
Article:
Karageorgis, G orcid.org/0000-0001-5088-6047, Foley, DJ, Laraia, L et al. (1 more author) (2020) Principle and design of pseudo-natural products. Nature Chemistry, 12 (3). pp. 227-235. ISSN 1755-4330

https://doi.org/10.1038/s41557-019-0411-x

Principle and Design of Pseudo-Natural Products

George Karageorgis*[,a], Daniel J. Foley*[,b], Luca Laraia*[,c], and Herbert Waldmann*

Abstract:

Natural products (NPs) are a significant source of inspiration towards the discovery of new bioactive compounds based on novel molecular scaffolds. However, there are currently only a small number of guiding synthetic strategies available to generate novel NP-inspired scaffolds, limiting both the number and types of compounds accessible. In this Perspective, we discuss a design approach for the preparation of biologically relevant small molecule libraries, harnessing the unprecedented combination of NP-derived fragments as an overarching strategy for the synthesis of new bioactive compounds. These novel 'pseudo-natural product' classes retain the biological relevance of NPs, yet exhibit structures and bioactivities not accessible to nature or through the use of existing design strategies. We also analyse selected pseudo-NP libraries using chemoinformatic tools, to assess their molecular shape diversity and properties. To facilitate the exploration of biologically relevant chemical space, we have identified design principles and connectivity patterns that would provide access to unprecedented pseudo-NP classes, offering new opportunities for bioactive small molecule discovery.

Main text

The search for small molecules with novel molecular scaffolds that display unprecedented or unexpected biological activity greatly benefits from insights gained from compound classes that are biologically relevant by definition, such as natural products (NPs). NPs have served as an inspiration and resource in drug discovery and chemical biology[1] and constitute chemical matter pre-validated by evolution.[2]

Libraries based on natural product structure or more relaxed chemical and structural considerations have been developed through different approaches. At one end of the spectrum NP-derived compound collections may be synthesised directly from readily available NPs, e.g. by derivatisations at pre-existing reactive sites, and by ring-distortion and/or -modification approaches ("complexity to diversity"; CtD) as first demonstrated by Hergenrother.[3,4] For instance, abietic acid was converted into a chemically diverse collection of complex compounds by means of 3-6 transformations, including ring opening-, expansion- and contraction sequences[4] (Figure 1a), whilst ring-distortion of the alkaloid quinidine yielded novel inhibitors of autophagy.[5] In an alternative approach complex polycyclic scaffolds that exhibit NP characteristics and properties are directly synthesised and distorted. For instance, tropane-containing compounds delivered multiple new scaffolds not found within NPs, several of which were found to be novel bromodomain binders.[6]

At the other end of the spectrum, Diversity Orinted Synthesis (DOS) [7] following the build/couple/pair strategy, employs the robust asymmetric intermolecular coupling of building blocks to induce stereochemical diversity, followed by intramolecular pairing of complementary functional groups, resulting in skeletal diversity (Figure 1b). DOS is not directly based on NP structures. However, the chemically diverse DOS compound classes can be considered NP-like due to their higher stereochemical content and higher fraction of sp^3-hybridised centers, when compared to commercial libraries. Incorporation of such NP-like characteristics positively contributes to biological performance of DOS libraries.

Biology-oriented synthesis (BIOS) draws inspiration from NPs to inform the synthesis of biologically relevant compound collections based on simplified NP scaffolds.[2] In BIOS, NP scaffolds are reduced in chemical complexity in a stepwise process, arriving at smaller intermediary scaffolds which then are employed as starting points for the synthesis of NP-inspired compound collections (Figure 1c). BIOS may be constrained by both chemical and biological limitations. Known and newly identified NPs occupy a large, definable, yet limited portion of chemical space.[8] However, NP-like chemical space, in general, is much bigger such that the focus on NP-scaffolds limits BIOS. The capacity to produce previously unknown NPs using biosynthetic technologies has been demonstrated, for instance, by the application of horizontal gene transfer techniques to assemble new artificial biosynthetic pathways,[9,10] and by the activation of transcriptionally silent biosynthetic pathways.[11,12] From a biological viewpoint, scaffolds prepared under the guidance of BIOS may retain the same kind of bioactivity as the guiding NPs, limiting the exploration of biological space.[9]

A design principle for bioactive compound discovery

Design of novel compound classes that broadly retain the biological relevance of NPs will benefit from strategies that enable the rapid exploration of biologically relevant chemical space, in particular fragment-based compound design.[13] The properties of NPs are retained in NP-derived fragments,[14] NPs may already be fragment-sized,[15] or can be converted into fragment-sized ring-systems.[16] Therefore, unprecedented combinations and fusions of NP fragments may provide access to novel scaffolds that may retain the chemical and biological characteristics of NPs, yet extend beyond the biologically relevant chemical space explored by nature (Figure 1d). These compounds would constitute ‘pseudo-natural products’, because these novel NP-inspired scaffolds would not be accessible through biosynthesis. Compound collections based on such molecular scaffolds, and, therefore, inspired by natural products may have different properties compared to both naturally occurring small molecules and BIOS collections, and may display unprecedented biological activities.

The term pseudo-natural product has been previously used by Suga et al. to describe cyclic peptides,[17,18] and by Oshima et al. for products of intercepted biosynthetic pathways.[19,20] We use the term for small molecule compound libraries based on scaffolds derived through the unprecedented combinations of NP fragments. Here we describe key principles to guide the design of such compound libraries and provide a first chemoinformatic analysis showing that pseudo-NPs may provide access to previously unexplored areas of biologically relevant chemical space, and show that pseudo-NPs can be endowed with novel bioactivity.

Figure 2 exemplifies the process for the design and actual synthesis of representative pseudo-natural product classes.[21–26] In an initial design step, fragments characteristic for different natural product classes are identified and combined in unprecedented arrangements (Figure 2a; for a discussion of the criteria guiding design and fragment combination types, see below), and then syntheses are developed and executed that give access to the desired pseudo-NP classes (Figure 2b-f). Finally, bioactivity of the novel pseudo-NPs is determined in different biochemical and biological assays.

Chromopynones **1** provide an illustrative example for pseudo-NP design and synthesis.[21] They embody a bridged-bicyclic scaffold which combines the fragments for widely occurring, bioactive chromane and tetrahydropyrimidinone NP-fragments. They were efficiently synthesised in a one-pot procedure involving a Biginelli reaction[27] as key step (Figure 2b). In a further example, the biologically prevalidated tropane and indole fragments were combined by means of an enantioselective intermolecular 1,3-dipolar cycloaddition to furnish a new, structurally-rich scaffold, **2** (Figure 2c).[22] Indoles and tropanes are characteristic of hundreds of NPs with widely differing bioactivities, but do not occur in combination in nature. Similarly, intermolecular 1,3-dipolar cycloadditions found use in the preparation of pyrrotropane pseudo-NPs **3** (Figure 2d), which fused the tropane and pyrrolidine fragment,[23] and a related intramolecular 1,3-dipolar cycloaddition proved powerful for the preparation of pseudo-NPs **4**, combining the pyrrolidine and quinolone fragments (Figure 2e).[24] The oxindole fragment and O-heterocycles which are characteristic substructures of iridoid terpenes were combined to spiro-configured indiridoids by means of a Au(I) catalyzed reaction cascade including a 6-endo-dig ene-yne cyclization followed by ring-opening and rearrangement. A one-pot process including the use of an enantioselective Pictet–Spengler reaction was employed to fuse indoles with piperidone components to provide the bridged ‘indopipenone’ scaffold **5** (Figure 2f).

