Bio-Based Solvents for Organic Synthesis

James Richard Sherwood

Submitted for the degree of Doctor of Philosophy

University of York

Department of Chemistry

July 2013

Abstract

Scrutiny over solvent selection in the chemical industry has risen in recent decades, popularising research into neoteric solvent systems such as ionic liquids and supercritical fluids. More recently bio-based solvent products have been considered as replacements for conventional petroleum derived solvents. Because they bear a close resemblance to existing solvent products, bio-based solvents can be readily absorbed into the fine chemical industries. This work develops a methodology for identifying reactions of concern with respect to current solvent selection practice, and then implementing a high performance bio-based solvent substitute.

In this thesis, kinetic studies of heteroatom alkylation, amidation, and esterification are documented, and the solvent effect dictating the rate of each reaction ascertained. With the ideal properties for the solvent known, bio-based solvent candidates were screened for suitability in each case study. This process, which employs computational tools, was also applied to model the productivity of the Biginelli reaction as a representative multi-component heterocycle synthesis. A strong case is made for limonene and *p*-cymene as bio-based solvents for all but heteroatom alkylation from the case studies listed above. Alkylations with nitrogen nucleophiles are instead suited to high polarity solvents, and to this end some bio-based amides were investigated.

Contents

List of figures 9

List of schemes 17

List of tables 21

Acknowledgements 27

Declaration 29

1 Introduction: A critical analysis of green and renewable solvents 33

1.1 Modern solvent use 34

1.2 Properties of solvents 37

1.3 Solvent selection 51

1.4 Analysis of bio-based solvents 63

2 Nucleophilic substitution 77

2.1 Solvents and nucleophilic substitution 77

2.2 Nucleophilic substitution results and discussion 81

2.3 Heteroatom alkylation summary 111

3 Amidation 117

3.1 Solvents and amidation 117

3.2 Amidation results and discussion 127

3.3 Amidation summary 153

4 Uncatalysed esterification 157

4.1 Solvents and esterification 157

4.2 Uncatalysed esterification results and discussion 161

4.3 Esterification summary 175

5 Catalysed carbonyl addition 179

5.1 Bio-based acid catalysts for organic chemistry 179

5.2 Combined solvent and catalytic effects in carbonyl additions 191

5.3 Catalysed carbonyl addition summary 204

6 Heterocycle synthesis: The Biginelli reaction 209

6.1 Introduction to Pietro Biginelli and his reaction 209

6.2 Standard Biginelli reaction solvent effects 214

6.3 Modified Biginelli reaction solvent effects 237

6.4 Biginelli reaction summary 240

7 Conclusion 245

7.1 Case study recapitulation 245

7.2 The future of bio-based solvents 248

8 Appendices 255

8.1 Experimental procedures 255

8.2 Supplementary data 261

Abbreviations 285

References 295

List of figures

Figure 1.1 Materials (by mass) required for the manufacture of a typical pharmaceutical product.

Figure 1.2 A comparison between reaction class frequency in process chemistry and manufacturing plant chemistry within the pharmaceutical industry.

Figure 1.3 An energy profile of 4-nitroaniline dissolving in acetic acid.

Figure 1.4 A comparison between relative permittivity and the Hildebrand solubility parameter.

Figure 1.5 Three dimensional Hansen plot based on the solubility of urea.

Figure 1.6 The solvent effect influencing the light absorbance of Reichardt's betaine dye.

Figure 1.7 The solvent polarity scale derived from Reichardt's dye.

Figure 1.8 The solvatochromism of *N*,*N*-diethyl-4-nitroaniline.

Figure 1.9 The UV-vis. spectrum of *N*,*N*-diethyl-4-nitroaniline in cyclohexane, 2-MeTHF, and DMSO.

Figure 1.10 A comparison between aniline dye absorbance maxima in different solvents.

Figure 1.11 Aprotic solvent polarity map.

Figure 1.12 Protic solvent polarity map.

Figure 1.13 A visual representation of the algorithmic solvent selection process.

Figure 1.14 Rule D of the solvent selection algorithm.

Figure 1.15 Rule C (R_{2a}) of the revised solvent selection algorithm when the solvent is opted to be recycled.

Figure 1.16 An arbitrary polarity map with rule G assignments from the revised solvent selection algorithm.

Figure 1.17 A hypothetical comparison between methods for identifying the permissible toxicity limits of solvents when the user defined limit is $\log(LD_{50}) = 3.5$.

Figure 1.18 A linear free energy diagram showing the acidity of benzoic acids in different solvents.

Figure 1.19 The rate of a Fischer esterification correlated to solvent polarity.

Figure 1.20 Aprotic bio-based solvent polarity map.

Figure 1.21 Protic bio-based solvent polarity map.

Figure 2.1 The natural logarithms of rate constants (relative to methanol) for the Menschutkin reaction.

Figure 2.2 The progression of the model Menschutkin reaction in DMSO as observed by ^{1}H-NMR spectroscopy.

Figure 2.3 The conversion to 1-decyl-2,3-dimethylimidazolium bromide with the Menschutkin reaction in DMSO and ethanol as determined by ^{1}H-NMR spectroscopy.

Figure 2.4 Integrated rate equations of Menschutkin reactions in DMSO and ethanol.

Figure 2.5 The polarity of the model Menschutkin reaction solvent set.

Figure 2.6 The SUS-HAS-ECO classifications of the model Menschutkin reaction solvent set.

Figure 2.7 The LSER describing the rate constant of the model Menschutkin reaction as a function of solvent dipolarity only.

Figure 2.8 A comparison between experimental and predicted S_N2 ln(**k**) values based on a LSER incorporating both $\boldsymbol{\pi^*}$ and $\boldsymbol{\alpha}$

Figure 2.9 Predicted ln(**k**) values from the LSER featuring the $\boldsymbol{\varepsilon}$ modification to $\boldsymbol{\alpha}$ compared to experimental Menschutkin reaction data.

Figure 2.10 A systematic error check of experimental rate constants for the model Menschutkin reaction.

Figure 2.11 A demonstration of the absence of a correlation between Reichardt's $\mathbf{E_T^N}$ parameter and ln(**k**) of the model Menschutkin reaction.

Figure 2.12 The relationship between relative permittivity and experimental values of ln(**k**) of the model Menschutkin reaction.

Figure 2.13 An arbitrary free energy diagram indicating Menschutkin reaction solvent effects in DMSO relative to chloroform and ethanol.

Figure 2.14 Menschutkin reaction solvent selection algorithm screenshot, step 1: Reaction components.

Figure 2.15 Menschutkin reaction solvent selection algorithm screenshot, step 2: Solvent class inclusion.

Figure 2.16 Menschutkin reaction solvent selection algorithm screenshot, step 3: Parameter input.

Figure 2.17 Menschutkin reaction solvent selection algorithm screenshot: Polarity matching diagram for estimating the solubility of 1,2-dimethylimidizole.

Figure 2.18 Menschutkin reaction solvent selection algorithm screenshot, step 4: Scoring system.

Figure 2.19 Menschutkin reaction solvent selection algorithm screenshot, step 5: Solvent effects.

Figure 2.20 Menschutkin reaction solvent selection algorithm screenshot, step 6: Solvent selection guide.

Figure 2.21 Menschutkin reaction solvent selection algorithm screenshot, step 7: Results preview.

Figure 2.22 Bio-based amide solvent yields.

Figure 2.23 The polarity of highly dipolar aprotic solvents including bio-based amides.

Figure 2.24 The performance of bio-based amide solvents and sulpholane in the Menschutkin reaction.

Figure 3.1 A reproduction and simplification of the Pfizer reagent selection Venn diagram for amide coupling protocols.

Figure 3.2 An example of a ^{1}H-NMR spectrum showing the partially complete model amidation in toluene.

Figure 3.3 Amidation reaction order determination in toluene.

Figure 3.4 A polarity map of solvents included in the initial screening of the model amidation reaction.

Figure 3.5 The environmental, health and safety of amidation solvents.

Figure 3.6 The LSER correlating the rate of the model amidation reaction with the hydrogen bond accepting ability of the solvent.

Figure 3.7 A multi-solvent Eyring relationship for amidation including a predicted crossover point of reaction rates at an iso-kinetic temperature.

Figure 3.8 The correlation between the enthalpy of activation and solvent hydrogen bond accepting ability (**β**).

Figure 3.9 The correlation between the entropy of activation and solvent hydrogen bond accepting ability (**β**).

Figure 3.10 The relationship between the activation parameters dictating the kinetics of the model amidation.

Figure 3.11 The generalised variable enthalpy of amidation illustrated in toluene and DMSO.

Figure 3.12 The toluene-DMSO binary solvent effect on the rate of amidation.

Figure 3.13 Amidation reaction solvent selection algorithm screenshot, step 3: Parameter input.

Figure 3.14 Amidation reaction solvent selection algorithm screenshot, step 6: Solvent selection guide.

Figure 3.15 A LSER indicating the performance of limonene and *p*-cymene in amidation reactions.

Figure 3.16 Estimated versus predicted ln(**k**) values of amidation.

Figure 3.17 Systematic error check in the amidation rate constants.

Figure 3.18 Isolated yields of *N*-benzyl-4-phenylbutanamide from different solvents.

Figure 3.19 Associated metrics of amidation reactions between 4-phenylbutanoic acid and benzylamine unless otherwise stated.

Figure 3.20 The comparison of PMI in different amidations.

Figure 4.1 The manipulation of a fluorous multiphasic system to enhance esterification yields.

Figure 4.2 A ^{1}H-NMR spectrum of the uncatalysed model esterification occurring in chloroform.

Figure 4.3 Polarity map of solvents used in the initial uncatalysed model esterification solvent set.

Figure 4.4 Uncatalysed model esterification solvent SUS-HAS-ECO classifications.

Figure 4.5 The LSER describing the uncatalysed model esterification but excluding acetonitrile.

Figure 4.6 Conversions to butyl butanoate in acetonitrile and propanenitrile at 323 K.

Figure 4.7 The estimated ln(**k**) values of the uncatalysed model esterification, including nitrile solvents, compared to experimental data.

Figure 4.8 Uncatalysed esterification solvent selection algorithm screenshot, step 3: Parameter input.

Figure 4.9 Uncatalysed esterification solvent selection algorithm screenshot, step 6: Solvent selection guide.

Figure 4.10 The estimated ln(**k**) values from the LSER describing the uncatalysed model esterification, including bio-based solvents, plotted against experimental data.

Figure 4.11 Systematic error check regarding the certainty of rate constant estimation in the uncatalysed model esterification.

Figure 4.12 A solvent polarity map highlighting limonene and *p*-cymene.

Figure 5.1 The comparison between Pd/C and $ZnCl_2$ catalysed oxidations of α-terpinene by ^{1}H-NMR spectroscopy.

Figure 5.2 An optimisation study concerning the loading of Pd/C and K-10 for the conversion of limonene to *p*-cymene at 373 K.

Figure 5.3 Product selectivity to *p*-cymene from limonene at different temperatures.

Figure 5.4 A ^{1}H-NMR spectroscopic analysis of the reaction mixture after limonene oxidation, largely to give *p*-cymene at 413 K, but less effective at 433 K.

Figure 5.5 The efficiency of *p*-cymene production utilising different sources of limonene.

Figure 5.6 The absorbance profile of 4-nitroaniline in the presence of varying concentrations of *p*-CSA.

Figure 5.7 The Hammett acidity of sulphonic acids.

Figure 5.8 Acid catalysed Fischer esterification to give benzyl acetate solvent selection algorithm screenshot, step 3: Parameter input.

Figure 5.9 A ^{1}H-NMR spectrum of the Fischer esterification to give benzyl acetate in *p*-cymene.

Figure 5.10 The solvent SUS-HAS-ECO classifications for the benzyl acetate case study.

Figure 5.11 The LSER describing Fischer esterification catalysed by *p*-TSA and *p*-CSA to give benzyl acetate, also including the combination of benzenesulphonic acid in *p*-cymene.

Figure 5.12 The difference between the $\ln(\mathbf{k})$ of *p*-TSA and *p*-CSA catalysed Fischer esterifications.

Figure 5.13 Combinations of different catalysts and solvents for the synthesis of ethyl levulinate.

Figure 5.14 Metrics associated with the conversion of limonene to *p*-cymene.

Figure 5.15 Process mass intensity of different *p*-cymene syntheses.

Figure 5.16 The manufacturing processes for arene sulphonic acids.

Figure 6.1 Combined catalytic and solvent effects.

Figure 6.2 Biginelli reaction solvent selection algorithm screenshot, step 3: Parameter input.

Figure 6.3 The polarity range of the Biginelli reaction solvent set.

Figure 6.4 The SUS-HAS-ECO classifications of solvents selected for the standard Biginelli reaction case study.

Figure 6.5 The relationship between reaction productivity and $\boldsymbol{\pi^*}$ to give methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate by HCl catalysis.

Figure 6.6 Solvent dependence of the tautomerisation equilibrium of methyl acetoacetate at the reaction concentration of 1.875 M.

Figure 6.7 HCl catalysed reactions showing the influence of $\ln(\mathbf{K}_T)$ on the yield of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate.

Figure 6.8 Reaction productivity to give methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate by EPZ-10 catalysis.

Figure 6.9 Correlating the influence of Brønsted acid and Lewis acid catalysis to the isolated yields of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate.

Figure 6.10 A comparison between the solvato-catalytic effects in the synthesis of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate under different conditions.

Figure 6.11 A comparison between the isolated yields of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate in acetic acid, ethanol, and *p*-cymene catalysed by either HCl or EPZ-10.

Figure 6.12 A comparison between the isolated yields of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate arising from different methods of catalysis in acetic acid over the course of 3 hours unless otherwise indicated.

Figure 6.13 A comparison between the isolated yields of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate in acetic acid, ethanol, and *p*-cymene using a variety of catalysts.

Figure 6.14 The comparison between the isolated yields of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate in aromatic solvents catalysed by their respective sulphonic acids.

Figure 6.15 The correlation between solvent hydrogen bond accepting ability and the productivity of the reaction to give 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione.

Figure 7.1 The polarity map of Figure 1.20 annotated with LSER trends.

Figure 8.1 The correlation between ^{1}H-NMR signal integrals and the proportion of amide product in solution based on the consumption of benzaldehyde.

Figure 8.2 A comparison between experimental and calculated amidation enthalpies of activation.

Figure 8.3 The comparison between experimental and calculated entropies of activation in the model amidation reaction.

Figure 8.4 Overlap of a pure *p*-cymene standard and an experimental ^{13}C-NMR spectrum of limonene converted to *p*-cymene at 413 K.

Figure 8.5 Identification of key proton signals in the ^{1}H-NMR spectrum of methyl acetoacetate in toluene.

List of schemes

Scheme 1.1 The relative permittivity of acetic acid and acetyl chloride for comparison.

Scheme 1.2 The solvent interactions of 4-nitroaniline and *N*,*N*-diethyl-4-nitroaniline as exemplified with DMSO.

Scheme 1.3 Some bio-based solvents accessible by the fermentation of glucose.

Scheme 2.1 Menschutkin's original reaction.

Scheme 2.2 The general mechanism of the Menschutkin reaction in which the activated complex is stabilised through an interaction with the solvent DMSO.

Scheme 2.3 The model Menschutkin reaction between 1,2-dimethylimidazole and 1-bromodecane.

Scheme 2.4 The typical X-H hydrogen bond donor ethanol compared to a C-H hydrogen bond donor chloroform as they interact with 1,2-dimethylimidazole.

Scheme 2.5 The manufacturing route to bio-based sulpholane.

Scheme 2.6 The potentially bio-derived ureas tetramethylurea and dimethyl-1,3-propyleneurea (DMPU).

Scheme 2.7 The synthesis of cyclic carbonate solvents from alcohols.

Scheme 2.8 The possibility of using a bio-based alcohol precursor for the synthesis of amides.

Scheme 2.9 The synthesis of a class of *N*-acylpyrrolidines from glutamic acid.

Scheme 2.10 The proposed bio-based amide solvents *N*-acetylpyrrolidine, *N*-propionylpyrrolidine, and *N*-laurylpyrrolidine.

Scheme 2.11 Generalised reaction scheme describing the synthesis of *N*-acylpyrrolidines.

Scheme 2.12 A fatty acid derived amide.

Scheme 2.13 The comparison between the mechanism of different types of S_N2 reaction.

Scheme 2.14 The non-Menschutkin S_N2 mechanism as assisted by DMSO.

Scheme 3.1 The commonly perceived competing amidation and salt forming equilibria.

Scheme 3.2 A proposed reactant stabilised activated complex of amidation.

Scheme 3.3 Nucleophilic attack on a carboxylic acid by either reactant in an amidation reaction.

Scheme 3.4 The hydrogen bond assisted amidation mechanism.

Scheme 3.5 Catalysis capable of enhancing amidation reactions, demonstrated in the synthesis of Efaproxiral.

Scheme 3.6 Amidation to give *N*-benzyl-4-phenylbutanamide.

Scheme 3.7 The synthesis of carboxylic acid anhydrides.

Scheme 3.8 The $A_{AC}2$ mechanism of amidation proceeding through a tetrahedral intermediate.

Scheme 3.9 The synthesis of *2-N*,*N*-diisopropyl-5-fluorobenzylaminoboronic acid.

Scheme 4.1 A proposal for an acid catalysed esterification solvent effect.

Scheme 4.2 The model uncatalysed esterification reaction.

Scheme 4.3 The $A_{AC}2$ mechanism specifically concerning the disturbance to the bulk solvent medium as illustrated with acetonitrile.

Scheme 4.4 The solvent-solute interactions dictating the rate of uncatalysed esterification as demonstrated with DMF.

Scheme 4.5 Manufacturing routes to solvents fit for esterification chemistry.

Scheme 5.1 The synthesis of *p*-CSA from limonene *via* *p*-cymene.

Scheme 5.2 A comparison between the synthetic routes linking the products of toluene and *p*-cymene.

Scheme 5.3 The mechanism of *p*-cymene sulphonation by sulphuric acid.

Scheme 5.4 A mechanistic proposal describing the process of converting limonene into *p*-cymene.

Scheme 5.5 The disproportionation of α-terpinene to *p*-cymene and *p*-menthane.

Scheme 5.6 The H_0 equilibrium between *p*-CSA and 4-nitroaniline.

Scheme 5.7 The synthesis of benzyl acetate by Fischer esterification.

Scheme 5.8 The combined roles of solvent and catalyst in the $A_{AC}2$ mechanism of a representative carbonyl addition.

Scheme 5.9 Fischer esterification to give ethyl levulinate.

Scheme 5.10 A chalcone synthesis by aldol condensation.

Scheme 5.11 Acetal protection using ethylene glycol.

Scheme 5.12 Three different sets of conditions that effect the transformation of limonene to *p*-cymene with varying success.

Scheme 5.13 A generalised enhanced leaving group strategy using the reagent *p*-cymenesulphonyl chloride.

Scheme 6.1 The original Biginelli reaction.

Scheme 6.2 A standard Brønsted acid catalysed Biginelli reaction and a proposed Lewis acid catalysed Biginelli reaction illustrated with a zinc cation.

Scheme 6.3 The model standard Biginelli reaction.

Scheme 6.4 The effect of acidic solvents on the tautomerisation equilibrium of methyl acetoacetate.

Scheme 6.5 The different reactivity of dicarbonyl compound tautomers.

Scheme 6.6 The nucleophilic moieties of methyl acetoacetate tautomers.

Scheme 6.7 Proposed solvent control over the generation of intermediates in the Biginelli reaction.

Scheme 6.8 The complex created by the addition of ethylene glycol to zinc chloride.

Scheme 6.9 The roles of metal cations as Lewis acids in the standard Biginelli reaction.

Scheme 6.10 A modified Biginelli reaction with 5,5-dimethyl-1,3-cyclohexanedione to give 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione.

Scheme 6.11 Differences in the solvent stabilisation and resulting tautomerisation equilibria of the 1,3-dicarbonyl compounds methyl acetoacetate and 5,5-dimethyl-1,3-cyclohexanedione exemplified with DMF.

Scheme 6.12 An example of the Hantzsch dihydropyridine synthesis.

List of tables

Table 1.1 Multi-company process development chemistry reaction frequency for transformations contributing to at least 1% of total chemistry.

Table 1.2 Data for manufacturing scale transformations at one Pfizer site for transformations contributing to at least 1% of total chemistry (1997-2002).

Table 1.3 A representation of the Pfizer medicinal chemistry solvent selection guide.

Table 1.4 An excerpt from the GSK solvent selection guide supporting table.

Table 1.5 The basis for assigning a renewability index (SUS) to bio-based solvents.

Table 1.6 A bio-based solvent selection guide.

Table 1.7 Solvent selection practices in reactions common to the pharmaceutical industry.

Table 2.1 A collection of LSER analyses for Menschutkin type reactions and the solvolysis of *t*-butyl chloride as reported in the original publications.

Table 2.2 Solvent polarity measurements and kinetics of the model Menschutkin reaction.

Table 2.3 An explanation of the polarisability term coefficient *d*.

Table 2.4 Solvent hits generated by the solvent selection algorithm for the Menschutkin reaction.

Table 2.5 The properties and reaction rates of the Menschutkin reaction in DMSO and optimal solvent candidates.

Table 2.6 A collection of LSERs describing S_N2 and S_NAr reactions.

Table 3.1 A selection of activating techniques for promoting amidation.

Table 3.2 Heterogeneously catalysed methodologies for the synthesis of carboxamides.

Table 3.3 Homogeneously catalysed methodologies for the synthesis of *N*-benzyl-4-phenylbutanamide.

Table 3.4 Amidation rate constants in different solvents and the polarity of those solvents.

Table 3.5 Activation parameters for the model amidation reaction including the Gibbs free energy at 373 K.

Table 3.6 Solvent hits generated by the solvent selection algorithm for amidation reactions.

Table 3.7 A comparison between high performance bio-based and petro-chemical amidation solvents.

Table 3.8 Properties of toluene and *p*-cymene.

Table 4.1 The acid catalysed mechanisms of ester hydrolysis.

Table 4.2 Uncatalysed model esterification rate constants and solvent properties.

Table 4.3 Uncatalysed esterification solvent hits from the solvent selection algorithm.

Table 4.4 A comparison between high performance bio-based and petro-chemical esterification solvents.

Table 5.1 A selection of properties defining *p*-TSA and *p*-CSA.

Table 5.2 Solvent properties concerning a Fischer esterification to give benzyl acetate.

Table 5.3 Comparison of yields in esterification, aldol condensation, and acetalisation chemistries obtained by *p*-TSA and *p*-CSA catalysis.

Table 6.1 A selection of catalysts that have been used to promote the standard Biginelli reaction in different solvents.

Table 6.2 Solvent hits generated by the first iteration of the solvent selection algorithm for the standard Biginelli reaction.

Table 6.3 First screening of the experimental solvent set through the solvent selection algorithm describing the standard Biginelli reaction.

Table 6.4 Solvent polarity and performance in the synthesis of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate by HCl catalysis.

Table 6.5 Solvent hits generated by the second iteration of the solvent selection algorithm for the standard Biginelli reaction.

Table 6.6 The proportions of the enol tautomer of 1,3-dicarbonyl compounds in different solvents.

Table 8.1 Bulk solvent polarity parameters.

Table 8.2 Values for the Kamlet-Taft solvent polarity scale.

Table 8.3 The full SUS-HAS-ECO classifications of solvent greenness.

Table 8.4 The statistical significance of LSER coefficients for the kinetics of the Menschutkin reaction in the form '$\ln(\mathbf{k}) = XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^{*}$'.

Table 8.5 The final LSER coefficients for the kinetics of the Menschutkin reaction in the form '$\ln(\mathbf{k}) = XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^{*}$' using only statistically significant parameters.

Table 8.6 The final LSER coefficients for the kinetics of the Menschutkin reaction in the form '$\ln(\mathbf{k}) = XYZ_0 + a(\boldsymbol{\alpha} + e\boldsymbol{\varepsilon}) + s\boldsymbol{\pi}^{*}$' using only statistically significant parameters.

Table 8.7 The LSER coefficients for the kinetics of the Menschutkin reaction in the form '$\ln(\mathbf{k}) = XYZ_0 + f\mathbf{E}_{\mathbf{T}}^{\mathbf{N}}$' or '$\ln(\mathbf{k}) = XYZ_0 + g\boldsymbol{\varepsilon}_{\mathbf{r}}$'.

Table 8.8 The statistical significance of LSER coefficients for the kinetics of amidation in the model reaction in the form '$\mathrm{XYZ} = XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^{*}$'.

Table 8.9 The final LSER coefficients for the kinetics of the model amidation in the form '$\mathrm{XYZ} = XYZ_0 + b\boldsymbol{\beta}$' using only statistically significant parameters.

Table 8.10 Mass utilisation in the extended synthesis of *N*-benzyl-4-phenylbutanamide and the catalyst *2-N*,*N*-diisopropyl-5-fluorobenzylaminoboronic acid.

Table 8.11 The statistical significance of LSER coefficients for the kinetics of an uncatalysed esterification in the form '$\ln(\mathbf{k}) = XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^{*}$'.

Table 8.12 Extended LSER relationships for the kinetics of the uncatalysed model esterification in the form '$\ln(\mathbf{k}) = XYZ_0 + b\boldsymbol{\beta}$' or '$\ln(\mathbf{k}) = XYZ_0 + b\boldsymbol{\beta} + h\boldsymbol{\delta}_{\mathbf{T}}^{2}$'.

Table 8.13 The statistical significance of LSER coefficients for the kinetics of Fischer esterification to give benzyl acetate expressed in the equation '$\ln(\mathbf{k}) = XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^{*}$'.

Table 8.14 The final LSER coefficients for the kinetics of Fischer esterification in the form '$\ln(\mathbf{k}) = XYZ_0 + b\boldsymbol{\beta}$' using only statistically significant parameters.

Table 8.15 Full yields and relevant solvent parameters for the synthesis of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate, including the tautomerisation equilibrium constants of methyl acetoacetate.

Table 8.16 Full yields and relevant solvent parameters for the synthesis of 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione, including the tautomerisation equilibrium constants of 5,5-dimethyl-1,3-cyclohexanedione.

Table 8.17 Solvent effects on the diketo-enol tautomerisation of 1,3-dicarbonyls and reaction productivity in the form of the LSER: $\mathrm{XYZ} = XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^*$.

Table 8.18 Correlations between reaction productivity and tautomerisation in the following form: $\ln(\mathbf{P}/\mathbf{R}) = XYZ_0 + t\cdot\ln(\mathbf{K}_\mathrm{T})$.

Acknowledgements

I would like to take this opportunity to thank the following people who made this work possible. Firstly to Prof. James Clark and Dr. Duncan Macquarrie, who have greatly influenced this work, and certainly for the better. I am also thankful to Prof. Rafiqul Gani, who allowed me to work with him at DTU in Copenhagen for a period of time during this work, and to those members of his group who have helped me with the computational solvent selection procedures. I am especially grateful for the help of the aptly named Dr. Merlin Alvarado-Morales. I must also extend my gratitude to Dr. Richard Henderson and my past associates at GSK who made the medicinal chemistry solvent selection guide possible, and to Prof. Tom Welton and Dr. Jason Hallett who first spurred my interest in solvent effects.

Of course I must also show my appreciation towards my colleagues at the University of York, and highlight where they have contributed towards this work. Paul Elliot conducted the thermal analysis studies used in this work to determine the thermal decomposition points of sulphonic acids. Karl Heaton, Andy Marriott, and Graham McAllister performed mass spectrometry and elemental analysis on my behalf. Some of the data concerning food waste was sourced from the literature by Lucie Pfaltzgraff. I would also like to thank Heather Fish for NMR spectroscopy training and assistance, and Anne-Kathrin Duhme-Klair for use of her laboratory while the Green Chemistry Centre of Excellence was on fire. Finally, it is with the most sincere gratitude that I acknowledge the input of Dr. Thomas Farmer, who in one of several insightful discussions provided me with the idea for using *p*-CSA as a catalyst.

Declaration

The research presented in this thesis was conducted at the Green Chemistry Centre of Excellence at the University of York, UK, between October 2009 and September 2012. To the best of the author's knowledge, all unreferenced work is original and of his own doing, except where other persons have been acknowledged. Part of the work disclosed herein has been published as the following articles. Except for the first entry, these articles were all written and submitted for publication by the author of this work.

1. 'Expanding GSK's solvent selection guide - Embedding sustainability into solvent selection starting at medicinal chemistry'

Richard K. Henderson, Concepción Jiménez-González, David J. C. Constable, Sarah R. Alston, Graham G. A. Inglis, Gail Fisher, James Sherwood, Steve P. Binks and Alan D. Curzons, *Green Chem.*, 2011, **13**, 854.

2. 'A quantitative comparison between conventional and bio-derived solvents from citrus waste in esterification and amidation kinetic studies'

James H. Clark, Duncan J. Macquarrie and James Sherwood, *Green Chem.*, 2012, **14**, 90.

3. '*p*-Cymenesulphonic acid: An organic acid synthesised from citrus waste'

James H. Clark, Emma M. Fitzpatrick, Duncan J. Macquarrie, Lucie A. Pfaltzgraff, and James Sherwood, *Catal. Today*, 2012, **190**, 144.

4. 'Green Solvents'

Simon W. Breeden, James H. Clark, Duncan J. Macquarrie and James R. Sherwood in *Green Techniques for Organic Synthesis and Medicinal Chemistry*, John Wiley and Sons, Chichester, 2012.

5. ‘The Combined Role of Catalysis and Solvent Effects on the Biginelli Reaction: Improving Efficiency and Sustainability’

James H. Clark, Duncan J. Macquarrie and James Sherwood, *Chem. Eur. J.*, 2013, **19**, 5174.

1. Introduction: A critical analysis of green and renewable solvents

Solvents are ubiquitous throughout synthetic chemistry and are also prevalent in professional and general consumer articles. The versatility of active ingredients when in solution is indispensible to our modern world. Still, the definition of solvent is fraught with complications. It is best to be lenient in order to avoid dismissing certain neoteric systems such as supercritical fluids [Oakes 2001]. Solvents can be defined by the roles which they serve, as in European legislation [EC 1999]. Alternatively, as in this work, a solvent can be regarded simply as a substance, or mixture of substances that has another substance dissolved in it, or it is the intention to dissolve another substance in it. The purpose of dissolving chemicals in an excess of fluid is usually to achieve homogeneity. A paint product for example must carry pigments, an adhesive, and biocides amongst other additives, which is achieved with a solvent. The solvent also modifies the viscosity of the product, imparting properties of its own onto the formulation. After application of the paint the solvent is no longer required, and allowed to evaporate. The same stages are present in synthetic chemistry. A solvent is introduced into a mixture, performs a given task during the reaction, and is then removed. Although the role of a solvent for synthesis is largely dissolution of the reaction components, it also includes thermal regulation and ease of material transfer [Adams 2004 page 3, Kerton 2009 page 1].

The renewed interest in solvents and solvent effects in organic chemistry is explored here at the beginning of this work, a phenomenon that is driven by the widespread desire to replace non-renewable chemicals with bio-based alternatives. This stance also provides an excellent opportunity to design safer solvents of lower toxicity and minimal environmental impact. Importantly, measurements of solvent performance, and the properties that define it, are also described. Protocols with which the performance of bio-based solvents can be established are presented throughout this chapter. The demand for green solvents and bio-based chemicals will only increase, and enhancing reaction performance with benign yet renewable solvents can only benefit people and the environment.

1.1 Modern solvent use

Solvents and green chemistry: The impact of large scale manufacturing practices on the environment has been recognised for many decades [Cue 2012 page 553, Lancaster 2002 page 24, Matlack 2001 page 1]. Numerous laws have been established in order to minimise pollution and maximise safety [Anastas 2002]. But only recently in the history of industrial chemistry is it more common to scrutinise the actual process in a pro-active manner rather than implement so-called ‘end of pipe’ solutions [Anastas 2010]. This is the central tenet of green chemistry, encompassing waste and energy efficiency as well as sustainability, toxicity, and health and safety [Clark 2005 page 3, Sheldon 2005]. The gradual act of discouraging the use of certain solvents and promoting others is now becoming common. Legislation, industry documents, and non-governmental organisation (NGO) reports have all published banned or restricted chemical lists which inevitably feature solvents [EC 2007, Kerton 2009 page 2, SubsPort 2013].

The use of solvents, especially in organic synthesis, is one of the largest areas of research conducted within the field of green chemistry. A lot of this research begins with the identification of an unusual solvent, often an ionic liquid, for which an application is developed as a means of demonstrating the prowess of that solvent. In this work, the opposite is true. Potential case studies will be scrutinised and those reactions that are of vital importance, whilst also commonly practiced and currently employing a less than desirable solvent will be considered for study. Legislative drivers and growing consumer pressure for safer and renewable chemical products means that novel bio-based solvents must be keenly considered in any case study considering solvent substitution. The pharmaceutical industry provides many of this type of case study [Constable 2007a, Dunn 2012, Jiménez-González 2012]. Highly dependent on solvents, the production of pharmaceuticals could benefit immensely from the identification of suitable bio-based solvents.

Solvent use in the pharmaceutical industry: The fine chemical industries rely on solvents, but none more than the pharmaceutical industry [Sheldon 2000]. The volume in which they use solvents dwarfs the amount of other chemical inputs. At manufacturing scale organic solvents are more than half the mass of material required to produce an active pharmaceutical ingredient (API) according to data from GlaxoSmithKline (GSK), rising to over 80% if including water (Figure 1.1) [Henderson 2011]. Accordingly much activity on solvent reduction and solvent replacement is occurring within the pharmaceutical industry, as it has been for a number of years now. Waste reduction and health and safety are important, as is the renewability of the solvents.

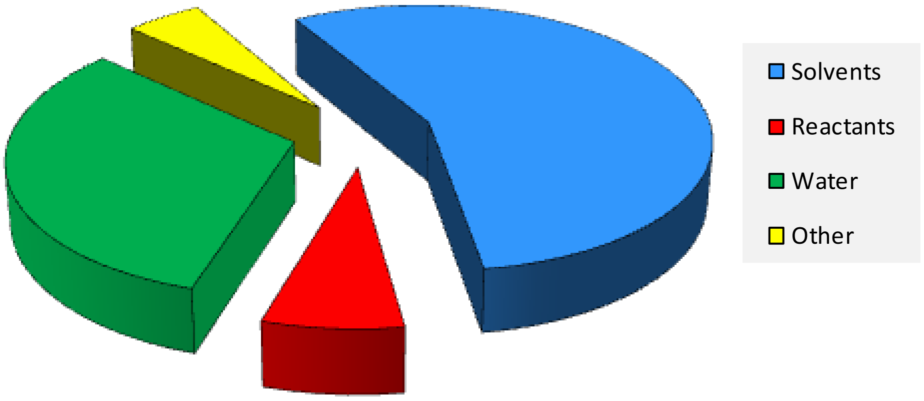


Figure 1.1 Materials (by mass) required for the manufacture of a typical pharmaceutical product.

The pharmaceutical industry utilises a wide range of transformations, adopting a distinct preference for particular solvents in each reaction. By identifying the most valuable, most practiced transformations, solvent replacement in favour of sustainable alternatives can be targeted to achieve the most benefit to the pharmaceutical industry and the environment. Recently pharmaceutical process development at AstraZeneca, GSK, and Pfizer collated the data from over 1000 reactions to provide an insight into the frequency at which they perform different classes of reaction (Table 1.1) [Carey 2006]. At this scale volumes of solvent become very significant. It should be noted that protection and deprotection reactions recorded in the original publication have been removed from the list provided in Table 1.1 as efforts should be made to avoid this practice rather than simply minimise its impact. Alkylation (nitrogen and oxygen nucleophiles) and amidation are clearly very important, followed by the synthesis of heterocycles and cross-coupling reactions. The most prevalent functional group interconversion is the synthesis of organohalides from alcohols. Other reactions are less practiced (2% of total chemistry or less each) and because of this should receive less attention when case studies are selected for the purpose of identifying alternative renewable solvents.

More reaction surveys have emerged in the last couple of years. GSK report that the top three reaction classes most practiced in their respiratory disease division of medicinal chemistry are alkylation, acylation, and palladium C-C cross coupling, each contributing to 17% of a total of 4800 reactions surveyed in 2005 [MacDonald 2010]. Heterocycle synthesis stands at 5% of the total reaction count. Solvent use is less per reaction in this department compared to process development because of the small scale of the work being conducted, but cumulatively it is still of great significance [Alfonsi 2008]. Another cross-company survey reports similar data from 2008, although this data set was gathered from literature sources and not in-house reports [Roughley 2011]. Regardless the picture is similar, with heteroatom alkylation (23%) and

Table 1.1 Multi-company process development chemistry reaction frequency for transformations contributing to at least 1% of total chemistry.

Rank	Transformation	Frequency
01	*N*-Substitution	10.8 %
02	*N*-Acylation to amide	7.9 %
03	*O*-Substitution	5.3 %
04	Heterocycle synthesis	3.4 %
05	Cross-coupling	2.4 %
06	Alcohol to halide	2.2 %
07	Nitrate to amine	2.0 %
08	Halogenation	1.6 %
=09	Claisen condensation	1.5 %
=09	*S*-Substitution	1.5 %
11	Amide to imidoyl chloride	1.4 %
=12	Imine/nitrile to amine	1.3 %
=12	Organometallic C-C bond formation	1.3 %
=12	Acid to acid chloride	1.3 %
=15	Alkene to alkane	1.1 %
=15	Friedel-Crafts alkylation/acylation	1.1 %
=15	*N*-Sulphonation to sulphonamide	1.1 %
=15	Nitration	1.1 %
=19	Dehydration	1.0 %
=19	Oxidation of sulphur	1.0 %

acylation (22%) dominating synthetic procedures. Unsurprisingly cross coupling and heterocycle synthesis feature prominently too, repeating the now familiar reaction hierarchy.

Data for manufacturing processes at Pfizer sees a fair correlation with the frequency of reactions performed on smaller scales (Table 1.2) [Dugger 2005]. Protection and deprotection reactions were not treated separately in this dataset and so they could not be removed as before. Amidation now tops nucleophilic substitution in terms of frequency, and with a quarter of drug molecules containing an amide bond it is no wonder that this is the case [Ghose 1999]. Esterification and the reverse hydrolysis reaction are prominent at manufacturing scale, probably mostly consisting of protection strategies given its absence in Table 1.1.

Grouping the transformations reviewed thus far by reaction class helps the comparison between process development and API manufacture. With the following classes the dominance

Table 1.2 Data for manufacturing scale transformations at one Pfizer site for transformations contributing to at least 1% of total chemistry (1997-2002).

Rank	Transformation	Frequency	Rank in Table 1.1
01	*N*-Acylation to amide	10.0 %	02
02	*O*-Acylation to ester/ester hydrolysis	4.5 %	n/a
03	Heterocycle synthesis	3.2 %	04
04	*N*-Substitution	3.1 %	01
05	Reductive amination	3.0 %	n/a
06	Reduction to amine	2.6 %	12
07	S_NAr (*N*-arylation)	2.2 %	n/a
08	*O*-Substitution	2.1 %	03
09	Cross-coupling	2.0 %	05
10	Aldol condensation	1.9 %	n/a
=11	Claisen condensation	1.7 %	09
=11	Enolate alkylation	1.7 %	n/a
13	Hydride reduction	1.6 %	n/a
=14	Lithium carbanion addition	1.4 %	12
=14	S_NAr (Ullmann and variants)	1.4 %	n/a
=14	Acid derivative reduction to amine	1.4 %	n/a
17	Imine/oxime/hydrazone formation	1.2 %	n/a
=18	Grignard addition	1.1 %	12
=18	Michael addition	1.1 %	n/a
=18	Friedel-Crafts acylation	1.1 %	15

of acylation and alkylation chemistries becomes very clear, from drug design to multi-kilogram global production: A, Acylation; B, heteroatom alkylation; C, oxidation; D, reduction; E, C-C bond formation; F, C-C bond formation (metal mediated); G, functional group interconversion; H, functional group addition; and I, heterocycle synthesis (Figure 1.2).

1.2 Properties of solvents

Historical solvent effect studies: The art of dissolution captured the minds of even the earliest chemists, the alchemists. Even Aristotle had something to say on the subject, remarking "*No coopora nisi fluida*", or "*No reaction occurs in the absence of solvent*" [Tanaka 2000]. Whether he

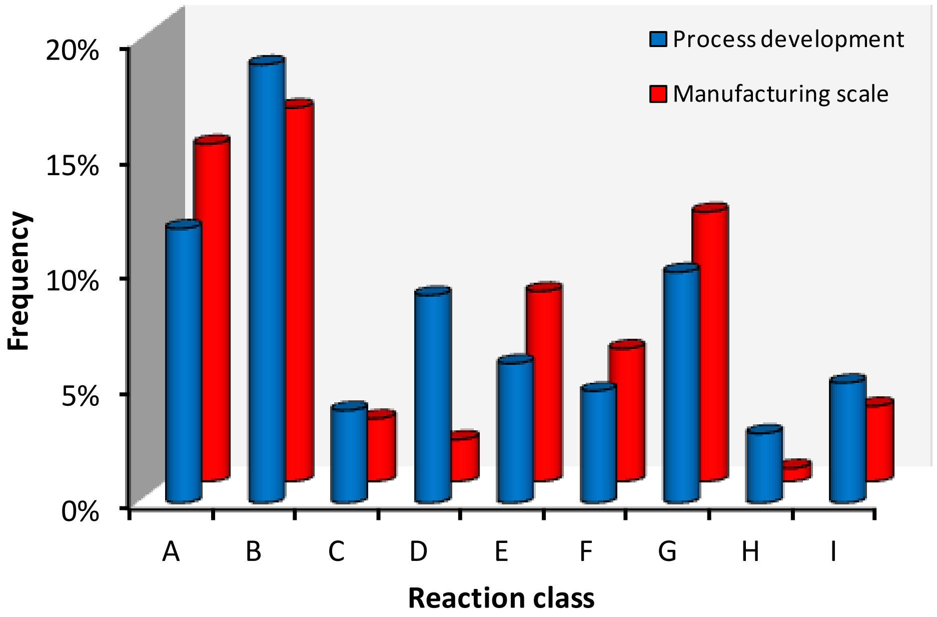


Figure 1.2 A comparison between reaction class frequency in process chemistry and manufacturing plant chemistry within the pharmaceutical industry.

was actually right or not is still debated to this day with the emergance of ‘solventless’ reactions [Mack 2012 page 297, Welton 2006]. Christian Reichardt gives a good account of these first endeavours towards understanding the interaction between solvent and solute, and so will not be repeated here in any great detail [Reichardt 2003 page 2]. The pioneering work of Berthelot and Péan de Saint-Gilles is worth discussing, who in 1862 documented the influence of the solvent on the esterification of acetic acid with ethanol. This was the first time that the role of the solvent on the rate of a chemical reaction was recognised. The other classic solvent effect study occurred some thirty years later, when Menschutkin noticed thousand fold rate enhancements in the reaction rate constant between amines and chloroalkanes depending on the choice of solvent. The influence of the solvent over chemical equilibria did not escape the chemists of the nineteenth century either because studies concerning diketo-enol tautomerisation conducted by Claisen and his contemporaries also revealed a dependence on the nature of the liquid medium in which the substrates were dissolved [Claisen 1896]. Perhaps surprisingly these three chemical systems feature within this work, revisited 100-150 years after they were first studied. The reason for this, in part, is a renewed interest in the origins of chemical products in light of concerns over the future security of crude oil and natural gas reserves. Having said this, with varying motivations these chemical systems have been studied periodically throughout the twentieth century and into the twenty first century, highlighting a continuing interest in the role of solvents in organic chemistry.

Dissolution: The number of solvents in common use might appear disproportional to their responsibility, which may be mistaken for a passive role at first glance. So why are so many different solvents used in synthesis? The use of solvents in purification highlights the reason for the abundance of solvents available. Purification by recrystallisation or column chromatography relies on the polarity of the solvent to match certain components of the mixture but not others at a given temperature. Selecting a solvent with an unsuitable polarity will result in poor separation of the desired components from each other and any impurities. Much the same applies in a reactive system where it is desirable to dissolve the reactants, stabilise intermediates along the reaction pathway, suppress formation of side-products, promote a favourable equilibrium position, and sometimes affect dissolution of the product for easy removal. All this depends on the polarity of the solvent [Adams 2004 page 6].

A complementary match between the polarities of solvent and solute will result in favourable mixing and high solubility (Figure 1.3). If the Gibbs free energy increases upon

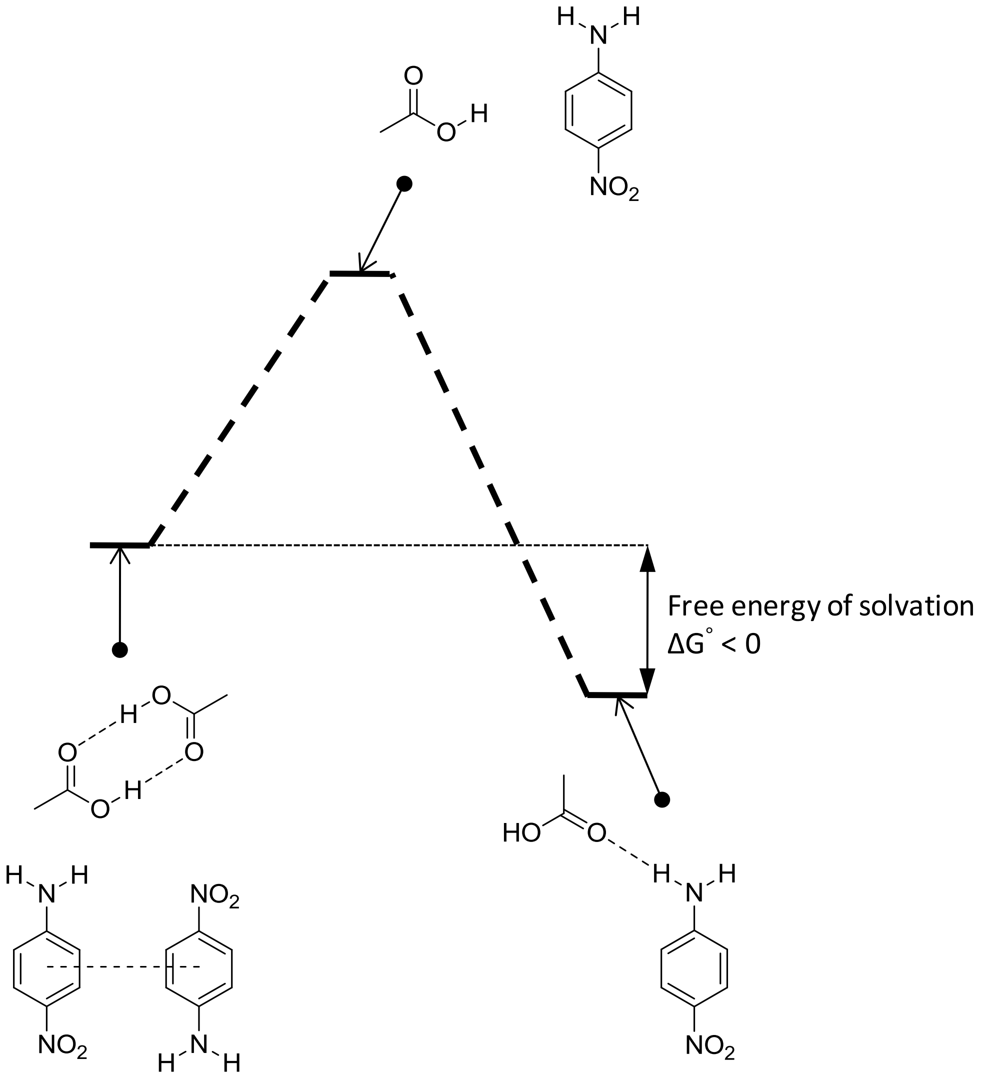
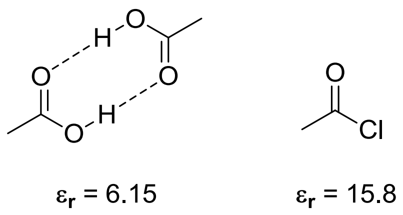


Figure 1.3 An energy profile of 4-nitroaniline dissolving in acetic acid.

breaking solvent-solvent bonds and solute-solute bonds and replacing these interactions with intermolecular solvent-solute bonds then the substrate will not be soluble. If the converse is true then dissolution will be favourable. This thermodynamic description has no bearing on the rapidity of the process, which is defined by kinetics.

In addition to solubility arguments, solvent melting and boiling points are often crucial to the chemistry they are applied to, along with other more application specific properties. Of course properties defining health, safety, and the environmental impact of solvents are also factored into solvent selection although they do not have a direct bearing on the chemistry. Although in total many solvent properties are relevant, giving rise to the multitude of solvents currently in use, a large number of physical properties such as boiling point and viscosity are a consequence of polarity (or more precisely the strength of intermolecular forces that are associated with polarity) which in turn arises from the molecular structure of the solvent. Even some modes of chemical toxicity can be partly attributed to polarity, with aquatic toxicity partially dependant on the lipophilicity of the molecule [Schultz 2006].

A term frequently associated with solvents when it comes to describing dissolution is polarity, although it is often only understood in a very qualitative and simplistic way. Polarity has been defined as "*the overall solvation capability for solutes which in turn depends on the action of all possible, nonspecific and specific, intermolecular interactions between solute ions or molecules and solvent molecules, excluding such interactions leading to definite chemical alterations of the ions or molecules of the solute*" [Muller 1994]. The number of approaches by which it is possible to measure polarity is probably detrimental to its understanding rather than useful. Relative permittivity (also known as dielectric constant, $\boldsymbol{\varepsilon}_r$) is widely used because of, and not despite of its age. It is a single parameter, 'non-intrusive' method of obtaining a measure of solvent polarity using a capacitor [Abboud 1999]. The scale of $\boldsymbol{\varepsilon}_r$ is essentially capped at the upper end by water ($\boldsymbol{\varepsilon}_r$ = 78.36) with hydrocarbons like heptane sitting at the lower end ($\boldsymbol{\varepsilon}_r$ = 1.92). Without being derived from the strength of a solvent-solute interaction like 'intrusive' methods of polarity determination are, the bulk behaviour of the medium is assessed rather than any specific interactions. The downfall of this attempt at understanding the solvent can be highlighted with acetic acid ($\boldsymbol{\varepsilon}_r$ = 6.15) [Lide 1991]. The obviously strong potential of acetic acid for hydrogen bonding is not accounted for by $\boldsymbol{\varepsilon}_r$ and so misrepresents the polarity of this acidic solvent. The specific reason for this is hydrogen bonding between pairs of acetic acid molecules, which persist even as a gas [Chocholoušová 2003]. This uses up the available intermolecular interactions, and in doing so creates an apparently non-polar environment (Scheme 1.1). The introduction of a solute defies this interpretation, with strong solvent-solute interactions readily formed in acetic acid solutions. To this end $\boldsymbol{\varepsilon}_r$ is poor when it comes to establishing relationships between molecular phenomena and the medium.



Scheme 1.1 The relative permittivity of acetic acid and acetyl chloride for comparison.

There is disagreement even between this type of single parameter non-intrusive (bulk) solvent polarity scale. Again hydrogen bonding is responsible. The Hildebrand solubility parameter ($\boldsymbol{\delta_T}$) is the square root of the cohesive energy density of the solvent [Hansen 2007 page 2, Reichardt 2003 page 9]. The conventional representation of $\boldsymbol{\delta_T}$ actually uses an H subscript but this conflicts with another parameter to be introduced shortly. It is a measure of the strength of solvent-solvent interactions, gauged by the enthalpy of vaporisation and the molar volume of the solvent in such a way to provide a term representing energy density:

Equation 1.1

$$\boldsymbol{\delta}_{\mathbf{T}}^2 = \frac{\Delta\mathrm{H}_{\mathrm{vap}} - \mathrm{RT}}{\mathrm{V}_{\mathrm{m}}}$$

A comparison between $\boldsymbol{\delta_T}$ and $\boldsymbol{\varepsilon_r}$ reveals a discrepancy because $\boldsymbol{\delta_T}$ accounts for the hydrogen bonding between solvent molecules more effectively (Figure 1.4). Accordingly protic solvents are given an extra emphasis by $\boldsymbol{\delta_T}$, distinguishing these solvents from aprotic solvents. Nevertheless the trends for protic and aprotic solvents when $\boldsymbol{\delta_T}$ is plotted against $\boldsymbol{\varepsilon_r}$ converge within the proximity of acetic acid ($\boldsymbol{\delta_T}$ = 21.4 $\mathrm{MPa}^{½}$) showing that both scales agree on the polarity of acetic acid as being relatively low. So is it in fact correct to conclude that acetic acid is a low polarity solvent? The answer to the successful modelling of solubility lies with removing the restrictions of a single parameter measurement of polarity.

To model solubility the Hansen solubility parameters of dispersion forces ($\boldsymbol{\delta_D}$), polarity ($\boldsymbol{\delta_P}$), and hydrogen bonding ($\boldsymbol{\delta_H}$) have proven to be successful [Hansen 2007 page 4]. These are derived by splitting the Hildebrand solubility parameter into these three constituent parts:

Equation 1.2

$$\boldsymbol{\delta}_{\mathbf{T}}^2 = \boldsymbol{\delta}_{\mathbf{D}}^2 + \boldsymbol{\delta}_{\mathbf{P}}^2 + \boldsymbol{\delta}_{\mathbf{H}}^2$$

The square of each Hansen solubility parameter is therefore an energy density term derived from the strength and nature of specific solvent-solvent interactions contained within the Hildebrand solubility parameter. Unlike the parent term $\boldsymbol{\delta_T}$ (and $\boldsymbol{\varepsilon_r}$), between them the three Hansen parameters give a good account of the strength of the different bulk solvent interactions that define polarity relevant to dissolution. Matching the Hansen solubility parameters of a solute to solvent candidates gives a fairly reliable method of predicting and rationalising solubility.

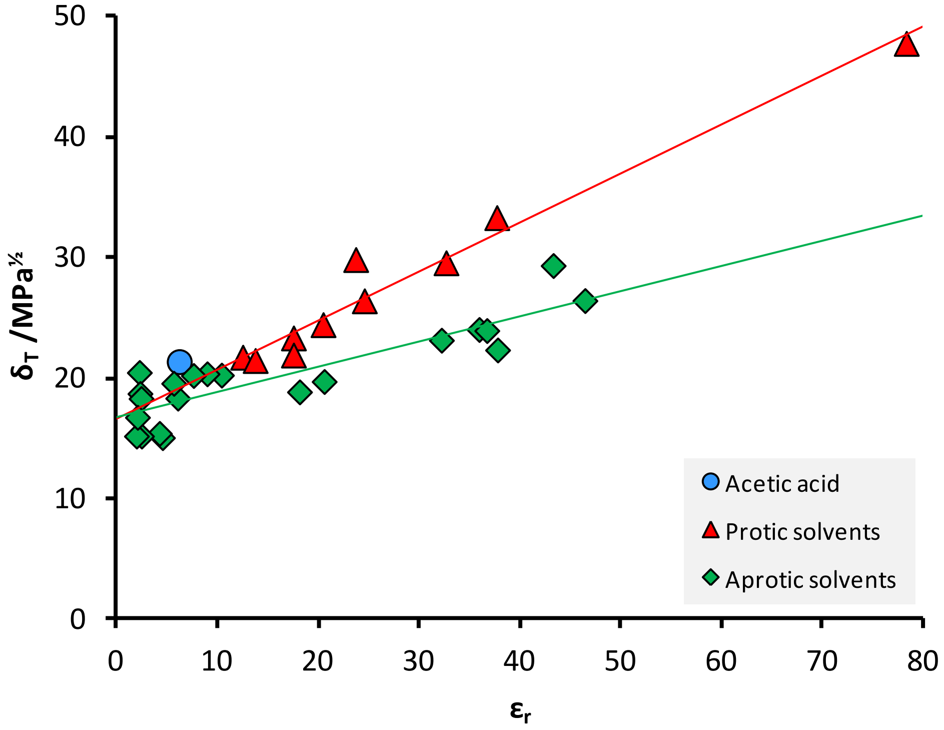


Figure 1.4 A comparison between relative permittivity and the Hildebrand solubility parameter.

A clear and often used demonstration of the Hansen solubility parameters is the differing solubility of ethanol and nitromethane in water [Hansen 2007 page 134]. Because they have similar $\boldsymbol{\delta}_T$ values this cannot be explained by this single parameter approach. The greater $\boldsymbol{\delta}_H$ value of ethanol ($\boldsymbol{\delta}_H$ = 19.4 MPa^{½}) matches the polarity of water ($\boldsymbol{\delta}_H$ = 42.3 MPa^{½}) much more closely than that of nitromethane ($\boldsymbol{\delta}_H$ = 5.1 MPa^{½}). The Hansen parameters are reliably calculated using computational methods [Hukkerikar 2012]. The Hansen solubility parameters can be represented three dimensionally relative to the polarity of a solute (Figure 1.5). Using urea as an example, positioned at the origin of Figure 1.5, solvents are to be found within a close proximity of the solute. Non-solvents have less closely related polarities. The solubility sphere is indicated with a dashed line, which is derived through empirical observation. Those fluids that dissolve urea fall inside the sphere and are indicated by green data points. Ethanol is one of these. Non-solvents are marked with red data points. The very high polarity of urea pushes most solvent candidates out of the solubility sphere.

Solvent polarity scales: The Hansen solubility parameters can only provide an explanation of dissolution. Intrusive measurements of solvent polarity that are based on solvent-solute interactions and not solvent-solvent interactions offer an insight into solvent effects in reactive systems. An example of such a quantitative polarity scale is that arrived at from the UV-vis.

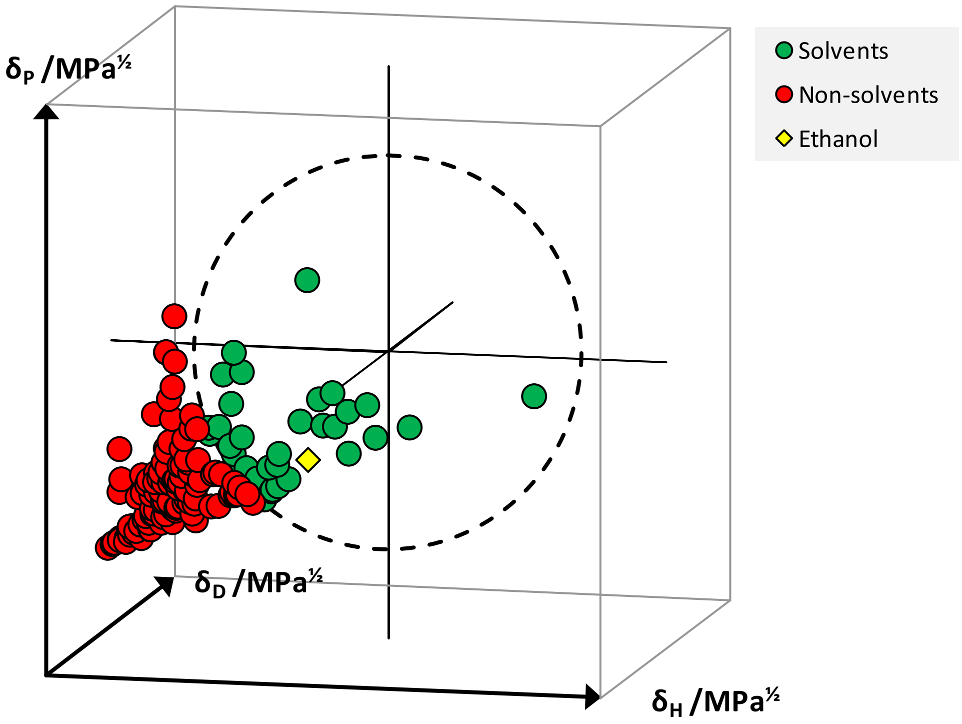


Figure 1.5 Three dimensional Hansen plot based on the solubility of urea.

spectra of Dimroth-Reichardt's betaine dye [Dimroth 1969, Reichardt 2003 page 411]. The ground state of this polarity probe is more polar than the excited state in which the negative charge of the zwitterion is spread over an extensive conjugated system and no longer localised on the phenoxide oxygen. This means that the ground state is more stable in highly polar solvents than it is low polarity solvents. The opposite is true for the excited state which is relatively stabilised in solvents of low polarity. This results in an energy gap between the two electronic states that is very susceptible to the medium in which the dye is dissolved. This phenomenon is known as solvatochromism [Reichardt 2003 page 330]. Dimroth-Reichardt's betaine dye is an example of a negatively solvatochromic (or hypsochromic) compound with the UV absorbance wavelength decreasing with increasing solvent polarity [Reichardt 1994]. This can be expressed as an energy term as follows:

Equation 1.3

$$\mathrm{E_T}\ /\mathrm{J\cdot mol^{-1}} = \mathrm{h\cdot c\cdot \tilde{\nu}\cdot N_A}$$

Accordingly the energy gap between the ground state and electronic excited state of the dye is very important (Figure 1.6). A simple scale of solvent polarity can be devised based on the observed absorption wavelength, normalised between tetramethylsilane (TMS, $\mathbf{E_T^N}$ = 0.000) and water ($\mathbf{E_T^N}$ = 1.000):

Equation 1.4

$$\mathbf{E_T^N} = \frac{\mathrm{E_T - E_T(TMS)}}{\mathrm{E_T(water) - E_T(TMS)}}$$

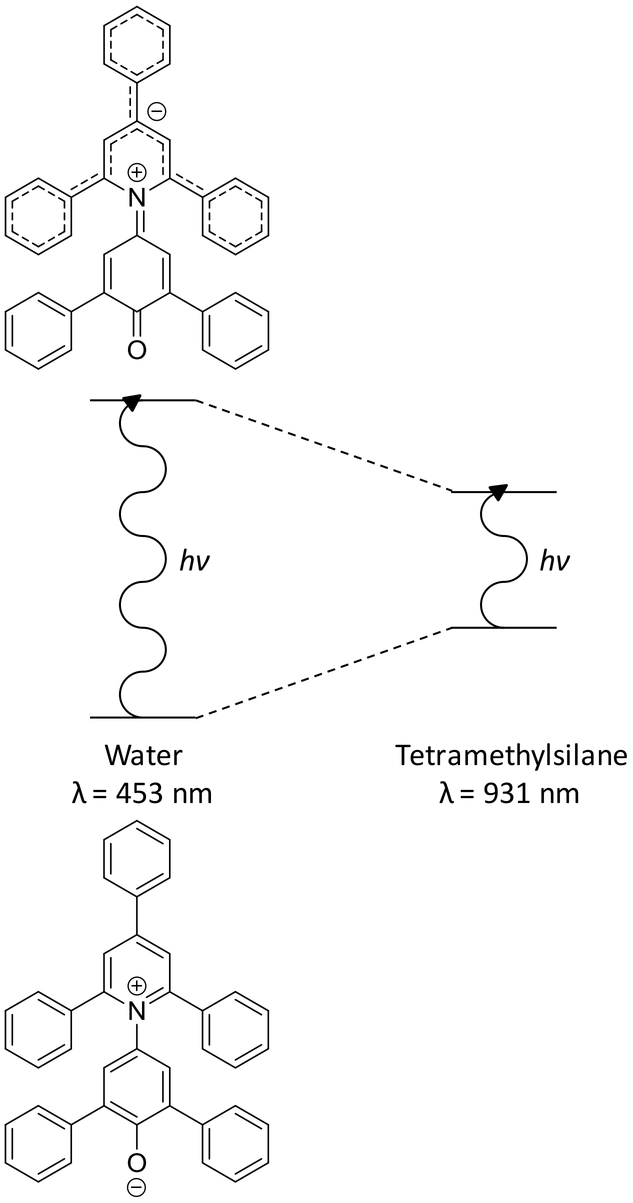


Figure 1.6 The solvent effect influencing the light absorbance of Dimroth-Reichardt's betaine dye.

The $\mathbf{E_T^N}$ scale is often used to represent solvent polarity in correlations with chemical kinetics and equilibria [Reichardt 1979]. Dipolarity is accounted for, but on this scale proticity has an almost additive effect, and this propels hydrogen bond donors into the high polarity regions of the scale. Whereas the most dipolar but aprotic solvents tend to hit a maximum at $\mathbf{E_T^N}$ = 0.500, protic solvents can have polarities double this value (Figure 1.7) [Mistry 2008 page 57]. Accordingly acetic acid is finally represented as a high polarity solvent ($\mathbf{E_T^N}$ = 0.648) [Reichardt 2003 page 422]. One drawback of this probe molecule is its poor solubility in low polarity solvents, and the protonation of the phenoxide in acidic solvents. Nile red is an alternative dye for lipophilic media, while the polarity of acidic solvents can be measured with Kosower's dye [Kosower 1958, Moog 2004].

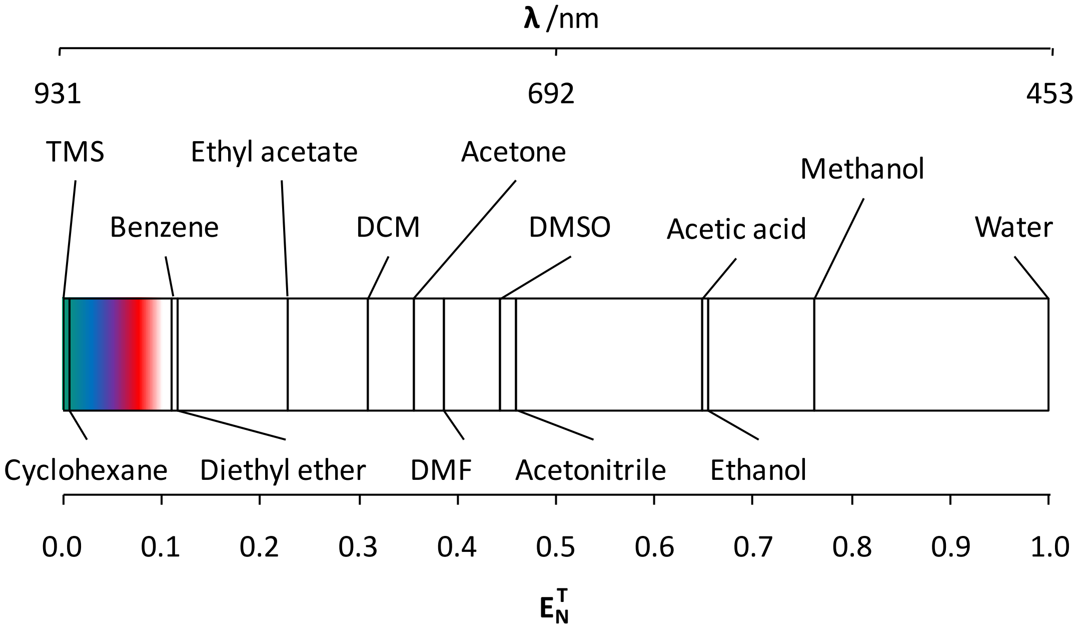


Figure 1.7 The solvent polarity scale derived from Reichardt's dye.

We have seen that single scales of polarity represent a mixture of interactions and are not always useful for constructing correlations with. The Kamlet-Taft solvatochromic scale of solvent polarity on the other hand offers a three parameter assessment consisting of hydrogen bond donating ability (**α**), hydrogen bond accepting ability (**β**), and dipolarity/polarisability (**π***) all based on solvent-solute intermolecular bonding strengths [Kamlet 1983]. The separation of these interactions is useful for resolving relationships between observed solvent effects and the polarity of the solvent. They are suited to this role because, like $\mathbf{E_T^N}$, the resultant solvent polarity scale is derived from the energy differences between two electronic states of a dye molecule. These solvent dependant energy differences mirror the variable energy levels of kinetic and equilibrium profiles in chemical transformations.

The UV-vis. absorbance maxima of *N*,*N*-diethyl-4-nitroaniline in solution, normalised between the observed wavelengths in cyclohexane and DMSO (although usually expressed as wavenumbers in units of cm^{-1}) is used to obtain values of **π*** [Kamlet 1977]. This is analogous in construction to the $\mathbf{E_T^N}$ polarity scale of Reichardt:

Equation 1.5

$$\boldsymbol{\pi}^* = \frac{\tilde{\nu} - \tilde{\nu}(\mathrm{CyH})}{\tilde{\nu}(\mathrm{DMSO}) - \tilde{\nu}(\mathrm{CyH})}$$

Alternative dyes can be used, and with normalisation give similar results in most cases [Kamlet 1979]. *N*,*N*-Diethyl-4-nitroaniline is sometimes replaced by its dimethyl- analogue or 4-nitroanisole because these dyes offer a sharper and more consistent band shape [Laurence 1994]. Sometimes average **π*** values are used to soften the effect of outliers unique to certain dyes [Marcus 1993]. However *N*,*N*-diethyl-4-nitroaniline was established as a solvatochromic probe even before the **π*** scale was established and understandably is the first choice in many studies of non-specific solvent effects [Crowhurst 2003, Kamlet 1977]. Regardless of the dye used, **π*** as a combined measure of dipolarity and polarisability relies on the difference in polarity between the electronic ground state and the excited state of the probe molecule (Figure 1.8). The opposite of Dimroth-Reichardt's betaine dye, the excited electronic state of *N*,*N*-diethyl-4-nitroaniline is more polar than the ground state, and so the energy gap between the two now decreases with increasing solvent polarity. Accordingly *N*,*N*-diethyl-4-nitroaniline is known as a positively solvatochromic (or bathochromic) compound. The range of wavelengths that can be obtained is not as broad as that of Dimroth-Reichardt's betaine dye but sufficient to produce a reliable scale (Figure 1.9).

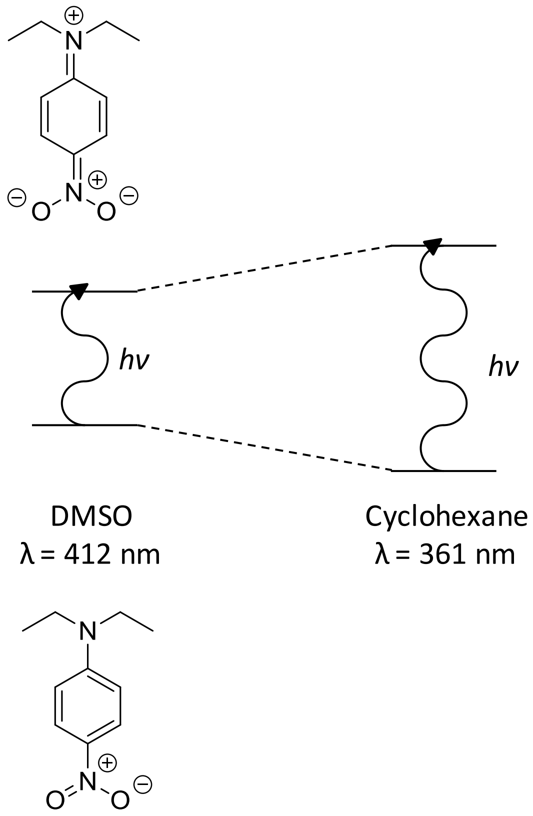


Figure 1.8 The solvatochromism of *N*,*N*-diethyl-4-nitroaniline.

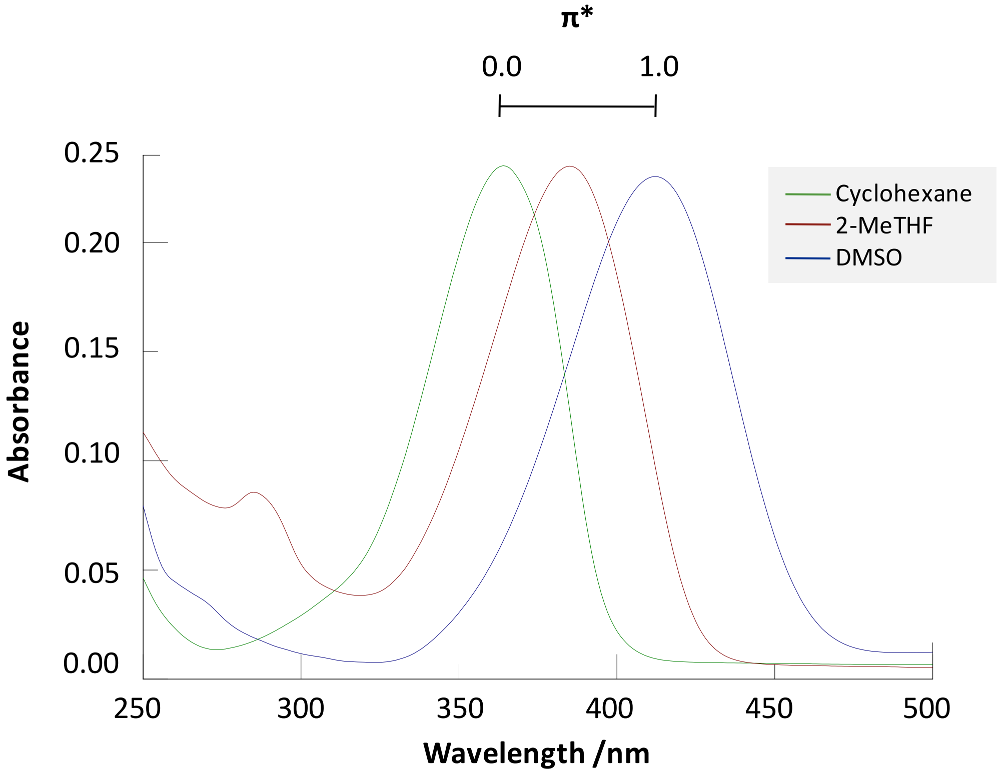
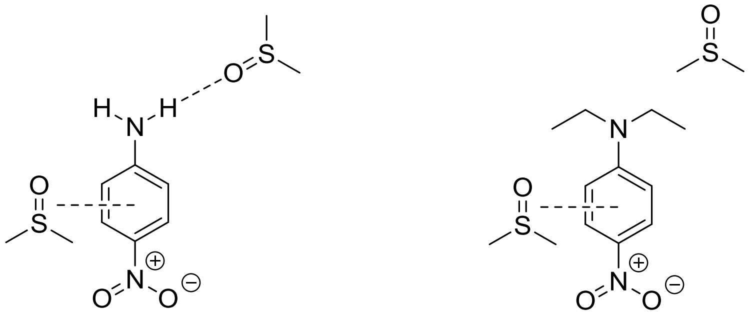


Figure 1.9 The UV-vis. spectrum of *N*,*N*-diethyl-4-nitroaniline in cyclohexane, 2-MeTHF, and DMSO.

Values of **β** representing solvent hydrogen bond accepting ability are obtained in conjunction with another dye, 4-nitroaniline [Kamlet 1976]. Solvents that cannot engage in hydrogen bonds as an acceptor (therefore **β** = 0.00) stabilise 4-nitroaniline with the same non-specific interactions that form the basis of interactions with *N*,*N*-diethyl-4-nitroaniline. The relationship between the absorbances of the two dyes can be represented in the following equation, derived from experimentation, where 4-nitroaniline is denoted as dye 1, and *N*,*N*-diethyl-4-nitroaniline as dye 2:

Equation 1.6 $$\tilde{\mathbf{v}}(1) = 1.00 \cdot \tilde{\mathbf{v}}(2) + 3.44 \cdot 10^{-3}$$ $R^2 = 0.994$

The departure of observed 4-nitroaniline UV absorbances from the response expected due to Equation 1.6 provides the basis of **β** [Nicolet 1986]. The magnitude of any deviation is proportional to the strength of the solvent's hydrogen bond accepting ability onto 4-nitroaniline. Of course this mode of stabilisation is not possible for *N*,*N*-diethyl-4-nitroaniline (Scheme 1.2). The excited state of 4-nitroaniline with its deshielded amino group will strive to generate stronger hydrogen bonds with the solvent than the ground state requires. This results in the variable energy gap between the electronic states of the dye needed for the polarity scale.



Scheme 1.2 The solvent interactions of 4-nitroaniline and *N*,*N*-diethyl-4-nitroaniline as exemplified with DMSO.

The upper end of the **β** scale was originally normalised with hexamethylphosphoramide (HMPA) set as **β** = 1.00 [Kamlet 1976]. This resulted in the polarity of DMSO being **β** = 0.74. The difficulties of using HMPA (namely toxicity and hydrolysis when exposed to air) mean that DMSO is an attractive upper bookend of the **β** scale for modern studies. Thus the equation for the calculation of **β** can be scaled to retain the historical values of hydrogen bond accepting ability:

Equation 1.7

$$\boldsymbol{\beta} = 0.74 \cdot \frac{\tilde{\nu}(\mathbf{1})^{\text{calculated}} - \tilde{\nu}(\mathbf{1})^{\text{observed}}}{\tilde{\nu}(\mathbf{1})^{\text{calculated for DMSO}} - \tilde{\nu}(\mathbf{1})^{\text{observed in DMSO}}}$$

There is a clear hierarchy separating alkanes and chloroalkanes, aromatic solvents, modestly basic oxygenated solvents and the highly polar amides and sulphoxides (Figure 1.10). The aromatic solvents have been incorporated into the baseline with the non-hydrogen bond accepting solvents in the past, but it is now usually accepted that they can engage with hydrogen bond donors *via* weak pi-orbital interactions [Nishio 2011].

Solvent hydrogen bond donating ability, the opposite effect to **β**, is represented as **α**. Calculation of **α** requires the same approach as for **β**, with Dimroth-Reichardt's betaine dye and 4-nitroanisole originally chosen as the homomorphic dye pairing [Taft 1976]. A shortcut for accessing **α** values is available, based on an empirical relationship developed to describe $\mathbf{E}_{\mathbf{T}}^{\mathbf{N}}$, **π***, and the parameter in question [Marcus 1991]:

Equation 1.8

$$\mathbf{E}_{\mathbf{T}}^{\mathbf{N}} = -0.02 + 0.49\boldsymbol{\alpha} + 0.38\boldsymbol{\pi}^{*}$$

The range of **α** values originally reported by Kamlet and Taft was capped by methanol (**α** = 1.00) [Taft 1976]. Alcohols and carboxylic acid solvents typically have high **α** values. A considerable number of solvents are aprotic and so understandably have zero **α** values. Of more interest is the range in-between these two extremes. Subtle effects of C-H acidity are recognised for the solvents chloroform (**α** = 0.20) and acetonitrile (**α** = 0.35) although the exact numbers vary between reports [Crowhurst 2003, Marcus 1993].

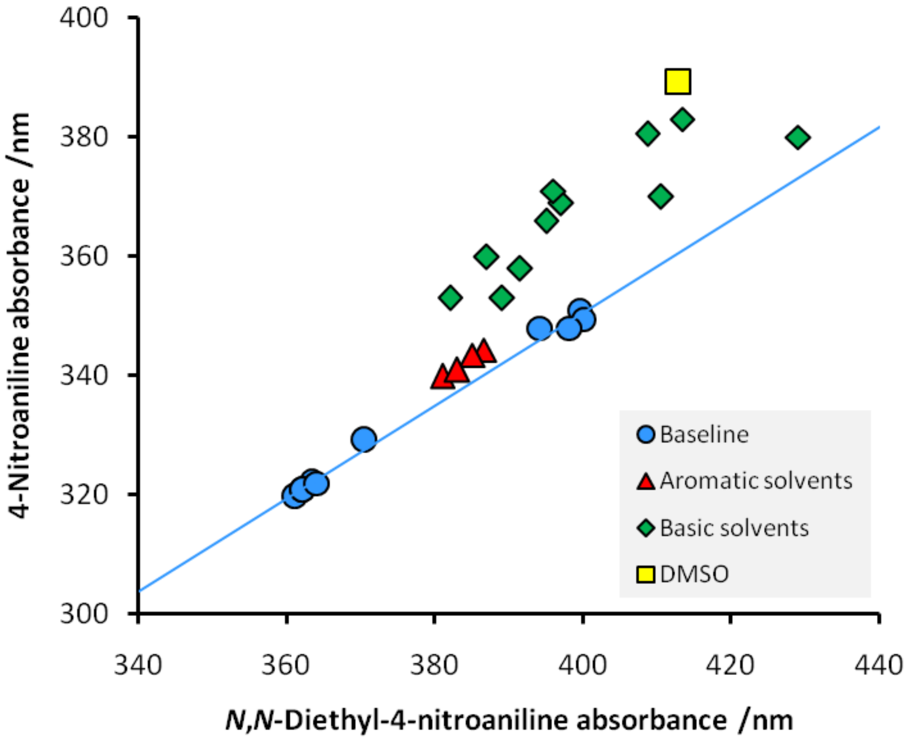


Figure 1.10 A comparison between aniline dye absorbance maxima in different solvents.

Populating the diagram: Collectively bio-based solvents will need to represent the same variety of polarity that we have come to rely on from traditional petroleum derived organic solvents. Protic and aprotic solvents can be presented in separate 2D graphs (by plotting **β** against **π*** in each case) which is regarded as a more helpful representation than a single three dimensional graph [Jessop 2011]. A cut-off point at **α** = 0.37 was chosen to define proticity so that acetonitrile (**α** = 0.35) falls with the aprotic solvents and *t*-butanol (**α** = 0.39) is considered as a protic solvent. Otherwise this treatment follows the same principle as Jessop's original tool for discerning the polarity of green solvents [Jessop 2011]. The resulting polarity maps can be neatly divided into a grid of nine regions which conveniently separate different types of solvent. The rule is not absolute but serves as a guide. In the aprotic solvent polarity map a distinction can be made between dialkyl ethers and cyclic ethers, and the same for organic carbonate solvents (Figure 1.11). Less systematic exceptions include pyridine, which although classed as an amine is more like a highly dipolar aprotic solvent, and the very dipolar but only weakly basic sulpholane. Chlorinated solvents, although inevitably only weakly hydrogen bonding, have dipolarities that vary from carbon tetrachloride (**π*** = 0.28) to DCM (**π*** = 0.82) depending on the overall dipole moment of the solvent [Abboud 1977]. What is most striking about this polarity map is that almost all the possible regions are represented. There are readily available solvents for all imaginable purposes, and presumably this is borne out of necessity and not luxury. Therefore aprotic bio-based solvents need to be found of all conceivable polarities.

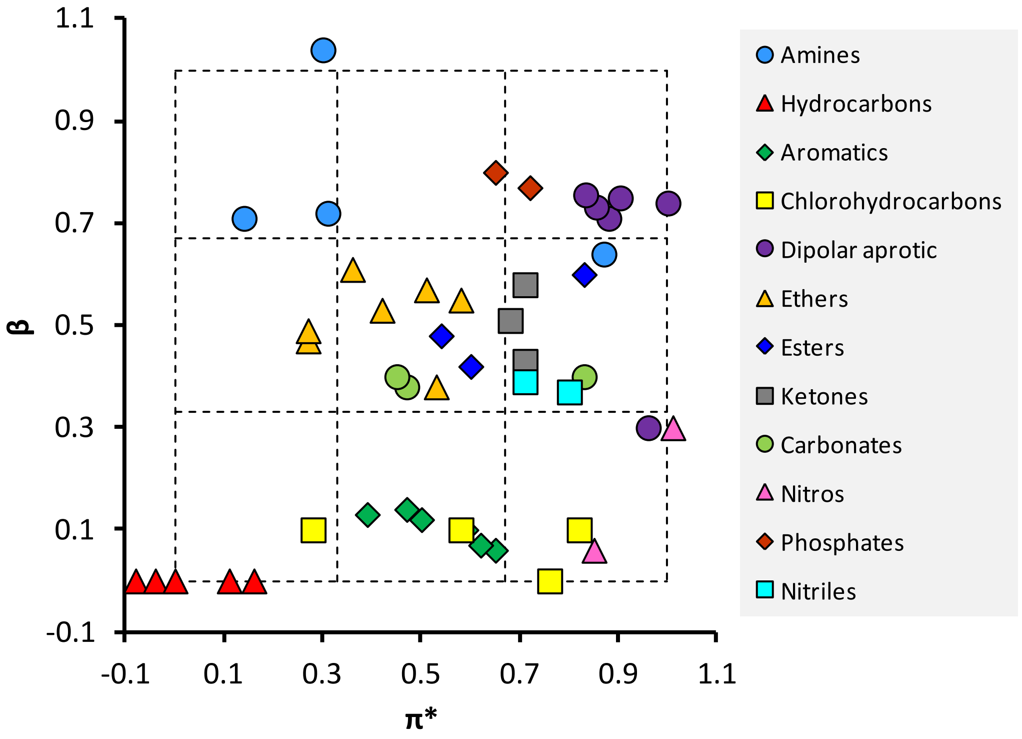


Figure 1.11 Aprotic solvent polarity map.

The protic solvents can also be mapped in the same way (Figure 1.12). Here there is no option for hydrogen bond donating yet non-dipolar solvents. This is not so surprising given that proticity must arise from some degree of bond polarisation. Fortunately there does not seem to be much demand for such a solvent. Weak hydrogen bond acceptors are provided in the guise of fluorinated alcohols. Alcoholic solvents appear to cover a wide spread of polarity, but this is slightly deceiving because the two obscure (but bio-based) solvents glycerol formal and ethyl lactate, both highly dipolar, extend the range usually occupied by more typical alcohols. These two solvents appear in the middle row-right column box rather than the top row-middle column of Figure 1.12 where the other alcohols reside. Lactic acid is much more dipolar than other acids, which appears outside of the boxed area in the proximity of water. A small selection of ionic liquids (1-butyl-3-methylimidazolium triflate, 1-butyl-3-methylimidazolium tetrafluoroborate, and 1-butylpyridinium tetrafluoroborate) are also represented in Figure 1.12 because a cation can give rise to a hydrogen bond donating effect as it stabilises the phenoxide anion of Dimroth-Reichardt's betaine dye (or an equivalent probe molecule) [Ab Rani 2011].

Many more solvents could have been represented on these polarity maps to accentuate what bio-based solvents need to replicate. Data for a wider range of solvents is included in the appendix (Table 8.2). What is clear is that a single bio-based solvent would not suffice as an alternative to even a few conventional solvents. Not only is polarity important of course, but

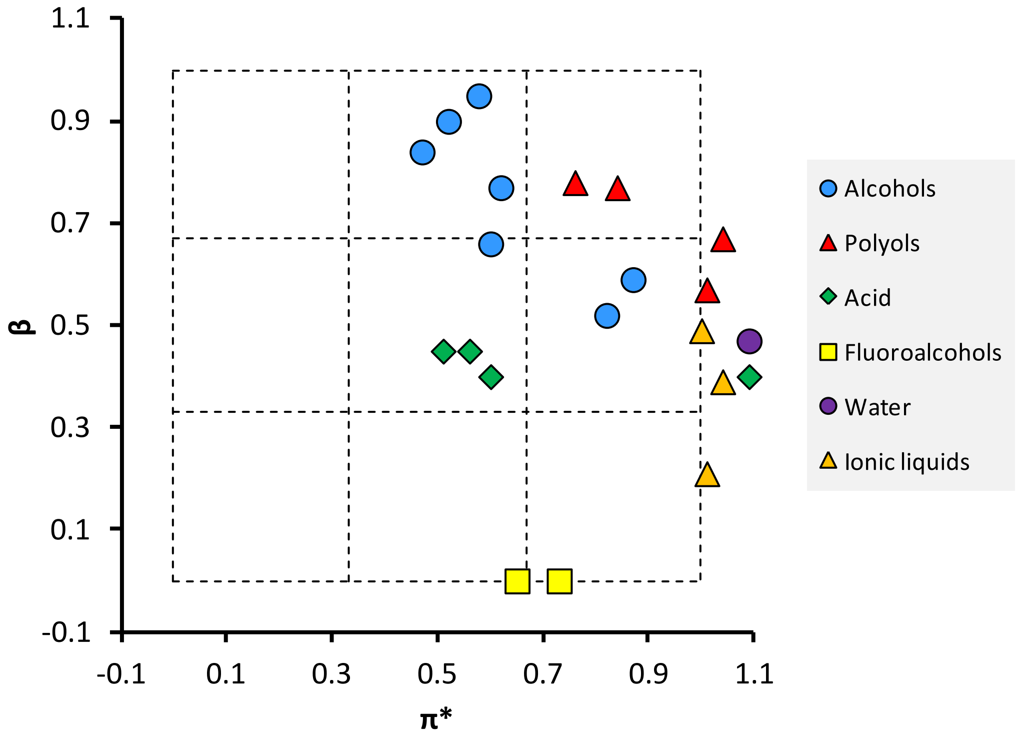


Figure 1.12 Protic solvent polarity map.

boiling point for example is massively significant. The reactivity of solvents is also a vital consideration and specific to the application. The use of Grignard reagents in solution requires a electron donating solvent to be able to stabilise the organometallic solute but if the solvent is also electrophilic it will react, and probably violently. Esters and cyclic ethers occupy a similar position in their polarity diagram (Figure 1.11) but only the latter are a viable option for this chemistry. These are the reasons for the present diversity of solvents. The daunting task required of bio-based solvents, if they are to ever completely supplant non-renewable solvents, is to approach the density and coverage of the solvents represented on Figure 1.11 and Figure 1.12. The tools available to tackle this considerable task are reviewed in the following section.

1.3 Solvent selection

Computational tools: An algorithm for solvent selection has been developed by the CAPEC research group at DTU, Denmark. The methodology was tested in collaboration with GSK and applied to actual chemical systems [Gani 2005]. One example is the aqueous enzymatic oxidation of toluene to its *cis*-glycol. A solvent is required to form a second phase and extract the unreacted toluene. As such the solvent needs to create a phase split with water, dissolve toluene, and be a liquid at the process temperature amongst other things. A toxicity limit was also set (LC_{50} > 10 mg/kg). The solvent selection algorithm refines a large dataset of solvent candidates into only those solvents that can provide the conditions and properties required. A score is attributed to each solvent that accumulates during the algorithm to help solvent selection. Instead of benzene and other viable but undesirable solvents, 2-heptanone was selected for further study based on its algorithm score. A multistep organic synthesis has also been optimised in this way [Gani 2008]. The reactants were not disclosed for confidentiality reasons but the method for selecting each solvent with the algorithm was reported. Along with other case studies, a proof of concept has been established, validating the usefulness of this solvent selection algorithm for improving examples of organic synthesis.

The original solvent selection algorithm refines solvent candidates largely based on the physical state of the solvent at the application temperature, its relative polarity with respect to certain solutes, with the possibly of environmental, health, and safety parameters being introduced as well (Figure 1.13). In this work the basic framework of the algorithm is expanded to encompass more solvents and a greater bias toward synthetic organic chemistry, rather than

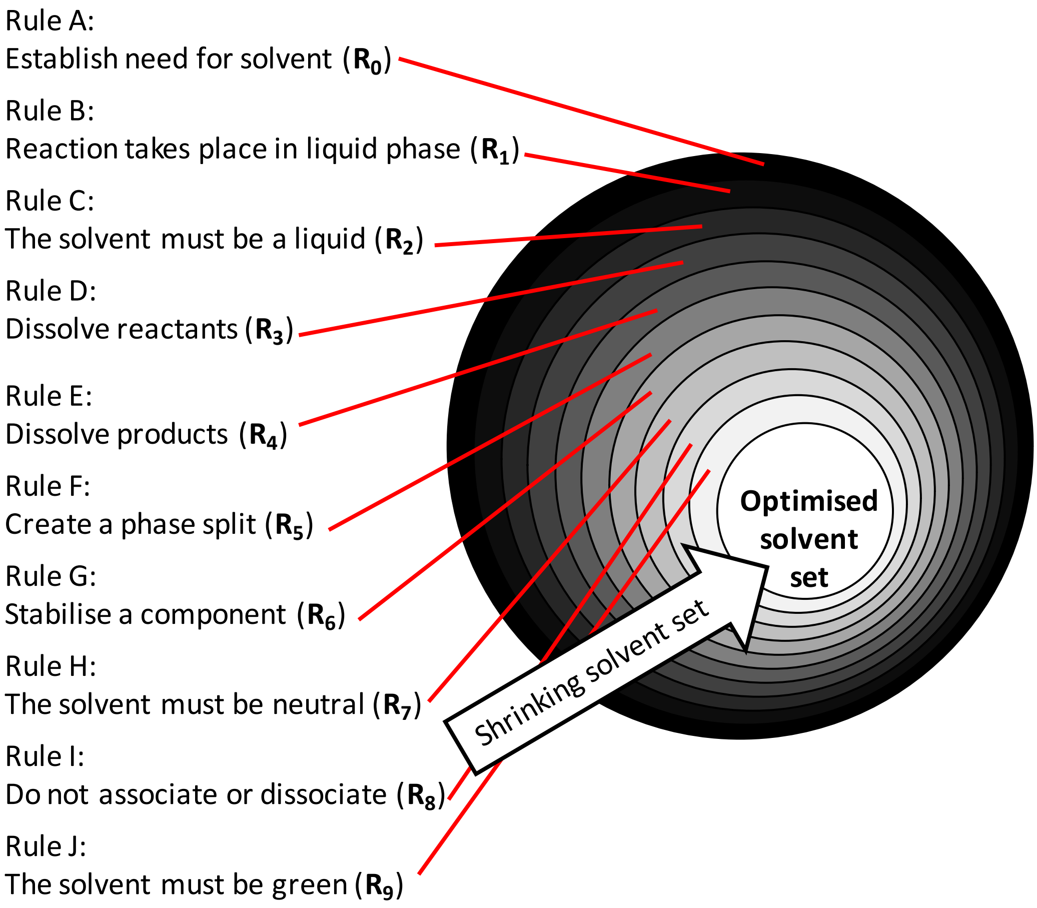

Rule A:
Establish need for solvent ($\mathbf{R_0}$)
Rule B:
Reaction takes place in liquid phase ($\mathbf{R_1}$)
Rule C:
The solvent must be a liquid ($\mathbf{R_2}$)
Rule D:
Dissolve reactants ($\mathbf{R_3}$)
Rule E:
Dissolve products ($\mathbf{R_4}$)
Rule F:
Create a phase split ($\mathbf{R_5}$)
Rule G:
Stabilise a component ($\mathbf{R_6}$)
Rule H:
The solvent must be neutral ($\mathbf{R_7}$)
Rule I:
Do not associate or dissociate ($\mathbf{R_8}$)
Rule J:
The solvent must be green ($\mathbf{R_9}$)
Optimised solvent set
Shrinking solvent set

Figure 1.13 A visual representation of the algorithmic solvent selection process.

chemical engineering as formerly intended. The revised solvent selection algorithm is set up in a computer spreadsheet (Microsoft Excel) which also contains a database of all the relevant solvent properties, either experimental values or predictions (obtained through recognised estimation models) where necessary [Hukkerikar 2012].

Each filtering step, or rule, is known as a reaction index (R_i). Adherence to each rule is typically assessed by way of a five tier assessment. These are called the reaction-solvent indices (RS_i). Each reaction-solvent index is converted into a score (S_i) for that reaction index (R_i), and the summation of the scores from each rule gives a numerical value for every solvent as a measure of its suitability. Scores (S_i) are associated with each reaction-solvent index (RS_i) as in the original solvent selection algorithm. However $RS_i = 5$ always results in $S_i = 1$ which is considered a fail. A solvent candidate only has to fail one reaction index to be removed from the solvent candidate optimisation process. The original scoring system is shown in the following figure, which exemplifies the generally applicable scoring system with rule D (Figure 1.14). Solubility of reaction components is evaluated with $\boldsymbol{\delta_T}$ (rule D and rule E). Rule G uses the Hansen solubility parameters in a more thorough polarity matching exercise.

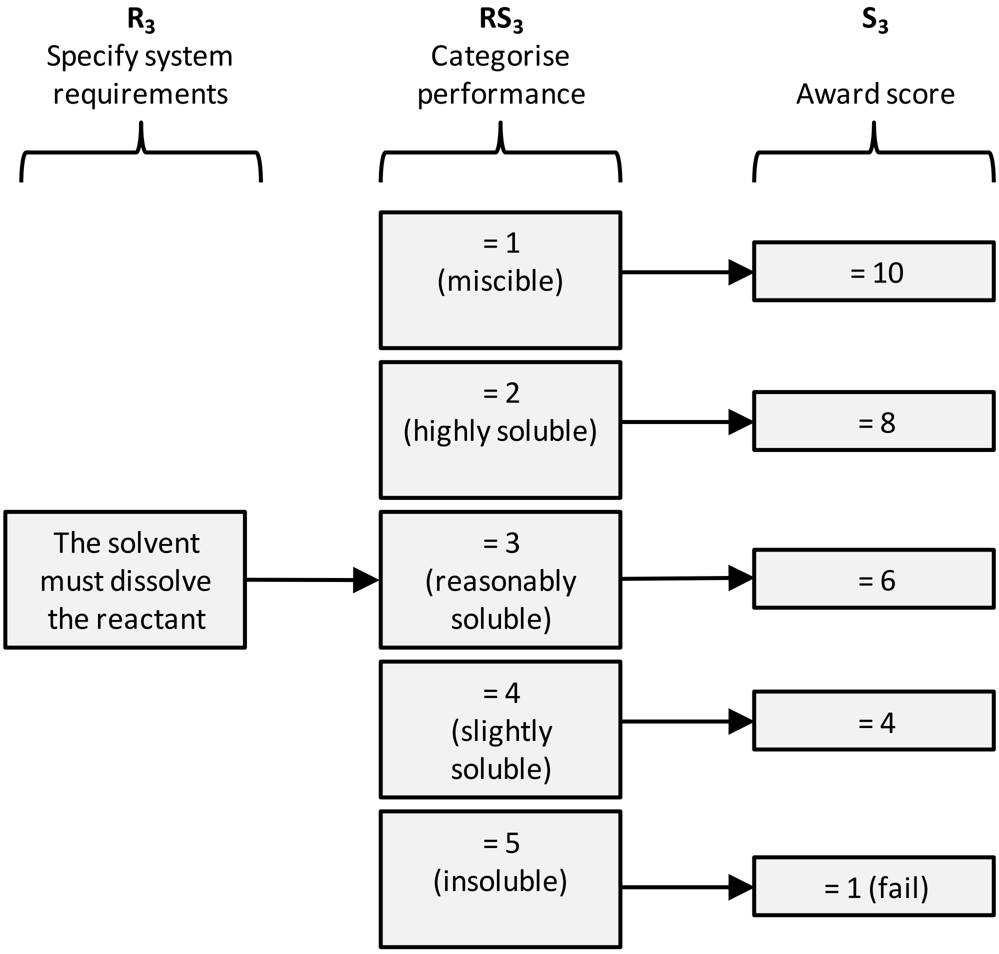

R3
Specify system requirements
RS3
Categorise performance
S3
Award score
The solvent must dissolve the reactant
= 1 (miscible)
= 2 (highly soluble)
= 3 (reasonably soluble)
= 4 (slightly soluble)
= 5 (insoluble)
= 10
= 8
= 6
= 4
= 1 (fail)

Figure 1.14 Rule D of the solvent selection algorithm.

Revised solvent selection algorithm: The greatest change to the original algorithm made in this work is to rule C as defined by R_2. In the original assessment a parameter denoted T_s is calculated as the midpoint between the boiling point and the melting point of a solvent. This is compared to the desired application temperature (T_r) and a score (S_2) awarded (*via* RS_2) based on the proximity of T_s to T_r. With a flexibility margin of only ±20 K this calculation will reliably ensure that only liquids are selected as solvent candidates. However the assessment is harsh and omits many potentially satisfactory solvents. It is common practice to operate under refluxing conditions which is not compatible with rule C of the original solvent selection algorithm.

In the revised algorithm rule C is divided into two sections. Firstly the user decides whether solvent recovery by distillation (large scale) or disposal (small scale) is preferable (R_{2a}). In the first instance solvents that are not gaseous at the desired reaction temperature, yet have low boiling points are favoured in the assessment (Figure 1.15). Solvents with increasingly higher boiling points (T_b) fair less well across a range of 60 K (T_x) above the reaction temperature (T_r), and those solvents with boiling points above that (*i.e.* $T_b \geq T_r + T_x$) are designated $RS_{2a} = 4$ (which

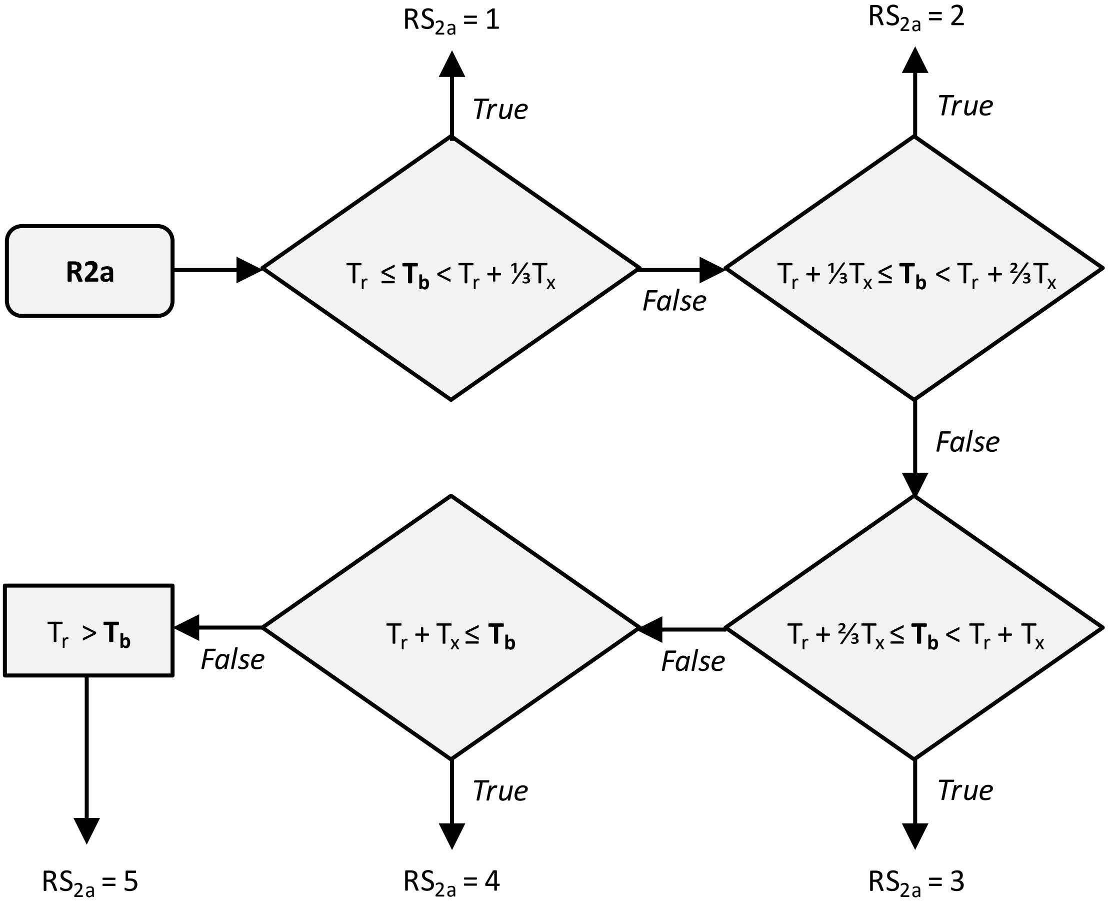


Figure 1.15 Rule C (R_{2a}) of the revised solvent selection algorithm when the solvent is opted to be recycled.

results in the lowest score barring a fail). Solvents that are a gas at the desired reaction temperature are removed from the final solvent set ($RS2_a = 5$). If the alternative solvent disposal scenario is chosen then the scoring gradient is reversed, and high boiling solvents are preferred to minimise losses to the atmosphere. This setting is also suitable for solvents that are to be reused without distillation, for example after induced crystallisation of the reaction components. Again the solvent must not be a gas at the desired reaction temperature. The second part of this rule (R_{2b}) concerns the melting point of the solvent to avoid recommended solvents being solids at the desired reaction temperature. It is generally desirable for a solvent to be liquid at room temperature for handling and purification purposes and so this is also factored into the assessment. Ultimately if the solvent is a liquid at the reaction temperature then it will pass this reaction index.

Rule G is also amended in the revised solvent selection algorithm. Previously stabilisation was determined with a comparison between a solute and each solvent candidate using their respective $\boldsymbol{\delta_P}$ and $\boldsymbol{\delta_H}$ values. This is plotted on a two-dimensional graph rather than the 3D Hansen chart seen earlier in Figure 1.5. This is because of the relatively small changes in the $\boldsymbol{\delta_D}$ parameter between solvents, but also the simplicity of a two dimensional graph is preferred. The maximum permissible discrepancy is ±20% of the magnitude of each Hansen solubility parameter describing the solute. Therefore if the solute has a $\boldsymbol{\delta_P}$ value of 10 $MPa^{½}$ then solvent candidates must have a $\boldsymbol{\delta_P}$ value between 8 $MPa^{½}$ and 12 $MPa^{½}$. This turns out to be hugely restrictive, especially because the $\boldsymbol{\delta_H}$ solubility parameter will impose constraints of its own (again ±20%) to create a small zone of acceptable solvent polarity within the vicinity of the solute (Figure 1.16). The use of a fixed percentage flexibility in the assessment is also unfair because it favours polar substrates with large Hansen solubility parameters. In the revised solvent selection algorithm a percentage leeway is still applied but it is now user defined to be as restrictive or lenient as required in order to approximate the solubility sphere. Solute stabilisation now only accounts for half of rule G in the revised solvent selection algorithm because destabilisation of a particular solute can now also be modelled. The scoring scale assigned to stabilisation (S_{6a}) is reversed for modelling destabilisation (S_{6b}). This may be useful in instances where precipitation of a product is useful.

