Supplementary information

Unifying principles for the design and evaluation of natural product-inspired compound collections

Frederik Simonsen Bro,[1] Luca Laraia[1*]

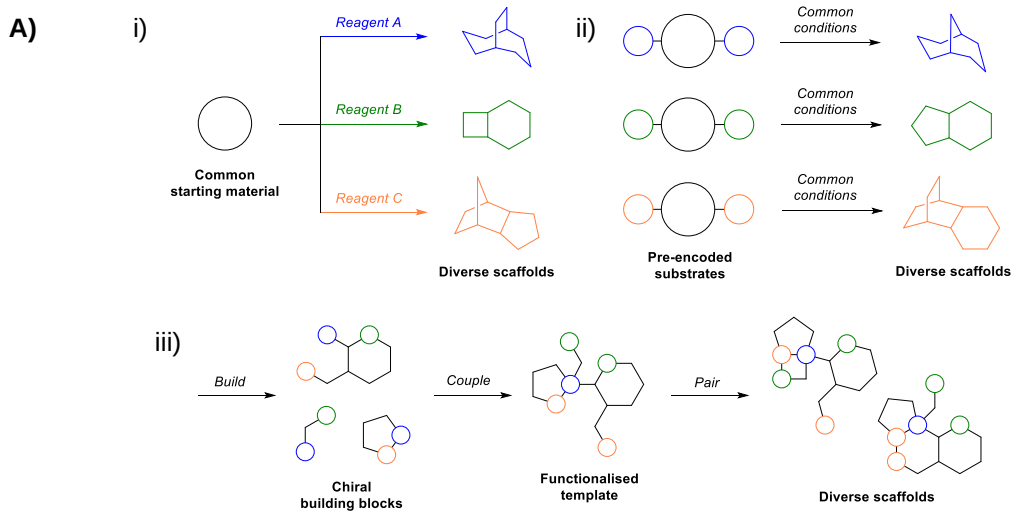
[1]Department of Chemistry, Technical University of Denmark, 2800, Kongens Lyngby, Denmark
*Correspondence: luclar@kemi.dtu.dk

Contents

Supplementary figures

Strategies

Diversity-oriented synthesis (DOS)



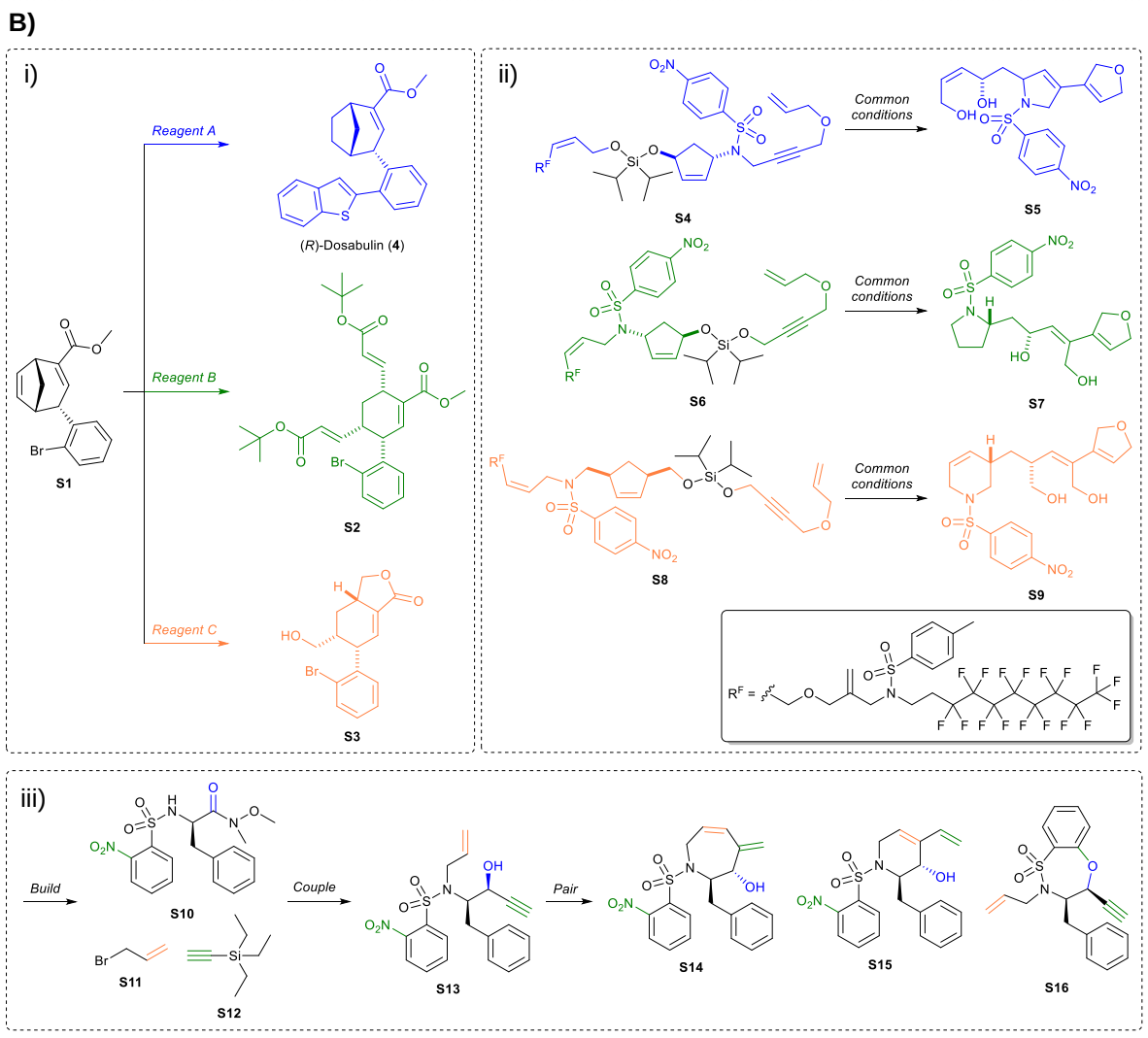


Figure S1: **A)** Graphical illustration of the DOS strategy and the three main underlying approaches: i) reagent-based DOS; ii) substrate-based DOS; and iii) build/couple/pair (B/C/P) DOS.[1–4] Figure inspired from references.[3,4] The DOS strategy seeks to make collections of structural diverse and complex compounds in an efficient way in order to cover as much of chemical space as possible. Compounds are designed using forward synthetic analysis to maximise complexity and diversity. In the reagent-based approach, a common starting material (SM) is transformed or paired intramolecularly by using different reagents to create different molecular skeletons. In the substrate-based DOS, common reaction conditions are used on different "pre-encoded" SMs that will give distinct molecular scaffolds in a folding approach. The B/C/P strategy is a three-step approach where the "build" step is the asymmetric synthesis of chiral building blocks, the "couple" step is the intermolecular coupling of the chiral building blocks, and the "pair" step is the intramolecular coupling of the incorporated functionalities in different combinations to provide skeletal diversity. **B)** Representative examples of the strategy. i) Reagent-based DOS to generate diverse scaffolds from the common starting material **S1**.[5] Substrate-based DOS to generate diverse scaffolds using common conditions from pre-encoded substrates.[6] B/C/P DOS to generate diverse scaffolds from the functionalised template **S13**.[7]