Chemoinformatic analyses of pseudo-natural products

To characterise the chemical space occupied by these previously reported pseudo-NPs,[21–26] we calculated the natural product-likeness scores (NP Score)[28] for the pyrrotropanes,[23] the pyrroquinolinones,[24] and the indopipenones.[26] We have reported the NP Score for chromopynones and indotropanes previously.[21,22] We compared the collective NP Score of these compound libraries to the scores of the NPs found in the ChEMBL database,[29] the set of approved drugs in DrugBank[30] representing marketed and experimental drugs and selected BIOS libraries (see Supplementary Figure 7 and Supplementary Tables 13 and 14). The NP likeness score measures the frequency at which fragments occur in natural products as compared to commercial non-NPs. It ranges from -5 to +5, and higher values indicate higher natural product-likeness. To calculate the score the corresponding molecule is divided into atom-centred fragments. These contribute to the score by their relative frequency of occurrence in NPs divided by their frequency of occurrence in commercial non-NPs. BIOS compounds display a broad NP-score distribution ranging from 0.0 to +4.0 (Figure 3a, red curve) and overlay significantly with the set of natural products in the ChEMBL database, as expected. The pseudo-NPs display more narrow distributions with a range from −1.0 to +1.0 (Figure 3a, black curve) and cover a portion of chemical space only sparsely occupied by NPs. These scores may be rationalised by the fact that these new NP fragment combinations cannot be found in known NP scaffolds.

Calculation of relevant molecular properties of pseudo-NPs revealed that the majority of the pseudo-NP collections fall within "Lipinski rule of 5" (Ro5) space[31] (60-100% for individual libraries, 86% for the collated collections; Figure 3b). Thus, unprecedented combinations of NP fragments may render pseudo-NP scaffolds to be fairly "drug-like" and pseudo-NP collections may, therefore, display advantageous physical properties for molecular discovery. This observation extends to other metrics of Ro5 space such as total polar surface area (62 – 109 $Å^2$), and H-bond donors and acceptors (0-3; and 0-10, respectively, see Supplementary Table 12 for further details). By comparison, only 48% of BIOS compounds fall within Lipinski-like space (Figure 3b), yet they have a higher Fsp^3 compared to the currently explored pseudo-NPs.

Assessment of the molecular shape diversity of pseudo-NPs by determination of the principal moment of inertia (PMI) plots revealed a systematic shift away from the rod-disk axis, towards more three-dimensional molecular shapes for the pseudo-NPs. Most synthetic molecules congest along the rod-disk axis,[32] as exhibited by the collection of approved drugs in the DrugBank database (Figure 3c).

Structural Analyses of Pseudo-Natural Products

NP-derived fragments may be combined using different connectivity patterns to generate new pseudo-NPs (Figure 4). A monopodal or linear connection could be used to connect two fragments, such as the phenylpyrrolidinones, **6** (Figure 4a) which have previously been explored as androgen receptor antagonoists.[33] By analogy, polypodal connections will be possible, and perhaps limited by synthetic accessibility. Patterns leading to the generation of stereogenic centres, and more structurally complex scaffolds should be prioritised to capture the high stereochemical content of NPs in general. For example, in indotropanes **2**, pyrrotropanes **3**, and pyrroquinolines **4**, two adjoining fragments are fused edge-on-edge, sharing two common atoms. This connectivity pattern is also observed in natural products e.g. the Murraya alkaloids **8** (Figure 4b).[34] Connection of two fragments through only one atom in a point-on-point manner will result in the formation of a spirocyclic centre such as reported for the diazaspirodecane core of selective dopamine D3 receptor antagonists **9**,[35] (Figure 4c) and the NP (−)-horsfiline, **10**.[36] Two fragments may share three or more common atoms, forming a bridged bicyclic scaffold (Figure 4d), as for instance in the chromopynones **1**, and in the NP sespenine, **11**.[37] Also, two adjoining fragments may not share any common atoms, but may even have one or two linking atoms as intervening connection points (Figure 4e). The adjoining fragments may be connected through non-sequential atoms resulting in the formation of a bridged bipodal[38] connection pattern. The indopipenone pseudo-NP class **5**, is such an example, and similarly this pattern can be observed in the NP aspergillin PZ, **12**[39] (For a more extended list of connection types and their examples in NPs and pseudo-NPs see Supplementary Scheme 1).

Pseudo-NP design principles

Alternative connective combinations of a given set of NP-fragments can yield markedly different pseudo-NP scaffolds, such that appropriate design of pseudo-NP structure may enable wider exploration of chemical space. For the design of such pseudo-NP classes, several basic criteria will be generally applicable:

(i) Since chirality is a defining property of many NPs and stereogenic content correlates with bioactivity,[40,41] NP fragments may be combined in complexity-generating reactions leading to novel three-dimensional scaffolds amenable to further derivatisation.

(ii) NP fragments with complementary heteroatom content should be combined, in particular nitrogen (more frequently found in drugs than in NPs) and oxygen (more frequently occurring in NPs).[42]

(iii) Combination of NP fragments derived from NPs with diverse bioactivities may maximise the biological relevance of the resulting pseudo-NP scaffold.

(iv) Combination of biosynthetically unrelated fragments may be beneficial for novel bioactivity, since biosynthetically different NPs will have been synthesised by different enzymes, encoding different structural features that may facilitate binding to different proteins.

Using these general criteria, we investigated regioisomers of the pyrrotropane **22**-**27**, and pyrroquinoline **28**-**35** as five related but different illustrative examples (Figure 5a) and calculated their NP scores. We found that compound classes **25**-**27**, and **34** and **35**, had been synthesised before,[6,43,44] while the rest of the investigated compound classes have not been reported. From this analysis three complementary, yet alternative principles emerged, which draw inspiration from scaffold hopping approaches in medicinal chemistry (Figure 5b-d):[45]

Combination of a fragment set in different connectivities. Scaffolds would be composed of frameworks that are different at the graph[46] level, implementing the connectivity patterns outlined above e.g. edge fusion (for pyrrotropanes: **22**; for pyrroquinolines: **28**), spiro fusion (for pyrrotropanes: **23**; for pyrroquinolines: **30**), or bridged fusion (for pyrroquinolines: **33**) (Figure 5a). Such frameworks exhibit remarkable differences between scores and offer an opportunity for the systematic exploration of new chemical space. For instance, there is almost one unit of difference in the NP-score between scaffolds **22** and **23**, and almost two units between scaffolds **28** and **36** (Figure 5a).

Combination of fragments with the same connectivity pattern, but different regioisomeric connectivity points. The connectivity pattern between fragments would be maintained (i.e. the framework is the same at the graph[46] level), but atoms within the resulting scaffold are subsequently alternated (at the graph-node[46] level). For example, **31** and **32** are both spiro-fused pyrroquinolines connected through the 2- or 3-position of the tetrahydroquinoline fragment respectively and the 3-position of the pyrrolidine fragment, but the pyrrolidine is rotated around the piperidine ring. Scaffolds **30** and **31** are both spiro-fused pyrroquinolines connected through the 2-position of the tetrahydroquinoline fragment, but through the 2- or 3-position of the pyrrolidine fragment (Fig. 5c), i.e. the nitrogen atom is rotated around the pyrrolidine ring. Scaffolds designed accordingly (e.g. for pyrrotropanes: **23**, **24**; for pyrroquinolines: **28**-**29**) tend to cluster together in the NP-score analysis (e.g. there are

only 0.09 units of difference between scaffolds **28** and **29**, Fig. 5a). However, despite exhibiting similar NP-scores, these regioisomeric scaffolds might exhibit different biological activities.