The final change made to the original solvent selection algorithm relates to how any environmental, health and safety (EHS) parameters are assessed (Rule J). Previously this was accounted for by selecting a target value for a relevant property (LD_{50} perhaps) and then proximity to this value within a limit of ±20% was rewarded [Gani 2005]. This means different solvent candidates may have either a higher or a lower value than the target and be awarded the same score. For example 10% below the target LD_{50} is obviously less favourable than a solvent with a LD_{50} 10% above the target value but these scenarios are rewarded equally. This oddity is

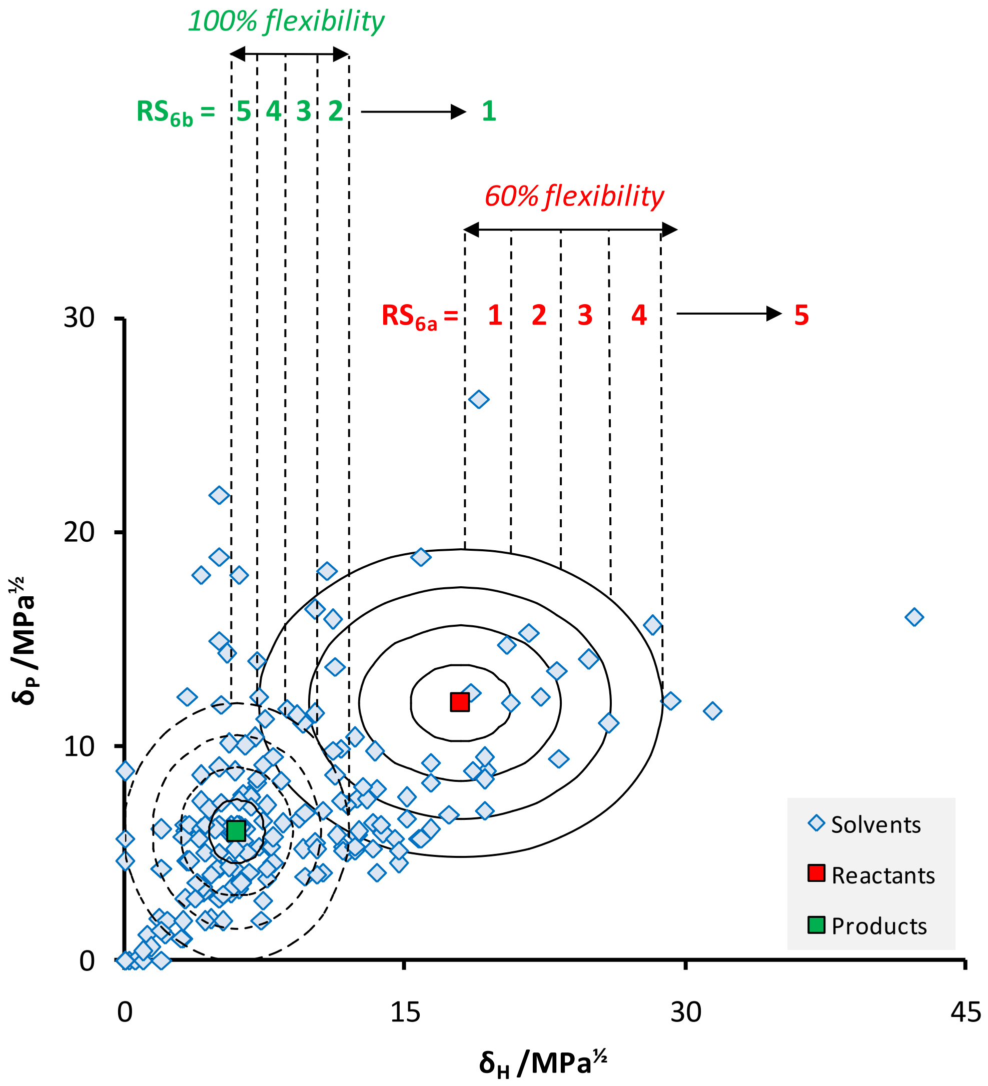


Figure 1.16 An arbitrary polarity map with rule G assignments from the revised solvent selection algorithm.

changed to the user having the option to either maximise or minimise a solvent EHS parameter in the course of solvent selection (Figure 1.17).

Linear solvation energy relationships: The major limitation of the existing solvent selection algorithm (referred to here as model A) is that there are few deviations from intuitive questions and answers, and even then these are limited to thermodynamic models. Kinetics is hugely important, and can used to differentiate between feasible solvent solutions and those solvents that would also pass the solvent selection algorithm but not facilitate the reaction within a meaningful time span. Predictions of solvent performance in a specific application can be made as an additional assessment (model B2), aligned with the revised solvent selection algorithm (now known as model B1). Scales of solvent polarity are known to permit correlations with parameters describing the efficiency of solution chemistry. A linear solvation energy relationship

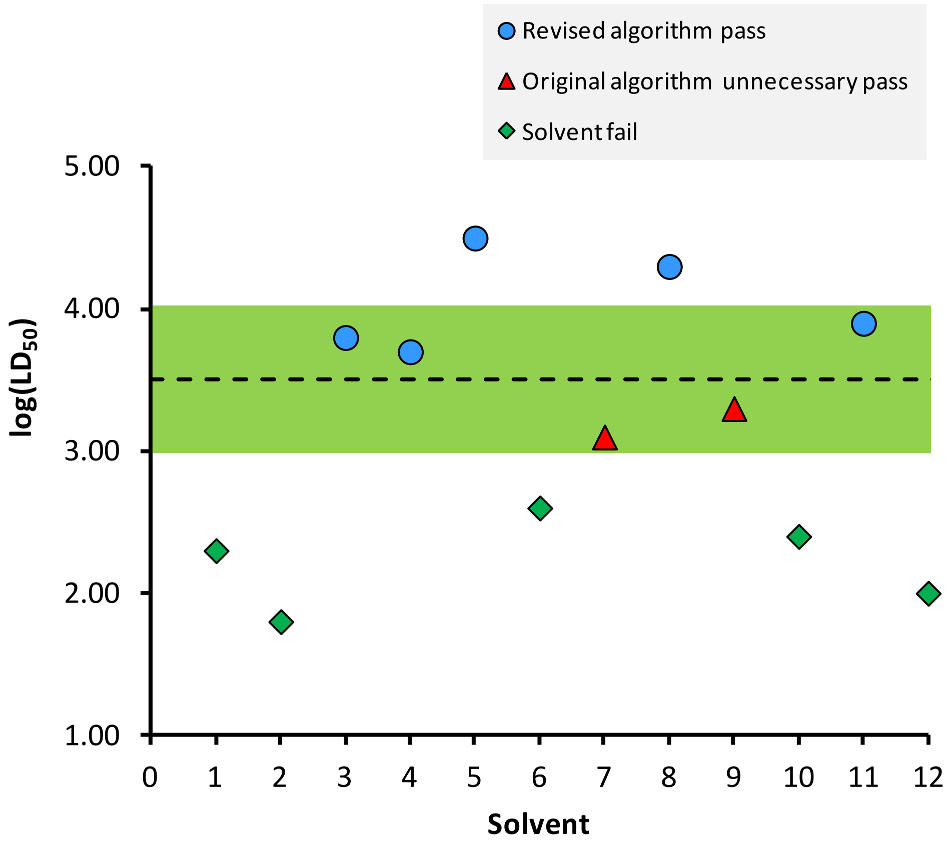


Figure 1.17 A hypothetical comparison between methods for identifying the permissible toxicity limits of solvents when the user defined limit is log(LD_{50}) = 3.5.

(LSER) is a correlation between an energy term (or a parameter proportional to energy such as the logarithm of a rate constant) and a property (or a summation of properties) of a solvent [Reichardt 1979]. Analogous to linear free energy relationships (LFERs), the original and primary purpose of the LSER (as with all linear free energy relationships) is to determine the mechanism of the chemical process being studied. Louis Hammett was at the forefront of developing the LFER approach in the 1930's, developing the use of benzoic acid dissociation equilibrium constants in water (K_a) as a reference system [Hammett 1933].

The Hammett equation models a chemical equilibrium or reaction rate (expressed as ln(**K**) or ln(**k**) respectively) as a function of reagent reactivity, as inferred from the electronic influence of substituents on K_a values ($\boldsymbol{\sigma}$) [Hammett 1937]:

Equation 1.9

$$\log\left(\frac{\mathbf{K}}{\mathbf{K_0}}\right) = \rho \cdot \boldsymbol{\sigma}$$

The value of ρ is determined by the nature of the activated complex [Hammett 1937]. Solvents are not accountable and so reactions are performed in the same solvent for a fair and valid comparison. However the influence of solvents can be seen in arguably the simplest Hammett relationship [Williams 2003 page 17, Hoefnagel 1989]. By plotting the acid dissociation equilibrium constants of benzoic acids in different solvent systems against pK_a, the intensity of the relationship changes, equating to the ability of the solvent to stabilise the charges created by deprotonation, in turn affecting the equilibrium position (Figure 1.18). Adding ethanol to aqueous solutions of benzoic acids reduces the extent of deprotonation, as expected from the introduction of an organic solvent that discourages the formation of ions. Accordingly the gradient of the LFER increases from unity when water is the solvent (effectively pK_a plotted against pK_a) to steeper relationships as the co-solvent concentration is increased. But the most significant difference is the change in intercept, reflecting the reluctance of benzoic acids to deprotonate when dissolved in organic solvent systems.

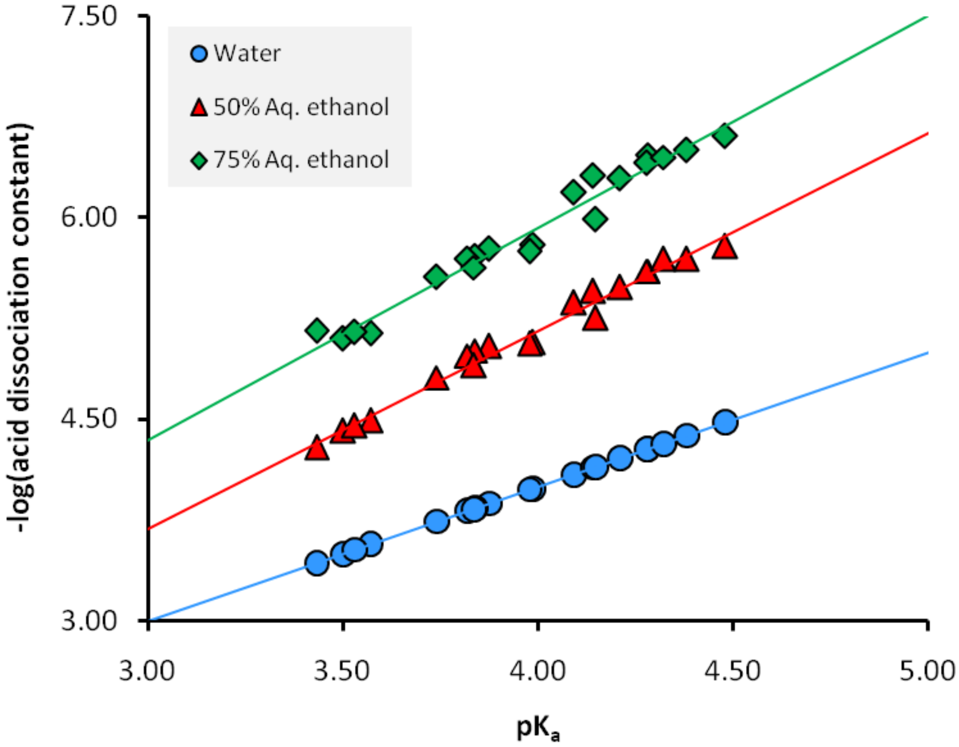


Figure 1.18 A linear free energy diagram showing the acidity of benzoic acids in different solvents.

The most reliable and commonly used versions of the LSER principle are based on the solvent-solute interactions described by the Kamlet-Taft solvatochromic scale of solvent polarity. It provides a reliable empirical means of correlating performance to the nature of the reaction medium, as required in order to ascertain solvent performance in organic synthesis [Williams 2003 page 35]. Although specifics will be introduced when the need arises, the general format of a LSER can be presented with the following equation:

Equation 1.10 $$XYZ = XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^*$$

The energy derived *XYZ* term will usually be ln(**k**) or ln(**K**), but selectivities, enthalpies and entropies are also applicable [Taft 1985]. Not every parameter will be relevant to each system, and there is no limit on the magnitude or sign of each coefficient. The solvent polarity parameter coefficients (*a*, *b*, and *s*) indicate the relative important of each type of solvent-solute interaction, and whether each is beneficial or not. But little further meaning can be gleaned from these coefficients as products of empirical observation. The normalisation of the Kamlet-Taft parameters affects the magnitude of their coefficients too, and so comparisons between terms are only qualitative. Even so an LSER is immensely useful in allocating a mechanism to a process. Indeed studies of solvatochromism and examples of LSERs are prevalent, a testament to their usefulness and validity [Reichardt 1994, Taft 1985]. In constructing one such relationship, Wells *et al*. found that the rates of the Fischer esterification between benzyl alcohol and 2-methoxyacetic acid are inversely proportional to the hydrogen bond accepting ability of the solvent (Figure 1.19) [Wells 2008]. Only **β** of the Kamlet-Taft parameters was statistically significant and the strength of the correlation was satisfactory (R^2 = 0.959). The relationship is indicative of an effect exerted by hydrogen bonding solvents that retards the progress of the reaction. A suggestion from the authors implicates deactivation of the acid catalyst by the solvent.

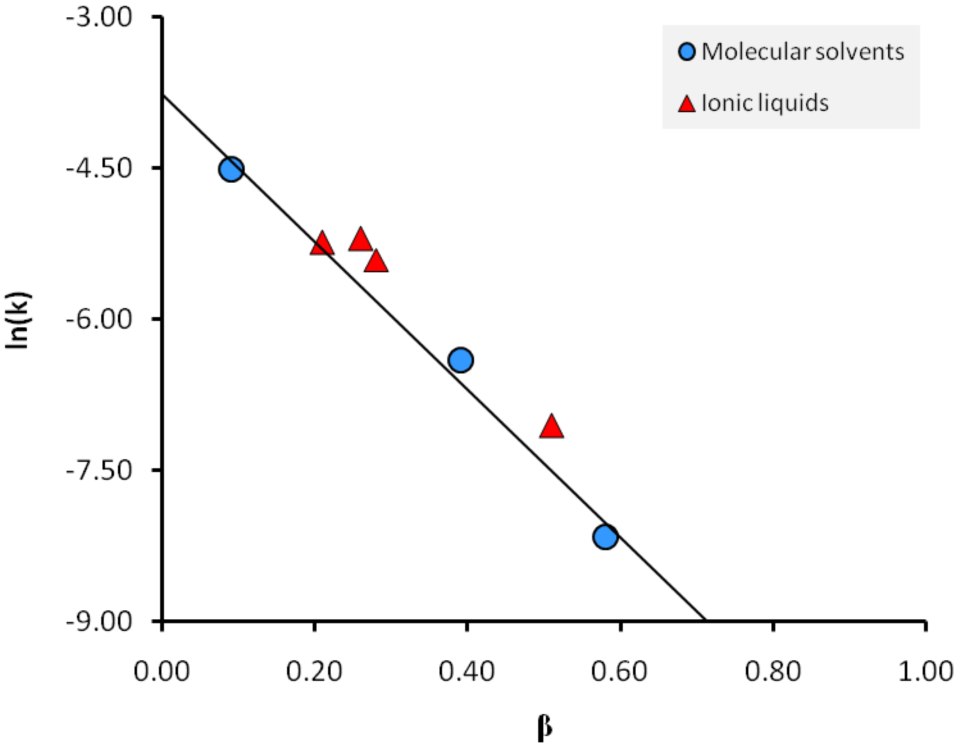


Figure 1.19 The rate of a Fischer esterification correlated to solvent polarity.

The other use of a LSER (the first being the study of mechanism) is to predict the performance of solvents in a given reaction and aid solvent selection on the basis of minimal experimentation. This is the role of the solvent performance assessment (model B2) which accompanies the revised solvent selection algorithm (model B1). In order to have a reliable correlation, the LSER must hold up to statistical tests of accuracy. Random error is a natural consequence of experimental science, although large random errors will disguise the underlying solvent effect. True deviations from a trend set by a LFER are in fact not uncommon, and are indicative of changes in mechanism or rate determining step [Sykes 1981, Williams 2003 page 129]. Once a case study has been represented by a LSER with enough solvents to be sure of the trend, and data obtained to a sufficient level of accuracy as exemplified with Figure 1.19, then it can be reinterpreted as a means of scoring solvents based on their performance. The solvent performance assessment (model B2) ranks solvents based on the magnitude their predicted *XYZ* value (probably ln(**k**) in most instances). Of the solvents in the dataset, the algorithm user chooses the number of solvents they wish to pass the assessment (model B2). The successful solvents are divided into a hierarchy of four quarters: the best performing solvents awarded the score associated with RS = 1, the second quarter of solvents assigned RS = 2, and so on. Those solvents outside the cut-off point receive the equivalent of RS = 5 which, as always, is considered a fail. The results of this assessment (model B2) can be combined with the revised solvent selection algorithm (model B1) with a weighting to enhance the role of solvent performance relative to the rules contained within the usual algorithm.

Solvent selection guides: A solvent selection guide can be used to further scrutinise the solvents that successfully negotiate the revised solvent selection algorithm. Solvent selection guides rate solvents according to their EHS properties to inform users of their risks to human health and the environment. The pharmaceutical industry has long recognised the significance of solvents in their drive to reduce costs and comply with tightening regulation, minimise waste, and lower energy requirements [Curzons 2001]. In order to address this, various prominent companies and institutes have developed solvent selection guides that sort solvents into classes of suitability. These in turn have been incorporated into a solvent greenness assessment (model B3).

The complexity of solvent selection guides is proportional to the scale of the chemistry for which they are intended. Beginning at the smaller end of the scale, tools intended for medicinal chemistry (*i.e*. sub-gram scale synthesis) take the form of colour coded diagrams in which solvents are ordered with respect to their health, safety, and environmental impact. The first solvent selection guide of this sort to be made openly available was developed by Pfizer [Alfonsi 2008]. Although published in 2008, the guide had been in circulation around the laboratories at Pfizer for several years by this time (Table 1.3). The purpose of the guide is to raise awareness about the issue of solvent selection, and reassure chemists about deviating from established procedures where a sound solvent substitution can be made to improve the EHS profile of the reaction. Unfortunately the rationale behind the assignments is not publicly available. Nevertheless most recommendations appear to be reasonable based on safety and toxicity arguments alone.

Table 1.3 A representation of the Pfizer medicinal chemistry solvent selection guide.

Preferred	**Usable**	**Undesirable**
Water	Cyclohexane	Pentane
Acetone	Heptane	Hexane(s)
Ethanol	Toluene	Diisopropyl ether
2-Propanol	Methylcyclohexane	Diethyl ether
1-Propanol	Methyl *t*-butyl ether	Dichloromethane
Ethyl acetate	Isooctane	Dichloroethane
Isopropyl acetate	Acetonitrile	Chloroform
Methanol	2-Methyltetrahydrofuran	DMF
Methyl ethyl ketone	Tetrahydrofuran	NMP
1-Butanol	Xylene(s)	Pyridine
t-Butanol	Dimethyl sulphoxide	DMAc
	Acetic acid	1,4-Dioxane
	Ethylene glycol	Dimethoxyethane
		Benzene
		Carbon tetrachloride

By the time the Pfizer solvent selection guide was published, a GSK process chemistry solvent selection guide of greater complexity had already been available for almost a decade [Curzons 1999]. The basis of this guide has recently been translated into a more useful tool for medicinal chemists, borrowing the colour coded presentation of the rival Pfizer solvent selection guide [Henderson 2011]. The GSK solvent selection guide consists of six constituent assessments (Table 1.4). A numerical scoring system is used to indicate the performance of every solvent in each category. Values range from one (serious issues) to ten (benign). The colour coding highlights the best scores (8-10) in green and the worst scores (1-3) in red. These scales are the amalgamation of various physical properties. For example, the waste category contains an evaluation of the ability to recycle the solvent, but incineration is also considered. A high calorific value is beneficial if the solvent is combusted to provide energy. Conversely, nitrogen and sulphur containing solvents will result in the production of the atmospheric pollutants NO_x and

SO_x when burnt, and are penalised accordingly. Other categories are environmental impact (‘E-Impact’ in Table 1.4), health, flammability (‘Fire’ in Table 1.4), reactivity and life cycle assessment (‘LCA’ in Table 1.4).

Table 1.4 An excerpt from the GSK solvent selection guide supporting table.

Solvent	Waste	E-Impact	Health	Fire	Reactivity	LCA
Acetone	3	9	8	4	9	7
Acetonitrile	2	6	6	6	10	3
1-Butanol	5	7	5	8	9	5
t-Butanol	3	9	6	6	10	8
Chloroform	3	6	3	6	9	6
1,2-DCE	4	4	2	6	10	7
DMSO	5	5	7	9	2	6
1,4-Dioxane	3	4	4	4	5	6
Ethanol	3	8	8	6	9	9
Ethyl acetate	4	8	8	4	8	6
Ethylene glycol	5	8	7	10	9	9
Hexane	5	3	4	2	10	7
2-MeTHF	4	5	4	3	6	4
NMP	5	6	3	9	8	4
Toluene	6	3	4	4	10	7
Triethylamine	4	5	3	4	8	7

Although solvent selection guides highlight the greenness of solvents, what these tools do not directly facilitate is the means to make an educated solvent substitution in favour of a preferable solvent which is also compatible with the application. Reaction specific solvent selection guides have recently been developed that overcome this hurdle, but then of course these guides have a very specialised and limited use [MacMillan 2013, McGonagle 2013]. The two general use solvent selection guides reviewed here have been incorporated into a more sophisticated tool (model B3) to enhance their focus and usefulness. A third guide for solvent selection developed by ETH is also included [Capello 2007]. Minimum acceptable levels of greenness can be set to refine the list of solvent candidates. Of those that remain a summation of the values in the GSK solvent selection guide can be used to establish a hierarchy of solvent greenness. Scores to reflect this are assigned in the same way that the predictions of an LSER were used to generate scores in the solvent performance assessment (model B2). Greater detail of this process, and all the algorithm procedures, will be given during the reaction case studies that follow in the subsequent chapters.

1.4 Analysis of bio-based solvents

A renewability index for solvents: Solvent selection guides adequately cover a range of issues stemming from health and safety legislation. They also represent a lot of environmental concerns resulting from solvent use. Yet all fail to consider the origin of the solvents they discuss. The closest any existing solvent selection guide has come to accounting for this is a tool developed by ETH that accounts for the energy used in the manufacturing process of a solvent [Capello 2007]. It too is incorporated into the greenness assessment (model B3) that accompanies the revised solvent selection algorithm (model B1). Although undoubtedly useful when considering whether it is practical to recycle a solvent, or instead to incinerate it after use and reclaim the energy of combustion, the origin of the solvent remains unaccountable. A new qualitative measure of renewability (denoted with the abbreviation SUS to indicate its relation to sustainability) can be implemented as part of a wider assessment of solvent greenness. A numerical score is assigned to each solvent based on its availability from a biomass feedstock (Table 1.5). Colour coding and the use of a numerical scale are directly taken from the presentation of the GSK solvent selection guide [Henderson 2011]. Solvents that cannot be obtained from biomass are assigned a score of zero. This system will be used in the following case studies to help identify the optimum solvent, given that the renewability of the solvent must be considered as an important factor in solvent selection.

Table 1.5 The basis for assigning a renewability index (SUS) to bio-based solvents.

Scenario	Score
Regarded as readily obtainable from biomass.	10
Available with 100% bio-carbon content but only produced on a small scale.	8
An alternative manufacturing process using biomass is being developed at pilot plant stage or is feasible based on existing academic research.	6
The feedstock of an existing manufacturing process could be directly replaced with an equivalent bio-based version.	4
The solvent is chlorinated but otherwise could be bio-derived based on its carbon and hydrogen content.	2
Not obtainable by any reasonable and known process beginning with biomass.	0

Debate surrounding the assignments in such a classification is inevitable. The assessment, because it is subjective, is difficult to justify beyond its role here as only a rough guide. For example, it is hard to appreciate the potential of an academic study to become a commercial enterprise. Chemical companies may readily produce press releases stating the imminent production of a bio-based solvent or a key platform molecule, but until manufacturing is live and sustained at a high enough volume to impact the market it is not possible to know the exact status and probable longevity of the process. Some decisions are easier, such as the assignment of bio-ethanol at the top of the scale and hexane at the bottom.

Tailoring solvent selection guides to compliment renewability: The GSK solvent selection guide combines physical property and toxicity data for solvents and establishes a useful hierarchy which can be used to supplement the SUS classification discussed previously. The six categories of the GSK solvent selection guide is probably too many for the chemist with a casual interest in the subject, but the Pfizer solvent selection guide is too vague with its poorly defined classes of greenness [Alfonsi 2008, Henderson 2011]. It is possible to combine the health, flammability and reactivity categories of the GSK solvent selection guide into a single health and safety classification system (denoted HAS, acting as an acronym for health and safety). The health score is given an enhanced weighting in the calculation, which although is subjective, reflects the importance of human toxicity:

Equation 1.11

$$\mathrm{HAS} = \frac{(4 \cdot Health) + Flammability + Reactivity}{6}$$

Similarly the waste, environmental impact and LCA categories of the GSK solvent selection guide can become an environmental classification system (denoted ECO, as an inference to the word ecological) by taking the average of the existing category scores, but weighted in favour of the arguably more important LCA category:

Equation 1.12

$$\mathrm{ECO} = \frac{Waste + Environmental\ impact + (4 \cdot LCA)}{6}$$

The resulting HAS and ECO classifications are scaled to enhance the spread of data. Without this treatment, an identical process to how the **π*** scale in Equation 1.5 is normalised, the resulting values of the HAS and ECO classifications tend to cluster together. With the scaling, equal numbers of solvents can be distributed across the complete range of possible scores (*i.e.* one to ten), and then a useful means of classification is provided.

A criticism of the SUS-HAS-ECO framework is because it relies on the published GSK data set, it has no predictive element. A total of 110 solvents make up the complete GSK solvent selection guide, but any additional bio-based solvents like limonene cannot be fully assessed. A small selection of these solvents have been reinterpreted with the SUS-HAS-ECO classifications in the following table (Table 1.6), and the rest tabulated in the appendix (Table 8.3). Although not applicable to the solvents presented in Table 1.6, some entries in the GSK solvent selection guide do not have any LCA data. In these instances the ECO classification is constructed from the mean average of the waste and environmental impact categories. This is noted where appropriate throughout this work. Nevertheless this interpretation appears to present sensible conclusions for the vast majority of solvents in the data set. Indeed the weightings of the categories in the HAS and ECO classifications (and the scaling that followed) were chosen because the appearance of the results they provided seemed the most sensible.

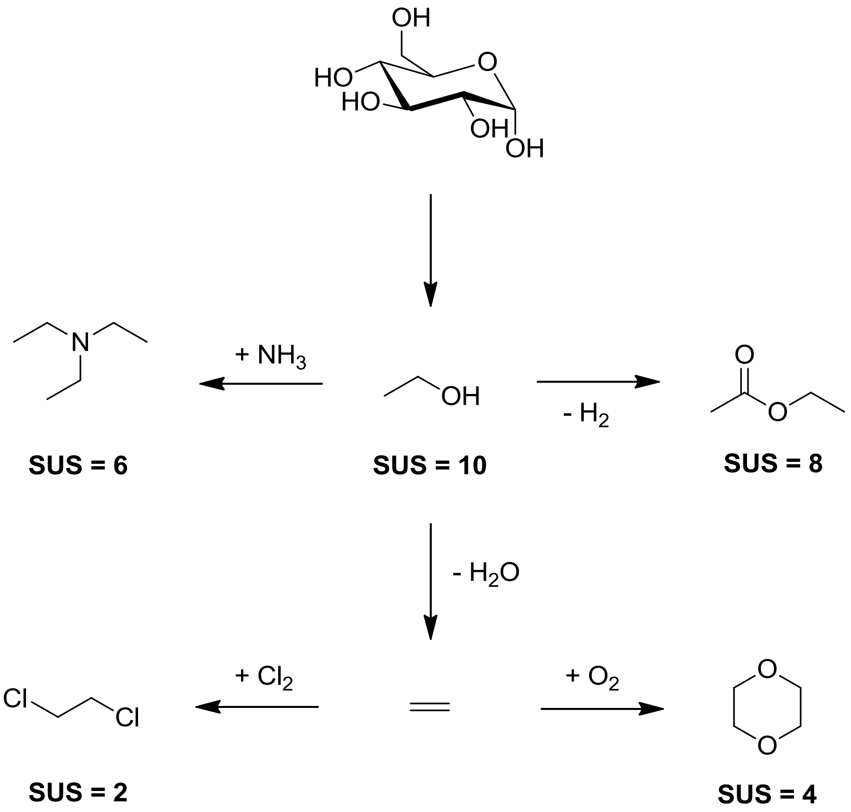
Table 1.6 A bio-based solvent selection guide.

Solvent	**SUS**	**HAS**	**ECO**
Acetone	8	9	9
Acetonitrile	4	7	1
1-Butanol	8	6	3
t-Butanol	4	7	10
Chloroform	2	1	3
1,2-DCE	2	1	6
DMSO	4	7	4
1,4-Dioxane	4	1	1
Ethanol	10	10	10
Ethyl acetate	8	9	6
Ethylene glycol	8	10	10
Hexane	0	1	6
2-MeTHF	10	1	1
NMP	6	2	1
Toluene	0	3	6
Triethylamine	6	1	6

It should be explained how each of the solvents in Table 1.6 came to receive their allocated SUS classification. Acetone was once made on an appreciable scale by fermentation, a technology that has only recently come back into consideration [Cathay Biotech 2013, Qureshi 2001, Rhodia 2013]. 1-Butanol is a co-product of this fermentation process. This results in a SUS classification of 8 assigned to acetone and 1-butanol. Acetonitrile is made from propene as a by-product of acrylonitrile production [McConvey 2012]. Several strategies with the purpose of producing bio-based propene are being explored by different organisations [Global Bioenergies

2012a, Hayashi 2013, Iwamoto 2013, Mizuno 2012, UOP 2013]. Accordingly a SUS classification of 4 is appropriate for any upstream products of propene, including acetonitrile. Remarkably isobutene has some history as a fermentation product, and a manufacturing process has been patented [Global Bioenergies 2012b, van Leeuwen 2012, Marlière 2010, Marlière 2011a, Marlière 2011b]. Hydration of bio-based isobutene would give *t*-butanol, hence the SUS classification of 4 [Weissermel 1997 page 69].

Chloroform could be obtained by replacing the natural gas used in the established chlorination process with methane from anaerobic digestion (biogas) [Lowenheim 1975 page 266]. Hence it has a SUS classification of 2, but it does not appear that anyone has been motivated to make chloroform in this way. The same can be said of 1,2-DCE which is made by reacting ethylene with chlorine (Scheme 1.3) [Lowenheim 1975 page 392]. A slightly better use of bio-based ethylene might be the synthesis of ethylene glycol *via* ethylene oxide [Lowenheim 1975 page 688, India Glycols 2013]. Although this intermediate presents toxicity issues, the manufacture of bio-based poly(ethylene terephthalate) (PET) plastic relies on this procedure [Kriegel 2010, Toray 2011]. 1,4-Dioxane is another solvent that is synthesised from ethylene oxide [Weissermel 1997 page 155]. Ethylene is now produced with the ethanol resulting from the fermentation of sugarcane [Fan 2013, Morschbacker 2009]. As such a SUS classification of 8 is



Scheme 1.3 Some bio-based solvents accessible by the fermentation of glucose.

appropriate for ethylene glycol, but hypothetical bio-1,4-dioxane is awarded a 4 instead. Bio-ethanol, being widely produced as a fuel attains the highest SUS classification, as does 2-methyltetrahydrofuran (2-MeTHF) because it is exclusively manufactured by the hydrogenation of furfural [Aycock 2007, Balat 2009, Pace 2012]. Ethyl acetate is available as the product of a ruthenium catalysed bio-ethanol dehydrogenation, although this is by no means a widespread practice [Ashley 2006, Colley 2004]. Another procedure that reacts ammonia with bio-ethanol to give triethylamine (mirroring the standard industry process) has been patented but not actually implemented [Gerlach 2006]. That is why triethylamine is assigned a SUS classification of 6 but bio-based ethyl acetate, which not only has been patented but then put into practice, is awarded an 8.

Both DMSO (SUS = 4) and NMP (SUS = 6) production utilises methanol which can be made from the syngas resulting from the steam reforming of biogas [Fukui 2002, Khadzhiev 2008, Lowenheim 1975 page 524, Weissermel 1997 page 102]. *N*-Methyl pyrrolidinone receives a higher SUS classification then DMSO because a procedure for making bio-based NMP has been developed, if perhaps of limited utility [Lammens 2010, Lammens 2011]. The fact that from a human health perspective DMSO is the superior of the two highly dipolar aprotic solvents, and thus a better target for bio-based product development, is accounted for by their respective HAS and ECO classifications.

Hexane and toluene are not considered to be bio-based solvents in this work. This is despite efforts to commercialise techniques converting carbohydrates to liquid aromatics using zeolite catalysts [Carlson 2009, Foster 2012, Huber 2005, Huber 2013]. If such a manufacturing process is in fact successful and economically viable, the product stream will probably be used entirely in fuel and platform molecule applications [Anellotech 2013]. The demand for aromatic solvents is decreasing in the face of legislative efforts aimed at discouraging their use [SubsPort 2013]. This is because of toxicity not feedstock reasons, and so aromatic bio-based solvents would appear to be redundant. So for reasons similar to those used to justify ostracising all chlorinated solvents, hydrocarbon solvents not naturally occurring in sources of biomass (or the direct products of related compounds) will not be considered as bio-based. This highlights that perhaps the distinctions made by the SUS classification will be short lived, and within a generation it is conceivable that all products of the oil refinery will also be obtained from bio-refineries on large scales [Saunders 2007]. Then the question will escalate to become one of environmental sustainability, which includes land use, water management, the preservation of bio-diversity, and many other factors beyond the scope of this work.

Repopulating the diagram: For the bio-based product economy to be successful and sustain our current standard of living, the roles of all petroleum derived chemicals will need to be accounted for with the products of renewable feedstocks. Using the aprotic (Figure 1.11) and protic (Figure 1.12) polarity maps developed earlier we are able to reveal the strengths and weaknesses of the bio-based solvent catalogue as it presently stands. Gaps in their collective polarity coverage will need to be resolved, aiding targeted bio-based solvent development for where demand is greatest [Jessop 2011]. Only solvents possessing a SUS classification of 6 or higher are regarded as being bio-based in this treatment. As such only the aprotic solvents acetone, ethyl acetate, 2-MeTHF, NMP and triethylamine from Table 1.6 are accounted for (Figure 1.20). Because the SUS classification is dependant only on academic articles, patents, and press releases it can be extended to the neoteric and sometimes obscure bio-based solvents such as the following diterpenes: limonene, α-pinene and cineole, all of which are present on Figure 1.20. Also included is *p*-cymene, which is synthesised from limonene and occasionally used as a solvent [D'hooghe 2008, Fessard 2007, Kelly 1997, Liu 2010, Marchais-Oberwinkler 2011, Ritter 2004]. To make a second aromatic bio-based solvent, a patented procedure for dehydrating the isobutanol formed through a fermentation process to give isobutene is followed by a dehydrocyclisation to give *p*-xylene [Gevo 2013, Peters 2011]. Finally γ-valerolactone, the product of levulinic acid hydrocyclisation is included as an ester solvent [Alonso 2013].

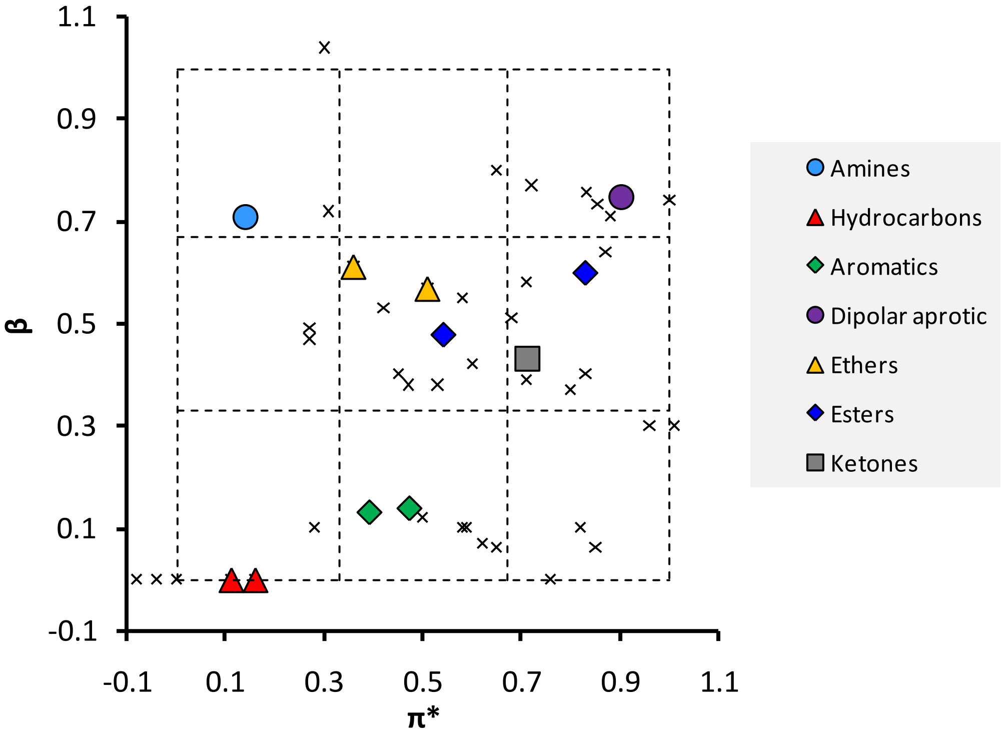


Figure 1.20 Aprotic bio-based solvent polarity map.

There are a total of eleven bio-based solvents on Figure 1.20. This leaves quite considerable gaps throughout the polarity map. The number of aprotic bio-based solvents would grow quite considerably if biogas and bio-based ethylene were used in conventional solvent manufacturing practices in place of their non-renewable analogues. More examples of highly dipolar aprotic solvents could be made from biogas, including DMF [Weissermel 1997 page 43]. Chlorination of biogas (bio-based methane) to DCM and chloroform would also be possible [Lowenheim 1975 page 266]. The modestly hydrogen bond accepting belt of the solvent polarity map (*i.e.* the middle row of Figure 1.20) could be enhanced by the presence of diethyl ether amongst others if ethylene was renewably sourced for the purpose of its synthesis [Lowenheim 1975 page 345]. Still this would not resolve the poor health and safety aspects of these solvents, which is why there is scope for introducing greener but unconventional solvents if they are produced from biomass. Sometimes there is no other reasonable option, like with the introduction of limonene and α-pinene as bio-based solvents in place of traditional hydrocarbon solvents. The difference in boiling point between *n*-hexane and limonene for example, because it is such an important property of the solvent, will limit the number of applications where limonene would actually be considered as a replacement for *n*-hexane, despite their similar polarity. So unfortunately in addition to the gaps in the polarity map of Figure 1.20, there are false positives where at first it appears that solvent substitution options are accounted for. Of course this depends on the nature of the application the solvent is being sought for; sometimes limonene is an excellent substitute for petroleum derived hydrocarbon solvents [Clark 2012, Veillet 2010]. This is why it is so important to combine a knowledge of solvent properties and greenness within the context of an application, as the solvent selection algorithm has been revised to do.

Because many fermentation products are alcoholic, the protic bio-based solvent polarity map is more densely populated than its aprotic equivalent (Figure 1.21). The ionic liquids of the original diagram did not qualify as bio-based, obviously nor does water. Although in the special case of water it would be acceptable to include it with the bio-based solvents if desired. Aside from the halogenated alcohols all the other solvents from Figure 1.12 are present in Figure 1.21 with the exception of *t*-butanol. Because of the number of bio-based alcohols available the impact of this loss should be minimal. Actually the number of bio-based solvents available greatly exceeds those indicted on Figure 1.20 and Figure 2.11, especially when blends for cleaning applications are taken into consideration [Datta 2005, Datta 2008]. But unless their Kamlet-Taft polarity parameters are available, bio-based solvents cannot be treated within the framework of this sort of polarity map assessment.

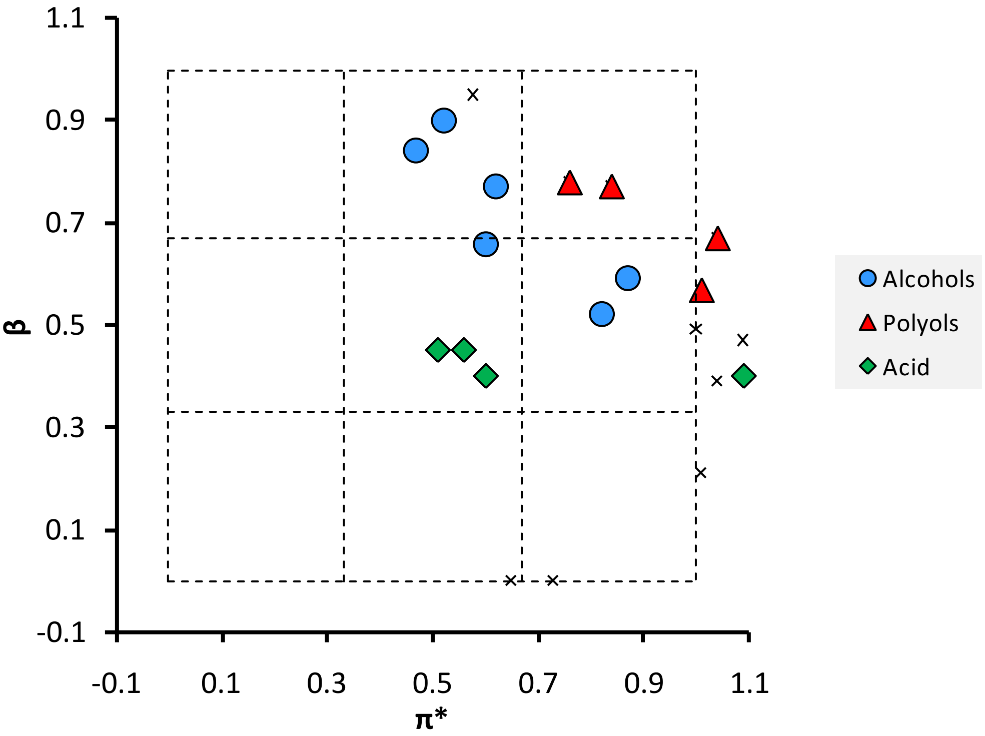


Figure 1.21 Protic bio-based solvent polarity map.

An earlier analysis of solvent polarity maps conducted by Jessop correctly identifies supercritical carbon dioxide as a useful addition to the bio-based solvents [Jessop 2011]. Carbon dioxide is obtained from various industrial waste-streams, including as a by-product of the brewing industry [Hunt 2010]. In its correct place on the aprotic polarity map, carbon dioxide would reside within the vicinity of the alkane solvents [Marcus 2005]. More interestingly, carbon dioxide can be mixed with alcoholic co-solvents to create either tuneable supercritical media for use in chromatographic applications or gas expanded liquids as solvents for synthesis [Jessop 2007, Rajendran 2012]. As expected, gas expanded liquids have an intermediary polarity between that of both components. The combination of carbon dioxide and methanol produces a medium with a **π*** value of just 0.37 [Abbott 2009]. Unfortunately **β** values were not available, although Jessop estimates it to be as low as 0.40 for carbon dioxide-methanol mixtures at 50 bar [Jessop 2011]. Using gas expanded liquids to explore the low dipolarity/polarisability region of Figure 1.21 (*i.e*. the left-most column) increases the polarity range described by protic solvent systems, whilst not relying on any non-renewable solvents.

Case study selection: Comparing reaction classes helped assess which reactions to target for solvent optimisation based purely on their frequency of use (Figure 1.2). Equally important are current trends in solvent selection, and this will dictate whether it is worth seeking an alternative solvent for each particular transformation. Each of the reactions surveyed will employ a solvent, except in a few instances where liquid reactants are used neat. This may change in the future when emerging technologies assist the application of solventless reactions (*e.g*. microwave and ball mill reactions) [Cave 2001, James 2012, Varma 1999]. The following exercise matches the popular (and seemingly essential) reactions discussed earlier with the regular solvent choice (Table 1.7). Solvent assignments were made primarily by using data describing the synthesis of the top 200 selling pharmaceutical products [Kleemann 2001, Mack 2009]. Only the most frequently used solvents are presented in Table 1.7. This assessment strengthens concerns over current solvent use with DCM and toluene frequently used to make esters and amides, and DMF and other highly dipolar aprotic solvents applied in alkylations. These solvents are known to be toxic and unsustainable, which makes these reactions ideal as case studies in order to push forward the art of solvent selection and the design and application of bio-based solvents.

Table 1.7 Solvent selection practices in reactions common to the pharmaceutical industry.

Transformation	Rank[a]	Popular solvents
N-Substitution	1-4	DMF and DMSO [Kleemann 2001]. Also acetone, acetonitrile, nitromethane and sulpholane [Reichardt 2003].
N-Acylation to amide	2-1	DCM and toluene [Kleemann 2001].
O-Substitution	3-8	DMF and THF [Kleemann 2001]
Heterocycle synthesis	4-3	Acetic acid, DMF, ethanol and toluene [Kleemann 2001]
Cross-coupling	5-9	DMF [Kleemann 2001]. Toluene, aqueous 1,4-dioxane and other aqueous ether or alcohol mixtures [Scifinder 2013a].
Alcohol to halide	6-n/a	DCM and DMF [Kleemann 2001].
Reduction to amine	7-6	Ethanol and methanol [Kleemann 2001].
O-Acylation to ester	n/a-2	DCM, THF, toluene and pyridine [Kleemann 2001].

[a] Transformation rankings displayed in the following format: process development-manufacturing scale.

Reduction to give an amine can be ruled out as a candidate for further study. Despite its popularity there is little point in trying to replace lower alcohols with unfamiliar solvents given that bio-ethanol is an established renewable product. All the other reactions in Table 1.7 rely on the use of undesirable solvents to some extent. However there must be good reason for these solvents being used in spite of their human toxicity and environmental impact. To appreciate this fully, solvent properties and how they influence organic reactions must be understood before a solvent substitution is attempted. The remaining reaction classes listed in Table 1.7 are all viable case studies, especially the much practiced *N*-substitution and amidation reactions. These two transformations are addressed in this work, as are esterifications (listed as '*O*-acylation to ester' in Table 1.7) and a heterocycle synthesis (specifically to make dihydropyrimidinones).

The synthesis of haloalkanes was not selected as a case study because in many cases the functional group interconversion is of an intermediary nature [Constable 2007b]. It would be better to devote research efforts towards the telescoping of reactions to avoid unnecessary halogenation [Risatti 2013]. The reason why *O*-substitution was omitted is because of its similarity to the more prevalent *N*-substitution. Originally it was imagined that an example of a cross-coupling reaction would follow the nucleophilic substitution case study because they traditionally rely on similarly dipolar aprotic solvents, and so studying nucleophilic substitution before cross coupling reflects a logical increase in complexity. Any bio-based solvent developed for one of these two applications would almost certainly be useful in the other, and strengthen the case for introducing a novel bio-based solvent if one could be found. But the results of the nucleophilic substitution case study, and the fact that (for the Suzuki reaction at least) aqueous alcohols have gained favour as a solvent system, meant this option became less attractive [Maegawa 2007].

Solvent effect screening methodology: What it means for a solvent to be bio-based, and the limits to how much we can elaborate this definition has been touched upon in the preceding pages. The reaction case studies in most need of a substitution in favour of a bio-based solvent have also been chosen. Now the experimental foundation for the assignment of solvent effects to help establish green yet bio-based replacement solvents must be equally understandable and reliable. This work will examine the benefits and flaws of the solvatokinetic method used independently by Wells and Schleicher [Schleicher 2009, Wells 2008]. This approach will be supplemented with the solvent selection algorithms and related assessments introduced previously. Most of the experiments described subsequently will have been conducted within the framework of this methodology, and the final conclusions of this thesis will be as much about the success of this approach in identifying high performance bio-based solvents as it will be about the solvents themselves.

The premise of the methodology is that a reaction can be conducted in a variety of solvents (under otherwise identical conditions) and the kinetics monitored by ^{1}H-NMR spectroscopy with an accuracy sufficient to then determine the underlying solvent effect with a LSER. As previously alluded to, the Kamlet-Taft solvatochromic parameters of solvent polarity are usually reliable for this purpose, and can be measured easily if required. If one of the reactants has an easily identifiable ^{1}H-NMR signal, and the corresponding moiety is observed in the product (albeit at a different chemical shift to avoid signal overlap) then the progress of the reaction can be derived from the relative intensities of the signal integrals. Conversion to the product at a given time (C_t) is assigned with the following equation, where $[A]_0$ is the initial concentration of the yield limiting reactant, $[B]_0$ is the initial concentration of a reactant with a distinctive spectrum signal, I_B is the signal integral of B, I_P is the signal integral of the product, H_B is the number of hydrogens responsible for reactant signal, and H_P is the number of hydrogens responsible for the product signal:

Equation 1.13
$$C_t = \frac{[\mathbf{B}]_0}{[\mathbf{A}]_0} \cdot \left[\frac{(I_P/H_P)}{(I_B/H_B)+(I_P/H_P)}\right]$$

If it is to be the yield limiting reactant that is monitored, then the equation simplifies to the following expression:

Equation 1.14
$$C_t = \left[\frac{(I_P/H_P)}{(I_B/H_B)+(I_P/H_P)}\right]$$

Calculated conversions at given times can be processed with an integrated rate equation to give the rate constant (**k**) for the reaction. Integrating the second order rate equation for example gives the following equation where $[P]_t$ is the concentration of product P at time t and the reactants are designated as A and B [Logan 1996]:

Equation 1.15
$$\mathbf{k}_2 t = \frac{1}{[\mathbf{A}]_0-[\mathbf{B}]_0} \ln\left(\frac{[\mathbf{B}]_0([\mathbf{A}]_0-[\mathbf{P}]_t)}{[\mathbf{A}]_0([\mathbf{B}]_0-[\mathbf{P}]_t)}\right)$$

As will be shown during the course of this work, no normalisation of ^{1}H-NMR signal intensities is necessary. This is an advantage over chromatographic methods of analysis. The ^{1}H-NMR spectrum is produced within ten minutes on a 400 MHz spectrometer which is faster than all but the most rapid HPLC systems. With either method each aliquot of the reaction mixture taken for analysis must be quenched. It was found that for reactions requiring elevated temperature, aliquots could simply be diluted in a suitable deuterated solvent at room temperature and not show any significant additional conversion. Analytical chromatography becomes more useful than NMR spectroscopy when the reaction components are very complex, resulting in no clear ^{1}H-NMR signals to follow as the reaction progresses. Similarly when the reactants and products are very alike it may be difficult to attribute signals in a ^{1}H-NMR spectrum to the correct component. The best choice of analysis will always be influenced by the type of chemistry of course, but ^{1}H-NMR spectroscopy will be regarded as the default option. If this analysis fails alternatives are available. Some reactions have been successfully monitored by UV-vis. spectroscopy and the kinetics interpreted with a LSER [Ranieri 2008]. Naturally this approach requires a chromophore to be present, which usually is the case with pharmaceutical products. Both UV-vis. and NMR spectroscopic methods also permit *in situ* monitoring of a dynamic system.

Rate constants need to be obtained in enough solvents to give statistically significant data. Typically at least six solvents are used in each of the following case studies, followed by verification in further solvents after the optimal solvent properties have been identified and then intepretated with the revised solvent selection algorithm and its associated assessments. In order to minimise error all experiments were repeated to ensure reproducibility, covering conversions in excess of at least 50% of the theoretical maximum where possible. The temperatures of the reactions were calibrated with internal temperature measurements and not from the settings on the external heater.

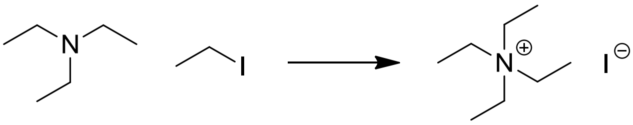
Solvent effects can be evaluated without much difficulty once the LSER coefficients have been calculated. The 'Data Analysis' tool in the Microsoft Excel spreadsheet software proves very useful for this job. Statistical significance of each Kamlet-Taft solvent polarity parameter is ascertained using p-values [Wells 2008]. A p-value below 0.01 is indicative of a given Kamlet-Taft parameter being responsible for the observed solvent effect. Of course other solvent polarity scales can be introduced when they provide a more fitting description of the system being studied. Once the LSER is completed to a satisfactory degree, the selection of possible solvent substitutes can commence. Extrapolation of an LSER trend should give a strong clue towards identifying the solvent properties required for maximal performance. Use of the solvent selection algorithms and associated solvent selection guides reviewed earlier will ensure any solvent candidates fulfil the requirements of a benign yet suitable solvent. If a high performance bio-based solvent candidate is identified, and experimentation confirms this prediction, then a comprehensive comparison between the traditional solvent option and the recommended bio-based substitute is provided in the following case studies. A wider discussion ensures that the broader implications of replacing a solvent with another, that might not otherwise be apparent from just conducting the reaction, are fully appreciated.

2. Nucleophilic substitution

Nucleophilic substitution, the introduction of vital functionality in place of sacrificial electron rich chemical groups, is currently an essential, if atom uneconomical transformation employed to build up complexity in drug candidate molecules. True to this, heteroatom alkylation is the most practiced class of transformation in the pharmaceutical industry [Carey 2006, Dugger 2005]. It seems that nucleophilic substitution was the most popular reaction system to study in the infancy of physical organic solvent chemistry. Reactions of both unimolecular (S_N1) and bimolecular (S_N2) mechanisms have been investigated in detail [Abraham 1985]. It is the more prevalent S_N2 mechanistic pathway that will be examined in this chapter, using a LSER to assess solvent performance and solvent selection tools to implement bio-based solvents.

2.1 Solvents and nucleophilic substitution

The dawn of solvent effects research: As long ago as 1890 Nikolai Menschutkin realised that a solvent and any reaction occurring within that solvent are inseparable [Menschutkin 1887, Menschutkin 1890a, Menschutkin 1890b, Menschutkin 1900, Reichardt 2003 page 2]. In concluding this, Menschutkin revealed much about the role of the solvent in the reaction between triethylamine and iodoethane to give the respective quaternary ammonium halide salt (Scheme 2.1). He attributed the variable rate constants of this reaction to the chemical properties and not the physical properties of the solvent [Reichardt 2003 page 3 and page 147]. Usually the solvent of a reaction will not undergo any chemical change, and so in the century that followed after Menschukin's discovery, attempts were made in the broader field of physical organic chemistry to understand how the physical properties of solvents could be related to observed changes in reaction rates or equilibrium positions. The inception of the LSER was the pinnacle if



Scheme 2.1 Menschutkin's original reaction.

not the culmination of this effort, which will be put to use here in combination with other complementary tools to suggest solvents with improved EHS profiles for nucleophilic substitution reactions.

Michael Abraham, before his collaborations with Kamlet and Taft, revisited the Menschutkin reaction with trimethylamine and both methyl iodide or *p*-nitrobenzyl chloride [Abraham 1969]. Although no LSER was constructed at this time, the rate constants for the reactions in a variety of solvents were reliably ascertained (Figure 2.1). The rate constants were adjusted against methanol as a reference. The pronounced variation in reaction rate, first recognised by Menschutkin, is obvious. It appears that highly dipolar solvents provide the best medium for accelerating the rate of reaction. Current practice usually employs solvents like DMF, DMSO, acetonitrile, and nitromethane in S_N2 alkylations, all solvents with large $\boldsymbol{\pi^*}$ values. [Kleeman 2001, Reichardt 2003 page 489].

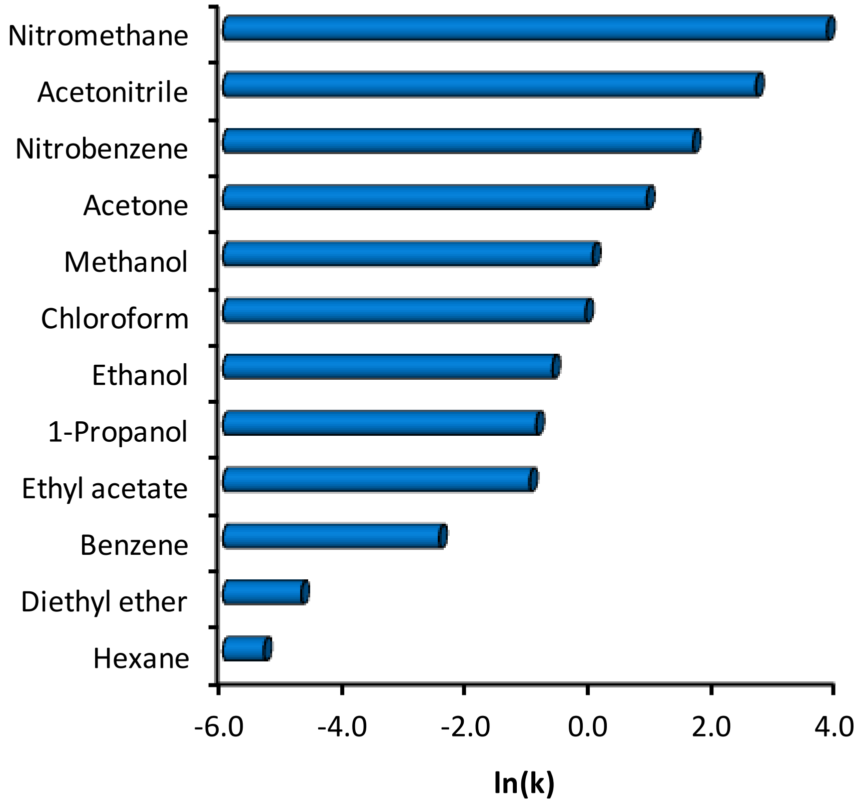
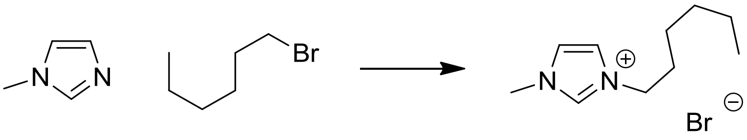


Figure 2.1 The natural logarithms of rate constants (relative to methanol) for the Menschutkin reaction.

Nucleophilic substitution and linear solvation energy relationships: To this day research is still dedicated to nucleophilic substitution solvent effects under the guise of topical reactions such as the synthesis of ionic liquid precursors. Schleicher observed that the non-specific polarity of the solvent is largely responsible for the observed rate of alkylation of imidazole derivatives [Schleicher 2009]. The rate of this S_N2 reaction is also inversely proportional to $\boldsymbol{\alpha}$. The $\boldsymbol{\beta}$ term claimed by Schleicher as significant is not actually so, as implied by the small magnitude of its coefficient (*b*) and verified with its insignificant p-value when recalculated here (Table 2.1). The work of Abraham on the rate of trialkylamine quarternisation also shows no dependence on $\boldsymbol{\beta}$, and although initially describing an LSER with the expected proportionality between $\ln(\mathbf{k})$ and $\boldsymbol{\pi}^*$ and inverse proportionality with respect to $\boldsymbol{\alpha}$, the latter term was later removed by the author on grounds of statistical insignificance [Abraham 1985]. It was not explained why solvent hydrogen bond donation was not relevant to the rate of this transformation, but is influential in near-identical reactions. Examples of both types of alkylation (of tertiary amines and imidazoles) with benzyl and alkyl halides are pesented in Table 2.1, indicating that the electrophile does not profoundly influence the solvent effect, only the magnitude of the coefficients obtained. A

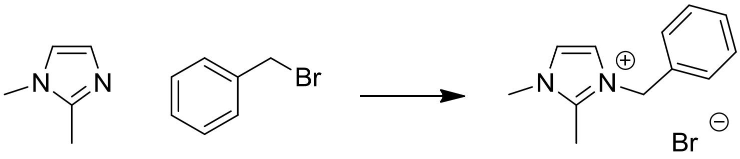
Table 2.1 A collection of LSER analyses for Menschutkin type reactions and the solvolysis of *t*-butyl chloride as reported in the original publications.

1. Alkylation of 1-methylimidazole [Schleicher 2009].



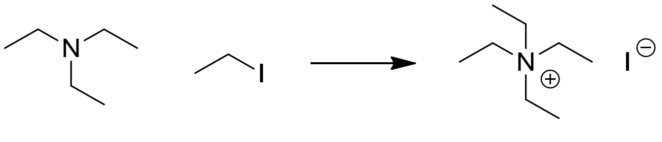
$\ln(\mathbf{k}) = -14.72 - 2.07\boldsymbol{\alpha} + 0.07\boldsymbol{\beta} + 4.99(\boldsymbol{\pi}^* - 0.20\boldsymbol{\delta})$ ($R^2 = 0.95$)

2. Alkylation of 1,2-dimethylimidazole [Skrzypczak 2004].



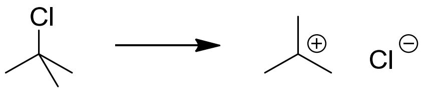
$\log(\mathbf{k}) = -4.95 - 1.22\boldsymbol{\alpha} + 2.61\boldsymbol{\pi}^*$ ($R^2 = 0.92$)

3. Alkylation of triethylamine [Abraham 1985].



$-\Delta\Delta G^{\ddagger} = -6.04 + 6.98\boldsymbol{\pi}^*$ ($R^2 = 0.980$)

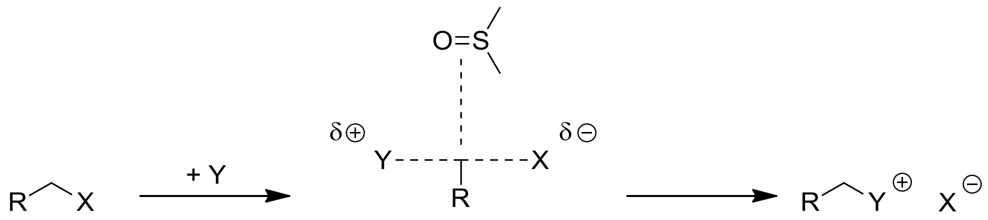
4. Solvolysis [Abraham 1985].



$-\Delta\Delta G^{\ddagger} = -8.36 + 6.87\boldsymbol{\alpha} + 8.76\boldsymbol{\pi}^* + 0.0007\boldsymbol{\delta}_{\mathrm{T}}^2$ ($R^2 = 0.996$)

comparison to the S_N1 solvolysis of *t*-butyl chloride reveals a distinct difference between the two mechanisms. The common use of alcoholic solvents in S_N1 heteroatom alkylations is indicative of the beneficial influence of hydrogen bond donation on the reaction [Kleemann 2001]. However the dipolarity of the solvent is equally vital as it is in determining the rate of S_N2 type reactions.

As discussed previously the role of LSER analysis is to deduce mechanism. In combination with other techniques, some computational, the linear activated complex of the S_N2 mechanism has become well understood [Abraham 1975, Castejon 1999]. Our current understanding of the Menschutkin reaction would suggest that an imidazole-solvent hydrogen bond reduces the nucleophilicity of the imidazole (or indeed another type of nucleophile). But most importantly the activated complex shows the beginnings of charge separation as the salt product is created, and it is this species that benefits most from the stabilising influence of solvent dipolarity (Scheme 2.2).

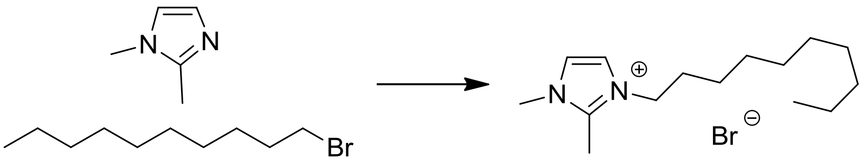


Scheme 2.2 The general mechanism of the Menschutkin reaction in which the activated complex is stabilised through an interaction with the solvent DMSO.

Generally no recommendations for an ideal solvent have been suggested in the case studies preceding this work . The goal of previous research was always the understanding of mechanism [Solà 1991]. Usually even yield maximisation was forsaken. The greatest achievements in the physical organic chemistry of solvent effects were conducted long before the principles of green chemistry were established [Anastas 1998]. To justify adding to this body of work, any contribution to the understanding of the mechanism will be greatly supplemented with an attempt to design solvents to enhance the productivity of the reaction. This approach also incorporates the solvent selection tools proposed in the Chapter 1. The advent of green chemistry has reinvigorated many of the chemical science disciplines, none more than solvent research.

2.2 Nucleophilic substitution results and discussion

Model reaction: The reaction between 1-bromodecane and 1,2-dimethylimidazole at 323 K was chosen as a model S_N2 nucleophilic addition (Scheme 2.3), which is analogous to the system studied by Schleicher [Schleicher 2009]. Such a reaction lends itself to this study because the rate of imidazole alkylation is within a measurable range at practical temperatures, making it a good choice of substrate. The use of a long chain haloalkane to give 1-decyl-2,3-dimethylimidazolium bromide is an attempt to impart wide ranging solubility to the salt product across a variety of solvents. Unfortunately the salt was not soluble in low polarity solvents such as arenes and alkanes and a biphasic system resulted as the reaction progressed. In other instances the reaction can be followed by ^{1}H-NMR spectroscopy.



Scheme 2.3 The model Menschutkin reaction between 1,2-dimethylimidazole and 1-bromodecane.

The initial reaction mixture consisting of 1,2-dimethylimidazole and a slight excess of 1-bromodecane in solution was stirred at 323 K, with aliquots removed from the mixture at regular intervals for ^{1}H-NMR spectroscopic analysis. Conversion to the product can be calculated using the CH_2X signal integrals of what begins as the 1-bromodecane reactant. Because 1-bromodecane is used in excess this must be accounted for when calculating conversions (Equation 1.13). The CH_3 signals of the methylimidazole group can also be used, or even the proton signals from the imidazole moiety itself. The variation of these NMR signal intensities can be seen in the results of a trial reaction in DMSO (Figure 2.2). Figure 2.2 shows ^{1}H-NMR spectra at four points during the course of a reaction, with the relative intensities of the signals associated with the product increasing with time relative to those representing the reactant.

The conditions of the reaction needed to be such that a large variation in rate constants could be obtained in different solvents within a reasonable time span. This depends on the concentration of the reaction and the temperature. Graphs of conversions (Figure 2.3) and integrated rate equations (Figure 2.4) as a function of time in DMSO and ethanol are drawn subsequently. It is obvious that the reaction proceeds with a much greater velocity in DMSO (a

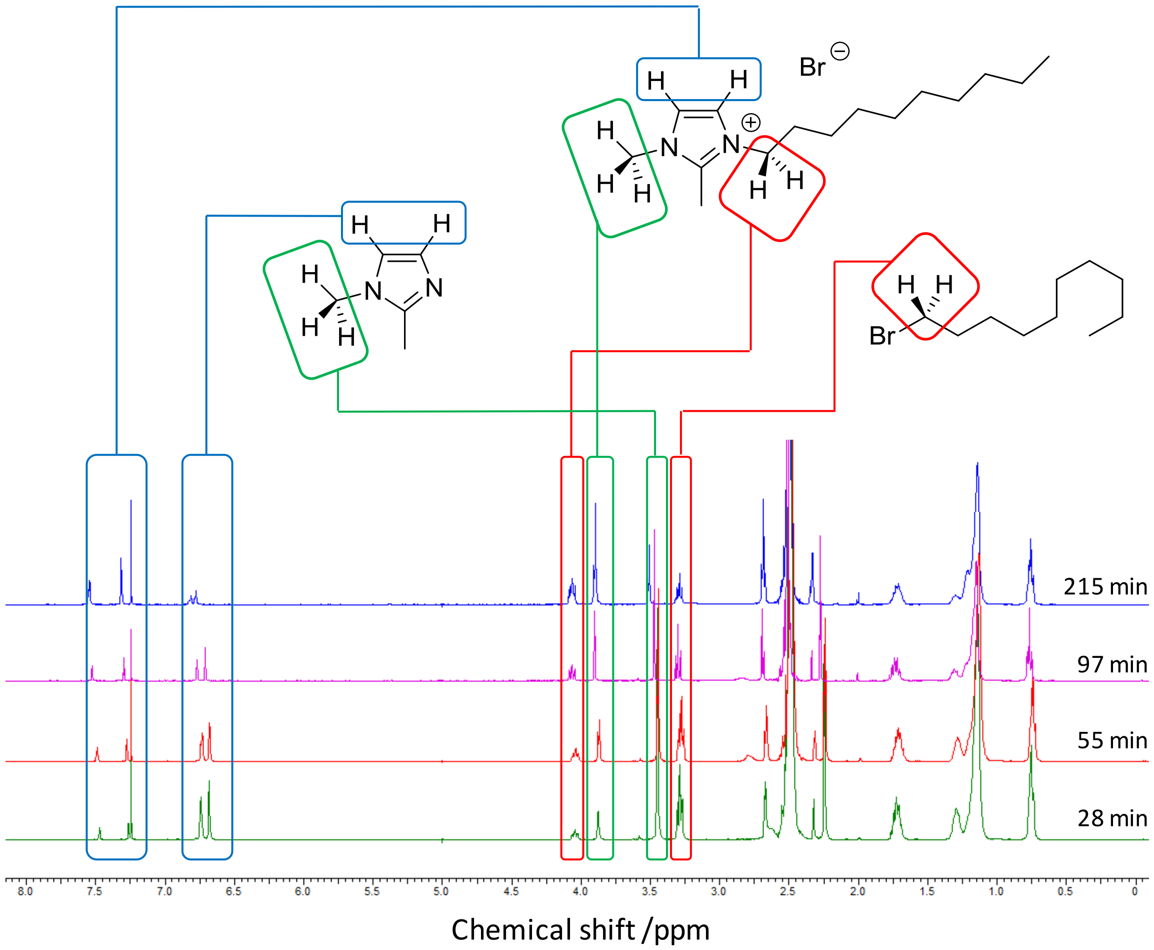


Figure 2.2 The progression of the model Menschutkin reaction in DMSO as observed by ^{1}H-NMR spectroscopy.

dipolar solvent) than ethanol (a protic solvent). It takes 2.5 hours for conversions in DMSO at 323 K to exceed 50%. In ethanolic solution more than 2.5 days are required. The extremes of this case study are indeed large enough to indicate a profound solvent effect, but manageable at the same time. It was not possible to greatly improve the rate of reaction in ethanol without making the reaction in DMSO impossible to accurately analyse. The integrated second order rate equation is equivalent to the rate constant multiplied by time (Equation 1.15), and so differentiation of each linear plot in Figure 2.4 provides the rate constant of the reaction.

Empirical parameters to describe the Menschutkin reaction solvent effects: The rate of the reaction needed to be studied in a greater variety of solvents in order to ascertain any solvent effect. Further solvents were selected to cover a range of polarities in order to help obtain strong correlations. Toluene and similar solvents have already been ruled out on the basis of their failure to dissolve the product. This is not always an issue for kinetic studies but the formation of a second, more polar, liquid phase meant that it could not be guaranteed that the reactants resided entirely in the intended solvent phase. Care was taken to avoid any inadvertent

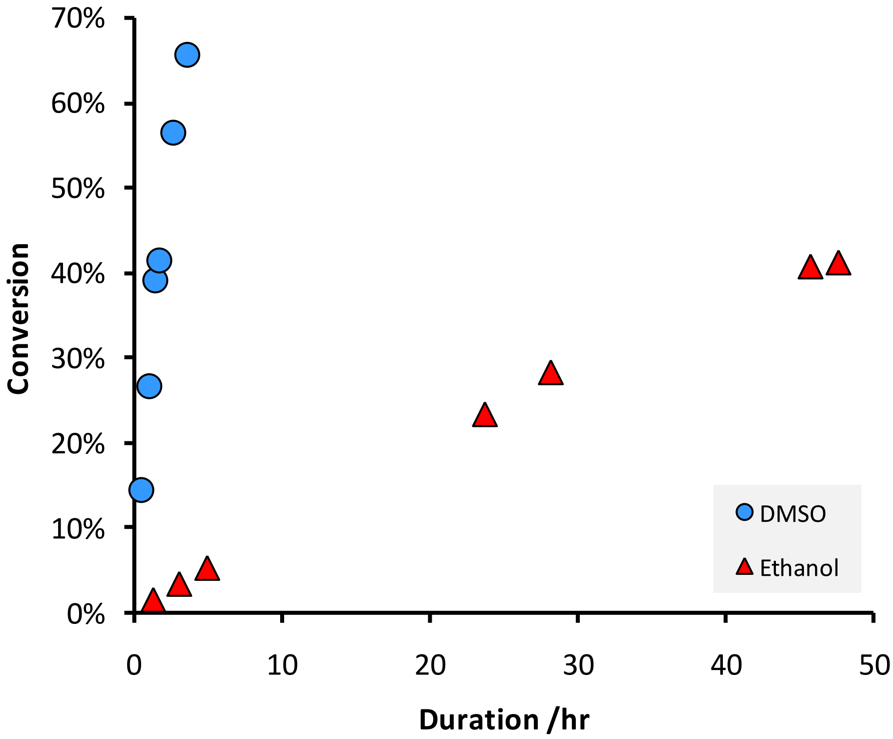


Figure 2.3 The conversion to 1-decyl-2,3-dimethylimidazolium bromide with the Menschutkin reaction in DMSO and ethanol as determined by ^{1}H-NMR spectroscopy.

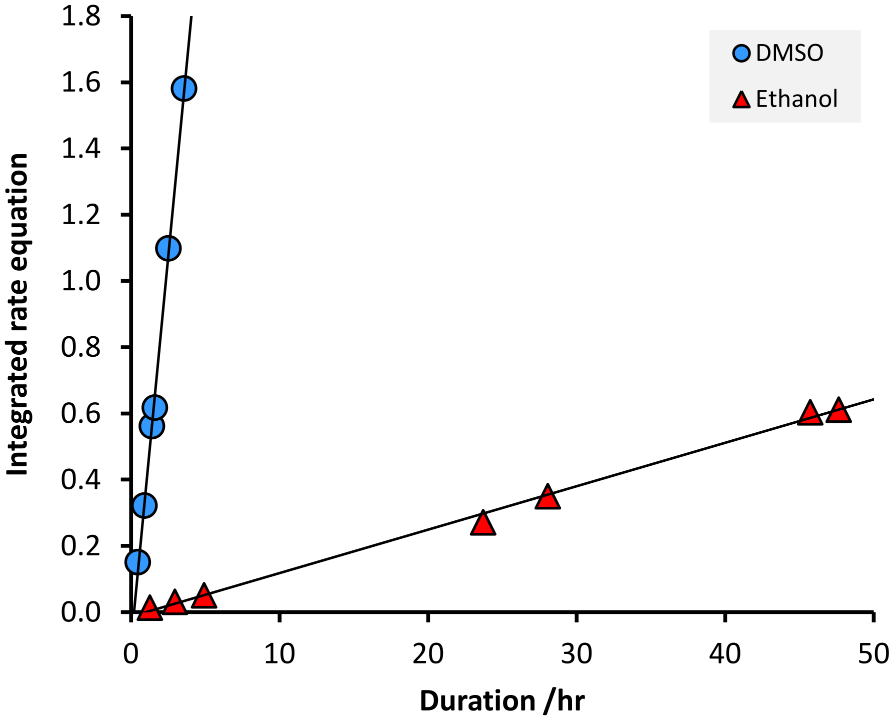


Figure 2.4 Integrated rate equations of Menschutkin reactions in DMSO and ethanol.

proportionality between the polarity parameters of the solvents in the data set in order to aid associating any solvent effect with only the responsible solvent-solute interactions. A bias towards highly dipolar aprotic solvents was adopted given that this class of solvent is the most popular for this transformation. A quick test to ensure an unbiased solvent set consists of plotting a graph of the Kamlet-Taft solvatochromic polarity parameters to visually assess whether a suitable distribution of polarity occurs (Figure 2.5). This can be simplified by just plotting **β** against **π*** as with the solvents polarity maps already featured in Chapter 1. The hydrogen bond donating ability of solvents is indicated visually.

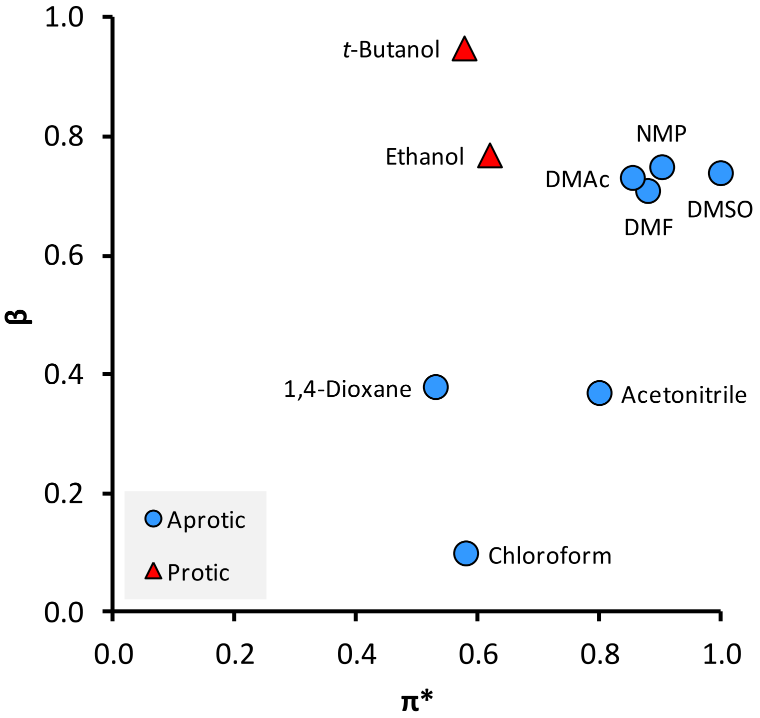


Figure 2.5 The polarity of the model Menschutkin reaction solvent set.

The greatest rate constant is obtained with DMSO as the solvent, closely followed by NMP and then the other highly dipolar aprotic solvents (Table 2.2). Ethanol and *t*-butanol can be written off as potential solvents for S_N2 reactions given the slow reaction kinetics. In fact only the most polar solvents can be considered as performing adequately given the long established precedent of solvents like DMSO in this reaction.

Aside from the alcoholic solvents that impart very slow reaction kinetics, DMSO has the best EHS profile in the solvent set (Figure 2.6). However this is not to say that DMSO is the ideal solvent for this class of reaction. Dimethyl sulphoxide is not currently made from a renewable feedstock. Economic and supply issues still prevent this from happening. Naturally occurring

Table 2.2 Solvent polarity measurements and kinetics of the model Menschutkin reaction.

Solvent	ln(**k**)	**α**	**β**	**π***	$\mathbf{E_T^N}$[a]	**ε_r**[a]
Acetonitrile	-9.87	0.35	0.37	0.80	0.460	35.94
t-Butanol	-11.88	0.39	0.95	0.58	0.389	12.47
Chloroform	-10.93	0.20	0.10	0.58	0.259	4.89
DMAc	-9.52	0.00	0.73	0.85	0.377	37.78
DMF	-9.58	0.00	0.71	0.88	0.386	36.71
DMSO	-9.01	0.00	0.74	1.00	0.444	46.45
1,4-Dioxane	-11.26	0.00	0.38	0.52	0.164	2.21
Ethanol	-12.53	0.83	0.77	0.62	0.654	24.55
NMP	-9.19	0.00	0.75	0.90	0.355	32.2

[a] Quoted as reported [Reichardt 2003 page 418 and page 472].

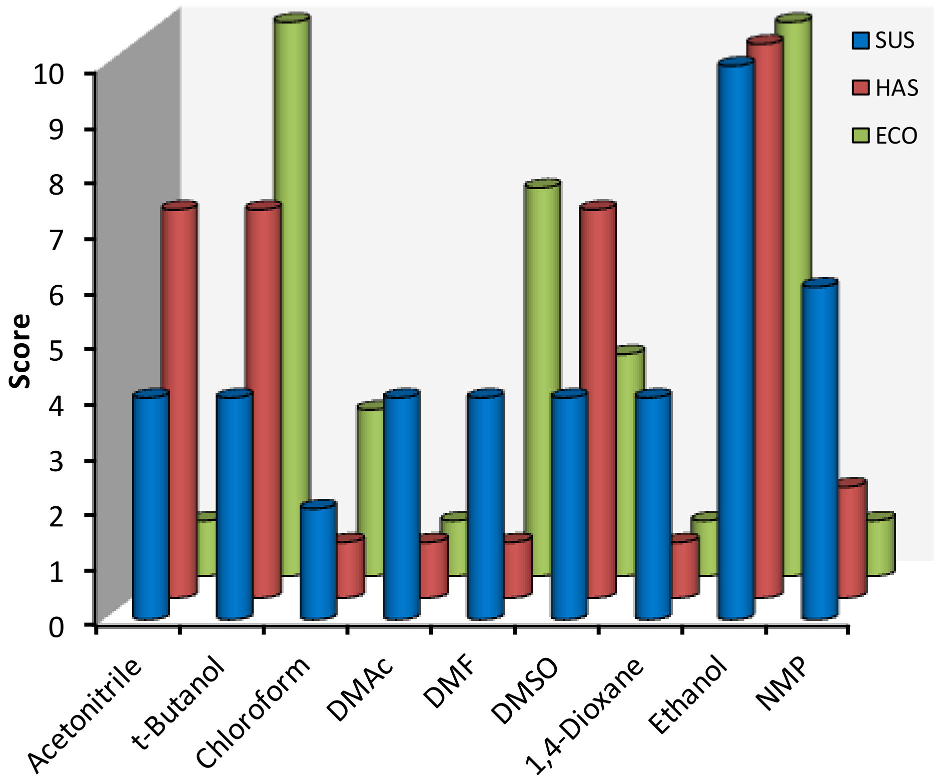


Figure 2.6 The SUS-HAS-ECO classifications of the model Menschutkin reaction solvent set.

dimethyl sulphide could be oxidised to DMSO also [Charlson 1987, Hussain 2012]. An issue that prevails over possible sustainability concerns and not taken into account in the health and safety (HAS) classification is the ability of DMSO to draw chemicals through the skin barrier of humans [Tilstam 2012]. In addition, the environmental profile of DMSO is less than satisfactory with an

ECO classification of four. This low score arises from problems caused by the presence of DMSO in waste streams. Being water miscible, incineration and recycling of DMSO are negatively impacted. It is also reactive at high temperatures and has a relatively low autoignition point [MSDS 2013]. As such, a replacement dipolar aprotic solvent is still a priority.

A correlation was developed describing the natural logarithm of the observed rate constants as a function of solvent polarity. As expected ln(**k**) is proportional to solvent polarity (**π***) for non-alcoholic solvents at least. Ethanol and *t*-butanol provide a less satisfactory medium for the reaction than predicted by the correlation with **π*** (Figure 2.7). The aprotic solvents appear to be closely correlated. A non-general LSER for aprotic solvents only can be represented by the following equation:

Equation 2.1 $$\ln(\mathbf{k}) = -13.75 + 4.85\boldsymbol{\pi}^{*} \qquad (R^2 = 0.986)$$

The discrepancy caused by alcoholic solvents can be corrected by introducing **α** into the LSER but this in turn displaces the C-H acids acetonitrile and chloroform from the new LSER (Figure 2.8):

Equation 2.2 $$\ln(\mathbf{k}) = -13.81 - 2.16\boldsymbol{\alpha} + 4.92\boldsymbol{\pi}^{*} \qquad (R^2 = 0.994)$$

As with the previous correlation, when Equation 2.2 is applied to those solvents for which it is valid, the resulting relationship is reliable. Note that **β** is not significant in this instance.

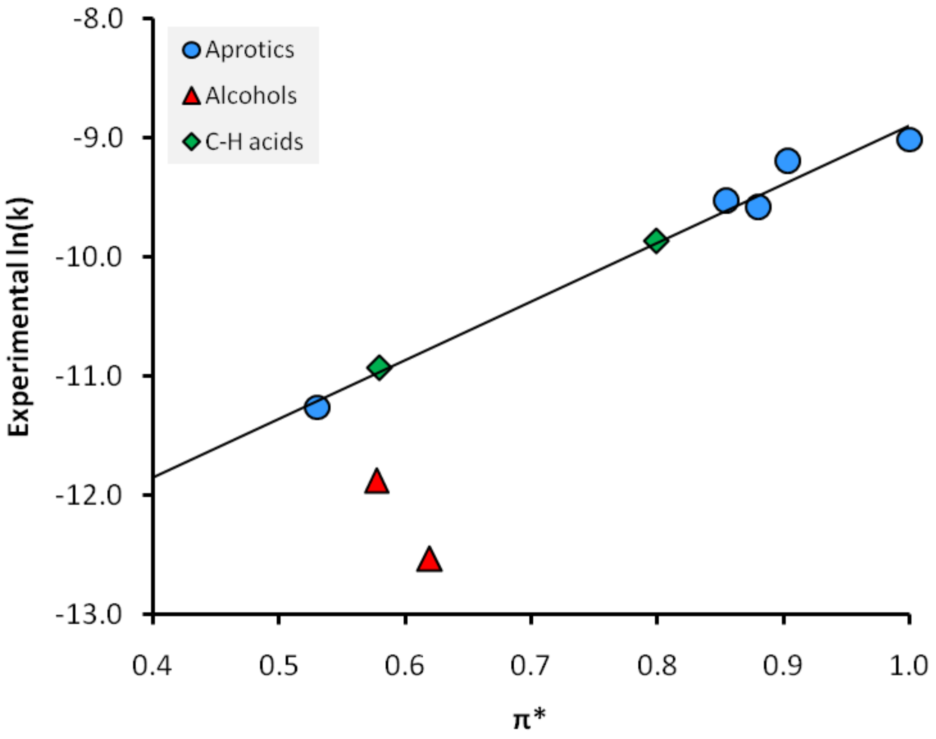


Figure 2.7 The LSER describing the rate constant of the model Menschutkin reaction as a function of solvent dipolarity only.

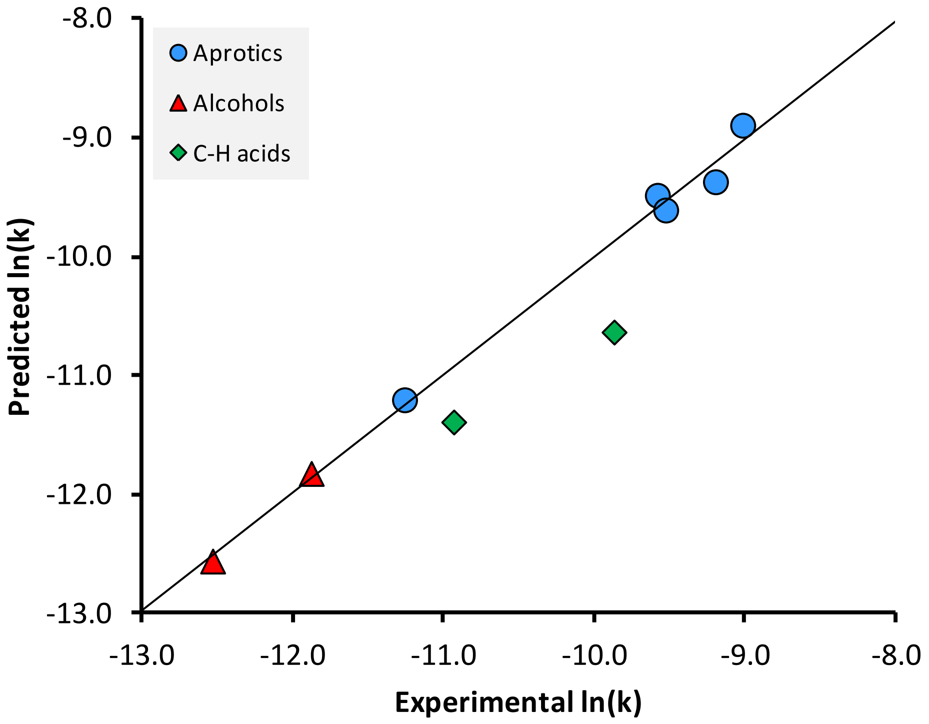
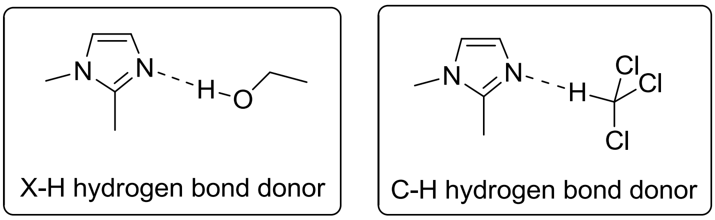


Figure 2.8 A comparison between experimental and predicted S_N2 ln(**k**) values based on a LSER incorporating both **π*** and **α**.

It appears necessary to consider the **α** term as separate contributions by X-H hydrogen bond acidity and the weaker C-H hydrogen bond acidity (Scheme 2.4). The former is applicable in modelling the kinetics of this reaction while the latter is not statistically significant. It can be imagined that the hydrogen bonding interaction between ethanol and an imidazole for example would be quite strong and influential in determining the kinetics of any subsequent reaction. Presented with the evidence it is not unconceivable that a similar interaction in a solution of chloroform might not occur at all, given that the electronic resonance of imidazoles will make hydrogen bonding less appealing then it would be with an amine.



Scheme 2.4 The typical X-H hydrogen bond donor ethanol (left) compared to a C-H hydrogen bond donor chloroform (right) as they interact with 1,2-dimethylimidazole.

Dividing a solvent polarity parameter into component parts is accepted for the **π*** term in cases where dipolarity and polarisability exert different strengths of interaction. A polarisability correction term (**δ**) is separated from **π*** so that the modified '*s*(**π*** + *d***δ**)' term now accounts for dipolarity and polarisability in a ratio suitable for the application that is being described [Taft 1981, Taft 1983]. The non-specific polarity of aromatic solvents is strongly dictated by polarisability, with **δ** equal to unity. Polychlorinated solvents are designated **δ** values half that of the aromatics. Other solvents are considered to be without polarisability. With few exceptions, the **π*** values of these unpolarisable solvents are linearly related to their dipole moments [Abboud 1977, Taft 1983]. Using the same procedure, C-H hydrogen bond acidity (designated as **ε**) can be separated from **α** so that a '*a*(**α** + *e***ε**)' term now accounts for the differences in the modes of hydrogen bond acidity. As with *d*, the *e* coefficient should always be of the opposite sign (+ or -) to its parent solvent polarity term, be it **α** or **π*** (Table 2.3). Note that **δ** is not used in this case study due to a lack of variation in the solvent set.

Table 2.3 An explanation of the polarisability term coefficient *d*.

Magnitude	Effect
$d > 1$	Not possible to account for more polarisability than is there.
$1 > d > 0$	Not possible for permanent and induced dipoles to have an opposing influences.
$d = 0$	Dipolarity and polarisability have an equivalent influence.
$0 > d > -1$	Polarisability has a diminished effect compared to dipolarity.
$d = -1$	Polarisability is statistically insignificant (dipolarity is not).
$d < -1$	Not possible to subtract more than the original amount.

With the introduction of **ε** the LSER can now account for all solvents in the initial set. This is the first reported observation of a noticeable difference in the mode of action by C-H and X-H hydrogen bond donors in a relationship formulated with the Kamlet-Taft solvent polarity parameters. The value of *e* is close to the negative of unity (*i.e.* complete removal of C-H acid contribution by **ε** from **α**), which goes some way to justify its introduction. As expected, the parameter with the largest coefficient is **π*** and there is a pronounced and negative **α** coefficient:

Equation 2.3
$$\ln(\mathbf{k}) = -13.79 - 2.17(\boldsymbol{\alpha} - 1.02\boldsymbol{\varepsilon}) + 4.90\boldsymbol{\pi}^{*} \quad (R^2 = 0.995)$$

The strength and hence utility of the correlation is most readily confirmed with R^2 values of data variance. Furthermore p-values were used to discriminate between statistically significant parameters and those which played no part in dictating the kinetics of the reaction. Error in the predictive capacity of the LSER was much less than 1% of the experimental ln(**k**) values in the solvent set (Figure 2.9). The related system studied by Schleicher cannot be confirmed as expressing a similar solvent effect in which **ε** is a meaningful parameter due to data scatter masking this subtle effect. The linearity of the trend in Figure 2.9 indicates that the rate determining step and the mechanism of the reaction is unchanged as the solvent is varied [Sykes 1981].

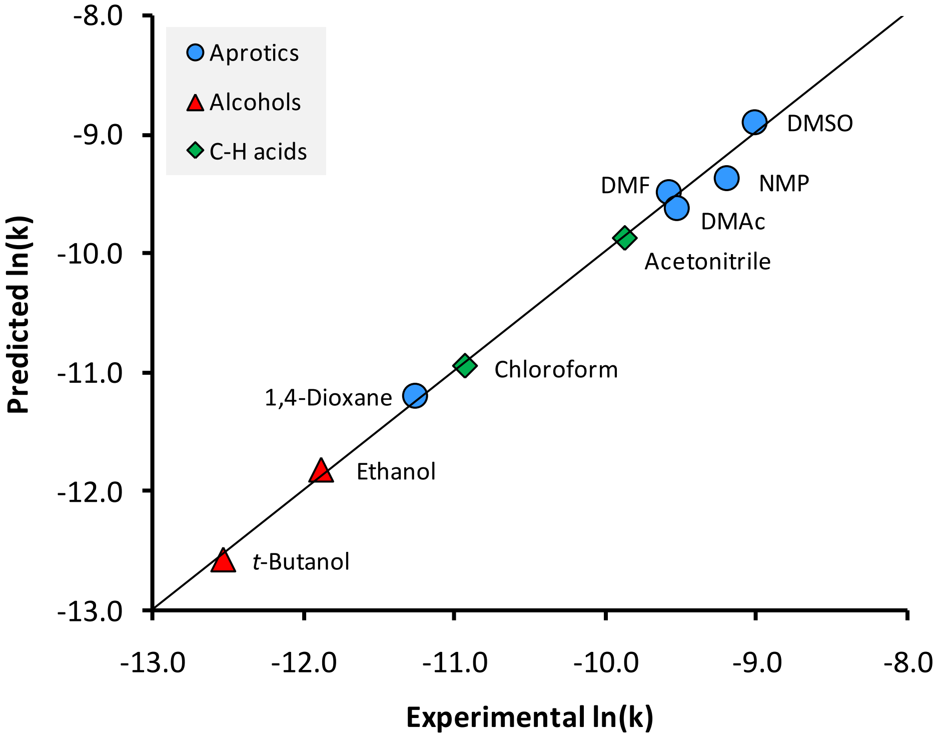


Figure 2.9 Predicted ln(**k**) values from the LSER featuring the **ε** modification to **α** compared to experimental Menschutkin reaction data (Equation 2.3).

Finally any systematic errors were sought by plotting the error in predictions against experimental ln(**k**) values. A correlation would indicate a systematic error. It does seem that the solvents producing faster kinetics have more error associated with the reaction rate measurements, but the error is not biased in one direction (Figure 2.10). This is likely to come from error introduced by removing an aliquot from the reaction mixture and diluting at room temperature. This was the only attempt to halt the progress of the reaction before analysis. Reassuringly the variation is not severe, as attested to by the data in Figure 2.9 in which little deviation from predicted performance is observed.

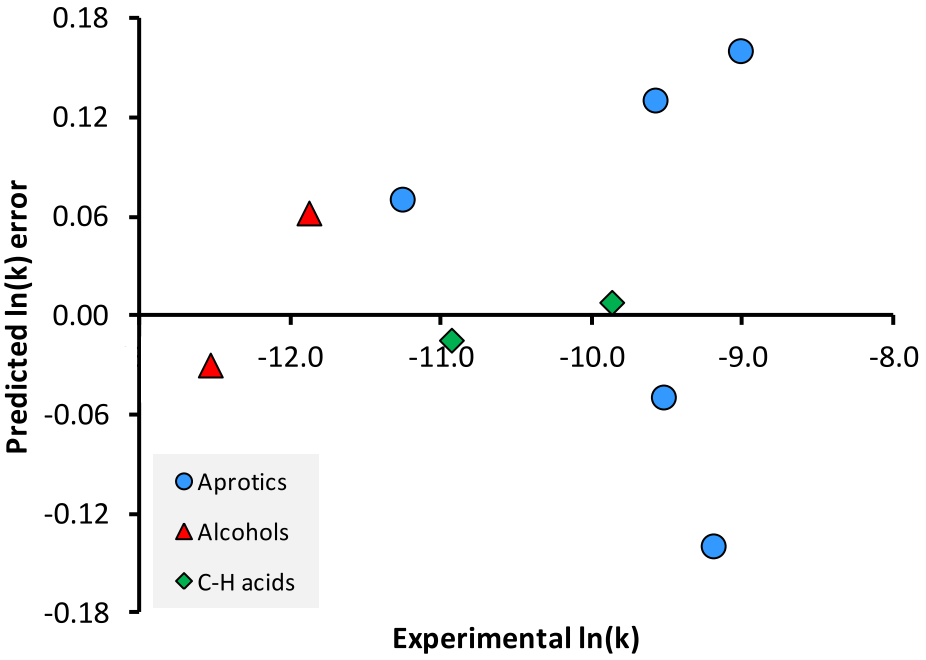


Figure 2.10 A systematic error check of experimental rate constants for the model Menschutkin reaction.

It has been shown that the Kamlet-Taft solvent polarity parameters, albeit with a modification to the term responsible for hydrogen bond donating ability, will successfully account the observed rates of reaction. The solvent polarity scale devised by Reichardt has also been used in this role, consisting of a single variable [Reichardt 2003 page 411]. There is an inherent proportionality between $\boldsymbol{\alpha}$ and $\boldsymbol{\pi^*}$ in the $\mathbf{E_T^N}$ scale, that unless it matches the system being studied, will introduce error. This error removes all traces of correlation between solvent polarity and the kinetics of the model reaction (Figure 2.11). Furthermore relative permittivity, once the favourite of the physical organic chemist, is shown to be no better at describing the system (Figure 2.12). Based on these assessments the Kamlet-Taft solvent polarity parameters will be used exclusively in further investigations.

Mechanistic insights: The refined LSER (Equation 2.3) meets the expectation that $\boldsymbol{\alpha}$ will generally impede the kinetic progress of the Menschutkin reaction, but $\boldsymbol{\pi^*}$ would chiefly determine it through stabilisation of the separated charge that is developed in the activated complex of the reaction. Strong solvent hydrogen bond donation will restrict the rate of reaction considerably regardless of how dipolar that solvent may be. In terms of free energy implications, the hindrance to the rate of reaction observed with X-H hydrogen bond donors most likely arises from a stabilising hydrogen bond interaction between 1,2-dimethylimidazole and the solvent that increases the energy gap between reactants and activated complex (Figure 2.13). The free energy

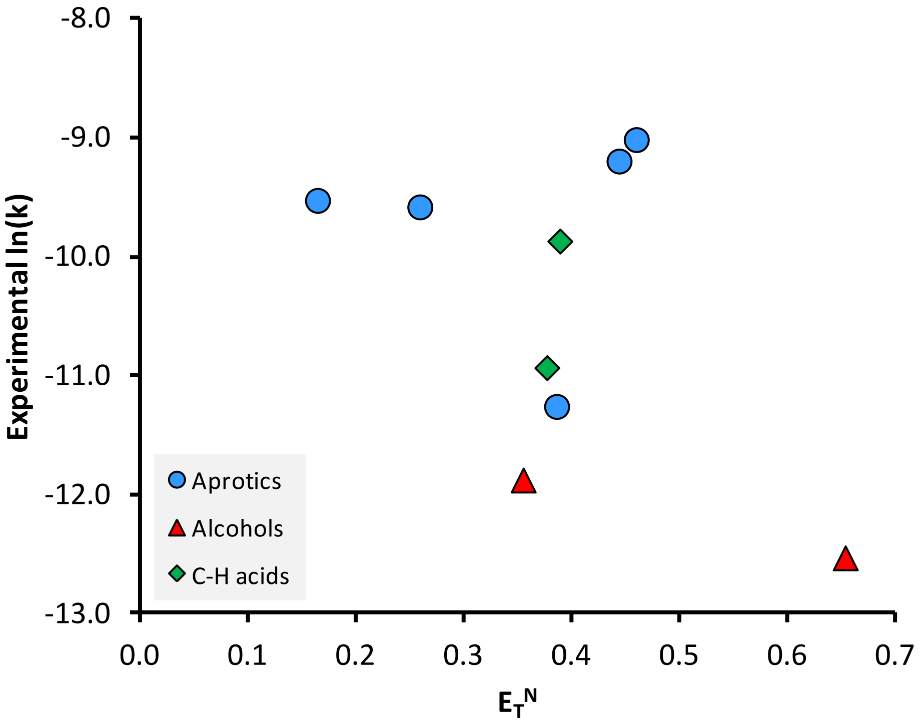


Figure 2.11 A demonstration of the absence of a correlation between Reichardt's $\mathbf{E_T^N}$ parameter and ln(**k**) of the model Menschutkin reaction.

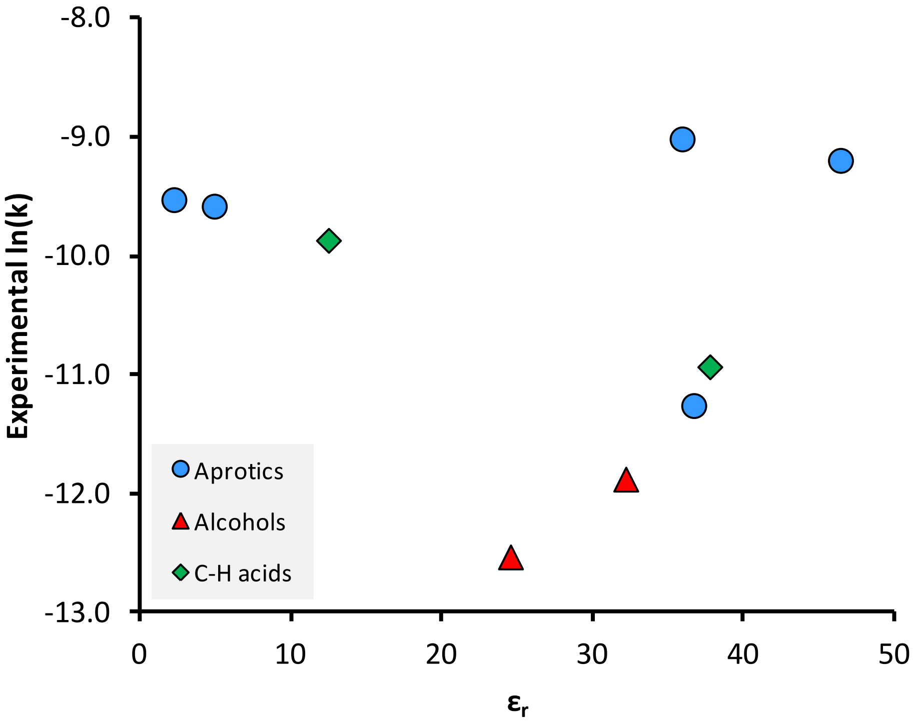


Figure 2.12 The relationship between relative permittivity and experimental values of ln(**k**) of the model Menschutkin reaction.

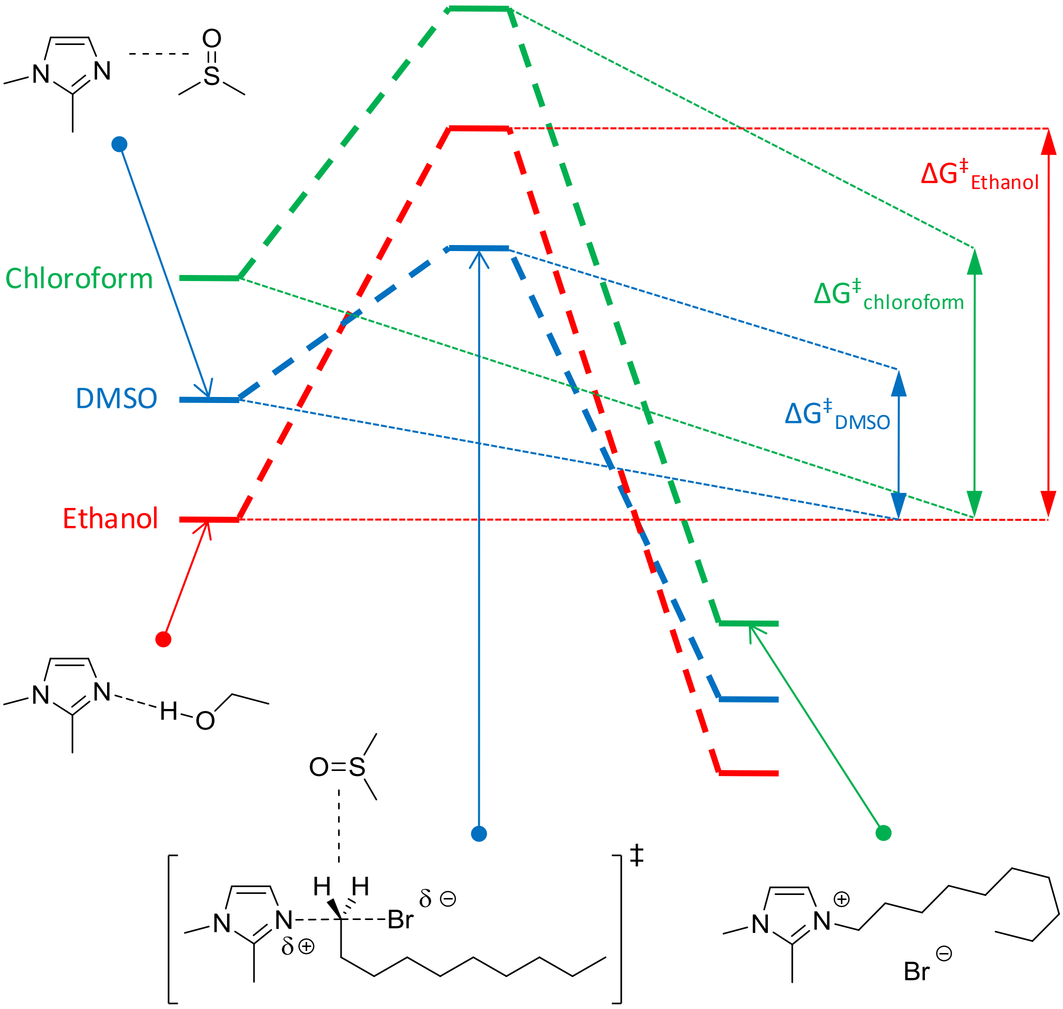


Figure 2.13 An relative free energy diagram indicating Menschutkin reaction solvent effects in DMSO relative to chloroform and ethanol.

diagram presented is a tool, and energy levels are not representative of actual chemical systems. It is the breaking of the hydrogen bond interaction between protic solvent and nucleophile that must occur to accommodate the introduction of the haloalkane in the activated complex, which leads to the longer reaction times reported earlier. Solvent dipolarity (**π***) will stabilise all reaction components, but the greater benefit is supplied to the most polarised species, in this case the activated complex. The activation energy of the reaction is increased upon the introduction of an alcoholic solvent but not so with C-H acidic solvents regardless of their non-zero **α** values. As this hydrogen bond interaction seems not to exist between 1,2-dimethylimidazole and C-H acids like acetonitrile and chloroform then the full benefit of their dipolarity can be exerted.

