Green Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/greenchem

Journal Name

RSCPublishing

ARTICLE

Life Cycle Inventory improvement in the pharmaceutical sector: assessment of the sustainability combining PMI and LCA tools

Cite this: DOI: 10.1039/xoxxoooooox

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/xoxxoooooox

www.rsc.org/

Daniele Cespi,[a,b] Evan S. Beach,[a] Thomas E. Swarr,[c] Fabrizio Passarini,[b,d] I. Vassura, [b,d] Peter J. Dunn[e] and Paul T. Anastas[a]

Pharmaceutical chemicals are complex, high value added products that typically impose significantly greater impacts on the environment per kilogram compared to basic chemicals. A variety of green metrics have been developed to guide the design of chemistries and processes that are more sustainable. Among these Process Mass Intensity (PMI) was selected by the American Chemical Society Green Chemistry Institute Pharmaceutical Roundtable as the key parameter to express sustainability. However, researchers were concerned that these metrics could miss relevant factors that would be addressed by a more comprehensive Life Cycle Assessment (LCA). Lack of inventory data for many chemicals poses a significant barrier to more extensive implementation of LCA for pharmaceuticals. A cradle-to-gate LCA of Viagra™ is used to present a practical approach for constructing inventories using patent and literature data. Details of the improved inventory data were presented for four chemicals to illustrate the methodology, and highlighted the importance of considering out-sourced processing of reagents used in pharmaceutical synthesis. A more comprehensive impact assessment was conducted using ReCiPe v1.11 at both midpoint and endpoint levels. Comparison of two synthesis routes compared well with results from the simpler green metrics. An area for future work is to address the lack of characterization factors for toxicity and other impact categories for many chemicals.

1. Introduction

Pharmaceutical chemicals are complex, high- value added products that are typically produced in batch processes, require multiple processing steps, and undergo subsequent formulation and purification processes to optimize pharmacological function. Researchers have shown that pharmaceutical chemicals impose significantly greater environmental impacts per kilogram compared to basic chemical products - Cumulative Energy Demand (CED) 20 times greater and a Global Warming Potential (GWP) which is 25 times greater.[1] This difference is not surprising given the greater structural complexity of pharmaceutical chemicals but it does indicate that there are some opportunities for improvement. A previous cradle- to- gate study indicated that solvent use was responsible for ~75% of the energy use.[2] The American Chemical Society Green Chemistry Institute Pharmaceutical Roundtable (the Roundtable)[3] evaluated several metrics to guide the design of chemistries and processes that are more sustainable and proposed process mass intensity (PMI) as a key mass- based green metric.[4,5] However, it was recognized that PMI did not address concerns regarding the toxicity or health and safety of the feedstock materials or wastes produced. The Roundtable identified the integration of Life Cycle Assessment (LCA) methodologies as one of the top ten priority areas for green engineering research.[4] There are numerous methodological challenges to LCA studies of fine chemicals and pharmaceuticals.[6] These chemicals are typically produced in multi-purpose dedicated pharmaceutical plants with shared equipment and facilities. One challenge is that there is no coherent framework for characterizing the toxicological impacts of chemicals[7] but the most significant barrier is the lack of inventory data.[1,5,8,9,10] A variety of approaches have been proposed for modelling the missing inventory data. A generic input- output scheme of chemical production was proposed estimating procedures and default values for most parameters.[11] Researchers at ETH Zürich used mass and energy data from the petrochemical production of 338 chemicals to develop models that could estimate key production and emission parameters and environmental impacts directly from molecular structure.[12] Thus, companies have to choose between the potential benefits of more complex LCA studies versus the use of more practical, but less comprehensive green metrics. The authors were interested in exploring the potential differences in evaluation of alternative chemical processes based on green metrics compared to more comprehensive LCA studies. Viagra™ was selected for a case study. Sildenafil citrate, the active pharmaceutical ingredient (API) in

Viagra™ constitutes a selective inhibitor of phosphodiesterase- 5 (PDE5) and was the first agent with this mode of action to treat male erectile dysfunction.[13] A short summary of the main background information concerning the sildenafil citrate is reported in the supplementary information. The medicine was very successful commercially, creating pressure for both economic and environmental reasons to improve the efficiency of the production process. An environmentally benign process for sildenafil citrate was developed and characterized with an E- factor of 6 kg of waste/kg of API, which is comparable to basic chemicals.[14] Recently Sheldon along with chemists from Boehringer Ingelheim has independently calculated the E-factor for sildenafil citrate based on published procedures and arrived at a very similar number of 6.4.[15] Particularly noteworthy is the 99.7 % reduction in solvent use from 1540 to 5Kg per Kg API[16] which is in line with the first and fifth principles of Green Chemistry.[17] These dramatic improvements have reduced the impacts attributable to solvents and energy use, which could make other impact categories more relevant. This paper investigates other assessment methodologies, such as LCA to characterize a more comprehensive set of impacts.

2. Methodology

Several tools able to determine sustainability in the chemical sector have been developed. One of the first to be published was the previously mentioned E-factor[18,19] which led to the development of a new and fast way to determine the greenness of a process by the evaluation of the waste per unit of product. The introduction of the Atom Economy by Trost[20] helped scientists to reach targets of sustainability underlined by the second and eighth principles of Green Chemistry.[17] More recently, the Roundtable developed the PMI metric[9] which is defined as the ratio between the quantity of raw materials involved in the process (including water) and the amount of product (API), both expressed in terms of mass (**Eq. 1**).

$$PMI = \frac{total\ quantity\ of\ raw\ materials\ input\ (kg)}{quantity\ of\ product\ (kg)} \quad (1)$$

Pfizer published its original Viagra™ metrics paper (which covered several metrics including E-Factor and atom economy)[14] in 2004 before the PMI metric was developed, hence the authors wanted to use PMI to quantify the environmental benefits of moving from the medicinal chemistry route to the optimised medicinal chemistry process and then to the commercial process. In addition, the authors wanted to investigate a life cycle approach for the final two processes. Such an approach yields supporting information on human health and environmental impacts. There have been calls for life cycle approaches to be considered by designers[21] in order to better understand the overall sustainability of chemical and pharmaceutical processes.[16] A life cycle approach guarantees a holistic perspective, making the chemist more conscious of his or her own actions which go beyond the immediate responsibility.[22] LCA is a standardized methodology,[23,24] recognized as an international tool able to identify hot spots and to express sustainability. Standardization provides a general framework of the LCA, dividing the methodology in four conceptual phases:

- the *goal and scope definition*, represents the first stage in which the aim of the study is defined, identifying system boundaries (as geographical, technological and temporal) and the functional unit necessary to refer each input and output data;
- the *life cycle inventory (LCI)*, generally recognized as the more time-consuming phase of the entire LCA. It is characterized by data searching and elaboration in order to create models that are subsequently analysed;
- the *life cycle impact assessment* (*LCIA*), results in terms of ecosystem quality, human health and resources consumption are visualized in this phase, analysing each model with an appropriate and standardized method;
- the *results interpretation and improvement*, different from the others, which follow each other in a predetermined order, this is a transversal phase that occurs during the entire methodology in order to pursue continuous improvement. The results obtained by LCIA are discussed by researchers in order to detect the crucial hot-spots that should be modified to improve the process.

Although the application of LCA methodology to the chemical sector is not new,[25,26,27,28,29] its usage in the pharmaceutical area is still not a widespread practice.[10,16,30] However, different studies have already been published which demonstrate that is a consolidated approach, able to be used as screening and forecasting tool for companies and researchers involved in this sector.[1,2,30,31,32,33] Therefore, a life cycle approach was proposed in order to assess the burdens related to the environmental and human health associated with both the different routes for the production of Sildenafil citrate. LCA analysis was limited to the synthesis stage only, considering a *cradle-to-synthesis* approach. A detailed system boundaries scheme is reported below in **Figure 1**. Much of the detailed information to complete the LCI stage for the production of 1 kg sildenafil citrate was available in the literature[13,14] though some additional unpublished information was provided by Pfizer. Details of the medicinal chemistry route were taken from the patent literature.[34] Despite all of the efforts to develop accredited and viable source of data (e.g. Ecoinvent),[35] the increasing number of chemicals developed and commercialized each year makes it impossible to have a data base which is up to date and complete. This results in a lack of data to complete the LCI phase and it is well known that data quality is crucial to guarantee reliable analysis and results. In the pharmaceutical sector, in which a lot of complex organic chemicals are involved, the lack of data is particularly acute.[1,9,10] So for this reason we propose an innovative approach able to complete the *cradle-to-gate* inventories of chemicals not included in the Ecoinvent database (v.2.2)[35] which were needed to complete the LCI of the API under investigation. In order to

explain the methodology developed, four different chemicals were selected from the list of substances included in the Sildenafil citrate syntheses and used as case study: tert-butyl alcohol (TBA), *N*-methylpiperazine (NMP), pyridine (PRD) and 2-ethoxybenzoic acid (EBA). The LCI improvement can be essentially divided in two phases. First, gaps in terms of mass flows implicated in the production of each chemical were filled using approaches which have already been proposed in the literature (e.g. average process yield, amounts of emissions). Secondly, in order not to neglect energy involved during the production stage the energy consumptions were estimated by the combination of data reported in literature with results obtained from a model able to predict different parameters from molecular structure. A detailed description of each phase is reported in the next two paragraphs. **Figure 2** reported briefly the LCI improvements approach.

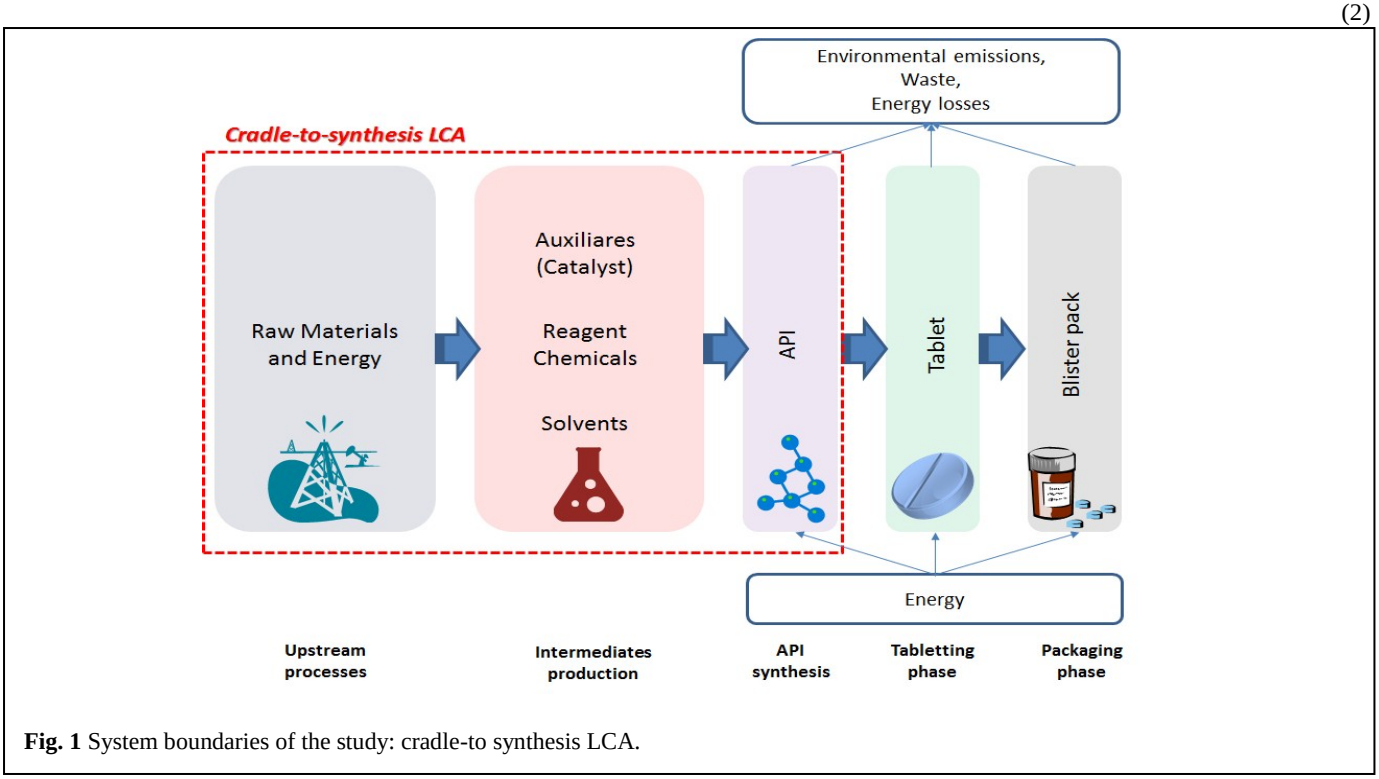
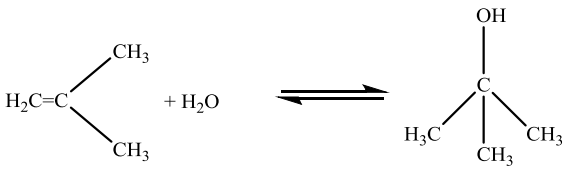


Fig. 1 System boundaries of the study: cradle-to synthesis LCA.

