*Teaser Two-thirds of drugs have more-potent **in vitro** affinities for their primary targets than those of the corresponding endogenous ligands, which define a baseline to estimate off-target safety margins.*

The human endogenous metabolome as a pharmacology baseline for drug discovery

Andreu Bofill[1,‡], Xavier Jalencas[1,‡], Tudor I. Oprea[2,3,4,5] and Jordi Mestres[1]

[1] Research Group on Systems Pharmacology, Research Program on Biomedical Informatics (GRIB), IMIM Hospital del Mar Medical Research Institute and University Pompeu Fabra, 08003 Barcelona, Catalonia, Spain
[2] Department of Internal Medicine, University of New Mexico School of Medicine, Albuquerque, NM, USA
[3] UNM Comprehensive Cancer Center, Albuquerque, NM, USA
[4] Department of Rheumatology and Inflammation Research, Institute of Medicine, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden
[5] Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

We have limited understanding of the variation in *in vitro* affinities of drugs for their targets. An analysis of a highly curated set of 815 interactions between 566 drugs and 129 primary targets reveals that 71% of drug–target affinities have values above that of the corresponding endogenous ligand, 96% of them fitting within a range of two orders of magnitude. Our findings suggest that the evolutionary optimised affinity of endogenous ligands for their native proteins can serve as a baseline for the primary pharmacology of drugs. We show that the degree of off-target selectivity and safety risks of drugs derived from their secondary pharmacology depend very much on that baseline. Thus, we propose a new approach for estimating safety margins.

Introduction

Drugs exert their therapeutic action by interacting with one (or more) disease-relevant protein(s). The final potency of this interaction is the result of a long optimisation process, in which multiple pharmacodynamic (PD) and pharmacokinetic (PK) parameters need to be taken into consideration within the context of a complex neighbourhood of activities [1,2]. In this respect, a common strategy followed during lead optimisation is to maximise *in vitro* binding affinity for the primary (mechanism of action) target(s) of the drug as a means to improve target selectivity and reduce secondary (off-target) affinities potentially linked to safety issues [3]. Compound potency is generally expected to decrease in an *in vivo* environment because of several factors, mostly associated with the compound absorption, distribution, and metabolism [4,5]. Accordingly, the finally selected drug candidate is unlikely to be the molecule having the strongest affinity value but the one showing an optimal balance between all properties [6].

Corresponding author:
‡ These authors contributed equally to this work.

Andreu Bofill is a PhD candidate in the Research Group on Systems Pharmacology under the supervision of Jordi Mestres. He studied human biology and later obtained a MSc in bioinformatics for health science at University Pompeu Fabra (Barcelona, Spain). His principal research interests include drug interactions, drug-induced adverse event detection, and personalised medicine.

Xavier Jalencas is a research scientist at Chemotargets. He has a PhD in biomedicine from the University Pompeu Fabra (Barcelona, Spain) for his research on structure-based approaches to polypharmacology, performed in the Research Group on Systems Pharmacology under the supervision of Jordi Mestres. After a postdoctoral period in the same group, he joined Chemotargets in 2019. His current research interests include the development of novel fragment-based approaches to structure-based virtual screening and hit identification.

Tudor Oprea is professor of medicine and pharmaceutical sciences, and head of the Translational Informatics Division, in the Department of Internal Medicine, University of New Mexico School of Medicine, Albuquerque (NM, USA). He is also guest professor at the Institute of Medicine, Gothenburg University (Sweden) and at the Center for Protein Research, University of Copenhagen (Denmark). To date, Dr Oprea has co-authored over 200 publications and book chapters, and eight US patents. Since 2014, Dr Oprea has been a principal investigator for the Illuminating the Druggable Genome Knowledge Management Center, a NIH Common Fund initiative. Dr Oprea's work has led to drug-repurposing clinical trials at UNM in two types of cancer, most notably R-ketorolac, which shows promise in an open-label clinical trial for ovarian cancer. His most current research is on the development of validated machine-learning and artificial intelligence models for target and drug discovery, by combining numerical and free-text information to model human health.

Jordi Mestres holds a PhD in computational chemistry from the University of Girona (Spain). After a postdoctoral stay at Pharmacia&Upjohn in Kalamazoo (Michigan, USA), he joined the Molecular Design & Informatics Department at N.V. Organon in Oss (The Netherlands) in 1997 and, in 2000, he was appointed head of computational medicinal chemistry at Organon Laboratories in Newhouse (UK). In 2003, he took on his current position as head of the Research Group on Systems Pharmacology, within the Research Program on Biomedical Informatics (GRIB) at the IMIM Hospital del Mar Medical Research Institute in Barcelona. He is also associate professor at the University Pompeu Fabra (UPF). In 2006, he founded Chemotargets as a spin-off company of his research group. He is also the recipient of the 2006 Corwin Hansch Award from the QSAR and Modelling Society and the 2007 Technology Transfer Award from the UPF. In 2018, he was admitted as a Fellow of the Royal Society of Chemistry. He is the author of over 150 publications, including 10 patents.

https://doi.org/10.1016/j.drudis.2019.06.007

This might explain, in part, the substantial variation of binding affinities observed for drugs targeting not only the same protein but also across different proteins [7]. For example, the serotonin receptor subtype 1A (5HT1A) is one of the protein targets involved in the mechanisms of action of brexpiprazole and vortioxetine [7]. However, whereas brexpiprazole binds to this receptor with subnanomolar affinity ($pK_i = 9.9$), the corresponding *in vitro* binding affinity for vortioxetine is more than two orders of magnitude lower ($pK_i = 7.8$). The variation in physicochemical properties (as estimated by the octanol-water partition coefficient and aqueous solubility), which is less than a log unit, does not offer a valid explanation for this difference. This is in stark contrast to the micromolar affinity ($pK_i = 5.6$) of theophylline for the adenosine receptor 2b (AA2BR), which is one of its mechanism of action targets [7]. Gaining a deeper understanding as to why drugs might require achieving certain levels of affinity for their primary target (s) and why these levels of affinity could be considerably different because of efficacy and receptor reserve across primary target(s) is at the core of modern preclinical drug discovery.

In humans, most drug targets are proteins the function of which is regulated by endogenous ligands or metabolites, understood here not as the products of drug metabolism [8] but as those naturally occurring small molecules of human metabolism [9–11]. The human endogenous metabolome is estimated to contain a few thousand chemical species [12]. Each endogenous ligand binds to its native protein with a certain affinity that has been sensitively optimised by evolution and that might subtlety vary across individuals [13]. For example, the subnanomolar affinity ($pK_i = 9.1$–9.7) of serotonin for 5HT1A contrasts with the low micromolar affinity ($pK_i = 4.82$) of adenosine for AA2BR [14]. Interestingly, the natural affinities between these two endogenous ligands and their respective native proteins compare well with some of the designed affinities for the drugs having those native proteins as primary targets (see earlier). These observations prompted us to investigate whether there is a general trend across drug targets relating the pharmacology of drugs and human endogenous ligands. In addition, we examined also whether the same holds true for the catalytic activities of human endogenous substrates and drugs for their respective enzyme targets.

The results obtained were consistent with most drugs having *in vitro* affinities for their primary target(s) above those of the corresponding metabolite and/or substrate–target interaction. Here, we also discuss the implications for secondary pharmacology and off-target safety margins derived from it [15]. In particular, we analyse in detail the case of drug-induced valvular heart disease when the small molecule drug is an agonist of the serotonin 5HT2B receptor.

Primary pharmacology of drugs and endogenous ligands

In vitro binding affinities for drug–target interactions linked to the mechanism of action of the drug were retrieved from the DrugCentral repository [7]. The list of unique human primary targets involved in those interactions was then used to interrogate the Guide to Pharmacology database (GtoPdb) to identify the native endogenous ligand (metabolite) of each one of them and extract the maximum affinity value reported [14]. For the sake of comparison, drug–target–metabolite triad associations were only accepted if they had been measured using the same units for the binding affinity (pK_i or pK_d) or, in its absence, for the potency (pIC_{50} or pEC_{50}) (see Box 1 for further methodology details). Although measurements of IC_{50} and EC_{50} values are assay specific, it has been shown that independent IC_{50} data show similar dispersions to K_i data in ChEMBL [16]. After careful curation, a total of 442 drug–target–metabolite triads were compiled, involving 293 drugs, 79 targets, and 43 endogenous ligands (Table S1 in the supplemental information online). Most triads (90%) involved K_i binding affinities.

The 442 pairs of drug and metabolite affinities for the same protein are plotted in Fig. 1. On average, drug affinities were found to be over 20-fold higher than the corresponding metabolite affinities, with median negative logarithm affinities of 8.52 and 7.20, respectively. If there was a perfect correlation between drug and metabolite affinities, data points would follow the diagonal solid line. If the affinities of the endogenous metabolites were stronger than the corresponding drug affinities, data points would be expected to gather below the diagonal line, in the lower triangle region. However, the density plot derived from the data distribution reveals a clear accumulation of points above the diagonal line, in the upper triangle region. Indeed, 67% of the 442 drug affinities for their primary target(s) were higher than the corresponding metabolite–target interactions (solid-grey line in Fig. 1). This percentage increased to 85% and 96% when considering drug affinities one order and two orders of magnitude (dashed-grey line in Fig. 1) below metabolite affinities, respectively.

Only 4% of drug affinities for primary target(s) were found to be over 100-fold lower than the corresponding metabolite affinities (Table S2 in the supplemental information online). A close examination of these 19 drug–target interactions suggests that the assignment of the interacting protein as primary target of the drug would benefit from a careful revision. For example, oxymorphone is a semisynthetic opioid analgesic that has all three opioid receptors (μ, δ, and κ) assigned as mechanism of action targets with binding affinities (pK_i) of 9.44, 7.30, and 6.83, respectively [7]. However, the affinity of oxymorphone for the κ opioid receptor is almost four orders of magnitude lower than the affinity for dynorphin A ($pK_i = 10.80$), its native endogenous metabolite. This suggests that the analgesic effect of oxymorphone is unlikely to be conducted through the activation of the κ opioid receptor, but mainly through its action on the μ opioid receptor, as some literature suggests [17–19]. Another drug found to be an outlier is tramadol, believed to exert its analgesic function via the μ opioid receptor. Yet, its binding affinity ($pK_i = 5.8$) is almost over three orders of magnitude lower than the affinity of the corresponding endogenous ligand, dynorphin B ($pK_i = 8.5$), suggesting that its therapeutic action is unlikely to be the result solely of its interaction with the μ opioid receptor. Indeed, tramadol is not a singular opioid drug, but an analgesic with several contributing components coming from its rich polypharmacology for the sodium-dependent serotonin, noradrenaline, and dopamine transporters, among others [7]. Interestingly, some studies have suggested that the tramadol-induced analgesic effect is produced, at least in part, by one of its metabolites, which binds with higher affinity to the μ opioid receptor ($pK_i = 6.82$) well within two orders of magnitude from that of the native endogenous metabolite [20]. Along the same lines, the assignment of the prostaglandin F2-alpha (PGF2a) receptor as the mechanism of action target of

BOX 1

Data and methods

in vitro **pharmacology of drugs and primary targets**

We explored public pharmacology databases in the search for interactions of drugs with their primary targets where quantitative affinity data were available for both the interactions between the drug and the target and between the target and its main endogenous metabolite. Accordingly, drug–target interactions labelled as being involved in the mechanism of action of the drug with defined activity values were retrieved from DrugCentral [7]. Of the 4486 drugs available in DrugCentral (downloaded on May 2017), 1862 had defined affinity values against a given target, 936 of them having at least one interaction against a human target labelled as being involved in the mechanism of action of the drug (i.e., its primary targets). A total of 1769 interactions were retrieved between those 936 drugs and 403 targets. In a second stage, two additional sources of affinity data were searched for quantitative affinities of drug–target interactions, namely Guide to Pharmacology database (GtoPdb) [14] and ChEMBL [28]. GtoPdb contributed with 76 additional interactions between 61 drugs and 40 proteins, whereas ChEMBL added 71 interactions between 47 drugs and 28 proteins.

