1

Introduction

Facets of Polypharmacology – a Janus-Headed Concept for Drug Discovery

Jürgen Bajorath

1.1 Origins

Since the 1980s, target-centric approaches have dominated drug discovery efforts, triggered by the molecular-biology-driven reductionist approach [1] and leading to the "one drug, one target," or "drug specificity" paradigm [2]. Molecular reductionism aimed at "dissecting biological systems into their constituent parts" [1]. Different from the preceding more holistic and pharmacology-oriented era in drug discovery, molecular sciences and the single-target (ST) focus took the centerstage and shaped drug discovery efforts for many years to come [1, 2]. These developments were paralleled by advances in X-ray crystallography and molecular graphics catalyzing a wave of structure-based ("rational") drug design efforts [3, 4], which further emphasized the focus on target-specific compounds in drug discovery.

In the early 2000s, systems biology emerged [5] and also entered the drug discovery arena [6] introducing, for example, network representations of biological systems, pathway modeling, and computational disease models. These developments originating from bioinformatics also altered the view of traditional disciplines such as pharmacology, giving rise to a network perception of physiological processes and increasing the notion of their interdependence [7]. In pharmacological networks, multiple signaling and metabolic pathways establish functional links and dependencies between different target proteins [7, 8]. If pathways in such networks are perturbed or regulatory and control mechanisms compromised, different types of multifactorial diseases might arise, including various forms of cancer, complex diseases of the central nervous system, or metabolic diseases [9–12]. Such diseases could most likely not be effectively treated by therapeutic intervention of individual targets, but required multi-target (MT) engagement instead, thus departing from the target specificity paradigm in drug discovery. MT activity of drugs was not unknown and probably first observed for anti-psychotics and antiepileptics beginning in the late 1980s [12, 13].

In 2006, as a consequence of the increasing notion of pharmacological networks, the concept of polypharmacology was introduced [14], focusing on MT agents for the treatment of multifactorial diseases: "*Contrary to the dogma that the 'rational' way to discover drugs is to design exquisitely selective ligands for single molecular targets, a rival hypothesis proposes polypharmacology or the promiscuous modulation of several molecular targets*" [14]. In 2014, a formal definition of polypharmacology appeared in the US National Library of Medicine (NLM) as "*the design or use of pharmaceutical agents that act on multiple targets or disease pathways*." Polypharmacology also encompasses the pharmacological effects resulting from the use of MT compounds (MT-CPDs), consistent with the principles of network pharmacology. MT activity of drugs and other bioactive compounds is often also referred to as "promiscuity" (not to be confused with nonspecific compound–protein interactions).

1.2 Pros and Cons

Following its inception, polypharmacology emerged as an alternative to reductionist approaches and rational drug design and further evolved into a multifaceted drug discovery strategy [15–17], albeit "Janus-headedly." In Roman mythology, Janus, the god of the beginnings, passages, and endings, had two opposing faces. Accordingly, the "Janus head" became a symbol of duality and ambivalence – which exactly applied to the polypharmacology concept: on the one hand, MT activity of drugs is a prerequisite for therapeutic efficacy in the treatment of multifactorial diseases; on the other, it is responsible for unwanted (adverse) side effects [15, 18, 19]. While adverse side effects can be elicited by the engagement of a primary target, they are more frequently caused by inhibiting so-called anti-targets

Polypharmacology: Strategies for Multi-Target Drug Discovery, First Edition. Edited by Jens-Uwe Peters.

such as cardiac ion channels (hERG), drug-metabolizing cytochrome P450 isoforms, or G-protein-coupled neurotransmitter receptors [15, 16]. Furthermore, side effects of MT-CPDs might also be caused by interacting with other targets not implicated in immediate toxicity, due to pathway modulations. In the pharmaceutical industry, potential liabilities as a consequence of interactions with anti-targets are a major concern, for example, leading to the assessment of newly identified candidate compounds in various safety screens for activity against such targets. However, not all unexpected side effects are undesired, taking into consideration that MT activity also provides the basis for drug repurposing [20]. Benefits of MT activity of drugs were often discovered post hoc. For example, adenosine triphosphate (ATP)-site-directed kinase inhibitors used in cancer therapy were originally thought to be kinase-selective, based on reductionist assessment, before it was discovered that their clinical efficacy depended on multi-kinase activity and simultaneous interference with multiple deregulated signaling pathways [21]. This also applied to imatinib, the first kinase inhibitor marketed as a drug [21].