Biology-oriented synthesis (BIOS)

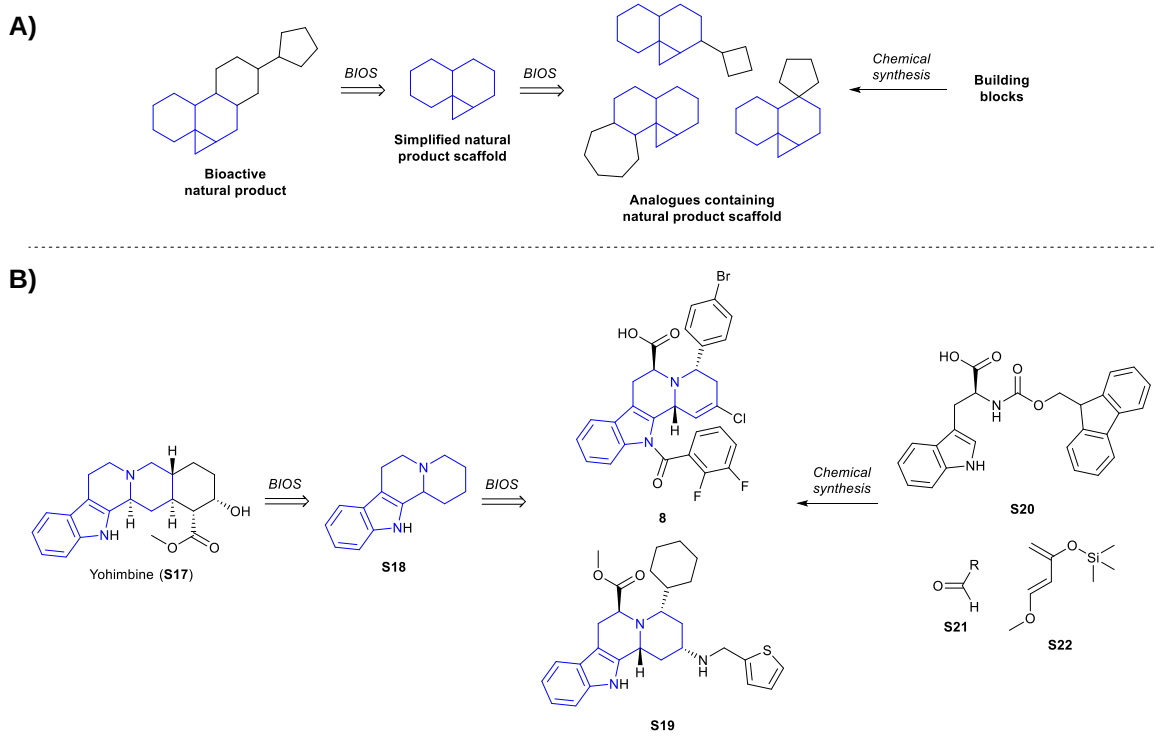


Figure S2: **A)** Graphical illustration of the BIOS strategy.[8–11] The BIOS strategy seeks to design compounds and libraries that cover more biologically relevant areas of chemical space by bridging the gap between synthetic chemistry and natural products (NPs). It takes advantage of the inherent biological relevance of NPs and the structural conservatism found in NPs and protein binding sites. NPs are simplified to NP-inspired scaffolds while maintaining biological relevance of the guiding NP. Thus, BIOS aim to synthesise compound collections where the core scaffold from a guiding NP is retained in the compounds. The assumption is that a high degree of the bioactivity of the guiding NP should be preserved, though not necessarily against the same target(s). **B)** Representative example of the strategy. Synthesis of substituted indolo[2,3-*a*]quinolizidines (**S18**).[8,12]

Pseudo-natural product (PNP)

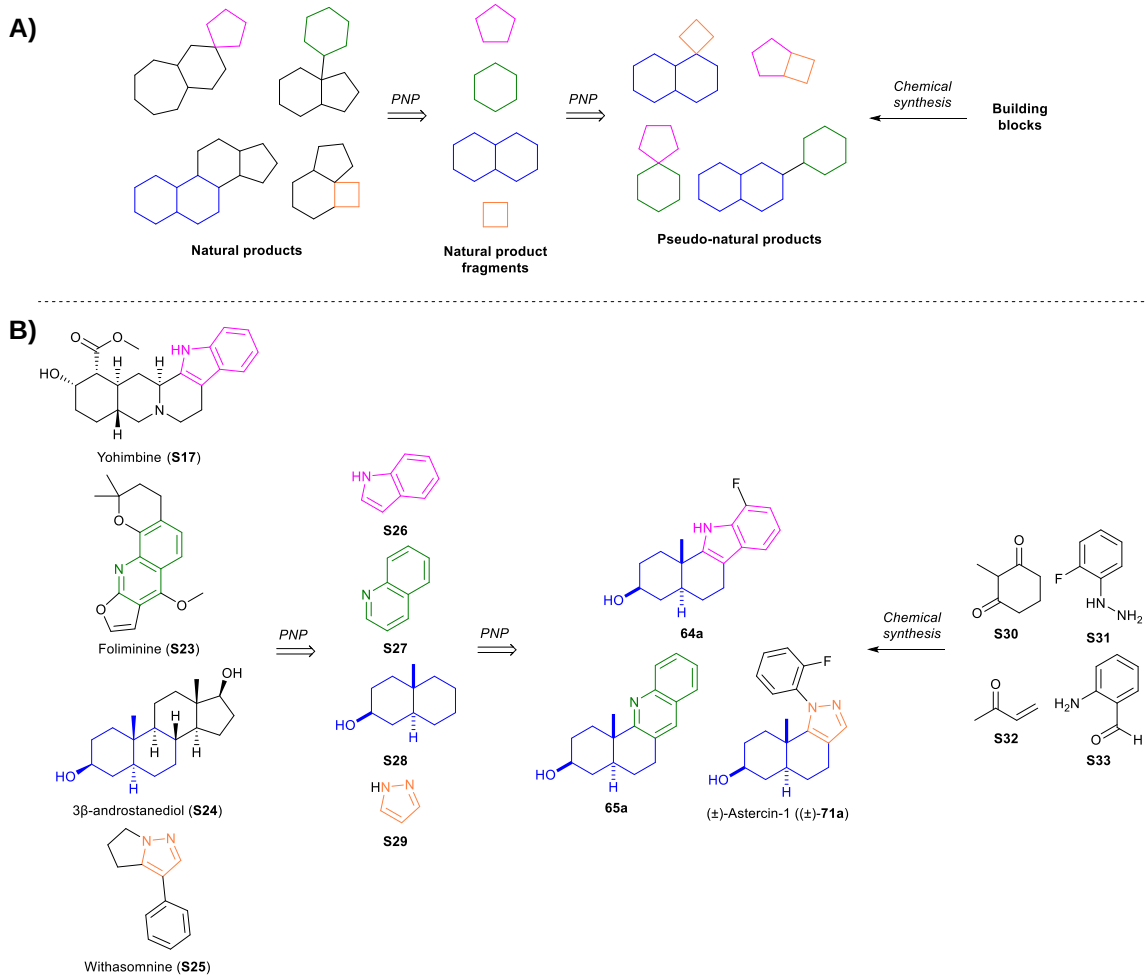


Figure S3: **A)** Graphical illustration of the PNP strategy.[13–15] The PNP strategy is a fragment-based approach that takes advantage of the biologically pre-validated NPs. The compounds in the library are designed by identifying NP fragments and combining these in unprecedented ways not found in NPs and unavailable by biosynthetic pathways. In this way, chemical space not covered by NPs is reached and compounds with potentially new bioactivities can be identified. **B)** Representative example of the strategy. Synthesis of sterol-inspired compounds.[16]

