JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau

Perspective

The Time and Place for Nature in Drug Discovery

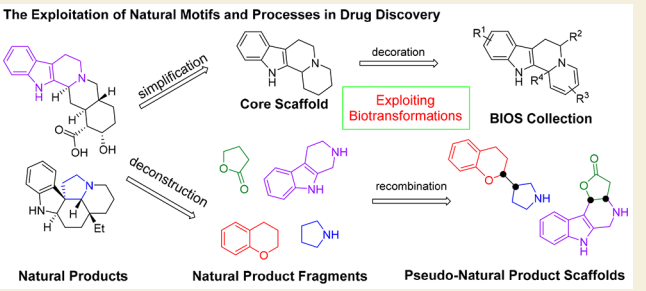
Robert J. Young,* Sabine L. Flitsch, Michael Grigalunas, Paul D. Leeson, Ronald J. Quinn, Nicholas J. Turner, and Herbert Waldmann

Cite This: *JACS Au* 2022, 2, 2400−2416

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: The case for a renewed focus on Nature in drug discovery is reviewed; not in terms of natural product screening, but how and why biomimetic molecules, especially those produced by natural processes, should deliver in the age of artificial intelligence and screening of vast collections both in vitro and in silico. The declining natural product-likeness of licensed drugs and the consequent physicochemical implications of this trend in the context of current practices are noted. To arrest these trends, the logic of seeking new bioactive agents with enhanced natural mimicry is considered; notably that molecules constructed by proteins (enzymes) are more likely to interact with other proteins (e.g., targets and transporters), a notion validated by natural products. Nature's finite number of building blocks and their interactions necessarily reduce potential numbers of structures, yet these enable expansion of chemical space with their inherent diversity of physical characteristics, pertinent to property-based design. The feasible variations on natural motifs are considered and expanded to encompass pseudo-natural products, leading to the further logical step of harnessing bioprocessing routes to access them. Together, these offer opportunities for enhancing natural mimicry, thereby bringing innovation to drug synthesis exploiting the characteristics of natural recognition processes. The potential for computational guidance to help identifying binding commonalities in the route map is a logical opportunity to enable the design of tailored molecules, with a focus on "organic/biological" rather than purely "synthetic" structures. The design and synthesis of prototype structures should pay dividends in the disposition and efficacy of the molecules, while inherently enabling greener and more sustainable manufacturing techniques.



KEYWORDS: *Physicochemical properties, absorption, transporters, pseudo-natural products, natural motifs, biotransformations, green manufacturing*

INTRODUCTION

This Perspective presents rationales and prospects for an increased focus on natural motifs and natural processes in drug discovery, bringing together concepts, observations, and opportunities to support and implement the hypothesis. Such methods are perhaps at odds with contemporary drug discovery practices, where practitioners remain heavily focused on a numbers-driven process through screening of molecules that can be made rapidly and cheaply,[1−4] contrasting to former practices of targeting designed natural product-like or biomimetic structures that often required more complex syntheses.[5] The following discussion is not about natural products per se, a subject well-covered in recent reviews,[6−8] rather the contextualisation of the importance of the features and motifs inherent to physiological molecules and bioactive exogenous natural products is presented.[9] This notion is developed to include variations on Nature's themes through pseudo-natural products and the inherent opportunities presented through an increased employment of biotransformations in synthetic approaches toward bioactive molecules.

It is evident that the propensity for "natural product" likeness[10] and natural mimicry in drug molecules is diminishing,[11] as illustrated in Figure1a, a temporal comparison of natural product likeness of approved drugs assessed by the Natural Product Scout algorithm[12] versus the year of the first disclosure of the drug. This change begins to take hold in around 1990, followed by a sharp decline after 2000. Leading up to the early 1990s, many drugs and drug families originated from "prototype" structures which were predominantly of natural origin, with fewer found from screening or "serendipity".[13,14] By 1990, a major change had begun in drug discovery with the advent of rapid primary screening employing cloned human protein targets. This, together with developments in

Received: July 26, 2022
Revised: October 6, 2022
Accepted: October 6, 2022
Published: October 14, 2022

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

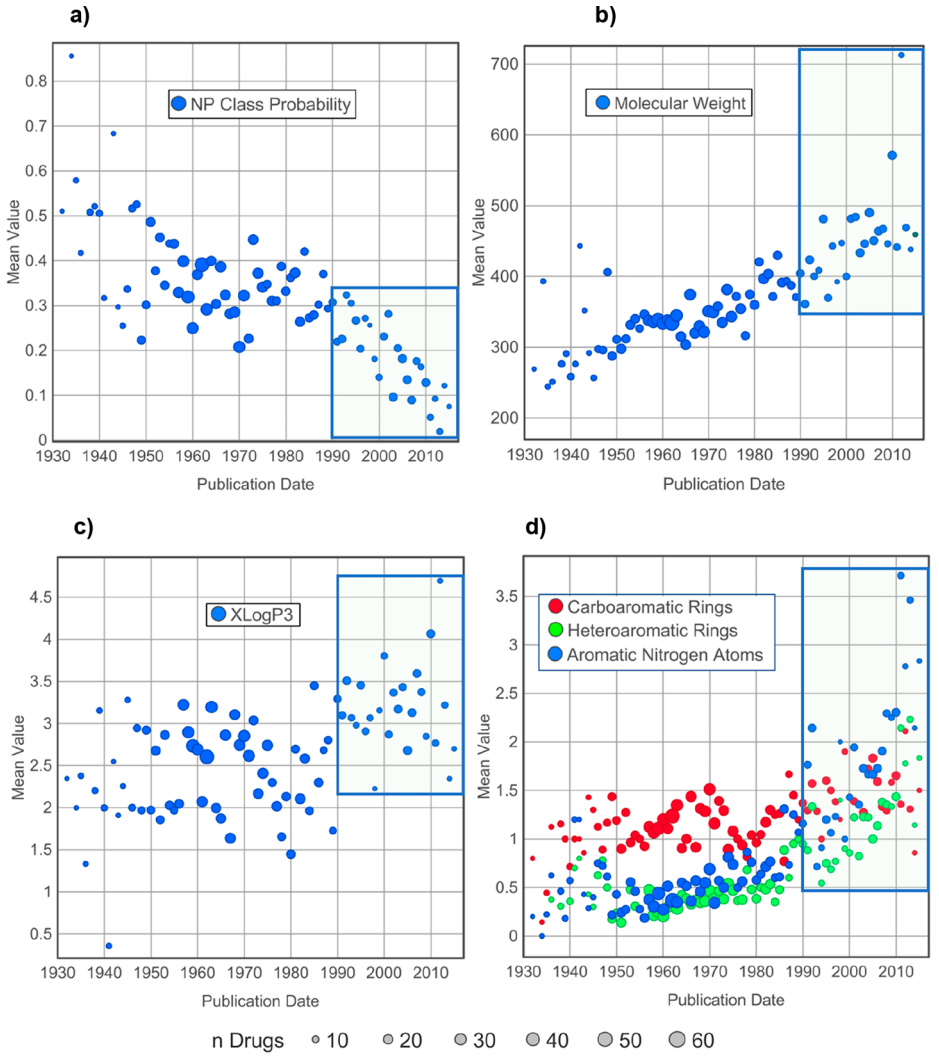


Figure 1. Mean oral drug properties versus year of first publication (mostly the first patent) as found in CAS Scifinder (CAS SciFinder | CAS) from 1930 onward. Years with <5 drugs are excluded. (a) Natural product class probability.[8] (b) Molecular weight. (c) Lipophilicity (XLogP3). (d) Carbo- and heteroaromatic ring counts and aromatic nitrogen atom counts. Drugs published post-1990 are boxed. The oral drug data set was manually updated to September 2022 approvals, from where previously disclosed.[25] Property calculations were done with Swiss ADME (SwissADME) and DataWarrior (www.openmolecules.org) and graphs created with DataWarrior. The higher mean values of molecular weight and XLogP3 in the years 2010 and 2012 are a consequence of several Hepatitis C Virus inhibitors, which comprise 5 of 23 drugs in 2010 and 4 of 9 drugs in 2012, all having molecular weight > 750 and XLogP3 > 4.5.

synthesis of compound libraries using parallel chemistry, facilitated the optimization of in vitro affinity of screening hits. These changes in practice help to explain the trend in Figure1a.

The time dependency of other physical properties in Figures1b–d is also consistent with a turning point occurring in around 1990. The overall increase in molecular weight in drugs over time is well-known,[15−19] but the detailed annual analysis in Figure1b shows that, after a period of little change in the 1980s, a further increase starts from around 1990. The lipophilicity of drug molecules is a key attribute and is increasing less over time than is molecular weight, and indeed in early studies it appeared relatively unchanged.[15,17,18] However, the extended time frame from 1990 shows an overall increase in XLogP3 of about 0.5 units has occurred since then[16] (Figure 1c).

The increase in molecular weight since 1990 is accompanied by increases in aromatic ring count, with a notable shift toward increased use of aromatic heterocycles (Figure 1d). The changes in nonaromatic rings over time are less substantial in comparison (not shown). Controlling carboaromatic ring count is consistent with the need to manage lipophilicity and overall developability.[16,20,21] The large increase in aromatic nitrogen atom count (Figure 1d) occurring since 1990 helps to explain the reduced natural product score, because nitrogen atoms are relatively uncommon in natural products versus synthetic compounds.[8]

Overall, increased molecular weight, lipophilicity, and heteroaromaticity co-occur along with the decline of computed natural product probability in oral drugs invented from 1990 onward. It is notable that novel ring systems[22] and frameworks[23] are being continuously introduced into drugs over time, but these appear to be having little impact on the natural product score in recent drugs. Combinations of natural product motifs in nonbiogenetic patterns, called pseudo-natural products,[24] provide additional insight into the impact of natural product structure in drug discovery, and are discussed in subsequent sections.

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

Synthetic innovation in drug discovery is often interpreted as the need for the employment of a wider array of techniques, methodologies, and technologies to furnish libraries of molecules,[26] providing the bedrocks that enable contemporary screening practices. The notion of defining,[27] exploring,[28] and extending[29] "chemical space" leads to various interpretations of what is needed for bioactivity[30] − as such, space is a nebulous term that might be described in various ways ([Box 1), as could the term "drug-likeness" (Box 2). Combinations of size, structure, shape, lipophilicity, and other physical characteristics can all be used to define space[31] or even devise rules for drug-likeness; but one size clearly does not fit all.

Box 1. Chemical Space

Compounds can be characterized by combinations of various *descriptors*, such as their size, mass, lipophilicity, charge, and topological features. Multidimensional combinations of these could be said to describe chemical space and drugs may disproportionately occupy regions defined by some combinations, yet frequent outliers clearly exist.

Box 2. Drug-like

This rather nebulous term is often misused—as there are no overarching descriptions that might differentiate drug molecules from other natural or synthetic compounds (without recourse to their activity and efficacy). Combinations of physical descriptors such as lipophilicity, weight, size, aromatic ring count etc that are used to define ***Chemical Space***, may be used to define ***Drug-like***, but different targets and routes of administration necessitate distinct combinations of properties.

To fuel the capacity of high throughput screens, most corporate collections are populated with inexpensive molecules made using readily available building blocks[32] and privileged structures (Box 3); a paradigm that delivers hits. Such practices may not be the panacea or contribute proportionately to the costs and quality of output; indeed, different organizations experience contrasting outcomes.[33] That the makeup and composition of corporate collections differ from both drugs[34] and natural products[35] is recognized, leading to questions about the biological relevance of synthetic collections versus natural compounds. An outcome of the review of *Nature versus Nurture in drug discovery* was to suggest a sweet spot (Figure 2) for drug design based on lipophilicity and weight,[36] perhaps relevant for synthetic molecules of but clearly unrepresentative of the physicochemical diversity of natural products or approved drug molecules (vide infra). The size-lipophilicity singularity of the "sweet spot" is probably reflective of synthetic origins and the chemistries employed in contemporary practice; such narrowness is clearly challenged by naturally derived molecules, and such a narrow range of properties is unlikely to be apposite for all targets. In particular, drugs with

Box 3. Privileged Structure

A chemical structure that is predisposed to furnish bioactive compounds based on differing substitution patterns appropriate for distinct binding sites.

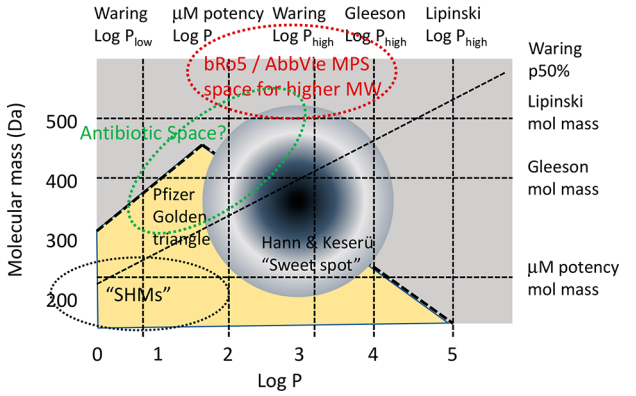


Figure 2. Adapted version of the drug discovery "sweet spot" defined in the molecular mass−log P space proposed by Hann and Keseru taken from ref 43. The diagonal line "Waring p50%" is the proposed line for 50% chance of achieving reasonable permeability.[44] Additional regions for "small hydrophilic molecules" (SHMs) with potential paracellular routes of absorption and the Pfizer "golden triangle,"[45] beyond the Rule of 5/AbbVie MPS,[46] and common antibiotic space are highlighted.

increased molecular weight and higher lipophilicity are emerging, contrasting practices that delivered more hydrophilic drugs with lower molecular weights. Current developments in computational methods allow for the virtual screening of millions of compounds to identify smaller numbers of molecules for screening, with demonstrable success.[37] Furthermore, the imaginative exploration[38] of vast virtual databases such as the innovative Enamine Real[39] has enabled the rapid and more cost-effective procurement of smaller numbers in silico hits for in vitro validation, demonstrably changing screening practices and quality.[32] Using such methods or well-designed indexed HTS collections, potentially rapid SAR expansion is thus feasible. DNA encoded Libraries (DELs) can be constructed with similar design principles to furnish similarly vast numbers for screening, limited only by synthetic feasibility.[40,41] The success of these approaches in delivering synthetic hits is proven, but how they might better be exploited in future will be considered, with opportunities for more imaginative monomer design with simple chemistries.[42]

The ensuing three sections consider and contextualize (i) the properties of drug molecules and how these relate to their disposition, noting the importance of molecular recognition in this; (ii) the motifs and structures in natural and pseudo-natural products, and (iii) the relevance and opportunities for the employment of biotransformations in drug discovery.

