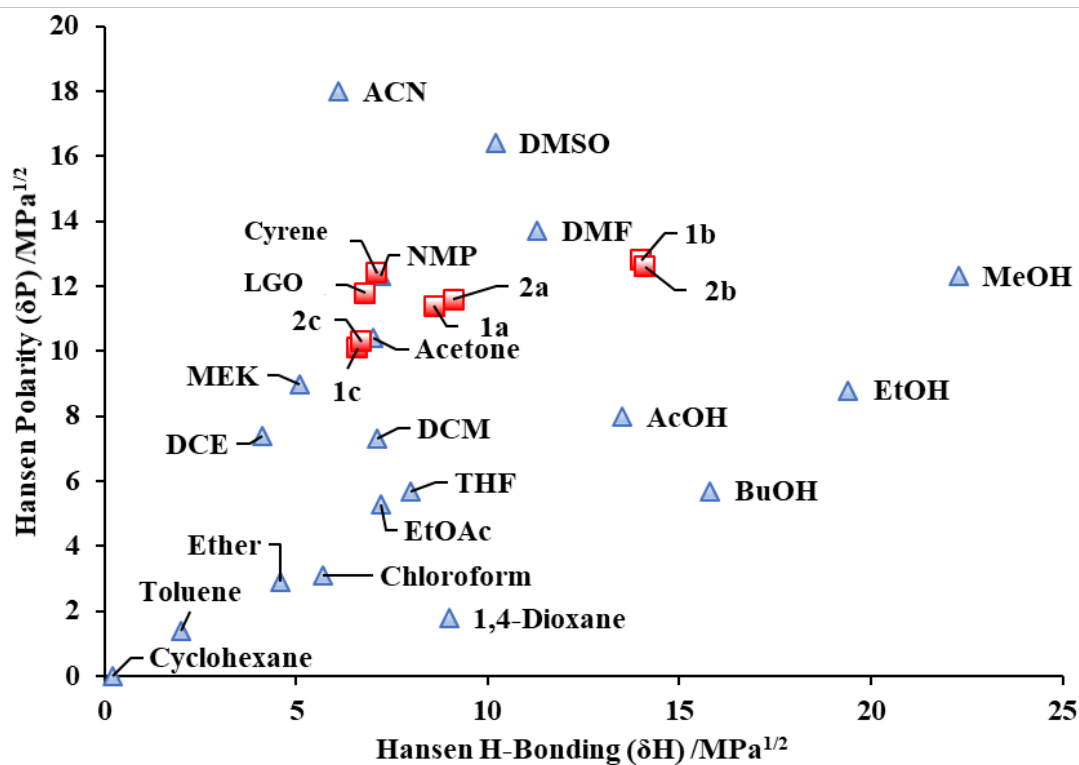
(b) 2D representation of δP against δH featuring compounds **1-1c** and **2-2c** among traditional solvents. Two other perspectives are given in appendix III.

Figure 3.10: HSPiP predictions: (a) 3D perspective (b) 2D perspective.

axis has been expanded for clarity. The reason such focus was given to the position of NMP relative to the lactones is that, as previously explained, dipolar aprotics generally have inherent toxicity issues and few alternatives from which to choose. It is difficult to mimic their properties without the incorporation of heteroatoms, such as nitrogen and sulfur, which are often associated with those risks to health.

3.3.2   COSMO-RS

A further modelling tool was used to predict the chemical potentials of each lactone structure. This was done as a complementary study to HSPiP using COSMO-RS, which is a highly structure-based prediction that takes even the molecular conformation into account. The COSMO-RS based software COSMOconf and COSMOtherm were used to determine the σ-surfaces of each of lactones **1a**, **2a**, **1b**, **2b**, **1c** and **2c**, as shown in figure 3.11.

The most striking thing about the images in figure 3.11 is the clear presence of positive charge density, represented by the bright blue colour in images 3.11c and 3.11d. This is, expectedly, due to the O-H bonds present in lactones **1b** and **2b**, which are obvious hydrogen bonding donors. Qualitatively, what can be observed is that lactones **1a** and **2a** seem to be similar in charge distribution to lactones **1c** and **2c**, barring the formate ester carbonyl. This corresponds well with the HSPiP prediction, which estimated that they would be similar in polarity, with lactones **1c** and **2c** having slightly more H-bonding character.

The σ-surfaces were then used by COSMOtherm to produce σ-profiles of each lactone. The lactones have been represented in pairs alongside NMP, with a separate plot for some representative solvents: hydrocarbon hexane, alcohol ethanol, chlorinated hydrocarbon chloroform, dipolar aprotic NMP, and ester ethyl acetate, as shown in figures 3.12 and 3.13. The profiles have been separated in this way to avoid crowding as the lines can become confused.

At a glance, it can be seen that the O-methylated lactones **1c** and **2c** are the closest mimics of NMP in with respect to their σ-potentials. In particular, the main peak position of lactone **2c**, at about -0.005 e $\mathring{A}^{-2}$, is close in position and intensity. Both lactones **1c** and **2c** have a similar peak in the hydrogen-bond acceptor region. As this peak is caused by the lactone carbonyl group, it appears very similar in each

(a) Lactone **1a**

(b) Lactone **2a**

(c) Lactone **1b**

(d) Lactone **2b**

(e) Lactone **1c**

(f) Lactone **2c**
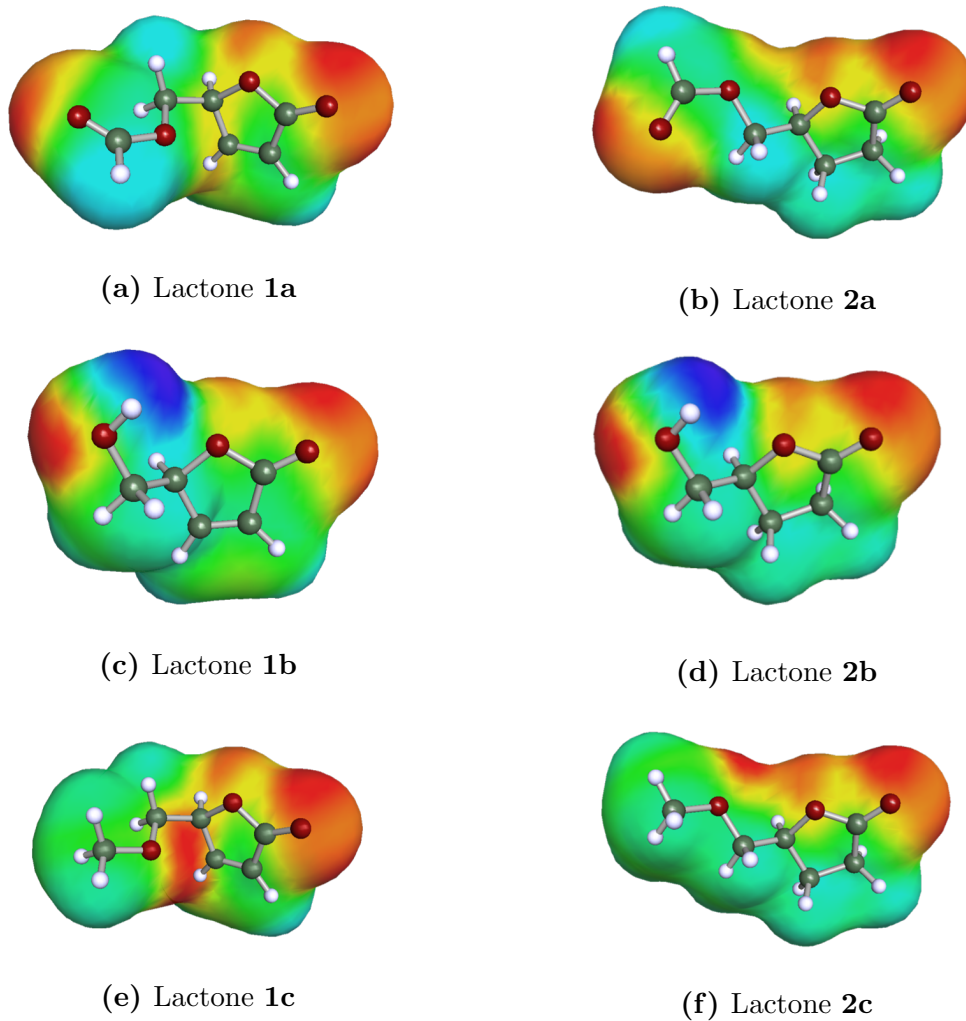
Figure 3.11: σ-Surface of each lactone, as rendered by COSMOtherm.
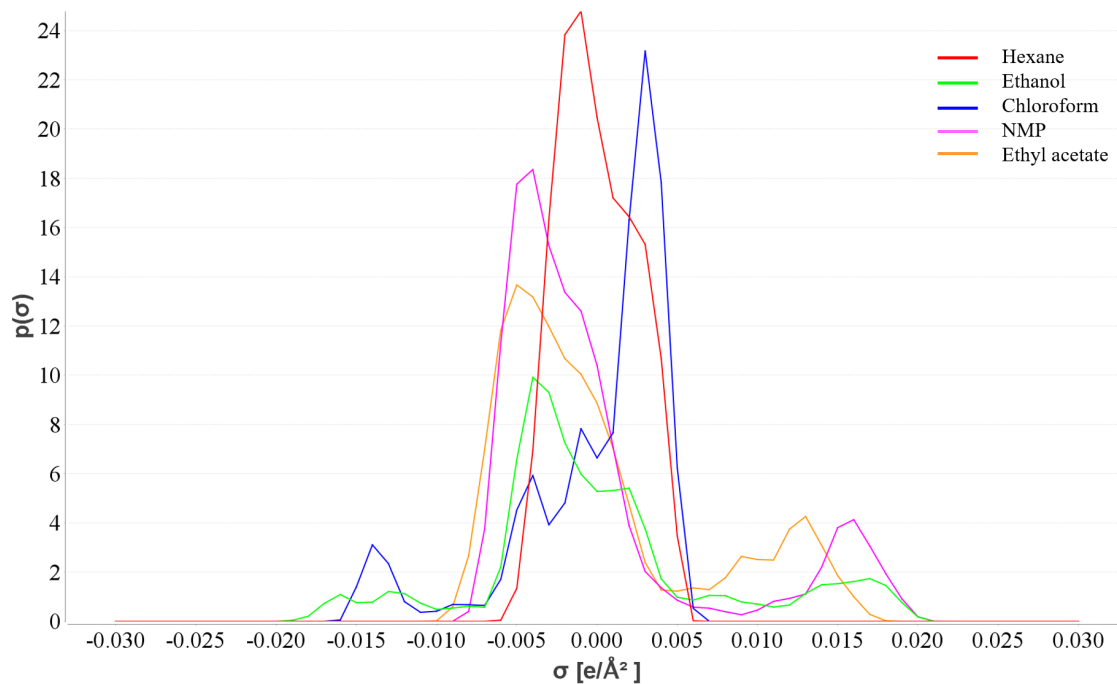
(a) Five representative solvents
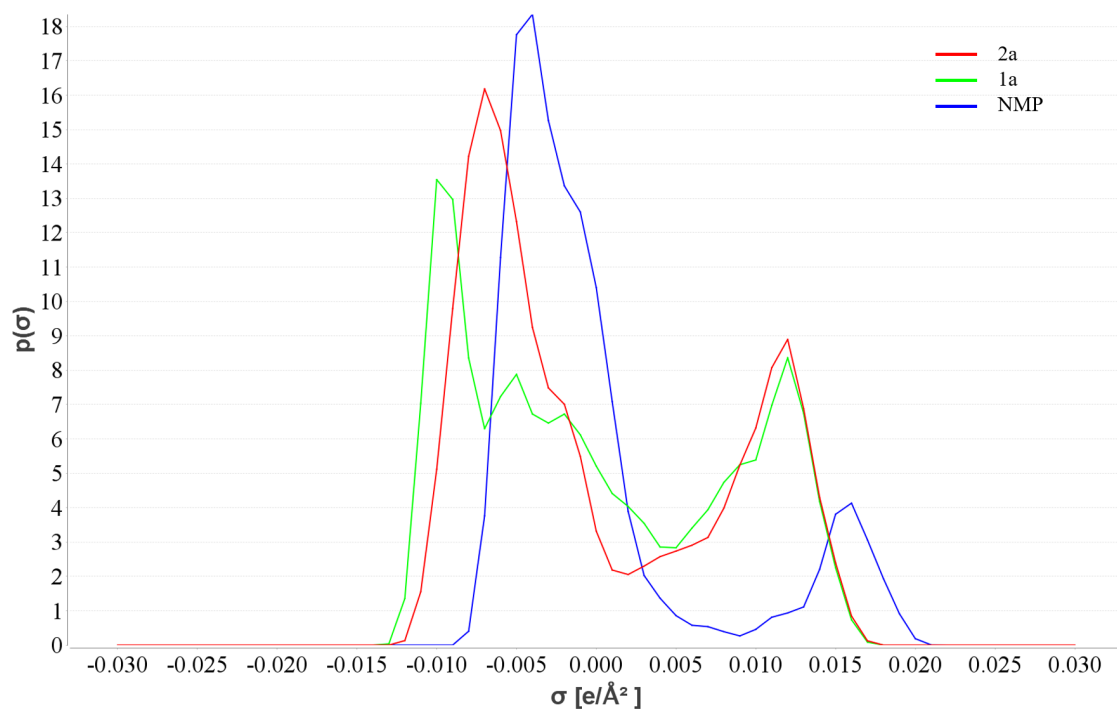
(b) Lactones **1a** and **2a**

Figure 3.12: σ-profiles of (a) 5 traditional solvents and (b) formate ester lactones **1a** and **2a**. Generated by COSMOtherm.
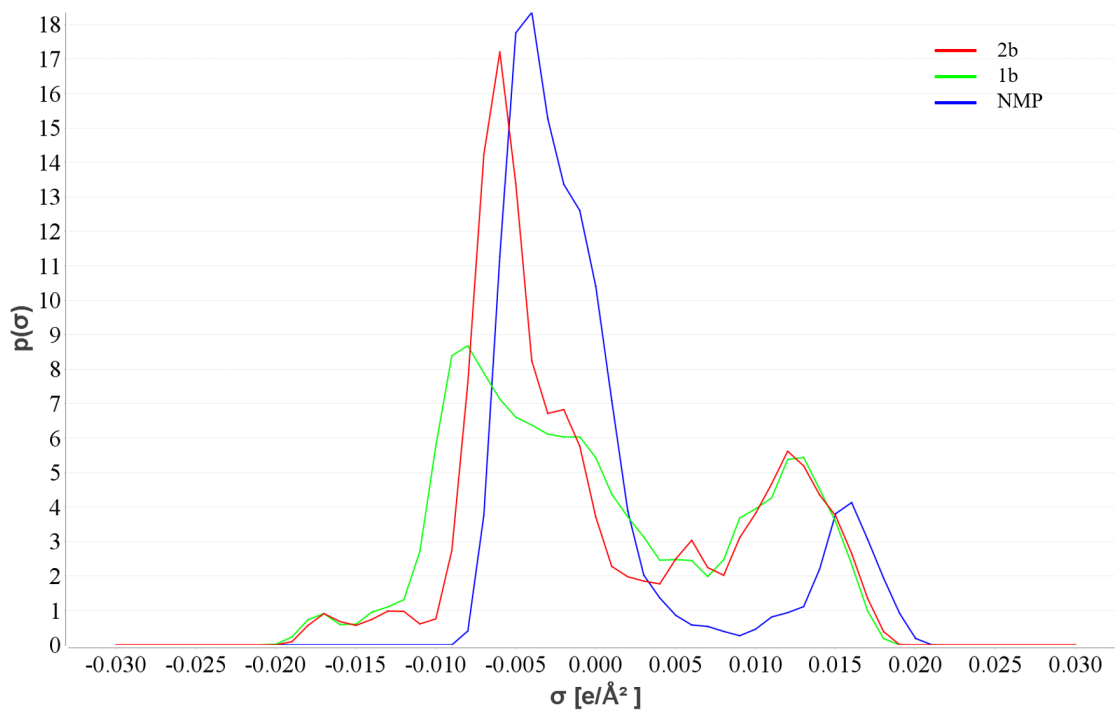
(a) Lactones **1b** and **2b**
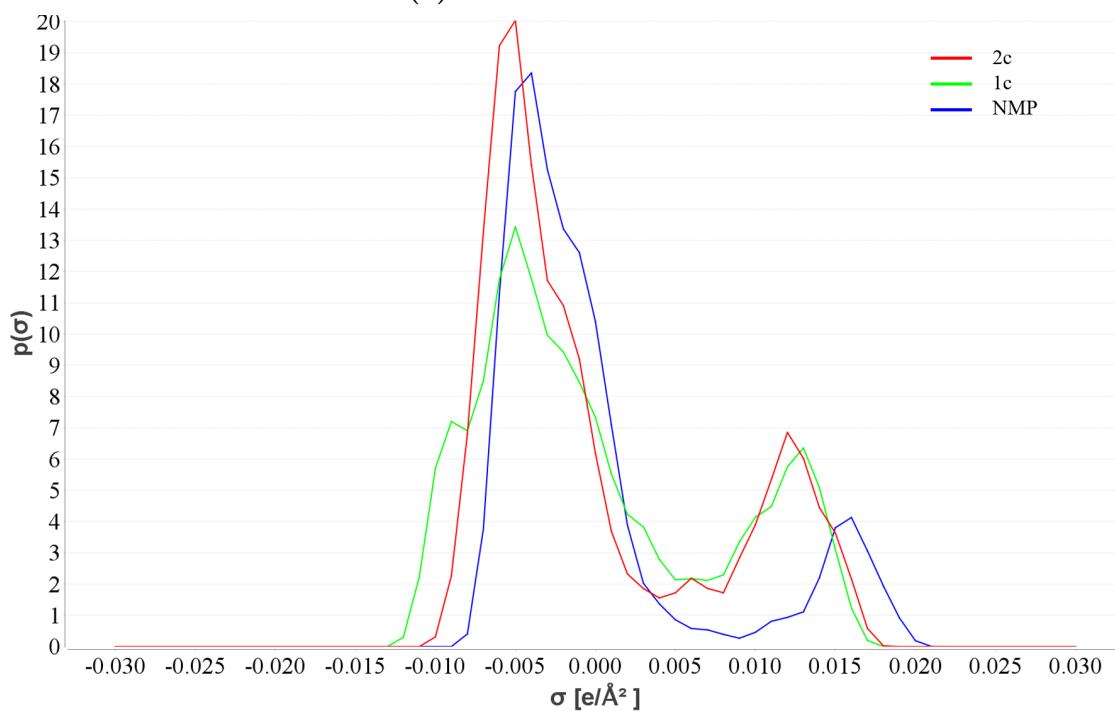
(b) Lactones **1c** and **2c**

Figure 3.13: σ-profiles of (a) lactones **1b** and **2b** and (b) O-methylated lactones **1c** and **2c**. Generated by COSMOtherm.

lactone profile, and doesn't quite match the position of NMP, which is effected by its cyclic nitrogen atom.

Along with σ-profiles, the σ-potential of each lactone was calculated. The output is shown in figures 3.14, 3.15 and 3.16.
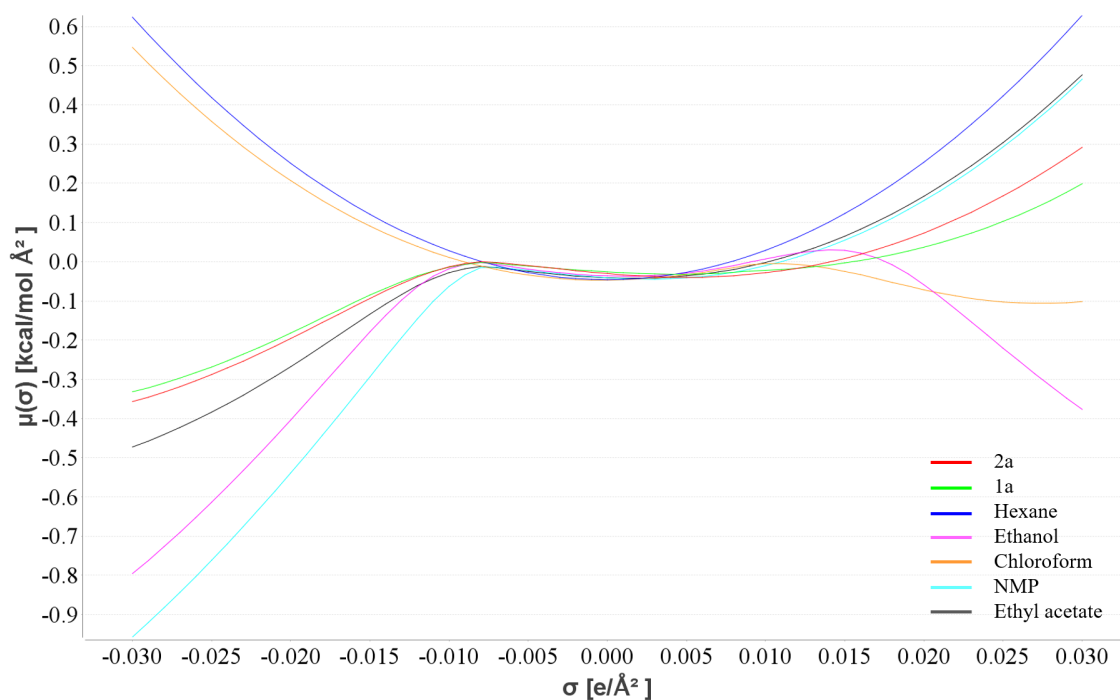
Figure 3.14: σ-Potential of formate ester lactones among known solvents.
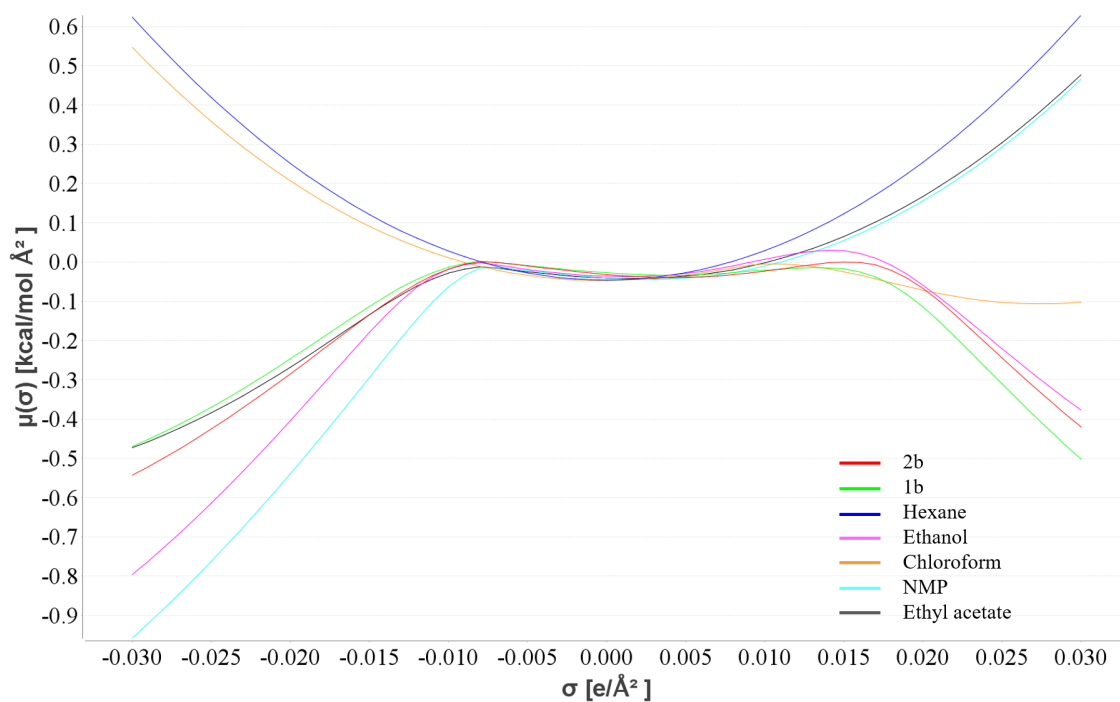
Figure 3.15: σ-Potential of hydroxymethyl lactones among known solvents.
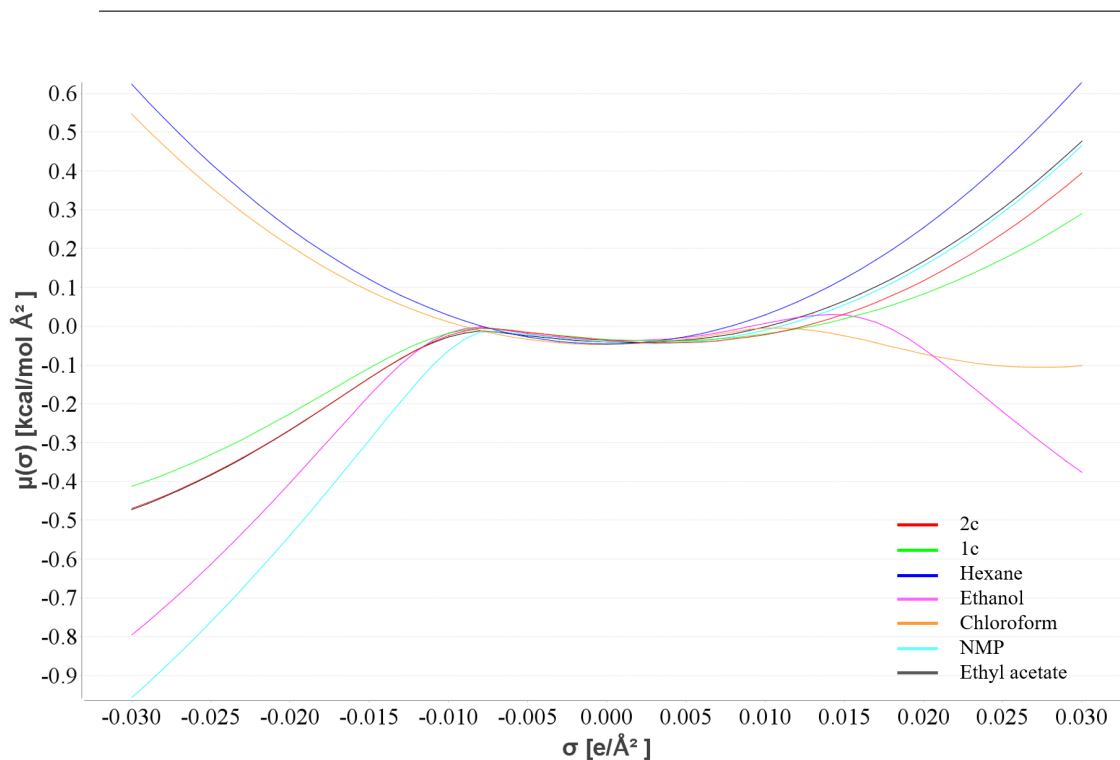
Figure 3.16: σ-Potential of methoxy lactones among known solvents.

As explained in section 1.2.2, the energetically less favourable reactions are represented by an upturned curve. For example, hexane is represented as a U-shape because it is a hydrocarbon, therefore contact with a significantly positive or negative charge density would be disadvantageous. Lactones **1b** and **2b** in figure 3.15 are unsurprisingly the opposite, since they can form energetically favourable interactions with both positive and negative charge densities. All lactones **1a**, **2a**, **1c** and **2c** have the same trend as NMP, i.e. will interact favourably with a positive charge density and not with a negative charge density. However, the curve is much more pronounced for NMP, signifying the particular characteristics of dipolar aprotics that are so hard to replicate.

3.4 Accelerating and Scaling Up the Baeyer-Villiger Reaction

After predicting the areas of solvent space covered by lactones **1a**, **2a**, **1b**, **2b**, **1c**, and **2c** described in section 3.3, the intention was to pursue greener synthetic methods over the mCPBA reagent process described in section 3.2.1 and build on the literature methods described in section 3.2.4. The syntheses preempt the ultimate

goal of testing the value of these substances as solvents and comparing their predicted properties to those experimentally determined.

3.4.1 Investigation of L-Lysine Method

25 °C, 1-6 day tests

A set of reactions were run using the L-lysine catalysed Baeyer-Villiger oxidation method described in section 3.2.5 (see scheme 3.8), this time at 25 °C rather than room temperature to ensure comparability between each run. The reactions were run over a period of one to six days and analysed by NMR spectroscopy to assess the changes occurring over time. The results showed that hydrolysis of the formate ester of lactone **1a** was slow and so a mixture of lactones **1a** and **1b** was observed during the early stages. An example proton NMR spectrum is given in figure 3.17.

It is clear from the proton NMR that there was still unhydrolysed formate ester present. After two days, it was present at about 15%. The selectivity of lactone **1b** was ascertained at the end of each reaction by comparing the integrals of signals 1d and 2d, as shown on the spectrum. The selectivities are given in figure 3.18.

Each reaction, with the exception of the four day reaction, was run three times, with close agreement of the results, as evident in the graph. The proportion of **1b** appears to reach a maximum between three and four days at 25 °C and there is virtually no change between four and six days. The selectivity plateaued at 97-98%, suggesting that the hydrolysis had reached equilibrium.

Scale-up

One of the main objectives of this project was to create a significant amount of the target material, since it was ultimately intended for use as a solvent. As the L-lysine catalysed reaction had proven to work effectively, a trial scale-up was undertaken. Initially, the reaction was scaled up by three times quantitatively. The approximate yields of each repeat are given in table 3.3.

It was unclear why the yield varied so much considering that the same method was applied to each reaction. It was hypothesised that the lactone could form an azeotrope with water, causing some to be lost during rotary evaporation. To
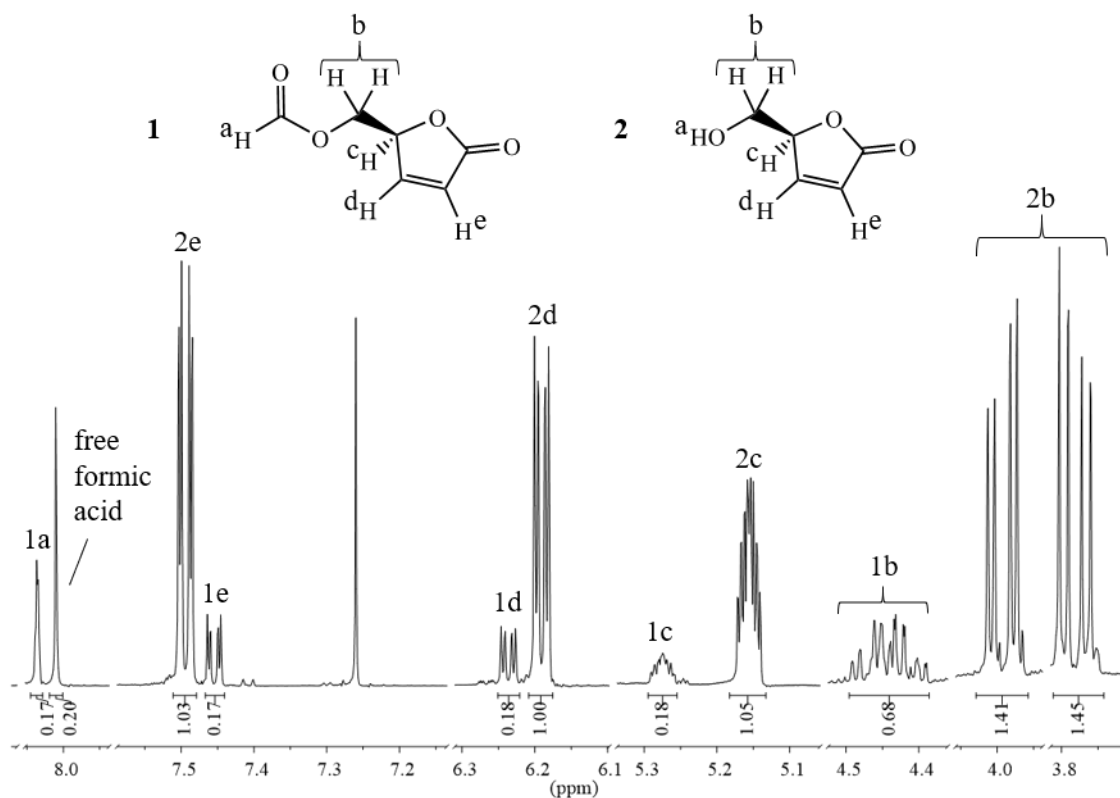
Figure 3.17: Proton (400 MHz) NMR spectrum of product from LGO oxidation to lactone **1b** (**2** above). Run in CDCl$_3$. By assignment of the proton signals, lactone **1a** (**1** above) was confirmed as a secondary product and, by integration, the selectivity was ascertained. It should be noted that the integration is unexpectedly high in the upfield region due to an underlying broad -OH signal. The full spectrum is provided in appendix XVII.

Table 3.3: Results of scale-up reactions run at 25 °C.

Reaction	Mass LGO /g	Reaction time /days	% Conversion[a]
1	3.8	6	71%
2	3.8	6	96%
3	3.8	6	73%
4[b]	7.6	5	93%

[a]Approximate: accounts for remaining acetonitrile by ^{1}H NMR peak integration. Minor impurities not accounted for.

