BioLogicTool: A Simple Visual Tool for Assisting in the Logical Selection of Pathways from Biomass to Products

Yann Lie,[a] Pablo Ortiz,[b] Richard Vendamme,[b] Karolien Vanbroekhoven[b] and Thomas J Farmer[a]*

[a] Green Chemistry Centre of Excellence, Department of Chemistry, University of York, Heslington, York, YO10 5DD, UK

[b] Flemish Institute for Technological Research – VITO, Separation & Conversion Technology, Boeretang 200, 2400 Mol, Belgium

KEYWORDS

Bio-based chemicals, van-Krevelen diagram, green metrics, platform molecules, lignin

ABSTRACT

The current chemical industry has been heavily optimised for the use of petroleum-derived base chemicals as its primary source of building-blocks. However, incorporation of heteroatoms, absent in the base chemicals, is necessary to meet the different property requirements in the plethora of products the industry makes such as plastics, cosmetics and pharmaceuticals. As global oil reserves

deplete, a shift towards renewable bio-derived building-blocks, so called platform molecules, will become a necessity. Bio-derived platform molecules are typically rich in heteroatoms as a result of their biomass feedstock also being heteroatom rich, and it would therefore seem logical to carry these heteroatoms through to the aforementioned products. A tool was herein developed to assess the rationality of a synthetic route from feedstock to product, designed specifically to give a visual representation of the pathways and options available. *BioLogicTool* plots (%heteroatom by mass vs M) are an alternative to the conventional van Krevelen diagram, and are designed to better consider the diversity in heteroatom content encountered in bio-based chemicals. The tool can rapidly help its user to design more logical multi-step synthetic routes and enhance the mass efficiency of pathways. Several examples were selected to demonstrate the power and limitations of the *BioLogicTool*, but it was clear from these examples that removing heteroatoms from platform molecules to reincorporate them later in the final product is, in most cases, not logical in mass efficiency sense.

INTRODUCTION

The chemical industry of the 20th Century was a success story, through the century countless advances were achieved that brought a plethora of products to market that have come to define humankind. However, the vast majority of the products created by the chemical industry are derived from non-renewable fossil resources (i.e. petroleum-derived). Additional to issues around long-term supply, petroleum-derived chemicals are inherently linked to negative environmental problems brought about by damage caused during their production, use and end-of-life. As such the 21st Century chemical industry will embrace a far greater prevalence of biomass as a feedstock in a hope of overcoming long-term supply issues and environmental damage linked to the use of fossil resources.

The vast majority of petroleum-derived products are reliant on just a handful of simple building-blocks, the so-called base chemicals. These base chemicals (Figure 1, further information for each is available in Table S1) have annual productions of a monumental scale, all exceeded 10 million tonnes with ethene being produced in excess of 100 million tonnes. They each also contain a functional group (phenyl ring, alkene or hydroxyl) that acts as the first site for further functionalization.
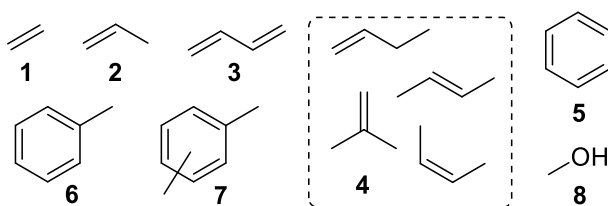

Figure 1. Petroleum-derived base chemicals: ethene **1**; propene **2**; butadiene **3**; butene (isomers, within dashed box) **4**; benzene **5**; methylbenzene **6**; dimethylbenzene (isomers) **7**; methanol **8**

A highly-optimized chemical industry takes these base chemicals as starting point for the production of bulk and commodity chemicals, and eventually leading to the many products we are reliant upon, such as plastics, textiles, pharmaceuticals, agrochemicals, dyes and detergents. It is therefore logical that any future bio-based chemical industry likewise seeks to select a range of large-scale building-block chemicals to use precursors to the desired products. Indeed, over the last 20 years the use of bio-based building-block chemicals has re-emerged as an area of significant academic and industrial interest. Although several terms are used for bio-based building-blocks, the most widely used is that of *platform molecules*. A platform molecule is defined as:[1]

"…a chemical compound whose constituent elements originate wholly from biomass (material of biological origin, excluding fossil carbon sources), and that can be utilised as a building block for the production of other chemicals."

For sake of further discussions in this article the authors compiled an extensive list of platform molecules (Figure 2, **9-59**) that match this definition. Further information for each compound in this list is available in Table S2. The list contains platform molecules that are derived from several different constituent parts of the biomass (e.g. cellulose, lignin, sugars, protein) and also from different processing technologies (chemocatalytic, biological, thermochemical and extraction). Although the list does not contain every single possible platform molecule it is a set that the authors believe adequately represents the possible options. An additional 5 compounds (Table S3, **60-64**) were initially considered but removed from later discussions as a cap on the molecular mass of 220 g/mol was deemed appropriate.
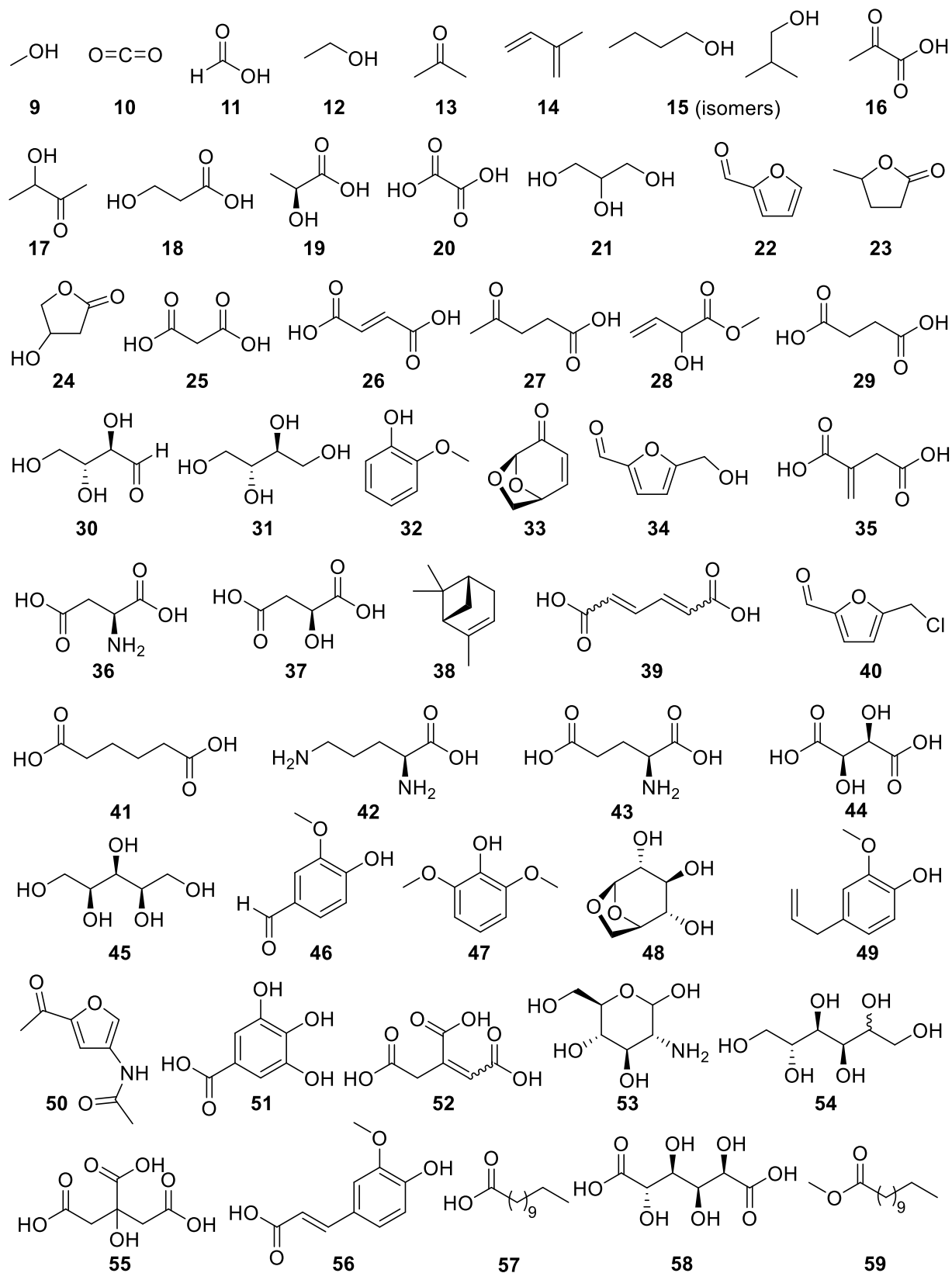
9 10 11 12 13 14 15 (isomers) 16

17 18 19 20 21 22 23

24 25 26 27 28 29

30 31 32 33 34 35

36 37 38 39 40

41 42 43 44

45 46 47 48 49

50 51 52 53 54

55 56 57 58 59

Figure 2. Biomass-derived Platform Molecules: methanol **9**; carbon dioxide **10**; formic acid **11**; ethanol **12**; acetone **13**; isoprene **14**; butanol (n-, iso-) **15**; pyruvic acid   **16**; acetoin **17**; 3-hydroxypropionic acid **18**; lactic acid **19**; oxalic acid **20**; glycerol **21**; furfural **22**; γ-valerolactone **23**; 3-hydroxybutyrolactone **24**; malonic acid **25**; fumaric acid **26**; levulinic acid **27**; methyl vinyl glycolate **28**; succinic acid **29**; erythrose **30**; erythritol **31**; guaiacol **32**; levoglucosenone **33**; 5-(hydroxymethyl)furfural **34**; itaconic acid **35**; aspartic acid **36**; malic acid **37**; D-limonene and pinenes **38**; muconic acid **39**; 5-(chloromethyl)furfural **40**; adipic acid **41**; L-lysine **42**; glutamic acid **43**; tartaric acid **44**; xylitol and arabitol **45**; vanillin **46**; syringol **47**; levoglucosan **48**; eugenol **49**; 3-acetamido-5-acetylfuran **50**; gallic acid **51**; aconitic acid **52**; glucosamine **53**; sorbitol and mannitol **54**; citric acid **55**; ferulic acid **56**; fatty acids (e.g. lauric acid) **57**; glucaric **58**; fatty acid alkyl esters (e.g. methyl laurate) **59**

A glance at Figures 1 and 2 already highlight the major difference between a petroleum- or bio-derived chemical industry in terms of the building blocks. The fossil feedstock of the current chemical industry is heavily depleted of heteroatoms (oxygen, nitrogen, halogens etc.) and is instead predominately comprised of hydrocarbons. This is clearly seen in the molecular structures of the base chemicals (Figure 1), where only methanol (**8**) contains a heteroatom and this oxygen is sourced from atmospheric gases and not from the crude oil or natural gas feedstock. The current chemical industry therefore adds further functionality to simple base chemical building blocks, this often includes the incorporation of heteroatoms. Biomass on the other hand is rich in heteroatoms, especially oxygen and nitrogen, and therefore the platform molecules are more often than not also very rich in heteroatoms. As such it is anticipated that any future bio-based chemical

industry will have to either remove this high heteroatom content if targeting drop-in replacements to petroleum-derived product, or to embrace new products with a higher heteroatom content.

