The Impact of the human endogenous metabolome on drug pharmacology and safety

Andreu Bofill Pumarola

TESI DOCTORAL UPF / ANY 2020

DIRECTORS DE LA TESI:

Dr. Jordi Mestres
Dr. Xavier Jalencas

DEPARTAMENT DE CIÈNCIES EXPERIMENTALS I DE LA SALUT

Universitat Pompeu Fabra Barcelona

The research in this Thesis has been carried out at the Systems Pharmacology Group, within the Research Programme on Biomedical Informatics (GRIB) at the Parc de Recerca Biomèdica de Barcelona (PRBB).

The research presented in this Thesis has been supported by the European H2020 ESCAPE-NET Project (Grant Agreement 733381) and by a RETOS project from the Spanish Ministerio de Ciencia, Innovació y Universidades (SAF2017-83614-R).

Agraïments

Primer de tot i indubtablement, gràcies **mare**. Gràcies **pare**.
Us ho dec tot.

Agrair especialment al **Jordi** tota la dedicació, els consells, els patiments, les bromes i el temps invertit. Moltes gràcies per oferir-me aquesta oportunitat i, sobretot i de tot cor, per haver confiat en mi. **Xavi**, moltes gràcies per tot el que m'has ensenyat tant a nivell científic com personal. Gràcies per la proximitat amb la que m'has tractat durant tots aquests anys. Has estat un pilar sense el qual, molt probablement, no hauria acabat aquesta tesis.

Agrair també a totes les persones que han passat pel grup, pel GRIB, per Chemotargets i pel PRBB i que m'han ajudat tant laboralment com personal durant tota aquesta etapa. **Coté**, ha estat un plaer compartir amb tu tots aquests anys i poder-nos escoltar i ajudar amb tot allò que hem pogut. **Albert**, **Cristina**, **Irene**, **Joaquim, Viktoria** i **Jordi**, gràcies per fer més suportables tots els anys al grup. **Alfons**, moltíssimes gràcies per totes les converses, consells i suport. Espero que segueixis fent de tòtem d'altres estudiants tant com ho has estat per mi. **Eric, Bet, Juanma, Jose Carlos, Oriol, Nacho, Biel**, molta, moltíssima nostàlgia d'aquells migdies irrepetibles. **Mariona, Judith**, **Adri, Karolina, David, Quim** i tants d'altres, molt agraït d'haver compartit amb vosaltres tants moments. Moltes gràcies **Ricard**, tant per la paciència com per l'ajuda durant tots aquests anys. **Carina**, gràcies per tota la feinada feta i el suport que m'has donat des d'un bon inici.

Al Burning Servers Team! **Will, Jana, Jone, Simone, Ignasio, Eva** i tots els que han passat pel grup de Volleywood, gràcies per les magnífiques tardes, partits, sopars i festes.

Uxue! Quantes coses et podria dir i quantes me'n deixaria al darrere. Moltes gràcies per ser-hi sempre. Nigan beti izango duzu lagun bat.

Leire, Leire, Leire… Best "rompeparejas team". Ets una autèntica joia que m'enduc i que espero poder seguir polint. Inacabables gràcies per escoltar-me, ajudar-me i comprendre'm tant.

Ángela… D'un somni frustrat d'una aclaparadora victòria al 2020 a aparegut una gran amiga. Espero poder-te seguir gaudint durant molts més anys. **Sandra**! Masses anys vivint vides paral·leles… Ets una troballa massa tardana. **Sonia**!! Gràcies per totes les converses, moments i distraccions que m'has donat! **Marcos, Davide, Jesse, Raquel, Anna, Sami, Jeroen, Ezequiel, Jenni, Marc**, gràcies per tantíssimes tardes de vòlei, tornejos, festes, teatres, balls, cerveses, bromes, terrasses i especialment per aquest darrer estiu, més estrany que mai, que m'heu fet viure.

Mar, Alberto, Paula, Laura. Us ho prego... seguiu-me parlant de kefirs, realfood, llenties o companys de feina. Ha estat un plaer viure-us tot aquest temps. Aquesta etapa hagués estat bastant més grisa sense vosaltres. **Natàlia**, Trobaré a faltar tantes quedades, cafès i posades al dia!

Externament al PRBB, també hi ha ple de persones que m'han ajudat i que he d'agrair tota l'estima que m'han donat.

