Organic Process Research & Development

Article

pubs.acs.org/OPRD

Route Design in the 21st Century: The IC*SYNTH* Software Tool as an Idea Generator for Synthesis Prediction

Anders Bøgevig,[‡,∥] Hans-Jürgen Federsel,*[,†] Fernando Huerta,[‡,∥] Michael G. Hutchings,*[,§] Hans Kraut,[§] Thomas Langer,[†] Peter Löw,[§] Christoph Oppawsky,[§] Tobias Rein,[‡,∥] and Heinz Saller[§]

[†]Chemical Development, AstraZeneca R&D, Silk Road Business Park, Macclesfield, SK10 2NA Cheshire, U.K.

[‡]Chemnotia AB, Forskargatan 20 J, 151 36 Södertälje, Sweden

[§]InfoChem GmbH, Landsberger Straße 408/V, D-81241 München, Germany

Supporting Information

ABSTRACT: The new computer-aided synthesis design tool IC*SYNTH* has been evaluated by comparing its performance in predicting new ideas for route design to that of historical brainstorm results on a series of commercial pharmaceutical targets, as well as literature data. Examples of its output as an idea generator are described, and the conclusion is that it adds appreciable value to the performance of the professional drug research and development chemist team.

INTRODUCTION

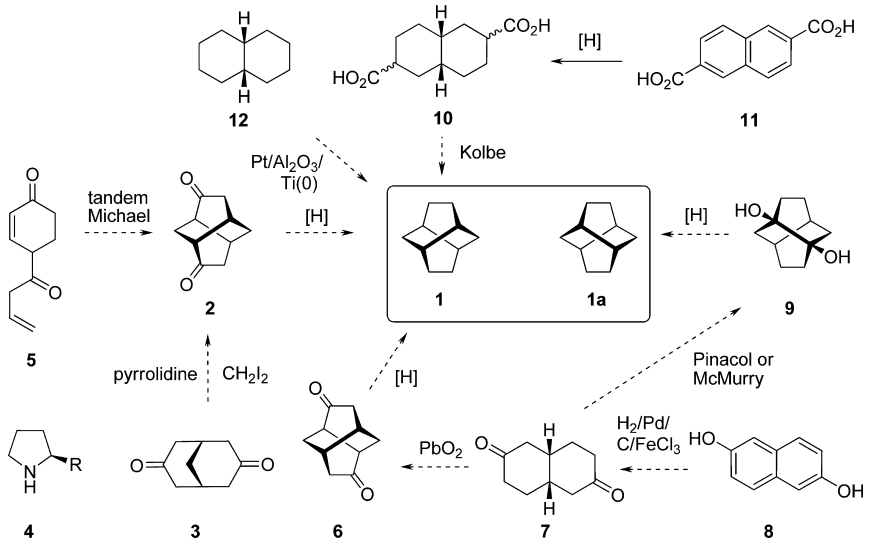
Process R&D in the pharmaceutical industry has to pay heed to a range of criteria, for example, availability of commercial quantities of starting material, chemical safety concerns and potential hazards, toxicity, environmental considerations and sustainability, cost of goods, quality criteria, prior art and the intellectual property situation, to mention some of the foremost,[1] but at the very core of the task lies route design.[2] Success in this key activity relies heavily on the skill of process chemists in applying their knowledge of the near-boundless chemical literature, which, at this stage of the 21st century, continues to expand to levels beyond normal comprehension. Modern topics such as organocatalysis, C−H activation, and new approaches to organofluorines are just three examples of whole areas of chemistry that have evolved relatively recently. Nevertheless, the chemist expects to be able to access the versatility encompassed by this gigantic toolbox. Of course, recourse to modern, electronically searchable databases of chemical structures and reactions such as SciFinder,[3] Spresi,[4] or Reaxys[5] expedites this task. But there remains room for more advanced tools to complement chemist knowledge and derive added value from the mass of information deposited in electronic databases. This article concerns the retrosynthesis design tool IC*SYNTH* and especially its application in various process R&D case studies. It goes beyond a pure database search engine; rather, its role is that of a new idea generator.

A Brief Background. Computer aided synthesis design (CASD) systems for retrosynthetic analysis have existed since E.J. Corey's ground-breaking LHASA[6] in the 1960s began CASD development amongst the organic chemical community. Philip Judson's excellent book from 2009 gives a comprehensive history and description of CASD evolution since the early times.[7] Yet, while predictive tools are now routinely deployed in other aspects of molecular design and chemistry-based R&D, organic synthesis in general and route design in particular in industrial process R&D do not have this type of support. Judson also reviews some of the reasons why the wide-ranging chemical and pharmaceutical industry adoption of CASD in the 1970s and 80s generally failed.[8] Our own experience in ICI[9] acts as a paradigm for those times: unreliable software, cumbersome hardware, insufficient underlying chemical reaction data, injudicious overoptimism within industry, and no common standards or user interfaces for the different systems available at the time, in combination eventually and critically leading to insufficient useful results. All of these converted the initial industrial optimism into indifference and the resultant neglect and eventual disappearance of the tool in industrial laboratories.[8] This outcome was expedited by the somewhat later but overlapping development of the first computer-searchable reaction databases, which were more intuitive to use, and, in any case, successfully addressed many of the questions initially being directed at CASD. In the intervening 20 or so years, the technical background has changed substantially: cheaper, powerful, and far more convenient hardware and peripherals, fast search engines, centralized computing distributed over the WWW, successful reaction mapping and classification,[10] and other chemical algorithm development, leading to well-organized and more easily curated reaction databases that, in turn, encapsulate extensive historical and modern chemistry, all contributing to new optimism. A further factor was fresh encouragement provided by a new generation of process R&D chemists, in the current instance from AstraZeneca (AZ), who were not only unbiased by earlier CASD experiences but also who had the enthusiasm and optimism to believe CASD must be able to contribute to their process R&D goals. In combination, these led to InfoChem's development of its retrosynthesis tool IC*SYNTH*, which went through various experimental and ultimately commercialized versions[11] between 2005 and 2013 centered on Java applet technology. The current 2014 version has been further developed and substantially re-engineered as an HTML5

Received: November 28, 2014
Published: January 22, 2015

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357−368

Scheme 1. IC*SYNTH* Suggestions for Novel Syntheses of Twistane[a]



[a]Note that in this scheme and all of those that follow, the convention is that solid arrows represent reactions known to work in practice, whereas dotted arrows represent speculative suggestions, either direct from the computer or extrapolated from its output by the chemist-user.

application. It is not the goal of this article to delve into its underpinning technology, although overviews of user features and technical design are to be found in Supporting Information. Rather, the main aim is to provide evidence for its successful performance as an idea generator including its use in real process R&D problems, and this is what we now turn to.

EVALUATION OF IC*SYNTH* AS AN IDEA GENERATOR

The best way to demonstrate the system's utility as an idea generator is through examples where it has produced unexpected retrosynthesis suggestions beyond commonly anticipated experience. We know that it has directly answered synthetic problems of industrial chemists that have been implemented in ongoing route design work. Similarly, demonstrations to senior academics have resulted in novel suggestions for access to their synthetic targets beyond professorial expertise, which are now being incorporated into their active research programs. A complication is that what is a new idea to one chemist may be claimed as common knowledge or a routine suggestion by another: novelty and originality are clearly subjective. In order to demonstrate the tool's capabilities, we have set up what we regard as a fair test. Besides the open literature, we have access to detailed brainstorm proposals and in-house experience for routes to various medicinal targets that have been derived by teams of AZ chemists in the course of commercial drug projects. These are relatively recent but no longer active. The overarching goal has been to explore to what extent IC*SYNTH* complements professional synthesis route proposals. If the software was not to add to the results of the brainstorms, then it would not fulfill its role as an idea generator. If, on the other hand, it augments the chemist team's suggestions, then it clearly has a role to play in designing the optimal synthetic route alongside conventional chemist expertise. We have not sought to determine if it is able to find all earlier chemist team suggestions. The key question is whether it can predict new and potentially useful routes and thus can provide a contribution in concert with the chemist. These tests have been carried out blind, in that the computer's searches on defined AZ target molecules were run in the absence of knowledge of the earlier brainstorm suggestions. Several case studies are discussed below, after a more academic teaching molecule is first described and compared with the open literature as an introduction to the sort of results that can be expected. As will become clear, IC*SYNTH* does indeed suggest realistic and potentially valuable and novel synthetic schemes in all of these case studies.

Case Study 1: Twistane. The cage hydrocarbon twistane (**1**; Scheme 1) was recognized 50 years ago as an interesting problem in synthesis[12] and has been a topic in many organic chemistry courses since to teach aspects of synthesis. It has also been used routinely as a target molecule during our development and evaluations of IC*SYNTH*.[13] While twistane is of no interest as a process development target, some of the results generated now serve as a vehicle to introduce aspects of the notion of idea generation as we intend it and to illustrate some of the user features of the system. In this case study, the basis of the test is the open literature. Three of the first five suggestions at level 1, that is, one reaction step from the target, correspond to key intermediates in published syntheses (twistene and both possible twistanones). Some other known twistane precursors appear at lower priority in the synthesis tree. Recovering known chemistry is largely irrelevant in the search for new ideas; this is the main purpose of pure database search tools. However, seeing known routes is nevertheless gratifying; indeed, it would be surprising and disappointing if such suggestions were *not* found. We would have to conclude that there is something deficient with the databases of reactions underpinning the system or with the search and output evaluation (ranking) algorithms. Furthermore, this rediscovery of known chemistry also serves as reassurance for new users. Much more interesting are some of the suggestions for new twistane syntheses, which, as far as we are aware, have not been reported. (Additionally, we have not had the opportunity to follow up any of these in the lab ourselves.) A few are summarized in Scheme 1. Besides showing relevant aspects of chemistry, they are all only a few steps from commercially available starting materials. The shortest literature routes for

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357−368

twistane are 3 steps,[12] and we are arbitrarily using this as the criterion to judge the value of the computer's output.