Despite the Janus-headed nature of polypharmacology and the risks associated with potential adverse side effects resulting from the MT activity of drugs, the positive impact of polypharmacology on drug discovery and development is undeniable, as demonstrated by the continuous occurrence of MT agents among newly approved drugs. For example, 10 of 49 European Medicines Agency (EMA)-approved drugs marketed in Germany in 2022 were annotated with two or more targets [22]. Of course, despite the strong impact of polypharmacology, the development of compounds with target selectivity or specificity continues to be a pillar of drug discovery and development. For example, for long-term treatment of chronic and non-life-threatening diseases, drug side effects must inevitably be minimized, rendering target-selective compounds highly desirable.

1.3 Discovery and Design

Similar to coincidental findings that side effects of drugs originally thought to be specific were caused by previously unknown secondary targets, new MT-CPDs are often discovered serendipitously, for example, in screening campaigns or target deconvolution of active compounds from phenotypic assays. Given the high interest in compounds with defined MT activity in different therapeutic areas, prospective design of such compounds is also a topical issue in drug discovery [23, 24]. However, consistent with findings that characteristic structural features of MT-CPDs generally depend on target combinations, as further discussed below, the prospective design of MT-CPDs with desired activity is challenging, mostly carried out on a case-by-case basis in medicinal chemistry and far from being routine. For all practical purposes, prospective design of MT-CPDs for polypharmacology is limited to two or at most three targets. To this end, combining or merging target-dependent pharmacophores is a popular knowledge-based approach for MT-CPD design that is readily applicable in the practice of medicinal chemistry and does not require sophisticated computations [23–25]. Pharmacophore fusion attempts can be further extended by screening of test compounds using pharmacophore models for different targets and follow-up analysis of shared hits [26]. As an alternative to pharmacophore modeling, scaffolds isolated from compounds with known activity against different targets can also be used as templates for MT-CPD design, as further discussed below.

In addition to knowledge-based design strategies that are close to practical medicinal chemistry, other ligand- or target-structure-based computational approaches have been applied to identify compounds for polypharmacology [27, 28]. For example, various machine learning (ML) models have been reported to distinguish between compounds with MT activity and corresponding ST activity (typically achieving reasonable to high prediction accuracy). Furthermore, ML models have been used for computational target profiling. Here, test compounds are virtually screened using large numbers of individually derived target-based models to predict MT-CPDs. As a deep learning alternative, multitask models have also been developed to predict compounds with activity against related targets. At the structural level, similarities of binding sites in different targets have been quantified as an indicator of polypharmacology potential at the target level. In addition, parallel docking campaigns or cross-docking screens have been carried out for structure-based target profiling. Furthermore, ligands bound to different proteins have been systematically compared to identify compound pairs with the highest shape similarity to prioritize and evaluate putative cross-target activities [28].

1.4 Structural Data

In addition to its relevance for polypharmacology, the study of MT-CPDs is also of interest from a basic scientific perspective. For example, exploring the mechanisms by which small molecules “multi-specifically” or “pseudo-specifically” interact with different targets helps to better understand these special molecular recognition phenomena. To this end, currently available X-ray structures of complexes formed by MT-CPDs and different proteins provide substantial information. For example, in 2018, we identified 1418 crystallographic MT-CPDs

(>300 Da) in X-ray structures of complexes with different targets available in the Protein Data Bank (PDB) [29, 30]. These MT-CPDs included 702 ligands forming complexes with targets from different protein families (termed multifamily ligands) [30]. Bound conformations of multifamily ligands available in complexes with unrelated targets were compared in detail, revealing a variety of ligand binding modes [31]. In some instances, these ligands conformationally adapted to binding sites having different architectures and chemical features and displayed different binding modes; in others, binding modes were surprisingly conserved in differently shaped active sites. If binding modes of multifamily ligands were conserved, characteristic interaction patterns emerged for targets from a given family that differed from others, hence providing a possible rationale for the conservation of binding modes [31].

