Computational synthesis design for controlled-degradation and revalorization.

Anna Żądło-Dobrowolska[1*], Karol Molga[1,2], Olga O. Kolodiazhna[1], Sara Szymkuć,[1,2] Martyna Moskal[2], Bartosz A. Grzybowski[1,2,3,4*]

[1] Institute of Organic Chemistry, Polish Academy of Sciences, ul. Kasprzaka 44/52, Warsaw 01-224, Poland

[2] Allchemy, Highland, IN, USA

[3] IBS Center for Soft and Living Matter and

[4] Department of Chemistry, UNIST, 50, UNIST-gil, Eonyang-eup, Ulju-gun, Ulsan, 689-798, South Korea

*Correspondence to: anna.zadlo@icho.edu.pl or nanogrzybowski@gmail.com

ABSTRTACT

Degradation of larger and undesired/harmful molecules into smaller and, ideally, value-added products is one of the important facets of circular chemistry. This task may be cumbersome to chemists who are accustomed to plan syntheses using bond-forming rather than bond-breaking methodologies. This work describes a forward-synthesis algorithm that can guide such degradation-oriented analyses. This algorithm uses a broad knowledge-base of degradative and related reactions and applies them to arbitrary small-molecule feeds to generate large synthetic networks within which it then traces degradative pathways that are chemically sound and lead to value-added products. Predictions of the algorithm are validated by proof-of-concept experiments entailing degradation and revalorization of two biomass feeds, D-glucose and quinine.

INTRODUCTION

Recent years have brought remarkable progress in the use of computers in various aspects[1-23] of synthesis design. Most attention has been centered on the classic problem of retrosynthesis[1-9], culminating in the demonstration of our Chematica (now Synthia™) algorithm autonomously planning total syntheses of complex natural products.[2,4] In parallel, we[19-21] and others[22-25] have studied an "inverse," forward-synthesis problem, in which the machine provided with a set of substrates is tasked with mapping out and productively exploring the space of products synthesizable from this starting set. We have recently used this forward synthesis approach to plan multistep routes converting industrial-scale chemical wastes into over 300 drugs and agrochemicals.[20] However, this work centered on making larger and/or more complex products from simple starting materials and, as such, did not address another important challenge of circular chemistry[26,27] – namely, how to selectively *degrade* larger feedstocks and, desirably, revalorize them into more benign and/or value-added smaller molecules. This is a timely question since traditional degradation methodologies relying on thermochemical decomposition often lack selectivity and yield complex mixtures of undefined products with unknown toxicity.[28,29] Not surprisingly, there has been much interest in developing more selective alternatives, albeit largely centered on a relatively small number of feedstocks like polymers,[30-32] haloalkanes,[33,34] polychlorinated aromatics,[35-37] some toxins[38], dyes[39] or certain types of biomass[40-43]. One of the difficulties in planning degradative routes for other "input" molecules may be that chemists are generally taught to think how to make bonds whereas degradation entails bond-breaking and molecule-fragmenting methodologies. Some of these degradative reactions are obvious (e.g., hydrolysis, decarboxylation) but some are perhaps less familiar and harder to spot as applicable to non-trivial, small-molecule feeds (say, Bauer-Haller cleavage or Eschenmoser-Tanabe

fragmentation). Here, we describe a forward-synthesis algorithm that uses a broad knowledge-base of such degradative and related reactions and applies them to arbitrary small-molecule feeds. This algorithm rapidly generates large networks of synthetic possibilities within which it traces degradative pathways that are chemically sound and, additionally, can be evaluated against various by-product, process and pricing criteria. We validate several of such syntheses in proof-of-concept experiments in which inexpensive biomass molecules are degraded into value-added specialty chemicals. We suggest that algorithms like ours can help foster the circular chemistry agenda and guide the discovery of new routes for the detoxification and revalorization of chemical wastes and biomass.

RESULTS AND DISCUSSION

Reaction rules. Our computational analyses were based on reaction rules expert-coded in SMARTs notation as detailed in our earlier works[1-3,14]. Briefly, each of the rules is much broader than any individual reaction precedent and spans "core" atoms characteristic of a given reaction type as well as substituents admissible at different flanking positions. To detect impossible reactions and eliminate false-positive predictions afflicting purely-data-driven AI approaches[14], each rule is accompanied by a list of groups that are incompatible with this reaction. These groups are chosen from a static list of 600 and, on average, some 130 incompatibilities are assigned to each reaction transform. Another aspect of the rules – not covered in our previous works – is that they are coded to account for the key reaction byproducts whose nature may be important in circular chemistry applications (e.g., to flag highly toxic ones). The rules also specify conditions typical to a given reaction type (solvent, reagents, temperature), suggest "greener" alternatives if available (this is done based on the GSK criteria)[44-46], and provide links to illustrative references.

The problem of chemical degradation requires judicious selection of pertinent reaction rules – in our choices, we prioritized popular and typically high-yielding methodologies that cover broad substrate scopes, involve no overly toxic reagents, and are potentially scalable. Of note, automatic extraction from reaction databases was problematic not only because of the low quality of such rules (cf. above and discussion in ref [14]) but also because it was hard to resolve which examples lowering mass or heavy-atom count were legitimate degradative methodologies vs. standard deprotections, substitutions, etc. Within this general scheme, one set of rules comprises reactions that disconnect (in the forward direction) the substrates into smaller molecules. We encoded 763 such strictly degradative transforms including different variants of simple hydrolysis, decarboxylation, retro-Diels-Alder, retro-Michael, retro-aldol, Cope elimination, Beckmann fragmentation, retro-hydroamidocarbonylation, etc. This set was then supplemented by 1161 "auxiliary" transforms: functional group interconversions, some popular rearrangements (e.g., Cope, Wittig, Beckmann), eliminations (water, hydrogen chloride, hydrogen cyanide, epoxide openings, etc.). The role of these auxiliary reactions is to generate intermediates that broaden the applicability of the degradative rules (which, by themselves, often present few options or may not even match a given starting material at all). Specific reaction classes are listed in the Supplementary **Section S1.3.**

Reaction networks. The starting point of the searches is a user-specified small organic molecule (the algorithm does not aim to degrade polymers) which we will refer to, per notation from earlier works[19,20], as the zeroth synthetic generation, G_0. To this substrate, the reaction transforms are applied resulting in the first generation, G_1, products, whose further reactions give generation G_2, and so on. The process is iterated until some user-specified generation G_n, and produces a reaction network in which the nodes correspond to molecules and connections/"edges", to reactions in

which these molecules participate. Because the networks expand very rapidly[20] with *n*, the searches described here are up to $n = 7$ with typical calculation times in minutes and generating rapidly expanding and densely connected reaction networks.

"Value" analyses of different molecules within these large networks are facilitated by coloring the nodes corresponding to commercially available chemicals by their prices (**Figure 1**). Here, we use a specialty-chemicals catalog kindly provided to us by Mcule[47] and comprising ca. 2.4 million molecules with prices standardized to per-gram (of course, in industrial applications, larger-scale units and different price catalogs would be desirable). Synthetic path(s) producing any given compound in the network are displayed by clicking on the corresponding node. As illustrated in **Figure 2**, these routes can be scored according to process criteria discussed in our earlier works[20]. This algorithm is part of our Allchemy platform[19,20] and is available for testing by the academic community (see code availability statement; for User Manual see Supplementary **Section S1.2**).

Specific feeds and experimental validations. We performed calculations for 12 feeds of different properties and origins: hazardous substances that need to be detoxified and utilized; industrial wastes that can be transformed into simpler but still useful reagents; renewable compounds of plant or animal origin that are produced in large quantities and, if degraded into value-added molecules, could reduce dependence on petrochemical-derived raw materials[48]. 19 degradative routes starting from these feeds are detailed in the Supplementary **Section S2.1** and 15 of them are summarized in **Figure 3**.

From amongst these options, we selected for experimental, laboratory-scale validation pathways originating from two biomass-derived substances, very inexpensive D-glucose (costing less than one cent per gram) and relatively inexpensive quinine (5.98 $/g at Mcule and 4.64 $/g at Millipore-Sigma). D-Glucose is produced from plants (corn, rice, wheat, barley, sweet potato) but can also

be recovered from textile, cardboard, or food industrial waste.[49-50] Quinine is isolated from the bark of a cinchona tree, is known for its antimalarial activity, and is a component of tonic water.[51] For both feeds, we were interested in products that would give significant increase in price, here taken as at least 30-fold.