[b]H$_2$O$_2$ entered in 1 ml aliquots. After each addition, temperature was monitored for 30 seconds to check for an exotherm. The temperature remained constant at 25 °C.
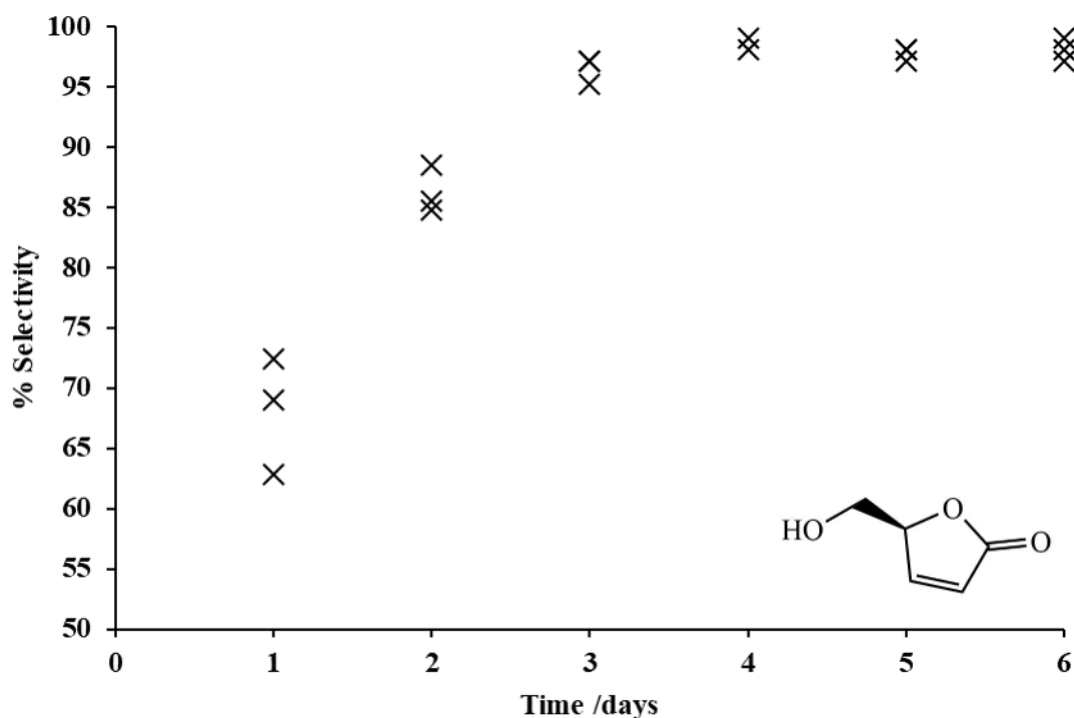
Figure 3.18: Increasing selectivity towards lactone **1b** between 1 and 6 days during the Baeyer-Villiger oxidation of LGO with H_2O_2 and L-lysine catalyst.

determine whether there was any validity in this, a sample of the waste solvent after rotary evaporation was tested by proton NMR spectroscopy. This showed no presence of **1b** and so it was assumed to be an insignificant loss. At this stage, the repeatability remained a concern, since the method contained an undesirable step of trituration with acetonitrile, which caused a sticky residue that retained a portion of the product. However, optimisation was not pursued until later, as purification of lactone **1b** was the primary concern.

3.4.2 Determination of Side-Products of L-Lysine Catalysed Reaction

During the course of this research project, some related work was undertaken by a Master's student under the supervision of the author. This section is a summary of the most relevant results obtained by Martyna Kundrotaite [Kundrotaite 2016]. This research higlights the complexity of the L-lysine catalysed Baeyer-Villiger oxidation detailed in section 3.4 and motivated the move to simplify the procedure, described

in section 3.6.

While probing the Baeyer-Villiger oxidation of LGO by altering the conditions, including temperature and catalyst, three side-products were identified and reported. The structures of these are given in figure 3.2.

(a)   (b)   (c)
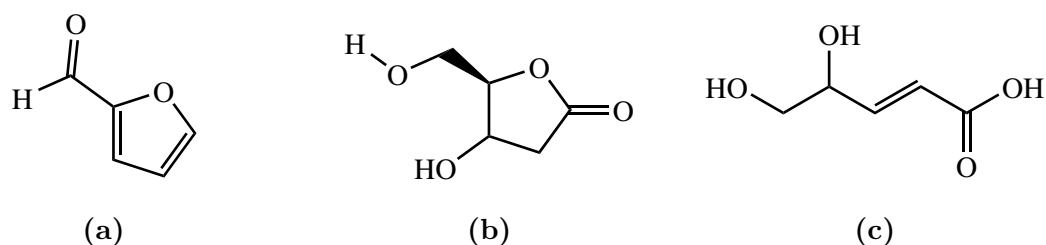
Figure 3.19: Stuctures of side-products reportedly emerging from the amino acid-catalysed Baeyer-Villiger oxidation of LGO: (a) furfural, (b) 2-deoxy-D-ribono-1,4-lactone, and (c) (S)-4,5-dihydroxy-(2E)-pentenoic acid.

Hydrolysis products **(b)**, 2-deoxy-D-ribono-1,4-lactone, and **(c)**, (S)-4,5-dihydroxy-(2E)-pentenoic acid, are likely outcomes, due to the aqueous conditions. The Baeyer-Villiger oxidation of LGO was tested with two different amino acid catalysts, L-lysine, glycine and amine, n-butylamine. The intention was to test whether the reaction is significantly affected by the extra amine group on L-lysine.

Close selectivities of the desired product were an indication that glycine should be pursued in future work as a simpler, less bulky alternative to L-lysine.
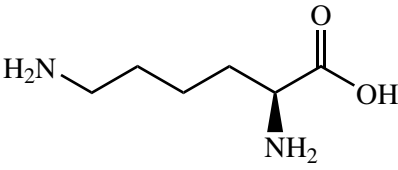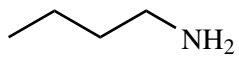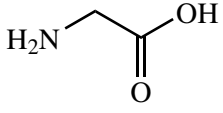
In addition to the alternative catalyst reactions, a set of experiments were run at varying temperatures. The relative selectivities of each compound are shown in figure 3.20.

Clearly, the L-lysine promoted reaction works best when running at 40 °C or below. Above this and the selectivity tends towards the ring-opened product (C). These results were used to inform the continuation of L-lysine catalysed Baeyer-Villiger oxidations that follow, carried out by the author.

Solvent-Free Reaction

A "solvent-free" reaction was run using L-lysine and H_2O_2 (30 wt%). The reaction was not free of water due to the presence of aqueous peroxide, but no excess water was added. The reaction was run with 1.26 g levoglucosenone, as in the earlier

Table 3.4: Selectivities arising from the Baeyer-Villiger oxidation of LGO with three different catalysts.

Reaction[a]	Catalyst[b]	Selectivity			
		1b	**(a)**	**(b)**	**(c)**
1	L-Lysine	57%	0%	3%	30%
2	n-Butylamine	44%	0%	3%	36%
3	Glycine	56%	7%	0%	36%

[a]15.8 mL H_2O, 3 equivalents of H_2O_2, 50 °C, 24 h.
[b]0.1 equivalents.

L-lysine catalysed reactions. It was run for 24 h, then quenched with PtAl. Proton NMR spectroscopy showed that a selectivity of nearly 93% of lactone **1b** was achieved over lactone **1a**, considerably higher than the 63-72% achieved by the one day reaction run in water. An exotherm of 10 °C was observed on addition of the peroxide.

The reaction was subsequently scaled up to 3.77 g with no extra water and achieved the same yield.

3.4.3   Catalyst-Free Baeyer-Villiger Oxidation

Preliminary test on LGO (1)

Baeyer-Villiger oxidations were tested without L-lysine to ascertain whether the reaction would proceed with no catalyst. A reaction was run on the gram scale at 25 °C for 6 days, to be comparable with the most effective L-lysine reactions. The
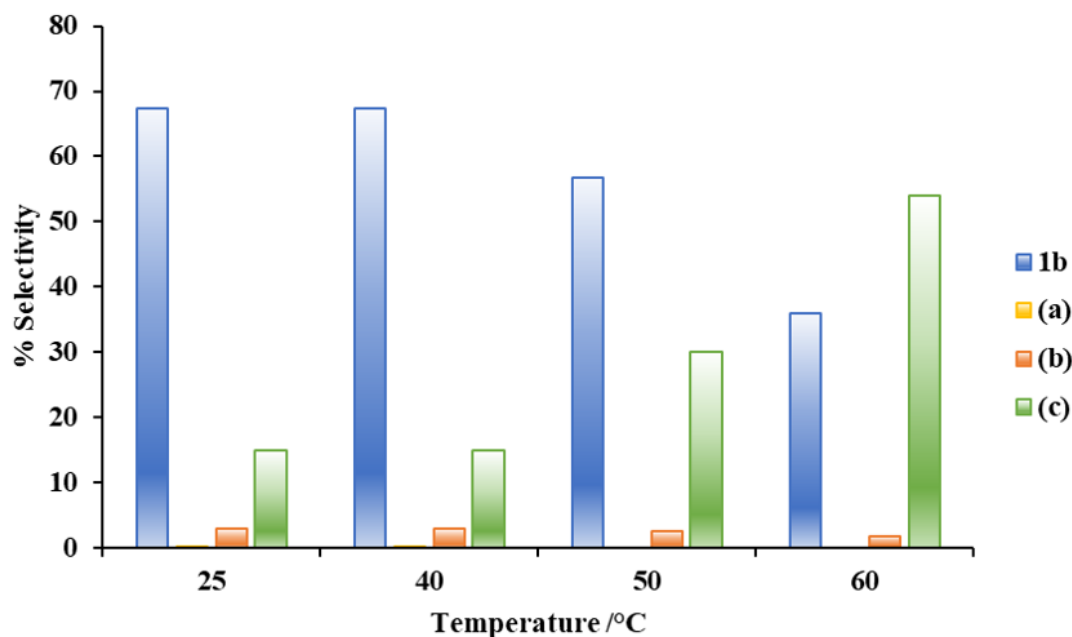
Figure 3.20: Effect of temperature on L-lysine catalysed Baeyer-Villiger oxidation of LGO. (a) furfural, (b) 2-deoxy-D-ribono-1,4-lactone, and (c) (S)-4,5-dihydroxy-(2E)-pentenoic acid. Conditions: L-lysine (0.1 equiv.), H_2O_2 (30%, 3 equiv.), water (15.8 mL), 24 h.

conditions are described in scheme 3.9.
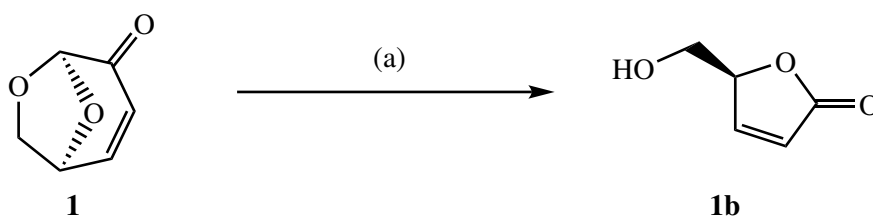
Scheme 3.9: Conditions (a): levoglucosenone (10 mmol) was dissolved in water (16 mL). H_2O_2 (30 wt%, 3 equiv.) added dropwise and stirred at RT, 6 days.

As such, all other parameters were kept the same, including trituration with acetonitrile, although this proved to be unnecessary as the remaining components were fully soluble in acetonitrile when no L-lysine was present. 3 equivalents of H_2O_2 were used as in previous reactions. On analysis, the L-lysine-free reaction appeared to have gone ahead without difficulty. Figure 3.21 shows how the proton NMR spectrum of the crude product compared to LGO, **1**.

Full conversion of **1** was observed, along with the presence of minor product **1a**.
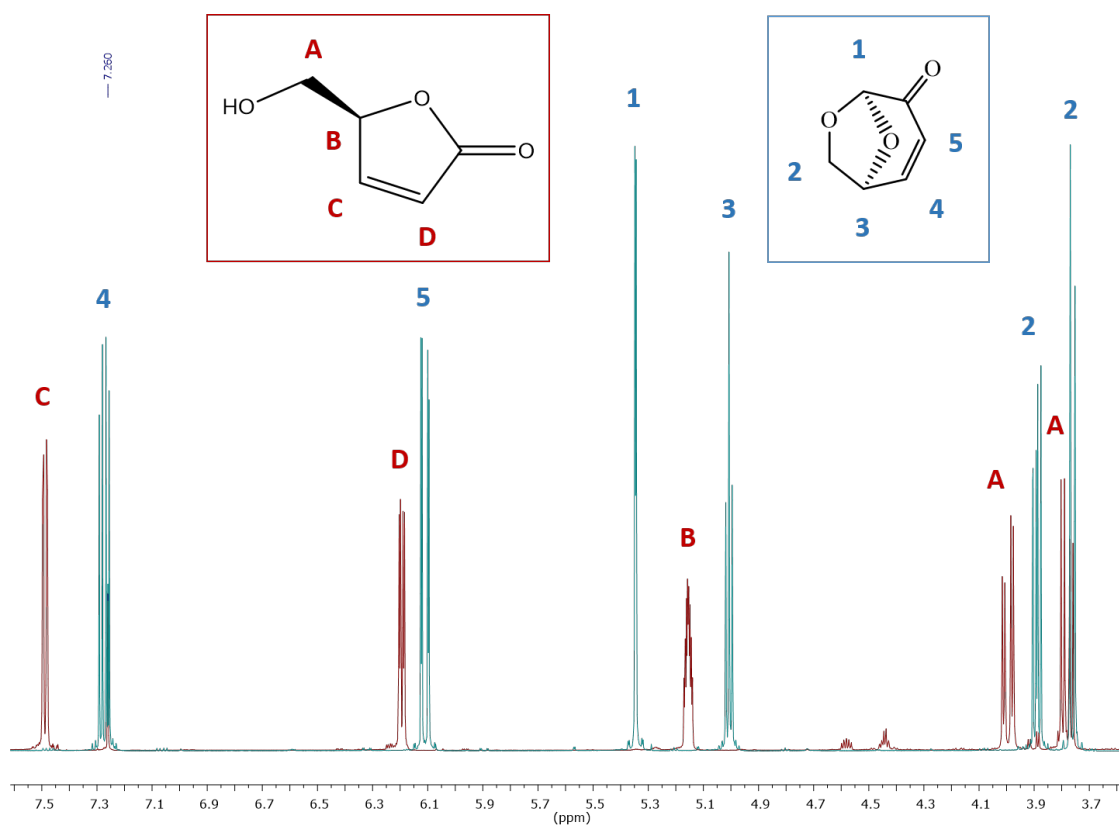
Figure 3.21: ^{1}H NMR of product, **1b** (red) compared to levoglucosenone, **1** (blue). Expanded for clarity. A broad -OH signal is also observed at δ 2.7 ppm.

Although a calibraton could not be obtained without isolation of the side products, a yield of approximately 65% was determined by GC analysis, assuming equivalent responses and accounting for the mass of the crude product.

Literature method

Subsequent to the running of this initial test, a communication was published by another research group detailing the observation that the Baeyer-Villiger oxidation of LGO and Cyrene can proceed using only H_2O_2 (30%) and no additional catalyst [Bonneau 2018]. In this method, H_2O_2 (30%) was added dropwise to a slight excess of **1** in water, over an ice-bath, and under a N_2 atmosphere. Dropwise addition was necessary to avoid a severe exotherm and to minimise the amount of Michael-addition in the case of LGO. On a kilogram scale, the dropwise addition took approximately 4 h. After addition, the mixture was heated to 50 °C for a further 6 h. The reason for increasing the temperature can be attributed to speeding up hydrolysis of the formate ester functional group of **1a**. In their optimized kilo-scale procedure, the authors achieved a 71% yield of **1b** and 72% of **2b**, but also demonstrated that conversion could be pushed to 100% with 2.45 or more equivalents of H_2O_2 within 24 h. However, this left a large excess of peroxide to be quenched at the end of the reaction, which created an additional hazard during the workup, not to mention further waste in the form of sodium sulfite, sodium bisulfite, metabisulfite or catalase which were used as quenching agents. Although the reaction worked well solvent-free (i.e. no added water besides the peroxide solution), Cyrene oxidation, in particular, proved to be highly exothermic and so water was used as a solvent to help dissipate heat.

Application to Cyrene (2)

In aiming to progress research into either lactone **1b** or **2b** as a possible solvent, or indeed their O-methylated counterparts, it was necessary to produce material at a much quicker rate. As such, the L-lysine-free method was prepared for scale-up using an amalgamation of the previously described work and the literature method, producing a less wasteful and cleaner way to increase the size of the reaction. The reaction was performed on a scale of 3 g Cyrene 0 °C and subsequently heated to 50

°C and a N$_2$ atmosphere was used. Preliminary GC-MS and proton NMR suggest that the expected product (**2**) had formed along with a secondary product.

This reaction proved to be highly expthermic and, due to a sudden and violent eruption, a third reaction was run, using a flow of nitrogen from a cylinder, with a bubbler to release pressure. As before, the peroxide was added drop-wise at 0 °C controlled by an ice-bath. The heating phase was monitored closely and filmed to judge the severity of the expansion. At 46 °C mantle temperature, the reaction mixture suddenly erupted and pushed up through the bubbler, as pictured in figure 3.22.

Figure 3.22: The moment of eruption during the Baeyer-Villiger oxidation using H$_2$O$_2$ (30%) and water only. H$_2$O$_2$(aq) was initially added at 0 °C and then heated. When the mantle temperature reached 46 °C, a sudden and violent expansion occurred. Arrow: the cloudy reaction mixture can be seen making its way through the bubbler tubing.

Due to the violence of the eruption, it was clear that the reaction posed a risk to safety due to to the rapid exotherm. The most logical explanation for this occurrence is that the H$_2$O$_2$ was allowed to build up during the addition stage due to the cold temperature. As the temperature was raised, this initiated a runaway reaction that could not be contained within the volume of the reaction vessel. The danger of this exotherm was a large concern with respect to increasing the scale

of the reaction. To mitigate the risks associated with the exotherm, the reaction was repeated with the addition of peroxide occurring at 50 °C rather than 0 °C and the exotherm was monitored, not allowing an increase of more than 10 °C above the starting temperature. Running the reaction at 50 °C meant that the initial oxidation step occurred quickly during addition so there was no significant build up of peroxide which permitted confidence that the concentration of peroxide would never be dangerously high. This allowed the oxidation to happen slowly and no further expansion problems were observed. A slight darkening of the reaction mixture when the initial heating step was used, as shown in figure 3.23, suggesting slight decomposition, but there was no significant difference observed in the product on analysis.
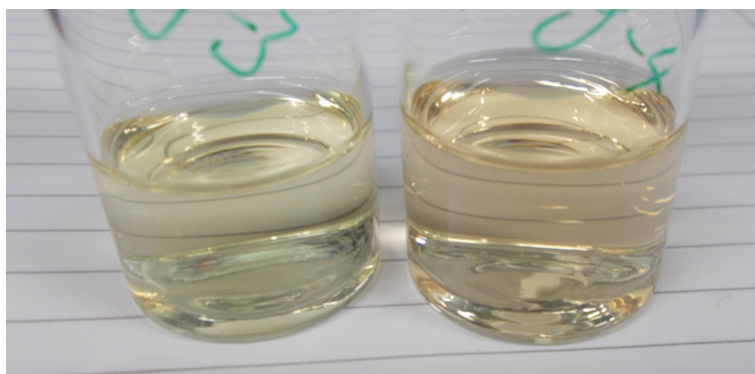
Figure 3.23: Colour of reaction mixtures with addition of H_2O_2 at 0 °C (left) and at 50 °C (right).

Both GC and proton NMR analysis showed that on a small scale, these conditions produced full conversion of Cyrene and yielded compound **2b** at 99%, both using the literature method and with initial heating. Even a small-scale test run at 25 °C for 6 h achieved a 98% yield, although it should be noted that the initial temperature of this reaction peaked at 55 °C due to the strong exotherm.

3.4.4 Scale-up

Having proven the more manageable nature of the pre-heated oxidation reaction on Cyrene, **2**, a set of 3 scale-up reactions were carried out. Sequentially, 17 g, 34 g, and 175 g of starting material were used to perform the oxidations and, in each case, 1.2 molar equivalents of H_2O_2 were used to drive the reaction to completion with

the knowledge that the excess peroxide could be quenched safely with platinum on alumina pellets. Table 3.5 provides a comparison of each reaction condition.

Table 3.5: Comparison of scale-up reaction conditions.

Reaction no.	1	2	3
Mass Cyrene /g	17	34	175
Volume H_2O /mL	20	40	200
H_2O_2 equiv. (molar)	1.2	1.2	1.2
$T_{(initial)}$ /°C	50	50	20
$T_{(max.)}$ /°C	60	60	70
H_2O_2 addition time /h	2	3	7
% Selectivity[a]	-	93	95

[a]Determined by GC peak area assuming equivalent responses to FID. Each reaction achieved full conversion of Cyrene.

Running the Baeyer-Villiger oxidation on the 17 g and 34 g scale compared well with the 3 g scale reactions in terms of manageability. Total addition time of the peroxide was increased on the larger scale due to the added waiting time for heat to dissipate during the exothermic stage. This was expected because of the increased ratio of volume to surface area, since the concentration remained the same.

Similarly, the 175 g scale reaction exhibited more variance in temperature during the exothermic stage as the larger volume caused a delay in the temperature spike observed on addition of each aliquot of peroxide. The reaction mixture was also more concentrated due to the need to keep it on a lab scale. This reaction was run in a 2 L flask using a heating mantle rather than a DrySyn, which caused difficulty in removing the heat from the reaction as the exotherm began to rise. Lowering the mantle away from the flask also impeded stirring, which was liable to cause hotspots in the reaction. To mitigate this problem, the reaction was not heated initially. Instead, the natural exotherm was allowed to increase until the reaction mixture reached 50 °C and, henceforth, kept within a 10 °C window.

Distillation

Earlier in the chapter, difficulties were described when trying to purify lactone **2b** by Kugelrohr distillation. It was clear that a high vacuum would need to be sustained in order to distil the product without significant degradation. Once the reaction was run on a larger scale, there was enough material to set up a static distillation system, which was more robust and less likely to leak due to rotating joints. Additionally, an Edwards V oil pump was used to draw a stronger vacuum, reducing the necessary heat input. Due to the problematic separation, a 36 cm Vigreaux column was used to lengthen the path of distillation. The product consistently distilled at 104 °C under a pressure of 0.3 mbar, resulting in approximately 129 g of distilled product.
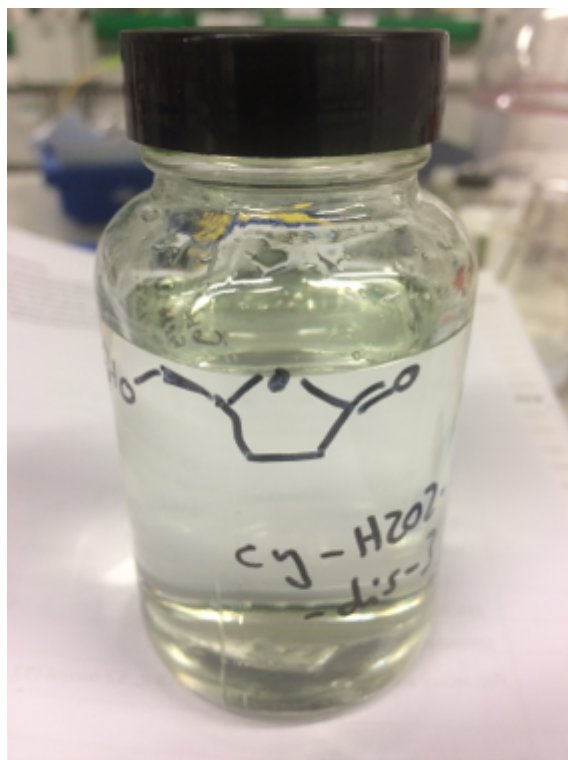
Figure 3.24: Large batch of lactone **2b** after distillation.

It was not possible to completely purify lactone **2b**, even during a very slow distillation, due to the apparent proximity in boiling point and vapour pressure of a side-product in the mixture. However, the quality of the product was significantly better than was observed in products of the mCPBA and L-lysine reactions, as is illustrated by the clarity of the -CH_2 signal at 2.55 ppm in the NMR.

This previously unresolved signal further confirms the identity of lactone **2b**, since two ddd signals can clearly be seen slightly overlapping, with pronounced
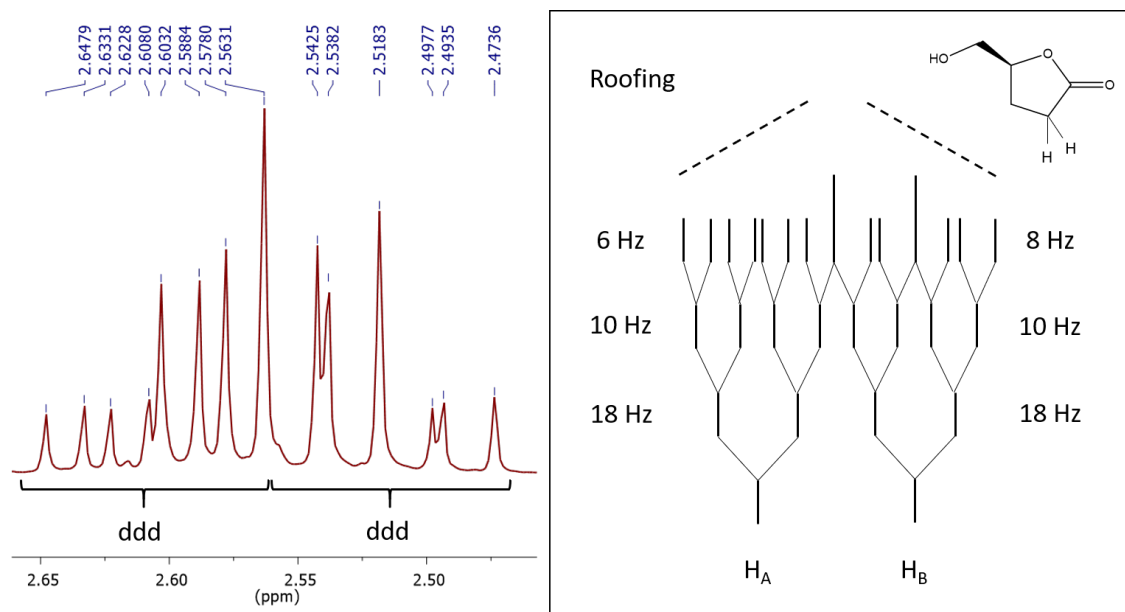
Figure 3.25: ^{1}H NMR signal and splitting tree of the CH$_2$ protons on the alpha carbon. Sample run in CDCl$_3$ on a 400 MHz spectrometer.

roofing. This is consistent with two non-equivalent CH$_2$ proton environments on the 5-membered ring, coupling to each other (18 Hz) and to two protons on the adjacent carbon.

A minor product was consistently seen throughout these reactions, which co-distilled with lactone **2b**. From the smaller signals present in the ^{1}H NMR, this product appeared to align with the signals expected from formate ester **2a**, as shown in figure 3.26.

In an attempt to fully hydrolyse the residual formate ester, a 2 mL aliquot was removed from the Baeyer-Villiger reaction after 17 h and dried by rotary evaporation. A portion of the residue was then dissolved in D$_2$O and heated on a hotplate at 50 °C for 2.5 h, to ascertain whether altering the equilibrium would encourage further hydrolysis. The result was a minimal reduction in the integration of the pendant CH$_2$ signals at δ 4.34 and 4.17, from 3% to 2%. However, new CH$_2$ signals appeared at δ 3.44 and 3.33, which have not yet been identified. Given the conditions of heating in aqueous solution and that hydrolysis of the formate ester happens fairly readily, it is surprising that this side-product remained even after a distillation, which required some considerable heating. One speculative reason for this could be that these signals represent a 6-membered ring formate ester lactone, **22**, which is
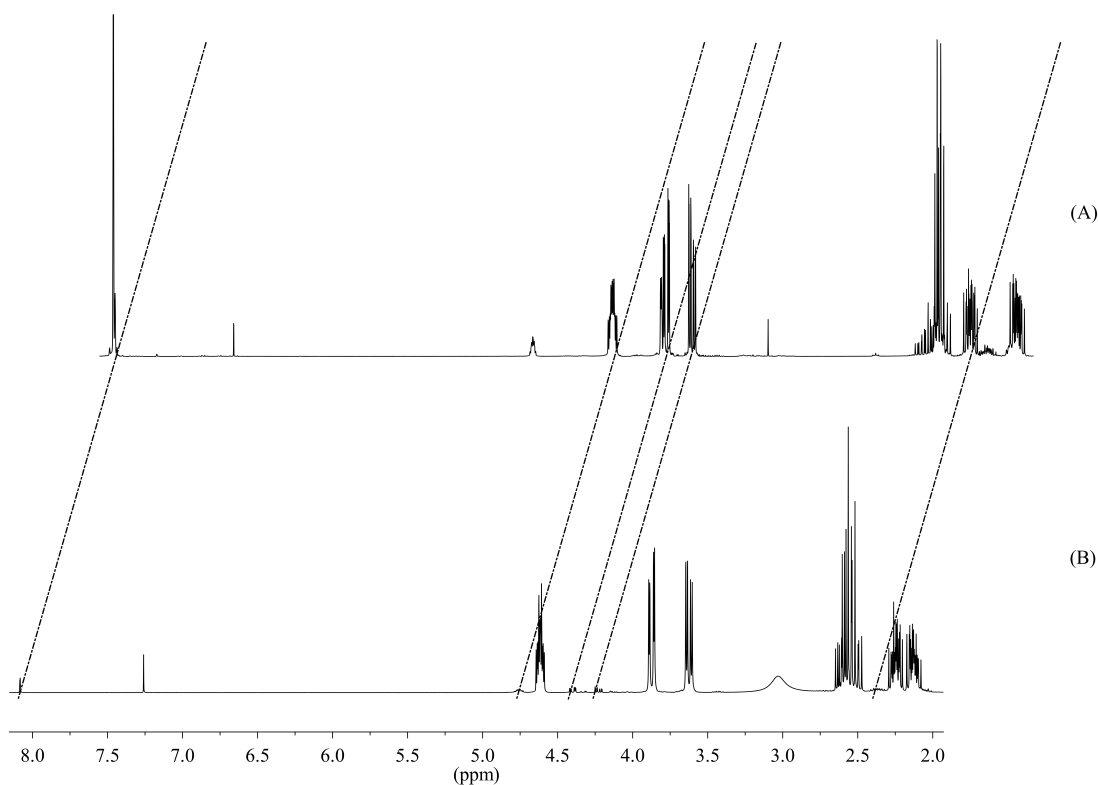
Figure 3.26: Proton (400 MHz) NMR spectrum for **2a** (A) stacked on **2b** (B). Both run in CDCl$_3$. Dashed lines indicate presence of **2a** signals in distilled **2b**

much less favoured by the Baeyer-Villiger reaction. It is unclear at this stage why such a structure should be more resistant to hydrolysis. The structure of **22** is given in figure 3.27.
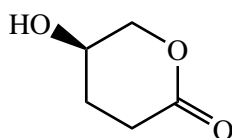
Figure 3.27: Structure of speculative contaminant **22**.

As both the **2a** and **22** are isomers of the Baeyer-Villiger product **2**, both have a relative molar mass of 144 g mol^{-1}. The use of electrospray ionisation (ESI) mass spectrometry confirmed that the molecular ion peak of **2a** or **22** was present in the mixture.

Field ionisation (FI) used on the same sample showed the presence of a product ion peak at m/z 233. As the mass of the major product is 116 g mol^{-1}, this peak suggested the presence of a dimer. If dimerisation occurred, it is unlikely that this happened prior to GC-MS since the two signals observed in the chromatogram eluted

very close together (within two minutes). A speculative answer for the presence of a dimer in the spectrum is that one of the more acidic alpha-protons was dislodged in the collision cell and followed by a Claisen condensation, as shown in scheme 3.10. The analagous aldol condensation product of Cyrene was shown in section 3.2c [Wilson 2016].
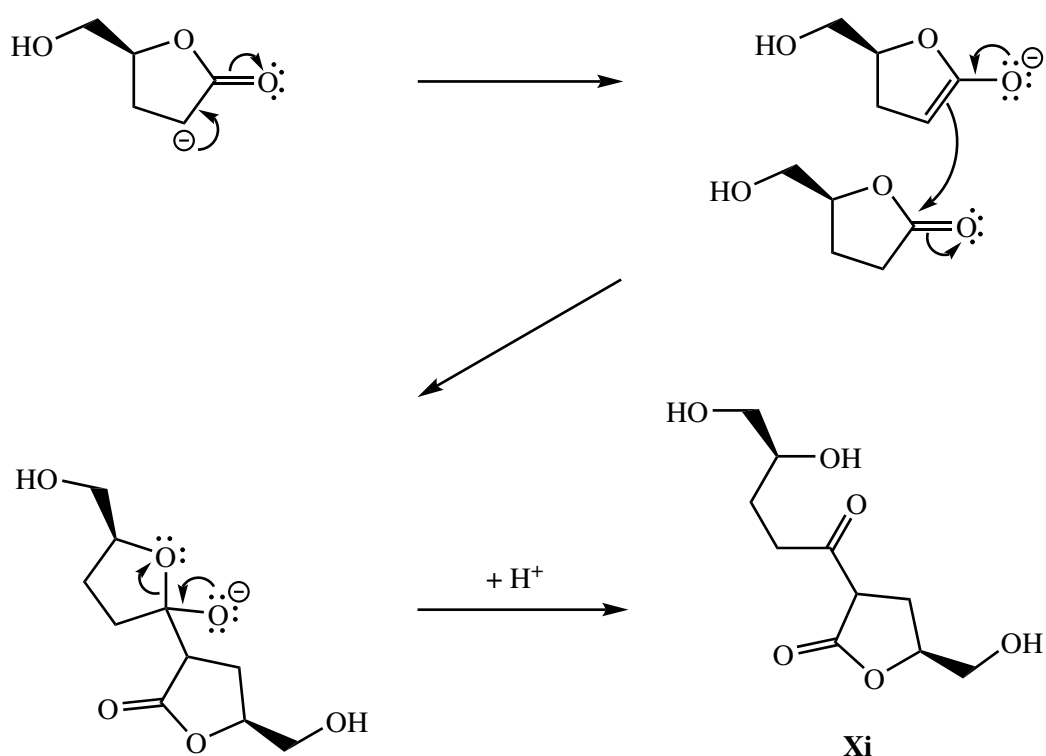
Scheme 3.10: Proposed Claisen reaction mechanism of **2b**

In order to ensure that separation could not be achieved by any of the available means of distillation, a small sample of the distillate was submitted for further purification by revisiting Kugelrohr distillation. The Sigma-Aldrich nomograph was used to predict the required temperature by inputting the previous parameters (104 °C at 0.3 mbar). As the pressure gauge for the Kugelrohr read 0.9 mbar, the product was expected to distil at 163.8 °C. The temperature was taken to 130 °C and raised in 5 °C increments until 160 °C, after which 1 °C increases were used. A colourless liquid began to condense in the cold bulb once the oven temperature was at 164 °C, fitting well with the prediction. Since the standard vacuum distillation and Kugelrohr distillation were run using different vacuum pump/gauge systems, this added validity to the pressure and temperature readings recorded for both systems.

