Journal of Medicinal Chemistry

pubs.acs.org/jmc

Perspective

Cheminformatic Analysis of Core-Atom Transformations in Pharmaceutically Relevant Heteroaromatics

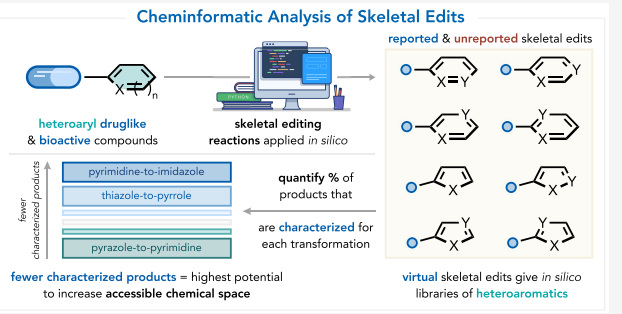
G. Logan Bartholomew,[†] Lucas J. Karas,[†] Reilly M. Eason, Charles S. Yeung,* Matthew S. Sigman,* and Richmond Sarpong*

Cite This: https://doi.org/10.1021/acs.jmedchem.4c02839

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Heteroaromatics are the basis for many pharmaceuticals. The ability to modify these structures through selective core-atom transformations, or "skeletal edits", can dramatically expand the landscape for drug discovery and development. However, despite the importance of core-atom modifications, the quantitative impact of such transformations on accessible chemical space remains undefined. Here, we report a cheminformatic platform to analyze which skeletal edits would most increase access to novel chemical space. This study underscores the significance of emerging single and multiple core-atom transformations of heteroaromatics in enhancing chemical diversity, for example, at a late-stage of a drug discovery campaign. Our findings provide a quantitative framework for prioritizing core-atom modifications in heteroaromatic structural motifs, calling for the development of new methods to achieve these types of transformations.



SIGNIFICANCE, IMPACT, AND INNOVATION

This work introduces a cheminformatic framework to quantify the impact of core-atom transformations in heteroaromatics and highlights which skeletal edits could most significantly increase access to chemical space. Additionally, this work calls for the development of new synthetic methods to enable prioritized core-atom modifications.

INTRODUCTION

Heterocycles are prevalent in bioactive molecules and ubiquitous in modern synthetic chemistry.[1−4] As such, developing methods for the selective functionalization or modification of heterocycles remains a pressing aim for pharmaceutical and agrochemical discovery campaigns.[5] Transformations that achieve single core-atom changes to heterocyclic structural motifs at a late-stage—increasingly termed "skeletal edits"—have emerged at a rapid pace in recent years.[6] These transformations complement more established peripheral diversification methods and enable efficient access to less explored regions of chemical space without requiring *de novo* synthesis.

As the field of single-atom skeletal editing has evolved, a key question has emerged: which single-atom transformations would be the most impactful if developed? We have sought to answer this question by employing computational tools to screen virtual chemical space for the novelty and diversity introduced by single-atom transformations.[7−10] Advances in software and hardware over the past decade have enabled researchers to navigate virtual chemical spaces containing billions of structures that can be made on demand using cheminformatics.[9,11] Virtual comparative analysis can identify areas of chemical space that are currently inaccessible or underexplored. This may assist in the identification of synthetic challenges and opportunities for reaction development that could provide access to these unique structural motifs. Indeed, the use of computational tools is closely linked to the development of new synthetic methodologies as well as the discovery and optimization of biologically active compounds.[12−20]

Furthermore, cheminformatics provides valuable tools for quantifying the impact of chemical transformations on large compound libraries. In particular, cheminformatic tools can systematically assess how specific molecular parameters are quantitatively influenced by different types of chemical reactions. For example, Cernak, Ma, Glorius, and co-workers recently developed a cheminformatic platform for the virtual analysis of late-stage arene hydrogenation in the context of

Received: November 20, 2024
Revised: January 26, 2025
Accepted: February 27, 2025

Downloaded via UNIV OF CALIFORNIA BERKELEY on March 8, 2025 at 19:36:32 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ACS Publications

© XXXX American Chemical Society



Figure 1. Percentage of the total data set that each type of analyzed heterocycle represents in the (A) Enamine 350/3, (B) Mcule, and (C) ChEMBL data sets of druglike compounds or screening fragments.

drug discovery.[21] To the best of our knowledge, there has been no quantitative analysis of the potential impact that a given skeletal edit may have on accessible chemical space. Powerful methods for skeletal editing of saturated heterocycles and carbocycles have emerged;[22−29] however, the majority of recently developed skeletal editing reactions focus on aromatic heterocycles (i.e., heteroarenes).[30−44] Because of the prevalence of heteroarenes in FDA-approved drugs,[2,4] we sought to develop a cheminformatic platform to analyze the potential impact of "virtual" skeletal edits by evaluating the number of uncharacterized molecules generated when both single-atom and multiple-atom skeletal transformations are applied to the most prevalent heteroaromatic motifs in pharmaceutical compounds. This platform enabled us to identify transformations for converting one heteroaromatic unit to another that would most greatly increase accessible chemical space if developed.

METHODS

We extracted molecules containing 11 distinct heteroaromatic rings from the Enamine 350/3 lead-like library, a subset of the REAL database containing over 613 million synthetically feasible molecules.[45] In total, over 275 million of the 613 million molecules contain at least one of the 11 heteroaromatic motifs we extracted. We selected these heteroaromatic cores (shown in Figure 1) because of their prevalence in FDA-approved drug molecules.[2−4] Compounds containing fused heteroaromatic rings were also included; for example, quinolines and *aza*-indoles were classified as pyridines, quinazolines were classified as pyrimidines, indoles were classified as pyrroles, etc. The ring-fusion element of benzo-fused heteroaromatics was maintained throughout each virtual skeletal edit transformation (i.e., the fusing carbon atoms were excluded as viable points of transformation). The same extraction protocol was also applied to the Mcule[46] and ChEMBL[47,48] databases, which feature purchasable molecules

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

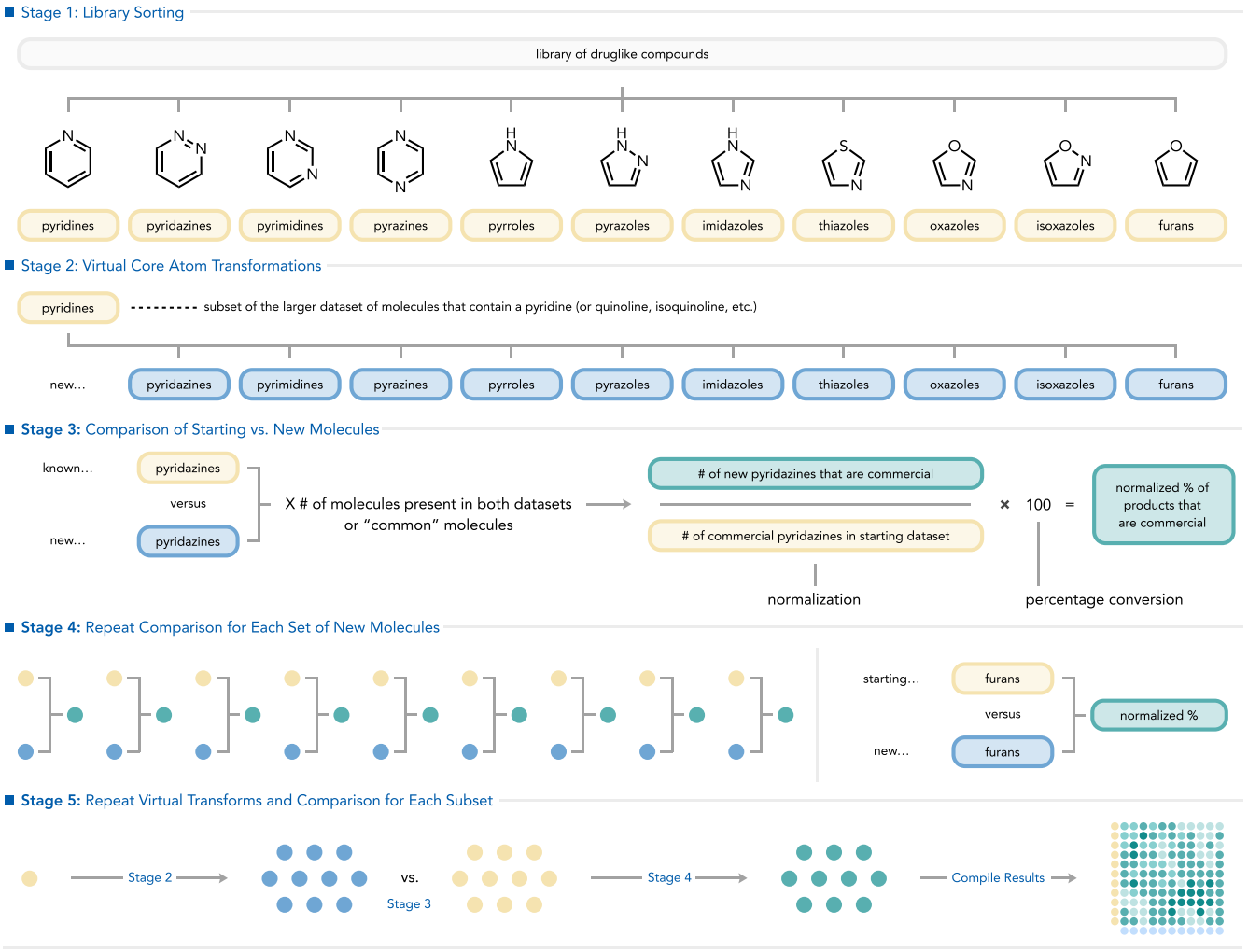


Figure 2. Workflow summary for sorting the large data set, performing heterocycle-swapping reactions, and counting the number of known compounds in the new products.

designed for screening against biological targets. Notably, the ChEMBL database of bioactive molecules is manually curated to suit this purpose. The heteroaromatics analyzed, as well as the percentages of the total data set they represent in each database, are shown in Figure 1.