More than two fragments can be combined in several distinct connectivity patterns. Although combinations of several fragments will lead to higher molecular weight compounds, such pseudo-NPs may still be very valuable. An example for a combination of an oxindole, a pyrrolidine, and a dihydroquinolinone fragment using spiro and edge fusions is shown in Figure 5d (**37**).

For validation of these design criteria, a full set of isomeric scaffolds would have to be synthesised and broadly evaluated for bioactivity. For the isomeric compound classes shown in Figure 4 a literature search revealed that the corresponding data are not availalble. However, pyrrotropanes **26** have been identified as bromodomain binders,[6] whilst pyrroquinolines **34** are active against the retinoid orphan receptor-γ (ROR-γ),[44] showing that different fragment combinations may lead to different bioactivity. Further analysis of the literature indicates that pseudo-NP structures have indeed been synthesised, but not yet in a consistent and complementary manner. Examples include indole-bridged chroman oxindoles by Bu et al.[47] and 3,4-dihydropyrrolopyrazinones by Zhang et al.[48] We note that the synthesis of such complementary pseudo-NP libraries will require and be facilitated by the development of novel, innovative synthesis methodologies. The pseudo-NPs synthesised in our laboratories will be stored and curated by the Compound Management and Screening Center of the Max Planck Gesellschaft (COMAS, see http://comas.mpi-dortmund.mpg.de/index.php) through which they are accessible in collaborative scientific projects.

Biological evaluation of pseudo-natural products

Pseudo-NPs may display unexpected bioactivities differing from the activities of the NPs from which their fragments are derived which calls for investigation of these compounds in assays monitoring wide biological space. For instance, the pseudo-NPs described in Figure 2 were investigated[21,22,26] in cell-based assays that monitor progression through different signaling cascades (e.g. Wnt, Hedgehog) and phenotypic changes (e.g. lipid droplet formation, reactive oxygen species induction, autophagy). Such screens monitor changes in whole systems rather than individual biological targets and, due to their unbiased nature, can lead to the identification of novel protein targets.[49]

A phenotypic assay monitoring glucose uptake identified chromopynones **1** as novel glucose uptake inhibitors. Glucose metabolism is deregulated in cancerous tissue.[50] Chromopynones exhibited a desirable dual GLUT-1,-3 selectivity profile.[21] Notably, compounds containing either of the underlying NP-fragments (chromane or tetrahydropyrimidinone), did not exhibit glucose uptake inhibitory activity. Thus, the chromopynone scaffold retains the biological relevance of the individual fragments, yet displays a different bioactivity profile. NPs containing both a chromane and a tetrahydropyrimidinone fragment have not been reported, suggesting that these fragments may be biosynthetically unrelated. The indotropane Myokinasib was identified as ATP-competitive/mixed type myosin light chain kinase (MLCK1) inhibitors with an unprecedented inhibitor chemotype.[22] MLCK1 is involved in inflammatory disease, myocardial infarction, as well as tumorigenesis.[51] Similar to chromopynones, the indotropanes emerge from a combination of biosynthetically unrelated NP-fragments, leading to a novel pseudo-NP class with unprecedented biological activity. Indopipenones **5** were revealed to be inhibitors of MptpB tyrosine phosphatase[26] and defined a novel phosphatase inhibitor chemotype. This enzyme plays a key role in the infection and proliferation of mycobacteria and it is a potential biological target for the treatment of tuberculosis.[52]

Beyond assays monitoring establishment of individual phenotypes, recently developed multiparametric high-content image screening technologies,[53] which quantify the changes of a wide number of cellular morphology characteristics, will be particularly suited to explore the biological activity of novel pseudo-NP classes. They combine the unbiased nature of phenotypic setups with the high information volume provided by high-content microscopy, and can be used to generate class- or compound specific phenotypic 'fingerprints'. The fingerprint information may subsequently be used to extrapolate information about the biological target or mode-of-action of a small molecule, and may facilitate the realisation of pseudo-NPs as a more general design strategy for the discovery of bioactive compounds.

Outlook

Natural products continue to guide the design of new biologically relevant compound libraries, and NP structure has provided inspiration for different approaches aimed at the development of NP-like compound collections. DOS is not based on NP structures but mainly built on chemical and structural considerations. It employs robust transformations to build up stereochemically diverse compound libraries with high scaffold diversity and number. These are considered NP-like but their biological relevance is initially unclear. BIOS, pseudo-NPs and the CtD strategy, on the other hand, are built on genuine NP structures, and, therefore, are endowed with biological relevance

derived from their creation in evolution. Due to their biosynthetic origin and biological functions they encode structural properties that enable binding to proteins. In this sense, chemically diverse DOS libraries and compound collections directly delineated from existing NPs primarily differ in their origin and can be considered complementary. In this continuum pseudo-NPs occupy an intermediary position. On the one hand they are endowed with biological relevance (as are BIOS collections), on the other hand their fragment-based synthesis enables coverage of a larger area of chemical space (as do DOS libraries).

Biological analysis of BIOS compound collections demonstrated that structurally simplified NP-inspired compounds can be endowed with diverse and interesting biological activities. Yet, BIOS is intrinsically limited both chemically and biologically due to its focus on specific NP scaffold structures. Combining BIOS with principles derived from fragment-based ligand discovery may enable the rapid exploration of biologically relevant chemical space beyond the areas accessible by known NP scaffolds, and give rise to new compound classes which are distinct from BIOS libraries. These synthetic, small molecule pseudo-natural products contain unprecedented combinations of NP-fragments.

Here, we have analysed biologically active pseudo-NP libraries based upon scaffolds comprising novel combinations of NP-fragments. Chemoinformatic analyses employing the NP-score tool demonstrate that pseudo-NP libraries occupy areas of chemical space not covered by NPs and BIOS libraries. The explored pseudo-NP libraries appear to have favourable physiochemical properties for molecular discovery.

Using five related libraries[21–26] as representative examples we have developed a set of design principles that can inform future syntheses of pseudo-NPs. A ‘thought experiment’ suggested potential pseudo-NPs that would arise though unprecedented combinations of some simple NP-fragments (e.g. the pyrrolidine and quinoline ring-systems). In general, multiple combinations of NP-fragments are possible, and pseudo-NPs have likely already been unknowingly prepared, but have yet to be studied in a systematic manner. Their synthesis is limited only by creativity and the tools available for their preparation. Therefore, the introduction of enabling synthetic methodologies that grant access to de novo combinations of NP-fragments will greatly facilitate pseudo-NP preparation and study.

The examples of pseudo-NPs discussed here and their bioactivity profiles provide evidence that pseudo-NPs may display novel and unexpected bioactivities, different from their guiding NPs. However, for full analysis and exploration of pseudo-NP bioactivity, wider profiling is required in unbiased phenotypic screening platforms that

facilitate the identification of novel biological targets and modes of action. Use of such phenotypic and multiparametric screening platforms, e.g. the "cell-painting assay"[53], in combination with powerful target identification methods, such as proteome-wide thermal shift assays,[54] or target degradation approaches,[55] may facilitate target deconvolution for novel biological phenomena induced by compound treatment, and may realise the full value of pseudo-NPs towards bioactive compound discovery. We expect that new pseudo-NP classes will have significant value towards the exploration of biologically relevant chemical space, and ultimately towards molecular discovery programmes aimed at the modulation of disease.

References:

1. Li, J. W.-H. & Vederas, J. C. Drug Discovery and Natural Products: End of an Era or an Endless Frontier? Science **325**, 161–165 (2009).

2. Wetzel, S., Bon, R. S., Kumar, K. & Waldmann, H. Biology-oriented synthesis. Angew. Chem. Int. Ed. **50**, 10800–10826 (2011).