Privileged-substructure-based diversity-oriented synthesis (pDOS)

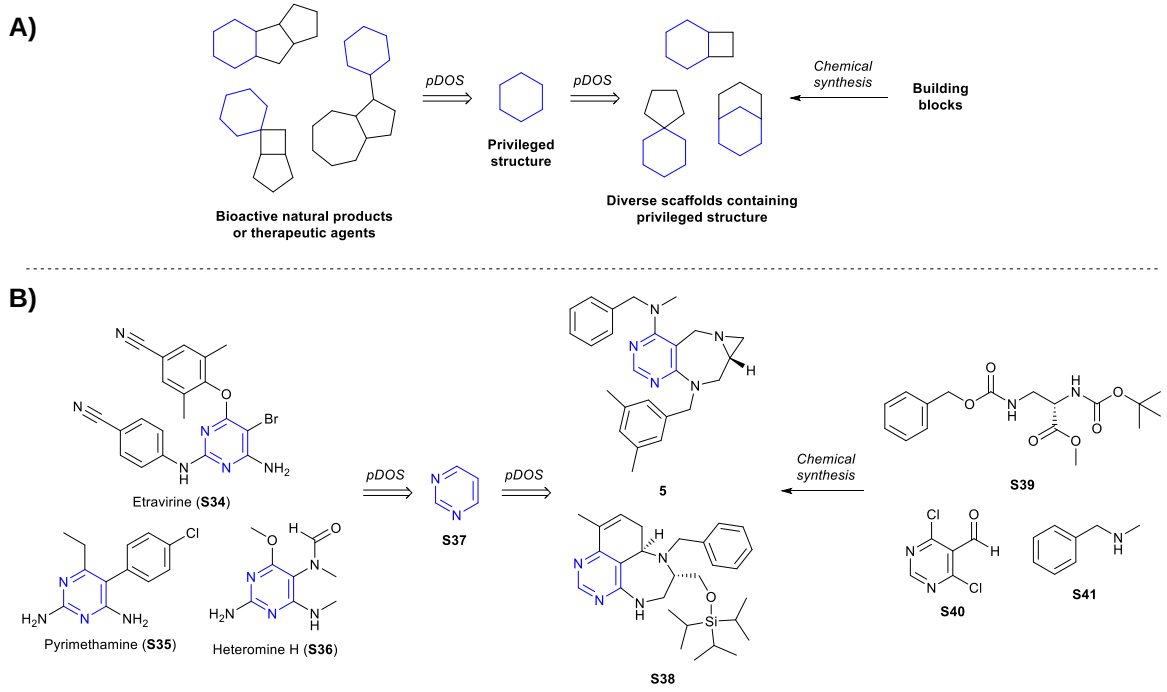


Figure S4: **A)** Graphical illustration of the pDOS strategy.[17,18] Figure inspired from references.[17,18] The pDOS strategy aims to maximise diversity within biologically relevant chemical space. This is achieved by employing so-called "privileged structures" and DOS strategies to create libraries of compounds with diverse polyheterocyclic scaffolds containing the privileged structure. The strategy seeks to construct complex and diverse scaffolds with the privileged structure incorporated. **B)** Representative example of the strategy. Synthesis of substituted pyrimidines (**S37**).[19]

Complexity-to-diversity (CtD)

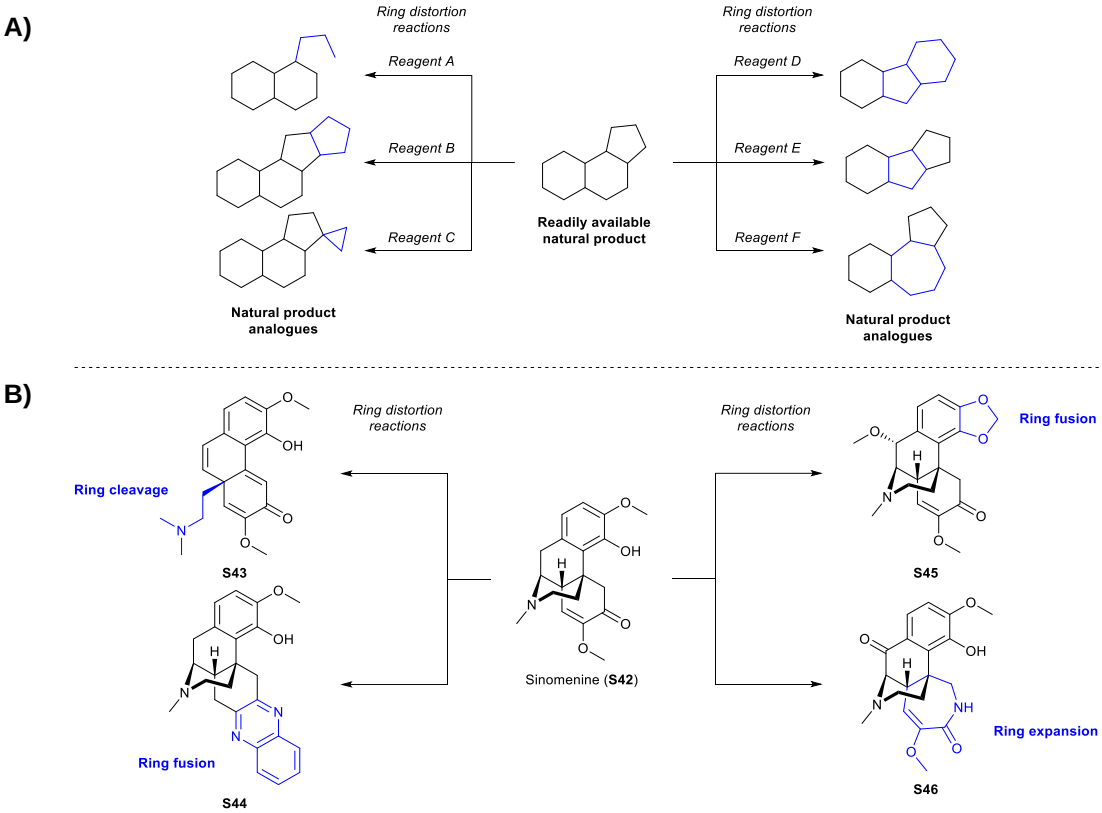


Figure S5: **A)** Graphical illustration of the CtD strategy.[20] Figure inspired from references.[21,22] In this approach, readily available complex NPs are subjected to various ring distortions and/or ring modifications to give NP analogues with high scaffold and stereochemical diversity. The strategy uses ring distortions such as ring cleavage/opening, ring formation/closure, ring egde/spiro-fusion, ring expansion, ring contraction, ring rearrangement, ring (de)aromatisation, and combinations of the above to create a collection of diverse molecules derived from the parent NP. **B)** Representative example of the strategy. Synthesis of sinomemine (**S42**) analogues.[23]