A pictorial summary of our workflow is shown in Figure 2. Using the Chemical Reaction functionality of RDKit,[49] we used the SMARTS language[50] to exchange the heteroaromatic unit of each molecule in a given subset for each of the other types of heteroaromatics listed in Figure 1. We refer to these heteroaromatic-exchanging transformations as "virtual skeletal edits", since reactions to achieve these transformations may or may not exist (for example, the conversion of pyrimidines into pyrazines is not yet reported), and the transformations are performed *in silico*. After generating data sets of the products of these virtual transformations, we compared the newly generated molecules against the same heterocycle type in the starting library of molecules to determine which of the computationally generated products are synthetically feasible (or commercially available in the cases of the Mcule and ChEMBL libraries).

RESULTS

Our initial analysis involved evaluating the number of uncharacterized compounds generated in each transformation. To describe the presence or absence of compounds in the data sets that were evaluated, we adopted a consistent terminology to avoid potential ambiguities associated with other descriptors. The terms "commercial" and "not commercial" were not appropriate, as some data sets included compounds that are not commercially available but are synthetically feasible. We also avoided the terms "known" and "unknown", since these terms are typically used to denote compounds which are or are not reported in the literature, respectively. The terms "known" and "unknown" do not indicate whether a compound is commercially available (immediately or on-demand). Instead, we defined compounds as "characterized" if they existed within any of the data sets and "uncharacterized" if they did not, providing a clear distinction relevant to our analysis.

The percentage of generated products that are new (i.e., not previously characterized in any of the data sets that were evaluated) for each virtual heteroaromatic-to-heteroaromatic transformation in the Enamine, Mcule, and ChEMBL libraries are shown in Figure 3**A**–**C**, respectively. We chose to analyze these data sets separately, since the molecules in each database

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

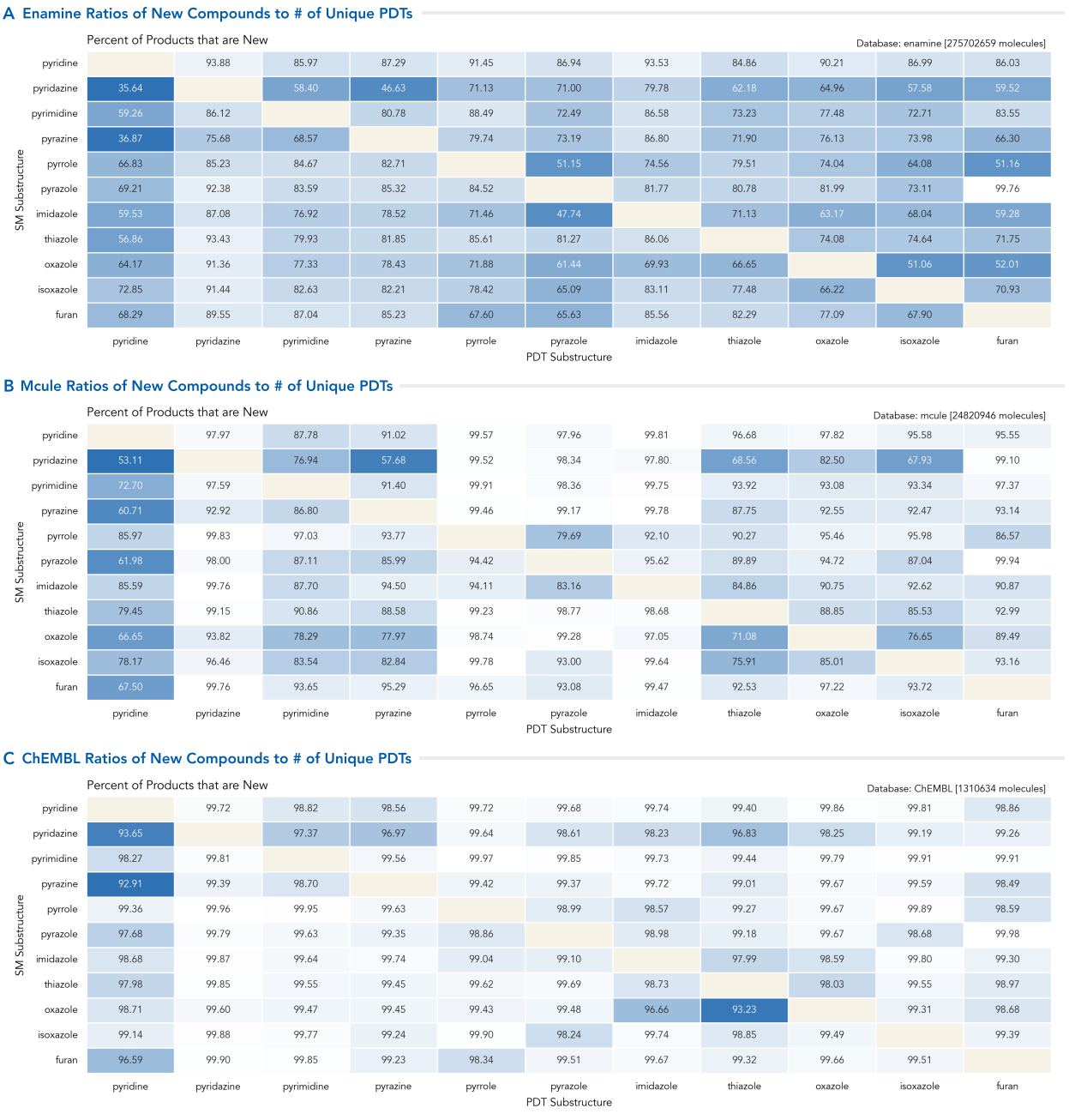


Figure 3. Percentages of the products generated for each skeletal transformation that are new or previously uncharacterized molecules for the (A) Enamine, (B) Mcule, and (C) ChEMBL data sets.

represent different levels of commercial availability. For example, the Enamine REAL database contains "synthetically feasible" or "make-on-demand" molecules, whereas the Mcule and ChEMBL databases contain existing, commercial compounds, with the latter being a manually curated database of bioactive compounds. The numerical values represent the percentage of computationally generated products that are new in relation to the starting data set. This metric identifies which transformations, if developed, would generate the greatest number of uncharacterized compounds.

Since established methods for heteroaromatic skeletal editing generally rely on the inherent functionality of a heteroatom in the ring, skeletal transformations typically occur at positions adjacent to these reactive heteroatoms.[30−38,43,51,52] Therefore, we chose to effect each virtual skeletal editing transformation by only modifying the atoms directly adjacent to the relevant heteroatom in each of the heteroaromatic units. For example, pyridines were converted to pyrroles by removing the carbon at position 2 (i.e., C2). For heterocycles with more than one heteroatom, the heteroatom with the lower atom number in the numbering of the heterocycle was designated as the priority to eliminate ambiguity. For example, the virtual pyrimidine-to-imidazole conversion is accomplished by removing C6 of the pyrimidine, which is adjacent to N1 (instead of removing C4, which is adjacent to N3).

On average, transformations which convert any of the other heteroaromatics to pyridazines generate the greatest percentage of uncharacterized compounds. This result underscores the

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

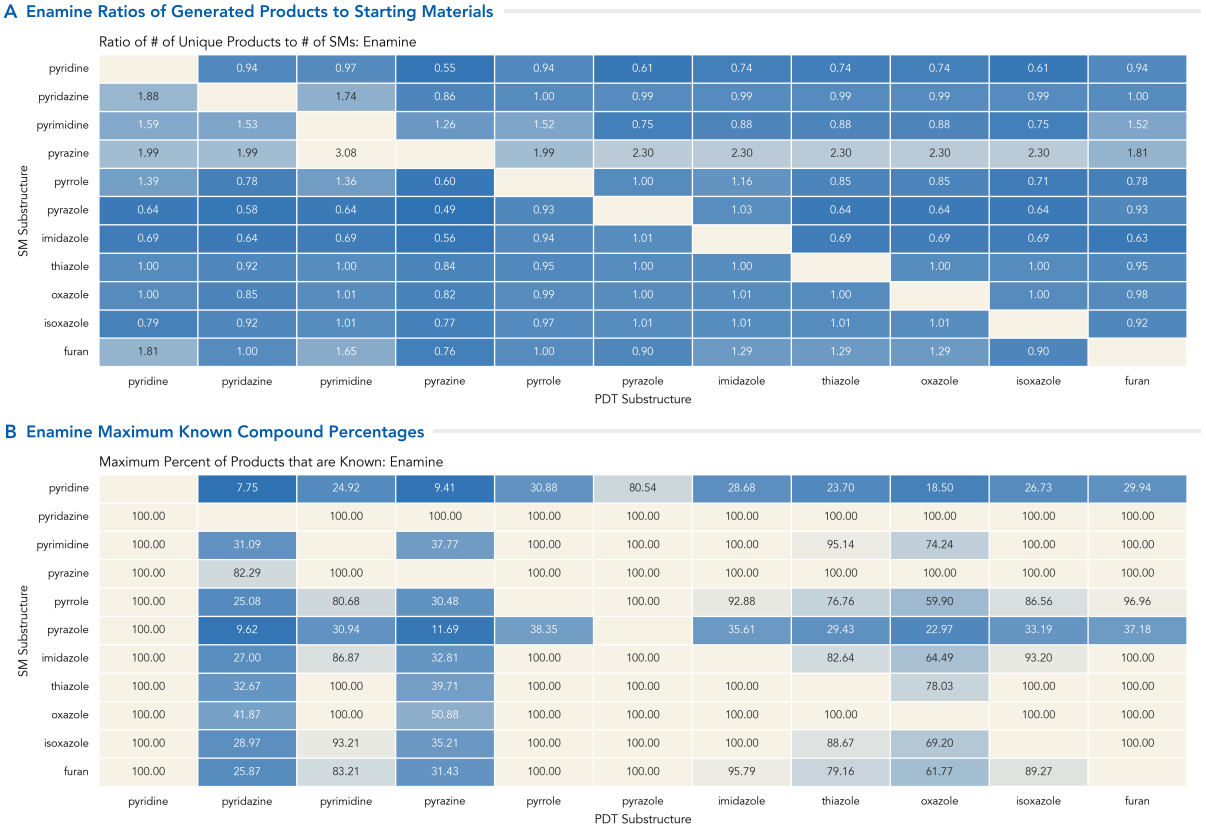


Figure 4. (A) The ratios of the number of products generated to the number of starting molecules for each transformation in the Enamine data set. (B) The maximum possible percentage of generated products that are uncharacterized for each transformation (i.e., the ratio of the known comparison molecules to the starting molecules).

