

In Silico Prediction of Pharmaceutical Degradation Pathways: A Benchmarking Study Using the Software Program Zeneth

Rachel Hemingway,* Steven W. Baertschi, David Benstead, John M. Campbell, Maggie Coombs, Zongyun Huang, Ryan Khalaf, Martin A. Ott, David J. Ponting, Darren L. Reid, Neil G. Stevenson, Samuel J. Webb, Angela White, and Todd Zelesky



Cite This: *Org. Process Res. Dev.* 2024, 28, 674–692



Read Online

ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Zeneth, a software application for the prediction of chemical degradation of small organic molecules, incorporates a knowledge base of rules to predict degradation pathways. In addition, the knowledge base contains property predictors that modulate the predicted likelihood of a given degradation product. In this study, a C–H bond dissociation energy (C–H BDE) predictor, which has been integrated into the software, was utilized. To determine this software's predictive capabilities [using its knowledge base (2020.1.0 KB)], experimentally derived degradation profiles for 25 drug substances were compared to Zeneth predictions. These degradation profiles were derived from forced degradation studies, including accelerated and long-term stability studies, aligned with International Council for Harmonisation (ICH) guidelines. In addition, two case studies highlighting how prediction data can be utilized to confirm experimental data or assist with the identification of unknown degradation products have been presented. The specificity of prediction results was evaluated; transformation types that often predict degradation products not observed experimentally were identified, and an assessment of the causes is presented. The sensitivity for the study group was also evaluated using a historic knowledge base (2012.2.0 KB), enabling an assessment of how the predictive capabilities have improved over this period; the comparison demonstrated a 40% increase in sensitivity. This study has demonstrated that the ongoing expansion and optimization of this in silico tools knowledge base continues to result in improvements in its predictive capability and its ability to impart insight into the drug degradation knowledge space to aid pharmaceutical development.

KEYWORDS: *chemical stability, forced degradation, in silico prediction, degradation products*

INTRODUCTION

Determining the Stability of Drugs under Development. Understanding the chemical stability risks associated with a drug substance (DS) throughout pharmaceutical development is key to ensuring drug product (DP) quality and safety. Forced degradation studies (also known as stress testing) are conducted to understand the intrinsic stability of a DS. Well-designed forced degradation studies can be used to support method development, validate stability-indicating methods, identify potential degradation pathways, and troubleshoot stability issues. Degradation information collected on the DS and DP provides a comprehensive degradation profile, which can inform formulation development and packaging needs (Figure 1). This can impact the selection of storage conditions and shelf life for the final DP.

Requirements for establishing potential degradation products (including consideration of potential intermediates in the degradation pathway) of a DS through forced degradation studies are outlined in several guidelines.^{1–5} Often these guidelines do not detail exactly how these studies should be conducted or the specifics of stress conditions that should be employed, making only general recommendations. Determining the set of stress conditions and experimental execution specifics needed to induce the potential degradation pathways most relevant to pharmaceutical storage, distribution, and use requires consideration of several factors, not least, the

physicochemical properties of the DS under study. Making the design of these studies and interpretation of the results a complex activity.^{6,7} Stress testing is often first initiated in Discovery (i.e., preclinical) and is typically repeated at least once during the clinical phase of development, in addition to excipient compatibility, and long-term DS and DP stability studies.^{8,9}

Due to the complexity of organic reaction chemistry and the diverse chemical environments pharmaceuticals experience, identifying the degradation mechanism(s), root causes, and control strategies is challenging. Knowledge of chemical pathways usually belongs in the domain of the practicing synthetic or organic chemist and may be obtained from chemistry textbooks, but it is not trivial to apply this knowledge to scientific problems in forced degradation and, in particular, structure determination of degradation products or pathway elucidation (Figure 1). Given the diverse backgrounds of typical pharmaceutical scientists, it is

Received: September 29, 2023

Revised: December 27, 2023

Accepted: January 22, 2024

Published: February 19, 2024



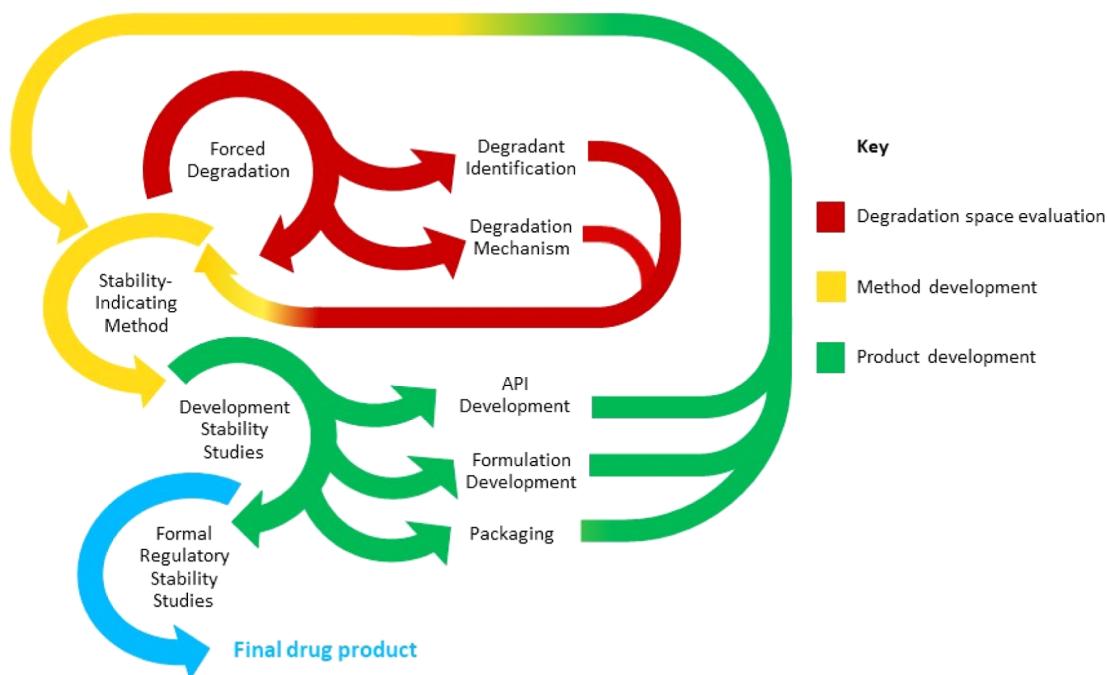


Figure 1. Generalized workflow outlining steps to develop a final drug product.

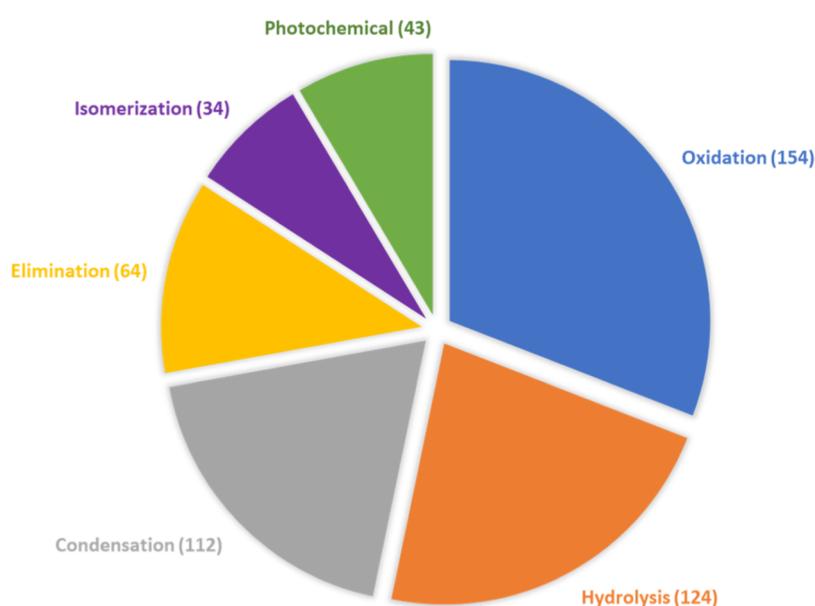


Figure 2. Coverage by reaction category of the knowledge base (2020.1.0 KB).

important that they have the proper tools to identify and work through stability related problems. In silico systems are one such tool that can be utilized early in the drug development workflow to identify theoretical reactivity risks and potential mechanisms by which degradation products can form. This information can then be used as a starting point to interpret the stress testing information generated experimentally.

Computer-Aided Prediction Tools. Computer-aided programs aimed at chemical degradation prediction represent a means of capturing and utilizing this vast, complex information to serve not only as an educational tool but also as a repository for historical information to support and supplement knowledge. Historical examples of computer programs aimed at chemical degradation prediction can be

found in the literature^{10–12} but all have been discontinued for various reasons. To the authors' knowledge, there are three active software programs focused on predicting chemical degradation pathways. The Environmental Protection Agency have produced two of these tools: one predicts hydrolysis products of organic chemicals¹³ and the other is a reaction library for predicting direct photochemical products of environmental organic contaminants.¹⁴ The third tool is Lhasa Limited's Zeneth.¹⁵

While not predicting chemical degradation directly, there are computational machine-learning models and synthesis prediction systems that can aid the field of degradation. Dana et al. have used kinetic models to predict likely oxidation pathways for DSs,^{16,17} while Coley et al. have developed a supervised

learning approach to predict the products of organic reactions.¹⁸ In addition, a retrosynthetic tool using deep neural networks and symbolic artificial intelligence has been described by Segler et al.¹⁹ For more comprehensive overviews on the use of artificial intelligence and machine-learning methods to develop computational reaction modeling methodologies, the reader is directed elsewhere.^{20,21}

In silico Prediction System Zeneth. Zeneth is a commercially available in silico software program for the prediction of chemical degradation pathways of organic DSs and DPs. The software couples a chemical transformation engine with a knowledge base developed mainly from public sources but also from (anonymized) proprietary chemical degradation data to generate predictions of degradation products based on functional group matching.^{22,23} If a functional group within the structure of the DS or DP matches a degradation reaction outlined in the knowledge base, a degradation product will be predicted, and that prediction will be associated with a likelihood score.

In this regard, the in silico assessment can provide knowledge and mechanistic understanding of theoretically possible chemical degradation pathways under defined reaction conditions (temperature, pH, water, oxygen, peroxide, light, metal ions, and radical initiator); these are typically used in forced degradation studies. Excipients (and their known impurities) can also be evaluated against a DS to assess DS-excipient reactivity risk using a built-in excipient database.

A detailed description of how the software works has been outlined in previous publications;^{21,22} the following sections aim to provide an overview of the key aspects of the software relevant to this study.

Knowledge Gathering. The contents of the knowledge base comes from several sources including books focused on chemical degradation,^{24–26} general chemistry textbooks,²⁷ and primary literature sources, as well as proprietary data through ongoing data-sharing initiatives with Zeneth users. The knowledge base is continuously refined and expanded as new knowledge is garnered from these sources; 2020.1.0 KB used in this study contains 531 transformation types and covers six degradation reaction categories (Figure 2). It should be noted that a given transformation will cover a “type” of chemical reaction, not a single reaction. For example, a substitution reaction may cover several different leaving groups, or an oxidation may cover various related functional groups as substrates.

Manually building a knowledge base necessarily starts with the most common reaction types, and as time progresses, increasingly less common and more idiosyncratic reaction types are added. The degradation literature regularly provides examples of reaction types that are not covered, and an important impetus and focus for adding new knowledge comes from studies such as the present one. Eventually, the knowledge base will approach “completion” but it will only do so asymptotically.

Likelihood Scores. Zeneth does not perform calculations such as reaction kinetics or energies of transition states nor does it take concentration or reaction time into account. Proper kinetics data and/or adequate data on the substrate dependence of reactivity are often hard to come by. On the other hand, while reactivity data can be found in the synthesis literature, likelihood scoring in Zeneth aims to be relevant to (forced) degradation studies. These studies often take place at room temperature or somewhat higher (rarely above 80 °C)

with time scales from minutes to weeks; moreover, they generally aim at 5–20% degradation of the DS,⁸ whereas synthesis studies aim at high yields, making it difficult to compare reactivities. All of the above make scoring predicted degradation reactions a challenge.

The likelihood score of a degradation product aims to estimate the probability of that degradation reaction occurring. Likelihood scores are influenced by the presence or absence of a given reaction condition such as pH or metal ions, as well as C–H BDE (bond dissociation energy) calculations. Likelihood scores range from 0 to 1000, where 0 is impossible, 1000 is certain, and any score above 599 is regarded as likely. While previous versions of the software categorized the calculated likelihood scores into one of seven levels, the new scoring system is on a continuous scale to give more granularity to prediction results (cf. Figure 3).

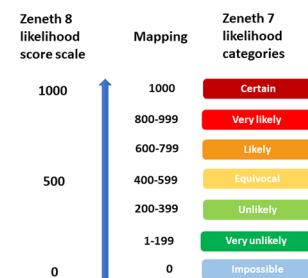


Figure 3. Boundaries of the likelihood score scale used in Zeneth 8 (0–1000) and their correspondence to categories used in Zeneth 7.

The approach taken in Zeneth is that an experimentally observed degradation product and/or types of degradation commonly observed in the literature should be predicted with a score of at least 500 (equivocal). For instance, ester hydrolyses are commonly predicted with a score of 700 (likely) around pH 7 and 900 (very likely) at any other pH. However, substrate dependency (e.g., the effect of steric hindrance) is not currently considered because this would require large amounts of data (to which we do not have access) and/or resource-intensive computational assessments. Assigning a score to a degradation product is based on the quality and quantity of literature available (degradation, physicochemical, and synthesis) and the expertise of scientists at Lhasa Limited and users participating in data sharing initiatives. Currently, scores such as 500, 700, and 900 have a high incidence in predictions as they are based on non-numerical likelihood assignments of “equivocal”, “likely”, and “very likely”, respectively, in all prior versions of the software.^{21,22} For reaction types involving H atom abstraction as the initiating step, predicted BDEs are used (*vide infra*) to modulate likelihood scores. It is expected that in the future, more predictors of physicochemical properties will be used if their applications can be validated with experimental data. Such developments could increase the accuracy and granularity of likelihood scoring and will also lower the incidence of “round” scores, e.g., 500 and 700.