THE PHYSICOCHEMICAL CHARACTERISTICS OF DRUGS

A principle advocated by Hansch that *drug molecules should be made as hydrophilic as possible without loss of efficacy*[47] is commonly expressed and utilized as Lipophilic Ligand Efficiency (LLE).[48] This metric, widely accepted and exploited in drug discovery as a key metric in optimization, is expressed on a log scale as activity (e.g., $-\log_{10}[XC_{50}]$) minus a lipophilicity term (typically the Partition coefficient or $\log_{10}$ P or sometimes $\log D_{7.4}$).[49] The impact of lipophilicity on efficacy needs to be considered in the context that reducing lipophilicity (equating to increasing hydrophilicity) will generally increase the solubility, reduce the metabolism, and reduce the promiscuity of a given compound in a series.[50]

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

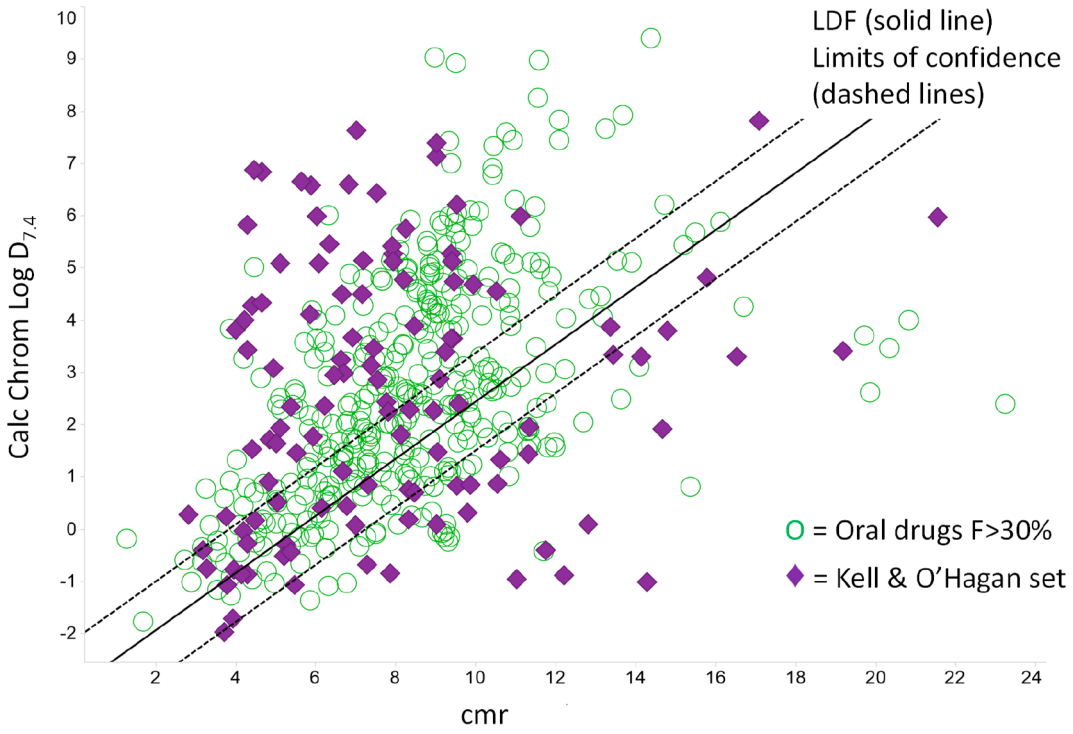


Figure 3. Natural products are found across most size lipophilicity combinations, as exemplified in a representative set designed and compiled by O'Hagan and Kell[64] superimposed on the Chrom log $D_{7.4}$ vs cmr training set of compounds with >30% bioavailability.[51]

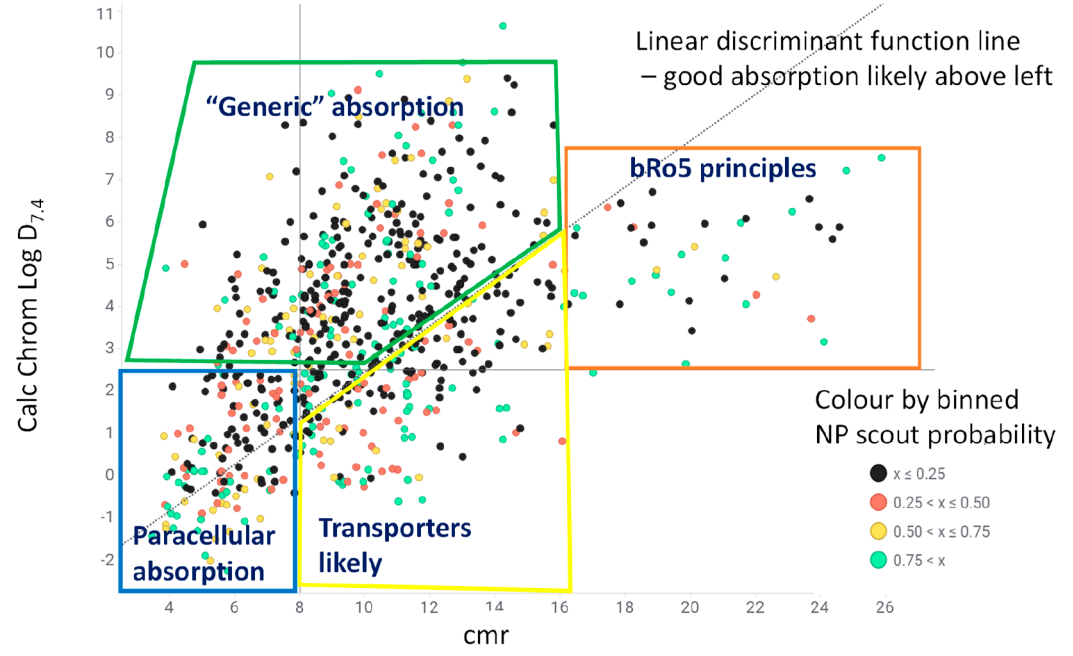


Figure 4. Proposed regions of size/lipophilicity space for an oral drug set,[51] using the effectual combination of Chrom Log $D_{7.4}$ vs calculated molar refraction (cmr) as a description of chemical space. The highlighted regions suggest likely absorption mechanisms, based on ref 65 with compounds colored by binned NPScout probability scores. Below the LDF line, then mean NPScout score is 0.45, (median 0.33) and above it (indicative of likely oral exposure) the mean is 0.31 and median 0.17 ($p < 0.01$).

Conversely, such a change will often reduce the inherent potency of the molecule and potentially compromise permeability.[51] Optimization in drug discovery is a necessary process of compromises between such conflicting demands, wherein intrinsically less potent molecules can prove more efficacious with improved pharmacokinetic exposure/higher free fraction.

The crossing of biological membranes by drug molecules is a necessary process to enable efficacy for most drugs, important in, inter alia, absorption from the gut, reabsorption in the kidneys, traversing the blood-brain barrier, and achieving therapeutic concentrations at intracellular targets. That these processes are sometimes facilitated by proteins[52] (solute carriers or transporters) is widely demonstrated,[53] but

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

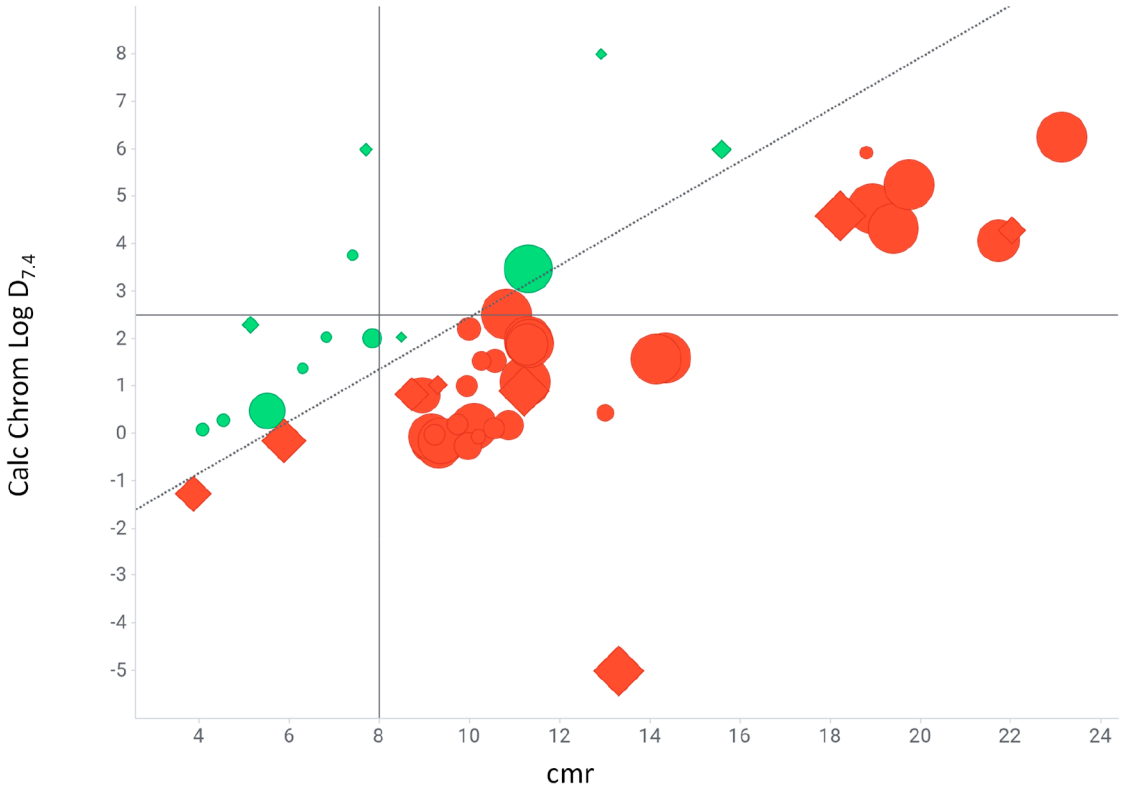


Figure 5. Illustration of antibiotic drug space, expressed as Calculated Chrom Log $D_{7.4}$ vs cmr adapted from data in ref 65 colored by antibiotics (circles) and TB drugs (diamonds) which are sized by NP class probabilities and colored by prediction of likelihood of oral exposure (either side of the diagonal "linear discriminant function line" so to be oral, transporters a likely mechanism for the red colored compounds, which mostly have a high NPScout score). Vertical (cmr < 8) and horizontal lines (Chrom Log $D_{7.4}$ < 2.5) together represent likely boundaries for paracellular absorption.

conflicting hypotheses regarding the extent of such processes remain subjects of contention.[54] Whether all molecules cross membranes via protein-mediated processes rather than a proportion entering by a partitioning process through the lipid bilayer structure is the essence of the debate.[55] Noting that carrier proteins facilitate the passage of molecules suggests that recognition processes are important in some (or potentially all) permeation events.[56] The evolutionary selection for cells expressing carrier proteins was driven by the necessary functions or beneficial impact of the transported molecules. The general levels of similarity between oral drugs and exogenous metabolites,[57] the major class of substrates for the carriers, led to the proposal of a "Rule of 0.5" by the Kell group.[58]

The impact of suboptimal physicochemical properties and the potency driven culture in drug discovery,[59] fuelled by high throughput screening practices are subjects for many reviews,[15] dissections and formulations of rules and principles.[60,61] Indeed, as molecules became more lipophilic, the consensus method to estimate lipophilicity itself was shown to be flawed when chromatographic methods and solubility were considered.[50] The bilinear relationship between permeation and lipophilicity[52] was rediscovered and the implications of this illustrated the necessary compromises needed to optimize intracellular[62] and/or oral drugs. Two useful guides to predict oral exposure, formulated from empirical observations, are the GSK model based on log $D_{7.4}$ vs calculated molar refraction (cmr, which largely correlates with MW)[51] and the AbbVie MPS[46] (AV MPS = |clog D7.4 − 3| + #Ar + #Rot; developed for high MW compounds, but the principles are applicable to less heavy ones too[51]) are consistent with these observations. In their groundbreaking paper on permeability and solubility that introduced the Rule of 5 (Ro5), Lipinski and colleagues noted the anomalies represented by compounds of classes known to be substrates for transporters.[63] Such exceptions, commonly at extremes of lipophilicity and or higher MW/size, which nonetheless achieve oral exposure, are almost invariably recognizable natural products or close analogues. The distribution of marketed oral drugs in terms of their lipophilicity and size, shows a remarkably similar distribution to the set of compounds designed by Kell as a representative set of natural products to investigate carrier mechanisms (Figure 3).[64] Natural mimics with these size/lipophilicity combinations offer genuine means of exploring expanded regions of space away from the perceived singularity described earlier. Analysis of oral drugs in Figure 4 shows the distribution of NP Scout probability scores, which is significantly higher (p < 0.01) for those orally available compounds lying below the linear discriminant function line than those above it (in spite of natural products showing ubiquitous distribution as illustrated in Figure 3). While this is a simplification, the ability for natural products to bend the rules, most likely due to the exploitation of transporters, is evident in the observed patterns.

The disconnect between the properties of tuberculosis drugs in particular[65] and antibiotics in general from approved drugs and screening collections was described in these terms (Figure 4).[66] That the majority of antibiotics have high NP likeness scores (indeed, several semisynthetic beta lactam structures have low scores) and lie in a hydrophilic region clearly differentiated from the majority of oral drugs is illustrated in

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

Figure 5. Nonetheless, most of these compounds are oral agents and not susceptible to microbial efflux defense mechanisms, so it is logical and reasonable to invoke the importance of molecular recognition and natural transporters in their activity.[65] This has profound implications for the physical descriptors to target for future antimicrobial research[67] and reflects observations of poor success rates in the field from high throughput screening of corporate collections.[68] In this vein, a recent notable antibiotic success story focused on porin permeation[69] in addition to biochemical potency during optimization.[70]

In addition to lipophilicity, higher counts of aromatic rings (#Ar)[20] and sp^2 hybridized atoms[71] are associated with increased risks in developability assays and lesser chances of clinical progression[72] (Escape from Flatland, Box 4).

Box 4. Escape from Flatland, Aromatic Ring Count (#Ar)

A concept coined to reflect improved chances of favorable outcomes for compounds with a higher ratio of sp^3- vs sp^2-hybridized atoms in their structure, alternatively (and more tangibly) expressed as the simplistic count of aromatic rings (#Ar) in a structure. Property and Solubility Forecast indices are based on the addition of #Ar to the logarithm of the partition or distribution coefficients; the higher the summation, the bigger risk of undesirable outcomes in developability assays is the principle.