2.1 LCI improvement – Phase I

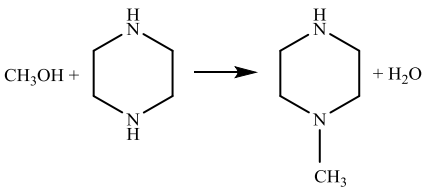
As previously discussed, the first phase of the improvement was represented by the identification of all mass flows involved in the synthesis of each substance. This first stage was carried out using standard approach and estimation already published: a process efficiency of 95% over the entire stoichiometric equation, air emissions estimated to be 0.2% of the input, and the water emissions were evaluated as a difference between unreacted reagents and air releases.[36] Also, in line with literature,[8] the use of average data proposed by Hischier et al.[36] is justified by the fact that the four substances being studied are basic and not advanced chemicals, which would inevitably require further details. In order to select the most consolidated production method for each chemical, different references including encyclopedia (e.g. Ullmann and Kirk-Othmer),[37,38,40,42,43] patents and database (SciFinder®)[41] were consulted. A summary description of each process is now reported.

- *tert-Butyl alcohol*, due to the impossibility to produce tertiary alcohol via oxo-synthesis, the main process involves the direct hydration of 2-methylpropene at 190-245°C using metal oxide catalysts (e.g. tungsten oxide).[37]
  The equilibrium is reported below (**Eq. 2**).



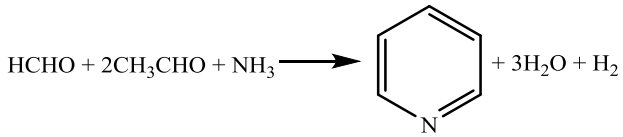
(2)

- *N-Methylpiperazine*, commonly produced from alkylation of piperazine with methanol in the presence of reductive amination catalyst (e.g. titania modified with phosphorus).[38,39] The reaction (**Eq. 3**) is usually conducted between 290-340°C and 200-1500psig.[39]



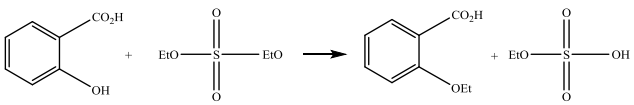
(3)

- *Pyridine*, first obtained from tar, is nowadays produced by the reaction of acetaldehyde and formaldehyde with ammonia at 350-550°C in the presence of silica-alumina catalyst (**Eq. 4**).[40]



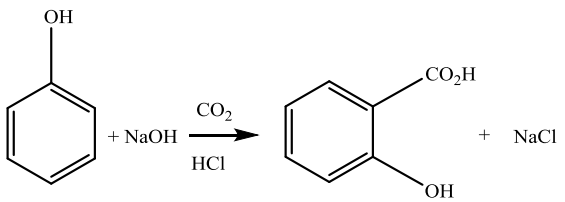
(4)

- *2-Ethoxybenzoic acid*, SciFinder® database[41] shows several production routes derived from salicylic acid and diethyl sulfate. A general framework for the reaction is reported below (**Eq. 5**).



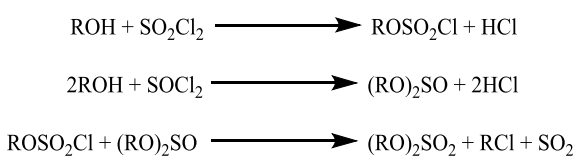
(5)

However as both salicyclic acid and diethyl sulfate were also absent from the data base we had to consider how these materials were made. Salicylic acid is commonly produced by the Kolbe-Schmitt process, which involves the reaction between phenol and carbon dioxide (**Eq. 6**).[42]



(6)

Dialkyl sulfates can be synthesized starting from alcohols following the reactions scheme reported below (**Eq. 7**).[43]



(7)

As reported in previous studies[26,27] the impact of catalyst production is relatively small, especially for these inorgainc type catalysts and hence was not included in this study. Also proxy data regarding water consumption, infrastructure and transportation were neglected. A detailed description of the LCI for each substance is showed in the supporting information section (**Tab. S1**).

2.2 LCI improvement – Phase II

Although getting reasonable data on mass flows is possible from the patent and publication literature, obtaining reliable estimates of the energies used is much more difficult and hence different solutions have been proposed, as for example: i) models able to simulate parts of chemical reactions,[44,45] that need detailed information on equipment and plants conditions;[8] ii) a standard approach able to predict average energy consumption involved per kg of base chemical produced (e.g. of 2MJ in the form of steam and 0.333kWh of electricity).[36] The latter was developed estimating average data collected in the internal environmental report of a large chemical plant situated in Germany.[46] Nevertheless collaboration with companies seems to be the best solution in order to carry out reliable life cycle analysis, but this route is also not always easy as it may be difficult to access the desired reports such as the Process Economics Program (PEP)[47] and the Process Economics and Research Planning (PERP)[48] reviews. So for this reason and in order to complete the LCI for the substances avoiding a cut-off principle for the energy consumption, a new approach to estimate these parameters combining literature with information extrapolated from a Molecular Structure-based Model. The FineChem tool developed by Gregor Wernet was selected in this study. The model, based on artificial neural networks is able to estimate direct correlation between molecular structures and some key production and emissions parameters.[8,12,49] The developers intent was not to replace traditional inventory analysis,[49] but to provide a screening tool in order to support LCA practitioners. The LCI is derived by the use of ten chemical descriptors:[12]

- Molecular weight
- Number of functional groups
- Number of oxygen atoms in carbonyl groups (keto and aldehyde)
- Number of oxygen atoms except those in carbonyl groups
- Number of nitrogen atoms
- Number of halogen atoms
- Number of rings (both aromatic and aliphatic)
- Number of tertiary and quaternary carbon atoms
- Number of heteroatoms in rings
- Number of unique substitutes on aromatic rings

FineChem is able to predict the Cumulative Energy Demand (CED) and Global Warming Potential (GWP) values per kg of substance analysed. However, while predictions of the CED are quite accurate the GWP model performs less well.[12] For that reason, LCI filling of chemicals was carried out just taking into account the CED values in the improvements, neglecting the carbon emissions in the inventories. In a recent study,[30] the CED values were extrapolated to provide a screening indicator predicting the environmental burdens. The work described in this paper takes a different approach and the CED values are used to obtain a more complete LCI for each substance investigated. All the inputs necessary to run the FineChem

tool and the outputs obtained from the simulation, are reported in the supporting information section (**Tab. S2**). However, as reported in the literature,[50] the CED values refer to "*the entire demand, valued as primary energy, which arises in connection with the production, use and disposal of an economic good*". Therefore, in line with system boundaries proposed for the LCI improvements just the energies involved in the cradle-to-gate of each substance should be included in the study. Literature[51] estimates indicated that the contribution of energy consumption involved in the upstream processes (raw material extraction and substance production) for organic chemical sector is around 45-69% of the total CED value. Therefore, the average 57% value of the CED values (predicted by Finechem) was taken into account as of average energy consumption in the manufacturing of chemicals. Also, in order to refine this evaluation, an average energy mix in a European chemical plant[46] was considered to split values previously obtained: 50% natural gas, 38% electricity and 12% steam. The same energy mix was taken into account by the developers of Ecoinvent database in order to simulate energy consumption for chemicals products without primary information.[52] The values considered as average simulation of energy involved in the production of each substance are listed in **Tab. S3**. Full details of the methodology used to generate the improved LCIs are given in the supporting information using N-methylpiperazine as an example.

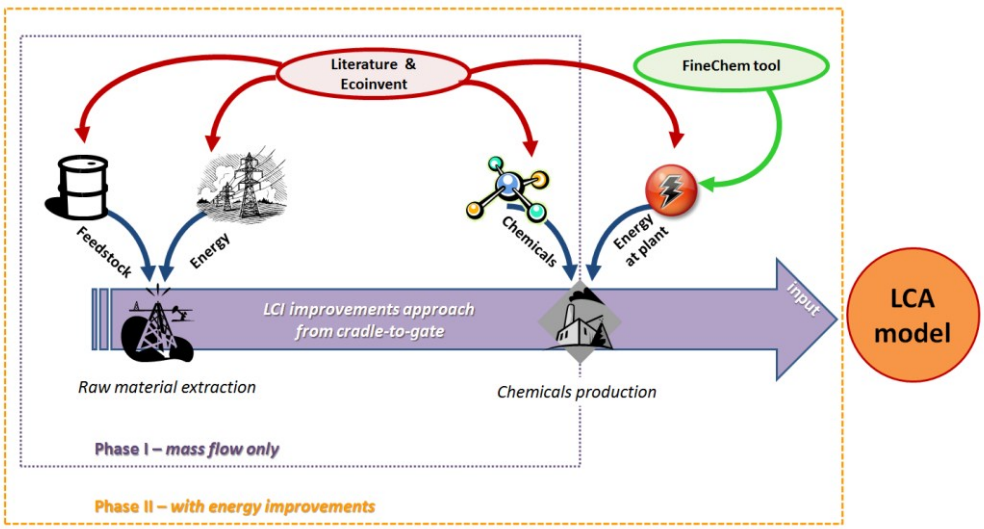


Fig. 2 Approach developed to fill the lack of data.

3. Results and discussion

3.1 Process Mass Intensity evaluation

A PMI evaluation was conducted to validate the life cycle model and provide a baseline assessment of improvements achieved by Pfizer with the development of the Commercial route compared to an Optimised Medicinal Chemistry route. The objective is to determine how well the PMI assessment aligns with a more comprehensive life cycle impact assessment (LCIA.) The initial medicinal synthesis route was extremely inefficient, with a PMI of 17,675. This is typical of preclinical drug development and does not provide a reasonable baseline for evaluating greener alternatives.[15] Four years of internal work yielded the Optimised Medicinal Chemistry process which provided material to fuel phase I, phase II and early phase III clinical trials. This process had a PMI of 116. A further reduction in PMI was achieved by introducing a new synthetic route (the Commercial route) which had a PMI of 60, a reduction of almost 50 %. This results is in agreement with an independent analysis which has recently been reported in the literature.[15] The PMI figure of 60 is much higher than the reported E-Factors of 6[14] and 6.4[15] and there are three reason for this:

- PMI includes water use whereas the two E-Factors do not.[1]
- PMI includes the mass of the product whereas E-Factor does not (so the perfect process with no waste has a PMI of 1 and an E-Factor of zero).
- The reported E-Factors take into account solvent recovery as recommended by Sheldon[53] whereas PMI does not.

However, the result is in reasonable agreement with the complete E factor, or cEF, which includes solvents and water.