Indeed, the selection between targeting a bio-based drop-in replacement versus a new chemical/material that offers similar properties to the incumbent petroleum-derived is a key question that must be answered. For example, when seeking to substitute petroleum-derived poly(ethene) with a bio-based equivalent, is the best approach to simply make the ethene monomer (**1**) from biomass feedstock or should an alternative polymer with a different molecular structure be sought? This question is particularly important when one considers the differences in the heteroatom content of a biomass and petroleum feedstock relative to the desired product. It is when considering this issue that we undertook a study to see if a simple calculation could be applied prior to any physical lab work taking place that would allow for the assessment of the proposed routes to be compared. Some previously established metrics showed initial promise in the regard. For example Atom Economy (AE) is a facile calculation that can be performed by simply understanding a reaction scheme and correctly identifying the reactants. AE is the percentage of the reactants that ends up in the product. For the example of poly(ethene) above, AE already begins to give a useful guide to the option clearly the use of a carbohydrate feedstock is somewhat illogical as all the heteroatom content is removed. However, there are several issues with AE:

1) It does not consider what reactants are bio-derived or not

2) It requires a sound grasp of the reaction mechanism so that the stoichiometric equivalents can be correctly determined (e.g. it may not be obvious to all that two of the carbon atoms in the glucose are lost as CO_2 during the fermentation)

3) It does not offer a visual representation of a pathway

Biomass Utilization Efficiency (BUE) is another metric of merit when considering a pathway from feedstock through to product. BUE seeks to determine how effective a chosen route is for converting 1 kg of feedstock through to product, and can be extended to take into consideration the yields of any reaction in question. Unlike AE, BUE specifically assigns atoms in the product as being bio-based or not, though again this requires an even better grasp of the reaction mechanism than required for AE. Despite the promise of the two metrics above we still wanted to have at our disposal a tool that visually represented the differences in the reaction pathways and that did not require an exact grasp of the reaction mechanisms.

MAPPING THE HETEROATOM CONTENT OF PLATFORM MOLECULES

The first obvious visual representation to use is that of the van-Krevelen diagram, as shown for the base chemicals and platform molecules in Figure 3**A**. A van-Krevelen diagram is a plot of the atomic ratios of H:C (y-axis) versus O:C (x-axis), and was developed as a mean of comparing the composition of kerogen and petroleum.[2] The approach of using a van Krevelen diagram for biomass feedstocks has already been demonstrated and certainly shows some merit, especially when plotting some biomass feedstocks against fossil resources.[3–5] However, the conventional van Krevelen plot has short-comings when used for considering pathways to and the use of bio-based building-block chemicals. Dusselier *et al.* also saw the need for a refined version of the van Krevelen diagram more appropriate for a bio-based chemical industry, and in their approach they replaced either axis with a functionality ratio (F:C).[6] The Dusselier approach is an excellent method that resulted in some important "key lesson" that those working in bio-based chemicals should remain aware of. Herein we sought to modify the van Krevelen diagram further into simple

but powerful tool for bio-based product pathway selection as well as leading to some broader general lessons around bio-based chemicals.
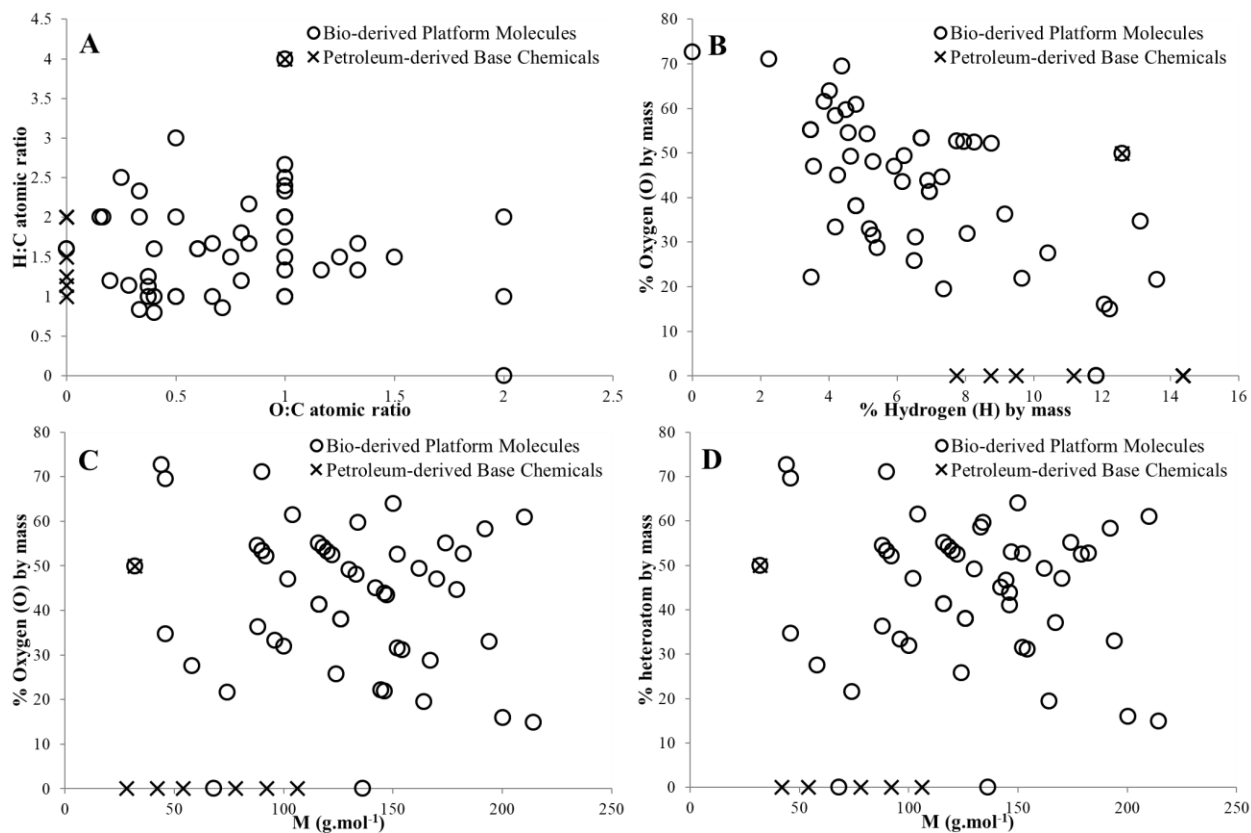

Figure 3. Iterations of heteroatom content plots for petroleum-derived base chemicals (X) and biomass-derived platform molecules (O): **A**) van Krevelen diagram (H:C versus O:C atomic ratios); **B**) modification of van Krevelen diagram with axes flipped and ratios changed to % oxygen and % hydrogen by mass; **C**) Enhanced plot with x-axis of **B** altered to be molecular mass (M, g.mol^{-1}); **D**) Further enhanced plot with y-axis of **C** altered to %heteroatom (%weight). A version of **D** with full assignments of all points is shown the supporting information (Figure S1)

Changing the x- and y-axis to a percentage of H and O content respectively (as %weight, Figure 3**B**) rather than the atomic ratio was preferential as it widened the number of points on each axis,

for example one can clearly see from Figure 3**A** how several molecules line up with an O:C ratio of 1. Additionally, using a %weight value also takes into consideration the more significant role the heteroatoms make relative to hydrogen towards the overall mass of the compounds. As such, the plot is further enhanced by using the molecular mass (M) for the x-axis, as shown in Figure 3**C**. When considering pathways from the feedstock through to a final product the M is logically a more useful component to include in the plot as this has a significant bearing of the mass yields of reactions (linked somewhat to atom economy), e.g. elimination reactions clearly lower mass. The use of M also highlights another key difference for the base chemicals and platform molecules, which is that most platform molecules have a higher molecular mass. Indeed we chose to remove those platform molecules whose M was >220 g/mol to refine the plots, details of these molecules (**60-64**) and extended *BioLogicTool* plots containing them can be found in the supporting information. The final iteration of the plot is made by converting the %oxygen of the y-axis to become %heteroatoms (Figure 3**D**). This final iteration has no effect on the base chemicals, and it is for this reason that such a change was never applied to van Krevelen diagrams, but is of importance to some platform molecules and for many other bio-based chemical, a point recently highlighted by Hüsley *et al.*[7] From our list of platform molecules, 6 contain heteroatoms other than oxygen, these being 3 example amino acids (lysine **42**, aspartic **36** and glutamic acid **43**, these being selected as commonly occurring amino acids), two compounds derivable from chitin (3-acetamido-5-acetylfuran **50** and glucosamine **53**) and 5-(chloromethyl)furfural **40**. It would be anticipated that other heteroatom containing platform molecules may become more prevalent in the future, including compounds containing sulfur (e.g. cysteine and methionine), phosphorous (e.g. phosphorolipids) or other halogens. Additionally, as the plots are planned to be used to assess pathways to products it is likely a greater consideration of the role of heteroatoms is required.

Rivas-Ubach *et al*. took a similar approach when refining the van-Krevelen plot for classifications of organic matter in environmental studies, they too appreciated the need to consider heteroatoms such as N and P in biological samples.[8] However, Rivas-Ubach's classification was still based on atomic ratios (O:C, H:C, N:C, P:C and N:P) and as shown from the discussion above there is additional merit in the use of %mass over atomic ratio.