3.5 Hydrogenation of Unsaturated Lactone

The primary concern regarding both levoglucosenone and each of the unsaturated lactones, **1a**, **1b**, and **1c** was the reactivity of the α, β-unsaturated C=C bond. By removing the likelihood of Michael-addition to the β-carbon and improving the stability of the molecules, it was expected that reducing this double-bond would improve the structures usefulness as solvents.

3.5.1 Initial Tests

Method

The combined crude products of lactone **1b** from L-lysine catalysed Baeyer-Villiger oxidation (section 3.4.1) were used in several small scale hydrogenation reactions as shown in scheme 3.11.
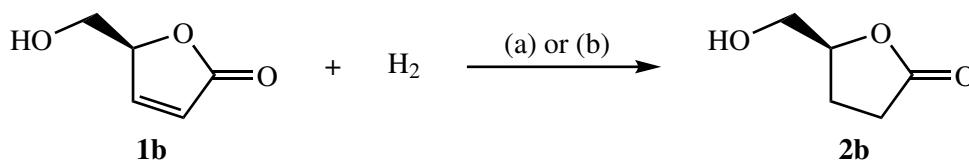
Scheme 3.11: (a)**1b** (0.6 mmol), Pd/C (5 wt%, 0.05 g), EtOH (4 mL), H_2 (70 bar), RT, 18 h; (b) **1b** (30 mmol), Pd/c (5 w%, 0.05 g), H_2 (70 bar), RT, 18 h.

Crude **1b** was used as opposed to purified lactone as it was hoped that the hydrogenation would go to completion regardless, removing the need for an extra, potentially energy intensive step such as distillation, or wasteful step, in the case of column chromatography. As lactone **1b** is a liquid, it was thought that the reaction might be successful with no additional solvent, thus improving the green metrics of the reaction. In order to test this, three repeats were run each of hydrogenations both without solvent, and including an ethanol solvent. The hydrogenations were carried out in a 6 chamber pressurised reactor, shown in figure 3.28, under a 70 bar atmosphere of hydrogen.

A palladium on carbon catalyst (5 wt.% Pd/C) was chosen as is common for alkene hydrogenations and has previously been shown to work in the conversion of LGO to Cyrene [Sherwood 2014]. Each reaction mixture was filtered through celite
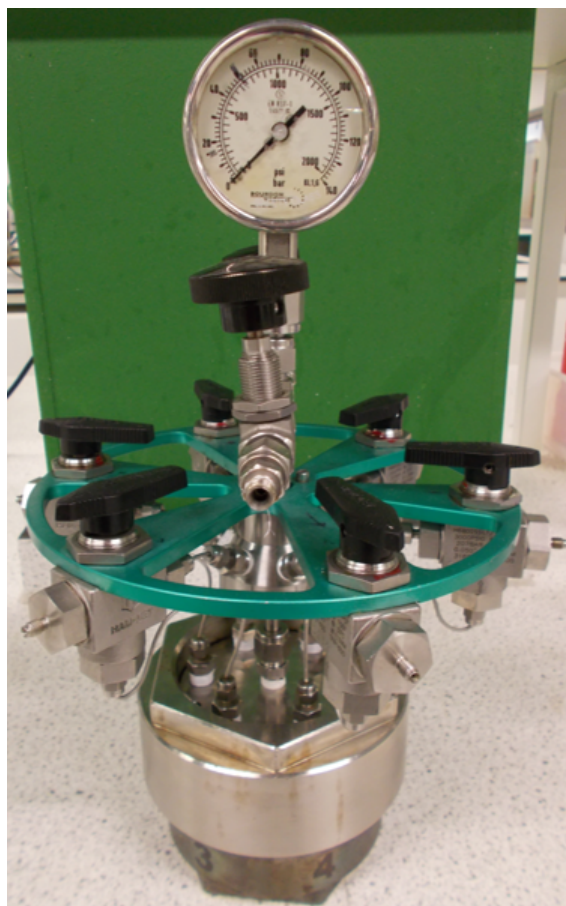
Figure 3.28: Autoclave reactor with 6 x 10 mL chambers and gas inlet used for hydrogenation.

to remove the Pd/C catalyst before using GC to determine the conversion of lactone **1b** to **2b**.

Results and discussion

6 small-scale hydrogenations were run simultaneously in the 6-point autoclave shown in figure 3.28. The reactor was charge with hydrogen, first expelling air, then increasing the pressure to 70 bar. Three repeats of the solvent-free reaction were run with a larger quantity of lactone **1b** since the chambers were each 10 mL in volume and it was more practical for efficient stirring to keep the chambers approximately half full. The reactions dissolved in ethanol were limited by the size of the chamber as a head-space needed to be left for the H_2 gas.

There were some practical problems with this first run as the hydrogen pressure dropped further to 5 bar in chambers 1 and 5 than the rest of the chambers, which

Table 3.6: Initial hydrogenation data.

Reaction[a]	Solvent	Mass **1b**	Final H_2 pressure /bar	Conversion[b]
1	EtOH	0.72	20	50%
2	EtOH	0.73	5	97%
3	EtOH	0.73	20	98%
4	None	3.48	20	42%
5	None	3.49	20	26%
6	None	3.50	5	38%

[a]All run using 5 wt% Pd/C catalyst (0.05 g), for 18 h at RT.
[b]Calculated by integration of GC peaks.

only dropped to 20 bar from the original 70 bar. It was attested that these two chambers had leaked and the significant decrease in conversion shown in the table above supports this assertion. Besides this set-back, what was clear was that the reactions run in ethanol performed much better than those with no solvent. It was speculated that the high viscosity of the starting material caused a slow rate of mixing and diffusion of H_2. Furthermore, the palladium on carbon catalyst had to be filtered off after the reaction, which proved to be very difficult when the neat lactone was used. Therefore, it was concluded that solvent-free hydrogenation of **1b** was not an ideal method.

3.5.2   Separation

Kugelrohr

In order to increase the conversion to product **2b**, the crude products from the previous reactions were combined and hydrogenated again using the proven ethanol-based method. An additional step of centrifugation was added prior to filtration, which greatly eased removal of the catalyst. As stated in previous sections, the ultimate goal was to produce lactone **2b** on a larger scale, so that it could be tested as a solvent. As such, distillation was the preferred method of separation, so that it could be achieved quickly and in large volumes. Kugelrohr short path distillation was initially tested as a small volume separation technique on the crude lactone, post-

hydrogenation. The earliest attempts had limited success due to a poor vacuum caused by inefficient pumps and leaks in the system. It was evident that, due to the constant rotation of the equipment, the system struggled to hold a low pressure and tended to increase over time. The result was that, at an oven temperature of 160 °C and above, crude **2b** turned brown/black and became increasingly viscous, but no distillate was observed.

On improvement of the vacuum strength down to 3 mbar, a colourless condensate was observed in the second Kugelrohr bulb. Unfortunately, the pressure could not be stabilised during this distillation and further product could not be isolated. However, proton NMR analysis of the condensate showed that lactone **2b** had been distilled and produced a much cleaner product than previously observed (NMR spectrum provided in appendix XVIII). Despite the practical difficulties, this result led to confidence that distillation could be optimised leading to purification without degrading the product, a notion that has been described further in section 3.6.

Silica plug

It seemed unlikely that lactone **2b** would spontaneously decompose at the temperatures used for Kugelrohr and more likely that contaminants in the crude product were either decomposing themselves or causing a reaction. To improve the purity of the crude product with relative simplicity, it was absorbed onto silica and washed through several times with a mixture of cyclohexane and ethyl acetate (50:50) to remove particularly polar compounds such as residual L-lysine, which had an R_F value of 0. On re-running of the Kugelrohr distillation, however, the same discolouration was observed.

DMC trituration

Difficulty having been encountered during Kugelrohr distillation, 1 g of crude **2b** was triturated and sonicated in dimethyl carbonate (DMC, 10 mL), then filtered and dried to see if extraction with another green solvent could aid in purification. The extract and residue were analysed by proton NMR spectroscopy and are compared in figure 3.29.

Aside from the ethanol signals at δ 1.2 and 6.1 ppm (A), which were removed
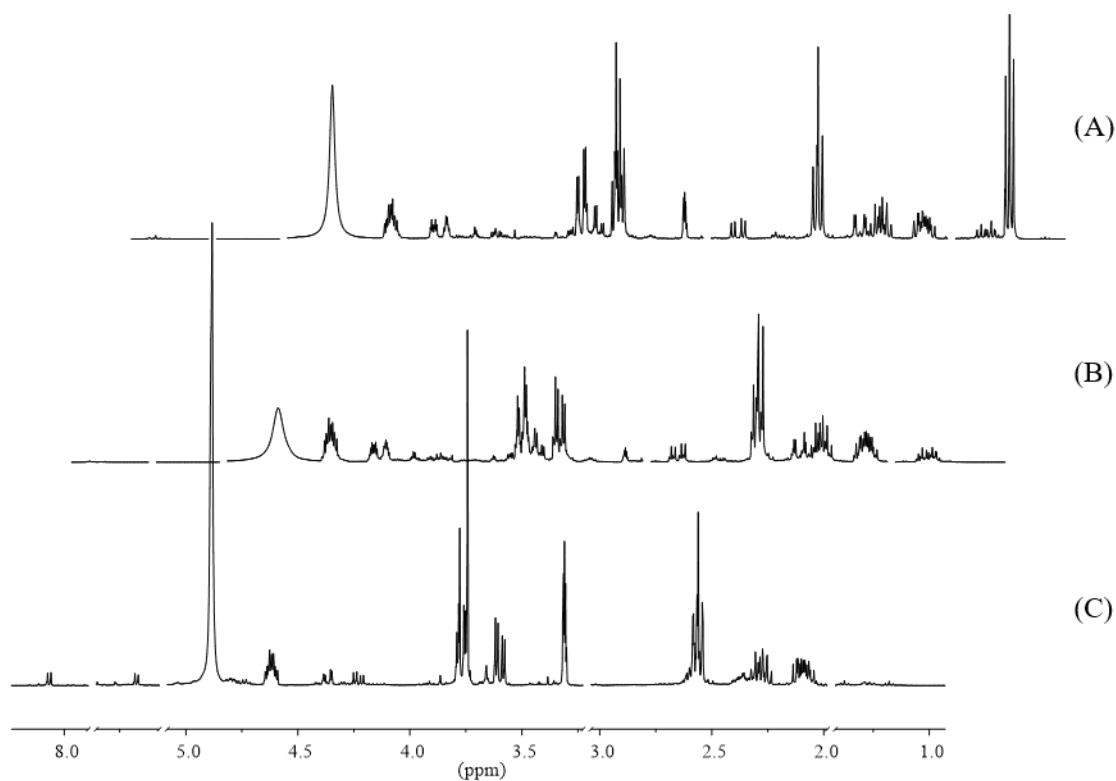
Figure 3.29: Proton NMR spectra run on 400 MHz spectrometer in CD_3OD of: (A) crude product of hydrogenation reaction; (B) product from DMC trituration; (C) residue from DMC trituration.

by drying, trituration with DMC left a slightly cleaner product in the residue (C) removing some of the signals just below 3.0 ppm and at 3.75 ppm (B). Unfortunately, the residue constituted <27% of the mass and there was no definitive separation. Overall, the result was not significant and added to the conclusion that separation problems were occurring due to close equivalencies in polarity between the product and its contaminants.

3.5.3 Scale-up

The hydrogenation method run using ethanol as a solvent was subsequently increased in scale to 15 g lactone, using 0.5 g Pd/C and, again, 70 bar H_2 gas. The conditions for this reaction are shown in scheme 3.12.

The reaction achieved a conversion of approximately 93% (from GC analysis). This was a promising step in the endeavour to produce enough of lactone **2b** for further testing, but required the use of a larger autoclave (see figure 3.30).
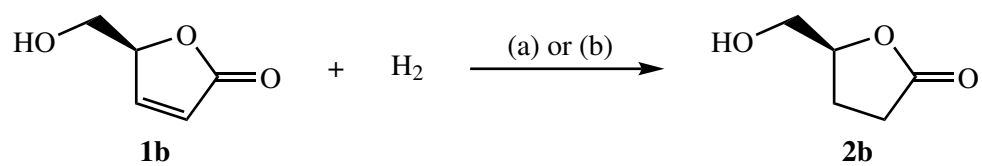
Scheme 3.12: (a)**1b** (0.14 mol), Pd/C (5 wt%, 0.5 g), EtOH (30 mL), H_2 (70 bar), RT, 18 h

Figure 3.30: Larger size autoclave used for scale-up hydrogenation with a maximum safe pressure of 200 bar.

Due to the high pressure safety concerns of this reaction, it was considered whether the same results could be achieved by using a simpler method. This is explored in the following section.

3.5.4   Alternative Hydrogenation Method

A hydrogenation reaction was carried out similarly to that in section 3.5.1, this time using an ethyl acetate solvent and hydrogen gas contained in balloons rather than a cylinder, the setup of which is shown in figure 3.31.

Figure 3.31: A simple set up for the hydrogenation of lactone **1b**.

The reaction was run for 65 h and the balloons were not replaced during that time. A conversion of approximately 87% (determined by GC analysis) was reached during this time. Considering that the conversion was 10-11% lower than that observed during the most successful reactions reported in table 3.6, it could be argued that this method was less successful. It is notable, however, that the balloons applied did not retain pressure throughout the reaction time and, although the reaction may have been aided by a higher pressure due to the decrease in entropy during the addition, it remains true that the reaction was fairly successful with much

less sophisticated equipment. Running any reaction without the need for excessive pressures is also safer.

It was held that this method of hydrogenating lactone **1b** after the Baeyer-Villiger oxidation of LGO was a viable one, but once the simpler method of performing the Baeyer-Villiger oxidation on Cyrene (described in section 3.6) was known, it became more practical to pursue lactone **2b** via that route. This was largely because Cyrene became more readily available from the supplier, Circa, and their hydrogenation process had been better optimised, as could be reasonably expected for a pilot plant.

3.6 Methylation

Lactones **1** and **2b** formed by the Baeyer-Villiger reaction, last described in section , both contain an -OH functional group that contributes towards hydrogen-bonding, but also leaves the structure susceptible to deprotonation by strong bases. In section 3.3, it was predicted that these alcohols would be more than twice as hydrogen-bonding as their methylated counterparts (14.0 $MPa^{1/2}$ compared to 6.6 $MPa^{1/2}$ according to HSPiP), putting them closer in quality to other protic solvents such as butanol. The methylated analogues **1c** and **2c** were positioned much closer to the dipolar aprotic solvents dichloromethane and N-methylpyrrolidone. Considering these were the main targets for replacement due to toxicity concerns (see table 3.1), it was deemed important to pursue the synthesis of the methylated compounds. This was expected to beneficially reduce the reactivity of the pendant chain.

3.6.1 MeI Method

Initial tests

In order to reach the target compound, **2c**, as quickly as possible, an established Williamson Ether synthetic method of methylation was attempted first [Nemoto 1984]. Iodomethane (MeI) was used as the methylating agent and sodium hydride (NaH) to deprotonate the alcohol, as shown in scheme 3.13.

Initially, this was carried out in a THF solvent under air, using a drying tube filled with $CaCl_3$ to prevent excess moisture entering the reaction. It was thought
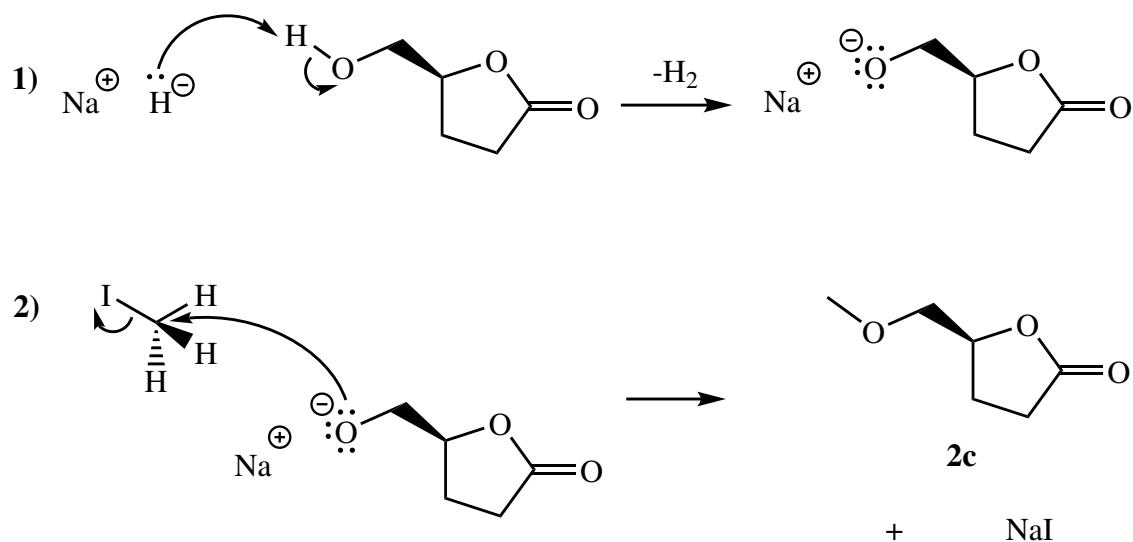
Scheme 3.13: Mechanism for methylation using the following conditions: 1) NaH (60% dispersion in oil, 9 mmol) and DMF (anhydrous, 20 mL) stirred at 0 °C under N_2, to which lactone **2b** (4 mmol) was added dropwise and stirred for 1 h at RT; 2) MeI (12 mmol) added dropwise and stirred for 2 h at RT.

that, if the reaction ran well without resorting to an inert atmosphere, this would demonstrate the robustness of the system and negate the need for extra inputs. The reaction yielded methylated lactone **2c** at 78%.

3.6.2 Scale-up

The reaction was subsequently scaled up to a scale of 10 g, then 25 g, Using a flange flask and overhead stirrer with a N_2 line attached and a gas bubbler as shown in figure 3.32.

It was noted from the 10 g scale reaction that the solid precipitate that formed tended to aggregate and cause difficulty of mixing. For this reason, an overhead stirrer was used in the larger scale reaction and provided much more effective mixing. The product was distilled using a vacuum distillation apparatus with a Vigreux column. With a pressure of 3 mbar, the product distilled at 100 °C, according to the vapour temperature giving a colourless liquid, visibly lower in viscosity to lactone **2b**.
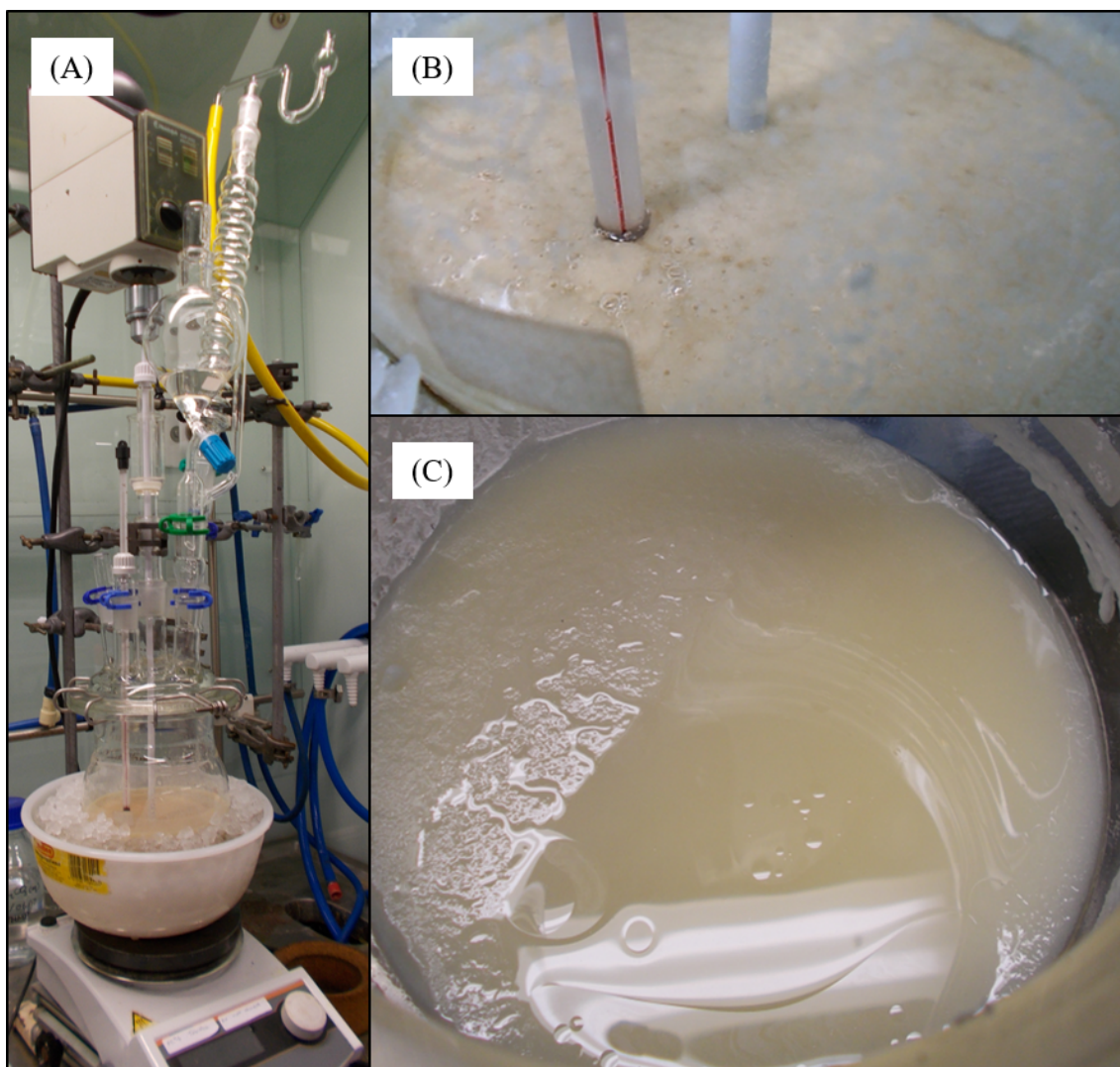
Figure 3.32: (A) Scale-up apparatus for MeI methylation of lactone **2b**; (B) precipitate formed demonstrating need for mechanical stirrer; (C) thick residue of salts leftover.

3.6.3   Greening the Reaction

It is undeniable that the given method for methylating the alcohol has a number of flaws in the eyes of a green chemist. By nature, a methylation is an unpleasant reaction due to the reactivity of the agents used to perform it and their propensity to react in the human body. Iodomethane, while a highly effective methylating agent, is toxic in the short-term and readily absorbed through the skin due to its size. In the long-term, it is a suspected carcinogen. Other risks associated with this method include the reaction of water with sodium hydride, which produces hydrogen, and can cause fire or explosion in the event of a spillage, especially on a large scale. In light of this understanding, the previous method has only been used for proof-of-concept, with the intention to seek out new, more acceptable ways of achieving the same product.
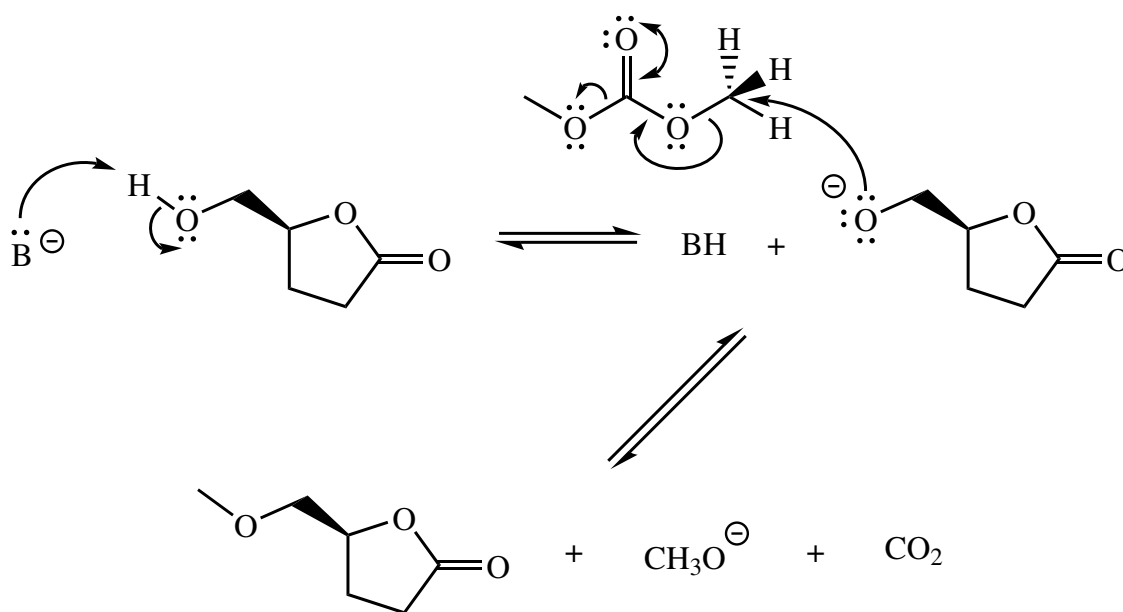
Scheme 3.14: Mechanism for methylation using the following conditions: lactone **2b** (0.4 mmol), dimethylcarbonate (5 mL, in excess) and base (20 mol % of either AlCl$_3$, K$_2$CO$_3$, hydrotalcite, or polymer supported p-toluenesulfonic acid) were heated to 170 °C at autogenous pressure in an autoclave for 15 h.

In an effort to explore ways to achieve this, five small-scale reactions were run concurrently using a more modern method of methylation with dimethyl carbonate (DMC) acting as both reactant and solvent [Tundo 2002]. Four different catalysts were tested: Lewis acid aluminium chloride, Bronsted acid para-toluenesulfonic acid

supported on polymer (*p*-TSA), anion exchange mineral hydrotalcite (htal), and basic potassium carbonate (K$_2$CO$_3$). A catalyst-free reaction was also run to differentiate catalyst-assisted transformations. The reactions were run in a multipoint autoclave, each chamber having a capacity of 10 mL (5 mL safe volume).

To overcome the lower activity of DMC as a methylating agent and to steer the reaction towards methylation of the alcohol, conditions of high temperature (170 °C) and autogenous pressure were used. DMC is a low toxicity, non-volatile solvent. It is necessary to assess the payoff of using less harmful chemicals with the amount of energy needed to perform the reaction.

One of the possible side reactions from using DMC as a methylating agent is the carboxymethylation of the alcohol, as shown in scheme 3.15.
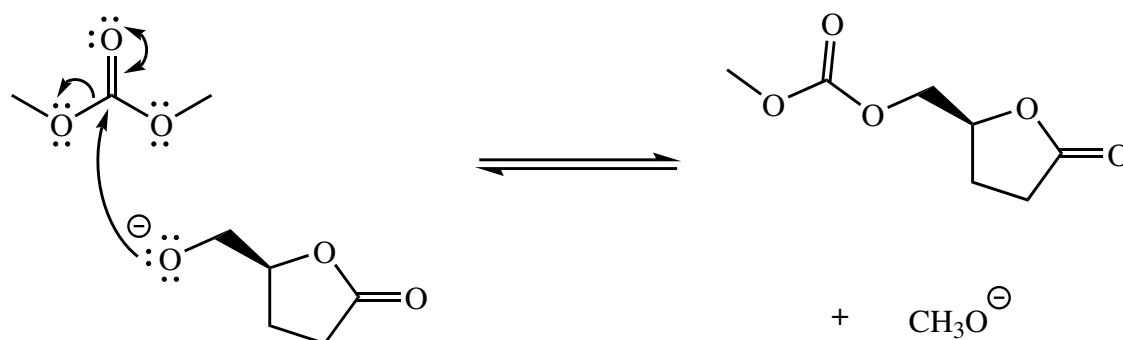
Scheme 3.15: Mechanism for carboxymethylation using the following conditions:

Unfortunately, the reactions were not very successful, although a small amount of product was observed by GC-MS (EI) in a reaction using a *p*-toluenesulfonic acid catalyst supported on polystyrene, which shows promise for future investigation into improving the greenness of this reaction.

3.7 Preliminary Stability Tests on Lactones

3.7.1 Base Sensitivity Tests

One of the problems that has been identified in the use of Cyrene as a solvent is its base-sensitivity, as surveyed by Wilson et al. who highlighted its reaction with most bases at 50 °C or above [Wilson 2016]. Cyrene was shown to be particularly sensitive to inorganic bases. It was, therefore, necessary to probe this issue with

regard to the present lactone species.

Method

As yet, ideal purity has not been achieved for either of the saturated lactones, **2b** and **2c**, so preliminary base sensitivity tests were run in the knowledge that there might be some interference from impurities. The base sensitivity tests were run at this point as this was the first juncture at which sufficient quantities of lactones **1b**, **2b**, **1c**, and **2c** were available post-synthesis.

In order to be comparable with the results of Wilson et al., the published method was used in this test: 0.5 mL of lactones **2b** and **2c** were stirred for 24 h at the specified temperature with 0.07 mmol of the given base. After 24 h, samples of each mixture were tested by TLC and NMR spectroscopy. Three representative bases were chosen for the study, pyridine, triethylamine and potassium carbonate. These were examples of bases which reportedly caused low, medium and high sensitivity in Cyrene.
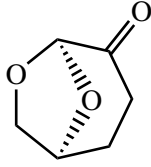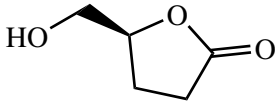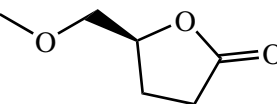
Results

No new spots appeared on any of the TLC plates which were not present in the starting material. Where reactions were observed by changes in the NMR spectra, the results have been given in table 3.7.

It is worth noting that no change was observed in any of the NMR spectra associated with the K_2CO_3 reaction with **2c**, but there was a milky, opaque precipitate that formed over the course of each reaction. The product was apparently insoluble in the NMR solvent and will require further testing to identify. The reason there was an apparent reaction with methoxy-lactone **2c** but not hydroxy-lactone **2b** is unclear, but it can be hypothesised that K_2CO_3 deprotonated the alcohol group of **2b** leaving no base for further reaction, which would not have been the case for **2c**. It is evident from these results that both lactones **2b** and **2c** display a greater resistance towards organic bases than Cyrene, however, **2c** was not stable towards inorganic base, K_2CO_3.

A comparison of the proton NMR spectra from analysis of each base sensitivity test of **2b** is given in figure 3.33.

Table 3.7: Base sensitivity test results: **2b** and **2c**.

Compound[a]	Base[b]	Temperature /°C	Reaction
		25	N
	Pyridine	50	Y
		100	Y
		25	N
	Et$_3$N	50	N
		100	Y
		25	Y
Cyrene, **2**[c]	K$_2$CO$_3$	50	Y
		100	Y
		25	N
	Pyridine	50	N
		100	N
		25	N
	Et$_3$N	50	N
		100	N
		25	N
2b[d]	K$_2$CO$_3$	50	N
		100	N
		25	N
	Pyridine	50	N
		100	N
		25	N
	Et$_3$N	50	N
		100	N
		25	Y
2c	K$_2$CO$_3$	50	Y
		100	Y

[a]0.5 mL of each compound used per run.

[b]0.07 mmol base used per run.

[c]Results published by Wilson et al. [Wilson 2016]

[d]Reactions not including deprotonation of the alcohol.
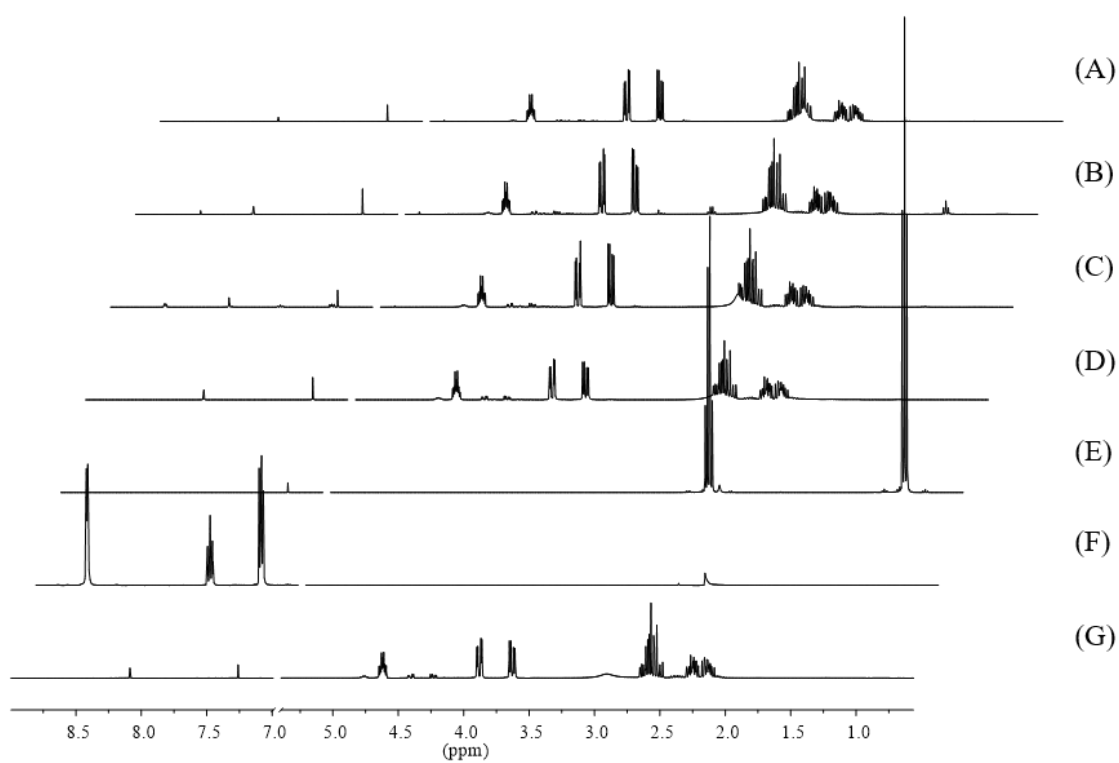
Figure 3.33: Proton NMR (400 MHz) spectrum comparison between base sensitivity test products of lactone **2b** at 100 °C with: (A) K_2CO_3, (B) Et_3N, (C) pyridine, (D) no base. Also included: (E) Et_3N, (F) pyridine, (G) **2b** before heating.