As a representative example, Figure 1.1 shows conserved and variable binding modes in different active sites for indomethacin, a nonsteroidal anti-inflammatory drug (NSAID) with known polypharmacology used for the treatment of acute pain and symptoms of osteoarthritis and rheumatoid arthritis.

For 243 of the 702 multifamily ligands, 168 analogue series were detected in the ChEMBL database [32]. These series consisted of a total of 4829 compounds, covered 190 additional targets, and yielded 133 unique analogue series-based scaffolds [30]. Figure 1.2 shows an exemplary scaffold. All analogue series scaffolds were annotated with different target combinations, providing a knowledge base of MT template compounds.

1.5 Activity Data

Rapidly growing volumes of compound activity data provide another information-rich resource for the study of MT-CPDs and polypharmacology. Since the analysis of MT activity is particularly vulnerable to false-positive activity annotations, compound activity data should be carefully curated and potential assay interference compounds [33, 34] or colloidal aggregators [35] should be removed. Indeed, results of MT activity analysis strongly depend on applied data confidence criteria [36], as illustrated in Figure 1.3 for imatinib, suggesting to restrict the assessment of MT-CPDs to high-confidence activity data [36].

There are different facets of MT activity. For instance, it is not very surprising that some active compounds exhibit a tendency to interact with more than one closely related target, such as ATP-site-directed protein kinase inhibitors. By contrast, compounds binding to structurally and functionally unrelated proteins are rather unexpected, but of special interest, from both a basic scientific and a polypharmacology perspective. For example, such compounds might interfere with distinct physiological functions and elicit

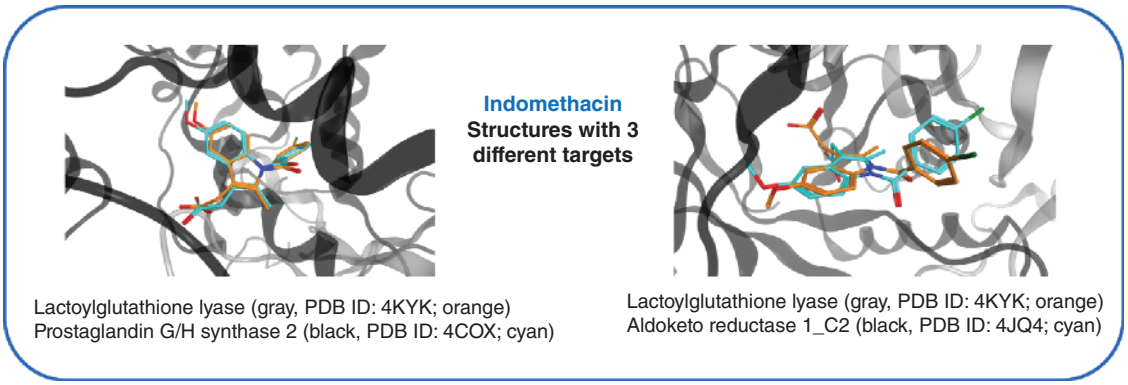
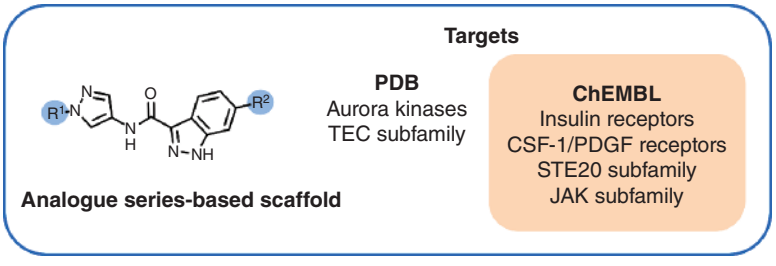


Figure 1.1 X-ray structures of indomethacin in complex with three distinct targets. On the left, and right, pairwise superpositions of bound ligand conformations are shown, revealing conserved (left) and variable binding modes (right) in different protein environments.