3. Huigens, R. W. et al. A Ring Distortion Strategy to Construct Stereochemically Complex and Structurally Diverse Compounds from Natural Products. Nat. Chem. **5**, 195–202 (2013).

4. Rafferty, R. J., Hicklin, R. W., Maloof, K. A. & Hergenrother, P. J. Synthesis of Complex and Diverse Compounds through Ring Distortion of Abietic Acid. Angew. Chem. Int. Ed. **53**, 220–224 (2014).

5. Laraia, L. et al. Discovery of Novel Cinchona-Alkaloid-Inspired Oxazatwistane Autophagy Inhibitors. Angew. Chem. Int. Ed. **56**, 2145–2150 (2017).

6. Foley, D. J. et al. Synthesis and Demonstration of the Biological Relevance of sp3-rich Scaffolds Distantly Related to Natural Product Frameworks. Chem. – A Eur. J. **23**, 15227–15232 (2017).

7. Nielsen, T. E. & Schreiber, S. L. Towards the Optimal Screening Collection: A Synthesis Strategy. Angew. Chem. Int. Ed. **47**, 48–56 (2008).

8. 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. **114**, 5601–5606 (2017).

9. Crane, E. A. & Gademann, K. Capturing Biological Activity in Natural Product Fragments by Chemical Synthesis. Angew. Chem. Int. Ed. **55**, 3882–3902 (2016).

10. Klein, J. et al. Yeast Synthetic Biology Platform Generates Novel Chemical Structures as Scaffolds for Drug Discovery. ACS Synth. Biol. **3**, 314–323 (2014).

11. Scherlach, K. & Hertweck, C. Triggering cryptic natural product biosynthesis in microorganisms. Org. Biomol. Chem. **7**, 1753–1760 (2009).

12. Khaldi, N. et al. SMURF: Genomic mapping of fungal secondary metabolite clusters. Fungal Genet. Biol. **47**, 736–741 (2010).

13. 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 Discov. **15**, 605–619 (2016).

14. Over, B. et al. Natural-product-derived fragments for fragment-based ligand discovery. Nat Chem **5**, 21–28 (2013).

15. Vu, H. et al. Fragment-Based Screening of a Natural Product Library against 62 Potential Malaria Drug Targets Employing Native Mass Spectrometry. ACS Infect. Dis. **4**, 431–444 (2018).

16. Prescher, H. et al. Construction of a 3D-shaped, natural product like fragment library by fragmentation and diversification of natural products. Bioorg. Med. Chem. **25**, 921–925 (2017).

17. Goto, Y., Ito, Y., Kato, Y., Tsunoda, S. & Suga, H. One-Pot Synthesis of Azoline-Containing Peptides in a Cell-free Translation System Integrated with a Posttranslational Cyclodehydratase. Chem. Biol. **21**, 766–774 (2014).

18. Ozaki, T. et al. Dissection of goadsporin biosynthesis by in vitro reconstitution leading to designer analogues expressed in vivo. Nat. Commun. **8**, 14207 (2017).

19. Asai, T. et al. Use of a biosynthetic intermediate to explore the chemical diversity of pseudo-natural fungal polyketides. Nat Chem **7**, 737–743 (2015).

20. Kikuchi, H. et al. Monoterpene Indole Alkaloid-Like Compounds Based on Diversity-Enhanced Extracts of Iridoid-Containing Plants and Their Immune Checkpoint Inhibitory Activity. Org. Lett. **18**, 5948–5951 (2016).

21. Karageorgis, G. et al. Chromopynones are pseudo natural product glucose uptake inhibitors targeting glucose transporters GLUT-1 and -3. Nat. Chem. **10**, 1103–1111 (2018).

22. Schneidewind, T. et al. The Pseudo Natural Product Myokinasib Is a Myosin Light Chain Kinase 1 Inhibitor with Unprecedented Chemotype. Cell Chem. Biol. **26**, 512–523 (2019).

23. Xu, H., Golz, C., Strohmann, C., Antonchick, A. P. & Waldmann, H. Enantiodivergent Combination of Natural Product Scaffolds Enabled by Catalytic Enantioselective Cycloaddition. Angew. Chem. Int. Ed. **55**, 7761–7765 (2016).

24. Vidadala, S. R., Golz, C., Strohmann, C., Daniliuc, C.-G. & Waldmann, H. Highly Enantioselective Intramolecular 1,3-Dipolar Cycloaddition: A Route to Piperidino-Pyrrolizidines. Angew. Chem. Int. Ed. **54**, 651–655 (2015).

25. Lee, Y.-C. et al. A ligand-directed divergent catalytic approach to establish structural and functional scaffold diversity. Nat. Commun. **8**, 14043 (2017).

26. Nören-Müller, A. et al. Discovery of protein phosphatase inhibitor classes by biology-oriented synthesis. Proc. Natl. Acad. Sci. U. S. A. **103**, 10606–10611 (2006).

27. Oliver Kappe, C. 100 years of the biginelli dihydropyrimidine synthesis. Tetrahedron **49**, 6937–6963 (1993).

28. Ertl, P., Roggo, S. & Schuffenhauer, A. Natural Product-likeness Score and Its Application for Prioritization of Compound Libraries. J. Chem. Inf. Model. **48**, 68–74 (2008).

29. Vanii Jayaseelan, K., Moreno, P., Truszkowski, A., Ertl, P. & Steinbeck, C. Natural product-likeness score revisited: an open-source, open-data implementation. BMC Bioinformatics **13**, 106–112 (2012).

30. Law, V. et al. DrugBank 4.0: shedding new light on drug metabolism. Nucleic Acids Res. **42**, D1091–

D1097 (2014).

31. 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 Deliv. Rev. **23**, 3–25 (1997).

32. Sauer, W. H. B. & Schwarz, M. K. Molecular Shape Diversity of Combinatorial Libraries: A Prerequisite for Broad Bioactivity. J. Chem. Inf. Comput. Sci. **43**, 987–1003 (2003).

33. Lack, N. A. et al. Targeting the Binding Function 3 (BF3) Site of the Human Androgen Receptor through Virtual Screening. J. Med. Chem. **54**, 8563–8573 (2011).

34. Kureel, S. P., Kapil, R. S. & Popli, S. P. Terpenoid alkaloids from Murraya koenigii spreng. - II.: The constitution of cyclomahanimbine, bicyclomahanimbine, and mahanimbidine. Tetrahedron Lett. **10**, 3857–3862 (1969).

35. Reilly, S. W. et al. Highly Selective Dopamine D3 Receptor Antagonists with Arylated Diazaspiro Alkane Cores. J. Med. Chem. **60**, 9905–9910 (2017).

36. Palmisano, G., Annunziata, R., Papeo, G. & Sisti, M. Oxindole alkaloids. A novel non-biomimetic entry to (−)-Horsfiline. Tetrahedron: Asymmetry **7**, 1–4 (1996).

37. Ding, L., Maier, A., Fiebig, H.-H., Lin, W.-H. & Hertweck, C. A family of multicyclic indolosesquiterpenes from a bacterial endophyte. Org. Biomol. Chem. **9**, 4029–4031 (2011).

38. MacLellan, P. & Nelson, A. A conceptual framework for analysing and planning synthetic approaches to diverse lead-like scaffolds. Chem. Commun. **49**, 2383–2393 (2013).

39. Canham, S. M., Overman, L. E. & Tanis, P. S. Identification of an unexpected 2-oxonia[3,3]sigmatropic rearrangement/aldol pathway in the formation of oxacyclic rings. Total synthesis of (+)-aspergillin PZ. Tetrahedron **67**, 9837–9843 (2011).