These results also support the finding of Roschangar et al[15] that outsourcing of some input materials can significantly affect the calculated green metrics. Their case study of Viagra™ showed the cEF increased from 50.3 to 85.5 when considering the impacts of sourced chemicals. In order to verify which class of substances contributes mostly in terms of process sustainability, the cumulative PMI values were split in its major components: reagents, solvents and water and the results are shown in **Figure 3**. The bar chart shows improvements in all classes, but solvents appear to be the major

contributor. The Commercial route reduced solvents usage by 66% in terms of PMI. This is consistent with data already published by Pfizer, in which a reduction of 88 liters of solvent per kg of API was estimated.[14] The relative improvements are shown more clearly by the pie chart in **Figure 4**. Analysis conducted on the Optimised Medicinal Chemistry route confirms the higher contribution of solvents to the cumulative PMI (50%). The ratios of solvent (50 %), water (31 %) and reagents (19 %) are broadly similar to those reported from a survey of several pharmaceutical companies by the Pharmaceutical Roundtable.[9] In contrast the results for the Commercial route show the percentage solvent usage is lower (34 %), percentage water use is higher (44 %) and reagents slightly higher (22 %) though of course much lower in absolute terms. This relative switch from solvent to water use, reduces the environmental impact of the process. Solvents will have a typical carbon footprint of 1-3 Kg CO_2 eq. per Kg[35] whereas water will have a carbon footprint close to zero. Subtracting water from the PMI results yields values of 80.2 for the Optimised Medicinal route and 33.3 for the Commercial route enabling further comparison with values previously reported in literature.[54]

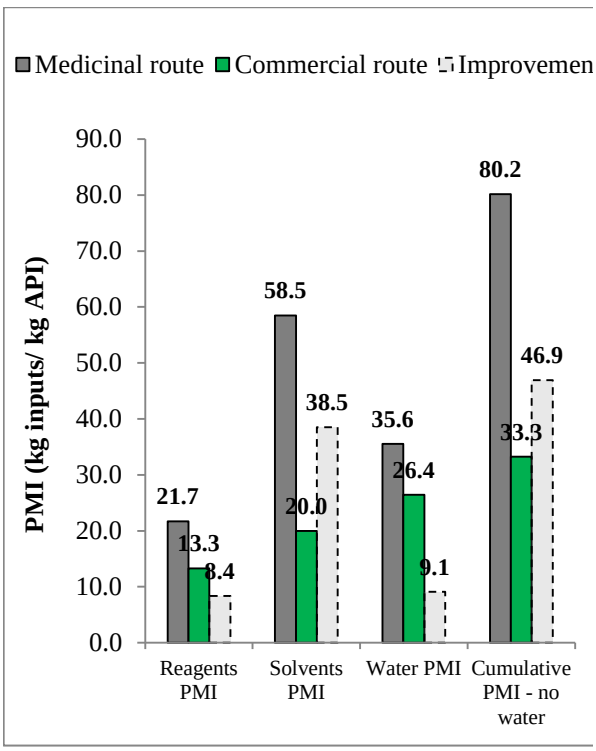


Fig. 3 PMI trend for both the Optimised Medicinal Chemistry and the Commercial routes: components of the cumulative value and scores not considering water in the calculation.

Also, as a confirmation of results obtained, cumulative PMI values for both routes to synthesize sildenafil citrate (Medicinal and Commercial) were plotted and compared with other scores already reported in literature (**Figure 5**). PMI values evaluated by ACS Roundtable[3] and those already published by Constable et al.[54] were selected as a basis for comparison with the sildenafil case study. The category reported as "specific chemistries" in the figure refers to the PMI values estimated by Constable et al.[54] and it includes a wide range of different single-step reactions (e.g. hydrogenation,

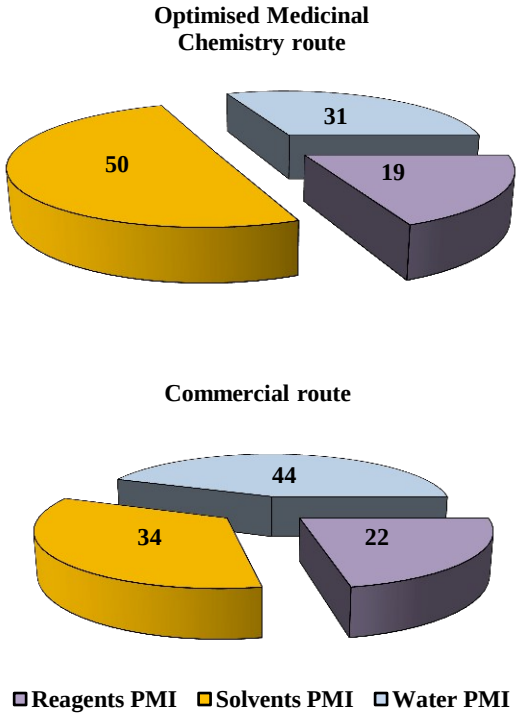


Fig. 4 Percentage contribution to the cumulative PMI for both routes.

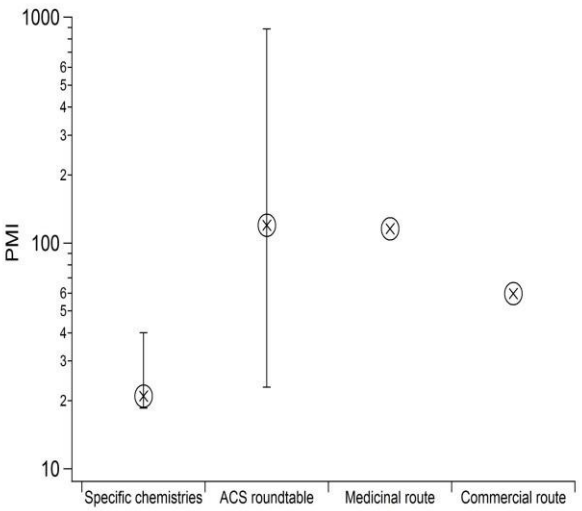


Fig. 5 Comparison of different PMI values for several products.
As a comparison, PMIs values (not including water) for different chemistries already published were selected.[54] Other values represents the cumulative PMIs estimated by ACS Roundtable for different APIs. In both case the range and the average score were considered.[3]

esterification, alkylation, chlorination, etc.). The figure shows direct correlation between molecular complexity, expressed in terms of several stages involved in the synthesis, and PMI values. PMI scores obtained from assessment of several one-step reactions are significantly lower with a smaller interval, while values estimated for different API show a much wider range. The relatively high PMIs estimated for both routes developed by Pfizer can be explained by the fact they are the cumulative values (including water) and refer to the entire synthesis chain, which includes chemical reactions to obtain the target molecule.

Moreover, the most significant aspect emerged from the figure is that both scores for sildenafil citrate are included in the range evaluated by the Roundtable for different APIs. This can be considered a satisfying verification of the reliability of the analysis carried out and of the method's robustness.

3.2 LCI improvements and LCA analysis

Although PMI showed a substantial improvement for the Commercial route, there is concern that the simple metrics could miss impacts shifted to other environmental concerns or to upstream suppliers. A better understanding can be reached using a life cycle approach. A cradle- to- gate model was constructed using the latest version of SimaPro software released (v.8.0.4.26).[55] Impacts were analyzed using the ReCiPe Endpoint (I) / World (I/A) (v.1.11)[56], the IPCC 2013 (v.1.00)[57] and the CED (v.1.09).[58] IPCC 2013 and CED are single issue assessments that could be considered comparable to the simplified green metrics. These would be expected to show similar trends to green metrics based on mass and energy. ReCiPe is a more comprehensive impact assessment method that models 18 impact categories at the midpoint level. These are mapped to endpoint damages in three areas of protection - human health, ecosystems and resources.

The IPCC (Intergovernmental Panel on Climate Change) and CED (Cumulative Energy Demand) were selected as good indicators of environmental sustainability for two reasons: i) they are well recognized and easy understandable to wider audience; ii) environmental sustainability of the chemical and pharmaceutical sectors should be evaluated in terms of global perspective. A 20-years time horizon was used (IPCC 2013 GWP 20a) to better reflect the economic pressures on corporate sustainability initiatives. As previously discussed, the first step of conducting a *cradle to synthesis* life cycle analysis for sildenafil citrate was to improve the LCI of some chemicals by adding the energy component, these results are reported in **Figure 6**. The pink bars represent results achieved in terms of GWP and CED just considering the first stage of improvements, which includes all the mass flows involved in each process evaluated by the average parameters developed by Hischier et al.[36] On the other hand, the blue bars show the entire LCI improvement stage, with the estimation of energy consumptions during the production phase. This visualization is able to show the effect of LCI improvements for each substance, showing how scores change if energy consumption is included or not in the inventory. The increase in impacts is dramatic from 120.7 to 363.2% for the GWP and from 100.8 to 227.5% for the CED, as a confirmation of the importance of including complete inventories in LCA studies regarding chemicals, class of substances for which the production phase represents the most relevant stage from an environmental and human health perspectives.[1] As shown the trend in improvement is not the same for each substance, but it is strictly related to first stage of improvement: the higher the amout of mass involved in the synthesis, due to process complexity (and also to the lack of data in the database) the lower will be the increase due to the energy consumptions (e.g. in the case of EBA). Error bars are also reported in the figure in order to show how scores from the entire improvement can change considering the range of each value provided by FineChem tool. These bars are a further confirmation of the importance in considering the entire LCI of chemicals. Indeed, as shown by the results, even considering just the minimum values for each substances scores reached in terms of GWP and CED are significantly higher than

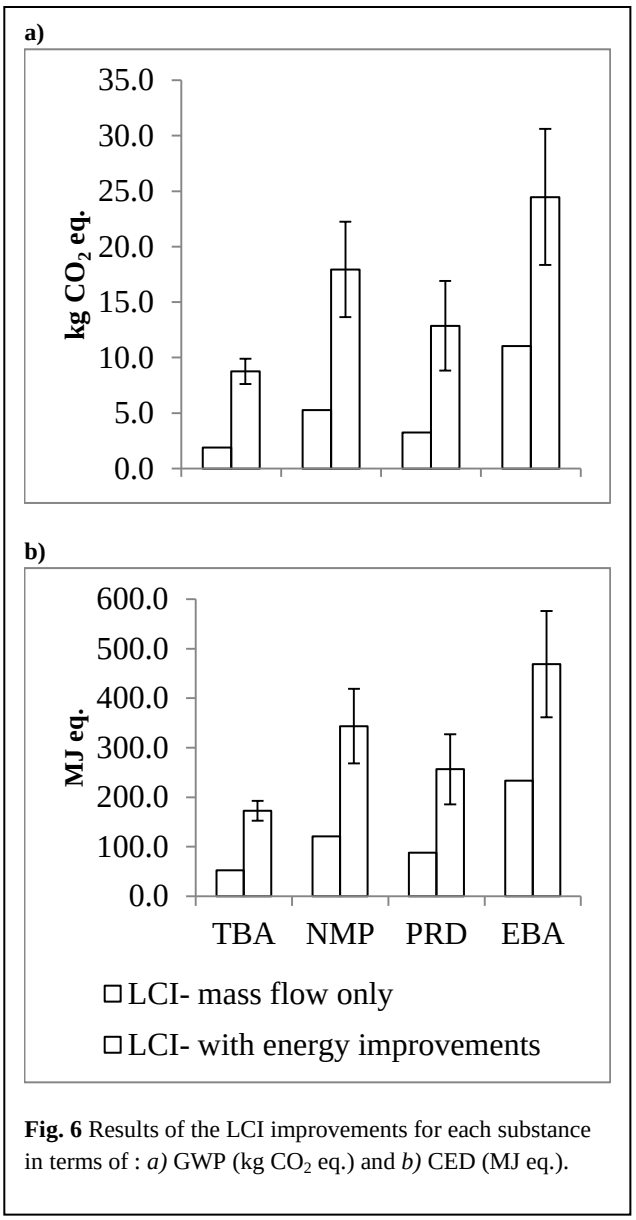


Fig. 6 Results of the LCI improvements for each substance in terms of : *a)* GWP (kg CO_2 eq.) and *b)* CED (MJ eq.).

results from the first stage of improvements. LCIA results for each substance, as well as percentage increase and MIN/MAX values, are collected below in **Table 1**.

