ELSEVIER

In silico repositioning of approved drugs for rare and neglected diseases

Sean Ekins[1,2,3,4], Antony J. Williams[5], Matthew D. Krasowski[6] and Joel S. Freundlich[7]

[1] Collaborations in Chemistry, 601 Runnymede Avenue, Jenkintown, PA 19046, USA
[2] Collaborative Drug Discovery, 1633 Bayshore Highway, Suite 342, Burlingame, CA 94010, USA
[3] Department of Pharmaceutical Sciences, University of Maryland, College Park, MD 21201, USA
[4] Department of Pharmacology, University of Medicine & Dentistry of New Jersey, Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, NJ 08854, USA
[5] Royal Society of Chemistry, 904 Tamaras Circle, Wake Forest, NC 27587, USA
[6] Department of Pathology, University of Iowa Hospitals and Clinics, Iowa City, IA 52242, USA
[7] Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA

One approach to speed up drug discovery is to examine new uses for existing approved drugs, so-called 'drug repositioning' or 'drug repurposing', which has become increasingly popular in recent years. Analysis of the literature reveals many examples of US Food and Drug Administration-approved drugs that are active against multiple targets (also termed promiscuity) that can also be used to therapeutic advantage for repositioning for other neglected and rare diseases. Using proof-of-principle examples, we suggest here that with current *in silico* technologies and databases of the structures and biological activities of chemical compounds (drugs) and related data, as well as close integration with *in vitro* screening data, improved opportunities for drug repurposing will emerge for neglected or rare/orphan diseases.

Introduction

Neglected diseases are primarily tropical infections common in Africa, Asia and the Americas. Infections with *Mycobacterium tuberculosis* (Mtb) or *Plasmodium* spp. are often included as neglected diseases and are estimated to kill over two million people annually [1]. Recent studies also suggest that over two billion individuals are infected with Mtb alone [2] and this represents approximately one-third of the global population. These statistics highlight the enormous economic and healthcare challenges for the countries and governments affected.

There are also thousands of diseases that occur in small patient populations and are not addressed by any existing treatments (http://rarediseases.info.nih.gov/Resources/Rare_Diseases_Information.aspx). These diseases are classified as rare or orphan diseases. Traditionally, such diseases have not been the focus of big pharmaceutical company research as they have small patient populations in industrialized countries that make it difficult to market drugs that recoup

Sean Ekins

Sean Ekins is Principal Consultant for Collaborations in Chemistry and Collaborations Director at Collaborative Drug Discovery Inc. He has written over 160 papers and book chapters on topics including drug–drug interaction screening, computational absorption, distribution, metabolism, and excretion (ADME)/Tox, collaborative computational technologies and neglected disease research. Dr Ekins graduated from the University of Aberdeen, where he received his MSc, PhD and DSc.

Antony J. Williams

Antony J. Williams is Vice President, Strategic development for ChemSpider at the Royal Society of Chemistry. Dr Williams has written chapters for many books and has authored or peer reviewed over 100 papers and book chapters on nuclear magnetic resonance (NMR), predictive ADME methods, internet-based tools, crowd-sourcing and database curation. He is an active blogger and participant in the internet chemistry network. Dr Williams graduated with a PhD in chemistry as an NMR spectroscopist.

Matthew D. Krasowski

Matthew D. Krasowski is a board-certified clinical pathologist and currently is Clinical Assistant Professor and Assistant Director of Clinical Laboratories at the University of Iowa Hospitals and Clinics. His research interests include nuclear hormone receptor regulation of drug metabolism and computational approaches to modeling receptors and immunoassay cross-reactivity. He graduated with an MD and PhD in Neurobiology from the University of Chicago.

Joel S. Freundlich

Joel S. Freundlich is a senior research scientist in Biochemistry and Biophysics at Texas A&M University and a visiting professor in Medicinal Chemistry at Rutgers University. His research interests focus on seeding tuberculosis and malaria drug discovery through the study of small molecules that modulate essential biochemical targets. He received his PhD in organic chemistry from the Massachusetts Institute of Technology and then spent ten years in the pharmaceutical industry before returning to academia.

Corresponding author. Ekins, S. (ekinssean@yahoo.com), (sekins@collaborativedrug.com)

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2011.02.016

the cost of research and development and that are then profitable over the long term. Consequently, drug discovery for neglected and rare diseases has occurred mainly in biotech companies and academia. Rare diseases usually have small patient populations, although there is no global agreement on what this size is. In the USA, a rare disease is described as one that affects less than 200 000 people. Some estimates suggest that this represents over 7000 rare diseases affecting 25–30 million people [3] or 5000 patients per orphan disease, with approximately 4000 orphan diseases needing treatment [4]. Such a 'small' market size would make drugs for these diseases less marketable compared with common diseases, such as cancers, cardiovascular disease and diabetes, with sufferers treated numbering in the millions annually. However, some have suggested that profits can be made on smaller patient populations in a personalized medicine strategy and have called for more academia–pharma collaborations that are focused on rare diseases [4].

There are considerable challenges with regards to clinical research applied to rare diseases. Even though over 300 orphan drugs have been approved since the passage of the US Orphan Drugs Act in 1983, there is still a long way to go until most rare diseases have a treatment [3,4].

Neglected and rare diseases as an attractive area for pharmaceutical companies

Pharmaceutical companies are beginning to view rare or neglected diseases as an opportunity to bring in more revenue as well as to improve public relations. Developing treatments for rare or orphan diseases might necessitate a smaller investment upfront as, for example, in-licensing deals for an advanced therapeutic candidate targeting this area are usually less costly than the typical US$100s of millions for licensing drugs for other diseases (http://www.crdnetwork.org/blog/big-pharma-moves-from-blockbusters-to-niche–busters/). Recently, GlaxoSmithKline (GSK) made some relatively small investments in rare diseases (http://cenblog.org/the-haystack/2010/10/gsk-highlights-rare-diseases–approach/); Pfizer (http://www.xconomy.com/boston/2010/09/01/pfizer-gobbles-foldrx-in-big-pharmas-latest-rare-disease-play-in-boston–area/) and several other large pharma companies, as well as the World Health Organization, have been working together, investing US$150 million in research into neglected disease treatments (http://thebigredbiotechblog.typepad.com/the-big-red-biotech-blog/2010/10/big-pharma-and-governments-put-up-150-m-to-fight-neglected–diseases.html).

These efforts might only be the tip of the iceberg, and more substantial investments are likely to follow in the near future to solidify the trend. These investments by pharma for rare diseases are in addition to their significant investments in neglected or tropical diseases represented by the GSK Tres Cantos facility (http://www.gsk.com/collaborations/tres-cantos.htm), the Novartis Institute for Tropical Diseases in Singapore (http://www.novartis.com/research/nitd/index.shtml), the Lilly MDR-TB Partnership (http://www.lillymdr-tb.com/), the Lilly TB Drug Discovery Initiative (http://www.tbdrugdiscovery.org/) and The Critical Path to TB Drug Regimens (http://www.tballiance.org/cptr/).

Drug repositioning

One approach to speeding up drug discovery is to find new uses for existing approved drugs. This is termed 'drug repositioning' or 'drug repurposing', and traditionally has occurred by serendipity [5]. Another strategy is to look at combinations of approved drugs in the hope of finding synergy [6,7], an approach that has found some success in cancer, HIV and Mtb treatments. In the neglected and rare disease space, predominantly academic researchers have looked at repositioning compounds that are already approved for other indications (see references in Table 1 and Table 2). Drug repositioning has been reviewed extensively in the context of finding uses for drugs applied to major diseases, such as obesity and Parkinson's disease [4]. Well-known examples include drugs such as thalidomide, sildenafil, bupropion and fluoxetine, which found new uses beyond their initially approved therapeutic indications [5]. The example of thalidomide specifically suggests that drugs that were originally withdrawn by manufacturers or removed by the US Food and Drug Administration (FDA), or other regulatory organizations, can be resurrected. Thalidomide was notorious for causing birth defects if taken during the first trimester of pregnancy. However, this adverse effect is not a major issue in the novel use of thalidomide in treating multiple myeloma, a disease that is not common in women of child-bearing age.

Benefits for pharma

For pharmaceutical companies, repositioning has significant commercial value as it extends the markets for a compound and finds new uses for shelved compounds at lower financial risk and in a shorter time [8]. There has also been much discussion about how different approaches to repositioning could work, but these have not focused specifically on neglected diseases [5,9]. Others have proposed that repurposing could be an invaluable tool for neglected diseases [10]. The benefits of repositioning include: working on known druggable targets, the availability of materials and data (such as on long-term toxicology studies) that can be used and presented to regulatory authorities; and, as a result, the potential for a significantly more time- and cost-effective research and development effort than typically seen with bringing a new molecular entity to market.