Carboaromatics are most notably risky,[21] noted from data analyses[16] and the principles of the Property Forecast Index, which equates to an aromatic ring of having an impact akin to another log unit (10-fold increase) in lipophilicity.[50] Structural constraint delivers enhancements to physiological interactions and activity by minimizing entropy loss[73] (Freire principles,[74] Box 5), so achieving this with sp^3-rich structures and minimal aromaticity is a desirable tenet in design.[75] Nature has optimized such interactions over hundreds of millions of years. Minimizing rotatable bond count is another key element of the AbbVie MPS and is an established design principle. Hydrophilic character in molecules is introduced by polarization of bonds with heteroatoms (i.e., noncarbon atoms, most commonly oxygen, nitrogen, sulfur) that are key to polar (hydrogen bonding or charged-based) molecular interactions. Forming such specific and directional interactions contributes to both specificity and selectivity given the energetic cost of desolvation between aqueous solution and a bound state (the enthalpy change of a productive polar interaction usually offsets the desolvation cost on removing the motif from aqueous solution). Noting the differing distribution of, inter alia, rings, rotatable bonds, lipophilicity, and nitrogens, Feher and Schmidt suggested that mimicking these patterns could produce collection of substantially more diverse compounds with greater biological relevance.[35]

Box 5. Freire Principles for Enthalpy-Driven Design

Work on characterizing the thermodynamic signatures of drug binding by Freire[74] and co-workers led to the following four principles to maximize the enthalpic contribution to the free energy (activity) signature associated with binding.

1. Polar groups should establish strong hydrogens bonds with structured groups in target.
    a. Enthalpy of hydrogen bond formation usually offsets desolvation of ligand; make a H-bond = retain activity
2. Eliminate buried polar groups that do not establish hydrogen bonds with the target.
    a. Polar group making no interaction = loss of activity (can be very important in achieving selectivity)
3. Nonpolar groups should fit tightly in binding pocket cavities
    a. Maximise filling of pockets; nature abhorring a vacuum and the positive effect of van der Waals interactions are important here.
4. Losses in conformational degrees of freedom should be minimized.
    a. Similarity of bound and solution conformations (or ready interconversion between them) = lesser entropy loss on binding

A logical extension of this hypothesis would be to consider recognition processes with natural molecules, which are likely to have discrete interactions with carrier proteins and therapeutic targets. Small molecule drugs are noted to be relatively promiscuous, so making interactions with several proteins is a likely event.[76] It similarly is logical to consider that a molecule made by a recognition process in a catalytic enzyme may also interact with another protein in a similar manner.[77]

The ensuing sections will consider these physicochemical observations in the context of (i) the finite number of substructures present in natural products, progress in the synthesis of close analogues (or pseudo-natural products), (ii) why natural products are validated via their interaction with protein surfaces during biosynthesis, and (iii) where and how biosynthetic mechanisms and methods offer sustainable methods for the discovery and manufacture of more rationally designed drugs in future.

DESIGN OF COMPOUND COLLECTIONS WITH HIGHER NATURAL PRODUCT-LIKENESS

Typical screening collections are populated by molecules that are structurally related to bioactive synthetic compounds with purported drug-like or lead-like chemical properties.[78] Nevertheless, their criteria represent a narrow portion of chemical space that may not overlap with many potential biological targets.[8] Natural products (NPs) are the result of Nature's exploration of biologically relevant chemical space through evolution and are a valuable source of structures[7,79] that expand into areas of chemical space that are underrepresented in screening libraries.[80,81] Yet, NPs have limitations in the scope of drug discovery. Natural evolution is slow and restricted by selectivity pressures. These evolutionary constraints have resulted in NPs occupying only a small portion of NP-like chemical space and having a limited number of scaffolds due to conserved biosynthetic pathways.[82] Additionally, the supply of NPs of interest may be scarce due to low yielding biosynthetic pathways and/or low synthetic tractability. Therefore, employing logic derived from evolution to design new small molecules may provide guidance for the exploration of NP-like biologically relevant chemical space in a synthetically more feasible manner.[83]

Biology-Oriented Synthesis

As NP biosynthesis and protein modulation are intertwined, so are protein and NP structures. Through evolutionary pressures,

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

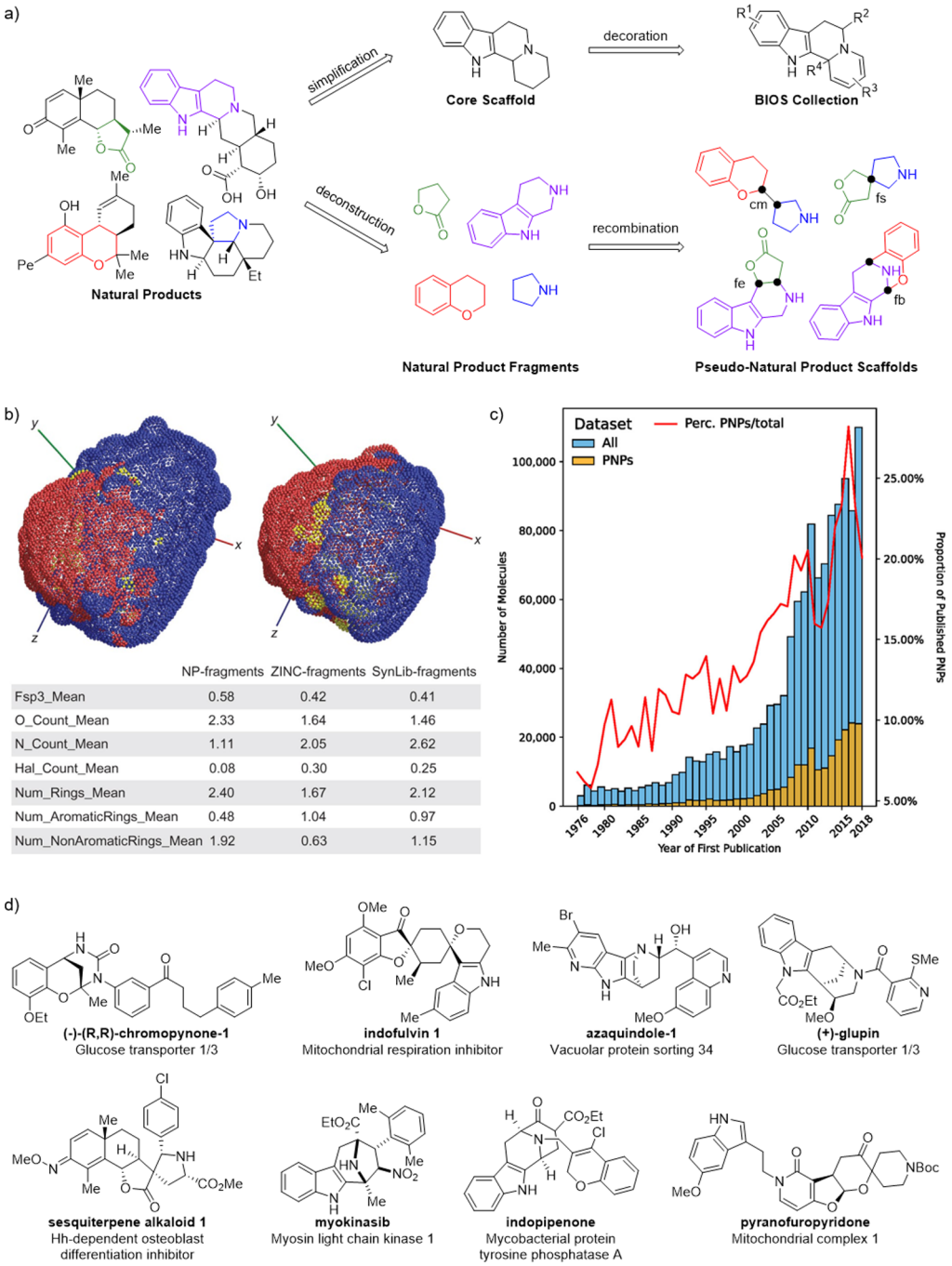


Figure 6. Natural product-like compound collections. (a) Design workflow for Biology-Oriented Synthesis and Pseudo-Natural Product compound libraries (Pe = nPentyl). (b) Comparison of the chemical space of natural product fragments and fragments derived from commercially available compounds. Each dot is a representative cluster of about 25 fragments with similar molecular features. Top left (b): Principal component analysis of NP fragments and ZINC fragments. Clusters with two-thirds or more of NP fragment content (red), two-thirds or more of ZINC fragment content (blue). Clusters that are in between are in yellow. Top right (b): Principal component analysis of NP fragments and SynLib fragments. Clusters with two-thirds or more of NP fragment content (red), two-thirds or more of SynLib fragment content (blue). Clusters that are in between are in yellow. Adapted from 10.1038/nchem.1506.[22] (c) The number of bioactive molecules published according to ChEMBL and the proportion that are pseudo-natural products over time. Adapted from ref 105. Copyright The Authors. (d) Selected pseudo-natural products for which mode of actions and/or biological targets have been identified.

a minute fraction of possible proteins has been selected by Nature that are fairly conserved as are the structures of their binding pockets.[31,84] Protein Structure Similarity Clustering (PSSC)[85] can identify the ligand sensing cores of proteins and sort them into protein similarity clusters based on structures and not amino acid sequences. Similarly, NPs have scaffolds that are conserved but can have high variability in their appendages. The variability of substitutions of a common scaffold can result in different protein targets and therefore different bioactivities. The Structural Classification of NPs (SCONP)[86] charts the regions of chemical space explored through evolution through systematic truncation of NP structures to provide a scaffold tree that is representative of all NPs with ring-based scaffolds.

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

Biology-Oriented Synthesis (BIOS) merges the idea that (1) the structure of ligand sensing cores of proteins and scaffolds of NPs are conserved in Nature and (2) deviations in amino acid side chains or scaffolds decorations can greatly influence binding and selectivity (Figure 6a).[87−89] Using PSSC and SCONP in tandem,[90] structurally similar protein binding pockets can be matched with truncated scaffolds in the SCONP tree whose parent NPs show activity for one or more proteins in the cluster. The identified scaffolds can then be decorated to afford compound collections that may be enriched in selective bioactive molecules for the entire protein cluster and may be more synthetically tractable than their parent NPs. In a sense, the workflow of BIOS can be considered the matching of proteomic space with biologically prevalidated chemical space.

BIOS compounds are inspired by NP structure and can efficiently explore biologically relevant chemical space occupied by NPs which inherently brings about limitations. BIOS relies on NP scaffolds as starting points; however, Nature has only explored a fraction of NP-like chemical space which is represented in its limited NP scaffolds. Furthermore, the scaffolds of BIOS compounds are directly derived from NPs and are therefore likely target the same protein cluster as the parent NPs, hindering the exploration of new bioactivities.

Natural Product Fragments in Fragment-Based Drug Discovery

Fragment-based drug discovery (FBDD) can be employed to rapidly explore large areas of chemical space for starting points of molecular design.[91−93] However, most FBDD libraries are composed of privileged substructures of known synthetic drugs and drug candidates and populate already well-explored areas of chemical space,[94−96] often through the use of fragments with high sp^2-character.[97] Underexplored areas of chemical space can be rapidly explored by employing fragments derived from NPs that are already biologically prevalidated by evolution.

Through the deconstruction of more than 180 000 NPs, 2000 structurally diverse NP fragments have been identified that retain the molecular features of NPs and populate areas of chemical space not occupied by representative synthetic fragment collections (Figure 6a and b).[98] Employing NP fragment collections that are composed of intrinsically different characteristics to typical collections for FBDD programs may lead to structurally novel ligand types for known biological targets with more desirable properties.

Pseudo-natural Products

The design principle of pseudo-NPs (PNPs) merges the biological relevance of NPs with the rapid access to diverse chemical space offered by fragment-based discovery. The PNP concept aims to recombine NP fragments or fragment-sized NPs[98−100] in arrangements that are not observed in Nature to afford compounds that are NP-like but are not obtainable through current biosynthetic pathways (Figure 6a).[101] The compounds therefore retain the biological relevance of NPs but may explore new areas of biologically relevant chemical space that may lead to new structural classes for known targets or the discovery of new targets in conjunction with broad and/or unbiased screening assays.

Through different NP fragment combinations[102] and arrangements,[103] several chemically distinct PNP scaffolds can be envisioned. Combinations that are particularly of interest are those that do not appear together in Nature and/or have biosynthetically unrelated origins.[104] Further points of scaffold diversification can come from different fragment connectivity types, such as monopodal (cm), edge fusion (fe), bridged fusion (fb), and spirocyclic fusion (fs) (Figure 6a). The type of connectivity pattern can have a significant impact on the shape of the resulting molecule. In synthetic compound collections, fragments are commonly connected in a monopodal fashion to give a linear linkage with a rotatable bond and may be due to synthetic ease. In Nature, biosynthetic pathways more commonly connect fragments through more complex connectivity patterns of two or more bonds.[105] The resulting structures may then be more rigid and/or three-dimensional than their synthetic monopodal counterparts. To more accurately mimic and retain the properties of NPs, complex connectivities should be employed and can be efficiently obtained through complexity-generating reactions. Beyond different fusion patterns, different regioisomeric arrangements of fragments can alter the PNP scaffold.

The concept of combining different NP fragments or fragment-sized NPs is not exclusive to PNP design. Nature frequently employs this tactic to afford NP hybrids by combining biosynthetically unrelated metabolic units[106] or by hetero- or homodimerization of NPs.[107] NP hybrid strategies employed by Nature and the PNP concept share the common goal of exploring biologically relevant chemical space; however, these two strategies differ in the design, preparation, and evaluation of new molecules.

NP hybrids are produced by enzymatic cascades that can be summarized in biosynthetic pathways. Through evolution, Nature has the potential of mutation and recombination to provide new biosynthetic machinery and pathways that can produce NPs with novel scaffolds. The biological relevance of the resulting NPs is evaluated by selectivity pressures in terms of the producing organism's survival and reproduction. Reiterations of this cycle are the natural evolution of NP structure.[108]

PNPs are designed by the cheminformatic deconstruction of NPs to NP fragments then recombined and decorated via synthetic reactions. The resulting PNPs are evaluated in biological assays and compounds of biological relevance are selected for reiteration. The logic of the PNP concept shares similarities to natural evolution and may be considered the chemical evolution of NP structure.

While both approaches can afford biologically relevant compounds, the natural process of mutation and selection is slow. Synthesis allows the mutation process to be accelerated and offers flexibility and opportunities for new NP fragment combinations by utilizing reaction pathways that are not possible in biosynthesis.