Study Objectives. In 2014, a benchmarking study was published in collaboration with several pharmaceutical companies.²⁸ The aim of that study was to establish how well the knowledge base within Zeneth could predict a data set of experimentally observed degradation products, i.e., the sensitivity, as a direct result of software improvements and updates to the knowledge base. An increase in sensitivity from

31 to 54% over a 3 year period (2009–2012) based on a study pool of 191 experimentally observed degradation products from 27 proprietary DSs was determined. It should be noted that a few (<10) degradation products arising from DS-excipient interactions were included in the 2014 study, but such interactions have not been included in this study. Continued updates to the knowledge base (since 2014) coupled with the release of an upgraded Zeneth application (in 2020) provides an opportunity to evaluate aspects of the software again through a new benchmarking study.

The present study is also primarily concerned with sensitivity with the aim of evaluating the knowledge base (2020.1.0 KB) against experimental data: degradation products observed from forced degradation studies, accelerated, and long-term stability studies. For clarity, in this study, sensitivity is defined as the number of experimentally observed degradation products correctly predicted or, to put it another way, how well can the test (Zeneth) identify a true positive (an experimentally observed degradation product). This metric can help assess how the program has progressed in the field of predictive degradation chemistry and whether the coverage of the knowledge base aligns to pharmaceutically relevant degradation pathways, i.e., real-world degradation chemistry. In addition, a statistical analysis will be performed using an older knowledge base (2012.2.0 KB) to determine whether there has been an improvement in the sensitivity over time, and a bootstrapping analysis will be used to determine if any improvement is “real” and not the result of random chance. This study also aims to demonstrate how *in silico* predictions can be used to confirm experimental data or assist with identification of unknown degradation products using non-proprietary experimentally derived forced degradation profiles, in the form of two case studies. Theoretical prediction of degradation pathways will naturally lead to an overprediction of possibilities. Hence, a further aim of this study is to assess the specificity of predictions to try and understand causation, with a longer-term objective of decreasing overprediction (the number of predicted degradation products that are not observed experimentally) to improve the accuracy of predictions. It should be noted that from a pharmaceutical development perspective, maximizing sensitivity is in general a greater priority than minimizing overprediction. This represents a more conservative position and helps minimize the chance of an unexpected degradation pathway not being identified during a compound’s development lifecycle. Thus, failure to predict an actual degradation pathway is more of a concern than predicting pathways that do not occur.

■ EXPERIMENTAL SECTION

Study Pool. Five pharmaceutical companies each provided between 3 and 8 small-molecule, organic DSs to the benchmarking study, totaling 25 DSs with 128 experimentally observed degradation products. It should be noted that all these DSs were naïve to the knowledge base. It should also be noted that the data from the 2014 benchmarking study was used to develop new transformation types for the knowledge base which is why a comparison using this data (i.e., DSs and associated degradation products) has not been made. The data for the 128 experimentally observed degradants were derived from forced degradation studies and included accelerated and long-term stability studies, aligned with International Council for Harmonisation (ICH) guidelines.^{1–4} In most cases, structures of the experimentally observed degradation products

were assigned using liquid chromatography–mass spectrometry (LC–MS) techniques without isolation from the stressed samples. In such cases, the chemical stress conditions applied could be used to infer degradation pathways and increase the confidence of the structural assignment. Where LC–MS data and chemical inference alone did not permit confident structural assignment of the degradation products, degradation products were isolated, and nuclear magnetic resonance (NMR) and/or other spectroscopic techniques were utilized. Decisions regarding which degradation products to characterize, as well as the thoroughness of the characterization, were made based on scientific judgment combined with each individual company’s internal practice. In rare cases, authentic markers prepared synthetically or purchased commercially were available and used to confirm degradation product structures by comparing the high-performance liquid chromatography (HPLC) retention times and other spectroscopic data of the markers and degradation products. All structural assignments were made prior to the initiation of this benchmarking study. Of the 25 DSs, 22 are proprietary and so their structures will not be disclosed; three are publicly available, and for two of these, their structures and degradation products (both experimentally observed and predicted) will be discussed in detail later. This data set was selected over other possible sources due to the availability of degradation data under relevant stability conditions and its relevance to current pharmaceutical development pipelines which continue to increase in complexity.²⁹ To demonstrate the diversity of chemistry this data pool covered, Table 1 outlines the cumulative count of functional groups and ring systems that were seen across the 25 DSs, more than a quarter of which were not covered in the 2014 benchmarking study data set.²⁸

Moreover, a comparison of their physicochemical properties to an analysis of the physicochemical properties of small molecules contained in the drug repurposing hub, recently conducted by Agarwal et al., shows their relevance to a drug-like space.³⁰ For example, Agarwal et al. noted that the molecular weight, polar surface area, hydrogen bond acceptors and number of rotatable bonds have increased for drugs approved in the past decade (2011–2020).³⁰ Similar observations were seen for the 25 DSs used in this study versus the data set from the 2014 benchmark study (Figure 4). Although average cLogP was slightly lower, and hydrogen bond donors were higher than the 2014 data set, the averages were also consistent with approved drugs in the past decade.

All descriptors except cLogP were calculated in KNIME³¹ (version 4.4)³² using RDKit³³ nodes. cLogP was calculated using the tool provided by BioByte Corp. (version 5.9), as made available by Lhasa Limited in Derek Nexus³⁴ (version 6.2.1).

Software Parameters. A query compound needs to be inputted, i.e., the structure of the DS under study. There are then a series of conditions that can be controlled by the user to reflect forced degradation conditions the user wishes to simulate, as well as parameters to control the number and type of degradation products predicted.²³

Parameters for Conditions and Processing Constraints. Table 2 outlines the conditions and processing constraints selected for this study. These conditions and constraints were chosen to align (as closely as possible) with the parameters used in the 2014 study to allow for fair comparison. The parameters described aim to mimic forced degradation conditions while allowing a sufficient assessment

Table 1. Cumulative Count of the Functional Groups and Ring Systems for 25 DSs. Note: Specific Aromatic Rings Have Been Counted as Isolated Rings (e.g., Benzene and Pyrrole) and Separated from Fused Aromatic Ring Systems (e.g., Indole, Naphthalene, and “Other”)

functional group	count	functional group	count
3,8-diazabicyclo[3.2.1]octane	1	fluoride—aliphatic	1
acrylamide	1	fluoride—aromatic	10
alcohol—aliphatic, primary	2	indole	4
alcohol—aliphatic, secondary	5	ketone—aliphatic-aromatic	1
alcohol—aliphatic, tertiary	1	morpholine	1
alcohol—phenol	2	naphthalene	1
alkene—1,2-disubstituted	1	nitrile	2
alkene—monosubstituted	1	other acids	1
amide—primary	1	other fused aromatic rings	5
amide—secondary	7	oxadiazole	1
amide—tertiary	3	piperazine	7
amine—aliphatic, primary	2	piperidine	5
amine—aliphatic, secondary	2	pyrazole	4
amine—aliphatic, tertiary	4	pyridine	9
amine—aromatic, secondary	4	pyrimidinone	1
amine—aromatic, tertiary	2	pyridinone	1
benzene	6	pyrimidine	3
benzimidazole	1	pyrrolidine	1
carboxylic acid	3	quinazoline	1
chloride—aromatic	8	sulfanone—imino	1
cyclobutane	2	sulfonamide	6
cyclohexane and larger	7	sulfone	2
cyclopentane	2	thioether—aliphatic—aliphatic	1
cyclopentenone	1	thioether—aliphatic-aromatic	2
cyclopropane	5	triazoles	2
ester	2	trifluoromethyl	2
ether—aliphatic—aliphatic	4	trifluoromethylsulfonyl benzene	1
ether—aliphatic-aromatic	9	urea	3

of the breadth and depth of the knowledge base. The parameters are “typical” for the prediction system, and many are defaults, ergo representative of what users of the software would also be utilizing. It should be noted that the 80 °C

Table 2. In silico Conditions and Processing Constraints Selected for This Study

	Parameter	Settings
conditions	temperature	80 °C
	pH	1 and 13
	water	present
	oxygen	present
	metal (i.e., Fe ³⁺ or Cu ²⁺)	present
	radical initiator	present
	peroxide	present
	light	present
processing constraints	maximum number of degradation products	400
	pathway length	2 steps (4 steps for additional confirmation if required)
	unimolecular and bimolecular reactions ^a	Q; Q + A; Q + Q; Q + D; D; D + A; D + D
	minimum score	400
	pathway calculation method	lowest step likelihood and multiplied step likelihood

^aQ = primary query structure, i.e., drug substance; D = degradation product(s); A = secondary query structure, i.e., excipient.

setting was chosen to invoke the thermally driven transformation types within the knowledge base and provide a worst-case prediction from a likelihood score perspective.

Within the software, condition parameters are indicated as either present or absent, except for temperature and pH which are set numerical values within defined boundaries: 0–14 for pH and 0–100 °C for temperature. Note that, at the point of processing, all parameters specified are run as a “set” of conditions. Different sets of condition parameters require another run of the program. Two predictions were performed for each compound to determine the degradation products arising due to acidic, pH = 1, and basic, pH = 13 conditions. The processing constraints control all other parameters of the prediction: the number of degradation products that can be produced for a given prediction, the number of reaction steps (length of degradation pathway), the minimum likelihood score for a given degradation product, the reaction modes, i.e., which unimolecular or bimolecular reactions are allowed, and the calculation method for determining the likelihood score. Conditions and processing constraints were kept constant

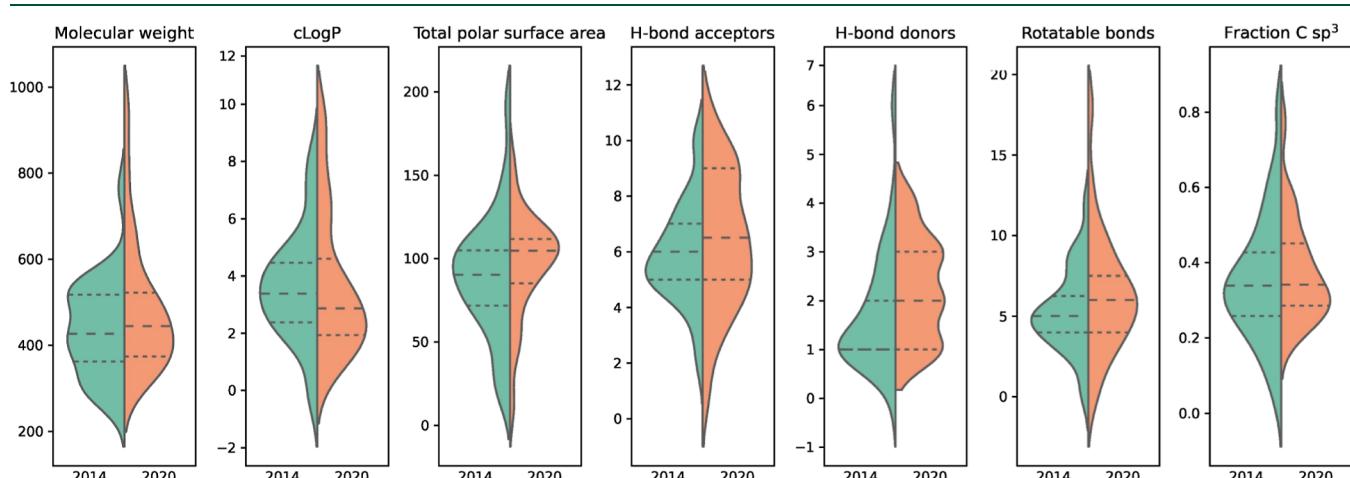


Figure 4. Violin plot showing the distribution of the physicochemical properties of compounds from the 2014 benchmarking study²⁶ (green) and the 25 DSs in this study (orange). The dashed lines are the mean values, and the dotted lines are the upper and lower quartiles.

