Received 00th January 20xx,

1. Department of Chemistry, The University of York, Heslington, York, YO10 5DD, United Kingdom
2. Materials Chemistry Research Center, Department of Chemistry and Center of Excellence for Innovation in Chemistry, Faculty of Science, Khon Kaen University, Khon Kaen, 40002, Thailand, E-mail: andrew@kku.ac.th

\* Corresponding author

Electronic Supplementary Information (ESI) available: Experimental details, characterisation data and single crystal XRD structures are presented. See DOI: 10.1039/x0xx00000x

Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Development of pharmaceutically relevant bio-based intermediates though aldol condensation and Claisen-Schmidt reactions of dihydrolevoglucosenone (Cyrene®)

Liam Hughes,a Con R. McElroya Adrian C. Whitwooda and Andrew J. Huntb\*

Dihydrolevoglucosenone (Cyrene®) has been successfully utilised as a bio-based platform molecule for the **synthesis of pharmaceutically relevant intermediates through aldol condensation reactions. Utilising sustainable synthetic methodologies, the self-aldol condensations reaction of Cyrene was achieved in high purity, with isolated yields of 81.3**%. Claisen-Schmidt **reactions with a range of aromatic and heteroaromatic aldehydes yielded novel cyrene-based compounds, characterised by single-crystal X-ray diffraction, FT-IR, NMR and MS.**

Dihydrolevoglucosenone (Cyrene) has received significant attention as a bio-based solvent for the replacement of traditional petroleum derived and toxic dipolar aprotic solvents such as NMP or DMF.1 This bio-based molecule is produced through the hydrogenation of levoglucosenone, which itself is obtained *via* the acid-catalysed pyrolysis of cellulose.2 Katz *et al.* investigated the use of cyrene as a green solvent in the synthesis of metal organic frameworks (MOFs) including ZIF-8 MOF.3 Analysis of reaction media pointed to the formation of a by-product in Cyrene systems that was not present when dimethylformamide (DMF) was applied as the solvent. 1H NMR spectroscopy and single-crystal X-ray diffraction confirmed the formation of a base catalysed cyrene self-aldol condensation product.3

While investigating the use of Cyrene as a solvent for the Sonogashira reaction, Wilson *et al.* identified that Cyrene showed a particular sensitivity to inorganic bases.4 It was demonstrated that cyrene can undergo a self-aldol condensation even at room temperature. The use of Et3N at low temperatures avoided the aldol reaction occurring, leading to a 100% conversion and a 96% yield for the reaction of 1-iodo-4-fluorobenzene and phenylacetylene.4 Despite such observations, to date limited research has focussed on the use of Cyrene as a bio-based platform molecule and its use in aldol or Claisen-Schmidt reactions.

The aldol condensation reaction is a key tool for in the formation of new carbon–carbon bonds.5 Crossed-aldol reactions between cyclic compounds with a similar structure to cyrene, such as cyclohexanone and cyclopentanone have been widely reported before.5-8 In order to minimise the self-condensation reaction, aldehydes are used as they are more reactive than ketones, with benzaldehydes and derivatives being most commonly utilised. Benzaldehyde lacks α-hydrogens and thus can only act as an acceptor within the reaction, thus decreasing the number of possible products. Due to the inherent chirality of cyrene (resulting from the anhydro bridge), aldol reactions could result in the formation of a range of novel bio-derived chiral platform molecules, which could potentially be used as frameworks in pharmaceutical synthesis.