Repositioning for neglected infectious diseases

In both the major-market and neglected infectious disease realms, the rapid emergence of multidrug-resistant strains of pathogenic microorganisms provides a sense of urgency to identify new scaffolds for antibiotics quickly. This is likely to require the exploration of chemical space beyond known active antimicrobial compounds. Pharma urgently needs new hits to initiate compound optimization studies. However, productivity of novel antibiotic classes over the past 30–40 years has been extremely low and this is exacerbated by the relatively low hit rates from high-throughput screening (HTS) and secondary screens [11]. Several new scaffold search efforts have been recently reviewed [12]. For example, Pfizer has shown that pyridopyrimidine compounds derived from a eukaryotic protein kinase inhibitor pharmacophore were effective against gram-negative pathogens following whole-cell screening [13]. The approach is an example of screening library repurposing (counterbalancing the pessimism derived from recently reported antibacterial-targeted screening efforts [11]) and illustrates the pursuit of bacterial targets with high sequence or structural similarity to eukaryotic targets, in this case the bacterial and eukaryotic kinomes. The Pfizer researchers proposed

TABLE 1

Examples of approved drug molecules identified using low-throughput screening methods as having effects against diseases other than the original target[a]

Molecule	Original use	New use	Method of discovery	Refs
Aprepitant	Nausea: NK-1 receptor antagonist	Drug-resistant HIV-1 infection: downregulates CCR5 in macrophages	Initial hypothesis tested with another NK-1 receptor antagonist *in vitro*	[99,100]
		Cryptosporidiosis in immunosuppressed hosts	Tested *in vivo* in immunosuppressed mice infected with *Cryptosporidium parvum*; decreased substance P levels	[101]
Amiodarone	Class III anti-arrhythmic	Chagas disease: blocks ergosterol biosynthesis	Literature search	[102]
Glybenclamide	Antidiabetic	Antithrombotic activity in mouse models IC_{50} 9.6 μM	Common pharmacophore with an experimental TP receptor antagonist SQ29,548	[103]
Tamoxifen	Antiestrogen	Anti-protozoal: *Leishmania amazonensis* IC_{50} 11.1–16.4 μM	Focused screening to test hypothesis and *in vivo* mice studies	[104,105]
Trimetrexate	Antifolate used in *Pneumocystis carinii* infection in patients with AIDS	Inhibitor of *Trypanosoma cruzi* DHFR IC_{50} 6.6 nM	Enzyme activity and antiparasite activity assays for one compound	[106]
Riluzole	Amyotrophic lateral sclerosis: inhibits glutamate release and reuptake	Currently in clinical trials for treating melanoma, but might have activity against other cancers	Treatment of GRM1-positive human melanoma cells reduced levels of released glutamate, suppressed melanoma cell growth and also suppressed tumor growth in xenograft model; induced cell cycle arrest, leading to apoptosis	[107]
Sertraline	Antidepressant (selective serotonin reuptake inhibitor)	Neuroprotective, prolongs survival, improves motor performance and ameliorates brain atrophy in the R6/2 HD model	Previously shown that another SSRI was neuroprotective	[108]

[a] Abbreviations: CCR5, chemokine receptor 5; DHFR, dihydrofolate reductase; GRM1, glutamate receptor, metabotropic 1; NK-1, neurokinin-1 receptor; SSRI, selective serotonin reuptake inhibitors.

that targets with high sequence and structural homology to known human drug targets are more likely to find inhibitors in the compound libraries. Others have suggested that the libraries of inhibitors for ion channel and prenyltransferases would be a good starting point for such library repurposing [14] and for finding chemotypes for novel antimicrobials.

It is unclear how extensively approved drugs are screened against multidrug-resistant strains of bacteria and it might be possible to find new acceptable treatments among them. Clearly, more could be done to reposition existing FDA-approved drugs, and the following sections survey these efforts to find new activities. To date, these studies have traditionally focused on *in vitro* screening; however, computational screening ('*in silico*' [15]) methods might also be applicable. Hence, it is proposed that a combined *in silico–in vitro* approach leveraging databases of molecular structures and their related information from the literature [such as absorption, distribution, metabolism, and excretion (ADME)/Tox [16], targets, clinical trials, etc.] could be a viable strategy for accelerating research in the treatment and prevention of rare, neglected and common diseases.

Searching FDA-approved drugs for new activities

Using HTS

It is suggested that there are over 10 000 drugs that have been tested in clinical medicine. This could be reduced to approximately 9000, given that many represent combinations of other drugs, different salt forms of the same molecule, or biologics (large proteins or antibodies) [17]. However, a physical library of this size does not exist for known drugs that could be screened and a virtual library of these compounds is also lacking (to our knowledge). Such a virtual library could be assembled using some of the public domain databases.

Some companies, such as Cerep (http://www.cerep.fr), have screened 2500 of the FDA drugs and reference compounds against 159 enzymes, receptors, ion channels and transporters, and have created a database called BioPrint [18], which is a commercial product with a cost that is likely to be out of reach of most academic researchers. To date, multiple groups have screened 1000–2000 drugs against different targets or cell types. The John Hopkins Clinical Compound Library (JHCCL) consists of plated compounds available for screening at a relatively small charge and has been used by some groups [19]. For example, 17 novel inhibitors of Mtb were found after screening 1514 compounds from the JHCCL [19]. Several new uses for FDA-approved drugs have been identified by screening this or other commercially available libraries of drugs or off-patent molecules (e.g. the Microsource US Drug Collection and Prestwick Chemical library) (Table 2). The accumulation of large databases of published data and compounds screened against G-protein coupled receptors (GPCRs), such as the psychoactive drug screening program (PDSP) receptorome profile, represent good starting points for finding compounds that are active against receptors of interest. One example described is a

TABLE 2

Examples of approved drug molecules identified using HTS or *in silico* screening methods as having effects against diseases other than original target[a]

Molecule	Original use	New use	Method of discovery	Refs
Itraconazole	Antifungal: lanosterol 14α-demethylase inhibitor	Inhibition of angiogenesis by inhibiting human lanosterol 14α-demethylase; IC_{50} 160 nM	*In vitro* HUVEC proliferation screen against FDA-approved drugs (JHCCL)	[109]
Astemizole	Non-sedating antihistamine (removed from US market by FDA in 1999)	Antimalarial IC_{50} 227 nM against *Plasmodium falciparum* 3D7	*In vitro* screen for *P. falciparum* growth of 1937 FDA-approved drugs (JHCCL)	[110]
Mycophenolic acid	Immunosuppressive drug: inhibits guanine nucleotide biosynthesis	Inhibition of angiogenesis by targeting type 1 inosine monophosphate dehydrogenase; IC_{50} 99.2 nM	*In vitro* HUVEC proliferation screen of 2450 FDA- and foreign-approved drugs (JHCCL)	[111]
Entacapone and tolcapone	Parkinson's Disease: catechol-*O*-methyltransferase inhibitors	Antitubercular: entacapone inhibits InhA; IC_{50} 80 μM	Used a chemical systems biology approach	[77]
Nitazoxanide	Infections caused by *Giardia* and *Cryptosporidium* spp.	Antitubercular: multiple potential targets	Screens against replicating and non-replicating Mtb	[112]
(±)-2-amino-3-phosphonopropionic acid	Human metabolite, mGluR agonist	Antimalarial: inhibits HSP-90; IC_{50} 0.06 μM against *P. falciparum* 3D7	HTS screening of 4000 compounds	[113]
Acrisorcin	Antifungal	Antimalarial: inhibits HSP-90; IC_{50} 0.05 μM against *P. falciparum* 3D7	HTS screening of 4000 compounds	[113]
Harmine	Anticancer	Antimalarial: inhibits HSP-90; IC_{50} 0.05 μM against *P. falciparum* 3D7	HTS screening of 4000 compounds	[113]
Acetophenazine, fluphenazine and periciazine	Antipsychotics–D2 and 5-HT_2 inhibitors	Human androgen receptor antagonists acetophenazine (K_i 0.8 μM), fluphenazine(K_i 0.8 μM), periciazine (K_i 3.0 μM)	Docking of known drugs into androgen receptor followed by *in vitro* screening	[96]
Levofloxacin, gatifloxacin, sarafloxacin, moxifloxacin and gemifloxacin	DNA gyrase	Active against ATCC17978; inactive against BAA-1605 MIC ≤0.03–0.04 (mg/l)	Screening of 1040 drugs from microsource drugs library versus *Acinetobacter baumannii*	[114]
Bithional, bortezomib, cantharidin, chromomycin A3, duanorubicin, digitoxin, ectinascidin 743, emetine, fluorosalen, manidipine HCl, narasin, lestaurtinib, ouabain, sorafenib tosylate, sunitinib malate, tioconazole, tribromsalen, triclabendazolum and zafirlukast	Various	NF-κB inhibitors; IC_{50} 0.02–39.8 μM	Screening of NCGC pharmaceutical collection of 2816 small molecules *in vitro*	[115]
Pyrvinium pamoate	Anthelmintic	Antitubercular: Alamar blue assay MIC 0.31 μM	*In vitro* screen against 1514 known drugs; many other previously unidentified hits found	[19]
		Anti-protozoal: *Cryptosporidium parvum* IC_{50} 354 nM	*In vitro* screen for *P. falciparum* growth of 1937 FDA-approved drugs hypothesized to be active because they are confined to intestinal epithelium	[116]
		Anti-protozoal: against *Trypanosoma brucei*;IC_{50} 3 μM	Screening of 2160 FDA-approved drugs and natural products from Microsource; 15 other drugs active; IC_{50} 0.2–3.0 μM	[117]
Riluzole	ALS: inhibits glutamate release and reuptake	Enhanced Wnt/β-catenin signaling in both the primary screen in HT22 neuronal cells and in adult hippocampal progenitor cells; GRM1 regulates Wnt/β-catenin signaling	Screening of 1857 compounds (1500 unique) *in vitro*; treating melanoma cells with riluzole *in vitro* enhanced the ability of WNT3A to regulate gene expression	[118]
Closantel	A veterinary anthelmintic with known proton ionophore activities	Onchocerciasis (river blindness); IC_{50} 1.6 μm; K_i 468 nM	Screening of 1514 FDA-approved drugs (JHCCL) against the chitinase OvCHT1 from *Onchocerca volvulus*	[119]

TABLE 2 (*Continued*)