fact that pyridazines are generally difficult to access and are only sparsely represented in pharmaceutical libraries. The difficulty associated with the synthesis of pyridazines is reflected in the relatively low number of commercially available pyridazine-containing molecules, leading to the high number of new compounds generated in these transformations. Since pyridazines are poorly represented in FDA-approved drugs (see the low prevalence of pyridazines in Figure 1), the high degree to which transformations generate uncharacterized pyridazine heterocycles should be considered in this context.[2,3,53]

In general, skeletal edits of pyridine give rise to the greatest percentages of uncharacterized compounds. In contrast, transformations that convert other heteroaromatics into pyridines give the lowest percentage of uncharacterized compounds. Additionally, both ring-expansions (e.g., pyrrole-to-pyrimidine) and ring-contractions (e.g., pyridine-to-pyrrole) of nitrogen heteroaromatics lead to high percentages (>80%) of previously uncharacterized structures. However, the percentage of the generated products that are uncharacterized is generally higher for ring-contracting transformations than for ring-expanding variants. Overall, this trend is consistent for each of the data sets; however, the Mcule and ChEMBL databases are much smaller than the Enamine database. Therefore, when using the Mcule or ChEMBL databases, transformations that generate heteroaromatics that are less represented in these databases are likely to be skewed toward a higher percentage of new compounds because there are fewer compounds represented in these databases. As such, the results from analysis of the Enamine data set should be prioritized to obtain a relatively unbiased perspective. Given the differences in the total numbers of molecules present in each data set, the discrepancy between ring contractions and ring expansions is pronounced in the Mcule data set, where contractions of pyridines, pyridazines, pyrimidines, and pyrazines into pyrroles, pyrazoles, and imidazoles, respectively, give rise to high percentages (>95%) of uncharacterized compounds as shown in Figure 1B.

While this initial analysis provided a metric for the transformations that would give the greatest proportion of new products relative to the total number of products generated, it cannot be used alone to judge the value of each skeletal edit. Specifically, because some transformations may be accomplished by inserting, removing, or changing the identity of the same type of atom at different positions in the ring, more than one product can be generated for each starting molecule for some transformations. This overemphasizes the potential benefit of transformations that generate more than one product molecule for each starting compound. For example, a 2-substituted pyridine can be converted into two different pyrimidines by exchanging either C3 or C5 for a nitrogen atom. The same is true of the ring contraction of pyridines to pyrroles, which could be accomplished by removing either C2 or C6 in the pyridine ring.

However, even limiting virtual transformations to ring atoms adjacent to heteroatoms, some transformations afford more than one product, on average, per starting material. The ratio of unique products generated to the number of starting

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

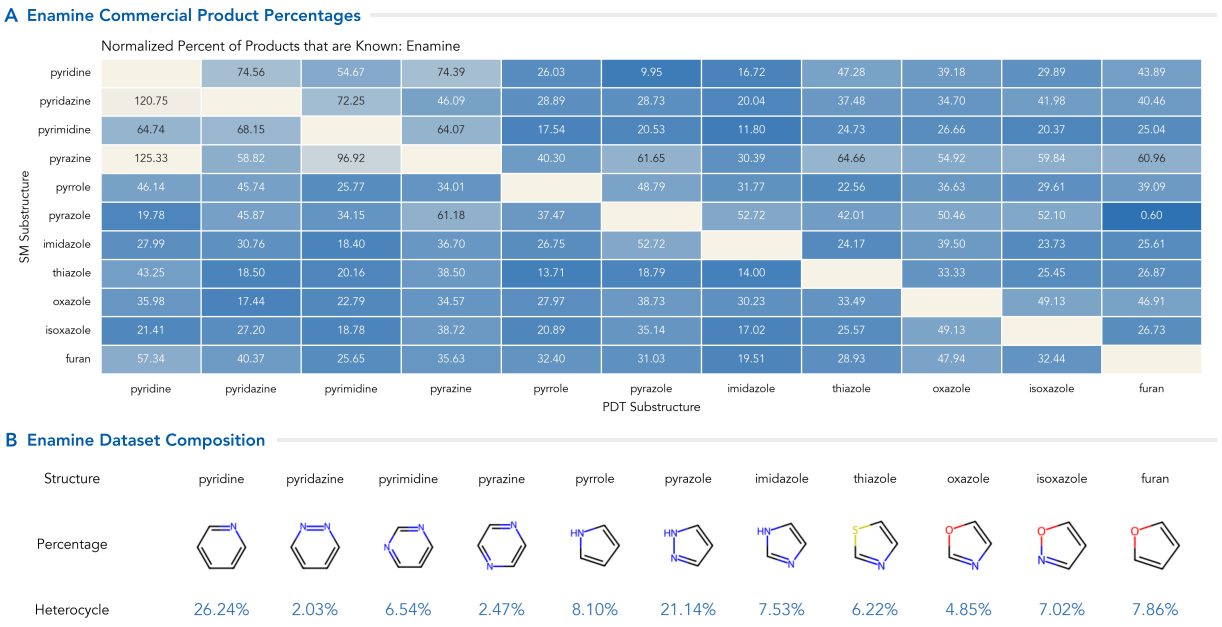


Figure 5. (A) Results from converting each starting heterocycle type into each other type of heterocycle and analyzing the number of known compounds generated in the Enamine 350/3 lead-like library. (B) Enamine data set composition.

molecules for each transformation was determined as shown in the heatmap in Figure 4A for the Enamine database. This increases the number of both characterized and uncharacterized products generated in these transformations. However, the increase in the number of uncharacterized products generated is generally higher than the increase in the number of characterized products in these cases. As such, the number of uncharacterized compounds a transformation generates can be a misleading metric. With this in mind, we sought to identify a more objective metric of the value of different virtual skeletal edits.

Toward this goal, we concluded that a metric derived from the number of *characterized* compounds generated through virtual skeletal edits would be more useful for analyzing the potential benefit of each transformation. This metric is less affected by transformations that generate more products than there are starting materials. However, the absolute number of characterized products a transformation generates cannot be used directly, since this absolute number is limited by the total number of characterized molecules of the same heterocycle type in the starting data set. As such, we cannot derive an accurate representation of the percentage of the products a transformation generates that are characterized by subtracting the percentage of those products that are *uncharacterized* from 100%.

For example, assuming 1000 pyridines exist in the starting data set, and 1000 pyrroles are generated from the 1000 pyridines, if there are only 500 pyrroles in the starting data set, a maximum of 500 of the 1000 generated pyrroles would be designated as characterized compounds. In this case, where the total number of products exceeds the number of characterized compounds in the starting data set, the number of characterized products is normalized to the total number of characterized molecules containing the same heterocycle type in the starting data set. Additionally, the opposite case is possible, where fewer molecules of a certain heterocycle type are generated than exist in the starting data set. For example, assume 500 pyrimidines exist in the starting data set and 500 imidazoles are generated from those 500 pyrimidines. If 1000 compounds characterized as imidazoles exist in the starting data set, the number of starting pyrimidines becomes limiting. If 100 of the 500 generated imidazoles are characterized in the starting data set, we normalize the 100 characterized imidazole products to the 500 starting pyrimidines. Both normalization tactics lead to a limit of 1.00 for each transformation (see Figure 4B for normalization of the Enamine 350/3 database). Importantly, because of these normalization steps, the products a transformation generates that are previously characterized (expressed as a percentage) is **not** the difference between the percentage of those products that are uncharacterized and 100%.

The results obtained by normalizing the number of characterized products in each transformation are shown in Figure 5A. A *smaller* percentage (darker blue) signifies that fewer of the generated compounds already exist, indicating more impactful skeletal edits. This analysis reveals that the pyrazole-to-furan conversion (tandem nitrogen-to-carbon and nitrogen-to-oxygen atom transmutations) is the transformation with the greatest potential to access the largest number of uncharacterized compounds. Other potential transformations that would be highly impactful include the pyrazole-to-pyridazine and thiazole-to-pyridazine transformations.