Using the PNP algorithm, several PNP collections have been prepared over the past decade. By employing assays and subsequent follow-up biological characterizations, several PNP classes have been identified to be enriched in bioactivity that affect therapeutically relevant pathways, processes, and targets (Figure 6d).[24,109−115] The frequent success of the PNP algorithm suggests that it is a valid design principle for the exploration of biologically relevant chemical space.

While the design principle of PNPs has only recently been reported,[101] compounds that meet the criteria of PNPs, i.e., two or more NP fragments in arrangements not found in Nature, have been synthesized for decades. A recent cheminformatic study found that 23% of synthetic compounds in the ChEMBL library, i.e., bioactive compounds, are

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

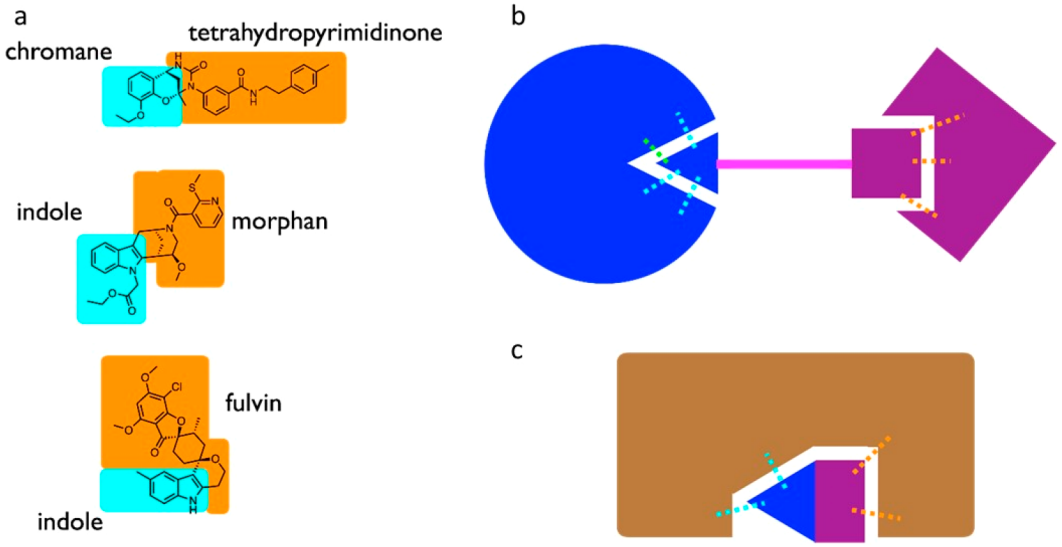


Figure 7. (a) Three examples of pseudo-NPs with different biological activity to the starting fragments. (b) Linking of two NPs allows each NP to recognize its original interacting partner. (c) Pseudo-NPs retain some of the interacting vectors to bind to novel protein partners.

PNPs.[105] Additionally, the percentage of PNPs being added to the database per year has been increasing over time (Figure 6c), indicating that a significant number of bioactive PNPs have been unintentionally synthesized without a guiding principle.

Although NPs by definition are biologically relevant, there is a significant number of biologically active and therapeutically relevant small molecules that are either partially or entirely composed of fragments not found in Nature.[105,116,117] Therefore, BIOS and PNP concepts may be extended to incorporate all biologically relevant molecules into their design algorithms. In turn, this may lead from these concepts exploring NP-like chemical space to more general biologically relevant chemical space.

MOLECULAR RECOGNITION

That the embedded recognition of natural products for proteins correlates with recognition of the biosynthetic enzyme is an increasingly validated concept.[118−120] The biosynthetic imprint translates to recognition of other proteins using similar interactions. For example, the analysis of protein structures of 38 biosynthetic enzymes gave 64 potential targets for 25 natural products.[121] Retrospective analysis validated the predicted targets for tetrahydrocannabinolic synthetase (THCA), acyltransferase lovD and isopenicillin N synthase (IPNS).[121] In the case of a pseudo-natural product, it might be expected that interactions would not correlate with either of the biosynthetic enzymes from which its fragments are derived, so, in effect, new protein recognition sites will be generated.

Three examples of pseudo-natural products with new biological activity not displayed by either of the starting fragments are presented in Figure 7a. Chromopynones selectively inhibit glucose uptake in cells; however, other compounds that contain only either the chromane or THPM fragments did not inhibit glucose uptake.[122] The indomorphan[123] and indofulvin[109] classes inhibit glucose uptake and starvation-induced autophagy, respectively. None of the individual fragments retained the bioactivities of the pseudo-NPs. These results suggest that the novel bioactivities of pseudo-NPs are not related to their individual fragments but rather the result of the combination of NP fragments.

If the two natural products where tethered, each natural product might retain their original protein interaction. This was demonstrated to be the case for PROTACs and should apply to other chimeric molecules that are intended to have polypharmacology (Figure 7b).[124−126] Complex fusion of fragments generates new scaffolds bringing H-bonding interactions into close proximity and appears to retain the ability of protein recognition but results in binding to a new/novel protein surface. (Figure 7c).

An understanding if the importance of chimeric behavior in larger molecules is increasingly emerging, beyond the empirical observations that led to the AbbVie MPS. As molecules get bigger and intramolecular interactions[127] become increasingly feasible and important, new insight into their constraint in macrocycles and/or the role of environment-dependent alterative conformations[128] is emerging, quantified by solution NMR observations[128] or using supercritical fluid chromatography methods to understand the effective polar surface area (EPSA) presented by the hydrophobic form.[129] It is reasonable to consider natural macrocyclic structures to be excellent templates for better exploiting such phenomena, given the role of evolution in selecting their structures.

BIOCATALYSIS IN DRUG DISCOVERY

Natural products and the enzymes that are responsible for their synthesis are intrinsically linked. Historically, the screening of natural products has provided a rich source of lead compounds for further development as natural product or semisynthetic drugs. Prime examples here include polyketides, beta-lactam antibiotics, peptides, steroids, and opioids. In some cases, the structural complexity of these compounds resulted in considerable effort being invested into understanding their biosynthetic origin, paving the way for the production of these bioactive compounds using microbial fermentation. For example, studies on the biosynthesis of penicillins and cephalosporins ultimately led to commercial processes for their production on scale.[130] Similarly, a detailed understanding of the biosynthetic pathways leading to polyketides

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

facilitated their production using microbial hosts.[131] In this approach, pathways comprising multiple genes/biocatalysts are assembled to convert simple feedstocks into both the natural product and also analogues via mutasynthesis.[132] However, these "biosynthetic" enzymes tend to be characterized by relatively low turnover rates and also narrow substrate scope, precluding their more general application as biocatalysts for target molecule synthesis.

To address this challenge, a complementary approach has been recently explored, namely, to create new engineered biocatalytic platforms that contain enzymes with broader substrate scope and hence wider application in the synthesis of natural product-like molecules.[133,134] These enzymes (e.g., transaminases, ketoreductases, imine reductases, aldolases) are typically recruited from primary metabolic pathways and hence have inherently higher catalytic activities, and broader substrate scope, than their counterparts from secondary metabolic biosynthetic pathways. Importantly, as a result of spectacular advances in protein engineering and directed evolution, these enzymes can be engineered to further increase their activity, selectivity, and substrate scope.[135] When coupled with equally important breakthroughs in enzyme discovery via metagenomics and bioinformatics, the ability to rapidly generate a broad toolbox of new biocatalysts becomes a powerful approach for organic synthesis and drug discovery. Exploiting biocatalytic methods using products from new synthetic developments is another emerging area with much potential.[136]

The recent availability of new biocatalytic chemistry platforms, with broad substrate scope, has impacted drug discovery programs in two important ways. First there is an increasing awareness that engineered biocatalysts have a key role to play in API manufacture since they often lead to a more efficient synthetic route, with reduced cost of goods, lower environmental impact, and importantly higher sustainability.[137] To assist with this process of where biocatalysts can be applied, there are now tools available for retrosynthetic planning (e.g., RetroBioCat) which incorporate enzymes in the route selection and planning process.[138] The second, and more recent, impact is in the earlier stage screening process for lead identification. Panels of engineered biocatalysts can be used to create either building blocks, or key intermediates, for biological screening. In this way biocatalysis can provide access to segments of chemical space that otherwise would be difficult to access via alternative synthetic approaches.[139] Examples here include hydroxylation (P450 monooxygenases), halogenation (halogenases), methylation (alkyltransferases), and amide bond formation (amide ligases). Biocatalysts also have an increasing role to play in Late-Stage Functionalization (LSF).[140]

A key issue here is that not only can biocatalysis provide access to alternative, difficult to access, chemical space but maybe it can also generate lead compounds that are more natural product-like, given their origins. As a rule, substrates and products for biocatalysts are characterized by the following properties: (i) MW 100−500 Da, (ii) good to high water solubility, and (iii) presence of one or more polar functional groups. At the lower sizes and with good solubility, biocatalysis is therefore a potentially very useful tool for Fragment Based Drug Discovery.[141] An analysis of the various synthetic platforms currently available through biocatalysis reveals that particular types of transformations are well represented, in particular reduction, oxidation, C−X bond formation, and hydrolysis/reverse hydrolysis. The reduction of ketones to secondary alcohols using KREDs is now well established but recently this has been expanded to encompass reduction of carboxylic acids to aldehydes and C═C to C−C. Selective oxidation represents a significant opportunity with enzymes now available for oxidation of alcohols, amines, alkenes, C−H bonds etc. Advances in C−X bond formation include C−C (aldolases, lyases, cyclases), C−N (imine reductases, transaminases), and C−Hal (halogenases). Hydrolytic enzymes, which have long been used for ester/amide hydrolysis, are now being repurposed particularly to address the challenge of generic methods for amide bond synthesis.

Another emerging theme in biocatalysis, which is relevant to drug discovery, is the development of multienzyme cascade processes. In the same way that Nature uses cascades of enzymes to generate biosynthetic pathways for natural product synthesis, researchers are increasingly trying to generate new multienzyme cascades for the synthesis of APIs and advanced intermediates. A prime example of this concept is the Merck synthesis of Islatravir which involves a total of 10 different biocatalysts deployed in three sequential steps.[142] Cascade processes have also been reported for the synthesis of key natural product-inspired building blocks including piperidines, tetrahydroisoquinolines, quinolines, and similar molecules. In each of these examples, biocatalysts are used to not only catalyze key bond-forming processes but also to recycle essential cofactors (e.g., NADPH, ATP) to ensure that the whole processes is synthetically efficient.[143] Multienzyme cascade processes serve to highlight some of the attractive features of biocatalysis including (i) no requirement for protecting groups, (ii) use of water as a common reaction medium, and (iii) minimum isolation of intermediates en route to product formation. Furthermore, the promise of enzymatic reactions for on-DNA synthesis has thus far received little attention,[144] perhaps surprisingly, given the necessity of an aqueous environment for their construction. An isolated example using glycosyl transferases and galactose oxidase illustrated the potential.[145]

Moving forward, the way in which engineered biocatalysts are being deployed as part of the overall drug discovery process is changing rapidly. By its very nature, biocatalysis is a "modular" technology which lends itself very well to automation and the application of high-throughput techniques. At each stage of the pipeline from enzyme discovery, via enzyme engineering, to enzyme screening, it is now possible to leverage the combined full power of bioinformatics, computational design, robotic screening, protein structure prediction, modeling, and in vitro screening. Enzyme discovery has benefitted enormously from interrogation of sequenced metagenomes which provide an almost unlimited supply of new enzymes. These enzymes can be rapidly expressed, either using conventional *E. coli* hosts or in vitro translation, and screened in MTP format to identify candidate hits which are then subjected to rounds of directed evolution and protein engineering to improve their characteristics. Here there is an increasing focus on applying computational design/AI methods to reduce library size and hence more rapidly optimize protein/sequence space for the desired application. Optimized biocatalysts that emerge from this pipeline can then rapidly and predictably be scaled for the delivery of increasing quantities of material required in clinical trials.[146]

Much remains to be done to fully exploit the opportunities for integrating biocatalysis more deeply into the drug discovery

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

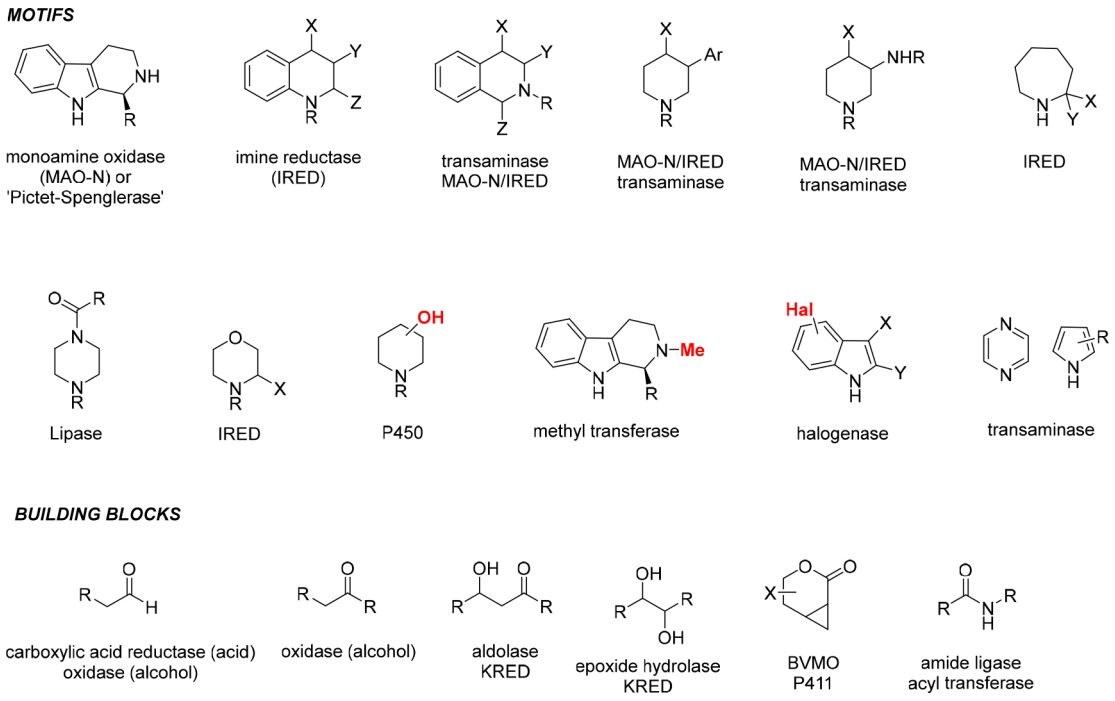


Figure 8. Example structures and strategic application of enzyme classes that enable the synthesis of biorelevant molecules and building blocks with a pedigree of already having been recognized by a protein structure.