Molecule	Original use	New use	Method of discovery	Refs
Nitroxoline	Antibiotic used outside USA for urinary tract infections	Antiangiogenic agent inhibits MetAP2 (IC_{50} 54.8 nM) and HUVEC proliferation; also inhibits sirtuin 1 (IC_{50} 20.2 μM) and sirtuin 2 (IC_{50} 15.5 μM)	Screening of 2687 FDA-approved drugs (JHCCL) for inhibition of HUVEC cells; also found the same compound in HTS of 175 000 compounds screened against MetAP2; active in mouse and human tumor growth models	[120]
Glafenine	Analgesic	Inhibits ABCG2 (IC_{50} 3.2 μM); could be used with chemotherapeutic agents to counteract tumor resistance	Screening of FDA-approved drugs (JHCCL) with bioluminescence imaging HTS assay; discovered 37 previously unknown ABCG2 inhibitors	[121]
Tiagabine	Antiepileptic (enhances gamma-aminobutyric acid activity)	Neuroprotective in N171-82Q and R6/2 mouse models of HD	Initial screen of NINDS Microsource database of drugs (1040 molecules) against PC12 cell model of HD found nipecotic acid, which is related to tiagabine	[122]
Digoxin, oubain and proscillardin A	Cardiac glycosides used to treat congestive heart failure and arrhythmia	Anticancer: inhibition of hypoxia-inducible factor 1; $IC_{50} \leq 400$ nM	3120 FDA-approved drugs (JHCCL) screened against reporter cell line Hep3B-c1; digoxin also tested in *in vivo* xenograft models	[123]
Tacrine, carvedilol, hexamethyleneamiloride and phenoxybenzamine	Acetylcholinesterase inhibitor, β_2-adrenergic blocker, diuretic, α_1-adrenergic blocker, respectively	Prevention of hearing loss: lowest dose tested that shows protection is 10 μM	Initial screen of NINDS Microsource database of drugs (1040 molecules) against neomycin induced hair cells in zebrafish. Tacrine was also active in mouse utricle	[124]
Ceftriaxone	β-lactam antibiotic	Neuroprotection. ALS: increases GLT1 expression; EC_{50} 3.5 μM. Other β-lactams also active	Screen of NINDS Microsource database of drugs (1040 molecules) against rat spinal cord cultures followed by immunoblot for GLT1 protein expression. Also tested in ALS mouse model: delayed neuron loss and increased survival	
Flufenamic acid	Non-steroidal anti-inflammatory drug	Familial amyloid polyneuropathy: inhibits transthyretin	Screening library not described	[125]

[a] Abbreviations: ABCG2, ATP-binding cassette sub-family G member 2; ALS, amyotrophic lateral sclerosis; GLT1, glutamate transporter 1; GRM1, metabotropic glutamate receptor; HSP-900, heat shock protein 90; HD, Huntingdon's disease; HUVEC, human umbilical vein endothelial cells; InHa, inhibin, alpha; MetAP2, type 2 methionine aminopeptidase; mGluR, metabotropic glutamate receptor; NCGC, National Clinical Guideline Centre; NF, nuclear factor; NINDS, National Institute of Neurological Disorders and Stroke.

potent 5-HT_{2A} ligand that could block the JC virus [20], which can cause the neurologic disease progressive multifocal leukoencephalopathy if untreated. A second example suggested side effects for known drugs mediated by the 5-HT_{2B} receptor [20]. The number of examples of groups finding new uses for approved drugs by HTS appears to be growing (Table 2) on a laboratory-by-laboratory basis. It is intriguing to ponder whether an organized effort to screen experimentally the set of all known drugs against all known targets validated for a given disease would be feasible. Certainly, the potential for success with one disease, let alone many human diseases, appears to be significant.

Using in silico *methods*

In silico methods, including target- and ligand-based strategies, are an excellent complement to experimental techniques, and are widely used in industry and academia [15,21]. There have been many studies establishing relationships between ligand molecular structures and broad biological activities, both on and off target [22–25]. Several examples using pharmacophore-based studies and searching databases of FDA drugs [26] to find new transporter inhibitors *in vitro*, represent attempts at understanding off-target effects, which is analogous to drug repositioning. For example, pharmacophores for various transporters, such as the human peptide transporter 1 (hPEPT1) [27], P-glycoprotein (P-gp) [28], the human organic cation/carnitine transporter (hOCTN2) [29,30] (Fig. 1) and the human apical sodium-dependent bile acid transporter (ASBT) [31], have been used to search a subset of FDA-approved drugs compiled from *A Small Physician's Handbook* (SCUT, structures available as a supplemental file) [26] and to identify previously unknown inhibitors based on *in vitro* testing (Table 3). Interestingly, for each transporter, inhibitors were found that belonged to different therapeutic classes and these represented molecules with overlapping pharmacophores. What has not been examined to date is whether the distinct therapeutic class hits for a single transporter are also shared by other common biological activities. These transporters were selected because of the inhibition of hPEPT1 or P-gp involved in drug–drug interactions [28], the putative role of hOCTN2 inhibition in rhabdomyolysis [29,30] and the potential for drugs inhibiting ASBT to promote several adverse drug reactions (ADRs), including colon

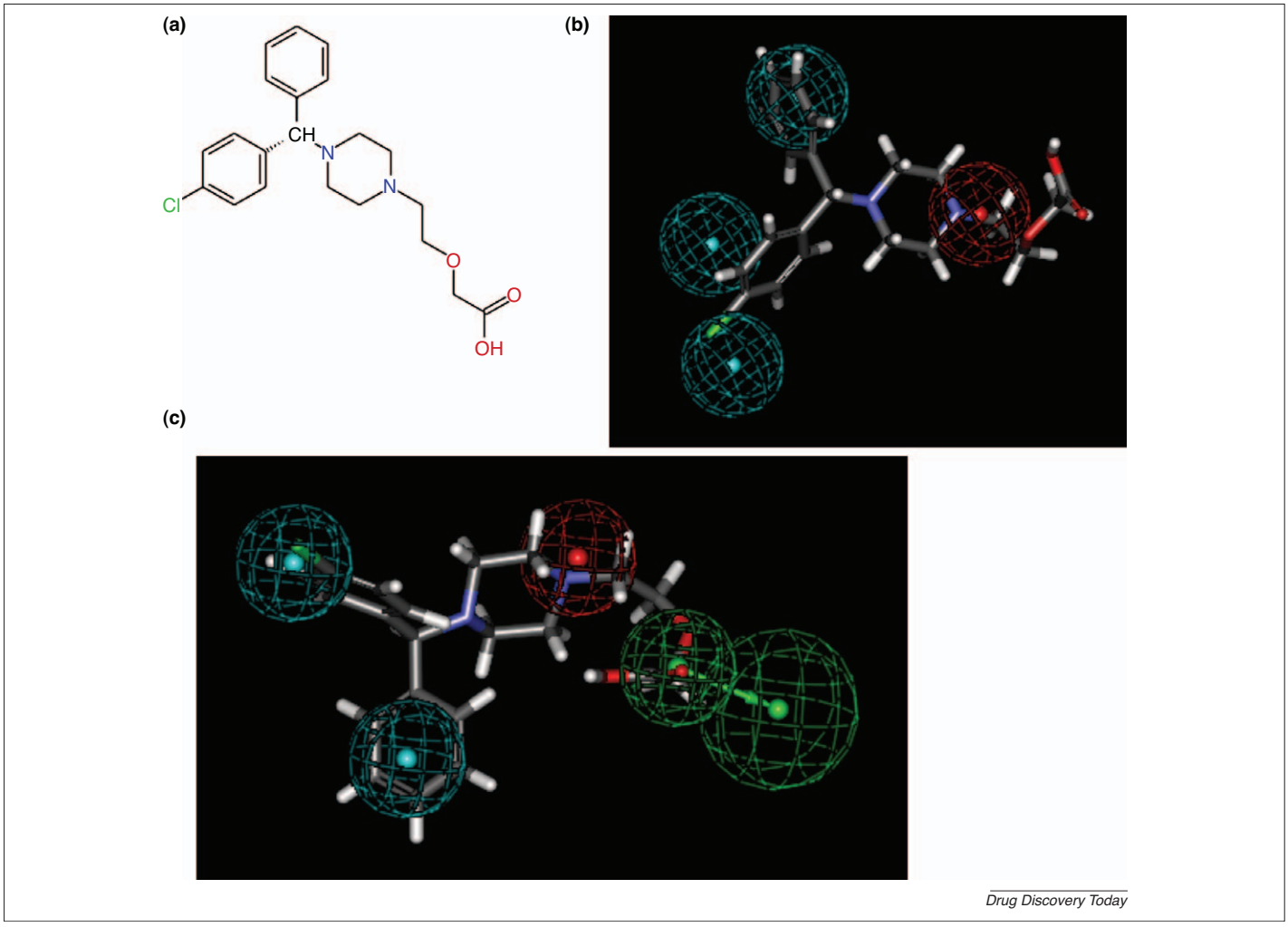

FIGURE 1

Transporter pharmacophores for hOCTN2. **(a)** Cetirizine. **(b)** Cetirizine mapped to a catalyst pharmacophore based on three actives and two inactives for OCTN2 [30]. The pharmacophore contains two hydrophobic features (cyan) and a positive ionizable feature (red). **(c)** Cetirizine mapping to a catalyst pharmacophore derived from 22 drugs with K_i data for OCTN2 (observed and predicted data described in [29]). The pharmacophore contains two hydrophobic features (cyan), a hydrogen-bond acceptor (green) and a positive ionizable feature (red).