A high ranking suggestion is one of the six possible diones based on the twistane skeleton (**2**; Scheme 1). Its [3.3.1]-bicyclic dione precursor **3** is commercially available or may be synthesized readily by condensing 2 equiv of 3-ketoglutarate and malondialdehyde followed by full decarboxylation.[14] The new route is a transannular α,α'-double ketone alkylation of **3** with CH_2I_2. This chemistry is a simple extrapolation from the same reaction reported for the construction of the isomeric adamantane skeleton.[15] The alkylation is promoted by pyrrolidine, implying enamine intermediates. While the starting dione **3** is achiral, twistane is chiral (D_2 symmetry; the enantiomers are represented as **1** and **1a** in Scheme 1), so if one of the several known enantiomerically pure proline-based pyrrolidines **4** were to be used as enamine base, then it can be inferred that an enantioselective synthesis of twistanedione **2**, and thus twistane itself, might result.[16] An alternative approach to dione **2** is based on the acylated cyclohexenone **5**. Here, two classical Michael additions in tandem are suggested, based on a single literature analogy, where an allylphosphonate ester has dialkylated cyclohexenone via double Michael reaction.[17] At first sight, this suggestion appears sensible. The first step would lead to an unsaturated decalin diketone which is set up, in principle, to undergo a second transannular Michael addition. However, an immediate first question concerns the stereochemistry to be expected for the newly formed decalin ring junction: *cis*-stereochemistry is required to permit the second Michael addition step. Even if unwanted *trans*-decalin is initially preferred, one of the ring junction C atoms is formally enolizable, so the required *cis*-geometry might be achievable under equilibrating conditions. A second question results from the observation that the two ketones in **5** are in a vinylogous 1,3-diketone relationship and thus the endocyclic CH alpha to the exocyclic carbonyl should be the more acidic. Requisite deprotonation of the allylic CH_2 unit may, therefore, be disfavored and inhibit or even eliminate the possibility of the first Michael addition. As an alternative to deprotonation, the chemist user applied their knowledge to adapt the idea to a manifold of ketone-enamine equilibria based on **5** (not shown). This includes species set up for the required intramolecular Michael additions, so the idea suggested to use **5** to access **2** may still be realistic. No further searching for access to the level 2 diketone **5** was carried out.

An isomeric twistane diketone **6** has also been suggested by IC*SYNTH* as a novel precursor, based on intramolecular α,α'-oxidative coupling across *cis*-decalindione **7**. Various oxidants have been reported in the literature for such intermolecular couplings, albeit with varying yields.[18] Dione **7** can, in principle, be derived directly from available 2,6-dihydroxynaphthalene **8** by appropriate partial ring-hydrogenation conditions. For example, in a close analogy, $FeCl_3$-modified Pd/C is reported to reduce 1-naphthol to *cis*-1-decalone.[19] Alternatively, dione **7** could undergo reductive transannular coupling to a bridgehead diol of twistane **9** by a pinacol reaction or similar. McMurry coupling may also terminate at this oxidation state, as the usual full reduction to alkene in this bridgehead case is inhibited.[20] Presumably, known deoxidation conditions for tertiary alcohols could be applied to diol **9** to reach twistane. An interesting observation is that these two different methods for formation of twistane skeletons from the same decalindione **7** lead to opposite product chiralities. This is of no consequence in the case of racemic **7**, but if a homochiral isomer of **7** is subjected to the two coupling chemistries, then enantiomeric twistane skeletons in **6** and **9** result, with deoxidation leading to the twistane enantiomers **1** and **1a**, respectively, as depicted in Scheme 1.

A further series of related suggestions relies on intramolecular diradical couplings across the 2,6-positions of various *cis*-decalins. Decomposition of Barton's thiohydroxamate diesters[21] of the decalin dicarboxylic acid **10** in the absence of an external radical trapping agent comprises one possibility. However, a more direct option is Kolbe electrolysis of this diacid. Of course, success depends on avoiding (di)radical rearrangements, external trapping, hydrogen shifts, intermolecular coupling, etc., but its simplicity is attractive. Two steps are implied from the multitonne polymer precursor 2,6-naphthalene dicarboxylic acid **11**.[22] In fact, conceivably, the solution of **10** resulting from hydrogenation of **11** could be filtered to remove catalyst and then electrolyzed directly in a telescoped one-pot synthesis of twistane from the naphthalene diacid **11**, effectively a one-step process.

The final synthesis suggestion included in Scheme 1 is simply transannular dehydrogenation of *cis*-decalin **12**. The precedents invoked are hydrocarbon dehydrogenations, most relevantly of relatively strained medium rings, such as cyclooctane to [3.3.0]-bicyclooctane, and conversion of seco-dodecahedrane to dodecahedrane, using alumina-supported platinum and titanium.[23] While there are reasons to be pessimistic regarding the use of decalin as substrate, the temptation of such a simple one-step route to twistane is hard to ignore.

Most of the new routes to twistane outlined in Scheme 1, as well as the shortest of the literature routes, involve new transannular C−C bond formation across decalin precursors. This is not surprising, as this bond has been historically identified as being the most strategic in retrosynthetic considerations of twistane.[12] Within IC*SYNTH*, a molecular complexity measure[24] is applied to identify the bond (or bonds) that leads to the greatest molecular simplification in the retro direction (bond-breaking), and in the twistane case, the bond ranked as being by far the most favorable of the four possible single C−C breakages results in a decalin skeleton precursor. This, coupled with reactions that result in multiple bond formation, is one factor that helps to determine which precursor suggestions are given highest priority by the search and evaluation algorithms. Lower priority suggestions invoke less attractive bond-making combinations; these appear to be more complex and are likely to result in syntheses from available starting materials comprising more than the target three steps.

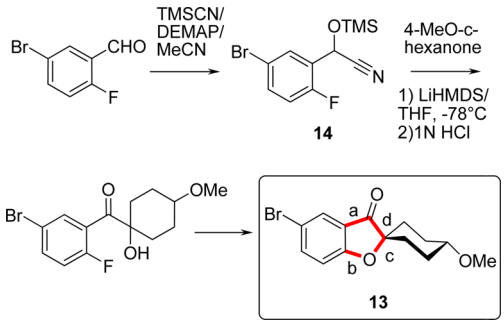
We now move onto synthesis targets more clearly identified with process development objectives. In practice, and counter-intuitively, many of these are small molecules. First, there remain many prospective intermediates that are difficult to access. Known routes that are acceptable in the lab may be unacceptable on scale-up for a variety of reasons, some of which have already been mentioned in the first sentence of the Introduction, or there may even be no known access route. Second, although a final target may be recognized as a complex molecule, the process chemist may be readily able to identify a route that is favorable, apart from its dependence on an unknown key conversion or intermediate. The retrosynthesis problem can then be limited (and simplified) according to the conclusions predefined by the chemist, and IC*SYNTH* searches can be set up accordingly. A particular strength of software-aided route design, as we have already seen, is the unbiased

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357−368

identification of sometimes unconventional sequences easily overlooked even by the trained mind of organic chemists because they involve counterintuitive disconnections or reagents that appear at first sight to be incompatible. However, in particular for process chemists, such suggestions can provide enormous value if the suggested sequence is significantly shorter or otherwise advantageous, e.g., operationally simpler or starting from cheaper commercial raw materials. The following is an example from AZ that prompted the experimental solution to a very real process development problem.

Case Study 2: Oxaspiroketone. Development of efficient syntheses of key intermediates can represent challenges when moving forward from typical medicinal chemistry syntheses to process development. As part of a lead optimization project, the Research Scale-Up Lab at AstraZeneca R&D, Södertälje, Sweden, had the assignment of finding a suitable scale-up route for the oxaspiroketone **13** (Scheme 2), a late intermediate

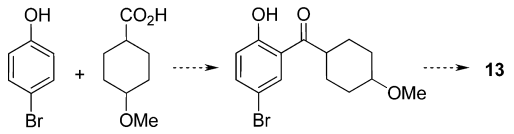
Scheme 2. Initial Route to Target Oxaspiroketone 13 and the Identification of Four Strategic Bonds



on the way to a series of potential Alzheimer's treatments.[25] A synthesis of **13** had already been reported,[26] and this exact route was used by the medicinal chemistry team to prepare the first batch of this target molecule (Scheme 2). Upon evaluation of the results, it was clear that several aspects of this route were not suitable for scale-up. First, the reported yields were not easily reproducible, and, in general, not more than 10% overall yield was obtained. Some of the reagents used for this sequence were not considered to be optimal for scale-up (TMSCN, LiHMDS), and, finally, complex reaction mixtures (e.g., in the hydroxyketone formation) and the resultant tedious chromatographic separations were not desired. An additional important aspect of the task was to find a method to generate the correct stereoisomer of **13** (which was critical for the biological activity of the final compound) or at least a method to isolate the desired isomer easily (e.g., by selective crystallization).

In a retrosynthetic analysis of the target molecule, the four disconnections illustrated for the target **13** (Scheme 2) were considered. Of these, disconnections ***a*** and ***b*** at first sight seemed more appealing since both implied the use of very well-known reactions: in ***a***, an acylation of an aromatic ring, and in ***b***, an aromatic nucleophilic substitution. But a drawback common to these alternatives is that both would involve the use of a cyanohydrin or a related α-hydroxyacyl derivative in one way or another, and, on the basis of our experience of the reported synthesis, we wanted to avoid using intermediates similar to **14**. One particular suggestion from an IC*SYNTH* search that caught the attention of the team was a synthesis reliant on the nonintuitive disconnection ***c*** as the last step (Scheme 3). Figures 1 and 2 are screen shots of the actual output produced.

Scheme 3. Unconventional Construction of Oxaspiroketone 13 Suggested by IC*SYNTH*



The overall synthesis suggested comprises a two-step sequence: first, a Friedel–Crafts acylation between *p*-bromophenol (**15**; node 23 in Figure 2) and 4-methoxy cyclohexyl carboxylic acid **16**, followed by cyclization of the hydroxyaryl ketone intermediate **17** (node 17 in Figure 2).