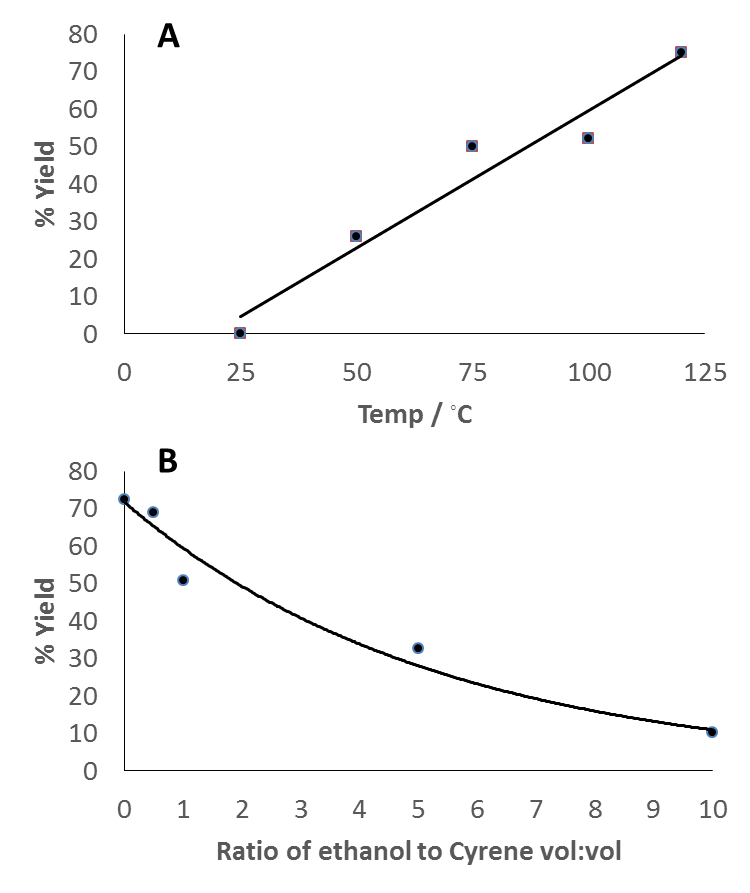
Herein, Cyrene is utilised as a bio-based platform molecule for the synthesis of pharmaceutically relevant intermediates through aldol condensation reactions. The formation of self-aldol product of cyrene was optimised and purified using a facile sustainable synthetic method. Optimisation was achieved with respect to base catalysts, temperature and solvent volume. The crossed aldol condensation reactions between Cyrene and a range of aromatic and heteroaromatic aldehydes were also catalysed by K3PO4 to synthesise eighteen novel Cyrene based compounds. Characterisation of the products was carried out by 1H and 13C NMR spectroscopy, GC-FID, GC-MS, IR spectroscopy and where possible, single-crystal X-ray diffraction.

Firstly, the synthetic route to the Cyrene self-Aldol condensation product was optimised with respect to the base catalyst, temperature, ratio of solvent and reaction time. Previous work reported that Cyrene exhibited high sensitivity to inorganic bases and identified K3PO4 as a promising catalyst.4 A number of bases were screened for formation of this product, with nearly all reporting self-aldol in reactions conducted at 50 °C or higher over a 24 hour period, but with no yields give. Therefore in order to advance the current knowledge of this reaction, it was important to fully investigate base selection (table 1).

**Table 1:** Base catalyst screening in the self-condensation reaction.



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Entry** | **Catalyst** | **pKaa** | **Isolated Yield (%)** | **GSK reagent assessment 9** |
| 1 | NEt3 | 10.75 | 0.0 | Some issues |
| 2 | Pyridine | 5.25 | 0.0 | Few issues |
| 3 | KOAc | 4.76 | 0.0 | Few issues |
| 4 | CaCO3 | 9.00 | 0.0 | N/A |
| 5 | K2CO3 | 10.25 | 32.3 | Few issues |
| 6 | KOH | 15.7 | 72.5 | Few issues |
| 7 | t-BuOK | 17 | 81.0 | Some issues |
| 8 | K3PO4 | 12.32 | 81.3 | Few issues |

a In water. Reaction conditions: Cyrene (97.6 mMol), base (4.87 mMol) and ethanol were heated to 120 ◦C for 40 minutes. N/A indicates that this base is not present in the GSK acid-base reagent guide, however, CaCO3 is likely to be a green base.

GC-FID showed no conversion after 40 minutes for the two organic bases (entries 1 and 2, table 1). Neither organic bases employed catalysed the reaction, this is particularly interesting in the case of triethylamine which has a similar pKa to potassium carbonate, which does result in the formation of the Aldol product (table 1). The reason for this result is currently unknown but is consistent with previously reported studies.4 No solid product was observed from the reaction using potassium acetate however, GC-FID indicated the initial signs of conversion (entry 3, table 1). Calcium carbonate was also unsuccessful in catalysing this reaction, although the stronger base potassium carbonate produced the lowest yield of the successful inorganic bases (entries 4 and 5, table 1). Previous studies did indicate the presence of the Cyrene self-Aldol condensation product when reacting with either KOAc or CaCO3 for 8 hours at 100 °C, however no % yields were reported.4 It is likely that the shorter reaction times of 40 minutes used in this current study were not long enough to generate any product with these weaker inorganic bases. Inorganic bases with a pKaa over 10 all led to the formation of self-Aldol condensation product after 40 minutes (table 1). The remaining bases performed comparatively well, with K3PO4 and *t*-BuOK giving essentially the same conversion although the latter reached completion within 20 minutes. In order to determine which base to carry forward, the greenness and safety of the two were considered. Applying the GSK acid-base selection guide, potassium *tert*-butoxide scores lower than potassium-phosphate in both EHS and greenness categories, it is also highly flammable and can cause severe skin burns.9