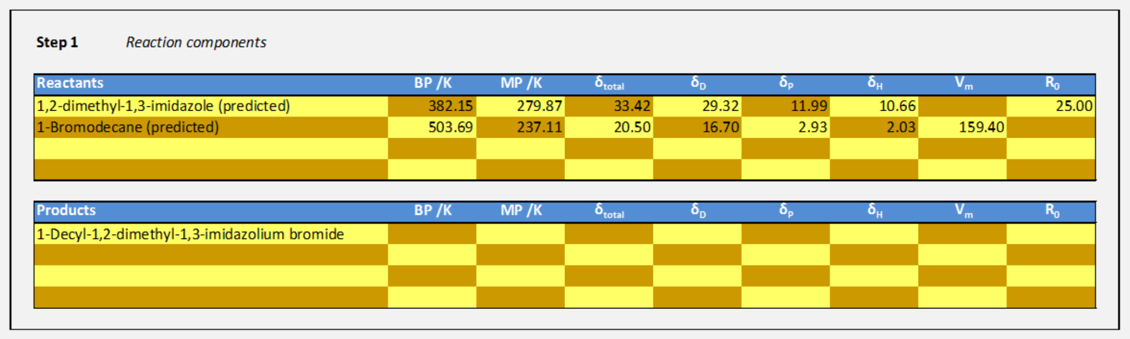
To tie these observations into existing studies, the solvent effect deduced here appears to be consistent with the enthalpy of activation and entropy of activation of the Menschutkin reaction in various mixtures of acetonitrile and methanol [Kondo 1984]. The enthalpy of activation is minimised in highly dipolar aprotic solvents. The increased ordering of the solvent around the activated complex entropically favours less dipolar solvents (or solvent mixtures for that matter) but this benefit is overwhelmed by the increase in enthalpy.

The magnitude of the LSER coefficients can sometimes aid the assignment of either an early or a late transition state. In Equation 2.3 the *s* coefficient that represents the importance of **π*** is neither extraordinarily small nor large for this type of correlation, and as such does not reveal the nature of the activated complex. Abraham suggested a dipole moment of 7.6 D exists in the activated complex of the reaction between methyl iodide and tripropylamine [Abraham 1975]. This corresponds to a late transition state and accordingly a sensitive solvent effect. Later Abraham claims to resolve the dispute over the nature of the activated complex and decided on an early transition state [Abraham 1981]. Nevertheless an appreciable dipole must be created during the course of the model reaction, even if the activated complex is 'reactant-like' as opposed to resembling the ionic product, resulting in the observed solvent effect.

Data entry for the solvent selection algorithm: Immediately, the LSER of the Menschutkin reaction strongly indicates that the use of dipolar aprotic solvents will enhance reaction kinetics. It would in theory be possible to select other dipolar solvents (aside from the usual basic amides and sulphur containing solvents) with the guidance of the polarity maps presented in the preceding chapter. However it is unlikely that, of the possible alternatives (a ketone or a nitrile perhaps), any will challenge the **π*** magnitude of DMF or DMSO. A notable exception is the exceedingly dipolar but unsustainable nitrobenzene.

Confirmation of bio-based solvent suitability can be obtained with the solvent selection algorithm. The basic outputs of the method will be examined here in some detail, this being the first case study, advancing from the principles set out in the general introduction. In later case studies the description will be truncated and only explored in greater depth where required. Because many bio-based solvents are novel reaction media their physical parameters are not necessarily known and will have to be estimated. A computational method for the estimation of the Kamlet-Taft solvatochromatic polarity scales has been developed, but without access to this, solvent polarities could not be approximated [Lamarche 2001, Platts 2000a, Platts 2000b].

The solvent selection algorithm (both the original and revised models) was set up to operate in a computer spreadsheet (Microsoft Excel). The initial input screen of the spreadsheet is dedicated to the properties of the solutes (Figure 2.14). Cells coloured either yellow or brown are available to the user to modify and enter data. Properties can be predicted using external software, or so called group contribution calculation methods if otherwise unavailable [Hansen 2007 page 6, Hukkerikar 2012]. The Hildebrand and Hansen solubility parameters are necessary as a means of assessing solubility, with cells provided for other optional properties. The product of the Menschutkin reaction is a salt and as such the prediction of physical properties was not possible using the group contribution methods of property estimation currently available. This has minor ramifications later on which should be accounted for, but within this solely kinetic analysis the properties of the product are not vital.



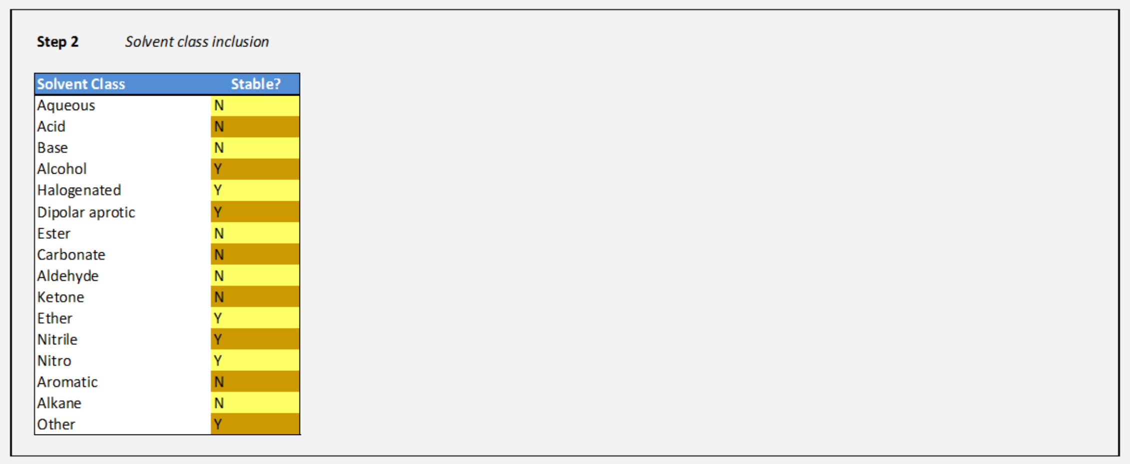
Step 1 *Reaction components*

Reactants	BP /K	MP /K	δ_{total}	δ_D	δ_P	δ_H	V_m	R_0
1,2-dimethyl-1,3-imidazole (predicted)	382.15	279.87	33.42	29.32	11.99	10.66		25.00
1-Bromodecane (predicted)	503.69	237.11	20.50	16.70	2.93	2.03	159.40	

Products	BP /K	MP /K	δ_{total}	δ_D	δ_P	δ_H	V_m	R_0
1-Decyl-1,2-dimethyl-1,3-imidazolium bromide								

Figure 2.14 Menschutkin reaction solvent selection algorithm screenshot, step 1: Reaction components.

The next stage refines the database of 195 solvents to include only those that have acceptable functional groups (Figure 2.15). This may have to be intuitive at the first pass, but after preliminary experiments have been conducted (as is the case here) this can be amended confidently and as necessary. For a nucleophilic substitution, acids and bases (*e.g.* amines) are not suitable solvents because competing side reactions are likely to occur. Esters, carbonates, aldehydes, and ketones were removed for the same reason. Hydrocarbon solvents were excluded



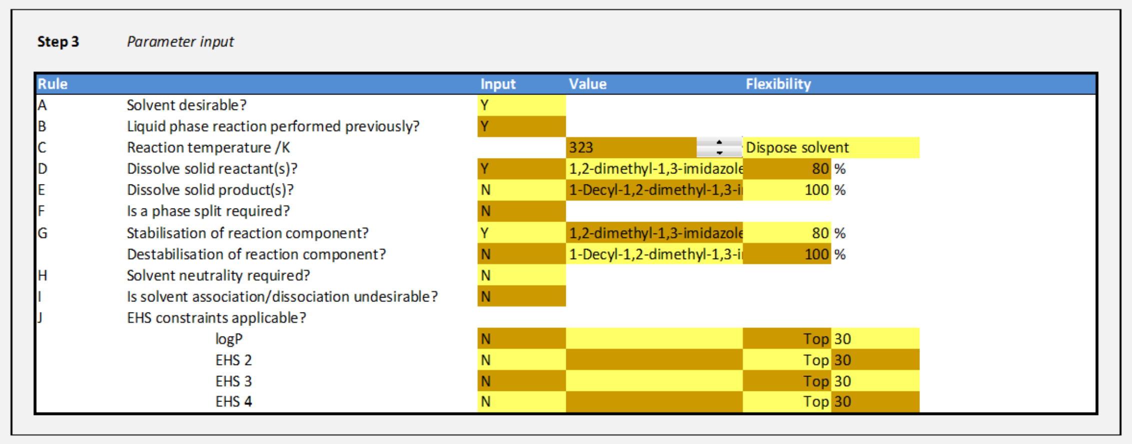
Step 2 *Solvent class inclusion*

Solvent Class	Stable?
Aqueous	N
Acid	N
Base	N
Alcohol	Y
Halogenated	Y
Dipolar aprotic	Y
Ester	N
Carbonate	N
Aldehyde	N
Ketone	N
Ether	Y
Nitrile	Y
Nitro	Y
Aromatic	N
Alkane	N
Other	Y

Figure 2.15 Menschutkin reaction solvent selection algorithm screenshot, step 2: Solvent class inclusion.

because they are unable to dissolve the reaction components, although they do not necessarily have to be removed here because the following polarity matching reaction indices will discount the low polarity solvents anyway.

Then the full set of reaction indices familiar to the original method of Gani are introduced under the heading of rules A to J [Gani 2005]. The parameter input section of the spreadsheet contains the rules of the original algorithm with the additions and alterations described in Chapter 1 (Figure 2.16). Although the data entry is the same from the perspective of the user, differences become apparent when the results are generated. The first two rules must be responded to with a 'Y' (yes) otherwise the rest of the assessment becomes redundant. The reaction temperature (rule C) is important to the algorithm because this is used to discount the majority of unsuitable solvents. The desired reaction temperature of 323 K was entered. An important distinction to remember between the original (model A) and revised (model B1) solvent selection algorithms is that in the revised algorithm no solvent that is a liquid at the reaction temperature will be discarded at this early stage (Figure 1.15).



Step 3 *Parameter input*

Rule		Input	Value	Flexibility	
A	Solvent desirable?	Y			
B	Liquid phase reaction performed previously?	Y			
C	Reaction temperature /K		323	Dispose solvent	
D	Dissolve solid reactant(s)?	Y	1,2-dimethyl-1,3-imidazole	80	%
E	Dissolve solid product(s)?	N	1-Decyl-1,2-dimethyl-1,3-i	100	%
F	Is a phase split required?	N			
G	Stabilisation of reaction component?	Y	1,2-dimethyl-1,3-imidazole	80	%
	Destabilisation of reaction component?	N	1-Decyl-1,2-dimethyl-1,3-i	100	%
H	Solvent neutrality required?	N			
I	Is solvent association/dissociation undesirable?	N			
J	EHS constraints applicable?				
	logP	N		Top	30
	EHS 2	N		Top	30
	EHS 3	N		Top	30
	EHS 4	N		Top	30

Figure 2.16 Menschutkin reaction solvent selection algorithm screenshot, step 3: Parameter input.

If certain solutes should be dissolved then this can also be decided in step 3 (Figure 2.16). The ability of a solvent to dissolve a solute is gauged by the closeness of their respective Hildebrand solubility parameters (rule D and rule E). Only rule D was applied here, as solubility of the product is not a requisite, but ultimately because the polarity parameters of the product are not known. The stabilisation or destabilisation of solutes is determined by the Hansen polarity and hydrogen bonding solubility parameters (rule G). This is a more sophisticated version of rule D and rule E. The flexibility in the solute stability assessment is user defined, and was set at ±80% of the polarity of 1,2-dimethylimidazole. In the original solvent selection algorithm (model A) it is fixed at ±20%. In accordance with the five reaction-solvent indices explained in the introduction (Figure 1.14), proximity to the polarity of the solute results in a higher score being attributed to that solvent. A diagram is available in the spreadsheet indicating the polarity of solvent candidates and whether they reside within the user defined limits (Figure 2.17). In this case study DMF is within the highest category (RS_{6a} = 1) while toluene fails the solvent selection algorithm assessment based on its poor polarity relative to 1,2-dimethylimidazole (RS_{6a} = 5). All solvents that do not reside within these limits are deemed unsuitable and removed from the solvent set. Destabilisation of any reaction components was not deemed necessary and so this part of rule G (RS_{6b}) was not applied.

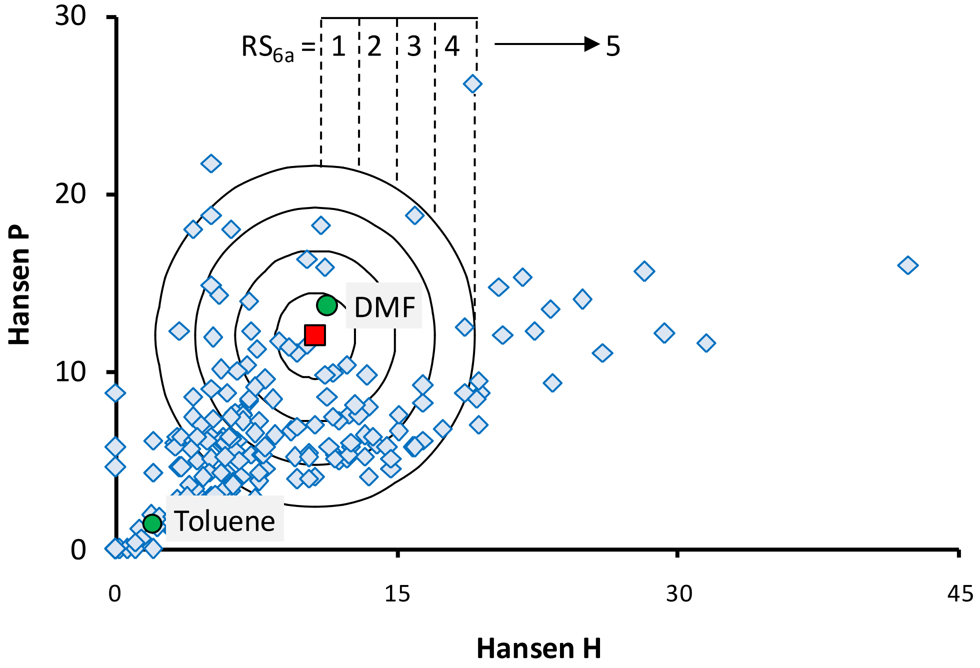
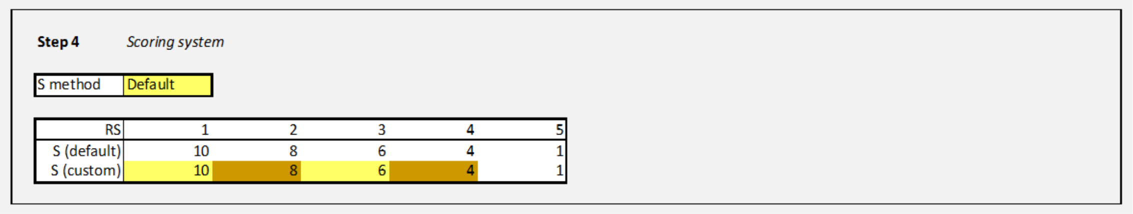


Figure 2.17 Menschutkin reaction solvent selection algorithm screenshot: Polarity matching diagram for estimating the solubility of 1,2-dimethylimidizole.

Solvent neutrality and association and/or disassociation can also be addressed, a remnant of the original method that is not vital in the revised version of the solvent selection algorithm, in which the emphasis is shifted from chemical engineering to laboratory organic synthesis. As such these rules are not applied in this case study, and neither is rule J because greenness and associated metrics are now assessed with the additional solvent greenness assessment based on solvent selection guides instead (model B3).

The original scoring method that translates user defined parameters (the reaction indices, R_i) into the relevant reaction-solvent index (RS_i) and then into a useful score (S_i) can be amended by the user if desired. The original method was kept for this case study (Figure 2.18). The number of solvents passing the assessment is not changed, only the score associated with each solvent because the score given for RS_5 cannot be changed from 1. Any solvent scoring a 1 in any reaction index in the solvent selection algorithm results in a fail and is excluded from the final list of solvent candidates.



Step 4 *Scoring system*

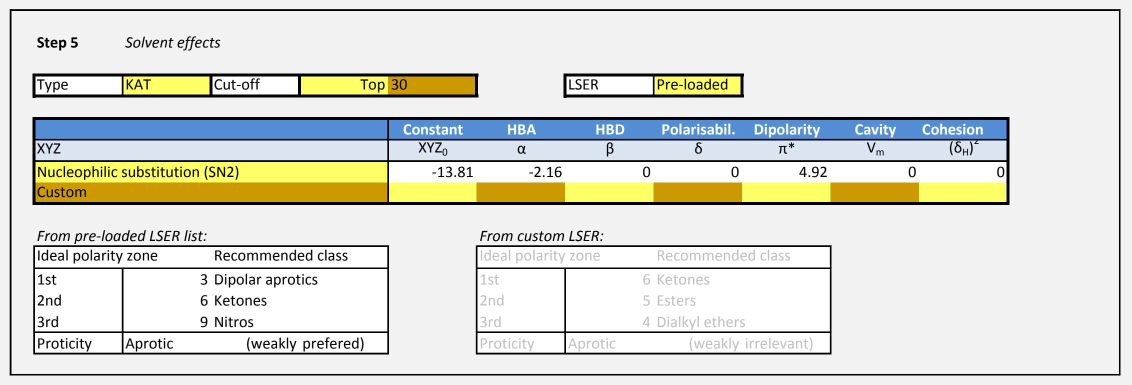
S method: Default

RS	1	2	3	4	5
S (default)	10	8	6	4	1
S (custom)	10	8	6	4	1

Figure 2.18 Menschutkin reaction solvent selection algorithm screenshot, step 4: Scoring system.

At this stage, data entry for the original solvent selection algorithm (model A) and the thermodynamic aspect of the revised solvent selection algorithm (model B1) is complete. These can be compared and questioned, but generally the conclusions are the same; unsurprising given they share much of their algorithmic structure. Before the results are considered, the extension to the revised solvent selection algorithm concerning solvent effects, the solvent performance assessment (model B2), should be completed. When a LSER is available describing a parameter such as the rate of reaction, the performance of solvent candidates can be estimated prior to any experimental work. Because it seems that generally solvent effects are of the same nature across a reaction class (heteroatom alkylation in this instance), if data is available for a related system it can be used in the absence of reaction specific data as an approximation. A custom LSER can also be entered based on laboratory data. This aspect of the assessment is heavily dependent on solvatochromic polarity parameters (*e.g.* Kamlet-Taft solvent polarity parameters) and so many solvents will fail this part of the solvent selection algorithm purely due to a lack of data. This subset of solvent performance results (model B2) are presented separately so if a solvent does well in the other solvent selection algorithm models that will not be lost because of undetermined polarity measurements. Either Kamlet-Taft or the related Catalán parameters can be used [Catalán 2009]. Although analogous in their construction, the latter are newer and unproven. As such the Kamlet-Taft solvent polarity parameters were used in this instance. These were sourced from a variety of publications and novel data featured throughout this work (Table 8.2). Problems with inconsistencies do occur as often different dye pairs or even averaged values are used in the determination of the Kamlet-Taft solvent polarity scales. This will be mentioned in instances where data might be skewed. Only the top 30 scoring solvents were allowed to pass the assessment, although up to 50 are permitted depending on the wishes of the user (Figure

2.19). Predicted ln(**k**) values of solvent candidates were calculated from the Menschutkin S_N2 LSER (Equation 2.2) which is preset into the spreadsheet along with other relationships from the literature and the studies contained in this work. The new polarity parameter **ε** is not represented in the solvent selection algorithm spreadsheet, which is why Equation 2.3 was not used.



Step 5 *Solvent effects*

Type	KAT	Cut-off	Top 30

LSER	Pre-loaded

	Constant	HBA	HBD	Polarisabil.	Dipolarity	Cavity	Cohesion
XYZ	XYZ_0	α	β	δ	π^*	V_m	$(\delta_H)^2$
Nucleophilic substitution (SN2)	-13.81	-2.16	0	0	4.92	0	0
Custom							

From pre-loaded LSER list:

Ideal polarity zone	Recommended class	
1st	3	Dipolar aprotics
2nd	6	Ketones
3rd	9	Nitros
Proticity	Aprotic	(weakly prefered)

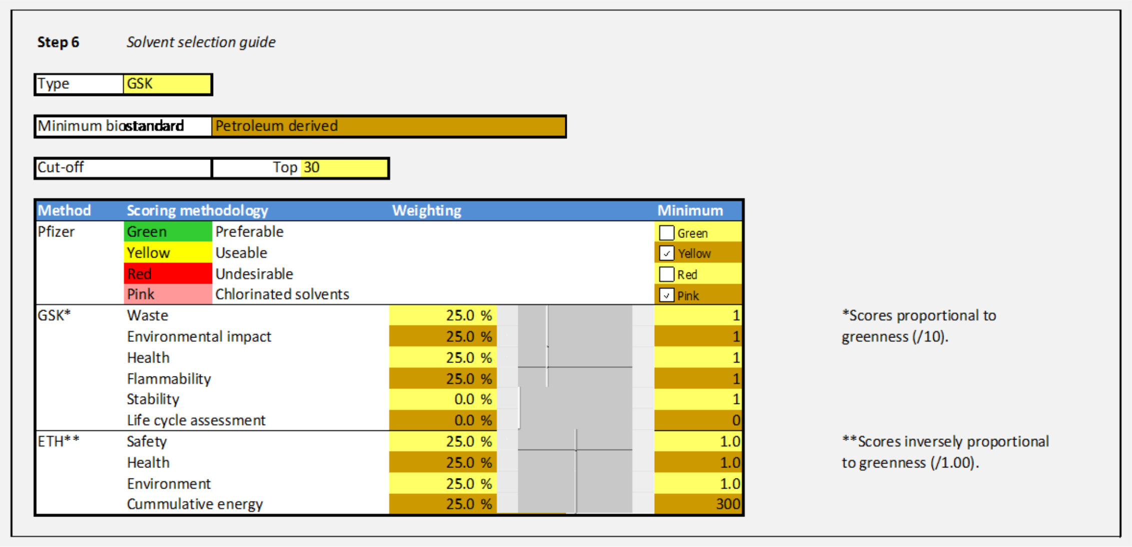
From custom LSER:

Ideal polarity zone	Recommended class	
1st	6	Ketones
2nd	5	Esters
3rd	4	Dialkyl ethers
Proticity	Aprotic	(weakly irrelevant)

Figure 2.19 Menschutkin reaction solvent selection algorithm screenshot, step 5: Solvent effects.

In the same data entry step (Figure 2.19), the spreadsheet will also suggest which general classes of solvent may be useful to maximise reaction rates based on their polarity, which can be fed into step 2 to narrow the number of solvents passing the assessment (model B1) to only those that will provide an effective reaction medium. As with the preset LSER equations, a custom LSER can also be used to generate a list of the best three solvent classes predicted to most accelerate the rate of reaction. The polarity zones are the same as those used in Chapter 1 to characterise solvent properties (Figure 1.20). This may be helpful to flesh out an understanding of solvent effects from a small preliminary study. The LSER for the Menschutkin reaction was converted into the suggestion that highly dipolar aprotic solvents are the ideal class of solvent for accelerating the rate of reaction. Secondly ketones are recommended for their relatively high dipolarity, and thirdly the less desirable nitroalkane and nitroarenes are suggested (Figure 2.19). This process can also suggest whether aprotic or protic solvents are preferable. A weak preference for aprotics is suggested because the coefficient for **α** is negative but smaller in magnitude than the positive coefficient of **π***. In actual fact the experimental data reveals that proticity is actually rather damaging to the rate of reaction although admittedly less important than dipolarity and polarisability.

The other novel section to the solvent selection algorithm spreadsheet comprises a greenness assessment of solvent candidates (model B3). The conclusions of three solvent selection guides are contained within the solvent database which can be used to refine the final list of recommended solvent candidates to only those which are relatively benign. For this case study the GSK solvent selection guide was selected, and only the top 30 performing solvents permitted to pass (Figure 2.20). More lenient users may wish to allow up to 50 solvent candidates to pass, or conversely as little as 10. Unlike the SUS-HAS-ECO classifications devised earlier (Table 1.6), a single metric of greenness is required in order to attribute a final overall score to each solvent. Although this is much less informative than a triple parameter assessment it is a simple method of differentiating between solvents. Further detail is always available by referring back to the original solvent selection guide. The contribution of the GSK solvent selection guide categories can be weighted to suit the user. For this case study the GSK solvent selection guide 'stability' category was removed because of the mild reaction conditions. The LCA category was removed because many of the solvents examined by GSK did not have full LCA data available. A minimum score for each category can be implemented, as can a minimum bio-standard to eliminate the least desirable solvent candidates (Table 1.5). Because of the strict cut-off (only the top 30 solvents pass), and the fact that many obvious bio-based solvents were excluded from the start because of potential reactivity, solvents of all sources were permitted. In other circumstances it is advisable to select from only established bio-based solvents, or solvents that will probably be available in the near future from renewable feedstocks. All the necessary SUS classifications are contained in the database of solvent properties. If the user wishes to use the Pfizer solvent selection guide then a minimum rank from the colour coded guide is selected. A new colour, pink, is introduced which encompasses all chlorinated solvents otherwise not present in the small solvent set of the Pfizer solvent selection guide. This is in order to remove them, as is often desirable.



Step 6 *Solvent selection guide*

Type: GSK

Minimum biostandard: Petroleum derived

Cut-off: Top 30

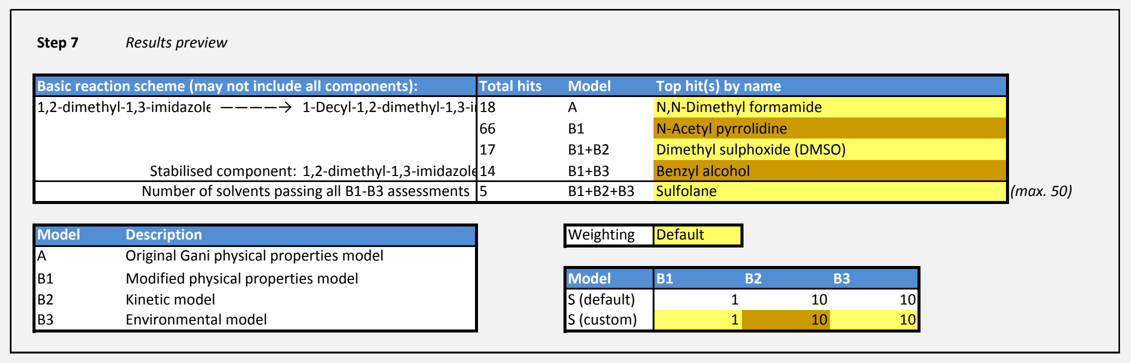
Method	Scoring methodology		Weighting	Minimum
Pfizer	Green	Preferable		☐ Green
	Yellow	Useable		☑ Yellow
	Red	Undesirable		☐ Red
	Pink	Chlorinated solvents		☑ Pink
GSK*	Waste		25.0 %	1
	Environmental impact		25.0 %	1
	Health		25.0 %	1
	Flammability		25.0 %	1
	Stability		0.0 %	1
	Life cycle assessment		0.0 %	0
ETH**	Safety		25.0 %	1.0
	Health		25.0 %	1.0
	Environment		25.0 %	1.0
	Cummulative energy		25.0 %	300

*Scores proportional to greenness (/10).

**Scores inversely proportional to greenness (/1.00).

Figure 2.20 Menschutkin reaction solvent selection algorithm screenshot, step 6: Solvent selection guide.

Results of the solvent selection algorithm: At this point the suitability of solvent candidates has been established. The primary visualisation of the results takes the form of a list of solvents recommended for use by the original solvent selection algorithm (model A), the revised version of this method (model B1), then the revised algorithm in combination with the solvent performance assessment (model B2), and finally the revised solvent selection algorithm in combination with the greenness calculation (model B3) (Figure 2.21). The scores accumulated by each solvent are not shown in this step. The scores can be accessed later and used to prioritise solvent selection if multiple solvents are recommended. Only three scoring reaction indices are active in the original solvent selection algorithm method and so the maximum score is 30 (specifically rule C, rule D, and rule G). There are no scores associated with rule A or rule B. In the revised solvent selection algorithm (model B1) the boiling point and melting point of each solvent candidate is treated separately and so the maximum score is 40. This should be taken into account if trying to compare across the two methods, which is why it is recommended to deal with the scores in a less than quantitative manner.



Step 7 *Results preview*

Basic reaction scheme (may not include all components):	Total hits	Model	Top hit(s) by name
1,2-dimethyl-1,3-imidazol ————→ 1-Decyl-1,2-dimethyl-1,3-i	18	A	N,N-Dimethyl formamide
	66	B1	N-Acetyl pyrrolidine
	17	B1+B2	Dimethyl sulphoxide (DMSO)
Stabilised component: 1,2-dimethyl-1,3-imidazol	14	B1+B3	Benzyl alcohol
Number of solvents passing all B1-B3 assessments	5	B1+B2+B3	Sulfolane

(max. 50)

Model	Description
A	Original Gani physical properties model
B1	Modified physical properties model
B2	Kinetic model
B3	Environmental model

Weighting	Default

Model	B1	B2	B3
S (default)	1	10	10
S (custom)	1	10	10

Figure 2.21 Menschutkin reaction solvent selection algorithm screenshot, step 7: Results preview.

The solvent performance (model B2) and greenness (model B3) assessments both consist of a single reaction index and so a maximum score of 10 is achievable. The usual score hierarchy is applied based on the rank of each solvent within the allowed cut-off, which in this case is thirty solvents in both models. The top quarter of solvents are awarded a score of 10 (which equates to 8 solvents here), the following quarter a score of 8, then a score of 6, and then a score of 4 for the final quarter. Solvent candidates outside the cut-off are designated with a score of 1 which equates to a fail (Figure 1.14).

The score of each solvent candidate is multiplied by the default weighting of 10 chosen for this case study, and added to the score obtained in the revised solvent selection algorithm (model B1). This results in a maximum score of 140. For example, a solvent candidate achieving a score of 34 in model B1 and just passing the requirement of model B2 would have an overall score of 74 in the combined assessment. The default scoring system favours the performance and greenness assessments over the basic thermodynamic model ten to one because model B1 (or model A) is not terribly important aside from the pass or fail aspect. If a solvent is a liquid at the reaction temperature and will dissolve the necessary solutes then this is generally enough as a starting point. It is the greenness of that solvent and how rapid the desired reaction will progress in that solvent that is prioritised here.

The original method returned 18 hits of which DMF was the highest scoring. Although this is a quite valid suggestion using traditional metrics, DMF is not a solvent that should be recommended for chemistry given its chronic toxicity issues [MSDS 2013]. The revised solvent selection algorithm does not fare much better. Again DMF is the top scoring solvent with a total 66 solvents passing the assessment. With so many candidates in the final solvent set the outcome is not any better than an educated judgement. This solvent set contains many of the candidates previously discounted because their melting points and boiling points are not symmetrically distributed around the reaction temperature (reaction index R_2).

When the revised solvent selection algorithm (model B1) is used in conjunction with the assessments concerning performance (*via* LSER, model B2) and greenness (*via* solvent selection guides, model B3), the number of recommended solvents is refined to a smaller, more manageable and useful amount. Those solvents passing the assessments are presented in the following table, accompanied by their scores (Table 2.4). Five solvents appear to satisfy all criteria: anisole, benzyl alcohol, dimethyl-1,3-propylene urea (DMPU), DMSO, and sulpholane. But benzyl alcohol with its proticity and anisole's modest dipolarity leave them ranked 26th and 27th respectively in the solvent set according to predicted reaction rate constant (model B2). Whereas anisole and benzyl alcohol narrowly made the top 30 cut in the solvent performance assessment (model B2), DMSO and sulpholane were ranked an impressive 2nd and 3rd respectively.

The solvent in which the velocity of the reaction is predicted to be maximised is nitrobenzene ($\boldsymbol{\pi^*}$ = 1.01). This solvent should give a marginal improvement over DMSO in this regard but presents other issues. It is known that nitrobenzene is toxic, and also considered as a carcinogen [MSDS 2013]. The production of nitrobenzene requires benzene as the substrate, reacted with nitric acid in sulphuric acid which is incredibly exothermic and needs to be carefully controlled [Lailach 1988]. Nitrobenzene is not present in the GSK solvent selection guide, the Pfizer solvent selection guide, or the ETH solvent selection guide and so cannot take part in the greenness assessment (model B3). By contrast sulpholane is thought of as a much more benign

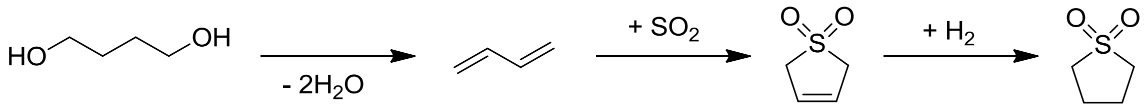
Table 2.4 Solvent hits generated by the solvent selection algorithm for the Menschutkin reaction.

Solvent	Score: Model B1 + model B2	Score: Model B1 + model B3
N-Acetylpiperidine	96	No data
N-Acetylpyrrolidine	116	No data
Anisole	72	72
Benzyl alcohol	74	114
Cyclohexanol	Fail	86
1,2-Dichloroethane	88	Fail
N,*N*-Diethyl acetamide	116	No data
DGME	No data	114
DMAc	136	Fail
DMF	138	Fail
DMPU	138	98
DMSO	136	76
Ethoxybenzene	Fail	112
2-Ethylhexanol	No data	112
Isoamyl alcohol	Fail	92
Methyl lactate	No data	76
NMP	136	Fail
Nitrobenzene	132	No data
Nitromethane	90	Fail
2-Pentanol	No data	74
2-Propanol	Fail	68
N-Propionylpyrrolidine	94	No data
Piperylene sulphone	108	No data
Sulpholane	130	130
Tri(ethylene glycol)	No data	134
Triethylphosphate	78	No data
Total hits	17	14

solvent, only bettered by tri(ethylene glycol) in the greenness assessment. Amongst those solvents passing the performance assessment (model B2) but not the greenness assessment (model B3) are NMP and DMF.

Implementing solvent selection: From the LSER it can be deduced that the ideal solvent for accelerating the rate of reaction will be very dipolar and aprotic. From the experimental solvent set, only DMSO fulfils this requirement and fully passes the revised solvent selection algorithm. Although DMSO is a clear favourite from this group there is scope for the introduction of a new highly dipolar aprotic solvent. This new solvent should address the issues of existing dipolar aprotic solvents, especially with an eye on improving the EHS benchmark set by DMSO; currently presenting the least issues of any solvent widely used in its class.

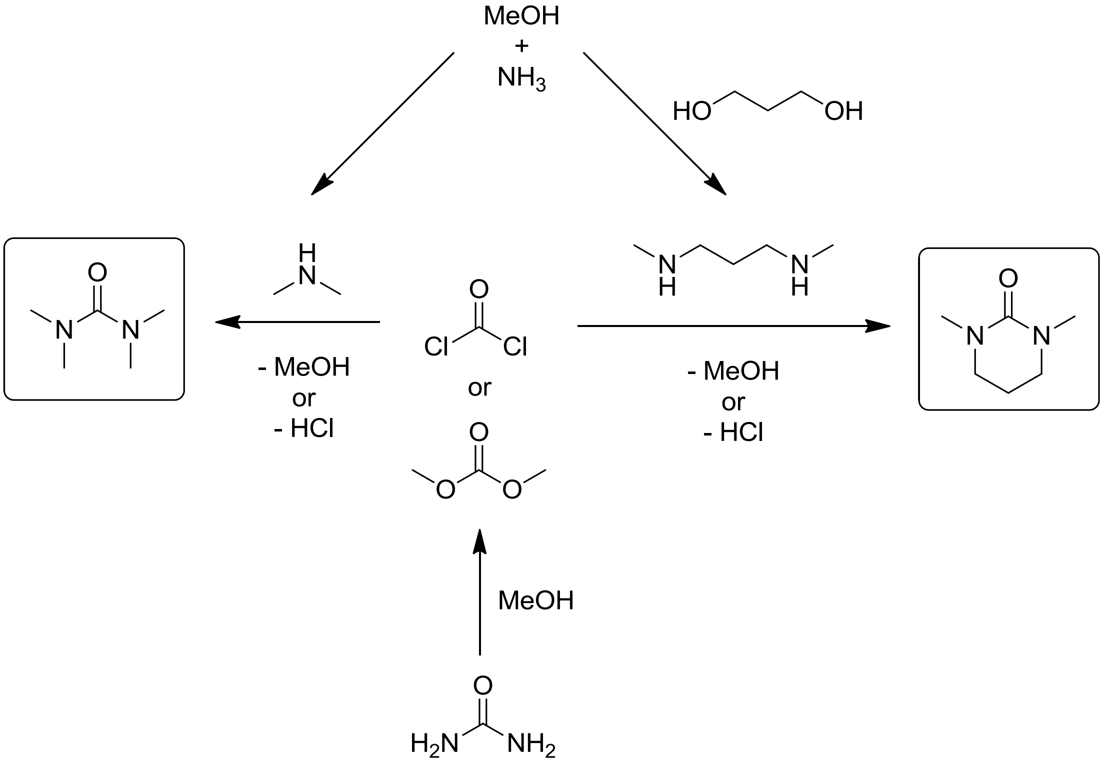
Within the solvent selection algorithm only sulpholane performs consistently well in all assessments, although dimethyl-1,3-propylene urea (DMPU) and DMSO both outscore sulpholane in the solvent performance assessment (model B2). Sulpholane is more toxic than DMSO (about same order as DMF and DMAc) but has much lower skin permeability than all other dipolar aprotics [Tilstam 2012]. Sulpholane, like DMSO, is not a reproductive toxicant like DMF. In terms of safety (flash point, vapour pressure) sulpholane is better positioned than other dipolar aprotics too. It could be manufactured from bio-1,4-butadiene (a bio-derived platform molecule), which aside from a necessary preliminary dehydration to butadiene, would not cause a deviation from the existing manufacturing infrastructure (Scheme 2.5) [BASF 2013, BioAmber 2013, Myriant 2013].



Scheme 2.5 The manufacturing route to bio-based sulpholane.

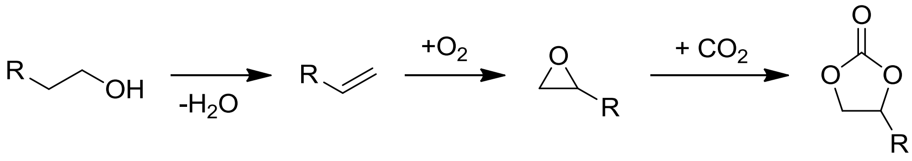
Ureas are believed to be safer dipolar aprotics to replace amides (Scheme 2.6). Dimethyl-1,3-propylene urea (DMPU) featured 6th in the list of solvents passing the solvent performance assessment (model B2) of the solvent selection algorithm. Interest in solvent substitution in favour of DMPU was ruled out during the earliest years of the green chemistry movement because of suspected mutagenicity [Lo 1990]. Still, DMPU is not more toxic than NMP or DMF. This is implied by the score of 4 DMPU receives in the GSK solvent selection guide health category compared to the 2 out of 10 scored by DMF, and the 3 of NMP [Henderson 2011]. Equally worrying is the LCA category of the GSK solvent selection guide, in which DMPU only scores 3 out of 10. This is lower than DMF (7) and NMP (4). The LCA category was not considered in the solvent selection algorithm because of gaps in the data set makes for an unfair comparison. But given that a life cycle of DMPU does exist, it should be used to better understand the environmental impact of that solvent.

The reaction between phosgene and amines seems likely to have been the most economical production method for urea solvents [Lüttringhaus 1964]. If alkylation of urea is not directly feasible then dimethyl carbonate can be made (albeit inefficiently) from urea and methanol [Zhang 2012], which in turn could be reacted with an amine to give a choice of alkylated urea derivatives (Scheme 2.6). A lack of complete toxicity, biodegradability, and physical property data will always be a disconcerting problem with new products, including bio-based solvents, and ureas are no different. Because of the issues with the suspected mutagenicity of DMPU, and its significant environmental impact, it is sensible not to proceed into experimental verification [Henderson 2011]. Nevertheless we cannot rule out all liquid urea derivatives on this basis, and tetramethylurea has been suggested as another viable urea solvent which may be more promising in other applications [Lüttringhaus 1964].



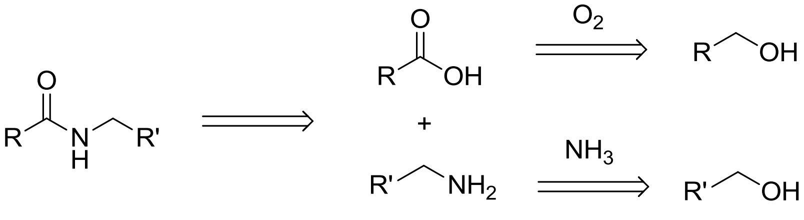
Scheme 2.6 The potentially bio-derived ureas tetramethylurea (left) and dimethyl-1,3-propyleneurea (DMPU, right).

Acetonitrile has been proposed as a possible replacement (in certain favourable instances) for dipolar aprotics like DMF [Alfonsi 2008]. But supply shortages and a significant environmental impact associated with its disposal have tainted the perception of this solvent [McConvey 2012]. In fact efforts are now being made to replace acetonitrile in chromatographic applications [Brettschneider 2010]. Another attractive alternative to amide solvents are cyclic carbonates (Scheme 2.7). The cyclic structure creates high dipolarity without resorting to nitrogen or sulphur containing functional groups. Therefore no NO_x or SO_x will be released upon incineration. Alcoholic fermentation products could serve as a feedstock in combination with carbon dioxide. Despite all these positive factors, propylene carbonate and ethylene carbonate, although scoring admirably in the GSK solvent selection guide, are electrophilic and may react with the nucleophilic 1,2-dimethylimidazole, and so for that reason carbonate solvents were excluded from the solvent selection algorithm.



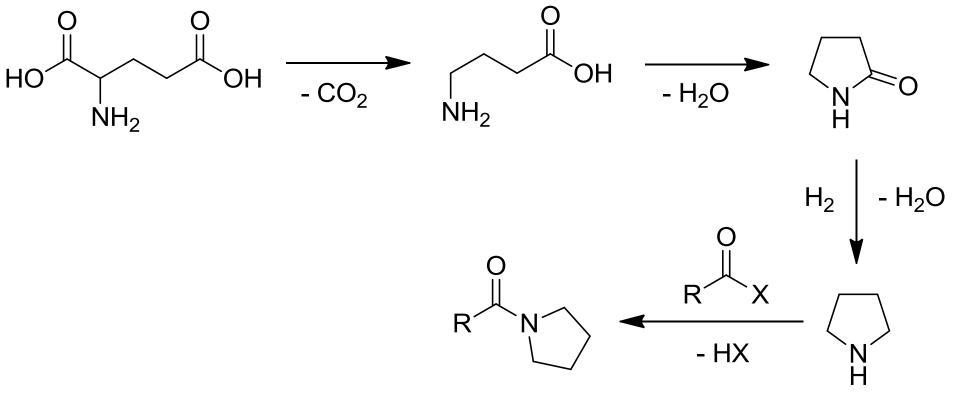
Scheme 2.7 The synthesis of cyclic carbonate solvents from alcohols.

With some competition provided by the aforementioned oxides of organosulphur compounds, ureas, and possibly also alkyl phosphates, amides appear to be the most likely class of solvent capable to serve as highly dipolar, basic, and aprotic reaction media. The retrosynthesis of an amide indicates an amine and a carboxylic acid as starting materials (Scheme 2.8). The latter are common renewable platform molecules, with acetic acid and succinic acid for example directly accessible by fermentation [Okino 2008, Yamada 2008]. Triglycerides are composed of three fatty acid moieties [Turley 2008 page 26]. The nitrogen containing portion of an amide could come from the alkylation of ammonia using bio-based alcohols. This would mirror current, if unsustainable, production methods for DMF and DMAc that utilise dimethyl amine [Weissermel 1997 page 43]. DMF is produced by the reaction of dimethylamine with carbon monoxide and so does not fit this retrosynthesis exactly.



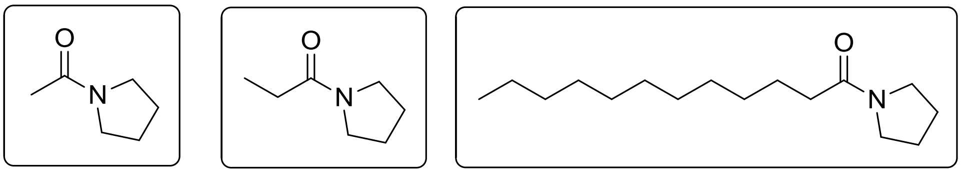
Scheme 2.8 The possibility of using a bio-based alcohol precursor for the synthesis of amides.

An alternative route to a bio-based amide solvent would be to modify an amino acid derivative. A catalogue of research has recently arisen on the use of glutamic acid as a platform molecule, including its role in the synthesis of NMP *via* pyrrolidinone [Lammens 2010]. Rather than alkylating pyrrolidinone to give the undesirable NMP or its analogues, reduction would give the amine pyrrolidine, which in turn could be acylated to give an amide (Scheme 2.9). The proposed synthetic route appears to be uncomplicated with no harmful by-products if an acid or ester is used in the final step of the reaction with pyrrolidine. It would seem to fit the requirement of feasible scale-up and solvent production. However the catalytic reduction of pyrrolidinone to pyrrolidine by hydrogen is perhaps slightly fanciful in its optimism [Núñez-Magro 2007]. Hydride reduction would be more probable, creating the hazardous and stoichiometric waste associated with lithium aluminium tetrahydride and other sources of the hydride anion. Otherwise the chemistry, which begins with an enzymatic decarboxylation, is straightforward.



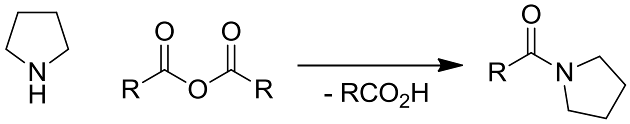
Scheme 2.9 The synthesis of a class of *N*-acylpyrrolidines from glutamic acid.

Examples of pyrrolidine derived bio-based amide solvents are *N*-acetylpyrrolidine, *N*-propionylpyrrolidine, and *N*-laurylpyrrolidine (Scheme 2.10). Acetic acid and propionic acid are derivatives of fermentation products. Lauric acid is the most abundant fatty acid moiety in the triglyceride oil of coconuts [Laureles 2002]. It has a short alkyl chain relative to most fatty acids which will keep the melting point of the resulting amide solvent as low as possible. Lauric acid is also saturated, reducing the ways in which it can easily be oxidised or otherwise degraded. These liquids have never been applied as solvents in the past. The amide of lauric acid and dimethylamine has been used as a solvent but the intended reaction did not occur [Pérez-Sánchez 2012]. None of these bio-based amides succeeded in negotiating the original version of the solvent selection algorithm (model A). They were predicted to dissolve and stabilise the solute 1,2-dimethylimidazole but failed reaction index R_2: the solvent must be a liquid at the reaction temperature. All three are actually liquid at 323 K, but because of the way the original selection algorithm (model A) takes the mid-point of melting point and boiling point and compares it to the application temperature these three solvents all fail the assessment. When ordered by predicted performance (model B2) *N*-laurylpyrrolidine ranks 37th in the set of solvent candidates, placing out of the required top 30 in order to be recommended by the solvent selection algorithm. The other two bio-based amides do feature in the top 30 passes. A lack of data prohibits any judgements on environmental impact or health and safety.



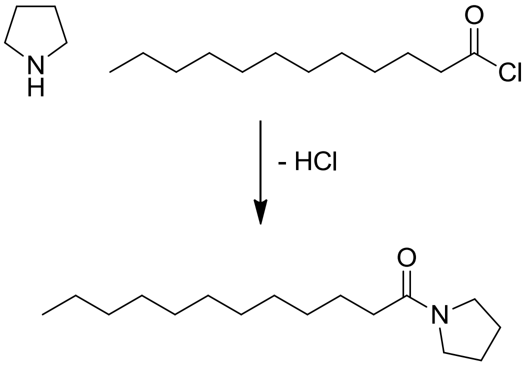
Scheme 2.10 The proposed bio-based amide solvents *N*-acetylpyrrolidine (left), *N*-propionylpyrrolidine (centre), and *N*-laurylpyrrolidine (right).

Bio-based amide solvent synthesis and characterisation: Acetyl and propionyl amides of pyrrolidine were synthesised from their respective acid anhydrides (Scheme 2.11). Moisture and surplus anhydride and carboxylic acid by-product was removed by addition of sodium hydroxide, magnesium sulphate, and dilution in DCM [Beamson 2010]. Using an auxiliary petroleum derived solvent in the workup of the synthesis of a bio-based solvent negates the point of performing the procedure. An alternative purification reported in the literature for *N*-acetylpiperidine using vacuum distillation was recognised as poor by the authors and the workup employed here was used instead [Beamson 2010]. The synthesis of solvents is of high importance. Ultimately simple and facile transformations to give high yields are essential. If manufacturing costs are too high then the solvent will not compete in the marketplace. A procedure that reacts carboxylic acids directly with pyrrolidine without resorting to using a coupling agent or solvent would be much more preferable to the procedure documented here. To this end silica based catalysts may be applicable [Comerford 2009]. If and when this new class of bio-amide solvent is proven to be adequate in a prominent application then it will make sense to investigate a reasonable manufacturing process.



Scheme 2.11 Generalised reaction scheme describing the synthesis of *N*-acylpyrrolidines. R = Me- or Et-.

Even less appealing was the synthesis of *N*-laurylpyrrolidine, which was achieved by the addition of lauryl chloride to a solution of excess pyrrolidine in 2-MeTHF (Scheme 2.12). The single point of merit is the replacement of THF typical of these procedures with bio-based 2-MeTHF [Kolocouris 1994]. Upon completion of the reaction to give *N*-laurylpyrrolidine, sodium hydroxide was added and the amide isolated as before. These procedures offered yields below ideal efficiency but quantities of product were enough to apply them as solvents in the model S_N2 reaction (Figure 2.22). Full procedures and yields can be found in the relevant appendix.



Scheme 2.12 A fatty acid derived amide.

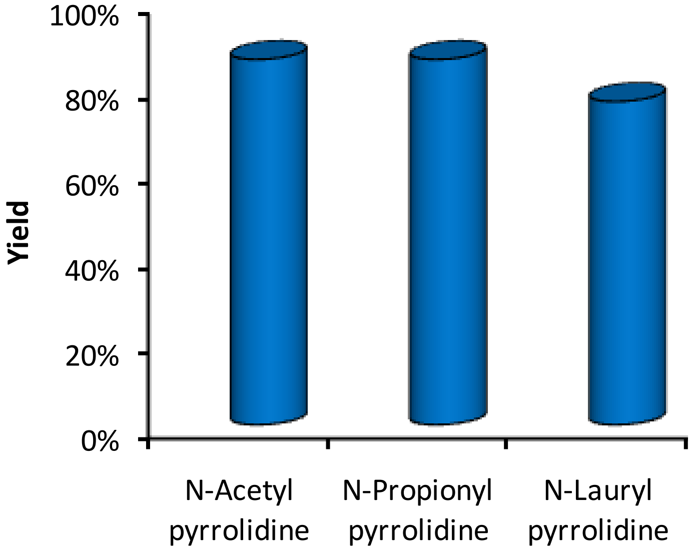


Figure 2.22 Bio-based amide solvent yields.

The polarity of these new solvents was established using the method described by Kamlet and Taft [Kamlet 1983]. From the following figure it appears that *N*-acetylpyrrolidine and *N*-propionylpyrrolidine are marginally less dipolar than the other amide solvents DMF, DMAc, and NMP (Figure 2.23). The larger carbon atom to heteroatom ratio in the bio-based amides has resulted in less dipolar character. This is exaggerated in the case of *N*-laurylpyrrolidine which has sixteen carbon atoms compared to the three of DMF. The polarities of tetramethylurea and DMPU were also examined. As expected the latter cyclic solvent is more dipolar than the former. Sulpholane is more akin to a ketone in terms of its polarity profile than an amide or DMSO but it has a strong dipole resulting in a **π*** value of 0.98. This makes it a strong candidate for solvent selection, probably more so than the less dipolar bio-based amide solvents.

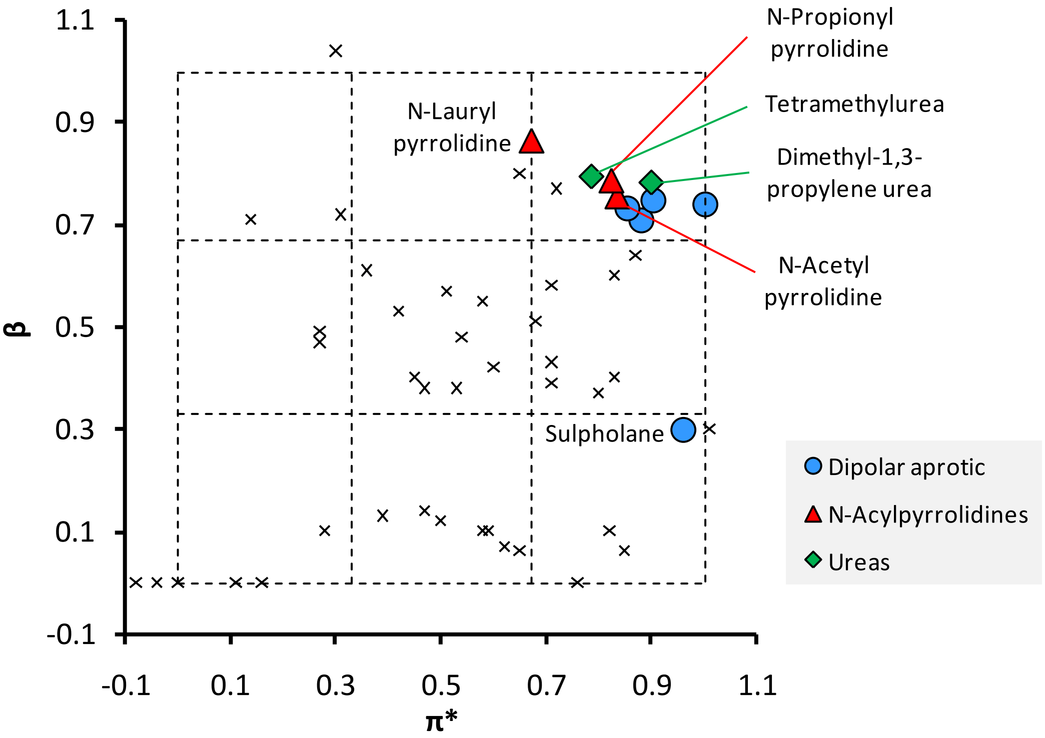


Figure 2.23 The polarity of highly dipolar aprotic solvents including bio-based amides.

Application of bio-based solvents: Leading on from the results of the solvent selection algorithm the bio-based amide solvents were applied in the Menschutkin reaction along with sulpholane. What is immediately obvious is that none of these new bio-based amides are able to surpass the existing highly dipolar aprotic solvents in terms of accelerating the reaction rate (Table 2.5). With dipolarity (**π***) crucial to the rate of the reaction, *N*-laurylpyrrolidine is less satisfactory than even acetonitrile. Sulpholane performs better in terms of kinetics, slightly more than predicted. It should be made clear that no LCA data is available for sulpholane [Henderson 2011]. This has an impact on the reported greenness of sulpholane with the maximum ECO score of 10 (that would normally include the LCA category) very tentative indeed. The method of calculating the ECO classification is stated in Chapter 1 (Equation 1.12). The bio-based amides are not present in the

GSK solvent selection guide and so no comment can be made regarding their greenness within the SUS-HAS-ECO framework.

Table 2.5 The properties and reaction rates of the Menschutkin reaction in DMSO and optimal solvent candidates.

Solvent	ln(**k**)	α	β	π*	SUS	HAS	ECO
DMSO	-9.01	0.00	0.74	1.00	4	7	4
Sulpholane	-8.81	0.00	0.30	0.96	4	9	10[a]
N-Acetylpyrrolidine	-9.59	0.00	0.76	0.83	6	n/a	n/a
N-Propionylpyrrolidine	-9.61	0.00	0.79	0.82	6	n/a	n/a
N-Laurylpyrrolidine	-10.55	0.00	0.86	0.67	6	n/a	n/a

[a]No LCA score available in the GSK solvent selection guide [Henderson 2011].

The predicted rate constants fit very well with experimentally obtained data. A comparison between experimental and predicted ln(**k**) can be updated to reflect this (Figure 2.24). Sulpholane produced the greatest reaction acceleration of all the solvents. It is not the

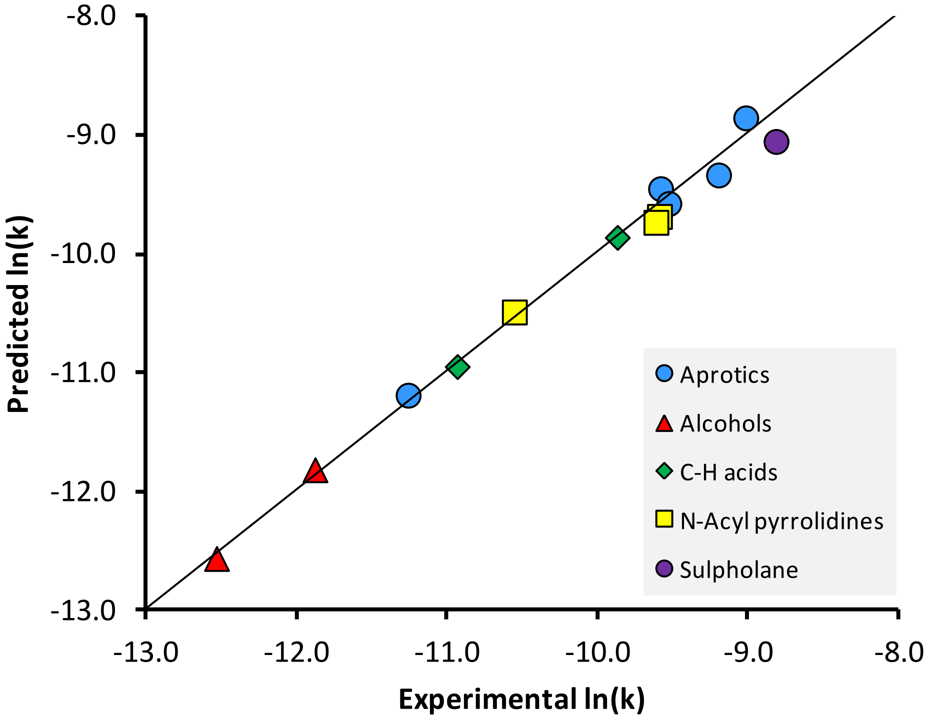


Figure 2.24 The performance of bio-based amide solvents and sulpholane in the Menschutkin reaction.

most dipolar solvent according to its **π*** value, this status is reserved for DMSO. The disparity is only small however, and this role reversal is probably due to experimental error. Attempts to improve the correlation with the Hildebrand solubility parameter failed. *N*-Acetylpyrrolidine and *N*-propionylpyrrolidine are competitive but not outstanding in terms of reaction kinetics, nestling at the lower end of the region defined by the highly dipolar aprotic solvents in Figure 2.24.

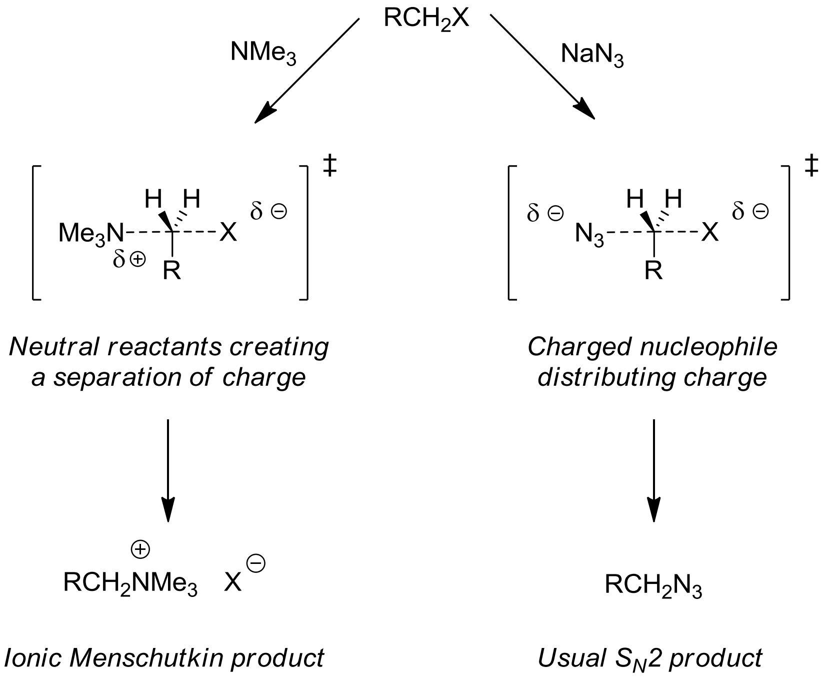
2.3 Heteroatom alkylation summary

The success and implications of solvent selection: The solvent effect determining the rate of the model Menschutkin reaction demands a highly dipolar and aprotic solvent for efficient conversion to the product. The precise reason for the rate enhancement in dipolar solvents can be attributed to a reduction in the enthalpy of activation, which outweighs a reduction in the entropy of activation [Kondo 1984]. The drawbacks of the optimal performing solvents (water miscibility, high boiling point, teratogenicity, depleting feedstock) have not been resolved through solvent selection. The health issues associated with solvents bearing an amide functionality would seem to be inseparable from their high dipolarity, both arising from the carboxamide moiety. Some potentially bio-derived amide solvents were synthesised, but the looming uncertainty over their EHS profiles means that there is no secure reason by which they can be recommended as replacements for established highly dipolar aprotic solvents, casting doubt on their suitability as solvents.

As identified using the solvent selection algorithm, no performance is lost by using more benign solvents such as sulpholane. Sulpholane resolves previous compromises between reaction performance and environmental impact, and could be produced from renewable feedstocks (Scheme 2.5). Sulpholane (£44.20, 500 g, Sigma-Aldrich, 99% purity, as of 21st June 2013) is reasonably comparable in price to DMSO (£35.00, 500 mL, Sigma-Aldrich, 99.5% purity, as of 21st June 2013).

To their detriment, sulpholane and nitrogen containing solvents will present end of life air pollution issues if incinerated. Cyclic carbonates are sufficiently dipolar to serve in nucleophilic substitution reactions without the possibility of producing SO_x or NO_x emissions upon incineration. This presents an opportunity for future work in this area. However for this specific case study carbonate solvents were excluded from the solvent selection algorithm for fear of unwanted nucleophilic addition. Like sulpholane, cyclic carbonates are probably good solvent candidates to consider for the wider range of chemistries reliant on highly dipolar yet aprotic solvents. This includes cross coupling reactions and some types of hydrogenation.

Broader appeal: It has been said of the Menschutkin reaction that it is "*a special kind of S_N2 reaction where the reactants are uncharged, in contrast to the most* [*sic*] *usual S_N2 reactions where one of the reactants is charged*" [Solà 1991]. The implication is that the separation of charge in the Menschutkin reaction will respond differently to the polarity of the solvent than the transfer of charge from nucleophile to leaving group in other S_N2 reactions (Scheme 2.13). This could mean that the case study presented here is not of broad interest. The Hughes-Ingold rules state that because the activated complex of the Menschutkin reaction generates a separation of charge in previously neutral molecules, the Gibbs free energy of activation will be lowered in polar solvents [Hughes 1935]. In this instance 'polar' is vaguely used to imply dipolarity. The more usual S_N2 reaction of an anion displacing a leaving group from an uncharged species has a transition state at which the activated complex features a shared charge across two electronegative centres. The reduction in point charge means that 'polar' solvents now discourage the reaction according to the Hughes-Ingold rules. Instead stabilisation of the ionic reactant (and the ionic product) is now preferentially favourable. The Gibbs free energy of activation should be reduced by using 'non-polar' solvents to destabilise the ionic species.

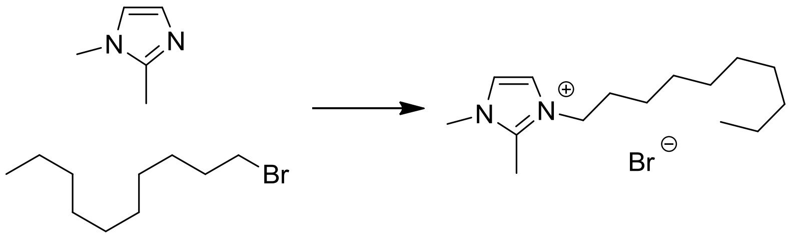

RCH₂X
NMe₃
NaN₃
Me₃N---C---X
δ⊕
δ⊖
N₃---C---X
H H
R
‡
Neutral reactants creating a separation of charge
Charged nucleophile distributing charge
RCH₂NMe₃ X
RCH₂N₃
Ionic Menschutkin product
Usual S_N2 product

Scheme 2.13 The comparison between the mechanism of different types of S_N2 reaction.

To resolve this potential hurdle, classic kinetic solvent effect studies concerning the more usual S_N2 reaction between an anion and an alkyl halide (or pseudo-halide) were reinterpreted using the Kamlet-Taft solvent polarity parameters to construct a LSER (Table 2.6). The reaction of

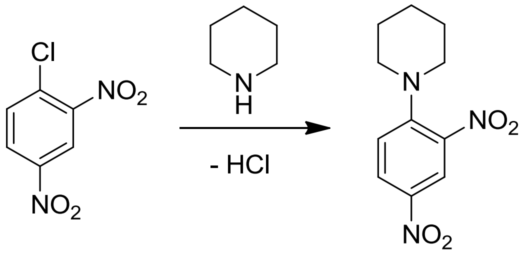
Table 2.6 A collection of LSERs describing S_N2 and S_NAr reactions.

1. Menschutkin reaction.



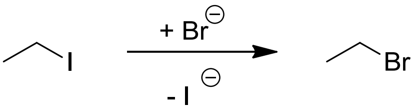
$\ln(\mathbf{k}) = -13.79 - 2.17(\boldsymbol{\alpha} - 1.02\boldsymbol{\varepsilon}) + 4.90\boldsymbol{\pi}^*$ $(R^2 = 0.994)$

2. Aromatic nucleophilic substitution reaction [Mancini 1986].



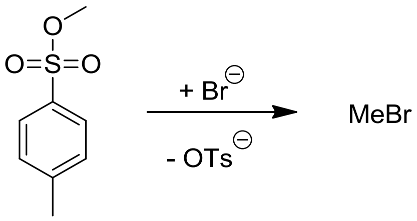
$\ln(\mathbf{k}) = -1.74 - 1.21\boldsymbol{\alpha} + 1.86\boldsymbol{\pi}^*$ $(R^2 = 0.934)$

3. Nucleophilic substitution [Kondo 1982].



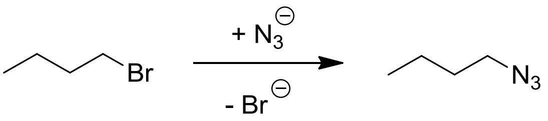
$\ln(\mathbf{k}) = -5.86 - 11.04\boldsymbol{\alpha} + 7.18\boldsymbol{\beta}$ $(R^2 = 0.981)$

4. Nucleophilic substitution [Müller 1972].



$\ln(\mathbf{k}) = -3.62 - 7.08\boldsymbol{\alpha} + 7.52\boldsymbol{\beta}$ $(R^2 = 0.927)$

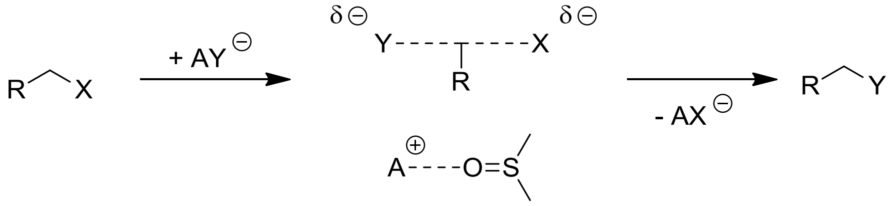
5. Nucleophilic substitution [Delpuech 1965].



$\ln(\mathbf{k}) = -6.58 - 6.25\boldsymbol{\alpha} + 5.10\boldsymbol{\beta}$ $(R^2 = 0.794)$

ethyl iodide and bromide ion at 303 K appears to be kinetically dependent on the **α** and **β** value of the solvent [Kondo 1982]. The familiar inverse proportionality with **α** was probably expected, but this is counterbalanced by an equally strong but opposing dependence on **β**. The same is true of the kinetics of methyl tosylate reacting with tetrabutylammonium bromide [Müller 1972]. The nucleophilic attack of an azide ion on 1-bromobutane follows these same rules, if a little weaker in terms of the strength of the resulting correlation [Delpuech 1965]. This means that instead of searching for dipolar solvent substitutes as with the Menschutkin reaction, strongly hydrogen bond accepting yet aprotic solvents should be sought in order to accelerate more typical heteroatom alkylation reactions. The choice consists of amines, reactive under these conditions, and the dipolar aprotic amides and sulphoxides already investigated. So it is still correct to state (if only by coincidence) that the kinetics of all S_N2 reactions are hastened by highly dipolar aprotic solvents. The kinetics of S_NAr reactions follow the familiar Menschutkin solvent dependence (Table 2.6).

This raises a question over the mechanism of non-Menschutkin S_N2 reactions, which apparently disobey the Hughes-Ingold rules. The negative influence of **α** on the kinetics of the reaction is presumably of the same variety that retards Menschuktin reactions. The beneficial effect of a solvent engaging in hydrogen bonds as an acceptor but not as a donor is not easy to pinpoint from the reaction pathway. It may be as simple as the solvent interacting with the counter ion of the reactive anion, unmasking the nucleophilicity of the reactant (Scheme 2.14). This effect might be more important than stabilising the activated complex, and so the suggestion from the Hughes-Ingold rules to use low polarity solvents is not realised experimentally.



Scheme 2.14 The non-Menschutkin S_N2 mechanism as assisted by DMSO.

The alkylation of 1,2-dimethylimidazole provides access to a few potentially useful products, but the purpose of this case study was to demonstrate a means of synthesising a great variety of compounds by heteroatom alkylation, this being just one example. Products formed from the *N*,*N′*-polyalkylated imidazolium moiety do not feature in medicinal products, but work using these compounds can be translated to other chemistries of wider utility, including singularly substituted imidazoles such as those that are used to treat fungal infection [Karakurt 2001, Kathiravan 2012, Yang 2012a]. Obviously ionic liquids are one product that utilises the dialkylimidazolium moiety directly [Welton 1999, Hallett 2011]. It has been shown that the

Menschutkin reaction offers little in the way of understanding the general S_N2 mechanism of charged nucleophiles but the optimum class of solvent is the same regardless.

Recapitulation and future work: Sulphur containing dipolar aprotic solvents seem superior to nitrogen containing dipolar aprotics in terms of performance and health and safety. End of life air pollution concerns would be very much reduced if neither nitrogen nor sulphur were present in the solvent, and so cyclic carbonates also appear to be an attractive alternative. However they were presumed to be reactive under the conditions of this case study and more research would need to be conducted in this respect. There does not seem to be any significant issues with the long term sourcing of feedstocks for these types of solvents, which are predominately constructed from low molecular weight fragments. Ultimately a LCA for sulpholane would be needed, and compared to those of other solvents such as propylene carbonate to arrive at any definitive conclusion.

The solvent selection algorithm and its associated assessments have successfully enhanced the understanding of the reaction and placed it within a firmer context amongst the other S_N2 reactions. Although issues clearly remain, sulpholane offers an improvement over many other seemingly similar solvents. What has become very clear is that any mechanistic knowledge of the reaction in itself is a massively useful tool when it comes to solvent selection. All of what has been learnt from studying this model reaction can be applied in the subsequent case studies to further develop the art of solvent selection.

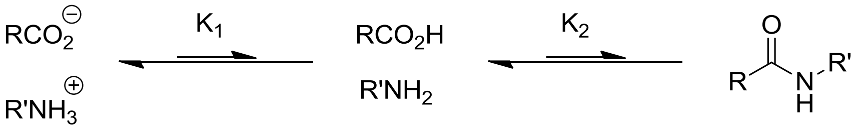
3. Amidation

The creation of amide linkages is fundamental in as much as amino acids combine in this way to form proteins. The synthesis of amide functionalised chemical products is prevalent in the fine chemical and pharmaceutical industries. One in four anti-cancer drugs contains a carboxamide functionality [Ghose 1999]. Amidation is therefore not simply unavoidable, but prevalent. A recent survey of process development in the pharmaceutical industry places amidation as the most common acylation procedure and second only to *N*-alkylation chemistries overall in terms of the most frequently practiced synthetic procedure [Carey 2006]. Acid chloride (44%) and coupling agent (25%) facilitated routes dominate, even though the health and safety implications are a concern, and the waste streams associated with these procedures of a significant volume. Amidation solvent effects have not been studied to the detail that those in the Menschutkin reaction have, but attempts to interpret what data is available have been conducted in the past [Charville 2010]. The following introductory passages will explain the findings of existing research that is relevant to the understanding of solvent effects in amidation chemistries. The following results help understand the reaction mechanism and apply this knowledge to solvent selection.

3.1 Solvents and amidation

Current mindset and practice: Carbonyl additions feature heavily in synthetic organic chemistry and have widely attracted the attention of those practicing contemporary green chemistry. Traditional methods of ester and amide synthesis rely on aromatic or chlorinated solvents to achieve high yields [Kleeman 2001]. Not much more is known about the role of the solvent. Typically the reaction is assisted by a coupling agent or the carboxylic acid is converted to a more reactive species [Smith 2007 page 1427]. To illustrate the point, one only needs to return to the previous chapter to find evidence of amide forming reactions that are less than ideal in terms of the reagents and conditions used.

Thus the synthesis of amides can be far from the ideal process, in which a carboxylic acid is reacted directly with an amine. Amidation by the liberation of water as the sole by-product, yet without the use of stoichiometric auxiliaries is rarely practiced by synthetic chemists. Although this approach can be adopted for the analogous formation of esters, the basicity of amines is often believed to lead to the unassailably preferential formation of an ammonium carboxylate salt with the carboxylic acid reactant (Scheme 3.1). The equilibrium position is thought to heavily favour the salt over the free reactants [Montalbetti 2005]. The term 'free reactants' will be used to distinguish the molecular species from the products of any acid-base equilibrium involving these substrates.



Scheme 3.1 The commonly perceived competing amidation and salt forming equilibria.

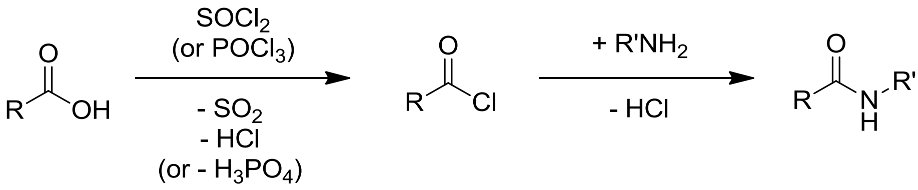
Amide hydrolysis results in the loss of the stabilising carboxamide resonance. As such the amide product is thermodynamically favoured over the free reactants [Ulijn 2002]. Consequently the equilibrium position of the amidation (denoted as $\mathbf{K_2}$ in Scheme 3.1) is not troublesome, the problem lies with the salt formation ($\mathbf{K_1}$). Instead of negotiating the thermodynamic pitfall of the ammonium carboxylate salt, the carboxylic acid reactant tends to be either activated, or a dehydrating reagent used to form an acid anhydride *in situ* for high yielding, but somewhat wasteful procedures (Table 3.1). The E-factor and associated metrics all indicate the flaws of these methodologies [Comerford 2009]. Because of the prevalence of these techniques, the literature is not lacking in comprehensive accounts appraising developments in carbodiimide and other coupling agent technologies [Han 2004, Mikolajczyk 1981, Valeur 2009]. Activation to form an intermediate acid chloride is a staple of discovery and medicinal level chemistry (*i.e.* the milligram scale synthesis of pharmaceutical drug candidates), which should not escape scrutiny for it also requires stoichiometric auxiliaries, in turn creating harmful by-products.

Macmillan *et al.* recently studied the rate of amide coupling assisted by various coupling agents in order to replace what they had identified as an overdependence on the solvents DMF and DCM [Macmillan 2013]. Such is the velocity of amide coupling in the presence of HATU, COMU, *etc.* that of the solvents investigated only *t*-butyl methyl ether and cyclopentyl methyl ether of their solvent set were poor substitutes. Ethyl acetate, 2-MeTHF, dimethyl carbonate, and often 2-propanol were suitable across a wide range of substrates. Any solvent effect is difficult to identify with a lack of kinetic variation across the data set.

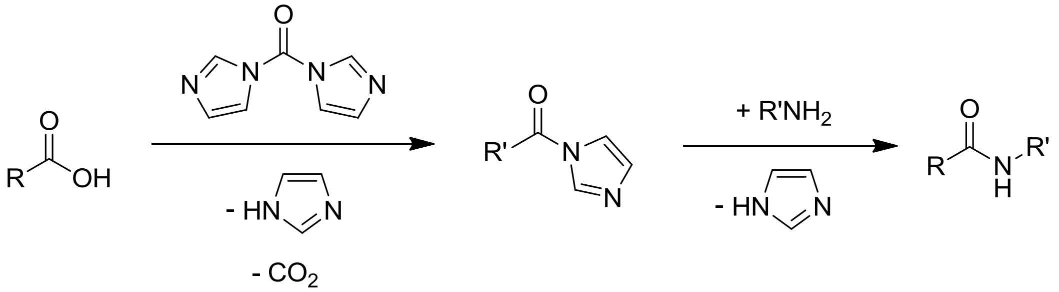
Current practice regarding coupling agents in amidation chemistry was considered as a suitable target for the work of the ACSGCI Pharmaceutical Roundtable [Constable 2007b].

Table 3.1 A selection of activating techniques for promoting amidation.

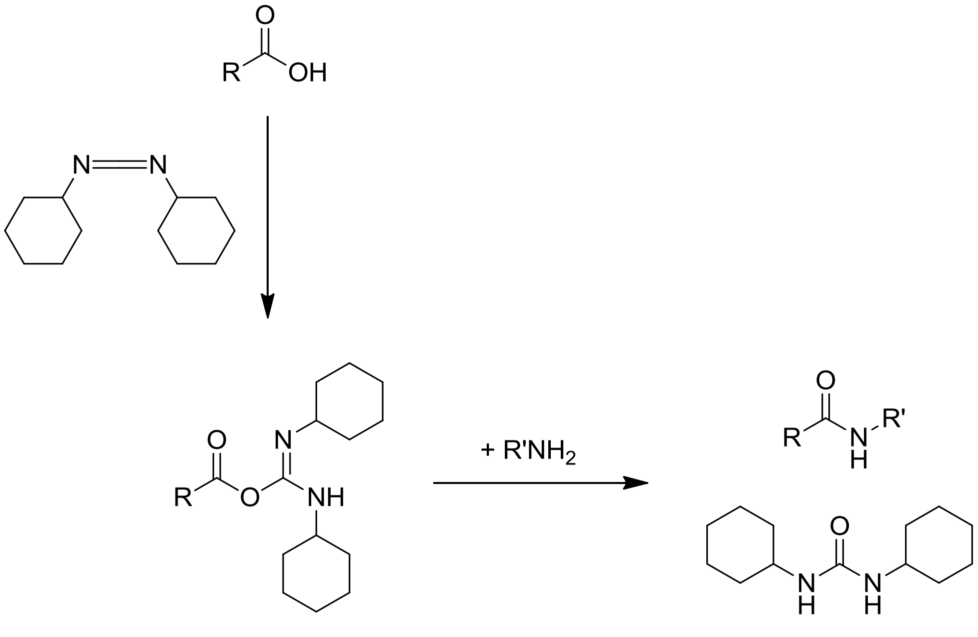
1. Two step acid activation [Comerford 2009, Montalbetti 2005].



2. *In situ* acid activation by CDI [Vaidyanathan 2004].



3. Acid dehydration DCC [Comerford 2009, Montalbetti 2005].



The committee decided that amide formation is the chief priority in the greening of pharmaceutical manufacturing and medicinal chemistry. In response, Pfizer have identified two activators in *N*,*N*′-carbonyldiimidazole (CDI) and isobutyl chloroformate which satisfy their understanding of wide utility, scalability, and greenness (Figure 3.1) [Alfonsi 2008]. It would seem that conversion of carboxylic acids into their respective chlorides is not considered 'green' but catalysis as it currently exists does not have wide utility. The use of traditional stoichiometric coupling agents, essentially dehydrating agents, cannot be considered as green or sustainably scalable. However CDI, the *in situ* activator covered in Table 3.1 and the similar isobutyl chloroformate still appear to present significant waste issues, and the likely use of toxic phosgene in their synthesis is also at odds with the concept of green chemistry. The authors of this assessment acknowledge that their greenness criteria rely very strongly on the context established by alternative methods of amidation. This being the case, any progress made in this regard should be gratefully received by the medicinal chemistry community.

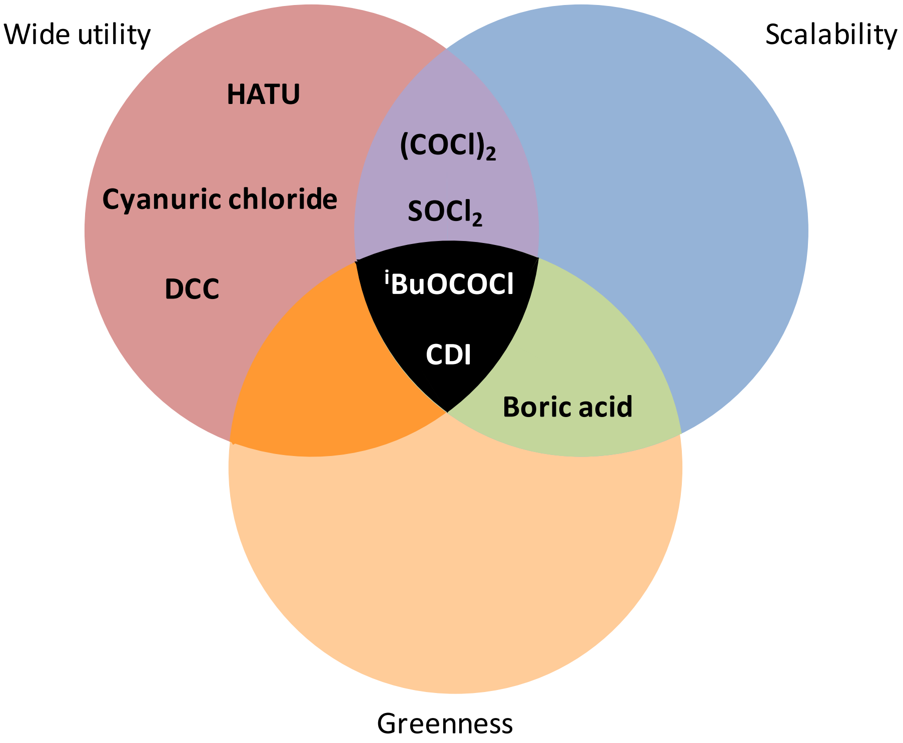
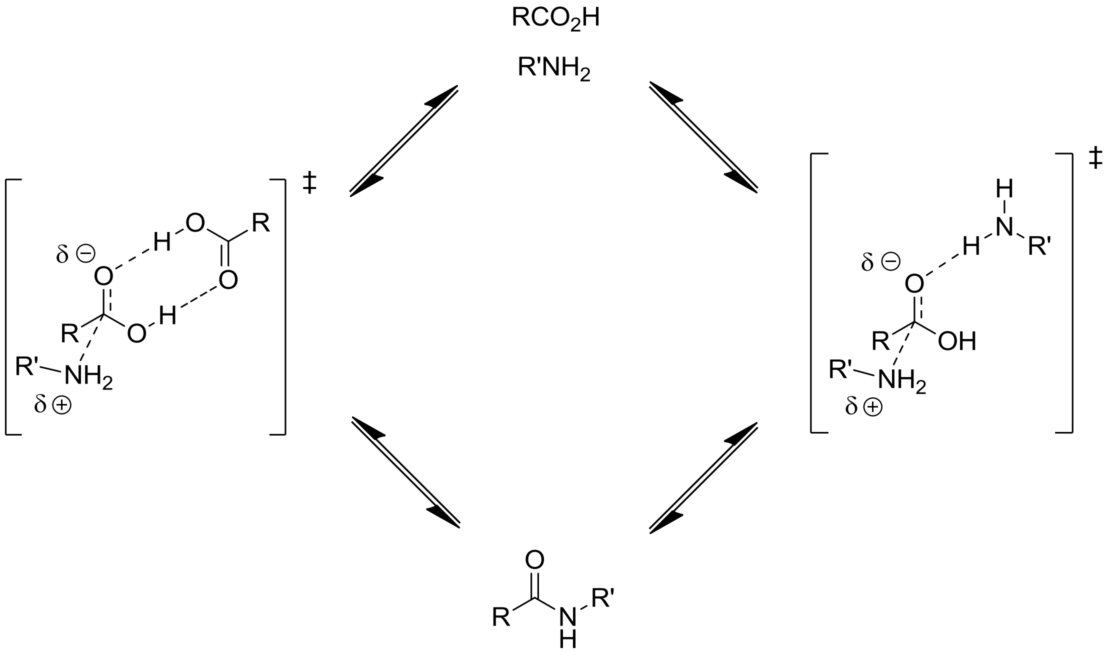


Figure 3.1 A reproduction and simplification of the Pfizer reagent selection Venn diagram for amide coupling protocols [Alfonsi 2008].

Solvent-free amidations: Despite the common use of coupling agents and activation by auxiliary compounds, with the application of high temperatures it is possible to form the desired carboxamide from the precursor salt, although this approach is incompatible with the delicate functionalities common to specialty chemicals [Al-Zoubi 2008]. The reaction between some carboxylic acids and amines can be thermally activated with temperatures above 433 K, without any requirement for an auxiliary solvent [Cossy 1989, Goßen 2009]. The reaction proceeds equally as well with or without molecular sieves, and so the reaction can be performed in the absence of any chemical species that are not incorporated into the product. The sole by-product of water suggests that (high temperatures aside) the procedure is an improvement over the use of coupling agents, with yields typically in excess of 75%. Mass utilisation is expected to be enhanced over some of the higher yielding protocols discussed previously.

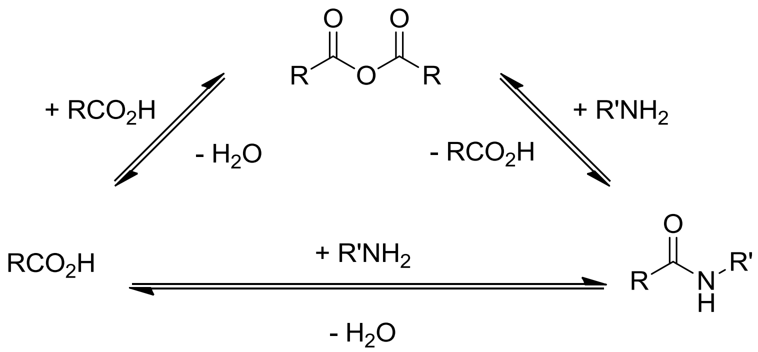
Microwave technology has provided surprisingly good amidation results. The microwave facilitated, direct reaction between acetic acid and a variety of amines begins to occur above a temperature of 363 K [Ferroud 2008]. The ammonium carboxylate salt will absorb microwave energy far better than a molecular organic solvent, and so the combination of an auxiliary and solvent-free microwave assisted synthesis seems the ideal marriage. A detailed exploration into the parameters upon which the microwave assisted reaction is dependent has been conducted, offering a useful insight prior to further studies [Perreux 2002]. The proposed rate determining step is bimolecular, and is equivalent mechanistically to the tetrahedral intermediates of other carbonyl additions, such as those associated with the synthesis or hydrolysis of esters [Smith 2007 page 1402]. The situation is believed to be complicated by the ability of a molecule of either reactant to stabilise the activated complex of the reaction *via* the formation of a hydrogen bonded complex (Scheme 3.2). This might be defined as a self assistance mechanism, a concept that will reappear throughout this chapter. The rate of related acid-catalysed esterifications have been shown to be decelerated by reaction solvents capable of blocking similar interactions when a catalyst is providing this type of hydrogen bonding stabilisation [Wells 2008]. The bimolecular hypothesis has other advocates [Gooßen 2009], as well as opponents [Arnold 2006]. Indeed the experimental data available to date has not resolved this debate, and so it is reasonable to make alternative suggestions based on new eveidence.



Scheme 3.2 A proposed reactant stabilised activated complex of amidation.

Mechanism: The rationale provided by Perreux and co-workers for the presence of the self assistance interactions shown in Scheme 3.2 comes from their own studies in which an excess of either reactant seems to be beneficial to the resultant yield, although this could just be because more substrate is available to react [Perreux 2002]. Complementary microwave studies show that the addition of an auxiliary base (imidazole rather than a classical amine) accelerates rates of amidation [Nezhad 2003, Baldwin 1996]. In the analogous synthesis of esters from an acid anhydride, *N*-methylimidazole performs almost equally as well in the role of an auxiliary base as 2-methylimidazole does [Kingston 1969]. This being the case it appears that the additive does not have to be protic to accelerate the rate of carbonyl additions. Hydrogen bond stabilisation of the reaction components is most likely to be offered by solvent molecules instead, owing to their relative prevalence in the reaction mixture. Furthermore, the observed reaction rate is much slower using triphenylimidazole as a catalyst, suggesting a nucleophilic mode of action, as opposed to an acid-base interaction.

If the possibility of nucleophilic catalysts is extended to include the reactants, this hypothetical exercise would lead either directly to the product if the amine reactant is the nucleophile (*i.e*. no catalysis) or an acid anhydride if a second carboxylic acid is the nucleophile (Scheme 3.3). Following the latter line of pursuit, in his investigations into the catalytic effect of boronic acids on amidations, Arnold proposes a mechanism for the uncatalysed reaction between a carboxylic acid and an amine by supposing that the carboxylic acid dehydrates to form an acid anhydride [Arnold 2006]. The nucleophilic attack of the amine would be so rapid by comparison that the rate equation describing the reaction would be zero order with respect to amine concentration. In buffered aqueous solutions, it is known that a dicarboxylic acid will proceed *via* an anhydride intermediate in its reaction with an amine to give the carboxamide [Higuchi 1963]. However it should be noted that an intramolecular reaction would presumably be more feasible than its intermolecular counterpart. Furthermore, the rate of anhydride formation decelerates as the pH of the solution is increased, dropping to zero in neutral media. At pH 7 the vast majority of

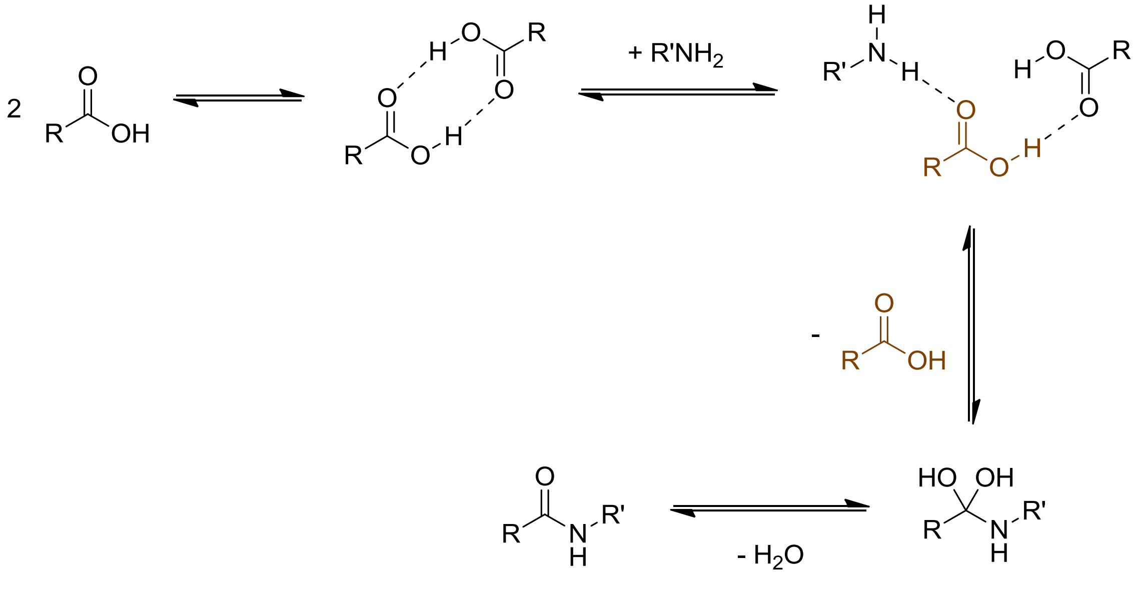


Scheme 3.3 Nucleophilic attack on a carboxylic acid by either reactant in an amidation reaction.

the carboxylic acid may well be deprotonated, resulting in the loss of a viable leaving group. In their investigations into phenyl boronic acid derived amidation catalysts, Al-Zoubi *et al*. were able to eliminate the possibility of an intermediate carboxylic acid anhydride [Al-Zoubi 2008].

The absence of any correlation in Perreux's work between the pKa values of the reactants with the production of the respective carboxamide precedes the inevitable conclusion that the competing equilibrium between the free reactants and their ammonium carboxylate salt is not of great significance in attempts to enhance the reaction of the free carboxylic acid and amine [Perreux 2002]. Instead satisfactory correlations describing the reactivity of different carboxylic acid and amine partners were based on the relative energy differences of the HOMO belonging to the nucleophilic amine and the vacant antibonding $\pi^*_{C=O}$ orbital of the carboxylic acid [Perreux 2002]. A bimolecular mechanism involving both reactants is therefore implicated, certainly proceeding *via* a tetrahedral intermediate.

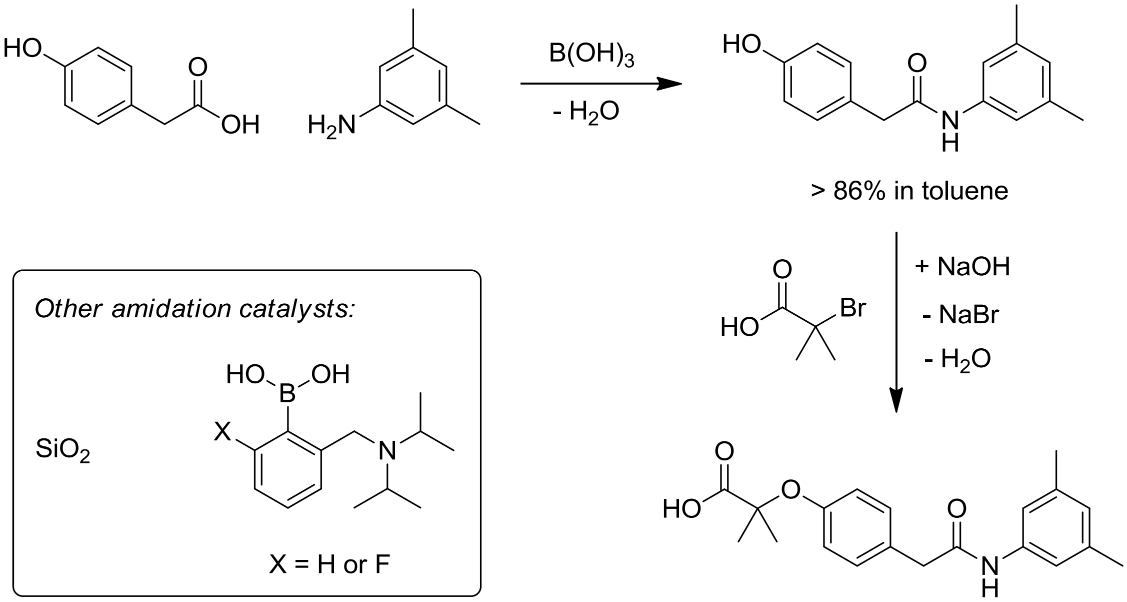
Whiting and co-workers have recently reported the results of a jointly computational and calorimetric study of the uncatalysed reaction of carboxylic acids and amines, which can be used to summarise the state of the art built upon the preceding mechanistic proposals [Charville 2011]. This supersedes the suggestion of an intermediate acid anhydride species proposed a few years earlier by this same group [Arnold 2006]. With new evidence they now conclude that acid catalysis and base catalysis, feasibly provided by the reactants themselves, does not occur [Charville 2011]. More importantly the zwitterionic intermediate created by the nucleophilic attack of the amine upon the carboxylic acid is unstable according to computational calculations. The remaining intermediate possibility is a neutral intermediate that must be accessed without passing through its zwitterionic analogue. Stabilisation by carboxylic dimerisation (not



Scheme 3.4 The hydrogen bond assisted amidation mechanism.

protonation) is called upon to avoid the formation of charge (Scheme 3.4). This proposal is a more specific version of the self assisted mechanism, beginning with a carboxylic acid dimer. Such a mechanism helps to explain the effect of surplus carboxylic acid on the rate and order of the reaction, but not the influence of excess amine as described in Scheme 3.2.

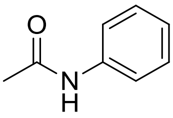
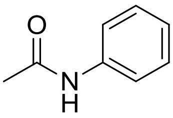
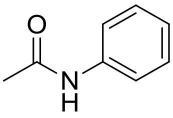
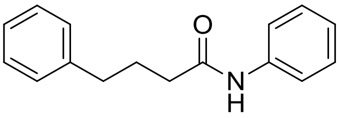
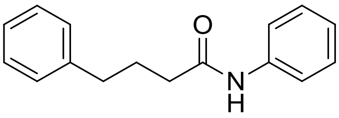
Catalysis: The mechanism should be used to address the fundamental problem that the direct reaction between a carboxylic acid and an amine is slow. One response to the slow rates and high temperatures required in the absence of coupling agents is the use of catalysis. Unfortunately the development of an efficient catalyst has been somewhat elusive. Boric acid has been demonstrated as catalysing the amide forming step in the synthesis of Efaproxiral, once tested as a potential treatment for brain metastasis (Scheme 3.5) [Anderson 2006]. A cheap but toxic chemical, boric acid is usually let down by its modest catalyst effect. Attempts to build upon the activity of boric acid have resulted in quite elaborate phenylboronic acid derivatives, the benefits of which must be weighted against their arduous preparation [Arnold 2006, Arnold 2008, Georgiou 2009]. Heterogeneous catalysts have also been successfully developed. The use of activated silica as a recyclable catalyst for example offers many advantages over the boronic acid derivatives in terms of greenness, as the associated metrics have shown [Comerford 2009].



Scheme 3.5 Catalysis capable of enhancing amidation reactions, demonstrated in the synthesis of Efaproxiral.

Aniline is often chosen as a relatively unreactive reagent to prove the worth of amidation catalysts. Because of its poor nucleophilicity, uncatalysed amidation reactions have been shown to be negligible over the course of a 24 hour reaction (0-6% conversion when aniline is reacted with 4-phenylbutanoic acid for example) [Comerford 2009]. High temperatures seem necessary for the heterogeneous catalysis of aniline acylation with carboxylic acids (Table 3.2). Acetylations are often conducted with a large excess of acetic acid in the dual role of reactant and solvent. Heterogeneous zeolite and clay catalysts can coerce high amide productivity under this type of reaction conditions [Choudary 2001, Narender 2000]. Acetic acid is not a terribly hazardous solvent, performing admirably in the GSK solvent selection guide [Henderson 2011], and has an established fermentative bio-synthesis [Fukaya 1992, Yamada 2008]. When also considering the low price of acetic acid there seems to be little need for further scrutiny of this procedure, aside perhaps from the high temperatures currently required. An acidified carbon (Starbon-400-SO_3H) appears to have wide utility as a heterogeneous catalyst [Luque 2009]. The catalytic effect exerted by activated silica is also impressive, especially given that unlike the other transformations covered in Table 3.2 equimolar quantities of each reactant were used [Comerford 2009].

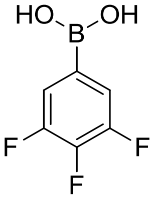
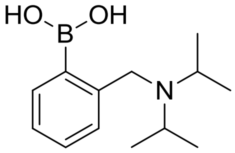
Table 3.2 Heterogeneously catalysed methodologies for the synthesis of carboxamides.

	Catalyst[a]	Product and conditions	Yield
1.	HY zeolite (0.15 g/mmol)	 6 hr, 389 K in excess acetic acid [Narender 2000].	99%
2.	Iron-exchanged montmorillonite (0.002 g/mmol)	 3 hr, 389 K in excess acetic acid [Choudary 2001].	98%
3.	Starbon-400-SO_3H (0.05 g/mmol)	 10 min, 403 K, neat, µW [Luque 2009].	87%
4.	Starbon-400-SO_3H (0.05 g/mmol)	 10 min, 403 K, neat, µW [Luque 2009].	92%
5.	K60 activated silica (0.05 g/mmol)	 24 hr, 383 K in toluene [Comerford 2009].	74%

[a]Loading relative to amine given in parentheses.

What is important to the fine chemical industries is that the high yielding synthesis of an amide from two high value (and often solid) reactants can be guaranteed. Therefore large reagent excesses are not feasible and an auxiliary solvent is required. Homogeneous catalysis often provides high yields but the conditions are no less harsh than in heterogeneously catalysed reactions (Table 3.3). For the synthesis of *N*-benzyl-4-phenylbutanamide, it seems that non-renewable aromatic solvents heated to reflux are the favoured reaction medium, highlighting the importance of solvent selection in favour of greener alternatives. Considering the solvent orientated context of this work, it is intriguing to note that Georgiou remarks that fluorobenzene offers some advantages over toluene regarding catalyst stability [Georgiou 2009]. The nature of the catalyst adds another layer of complexity onto the solvent selection dilemma. The bottom line is this: if the reactants are of high value, then maximal conversion to the amide product will undoubtedly be pursued for economic purposes in preference to E-factor minimisation. This means that these catalysed methodologies, be it homogenous or heterogeneous in nature, will have to compete with the established coupling agents and activators. The boron based catalysts

Table 3.3 Homogeneously catalysed methodologies for the synthesis of *N*-benzyl-4-phenylbutanamide.

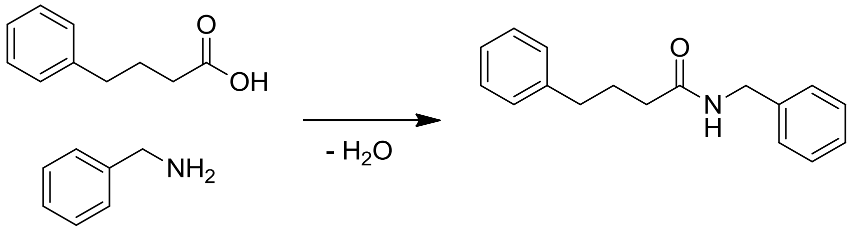
	Catalyst[a]	Conditions	Yield
1.	$B(OH)_3$ (1 mol%)	16 hr, 383 K in toluene [Tang 2005].	91%
2.	 3,4,5-Trifluoro-benzene boronic acid (1 mol%)	18 hr, 383 K in toluene [Ishihara 1996].	96%
3.	 *o*-*N*,*N*-diisopropyl-benzylaminoboronic acid (10 mol%)	24 hr, 358 K in fluorobenzene [Arnold 2008].	68%

[a]Loading relative to amine.

will often provide high enough amide yields, but the reaction duration is too long and the temperatures too high when compared to the rapid room temperature, coupling agent facilitated procedures that are still commonplace.

3.2 Amidation results and discussion

Model reaction: The focus of this work will be an examination of the solvent effects and mechanism of an uncatalysed amidation, without the complications of coupling agents. Ultimately the goal is to deduce an effective yet renewable solvent. The reaction between 4-phenylbutanoic acid and benzylamine is known to proceed without the use of any chemical auxiliaries, having already been established as a model reaction in previous work (Scheme 3.6) [Arnold 2006]. The pairing of these two reactants in toluene does not result in the precipitation of their ammonium carboxylate salt. This is important for accurately measuring the rate of reaction. Instead, mixing 4-phenylbutanoic acid and benzylamine in toluene has been shown by ^{1}H-NMR spectroscopy to create a 2:1 molar ratio of the expected salt (in solution) and a hydrogen bonded pair of the solutes [Charville 2011]. Using a carboxylic acid with a lower pKa value would result in a greater amount of salt formation and inevitably some, if not complete, precipitation. The order of the reaction will be clarified and the mechanism confirmed by observing solvent effects. By contributing this study, practitioners of amidation chemistries may take it upon themselves to utilise its findings in combination with catalysts of their own invention, so that the solvent may be complimentary yet also environmentally benign.



Scheme 3.6 Amidation to give *N*-benzyl-4-phenylbutanamide.

Experimental procedure: As a benchmark, Arnold observed a little less than 60% conversion in refluxing toluene for the uncatalysed reaction of 4-phenylbutanoic acid and benzylamine within approximately 30 minutes [Arnold 2006]. A reaction temperature of 373 K was chosen based on this existing precedent. Each experiment consisted of preheating 4-phenylbutanoic acid in the chosen solvent, to which benzylamine was added in a single dose. This marked the start of the reaction. For the determination of reaction kinetics a small aliquot of the reaction mixture was diluted in deuterated chloroform at selected intervals during the course of the reaction. Rate constants were calculated from the ratio of the ^{1}H-NMR signal integrals corresponding to the N-CH_2 moiety of the product, *N*-benzyl-4-phenylbutanamide, and the reactant, benzylamine (Figure 3.2). The spectrum in Figure 3.2 shows that the progress of the reaction can be followed without interference from other signals. These ratios can be converted into a rate constant for the reaction (Equation 1.13 and Equation 1.15). The ratio of NMR signal intensities were found to accurately represent the conversion when compared to standards of benzylamine and *N*-benzyl-4-phenylbutanamide in solution, as presented in the appendix (Figure 8.1).