Figure 1.2 Scaffold of a multifamily ligand with kinase activity representing an analogue series. For the ligand, crystal structures of complexes with Aurora and TEC kinases were available (PDB) and structural analogues found in ChEMBL were active against additional kinase targets from other families.



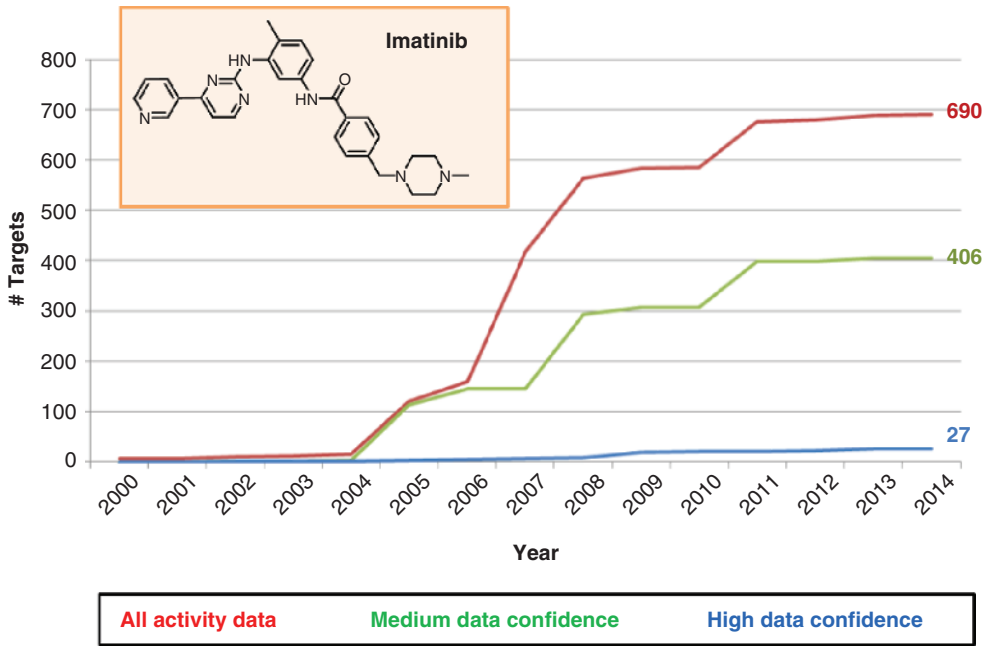


Figure 1.3 Target annotations of imatinib. Based on increasing volumes of activity data from ChEMBL collected over time, the number of targets reported for imatinib is monitored at three different confidence levels: all activity data (no confidence criteria were applied) medium- and high-confidence data. Adapted from Hu and Bajorath [36]. The number of target annotations based on all activity data and medium-confidence data (690 and 406, respectively) is unrealistic.

unexpected pharmacological effects. Systematic analysis of compound activity data helps to estimate the frequency with which MT-CPDs occur and the number of targets they are active against. Especially for candidate compounds and drugs, such estimates are relevant to balance often articulated expectation values that are largely unsubstantiated (e.g., "most drugs bind to 10 or 20 targets …"). In addition, careful analysis of available compounds and activity data also helps to gauge predictions of MT-CPDs and their target numbers, for example, from computational target profiling (*vide supra*). Notably, compound-data-driven analysis principally underestimates MT activity due to data incompleteness, given that not "all compounds have been tested against all targets" (the ultimate goal of chemogenomics). This must be taken into consideration. On the other hand, analysis of the large and rapidly growing volumes of activity data available in the public domain should reveal some statistically sound trends [36]. For instance, in 2019, we carried out a large-scale analysis of biological screening data from PubChem [37] in the search for compounds with activity against targets from different classes [38]. A total of 1063 compounds were identified that were tested in assays for at least 100 human target proteins and were active against at least 10 targets from more than one class [38]. These findings showed that MT-CPDs with activity against distantly or unrelated targets occurred rather frequently.