K3PO4 is a strong inorganic base (pKa = 12.32, Table 1), but, is crucially non-toxic, inexpensive, commercially available, soluble in many organic solvents and has a high decomposition temperature. K3PO4 has also been utilised as a non-nucleophilic base in several reactions.10 The choice of K3PO4 as a “green” base was based on several factors including those highlighted in GSK acid-base selection guide.9 These included having favourable scores for EHS, clean synthesis and greenness. On completion of the reaction the base can be washed out of the product using water. Importantly, the recovered base could potentially be recycled within the reaction or utilised for further downstream applications including its use a fertilisers. Therefore, although *t*-BuOK catalyses the reaction at a greater rate, K3PO4 was selected as the optimal base to take forward.

In addition to the base, it was also important to determine if a lower temperature of reaction could be applied whilst maintaining yield. The maximum temperature of reaction employed was 120 °C, as the MSDS for Cyrene advises staying below 140 °C due to potential degradation.11 As can be seen in Figure 1A, there is a strong correlation between temperature and yield meaning one cannot be lowered without adversely affecting the other.

**Figure 1.** A) Optimisation of temperature (Cyrene (97.6 mMol), K3PO4 (4.87 mMol) and ethanol (0.5 ml)) for 30 minutes and B) Optimisation of solvent ratio (Cyrene (97.6 mMol), K3PO4 (4.87 mMol) at 120 ◦C for 30 minutes.

Finally, optimisation of the ethanol co-solvent to Cyrene ratio was conducted to better understand the role of the solvent. Importantly, Cyrene is actually both the reagent and solvent in this reaction. Cyrene has been recognised as a bio-based alternative dipolar aprotic solvent, with similar in properties to NMP or DMF.1 High yields of the self-aldol product can actually be obtained when the reaction is run in the absence of ethanol (Figure 1,B). This is presumably due to the higher concentration of base, resulting in a greater number of interactions between it and Cyrene.

In the absence of any co-solvent, the product quickly degrades, resulting in the formation of a tar and char. This results in extensive solvent usage to further wash and purify the Cyrene self-Aldol condensation product. Such purification steps are resource and energy intensive, in addition the negative impact of potential losses of product during such processes. Therefore the optimal solvent ratio was set at 0.5 mL of ethanol per 1 mL of Cyrene. Ethanol was selected as a co-solvent based on a number of factors including its miscibility with Cyrene, green credentials as a bio-based solvent, cost, and importantly, at the ratio of Cyrene to ethanol it had no negative impact on the solubility of the bases. Another suitable solvent tested included methanol however, ethanol was selected based on its bio-based greener credentials.

The K3PO4 catalysed self-aldol condensation of Cyrene at 120 °C and 0.5 vol/vol of ethanol gave product with an isolated yield of 81.3%. The reaction was completed within a short time (40 min) and with the minimum amount of solvent required. This reaction time was a considerable improvement over the previous reported synthesis of this molecule (previously 8 hours heating in a microwave was reported).4 The product was characterised by 1H and 13C NMR spectroscopy, GC-FID, GC-MS and IR spectroscopy (See Electronic supporting information). Additional experiments were undertaken in a pressure reactor at 120 °C for 40 minutes, however, these demonstrated comparable yields to those undertaken at ambient pressure over the same time. Consequently, the reaction at ambient pressure was scaled up by a factor of 10 which had no detrimental effect on the yield or the purity (GC-FID). Recrystallisation of the product in hot acetonitrile has been shown to produce large, white crystals (see Electronic supporting information). Other green metrics for the optimised self-aldol condensation of Cyrene indicated an atom economy of 93%, a reaction mass efficiency of 76.6% and a PMI of 2.78.