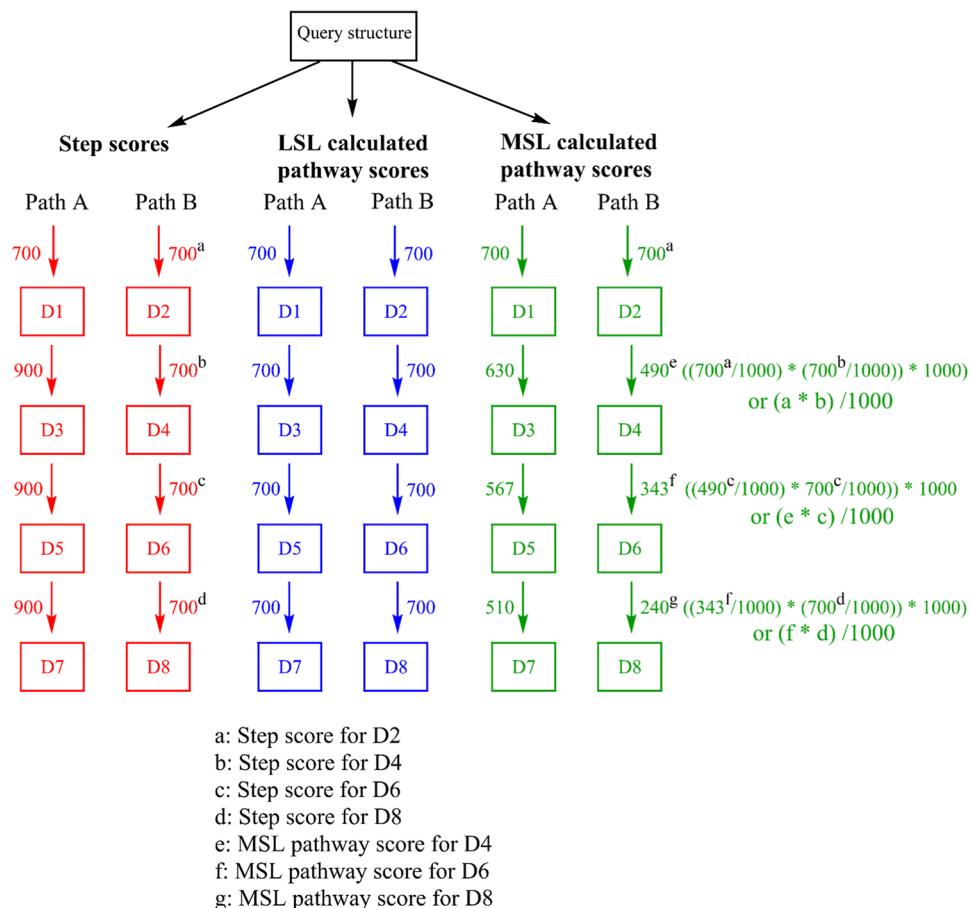


Figure 5. How the LSL and MSL calculation methods are used to generate likelihood scores for a given pathway. The LSL method yields a pathway score equal to the lowest of the individual step scores, whereas the MSL method turns step scores into probabilities and multiplies them together.

across all predictions to ensure consistent analysis across data sets, as well as for optimizing sensitivity.

The number of reaction steps controls the length of the degradation pathway. The program finds all applicable transformation types within its knowledge base and generates a first-generation of degradation products. Generation refers to the step (or level) at which the degradant is from the query compound. Each predicted first-generation degradation product was then processed in the same way as the query compound until the number of steps defined had been reached. Note that each transformation type in the knowledge base has its own condition requirement. Thus, when in step 1, a degradation product is formed from a single transformation type (because of a particular condition), that same degradation product may react in step 2 as if directly exposed to any of the conditions.

Degradation pathways were limited to 400 structures and two steps. If more than 400 degradation products were predicted, the application stopped further prediction (truncation of the degradation tree). Prediction results are presented in a tree-shaped diagram showing degradation pathways branching out from the query; often referred to as a “tree” of results. Truncation is used as a practical limitation, based on the user’s ability to interrogate the output. However, the effect of these study constraints was evaluated by allowing the maximum number of degradation products to be 1000 and the number of steps in the tree to be three. These “wider” constraints meant Zeneth was able to predict 2 additional

experimentally derived degradation products. These two products, while acknowledged, have not been used in any analysis.

Reaction modes allow the query compound (Q) to react with its environment, with the same being true for a given degradation product (D). They also control the inclusion or exclusion of dimerization reactions ($Q + Q$; $D + D$) as well as the possibility of the query compound reacting with degradation products or excipients (or their impurities) (A) ($Q + D$; $Q + A$) and indeed degradation products reacting with excipients (or their impurities) ($D + A$). It should be noted that excipients were not considered for this study, and therefore, the ‘ $Q + A$ ’ and ‘ $D + A$ ’ options did not yield additional predictions.

A minimum score threshold of 400 was set; this equates to a threshold level of “equivocal” in the previous Zeneth platform (Zeneth 7) (cf. Figure 3). This was the minimum threshold used in the 2014 study,²⁷ enabling a fair comparison. There are two methods available for calculating the pathway score; see the next section. Predictions at pH 1 and 13 were run using both methods for each DS.

Pathway Likelihood Calculations. The likelihood score of a degradation pathway (“pathway score”) is calculated from the score of the individual steps (“step scores”) leading to the degradation product. A step score is determined via an evaluation of an individual transformation type and depends only on the specified conditions and structural features of the query compound that affect the chemical reactivity. Two

calculation methods are available to determine a pathway score from step scores. The simplest is the Lowest Step Likelihood (LSL), which is the default method used in the software, where the pathway score is set to the score of the least likely step in the pathway. The other is Multiplied Step Likelihood (MSL), where the step scores are interpreted as probabilities (i.e., divided by 1000), and the pathway score is calculated as the product of these probabilities, which is then multiplied by 1000 (Figure 5).

The LSL method is straightforward but does not provide much granularity with longer pathways. For example, a pathway (B) of four steps, each having a score of 700, yields a pathway score of 700; while a different pathway (A) of four steps, the first having a score of 700 and the next ones a score of 900 (very likely), also yields a pathway score of 700. However, pathway B, a concatenation of likely steps, seems less likely than pathway A. The MSL method aims to overcome this limitation. Using this method, pathway B yields a pathway score of 240 (well below any threshold of interest), while pathway A yields a pathway score of 510 (Figure 5). In practice, predictions that allow longer pathways (e.g., four steps) tend to generate large trees that can be unwieldy and can contain many doubtful pathways when the LSL method is applied. Using the MSL method, the scores of longer pathways soon end up below the default threshold (400), retaining pathways that contain cascades of very likely steps. The difference between the methods is small with “shallow” (one or two steps) trees, but the MSL method is deemed especially useful with “deeper” trees (\geq three steps). This principle has not yet been validated, but data from the present study will provide an indication of this theory.

Knowledge Base Versions. The knowledge bases used in this study were the 2020.1.0 and the 2012.2.0 KB. The 2020.1.0 KB contains 531 transformation types; see Figure 2 for a breakdown of the degradation chemistry contained within it. This knowledge base is used within the Zeneth 8 platform. The 2012.2.0 KB contains 277 transformation types and was used within a previous software platform (Zeneth 7).

Bootstrapping Analysis. The difference in sensitivity between the 2020.1.0 and the 2012.2.0 KB was assessed using a bootstrapping methodology to establish whether this difference was “true” or the result of random chance. Bootstrapping, also known as random sampling with replacement, takes a data set of size ‘n’ and generates ‘m’ new training sets, each of size ‘n’, by sampling uniformly and with replacement. This means all samples are the same size, but the same data point can be used multiple times within the same sample, meaning some are not used at all. Sampling with replacement ensures each “bootstrap” is independent from its peers, as it is independent of previous samples when sampling, i.e., samples are randomly generated. All experimentally observed degradation products were classified (using a binary classification) as being predicted (classification 1) or not (classification 0) by the two different knowledge bases. KNIME³¹ (version 4.3)³² was used to perform the analysis. 250 “bootstraps” were generated to obtain distribution graphs of the sensitivity levels. The difference in sensitivity between the 2020.1.0 and the 2012.2.0 KB for each bootstrap was also plotted to visualize the difference in performance; see Supporting Information.

Overprediction Analysis. Overprediction levels were assessed for the 25 DSs using prediction results from both LSL and MSL calculation methods. Structure comparisons

(between a prediction result and experimental degradation products as well as between different prediction results) were performed using InChI³³ strings. An InChI string is a standardized way of representing a chemical structure as a text string. However, experimental data may include individual stereoisomeric products, but when there are several possible stereochemical outcomes, Zeneth will only generate one product structure with undefined stereocenters. Therefore, to enable effective structure comparisons and identification of duplicates, InChI strings were generated by using the SNon option throughout. This option causes stereo layer(s) to be omitted from the InChI string, meaning that stereochemical information is completely ignored.³³

The parent compounds were processed through Zeneth with a pH condition value of 1 and 13 generating a pair of result files encoded in structure data format (SDF).³⁴ Using KNIME³¹ (version 4.5),³² for each parent compound in the study, both SDFs were read in, where appropriate transformation sequences were generated, i.e., for second-generation degradation products, and duplicates (identified using InChI strings) were removed. For the reaction mode analysis (cf. Table 5), the analysis is based upon identification of the degradation product regardless of the transformation pathway. If a degradation product is generated by ‘n’ transformation pathways, it will be counted ‘n’ times in this study; however, this duplication check is agnostic about the conditions, so a degradation product generated at pH 1 and 13 by the same transformation or transformation pathway is only counted once. Experimentally observed degradation products, correctly predicted, were filtered out by removing rows from the prediction tables, where the InChI string and parent ID are found in the list of experimentally observed degradation products.

The above process was run using the LSL and MSL methods. For each method, each overpredicted degradation product (predicted but not observed experimentally) was assigned a unique identifier of the form {identifier} – {inchi} – {sequence}. Each of these degradation products was then assigned as “common”, “lowest step only”, or “multiplied step only” based on the following:

1. Common: unique identifier found in both sets
2. Lowest step only: unique identifier only found in the results from the LSL SDF
3. Multiplied step only: unique identifier only found in the results from the MSL SDF

■ RESULTS AND DISCUSSION

Sensitivity Analysis. Sensitivity is a measure of how well a test can identify a true positive. In the case of this study, the sensitivity is a measure of whether Zeneth can predict a degradation product that was observed experimentally. Predictions for the 25 DSs under study, using the LSL calculation method, were obtained using the conditions and processing constraints outlined in the Experimental Section. The structures of the experimentally observed degradation products were then cross-referenced in Zeneth (using an exact-mass search coupled with a manual review of the structure) to determine whether they were predicted by 2020.1.0 KB (Table 3 and Figure 6). Of the 128 experimentally observed degradation products, 90 were correctly predicted, equating to a sensitivity level of 70%. For eight of the DSs, all of their experimental degradation products were correctly predicted i.e.

Table 3. Number of Experimentally Observed Degradation Products Correctly Predicted Using 2020.1.0 KB or 2012.2.0 KB, Respectively. Note: For Compound IDs 15 and 19 and for the Total Count, the Number in Parentheses Represents the Number of Degradation Products That Would be Correctly Predicted if the Number of Steps in the Pathway was Increased to Three and the Tree was Not Truncated at 400

compound ID	experimentally observed degradation products	predicted degradation products from 2020.1.0 KB	predicted degradation products from 2012.2.0 KB
1	8	4	4
2	10	5	3
3	7	6	3
4	3	3	1
5	2	2	1
6	6	5	5
7	5	4	3
8	3	3	3
9	3	3	2
10	4	2	1
11	3	1	1
12	6	5	3
13	6	4	4
14	6	2	2
15	4	3 (4)	2 (3)
16	5	4	2
17	3	3	3
18	7	5	4
19	11	6 (7)	4 (5)
20	4	4	4
21	2	2	2
22	3	2	2
23	8	7	2
24	3	3	1
25	6	2	2
total	128	90 (92)	64 (66)
sensitivity		70% (72%)	50% (52%)

100% coverage. Two further degradation products were predicted when the conditions and processing constraints were widened, one being observed in a three-step pathway and another when the tree was not truncated at 400 degradation products. Inclusion of these two degradation products would increase the sensitivity level to 72% (cf. Table 3).

A comparison was performed using the knowledge base from the 2014 benchmarking study (2012.2.0 KB), using the same conditions and processing constraints (to allow for a fair comparison). The 25 DSs in this study were processed through 2012.2.0 KB to see which experimentally observed degradation products would have been correctly predicted at this point in the development history of the knowledge base. Of the 128 experimentally observed degradation products, 64 were correctly predicted by 2012.2.0 KB, equating to a sensitivity level of 50%. This would increase to 52% if the two degradation products seen outside of the defined study parameters were included.

Therefore, for this data pool, there has been a 40% increase in sensitivity going from 2012.2.0 KB (50% sensitivity value) to 2020.1.0 KB (70% sensitivity value). This demonstrates that development of the knowledge base from 2012 to present day has increased its coverage of degradation chemistry in a

realistic manner and that Zeneth was able to provide valid predictions. A bootstrap analysis was performed to determine if the percentage increase in sensitivity was a “true” difference or attributable to noise or random chance. The analysis demonstrated little overlap between the distributions of each data set and that the difference in sensitivity is in fact “real”, see Supporting Information.

Score Distribution. Assessment of the results using the LSL method shows that the majority of experimentally observed degradation products were predicted as first-generation; 76 out of 90, the other 14 were predicted as second-generation (Figure 7). Of the 90 experimentally observed degradation products correctly predicted, 84 of them had a score ≥ 600 (cf. Figure 3). This indicates a high level of confidence in the chemistry underpinning these predictions.

Experimentally observed degradation products were not predicted for two main reasons. Either the chemistry was simply unknown so new transformation types would need to be added to the knowledge base or the chemistry was known, but a structural aspect of that DS was not within the scope of the relevant transformation type, resulting in no pattern match. Being out of scope was the case for around 15 of the experimentally observed degradation products not predicted. Allowing a heteroatom alpha to the reacting center was a common modification required and suggests more consideration of this factor when determining the scope of a new transformation type.

Case Study 1: Degradation Product Identification for Beclabuvir. One use case of Zeneth is to help elucidate the mechanistic pathways and structures of potential degradation products observed in forced degradation and stability studies for which there is already some understanding of their molecular identity from structural characterization data, for example, mass spectrometry and/or NMR. Scheme 1 outlines the experimentally observed degradation products seen for beclabuvir, (compound ID 23 in this study).³⁷ Eight pharmaceutically relevant degradation products were observed from forced degradation studies aligned with ICH Q1A and Q1B guidelines and were characterized using NMR, high-resolution mass spectrometry, and HPLC techniques. These degradation products were observed across light, oxidative, and acid/base hydrolysis forced degradation study experimental conditions. Table 4 outlines which experimentally observed degradation products were predicted with detail about corresponding predictive conditions. It should be noted that three other degradation products were observed from a solvent interaction with methanol (so-called pseudodegradation products³⁸). However, these are outside the scope of this study and are not discussed further.