	TBA	NMP	PRD	EBA
GWP - kg CO_2 eq.				
LCI- mass flow only	1.9	5.3	3.3	11.1
LCI- with energy improvements	8.8	17.9	12.9	24.5
Increase (%)	363.2	237.7	290.9	120.7
MIN - LCI- with energy improvements	7.6	13.6	8.8	18.4
MAX - LCI- with energy improvements	9.9	22.3	16.9	30.6
CED - MJ eq.				
LCI- mass flow only	52.8	121.1	87.8	233.5
LCI- with energy improvements	172.9	343.5	256.6	468.8
Increase (%)	227.5	183.6	192.2	100.8
MIN - LCI- with energy improvements	152.8	268.1	185.7	361.5
MAX - LCI- with energy improvements	192.9	419.1	327.1	576.4

Table 1 LCIA results for each substance LCI improvement phases.

Moreover, in order to evaluate the contribution of the different energy sources involved in each production process the CED values were disaggregated as reported in the **Figure 7**. Results of the analysis confirm the large contribution of the non-renewable energies among the entire production chains. In particular fossil fuels and nuclear together represent more than 90% of the total requirements for each substance, due mainly to the mix in order to provide electricity and thermal energy at the chemical plant. European electricity production mix was considered in order to simulate the electric utilities of a common chemical plant located everywhere in Europe (see paragraph 2.2). However, should be noticed that each country chooses its own energy mix, which can result in a variation of the scores depicted in **Figure 7**. A detailed analysis of the resources availability and connections between European member states (e.g. import and export) seems necessary to deal with all the three spheres of sustainability: economic, social and environment. In order to help the reader to better understand this point, a further contribution analysis of the disaggregated CED values for the different energy mixes for some European countries was ran showing results in the supporting information (**Fig. S1**).

Unfortunately, as shown by the **Figure 7** the use of renewable sources is quite limited, because the chemical sector requires large amounts of energy provided continuously. The contribution of alternative sources is mainly limited to the use of hydro and biomass for the production of electricity, while other sources such as solar and wind do not reach relevant levels significant enough to show in the figure. However, switching to the renewables, either through the installation of integrated systems or the use of alternative service providers for the electricity, could lead substantial benefits both to the environment (e.g. resources preservation and climate change mitigation) as well as to the human health (see discussion below). Furthermore, the usage of fossil fuel based resources in the chemical sector is also linked with the raw materials fed into processes which are mainly based on the petrochemical industry.

The ReCiPe method was selected to provide a more comprehensive assessment of life cycle impacts, because it models both at midpoint and endpoint levels. Also in this case a 20-years time horizon was used (World ReCiPe I/A). A comparison of the different synthesis routes at the midpoint level provides a quick assessment of potential burden shifting (either to other impact categories or to different parts of the product system). The midpoint characterization results are shown in **Figure 8**. The improved Commercial route dominates the comparison across all impact categories. In fact, in most categories, the relative improvement was greater than indicated by the PMI metric.

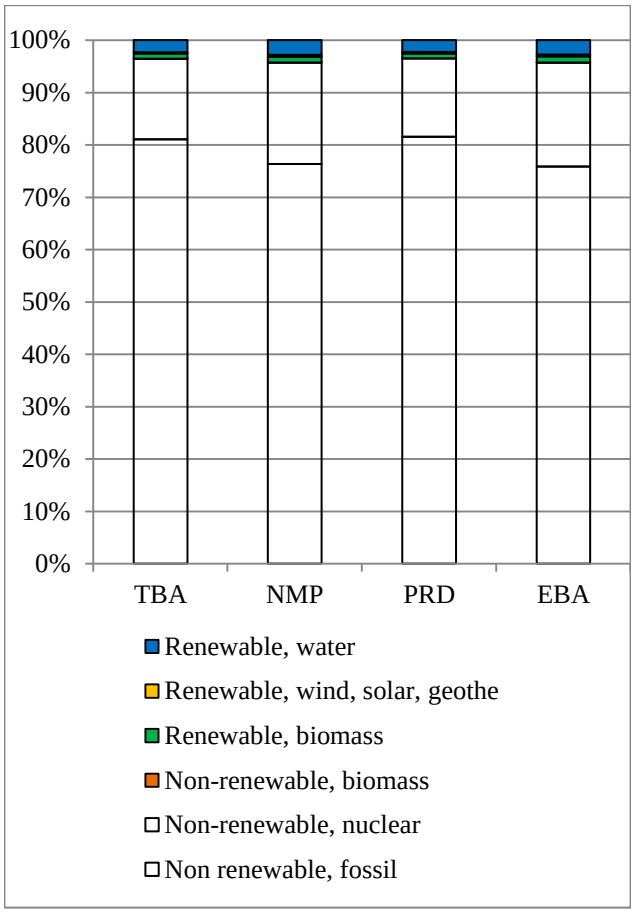


Fig. 7 LCI-with energy improvement, contribution to the CED values for each substance.

ReCiPe aggregates results for ozone depletion, human toxicity, ionizing radiation, photochemical ozone formation (smog), particulates and climate change to assess damage to human health in units of disability adjusted life years (DALYs). DALYs represent the sum of years of life lost (YLL) and years of life

lived as disabled (YLD) due to the onset of a disease (e.g. cancer). Human health was selected as the area of protection of highest priority to the pharmaceutical sector, and therefore, an appropriate indicator of sustainability. Previous researchers have noted that companies focus on improving the operations within their own facilities and could underestimate the impacts occurring in upstream suppliers. It is important to verify that green chemistry initiatives do not simply shift the burden on to other parts of the product system. The results shown in **Figure 8** above do not indicate any such shifting.

A final comparison of the alternative synthesis routes is provided in **Figure 9**, comparing the PMI metric, the single issue indicators, and the results for human health (HH) as potential indicators of sustainability. CED shows the closest agreement with PMI, which is to be expected since the metric includes the embodied energy of materials and process energy. The improved inventory can explain the observed difference in that the mass metric underestimates the benefit of energy improvements.

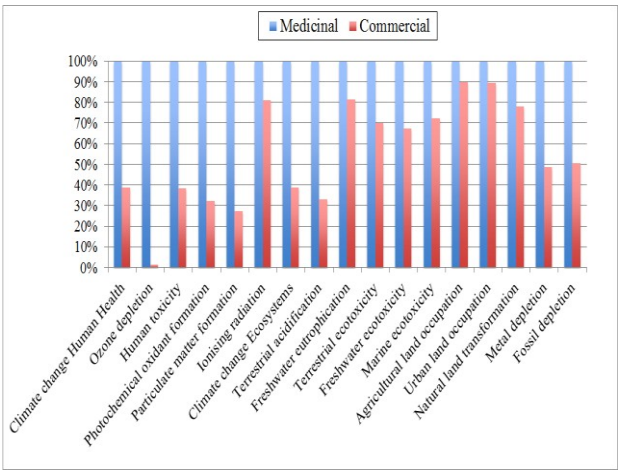


Fig. 8 Midpoint Comparison of Synthesis Routes.

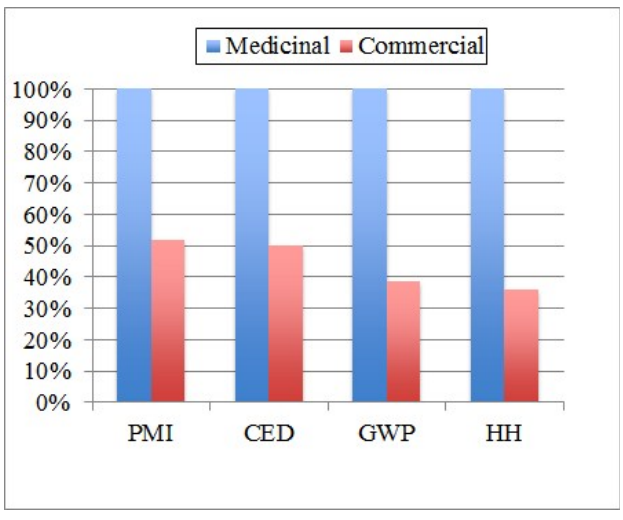


Fig. 9 Comparison of Sustainability Indicators

The GWP metric would depend in part on assumptions about the energy mix, and therefore would need to be assessed based on the specifics of a particular product system.

Although each substance could contribute to the HH in a different way, a detailed analysis carried out using the network tool provided by the SimaPro software, revealed that energy inputs for each chemical are the key drivers in order to evaluate sustainability of the production phase. The chlorine supply chain was also a significant contributor for Cl containing compounds. See the network diagrams reported in the supporting information section (**Fig. S2-S5**).

The results for HH, however, could be an artifact of missing characterization factors. There are many emissions in the inventory that were not included in the results because of these gaps. In particular, it was not possible to adequately address the potential impact associated with a release of chemicals during the synthesis of the API (e.g. air, water and soil emissions) consistent with the system boundaries indicated in **Figure 1**. In fact, as highlighted in the introduction and confirmed by literature,[1] most of pharmaceutical compounds are obtained using batch procedures which result in a difficult collection of data for several products and much more for each step involved in the production process. This is an area for future work to fill these data gaps, but the results do not suggest any shifting of the impacts to other parts of the system.

A detailed analysis conducted by Pfizer showed that the commercial process was mainly improved in terms of solvents usage.[14] In particular main efforts regarded the substitution of dichloromethane and acetone, as well as the substantial reduction of toluene involved in the synthesis. These trend seems to be in line with the results emerged from the contribution analysis carried out for the Optimised Medicinal Chemistry route. This evaluation confirms that the solvents involved in the synthesis are the main contributors to all impact categories selected. In particular, as shown in **Table 2**, dichloromethane, acetone and toluene seem to have the greatest negative effects on the ecosystem, expressed in terms of higher percentage contribution to the impact categories. These results also are a confirmation of what was obtained previously from the detailed analysis of PMI,

Optimised Medicinal Chemistry Route			
Percentage contribution to:	HH	GWP	CED
Dichloromethane	45	40	23
Acetone	14	17	20
Toluene	6	8	13
Remaining processes	35	35	44

Table 2 Optimised Medicinal Chemistry Route, main substances contribute to all three impact categories.

showing that the choice of solvent and its amount in the process is crucial for companies to reach sustainability targets. The significance of energy inputs and solvent usage provide a further confirmation of the importance to include the CED evaluation as additional method to evaluate the LCA model. Moreover, several studies proposed CED as a good tool to assess sustainability in the chemical sector both in the case of raw materials assessment,[59] but also as good indicator of the overall impact[51] showing wide range of environmental impacts.[60]

Conclusions

As suggested by The Roundtable, PMI was used in order to express sustainability in the production of sildenafil citrate comparing two alternative routes. In particular, the study highlights its importance as a screening tool, able to predict the hot spots in the API synthesis and pointing out where the main efforts should be directed in order to minimize environmental impact. The study also showed a practical approach to constructing more comprehensive life cycle models using patent, literature data and a Molecular Structure-based Model tool. The results highlighted the importance of a life cycle perspective for PMI assessments to adequately account for the downstream processes involved in the drug synthesis. In addition, a LCA perspective is necessary to obtain a holistic vision of the entire process to effectively identify priority areas for procedures to reduce loads and reach targets of sustainability expressed by the green chemistry principles. Different approaches were developed to address data gaps. Results from this study show that reasonable models can be constructed based on available literature and expert judgment. The initial results support the value of the simplified green metrics as a screening tool. The preliminary LCA models can also be a valuable tool to screen potential sourcing alternatives.
As already outlined, this approach still has some gaps due to the lack of characterization factors for the substances not contained in the database; this aspect will be the subject of follow-up studies. The preliminary model can provide valuable information to focus effort on developing the most relevant characterization factors. Thus, the absolute values for the toxicity impact categories must be viewed with caution, and the model will need to be further evaluated to better understand which emission streams are being excluded.

Abbreviations

API	Active Pharmaceutical Ingredient
CED	Cumulative Energy Demand
DALY	Disability Adjusted Life Years
EBA	2-Ethoxybenzoic acid
ETH	Eidgenössische Technische Hochschule
GWP	Global Warming Potential
HH	Human Health
IPCC	Intergovernmental Panel on Climate Change
LCA	Life Cycle Assessment
LCI	Life Cycle Inventory
LCIA	Life Cycle Impact Assessment
NMP	*N*-Methylpiperazine
PEP	Process Economics Program
PERP	Process Economics and Research Planning
PDE5	Phosphodiesterase-5
PMI	Process Mass Intensity
PRD	Pyridine
TBA	Tert-butyl alcohol
YLD	Years of Life lived as Disabled
YLL	Years of Life Lost

Acknowledgements

We would like to thank Dr. Stavros Papadokonstantakis and Dr. Sara Badr from the ETH Zürich - Institute for Chemical and Bioengineering, for their support with the FineChem tool.