Sense cap mena de dubte, la parella que m'ha aguantat més durant tot aquest trajecte, **Adri, Marina**, us dec un monument. Sou la meva petita segona família. Gràcies per aguantar-me (no és fàcil), per escoltar-me i per ser-hi. **Aina, Meri, Maria, Sara** gràcies per formar part d'aquesta experiència. **Adri (1)**, no sabria ben bé on posar-te. Per aquí i per allà, i per tot arreu. Un enormíssim plaer gaudir tantes coses plegats.

Bertran, **Mariangeles**, **Marc**. Amor incondicional. Realment, no entendria el què sóc sense vosaltres (excepte per en Bertran, que segur que no hi està d'acord, i per tant espero que s'autoexclogui del 'vosaltres' (L)). Us estimo molt.

Bernat, Aaron, Nil, Ferran, Víctor. Estimats. Gràcies per l'escalf i la proximitat (a vegades llunyana). Us porto al cor!

Indubtablement, **Júlia**, ha estat una meravella haver-te conegut.

Natalia! Gràcies pel suport, la ajuda i per haver crescut plegats durant tants anys.

Joan, Pau, Vila, Edu, Joan, Enric, Íñigo, Dani, Nil, Jawad, Eli. Sou una part inesborrable de mi. Gràcies per ser-hi.

David, formes part de cada pensament, cada acció i cada minúscula passa que faci a la meva vida.

Una enorme abraçada a l'**àvia Marta**, amb una sincera i especial dedicatòria, que vull compartir amb l'àvia Montsa i als avis Joan. **Adela, Joan** i a la resta de família, gràcies de tot cor.

Clara, Adrià, malauradament no crec que us pugui arribar a expressar mai tot el que signifiqueu a la meva vida. Infinites gràcies.

Per a tu, un mar sencer. Gràcies per somiar amb mi, **Paula**.

"Això és tot el que us penso explicar. [...] Em sap greu haver-ho explicat a tanta gent. L'única cosa que sé és que trobo a faltar tota la gent que us he explicat. Fins i tot l'Stradlater i l'Ackley, per exemple. Em sembla que fins i tot trobo a faltar aquell desgraciat d'en Maurice. Té gràcia. No expliqueu mai res a ningú. Si ho feu, començareu a trobar a faltar tothom."

J.D. Salinger "The Catcher in the Rye"

Abstract

A great deal of efforts has been made to improve and expand pharmacological testing and optimisation of drug candidates during the discovery process. However, we are yet to understand fully why drugs require certain levels of affinity for their mechanism of action targets to exert their therapeutic effect. In parallel, the mechanistic understanding of the endogenous metabolome has been demonstrated to have incredible implications in drug discovery, being potentially one of the pillars of precision medicine. Endogenous metabolites are small molecules evolutionally optimised to interact in an appropriate way with its native protein with a specific level of affinity. This Thesis provides evidences that the level of affinity required by a drug to interact with its primary target and produce a therapeutic effect is related to the affinity of the native endogenous metabolite for that target. In addition, this Thesis also highlights the implications that the human endogenous metabolome could have to assess a more realistic drug polypharmacology landscape and the risk of safety events linked to it. Finally, an analysis of natural compound pharmacology on the perspective of the endogenous metabolome is also performed. All results indicate that understanding the role of the human metabolome and its implications on drug bioactivity could offer a new useful perspective in drug efficacy and safety.

Resum

S'han realitzat grans esforços per millorar i ampliar l'avaluació farmacològica i l'optimització de les molècules candidates durant el procés de descobriment de fàrmacs. Tot i això, encara no es coneix amb exactitud per què els fàrmacs necessiten certs nivells d'afinitat amb les seves dianes d'acció per tal de produir l'efecte terapèutic buscat. Paral·lelament, s'ha demostrat que la comprensió del metaboloma endogen té grans implicacions en el descobriment de nous fàrmacs i és potencialment un pilar de la medicina de precisió. Els metabòlits endògens són molècules petites optimitzades evolutivament per interaccionar de forma apropiada amb la seva proteïna nativa amb un nivell d'afinitat específic. Aquesta Tesi aporta evidències que l'afinitat necessària per tal que un fàrmac interaccioni amb la seva diana primària i produeixi un efecte terapèutic concret està relacionada amb l'afinitat del metabòlit endogen natiu amb aquesta mateixa diana. A més, aquesta Tesi també posa de manifest les implicacions que pot tenir el metaboloma endogen humà per tal d'avaluar, de forma més realista, la polifarmacologia dels fàrmacs i els seus efectes secundaris associats. Finalment, també s'ha realitzat una anàlisi de la farmacologia dels compostos naturals incorporant la informació del metaboloma endogen. Tots aquests resultats indiquen que entendre el rol del metaboloma humà i les seves implicacions en la bioactivitat dels fàrmacs pot oferir una perspectiva nova i útil per a millorar l'eficàcia i la seguretat dels fàrmacs.