Degradation of D-glucose. The network propagated from D-glucose up to G_7, comprises 6569 molecules nodes, of which 1490 are detected as commercially available and 1303 have Mcule list prices >30 times higher than the price of the feed (for clarity, only subnetwork up to G_4 is shown in **Figure 4a**; for the full network, see Supplementary **Section S.2.2**). With reference to this figure, degradative routes leading 1,4-diaminobutane, **5** (8.87 $/g at Mcule and 2.37 $/g at Millipore-Sigma as dihydrochloride salt) appeared both straightforward and interesting since **5** is (i) a building block useful in the syntheses of various biologically active substances[52-55] and (ii) although relatively inexpensive (8.87 $/g at Mcule and 0.5 $/g at Millipore-Sigma), it is produced at large scales from non-renewable, petroleum-derived feeds[56,57] which, at least in long term, are undesirable for circular chemistry.

The synthesis plans are detailed in **Figure 4b**. As seen therein, despite being trained on general reaction rules rather than specific literature precedents, the algorithm re-discovered (*blue* reaction arrows) all the literature reported steps (*black* arrows), and provided additional suggestions for unprecedented steps (*red* arrows). In particular, the four-step D-glucose-**1**-**2**-**3**-**5** route involving straightforward dehydration[58], oxidation[59], amidation[60] and direct reduction of succinamide **3** into **5** was suggested as an alternative to the literature sequence D-glucose-**1**-**2**-**3**-**4**-**5** requiring five steps and offering only 3% overall yield. In both of these routes, the bottleneck was the low-yielding (9%) amidation **2**-**3**[60] employing hazardous ammonia and high temperatures, >200 ^{o}C. For this step, the algorithm suggested the use of the popular, more benign and inexpensive

(0.83$/g) 1,1′-carbonyldiimidazole (CDI) coupling reagent in DMF solvent. This recommendation worked very well, and succinamide (**3**) synthesis was substantially improved, giving 84% isolated yield. For the previously unreported **3**-**5** reduction, popular lithium aluminium hydride (LAH) – considered greener than alternatives such as borane-THF or H_2/Ni – was suggested and gave amine **5** in 40% yield, translating into the pathway yield of 16%. The program also suggested improvements for the literature route. In particular, the **3**-**4** dehydration step was previously reported using a highly toxic $Pb(OAc)_2$ catalyst[61] or a sophisticated apparatus[62]. In this case, the algorithm recommended the use of less problematic trifluoroacetic anhydride (TFAA) as a dehydrating agent. In just one attempt (without optimization), succinonitrile **4** was isolated in 75%, translating into the pathway yield of 19%. In terms of pricing, the estimated cost of preparing 1 gram of **5** dihydrochloride along the shorter, D-glucose-**1**-**2**-**3**-**5** sequence (including ~~all~~ solvents and reagents) was ~$15, which compares favorably against the 82 $/g pricing for the optimized literature, D-glucose-**1**-**2**-**3**-**4**-**5** route but, expectedly, is not competitive against the abovementioned market prices of **5** derived at scale from petroleum feedstocks.

Of note, the network shown in **Figure 4a** also contained conversion of D-glucose to D-fructose, allowing for the implementation of the routes in **Figure 4b** from an alternative and only slightly more expensive biomass feed (16.5 PLN/kg or 0.004 $/g in a local grocery in Poland; 0.058 $/g at Millipore-Sigma).

Degradation of quinine. As the second example, we considered a more structurally advanced feed, quinine. A fragment of the network, up to G_4, propagated from this starting materials is shown in **Figure 5a**. The full G_7 network (Supplementary **Section S2.3**) comprises 6702 molecules, of which 108 are identified as commercially available and, of these, 29 have the per-gram prices at least 30 times higher than the feed. We were particularly interested in the degradative routes

leading to 6-methoxyquinoline-4-carbaldehyde, **10** (**Figure 5b**) (151.99 $/g at Mcule, 595 $/g at Millipore-Sigma) and 4-amino-6-methoxyquinoline, **12** (202.65 $/g/ at Mcule and 255 $/g at Millipore-Sigma) as both are building blocks used in syntheses of various antibiotics and other biologically active compounds.[63-67] Additionally, we were interested in 2-((3*R*,4*R*)-3-vinylpiperidin-4-yl)acetaldehyde, **8** – although this particular product is not commercially available, its Boc-protected derivative is used for stereocontrolled synthesis of drugs active against malaria parasites, including analogs of quinine and quinidine.[68-71]

The computer-designed plans are detailed in **Figure 5b**, where **8**, **10** and **12** share the first three synthetic steps: oxidation, condensation, and Beckmann fragmentation. Oxidation of quinine and condensation of the resulting ketone with hydroxylamine were previously described[72,73], although the latter gave only moderate, 18% yield and could thus benefit from optimization of conditions. The key Beckmann fragmentation lacks a literature precedent as is interesting because it can yield two valuables at once: the aforementioned **8** and 6-methoxyquinoline-4-carbonitrile **9**. Compound **9** can then be converted into one of our targets, aldehyde **10**, via previously unreported reduction or into **12** via straightforward (though lacking direct literature precedent) hydrolysis followed by Hoffman rearrangement (previously reported in ref. [74]).

In experiment, quinine was oxidized under literature, Oppenauer-Woodward conditions (benzophenone, t-BuOK, toluene) giving a mixture of two diastereomeric products, quininone and quinidinone (**6**) in 50:50 ratio and 96% yield. Then, mixture of (*E*) and (*Z*)-oximes (**7**) was synthesized by condensation of corresponding ketones with hydroxylamine. Computer-suggested conditions (NH_2OH, NaOAc, EtOH) gave 71% yield which was increased to 86% using NaOH instead of NaOAc. Subsequently, **7** was subjected to Beckmann fragmentation and both molecules 2-((3*R*,4*R*)-1-tosyl-3-vinylpiperidin-4-yl)acetaldehyde (*p*Ts-protected-**8**, 34%) and

methoxyquinoline-4-carbonitrile **9** (80%) were obtained. During solvent screen for this step, a change from EtOH into DCM resulted in Beckmann rearrangement product **13** which we also isolated in 30% yield. Treatment of **9** with DIBAL-morpholine led to aldehyde **10** in 80% yield (DIBAL is not a green reagent but does not have any alternatives with comparable versatility; we decided to use DIBAL-morpholine system as it is more effective, proceeds under ambient temperature and shorter reaction times compared to DIBAL alone) while oxidative hydrolysis (H_2O_2, LiOH, THF:H_2O) led to 6-methoxyquinoline-4-carboxamide (**11**) in 97% yield. Also, **13** was transformed to amide **11** under the same oxidative conditions (77% yield). Moreover, we established that one-pot, four-step sequence is possible, starting with ketones **6** and giving amide **11** in overall 66% yield and eliminating the need for purification of intermediates (for details see Supplementary **Section S3.4.2**). The last step is formal and, according to literature procedure[74], converts **11** into amine **12** in 68% yield.

Comparing to existing pathways, our 4-step computer-designed route to aldehyde **10** is a more effective alternative (53% overall yield) to either the low-yielding (3%) 5-step procedure starting from quinine (and using toxic DBU and OsO_4),[75,76] or to the 2-step approach offering 46% yield but starting from toxic, not renewable *p*-metoxyaniline[77]. Estimated cost of 1 gram of **10** being synthesized via our method – including cost of reagents and solvents from Sigma-Millipore – is $51.11, compared to a market price of $152.

In the case of amine **12** synthesis, our methodology (overall yield 44%) also appears advantageous to existing pathways, either the one including oxidative cleavage of ketone followed by *N*-tert-butoxycarbamate synthesis (with toxic diphenyl phosphoryl azide) and Curtius rearrangement (19%)[72,78], or the second procedure starting from the previously mentioned toxic *p*-methoxyaniline

(13%).[68] Calculated production cost including solvents and reagents is estimated at 37.39 $/g, compared to the market price of 203 $/g.

In case of **8**, the overall yield along our route, 28% is lower than in existing methods (38%[71,79]) though our route is two-steps shorter which may be significant process-wise. We estimate the cost of this route at 40.2 $/g but cannot compare it with the pricing of **8** since, as mentioned before, it is not commercially available.

CONCLUSIONS

In summary, we described computer-driven design of degradative synthetic pathways to arbitrary small-molecule targets. This forward-synthesis network algorithm aggregates and applies broad knowledge of degradative chemistries that may be less often used by and thus less familiar to synthetic chemists. We see this algorithm as a generator of plausible synthetic ideas: the validations we performed indicate that majority of these ideas fare well in experiment but there were also some cases in which conditions needed to be adjusted by human chemists. The software itself is a work-in-progress and can certainly benefit from further improvements (addition of new transforms, fine-tuning of conditions' information and, pending access to industrial data, scale-up, and process predictions). This said, we believe that even in its current state, it can find practical applications in closing the loop of circular chemistry, suggesting new ideas how to degrade unwanted chemicals which we nowadays produce and accumulate at alarming quantities.