Notes and references

[a] Center for Green Chemistry and Green Engineering, Yale University, 225 Prospect Street, New Haven, CT, United States.
[b] Department of Industrial Chemistry "Toso Montanari", Bologna University, Viale del Risorgimento 4, 40136 Bologna (BO), Italy.
[c] School of Forestry & Environmental Studies, Yale University, 195 Prospect Street, New Haven, CT, United States.
[d] Centro Interdipardimentale di Ricerca Industriale - Energia e Ambiente, Via Angherà 22, 47900 Rimini (RN), Italy.
[e] Pfizer Global Supply, Discovery Park, Sandwich, Kent, United Kingdom.

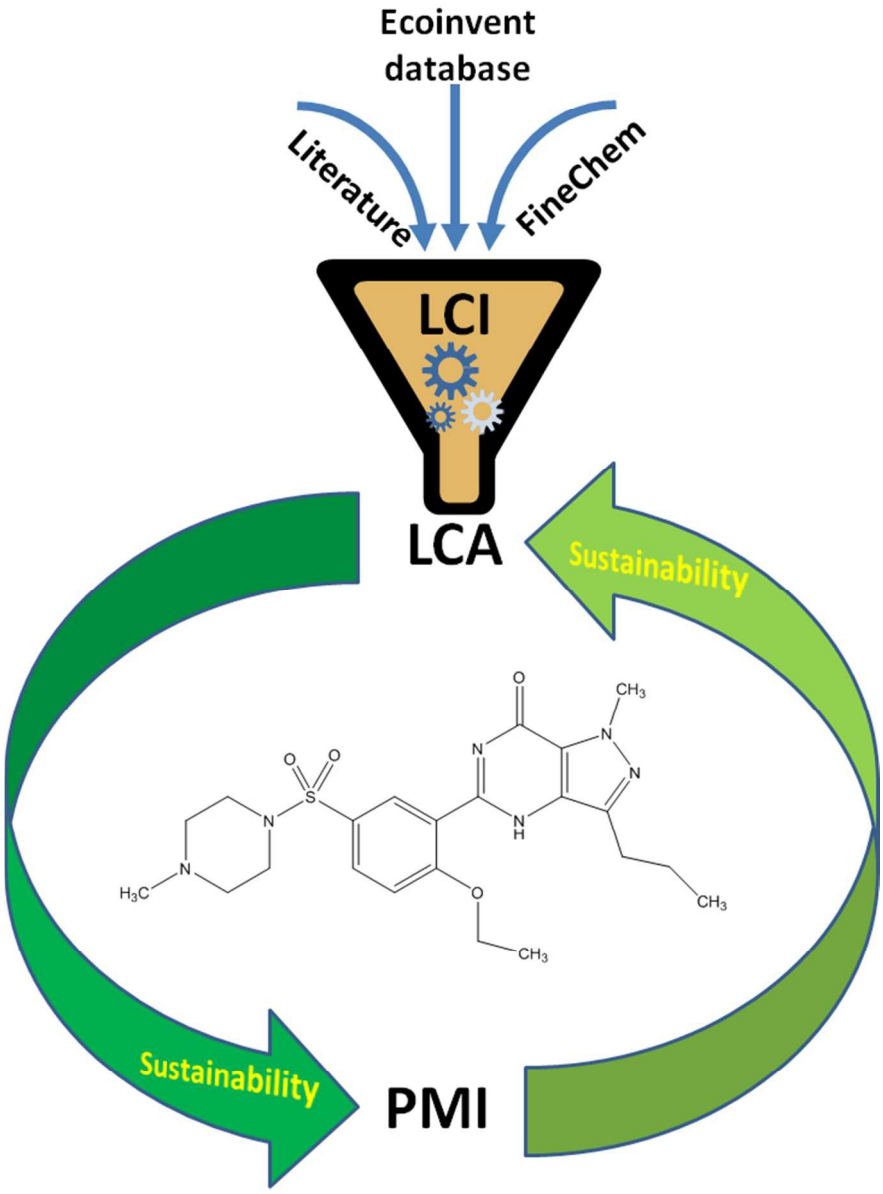
References

1. G. Wernet, S. Conradt, H. P. Isenring, C. Jiménez-González, K. Hungerbühler, *Int. J. Life Cycle Assess.*, 2010, **15**, 294-303.
2. C. Jiménez-González, A. D. Curzons, D. J. C. Constable, V. L. Cunningham, *Int. J. Life Cycle Assess.*, 2004, **9**, 114-121.
3. ACS GCIPR - American Chemical Society, Green Chemistry Institute, Pharmaceutical Roundtable website www.acs.org/gcipharmaroundtable, 2008.
4. C. Jiménez-González, P. Poechlauer, Q. B. Broxterman, B. S. Yang, D. am Ende, J. Baird, C. Bertsch, R. E. Hannah, P. Dell'Orco, H. Noorman, S. Yee, R. Reintjens, A. Wells, V. Massonneau, J. Maley, *Org. Process Res. Dev.*, 2011, **15**, 900-911.
5. C. Jiménez-González, C. Ollech, W. Pyrz, D. Hughes, Q. B. Broxterman, N. Bhathela, *Org. Process Res. Dev.*, 2013, **17**, 239-246.
6. L. M. Tufvesson, P. Tufvesson, J. Woodley, P. Börjesson, *Int. J. Life Cycle Assess.*, 2013, **18**, 431-444.

7. H. Baumann, A.-M. Tillman, *The Hitch Hiker's Guide to LCA: An Orientation in Life Cycle Assessment Methodology and Application*, Studentlitteratur AB , Lund, Sweden, 2004.
8. G. Wernet, S. Hellweg, K. Hungerbühler, *Int. J. Life Cycle Assess.*, 2012, **17**, 720-728.
9. C. Jiménez-González, C. S. Ponder, Q. B. Broxterman, J. B. Maley, *Org. Process Res. Dev.*, 2011, **15**, 912-917.
10. C. Jiménez-González, M. R. Overcash, *Green Chem.*, 2014, **16**, 3392-3400.
11. G. Geisler, T. B. Hofstetter, K. Hungerbühler, *Int. J. Life Cycle Ass.*, 2004, **9**, 101-113.
12. G. Wernet, S. Papadokonstantakis, S. Hellweg, K. Hungerbühler, *Green Chem.*, 2009, **11**, 1826-1831.
13. D. J. Dale, P. J. Dunn, C. Golightly, M. L. Hughes, P. C. Levett, A. K. Pearce , P. M. Searle, G. Ward , A. S. Wood, *Org. Process Res. Dev.*, 2000, **4**, 17-22.
14. P. J. Dunn, S. Galvin, K. Hettenbach, *Green Chem.*, 2004, **6**, 43-48.
15. F. Roschangar, R. A Sheldon, C. H. Senanayake, *Green Chem.*, Advance article. DOI 10.1039/c4g01563k.
16. P. J. Dunn, A. S. Wells, M. T. Williams, *Green Chemistry in the Pharmaceutical Industry*, Wiley - VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2010.
17. P. T. Anastas, J. C. Warner, *Green Chemistry: Theory and Practice*, Oxford University Press, Oxford, 1998.
18. R. A. Sheldon, *Chem. Ind.*, 1992, 903-906.
19. R. A. Sheldon, *Chem. Ind.*, 1997, 12-15.
20. B.M. Trost, *Science*, 1991, **254**, 1471-1477.
21. P. T. Anastas, J. B. Zimmerman, *Environ. Sci. Techol.*, 2003, **37**, 94-101.
22. C. Jiménez-González, D. J. C. Constable, C. S. Ponder, *Chem. Soc. Rev.*, 2012, **41**, 1485-1498.
23. DIN EN ISO 14040 *Environmental Management, Life Cycle Assessment, Principles and Framework*, DIN (Deutsches Institut für Normung e.V.), Beuth Verlag, Berlin, Germany, 2006.
24. DIN EN ISO 14044 *Environmental Management, Life Cycle Assessment, Requirements and Guidelines*, DIN (Deutsches Institut für Normung e. V.), Beuth Verlag, Berlin, Germany, 2006.
25. P.T. Anastas, R.L. Lankey, *Green Chem.*., 2000, **2**, 289-295.
26. D. Cespi, F. Passarini, E. Neri, I. Vassura, L. Ciacci, F. Cavani, *J. Clean. Prod.*, 2014, **69**, 17-25.
27. D. Cespi, F. Passarini, G. Mastragostino, I. Vassura, S. Larocca, A. Iaconi, A. Chieregato, J.-L. Dubois, F. Cavani, *Green Chem.*, 2015,17, 343-355.
28. N. von der Assen, A. Bardow, *Green Chem.*, 2014, **16**, 3272-3280;
29. H. A. van Kalkeren, A. L. Blom, F. P. J. T. Rutjies, M. A. J. Huijbregts, *Green Chem.*, 2013, **15**, 1255-1263.
30. D. Ott, D. Kralisch, I. Denčić, V. Hessel, Y. Laribi, P. D. Perrichon, C. Berguerand, L. Kiwi-Minsker, P. Loeb, *ChemSusChem*, 2014, **12**, 3521-3533.
31. C. Jiménez-González, A. D. Curzons, D. J. C. Constable, V. L. Cunningham, *J. Clean. Technol. Environ. Policy*, 2004, **7**, 42-50.
32. G. Jödicke, O. Zenklusen, A. Weidenhaupt, K. Hungerbühler, *J. Clean. Prod.*, 1999, **7**, 159-166.
33. S. Kim, C. Jiménez-González, B.E. Dale, *Int. J. Life Cycle Assess.*, 2009, **14**, 392–400.
34. A. S. Bell, D. Brown, N. K. Terrett, *EU Pat.*, 0 463 756 A1, 1991.
35. Ecoinvent Centre (formerly Swiss Centre for Life Cycle Inventories) (2009) Ecoinvent 2.2 Database.
36. R. Hischier, S. Hellweg, C. Capello, A. Primas, *Int. J. Life Cycle Assess.*, 2005, **10**, 59-67.
37. H.-D. Hahn, G. Dämbkes, N. Rupprich, H. Bahl H., *Butanols*, Ullmann's Encyclopedia of Industrial Chemistry. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2010, DOI: 10.1002/14356007.a04_463.pub2.
38. S. Sridhar, R.G. Carter, *Diamines and Higher Amines, Aliphatic*, Kirk-Othmer Encyclopedia of Chemical Technology. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2001, DOI: 10.1002/0471238961.0409011303011820.a01.pub2.
39. *US Pat.*, 4 727 143, 1988.
40. S. Shimizu, N. Watanabe, T. Kataoka, T. Shoji, N. Abe, S. Morishita, H. Ichimura *Pyridine and Pyridine Derivatives*, Ullmann's Encyclopedia of Industrial Chemistry. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000, DOI: 10.1002/14356007.a22_399.
41. SciFinder® https://scifinder.cas.org (accessed September 2014).
42. M. R. Thomas, *Salicylic Acid and Related Compounds*, Kirk-Othmer Encyclopedia of Chemical Technology. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000, DOI: 10.1002/0471238961.1901120920081513.a01.
43. W. B. McCormack, B. C. Lawes, *Sulfuric and Sulfurous Esters*, Kirk-Othmer Encyclopedia of Chemical Technology. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000, DOI: 10.1002/0471238961.1921120613030315.a01.
44. A. Szïjjarto, S. Papadokonstantakis, U. Fischer, K. Hungerbühler, *Ind. Eng. Chem. Res.*, 2008, **47**, 7323-7334.
45. G. Van der Vorst, J. Dewulf, W. Aelterman, B. De Witte, H. Van Langenhove, *Ind. Eng. Chem. Res.*, 2009, **48**, 5344-5350.
46. Gendorf, Umwelterklärung 2000, Industrial Park Werk Gendorf, Burgkirchen, 2000.
47. HIS - Process Economics Program (PEP), http://chemical.ihs.com/PEP/yearbook.htm, (accessed February 2015).
48. Nexant - Process Economics and Research Planning (PERP), http://thinking.nexant.com/program/process-evaluationresearch-planning, (accessed February 2015).
49. G. Wernet, S. Hellweg, U. Fischer, S. Papadokonstantakis, K. Hungerbühler, *Environ. Sci. Technol.*, 2008, **42**, 6717-6722.
50. VDI – Verein Deutscher Ingenieure, *VDI-richtlinie 4600: Cumulative energy demand, terms, defiitions, methods of calculation*, Düsseldorf, Germany, 1997.
51. G. Wernet, C. Mutel, S. Hellweg, K. Hungerbühler, *J. Ind. Ecol.*, 2011, **15**, 96-107.
52. H.J. Althaus, M. Chudacoff , R. Hischier, N. Jungbluth, M. Osses, A. Primas *Life Cycle Inventories of Chemicals*. *Ecoinvent report No. 8, v2.0*, EMPA Dübendorf, Swiss Centre for Life Cycle Inventories, Dübendorf, CH, 2007.
53. R. A. Sheldon, *Green Chem.*, 2007, **9**, 1273-1283.
54. D. J. C. Constable, A. D. Curzons and V. L. Cunningham, *Green Chem.*, 2002, **4**, 521-527.
55. PRé Consultants, SimaPro7, PhD version 7.3.3, Amersfoort, The Netherlands, 2013.
56. M. Goedkoop, R. Heijungs, M. Huijbregts, A. De Schryver, J. Struijs, R. van Zelm, *ReCiPe 2008 – A life cycle impact assessment method which comprises harmonised category indicators at the midpoint and the endpoint level, First edition (version 1.08)*, Ministry of Housing, Spatial Planning and the Environment (VROM), Netherlands, 2013.
57. IPCC - Intergovernmental Panel on Climate Change, *Climate Change 2013. The Physical Science Basis. Fifth Assessment Report.*, Cambridge University Press, United Kingdom, 2013.
58. R. Frischknecht,N. Jungbluth, H.-J. Althaus, C. Bauer, G. Doka, R. Dones, R. Hischier, S. Hellweg, S. Humbert, T. Köllner, Y.