in vitro **pharmacology of endogenous ligands and primary targets**

To include data on endogenous ligands, we searched GtoPdb [14] first for those 403 drug targets identified in the previous step, retrieving all their interactions with ligands labelled as being the principal human endogenous ligand. In total, 179 interactions between 86 metabolites and 95 proteins were extracted. GtoPdb collects the highest and lowest affinity values reported for each metabolite–protein interaction, which reflects the inherent variability of independent affinity measures, and provides the detail of the units used to measure binding. On average, pK_i affinity ranges extracted from GtoPdb spanned 1.0 log units, with a standard deviation of 0.59 log units. However, there were some extreme cases, such as the affinity of dopamine for the dopamine D2 receptor, for which pK_i values vary largely (4.7–7.2). To partially alleviate the effects of data variability, the analysis was done using always the highest pK_i affinity value reported in GtoPdb for each endogenous ligand. By taking the least favourable scenario when a range of affinities was provided, we aimed to increase the robustness of our analysis. Then, for each protein, the main endogenous metabolite with highest affinity was selected as the native metabolite. In a second stage, the ChEMBL database [28] was also searched for additional metabolite affinities on any of those 403 drug targets. An additional set of 121 interactions between 22 metabolites and 34 proteins were extracted.

in vitro **pharmacology mapping of drugs and endogenous ligands**

To allow for affinity comparisons, drug–receptor–metabolite associations were made only when both drug and metabolite had a described affinity for the target with the same affinity unit (K_i, Kd, IC_{50}, or EC_{50}), using the highest metabolite affinity available. We identified a total of 442 drug–receptor–metabolite triads involving 293 drugs, 79 proteins, and 43 metabolites (Table S1 in the supplemental information online). Among them, 404 (90.4%) were pK_i values. In a second stage, we collected an additional set of 202 drug–receptor–metabolite triads involving 145 drugs, 55 targets, and 34 metabolites (Table S3 in the supplemental information online). Of them, 116 (57.4%) were pK_i values.

in vitro **binding affinities of 5HT2B agonists and risk of VHD**

A total of 153 drugs with affinity data for 5HT2B were identified in DrugCentral [7]. The PDSP K_i database [48] was then searched to complement with Ki data those interactions for which activities in other units (K_d, IC_{50} or EC_{50}) were available in DrugCentral. Of these, two drugs were annotated as 5HT2B agonists in DrugCentral [7]. An additional eight drugs were found annotated as 5HT2B agonists in ChEMBL [28]. A manual literature search of the remaining drugs confirm 11 other drugs as 5HT2B agonists and provided evidence supporting or rejecting their associated risk of causing VHD. A final set of 21 drugs with pK_i affinity values, confirmed agonist action to the 5HT2B receptor, and evidence of VHD risk was collected [37–45]. Among them, 12 were recognised valvulopathic drugs. An additional set of three 5HT2B agonist drugs for which pK_i values were available but no information on VHD could be found was also included. The final list of 24 5HT2B agonist drugs considered in this study is provided in Table S5 in the supplemental information online.

bimatoprost, a prostaglandin analog used in the treatment of glaucoma and ocular hypertension, is open to debate. The affinity of bimatoprost for that receptor (pK_i = 5.30) is over three orders magnitude lower than the affinity reported for the PGF2a endogenous ligand (pK_i = 8.5), suggesting that the PGF2a receptor is not the primary target responsible for the therapeutic effect of bimatoprost. Along these lines, some studies have reported no meaningful activity of bimatoprost for the prostaglandin receptors and proposed a novel prostamide-sensitive receptor as the probable primary and functional target for bimatoprost [21–23]. Similarly, the primary targets assigned to methoxamine are the α_{1A}, α_{1B}, and α_{1D} adrenergic receptors with affinities of 5.1, 4.0, and 4.9, respectively [7]. Yet, the binding affinity of this drug for α_{1B} is two and a half orders of magnitude lower than the affinity of (-)-noradrenaline (pK_i = 6.5), the native endogenous ligand, suggesting that the α_{1B} adrenoceptor is unlikely to have a major role in the mechanism of action of methoxamine [24].

Analysis of the functional action of the drugs on each target allows for deconvoluting the 442 drug–target interactions depicted in Fig.1 into two sets of 160 and 260 interactions that involve drugs acting as agonists and antagonists on their targets, respectively (Fig. 2). The 160 drug agonist interactions (Fig. 2a) involve 109 drugs, 63 targets, and 37 endogenous ligands, whereas 181 drugs, 42 targets, and 22 endogenous ligands define the 260 interactions from drug antagonists (Fig. 2b). The remaining 22 interactions correspond to drugs being annotated as partial agonists (14) or inverse agonists (eight). Interestingly, there are six cases in which the same drug acts as an agonist and an antagonist on different targets. This is the case, for example, of agomelatine, a high-affinity agonist of melatonin MT_1 and MT_2 receptors but an antagonist on the serotonin 5HT2C receptor [25], and of flibanserin, a serotonin 5HT1A agonist and 5HT2A antagonist [26].

The shape of the density plot derived from the 160 interactions involving drug agonists (Fig. 2a) follows the diagonal line. Compared with Fig. 1, not even half (45%) of the 160 drug agonist affinities for their primary target(s) are higher than the corresponding metabolite–target interactions (solid-grey line in Fig. 2a), although this percentage increases up to 91% if a drug–metabolite affinity window of minus two orders of magnitude is considered (dashed-grey line in Fig. 2a). This is a reflection of the fact that drug agonists and endogenous ligands tend to have rather similar affinities for their common protein targets across the full

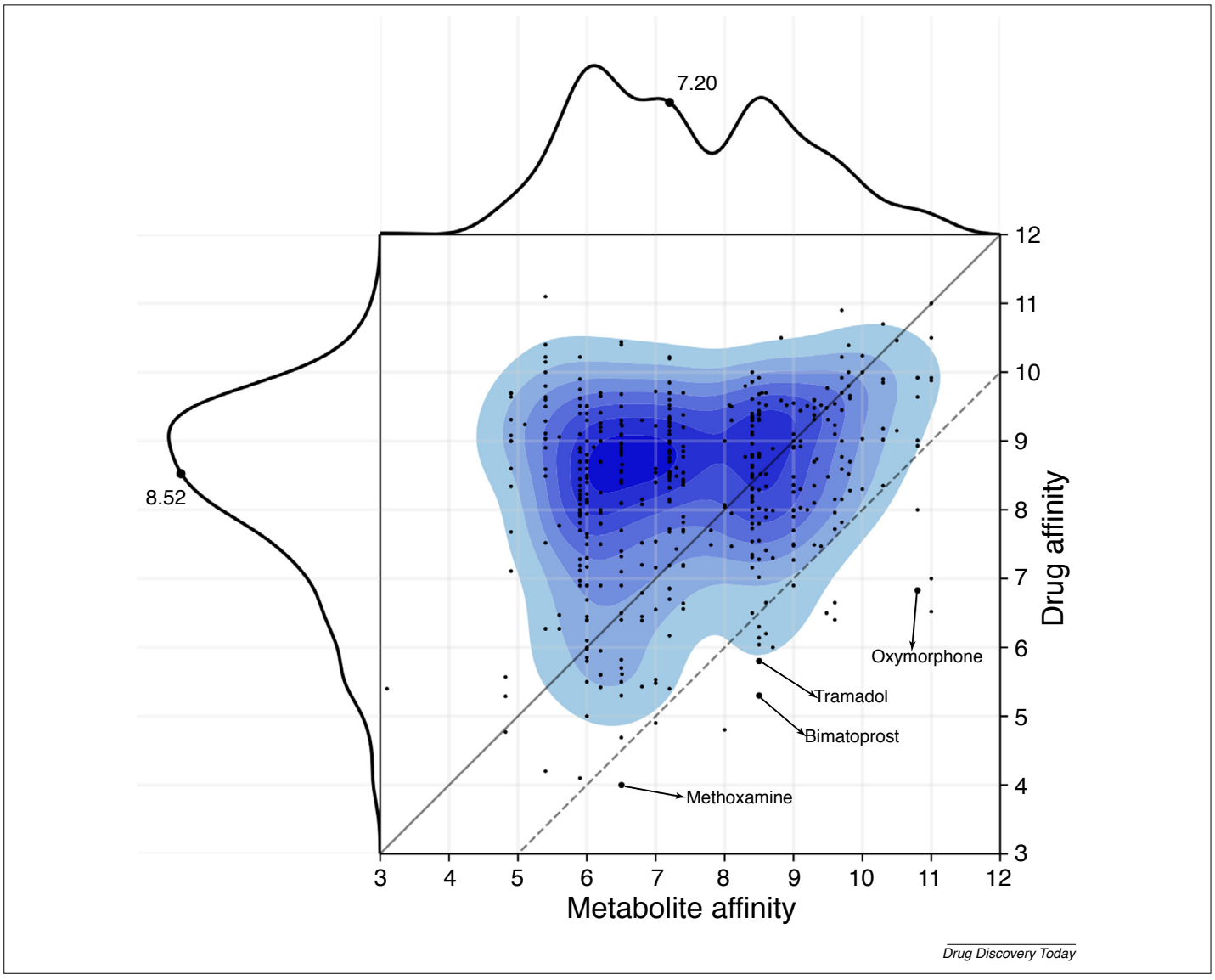

FIGURE 1

Density plot of the 442 drug–target–metabolite triads. Drug affinities for their primary target(s) are plotted against the corresponding endogenous metabolite affinities. A kernel density estimation is shaded in blue tones, darkest blue corresponding to highest density regions. Solid and dashed lines correspond to drug affinities being equal to and two orders of magnitude lower than the corresponding metabolite affinities for the same protein, respectively. Also included are the distributions of drug affinities (left) and metabolite affinities (top) with median negative logarithm values of 8.52 and 7.20, respectively. The position and name of the four outlier drugs discussed in the main text are also indicated.

range of affinity values. By contrast, the 260 interactions from drug antagonists appear concentrated well above the diagonal line, with a clear shift towards targets with low-affinity endogenous ligands compared with drug agonists. In this case, over three-quarters (79%) of the 260 drug antagonist affinities for their mechanism of action target(s) already have higher affinity values than the corresponding metabolite–target interactions (solid-grey line in Fig. 2b), with 98% of them fitting within a range of two orders of magnitude below the endogenous ligand affinities (dashed-grey line in Fig. 2b). In fact, almost two-thirds of the antagonist drug–target interactions (61.5%) have affinity values one order of magnitude above the metabolite baseline compared with only one-sixth of the agonist drug–target interactions (16.5%). Based on data currently available, this suggests that drug antagonists tend to require higher affinities than the endogenous ligands binding to the same target. One possible explanation for this clear difference could be that agonist drugs do not need to have much stronger affinities than endogenous ligands because they essentially seek to mimic their behaviour to persistently activate the receptor. By contrast, antagonist drugs do need higher affinities to completely block the action of endogenous ligands when released.