Activity-directed synthesis (ADS)

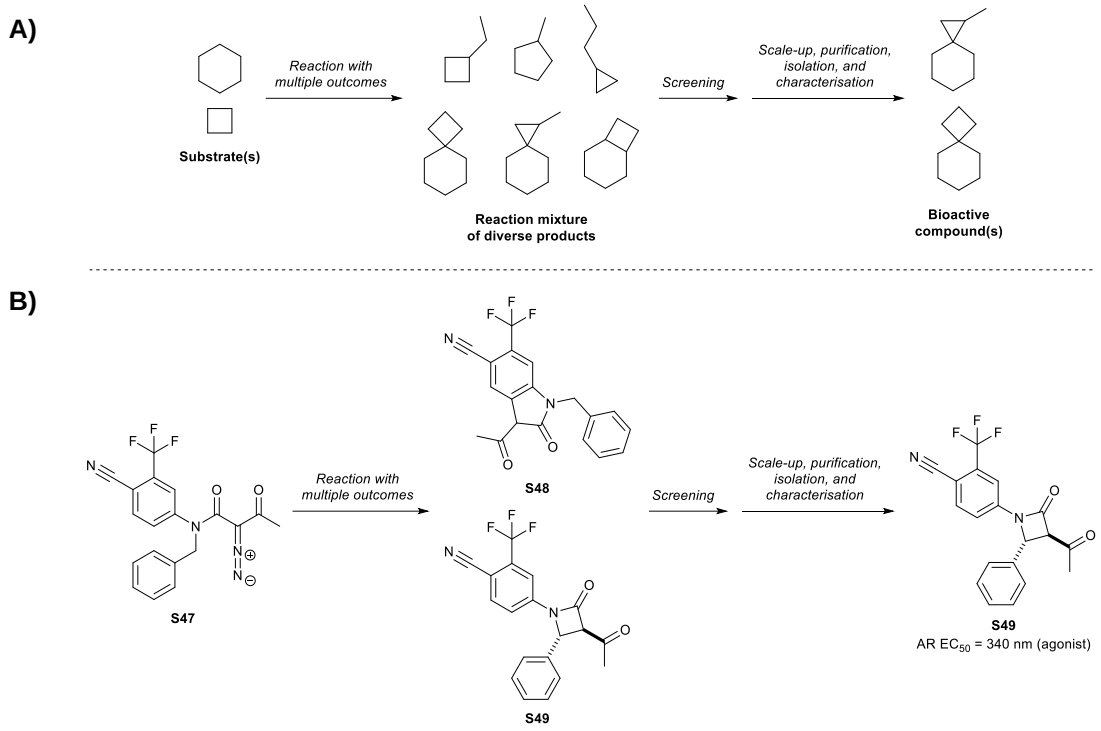


Figure S6: **A)** Graphical illustration of the ADS strategy.[24,25] In the ADS strategy, the compounds of the "library" are not purified or isolated initially. It is a function-driven approach with focus on molecular scaffold diversity of the compounds. It takes advantage of promiscuous reactions with multiple possible outcomes to generate mixtures of diverse products. Several rounds of reaction arrays can be performed by changing reaction variables such as substrates, co-substrates, catalysts, solvents etc. to give different products. The crude reaction product mixtures are screened, and the most promising hit reactions are scaled-up and purified to isolate and characterise the products and identify the active compound(s). **B)** Representative example of the strategy. Identification of androgen receptor (AR) agonist (**S49**).[24]

Diverse pseudo-natural product (dPNP)

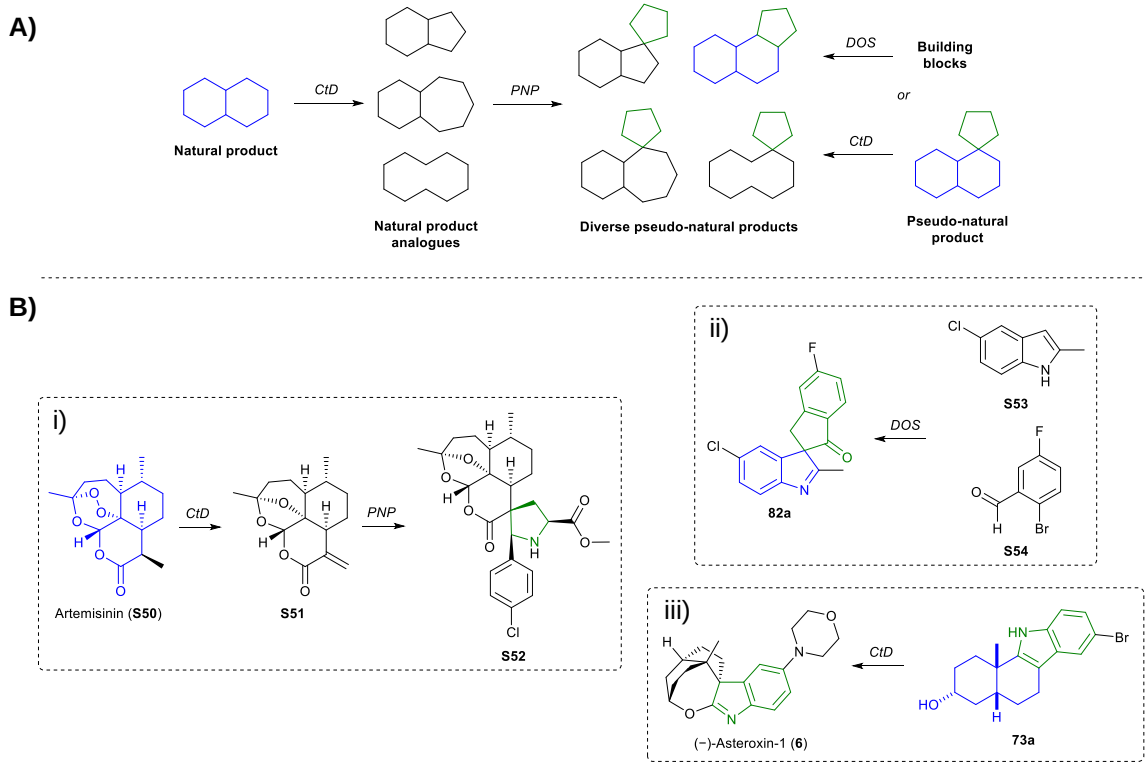


Figure S7: **A)** Graphical illustration of the dPNP strategy.[26] The dPNP strategy is defined as the combination of the diversification strategies such as DOS and CtD with the PNP strategy to generate compounds with high scaffold diversity and biological relevance. There are currently no strict rules on how the strategies should be combined. **B)** Representative examples of the strategy. i) Using CtD on artemisinin (**S50**) followed by fusion with NP fragments.[27] ii) Using DOS to generate PNPs.[26] iii) Using PNPs in a CtD setting.[28]

Function-oriented synthesis (FOS)