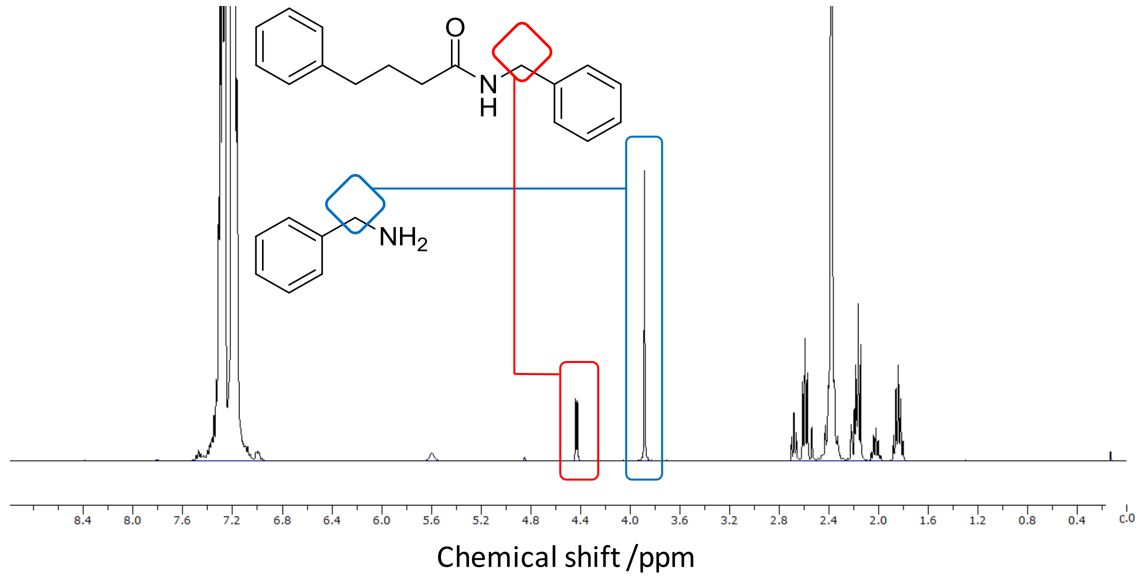


Figure 3.2 An example of a ^{1}H-NMR spectrum showing the partially complete model amidation in toluene. Full solvent signals are not shown to improve resolution of the solute signals.

Determining the reaction order: Prior to determining rate constants in a greater variety of solvents, the order of the reaction was confirmed (at least in toluene) as being second order. The actual order with respect to the individual reactants is slightly larger than unity, measured at 1.1 for both amine and carboxylic acid. This result was obtained by calculating the gradient of a graph plotting the natural logarithm of the initial rate of reaction against the natural logarithm of reactant concentration (Figure 3.3) [Moore 1981 page 65, Pilling 1995 page 13]. The initial concentration of one reactant was varied while the other was kept constant across a series of kinetic experiments. A modest excess of amine created pseudo-first order conditions. With the ammonium carboxylate salt present in solution, a small excess of either reactant becomes very significant relative to the amount of the other free reactant able to exist at any given time.

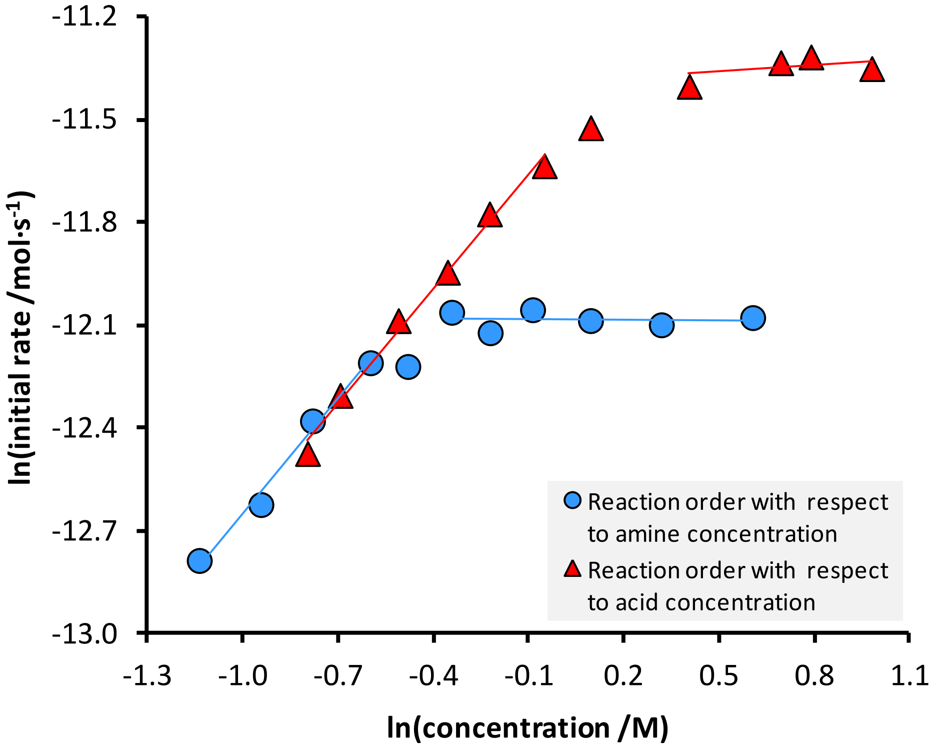


Figure 3.3 Amidation reaction order determination in toluene.

The reaction order of slightly greater than unity with respect to either reactant lends some credence to the Perreux self assistance mechanism (Scheme 3.2). While the mechanism is, by and large, of the obvious bimolecular variety the interaction of a third reactant molecule has increased the experimentally determined reaction order coefficients slightly above unity (Scheme 3.4). An excess of benzylamine results in a reaction that is zero order in terms of the amine, not inconsistent with the acid anhydride mechanism [Arnold 2006]. But why this mechanism should be active only when virtually all the acid is deprotonated by excess amine is counterintuitive. While a modest excess of benzylamine removes rate dependence dramatically, the change from first to zero order kinetics with respect to 4-phenylbutanoic acid is much less sensitive, occurring gradually. Once again invoking the Perreux self assistance mechanism, it is possible that the acid reactant is a superior chelator to the amine, although this is not borne out of the observed reaction orders which are the same. The acid is probably able to form the same sort of dimer with the activated complex that results in the perceived polarity of acetic acid being lower than expected on the scale of relative permittivity [Chocholoušová 2003]. Although just conjecture at this time, this may be the cause of the superior maximum reaction rate ceiling achieved in excess 4-phenylbutanoic acid compared to excess benzylamine conditions (Figure 3.3).

Solvent effects to describe the kinetics of amidation: The concentration of 4-phenylbutanoic acid and benzylamine chosen for the following solvent effect study was the upper end of the bimolecular mechanism region (Figure 3.3). Therefore the reaction was performed under conditions in which the kinetics can be described as first order with respect to each reactant, and second order overall. The solvent study at the temperature of 373 K permitted a useable range of reaction rate constants to be derived, but restricted the number of solvents available to study. A result of this was a smaller range of solvent polarities to explore and the solvent set does show some equivalence in the values of **β** and **π*** (Figure 3.4). In hindsight greater diversity might have been provided with a high boiling dialkyl ether solvent such as dibutyl ether. Hydrogen bond donating solvents were not included in the solvent set and so all solvents had **α** values of zero. This is because C-H acids such as acetonitrile and chloroform are gases at the reaction temperature, and stronger hydrogen bond donators, specifically alcohols and carboxylic acids, will result in competitive reactions to give esters and alternative amides respectively. Chlorobenzene, cyclohexanone, DMF, DMSO, 1,4-dioxane, *n*-octane, toluene, *p*-xylene formed the initial solvent screen.

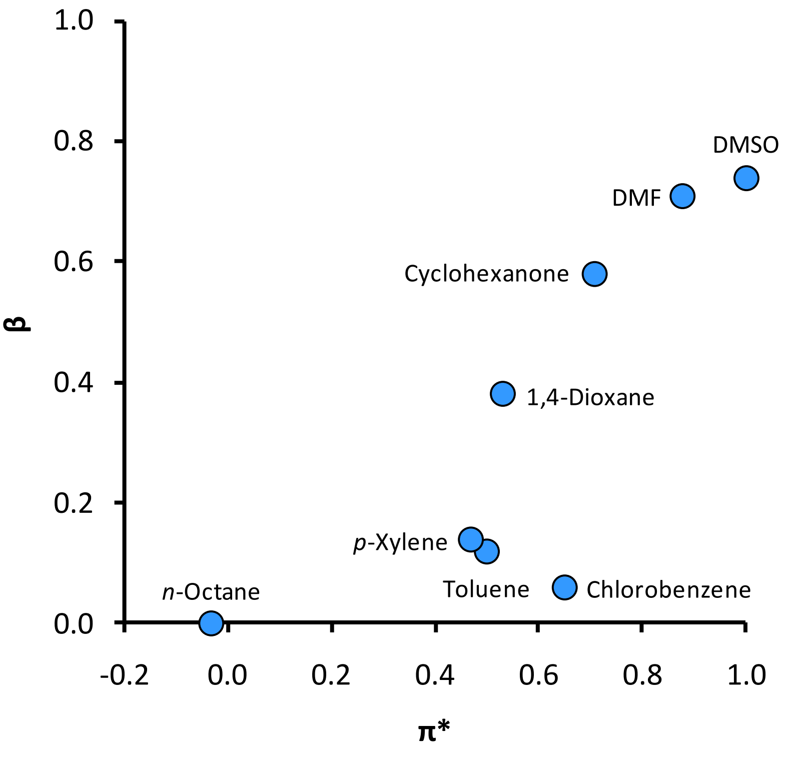


Figure 3.4 A polarity map of solvents included in the initial screening of the model amidation reaction.

The reaction progressed at a measurable rate in seven of the original eight solvents (Table 3.4). Reactions attempted in *n*-octane failed to occur. At temperatures above 313 K 4-phenylbutanoic acid becomes soluble in *n*-octane, a necessary prerequisite, but addition of benzylamine immediately resulted in the white precipitate of the ammonium carboxylate salt of the two reactants. The heterogeneous reaction mixture was not suitable for study, and the reaction abandoned. The fastest rates of reaction occurred in aromatic solvents. The possible formation of an imine adduct (a secondary ketimine to be precise) between cyclohexanone and benzylamine was suggested by the red colour of the reaction mixture rather than the pale yellow achieved in other solvents. This did not appear to disrupt the progression of the reaction (as followed by ^{1}H-NMR spectroscopy) but the solvent was withdrawn from further studies anyway in case of misleading results.

Table 3.4 Amidation rate constants in different solvents and the polarity of those solvents.

Solvent	ln(**k**)	**α**	**β**	**π***
Chlorobenzene	-10.51	0.00	0.06	0.65
Cyclohexanone	-11.06	0.00	0.58	0.71
DMF	-11.30	0.00	0.71	0.88
DMSO	-11.42	0.00	0.74	1.00
1,4-Dioxane	-10.97	0.00	0.38	0.52
Toluene	-10.65	0.00	0.12	0.50
p-Xylene	-10.63	0.00	0.14	0.47

Examination of the data suggests that chlorobenzene not only accelerates the reaction beyond the rate obtained in toluene, but is actually better for the environment (Figure 3.5). Chorobenzene is not without its flaws however and appears on restricted chemical lists just as toluene does [SubsPort 2013]. The HAS classification of chlorobenzene is shaped by a toxicity (rat oral LD_{50}) on a par with toluene. It is still toxic to aquatic organisms and bioaccumulating [ECOTOX 2013], but the reduced flammability of chlorobenzene makes the biggest difference in a comparison with toluene. *p*-Xylene has the best HAS profile of the aromatic solvents, which is marginally less flammable and less toxic than toluene. 1,4-Dioxane and DMF have serious safety and health issues respectively, but like cyclohexanone and DMSO they were not very good at accelerating the rate of reaction anyway. The solvent set does not feature any high performance and green solvents, providing the justification for an alternative bio-based solvent.

Natural logarithms of the second order rate constants obtained from the initial screening could be correlated to the hydrogen bond accepting ability of the solvents (**β**):

Equation 3.1

$$\ln(\mathbf{k}) = -10.47 - 1.19\boldsymbol{\beta} \qquad R^2 = 0.974$$

The coefficient of determination (R^2 value) rises to 0.991 when cyclohexane is removed from the

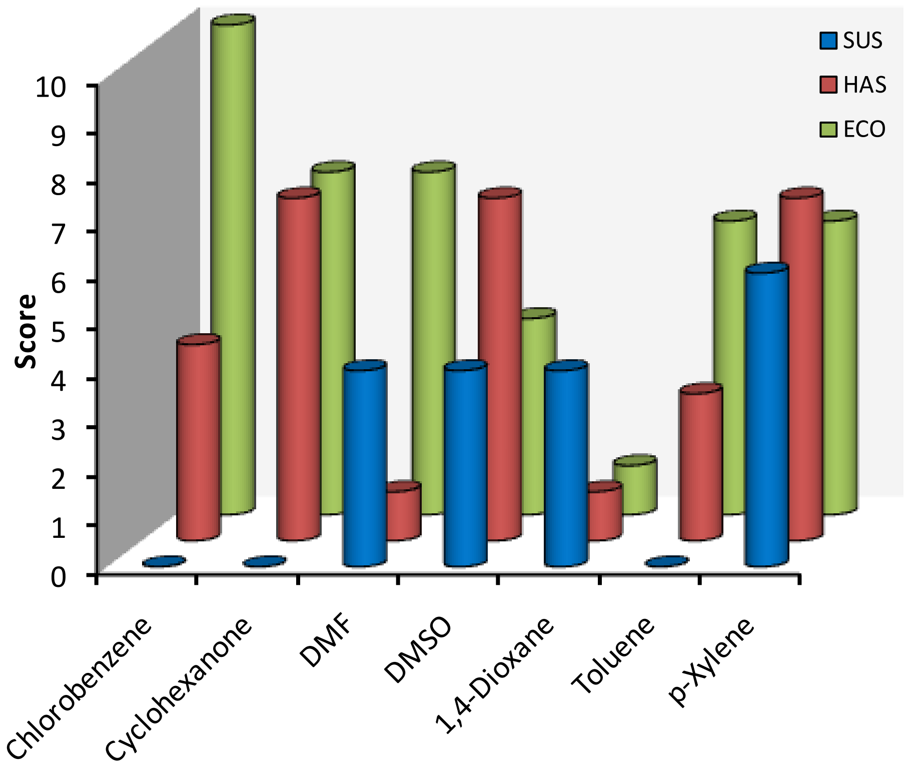


Figure 3.5 The environmental, health and safety of amidation solvents.

solvent set. Refer to the appendix for more details (Table 8.9). The inverse proportionality of this relationship indicates that low polarity solvents enhance the rate of reaction (Figure 3.6). Other solvent polarity parameters were found to be insignificant, or in the case of **α** for example, not enough variation in the data set was present to construct a meaningful relationship. This observation is convincing from the strength of the data fit but not in terms of its consequences. The reaction progresses *via* an activated complex in which the partial forming and breaking of bonds is polarising. As such polar solvents are expected to stabilise the activated complex and accelerate the reaction according to the same principles established for S_N2 type reaction mechanisms. However the opposite effect is observed. To understand the role of the solvent, the observed rate constants must be separated into their enthalpic and entropic contributions.

Determination of the activation parameters of amidation: It is now established that at a temperature of 373 K, poor hydrogen bond accepting solvents provide a superior environment for amidation [Clark 2012]. The reactions were repeated at various temperatures and the resulting rate constants interpreted with the linear form of the Eyring equation:

Equation 3.2

$$\ln\left(\frac{\mathbf{k}}{\mathrm{T}}\right) = \ln\left(\frac{\mathrm{k_B}}{\mathrm{h}}\right) - \frac{\Delta \mathrm{H}^{\ddagger}}{\mathrm{RT}} + \frac{\Delta \mathrm{S}^{\ddagger}}{\mathrm{R}}$$

In graphical form, the Erying equation reveals a temperature at which all solvents are predicted to provide an equally suitable reaction medium (Figure 3.7). However this iso-kinetic

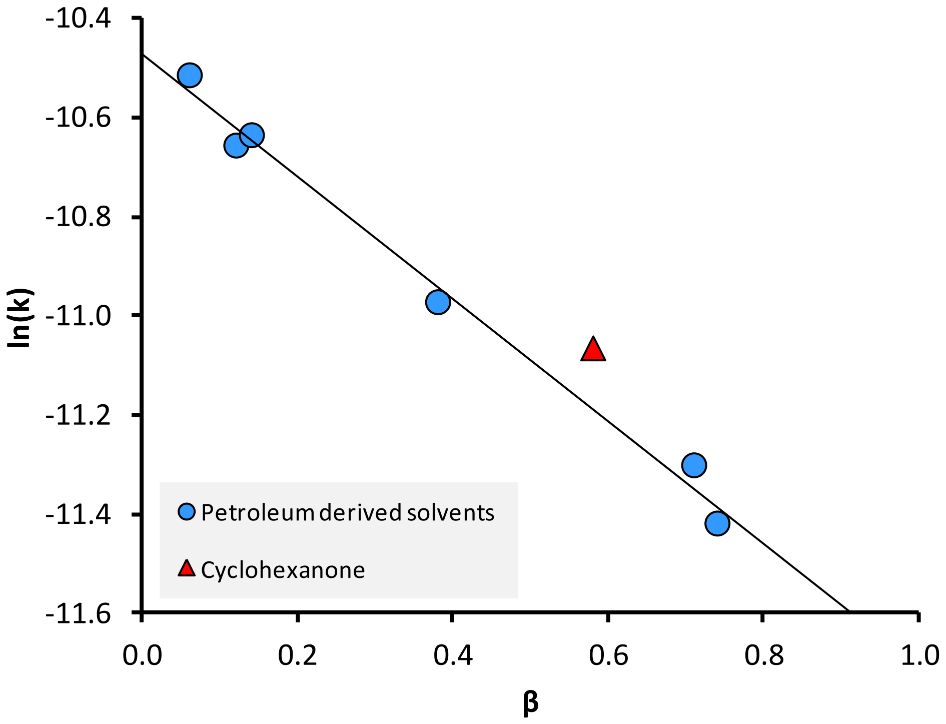


Figure 3.6 The LSER correlating the rate of the model amidation reaction with the hydrogen bond accepting ability of the solvent.

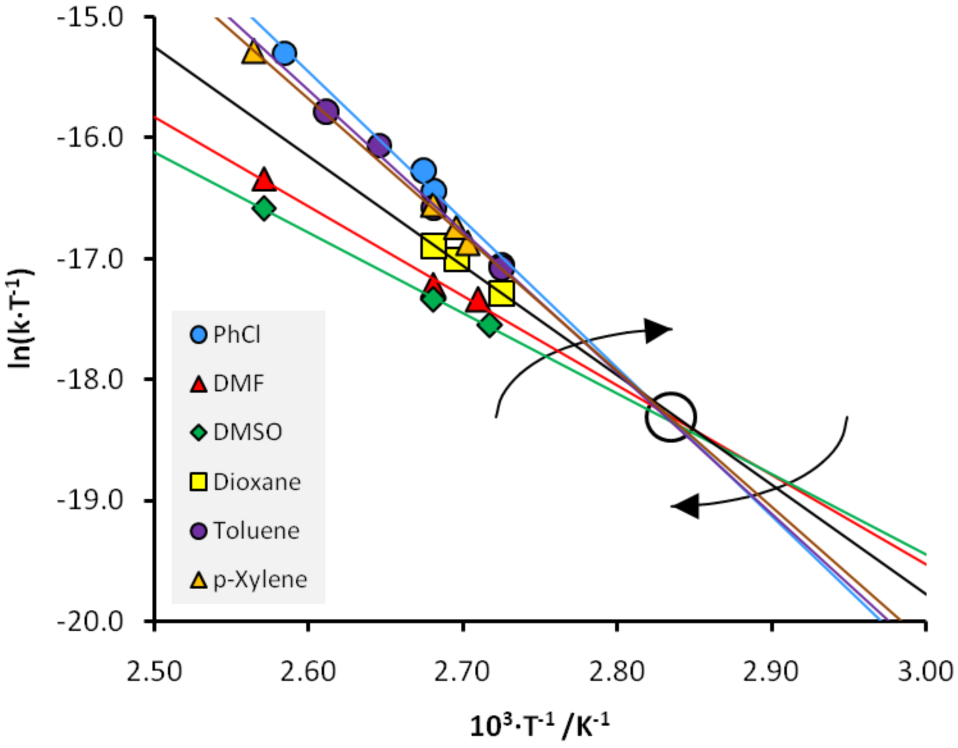
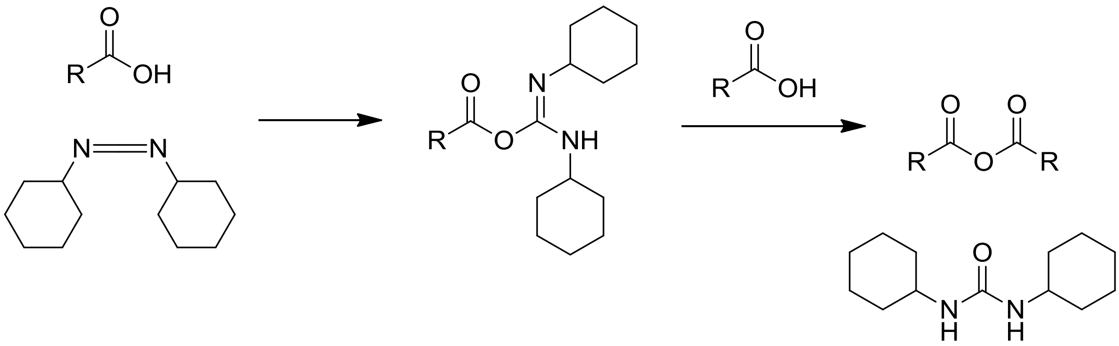


Figure 3.7 A multi-solvent Eyring relationship for amidation including a predicted crossover point of reaction rates at an iso-kinetic temperature.

temperature corresponds to a point at which the uncatalysed reaction is unmeasurably slow to complete (approximately 353 K). This means that realistically the point of coalescence can only be extrapolated from data obtained at higher temperatures. This data itself is somewhat limited by the boiling points of the solvents but it gives an indication of the solvent effect in operation.

The existence of the inversion point at the iso-kinetic temperature suggests that this case study may not be directly relevant to room temperature amidations which proceed assisted by a coupling agent. Reassuringly a reaction case study on the dehydration of carboxylic acids to their respective anhydride (using DCC as a coupling agent at 303 K) notes that the rate of the initial addition of the acid to DCC is inversely proportional to **β**, consistent with this work (Scheme 3.7) [Balcom 1989]. This suggests that all carbonyl additions, unassisted or otherwise, may be kinetically governed by the same empirical solvent effects revealed here.



Scheme 3.7 The synthesis of carboxylic acid anhydrides.

The iso-kinetic temperature demonstrates that enthalpy and entropy are exerting opposing influences. This can be seen in the Gibbs free energy equation that relates the two:

Equation 3.3

$$\Delta G^{\ddagger} = \Delta H^{\ddagger} - T \cdot \Delta S^{\ddagger}$$

The temperature-entropy term of Equation 3.3 becomes smaller than the enthalpy of activation below the trend inversion in Figure 3.7, reversing the solvent effect when the magnitude of the former effect is superseded by the latter. The gradient of the Eyring plots are proportional to the negative of the enthalpy of activation, as deducible from the actual Eyring equation (Equation 3.2). The intercept is determined by the magnitude of the entropy of activation, which becomes more negative with increasing solvent polarity. Together these two parameters give the Gibbs free energy of activation term which in turn is related to ln(**k**) (Equation 3.3). Equation 3.1 suggests that the enthalpy of activation and the entropy of activation will also have a strong dependence on **β**. The enthalpy of activation is indeed inversely proportional to hydrogen bond accepting ability (Figure 3.8). The same is true for the entropy of activation (Figure 3.9):

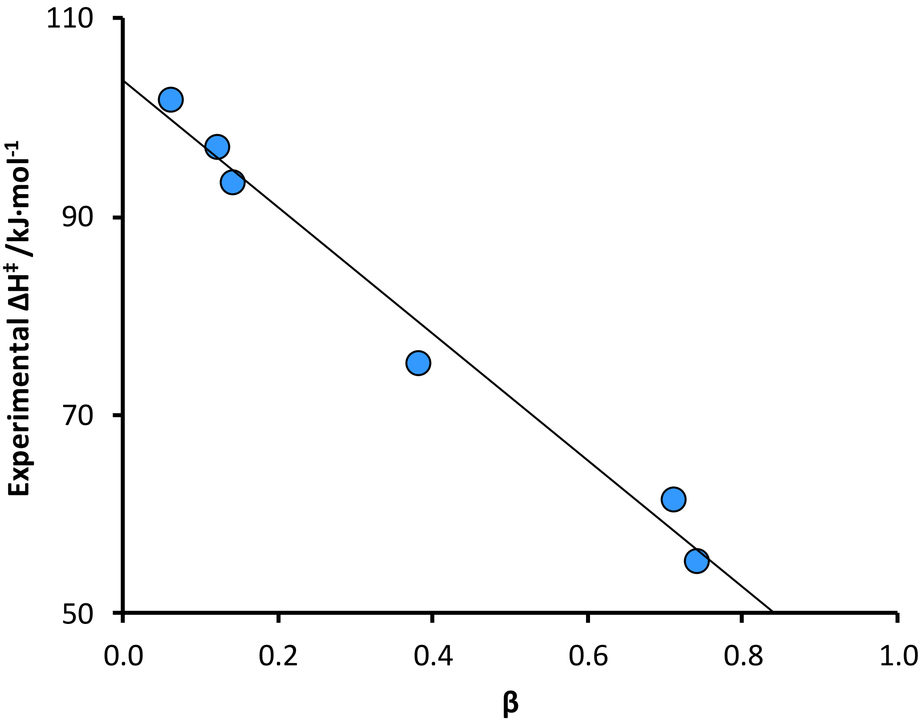


Figure 3.8 The correlation between the enthalpy of activation and solvent hydrogen bond accepting ability (**β**).

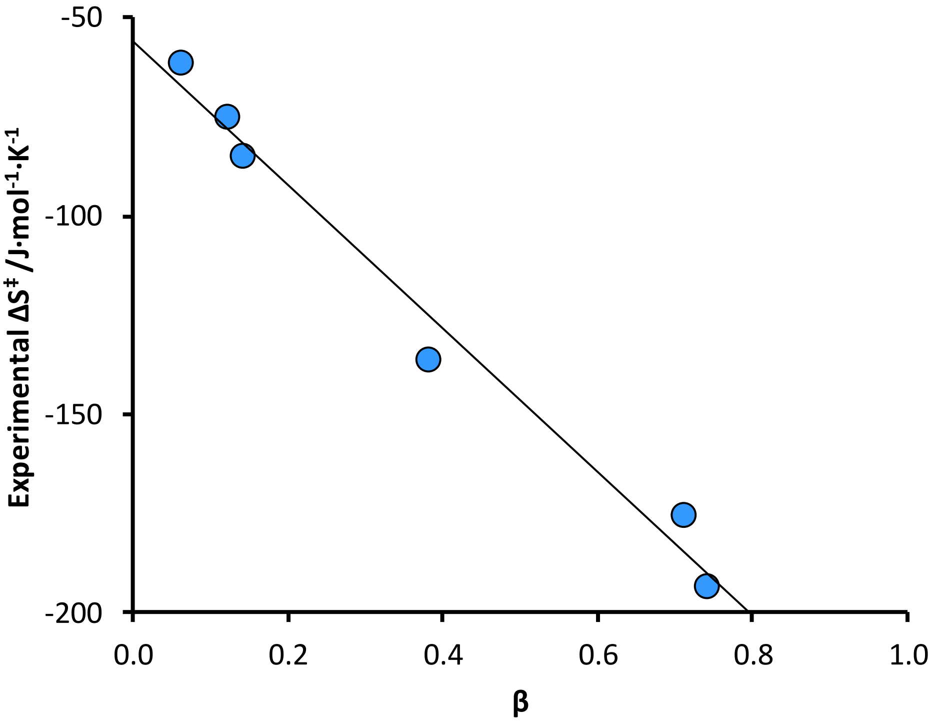


Figure 3.9 The correlation between the entropy of activation and solvent hydrogen bond accepting ability (**β**).

Equation 3.4 $$\Delta H^{\ddagger}/\mathrm{kJmol^{-1}} = 104 - 63.8\boldsymbol{\beta}$$ $R^2 = 0.982$

Equation 3.5 $$\Delta S^{\ddagger}/\mathrm{Jmol^{-1}K^{-1}} = -55.9 - 181\boldsymbol{\beta}$$ $R^2 = 0.982$

Both Equation 3.4 and Equation 3.5 present themselves identically but when combined in Equation 3.3 it is obvious they will exert opposing effects because the entropic term is subtracted from the enthalpy of activation. As solvents with ever increasing polarity (as measured by **β**) are selected to support the amidation, the enthalpy of activation is favourably decreased, as expected on the basis of hydrogen bonding arguments. However at the reaction temperature of 373 K the entropic contribution dominates ($T{\cdot}\Delta S^{\ddagger} > \Delta H^{\ddagger}$), and an increase in **β** further enhances the system order as the reactants go on to form the activated complex. The unfavourable decrease in entropy caused by greater hydrogen bonding in polar solvents is ultimately not beneficial to the kinetics of the reaction above the iso-kinetic temperature.

The considerable effect the solvent exerts on the rate of reaction can be further demonstrated by examining the data in more detail. The enthalpy of activation stands at 97.2 $\mathrm{kJmol^{-1}}$ in toluene but a relatively low 55.4 $\mathrm{kJmol^{-1}}$ in DMSO (Table 3.5) The analogous Arrhenius parameters have also been derived for comparison and show a similar trend:

Equation 3.6 $$\ln(\mathbf{k}) = \ln(\mathrm{A}) - \frac{\mathrm{E_a}}{\mathrm{RT}}$$

Conversely the entropy of activation is -74.7 $\mathrm{Jmol^{-1}K^{-1}}$ in toluene, indicating an increase in the order of the system as the activated complex is formed from the two reactants. A significant entropy change of -193.1 $\mathrm{Jmol^{-1}K^{-1}}$ occurs in DMSO. The enthalpy and entropy terms, in combination at 373 K, create less than a 3 $\mathrm{kJmol^{-1}}$ difference between the Gibbs free energies of activation in these two solvents (Equation 3.3). Although the change in Gibbs free energy is small,

Table 3.5 Activation parameters for the model amidation reaction including the Gibbs free energy at 373 K.

Solvent	$\Delta H^{\ddagger}$ /$\mathrm{kJmol^{-1}}$	$\Delta S^{\ddagger}$ /$\mathrm{Jmol^{-1}K^{-1}}$	$\Delta G^{\ddagger}$ /$\mathrm{kJmol^{-1}}$	ΔE_a /$\mathrm{kJmol^{-1}}$	A /$\mathrm{dm^3mol^{-1}s^{-1}}$
PhCl	101.9	-61.0	124.7	105.1	$1.39*10^{10}$
DMF	61.6	-175.1	126.9	64.7	$1.53*10^{4}$
DMSO	55.4	-193.1	127.4	58.5	$1.76*10^{3}$
1,4-Dioxane	75.4	-135.9	126.0	78.4	$1.67*10^{6}$
Toluene	97.2	-74.7	125.0	100.3	$2.67*10^{9}$
p-Xylene	93.6	-84.5	125.1	96.8	$8.29*10^{8}$

it makes a striking difference to the conversions achievable in dissimilar solvents. This effect, when proportionality between enthalpy and entropy almost cancel out when combined into the Gibbs free energy, is known as enthalpy-entropy compensation [Boots 1989, Liu 2001, Perez-Benito 2013].

The relationship between the enthalpy of activation and the entropy of activation reported here is perfectly linear ($R^2 = 1.000$), a consequence of the iso-kinetic temperature (353 K) which manifests itself as the gradient in this relationship (Figure 3.10). This type of plot is regularly found to be linear, but only valid if a single iso-kinetic temperature point exists such as that shown in Figure 3.7 [Petersen 1964]. Not only are enthalpy and entropy exactly proportional, but furthermore, as a consequence of this Figure 3.8 must be identical in appearance to Figure 3.9 ($R^2 = 0.988$ in both cases). The relationship between the two activation parameters is partly a product of the manner in which they are deduced from a single relationship. Enthalpy is calculated from the gradient of the trends in Figure 3.7 which is quite acceptable. The entropy of activation is obtained by extrapolating a trend to its interception of the y-axis far from the range where the experimental data resides. This treatment is often thought of as being susceptible to error, but this idea has been discredited. It is possible to reconstruct the Eyring equation so that graphically the gradient of the trend corresponds to the entropy and the intercept is proportional to the enthalpy [Lente 2005]. Either way, enthalpy and entropy are derived from the same ln(**k**)

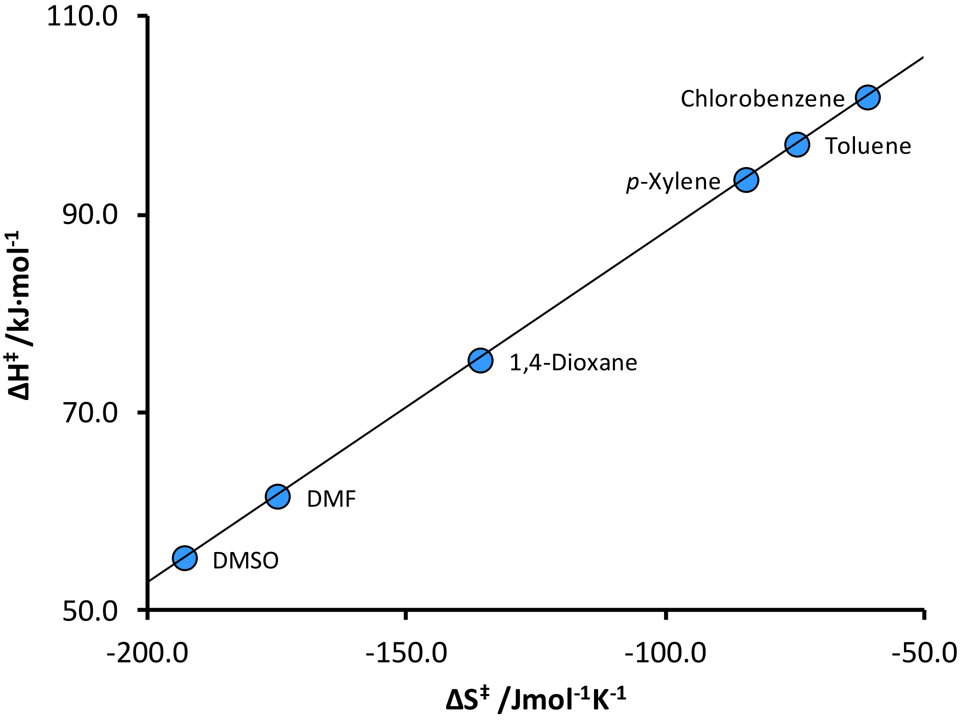


Figure 3.10 The relationship between the activation parameters dictating the kinetics of the model amidation.

measurements (*via* the Eyring equation) and are therefore linked. The experimental errors associated with the measurement of rate constants are unavoidably translated onto both the enthalpy of activation and the entropy of activation, giving rise to the similarity in their presentation.

This does not now mean that the general trends required from this analysis, namely that the hydrogen bond accepting ability of the solvent is proportional to the enthalpy of activation (Equation 3.4) and entropy of activation (Equation 3.5), are flawed. This work has shown a precise iso-kinetic temperature of amidation, and therefore it must be so that enthalpy and entropy are closely related in a well defined manner. In fact, one of the prerequisites for a valid LFER or LSER such as Equation 3.1, Equation 3.4, or Equation 3.5 is that if one of enthalpy or entropy is not constant, then both terms must be proportional to each other [Exner 1964, Exner 1972 page 8, Leffler 1955]. Without this stipulation there is no basis for a relationship between ln(**k**) and **β**, or the variety of other parameters used to construct free energy relationships. The dye responses used to determine **β** are necessarily the result of proportional enthalpy and entropy contributions, not dissimilar to the same hydrogen bonding arguments that describe the mechanism of amide formation.

The mechanism of amide formation: We can now appreciate the influence of the solvent upon this model amidation reaction. For polar solvents, the enthalpic benefit of stabilising the activated complex of the reaction is overridden by the large entropic penalty of arranging the necessary solvent-solute hydrogen bonds at the high temperatures required to see any progress in the reaction. The enthalpy of activation is still relevant and should not be completely dismissed, especially at lower temperatures where hopefully future catalysts will be effective. The following enthalpy diagram summarises the important solvent-solute interactions and also provides a clue as to why entropic effects trump enthalpy at higher temperatures (Figure 3.11). This is not a free energy diagram, which by including entropy would show the opposite trend. For simplicity the ammonium carboxylate salt is ignored in this treatment, after it was found not to be influential in a previous study [Charville 2011]. The self assistance mechanism at the transition state is not shown either, partly for simplicity, partly because the order of the reaction has been shown to be only slightly greater than unity with respect to each reactant. The free reactants will be heavily stabilised by DMSO and other hydrogen bond accepting solvents. These interactions are only strengthened as the activated complex polarises the reactants, which is the cause of the solvent effect governing the enthalpic profile of the reaction. The entropy of activation and free energy of activation also stem from this phenomena. The kinetics of the reaction are of course entropically controlled at 373 K and some indication of why this is so can be seen from the ordered structures of Figure 3.11.

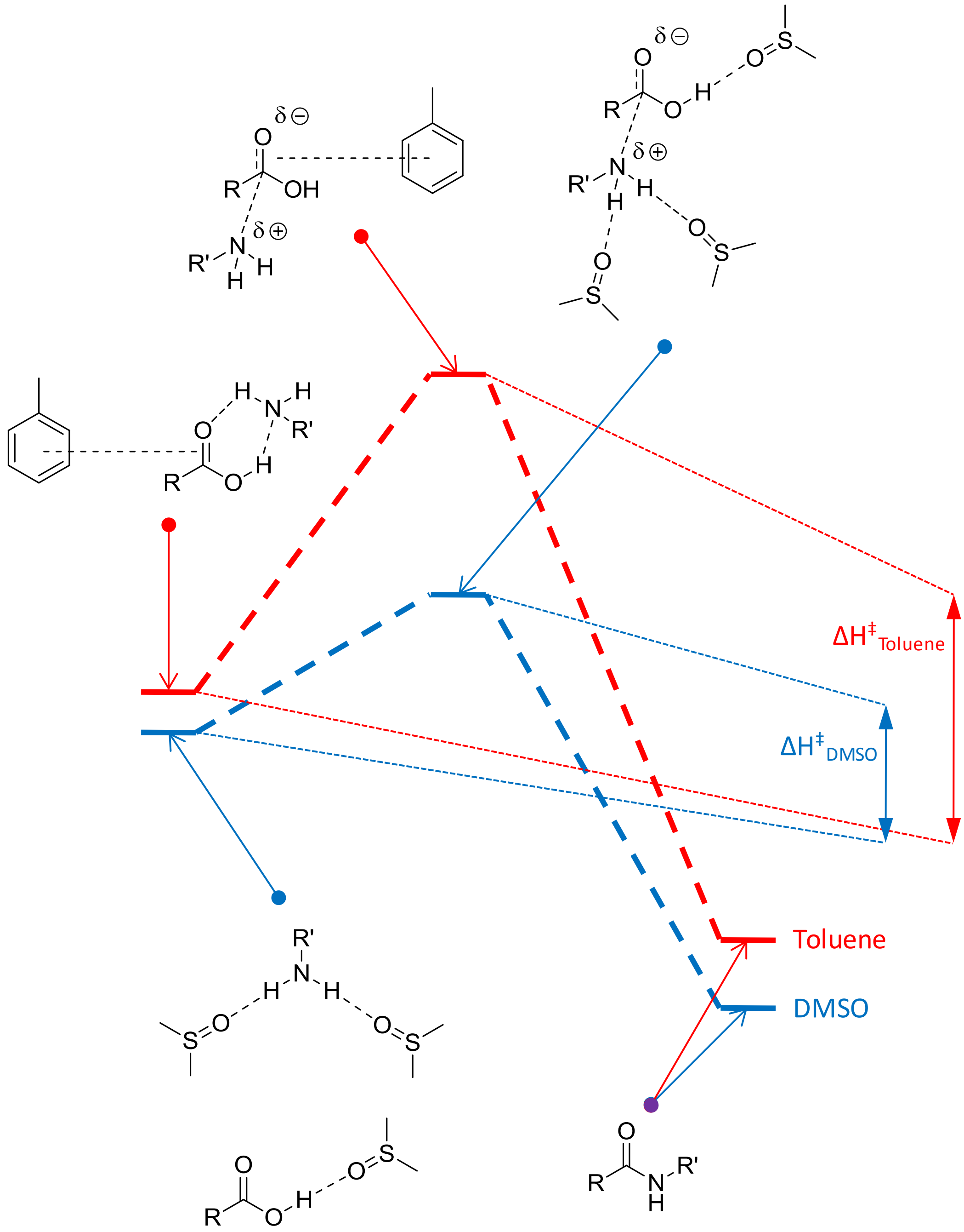
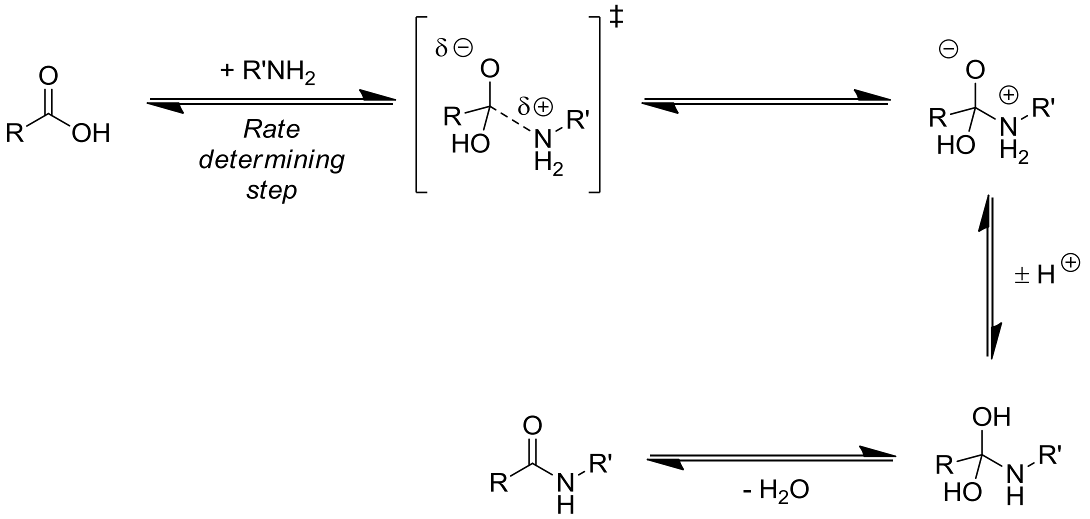


Figure 3.11 The generalised variable enthalpy of amidation illustrated in toluene and DMSO.

Toluene is weakly hydrogen bond accepting, as indicated by its small **β** value. It is quite probable that the free reactants might help themselves to each others' hydrogen bonding functionalities in non-hydrogen bonding solvents like toluene. This could be in the form of a carboxylic acid dimer (Scheme 3.4), or the alternative acid-base partnership indicated in Figure 3.11. This will increase the enthalpic gap between the reactants and the activated complex in non-hydrogen bonding solvents, making the counteractive entropic forces even more impressive. If a carboxylic acid molecule is hydrogen bonded to another reactant, the stabilising hydrogen bond of the self assistance mechanism is installed prior to reaction. Strongly hydrogen bond accepting solvents may disrupt the formation of the hydrogen bond dimer prior to the reaction (as indicated in the comparison between DMSO and toluene in Figure 3.11) but because there is no net loss in hydrogen bonding no energy difference penalty is added to the enthalpy of activation. This conclusion is borne out of experimental observations that indicate the enthalpy of activation is reduced in polar solvents. Finally, their claim that a zwitterionic intermediate species is unstable led Charville *et al*. to rely on a modification of the self assistance mechanism (Scheme 3.2), utilising a trimolecular reaction to avoid the formation of any charged species (Scheme 3.4) [Charville 2011]. Given that the order of the reaction has been established and does not necessitate the presence of a third reactant, it seems more plausible that amidation shares a mechanism with $A_{AC}2$ esterification, proceeding *via* the same intermediates but without acid catalysis (Scheme 3.8) [Smith 2007 page 1402].



Scheme 3.8 The $A_{AC}2$ mechanism of amidation proceeding through a tetrahedral intermediate.

Binary solvents for amidation: The role of the solvent was explored further by using a binary solvent mixture of toluene and DMSO. It is known that Kamlet-Taft solvent polarity parameters do not vary linearly as a function of solvent composition in binary mixtures [Marcus 1994]. It was hoped that **β** would decrease in a binary mixture and accelerate the reaction. The outcome was synergetic but not beneficial if such juxtaposition exists (Figure 3.12). A significant drop in the velocity of the reaction occurred. Starting with a wholly DMSO based system, the addition of the preferable solvent toluene would be expected to increase the reaction rate. Instead a reduction in productivity occurs. At 373 K it has been already been shown that entropic effects dictate the rate of reaction. Of the two solvents, DMSO with its superior polarity (gauged by the Hansen solubility parameters) is assumed to be the preferred candidate to fill the cybotactic region (in other words, the solvation sheath), suggesting DMSO preferentially solvates the solutes. As such we are presented with a scenario in which the inferior solvent in terms of kinetic enhancement but the superior solvent in terms of solvating ability is ordered around the reaction component, further increasing entropic penalties. The result is a drop in reaction rate that is gradually rectified as more toluene is introduced, displacing DMSO due to its numerical advantage.

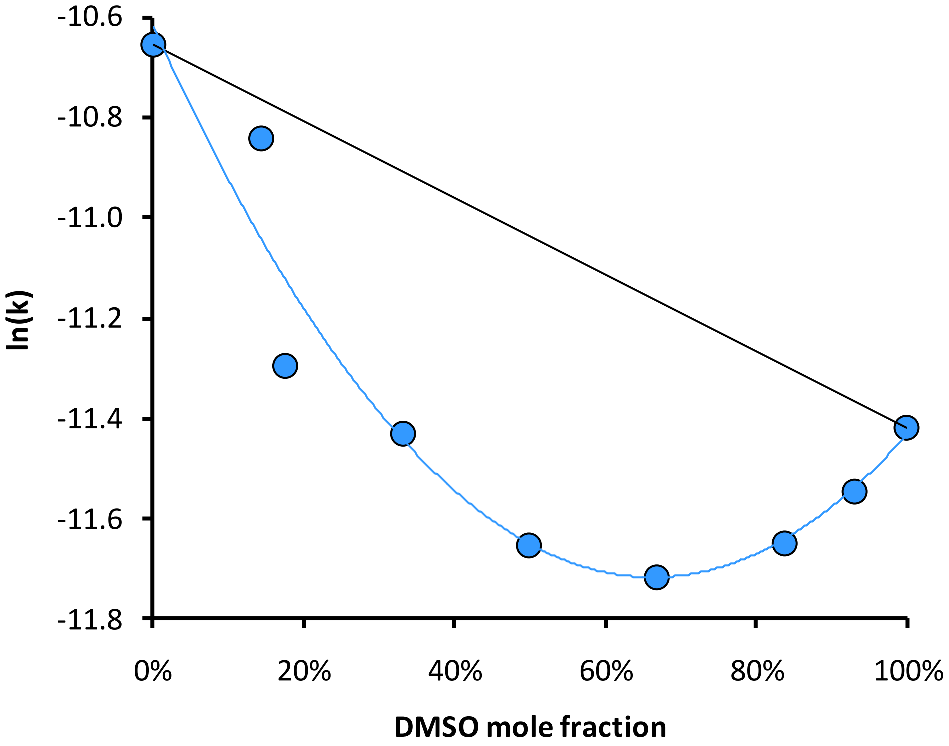
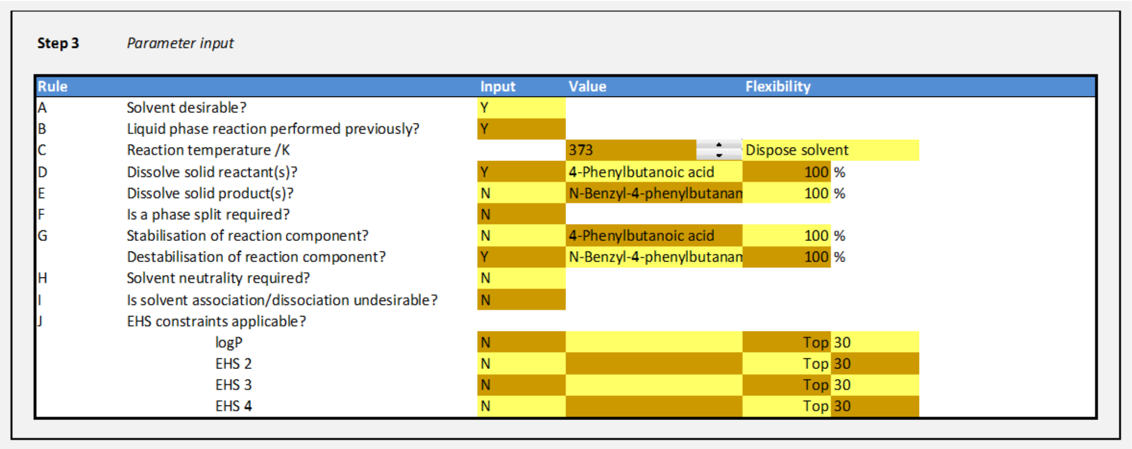


Figure 3.12 The toluene-DMSO binary solvent effect on the rate of amidation.

Solvent selection: Although hydrocarbons are clearly ideal solvents within the narrow scope of rate enhancement, their EHS profile and sustainability are of concern. The GSK solvent selection guide and the SUS-HAS-ECO interpretation of it highlight this. Aside from being non-renewable, most alkane and aromatic solvents are highly flammable and pose environmental issues of bioaccumulation and toxicity [Curzons 1999, Henderson 2011]. The representative alkane *n*-octane was found to be unsuitable because it fails to homogenise the reaction mixture or produce any amide. Aromatic solvents were successful at enhancing reaction productivity but the use of toluene and *p*-xylene presents risks worth avoiding. Mesitylene is marginally less worrying than either toluene or *p*-xylene in terms of flammability and toxicity and is expected to perform equally as well in the model amidation reaction. However it is still a product of the petroleum industry and as such does not satisfy the need to establish bio-based solvents for the long term security of organic synthesis. The other common class of solvent for synthesising amides are the halogenated solvents, which even though they fit the polarity requirements, do not improve upon the aromatic solvents for environmental, health and safety properties. Many are highly regulated or on restricted lists because they are carcinogens [SubsPort 2013].

The solvent selection algorithm was used to verify the application of alternative hydrocarbon solvents. Dissolution of 4-phenylbutanoic acid (rule D) and destabilisation of *N*-benzyl-4-phenylbutanamide (rule G) were required to permit the reaction yet facilitate removal of the product from high boiling solvents by promoting precipitation (Figure 3.13). The reaction temperature was set at 373 K. Limonene, along with toluene, fails to complete the demands of the original algorithm (model A) based on these inputs due to the manner in which the solvent is deemed to be a liquid at the reaction temperature (rule C). *p*-Cymene passes the original solvent selection algorithm model, and along with limonene and toluene also completes the revised solvent selection algorithm (model B1). (*R*)-(+)-Limonene is contained in waste orange peel, but infrequently used as a solvent. However potential solvent applications for limonene and its derivatives are not insignificant. Limonene is a common component in cleaning and degreasing products [Henneberry 2004], and has also been used as a solvent for synthesis. The reactivity of the limonene terminal double bond and its chirality has been exploited in both ring opening polymerisations and the construction of optically active polymers respectively. [Kawagoe 2010, Mathers 2006] The sequential isomerisation and dehydrogenation of limonene to the aromatic compound *p*-cymene provides a route to a more robust solvent [Martin-Luengo 2010]. *p*-Cymene is also naturally occurring in trace amounts [Banthorpe 1972]. The current uses of *p*-cymene include, most recognisably to synthetic chemists, use as a ligand for ruthenium complexes devoted to catalysis [Castarlenas 2006]. Recent cameos as a solvent include dehydrogenation [Marchais-Oberwinkler 2011], decarbonylation reactions [Fessard 2007], and others [D'hooghe 2008, Kelly 1997].

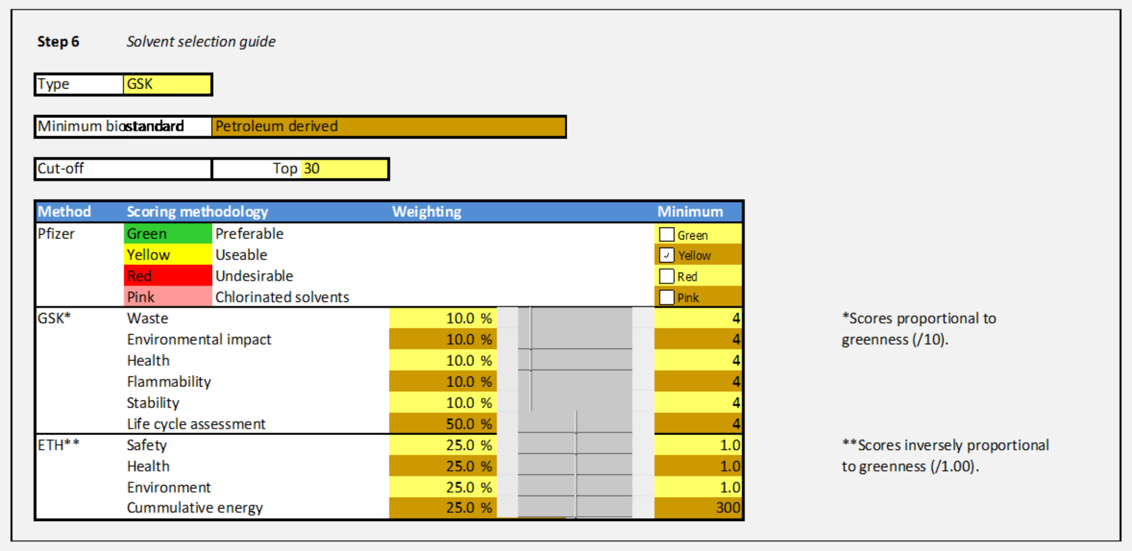


Step 3 *Parameter input*

Rule		Input	Value	Flexibility
A	Solvent desirable?	Y		
B	Liquid phase reaction performed previously?	Y		
C	Reaction temperature /K		373	Dispose solvent
D	Dissolve solid reactant(s)?	Y	4-Phenylbutanoic acid	100 %
E	Dissolve solid product(s)?	N	N-Benzyl-4-phenylbutanan	100 %
F	Is a phase split required?	N		
G	Stabilisation of reaction component?	N	4-Phenylbutanoic acid	100 %
	Destabilisation of reaction component?	Y	N-Benzyl-4-phenylbutanan	100 %
H	Solvent neutrality required?	N		
I	Is solvent association/dissociation undesirable?	N		
J	EHS constraints applicable?			
	logP	N		Top 30
	EHS 2	N		Top 30
	EHS 3	N		Top 30
	EHS 4	N		Top 30

Figure 3.13 Amidation reaction solvent selection algorithm screenshot, step 3: Parameter input.

The LSER modelling ln(**k**) as a function of **β** (Equation 3.1) was entered into the solvent selection algorithm and the top 30 solvent candidates permitted to pass the solvent performance assessment (model B2). To determine the greenness of solvent candidates a stricter approach was taken to that used in the nucleophilic substitution case study. Again using the GSK solvent selection guide as its basis, a minimum score of 4 out of 10 was required in each category (Figure 3.14). This included the LCA category that was omitted from the nucelophilic substitution case study. Furthermore the LCA category was given a much greater weighting in the assessment than the other categories, accounting for half of the eventual greenness assessment score (model B3). As per usual this does not affect the number of solvent candidates passing the assessment only the score attributed to them.



Step 6 *Solvent selection guide*

Type: GSK

Minimum biostandard: Petroleum derived

Cut-off: Top 30

Method	Scoring methodology		Weighting	Minimum
Pfizer	Green	Preferable		☐ Green
	Yellow	Useable		☑ Yellow
	Red	Undesirable		☐ Red
	Pink	Chlorinated solvents		☐ Pink
GSK*	Waste		10.0 %	4
	Environmental impact		10.0 %	4
	Health		10.0 %	4
	Flammability		10.0 %	4
	Stability		10.0 %	4
	Life cycle assessment		50.0 %	4
ETH**	Safety		25.0 %	1.0
	Health		25.0 %	1.0
	Environment		25.0 %	1.0
	Cummulative energy		25.0 %	300

*Scores proportional to greenness (/10).

**Scores inversely proportional to greenness (/1.00).

Figure 3.14 Amidation reaction solvent selection algorithm screenshot, step 6: Solvent selection guide.

Because limonene and *p*-cymene do not feature in any solvent selection guides they cannot be fully assessed. However both solvent candidates successfully negotiate the revised solvent selection algorithm (model B1) and the solvent performance assessment (model B2). The only three solvents that pass the revised algorithm and the two associated solvent performance and greenness assessments are 1,2-dichlorobenzene, chlorobenzene, and cumene (Table 3.6). The maximum score was 140, just as in the previous case study. None of these solvents immediately appear to be what could be considered as green solvents, but all three do not have any scores in the GSK solvent selection guide below 4, which was set as the minimum requirement. These solvents actually perform quite well in the GSK solvent selection guide which puts them at an advantage over limonene and *p*-cymene which are not present in any publicly available solvent selection tool. Toluene scores a 3 in the environmental impact category of the GSK solvent selection guide, eliminating it from the final solvent candidate list [Henderson 2011]. Due to the confines of the solvent selection algorithm's solvent database, and the demands of the model reaction, it is important to note the lack of bio-based suggestions for an improved amidation solvent.

Table 3.6 Solvent hits generated by the solvent selection algorithm for amidation reactions.

Solvent	Score: Model B1 + model B2	Score: Model B1 + model B3
1,2,4-Trichlorobenzene	No data	116
1,2-Dichlorobenzene	116	116
Chlorobenzene	92	112
cis-Decalin	100	Fail
Cumene	78	78
Cyclohexanone	Fail	92
Cyclopentanone	Fail	96
Diphenyl ether	68	No data
Limonene	136	No data
Mesitylene	80	Fail
m-Xylene	72	No data
Nitromethane	94	Fail
p-Cymene	78	No data
p-Xylene	74	Fail
Tetrachloroethylene	116	No data
Toluene	72	Fail
Total hits	13	6

Bio-based amidation solvent implementation: Predicted rates of reaction are essentially the same across the aromatic and chloroarene solvents but with limonene (**β** = 0.00) emerging as a marginal favourite. The certainty in the apparent greenness of the petroleum based solvents was overridden by the desire to implement bio-based solvents. As such limonene and *p*-cymene were taken forward and tested in the model amidation. Reaction rates were expectedly high (Table 3.7). But prolonged heating of limonene resulted in some discolouration, suggesting that *p*-cymene, a more robust solvent, would be preferable given a choice of the two. Degradation of limonene by oligomerisation, isomerisation, or oxidation could all be possible [Thomas 1989].

Aside from a risk of hydroperoxidation in the presence of a base at elevated temperature (resulting in *p*-cresol and acetone) *p*-cymene is inert under most realistic reaction conditions [Fiege 1995 page 33, Makgwane 2010].

Table 3.7 A comparison between high performance bio-based and petro-chemical amidation solvents.

Solvent	ln(**k**)	α	β	π*	SUS	HAS	ECO
Chlorobenzene	-10.51	0.00	0.06	0.65	0	4	10
p-Cymene	-10.62	0.00	0.13	0.39	6	n/a	n/a
Limonene	-10.50	0.00	0.00	0.16	10	n/a	n/a
Toluene	-10.65	0.00	0.12	0.50	0	3	6

Limonene and *p*-cymene fit the trend described by the LSER (Equation 3.1) with excellent precision (Figure 3.15). As expected limonene improves upon the maximum kinetics obtained in non-renewable solvents. The relationship between ln(**k**) values predicted with the LSER and experimental data is also strong, aiding the hypotheses presented earlier in the chapter (Figure 3.16). *p*-Xylene has not been designated as a bio-based solvent for the purpose of Figure 3.15 or

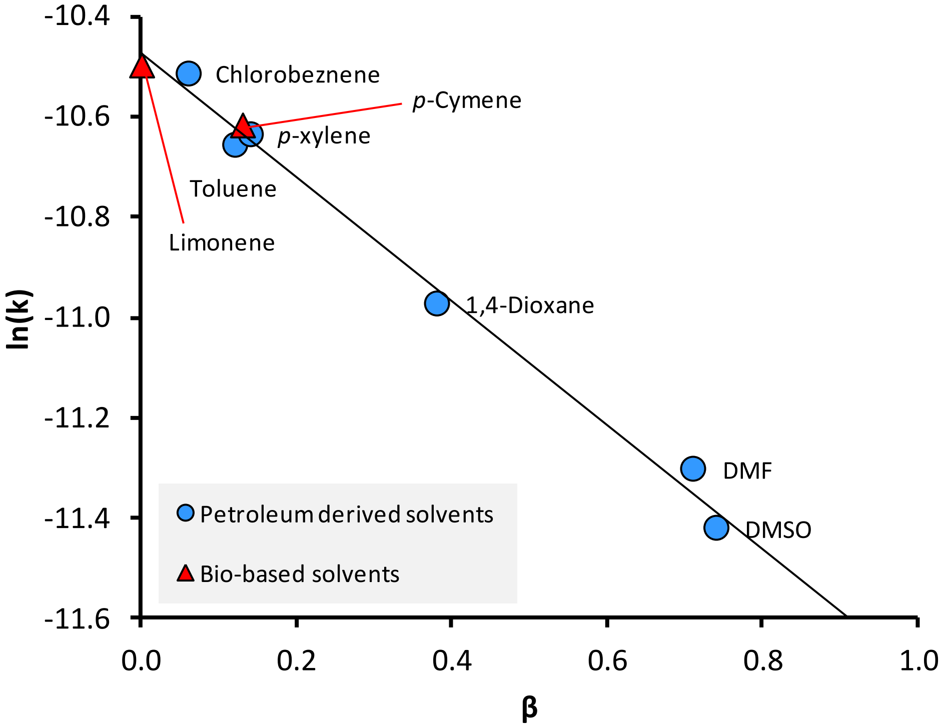


Figure 3.15 A LSER indicating the performance of limonene and *p*-cymene in amidation reactions.

in subsequent analyses. Although *p*-xylene was awarded a SUS classification of 6, the lower limit for recognition as a bio-based solvent as established in Chapter 1, a clear comparison between the citrus derived solvents and the original solvent set was preferable in this instance. However this does raise an interesting point. It is quite likely that given the choice between *p*-cymene and p-xylene, both of renewable sources, an organic chemist would opt for *p*-xylene because of its familiarity. On top of this the lower boiling point of *p*-xylene will usually work in its favour too. For the purpose of this work, with its emphasis strongly on the feedstock from which the solvent is made, *p*-cymene will be favoured over *p*-xylene. Whereas *p*-xylene is made by the fermentation of edible sugars, *p*-cymene originates from a voluminous waste-stream [Lohrasbi 2010, Martin-Luengo 2010].

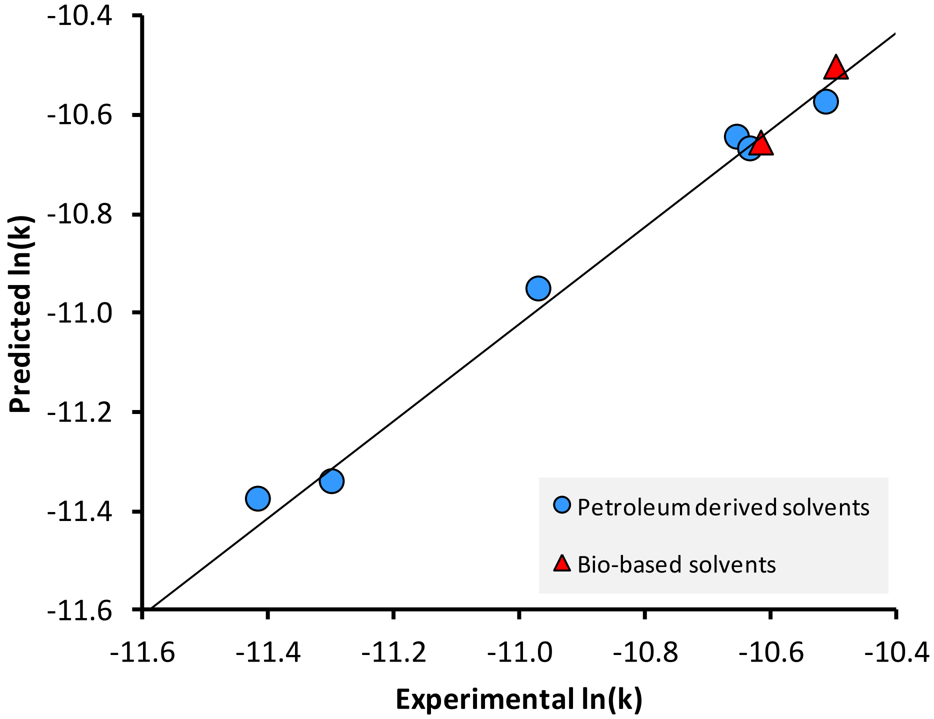


Figure 3.16 Experimental versus predicted ln(**k**) values of amidation.

There seems to be no major systematic errors across the solvent set (Figure 3.17). The range of error is below that found in the previous Menschutkin reaction case study (cyclohexanone is the most erroneous based on the proposed correlation with **β**). Given the slower reaction progress in the amidation compared to the model Menschutkin reaction, this adds further credence to the proposal that the velocity of the reaction is responsible for the achievable accuracy.

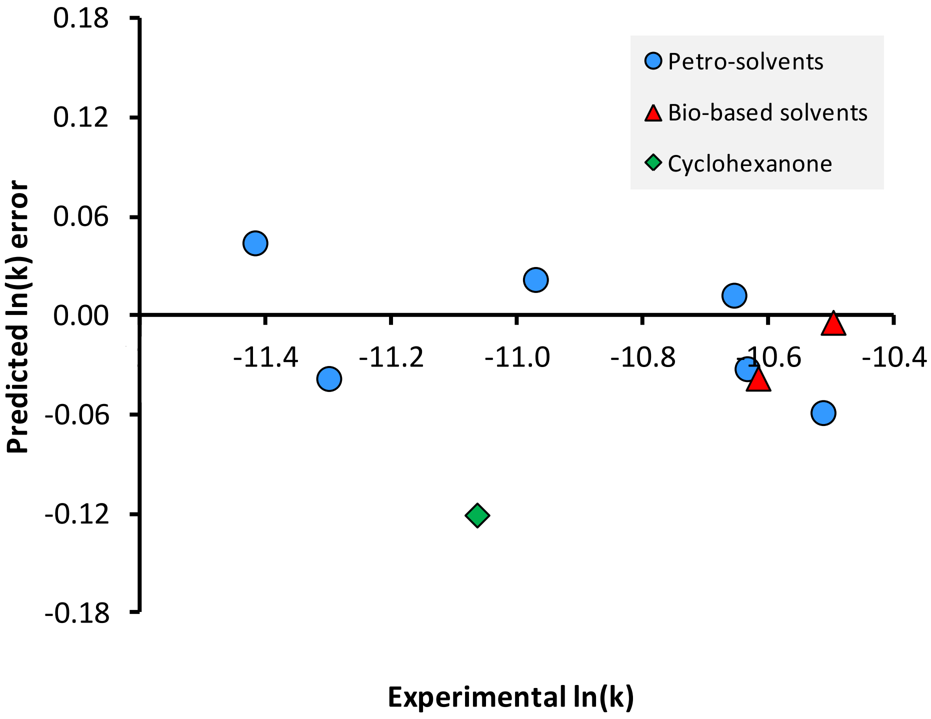


Figure 3.17 Systematic error check in the amidation rate constants.

Preparative scale amidation and analysis: Monitoring the kinetics of this model amidation serves a purpose only if the synthesis and isolation of the product is realised. Bearing in mind that this is an uncatalysed procedure, also devoid of coupling agents, the reaction was conducted at 373 K for 24 hours in toluene, *p*-cymene, limonene, cold pressed orange oil, and for 48 hours in DMSO. Cold pressed orange oil was obtained from Sigma-Aldrich chemical suppliers as an orange coloured liquid. The steam distilled essential oil of orange is colourless, indicating that the non-volatile compounds existing in the oil are retrieved when extracting the oil from the food waste by force. The cold pressed orange oil is still more than 90% limonene.

The carboxamide product was recystallised until thin, needle-like crystals were obtained. Characterisation was consistent with the literature [Verma 1998]. As expected toluene and *p*-cymene are similarly productive, with an average of 74% yield obtained in both instances (Figure 3.18). With the boiling point of *p*-cymene being significantly higher than the reaction temperature there is scope for further yield improvement not applicable to toluene. With some disappointment the use of either limonene or cold pressed orange oil did not result in the yields of amide expected from the rate constant provided by limonene in earlier reactions. The gradual decomposition of limonene at 373 K may be to blame. The slight difference in amide yields suggests that the other minor components of cold pressed orange oil do not interfere with the reaction to any significant extent. This is probably not a universal rule, which will have to be considered in the future. Dimethyl sulphoxide was confirmed to be a poor choice of solvent with an average product yield of 24%, far below that required for a fairly uncomplicated organic synthesis protocol. Furthermore, the product was quick to precipitate from solutions of toluene and *p*-cymene upon cooling, aiding isolation.

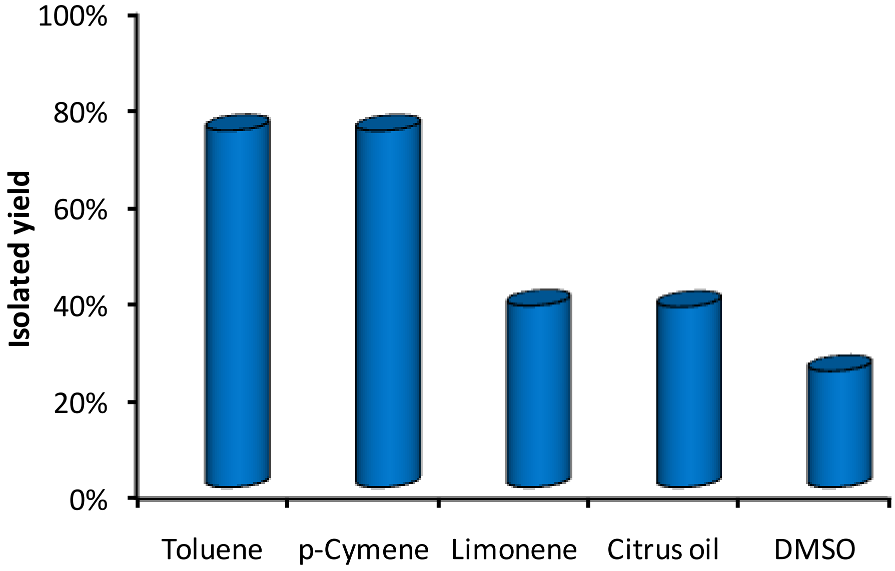


Figure 3.18 Isolated yields of *N*-benzyl-4-phenylbutanamide from different solvents.

In terms of toxicity *p*-cymene is on a par with toluene [MSDS 2013]. The same can be said when considering their aquatic toxicity (Table 3.8). However, unlike toluene its flash point (closed cup) is above room temperature and has a lower vapour pressure at ambient temperatures. Given that *p*-cymene can also be made from a renewable feedstock there are sufficient benefits to justify the substitution of toluene with *p*-cymene. The chief drawback to *p*-cymene that will be

Table 3.8 Properties of toluene and *p*-cymene.

Property	Toluene	*p*-Cymene
Melting point /K	178 [Gani 2005]	204 [ChemIDplus 2013]
Boiling point /K	384 [Gani 2005]	449 [MSDS 2013]
δ_T /MPa$^{0.5}$	18.3 [Gani 2005]	17.3 [Hukkerikar 2012]
Flash point /K	277 [MSDS 2013]	320 [MSDS 2013]
Autoignition point /K	808 [MSDS 2013]	709 [MSDS 2013]
LD_{50} (Rat oral) /mgkg^{-1}	>5580 [MSDS 2013]	4750 [MSDS 2013]
logP	2.58 [Gani 2005]	4.1 [ChemIDplus 2012]
LC_{50} (Sheepshead minnow) /mgL^{-1} (4 days)	280 [ECOTOX 2013]	48 [ECOTOX 2013]
EC_{50} (Daphnia magna) /mgL^{-1} (24 hr)	8.0 [MSDS 2013]	9.4 [ECOTOX 2013]

felt by chemists is its higher boiling point when compared to toluene. Because many bio-based solvents will be non-volatile this may result in the need for a general departure from distillation as a means of solvent removal. Avoidance of distillation has already been identified as a priority for contemporary solvent research, and moving towards low volatility solvents will reduce loses to the atmosphere [Jessop 2011].

Metrics can be considered to further enhance our understanding of the reaction as a preparative procedure. The selection of reactions surveyed earlier in this chapter have been complied and their reported yields listed, along with calculated atom economies and reaction mass efficiencies [Constable 2002, Curzons 2001]. These can be compared to the results of the uncatalysed reactions performed in this work (Figure 3.19). The key to Figure 3.19 is as follows: A, uncatalysed reaction in *p*-cymene; B, uncatalysed reaction in limonene; C, thermally activated solventless reaction [Gooßen 2009]; D, thionyl chloride activation [Comerford 2009, Pearson 1999 page 370]; E, CDI activation [Vaidyanathan 2004]; F, DCC activation [Comerford 2009, Sheehan 1955]; G, catalysis by boric acid [Tang 2005]; H, catalysis by 3,4,5-trifluorobenzoic acid [Ishihara 1996]; I, catalysis by *2-N*,*N*-diisopropylbenzylaminoboronic acid [Arnold 2008]; J, as I but with an approximation of the catalyst synthesis based on the procedure for making *2-N*,*N*-diisopropyl-5-fluorobenzylaminoboronic acid [Arnold 2008]; K, catalysis by activated silica [Comerford 2009].

With the exception of the limonene facilitated reaction (entry B, Figure 3.19), yields are good to excellent in all other cases. With water as the only necessary by-product, atom economies only fall below satisfactory levels when activators are introduced and treated as reactants. So-called solventless reactions (*e.g.* entry C, Figure 3.19) do not present any extra benefit to waste minimisation in this assessment because solvents are not included in these metrics. The best scores are associated with catalytic processes utilising boric acid (entry G, Figure 3.19) and 3,4,5-trifluorobenzoic acid (entry H, Figure 3.19). This said, a word of caution must be exercised regarding catalysts. The most striking change to reaction mass efficiency (RME) is observed when the synthesis of an elaborate catalyst is included in the calculation (entry J, Figure 3.19). With *2-N*,*N*-diisopropyl-5-fluorobenzylaminoboronic acid as an example, RME falls from 64% to 19% [Arnold 2008]. Catalyst loadings lower than 10 mol% would reduce the impact of the catalyst which is why high catalytic activity is desperately needed from amidation catalysts. The synthesis of *2-N*,*N*-diisopropyl-5-fluorobenzylaminoboronic acid was approximated from available data (Scheme 3.9). A discussion of this calculation can be found in the appendix.

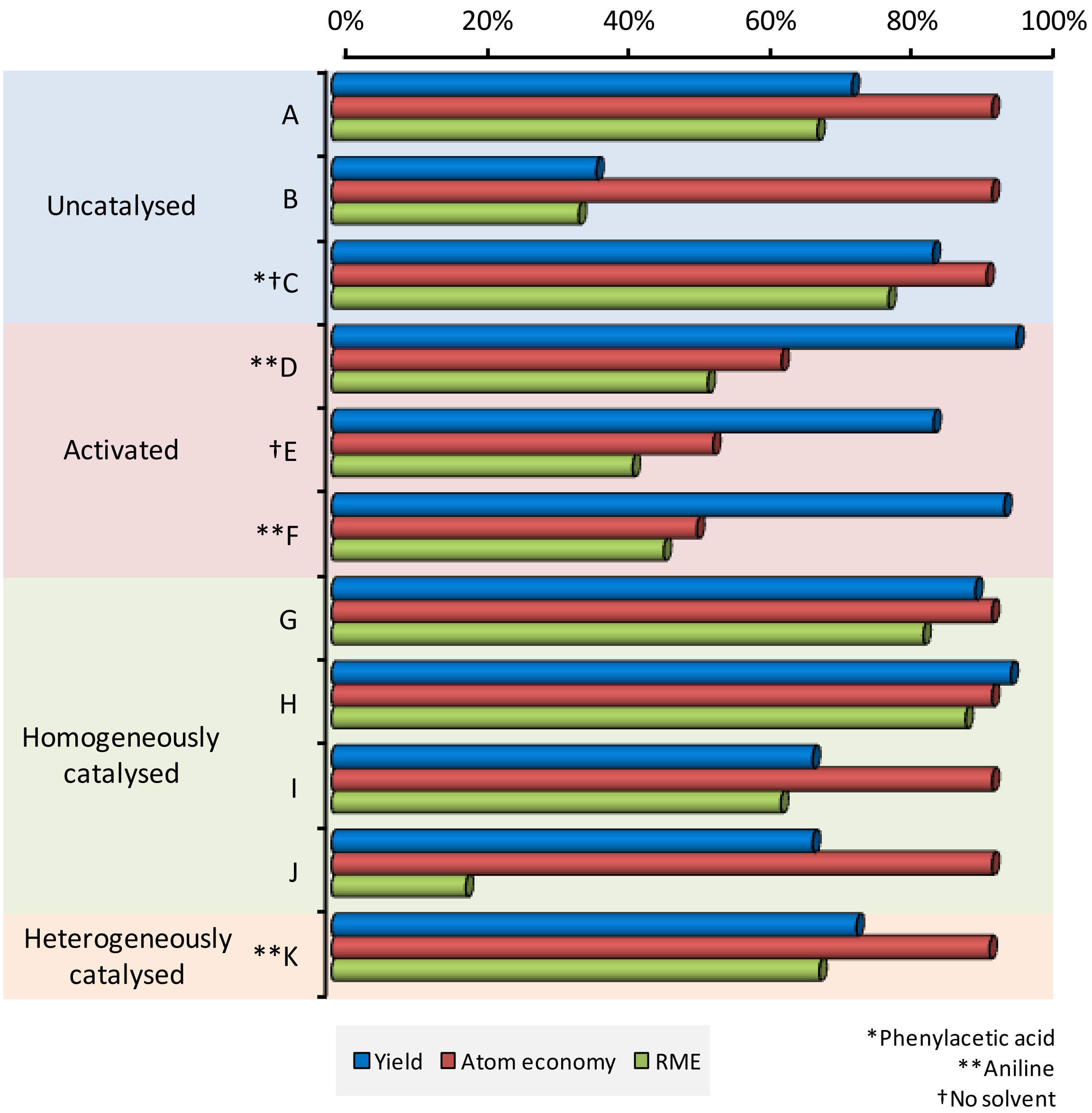
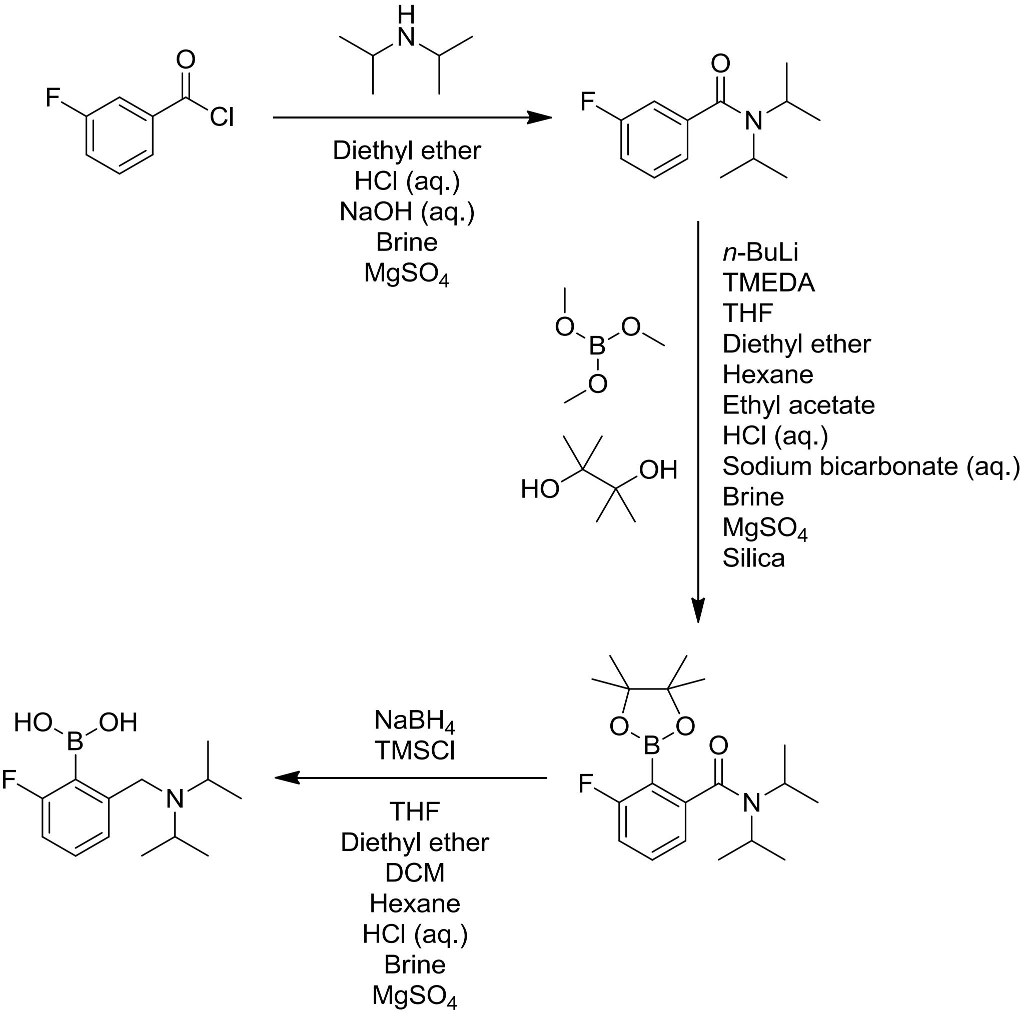


Figure 3.19 Associated metrics of amidation reactions between 4-phenylbutanoic acid and benzylamine unless otherwise stated.

Activated silica is a strong candidate for a catalyst when considering the reaction used in this demonstration involved the less reactive aniline and not benzylamine [Comerford 2009]. To its detriment, reactivation of the catalyst requires a high temperature which is not accounted for in these mass utilisation metrics. Boric acid in certain instances will offer very high yields, and appears to operate successfully in aromatic solvents. Perhaps activated silica and boric acid can be considered as effective, benign catalyst options to complement *p*-cymene as the choice of solvent in future studies. It should be reiterated that the uncatalysed reaction can be satisfactory too, with or without solvent, but it is very substrate specific.


Diethyl ether
HCl (aq.)
NaOH (aq.)
Brine
MgSO4
n-BuLi
TMEDA
THF
Diethyl ether
Hexane
Ethyl acetate
HCl (aq.)
Sodium bicarbonate (aq.)
Brine
MgSO4
Silica
NaBH4
TMSCl
THF
Diethyl ether
DCM
Hexane
HCl (aq.)
Brine
MgSO4

Scheme 3.9 The synthesis of *2-N*,*N*-diisopropyl-5-fluorobenzylaminoboronic acid.

Another metric worthy of attention is process mass intensity (PMI) [Jiménez-González 2011]. As the ratio of input to output materials by way of a mass balance it goes further than the previous metrics, now accounting for the use of solvent (Figure 3.20). Solvents are responsible for almost all of the mass input in this sample of reactions, highlighting the limitations of the previous collection of metrics. The solventless reactions can now prove their worth, although both examples used a solvent in their work-up procedures (entries C and E, Figure 3.20). The old argument of quality over quantity can be applied here, toluene being the usual choice of solvent. Replacing toluene with limonene or *p*-cymene may not decrease the volume of waste solvent but it will alleviate the burden somewhat on limited petroleum resources. The counter-claim against this argument is that *p*-cymene is probably harder to recycle than toluene using conventional means because of its higher boiling point.

The mass utilisation issues with the associated synthesis of elaborate amidation catalysts (entry J, Figure 3.20) can be trumped by the solvent intensity of work-up procedure, which in the case of homogeneously catalysed reactions can be quite severe. One amidation protocol calls for

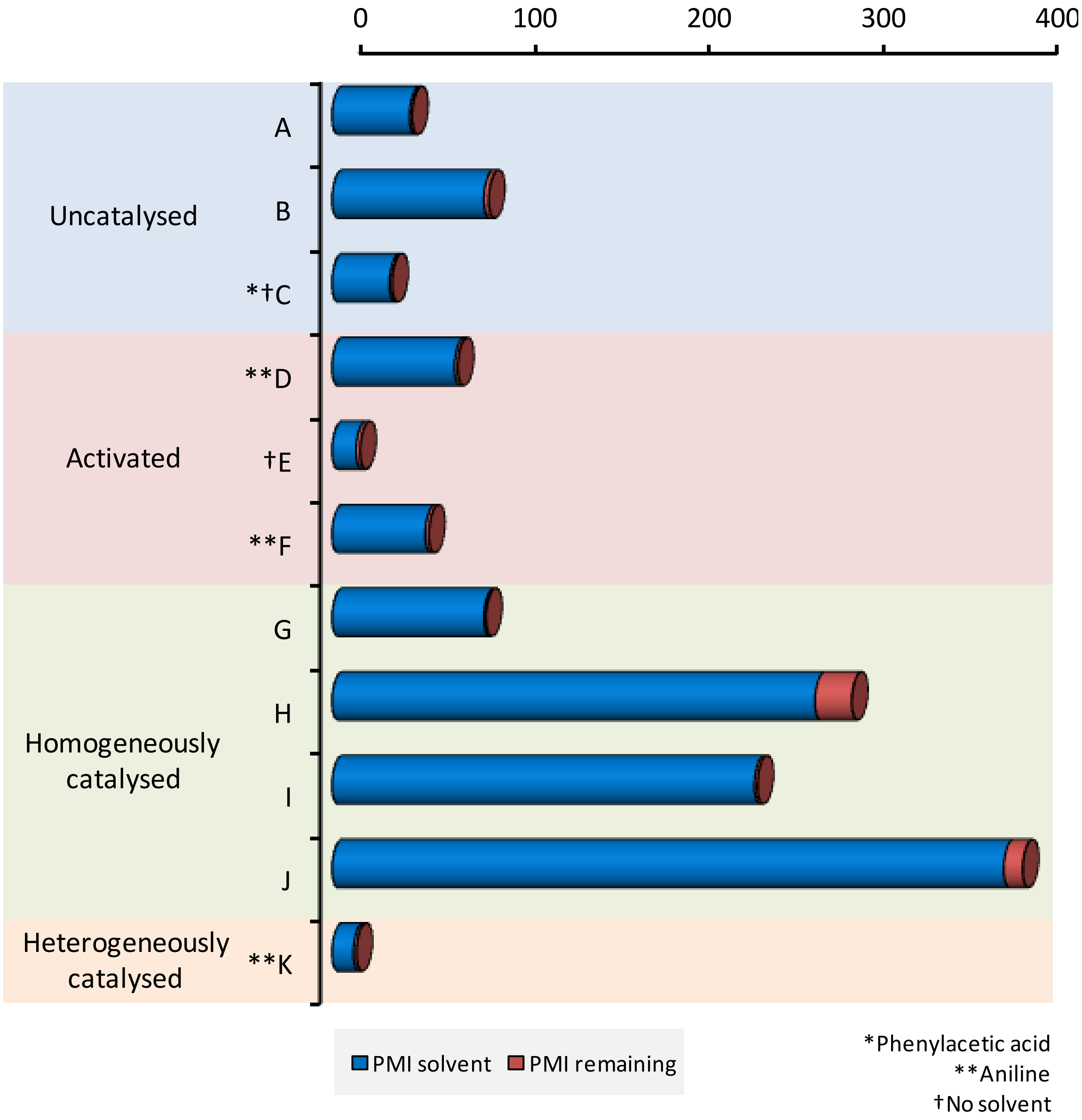


Figure 3.20 The comparison of PMI in different amidations. The key for Figure 3.19 also applies here.

the use of column chromatograph to isolate the product (entry H, Figure 3.20) [Ishihara 1996]. This also adds to the non-solvent waste in the form of spent chromatographic silica. The large PMI values found elsewhere are caused by solvent washes and recystallisation solvents. The least wasteful procedure is the silica catalysed amidation (entry K, Figure 3.20). The catalyst is washed with an additional aliquot of toluene before the product is allowed to crystallise from the filtered solution. This is close to an ideal situation and the catalyst can be recycled without much further issue.

3.3 Amidation summary

The success and implications of solvent selection: It appears that the rate of amidation is accelerated by solvents with low **β** values, based on the case study of benzylamine reacting with 4-phenylbutanoic acid. In addition to the measurement of rate constants, the enthalpy of activation and the entropy of activation were determined. Both these parameters are also dependant on solvent hydrogen bond basicity. The explanation for this phenomenon is founded on the enthalpic benefits of stabilising the activated complex of the reaction being overridden by the large entropic penalty created by organising the solvent to form these interactions. Hydrogen bond acidity may be influential too, but in this case study no protic solvents were used.

There has been, at this fundamental level, disagreement over the reaction mechanism. The results presented in this work do not suggest the trimolecular mechanism of Scheme 3.4 is absolute. With a reaction order magnitude of 1.1 with respect to 4-phenylbutanoic acid, it is simply beneficial when it does occur. The polarity of the solvent possibly has a role in disrupting the carboxylic acid dimer, with strong hydrogen bond acceptors like DMSO able to provide their own source of intermolecular stabilisation. The same reaction order has been attributed to benzylamine, suggesting that stabilisation of the activated complex through the self assistance mechanism can occur with either acid or amine as the third party without bias (Scheme 3.2).

Given that perhaps the ideal conventional solvent for this class of chemistry is toluene, limonene and more obviously *p*-cymene offer a sustainable alternative. The synthesis of amides gives products that are likely in most cases to precipitate out of a *p*-cymene solution. This greatly simplifies their purification because the high boiling point of *p*-cymene will often be an issue in this regard. Currently *p*-cymene is synthesised by the alkylation of toluene, the solvent it has the intention of replacing [Strohmeyer 1971]. Consequently *p*-cymene (£18.33/L, Sigma-Aldrich, 99% purity, as of 21st June 2013) is more expensive than toluene (£11.86/L, Sigma-Aldrich, >99.5% purity, as of 21st June 2013) which will be important industrially. Until a sustainable manufacturing process for *p*-cymene is operational (and necessarily economically viable), persuasive arguments advocating the use of *p*-cymene must be accumulated until demand is satisfactory to warrant commercial production from citrus waste. An LCA comparing toluene and bio-based *p*-cymene will be required to prove the transition from the former to the latter is indeed worthwhile, but reassuringly the two step isomerisation and dehydrogenation of limonene to afford *p*-cymene is not drastically out of line from current protocols in the fine chemical and bulk chemical sectors.

Broader appeal: Amidation is prominent in organic synthesis, and so the appeal of work concerning amide forming reactions is necessarily far reaching. But if current practice prevails, and coupling agents remain widely utilised, then the results of this specific work become much less significant. MacMilllan has recently shown that a range of different solvents can be used without too much change in the rates of coupling agent assisted amidations at room temperature [MacMillan 2013]. However the work of Balcom highlights that carbonyl addition solvent effects can still be important at room temperature [Balcom 1989], and indeed the point at which catalysis becomes appealing is at similarly ambient temperatures and for equally short durations.

The assumption that amidation catalysis will develop to the point where catalysts offer a reasonable alternative to stoichiometric auxiliaries is vital. At this point the interplay between solvent and catalyst becomes important. Future work should definitely address the optimal pairing of solvent and catalyst across a larger range of substrates. *p*-Cymene should make a complimentary partner to most catalysts, and maybe the design of catalysts should be guided by the likely choice of performance enhancing solvent. Activated silica has already been shown to be an effective amidation catalyst compatible with aromatic solvents and flow reactors, making it the ideal companion for *p*-cymene [Comerford 2012]. The silica acts as a stationary phase, and the reaction solution can later be cooled down which should cause the product to precipitate. The *p*-cymene could then be reused after the depleted reactant concentrations are rejuvenated. The challenge facing amidation chemistry is equally about its perception as it is about the development of greener technologies. To this end, attention should be drawn to the use of coupling agents, and this practice discouraged on grounds of mass efficiency whilst promiting burgeoning catalytic technologies.

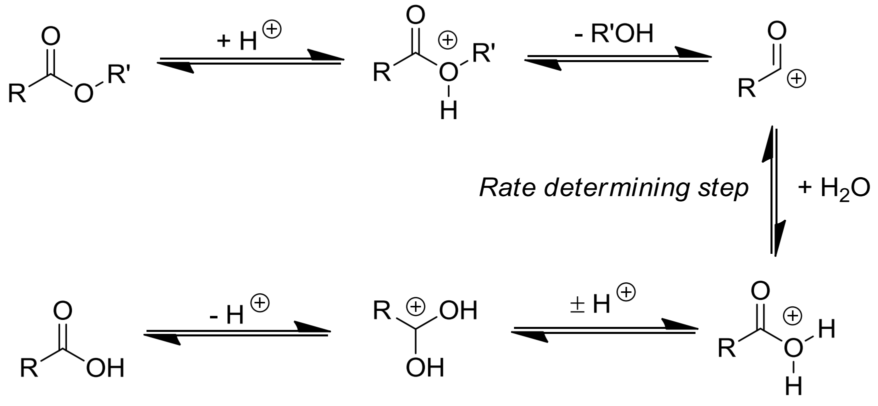
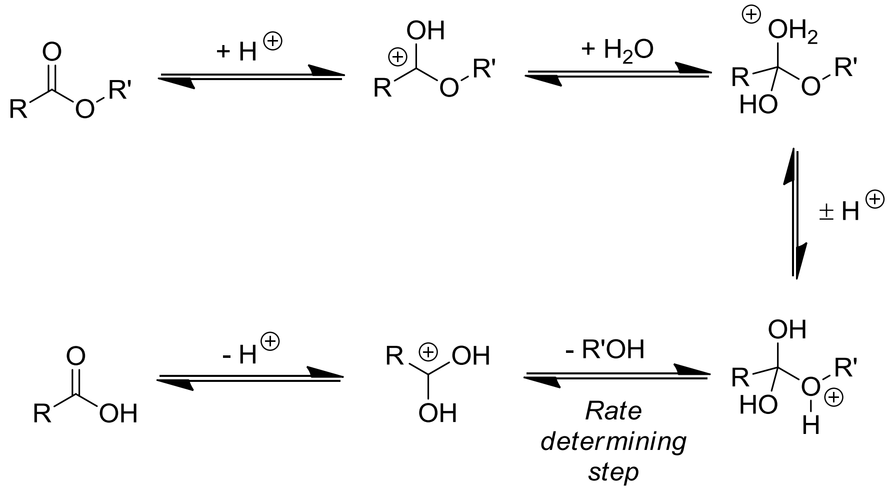
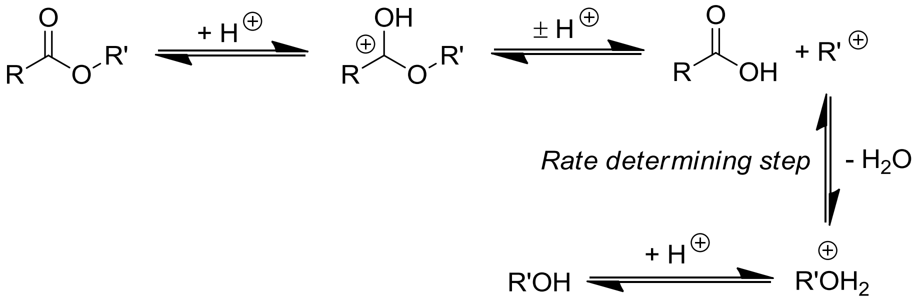
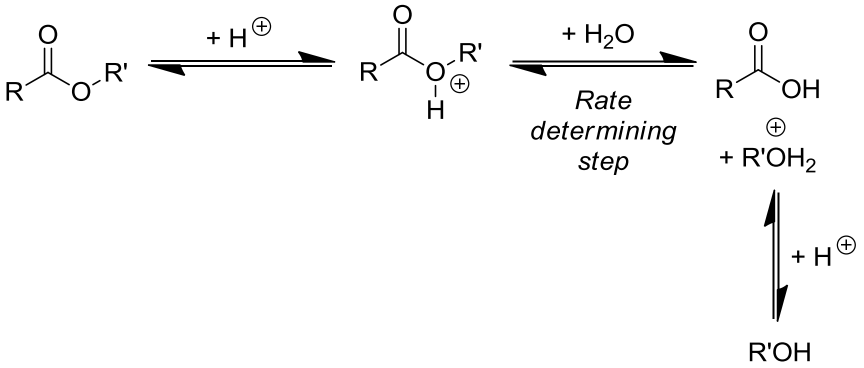
4. Uncatalysed esterification

The very first examination of solvents influencing the rate of a reaction was Berthalot's nineteenth century observations on the synthesis of ethyl acetate, preceding Menschutkin's work on solvent effects by a few decades [Reichardt 2003 page 2]. Esters, like ethyl acetate, are useful products for a variety of applications, including fragrances and pharmaceuticals [Liu 2006, Yadav 2004]. Many solvents are esters. Acylations account for 12% of the chemistry performed in pharmaceutical research and development [Carey 2006]. The majority are implemented for the synthesis of amides, but esterification and its reverse reaction are prominent in their own right at manufacturing scale [Dugger 2005]. Esterification features heavily in the literature when green chemistry is concerned, usually with the catalyst under scrutiny. Such is the importance of catalysis, the interplay between solvent and catalyst is addressed the following chapter. Prior to investigations into catalysis, the role of the solvent shall be analysed here in the absence of a catalyst, and the two will be introduced together when appropriate. Given the previous case study on amidation and its superficial similarity to esterification, one set of results should strengthen or disprove the hypotheses formulated in the other to the benefit of the whole research.

4.1 Solvents and esterification

Mechanism and protocol: Concerning the reaction of carboxylic acids with alcohols, eight possible reaction mechanisms have been proposed for the hydrolysis of esters, of which half (only the acid catalysed mechanisms) are applicable to the reverse reaction [Ingold 1969 page 1129, Smith 2007 page 1402]. The most commonly observed of these reaction mechanisms proceeds *via* a tetrahedral intermediate and has been given the abbreviated name $A_{AC}2$, representing a second order, acid catalysed acylation or ester hydrolysis (Table 4.1). The rate determining step is the combination of the protonated acid and alcohol to form a tetrahedral intermediate. As such this reaction is second order and probably similar to what was established for amidation reactions previously in this work.

Table 4.1 The acid catalysed mechanisms of ester hydrolysis.

Code	Mechanism	Class
$A_{AC}1$	 + H⊕ - R'OH Rate determining step + H_2O ± H⊕ - H⊕	S_N1
$A_{AC}2$	 + H⊕ + H_2O ± H⊕ - R'OH Rate determining step - H⊕	Tetrahedral
$A_{AL}1$	 + H⊕ ± H⊕ + R'⊕ Rate determining step - H_2O + H⊕	S_N1
$A_{AL}2$	 + H⊕ + H_2O Rate determining step + R'OH$_2$⊕ + H⊕	S_N2

There are four traditional methods by which the acylation of alcohols with carboxylic acids can be performed; employing an excess of one reactant to drive the equilibrium forward, removal of water by azeotropic distillation, removal of water using dehydrating agents, or continuous extraction of the ester product by distillation [Smith 2007 page 1414]. The common theme is the separation of the products, the ester and water, in order to overcome the yield limiting equilibrium position. Dean-Stark apparatus is suitable for this purpose, and limonene has already been demonstrated as being a solvent compatible with this technology [Veillet 2010]. Of course, solvents are also important, not only to influence the equilibrium position but also for stabilising the reaction intermediates. Two of the four acid catalysed esterification (or hydrolysis) mechanisms in Table 4.1 are designated as S_N1 reactions which should favour highly dipolar or polarisable protic solvents as alluded to briefly in Chapter 2 (Table 2.1). The one S_N2 mechanism pathway in Table 4.1 involves a charged reactant and as such will not adhere to the LSER devised for the model Menschutkin reaction, but rather this particular mechanistic pathway should be accelerated by aprotic hydrogen bond accepting solvents (Table 2.6). However by far the most common mode of esterification is that described by the $A_{AC}2$ mechanism [Bender 1951]. This is the same as the mechanism of amidation and as such the solvent dependence will be different to the other three esterification mechanisms (but possibly the same as for amide forming reactions). The solvent effects controlling esterification deduced later on can help identify the mechanism of the reaction more precisely.

Another option is to employ the acid anhydride of the carboxylic acid reactant. The reaction between sterically unhindered reactants and an acid anhydride does not require an additional catalyst, but in the reaction with tertiary alcohols catalytic quantities of a catalyst such as DMAP are usually employed [Hölfe 1987, Reicheneder 2004]. *N*,*N*-Dimethyl-4-aminopyridine (DMAP) is a nucleophilic catalyst, with a stoichiometric amount of an auxiliary base usually added to prevent protonation of the catalyst.

The application of solvents in esterification: A great deal of work has been done on refining the acylation of alcohols to develop a protocol that is consistent with green chemistry principles [Otera 2001]. The ultimate goal is the use of equimolar quantities of each reactant to give 100% conversion. To this end, recycling of the catalyst and solvent would result in zero waste. The work lead by Otera has come close to this ideal scenario. By using a catalyst with fluorinated appendages in a liquid fluorous phase, the reaction medium (solvent plus catalyst) can be recycled, reducing the environmental impact of the tin based catalyst and perfluorohexane solvent. An elevated reaction temperature is of kinetic assistance, but heating is also necessary to dissolve the carboxylic acid and alcohol reactants [Otera 2002]. Upon cooling the product is immiscible with the fluorous phase and hence easily separated (Figure 4.1). Using the equivalent acid anhydride, the reaction will proceed at room temperature [Otera 2005]. An additional issue is that the recycling of the perfluorinated solvent demands the use of a traditional organic solvent, in this case toluene, to retrieve the solutes. Although high yielding as it stands, there is definitely scope for the application of safer, renewable, and more environmentally sound bio-based solvents in place of this procedure.