Comparing primary pharmacology across drug targets

A target-centred analysis of the set of 442 interactions provides another perspective of how drug affinities compare with endogenous ligand affinities across 79 proteins, 70 G-protein-coupled receptors (GPCRs), and nine nuclear hormone receptors. The height of each column in the circular plot of Fig. 3 reflects the affinity between a drug target and its main endogenous ligand. As

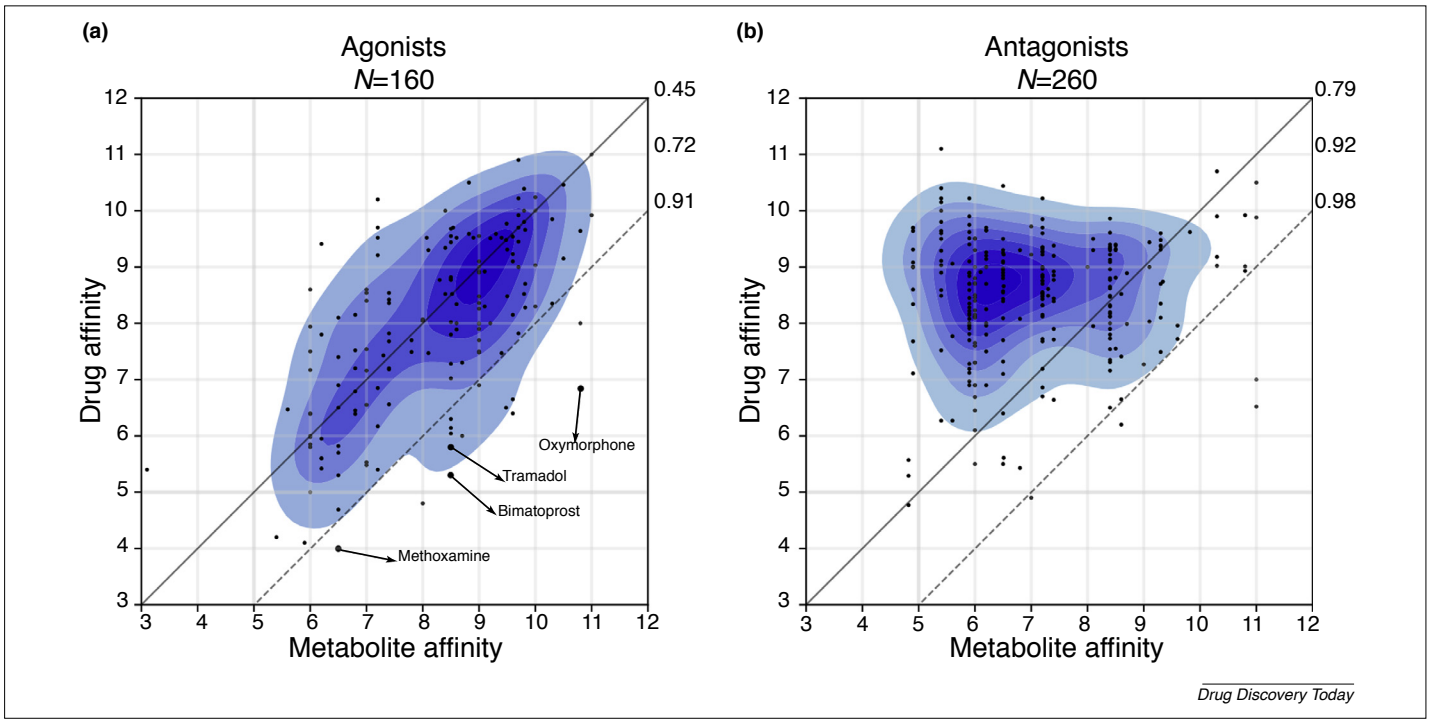

FIGURE 2

Density plots of the 160/260 drug agonist and/or antagonist–target–metabolite triads. Affinities values of drug agonists (a) and antagonists (b) for their primary target(s) are plotted against the corresponding endogenous metabolite affinities. A kernel density estimation is shaded in blue tones, darkest blue corresponding to highest density regions. Values on the top-right side of each graph reflect the percentage of drug–target interactions having affinity values equal to (solid line), one order, and two orders (dashed line) of magnitude more potent than the corresponding metabolite affinities for the same protein.

can be observed, there are some clear differences between the endogenous metabolite affinities across protein families. Even within a particular family, subtle variations exist. For example, the endogenous ligands of the serotonin, opioid, somatostatin, vasopressin, and melatonin receptor families show *in vitro* binding affinities in the nanomolar range. By contrast, micromolar affinities appear to be sufficient for the endogenous ligands of the acetylcholine (muscarinic), dopamine, and adenosine receptor families. Thus, it appears clear that different proteins have evolved to interact with their native metabolites at different levels of affinity, which might in turn translate into different lower-bound affinity criteria for any potential drug targeting them.

As illustrated in Fig. 3, the endogenous ligand affinity for each protein sets a baseline above which drug affinities might spread. In principle, the lower the affinity baseline, the wider the affinity window allowed for drugs to optimise other PD and PK parameters. This is well reflected by the drug affinity distributions for targets having different endogenous ligand affinities. For example, serotonin and 17β-estradiol both have *in vitro* binding affinities well into the subnanomolar range for 5HT1A (pK_i = 9.7) and the estrogen receptor alpha (ESR1; pK_i = 9.8), respectively. Among the set of 442 interactions, there are six drugs that have 5HT1A as a primary target and their pK_i values vary from 7.8 (right within the 100-fold difference from the serotonin affinity for 5HT1A) to 9.9. There are also six drugs having ESR1 assigned as a mechanism-of-action target and, most interestingly, their affinities cover a similar range of pK_i values, from 7.5 (close to the 100-fold difference from the 17β-estradiol affinity for ESR1) to 9.7. By contrast, the affinity of serotonin for the serotonin receptor 2A (5HT2A) is approximately one order of magnitude lower than that for 5HT1A (pK_i = 8.4). This is translated in a distribution of pK_i values for the 28 drugs having 5HT2A as primary target ranging from 6.5 (within two orders of magnitude from the serotonin affinity for 5HT2A) to 9.9. Comparably, the case of dopamine for the dopamine D2 receptor (DRD2) sets a relatively lower metabolite affinity baseline for this protein (pK_i = 7.2). There is a total of 46 drugs having DRD2 as a primary target, covering a wide affinity window of pK_i values from 5.4 (within the 100-fold difference from the dopamine affinity for DRD2) to 10.2. Finally, (-)-noradrenaline has micromolar affinity (pK_i = 6.0) for the β1 adrenergic receptor (ADRB1). The pK_i values of the 25 drugs that were found to have ADRB1 as primary target range from 5.0 (just tenfold below the (-)-noradrenaline affinity for ADRB1) to 9.5.

Figure 4 illustrates these trends for the set of 26 targets for which pK_i data are available for more than three drugs (Table S1 in the supplemental information online). It is shown that there is a directly proportional relationship (Pearson correlation coefficient = 0.71; P = 4.52E-5; r^2 = 0.51) between the affinity of an endogenous ligand and the minimum affinity of a drug for its primary target (Fig. 4a). This suggests that the *in vitro* binding affinity of the endogenous ligand for its native protein is a good reference baseline for the objective primary pharmacology of drugs in preclinical research. It is also shown that there is an inversely proportional relationship (Pearson correlation coefficient = –0.64; P = 3.8E–4; r^2 = 0.41) between endogenous ligand affinities and the range of drug affinities for a given target (Fig. 4b). This reflects the fact that, for every target, the upper-bound drug affinities reach always subnanomolar potencies, whereas the

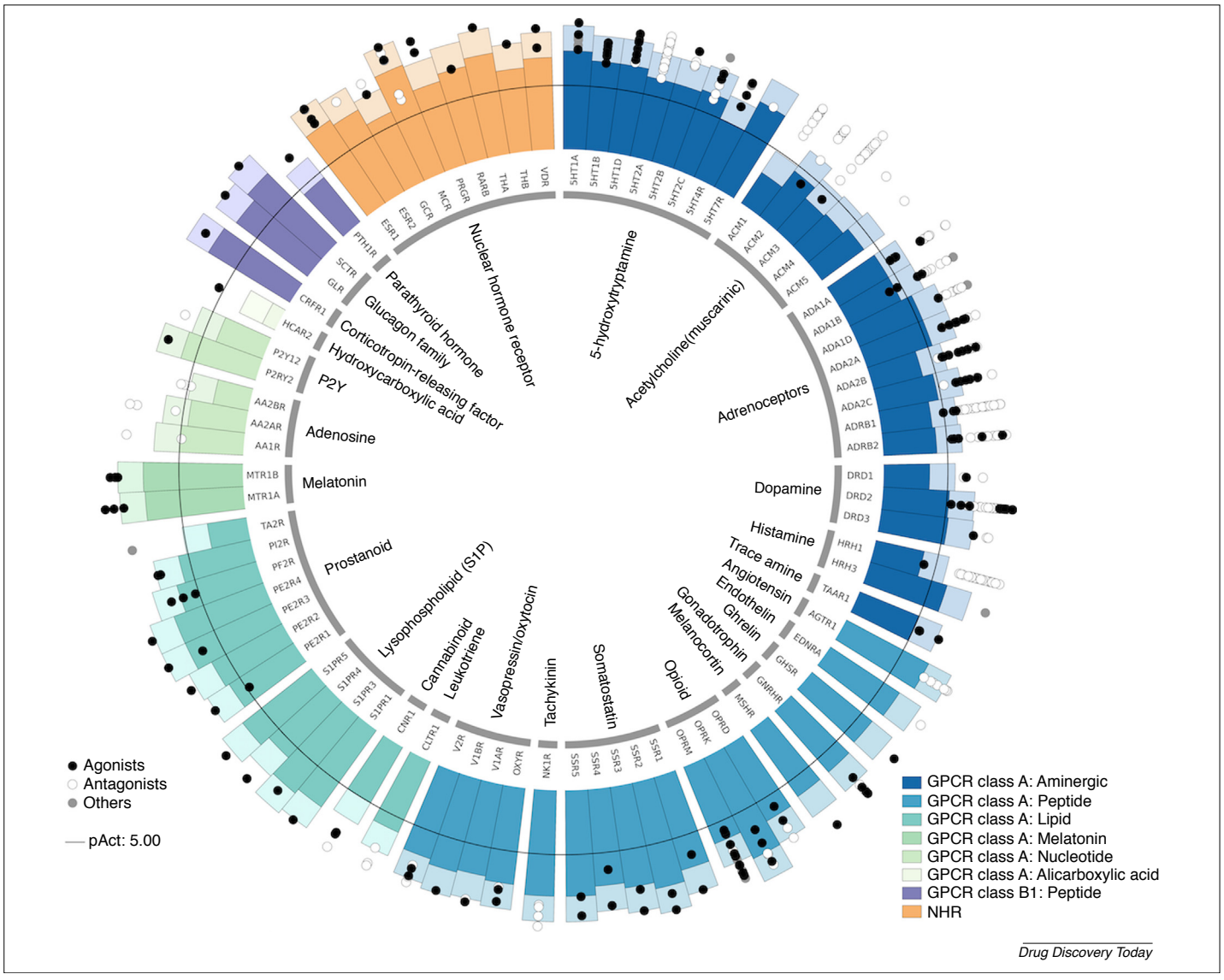

FIGURE 3

Target-centred analysis of the 442 drug–receptor–metabolite triads. Each one of the 79 drug targets, 70 G-protein-coupled receptors and 9 nuclear hormone receptors, is represented by a column in this circular plot. The height of the column reflects the affinity between a drug target and its main endogenous ligand. The tip of each column, in lighter colour, marks the two order of magnitude window below the affinity of the endogenous metabolite. Drug affinities for their primary target(s) are displayed as circles, filled or open depending on its functional action. The solid-black line crossing all columns serves as the 10-μM reference affinity level for all drug and endogenous metabolite affinities.

acceptable lower-bound drug affinities decrease relative to the endogenous ligand affinity for the drug target.