40. Lee, M.-L. & Schneider, G. Scaffold Architecture and Pharmacophoric Properties of Natural Products and Trade Drugs: Application in the Design of Natural Product-Based Combinatorial Libraries. J. Comb. Chem. **3**, 284–289 (2001).

41. Lovering, F., Bikker, J. & Humblet, C. Escape from Flatland: Increasing Saturation as an Approach to Improving Clinical Success. J. Med. Chem. **52**, 6752–6756 (2009).

42. Grabowski, K., Baringhaus, K.-H. & Schneider, G. Scaffold diversity of natural products: inspiration for combinatorial library design. Nat. Prod. Rep. **25**, 892–904 (2008).

43. Muhuhi, J. & Spaller, M. R. Expanding the Synthetic Method and Structural Diversity Potential for the Intramolecular Aza Diels−Alder Cyclization. J. Org. Chem. **71**, 5515–5526 (2006).

44. Jingwu Duan, et al. Tricyclic sulfones as RORγ modulators and their preparations - WO 2016179460 A1, **US15148209**, (2016).

45. Sun, H., Tawa, G. & Wallqvist, A. Classification of scaffold-hopping approaches. Drug Discov. Today **17**, 310–324 (2012).

46. Lipkus, A. H. et al. Structural Diversity of Organic Chemistry. A Scaffold Analysis of the CAS Registry. J. Org. Chem. **73**, 4443–4451 (2008).

47. Guo, J., Bai, X., Wang, Q. & Bu, Z. Diastereoselective Construction of Indole-Bridged Chroman Spirooxindoles through a TfOH-Catalyzed Michael Addition-Inspired Cascade Reaction. J. Org. Chem. **83**, 3679–3687 (2018).

48. Sandoval, C., Lim, N.-K. & Zhang, H. Two-Step Synthesis of 3,4-Dihydropyrrolopyrazinones from Ketones and Piperazin-2-ones. Org. Lett. **20**, 1252–1255 (2018).

49. Mullard, A. The phenotypic screening pendulum swings. Nat. Rev. Drug Discov. **14**, 807–809 (2015).

50. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. Science **324**, 1029–1033 (2009).

51. Rigor, R. R., Shen, Q., Pivetti, C. D., Wu, M. H. & Yuan, S. Y. Myosin Light Chain Kinase Signaling in Endothelial Barrier Dysfunction. Med. Res. Rev. **33**, 911–933 (2013).

52. Butler, D. New fronts in an old war. Nature **406**, 670–672 (2000).

53. Caicedo, J. C. et al. Data-analysis strategies for image-based cell profiling. Nat. Methods **14**, 849–863 (2017).

54. Jafari, R. et al. The cellular thermal shift assay for evaluating drug target interactions in cells. Nat. Protoc. **9**, 2100–2122 (2014).

55. Chessum, N. E. A. et al. Demonstrating In-Cell Target Engagement Using a Pirin Protein Degradation Probe (CCT367766). J. Med. Chem. **61**, 918–933 (2018).

Affiliations

*Max-Planck Institute for Molecular Physiology, Dept. of Chemical Biology, Dortmund Germany

[a]Current address: School of Chemistry, University of Leeds, U.K.

[b]Current address: College of Science, University of Canterbury, New Zealand.

[c]Current address: Dept. of Chemistry, Technical University of Denmark, Copenhagen, Denmark.

Dr George Karageorgis ORCID: 0000-0001-5088-6047

Dr Daniel J. Foley ORCID: 0000-0001-6221-3333

Dr Luca Laraia ORCID: 0000-0001-6692-3412

Prof. Dr Herbert Waldmann ORCID: 0000-0002-9606-7247

Acknowledgments

This research was supported by the Max-Planck-Gesellschaft. G. K. is grateful to the Alexander von Humboldt Foundation for a fellowship. D. J. F. is grateful to the European Commission for a Marie Skłodowska-Curie Fellowship (Grant 794259). L. L. is grateful to the Novo Nordisk Foundation and DTU for funding.

Author Contributions

G. K. performed the cheminformatic analyses, and derived the connection types and design principles for novel pseudo-NP classes. G. K., D. J. F, L. L., and H. W. wrote the paper.

Corresponding Author

Correspondence to Prof. Dr. Herbert Waldmann.

Competing Interests

The authors declare no competing interests.

Data Availability

The material and data reported in this study are available upon request from Prof. Dr. Herbert Waldmann.

Table of Contents Graphic:

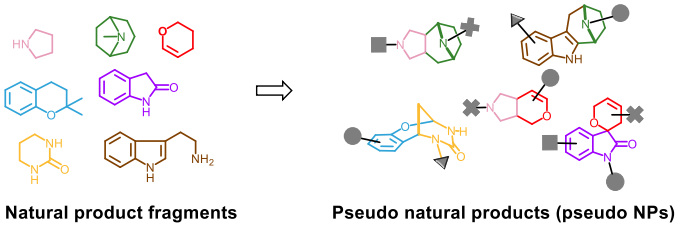


Table of Contents Summary:

In this Perspective we outline the principles for the design of novel and distinct compound classes emerging from de novo combinations of NP-derived fragments. These compounds are termed Pseudo-NPs, facilitate the exploration of biologically relevant chemical space and may have significant value for molecular discovery programs aimed at the modulation of disease.

Figure Captions:

Figure 1: Approaches to the design of biologically relevant compound collections. Panel a: Principle of the build-couple-pair approach in Diversity Oriented Synthesis (DOS). This strategy results in compounds with high stereogenic/sp^3 content displaying some NP characteristics; Panel b: Synthesis of natural product derived compounds by scaffold manipulation and decoration following the Complexity to Diversity (CtD) approach results in novel NP-derived compounds; Panel c: Scaffold synthesis and decoration following the principle of Biology Oriented Synthesis (BIOS) leads to NP-inspired compounds; Panel d: Design of pseudo-natural product collections by de novo combination of natural product fragments leads to NP-inspired compounds.

Figure 2: Development of pseudo-natural product collections. Panel a: Rationale for the design of pseudo-natural product compound classes. Natural product fragments can be combined in unprecedented combinations to afford novel structures, e.g. indole with tropane, indole with piperazine etc. The individual fragment structures combined in the pseudo-NP scaffolds are colour coded. Panel b: Synthesis of chromopynones[21] combining chromane and tetrahydropyrimidinone NP fragments. They are glucose uptake inhibitors with dual GLUT-1/-3 selectivity. Panel c: Synthesis of Indotropanes[22] combining the indole and tropane NP fragments. They are mixed type MLCK1 inhibitors. Panel d: Synthesis of Pyrrotropanes[23] combining the pyrrolidine and tropane NP fragments. Panel e: Synthesis of Pyrroquinolinones[24] combining the pyrrolidine and quinolinone NP fragments. Panel f: Synthesis of Indopipenones[26] combining the indole and piperazine NP fragments. They are tyrosine phosphatase inhibitors. The individual fragments structures combined in the pseudo-NP scaffolds are colour coded.