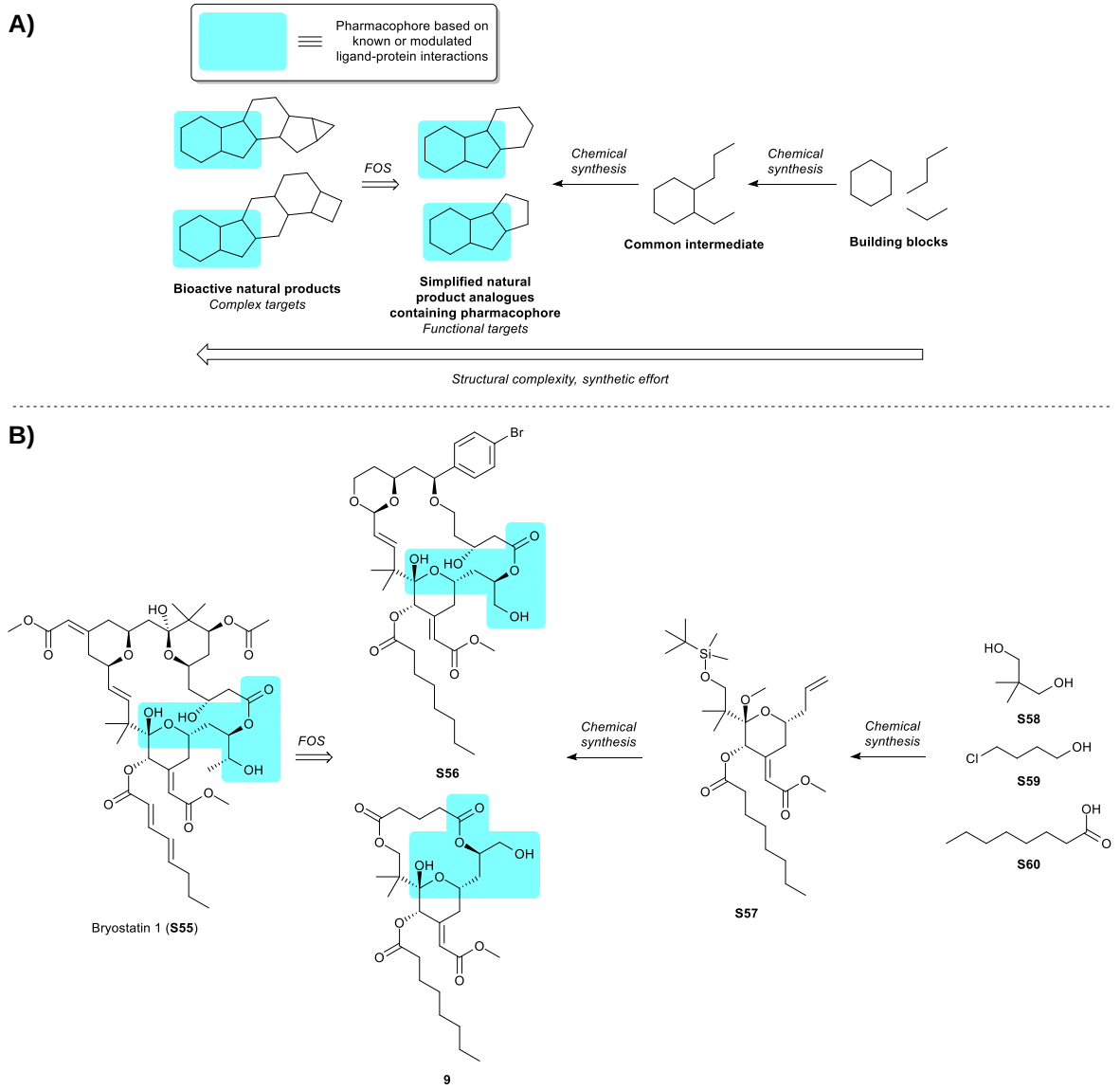


Figure S8: **A)** Graphical illustration of the FOS strategy.[29] Figure inspired from references.[30,31] The FOS strategy aims to bring function of a NP in to the retrosynthetic analysis (RA). This is achieved by targeting the function of the NP and not the structure using "retrofunction" analysis in the forward synthetic design. Thus, the NP is not necessarily seen as a synthetic target but as useful information that combines structure and function (biological activity). The key point of FOS is that by making simplified NP analogues with activity-determining features incorporated, function can be emulated, tuned, or improved. **B)** Representative example of the strategy. Synthesis of bryostatin 1 (**S55**) analogues **S56** and **9**.[32–34]

Pharmacophore-directed retrosynthesis (PDR)

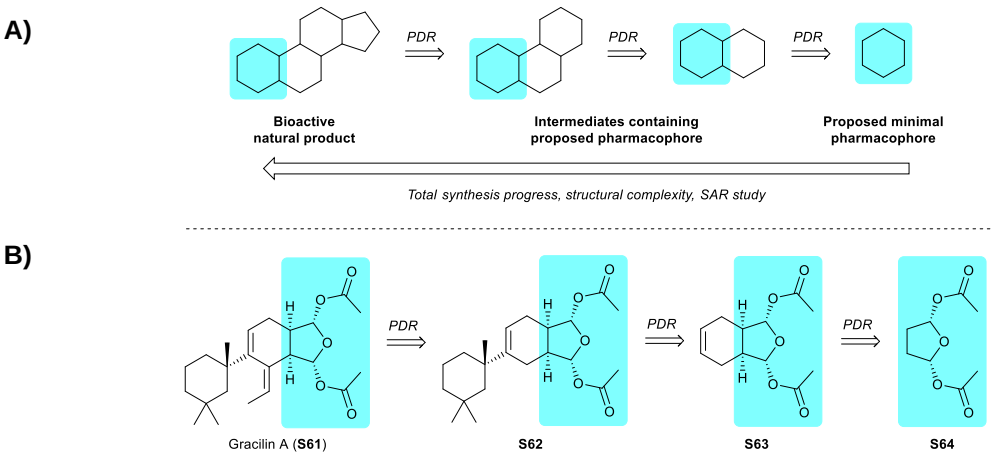


Figure S9: **A)** Graphical illustration of the PDR strategy.[30] Figure inspired and from references.[30,31] In the PDR strategy a proposed minimal pharmacophore of a guiding NP is the synthetic target. This minimal pharmacophore should be accessed as soon as possible, thus guiding the RA. The modified RA results in simplified intermediates containing the proposed pharmacophore with increasing complexity *en route* to the actual NP. It enables structure-activity relationship (SAR) studies as the total synthesis progresses. **B)** Representative example of the strategy. Synthesis of gracillin A (**S61**) analogues.[30]

Dynamic retrosynthetic analysis (DRA)