Data availability

User Manual and login details are available in Supplementary **Section S1**. All degradation pathways described in this work and experimental validation data supporting this study are available in Supplementary **Sections S2** and **S3**.

Code availability

The interactive Allchemy web application is freely available for academic users at https://degradation.allchemy.net (given server capacity, to five concurrent academic users on a rolling basis and two-week slots). Reaction rules and the source code of Allchemy are proprietary. To obtain access, please send an email (from your academic address) to admin@allchemy.net.

Acknowledgements

Development of the Degradation module within the Allchemy platform (by K.M., S.S., M.M. and B.A.G.) has been supported by internal funds of Allchemy, Inc. Calculations, analysis of pathways, and laboratory experiments (by A.Ż-D. and O.O.K.) were supported by the National Science Centre, Poland (award 2020/39/D/ST4/01890 to A.Ż-D.). Analysis of results and writing of the paper by B.A.G. was supported by the Institute for Basic Science, Korea (project code IBS-R020-D1). The authors thank the Mcule team for providing access to their catalog and standardizing the prices information therein to per-gram.

Ethics declarations

Competing interests

K.M., S.S., M.M. and B.A.G. are consultants and/or stakeholders of Allchemy, Inc. Allchemy software is the property of Allchemy, Inc., USA.

Contributions

K.M., S.S., M.M. and B.A.G. designed and developed the Allchemy platform. A.Ż-D. performed analyses and computational calculations described in the paper, most syntheses described in Fig. 4 (D-glucose-1-2-3-4) and all syntheses described in Fig. 5.; O.O.K performed synthesis 3-5

described in Fig 4. B.A.G. conceived and supervised the research and wrote the paper with help from A.Ż-D.

References:

1. Szymkuć, S. et al. Computer-assisted synthetic planning: The end of the beginning. *Angew. Chem. Int. Ed.* **55**, 5904–5937 (2016).

2. Mikulak-Klucznik, B. et al. Computational planning of the synthesis of complex natural products. *Nature* **588**, 83–88 (2020).

3. Klucznik, T. et al. Efficient syntheses of diverse, medicinally relevant targets planned by computer and executed in the laboratory. *Chem* **4**, 522–532 (2018).

4. Lin, Y., Zhang, R., Wang, D., & Cernak, T. Computer-aided key step generation in alkaloid total synthesis. *Science* **379**, 453-457 (2023).

5. Segler, M. H. S., Preuss, M. & Waller, M. P. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **555**, 604–610 (2018).

6. Liu, B. et al. Retrosynthetic reaction prediction using neural sequence-to-sequence models. *ACS Cent. Sci.* **3**, 1103–1113 (2017).

7. Coley, C. W. et al. A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science* **365**, eaax1566 (2019).

8. Schwaller, P. et al. Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy. *Chem. Sci.* **11**, 3316-3325 (2020).

9. Rother, D., & Malzacher, S. Computer-aided enzymatic retrosynthesis. *Nat. Catal.* **4**, 92-93 (2021).

10. Coley, C. W., Green, W. H., Jensen, K. F. Machine learning in computer-aided synthesis planning. *Acc. Chem. Res.* **51**, 1281-1289 (2018).

11. de Almeida, A. F., Moreira, R. & Rodrigues, T. Synthetic organic chemistry driven by artificial intelligence. *Nat. Rev. Chem.* **3**, 589–604 (2019).

12. Schwaller, P. et al. Machine intelligence for chemical reaction space. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **12**, e1604 (2022).

13. Strieth-Kalthoff, F., Sandfort, F., Segler, M. H. S. & Glorius, F. Machine learning the ropes: principles, applications and directions in synthetic chemistry. *Chem. Soc. Rev.* **49**, 6154–6168 (2020).

14. Molga, K., Gajewska, E. P., Szymkuć, S., & Grzybowski, B. A. The logic of translating chemical knowledge into machine-processable forms: a modern playground for physical-organic chemistry. *React. Chem. Eng.* **4**, 1506-1521 (2019).

15. Molga, K., Szymkuć, S., Gołębiowska, P., Popik, O., Dittwald, P., Moskal, M., Roszak R., Młynarski J. & Grzybowski, B. A. A computer algorithm to discover iterative sequences of organic reactions. *Nat. Synth.* **1**, 49-58 (2022).

16. Roszak, R., Beker, W., Molga, K., & Grzybowski, B. A. Rapid and accurate prediction of pK_a values of C–H acids using graph convolutional neural networks. *J. Am. Chem. Soc.* **141**, 17142-17149 (2019).

17. Beker, W., Gajewska, E. P., Badowski, T., & Grzybowski, B. A. Prediction of major regio-, site-, and diastereoisomers in Diels–Alder reactions by using machine-learning: The importance of physically meaningful descriptors. *Angew. Chem. Int. Ed.* **58**, 4515-4519 (2019).

18. Guan, Y. et al. Regio-selectivity prediction with a machine-learned reaction representation and on-the-fly quantum mechanical descriptors. *Chem. Sci.* **12**, 2198–2208 (2021).

19. Wołos, A. et al. Synthetic connectivity, emergence, and self-regeneration in the network of prebiotic chemistry. *Science* **369**, eaaw1955 (2020).

20. Wołos, A., et al. Computer-designed repurposing of chemical wastes into drugs. *Nature* **604**, 668-676 (2022).

21. Skoraczyński, G. et al. Predicting the outcomes of organic reactions via machine learning: are current descriptors sufficient? *Sci. Rep.* **7**, 3582 (2017).

22. Gao, W., Mercado, R. & Coley, C.W. Amortized tree generation for bottom-up synthesis planning and synthesizable molecular design. *arXiv:2110.06389* (2021).

23. Granda, J. M., Donina, L., Dragone, V., Long, D. L., & Cronin, L. Controlling an organic synthesis robot with machine learning to search for new reactivity. *Nature* **559**, 377-381 (2018).

24. Robinson, W.E., Daines, E., van Duppen, P. *et al.* Environmental conditions drive self-organization of reaction pathways in a prebiotic reaction network. *Nat. Chem.* **14**, 623–631 (2022).

25. Coley, C. W., Barzilay, R., Jaakkola, T. S., Green, W. H., & Jensen, K. F. Prediction of organic reaction outcomes using machine learning. *ACS Cent. Sci.* **3**, 434-443 (2017).

26. Keijer, T., Bakker, V., & Slootweg, J. C. Circular chemistry to enable a circular economy. *Nat. Chem.* **11**, 190-195 (2019).

27. Kümmerer, K., Clark, J. H., & Zuin, V. G. Rethinking chemistry for a circular economy. *Science* **367**, 369-370 (2020).

28. Zhang, Y., Qi, M. Y., Tang, Z. R., & Xu, Y. J. Photoredox-catalyzed plastic waste conversion: nonselective degradation versus selective synthesis. *ACS Catal.* **13**, 3575-3590 (2023).

29. Sobol, Ł., Dyjakon, A., & Soukup, K. Dioxins and furans in biochars, hydrochars and torreficates produced by thermochemical conversion of biomass: a review. *Environ. Chem. Lett.* 1-25 (2023).

30. Jehanno, C. et al. Critical advances and future opportunities in upcycling commodity polymers. *Nature* **603**, 803-814 (2022).

31. Coates, G. W., & Getzler, Y. D. Chemical recycling to monomer for an ideal, circular polymer economy. *Nat. Rev. Mater.* **5**, 501-516 (2020).

32. Zhang, F. et al. Polyethylene upcycling to long-chain alkylaromatics by tandem hydrogenolysis/aromatization. *Science* **370**, 437-441 (2020).

33. Trang, B., Li, Y., Xue, X. S., Ateia, M., Houk, K. N., & Dichtel, W. R. Low-temperature mineralization of perfluorocarboxylic acids. *Science* **377**, 839-845 (2022).

34. He, J., Ritalahti, K. M., Yang, K. L., Koenigsberg, S. S., & Löffler, F. E. Detoxification of vinyl chloride to ethene coupled to growth of an anaerobic bacterium. *Nature* **424**, 62-65 (2003).