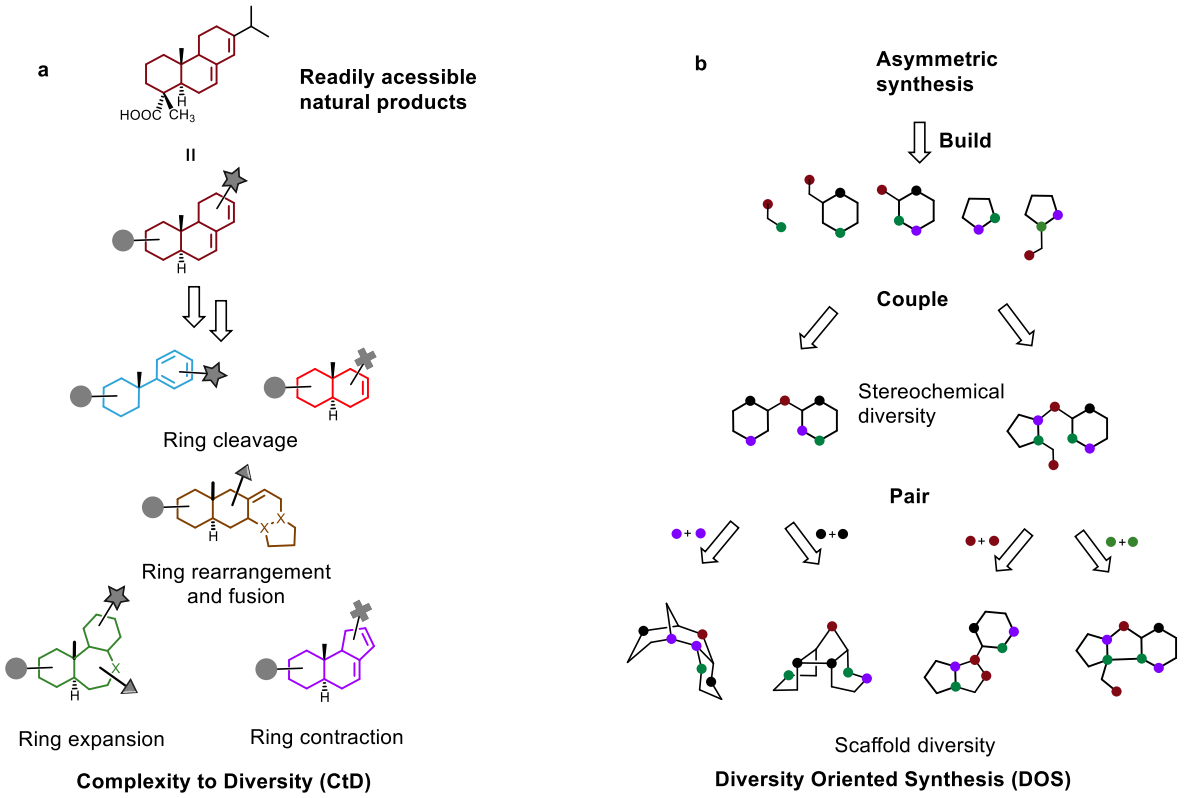
Figure 3: Chemoinformatic analyses of pseudo-NPs. a: The NP-score for pseudo-NPs (black curve), the set of NPs from ChEMBL (red curve), approved drugs in the DrugBank database (blue curve) and selected BIOS libraries (green curve). Pseudo-NPs occupy a different area of the NP Score graph compared to NPs and fall under the distribution of compounds in DrugBank. b: Molecular properties analysis (ALogP vs MW) of pseudo-NPs (black squares) and BIOS libraries (green dots). A higher portion of Pseudo-NPs fall within the limits of Lipinski space compared to BIOS libraries. c: PMI plot of pseudo-NPs (black squares) and DrugBank compounds (green dots). Pseudo-NPs display a wide distribution of PMI values and move further away from the rod-disk axis of the graph compared to DrugBank compounds. For NP Scores, molecular properties analyses and PMI plots for individual Pseudo-NP classes, BIOS libraries NPs in ChEMBL and compounds in DrugBank please see SI, Supplementary Figures 1-7 and Supplementary Tables 1-17.

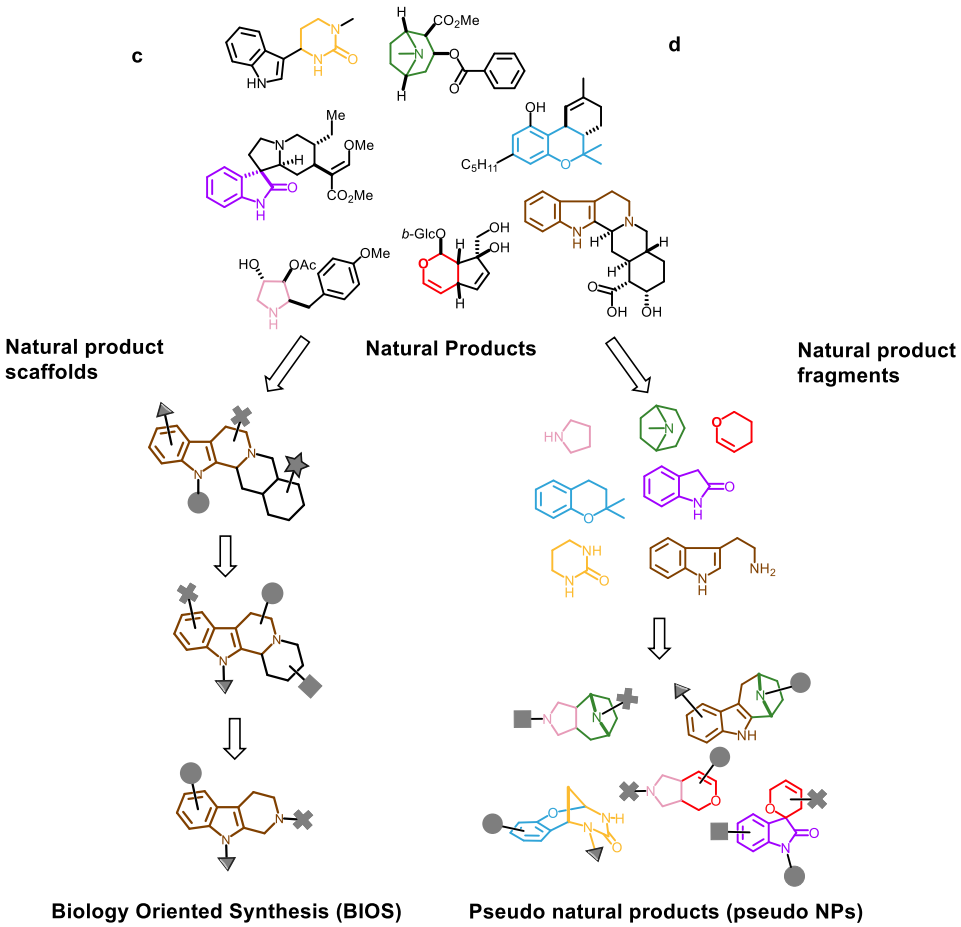
Figure 4: Illustration of possible NP fragment connectivities to guide synthesis and design of pseudo-NPs. These connectivity patterns are also found in natural products, and characteristic examples are shown. a: Monopodal Connection: observed when two fragments are connected through a distinct atom in each fragment. b: Edge Fusion: Two fragments can be connected through two common atoms. c: Spiro Fusion: Connecting two fragments through the same atom. d: Bridged Fusion: two fragments are connected through three or more fragments which form a bridged scaffold. e: Bridged Bipodal Connection: two connected fragments may not have any common atoms, and may be linked through intervening connection points resulting in a bridged scaffold. Black dots denote connectivity points. Individual fragments are indicated in either blue or green. Varied substituents or extended parts of NPs have been greyed for clarity.