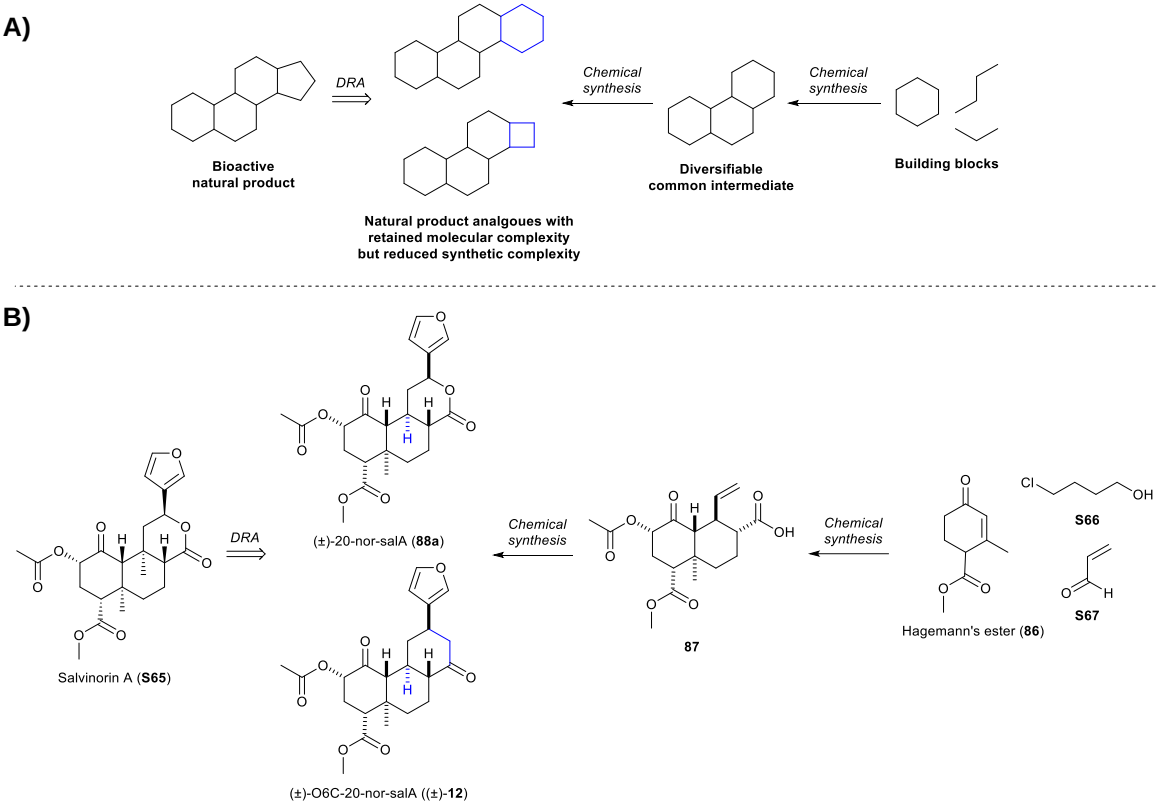


Figure S10: **A)** Graphical illustration of the DRA strategy.[35–37] In the DRA strategy, a NP is treated as a dynamic target and not static as in traditional RA. It is very useful for bioactive NPs where their syntheses are limited by chemical or metabolic instability and/or synthetic methodological gaps. The goal is to retain molecular complexity, reduce synthetic complexity, and retain bioactivity through perturbation of the initial NP target. The initial NP target serve as inspiration and a variable (not a constant) that can be changed to give the shortest and most efficient RA and synthesis of NP analogues with maintained complexity. Minimal and rational changes in the RA to improve chemical stability can simplify the synthesis of NP analogues. **B)** Representative example of the strategy. Synthesis of salvinorin A (**S65**) analogues (±)-20-nor-salA (**88a**) and (±)-O6C-20-nor-salA ((±)-**12**).[38,39]

Diverted total synthesis (DTS)

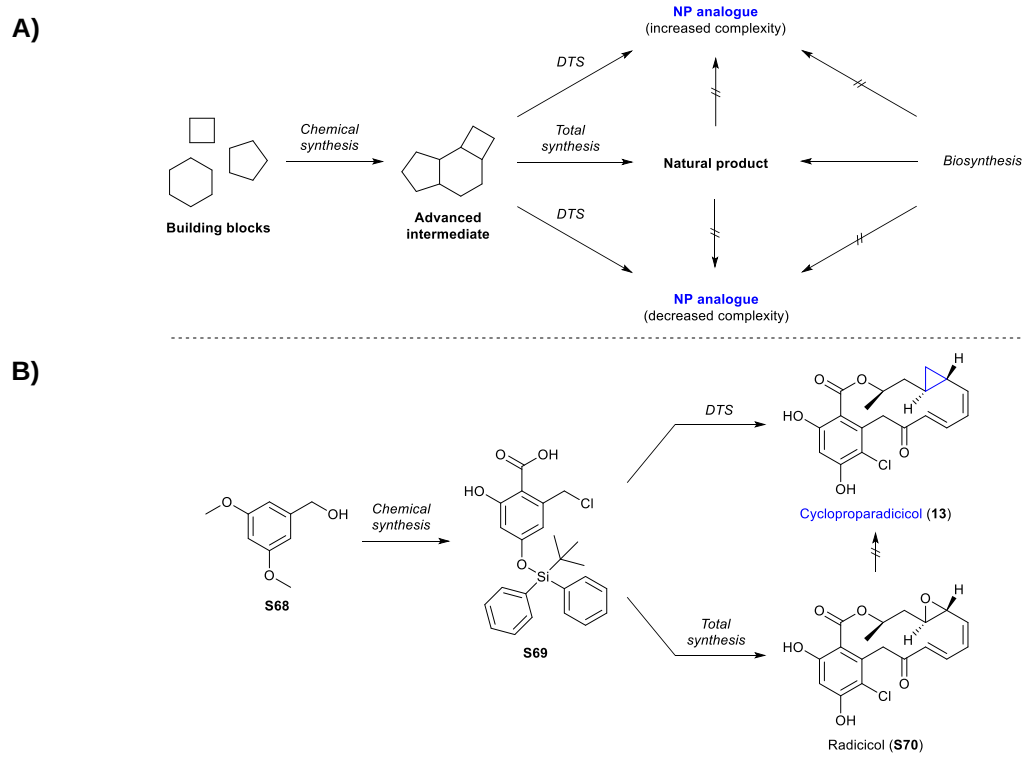


Figure S11: **A)** Graphical illustration of the DTS strategy.[40,41] Figure inspired from reference.[40] The DTS strategy seeks to investigate the chemical space around a parent NP through the synthesis of analogues of the parent NP. The analogues are synthesised from an advanced intermediate. The advanced intermediate is on the route to the NP but before reaching the NP the intermediate is derivatised to access analogues with higher or lower chemical complexity than the parent NP. A central point in DTS is that the analogues in most cases cannot be accessed directly from the NP due to chemical limitations. **B)** Representative example of the strategy. Synthesis of radicicol (**S70**) and analogue cyclopropparadicicol (**13**).[42–44]

Analogue-oriented synthesis (AOS)