Although it might be expected for the lactone alcohol group to be deprotonated by a strong enough base, there is no indication that the lactone structure had broken down.

A comparison of the proton NMR spectra from analysis of each base sensitivity test of **2c** is given in figure 3.34.
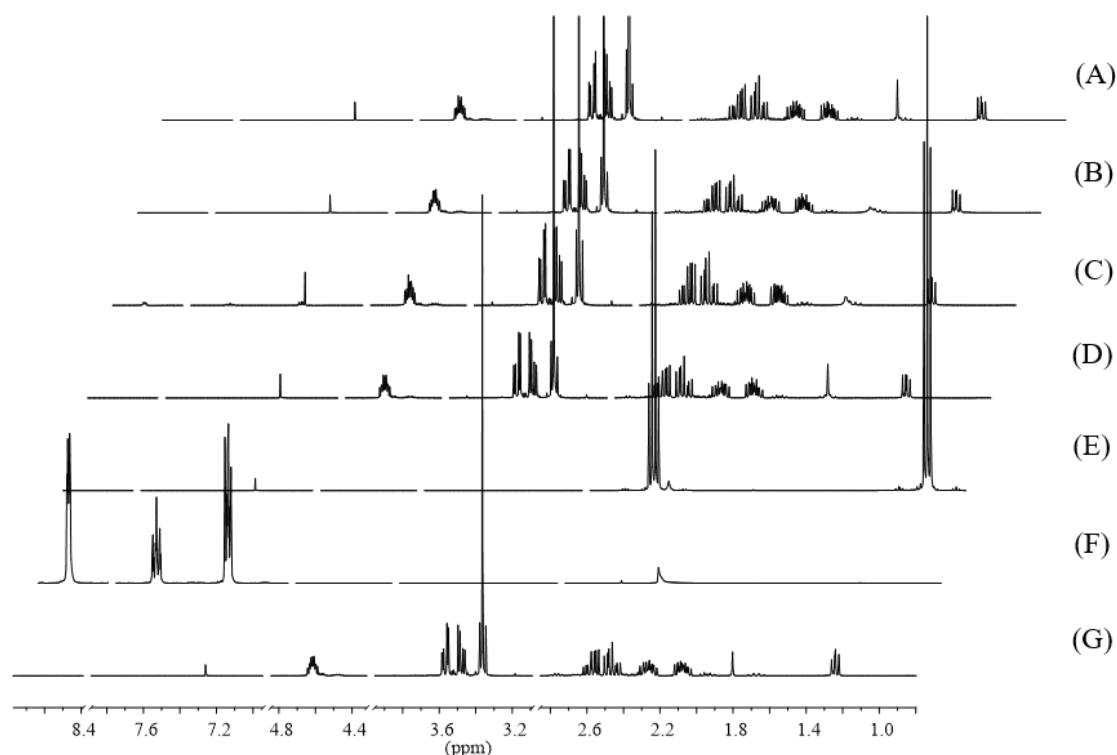
Figure 3.34: Proton NMR (400 MHz) spectrum comparison between base sensitivity test products of lactone **2c** at 100 °C with: (A) K_2CO_3, (B) Et_3N, (C) pyridine, (D) no base. Also included: (E) Et_3N, (F) pyridine, (G) **2c** before heating.

It should be noted that these base sensitivity tests were run without the addition of water, in accordance with the method published by Wilson et al. [Wilson 2016]. Due to the possibility of hydrolysis reaction, it would be sensible to do an aqueous test in the future. Also, due to the difficulty experienced in fully purifying either lactone, these tests have not been run on completely pure samples and it would be prudent to repeat them when pure lactones become available.

3.8 Concluding Remarks

Formate ester lactones **1b** and **2b** have successfully been produced by Baeyer-Villiger oxidation of LGO and Cyrene, respectively, using a green L-lysine catalysed reaction, and subsequently using a simplified, hydrogen peroxide only, reaction. This reaction has been safely scaled up to 175 g using a modified literature method allowing for a medium scale distillation of the product, with reasonable effectiveness. Lactone **2b** was found to distil with its byproduct **2a**, and separation work is ongoing.

Lactone **2b** has been successfully O-methylated using a literature procedure and scaled up for distillation. Although neither lactone has been produced in total purity, the purification has been improved significantly. Both lactones performed well in base sensitivity tests, with the only reaction observed being between K_2CO_3 and lactone **2c**.

HSPiP has been used to model each lactone in HSP space and both formate esters and O-methylated compounds to NMP, suggesting that they could be investigated as potential replacements for undesirable dipolar aprotic solvents. COSMO-RS has also been used to model the σ-potentials of each lactone, with the result that concurs with HSPiP. Lactones **1a**, **2a**, **1c** and **2c** have similar reponses to positive and negative charge densities but do not have as high an affinity as NMP, so this is a challenge to bear in mind for finding a replacement.

3.9 Further work

3.9.1 Physical Properties

Lactones

One of the goals of this research will eventually be to collect experimental data on the physical properties of the pure lactones. A selection of physical properties have been predicted using both COSMOtherm and HSPiP, shown in table 3.8.

Table 3.8: Predicted physical properties

Software	Lactone	b.p. /°C	Density /g mL^{-1}	Viscosity /cP	Flash point /°C
HSPiP	2b	276	1.2	52.6	137
HSPiP	2c	234	1.1	3.7	95
COSMOtherm	2b	344	1.3	12.8	183
COSMOtherm	2c	360	1.1	3.3	180

It will be the next step to determine these properties experimentally and compare the results to the predictions to assess their accuracy and to gain information on the usefulness of the solvent candidates.

Once the pure lactones are available in larger quantities, the aim will be to study the use of those lactones as solvents in real reactions. It would be useful to examine the Kamlett-Abboud-Taft parametersas they are useful for kinetics where HSPiP is based only on thermodynamics [Sherwood 2014.]

3.9.2 Baeyer-Villiger

In the future, the Baeyer-Villiger reaction may be improved further than the use of hydrogen peroxide, which would be favourable due to the potentially explosive hazards of peroxides. A recent publication demonstrated an alternative Baeyer-Villiger oxidation of cyclohexanone using a non-metallic SiO$_2$ catalyst, in many cases with >99 % yield [Zhang 2018].

In fact, the Baeyer-Villiger oxidation was recently successfully performed by Dr. Fergal Byrne, this time at reflux. Although high temperatures had previously been avoided, there is a potential benefit to running an exothermic reaction close to

100 °C, which in this case was controlled by using water as the solvent. From an engineering standpoint, this is an ideal temperature for energy recovery by the use of steam driven turbines, since the exotherm is self-sustaining.

3.9.3 t-Butylation of Lactone 2b

Recent work done in collaboration with Dr. Fergal Byrne, following large-scale production of lactone **2b** showed that *t*-butylation could be successfully performed as shown in scheme 3.16.
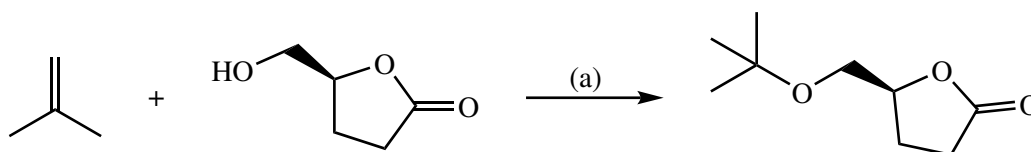

Scheme 3.16: *t*-Butylation of **2b**: (a) **2b** (10 mL), isobutene (16 mL), β-zeolite (30:1 Si/Al, 200 mg), DCM (30 mL).

Using the conditions described in scheme 3.16, an isolated yield of 68 % was achieved, although GC analysis suggested that the result could be significantly improved by optimisation of separation techniques. Although the reaction was run in DCM, smaller scale tests suggested that it could also be done in safer dioxolane. Due to this initial success, it is expected that future research on this molecule will involve further derivatisation of the pendant alcohol, and possibly the C=C double bond of unsaturated lactone, **1b** via Michael addition. The structures will then be screened as potential candidate molecules as in this thesis. It is pertinent to point out that, as more steps are added to a synthesis, the less green a candidate molecule becomes and less suitable as a solvent. However, there are related avenues to explore, such as the potential polymerisation of the unsaturated lactones via the double bond, or via ring-opening metathesis polymerisation (ROMP), thus creating a set of new bio-derived polymers.

3.9.4 Cyclic Carbonate

One of the unexplored products from the Baeyer-Villiger oxidation using *m*-CPBA reported by Koseki et al. is a cyclic carbonate from LGO, shown in scheme 3.17 [Koseki 1990].
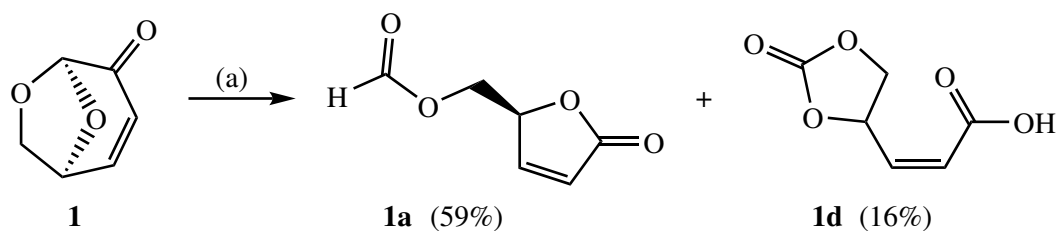
Scheme 3.17: Formation of a cyclic carbonate (**1d**) during the Baeyer-Villiger oxidation of LGO [Koseki 1990]. Conditions (a): LGO (4 mmol), *m*-CPBA (1.2 equiv.) stirred in dichloromethane, RT, 3 days.

Koseki et al. showed that the cyclic carbonate had been produced in a 16% yield. It is speculated that this could be increase by using a higher quantity of *m*-CPBA, and would be interesting to find out the properties of this molecule or its hydrogenated analogue for solvent or polymer use. HSPiP predictions are shown in figure 3.36 and suggest both carbonates, derived from LGO (**1d**) and Cyrene (**2d**), are much higher in polarity that the lactones.
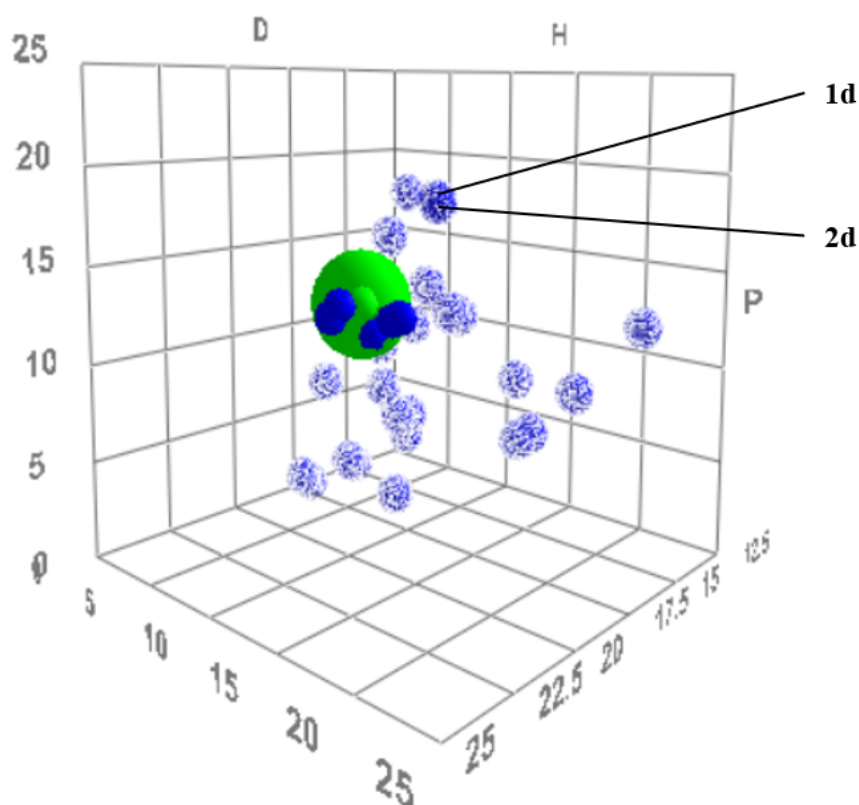
Figure 3.35: 3D HSPiP of lactones and carbonate

Both other 2D HSPiP images are given in appendix III.

Figure 3.36: 2D HSPiP of lactones and carbonate

The σ-surface produced in COSMO is shown in figure 3.37.

Figure 3.37: σ-Surface of cyclic carbonate **2d** produced in COSMOtherm.
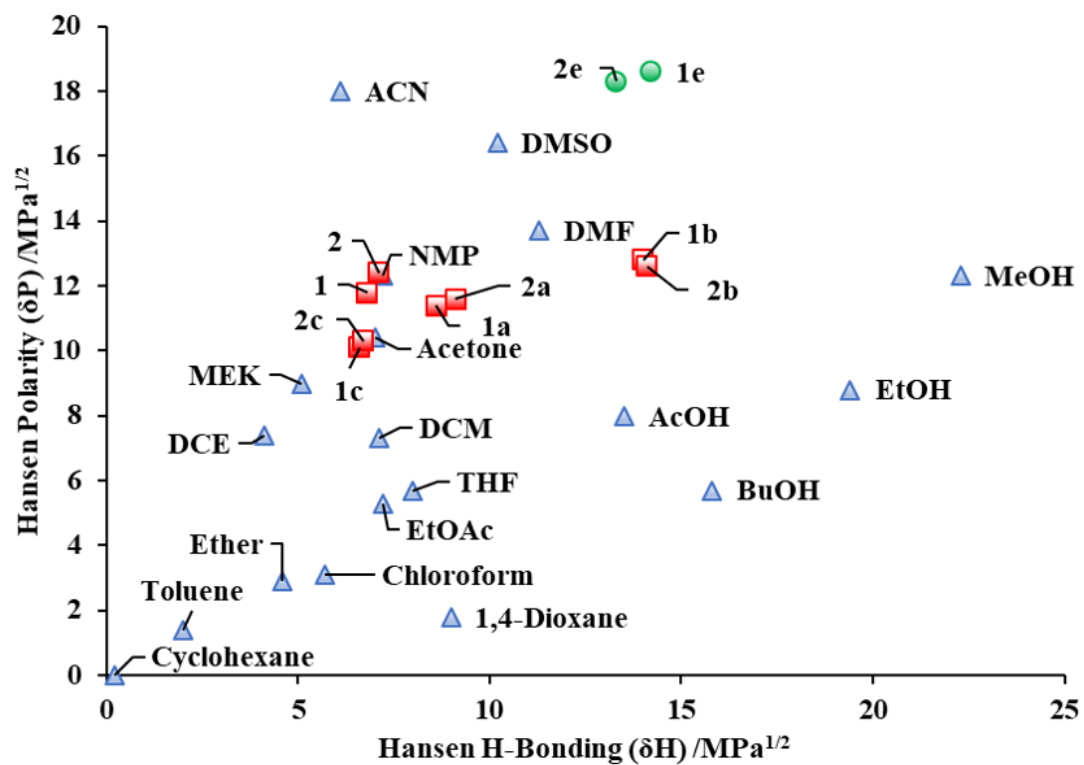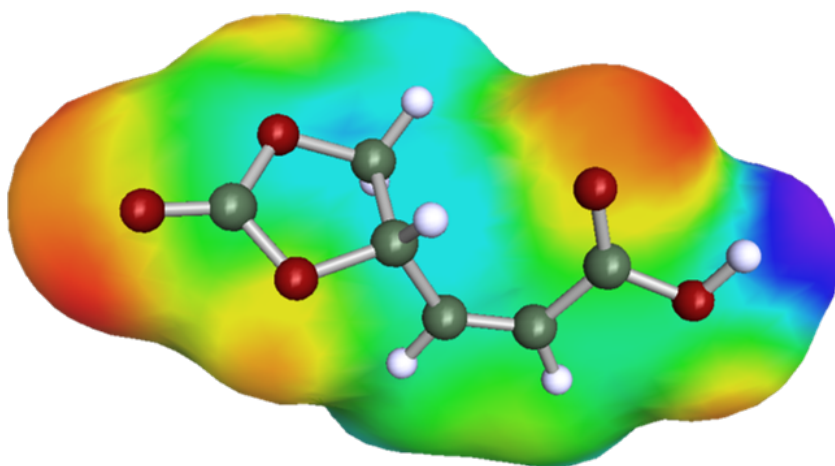
Chapter 4

Conclusions

As addressed in this thesis, solvents play a varied and important role in the chemical industry. Considering many valid concerns about the unsustainable sources and health hazards of many traditional solvents, there is increasing pressure on chemists to find alternatives. The primary aim of this work was to expand the body of available green and sustainable solvents.

One of the main objectives of this research was to find an alternative solvent to toluene for use in direct amidation reactions, with a previously identified green heterogeneous catalyst. In doing so, the intention was to use computer modelling techniques to help inform the strategy development and, thereafter, qualify the usefulness of predictive modelling for this purpose.

The second line of research described was intended to create a new set of green solvent candidates from bio-based platform molecule, levoglucosenone. The aim was to use computer modelling to predict the physical and solvent properties of the candidates and compare them to experimental results. In addition, a greener synthetic method was sought.

p-Cymene was chosen as a potential replacement for toluene in amidation reactions due to chemical similarities and to existing knowledge that low polarity solvents tend to out-perform more polar solvents in this reaction. The methodology used required that the amide products would be soluble enough at high temperatures so that the catalyst could be removed by filtration. Following this, the amide needed to be insoluble enough to recover by filtration at low temperatues. In this work, a range of amides were chosen using HSPiP and the Yalkowsky approximation

to predict the likelihood that they would fulfil these criteria. The method was then transferred to a flow system to improve efficiency.

The focus of candidate molecules derived from levoglucosenone was on using the Baeyer-Villiger reaction to produce 5-membered ring lactones. These structures were modelled using HSPiP and COSMO-RS to map their positions in solvent space relative to other known solvents. Subsequently, the strategy was to improve the purification and scale up the syntheses in order to produce enough product for experimental testing of their properties.

In this work, it has been shown that p-cymene, a potentially bio-based molecule derived from D-limonene, can be used as an effective substitute for toluene as a solvent in direct amidation reactions. In most cases, the heightened temperature available to p-cymene at atmospheric pressure provided some, even considerable improvements to the yield. The activated silica catalyst was shown to make significant improvements to the yields in cases where anilines were used, and much less so for aliphatic amines. Practical testing showed good agreement with the Yalkowsky approximation, but amide solubility did not quite correlate with HSP distance. It is concluded that there may be some flaws in the group contribution method, which may be biased towards certain functional groups. Additionally, using HSPiP to determine the relative solubility of a set of solutes is not a perfect model as the sphere of solvency around each solute will have a specific radius. However, the general trend in solubility data that HSPiP provides is convenient when used with caution and saves considerable lab time. Flow reactions were demonstrative of the possibility of improving the greenness of the reaction in the future, although the methodology requires optimisation as yields were low even with recirculation.

It has also been shown that levoglucosenone can be used as the basis for a range of derivative lactone molecules. Both polar protic and dipolar aprotic type structures were covered by these derivatives. The successful Baeyer-Villiger oxidation of levoglucosenone and Cyrene have been reported by three methods, using a known m-CPBA catalyst in a non-green reaction, using amino acid L-lysine as a catalyst in combination with H_2O_2, and by using H_2O_2 alone, which has also been reported in the literature. The final method was less wasteful and made purification easier and was scaled up successfully to 175 g. Lactone **2b** was subsequently methylated

via a literature procedure and scaled up to 25 g. Both of the lactones performed well in comparison to Cyrene when subjected to base sensitivity tests. Only **2c** had a clear reaction with K_2CO_3 at all temperatures tested.

Modelling software, HSPiP was used to determine the predictive Hansen solubility parameters for each lactone and compared them to a number of traditional solvents. Lactones **1a**, **2a**, **1c** and **2c** were all predicted to lie relatively close in HSP space to NMP, a concept that was confirmed by COSMO-RS modelling to produce the σ-surface and, via that, the σ-potential. Although none of the lactones have the extent of polarity which NMP does, they do resemble its characteristics, which is a step forward in the replacement of toxic dipolar aprotics.

To increase the relevance of this work it would be desirable to expand the dataset of amides chosen for synthesis to include heteroaromatic starting materials. In addition, proof that this method can be applied to a real-world pharmaceutical synthesis is strongly desirable. Using the HSPiP modelling method with another system, such as esterification, would show whether this is a good general starting point. Moreover, improving the synthesis of *p*-cymene would increase the possibility of it being financially viable as a solvent. Since the synthesis of Cyrene derived lactones, **2b** and **2c**, have been successfully scaled up, the next step will be to obtain a full set of physical data to compare with the predictive modelling. This will require an improved purification. The end point of this will be to test the usefulness of these lactones as solvents.

Chapter 5

Experimental Section

5.1 General Notes

5.1.1 Analytical Methods

CHN analysis was conducted by using an Exeter Analytical Inc CE-440 analyser and the samples were weighed by using a Sartorius SE-2 analytical balance. Gas chromatography was run using flame ionisation detection (FID) on an Agilent Technologies 6890N Network GC System with combined Agilent Technologies 7683B Series Injector autosampler and Peak Scientific PH600 hydrogen generator. Infrared Spectroscopy was run on a Perkin Elmer FTIR/FTNIR Spectrum 400 Spectrophotometer. Mass spectra by EI were obtained on a Perkin-Elmer Clarus 560 S mass spectrometer coupled to an Perkin-Elmer Clarus 500 gas chromatograph (General 1 method: 30 m Rxi-5HT column, 0.25 mm ID, 0.25 μm d^f. Carrier gas: helium. Flow rate: 2 mL min^{-1}. Injection volume: 0.5 μL. Inlet temperature: 300 °C. Oven temperature: initial, 50 °C; ramp, 30 °C min^{-1}; final, 300 °C; hold 5 min). NMR spectra were obtained using a Bruker 400 MHz spectrometer.

Automated column chromatography was run on a Biotage Isolera 4 Flash Purification System, with UV-detection, using 25, 50 or 100 g column inserts (not pre-filled) and K60 silica gel. Amidations were carried out on a Radleys 6-position multipoint hot-plate and condenser. Flow reactions were run on a Uniqsis Flowsyn continuous flow reactor.

5.1.2 Chemicals Used

Carboxylic acids and amines were obtained from Sigma-Aldrich. mCPBA (less than or equal to 77 %) and 2,4,6-tri-tert-butylphenol, were also sourced from Sigma-Aldrich, as were dichloroethane, platinum on alumina pellets (1 wt% loading), L-lysine, iodomethane, sodium hydride, and dimethyl carbonate. Solvents were purchased from VWR chemicals apart from Cyrene and Levoglucosenone, which were donated by Circa. Sodium sulfite, sodium hydrogen carbonate, sodium chloride, magnesium sulfate and hydrogen peroxide 30% were bought from Fisher Scientific.

5.1.3 COSMO-RS

The 3-dimensional molecular geometry of lactones in Chapter 3 was drawn using ArgusLab (version 4.0.1, M. Thompson and Planaria Software LLC, Seattle, WA). Conformational analysis was then performed using the COSMOconf program (COSMOconfX Version 4.0 COSMOlogic GmbH & Co.KG, Germany) that involves semi-empirical AM1 (Austin Model 1) calculations followed by a more accurate DFT (Density Functional Theory) treatment of the most important AM1 conformers. σ-surfaces were generated using the COSMOtherm software (F. Eckert and A. Klamt, COSMOthermX, Version C30_1705; COSMOlogic GmbH & Co. KG, Leverkusen, Germany, 2013).

5.1.4 HSPiP

Hansen solubility parameter predictions (δD, δP, δH) were obtained using Hansen Solubility Parameters in Practice (HSPiP) 5th Edition (version 5.0.04, S. Abbott and H. Yamamoto, HSPiP Computer Program, 2013) using the Y-MB method.
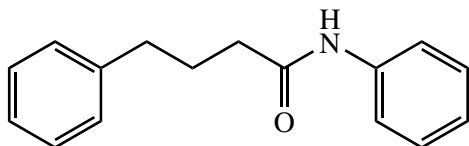
5.2 Optimised Synthetic Methods and Characterisation

5.2.1 Amides

Activation of silica catalyst

K60 silica gel was activated at 700 °C for 4 h in a furnace under ambient pressure, with a ramp rate of 10 °C min^{-1}. The catalyst was allowed to cool naturally to room temperature.
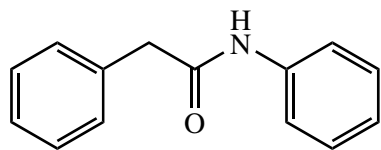
N-Phenyl-4-phenylbutanamide (3)

4-Phenylbutyric acid (1.97 g, 12 mmol), activated K60 silica (0.62 g, 20 wt%), hexadecane (0.15 g) and 20 mL of solvent (toluene or p-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, aniline (1.12 g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or p-cymene. The filtrate was left in a refrigerator (4 °C). The product was filtered and washed with cold reaction solvent (30 mL) then cyclohexane (30 mL) and dried under vacuum. Yield: 61% (toluene, 111 °C), 37% (p-cymene, 111 °C), 92% (p-cymene, 177 °C); Literature m.p. 94-96 °C [McDonald 2016]; m.p. 97-99 °C. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 7.51 (d, J = 7.9 Hz, 2H; Ar), 7.39 (br s, 1H; NH), 7.33-7.28 (m, 4H; Ar), 7.20 (t, J = 8.8 Hz, 3H; Ar), 7.10 (t, J = 7.4 Hz, 1H; Ar), 2.70 (t, J = 7.5 Hz, 2H; CH$_2$), 2.34 (t, J = 7.5 Hz, 2H; CH$_2$), 2.06 (m, J = 7.5 Hz, 2H; CH$_2$). ^{13}C NMR (100 MHz, CDCl$_3$): δ = 171.2, 141.5, 138.0, 129.1, 128.6, 128.6, 126.2, 124.3, 120.0, 36.8, 35.2, 27.0. GC-MS (GC-EI) m/z: [M]$^+$ Calcd for C$_{16}$H$_{17}$NO 239.13; Found 239.27. IR: $\tilde{\nu}$ = 3317 (NH), 1663 cm^{-1} (C=O).
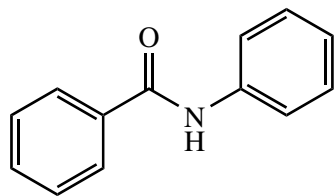
2,N-Diphenylacetamide (4)

Phenylacetic acid (1.63 g, 12 mmol), activated K60 silica (0.55 g, 20 wt%), hexadecane (0.15 g) and 20 mL of solvent (toluene or p-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, aniline (1.12 g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or p-cymene. The filtrate was left in a refrigerator (4 °C). The product was filtered and washed with cold reaction solvent (30 mL) then cyclohexane (30 mL) and dried under vacuum. Yield: 80% (toluene, 111 °C), 57% (p-cymene, 111 °C), 88% (p-cymene, 177 °C); Literature m.p.: 119-120 °C [Yan 2017]; m.p. 118-122 °C. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 7.44-7.26 (m, 10H; Ar + NH), 7.09 (t, J = 7.4 Hz, 1H; Ar), 3.72 (s, 2H; CH$_2$). ^{13}C NMR (100 MHz, CDCl$_3$): δ = 169.3, 137.8, 134.6, 129.6, 129.3, 129.0, 127.7, 124.6, 120.0, 44.9. GC-MS (GC-EI) m/z: [M]$^+$+ Calcd for C$_{14}$H$_{13}$NO 211.10; Found 211.24. IR: $\tilde{\nu}$ = 3254 (NH), 1655 cm^{-1} (C=O).
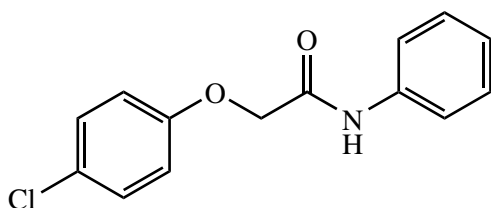
N-Phenylbenzamide (5)

Benzoic acid (1.47 g, 12 mmol), activated K60 silica (0.52 g, 20 wt%), hexadecane (0.15 g) and 20 mL of solvent (toluene or p-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, aniline (1.12 g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the

catalyst washed with 10 mL of hot toluene or *p*-cymene. The filtrate was left in a refrigerator (4 °C). The product was filtered and washed with cold reaction solvent (30 mL) then cyclohexane (30 mL) and dried under vacuum. Product purified by biphasic separation in dichloromethane, using 1M HCl(aq), 1M Na_2CO_3 then brine. The organic layer was dried over $MgSO_4$. Yield: 57% (toluene, 111 °C), 10% (*p*-cymene, 111 °C), 69% (*p*-cymene, 177 °C); Literature m.p. 163-165 °C [Xie 2017]; m.p. 164-166 °C. ^{1}H NMR (400 MHz, $(CD_3)_2CO$): δ = 7.87 (m, 3H; Ar + NH), 7.65 (d, J = 7.8 Hz, 2H; Ar), 7.55 (t, J = 7.4 Hz, 1H, Ar), 7.49 (t, J = 7.4 Hz, 2H; Ar), 7.38 (t, J = 7.1 Hz, 2H; Ar), 7.16 (t, J = 7.5 Hz, 1H; Ar). ^{13}C NMR (400 MHz, $(CD_3)_2CO$): δ = 166.3, 140.3, 136.3, 132.3, 129.5, 128.3, 124.5, 121.0. GC-MS (GC-EI) m/z: $[M]^+$ Calcd for $C_{13}H_{11}NO$ 197.08; Found 197.19. IR: $\tilde{\nu}$ = 3344 (NH), 1654cm^{-1} (C=O).

**2-(4-Chlorophenoxy)-*N*-phenylacetamide (6)**

4-Chlorophenoxyacetic acid (2.24 g, 12 mmol), activated K60 silica (0.67 g, 20 wt%), tetradecane (0.15 g) and 20 mL of solvent (toluene or *p*-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, aniline (1.12 g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or *p*-cymene. The filtrate was left in a refrigerator (4 °C). The product was filtered and washed with cold reaction solvent (30 mL) then cyclohexane (30 mL) and dried under vacuum. Yield: 74% (toluene, 111 °C), 75% (*p*-cymene, 111 °C), 77% (*p*-cymene, 177 °C); Literature m.p. 128-129 °C [Comerford 2009]; m.p. 128-130 °C. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 8.22 (br s, 1H; NH), 7.58 (d, J = 7.7 Hz, 2H; Ar), 7.36 (t, J = 8.0 Hz, 2H; Ar), 7.31 (m, 2H; Ar), 7.17 (t, J = 7.4 Hz, 1H; Ar), 6.93 (m, 2H; Ar), 4.58 (s, 2H; CH$_2$). ^{13}C NMR

(100 MHz, CDCl$_3$): δ = 165.9, 155.7, 136.8, 130.0, 129.3, 127.6, 125.1, 120.3, 116.3, 68.0. GC-MS (GC-EI) m/z: [M]$^+$ Calcd for C$_{14}$H$_{12}$ClNO$_2$ 261.06; Found 261.08. IR: $\tilde{\nu}$ = 3187 (NH), 1660 cm^{-1} (C=O).