By using the %heteroatom versus M plots, an even more striking comparison of a petroleum- and biomass-derived chemical industry was possible by plotting a range of key products from the current chemical industry onto the plot developed above. For this a selection of fourteen products that represent examples of polymers of various classes, solvents, pharmaceuticals, synthetic preservatives, organocatalysts, herbicides, nematic liquid crystals, dyes and surfactants was used. The outcome of this approach, shown in Figure 4, highlights how important the introduction of heteroatoms in the chemical industry is, whilst also further showing the necessary increase in molecular mass to prepare the desired products (note: the mass of the constitutional repeat unit was used for the example polymers). Only the plastic poly(ethene) **65** remains devoid of heteroatoms, all of the other products selected include some heteroatom content, with poly(vinylchloride) **66** and the herbicide 2,4-dichlorophenoxyethanoic acid **72** both containing >50%weight heteroatom. Nine (64%, **66-68**, **70**, **72-74**, **76-77**) of the selected products contain heteroatoms other than oxygen, this reinforcing the need to use %heteroatom and not simply %oxygen for the comparative plots. Figure 4 also highlights another major consideration, this being that a petroleum-derived chemical industry is essentially heavily reliant on oxidation processes, where oxidation is viewed as being the increase in oxidation number of the carbons present and therefore means inclusion of all heteroatoms. Instead, many of the platform molecules already have levels of oxidation similar to the desired products and therefore a biomass-derived

chemicals industry will be less reliant on oxidation, perhaps instead favouring preservation of the heteroatom content from the platform molecules, this ensuring better overall mass efficiencies for the conversion of feedstock through to product. This important aspect will be discussed in greater detail through selected examples below, and is where the use of such %heteroatom versus M becomes most valuable. Using these plots as a tool to assess what pathways from biomass feedstock through to product are most logical has been abbreviated by the authors to *BioLogicTool*.
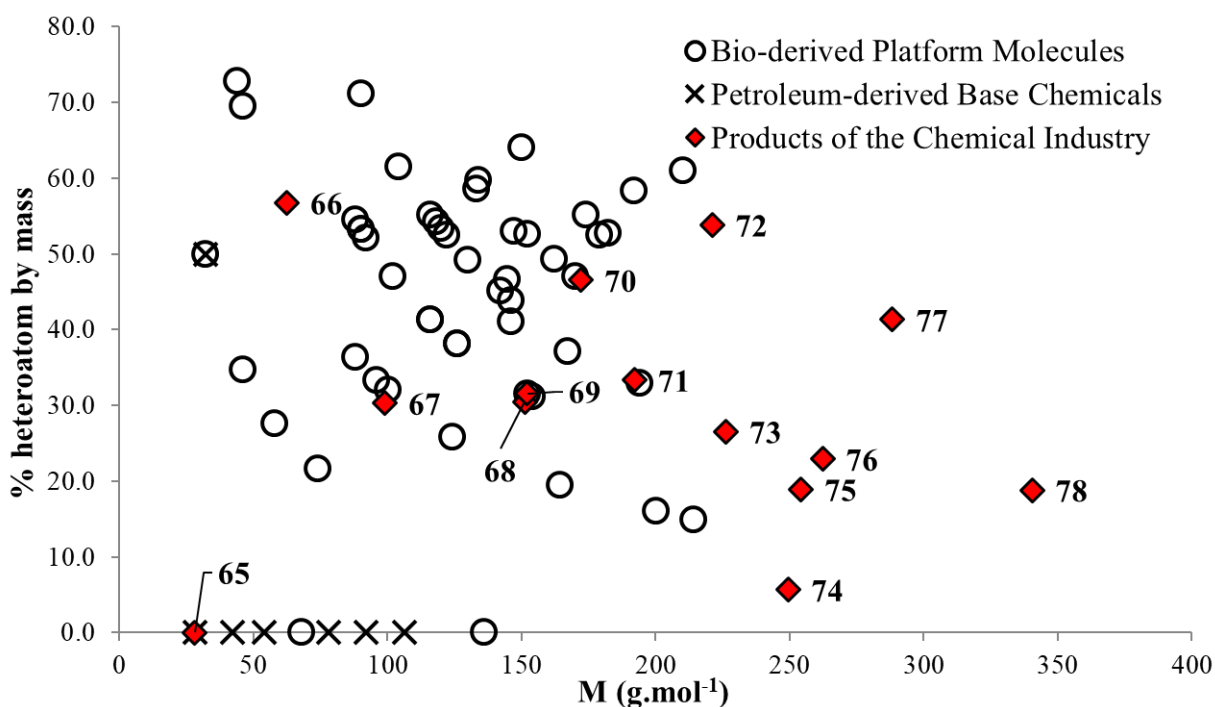

Figure 4. Mapping major products from the chemical industry against the base chemicals and platform molecules on a plot of %heteroatom (by mass) versus molecular mass (M, g.mol^{-1}). Products: poly(ethene) **65**; poly(vinyl chloride) **66**; N-methyl-2-pyrolidone (solvent) **67**; N-acetyl-p-aminophenol (paracetamol) **68**; methyl paraben (preservative) **69**; p*ara*-toluenesulfonic acid (organocatalyst) **70**; poly(ethylene terephthalate) (polyester) **71**; 2,4-dichlorophenoxyethanoic acid (herbicide) **72**; nylon-6,6 (polyamide) **73**; 4-cyano-4'-penylbiphenyl (liquid crystal) **74**; poly(bisphenol-A) carbonate (polycarbonate) **75**; 2,2'-Bis(2,3-dihydro-3-oxoindolyliden) (indigo

blue dye) **76**; sodium dodecylsulfate (surfactant) **77**; bisphenol A diglycidyl ether (resin precursor) **78**. Note: for polymers the mass of the constitutional repeat unit (CRU) is used for the M.

BIOLOGICTOOL: ASSESSING PATHWAYS FROM FEEDSTOCK TO PRODUCTS BY SUMMATION OF ROUTE VECTOR LENGTHS

The *BioLogicTool* created by modification of the van Krevelen diagram allows all feedstock (A), intermediates (B, C) and final products (D) of a synthetic route to be plotted on the same graphical object. This gives the opportunity to 'visualise" reaction pathways by plotting vectors beginning from starting materials (or intermediates) and ending to products (or intermediates) (Figure 5).

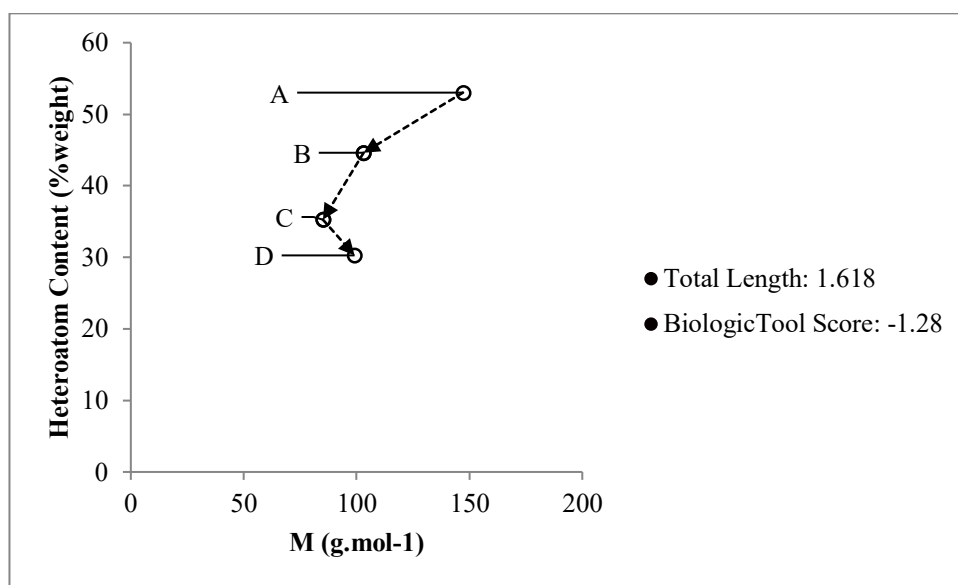**Figure 5.** Visualisation of a synthetic route starting from feedstock (A) and ending to product (D) via intermediates (C & D) using the heteroatom content (%weight) for the y-axis and the molecular mass M (g.mol^{-1}) for the x-axis.

The tortuosity of the route formed by the vectors was deemed a good representation of the logicality of said synthetic path (within the limits of the variables used here). However, tortuosity

can be laborious to calculate and may require long computing times. Instead, the lengths (norms) of all individual vectors were calculated (l_i, eq. ii) and summed (iii) to give the *Total length* of a synthesis (based on its vectorial representation). In the example Figure 5, l_1 would be the length of the vector formed by points A and B, l_2 the length of the vector formed by points B & C and so on. Prior to any calculation, units of the data were removed by normalising them following equation (i).

$$Xstd = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad \text{(i)}$$

$$l_i = \sqrt{(x_{ib} - x_{ia})^2 + (y_{ib} - y_{ia})^2} \qquad \text{(ii)}$$

$$\forall i, n \in \mathbb{N}, \; Total \; length = \sum_{i=1}^{n} l_i \qquad \text{(iii)}$$

Equations i, ii & iii. (i) Formula used to standardised data (ii) Formula used for the calculation of the norm of a vector formed by the points A (xa,ya) and B (xb,yb) (iii) Equation for the *Total length* of a synthetic pathway. n is the number of vectors for a reaction pathway and i the variable attributed to a vector

Finally, the *Total length* alone was not sufficient to give a proximate of the tortuosity of a reaction pathway. During the computation, a final vector was taken into account: starting from the feedstock and ending to the final product namely this vector was called the *direct vector*. In the above example, the *direct vector* would be starting from A and ending to D but not via B or C. This direct vector was also given a positive sign if the difference in %Heteroatom was positive (i.e. heteroatom content increased) or a negative sign in the other case (i.e. heteroatom content is reduced).

Dividing the *Total length* of a reaction route by the length of the *direct vector* gives the *BioLogicTool scores* which represent the logicality of a synthetic pathway. The more "logical" a reaction is the less tortuous its vector path and the closer to 1 the *BioLogicTool score* should be. Explanation for interpreting the *BioLogicTool score* value and sign is given in Table 1.

Table 1. Interpretation of high/low *BioLogicTool* score value and positive of negative sign

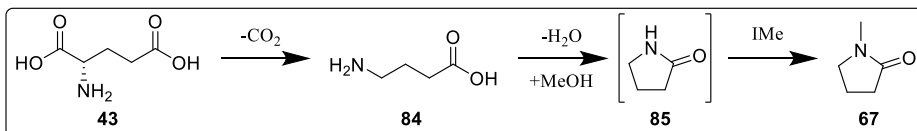|		*BioLogicTool score* sign	
		+	-
BioLogicTool score value	>> 1	Gain of heteroatom, more tortuous route	Loss of heteroatom, more tortuous route
	~ 1	Gain of heteroatom, more direct route	Loss of heteroatom, more direct route

A spreadsheet tool was designed to allow any researcher interested in calculating the *BioLogicTool score* of a reaction pathway and plotting it. This tool is available via the supporting information.

