NEWS & VIEWS

DRUG DISCOVERY

Predicting promiscuity

Andrew L. Hopkins

Computational methods that reliably predict the biological activities of compounds have long been sought. The validation of one such method suggests that *in silico* predictions for drug discovery have come of age.

The metaphor of drugs acting as 'magic bullets', selectively binding to specific physiological targets, dates back to Paul Ehrlich and the foundation of modern drug discovery. But in recent years, the observation that drugs often bind to more than one molecular target — that they exhibit polypharmacology — has gained attention[1], suggesting that a 'magic shotgun' analogy might be more apt (Fig. 1). To fully understand the actions of a drug, knowledge of its polypharmacology is clearly essential. On page 175 of this issue, Keiser *et al.*[2] report a computational tool that generates predictions of the pharmacological profile of drugs, and provide experimental validation of their method.

The interactions of drugs with off-target proteins have conventionally been viewed as undesirable 'promiscuity', responsible for unwanted side effects. But in many cases — ranging from certain older psychiatric drugs to modern anticancer therapies — this promiscuity is intrinsic to the drug's therapeutic efficacy. An understanding of polypharmacology can therefore explain why some drugs work better than expected, or why different drugs that supposedly act on the same target vary in their side effects. Perhaps more importantly, such insight offers opportunities to develop drugs for diseases that don't yet have therapies. But screening drugs against all the proteins expressed by the 21,000-plus genes in the human genome is currently unfeasible. A long-time goal of chemoinformatics has therefore been to develop computational techniques that predict the proteins to which drugs are likely to bind.

Various *in silico* methods for predicting the pharmacological profile of drugs are in development, the most well known of which is to 'dock' the three-dimensional structure of a compound virtually into the structure of a protein. But among the limitations of docking methods is the need for high-resolution X-ray crystal structures of proteins. These are particularly difficult to obtain for membrane-bound proteins, which account for 60% of drug targets. An alternative approach has therefore been developed that does not require protein structures. This approach works by analysing the chemical structures of ligand molecules that are known to bind to drug targets, to identify the structural motifs responsible for the binding.

Several groups are also developing machine-learning methods that identify molecular binding motifs that might relate compounds that act at one protein to those that act at another, as a means of inferring the unknown polypharmacology of a drug. These methods are often based on Bayesian statistical analyses[3]. Keiser and colleagues' technique[2] is also ligand-based, but it uses a new algorithm to build up profiles of the patterns (or fingerprints) of chemical structures for all the reported ligands of a drug target. These profiles can then be used to calculate the probability that another structurally unrelated compound contains the same chemical patterns[4]. The underlying mathematics of this similarity ensemble approach[2] (SEA) is inspired by the methods used to compare the similarity of amino-acid sequences in proteins.

Keiser *et al.* used their technique to analyse more than 3,600 compounds, each of which was either an approved drug or a compound that had been investigated as a possible drug. They computationally screened the chemical structures of these compounds against databases of hundreds of thousands of biologically active compounds, for which the binding profiles to a panel of more than 1,400 proteins are known. This yielded a massive network of predicted drug–target interactions for the 3,600 compounds in question. Nearly 7,000 of these predicted interactions had sufficiently high probabilities to warrant further study.

The authors investigated the most likely predictions of drug polypharmacology for several compounds, both retrospectively (by database and literature searching) and prospectively (by *in vitro* and *in vivo* screening). Although only a tiny fraction of all the predictions was tested, the authors' initial validation data confirmed that two out of every three high-confidence predictions were correct. This implies that the polypharmacology network between drugs