Analysis of the literature precedent[27] for the final step leading to the target **13** (see also the right-hand screen shot in Figure 2) showed an interesting and unusual reaction pathway (Scheme 4), in which intramolecular triflate migration in **18** is followed by ring closure of enol triflate **19** to give cyclic ketone **20**, where the carbon α to the ketone has been oxidized and the sulfur atom of the triflate, reduced. As far as we can determine, this chemistry is restricted to a single report.[27]

Encouraged by the discovery of this unusual last step and the examples reported using this methodology, efforts were directed toward the synthesis of the intermediate hydroxyaryl ketone **17**. The first step suggested by the computer seemed to be straightforward enough to evaluate in the lab, as both starting materials (**15** and **16**) were commercially available. Unfortunately, the Friedel–Crafts acylation did not work quite as well as hoped, and only by using polyphosphoric acid could the corresponding aryl cyclohexyl ketone be obtained. Obviously, this method was far from optimal for scale-up purposes, so other alternatives were sought. In practice, the final version of the synthesis was performed by reacting the Weinreb amide **21** derived from a mixture of *cis*- and *trans*-4-methoxycyclohexyl carboxylic acid **16** with 2,4-dibromomethoxybenzene pretreated with *n*-Bu_3LiMg **22**.[28] The resulting crude reaction mixture was treated with $AlCl_3$ in dichloromethane, yielding the hydroxyketone intermediate **17** that was purified by crystallization from aqueous methanol. Finally, triflate formation and reaction with DBU in 2-methyltetrahydrofuran gave the desired compound in good yield and a 10:1 diastereomer ratio in favor of the target **13** after the final crystallization (Scheme 5).

In summary, the initial retrosynthetic analysis had not identified disconnection ***c*** (Scheme 2), but this option was quickly found when IC*SYNTH* was used to enhance the idea generation. Once the precedent reported by Coe et al.[27] came to light, the new route was readily developed. This example thus illustrates that the new system can support synthesis planning by identifying unconventional and unusual transformations that are highly relevant for the specific case at hand.

Case Study 3: Aminoalkylpyrimidine. The particular challenge to prepare this required pyrimidine-based chiral amine **23** stemmed from the rather limited supply of suitable 5-fluoropyrimidine precursors. Hence, the medicinal chemistry route involved a sequence of selective functional group interconversions starting from 2,4-dichloro-5-fluoropyrimidine (Scheme 6), a commercial precursor with the required

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357–368

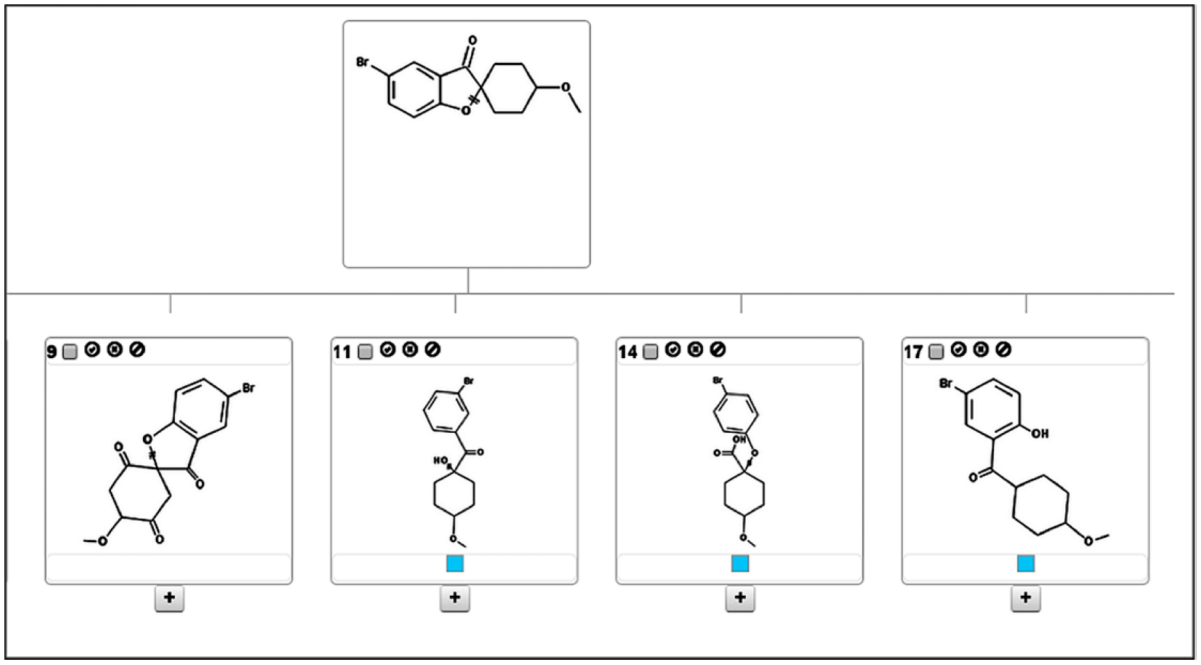


Figure 1. Part of a screen shot of an IC*SYNTH* output tree for a retrosynthesis search of oxaspiroketone **13**. The target and some of the level 1 precursors are shown. Node 17 led to the route reduced to practice described in the text.

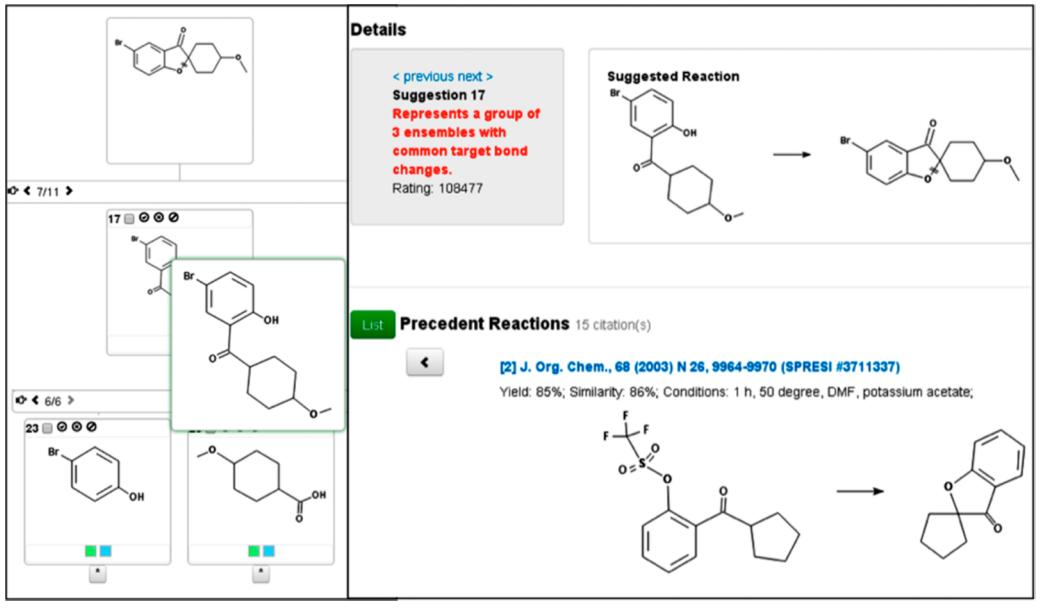
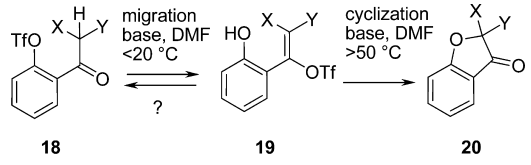


Figure 2. (Left) Alternative view of a result tree for oxaspiroketone **13**, where just one precursor is displayed at each level (2 precursor levels are shown) and others are accessed by scrolling. Hovering the mouse over a node (#17) magnifies it, and clicking leads to another screen (right), which displays precedent chemistry. The literature citation is an active link that leads to the original literature (with an appropriate license).

Scheme 4. Mechanism of Oxaspiroketone Formation, Involving Oxidation at the α-Ketone Site and Triflate Reduction

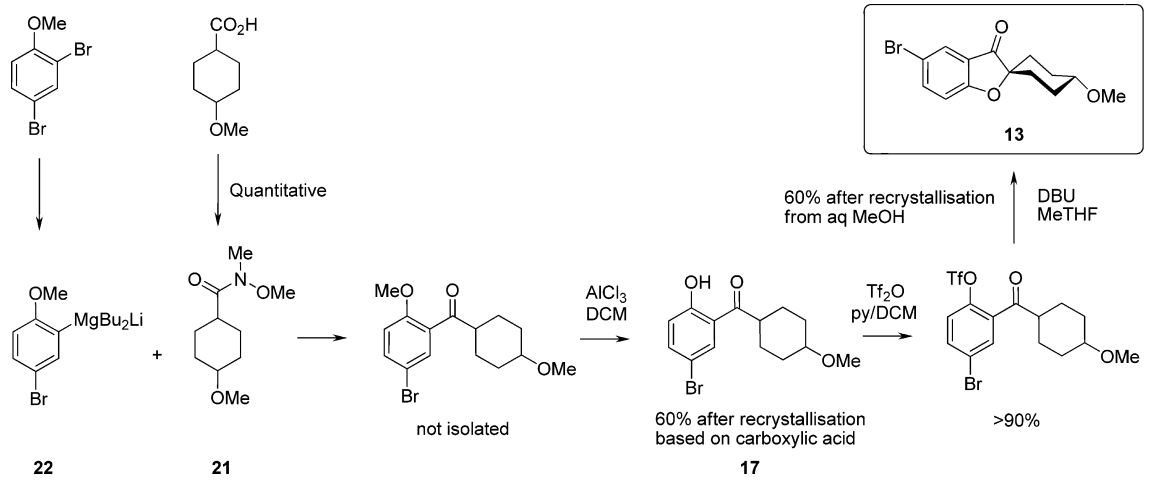


preinstalled functionality.[29] However, the use of multiple (transition) metals with the associated laborious and lengthy workup and subsequent long residence times in the pilot plant rendered this sequence unfavorable for scale-up. Furthermore, the use of toxic cyanide simply to introduce a carbonyl function did not seem to be justified. Various alternative routes to pyrimidine **23** had been suggested in brainstorm sessions to circumvent or simplify at least some of the challenges (Scheme 7).