Overall, transformations involving multiple-atom changes to the heteroaromatic core give, on average, the lowest percentages of characterized products (Figure 5) and larger percentages of novel compounds (Figure 3). Therefore, these multiple-atom transformations would be designated as highly impactful in providing access to novel chemical space. However, single-atom changes to the heteroaromatic core are logical from a retrosynthetic standpoint, and the precision associated with single-atom skeletal edits make them highly valuable in the context of drug discovery.[6]

A modified heatmap containing only single-atom changes is shown in Figure 6. If we only analyze transformations involving

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

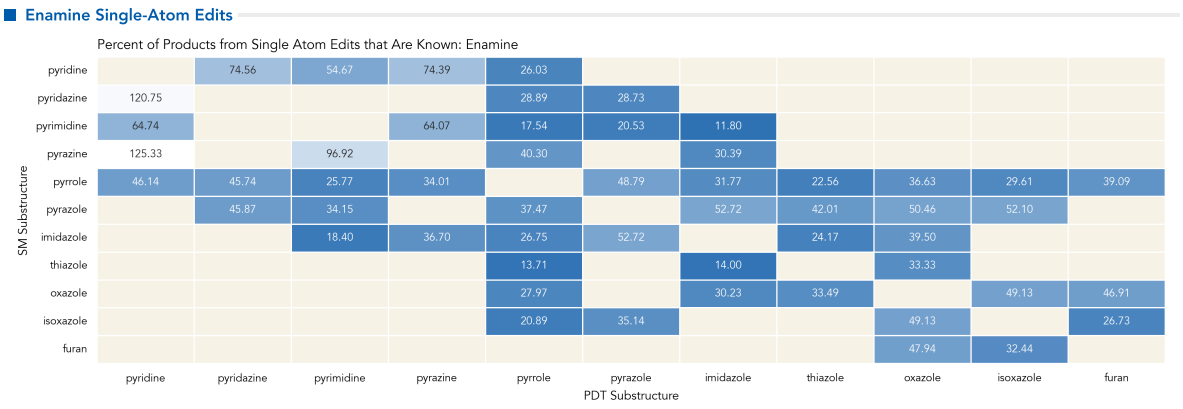


Figure 6. Truncated results for the Enamine library including only single atom transformations.

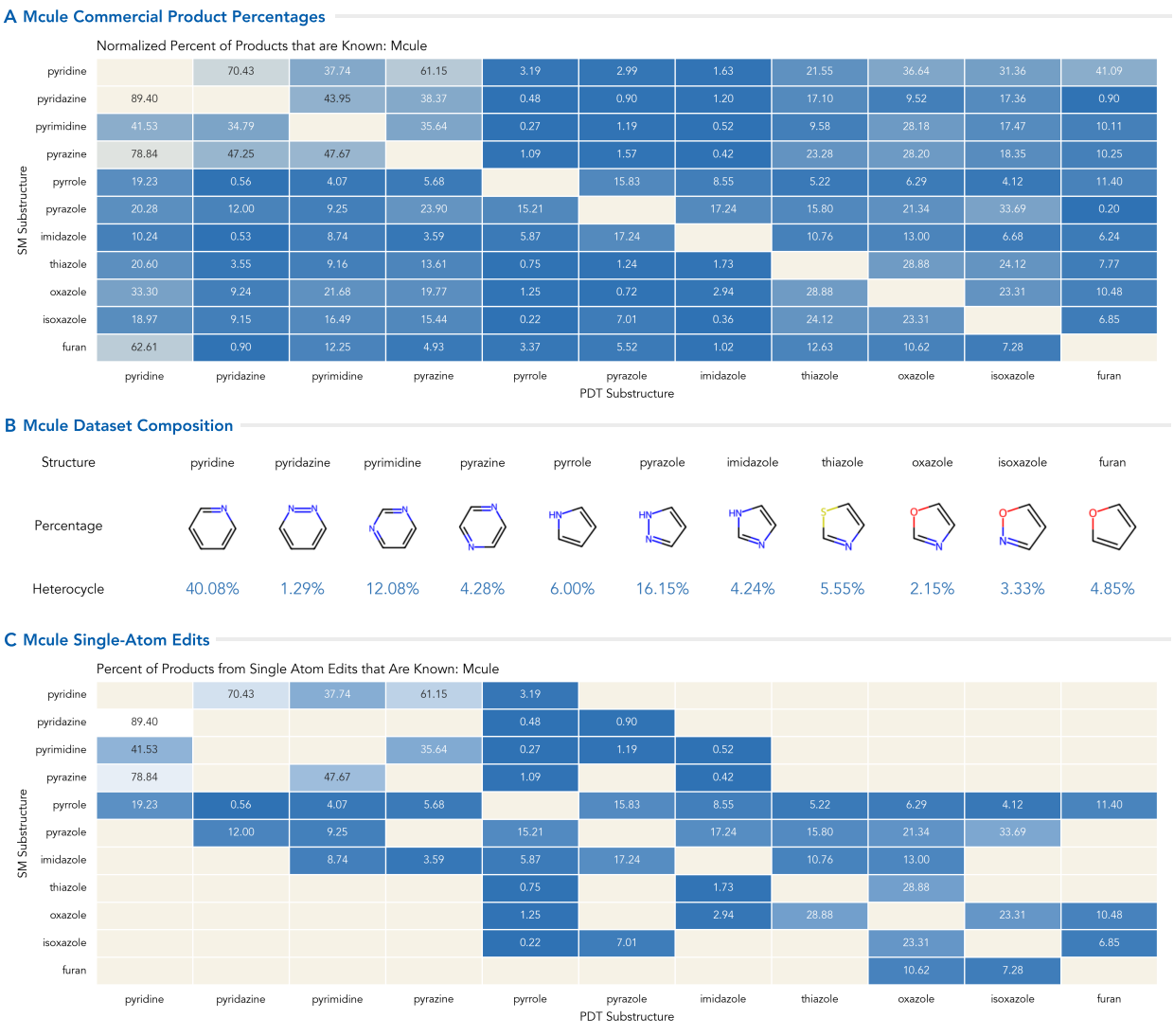


Figure 7. Normalized percentages of the products generated in each transformation that are **commercial** for the (A) full Mcule data set and (C) data set of only single-atom edits. The overall data set composition is shown in (B).

single-atom changes to the starting heteroaromatic unit, the pyrimidine-to-imidazole conversion (a single *carbon* deletion) emerges as the leading transformation for accessing novel chemical space (only 11.8% characterized). Interestingly, the opposite of this transformation—the imidazole-to-pyrimidine ring expansion—follows closely, with only 18% of the products having been previously characterized. The analogous pyrimidine-to-pyrrole transformation (a *nitrogen* deletion from the