35. Liang, X., Fu, D., Liu, R., Zhang, Q., Zhang, T. Y., & Hu, X. Highly efficient $NaNO_2$-catalyzed destruction of trichlorophenol using molecular oxygen. *Angew. Chem. Int. Ed.* **44**, 5520-5523 (2005).

36. Kumamaru, T., Suenaga, H., Mitsuoka, M. *et al.* Enhanced degradation of polychlorinated biphenyls by directed evolution of biphenyl dioxygenase. *Nat. Biotechnol.* **16**, 663–666 (1998).

37. Meunier, B. Catalytic degradation of chlorinated phenols. *Science* **296**, 270-271 (2002).

38. Smith, B. M. Catalytic methods for the destruction of chemical warfare agents under ambient conditions. *Chem. Soc. Rev.* **37**, 470-478 (2008).

39. Rathi, B. S., & Kumar, P. S. Sustainable approach on the biodegradation of azo dyes: a short review. *Curr. Opin. Green Sustain. Chem.* **33**, 100578 (2022).

40. Antonetti, C., Licursi D., Fulignati, S., Valentini, G. & Raspolli Galletti, A.M. New frontiers in the catalytic synthesis of levulinic acid: from sugars to raw and waste biomass as starting feedstock. *Catal.* **6**, 196 (2016).

41. Liu, F. et al. Continuously processing waste lignin into high-value carbon nanotube fibers. *Nat Commun* **13**, 5755 (2022).

42. Sun, Z., Balint, F., de Santi, A., Saravanakumar, E. & Barta, K. Bright side of lignin depolymerization: toward new platform chemicals. *Chem. Rev.* **118**, 614–678 (2018).

43. Lee, K., Jing, Y., Wang, Y., & Yan, N. A unified view on catalytic conversion of biomass and waste plastics. *Nat. Rev. Chem.***6**, 635-652 (2022).

44. Adams, J. P. et al. Development of GSK's reagent guides – embedding sustainability into reagent selection. *Green Chem.* **15**, 1542 (2013).

45. Henderson, R. K., Hill, A. P., Redman, A. M. & Sneddon, H. F. Development of GSK's acid and base selection guides. *Green Chem.* **17**, 945–949 (2015).

46. Alder, C. et al. Updating and further expanding GSK's solvent sustainability guide. *Green Chem.* **18**, 3879-3890 (2016).

47. MCule (accessed 30 June 2023); https://mcule.com/

48. Zhu, Y., Romain, C., & Williams, C. K. Sustainable polymers from renewable resources. *Nature* **540**, 354-362 (2016).

49. Zebec, Ž., Poberžnik, M., & Lobnik, A. Enzymatic hydrolysis of textile and cardboard waste as a glucose source for the production of limonene in Escherichia coli. *Life* **12**, 1423 (2022).

50. Pleissner, D., & Lin, C. S. K. Valorisation of food waste in biotechnological processes. *Sustain. Chem. Process.* **1**, 1-6 (2013).

51. Woodland, J. G., & Chibale, K. Quinine fever. *Nat. Chem.* **14**, 112-112 (2022).

52. GlaxoSmithKline Intellectual Property Development Ltd. Chemical compounds. W.O. patent 2020/202091 (2020).

53. Trillium Therapeutics Inc. Heteroaromatic-fused imidazolyl amides, compositions and uses thereof as sting agonists. W. O. patent 2020/10451 (2020).

54. Hunter Fleming Ltd. Steroid-hormone conjugates with polyamides and their therapeutic use as anti-cancer and angiostatic agents. W. O. patent 2005/116050 (2005).

55. Pfizer INC. Adenosine A2A receptor antagonists. U. S. patent 2008/242672 (2008).

56. Cespi, D., Passarini, F., Neri, E., Vassura, I., Ciacci, L., & Cavani, F. Life Cycle Assessment comparison of two ways for acrylonitrile production: the SOHIO process and an alternative route using propane. *J. Clean. Prod.* **69**, 17-25 (2014).

57. Diaminobutane (DAB), unique large scale production (AnQuore, accessed 7 July 2023); https://www.anqore.com/en/products/Diaminobutane

58. Qi, L., Mui, Y. F., Lo, S. W., Lui, M. Y., Akien, G. R., & Horváth, I. T. Catalytic conversion of fructose, glucose, and sucrose to 5-(hydroxymethyl) furfural and levulinic and formic acids in γ-valerolactone as a green solvent. *ACS Catal.* **4**, 1470-1477 (2014).

59. Dutta, S., Wu, L., & Mascal, M. Efficient, metal-free production of succinic acid by oxidation of biomass-derived levulinic acid with hydrogen peroxide. *Green Chem.* **17**, 2335-2338 (2015).

60. Battelle Memorial Institute. Process for producing N-methyl succinimide. W. O. patent 2004/58708 (2004).

61. Ruan, S., Ruan, J., Chen, X., & Zhou, S. Facile dehydration of primary amides to nitriles catalyzed by lead salts: The anionic ligand matters. *Mol. Catal.* **499**, 111313 (2021).

62. Chemistry and Chemical ENG Guangdong Laboratory; Beijing Yuji Science & Technology CO LTD. Method for preparing nitrile through gas phase dehydration. Chinese patent 114057605 (2022).

63. Zhou, X. et al. Discovery of novel inhibitors of human phosphoglycerate dehydrogenase by activity-directed combinatorial chemical synthesis strategy. *Bioorg. Chem.* **115**, 105159 (2021).

64. Surivet, J. P. et al. Design, synthesis, and characterization of novel tetrahydropyran-based bacterial topoisomerase inhibitors with potent anti-gram-positive activity. *J. Med. Chem.* **56**, 7396-7415 (2013).

65. Agouridas, C. et al. Synthesis and antibacterial activity of ketolides (6-O-methyl-3-oxoerythromycin derivatives): a new class of antibacterials highly potent against macrolide-resistant and-susceptible respiratory pathogens. *J. Med. Chem.* **41**, 4080-4100 (1998).

66. GlaxoSmithKline Beecham PLC. Nitrogen-containing bicyclic heterocycles for use as antibacterials. W. O. patent 2003/87098 (2003).

67. Gemma, S. et al. Optimization of 4-aminoquinoline/clotrimazole-based hybrid antimalarials: Further structure–activity relationships, in vivo studies, and preliminary toxicity profiling. *J. Med. Chem.* **55**, 6948-6967 (2012).

68. Beuchel, A. et al. Structure–Activity Relationship of Anti-Mycobacterium abscessus Piperidine-4-carboxamides, a New Class of NBTI DNA Gyrase Inhibitors. *ACS Med. Chem. Lett.* **13**, 417-427 (2022).

69. Hoffmann La Roche. Process for piperidine intermediates for quinine, quinidine and analogs thereof. U. S. patent 3931192 (1976).

70. Igarashi, J., Katsukawa, M., Wang, Y. G., Acharya, H. P., & Kobayashi, Y. Stereocontrolled synthesis of quinine and quinidine. *Tetrahedron Lett.* **45**, 3783-3786 (2004).

71. Breman, A. C., Ruiz-Olalla, A., van Maarseveen, J. H., Ingemann, S., & Hiemstra, H. Synthesis of Quinuclidines by Intramolecular Silver-Catalysed Amine Additions to Alkynes. *Eur. J. Org. Chem.* **2014**, 7413-7425 (2014).

72. Illa, O., Arshad, M., Ros, A., McGarrigle, E. M., & Aggarwal, V. K. Practical and highly selective sulfur ylide mediated asymmetric epoxidations and aziridinations using an inexpensive, readily available chiral sulfide. Applications to the synthesis of quinine and quinidine. *J. Am. Chem. Soc.* **132**, 1828-1830 (2010).

73. Kapoor, V. K., & Karwayun, R. 4-Cyano-6-Methoxyquinoline from Quinine. *Indian J. Pharm. Sci*. **57**, 237 (1995).

74. Elderfield, R. C., Gensler, W. J., Birstein, O., Kreysa, F. J., Maynard, J. T., & Galbreath, J. Synthesis of Certain Simple 4-Aminoquinoline Derivatives1. *J. Am. Chem. Soc.* **68**, 1250-1251 (1946).

75. von Riesen, C., Jones, P. G., & Hoffmann, H. M. R. From quinidine to new enantiopure materials–tricyclic allylic N,O-acetals and a stereospecific, onepot conversion of 1, 2-secondary, tertiary diols into spiroepoxides. *Eur. J. Chem*. **2**, 673-679 (1996).

76. von Riesen, C., & Hoffmann, H. M. R. A tricyclic dehydrorubanone and new isomers of the major quinidine metabolite. *Eur. J. Chem.* **2**, 680-684 (1996).