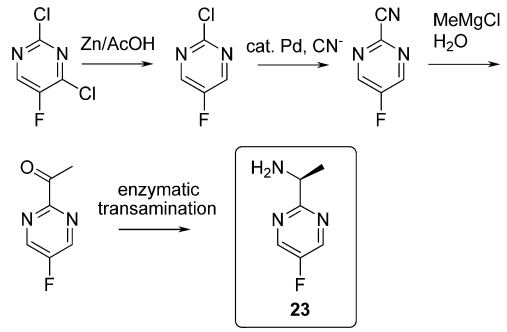
Scheme 7 represents a selection of route suggestions that were considered at the time and reflect just a fraction of creative options and variants to access the target molecule in different ways. However, all proposals are based on two principal concepts: either to construct the pyrimidine ring from scratch

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357−368

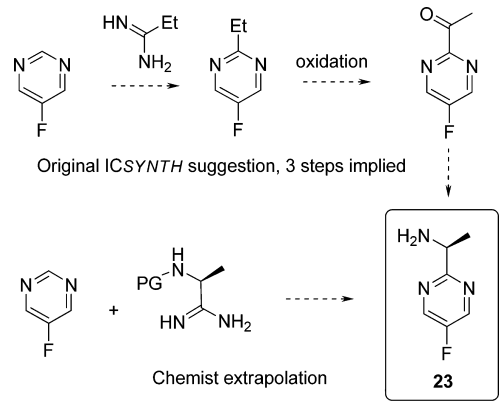
Scheme 5. Developed Route to Spiroketone 13 Based on IC*SYNTH* Prediction



Scheme 6. Medicinal Chemistry Route to Target Pyrimidine 23

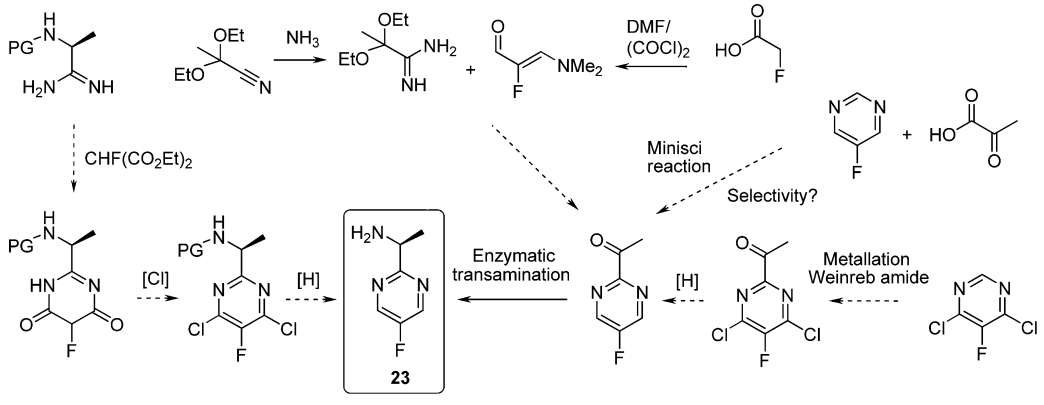


Scheme 8. IC*SYNTH* Proposal and Further Chemist Extrapolation



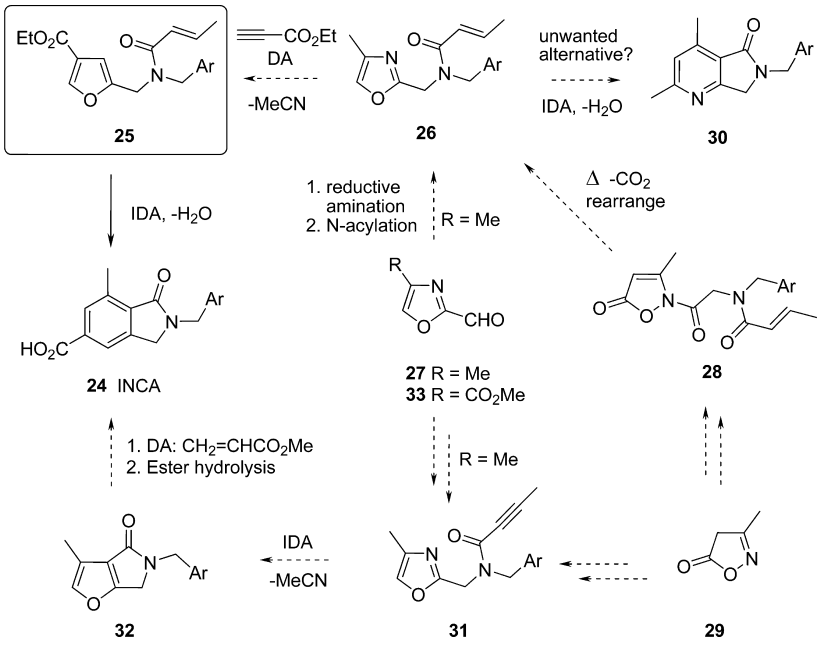
or to transform functional groups of a commercially available and suitably substituted precursor. When **23** was submitted as target to an IC*SYNTH* search, the program returned various suggestions already highlighted by colleagues in the brainstorming session, but, additionally, it returned one unusual but highly attractive proposal: a degenerate ring transformation starting from inexpensive and readily available 5-fluoropyrimidine (Scheme 8). This is conceptually complementary to previous suggestions, as the sequence involves fragmentation of the precursor and reassembly of the ring replacing one N−C fragment with another (the transformation is degenerate, as the same ring system is generated). This sequence is reported[30] to provide higher yields with branched amidines, does not require transition metals, and is shorter. Furthermore, the precedent reaction is reported to work equally well in the absence of the 4-nitro group present in the pyrimidine cited as the literature precedent. It is thus inferred that the required 5-fluoro substituent in the starting material would be tolerated. Finally, as chiral amidines are readily accessible from *N*-protected chiral amino acid esters,[31] a direct and very short route to **23** might be feasible. In the event, this project was terminated before the idea could be put into practice. It is fair to say that, as a concept,

Scheme 7. Alternative Routes to Pyrimidine 23 Suggested in Chemist Brainstorming Sessions



DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357−368

Scheme 9. Cascade Cycloaddition Routes to INCA 24[a]



[a]Ar = 4-$CF_3OC_6H_4$.

degenerate ring transformations are less familiar to organic chemists, and this demonstrates the value that an unbiased, computer-aided approach can offer.

Case Study 4: Furan Precursor of INCA. The *N*-benzylated isoindolinone (**24**; INCA) was required as an intermediate for a final active medicinal ingredient. Various conventional functional group interchanges around a benzene ring had been developed, but none were considered to be ideal for scale-up. Process research of several route alternatives had concluded that a successful intramolecular Diels–Alder reaction[32] (IDA) from furan **25** offered the best way forward. Routes into furan **25** were devised and explored experimentally, including starting with formylation of 3-furanoate esters under various conditions. This rather straightforward approach from very cheap raw materials[33] proved to be unexpectedly complex in the lab: the Vilsmeier reaction with ethyl furan-3-carboxylate was so slow that the process had to be carried out solvent free in neat $POCl_3$/DMF at elevated temperature close to the point of uncontrolled self-heating of the mixture in order to achieve appreciable reaction rates. Development of a safe failure mode eventually allowed accommodation of this process in a conventional plant, and manufacture of several batches proceeded without problems. The project was terminated prior to establishment of the final manufacturing route, but the inherently hazardous nature of the process and generation of the highly carcinogenic *N*,*N*-dimethylcarbamoyl chloride (Me_2NCOCl) warranted further brainstorm analysis at the time, and more recently, intermediate **25** has become the target of IC*SYNTH* searches. Many routes were forthcoming from the computer, including many previously suggested by the chemist team. Some of the more novel and noteworthy are now considered further.

The chemistry in Scheme 9 revolves around cascades of Diels–Alder (DA) and other thermal reactions. Thus, IC*SYNTH* suggested an intermolecular DA reaction between propiolate ester and oxazole **26**, which not only brings about the required [4 + 2] addition but also spontaneously causes elimination of acetonitrile in a reverse hetero-Diels–Alder reaction.[34] Amongst other suggestions, the oxazole **26** is probably best accessed by reductive amination of commercially available formyloxazole **27** and *N*-crotonoylation. An alternative suggestion from the computer is vapor-phase thermolysis or photolysis of *N*-acylated isoxazolidinone **28**, itself derived by straightforward *N*-acylation routes from commercial isoxazolidinone **29** (from acetoacetate and hydroxylamine).[35] The decarboxylative rearrangement is presumably driven by homolysis of the relatively weak N–O bond in **28** followed by decarboxylation and alternative ring closure of a formal diradical. Thus, the process routes **27** or **29** → **26** → **25** → **24** can be purely thermal and, in principle, could comprise attractive one-pot cascades. However, chemist evaluation of the scheme rather rapidly revealed a first potential flaw, in that oxazole **26** could alternatively undergo IDA itself, which likely would be followed by H_2O elimination, terminating in the tetra-substituted and wholly useless pyridine **30**. A simple, if inelegant, workaround would be to replace the crotonoyl group in **26** and its predecessors by H or an *N*-protecting group and to install it only into the penultimate furan intermediate **25**. However, alternatives that retain any elegance present in the potential cascade processes take advantage of the initial IDA step by inverting the unsaturated units central to the various DA steps. Thus, the crotonoyl amide unit is replaced by tetrolic acid amide in **31**, derived from the same commercially available materials **27** or **29**. IDA with acetonitrile elimination delivers the annelated bicyclic furan **32** ready for conventional DA addition to acrylate as dienophile.[36] Dehydration of the [4 + 2] intermediate and ester hydrolysis (*in situ*?) would then give the required INCA **24**. Again, all of this could be carried out in a thermal, one-pot cascade process. However, a second question concerns the regioselectivity of the dienophile DA addition to oxazole **26** and furan **32**. Oxazole plus acetylene DA addition appears to be influenced by the nature of the oxazole 4-substituent, where increasingly electron-donating groups favor the undesired 2,3-disubstituted product regioisomer.[37] It