As an example, the comparison between the bio-based and the oil-based route to NMP (**67**) will be discussed (Scheme.1). NMP is a polar non-protic solvent used in large amount in the formulation and processing paints and coatings, pharmaceuticals, plastics etc., its production is estimated to be 200-300 kt/a. The current industrial process uses acetylene (**79**) as feedstock, this condensed with formaldehyde to 1,4-butyndiol (**80**). A hydrogenation step forms 1,4-butanediol (**81**) which is oxidised to γ-hydroxybutyric acid (**82**). **82** is cyclised to form the γ-butyrolactone (**83**) and a lactone-lactam conversion with methylamine converts **83** to NMP (Scheme 1).
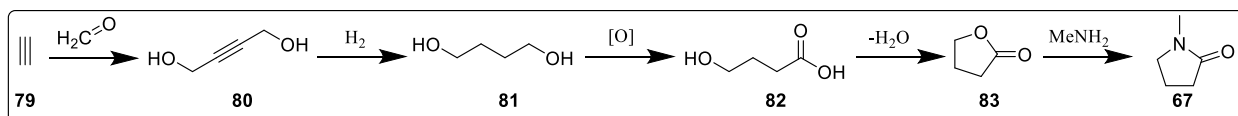
Scheme 1. Petroleum-derived and bio-derived route to NMP (**67**)
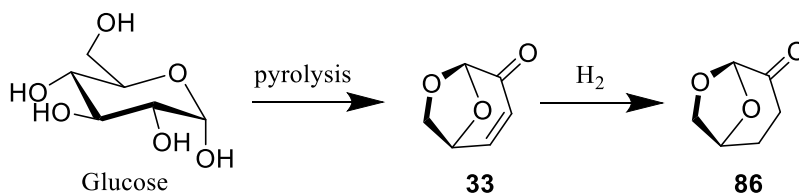
Scheme 2. Route to Cyrene (**86**) using glucose as the repetitive unit of cellulose

Figure 6. A *BioLogicTool* plot for the oil-based route to NMP; **B** *BioLogicTool* plot for a bio-based route to NMP; **C** *BioLogicTool* plot for bio-based route to NMP without isolation of **85**; **D** *BioLogicTool* plot for synthetic pathway to Cyrene®

The bio-based route to NMP starts from L-glutamic acid (**43**), one of the most abundant amino acid in nature and is currently produced on very large scale by fermentation (2 million tons/year).[9] The first step to NMP invo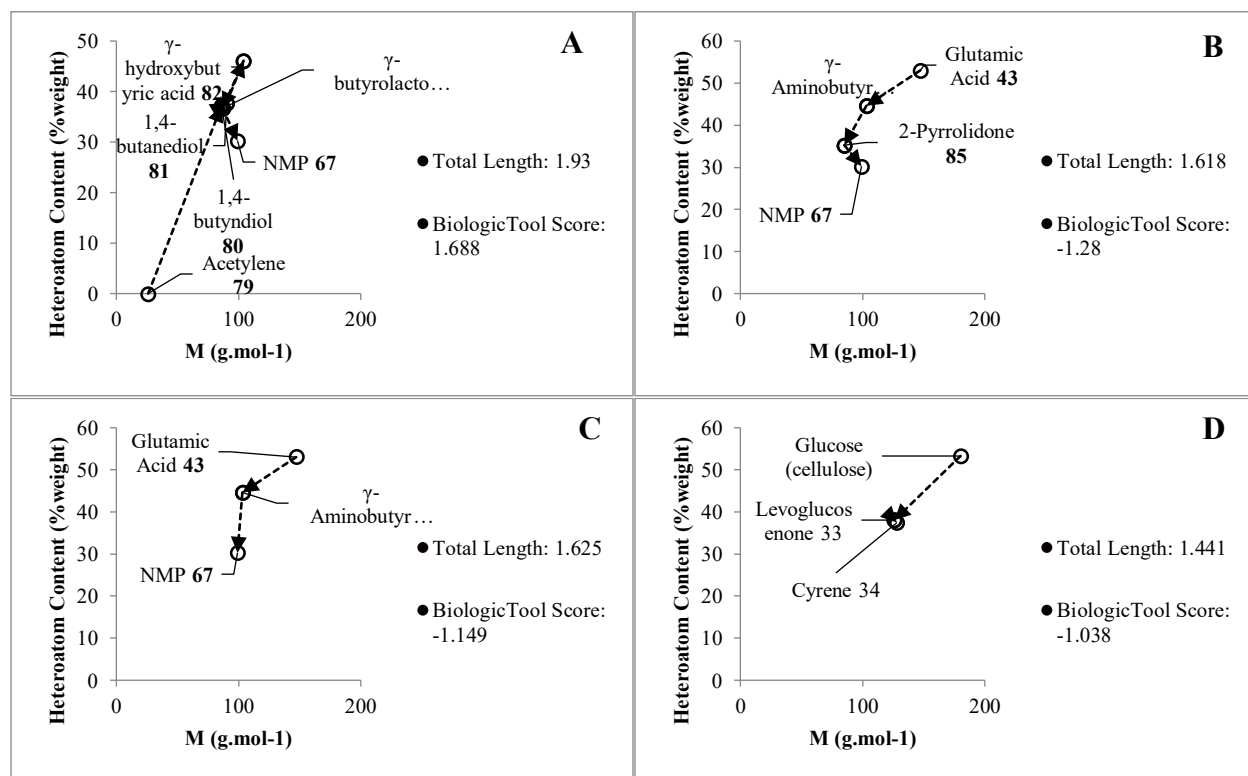lves the decarboxylation of glutamic acid to γ-aminobutyric acid (GABA, **84**). GABA is then cyclised to 2-pyrrolidone (**85**) which is eventually methylated to form NMP. Alternatively, 2-pyrrolidone may not be isolated, instead, NMP can be directly generated by a one-pot cyclisation/methylation of GABA. [10]

Using the *BioLogicTool* to visualise these pathways highlights the differences in terms of complexity and rationality (Figure 6). Many intermediates with close M/%weight Heteroatoms are involved in the oil-based route which necessarily increases the *Total length* of the reaction (1.930) and thus, its *BioLogicTool score* (+1.688) (Figure 6**A**). On the other hand, the bio-based route, with only 3 reaction steps, has a cumulative vector length of 1.62 and a *BioLogicTool score of -*1.280 (Figure 6**B**). If the isolation is avoided and GABA is directly cyclised and methylated the cumulative vector length does not change (1.62 due to the normalisation of data) but the *BioLogicTool* scores drops to -1.15 (Figure 6**C**). Despite the seemingly small difference between these scores they illustrate well how the *BioLogicTool* can be used as a tool to discriminate between two routes to an identical product but that both cumulative vector length and *BioLogicTool score* should be taken into account to make an informed decision. Nevertheless, the bio-based route to NMP is a drop-in replacement which does not overcome the initial issues caused by NMP. Indeed, despite its good biodegradability and low aquatic life toxicity, NMP is an irritant solvent that is toxic upon chronic exposure. Hence, using a safer, greener replacement would be preferable. Recently, an alternative solvent, Cyrene®, was developed which is likely an excellent replacement for polar aprotic solvent such as NMP. Cyrene® is produced from raw biomass by pyrolysis, forming levoglucosenone (**33**) and a single step hydrogenation produces Cyrene® (**86**) (Scheme

2).[11,12] Because of the simple synthetic route (Figure 6**D**), with the only intermediate **33** almost being on the *direct vector* (From glucose (cellulose) to Cyrene®), the *BioLogicTool score* of this pathway is -1.038, showing very low tortuosity.

These chosen examples illustrate well how the *BioLogicTool* presented here can help in assessing the rationality of a reaction scheme using readily available data. The high-heteroatom content of bio-derived molecules make them attractive to use as raw materials for commodity products. Nonetheless, it is evident that in some cases the bio-based route to a product might not always be the preferred one, especially when low/no heteroatoms are present in the final product.

LOW/NO HETEROATOM CONTENT PRODUCTS FROM BIOMASS

It will not always be the case that the final products of the chemical industry will need to have a high heteroatom content, there will still be a necessity in a future bio-economy for chemicals of very low polarity (e.g. hydrophobes of surfactants, lubricants, solvents for cleaning). Therefore, practical solutions for the production of low/no heteroatom content products from a biomass-derived chemical industry will need to be developed. The first option is to consider which platform molecules of low heteroatom content will be of most practical use. The terpenes (isoprene **14**, pinenes and D-limonene **38**) were the only platform molecules on the list with no heteroatom content. In fact, terpene chemistry is already widely developed and exploited, especially in the preparation of flavours and fragrances and is a major contributor to proportion of our current chemical industry that is already bio-derived. However, terpenes are not available on a scale to match base chemicals, data is limited for turpentine oil (a mixture rich in terpenens) production but is estimated to have been around 360-370 ktonnes in 2017.[13]

An alternative option would be plant-derived oleochemicals, the examples used in the platform molecules list being lauric acid **57** and methyl laurate **59**. Again, oleochemistry is already industrially well established, for example many surfactants are reliant on the C12 lauric acid from palm kernel oil as the hydrophobic component. Other important fatty acids as chemical precursors include oleic **60** and ricinoleic acid **61**, both widely used as their C=C (both) and –OH (latter) functionality allow for further chemical modification. However, both have a M of >280 g/mol and were therefore omitted from the list of platform molecules in Figure 2 but are included in the extended plots in the supporting information. Nevertheless, oleochemicals will likely remain a major contributor to the low heteroatom content products of a bioeconomy though more sustainable pathways to their manufacture, perhaps via fermentation, will most certainly be needed.

Other low heteroatom content compounds are extractable directly from biomass, especially secondary metabolites such as waxes, carotenoids **63-64** or phenolics like anacardic acid **62**. However, in all these instances the global annual production will be too low to substitute the many millions of tonnes of base chemicals produced.

The final option for a biomass-fed chemical industry will be to remove heteroatom content from those platform molecules that are more easily producible a mega-scale (i.e. those available from lignocellulose). An example of this possible approach is given below for the comparison of petroleum-derived poly(ethene) versus the same polymer but produced from glucose as feedstock (Scheme 3). As petroleum-derived poly(ethene) is simply the addition polymerization of the base chemical ethene **1**, and the polymers constitutional repeat unit (CRU, atoms within the brackets) is identical to the M of **1** then no path on the plot is possible as neither M nor %heteroatom (0%) changes. This also means that the *BioLogicTool score* (tortuosity) is 1 for the petrochemical route.