Nonetheless, some outliers from these trends are detected, namely, the muscarinic acetylcholine receptor M1 (ACM1), the κ opioid receptor (OPRK), the μ opioid receptor (OPRM) and the α_{1B} adrenergic receptor (ADA1B). Given the affinity of acetylcholine for ACM1 (pK_i = 4.9), one would expect that, among the 12 drugs identified as having ACM1 as primary target, the minimum drug affinity would be close to the 10-μm level. However, the minimum affinity for ACM1 corresponds currently to diphenidol (pK_i = 7.1). The existence of a highly conserved acetylcholine binding site among the five muscarinic receptor subtypes could explain higher than expected affinities in the search for selective ACM1 drugs. Those differences might also reflect the lack of completeness of pharmacological data [27]. Despite the fact they are not represented in the data we analysed, drugs having ACM1 as primary target and with pK_i affinities around 4 might exist. By contrast, the case of OPRK is completely different. If one disregards oxymorphone (discussed earlier), the drug with minimum affinity is nalbuphine (pK_i = 8.0), which places OPRK right where one would expect. Similarly, discarding the discussed cases of tramadol and methoxamine for OPRM and ADA1B, respectively, also pushes those targets up into, or much closer to, the grey zone that one would expect. Accordingly, removing ACM1 from the set and placing OPRK, OPRM, and ADA1B to the corresponding position after discarding the oxymorphone, tramadol, and methoxamine results in stronger correlations between endogenous ligand affinities and minimum drug affinities (Pearson correlation coefficient = 0.87; P = 1.59E–8; r^2 = 0.76) on the one hand, and range of drug affinities (Pearson correlation coefficient = –0.79; P = 3.28E-6; r^2 = 0.62) on the other hand.

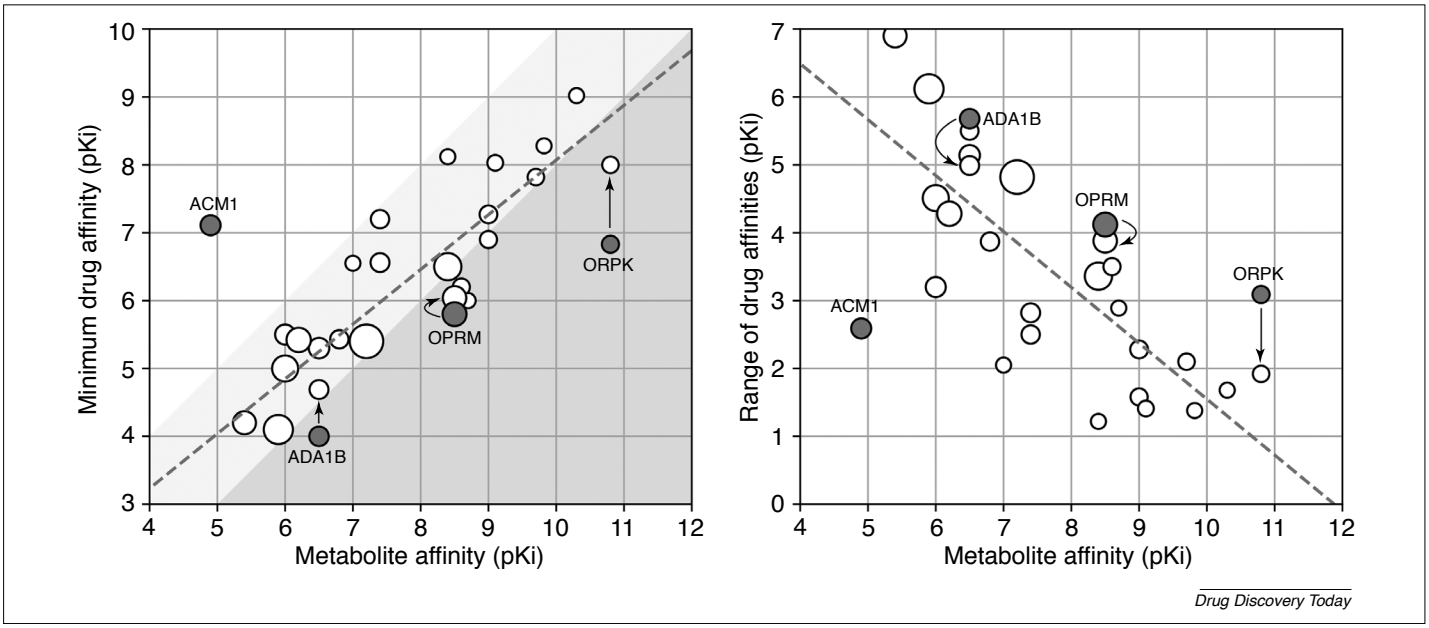

FIGURE 4

Metabolite affinity as reference baseline for the primary pharmacology of drugs. Each circle represents one of the 26 targets for which pK_i data are available for more than three drugs. The size of the circle is proportional to the number of drugs associated with each target. Trends show that endogenous metabolite affinity is (i) directly proportional to the minimum affinity of all drugs having the protein as primary target; and (ii) inversely proportional to the range of drug affinities. Marked outliers (grey circles) are the muscarinic acetylcholine receptor M1 (ACM1), the κ opioid receptor (OPRK), the μ opioid receptor (OPRM) and the α1B adrenergic receptor (ADRA1B). For OPRK, OPRM, and ADRA1B, their positions after discarding oxymorphone, tramadol, and methoxamine, respectively, are also included and indicated with an arrow from the original point. Dotted lines in both cases reflect the existing direct and inverse linear correlations.

Extended primary pharmacology of drugs and endogenous ligands

To assemble the first set of 442 drug–receptor–metabolite triad associations from which trends between the affinities of drugs and endogenous ligands for the same target were derived, we relied exclusively on two public sources of highly curated data, namely, DrugCentral [7] for drug–target interactions and GtoPdb [14] for metabolite–target interactions. To assess the general validity of those trends beyond the set of pharmacological data from which they were derived, we collected a second set of 202 additional drug–receptor–metabolite triad associations (Table S3 in the supplemental information online) with affinities for drug–target interactions available in DrugCentral [7], GtoPdb [14], and ChEMBL [28], and metabolite–target affinities available in GtoPdb [14] and ChEMBL [28]. By including new data sources, we mostly added new drug–target interactions for targets already present in the original set. However, from the total number of 145 drugs, 55 targets and 34 metabolites involved in those 202 new triads, 110 drugs, nine targets, and nine metabolites were not represented in the original set.

The density plot of this extended set of 202 drug–receptor–metabolite triads (Fig. S1 in the supplemental information online) is similar to the distribution displayed in Fig. 1, with 60% of drug affinities being higher than the affinities of the endogenous ligand for the same target, and 93% of them fitting within a range of two orders of magnitude. When these 202 new triads were added to the original 442 triads, the list of targets for which pK_i data for more than three drugs were available increased from 26 to 37. Accordingly, Fig. 5 represents an expanded version of Fig. 4, in which the 11 new targets are marked as white circles with thick-black borders.

As can be observed in Fig. 5a, besides the special cases of ADA1B, OPRM, and OPRK (drawn as grey circles) already discussed in Fig. 4, there are now some other targets inside the 'forbidden' zone of minimum drug affinities two orders of magnitude below the endogenous metabolite affinities for the same target. One of them is the α_{1D} adrenoceptor (ADA1D). After removal of methoxamine (one of the outlier drugs discussed earlier), its position moves up significantly. Other targets are the prostaglandin F2-alpha receptor (PF2R), which includes the outlier drug bimatoprost, and the androgen receptor (ANDR), which includes the minimum-affinity drug flutamide. Flutamide is a nonsteroidal molecule acting as a selective antagonist of ANDR. However, its binding affinity to ANDR ($pK_i = 5.89$) is almost four orders of magnitude lower than the affinity of dihydrotestosterone ($pK_i = 9.7$), the native endogenous metabolite for ANDR. Interestingly, it has long been proved that the action of flutamide is conducted through one of its main metabolites, 2-hydroxyflutamide [29], which has a significantly higher binding affinity for ANDR ($pK_i = 7.65$) [28]. The removal of flutamide brings ANDR closer to the region of accepted minimum drug affinity for this target. Another target marked as outlier in Fig. 4, ACM1, now fits well within the expected region as a new drug, cevimeline, with a pK_i of 4.9, was incorporated in the set. These observations emphasise the importance of data completeness in this type of analysis [27].

After repositioning of those target outliers in Fig. 5a, endogenous ligand affinities for a given protein were directly proportional to the currently known minimum affinities of drugs for that protein, with a similar correlation to that observed already in Fig. 4 (Pearson correlation coefficient = 0.83; *P* value = 5.49E–10; $r^2 = 0.68$). By contrast, the inverse correlation between

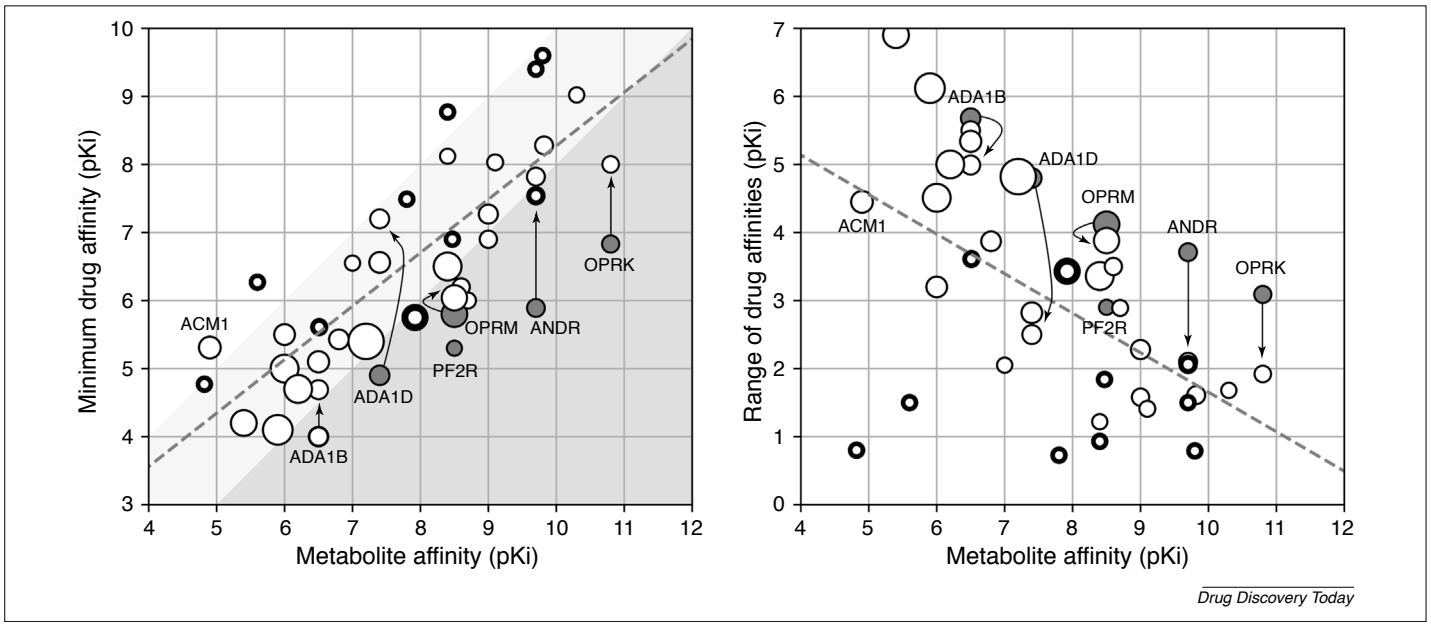

FIGURE 5

Endogenous metabolite affinity as reference baseline for the primary pharmacology of drugs. Each circle represents one of the 37 targets for which pK_i data are available for more than three drugs. The size of the circle is proportional to the number of drugs associated with each target. Newly incorporated targets with respect to Fig. 4 (main text) appear as circles with thick-black borders. The same trends observed in Fig. 4 (main text) hold true: the endogenous ligand affinity for a protein is: (i) directly proportional to the minimum affinity of the drug for that protein; and (ii) inversely proportional to the range of drug affinities for that protein. Targets with outlier drugs discussed in the main text and removed from the correlations are shown in grey and their corrected position indicated by an arrow.

endogenous ligand affinities and the range of drug affinities for a given target (Fig. 5b) is still significant (Pearson correlation coefficient = –0.57; *P* value = 3.01E–4; $r^2 = 0.32$), but some of the new targets with just three representative drugs show drug affinity ranges clearly below the expected values for their respective targets. As already mentioned, this type of analyses is especially sensitive to data completeness and the range of drug affinities are susceptible to variations as new drugs are included. As can be observed, most of the targets located well below the correlation line are in fact targets with a small number of drugs (three), whereas targets with more representative drugs cover the full range of drug affinities from the baseline set by the endogenous ligand affinity up to nanomolar potencies.