The carboxylic acid D1 results from hydrolysis of an amide bond observed under acid and base conditions. Zeneth also predicts this degradation product under these conditions, as well as the amine “fragment” of this hydrolysis [(1*R*,5*S*)-3-methyl-3,8-diazabicyclo[3.2.1]octane], which is not explicitly mentioned by Ye et al.³⁷ However, it is often important to understand all of the products that occur during a given degradation pathway, not just the ones detected by an analytical method, to mitigate missing possible issues. D2 is formed via an oxidative pathway for which Ye et al. propose a reaction mechanism³⁷ and Zeneth mirrors these observations, displaying the formation of the dicarbonyl product via [2 + 2] cycloaddition leading to an intermediate 1,2-dioxetane (Figure

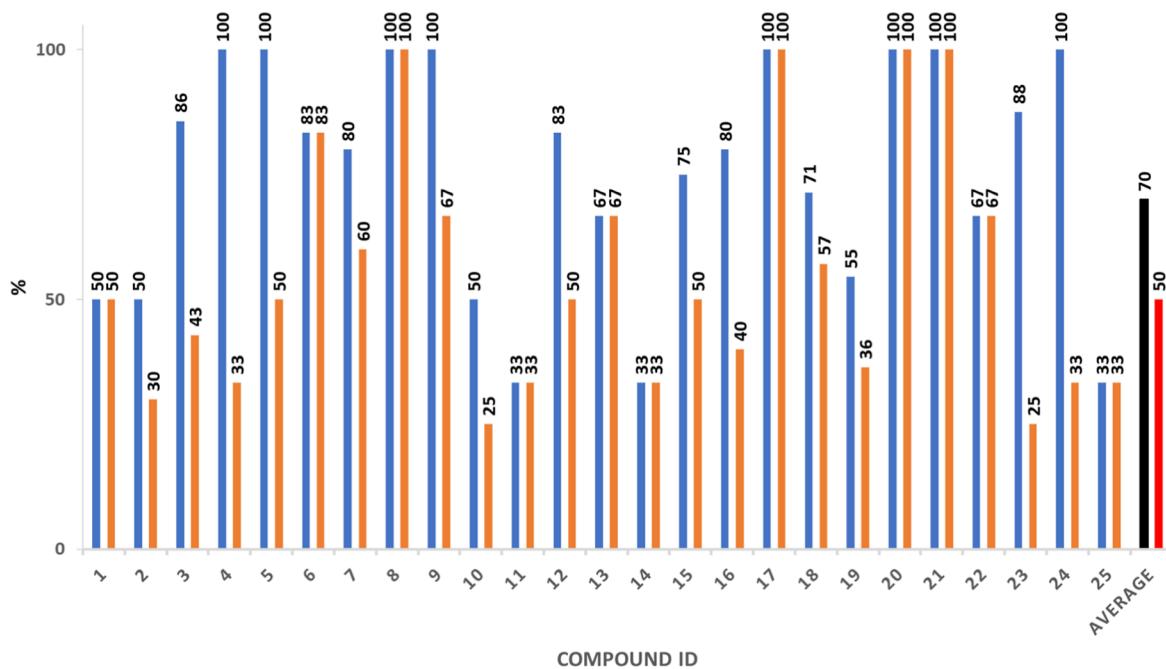


Figure 6. Percentage of correctly predicted degradation products for the 25 DSs using 2020.1 KB (blue) and 2012.2.0 KB (orange). The black (2020.0.1 KB) and red (2012.2.0 KB) bars represent the average value calculated using the equation:

$$\frac{1}{25} \sum_{i=1}^{25} \frac{(\text{Predicted degradation products})_i}{(\text{Observed degradation products})_i} \times 100\%$$

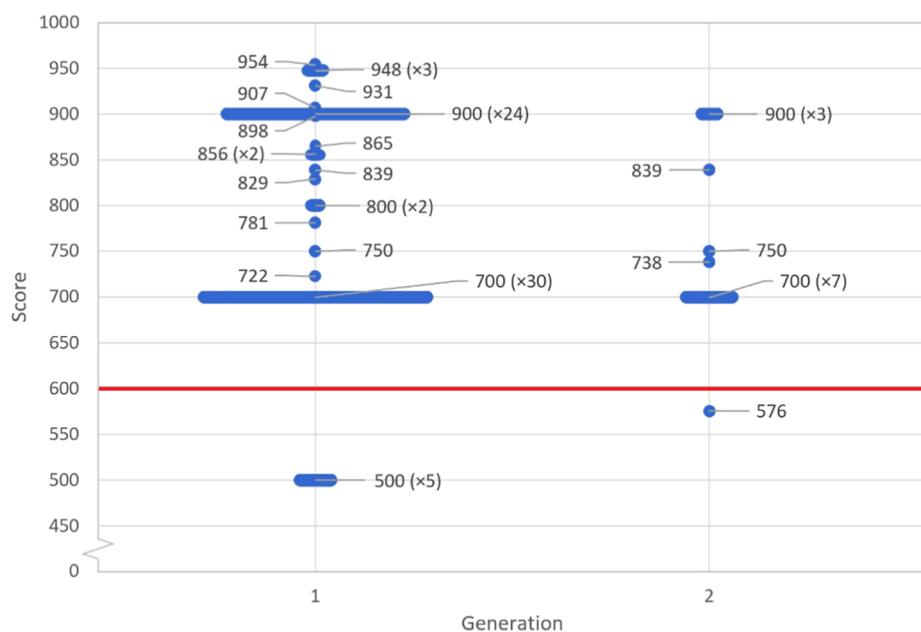


Figure 7. Distribution of likelihood scores vs generation level for experimental degradation products correctly predicted (2020.1.0 KB), using the LSL method.

8). D3 and D5 are diastereoisomers, and it should be noted that only one degradation product (with partially undefined stereochemical configuration) will be predicted, not the individual diastereoisomers. Again, Zeneth would also predict the proposed mechanism outlined by Ye et al.,³⁷ formation of the alpha-hydroxy iminium ion (Int 3) which then undergoes a pinacol-type rearrangement (1,2-shift) as a possibility to generate the dihydropyrrol-3-one degradation product. How-

ever, Zeneth would predict that the first step in this pathway was 2,3-epoxide formation (Int 2).³⁹ This epoxide intermediate is not part of the mechanistic pathway suggested by Ye et al.³⁷

D4 is formed via hydrolysis of the *N*-acylsulfamide bond; this degradation product was not predicted, but related chemistry is in the knowledge base, namely, hydrolysis of sulfamides and of *N*-acylsulfonamides. Scope expansion of the hydrolysis of *N*-acylsulfonamides to include *N*-acylsulfamides

Scheme 1. Experimentally Observed Degradation Products of Beclabuvir Observed During Forced Degradation Experiments. Key: Green = Correctly Predicted. Purple = Not Predicted (Chemistry Not Captured in the Scope of an Existing Transformation Type)

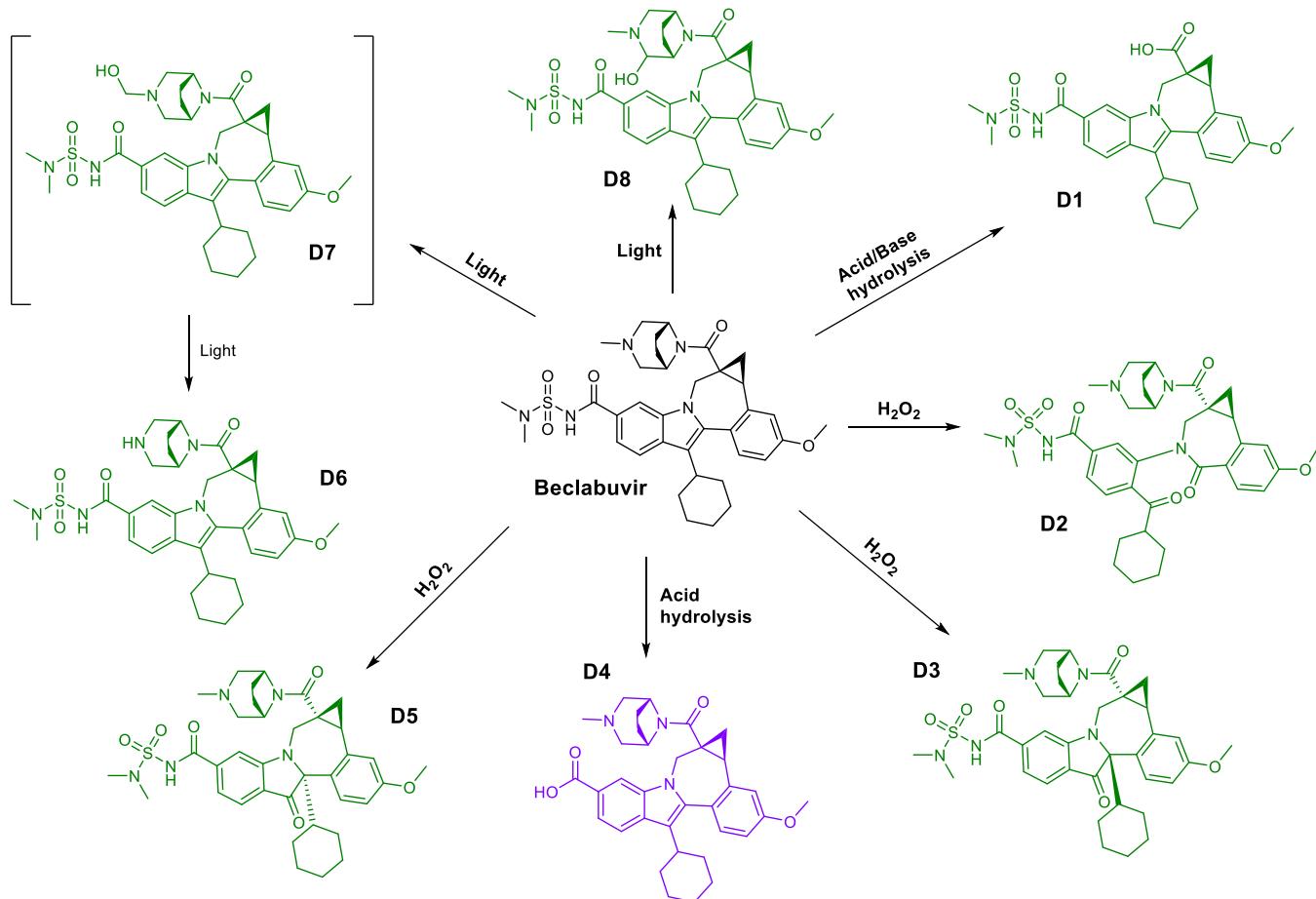


Table 4. Experimental (Forced Degradation Studies) and Predicted (In silico Assessment) Details for the Degradation Products of Beclabuvir

degradant ID	experimental condition	predicted by Zeneth	predicted condition(s)	score	generation
D1	acid/base hydrolysis	yes	pH and water	700	first
D2	hydrogen peroxide	yes	oxygen and light	700	first
D3	hydrogen peroxide	yes	oxygen or peroxide	700	first
D4	acid hydrolysis	no	n/a	n/a	n/a
D5	hydrogen peroxide	yes	oxygen or peroxide	700	first
D6	light	yes	pH and oxygen and (metal or light or radical)	700	first
D7	light	yes	pH and oxygen and (metal or light or radical)	700	first
D8	light	yes	pH and oxygen and (metal or light or radical)	722	first

could be an appropriate modification. Using these experimental data, these modifications will provide a greater depth of knowledge to this transformation type. Ye et al. deduce that D7, produced from photochemical degradation, is an unstable hemiaminal intermediate that releases a molecule of form-

aldehyde to form the demethyl degradation product D6.³⁷ D6 was predicted along with the hemiaminal intermediate D7 and the iminium intermediate suggested by Ye et al. (Figure 9) who proposed this happens via N-oxide formation, followed by dehydration.³⁷ This is a well-supported mechanism²⁵ that Zeneth also predicts but at a low likelihood level (score = 300). The N-oxidation step is deemed slow in acid²⁵ and the dehydration step slow in base, yielding low scores over two steps at extreme pH values (i.e., pH 1 and 13, those used in the study). At pH values around 7, the predicted two-step score is higher (but never more than 500). From the foregoing, two conclusions can be drawn: 1) a future benchmarking study should consider running a prediction at pH 7 as well as 1 and 13; 2) the pH-dependencies of the transformation types for the N-oxidation step and the dehydration step may require a re-evaluation. An alternative but also well-supported mechanism is predicted as well (*cf.* Figure 9). Generation of the amine alpha-radical, either via direct H-abstraction or single-electron transfer followed by deprotonation, and subsequent rapid reaction with molecular oxygen gives a peroxy radical, which abstracts a hydrogen atom from any donor present, yielding an alpha-hydroperoxy amine. This hydroperoxyde, being structurally related to a hemiaminal, can lose hydrogen peroxide to give an iminium compound, which then adds water to furnish a hemiaminal. The hemiaminal finally breaks down into the products, an amine, and a carbonyl compound. This sequence received a score of 700. This demonstrates that Zeneth can