1.6 Drug Target Estimates

Systematic experimental determination of the targets that drugs are active against is far from being an easy task. Accordingly, insights into drug target numbers are typically confined to case-by-case proteomic analysis or statistics from target panel assays such as kinome screens [39]. However, based on compound data analysis, different estimates of target numbers for drugs and other active compounds have been reported.

Early attempts to predict drug targets used network representations of drug–target interactions [40]. From different databases, drugs, targets, and interaction data were collected and analyzed in drug–target networks. From such network representations, it was estimated that a drug on average interacted with six targets. Depending on the data used, targets per drug ranged from approximately 3 to 13 [40]. Comparable estimates were obtained when approved and experimental drugs taken from DrugBank [41] were mapped to ChEMBL and drug data and targets were monitored over a 15-year period [42]. For bioactive compounds from screening assays, different target numbers were determined. In an early analysis of PubChem [37], MT activities were analyzed on the basis 600+ assays [43]. It was found that approximately 58% of active screening compounds only displayed ST activity in combined primary and confirmatory assays. In addition, based on high-confidence

activity data extracted from ChEMBL, it was determined that a bioactive compound on average interacted with only one to two targets, with no significant variations across different families [44, 45]. Different from PubChem data, ChEMBL does not contain test frequencies for compounds taken from the literature (where negative data is typically not reported). Hence, data incompleteness generally plays a greater role in target analysis based on ChEMBL. On the other hand, high-confidence activity data can best be obtained from ChEMBL. In 2016, approximately 430,000 extensively assayed compounds were extracted from PubChem (tested in both primary and confirmatory assays), with mean and median values of 411 and 437 assays per compound, respectively [46]. Most extensively assayed screening hits were on average active against 2.5 targets [46], thus only slightly exceeding the mean number of targets determined for ChEMBL compounds based on high-confidence activity data (in the presence of data sparseness).

Global estimates of drug targets are generally affected by the presence of small subsets of highly promiscuous drugs. For example, during a period of strong growth of public compound activity data (2012–2016), the mean number of targets per drug increased from 5.9 (2012) to 7.3 (2016) for approximately 1600 drugs available in DrugBank. When the top-50 most promiscuous drugs were omitted, the mean (median) value decreased from 7.3 (5.0) to 6.4 (4.5) targets per drug; when the top-100 drugs were omitted, it further decreased to 6.0 (4.0) [36]. Our most recent in-house calculations for drugs with available high-confidence activity data indicate that the mean number of drug targets has essentially remained constant over the past five years, with approximately four targets per drug. Of note, computational predictions of drug targets tend to be higher than data-driven estimates.

1.7 Explainable Machine Learning

In addition to activity or target predictions, ML can also be used for hypothesis testing in polypharmacology research, quasi as a "diagnostic" approach. We have carried out a large-scale ML analysis to address the questions (i) if structural features exist that systematically distinguish MT-CPDs from compounds with corresponding ST activity (ST-CPDs) and (ii) if such structural features might be shared by compounds with activity against different targets. Therefore, a compound test system was generated by systematically searching ChEMBL for compounds with activity against different target pairs (dual-target compounds, DT-CPDs) and corresponding ST-CPDs that were available in sufficient numbers for ML. The search identified 170 target pairs (involving 137 unique targets) for which at least 100 MT- and (50 + 50) ST-CPDs with high-confidence activity annotations were available [47]. Many more target pairs with less than 100 DT-CPDs were detected. These findings also demonstrated that a wealth of DT-CPDs with activity against different targets is publicly available, from which analogue series and representative scaffolds can be isolated and considered as design templates for polypharmacology (*vide supra*).