77. Friestad, G. K., Ji, A., Baltrusaitis, J., Korapala, C. S., & Qin, J. Scope of stereoselective Mn-mediated radical addition to chiral hydrazones and application in a formal synthesis of quinine. *J. Org. Chem.* **77**, 3159-3180 (2012).

78. SmithKline Beecham Corporation. Quinoline derivatives as antibacterials. U. S. patent 2004053928A1 (2004).

79. Clark, J. S., Townsend, R. J., Blake, A. J., Teat, S. J., & Johns, A. A concise enantioselective synthesis of the AB ring system of the manzamine alkaloids by ring-closing enyne metathesis. *Tetrahedron Lett.* **42**, 3235-3238 (2001).

80. Wan Station Ltd.; Tokyo Institute of Technology; Fukuoka University; Kinki University Rotaxane compound and antitumor agent E. P. patent 2348021 (2011).

81. Peters, J. U. et al. Discovery of potent, balanced and orally active dual NK1/NK3 receptor ligands. Bioorg. Med. Chem. Lett. **20**, 3405-3408 (2010).

82. Levi, N., Khenkin, A. M., Hailegnaw, B., & Neumann, R. Depolymerization of cellulose in water catalyzed by phenylboronic acid derivatives. *ACS Sustain. Chem. Eng.* **4**, 5799-5803 (2016).

83. Zhang, Z. *et al.* Conversion of carbohydrates into 5-hydroxymethylfurfural using polymer bound sulfonic acids as efficient and recyclable catalysts. *RSC Adv.* **3**, 9201-9205 (2013).

84. Li, C., Zhang, Z., & Zhao, Z. K. Direct conversion of glucose and cellulose to 5-hydroxymethylfurfural in ionic liquid under microwave irradiation. *Tetrahedron Lett.* **50**, 5403-5405 (2009).

85. Li, X., Ho, B., Lim, D. S., & Zhang, Y. Highly efficient formic acid-mediated oxidation of renewable furfural to maleic acid with H_2O_2. *Green Chem.* **19**, 914-918 (2017).

86. Tirsoaga, A., El Fergani, M., Parvulescu, V. I., & Coman, S. M. Upgrade of 5-hydroxymethylfurfural to dicarboxylic acids onto multifunctional-based Fe_3O_4@ SiO_2 magnetic catalysts. *ACS Sustain. Chem. Eng.* **6**, 14292-14301 (2018).

87. Kantam, L., Parsharamulu, T., & Manorama, S. V. Layered double hydroxides supported nano palladium: An efficient catalyst for the chemoselective hydrogenation of olefinic bonds. *J. Mol. Catal. A: Chem.* **365**, 115-119 (2012).

88. McMaster, L., & Langreck, F. B. On the preparation of fumaric nitrile. The action of hydroxylamine on fumaric nitrile. *J. Am. Chem. Soc.* **40**, 970-973 (1918).

89. Blomquist, A. T., & Winslow, E. C. Unsaturated nitriles as dienophiles in the diene synthesis. *J. Org. Chem.* **10**, 149-158 (1945).

90. Chang, F., Kim, H., Lee, B., Park, S., & Park, J. Highly efficient solvent-free catalytic hydrogenation of solid alkenes and nitro-aromatics using Pd nanoparticles entrapped in aluminum oxy-hydroxide. *Tetrahedron Lett.* **51**, 4250-4252 (2010).

91. Khan, H. A., & Robins, D. J. Pyrrolizidine alkaloid biosynthesis. Synthesis of ^{13}C-labelled putrescines and their incorporation into retronecine. *J. Chem. Soc., Perkin Trans.* **1**, 101-105 (1985).

92. Chatupheeraphat, A., Liao, H. H., Lee, S. C., & Rueping, M. Nickel-catalyzed C–CN bond formation via decarbonylative cyanation of esters, amides, and intramolecular recombination fragment coupling of acyl cyanides. *Org. Lett*. **19**, 4255-4258 (2017).

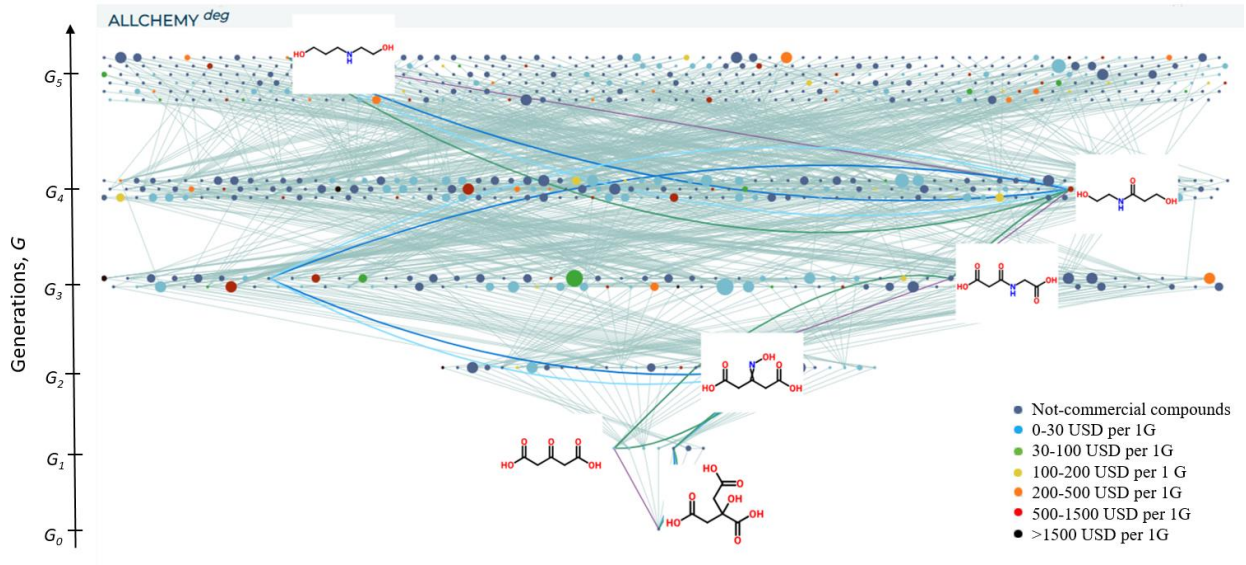
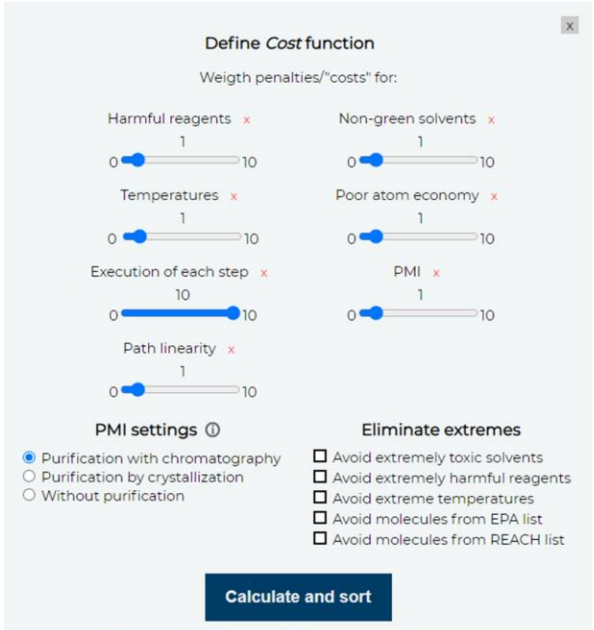
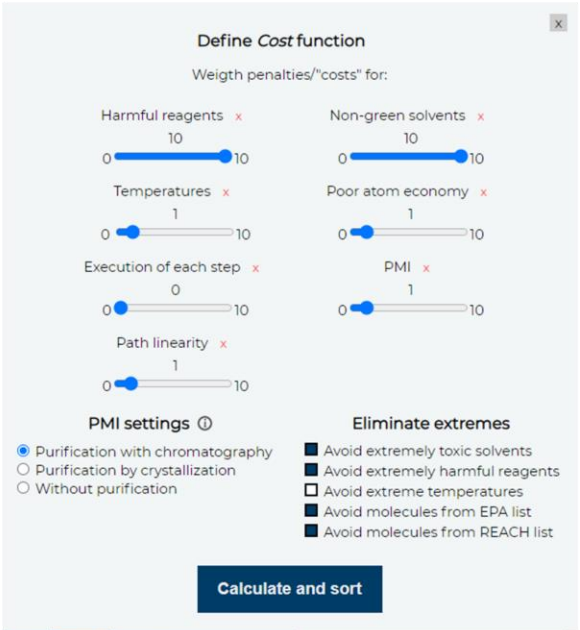