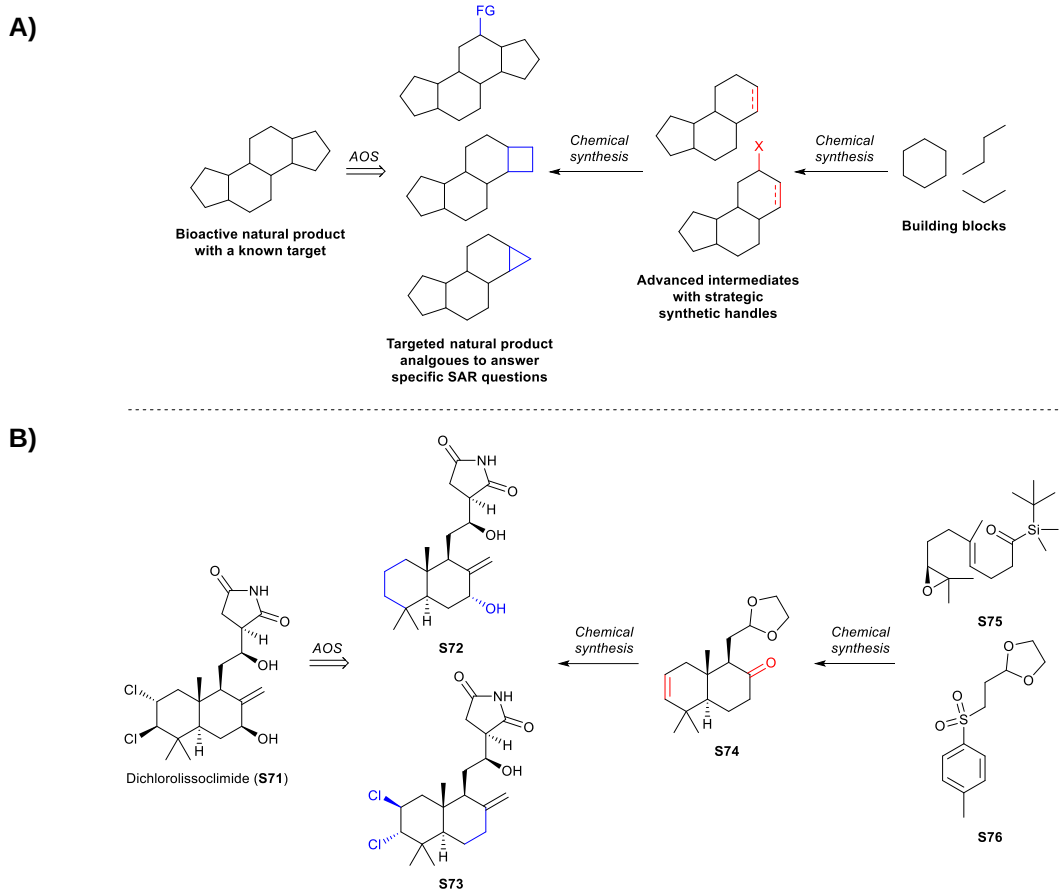


Figure S12: **A)** Graphical illustration of the AOS strategy.[45] Figure inspired from reference.[31] The AOS strategy aims to synthesise targeted and specific analogues of a NP with a known target in order to answer specific SAR questions. Thus, the SAR questions influence the design of the NP analogues. The NP analogues are synthesised through advanced intermediates with strategically placed functionalities and synthetic handles/vectors. **B)** Representative example of the strategy. Synthesis of lissoclimide analogues.[45]

Two-phase synthesis (TPS)

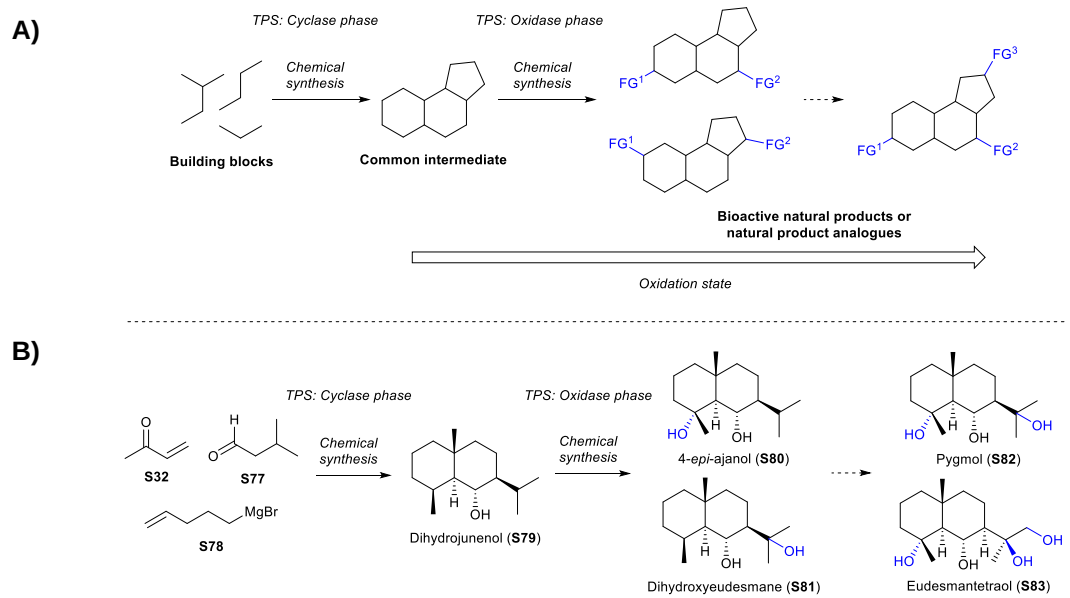


Figure S13: **A)** Graphical illustration of the TPS strategy.[46] Figure inspired from reference.[31] The TPS strategy is inspired by nature's biosynthetic assembly of some NPs. It consists of a "cyclase" and a "oxidase" phase. In the "cyclase" phase the carbon skeleton is constructed to give a common intermediate of the lowest oxidation state possible *en route* to the NP(s). In the "oxidase" phase the common intermediate is oxidised to afford the targeted NPs and their analogues. **B)** Representative example of the strategy. Synthesis of eudesmane terpenes.[47,48]