The possibility for production of bio-derived poly(ethene) has received much press lately, but clearly the *BioLogicTool* plot (Figure 7) highlights a major issue in this approach, the total removal of all the heteroatoms from original glucose feedstock means that a very significant mass is lost overall during the process, this being in the form of waste CO_2 and H_2O that carries little or no value as by-product. Evidently, a far more suitable bio-derived replacement for poly(ethene) must be sought in the future, this appears even more obvious before one considers the negative environmental contribution made by poly(ethene) due to its low biodegradability.
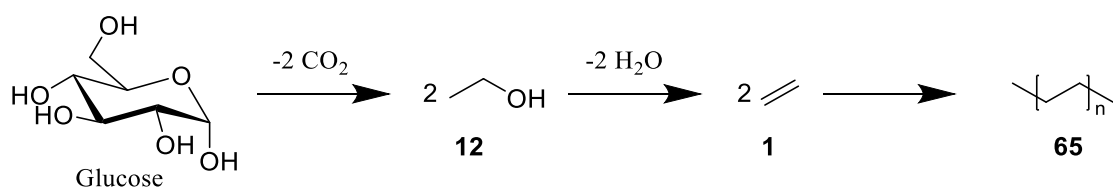
Scheme 3. Route to biomass-derived poly(ethene) **65** from glucose via the platform molecule ethanol **12**
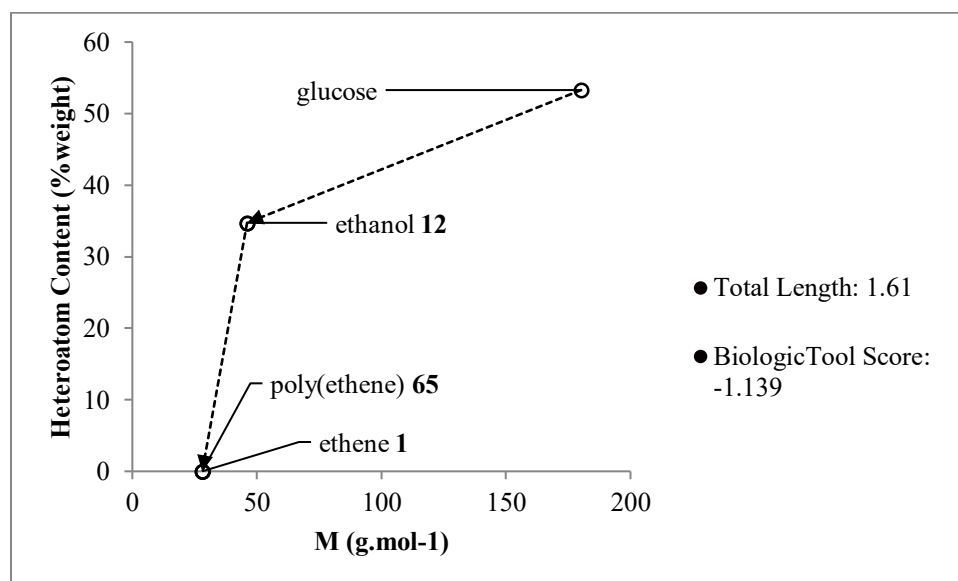

Figure 7. *BioLogicTool* plot (heteroatom content by %weight vs. molecular mass M) for the production of biomass-derived poly(ethene) from a glucose feedstock

POLY(ETHYLENE TEREPHTHALATE) – ASSESSING THE OPTIONS

Another plastic whose production from biomass is being extensively investigated is that of poly(ethene terephthalate) **71** (PET). The production of 1,2-ethanediol from biomass has been previously developed and is already commercial, and therefore the current challenge relates primarily to finding a viable route to bio-based terephthalic acid **87** or its corresponding alkyl diesters. Leading potential pathways from biomass to terephthalic acid are shown in Scheme 4, alongside the current route from petroleum-derived *para*-xylene (Route **A**). *BioLogicTool* plots were created for each pathway, these plots are presented in Figure 8, and the values for the cumulative route vectors and *BioLogicTool score* (tortuosity) are collated in Table 2. A glance at Scheme 4 reaffirms the earlier discussion that a move to a biomass fed chemical industry will result in a shift to higher hetereoatom content building-blocks, especially seen in Routes C-F where sugars are the feedstock for each pathway.

The first bio-derived Route (**B**) is via the D-limonene **38**, as previously stated this is a rare example of a platform molecule that is devoid of heteroatoms. Route 2 requires first isomerization and dehydrogenation to yield para-cymene **88**, this subsequently oxidized using nitric acid and potassium permanganate to form **87**.[14] However, as previously stated the global production of D-limonene of *circa* 70,000 tonnes would be orders of magnitude too small to meet demand of **87** and the use of such a harsh oxidation is best avoided.[15] Nevertheless, Route B (Figure 8**B**) does have a very short cumulative vector path (1.48, Table 2) and a very low tortuosity score (-1.08) by virtue of only having two steps. Route **C** is via the glucose-derived platform molecule isobutanol **15**, this dehydrated to isobutene gas **89** followed by multiple steps of dimerization, cyclisation and dehydrogenation to yield bio-based *para*-xylene **7** that is oxidized as per Route **A**.[16] The *BioLogicTool* plot for this pathway (Figure 8**C**) instantly highlights an issue, which is the total

removal of all the oxygen from the initial glucose is followed by its reintroduction at the oxidation step of **7**. As such Route **C** results in the highest cumulative vector (2.84) and the highest score for tortuosity (-9.50), and on this basis would seem an illogical pathway to select for the production of bio-based terephthalic acid. Route **D** goes via the platform molecule 5-(hydroxymethyl)furfural **34**, and uses the ability of furan units to undergo Diels-Alder additions as a diene.[17] The Diel-Alder addition of ethene **1** to **34** is problematic due to extensive side reaction as a result of this platform molecules instability, and instead the addition and concurrent aromatization is carried out on 2,5-dimethylfuran **91**. This pathway produces bio-derived *para*-xylene **7** that, as for Route **C**, can be oxidized to **87** under the same conditions current used for petrochemical terephthalic acid, thus representing a bio-based drop-in replacement.[18] Route **D** also requires total removal of all the original heteroatom content of the feedstock followed by extensive reoxidation to **87**, and much like Route **C**, this results in a very high cumulative vector (2.59) and *BioLogicTool score* (-8.01) suggesting this pathway is not ideal. However, this route is already heavily optimised and the Diels-Alder addition and concurrent dehydration occurring at very high conversions.[17,19,20] An alteration to the Route **D** has recently been demonstrated whereby some of the original oxygen can be maintained in the product diethyl terephthalate **94** (Route **E**), this is achieved through the Diels-Alder addition of ethene to diethylfuranoate **93** (more stable than **34**) with ethene over a clay catalyst.[21] By avoiding total removal of heteroatom content the *BioLogicTool score* for this path from feedstock to product is much lower (Figure 8E), -2.42. Evident from the comparison of Routes **D** and **E** is that there is clear merit in avoiding total heteroatom removal in the pathway to **87**.
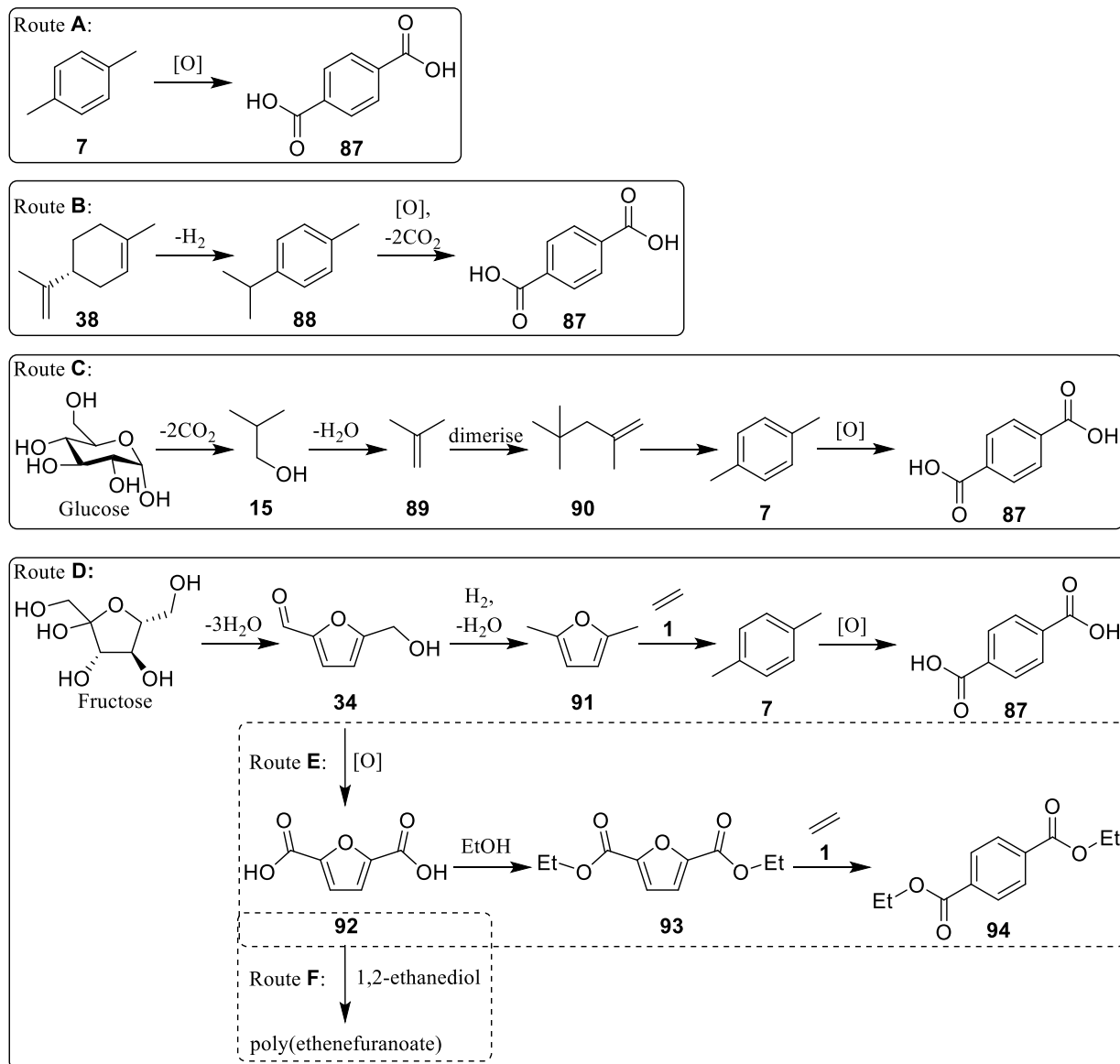
Scheme 4. Routes to terephthalic acid **87** from petroleum and biomass feedstock Route **A**: oil-based route to **87** using p-xylene **7** as intermediate; Route **B**: Bio-derived route to **87** from D-Limonene **38**; Route **C**: Bio-derived route to **87** via complete deoxygenation of glucose to **7** then oxidation to **87**; Route **D**: bio-derived route to **87** via complete deoxygenation of fructose then oxidation of **7** to **87**; Route **E**: Bio-derived route to alternative replacement of **87**, diethylterephthalate with partial deoxygenation of fructose; Route **F**: Route to potential bio-derived replacement of PET (hence **87)**, poly(ethenefuranoate) (PEF)
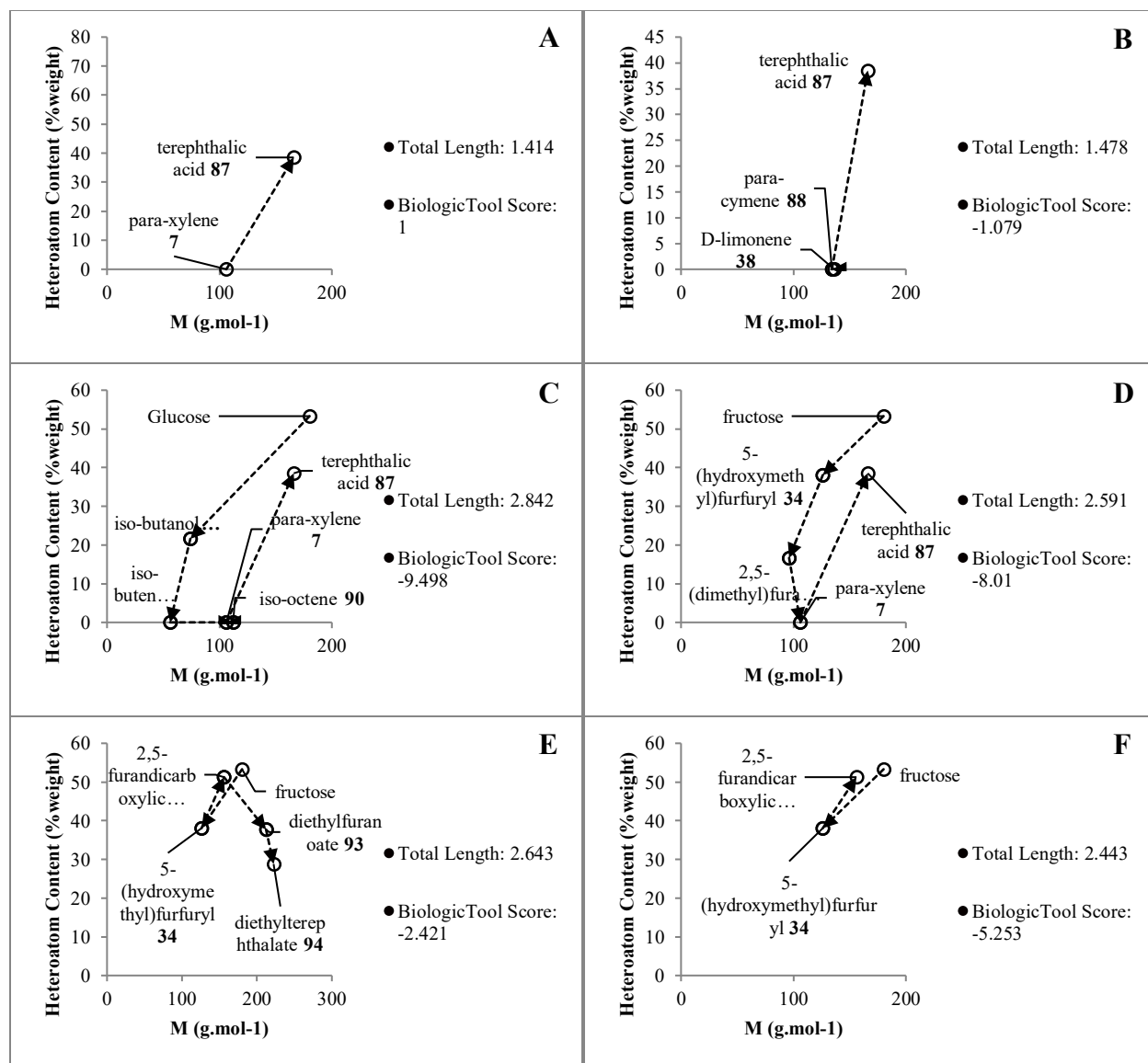
Figure 8. *BioLogicTool* plots (heteroatom content by %weight vs. molecular mass M) for the production of terephthalic acid 87 via: **A**) petroleum-derived base chemical para-xylene **7**; **B**) platform molecule D-limonene **38**; **C**) platform molecule iso-butanol **15**; **D**) platform molecule 5-(hydroxymethyl)furfuryl **34** and 2,5-dimethylfuran **91** as intermediate; **E**) platform molecule 5-(hydroxymethyl)furfuryl **34** and diethylfuranoate **93** as intermediate. **F**) Production of **92** from fructose via **34**. Note: the biomass-derived routes all start from a relevant biomass feedstock (sugar or D-limonene)