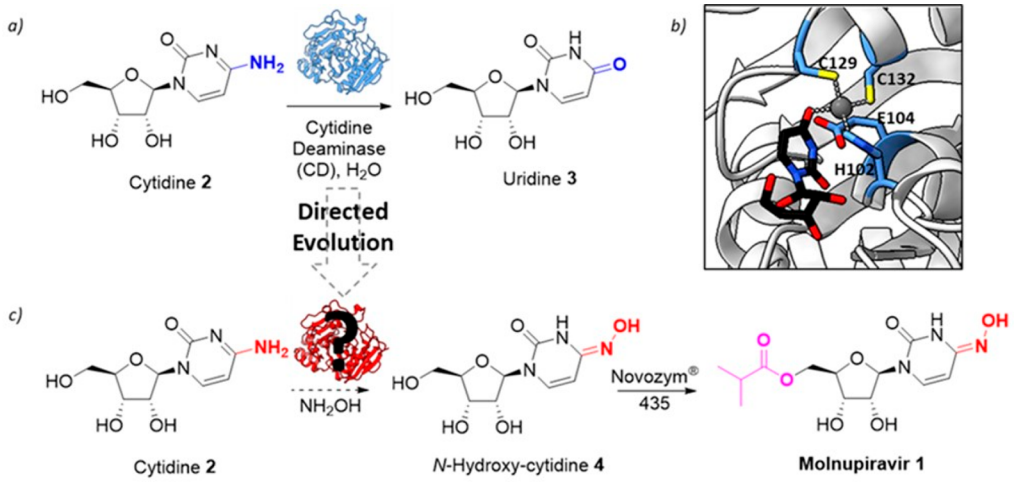


Figure 9. Biocatalytic route to Molnupiravir. (a) Wild-type cytidine deaminase (CD) catalyzes the hydrolysis of **2** to **3**. (b) Active site of CD with uridine bound (PDB code: 1AF2). The Zn^{2+} ion is shown in gray. His102, Cys129, and Cys132 and catalytic Glu104 are shown as atom-colored sticks with blue carbons. Uridine ligand is shown as atom-colored sticks with black carbons. (c) Cytidine **2** is converted to *N*-hydroxycytidine **4** by an engineered cytidine aminotransferase followed by acylation using Novozym-435 to give Molnupiravir **1**.

process. First and foremost, the range of synthetic chemistries that can be catalyzed using engineered biocatalysts needs to continue to expand and at a greater rate than is currently the case. In this respect, many approaches are currently being simultaneously explored including (i) genetic code expansion for introduction of unnatural amino acids (e.g., Bayliss-Hillmanase),[147] (ii) repurposing P450 enzymes for carbene/nitrene generation (e.g., cyclopropanation; *N*-insertion in C–H bonds),[148] (iii) hybrid artificial metalloenzymes (e.g., asymmetric imine reduction),[149] and (iv) mining of bio-synthetic pathways for new enzymes (e.g., halogenation).[150] Second, the discovery of new enzyme activities, as outlined above, needs to be coupled closely with smarter methods for reaction screening in order to both initially locate the activity and then subsequently enhance through rounds of protein engineering. The ability to produce enzymes on a small scale via in vitro translation enables panels of biocatalysts to be rapidly "printed" in a format that enables rapid screening via MS-based approaches.[151] Figure 8 summarizes generic examples of structures of where and how biocatalysis can influence biorelevant syntheses.

The recent COVID-19 pandemic has highlighted not only the need to rapidly develop new therapeutics to combat global diseases, but also the ability to manufacture these drugs at a cost that enables all patient populations, including access of low-income economies to the medicine. Using biocatalysts for API manufacture has been shown to reduce costs, particularly for second and third generation manufacture of established

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400–2416

drugs (e.g., Atorvastatin,[152] Paroxetine,[153] Sitagliptin[154]).[155] Biocatalysts are increasingly being used for first generation processes to both accelerate the launch of the drug (e.g., Sitagliptin) as well as provide lower cost routes for manufacture by generic manufacturers under license (e.g., Molnupiravir) (Figure 9).[156]

CONCLUSIONS AND PROSPECTS FOR FUTURE DEVELOPMENT

There is much current interest in the impact of synthesis in drug discovery, but little focus on how (and why) the employment of natural motifs and methods might better be addressed, in spite of the inherent advantages these truly organic/biosynthetic molecules engender. Working with such molecules to better understand and truly expand chemical space should pay dividends in many areas. Not the least, in connecting commonalities in the structure and recognition between biosynthetic, carrier, and target proteins. This should allow the exploration of valuable targets with molecules with properties differentiated from the apparent singularity where "generic passive permeation" (be it via the membrane bilayer or generic protein carriers) appears feasible. It is clear that exogenous natural products, many with important bioactivities, have properties described in a wider chemical space described by size and lipophilicity. Larger, more hydrophilic structures with a propensity for transporter recognition present a validated and precedented paradigm for future antimicrobial research where high throughput methods in "singular space" have failed.[68] Appropriate synthetic and semisynthetic diversity can be achieved through pseudo-natural products and building blocks accessed through biotransformations.

Opportunities exist within current and future developments in computational methods to better research and exploit the potential of natural motifs in design—increased confidence in the predictive outcomes would give impetus to invest time in synthesis. This should be impactful in a variety of ways, including (i) recognizing commonality in protein structures between binding sites in biosynthetic enzymes, carriers, and targets, to (ii) database searching of known ligands or recognition for given/recognized sequences, (iii) prioritization of higher natural product (and PNP) likeness in library generation (exploiting biosynthesised monomers) and virtual searching, and (iv) effective structure hopping and introduction of natural isosteres in hit to lead and lead optimization if hits are from fully synthetic origins. An ultimate goal might be designer drugs from engineered bugs, whereby the drug substance or an advanced intermediate can be biosynthesised from simple feedstocks by design.

The preceding sections fall into the protein-centric world that dominates drug discovery; yet as new modalities emerge,[157] the essence of many of such approaches is based on natural products. Targeting RNA in the Ribosome[158] is an area where natural and semisynthetic drugs have a long pedigree, in particular in the antibiotic field.[159,160] The importance of sugar chemistry and glycomics in disease is receiving increasing attention,[161−163] and the inherent advantages of carbohydrates due to their solubility, specificity, and transporter recognition make this a field ripe for exploitation with natural methods,[164] be it through molecular recognition of functionalized glycosides for activity or transporter recognition or a potential new world where the reading/writing/erasing of glycosyl functionality is exploited.[165]

Understanding and recognizing the common features of binding sites, transporters, and synthesizing enzymes should offer opportunities for defining motifs to pursue, be this validating a target with a natural product or using a natural fragment as the basis for further design. In doing this, the intangible universe of 10^{70} potential molecules could be reduced to a more manageable number, driven by focus on more natural shapes and connectivity more within the constraints of natural building blocks.[166] Paradoxically, despite a lesser number of building blocks and reactions, the inherent complexity and diversity of natural structures would offer validated opportunities to truly explore wider regions of chemical space. This is a manifestation of the Walsh "Think Biologically, Act Chemically" idiom, recently compared to the practices of medicinal chemistry.[167] The logic laid out herein is based on precedent and the ongoing influence of natural products and structures in drugs;[7] there are likely opportunities to test hypotheses in curated sets of molecules and accompanying data within Structure Property Relationships and their impact on efficacy, an area in which computational power should provide insight beyond our quantitative observations with drugs and structural classifications.

The persistence and environmental impact of pharmaceuticals is a topic receiving increasing attention and the wish to "develop pharmaceuticals that better degrade as they reach the environment or break down in the sewage treatment plant"[168] could better be addressed with more natural substances that are better recognized by nature's agents. Microbacteria evolve ways to feast on Xenobiotics, but if this process can be expedited with motifs they inherently recognize, then the 30 years it took to evolve new mechanisms to respond to a molecule such as the herbicide atrazine could considerably be shortened.[169]

The ultimate manifestation of the discussion herein may be theoretical or even hypothetical in the confines of current drug discovery practices, but component concepts have clear validity so the notions and their implications we believe will be impactful in various ways toward the design and development of higher quality molecules with reduced aromaticity and modulated lipophilicity. More natural molecules will increase quality through their inherently improved permeability and solubility; this is a case of investing time and effort in the early stages of drug discovery to reap rewards with improvements in the later stages through more predictability in trials (and thus a greater chance of success, where quality rather than speed demonstrably impacts[170]) and more sustainable manufacturing methods driven by the *transformative power of biocatalysis*.[171] Investing in quality in the early stages of discovery (paying now) will pay dividends with higher quality carried through the process whence the costs of paying later are manifested in poorer translation and higher potential for failure with suboptimal molecules.

AUTHOR INFORMATION

Corresponding Author

Robert J. Young − *Blue Burgundy Ltd, Ampthill, Bedfordshire MK45 2AD, U.K.;* orcid.org/0000-0002-7763-0575; Email: rob@blue-burgundy.co.uk

Authors

Sabine L. Flitsch − *Department of Chemistry, University of Manchester, Manchester Institute of Biotechnology,*

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

Manchester M1 7DN, United Kingdom; orcid.org/0000-0003-3974-646X

Michael Grigalunas − *Department of Chemical Biology, Max-Planck-Institute of Molecular Physiology, 44227 Dortmund, Germany;* orcid.org/0000-0003-0419-8168

Paul D. Leeson − *Paul Leeson Consulting Limited, Nuneaton, Warwickshire CV13 6LZ, U.K.;* orcid.org/0000-0003-0212-3437

Ronald J. Quinn − *Griffith Institute for Drug Discovery, Griffith University, Nathan, Queensland 4111, Australia;* orcid.org/0000-0002-4022-2623

Nicholas J. Turner − *Department of Chemistry, University of Manchester, Manchester Institute of Biotechnology, Manchester M1 7DN, United Kingdom;* orcid.org/0000-0002-8708-0781

Herbert Waldmann − *Department of Chemical Biology, Max-Planck-Institute of Molecular Physiology, 44227 Dortmund, Germany; Faculty of Chemistry and Chemical Biology, Technical University of Dortmund, 44227 Dortmund, Germany;* orcid.org/0000-0002-9606-7247

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.2c00415

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The Australian Research Council Discovery Project Grant DP130102400 to R.J.Q. is gratefully acknowledged.

■ REFERENCES

(1) Brown, D. G.; Bostrom, J. Where Do Recent Small Molecule Clinical Development Candidates Come From? *J. Med. Chem.* **2018**, *61* (21), 9442−9468.
(2) Roughley, S. D.; Jordan, A. M. The medicinal chemist's toolbox: an analysis of reactions used in the pursuit of drug candidates. *J. Med. Chem.* **2011**, *54* (10), 3451−79.
(3) Schneider, N.; Lowe, D. M.; Sayle, R. A.; Tarselli, M. A.; Landrum, G. A. Big Data from Pharmaceutical Patents: A Computational Analysis of Medicinal Chemists' Bread and Butter. *J. Med. Chem.* **2016**, *59* (9), 4385−402.
(4) Voight, E. A.; Greszler, S. N.; Kym, P. R. Fueling the Pipeline via Innovations in Organic Synthesis. *ACS Med. Chem. Lett.* **2021**, *12* (9), 1365−1373.
(5) Brown, D. G.; Bostrom, J. Analysis of Past and Present Synthetic Methodologies on Medicinal Chemistry: Where Have All the New Reactions Gone? *J. Med. Chem.* **2016**, *59* (10), 4443−58.
(6) Saldivar-Gonzalez, F. I.; Aldas-Bulos, V. D.; Medina-Franco, J. L.; Plisson, F. Natural product drug discovery in the artificial intelligence era. *Chem. Sci.* **2022**, *13* (6), 1526−1546.
(7) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83* (3), 770−803.
(8) Bauer, R. A.; Wurst, J. M.; Tan, D. S. Expanding the range of 'druggable' targets with natural product-based libraries: an academic perspective. *Curr. Opin. Chem. Biol.* **2010**, *14* (3), 308−314.
(9) Ertl, P. Substituents of life: The most common substituent patterns present in natural products. *Biorg. Med. Chem.* **2022**, *54*, 116562.
(10) Ertl, P.; Roggo, S.; Schuffenhauer, A. Natural Product-likeness Score and Its Application for Prioritization of Compound Libraries. *J. Chem. Inf. Model.* **2008**, *48* (1), 68−74.
(11) Ertl, P.; Altmann, E.; McKenna, J. M. The Most Common Functional Groups in Bioactive Molecules and How Their Popularity Has Evolved over Time. *J. Med. Chem.* **2020**, *63* (15), 8408−8418.
(12) Chen, Y.; Stork, C.; Hirte, S.; Kirchmair, J. NP-Scout: Machine Learning Approach for the Quantification and Visualization of the Natural Product-Likeness of Small Molecules. *Biomolecules* **2019**, *9* (2), 43.
(13) Sneader, W. *Drug Prototypes and their Exploitation*; John Wiley and Sons Ltd: Chichester, UK, 1996; p 788.
(14) Sneader, W. *Drug Discovery: A History*; Wiley: Chichester, UK, 2005; p 472.
(15) Leeson, P. D.; Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov* **2007**, *6* (11), 881−90.
(16) Leeson, P. D.; Bento, A. P.; Gaulton, A.; Hersey, A.; Manners, E. J.; Radoux, C. J.; Leach, A. R. Target-Based Evaluation of "Drug-Like" Properties and Ligand Efficiencies. *J. Med. Chem.* **2021**, *64* (11), 7210−7230.
(17) Leeson, P. D.; Davis, A. M. Time-related differences in the physical property profiles of oral drugs. *J. Med. Chem.* **2004**, *47* (25), 6338−48.
(18) Proudfoot, J. R. The evolution of synthetic oral drug properties. *Bioorg. Med. Chem. Lett.* **2005**, *15* (4), 1087−90.
(19) Shultz, M. D. Two Decades under the Influence of the Rule of Five and the Changing Properties of Approved Oral Drugs. *J. Med. Chem.* **2019**, *62* (4), 1701−1714.
(20) Ritchie, T. J.; Macdonald, S. J. The impact of aromatic ring count on compound developability–are too many aromatic rings a liability in drug design? *Drug Discov Today* **2009**, *14* (21−22), 1011−20.
(21) Ritchie, T. J.; Macdonald, S. J.; Young, R. J.; Pickett, S. D. The impact of aromatic ring count on compound developability: further insights by examining carbo-and hetero-aromatic and-aliphatic ring types. *Drug Discovery Today* **2011**, *16* (3), 164−171.
(22) 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* (13), 8699−8712.
(23) Wills, T. J.; Lipkus, A. H. Structural Approach to Assessing the Innovativeness of New Drugs Finds Accelerating Rate of Innovation. *ACS Med. Chem. Lett.* **2020**, *11* (11), 2114−2119.
(24) Karageorgis, G.; Foley, D. J.; Laraia, L.; Brakmann, S.; Waldmann, H. Pseudo Natural Products—Chemical Evolution of Natural Product Structure. *Angew. Chem., Int. Ed.* **2021**, *60* (29), 15705−15723.
(25) Leeson, P. D. Impact of Physicochemical Properties on Dose and Hepatotoxicity of Oral Drugs. *Chem. Res. Toxicol.* **2018**, *31* (6), 494−505.
(26) Bostrom, J.; Brown, D. G.; Young, R. J.; Keseru, G. M. Expanding the medicinal chemistry synthetic toolbox. *Nat. Rev. Drug Discov* **2018**, *17* (12), 922.
(27) Reymond, J. L. The chemical space project. *Acc. Chem. Res.* **2015**, *48* (3), 722−30.
(28) Chow, S. Y.; Nelson, A. Embarking on a Chemical Space Odyssey. *J. Med. Chem.* **2017**, *60* (9), 3591−3593.
(29) Goodnow, R. A., Jr.; Dumelin, C. E.; Keefe, A. D. DNA-encoded chemistry: enabling the deeper sampling of chemical space. *Nat. Rev. Drug Discov* **2017**, *16* (2), 131−147.
(30) Irwin, J. J.; Gaskins, G.; Sterling, T.; Mysinger, M. M.; Keiser, M. J. Predicted Biological Activity of Purchasable Chemical Space. *J. Chem. Inf. Model.* **2018**, *58* (1), 148−164.
(31) Dobson, C. M. Chemical space and biology. *Nature* **2004**, *432* (7019), 824−828.
(32) Volochnyuk, D. M.; Ryabukhin, S. V.; Moroz, Y. S.; Savych, O.; Chuprina, A.; Horvath, D.; Zabolotna, Y.; Varnek, A.; Judd, D. B. Evolution of commercially available compounds for HTS. *Drug Discovery Today* **2019**, *24* (2), 390−402.
(33) Dragovich, P. S.; Haap, W.; Mulvihill, M. M.; Plancher, J. M.; Stepan, A. F. Small-Molecule Lead-Finding Trends across the Roche and Genentech Research Organizations. *J. Med. Chem.* **2022**, *65* (4), 3606−3615.