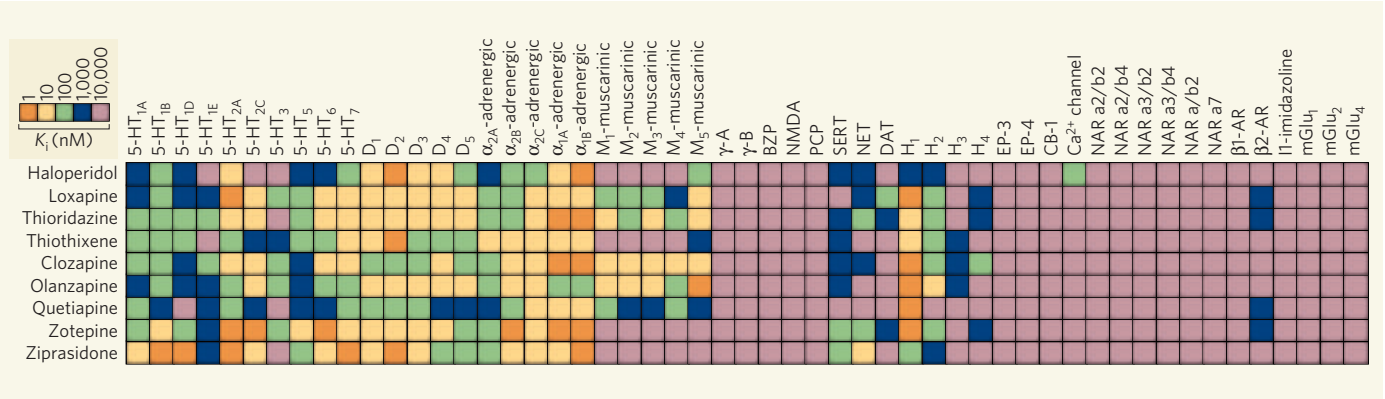

Figure 1 | One drug, many targets. Drugs often exhibit polypharmacology — they bind at more than one molecular site. The chart shows the known affinity (K_i) values of several antipsychotic drugs for a panel of receptors. Low K_i values correspond to high affinities. Keiser *et al.*[2] report and validate a computational method that predicts the molecular sites to which a drug will bind. (Adapted from ref. 1.)

© 2009 Macmillan Publishers Limited. All rights reserved

NATURE|Vol 462|12 November 2009

and drug targets is far denser than is currently suggested by the published data[5] (see Fig. 1 on page 176).

Several of Keiser and colleagues' predictions[2] provide fascinating insights into the previously unknown pharmacology of well-known drugs. For example, the antidepressants fluoxetine (Prozac) and paroxetine (Paxil) are selective serotonin reuptake inhibitors (SSRIs) — they work by inhibiting the uptake of the neurotransmitter serotonin into presynaptic neurons. But the authors' computational method strongly predicted that these drugs also act as beta-blockers (they bind to β-adrenergic receptors in, for example, blood vessels and heart muscle), a prediction that was confirmed *in vitro*. This result suggests that some of the side effects that occur when patients stop taking SSRI antidepressants might be caused by a rebound in β-adrenergic signalling.

As serotonin and β-adrenergic receptors bind to chemically similar, naturally occurring ligands, the beta-blocker activity of SSRIs might not be too unexpected. A much tougher challenge is to predict a compound's polypharmacology for drug targets that share no discernible similarity in terms of amino-acid sequence, protein structure or endogenous ligands. Again, Keiser and colleagues' technique comes up with the goods: the authors confirmed the SEA prediction that delavirdine (Rescriptor, an HIV-1 reverse transcriptase inhibitor) also binds to histamine H_4 receptors. These drug targets have no biological similarity whatsoever, yet the structure of delavirdine 'encodes' its ability to recognize both reverse transcriptase and H_4 receptors. The binding of delavirdine to these receptors might explain why it sometimes causes painful skin rashes as a side effect.

Although computational techniques for drug discovery have been around for some time, confidence in their use has often been lacking. But with the growing availability of free databases of pharmacological information, such as ChEMBL[6] and PubChem[7], the field of chemoinformatics finally has sufficiently large quantities of data for machine-learning methods to be developed that could reliably predict drug activity. Keiser and colleagues' experimental validation[2] of their SEA for predicting polypharmacology from chemical structure shows that such methods are truly beginning to bear fruit. ■

Andrew L. Hopkins is in the Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee DD1 5EH, UK.
e-mail: a.hopkins@dundee.ac.uk

1. Roth, B. L., Sheffler, D. J. & Kroeze, W. K. *Nature Rev. Drug Discov.* **3,** 353–359 (2004).
2. Keiser, M. J. *et al. Nature* **462,** 175–181 (2009).
3. Nigsch, F., Bender, A., Jenkins, J. L. & Mitchell, J. B. *J. Chem. Inf. Model.* **48,** 2313–2325 (2008).
4. Keiser, M. J. *et al. Nature Biotechnol.* **25,** 197–206 (2007).
5. Yildirim, M. A. *et al. Nature Biotechnol.* **25,** 1119–1126 (2007).
6. www.ebi.ac.uk/chembl
7. http://pubchem.ncbi.nlm.nih.gov

ASTROPHYSICS

A fossil record for exoplanets

Marc Pinsonneault

Stars that host planets experience more mixing of their internal elements than do stars that lack such companions. This correlation may serve as a useful diagnostic in the search for planets around stars other than the Sun.