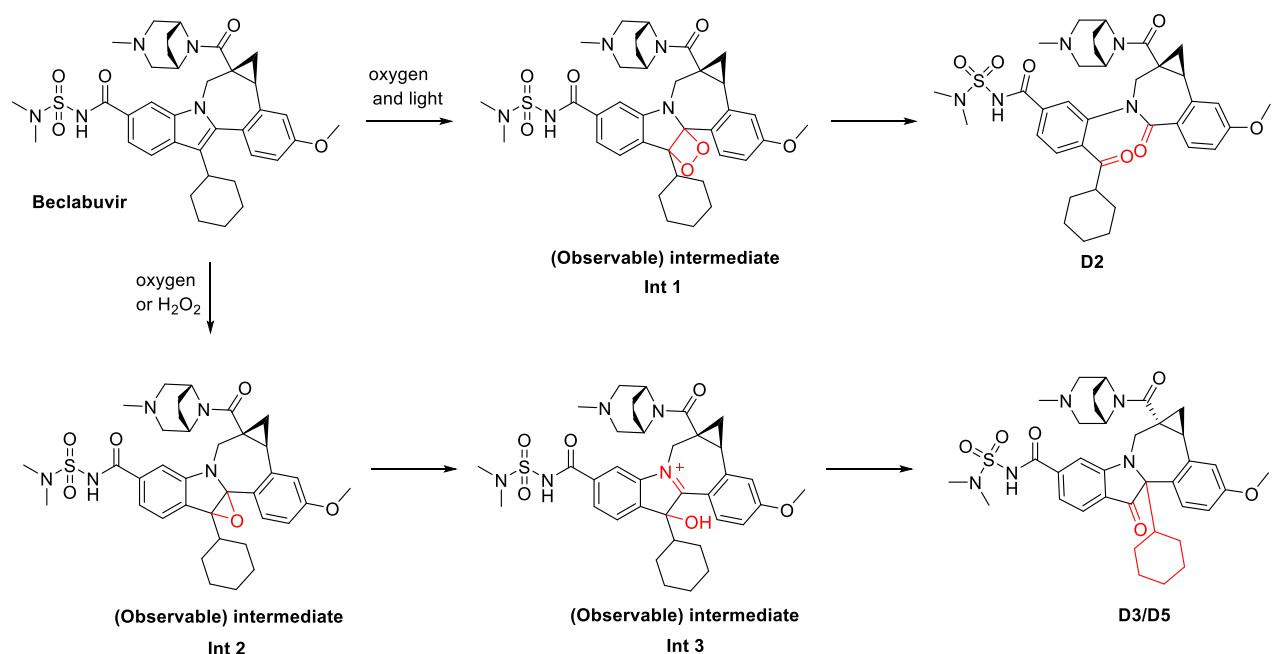


Figure 8. Predicted degradation pathway for the formation of D2 and D3/D5 for beclabuvir.

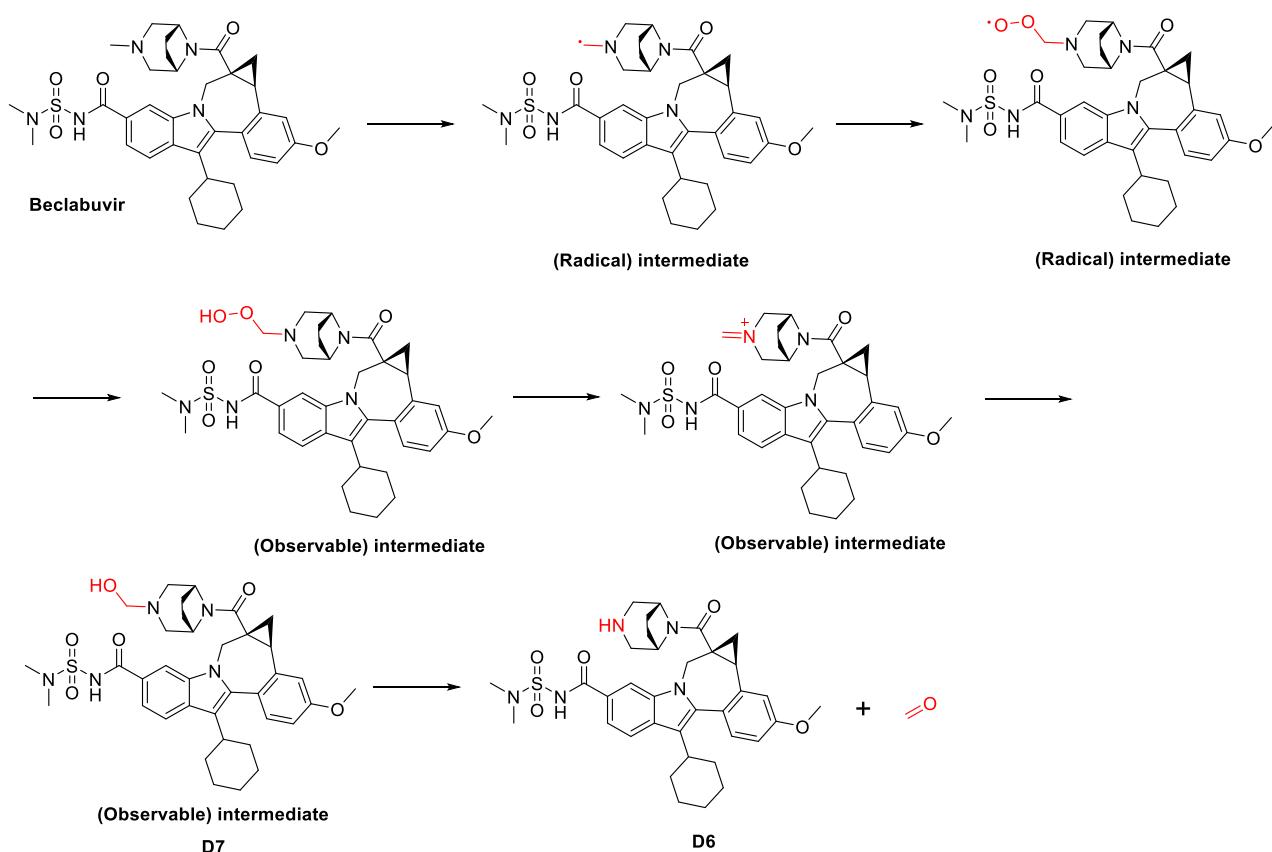


Figure 9. Predicted degradation pathway for D6 and hemiaminal intermediate D7 of beclabuvir.

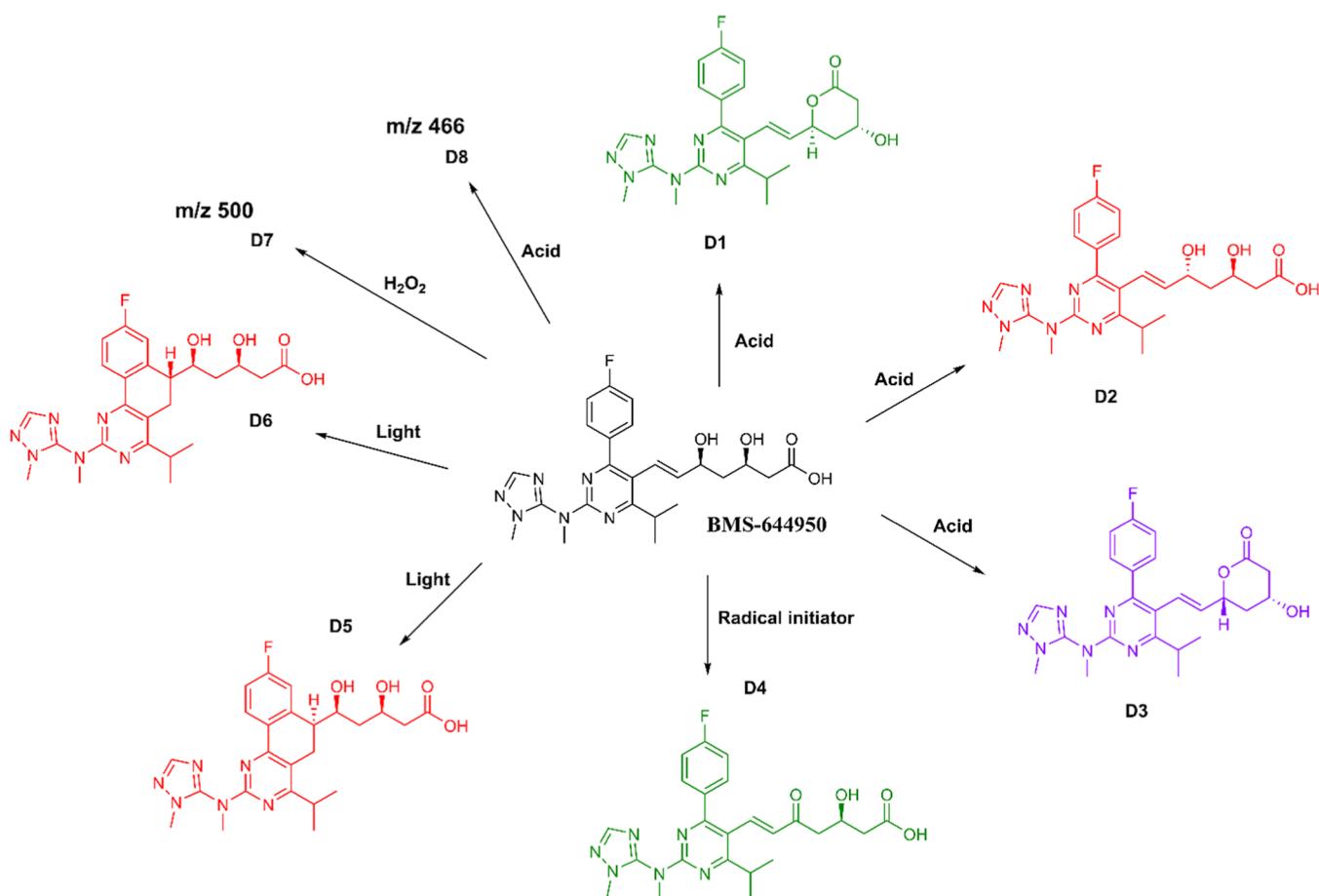
propose alternative mechanisms for consideration that can provide insight into potential control strategies, based on the reaction conditions (e.g., radical initiator vs peroxide), or structural modification.

Formation of D8 follows a mechanistic path similar to that of D6 and D7. Ye et al. suggest N-oxide formation, dehydration, formation of an iminium 'b', and finally, addition

of water to give the hemiaminal D8.³⁷ As described above, Zeneth can also predict this mechanistic pathway, albeit at a low likelihood level for the study conditions, favoring instead a radical pathway via direct H-abstraction or single-electron transfer.

Overall, seven out of the eight experimentally observed degradation products were correctly predicted for beclabuvir,

Scheme 2. Experimentally Observed Degradation Products of BMS-644950 Observed During Forced Degradation Experiments. Key: Green = Correctly Predicted. Purple = Not Predicted (Chemistry Not Captured in the Scope of an Existing Transformation Type). Red = Not Predicted (Chemistry Missing from the Knowledge Base)



all as first-generation products and all with scores ≥ 700 . Many of the predicted mechanistic steps matched those proposed in the Ye et al. publication,³⁷ but prediction results can also provide insight where competing mechanisms may result in the same degradation product, which may aid decision-making when developing control strategies.

Case Study 2: Identification of Unknown Degradation Products for BMS-644950. An additional use case for Zeneth is to aid in the structure elucidation of unknown degradation products that result from forced degradation studies. Scheme 2 shows the degradation profile for BMS-644950,^{40,41} (compound ID 25 in this study). Conditions, outlined in ICH Q1A and Q1B guidelines, were used to degrade this DS and a high-performance liquid chromatography-photo diode array-mass spectrometry method was developed to detect, resolve, and elucidate the resulting degradation products.⁴²

Experimentally, two stereoisomeric lactone products were seen under acidic conditions; D1 and D3, along with an epimer of BMS-644950; D2. Tattersall and Ruan propose that BMS-644950 undergoes an acid-catalyzed dehydration forming a carbocation intermediate which can cyclize to form the two stereoisomeric lactones.⁴⁰ The carbocation intermediate can also add water to form D2. All these reaction steps are considered reversible.⁴⁰ The enone D4 was seen in an oxidative study in the presence of a radical initiator (AIBN). Two diastereomeric products, D5 and D6, generating a new

carbocyclic ring, were seen under UV light (solution and solid state). Two additional peaks in the chromatogram of the stressed sample were observed: one when peroxide was present and one under acidic conditions, D7 and D8, respectively, for which structures could not be elucidated. For clarity, the two unknown peaks have not been counted as observed experimental degradation products nor have they been counted as predicted by Zeneth.

The enone D4 is predicted as a first-generation degradation product with a high likelihood score (954) via a radical initiated oxidative pathway (transformation 062; oxidation of alcohol to a carbonyl compound). The chemistry to generate lactones D1 and D3 is within the knowledge base (transformation 066; lactonization of hydroxy acid or derivative), but the epimerization to generate D2 is not. This results in the prediction of lactone D1 as a first-generation degradation product with a score of 900, but not of lactone D3. Inclusion of the missing epimerization chemistry, which proceeds via an allylic carbonium ion, would result in both D2 and D3 being predicted. Moreover, both lactones D1 and D3 can be formed directly through intermolecular cyclization of the carbonium ion with the carboxylic acid; this chemistry would require an additional lactonization transformation type with a different mechanism and condition dependency. Note that transformation 066 proceeds via nucleophilic attack of the delta-hydroxy group on the carboxylic acid through the “normal” tetrahedral mechanism for lactone formation.