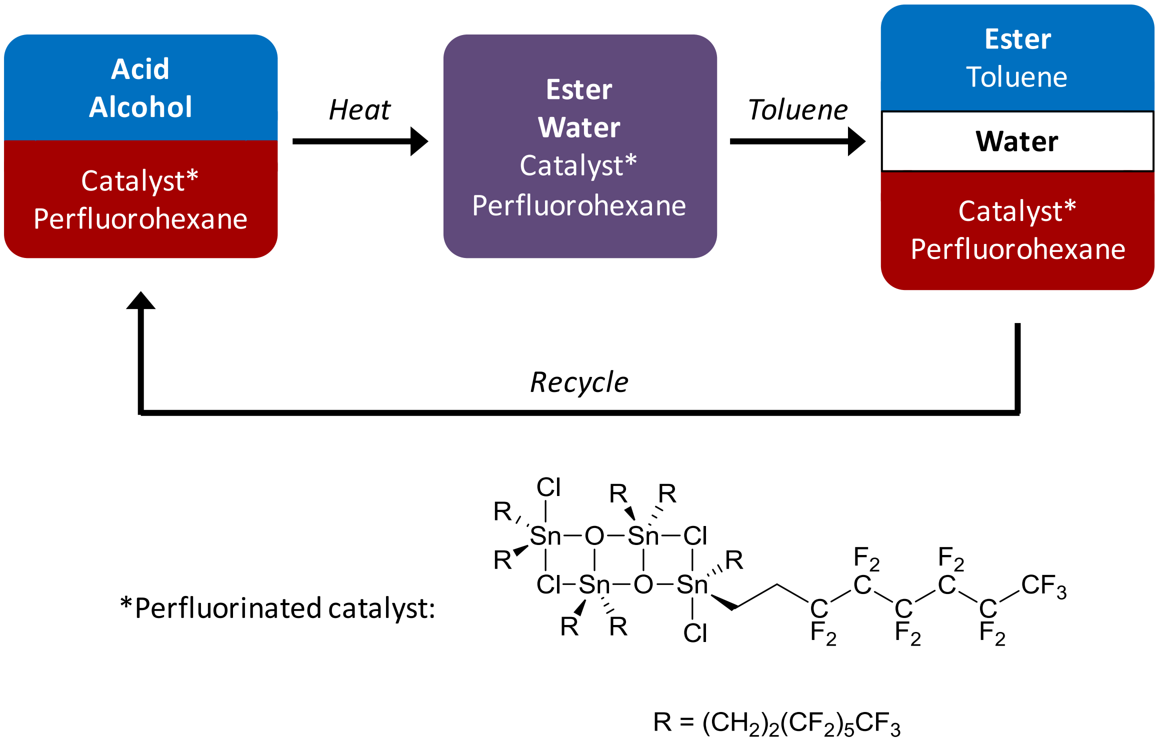
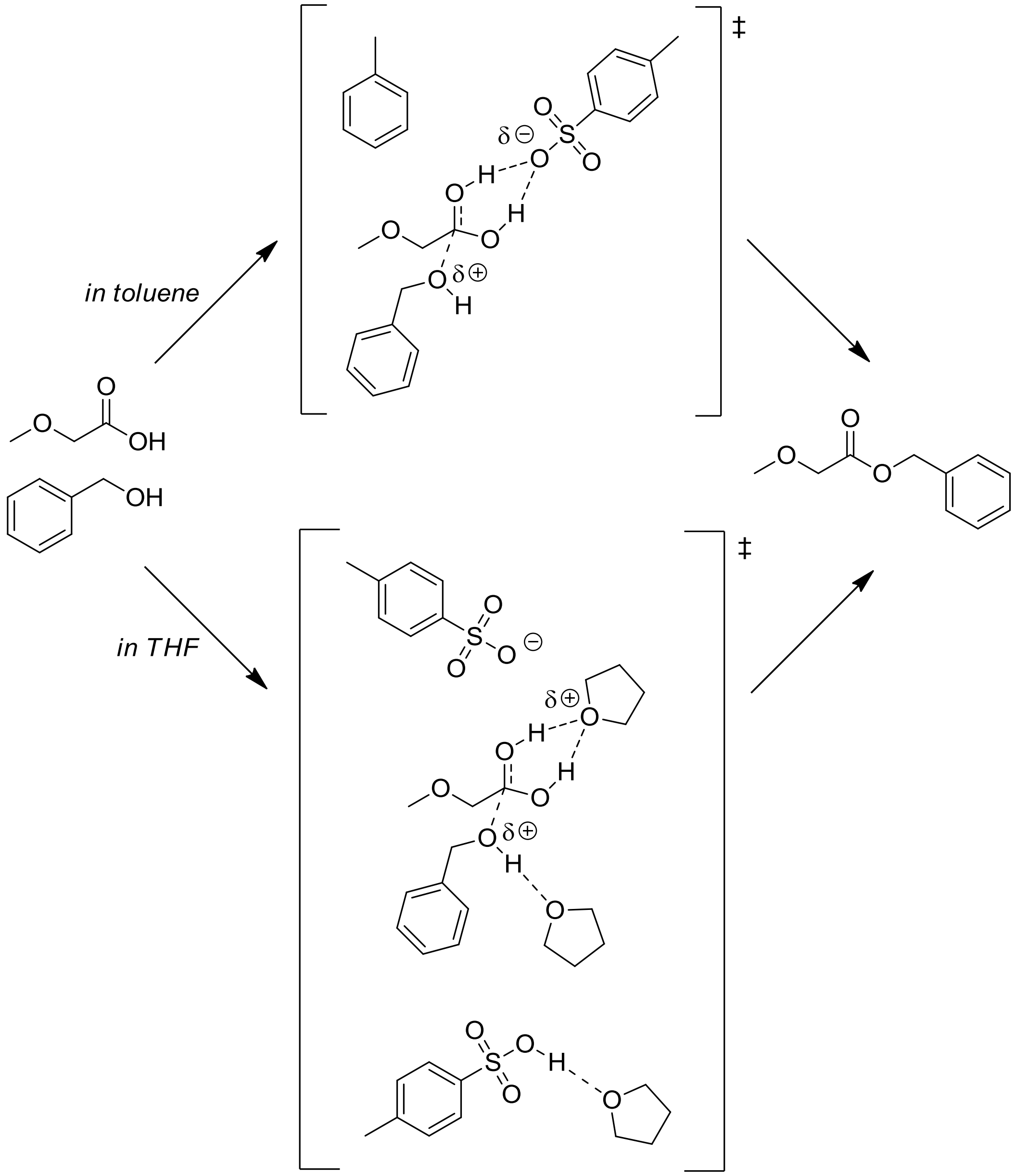

Acid
Alcohol
Catalyst*
Perfluorohexane
Heat
Ester
Water
Catalyst*
Perfluorohexane
Toluene
Ester
Toluene
Water
Catalyst*
Perfluorohexane
Recycle
*Perfluorinated catalyst:
R = (CH2)2(CF2)5CF3

Figure 4.1 The manipulation of a fluorous multiphasic system to enhance esterification yields.

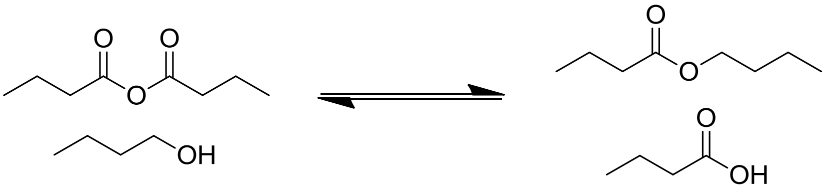
The rate of Fischer esterification between methoxyacetic acid and benzyl alcohol in various molecular and ionic liquid solvents has been reported as expressing an inverse relationship with **β** (Figure 1.19), mirroring the amidation case study previously described (Equation 3.1) [Wells 2008]. The kinetic benefit of hydrogen bond accepting solvents was attributed to the avoidance of a deactivating acid-base interaction between the catalyst (*p*-TSA) and the solvent (Scheme 4.1). Although this is likely, it seems plausible that a more general solvent effect, the same as that observed in amidation reactions, is also in operation. It was reported by Wells *et al*. that the reaction is bimolecular, and of the $A_{AC}2$ tetrahedral mechanism. This is consistent with the observed solvent effect, and strengthens the similarity with the kinetics of uncatalysed amide synthesis [Wells 2008].



Scheme 4.1 A proposal for an acid catalysed esterification solvent effect.

4.2 Uncatalysed esterification results and discussion

Model reaction: In order to confirm the solvent effect present in esterifications and identify the ideal renewable solvent, an uncatalysed reaction between butanoic anhydride and 1-butanol to give butyl butanoate was attempted (Scheme 4.2). Acid catalysed Fischer esterifications can be found in the following chapter. If the same inverse dependence on solvent hydrogen bond accepting ability (**β**) is observed in this case study as it was for amidations and Well's catalysed esterification, it would constitute as good evidence of an analogous mechanism being in operation.



Scheme 4.2 The model uncatalysed esterification reaction.

The kinetics of the reaction at 323 K to give butyl butanoate were again monitored by ^{1}H-NMR spectroscopy. This time the signal integrals of the O-CH$_2$ moiety of 1-butanol and the equivalent group present in its ester upon reaction with butanoic anhydride were used to deduce the progress of the reaction (Figure 4.2). The data obtained from the ^{1}H-NMR spectra was processed in an identical way to the amidation case study, assuming a bimolecular reaction mechanism (Equation 1.14 and Equation 1.15).

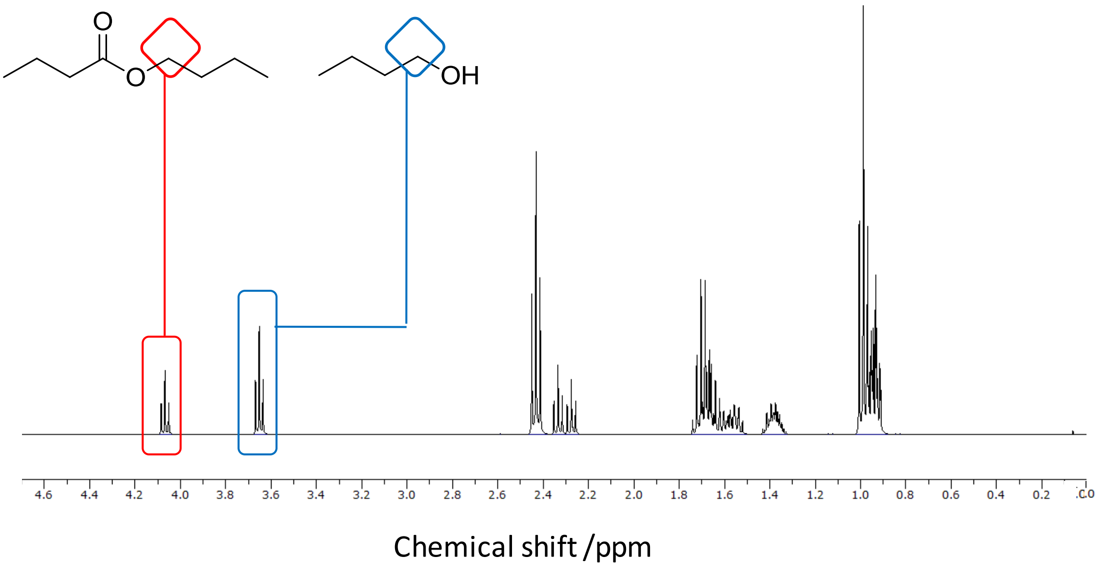


Figure 4.2 A ^{1}H-NMR spectrum of the uncatalysed model esterification occurring in chloroform.

The kinetics of uncatalysed esterification: The model reaction to give butyl butanoate was attempted in the following solvents: acetonitrile, butanone, chlorobenzene, chloroform, cyclohexane, DMF, 1,4-dioxane, and toluene. Weak C-H acids were selected to observe the influence of hydrogen bond donors on the rate of esterification, also adding much needed variety to the polarities covered by the solvent set (Figure 4.3). Obviously alcoholic solvents would tend to interfere with the reaction, with the exception of *t*-butanol perhaps.

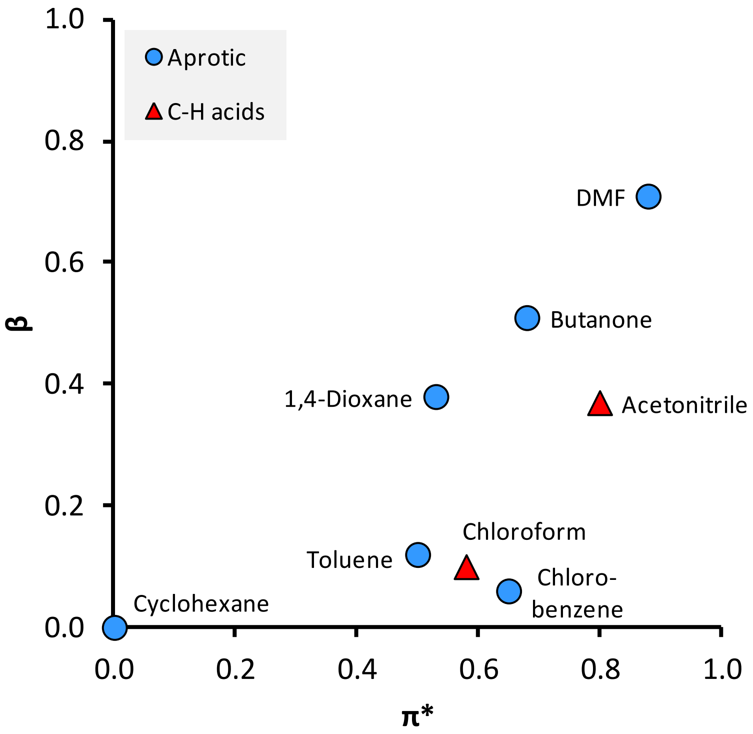


Figure 4.3 Polarity map of solvents used in the initial uncatalysed model esterification solvent set.

As expected from previous work concerning amidation, the aromatic and chlorinated solvents provide a superior reaction medium to oxygenated solvents (Table 4.2). Unfortunately, as we have come to know, this generally creates a trade-off between solvent greenness and solvent performance, although plenty of members of the solvent set are toxic, unsafe, and damaging to the environment regardless of their performance in the reaction (Figure 4.4). 1,4-Dioxane is especially bad in this regard within the SUS-HAS-ECO classification framework, closely

Table 4.2 Uncatalysed model esterification rate constants and solvent properties.

Solvent	ln(**k**)	α	β	π*	δ_T
Acetonitrile	-11.81	0.35	0.37	0.80	24.2
Butanone	-11.46	0.00	0.51	0.68	19.0
Chlorobenzene	-9.73	0.00	0.06	0.65	19.4
Chloroform	-9.81	0.20	0.10	0.58	18.9
Cyclohexane	-9.18	0.00	0.00	0.00	16.8
DMF	-12.64	0.00	0.71	0.88	24.0
1,4-Dioxane	-10.88	0.00	0.38	0.52	20.5
Toluene	-9.68	0.00	0.12	0.50	18.2

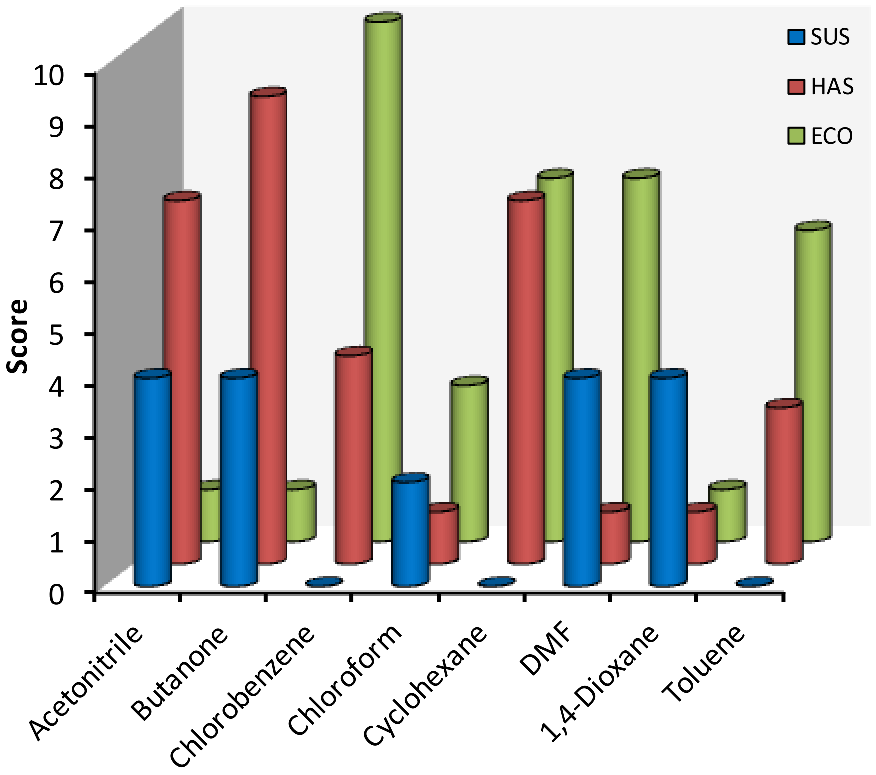


Figure 4.4 Uncatalysed model esterification solvent SUS-HAS-ECO classifications.

followed by chloroform. The former is consistently low scoring throughout the GSK solvent selection guide and chloroform is considered toxic whilst also presenting waste issues [Henderson 2011]. The two best solvent options, all things considered, appear to be cyclohexane and chlorobenzene, leaving room for improvement. Cyclohexane is able to do what *n*-octane could not in the model amidation reaction and support the reaction, providing the greatest reaction rate acceleration of the eight solvents. Of the potential mechanisms presented in Table 4.1, the kinetic benefit of a non-polar solvent in this reaction is only consistent with the tetrahedral intermediate $A_{AC}2$ mechanism.

As anticipated the relationship between solvent polarity and ln(**k**) is inversely correlated to $\boldsymbol{\beta}$ (Figure 4.5). An exception is acetonitrile, which underperforms based on the expectation of its polarity. This does not appear to be caused by its modest $\boldsymbol{\alpha}$ value because chloroform adheres to the trend set by the rest of the solvents. Dipolarity and polarisability (as represented by $\boldsymbol{\pi^*}$ and $\boldsymbol{\delta}$) along with $\boldsymbol{\alpha}$ were statistically insignificant. The LSER (ignoring acetonitrile for now) is as follows:

Equation 4.1 $$\ln(\mathbf{k}) = -9.25 - 4.57\boldsymbol{\beta} \qquad R^2 = 0.987$$

Error bars indicating the range of ln(**k**) values obtained in each solvent have been marked on Figure 4.5, but none are longer in length than the data points representing the average value of at least three experiments. The precision of monitoring the reaction by NMR spectroscopy alleviates fears that the outlying acetonitrile is caused by experimental error.

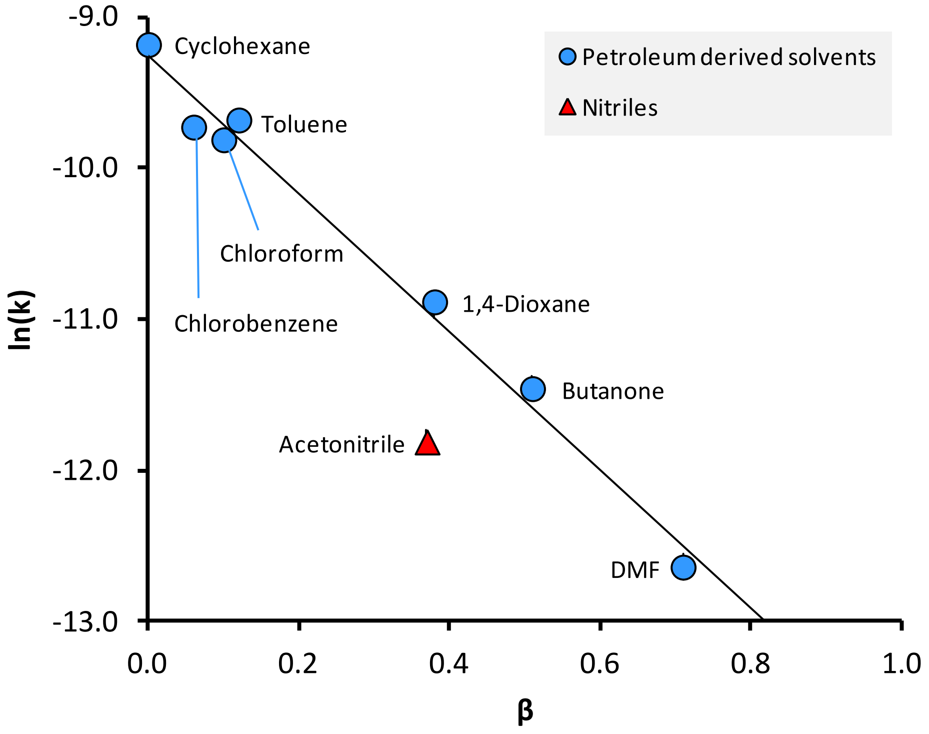


Figure 4.5 The LSER describing the uncatalysed model esterification but excluding acetonitrile.

The notable anomaly (that being acetonitrile) was not present in the amidation case study, and so this deviation from ideality, less efficient at promoting the reaction than expected from a trend based on **β** alone, cannot yet be properly described. Anhydrous grade acetonitrile did not offer improved results to other products, nor did the addition of drying agents. For clarification propanenitrile (**β** = 0.39) was used as a solvent too, but behaved as predicted by estimation with Equation 4.1. Both nitriles have similar **β** values and would be expected to perform near equally as well. The difference in the experimentally determined rate constants is too large to be casually attributed to error (Figure 4.6). This suggests that acetonitrile is a special case and not an effect arising from the nitrile functional group. Instead, it might be concluded that another parameter, as of yet unaccounted for, is jointly responsibly in combination with **β** for the observed rates of reaction.

The introduction of the Hildebrand solubility parameter ($\boldsymbol{\delta}_T$) rectifies the discrepancy between the nitrile solvents. Accounting for the cohesive energy density of the solvent by taking the square of this parameter, a negative coefficient for $\boldsymbol{\delta}_T^2$ signifies that strongly self-associating solvents, of which acetonitrile and DMF are clearly more inclined towards than the other solvents

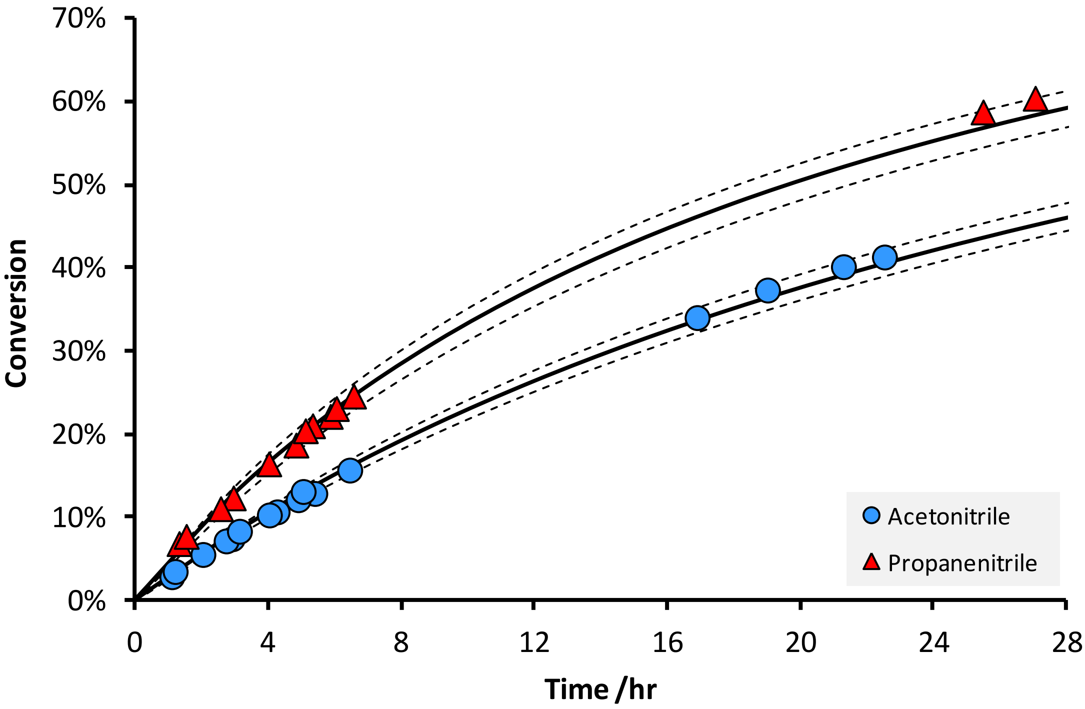


Figure 4.6 Conversions to butyl butanoate in acetonitrile and propanenitrile at 323 K. The dashed lines indicate ±1 standard deviation.

in this set, are disadvantaged with respect to accelerating the rate of reaction:

Equation 4.2
$$\ln(\mathbf{k}) = -8.06 - 3.48\boldsymbol{\beta} - 0.00391\boldsymbol{\delta}_{\mathbf{T}}^{2} \qquad R^2 = 0.988$$

The upgraded LSER (Equation 4.2) appears to have a small coefficient of $\boldsymbol{\delta}_{\mathbf{T}}^{2}$, which raises questions over its necessity. It should be remembered that $\boldsymbol{\delta}_{\mathbf{T}}^{2}$ is itself typically three orders of magnitude larger than the Kamlet-Taft solvent polarity parameters and so LSER coefficients will be proportioned appropriately to account for this. Importantly both **β** and $\boldsymbol{\delta}_{\mathbf{T}}^{2}$ are statistically significant, and only when they are combined, as in Equation 4.2, is a satisfactory predictive element to the resulting LSER achieved (Figure 4.7).

The simplest explanation for the partial dependence of the reaction rate on $\boldsymbol{\delta}_{\mathbf{T}}$ can be thought of as the activated complex of the reaction occupying a larger volume than the sum of the reactants in solution (Scheme 4.3) [Taft 1985]. The implication being that some solvent-solvent interactions must be broken, and if these are strong intermolecular bonds the reaction becomes less favourable than what otherwise would be appreciated from the Kamlet-Taft solvent polarity parameters alone. In a more precise sense, it may not strictly be the size of the cavity occupied by the reaction components that is relevant but the relative quantities of solvent-solvent interactions in the vicinity that are lost or gained as the reaction progresses. The variation of $\boldsymbol{\delta}_{\mathbf{T}}$ in the initial solvent set of the related amidation reaction was sufficient to highlight its role,

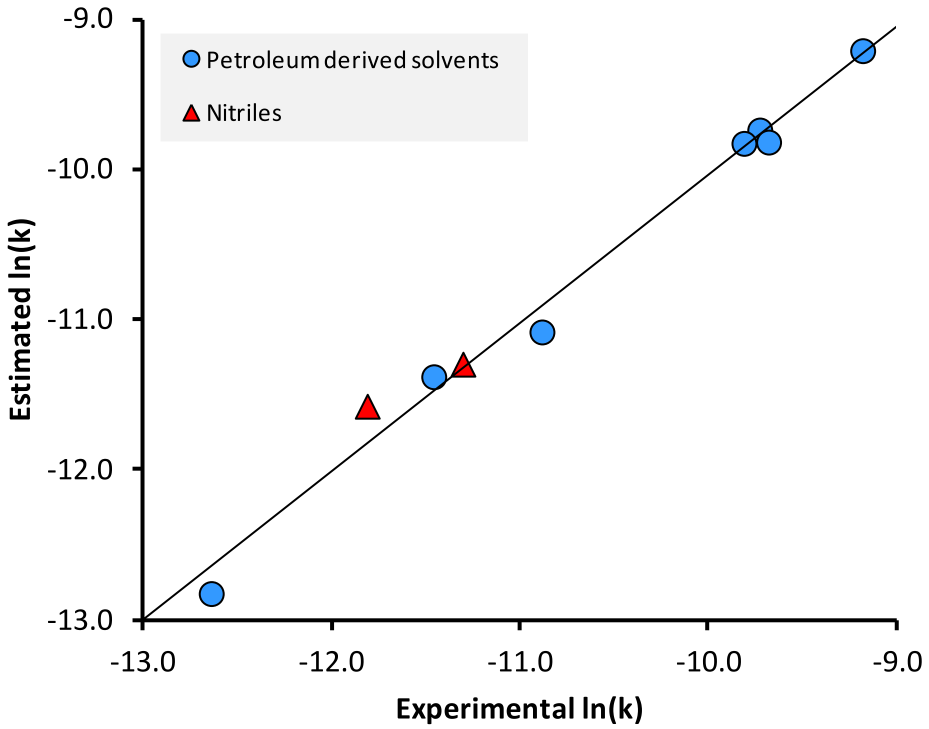
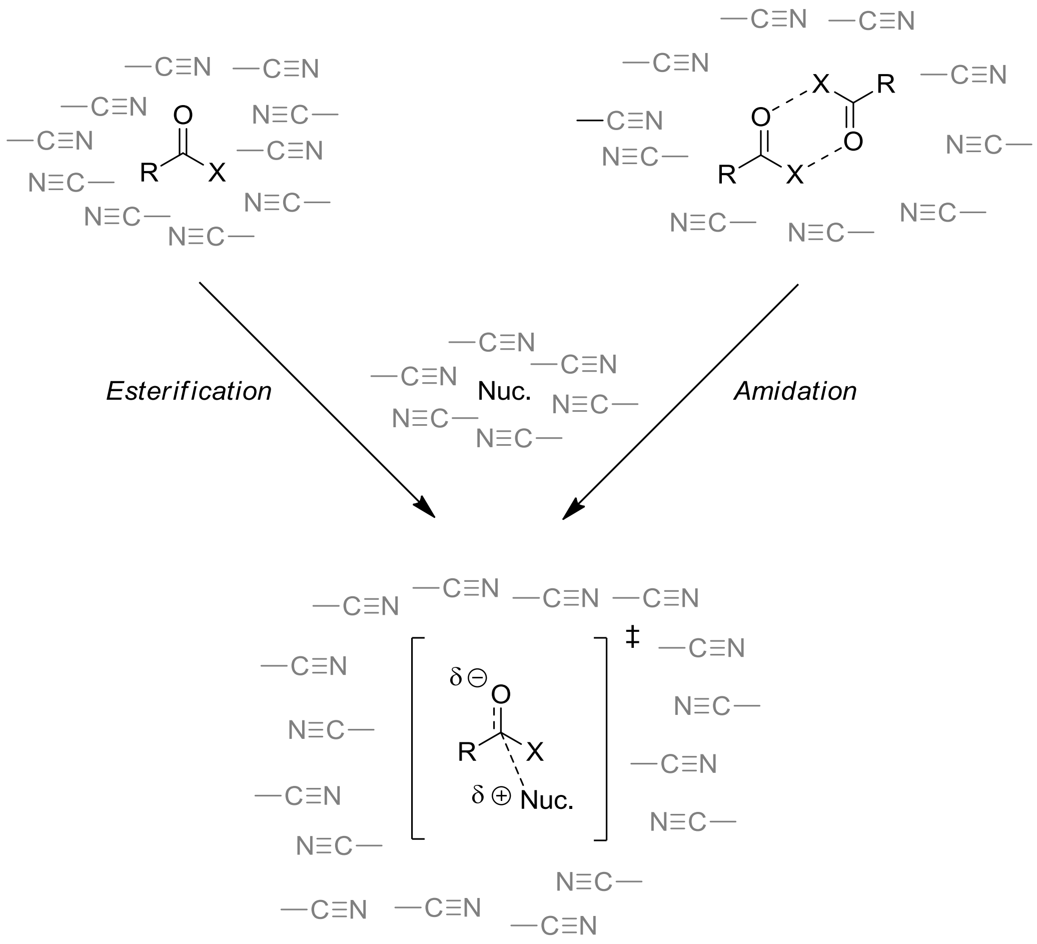


Figure 4.7 The estimated ln(**k**) values of the uncatalysed model esterification, including nitrile solvents, compared to experimental data.

if any, in the kinetics of the reaction. The fact that it was found to have no influence means that the activated complex of the amidation should be approximately as voluminous as the two reactants. If hydrogen bonded pairs of reactants are considered as the precursors to amidation reactions (Scheme 3.4), the experimental data requires a similar sized solvent cavity to that made by the activated complex. The reactants of the uncatalysed model esterification (an alcohol and a carboxylic acid anhydride) do not form strong hydrogen bonds between themselves. If the tetrahedral activated complexes of esterifications are assumed to require a similar volume within the bulk solvent to that of analogous amidations, then the individual esterification reactants must reside in a smaller cavity, or at least permit more intermolecular solvent-solvent bonding, than the reactants in an amidation reaction. If this is the case, creating the activated complex of esterification would result in a significant loss in solvent-solvent bonds, consistent with the observed solvent effect represented with a statistically significant negative coefficient of $\boldsymbol{\delta}_{\mathbf{T}}^{2}$, as in Equation 4.2 (Scheme 4.3).

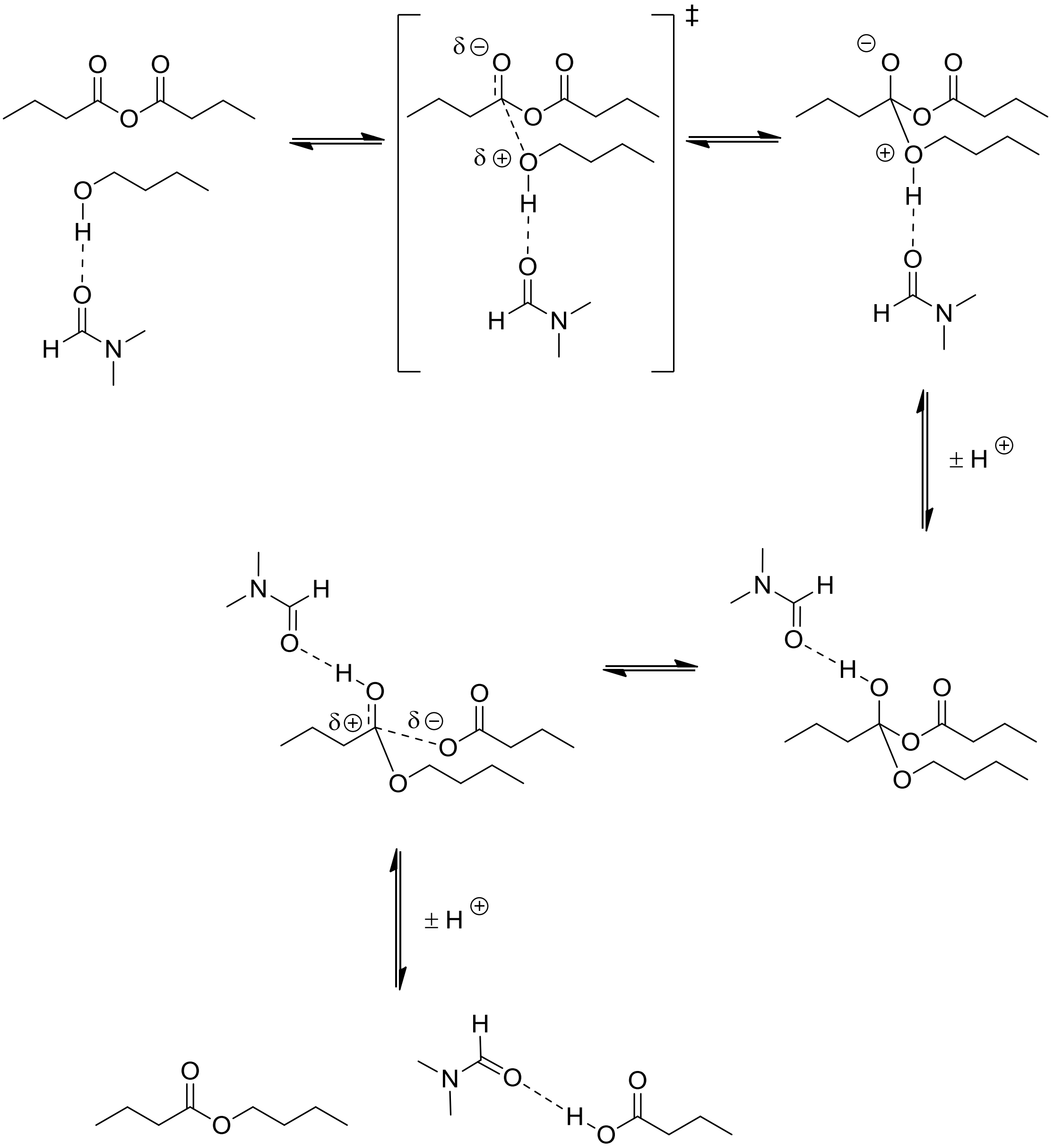
An explanation for these solvent cavity volumes could stem from the previously visited anomaly regarding the relative permittivity of acetic acid (Scheme 1.1). Carboxylic acid dimers have a surprisingly low relative permittivity, appreciably less than even their respective anhydrides or acyl chlorides [Lide 1991, Reichardt 2003 page 472]. This would imply hydrogen



Scheme 4.3 The $A_{AC}2$ mechanism specifically concerning the disturbance to the bulk solvent medium as illustrated with acetonitrile.

bonded solutes demand less interaction with the solvent and a 'looser' solvation sphere results. This could be similar in volume to the cybotactic region of the activated complex of carbonyl addition. The 'tighter' solvation sphere of carboxylic acid anhydrides caused by stronger solvent-solute intermolecular interactions would have to be disrupted to accommodate the activated complex of any reaction undergone by the solutes.

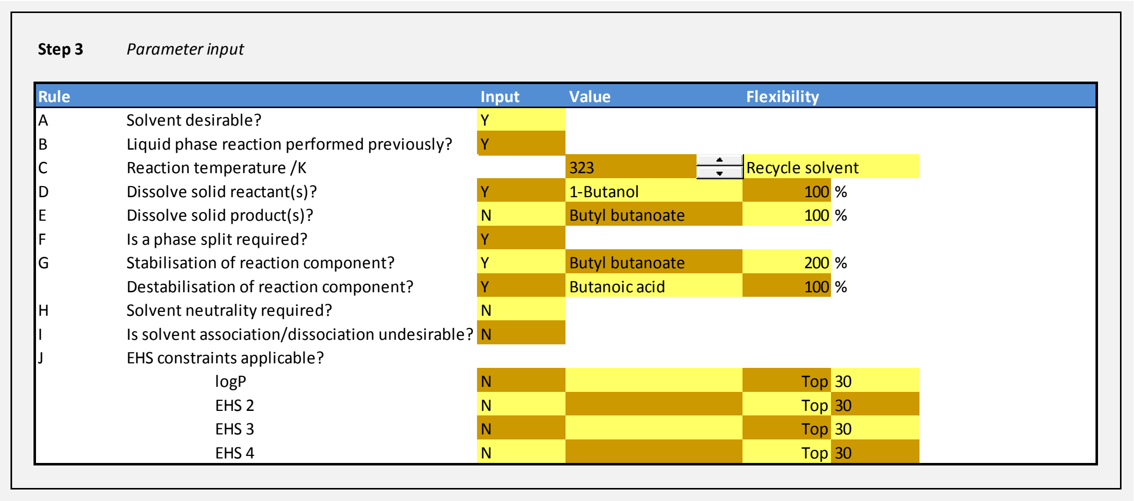
This result implies that, generally speaking, the same type of solvent effect as observed for amide forming reactions is in place for esterifications, dictating the rate of reaction. This requires entropic control to dominate the contribution of enthalpy towards the Gibbs free energy of activation at a much lower temperature than is the case for amidation. Without the full Eyring treatment it is not possible to say if this is true. But it is plausible, depending on the relative magnitude of the activation parameters. The added influence of solvent cohesive energy density would suggest that solvent-solvent ordering is of crucial importance to the rate of reaction in a given solvent, complimenting this hypothesis. These suggestions are consistent with experimentally derived terms for the enthalpy of activation and entropy of activation in various esterifications [Bamford 1972, Bankole 2011 page 69]. These state that the enthalpy of activation is positive, as expected, across a wide range of substrates. Meanwhile the entropy of activation in esterifications is negative, describing a change towards a more ordered system as the activated complex is created from the combining of reactants. Because of this, hydrogen bond accepting solvents might stabilise the activated complex, but in doing so will create additional structural ordering and increase further still the entropic penalty associated with the kinetics of the reaction. A full proposal for the mechanism of $A_{Ac}2$, uncatalysed esterification with an emphasis on solvent effects can now be established (Scheme 4.4). This mechanism does not need to rely on the acidity-blocking prowess of the solvent. Although only one specific solvent-solute



Scheme 4.4 The solvent-solute interactions dictating the rate of uncatalysed esterification as demonstrated with DMF.

hydrogen bond exists throughout, this will get stronger (*i.e*. more stabilisation provided by the solvent) as the reaction precursors become polarised in the activated complex at the high energy transition state.

Solvent selection: In the solvent selection algorithm the polarity characteristics of solvent candidates were targeted so as to stabilise the product butyl butanoate. Destabilisation of the by-product butanoic acid was also demanded from rule G in an effort to perhaps aid separation of the two compounds. A phase split with water was also asked of the solvent because unlike amides, ester products will hydrolyse back to their constituent reactants. Otherwise, aside from the lower reaction temperature, the data input is much the same as it was for the amidation case study (Figure 4.8). This is also true of the solvent performance (model B2) and greenness assessments (model B3), except in the latter the weighting of the GSK solvent selection guide categories have been adapted slightly (Figure 4.9).

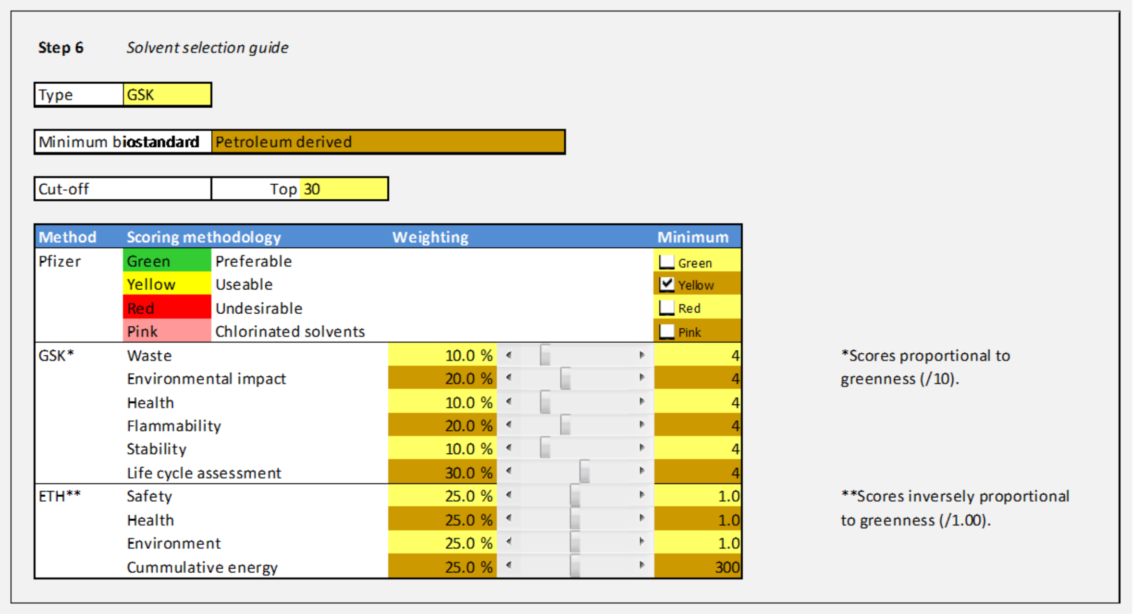


Step 3 *Parameter input*

Rule		Input	Value	Flexibility
A	Solvent desirable?	Y		
B	Liquid phase reaction performed previously?	Y		
C	Reaction temperature /K		323	Recycle solvent
D	Dissolve solid reactant(s)?	Y	1-Butanol	100 %
E	Dissolve solid product(s)?	N	Butyl butanoate	100 %
F	Is a phase split required?	Y		
G	Stabilisation of reaction component?	Y	Butyl butanoate	200 %
	Destabilisation of reaction component?	Y	Butanoic acid	100 %
H	Solvent neutrality required?	N		
I	Is solvent association/dissociation undesirable?	N		
J	EHS constraints applicable?			
	logP	N		Top 30
	EHS 2	N		Top 30
	EHS 3	N		Top 30
	EHS 4	N		Top 30

Figure 4.8 Uncatalysed esterification solvent selection algorithm screenshot, step 3: Parameter input.

Just like the amidation case study, 1,2-dichlorobenzene, chlorobenzene, and cumene are the only three solvents to pass all of the requirements of the revised solvent selection algorithm and the associated models (Table 4.3). The maximum score when the revised solvent selection algorithm (model B1) is combined with the solvent performance assessment (model B2) or the solvent greenness assessment (model B3) is 160 (with a 1:10 weighting between models as before). The returned solvent candidates are much the same as what was recommended for amidation reactions. Unfamiliar additions to the list of solvent hits are due to the lower reaction temperature of the uncatalysed esterification compared to the hotter amidation conditions. After a promising start, cyclohexane failed to meet the requirements of the greenness assessment



Step 6 *Solvent selection guide*

Type: GSK

Minimum biostandard: Petroleum derived

Cut-off: Top 30

Method	Scoring methodology		Weighting	Minimum
Pfizer	Green	Preferable		☐ Green
	Yellow	Useable		☑ Yellow
	Red	Undesirable		☐ Red
	Pink	Chlorinated solvents		☐ Pink
GSK*	Waste		10.0 %	4
	Environmental impact		20.0 %	4
	Health		10.0 %	4
	Flammability		20.0 %	4
	Stability		10.0 %	4
	Life cycle assessment		30.0 %	4
ETH**	Safety		25.0 %	1.0
	Health		25.0 %	1.0
	Environment		25.0 %	1.0
	Cummulative energy		25.0 %	300

*Scores proportional to greenness (/10).

**Scores inversely proportional to greenness (/1.00).

Figure 4.9 Uncatalysed esterification solvent selection algorithm screenshot, step 6: Solvent selection guide.

(model B3). This was because of a low flammability score. Perfluoroalkane solvents appear in Table 4.3, and after their success in facilitating esterifications this is not surprising [Otera 2005]. However their greenness scores in the GSK solvent selection guide are their downfall, with no data available for perfluoromethylcyclohexane, and perfluorohexane scoring poorly throughout. Critically, a score of 3 in the health category, and no LCA data resulted in a fail for perfluorohexane. Both limonene and *p*-cymene appear as recommended solvent candidates in the solvent performance assessment (model B2), and it seemed fitting that they should be applied in the uncatalysed esterification case study.

Application of bio-derived solvents in uncatalysed esterification: As with the model amidation, limonene and *p*-cymene would appear to be excellent solvent candidates for esterification reactions. The experimentally determined rates of reaction are very high (Table 4.4). Limonene and *p*-cymene follow the trend dictated by $\boldsymbol{\beta}$ and $\boldsymbol{\delta}_T$ as expected (Equation 4.2). Updating the LSER with these solvents gives the following equation:

Equation 4.3 $$\ln(\mathbf{k}) = -8.13 - 3.45\boldsymbol{\beta} - 0.00379\boldsymbol{\delta}_{\mathbf{T}}^{2} \qquad R^2 = 0.989$$

Limonene appears to be slightly superior to cyclohexane as a reaction medium, although with identical $\boldsymbol{\beta}$ values this is not a significant improvement. Nevertheless limonene (and *p*-cymene not far behind) again demonstrates the ability of bio-based solvents to easily substitute petroleum derived solvents whilst maintaining high rates of reaction.

Table 4.3 Uncatalysed esterification solvent hits from the solvent selection algorithm.

Solvent	Score: Model B1 + model B2	Score: Model B1 + model B3
1,1,1-Trichloroethane	152	No data
1,2,4-Trichlorobenzene	No data	128
1,2-Dichlorobenzene	128	128
1,2-Dichloroethane	114	Fail
Benzene	112	Fail
Carbon tetrachloride	114	Fail
Chlorobenzene	110	130
cis-Decalin	110	Fail
Cumene	86	86
Cyclohexane	152	Fail
Cyclohexanone	Fail	130
Cyclopentanone	Fail	108
Diphenyl ether	82	No data
Fluorobenzene	114	Fail
Heptane	150	Fail
Hexane	154	Fail
Limonene	148	No data
Mesitylene	90	Fail
m-Xylene	90	No data
p-Cymene	88	No data
Perfluorohexane	154	Fail
Perfluoromethylcyclohexane	152	No data
p-Xylene	90	Fail
Tetrachloroethylene	130	No data
Toluene	90	Fail
Trichloroethylene	134	No data
Total hits	23	6

Table 4.4 A comparison between high performance bio-based and petro-chemical esterification solvents.

Solvent	ln(**k**)	**α**	**β**	**π***	δ_T	SUS	HAS	ECO
PhCl	-9.73	0.00	0.06	0.65	19.4	1	4	10
p-Cymene	-9.38	0.00	0.13	0.39	17.4	6	n/a	n/a
Limonene	-9.15	0.00	0.00	0.16	15.1	10	n/a	n/a
Toluene	-9.68	0.00	0.12	0.50	18.2	1	3	6

Both the citrus waste derived solvents can be incorporated into the kinetic LSER (Equation 4.3) and represented successfully in terms of the correlation between experimental and predicted ln(**k**) values (Figure 4.10). There does not appear to be any systematic error, although the data scatter is slightly greater than that seen in the amidation and nucleophilic substitution reaction case studies (Figure 4.11).

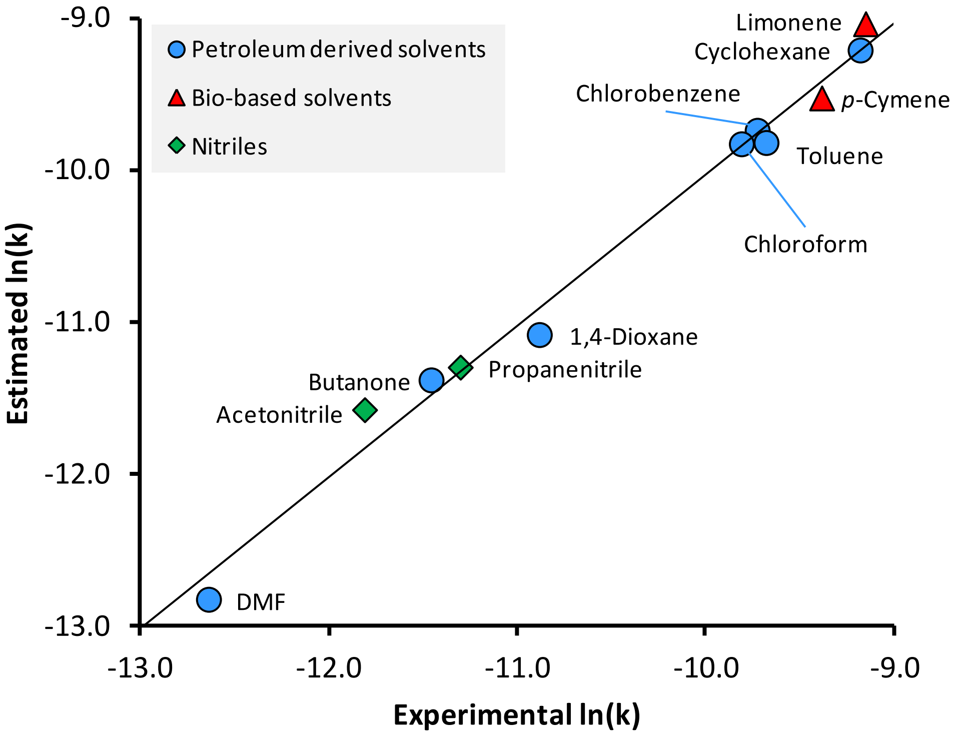


Figure 4.10 The estimated ln(**k**) values from the LSER describing the uncatalysed model esterification, including bio-based solvents, plotted against experimental data.

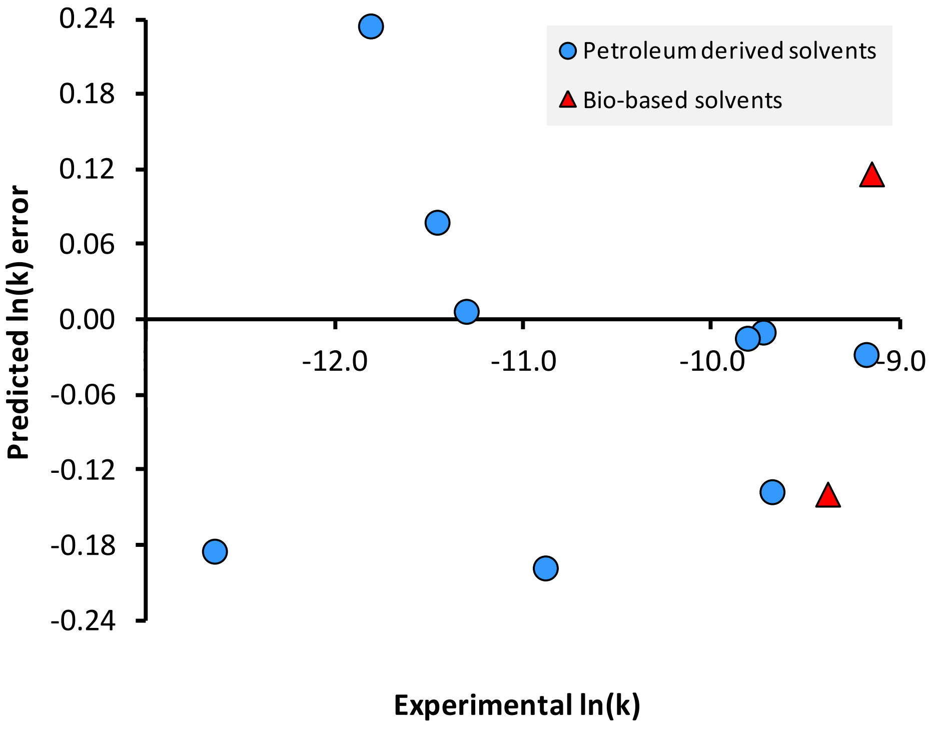


Figure 4.11 Systematic error check regarding the certainty of rate constant estimation in the uncatalysed model esterification.

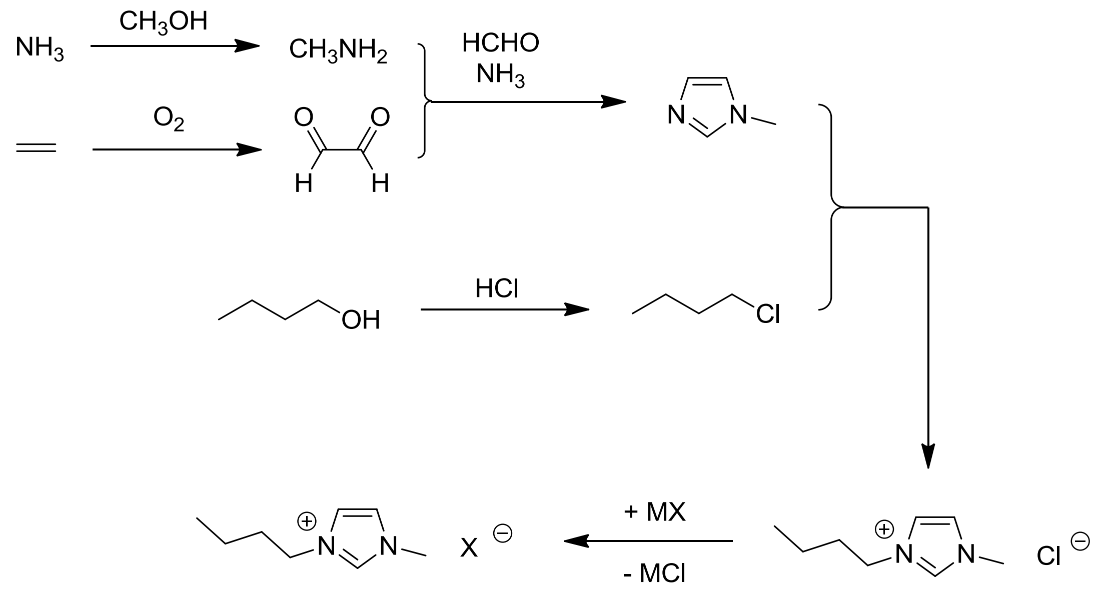
Unsurprisingly the same arguments for and against the introduction of bio-based solvent can be raised as before in Chapter 3. A lack of data concerning health, safety, and environmental considerations is not acceptable just because the product is renewable. Caution must be applied in this circumstance. Having said this, the prevalence of limonene in consumer products is reassuring of its low potential to be toxic, but aquatic toxicity data suggests limonene pollution is a burden on the environment. Due to its lipophilicity, and resulting aquatic toxicity and bioaccumulation potential, limonene may not be a holistically beneficial solvent substitute in all instances. Of course the damage that limonene can do to its environment depends on the way in which it is handled. If disposal is carried out correctly, or better still, the solvent is recycled, then this argument loses some of its power. If limonene is not being extracted from citrus waste to be used in synthetic chemistry, then it is a burden on the environment regardless as a component of citrus waste, without having been made to work for that privilege.

4.3 Esterification summary

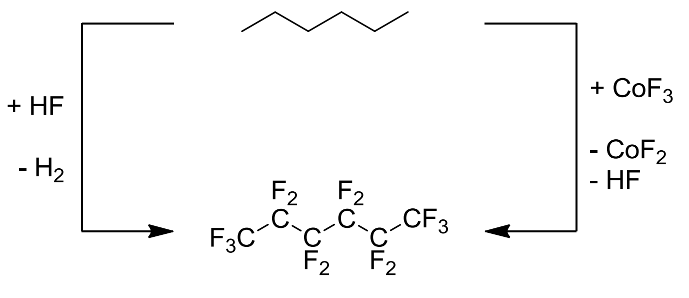
The success and implications of solvent selection: Considering the esterification presented here, and further evidence from other solvent studies on related systems, it appears (perhaps unsurprisingly) that a universal relationship between the rate of $A_{AC}2$ carbonyl additions and solvent hydrogen bond accepting ability is likely to exist. This class of reaction is accelerated by solvents with low **β** values. The explanation for this solvent effect is that the enthalpic benefits of stabilising the activated complex of the reaction are overridden by the large entropic penalty created by organising the solvent in order to establish these interactions. The influence of the Hildebrand solubility parameter on the kinetics of butyl butanoate synthesis is complimentary to the impeding function of solvent hydrogen bond accepting ability (**β**). For these reasons, *p*-cymene can be considered a renewable alternative solvent for carbonyl additions. Limonene or even citrus oils directly can be used as solvents for synthesis, although the reactivity of limonene is sometimes a major and usually unsurpassable barrier to its application as a solvent. *p*-Cymene is also compatible with existing methodologies for managing less routine carbonyl additions. A 1 mol% loading of DMAP in *p*-cymene at 323 K provided complete conversion to butyl butanoate within one hour, as judged by ^{1}H-NMR spectroscopy of the reaction mixture.

It is also worth addressing the relative merits of esterification reactions performed in perfluorinated solvents, ionic liquids, and of course more typical reaction conditions using limonene or *p*-cymene as renewable reaction media. The LCA of typical ionic liquids indicates a relatively large environmental impact for a solvent [Zhang 2008]. Perfluorinated solvents pose a threat in the environment, but also have an ungreen manufacturing process, like ionic liquids, but for different reasons (Scheme 4.5). With fluorine inevitably comes the demand or emission of hydrogen fluoride [Barbour 1952, Pearlson 1986]. A more detailed description of ionic liquids and perfluorinated solvents is available elsewhere [Breeden 2012]. Needless to say, bio-based solvents, accessible in two or less synthetic transformations from a renewable feedstock, will significantly reduce the associated energy and waste issues of chemical manufacture attributed to the products in which they are synthesised compared to ionic liquid or perfluorinated solvent based processes.

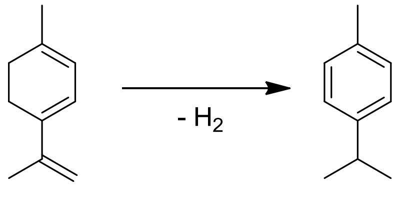
Broader appeal and future work: There are a variety of roles in which limonene and *p*-cymene could be adopted as bio-based alternatives, but some important applications will be off-limits, either because the polarity of these solvents is not suitable or that their high boiling points are irreconcilable with the design of the application. The high boiling points of the citrus oil derived solvents also mean they have no chromatographic use. This is unfortunate given the volumes of solvent required for column chromatography, dwarfing that of the actual reaction solvent.


NH3
CH3OH
CH3NH2
HCHO
NH3
O2
O
O
H
H
N
N
HCl
OH
Cl
+ MX
- MCl
N
N
X
Cl

Synthesis of an imidazolium ionic liquid


+ HF
- H2
+ CoF3
- CoF2
- HF
F2
F2
F3C
C
C
C
C
CF3
F2
F2

Synthesis options for perfluorohexane



Synthesis of p-cymene from limonene

Scheme 4.5 Manufacturing routes to solvents fit for esterification chemistry.

Considering the huge number of solvents in use today, it would be impossible for one or two new bio-based solvents to cover all applications once reliant on petroleum derived solvents. The quest to discover and validate more sustainable solvent solutions will continue indefinitely as scientists seek to improve and investigate solvent phenomena. What is now known is that limonene and *p*-cymene can be applied as solvents in an assortment of commercially attractive acylation chemistries relied on by all chemical sectors. By being of low polarity, both limonene and *p*-cymene help populate a niche in the range of possible solvent polarity that typical oxygenated bio-based solvents do not satisfy (Figure 4.12). Expanding the polarity range of bio-based solvents is probably the strength of the citrus waste derived solvents.

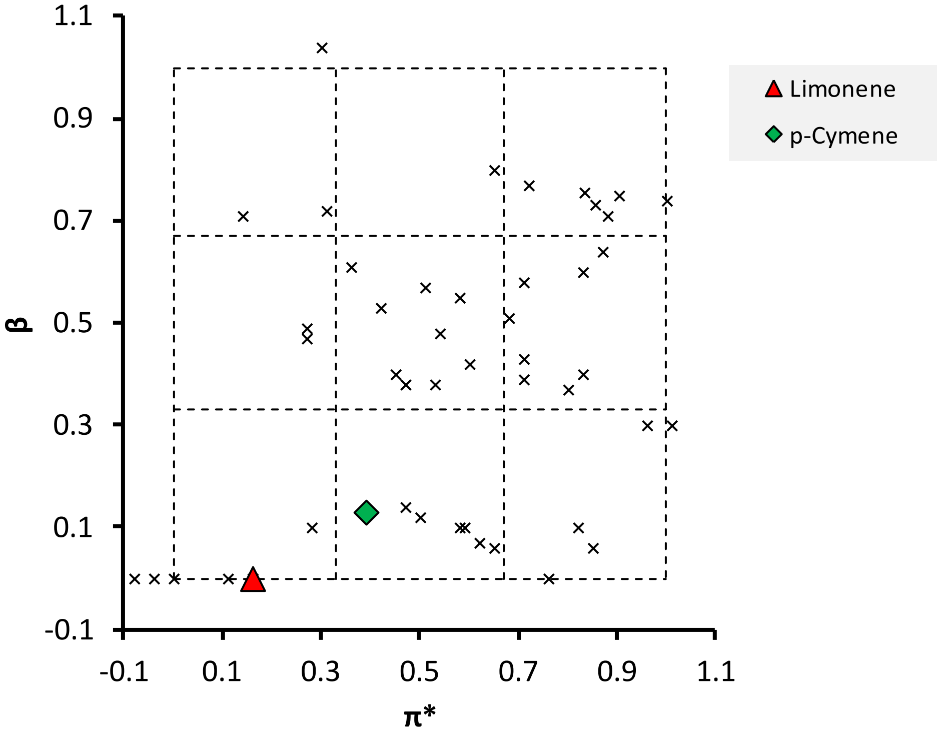


Figure 4.12 A solvent polarity map highlighting limonene and *p*-cymene.

A reasonable criticism of this work is that an abstract model reaction is not necessarily versatile in making a case for a particular solvent selection across a whole class of reaction. The reaction between butanoic anhydride and 1-butanol serves a purpose as an easy to follow kinetic study, but bears little resemblance to typical esterification strategies. The solvent effects governing a Fischer esterification would be more useful in this respect, and a study of this kind is already available [Wells 2008]. The uncatalysed esterification developed here partners this existing research and reveals that the role of the acid catalyst in Fischer esterification is not necessarily dependent on the solvent. Whether limonene and *p*-cymene can also be used in catalysed reactions is investigated in the following chapter, where the range of esterifications studied is significantly broadened, as is the capacity of limonene as a platform molecule.

5. Catalysed carbonyl addition

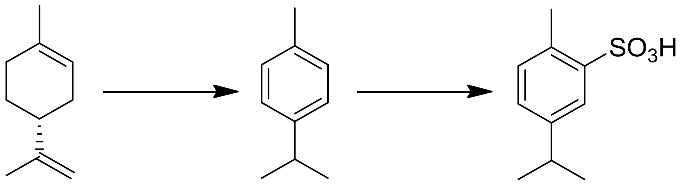
The imposing solvent effects discussed in the previous chapters may conjure a likeness between the rate enhancements provided by solvents and those imparted by catalysis. In fact solvents will also modify reaction equilibria in addition to adapting kinetic profiles. Solvent effects are not equated to catalytic effects because solvents are applied in a large excess and are responsible for the environment surrounding reaction components. Although solvent effects often involve a specific solute-solvent interaction the effect is usually not attainable if the solvent were to be applied in catalytic amounts. However it is clear that when optimising a reaction system the choice of solvent and catalyst is important for similar reasons, and to a certain extent the performance of each is dependant on the other. It is sensible to examine how all the reaction auxiliaries perform in combination, and which precise combination is optimal. The bio-based content and suitability of catalysts can be scrutinised in the same way that solvents have been treated thus far.

5.1 Bio-based acid catalysts for organic chemistry

Acid catalysis in Fischer esterification: All the Fischer esterification methods and mechanisms discussed previously require an acid catalyst. Bases unfavourably deprotonate the carboxylic acid and are only useful for ester hydrolysis. The favoured choice of acid catalyst for the manufacturing of typical ester bulk chemicals is either a mineral acid or quite often *p*-toluenesulphonic acid (*p*-TSA) [Smith 2007, page 1414]. Alternatively a number of heterogeneous catalysts have recently been applied to esterifications in an attempt to reduce waste [Barbosa 2006, Kirumakki 2004, Sawant 2007, Shanmugam 2004]. Lewis acids are presented as less corrosive alternatives to *p*-TSA and other Brønsted acids [Chakraborti 2003, Chandra 2002, Chandrasekhar 2002a, Corma 2003, Mihara 2010, Orita 2001]. Alternatively activated carboxylic derivatives can be used; either generated *in situ* or as isolated intermediates. This approach obviously mirrors the attitude widely adopted in the synthesis of amides, and in the previous chapter regarding uncatalysed esterification with acid anhydrides. Acid chlorides are often used too, and also dialkylcarbodiimide coupling agents [Smith 2007 page 1417]. Both solutions either avoid the formation of, or remove the stoichiometric water formed during the acylation. But in order to do so atom economy is sacrificed.

Introducing *p*-cymenesulphonic acid: The consensus amongst contempory chemists and chemical engineers is that an increased use of heterogeneous, non-corrosive, recyclable catalysts is a good approach for the future of sustainable chemistry [Kaneda 2006, Martin 2002 page 321]. Heterogeneous catalysis is already prevalent in the bulk chemical manufacturing sector [Tanabe 1999]. However the more delicate speciality chemicals produced on a smaller scale are more frequently produced with the assistance of homogeneous catalysis. Homogeneous catalysis is ubiquitous throughout the various stages of pharmaceutical research and development. Although heterogeneous catalysts can offer greater thermal stability, less contamination of products and improved reusability, this practice is not established in the fine chemical sector, unable as of yet to broadly match homogeneous catalysis [Hagen 2006 page 10].

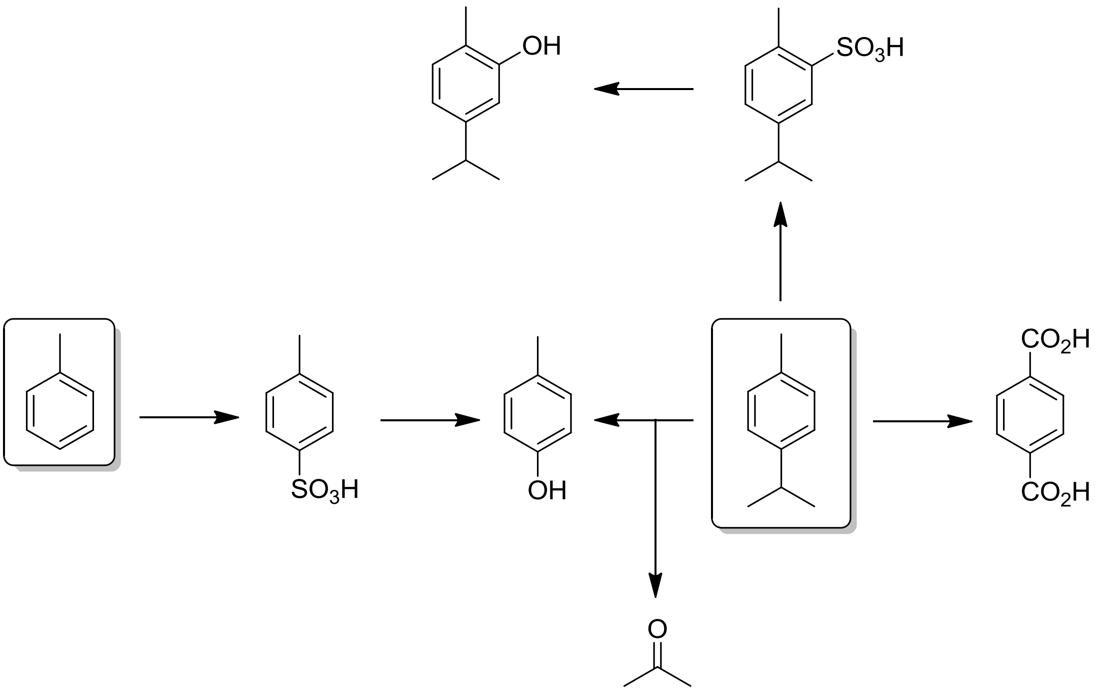
As a strong acid that is also soluble in organic solvents, *p*-TSA is widely used as a Brønsted acidic catalyst for Fischer esterifications and similar reactions [Baghernejad 2011]. Historically *p*-TSA has been synthesised by the sulphonation of toluene as an intermediate in the production of *p*-cresol [Hoff 1979 page 268]. Because this valuable process is based on a petrochemical substrate it is beneficial to seek a sustainable alternative. Having suggested that *p*-cymene made from limonene is a capable solvent replacement for toluene in carbonyl addition chemistry, by extension it seems probable that *p*-TSA could be supplanted for the sulphonic acid of *p*-cymene. Treatment of *p*-cymene with sulphuric acid will give *p*-cymene-2-sulphonic acid (*p*-CSA) (Scheme 5.1) [Hixson 1918, Phillips 1924].



Scheme 5.1 The synthesis of *p*-CSA from limonene *via* *p*-cymene.

The sulphonation of *p*-cymene contained within sulphite turpentine provided a means of obtaining carvacrol in the past [Hixson 1918]. The purpose of making *p*-CSA has never been for catalyst applications, and so the work described here is the first time that *p*-CSA has been used as an acid catalyst. Sulphonation occurs primarily at the 2-position relative to the methyl group, and this isomer can be selectively recrystallised from the reaction mixture [Le Fèvre 1934, Hixson 1918, Phillips 1924, Schorger 1918, Spica 1881]. To obtain the alternative 3-isomer a series of protection strategies need to be employed to block the less sterically hindered 2-position during sulphonation [Philips 1920].

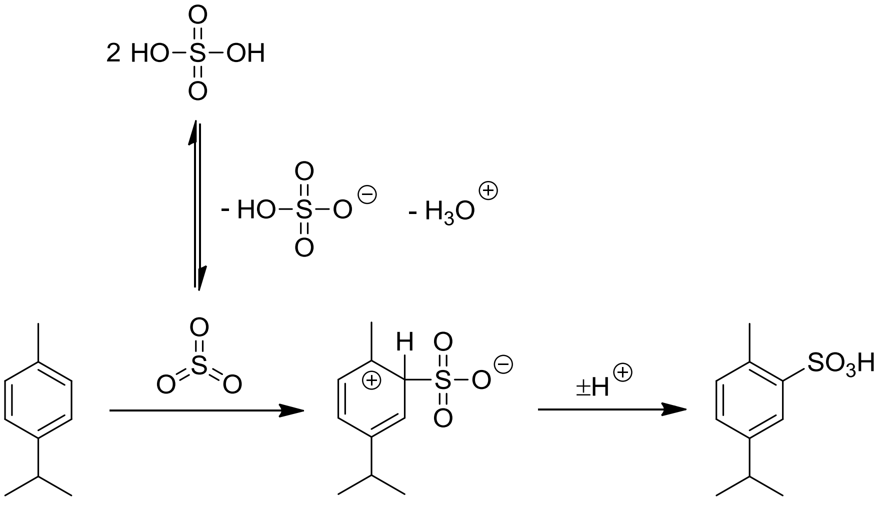
This all contributes to the argument in favour of replacing toluene across all of its current diverse uses with *p*-cymene as far as can possibly be achieved. The greater and broader the demand, the sooner and more economically viable the supply is likely to become. The most important of these substitutions might be the synthesis of terephthalic acid from *p*-cymene, replacing *p*-xylene in this instance. This is currently being explored as part of the movement towards 100% bio-based poly(ethylene terephthalate) (PET) [Berti 2010]. Although outside of the scope of this work, this would probably be the greatest driver in establishing *p*-cymene as a platform molecule. This network of chemical transformations can be likened to that which stems from toluene as a platform molecule (Scheme 5.2). Both substrates can be used to synthesise *p*-cresol [Shinohara 1973], which in turn can be alkylated to give the antioxidant BHT [Hoff 1979 page 268].



Scheme 5.2 A comparison between the synthetic routes linking the products of toluene and *p*-cymene.

In light of previous discussions, an alternative process for obtaining *p*-CSA beginning with the oxidation of the limonene present in the peel of citrus fruits is more suitable for the purposes of this investigation [Martin-Luengo 2008, Martin-Luengo 2010]. Although this route was alluded to in the preceding chapters, currently the production of *p*-cymene is the result of the alkylation of toluene [Simons 1944, Zupp 2012]. This is despite attempts at the synthesis of *p*-cymene from limonene being high yielding on a laboratory scale and the toluene alkylation route not being selective towards the *para*- isomer [Hoff 1979 page 251]. It is the aim of this chapter to demonstrate the synthesis of *p*-CSA from a renewable citrus waste feedstock and its application as an organic acid catalyst. Along the way solvent effects and the interplay between solvent and catalyst will be discussed when the need arises.

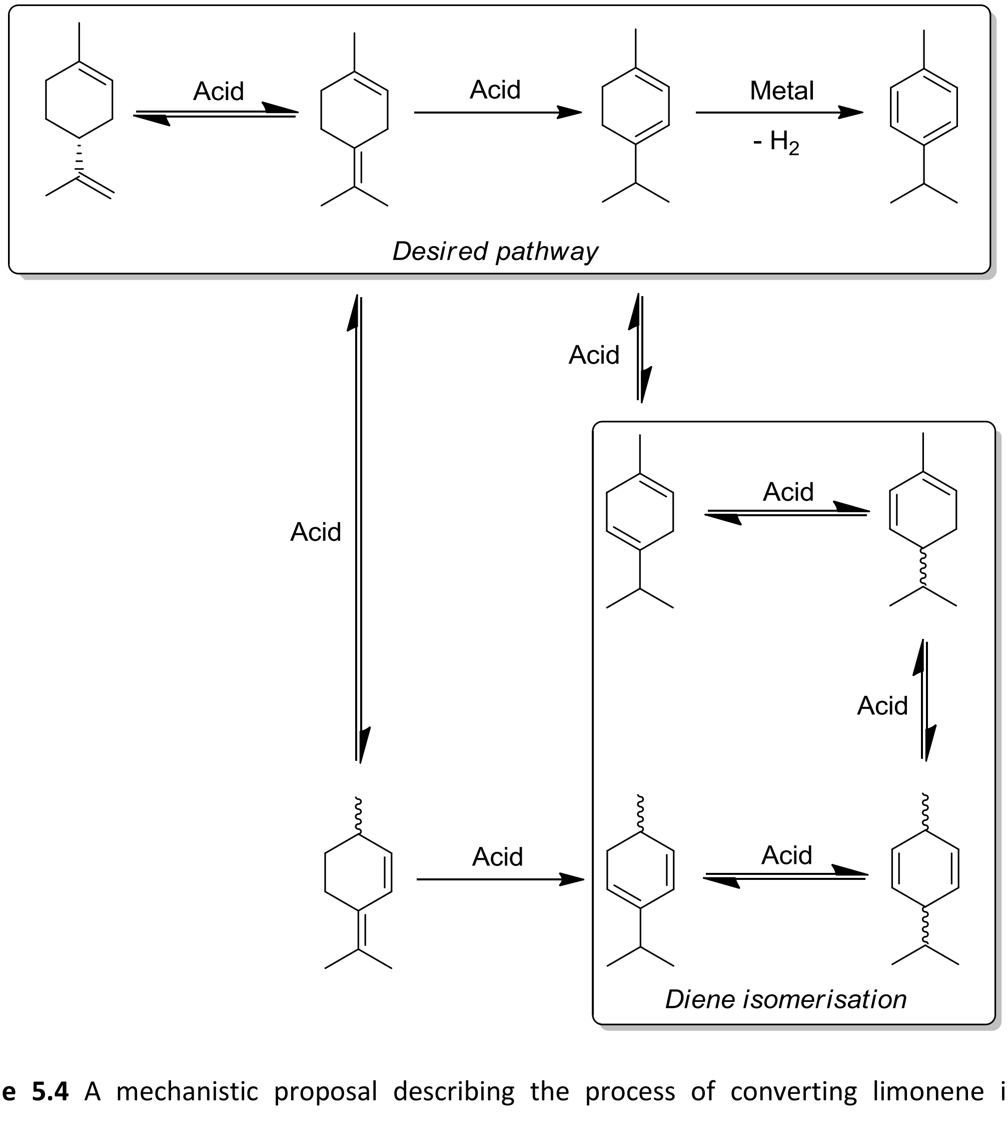
Optimisation of the processes required for the synthesis of *p*-CSA: Given the success of experiments in the literature showing the production of bio-derived *p*-cymene, the transformation of *p*-cymene into its sulphonic acid was explored to obtain the desired product before a complete synthesis from citrus waste was attempted. The substrate *p*-cymene was purchased from a chemical supplier (Sigma-Aldrich), and as mentioned is likely to be manufactured from toluene rather than limonene or another terpenoid source. The action of concentrated sulphuric acid on *p*-cymene at 373 K for 4 hours produced an unsatisfactory 30% yield of *p*-CSA. The active electrophile in this reaction is actually sulphur trioxide (which can be thought of as dehydrated sulphuric acid) or higher dehydrated species, and so a better source of the electrophile was sought to improve the yield (Scheme 5.3) [Smith 2007 page 695]. The use of 20 w/w% fuming sulphuric acid at room temperature increased the sulphonation yield significantly [Le Fèvre 1934, Schorger 1918]. The crude product could be isolated from the reaction by the addition of water, which induces the solidification of the product [Hixson 1918, Schorger 1918]. Recrystallisation from concentrated hydrochloric acid results in a white crystalline solid, previously described in the literature as the dihydrate of *p*-CSA in yields exceeding 90% of the theoretical maximum. As expected only the 2-isomer was present after recrystallisation, fitting reports from previous syntheses, which could be characterised by two



Scheme 5.3 The mechanism of *p*-cymene sulphonation by sulphuric acid.

dimensional NMR spectroscopy (HSQC and HMBC).

It was then wise to confirm that *p*-cymene could indeed be produced from citrus waste. Optimisation of the procedure was in the first instance conducted with limonene rather than citrus oil. The mechanism of this transformation proceeds *via* two distinct stages (Scheme 5.4). Firstly isomerisation of the exo-cyclic double bond of limonene results in α-terpinene amongst its other isomers [Martin-Luengo 2008]. Earlier reports have stated that the isomerisation of limonene to terpinolene is rapid, then the isomerisation onto α-terpinene, although slower, is considered irreversible [Derfer 1979 page 720]. This process is followed by dehydrogenation to give the oxidised product *p*-cymene.



Scheme 5.4 A mechanistic proposal describing the process of converting limonene into *p*-cymene.

It was found that when mineral acids come into contact with limonene they cause its decomposition, probably into polymerisation products as suggested by the resulting black tar.

Instead the isomerisation was promoted by K-10 montmorillonite, an acidic clay known for its ability to isomerise terpenes [Frenkel 1983]. Stirring a wet slurry of limonene and K-10 montmorillonite at 373 K for one hour resulted in only trace amounts of aromatic products. Instead a variety of olefinic species were formed by acid catalysed alkene isomerisation. A slurry of *p*-cymene and K-10 montmorillonite subjected to the same conditions did not affect any decomposition, suggesting that the catalyst provides poor selectivity towards aromatic products and not that the product is unstable.

The use of palladium on activated carbon (or palladium acetate) also failed to convert limonene to *p*-cymene under similar conditions. However it was found that palladium catalysts will oxidise α-terpinene, the key intermediate in the synthesis of *p*-cymene from limonene, to the desired product (Scheme 5.4). The chloride salts of other metals were also screened, including copper, zinc, iron, nickel, and platinum. Polymerisation catalysts, such as aluminium trichloride, that are used to make terpeniod resins were avoided [Derfer 1979 page 749] Although they all showed some evidence of catalysing the dehydrogenation of α-terpinene, none of the metal salts compared to palladium, often providing less than complete conversion of the substrate and a variety of products. By contrast, 5 mol% of palladium supported on activated carbon was able to convert all of the α-terpinene at 413 K, and largely into *p*-cymene (Figure 5.1).

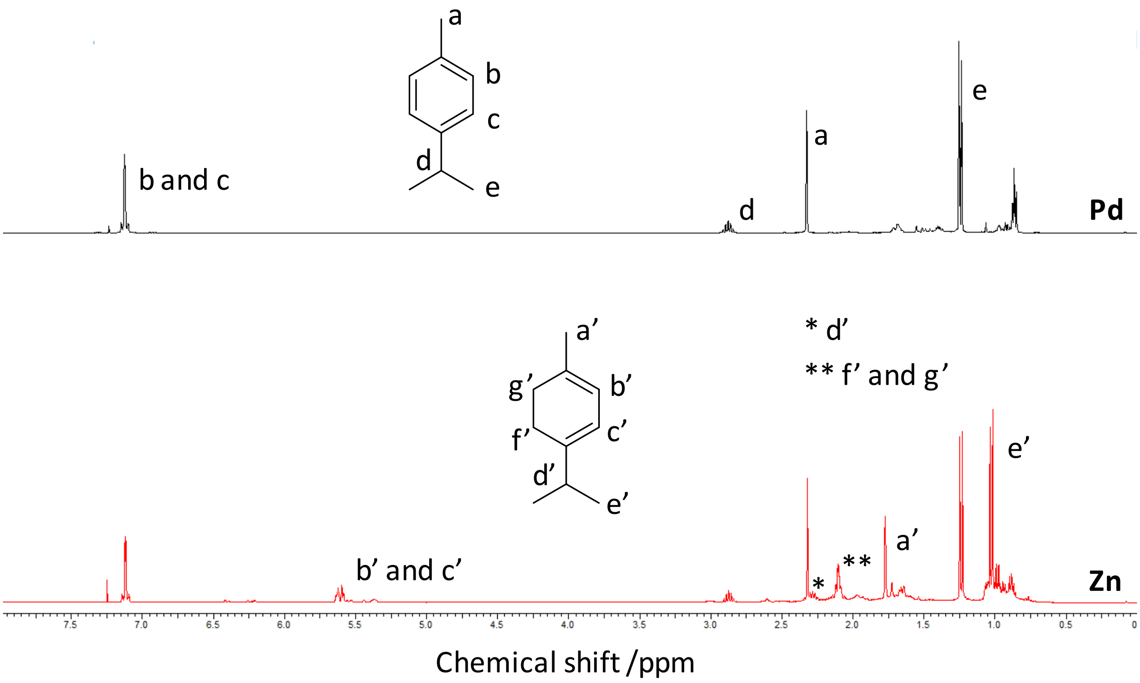


Figure 5.1 The comparison between Pd/C and $ZnCl_2$ catalysed oxidations of α-terpinene by ^{1}H-NMR spectroscopy.

It is apparent that the first stage of the transformation from limonene to *p*-cymene can be achieved with a mildly acidic catalyst to generate the intermediate olefins. It is then necessary that a metal abstracts dihydrogen from the intermediate α-terpinene to give *p*-cymene. It appears that only a combination of these catalyst types will suffice to see this mechanistic pathway to fruition (Scheme 5.4). That is not to say that another route circumnavigating these stages is not feasible, but harder to imagine. A suitable catalyst, the so-called mechanical mixture of palladium on activated carbon (Pd/C) and K-10 montmorillonite can be created simply by shaking a vial containing both components until homogeneity is apparent by eye. This combination of solid catalysts has been used before, and again dehydrogenation was a key step [Kulkarni 2009]. Limonene was cautiously added dropwise to the catalyst mixture, preheated to 373 K. An acid loading of at least 1 mol% (based on the number of acid sites in the clay) was required to obtain any *p*-cymene at all, which corresponds to 0.12 g of clay per millilitre of limonene [Gonçalves 2008]. Raising the acid fraction higher had no effect on the yield (Figure 5.2). An increase in selectivity could be obtained by increasing the quantity of 10 wt% Pd/C up to 1 mol% but further palladium had no significant effect, especially when considering the monetary price associated with doing so. The loading of each catalyst component can therefore be applied at levels equating to 1 mol% each in the mechanical mixture to give a little better than 50% conversion to *p*-cymene at 373 K within 1 hour.

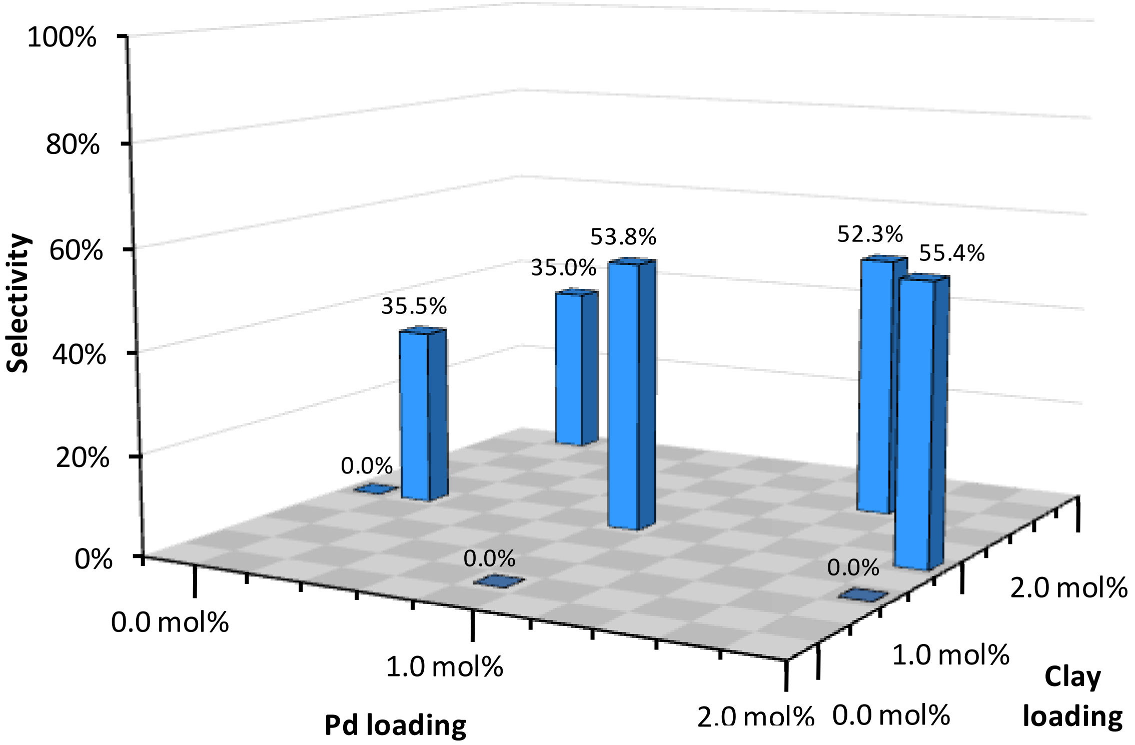


Figure 5.2 An optimisation study concerning the loading of Pd/C and K-10 for the conversion of limonene to *p*-cymene at 373 K.

Although increasing the catalyst loadings beyond 1 mol% had minimal benefit, raising the reaction temperature from 373 K to 413 K resulted in an improvement to 71% selectivity to *p*-cymene (Figure 5.3). The trend of selectivity dependence on temperature is not ideal, with a modest change in selectivity if a reaction temperature of 393 K rather than 373 K was applied. Then a jump in selectivity occurs when reactions are conducted 20 K hotter still. Higher temperatures were less promising, with a drop in selectivity observed. The reason, as inferred from ^{1}H-NMR spectroscopy, is that reactions at 413 K show a single aromatic product, *p*-cymene, but other aromatic signals are present at higher reaction temperatures (Figure 5.4). This implies some cracking of the terpenoid structure is occurring to our disadvantage. At lower temperatures complete conversion of the limonene was still achieved, and a single aromatic signal accompanied by the other three distinctive *p*-cymene signals observed in the ^{1}H-NMR spectrum of the reaction mixture. Complete consumption of limonene and no olefin by-products were always observed when Pd/C was the catalyst, confirmed with NMR spectroscopy and GC-MS, which means the selectivities reported in Figure 5.2 and Figure 5.3 should translate to yields.

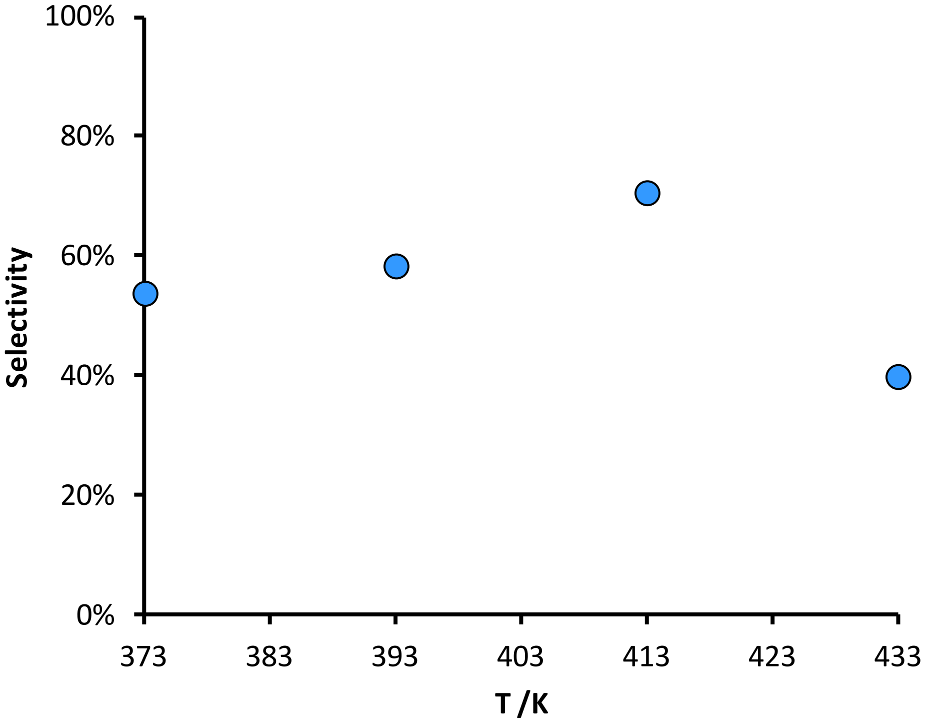


Figure 5.3 Product selectivity to *p*-cymene from limonene at different temperatures.

The limit to *p*-cymene selectivity that is being encountered can be attributed to the co-synthesis of the completely hydrogenated C_{10} hydrocarbon *p*-menthane (1-methyl-4-isopropylcyclohexane). The maximal proportion of *p*-cymene obtained at 413 K approximates the

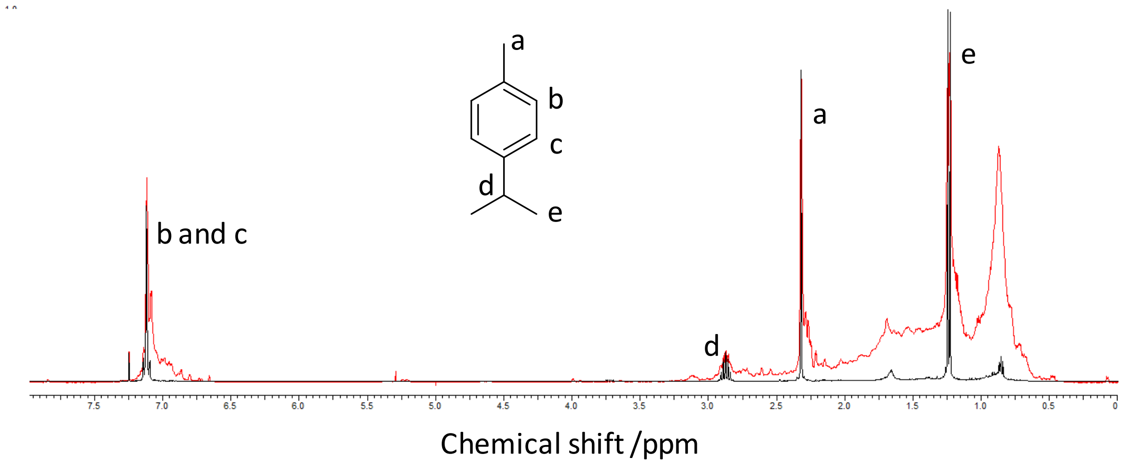
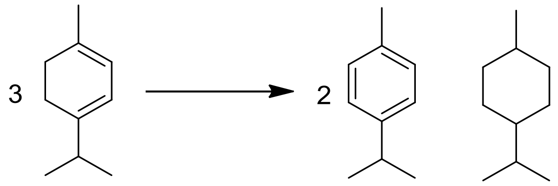


Figure 5.4 A ^{1}H-NMR spectroscopic analysis of the reaction mixture after limonene oxidation, largely to give *p*-cymene at 413 K (black), but less effective at 433 K (red).

2:1 molar ratio of the two disproportionation products (Scheme 5.5) [Lesage 1996]. In fact a product selectivity just above 67% has been achieved, hopefully suggesting disproportionation does not impose an impassable ceiling for the highest achievable yield. The yield of *p*-cymene from α-terpinene was even higher at 82% (Figure 5.1). This implies that the role of the acidic clay, by promoting multiple alkene isomerisations, is also to blame for the less than quantitative yield. Because no other dehydrogenation catalyst from the metal salt screening could match this obviously flawed disproportionating mode of catalysis, palladium still proves to be the most useful metal catalyst for this reaction. It is possible that K-10 montmorillonite could be supplanted for a more efficient acid to help increase the yield of *p*-cymene. Alternatively it is known that *p*-menthane will oxidise to *p*-cymene in the presence of palladium above 530 K, but this was not attempted here [Littmann 1942].



Scheme 5.5 The disproportionation of α-terpinene to *p*-cymene and *p*-menthane.

The reusability of the co-catalysts as their mechanical mixture was poor. Virtually no *p*-cymene was produced upon the second use of the catalyst under otherwise identical conditions, although no pre-treatment was applied beyond the drying of the catalyst. Attempts to use hydrogen acceptors have been applied in the past, to this and similar reactions, but 1-decene and levulinic acid were found to actually inhibit the reaction with no *p*-cymene identified in the reaction mixture [Lesage 1996, Wise 2007]. Finally, diluting the reaction using *p*-cymene as a solvent could not improve the selectivity of the reaction either. So it seems an improvement in *p*-cymene yield might require a completely different approach. In order to progress with the synthesis of *p*-CSA, what *p*-cymene could be made by this method was taken forward.

The synthesis of bio-based *p*-cymenesulphonic acid: This procedure for making *p*-cymene was then transposed onto citrus oil feedstocks. Steam distillation of orange peel afforded the colourless essential oil in yields equating to a little over 1 g of essential oil for every three fruits. Each fruit, Uruguyan Navel Late, weighed approximately 80 g. The orange oil gave similar results to neat limonene in the synthesis of *p*-cymene (Figure 5.5). The final reaction mixture containing *p*-cymene could be purified with a second steam distillation stage, although *p*-cymene and *p*-menthane are co-distillates. Attempts to separate these compounds by distillation have been fraught with difficulty [Berg 1992]. Resorting to chromatography would be an unacceptable move considering the product is intended for use as a commodity solvent as well as a chemical intermediate. Also, using additional solvent to purify *p*-cymene would defeat the point of synthesising the *p*-cymene in the first place. Thankfully in the following sulphonation procedure *p*-menthane is unreactive, and upon formation of the highly polar sulphonic acid, the hydrocarbon component of the resultant mixture can be decanted.

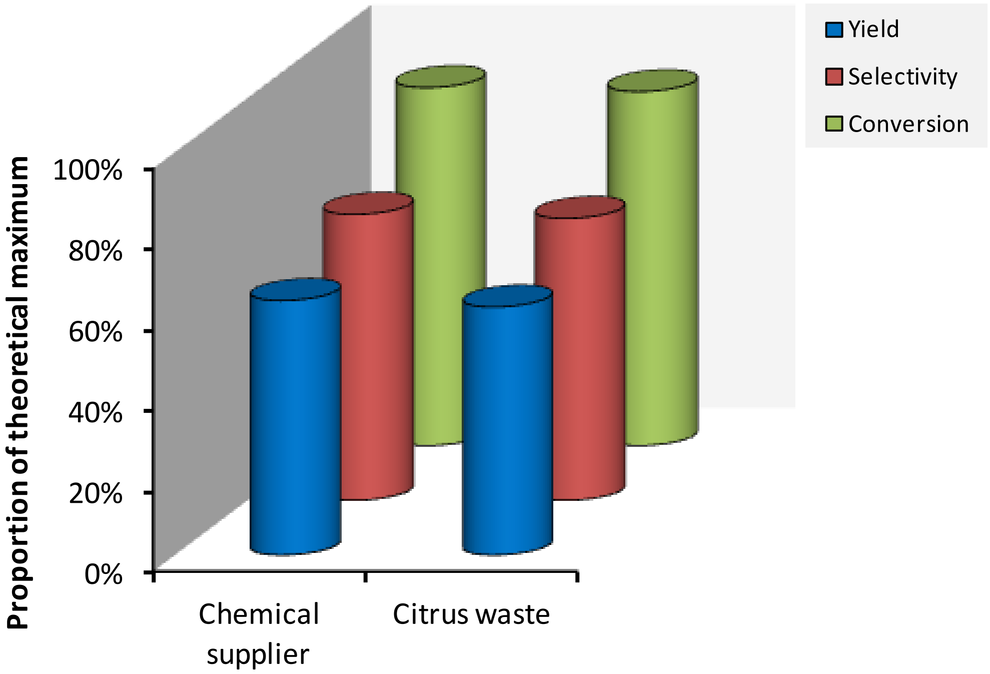


Figure 5.5 The efficiency of *p*-cymene production utilising different sources of limonene.

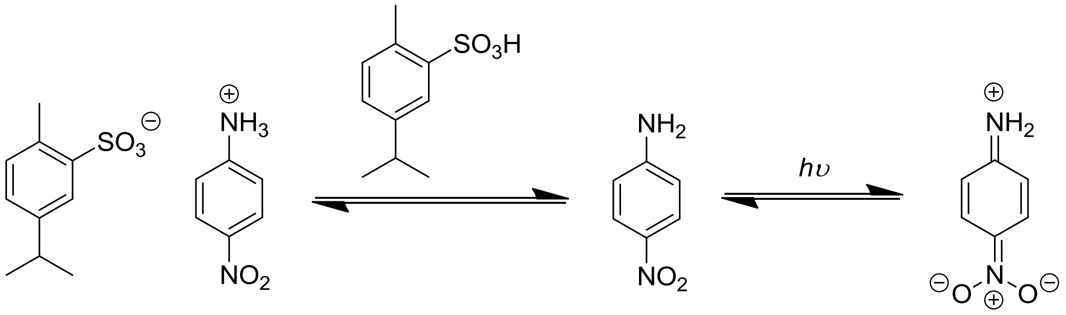
Unfortunately sulphonation of the mixture of *p*-cymene and *p*-menthane was not nearly as high yielding as observed in the previous optimisation study using pure *p*-cymene. A *p*-CSA yield of 28% of the theoretical maximum based on the *p*-cymene content of the substrate was isolated, corresponding to a 16% yield based on the limonene content of the citrus waste. Nevertheless the bio-based product was identical to *p*-CSA made from pure *p*-cymene. Either could be used to catalyse a variety of reactions equally well.

The properties of *p*-cymene sulphonic acid: Prior to its application as an acid catalyst it was necessary to determine the acidity of *p*-CSA. The Hammett acidity function (H_0) is an equivalent measurement to pH for strong acids [Hammett 1932]. As a method of calculating the ability of strong acids to protonate a weak base by UV–vis. spectroscopy, it is now routinely used as part of the characterisation of novel sulphonic acids [Tao 2011]. The calculation is equivalent to that describing pH (Henderson-Hasselbalch equation) but water is generalised to a base:

Equation 5.1

$$\mathrm{H_0 = pK_{BH^+} + \log_{10} \frac{[B]}{[BH^+]}}$$

Familiar from Kamlet-Taft solvatochromism studies, 4-nitroaniline was selected as the base partner for determining H_0 (Scheme 5.6).



Scheme 5.6 The H_0 equilibrium between *p*-CSA and 4-nitroaniline.

Increasing the concentration of the sulphonic acid will adjust the equilibrium, generating a greater quantity of the conjugate acid. This protonated species does not absorb light as the free base does, and so the absorbance recorded by a UV spectrophotometer is proportional to the equilibrium position (Figure 5.6). As expected from electronic arguments, the acid strength of alkyl functionalised arenesulphonic acids is inversely proportional to the degree of substitution present, and so *p*-CSA (H_0 = 1.26) is a slightly weaker acid than *p*-TSA (H_0 = 1.22) at a concentration of 80 mmol·dm^{-3} (Figure 5.7). To put this in context the difference is less than that between *p*-TSA and benzenesulphonic acid (BSA, H_0 = 1.12) implying that little influence would be felt in terms of performance upon substituting one sulphonic acid for the other.

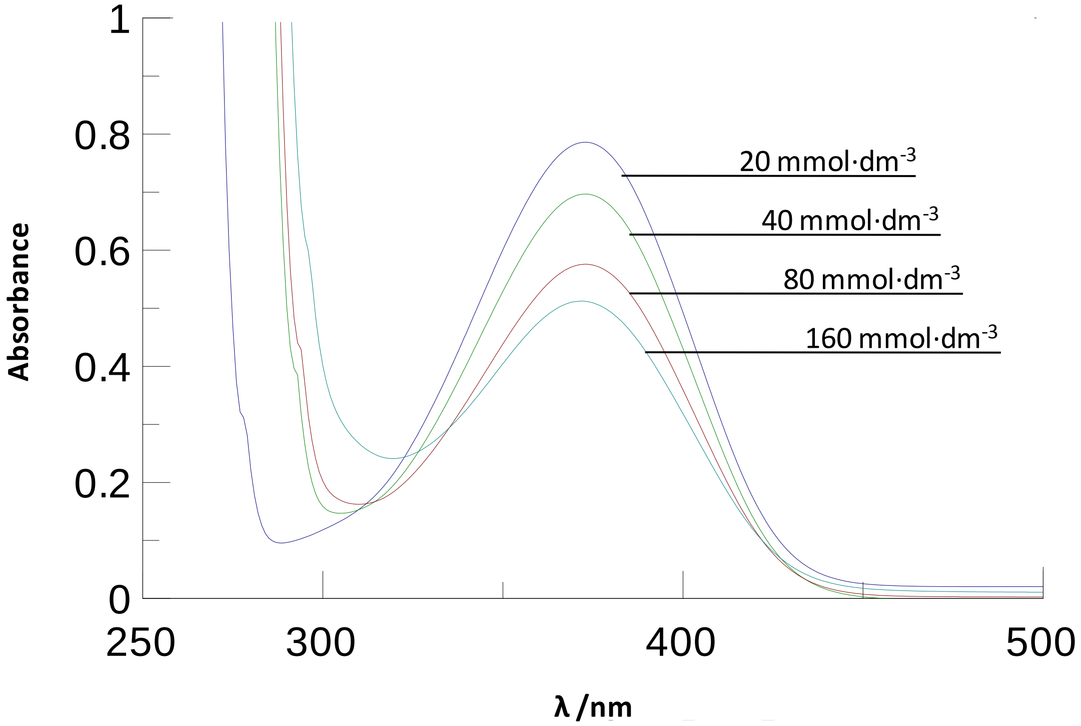


Figure 5.6 The absorbance profile of 4-nitroaniline in the presence of varying concentrations of *p*-CSA.

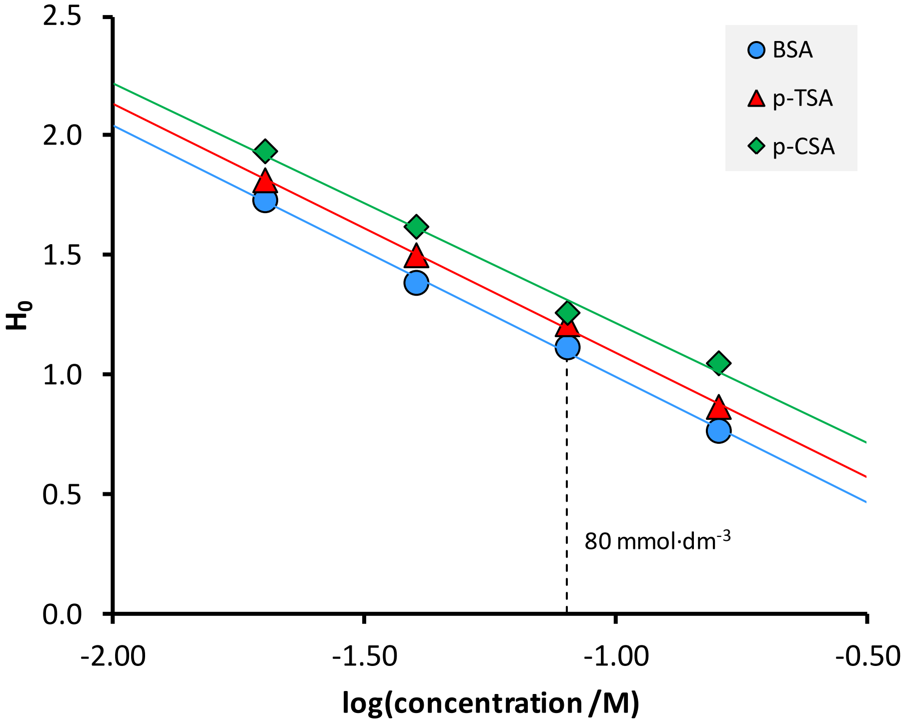


Figure 5.7 The Hammett acidity of sulphonic acids.

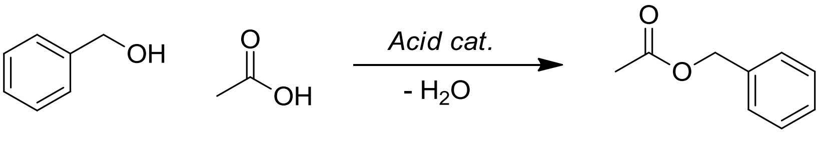
A key benefit of using *p*-TSA as a catalyst in synthesis is that it can be easily washed out of an organic phase with water. Biphasic systems of ethyl acetate (20 mL) and water (20 mL) containing 0.05 mmol of either *p*-TSA or *p*-CSA were prepared to represent a post-reaction work-up. Neither organic layer contained any sulphonic acid after mixing and separation of the phases upon standing, as determined by ^{1}H-NMR spectroscopy. It is noticeable that the system containing *p*-CSA is slower to separate into two phases after agitation, but addition of sufficient water allowed the complete removal of the acid from the organic phase. The thermal stability of both sulphonic acids is also comparable, as evidenced by their comparable decomposition temperatures (T_d). These observations are summarised in the following table, along with their melting points (T_m) (Table 5.1).

Table 5.1 A selection of properties defining *p*-TSA and *p*-CSA.

Property	*p*-TSA	*p*-CSA
T_m /K	379-380 [Shultz 1979 page 60]	323-324
T_d /K	479	476
H_0 at 80 mmol.dm^{-3}	1.22	1.26

5.2 Combined solvent and catalytic effects in carbonyl additions

Kinetics of Fischer esterification assisted by sulphonic acid catalysts: The first application selected as a test of the catalytic prowess of *p*-CSA was the Fischer esterification of acetic acid and benzyl alcohol (Scheme 5.7). The reaction was attempted in a variety of solvents in combination with both *p*-TSA and *p*-CSA. A similar case study evaluating the synthesis of benzyl methoxyacetate can be found in the literature, and was referred to earlier as a counterpoint to the study of 1-butanol reacting with butanoic anhydride [Wells 2008]. The purpose of this study, beyond confirming the solvent effect, is to further develop an understanding of solvent-catalyst synergy, and of course to determine if *p*-CSA is a suitable replacement for *p*-TSA or not.