Table 2. Comparison *BioLogicTool* Scores for the Various Pathways to Terephthalic acid

Route	Feedstock	Key intermediate	Number of steps	Cumulative Pathway Length	*BioLogicTool Score* (tortuosity)
A	Petroleum	**7**	1	1.41	1.00
B	D-limonene	**88**	2	1.48	-1.08
C	Glucose	**15**	5	2.84	-9.50
D		**91**	4	2.59	-8.01
D*		**92**	4	3.30	-3.29
E[1]	Fructose	**93**	5	2.64	-2.42
F[c]		**34** and **92**	2	2.44	-5.25

[1]final product is diethylterephthalate; [b]same as D but **92** replaces **91** for Diels-Alder addition; [c]final product is 2,5-furandicarboxylic acid **92**, this a potential replacement for terephthalic acid in polyesters such as poly(ethene furanoate) (PEF)

If the direct Diels-Alder addition of diacid **92** with **1** could be achieved in high yields then this pathway becomes even better (Route **D***, Table 2) with the cumulative vector reducing to just 3.3, and low lower tortuosity retained at -3.29. This reaction has yet to be optimised, with only very small yields quoted in patents from 2008 and 2014.[18,22] A final option for maximum retention of feedstock heteroatom content is to consider replacing terephthalic acid entirely, one choice being selecting poly(ethenefuranoate) (PEF) as an alternative to PET. This pathway, Route **F**, has one of the shortest cumulative pathways (2.44) and a relatively low *BioLogicTool score* (-5.253), higher than **D*** due to the normalisation of the data. The low cumulative pathway for PEF together with the relatively low tortuosity score also represents a conclusion that was anticipated at the end of

the introduction: the *BioLogicTool* can highlight how important retention of feedstock %heteroatom content is in maximizing the mass efficiency of converting biomass to products.

OPPORTUNITIES FOR BIO-BASED AROMATICS FROM LIGNIN

The replacement of petroleum-derived aromatics by bio-based alternatives is one of the biggest challenges of the transition towards the use of renewable feedstocks in chemical production. The case of terephthalic acid above is only one of many examples of aromatic compounds extensively used in the chemical industry. Currently, 40% of the bulk chemicals belong to the aromatic category, and a myriad of products are derived from benzene, toluene and xylenes. Although some renewable aromatic feedstocks such as tannins or cardanols derived from cashew oil have already a place in the chemical industry, their abundance is not adequate to replace all petroleum-derived aromatic compounds. Fortunately, there is a large reservoir of aromatics in the support tissue of plants. Lignin, being the second most abundant biopolymer on earth, is the only source that can replace petroleum-derived aromatics at scale. However, its complexity, heterogenicity, and recalcitrance have complicated its application. Currently, only a small proportion of lignin is used to produce chemicals, vanillin and ferulic acid being the only products produced at industrial scale from it (Figure 9**B**).[23] The production of vanillin is an illustrative example of the current status of bio- *vs* petroleum-based industry: 85% of world's vanillin production comes from petroleum, while only 15% has its origin from woody biomass.[23] Moreover, the only reason why this figure is not even smaller is the premium that is paid for "natural" vanilla as opposed to "synthetic" or "artificial" vanilla. Cheap oil prices ensure that many commodity chemicals are currently produced at a price that is difficult to obtain from bio-based sources. A look at the path from benzene to vanillin shows, though, that other parameters (atom economy, greenness, sustainability etc.) favour

the isolation of vanillin from woody biomass (Figure 9A). In this case, the bio-based route is less tortuous than the petroleum-based route. Having said that, vanillin has to be obtained from depolymerization of lignin. Unfortunately, lignin is a complex polymer with a non-defined structure, and it is not possible to apply the *BioLogicTool* for this particular feedstock.
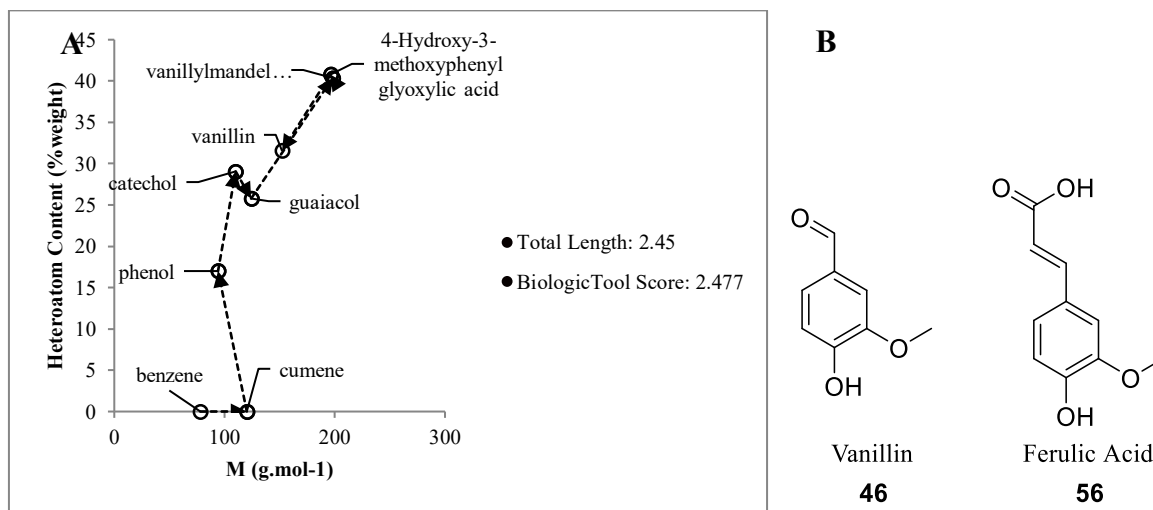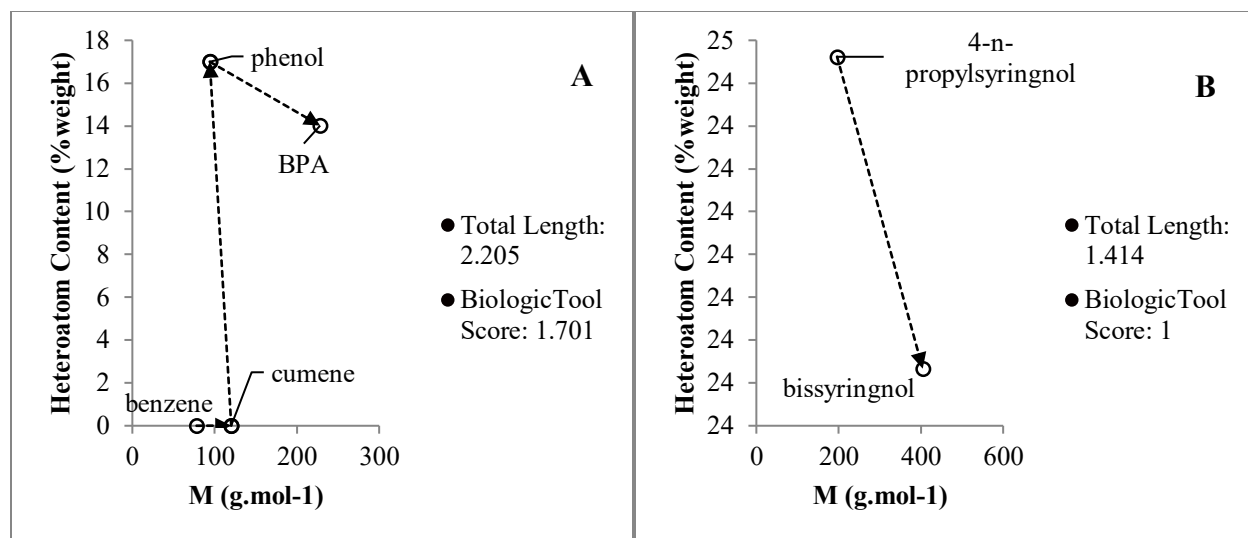

Figure 9. A: *BioLogicTool* plot for the oil-based production of vanillin **46**; **B** Monomers currently produced from lignin.