TABLE 3

FDA-approved compounds found by an *in silico–in vitro* approach to inhibit transporters[a]

Compounds	Transporter	Biological effect	Pharmacophore features	Refs
Aspartame, fluvastatin and repaglinide	hPEPT1	Inhibit uptake of natural substrates and other drugs that are substrates	Two hydrophobic and one hydrogen bond acceptor; one hydrogen bond donor; one negative ionizable feature	[27]
Acitretin, cholecalciferol, misoprostol, nafcillin, repaglinide, salmeterol and telmisartan	P-gp	Decrease clearance of drugs by inhibiting efflux into intestine of P-gp substrates.	Three hydrophobic features and two hydrogen bond acceptor features	[28]
Thioridazine, vinblastine, clozapine, amlodipine, gefitinib, trifluoperazine, dibucaine, tamoxifen, amiodarone, atracurium, nefazodone, argatroban, nelfinavir, proclorperazine, raloxifene, metoclopramide, desloratidine, duloxetine, carvedilol, olanzapine, amitriptyline, imatinib, desipramine, quinine, quinidine, haloperidol and bromocriptine	OCTN2	Inhibition may cause rhabdomyolysis.	Three hydrophobic features and one positive ionizable feature Two hydrophobic features, one positive ionizable feature and one hydrogen bond acceptor.	[30] [29]
Nimodipine, fluvastatin, latanoprost, lovastatin, pentamidine, simvastatin, pioglitazone and tioconazole	ASBT	ASBT inhibition can cause diarrhea, hyperglyceridemia, gallstone disease and colon cancer.	Two hydrophobic features, two hydrogen bond acceptors and shape restriction around mesoridazine	[31]

[a] Additional examples of transporter pharmacophore searches can be found in [126].

cancer [31]. The transporters also represent a class of proteins for which *in vitro* models might be limited in throughput and where *in vivo* study is even more complicated owing to the presence of multiple transporters with overlapping substrate specificities. Therefore, the *in silico–in vitro* approach has value in targeting compounds with a high probability of activity.

Computational pharmacophores and molecular similarity methods for drug repositioning

Pharmacophores and 3D database searching could be readily used for drug repositioning. 2D approaches might, however, be more readily available for both similarity and substructure searching and have been used with success for finding metabolite mimics for Mtb [32] and in studies to predict the cross-reactivity of drugs and drug metabolites with immunoassays used in clinical medicine [33–36]. Common applications of immunoassays include drug of abuse (DOA) screening, endocrinology testing and therapeutic drug monitoring (TDM). Immunoassays can be limited by the occurrence of false positives (or 'cross-reactive' compounds). For example, drugs with structural similarity to amphetamine and methamphetamine, such as ephedrine and pseudoephedrine, can cross-react with DOA screening assays designed to detect the presence of amphetamine or methamphetamine. Diagnostic companies manufacturing clinically used immunoassays often test a limited number of compounds for cross-reactivity against their immunoassay, although there is a potentially large array of compounds (metabolites, herbals and environmental chemicals) that could possibly interact. Consequently, cross-reacting compounds are discovered on a case-by case basis [33,34].

Similarity searching examples

Computational 2D similarity (using the MDL public keys fingerprint descriptors) of test compounds to that of the antigen used in immunoassays, has been used to predict cross reactivity [33–36]. The SCUT database of frequently used FDA-approved drugs was used for similarity searching and was supplemented with some metabolites of drugs (see Online Supplementary Information). This relatively simple computational approach showed a statistically significant separation between cross-reactive and non-cross-reactive compounds for TDM immunoassays [33,34] and DOA/toxicology immunoassays [35,36]; the approach was further used to identify novel inhibitors of DOA/toxicology immunoassays [36]. These examples show how *in silico* methods can build on existing data and focus *in vitro* testing.

The examples above also illustrate how 2D similarity alone might be useful for finding compounds that could have pharmacophore features that are similar to those of other drugs. This raises the question of whether such similar molecules might share overlapping biological activities. Simple similarity searching could be a component of a compound-repositioning strategy, using computational methods to predict probable cross-reactive compounds by similarity followed by a quick confirmation with immunoassays that are commercially available. Other computational approaches comparable to searching by similarity, such as those involving LASSO descriptors [37], can make use of large, publicly available, databases, such as ChemSpider [38], to compare existing drugs with virtual libraries. Comparable methods, such as PASS (prediction of activity spectrum for substances) could also be used to predict potential new bioactivities for existing drugs [39]. Computational methods that account for molecular shape might be generally useful for searching for compounds with common bioactivity [40]. Molecular docking is one example that has been used successfully to find molecules with complementary shape and electrostatic interactions with known protein active sites. For example, docking approaches have been used to dock 1055 known drugs (from DrugBank) into 78 unique human drug targets and the authors found 52 interactions of interest (although no experimental verification was reported) [41].

Using networks and systems biology for drug repositioning

During the past decade, understanding of biological mechanisms has been significantly enhanced by the curation of vast ligand– and protein–protein interaction databases and the use of top-down and bottom-up network modeling leading to a systems biology approach [42–46]. During the past five years alone, 2D ligand-based approaches have been increasingly used along with sophisticated algorithms and networks. This approach has been used for drug repositioning and for understanding the off-target effects of drugs. Fliri *et al.* used biological spectra for a cross section of the proteome [47]. They implemented hierarchical clustering of the biological activity spectra similarity and created a relationship between structure and bioactivity before extending this to identify receptor agonist and antagonist profiles [48]. The same group from Pfizer took this concept further and applied a probabilistic approach to link adverse effects for drugs with biological spectra (similar molecules had overlapping profiles, in the same way that they had similar biological spectra), thus linking preclinical with clinical effects [49].

Promiscuity networks and insights for repositioning

There have been many efforts to look at compound or protein promiscuity or polypharmacology that could lead to the discovery of new uses for existing molecules. Specifically, there has been considerable discussion of predicting undesirable drug interactions with promiscuous proteins *in silico*. This is a particular issue for hydrophobic compounds that might bind to cytochrome P450 (CYP) 3A4, the pregnane X receptor (PXR), P-gp or the human ether-à-go-go-related gene (hERG) [50]. Quantitative structure–activity relationship (QSAR) models for these proteins have been used to predict potential molecule protein interactions and then visualize this as a node on a network, simultaneously showing other endogenous and exogenous ligand–protein interactions [45,46,51] as well as allowing overlay of any gene expression or other high content data [52–54]. Such an approach could be useful for ensuring that repurposed compounds do not have negative effects on biological networks through binding other off-targets. A global mapping of pharmacological space focusing on a polypharmacology network of 200 000 molecules with activity against 698 proteins has also been produced [55]. A further published study created a drug–target network of approved drug data for 890 molecules from DrugBank [56] and OMIM (http://www.ncbi.nlm.nih.gov/omim), with over half of these molecules forming the largest connected network with multiple target proteins (also illustrating polypharmacology or promiscuity) [57]. Such networks might help understand probable nodes involved in toxicity

and add to the similarity maps for enzymes and receptors [58] and human polypharmacology networks [55] that have also been developed to date. A recent study from Abbott introduced a sequence-independent kinome inhibitor interaction map [59], whereas another study established links between over 44 000 small molecules and 160 cardiovascular targets, with kinases having, on average, seven connections to cardiovascular targets [60]. An example from Berg *et al*. has merged chemical target and pathway toxicity mechanisms that can be defined from profiling in primary human cell systems covering many readouts and enabling known reference compounds to be mapped by functional similarity [61].

Using chemical substructures to understand side effects and assist repositioning

A complimentary approach taken by a group at Novartis uses chemical substructures relevant for toxicology-related adverse effects [62] for approximately 4000 molecules with side-effect information from the World Drug Index. The same group related promiscuity of compounds to their safety [63]: for a given compound, the number of biological targets inhibited to a significant extent typically correlates with a higher incidence of effects. More recently, the group has related over 4000 MedDRA (http://www.meddramsso.com/) terms for ADRs for over 1800 drugs using the ECFP_4 descriptors and Bayesian models [64,65]. This resulted in a map of ADRs in chemical space and an approach that could be used to predict, *in silico,* the ADR likelihood for new molecules based on substructures. Interestingly, the recent similarity ensemble analysis described by Keiser *et al*. also used the ECFP_4 descriptors and Bayesian models to predict off-target effects of 3665 FDA-approved drugs and investigational compounds [66]. This study clearly showed the promiscuity of many compounds. Their *in vitro* validation of the computational predictions focused on compounds with predicted GPCR activity other than the known targets. The approach could be particularly useful for understanding the potential targets for compounds where these have been previously unknown.

Using machine learning and databases for drug repositioning

Machine-learning models have also been applied with various types of literature data on drugs that could also assist in their repositioning. Decision tree induction has been used to predict the adverse drug reactions for 507 drugs from the Swiss Drugs Registry, and resulted in models that looked internally predictive [67]. A machine-learning method has also been used with a set of 390 drugs to demonstrate that anatomical therapeutic chemical classification, a system used for drug repurposing, can be predicted by using a binary feature vector derived from extraction of drug property data from text alone [68]. Chiang and Butte compiled a drug-disease knowledge base (DrDKB) to capture the 3517 FDA-approved drug indications and 8130 off-label uses of 2022 distinct drugs used to treat each of 726 diseases [69]. They were able to make 57 542 unique novel drug use suggestions and, leaving out 10–20% of the data as a test set, resulted in over 85% recovery of the drug uses [69]. Others have generated a database called PROMISCOUS (http://bioinformatics.charite.de/promiscuous/index.php?site=drugdev) representing a set of 25 000 withdrawn or experimental drugs annotated with 21 500 drug–protein and 104 000 protein–protein relationships, using public resources (e.g. DrugBank, SuperTarget, etc.) and text or data mining [70]. These data can be searched using a network visualization tool and several anecdotal examples were provided of molecule or side-effect similarity, although no prospective testing was described [70]. Another tool suggested to be useful for drug repositioning is IDMap, which integrates the Elsevier MDDR database, Asinex compounds, PASS and molecular descriptors from Cerius^2 [71]. Text mining was used to compare PASS and MDDR bioactivity and provide a co-occurrence frequency, although, again, no prospective testing was shown [71]

Integration of methods for repositioning

By connecting data on drugs, proteins and diseases, these various databases, networks and computational methods might be useful not only for understanding and identifying promiscuity, polypharmacology and toxicity mechanisms, but also potentially for repurposing molecules for new uses that could focus and accelerate *in vitro* screening efforts [17,20,72–74] as previously described with transporters [27–31]. For some researchers, finding molecules with manageable ADRs might be useful and lead to new indications. Many of these examples illustrate how molecules can be put into a biological context through networks. The integration of different computational and experimental approaches along with published data could lead to a more complete understanding than using a single approach in isolation and could enable network-based drug discovery described elsewhere [75,76]. Others have also suggested that data integration platforms for systems biology (whether using ligand [58,66] or binding site similarity [77]) could support repositioning and drug discovery, although no solid examples of bringing new treatments into the clinic have been provided as yet [78].