N-(2-Chlorophenyl)-2-phenylacetamide (7)
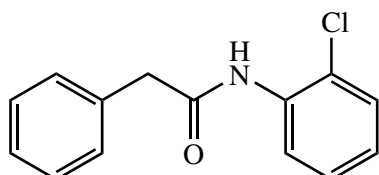
Phenylacetic acid (1.63 g, 12 mmol), activated K60 silica (0.63 g, 20 wt%), hexadecane (0.15 g) and 20 mL of solvent (toluene or *p*-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, 2-chloroaniline (1.53 g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or *p*-cymene. The filtrate was left in a refrigerator (4 °C). The product was filtered and washed with cold reaction solvent (30 mL) then cyclohexane (30 mL) and dried under vacuum. Yield: 18% (toluene, 111 °C), 6% (*p*-cymene, 111 °C), 55% (*p*-cymene, 177 °C); Literature m.p. 129 [Yamagami 1984], 120 °C [Jenkins 1933]; m.p. 121-123 °C. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 8.38 (d, J = 8.2 Hz, 1H; Ar), 7.67 (br s, 1H; NH), 7.45-7.34 (m, 5H; Ar), 7.29-7.22 (m, 2H; Ar), 7.00 (td, J = 7.9 Hz, 1.6, 1H; Ar), 3.80 (s, 2H; CH$_2$). ^{13}C NMR (100 MHz, CDCl$_3$): δ = 169.2, 134.5, 134.0, 129.8, 129.5, 129.0, 128.0, 127.8, 124.8, 122.8, 121.3, 45.3. GC-MS (GC-EI) m/z: [M]$^+$ Calcd for C$_{14}$H$_{12}$ClNO 245.06; Found 245.05. IR: $\tilde{\nu}$ = 3261 (NH), 1660 cm^{-1} (C=O).
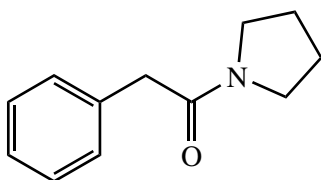
N-(2,6-Dimethylphenyl)-2-phenylacetamide (8)
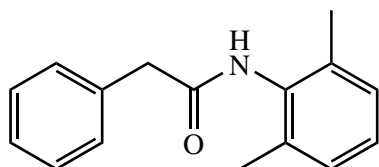
Phenylacetic acid (1.63 g, 12 mmol), activated K60 silica (0.62 g, 20 wt%), hexadecane (0.15 g) and 20 mL of solvent (toluene or p-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, 2,6-dimethylaniline (1.45 g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or p-cymene. The filtrate was left in a refrigerator (4 °C). The product was filtered and washed with cold reaction solvent (30 mL) then cyclohexane (30 mL) and dried under vacuum. Product separated by column chromatography using 95:5 DCM-ethyl acetate. Yield: 20% (toluene, 111 °C), 2% (p-cymene, 111 °C), 69% (p-cymene, 177 °C); Literature m.p. 145-148 °C [Diago-M 1980]; m.p. 146-148 °C. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 7.44-7.01 (m, 8H; Ar), 6.59 (br s, 1H; NH), 3.79 (s, 2H; CH$_2$), 2.11 (s, 6H; CH$_3$). ^{13}C NMR (100 MHz, CDCl$_3$): δ = 169.6, 135.2, 133.8, 135.3, 129.6, 129.3, 128.2, 127.7, 127.4, 44.1, 18.3. GC-MS (GC-EI) m/z: $[M]^+$ Calcd for C$_{16}$H$_{17}$NO 239.13; Found 238.98. IR: $\tilde{\nu}$ = 3250 (NH), 1641 cm^{-1} (C=O). Anal. Calcd for C$_{16}$H$_{17}$NO: C, 80.30; H, 7.16; N, 5.85. Found: C, 80.32; H, 7.07; N, 5.86.
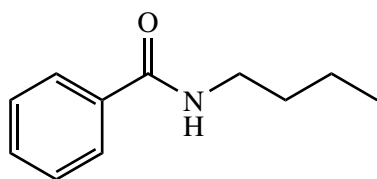
Phenylacetylpyrrolidine (9)

Phenylacetic acid (1.63 g, 12 mmol), activated K60 silica (0.50 g, 20 wt%), hexadecane (0.15 g) and 20 mL of solvent (toluene or p-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, pyrrolidine (0.85 g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or p-cymene. The filtrate was left in a refrigerator (4 °C). Product purified by distillation of the solvent, followed by dry column vacuum chromatography using increasingly polar fractions of petroleum spirit-ethyl acetate. Yield: 52% (toluene, 111 °C), 28% (p-cymene, 111 °C), 22%

(p-cymene, 177 °C); Literature m.p. 48 °C [Mukherjee 1971]; m.p. 27-38 °C. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 7.30-7.18 (m, 5H; Ar), 3.62 (s, 2H; CH$_2$), 3.45 (t, J = 6.8 Hz, 2H; CH$_2$), 3.38 (t, J = 6.8 Hz, 2H; CH$_2$), 1.91-1.76 (m, 4H; CH$_2$). ^{13}C NMR (100 MHz, CDCl$_3$): δ = 169.6, 135.0, 129.0, 128.7, 126.8, 47.0, 46.0, 42.4, 26.2, 24.4. GC-MS (GC-EI) m/z: [M]$^+$ Calcd for C$_{12}$H$_{15}$NO 189.12; Found 188.63. IR: $\tilde{\nu}$ = 1619 cm^{-1} (C=O).
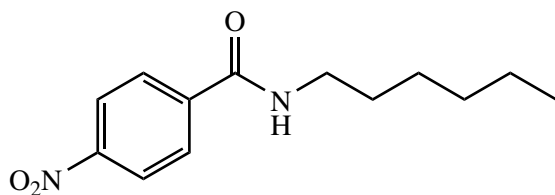
N-Butylbenzamide (10)

Benzoic acid (1.47 g, 12 mmol), activated K60 silica (0.47 g, 20 wt%), hexadecane (0.15 g) and 20 mL of solvent (toluene or p-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, n-butylamine (0.88 g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or p-cymene. The filtrate was left in a refrigerator (4 °C). Product purified by distillation of the solvent, followed by an automated chromatography column on a Biotage Isolera 4 using 12% ethyl acetate in petroleum spirit rising to 100% ethyl acetate. Subsequent biphasic separation in dichloromethane, using 1M HCl(aq), 1M NaOH(aq) then brine used to remove excess acid. The organic layer was dried over MgSO$_4$. Yield: 0% (toluene, 111 °C), 0% (p-cymene, 111 °C), 55% (p-cymene, 177 °C); Oil. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 7.76 (d, J = 7.1 Hz, 2H; Ar), 7.47 (t, J = 7.4 Hz, 1H; Ar), 7.40 (t, J = 7.0 Hz, 2H; Ar), 6.27 (br s, 1H; NH), 6.88 (td, J = 7.2, 5.9 Hz, 2H; CH$_2$), 1.63-1.55 (m, 2H; CH$_2$), 1.45-1.35 (m, 2H; CH$_2$), 0.94 (t, J = 7.4 Hz, 3H; CH$_3$). ^{13}C NMR (100 MHz, CDCl$_3$): δ = 167.7, 135.0, 131.4, 128.6, 127.0, 39.9, 31.9, 20.3, 13.9. GC-MS (GC-EI) m/z: [M]$^+$ Calcd for C$_{11}$H$_{15}$NO 177.12; Found 177.11. IR: $\tilde{\nu}$ = 3308 (NH), 1633 cm^{-1} (C=O).
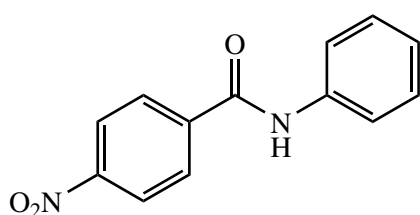
N-Hexyl-4-nitrobenzamide (11)

4-Nitrobenzoic acid (2.01 g, 12 mmol), activated K60 silica (0.65 g, 20 wt%), tetradecane (0.15 g) and 20 mL of solvent (toluene or p-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and subaseal. Once reflux was reached, hexylamine (1.21g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or p-cymene. The filtrate was left in a refrigerator (4 °C). The product was filtered and washed with cold reaction solvent (30 mL) then cyclohexane (30 mL) and dried under vacuum. Product purified by biphasic separation in ethyl acetate, using 1M HCl(aq), 1M Na_2CO_3 then brine. The organic layer was dried over $MgSO_4$. Yield: 0% (toluene, 111 °C), 0% (p-cymene, 111 °C), 34% (p-cymene, 177 °C); Literature m.p. 82-83 °C [Nammalw. 2015]; m.p. 82-85 °C. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 8.28 (d, J = 8.9 Hz, 2H; Ar), 7.92 (d, J = 8.7 Hz, 2H; Ar), 6.23 (br s, 1H; NH), 3.47 (td, J = 7.2, 6.0 Hz, 2H; CH$_2$), 1.67-1.59 (m, 2H; CH$_2$), 1.42-1.30 (m, 6H; CH$_2$), 0.89 (t, J = 7.0 Hz, 3H; CH$_3$). ^{13}C NMR (100 MHz, CDCl$_3$): δ = 165.7, 149.6, 140.5, 128.2, 123.9, 40.6, 31.6, 29.6, 26.8, 22.7, 14.1. GC-MS (GC-EI) m/z: [M]$^+$ Calcd for C$_{13}$H$_{18}$N$_2$O$_3$ 250.13; Found 250.07. IR: $\tilde{\nu}$ = 3305 (NH), 1637 cm^{-1} (C=O).
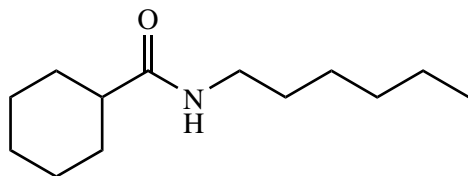
4-Nitro-N-phenylbenzamide (12)

4-Nitrobenzoic acid (2.01 g, 12 mmol), activated K60 silica (0.63 g, 20 wt%), tetradecane (0.15 g) and 20 mL of solvent (toluene or p-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, aniline (1.12 g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or p-cymene. The filtrate was left in a refrigerator (4 °C). Product purified by biphasic separation in ethyl acetate, using 1M HCl(aq), 1M Na_2CO_3 then brine. The organic layer was dried over $MgSO_4$. Yield: 8% (toluene, 111 °C), <1% (p-cymene, 111 °C), 11% (p-cymene, 177 °C); Literature m.p. 215-217 °C [Liu 2017]; m.p. 216-217 °C. ^{1}H NMR (400 MHz, CDCl$_3$): $\delta =$ 9.85 (br s, 1H; NH), 8.37 (d, $J = 8.8$ Hz, 2H; Ar), 8.24 (d, $J = 8.8$, 2H; Ar), 7.84 (d, $J = 7.6$, 2H; Ar), 7.38 (t, $J = 8.0$, 2H; Ar), 7.15 (t, $J = 7.4$, 1H; Ar). ^{13}C NMR (100 MHz, CDCl$_3$): $\delta =$ 164.0, 149.2, 140.7 138.7, 129.3, 128.8, 124.3, 123.6, 120.6. GC-MS (GC-EI) m/z: [M]$^+$ Calcd for $C_{13}H_{10}N_2O_3$ 242.07; Found 242.11. IR: $\tilde{\nu} =$ 3320 (NH), 1650 cm^{-1} (C=O).
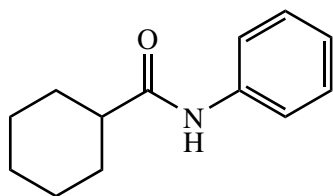
N-Hexylcyclohexanecarboxamide (13)

Cyclohexanecarboxylic acid (1.54 g, 12 mmol), activated K60 silica (0.55 g, 20 wt%), hexadecane (0.15 g) and 20 mL of solvent (toluene or p-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, hexylamine (1.21g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or p-cymene. The filtrate was left in a refrigerator (4 °C). Product purified by distillation of the solvent followed by chromatography column using 60:40 petroleum spirit:ethyl acetate, then by biphasic separation in dichloromethane, using 1M HCl(aq), 1M Na_2CO_3 then brine. The organic layer was dried over $MgSO_4$. Yield: 9% (toluene, 111 °C), 16% (p-cymene,

111 °C), >99% (*p*-cymene, 177 °C); Literature m.p. 69-70 °C [Ohshima 2012]; m.p. 66-68 °C. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 5.41 (br s, 1H, NH), 3.22 (td, J = 7.1, 5.9 Hz, 2H; CH$_2$), 2.04 (tt, J = 11.8, 3.5 Hz, 1H; CH), 1.86-1.76 (m, 4H; CH$_2$), 1.67-1.64 (m, 1H; CH$_2$), 1.49- 1.37 (m, 4H; CH$_2$), 1.34- 1.22 (m, 9H; CH$_2$), 0.88 (t, J = 7.0, 3H; CH$_3$). ^{13}C NMR (100 MHz, CDCl$_3$): δ = 176.2, 45.8, 39.4, 32.0, 29.9, 29.8, 26.7, 25.9, 22.7, 14.1. GC-MS (GC-EI) m/z: [M]$^+$ Calcd for C$_{13}$H$_{25}$NO 211.19; Found 211.24. IR: $\tilde{\nu}$ = 3295 (NH), 1636 cm^{-1} (C=O).
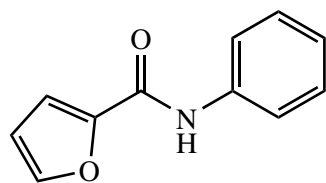
N-Phenylcyclohexanamide (14)

Cyclohexanecarboxylic acid (1.54 g, 12 mmol), activated K60 silica (0.55 g, 20 wt%), hexadecane (0.15 g) and 20 mL of solvent (toluene or *p*-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, aniline (1.12g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or *p*-cymene. The filtrate was left in a refrigerator (4 °C). The product was filtered and washed with cold reaction solvent (30 mL) then cyclohexane (30 mL) and dried under vacuum. Yield: 11% (toluene, 111 °C), 5% (*p*-cymene, 111 °C), 70% (*p*-cymene, 177 °C); Literature m.p. 149-151 °C [Lebedyeva 2014]; m.p. 148-149 °C. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 7.54 (d, J = 8.0 Hz, 2H; Ar), 7.47 (br s, 1H; NH), 7.29 (t, J = 7.9 Hz, 2H; Ar), 7.08 (t, J = 7.4 Hz, 1H; Ar), 2.23 (tt, J = 11.8, 3.5, 1H; CH), 1.94 (d, J = 13.3, 2H; CH$_2$), 1.84-1.80 (m, 2H; CH$_2$), 1.70-1.68 (m, 1H; CH$_2$), 1.58-1.49 (m, 2H; CH$_2$), 1.34-1.21 (m, 3H; CH$_2$). ^{13}C NMR (100 MHz, CDCl$_3$): δ = 174.7, 138.3, 129.0, 129.2, 129.9, 46.6, 29.8, 25.8. GC-MS (GC-EI) m/z: [M]$^+$ Calcd for C$_{13}$H$_{17}$NO 203.13; Found 203.31. IR: $\tilde{\nu}$ = 3311 (NH), 1650 cm^{-1} (C=O).

Furan-2-carboxylic acid phenylamide (15)

2-Furoic acid (1.35 g, 12 mmol), activated K60 silica (0.50 g, 20 wt%), tetradecane (0.15 g) and 20 mL of solvent (toluene or *p*-cymene) were heated to reflux (111 °C/177 °C) in a two-necked round bottom flask equipped with a condenser and suba-seal. Once reflux was reached, aniline (1.12g, 12 mmol) was injected. After 24 hours, the reaction mixture was hot-filtered through a sintered glass funnel and the catalyst washed with 10 mL of hot toluene or *p*-cymene. The filtrate was left in a refrigerator (4 °C). The product was filtered and washed with cold reaction solvent (30 mL) then cyclohexane (30 mL) and dried under vacuum. Yield: 0% (toluene, 111 °C), 0% (*p*-cymene, 111 °C), 78% (*p*-cymene, 177 °C); Literature m.p. 123-125 °C [Yan 2017]; m.p. 123-124 °C. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 8.08 (br s, 1H; NH), 7.64 (dd, J = 8.7, 1.0 Hz, 2H: Ar), 7.50 (dd, J = 1.7, 0.8 Hz, 1H; CH furan), 7.35 (t, J = 8.0 Hz; 2H; Ar), 7.23 (dd, J = 3.5, 0.8 Hz, 1H; CH furan), 7.13 (t, J = 7.4 Hz, 1H; Ar), 6.55 (dd, J = 3.5, 1.8 Hz, 1H; CH furan). ^{13}C NMR (100 MHz, CDCl$_3$): δ = 156.2, 147.9, 144.3, 137.5, 129.2, 124.6, 120.0, 115.4, 112.7. GC-MS (GC-EI) m/z: [M]$^+$ Calcd for C$_{11}$H$_9$NO$_2$ 187.06; Found 187.09. IR: $\tilde{\nu}$ = 3278 (NH), 1651 cm^{-1} (C=O).

Flow system

4-Phenylbutyric acid (1.97 g, 12 mmol), aniline (1.14 g, 12 mmol), hexadecane (0.15 g) and 20 mL of *p*-cymene were weighed into a vessel, which was attached to the flow system and pumped at 0.3 mL min^{-1} through a heated chamber (150 °C) containing activated K60 silica (2.5 g). Yields were obtained by reheated the contents of the flask and running a sample on GC (Rxi-5HT column, General 1 method).
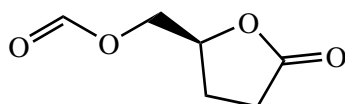
5.2.2 Lactones

(S)-γ-Formyloxymethyl-α,β-butenolide (1a)
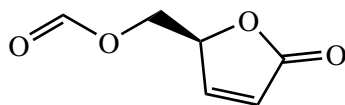
Levoglucosenone (**1**) (1.01 g, 8 mmol), tri-tert-butylphenol (0.11 g, 4 % wt. wrt m-CPBA) and dichloroethane (32 mL) were entered into a flask and stirred at room temperature. To the solution was added *meta*-chloroperoxybenzoic acid (m-CPBA, 2.13 g, 12 mmol). Stirring continued for 24 h. Reaction mixture subsequently diluted with DCM (20 mL) and washed with saturated Na_2SO_3 (50 mL). Peroxide quenched with 3.2 mm pellet of platinum on alumina (1 wt.% loading) then $NaHCO_3$ powder was added with shaking until no more gas evolved. Product obtained by automated column chromatography using a 25 g Biotage cartridge with K60 silica. Method: (Part 1) silica cartridge: 25 g; solvent system: cyclohexane (A)/ethyl acetate (B); gradient: B = 2% for 1 CV, B = 2% rising to 20% over 10 CV, B = 20% for 2 CV. (Part 2) Flush through; silica cartridge; 25 g; solvent system: cyclohexane (A)/ethyl acetate (B); gradient: B = 0% for 10 CV, B = 0% rising to 100% over 5 CV, B = 100% for 10 CV; flow rate: 25 mL min^{-1}. Yield = 14%. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 8.01 (s, 1H; CH formate), 7.43 (dd, J = 5.8, 1.6 Hz, 1H; CH$_2$), 6.2 (dd, J = 5.7, 2.1 Hz, 1H; CH$_2$), 5.27-5.23 (m, 1H), 4.45 (dd, J = 12.1, 3.9 Hz, 1H; CH), 4.40 (dd, J = 12.1, 4.9 Hz, 1H; CH). ^{13}C NMR (100 MHz, CDCl$_3$, DEPT): δ = C 172.1, CH 160.2, 152.2, 123.6, 80.4, CH$_2$ 61.9.

(S)-γ-Formyloxymethyl-α,β-butyrolactone (2a)


Cyrene (**2**) (1.02 g, 8 mmol), tri-tert-butylphenol (0.11 g, 4 % wt. wrt m-CPBA) and dichloroethane (32 mL) were entered into a flask and stirred at room temperature. To the solution was added *meta*-chloroperoxybenzoic acid (m-CPBA, 2.13 g, 12

mmol). Stirring continued for 24 h. Reaction mixture was subsequently diluted with DCM (20 mL) and washed with saturated Na_2SO_3 (50 mL). Peroxide quenched with 3.2 mm pellet of platinum on alumina (1 wt.% loading) then $NaHCO_3$ powder was added with shaking until no more gas evolved. Product obtained by automated column chromatography using a 50 g Biotage cartridge with K60 silica. Solvent system: 2-20% ethyl acetate in cyclohexane, 25 mL min^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 8.05 (s, 1H), 4.75-4.69 (m, 1H), 4.36 (dd, J = 12.3, 3.2 Hz, 1H), 4.19 (dd, J = 12.0, 5.2 Hz, 1H), 2.71-2.47 (m, 2H), 2.38-2.29 (m, 1H), 2.10-1.98 (m, 1H). ^{13}C NMR (100 MHz, CDCl$_3$, DEPT): δ = C 176.5, CH 160.4, 76.9, CH$_2$ 64.6, 28.2, 23.9. MS (ESI) m/z: [M]$^+$ Calcd for $C_6H_8NaO_4$ 167.0317; Found 167.0315. IR: $\tilde{\nu}$ = 2939 (C-H), 1771, 1714 (C=O), 1151 (C-O) cm^{-1}.

(S)-γ-Hydroxymethyl-α,β-butenolide (1b)
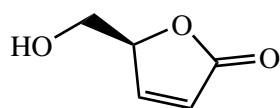
L-lysine catalysed Baeyer-Villiger oxidation: Levoglucosenone (**1**) (1.26 g, 10 mmol), L-lysine (0.15 g, 1 mmol) and water (15.8 mL) were stirred at room temperature. To the solution was added H_2O_2 (3.4 mL of a 30% solution in water). Stirring continued for 24 h. Peroxide quenched with four pellets of platinum on alumina (1 wt.% loading) and left to degas. After removal of pellets, the H_2O was removed by rotary evaporation. The reaction mixture was triturated and sonicated in acetonitrile (10 mL, 30 mins) then filtered. The liquid portion was dried by rotary evaporation to collect the product. Yield = 72%. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 7.50 (dd, J = 5.8, 1.5 Hz, 1H; CH), 6.18 (dd, J = 5.7, 2.0 Hz, 1H; CH), 5.16-5.13 (m, 1H; CH), 3.98 (dd, J = 12.4, 3.8 Hz, 1H; C**H**$_2$OH), 3.77 (dd, J = 12.3, 4.9 Hz, 1H; C**H**$_2$OH). ^{13}C NMR (100 MHz, CDCl$_3$, DEPT): δ = C 173.8, CH 154.2, 122.9, 84.5, CH$_2$ 62.1. MS (ESI) m/z: [M]$^+$ Calcd for $C_6H_8NaO_4$ 137.0210; Found 137.0209. IR: $\tilde{\nu}$ = 3387 (O-H), 3094 (C-H alkene), 2920 (C-H alkane), 1734 (C=O) cm^{-1}. IR: $\tilde{\nu}$ = 3404 (OH), 1782 cm^{-1} (C=O),

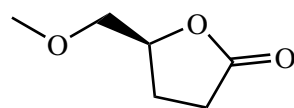## (S)-γ-Hydroxymethyl-α,β-butyrolactone (2b)
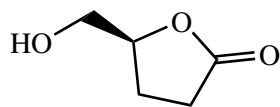
Large-scale Baeyer-Villiger oxidation of Cyrene: Cyrene (**2**) (175 g, 1.37 mol) and H$_2$O (200 mL) were stirred at room temperature. H$_2$O$_2$ (170 mL, 30 wt%, 1.22 equiv.) was added dropwise over 7 h with temperature monitoring. The temperature was maintained between 50-70 °C. The reaction was stirred at 50 °C for 12 h. Reaction was quenched with platinum on alumina (40 pellets) and allowed to gas off for 4 days. H$_2$O removed by rotavap. Product was recovered by vacuum distillation (0.3 mbar, 104 °C). Yield = 94%.

Hydrogenation of **1b**: (S)-γ-Formyloxymethyl-α,β-butyrolactone (**1b**) (15.52 g, 0.14 mol), Pd/C (5 wt% loading, 0.50 g, 3 wt% wrt. **1b**) and ethanol (30 mL) were mixed in an autoclave. The reaction was stirred under 70 bar H$_2$ for 18 h. The catalyst was removed by centrifuge at 3500 rpm, for 20 min. The supernatant was filtered through celite (3 mL) and ethanol removed by rotary evaporation. Conversion = 93% (determined by GC).

^{1}H NMR (400 MHz, CDCl$_3$): δ = 4.64-4.59 (m, 1H; CH), 3.87 (dd, J = 12.6, 2.8 Hz, 1H; C**H**$_2$OH), 3.62 (dd, J = 12.5, 4.6 Hz, 1H; C**H**$_2$OH), 3.03 (brs, 1H; OH), 2.65-2.47 (m, 2H; CH$_2$), 2.29-2.20 (m, 1H; CH$_2$), 2.17-2.08 (m, 1H; CH$_2$). ^{13}C NMR (100 MHz, CDCl$_3$, DEPT): δ = C 178.1, CH 160.6, 81.1, CH$_2$ 64.0, 28.8, 23.2. MS (ESI) m/z: [M]$^+$ Calcd for C$_6$H$_{10}$NaO$_4$ 139.0362; Found 139.0366. IR: $\tilde{\nu}$ = 3404 (OH), 1782 cm^{-1} (C=O).

(S)-γ-Methoxymethyl-α,β-butyrolactone (2c)

For the large scale reaction, (S)-γ-Hydroxymethyl-α,β-butyrolactone (**2b**) (175 g, 1.5 mol) was added dropwise to a solution of sodium hydride (30% dispersion in oil,

11 g, 2 equiv.) in dimethylformamide (anhydrous, 220 mL), over an ice-bath. The reaction was stirred at RT for 1h. Iodomethane was added dropwise (3 h) over an ice-bath. Reaction stirred overnight whilst allowing ice to melt 18 h. Reaction mixture was quenched with 15 g NH_4Cl(aq) and washed with chloroform (3 x 100 mL). Chloroform was removed by rotavap. The remaining liquid was distilled under vacuum using a Vigreux column (0.4 mbar, 75 °C). Conversion 78% from GC. ^{1}H NMR (400 MHz, CDCl$_3$): δ = 4.63 (m, 1H; CH), 3.60 (dd, J = 10.8, 3.4 Hz, 1H; C**H**$_2$OMe), 3.50 (dd, J = 10.8, 4.4 Hz, 1H; C**H**$_2$OMe), 3.39 (s, 3H; CH$_3$), 2.65-2.44 (m, 2H; CH$_2$), 2.33-2.23 (m, 1H; CH$_2$), 2.14-2.05 (m, 1H; CH$_2$). IR: $\tilde{\nu}$ = 1766 cm^{-1} (C=O).

5.3   Analytical Tests

p-Cymene stability test with activated silica

Phenylacetic acid, (1.6 g, 12 mmol) stirred in *p*-cymene for 1 h at 111 °C or reflux under an air atmosphere.

p-Cymene stability test with phenylacetic acid

Activated silica, (0.5 g) stirred in *p*-cymene for 1 h at 111 °C or reflux under an air atmosphere. Filtered through sintered glass funnel to remove silica.

p-Cymene stability test under argon

p-Cymene (20 mL) degassed with argon overnight. Phenylacetic acid, (1.6 g, 12 mmol) stirred in *p*-cymene for 1 h at reflux under an argon atmosphere.

Amide solubility tests

A saturated solution of each amide was formed in both toluene and *p*-cymene by stirring amide (0.1-0.2 g) in solvent (4-10 mL) at 25 °C for 24 h. Quantities were based on ensuring that there was still solid visible at the end of the 24 h. Stirring was subsequently stopped and the solutions kept at 25 °C for a further 24 h to allow all solid particles to settle. Approximately 0.5 mL of the upper liquid layer

was removed by a pre-heated Pasteur pipette and weighed into a vial along with a known amount of internal standard (tetradecane or hexadecane) in ethyl acetate or acetone depending on the solubility of the amide. Solution concentrations were determined by GC analysis.
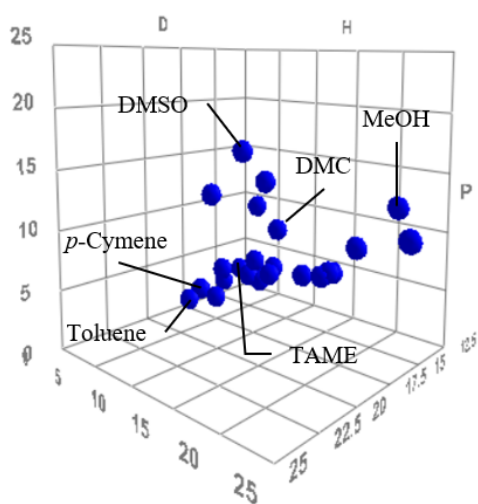
Lactone base sensitivity tests

Each lactone **2b** and **2c** were put into a vial (5 mmol) with base (0.07 mmol) and strred for 24 h at 25, 50 and 100 °C respectively. Bases used: pyridine, triethylamine, K_2CO_3 and blank with no base added. Results were obtained by TLC and NMR (see table 3.7).
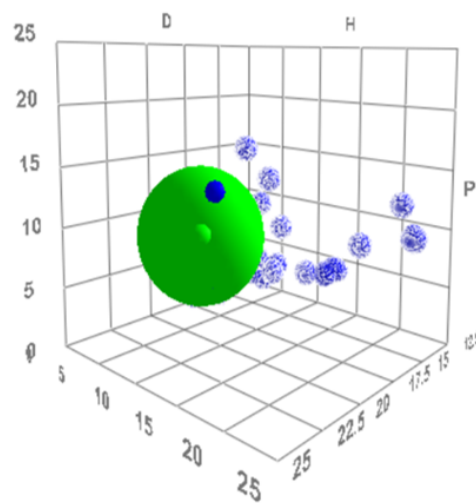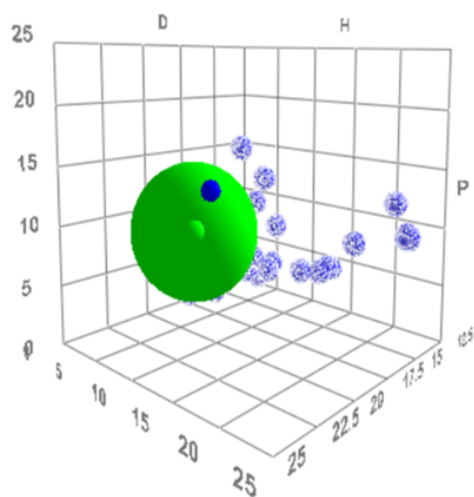
Appendices
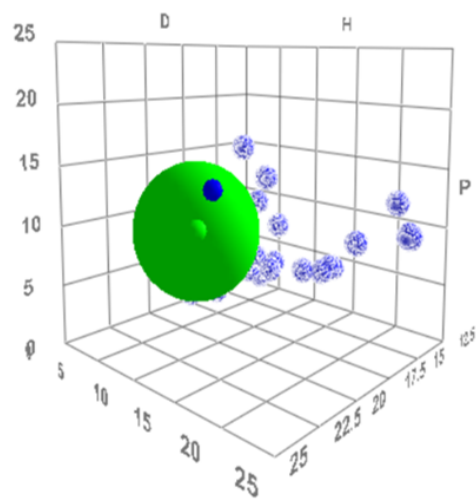
I HSP predictions for amides
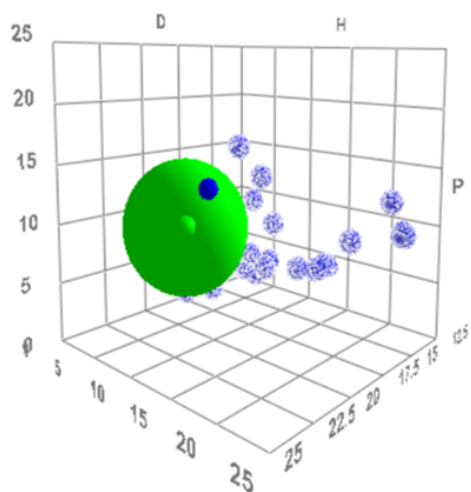

CHEM21 solvents + DMF and DMSO

2,*N*-Diphenylacetamide (**4**)

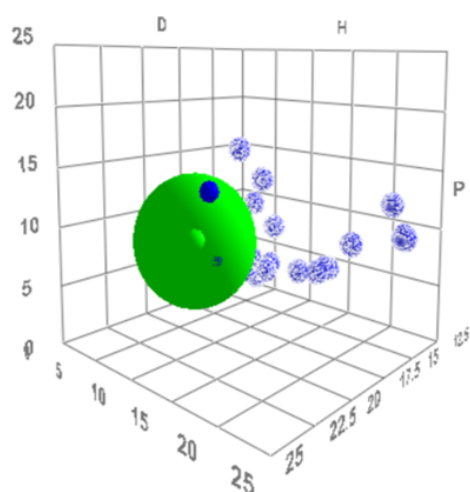*N*-Phenylbenzamide (**5**)

2-(4-Chlorophenoxy)-*N*-phenylacetamide (**6**)

N-(2-Chlorophenyl)-2-phenylacetamide (**7**)

N-(2,6-Dimethylphenyl)-2-phenylacetamide (**8**)
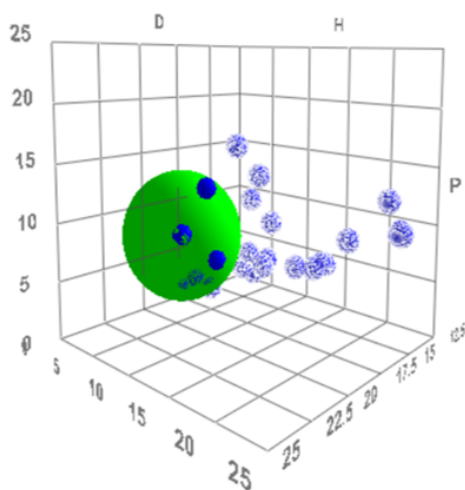
Phenylacetylpyrrolidine (**9**)
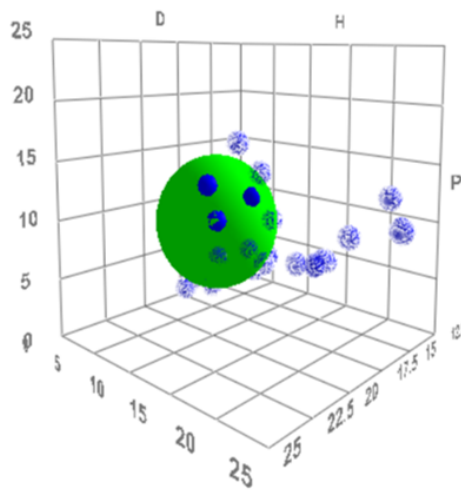
N-Butylbenzamide (**10**)
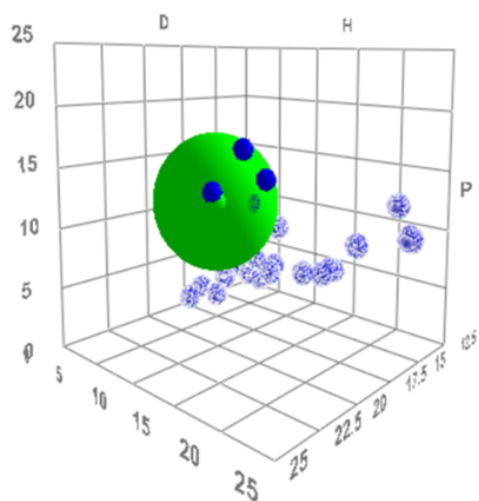
N-Hexyl-4-nitrobenzamide (**11**)
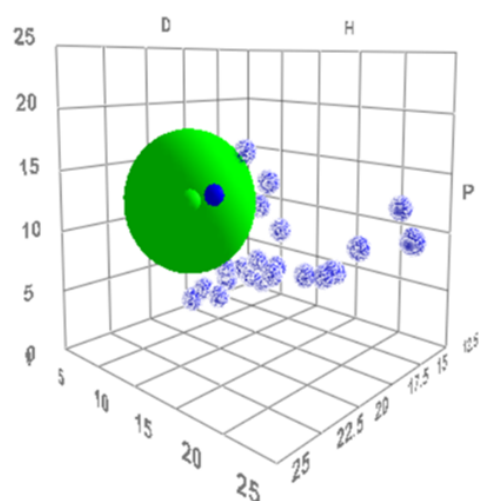
4-Nitro-N-phenylbenzamide (**12**)
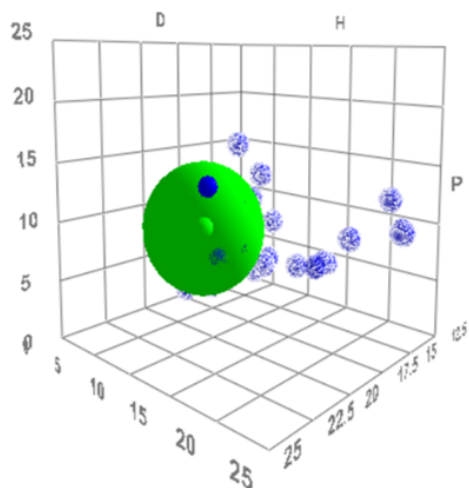
4-Nitro-N-phenylbenzamide (**12**)
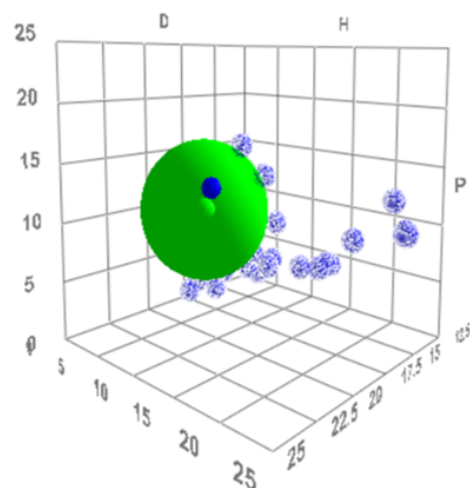
N-Hexylcyclohexanecarboxamide (**13**)