Due to the ease by which Cyrene undergoes a self-Aldol condensation reaction, a range of crossed-aldol reactions were attempted in order to design novel Cyrene based compounds. It was important to select substrates which facilitated the crossed reaction and minimise any self-condensation. To eliminate self-condensation, the Cyrene enolate is required to be more reactive to the acceptor than to another Cyrene molecule. In addition, the acceptor must not easily form an enolate and react with another acceptor or Cyrene molecule. If these conditions are not fulfilled, a whole range of compounds can be synthesised. To meet these conditions, the chosen substrates must not contain any α-protons thus ensuring it does not enolise and must be an aldehyde as they are more reactive than ketones. Benzaldehyde and a range of additional aryl and heteroaryl aldehydes were therefore chosen. The reaction between a ketone, such as Cyrene, and an aromatic aldehyde in a crossed aldol reaction is referred to as the Claisen-Schmidt reaction.12 The reactions attempted with corresponding GC-conversions and isolated yields obtained are summarised in Table 2.

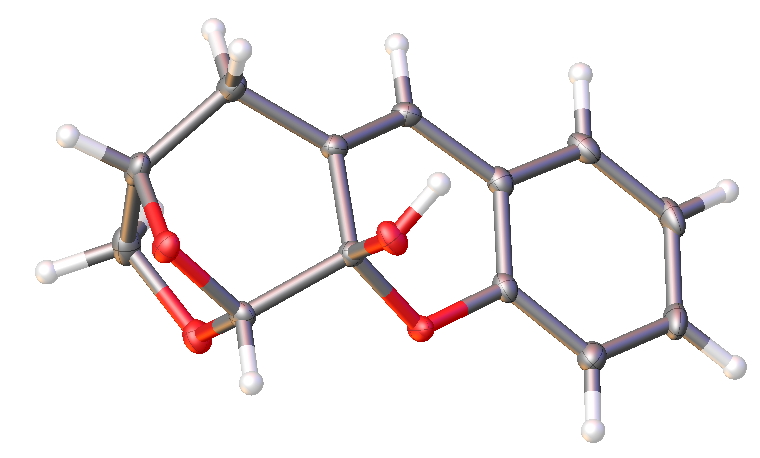
**Table 2.** Claisen-Schmidt reaction between Cyrene and substrate



|  |  |  |  |
| --- | --- | --- | --- |
| **Entry** | **Substrate (R)** | **Isolated Yield (%)** | **GC-FID Conversion (%)** |
| 1 | Benzaldehyde | 1.8 | 12.3 |
| 2 | 4-Cl Benzaldehyde | 30.7 | 85.7 |
| 3 | Salicylaldehyde | 83.1 | 95.0 |
| 4 | Furfural | 63.2 | 94.9 |
| 5 | 2-Thiophene Carboxaldehyde | 77.3 | 97.7 |
| 6 | Trans-Cinnamaldehyde | / | 19.2 |
| 7 | 4-Methyl Benzaldehyde | 40.9 | 93.4 |
| 8 | 4-Nitro Benzaldehyde | 94.7 | 100.0 |
| 9 | 3-Nitro Benzaldehyde | 43.8 | 83.7 |
| 10 | 2-Nitro Benzaldehyde | 58.2 | 62.4 |
| 11 | 4-Fluoro Benzaldehyde | 31.5 | NA |
| 12 | 3-Hydroxy Benzaldehyde | 8.3 | 77.9 |
| 13 | 2-Hydroxy, 5-Nitro Benzaldehyde | 1.2 | 60.6 |
| 14 | 4-Pyridine Carboxaldehyde | 13.8 | 89.7 |
| 15 | 3-Pyridine Carboxaldehyde | 2.7 | 93.6 |
| 16 | 2-Pyridine Carboxaldehyde | 13.0 | 91.0 |
| 17 | 2,4-Dichloro Benzaldehyde | / | 96.7 |
| 18 | 2,4-Dimethoxy Benzaldehyde | / | 91.2 |
| 19 | Benzaldehyde\* | 67.3 | 99.3 |
| 20 | Trans-Cinnamaldehyde\* | 47.2 | 99.0 |

\* Heated to 120 °C in a closed pressure vessel with stirring (550 rpm) for 18 hours, NA indicates the likely decomposition of the product in the injection liner of the GC.