For testing via ML, the following hypothesis was formulated: if structural features exist that systematically distinguish DT- and corresponding ST-CPDs, then ML models derived based on chemical structures should detect these features and accurately classify MT- and corresponding DT-CPDs; if such structural features do not exist, the models could not possibly be predictive [47]. Accordingly, for each target pair, ML models were derived using different algorithms and the same structural fingerprints. These models were then used to predict test compounds in "native" predictions (i.e., using test compounds of the target pair for which the models were derived) and for all other 169 target pairs in "cross-pair" predictions (i.e., using test compounds of all target pairs for which the models were not derived). Figure 1.4 summarizes the composition of the target-pair-based test system and the ML prediction strategy.

The 170 native predictions using different ML models yielded overall consistently accurate predictions with a median accuracy higher than 80% and only a few target pair datasets with an accuracy higher than 60% (statistical outliers of the distribution). By contrast, the vast majority of cross-pair predictions failed, with a median accuracy of approximately 50% (corresponding to random predictions) [47]. Taken together, these findings demonstrated

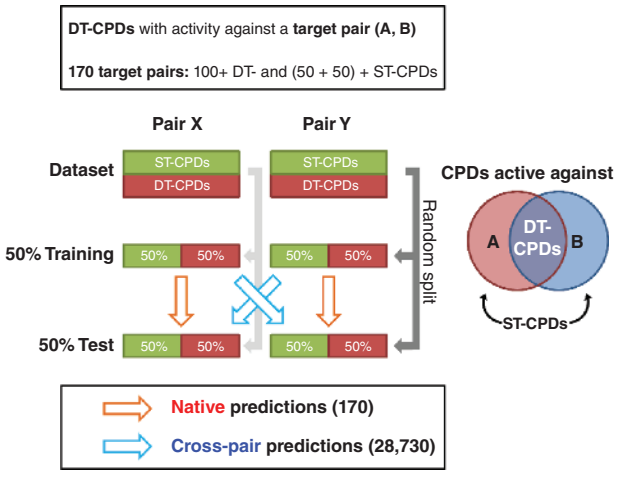


Figure 1.4 Target-pair-based compound test system for diagnostic ML investigating potential structural signatures of DT-CPDs. Adapted from Feldmann and Bajorath [47].

that structural features distinguishing between DT- and corresponding ST-CPDs existed and depended on target pairs. Hence, distinguishing features were not generalizable across target pairs.

Subsequently, structural features determining accurate predictions were identified for exemplary pairs of unrelated targets [48] using the SHapley Additive exPlanations (SHAP) approximation of the Shapley value concept from game theory [49] that was adapted as an explainable artificial intelligence (XAI) approach for rationalizing ML predictions [49, 50]. Features determining accurate predictions were characteristic of DT-CPDs and absent in ST-CPDs. In DT-CPDs, these structural features formed coherent and well-defined substructures [48]. These findings showed that explainable ML can be used to identify structural motifs implicated in MT activity.

1.8 Conclusion

In this introductory chapter, a brief account of the origins and history of the polypharmacology concept is provided and the Janus-headed nature of polypharmacology is emphasized. Despite potential pitfalls, polypharmacology has evolved into a viable therapeutic strategy for the treatment of multifactorial diseases that is here to stay and will further expand. MT activity of small molecules is at the root of polypharmacology and, in addition to its relevance for drug discovery, of considerable interest from a basic research perspective, as also discussed. Despite increasing insights into different facets of polypharmacology, a number of open questions remain to be addressed. For example, the magnitude of polypharmacology among current drugs still needs to be quantified, beyond expectations and case-by-case studies, which might require Herculean efforts (or novel experimental concepts). Furthermore, serendipity continues to play a major role in polypharmacology-oriented drug discovery, leaving much room for the development of further advanced drug design strategies. These and other questions related to polypharmacology provide ample opportunities for attractive future research.

References

1 Regenmortel, M.H.V. (2004). Reductionism and complexity in molecular biology: scientists now have the tools to unravel biological complexity and overcome the limitations of reductionism. *EMBO Rep.* 5: 1016–1020.

2 Morphy, R., Ka, C., and Rankovic, Z. (2004). From magic bullets to designed multiple ligands. *Drug Discov. Today* 9: 641–651.