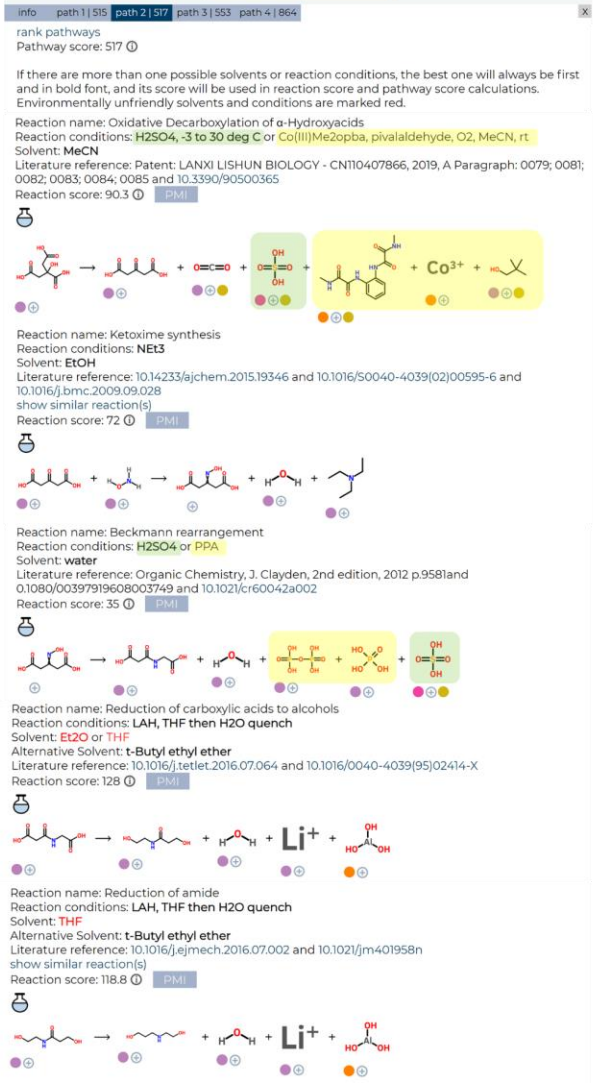
Figure 1. Degradative reaction networks. Screenshot from Allchemy showing a network of degradative reactions originating from inexpensive (< 1 $/g) biomass, citric acid feed (node at the bottom of the network, in the "zeroth generation" marked G_0). Iterative application of reaction rules yields molecules corresponding to nodes in generations G_1 to G_5. Here, the sizes of the nodes correspond to the numbers of "incoming" connections (i.e., how many reactions lead to a given node) whereas colors correspond to prices per gram according to the Mcule catalog. Price ranges are indicated in the legend. One of the value-added molecules in G_5 is 3-[(2-hydroxyethyl)amino]propan-1-ol. This molecule is useful as a building block in pharmaceutical synthesis[80-81] and costs 266.76 $/g according to Mcule. Colorful arcs trace multiple degradative routes to this target. Note that the arcs have several common nodes corresponding to common intermediates in the pathways (shown as miniatures next to these "hubs"). For path details corresponding to this example, see **Figure 2**. For details of network generation and analysis, see User Manual in Supplementary **Section S1.2**.

a


Define *Cost* function
Weigth penalties/"costs" for:
Harmful reagents
Non-green solvents
Temperatures
Poor atom economy
Execution of each step
PMI
Path linearity
PMI settings
Purification with chromatography
Purification by crystallization
Without purification
Eliminate extremes
Avoid extremely toxic solvents
Avoid extremely harmful reagents
Avoid extreme temperatures
Avoid molecules from EPA list
Avoid molecules from REACH list
Calculate and sort


Define *Cost* function
Weigth penalties/"costs" for:
Harmful reagents
Non-green solvents
Temperatures
Poor atom economy
Execution of each step
PMI
Path linearity
PMI settings
Purification with chromatography
Purification by crystallization
Without purification
Eliminate extremes
Avoid extremely toxic solvents
Avoid extremely harmful reagents
Avoid extreme temperatures
Avoid molecules from EPA list
Avoid molecules from REACH list
Calculate and sort

b


info
path 1 | 515
path 2 | 517
path 3 | 553
path 4 | 864
rank pathways
Pathway score: 517
If there are more than one possible solvents or reaction conditions, the best one will always be first and in bold font, and its score will be used in reaction score and pathway score calculations. Environmentally unfriendly solvents and conditions are marked red.
Reaction name: Oxidative Decarboxylation of α-Hydroxyacids
Reaction conditions: H2SO4, -3 to 30 deg C or Co(III)Me2opba, pivalaldehyde, O2, MeCN, rt
Solvent: MeCN
Literature reference: Patent: LANXI LISHUN BIOLOGY - CN110407866, 2019, A Paragraph: 0079; 0081; 0082; 0083; 0084; 0085 and 10.3390/90500365
Reaction score: 90.3
PMI
Reaction name: Ketoxime synthesis
Reaction conditions: NEt3
Solvent: EtOH
Literature reference: 10.14233/ajchem.2015.19346 and 10.1016/S0040-4039(02)00595-6 and 10.1016/j.bmc.2009.09.028
show similar reaction(s)
Reaction score: 72
PMI
Reaction name: Beckmann rearrangement
Reaction conditions: H2SO4 or PPA
Solvent: water
Literature reference: Organic Chemistry, J. Clayden, 2nd edition, 2012 p.9581and 0.1080/00397919608003749 and 10.1021/cr60042a002
Reaction score: 35
PMI
Reaction name: Reduction of carboxylic acids to alcohols
Reaction conditions: LAH, THF then H2O quench
Solvent: Et2O or THF
Alternative Solvent: t-Butyl ethyl ether
Literature reference: 10.1016/j.tetlet.2016.07.064 and 10.1016/0040-4039(95)02414-X
Reaction score: 128
PMI
Reaction name: Reduction of amide
Reaction conditions: LAH, THF then H2O quench
Solvent: THF
Alternative Solvent: t-Butyl ethyl ether
Literature reference: 10.1016/j.ejmech.2016.07.002 and 10.1021/jm401958n
show similar reaction(s)
Reaction score: 118.8
PMI