(34) Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. The Design of Leadlike Combinatorial Libraries. *Angew. Chem., Int. Ed. Engl.* **1999**, *38* (24), 3743−3748.
(35) Feher, M.; Schmidt, J. M. Property Distributions: Differences between Drugs, Natural Products, and Molecules from Combinatorial Chemistry. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (1), 218−227.
(36) Hann, M. M.; Keseru, G. M. Finding the sweet spot: the role of nature and nurture in medicinal chemistry. *Nat. Rev. Drug Discov* **2012**, *11* (5), 355−65.
(37) Steadman, D.; Atkinson, B. N.; Zhao, Y.; Willis, N. J.; Frew, S.; Monaghan, A.; Patel, C.; Armstrong, E.; Costelloe, K.; Magno, L.; Bictash, M.; Jones, E. Y.; Fish, P. V.; Svensson, F. Virtual Screening Directly Identifies New Fragment-Sized Inhibitors of Carboxylesterase Notum with Nanomolar Activity. *J. Med. Chem.* **2022**, *65* (1), 562−578.
(38) Sadybekov, A. A.; Sadybekov, A. V.; Liu, Y.; Iliopoulos-Tsoutsouvas, C.; Huang, X.-P.; Pickett, J.; Houser, B.; Patel, N.; Tran, N. K.; Tong, F.; Zvonok, N.; Jain, M. K.; Savych, O.; Radchenko, D. S.; Nikas, S. P.; Petasis, N. A.; Moroz, Y. S.; Roth, B. L.; Makriyannis, A.; Katritch, V. Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* **2022**, *601* (7893), 452−459.
(39) Grygorenko, O. O.; Radchenko, D. S.; Dziuba, I.; Chuprina, A.; Gubina, K. E.; Moroz, Y. S. Generating Multibillion Chemical Space of Readily Accessible Screening Compounds. *iScience* **2020**, *23* (11), 101681−101681.
(40) Zhang, Y.; Clark, M. A. Design concepts for DNA-Encoded Library synthesis. *Biorg. Med. Chem.* **2021**, *41*, 116189.
(41) Shi, Y.; Wu, Y.-r.; Yu, J.-q.; Zhang, W.-n.; Zhuang, C.-l. DNA-encoded libraries (DELs): a review of on-DNA chemistries and their output. *RSC Adv.* **2021**, *11* (4), 2359−2376.
(42) Tomberg, A.; Boström, J. Can easy chemistry produce complex, diverse, and novel molecules? *Drug Discovery Today* **2020**, *25* (12), 2174−2181.
(43) Young, R. J. Physical Properties in Drug Design. In *Tactics in Contemporary Drug Design*; Meanwell, N. A., Ed.; Springer: Berlin, Heidelberg, 2014; pp 1−68.
(44) Waring, M. J. Defining optimum lipophilicity and molecular weight ranges for drug candidates-Molecular weight dependent lower logD limits based on permeability. *Bioorg. Med. Chem. Lett.* **2009**, *19* (10), 2844−51.
(45) Johnson, T. W.; Dress, K. R.; Edwards, M. Using the Golden Triangle to optimize clearance and oral absorption. *Bioorg. Med. Chem. Lett.* **2009**, *19* (19), 5560−4.
(46) DeGoey, D. A.; Chen, H. J.; Cox, P. B.; Wendt, M. D. Beyond the Rule of 5: Lessons Learned from AbbVie's Drugs and Compound Collection. *J. Med. Chem.* **2018**, *61* (7), 2636−2651.
(47) Hansch, C.; Bjorkroth, J. P.; Leo, A. Hydrophobicity and central nervous system agents: on the principle of minimal hydrophobicity in drug design. *J. Pharm. Sci.* **1987**, *76* (9), 663−87.
(48) Hopkins, A. L.; Keseru, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discov* **2014**, *13* (2), 105−21.
(49) Johnson, T. W.; Gallego, R. A.; Edwards, M. P. Lipophilic Efficiency as an Important Metric in Drug Design. *J. Med. Chem.* **2018**, *61* (15), 6401−6420.
(50) Young, R. J.; Green, D. V.; Luscombe, C. N.; Hill, A. P. Getting physical in drug discovery II: the impact of chromatographic hydrophobicity measurements and aromaticity. *Drug Discov Today* **2011**, *16* (17−18), 822−30.
(51) Tinworth, C. P.; Young, R. J. Facts, Patterns, and Principles in Drug Discovery: Appraising the Rule of 5 with Measured Physicochemical Data. *J. Med. Chem.* **2020**, *63* (18), 10091−10108.
(52) Sugano, K.; Kansy, M.; Artursson, P.; Avdeef, A.; Bendels, S.; Di, L.; Ecker, G. F.; Faller, B.; Fischer, H.; Gerebtzoff, G.; Lennernaes, H.; Senner, F. Coexistence of passive and carrier-mediated processes in drug transport. *Nat. Rev. Drug Discov* **2010**, *9* (8), 597−614.
(53) Dobson, P. D.; Kell, D. B. Carrier-mediated cellular uptake of pharmaceutical drugs: an exception or the rule? *Nat. Rev. Drug Discov* **2008**, *7* (3), 205−20.
(54) Kell, D. B.; Dobson, P. D.; Oliver, S. G. Pharmaceutical drug transport: the issues and the implications that it is essentially carrier-mediated only. *Drug Discov Today* **2011**, *16* (15−16), 704−14.
(55) Di, L.; Artursson, P.; Avdeef, A.; Ecker, G. F.; Faller, B.; Fischer, H.; Houston, J. B.; Kansy, M.; Kerns, E. H.; Kramer, S. D.; Lennernas, H.; Sugano, K. Evidence-based approach to assess passive diffusion and carrier-mediated drug transport. *Drug Discov Today* **2012**, *17* (15−16), 905−12.
(56) Superti-Furga, G.; Lackner, D.; Wiedmer, T.; Ingles-Prieto, A.; Barbosa, B.; Girardi, E.; Goldmann, U.; Gurtl, B.; Klavins, K.; Klimek, C.; Lindinger, S.; Lineiro-Retes, E.; Muller, A. C.; Onstein, S.; Redinger, G.; Reil, D.; Sedlyarov, V.; Wolf, G.; Crawford, M.; Everley, R.; Hepworth, D.; Liu, S.; Noell, S.; Piotrowski, M.; Stanton, R.; Zhang, H.; Corallino, S.; Faedo, A.; Insidioso, M.; Maresca, G.; Redaelli, L.; Sassone, F.; Scarabottolo, L.; Stucchi, M.; Tarroni, P.; Tremolada, S.; Batoulis, H.; Becker, A.; Bender, E.; Chang, Y. N.; Ehrmann, A.; Muller-Fahrnow, A.; Putter, V.; Zindel, D.; Hamilton, B.; Lenter, M.; Santacruz, D.; Viollet, C.; Whitehurst, C.; Johnsson, K.; Leippe, P.; Baumgarten, B.; Chang, L.; Ibig, Y.; Pfeifer, M.; Reinhardt, J.; Schonbett, J.; Selzer, P.; Seuwen, K.; Bettembourg, C.; Biton, B.; Czech, J.; de Foucauld, H.; Didier, M.; Licher, T.; Mikol, V.; Pommereau, A.; Puech, F.; Yaligara, V.; Edwards, A.; Bongers, B. J.; Heitman, L. H.; AP, I. J.; Sijben, H. J.; van Westen, G. J. P.; Grixti, J.; Kell, D. B.; Mughal, F.; Swainston, N.; Wright-Muelas, M.; Bohstedt, T.; Burgess-Brown, N.; Carpenter, L.; Durr, K.; Hansen, J.; Scacioc, A.; Banci, G.; Colas, C.; Digles, D.; Ecker, G.; Fuzi, B.; Gamsjager, V.; Grandits, M.; Martini, R.; Troger, F.; Altermatt, P.; Doucerain, C.; Durrenberger, F.; Manolova, V.; Steck, A. L.; Sundstrom, H.; Wilhelm, M.; Steppan, C. M. The RESOLUTE consortium: unlocking SLC transporters for drug discovery. *Nat. Rev. Drug Discovery* **2020**, *19* (7), 429−430.
(57) O'Hagan, S.; Kell, D. B. Consensus rank orderings of molecular fingerprints illustrate the most genuine similarities between marketed drugs and small endogenous human metabolites, but highlight exogenous natural products as the most important 'natural' drug transporter substrates. *ADMET and DMPK* **2017**, *5* (2), 85−125.
(58) O'Hagan, S.; Swainston, N.; Handl, J.; Kell, D. B. A 'rule of 0.5' for the metabolite-likeness of approved pharmaceutical drugs. *Metabolomics* **2015**, *11* (2), 323−339.
(59) Hann, M. M. Molecular obesity, potency and other addictions in drug discovery. *Med. Chem. Comm.* **2011**, *2*, 349−355.
(60) Meanwell, N. A. Improving drug candidates by design: a focus on physicochemical properties as a means of improving compound disposition and safety. *Chem. Res. Toxicol.* **2011**, *24* (9), 1420−56.
(61) Meanwell, N. A. Improving Drug Design: An Update on Recent Applications of Efficiency Metrics, Strategies for Replacing Problematic Elements, and Compounds in Nontraditional Drug Space. *Chem. Res. Toxicol.* **2016**, *29* (4), 564−616.
(62) Hann, M. M. Lessons in Transcellular Membrane Transport Re-Learned. *J. Pharm. Sci.* **2021**, *110* (1), 548−551.
(63) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3−25.
(64) O'Hagan, S.; Kell, D. B. Generation of a Small Library of Natural Products Designed to Cover Chemical Space Inexpensively. *Pharmaceutical Frontiers* **2019**, *1* (1), e190005.
(65) Fullam, E.; Young, R. J. Physicochemical properties and Mycobacterium tuberculosis transporters: keys to efficacious anti-tubercular drugs? *RSC Medicinal Chemistry* **2021**, *12* (1), 43−56.
(66) Motamen, S.; Quinn, R. J. Analysis of Approaches to Anti-tuberculosis Compounds. *ACS Omega* **2020**, *5* (44), 28529−28540.
(67) Hutchings, M. I.; Truman, A. W.; Wilkinson, B. Antibiotics: past, present and future. *Curr. Opin. Microbiol.* **2019**, *51*, 72−80.
(68) Tommasi, R.; Brown, D. G.; Walkup, G. K.; Manchester, J. I.; Miller, A. A. ESKAPEing the labyrinth of antibacterial discovery. *Nat. Rev. Drug Discovery* **2015**, *14* (8), 529−542.
(69) Iyer, R.; Sylvester, M. A.; Velez-Vega, C.; Tommasi, R.; Durand-Reville, T. F.; Miller, A. A. Whole-Cell-Based Assay To