Preface

Design and development of new drugs is one of the pillars of our welfare society. We want drugs to be able to exert its therapeutic effects while producing as few adverse events as possible. However, the reality of drug polypharmacology offers an incredibly complex scenario where the concept of drugs interacting selectively with one single target to produce the desired therapeutic effect without any other implication is almost unimaginable. According to this new perspective, several approaches are being developed in order to analyse drug safety and its pharmacological activity in a more holistic manner, that can be translated into immediate therapeutic implications.

The impact of metabolomics on several scientific disciplines is indubitable. This almost new discipline helps to understand biological systems from an entire new perspective, emphasizing the key role of metabolites in the chemical, biological and physiological processes. In this regard, metabolomics is a promising discipline that could bring some new viewpoints in drug pharmacology and drug safety. Indeed, the integration of metabolomics in drug discovery could be one of the key steps towards a more individualized medicine.

The principal aim of this Thesis is to study the role of the human endogenous metabolome and analyse its potential implications

for both drug efficacy and safety. This Thesis is divided into five different sections. First, a historical overview of drug discovery is provided, followed by a general perspective on metabolomics and drug safety, with special emphasis on a particular safety event, sudden cardiac arrest. The main objectives of the Thesis are then provided, followed by the results achieved by this Thesis, including a published manuscript. The results section leads to a global discussion that tries to address the impact of the human endogenous metabolome on drug discovery and natural compound administration. Finally, the main conclusions derived from this Thesis are exposed. The document is completed with a bibliographic section containing all cited references.

TABLE OF CONTENTS

ABSTRACT ..... **XI**

PREFACE ..... **XIII**

INTRODUCTION **1**

1. DRUG DISCOVERY **3**

Origins of drug discovery 3

Change of paradigm: the drug receptor theory 6

Target-based drug discovery 9

Drug polypharmacology 13

In silico drug discovery 16

Pharmaceutical business 22

Drug discovery and development process 24

2. METABOLOMICS **31**

Metabolomics: origin and evolution of a new omics discipline 31

Metabolomics, drug discovery and personalized medicine 35

3. DRUG SAFETY **39**

Origin of pharmacovigilance and safety pharmacology 39

Polypharmacology and safety pharmacology: A new pharmacovigilance paradigm 43

Off-target interactions as a cause of adverse drug reactions. 45

4. SUDDEN CARDIAC ARREST **49**

Epidemiology of Sudden cardiac arrest 49

Risk factors identification. First step for the detection and prevention of SCA 50