Examples using *in silico* methods for drug repositioning in neglected infectious diseases

As a proof of principle that computational methods could help accelerate neglected disease research, a machine-learning method has been used and validated with multiple data sets. Bayesian classifier models are computationally fast and have been used widely for several drug discovery applications in recent years, including with Mtb [79]. Bayesian classification methods [80] have been previously used for CYP, transporter and toxicity models [31,81–84] as well as to identify substructures that are important in recent TB screening data sets [85]. The Mtb Bayesian models (training sets from 2000 to $\geq$200 000 molecules) have been validated with external compounds using the published National Institute of Allergy and Infectious Diseases (NIAID), GVKbio data sets (which include known drugs and other experimental compounds) and a set of 102 000 compounds [Tuberculosis Antimicrobial Acquisition and Coordinating Facility (TAACF)-NIAID cannabinoid receptor 2 (CB2)] containing 1702 molecules with $\geq$90% inhibition at 10 μm (representing a hit rate of 1.66%) [86]. Tenfold enrichments were shown in finding active compounds in the top-ranked 600 molecules for the TAACF-NIAID CB2 [86], which came from the same source [84,87] as the training sets used in the original models and represents an ideal scenario from modeling to limit any experimental variability. The three test sets ranged from 2880 to over 102 000 compounds. The largest test set also contained a more realistic percentage of hits representative of HTS screens.

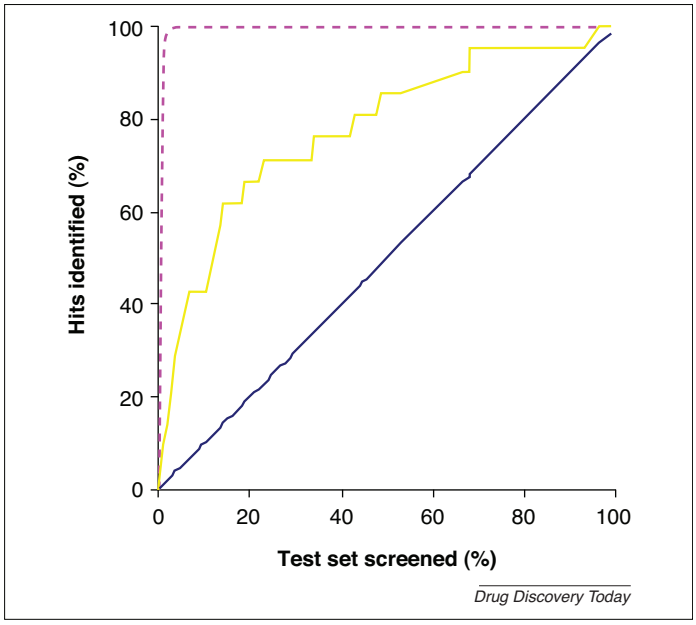

FIGURE 2

Receiver operator characteristic plot for the FDA-approved Mtb hits ($n = 21$) used as a test set ($n = 2108$) for a previously published Bayesian model [89]. Key: purple, best rate of finding hits; yellow, Bayesian model; blue, random rate of finding hits.

More recently, the JHCCL set of 1514 known drugs were used to screen experimentally against Mtb and the minimum inhibitory concentration (MIC) values determined using the Alamar blue susceptibility assay (published by others [19]). Of the actives identified, 21 were used as a test set in a larger set of 2108 FDA-approved molecules downloaded from the Collaborative Drug Discovery database (CDD) database. After removal of compounds that were also in the Bayesian models, it was shown that the Bayesian models initially had approximately tenfold enrichments. One model identified ≥60% of the drug hits in the top 14% of compounds (S. Ekins and J.S. Freundlich, in press; Fig. 2). The Bayesian models were also used to suggest drugs with a high probability of predicted Mtb activity that could be tested *in vitro* in future (S. Ekins and J.S. Freundlich, in press).

Resources for *in silico* repositioning of molecules for neglected and rare diseases

CDD

If researchers are going to accelerate rare/orphan and neglected disease research *in silico*, what resources are currently available and what are still needed? One accessible tool is the CDD database [88] with a focus on neglected diseases, which has been recently described in detail [86,89,90]. Chris Lipinski (Melior Discovery) provided a database of 1055 FDA-approved drugs with designated orphan indications, sponsor name and chemical structures. In addition, David Sullivan (Johns Hopkins University) collated and provided a database of 2815 FDA-approved drugs. Bryan Roth (University of North Carolina) provided the PDSP database, which currently consists of nearly 1500 molecules structures that have been screened against an array of GPCRs [20,58,73,74]. These data, in addition to the >20 screening data sets for malaria and TB (Table 4), have enabled recent analysis of the physicochemical properties of active compounds [86,91,92] and filtering with readily available substructure alerts or 'filters' [86,91,92]. All these data sets allow for free access of substructure, similarity or Boolean searches upon registration (e.g. http://www.collaborativedrug.com/register). The data have also been used for validating similarity searching and pharmacophore approaches to find mimics of essential metabolites for Mtb [32].

In addition, a license to CDD can enable download of data sets that are not freely available. This might be advantageous if they need to be searched with third-party cheminformatics software (e.g. pharmacophore models or QSAR methods, etc.) (Fig. 3). This suggests an additional approach for repurposing using *in silico*

TABLE 4

A subset of the ≥20 CDD publicly available antimalarial and TB data sets[a]

Database name/source	Description	Molecules
US Army survey	An extensive collection of antimalarial drug animal SAR data, including structures, bioactivity etc., published originally by the US Army in 1946	12 318
St Jude Children's Research Hospital	Supplemental data for [127]: structures tested in a primary screen, with additional data in eight protocols: Bland–Altman analysis, calculated ADMET properties, phylochemogenetic screen, sensitivity, synergy and enzyme assays, as well as a thermal melt analysis	1524
Novartis Malaria	Data from [128] *Plasmodium falciparum* strains 3D7 (drug-susceptible) and W2 (chloroquine-, quinine-, pyrimethamine-, cycloguanil- and sulfadoxine-resistant), obtained from MR4, were tested in an erythrocyte-based infection assay for susceptibility to inhibition of proliferation by selected compounds	5695
Johns Hopkins-Sullivan	Percent inhibition of approved drugs at 10 μM	2693
MLSMR	A diverse collection tested by the Southern Research Institute against Mtb H37Rv; the most active compounds have dose-response and cytotoxicity data	214 507
TB efficacy data from the literature	TB efficacy data from >300 published literature sources; data include PubMed citations, targets, cells and organisms tests, MIC, % inhibition, EC_{50}, IC_{50}, etc.	6771
TAACF–NIAID–CB2	Results of a commercial compound library screening by the Southern Research Institute to inhibit the growth of *Mycobacterium tuberculosis* strain H37Rv	102 634
Novartis Mtb	Aerobic and anaerobic hits versus *M. tuberculosis*	283

[a] Abbreviations: ADMET, absorption, distribution, metabolism, excretion, and toxicity; MR4, Malaria Research and Reference Reagent Resource Center.