Loerincik, M. Margni, T. Nemecek, *Implementation of Life Cycle Impact Assessment Methods. ecoinvent report No. 3*, v2.0. Swiss Centre for Life Cycle Inventories, Dübendorf, 2007.

59. H. Sugiyama, U. Fischer, K. Hungerbuhler, M. Hirao, *AIChE J.*, 2008, **54**, 1037–1053.
60. A.D. Patel, K. Meesters, H. den Uil, E. de Jong, K. Blok, M. K. Patel, *Energy Environ. Sci.*, 2012, **5**, 8430-8444.



192x263mm (96 x 96 DPI)

Supporting information

Background information about sildenafil citrate

Sildenafil Citrate, the API in Viagra™, constitutes a selective inhibitor of PDE5 used to prevent the male erectile dysfunction. Initially the drug was targeted for Angina and the active ingredient for the development program was produced using an optimized medicinal chemistry route. The improvements to the Medicinal Chemistry route are fully detailed in literature,[*] the main changings regarded: i) the substitution of tin based reducing agent ($SnCl_2$) with a catalytic hydrogenation using palladium, ii) the replacement of the old hydrogen peroxide-based cyclization method with the use of KOBut/tBuOH (100% yield), iii) the introduction of a solvent into the exothermic methylation reaction (molecule 3) and iv) the reduction of thionyl chloride amount (from 1.6 to 1.2 equivalent) by the use of toluene.[*] However, when the indication moved from Angina to Male Erectile Dysfunction it was difficult to keep up with the material requirements for a rapid clinical and then market expansion and this led Pfizer to investigate a new chemical route. Main improvements concerned (i) introduction of a convergent synthesis (ii) the sole use of water as a solvent in the conversion of 2-ethoxybenzoic acid into sulfonamide derivate (compound 9), (iii) the activation of this intermediate using a relatively cheap reagent such as *N,N'* –carbonyldiimidazole and (iv) improving the yield in the citrate salt forming step to 100% using a statistical design approach. Literature[*,16] reports that transition from the medicinal to the commercial route in the case of sildenafil citrate lead to substantial resources reduction which produces an E-factor of 8, lower than average estimated 25 to 100 for the pharmaceutical sector.[18,19] Moreover, all the efforts made in order to increase the process yield while reducing the solvents amount (around 99.7%, from 1540 to 5 kg/kg API) led in 2003 to the win of the Crystal Faraday Award for green chemical technology by the Institute of Chemical Engineers. As reported previously, a detailed description of all the chemical mechanisms and substances involved in both pathways was already reported in literature.

References

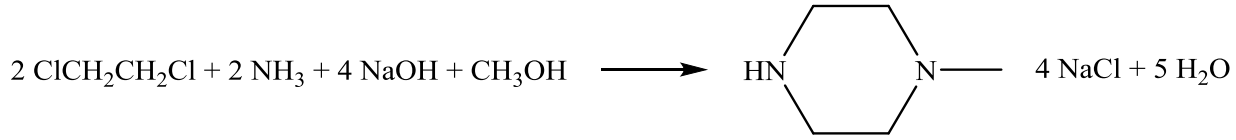
[*] P.J. Dunn, in *Process Chemistry in the Pharmaceutical Industry*, ed. K. Gadamasetti and T. Braisch, CRC Press, Taylor & Francis Group, Boca Raton, Florida, 2008, vol. 2, ch. 16, pp. 267-277.

LCI improvements: detailed description of the N-methylpiperazine pathway

LCI improvement – Phase I

As reported in manuscript, the first step to fill the data gap in the inventory of chemical is select the main pathway for the target molecule. Organic chemistry suggests that N-methylpiperazine (NMP) is commonly produced by the methylation reaction of the piperazine, $HN(CH_2\ CH_2)_2NH$, using methanol as the source of the methyl group.[39] Then the second step is to verify if LCI libraries contain default processes which simulate the synthesis of reagents used in the reaction. Unfortunately, Ecoinvent database did not provide any information regarding the industrial synthesis of the piperazine. However, literature[38] suggests that piperazine is obtained as co-product in the ammoniation of 1,2-dichloroethane, $ClCH_2CH_2Cl$. The reaction with aqueous ammonia, carried out at high temperature and pressure, results in piperazine and ethyleneamine co-products in the form of amine hydrochloride salts, which is then neutralized by sodium hydroxide with the release of NaCl and free amines. [38]

Below the total stoichiometric equation which describes the entire pathway is reported



As stated in the manuscript, a *cradle-to-gate* boundaries were considered using an already published standard approach in order to fill the inventory gap for chemicals, assuming: a process efficiency of 95% over the entire stoichiometric equation, air emissions estimated to be 0.2% of the input and the water emissions were evaluated as a difference between unreacted reagents and air releases.[36] Therefore, neglecting NaCl and H_2O in waste streams due to their no environmental relevance, the first stage of the inventory was completed using the mass balance reported below calculated assuming the production of 1kg of NMP (100.16 kg/kmol) as a functional unit, which corresponds to almost 9.98E-03kmoles.

Molecular weights (kg/kmol)

1,2-dichloroethane: 98.96

NaOH: 39.99

NH_3: 17.03

CH_3OH: 32.04

System input, process efficiency of 95% over the entire stoichiometric equation

kmol 1,2-dichloroethane = (9.98E-03/0.95) * 2 = 2.10E-02

kg 1,2-dichloroethane = 2.10E-02 * 98.96 = 2.08

kmol NH_3 = (9.98E-03/0.95) * 2 = 2.10E-02

kg NH_3 = 2.10E-02 * 17.03 = 3.58E-03

kmol NaOH = (9.98E-03/0.95) * 4 = 4.20E-02

kg NaOH = 4.20E-02 * 39.99 = 1.68

kmol CH_3OH = (9.98E-03/0.95) = 1.05E-02

kg CH_3OH = 1.05E-02 * 32.04 = 3.37E-01

System outputs

kg 1,2-dichloroethane in air = 2.08 * 0.2% = 4.16E-03

kg 1,2-dichloroethane in water = (2.08 * 0.05) - 4.16E-03 = 9.98E-02

kg NH_3 in air = 3.58E-03 * 0.2% = 7.16E-04

kg NH_3 in water = (3.58E-03 * 0.05) - 7.16E-04 = 1.72E-02

kg NaOH in air = 1.68 * 0.2% = 3.36E-03

kg NaOH in water = (1.68 * 0.05) - 3.36E-03 = 8.07E-02

kg CH_3OH in air = 3.37E-01 * 0.2% = 6.73E-04

kg CH_3OH in water = (3.37E-01 * 0.05) - 6.73E-04 = 1.62E-02

LCI improvement – Phase II

Completed the first stage of improvement, the second step concerned the introduction of energy usages in the synthesis of the four substances. As described in the manuscript, a combination of the Finechem tool and information already published in literature was used to fill the data lack. A detailed description of the tool developed by ETH was already published. For more details please consult literature.[12,49] In order to run Finechem, all the procedure reported on http://www.sust-chem.ethz.ch/tools/finechem was followed. First the R project for statistical computing was downloaded. Then, in order to run the analysis, was necessary completed the information requested by the bullet points for each chemical. This operation has been facilitated through the use of the detailed information listed as a support of the previous literature.[12] In the case of NMP, as well as for all the other chemicals, the data used are depicted below in Table S2. After the analysis, the tool provides results in terms of CED, GWP and Ecoindicator 99. In order to simulate the energy requirements for NMP the CED value and its uncertainty were taken into account. As suggested by literature,[51] the contribution of energy consumption involved in the upstream processes (raw material extraction and substance production) for the sector of organic chemical is around 57% of the total CED value. Therefore, the value predicted by Finechem was multiplied for 57% and then split using an average energy mix of European chemical plant[46] (Table S3). In the case of 1kg of N-methylpiperazine the following calculations should be considered:

CED estimation by Finechem - MJ eq. = 193.8 ±65.6

Average energy consumption in the manufacturing - MJ eq. = 193.8 * 0.57 = 110.4

50% Natural gas - MJ eq. = 110.4 * 0.5 = 55.2

38% Steam - MJ eq. = 110.4 * 0.38 = 42.0

12% Electricity - MJ eq. = 110.4 * 0.12 = 13.3

Life cycle impact assessment

After the completion of the LCI improvement stage, two different scenarios were created using SimaPro software: the first, which only includes the mass flows involved in the NMP synthesis and the second, which simulates the entire process filled with the energy requirements. Then, each of them was analyzed separately in order to show the improvements in terms of results achieved per method considered (CED and IPCC 2013), as reported in the manuscript (Figure 6).

Substance	Production Process	Process Input	Amount (kg)	Process Output	Released in air (kg)	Released in water (kg)
TBA	from olefin hydration [i]	Butene, mixed, at plant/RER U	7.97E-01	Butene	1.59E-03	3.82E-02
		Water, decarbonised, at plant/RER U	2.56E-01	-	-	-
NMP	from piperazine and methanol [ii]	Ethylene dichloride, at plant/RER U	2.08E+00	Ethane, 1,2-dichloro-	4.16E-03	9.98E-02
		Ammonia, steam reforming, liquid, at plant/RER U	3.58E-01	Ammonia	7.16E-04	1.72E-02
		Sodium hydroxide, 50% in H2O, production mix, at plant/RER U	7.16E-01	Sodium hydroxide	1.43E-03	3.44E-02
		Methanol, at plant/GLO U	3.37E-01	Methanol	6.73E-04	1.62E-02
PRD	from aldehydes with ammonia [iii]	Acetaldehyde, at plant/RER U	1.17E+00	Acetaldehyde	2.34E-03	5.63E-02
		Formaldehyde, production mix, at plant/RER U	4.00E-01	Formaldehyde	7.99E-04	1.92E-02
		Ammonia, liquid, at regional storehouse/RER U	2.27E-01	Ammonia	4.53E-04	1.09E-02
EBA	from salicylic acid with ethyl-sulfate [iv]	Ethanol from ethylene, at plant/RER U	8.76E-01	Ethanol	1.75E-03	4.20E-02
		Sulphur dioxide, liquid, at plant/RER U	4.06E-01	Sulfur dioxide	8.12E-04	1.95E-02
		Chlorine, liquid, production mix, at plant/RER U	4.49E-01	Chlorine	8.98E-04	2.16E-02
		Thionyl chloride * **	7.54E-01	Thionyl chloride	1.51E-03	3.62E-02
		Phenol, at plant/RER U	5.96E-01	Phenol	1.19E-03	2.86E-02
		Sodium hydroxide, 50% in H2O, production mix, at plant/RER U	5.06E-01	Sodium hydroxide	1.01E-03	2.43E-01
		Carbon dioxide liquid, at plant/RER U ***	2.79E-01	Carbon dioxide	1.39E-02	-
		Hydrochloric acid, from the reaction of hydrogen with chlorine, at plant/RER U	2.31E-01	Hydrogen chloride	4.62E-04	1.11E-02

Tab. S1 LCI - Mass flows involved in the production of chemicals, referred to the production of 1kg of each substance

[i] H.-D. Hahn, G. Dämbkes, N. Rupprich, H. Bahl H., *Butanols*, Ullmann's Encyclopedia of Industrial Chemistry. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2010, DOI: 10.1002/14356007.a04_463.pub2.