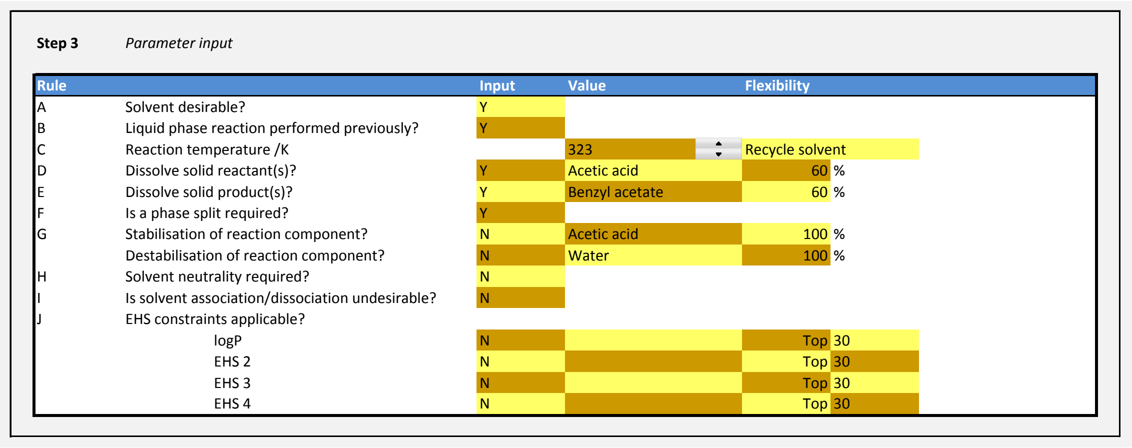
Scheme 5.7 The synthesis of benzyl acetate by Fischer esterification.

The solvent selection algorithm can be applied prior to any experimental work, because thanks to various studies in this work and elsewhere it is well established that carbonyl addition is inversely proportional to **β**. This being the case the LSER generated by Wells *et al*. was used [Wells 2008]:

Equation 5.2

$$\ln(\mathbf{k}) = -3.59 - 7.25\boldsymbol{\beta}$$

With this single parameter relationship, candidates are ranked within the solvent performance assessment (model B2) in accordance to their inverse proportionality with **β**. Few changes to the inputs used in the previous case study concerning uncatalysed esterification were required. No extra considerations were made given the introduction of an acid catalyst, whether that is *p*-TSA or *p*-CSA. Alkene functionalised solvents like limonene are certain to be incompatible with a sulphonic acid catalyst and so were not included in the solvent set. Amines were also excluded in order to limit interference with the catalyst and prevent side reactions. Homogeneity of acetic acid and benzyl acetate was demanded, with the addition of a phase split to eliminate water from the reaction medium, thereby discouraging an equilibrium that would limit the eventual yield (Figure 5.8). The GSK solvent selection guide was again used as the basis of the greenness assessment (model B3), with a score of 4 set as the minimum permissible value in each category, including the LCA category [Henderson 2011].



Step 3 *Parameter input*

Rule		Input	Value	Flexibility
A	Solvent desirable?	Y		
B	Liquid phase reaction performed previously?	Y		
C	Reaction temperature /K		323	Recycle solvent
D	Dissolve solid reactant(s)?	Y	Acetic acid	60 %
E	Dissolve solid product(s)?	Y	Benzyl acetate	60 %
F	Is a phase split required?	Y		
G	Stabilisation of reaction component?	N	Acetic acid	100 %
	Destabilisation of reaction component?	N	Water	100 %
H	Solvent neutrality required?	N		
I	Is solvent association/dissociation undesirable?	N		
J	EHS constraints applicable?			
	logP	N		Top 30
	EHS 2	N		Top 30
	EHS 3	N		Top 30
	EHS 4	N		Top 30

Figure 5.8 Acid catalysed Fischer esterification to give benzyl acetate solvent selection algorithm screenshot, step 3: Parameter input.

Being so closely related to the previous two case studies, it is no surprise that *p*-cymene is offered as an option for a renewable solvent by the solvent selection algorithm. Of course the greenness assessment (model B3) can not be applied to *p*-cymene because of an absence of the necessary data. Otherwise it is competitive. As with the previous esterification case study, 1,2-

dichlorobeznene, chlorobenzene, and cumene are the only solvents to pass all the assessments in the solvent selection algorithm. Rather than overload the experimental solvent set with obscure petroleum-derived aromatics, toluene was chosen as the only hydrocarbon solvent to compliment *p*-cymene. Acetonitrile, butanone, chloroform, 2-MeTHF, and THF were also selected as solvents for experimentation. Rate constants were determined with a ^{1}H-NMR spectroscopic analysis (Equation 1.14 and Equation 1.15), following the benzyl group of benzyl alcohol moiety much as benzylamine in its amidation with 4-phenylbutanoic was before (Figure 5.9).

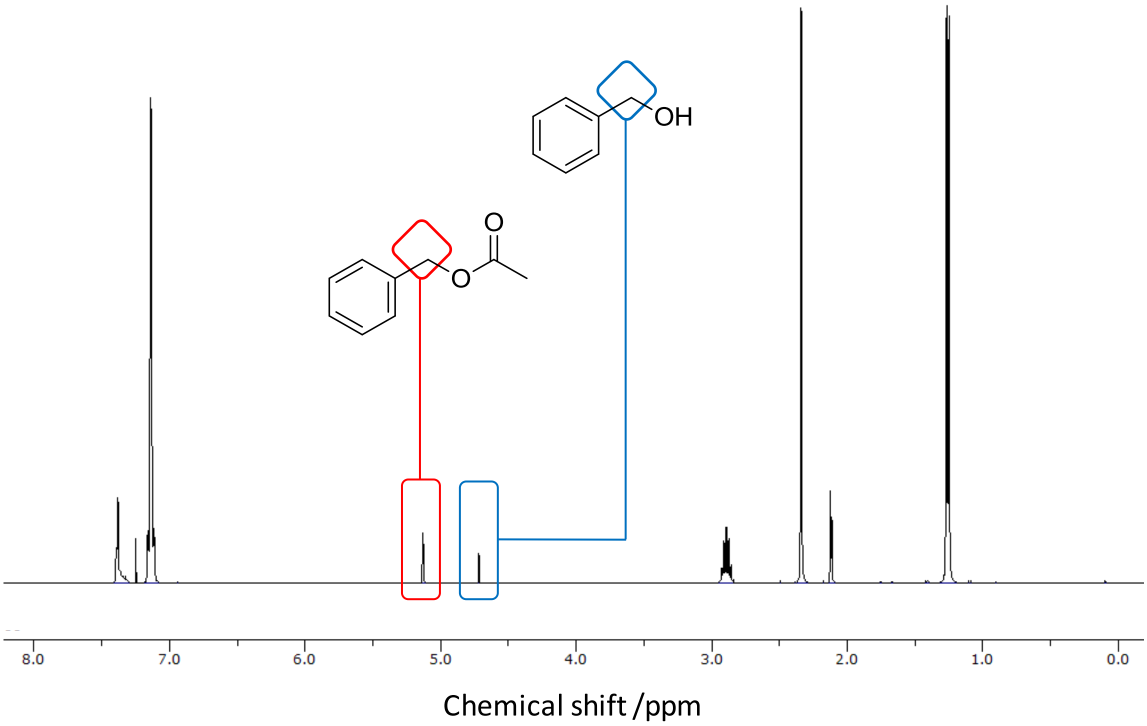


Figure 5.9 A ^{1}H-NMR spectrum of the Fischer esterification to give benzyl acetate in *p*-cymene.

As expected, toluene and *p*-cymene, the precursors to the sulphonic acid catalysts, provide an environment that accelerates the rate of reaction in combination with either catalyst (Table 5.2). None of these solvents are considered green, but of course *p*-cymene is an unknown in this regard (Figure 5.10). Joining *p*-cymene in this case study is 2-MeTHF, another bio-based solvent [Aycock 2007]. Unfortunately this solvent has a medium hydrogen bond accepting ability and will not provide a great environment for esterifications. It also fares poorly in the SUS-HAS-ECO classifications, being outscored by chloroform in two categories (Figure 5.10). This is because of consistent mediocrity rather than one severe shortcoming, although 2-MeTHF is regarded as a dangerous peroxide former [Fábos 2009]. With no data for *p*-cymene, none of the solvent set provide an appealing reaction medium.

Table 5.2 Solvent properties concerning a Fischer esterification to give benzyl acetate.

Solvent	ln(**k**) *p*-TSA	ln(**k**) *p*-CSA	α	β	π*
Acetonitrile	20.5	18.0	0.35	0.37	0.80
Butanone	11.0	10.4	0.00	0.51	0.68
Chloroform	85.1	102	0.20	0.10	0.58
p-Cymene	92.9	104	0.00	0.13	0.39
2-MeTHF	5.60	4.02	0.00	0.57	0.51
THF	3.35	2.37	0.00	0.55	0.58
Toluene	84.5	111	0.00	0.12	0.50

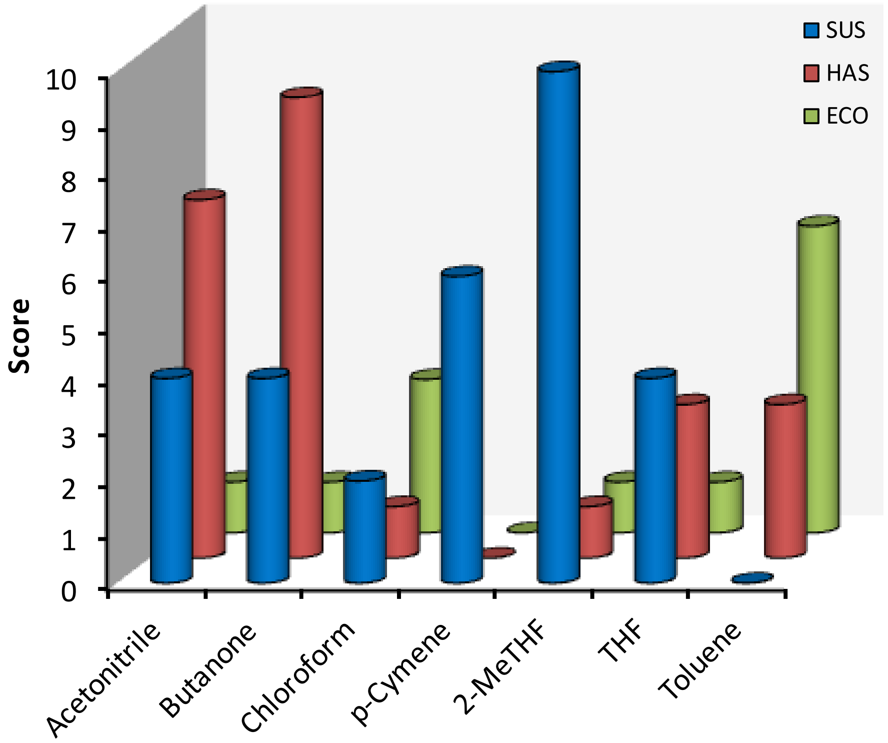


Figure 5.10 The solvent SUS-HAS-ECO classifications for the benzyl acetate case study.

With the experimental data now obtained the LSER can be constructed. As presumed only **β** is influential in determining the rate of esterification (Figure 5.11). The weak hydrogen bond donating ability of acetonitrile and chloroform has no effect on the reaction, which is not surprising given that this effect will be relatively weak by comparison in the presence of a strong acid catalyst. Dipolarity (**π***) was also statistically insignificant. Reactions were performed using benzenesulphonic acid, *p*-TSA, and *p*-CSA as catalysts in *p*-cymene. The resultant reaction rates

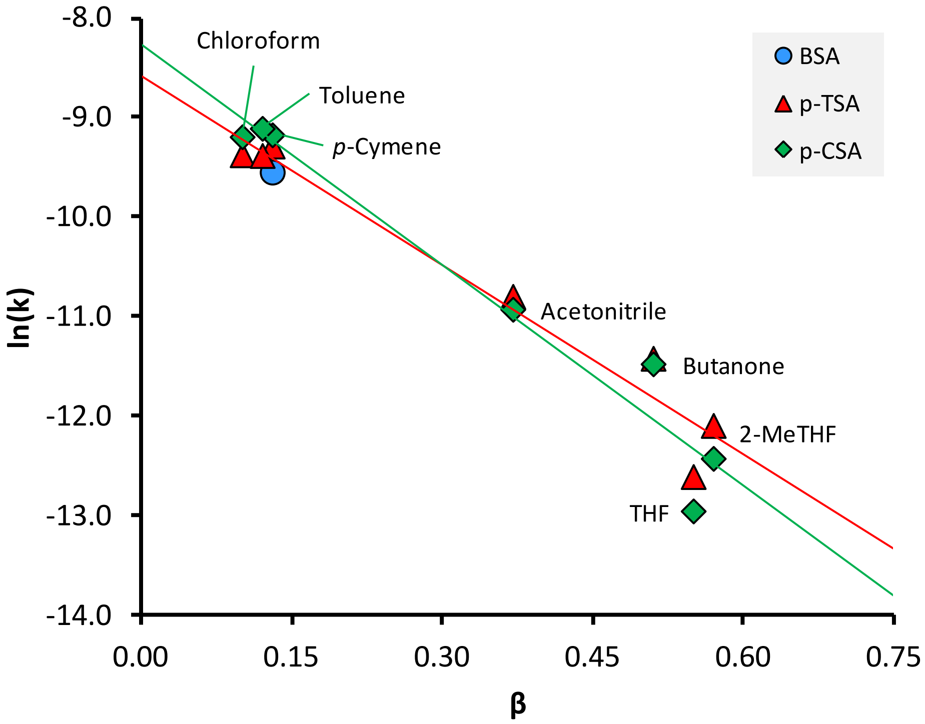


Figure 5.11 The LSER describing Fischer esterification catalysed by *p*-TSA and *p*-CSA to give benzyl acetate, also including the combination of benzenesulphonic acid in *p*-cymene.

were barely distinguishable from each other (Figure 5.11). Equations describing these LSERs can be found in the appendix (Table 8.14)

There is an indication that *p*-TSA is superior to *p*-CSA in high polarity solvents, while *p*-CSA marginally enhances reaction rates in low polarity solvents. Given the slightly superior acidity of *p*-TSA, the former observation is expected. The latter observation is most probably due to differences in the solubility of the catalysts, or more precisely the stability of their conjugate bases in solution during protonation of the reactive substrate. The upshot of this is that *p*-CSA will outperform *p*-TSA in solvents whose polarity is suited to esterification. The relative difference in ln(**k**) between *p*-TSA and *p*-CSA catalysed esterifications can also be interpreted as a function of **β** (Figure 5.12). The correlation is weaker than what would be liked (R^2 = 0.856) but is still significant, suggesting this observation is a true consequence of the structures of the sulphonic acids. Removing butanone from the solvent set increases the coefficient of determination to a more satisfactory level (R^2 = 0.963).

This study, beyond the confirmation that *p*-cymene is an effective bio-based solvent for this class of transformation, aids the hypothesis that the same solvent effect seems to dictate the rate of all carbonyl addition chemistries. To reiterate, the role of the solvent appears to be one of

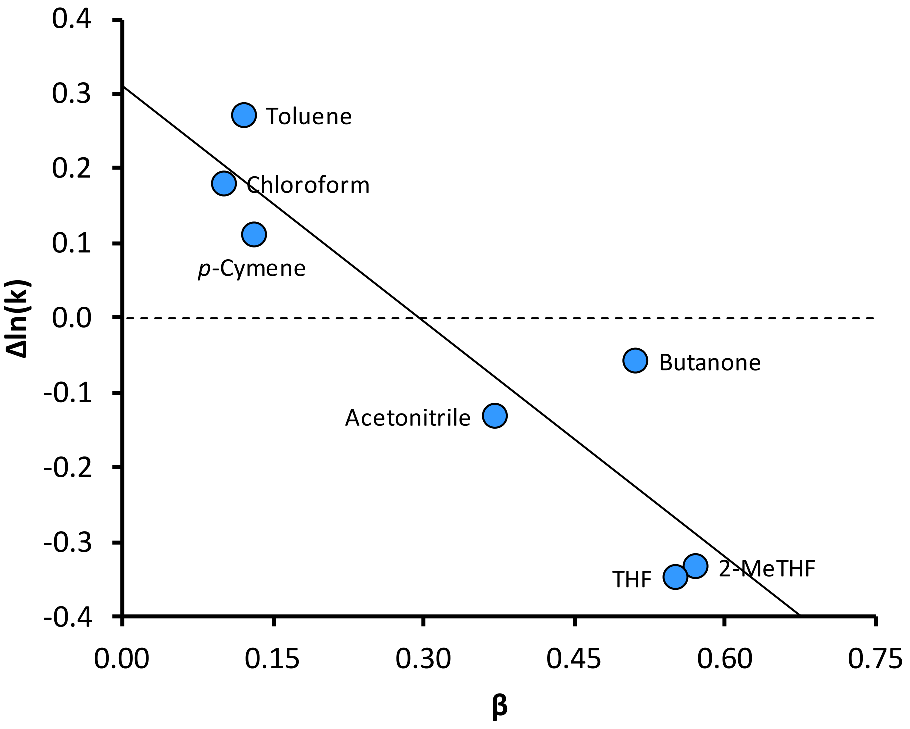
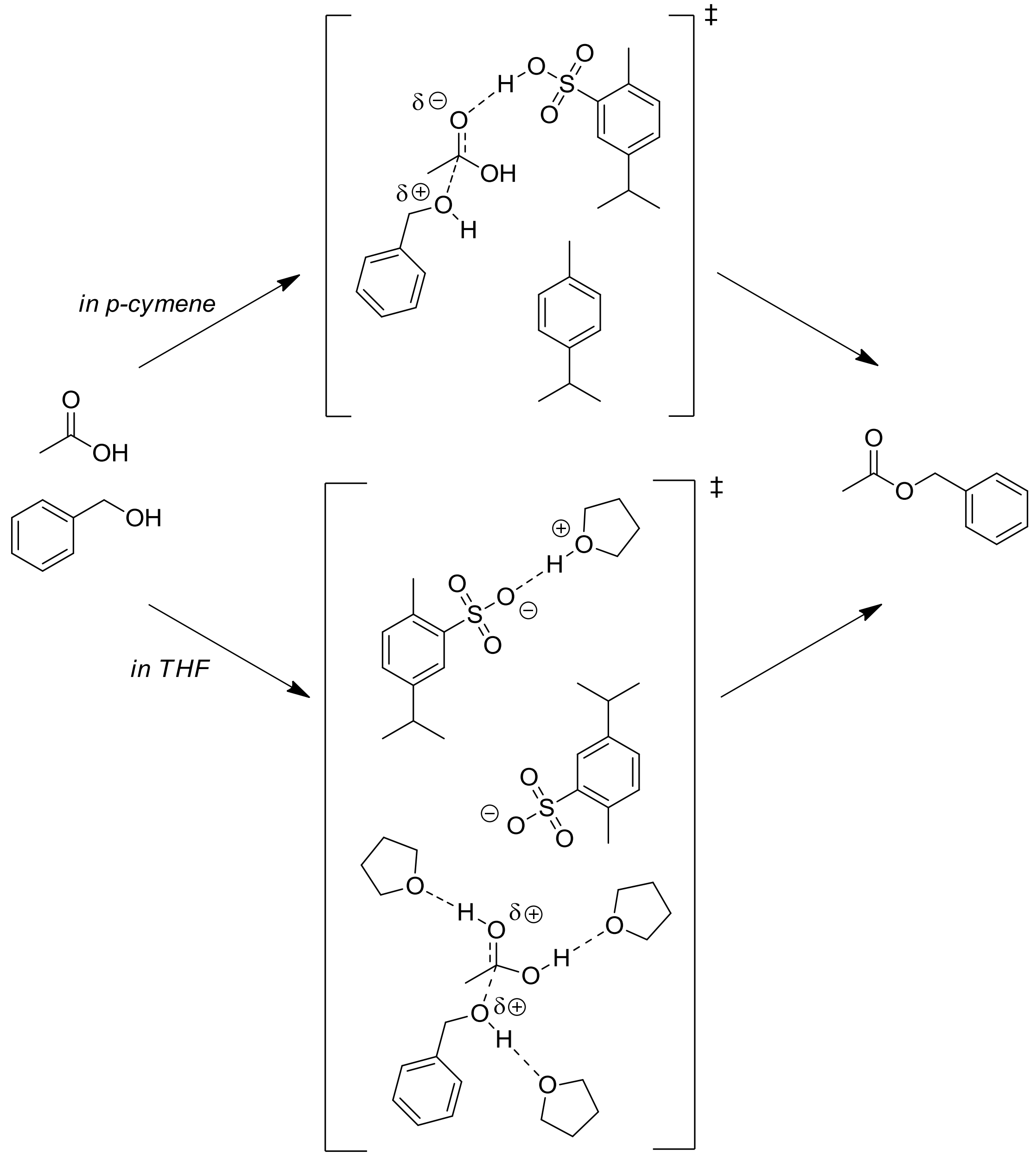


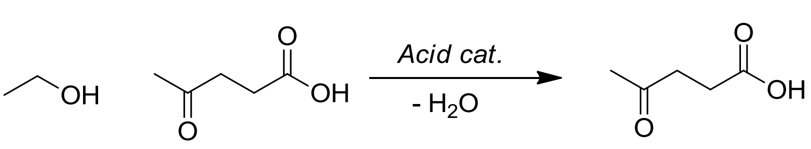
Figure 5.12 The difference between the ln(**k**) of *p*-TSA and *p*-CSA catalysed Fischer esterifications.

enthalpy stabilisation at the transition state but at a large entropic cost. Once more, there is no need for the explanation previously presented by Wells for the action of the solvent consisting solely of a solvent-catalyst hydrogen bond (Scheme 4.1) [Wells 2008]. In many cases a catalyst such as *p*-CSA would be expected to protonate the solvent (Scheme 5.8). As long as the proton is transferred to the carbonyl moiety of the reactive substrate then this apparently makes no difference. Because the correlations are derived from experimental data and do not rely on a correction factor to account for the exact nature of the proton donor, there is no need to make any special considerations for catalyst-solvent interactions or even salt formation between the two.

Preparative Fischer esterifications catalysed by sulphonic acids or metal salts: It is apparent from previous results that the *p*-cymene/*p*-CSA system should be very well suited to promoting a variety of esterifications and related transformations. To verify this, another Fischer esterification was undertaken, but with the objective of isolating the product rather than monitoring the rate of the reaction. Accordingly ethyl levulinate, itself identified as a possible bio-based solvent, was synthesised in *p*-cymene, toluene and 2-MeTHF using a variety of catalysts (Scheme 5.9). Combinations of *p*-cymene, toluene and 2-MeTHF with *p*-TSA (1 mol%), *p*-CSA (1 mol%), In(OTf)$_3$ (5 mol%), InCl$_3$ (5 mol%), or FeCl$_3$ (5 mol%) as catalysts were evaluated. Both the Brønsted acid catalysts were effective, with *p*-CSA indistinguishable from *p*-TSA (Figure 5.13). In fact the performances of all the catalysts were fairly similar, although the Lewis acids had to be


in p-cymene
in THF
‡
‡
δ⊖
δ⊕
OH
OH
OH

Scheme 5.8 The combined roles of solvent and catalyst in the $A_{AC}2$ mechanism of a representative carbonyl addition.


Acid cat.
- H$_2$O
OH
OH
OH

Scheme 5.9 Fischer esterification to give ethyl levulinate.

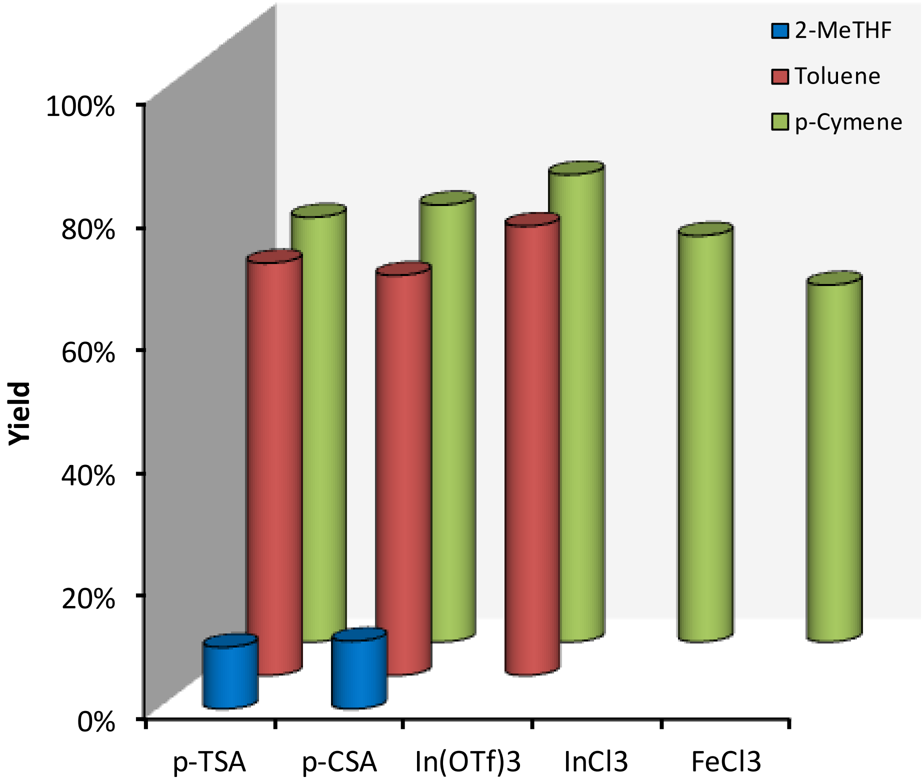
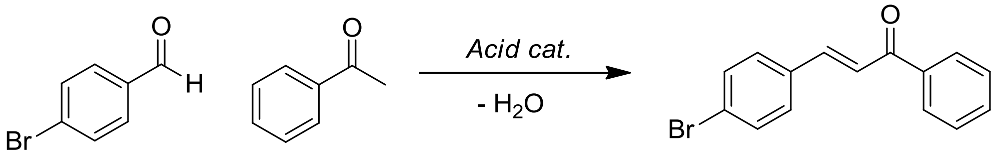


Figure 5.13 Combinations of different catalysts and solvents for the synthesis of ethyl levulinate.

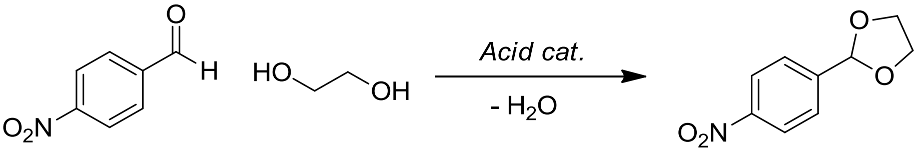
used in higher loadings to compete with the sulphonic acids. The comparable yields suggest that the equilibrium position of the reaction was being approached, which is not an unreasonable assumption given that equimolar quantities of the reactants were used. There is nothing to suggest from these results that Lewis acidic compounds make especially effective catalysts in condensation chemistry but can be made competitive with Brønsted acids at higher loadings.

The most significant practical drawback to the use of *p*-cymene instead of toluene (despite of the slightly superior yields) becomes apparent in this case study. The high boiling point of *p*-cymene (450 K) means that removal of the solvent is not routine as it is for toluene. Toluene and 2-MeTHF could be removed *in vacuo* to give the product. In contrast, *p*-cymene was removed by running the reaction mixture through a column of silica, resulting in extra solvent waste and silica waste far outweighing the benefits of substituting toluene with *p*-cymene in the first place. Removal of *p*-cymene by steam distillation was deemed inappropriate given the nature of the product, although this approach may be suitable in a number of other reactions. In real terms this restricts the use of *p*-cymene as a solvent to chemistry in which the product can be precipitated or otherwise extracted from solution. Although this is the ideal way of claiming the product from a reaction mixture it is not always feasible. On the other hand, the higher boiling point of *p*-cymene when compared to toluene imparts the ability to perform reactions at higher temperatures than those attainable using toluene, xylene(s), or even mesitylene. Although this is not necessarily enough of a benefit to warrant the complete phasing out of toluene on this basis alone it will be advantageous on occasion.

Further *p*-CSA catalysed condensations: In order to establish *p*-CSA as a catalyst for condensation chemistry more examples of successful transformations are required to complement the esterification studies. Accordingly the synthesis of 4-bromochalcone *via* an aldol condensation (Scheme 5.10), and the acetal protection of 4-nitrobenzaldehyde with ethylene glycol were performed as further demonstrations of acid catalysis by *p*-CSA (Scheme 5.11). Both these types of procedure have been used previously as demonstrations of novel sulphonic acid catalysis [Liang 2008, Xu 2008].



Scheme 5.10 A chalcone synthesis by aldol condensation.



Scheme 5.11 Acetal protection using ethylene glycol.

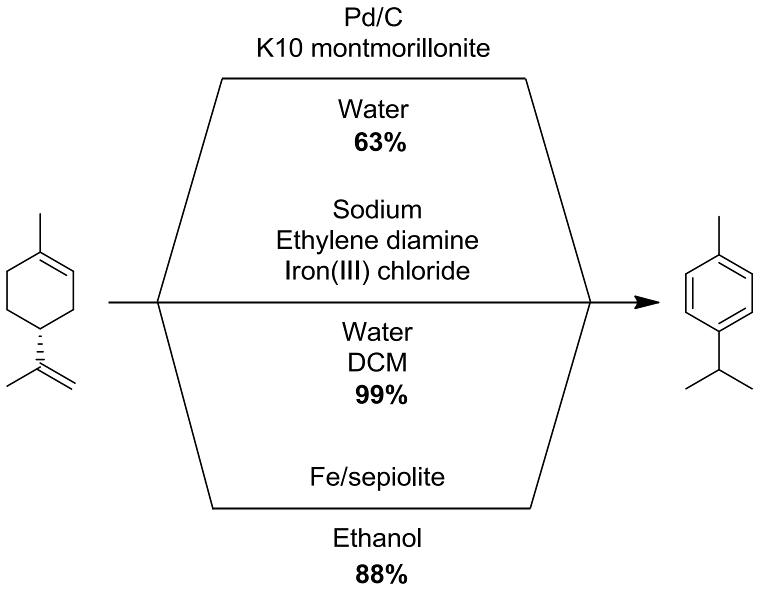
The aldol condensation was performed without an auxiliary solvent, and the product was recrystallised from ethanol. After refluxing in cyclohexane, the acetalisation reaction product precipitated from the reaction mixture upon cooling. The comparison shows that *p*-TSA fairs no better than *p*-CSA in its role as an acid catalyst despite its ever so slightly greater acidity (Table 5.3). After performing each reaction in triplicate no difference in the mean yield between *p*-TSA and *p*-CSA catalysed reactions is observed. In the synthesis of ethyl levulinate only a 2% variation in yields was observed. Generally it appears that no major allowances have to be made or additional considerations accounted for if replacing *p*-TSA with *p*-CSA. Plus the benefit of a renewable reaction auxiliary is an important attribute in modern synthetic chemistry.

Life cycle aspects of *p*-cymene and *p*-CSA: It is important to address the environmental impact of sulphonic acids pre-application and post-application and not just during the reaction. To begin, the synthesis of *p*-cymene from limonene must be addressed. Several procedures are in the

Table 5.3 Comparison of yields in esterification, aldol condensation, and acetalisation chemistries obtained by *p*-TSA and *p*-CSA catalysis.

Product	Solvent	Yield (*p*-TSA)	Yield (*p*-CSA)
Ethyl levulinate	*p*-Cymene	71%	69%
Ethyl levulinate	Toluene	65%	67%
Ethyl levulinate	2-MeTHF	11%	10%
4-Bromochalcone	n/a	73%	73%
2-(4-Nitrophenyl)-1,3-dioxolane	Cyclohexane	92%	92%

literature describing highly efficient protocols (Scheme 5.12). These arguments were used in the previous chapters to validate the use of *p*-cymene as a viable bio-based solvent. Now that the conversion of limonene to *p*-cymene has actually been attempted, and demand for *p*-cymene as a chemical intermediate has been established, it becomes wise to assess the benefits of each route to *p*-cymene. The iron doped sepiolite clay catalyst of Martin-Luengo can give complete conversion of limonene into *p*-cymene [Martin-Luengo 2010]. Secondly, the patented procedure of Berti requires sodium metal, iron trichloride, and ethylene diamine to oxidise limonene [Berti 2010]. Both these procedures are more productive than the route proposed in this work and so deserve to be considered as alternatives.



Scheme 5.12 Three different sets of conditions that effect the transformation of limonene to *p*-cymene with varying success.

The effectiveness of each route can be illustrated with the appropriate metrics where A refers to this work; B, iron-sodium oxidation [Berti 2010]; C, small scale iron sepiolite microwave experiment [Martin-Luengo 2010]; D, larger scale iron sepiolite microwave experiment; E, large scale iron sepiolite thermally heated experiment (Figure 5.14). Although the yield of the palladium catalysed reaction from this work is lower yielding the other limonene oxidations, it benefits from the low quantity of catalyst required. Perhaps if more effective heating could be imparted to the reaction mixture by microwave technology, the yields might be improved in the same way those of Martin-Luengo's process are. To its disadvantage, the iron doped clay catalyst of Martin-Luengo is applied in super-stoichiometric amounts (based on its iron content) in the most effective protocol. However the worst reaction mass efficiency actually belongs to the sodium and iron trichloride assisted *p*-cymene synthesis, which is not helped by the 20 mol% of sodium metal and the excess ethylene diamine required.

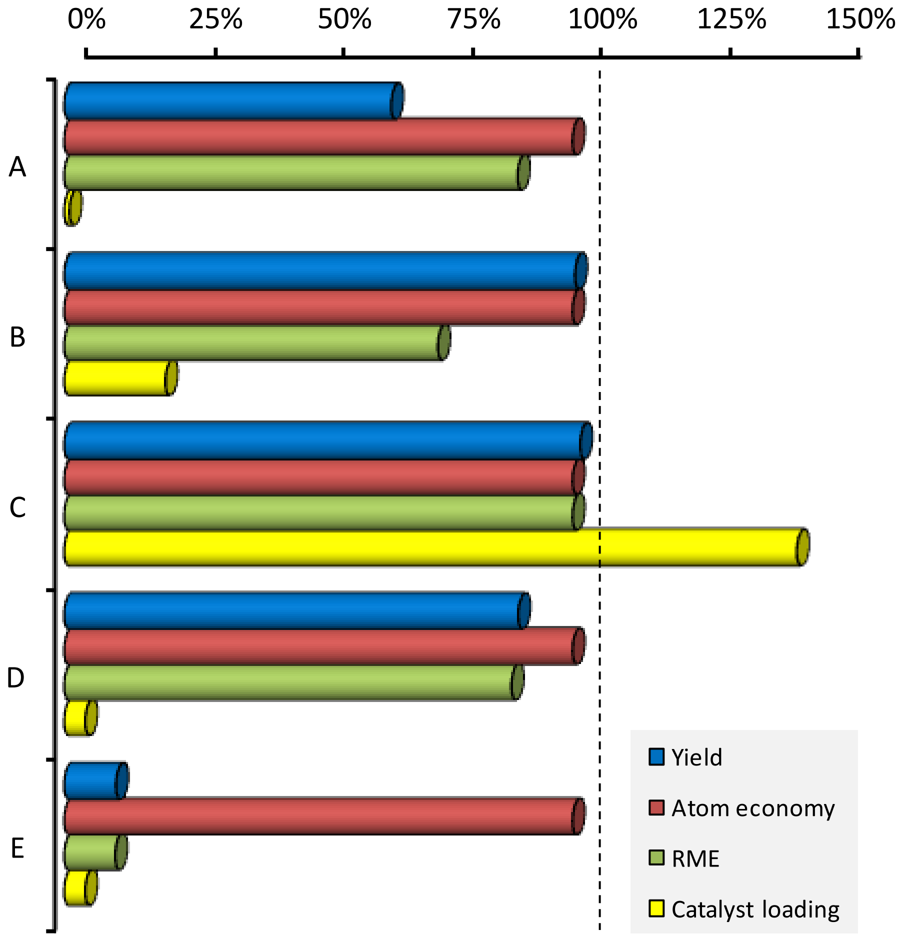


Figure 5.14 Metrics associated with the conversion of limonene to *p*-cymene.

Concerning the patented method of Berti and co-workers, the PMI also suffers from the large amount of solvent used for purification, half of which is DCM (Figure 5.15) [Berti 2010, Colonna 2011]. The same issue applies to the route proposed in this work, but in this case all of the auxiliary solvent is the water required for steam distillation. Unlike the procedure devised by Berti, no auxiliary organic compounds are required but the process of steam distillation introduced a small loss of product, further affecting the yield and associated metrics. The PMIs of Martin-Luengo's work varies according to the scale of the reaction [Martin-Luengo 2010]. At almost analytical scale the minute amount of product is recovered with ethanol after the reaction, the auxiliary solvent contributing massively to the PMI. Thermally heated reactions with improved substrate to catalyst ratios (4 mol% iron loading) fail to provide a reasonable conversion to *p*-cymene, and so the PMI is now dictated by unreacted limonene and by-products. However the yields and the quantity of waste produced from the procedure are acceptable when the microwave activated reaction is conducted with the lower loading of iron doped clay, and no solvent is required at all. It seems that overall, each procedure has faults and that makes it is hard to decide upon the best option. The literature protocols will maximise the yield [Colonna 2011, Martin-Luengo 2010], which is the greatest hurdle in the palladium catalysed route described here. But in actuality, if the water used in the steam distillation was recycled then the waste associated with the K-10 montmorillonite and Pd/C catalysed synthesis of *p*-cymene would

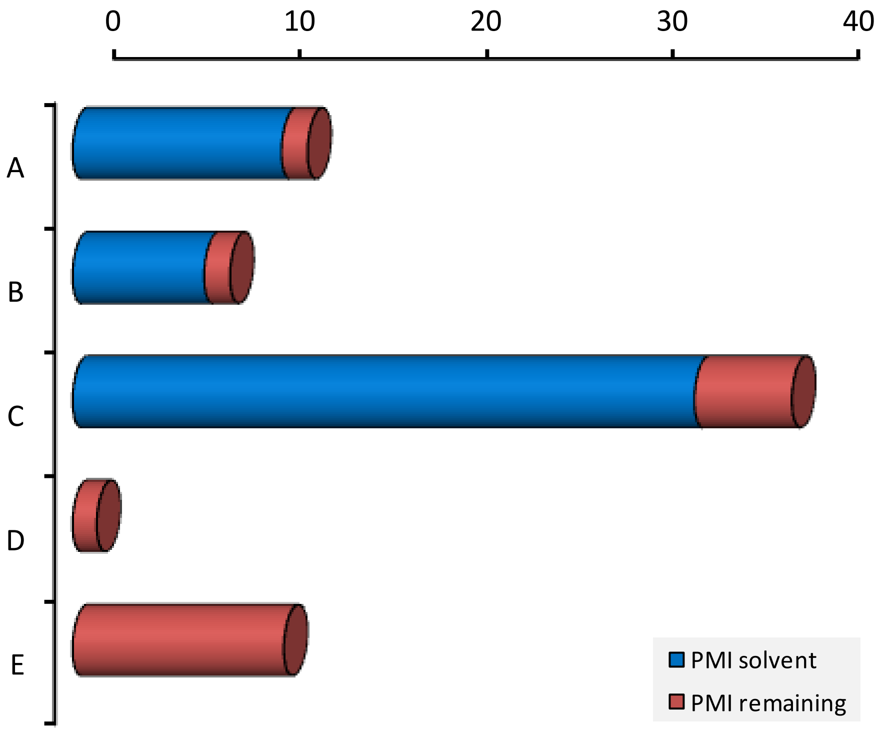


Figure 5.15 Process mass intensity of different *p*-cymene syntheses. Key: Same as that stated for Figure 5.14.

be very little indeed. Ultimately the full conversion of limonene to *p*-cymene would be desirable, probably essential, for larger scale production purposes.

Of course it is not just the synthesis of *p*-cymene that needs to be addressed. The subsequent sulphonation to give *p*-CSA presents some issues. The manufacturing process for *p*-TSA uses an excess of toluene that is constantly recycled by distillation as the reaction proceeds by reflux [Knaggs 1979 page 7]. The higher boiling point of *p*-cymene will certainly prevent the same approach being used (Figure 5.16). Instead, an excess of sulphuric acid was used in this work to sulphonate *p*-cymene. Unlike *p*-TSA production, this means that more sulphuric acid is used in the synthesis of *p*-CSA than can be accounted for by the resultant sulphonic acid. This has obvious economic implications, but more importantly (it might be argued) it is a waste of resources given the catalytic properties of sulphuric acid.

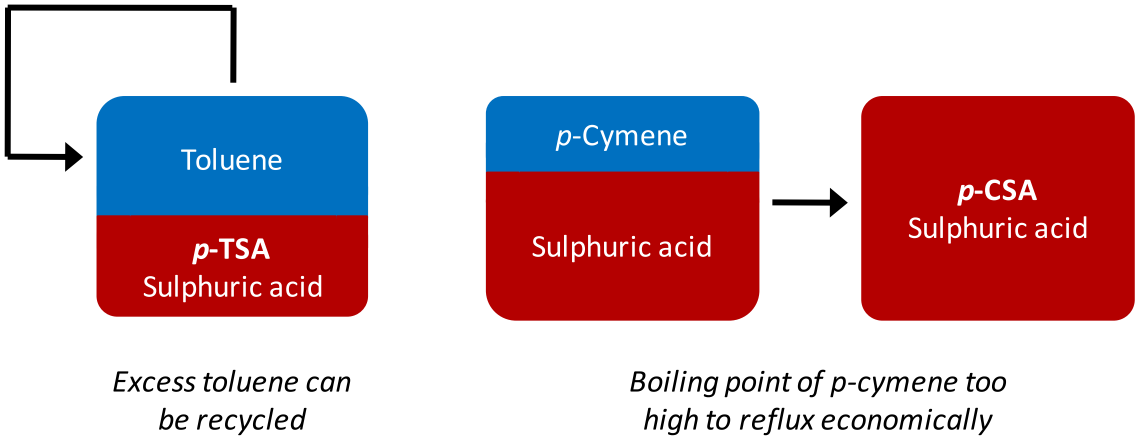

Toluene
***p*-TSA**
Sulphuric acid
p-Cymene
Sulphuric acid
***p*-CSA**
Sulphuric acid
Excess toluene can be recycled
Boiling point of p-cymene too high to reflux economically

Figure 5.16 The manufacturing processes for arene sulphonic acids.

Since sulphonic acids are products of a reaction between an arene and sulphuric acid, why is sulphuric acid not used directly as a catalyst? Unless sulphonic acids resolve issues in synthesis caused by the use of sulphuric acid then converting sulphuric acid into a sulphonic acid appears to have little use. Sulphonic acids have the advantage of not corroding steel, steel being the material that many industrial reaction vessels are made of [Avdeev 2007]. In fact steel corrosion protection agents sometimes feature the sulphonic acid moiety [Srivastava 2010, Zhao 1999]. The milder acidity of sulphonic acids may improve reaction selectivity in cases where sulphuric acid causes the unwanted dehydration of the substrate. However it appears that solubility is the main reason that sulphonic acids are used instead of mineral acids. This may have more to do with perception rather than actual evidence, but nevertheless in small scale preparatory chemistry *p*-TSA is used more frequently than sulphuric acid [Scifinder 2013b]. This bodes well for the introduction of *p*-CSA as a bio-based analogue of *p*-TSA.

We are now left to consider post-application concerns regarding the use of *p*-CSA. Sulphonic acids have been detected in industrial waste waters, but being of high polarity, their bioaccumulation potential is much less than that of their precursor hydrocarbons [Alonso 1999]. Some bacteria can desulphonate *p*-TSA to *p*-cresol but then the environmental impact of phenols would have to be considered [Kertesz 1994]. Nothing can be concluded with any certainty until a full LCA is calculated. What may be said of the systems in this work is that the quantity of *p*-CSA used has always been small (it was 2.5 g per mole of substrate in the synthesis of ethyl levulinate for example) and so barring any intensely acute toxicity problems its impact on the environment should be minimal. If at any point in the future *p*-CSA emerges as an acid catalyst and is widely adopted for this purpose it would be subjected to tighter controls and greater scrutiny.

Although it seems obvious that the sustainability of *p*-CSA will be greater than petroleum derived *p*-TSA this does not certainly have to be the case. Sustainability is a complex entity incorporating social and economic implications in addition to environmental issues [UN 1987]. One consequence of this is that the manufacturing process of *p*-CSA would have to be profitable and fulfil an ongoing market demand. An impending fossil fuel reserve crisis will probably force this issue at some point in the not too distant future, which makes the assessment of bio-based product performance a wise endeavour.

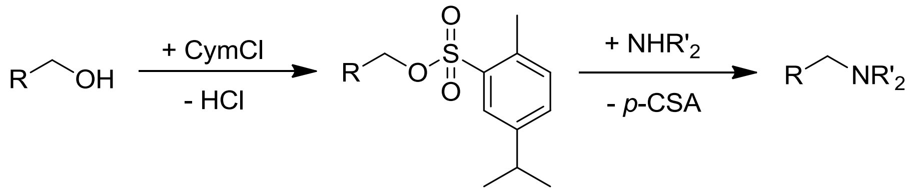
5.3 Catalysed carbonyl addition summary

General remarks on acid catalysed esterification: The solvent effect dictating the rate of an acid catalysed acetylation of benzyl alcohol mirrored an earlier study by Wells on the Fischer esterification of 2-methoxyacetic acid and benzyl alcohol [Wells 2008]. The inverse proportionality between ln(**k**) and **β** is no different to the solvent effect observed when 1-butanol reacts with butanoic anhydride, or in an uncatalysed amidation to give *N*-benzyl-4-phenylbutanamide. This amount of evidence indicates that carbonyl addition chemistry of this type proceeds rapidly in low basicity solvents because of a lack of entropic interference. The bio-based condensation reaction auxiliaries *p*-cymene and *p*-CSA in combination are ideal for accelerating carbonyl addition chemistry through optimal solvent stabilisation and acid catalysis together.

General remarks on *p*-CSA as an acid catalyst: The organic acid *p*-TSA and its parent compound toluene have become a standard system for a variety of chemical transformations. However in light of recent attitudes and ensuing legislation, toluene may soon no longer be a permissible solvent for manufacturing scale chemistry. Some effort has already been dedicated to phasing this solvent out of use [Constable 2007a, SubsPort 2013]. At best its use is likely to be significantly restricted. Toluene and its sulphonic acid are unsustainable, albeit useful and versatile chemicals. It seems that *p*-CSA is an adequate replacement for *p*-TSA, indeed more so than *p*-cymene is for toluene it might be said. Successful acid catalysed reactions consist of esterification, aldol condensation, and acetalisation. No further considerations need be made within the context of the synthetic application when substituting *p*-TSA for *p*-CSA. This will be of much reassurance to the medicinal and discovery chemists whose primary objective is high throughput, and the best way to achieve this is with established and reliable methods.

The synthetic route to *p*-CSA has potential for improvement, a certain amount of optimisation having already been conducted. Palladium was selected as the best choice in dehydrogenation catalyst, but more detailed studies might find a way of circumventing the use of rare and expensive metal catalysts for what should be an uncomplicated high yielding reaction. In turn it was found that oleum was more effective at sulphonating *p*-cymene than sulphuric acid at 373 K to give the desired product, although it is more hazardous.

Broader appeal: The role of *p*-CSA does not necessarily end here. Many other reactions currently rely on *p*-TSA or chemicals derived from it. A simple substitution to *p*-CSA would be beneficial in most instances. Although leaving groups and protecting groups are discouraged by the principles of green chemistry [Anastas 1998], they are still used in order to realise the total synthesis of natural products and pharmaceuticals [Wuts 2007]. To this end the tosylate protecting/leaving group could be supplanted by the *p*-cymenesulphonate moiety (Scheme 5.13). Procedures for the synthesis of *p*-cymenesulphonyl chloride (CymCl) already exist in the literature [Brown 1950, Huntress 1941]. Although waste problems are not addressed with this approach, using bio-based reaction auxiliaries and protecting groups contributes to a chemical industry independent of non-renewable materials.



Scheme 5.13 A generalised enhanced leaving group strategy using the reagent *p*-cymenesulphonyl chloride.

To establish *p*-CSA as a bio-based product, it would also seem prudent to follow up on some of the previous suggestions for realising the broader appeal of *p*-CSA and its derivatives.

Revisiting some historical procedures, *p*-CSA could be used as an intermediate in the synthesis of bio-derived compounds such as sulphonamides or even phenols and acetone. There is a whole network of chemical transformations that utilise *p*-cymene and limonene as a feedstock, and now with the onset of a heightened interest in bio-based products, *p*-cymene and its derivatives should be able to find a greater role in modern organic chemistry.

6. Heterocycle synthesis: The Biginelli reaction

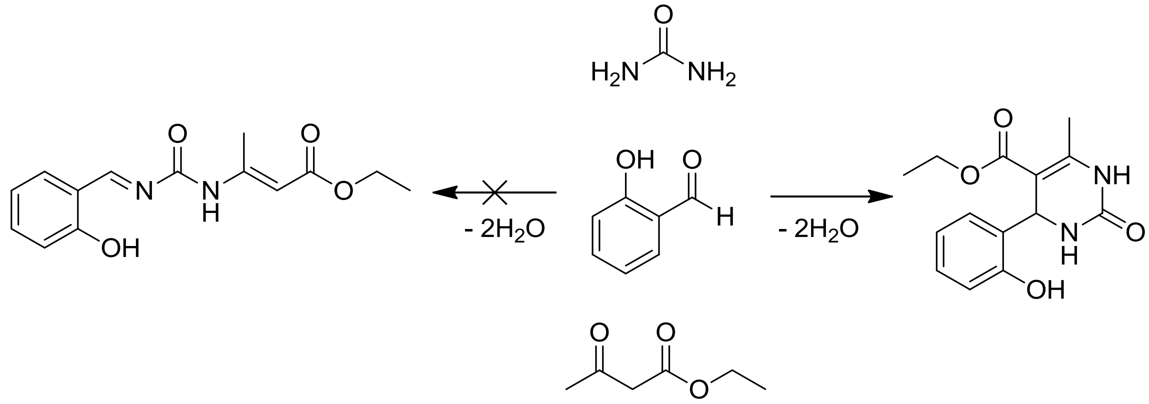
Heterocycles are fundamental structural motifs for drug molecule design. Medicinal chemistry is conducted in a systematic way by functionalising, each in turn, the positions of a core heterocyclic template. The substitutents at the key positions are then substituted and new molecules synthesised to seek out more active compounds. The synthesis of heterocycles is ranked within the top four most prevalent reactions at both process development and manufacturing scale [Carey 2006, Dugger 2005]. The complexity of these (often multicomponent) transformations represents an increase in difficulty from the transformations previously assessed. This is a sterner test of the solvent selection algorithm and should push the art of identifying solvent effects beyond typical case studies. The Biginelli reaction has been selected for this purpose.

6.1 Introduction to Pietro Biginelli and his reaction

Dihydropyrimidinones: One interesting class of heterocycles are the pyrimidines, amongst the most common heterocycles in pharmaceutical process development [Carey 2006]. The Biginelli reaction to give related dihydropyrimidinones consists of the double condensation of urea (or one of its derivatives), an aldehyde, and a 1,3-dicarbonyl compound [Kappe 1993, Tron 2011]. The classical synthetic method requires a large quantity of hydrochloric acid as a catalyst (10 mol%) in either ethanol or sometimes acetic acid as the solvent. Initiated by the first revisit to this reaction by Folkers *et al*. [Folkers 1932], research dedicated to catalysis has been considerable, driven on by the biological activity of Biginelli products [Kappe 1993, Seresh 2012]. Antihypertensive agents, potassium channel antagonists, anti-inflammatories, anti-malarials, anti-bacterials, epilepsy medicines, and other applications are known [Seresh 2012].

Discovery and mechanism: Pietro Biginelli discovered the reaction that is named in his honour in 1891 when working for Hugo Schiff, who had previously condensed urea and salicylaldehyde [Tron 2011]. Although at first Biginelli himself thought a linear compound was created upon the mixing of urea, salicylaldehyde and now also ethyl acetoacetate in acidified ethanol, it is now understood that a cyclised dihydropyrimidinone product had been formed (Scheme 6.1) [Biginelli

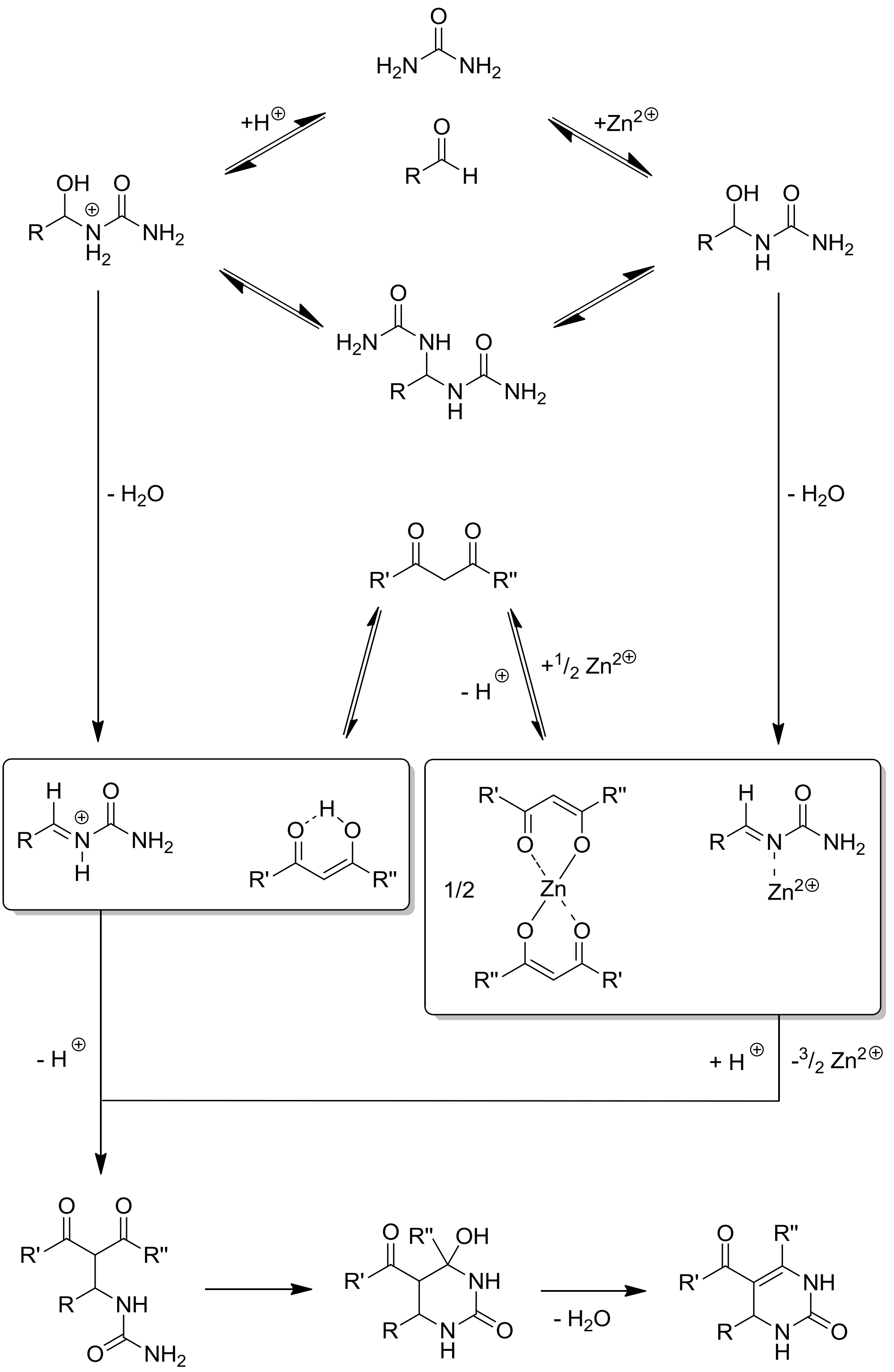
1891a, Biginelli 1891b]. A testament to the usefulness of this reaction are the multiple review articles that document hundreds of conditions by which dihydropyrimidinones can be synthesised [Kappe 2000, Kolosov 2009, Seresh 2012, Wan 2010]. The state of the art is such that it appears that any acid may act as a reliable catalyst if applied in sufficient quantities. Base catalysis is less developed but also viable [Debache 2008, Raj 2011, Shen 2010].



Scheme 6.1 The original Biginelli reaction.

An increase in the depth of knowledge regarding the mode of catalysis and how this differs between Brønsted and Lewis acids has not grown proportionally with the number of articles proposing novel reaction conditions. Typically Lewis acid catalysis is explained with an analogous pathway to the mechanism of Brønsted acid catalysis experimentally derived by Kappe [Kappe 1997], verified by De Souza [De Souza 2009], but proposed some time before by Folkers and Johnson [Folkers 1932, Folkers 1933]. Despite this, matters are complicated by reports of different mechanisms for specific but unusual reaction conditions [Cepanec 2007, Seresh 2012, Shen 2010]. In this work the standard Folkers mechanism will be used unless such evidence is amassed to warrant a new hypothesis (Scheme 6.2). The rate determining step in Brønsted acid catalysed Biginelli reactions is considered to be the initial carbonyl addition between urea and the aldehyde. In proposing Lewis acid promoted mechanisms, the rate determining step is frequently given less attention than metal enolate and metal acylimine complexes at later stages of the reaction [Adibi 2007, Bose 2003, Russowsky 2004].

Solvent effects in the synthesis of dihydropyrimidinones: Considering the large number of investigations into catalysis, by contrast it is surprising that little work has been conducted in order to describe the role of the solvent in the Biginelli reaction, and none of it quantitative [Debache 2008, Dilmaghani 2009, Ghassamipour 2010, Lee 2004]. Deducing the role of the solvent from existing studies is hampered because reactions are often conducted at the boiling points of each solvent and for unequal durations. With the vast number of catalysts available it is



Scheme 6.2 A standard Brønsted acid catalysed Biginelli reaction (left path) and a proposed Lewis acid catalysed Biginelli reaction illustrated with a zinc cation (right path).

possible that different solvent effects could be observed depending on the exact conditions employed. The overwhelming favourite choice of solvent is ethanol, but perhaps only for historical reasons. The yields possible in ethanol are rarely outstanding and so there is scope to improve the productivity of the Biginelli reaction *via* solvent selection protocols. A word of caution however, because ethanol, being a bio-based solvent with few environmental issues, should not be replaced without careful consideration.

Of the work that does exist concerning solvent selection in the Biginelli reaction there seems no clear agreement over a definitive solvent. The yields from many different solvent-catalyst combinations have been compared (Figure 6.1). The five Brønsted acid catalysed reactions, and one each for Lewis acid catalysis, Brønsted base catalysis, and Lewis base catalysis have been summarised in the following table (Table 6.1). Because the reaction temperatures are not consistent and the catalyst loadings not necessarily the same, it is not unexpected that no obvious correlation exists. The objective of all these works was the development of a catalyst which explains the haphazard approach to solvent selection. The essentially random yields across the study make it hard to decide whether there even is a solvent effect governing the productivity of the Biginelli reaction at all. Surprisingly ethanol rarely provides the best yield

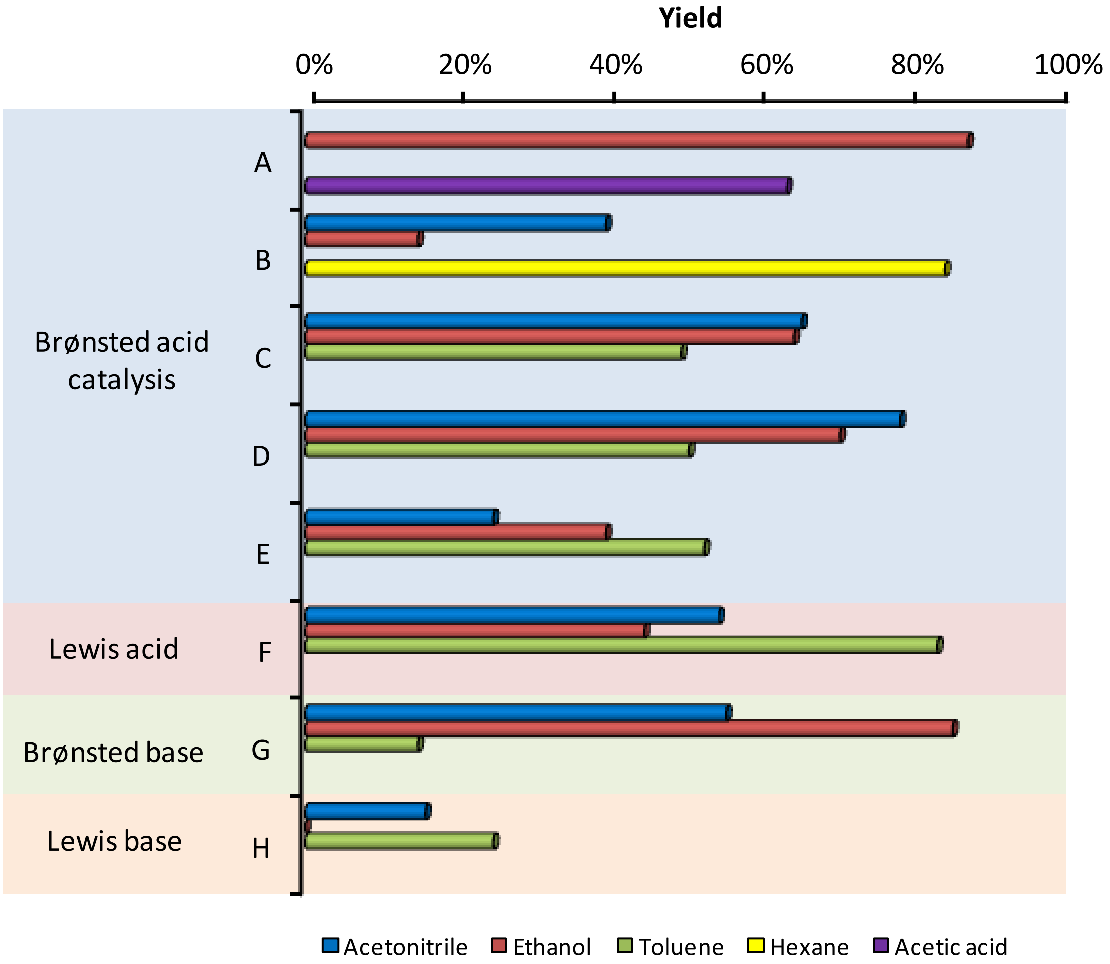
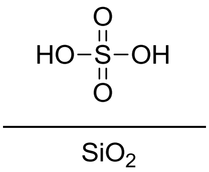
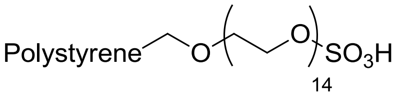
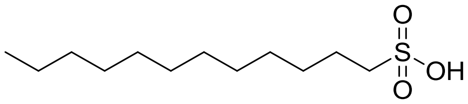
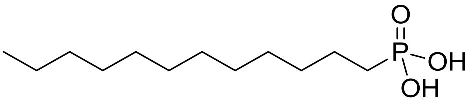
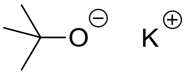
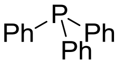


Figure 6.1 Combined catalytic and solvent effects. Letter coding refers to the entries in Table 6.1.

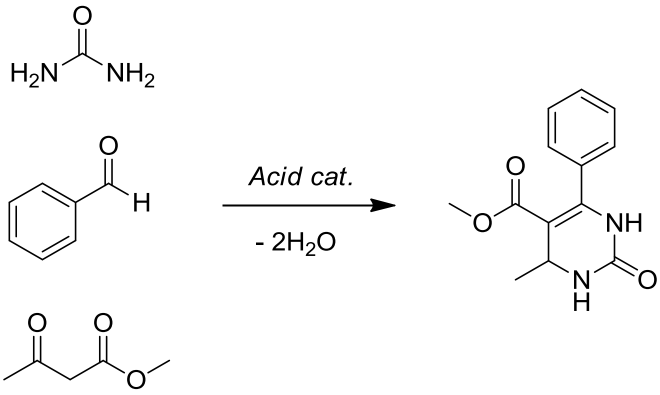
Table 6.1 A selection of catalysts that have been used to promote the standard Biginelli reaction in different solvents.

	Catalyst	Conditions
A	HCl (Hydrochloric acid)	Eight drops of HCl in 40 mL of solvent were refluxed for 3 hours [Folkers 1932].
B	 $\frac{HO-S(=O)_2-OH}{SiO_2}$ (Supported sulphuric acid)	1.5 mol% catalyst, refluxed for 5-7 hours [Dilmaghani 2009].
C	 (Supported PEG-SO_3H)	300 g of catalyst per mole of benzaldehyde was heated for 10 hours at 353 K [Quan 2009].
D	 (Dodecyl sulphonic acid)	10 mol% catalyst for between 4-10 hours under standard conditions [Sharma 2007].
E	 (Dodecyl phosphonic acid)	10 mol% catalyst, refluxed for 3-6 hours [Ghassamipour 2010].
F	$\frac{ZnCl_2}{(Na, Ca)_{0.33}(Al, Mg)_2(Si_4O_{10})(OH)_2}$ (EPZ-10)	10 mol% catalyst, refluxed for 6-48 hours [Lee 2004].
G	 (*t*-BuOK)	Equimolar base at ambient temperature was stirred over 12 hours [Shen 2010].
H	 (Triphenylphosphine)	10 mol% catalyst, refluxed for 18 hours [Debache 2008].

across these case studies. Only trace amounts of product were formed in ethanol with triphenylphosphine [Debache 2008]. What can be gleaned from these examples is that there is no real reason to stray from the usual Brønsted acid catalysed reaction conditions, which work perfectly well, but more might be gained from optimising the solvent.

6.2 Standard Biginelli reaction solvent effects

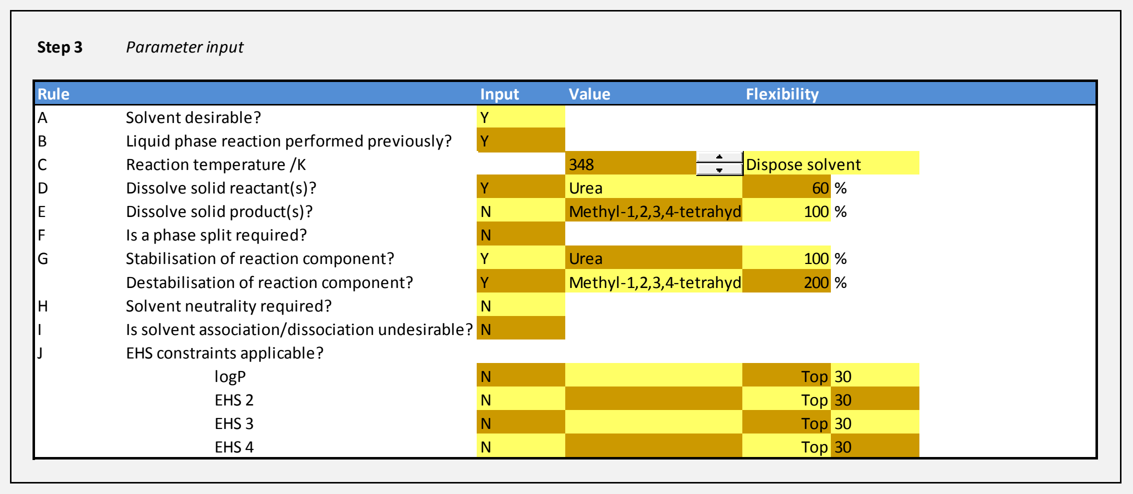
Model reaction: For the initial study a hydrochloric acid catalysed Biginelli reaction was attempted. Urea, benzylaldehyde and methyl acetoacetate were chosen as the reactants to give methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate (Scheme 6.3). A uniform reaction temperature of 348 K was chosen to approximate the boiling point of ethanol, and more importantly for a fair comparison between solvents. In order to assess the interplay between solvent and catalyst, zinc chloride supported on montmorillonite clay (EPZ-10) was also selected as a catalyst for a subsequent set of complementary experiments to come. The comparison between a homogeneous Brønsted acid catalyst and a heterogenous Lewis acidic catalyst should cover close to the extremes of the wide range of materials applied as catalysts in this transformation to date. Instead of the usual kinetic analysis it was expected that the product would precipitate from solution and so yields were recorded instead.



Scheme 6.3 The model standard Biginelli reaction.

Solvent selection: For this case study the solvent selection algorithm will be used prior to any experimental work in order to assess how well the original and revised solvent selection algorithms work without an overreliance on a LSER. Because of this, the solvent performance (model B2) was based on a generic LSER for the rate of carbonyl addition to offer preliminary results (Equation 5.2) [Wells 2008]. It was not known whether this correlation would be appropriate, although the initial nucleophilic attack on benzaldehyde by urea is generally considered to be the rate determining step of this synthesis. The rate of carbonyl addition is by now well established as being inversely proportional to **β**. A lack of any dependence on **β** by the performance of this reaction would strongly suggest that either the wrong rate determining step has been identified, or potentially the kinetics of the reaction are not responsible for the final yield at all. The LSER can be changed in a second reiteration of the solvent selection algorithm to a more suitable alternative once experimental data is available. For this reason no solvent could be absolutely ruled out at this early stage.

Of the candidates in the solvent selection algorithm only basic solvents were omitted (to avoid competition with urea during the reaction). The ability to dissolve and stabilise urea was demanded of the remaining solvent candidates. The radius of the solubility sphere surrounding urea was roughly approximated with a maximum flexibilty of 60% in Rule D, as previously described in the introduction (Figure 1.5). Precipitation (destabilisation) of the product was also incorporated in Rule G as this has been a useful facet of the reaction in the past for retrieving the product (Figure 6.2). Those solvents that adhere to these polarity requirements are unfortunately not predicted to be excellent solvents in terms of accelerating the rate of reaction. Accordingly the top 50 solvents in the solvent performance assessment (model B2) were selected to pass. Only the top 30 entries in the solvent greenness assessment (model B3) were permitted to pass, although no minimum requirements were established with regards to the GSK solvent selection guide categories, only that they existed (*i.e.* a minimum score of 1 in each of the six categories, including LCA). Model B2 and model B3 scores have been double weighted to make 70 the highest possible score when either are combined with the revised solvent selection algorithm (model B1). Previously it was a 1:10 ratio. This change reflects both the cautiousness of applying a LSER to discriminate between solvents before experiments have been conducted, and the greater attention now being paid to the solubility characteristics of the solutes.



Step 3 *Parameter input*

Rule		Input	Value	Flexibility
A	Solvent desirable?	Y		
B	Liquid phase reaction performed previously?	Y		
C	Reaction temperature /K		348	Dispose solvent
D	Dissolve solid reactant(s)?	Y	Urea	60 %
E	Dissolve solid product(s)?	N	Methyl-1,2,3,4-tetrahyd	100 %
F	Is a phase split required?	N		
G	Stabilisation of reaction component?	Y	Urea	100 %
	Destabilisation of reaction component?	Y	Methyl-1,2,3,4-tetrahyd	200 %
H	Solvent neutrality required?	N		
I	Is solvent association/dissociation undesirable?	N		
J	EHS constraints applicable?			
	logP	N		Top 30
	EHS 2	N		Top 30
	EHS 3	N		Top 30
	EHS 4	N		Top 30

Figure 6.2 Biginelli reaction solvent selection algorithm screenshot, step 3: Parameter input.

Only dimethyl carbonate and propionic acid passed all aspects of the solvent selection algorithm (Table 6.2). Solvents scoring highly in the greenness assessment (model B3) such as ethanol were not competitive in the solvent performance assessment (model B2) due to their high polarities. Acetic acid, the other traditional solvent choice, fails the revised solvent selection algorithm (model B1) because it is predicted to dissolve the product which is forbidden in rule E.

Table 6.2 Solvent hits generated by the first iteration of the solvent selection algorithm for the standard Biginelli reaction.

Solvent	Score: Model B1 + model B2	Score: Model B1 + model B3
1,2-Dichloroethane	42	Fail
1,2-Propanediol	Fail	56
1,3-Dioxolane	36	No data
1,3-Propanediol	Fail	56
1,4-Butanediol	No data	58
2-Ethylhexanol	No data	48
3-Pentanone	36	Fail
Acetonitrile	36	Fail
Benzyl alcohol	Fail	42
Bis(2-methoyethyl) ether	40	Fail
Butyric acid	40	No data
Cyclohexanol	Fail	48
Cyclohexanone	Fail	48
Cyclopentanone	Fail	48
Di(ethylene glycol)	No data	60
Diethyoxymethane	34	No data
DGME	No data	40
Dimethyl carbonate	38	42
Ethanol	Fail	46
Ethoxybenzene	44	No data
Ethyl acetate	34	Fail
Ethylene glycol	Fail	64
Glycerol	Fail	64
Isoamyl alcohol	Fail	48

Table 6.2 Solvent hits generated by the first iteration of the solvent selection algorithm for the standard Biginelli reaction (continued).

Solvent	Score: Model B1 + model B2	Score: Model B1 + model B3
Nitrobenzene	46	No data
Nitromethane	46	No data
Propanenitrile	38	No data
Propanoic acid	40	40
Propylene carbonate	44	No data
t-Butyl acetate	No data	48
t-Butanol	Fail	34
Tri(ethylene) glycol	No data	60
Water	Fail	60
Total hits	15	20

An initial screening of eight solvents was undertaken so that a comparison between prediction and experiment could be established. The selection was chosen to represent a wide range of solvent polarity as gauged by the Kamlet-Taft solvent polarity parameters (Figure 6.3). The preliminary results of the solvent selection algorithm were not used to select solvents but instead confirm its success retrospectively. Four protic and four aprotic solvents were included, as well as an equal number of solvents sourced from petroleum and (potentially) renewable feedstocks. Acetic acid [Gorbanev 2012], ethanol [Balat 2009, Hahn-Hägerdal 2006], ethyl acetate [Colley 2004], and ethylene glycol from bio-ethylene [Morschbacker 2009], all have straightforward syntheses beginning with the fermentation of biomass. *t*-Butanol, 1,2-dichloroethane (1,2-DCE), DMF and toluene make up the remainder of this solvent set. This is the most elaborate diversity of solvent polarities selected yet for a case study in this work, helped by the compatability of protic, hydrogen bond donors in the Biginelli reaction.

The predictions of the solvent selection algorithm regarding these eight solvents were compared to the rest of the solvent candidates. Of the solvent set, toluene (poor solubility of urea), DMF (dissolves product) and acetic acid (dissolves product) all failed the revised solvent selection algorithm (model B1) and could not be carried forward into the additional assessments (Table 6.3). Ethanol fails the solvent performance assessment (model B2) because of its relatively high $\boldsymbol{\beta}$ value. This is inconsistent with the prevalent use of ethanol as a solvent in this reaction, casting doubts over the use of this LSER. All eight of the solvents were used in the reaction anyway as a means of determining how successful the solvent selection algorithm was in

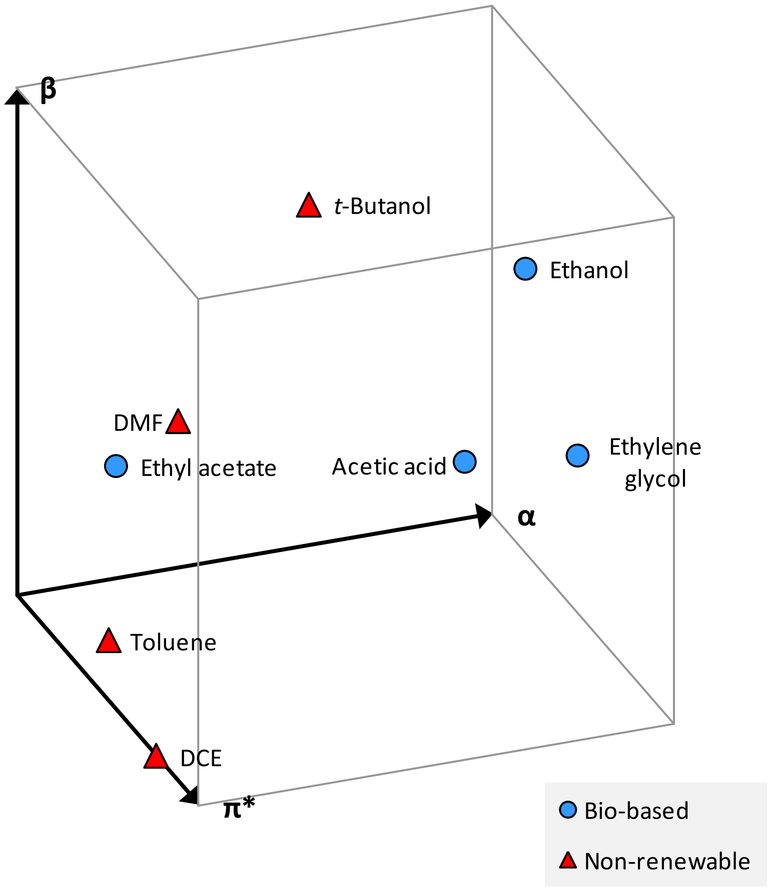


Figure 6.3 The polarity range of the Biginelli reaction solvent set.

Table 6.3 First screening of the experimental solvent set through the solvent selection algorithm describing the standard Biginelli reaction.

Solvent	Score: Model B1 + model B3 (failed assessment in parenthesise)	Score: Model B1 + model B3 (failed assessment in parenthesise)
Acetic acid	Fail (model B1)	Fail (model B1)
t-Butanol	Fail (model B2)	34
1,2-DCE	42	Fail (model B3)
DMF	Fail (model B1)	Fail (model B1)
Ethanol	Fail (model B2)	46
Ethyl acetate	34	Fail (model B3)
Ethylene glycol	Fail (model B2)	64
Toluene	Fail (model B1)	Fail (model B1)

differentiating between the solvent candidates. After the preliminary experiments the solvent selection algorithm inputs can be amended as part of the iterative process of optimisation.

The SUS-HAS-ECO classifications define more acceptable solvents in this solvent set than in the previous case studies (Figure 6.4). This is because alcoholic fermentation products and their close derivatives tend to be fairly ammenable in terms of health and safety and environmental impact, as well of course as being bio-based. The same cannot be said for 1,2-DCE, DMF, and toluene. Ethanol and ethylene glycol are assigned the highest possible classifications, which is made feasible because the data is normalised to give a good spread of results. Inevitably this means the select few of especially high performing solvents get bunched up at the top of each classification scale. In reality occasional issues such as mild toxicity or flammability still exist with these solvents.

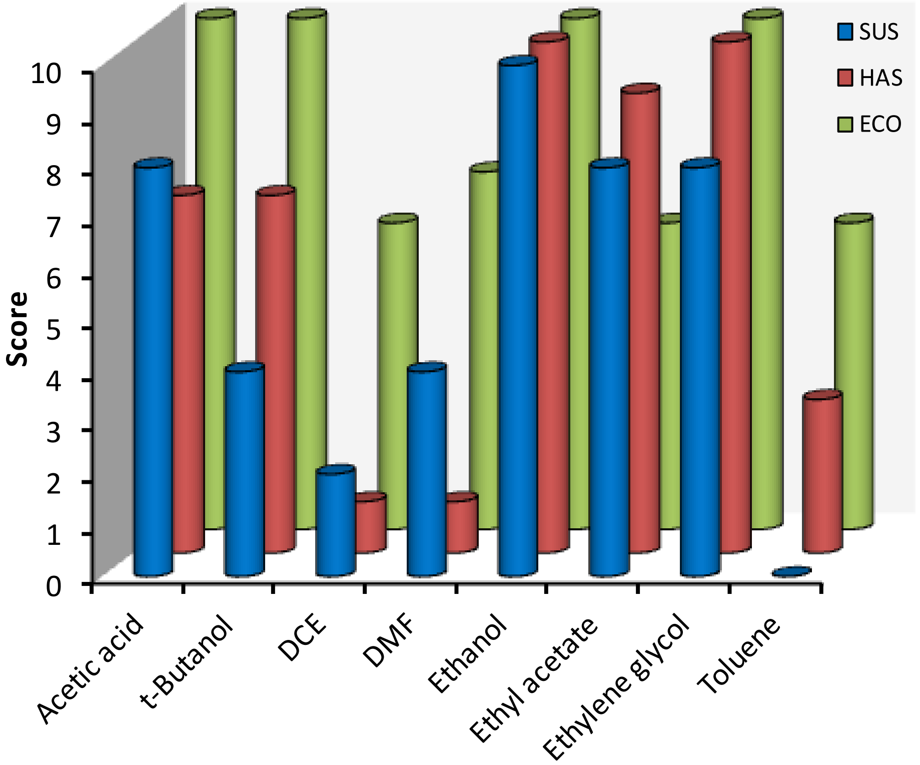


Figure 6.4 The SUS-HAS-ECO classifications of solvents selected for the standard Biginelli reaction case study.

Homogeneous Brønsted acid catalysed Biginelli reaction: Upon addition of hydrochloric acid (10 mol%) to the other reaction components each mixture became homogeneous. This means that the requirement of the solvent selection algorithm for the solvent to dissolve urea is unnecessarily restrictive. Reactions were stirred for 3 hours, but dissolution of the product occurred very soon after addition of the catalyst, often in less than 5 minutes (Table 6.4). Reactions in acetic acid and DMF were the exception to this, remaining in a single phase as predicted. Ethylene glycol created a suspension of the product rather than a distinct precipitate, but the addition of a small amount of water after the reaction produced satisfactory separation. Filtration of the resultant solid after 3 hours was followed by recrystallisation from ethanol. Toluene provides a marginal improvement over ethanol in terms of the eventual yield. But, as warned previously, the proven and commercially successful production of bio-ethanol to give a sustainable solvent cannot be overlooked in favour of marginally enhanced yields. Toluene has been observed as a product of the catalytic pyrolysis of bio-derived feedstocks [Hoang 2009], but at the time of writing bio-derived toluene is not a commodity product.

Table 6.4 Solvent polarity and performance in the synthesis of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate by HCl catalysis.

Solvent	Yield	Solubility (Y/N)		**α**	**β**	**π***
		Urea	*Product*			
Acetic acid	35%	Y	Y	0.71	0.40	0.60
t-Butanol	55%	N	N	0.39	0.95	0.58
1,2-DCE	44%	N	N	0.00	0.00	0.76
DMF	37%	Y	Y	0.00	0.71	0.88
Ethanol	56%	Y	N	0.83	0.77	0.62
Ethyl acetate	51%	N	N	0.00	0.48	0.54
Ethylene glycol	38%	Y	N	0.79	0.57	1.01
Toluene	59%	N	N	0.00	0.12	0.50

The range of product yields from the initial screening provided enough dissimilarity to attempt a correlation between reaction productivity and the nature of the solvent. The yield, although dimensionless, is not a suitable term for a LSER expression. For this purpose a quantity proportional to an energy change is required, with ln(**k**) proportional to the Gibbs free energy of activation for example. Given the large quantity of catalyst and the observation that the reactions appear to be complete before the designated 3 hours suggests that the resulting yields are dictated by the thermodynamics of the system rather than kinetics. An equilibirum constant is not obtainable for the Biginelli reaction, which even if precipitation of the product does not occur, involves an irreversible cyclisation step. For quantification purposes, reaction productivity will refer herein to the natural logarithm of the molar ratio of isolated product (**P**) to unincorporated urea (**R**) as the yield limiting reactant. The form of an equilibrium constant, *e.g.*

ln(**P**/**R**), is used for this metric *in lieu* of an actual equilibrium measurement. Other candidate expressions did not provide a correlation with the polarity of the solvent. The only reasonable LSER correlation was achieved by expressing the productivity of the reaction as a function of **π*** and not **β**, which along with **α** was found to be statistically insignificant (Figure 6.5). The clear exception to this correlation is acetic acid which is less productive than predicted. Without acetic acid the LSER can be described as the following:

Equation 6.1 $$\ln\left(\frac{\mathbf{P}}{\mathbf{R}}\right) = 1.17 - 1.76\boldsymbol{\pi}^{*} \qquad R^2 = 0.862$$

The lower than anticipated R^2 value may be due to the nature of the phenomenon, or it could be indicative of the limitations of the dependant variable. Regardless, when coupled with the fact that the product is made to precipitate, the solvent effect justifies the common choice of ethanol as a solvent. It also nullifies the results of the solvent selection algorithm, highlighting the importance of the combination of experimental observations and a theoretical framework by which to intepret them.

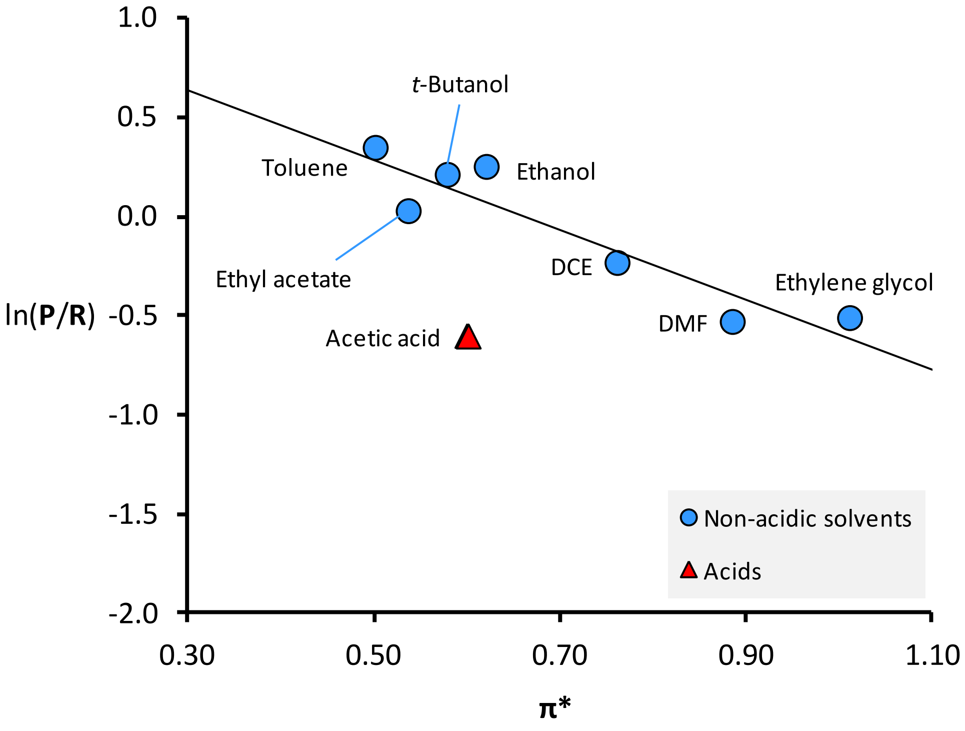


Figure 6.5 The relationship between reaction productivity and **π*** to give methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate by HCl catalysis.

It is noticeable that the inverse proportionality between reaction productivity and **π*** followed by non-acidic solvents mirrors the solvent dependence of the diketo-enol tautomerisation equilibrium constant (**K_T**) of methyl acetoacetate at the concentration applied in the reaction (Figure 6.6). Other alkyl acetoacetates express the same solvent effect trend [Mills 1985, Moriyasu 1986]. These tautomerisation equilibria constants can be determined by UV-vis. spectroscopy or NMR spectroscopy (Figure 8.5). Acetic acid and propanoic acid betray this otherwise strong relationship just as acetic acid does in the relationship shown in Figure 6.5. The reason for this is not clear from previous studies. The ability of acidic solvents to protonate the diketo form of the acetoacetate ester offers the stability of an intramolecular hydrogen bond and may supplant the role of the enol tautomer. Despite being a stronger acid than acetic acid or propanoic acid, tautomerisation of methyl acetoacetate in lactic acid obeys the relationship with **π^***. The low pKa of lactic acid is due to the stabilising intramolecular hydrogen bond character of its conjugate base between the alcoholic and carboxylic moieties of the anion [Losada 2008]. This implies that the conjugate base of an acidic solvent may be responsible for diminishing the enol concentration and not the donated proton. Without their own intramolecular stabilisation, the 'naked' conjugate bases of acetic acid and propanoic acid may interfere with the enol hydrogen bonding system making this tautomer less energetically favorable than otherwise expected (Scheme 6.4).

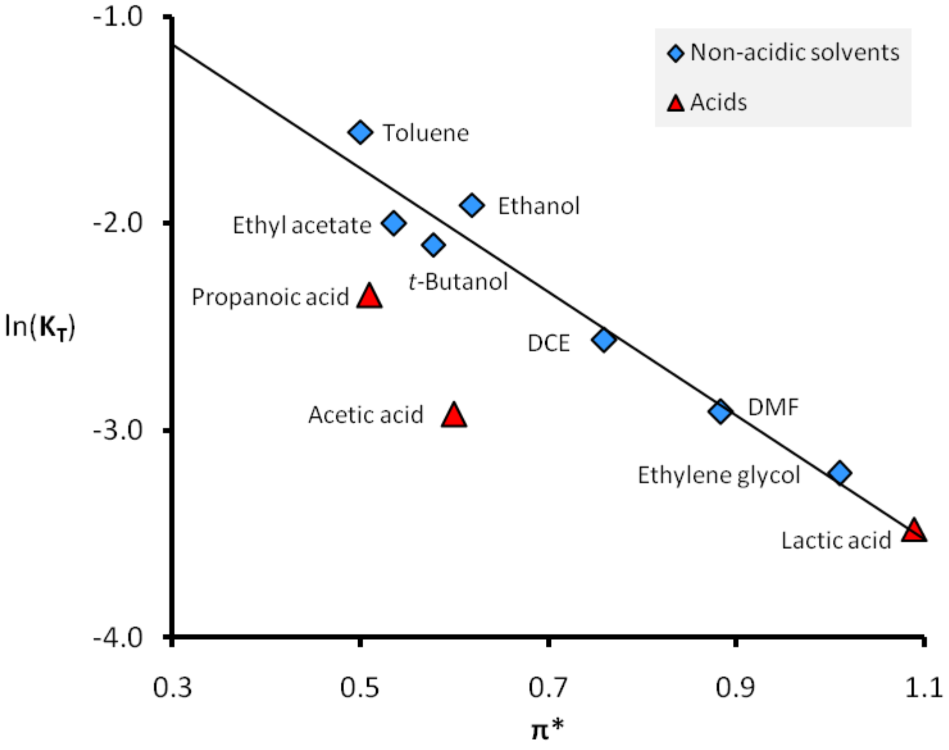
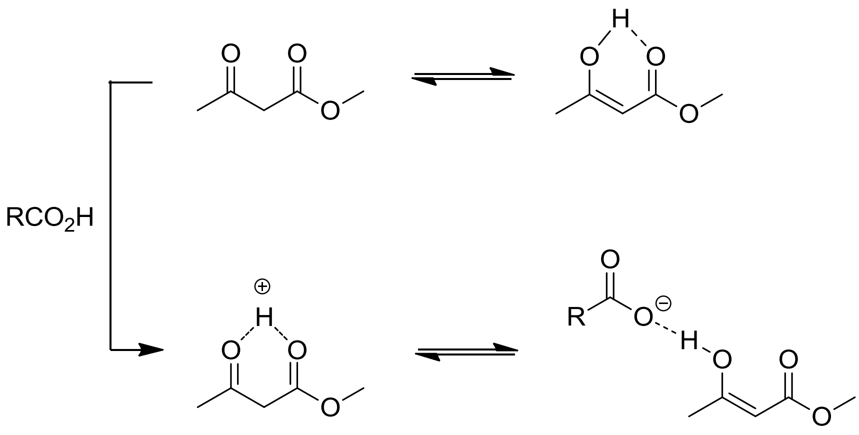


Figure 6.6 Solvent dependence of the tautomerisation equilibrium of methyl acetoacetate at the reaction concentration of 1.875 M.



Scheme 6.4 The effect of acidic solvents on the tautomerisation equilibrium of methyl acetoacetate.

A more suitable relationship in which reaction productivity is expressed as a function of the tautomerisation equilibrium (and not **π***) can be derived, and the predictive capacity of this correlation tested. Predictions should be improved from those obtainable from the previous relationship because of the better data fit:

Equation 6.2
$$\ln\left(\frac{\mathbf{P}}{\mathbf{R}}\right) = 1.59 + 0.74 \cdot \ln(\mathbf{K_T}) \qquad R^2 = 0.925$$

If it is correct to suppose that solvents which promote the enol form of methyl acetoacetate boost the eventual yield, then non-acidic solvents with little dipolarity or polarisability will be ideal. The solvent selection algorithm was applied with the following ammendments: The requirement to dissolve urea was removed (rule D) and the LSER changed to Equation 6.1. Tautomerisation equilibrium constants are not available in the solvent selection algorithm database and so Equation 6.2 could not be used. With the loss of one reaction index the maximum score is now 60 for the revised solvent selection algorithm (model B1) in combination with either the performance assessment (model B2) or the solvent greenness assessment (model B3). Ethanol now reassuringly passes all assessments in the solvent selection algorithm (Table 6.5). The biggest change is to the solvent candidates now suggested as attaining the required level of solvent performance (model B2). *p*-Cymene is now recommended, and the solvent selection algorithm suggests a variety of alcohols, esters and ethers as possible solvents based (indirectly) on the diketo-enol tautomerisation equilibrium position of methyl acetoacetate with Equation 6.1. Although many of the alcoholic solvents narrowly missed the cut-off in the greenness assessment (model B3) there is a case for reinvestigating them as solvent candidates, especially the higher alcohols with low **π*** values, and therefore presumably larger $\mathbf{K_T}$ constants.

Table 6.5 Solvent hits generated by the second iteration of the solvent selection algorithm for the standard Biginelli reaction.

Solvent	Score: Model B1 + model B2 (if different, previous score in parenthesise)	Score: Model B1 + model B3 (if different, previous score in parenthesise)
1,2-Propanediol	Fail	48 (56)
1,3-Propanediol	Fail	48 (56)
1,4-Butanediol	No data	48 (58)
1,4-Dioxane	32 (n/a)	Fail (n/a)
1-Butanol	36 (n/a)	Fail (n/a)
1-Propanol	34 (n/a)	Fail (n/a)
2-Butanol	42 (n/a)	Fail (n/a)
2-MeTHF	30 (n/a)	Fail (n/a)
2-Propanol	32 (n/a)	Fail (n/a)
Benzyl alcohol	Fail	38 (42)
Butyl acetate	38 (n/a)	46 (n/a)
Cineole	44 (n/a)	No data (n/a)
Di(ethylene glycol)	No data	52 (60)
Dibutoxymethane	48 (n/a)	No data (n/a)
Dibutyl ether	44 (n/a)	Fail (n/a)
Diethyoxymethane	38 (34)	No data
Diethyl carbonate	38 (n/a)	No data (n/a)
DGME	No data	36 (40)
Dimethyl carbonate	30 (38)	38 (42)
Ethanol	32 (Fail)	40 (46)
Ethyl acetate	30 (34)	Fail
Ethylene glycol	Fail	56 (64)
Glycerol	Fail	56 (64)
Isoamyl alcohol	44 (Fail)	44 (48)
Isobutanol	42 (n/a)	No data (n/a)
Isopropyl acetate	30 (n/a)	30 (n/a)
Limonene	50 (n/a)	No data (n/a)
p-Cymene	46 (n/a)	No data (n/a)
α-Pinene	46 (n/a)	No data (n/a)

Table 6.5 Solvent hits generated by the second iteration of the solvent selection algorithm for the standard Biginelli reaction (continued).

Solvent	Score: Model B1 + model B2 (if different, previous score in parenthesise)	Score: Model B1 + model B3 (if different, previous score in parenthesise)
Propanoic acid	Fail (40)	36 (40)
p-Xylene	42 (n/a)	Fail (n/a)
Tri(ethylene) glycol	No data	52 (60)
Water	Fail	54 (60)
Total hits	22	16

Cyclohexane and *p*-cymene were selected as candidates for optimal solvents in this model Biginelli reaction, providing yields of 68% and 66% respectively in combination with HCl catalysis. Cyclohexane is produced from a non-renewable feedstock but has a marginally less detrimental environmental, health and safety profile to hydrocarbons of equally low polarity [Henderson 2011]. As we know the limonene in the essential oil of citrus fruits can be converted into the aromatic compound *p*-cymene [Martin-Luengo 2010]. The yield obtained with *p*-cymene acting as the reaction solvent is in good agreement with that predicted from the relationship with tautomerisation equilibrium (Figure 6.7). It is also an improvement over the previous best yield using toluene and HCl as the reaction auxiliaries. Cyclohexane gave a marginally improved yield over *p*-cymene, but deviated significantly from the expected yield. This is likely to be due to solubility issues often associated with a solvent of such low polarity. Reactions in dimethyl carbonate were not investigated because of its hydrogen bond accepting ability ($\boldsymbol{\beta}$ = 0.32) and the electrophilicity of the solvent.

The correlation between tautomerisation equilibrium and productivity is more satisfactory than that using $\boldsymbol{\pi^*}$ because all solvents (with the exception of cyclohexane) can now be accounted for. Observations suggesting that the reaction is quick to complete are essential to the hypothesis that the tautomerisation equilibrium, also a rapid process, is responsible for the reaction productivity. If the reaction was sluggish this would give an opportunity for the reserves of enol to be replenished as the tautomerisation equilibrium seeked to re-establish itself. In doing so the observed solvent effect would be significantly diminished. The formation of intermediate products, as explored by Kappe for example, might explain the yields of less than the theoretical maximum (Scheme 6.5) [Kappe 1997, De Souza 2009]. Side reactions preferentially proceeding *via* the nucleophilic attack of the diketo tautomer and not through the

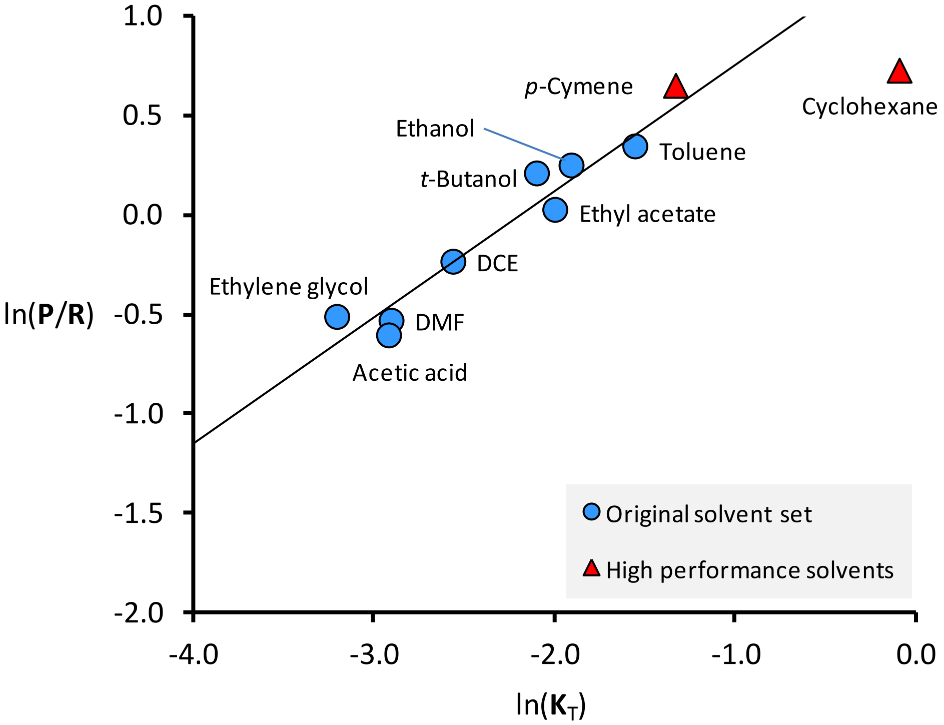
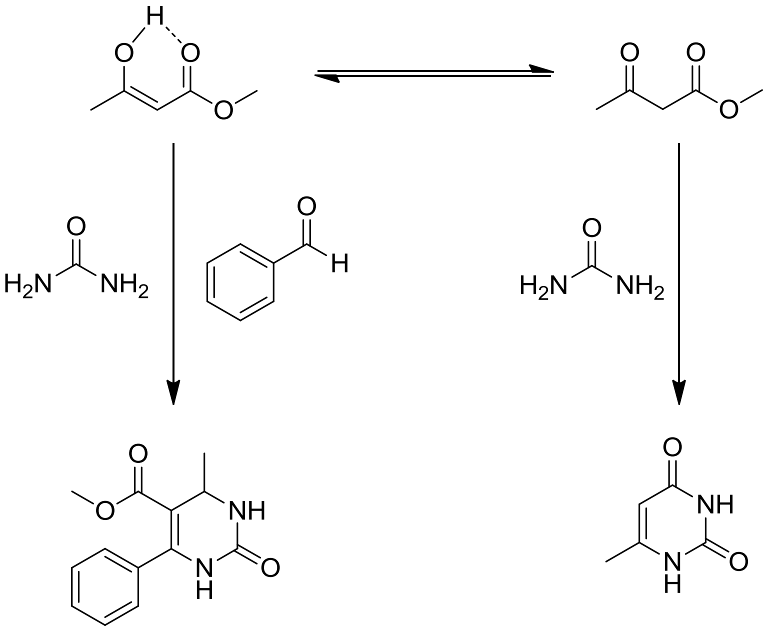


Figure 6.7 HCl catalysed reactions showing the influence of ln($\mathbf{K}_T$) on the yield of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate.



Scheme 6.5 The different reactivity of dicarbonyl compound tautomers.

enol would be required in order to be consistent with this hypothesis. Urea is known to react with acetoactate esters in the presence of an acid catalyst to give a uracil derivative, or at least an imine condensation product [Burgula 2012, Cepanec 2007, Kraljević 2010]. If the reactants are quickly expended by an alternative competing pathway such as this then the amount of enol available at any one time will be vital in dictating the final yield.

Heterogeneous Lewis acid catalysed Biginelli reaction: To clarify both the solvent and catalytic effects on the system the same set of eight solvents was then examined in combination with a heterogeneous Lewis acidic catalyst: EPZ-10, which is simply zinc chloride on a montmorillonite clay support [Clark 1989, Clark 1996, Shaikh 2011]. Given that they both catalyse the reaction, the differences between HCl and EPZ-10 are considerable. A previous study on the EPZ-10 catalysed Biginelli reaction exists, featuring a simple solvent screening study comparing the yields obtained in five refluxing solvents [Lee 2004]. It presented toluene as the optimum solvent (Figure 6.1), akin to the results presented here for the equivalent HCl catalysed procedure.

The catalyst was dried prior to the reaction but otherwise the reaction conditions were unchanged. After three hours the catalyst was filtered from the reaction and washed with acetic acid. The combined organic phase became homogeneous once acetic acid was introduced. Water was then added to induce precipitation of the product, which was subsequently isolated with a second filtration. The productivity of the reaction was generally lower than the analogous HCl catalysed reactions but still dependant on the tautomerisation of methyl acetoacetate as dictated by the solvent (Figure 6.8).

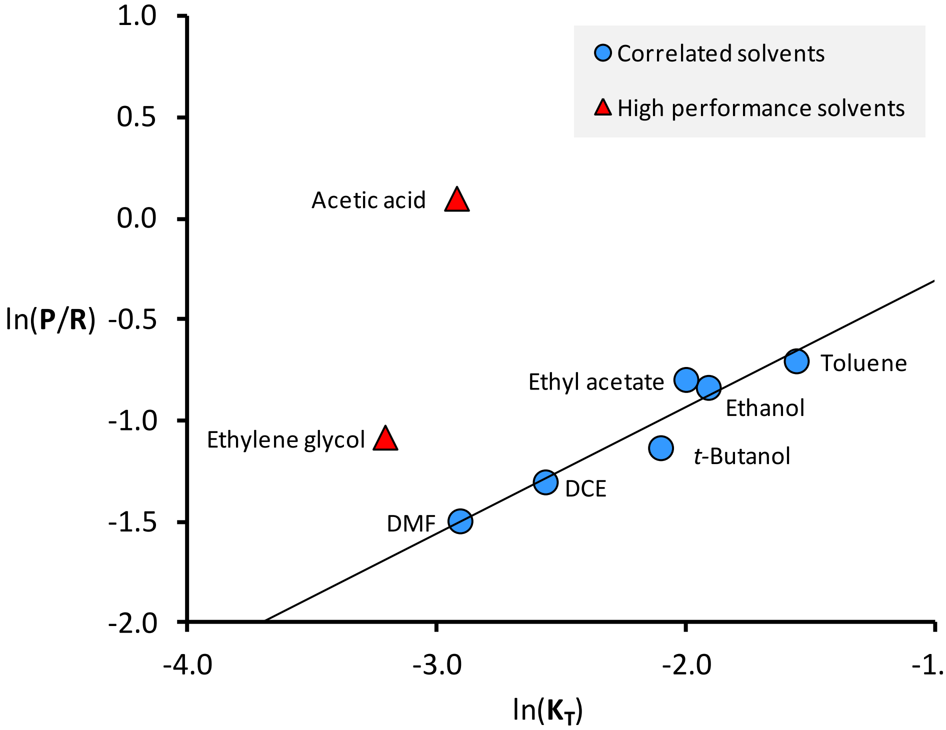


Figure 6.8 Reaction productivity to give methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate by EPZ-10 catalysis.

Interestingly ethylene glycol, but more significantly acetic acid are more productive than expected from the tautomerisation equilibrium constant of methyl acetoacetate in these solvents. The yield achieved in acetic acid was 53%, which is appreciably higher than the 35% yield arising from HCl catalysis in the same solvent. Toluene on the other hand, by adhering to the expected relationship with ln($\mathbf{K}_T$) gave a lower yield of 33% compared to 59%. The reason for enhanced productivity in acetic acid and ethylene glycol is not immediately obvious. Introducing propanoic acid and lactic acid demonstrates that an enhancement to reaction productivity in the presence of EPZ-10 is common to all carboxylic acid solvents (Figure 6.9). The precise trend set by the acidic solvents in Figure 6.9 is not strong, and would require a larger set of solvents in order to correctly attribute a trendline. The increase in yield occurring with ethylene glycol and EPZ-10 in combination is too modest to be regarded with the same esteem as the acidic solvent yield enhancements.