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357–368

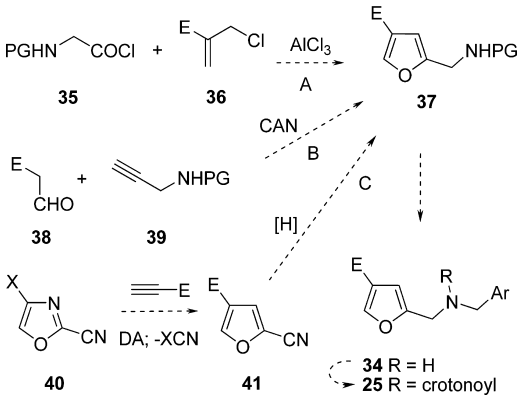
therefore seems that inclusion of an electron-withdrawing group at the oxazole 4-position, as in **33** (Scheme 9), may well drive the DA cycloaddition in the desired direction implied by the 2,4-disubstitution of furan **25** rather than its unwanted 2,3-disubstituted isomer.

Cascade reactions involving skeletal rearrangements with extrusion of molecular fragments are notoriously difficult to recognize in retrosynthesis, and none of this chemistry had been considered in the earlier experimental or brainstorming work, but it was readily accepted *post facto* by chemists involved as a complementary concept and attractive development option.

It is worth emphasizing the respective roles of chemist and computer in developing the suggestions in Scheme 9. The direction of the process, centered on IDA of a furan precursor, emanated from the chemists. The computer then provided ideas for the core of the cascade of thermal processes to access this furan. It did not, however, recognize the possibility of unwanted alternative chemistry. (The system does, in fact, identify unacceptably strained precursors, functional group conflicts, and some selectivity issues, but not in this case.) The chemist's experience saw this and was able to suggest workarounds that retain the basic route ideas. Thus, we are not proposing that IC*SYNTH* can or should replace expert synthesis chemists. However, this case study demonstrates again that it can add appreciable value as a member of a process chemist team.

The additional novel computer-generated ideas in Scheme 10 are relatively short, judged to be realistic, and have the added

Scheme 10. Further IC*SYNTH* Routes to Furan 25[a]

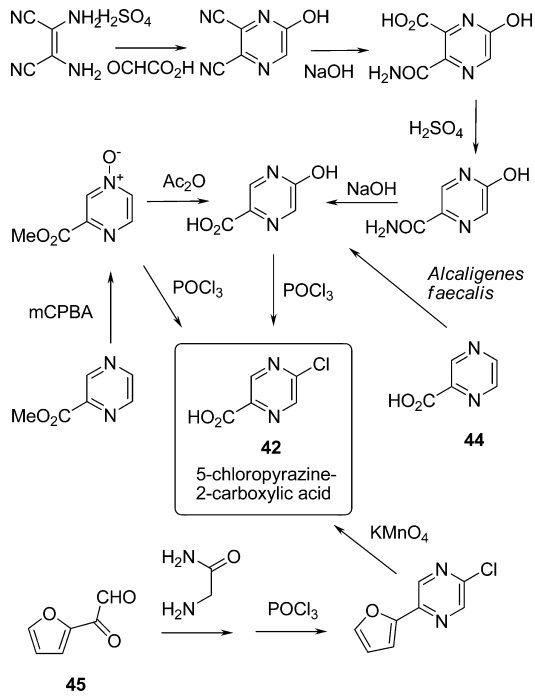


[a]E = CO_2Et; Ar = 4-$CF_3OC_6H_4$; PG = protecting group or H; see text

attraction they could conceivably be telescoped. (A) $AlCl_3$-promoted reactions between acyl halides and methallyl halides that give 2,4-disubstituted furans are reported.[38] Extrapolating this chemistry to the *N*-protected acyl halide of glycine **35** and commercially available chloromethacrylate **36** leads to *N*-protected furan **37** in one step. Conversion to INCA precursor **25** via **34** is then anticipated to be routine. (B) The same product **37** could result from oxidative addition of formyl acetate **38** to *N*-protected propargylamine **39**. Ceric ammonium nitrate has been applied as oxidant in precedent chemistry,[39] but, conceivably, other oxidizing metal ions could also be effective. (C) The final entry in this reaction menu derives from simple hydrogenation of the 2-cyanofuran **41** to give unprotected **37** (PG = H) directly. Although **41** is commercially available, it can be alternatively accessed by DA addition of propiolate to cyanooxazole **40**. Again, regiochemistry of the DA addition is likely to give mixtures for the 4-methyl oxazole (**40**; X = Me), but the possibility of biasing the cycloaddition toward the required regioisomer by a 4-EWG in **40** (e.g., X = CO_2Me, leading to elimination of Mander's reagent $NCCO_2Me$ from the initial [4 + 2] adduct) is worth investigation.[40]

Case Study 5: Unsymmetrically 2,5-Disubstituted Pyrazine. There are occasions when an unbiased computer-aided retrosynthetic analysis can even complement route suggestions for simple molecules that are well-described in the literature and commercially available, as this case study illustrates. Unsymmetrically substituted pyrazine **42** (Scheme 11) was required in bulk quantities for further reaction with a

Scheme 11. Published Routes to 2-Chloropyrazine-5-carboxylic Acid 42

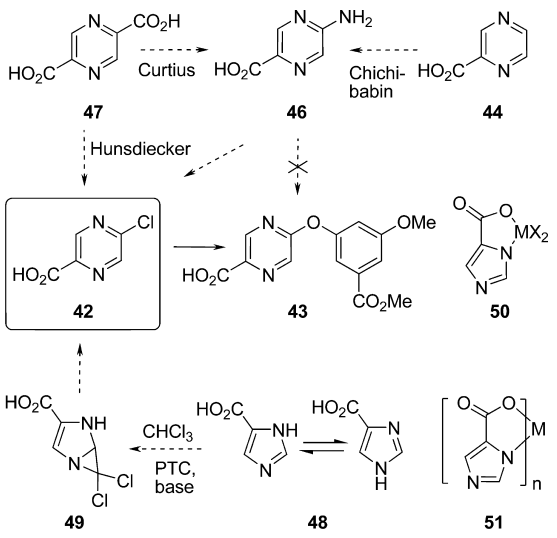


trisubstituted phenol to give target **43** (see below). In bulk quantities, >100 kg, compound **42** was surprisingly expensive, and the quality varied depending on supplier: very pure (but most expensive) material was derived from a patented enzymatic oxidation[41] of carboxypyrazine **44**, whereas cheaper but less clean material was manufactured according to a published 4-step sequence starting from condensation of 1,2-diaminomaleodinitrile with glyoxylic acid.[42] Amongst alternatives, pyrazine-*N*-oxides are thermally unstable compounds (some are known to be explosive), so the short sequence of oxidation of pyrazine-2-carboxylic acid followed by reaction with a chlorinating reagent[43] did not lend itself to safe scale-up. Similarly, limited access to 2-furyl glyoxal **45** and oxidative degradation of a large part of an intermediate in the last step were not regarded as an atom-efficient long-term supply route.[44]

A fair number of alternative routes had been suggested by AZ chemists, but when compound **42** was subjected to analysis by IC*SYNTH*, the following additional interesting concepts were identified (Scheme 12). The suggested Chichibabin reaction applied to pyrazine carboxylic acid **44** is unlikely to give the

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357−368

Scheme 12. IC*SYNTH* Routes to Unsymmetrically Disubstituted Pyrazine 42



desired target compound **46** in appreciable selectivity. This proposal, amongst others, highlights the necessity for further evaluation of suggestions by a synthetic chemist with the benefit of mechanistic understanding currently absent from the system. However, conversion of pyrazine diacid **47** to **46** is a realistic proposal, and, by taking advantage of recent developments in flow chemistry allowing the use of DPPA at scale in a safer environment than batch mode, the suggested desymmetrization through mono-Curtius reaction is an interesting complementary concept in this context. The analogous desymmetrization of **47** by Hunsdiecker reaction or its Kochi modification[45] is, in its proposed form, of little interest for process chemists due to use of silver and lead salts. However, modern variants, such as the Barton thiohydroxamate reaction, might offer metal-free alternatives.[20] Diazotization of the primary amine of **46** and chloride replacement to generate **42** is well-precedented in the literature, either by Sandmeyer reaction or diazonium ion hydrolysis and $POCl_3$ chlorination, rendering this sequence a realistic and shorter access to **42**. Unfortunately, an attractive option involving direct linkage between the diazonium salt from amine **46** and the OH of phenol in a Buchwald-type C−O coupling to give aryl ether **43** (the actual target in this development project) is unknown.[46] Classical arylazo coupling is inevitably kinetically favored.