[ii] S. Sridhar, R.G. Carter, *Diamines and Higher Amines, Aliphatic*, Kirk-Othmer Encyclopedia of Chemical Technology. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2001, DOI: 10.1002/0471238961.0409011303011820.a01.pub2; *US Pat.*, 4 727 143, 1988.

[iii] S. Shimizu, N. Watanabe, T. Kataoka, T. Shoji, N. Abe, S. Morishita, H. Ichimura *Pyridine and Pyridine Derivatives*, Ullmann's Encyclopedia of Industrial Chemistry. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000, DOI: 10.1002/14356007.a22_399.

[iv] SciFinder® - The choice for chemistry research™ https://scifinder.cas.org (accessed February 2015); M. R. Thomas, *Salicylic Acid and Related Compounds*, Kirk-Othmer Encyclopedia of Chemical Technology. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000, DOI: 10.1002/0471238961.1901120920081513.a01; W. B. McCormack, B. C. Lawes, *Sulfuric and Sulfurous Esters*, Kirk-Othmer Encyclopedia of Chemical Technology. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000, DOI: 10.1002/0471238961.1921120613030315.a01.

* Process not included in Ecoinvent database. Modeled based on process reported in literature (W. B. McCormack, B. C. Lawes, *Sulfuric and Sulfurous Esters*, Kirk-Othmer Encyclopedia of Chemical Technology. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000, DOI: 10.1002/0471238961.1921120613030315.a01.) and standard estimation method developed by Hischier et al. 2005: a process efficiency of 95% over the entire stoichiometric equation; 0.2% of input released as air emissions; the difference between unreacted input and air emission is released in water; average energy consumption (e.g. 2MJ in the form of steam and 0.333kWh of electricity).

** Just added as kg amount in the air and water emissions

*** Assumed that all unreacted is released in air.

	TBA	NMP	EBA	PRD
Chemical descriptors				
Molecular weight (g/mol)	74.12	100.16	166.17	79.10
Number of N atoms	0	2	0	1
Number of halogen atoms	0	0	0	0
Number of rings (both aromatic and aliphatic)	0	1	1	1
Number of tertiary and quaternary C atoms	1	0	1	0
Number of heteroatoms within the rings	0	2	0	1
Number of unique substituents on aromatic ring systems in the molecule	0	0	2	0
Number of functional groups	1	2	3	1
Number O atoms in carbonyl groups	0	0	0	0
Number O atoms except those in carbonyl groups	1	0	2	0
Results from FineChem				
CED (MJ eq./ kg)	104.8±17.5	193.8±65.6	205.4±93.6	147.1±61.8

Tab. S2 FineChem tool - input and output values

	Ecoinvent process	TBA	NMP	EBA	PRD
Energy values (MJ)					
Average CED values - evaluated by Finechem	-	104.8	193.8	205.4	147.1
Average energy consumption in the manufacturing [i]	-	59.7	110.4	117.1	83.8
Energy mix considered in the production process [ii]	Heat, natural gas, at industrial furnace low-NOx >100kW/RER U	29.9	55.2	58.5	41.9
	Electricity, production mix RER/RER U	22.7	42.0	44.5	31.9
	Heat, unspecific, in chemical plant/RER U	7.2	13.3	14.0	10.1

Tab. S3 LCI – Energy consumption involved in the production of chemicals, referred to the production of 1kg of each substance

[i] based on: G. Wernet, C. Mutel, S. Hellweg, K. Hungerbühler, *J. Ind. Ecol.*, 2011, **15**, 96-107.

[ii] based on: Gendorf, Umwelterklärung 2000, Industrial Park Werk Gendorf, Burgkirchen, 2000.

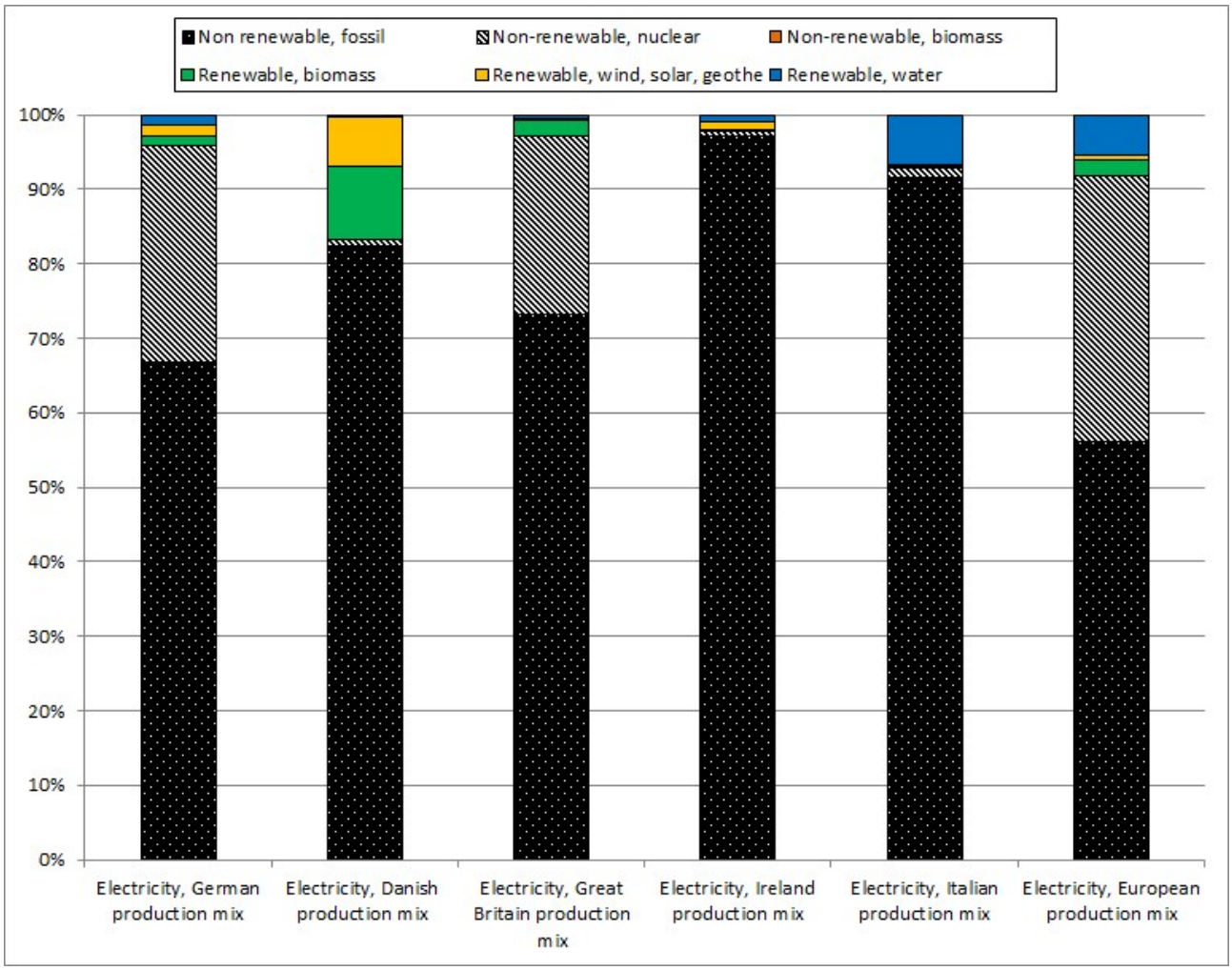


Fig. S1 Contribution analysis in terms of disaggregated CED values of the differences country energy mixes (1kWhe)

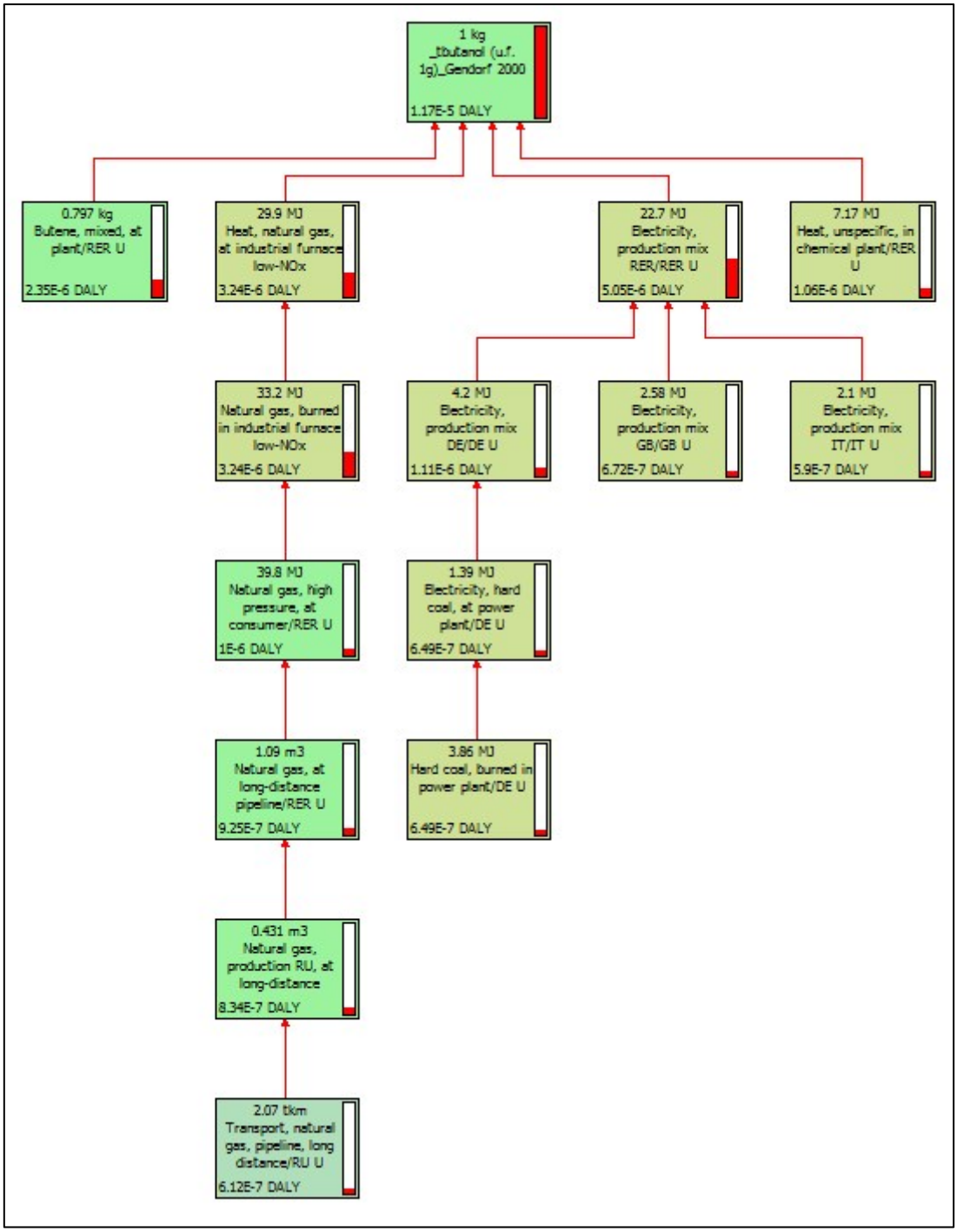


Fig. S2 Network tool results for TBA: processes with higher contribution for the HH category (5% cut-off).