Despite its low numbers, the fact that vanillin is at present produced from lignin has driven the interest of scientists to use it as a monomer for constructing renewable polymers from it.[24] Other monomers potentially derived from lignin have also been used for the synthesis of bio-based polymers, and this has been recently reviewed.[25–28] Nonetheless, all these examples use model compounds, that is, molecules that "could" be obtained from lignin but that are currently derived from petroleum. This trick allows preparing the polymers in similar conditions to the standard petrochemical settings: high purity monomers. Unfortunately, this is not representative of what is actually obtained from lignin depolymerization, where purity is typically considerably lower.[25,29,30]

One of the most controversial aromatic petroleum derivates is bisphenol A (BPA, Figure 10**C**), a precursor for most polycarbonates and epoxy resins. Many potential replacements for BPA have been reported based on *potentially* lignin-derived monomers.[31–37] Yet, to the best of our knowledge, there is only one example where an actual lignin-derived monomer (4-*n*-propylsyringnol, Figure 10**C**) was used.[38] Moreover, the authors conducted an early-stage economic assessment which showed that the scale-up of the bissyringol (Figure 10**C**) could be economically feasible. The comparison of BPA with the lignin-based BPA replacement shows two interesting features. On one hand, that the use of lignin-derived fragments is advantageous for phenolic derivatives, as the hydroxyl group is already present. The cumulative pathway length is 1.42 for the bissryingol synthesis (Figure 10**B**) versus 2.21 for BPA (Figure 10**A**). However, as already stated, the process for obtaining lignin-derived monomer is not computed in the *BioLogicTool*. In any case, the formation of aromatics from lignin seems more logical than from sugars (vide supra). The other feature that stands out from the comparison of the BPA against its bio-based replacement is the extra functionality in the form of methoxy groups in the latter. The extra functionality has, of course, some impact on the properties of the materials.[39] Sometimes it will be desired and sometimes undesired but it can be certainly stated that lignin-derived polymers will have unique properties. In the current situation of dominant petroleum-derived polymers, these unique properties could confer an advantage to lignin alternatives and offer a unique selling point to their production.
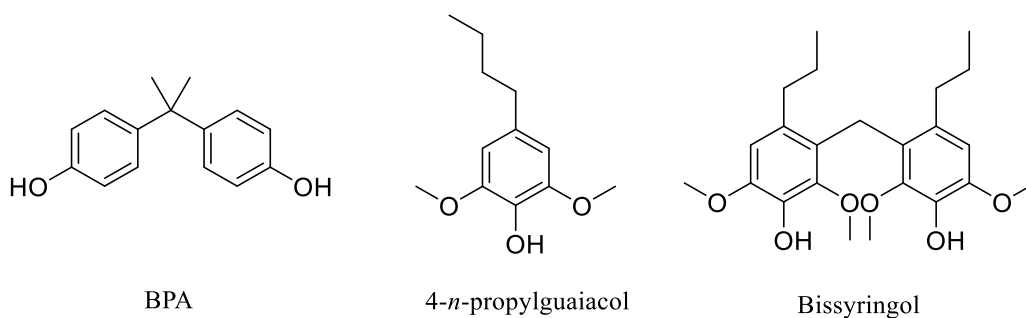
Figure 10. Comparison of BPA and the lignin-derived potential alternatives. **A**: *BioLogicTool* plot of the oil-based route to BPA; **B**: *BioLogicTool* plot of a potential bio-derived replacement for BPA, bissynrigol; **C**: Chemical structures of plotted molecules BPA, 4-*n*-propylguaiacol, bissyringol and representative dimers derived from lignin depolymerization.

Besides polymers, lignin-derived monomers can serve as a starting point for drugs and natural products.[25] An example of applying the *BioLogicTool* to consider the production of L-dopa from lignin can be found in the supplementary information.

LIMITATIONS OF THE *BIOLOGICTOOL*

The tool was specifically developed to be simple to use, giving a visual representation of data supported with relevant numerical values (*Total length* and *BioLogicTool score*) and to also be accessible prior to experimental work. However, users should also be aware of a number of potential limitations to the tool, some of which are listed below:

- *BioLogicTool Total length* and *score* (tortuosity) do not take into account the yields of the various reaction steps in any path. Users can insert values for yields in the available tool and this will colour code each arrow for a reaction based on the yield quoted (example given in Supporting Information, figure S5). Note the values generated (cumulative vector and *BioLogicTool* score) do not take these yields into account, only the visual representation of the plot is altered

- The tool does not highlight where excess reagents or reactants are required. Other metrics such Process Mass Intensity are more appropriate for this but they require the reaction to be performed in the first instance and for the user to gather data for the mass of each regent used

- The tool does not consider energetics of the reaction steps and overall pathways. It is possible that a third axis could be included to address this but this approach would require additional computational work and remove the ease to which the tool can be used

- The tool does not give direct information on whether one process is more or less environmentally friendly or economically viable to another, many additional parameters are required to make such judgements. However, there is likely to be some correlation between improved resource efficiency and a reduction in the number of steps to lower both cost and environmental impact, therefore the tool may offer some initial general assessment of this

CONCLUSION

Iterative steps in the modification of the van Krevelen diagram led to development of a new visual representation of building-block chemicals, one that is more suited to making a comparison between petrochemical and bio-based routes to desired products of the chemical industry. The higher heteroatom content in platform molecules were highlighted by the so-called *BioLogicTool* plots (%heteroatom by mass vs. molecular mass). Whether the chemical industry should seek drop-in replacements to conventional petrochemicals or embrace a whole new range of products is thus assisted by visualisation through application of the tool. If logic prevails then the future bio-based economy must seek to keep incorporated into its final products those original heteroatoms present from the starting biomass feedstock. The capabilities of the *BioLogicTool* was enhanced further by summation of the vectors for multi-step pathways from feedstock to product, whilst a simplified tortuosity score could also be easily calculated. In doing so the *BioLogicTool* allows for hypothetical routes to be assessed before any laboratory reactions takes place. The approach of the *BioLogicTool* has been demonstrated through some selected examples including options for replacing NMP, bio-based routes to terephthalic acid and the importance lignin holds in accessing aromatic compounds (especially ones that contain heteroatoms). As such the simple approach herein offers a facile and rapid means to compare established and new routes from feedstock to product, with that feedstock being either the incumbent fossil resources (petroleum, gas) or biomass.

ASSOCIATED CONTENT

Supporting Information contains: full tables of the base chemicals and platform molecules with all the values (M, O:C and H:C, %oxygen, %heteroatom) used to prepare the *BioLogicTool*

plots; a larger version of Figure 3D with assignment of the all data points; the structures of the platform molecules with an M >220 g/mol*; BioLogicTool* plots extended to contain the platform molecules with M >220 g/mol; discussion on potential benefits of the synthesis of L-dopa from lignin as highlighted by the *BioLogicTool*; an example of a colour-coded plot based on user defined yields for each step. The spreadsheet form of the tool is available as part of the supplementary information and also accessible through via DOI 10.15124/64d80432-9929-4b74-bcb7-8e3226ec03ce

AUTHOR INFORMATION

Corresponding Author

*Dr Thomas J Farmer; thomas.farmer@york.ac.uk; Green Chemistry Centre of Excellence, Department of Chemistry, University of York, Heslington, York, YO10 5DD, UK

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. T.J.F. perceived the initial idea for the tool, T.J.F. and Y.L. developed the tool and demonstrated it uses for some examples, P.O., R.V. and K.V. applied the tool to the use of lignin as feedstock.

REFERENCES

(1)   Farmer, T. J.; Mascal, M. Platform Molecules. In *Introduction to Chemicals from Biomass*; John Wiley & Sons, Ltd: Chichester, UK, 2014; pp 89–155. https://doi.org/10.1002/9781118714478.ch4.

(2)   Van Krevelen, D. W. Graphical-Statistical Method for the Study of Structure and Reaction Processes of Coal. *Fuel* **1950**, *29*, 269–284.

(3)   Hertkorn, N.; Frommberger, M.; Witt, M.; Koch, B. P.; Schmitt-Kopplin, P.; Perdue, E. M. Natural Organic Matter and the Event Horizon of Mass Spectrometry. *Anal. Chem.* **2008**, *80* (23), 8908–8919. https://doi.org/10.1021/ac800464g.

(4)   Meija, J. Mathematical Tools in Analytical Mass Spectrometry. *Anal. Bioanal. Chem.* **2006**, *385* (3), 486–499. https://doi.org/10.1007/s00216-006-0298-4.

(5)   Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram. *Anal. Chem.* **2003**, *75* (20), 5336–5344. https://doi.org/10.1021/ac034415p.

(6)   Dusselier, M.; Mascal, M.; Sels, B. F. Top Chemical Opportunities from Carbohydrate Biomass: A Chemist's View of the Biorefinery. In *Selective Catalysis for Renewable Feedstocks and Chemicals*; 2014; pp 1–40. https://doi.org/10.1007/128_2014_544.

(7)   Hülsey, M. J.; Yang, H.; Yan, N. Sustainable Routes for the Synthesis of Renewable Heteroatom-Containing Chemicals. *ACS Sustain. Chem. Eng.* **2018**, *6* (5), 5694–5707. https://doi.org/10.1021/acssuschemeng.8b00612.