The reactions were carried under the optimised conditions used for the self-condensation except reaction time was extended to 18 hours (Table 2). The only exception being reaction 8, which was run for only 20 minutes and resulted in a solid product. As observed in table 2, the conversions for the Claisen-Schmidt reaction are high (typically >80%). In keeping with qualitative metrics developed for the pharmaceutical sector, a green work up methodology was used in the form of recrystallisation.13 On the small scale employed within this study, the work up of some molecules showed a much lower recovered yield in comparison to GC data. Further optimisation at scale would undoubtedly result in far higher recovery. Such procedures could lead to the generation of pharmaceutical scaffolds at scale for further functionalisation, although this is outside the scope of the current work. In addition the use of chromatography could also aid in improving the yields of recovery but would result in extensive solvent and silica waste, leading to poor green credentials for the process. As such was not considered an appropriate purification process for this work. According to GC and GC-MS analysis the selectivity towards the product is also exceptionally high (in many cases the only product generated in the reaction), however, the isolated yields have been negatively affected due to difficulties in working up (table 2). Therefore, conversion was used to compare the differences between reactions.

 It can be seen that the reactions using benzaldehyde and trans-cinnamaldehyde (entries 1 and 6, Table 2) resulted in the lowest conversion by far. Trans-cinnamaldehyde gave slightly higher, potentially due to less steric hindrance around the carbonyl. It is unclear why these two performed the worst as even electron donating substituents (entries 3, 7 and 12, Table 2) which would give less electrophilic carbonyls produced higher conversions, however when repeated within a high-pressure reactor the conversions increased significantly(entry 19 and 20, Table 2). Previous studies have reported that the reaction of cyclohexanone with benzaldehyde also gives a slightly lower yield.8 Lower yields were not observed with benzaldehyde when the catalysts used co-ordinated with the reagents, thus bringing them in close proximity to each other and resulting in higher yields.6

The GC conversions were high for the remaining reactions. For compounds where a large substituent, such as a nitro group, is in the 2 and/or 5 positions (entries 10 and 13, Table 2) the conversions are lower which can be attributed to steric hindrance between the enolate and the aldehyde. The extremely fast rate and quantitative conversion of the reaction using 4-nitrobenzaldehyde compared to all the other reactions is due to the strong electron withdrawing effect of the nitro group and its conjugation with the aldehyde in the 4-position (entry 8, Table 2). 3-nitrobenzaldehyde reacted slower as the substituent is not conjugated with the aldehyde.

Where electronic effects could play a part is in the reactions using heteroaromatics. In furfural and 2-thiophene carboxaldehyde, the lone pair of electrons in the sp2 orbital are in the same plane as the orbitals that contribute to the aromaticity of the ring. This means the electrons are donated in to the ring rather than causing electrostatic repulsion with the lone pairs on the oxygen of the carbonyl, which would decrease its electrophilicity.

The less electronegative sulphur atom will donate its lone pair towards the ring more than the oxygen will. This could explain the higher conversion for entry 5 than for entry 4 (Table 2). Within 2-pyridine carboxaldehyde, the lone pair of the nitrogen is within a sp2 orbital in a different plane to the orbitals of the ring. This means they are not donated and the pyridine ring is described as π-deficient. The electrons then experience repulsion with the carbonyl lone pair electrons and decrease its electrophilicity. This could explain the slightly lower conversions for entry 16, Table 2. Literature data has reported this effect previously.6 The steric effect due to the smaller 5-membered ring could also contribute to this.

These products were characterised by single-crystal X-ray diffraction, FT-IR, NMR and MS. Analysis by XRD lead to production of six crystal structures. Full details of the structures can be found in electronic supporting information. The XRD confirmed the formation of a single crystal structure with one orientation. The chirality from the starting material has been retained and the crystal system is shown to be orthorhombic with a space group of P212121. However, the expected structure for salicylaldehyde (Table 2, entry 3) was not obtained. In this case, following the aldol condensation, the phenoxy oxygen of the salicylaldehyde unit has attacked the carbonyl of Cyrene, resulting in a three-ringed system (Figure 2). This structure is also confirmed via 13C NMR and is potentially more interesting due to the widespread use of fused ring compounds in the pharmaceutical industry.14 Product isolation for 2,4-dichloro benzaldehyde and 2,4-dimethoxy benzaldehyde proved to be challenging and as such limited product was isolated from reactions with these substrates. All Claisen-Schmidt reactions demonstrated atom economies in excess of 90%. Further work will focus on improving product isolation for all Claisen-Schmidt reaction products of cyrene.