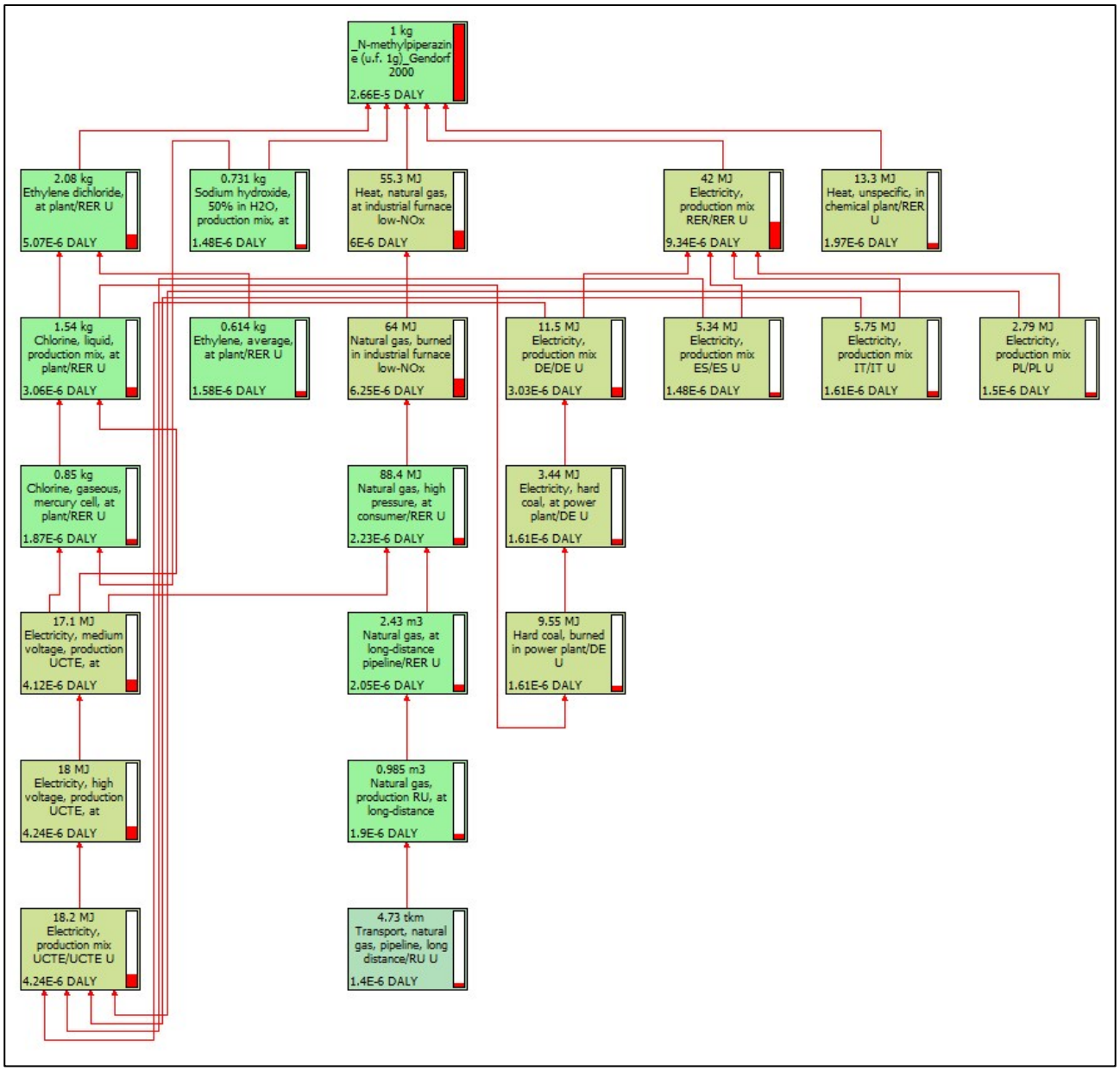


Fig. S3 Network tool results for NMP: processes with higher contribution for the HH category (5% cut-off).

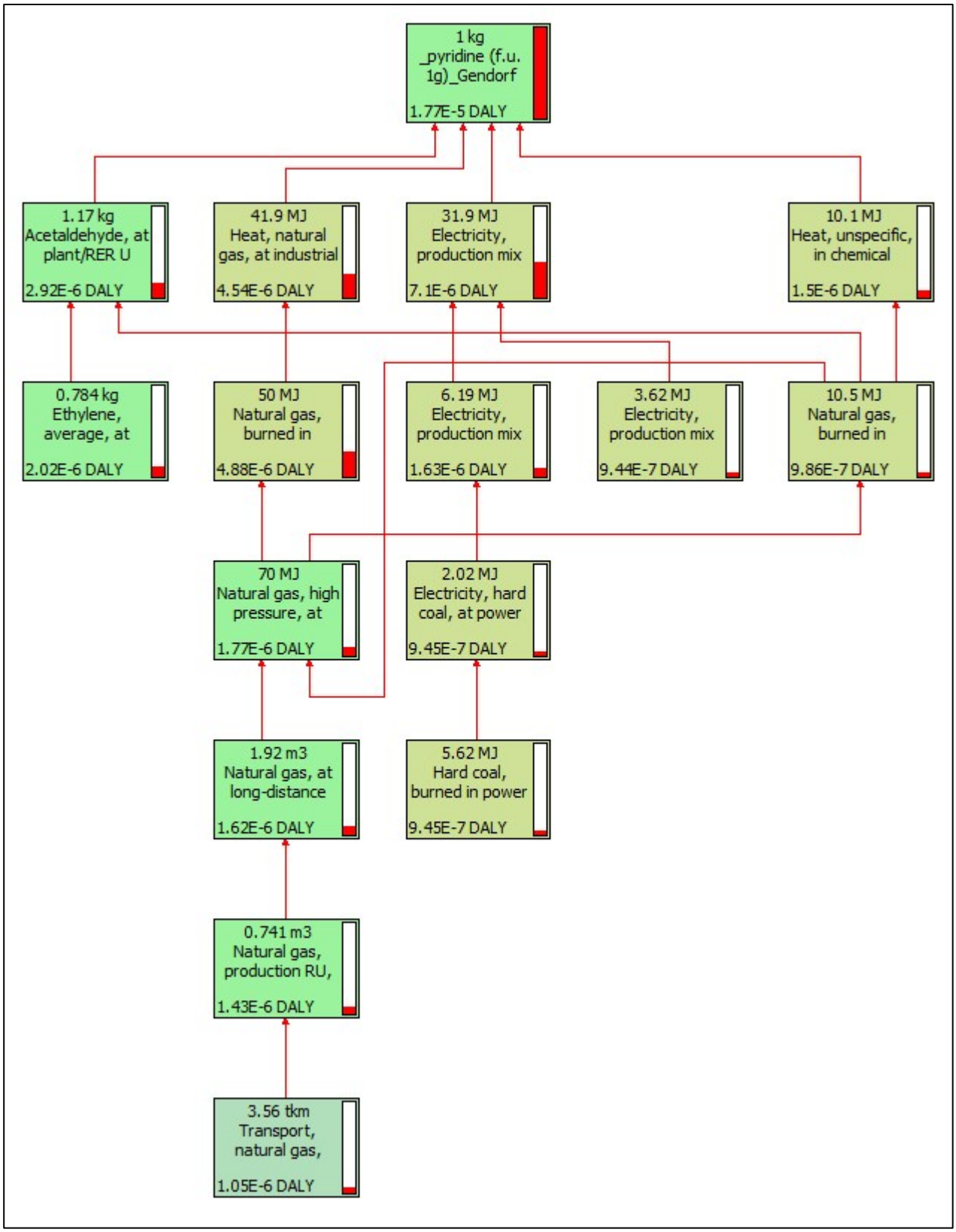


Fig. S4 Network tool results for PRD: processes with higher contribution for the HH category (5% cut-off).

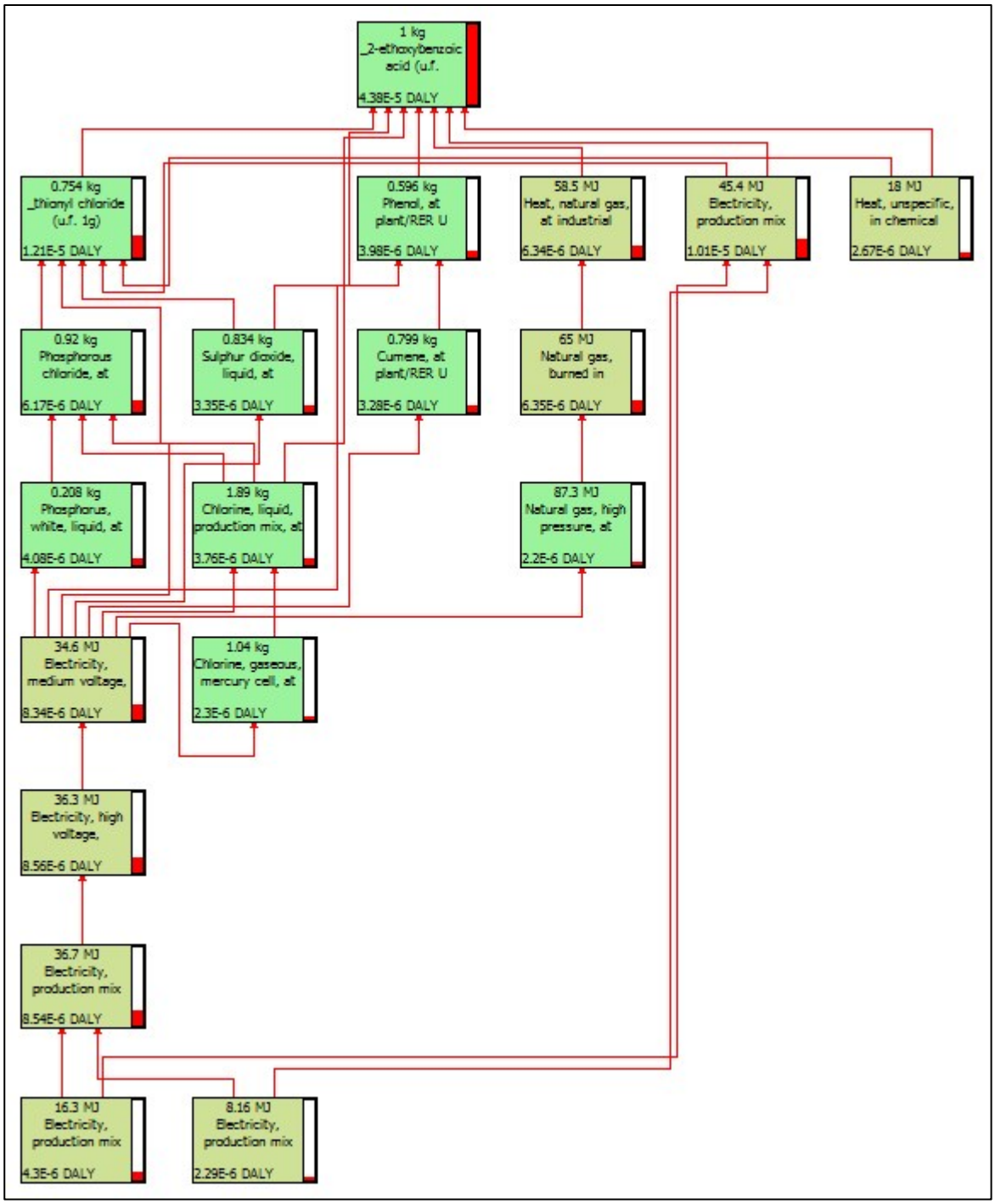


Fig. S5 Network tool results for EBA: processes with higher contribution for the HH category (5% cut-off).