Extended primary pharmacology of drugs and endogenous substrates

To assess the validity of trends observed for receptors on other protein families, we compiled a second external biochemistry dataset of *in vitro* binding affinities of drugs and endogenous substrates for enzymes. Accordingly, drug–enzyme interactions were retrieved from the DrugCentral repository [7]. The list of unique human enzymes involved in the mechanism of action of drugs was then used to interrogate the BRaunschweig ENzyme DAtabase (BRENDA) [30] to identify the endogenous substrate for each enzyme target and to extract its corresponding maximum Michaelis constant value, $K_M = (k_{-1} + k_2)/k_1$, where $(k_{-1} + k_2)$ is the rate of breakdown and k_1 is the rate of product formation. For the sake of simplicity, we consider K_M as an estimate of the dissociation constant for the enzyme–substrate complex when $k_2 << k_{-1}$ (thus, $K_M \sim k_{-1}/k_1$). Under these conditions, K_M can be directly compared with K_i values (i.e., with the dissociation constants of the enzyme–inhibitor complex). To achieve effective inhibition, K_i values for competitive inhibitors would need to be higher than K_M values, on the negative log(molar) scale, to overcome the effect of accumulating substrate. This is not relevant for covalent inhibitors, which maintain effectiveness regardless of increasing substrate concentration, and for allosteric inhibitors, which bind to a noncanonical site and, thus, do not compete with the endogenous substrate.

A total of 222 drug–enzyme–substrate triad associations were collected, involving 164 approved drugs, 41 mechanism of action enzymes, and 32 human endogenous substrates (Table S4 in the supplemental information online). The density plot of drug and endogenous substrate affinities for the same enzyme shows a clear accumulation of interactions above the diagonal solid line (Fig. S2 in the supplemental information online). From a target-centred perspective, analogous to Fig. 3 for receptors, Fig. 6 allows for assessing how the 222 drug affinities compare with the corresponding endogenous substrate affinities across all 41 enzymes. The darker tip of each column in the circular plot of Fig. 6 reflects the affinity between the enzyme and its main endogenous substrate. Compared with earlier findings on drug–receptor interactions, for which 96% of drug affinities for their primary targets are two orders of magnitude lower than the corresponding endogenous metabolite affinities, we found that 88% of the drug–enzyme interactions have affinity values above the affinity of the natural substrate of the enzyme, with 78% and 60% of drug–enzyme affinities being at least one and two orders of magnitude higher than the corresponding substrate–enzyme affinities, respectively. This represents a two order of magnitude shift for drug–enzyme affinities relative to their endogenous substrates compared with drug–receptor affinities relative to their endogenous ligands.

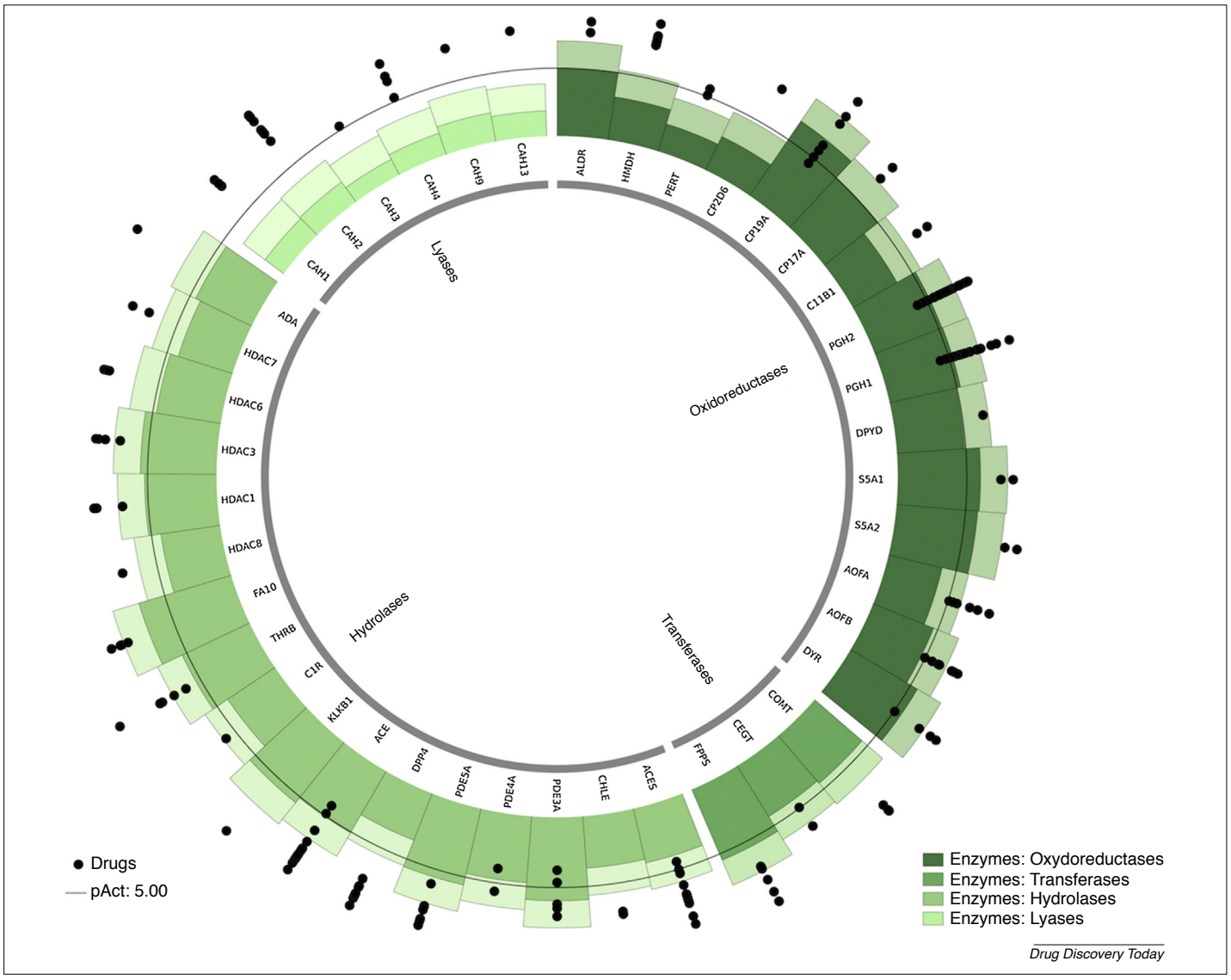

FIGURE 6

Target-centred analysis of the 222 drug–enzyme–substrate triads. Each of the 41 enzymes is represented by a column in this circular plot. In contrast to the receptor families, the height of the column reflects the two order of magnitude window, in lighter colour, above the affinity of the endogenous substrate for its native enzyme. The darker height of each column marks the actual affinity of the endogenous substrate. Drug affinities for their primary target(s) are displayed as black circles. The solid-black line crossing all columns serves as the 10-μM reference affinity level for all drug and substrate affinities.

Impact on secondary pharmacology: the case of 5HT2B

The difference of two orders of magnitude below the endogenous ligand affinity for a drug affinity to be biologically relevant can have important implications beyond primary pharmacology. In this respect, there is currently ample evidence that drugs bind to multiple proteins [31]. This polypharmacology is of particular concern for drugs targeting GPCRs [32], given that it has been estimated that, on average, they can have biologically relevant binding affinities for up to ten members of this protein family [33]. Although some of this secondary pharmacology might be necessary for the efficacy of drugs addressing complex diseases [34], binding to certain off-targets can lead to serious drug safety issues [35].

One of these red-flag off-targets is the serotonin receptor 5HT2B. Small-molecule drugs acting as 5HT2B agonists have long been associated with valvular heart disease (VHD) [36] and most new drug submissions to regulatory agencies require now both binding and functional testing to assess 5HT2B agonist activity. In this respect, a recent report from a regulatory agency on the use of *in vitro* secondary pharmacology to assess the risk of VHD concluded that safety margins based on *in vitro* binding affinities (pK_i), or those relative to serotonin, appear to be a better predictor for determining the risk of a 5HT2B agonist to produce VHD compared with measures involving the maximum therapeutic free plasma drug concentration *in vivo* [37]. Most interestingly, they observed that nonvalvulopathic 5HT2B agonist drugs have a pK_i value over two orders of magnitude lower than that of serotonin [37]. Even though their analysis was limited to nine drugs, the suggestion of a 100-fold difference in K_i values between the *in vitro* affinities of the endogenous ligand and the drug for the off-target agrees well with the trends observed in the current on the basis of 442 interactions for 293 drugs.

To assess how the link between *in vitro* binding affinities of 5HT2B agonists and VHD would fit within the framework established here, we extended from nine to 24 the set of 5HT2B agonists with VHD information. Among them, 12 are known valvulopathic drugs [37–45], whereas the other 12 are assumed to be nonvalvulopathic, given that no bibliographic evidence of links to VHD was found. All binding (pK_i) and VHD risk data for these 24 5HT2B agonist drugs are provided in Table S5 in the supplemental information online. The distribution of *in vitro* binding affinities of the 24 5HT2B agonists is provided in Fig. 7, which is analogous to Fig. 1 but centred solely on the serotonin affinity for 5HT2B (pKi = 8.4).

As can be observed, there is a clear separation between valvulopathic (in red) and nonvalvulopathic (in green) drugs. Of the 12 valvulopathic drugs, five (42%) have pK_i values equal to or larger than the serotonin affinity, seven (58%) within an order of magnitude of the serotonin affinity, and 11 (92%) within two orders of magnitude of the serotonin affinity. Only one of them, dexfenfluramine, has a pK_i value clearly below this two-order of magnitude window, although its racemic mixture, fenfluramine, is right at the edge of it (pK_i = 6.4). In fact, it was the case of fenfluramine that alerted researchers almost 20 years ago to the fact that its more potent 5HT2B agonist metabolite, norfenfluramine, could be responsible for its associated risk for VHD [46,47]. Indeed, the binding affinity of norfenfluramine for 5HT2B (pK_i = 7.28) puts this drug metabolite well into the risk zone for VHD. By contrast, of the 12 nonvalvulopathic drugs, seven (58%) have pK_i values below the two order of magnitude mark of the serotonin affinity. The five nonvalvulopathic drugs found within two orders of