Drug-induced sudden cardiac arrest 55

OBJECTIVES **59**

RESULTS **63**

1. THE HUMAN ENDOGENOUS METABOLOME AS A PHARMACOLOGY BASELINE FOR DRUG DISCOVERY **65**

2. THE IMPACT OF THE HUMAN ENDOGENOUS METABOLOME ON THE SECONDARY PHARMACOLOGY OF DRUGS **125**

3. THE IMPACT OF THE HUMAN ENDOGENOUS METABOLOME ON NATURAL COMPOUNDS **151**

4. DRUG-INDUCED SUDDEN CARDIAC ARREST: AN ANALYSIS OF DRUG REPORT DATABASES **173**

DISCUSSION **207**

THE HUMAN ENDOGENOUS METABOLOME AS A BASELINE FOR PRIMARY PHARMACOLOGY OF DRUGS. 209

IMPACT OF THE HUMAN ENDOGENOUS METABOLOME INTO SECONDARY PHARMACOLOGY OF DRUGS. 211

PHARMACOLOGICAL VIEW OF NATURAL COMPOUNDS. THE IMPLICATIONS OF HUMAN ENDOGENOUS METABOLOME **214**

ANALYSIS OF POSTMARKETING ADVERSE EVENT REPORTS. A REVIEW OF DRUGS WITH A RISK OF INDUCING SUDDEN CARDIAC ARREST **215**

IMPLICATIONS OF METABOLOMICS IN DRUG DISCOVERY: INSIGHTS AND LIMITATIONS **216**

CONCLUSIONS **219**

REFERENCES **223**

INTRODUCTION

The discovery and development of drugs represent, nowadays, one of the cornerstones of our society and the therapeutical uses of drugs are widespread globally. In fact, the use of drugs dates back to the first days of human history.[1] Although several historians suggested that drug usage goes back to before 3000 BC, one of the first written records of medical therapeutics dates can be traced back to the days of ancient Egypt (1500 BC).[2] The document is named *Ebers papyrus* and it is a 20-meter long papyrus containing the description of many ancient remedies for a wide range of discomforts and ailments such as birth control, depression or diabetes mellitus.[3] The first compounds to be used as a remedy were derived from plants and could be enriched by mineral or animal substances. However, the discovery of new therapeutic compounds remained an empirical process, where the observation of the natural effects of these compounds on human bodies was crucial.[4] It was not until the early 19th century that drug discovery underwent a major transformation.

1. DRUG DISCOVERY

Origins of drug discovery

It has been established that drug research as such was born as an interdisciplinary science at the beginning of the nineteenth century with the expansion of the industrial

process.[5] The degree of maturity reached by the chemical sciences and the consolidation of pharmacology as a well-defined discipline established a solid background for the appearance of rational approaches to drug discovery.[6] Chemical advances like Avogradro's atomic principle, hypothesized in 1812 and, more importantly, the organization of chemical elements as a periodic table in 1869 by Dmitri Mendelev, allowed to classify small molecules and to universalize the research in the scientific field.[7] Within this context, it ought to be stressed that analytic chemistry emerged some years before and it can be considered the precursor of the 19th century medicinal chemistry.[8] Some active entities of natural medicinal plants had already been purified, such as the isolation of morphine (1815) or papaverin (1848) from the opium plant (Fig. 1).[9] The discovery and the possible

Figure 1: Opium poppy. Extracted from (Chast, 2008).[4]

purification of these active compounds from natural plants allowed some of the first pharmacies to provide standardized preparations.[5]

In 1865, the benzene theory was formulated by August Kekulé. This fact was extremely relevant for the coal-tar derivatives such as dyes.[5,10] The great abundance of coal-tar due to the industrialization process incited an early process of drug discovery based on the screening of compounds (mainly aromatic or aliphatic blocks) synthesized from the coal-tar derivatives that were the base of medicinal chemistry.[10,11] In this regard, Paul Ehrlich studied the selective affinity of dyes for specific biological tissues. He postulated the existence of "chemoreceptors" and observed how some compounds could bind extremely differently on the chemoreceptors of some microorganisms or parasites compared to the analogous structures on their hosts, introducing a new therapeutic horizon.[5] In many aspects, his studies can be considered the basis of chemotherapy. All these processes revealed the therapeutic potential of non-natural molecules obtained by synthetic organic chemistry.

In the last decades of the nineteenth century, new institutions and companies emerged from both pharmacies and dye companies in order to support this interdisciplinary drug discovery and to create a new way of finding, characterizing and developing novel synthetic medicines.[5]

Finally, drug discovery experienced an exponential growth during the twentieth century as a result of a quick expansion of the entire scientific research (Fig. 2). In this respect, disciplines such as biochemistry, physiology or microbiology gained an essential role in pharmacological research.

Change of paradigm: the drug receptor theory

In 1905, a British physiologist, John N. Langley, formulated the concept of "receptive substances" that mediates the drug