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

Evaluate Structure Permeation Relationships for Carbapenem Passage through the Pseudomonas aeruginosa Porin OprD. *ACS Infectious Diseases* **2017**, *3* (4), 310−319.
(70) Durand-Reville, T. F.; Miller, A. A.; O'Donnell, J. P.; Wu, X.; Sylvester, M. A.; Guler, S.; Iyer, R.; Shapiro, A. B.; Carter, N. M.; Velez-Vega, C.; Moussa, S. H.; McLeod, S. M.; Chen, A.; Tanudra, A. M.; Zhang, J.; Comita-Prevoir, J.; Romero, J. A.; Huynh, H.; Ferguson, A. D.; Horanyi, P. S.; Mayclin, S. J.; Heine, H. S.; Drusano, G. L.; Cummings, J. E.; Slayden, R. A.; Tommasi, R. A. Rational design of a new antibiotic class for drug-resistant infections. *Nature* **2021**, *597* (7878), 698−702.
(71) Lovering, F.; Bikker, J.; Humblet, C. Escape from flatland: increasing saturation as an approach to improving clinical success. *J. Med. Chem.* **2009**, *52* (21), 6752−6.
(72) Lovering, F. Escape from Flatland 2: complexity and promiscuity. *MedChemComm* **2013**, *4* (3), 515−519.
(73) Freire, E. Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov Today* **2008**, *13* (19−20), 869−74.
(74) Ladbury, J. E.; Klebe, G.; Freire, E. Adding calorimetric data to decision making in lead discovery: a hot tip. *Nat. Rev. Drug Discov* **2010**, *9* (1), 23−7.
(75) Subbaiah, M. A. M.; Meanwell, N. A. Bioisosteres of the Phenyl Ring: Recent Strategic Applications in Lead Optimization and Drug Design. *J. Med. Chem.* **2021**, *64* (19), 14046−14128.
(76) Chakravorty, S. J.; Chan, J.; Greenwood, M. N.; Popa-Burke, I.; Remlinger, K. S.; Pickett, S. D.; Green, D. V. S.; Fillmore, M. C.; Dean, T. W.; Luengo, J. I.; Macarron, R. Nuisance Compounds, PAINS Filters, and Dark Chemical Matter in the GSK HTS Collection. *SLAS Discov* **2018**, *23* (6), 532−545.
(77) Kellenberger, E.; Hofmann, A.; Quinn, R. J. Similar interactions of natural products with biosynthetic enzymes and therapeutic targets could explain why nature produces such a large proportion of existing drugs. *Nat. Prod Rep* **2011**, *28* (9), 1483−92.
(78) Hert, J.; Irwin, J. J.; Laggner, C.; Keiser, M. J.; Shoichet, B. K. Quantifying biogenic bias in screening libraries. *Nat. Chem. Biol.* **2009**, *5* (7), 479−483.
(79) Newman, D. J.; Cragg, G. M. Natural Product Scaffolds of Value in Medicinal Chemistry. In *Privileged Scaffolds in Medicinal Chemistry: Design, Synthesis, Evaluation*; The Royal Society of Chemistry: 2016; Ch.13, pp 348−378.
(80) Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov* **2015**, *14* (2), 111−29.
(81) Wilson, B. A. P.; Thornburg, C. C.; Henrich, C. J.; Grkovic, T.; O'Keefe, B. R. Creating and screening natural product libraries. *Nat. Prod. Rep.* **2020**, *37* (7), 893−918.
(82) Pye, C. R.; Bertin, M. J.; Lokey, R. S.; Gerwick, W. H.; Linington, R. G. Retrospective analysis of natural products provides insights for future discovery trends. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (22), 5601−5606.
(83) Grigalunas, M.; Burhop, A.; Christoforow, A.; Waldmann, H. Pseudo-natural products and natural product-inspired methods in chemical biology and drug discovery. *Curr. Opin. Chem. Biol.* **2020**, *56*, 111−118.
(84) Gao, M.; Skolnick, J. A Comprehensive Survey of Small-Molecule Binding Pockets in Proteins. *PLOS Computational Biology* **2013**, *9* (10), e1003302.
(85) Koch, M. A.; Wittenberg, L.-O.; Basu, S.; Jeyaraj, D. A.; Gourzoulidou, E.; Reinecke, K.; Odermatt, A.; Waldmann, H. Compound library development guided by protein structure similarity clustering and natural product structure. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (48), 16721−16726.
(86) Koch, M. A.; Schuffenhauer, A.; Scheck, M.; Wetzel, S.; Casaulta, M.; Odermatt, A.; Ertl, P.; Waldmann, H. Charting biologically relevant chemical space: A structural classification of natural products (SCONP). *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (48), 17272−17277.
(87) Wetzel, S.; Bon, R. S.; Kumar, K.; Waldmann, H. Biology-Oriented Synthesis. *Angew. Chem., Int. Ed.* **2011**, *50* (46), 10800−10826.
(88) van Hattum, H.; Waldmann, H. Biology-Oriented Synthesis: Harnessing the Power of Evolution. *J. Am. Chem. Soc.* **2014**, *136* (34), 11853−11859.
(89) Karageorgis, G.; Waldmann, H. Guided by Evolution: Biology-Oriented Synthesis of Bioactive Compound Classes. *Synthesis* **2019**, *51* (01), 55−66.
(90) Arve, L.; Voigt, T.; Waldmann, H. Charting Biological and Chemical Space: PSSC and SCONP as Guiding Principles for the Development of Compound Collections Based on Natural Product Scaffolds. *QSAR & Combinatorial Science* **2006**, *25*, 449−456.
(91) Murray, C. W.; Rees, D. C. The rise of fragment-based drug discovery. *Nat. Chem.* **2009**, *1* (3), 187−192.
(92) Roughley, S. D.; Hubbard, R. E. How Well Can Fragments Explore Accessed Chemical Space? A Case Study from Heat Shock Protein 90. *J. Med. Chem.* **2011**, *54* (12), 3989−4005.
(93) Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discovery* **2016**, *15* (9), 605−619.
(94) Babaoglu, K.; Shoichet, B. K. Deconstructing fragment-based inhibitor discovery. *Nat. Chem. Biol.* **2006**, *2* (12), 720−723.
(95) Siegel, M. G.; Vieth, M. Drugs in other drugs: a new look at drugs as fragments. *Drug Discovery Today* **2007**, *12* (1), 71−79.
(96) Hubbard, R. E.; Chen, I.; Davis, B. Informatics and modeling challenges in fragment-based drug discovery. *Curr. Opin. Drug Discovery Dev.* **2007**, *10* (3), 289−97.
(97) Hung, A. W.; Ramek, A.; Wang, Y.; Kaya, T.; Wilson, J. A.; Clemons, P. A.; Young, D. W. Route to three-dimensional fragments using diversity-oriented synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (17), 6799−6804.
(98) Over, B.; Wetzel, S.; Grutter, C.; Nakai, Y.; Renner, S.; Rauh, D.; Waldmann, H. Natural-product-derived fragments for fragment-based ligand discovery. *Nat. Chem.* **2013**, *5* (1), 21−8.
(99) Vu, H.; Pedro, L.; Mak, T.; McCormick, B.; Rowley, J.; Liu, M.; Di Capua, A.; Williams-Noonan, B.; Pham, N. B.; Pouwer, R.; Nguyen, B.; Andrews, K. T.; Skinner-Adams, T.; Kim, J.; Hol, W. G. J.; Hui, R.; Crowther, G. J.; Van Voorhis, W. C.; Quinn, R. J. Fragment-Based Screening of a Natural Product Library against 62 Potential Malaria Drug Targets Employing Native Mass Spectrometry. *ACS infectious diseases* **2018**, *4* (4), 431−444.
(100) Pascolutti, M.; Campitelli, M.; Nguyen, B.; Pham, N.; Gorse, A.-D.; Quinn, R. J. Capturing Nature's Diversity. *PLoS One* **2015**, *10* (4), No. e0120942.
(101) Karageorgis, G.; Foley, D. J.; Laraia, L.; Waldmann, H. Principle and design of pseudo-natural products. *Nat. Chem.* **2020**, *12* (3), 227−235.
(102) Grigalunas, M.; Burhop, A.; Zinken, S.; Pahl, A.; Gally, J.-M.; Wild, N.; Mantel, Y.; Sievers, S.; Foley, D. J.; Scheel, R.; Strohmann, C.; Antonchick, A. P.; Waldmann, H. Natural product fragment combination to performance-diverse pseudo-natural products. *Nat. Commun.* **2021**, *12* (1), 1883.
(103) Liu, J.; Cremosnik, G. S.; Otte, F.; Pahl, A.; Sievers, S.; Strohmann, C.; Waldmann, H. Design, Synthesis, and Biological Evaluation of Chemically and Biologically Diverse Pyrroquinoline Pseudo Natural Products. *Angew. Chem., Int. Ed.* **2021**, *60* (9), 4648−4656.
(104) Ertl, P.; Schuhmann, T. Cheminformatics Analysis of Natural Product Scaffolds: Comparison of Scaffolds Produced by Animals, Plants, Fungi and Bacteria. *Mol. Inform.* **2020**, *39* (11), 2000017.
(105) Gally, J.-M.; Pahl, A.; Czodrowski, P.; Waldmann, H. Pseudonatural Products Occur Frequently in Biologically Relevant Compounds. *J. Chem. Inf. Model.* **2021**, *61* (11), 5458−5468.
(106) Suzuki, K.; Yasui, Y. Hybrid Natural Products. In *Natural Products in Medicinal Chemistry*; Wiley, 2014; pp 441−472.
(107) Tietze, L. F.; Bell, H. P.; Chandrasekhar, S. Natural Product Hybrids as New Leads for Drug Discovery. *Angew. Chem., Int. Ed.* **2003**, *42* (34), 3996−4028.

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

(108) Grigalunas, M.; Brakmann, S.; Waldmann, H. Chemical Evolution of Natural Product Structure. *J. Am. Chem. Soc.* **2022**, *144* (8), 3314−3329.
(109) Burhop, A.; Bag, S.; Grigalunas, M.; Woitalla, S.; Bodenbinder, P.; Brieger, L.; Strohmann, C.; Pahl, A.; Sievers, S.; Waldmann, H. Synthesis of Indofulvin Pseudo-Natural Products Yields a New Autophagy Inhibitor Chemotype. *Adv. Sci.* **2021**, *8* (19), 2102042.
(110) Liu, J.; Flegel, J.; Otte, F.; Pahl, A.; Sievers, S.; Strohmann, C.; Waldmann, H. Combination of Pseudo-Natural Product Design and Formal Natural Product Ring Distortion Yields Stereochemically and Biologically Diverse Pseudo-Sesquiterpenoid Alkaloids. *Angew. Chem., Int. Ed.* **2021**, *60* (39), 21384−21395.
(111) Niggemeyer, G.; Knyazeva, A.; Gasper, R.; Corkery, D.; Bodenbinder, P.; Holstein, J. J.; Sievers, S.; Wu, Y.-W.; Waldmann, H. Synthesis of 20-Membered Macrocyclic Pseudo-Natural Products Yields Inducers of LC3 Lipidation. *Angew. Chem., Int. Ed.* **2022**, *61* (11), e202114328.
(112) Wu, G.; Yu, G.; Yu, Y.; Yang, S.; Duan, Z.; Wang, W.; Liu, Y.; Yu, R.; Li, J.; Zhu, T.; Gu, Q.; Li, D. Chemoreactive-Inspired Discovery of Influenza A Virus Dual Inhibitor to Block Hemagglutinin-Mediated Adsorption and Membrane Fusion. *J. Med. Chem.* **2020**, *63* (13), 6924−6940.
(113) Whitmarsh-Everiss, T.; Olsen, A. H.; Laraia, L. Identification of Inhibitors of Cholesterol Transport Proteins Through the Synthesis of a Diverse, Sterol-Inspired Compound Collection. *Angew. Chem., Int. Ed.* **2021**, *60* (51), 26755−26761.
(114) Cunha, M. R.; Bhardwaj, R.; Carrel, A. L.; Lindinger, S.; Romanin, C.; Parise-Filho, R.; Hediger, M. A.; Reymond, J.-L. Natural product inspired optimization of a selective TRPV6 calcium channel inhibitor. *RSC Medicinal Chemistry* **2020**, *11* (9), 1032−1040.
(115) Zhu, H.; Cai, Y.; Ma, S.; Futamura, Y.; Li, J.; Zhong, W.; Zhang, X.; Osada, H.; Zou, H. Privileged Biorenewable Secologanin-Based Diversity-Oriented Synthesis for Pseudo-Natural Alkaloids: Uncovering Novel Neuroprotective and Antimalarial Frameworks. *ChemSusChem* **2021**, *14* (23), 5320−5327.
(116) Taylor, R. D.; MacCoss, M.; Lawson, A. D. G. Rings in Drugs. *J. Med. Chem.* **2014**, *57* (14), 5845−5859.
(117) Ertl, P. Magic Rings: Navigation in the Ring Chemical Space Guided by the Bioactive Rings. *J. Chem. Inf. Model.* **2022**, *62*, 2164.
(118) McArdle, B. M.; Campitelli, M. R.; Quinn, R. J. A Common Protein Fold Topology Shared by Flavonoid Biosynthetic Enzymes and Therapeutic Targets. *J. Nat. Prod.* **2006**, *69* (1), 14−17.
(119) Sturm, N.; Quinn, R. J.; Kellenberger, E. Similarity between Flavonoid Biosynthetic Enzymes and Flavonoid Protein Targets Captured by Three-Dimensional Computing Approach. *Planta Med.* **2015**, *81* (6), 467−73.
(120) Sturm, N.; Rognan, D.; Quinn, R. J.; Kellenberger, E. Comparing atom-based with residue-based descriptors in predicting binding site similarity: do backbone atoms matter? *Future Med. Chem.* **2016**, *8* (15), 1871−1885.
(121) Sturm, N.; Quinn, R. J.; Kellenberger, E. Structural Searching of Biosynthetic Enzymes to Predict Protein Targets of Natural Products. *Planta Med.* **2018**, *84* (05), 304−310.
(122) Karageorgis, G.; Reckzeh, E. S.; Ceballos, J.; Schwalfenberg, M.; Sievers, S.; Ostermann, C.; Pahl, A.; Ziegler, S.; Waldmann, H. Chromopynones are pseudo natural product glucose uptake inhibitors targeting glucose transporters GLUT-1 and −3. *Nat. Chem.* **2018**, *10* (11), 1103−1111.
(123) Ceballos, J.; Schwalfenberg, M.; Karageorgis, G.; Reckzeh, E. S.; Sievers, S.; Ostermann, C.; Pahl, A.; Sellstedt, M.; Nowacki, J.; Carnero Corrales, M. A.; Wilke, J.; Laraia, L.; Tschapalda, K.; Metz, M.; Sehr, D. A.; Brand, S.; Winklhofer, K.; Janning, P.; Ziegler, S.; Waldmann, H. Synthesis of Indomorphan Pseudo-Natural Product Inhibitors of Glucose Transporters GLUT-1 and −3. *Angew. Chem., Int. Ed.* **2019**, *58* (47), 17016−17025.
(124) Sternicki, L. M.; Nonomiya, J.; Liu, M.; Mulvihill, M. M.; Quinn, R. J. Native Mass Spectrometry for the Study of PROTAC GNE-987-Containing Ternary Complexes. *ChemMedChem.* **2021**, *16* (14), 2206−2210.
(125) Beveridge, R.; Kessler, D.; Rumpel, K.; Ettmayer, P.; Meinhart, A.; Clausen, T. Native Mass Spectrometry Can. Effectively Predict PROTAC Efficacy. *ACS Central Science* **2020**, *6* (7), 1223−1230.
(126) Song, J. H.; Wagner, N. D.; Yan, J.; Li, J.; Huang, R. Y. C.; Balog, A. J.; Newitt, J. A.; Chen, G.; Gross, M. L. Native mass spectrometry and gas-phase fragmentation provide rapid and in-depth topological characterization of a PROTAC ternary complex. *Cell Chemical Biology* **2021**, *28* (10), 1528−1538.
(127) Caron, G.; Kihlberg, J.; Ermondi, G. Intramolecular hydrogen bonding: An opportunity for improved design in medicinal chemistry. *Med. Res. Rev.* **2019**, *39* (5), 1707−1729.
(128) Danelius, E.; Poongavanam, V.; Peintner, S.; Wieske, L. H. E.; Erdélyi, M.; Kihlberg, J. Solution Conformations Explain the Chameleonic Behaviour of Macrocyclic Drugs. *Chemistry − A European Journal* **2020**, *26* (23), 5231−5244.
(129) Goetz, G. H.; Philippe, L.; Shapiro, M. J. EPSA: A Novel Supercritical Fluid Chromatography Technique Enabling the Design of Permeable Cyclic Peptides. *ACS Med. Chem. Lett.* **2014**, *5* (10), 1167−72.
(130) Galanie, S.; Entwistle, D.; Lalonde, J. Engineering biosynthetic enzymes for industrial natural product synthesis. *Nat. Prod. Rep.* **2020**, *37* (8), 1122−1143.
(131) Kalkreuter, E.; Williams, G. J. Engineering enzymatic assembly lines for the production of new antimicrobials. *Curr. Opin. Microbiol.* **2018**, *45*, 140−148.
(132) Hohlman, R. M.; Newmister, S. A.; Sanders, J. N.; Khatri, Y.; Li, S.; Keramati, N. R.; Lowell, A. N.; Houk, K. N.; Sherman, D. H. Structural Diversification of Hapalindole and Fischerindole Natural Products via Cascade Biocatalysis. *ACS Catal.* **2021**, *11* (8), 4670−4681.
(133) Devine, P. N.; Howard, R. M.; Kumar, R.; Thompson, M. P.; Truppo, M. D.; Turner, N. J. Extending the application of biocatalysis to meet the challenges of drug development. *Nature Reviews Chemistry* **2018**, *2* (12), 409−421.
(134) Bell, E. L.; Finnigan, W.; France, S. P.; Green, A. P.; Hayes, M. A.; Hepworth, L. J.; Lovelock, S. L.; Niikura, H.; Osuna, S.; Romero, E.; Ryan, K. S.; Turner, N. J.; Flitsch, S. L. Biocatalysis. *Nature Reviews Methods Primers* **2021**, *1* (1), 46.
(135) Arnold, F. H. Directed Evolution: Bringing New Chemistry to Life. *Angew. Chem., Int. Ed.* **2018**, *57* (16), 4143−4148.
(136) Ascaso-Alegre, C.; Mangas-Sanchez, J. Construction of Chemoenzymatic Linear Cascades for the Synthesis of Chiral Compounds. *Eur. J. Org. Chem.* **2022**, *2022* (22), e202200093.
(137) Sheldon, R. A.; Brady, D. Streamlining Design, Engineering, and Applications of Enzymes for Sustainable Biocatalysis. *ACS Sustainable Chem. Eng.* **2021**, *9* (24), 8032−8052.
(138) Finnigan, W.; Hepworth, L. J.; Flitsch, S. L.; Turner, N. J. RetroBioCat as a computer-aided synthesis planning tool for biocatalytic reactions and cascades. *Nature Catalysis* **2021**, *4* (2), 98−104.
(139) Truppo, M. D. Biocatalysis in the Pharmaceutical Industry: The Need for Speed. *ACS Med. Chem. Lett.* **2017**, *8* (5), 476−480.
(140) Romero, E.; Jones, B. S.; Hogg, B. N.; Rue Casamajo, A.; Hayes, M. A.; Flitsch, S. L.; Turner, N. J.; Schnepel, C. Enzymatic Late-Stage Modifications: Better Late Than Never. *Angew. Chem., Int. Ed. Engl.* **2021**, *60* (31), 16824−16855.
(141) Ramsden, J. I.; Cosgrove, S. C.; Turner, N. J. Is it time for biocatalysis in fragment-based drug discovery? *Chemical Science* **2020**, *11* (41), 11104−11112.
(142) Huffman, M. A.; Fryszkowska, A.; Alvizo, O.; Borra-Garske, M.; Campos, K. R.; Canada, K. A.; Devine, P. N.; Duan, D.; Forstater, J. H.; Grosser, S. T.; Halsey, H. M.; Hughes, G. J.; Jo, J.; Joyce, L. A.; Kolev, J. N.; Liang, J.; Maloney, K. M.; Mann, B. F.; Marshall, N. M.; McLaughlin, M.; Moore, J. C.; Murphy, G. S.; Nawrat, C. C.; Nazor, J.; Novick, S.; Patel, N. R.; Rodriguez-Granillo, A.; Robaire, S. A.; Sherer, E. C.; Truppo, M. D.; Whittaker, A. M.; Verma, D.; Xiao, L.; Xu, Y.; Yang, H. Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **2019**, *366* (6470), 1255−1259.