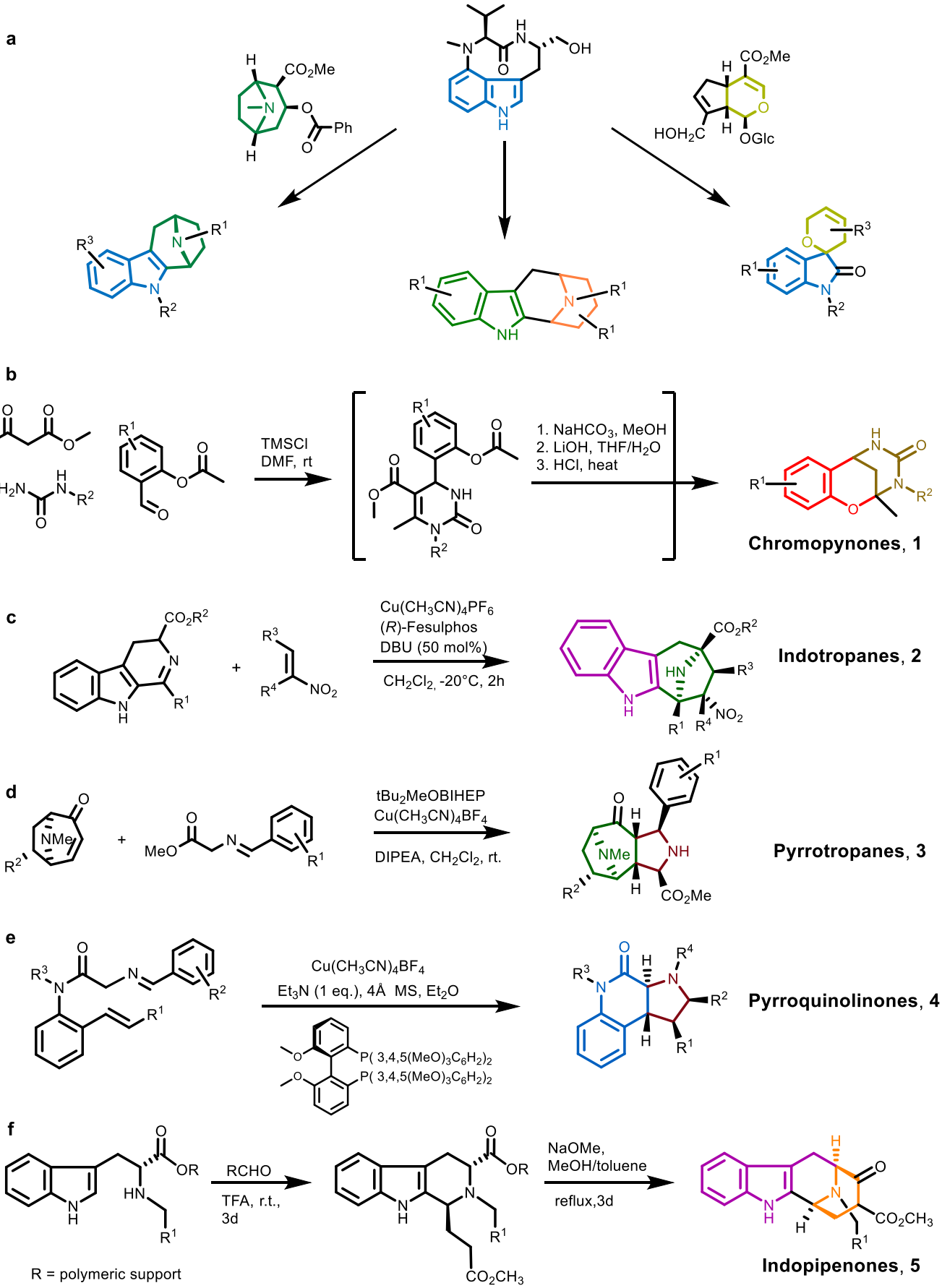
Figure 5: NP Scores of regioisomeric pyrrotropane and pyrroquinoline scaffolds. a: Designed (black) and reported[6,43,44] (blue) pseudo-natural product scaffolds stemming from the combination of pyrrolidine, tropane, and tetrahydroquinoline fragments. The variable NP-scores demonstrate that pseudo-NPs may facilitate the exploration of chemical space. Black dots are used to denote the connectivity points. b: Illustrative examples of Design Principle 1: Diverse Pseudo-NP scaffolds can be derived by combining two NP fragments using different connectivity patterns; **22**: Pyrrotropane scaffold by edge fusion; **23**: Pyrrotropane scaffold by spiro fusion c: Design Principle 2: Pseudo-NP scaffolds can be derived by combining two NP fragments by maintaining the same connectivity pattern e.g. spiro fusion, while changing the connectivity points between fragments; **24**: Pyrrotropane scaffold by spiro fusion through C2 atom of pyrrolidine; **30**: Pyrroquinoline scaffold by spiro fusion through C2 atom of tetrahydroquinoline and C3 atom of pyrrolidine; **31**: Pyrroquinoline scaffold by spiro fusion through C2 atom of

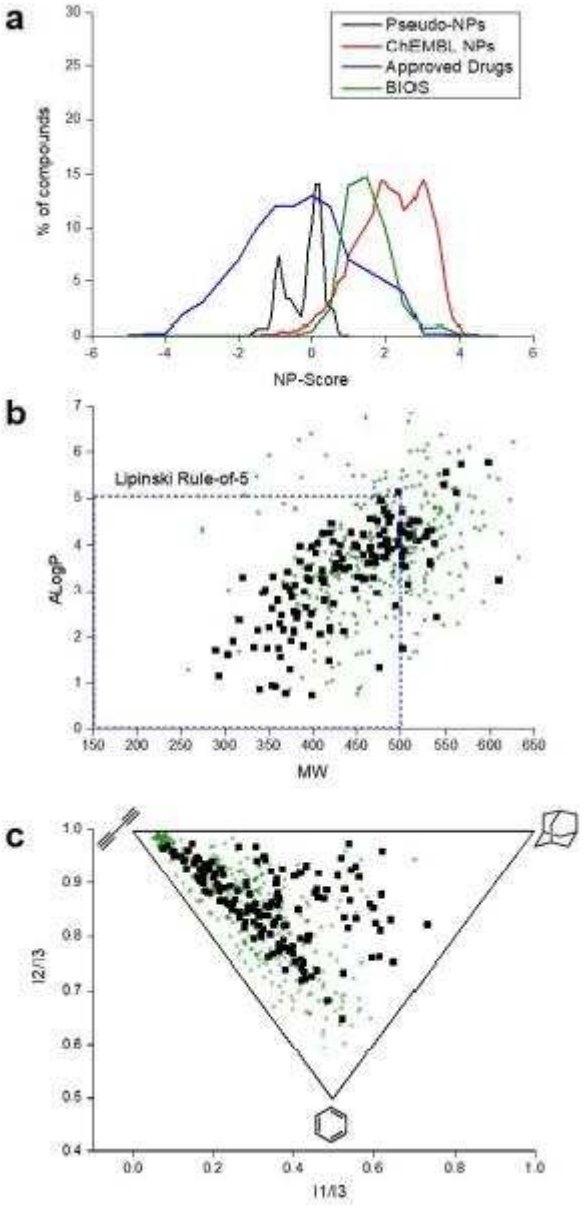
tetrahydroquinoline and C2 atom of pyrrolidine; d: Design Principle 3: Pseudo-NP scaffolds can be derived by combining more than two NP fragments using different connectivity patterns; **37**: Indopyrroquinolinone scaffold by spiro fused oxindole and pyrrolidine, and edge fused dihydroquinolone.