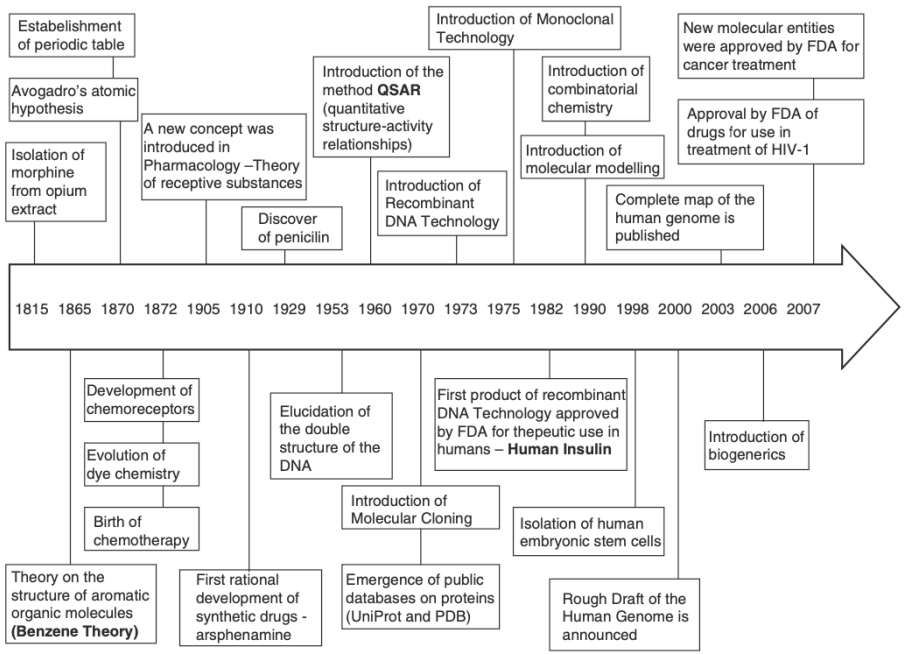


Figure 2: Timeline of some important moments on drug discovery history. Figure extracted from (Pina *et al.*, 2009)[8]**.**

action.[12] It transformed the entire concept of drug interaction, postulating that receptors are not only capable of accepting drug signals but they also generate responses that are going to be directly translated into specific biological effects.[4,8] This was the first receptor theory of drug action. A few years later, in 1933, Alfred Joseph Clark took a step forward and he was the first one to quantify biological responses produced by the administration of drugs.[4,13] He observed that for several drugs the relationship between their concentration and biological effect follows a hyperbolic curve.[4,13] Furthermore, he proposed that this hyperbolic curve (known as dose-response curve) is the result of the directly proportional relationship between the number of receptors occupied by a specific drug and the pharmacological effect produced.[4] This concept proposed by Clark will be named as occupancy theory.

Around the same time, penicillin was discovered in 1929 by Alexander Fleming, although its therapeutic use did not occur until 1940.[4] The discovery of penicillin offered an important advance in bacterial research and microbial drug discovery, allowing the development of many of the current antibiotics during the first half of the twentieth century.[5]

The occupancy theory was adjusted by Everhardus J. Ariëns and Robert P. Stephenson in 1954 and 1956, respectively. Ariëns added the concept of ‘intrinsic activity’ that allows the existence of partial agonists to produce a reduced response in

a full receptor occupancy situation (Fig. 3).[14] Stephenson introduced the concept of 'intrinsic efficacy'. He proposed that not all agonist drugs stimulate the receptor in the same way. So the response and the receptor occupancy could be completely dissociated (Fig. 3).[14] These adjustments on the occupancy theory clearly differentiated the affinity, which is the attraction between a drug and a receptor, from the efficacy, which, as mentioned above, is the capacity of the molecule-receptor binding to produce a specific biological response.[4]

About the same time, a crucial step in drug discovery was taken by Raymond P. Ahlquist. Following the Clark theory introduced a few years before, he suggested in 1948 that there are two types of adrenergic receptors, alpha and beta, and he proposed that they can be separately activated and, consequently, produce a different effect.[5,15] However, it will not

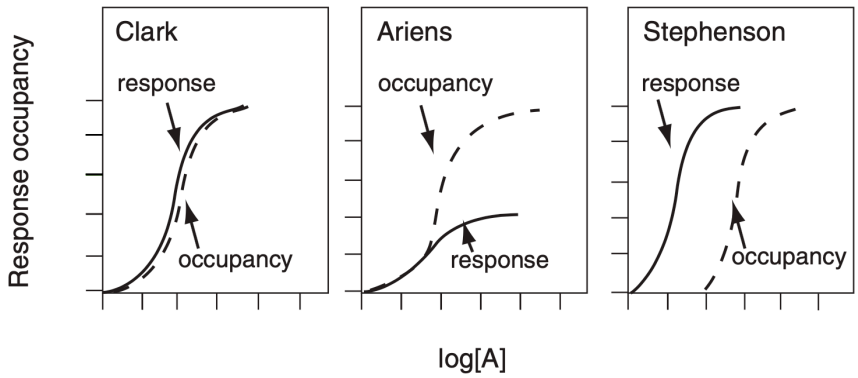


Figure 3: Illustration of different occupancy-response relationships based on the different contributions of Clark, Ariëns and Stephenson. Figure extracted from (Kenakin, 2008)[14].