References

1 S. L. Schreiber, *Science*, 2000, **287**, 1964–1969.
2 M. D. Burke and S. L. Schreiber, *Angew. Chem. Int. Ed.*, 2004, **43**, 46–58.
3 T. E. Nielsen and S. L. Schreiber, *Angew. Chem. Int. Ed.*, 2008, **47**, 48–56.
4 W. R. J. D. Galloway, A. Isidro-Llobet and D. R. Spring, *Nat. Commun.*, 2010, **1**, 80.
5 B. M. Ibbeson, L. Laraia, E. Alza, C. J. O' Connor, Y. S. Tan, H. M. L. Davies, G. McKenzie, A. R. Venkitaraman and D. R. Spring, *Nat. Commun.*, 2014, **5**, 3155.
6 D. Morton, S. Leach, C. Cordier, S. Warriner and A. Nelson, *Angew. Chem. Int. Ed.*, 2009, **48**, 104–109.
7 D. Pizzirani, T. Kaya, P. A. Clemons and S. L. Schreiber, *Org. Lett.*, 2010, **12**, 2822–2825.
8 A. Nören-Müller, I. Reis-Corrêa, H. Prinz, C. Rosenbaum, K. Saxena, H. J. Schwalbe, D. Vestweber, G. Cagna, S. Schunk, O. Schwarz, H. Schiewe and H. Waldmann, *Proc. Natl. Acad. Sci.*, 2006, **103**, 10606–10611.
9 R. S. Bon and H. Waldmann, *Acc. Chem. Res.*, 2010, **43**, 1103–1114.
10 S. Wetzel, R. S. Bon, K. Kumar and H. Waldmann, *Angew. Chem. Int. Ed.*, 2011, **50**, 10800–10826.
11 H. van Hattum and H. Waldmann, *J. Am. Chem. Soc.*, 2014, **136**, 11853–11859.
12 I. R. Corrêa Jr., A. Nören-Müller, H.-D. Ambrosi, S. Jakupovic, K. Saxena, H. Schwalbe, M. Kaiser and H. Waldmann, *Chem. Asian J.*, 2007, **2**, 1109–1126.
13 G. Karageorgis, E. S. Reckzeh, J. Ceballos, M. Schwalfenberg, S. Sievers, C. Ostermann, A. Pahl, S. Ziegler and H. Waldmann, *Nat. Chem.*, 2018, **10**, 1103–1111.
14 G. Karageorgis, D. J. Foley, L. Laraia and H. Waldmann, *Nat. Chem.*, 2020, **12**, 227–235.
15 G. Karageorgis, D. J. Foley, L. Laraia, S. Brakmann and H. Waldmann, *Angew. Chem. Int. Ed.*, 2021, **60**, 15705–15723.
16 T. Whitmarsh-Everiss, A. H. Olsen and L. Laraia, *Angew. Chem. Int. Ed.*, 2021, **60**, 26755–26761.
17 S. Oh and S. B. Park, *Chem. Commun.*, 2011, **47**, 12754–12761.
18 J. Kim, H. Kim and S. B. Park, *J. Am. Chem. Soc.*, 2014, **136**, 14629–14638.
19 J. Kim, J. Jung, J. Koo, W. Cho, W. S. Lee, C. Kim, W. Park and S. B. Park, *Nat. Commun.*, 2016, **7**, 13196.
20 R. W. Huigens III, K. C. Morrison, R. W. Hicklin, T. A. Flood Jr, M. F. Richter and P. J. Hergenrother, *Nat. Chem.*, 2013, **5**, 195–202.
21 S. Yi, B. V. Varun, Y. Choi and S. B. Park, *Front. Chem.*, 2018, **6**, 507.
22 T. G. Adessi, P. M. Wagner, F. R. Bisogno, V. E. Nicotra, M. E. Guido and M. E. García, *Sci. Rep.*, 2024, **14**, 14229.
23 A. Garcia, B. S. Drown and P. J. Hergenrother, *Org. Lett.*, 2016, **18**, 4852–4855.
24 G. Karageorgis, S. Warriner and A. Nelson, *Nat. Chem.*, 2014, **6**, 872–876.
25 G. Karageorgis, S. Liver and A. Nelson, *ChemMedChem*, 2020, **15**, 1776–1782.
26 S. Bag, J. Liu, S. Patil, J. Bonowski, S. Koska, B. Schölermann, R. Zhang, L. Wang, A. Pahl, S. Sievers, L. Brieger, C. Strohmann, S. Ziegler, M. Grigalunas and H. Waldmann, *Nat. Chem.*, 2024, **16**, 945–958.
27 J. Liu, J. Flegel, F. Otte, A. Pahl, S. Sievers, C. Strohmann and H. Waldmann, *Angew. Chem. Int. Ed.*, 2021, **60**, 21384–21395.
28 F. S. Bro, L. Depta, N. J. Dekker, H. P. Bryce-Rogers, M. L. Madsen, K. F. Præstegaard, T. Petersson, T. Whitmarsh-Everiss, M. Kubus and L. Laraia, *ACS Cent. Sci.*, 2025, **11**, 136–146.
29 P. A. Wender, V. A. Verma, T. J. Paxton and T. H. Pillow, *Acc. Chem. Res.*, 2008, **41**, 40–49.
30 M. E. Abbasov, R. Alvariño, C. M. Chaheine, E. Alonso, J. A. Sánchez, M. L. Conner, A. Alfonso, M. Jaspars, L. M. Botana and D. Romo, *Nat. Chem.*, 2019, **11**, 342–350.
31 N. J. Truax and D. Romo, *Nat. Prod. Rep.*, 2020, **37**, 1436–1453.
32 P. A. Wender, J. L. Sloane, Q. H. Luu-Nguyen, Y. Ogawa, A. J. Shimizu, S. M. Ryckbosch, J. H. Tyler and C. Hardman, *J. Org. Chem.*, 2020, **85**, 15116–15128.

33 P. A. Wender and J. C. Horan, *Org. Lett.*, 2006, **8**, 4581–4584.
34 P. A. Wender, J. L. Baryza, C. E. Bennett, F. C. Bi, S. E. Brenner, M. O. Clarke, J. C. Horan, C. Kan, E. Lacôte, B. Lippa, P. G. Nell and T. M. Turner, *J. Am. Chem. Soc.*, 2002, **124**, 13648–13649.
35 J. J. Roach, Y. Sasano, C. L. Schmid, S. Zaidi, V. Katritch, R. C. Stevens, L. M. Bohn and R. A. Shenvi, *ACS Cent. Sci.*, 2017, **3**, 1329–1336.
36 B. J. Huffman and R. A. Shenvi, *J. Am. Chem. Soc.*, 2019, **141**, 3332–3346.
37 S. Woo and R. A. Shenvi, *Acc. Chem. Res.*, 2021, **54**, 1157–1167.
38 S. J. Hill, N. Dao, V. Q. Dang, E. L. Stahl, L. M. Bohn and R. A. Shenvi, *ACS Cent. Sci.*, 2023, **9**, 1567–1574.
39 S. Hirasawa, M. Cho, T. F. Brust, J. J. Roach, L. M. Bohn and R. A. Shenvi, *Bioorg. Med. Chem. Lett.*, 2018, **28**, 2770–2772.
40 J. T. Njardarson, C. Gaul, D. Shan, X.-Y. Huang and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2004, **126**, 1038–1040.
41 R. M. Wilson and S. J. Danishefsky, *J. Org. Chem.*, 2006, **71**, 8329–8351.
42 R. M. Garbaccio, S. J. Stachel, D. K. Baeschlin and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2001, **123**, 10903–10908.
43 K. Yamamoto, R. M. Garbaccio, S. J. Stachel, D. B. Solit, G. Chiosis, N. Rosen and S. J. Danishefsky, *Angew. Chem. Int. Ed.*, 2003, **42**, 1280–1284.
44 X. Lei and S. J. Danishefsky, *Adv. Synth. Catal.*, 2008, **350**, 1677–1681.
45 Z. A. Könst, A. R. Szklarski, S. Pellegrino, S. E. Michalak, M. Meyer, C. Zanette, R. Cencic, S. Nam, V. K. Voora, D. A. Horne, J. Pelletier, D. L. Mobley, G. Yusupova, M. Yusupov and C. D. Vanderwal, *Nat. Chem.*, 2017, **9**, 1140–1149.
46 Y. Ishihara and P. S. Baran, *Synlett*, 2010, **12**, 1733–1745.
47 K. Chen and P. S. Baran, *Nature*, 2009, **459**, 824–828.
48 K. Chen, Y. Ishihara, M. M. Galán and P. S. Baran, *Tetrahedron*, 2010, **66**, 4738–4744.