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416

(143) Schrittwieser, J. H.; Velikogne, S.; Hall, M.; Kroutil, W. Artificial Biocatalytic Linear Cascades for Preparation of Organic Molecules. *Chem. Rev.* **2018**, *118* (1), 270−348.
(144) Fair, R. J.; Walsh, R. T.; Hupp, C. D. The expanding reaction toolkit for DNA-encoded libraries. *Bioorg. Med. Chem. Lett.* **2021**, *51*, 128339.
(145) Thomas, B.; Lu, X.; Birmingham, W. R.; Huang, K.; Both, P.; Reyes Martinez, J. E.; Young, R. J.; Davie, C. P.; Flitsch, S. L. Application of Biocatalysis to on-DNA Carbohydrate Library Synthesis. *ChemBioChem.* **2017**, *18* (9), 858−863.
(146) Fryszkowska, A.; Devine, P. N. Biocatalysis in drug discovery and development. *Curr. Opin. Chem. Biol.* **2020**, *55*, 151−160.
(147) Crawshaw, R.; Crossley, A. E.; Johannissen, L.; Burke, A. J.; Hay, S.; Levy, C.; Baker, D.; Lovelock, S. L.; Green, A. P. Engineering an efficient and enantioselective enzyme for the Morita−Baylis−Hillman reaction. *Nat. Chem.* **2022**, *14* (3), 313−320.
(148) Chandgude, A. L.; Ren, X.; Fasan, R. Stereodivergent Intramolecular Cyclopropanation Enabled by Engineered Carbene Transferases. *J. Am. Chem. Soc.* **2019**, *141* (23), 9145−9150.
(149) Davis, H. J.; Ward, T. R. Artificial Metalloenzymes: Challenges and Opportunities. *ACS Central Science* **2019**, *5* (7), 1120−1136.
(150) Fisher, B. F.; Snodgrass, H. M.; Jones, K. A.; Andorfer, M. C.; Lewis, J. C. Site-Selective C−H Halogenation Using Flavin-Dependent Halogenases Identified via Family-Wide Activity Profiling. *ACS Central Science* **2019**, *5* (11), 1844−1856.
(151) Schmid-Dannert, C.; López-Gallego, F. Advances and opportunities for the design of self-sufficient and spatially organized cell-free biocatalytic systems. *Curr. Opin. Chem. Biol.* **2019**, *49*, 97−104.
(152) Luo, Y.; Chen, Y.; Ma, H.; Tian, Z.; Zhang, Y.; Zhang, J. Enhancing the biocatalytic manufacture of the key intermediate of atorvastatin by focused directed evolution of halohydrin dehalogenase. *Sci. Rep.* **2017**, *7* (1), 42064.
(153) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the third wave of biocatalysis. *Nature* **2012**, *485* (7397), 185−194.
(154) Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; Devine, P. N.; Huisman, G. W.; Hughes, G. J. Biocatalytic Asymmetric Synthesis of Chiral Amines from Ketones Applied to Sitagliptin Manufacture. *Science* **2010**, *329* (5989), 305−309.
(155) Ar, A. Biotransformations in Drug Synthesis: A Green and Powerful Tool for Medicinal Chemistry. *Journal of Medicinal Chemistry and Drug Design* **2018**, *1* (1), 1−7.
(156) Burke, A. J.; Birmingham, W. R.; Zhuo, Y.; Thorpe, T. W.; Zucoloto da Costa, B.; Crawshaw, R.; Rowles, I.; Finnigan, J. D.; Young, C.; Holgate, G. M.; Muldowney, M. P.; Charnock, S. J.; Lovelock, S. L.; Turner, N. J.; Green, A. P. An Engineered Cytidine Deaminase for Biocatalytic Production of a Key Intermediate of the Covid-19 Antiviral Molnupiravir. *J. Am. Chem. Soc.* **2022**, *144* (9), 3761−3765.
(157) Valeur, E.; Gueret, S. M.; Adihou, H.; Gopalakrishnan, R.; Lemurell, M.; Waldmann, H.; Grossmann, T. N.; Plowright, A. T. New Modalities for Challenging Targets in Drug Discovery. *Angew. Chem., Int. Ed. Engl.* **2017**, *56* (35), 10294−10323.
(158) Hermann, T. Drugs targeting the ribosome. *Curr. Opin. Struct. Biol.* **2005**, *15* (3), 355−366.
(159) Wilson, D. N. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nature Reviews Microbiology* **2014**, *12* (1), 35−48.
(160) Zhang, L.; He, J.; Bai, L.; Ruan, S.; Yang, T.; Luo, Y. Ribosome-targeting antibacterial agents: Advances, challenges, and opportunities. *Med. Res. Rev.* **2021**, *41* (4), 1855−1889.
(161) Crich, D. En Route to the Transformation of Glycoscience: A Chemist's Perspective on Internal and External Crossroads in Glycochemistry. *J. Am. Chem. Soc.* **2021**, *143*, 17−34.
(162) Baum, L. G.; Crocker, P. R. Glycoimmunology: ignore at your peril! *Immunol. Rev.* **2009**, *230* (1), 5−8.
(163) Landhuis, E. Glycobiology: Sweet success. *Nature: Nature* **2017**, *547*, 127−129.
(164) Hevey, R. Strategies for the Development of Glycomimetic Drug Candidates. *Pharmaceuticals* **2019**, *12* (2), 55.
(165) Dedola, S.; Rugen, M. D.; Young, R. J.; Field, R. A. Revisiting the Language of Glycoscience: Readers, Writers and Erasers in Carbohydrate Biochemistry. *ChemBioChem.* **2020**, *21* (3), 423−427.
(166) Bolcato, G.; Heid, E.; Bostrom, J. On the Value of Using 3D Shape and Electrostatic Similarities in Deep Generative Methods. *J. Chem. Inf. Model.* **2022**, *62* (6), 1388−1398.
(167) LaMarche, M. J. Think Biologically. *Act Chemically. J. Med. Chem.* **2022**, *65* (17), 11478−11484.
(168) Wilkinson, J. L.; Boxall, A. B. A.; Kolpin, D. W.; Leung, K. M. Y.; Lai, R. W. S.; Galbán-Malagón, C.; Adell, A. D.; Mondon, J.; Metian, M.; Marchant, R. A.; Bouzas-Monroy, A.; Cuni-Sanchez, A.; Coors, A.; Carriquiriborde, P.; Rojo, M.; Gordon, C.; Cara, M.; Moermond, M.; Luarte, T.; Petrosyan, V.; Perikhanyan, Y.; Mahon, C. S.; McGurk, C. J.; Hofmann, T.; Kormoker, T.; Iniguez, V.; Guzman-Otazo, J.; Tavares, J. L.; Gildasio De Figueiredo, F.; Razzolini, M. T. P.; Dougnon, V.; Gbaguidi, G.; Traoré, O.; Blais, J. M.; Kimpe, L. E.; Wong, M.; Wong, D.; Ntchantcho, R.; Pizarro, J.; Ying, G.-G.; Chen, C.-E.; Páez, M.; Martínez-Lara, J.; Otamonga, J.-P.; Poté, J.; Ifo, S. A.; Wilson, P.; Echeverría-Sáenz, S.; Udikovic-Kolic, N.; Milakovic, M.; Fatta-Kassinos, D.; Ioannou-Ttofa, L.; Belušová, V.; Vymazal, J.; Cárdenas-Bustamante, M.; Kassa, B. A.; Garric, J.; Chaumot, A.; Gibba, P.; Kunchulia, I.; Seidensticker, S.; Lyberatos, G.; Halldórsson, H. P.; Melling, M.; Shashidhar, T.; Lamba, M.; Nastiti, A.; Supriatin, A.; Pourang, N.; Abedini, A.; Abdullah, O.; Gharbia, S. S.; Pilla, F.; Chefetz, B.; Topaz, T.; Yao, K. M.; Aubakirova, B.; Beisenova, R.; Olaka, L.; Mulu, J. K.; Chatanga, P.; Ntuli, V.; Blama, N. T.; Sherif, S.; Aris, A. Z.; Looi, L. J.; Niang, M.; Traore, S. T.; Oldenkamp, R.; Ogunbanwo, O.; Ashfaq, M.; Iqbal, M.; Abdeen, Z.; O'Dea, A.; Morales-Saldaña, J. M.; Custodio, M.; de la Cruz, H.; Navarrete, I.; Carvalho, F.; Gogra, A. B.; Koroma, B. M.; Cerkvenik-Flajs, V.; Gombač, M.; Thwala, M.; Choi, K.; Kang, H.; Ladu, J. L. C.; Rico, A.; Amerasinghe, P.; Sobek, A.; Horlitz, G.; Zenker, A. K.; King, A. C.; Jiang, J.-J.; Kariuki, R.; Tumbo, M.; Tezel, U.; Onay, T. T.; Lejju, J. B.; Vystavna, Y.; Vergeles, Y.; Heinzen, H.; Pérez-Parada, A.; Sims, D. B.; Figy, M.; Good, D.; Teta, C. Pharmaceutical pollution of the world's rivers. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119* (8), e2113947119.
(169) Udiković-Kolić, N.; Scott, C.; Martin-Laurent, F. Evolution of atrazine-degrading capabilities in the environment. *Appl. Microbiol. Biotechnol.* **2012**, *96* (5), 1175−1189.
(170) Bender, A.; Cortés-Ciriano, I. Artificial intelligence in drug discovery: what is realistic, what are illusions? Part 1: Ways to make an impact, and why we are not there yet. *Drug Discovery Today* **2021**, *26* (2), 511−524.
(171) Zetzsche, L. E.; Chakrabarty, S.; Narayan, A. R. H. The Transformative Power of Biocatalysis in Convergent Synthesis. *J. Am. Chem. Soc.* **2022**, *144* (12), 5214−5225.

https://doi.org/10.1021/jacsau.2c00415
JACS Au 2022, 2, 2400−2416