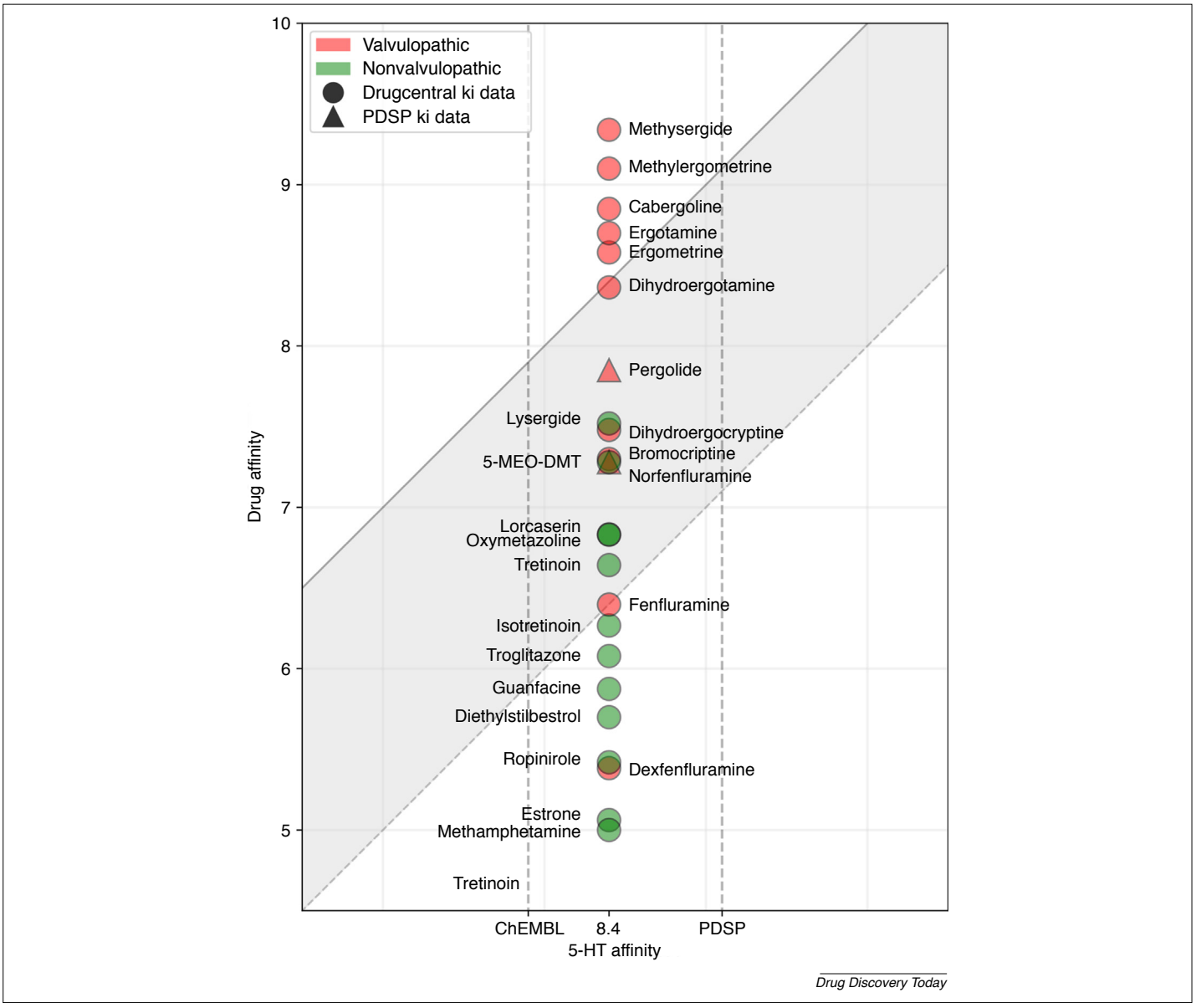

FIGURE 7

Distribution of binding affinities (pK_i) for 24 serotonin receptor subtype 2B (5HT2B) agonists. The set includes 12 valvulopathic drugs (in red) and 12 nonvalvulopathic drugs (in green). Drug affinities are aligned at the maximum binding affinity of serotonin for 5HT2B found in Guide to Pharmacology database (GtoPdb) (8.4). For the sake of comparison, the corresponding serotonin affinities reported in ChEMBL [28] and PDSP [48] (vertical dashed lines) are 7.9 and 9.1. The diagonal grey area marks the region where drug affinities lie between the serotonin affinity (solid line) and the 100-fold window (dashed line).

magnitude of the serotonin affinity are tretinoin, oxymetazoline, lorcaserin, 5-MEO-DMT, and lysergide. Interestingly, clinical monitoring on the risk of VHD has been already recommended for lorcaserin [37].

Overall, our work confirms and strengthens the early recommendation that a 100-fold difference in K_i values between the serotonin and drug affinities for 5HT2B is a reasonable criterion to discriminate valvulopathic from nonvalvulopathic drugs [37]. This notwithstanding, one ought to consider the effects of data dispersion in binding affinities, which could alter, in one direction or another, the final ability of a drug to produce VHD. As an example, we took as a reference value from Fig. 7 the maximum affinity of serotonin for 5HT2B reported in GtoPdb ($pK_i = 8.4$). However, the corresponding binding affinities found in other public sources differ slightly. For example, the pK_i values in ChEMBL and PDSP are 7.9 and 9.1, respectively [28,48]. Considering these pK_i values in Fig. 7, the serotonin affinity in ChEMBL ($pK_i = 7.9$) would move the entire drug affinity distribution to the left, which would cause the nonvalvulopathic drugs isotretinoin and troglitazone to enter the grey zone. However, the serotonin affinity in PDSP ($pK_i = 9.1$) would move the entire drug affinity distribution to the right, which would then result in ten out of the 12 nonvalvulopathic drugs being safely placed below the two-order of magnitude window. Thus, data robustness and dispersion are important aspects to consider in this type of analysis.

Implications for initial assessments of safety margins

Many drug safety events are caused by the interaction of drugs with their own primary targets or with secondary (off-)targets. To lower the chance for undesirable off-target effects, *in vitro* safety pharmacology profiling is now an integral part of preclinical drug discovery [3] and its utility for the early risk assessment of drug-induced safety events has been well recognised [36]. Normally, off-target safety margins are defined by dividing the concentration of the drug that is required for 50% off-target inhibition *in vitro* (IC_{50}) by the maximum plasma concentration of the drug *in vivo* (C_{max}) [49]. However, *in vivo* C_{max} values are unlikely to be collected for large numbers of compounds during the early stages of a drug discovery project and, in addition, it was recently shown that, at least for 5HT2B, safety margins based on *in vitro* binding affinities (K_i values), or those relative to the endogenous ligand, are better safety predictors than the use of *in vivo* C_{max} values [36]. This prompted us to elaborate on this aspect based on the results presented earlier.

Strategies to widen the off-target safety margins involve increasing the potency of the drug for the primary target and decreasing the potency of the drug for the off-target [49]. In this respect, the objective is usually to achieve a 30- to 100-fold selectivity between the affinities for the primary target, $pK_i(pT)$, and the off-target, pKi (oT) [36,49]. However, the outcomes presented earlier could change the perspective from which off-target safety margins have been traditionally regarded. The concept is illustrated in Fig. 8a for two hypothetical drugs (labelled as 1 and 2) representing two opposite case scenarios: on the one hand, the affinity of drug 1 for its primary target (circle labelled as pT_1) is 100-fold higher than the affinity for one of its off-targets (orange square labelled as oT_1), yet both values are above the affinities of the respective endogenous ligands for pT_1 and oT_1 and, thus, both are likely to be biologically relevant; on the other hand, the affinity of drug 2 for its primary target (circle labelled as pT_2) is tenfold lower than that for one of its off-targets (green square labelled as oT_2), and the latter is much lower than the affinity of the corresponding endogenous ligand for oT_2, thus below the grey zone of potential safety risk associated with the oT_2 interaction (see Fig. 7). The result is that, despite its 100-fold selectivity, drug 1 might be at risk of producing the safety issue associated with affinity to oT_1, whereas the lack of selectivity might not be an issue for drug 2 to be safe of the adverse event linked to oT_2.

Among the list of 5HT2B agonist drugs discussed earlier (Fig. 7), bromocriptine and troglitazone are case studies resembling the hypothetical examples of drugs 1 and 2 in Fig. 8a. The corresponding safety diagram for these two drugs is presented in Fig. 8b. The pK_i of serotonin for 5HT2B is 8.40 [28], which establishes the baseline against which drug affinities for 5HT2B would need to be evaluated. The affinity of bromocriptine for its primary target (DRD2), $pK_i(pT) = 9.70$ [28], is over 100-fold higher than that for 5HT2B, $pK_i(oT) = 7.30$ [28], yet its off-target affinity is approximately tenfold lower relative to the corresponding affinity for the endogenous ligand, placing this drug within the safety-risk grey zone. Bromocriptine exemplifies the hypothetical drug 1 case: despite an ample pKi-based off-target safety margin, $pK_i(pT) - pK_i(oT) = 2.40$, the difference between the affinities of the endogenous ligand and the drug, $pK_i(\text{serotonin}) - pK_i(\text{bromocriptine}) = 1.10$, foreshadows its potential risk for VHD. By contrast, the affinity of troglitazone for its primary target (peroxisome proliferator activated receptor γ; PPARG), $pK_i(pT) = (5.42,6.52)$ [28], does not differ much from that for 5HT2B, $pK_i(oT) = 6.08$ [28], yet the latter value is over 100-fold lower than the corresponding affinity of serotonin, placing it well below the safety-risk grey zone. Troglitazone embodies the hypothetical drug 2 case: despite poor selectivity compared with the off-target, pKi(pT) – pKi(oT) = (−0.66, + 0.44), the difference between the affinities of the endogenous ligand and the drug, $pK_i(\text{serotonin}) - pK_i(\text{troglitazone}) = 2.32$, is a better predictor for its low risk in causing VHD.

Overall, these results would favour a K_i-based safety margin (SM) defined as the difference in off-target binding affinities between the endogenous metabolite (M), $pK_i^M(oT)$, and the drug (D), $pK_i^D(oT)$ (Eq. 1):

$$SM = pK_i^M(oT) - pK_i^D(oT) \quad (1)$$

in contrast to the traditional definition based on the difference in binding affinities of the drug between the target and the off-target (Eq. 2):

$$SM = pK_i^D(pT) - pK_i^D(oT) \quad (2)$$

Values of $SM > 2$ would be recommended.

Some practical limitations

However, there are some aspects of the present study that merit further consideration. As already highlighted, there are limitations associated with data completeness and bias, always present in this type of analysis [27]. Among the set of 293 drugs collected in our set, there are representatives of 11 out of the 14 topmost levels of the Anatomical Therapeutic Chemical classification system of drugs [50] but almost 25% of the 442 interactions implicate drugs of the nervous system (N level). In addition, even though there are