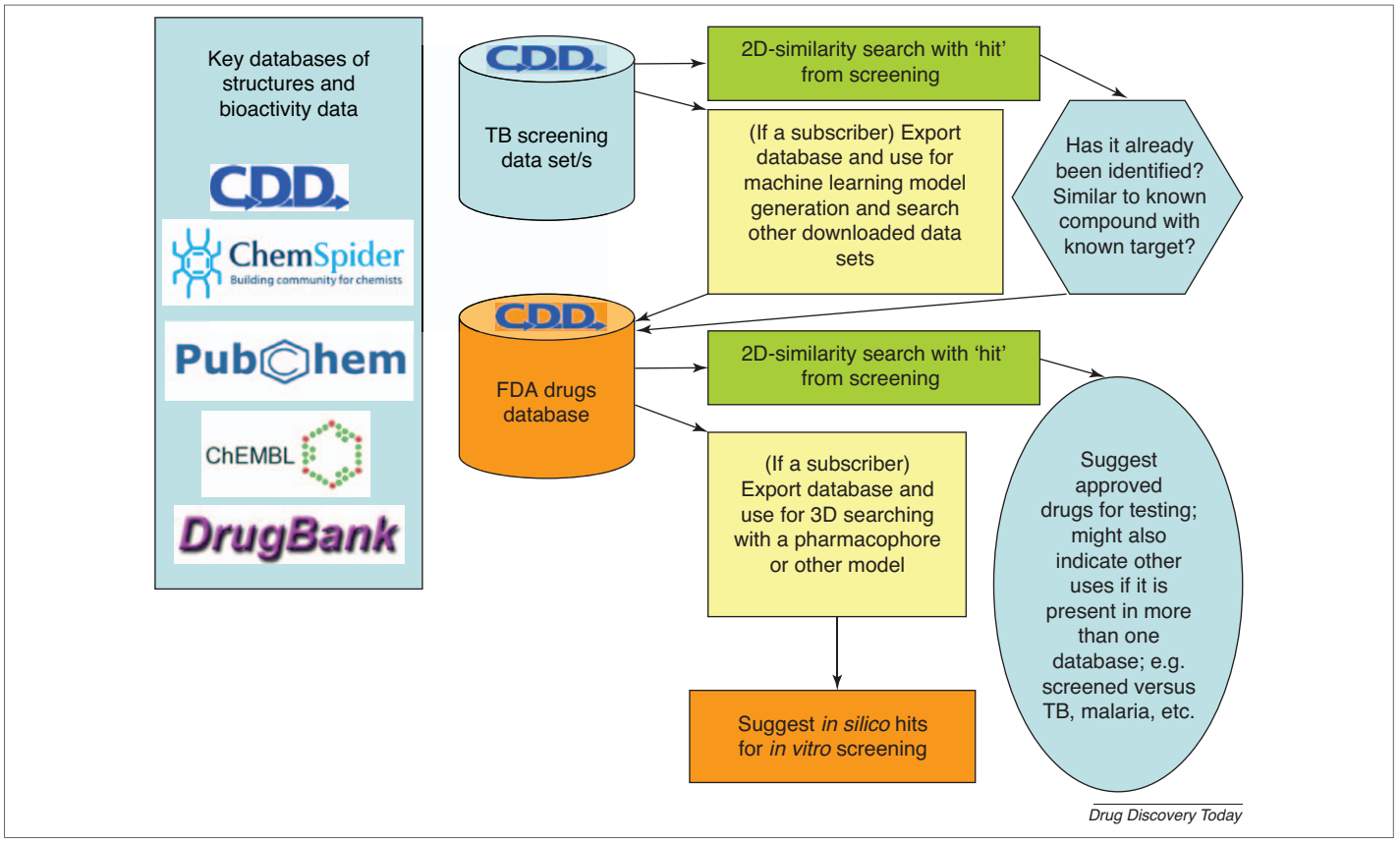

FIG. 3

A repositioning strategy using the CDD (http://www.collaborativedrug.com/register), ChemSpider (http://www.chemspider.com) or other databases (bioactivity data for target or disease of interest and FDA drug data set) in combination with computational methods (pharmacophore, similarity assessment, machine learning, etc.).

models to find compounds of interest in the FDA-approved drugs set. For instance, models generated with data from one or more public data sets (or the user's own private data) could be used to search other data sets and find new molecules for screening (Fig. 3).

Other tools

The same strategy described previously could be readily taken with other databases and software tools taking advantage of freely available content and tools in databases such as ChemSpider [38], PubChem [93], DrugBank [56] and ChEMBL (http://www.ebi.ac.uk/chembldb/index.php) or others [16,94,95]. This overall approach is analogous to the pharmacophore approach taken with transporters searching the SCUT database of commonly used drugs (Table 3), similarity searching for drugs cross-reactive with DOA and TDM immunoassays [33–36] and with Mtb Bayesian models [86,91] for searching the FDA-approved drugs. Recent efforts to validate the Bayesian models with data from other laboratories (described above) would indicate that the *in silico* approach certainly has merit for neglected diseases.

The missing piece

What is still needed is a single comprehensive resource that has validated chemical structures (and properties) of both FDA- and internationally approved drugs, as well as those that are either no longer used or are removed from the market. A database containing information on studies in which these compounds show activity (e.g. enzyme, receptor, whole cell data, etc. similar to Tables 1 and 2) as well as clinical data would be invaluable. Such a database could then be linked with other mining tools that enable 1D–3D similarity searching. Once created it could be used as the authoritative virtual screening database for repurposing before testing physical compounds in whole cells or target assays.

Summary

Analysis of the literature suggests that, by using HTS, there are many examples of FDA-approved drugs that are active against additional targets that can be used to therapeutic advantage for repositioning. For example, there are several examples for neglected diseases, including compounds with antimalarial, antitubercular, trypanosomal and Chagas disease activity (Table 2). To date, there are fewer such examples where *in silico* approaches have derived new uses for approved drugs (Table 2) [77,96]. However, with current technologies and databases, as well as a close integration with *in vitro* screening, this will change. Although computational approaches, such as ligand- and structure-based methods, have been widely used for searching libraries of commercial compounds for neglected diseases [97], few have tried to use already existing drugs with computational methods [77]. A recently described apparent gap has been noted in the Mtb community between the generation and utilization of computational models for drug discovery [98]. These *in silico* models are not well disseminated and certainly not widely used for repositioning FDA-approved drugs. This situation needs to be rectified. Another important consideration should be the quality of the structures

in the databases used, whether of FDA drugs or other molecules, as these will impact the *in silico* results [94]. If neglected diseases can benefit from *in silico* methods so too can rare or orphan diseases as well as more common diseases. Repositioning approved drugs brings with it other incentives, such as seven-year market exclusivity [98], whereas new approved drugs or vaccines for a neglected disease can qualify for an FDA priority review voucher (US Medical Device User Fee and Modernization Act). In our opinion, some or all of the aforementioned *in silico* approaches should be used alongside *in vitro* methods to drug repurposing, if for no other reason than to speed up the process of drug discovery at little additional cost.

Conflicts of interest

S.E. consults for Collaborative Drug Discovery, Inc on a Bill and Melinda Gates Foundation Grant#49852 'Collaborative drug discovery for TB through a novel database of SAR data optimized to promote data archiving and sharing'.

Supplementary information

Supplemental information is available at http://www.4shared.com/account/file/MzCpwWw-/SCUT_Monkey_CLEANED.html or available from the corresponding author upon request. The updated SCUT database is provided as an sd file as used in recent similarity studies and pharmacophore searches [30,31,33,35,36].

Acknowledgements

S.E. gratefully acknowledges the many groups that have provided data sets, including that of David Sullivan, and also Barry A. Bunin and other colleagues at CDD for developing the software, for assistance with large data sets, and for collaborations. S.E. also acknowledges collaborators including Maggie A.Z. Hupcey, Sandhya Kortagere, Peter W. Swaan, James E. Polli and their students for collaborations summarized in Table 3. M.D.K. received support from grant K08-GM074238 from the National Institutes of Health. Accelrys Inc. are thanked for providing Discovery Studio.