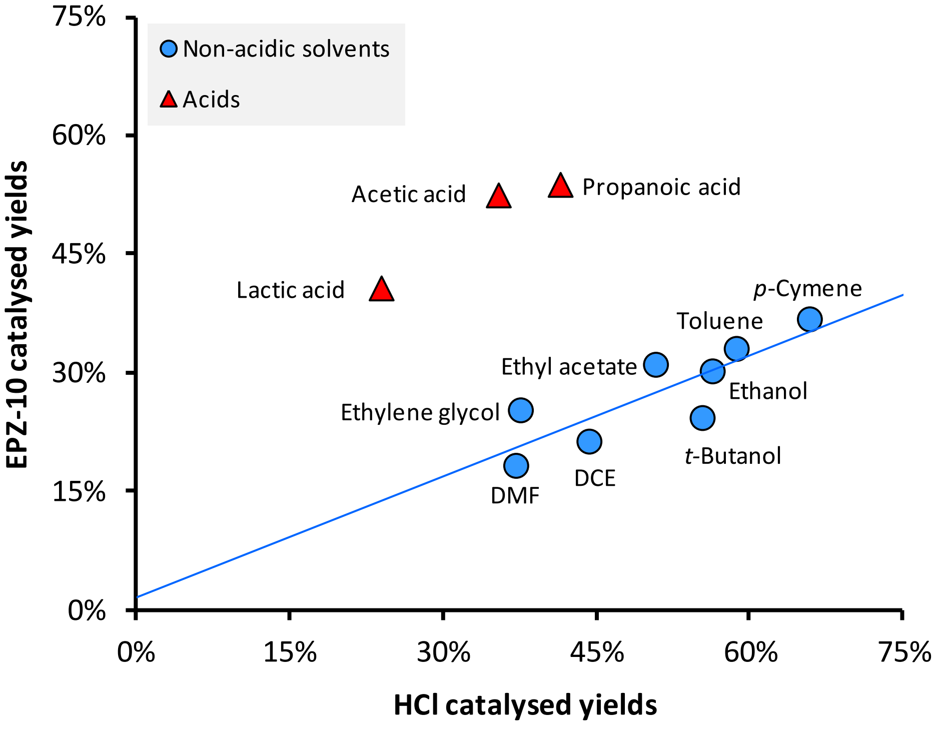


Figure 6.9 Correlating the influence of Brønsted acid and Lewis acid catalysis to the isolated yields of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate.

The productivity trends of catalysis by HCl, EPZ-10 in non-acidic solvents, and EPZ-10 in acidic solvents appear to operate in parallel (Figure 6.10). Exceptions are the ethylene glycol and EPZ-10 system already mentioned, and also the poor yield obtained from lactic acid in combination with HCl. The latter does not seem to be indicative of a special effect that should be considered in more detail because lactic acid performs as expected of an acidic solvent in reactions catalysed by EPZ-10. The trends in Figure 6.10 suggest that the underlying influence of the tautomerisation equilibrium is in effect for each of these three scenarios, but acidic solvents offer a synergetic enhancement that compliments catalysis by EPZ-10.

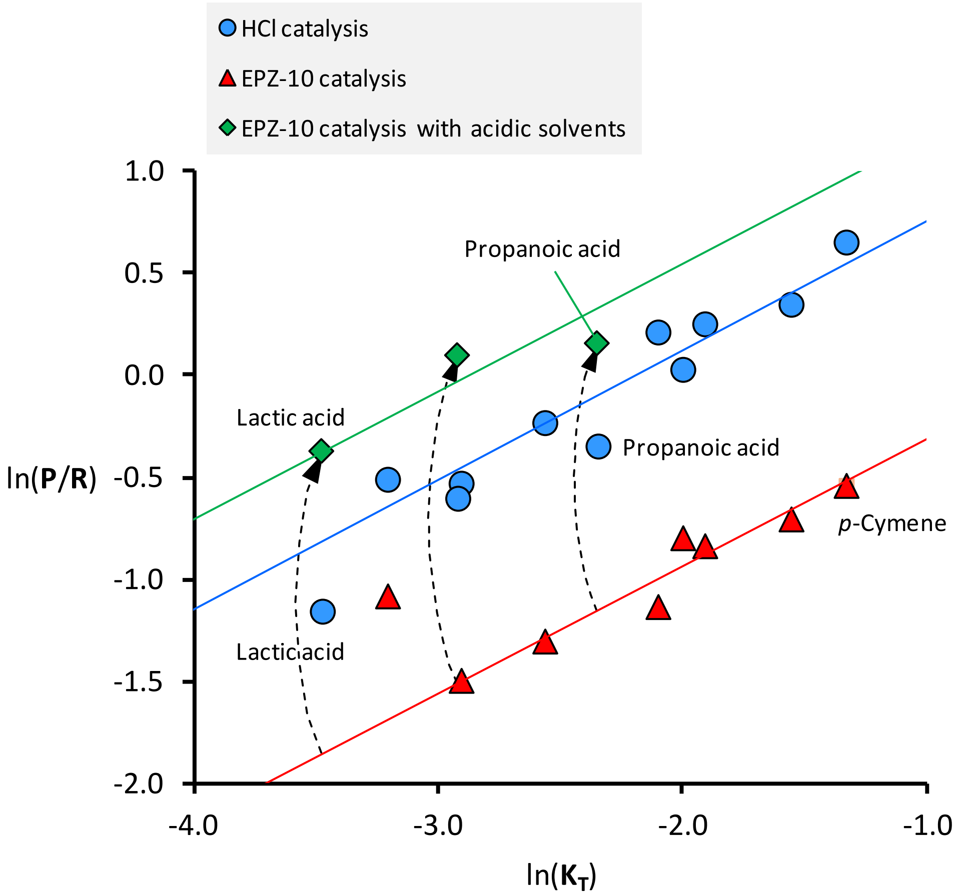
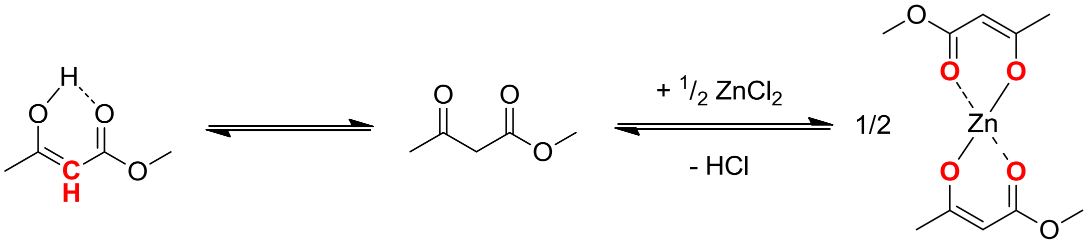


Figure 6.10 A comparison between the solvato-catalytic effects in the synthesis of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate under different conditions. New solvent entries are labelled.

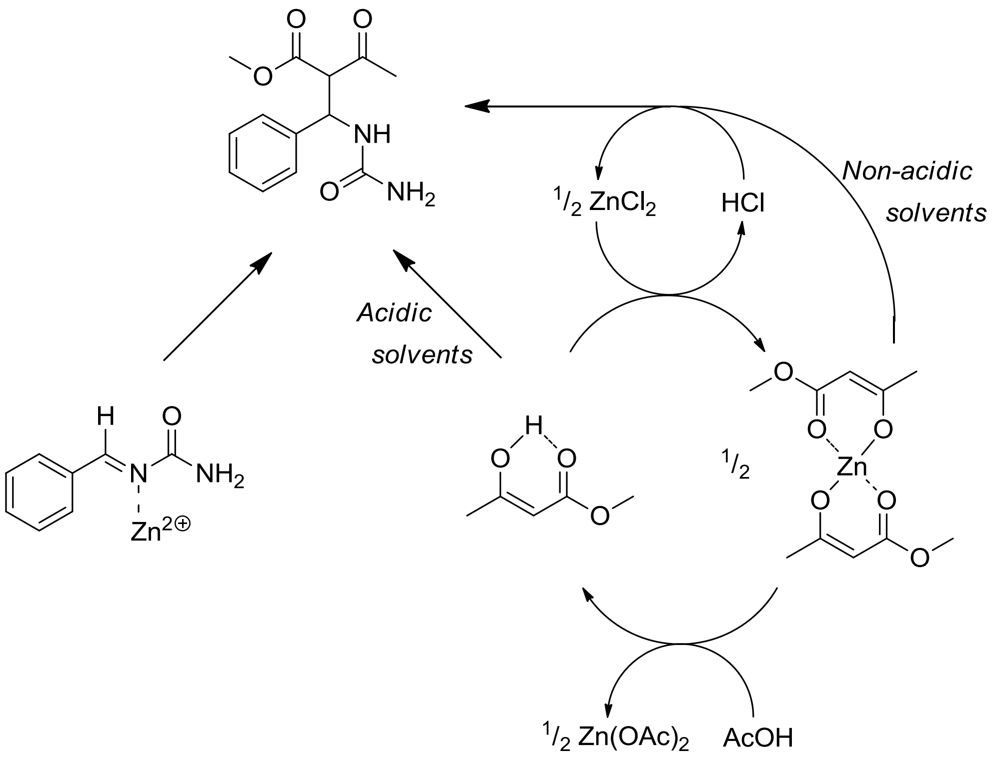
The result of Lewis acid catalysis can be explained by considering the interaction between the metal cation and the chelating methyl acetoacetate. Yields obtained from EPZ-10 catalysis are generally lower than those when HCl is the catalyst. The assumption applied here that thermodynamics and not kinetics determines the yield is not necessarily inconsistent with this observed catalyst effect. Even though catalysis does not modify equilibrium positions like solvents do, the introduction of a stable zinc enolate complex with methyl acetoacetate creates a new reaction pathway (Scheme 6.2). Such a complex is expected to form with all the available enol tautomer but reduce its reactivity. Formation of a metal enolate complex will focus electron density onto the oxygen atoms and not the α-carbon of the ligand, decreasing the desired nucleophilicity of this species. This reduction in nucleophilicity may be responsible for the lower yields in most solvents (Scheme 6.6). The 10 mol% loading of zinc chloride in each reaction (based on 12 wt% in EPZ-10) is enough to exert an appreciable influence, because only *p*-cymene and cyclohexane of the solvents examined will permit more than 20% enol at equilibrium at the concentration applied here. Presuming a 2:1 reaction ratio between the enol and zinc cation means that generally most of the methyl acetoacetate (1.5 equivalents in each reaction) could be suppressed in this manner.



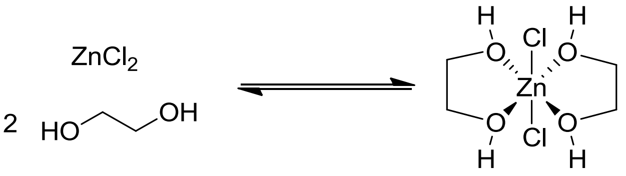
Scheme 6.6 The nucleophilic moieties of methyl acetoacetate tautomers (marked in red).

In acidic solvents the interaction between solvent and zinc cation may displace the enolate ligand from its complex allowing greater reactivity to return (Scheme 6.7). Formation of the zinc enolate liberates an equivalent amount of HCl which would be expected to assist the reaction. However in non-acidic solvents this hydrochloric acid is used up to form the subsequent intermediate, whereas in acidic solvents the hydrochloric acid survives and presumably assists the reaction, perhaps in combination with the resultant zinc acetate. Ethylene glycol produces a similar if less profound effect, which can be explained by a more liable complex between solvent and metal cation as observed in related systems with zinc [Labadi 1993], and nickel [Nylander 1970] (Scheme 6.8). Although the zinc enolate may be broken down in the presence of ethylene glycol, HCl will not be liberated by creating the expected additive complex.

The Brønsted acidity of acetic acid, propanoic acid, and lactic acid may be assisting the reaction directly, although in the presence of HCl this effect should not be significant. Acetic acid and lactic acid have actually been used as catalysts for the Biginelli reaction in place of HCl and so their proticity may become influential in combination with Lewis acidic catalysts [Seresh 2009, El-Hamouly 2006]. However the enhancement to reaction productivity in ethylene glycol, a non-acidic solvent, in combination with EPZ-10 belies this hypothesis. Instead it reaffirms the proposal of a breakdown of the metal enolate complex being greater in importance than contributions from the acidity of the solvent.



Scheme 6.7 Proposed solvent control over the generation of intermediates in the Biginelli reaction.



Scheme 6.8 The complex created by the addition of ethylene glycol to zinc chloride.

What must also be addressed is the possibility of a change in mechanism upon replacing HCl with a Lewis acid catalyst. The analogous solvent effect suggests this is not the case but does not provide comprehensive evidence either way. In reports of antimony trichloride being used as a Lewis acid catalyst the authors claimed that urea would not react with benzaldehyde in its presence, seemingly ruling out the mechanistic pathway of Kappe which relies on this carbonyl addition as the first (and rate determining) step of the reaction [Cepanec 2007]. However, a reaction between urea and benzaldehyde was observed in ethanolic reactions catalysed by EPZ-10, consistent with the mass spectrometry evidence gathered by De Souza *et al*. and the ^{1}H-NMR spectroscopic analysis of the Biginelli reaction by Kappe [De Souza 2009, Kappe 1997]. Hence we find that the mechanism of the Biginelli reaction appears to be the same for both Brønsted and Lewis acid catalysed pathways.

The effect of catalysis observed thus far can be summarised using experiments in acetic acid, ethanol, and *p*-cymene (Figure 6.11). It was found that further yield of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate is minimal when extending the reaction duration from 3 hours to 16 hours, again suggesting predominately thermodynamic control. Interestingly, over 16 hours in the absence of a catalyst a product yield of 25% is obtained in acetic acid. Recalling that acetic acid can be used as the catalyst in the Biginelli reaction, this is not unexpected [El-Hamouly 2006]. In non-acidic solvents uncatalysed yields are more modest over the same time period. Clearly EPZ-10 is superior to HCl as a catalyst only when in combination with acidic solvents, but the combination of HCl and *p*-cymene afforded the highest yield of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate recorded in this case study.

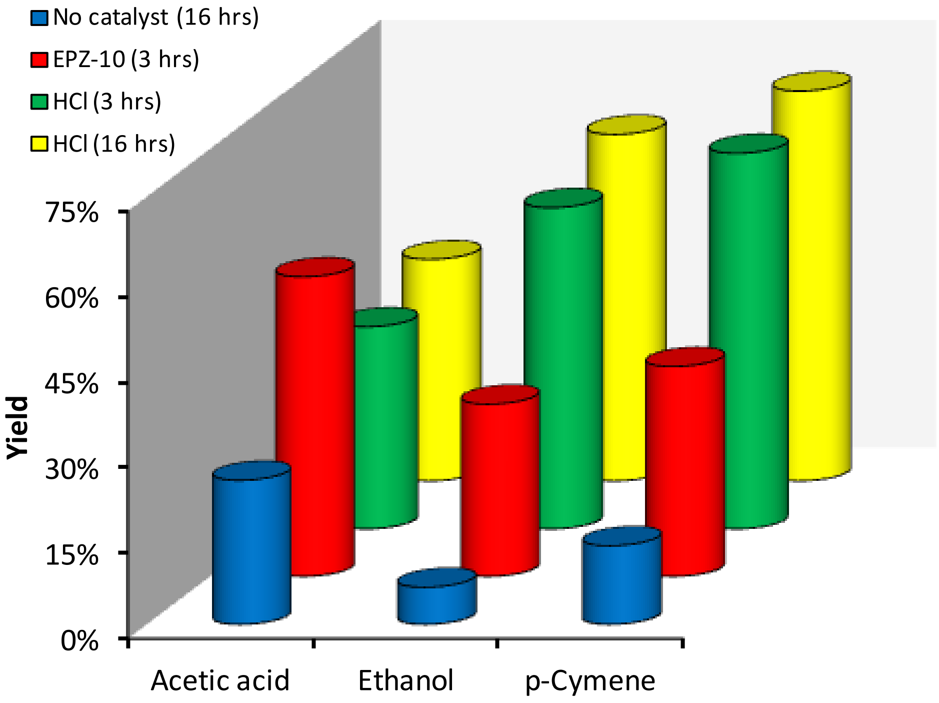


Figure 6.11 A comparison between the isolated yields of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate in acetic acid, ethanol, and *p*-cymene catalysed by either HCl or EPZ-10.

Further solvato-catalytic effects with Lewis acidic catalysts: A more detailed investigation of catalysis in acetic acid provided results not inconsistent with earlier propositions (Figure 6.12). Drying of EPZ-10 was useful but only offered a minimal increase in yield. This catalyst could not be reused successfully which probably means the zinc chloride (which is not intercalated into the clay) leeches off the support during the reaction. The similarity between yields using EPZ-10 and zinc chloride suggests that the clay has a minimal, if any role in the reaction. In acetic acid, zinc acetate is a poorer catalyst than both zinc chloride and EPZ-10, giving credence to the influence of liberated hydrochloric acid with the latter two catalysts (Scheme 6.7).

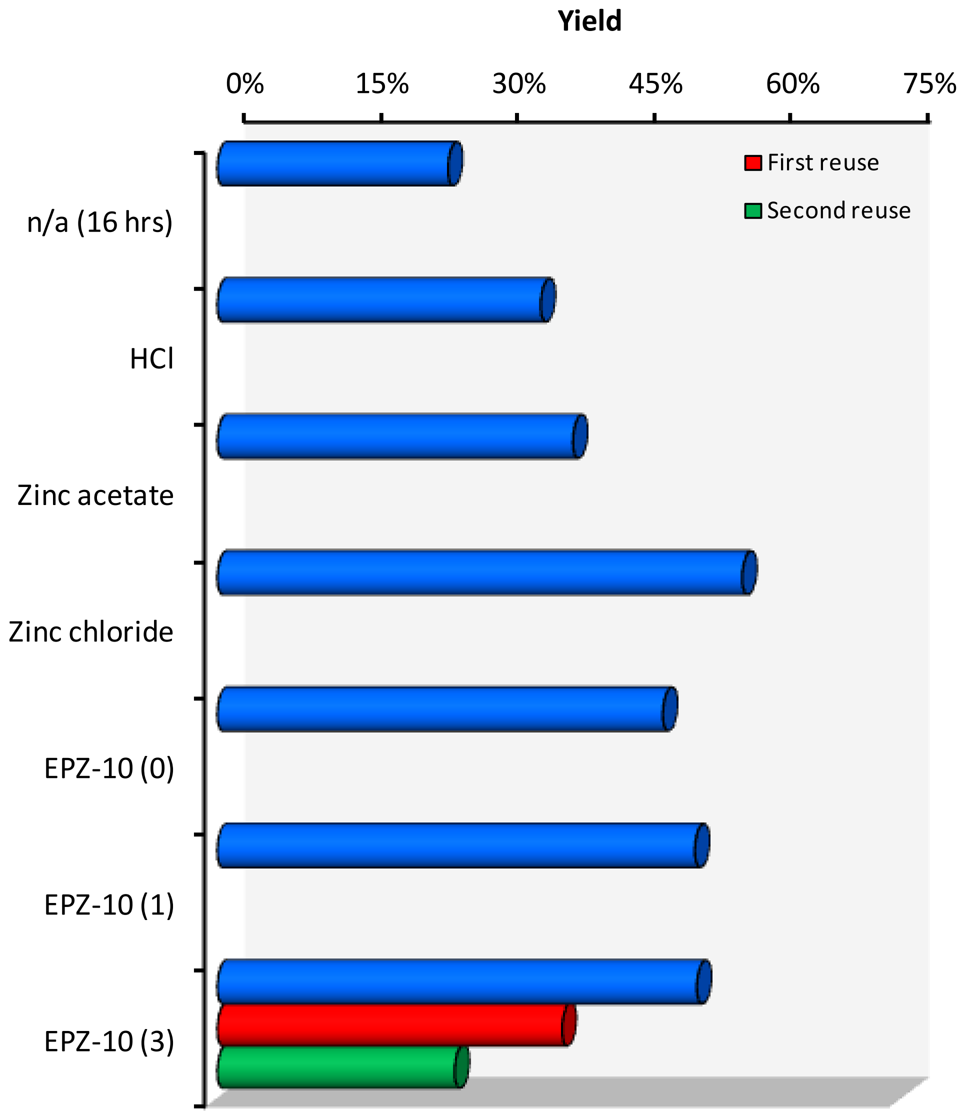
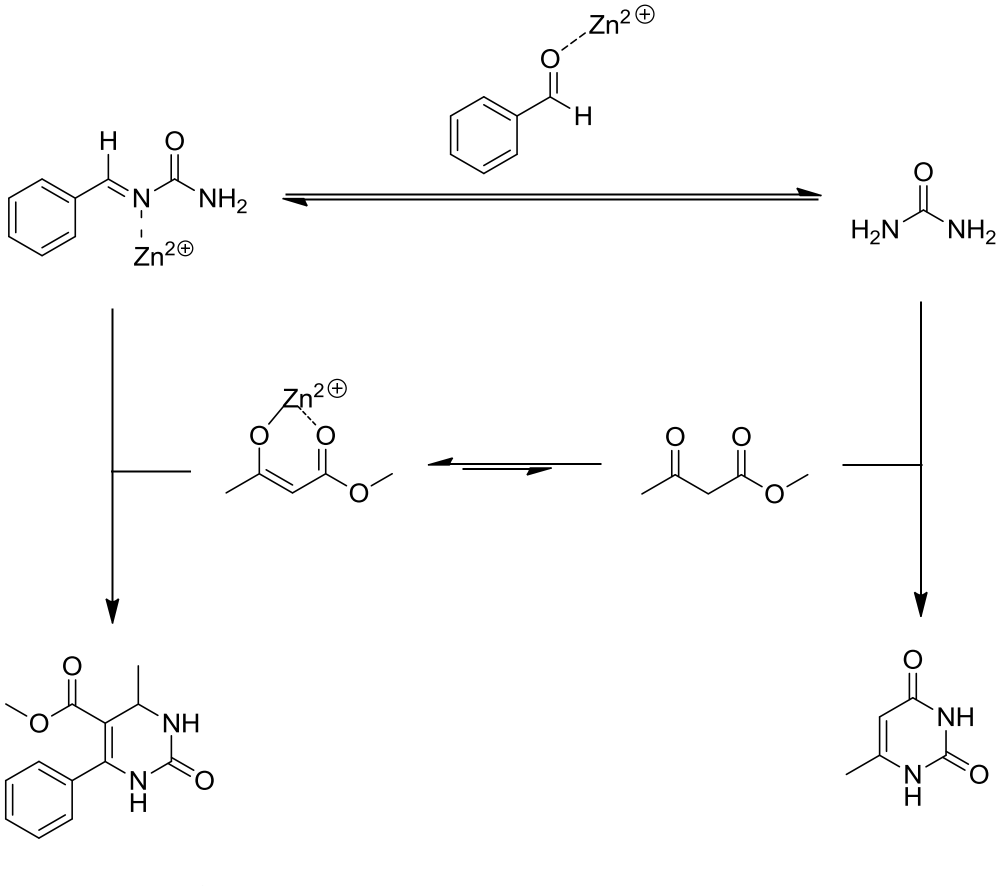


Figure 6.12 A comparison between the isolated yields of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate arising from different methods of catalysis in acetic acid over the course of 3 hours unless otherwise indicated. Activation times of EPZ-10 in parenthesise.

Because reactions in acetic acid without any catalyst are less productive than those catalysed by zinc acetate, the zinc cation must also be influential in assisting the reaction. Of course zinc acetate will not lead to the evolution of hydrochloric acid and so the metal must be actively involved (Scheme 6.7). A likely source of this stabilisation is in the carbonyl addition steps of the reaction, especially the first step of the reaction to give the adduct of urea and benzaldehyde. Dismissing the zinc enolate as a source of catalytic enhancement, the other remaining alternative is the imagined metal-acylimine interaction (Scheme 6.2). This species would seem only to provide a weak contribution to enhancing electrophilicity given the influence of the neighbouring carbonyl group next to the proposed nitrogen donor atom. However these are kinetic effects, and if the hypothesis of thermodynamic control is correct they would be irrelevant. The reaction productivity, expressed as ln(**P**/**R**), may actually be indicative of selectivity (a kinetic phenomena) and not thermodynamics. In this case the unreacted starting material would be incorporated into a competing by-product, giving the equivalent expression ln($\mathbf{P_1}$/$\mathbf{P_2}$). The dual reactivity of methyl acetoacetate, depending on the tautomer in question, will convert urea, our designated yield limiting reactant (**R**), to the desired dihydropyrimidinone (**P**) or a uracil derivative as a result of its reaction with the enol or diketo tautomer respectively (Scheme 6.5). Metal cations diminish the reactivity of the enol by forming a stable complex, but in doing so they might also reduce the concentration of the diketo tautomer (Scheme 6.9). By enhancing the tautomerisation equilibrium, zinc acetate may be able to exert a yield enhancing effect, if not strictly speaking a catalytic one.



Scheme 6.9 The roles of metal cations as Lewis acids in the standard Biginelli reaction.

Reactions with iron(III) chloride were conducted and compared to the performance of zinc chloride. Generally iron(III) chloride was found to be superior to zinc chloride (Figure 6.13). An additional benefit to the iron salt is that a purple colour is observed in acetic acid and ethanol indicative of the metal enolate intermediate [Schüttler 1972, Starke 1963]. As the reaction progresses in acetic acid, the colour fades to a pale yellow (typical of many organic reaction mixtures) as expected if the complex were to be broken down by the solvent. Tellingly, the purple colour remains in ethanol. Co-catalysis with a Lewis acid and a Brønsted acid features infrequently in Biginelli reaction studies [Zorhun 2006], but serves a purpose in this study to highlight the roles of the catalysts. Employing HCl and a metal salt in a non-acidic solvent is less productive than using HCl as the sole catalyst, although the improvement in yield over that obtained with the Lewis acid alone is quite significant. Therefore in this so-called co-catalysis,

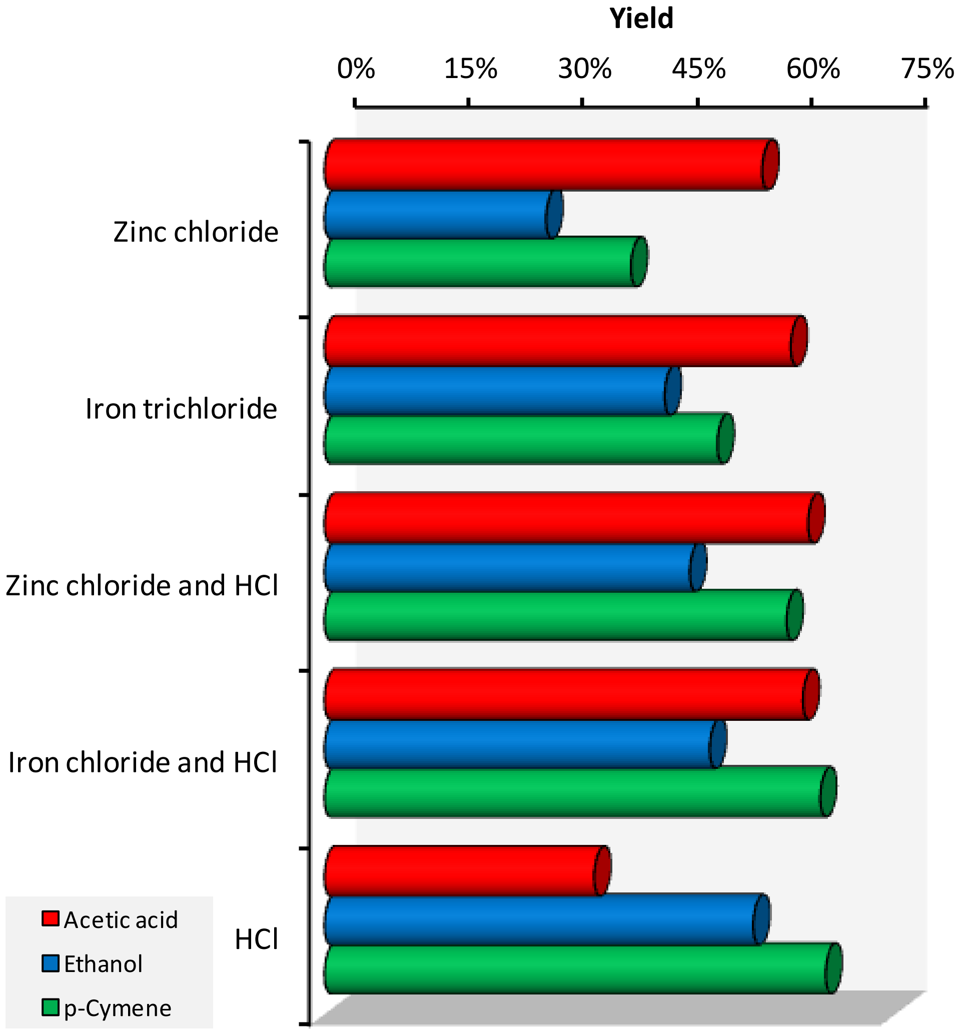


Figure 6.13. A comparison between the isolated yields of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate in acetic acid, ethanol, and *p*-cymene using a variety of catalysts.

while HCl is promoting the reaction, the Lewis acid is probably deactivating the enol by creating the now familiar metal enolate complex. Co-catalysis in acetic acid is underwhelming. As suggested by the observed yields, the liberation of HCl postulated previously with Lewis acid catalysts means the addition of another 10 mol% of HCl is unlikely to be of much benefit in this thermodynamically controlled reaction.

Further solvato-catalytic effects with Brønsted acidic catalysts: In addition to the plethora of examples of Lewis acids being applied to the Biginelli reaction, novel sulphonic acids and sulphonated solids also appear as candidate catalysts [Bose 2004, Gupta 2006, Jin 2002, Konkala 2012, Quan 2009, Sharma 2007]. Using the solvents toluene and *p*-cymene for a case study, a comparison between the yields obtained using their respective sulphonic acids and HCl catalysis was made (Figure 6.14). Although the *p*-cymene based reactions outperform those conducted in toluene, the highest yield achieved in this work still arises from the combination of *p*-cymene and HCl, and not by the application of more complicated catalysts. The use of *p*-TSA in ethanol has already been reported as giving much higher yields than HCl in ethanol [Jin 2002]. However the authors did not make clear that they were comparing the historical Biginelli reactions of Folkers catalysed by 10 mol% of HCl to their own in the presence of 15 mol% *p*-TSA [Folkers 1932]. Accordingly HCl can still be considered as the slightly superior of these two catalysts.

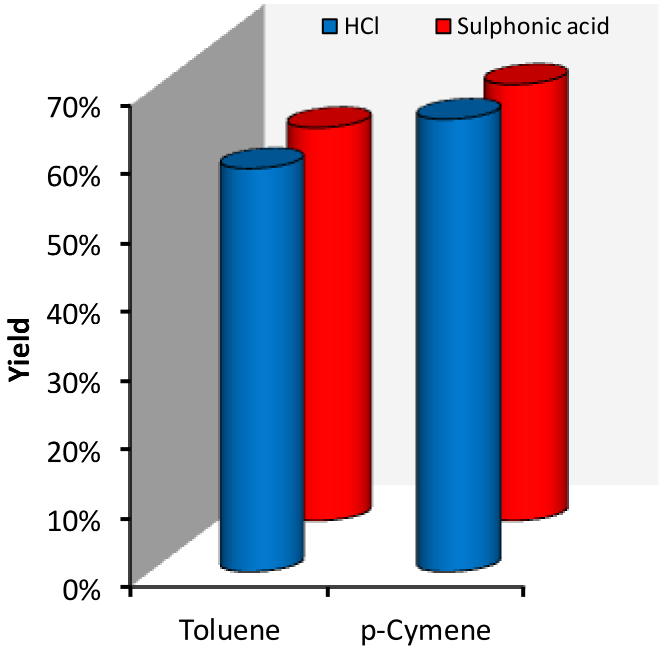
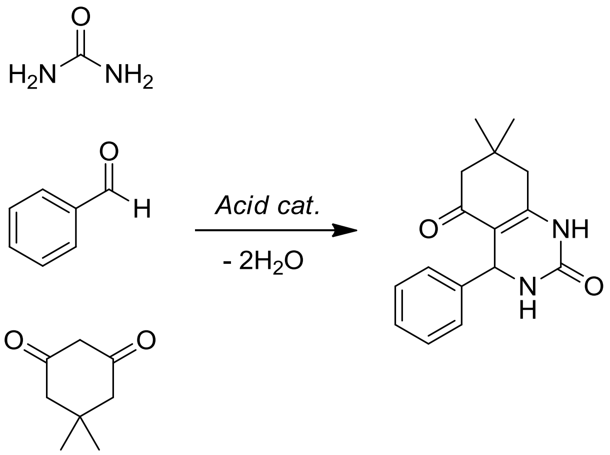


Figure 6.14 The comparison between the isolated yields of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate in aromatic solvents catalysed by their respective sulphonic acids.

6.3 Modified Biginelli reaction solvent effects

Solvent effects in a modified Biginelli reaction: The hypotheses developed from studying the standard Biginelli reaction hinge on the presumption that the tautomerisation equilibrium of methyl acetoacetate determines the productivity of the reaction. In order to provide stronger evidence for this essential premise a modified version of the Biginelli reaction was attempted. Methyl acetoacetate can be replaced with 5,5-dimethyl-1,3-cyclohexanedione to give 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione from the usual Biginelli reaction conditions (Scheme 6.10) [Konkala 2012, Yarım 2003].

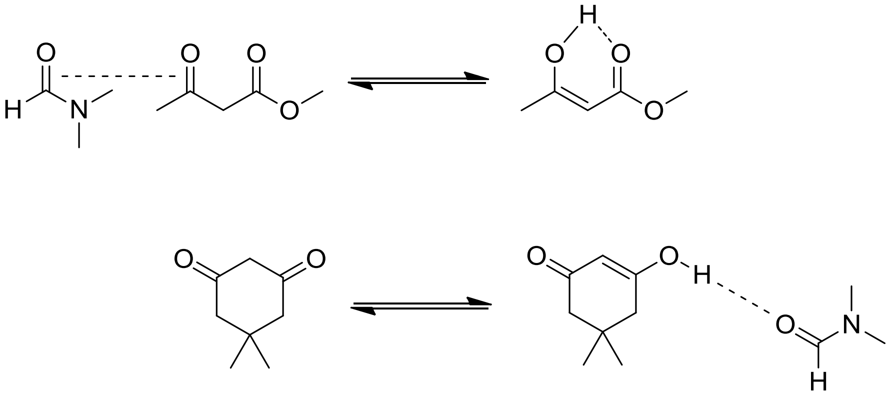


Scheme 6.10 A modified Biginelli reaction with 5,5-dimethyl-1,3-cyclohexanedione to give 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione.

The dependence of the tautomerisation equilibrium constant of 5,5-dimethyl-1,3-cyclohexanedione on the solvent differs markedly from that of methyl acetoacetate (Scheme 6.11). The enol form of 5,5-dimethyl-1,3-cyclohexanedione is increasingly favored as the hydrogen bond basicity (**β**) of the solvent rises [Mills 1985]. Due to conformational constraints this cyclic diketone is unable to create an intramolecular hydrogen bond as methyl acetoacetate does. In the absence of this interaction, the enol tautomer is highly dependent on the stability provided by a solvent-solute hydrogen bond. The range of tautomerisation equilibrium constants of 5,5-dimethyl-1,3-cyclohexanedione vary over a much wider range to that expressed by methyl acetoacetate, with less than a 10% proportion of enol in toluene but over 99% in ethanol (Table 6.6) [Mills 1985]. With the exception of water, the correlation between ln($\mathbf{K}_T$) of 5,5-dimethyl-1,3-cyclohexanedione and **β** is very strong, and so this was used to predict the values of ln($\mathbf{K}_T$) for solvents where data was not available:

Equation 6.3 $$\ln(\mathbf{K_T}) = -3.74 + 11.77\boldsymbol{\beta}$$ $R^2 = 0.962$

The difference in dipolarity between the diketo and enol tautomers that previously gave rise to the dependence on **π*** is lost, meaning that a different solvent effect proportional to **β** should now dictate the performance of the Biginelli reaction.



Scheme 6.11 Differences in the solvent stabilisation and resulting tautomerisation equilibria of the 1,3-dicarbonyl compounds methyl acetoacetate and 5,5-dimethyl-1,3-cyclohexanedione exemplified with DMF.

Table 6.6 The proportions of the enol tautomer of 1,3-dicarbonyl compounds in different solvents.

Solvent	Methyl acetoacetate	5,5-Dimethyl-1,3-cyclohexanedione
Ethanol	12.9%	99.4%
DMF	5.2%	98.8%
Toluene	17.4%	7.4%
Water	2.0%	95.0%

The experimental procedure was similar to that for the standard Biginelli reaction with catalysis provided by HCl (10 mol%). Recrystallisation from ethanol enabled purification of the product. A correlation between solvent polarity and reaction productivity was found using **β** or equally ln($\mathbf{K_T}$), with **π*** now statistically insignificant:

Equation 6.4 $$\ln\left(\frac{\mathbf{P}}{\mathbf{R}}\right) = -3.02 + 3.33\boldsymbol{\beta}$$ $R^2 = 0.899$

Equation 6.5 $$\ln\left(\frac{\mathbf{P}}{\mathbf{R}}\right) = -1.95 + 0.28 \cdot \ln(\mathbf{K_T})$$ $R^2 = 0.902$

This proportionality with **β** is of course indicative of a more fundamental correlation with ln(**K**$_T$) (Figure 6.15). Remembering that the kinetics of carbonyl addition is inversely proportional to **β** (the opposite of the observed effect) suggests that kinetics are still not responsible for the productivity of the reaction. This is strong evidence for the productivity of the Biginelli reaction being dictated only by the tautomerisation equilibrium of the dicarbonyl reactant when in the presence of sufficient catalyst. The solvent effect now in place denies *p*-cymene with its poor hydrogen bond accepting ability being a feasible option. Water performs better than expected from its already favorable tautomerisation equilibrium position. This suggests that although this is a condensation reaction the heterogeneous nature of the system is assisting the reaction. Urea dissolves in water but the other two reaction components have very restricted aqueous solubilities. Accordingly the reaction might only occur at the boundary between phases which forces the reaction to occur in a very concentrated region, enhancing productivity [Chanda 2009]. This hypothesis cannot be confirmed at this stage, but whatever the reason this does mean that a benign solvent such as water is an excellent choice of solvent for this specific transformation.

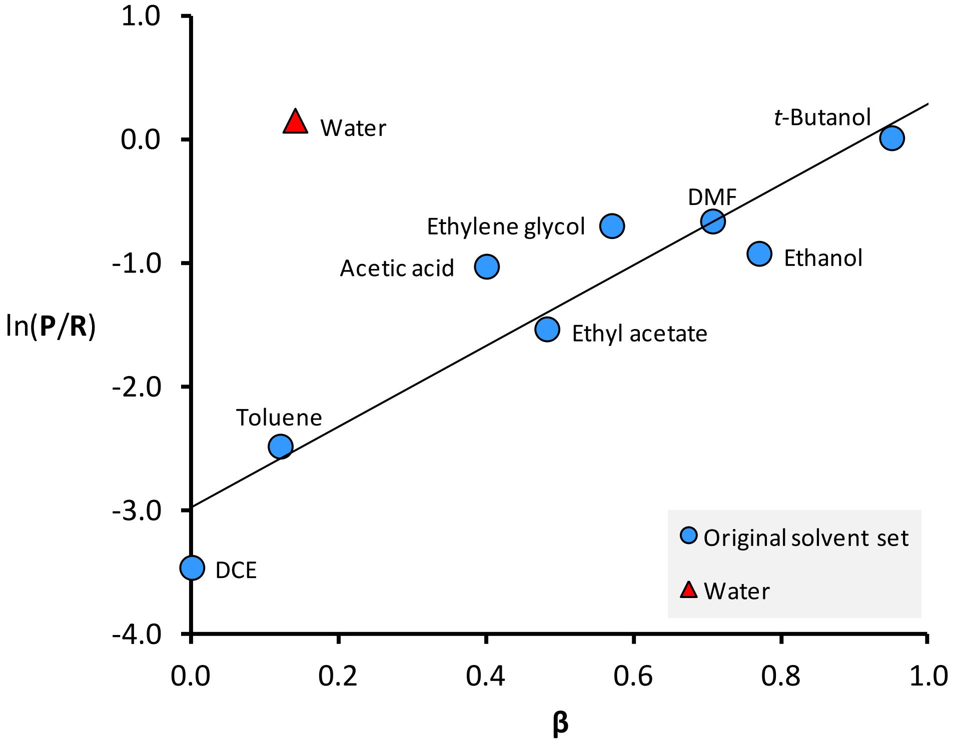


Figure 6.15 The correlation between solvent hydrogen bond accepting ability and the productivity of the reaction to give 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione.

6.4 Biginelli reaction summary

Solvent optimisation and mechanism determination: Overall it seems that the Biginelli reaction is very sensitive to the choice of solvent, catalyst, and the 1,3-dicarbonyl reactant. However all facets of the reaction can be accounted for by the quantity of available enol in the reaction mixture, a phenomenon not recognised in previous studies. The apparent synergy between Lewis acid catalysts and Brønsted acidic solvents may not be synergetic in the most precise sense but instead the liberation of HCl assists the action of the existing Lewis acid catalyst and activates the enol. After elucidating the combined influence of solvent and catalyst with the intention of ensuring that reaction auxiliaries of low environmental impact could be justified as efficient components in the Biginelli reaction, it is hard to justify moving away from HCl as a catalyst. Unless an alternative is equally active at lower loadings (and reusable too) there is little to gain from deviating from the traditional catalyst. More benefit is to be gained from optimising the solvent.

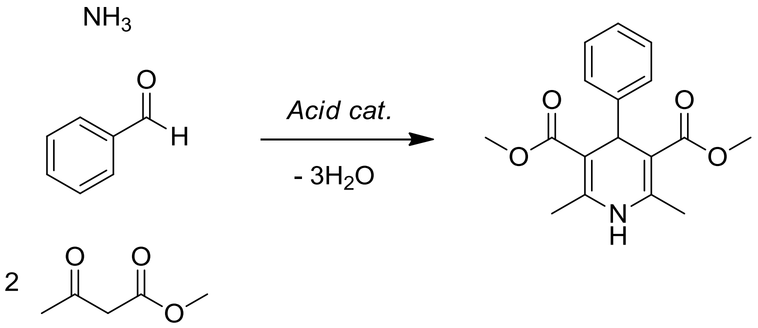
Replacing ethanol, the most popular solvent for the Biginelli reaction with *p*-cymene using HCl catalysis raises the yield obtained from 56% to 66%, an appreciable increase. In addition the greater boiling point of *p*-cymene (440 K) compared to ethanol permits higher temperature reactions if deemed desirable. Water is a very favorable option for a solvent if a cyclic dicarbonyl reactant is employed. Together these two solvents provide a renewable option for maximising the productivity of the Biginelli reaction. Some catalyst development studies favour the use of hydrocarbon solvents over ethanol to an even greater extreme than shown in this work, providing more justification for a move towards alternative solvent systems [Dilmaghani 2009, Zeynizadeh 2009].

Different mechanisms have been continually proposed since a wide interest in the Biginelli reaction evolved. The observation of solvent effects has been a disappointingly underused tool in the past for this purpose. The actual mechanism, at least the one that prevails under the conditions applied here, is consistent with the original suggestion confirmed by Kappe [Kappe 1997]. As revealed by this case study the thermodynamics of the system is probably more important than the kinetic profile. Under different experimental conditions, perhaps with less catalyst, this may change. Although the revelation that diketo-enol tautomerisation is very important in determining the yield of the Biginelli reaction, this does not tell us very much about the actual mechanism of how the product comes to be. What does appear to be the case is that an electrophilic attack on the enol occurs, which has been disputed [Cepanec 2007]. One facet of the mechanism clarified in this work is the role of Lewis acid catalysts. Usually suggested as stabilising various intermediates, it now seems that the little help Lewis acidity provides in the transition state is overwhelmed by its disadvantageous interactions with methyl acetoacetate.

Broader appeal and future work: The appeal of the Biginelli reaction is already wide ranging and so the solvent effects disclosed in this work should be of great use to a number of interested parties. The development of new catalysts for the Biginelli reaction will remain a topic of interest for some time yet. Even sodium chloride has ironically been proposed as a catalyst [Kolosov 2009]. What the authors of this piece of work failed to realise is that the role of a catalyst is to lower the energy requirement of the reaction, and by 'catalysing' the reaction with sodium chloride at 500K ruined their otherwise good point that too much time has been invested researching and developing new catalysts for the Biginelli reaction without improving upon hydrochloric acid. Instead they have emphasised the point that catalysis is the key to making the Biginelli reaction viable, practical and efficient.

This work should be very useful as a tool for the design and implementation of further catalysts, but as stated previously this is superfluous to requirements given the numerous options currently available. A better use of resources would be the optimisation of the synthesis of pharmaceutical products utilising the pyrimidine structural motif. The use of the 'correct' solvent (as defined by the structure of the dicarbonyl reactant) might help facilitate higher efficiency and make the synthesis of these products more appealing from a mass balance and economic perspective. Further solvents, not tested in this work but maybe of some benefit might be fatty acids hydrolysed from triglycerides. Acidic solvents have already proven to be beneficial in the standard Biginelli reaction but are usually too dipolar to be considered as optimal for this task. The substantial hydrophobic region of a fatty acid will lower the **π*** value of the solvent below that of acetic acid. The only issue with these compounds are their relatively high melting points which will hamper isolation of the product. Lauric acid and palmitic acid for example are common, saturated fatty acids that melt at temperatures lower than the reaction temperature of 348 K but are solids at room temperature [Turley 2008 page 28].

The current case study could easily be extended to further analyse a variety of other phenomena. More reactions involving different alkyl acetoacetates and diketone deriviatives would help promote the central hypothesis to this work concerning tautomerisation. The solvent effects in the base catalysed reaction were not considered at all and so great potential lies here as well. The use of lower catalyst loadings might allow a kinetic investigation to be held, providing a way to resolve the mechanistic conundrum that still exists. The most interesting option for further work might well be comparing this case study to an equivalent investigation concerning the Hantzsch synthesis of dihydropyridines for example (Scheme 6.12). The requirement of two equivalents of an acetoacetate reagent should mean that the solvent effects observed here should at least not be diminished, and maybe even enhanced in this related transformation.


NH3
2
Acid cat.
- 3H2O

Scheme 6.12 An example of the Hantzsch dihydropyridine synthesis.

7. Conclusion

The results of each case study have been reviewed in some detail within their individual chapters. Aside from a brief summary of each transformation, what is left to describe here is the success of the methods applied in order to obtain those conclusions, and what improvements can be made in preparation for future studies. This is also an opportunity to develop an idea of how solvents might present themselves in the bio-based economy of the future.

7.1 Case study recapitulation

Carbonyl addition reaction case studies: A strong case for the wider acceptance of limonene, but especially *p*-cymene as bio-based solvents has been made over the course of this work. It was coincidental that three of the four transformations examined favoured one or both of these solvents, but this common theme is beneficial. To justify the adoption of an unfamiliar solvent in only one, perhaps quite specific, reaction might not warrant the effort required, but a widely applicable solvent should achieve general approval much more readily. Whereas the amidation and esterification studies favoured *p*-cymene as a solvent because of its low hydrogen bond accepting ability, the Biginelli reaction is enhanced by replacing ethanol with *p*-cymene because of the weak polarisability of the substitute solvent.

Citrus waste derived limonene and *p*-cymene occupy a useful region of the polarity maps evaluated in the introductory chapter, because most other bio-based solvents are dipolarised oxygenated compounds with a disposition for hydrogen bonding (Figure 1.20). A notable exception is *p*-xylene, which can also be considered as a bio-based solvent when its synthesis from isobutanol is commercialised [Gevo 2013, Peters 2011]. The high boiling points of limonene, *p*-cymene, and *p*-xylene will leave some applications usually suited to hydrocarbon solvents off limits. As such the need for low polarity yet also renewable solvents has not been fulfilled just yet. Generally speaking the need for new bio-based solvents will never be completely satisfied. There will always be demand for improved products and greater consumer choice, and renewable solvents are no different in this respect.

Heteroatom alkylation: The first case study, presented in Chapter 2, was a Menschutkin-type nucleophilic substitution. It was already known prior to this work that this class of reaction favours the highly dipolar and aprotic solvents that are able to stabilise the electronic charge generated in the intermediary stages of the reaction [Schleicher 2009]. This is a class of transformation that hydrocarbon solvents like limonene cannot adequately serve, and to a certain degree this is the reason why it was studied. The typical oxygenated bio-based solvents (acetone, ethanol, glycerol, 2-MeTHF, *etc*.) are either not inert to the reaction components, eliminate the nucleophilicity of the amine reactant through hydrogen bonds, or simply are not dipolar enough to deserve consideration. Renewable alternatives to the current crop of highly dipolar aprotic amides and sulphur functionalised solvents are few in number. An elaborate synthetic procedure for bio-based NMP has been developed [Lammens 2010]. In a not dissimilar approach, the amides of pyrrolidine were tested as solvent candidates here in this work (Scheme 2.9). However their uncertain toxicity and undetermined environmental impact lead to some hesitance regarding further study.

By introducing bio-gas to replace natural gas in conventional solvent manufacturing processes, DMF, DMSO, and other solvents could perhaps be made in a sustainable fashion [Weissermel 1997]. Recreating highly dipolar aprotic solvents from biomass however does nothing help to minimise their health issues, which arguably trump concerns over feedstock security. It is probable that some of the more amenable highly dipolar aprotic solvents like DMSO will survive tightening legislative measures, and solvent selection options will be supplemented with some new, bio-derived oxygenated solvents in the future. Poly(ethylene glycol) and cyclic carbonates have shown some promise as dipolar aprotics in the applications dominated by DMF and similar solvents [Chandrasekhar 2002b, Pieber 2013]. The useful polarity profile of amides and the oxides of organosulphur compounds will continue to attract synthetic organic chemists wishing to use them as solvents, and green chemists attempting to make less toxic replacements for them. Even though the amide solvents synthesised for this work were not ideal in this respect, the information contained within may help progress towards low toxicity solvent substitutes in the future.

Solvent-catalyst synergy: During the course of these studies the role of catalysis grew in importance to the point where the reaction system (solvent plus catalyst) were being treated together as an inter-related system. The synergy between acid catalysis and the choice of solvent is very important to the function of both. Work in this area stemmed from the novel use of *p*-CSA as an acid catalyst, something that was not anticipated at first, but complimented the use of *p*-cymene as a renewable solvent. Once the combined *p*-cymene and *p*-CSA condensation reaction system was developed it was found to be of equally broad utility as its toluene derived cousin.

Success of the solvent effect screening methodology: This work has documented several attempts to optimise the solvent in organic transformations of use to the pharmaceutical industry. The approach taken in order to fulfil this task was conscientiously systemised, adopting computational tools, a simple method of quantifying reaction performance, and above all a focus on replacing conventional non-renewable solvents with bio-derived substitutes. The solvent selection algorithm and its associated assessments were very successful at presenting viable solvent options for a given transformation, once a preliminary solvent study had been conducted. When used prior to any experimental work the results again appeared to be satisfactory. In the case of the standard Biginelli reaction however, the solvent selection algorithm was unable to arrive at the unintuitive but accurate conclusion that hydrocarbon solvents are ideally suited to maximise the product yield. In fact this conclusion is not surprising at all upon completion of a short solvent screening exercise. If we consider the LSER and its interpretation, *i.e.* the solvent performance assessment (model B2), together with the greenness assessment (model B3), then it was these two tools that were of most use for solvent optimisation. The original solvent selection guide (model A) was criticised in the introduction to this work as not progressing beyond simple, often intuitive criteria that could be scrutinised by most chemists without relying on any form of computation. Although how these rules were judged was improved in the revised solvent selection algorithm (model B1), the underlying system was no better. Still, there is a use for the revised solvent selection algorithm. Optimising the performance and greenness of the reaction medium relies on the databank of physical property data contained within the solvent selection algorithm. So if only as a means of processing this data, the computational aspect of solvent selection still remains relevant.

The most rewarding exercise within each case study was the construction of an LSER, first to elucidate the underlying influence of the solvent, and then to extrapolate the relationship to suggest high performance solvents from a renewable source. This solvatokinetic assessment would not have been of any great utility had it not been for a reliable method of calculating the rate of reaction. The use of ^{1}H-NMR spectroscopy, and its advantages and disadvantages compared to other methods of analysis were dealt with in Chapter 1. In practice these attributes of NMR spectroscopy, *pro et contra,* were realised without notable issues arising. Reactants for each case study were chosen with the preferred analysis in mind, and certain solvents could not be used without masking key signals in the spectra, but otherwise the use of ^{1}H-NMR spectroscopy imposed few limits on the chemistry being undertaken.

7.2 The future of bio-based solvents

The bio-based economy: Ultimately the use of fossil fuel feedstocks for chemical production will become impractical. Either the current power structure and manufacturing infrastructure will survive to some extent, with biomass converted to hydrocarbon platform molecules, or those compounds found in biomass that might currently be considered only as exotic curiosities to the platform chemical industry could be used more effectively. Beginning with the simplest argument, the synthesis of hydrocarbons from biomass robs the feedstock of its pre-existing functionality. Subtleties such as chirality will also be lost. Preserving the effort already made in bio-synthesis should be appealing to the producers of chemicals. Conversely, by continuing to use the established methods of chemical production, less research effort will be required in order to continue producing high demand products in high volumes. A further bonus is that there is no need to persuade consumers to use replacement bio-based products when the products are essentially the same as the originals.

Inevitably both scenarios will share the marketplace to some degree, but it remains to be seen how much the landscape of the bulk chemical industry can be changed. However, to what extent chemical manufacturing adapts and evolves in the future is certainly beyond the influence of the demand for bio-based solvents. Examining the sizes of the different chemical product markets can only lead to the conclusion that it must be the bio-plastics and fuel sectors that will determine the nature of the chemical industry to come. The primary uses of crude oil at present (with about 13 billion litres consumed every day) are as fuel (90%) and for making plastic (6%) [Achema 2012, BP 2013]. The remaining 4% of the annual crude oil crop is enough to satisfy society's present need for solvents and all other petroleum derived products.

The global solvent market is 20 million tonnes per year [Achema 2012, Kerton 2009 page 2]. Therefore plastics (265 million metric tonnes per year) are an order of magnitude more important than solvents in terms of production effort. Even recognisable solvent products such as 1-butanol are more frequently used to make plastics and plasticisers than they are used as solvents [Chemical Strategies Group 2013]. Consequently the availability and precise structure of bio-based solvents will depend on other manufacturing processes. This is already true of most contemporary solvents, and must continue this way for economic reasons. This will affect bio-based solvent availability whenever a new manufacturing process is established. Novel synthetic routes unfamiliar to the oil refinery and its associated industries, that in turn generate unfamiliar products, will be a source of future bio-based solvents. Although the exact nature of these future chemicals cannot be foretold with certainty, a contemporary example may illustrate the point. Production volumes of bio-diesel are generally increasing in countries across the world

[EurObser'ER 2012]. The rise of bio-diesel as an alternative fuel has provided glycerol an opportunity to become a successful bio-based solvent and platform molecule [Bauer 2013, Gu 2010]. If it were not an unavoidable by-product of a more valuable process then glycerol would not be studied seriously as a potentially green solvent. This is perhaps at odds with the philosophy purported throughout this work, namely that solvents derived from wastes should be selected because they enhance a reaction and not simply because they are available, but there is no harm in understanding the capabilities of this abundant resource. Before the emergence of green chemistry, glycerol was rarely considered as a solvent due to its high viscosity hindering any reaction that might have occurred within it. Glycerol also has a prohibitively high boiling point for any useful means of removal from the reaction mixture. But now several research groups have dedicated a significant portion, if not all of their resources into studying its application in organic synthesis [Delample 2010, Diaz-Álverez 2013, Wolfson 2009].

It may not be long before a biorefinery based on citrus waste becomes operational too. A manufacturing plant that extracts limonene from citrus waste, perhaps in order to make *p*-cymene, could never be economically viable. But a process that makes ethanol by fermenting the sugars in citrus waste, digesting the resulting residues to produce another fuel product in the form of methane, whilst also presenting limonene and pectin as secondary products is a realistic enterprise [Lorasbi 2010]. Again it is the prospect of selling fuel, and maybe expensive niche chemicals that makes the processing of food waste an attractive prospect, not the production of solvents.

Ethyl lactate is another bio-based solvent that is gaining popularity [Pereira 2011]. The only reason why ethyl lactate has aroused an interest within solution chemistry is because it shares a precursor, lactic acid, with poly(lactic acid) (PLA). In fact lactic acid, from the anaerobic fermentation of corn starch, has itself been successfully implemented as a solvent in a number of transformations [Yang 2012b]. Annual production of PLA is expected to grow, but with a diminishing percentage share of global bio-plastics sales because of the rapidly expanding bio-PET market [European Bio-plastics 2013]. Without the bio-plastic manufacturing process, little or no effort would have been made to obtain bio-based ethyl lactate for use as a solvent.

Instead of solely relying on niche solvent products from a variety of differing bio-refineries (which would be welcome on an appropriate scale of course), gaining access to the traditional solvents *via* the processes established by the bulk chemical industries would be appealing to most end-users. Obviously those companies holding intellectual property in this area, and with the specialised facilities needed to execute it, find this approach more attractive still. The huge quantities of bio-ethanol fuel produced in Brazil, the USA, and increasingly elsewhere provides a platform molecule just one dehydration reaction away from the established oil derived chemical manufacturing hierarchy [Angelici 2013, Balat 2009]. In a crucial advance for the bio-based economy, bio-ethanol is being upgraded to make poly(ethylene) and PET *via* ethylene. It is here that the strongest short term potential lies for bio-based solvents. Ethylene provides access to ethers and other familiar solvents [Angelici 2013, Fan 2013]. These could be blended into (probably) cheaper but non-renewable versions of themselves in order to ease consumers into what will probably be a more expensive bio-based economy at first, until production volumes increase and bring prices down.

Other attempts to retain the principles of the oil refinery include synthesising liquid aromatics from carbohydrate [Anellotech 2013, Huber 2013]. This technology complements the use of bio-gas and bio-ethanol as platform molecules. To complete the full array of petroleum platform molecules, other potential bio-based hydrocarbon manufacturing operations (specifically those for isobutene and propene) will also have to progress beyond the pilot plant [Global Bioenergies 2013, UOP 2013]. Whether this is the dawn of a bio-based industry developed specifically to mirror that built on the oil refinery and its downstream products is yet to be seen. Although there are arguments against it, moving from non-renewable hydrocarbon platform molecules to their bio-based analogues will at least change the geopolitical dynamic of the world. A global bias in energy reserves will not be eliminated, but instead regions rich in biomass and its wastes will replace the dominant oil producing nations when it comes to monopolies over energy resources. However one can envisage a less extreme scenario to the one we currently operate within [Woolsey 2013].

So could bio-refineries producing bio-diesel from used cooking oil and methane from citrus waste present the beginning of a greater overhaul of the bulk chemical industry? Or is the scenario described in the previous paragraph an inevitable eventuality? It is an unavoidable consequence of the world socio-economic model that available resources will be exploited on the basis of price, be it food waste to make new chemicals, virgin plant feedstocks to replicate petroleum derived chemicals, or crude oil while it is still profitable to refine. As long there is an economic stimulus it will be done. Of course the vision of a truly free market that this statement invokes has never actually existed, and governmental initiatives, either incentives or prohibitive legislation, will have to be the means with which to dictate and direct the future bio-based economy. Hopefully this will provide the motivation for a network of smaller independent organisations to continue to operate bio-refineries, processing local waste-streams in a cost effective manner to produce inventive product streams. Larger hydrocarbon manufacturing plants will never be eliminated it seems, but the petroleum lead energy and chemicals industry can be adapted.

Bio-based solvent strategy: The current bio-based solvent market is probably less than 2% of the total solvent market, although estimating this value is difficult [Achema 2012]. The majority of bio-based solvents are plant oils and simple fermentation products in the form of cleaning products. The most exotic of these will be no more than one synthetic step from a fermentation product. Ethyl lactate is one example, currently blended with either citrus oils or FAMEs and marketed as a cleaning solvent [Henneberry 2004]. In terms of solvents of interest to the synthetic organic chemistry industries, acetone [Qureshi 2001], ethanol [Hahn-Hägerdal 2006], and ethyl acetate are amongst the most obvious bio-based examples [Colley 2004]. The types of bio-based solvent that are most likely to be introduced in the near future are still those that directly replace familiar solvents of a petrochemical origin. A chemist will not necessary know that their acetone or ethyl acetate is bio-derived in the future, or indeed a blend of a bio-based solvent and the same petroleum sourced solvent. 2-Methyltetrahydrofuran is a slight oddity in this respect, given that it possesses an extra carbon atom compared to its most obvious substitute, THF [Aycock 2007]. Because 2-MeTHF is so structurally similar to THF, few objections have been raised, and the growing presence of 2-MeTHF is a testament to the willingness of chemists to adapt to the problems presented by feedstock security [Pace 2012]. This is not to say that 2-MeTHF is inferior to THF, quite the opposite in fact, and this must be true of all bio-based solvents if they are to be successful.

Whereas the phasing out of THF in favour of 2-MeTHF is a slow but unproblematic task, the same might not be true of future changes in solvent use habits. The day may come when legislation or the management structure overseeing medicinal chemistry (one of several relevant examples) rule against the use of a vital solvent. Given current attitudes and the influence of REACH, solvents such as NMP and toluene seem particularly vulnerable [EC 2007]. So whereas replacing non-renewable acetic acid, ethanol, ethyl acetate, or ethylene glycol ‘like for like’ will not be an issue because of their reasonable environmental, health and safety profiles, other solvents will need to be phased out because of legislative measures long before feedstock security becomes an issue. This means that although the synthesis of NMP from the glutamic acid in plant protein wastes has been demonstrated, this will not save it from being tightly restricted [Lammens 2010]. Toluene is less obviously bio-based, but this is of little relevance given the toxicity of these chemicals will always remain the central issue.

To avoid the crisis point caused when a solvent is suddenly made unavailable, more bio-based solvents must be implemented. If needs be, they should deviate from the traditional solvent structures in order to possess more agreeable physical properties, without being toxic or otherwise a danger to humans or the environment. *p*-Cymene, as an alternative to toluene, has been suggested throughout this work as one such possibility. By phasing in new bio-based solvents as early (and gradually) as possible, less of a ‘culture shock’ will result. Once a greater variety of bio-based hydrocarbon solvents are established, and the quest for greener dipolar aprotics is realised, then the oxygenated bio-based solvents, ethyl acetate and ethanol to name but two, will satisfy demand for most other solvent applications. Reviewing the LSERs featured in this work makes the need for alternative solvents even more apparent (Figure 7.1). By annotating the aprotic bio-based solvent polarity map (Figure 1.20) with the conclusions of the solvent effect studies, it becomes obvious that only when the extremes of solvent polarity have been accounted for with bio-based solvents can we be comfortable that reaction performance will not become impaired without the use of unsustainable solvents. In what guise these solvents will present themselves is unknown, but to replicate the chlorinated solvents or the highly dipolar aprotic solvents with benign substitutes will require a great leap of imagination and an equal measure of support from the synthetic organic chemistry community. We are beginning to realise that a combination of theoretical tools and experiment can provide an answer to current worries

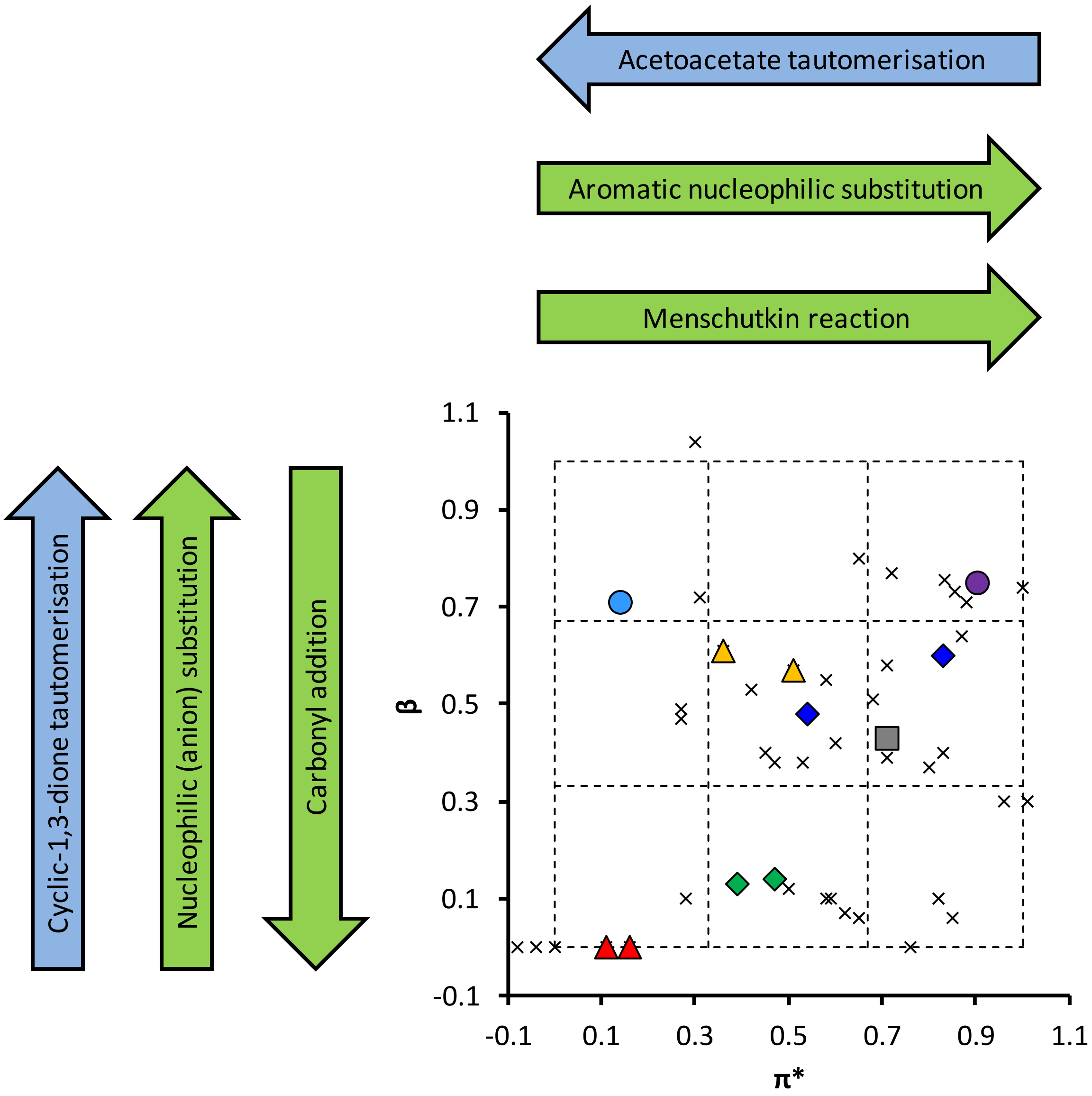


Figure 7.1 The polarity map of Figure 1.20 annotated with LSER trends (reaction rates in green, equilibria in blue).

over solvent safety, environmental impact, and renewability. Amongst the plethora of novel solvents that will inevitably appear in the near future, a few vital solvent substitutes will surely be found to ease concerns and advance the art of green solvent selection [Moity 2012].

8. Appendices

8.1 Experimental section

General notes: All reactions were conducted under an ambient atmosphere and generally without any purification of reagents prior to use. Hammett acidity functions and solvent polarity measurements were determined using a Jasco V-550 UV-vis. spectrophotometer. Thermal decomposition temperatures were obtained using a PL Thermal Sciences system (STA 625). Mass spectra by ESI detection were obtained with a Bruker MicrOTOF mass spectrometer, while EI mass spectra were obtained with a Perkin-Elmer Clarus 560 S mass spectrometer coupled to an Perkin-Elmer Clarus 500 gas chromatograph. Elemental analysis was determined using an Exeter Analytical CE 440 elemental analyser. All NMR spectra were obtained with a Bruker 400 MHz spectrometer, and calibrated against the residual solvent signal. No calibration of the NMR proton signals was found to be necessary. Characterisation of reaction products was consistent with literature data or authentic samples where available.

Determination of the Kamlet-Taft solvatochromic parameters: The determination of the **β** Kamlet-Taft solvatochromic parameter was performed in the same manner as originally described with 4-nitroaniline and *N*,*N*-diethyl-4-nitroaniline [Kamlet 1976]. Similarly values of **π*** were obtained by applying the absorbance maxima wavelengths of *N*,*N*-diethyl-4-nitroaniline to Equation 1.5 [Kamlet 1977]. When **α** values were needed spectroscopic data from Dimroth-Reichardt's betaine dye was interpreted with Equation 1.8. A Jasco V-550 UV-vis. spectrophotometer was used to obtain the required absorbance maxima wavelengths of each dye in solution. Aside from drying, generally no purification of the solvents was performed before the measurements. In the case of *p*-cymene a distillation was performed prior to analysis.

1-Decyl-2,3-dimethylimidazolium bromide: To a solution of 1,2-dimethylimidazole (0.288 g, 3.00 mmol) preheated to 323 K in the chosen solvent (3 mL) was added 1-bromodecane (0.736 g, 3.33 mmol) in a single aliquot. The progression of the reaction was monitored by ^{1}H-NMR spectroscopy, ideally until over 50% conversion had been achieved (Figure 2.2). No further analysis or isolation of the product was attempted.

***N*-Acetylpyrrolidine:** An excess of acetic anhydride (36.72 g, 0.36 mol) was added dropwise to pyrrolidine (21.30 g, 0.30 mol) cooled to 273 K. The reaction mixture was then allowed to warm to room temperature with stirring. To the solution was added solid sodium hydroxide (22 g), magnesium sulphate (22 g), and DCM (100 mL). The solution was retrieved by filtration and the reaction concentrated *in vacuo* to give *N*-acetylpyrrolidine (29.08 g, 86%). NMR: δ_H (400 MHz, $CDCl_3$) 1.79 (2H, m, CH_2CH_2N), 1.89 (2H, m, CH_2CH_2N), 1.97 (3H, s, CH_3CO), 3.36 (4H, m, CH_2N) /ppm; δ_C (100 MHz, $CDCl_3$) 22.5 (CH_3CO), 24.6 (CH_2CH_2N), 26.1 (CH_2CH_2N), 45.5 (CH_2N), 47.5 (CH_2N), 169.3 (CO) /ppm. MS (+EI): *m/z* 114 ($M+H^+$). Consistent with characterisation in the literature [Mucsi 2008].

***N*-Propionylpyrrolidine:** An excess of propionic anhydride (42.63 g, 0.36 mol) was added dropwise to pyrrolidine (21.30 g, 0.30 mol) at 273 K. The reaction was conducted and purified according to the previous procedure to give *N*-propionylpyrrolidine (33.0 g, 86%). NMR: δ_H (400 MHz, $CDCl_3$) 1.10 (3H, t, ^3J = 7.5 Hz, CH_3CH_2CO), 1.80 (2H, m, CH_2CH_2N), 1.91 (2H, m, CH_2CH_2N), 2.24 (2H, q, ^3J = 7.5 Hz, CH_3CH_2CO), 3.35 (2H, m, CH_2CH_2N), 3.41 (2H, m, CH_2CH_2N) /ppm; δ_C (100 MHz, $CDCl_3$) 9.0 (CH_3CH_2CO), 24.2 (CH_2CO), 26.0 (CH_2CH_2N), 27.8 (CH_2CH_2N), 45.4 (CH_2N), 46.4 (CH_2N), 172.2 (CO) /ppm. MS (+EI): *m/z* 128 ($M+H^+$).

***N*-Laurylpyrrolidine:** Lauryl chloride (6.527 g, 86 mmol) was added dropwise to a stirred solution of excess pyrrolidine (1.448 g, 0.26 mol) in 2-MeTHF (10 mL). Upon completion of the reaction solid sodium hydroxide (7 g) was added, filtered, and the solvent removed *in vacuo* to give *N*-laurylpyrrolidine. (16.7 g, 76%) NMR: δ_H (400 MHz, $CDCl_3$) 0.84 (3H, t, ^3J = 6.6 Hz, CH_3), 1.11-1.34 (16H, m, CH_2), 1.61 (2H, m, CH_2CH_2CO), 1.81 (2H, m, CH_2CH_2N), 1.91 (2H, m, CH_2CH_2N), 2.22 (2H, t, ^3J = 8.0 Hz, CH_2CO), 3.38 (2H, t, ^3J = 6.8 Hz, CH_2CH_2N), 3.43 (2H, t, ^3J = 6.8 Hz, CH_2CH_2N) /ppm; δ_C (100 MHz, $CDCl_3$) 14.13 (CH_3), 22.70 (CH_3CH_2), 24.44 (CH_2CH_2N), 25.00 (CH_2CH_2CO), 26.16 (CH_2CH_2N), 29.35 (CH_2), 29.50 (CH_2), 29.55 (CH_2), 29.57 (CH_2), 29.64 (CH_2), 29.66 (CH_2), 31.93 ($CH_2CH_2CH_2CO$), 34.88 (CH_2CO), 45.59 (CH_2N), 46.64 (CH_2N), 171.91 (CO) /ppm. MS (+EI) *m/z*: 254 [$M+H^+$].

***N*-Benzyl-4-phenylbutanamide kinetic experiments:** To a solution of 4-phenylbutanoic acid (0.328 g, 2.0 mmol) in the chosen solvent (4 mL) preheated to 373 K was added benzylamine (0.235 g, 2.2 mmol) in a single aliquot. The progression of the reaction was monitored by ^{1}H-NMR spectroscopy, ideally until over 50% conversion had been achieved (Figure 3.2). This process was conducted for binary solvent mixtures as well as single solvents. The precision executed during the experimental practice is listed here as follows: solvent volume by volumetric flask (± 0.05 mL), weighing of chemical reactants (± 0.001 g), reaction temperature calibrated by internal measurement of the liquid (rather than the heating apparatus) within ±1 K, sensitivity of NMR integration reported to 4 decimal places using *Spinworks* software. This means the concentration of the rate limiting benzylamine will reside within limits of 0.500 ±0.008 M. The kinetics of chemical reactions are temperature dependant, and a ±1 K (0.3% of absolute) accuracy regarding temperature control will introduce an error into the calculated Gibbs free energy of the reaction of 125.0 ±0.4 $kJmol^{-1}$ (also 0.3% error).

***N*-Benzyl-4-phenylbutanamide kinetic activation parameter determination:** The above procedure was repeated at various temperatures in different solvents and modelled with the Eyring equation (Equation 3.2) to provide experimental values of $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$.

***N*-Benzyl-4-phenylbutanamide reaction order determination:** The model amidation was repeated in toluene with different concentrations of both reactants ranging between 0.32 M and 1.83 M of benzylamine (0.5 M 4-phenylbutanoic acid) and between 0.45 M and 2.66 M of 4-phenylbutanoic acid (benzylamine concentration held and 0.55 M). The initial rate of reaction was determined by ^{1}H-NMR spectroscopy, taking multiple measurements during the first 10% of the total possible conversion to the product.

***N*-Benzyl-4-phenylbutanamide NMR signal calibration:** Several standards were prepared using a known amount of benzylamine and recrystallised *N*-benzyl-4-phenylbutanamide dissolved in toluene. For example, 30% conversion was approximated by dissolving 0.152 g (0.6 mmol) of *N*-benzyl-4-phenylbutanamide and 0.171 g of benzylamine (1.6 mmol) in 4 mL of toluene. A hypothetical 10 mol% excess of benzylamine was used to mirror the conditions used in kinetic experiments, and so 1.6 mmol of benzylamine is required to represent 30% conversion and not 1.4 mmol. 4-Phenylbutanoic acid was omitted from this exercise because it was not required to determine the conversion *in situ* during the reaction.

***N*-Benzyl-4-phenylbutanamide preparative experiments:** To a solution of 4-phenylbutanoic acid (0.493 g, 3.0 mmol) in *p*-cymene (4 mL), preheated to 373 K was added benzylamine (0.324 g, 3.0 mmol) in a single aliquot. The reaction was stirred for 24 hours, at which point the solution was allowed to cool to ambient temperature. The product crystallised from the solution when refrigerated. Recystallisation from aqueous acetone gave long white needle-like crystals of *N*-benzyl-4-phenylbutanamide (0.56 g, 74%). NMR: δ_H (400 MHz, $CDCl_3$) 1.62 (2H, m, $CH_2CH_2CH_2CONH$), 1.83 (2H, t, ^3J = 7.1 Hz, $CH_2CH_2CH_2CONH$), 2.29 (2H, t, ^3J = 7.1 Hz, $CH_2CH_2CH_2CONH$), 4.04 (2H, d, ^3J = 5.7 Hz, CH_2NHCO), 5.30 (1H, bs, NHCO), 6.69-7.03 (10H, m, aromatic protons) /ppm; δ_C (100 MHz, $CDCl_3$) 27.3 ($CH_2CH_2CH_2CONH$), 35.3 ($CH_2CH_2CH_2CONH$), 36.1 ($CH_2CH_2CH_2CONH$), 43.8 (CH_2NHCO), 126.1 ($C_{ar}H$), 127.7 ($C_{ar}H$), 128.0 ($C_{ar}H$), 128.5 ($C_{ar}H$), 128.6 ($C_{ar}H$), 128.88 ($C_{ar}H$), 138.4 ($C_{ar}CH_2N$), 141.6 ($C_{ar}CH_2CH_2CH_2CONH$), 172.7 (CONH) /ppm. MS (+ESI): *m/z* 254 ($M+H^+$). Elemental analysis calculated (%) for $C_{17}H_{19}NO$: C 80.60%, H 7.56%, N 5.53%; found C 80.32%, H 7.59%, N 5.51%.

Butyl butanoate kinetic experiments: To a solution of 1-butanol (0.373 g, 5.0 mmol) in the chosen solvent (5 mL), preheated to 323 K, was added butanoic anhydride (0.967 g, 5.5 mmol) in a single aliquot. The progression of the reaction was monitored by ^{1}H-NMR spectroscopy, ideally until over 50% conversion had been achieved (Figure 4.2). Processing of data into rate constants was performed in the same way as the amidation data was previously.

Synthesis of *p*-cymene from limonene: A mixture of 10 wt% Pd/C (0.788 g, 1 mol%) and K-10 montmorillonite clay (1.48 g) was heated to 413 K, and limonene (10.10 g, 74.1 mmol) slowly added dropwise. After stirring for one hour, the reaction was cooled to room temperature and water (100 mL) added prior to steam distillation. The organic phase of the distillate was dried to give *p*-cymene as a colourless liquid (8.87 g, 89% yield, 71% selectivity as determined by GC and ^{1}H-NMR spectroscopy). NMR: δ_H (400 MHz, $CDCl_3$) 1.22 (6H, d, ^3J = 7.0 Hz, $CHCH_3$), 2.31 (3H, s, $C_{ar}CH_3$), 2.87 (1H, m, $CHCH_3$), 7.11 (4H, m, $C_{ar}H$) /ppm; δ_C (100 MHz, $CDCl_3$) 21.1 ($C_{ar}CH_3$), 24.3 ($CHCH_3$), 33.9 ($CHCH_3$), 126.5 ($C_{ar}HC_{ar}CH$), 129.2 ($C_{ar}HC_{ar}CH_3$), 135.3 ($C_{ar}CH_3$), 146.1 ($C_{ar}CH$) /ppm. MS (+EI): *m/z* 134 (M^+).

Sulphonation of *p*-cymene: To *p*-cymene (4.300 g, 32.0 mmol) was slowly added 20% fuming sulphuric acid (5 mL). The reaction was stirred at room temperature for 4 hours. After this time had elapsed, stirring was stopped and water (6 mL) carefully added to avoid the mixture becoming hot. The diluted mixture was left to stand in a refrigerator overnight to produce a solid, which could be recrystallised from concentrated hydrochloric acid to give *p*-cymene-2-sulfonic acid dihydrate (7.29 g, 91%). Melting point: T_m 323-324 K. NMR: δ_H (400 MHz, DMSO-d_6) 1.14 (6H, d, ^3J = 6.9 Hz, $CHCH_3$), 2.46 (3H, s, $C_{ar}CH_3$), 2.83 (1H, m, $CHCH_3$), 7.05 (1H, d, ^3J = 7.8 Hz, $C_{ar}HC_{ar}CH_3$), 7.16 (1H, d, ^3J = 7.8 Hz, ^4J = 1.7 Hz, $CHC_{ar}C_{ar}HC_{ar}H$), 7.70 (1H, d, ^4J = 1.7 Hz, $C_{ar}HC_{ar}SO_3H$), 9.56 (1H, bs, $C_{ar}SO_3H$) /ppm; δ_C (100 MHz in DMSO-d_6): 19.9 ($C_{ar}CH_3$), 24.3 ($CHCH_3$), 33.2 ($CHCH_3$), 124.7 ($C_{ar}HC_{ar}CH_3$), 127.4 ($CHC_{ar}C_{ar}HC_{ar}H$), 131.2 ($C_{ar}SO_3H$), 133.1 ($C_{ar}HC_{ar}SO_3H$), 145.1 ($C_{ar}CH_3$), 145.3 ($C_{ar}CH$) /ppm. MS (-EI): *m/z* 213 (M-H^-).

Complete synthesis of *p*-CSA from the essential oil of oranges: The flavedo (outer peel) was separated from the albedo (inner peel) of sixteen oranges (Navel late variety, diameter of 70-80 mm). Using a fine grater gave 110 g of wet citrus waste. To the separated flavedo was added 150 mL of water and the suspension distilled for 1 hour. The organic phase of the distillate was dried to give 5.71 g of the essential oil (5.2 wt%). The steam extracted citrus oil was added dropwise to a mixture of 10 wt% Pd/C (0.434 g, 1 mol%) and K-10 montmorillonite clay (0.815 g), pre-heated to 413 K, and stirred for an hour once addition of the citrus oil was complete. The reaction was then allowed to cool and 100 mL of water added. The solution was distilled, and the organic phase of the distillate dried to give crude *p*-cymene (4.43 g, 88% yield, 70% selectivity). To this colourless liquid was carefully added 3.5 mL of 20% fuming sulphuric acid. The reaction was stirred at room temperature for 4 hours. After this time had elapsed, stirring was stopped and water (4.2 mL) carefully added to avoid the mixture becoming hot. At this stage the *p*-menthane co-product caused by limonene disproportionation can be decanted. The diluted mixture was left to stand in a refrigerator overnight to solidify. The solid was retrieved by filtration to give *p*-cymene-2-sulfonic acid dihydrate (1.55 g, 27% yield based on the *p*-cymene content of the reaction distillate, 16% total yield based on the limonene content of the citrus oil, and 27 wt% based on the mass of citrus oil).

Benzyl acetate kinetic experiments: To a solution of benzyl alcohol (0.541 g, 5.0 mmol) and the acid catalyst (*p*-TSA or *p*-CSA, 0.05 mmol) in the chosen solvent (5 mL), stirred at 323 K, was added acetic acid (0.330 g, 5.5 mmol). Aliquots of the reaction mixture were removed at convenient intervals and diluted with deuterated chloroform to allow the reaction progress to be monitored by ^{1}H-NMR spectroscopy by the same method described by Welton and co-workers [Wells 2008]. Reactions were typically allowed to proceed beyond 50% conversion to guarantee accuracy in the calculation of rate constants (Figure 5.9).

Ethyl levulinate: To a solution of levulinic acid (0.581 g, 5.0 mmol) in the chosen solvent (5 mL) was added either a Brønsted acid (*p*-TSA or *p*-CSA, 0.05 mmol) or a Lewis acid ($In(OTf)_3$, $InCl_3$, or $FeCl_3$, 0.25 mmol), followed by the addition of ethanol (0.230 g, 5.0 mmol). The reaction mixture was stirred at 323 K for 20 hours. After this time, the reaction was cooled and potassium carbonate added. Filtration of the reactions performed in either toluene or 2-MeTHF gave a filtrate that could be concentrated *in vacuo* to give the desired product, ethyl levulinate, in yields of up to 73% of the theoretical yield depending on the conditions used. Reactions conducted in *p*-cymene were similarly filtered and purified by column chromatography (hexane:ethyl acetate) to give ethyl levulinate (up to 76% yield depending on the conditions used). NMR: δ_H (400 MHz, $CDCl_3$) 1.24 (3H, t, ^3J = 7.1 Hz, CH_3CH_2O), 2.18 (3H, s, CH_3CO), 2.56 (2H, t, ^3J = 6.6 Hz, CH_2CO_2Et), 2.74 (2H, t, ^3J = 6.6 Hz, $COCH_2$), 4.12 (2H, q, ^3J = 7.1 Hz, CH_3CH_2O) /ppm; δ_C (100 MHz, $CDCl_3$) 14.23 (CH_2CH_3), 28.11 ($CH_3COCH_2CH_2$), 29.90 (CH_3CO), 38.04 (CH_3COCH_2), 60.64 (CH_2CH_3), 172.88 ($COCH_2CH_3$), 206.81 (CH_3CO) /ppm. MS (+EI): *m*/*z* 144 (M^+).

4-Bromochalcone: A mixture of acetophenone (0.258 g, 2.03 mmol), 4-bromobenzaldehyde (0.562 g, 3.04 mmol, and the acid catalyst (*p*-TSA or *p*-CSA, 0.10 mmol) was stirred at 393 K for 24 hours. The reaction was then allowed to cool, giving rise to fine needles of a white solid and an amorphous orange solid. The latter could be removed by recrystallisation with ethanol to give 4-bromochalcone (0.44 g, 73%). NMR: δ_H (400 MHz, $CDCl_3$) 7.44-7.61 (8H, m, $C_{ar}H$, CH=CHCO), 7.72 (1H, d, ^3J = 15.7 Hz, CH=CHCO), 8.00 (2H, m, $C_{ar}HC_{ar}CO$) /ppm; δ_C (100 MHz, $CDCl_3$) 122.7 (CH=CHCO), 124.9 ($C_{ar}Br$), 128.6 ($C_{ar}HC_{ar}CO$), 128.8 ($C_{ar}HC_{ar}HC_{ar}CO$), 130.0 ($C_{ar}HC_{ar}HC_{ar}Br$), 132.4 ($C_{ar}HC_{ar}Br$), 133.1 ($C_{ar}HC_{ar}HC_{ar}HC_{ar}CO$), 133.9 ($C_{ar}C$=CH), 138.1 ($C_{ar}CO$), 143.5 (CH=CHCO), 190.5

(CO) /ppm. MS (+EI): *m/z* 287, 289 ($M+H^+$).

2-(4-Nitrophenyl)-1,3-dioxolane: A suspension of 4-nitrobenzaldehyde (1.511 g, 10.0 mmol) and the acid catalyst (*p*-TSA or *p*-CSA, 0.20 mmol) in cyclohexane was heated to reflux in Dean-Stark apparatus, at which point the mixture became homogeneous. Ethylene glycol (0.621 g, 10.0 mmol) was then added and the reaction stirred at reflux for 5 hours. The reaction was the left to cool, allowing the product, 2-(4-nitrophenyl)-1,3-dioxolane, to precipitate and be isolated by filtration (1.92 g, 92%). NMR: δ_H (400 MHz, DMSO-d_6) 4.03 (4H, m, OCH_2), 5.89 (1H, s, CH), 7.71 (2H, d, ^3J = 8.7 Hz, $C_{ar}\mathit{H}C_{ar}HC_{ar}NO_2$), 8.26 (2H, d, ^3J = 8.7 Hz, $C_{ar}HC_{ar}NO_2$) /ppm; δ_C (100 MHz, DMSO-d_6) 65.10 (OCH_2), 101.44 (CH), 123.61 ($\mathit{C}_{ar}HC_{ar}NO_2$), 127.93 ($\mathit{C}_{ar}HC_{ar}HC_{ar}NO_2$), 145.14 ($C_{ar}$), 147.99 ($C_{ar}NO_2$) /ppm. MS (+EI): *m/z* 136 ($M+H^+$).

Synthesis of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate with HCl: Urea (0.300 g, 5.00 mmol) dissolved or suspended in the chosen solvent (4 mL) was heated to 348 K. Upon reaching thermal equilibrium benzaldehyde (0.532 g, 5.00 mmol), methyl acetoacetate (0.872 g, 7.50 mmol), and finally concentrated hydrochloric acid (10 mol%) were added to the mixture. The reaction was stirred for a duration of 3 hours unless otherwise stated in Chapter 6. Upon completion of the reaction, the mixture was allowed to cool to ambient temperature. If acetic acid or DMF were used as the solvent, water was added to effect precipitation of the product. The resultant solid was separated from the reaction mixture by filtration, washed with 50% aqueous ethanol, and recrystallised from ethanol to give methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate as a white crystalline solid. NMR: δ_H (400 MHz, DMSO-d_6) 2.25 (3H, s, $C\mathit{H}_3CNH$), 3.53 (3H, s, CO_2CH_3), 5.15 (1 H, d, ^3J(N,H) = 3.4 Hz, C*H*NH), 7.35-7.20 (5H, m, $C_{ar}H$), 7.76 (1H, bs, CHN*H*), 9.22 (1H, s, $CH_3CN\mathit{H}$) /ppm; δ_C (100 MHz, DMSO-d_6) 17.9 ($\mathit{C}H_3CNH$), 50.4 (CHNH), 54.6 ($CO_2\mathit{C}H_3$), 99.9 ($\mathit{C}CO_2$), 126.0 ($C_{ar}H$), 127.1 ($C_{ar}H$), 128.0 ($C_{ar}H$), 143.8 (C_{ar}), 147.2 ($CH_3\mathit{C}NH$), 153.2 (NH*C*ONH), 165.8 ($\mathit{C}O_2CH_3$) /ppm. MS (+ESI): 247 ($M+H^+$). Elemental analysis calculated (%) for $C_{13}H_{14}N_2O_3$: C 63.40%, H 5.73%, N 11.38%; found C 63.44%, H 5.77%, N 11.44%. Consistent with characterisation in the literature [Gangadasu 2006].

Synthesis of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate with EPZ-10: Typically EPZ-10 (0.57g, equivalent to 10 mol% zinc chloride) was pre-heated to 398 K for 3 hours prior to the reaction. In some experiments this was modified, as represented in Figure 6.12. After this time urea (0.300 g, 5.00 mmol) and the chosen solvent (4 mL) were added to the catalyst and heated at 348 K. Upon reaching thermal equilibrium benzaldehyde (0.532 g, 5.00 mmol), and methyl acetoacetate (0.872 g, 7.50 mmol) were added to the mixture. The reaction was stirred for a duration of 3 hours, and then allowed to cool to ambient temperature. The mixture was filtered and the solid washed with acetic acid. To the combined organic phase was added water to induce precipitation of the product. The resultant solid was separated by filtration and recrystallised from ethanol to give a white crystalline solid.

Synthesis of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate with other Lewis acids: Urea (0.300 g, 5.00 mmol), the catalyst (10 mol%) and the chosen solvent (4 mL) were heated to 348 K. Upon reaching thermal equilibrium benzaldehyde (0.532 g, 5.00 mmol) and methyl acetoacetate (0.872 g, 7.50 mmol) were added. The reaction was stirred for a duration of 3 hours. Upon completion of the reaction the mixture was allowed to cool to ambient temperature, and then water added to effect dissolution of the product. The resultant solid was separated from the reaction mixture by filtration and recrystallised from ethanol to give a white crystalline solid as previously obtained.

Synthesis of 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione with HCl: Urea (0.300 g, 5.00 mmol), 5,5-dimethyl-1,3-cyclohexanedione (1.05 g, 7.50 mmol), and the chosen solvent (12 mL) were heated to 348 K. Upon reaching thermal equilibrium benzaldehyde (0.532, 5.00 mmol) and concentrated hydrochloric acid (10 mol%) were added to the mixture. The reaction was stirred for a duration of 24 hours. Upon completion of the reaction, the mixture was allowed to cool to ambient temperature. Water was then added to ensure complete dissolution of the product. The resultant solid was separated from the reaction mixture by filtration, washed with 50% aqueous ethanol and recrystallised from ethanol to give 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione as a white needle-like crystals. NMR: δ_H (400 MHz, DMSO-d$_6$) 0.88 (3H, s, CH_3) 1.00 (3H, s, CH_3), 2.10 (2H, m, CH_2), 2.34 (2H, m, CH_2CO), 5.14 (1H, d, ^3J(H,N) = 2.8 Hz, CH), 7.34-7.18 (5H, m, $C_{ar}H$), 7.77 (1H, bs, CHN*H*), 9.47 (1H, s, CH_2CN*H*) /ppm; δ_C (100 MHz, DMSO-d$_6$) 26.9 (CH_3), 28.8 (CH_3), 32.3 (*C*CH_2CO), 40.3 (*C*H_2CNH), 49.8 (*C*H_2CO), 52.0 (CH), 107.4 (CO*C*=C), 126.3 ($C_{ar}H$), 127.2 ($C_{ar}H$), 128.4 ($C_{ar}H$), 144.7 (C_{ar}), 152.0 (COC=*C*), 152.5 (NHCONH), 192.2 (CH_2*C*O) /ppm. MS (+ESI): 271 (M+H$^+$). Elemental analysis calculated (%) for $C_{16}H_{19}N_2O_2$: C 71.09%, H 6.71%, N 10.36%; found C 71.07%, H 6.69%, N 10.27%. Consistent with characterisation in the literature [Yarım 2003].

8.2 Supplementary data

Solvent properties: The non-intrusive measurements of solvent polarity, namely relative permittivity, the Hildebrand solubility parameter, and the Hansen solubility parameters are tabulated here (Table 8.1) [Abboud 1999, Reichardt 2003 page 472, Hansen 2007 page 347]. The Kamlet-Taft polarity parameters for relevant solvent are also reported (Table 8.2). Instances when literature values have not been calculated using the single dye set consisting of 4-

nitroaniline, *N*,*N*-diethyl-4-nitroaniline, and Dimroth-Reichart's betaine dye have been noted. This is most useful for haloalkanes that otherwise express **β** values of zero, and an average of dye sets is used to correct this.

Table 8.1 Bulk solvent polarity parameters.

Solvent	ε_r	δ_T /MPa$^{1/2}$	δ_D /MPa$^{1/2}$	δ_P /MPa$^{1/2}$	δ_H /MPa$^{1/2}$
Acetic acid	6.17[a]	21.38	14.5	8.0	13.5
Acetone	20.56	19.73	15.5	10.4	7.0
Acetonitrile	35.94	24.2	15.3	18.0	6.1
N-Acetyl pyrrolidine	n/a	17.00[b]	18.3[b]	9.1[b]	7.4[b]
Benzene	2.27	18.73	18.4	0.0	2.0
Butanoic acid	22.59[b]	18.74[c]	14.9	4.1	10.6
1-Butanol	17.51	23.35	16.0	5.7	15.8
t-Butanol	12.47	21.75	15.2	5.1	14.7
Butanone	18.11	19.0	16.0	9.0	5.1
n-Butyl acetate	5.01[a]	17.41[c]	15.8	3.7	6.3
Butylamine	5.4	18.62[c]	16.2	4.5	8.0
t-Butyl methyl ether	4.5[a]	15.07	14.8	4.3	5.0
Carbon tetrachloride	2.24	18.11[c]	16.1	8.3	0.0
Chlorobenzene	5.62	19.4	19.0	4.3	2.0
Chloroform	4.89	18.9	17.8	3.1	5.7
Cineole	4.84[a]	17.65[c]	16.7	4.6	3.4
Cyclohexane	2.02[a]	16.76	16.8	0.0	0.2
Cyclohexanone	15.50	19.56[c]	17.8	6.3	5.1
Cyclopentyl methyl ether	4.47[b]	18.82[b]	14.4[b]	3.2[b]	4.3[b]
p-Cymene	4.68[b]	17.4[d]	18.5[b]	1.8[b]	2.3[b]
Cumene	2.38[a]	18.18[c]	18.1	1.2	1.2
1,2-Dichlorobenzene	9.93	20.47[c]	19.2	6.3	3.3
1,2-DCE	10.36	20.26	19.0	7.4	4.1
DCM	8.93	20.37	18.2	6.3	6.1
Diethyl carbonate	2.82[a]	16.73[c]	15.1	6.3	3.5
Diethyl ether	4.20	15.42	14.5	2.9	5.1
Diisopropyl ether	4.04[a]	14.43[c]	13.7	3.9	2.3
DMAc	37.78	22.35	16.8	11.5	10.2
Dimethyl carbonate	3.17[a]	18.70[c]	15.5	3.9	9.7
DMF	36.71	23.96	17.4	13.7	11.3
DMSO	46.45	26.45	18.4	16.4	10.2
1,4-Dioxane	2.21	20.48	19.0	1.8	7.4
Ethanol	24.55	26.43	15.8	8.8	19.4
Ethyl acetate	6.02	18.35	15.8	5.3	7.2

Table 8.1 Bulk solvent polarity parameters (continued).

Solvent	ε_r	δ_T /MPa^{½}	δ_D /MPa^{½}	δ_P /MPa^{½}	δ_H /MPa^{½}
Ethyl lactate	16.54[b]	21.68[c]	16.0	7.6	12.5
Ethylene glycol	37.70	33.34	16.9	11.1	26.0
Fluorobenzene	5.55[a]	19.77[c]	18.7	6.1	2.0
Glycerol	42.5	36.16[c]	17.4	12.1	29.3
Glycerol formal	n/a	28.79[b]	19.9[b]	12.0[b]	16.6[b]
n-Heptane	1.92[a]	15.20	15.3	0.0	0.0
Hexafluoroisopropanol	n/a	23.07[c]	17.2	4.5	14.7
n-Hexane	1.88	14.90[c]	14.9	0.0	0.0
Isoamyl alcohol	15.19	21.30[c]	15.8	5.2	13.3
Lactic acid	39.52[b]	34.12[c]	17.0	8.3	28.4
Limonene	2.3	15.1[e]	17.2[b]	1.8[b]	4.3[b]
Methanol	32.66	29.59	15.1	12.3	22.3
Methyl acetate	6.68	18.70[c]	15.5	7.2	7.6
2-MeTHF	6.97	18.14[c]	16.9	5.0	4.3
NMP	32.2	23.16	18.0	12.3	7.2
Nitrobenzene	34.79	22.15[c]	20.0	8.6	4.1
Nitromethane	35.87	25.08[c]	15.8	18.8	5.1
α-Pinene	2.7	17.28[c]	16.9	1.8	3.1
Piperidine	9.74[b]	20.23[c]	17.6	4.5	8.9
1,2-Propanediol	29.35[b]	30.22[c]	16.8	9.4	23.3
1,3-Propanediol	33.68[b]	31.67[c]	16.8	13.5	23.2
Propanenitrile	28.26	21.65[c]	15.3	14.3	5.5
Propanoic acid	25.91[b]	19.95[c]	14.7	5.3	12.4
1-Propanol	20.45	24.60[c]	16.0	6.8	17.4
2-Propanol	19.92	23.58[c]	15.8	6.1	16.4
Propylene carbonate	64.92	27.22[c]	20.0	18.0	4.1
Pyridine	12.91	21.75[c]	19.0	8.8	5.9
Sulpholane	43.3	29.36	20.3	18.2	10.9
THF	7.58	20.23	16.8	5.7	8.0
Toluene	2.38	18.2	18.0	1.4	2.0
Tributyl phosphate	8.29[a]	18.00[c]	16.3	6.3	4.3
Triethylamine	2.42[a]	15.20	17.8	0.4	1.0
Triethyl phosphate	13.01[a]	22.21[c]	16.7	11.4	9.2
Trifluoroethanol	n/a	23.98[c]	15.4	8.3	16.4

Table 8.1 Bulk solvent polarity parameters (continued).

Solvent	ε_r	δ_T /MPa$^{1/2}$	δ_D /MPa$^{1/2}$	δ_P /MPa$^{1/2}$	δ_H /MPa$^{1/2}$
γ-Valerolactone	7.08[b]	20.02[b]	15.2[b]	10.1[b]	5.6[b]
Water	78.36	47.81	15.5	16.0	42.3
p-Xylene	2.27[a]	17.90[c]	17.6	1.0	3.1

[a] Recorded at 293 K rather than 298 K.

[b] Predicted values using ProPred property estimation software.

[c] Calculated with Equation 1.2.

[d] Reference: Sagadeev 2006.

[e] Reference: Hazra 2002.

Table 8.2 Values for the Kamlet-Taft solvent polarity scale.

Solvent	α	β	π^*	Ref.
Acetic acid	0.71	0.40	0.60	Clark 2013, Taft 1976
Acetone	0.00	0.51	0.70	Kamlet 1976
Acetonitrile	0.35	0.37	0.80	Crowhurst 2003
N-Acetyl pyrrolidine	0.00	0.76	0.83	This work
Benzene[a]	0.00	0.10	0.59	Marcus 1993
Butanoic acid[a]	1.10	0.45	0.56	Marcus 1993
1-Butanol	0.73	0.85	0.61	Kamlet 1976, Taft 1976
t-Butanol	0.39	0.95	0.58	Kamlet 1976, Taft 1976
Butanone	0.00	0.51	0.68	Kamlet 1976, Reichardt 1994
n-Butyl acetate[a]	0.00	0.45	0.46	Marcus 1993
Butylamine[a]	0.00	0.72	0.31	Marcus 1993
Carbon tetrachloride[a]	0.00	0.10	0.28	Marcus 1993
Chlorobenzene	0.00	0.06	0.65	Clark 2012
Chloroform[a]	0.20	0.10	0.58	Marcus 1993
Cineole	0.00	0.61	0.36	Jessop 2012
Cyclohexane	0.00	0.00	0.00	This work
Cyclohexanone	0.00	0.58	0.71	Clark 2012
Cyclopentyl methyl ether	0.00	0.49	0.41	This work
p-Cymene	0.00	0.13	0.39	Clark 2012
Cumene	0.00	0.11	0.43	This work
1,2-Dichlorobenzene[a]	0.00	0.03	0.80	Marcus 1993
1,2-DCE	0.00	0.00	0.76	This work

Table 8.2 Values for the Kamlet-Taft solvent polarity scale (continued).

Solvent	**α**	**β**	**π***	Ref.
DCM[a]	0.13	0.10	0.82	Marcus 1993
Diethyl carbonate[a]	0.00	0.40	0.45	Marcus 1993
Diethyl ether	0.00	0.51	0.28	Kamlet 1976
Diisopropyl ether[a]	0.00	0.49	0.27	Marcus 1993
DMAc	0.00	0.73	0.85	Kamlet 1976, Reichardt 1994
Dimethyl carbonate	0.00	0.32	0.55	This work
DMF	0.00	0.71	0.88	Kamlet 1976, Reichardt 1994
DMSO	0.00	0.74	1.00	Kamlet 1976, Reichardt 1994
1,4-Dioxane	0.00	0.38	0.52	Kamlet 1976, Reichardt 1994
Ethanol	0.83	0.77	0.62	Kamlet 1976, Taft 1976
Ethyl acetate	0.00	0.48	0.54	Kamlet 1976, Reichardt 1994
Ethyl lactate	0.69	0.52	0.82	Jessop 2012
Ethylene glycol	0.79	0.57	1.01	Clark 2013, Taft 1976
Fluorobenzene[a]	0.00	0.07	0.62	Marcus 1993
Glycerol	0.93	0.67	1.04	Jessop 2012
Glycerol formal	0.59	0.59	0.87	Jessop 2012
n-Heptane	0.00	0.00	-0.03	This work
Hexafluoroisopropanol[a]	1.96	0.00	0.65	Marcus 1993
n-Hexane	0.00	0.00	-0.05	This work
Isoamyl alcohol[a]	0.84	0.86	0.40	Marcus 1993
Lactic acid	n/a	0.40	1.09	This work
Limonene	0.00	0.00	0.16	Clark 2012
Methanol	1.00	0.65	0.69	This work
Methyl acetate[a]	0.00	0.42	0.60	Marcus 1993
2-MeTHF	0.00	0.57	0.51	This work
NMP	0.00	0.75	0.90	This work
Nitrobenzene[a]	0.00	0.30	1.01	Marcus 1993
Nitromethane[a]	0.22	0.06	0.85	Marcus 1993
α-Pinene	0.00	0.00	0.11	Jessop 2012
Piperidine[a]	0.00	1.04	0.30	Marcus 1993
1,2-Propanediol	0.83	0.78	0.76	Jessop 2012
1,3-Propanediol	0.80	0.77	0.84	Jessop 2012
Propanenitrile	0.00	0.39	0.72	This work
Propanoic acid[a]	1.12	0.28	0.51	This work, Marcus 1993

Table 8.2 Values for the Kamlet-Taft solvent polarity scale (continued).

Solvent	α	β	π*	Ref.
1-Propanol[a]	0.84	0.90	0.52	Marcus 1993
2-Propanol	0.66	0.92	0.61	Kamlet 1976, Taft 1976
Propylene carbonate[a]	0.00	0.40	0.83	Marcus 1993
Pyridine	0.00	0.67	0.85	Kamlet 1976
Sulpholane	0.00	0.30	0.96	This work
THF	0.00	0.55	0.58	Kamlet 1976, Reichardt 1994
Toluene	0.00	0.12	0.50	Clark 2012
Tributyl phosphate[a]	0.00	0.80	0.65	Marcus 1993
Triethylamine	0.00	0.70	0.08	Kamlet 1976
Triethyl phosphate	0.00	0.79	0.71	Kamlet 1976
Trifluoroethanol[a]	1.51	0.00	0.73	Marcus 1993
γ-Valerolactone	0.00	0.60	0.83	Jessop 2012
Water	1.05	0.18	1.28	This work, Taft 1976
p-Xylene	0.00	0.14	0.47	Clark 2012

[a] Polarity measurements derived from the average value of different dyes.

Solvent greenness: The SUS-HAS-ECO classifications for all the solvents in the original GSK solvent selection guide have been calculated (Table 8.3) [Henderson 2011]. The SUS classifications were assigned in the same way as indicated in the main text (Table 1.5).