The planets that have been discovered beyond our Solar System — exoplanets — are extremely diverse. As the field of exoplanetary research has matured, systems containing multiple planets have been detected and classified, and our detection sensitivity has extended to planets with lower masses and longer orbital periods. Planet formation is highly sensitive to the properties of the accretion disks of gas around the protostars, and the masses and lifetimes of these disks are known to vary dramatically from system to system. The discovery that very different types of exoplanetary system exist is therefore not surprising, and competing theoretical models for their formation have emerged. However, evidence of the formation process around stars such as the Sun is typically erased by the passage of time, and directly correlating planetary properties with formation scenarios has not been possible. In this issue, Israelian *et al.*[1] (page 189) demonstrate that the abundance of lithium — a diagnostic of the internal mixing of elements in stars — differs significantly between stars that host planets and those that do not.

The depletion of lithium on the surface of a star is a sign either that the matter at the surface layers has been exposed to temperatures greater than 2.5 million kelvin or that the lithium has settled — like sugar in an unstirred glass of iced tea — below the star's surface. The latter process has a long timescale for Sun-like stars. Standard stellar-evolution theory predicts that protostars have a thick, turbulent, convective envelope of gas with a temperature at its base that is sufficiently high to burn some, but not all, of the lithium. This envelope retreats as the star stabilizes and reaches a Sun-like 'main-sequence' state, in which energy is created through the fusion of hydrogen at the star's core. Lithium depletion is progressively more severe at all stages for lower-mass stars and those with a higher concentration of heavy elements. However, no main-sequence depletion of lithium is predicted for stars more massive than about 0.9 solar masses, and Sun-like stars are observed to arrive at the main-sequence state with relatively high surface lithium abundances.

This basic picture of the thermal structure of stars is confirmed by seismic measurements of the thickness of the Sun's convection zone, which is too thin and cold at its base to destroy lithium[2]. However, the Sun is depleted in lithium by a factor of more than 100 relative to the value measured in meteorites, and by a factor of more than 50 relative to young Sun-like stars[3]. This is a general phenomenon: we observe that, for Sun-like stars, lithium abundance declines with age during the main-sequence stage and, furthermore, that the degree of depletion differs for stars of the same mass, composition and age[4]. An additional physical mechanism is thus required to explain this complex observational pattern.

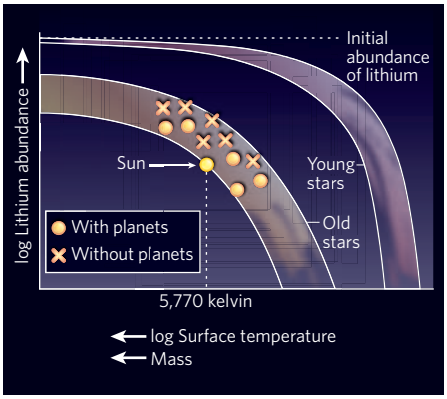

Figure 1 | Stellar lithium depletion. Lithium is depleted in stars early in their evolution, when they have a thick, turbulent, convective envelope of gas; turbulence cycles the fragile element to the stars' centre, where it is destroyed by nuclear reactions. This turbulent mixed envelope becomes thinner once hydrogen fusion is established in the stars' cores (at which point they become 'main-sequence' stars), but young stars retain a record of their early history (top band in this schematic illustration). The lithium depletion rate in this early stage increases with decreasing mass and surface temperature. After the main-sequence state is reached, stellar rotation can further enhance lithium depletion by slowly mixing it into the stars' interiors. This extra degree of mixing varies from star to star as they age, yielding a dispersion in lithium abundances in old stars (bottom band). Israelian *et al.*[1] find that the Sun and Sun-like stars that host planets (circles) have lower lithium abundances than those that do not (crosses). This indicates that planet formation must have influenced the lithium depletion process.

Israelian and colleagues[1] demonstrate that lithium depletion is more severe in Sun-like stars that host planets than in those that do not (Fig. 1). They choose a sample of stars with similar surface temperatures, and therefore masses, so as to minimize the known mass dependence of lithium depletion. They find no convincing evidence that the relative

© 2009 Macmillan Publishers Limited. All rights reserved