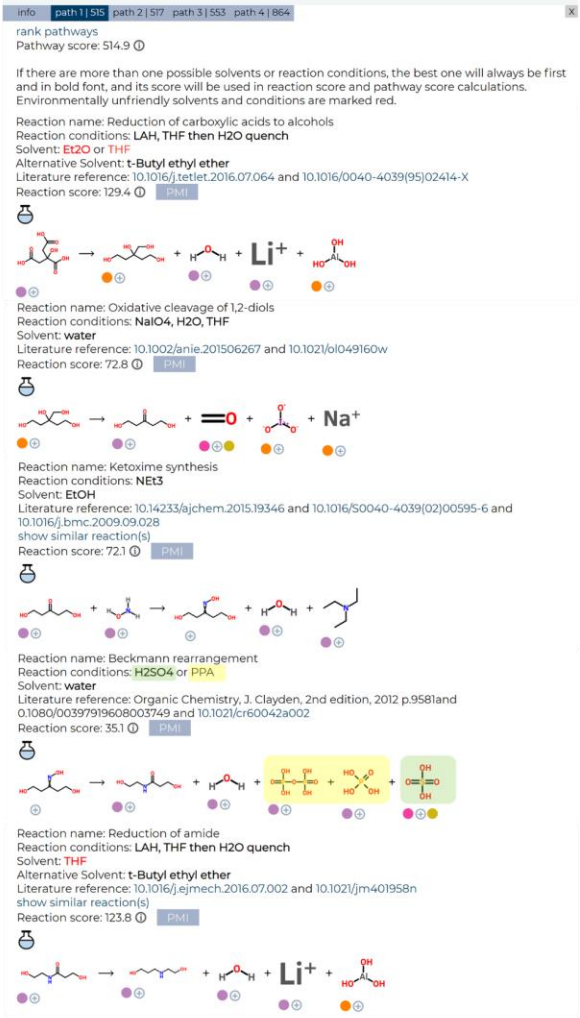

info
path 1 | 515
path 2 | 517
path 3 | 553
path 4 | 864
rank pathways
Pathway score: 514.9
If there are more than one possible solvents or reaction conditions, the best one will always be first and in bold font, and its score will be used in reaction score and pathway score calculations. Environmentally unfriendly solvents and conditions are marked red.
Reaction name: Reduction of carboxylic acids to alcohols
Reaction conditions: LAH, THF then H2O quench
Solvent: Et2O or THF
Alternative Solvent: t-Butyl ethyl ether
Literature reference: 10.1016/j.tetlet.2016.07.064 and 10.1016/0040-4039(95)02414-X
Reaction score: 129.4
PMI
Reaction name: Oxidative cleavage of 1,2-diols
Reaction conditions: NaIO4, H2O, THF
Solvent: water
Literature reference: 10.1002/anie.201506267 and 10.1021/ol049160w
Reaction score: 72.8
PMI
Reaction name: Ketoxime synthesis
Reaction conditions: NEt3
Solvent: EtOH
Literature reference: 10.14233/ajchem.2015.19346 and 10.1016/S0040-4039(02)00595-6 and 10.1016/j.bmc.2009.09.028
show similar reaction(s)
Reaction score: 72.1
PMI
Reaction name: Beckmann rearrangement
Reaction conditions: H2SO4 or PPA
Solvent: water
Literature reference: Organic Chemistry, J. Clayden, 2nd edition, 2012 p.9581and 0.1080/00397919608003749 and 10.1021/cr60042a002
Reaction score: 35.1
PMI
Reaction name: Reduction of amide
Reaction conditions: LAH, THF then H2O quench
Solvent: THF
Alternative Solvent: t-Butyl ethyl ether
Literature reference: 10.1016/j.ejmech.2016.07.002 and 10.1021/jm401958n
show similar reaction(s)
Reaction score: 123.8
PMI

Figure 2. Scoring of degradative pathways according to process criteria. Allchemy screenshots of two out of 17 synthetic paths commencing from citric acid and leading to 3-[(2-hydroxyethyl)amino]propan-1-ol). The routes shown are from amongst those traced by colorful arcs in **Figure 1**. Top panels are interactive scoring functions assigning penalties to different parameters. In **a,** the sliders (for details, see ref [20] and User Manual in Supplementary **Section S1.2**) assign only moderate penalties for the use of non-green conditions, poor atom economy, etc., and the highest ("10") penalty for the "cost" of each step being performed. Under this scoring scheme, the route shown beneath entails several non-green steps. For example, oxidative decarboxylation of α-hydroxyacids uses MeCN which is flagged (in *red* font) according to the GSK criteria.[46] The program does not suggest any greener conditions for this step. The two reductions are also flagged for non-green conditions (Et_2O or THF solvents) but here the program suggests an alternative, greener *t*-butyl ethyl ether solvent. By contrast, the scoring scheme in **b,** penalizes heavily the use of harmful reagents or non-green solvents. In effect, the highest scoring path rests on the oxidative cleavage of 1,2-diols instead of the environmentally problematic oxidative decarboxylation of α-hydroxyacids. Note that all reactions provide by-products (if alternative conditions are available, all by-products are listed). Also, the colored markers indicate: *orange* = known compound, *pink* = EPA-regulated compound, *violet* = buyable chemical, *yellow* = by-product that can be used as a substrate in other, large-scale, industrial reactions (of which Allchemy has a separate reaction-rules set not discussed here).

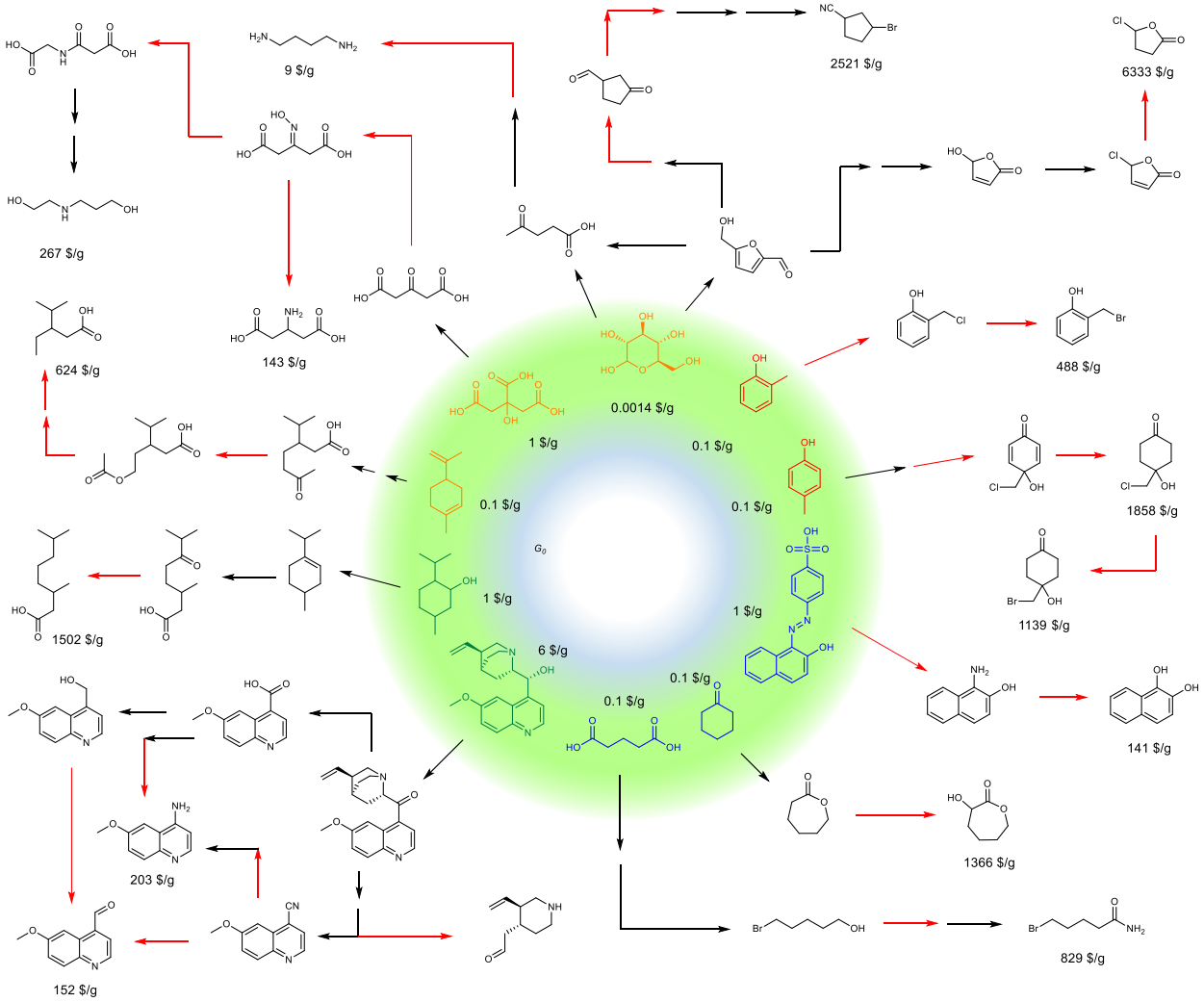


Figure 3. Examples of computer-designed degradation pathways starting from different feeds. The feeds are shown inside the *green* circle, and are from biomass (*green*-colored chemical structures), biomass-waste (*orange*), industrial waste (*blue*) and toxic waste (*pink*). Pathways are abbreviated but numbers of steps correspond to the numbers of arrows – for path details, see Supplementary **Section S2.1.** *Black* arrows indicate steps with literature precedents, *red* arrows indicate unprecedented steps. Prices per gram (from Mcule's catalog) are given next to the feeds and next to the value-added products.