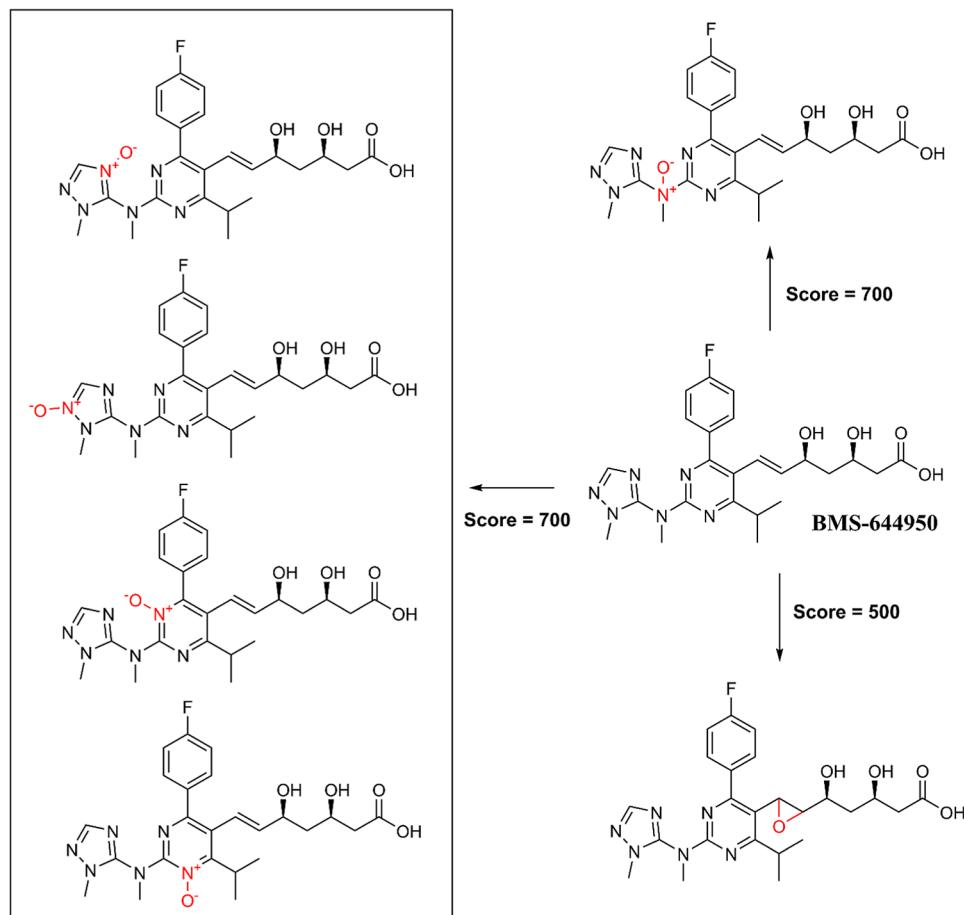


Figure 10. Theoretically predicted degradation products for BMS-644950 consistent with the nominal mass of an experimentally observed unknown degradation product; D7.

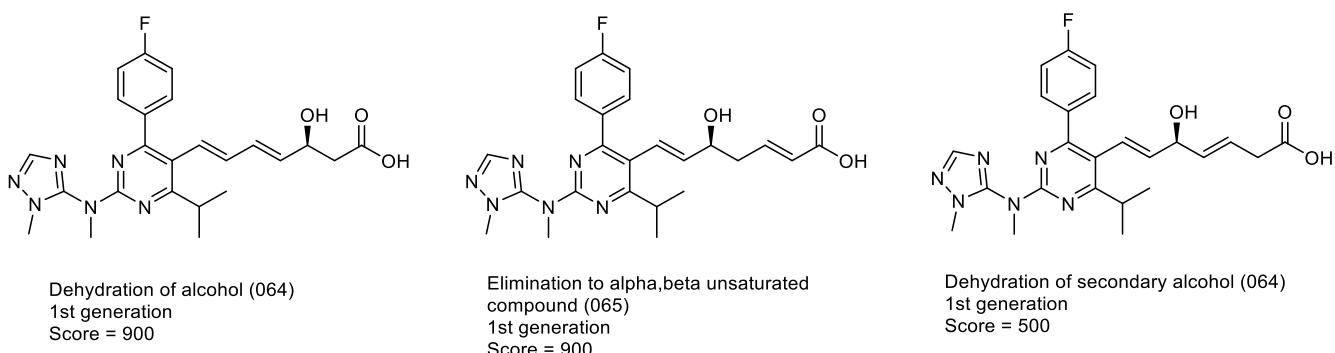


Figure 11. Theoretically predicted degradation products consistent with the nominal mass of an experimentally observed unknown degradation product; D8.

Interestingly, the chemistry to predict diastereoisomers D5 and D6 will be included in the next release of the knowledge base. The chemistry was found in an excellent review on photodegradation pathways by Burrows et al.,⁴³ independently of this benchmarking study. This chemistry was incorporated into the knowledge base following peer review between Lhasa and its industry partners. This example illustrates a key mechanism by which the knowledge base is expanded and kept relevant.

Elucidation of the structure of degradation products D7 and D8 could not be determined using the LC–MS methods employed. However, theoretical degradation products corre-

sponding to both masses were predicted, which, when coupled with the additional experimental data, may provide a means to aid structure elucidation.

For the unknown degradation product observed in the presence of peroxide, D7, a mass of 500 Da was determined. Looking at the first-generation level only, at high pH, six possibilities are predicted for this experimentally observed peak (Figure 10). Five N-oxide possibilities, each at a score of 700 and an epoxidation of the alkene predicted as less likely than the other possibilities (score of 500). One of these predicted possibilities may be correct (overprediction of degradation products will be discussed in a later section); hence, *in silico*

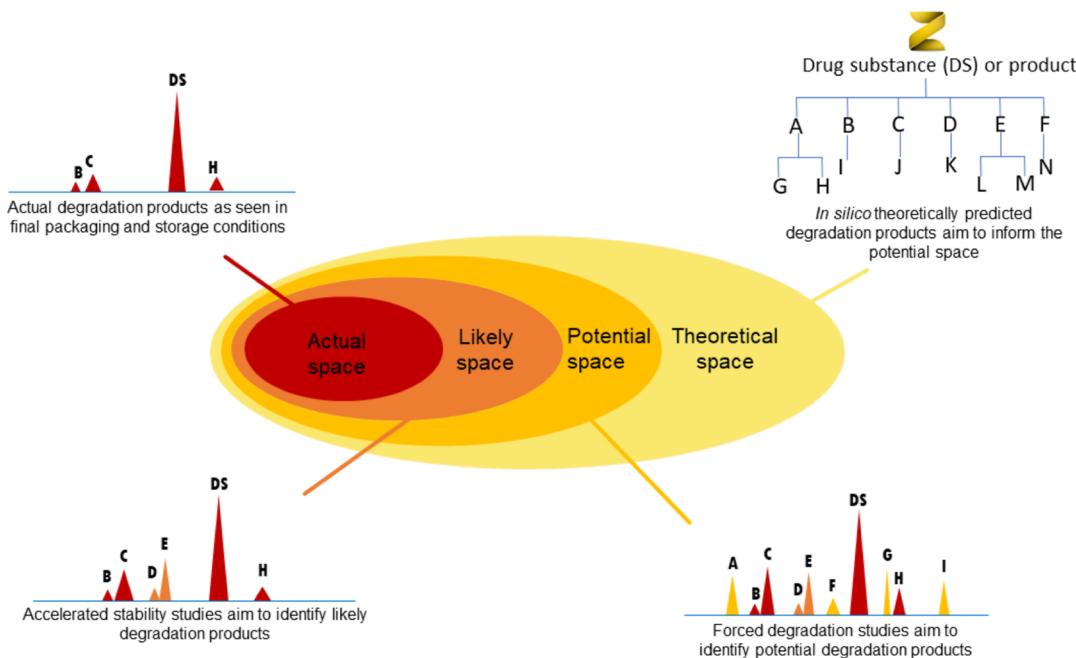


Figure 12. Knowledge of the theoretical space from an in silico prediction can be used to support identification of degradation products, which will ultimately be observed in final packaging and storage conditions (the actual space).

assessments can provide a starting point to map additional data and knowledge to come to an informed conclusion.

For the unknown degradation product observed in acidic conditions; D8, a mass of 466 Da was determined. Looking at only the first-generation level, three possibilities were predicted (Figure 11). Two of these as “very likely” (score = 900) because their formation leads to conjugated products, as opposed to dehydration of a secondary alcohol which forms an unconjugated product (score = 500). Theoretically predicted degradation products can be used to assist in structural elucidation studies.

Overprediction Analysis. Zeneth is designed to process a given chemical structure through its knowledge base of chemical transformation types. It predicts theoretical degradation products when a transformation type matches the conditions and parameters defined by the user. Degradation products are predicted based on chemical functionality present. Zeneth therefore predicts the entire theoretical degradation space for a given chemical structure (as defined by the conditions and rules in the knowledge base), generating all possibilities about which it has knowledge (limited by the contents of its knowledge base). Overprediction of degradation products is consequently present by design and is generally preferable to underprediction from a drug development perspective. For this benchmarking study, overprediction is defined as the number of predicted degradation products that are not observed experimentally. Overpredicted degradation products either did not form in the forced degradation experiments employed by the participating companies or were formed at levels insufficient to trigger their structural characterization.

Figure 12 is a representation of how theoretically predicted degradation products (i.e., the theoretical space) from in silico tools, and experimentally observed degradation products (i.e., the potential space, likely space, and actual space) from stress testing studies and stability studies can be integrated into a larger drug degradation knowledge space. As a new chemical

entity enters development, one of the first activities that may be performed to understand the chemical stability risks associated with a DS or a DP, is the use of in silico tools to generate theoretically predicted degradation products. A predictive system with 100% sensitivity would, by definition, provide coverage for all experimentally observed degradation products (cf. Figure 12). Outputs from a comprehensive forced degradation study design provide a potential degradation profile (i.e., potential space). This, in part, can underpin the optimization and validation of stability indicating methods and support the development of appropriate control strategies. Properly designed stress testing studies provide coverage for degradation products observed in both accelerated and real-time stability studies. The likely space represents degradation products observed from accelerated and long-term stability studies with no packaging controls in place (e.g., open dish), while the actual space represents the degradation products observed and controlled in a final packaged product over shelf life at storage conditions. The integrity of the likely and actual spaces as well as the power of the stability indicating analytical methods are dependent on the completeness and comprehensive nature of the forced degradation studies and degradation chemistry understanding of the drug candidate.²⁵

Although overprediction is a product of design, it can result in many degradation products being predicted, which can make analysis of the prediction result difficult. Improving the accuracy of a prediction by reducing the level of overprediction without losing any sensitivity is a challenge. Understanding which transformation types as well as which reaction modes, i.e., which unimolecular or bimolecular reactions are allowed (cf. Table 2) are contributing to excessive levels of overprediction are key steps in efforts to reduce overprediction.

Assessment of the overprediction levels for the 25 DSs in this study was performed using KNIME (for details of the methodology used, see the Experimental Section). Prediction results using the LSL calculation method identified a total of 5079 theoretical degradation products as being overpredicted.

An assessment of the reaction modes showed a 22% reduction (1121 theoretical degradation products) could be achieved if the (user-defined) reaction modes 'Q + D' and 'D + D' were removed (Table 5). It is agreed among this group of authors

Table 5. Proportion of Overpredicted Degradation Products Attributed to Specific Reaction Modes When Using the LSL Calculation Method^a

reaction mode type	number of experimentally observed degradation products	number of overpredicted degradation products	proportion of overpredicted degradation products
Q + D	1	538	10.59% (538/5079 × 100)
D + D	0	583	11.48% (583/5079 × 100)
Total	1	1121	22.07%

^aQ = query compound; D = degradation product(s).

that experimental degradation products arising from these types of reaction mode are less likely, and this is corroborated by only one of the 90 experimentally observed degradation products being predicted as a 'Q + D' reaction. Therefore, a potential measure to reduce overprediction is to disable the 'Q + D' and 'D + D' reaction modes, enabling only when a prediction warrants it.

To assess transformation types responsible for overprediction, a decision was made to analyze the first step ($n = 1$) only, to get a clear idea for major causes of overprediction. Analysis of two-step pathways would have made it difficult to derive clear causes due to the large numbers involved. Reducing the level of overprediction in the first step will naturally reduce the level of overprediction in the second step. The total number of overpredicted degradation products in the first step is 447. Oxidations are by far the most overpredicted category when looking at first-step degradation products (see Supporting Information). Looking at this analysis at a more granular level (Figure 13) by plotting the actual transformation numbers causing the overprediction at the first step. 37% of the overprediction observed is attributed to just six transformation types (transformations: 094, 159, 163, 180, 293, and 412 (see

Supporting Information)). Other overpredicting transformation types will not be discussed in further detail. Several of the observed experimental degradation products in this study are predicted by these six transformation types, suggesting that the chemistry itself is valid, but refinement may be required.

N-Oxidation of pyridine-type nitrogens (transformation 094) has a wide substrate scope which includes oxazoles, thiazoles, and pyrimidines. All N-oxides arising from this transformation are given the same likelihood score (cf. Figure 10). The prediction for BMS-644950 generated five possible N-oxides for an experimentally observed degradation product at a mass of 500 Da, all with a likelihood score of 700. The data from this study suggest that some refinement is needed. Either the scope needs to be modified so as not to allow so many heteroaromatic rings to match or the likelihood score needs to be based on the reactivity of the heteroaromatic nitrogen, thus differentiating the "type" of N-oxide predicted. Such scope modifications can be added into the knowledge base and may result in a less overpredicting transformation type. In a similar manner, the hydrolysis of 2-heterosubstituted pyridines (transformation 163) has a substrate scope that includes oxazoles, thiazoles, and pyrimidines as well as a variety of leaving groups. The above suggestions for scope refinement made to transformation 094 to reduce its overprediction may also benefit transformation 163.