be until 25 years later, in 1963, that James Black would discover propranolol, the first selective beta-blocker and also the first to be widely marketed. A structural modification of the early pronethalol, which consist in an insertion of an oxymethylene bridge into the arylethanolamine group, confers to propranolol a great affinity to beta-adrenoreceptors and a lack of affinity for alpha-adrenoreceptors. This was the first time that a wanted selectivity of a drug was preconceived and generated and it changed the entire pharmaceutical industry, introducing a new strategy approach to drug design: target-based drug discovery.[5,15]

Target-based drug discovery

Target-based drug discovery, also known as reverse pharmacology, is an hypothesis-based approach that tries to focus primarily into a defined molecular target for which a main role in some specific disease or condition is expected.[16] This was an important change of paradigm for the whole pharmaceutical industry, leaving behind the classical pharmacology (also known as forward pharmacology) with an observational approach, where the specific target interactions were not taken into account at the starting point of the drug discovery (Fig. 4).[17,18]

The scientific and technological advances on several disciplines, such as molecular biology and genetics, were extremely important in order to facilitate the consolidation of this new drug discovery paradigm. Molecular biology allowed to study disease processes at the molecular and genetic levels, leading us to determine which are the most relevant targets for a specific drug to exert the desired therapeutic effects.[5]

In 1972, Paul Berg produced the first recombinant DNA. It was based on the isolation of DNA sequences from specific species and the subsequent insertion into a vector or host cell. Moreover, this discovery was one of the firsts steps of molecular cloning, that will use this knowledge in order to make identical copies of a specific fragment of DNA as many times

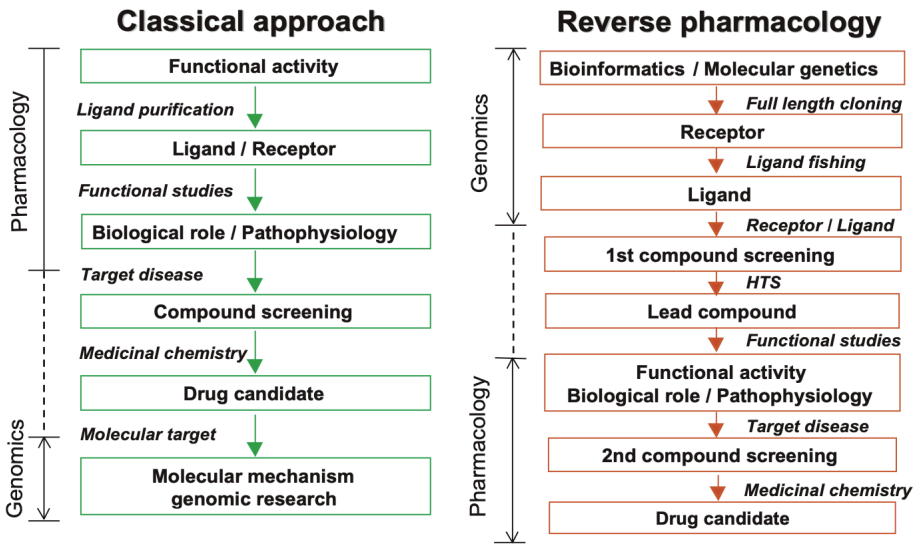


Figure 4: Comparison between Classical or forward, and reverse pharmacology approaches in drug discovery. Figure extracted from (Takenaka, 2001)[17].

as wanted.[19] Ten years later, in 1982, the first product of recombinant DNA for therapeutic use in humans, the human insulin, was produced.[20] In the same year, the polymerase chain reaction (PCR) technique was invented, allowing to automatize and facilitate this cloning process. It was a major breakthrough into molecular biology and therapeutic drug target detection, promoting a huge increase on the number of targets available.[21]

The impact of molecular biology on drug discovery and the consequent rise and improvement of target detection and structural studies techniques stimulated the emergence of structure-based drug design (SBDD).[22,23] One of the most used techniques was X-ray crystallography. This technique is based on Bragg's theory of X-ray diffraction[24] and was firstly observed