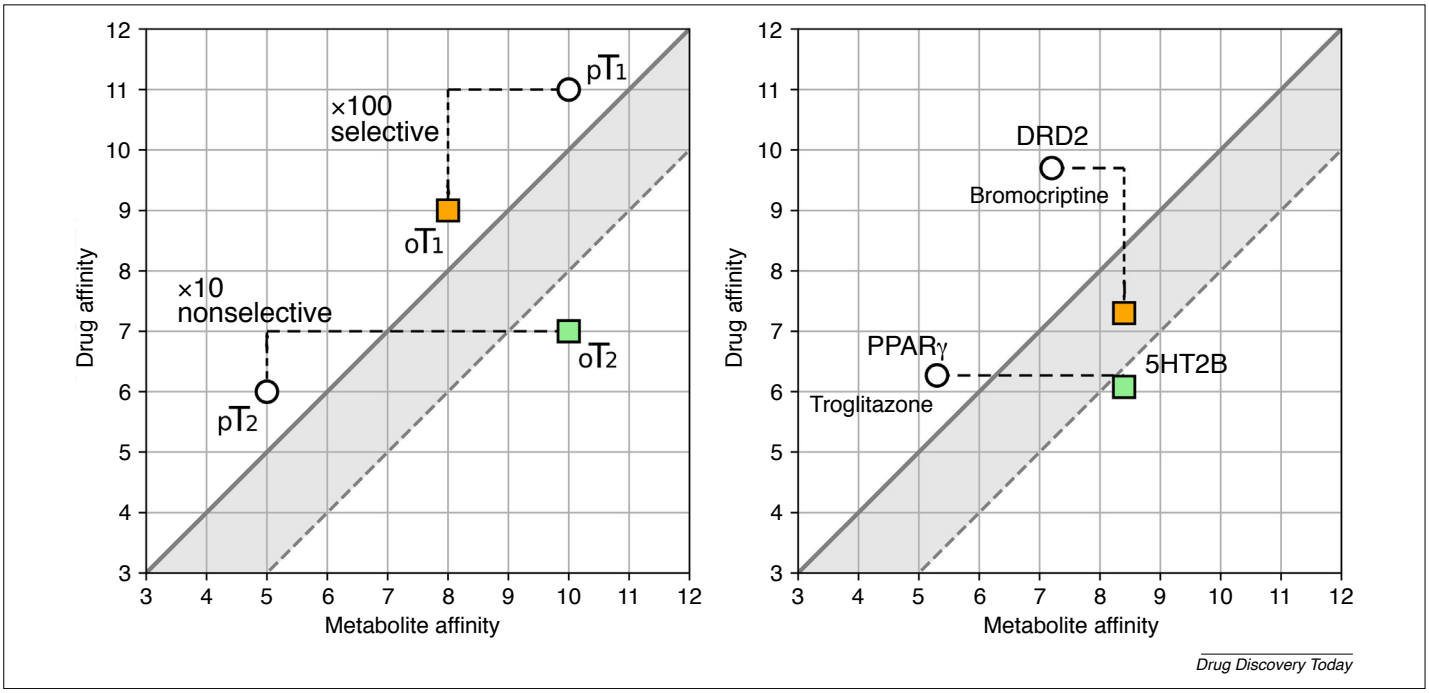

FIGURE 8

A new perspective on safety margins considering endogenous ligand affinities. (a) Schematic plotting endogenous metabolite affinities versus drug affinities for a set of four exemplary proteins: two primary targets (pT_1 and pT_2, represented by circles) and two off-targets (oT_1 and oT_2, represented by squares). Despite its 100-fold selectivity, drug 1 might be at risk of showing the safety issue linked to off-target oT_1 (orange square), whereas drug 2 might be safe, even though its affinity for off-target oT_2 (green square) is tenfold higher than the affinity for its primary target pT_2. The diagonal grey area marks the region where drug affinities lie within two orders of magnitude (dashed line) below the endogenous ligand affinity (solid line); (b) the same diagram plotting the cases of bromocriptine and troglitazone. Binding affinities (pK_i) for their respective primary targets dopamine [D2 receptor (DRD2) and peroxisome proliferator activated receptor γ (PPARG)] and the 5HT2B off-target are plotted against the corresponding affinities for the endogenous ligands, namely, dopamine for DRD2, linoleic acid for PPARG, and serotonin for serotonin receptor subtype 2B (5HT2B). See main text for discussion.

43 endogenous ligands assigned to those 442 interactions, a single one (serotonin) is involved in almost 15% of them. This reflects the fact that 70 of the 79 receptors for which affinities with the same units were found in public sources for both endogenous ligands and drugs are GPCRs. A similar situation is encountered in the enzyme data set, in which a single molecule among the 32 endogenous substrates, arachidonate, is involved in 27% of the 222 interactions. This data bias stresses our current limited knowledge on the pharmacology of the human endogenous metabolome. In this respect, to compile a reference repository of metabolite affinity baselines, it is important to coordinate worldwide to identify the main endogenous ligands of human proteins, measure their *in vitro* binding affinities, and characterise their complete pharmacological profile across multiple protein family members, making all data publicly available to the research community.

In addition, an underlying assumption of the entire analysis is that, for a fair comparison of affinity values, drugs and endogenous ligands interact with the same protein at the same binding site. For receptors, based on the well-established similarity between drugs and metabolites [51–54], one could take this assumption for granted, because similar small molecules are expected to bind to similar protein sites, but the ultimate proof would come from structural data. For enzymes, this aspect should be less crucial because most drugs and substrates are expected to bind at the same catalytic site. In all cases, orthosteric binding was assumed and the possibility of allosteric effects was not considered.

Unfortunately, if consistent affinity data of both metabolites and/or substrates and drugs for the same protein were scarce, structural data of the complex between the protein with metabolites and/or substrates and drugs are rarer still. This notwithstanding, we searched the Protein Data Bank (PDB) [55] for entries of proteins that were cocrystallised with both their endogenous metabolite and/or substrate and one of the drugs from our list. Two illustrative examples for GPCRs were found, namely, the β2-adrenergic receptor (ADRB2) and AA2AR. ADRB2 has been cocrystallised with its endogenous ligand, adrenaline (4ldo), and a drug antagonist, timolol (3d4s), whereas structures of AA2AR cocrystallised with its endogenous ligand, adenosine (2ydo), and two drug antagonists, theophylline (5mzj) and caffeine (3rfm), were also identified. The backbone superposition of the binding cavity of AA2AR with the metabolite and two drugs is show in Fig. 9a. For the enzyme data set, protein structures cocrystallised with both the endogenous substrate and a drug were found for ten out of the 222 drug–enzyme–substrate triads collected. In total, 43 PDB entries were identified, 38 of which involving carbonic anhydrase, three for 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase, and one for aldose reductase. The backbone superposition of the catalytic site of HMG-CoA reductase bound to its endogenous substrate, HMG-CoA (1dqn), and two drug inhibitors, atorvastatin (1hwk) and fluvastatin (1hwi) is illustrated in Fig. 9b. For both proteins, the alignment confirms that the endogenous metabolite and/or substrate and the drugs bind to the same site.

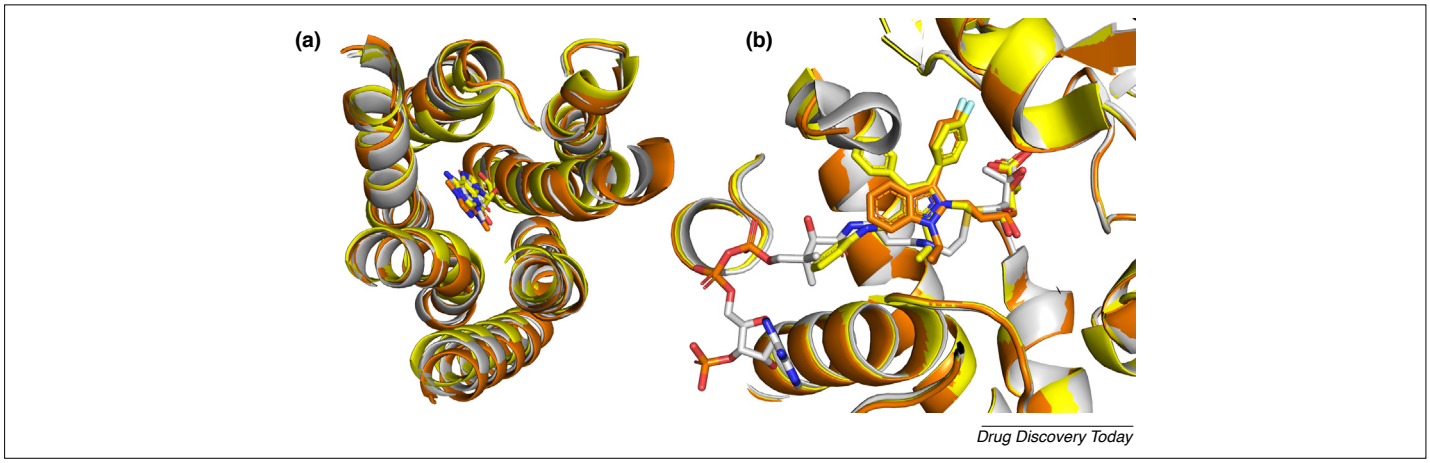

FIGURE 9

Pairs of metabolite and/or substrate and drugs binding at the same protein site. Backbone superpositions of (a) the adenosine 2A receptor cocrystallised with its endogenous ligand, adenosine [Protein Data Bank (PDB) 2ydo, in white), and two drug antagonists, theophylline (PDB 5mzj, in yellow) and caffeine (PDB 3rfm, in orange), and (b) the 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase enzyme cocrystallised with its endogenous substrate, HMG-CoA (PDB 1dqn, in white), and two drug inhibitors, atorvastatin (PDB 1hwk, in yellow) and fluvastatin (PDB 1hwi, in orange).

Accordingly, along the same lines expressed earlier for pharmacological data, more efforts to resolve crystal structures of protein–metabolite and/or substrate–drug complexes for which *in vitro* affinity data are available would be an informative addition to a reference human endogenous ligand repository.

Finally, we emphasise that there are multiple and complex components involved in both the biological action of endogenous ligands and the therapeutic effect of a drug. In particular, metabolite abundance (concentration of the endogenous ligand at the site of action) and drug exposure (concentration of the drug over time at the target tissue) are two major factors [49], the effect of which was not considered in our analyses. Metabolite abundance varies over time, cell type, tissue, and environmental conditions and depends largely on the individual endogenous ligand [56]. Likewise, drug exposure depends on multiple pharmacokinetic factors (e.g., half-life, distribution, and clearance) that are affected by interindividual variations (e.g., body mass, metabolism, drug–drug interactions, comorbidities, and other environmental factors). Receptor occupancy, the on/off rate captured by kinetic constants, as well as considerations of high- and low-affinity states for receptors and coexisting catalytic efficiency states for enzymes were also not taken into account. It was not the aim of this work to model the complexity of the process but to highlight simple trends that were identified using *in vitro* binding affinities of drugs relative to those of endogenous ligands as a contributing factor in the *in vitro* to *in vivo* translatability of drugs.

Concluding remarks

Understanding why a specific drug needs a certain level of affinity for its primary protein target(s) to exert its therapeutic action is essential for drug discovery. Based on a curated collection of 442 interactions between 293 drugs, 79 receptors, and 43 endogenous ligands, this study demonstrates that the affinity of an endogenous ligand for its native receptor can be used as a reference baseline for the primary pharmacology of drugs. Our findings reveal that 67% of all drug–receptor interactions have affinity values above the corresponding metabolite–receptor affinities and that up to 96% of those drug affinities have values within two orders of magnitude of metabolite affinities for the same protein. An analysis of the remaining 4% of drug–receptor interactions indicates that primary targets assigned to drugs with affinities below two orders of magnitude of metabolite affinities should be critically revised.

The relationship between the affinity of both the drug and the endogenous ligand for the same protein target was further validated on an external set of 202 interactions between 145 drugs, 55 receptors, and 34 endogenous ligands, which included 110 drugs, nine receptors, and nine metabolites that were not originally considered. The results confirmed the trends observed previously, with 60% of drug affinities being higher than the affinities of the endogenous ligand for the same target, and 93% of them fitting within a range of two orders of magnitude of the corresponding metabolite affinities.

Furthermore, a second external set of 222 interactions involving 164 drugs, 41 enzymes, and 32 endogenous substrates revealed that there is a two order of magnitude shift for drug–enzyme affinities relative to endogenous substrates compared with drug–receptor affinities relative to endogenous ligands, with 60% of enzyme drug inhibitors having affinity values over two orders of magnitudes higher than the corresponding substrate affinities. Overall, our analysis comprised a set of 815 highly-curated interactions between 566 drugs and 129 primary targets, associated with 79 endogenous ligands. In total, 71% of the drug-target affinities have values above that of the corresponding endogenous ligand, with 96% placed within two orders of magnitude.

Overall, our analysis comprised a set of 815 highly-curated interactions between 566 drugs and 129 primary targets, associated with 79 endogenous ligands. In total, 71% of the drug-target affinities have values above that of the corresponding endogenous ligand, with 96% placed within two orders of magnitude. In addition, our findings that the human endogenous metabolome could serve as a pharmacology baseline for drug *in vitro* affinities on proteins was strengthened by the observations that endogenous ligand affinities for their native proteins are, on the one hand, directly proportional to the minimum drug affinity among

all drugs optimised for a given primary target and, on the other hand, inversely proportional to the range of drug affinities for each primary target.

Besides the implications for the primary pharmacology of the drug, our results also highlight the impact that *in vitro* endogenous ligand affinities can have on assessing the risk of safety events linked to the secondary pharmacology of the drug. As an illustrative example, the case of VHD related to long-term agonist action on the 5HT2B receptor was presented. Our results confirm that a difference of two orders of magnitude below the *in vitro* affinity of serotonin for 5HT2B successfully separates valvulopathic drugs from drugs devoid of risk to produce VHD, in agreement with the recommendation formulated recently by a regulatory agency [37].