A C–H BDE predictor has been developed in-house at Lhasa, and integrated into the Zeneth software, using the data set published by St John et al.,^{44,45} to help interrogate reactions that are initiated by a hydrogen abstraction.^{46,47} It has the potential to derive more accurate predictions for radical-initiated oxidative degradation by using the BDE value of the breaking C–H bond to influence the likelihood score of a given degradation product. Likelihood scores are calculated for a given degradation product using an equation with an assumed BDE of 92 kcal/mol at the equivalence point (equating to a score of 500). This assumption may require some adjustment if the conversion to a likelihood score results in too much overprediction. Likelihood scores of the overpredicting transformation types 159, 180, 293, and 412 are controlled by the BDE predictor. A validation exercise using experimentally derived BDE values would be the ideal



Figure 13. Overpredicted first-generation degradation products broken down by transformation number.

next step to assess the accuracy of the BDE predictor within the software and fine-tune the parameters used to turn the BDE values into likelihood scores. Details of how the BDE predictor was developed as well as the equation for how the calculated BDE influences the likelihood score can be found in the [Supporting Information](#).

Analysis has so far focused on improving first-generation overprediction levels, since this will naturally prune second-generation overprediction levels too. One of the guiding principles behind this analysis is to find a balance between improving accuracy without losing any sensitivity. Establishing which transformation types are the main contributors to overprediction is key to finding this balance and provides a firm direction in which to focus efforts to decrease overprediction and improve the accuracy of predictions.

Pathway Likelihood Validation. A method that may also help to reduce the level of overprediction is to assign less weight to less likely degradation pathways. The MSL calculation method allows the user to let longer pathways happen without generating vast numbers of degradation products. It “prunes” unlikely pathways, giving less weight to less likely pathways more effectively than the LSL calculation method. This study aims to validate this concept. The primary aim is to demonstrate that there is no loss in sensitivity when using MSL in place of LSL. **Table 6** shows that the same number of experimentally observed degradation products is predicted by 2020.1.0 KB regardless of the calculation method employed.

While sensitivity levels remain the same, the level of overprediction is often less per compound when employing the MSL method, suggesting that more accurate predictions are possible ([Figure 14](#)). 660 (9%) fewer degradation products were predicted for 25 of the compounds in this study group when using the MSL method (see [Supporting Information](#)). While 9% may not be viewed as a large decrease, it validates putting effort into fine-tuning the parameters of the MSL calculation method. A future study comparing the calculation methods where “longer” ($n \geq 3$) degradation pathways are analyzed for overprediction would be an applicable next step to this workstream.

The principal aim of the MSL method is to apply less weight to less likely pathways (cf. [Figure 5](#)). Its effect can be observed further when a cutoff level is applied to the overprediction assessment. This study generated degradation products with a score ≥ 400 (when they also conformed to all other user-defined parameters). If a cutoff is applied at ≥ 700 , the amount of overprediction observed reduces dramatically for both calculation methods, but more so for MSL. [Figure 15](#) shows that likelihood scores shift from majority ≥ 700 (green) to majority < 700 (red + blue) for MSL. Likelihood scores ≥ 600 (red) have also been displayed to demonstrate the increased distribution of scores for the MSL method. When a cutoff of ≥ 700 was applied using the LSL method, six experimentally observed degradation products were “lost” i.e., filtered out of the prediction result; five first-generation and one second-generation (cf. [Figure 7](#)). If this same analysis is performed using the MSL method, 12 experimentally observed degradation products are lost; five first generation and seven second generation. All additional experimental degradation products are lost from the second-generation, because of the multiplication method utilized by MSL to derive pathway scores. This suggests that the probabilities used in the MSL method may need to be refined. In addition, a minimum score of ≥ 700

Table 6. Number of Experimental Degradation Products Correctly Predicted by 2020.1.0 KB Using LSL and MSL Calculation Methods

compound ID	experimentally observed degradation products	degradation products correctly predicted by 2020.1.0 KB	
		lowest step likelihood	multiplied step likelihood
1	8	4	4
2	10	5	5
3	7	6	6
4	3	3	3
5	2	2	2
6	6	5	5
7	5	4	4
8	3	3	3
9	3	3	3
10	4	2	2
11	3	1	1
12	6	5	5
13	6	4	4
14	6	2	2
15	4	3	3
16	5	4	4
17	3	3	3
18	7	5	5
19	11	6	6
20	4	4	4
21	2	2	2
22	3	2	2
23	8	7	7
24	3	3	3
25	6	2	2
Total	128	90	90

may be too harsh for MSL calculations; a lower threshold of 600, for example, may allow most observed degradation products to be predicted while still reducing overprediction more effectively than the LSL method. Obtaining a balance between sensitivity and accuracy is a difficult goal to achieve but one where efforts will continue to be focused.

CONCLUSIONS

Zeneth considers the reaction conditions that are used in forced degradation studies and is intended to predict theoretical degradation products that can arise for a given DS or DP. It is founded on an understanding of known chemistry (its knowledge base) supported by property predictors (in this study, a C–H BDE predictor). This study aimed to assess the sensitivity of the 2020.1.0 KB knowledge base vs a historical knowledge base (2012.2.0 KB) when evaluated against an experimental data set (of 25 DSs), as well as the level of overprediction seen in prediction results. Zeneth overpredicts by design because it is focused on assessing the theoretical possibility that a reaction can occur. However, the level of overprediction can be a hindrance to both the accuracy and interpretation of the prediction results. Improving the accuracy of predictions while maintaining a high sensitivity level is a challenge. Transformation types that significantly contributed to the level of overprediction were identified, and efforts to reduce the level of overprediction are an ongoing project. In addition, how *in silico* outputs can be utilized to confirm experimental data or assist with identification of

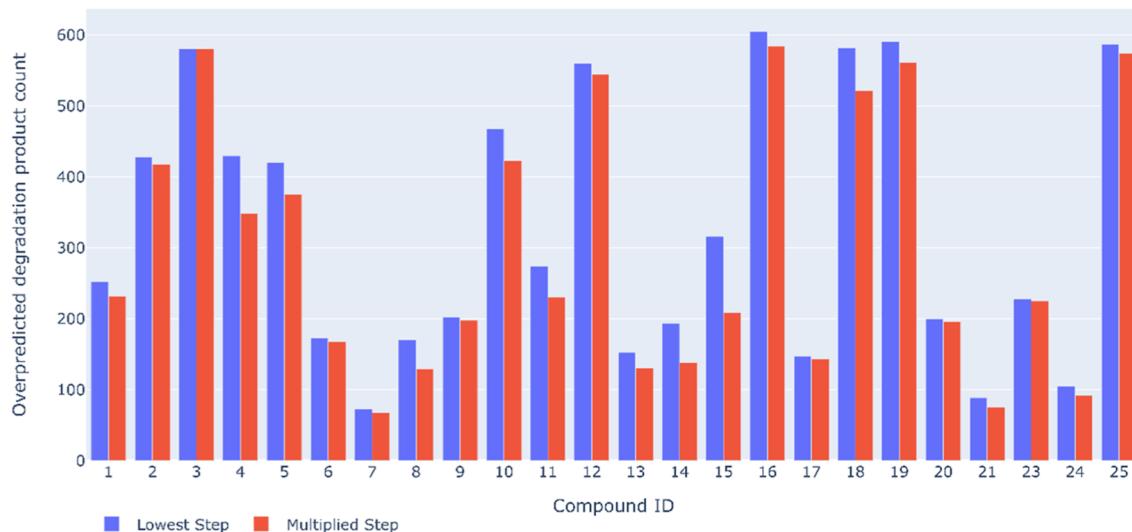


Figure 14. Overprediction levels (for pH 1 and 13 predictions, combined) when using LSL (blue) and MSL (red) calculation methods.

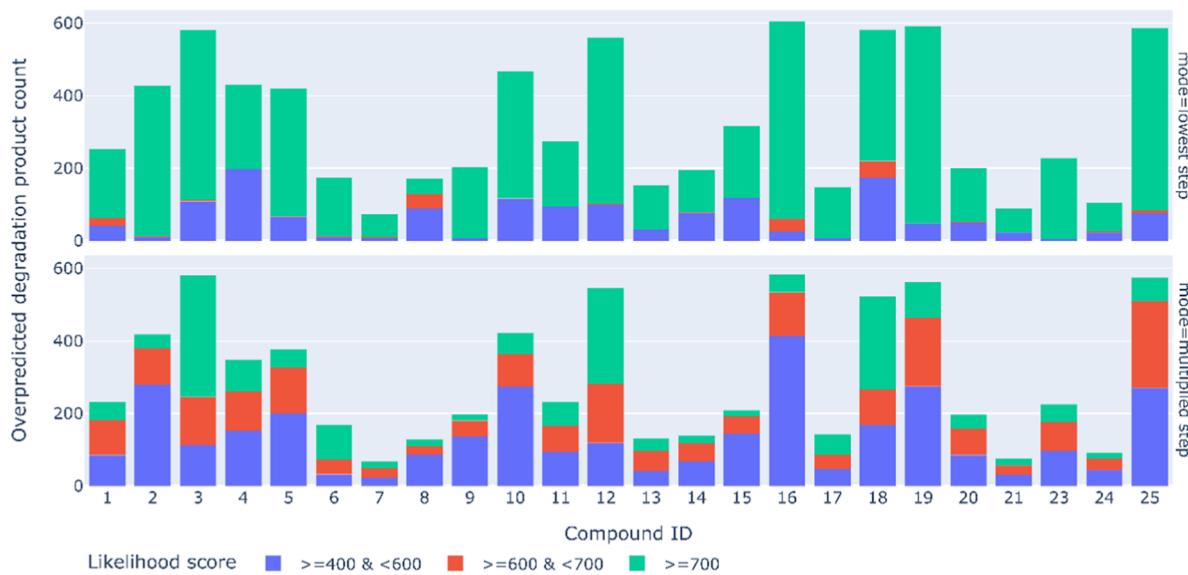


Figure 15. Overprediction levels (for pH 1 and 13 predictions, combined) when using LSL (top) and MSL (bottom) calculation methods, broken down by likelihood score.

unknown degradation products has been outlined in two case studies. Finally, an assessment of the 2020.1.0 KB (containing 531 transformation types) and the 2012.2.0 KB (containing 277 transformation types) was performed using 128 experimentally observed degradation products (from 25 DSs). These were obtained from forced degradation, accelerated, and long-term stability studies (aligned with ICH guidelines). A 40% increase in sensitivity was demonstrated on going from 2012.2.0 KB (50% sensitivity) to 2020.1.0 KB (70% sensitivity). Analysis of those experimentally observed degradation products not correctly predicted allows identification of gaps in the predictivity or areas where expansion of the scope of existing transformation types is required, thus further enhancing the knowledge base. Overall, these results demonstrate that steady expansion and refinement of transformation types have resulted in increased coverage of the experimentally observed degradation space associated with DS structures from several pharmaceutical companies.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.oprd.3c00344>.

Additional information on how the bootstrapping analysis was performed, additional overprediction data and detail on how the C–H BDE predictor was built and works, and Zeneth prediction outputs for LSL and MSL at pH 1 and 13 for beclabuvir and BMS-644950 ([PDF](#))

Prediction results for beclabuvir at pH 1 using the LSL method ([TXT](#))

Prediction results for BMS-644950 at pH 1 using the LSL method ([TXT](#))

Prediction results for beclabuvir at pH 13 using the LSL method ([TXT](#))

Prediction results for BMS-644950 at pH 13 using the LSL method ([TXT](#))

Prediction results for beclabuvir at pH 1 using the MSL method ([TXT](#))

Prediction results for BMS-644950 at pH 1 using the MSL method ([TXT](#))

Prediction results for beclabuvir at pH 13 using the MSL method [TXT](#)

Prediction results for BMS-644950 at pH 13 using the MSL method [TXT](#)

■ AUTHOR INFORMATION

Corresponding Author

Rachel Hemingway — *Granary Wharf House, Lhasa Limited, Holbeck, Leeds LS11 5PS, U.K.*;  [orcid.org/0000-0003-1198-479X](#); Email: rachel.hemingway@lhasalimited.org

Authors

Steven W. Baertschi — *Baertschi Consulting, LLC, Carmel, Indiana 46033, United States*

David Benstead — *Chemical Development, Pharmaceutical Technology & Development, Operations, AstraZeneca, Charter Way, Macclesfield SK10 2NA, U.K.*

John M. Campbell — *GSK, Collegeville, Pennsylvania 19426, United States*

Maggie Coombs — *Granary Wharf House, Lhasa Limited, Holbeck, Leeds LS11 5PS, U.K.*

Zongyun Huang — *Bristol-Myers Squibb, Drug Product Development, New Brunswick, New Jersey 08901, United States*

Ryan Khalaf — *Amgen Inc., Drug Product Technologies, Process Development, Thousand Oaks, California 91320, United States*; Present Address: *School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, United States*

Martin A. Ott — *Granary Wharf House, Lhasa Limited, Holbeck, Leeds LS11 5PS, U.K.*;  [orcid.org/0000-0002-3179-7329](#)

David J. Ponting — *Granary Wharf House, Lhasa Limited, Holbeck, Leeds LS11 5PS, U.K.*;  [orcid.org/0000-0001-6840-2629](#)

Darren L. Reid — *Amgen Inc., Drug Product Technologies, Process Development, Cambridge, Massachusetts 02142, United States*;  [orcid.org/0000-0003-4399-2452](#)

Neil G. Stevenson — *GSK, Medicines Research Centre, Stevenage SG1 2NY, U.K.*;  [orcid.org/0000-0001-5927-6755](#)

Samuel J. Webb — *Granary Wharf House, Lhasa Limited, Holbeck, Leeds LS11 5PS, U.K.*;  [orcid.org/0000-0003-1782-0921](#)

Angela White — *GSK, David Jack Centre for Research and Development, Ware SG12 0DP, U.K.*

Todd Zelesky — *Pfizer Inc., Analytical Research & Development, Groton, Connecticut 06340, United States*

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.oprd.3c00344>

Notes

The authors declare the following competing financial interest(s): Rachel Hemingway, Martin Ott, Maggie Coombs, David Ponting and Samuel Webb are employees of Lhasa Limited, which develops the Zeneth software.