II   HSPiP predictions for all amides
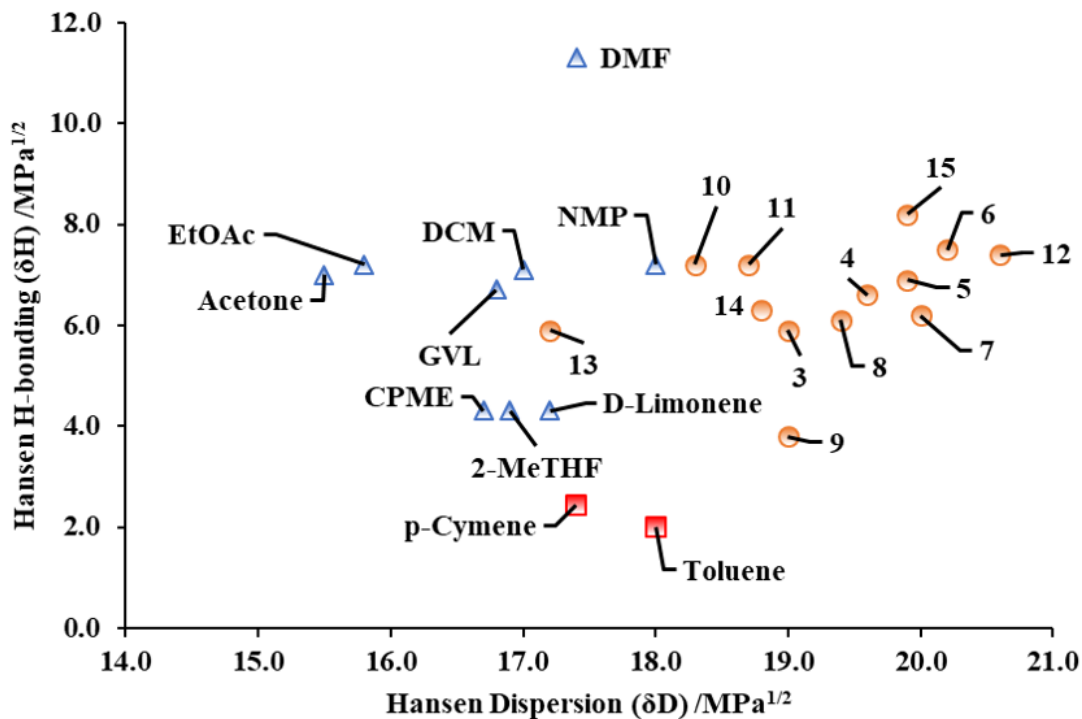
2D perspective HSPiP predictions δH vs. δD

2D perspective HSPiP predictions δP vs. δD

III   HSPiP predictions for all lactones
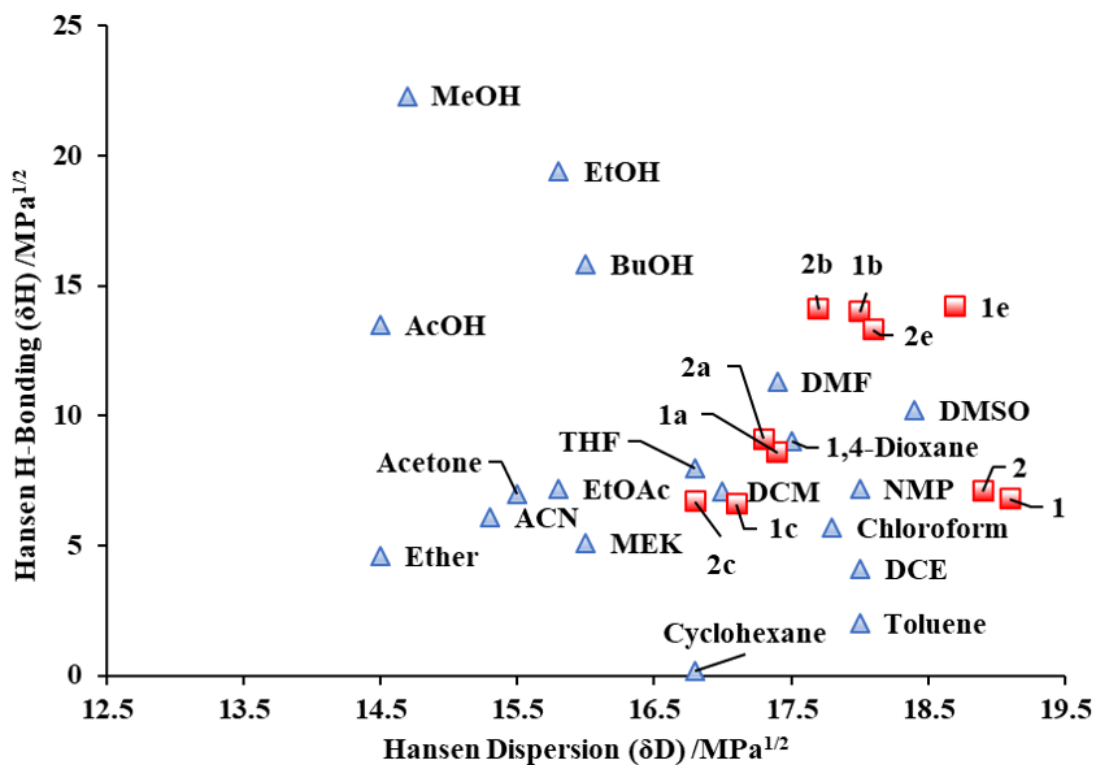
2D perspective HSPiP predictions δH vs. δD
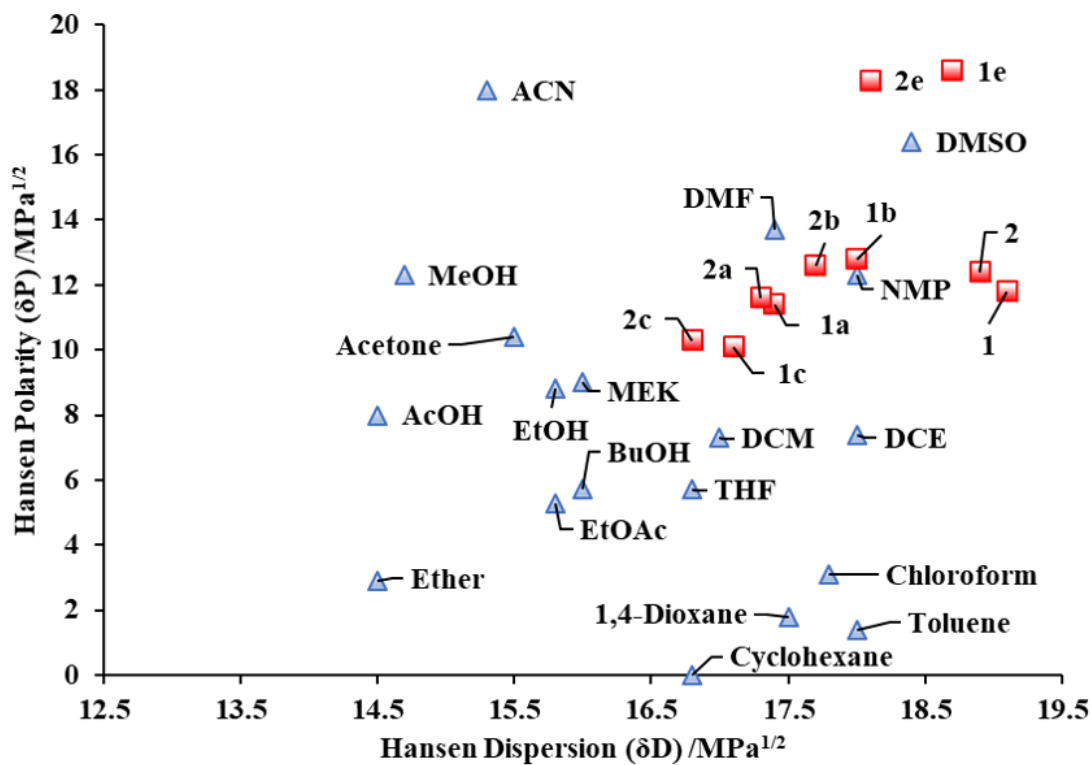
2D perspective HSPiP predictions δP vs. δD

IV Table of solubility data for amides 3-15
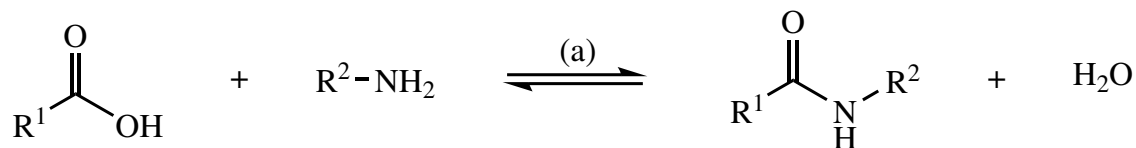
(a) Reaction conditions as described in experimental section 5.2.1. Amides isolated from amidation study and used in solubility testing.

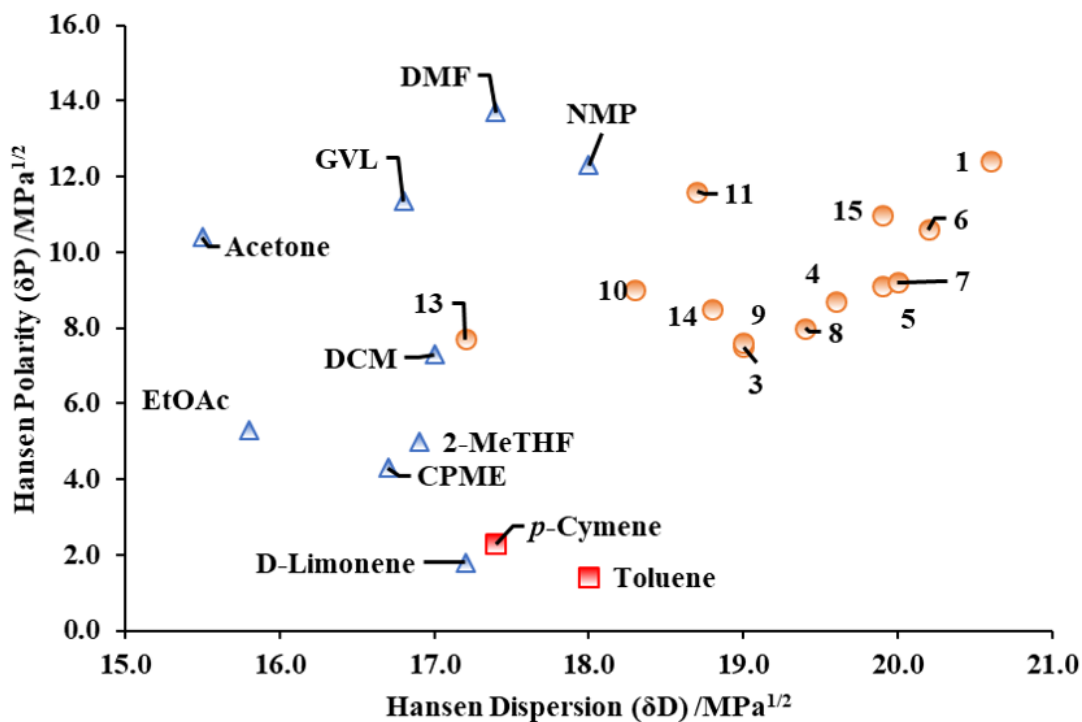| Amide	R^1	R^2	Amide solubility /mol kg^{-1} [a]	
			Toluene	p-Cymene
3	$C_6H_5CH_2CH_2CH_2$	C_6H_5	0.077^b	0.017
4	$C_6H_5CH_2$	C_6H_5	0.042	0.013
5	C_6H_5	C_6H_5	0.015	0.005
6	$ClC_6H_4OCH_2{}^c$	C_6H_5	0.073	0.021
7	$C_6H_5CH_2$	$ClC_6H_4{}^d$	0.085	0.045
8	$C_6H_5CH_2$	$C_6H_4(CH_3)_2{}^e$	0.029	0.000
9^f	$C_6H_5CH_2$	$C_4H_8{}^g$	high^h	high
10^i	C_6H_5	$CH_3CH_2CH_2CH_2$	high	high
11	$C_6H_4(NO_2)^j$	$CH_3(CH_2)_4CH_2$	0.052	0.004
12	$C_6H_4(NO_2)$	C_6H_5	0.000	0.000
13^f	$C_6H_{11}{}^k$	$CH_3(CH_2)_4CH_2$	high	high
14	C_6H_{11}	C_6H_5	0.031	0.011
15	$C_4H_3O^l$	C_6H_5	0.081	0.018

[a]Solubilities determined at 25 °C.
[b]Mean of three tests.
[c]4-Chlorophenoxyacetic acid.
[d]2-Chloroaniline
[e]2,6-Dimethylaniline.
[f]Non-precipitating, waxy solid.
[g]Pyrrolidine.
[h]Accurate measurement not possible by this method and with available quantities.
[i]Oil.
[j]4-Nitrobenzoic acid.
[k]Cyclohexane carboxylic acid.
[l]2-Furoic acid.

V GC calibration for amidation

GC calibration for the reaction between aniline and 4-phenylbutyric acid. FID response towards amide 3 calibrated against internal standard, hexadecane.

$y = 1.1032x$
$R^2 = 0.9998$
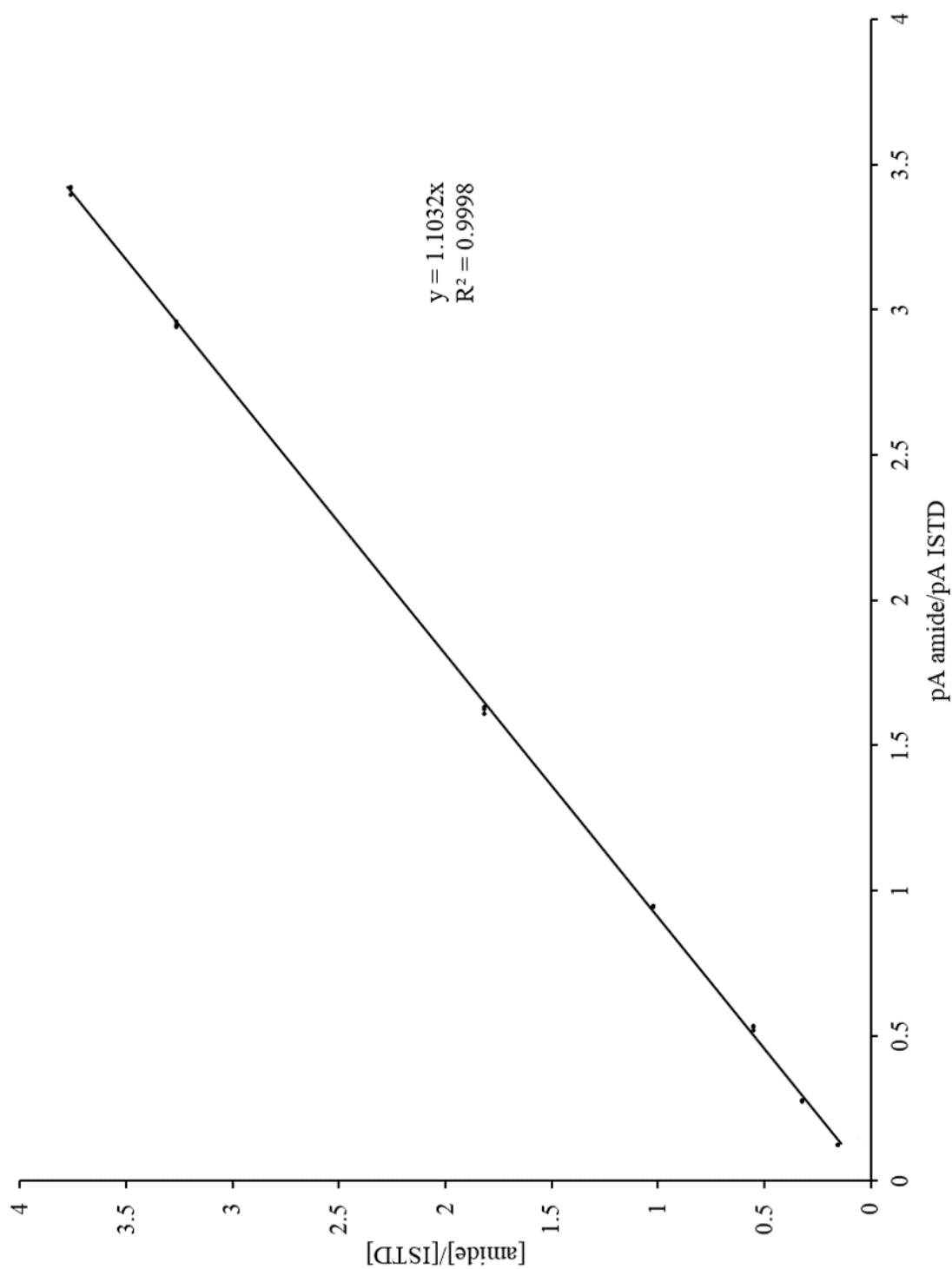
VI Gas Chromatogram of *N*-(2,6-Dimethylphenyl)-2-phenylacetamide (8)

Column: 30 m Rxi-5HT; method: General 1; c.f. section 2.3.2.
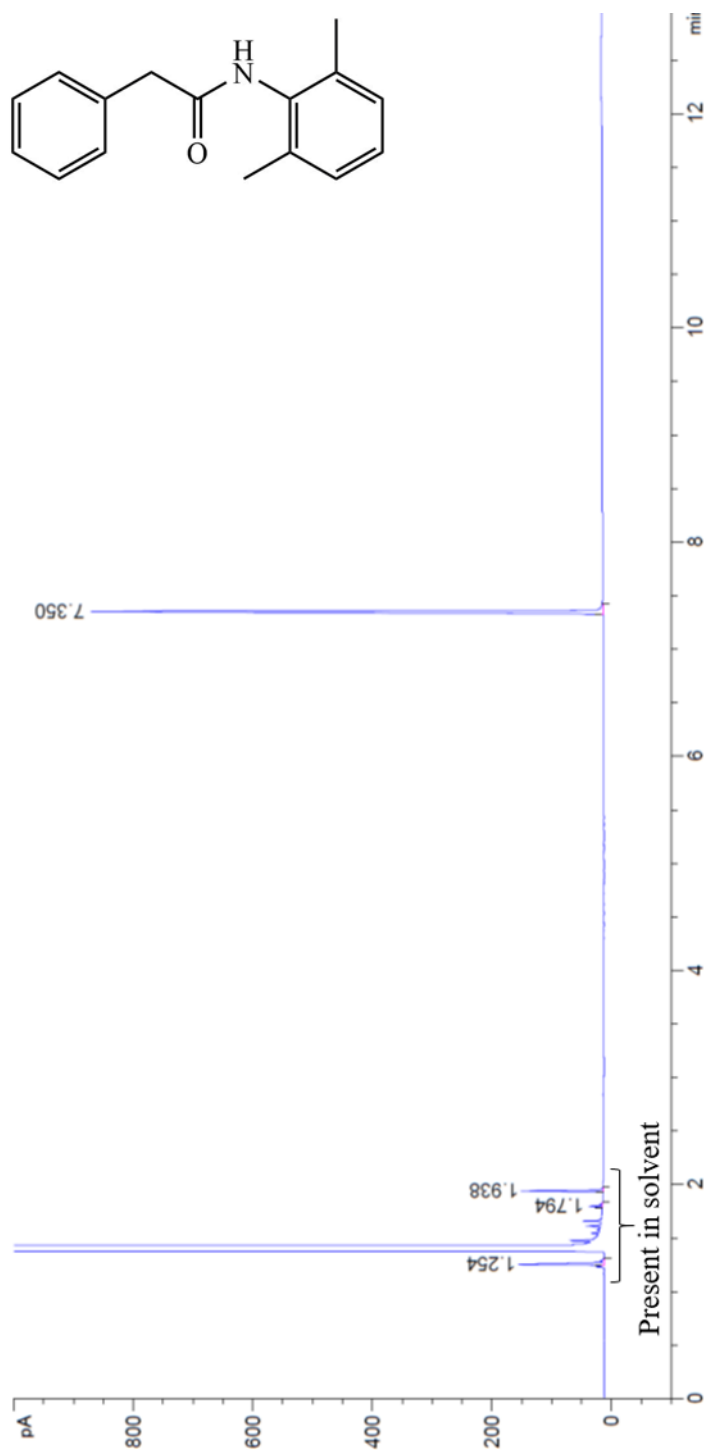
VII Gas Chromatogram of (S)-γ-Formyloxymethyl-α,β-butyrolactone (2a)

Crude product from m-CPBA oxidation of Cyrene. Column: 30 m Rxi-5HT; method: General 1; c.f. section 3.2.2.
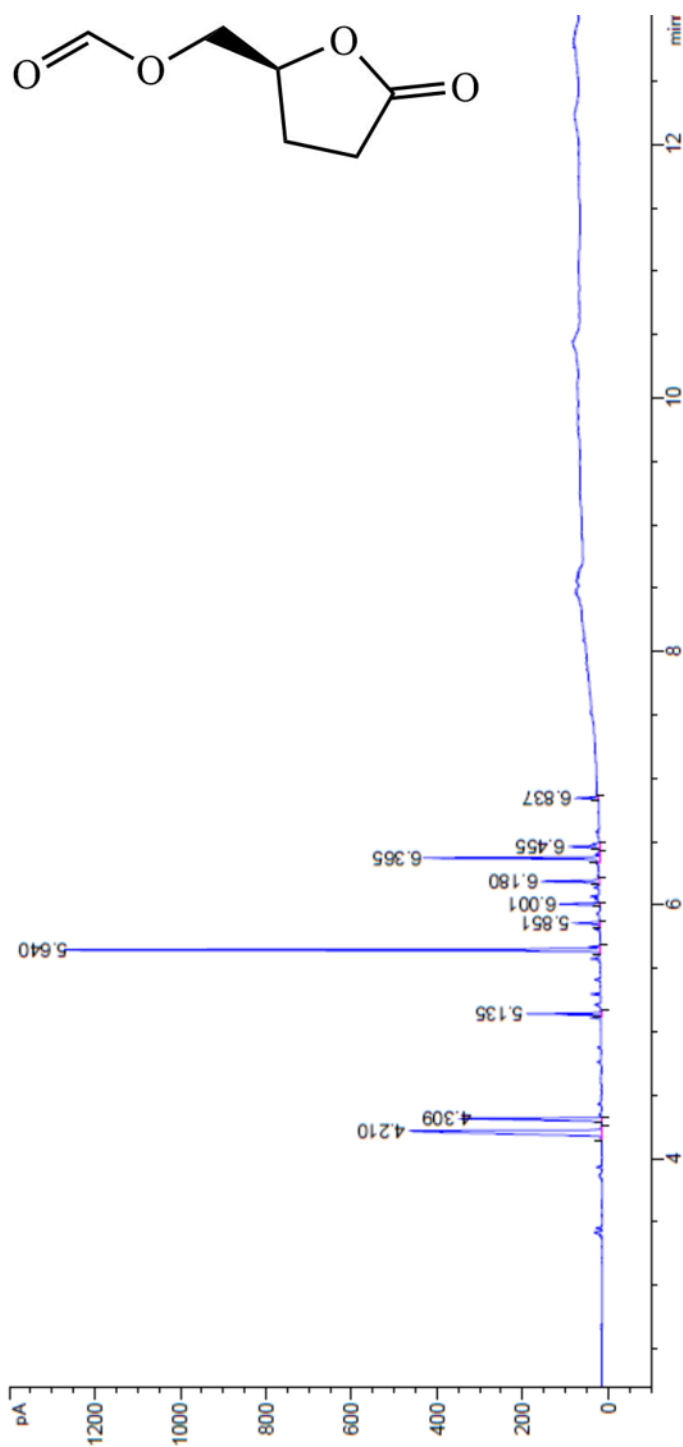
VIII Mass Spectrum (ESI) of (S)-γ-Formyloxymethyl-α,β-butyrolactone (2a)

c.f. section 3.2.2. ESI run by Karl Heaton, University of York Mass Spectrometry service.
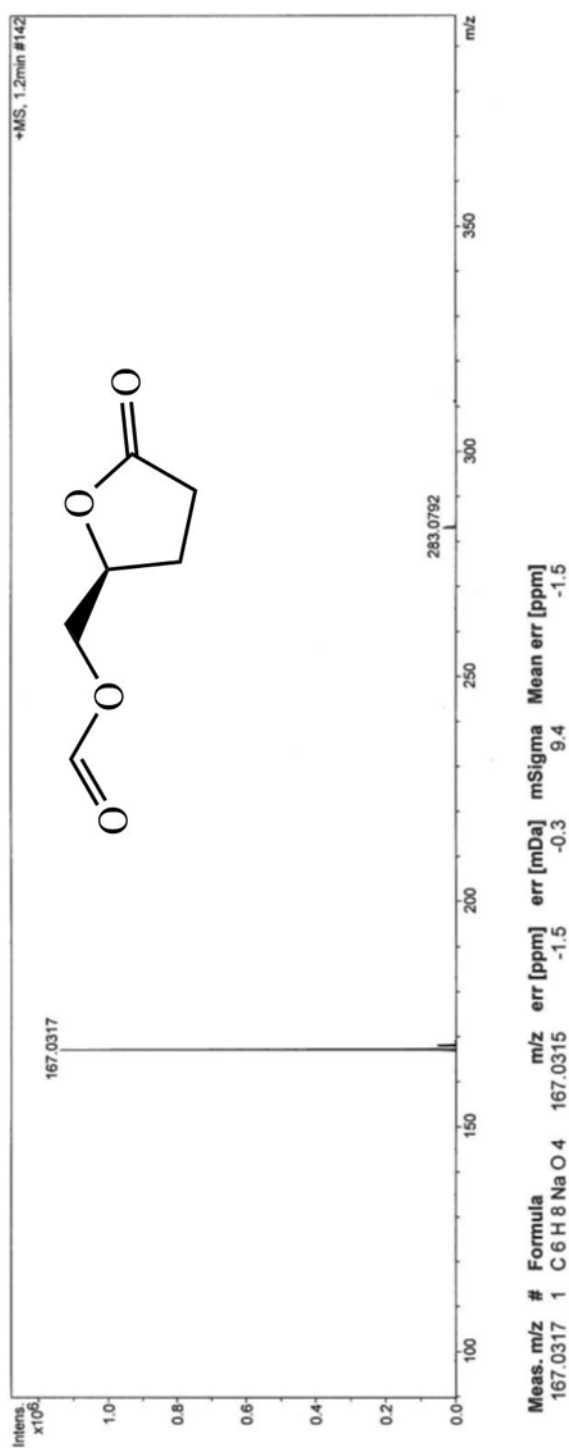
IX   Mass Spectrum (ESI) of S)-γ-Hydroxymethyl-α,β-butyrolactone (2b) after Kugelrohr distillation

c.f. section 3.6. ESI run by Karl Heaton, University of York Mass Spectrometry service.
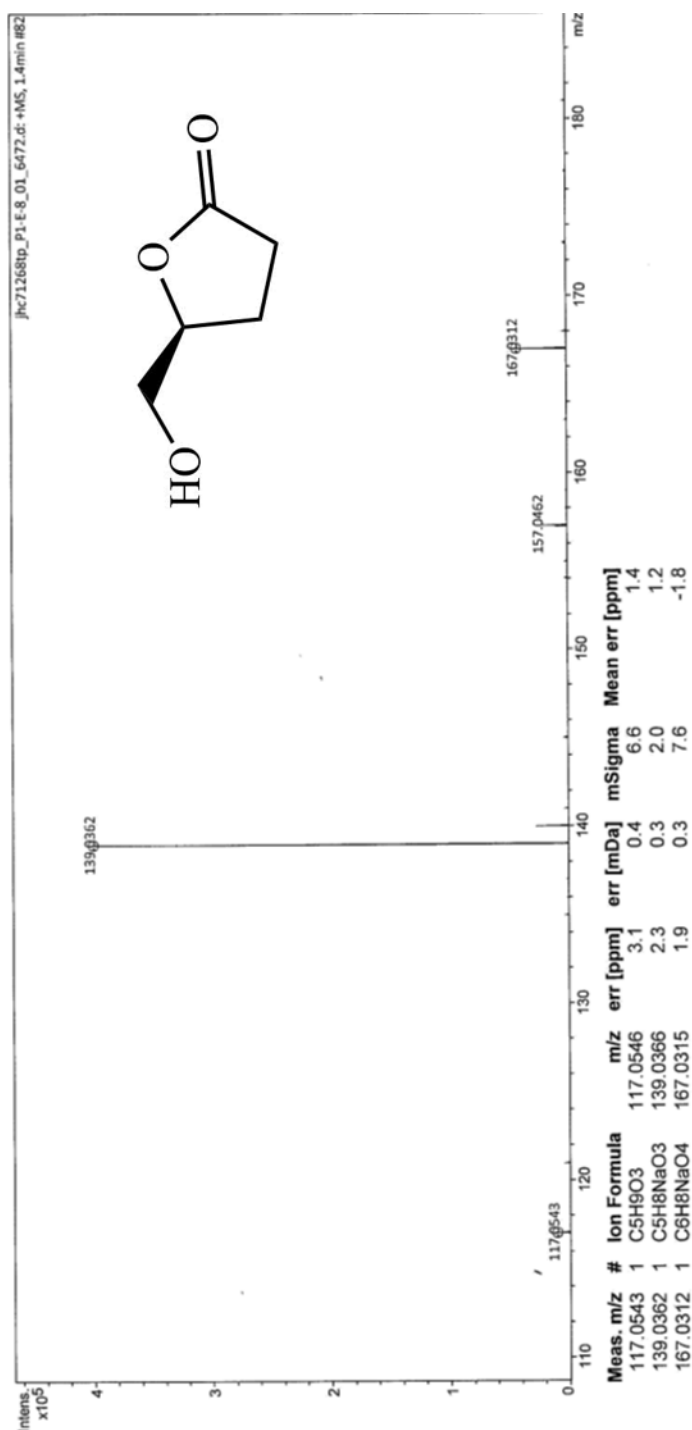
X   NMR Spectrum (^{1}H) of N-(2,6-Dimethylphenyl)-2-phenylacetamide (8)