(8)   Rivas-Ubach, A.; Liu, Y.; Bianchi, T. S.; Tolić, N.; Jansson, C.; Paša-Tolić, L. Moving beyond the van Krevelen Diagram: A New Stoichiometric Approach for Compound Classification in Organisms. *Anal. Chem.* **2018**, *90* (10), 6152–6160.

https://doi.org/10.1021/acs.analchem.8b00529.

(9)   Lie, Y.; Farmer, T. J.; Macquarrie, D. J. Facile and Rapid Decarboxylation of Glutamic Acid to γ-Aminobutyric Acid via Microwave-Assisted Reaction: Towards Valorisation of Waste Gluten. *J. Clean. Prod.* **2018**, *205*. https://doi.org/10.1016/j.jclepro.2018.09.151.

(10)   Lammens, T. M.; Franssen, M. C. R.; Scott, E. L.; Sanders, J. P. M. Synthesis of Biobased N-Methylpyrrolidone by One-Pot Cyclization and Methylation of Aminobutyric Acid. *Green Chem.* **2010**, *12* (8), 1430–1436. https://doi.org/10.1039/c0gc00061b.

(11)   Court, G. R.; Lawrence, C. H.; Warwick, D.; Anthony, J. Method For Converting Lignocellulosic Materials Into Useful Chemicals. WO2011/000030 Al, 2011.

(12)   Mouterde, L. M. M.; Allais, F.; Stewart, J. D. Enzymatic Reduction of Levoglucosenone by an Alkene Reductase (OYE 2.6): A Sustainable Metal- and Dihydrogen-Free Access to the Bio-Based Solvent Cyrene®. *Green Chem.* **2018**, *20* (24), 5528–5532. https://doi.org/10.1039/C8GC03146K.

(13)   Alma, M. H.; Salan, T. A Review On Novel Bio-Fuel From Turpentine Oil. *Process. Petrochemistry oil Refin.* **2017**, *18* (1), 1–12.

(14)   Colonna, M.; Berti, C.; Fiorini, M.; Binassi, E.; Mazzacurati, M.; Vannini, M.; Karanam, S. Synthesis and Radiocarbon Evidence of Terephthalate Polyesters Completely Prepared from Renewable Resources. *Green Chem.* **2011**, *13* (9), 2543–2548. https://doi.org/10.1039/c1gc15400a.

(15)   Kerton, F.; Marriott, R. *Alternative Solvents for Green Chemistry*, 2nd ed.; Green Chemistry Series; Royal Society of Chemistry: Cambridge, 2013. https://doi.org/10.1039/9781849736824.

(16)   Peters, M. W.; Taylor, J. D.; Jenni, M.; Manzer, L. E.; Henton, D. E. Integrated Process To

Selectively Convert Renewable Isobutanol To P-Xylene. US 2011/0087000 A1, 2011.

(17) Do, P. T. M.; McAtee, J. R.; Watson, D. A.; Lobo, R. F. Elucidation of Diels-Alder Reaction Network of 2,5-Dimethylfuran and Ethylene on HY Zeolite Catalyst. *ACS Catal.* **2013**, *3* (1), 41–46. https://doi.org/10.1021/cs300673b.

(18) Wang, B.; Gruter, G. J.; Dam, M. A.; Krigel, R. M. Process For The Preparation Of Benzene Derivatives From Furan Derivatives. WO 2014/065657Al, 2014.

(19) Williams, C. L.; Chang, C.-C.; Do, P.; Nikbin, N.; Caratzoulas, S.; Vlachos, D. G.; Lobo, R. F.; Fan, W.; Dauenhauer, P. J. Cycloaddition of Biomass-Derived Furans for Catalytic Production of Renewable p -Xylene. *ACS Catal.* **2012**, *2* (6), 935–939. https://doi.org/10.1021/cs300011a.

(20) Wang, D.; Osmundsen, C. M.; Taarning, E.; Dumesic, J. A. Selective Production of Aromatics from Alkylfurans over Solid Acid Catalysts. *ChemCatChem* **2013**, *5* (7), 2044–2050. https://doi.org/10.1002/cctc.201200757.

(21) Ogunjobi, J. K.; Farmer, T. J.; R., M. C.; Breeden, S. W.; Macquarrie, D. J.; Thornthwaite, D.; Clark, J. H. Synthesis of Bio-Based Diethyl Terephthalate via Diels-Alder Addition of Ethylene to 2,5-Furandicarboxylic Acid Diethyl Ester: An Enhanced Route to 100% Bio-Based Poly(Ethylene Terephthalate). *ACS Sustain. Chem. Eng.* **2019**, *accepted*.

(22) Gong, W. H. Terephthalic Acid Composition And Process For The Production Thereof. US 2009/0124829 A1, 2009.

(23) Fache, M.; Boutevin, B.; Caillol, S. Vanillin Production from Lignin and Its Use as a Renewable Chemical. *ACS Sustain. Chem. Eng.* **2016**, *4* (1), 35–46. https://doi.org/10.1021/acssuschemeng.5b01344.

(24) Fache, M.; Boutevin, B.; Caillol, S. Vanillin, a Key-Intermediate of Biobased Polymers.

Eur. Polym. J. **2015**, *68*, 488–502. https://doi.org/10.1016/j.eurpolymj.2015.03.050.

(25) Sun, Z.; Fridrich, B.; de Santi, A.; Elangovan, S.; Barta, K. Bright Side of Lignin Depolymerization: Toward New Platform Chemicals. *Chem. Rev.* **2018**, *118* (2), 614–678. https://doi.org/10.1021/acs.chemrev.7b00588.

(26) Delidovich, I.; Hausoul, P. J. C.; Deng, L.; Pfützenreuter, R.; Rose, M.; Palkovits, R. Alternative Monomers Based on Lignocellulose and Their Use for Polymer Production. *Chem. Rev.* **2016**, *116* (3), 1540–1599. https://doi.org/10.1021/acs.chemrev.5b00354.

(27) Llevot, A.; Grau, E.; Carlotti, S.; Grelier, S.; Cramail, H. From Lignin-Derived Aromatic Compounds to Novel Biobased Polymers. *Macromol. Rapid Commun.* **2016**, *37* (1), 9–28. https://doi.org/10.1002/marc.201500474.

(28) Isikgor, F. H.; Becer, C. R. Lignocellulosic Biomass: A Sustainable Platform for the Production of Bio-Based Chemicals and Polymers. *Polym. Chem.* **2015**, *6* (25), 4497–4559. https://doi.org/10.1039/C5PY00263J.

(29) Schutyser, W.; Renders, T.; Bosch, S. Van den; Koelewijn, S.-F.; Beckham, G. T.; Sels, B. F. Chemicals from Lignin: An Interplay of Lignocellulose Fractionation, Depolymerisation, and Upgrading. *Chem. Soc. Rev.* **2018**, *47* (3), 852–908. https://doi.org/10.1039/C7CS00566K.

(30) Rinaldi, R.; Jastrzebski, R.; Clough, M. T.; Ralph, J.; Kennema, M.; Bruijnincx, P. C. A.; Weckhuysen, B. M. Paving the Way for Lignin Valorisation: Recent Advances in Bioengineering, Biorefining and Catalysis. *Angew. Chemie Int. Ed.* **2016**, *55* (29), 8164–8215. https://doi.org/10.1002/anie.201510351.

(31) Zhao, S.; Abu-Omar, M. M. Renewable Epoxy Networks Derived from Lignin-Based Monomers: Effect of Cross-Linking Density. *ACS Sustain. Chem. Eng.* **2016**, *4* (11), 6082–

6089. https://doi.org/10.1021/acssuschemeng.6b01446.

(32) Meylemans, H. A.; Groshens, T. J.; Harvey, B. G. Synthesis of Renewable Bisphenols from Creosol. *ChemSusChem* **2012**, *5* (1), 206–210. https://doi.org/10.1002/cssc.201100402.

(33) Chen, Q.; Huang, W.; Chen, P.; Peng, C.; Xie, H.; Zhao, Z. K.; Sohail, M.; Bao, M. Synthesis of Lignin-Derived Bisphenols Catalyzed by Lignosulfonic Acid in Water for Polycarbonate Synthesis. *ChemCatChem* **2015**, *7* (7), 1083–1089. https://doi.org/10.1002/cctc.201500010.

(34) Ferrini, P.; Koelewijn, S.-F.; Van Aelst, J.; Nuttens, N.; Sels, B. F. Zeolites as Sustainable Catalysts for the Selective Synthesis of Renewable Bisphenols from Lignin-Derived Monomers. *ChemSusChem* **2017**, *10* (10), 2249–2257. https://doi.org/10.1002/cssc.201700338.

(35) Ménard, R.; Caillol, S.; Allais, F. Ferulic Acid-Based Renewable Esters and Amides-Containing Epoxy Thermosets from Wheat Bran and Beetroot Pulp: Chemo-Enzymatic Synthesis and Thermo-Mechanical Properties Characterization. *Ind. Crops Prod.* **2017**, *95*, 83–95. https://doi.org/10.1016/j.indcrop.2016.10.016.

(36) Janvier, M.; Hollande, L.; Jaufurally, A. S.; Pernes, M.; Ménard, R.; Grimaldi, M.; Beaugrand, J.; Balaguer, P.; Ducrot, P.-H.; Allais, F. Syringaresinol: A Renewable and Safer Alternative to Bisphenol A for Epoxy-Amine Resins. *ChemSusChem* **2017**, *10* (4), 738–746. https://doi.org/10.1002/cssc.201601595.

(37) Maiorana, A.; Reano, A. F.; Centore, R.; Grimaldi, M.; Balaguer, P.; Allais, F.; Gross, R. A. Structure Property Relationships of Biobased N-Alkyl Bisferulate Epoxy Resins. *Green Chem.* **2016**, *18* (18), 4961–4973. https://doi.org/10.1039/C6GC01308B.

(38) Koelewijn, S. F.; Cooreman, C.; Renders, T.; Andecochea Saiz, C.; Van Den Bosch, S.;

Schutyser, W.; De Leger, W.; Smet, M.; Van Puyvelde, P.; Witters, H.; Van Der Bruggen, B.; Sels, B. F. Promising Bulk Production of a Potentially Benign Bisphenol A Replacement from a Hardwood Lignin Platform. *Green Chem.* **2018**, *20* (5), 1050–1058. https://doi.org/10.1039/c7gc02989f.

(39) Harvey, B. G.; Guenthner, A. J.; Lai, W. W.; Meylemans, H. A.; Davis, M. C.; Cambrea, L. R.; Reams, J. T.; Lamison, K. R. Effects of O-Methoxy Groups on the Properties and Thermal Stability of Renewable High-Temperature Cyanate Ester Resins. *Macromolecules* **2015**, *48* (10), 3173–3179. https://doi.org/10.1021/acs.macromol.5b00496.