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

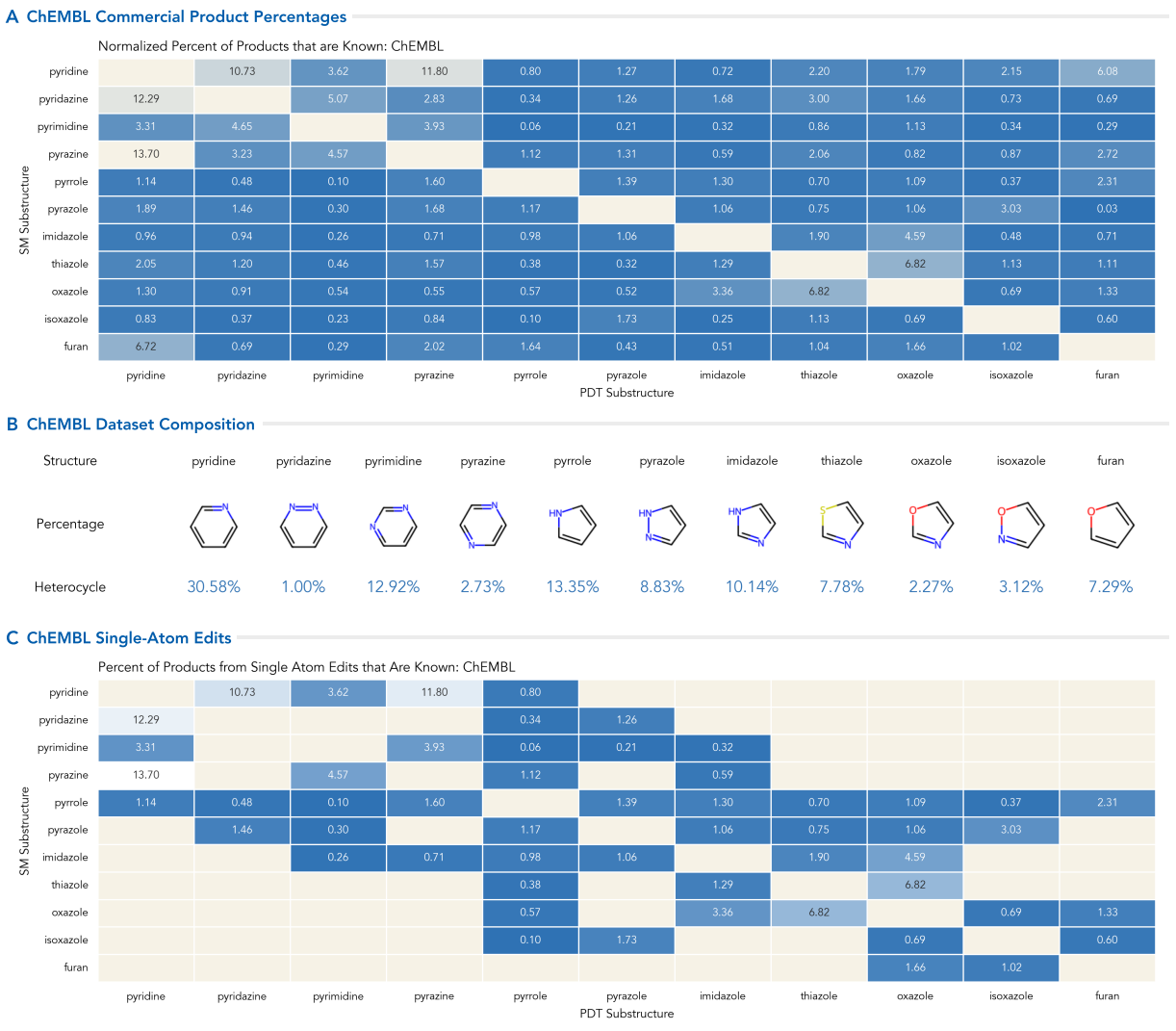


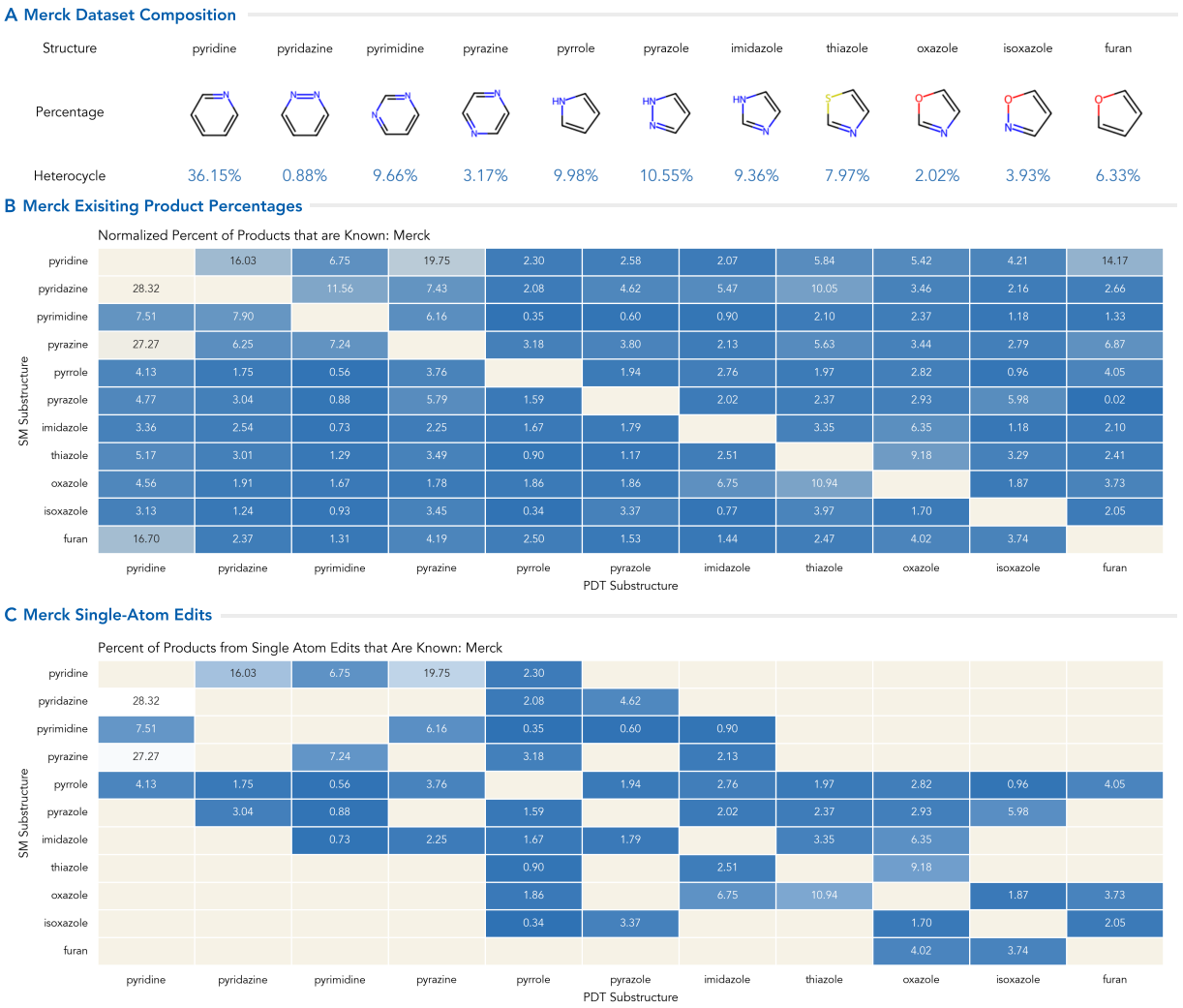
Figure 8. Normalized percentages of the products generated in each transformation that are **commercial** for the (A) full ChEMBL data set and (C) data set of only single-atom edits. The overall data set composition is shown in (B).

pyrimidine core) gives similarly few characterized products (17%). As such, the development of new methodologies for these deceptively simple transformations would most meaningfully impact late-stage access to novel chemical space. The result of this analysis, taken together with the large percentage of new compounds generated through these transformations (shown in Figure 3) highlights the importance of continued development of skeletal editing methods focused on the pyridine and pyrimidine heterocycles. Finally, atom *transmutations* of the thiazole heterocycle to give pyrroles or imidazoles by either sulfur-to-carbon or sulfur-to-nitrogen exchanges also give low percentages (13% and 14%, respectively) of characterized products and would therefore be impactful. More generally, both ring contracting and ring expanding transformations for simple nitrogen-containing heteroaromatics lead to relatively few characterized compounds (9−34%). As mentioned previously, ring contracting and expanding transformations of these types of heteroaromatics give products with the highest percentages of uncharacterized compounds (Figure 3).

In practice, there are fewer reported skeletal edits of five-membered heteroaromatics relative to those reported for larger heteroaromatic rings.[54−56] Most transformations of five-membered heteroaromatics rely on the reactivity of carbene and nitrene species with the weak π-bonds in azole and diazole scaffolds to achieve ring expansions.[30,31,33] According to our analysis, the virtual conversion of five-membered heteroaromatics into their expanded six-membered ring counterparts generally leads to a low percentage (18−34%) of commercially available compounds—indicative of high value transformations. However, according to the data presented in Figure 6, ring-contracting transformations of six-membered heteroaromatics give products with even lower percentages (11−28%) of characterized compounds. Given the general lack of ring-contracted analogues of the six-membered heterocycles in the Enamine 350/3 library,[45] there is clearly a need for new ring-contracting reactions of six-membered heteroaromatics.

Though we focused our analysis on the Enamine 350/3 database (*vide supra*), we sought to ascertain whether the trends that emerged are consistent with those from the smaller Mcule database presented in Figure 1. Despite the difference in size (by orders of magnitude) between the Enamine and Mcule databases, the same trends emerged in our analysis. The Mcule (Figure 7) database also indicates that the same trans-

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX



D Lowest-Percentage Transformations: Merck

Merck Dataset	Five Lowest % Characterized Transformations Overall					
	Full Dataset			Single Atom Skeletal Edits		
	SM	PDT	% Characterized	SM	PDT	% Characterized
	pyrazole	furan	0.02	isoxazole	pyrrole	0.34
	isoxazole	pyrrole	0.34	pyrimidine	pyrrole	0.35
	pyrimidine	pyrrole	0.35	pyrrole	pyrimidine	0.56
	pyrrole	pyrimidine	0.56	pyrimidine	pyrazole	0.60
	pyrimidine	pyrazole	0.60	imidazole	pyrimidine	0.73

Figure 9. (A) Overall composition of the Merck data set. (B) Normalized percentages of the products generated in each transformation that **exist** for the full Merck data set and (C) data set of only single-atom edits. (D) Transformations which give the five-lowest percentages of existing products in the Merck data set.

formations identified in our analysis of the Enamine database give the lowest percentages of commercial products. For example, the pyrazole-to-furan conversion gives products with the lowest percentages of commercial compounds in both databases (0.20%). Similarly, other multiple-atom transformations, such as the isoxazole-to-imidazole, oxazole-to-pyrazole, and furan-to-imidazole virtual skeletal edits, give lower percentages (0.36%−1.02%) of commercial products. However, the bias toward these multiple- atom transformations over single-atom skeletal edits is not as pronounced as in the Enamine database.

When we limit this analysis to single atom edits in the Mcule data set, the pyrimidine-to-pyrrole and pyrimidine-to-imidazole transformations give similarly low percentages of commercial compounds (0.27% and 0.52%, respectively). The same thiazole-to-pyrrole transformation identified in our analysis of the Enamine database to be of high value also emerges (giving only 0.75% commercial products). However, several additional transformations, such as the pyrazine-to-imidazole conversion (a *carbon* deletion giving 0.42% characterized products) and the isoxazole-to-pyrrole conversion (an *oxygen* transmutation giving 0.22% characterized products) emerge. The low percentages (0.22% to 2.94%) of commercial compounds that emerge from single-atom transformations of five-membered heteroaromatics further highlights the need for

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

A Enamine Top Transformations Per Heterocycle

	Lowest %Characterized Transformations for Each Heterocycle					
	Full Dataset			Truncated, Single Atom Edits		
	SM	PDT	% Characterized	SM	PDT	% Characterized
Enamine 350/3 Library	pyrazole	furan	0.60	pyrimidine	imidazole	11.80
	pyridine	pyrazole	9.95	thiazole	pyrrole	13.71
	pyrimidine	imidazole	11.80	imidazole	pyrimidine	18.40
	thiazole	pyrrole	13.71	isoxazole	pyrrole	20.89
	isoxazole	imidazole	17.02	pyrrole	thiazole	22.56
	oxazole	pyridazine	17.44	pyridine	pyrrole	26.03
	imidazole	pyrimidine	18.40	oxazole	pyrrole	27.97
	furan	imidazole	19.51	pyridazine	pyrazole	28.73
	pyridazine	imidazole	20.04	pyrazine	imidazole	30.39
	pyrrole	thiazole	22.56	furan	isoxazole	32.44
	pyrazine	imidazole	30.39	pyrazole	pyrimidine	34.15

B Enamine Top Ten Tranformations Overall

	Ten Lowest %Characterized Transformations Overall					
	Full Dataset			Truncated, Single Atom Edits		
	SM	PDT	% Characterized	SM	PDT	% Characterized
Enamine 350/3 Library	pyrazole	furan	0.60	pyrimidine	imidazole	11.80
	pyridine	pyrazole	9.95	thiazole	pyrrole	13.71
	pyrimidine	imidazole	11.80	thiazole	imidazole	14.00
	thiazole	pyrrole	13.71	pyrimidine	pyrrole	17.54
	thiazole	imidazole	14.00	imidazole	pyrimidine	18.40
	pyridine	imidazole	16.72	pyrimidine	pyrazole	20.53
	isoxazole	imidazole	17.02	isoxazole	pyrrole	20.89
	oxazole	pyridazine	17.44	pyrrole	thiazole	22.56
	pyrimidine	pyrrole	17.54	imidazole	thiazole	24.17
	imidazole	pyrimidine	18.40	pyrrole	pyrimidine	25.77

Figure 10. (A) Transformations for each heterocycle class that give the lowest percentage of previously characterized compounds and (B) the transformations overall that give products with the ten-lowest percentages of characterized compounds.

novel methodologies for single-atom point mutations of five-membered heteroaromatics.