3 Verlinde, C.L. and Hol, W.G. (1994). Structure-based drug design: progress, results and challenges. *Structure* 2: 577–587.

4 Anderson, A.C. (2003). The process of structure-based drug design. *Chem. Biol.* 10: 787–797.

5 Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* 2: 343–372.

6 Butcher, E.C., Berg, E.L., and Kunkel, E.J. (2004). Systems biology in drug discovery. *Nat. Biotechnol.* 22: 1253–1259.

7 Hopkins, A.L. (2007). Network pharmacology. *Nat. Biotechnol.* 25: 1110–1111.

8 Ainsworth, C. (2011). Networking for new drugs. *Nat. Med.* 17: 1166–1169.

9 Knight, Z.A., Lin, H., and Shokat, K.M. (2010). Targeting the cancer kinome through polypharmacology. *Nat. Rev. Cancer* 10: 130–137.

10 Tschöp, M.H., Finan, B., Clemmensen, C. et al. (2016). Unimolecular polypharmacy for treatment of diabetes and obesity. *Cell Metab.* 24: 51–62.

11 Benek, O., Korabecny, J., and Soukup, O. (2020). A perspective on multi-target drugs for Alzheimer's disease. *Trends Pharmacol. Sci.* 41: 434–445.

12 Weston-Green, K. (2022). Antipsychotic drug development: from historical evidence to fresh perspectives. *Front. Psychiatry* 13: 903156.

13 Löscher, W. (2021). Single-target versus multi-target drugs versus combinations of drugs with multiple targets: preclinical and clinical evidence for the treatment or prevention of epilepsy. *Front. Pharmacol.* 12: 730257.

14 Hopkins, A.L., Mason, J.S., and Overington, J.P. (2006). Can we rationally design promiscuous drugs? *Curr. Opin. Struct. Biol.* 16: 127–136.

15 Peters, J.U. (2013). Polypharmacology – foe or friend? *J. Med. Chem.* 56: 8955–8971.

16 Anighoro, A., Bajorath, J., and Rastelli, G. (2014). Polypharmacology: challenges and opportunities in drug discovery. *J. Med. Chem.* 57: 7874–7887.

17 Bolognesi, M.L. and Cavalli, A. (2016). Multitarget drug discovery and polypharmacology. *ChemMedChem* 11: 1190–1192.

18 Campillos, M., Kuhn, M., Gavin, A.C. et al. (2008). Drug target identification using side-effect similarity. *Science* 321: 263–266.

19 Berger, S.I. and Iyengar, R. (2011). Role of systems pharmacology in understanding drug adverse events. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 3: 129–135.

20 Pushpakom, S., Iorio, F., Eyers, P.A. et al. (2019). Drug repurposing: progress, challenges and recommendations. *Nat. Rev. Drug Discov.* 18: 41–58.

21 Cohen, P., Cross, D., and Jänne, P.A. (2021). Kinase drug discovery 20 years after imatinib: progress and future directions. *Nat. Rev. Drug Discov.* 20: 551–569.

22 Ryszkiewicz, P., Malinowska, B., and Schlicker, E. (2023). Polypharmacology: promises and new drugs in 2022. *Pharmacol. Rep.* 75: 755–770.

23 Proschak, E., Stark, H., and Merk, D. (2018). Polypharmacology by design: a medicinal chemist's perspective on multitargeting compounds. *J. Med. Chem.* 62: 420–444.

24 Li, X., Li, X., Liu, F. et al. (2021). Rational multitargeted drug design strategy from the perspective of a medicinal chemist. *J. Med. Chem.* 64: 10581–10605.

25 Morphy, R. and Rankovic, Z. (2005). Designed multiple ligands. An emerging drug discovery paradigm. *J. Med. Chem.* 48: 6523–6543.

26 Moser, D., Wisniewska, J.M., Hahn, S. et al. (2012). Dual-target virtual screening by pharmacophore elucidation and molecular shape filtering. *ACS Med. Chem. Lett.* 3: 155–158.