The IC*SYNTH* suggestion of direct conversion of imidazole-4-carboxylic acid **48** to target **42** is novel, attractively short, and relies on potentially highly scalable chemistry, but it would be useful only if selectivity on various levels could be achieved. The classical Ciamician−Dennstedt rearrangement (the abnormal Reimer−Tiemann reaction) proceeds through dichlorocarbene addition to a CC double bond.[47] Cyclopropyl ring opening at the endocyclic bond of the resulting fused dichlorocyclopropyl ring gives ring expansion of the substrate (e.g., pyrrole to pyridine). Conversely, opening of one of the two exocyclic bonds of the cyclopropyl ring leads, after hydrolysis, to formation of an aldehyde (the conventional Reimer−Tiemann reaction). Useful selectivity in favor of the desired ring expansion in moderate to good yields can be achieved by phase transfer catalysis.[48] Although dichlorocarbene as an electrophilic species usually adds to CC double bonds, the addition to CN double bonds in imidazoles has been reported, albeit at that time under non-PTC conditions.[49] Chemoselectivity in compound **48** in favor of the desired C═N attack may be possible, as the alternative C═C bond is comparatively electron deficient. Further regiochemical discrimination between the two possible CN bond additions due to tautomerism within the imidazole **48** must also be considered. A possible approach to avoid unwanted dichlorocarbene addition across the C(2)−N(3) bond (leading to the undesired isomer 6-chloropyrazine-2-carboxylic acid) and direct it to the required N(1)−C(2) bond highlighted in Scheme 12, to give carbene adduct **49**, is to lock in the required tautomer, for example, by using the 4-carboxy group to tether potential electropositive units at N(3) such as silicon (**50**; M = Si), borate (**50**; M = B^-), or a metal ion (**51**). Awareness of the susceptibility to decarboxylation of carboxyimidazoles[50] suggests that masking of the CO_2H group (ester, amide) may be necessary. Nevertheless, the opportunity for a one-step ring expansion under potentially mild and scalable PTC conditions would remain a high-priority option were further product development of this target desirable.

■ OVERVIEW AND CONCLUSIONS

We started by defining an objective to test IC*SYNTH* against previously known ideas for the synthesis of target molecules, the latter especially including the results of professional chemist brainstorms. In fact, after the 50 or so years of CASD development, we believe that this article constitutes the first published comparison, conducted under controlled conditions, of the relative performances of a CASD tool and organic chemist experts, each facing a series of synthesis targets.[51] The major conclusion is that in all cases the computer has been able to identify new ideas for defining routes to synthetic targets that go beyond known chemist-derived suggestions. However, we emphasize that this result in no way detracts from the continuing central importance of the chemist, both in their own generation of new route options (which may well go beyond what the computer suggests) and in the evaluation of the computer's suggestions. In fact, we also find that a computer-derived new idea can lead the open-minded chemist to further new ideas of his/her own. Thus, there is frequently a positive synergy between chemist and computer.

The five case studies we have described provide different types of highlighted solutions to synthesis targets. Case study 1, twistane, is a comprehensively studied and well-known literature molecule, for which new routes can still be suggested. Case study 2, an oxaspiroketone, shows that an unbiased search can lead to a nonintuitive solution to a synthesis problem, in this case followed up successfully in the development lab. Case study 3, a disubstituted pyrimidine, uncovered an uncommon and again nonintuitive degenerate ring synthesis. Case study 4, a highly substituted benzisoindolinone, can be accessed by a difficult-to-spot potential cascade of cycloaddition and fragmentation reactions, and case study 5, a commercially available disubstituted pyrazine, leads to a suggestion based on a poorly exemplified and low-yielding imidazole ring expansion via potentially attractive chemistry. In common for all of these is that the selected solutions appear only a few steps from the target molecules. Conversely, useful solutions can appear essentially anywhere across a level of the synthesis tree, which raises the question of prioritization of output and then the wider question of route evaluation. The overall route development workflow can be regarded as three distinct phases, each with a type of evaluation. The first is idea generation, by

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357−368

chemist and computer. Each suggestion in IC*SYNTH*'s output is automatically scored by a quantitative model reliant mainly on parameters that describe features of the target, suggested precursor, and interconnecting reactions. The order of appearance of the precursors across the tree is determined by this model. In the second phase, the chemist addresses the computer's suggestions with two simple questions: is this new? and might it be of value?, in effect the first step in a feasibility assessment. Follow up includes searches using other data mining tools to evaluate the scope and limitations of the IC*SYNTH* ideas. The results of the overall feasibility assessment, including route suggestions originating from both chemist and computer, are then prioritized for experimental follow-up. Only in the third phase are detailed quantitative route evaluations normally applied in the AZ protocol. These take into account all facets of process development, including those listed in the first sentence of this article. Such route metrics (sometimes, greenness metrics) are becoming more widely applied, both in AZ[52] and amongst others,[53] and, for meaningful application, they require at least preliminary experimental data.

The developments that IC*SYNTH* encapsulates relative to older systems that enable it to provide the positive results demonstrated include the underlying amount of reaction data (>4.4 million reactions and a correspondingly high number of derived transforms), which is orders of magnitude more than in the past, the fact that this data is not restricted to tried-and-tested chemistry but includes many rare reactions perhaps exemplified by just one published example, modern algorithms that automatically and rapidly derive reactions and transforms from abstracted chemistry and carry out fast searches, and central implementation enabling access by anyone with a link into the WWW (and a user id and password) and convenient central maintenance and management, all benefiting from modern hardware. Furthermore, in our experience, IC*SYNTH* is complementary to standard data mining tools. In particular, it enables fast and comprehensive idea generation, including identification of relevant unconventional chemistry as well as complex transformations that are difficult to spot by a manual analysis. Given enough time, some of these transformations could maybe have been identified by standard chemical data mining tools, but one strength is that this tool gives a comprehensive overview in just a few searches.

Currently, IC*SYNTH* has assumed a place as a unique predictive tool for route design in Chemical Development in AZ. While it is finding valuable commercial application in our own and others' hands, it remains a work in progress. For example, the vexing question of chemical noise amongst the results is a problem for some users.[54] Noise includes unwanted suggestions that, for various reasons, evade the chemistry algorithms that attempt to filter out unrealistic and otherwise undesirable output. Improvements in aspects of stereochemistry handling, better chemical selectivity and reaction conflict predictions, and algorithms to enable new starting material-based strategies are all under development. It is recognized that the case study subjects of this article are (intentionally) rather simple molecules (which, nevertheless, are representative of real pharmaceutical development targets). Applications to more challenging and complex synthetic targets are underway.

Finally, in a fundamentally different direction, the scope of IC*SYNTH* for the role of reaction prediction in new molecule design is proving to be particularly valuable and exciting.[55] For now, though, this remains the subject of a future account.

EXPERIMENTAL SECTION

All synthetic procedures described herein refer to case study 2, as depicted in Scheme 5.

Preparation of (5-Bromo-2-methoxyphenyl)(4-methoxycyclohexyl)methanone. *n*-BuLi (48.18 mmol) was added to a solution of butylmagnesium chloride (24.09 mmol) in 2-methyltetrahydrofuran at 0 °C, and the resulting mixture was stirred for 15 min before the addition of 2,4-dibromo-1-methoxybenzene (72.26 mmol) in 2-methyltetrahydrofuran (40 mL). The reaction was monitored by GC-MS, and when all dibromide had been transformed, *N*,4-dimethoxy-*N*-methylcyclohexanecarboxamide (**21**; 60.22 mmol) was added and the mixture was stirred at 15 °C (internal temperature) until it was completed as monitored by GC-MS. After quenching with saturated aqueous NH_4Cl, the phases were separated. The organic phase was dried and evaporated to give the title compound in 96% yield. *m*/*z*, 326 (M^+ + 2), 328 (M^+). ^{1}H NMR (500 MHz, $CDCl_3$) δ ppm 7.56 (s, 1H), 7.49 (m, 1H), 6.82 (m, 1H), 3.86 (s, 3H), 3.40 (m, 1H), 3.30 (s, 3H), 3.17 (m, 1H), 1.84 (m, 4H), 1.66 (m, 2H), 1.53 (m, 2H).

Preparation of (5-Bromo-2-hydroxyphenyl)(4-methoxycyclohexyl)methanone (17). Aluminum chloride ($AlCl_3$) (103.91 mmol) was added to a solution of (5-bromo-2-methoxyphenyl)(4-methoxycyclohexyl)methanone (25.98 mmol) in DCM (64.9 mL) at 0 °C. After stirring the resultant mixture at 0 °C for 2 h, more $AlCl_3$ was added (1 g), and the mixture stirred for an additional hour before it was quenched with water (50 mL) and HCl 1 M (50 mL) at 0 °C. The phases were separated, and the water phase was extracted twice with chloroform. The combination of the organic phases was dried and evaporated, and the product was purified by crystallization from MeTHF/heptane 1:1 to give the title compound in 55.9% yield. *m*/*z* 314 (M^+ + 2), 312 (M^+). ^{1}H NMR (500 MHz, $CDCl_3$) δ ppm 12.48 (s, 1H), 7.85 (s, 1H), 7.55 (m, 1H), 6.91 (m, 1H), 3.53 (m, 1H), 3.34 (s, 3H), 3.25 (m, 1H), 2.06 (m, 2H), 1.93 (m, 2H), 1.68 (m, 2H), 1.57 (m, 2H).

Preparation of 4-Bromo-2-(4-methoxycyclohexanecarbonyl)phenyl Trifluoromethanesulfonate. Pyridine (102.49 mmol) was added to a solution of (5-bromo-2-hydroxyphenyl)(4-methoxycyclohexyl)methanone (34.16 mmol) in DCM (52.0 mL). The resulting mixture was then cooled on an ice bath and stirred for 15 min until the internal temperature was around 0 °C. Triflic anhydride (41.00 mmol) was then slowly added (15 min addition time; fuming suspension formed), the resultant mixture was monitored by GC-MS, and, when conversion reached 90%, additional triflic anhydride (6.83 mmol) was added and the mixture was stirred overnight. The reaction mixture was directly filtered through a short silica plug and eluted with a heptane/DCM mixture, and the resulting solution was concentrated to give the title compound in 99% yield as a 7:3 mixture of cis/trans isomers. This product was used directly in the next step without further treatment. *m*/*z* 446 (M^+ + 2), 444 (M^+).