Table 8.3 The full SUS-HAS-ECO classifications of solvent greenness.

Solvent	SUS	HAS	ECO
Acetic acid	8	7	10
Acetic anhydride	4	3	6
Acetone	8	9	9
Acetonitrile	4	7	1
t-Amyl methyl ether	0	4	10
Anisole	0	7	3
Benzene	0	1	8
Benzyl alcohol	0	9	6
Bis(2-methoxyethyl) ether	4	1	3
1,4-Butanediol	6	10	1
1-Butanol	8	6	3

Table 8.3 The full SUS-HAS-ECO classifications of solvent greenness (continued).

Solvent	SUS	HAS	ECO
2-Butanol	4	10	4
t-Butanol	4	7	10
Butanone	4	9	1
n-Butyl acetate	4	10	4
t-Butyl acetate	4	10	10
t-Butyl ethyl ether	4	2	10
t-Butyl methyl ether	4	4	9
Carbon disulphide	4	1	10
Carbon tetrachloride	2	1	6
Chloroacetic acid	2	8	7
Chlorobenzene	0	4	10
Chloroform	2	1	3
Cumene	0	6	9
Cyclohexane	0	7	7
Cyclohexanol	0	10	10
Cyclohexanone	0	7	7
Cyclopentanone	0	8	6
Cyclopentyl methyl ether	0	2	1
cis-Decalin	0	7	7
Di(ethylene glycol)	4	10	10
Di(ethylene glycol) monobutyl ether	4	8	9
Dibutyl ether	4	1	1
1,2-Dichlorobenzene	0	8	10
DCM	2	3	6
1,2-DCE	2	1	6
Diethyl ether	4	1	3
Diisopropyl ether	4	4	10
1,2-Dimethoxyethane	4	1	6
DMAc	4	1	1
N,*N*-Dimethyl aniline	0	4	1
Dimethyl carbonate	4	9	10
Dimethyl ether	8	4	6
DMF	4	1	7
Dimethylpropylene urea	4	4	1

Table 8.3 The full SUS-HAS-ECO classifications of solvent greenness (continued).

Solvent	SUS	HAS	ECO
DMSO	4	7	4
1,4-Dioxane	4	1	1
Diphenyl ether[a]	0	3	8
Ethanol	10	10	10
Ethoxybenzene[a]	0	10	6
Ethyl acetate	8	9	6
Ethyl formate[a]	4	4	1
2-Ethyl hexanol	0	8	7
Ethyl lactate[a]	8	4	6
Ethyl propionate[a]	4	1	6
Ethylene carbonate[a]	4	7	8
Ethylene glycol	8	10	10
Fluorobenzene	0	6	1
Formamide	4	1	10
Glycerol	10	10	10
Heptane	0	9	6
Hexane	0	1	6
Isoamyl alcohol	6	10	6
Isooctane	0	9	7
ISOPAR G (C_{10-12} isoalkanes)[a]	0	10	1
Isopropyl acetate	4	8	9
Mesitylene	0	9	8
Methanol	8	5	10
2-Methoxyethanol	4	1	8
Methyl acetate	4	7	9
Methyl lactate	4	4	4
Methylcyclohexane	0	9	8
Methylcyclopentane	0	3	7
N-Methylformamide	4	1	7
Methylisobutyl ketone	4	7	1
2-Methylpentane	0	7	6
NMP	6	2	1
2-MeTHF	10	1	1
Nitromethane[a]	0	1	3

Table 8.3 The full SUS-HAS-ECO classifications of solvent greenness (continued).

Solvent	SUS	HAS	ECO
n-Octyl acetate	0	6	7
n-Pentane	0	9	8
2-Pentanol	0	7	6
2-Pentanone	0	7	1
3-Pentanone	0	9	1
Perfluorocyclic ether[a]	0	2	1
Perfluorocyclohexane[a]	0	1	1
Perfluorohexane[a]	0	1	1
Perfluorotoluene[a]	0	3	1
Petroleum spirit	0	1	6
1,2-Propanediol	6	10	1
1,3-Propanediol	6	10	1
Propanenitrile[a]	4	3	1
Propanoic acid	6	7	9
1-Propanol	6	6	8
2-Propanol	6	10	1
Propyl acetate	4	10	1
Propylene carbonate[a]	4	6	8
Pyridine	0	4	1
Sulpholane[a]	4	9	10
Tetrahydrofuran	4	3	1
Toluene	0	3	6
Tri(ethylene glycol)	4	8	10
Trichloroacetic acid	2	7	6
Trichloroacetonitrile[a]	2	7	3
1,2,4-Trichlorobenzene	0	5	10
Triethylamine	6	1	6
Trifluoroacetic acid[a]	2	7	1
2,2,2-Trifluoroethanol	2	1	6
Trifluorotoluene[a]	0	1	1
Water	10	10	10
p-Xylene	6	7	6

[a] No LCA data and so the ECO classification is based only on the waste and environmental impact categories of the GSK solvent selection guide.

Linear solvation energy relationship calculations from Chapter 2: Coefficients for the construction of all LSERs are deemed to be statisically significant if their associated p-value is below 0.01. The following tables document the initial Kamlet-Taft solvent polarity parameter screening (Table 8.4), and final relationships (Table 8.5). Also included are LSER calculations that include the $\boldsymbol{\varepsilon}$ term (Table 8.6) and redundant relationships featuring $\boldsymbol{\varepsilon_r}$ or $\mathbf{E_T^N}$ (Table 8.7).

Table 8.4 The statistical significance of LSER coefficients for the kinetics of the Menschutkin reaction in the form ‘$\ln(\mathbf{k}) = XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^*$’.

n	Coefficient (p-value)				R^2
	XYZ_0	a	b	s	
7 (no alcohols)	-13.64	0.46	0.51	4.32	0.988
	($2.5*10^{-5}$)	(0.60)	(0.55)	(0.014)	
7 (no C-H acids)	-13.79	-2.14	-0.04	4.94	0.995
	($3.1*10^{-5}$)	($4.7*10^{-3}$)	(0.93)	($1.7*10^{-3}$)	
9 (original solvent set)	-13.74	-1.65	-0.98	5.67	0.980
	($5.1*10^{-7}$)	($4.3*10^{-3}$)	(0.033)	($1.8*10^{-4}$)	

Table 8.5 The final LSER coefficients for the kinetics of the Menschutkin reaction in the form ‘$\ln(\mathbf{k}) = XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^*$’ using only statistically significant parameters.

n	Coefficient (p-value)				R^2
	XYZ_0	a	b	s	
7 (no alcohols)	-13.75	0.00	0.00	4.85	0.986
	($1.4*10^{-8}$)			($7.4*10^{-6}$)	
7 (no C-H acids)	-13.81	-2.16	0.00	4.92	0.994
	($9.2*10^{-7}$)	($3.4*10^{-4}$)		($1.1*10^{-4}$)	
9 (original solvent set)	-13.69	-1.99	0.00	4.90	0.948
	($6.4*10^{-7}$)	($5.4*10^{-3}$)		($7.0*10^{-4}$)	

Table 8.6 The final LSER coefficients for the kinetics of the Menschutkin reaction in the form '$\ln(\mathbf{k}) = XYZ_0 + a(\boldsymbol{\alpha} + e\boldsymbol{\varepsilon}) + s\boldsymbol{\pi}^*$' using only statistically significant parameters.

n	Coefficient (p-value)				R^2
	XYZ_0	a	e	s	
9 (original solvent set)	-13.79	-2.17	-1.02	4.90	0.995
	($1.8*10^{-8}$)	($4.2*10^{-5}$)	($1.0*10^{-3}$)	($8.0*10^{-6}$)	
13 (all solvents)	-13.79	-2.18	-1.03	5.12	0.991
	($3*10^{-13}$)	($4.2*10^{-7}$)	($2.1*10^{-4}$)	($1.1*10^{-8}$)	

Table 8.7 The LSER coefficients for the kinetics of the Menschutkin reaction in the form '$\ln(\mathbf{k}) = XYZ_0 + f\mathbf{E}_\mathbf{T}^\mathbf{N}$' or '$\ln(\mathbf{k}) = XYZ_0 + g\boldsymbol{\varepsilon}_\mathbf{r}$'.

n	Coefficient (p-value)			R^2
	XYZ_0	f	g	
9 (original solvent set)	-8.72	-4.39	0.00	0.220
	($2.4*10^{-4}$)	(0.20)		
9 (original solvent set)	-10.05	0.00	-0.014	0.031
	($9.6*10^{-6}$)		(0.65)	

Amidation reaction monitoring from Chapter 3: Standard solutions were prepared using the recrystallised carboxamide product, and combining with benzylamine in toluene to re-create the concentration of both components at the desired conversion. 4-Phenylbutanoic acid was omitted from this exercise because it is not required to determine the conversion *in situ* during the reaction. At room temperature these standard solutions could be prepared up to about a replication of 50% conversion. It was at this point that the saturation point of the amide was reached. Reactions were typically carried out until 50% conversion was achieved, sometimes further in the more accelerated reactions, but 50% conversion was deemed to provide enough data variation for reliable results [Moore 1981 page 37]. As such it seemed convenient to limit the calibration at this point, rather than change the concentration of the standards and risk introducing an error. The correlation suggests quantitative proportionality between yields and conversions estimated from the signal intensities of ^{1}H-NMR spectra (Figure 8.1). This considered, no actual calibration was applied during kinetic experiments. This gave the methodology of kinetic analysis by ^{1}H-NMR spectroscopy a firmer grounding than just the assumption that conversion is accurately accounted for. Similarly strong calibration relationships were obtained in other case studies.

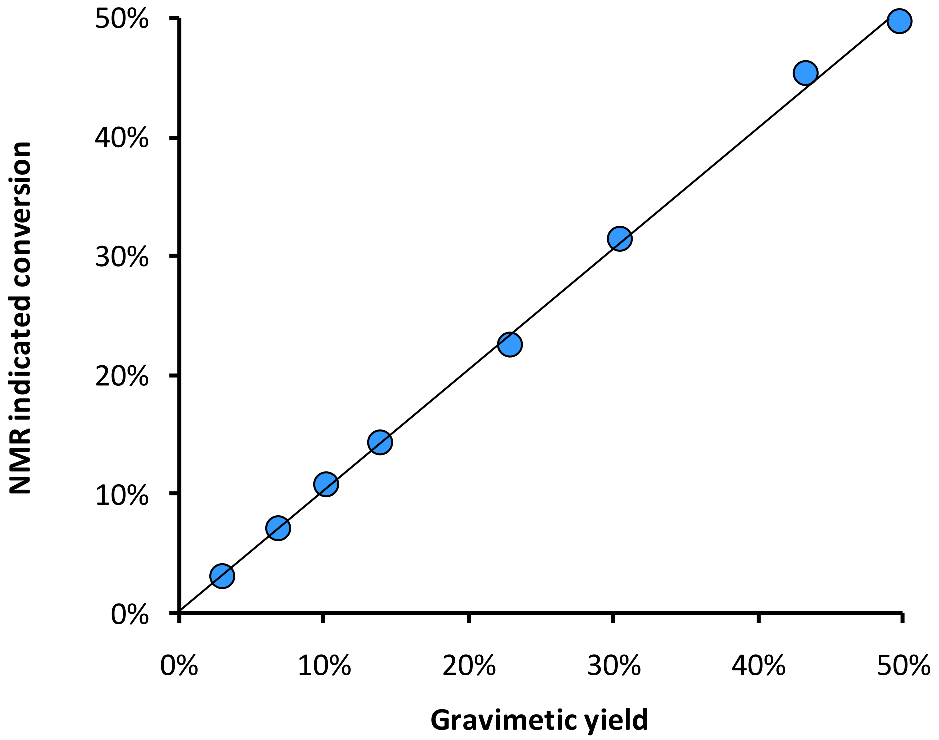


Figure 8.1 The correlation between ^{1}H-NMR signal integrals and the proportion of amide product in solution based on the consumption of benzaldehyde.

Linear solvation energy relationship calculations from Chapter 3: As previously, coefficients for the construction of all LSERs are deemed to be statistically significant if their associated p-value is below 0.01. The following tables document the initial parameter screening for correlations describing ln(**k**), $\Delta H^{\ddagger}$, and $\Delta S^{\ddagger}$ solvent dependant variables (Table 8.8), and final relationships (Table 8.9).

Table 8.8 The statistical significance of LSER coefficients for the kinetics of amidation in the model reaction in the form 'XYZ = $XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^*$'.

XYZ	n	Coefficient (p-value)				R^2
		XYZ_0	*a*	*b*	*s*	
ln(**k**)	7 (original solvent set)	-10.45 ($1.2*10^{-7}$)	n/a	-1.16 ($2.8*10^{-3}$)	-0.04 (0.88)	0.974
ln(**k**)	6 (without cyclohexanone)	-10.52 ($1.2*10^{-6}$)	n/a	-1.30 ($2.0*10^{-3}$)	0.10 (0.59)	0.991
ln(**k**)	8 (with bio-based solvents)	-10.50 ($1.6*10^{-11}$)	n/a	-1.29 ($3.7*10^{-5}$)	0.08 (0.47)	0.993
$\Delta H^\ddagger$	6 (without cyclohexanone)	97.47 ($3.4*10^{-4}$)	n/a	-72.4 ($2.7*10^{-3}$)	13.89 (0.29)	0.988
$\Delta S^\ddagger$	6 (without cyclohexanone)	-73.00 (0.016)	n/a	-204.35 ($2.7*10^{-3}$)	37.97 (0.31)	0.988

Table 8.9 The final LSER coefficients for the kinetics of the model amidation in the form 'XYZ = $XYZ_0 + b\boldsymbol{\beta}$' using only statistically significant parameters.

XYZ	n	Coefficient (p-value)		R^2
		XYZ_0	*b*	
ln(**k**)	7 (original solvent set)	-10.47 ($1.7*10^{-11}$)	-1.19 ($3.7*10^{-5}$)	0.974
ln(**k**)	6 (without cyclohexanone)	-10.47 ($2.8*10^{-10}$)	-1.24 ($3.3*10^{-5}$)	0.991
ln(**k**)	8 (with bio-based solvents)	-10.48 ($1.7*10^{-15}$)	-1.23 ($1.8*10^{-7}$)	0.992
$\Delta H^\ddagger$	6 (without cyclohexanone)	103.71 ($7.9*10^{-7}$)	-63.84 ($1.3*10^{-4}$)	0.982
$\Delta S^\ddagger$	6 (without cyclohexanone)	-55.92 ($5.4*10^{-4}$)	-180.84 ($1.2*10^{-4}$)	0.982

For verification purposes the enthalpy of activation and the entropy of activation can be demonstrated as being dependant on **β** by comparing the relationship between experimental values and those estimated with the respective LSER. The estimation of enthalpy (Figure 8.2) and entropy (Figure 8.3) through Equation 3.4 and Equation 3.5 respectively give satisfactory correlations.

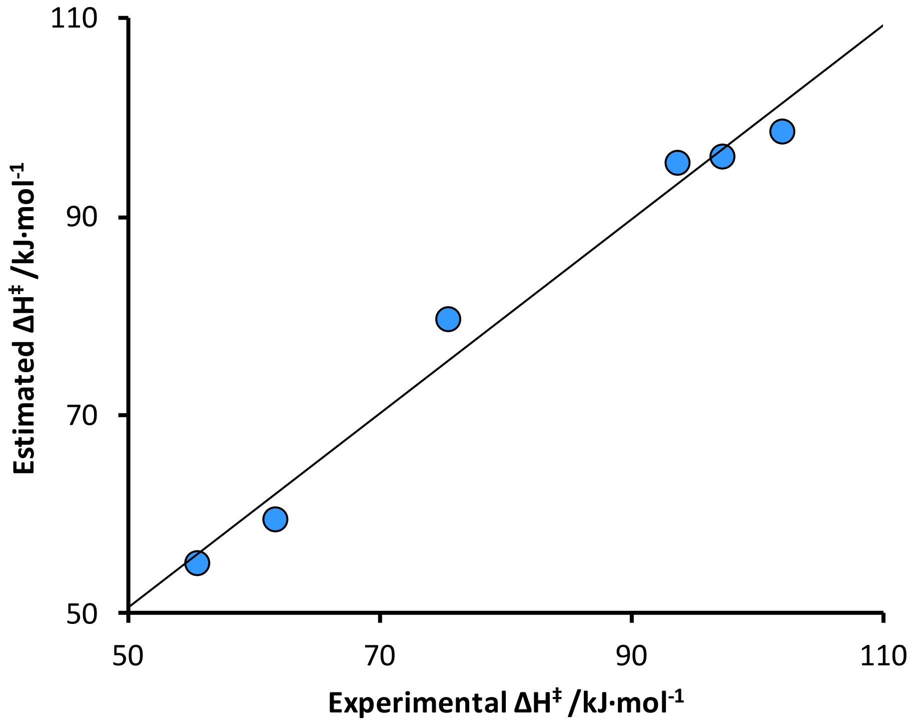


Figure 8.2 A comparison between experimental and calculated amidation enthalpies of activation.

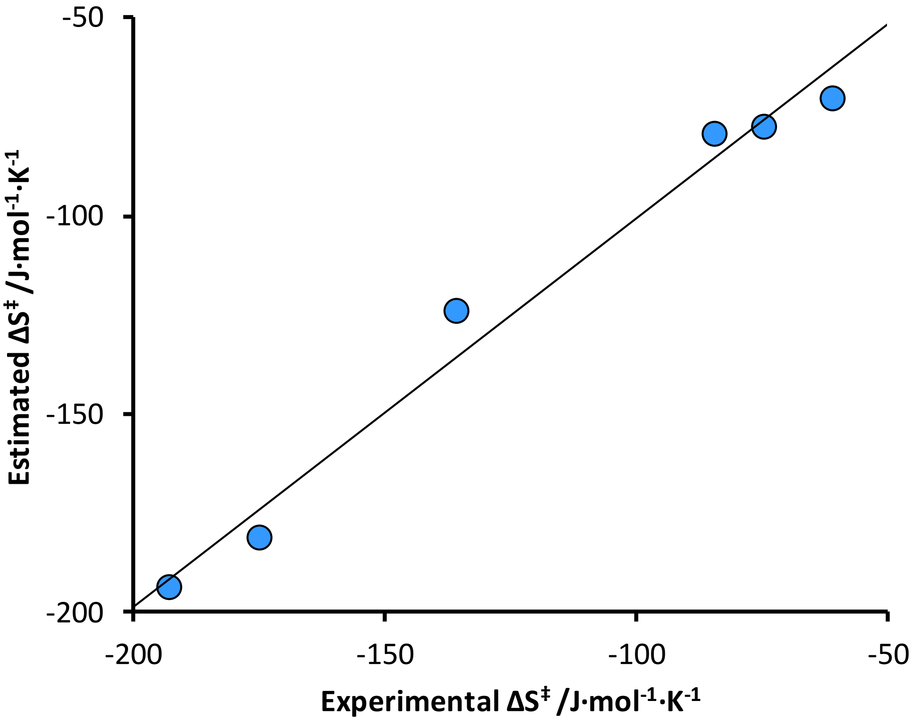


Figure 8.3 The comparison between experimental and calculated entropies of activation in the model amidation reaction.

Metric calculations from Chapter 3: The equations required for calculating atom economy, RME, and PMI are given here [Curzons 2001]:

Equation 8.1

$$\text{Atom Economy} = \frac{\text{RMM of product}}{\sum \text{RMM of all reactants}}$$

Equation 8.2

$$\text{RME} = \frac{\text{Mass of product}}{\sum \text{Masses of reactants}}$$

Equation 8.3

$$\text{PMI} = \frac{\text{Total mass of material inputs}}{\text{Mass of product}}$$

The synthesis mass metrics of *2-N*,*N*-diisopropyl-5-fluorobenzylaminoboronic acid were based on the published experimental procedure [Arnold 2008]. The quantities of solvent and silica used in column chromatography were not available and so these were estimated, employing a 20:1 mass ratio between silica and substrate and supposing that 20 litres of solvent are required per kilogram of silica. Strangely, in their publication Arnold *et al.* give the preparation of the fluorinated catalyst but record the amide yields assisted with its non-fluorinated analogue [Arnold 2008]. It is assumed that the yield of 68% would be the same if *2-N*,*N*-diisopropyl-5-fluorobenzylaminoboronic acid was applied as the catalyst in refluxing fluorobenzene for 24 hours instead of *2-N*,*N*-diisopropylbenzylaminoboronic acid. All mass data required to calculate the metrics used in this work are provided here (Table 8.10). The recrystallisation solvent for the catalyst was ignored, while the other chemicals are correctly scaled to provide enough catalyst for a 10 mol% loading in the amidation reaction to give *N*-benzyl-4-phenylbutanamide.

It should be noted that the metrics in this work are not a comprehensive analysis, which in order to be quantitatively comparable to other reaction protocols would require much more information than was made available. The additional contribution of energy usage should be considered, and combined with mass utilisation in a comparable form. Carbon dioxide equivalents are one way of achieving this, as used for carbon footprinting [Peters 2010]. Thinking in terms of energy equivalents would provide access to the wider scope of LCA [ISO 2006a, ISO 2006b]. Subjects not discussed in this work, including water use, land use, pollution and generally the long term sustainability of materials and processes can then be understood with more clarity.

Table 8.10 Mass utilisation in the extended synthesis of *N*-benzyl-4-phenylbutanamide and the catalyst *2*-*N*,*N*-diisopropyl-5-fluorobenzylaminoboronic acid.

Name	Mass /g	Purpose
Catalyst preparation		
3-Fluorobenzoyl chloride	3.94	Reagent
Diisopropylamine	6.25	Reagent
TMEDA	3.13	Reagent
n-BuLi (1.6 M solution in hexane)	11.00	Reagent
Trimethyl borate	2.80	Reagent
Pinacol	3.20	Reagent
Sodium tetraborahydrate	5.60	Reagent
TMSCl	32.16	Reagent
Diethyl ether	548.2	Work-up
THF	106.7	Work-up
Aqueous reagents	469.2	Work-up
Silica	103.2	Chromatography
Hexane	901.3	Chromatography
Ethyl acetate	617.1	Chromatography
Amidation		
4-Phenylbutanoic acid	17.58	Reagent
Benzylamine	11.47	Reagent
Fluorobenzene	1092	Work-up
DCM	653.1	Work-up
Aqueous reagents	2669	Work-up
Product		
N-Benzyl-4-phenylbutanamide	18.44 (68% yield)	Product

Linear solvation energy relationship calculations from Chapter 4: The following tables document the initial Kamlet-Taft solvent polarity parameter screening (Table 8.11), and final relationships that also include the square of the Hildebrand solubility parameter (Table 8.12).

Table 8.11 The statistical significance of LSER coefficients for the kinetics of an uncatalysed esterification in the form ‘$\ln(\mathbf{k}) = XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^*$’.

n	Coefficient (p-value)				R^2
	XYZ_0	a	b	s	
8 (initial solvent set)	-9.06	-1.91	-4.53	-0.31	0.983
	($1.3*10^{-6}$)	(0.048)	($6.8*10^{-4}$)	(0.54)	
7 (excluding acetonitrile)	-9.09	-0.22	-4.29	-0.42	0.992
	($8.4*10^{-6}$)	(0.83)	($1.3*10^{-3}$)	(0.31)	

Table 8.12 Extended LSER relationships for the kinetics of the uncatalysed model esterification in the form ‘$\ln(\mathbf{k}) = XYZ_0 + b\boldsymbol{\beta}$’ or ‘$\ln(\mathbf{k}) = XYZ_0 + b\boldsymbol{\beta} + h\boldsymbol{\delta}_{\mathbf{T}}^2$’.

n	Coefficient (p-value)			R^2
	XYZ_0	b	$0.001*h$	
8 (initial solvent set)	-9.31	-4.74	n/a	0.929
	($5.9*10^{-9}$)	($1.2*10^{-4}$)		
8 (initial solvent set)	-8.06	-3.48	-3.91	0.988
	($3.9*10^{-6}$)	($1.3*10^{-4}$)	($4.0*10^{-3}$)	
7 (excluding acetonitrile)	-9.25	-4.57	n/a	0.987
	($1.3*10^{-9}$)	($6.4*10^{-6}$)		
7 (excluding acetonitrile)	-8.59	-3.95	-2.16	0.996
	($3.9*10^{-6}$)	($1.3*10^{-4}$)	(0.048)	
11 (all solvents)	-9.20	-4.98	n/a	0.928
	($5.5*10^{-13}$)	($1.9*10^{-6}$)		
11 (all solvents)	-8.13	-3.45	-3.79	0.989
	($4.3*10^{-11}$)	($2.7*10^{-6}$)	($1.4*10^{-4}$)	

Linear solvation energy relationship calculations from Chapter 5: As previously, coefficients for the construction of all LSERs are deemed to be statistically significant if their associated p-value is below 0.01. The following tables document the initial parameter screening for correlations describing $\ln(\mathbf{k})$ of Fischer esterification to give benzyl acetate (Table 8.13), and then final relationships (Table 8.14).

Table 8.13 The statistical significance of LSER coefficients for the kinetics of Fischer esterification to give benzyl acetate expressed in the equation 'ln(**k**) = $XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi}^*$'.

XYZ	n	Coefficient (p-value)				R^2
		XYZ_0	*a*	*b*	*s*	
ln(**k**)	7 (*p*-TSA)	-9.10 ($3.5*10^{-3}$)	-0.82 (0.74)	-6.73 ($9.6*10^{-3}$)	1.23 (0.66)	0.959
ln(**k**)	7 (*p*-CSA)	-9.26 ($4.6*10^{-3}$)	-1.94 (0.50)	-8.21 ($7.6*10^{-3}$)	2.45 (0.45)	0.962
Δln(**k**)	7 (*p*-CSA minus *p*-TSA)	-0.16 (0.37)	-1.12 (0.040)	-1.48 ($2.9*10^{-3}$)	2.45 (0.044)	0.972

Table 8.14 The final LSER coefficients for the kinetics of Fischer esterification in the form 'ln(**k**) = $XYZ_0 + b\boldsymbol{\beta}$' using only statistically significant parameters.

XYZ	n	Coefficient (p-value)		R^2
		XYZ_0	*b*	
ln(**k**)	7 (*p*-TSA)	-8.58 ($3.1*10^{-7}$)	-6.33 ($1.4*10^{-4}$)	0.956
ln(**k**)	7 (*p*-CSA)	-8.27 ($9.5*10^{-7}$)	-7.38 ($1.7*10^{-4}$)	0.953
Δln(**k**)	7 (*p*-CSA minus *p*-TSA)	0.31 ($9.2*10^{-3}$)	-1.05 ($2.8*10^{-3}$)	0.856
Δln(**k**)	6 (*p*-CSA minus *p*-TSA, excluding butanone)	0.33 ($1.6*10^{-3}$)	-1.19 ($5.1*10^{-4}$)	0.963

Conversion of limonene into *p*-menthane and *p*-cymene: The ^{13}C-NMR spectroscopic analysis of the reaction mixture resulting from the action of Pd/C and montmorillonite clay on limonene at 413 K was compared to *p*-cymene (Figure 8.4). The ^{13}C-NMR spectrum in chloroform indicates a lack of olefins but extra alkane signals associated with *p*-menthane.

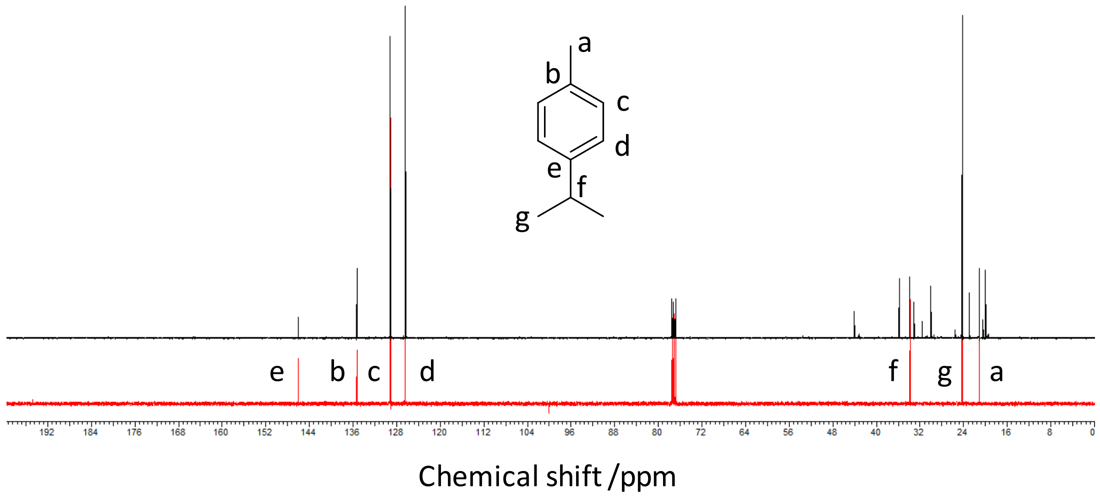


Figure 8.4 Overlap of a pure *p*-cymene standard (obtained from Sigma-Aldrich, red) and an experimental ^{13}C-NMR spectrum of limonene converted to *p*-cymene at 413 K (black).

Determination of tautomerisation equilibrium constants: Methyl acetoacetate (0.81 mL) was dissolved in the chosen solvent (4 mL). The neat solution was analysed by ^{1}H-NMR spectroscopy at 400 MHz in the absence of a deuterated solvent. The relative amounts of the enol and diketo tautomers was determined from the signal intensity ratio of the peaks belonging to the α-olefinic enol proton and the α-methylene diketo protons (Figure 8.5). In instances where solvent signals

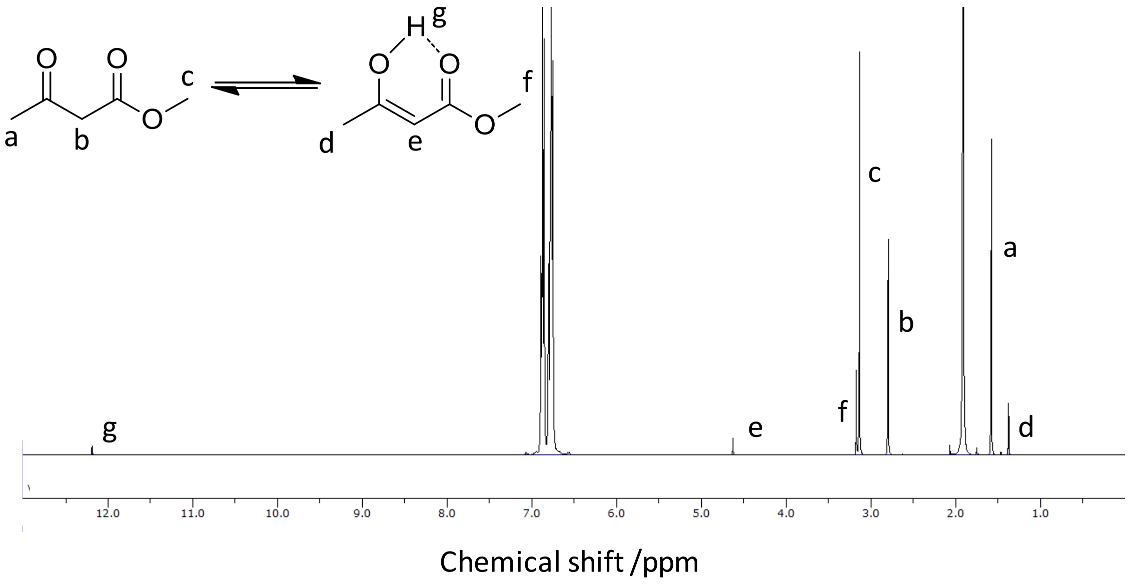


Figure 8.5 Identification of key proton signals in the ^{1}H-NMR spectrum of methyl acetoacetate in toluene. The complete height of the solvent signals is not shown to enhance resolution of the substrate signals.

overlapped with these signals the ratio of methyl group signals could be used. Experiments were conducted at 298 K and repeated to verify accuracy.

Further reaction data from Chapter 6: The full array of data collected from the Biginelli is presented below, including all average yields, tautomerisation equilibrium coefficients and correlations with solvent parameters (Table 8.15, Table 8.16, Table 8.17, and Table 8.18). Valid statistical significance was declared when p-values of coefficients were below 0.01, otherwise these coefficients were assumed to be zero [Wells 2008].

Table 8.15 Full yields and relevant solvent parameters for the synthesis of methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate, including the tautomerisation equilibrium constants of methyl acetoacetate.

Solvent	**π***	ln($\mathbf{K}_T$)	Yield		ln(**P**/**R**)	
			HCl	EPZ-10	HCl	EPZ-10
Acetic acid	0.60	-2.92	35%	53%	-0.60	0.10
t-Butanol	0.58	-2.10	55%	24%	0.21	-1.13
Cyclohexane	0.00	-0.09	68%	n/a	0.73	n/a
p-Cymene	0.39	-1.33	66%	37%	0.66	-0.54
1,2-DCE	0.76	-2.57	44%	21%	-0.23	-1.30
DMF	0.88	-2.91	37%	18%	-0.53	-1.49
Ethanol	0.62	-1.91	56%	30%	0.25	-0.83
Ethyl acetate	0.54	-2.00	51%	31%	0.03	-0.80
Ethylene glycol	1.01	-3.21	38%	25%	-0.51	-1.08
Lactic acid	1.09	-3.48	24%	41%	-1.16	-0.37
Propanoic acid	0.51	-2.35	42%	54%	-0.35	0.16
Toluene	0.50	-1.56	59%	33%	0.35	-0.70

Table 8.16 Full yields and relevant solvent parameters for the synthesis of 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione, including the tautomerisation equilibrium constants of 5,5-dimethyl-1,3-cyclohexanedione.

Solvent	ln($\mathbf{K}_\mathbf{T}$)	Yield	ln(**P**/**R**)
Acetic acid	1.55[a]	27%	-1.02
t-Butanol	7.43[a]	51%	0.02
1,2-DCE	-3.74[a]	3%	-3.46
DMF	4.44	34%	-0.65
Ethanol	5.13	29%	-0.91
Ethyl acetate	1.93[a]	18%	-1.53
Ethylene glycol	2.96[a]	33%	-0.69
Toluene	-2.53	8%	-2.48
Water	2.94	54%	0.16

[a]Predicted from Equation 6.3

Table 8.17 Solvent effects on the diketo-enol tautomerisation of 1,3-dicarbonyls and reaction productivity in the form of the LSER: XYZ = $XYZ_0 + a\boldsymbol{\alpha} + b\boldsymbol{\beta} + s\boldsymbol{\pi^*}$.

XYZ	n	Coefficient (p-value)				R^2
		XYZ_0	*a*	*b*	*s*	
ln($\mathbf{K}_\mathbf{T}$) of methyl acetoacetate	9	-0.13 (2.41×10^{-1})	0.00	0.00	-3.13 (2.09×10^{-7})	0.982
ln($\mathbf{K}_\mathbf{T}$) of 5,5-dimethyl-1,3-cyclohexanedione	9	-3.74 (8.19×10^{-5})	0.00	11.77 (3.05×10^{-6})	0.00	0.962
ln(**P**/**R**) w/ HCl[a]	7	1.17 (3.61×10^{-3})	0.00	0.00	-1.76 (2.53×10^{-3})	0.862
ln(**P**/**R**) w/ EPZ-10[a]	6	0.22 (2.96×10^{-1})	0.00	0.00	-1.97 (1.47×10^{-3})	0.889
ln(**P**/**R**) w/ HCl[b]	8	-3.02 (2.85×10^{-5})	0.00	3.33 (3.30×10^{-4})	0.00	0.899

[a]Product is methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate.

[b]Product is 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione.

Table 8.18 Correlations between reaction productivity and tautomerisation in the following form: $\ln(\mathbf{P}/\mathbf{R}) = XYZ_0 + t \cdot \ln(\mathbf{K}_T)$.

XYZ	n	Coefficient (p-value)		R^2
		XYZ_0	t	
ln(**P**/**R**) w/ HCl[a]	11	1.59 ($7.30*10^{-6}$)	0.74 ($2.34*10^{-6}$)	0.925
ln(**P**/**R**) w/ EPZ-10[a]	7	0.29 ($9.91*10^{-2}$)	0.62 ($2.88*10^{-4}$)	0.941
ln(**P**/**R**) w/ HCl[b]	8	-1.95 ($1.68*10^{-5}$)	0.28 ($3.03*10^{-4}$)	0.902

[a] Product is methyl-1,2,3,4-tetrahydro-6-methyl-2-oxo-4-phenyl-5-pyrimidinecarboxylate.

[b] Product is 4,6,7,8-tetrahydro-7,7-dimethyl-4-phenyl-2,5(1H,3H)-quinazolinedione.

Abbreviations and symbols

α	Kamlet-Taft scale of solvent hydrogen bond donating ability
β	Kamlet-Taft scale of solvent hydrogen bond accepting ability
$\Delta\Delta G^{\ddagger}$	Change in Gibbs free energy of activation
$\Delta G^{\ddagger}$	Gibbs free energy of activation
ΔG°	Gibbs free energy of formation
$\Delta H^{\ddagger}$	Enthalpy of activation
ΔH_{vap}	Enthalpy of vaporisation
$\Delta \ln(k)$	Relative change in the natural logarithm of the reaction rate constant
$\Delta S^{\ddagger}$	Entropy of activation
δ_C	Chemical shift of carbon nuclei signals in nuclear magnetic resonance spectra
δ_H	Chemical shift of hydrogen nuclei signals in nuclear magnetic resonance spectra
δ_P	Hansen solubility parameter of dipolarity
δ_D	Hansen solubility parameter of dispersion forces
δ_H	Hansen solubility parameter of hydrogen bonding
δ_T	Hildebrand solubility parameter
δ	Kamlet-Taft scale polarisability correction term
ε	Hydrogen bond donating ability correction term
ε_r	Relative permittivity
λ	Wavelength
μW	Microwave energy irradiation
ν	Frequency
$\tilde{\nu}$	Wavenumber

π^*	Kamlet-Taft scale of solvent dipolarity/polarisability
π^*	Pi anti-bonding orbital
ρ	Reaction constant
σ	Substituent constant
A	Arrhenius collision pre-factor
$A_{AC}2$	Bimolecular acyl cleavage reaction mechanism
$A_{AL}2$	Bimolecular alkyl cleavage reaction mechanism
$[A]_0$	Initial concentration of the reactant designated as A
$A_{AC}1$	Unimolecular acyl cleavage reaction mechanism
$A_{AL}1$	Unimolecular alkyl cleavage reaction mechanism
a	Coefficient of the Kamlet-Taft scale of solvent hydrogen bond donating ability in linear solvation energy relationships
ACSGCI	American chemical society green chemistry institute
API	Active pharmaceutical ingredient
Aq.	Aqueous
$[B]_0$	Initial concentration of the reactant designated as B
b	Coefficient of the Kamlet-Taft scale of solvent hydrogen bond accepting ability in linear solvation energy relationships
BHT	Butylated hydroxytoluene
bs	Broad singlet
BSA	Benzenesulphonic acid
Bu	Butyl
C_t	Conversion at time t
c	Light velocity
CAPEC	Computer assisted process-product engineering centre
Cat.	Catalyst
CDI	*N*,*N*'-Carbonyldiimidazole
cm	Centimetre

^{13}C-NMR	Carbon-13 nuclear magnetic resonance
COMU	(1-Cyano-2-ethoxy-2-oxoethylidenaminooxy)dimethylamino-morpholino-carbenium hexafluorophosphate
CSA	Cymenesulphonic acid
CyH	Cyclohexane
Cym	*p*-Cymenesulphonyl
D	Debye
d	Coefficient of the Kamlet-Taft polarisability correction term in linear solvation energy relationships
d	Doublet
DCC	Dicyclohexylcarbodiimide
DCE	Dichloroethane
DCM	Dichloromethane
DGME	Diethylene glycol monobutyl ether
dm	Decimetre
DMAc	*N*,*N*-Dimethyl acetamide
DMAP	*N*,*N*-Dimethyl-4-aminopyridine
DMF	*N*,*N*-Dimethyl formamide
DMPU	Dimethyl-1,3-propylene urea
DMSO	Dimethyl sulphoxide
DTU	Technical University of Denmark
E_a	Activation energy
$\mathrm{E_T}$	Dimroth-Reichardt’s betaine dye electonic transition energy
$\mathrm{E_T^N}$	Reichardt’s normalised scale of solvatochromism
e	Coefficient of the Kamlet-Taft hydrogen bond donating ability correction term in linear solvation energy relationships
EC	European commission
EC_{50}	Half maximal effective concentration

ECO	Environmental index
E-Factor	Environmental-factor metric
EHS	Environmental, health and safety
EI	Electron impact ionisation
E-Impact	Environmental impact
ESI	Electrospray ionisation
Et	Ethyl
ETH	Eidgenössische Technische Hochschule Zürich (Zurich Technical Institute)
f	Coefficient of Reichardt's normalised scale of solvatochromism in linear solvation energy relationships
g	Coefficient of relative permittivity in linear solvation energy relationships
g	Gramme
GC	Gas chromatography
GSK	GlaxoSmithKline
H_0	Hammett acidity function
H	Number of hydrogens responsible for a NMR signal
h	Planck constant
HAS	Health and safety index
HATU	*N*-[(Dimethylamino)-1*H*-1,2,3-triazolo-[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide
HMBC	Heteronuclear multiple-bond correlation
HMPA	Hexamethylphosphoramide
^{1}H-NMR	Hydrogen-1 nuclear magnetic resonance
HPLC	High performance liquid chromatography
hr	Hour
HSQC	Heteronuclear single quantum correlation
Hz	Hertz
I	NMR signal integral

J	Indirect dipole-dipole (J-coupling) constant
J	Joule
K	Equilibrium constant
K_a	Equilibrium constant of acid dissociation in water
K_0	Equilibrium constant (reference)
K	Kelvin
K_T	Tautomerisation equilibrium constant
k_B	Boltzmann constant
k	Rate constant
Kg	Kilogram
kJ	Kilojoule
L	Litre
LC_{50}	Lethal concentration affecting 50% of the sample population
LCA	Life cycle assessment
LD_{50}	Lethal dose affecting 50% of the sample population
LFER	Linear free energy relationship
ln	Natural logarithm function
log	Logarithm (base ten) function
LSER	Linear solvation energy relationship
M	Molar
m	Multiplet
Me	Methyl
MeOH	Methanol
2-MeTHF	2-Methyltetrahydrofuran
mg	Milligram
MHz	Megahertz
min	Minute
mL	Millilitre

mmol	Millimole
mol	Mole
MPa	Megapascals
MS	Mass spectrometry
MSDS	Material safety datasheet
N_A	Avogadro constant
N	No
n	*Normal*-
n/a	Not applicable
NGO	Non-governmental organisation
nm	Nanometre
NMP	*N*-Methyl pyrrolidinone
NMR	Nuclear magnetic resonance
NO_x	Nitrogen oxides
Nuc.	Nucleophile
OTf	Triflate anion, CF_3SO_3-
OTs	Tosylate anion, $CH_3C_6H_4SO_3$-
$[P]_t$	Concentration of the product at time t
P	Lipophilicity
P	Product
p	*Para*-
PEG	Poly(ethylene glycol)
PET	Poly(ethylene terephthalate)
Ph	Phenyl
pH	Negative logarithm of the activity (concentration) of oxonium ions
PhCl	Chlorobenzene
pK_a	Negative logarithm of the equilibrium constant of acid dissociation in water

$\mathrm{pK}_{\mathrm{BH}^+}$	Negative logarithm of the equilibrium constant of protonated base dissociation in water
PLA	Poly(lactic acid)
PMI	Process mass intensity
ppm	Parts per million
q	Quadruplet
R	Alkyl group
R^2	Coefficient of determination
R	Gas constant
R	Reactant
R_i	Reaction index i
REACH	Registration, evaluation, authorisation and restriction of chemicals
Ref.	Reference
RME	Reaction mass efficiency
RMM	Relative molecular mass
RS_i	Reaction-solvent index i
S_N1	First order nucleophilic substitution reaction mechanism
S_N2	Second order nucleophilic substitution reaction mechanism
s	Coefficient of the Kamlet-Taft scale of solvent dipolarity/polarisability in linear solvation energy relationships
s	Second
s	Singlet
S_i	Solvent score i
S_NAr	Second order aromatic nucleophilic substitution reaction mechanism
SO_x	Sulphur oxides
SUS	Renewability index
T_b	Boiling point
T_d	Decomposition temperature

T_m	Melting point
T_s	Mid-point between solvent melting point and boiling point
T_r	Reaction temperature
T	Temperature
T_x	Temperature flexibility in Rule C of the solvent selection algorithm
t	Coefficient of the natural logarithm of the tautomerisation equilibrium constant in linear solvation energy relationships
t	*Tertiary-*
t	Triplet
THF	Tetrahydrofuran
TMEDA	Tetramethylethylenediamine
TMS	Tetramethylsilane
TMSCl	Trimethylsilyl chloride
TSA	Toluenesulphonic acid
UV	Ultra violet
UV-vis.	Ultra violet-visible light
V_m	Molar volume
w/w	Weight ratio
wt	Weight
X	Electron rich atom or chemical moiety
XYZ	Linear solvation energy relationship dependant variable
XYZ_0	Linear solvation energy relationship proportionality constant
Y	Yes
Y	Electron donor

References

Abbott 2009	A. P. Abbott, E. G. Hope, R. Mistry and A. M. Stuart, *Green Chem.*, 2009, **11**, 1530.
Abboud 1977	J. L. Abboud, M. J. Kamlet and R. W. Taft, *J. Am. Chem. Soc.*, 1977, **99**, 8325.
Abboud 1999	J. –L. M. Abboud and R. Notario, *Pure Appl. Chem.*, 1999, **71**, 645.
Abraham 1969	M. H. Abraham, *Chem. Comm.*, 1969, 1307.
Abraham 1975	M. H. Abraham and R. J. Abraham, *J. Chem. Soc. Perkin Trans. II*, 1975, 1677.
Abraham 1981	M. H. Abraham and A. Nesehzadeh, *Chem Comm.*, 1981, 905.
Abraham 1985	M. H. Abraham, *Pure Appl. Chem.*, 1985, **57**, 1055.
Achema 2012	Achema, press release, *New Natural Resource Base in the Chemical Industry: Only a Matter of Time*, 2012. Available online at presse.achema.de/achema_presse/en/ ACHEMA+Press+Releases/Trend+Reports/tb_19_en+Biobased+Chemicals.print. Accessed 14th June 2013.
Adams 2004	D. J. Adams, P. J. Dyson and S. J. Tavener in *Chemistry In Alternative Reaction Media*, John Wiley and Sons, Chichester, 2004.
Adibi 2007	H. Adibi, H. A. Samimi and M. Beygzadeh, *Catal. Commun.* 2007, **8**, 2119.
Al-Zoubi 2008	R. M. Al-Zoubi, O. Marion and D. G. Hall, *Angew. Chem. Int. Ed.*, 2008, **47**, 2876.
Alfonsi 2008	K. Alfonsi, J. Colberg, P. J. Dunn, T. Fevig, S. Jennings, T. A. Johnson, H. P. Kleine, C. Knight, M. A. Nagy, D. A. Perry and M. Stefaniak, *Green Chem.*, 2008, **10**, 31.
Alonso 1999	M. C. Alonso and D. Barceló, *Anal. Chim. Acta*, 1999, **400**, 211.

Alonso 2013	D. M. Alonso, S. G. Wettstein and J. A. Dumesic, *Green Chem.*, 2013, **15**, 584.
Anastas 1998	P. T. Anastas and J. C. Warner in *Green Chemistry Theory and Practice*, OUP, Trowbridge, 1998.
Anastas 2002	P. T. Anastas and M. M. Kirchhoff, *Acc. Chem. Res.*, 2002, **35**, 686.
Anastas 2010	P. Anastas and N. Eghbali, *Chem. Soc. Rev.*, 2010, **39**, 301.
Anderson 2006	J. E. Anderson, R. Davis, R. N. Fitzgerald and J. M. Haberman, *Synthetic Commun.*, 2006, **36**, 2129.
Anellotech 2013	Anellotech, press release, *Anellotech Announces Ability to Produce Large Volume Product Development Samples of Biomass-Derived Benzene and Toluene*, 2013. Available online at www.anellotech.com/release2.html. Accessed 14th June 2013.
Angelici 2013	C. Angelici, B. M. Weckhuysen and P. C. A. Bruijnincx, *ChemSusChem*, 2013, advanced article, DOI: 10.1002/cssc.201300214. Accessed 14th June 2013.
Arnold 2006	K. Arnold, B. Davies, R. L. Giles, C. Grosjean, G. E. Smith and A. Whiting, *Adv. Synth. Catal.*, 2006, **348**, 813.
Arnold 2008	K. Arnold, A. S. Batsanov, B. Davies and A. Whiting, *Green Chem.*, 2008, **10**, 124.
Ashley 2006	M. Ashley*, Ingenia*, 2006, **29**, 46.
Avdeev 2007	Y. G. Avdeev, P. A. Belinski, Y. I. Kuznetsov and O. O. Zel, *Prot. Met.*, 2007, **43**, 587.
Aycock 2007	D. F. Aycock, *Org. Process Res. Dev.*, 2007, **11**, 156.
Baghernejad 2011	B. Baghernejad, *Curr. Org. Chem.*, 2011, **15**, 3091.
Balat 2009	M. Balat and H. Balat, *Appl. Energy*, 2009, **86**, 2273.
Balcom 1989	B. J. Balcom and N. O. Petersen*, J. Org. Chem.*, 1989, **54**, 1922.
Baldwin 1996	B. W. Baldwin, T. Hirose, and Z. –H. Wang, *Chem. Commun.*, 1996, 2669.
Bamford 1972	C. H. Bamford and C. F. H. Tipper in *Comprehensive Chemical Kinetics*, volume 10, Elsevier publishing company, Amsterdam, 1972.

Bankole 2011	K. S. Bankole in *Uncatalyzed Esterification of Biomass-Derived Carboxylic Acids*, PhD thesis, University of Iowa, 2011.
Banthorpe 1972	D. V. Benthorpe, B. V. Charlwood and M. J. O. Francis, *Chem. Rev.*, 1972, **72**, 115.
Barbosa 2006	S. L. Barbosa, M. J. Dabdoub, Gabriela R. Hurtado, Stanlei I. Klein, A. C. M. Baroni and C. Cunha, *Appl. Catal. A: Gen.*, 2006, **313**, 146.
Barbour 1952	A. K. Barbour, G. B. Barlow and J. C. Tatlow, *J. Appl. Chem.*, 1952, **2**, 127.
BASF 2013	BASF, press release, *BASF plans production of butanediol from renewable feedstock using Genomatica technology*, 2013. Available online at www.basf.com/group/pressrelease/P-13-268. Accessed 14th June 2013.
Bauer 2013	F. Bauer and C. Hulteberg, *Biofuels Bioprod. Bioref.*, 2013, **7**, 43.
Beamson 2010	G. Beamson, A. J. Papworth, C. Philipps, A. M. Smith and R. Whyman, *J. Catal.*, 2010, **269**, 93.
Bender 1951	M. Bender, *J. Am. Chem. Soc.*, 1951, **73**, 1626.
Berg 1992	L. Berg, US pat., 5106459, 1992.
Berti 2010	C. Berti, E. Binassi, M. Colonna, M. Fiorini, G. Kannan, S. Karanam, M. Mazzacurati, I. Odeh and M. Vannini, international pat., 078328, 2010.
Biginelli 1891a	P. Biginelli, *Ber.*, 1891, **24**, 1317.
Biginelli 1891b	P. Biginelli, *Ber.*, 1891, **24**, 2962.
BioAmber 2013	BioAmber, manufacturer information. Available online at www.bio-amber.com/products/en/products/bdo_1_4_ butanediol. Accessed 14th June 2013.
Blackmond 2007	D. G. Blackmond, A. Armstrong, V. Coombe and A. Wells, *Angew. Chem. Int. Ed.*, 2007, **46**, 3798.
Boots 1989	H. M. J. Boots and P. K. de Bokx, *J. Phys. Chem.*, 1989, **93**, 8240.
Bose 2003	D. S. Bose, L. Fatima and H. B. Mereyala, *J. Org. Chem.* 2003, **68**, 587.
Bose 2004	A. K. Bose, S. Pednekar, S. N. Ganguly, G. Chakraborty and M. S. Manhas, *Tetrahedron Lett.* 2004, **45**, 8351.

BP 2013 — British Petroleum, *Statistical Review of World Energy*, 2013. Available online at www.bp.com/en/global/corporate/about-bp/statistical-review-of-world-energy-2013.html. Accessed 14th June 2013.

Breeden 2012 — S. W. Breeden, J. H. Clark, D. J. Macquarrie and J. Sherwood in *Green Techniques for Organic Synthesis and Medicinal Chemistry*, edited by W. Zhang and B. W. Cue, John Wiley and Sons, Chichester, 2012.

Brettschneider 2010 — F. Brettschneider, V. Jankowski, T. Günthner, S. Salem, M. Nierhaus, A. Schulz, W. Zidek and J. Jankowski, *J. Chromatogr. B*, 2010, **878**, 763.

Brown 1950 — S. M. Brown and G. H. Thomson, *J. Chem. Soc.*, 1950, 1019.

Burgula 2012 — L. N. Burgula, K. Radhakrishnan and L. M. Kundu, *Tetrahedron Lett.* 2012, **53**, 2639.

Capello 2007 — C. Capello, U. Fischer and K. Hungerbühler, *Green Chem.*, 2007, **9**, 927.

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

Carlson 2009 — T. R. Carlson, G. A. Tompsett, W. C. Conner and G. W. Huber, *Top. Catal.*, 2009, **52**, 241.

Castarlenas 2006 — R. Castarlenas, C. Vovard, C. Fischmeister and P. H. Dixneuf, *J. Am. Chem. Soc.*, 2006, **128**, 4079.

Castejon 1999 — H. Castejon and K. B. Wiberg, *J. Am. Chem. Soc.*, 1999, **121**, 2139.

Catalán 2009 — J. Catalán, *J. Phys. Chem. B*, 2009, **113**, 5951.

Cathay Biotech 2013 — Cathay Industrial Biotech, manufacturer information. Available online at www.cathaybiotech.com/en/products/bioacetone. Accessed 21st June 2013.

Cave 2001 — G. W. V. Cave, C. L. Raston and J. L. Scott, *Chem. Commun.*, 2001, 2159.

Cepanec 2007 — I. Cepanec, M. Litvić, M. Filipan-Litvić and I. Grüngold, *Tetrahedron* 2007, **63**, 11822.

Chakraborti 2003 — A. K. Chakraborti and R. Gulhane, *Tetrahedron Lett.*, 2003, **44**, 6749.

Chanda 2009 — A. Chanda and V. V. Fokin, *Chem. Rev.* 2009, **109**, 725.

Chandra 2002	K. L. Chandra, P. Saravanan, R. K. Singh and V. K. Singh, *Tetrahedron*, 2002, **58**, 1369.
Chandrasekhar 2002a	S. Chandrasekhar, S. S. Sultana, C. Narsihmulu, J. S. Yadav, R. Gree and J. C. Guillemin, *Tetrahedron Lett.*, 2002, **43**, 8335.
Chandrasekhar 2002b	S. Chandrasekhar, C. Narsihmulu, S. S. Sultana and N. R. Reddy, *Org. Lett.*, 2002, **4**, 4399.
Charlson 1987	R. J. Charlson, J. E. Lovelock, M. O. Andreae and S. W. Warren, *Nature*, 1987, **326**, 655.
Charville 2010	H. Charville, D. Jackson, G. Hodges and A. Whiting, *Chem. Commun.*, 2010, **46**, 1813.
Charville 2011	H. Charville, D. A. Jackson, G. Hodges, A. Whiting and M. R. Wilson, *Eur. J. Org. Chem.*, 2011, 5981.
Chem. Strat. Gr. 2013	Chemical Strategies Group, market survey, *Bio-butanol: The Game Changer*, Informa Economics, 2013. Prospectus available online at www.informaecon.com/MCSBiobutanol2013.pdf. Accessed 21st June 2013.
ChemIDPlus 2013	US National Library of Medicine physical property and toxicity database, ChemIDPlus Advanced, 2013. Available online at chem.sis.nlm.nih.gov/chemidplus/. Accessed 14th June 2013.
Chocholoušová 2003	J. Chocholoušová, J. Vacek and P. Hobza, *J. Phys. Chem. A*, 2003, **107**, 3086.
Choudary 2001	B. M. Choudary, V. Bhaskar, M. L. Kantam, K. K. Rao and K. V. Raghavan, *Catal. Lett.*, 2011, **74**, 207.
Claisen 1896	L. Claisen, *Liebigs Ann. Chem.*, 1896, **291**, 25.
Clark 1989	J. H. Clark, A. P. Kybett, D. J. Macquarrie, S. J. Barlow and P. J. Landon, *Chem. Comm.* 1989, 1353.
Clark 1996	J. H. Clark, D. J. Macquarrie, *Chem. Soc. Rev.* 1996, **25**, 303.
Clark 2005	J. H. Clark in *Green Separation Processes*, edited by C. A. M. Afonso and J. G. Crespo, Wiley-VCH, Weinheim, 2005.
Clark 2012	J. H. Clark, D. J. Macquarrie and J. Sherwood, *Green Chem.*, 2012, **14**, 90.

Clark 2013	J. H. Clark, D. J. Macquarrie and J. Sherwood, *Chem. Eur. J.*, 2013, **19**, 5174.
Colley 2004	S. W. Colley, C. R. Fawcett, C. Rathmell and M. W. M. Tuck, US pat., 6809217, 2004.
Colonna 2011	M. Colonna, C. Berti, M. Fiorini, E. Binassi, M. Mazzacurati, M. Vanninia and S. Karanam, *Green Chem.*, 2011, **13**, 2543.
Comerford 2009	J. W. Comerford, J. H. Clark, D. J. Macquarrie and S. W. Breeden, *Chem. Commun.*, 2009, 2562.
Comerford 2012	J. W. Comerford, T. J. Farmer, D. J. Macquarrie, S. W. Breeden and J. H. Clark, *ARKIVOC*, 2012, 282.
Constable 2002	D. J. C. Constable, A. D. Curzons and V. L. Cunningham, *Green Chem.*, 2002, **4**, 521.
Constable 2007a	D. J. C. Constable, C. Jimenez-Gonzalez and R. K. Henderson, *Org. Process Res. Dev.*, 2007, **11**, 133.
Constable 2007b	D. J. C. Constable, P. J. Dunn, J. D. Hayler, G. R. Humphrey, J. L. Leazer, Jr., R. J. Linderman, K. Lorenz, J. Manley, B. A. Pearlman, A. Wells, A. Zaks and T. Y. Zhang, *Green Chem.*, 2007, **9**, 411.
Corma 2003	A. Corma and H. García, *Chem. Rev.*, 2003, **103**, 4307.
Cossy 1989	J. Cossy and C. Pale-Grosdemange, *Tetrahedron Lett.*, 1989, **30**, 2771.
Crowhurst 2003	L. Crowhurst, P. R. Mawdsley, J. M. Perez-Arlandis, P. A. Salter and T. Welton, *Phys. Chem. Chem. Phys.*, 2003, **5**, 2790.
Cue 2012	B. W. Cue in *Green Techniques for Organic Synthesis and Medicinal Chemistry*, edited by W. Zhang and B. W. Cue, John Wiley and Sons Ltd, Chichester, 2012.
Curzons 1999	A. D. Curzons, D. C. Constable, V. L. Cunningham, *Clean Products and Processes*, 1999, **1**, 82.
Curzons 2001	A. D. Curzons, D. J. C. Constable, D. N. Mortimera and V. L. Cunningham, *Green Chem.*, 2001, **3**, 1.
Datta 2005	R. Datta, US pat., 6890893, 2005.
Datta 2008	R. Datta, US pat., 0029740, 2008.

Debache 2008	A. Debache, M. Amimour, A. Belfaitah, S. Rhouati and B. Carboni, *Tetrahedron Lett.* 2008, **49**, 6119.
Delample 2010	M. Delample, N. Villandier, J. –P. Douliez, S. Camy, J. –S. Condoret, Y. Pouilloux, J. Barrault and F. Jérôme, *Green Chem.*, 2010, **12**, 804.
Delpuech 1965	J. J. Delpuech, *Tetrahedron Lett.*, 1965, **25**, 2111.
Derfer 1979	J. M. Derfer and M. M. Derfer in *Kirk–Othmer Encyclopedia of Chemical Technology*, volume 22, edited by M. Grayson, John Wiley and Sons, New York, 1979.
D'hooghe 2008	M. D'hooghe, J. Baele, J. Contreras, M. Boelens and N. De Kimpe, *Tetrahedron Lett*., 2008, **49**, 6039.
Diaz-Álverez 2013	A. E. Diaz-Álverez and V. Cadierno, *Appl. Sci.*, 2013, **3**, 55.
Dilmaghani 2009	K. A. Dilmaghani, B. Zeynizadeh and M. Yari, *Phosphorus Sulfur* 2009, **184**, 1722.
Dimroth 1969	K. Dimroth and C. Reichardt, *Liebigs Ann. Chem.*, 1969, **727**, 93.
Dugger 2005	R. W. Dugger, J. A. Ragan and D. H. B. Ripin, *Org. Process Res. Dev.*, 2005, **9**, 253.
Dunn 2012	P. J. Dunn, *Chem. Soc. Rev.,* 2012, **41**, 1452.
EC 1999	European commission, council directive 1999/13/EC, *on the Limitation of Emissions of Volatile Organic Compounds due to the Use of Organic Solvents in Certain Activities and Installations*, 1999.
EC 2007	European commission, council directive 1907/2006, *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*, 2007.
ECOTOX 2013	USA Environmental Protection Agency (EPA) toxicity database, *ECOTOX Release*, version 4.0, Available online at cfpub.epa.gov/ecotox/. Accessed 14th June 2013.
El-Hamouly 2006	W. S. El-Hamouly, A. –M. A. El-Khamry and E. M. H. Abbas, *Indian J. Chem.* 2006, **45B**, 2091.

EurObser’ER 2012	EurObserv'ER in *Biofuels Barometer*, 2012, **210**, 42. Available online at www.eurobserv-er.org/pdf/baro212.asp. Accessed 14th June 2013.
Euro. Bio-plastics 2013	European Bio-plastics organisation, market statistics. Available online at en.european-bioplastics.org/press/press-pictures/ labelling-logos-charts/. Accessed 14th June 2013.
Exner 1964	O. Exner, *Nature*, 1964, **201**, 488.
Exner 1972	O. Exner in *Advances in Linear Free Energy Relationships*, edited by N. B. Chapman and J. Shorter, Plenum Press, London, 1972.
Fábos 2009	V. Fábos, G. Koczó, H. Mehdi, L. Boda and I. T. Horváth, *Energy Environ. Sci.*, 2009, **2**, 767.
Fan 2013	D. Fan, D. –J. Dai and H. –S. Wu, *Materials*, 2013, **6**, 101.
Ferroud 2008	C. Ferroud, M. Godart, S. Ung, H. Borderies and A. Guy, *Tetrahedron Lett.*, 2008, **49**, 3004.
Fessard 2007	T. C. Fessard, S. P. Andrews, H. Motoyoshi and E. M. Carreira, *Angew. Chem. Int. Ed.*, 2007, **46**, 9331.
Le Fèvre 1934	R. J. W. Le Fèvre, *J. Chem. Soc.*, 1934, 1501.
Fiege 1995	H. Fiege in *Ullmann’s Encyclopedia of Industrial Chemistry*, fifth edition, volume A8, edited by B. Elvers, S. Hawkins and W. Russey, Wiley VCH, Weinheim, 1995.
Folkers 1932	K. Folkers, H. J. Harwood and T. B. Johnson, *J. Am. Chem. Soc.*, 1932, **54**, 3751.
Folkers 1933	K. Folkers and T. B. Johnson, *J. Am. Chem. Soc.*, 1933, **54**, 3784.
Foster 2012	A. J. Foster, J. Jae, Y. –T. Cheng, G. W. Huber and R. F. Lobo, *Appl. Catal. A: Gen.*, 2012, **154**, 423.
Frenkel 1983	M. Frenkel and L. Heller-Kallai, *Clays and Clay Miner.*, 1983, **31**, 92.
Fukaya 1992	M. Fukaya, Y. S. Park and K. Roda, *J. Appl. Bacteriol.*, 1992, **73**, 447.
Fukui 2002	Y. Fukui, K. Aburai and K. Sakamoto, US pat., 6414193, 2002.
Gangadasu 2006	B. Gangadasu, S. Palaniappan, C. A. Amarnath and V. J. Rao, *J. Appl. Polymer Sci.*, 2006, **102**, 1741.
Gani 2005	R. Gani, C. Jiménez-González and D. J. C. Constable, *Comput. Chem. Eng.*, 2005, **29**, 1661.

Gani 2008	R. Gani, P. A. Gómez, M. Folić, C. Jiménez-González and D. J. C. Constable, *Comput. Chem. Eng.*, 2008, **32**, 2420.
Georgiou 2009	I. Georgiou, G. Ilyashenko and A. Whiting, *Acc. Chem. Res.*, 2009, **42**, 756.
Gerlach 2006	T. Gerlach, F. Haese, A. Meier, J. –P. Melder and H. Rütter, world pat., 097468, 2006.
Gevo 2013	Gevo, press release, *Gevo Resumes Commercial Production of Isobutanol*, 2013. Available online at ir.gevo.com/phoenix.zhtml?c=238618&p=irol-newsArticle&ID=1830806&highlight=. Accessed 21st June 2013.
Ghassamipour 2010	S. Ghassamipour and A. R. Sardarian, *J. Iran. Chem. Soc.*, 2010, **7**, 237.
Ghose 1999	A. K. Ghose, V. N. Viswanadhan and J. J. Wendoloski, *J. Comb. Chem.*, 1999, **1**, 55.
Global Bioenergies 2012a	Global Bioenergies, press release, *Global Bioenergies Opens the Way to Bio-Sourced Propylene*, 2012. Available online at www.global-bioenergies.com/index.php?option=com_content&view=article&id=117&Itemid=200&lang=en. Accessed 14th June 2013.
Global Bioenergies 2012b	Global Bioenergies, press release, *Global Bioenergies on Track with the Development Schedule of its Bio-Isobutene Production Process*, 2012. Available online at www.global-bioenergies.com/index.php?option=com_content&view=article&id=117&Itemid=200&lang=en. Accessed 14th June 2013.
Global Bioenergies 2013	Global Bioenergies, press release, *Industrial Pilot Plant Supported by a €5.2m State Financing*, 2013. Available online at www.global-bioenergies.com/index.php?option=com_content&view=article&id=117&Itemid=200&lang=en. Accessed 14th June 2013.
Gonçalves 2008	V. L. C. Gonçalves, B. P. Pinto, J. C. Silva and C. J. A. Mota, *Catal. Today*, 2008, **133**–**135**, 673.
Gooßen 2009	L. J. Gooßen, D. M. Ohlmann and P. P. Lange, *Synthesis*, 2009, 160.

Gorbanev 2012 Y. Y. Gorbanev, S. Kegnæs, C. W. Hanning, T. W. Hansen and A. Riisager, *ACS Catal.*, 2012, **2**, 604.

Gu 2010 Y. Gu and F. Jérôme, *Green Chem.*, 2010, **12**, 1127.

Gupta 2006 R. Gupta, S. Paul and R. J. Gupta, *Mol. Catal. A: Chem.*, 2007, **266**, 50.

Hagen 2006 J. Hagen in *Industrial Catalysis: A Practical Approach*, second edition, Wiley VCH, Weinheim, 2006.

Hahn-Hägerdal 2006 B. Hahn-Hägerdal, M. Galbe, M. F. Gorwa-Grauslund, G. Lidén, G. Zacchi, *TRENDS Biotechnol.*, 2006, **24**, 549.

Hallett 2011 J. P. Hallett and T. Welton, *Chem Rev.*, 2011, **111**, 3508.

Hammett 1932 L. P. Hammett and A. J . Deyrup, *J. Am. Chem. Soc.*, 1932, **54**, 2721.

Hammett 1933 L. P. Hammett and H. L. Pfluger, *J. Am. Chem. Soc.*, 1933, **55**, 4079.

Hammett 1937 L. P. Hammett, *J. Am. Chem. Soc.*, 1937, **59**, 96.

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

Hansen 2007 C. M. Hansen in *Hansen Solubility Parameters: A User's Handbook*, second edition, CRC Press, Boca Raton, 2007.

Hayashi 2013 F. Hayashi and M. Iwamoto, *ACS Catal.*, 2013, **3**, 14.

Hazra 2002 A. Hazra, D. Dollimore and K. Alexander, *Thermochim. Acta*, 2002, **392–393**, 221.

Henderson 2011 R. K. Henderson, C. Jiménez-González, D. J. C. Constable, S. R. Alston, G. G. A. Inglis, G. Fisher, J. Sherwood, S. P. Binks and A. D. Curzons, *Green Chem.*, 2011, **13**, 854.

Henneberry 2004 M. Henneberry, J. A. Snively, G. J. Vasek and R. Datta, US pat., 6797684, 2004.

Higuchi 1963 T. Higuchi, T. Miki, A. C. Shah and A. K. Herd, *J. Am. Chem. Soc.*, 1963, **85**, 3655.

Hixson 1918 A. W. Hixson and R. H. McKee, *J. Ind. Eng. Chem.*, 1918, **10**, 982.

Hoang 2009 T. Q. Hoang, X. Zhu, T. Danuthai, L. L. Lobban, D. E. Resasco and R. G. Mallinson, *Energy Fuels*, 2010, **24**, 3804.

Hoefnagel 1989 A. J. Hoefnagel and B. M. Wepster, *J. Chem. Soc., Perkin Trans. II*, 1989, 977.

Hoff 1979 M. C. Hoff in *Kirk–Othmer Encyclopedia of Chemical Technology*, volume 23, edited by M. Grayson, John Wiley and Sons, New York, 1979.

Hölfe 1987 G. Hölfe, W. Steglich and H. Vorbrüggen, *Angew. Chem. Ed. Engl.*, 1987, **17**, 569.

Huber 2005 G. W. Huber, J. N. Chheda, C. J. Barrett and J. A. Dumesic, *Science*, 2005, **308**, 1446.

Huber 2013 G. W. Huber, Y-. T. Cheng, T. Carlson, T. Vispute, J. Hae and G. Tompsett, US Pat., 0023706, 2013.

Hughes 1935 E. D. Hughes and C. K. Ingold *J. Chem. Soc.,* 1935, 244.

Hukkerikar 2012 A. S. Hukkerikar, B. Sarup, A. Ten Kate, J. Abildskov, G. Sin and R. Gani, *Fluid Phase Equilibr.*, 2012, **321**, 25.

Hunt 2010 A. J. Hunt, E. H. K. Sin, R. Marriot and J. H. Clark, *ChemSusChem*, 2010, **3**, 306.

Huntress 1941 E. H. Huntress and J. S. Autenrieth, *J. Am. Chem. Soc.*, 1941, **63**, 3446.

Hussain 2012 S. Hussain, D. Talukdar, S. K. Bharadwaj and M. K. Chaudhuri, *Tetrahedron Lett.*, 2012, **53**, 6512.

India Glycols 2013 India Glycols Limited, product information. Available online at www.indiaglycols.com/product_groups/monoethylene_glycol.htm. Accessed 14th June 2013.

Ingold 1969 C. K. Ingold in *Structure and Mechanism in Organic Chemistry*, second edition, Cornell University Press, New York, 1969.

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

ISO 2006a International Organisation for Standardization, international standard, *Environmental Management Life Cycle Assessment Principles and Framework*, ISO 14040:2006, 2006.

ISO 2006b International Organisation for Standardization, international standard, *Environmental Management Life Cycle Assessment Requirements and Guidlines*, ISO 14044:2006, 2006.

Iwamoto 2013 M. Iwamoto, S. Mizuno and M. Tanaka, *Chem. Euro. J.*, 2013, **19**, 7214.

James 2012	S. L. James, C. J. Adams, C. Bolm, D. Braga, P. Collier, T. Friščić, 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.
Jessop 2007	P. G. Jessop and B. Subramamiam, *Chem Rev.*, 2007, **107**, 2666.
Jessop 2011	P. G. Jessop, *Green Chem.*, 2011, **13**, 1391.
Jessop 2012	P. G. Jessop, D. A. Jessop, D. Fu and L. Phan, *Green Chem.*, 2012, **14**, 1245.
Jiménez-González 2011	C. Jiménez-González, D. J. C. Constable, C. S. Ponder, Q. B. Broxterman and J. B. Manley, *Org. Proc. Res. Dev.,* 2011, **15**, 912.
Jiménez-González 2012	C. Jiménez-González, D. J. C. Constable and C. S. Ponder, *Chem. Soc. Rev.,* 2012, **41**, 1485.
Jin 2002	T. Jin, S. Zhang, T. Li, *Synthetic Commun.*, 2002, **32**, 1847.
Kamlet 1976	M. J. Kamlet and R. W. Taft, *J. Am. Chem. Soc.*, 1976, **98**, 377.
Kamlet 1977	M. J. Kamlet, J. L. Abboud and R. W. Taft, *J. Am. Chem. Soc.*, 1977, **99**, 6027.
Kamlet 1979	M. J. Kamlet, T. N. Hall, J. Boykin and R. W. Taft, *J. Org. Chem.*, 1979, **44**, 2599.
Kamlet 1983	M. J. Kamlet, J. L. Abboud, M. H. Abraham and R. W. Taft, *J. Org. Chem.*, 1983, **48**, 2877.
Kaneda 2006	K. Kaneda, K. Ebitani, T. Mizugaki and K. Mori, *Bull. Chem. Soc. Jpn.*, 2006, **79**, 981.
Kappe 1993	C. O. Kappe, *Tetrahedron*, 1993, **49**, 6937.
Kappe 1997	C. O. Kappe, *J. Org. Chem.*, 1997, **62**, 7201.
Kappe 2000	C. O. Kappe, *Acc. Chem. Res.*, 2000, **33**, 879.
Karakurt 2001	A. Karakurt, S. Dalkara, M. Özalp, S. Özbey, E. Kendi and J. P. Stables, *Eur. J. Med. Chem.*, 2001, **36**, 421.
Kathiravan 2012	M. K. Kathiravan, A. B. Salake, A. S. Chothe, P. B. Dudhe, R. P. Watode, M. S Mukta and S. Gadhwe, *Bioorgan. Med. Chem.*, 2012, **20**, 5678.

Kawagoe 2010	Y. Kawagoe, M. Fujiki and Y. Nakano, *New J. Chem.*, 2010, **34**, 637.
Kelly 1997	T. R. Kelly, Y. –J. Lee and R. J. Mears, *J. Org. Chem.*, 1997, **62**, 2774.
Kertesz 1994	M. A. Kertesz, P. Kölbener, H. Stockinger, S. Beil and A. M. Cook, *Appl. Environ. Microb.*, 1994, **60**, 2296.
Kerton 2009	F. M. Kerton in *Alternative Solvents for Green Chemistry*, RSC, Cambridge, 2009.
Khadzhiev 2008	S. N. Khadzhiev, N. V. Kolesnichenko and N. N. Ezhova, *Petrol. Chem.*, 2008, **48**, 325.
Kingston 1969	B. H. M. Kingston, J. J. Garey and W. B. Hellwig, *Anal. Chem.*, 1969, **41**, 86.
Kirumakki 2004	S. R. Kirumakki, N. Nagaraju and S. Narayanan, *Appl. Catal. A: Gen.*, 2004, **273**, 1.
Kleemann 2001	A. Kleemann and J. Engel, B. Kutscher and D. Reichert in *Encyclopedia of Pharmaceutical Substances*, fourth edition, Georg Thieme Verlag, Stuttgart, 2001.
Knaggs 1979	E. A. Knaggs, M. L. Nussbaum and A. Shultz in *Kirk–Othmer Encyclopedia of Chemical Technology*, volume 22, edited by M. Grayson, John Wiley and Sons, New York, 1979.
Kolocouris 1994	N. Kolocouris, G. B. Foscolos, A. Kolocouris, P. Marakos, N. Pouli, G. Fytas, S. Ikeda and E. De Clercq, *J. Med. Chem.*, 1994, **37**, 2896.
Kolosov 2009	M. A. Kolosov, V. D. Orlov, D. A. Beloborodov and V. V. Dotsenko, *Mol. Divers.*, 2009, **13**, 5.
Kondo 1982	Y. Kondo, M. Ittoh, and S. Kusabayashi, *J. Chem. Soc., Faraday Trans. I,* 1982, **78**, 2793.
Kondo 1984	Y. Kondo, M. Ogasa and S. Kusabayashi, *J. Chem. Soc. Perkin Trans. II*, 1984, 2093.
Konkala 2012	K. Konkala, N. M. Sabbavarapu, R. Katla, N. Y. V. Durga, V. K. T. Reddy, B. L. A. P. Devi and R. B. N. Prasad, *Tetrahedron Lett.*, 2012, **53**, 1968.
Kosower 1958	E. M. Kosower, *J. Am. Chem. Soc.*, 1958, **80**, 3253.

Kraljević 2010	T. G. Kraljević, S. Krištafor, L. Šuman, M. Kralj, S. M. Ametamey, M. Cetina and S. Raić-Malić, *Bioorgan. Med. Chem.*, 2010, **18**, 2704.
Kriegel 2010	R. Kriegel, X. Huang, M. W. Schultheis, D. A. Bippert, G. E. Insolia, B. Kolls and S. Summerville, US pat., 0028512, 2010.
Kulkarni 2009	A. Kulkarni, M. Abid, B. Török and X. Huang, *Tetrahedron Letters*, 2009, **50**, 1791.
Labadi 1993	I. Labadi, L. Parkanyi, G. Kenessey and G. Liptay, *J. Cryst. Spectrosc.* 1993, **23**, 333.
Lailach 1988	G. Lailach, R. Gerken, K. –H. Schultz, R. Hornung, W. Böckmann, W. Larbig and W. Dietz, US Pat., 4772757, 1988.
Lamarche 2001	O. Lamarche, J. A. Platts and A. Hersey, *Phys. Chem. Chem. Phys.*, 2001, **3**, 2747.
Lammens 2010	T. M. Lammens, M. C. R. Franssen, E. L. Scott and J. P. M. Sanders, *Green Chem.*, 2010, **12**, 1430.
Lammens 2011	T. M. Lammens, J. Potting, J. P. M. Sanders and I. J. M. De Boer, *Environ. Sci. Technol*., 2011, **45**, 8521.
Lancaster 2002	M. Lancaster in *Green Chemistry: An Introductory Text*, RSC, Cambridge, 2002.
Laureles 2002	L. R. Laureles, F. M. Rodriguez, C. E. Reaño, G. A. Santos, A. C. Laurena and E. M. T. Mendoza, *J. Agric. Food Chem.*, 2002, **50**, 1581.
Laurence 1994	C. Laurence, P. Nicolet, M. T. Dalati, J. –L. M. Abboud and R. Notario, *J. Phys. Chem.*, 1994, **98**, 5807.
Lee 2004	K. –Y. Lee, K. –Y. Ko, *B. Kor. Chem. Soc.* 2004, **25**, 1929.
van Leeuwen 2012	B. N. M. van Leeuwen, A. M. van der Wulp, I. Duijnstee, A. J. A. van Maris and A. J. J. Straathof, *Appl. Microbiol. Biotechnol.*, 2011, **93**, 1377.
Leffler 1955	J. E. Leffler, *J. Org. Chem.*, 1955, **20**, 1202.
Lente 2005	G. Lente, I. Fábián and A. J. Poë, *New J. Chem.*, 2005, **29**, 759.
Lesage 1996	P. Lesage, J. P. Candy, C. Hirigoyen, F. Humblot and J. M. Basset, *J. Mol. Catal. A: Chem.*, 1996, **112**, 431.
Liang 2008	X. Liang, S. Gao, G. Gong, Y. Wang and J. –G. Yang, *Catal. Lett.*, 2008, **124**, 352.

Lide 1991	*CRC Handbook of Chemistry and Physics*, seventy second edition, edited by D. R. Lide, CRC press, Boca Raton, 1991.
Littmann 1942	E. R. Littmann, US pat., 2400012, 1942.
Liu 2001	L. Liu and Q. -X Guo, *Chem. Rev.*, 2001, **101**, 673.
Liu 2006	Y. Liu, E. Lotero and J. G. Goodwin, Jr., *J. Catal.*, 2006, **242**, 278.
Liu 2010	K. Liu and W. –D. Woggon, *Eur. J. Org. Chem*., 2010, 1033.
Lo 1990	C. –C. Lo and P. –M. Chao, *J. Chem. Ecol.*, 1990, **16**, 3245.
Lohrasbi 2010	M. Lohrasbi, M. Pourbafrani, C. Niklasson and M. J. Taherzadeh, *Bioresour. Technol.*, 2010, **101**, 7382.
Logan 1996	S. R. Logan in *Fundamentals of Chemical Kinetics*, Longman Group Ltd., Harlow, 1996.
Losada 2008	M. Losada, H. Tran and Y. Xu, *J. Chem. Phys.*, 2008, **128**, 014508.
Lowenheim 1975	F. A. Lowenheim and M. K. Moran in *Industrial Chemicals*, fourth edition, John Wiley and Sons, New York, 1975.
Luque 2009	R. Luque, V. Budarin, J. H. Clark and D. J. Macquarrie, *Green Chem.*, 2009, **11**, 459.
Lüttringhaus 1964	A. Lüttringhaus and H. W. Dirksen, *Angew. Chem. Int. Ed.*, 1964, **3**, 260.
MacDonald 2010	T. W. J. Cooper, I. B. Campbell and S. J. F. MacDonald, *Angew. Chem. Int. Ed.*, 2010, **49**, 8082.
Mack 2009	D. J. Mack, M. Brichacek, A. Plichta and J. T. Njarđarson in *Top 200 Pharmaceutical Products by Worldwide Sales in 2009*, 2009. Available online at cbc.arizona.edu/ njardarson/group/sites/default/files/Top200PharmaceuticalProductsByWorldwideSalesin2009.pdf. Accessed 14th June 2013.
Mack 2012	J. Mack and S. Muthukrishnan in *Green Techniques for Organic Synthesis and Medicinal Chemistry*, edited by W. Zhang and B. W. Cue, John Wiley and Sons Ltd., Chichester, 2012.
MacMillan 2012	D. S. MacMillan, J. Murray, H. F. Sneddon, C. Jamieson and A. J. B. Watson, *Green Chem.*, 2012, **14**, 3016.
MacMillan 2013	D. S. MacMillan, J. Murray, H. F. Sneddon, C. Jamieson and A. J. B. Watson, *Green Chem.*, 2013, **15**, 596.

Maegawa 2007	T. Maegawa, Y. Kitamura, S. Sako, T. Udzu, A. Sakurai, A. Tanaka, Y. Kobayashi, K. Endo, U. Bora, T. Kurita, A. Kozaki, Y. Monguchi and H. Sajiki, *Chem. Eur. J.*, 2007, **13**, 5937.
Makgwane 2010	P. R. Makgwane, N. I. Harmse, E. E. Ferg and B. Zeelie, *Chem. Eng. J.*, 2010, **162**, 341.
Mancini 1986	R. D. Martinez, P. M. E. Mancini, L. R. Vottero and N. S. Nudelman, *J. Chem. Soc. Perkin Trans. II*, 1986, 1427.
Marchais-Oberw. 2011	S. Marchais-Oberwinkler, M. Wetzel, E. Ziegler, P. Kruchten, R. Werth, C. Henn, R. W. Hartmann and M. Frotscher, *J. Med. Chem.*, 2011, **54**, 534.
Marcus 1991	Y. Marcus, *J. Solution Chem.*, 1991, **20**, 929.
Marcus 1993	Y. Marcus, *Chem. Soc. Rev.*, 1993, **22**, 409.
Marcus 1994	Y. Marcus, *J. Chem. Soc., Perkin Trans II*, 1994, 1015.
Marcus 2005	Y. Marcus, *J. Phys. Org. Chem.*, 2005, **18**, 373.
Martin 2002	K. Martin in *Handbook of Green Chemistry and Technology*, Blackwell Science Ltd., Oxford, 2002.
Martin-Luengo 2008	M. A. Martin-Luengo, M. Yates, M. J. Martínez Domingo, B. Casal, M. Iglesias, M. Esteban and E. Ruiz-Hitzky, *Appl. Catal. B: Environ.*, 2008, **81**, 218.
Martin-Luengo 2010	M. A. Martin-Luengo, M. Yates, E. Saez-Rojo, D. Huerta-Arribas, D. Aguilar and E. Ruiz-Hitzky, *Appl. Catal. A: Gen.*, 2010, **387**, 141.
Marlière 2010	P. Marlière, international pat., 001078, 2010.
Marlière 2011a	P. Marlière, European pat., 2295593, 2011.
Marlière 2011b	P. Marlière, international pat., 076691, 2011.
Mathers 2006	R. T. Mathers, K. C. McMahon, K. Damodaran, C. J. Retarides and D. J. Kelley, *Macromolecules*, 2006, **39**, 8982.
Matlack 2001	A. S. Matlack in *Introduction to Green Chemistry*, Marcel Dekker, New York, 2001.
McConvey 2012	I. F. McConvey, D. Woods, M. Lewis, Q. Gan and P. Nancarrow, *Org. Process Res. Dev.*, 2012, **16**, 612.

McGonagle 2013	F. I. McGonagle, D. S. MacMillan, J. Murray, H. F. Sneddon, C. Jamieson and A. J. B. Watson, *Green Chem.*, 2013, **15**, 1159.
Menschutkin 1887	N. Menschutkin, *Z. Phys. Chem.*, 1887, **1**, 611.
Menschutkin 1890a	N. Menschutkin, *Z. Phys. Chem.*, 1890, **5**, 589.
Menschutkin 1890b	N. Menschutkin, *Z. Phys. Chem.*, 1890, **6**, 41.
Menschutkin 1900	N. Menschutkin, *Z. Phys. Chem.*, 1900, **34**, 157.
Mihara 2010	M. Mihara, T. Nakai, T. Iwai, T. Ito, T. Ohno and T. Mizuno, *Synlett*, 2010, **21**, 253.
Mikolajczyk 1981	M. Mikolajczyk and P. Kielbasinski, *Tetrahedron*, 1981, **37**, 233.
Mills 1985	S. G. Mills and P. J. Beak, *J. Org. Chem.*, 1985, **50**, 1216.
Mistry 2008	R. Mistry in *Characterisation and Application of CO_2-Expanded Solvents*, PhD thesis, University of Leicester, 2008.
Mizuno 2012	S. Mizuno, M. Kurosawa, M. Tanaka and M. Iwamoto, *Chem. Lett.*, 2012, **41**, 892.
Moity 2012	L. Moity, M. Durand, A. Benazzouz, C. Pierlot, V. Molinier and J. –M. Aubry, *Green Chem.*, 2012, **14**, 1132.
Montalbetti 2005	C. A. G. N. Montalbetti and V. Falque, *Tetrahedron*, 2005, **61**, 10827.
Moog 2004	R. S. Moog, D. D. Kim, J. J. Oberle and S. G. Ostrowski, *J. Phys. Chem. A*, 2004, **108**, 9294.
Moore 1981	J. W. Moore and R. G. Pearson in *Kinetics and Mechanism*, third edition, John Wiley and Sons, Hoboken, 2007.
Moriyasu 1986	M. Moriyasu, A. Kato and Y. Hashimoto, *J. Chem. Soc. Perkin Trans. II*, 1986, 515.
Morschbacker 2009	A. J. Morschbacker, *Macromol. Sci. Pol. R.*, 2009, **49**, 79.
MSDS 2013	Material safety datasheets obtained from Sigma-Aldrich®. Available online at www.sigmaaldrich.com/united-kingdom.html. Accessed 14th June 2013.
Müller 1972	P. Müller and B. Siegfried, *Helvitica Chimica Acta*, 1972, **55**, 2400.
Mucsi 2008	Z. Musci, G. A. Chass and I. G. Csizmadia, *J. Phys. Chem. B*, 2008, **112**, 7885.
Muller 1994	P. Muller, *Pure Appl. Chem*., 1994, **66**, 1077.

Myriant 2013	Myriant, press release, *Myriant Corporation and Johnson Matthey - Davy Technologies Produce Market Grade, Bio-Based Butanediol with Superior Carbon Efficiency and Reduced Carbon Footprint at a Competitive Cost Level*, 2013. Available online at www.myriant.com/media/press-releases/myriant-corporation-johnson-matthey-davy-technologies-produce-bio-based-butanediol-at-competitive-cost-level.cfm. Accessed 14th June 2013.
Narender 2000	N. Narender, P. Srinivasu, S. J. Kulkarni and K. V. Raghavan, *Green Chem.*, 2000, **2**, 104.
Nezhad 2003	A. K. Nezhad, B. Mokhtari and M. N. S. Rad, *Tetrahedron Lett.*, 2003, **44**, 7325.
Nicolet 1986	P. Nicolet and C. Laurence, *J. Chem. Soc. Perkin Trans. II*, 1986, 1071.
Nishio 2011	M. Nishio, *Phys. Chem. Chem. Phys.*, 2011, **13**, 13873.
Núñez-Magro 2007	A. A. Núñez-Magro, G. R. Eastman and D. J. Cole-Hamilton, *Chem. Commun.*, 2007, 3154.
Nylander 1970	L. R. Nylander and S. F. Pavkovic, *Inorg. Chem.*, 1970, **9**, 1959.
Oakes 2001	R. S. Oakes, A. A. Clifford and C. M. Rayner, *J. Chem. Soc. Perkin Trans. I*, 2001, 917.
Okino 2008	S. Okino, R. Noburyu, M. Suda, T. Jojima, M. Inui and H. Yukawa, *Appl. Microbiol. Biotechnol.*, 2008, **81**, 459.
Orita 2001	A. Orita, C. Tanahashi, A. Kakuda and J. Otera, *J. Org. Chem.*, 2001, **66**, 8926.
Otera 2001	J. Otera, *Angew. Chem. Int. Ed.*, 2001, **40**, 2044.
Otera 2002	J. Xiang, A. Orita and J. Otera, *Angew. Chem. Int. Ed.*, 2002, **41**, 4117.
Otera 2005	Z. Peng, A. Orita, D. Ana and J. Otera, *Tetrahedron Lett.*, 2005, **46**, 3187.
Pace 2012	V. Pace, P. Hoyes, L. Castoldi, P. D. de Maria and A. R. Alcantara, *ChemSusChem*, 2012, **5**, 1369.
Pearlson 1986	W. H. Pearlson, *J. Fluorine Chem.*, 1986, **32**, 29.
Pearson 1999	A. J. Pearson and W. R. Roush in *Handbook of Reagents for Organic Synthesis: Activating Agents and Protecting Groups*, Wiley, New York, 1999.

Pereira 2011	C. S. M. Pereira, V. M. T. M. Silva and A. E. Rodrigues, *Green Chem.*, 2011, **13**, 2658.
Perez-Benito 2013	J. F. Perez-Benito, *Monatsh. Chem.*, 2013, **144**, 49.
Pérez-Sánchez 2012	M. Pérez-Sánchez, M. Sandoval and M. J. Hernáiz, *Tetrahedron*, 2012, **68**, 2141.
Perreux 2002	L. Perreux, A. Loupy and F. Volatron, *Tetrahedron*, 2002, **58**, 2155.
Peters 2010	G. P. Peters, *Current Opinion in Environmental Sustainability*, 2010, **2**, 245.
Peters 2011	M. W. Peters, J. D. Taylor, M. Jenni, L. E. Manzer and D. E. Henton, US pat., 0087000, 2011.
Petersen 1964	R. C. Petersen, *J. Org. Chem*., 1964, **29**, 3133.
Phillips 1920	M. Phillips and H. B. Gibbs, *J. Ind. Eng. Chem.*, 1920, **12**, 733.
Phillips 1924	M. Phillips, *J. Am. Chem. Soc*., 1924, **46**, 686.
Pieber 2013	B. Pieber and C. O. Kappe, *Green Chem.*, 2013, **15**, 320.
Pilling 1995	M. J. Pilling and P. W. Seakins in *Reaction Kinetics*, Oxford Science Publications, Oxford, 1995.
Platts 2000a	J. A. Platts, *Phys. Chem. Chem. Phys.*, 2000, **2**, 973.
Platts 2000b	J. A. Platts, *Phys. Chem. Chem. Phys.*, 2000, **2**, 3115.
Quan 2009	Z. –J. Quan, Y. –X. Da, Z. Zhang and X. –C. Wang, *Catal. Commun.* 2009, **10**, 1146.
Qureshi 2001	N. Qureshi and H. P. Blaschek, *J. Ind. Microbiol. Biotechnol. Biot.*, 2001, **27**, 292.
Raj 2011	M. K. Raj, H. S. P. Rao, S. G. Manjunatha, R. Sridharan, S. Nambiar, J. Keshwan, J. Rappai, S. Bhagat, B. S. Shwetha, D. Hegde and U. Santhosh, *Tetrahedron Lett.*, 2011, **52**, 3605.
Rajendran 2012	A. Rajendran, *J. Chromatogr. A*, 2012, **1250**, 227.
Ab Rani 2011	M. A. Ab Rani, A. Brant, L. Crowhurst, A. Dolan, M. Lui, N. H. Hassan, J. P. Hallett, P. A. Hunt, H. Niedermeyer, J. M. Perez-Arlandis, M. Schrems, T. Welton and R. Wildinga, *Phys. Chem. Chem. Phys.*, 2011, **13**, 16831.

Ranieri 2008	G. Ranieri, J. P. Hallett and T. Welton, *Ind. Eng. Chem. Res.*, 2008, **47**, 638.
Reichardt 1979	C. Reichardt, *Angew. Chem. Int. Ed.*, 1979, **18**, 98.
Reichardt 1994	C. Reichardt, *Chem. Rev.*, 1994, **94**, 2319.
Reichardt 2003	C. Reichardt in *Solvents and Solvent Effects in Organic Chemistry*, third edition, Wiley-VCH, Weinheim, 2003.
Reicheneder 2004	S. Reicheneder and C. Unverzagt, *Angew. Chem. Int. Ed.*, 2004, **43**, 4353.
Rhodia 2013	Rhodia, press release, *AknoNobel and Solvay Establish a Partnership on Renewable Solvents*, 2013. Available online at www.rhodia.com.br/en/news_center/news_releases/Akzonobel_e_Solvay.tcm. Accessed 21st June 2013.
Risatti 2013	C. Risatti, K. J. Natalie, Z. Shi and D. A. Conlon, *Org. Process Res. Dev.*, 2013, **17**, 257.
Ritter 2004	T. Ritter, P. Zarotti and E. Carreira, *Org. Lett.*, 2004, **6**, 4371.
Roughley 2011	S. D. Roughley and A. M. Jordan, *J. Med. Chem.*, 2011, **54**, 3451.
Russowsky 2004	D. Russowsky, F. A. Lopes, V. S. S. da Silva, K. F. S. Canto, M. G. Montes D'Oca and M. N. Godoi, *J. Brazil. Chem. Soc.*, 2004, **15**, 165.
Sagadeev 2006	E. V. Sagadeev and V. V. Sagadeev, *High Temp.*, 2006, **44**, 530.
Saunders 2007	J. Saunders, E. Scott, R. Weusthuis and H. Mooibroek, *Macromol. Biosci.*, 2007, **7**, 105.
Sawant 2007	D. P. Sawant, A. Vinu, J. Justus, P. Srinivasu and S. B. Halligudi, *J. Mol. Catal. A: Chem.*, 2007, **276**, 150.
Schleicher 2009	J. C. Schleicher and A. M. Scurto, *Green Chem.*, 2009, **11**, 694.
Schultz 2006	T.W. Schultz, J. W. Yarbrough and S. K. Koss, *Cell Biol. Toxicol.*, 2006, **22**, 339.
Scifinder 2013a	Scifinder reaction database search, based on 4-bromotoluene and phenyl boronic acid reacting to form 4-phenyltoluene. Conducted on 27th May 2013.

Scifinder 2013b	As of 30th November 2011 the Scifinder® database cited 1,148,722 instances in which *p*-TSA or its monohydrate salt have been applied as a catalyst. By 25th March 2013 1,489,653 entries were listed in Scifinder®, a 30% increase. By comparison sulfuric acid has been recorded as a catalyst in only 353,595 different reactions. The Reaxys® database (more realistically) contained 966,786 reactions catalysed by *p*-toluenesulphonic acid and 854,058 reactions catalysed by sulphuric acid. Conducted on 25th March 2013.
Schorger 1918	A. W. Schorger, *J. Ind. Eng. Chem.*, 1918, **10**, 258.
Schüttler 1972	A. Schüttler, U. Kraatz, H. Wamhoff and F. Korte, *Tetrahedron*, 1972, **29**, 4871.
Seresh 2009	Seresh, A. Saini, D. Kumar and J. S. Sandhu, *Green Chemistry Letters and Reviews*, 2009, **2**, 29.
Seresh 2012	Suresh and J. S. Sandhu, *ARKIVOC*, 2012, **1**, 66.
Shaikh 2011	T. S. Shaikh, K. A. Undale, D. S. Gaikwad and D. M. Pore, *C. R. Chim.* 2011, **14**, 987.
Shanmugam 2004	S. Shanmugam, B. Viswanathan and T. K. Varadarajan, *J. Mol. Catal. A: Chem.*, 2004, **223**, 143.
Sharma 2007	S. D. Sharma, P. Gogoi and D. Konwar, *Green Chem.*, 2007, **9**, 153.
Sheehan 1955	J. C. Sheehan and P. G. Hess, *J. Am. Chem. Soc.*, 1955, **77**, 1067.
Sheldon 2000	R. A. Sheldon, *Pure Appl. Chem.*, 2000, **72**, 1233.
Sheldon 2005	R. A. Sheldon, *Green Chem.*, 2005, **7**, 267.
Shen 2010	Z. –L. Shen, X. –P. Xu and S. –J. Ji, *J. Org. Chem.* 2010, **75**, 1162.
Shinohara 1973	Y. Shinohara and T. Isaka, US pat., 3720716, 1973.
Shultz 1979	A. Shultz in *Kirk–Othmer Encyclopedia of Chemical Technology*, volume 22, edited by M. Grayson, John Wiley and Sons, New York, 1979.
Simons 1944	J. H. Simons and H. Hart, *J. Am. Chem. Soc.*, 1944, **66**, 1309.
Skrzypczak 2004	A. Skrzypczak and P. Neta, *Int. J. Chem. Kinet.*, 2004, **36**, 253.

Smith 2007	M. B. Smith and J. March in *March's Advanced Organic Chemistry*, sixth edition, John Wiley and Sons, Hoboken, 2007.
Solà 1991	M. Solà, A. Lledós, M. Duran, J. Bertrán and J. –L. M. Abboud, *J. Am. Chem. Soc.*, 1991, **113**, 2873.
De Souza 2009	R. O. M. A. De Souza, E. T. da Penha, H. M. S. Milagre, S. J. Garden, P. M. Esteves, M. N. Eberlin and O. A. C. Antunes, *Chem. Eur. J.*, 2009, **15**, 9799.
Spica 1881	P. Spica, *Ber.*, 1881, **14**, 653.
Srivastava 2010	V. Srivastava and M. M. Singh, *J. Appl. Electrochem.*, 2010, **40**, 2135.
Starke 1963	K. Starke, *J. Inorg. Nucl. Chem.*, 1963, **25**, 823.
Strohmeyer 1971	M. Strohmeyer, US pat., 3555103, 1971.
SubsPort 2013	SubsPort, restricted and priority substances database. Available online at www.subsport.eu/listoflists. Accessed 14th June 2013.
Sykes 1981	P. Sykes in *Linear Free Energy Relationships*, RSC Chemistry Cassettes, London, 1981. Available online at www.rsc.org/Membership/Networking/InterestGroups/EducationalTechniques/ChemistryCassettes/energyRelations.asp. Accessed 14th June 2013.
Taft 1976	R. W. Taft and M. J. Kamlet, *J. Am. Chem. Soc.*, 1976, **98**, 2886.
Taft 1981	R. W. Taft and M. J. Kamlet, *J. Am. Chem. Soc.*, 1981, **103**, 1080.
Taft 1983	R. W. Taft and M. J. Kamlet, *Inorg. Chem.*, 1983, **22**, 250.
Taft 1985	R. W. Taft, J. –L. M. Abboud, M. J. Kamlet and M. H. Abraham, *J. Solution Chem.*, 1985, **14**, 153.
Tanabe 1999	K. Tanabe and W. F. Hölderich, *Appl. Catal. A: Gen.*, 1999, **181**, 399.
Tanaka 2000	K. Tanaka and F. Toda, *Chem. Rev.*, 2000, **100**, 1025.
Tang 2005	P. Tang, *Organic Syntheses*, 2005, **81**, 262.
Tao 2011	D. –J. Tao, Y. –T. Wu, Z. Zhou, J. Geng, X. –B. Hu and Z. –B. Zhang, *Ind. Eng. Chem. Res.*, 2011, **50**, 1989.
Taygerly 2012	J. P. Taygerly, L. M. Miller, A. Yee and E. A. Peterson, *Green Chem.*, 2012, **14**, 3020.
Thomas 1989	A. F. Thomas and Y. Bessière, *Nat. Prod. Rep.*, 1989, **6**, 291.

Tilstam 2012	U. Tilstam, *Org. Process Res. Dev.*, 2012, **16**, 1273.
Toray 2011	Toray, press release, *Toray Succeeds in Production of the World's First Fully Renewable, Bio-Based Polyethylene Terephthalate (PET) Fibre*, 2011. Available online at www.toray.com/news/fiber/nr111115.html. Accessed 14th June 2013.
Tron 2011	G. C. Tron, A. Minassi and G. Appendino, *Eur. J. Org. Chem.*, 2011, 5541.
Turley 2008	D. B. Turley in *Introduction to Chemicals from Biomass*, edited by J. H. Clark and F. Deswarte, John Wiley and Sons, Chichester, 2008.
Ulijn 2002	R. V. Ulijn, B. D. Moore, A. E. M. Janssen and P. J. Halling, *J. Chem. Soc.*, *Perkin Trans. II*, 2002, 1024.
UN 1987	United Nations General Assembly in *Report of the World Commission on Environment and Development: Our Common Future*, *Annex to document A/42/427 - Development and International Co-operation: Environment* 1987. Available online at www.un-documents.net/wced-ocf.htm. Accessed 14th June 2013.
UOP 2013	UOP/Eni 'Ecofining'™ process schematic. Available online at www.uop.com/hydroprocessing-ecofining. Accessed 14th June 2013.
Vaidyanathan 2004	R. Vaidyanathan, V. G. Kalthod, D. P. Ngo, J. M. Manley and S. P. Lapekas, *J. Org. Chem.*, 2004, **69**, 2565.
Varma 1999	R. S. Varma, *Green Chem.*, 1999, **1**, 43.
Valeur 2009	E. Valeur and M. Bradley, *Chem. Soc. Rev.*, 2009, **38**, 606.
Veillet 2010	S. Veillet, V. Tomao, K. Ruiz and F. Chemat, *Anal. Chim. Acta*, 2010, **674**, 49.
Verma 1998	R. Verma and S. K. Ghosh, *J. Chem. Soc., Perkin Trans. I*, 1998, 2377.
Wan 2010	J. -P. Wan and Y. Liu, *Synthesis*, 2010, 3943.
Weissermel 1997	K. Weissermel and H. –J. Arpe in *Industrial Organic Chemistry*, third edition, Wiley VCH, Weinheim, 1997.
Wells 2008	T. P. Wells, J. P. Hallett, C. K. Williams and T. Welton, *J. Org. Chem.*, 2008, **73**, 5585.
Welton 1999	T. Welton, *Chem. Rev.*, 1999, **99**, 2071.

Welton 2006 T. Welton, *Green Chem.*, 2006, **8**, 13.

Williams 2003 A. Williams in *Free Energy Relationships in Organic and Bio-Organic Chemistry*, RSC, Cambridge, 2003.

Wise 2007 N. J. Wise and J. M. J. Williams, *Tetrahedron Lett.*, 2007, **48**, 3639.

Wolfson 2009 A. Wolfson, G. Litvak, C. Dlugy, Y. Shotland and D. Tavor, *Ind. Crop. Prod.*, 2009, **30**, 78.

Woolsey 2013 R. J. Woolsey in *Destroying Oil's Monopoly and OPEC's Cartel*, Geopolitics of Energy, US Energy Security Council, 2012. Available online at www.usesc.org/ energy_security/index.php/resources. Accessed 14th June 2013.

Wuts 2007 P. G. M. Wuts and T. A. W. Greene in *Greene's Protective Groups in Organic Synthesis*, fourth edition, John Wiley and Sons, Hoboken, 2007,

Xu 2008 Q. Xu, Z. Yang, D. Yin and F. Zhang, *Catal. Commun.*, 2008, **9**, 1579.

Yadav 2004 G. D. Yadav, *Top. Catal.*, 2004, **29**, 145.

Yamada 2008 Y. Yamada and P. Yukphan, *Int. J. Food Microbiol.*, 2008, **125**, 15.

Yang 2012a W. –C. Yang, J. Li, J. Li, Q. Chen and G. –F. Yang, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 1455.

Yang 2012b J. Yang, J. –N. Tan and Y. Gu, *Green Chem.*, 2012, **14**, 3304.

Yarım 2003 M. Yarım, S. Saraç, F. S. Kılıç and K. Erol, *Il Farmaco*, 2003, **58**, 17.

Zeynizadeh 2009 B. Zeynizadeh, K. A. Dilmaghani and M. Yari, *Phosphorus Sulfur*, 2009, **184**, 2465.

Zhang 2008 Y. Zhang, B. R. Bakshi and E. S. Demessie, *Environ. Sci. Technol.*, 2008, **42**, 1724.

Zhang 2012 C. Zhang, B. Lu, X. Wang, J. Zhao and Q. Cai, *Catal. Sci. Technol.*, 2012, **2**, 305.

Zhao 1999 T. Zhao and G. Mu, *Corros. Sci.*, 1999, **41**, 1937.

Zorhun 2006 I. S. Zorkun, S. Saraç, S. Çelebi and K. Erol, *Bioorgan. Med. Chem.*, 2006, **14**, 8582.

Zupp 2012 L. R. Zupp, V. L. Campenella, D. M. Rudzunski, F. Beland and R. Priefer, *Tetrahedron Lett.*, 2012, **53**, 5343.