When we analyze the ChEMBL database (the smallest of the databases presented in Figure 1), the pyrimidine-to-imidazole and pyrimidine-to-pyrrole transformations that emerged in the analysis of the Enamine 350/3 and Mcule libraries give similarly low percentages of characterized compounds (0.06% and 0.32%, respectively), once again indicative of their high value. These results are summarized in Figure 8. However, in this case, the pyrrole-to-pyrimidine transformation gives a remarkably small percentage of characterized products (0.10%). The pyridine- and pyrimidine-to-pyrrole transformations give fewer characterized compounds in this data set as well. Importantly, the pyrrole-to-pyrimidine conversion is the first transformation identified by this analysis that is, at least in part, reported in the chemical literature (for benzo-fused pyrroles, i.e., indoles to quinazolines).[33] The pyrimidine-to-pyrazole transformation, which has also been reported,[34] also gives a relatively low percentage of commercial products (0.21%) and is therefore an important transformation to add to the arsenal of methods for drug discovery. Much like the Enamine and Mcule data sets, several single-atom transformations of five-membered heteroaromatics featured in the ChEMBL data set, such as the isoxazole-to-pyrrole and thiazole-to-pyrrole conversion, give rise to very small percentages of commercially available compounds (0.10% and 0.38%, respectively).

Finally, we applied this analysis to a library of existing molecules proprietary to Merck & Co., Inc. The results of this analysis are summarized in Figure 9. The composition of this proprietary molecule database is summarized in Figure 9A and the percentage of the products generated in each transformation that are existing compounds for the Merck data set are shown in Figure 9B. This data is sorted for single-atom edits in Figure 9C, and a table of the five transformations that give the lowest percentages of existing molecules (i.e., the most potentially impactful skeletal edits) is shown in Figure 9D. This analysis reveals that the same virtual skeletal edits identified in the analyses of the other databases presented in the previous sections give the lowest percentages of existing products. For example, the pyrazole-to-furan conversion only gives 0.02% existing products. An interesting atom transmutation of a five-membered heteroaromatic—the isoxazole-to-pyrrole transformation—gives the lowest percentage of existing products out of the single-atom skeletal edits (0.34%). The same pyrimidine-to-pyrrole and pyrrole-to-pyrimidine transformations identified from the other data sets that were analyzed follow closely, giving 0.35% and 0.56% existing products, respectively. Likewise, the pyrimidine-to-pyrazole and imidazole-to-pyrimidine transformations also give low percentages of existing products.

Figure 10A lists the transformations that give the lowest percentage of characterized products for each starting heterocycle type for the Enamine database. In this table,

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

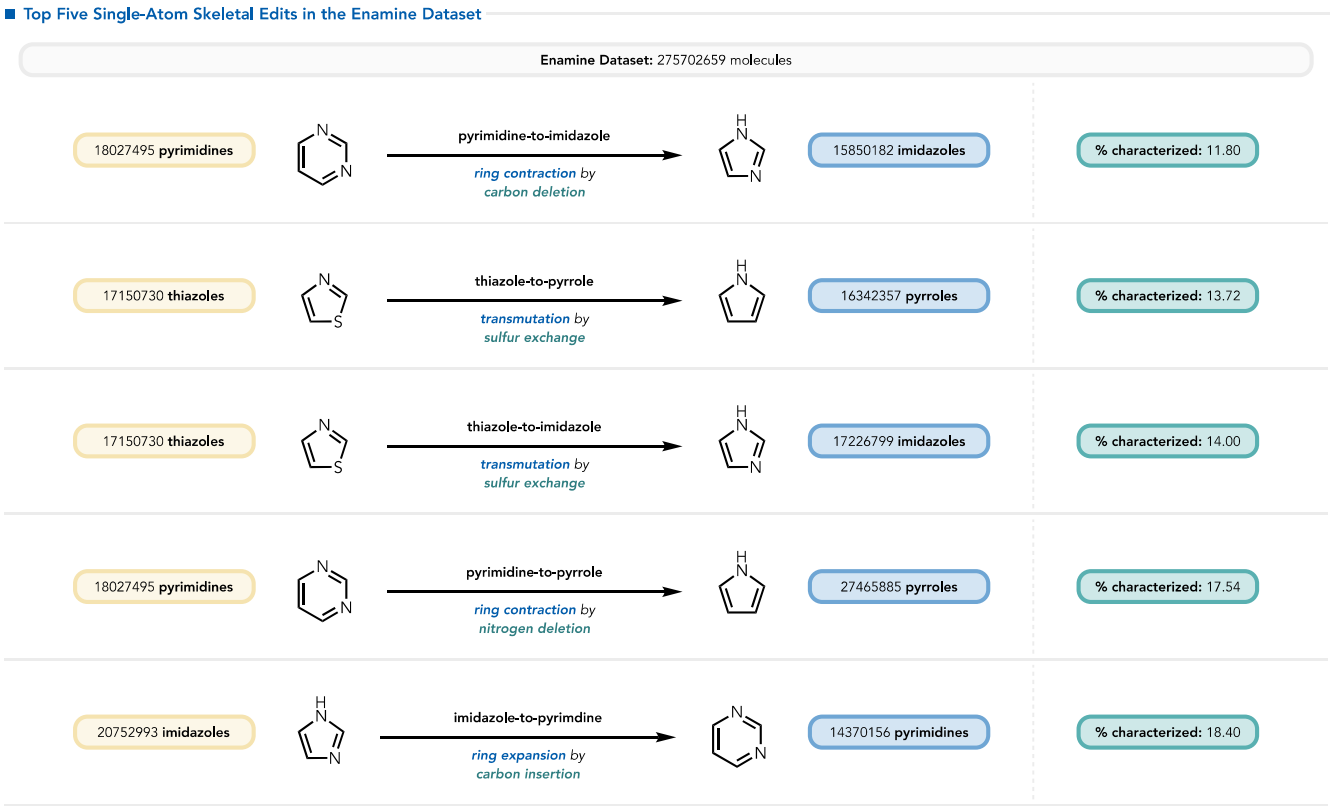


Figure 11. Graphic representation of the single-atom transformations with the five-lowest percentage of characterized, synthetically feasible products in the Enamine database.

each of the analyzed starting heteroaromatic substrates are represented exactly once. Figure 10B lists the ten transformations in the data set with the lowest percentages of characterized products regardless of the starting heterocycle type. In the tables, the transformations are listed in the order of increasing percentage of products that are characterized. The high potential impact of a pyrazole-to-furan transformation is evidenced by the lowest percentage (0.60%) of characterized compounds. In Figure 10B, some starting heterocycle types are listed more than once. Transformations involving ring contractions of the pyrimidine heterocycle (to give imidazoles, pyrroles, and pyrazoles) emerge as highly impactful. Three other heterocycle types—thiazoles, imidazoles, and pyrroles—appear more than once in the top ten transformations. However, pyrimidine is the only starting substructure that appears three times in the top ten transformations of the single-atom edits, highlighting the potential for single-atom skeletal edits of pyrimidines to greatly expand accessible heteroaromatic chemical space. Accordingly, McNally and co-workers recently reported a powerful strategy for the diversification of the pyrimidine scaffold into several types of other heterocyclic substructures using skeletal editing approaches.[35]

In this representation of the data, the potential value of methods to convert a single starting heteroaromatic into several different heterocycles is clear. However, our analysis suggests an orthogonal direction that could be equally valuable would be to focus on accessing previously inaccessible products, such as pyrroles, from a broader range of heterocycles rather than accessing a wide range of heterocycles from one starting substructure. For example, conversion of different types of starting substructures into pyrroles generally leads to the lowest percentages of characterized compounds, positioning pyrroles as an attractive target for continued development in skeletal editing. In fact, Park and co-workers recently reported a strategy for the skeletal editing of furans to pyrroles, an important development with this analysis in mind.[39] Finally, the five single-atom transformations with the lowest percentage of characterized products (i.e., of high value) from the Enamine 350/3 database are graphically represented in Figure 11.

CONCLUSION

Across large chemical databases including Enamine 350/3, Mcule, and ChEMBL, consistent trends emerge regarding which skeletal editing transformations stand to most significantly increase commercially accessible chemical space for heteroaromatic compounds. Despite variations in database size and composition, virtual ring contractions of heteroaromatics often give fewer characterized or commercially available products as compared to the range of products accessible using ring expansion or atom transmutation skeletal editing transformations. This quantitative analysis highlights the need for new synthetic methodologies that can efficiently modify heteroaromatics such as pyridines, pyrimidines, and pyrroles through precise single-atom changes in the cores of these molecules. Transformations of this type would significantly expand the structural diversity of the currently accessible chemical space for pharmaceutically relevant compounds and other applications. According to this analysis, skeletal edits of pyrimidines, thiazoles, and pyrroles would most significantly enhance access to novel chemical space. Additionally, pyrroles emerged as a valuable target product for skeletal editing reactions of several heteroaromatics.