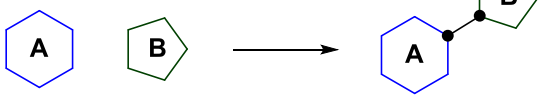
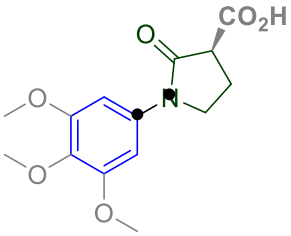
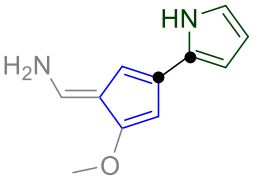
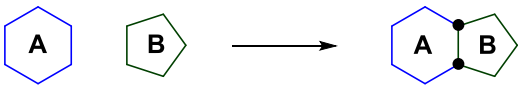
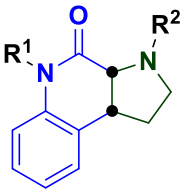
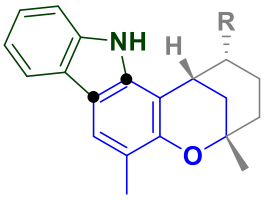
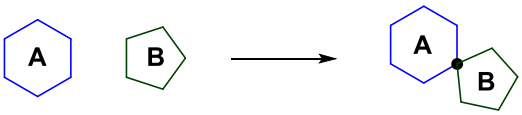
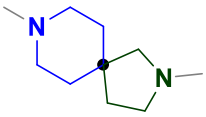
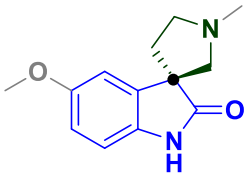
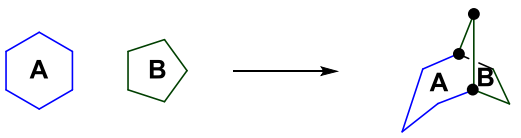
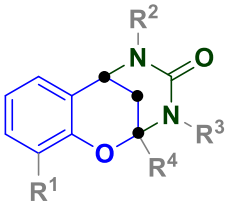
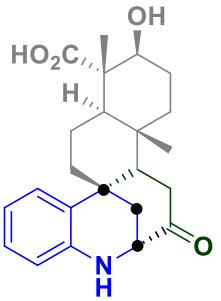
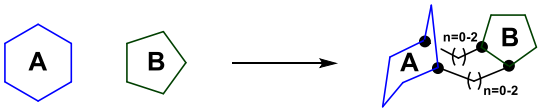
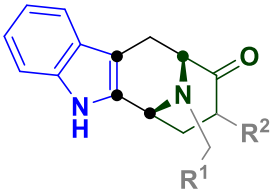
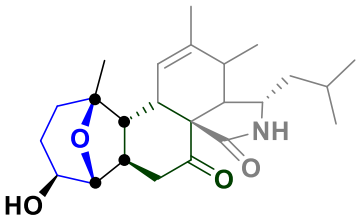
Figures:

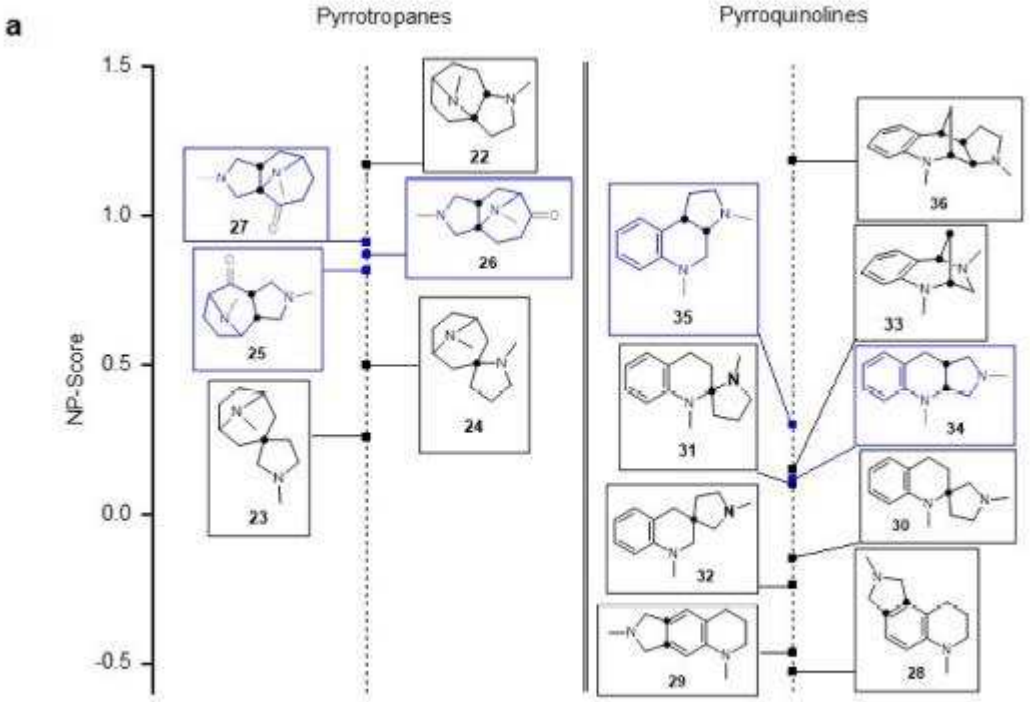

a
Readily acessible natural products
Ring cleavage
Ring rearrangement and fusion
Ring expansion
Ring contraction
Complexity to Diversity (CtD)
b
Asymmetric synthesis
Build
Couple
Stereochemical diversity
Pair
Scaffold diversity
Diversity Oriented Synthesis (DOS)


c
d
Natural product scaffolds
Natural Products
Natural product fragments
Biology Oriented Synthesis (BIOS)
Pseudo natural products (pseudo NPs)

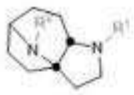

a
b
TMSCl
DMF, rt
1. NaHCO3, MeOH
2. LiOH, THF/H2O
3. HCl, heat
Chromopynones, 1
c
Cu(CH3CN)4PF6
(R)-Fesulphos
DBU (50 mol%)
CH2Cl2, -20°C, 2h
Indotropanes, 2
d
tBu2MeOBIHEP
Cu(CH3CN)4BF4
DIPEA, CH2Cl2, rt.
Pyrrotropanes, 3
e
Cu(CH3CN)4BF4
Et3N (1 eq.), 4Å MS, Et2O
P(3,4,5(MeO)3C6H2)2
P(3,4,5(MeO)3C6H2)2
Pyrroquinolinones, 4
f
RCHO
TFA, r.t.,
3d
NaOMe,
MeOH/toluene
reflux,3d
R = polymeric support
Indopipenones, 5


a
Pseudo-NPs
ChEMBL NPs
Approved Drugs
BIOS
% of compounds
NP-Score
b
Lipinski Rule-of-5
ALogP
MW
c
I2/I3
I1/I3

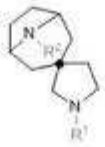
Connection Type	Pseudo Natural Product Example	Natural Product Example
a 	 Phenylpyrrolidinone, 6	 Tambjamine A, 7
b 	 Pyrroquinolinones, 4	 Murraya Alkaloids, 8
c 	 2,8-Diazaspiro[4.5]decane, 9	 (–)-Horsfiline, 10
d 	 Chromopynones, 1	 Sespenine, 11
e 	 Indopipenones, 5	 Aspergillin PZ, 12



b **Design Principle 1:**



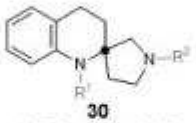
22
NP-Score = 1.18



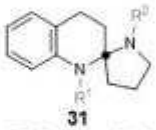
23
NP-Score = 0.26

c **Design Principle 2:**

24
NP-Score = 0.50

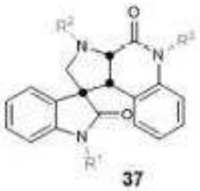


30
NP-Score = -0.15



31
NP-Score = 0.10

d **Design Principle 3:**



37