**Figure 2.** Single X-ray crystal structure for the Claisen-Schmidt reaction product of cyrene and salicylaldehyde.

In order to investigate potential pharmaceutical relevance of the compounds synthesised, they were subjected to the LLAMA protocol which allows lead-likeness of all of these structures to be assessed.15 Here, structures are entered in to a database and are then in turn subjected to a number of pharmaceutically relevant reactions to generate a virtual library of products. These are then assessed on key properties (heavy atom count, AlogP, number of aromatic rings and an undesirable functional group filter) with penalty points assigned depending on distance from the ideal space. This value is termed the lead-likeness (Figure S27-S29, supporting information). Scaffolds that had only one reaction (decoration) resulted in 84 new structures of which 35 came with a lead-likeness penalty of 0 and 12 with a penalty of 1. This suggests that these structures have potential as API scaffolds. Unsurprisingly all compounds had 0% likeness on the ZINC database *i.e.* entirely novel.

In conclusion, the yield and conversion of the self-aldol condensation reaction of cyrene was optimised with respect to temperature, co-solvent ratio and base catalyst, with K3PO4 performing the best in this study. Importantly, reaction time has been reduced from the previously reported 8 hours in a microwave, to a 40 minute reaction at reflux whilst achieving better or equivalent yields respectively. Analysis by 1H and 13C NMR, GC-MS and GC-FID demonstrated high product purity, thus enabling its use as a bio-derived platform molecule.

In addition eighteen previously unreported structures were synthesised through the Claisen-Schmidt reaction of cyrene and aldehydes. The X-ray crystal structures and use of the LLAMA protocol showed how these compounds have similar structures to current drug analogues and therefore it is conceivable that they could have similar uses. This work identifies potential new routes to bio-based molecules which may enable them to compete with petrochemical products in the future.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1. J. Sherwood, M. De bruyn, A. Constantinou, L. Moity, C. R. McElroy, T. J. Farmer, T. Duncan, W. Raverty, A. J. Hunt and J. H. Clark, *Chem. Commun.,* 2014, **50**, 9650-9652
2. A. Alves Costa Pacheco, J. Sherwood, A. Zhenova, C. R. McElroy, A. J. Hunt, H. L. Parker, T. J. Farmer, A. Constantinou, M. De bruyn, A. C. Whitwood, W. Raverty and J. H. Clark, *ChemSusChem*, 2016, **9**, 3503-3512
3. J. Zhang, G. White, M. Ryan, A. J. Hunt, and M. J. Katz, *ACS Sustainable Chem. Eng.*, 2016, **4**, 7186-7192
4. K. L. Wilson, A. R. Kennedy, J. Murray, B. Greatrex, C. Jamieson and A. J. B. Watson, *Beilstein J. Org. Chem.* 2016, **12***,* 2005–2011.
5. C. Palomo, M. Oiarbide, and J. M. Garcia, *Chem. Soc. Rev*., 2004, 33, 65
6. S. Bhagat, R. Sharma and A. K. Chakraborti, *J. Mol. Catal. A: Chem.*, 2006, **260**, 235-240
7. L. An, J. Zou and L. Zhang, *Catal. Commun.*, 2008, **9**, 349-354.
8. J. Li, W. Yang, G. Chen and T. Li, *Synth. Commun.*, 2003, **33**, 2619-2625.
9. R. K. Henderson, A. P. Hill, A. M. Redman and H. F. Sneddon, *Green Chem*., 2015,**17**, 945-949.
10. J.L.L. Beitia, *Synlett*, 2011, **1**, 139-140
11. (a) Data kindly provided by Circa Group Ltd., the manufacturer of Cyrene, by F. Hoffmann La Roche Ltd., hereby acknowledged as the source of the data. (b) Safety Data Sheet of Circa Group Pty Ltd. (c) R. Azzi-Hartmann, unpublished work.
12. (a) L. Claisen and A. Claparède, *Ber. Dtsch. Chem. Ges*., 1881, **14**, 2460; (b) J. Gustav Schmidt, *Ber. Dtsch. Chem. Ges*., 1881, **14**, 1459; (c) A. T. Nielsen and W .J. Houlihan, *Org. React*., 1968, **16**, 1.
13. C. R. McElroy, A. Constantinou, L. C. Jones, L. Summerton and J. H. Clark, *Green Chem.*, 2015, **17**, 3111-3121
14. H. K. Maurya, maebaldolreaction.blogspot.com, 2011, 1P, 1-29 (<http://maebaldolreaction.blogspot.in/>), accessed March 2018
15. I. Colomer, C. J. Empson, P. Craven, Z. Owen, R. G. Doveston, I. Churcher and S. P. Marsden, *Chem. Commun*., 2016, **52**, 7209-7212.