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

Overall, we also found that changes involving multiple atoms in the core structure often yield a greater number of previously uncharacterized compounds when compared to single-atom modifications. Transformations like pyrazole-to-furan are illustrative. The development of new methodologies to effect transformations of this type would enable access to areas of chemical space that are currently underexplored. Intuitive single-atom alterations, such as the conversions of pyridine-to-pyrrole or pyrimidine-to-pyrrole, tend to produce a slightly higher percentage of characterized compounds and would therefore be judged as of lesser value. Nonetheless, these transformations are extremely valuable in discovery campaigns due to the precision afforded by single-atom edits in the context of late-stage diversification. Importantly, the transformations identified herein would enable interconversions between many heteroaromatics that are prevalent in FDA approved drugs, especially in the context of the aforementioned ring expansion methodologies.[2,3,30,31,33]

Synthetic methods for transformations that are complementary to those identified as potentially impactful in this report are continuing to emerge.[30,33] Considered together, the now quantified potential impact of these transformations highlights the necessity of continued development of new skeletal editing reactions. The analysis presented here may serve as a useful starting point for the community to use quantitative metrics to quantify the impact of different core-atom transformations of heteroaromatics in compound diversification campaigns.

■ ASSOCIATED CONTENT

Data Availability Statement

The detailed workflow and data used to produce the reported results can be found online at https://github.com/ljkaras/skeletal_editing/releases/tag/v1.0 (DOI: 10.5281/zenodo.14077417).

■ AUTHOR INFORMATION

Corresponding Authors

Charles S. Yeung − *Discovery Chemistry, Merck & Co., Inc., Boston, Massachusetts 02115, United States;* orcid.org/0000-0001-7320-7190; Email: charles.yeung@merck.com

Matthew S. Sigman − *Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, United States;* orcid.org/0000-0002-5746-8830; Email: matt.sigman@utah.edu

Richmond Sarpong − *Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, United States;* orcid.org/0000-0002-0028-6323; Email: rsarpong@berkeley.edu

Authors

G. Logan Bartholomew − *Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, United States;* orcid.org/0000-0002-3023-4460

Lucas J. Karas − *Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, United States*

Reilly M. Eason − *Modeling & Informatics, Merck & Co., Inc., Boston, Massachusetts 02115, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.4c02839

Author Contributions

†G.L.B. and L.J.K. contributed equally.

Funding

R.S. and M.S.S. are grateful to the National Science Foundation (NSF) under the Center for Computer Assisted Synthesis (CCAS, NSF CHE 2202693) for financial support. R.S. also thanks Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA (MSD) for support. G.L.B. (NSF 2021294420) thanks the National Science Foundation for a graduate fellowship.

Notes

The authors declare the following competing financial interest(s): R.S. is a paid consultant for Merck & Co., Inc.

Biographies

G. Logan Bartholomew received his PhD from the University of California, Berkeley, in 2024 with Professor Richmond Sarpong. Logan joined Discovery Process Chemistry at Merck & Co., Inc. in South San Francisco as a Senior Scientist in March of 2025.

Lucas J. Karas earned his PhD in 2021 from the University of Houston under the guidance of Professor Judy Wu. He then completed postdoctoral training with Professor Matthew Sigman at the University of Utah. Since 2024, Lucas has been working as a Cheminformatics Scientist at Terray Therapeutics.

Reilly M. Eason received his PhD in 2016 from the University of Missouri. He has been a member of the Cheminformatics group at Merck & Co., Inc. since 2022.

Matthew S. Sigman has been a professor at the University of Utah since 1999. His work focuses on integrating data science with synthetic chemistry.

Charles S. Yeung received his PhD at the University of Toronto with Professor Vy Dong. Following postdoctoral training with Professor Eric Jacobsen, he has been part of Discovery Chemistry at Merck & Co., Inc., Boston., where he is currently a Principal Scientist.

Richmond Sarpong received his PhD from Princeton University in 2001 with Professor Martin F. Semmelhack. Following postdoctoral training with Professor Brian M. Stoltz at Caltech, he started his independent career in 2004 at UC Berkeley where he remains. His research focuses on the development of methods and strategies for the preparation of complex molecules including natural products.

■ ACKNOWLEDGMENTS

We thank Paul P. Sinclair (UC Berkeley) and Sojung F. Kim (UC Berkeley) for thorough reviews of the manuscript and helpful discussions. The support and resources from the Center for High Performance Computing at the University of Utah are gratefully acknowledged. We thank Michael Altman and Colin Lam (Merck & Co., Inc.) for thorough reviews of the manuscript and helpful discussions.

■ ABBREVIATIONS USED

FDA, Food and Drug Administration; REAL, REadily AccessibLe; ChEMBL, Chemistry European Molecular Biology Laboratory; RDKit, Rational Discovery Kit; SMARTS, SMILES ARbitrary Target Specification; SMILES, Simplified Molecular Input Line Entry System

■ REFERENCES

(1) Jampilek, J. Heterocycles in Medicinal Chemistry. *Molecules* **2019**, *24*, 3839.

(2) Vitaku, E.; Smith, D. T.; Njardarson, J. T. Analysis of the Structural Diversity, Substitution Patterns, and Frequency of Nitrogen