400 MHz, CDCl$_3$, c.f. section 2.3.2 (Identification of Rotamer)
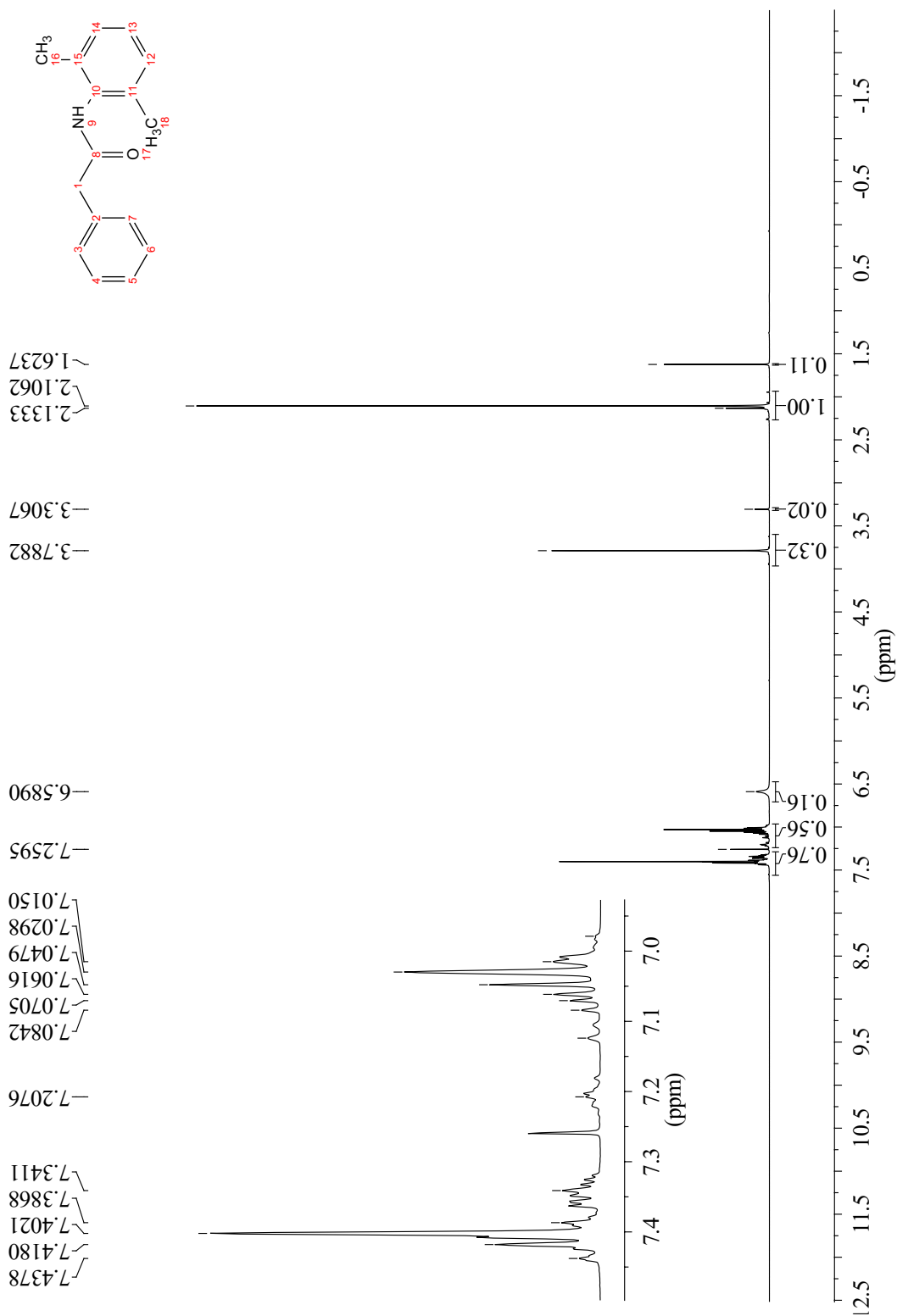
XI   NMR Spectrum (^{13}C) of *N*-(2,6-Dimethylphenyl)-2-phenylacetamide (8)

400 MHz, CDCl₃, c.f. section 2.3.2 (Identification of Rotamer)
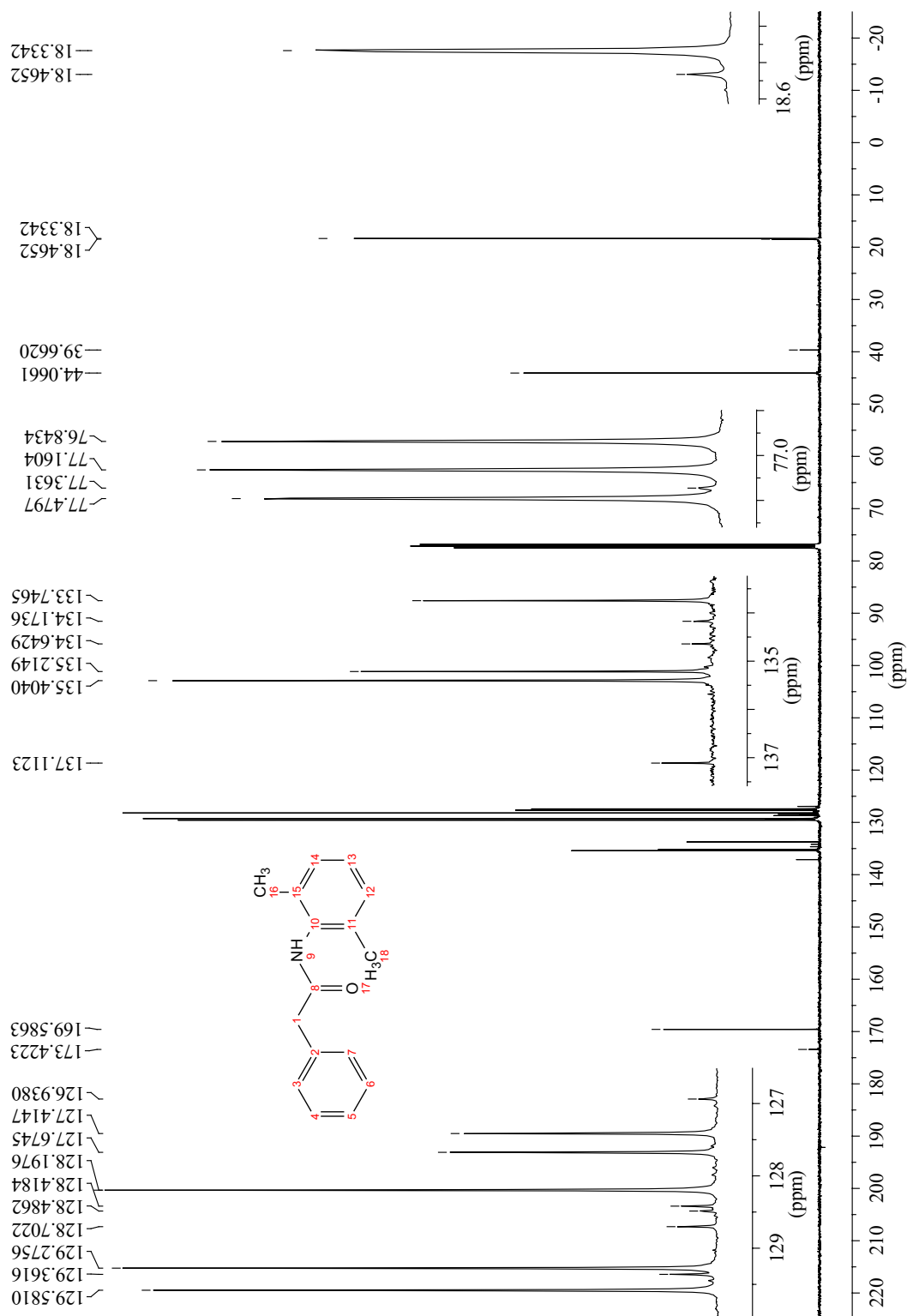
XII NMR Spectrum (^{1}H) of (S)-γ-Formyloxymethyl-α,β-butenolide (1a)

400 MHz, CDCl$_3$, c.f. section 3.2.1.
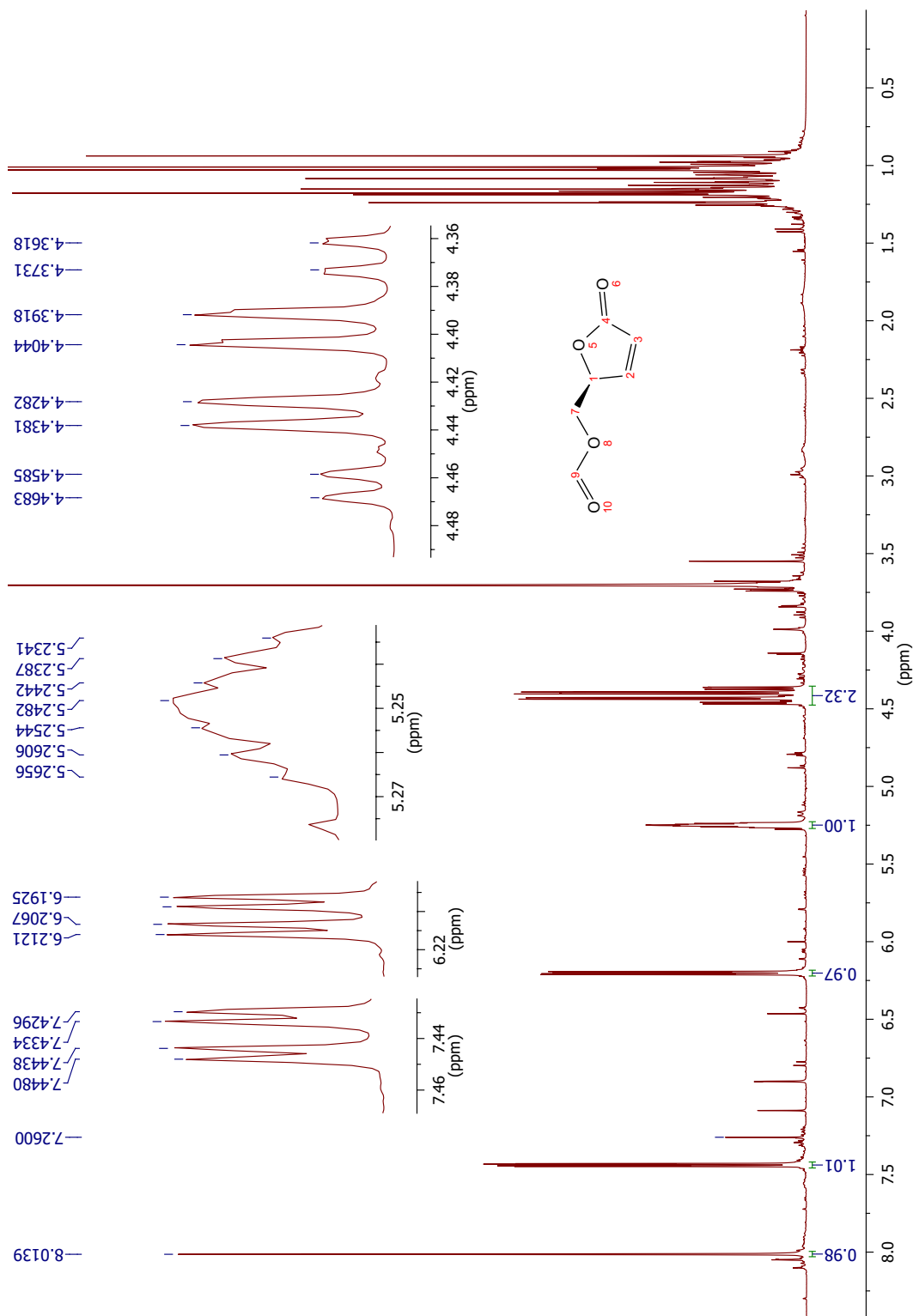
XIII NMR Spectrum (^{1}H) of (S)-γ-Formyloxymethyl-α,β-butyrolactone (2a)

400 MHz, CDCl$_3$, c.f. section 3.2.1.
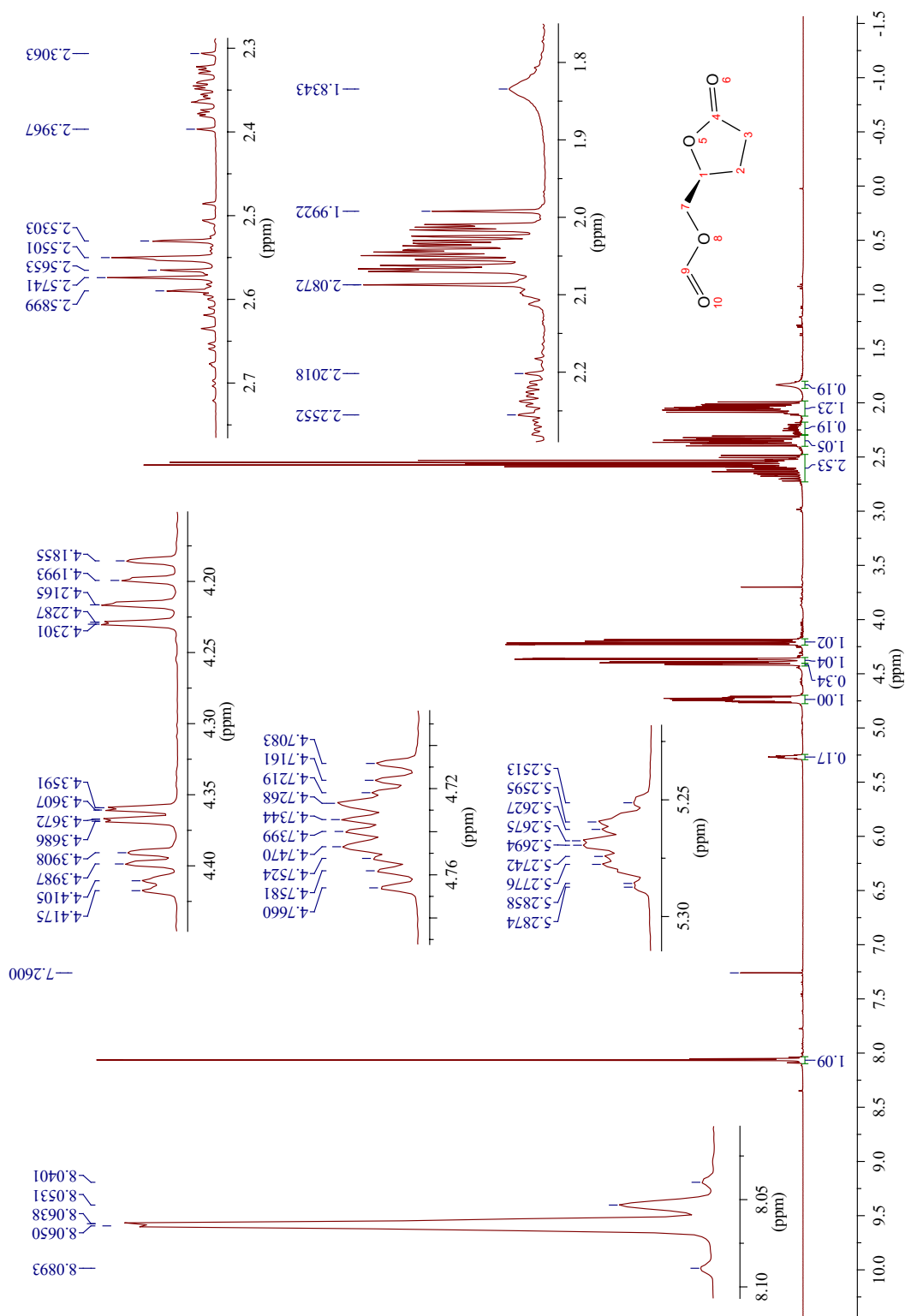
XIV   NMR Spectrum (HSQC) of (S)-γ-Formyloxymethyl α,β-butyrolactone 2a

400 MHz, CDCl$_3$, c.f. section 3.2.1.
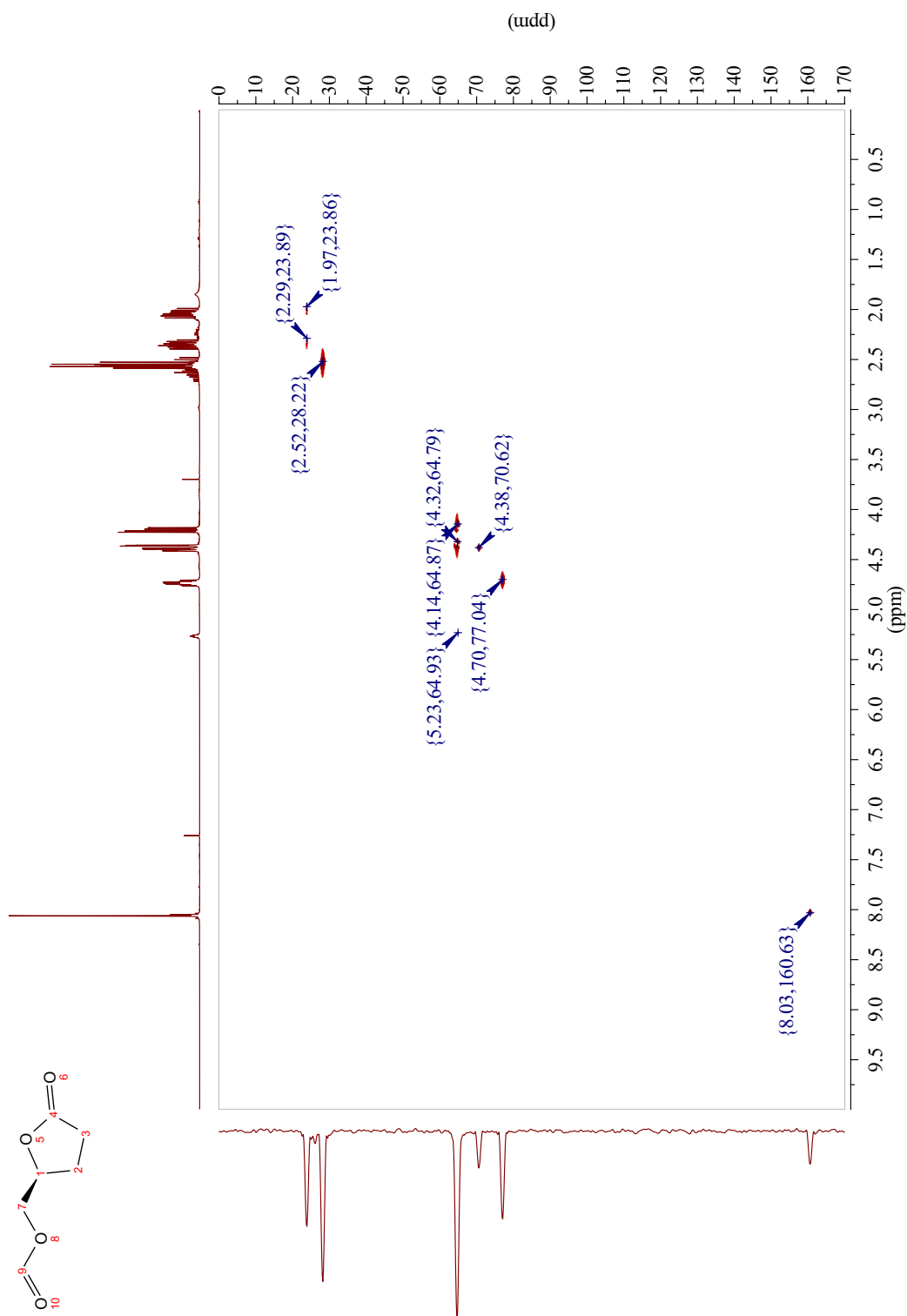
XV NMR Spectrum (^{1}H) of (S)-γ-Hydroxymethyl-α,β-butenolide (1b)

400 MHz, CDCl$_3$, c.f. section 3.2.1.
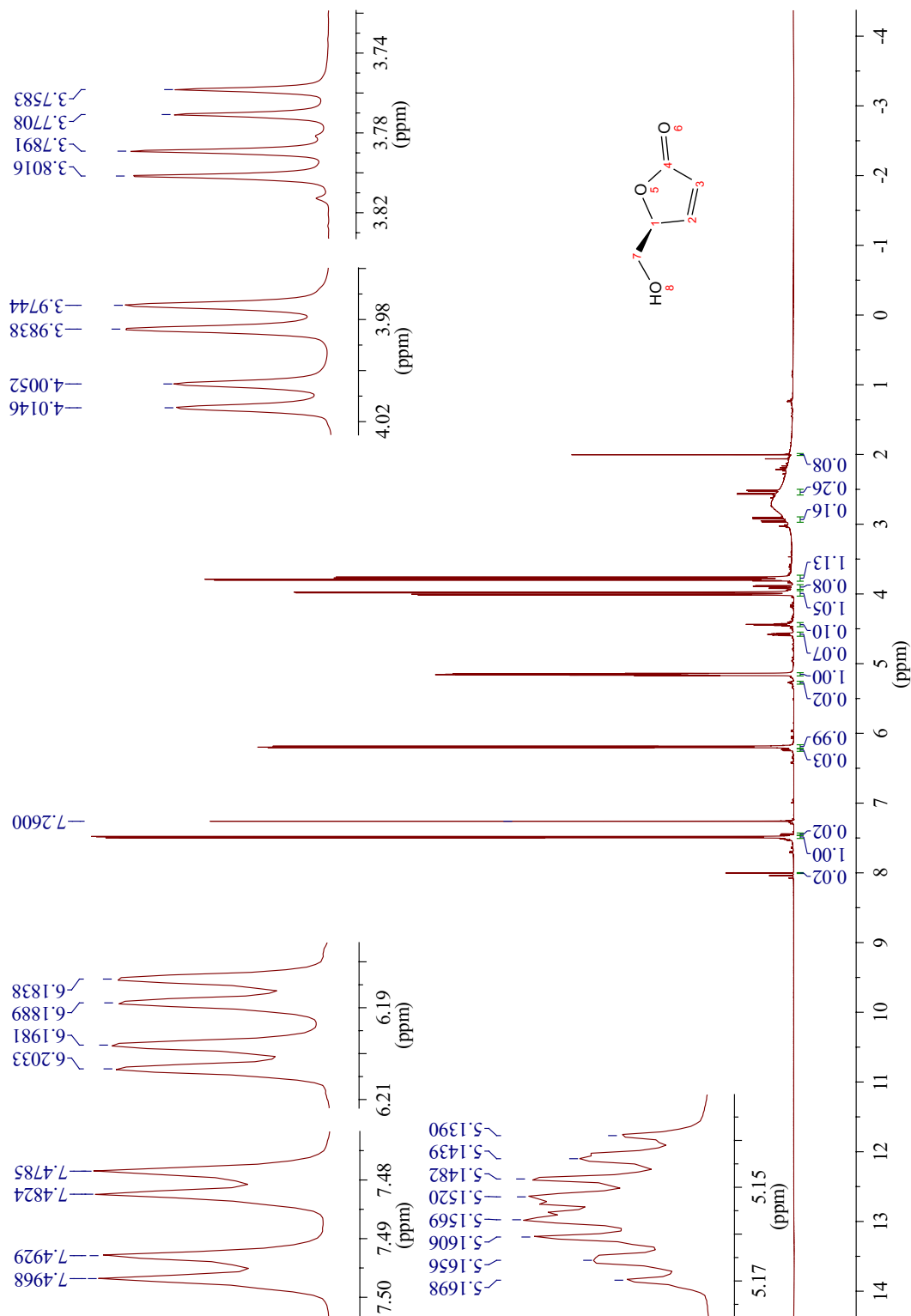
XVI   NMR Spectrum (^{1}H) of (S)-γ-Hydroxymethyl-α,β-butyrolactone (2b)

400 MHz, CDCl$_3$, c.f. section 3.2.1.
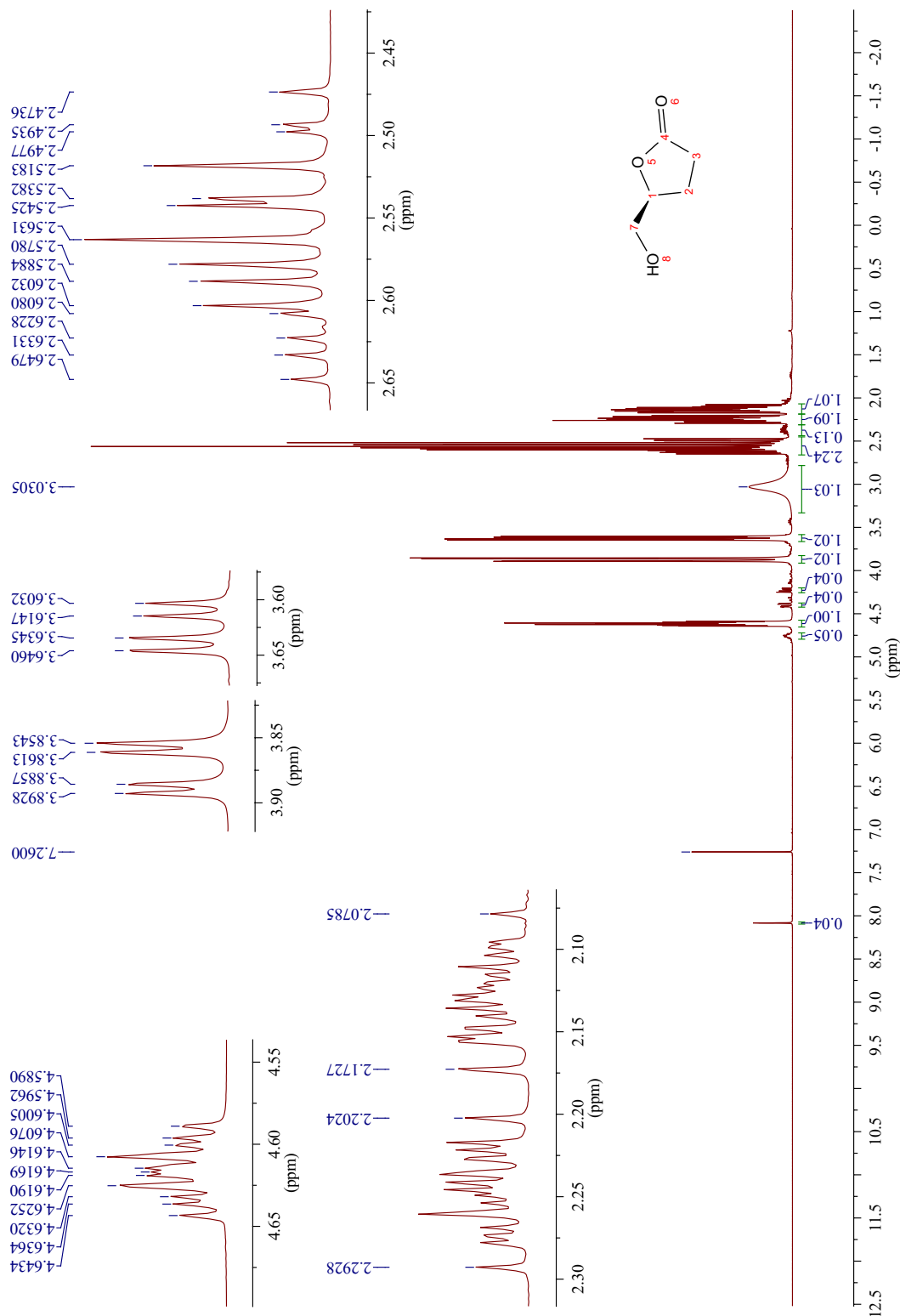
XVII NMR Spectrum (^{1}H) of 1b with 1a side-product

400 MHz, CDCl$_3$, product of L-lysine catalysed Baeyer-Villiger oxidation of LGO, 2 days at 25 °C, c.f. section 3.4.1.
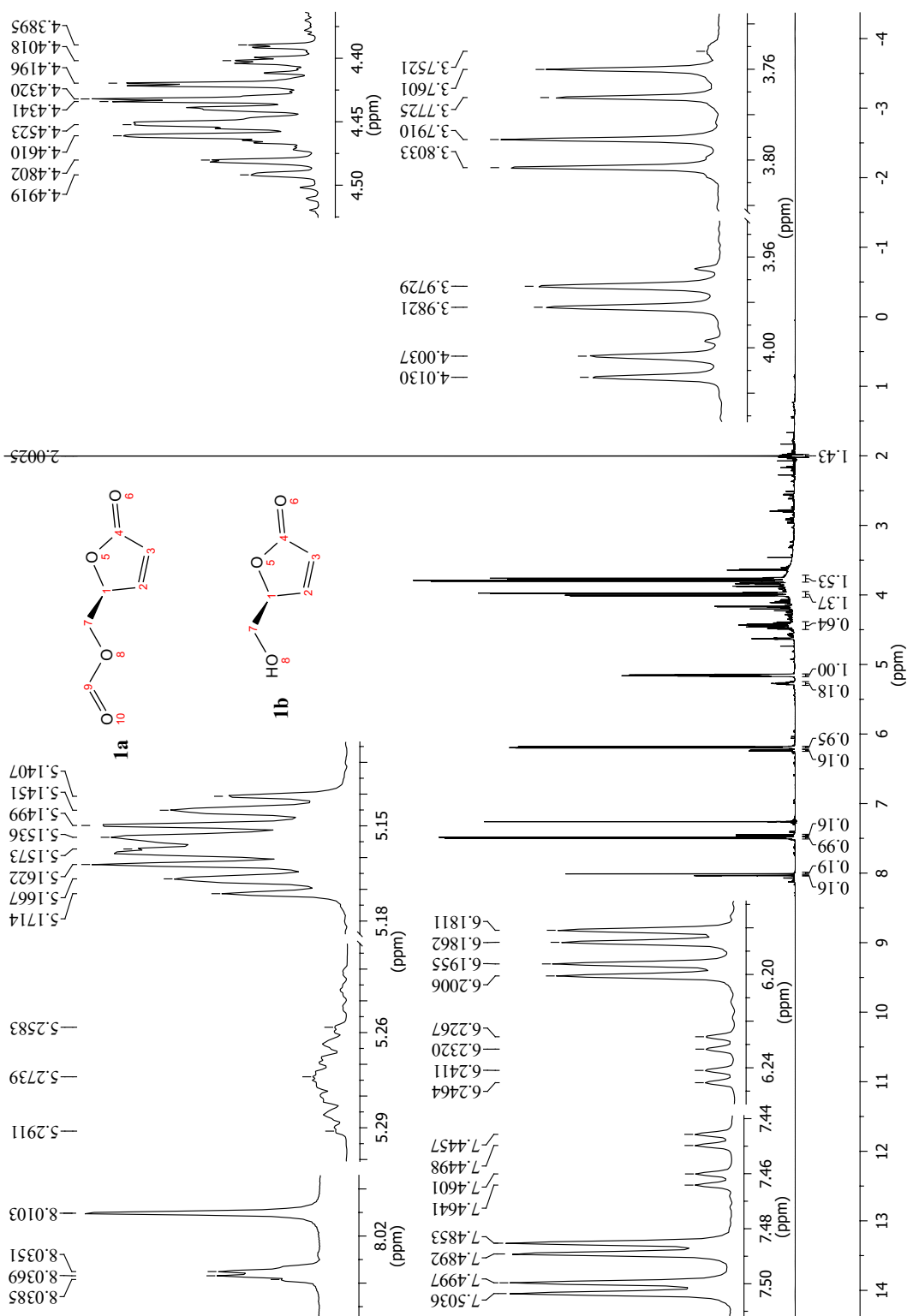
400 MHz, D$_2$O, product of hydrogenation of 1b, c.f. section 3.5.2.
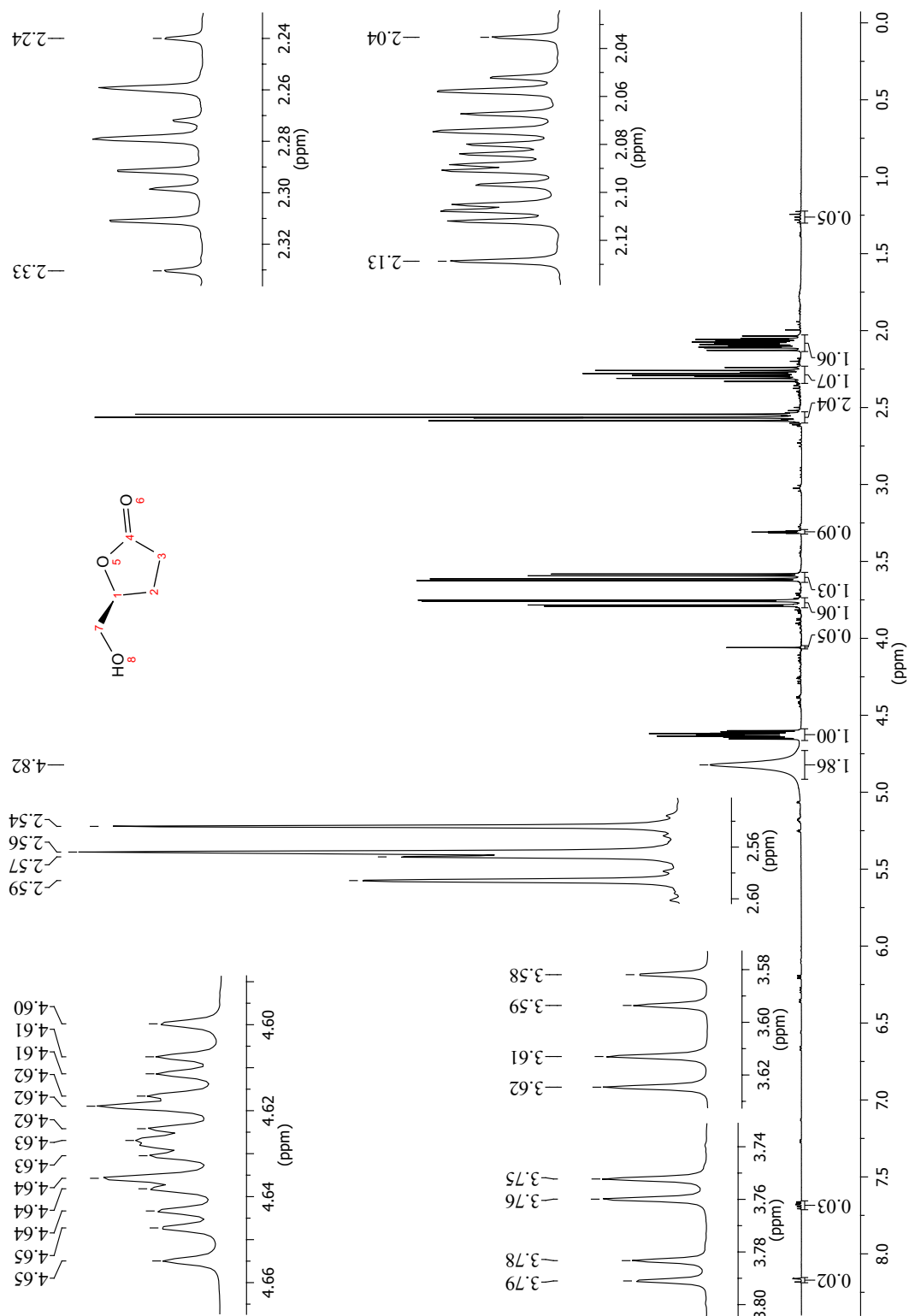
400 MHz, CDCl$_3$, c.f. section 3.2.1.
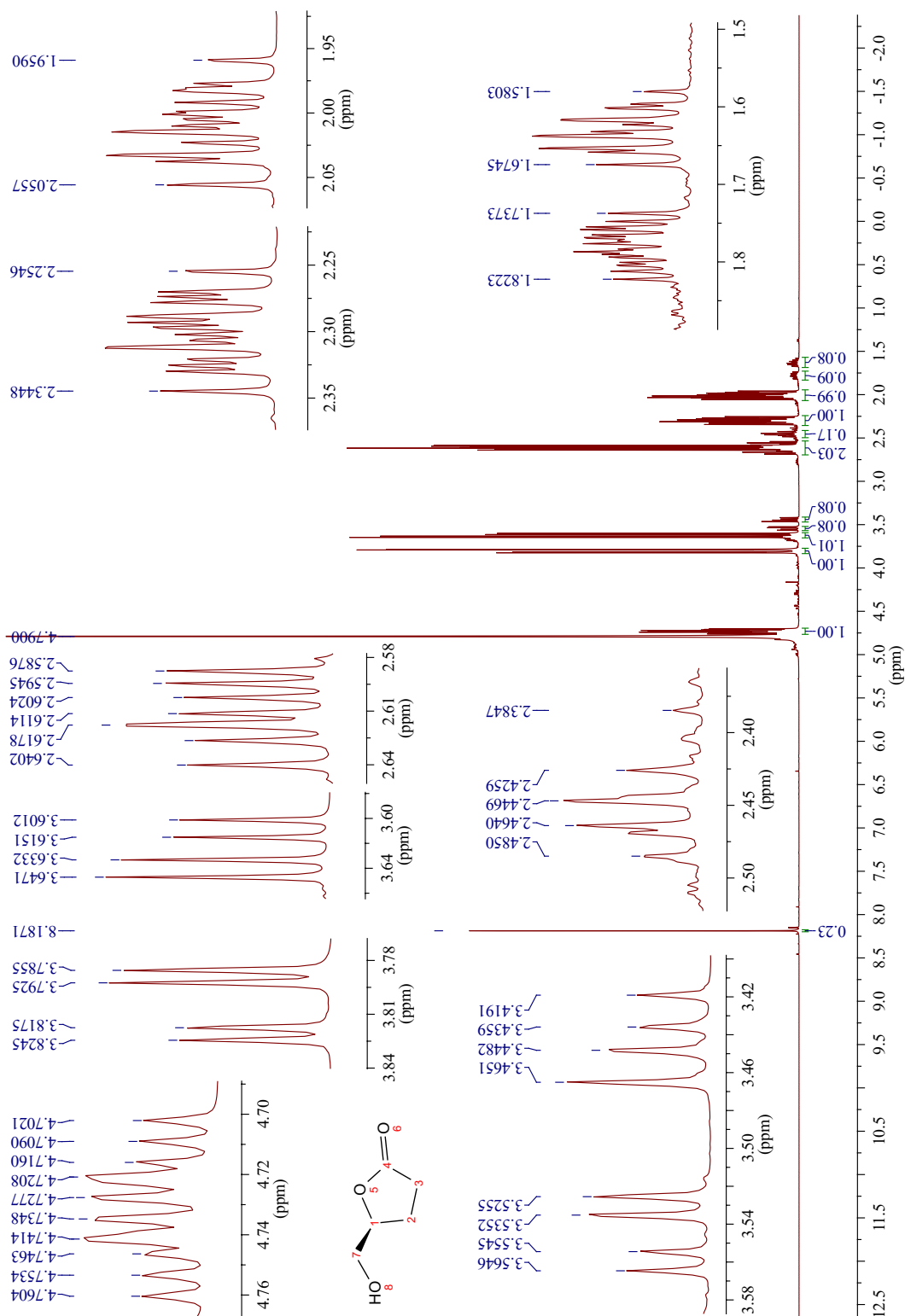
Abbreviations

c	Cohesive energy density
$CDCl_3$	Deuterated chloroform
CD_3OD	Deuterated methanol
CLFR	Crystal-liquid fugacity ratio
cm	Centimetre
cm^{-1}	Wavenumber
^{13}C NMR	Carbon-13 nuclear magnetic resonance
CV	Column volume
d	Doublet
DMAc	Dimethylacetamide
DMC	Dimethylcarbonate
DMF	Dimethylformamide
δ	Chemical shift
δD	Hansen dispersion
δH	Hansen H-bonding
δP	Hansen polarity
EI	Electron impact ionisation
ESI	Electrospray ionisation