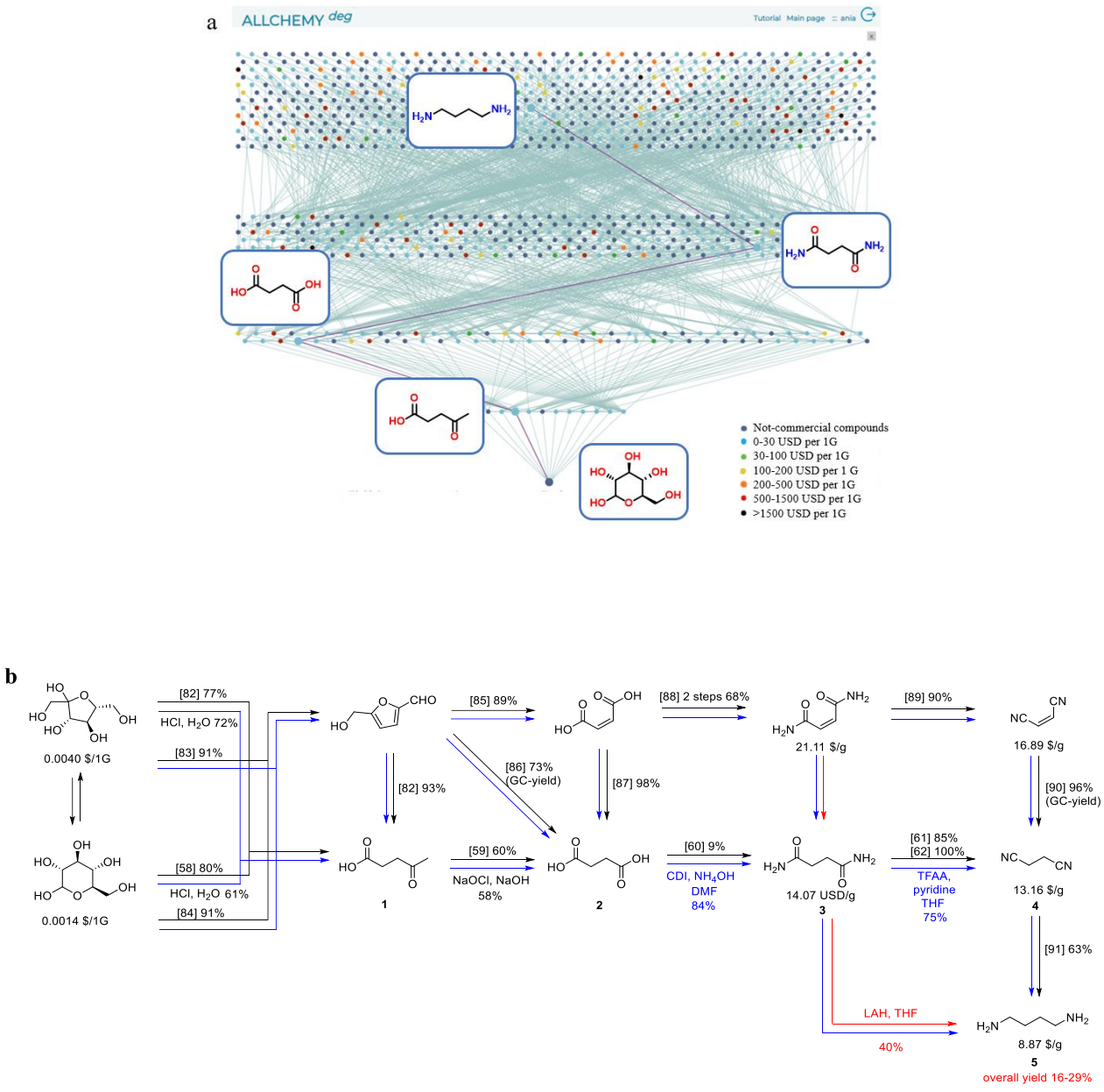


Figure 4. Degradation of D-glucose into value-added 1,4-diaminobutane, 5. **a,** A screenshot from Allchemy showing the network of 950 molecules that can be obtained from D-glucose within 4 synthetic generations. Mcule's prices per gram are indicated by node colors and are explained in the legend on the bottom right. **b,** Schemes of the reactions validated by experiment. *Blue* arrows = computer-planned reactions. *Black* arrows = reactions previously described in the literature

(numbers in square brackets are literature references). Note that in some cases, the literature-reported reactions had low yields or non-green conditions – in such cases, computer-proposed conditions in *blue* were applied. *Red arrows* highlight reactions that have no literature precedent. Most of these reactions are also computer-designed (steps also have *blue* arrows) but some were found useful by a posterior analysis of the network.

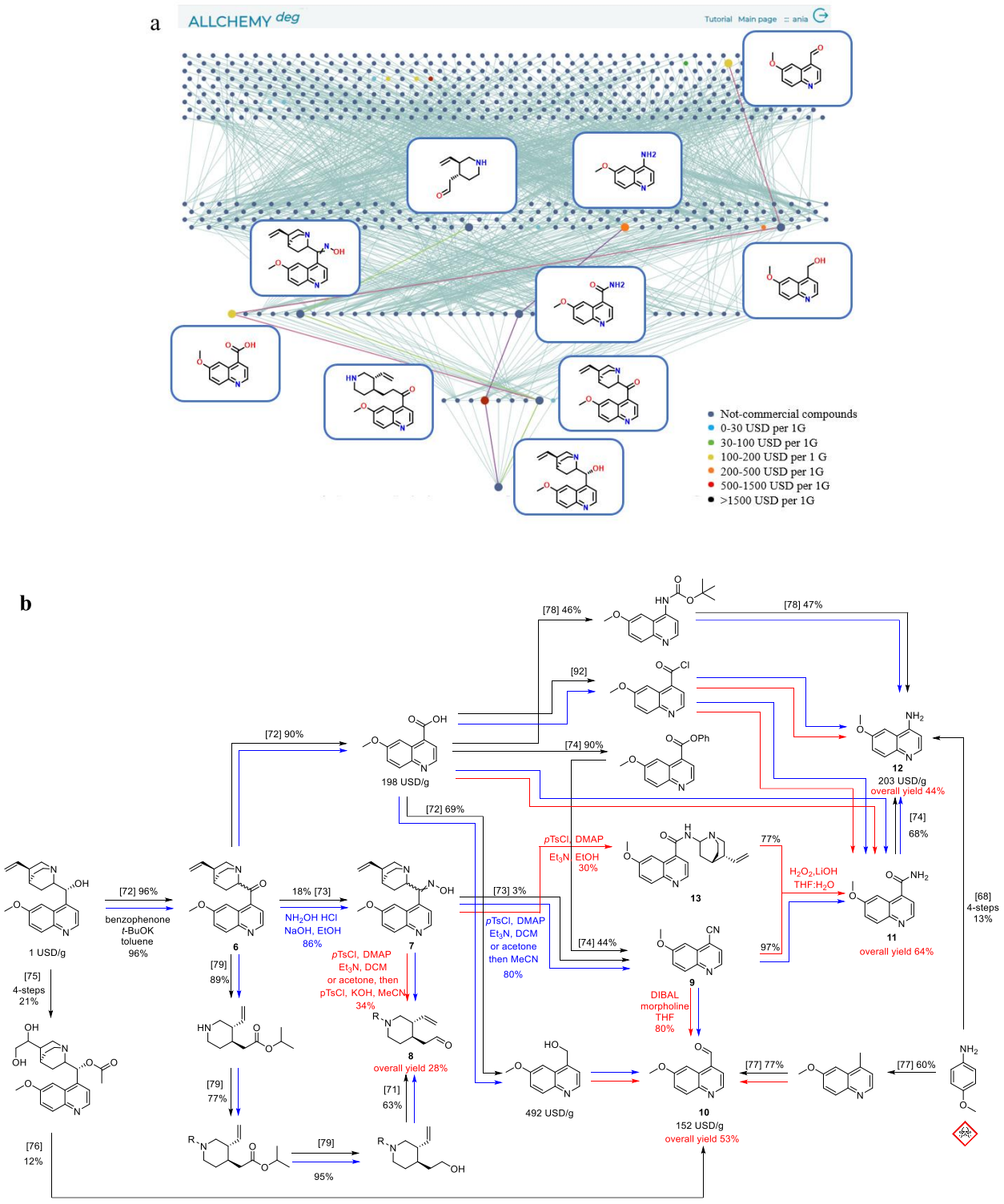


Figure 5. **Degradation of quinine into smaller but value-added chemicals. a,** A screenshot from Allchemy showing the network of 608 molecules that can be obtained by degrading quinine within 4 synthetic generations. Mcule's prices per gram are indicated by node colors and are explained in

the legend on the bottom right. **b,** Scheme detailing experimental syntheses of the molecules of interest, 2-((3*R*,4*R*)-3-vinylpiperidin-4-yl)acetaldehyde, **8**, 6-methoxyquinoline-4-carbaldehyde, **10**, 4-amino-6-methoxyquinoline, **12**. *Blue* arrows = computer-planned reactions. *Black* arrows = reactions previously described in the literature (numbers in square brackets are literature references). Note that in some cases, the literature-reported reactions had low yields or non-green conditions – in such instances, conditions in *blue* were applied. *Red arrows* and *red conditions* highlight reactions that have no literature precedent. Most of these reactions are also computer-designed (steps also have *blue* arrows) but some were found useful by a posterior analysis of the network.