The link between the *in vitro* binding of small molecules to certain disease-relevant proteins and its ultimate translation into an *in vivo* phenotypic outcome is one of the main pillars of drug discovery [57]. However, this *in vitro* to *in vivo* extrapolation remains challenging, because differences in the activity of compounds in biochemical assays and in cellular, tissue, or organism assays are common and difficult to understand in the context of a biological system, where multiple factors intervene [58]. Our findings suggest that the endogenous metabolome is one of those factors. The fact that simple trends between metabolite affinities and the primary and secondary pharmacologies of a drug could be derived indicates that more research should be devoted to further understand the true reach of the impact of the human endogenous metabolome on the efficacy and safety of drugs. Up to 79% of drug clinical failures remain to be attributable to safety or efficacy reasons [59]. In this respect, the recommendations outlined here based on preclinical *in vitro* pharmacology data could provide additional metrics to assess the risk of clinical failure and contribute to reduce drug attrition.

Conflicts of interest

T.I.O. was a former full-time employee at AstraZeneca (1996–2002). He has received honoraria, or consulted for, Abbott, AstraZeneca, Chiron, Genentech, Infinity Pharmaceuticals, Merz Pharmaceuticals, Merck Darmstadt, Mitsubishi Tanabe, Novartis, Ono Pharmaceuticals, Pfizer, Roche, Sanofi, and Wyeth. His spouse was a full-time employee of AstraZeneca (2002–2014) and is a full-time employee of Genentech. J.M. was a former full-time employee at Pharmacia & Upjohn (1996) and Organon (1997-2003). He is currently CEO of Chemotargets.

Acknowledgements

We are grateful to Rita Moita Santos and Anne Hersey for their kind assistance in the construction of the circular plots in Figs 3 and 6. Ki data for serotonin, pergolide and norfenfluramine were generously provided by the National Institute of Mental Health's Psychoactive Drug Screening Program, Contract HHSN-271-2013-00017-C (NIMH PDSP). This work was supported by the European H2020 ESCAPE-NET Project (Grant Agreement 733381), by a RETOS project from the Spanish Ministerio de Ciencia, Innovación y Universidades (SAF2017-83614-R) and by the National Institutes of Health Common Fund program (Grants U24 CA224370, U24 TR002278 and U01 CA239108).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.drudis.2019.06.007.

References

1 Hughes, J.P. *et al.* (2011) Principles of early drug discovery. *Br. J. Pharmacol.* 162, 1239–1249
2 Wagner, J. *et al.* (2018) A dynamic map for learning, communicating, navigating and improving therapeutic development. *Nat. Rev. Drug Discov.* 17, 150
3 Whitebread, S. *et al.* (2005) In vitro safety pharmacology profiling: an essential tool for successful drug development. *Drug Discov. Today* 10, 1421–1433
4 Smith, D.A. *et al.* (2010) The effect of plasma protein binding on *in vivo* efficacy: misconceptions in drug discovery. *Nat. Rev. Drug Discov.* 9, 929–939
5 Schwaid, A.G. and Cornella-Taracido, I. (2018) Causes and significance of increased compound potency in cellular or physiological contexts. *J. Med. Chem.* 61, 1767–1773
6 Dambach, D.M. *et al.* (2016) Safety lead optimization and candidate identification: integrating new technologies into decision-making. *Chem. Res. Toxicol.* 29, 452–472
7 Ursu, O. *et al.* (2017) DrugCentral: online drug compendium. *Nucleic Acids Res.* 45, D932–D939
8 Luffer-Atlas, D. (2008) Unique/major human metabolites: why, how, and when to test for safety in animals. *Drug Metab. Rev.* 40, 447–463
9 Rodríguez-Morató, J. *et al.* (2018) Targeting human urinary metabolome by LC-MS/MS: a review. *Bioanalysis* 10, 489–516
10 Metallo, C.M. and Vander Heiden, M.G. (2013) Understanding metabolic regulation and its influence on cell physiology. *Mol. Cell* 49, 388–398
11 Li, X. and Snyder, M. (2011) Metabolites as global regulators: a new view of protein regulation: systematic investigation of metabolite-protein interactions may help bridge the gap between genome-wide association studies and small molecule screening studies. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 33, 485–489
12 Wishart, D.S. *et al.* (2018) HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 46, D608–D617
13 Suhre, K. *et al.* (2011) Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 477, 54–60
14 Sharman, J.L. *et al.* (2018) Accessing expert-curated pharmacological data in the IUPHAR/BPS Guide to Pharmacology. *Curr. Protoc. Bioinforma.* 61, 1.34.1–1.34.46
15 Ahuja, V. and Sharma, S. (2014) Drug safety testing paradigm, current progress and future challenges: an overview. *J. Appl. Toxicol. JAT* 34, 576–594
16 Kalliokoski, T. *et al.* (2013) Comparability of mixed IC50 data – a statistical analysis. *PLoS ONE* 8, e61007
17 Steinfels, G.F. and Cook, L. (1986) Antinociceptive profiles of mu and kappa opioid agonists in a rat tooth pulp stimulation procedure. *J. Pharmacol. Exp. Ther.* 236, 111–117
18 Smith, H.S. (2009) Clinical pharmacology of oxymorphone. *Pain Med.* 10, S3–S10
19 Yoo, J.H. *et al.* (2017) Characterization of [3H] oxymorphone binding sites in mouse brain: quantitative autoradiography in opioid receptor knockout mice. *Neurosci. Lett.* 643, 16–21
20 Lai, J. *et al.* (1996) Tramadol, M1 metabolite and enantiomer affinities for cloned human opioid receptors expressed in transfected HN9.10 neuroblastoma cells. *Eur. J. Pharmacol.* 316, 369–372
21 Krauss, A.H.-P. and Woodward, D.F. (2004) Update on the mechanism of action of bimatoprost: a review and discussion of new evidence. *Surv. Ophthalmol.* 49 (Suppl. 1), S5–11
22 Woodward, D.F. *et al.* (2003) Pharmacological characterization of a novel antiglaucoma agent, Bimatoprost (AGN 192024). *J. Pharmacol. Exp. Ther.* 305, 772–785
23 Woodward, D.F. *et al.* (2001) The pharmacology of bimatoprost (Lumigan). *Surv. Ophthalmol.* 45 (Suppl. 4), S337–S345
24 Taniguchi, T. *et al.* (1999) Microphysiometric analysis of human alpha1a-adrenoceptor expressed in Chinese hamster ovary cells. *Br. J. Pharmacol.* 127, 962–968
25 Guardiola-Lemaitre, B. *et al.* (2014) Agomelatine: mechanism of action and pharmacological profile in relation to antidepressant properties. *Br J Pharmacol.* 171, 3604–3619
26 Borsini, F. *et al.* (2002) Pharmacology of flibanserin. *CNS Drug Rev.* 8, 117–142
27 Mestres, J. *et al.* (2008) Data completeness – the Achilles heel of drug-target networks. *Nat. Biotechnol.* 26, 983–984
28 Gaulton, A. *et al.* (2017) The ChEMBL database in 2017. *Nucleic Acids Res.* 45, D945–D954

29 Schulz, M. *et al.* (1988) The pharmacokinetics of flutamide and its major metabolites after single oral dose and during chronic treatment. *Eur. J. Clin. Pharmacol.* 34, 633–636
30 Schomburg, I. *et al.* (2017) The BRENDA enzyme information system – from a database to an expert system. *J. Biotechnol.* 261, 194–206
31 Jalencas, X. and Mestres, J. (2013) On the origins of drug polypharmacology. *MedChemComm* 4, 80–87
32 Sexton, P.M. and Christopoulos, A. (2018) To bind or not to bind: unravelling GPCR polypharmacology. *Cell* 172, 636–638
33 Mestres, J. *et al.* (2009) The topology of drug-target interaction networks: implicit dependence on drug properties and target families. *Mol. Biosyst.* 5, 1051–1057
34 Anighoro, A. *et al.* (2014) Polypharmacology: challenges and opportunities in drug discovery. *J. Med. Chem.* 57, 7874–7887
35 Whitebread, S. *et al.* (2016) Secondary pharmacology: screening and interpretation of off-target activities – focus on translation. *Drug Discov. Today* 21, 1232–1242
36 Roth, B.L. (2007) Drugs and valvular heart disease. *N. Engl. J. Med.* 356, 6–9
37 Papoian, T. *et al.* (2017) Utility of in vitro secondary pharmacology data to assess risk of drug-induced valvular heart disease in humans: regulatory considerations. *Toxicol. Pathol.* 45, 381–388
38 Smith, S.A. *et al.* (2009) Role of serotoninergic pathways in drug-induced valvular heart disease and diagnostic features by echocardiography. *J. Am. Soc. Echocardiogr.* 22, 883–889
39 Hutcheson, J.D. *et al.* (2011) Serotonin receptors and heart valve disease—it was meant 2B. *Pharmacol. Ther.* 132, 146–157
40 Boguszewski, C.L. *et al.* (2012) A comparison of cabergoline and bromocriptine on the risk of valvular heart disease in patients with prolactinomas. *Pituitary* 15, 44–49
41 Roth, B.L. (2007) Drugs and valvular heart disease. *N. Engl. J. Med.* 356, 6–9
42 Iverson, S.L. and Smith, D.A., eds (2016) *Metabolite Safety in Drug Development*, John Wiley & Sons
43 Tan, L.C.S. *et al.* (2009) Bromocriptine use and the risk of valvular heart disease. *Mov. Disord.* 24, 344–349
44 Elangbam, C.S. (2010) Drug-induced valvulopathy: an update. *Toxicol Pathol.* 38, 837–848
45 Huang, X.-P. *et al.* (2009) Parallel functional activity profiling reveals valvulopathogens are potent 5-hydroxytryptamine(2B) receptor agonists: implications for drug safety assessment. *Mol. Pharmacol.* 76, 710–722
46 Fitzgerald, L.W. *et al.* (2000) Possible role of valvular serotonin 5-HT(2B) receptors in the cardiopathy associated with fenfluramine. *Mol. Pharmacol.* 57, 75–81
47 Rothman, R.B. *et al.* (2000) Evidence for possible involvement of 5-HT(2B) receptors in the cardiac valvulopathy associated with fenfluramine and other serotonergic medications. *Circulation* 102, 2836–2841
48 Roth, B.L. *et al.* (2000) The multiplicity of serotonin receptors: uselessly diverse molecules or an embarrassment of riches? *Neuroscientist* 6, 252–262
49 Muller, P.Y. and Milton, M.N. (2012) The determination and interpretation of the therapeutic index in drug development. *Nat. Rev. Drug Discov.* 11, 751–761
50 WHO Collaborating Centre for Drug Statistics Methodology (2017) *Guidelines for ATC Classification and DDD Assignment.* WHO
51 O'Hagan, S. and Kell, D.B. (2017) Analysis of drug-endogenous human metabolite similarities in terms of their maximum common substructures. *J. Cheminformatics* 9, 18
52 O'Hagan, S. and Kell, D.B. (2015) Understanding the foundations of the structural similarities between marketed drugs and endogenous human metabolites. *Drug Metab. Transp.* 6, 105
53 Hert, J. *et al.* (2009) Quantifying biogenic bias in screening libraries. *Nat. Chem. Biol.* 5, 479–483
54 Gupta, S. and Aires-de-Sousa, J. (2007) Comparing the chemical spaces of metabolites and available chemicals: models of metabolite-likeness. *Mol. Divers.* 11, 23–36
55 Berman, H.M. *et al.* (2000) The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242
56 Bennett, B.D. *et al.* (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. Biol.* 5, 593–599
57 Copeland, R.A. and Boriack-Sjodin, P.A. (2018) The elements of translational chemical biology. *Cell Chem. Biol.* 25, 128–134
58 Hann, M.M. and Simpson, G.L. (2014) Intracellular drug concentration and disposition – the missing link? *Methods* 68, 283–285
59 Dowden, H. and Munro, J. (2019) Trends in clinical success rates and therapeutic focus. *Nat. Rev. Drug Discov.* Published online May 8, 2019 https://doi.org/10.1038/d41573-019-00074-z