Preparation of 5-Bromo-4′-methoxy-3*H*-spiro[benzofuran-2,1′-cyclohexan]-3-one (13). 4-Bromo-2-(4-methoxycyclohexanecarbonyl)phenyl trifluoromethanesulfonate (33.69 mmol) was dissolved in methyltetrahydrofuran; then, DBU (84.22 mmol) was added, and the resultant mixture was stirred at room temperature. The reaction was monitored by HPLC, and when no starting material was observed (ca. 2

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357−368

h), HCl (1 M) was added. The phases were separated, and the organic phase washed with KOH (1 M solution), brine, and water, dried, and concentrated to give the title compound in 66.8% yield. The required isomer was isolated by separation on a short silica column, eluting with heptane/Et_2O. After concentration, the single isomer precipitated in a 15% yield. m/z 312 (M^+ + 2), 310 (M^+). ^1H NMR (500 MHz, $CDCl_3$) δ ppm 7.76−7.79 (m, 1H), 7.70 (m, 1H), 7.04 (m, 1H), 3.43 (s, 3H), 3.29−3.37 (m, 1H), 2.08−2.18 (m, 2H), 1.64−1.89 (m, 6H).

■ ASSOCIATED CONTENT

Supporting Information

Section 1: User features of IC*SYNTH*. Section 2: Transform generation. Section 3: Technical design of IC*SYNTH*. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*(H.-J.F.) E-mail: Hans-Jurgen.Federsel@astrazeneca.com.
*(M.G.H.) E-mail: mghutchings@infochem.de.

Notes

The authors declare no competing financial interest.
‖(A.B., F.H., and T.R.) Initiated studies while an employee at AstraZeneca, Södertälje, Sweden.

■ ACKNOWLEDGMENTS

Collaboration between AstraZeneca and InfoChem was initiated by Dr. Adrian Clark, and he, alongside Dr. Jenny Ekegren and Ann-Sofie Krig (all ex-AZ, Södertälje), were instrumental in the development of aspects of IC*SYNTH* as well as in performing important evaluations. Fanny Irlinger (InfoChem) devised early strategies in IC*SYNTH*. Dr. Stephanie North (Allyl Consulting, UK) and business contacts in various companies carrying out user trials have contributed valuable ideas for improvements in the technical design and its user interface.

■ REFERENCES

(1) (a) Federsel, H.-J. *Nature Rev. Drug Discovery* **2003**, *2*, 654−664. (b) Federsel, H.-J. Chiral drug discovery and development−from concept stage to market launch. In *Comprehensive Medicinal Chemistry II*; Triggle, D. J., Taylor, J. B., Eds.; Elsevier: London, 2007; Vol 2, Chapter 2.17, pp 713−736. (c) Federsel, H.-J. *Acc. Chem. Res.* **2009**, *42*, 671−680. (d) *The Art of Process Chemistry*; Yasuda, N., Ed.; Wiley-VCH: Weinheim, Germany, 2011. (e) Federsel, H.-J. *Bioorg. Med. Chem.* **2010**, *18*, 5775−5794. (f) *Pharmaceutical Process Development: Current Chemical and Engineering Challenges*; Blacker, A. J., Williams, M. T., Eds.; RSC: Cambridge, 2011.

(2) Federsel, H.-J. *Green Chem.* **2013**, *15*, 3105−3115.

(3) SciFinder. http://www.cas.org/products/scifinder.

(4) SPRESI^{web}. http://www.spresi.com/.

(5) REAXYS. http://www.elsevier.com/online-tools/reaxys.

(6) Corey, E. J.; Wipke, W. T. *Science* **1969**, *166*, 178−192. LHASA was preceded by the experimental OCSS program.

(7) Judson, P. *Knowledge-Based Expert Systems in Chemistry: Not Counting on Computers*; RSC: Cambridge, 2009.

(8) A notable exception to industry's abandonment of early CASD implementations is the joint work by Bersohn and Sumitomo Chemical leading to SYNSUP. Starting in 1984, this was still being developed within Sumitomo until 2007, and was being used within the company at least until 2010 (information since that time is unavailable). (a) See the internal Sumitomo R&D report: Takabatake, T. *Sumitomo Kagaku*; Sumitomo Chemical Co., Ltd.; vol. *2009-II*; English translation: http://www.sumitomo-chem.co.jp/english/rd/report/theses/docs/20090204_6uf.pdf (downloaded January 5th 2015). (b) Tanaka, A.; Kawai, T.; Takabatake, T.; Okamoto, H.; Bersohn, M. *J. Comp. Aided Chem.* **2009**, *10*, 104−117. (c) Tanaka, A.; Okamoto, H.; Bersohn, M. *J. Chem. Inf. Model.* **2010**, *50*, 327−338.

(9) One of the authors (M.G.H.) spent more than 2 years with Prof. J. Gasteiger at the Technical University Munich as a visiting ICI fellow working on the CASD system EROS and was then, for about 6 years, heavily involved in introducing EROS, LHASA, CHIRON, and CAMEO throughout ICI and training industrial chemists in their use, before failure of CASD in the ICI environment was eventually accepted. In later positions, first in BASF and then a joint spin-off from Bayer and Hoechst, he met colleagues from those companies who previously had corresponding CASD responsibilities. All had parallel experiences of initial optimism followed by gradual loss of interest and withdrawal of support. Two other authors who had been a PDRF (P.L.) and Ph.D. student (H.S.) in the same TUM group tried to continue CASD commercialization when subsequently setting up their company, ChemoData, in 1986. The indifferent commercial appetite for CASD by that time induced them to move to other business opportunities.

(10) Kraut, H.; Eiblmaier, J.; Grethe, G.; Löw, P.; Matuszczyk, H.; Saller, H. *J. Chem. Inf. Model.* **2013**, *53*, 2884−2895.

(11) (a) Eigner-Pitto, V.; Eiblmaier J.; Grethe, G.; Kraut, H.; Saller, H.; Löw, P. *Comprehensive Synthesis Planning Using Multiple Reaction Search Algorithms*; ACS Annual Meeting, Atlanta, GA, 29th March, 2006. (b) Eigner-Pitto, V.; Eiblmaier, J.; Grethe, G.; Kraut, H.; Saller, H.; Löw, P. *Novel Approach to Retrosynthesis; Automatic Generation of Transform Libraries*; ICIC, Nimes, France, 23rd October, 2006. (c) Hutchings, M. G.; Saller, H.; Kraut, H.; Löw, P. *A View of Recent Computer Aided Synthesis Developments*; ICIC, Berlin, Germany, 17th October, 2012. (d) Rein, T. *Application of Predictive Tools to Solve Synthetic Challenges in Drug Discovery and Development*; Symposium on Predictive Chemistry in Drug Discovery and Development, Mölndal, Sweden, 25th September, 2014.

(12) (a) Corey, E. J.; Ohno, M.; Mitra, R. B.; Vatakencherry, P. A. *J. Am. Chem. Soc.* **1964**, *86*, 478−485. (b) Hamon, D. P. G.; Young, R. N. *Aust. J. Chem.* **1976**, *29*, 145−161.

(13) Twistane was the company logo of InfoChem's predecessor company, ChemoData (1986−1989).

(14) (a) Sands, R. D. *J. Org. Chem.* **1983**, *48*, 3362−3363. (b) Bertz, S. H. *J. Org. Chem.* **1985**, *50*, 3585−3592.

(15) Gerlach, H. *Helv. Chim. Acta* **1985**, *68*, 1815−1821.

(16) Nakazaki, M.; Chikamatsu, H.; Taniguchi, M. *Chem. Lett.* **1982**, *11*, 1761−1764.

(17) Phillips, A. M. M. M.; Modro, T. A. *J. Chem. Soc., Perkin Trans. 1* **1991**, 1875−1879.

(18) (a) PbO_2 oxidant: Szakal-Quin, G.; Graham, D. G.; Millington, D. S.; Maltby, D. A.; McPhail, A. T. *J. Org. Chem.* **1986**, *51*, 621−624. (b) LDA/$FeCl_3$ oxidant: Frazier, R. H., Jr.; Harlow, R. L. *J. Org. Chem.* **1980**, *45*, 5408−5411.

(19) (a) Bird, C. W.; Coffee, E. C. J.; Cooper, R. *Chem. Ind.* **1988**, *15*, 498. (b) Bird, C. W.; Butler, H. I.; Coffee, E. C. J.; James, L. M.; Schmidl, B. W. C. *Tetrahedron* **1989**, *45*, 5655−5666.

(20) Kürti, L.; Czakó, B. *Strategic Applications of Named Reactions in Organic synthesis*; Elsevier Academic Press: Boston, MA, 2005; p 276.

(21) See ref 20, p 44.

(22) This hydrogenation does not appear in the open chemical literature, but it has been reported in patent specifications, e.g., Machida, H.; Kedo, K.; Zaima, F. (Mitsubishi Gas Chemical Company, Inc.) Continuous preparation of a hydrogenation product of a high melting point aromatic carboxylic acid using a solid catalyst in a minimal amount of solvent, useful in the production of cyclohexanedicarboxylic acids in high yields. EP1219586A2, 3 July, 2002.

(23) Paquette, L. A.; Miyahara, Y.; Doecke, C. W. *J. Am. Chem. Soc.* **1986**, *108*, 1716−1718.