References

1 Fidock, D.A. (2010) Drug discovery: priming the antimalarial pipeline. *Nature* 465, 297–298
2 Balganesh, T.S. *et al.* (2008) Rising standards for tuberculosis drug development. *Trends Pharmacol. Sci.* 29, 576–581
3 Griggs, R.C. *et al.* (2009) Clinical research for rare disease: opportunities, challenges, and solutions. *Mol. Genet. Metab.* 96, 20–26
4 Brewer, G.J. (2009) Drug development for orphan diseases in the context of personalized medicine. *Transl. Res.* 154, 314–322
5 Ashburn, T.T. and Thor, K.B. (2004) Drug repositioning: identifying and developing new uses for existing drugs. *Nat. Rev. Drug Discov.* 3, 673–683
6 Lehar, J. *et al.* (2008) High-order combination effects and biological robustness. *Mol. Syst. Biol.* 4, 215
7 Borisy, A.A. *et al.* (2003) Systematic discovery of multicomponent therapeutics. *Proc. Natl. Acad. Sci. U.S.A.* 100, 7977–7982
8 Cavalla, D. (2009) APT drug R&D: the right active ingredient in the right presentation for the right therapeutic use. *Nat. Rev. Drug Discov.* 8, 849–853
9 Boguski, M.S. *et al.* (2009) Drug discovery. Repurposing with a difference. *Science* 324, 1394–1395
10 Uliana, S.R. and Barcinski, M.A. (2009) Repurposing for neglected diseases. *Science* 326, 935; author reply, 935
11 Payne, D.A. *et al.* (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* 6, 29–40
12 Fischbach, M.A. and Walsh, C.T. (2009) Antibiotics for emerging pathogens. *Science* 325, 1089–1093
13 Miller, J.R. *et al.* (2009) A class of selective antibacterials derived from a protein kinase inhibitor pharmacophore. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1737–1742
14 Walsh, C.T. and Fischbach, M.A. (2009) Repurposing libraries of eukaryotic protein kinase inhibitors for antibiotic discovery. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1689–1690
15 Ekins, S. *et al.* (2007) In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *Br. J. Pharmacol.* 152, 9–20
16 Ekins, S. and Williams, A.J. (2010) Precompetitive preclinical ADME/Tox data: set it free on the web to facilitate computational model building to assist drug development. *Lab Chip* 10, 13–22
17 Chong, C.R. and Sullivan, D.J., Jr (2007) New uses for old drugs. *Nature* 448, 645–646
18 Krejsa, C.M. *et al.* (2003) Predicting ADME properties and side effects: the BioPrint approach. *Curr. Opin. Drug Discov. Dev.* 6, 470–480
19 Lougheed, K.E. *et al.* (2009) New anti-tuberculosis agents amongst known drugs. *Tuberculosis* 89, 364–370
20 O'Connor, K.A. and Roth, B.L. (2005) Finding new tricks for old drugs: an efficient route for public-sector drug discovery. *Nat. Rev. Drug Discov.* 4, 1005–1014
21 Ekins, S. *et al.* (2007) In silico pharmacology for drug discovery: applications to targets and beyond. *Br. J. Pharmacol.* 152, 21–37
22 Kauvar, L.M. *et al.* (1995) Predicting ligand binding to proteins by affinity fingerprinting. *Chem. Biol.* 2, 107–118
23 Kauvar, L.M. and Laborde, E. (1998) The diversity challenge in combinatorial chemistry. *Curr. Opin. Drug Discov. Dev.* 1, 66–70
24 Kauvar, L.M. *et al.* (1998) Protein affinity map of chemical space. *J. Chromatogr. B* 715, 93–102
25 Ma'ayan, A. *et al.* (2007) Network analysis of FDA approved drugs and their targets. *Mt Sinai J. Med.* 74, 27–32
26 Gomella, L. and Haist, S. (2004) *Clinician's Pocket Drug Reference 2004*. McGraw-Hill
27 Ekins, S. *et al.* (2005) In vitro and pharmacophore based discovery of novel hPEPT1 inhibitors. *Pharm. Res.* 22, 512–517
28 Chang, C. *et al.* (2006) Rapid identification of P-glycoprotein substrates and inhibitors. *Drug Metab. Dispos.* 34, 1976–1984
29 Diao, L. *et al.* (2010) Quantitative structure activity relationship for inhibition of human organic cation/carnitine transporter. *Mol. Pharm.* 7, 2120–2131
30 Diao, L. *et al.* (2009) Novel inhibitors of human organic cation/carnitine transporter (hOCTN2) via computational modeling and in vitro testing. *Pharm. Res.* 26, 1890–1900
31 Zheng, X. *et al.* (2009) Computational models for drug inhibition of the human apical sodium-dependent bile acid transporter. *Mol. Pharm.* 6, 1591–1603
32 Lamichhane, G. *et al.* (2011) Essential metabolites of M. tuberculosis and their mimics. *Mbio.* 2, e00301–e00310
33 Krasowski, M.D. *et al.* (2009) Molecular similarity methods for predicting cross-reactivity with therapeutic drug monitoring immunoassays. *Ther. Drug Monit.* 31, 337–344
34 Krasowski, M.D. *et al.* (2010) Immunoassays for tricyclic antidepressants: unsuitable for therapeutic drug monitoring. In *Advances in Chromatographic Techniques for Therapeutic Drug Monitoring* (Dasgupta, A., ed.), pp. 179–190, CRC Press
35 Krasowski, M.D. *et al.* (2009) Chemoinformatic methods for predicting interference in drug of abuse/toxicology immunoassays. *Clin. Chem.* 55, 1203–1213
36 Krasowski, M.D. *et al.* (2009) Using molecular similarity to highlight the challenges of routine immunoassay-based drug of abuse/toxicology screening in emergency medicine. *BMC Emerg. Med.* 9, 5
37 Reid, D. *et al.* (2008) LASSO-ligand activity by surface similarity order: a new tool for ligand based virtual screening. *J. Comput. Aided Mol. Des.* 22, 479–487
38 Williams, A.J. (2008) Internet-based tools for communication and collaboration in chemistry. *Drug Discov. Today* 13, 502–506
39 Poroikov, V.V. *et al.* (2000) Robustness of biological activity spectra predicting by computer program PASS for noncongeneric sets of chemical compounds. *J. Chem. Inf. Comput. Sci.* 40, 1349–1355
40 Kortagere, S. *et al.* (2009) The importance of discerning shape in molecular pharmacology. *Trends Pharmacol. Sci.* 30, 138–147
41 Li, Y.Y. *et al.* (2006) A large-scale computational approach to drug repositioning. *Genome Inform.* 17, 239–247
42 Ekins, S. and Abramowitz, D.L. (2009) A systems-biology view of drug transporters. In *Transporters as Drug Carriers* (Ecker, G. and Chiba, P., eds), pp. 365–385, Wiley-VCH, Verlag GmbH &Co. KGaA

43 Ekins, S. and Giroux, C. (2006) Computers and systems biology for Pharmaceutical Research and Development. In *Computer Applications in Pharmaceutical Research and Development* (Ekins, S., ed.), pp. 139–165, John Wiley & Sons
44 Ekins, S. *et al.* (2006) Algorithms for network analysis in systems: ADME/Tox using the MetaCore and MetaDrug platforms. *Xenobiotica* 36, 877–901
45 Ekins, S. (2006) Systems-ADME/Tox: resources and network approaches. *J. Pharmacol. Toxicol. Methods* 53, 38–66
46 Ekins, S. *et al.* (2005) Techniques: application of systems biology to absorption, distribution, metabolism, excretion, and toxicity. *Trends Pharmacol. Sci.* 26, 202–209
47 Fliri, A.F. *et al.* (2005) Biological spectra analysis: linking biological activity profiles to molecular structure. *Proc. Natl. Acad. Sci. U.S.A.* 102, 261–266
48 Fliri, A.F. *et al.* (2005) Biospectra analysis: model proteome characterizations for linking molecular structure and biological response. *J. Med. Chem.* 48, 6918–6925
49 Fliri, A.F. *et al.* (2005) Analysis of drug-induced effect patterns to link structure and side effects of medicines. *Nat. Chem. Biol.* 1, 389–397
50 Ekins, S. (2004) Predicting undesirable drug interactions with promiscuous proteins in silico. *Drug Discov. Today* 9, 276–285
51 Ekins, S. *et al.* (2006) A combined approach to drug metabolism and toxicity assessment. *Drug Metab. Dispos.* 34, 495–503
52 Nikolsky, Y. *et al.* (2005) A novel method for generation of signature networks as biomarkers from complex high throughput data. *Toxicol. Lett.* 158, 20–29
53 Ekins, S. *et al.* (2005) Systems biology: applications in drug discovery. In *Drug discovery Handbook* (Gad, S., ed.), pp. 123–183, John Wiley & Sons
54 Ekins, S. *et al.* (2007) Pathway mapping tools for analysis of high content data. *Methods Mol. Biol.* 356, 319–350
55 Paolini, G.V. *et al.* (2006) Global mapping of pharmacological space. *Nat. Biotechnol.* 24, 805–815
56 Wishart, D.S. *et al.* (2008) DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* 36, D901–D906
57 Yildirim, M.A. *et al.* (2007) Drug-target network. *Nat. Biotechnol.* 25, 1119–1126
58 Keiser, M.J. *et al.* (2007) Relating protein pharmacology by ligand chemistry. *Nat. Biotechnol.* 25, 197–206
59 Metz, J.T. and Hajduk, P.J. (2010) Rational approaches to targeted polypharmacology: creating and navigating protein–ligand interaction networks. *Curr. Opin. Chem. Biol.* 14, 498–504
60 Cases, M. and Mestres, J. (2009) A chemogenomic approach to drug discovery: focus on cardiovascular diseases. *Drug Discov. Today* 14, 479–485
61 Berg, E.L. *et al.* (2010) Chemical target and pathway toxicity mechanisms defined in primary human cell systems. *J. Pharmacol. Toxicol. Methods* 61, 3–15
62 Bender, A. *et al.* (2007) Analysis of pharmacology data and the prediction of adverse drug reactions and off-target effects from chemical structure. *ChemMedChem* 2, 861–873
63 Azzaoui, K. *et al.* (2007) Modeling Promiscuity based on in vitro safety pharmacology profiling data. *ChemMedChem* 2, 874–880
64 Scheiber, J. *et al.* (2009) Gaining insight into off-target mediated effects of drug candidates with a comprehensive systems chemical biology analysis. *J. Chem. Inf. Model.* 49, 308–317
65 Scheiber, J. *et al.* (2009) Mapping adverse drug reactions in chemical space. *J. Med. Chem.* 52, 3103–3107
66 Keiser, M.J. *et al.* (2009) Predicting new molecular targets for known drugs. *Nature* 462, 175–181
67 Hammann, F. *et al.* (2010) Prediction of adverse drug reactions using decision tree modeling. *Clin. Pharmacol. Ther.* 88, 52–59
68 Gurulingappa, H. *et al.* (2009) Concept-based semi-automatic classification of drugs. *J. Chem. Inf. Model.* 49, 1986–1992
69 Chiang, A.P. and Butte, A.J. (2009) Systematic evaluation of drug–disease relationships to identify leads for novel drug uses. *Clin. Pharmacol. Ther.* 86, 507–510
70 von Eichborn, J. *et al.* (2011) PROMISCUOUS: a database for network-based drug-repositioning. *Nucleic Acids Res.* 39, D1060–1066 Epub 2010 Nov 10
71 Ha, S. *et al.* (2008) IDMap: facilitating the detection of potential leads with therapeutic targets. *Bioinformatics* 24, 1413–1415
72 Jensen, N.H. and Roth, B.L. (2008) Massively parallel screening of the receptorome. *Comb. Chem. High Throughput Screen* 11, 420–426
73 Strachan, R.T. *et al.* (2006) Screening the receptorome: an efficient approach for drug discovery and target validation. *Drug Discov. Today* 11, 708–716
74 Roth, B.L. *et al.* (2004) Screening the receptorome to discover the molecular targets for plant-derived psychoactive compounds: a novel approach for CNS drug discovery. *Pharmacol. Ther.* 102, 99–110
75 Schadt, E.E. *et al.* (2009) A network view of disease and compound screening. *Nat. Rev. Drug Discov.* 8, 286–295
76 Pujol, A. *et al.* (2010) Unveiling the role of network and systems biology in drug discovery. *Trends Pharmacol. Sci.* 31, 115–123
77 Kinnings, S.L. *et al.* (2009) Drug discovery using chemical systems biology: repositioning the safe medicine Comtan to treat multi-drug and extensively drug resistant tuberculosis. *PLoS Comput. Biol.* 5, e1000423
78 Cockell, S.J. *et al.* (2010) An integrated dataset for in silico drug discovery. *J. Integr. Bioinform.* 10.2390/biecoll-jib-2010-116
79 Prathipati, P. *et al.* (2008) Global Bayesian models for the prioritization of antitubercular agents. *J. Chem. Inf. Model.* 48, 2362–2370
80 Jones, D.R. *et al.* (2007) Computational approaches that predict metabolic intermediate complex formation with CYP3A4 (+b5). *Drug Metab. Dispos.* 35, 1466–1475
81 Zientek, M. *et al.* (2010) Integrated in silico-in vitro strategy for addressing cytochrome P450 3A4 time-dependent inhibition. *Chem. Res. Toxicol.* 23, 664–676
82 Pan, Y. *et al.* Identification and validation of novel hPXR activators amongst prescribed drugs via ligand-based virtual screening. *Drug Metab. Dispos.* (in press)
83 Ekins, S. *et al.* (2010) A predictive ligand-based bayesian model for human drug induced liver injury. *Drug Metab. Dispos.* 38, 2302–2308
84 Ekins, S. *et al.* (2009) Challenges predicting ligand–receptor interactions of promiscuous proteins: the nuclear receptor PXR. *PLoS Comput. Biol.* 5, e1000594
85 Ananthan, S. *et al.* (2009) High-throughput screening for inhibitors of Mycobacterium tuberculosis H37Rv. *Tuberculosis* 89, 334–353
86 Ekins, S. *et al.* (2010) Analysis and hit filtering of a very large library of compounds screened against Mycobacterium tuberculosis. *Mol. BioSyst.* 6, 2316–2324
87 Maddry, J.A. *et al.* (2009) Antituberculosis activity of the molecular libraries screening center network library. *Tuberculosis* 89, 354–363
88 Hohman, M. *et al.* (2009) Novel web-based tools combining chemistry informatics, biology and social networks for drug discovery. *Drug Discov. Today* 14, 261–270
89 Ekins, S. *et al.* (2010) A collaborative database and computational models for tuberculosis drug discovery. *Mol. BioSyst.* 6, 840–851
90 Ekins, S. *et al.* (2011) Pioneering use of the cloud for development of the collaborative drug discovery (cdd) database In *Collaborative Computational Technologies for Biomedical Research* (Ekins, S. et al., eds), pp. 335–361, John Wiley & Sons (in press)
91 Ekins, S. and Williams, A.J. (2010) Meta-analysis of molecular property patterns and filtering of public datasets of antimalarial 'hits' and drugs. *MedChemComm* 1, 325–330
92 Ekins, S. and Williams, A.J. (2010) When pharmaceutical companies publish large datasets: an abundance of riches or fool's gold? *Drug Discov. Today* 15, 812–815
93 Wang, Y. *et al.* (2010) An overview of the PubChem BioAssay resource. *Nucleic Acids Res.* 38, D255–D266
94 Williams, A.J. *et al.* (2009) Free online resources enabling crowdsourced drug discovery. *Drug Discov. World* 10, winter, 33–38
95 Louise-May, S. *et al.* (2009) Towards integrated web-based tools in drug discovery. *Touch Briefings: Drug Discov.* 6, 17–21
96 Bisson, W.H. *et al.* (2007) Discovery of antiandrogen activity of nonsteroidal scaffolds of marketed drugs. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11927–11932
97 Jenwitheesuk, E. *et al.* (2008) Novel paradigms for drug discovery: computational multitarget screening. *Trends Pharmacol. Sci.* 29, 62–71
98 Ekins, S. *et al.* (2011) Computational databases, pathway and cheminformatics tools for tuberculosis drug discovery. *Trends Microbiol* 65–74 Epub 2010 Dec 2
99 Manak, M.M. *et al.* (2010) Anti-HIV-1 activity of the neurokinin-1 receptor antagonist aprepitant and synergistic interactions with other antiretrovirals. *AIDS* 24, 2789–2796
100 Wang, X. *et al.* (2007) Neurokinin-1 receptor antagonist (aprepitant) inhibits drug-resistant HIV-1 infection of macrophages in vitro. *J. Neuroimmune Pharmacol.* 2, 42–48
101 Robinson, P. *et al.* (2008) Substance P receptor antagonism for treatment of cryptosporidiosis in immunosuppressed mice. *J. Parasitol.* 94, 1150–1154
102 Oldfield, E. (2010) Targeting isoprenoid biosynthesis for drug discovery: bench to bedside. *Acc. Chem. Res.* 43, 1216–1226
103 Ting, H.J. and Khasawneh, F.T. (2010) Glybenclamide: an antidiabetic with in vivo antithrombotic activity. *Eur. J. Pharmacol.* 649, 249–254
104 Miguel, D.C. *et al.* (2008) Tamoxifen is effective in the treatment of Leishmania amazonensis infections in mice. *PLoS Negl. Trop. Dis.* 2, e249
105 Miguel, D.C. *et al.* (2007) Tamoxifen is effective against Leishmania and induces a rapid alkalinization of parasitophorous vacuoles harbouring Leishmania (Leishmania) amazonensis amastigotes. *J. Antimicrob. Chemother.* 60, 526–534
106 Senkovich, O. *et al.* (2005) Lipophilic antifolate trimetrexate is a potent inhibitor of Trypanosoma cruzi: prospect for chemotherapy of Chagas' disease. *Antimicrob. Agents Chemother.* 49, 3234–3238
107 Namkoong, J. *et al.* (2007) Metabotropic glutamate receptor 1 and glutamate signaling in human melanoma. *Cancer Res.* 67, 2298–2305