■ ACKNOWLEDGMENTS

The authors would like to thank Dr Alex P. Harding (Lhasa Limited) for being instrumental in organizing and getting legal agreements signed. Dr Martyn Chilton (Lhasa Limited) for guiding the bootstrapping analysis, as well as helpful discussions around this. In addition, John Nightingale (AstraZeneca) and Laura E. Blue (Amgen Inc.) are thanked for their help in gathering relevant experimental degradation data for the study.

■ REFERENCES

- (1) ICH Q3A (R2): Impurities in New Drug Substances. October 25, 2006. Available at: <https://database.ich.org/sites/default/files/Q3A28R229Guideline.pdf> (accessed December 13, 2023).
- (2) ICH Q3B (R2): Impurities in New Drug Products. June 2, 2006. Available at: <https://database.ich.org/sites/default/files/Q3B28R229Guideline.pdf> (accessed December 13, 2023).
- (3) ICH Q1A (R2): Stability testing of new drug substances and products. February 6, 2003. Available at: <https://database.ich.org/sites/default/files/Q1A28R229Guideline.pdf> (accessed December 13, 2023).
- (4) ICH Q1B: Photostability testing of new active substances and medicinal products. November 6, 1996. Available at: <https://database.ich.org/sites/default/files/Q1BGuideline.pdf> (accessed December 13, 2023).
- (5) ANVISA Resolução De Diretoria Colegiada [Collegiate Board of Directors Resolution] RDC - N° 53. 'Estabelece parâmetros para a notificação, identificação e qualificação de produtos de degradação em medicamentos com substâncias ativas sintéticas e semissintéticas, classificados como novos, genéricos e similares, e dá outras providências' ['Establishes parameters for the notification, identification and qualification of degradation products in medicines with synthetic and semi-synthetic active substances, classified as new, generic and similar, and takes other measures']. December 4, 2015. Available at: http://antigo.anvisa.gov.br/documents/10181/3295768/8129RDC_53_2015_COMP.pdf/d38f507d-745c-4f6b-a0a6-bd250f2e9892 (accessed December 13, 2023).
- (6) Campbell, J. M.; Foti, C.; Wang, C.; Adams, N.; Allain, L.; Araujo, G.; Azevedo, R.; Franca, J. R.; Hicks, S. R.; Hostyn, S.; Jansen, P. J.; Kotoni, D.; Kuemannell, A.; Marden, S.; Rullo, G.; Santos, A. C. O.; Slaggett, G. W.; Zelesky, T.; Baertschi, S. W. Assessing the Relevance of Solution Phase Stress Testing of Solid Dosage Form Drug Products: A Cross-Industry Benchmarking Study. *J. Pharm. Sci.* **2022**, *111*, 298–305.
- (7) Zelesky, T.; Baertschi, S. W.; Foti, C.; Allain, L. R.; Hostyn, S.; Franca, J. R.; Li, Y.; Marden, S.; Mohan, S.; Ultramari, M.; Huang, Z.; Adams, N.; Campbell, J. M.; Jansen, P. J.; Kotoni, D.; Laue, C. Pharmaceutical forced degradation (stress testing) endpoints: A scientific rationale and industry perspective. *J. Pharm. Sci.* **2023**, *112*, 2948–2964.
- (8) Alsante, K. M.; Martin, L.; Baertschi, S. W. A Stress Testing Benchmarking Study. *Pharm. Technol.* **2003**, *27*, 60–72.
- (9) Alsante, K. M.; Ando, A.; Brown, R.; Ensing, J.; Hatajik, T. D.; Kong, W.; Tsuda, Y. The role of degradant profiling in active pharmaceutical ingredients and drug products. *Adv. Drug Delivery Rev.* **2007**, *59*, 29–37.
- (10) Pensak, D. A.; Corey, E. J. LHASA - logic and heuristics applied to synthetic analysis. In *Computer Assisted Organic Synthesis*; Wipke, W. T., Howe, W. J., Eds.; ACS: Washington, USA, 1997, pp 1–32.
- (11) Jorgensen, W. L.; Laird, E. R.; Gushurst, A. J.; Fleischer, J. M.; Gothe, S. A.; Helson, H. E.; Paderes, G. D.; Sinclair, S. CAMEO: a program for the logical prediction of the products of organic reactions. *Pure Appl. Chem.* **1990**, *62*, 1921–1932.
- (12) Pole, D. L.; Ando, H. Y.; Murphy, S. T. Prediction of drug degradants using DELPHI: an expert system for focusing knowledge. *Mol. Pharmaceutics* **2007**, *4*, 539–549.

- (13) Tebes-Stevens, C.; Patel, J. M.; Jones, J.; Weber, E. J. Prediction of hydrolysis products of organic chemicals under environmental pH conditions. *Environ. Sci. Technol.* **2017**, *51*, 5008–5016.
- (14) Yuan, C.; Tebes-Stevens, C.; Weber, E. J. Reaction library to predict direct photochemical transformation products of environmental organic contaminants in sunlit aquatic systems. *Environ. Sci. Technol.* **2020**, *54*, 7271–7279.
- (15) Lhasa Limited Zeneth. Available at: Impurity And Degradant Control | Lhasa Limited. (accessed December 13, 2023).
- (16) Grinberg Dana, A.; Wu, H.; Ranasinghe, D. S.; Pickard, F. C.; Wood, G. P. F.; Zelesky, T.; Slaggett, G. W.; Mustakis, J.; Green, W. H. Kinetic Modeling of API Oxidation: (1) The AIBN/H₂O/CH₃OH Radical “Soup. *Mol. Pharmaceutics* **2021**, *18*, 3037–3049.
- (17) Wu, H.; Grinberg Dana, A.; Ranasinghe, D. S.; Pickard, F. C.; Wood, G. P. F.; Zelesky, T.; Slaggett, G. W.; Mustakis, J.; Green, W. H. Kinetic Modeling of API Oxidation: (2) Imipramine Stress Testing. *Mol. Pharmaceutics* **2022**, *19*, 1526–1539.
- (18) Coley, C. W.; Jin, W.; Rogers, L.; Jamison, T. F.; Jaakkola, T. S.; Green, W. H.; Barzilay, R.; Jensen, K. F. A graph-convolutional neural network model for the prediction of chemical reactivity. *Chem. Sci.* **2019**, *10*, 370–377.
- (19) Segler, M.; Preuss, M.; Waller, M. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **2018**, *555*, 604–610.
- (20) Venkatasubramanian, V.; Mann, V. Artificial intelligence in reaction prediction and chemical synthesis. *Curr. Opin. Chem. Eng.* **2022**, *36*, 100749.
- (21) Griffin, D.; Coley, C. W.; Frank, S. A.; Hawkins, J. M.; Jensen, K. F. Opportunities for Machine Learning and Artificial Intelligence to Advance Synthetic Drug Substance Process Development. *Org. Process Res. Dev.* **2023**, *27*, 1868–1879.
- (22) Ali, M. A.; Hemingway, R.; Ott, M. A. In silico drug degradation prediction. In *Methods for Stability Testing of Pharmaceuticals. Methods in Pharmacology and Toxicology*; Bajaj, S., Singh, S., Eds.; Humana Press: New York, USA, 2018, pp 53–73.
- (23) Parenty, A. D. C.; Button, W. G.; Ott, M. A. An expert system to predict the forced degradation of organic molecules. *Mol. Pharmaceutics* **2013**, *10*, 2962–2974.
- (24) Albini, A.; Fasani, E. *Drugs, Photochemistry and Photostability*. Royal Society of Chemistry: London, UK, 1998.
- (25) Baertschi, S. W.; Alsante, K. M.; Reed, R. A. *Pharmaceutical Stress Testing: Predicting Drug Degradation*. Informa Healthcare: London, UK, 2011.
- (26) Min, L. *Organic Chemistry of Drug Degradation*. Royal Society of Chemistry: Cambridge, UK, 2012.
- (27) Smith, M. B.; March, J. *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*. Wiley: New Jersey, USA, 2001.
- (28) Kleinman, M. H.; Baertschi, S. W.; Alsante, K. M.; Reid, D. L.; Mowery, M. D.; Shimanovich, R.; Foti, C.; Smith, W. K.; Reynolds, D. W.; Nefliu, M.; Ott, M. A. In silico prediction of pharmaceutical degradation pathways: A benchmarking study. *Mol. Pharmaceutics* **2014**, *11*, 4179–4188.
- (29) Caille, S.; Cui, S.; Faul, M. M.; Mennen, S. M.; Tedrow, J. S.; Walker, S. D. Molecular Complexity as a Driver for Chemical Process Innovation in the Pharmaceutical Industry. *J. Org. Chem.* **2019**, *84*, 4583–4603.
- (30) Agarwal, P.; Huckle, J.; Newman, J.; Reid, D. L. Trends in small molecule drug properties: A developability molecule assessment perspective. *Drug Discovery Today* **2022**, *27*, 103366–103419.
- (31) Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Kotter, T.; Meinl, T.; Ohl, P.; Sieb, C.; Thiel, K.; Wiswedel, B. KNIME: The Konstanz Information Miner. In *Data Analysis, Machine Learning and Applications. Studies in Classification, Data Analysis, and Knowledge Organization*; Preisach, C., Burkhardt, H., Schmidt-Thieme, L., Decker, R., Eds.; Springer: Berlin, Germany, 2008, pp 319–326.
- (32) Available at: www.knime.org.(accessed December 13, 2023).
- (33) Available at: www.rdkit.org (accessed December 13, 2023).
- (34) Lhasa Limited Derek. In silico Mutagenicity Assessment | Lhasa Limited. Available at (accessed December 13, 2023).
- (35) Heller, S. R.; McNaught, A.; Pletnev, I.; Stein, S.; Tchekhovskoi, D. InChI, the IUPAC international chemical identifier. *J. Cheminf* **2015**, *7*, 23.
- (36) Dalby, A.; Nourse, J. G.; Hounshell, W. D.; Gushurst, A. K. I.; Grier, D.; Leland, B. A.; Laufer, J. Description of several chemical structure file formats used by computer programs developed at molecular design limited. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 244–255.
- (37) Ye, Q.; Huang, Y.; Grier, S.; Miller, S. A. Oxidative degradation pathways of beclabuvir hydrochloride mediated by hydrogen peroxide and UV/Vis light. *J. Pharm. Biomed. Anal.* **2019**, *172*, 388–394.
- (38) Singh, D. K.; Singh, S. Identification and control of pseudo degradation products during stress testing on drugs. *Spinco Biotechnol. Cutting Edge* **2019**, *26*–31.
- (39) Zhang, X.; Foote, C. S. Dimethyldioxirane oxidation of indole derivatives. Formation of novel indole-2,3-epoxides and a versatile synthetic route to indolinones and indolines. *J. Am. Chem. Soc.* **1993**, *115*, 8867–8868.
- (40) Tattersall, P.; Ruan, J. Poster Presentation, 232nd American Chemical Society National Meeting, San Francisco, September 10–14, 2006.
- (41) Ahmad, S.; Madsen, C. S.; Stein, P. D.; Janovitz, E.; Huang, C.; Ngu, K.; Bisaha, S.; Kennedy, L. J.; Chen, B. C.; Zhao, R.; Sitkoff, D.; Monshizadegan, H.; Yin, X.; Ryan, C. S.; Zhang, R.; Giancarli, M.; Bird, E.; Chang, M.; Chen, X.; Setters, R.; Search, D.; Zhuang, S.; Nguyen-Tran, V.; Cuff, C. A.; Harrity, T.; Darienzo, C. J.; Li, T.; Reeves, R. A.; Blanar, M. A.; Barrish, J. C.; Zahler, R.; Robl, J. A. (3R,SS,E)-7-(4-(4-Fluorophenyl)-6-isopropyl-2-(methyl(1-methyl-1H-1,2,4-triazol-5-yl)amino)pyrimidin-5-yl)-3,5-dihydroxyhept-6-enoic acid (BMS-644950): a rationally designed orally efficacious 3-hydroxy-3-methylglutaryl coenzyme-A reductase inhibitor with reduced myotoxicity potential. *J. Med. Chem.* **2008**, *51*, 2722–2733.
- (42) Ruan, J.; Tattersall, P.; Lozano, R.; Shah, P. The role of forced degradation studies in stability indicating HPLC method development. *Am. Pharm. Rev.* **2006**, *9*, 46–53.
- (43) Burrows, H. D.; Canle L, M.; Santaballa, J. A.; Steenken, S. Reaction pathways and mechanisms of photodegradation of pesticides. *J. Photochem. Photobiol. B* **2002**, *67*, 71–108.
- (44) St John, P. C.; Guan, Y.; Kim, Y.; Kim, S.; Paton, R. S. Prediction of organic homolytic bond dissociation enthalpies at near chemical accuracy with sub-second computational cost. *Nat. Commun.* **2020**, *11*, 2328.
- (45) St John, P. C.; Guan, Y.; Kim, Y.; Etz, B. D.; Kim, S.; Paton, R. S. Quantum chemical calculations for over 200,000 organic radical species and 40,000 associated closed-shell molecules. *Sci. Data* **2020**, *7*, 244.
- (46) Lienard, P.; Gavartin, J.; Boccardi, G.; Meunier, M. Predicting Drug Substances Autoxidation. *Pharm. Res.* **2015**, *32*, 300–310.
- (47) Andersson, T.; Broo, A.; Evertsson, E. Prediction of Drug Candidates' Sensitivity Toward Autoxidation: Computational Estimation of C-H Dissociation Energies of Carbon-Centered Radicals. *J. Pharm. Sci.* **2014**, *103*, 1949–1955.