(24) (a) Bertz, S. H. *J. Am. Chem. Soc.* **1981**, *103*, 3599−3601. (b) Bertz, S. H. *New J. Chem.* **2003**, *27*, 860−869. (c) Bertz, S. H. *New*

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357−368

J. Chem. **2003**, *27*, 870−879. (d) Rücker, C.; Rücker, G.; Bertz, S. H. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 378−386.
(25) See, e.g., Ghosh, A. K.; Osswald, H. L. *Chem. Soc. Rev.* **2014**, *43*, 6765−6813.
(26) (a) Dillard, L. W.; Yuan, J.; Leftheris, K.; Venkatraman, S.; Wu, G.; Jia, L.; Xu, Z.; Cacatian, S.; Morales-Ramos, A.; Singh, S.; Zheng, Y. (Vitae Pharmaceuticals Inc.) New azaspiro bicyclic compounds useful for treating e.g. Alzheimer's disease, cognitive impairment, Down's syndrome, senile dementia, cerebral amyloid angiopathy, and glaucoma. WO 2011106414 A1, 1 Jan, 2011. (b) Cacatian, S.; Claremon, D. A.; Dillard, L. W.; Fuchs, K.; Heine, N.; Jia, L.; Leftheris, K.; McKeever, B.; Morales-Ramos, A.; Singh, S.; Venkatraman, S.; Wu, G.; Wu, Z.; Xu, Z.; Yuan, J.; Zheng, Y. (Vitae Pharmaceuticals Inc.) New heterocyclic spiro compounds are beta-secretase enzyme inhibitors useful to treat e.g. Alzheimer's disease, cognitive impairment, cognitive decline, senile dementia, degenerative dementia and glaucoma. WO 2010105179 A2, 16 Sept, 2010.
(27) Coe, J. W.; Bianco, K. E.; Boscoe, B. P.; Brooks, P. R.; Cox, E. D.; Vetelino, M. G. *J. Org. Chem.* **2003**, *68*, 9964−9970.
(28) Iida, T.; Wada, T.; Tomimoto, K.; Mase, T. *Tetrahedron Lett.* **2001**, *42*, 4841−4844.
(29) (a) Meadows, R. E.; Mulholland, K. R.; Schürman, M.; Golden, M.; Kierkels, H.; Meulenbroeks, E.; Mink, D.; May, O.; Squire, C.; Straatman, H.; Wells, A. S. *Org. Process Res. Dev.* **2013**, *17*, 1117−1122. (b) Frodsham, L.; Golden, M.; Hard, S.; Kenworthy, M. N.; Klauber, D. J.; Leslie, K.; Macleod, C.; Meadows, R. E.; Mulholland, K. R.; Reilly, J.; Squire, C.; Tomasi, S.; Watt, D.; Wells, A. S. *Org. Process Res. Dev.* **2013**, *17*, 1123−1130.
(30) Barczynski, P.; Van der Plas, H. C. *J. Org. Chem.* **1982**, *47*, 1077−1080.
(31) Gielen, H.; Alonso-Alija, C.; Hendrix, M.; Niewoehner, U.; Schauss, D. *Tetrahedron Lett.* **2002**, *43*, 419−421.
(32) (a) Ball, M.; Boyd, A.; Churchill, G.; Cuthbert, M.; Drew, M.; Fielding, M.; Ford, G.; Frodsham, L.; Golden, M.; Leslie, K.; Lyons, S.; McKeever-Abbas, B.; Stark, A.; Tomlin, P.; Gottschling, S.; Hajar, A.; Jiang, J.-I.; Lo, J.; Suchozak, B. *Org. Process Res. Dev.* **2012**, *16*, 741−747. (b) Juhl, M.; Tanner, D. *Chem. Soc. Rev.* **2009**, *38*, 2983−92. (c) Funel, J.-A.; Abele, S. *Angew. Chem., Int. Ed.* **2013**, *52*, 3822−3863.
(33) Williams, D. H.; Faulkner, D. J. *Tetrahedron* **1996**, *52*, 4245−4256.
(34) Levin, J. I.; Laakso, L. M. Oxazole Diels−Alder reactions. In *Oxazoles: Synthesis, Reactions, and Spectroscopy, Part A*; Palmer, D. C., Ed; John Wiley & Sons, Inc.: Hoboken, NJ, 2003; Vol. *60*.
(35) Ang, K. H.; Prager, R. H.; Smith, J. A.; Weber, B.; Williams, C. M. *Tetrahedron Lett.* **1996**, *37*, 675−678.
(36) An analogous dihydrofuropyrrolone undergoes ready Diels-Alder cycloaddition with *N*-phenylmaleimide across the furan substructure to give a hexasubstituted benzene; see Janvier, P.; Bienayme, H.; Zhu, J. *Angew. Chem., Int. Ed.* **2002**, *41*, 4291−4294.
(37) (a) Jaworski, T.; Mizerski, T. *Polish J. Chem.* **1981**, *55*, 317−320. (b) Hari, Y.; Iguchi, T.; Aoyama, T. *Synthesis* **2004**, *9*, 1359−1362.
(38) Ibragimov, I. I.; Mamedov, E. I.; Ismailov, A. T.; Aliev, A. G.; Mextieva, S. Z.; Dzhafarov, V. G.; Belyaeva, V. I. *Zh. Org. Khim.* **1990**, *26*, 1648−1654.
(39) Kobayashi, K.; Tanaka, H.; Tanaka, K.; Yoneda, K.; Morikawa, O.; Konishi, H. *Synth. Commun.* **2000**, *30*, 4277−4291.
(40) Isomeric 4-methyl-5-cyanooxazole cycloaddition to propiolate ester gives two products, one of which is **41**. See ref 37a.
(41) (a) Kiener, A.; Roduit, J.-P.; Tschech, A.; Tinschert, A.; Heinzmann, K. *Synlett.* **1994**, *10*, 814−16. (b) Cho, J.; Lee, S.; Rhee, J.; Park, S. J.; Park, O. (LG Chemical Ltd.) Novel microorganism of Alcaligenes capable of catalyzing regioselective hydroxylation of nitrogen containing heterocyclic compound, useful for producing compounds such as 6-hydroxynicotinic acid by hydroxylation of nicotinic acid. WO 2007011156A1, 25 Jan, 2007.
(42) (a) Hosogai, T.; Nishida, T.; Stoi, K.; Takagi, T. (Kuraray Co., Ltd., Japan; Kyowa Gas Chemical Industry Co., Ltd.) 2-Substituted 5,6-dicyano-3-hydroxypyrazines. Jpn. Kokai Tokkyo Koho, JP 50059379 A 19750522, 1975; *Chem. Abst.*, **1975**, *83*, 193380a. (b) See also Nakamura, A.; Ataka, T.; Segawa, H.; Takeuchi, Y.; Takematsu, T. *Agric. Biol. Chem.* **1983**, *47*, 1555−1560.
(43) Okada, S.; Kosasayama, A.; Konno, T.; Uchimaru, F. *Chem. Pharm. Bull.* **1971**, *19*, 1344−1357.
(44) Sato, N.; Arai, S. *J. Heterocycl. Chem.* **1982**, *19*, 407−408.
(45) See ref 20, p 218.
(46) A comprehensive review of diazonium salts as substrates in palladium-catalyzed cross-coupling reactions mentions no example of C−O coupling with phenols: Roglans, A.; Pla-Quintana, A.; Moreno-Mañas, M. *Chem. Rev.* **2006**, *106*, 4622−4643.
(47) See ref 20, p 84.
(48) (a) Vompe, A. F.; Turitsyna, N. F.; Levkoev, I. I. *Dokl. Akad. Nauk SSSR* **1949**, *65*, 839−842. *Chem. Abstr.* **1949**, *43*, 6626h. (b) Vompe, A. F.; Turitsyna, N. F. *Zh. Obshch. Khim.* **1957**, *27*, 3282−3290. *Chem. Abstr.* **1958**, *52*, 9112d. (c) Makosza, M.; Fedorynski, M. *Russ. Chem. Bull. Int. Ed.* **2011**, *6*, 2141−2146.
(49) Busby, R. E.; Khan, M. R.; Parrick, J.; Shaw, C. J. G.; Iqbal, M. *J. Chem. Soc., Perkin Trans. 1* **1980**, 1427−1430.
(50) Grimmett, M. R. *Imidazole and Benzimidazole Synthesis*; Academic Press: San Diego, CA, 1997.
(51) On page 29 of ref 7, Judson outlines comparison of LHASA output with chemist suggestions for routes to synthesis targets. The conclusion was that both performed at about the same level, with the exception that LHASA came up with better rearrangement-based synthesis suggestions. However, as far as we are aware, none of these comparisons has been published.
(52) (a) Parker, J. S.; Moseley, J. D. *Org. Process Res. Dev.* **2008**, *12*, 1041−1043. (b) Moseley, J. D.; Brown, D.; Firkin, C. R.; Jenkin, S. L.; Patel, B.; Snape, E. W. *Org. Process Res. Dev.* **2008**, *12*, 1044−1059. (c) Parker, J. S.; Bower, J. F.; Murray, P. M.; Patel, B.; Talavera, P. *Org. Process Res. Dev.* **2008**, *12*, 1060−1077.
(53) (a) Unified suite of metrics: Andraos, J. *Org. Process Res. Dev.* **2013**, *17*, 175−192 and earlier papers in the series. (b) Andraos, J. *The Algebra of Organic Synthesis: Green Metrics, Design Strategy, Route Selection, and Optimization*; CRC Press: Boca Raton, FL, 2012. (c) Lapkin, A., Constable, D. C., Eds. *Green Chemistry Metrics: Measuring and Monitoring Sustainable Processes*; Blackwell Scientific: Oxford, 2008. (d) The E-factor: Sheldon, R. A. *Green Chem.* **2007**, *9*, 1273−1283. (e) The eight criteria of Boehringer Ingelheim: Dach, R.; Song, J. J.; Roschangar, F.; Samstag, W.; Senanayake, C. H. *Org. Process Res. Dev.* **2012**, *16*, 1697−1706. (f) Process mass intensity: Jiménez-González, C.; Ollech, C.; Pyrz, W.; Hughes, D.; Broxterman, Q. B.; Bhathela, N. *Org. Process Res. Dev.* **2013**, *17*, 239−246.
(54) The question of noise in the results and user acceptance is complex. The intuitive reaction is that questionable or chemically inappropriate suggestions are unwanted and should not appear. This is the default development aim, especially for plainly wrong suggestions that evade currently implemented filters. However, some users employ such results to set their imagination running and are happy to see them.
(55) Huerta, F.; Hutchings, M. G.; Saller, H.; Löw, P. *Knowledge-Based de Novo Molecular Design Using ICSYNTH FRP*, ICIC, Heidelberg, Germany, 17th October, 2014.

DOI: 10.1021/op500373e
Org. Process Res. Dev. 2015, 19, 357−368