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

Heterocycles among U.S. FDA Approved Pharmaceuticals: Miniperspective. *J. Med. Chem.* **2014**, *57*, 10257−10274.
(3) Shearer, J.; Castro, J. L.; Lawson, A. D. G.; MacCoss, M.; Taylor, R. D. Rings in Clinical Trials and Drugs: Present and Future. *J. Med. Chem.* **2022**, *65*, 8699−8712.
(4) Marshall, C. M.; Federice, J. G.; Bell, C. N.; Cox, P. B.; Njardarson, J. T. An Update on the Nitrogen Heterocycle Compositions and Properties of U.S. FDA-Approved Pharmaceuticals (2013−2023). *J. Med. Chem.* **2024**, *67*, 11622−11655.
(5) Campos, K. R.; Coleman, P. J.; Alvarez, J. C.; Dreher, S. D.; Garbaccio, R. M.; Terrett, N. K.; Tillyer, R. D.; Truppo, M. D.; Parmee, E. R. The Importance of Synthetic Chemistry in the Pharmaceutical Industry. *Science* **2019**, *363*, 0805.
(6) Jurczyk, J.; Woo, J.; Kim, S. F.; Dherange, B. D.; Sarpong, R.; Levin, M. D. Single-Atom Logic for Heterocycle Editing. *Nat. Synth* **2022**, *1*, 352−364.
(7) Ertl, P. Cheminformatics Analysis of Organic Substituents: Identification of the Most Common Substituents, Calculation of Substituent Properties, and Automatic Identification of Drug-like Bioisosteric Groups. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 374−380.
(8) Lenci, E.; Trabocchi, A. Smart Design of Small-Molecule Libraries: When Organic Synthesis Meets Cheminformatics. *ChemBioChem.* **2019**, *20*, 1115−1123.
(9) Dreher, S. D.; Krska, S. W. Chemistry Informer Libraries: Conception, Early Experience, and Role in the Future of Cheminformatics. *Acc. Chem. Res.* **2021**, *54*, 1586−1596.
(10) Wishart, D. S. Introduction to Cheminformatics. *Current Protocols in Bioinformatics* **2016**, *53*, 14.1.1−14.1.21.
(11) Muchmore, S. W.; Edmunds, J. J.; Stewart, K. D.; Hajduk, P. J. Cheminformatic Tools for Medicinal Chemists. *J. Med. Chem.* **2010**, *53*, 4830−4841.
(12) Prada-Gracia, D.; Huerta-Yépez, S.; Moreno-Vargas, L. M. Application of Computational Methods for Anticancer Drug Discovery, Design, and Optimization. *Boletín Médico Del Hospital Infantil de México (English Edition)* **2016**, *73*, 411−423.
(13) Manly, C. J.; Louise-May, S.; Hammer, J. D. The Impact of Informatics and Computational Chemistry on Synthesis and Screening. *Drug Discovery Today* **2001**, *6*, 1101−1110.
(14) Arrieta, A.; de la Torre, M.; de Cozar, A.; Sierra, M.; Cossio, F. Computational Chemistry; A Useful Tool for the Chemical Synthesis of Complex Molecules, Heterocycles and Catalysts. *Synlett* **2013**, *24*, 535−549.
(15) Cavasotto, C. N.; Aucar, M. G.; Adler, N. S. Computational Chemistry in Drug Lead Discovery and Design. *Int. J. Quantum Chem.* **2019**, *119*, No. e25678.
(16) Zhang, X.; Chung, L. W.; Wu, Y.-D. New Mechanistic Insights on the Selectivity of Transition-Metal-Catalyzed Organic Reactions: The Role of Computational Chemistry. *Acc. Chem. Res.* **2016**, *49*, 1302−1310.
(17) Patel, P.; Wilson, A. K. Computational Chemistry Considerations in Catalysis: Regioselectivity and Metal-Ligand Dissociation. *Catal. Today* **2020**, *358*, 422−429.
(18) Nandy, A.; Adamji, H.; Kastner, D. W.; Vennelakanti, V.; Nazemi, A.; Liu, M.; Kulik, H. J. Using Computational Chemistry To Reveal Nature's Blueprints for Single-Site Catalysis of C−H Activation. *ACS Catal.* **2022**, *12*, 9281−9306.
(19) Chen, B. W. J.; Xu, L.; Mavrikakis, M. Computational Methods in Heterogeneous Catalysis. *Chem. Rev.* **2021**, *121*, 1007−1048.
(20) Lan, J.; Li, X.; Yang, Y.; Zhang, X.; Chung, L. W. New Insights and Predictions into Complex Homogeneous Reactions Enabled by Computational Chemistry in Synergy with Experiments: Isotopes and Mechanisms. *Acc. Chem. Res.* **2022**, *55*, 1109−1123.
(21) Liu, D.-H.; Pflüger, P. M.; Outlaw, A.; Lückemeier, L.; Zhang, F.; Regan, C.; Rashidi Nodeh, H.; Cernak, T.; Ma, J.; Glorius, F. Late-Stage Saturation of Drug Molecules. *J. Am. Chem. Soc.* **2024**, *146*, 11866−11875.
(22) Tehrani, K. A.; Van Syngel, K.; Boelens, M.; Contreras, J.; De Kimpe, N.; Knight, D. W. Boron(III) Bromide-Induced Ring Contraction of 3-Oxygenated Piperidines to 2-(Bromomethyl)-Pyrrolidines. *Tetrahedron Lett.* **2000**, *41*, 2507−2510.
(23) Feraldi-Xypolia, A.; Gomez Pardo, D.; Cossy, J. Ring Contraction of 3-Hydroxy-3-(Trifluoromethyl)Piperidines: Synthesis of 2-Substituted 2-(Trifluoromethyl)Pyrrolidines. *Chemistry − A European Journal* **2015**, *21*, 12876−12880.
(24) Roque, J. B.; Kuroda, Y.; Göttemann, L. T.; Sarpong, R. Deconstructive Diversification of Cyclic Amines. *Nature* **2018**, *564*, 244−248.
(25) Wang, F.; He, Y.; Tian, M.; Zhang, X.; Fan, X. Synthesis of α-Formylated N-Heterocycles and Their 1,1-Diacetates from Inactivated Cyclic Amines Involving an Oxidative Ring Contraction. *Org. Lett.* **2018**, *20*, 864−867.
(26) Kennedy, S. H.; Dherange, B. D.; Berger, K. J.; Levin, M. D. Skeletal Editing through Direct Nitrogen Deletion of Secondary Amines. *Nature* **2021**, *593*, 223−227.
(27) Jurczyk, J.; Lux, M. C.; Adpressa, D.; Kim, S. F.; Lam, Y.; Yeung, C. S.; Sarpong, R. Photomediated Ring Contraction of Saturated Heterocycles. *Science* **2021**, *373*, 1004−1012.
(28) Kim, S. F.; Schwarz, H.; Jurczyk, J.; Nebgen, B. R.; Hendricks, H.; Park, H.; Radosevich, A.; Zuerch, M. W.; Harper, K.; Lux, M. C.; Yeung, C. S.; Sarpong, R. Mechanistic Investigation, Wavelength-Dependent Reactivity, and Expanded Reactivity of N−Aryl Azacycle Photomediated Ring Contractions. *J. Am. Chem. Soc.* **2024**, *146*, 5580−5596.
(29) Amber, C.; Göttemann, L. T.; Steele, R. T.; Petitjean, T. M.; Sarpong, R. Reductive Amination of Carbonyl C−C Bonds Enables Formal Nitrogen Insertion. *J. Org. Chem.* **2024**, *89*, 17655.
(30) Dherange, B. D.; Kelly, P. Q.; Liles, J. P.; Sigman, M. S.; Levin, M. D. Carbon Atom Insertion into Pyrroles and Indoles Promoted by Chlorodiazirines. *J. Am. Chem. Soc.* **2021**, *143*, 11337−11344.
(31) Hyland, E. E.; Kelly, P. Q.; McKillop, A. M.; Dherange, B. D.; Levin, M. D. Unified Access to Pyrimidines and Quinazolines Enabled by N−N Cleaving Carbon Atom Insertion. *J. Am. Chem. Soc.* **2022**, *144*, 19258−19264.
(32) Woo, J.; Christian, A. H.; Burgess, S. A.; Jiang, Y.; Mansoor, U. F.; Levin, M. D. Scaffold Hopping by Net Photochemical Carbon Deletion of Azaarenes. *Science* **2022**, *376*, 527−532.
(33) Reisenbauer, J. C.; Green, O.; Franchino, A.; Finkelstein, P.; Morandi, B. Late-Stage Diversification of Indole Skeletons through Nitrogen Atom Insertion. *Science* **2022**, *377*, 1104−1109.
(34) Bartholomew, G. L.; Carpaneto, F.; Sarpong, R. Skeletal Editing of Pyrimidines to Pyrazoles by Formal Carbon Deletion. *J. Am. Chem. Soc.* **2022**, *144*, 22309−22315.
(35) Uhlenbruck, B. J. H.; Josephitis, C. M.; de Lescure, L.; Paton, R. S.; McNally, A. A Deconstruction−Reconstruction Strategy for Pyrimidine Diversification. *Nature* **2024**, *631*, 87−93.
(36) Bartholomew, G. L.; Kraus, S. L.; Karas, L. J.; Carpaneto, F.; Bennett, R.; Sigman, M. S.; Yeung, C. S.; Sarpong, R. ^{14}N to ^{15}N Isotopic Exchange of Nitrogen Heteroaromatics through Skeletal Editing. *J. Am. Chem. Soc.* **2024**, *146*, 2950−2958.
(37) Nguyen, H. M. H.; Thomas, D. C.; Hart, M. A.; Steenback, K. R.; Levy, J. N.; McNally, A. Synthesis of 15N-Pyridines and Higher Mass Isotopologs via Zincke Imine Intermediates. *J. Am. Chem. Soc.* **2024**, *146*, 2944−2949.
(38) Tolchin, Z. A.; Smith, J. M. 15NRORC: An Azine Labeling Protocol. *J. Am. Chem. Soc.* **2024**, *146*, 2939−2943.
(39) Kim, D.; You, J.; Lee, D. H.; Hong, H.; Kim, D.; Park, Y. Photocatalytic Furan-to-Pyrrole Conversion. *Science* **2024**, *386*, 99−105.
(40) Woo, J.; Stein, C.; Christian, A. H.; Levin, M. D. Carbon-to-Nitrogen Single-Atom Transmutation of Azaarenes. *Nature* **2023**, *623*, 77−82.
(41) Pearson, T. J.; Shimazumi, R.; Driscoll, J. L.; Dherange, B. D.; Park, D.-I.; Levin, M. D. Aromatic Nitrogen Scanning by Ipso-Selective Nitrene Internalization. *Science* **2023**, *381*, 1474−1479.
(42) Roure, B.; Alonso, M.; Lonardi, G.; Yildiz, D. B.; Buettner, C. S.; dos Santos, T.; Xu, Y.; Bossart, M.; Derdau, V.; Méndez, M.; Llaveria, J.; Ruffoni, A.; Leonori, D. Photochemical Permutation of

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX

Thiazoles, Isothiazoles and Other Azoles. *Nature* **2025**, *637*, 860−867.
(43) Cheng, Q.; Bhattacharya, D.; Haring, M.; Cao, H.; Mück-Lichtenfeld, C.; Studer, A. Skeletal Editing of Pyridines through Atom-Pair Swap from CN to CC. *Nat. Chem.* **2024**, *16*, 741−748.
(44) Bartholomew, G. L.; Kim, S. F.; Oyamada, Y.; Sbordone, F.; Carroll, J. A.; Jurczyk, J. E.; Yeung, C. S.; Barner-Kowollik, C.; Sarpong, R. Phototransposition of Indazoles to Benzimidazoles: Tautomer-Dependent Reactivity, Wavelength Dependence, and Continuous Flow Studies. *Angew. Chem. Int. Ed* **2025**, No. e202423803.
(45) Enamine. *REAL Database*; https://enamine.net/compound-collections/real-compounds/real-database.
(46) Kiss, R.; Sandor, M.; Szalai, F. A. http://mcule.com: A Public Web Service for Drug Discovery. *Journal of Cheminformatics* **2012**, *4*, No. P17.
(47) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res.* **2012**, *40*, D1100−D1107.
(48) Zdrazil, B.; Felix, E.; Hunter, F.; Manners, E. J.; Blackshaw, J.; Corbett, S.; de Veij, M.; Ioannidis, H.; Lopez, D. M.; Mosquera, J. F.; Magarinos, M. P.; Bosc, N.; Arcila, R.; Kizilören, T.; Gaulton, A.; Bento, A. P.; Adasme, M. F.; Monecke, P.; Landrum, G. A.; Leach, A. R. The ChEMBL Database in 2023: A Drug Discovery Platform Spanning Multiple Bioactivity Data Types and Time Periods. *Nucleic Acids Res.* **2024**, *52*, D1180−D1192.
(49) Landrum, G. *RDKit: Open-Source Cheminformatics*; https://www.bibsonomy.org/bibtex/28d01fceeccd6bf2486e47d7c4207b108/salotz.
(50) *Daylight Theory: SMARTS - A Language for Describing Molecular Patterns*; https://www.daylight.com/dayhtml/doc/theory/theory.smarts.html.
(51) Zhong, H.; Egger, D. T.; Gasser, V. C. M.; Finkelstein, P.; Keim, L.; Seidel, M. Z.; Trapp, N.; Morandi, B. Skeletal Metalation of Lactams through a Carbonyl-to-Nickel-Exchange Logic. *Nat. Commun.* **2023**, *14*, 5273.
(52) Feng, M.; Norlöff, M.; Guichard, B.; Kealey, S.; D'Anfray, T.; Thuéry, P.; Taran, F.; Gee, A.; Feuillastre, S.; Audisio, D. Pyridine-Based Strategies towards Nitrogen Isotope Exchange and Multiple Isotope Incorporation. *Nat. Commun.* **2024**, *15*, 6063.
(53) Bel Abed, H.; Mammoliti, O.; Bande, O.; Van Lommen, G.; Herdewijn, P. Strategy for the Synthesis of Pyridazine Heterocycles and Their Derivatives. *J. Org. Chem.* **2013**, *78*, 7845−7858.
(54) Joynson, B. W.; Ball, L. T. Skeletal Editing: Interconversion of Arenes and Heteroarenes. *Helv. Chim. Acta* **2023**, *106*, No. e202200182.
(55) Liu, Z.; Sivaguru, P.; Ning, Y.; Wu, Y.; Bi, X. Skeletal Editing of (Hetero)Arenes Using Carbenes. *Chemistry − A European Journal* **2023**, *29*, No. e202301227.
(56) Ji, P.; Duan, K.; Li, M.; Wang, Z.; Meng, X.; Zhang, Y.; Wang, W. Photochemical Dearomative Skeletal Modifications of Hetero-aromatics. *Chem. Soc. Rev.* **2024**, *53*, 6600−6624.

https://doi.org/10.1021/acs.jmedchem.4c02839
J. Med. Chem. XXXX, XXX, XXX−XXX