108 Peng, Q. *et al.* (2008) The antidepressant sertraline improves the phenotype, promotes neurogenesis and increases BDNF levels in the R6/2 Huntington's disease mouse model. *Exp. Neurol.* 210, 154–163

109 Chong, C.R. *et al.* (2007) Inhibition of angiogenesis by the antifungal drug itraconazole. *ACS Chem. Biol.* 2, 263–270

110 Chong, C.R. *et al.* (2006) A clinical drug library screen identifies astemizole as an antimalarial agent. *Nat. Chem. Biol.* 2, 415–416

111 Chong, C.R. *et al.* (2006) Identification of type 1 inosine monophosphate dehydrogenase as an antiangiogenic drug target. *J. Med. Chem.* 49, 2677–2680

112 de Carvalho, L.P. *et al.* (2009) Nitazoxanide kills replicating and nonreplicating Mycobacterium tuberculosis and evades resistance. *J. Med. Chem.* 52, 5789–5792

113 Shahinas, D. *et al.* (2010) A repurposing strategy identifies novel synergistic inhibitors of Plasmodium falciparum heat shock protein 90. *J. Med. Chem.* 53, 3552–3557

114 Chopra, S. *et al.* (2010) Repurposing FDA-approved drugs to combat drug-resistant Acinetobacter baumannii. *J. Antimicrob. Chemother.* 65, 2598–2601

115 Miller, S.C. *et al.* (2010) Identification of known drugs that act as inhibitors of NF-kappaB signaling and their mechanism of action. *Biochem. Pharmacol.* 79, 1272–1280

116 Downey, A.S. *et al.* (2008) Efficacy of pyrvinium pamoate against *Cryptosporidium parvum* infection in vitro and in a neonatal mouse model. *Antimicrob. Agents Chemother.* 52, 3106–3112

117 Mackey, Z.B. *et al.* (2006) Discovery of trypanocidal compounds by whole cell HTS of *Trypanosoma brucei*. *Chem. Biol. Drug Des.* 67, 355–363

118 Biechele, T.L. *et al.* (2010) Chemical-genetic screen identifies riluzole as an enhancer of Wnt/beta-catenin signaling in melanoma. *Chem. Biol.* 17, 1177–1182

119 Gloeckner, C. *et al.* (2010) Repositioning of an existing drug for the neglected tropical disease Onchocerciasis. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3424–3429

120 Shim, J.S. *et al.* (2010) Effect of nitroxoline on angiogenesis and growth of human bladder cancer. *J. Natl. Cancer Inst.* 102, 1855–1873 Epub 2010 Nov 18

121 Zhang, Y. *et al.* (2009) Identification of inhibitors of ABCG2 by a bioluminescence imaging-based high-throughput assay. *Cancer Res.* 69, 5867–5875

122 Masuda, N. *et al.* (2008) Tiagabine is neuroprotective in the N171-82Q and R6/2 mouse models of Huntington's disease. *Neurobiol. Dis.* 30, 293–302

123 Zhang, H. *et al.* (2008) Digoxin and other cardiac glycosides inhibit HIF-1alpha synthesis and block tumor growth. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19579–19586

124 Ou, H.C. *et al.* (2009) Identification of FDA-approved drugs and bioactives that protect hair cells in the zebrafish (Danio rerio) lateral line and mouse (Mus musculus) utricle. *J. Assoc. Res. Otolaryngol.* 10, 191–203

125 Peterson, S.A. *et al.* (1998) Inhibiting transthyretin conformational changes that lead to amyloid fibril formation. *Proc. Natl. Acad. Sci. U.S.A.* 95, 12956–12960

126 Ekins, S. (2009) Drug transporter pharmacophores. In *Transporters as Drug Carriers: Structure, Function, Substrates*, (Vol. 44) (Ecker, G. and Chiba, P., eds) pp. 215–227, Wiley-VCH Verlag GmbH & Co KGaA

127 Guiguemde, W.A. *et al.* (2010) Chemical genetics of Plasmodium falciparum. *Nature* 465, 311–315

128 Gamo, F.-J. *et al.* (2010) Thousands of chemical starting points for antimalarial lead identification. *Nature* 465, 305–310