27 Rastelli, G. and Pinzi, L. (2015). Computational polypharmacology comes of age. *Front. Pharmacol.* 6: 157.

28 Feldmann, C. and Bajorath, J. (2022). Advances in computational polypharmacology. *Mol. Inform.* 41: 2200190.

29 Berman, H.M., Westbrook, J., Feng, Z. et al. (2000). The protein data bank. *Nucleic Acids Res.* 28: 235–242.

30 Gilberg, E., Stumpfe, D., and Bajorath, J. (2018). X-ray-structure-based identification of compounds with activity against targets from different families and generation of templates for multitarget ligand design. *ACS Omega* 3: 106–111.

31 Gilberg, E., Gütschow, M., and Bajorath, J. (2019). Promiscuous ligands from experimental structures, binding conformations, and protein family dependent interaction hotspots. *ACS Omega* 4: 1729–1737.

32 Gaulton, A., Hersey, A., Nowotka, M. et al. (2017). The ChEMBL database in 2017. *Nucleic Acids Res.* 45: D945–D954.

33 Baell, J.B. and Holloway, G.A. (2010). New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* 53: 2719–2740.

34 Bruns, R.F. and Watson, I.A. (2012). Rules for identifying potentially reactive or promiscuous compounds. *J. Med. Chem.* 55: 9763–9772.

35 Irwin, J.J., Duan, D., Torosyan, H. et al. (2015). An aggregation advisor for ligand discovery. *J. Med. Chem.* 58: 7076–7087.

36 Hu, Y. and Bajorath, J. (2017). Entering the 'big data' era in medicinal chemistry: molecular promiscuity analysis revisited. *Future Sci. OA* 3: FSO179.

37 Wang, Y., Bryant, S.H., Cheng, T. et al. (2017). PubChem BioAssay: 2017 update. *Nucleic Acids Res.* 45: D955–D963.

38 Feldmann, C., Miljković, F., Yonchev, D., and Bajorath, J. (2019). Identifying promiscuous compounds with activity against different target classes. *Molecules* 24: 4185.

39 Wang, Y. and Ma, H. (2015). Protein kinase profiling assays: a technology review. *Drug Discov. Today Technol.* 18: 1–8.

40 Mestres, J., Gregori-Puigjané, E., Valverde, S., and Solé, R.V. (2009). The topology of drug–target interaction networks: implicit dependence on drug properties and target families. *Mol. BioSyst.* 5: 1051–1057.

41 Law, V., Knox, C., Djoumbou, Y. et al. (2014). DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res.* 42: D1091–D1097.

42 Hu, Y. and Bajorath, J. (2014). Monitoring drug promiscuity over time. *F1000Research* 3: 218.

43 Han, L., Wang, Y., and Bryant, S.H. (2009). A survey of across-target bioactivity results of small molecules in PubChem. *Bioinformatics* 25: 2251–2255.

44 Hu, Y. and Bajorath, J. (2013). High-resolution view of compound promiscuity. *F1000Research* 2: 144.

45 Hu, Y., Jasial, S., and Bajorath, J. (2015). Promiscuity progression of bioactive compounds over time. *F1000Research* 4: 118.

46 Jasial, S., Hu, Y., and Bajorath, J. (2016). Determining the degree of promiscuity of extensively assayed compounds. *PLoS ONE* 11: e0153873.

47 Feldmann, C. and Bajorath, J. (2021). Machine learning reveals that structural features distinguishing promiscuous and non-promiscuous compounds depend on target combinations. *Sci. Rep.* 11: 7863.

48 Feldmann, C., Philipps, M., and Bajorath, J. (2021). Explainable machine learning predictions of dual-target compounds reveal characteristic structural features. *Sci. Rep.* 11: 21594.

49 Lundberg, S.M. and Lee, S. (2017). A unified approach to interpreting model predictions. *Adv. Neural Inf. Proces. Syst.* 30: 4766–4775.

50 Chen, H., Covert, I.C., Lundberg, S.M., and Lee, S.I. (2023). Algorithms to estimate Shapley value feature attributions. *Nat. Mach. Intell.* 5: 590–601.