EtOH	Ethanol
FI	Field ionisation
FID	Flame ionisation detection
g	Gram
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
h	Hour
^{1}H NMR	Hydrogen-1 nuclear magnetic resonance
HSP	Hansen Solubility Parameter
HSPiP	Hansen Solubility Parameters in Practice
IR	Infrared spectroscopy
J	Coupling constant
kg	Kilogram
LGO	Levoglucosenone
m-	Meta-
m	Multiplet
mAU	Milli absorbance unit
mbar	Millibar
m-CPBA	*meta*-Chloroperoxybenzoic acid
MeI	Methyl iodide
min	Minute
mg	Milligram
MHz	Megahertz

min	Minute
mL	Millilitre
mmol	Millimole
MOF	Metal-organic framework
mol	Mole
mp	Melting point
MPa	Megapascal
MT	Megatonne
NMP	N-Methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
NOx	Nitrous oxides
p-	Para-
ppm	Parts per million
RSD	Relative standard deviation
SD	Standard deviation
SOx	Sulfurous oxides
t	Triplet
t a^{-1}	Tonnes per annum
TLC	Thin layer chromatography
ttbp	Tri-$tert$-butylphenol
UV	Ultraviolet radiation
VT NMR	Variable temperature nuclear magnetic resonance
wt%	Percent by weight

Bibliography

Abdelmoty 1994   I. Abdelmoty, F. Albericio, L. A. Carpino, B. M. Foxman and S. A. Kates, *Lett. Pept. Sci.*, 1994, **1**, 57-67

Alfonsi 2008   K. Alfonsi, J. Colberg, P. Dunn, T. Fervig, S. Jennings, T. A. Johnson, H. P. Kleine, M. A. Nagy, D. A. Perry and M. Stefaniak, *Green Chem.*, 2008, **10**, 31-36

Allen 2011   C. L. Allen and J. M. J. Williams, *Chem. Soc. Rev.*, 2011, **40**, 3405-3415

Al-Warhi 2012   T. I. Al-Warhi, H. M. A. Al-Hazimi and A. El-Faham, *J. Saudi Chem. Soc.*, 2012, **16**, 97-116

Anastas 2013   P. T. Anastas, W. Leitner, P. G. Jessop, C.-J. Li, P. Wasserscheid, A. Stark, *Handbook of Green Chemistry - Green Solvents*, Wiley-VCH, Weinheim, 2013

Anderl 2018   F. Anderl, S. Größl, C. Wirtz and Alois Fürstner, *Angew. Chem. Int. Ed.*, 2018, **57**, 10712-10717

Arkhipenko 2018   S. Arkhipenko, M. T. Sabatini, A. S. Batsanov, V. Karaluka, T. D. Sheppard, H. S. Rzepad and Andrew Whiting, *Chem. Sci.*, 2018, **9**, 1058-1072

Ashcroft 2015   C. P. Ashcroft, P. J. Dunn, J. D. Hayler and A. S. Wells, *Org. Process Res. Dev.*, 2015, **19**, 740-747

Baer 1948   E. Baer and H. O. L. Fischer, *J. Am. Chem. Soc.*, 1948, **70**, 609-610

Berhal 2009   F. Berhal, S. Tardy, J. Pérard-Viret, and J. Royer, *Eur. J. Org. Chem.*, 2009, **2009**, 437-443

Bonneau 2018   G. Bonneau, A. A. M. Peru, A. L. Flourat and F. Allais, *Green Chem.*, 2018, **20**, 2455-2458

Breslow 2010   R. Breslow, in *Handbook of Green Chemistry Online*, ed. C.-J. Li, Wiler-VCH, Weinheim, 2010, vol. 5, ch. 1, pp. 1-29

Byrne 2016   F. P. Byrne, S. Jin, G. Paggiola, T. H. M. Petchey, J. H. Clark, T. J. Farmer, A. J. Hunt, C. R. McElroy and J. Sherwood, *Sustain. Chem. Process*, 2016, **4**, DOI 10.1186/s40508-016-0051-z

Byrne 2017   F. Byrne, B. Forier, G. Bossaert, C. Hoebers, T. J. Farmer, J. H. Clark and A. J. Hunt, *Green Chem.*, 2017, **19**, 3671-3678

Calvo-F. 2018   F. G. CalvoFlores, M. J. MonteagudoArrebola and J. A. Dobado, J. IsacGarcía, *Top. Curr. Chem.*, 2018, **18**, 376

Camp 2018   J. E. Camp, *ChemSusChem*, 2018, **11**, 3048-3055

Camps 1982   P. Camps, J. Cardellach, J. Font, R. M. Ortuño and O. Ponsati, *Tetrahedron*, 1982, **38**, 2395-2402

Carey 2006   J. S. Carey, D. Laffan, C. Thomson and M. T. Williams, *Org. Biomol. Chem.*, 2006, **4**, 2337-2347

Carotenuto 2013   G. Carotenuto, R. Tesser, M. Di Serio and E. Santacesaria, *Biomass Conv. Bioref.*, 2013, **3**, 55-67

Castro-Osma   J. A. Castro-Osma, J. W. Comerford, R. H. Heyn, M. North and E. Tangstad, *Faraday Discuss.*, 215, **183**, 19-30

Charville 2010   H. Charville, D. Jackson, G. Hodges and A. Whiting, *Chem. Commun.*, 2010, **46**, 1813-1823

Chu 1988   C. K. Chu, J. W. Beach, G. V. Ullas and Y. Kosugi, *Tetrahedron Lett.*, 1988, **29**, 5349-5352

Clark 2012	J. H. Clark, D. J. Macquarrie and J. Sherwood, *Green Chem.*, 2012, **14**, 90-93
Clark 2015	J. H. Clark, T. J. Farmer, A. J. Hunt and J. Sherwood, *Int. J. Mol. Sci.*, 2015, **16**, 17101-17159
Clark 2017	J. H. Clark, A. Hunt, C. Topi, G. Paggiola and J. Sherwood, *Sustainable Solvents*, Royal Society of Chemistry, London, 2017
Clarke 2018	C. J. Clarke, W.-C. Tu, O. Levers, A. Brohl and J. P. Hallett, *Chem. Rev.*, 2018, **118**, 747-800
Comba 2018	M. B. Comba, Y.-h. Tsai, A. M. Sarotti, M. I. Mangione, A. G. Suárez, and R. A. Spanevello, *Eur. J. Org. Chem.*, 2018, 590-604
Comerford 2009	J. W. Comerford, J. H. Clark, D. J. Macquarrie and S. W. Breeden, *Chem. Commun.*, 2009, 2562-2564
Comerford 2010	J. Comerford, PhD thesis, University of York, 2010
Comerford 2012	J. W. Comerford, T. J. Farmer, D. J. Macquarrie, S. W. Breeden and J. H. Clark, ARKIVOC, 2012, **7**, 282-293
Constable 2007	D. J. C. Constable, C. Jiminez-Gonzalez and R. K. Henderson, *Org. Process Res. Dev.*, 2007, **11**, 133-137
Coste 1990	J. Coste, D. Le-Nguyen and B. Castro, *Tetrahedron Lett.*, 1990, **31**, 205-208
Court 2012	G. R. Court, C. H. Lawrence, W. D. Raverty and A. J. Duncan, US pat., *US20120111714A1*, 2012
Cseri 2018	L. Cser, M. Razali, P. Pogany and G. Szekely, in *Green Chemistry: An Inclusive Approach*, ed. B. Török and T. Dransfield, Elsevier, Amsterdam, 1st edn, 2018, ch. 3.15, 513-553
Dàvila 2015	J. A. Dàvila, M. Rosenberg and C. A. Cardona, *Waste Biomass Valor*, 2015, **6**, 253-261

Diago-M. 1980   J. Diago-Meseguer, A. L. Palomo-Coll, J. R. Fernández-Lizarbe and A. Zugaza-Bilbao, *Synthesis*, 1980, **7**, 547–551

Dunetz 2016   J. R. Dunetz, J. Magano and G. A. Weisenburger, *Org. Process Res. Dev.*, 2016, **20**, 140-177

Durand 2011   M. Durand, V. Molinier, W. Kunz and J.-M. Aubry, *Chem. Eur. J.*, 2011, **17**, 5155-5164

EC 1999   Council Directive 1999/13/EC of 11 March 1999 on the limitation of emissions of volatile organic compounds due to the use of organic solvents in certain activities and installations

EC 2006   Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC

ECHA 2018a   Candidate list for Substances of Very High Concern (SVHC), http://echa.europa.eu/candidate-list-table (last accessed July 2018)

ECHA 2018b   Substance information, https://echa.europa.eu/substance-information/ (last accessed August 2018)

ECHA 2018c   Substances restricted under REACH, https://echa.europa.eu/substances-restricted-under-reach, (Last accessed September 2018)

Eckert 1996   C. A. Eckert, *Nature*, 1996, **383**, 313-318

El-Faham 2009   A. El-Faham, R. S. Funosas, R. Prohens and F. Alvericio, *Chem. Eur. J.*, 2009, **15**, 9404-9416

El-Faham 2011   A. El-Faham and F. Albericio, *Chem. Rev.*, 2011, **111**, 6557-6602

EPA 1998   EPA, *Federal Register*, National Volatile Organic Compound Emission Standards for Consumer Products, 1998, **23**, 48819-48847

EPA 2015   EPA Document# 740-R1-5002, TSCA Work Plan Chemical Risk Assessment, N-Methylpyrrolidone: Paint Stripper Use, 2015

EPA 2017   EPA Docket EPA-HQ-OPPT-2016-0231, Regulation of Certain Uses under Toxic Substances Control Act: Methylene Chloride and N-Methylpyrrolidone, 2017

Farmer 2013   T. J. Farmer, unpublished work

Fierz   H. Fierz, *Assessment of thermal safety during distillation of DMSO*, IChemE Symposium series no. 134, 1994

Flourat 2015   A. L. Flourat, A. A. M. Peru, A. R. S. Teixeira, F. Brunissen and F. Allais, *Green Chem.*, 2015 **17**, 404-412

FMI 2015   Future Market Insights, https://www.futuremarketinsights.com/reports/global-n-methyl-2-pyrrolidone-market (last accessed August 2018)

Fung 1972   D. P. C. Fung, Y. Tsuchiya and K. Sumi, *Wood Sci.*, 1972, **5**, 38-43

Ghandi 2014   K. Ghandi, *Green and Sustainable Chemistry*, 2014, **4**, 44-53

Gissi 2015   A. Gissi, A. Lombardo, A. Roncaglioni, D. Gadaleta, G. F. Mangiatordi, O. Nicolotti and E. Benfenati, *Environmental Research*, 2015, **137**, 398–409

Gu 2013   Y. Gu and F. Jérôme, *Chem. Soc. Rev.*, 2013 **42**, 9550-9570

Halpern 1973   Y. Halpern, R. Riffer and A. Broido, *J. Org. Chem.*, 1973, **38**, 204-209

Han 2004   S.-Y. Han and Y.-A. Kim, *Tetrahedron*, 2004, **60**, 2447-2467

Hansen 1967   C. M. Hansen, PhD thesis, The Technical University of Denmark, 1967

Hansen 2018   Hansen Solubility Parameters Online, https://www.hansen-solubility.com/, (last accessed September 2018)

Henderson 2011   R. K. Henderson, C. Jiminez-Gonzalez, D. J. C. Constable, S. R. Alton, G. G. A. Inglis, G. Fisher, J. Sherwood, S. P. Binks and A. D. Curzons, *Green Chem.*, 2011, **13**, 854-862

Herrero 2017   M. Herrero, J. A. Mendiola, E. Ibanez, *Current Opinion in Green and Sustainable Chemistry*, 2017, **5**, 24-30

Hildebrand 1936   J. H. Hildebrand, *Solubility of Non-electrolytes*, Reinhold Publishing Coorporation, New York, 1936

Hock 1944   H. Hock and S. Lang, *Chemische Berichte*, 1944, **77**, 257-264

Hulsbosch 2016   J. Hulsbosch, D. E. De Vos, K. Binnemans and R. Ameloot, *ACS Sustainable Chem. Eng.*, 2016, **4**, 2917-2931

Hunt 2018   A. J. Hunt and T. M. Attard, *Supercritical and Other High-Pressure Solvent Systems*, Royal Society of Chemistry, Camridge, 2018

IARC 2018   IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, https://monographs.iarc.fr/list-of-classifications-volumes, (Last accessed September 2015)

Ishihara 1996   K. Ishihara, S. Ohara and H. Yamamoto, *J. Org. Chem.*, 1996, **61**, 4196-4197

Ismalaj 2014   E. Ismalaj, G. Strappaveccia, E. Ballerini, F. Elisei, O. Piermatti, D. Gelman, and L. Vaccaro, *ACS Sustainable Chem. Eng.*, 2014, **2**, 2461-2464

IUPAC 2019a   IUPAC Compendium of Chemical Terminology, https://goldbook.iupac.org/html/S/S05744.html, (Last accessed February 2019)

IUPAC 2019b	IUPAC Compendium of Chemical Terminology, https://goldbook.iupac.org/html/S/S05746.html, (Last accessed February 2019)
Jabri-K 2016	I. Jabri Karoui, K. Msaada, M. Abderrabba and B. Marzouk, *J. Agr. Sci. Tech.*, 2016, **18**, 79-91
James 2012	S. L. James, C. J. Adams, C. Bolm, D. Braga, P. Collier, T. Friscić, F. Grepioni, K. D. M. Harris, G. Hyett, W. Jones, A. Krebs, J. Mack, L. Maini, A. G. Orpen, I. P. Parkin, W. C. Shearouse, J. W. Steed and D. C. Waddell, *Chem. Soc. Rev.*, 2012, **41**, 413-447
Jankowski 2008	M. D. Jankowski, C. S. Henry, L. J. Broadbelt and V. Hatzimanikatis, *Biophysical Journal*, 2008, **95**, 1487-1499
Jenkins 1933	S. S. Jenkins and E. M. Richardson, *J. Am. Chem. Soc.*, 1933, **55**, 1618–1621
Jeong 1993	L. S. Jeong, J. W. Beach and C. K. Chu, *J. Heterocyclic Chem.*, 1993, **30**, 1445-1452
Jessop 2005	P. G. Jessop, D. J. Heldebrant, L. Xiaowang, C. A. Eckert and C. L. Liotta, *Nature*, 2005, **436**, 1102
Jessop 2011	P. G. Jessop, *Green Chem.*, 2011, **13**, 1391-1398
Jessop 2012	P. G. Jessop, S. M. Mercera and D. J. Heldebrant, *Energy Environ. Sci.*, 2012, **5**, 7240-7253
Josephson 1929	K. Josephson, *Chem. Ber.*, 1929, **62**, 313-316
Kawakami 1990	H Kawakami, T. Ebata, K. Koseki, K. Matsumoto, H. Matsushita, Y. Naoi and K. Itoh, *HeteroCycles*, 1990, **31**, 2041-2054
Kerton 2013	F. M. Kerton and R. Marriott, *Alternative Solvents for Green Chemistry*, Royal Society of Chemistry, Cambridge, 2nd edn, 2013
Kim 2011	J. Kim, S. Jung, S. Park and S. Park, *Tetrahedron Lett.*, 2011, **52**, 2866-2868

Klamt 1993   A. Klamt and G. Schüürmann, *J. Chem. Soc. Perkin Trans. 2*, 1993, **0**, 799-805

Klamt 1995   A. Klamt, *J. Phys. Chem.*, 1995, **99**, 2224-2235

Klamt 2018   A. Klamt, *WIREs Comput Mol Sci*, 2018, **8**, 1-7

Koseki 1990   K. Koseki, T. Ebata, H. Kawakami, H. Matsushita, Y. Naoi and K. Itoh, *HeteroCycles*, 1990, **31**, 423 - 426

Kundrotaite 2016   M. Kundrotaite, MSc Dissertation, University of York, 2016

Laird 2012   T. Laird, *Org. Process Res. Dev.*, 2012 **16**, 1-2

Lam 2006   T. T. Lam, T. Vickery and L. Tuma, *J. Therm. Anal. Cal.*, 2006, **85**, 25-30

Lamarche 2017   M. Lamarche, M. T. Dang, J. Lefebvre, J. D. Wuest and S. Roorda, *ACS Sustainable Chem. Eng.*, 2017, **5**, 5994-5998

Lebedyeva 2014   I. O. Lebedyeva, S. Biswas, K. Goncalves, S. M. Sileno, A. R. Jackson, K. Patel, P. J. Steel and A. R. Katritzky, *Chem. Eur. J.*, 2014, **20**, 11695–11698

Le Cloirec 2012   P. Le Cloirec, *Rev. Environ. Sci. Biotechnol.*, 2012, **11**, 381-392

Liu 2000   Z.-Y. Liu, J.-X. Ji and B.-G. Li, *J. Chem. Soc., Perkin Trans.*, 2000, **1**, 3519-3521

Liu 2017   X. Liu, W. Xu, M.-T. Zeng, M. Liu, C.-Z. Chang, H. Zhu and Z.-B. Dong, *J. Chem. Res.*, 2017, **41**, 484–486

Lundberg 2014   H. Lundberg, F. Tinnis, N. Selander and H. Adolfsson, *Chem. Soc. Rev.*, 2014, **43**, 2714-2742

MacMillan 2013   D. S. MacMillan, J. Murray, H. F. Sneddon, C. Jamieson and A. J. B. Watson, *Green Chem.*, 2013, **15**, 596-600

Malag. 2012   G. Malaguarnera, E. Cataudella, M. Giordano, G. Nunnari, G. Chisari and M. Malaguarnera, *World J. Gastroenterol.*, 2012, **18**, 2756-2766

Marchese 2017   A. Marchese, C. R. Arciola, R. Barbieri, A. S. Silva, S. F. Nabavi, A. J. T. Sokeng, M. Izadi, N. J. Jafari, I. Suntar, M. Daglia and S. M. Nabavi, *Materials*, 2017, **10**, 947-961

Martin-L. 2010   M. A. Martin-Luengo, M. Yates, E. S. Rojo, D. Huerta Arribas, D. Aguilar and E. Ruiz Hitzky, *Appl. Catal. A Gen.*, 2010, **387**, 141-146

McDonald 2016   C. E. McDonald, J. D. Ramsey, C. C. McAtee, J. R. Mauck, E. M. Hale and J. A. Cumens, *J. Org. Chem.*, 2016, **81**, 5903–5914

Miftakhov 1994   M. S. Miftakhov, F. A. Valeev and I. N. Gaisina, *Russ. Chem. Rev.*, 1994, **63**, 869-882

Miller 1961   J. Miller and A. J. Parker, *J. Am. Chem. Soc.*, 1961, **83**, 117-123

Mistry 2017   L. Mistry, K. Mapesa, T. W. Bousfield and J. E. Camp, *Green Chem.*, 2017, **19**, 2123-2128

Moity 2012   L. Moity, M. Durand, A. Benazzouz, C. Pierlot, V. Molinier and J.-M. Aubry, *Green Chem.*, 2012, **14**, 1132-1145

Montalb. 2005   C. A. G. N. Montalbetti and V. Falque, *Tetrahedron*, 2005, **61**, 10827–10852

Mukherjee 1971   R. Mukherjee, *Chem. Commun.*, 1971, **625**, 1113–1114

Nammalw. 2015   B. Nammalwar, N. P. Muddala, F. M. Watts and R. A. Bunce, *Tetrahedron*, 2015, **71**, 9101–9111

Nemoto 1985   H. Nemoto, M. Nagai, and K. Fukumoto, *J. Org. Chem.*, 1985, **50**, 2764-2766

Nikishin 1972   G.N. Nikishin, V. G. Glukhovstev, M. A. Peikova and N. G. Maksimova, *Russ. Chem. Bull.*, 1972, 533-535

Ohshima 2012   T. Ohshima, Y. Hayashi, K. Agura, Y. Fujii, A. Yoshiyama and K. Mashima, *Chem. Commun.*, 2012, **48**, 5434–5436

Pace 2012   V. Pace, P. Hoyos, L. Castoldi, P. D. de María and A. R. Alcántra, *ChemSusChem*, 2012, **5**, 1369-1379

Pacheco 2016   A. A. C. Pacheco, J. Sherwood, A. Zhenova, C. R. McElroy, Andrew J. Hunt, H. L. Parker, T. J. Farmer, A. Constantinou, M. De bruyn, A. C. Whitwood, W. Raverty and J. H. Clark, *ChemSusChem*, 2016, **9**, 3503-3512

Paggiola 2016   G. Paggiola, S. V. Stempvoort, J. Bustamente, J. M. V. Barbero, A. J. Hunt and J. H. Clark, *Biofuels, Bioprod. Bioref.*, 2016, **10**, 686-698

Paris 2013   C. Paris, M. Moliner and A. Corma, *Green Chem.*, 2013, **15**, 2101-2109

Parker 1961   A. J. Parker, *J. Chem. Soc.*, 1961, **83**, 1328-1337

Parker 2014   H. L. Parker, J. Sherwood, A. J. Hunt and J. H. Clark, *ACS Sustainable Chem. Eng.*, 2014, **2**, 1739-1742

Parr 2018   Anton-Parr viscosity tables, https://wiki.anton-paar.com/en/toluene/, (Last accessed September 2018)

Pattab. 2011   V. R. Pattabiraman and J. W. Bode, *Nature*, 2011, **480**, 471-479

Pena-P. 2015   F. Pena-Pereira, A. Kloskowski and J. Namieśnik, *Green Chem.*, 2015, **17**, 3687-3705

Pirika 2018   Yamamoto-Molecular Break (Y-MB) in HSPiP, http://www.pirika.com/NewHSP/Y-MB/Y-MB.html, (Last accessed September 2018)

Piskun 2016   A. S. Piskun, J. E. de Haan, E. Wilbers, H. H. van de Bovenkamp, Z. Tang, and H. J. Heeres, *ACS Sustainable Chem. Eng.*, 2016, **4**, 2939-2950

Pham 2010   T. P. T. Pham, C.-W. Cho and Y.-S. Yun, *Water Res.*, 2010, **44**, 352-372

Pictet 1918   A. Pictet and J. Sarasin, *Helv. Chim. Acta*, 1918, **1**, 87-96

Prat 2013   D. Prat, O. Pardigon, H.-W. Flemming, S. Letestu, V. Ducandas, P. Isnard, E. Guntrun, T. Senac, S. Ruisseau, P. Cruciani and P. Hosek, *Org. Process Res. Dev.*, 2013, **17**, 1517-1525

Prat 2014   D. Prat, J. Hayler and A. Wells, *Green Chem.*, 2014, **16**, 4546-4551

Prat 2016   D. Prat, A. Wells, J. Hayler, H. Sneddon, C. R. McElroy, S Abou-Shehada and P. J. Dunn, *Green Chem.*, 2016, **18**, 288-296

Rajput 2018   P. Rajput and A. Sharma, *J Pharmacol Med Chem*, 2018, **2**, 22-31

Rama 2014   P. Rama-koteswararao, S. L. Tulasi and Y. Pavani, *JCHPS*, 2014, 132-135

RAPEX 2018   Rapid Alert System for dangerous non-food products, https://ec.europa.eu/consumers/consumers_safety/ safety_products/rapex/alerts/repository/content/ pages/rapex/index_en.htm, (Last accessed September 2018)

Ravid 1978   U. Ravid, R. M. Silverstein and L. R. Smith, *Tetrahedron*, 1978, **34**, 1449-1452

Reichardt 2011   C. Reichardt and T. Welton, *Solvents and Solvent Effects in Organic Chemistry*, Wiley-VCH, Weinheim, 4th edn, 2011

Roughley 2011   S. D. Roughley and A. M. Jordan, *J. Med. Chem.*, 2011, **54**, 3451-3479

Roy 2012   S. Roy, S. Roy and G. W. Gribble, *Tetrahedron*, 2012, **68**, 9867-9923

Royal 2012   *People and the Planet, The Royal Society Science Policy Centre report 01/12*, The Royal Society, London, 2012

Salavagione 2017 H. J. Salavagione, J. Sherwood, M. De bruyn, V. L. Budarin, G. J. Ellis, J. H. Clark and P. S. Shuttleworth, *Green Chem.*, 2017, **19**, 2550-2560

Shafizadeh 1979   F. Shafizadeh, R. H. Furneaux and T. T. Stevenson, *Carbohydrate Research*, 1979, **71**, 169-191

Sheldon 2000   R. A. Sheldon, *Pure Appl. Chem.*, 2000, **72**, 1233-1246

Sherwood 2014   J. Sherwood, M. De bruyn, A. Constantinou, L. Moity, C. R. McElroy, T. J. Farmer, T. Duncan, W. Raverty, A. J. Hunt and J. H. Clark, *Chem. Commun*, 2014, **50**, 9650-9652

Shuai 2015   L. Shuai, Y. M. Questell-Santiagoa and J. S. Luterbacher, *Green Chem.*, 2015, **18**, 937-943

Sigma 2018   The Sigma-Aldrich chemical synthesis catalogue, https://www.sigmaaldrich.com/chemistry/chemical-synthesis/ chemical-synthesis-catalog.html, (last accessed September 2018)

Snape 2010   T. J. Snape, A. M. Astles and J. Davies, *The Pharmaceutical Journal*, 2010, **285**, 416

Stefanis 2008   Emmanuel Stefanis and Costas Panayiotou, *Int. J. Thermophys.*, 2008, **29**, 568-585

Stockholm 2018   The Stockholm Convention, http://chm.pops.int/TheConvention/ Overview/TextoftheConvention/tabid/2232/Default.aspx, (Last accessed September 2018)

Su 2017   W. Su, L. Zhao and S. Deng, *Renew. Sust. Energ. Rev.*, 2017, **79**, 984-1001

Subram. 1997   B. Subramaniam, R. A. Rajewski and K. Snavely, *J. Pharm. Sci.*, 1997, **86**, 885-890

Sugisaki 2003   C. H. Sugisaki, Y. Ruland and M. Baltas, *Eur. J. Org. Chem.*, 2003, 672-688

Takano 1981a   S. Takano, E. Goto, M. Hirama and K. Ogasawara, *Heterocycles*, 1981, **16**, 381-385

Takano 1981b   S. Takano, E. Goto, M. Hirama and K. Ogasawara, *Heterocycles*, 1981, **16**, 951-954

Tamaru 1951   K. Tamaru, *Bull. Chem. Soc. Jap.*, 1951, **24**, 164-168

Tanaka 2003   K. Tanaka, *Solvent-free Organic Synthesis*, Wiley-VCH, Weinheim, 2003

Tang 2005   P. Tang, H. Krause and A. Fürstner, *Organic Syntheses*, 2005, **81**, 262-272

Teixeira 2016   A. R. S. Teixeira, A. L. Flourat, A. A. M. Peru, F. Brunissen and F. Allais, *Front. Chem.*, 2016, **4**, 1-11

Trost 1991   B. M. Trost, *Science*, 1991, **254**, 1471-1477

Tsuchiya 1970   Y. Tsuchiya and K. Sumi, *J. Appl. Polym. Sci.*, 1970, **14**, 2003-2013

Tundo 2002   P. Tundo and M. Selva, *Acc. Chem. Res.*, 2002, **35**, 706–716

UNEP 2018   United Nations Environment Programme, Still fresh at 30: Ozone hole healing, Montreal Protocol takes on climate change, https://www.unenvironment.org/news-and-stories/story/still-fresh-30-ozone-hole-healing-montreal-protocol-takes-climate-change, (Last accessed Sep 2018)

Valeur 2009   E. Valeur and M. Bradley, *Chem. Soc. Rev.*, 2009, **38**, 606-631

Vansant 1995   E. F. Vansant, P. Van Der Voort and K. C. Vrancken, *Characterization and Chemical Modification of the Silica Surface*, Elsevier Science, Amsterdam, 1995, ch. 3, 62–65.

Vidra 2018   A. Vidra and Á. Németh, *Periodica Polytech., Chem. Eng.*, 2018, **62**, 245-256

Vora 2003   B. V. Vora, J. A. Kocal, P. T. Barger, R. J. Schmidt and J. A. Johnson, *Kirk-Othmer Encyclopedia of Chemical Technology*, John Wiley & Sons, Inc., New Jersey, 2003

Wilson 2016	K. L. Wilson, A. R. Kennedy, J. Murray, B. Greatrex, C. Jamieson and A. J. B. Watson, *Beilstein J. Org. Chem.*, 2016, **12**, 2005-2011
Wilson 2017	K. L. Wilson, J. Murray, H. F. Sneddon, K. M. P. Wheelhouse and A. J. B. Watson, *Chem*, 2017, **3**, 365-368
Wilson 2018b	K. L. Wilson, J. Murray, C. Jamieson and A. J. B. Watson, *Synlett*, 2018, **29**, 650-654
Wilson 2018b	K. L. Wilson, J. Murray, C. Jamieson and A. J. B. Watson, *Org. Biomol. Chem.*, 2018, **16**, 2851-2854
Wodley 1971	F. A. Wodley, *J. Appl. Polym. Sci.*, 1971, **15**, 835-851
Xie 2017	C. Xie, X. Han, J. Gong, D. Li and C. Ma, *Org. Biomol. Chem.*, 2017, **15**, 5811–5819
Yalkowsky 2010	S. H. Yalkowsky and M. J. Wu, *Pharm. Sci.*, 2010, **99**, 1100-1106
Yamagami 1984	C. Yamagami, N. Takao, M. Tanaka, K. Horisaka, S. Asada and T. Fujita, *Chem. Pharm. Bull.*, 1984, **32**, 5003–5009
Yamamoto 2018	H. Yamamoto in *Hansen Solubility Parameters in Practice (HSPiP) e-Book*, ch. 30, https://www.pirika.com/ENG/HSP/E-Book/Chap30.html (Last accessed September 2018)
Yan 2017	Y. Yan, Z. Zhang, Y. Wan, G. Zhang, N. Ma and Q. Liu, *J. Org. Chem.*, 2017, **82**, 7957–7963
Zhang 2016	J. Zhang, G. B. White, M. D. Ryan, A. J. Hunt and M. J. Katz, *ACS Sustainable Chem. Eng.*, 2016, **4**, 7186-7192
Zhang 2018	X. Zhang, H. Yang, G. Yang, S. Li, X. Wang and J. Ma, *ACS Sustainable Chem. Eng.*, 2018, **6**, 5868-5876