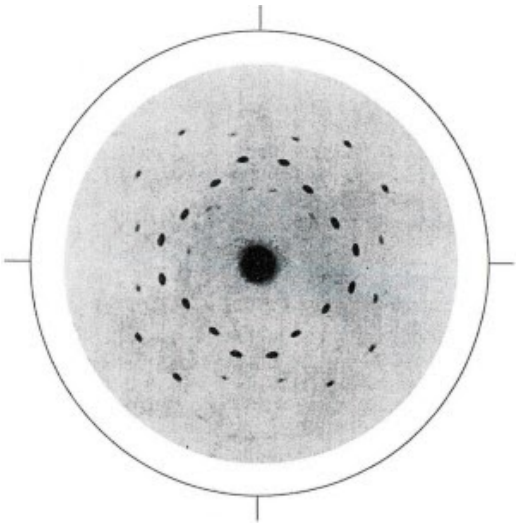


Figure 5: First photo of an X-ray diffraction from ZnS structure. It was shot by Max Laue. Photo extracted from (Thomas, 2012)[25].

by Max Laue a few years before it was widely used (Fig. 5).[25] SBDD puts the drug design focus on the structural analysis of the target and how a specific drug structurally interacts with it.[26] It allowed to study which are the best molecular properties for a drug to interact with a specific wanted target and produce the desired therapeutic effect. Structured-based drug design became the central paradigm of drug discovery for the next decades, allowing the exploration of more molecules and targets.

In parallel with the increase of the number of available targets, chemical synthesis of new compounds also underwent a great transformation. The improvements of structural techniques, such as X-ray crystallography, and the consequent increase of the number of available compounds allowed the creation of big chemical libraries with a large number of structurally diverse compounds.[21,27] These chemical libraries can be simultaneously screened for single interactions with targets. This was an extremely fast procedure that allowed the selection of active molecule hits more appropriately.[28] This new and faster approach was called combinatorial chemistry and it was introduced by Furka in the late 1980s.[29]

All these advances improved our knowledge of chemical space, both in terms of coverage and completeness, expanding the number of targets and compounds available and the information about their interaction. This knowledge

expansion led to the creation or enlargement of protein and chemical databases that, in the next decades, will grow exponentially allowing the consolidation of a new drug discovery era based on the rise of computation and big data analysis.

Drug polypharmacology

Drug receptor theory formulated and adjusted during the second half of the twentieth century brought the conception that drugs needed to be binding selectively to their physiological targets.[30] Although important advances on the absorption, distribution, metabolism and excretion (ADME) profiling of drug candidates had been achieved, the concept of "one drug - one target" did not change.[30] In this respect, although new evidence demonstrated that drug activity is modulated through some drug metabolism targets, it was still globally considered that ADME interactions did not interfere with the occurrence and the progression of the disease.[27,31]

The observed lack of selectivity and promiscuity of certain types of targets being able to interact with totally different drugs offered a new perspective regarding drug activity. In the early 20th century, a new concept emerged in order to describe the fact that drugs could often interact with more than one protein target.[32] This concept is referred to as polypharmacology and

it was first used by Julian Blagg in 1997 to refer to this lack of selectivity.[31] In fact, although some drugs are extremely selective for specific targets, nowadays we know that they are a minority and most small molecule therapeutics have a more complex spectrum, interacting with multiple targets.

Interestingly, drug polypharmacology was not only associated with off-target proteins (undesirable promiscuity), but in some cases this drug promiscuity is needed to obtain the therapeutic efficacy observed, which was only related with the primary target interaction until now.[30] One of the first scientists to associate the concepts of drug promiscuity with the therapeutic efficacy was Bryan Roth in 2004.[33] He observed that most common central nervous system disorders are polygenic and selective drugs were not effective. He proposed the design of selectively non-selective drugs in order to target several molecular proteins and treat these disorders in a more holistic way.[33] Accordingly, the classical metaphor of drugs acting as 'magic bullets' changed into a 'magic shotgun' analogy (Fig. 6).[33]

There have been several attempts to predict a global map of drug polypharmacology. However, there was an incredible lack of information of lots of non-interested targets, leading into a completely biased observation.[34] Until then, pharmacological profiles of drugs were based on these targets of interest for the study but not for the entire set of existent targets.[34] Moreover,

only an extremely low percentage of data generated in the scientific studies was available in a final publication, and most of the time with the same well-known targets. So, there was not only a lack of data available, but that data was completely scattered in the literature.[35] Accordingly, several computation

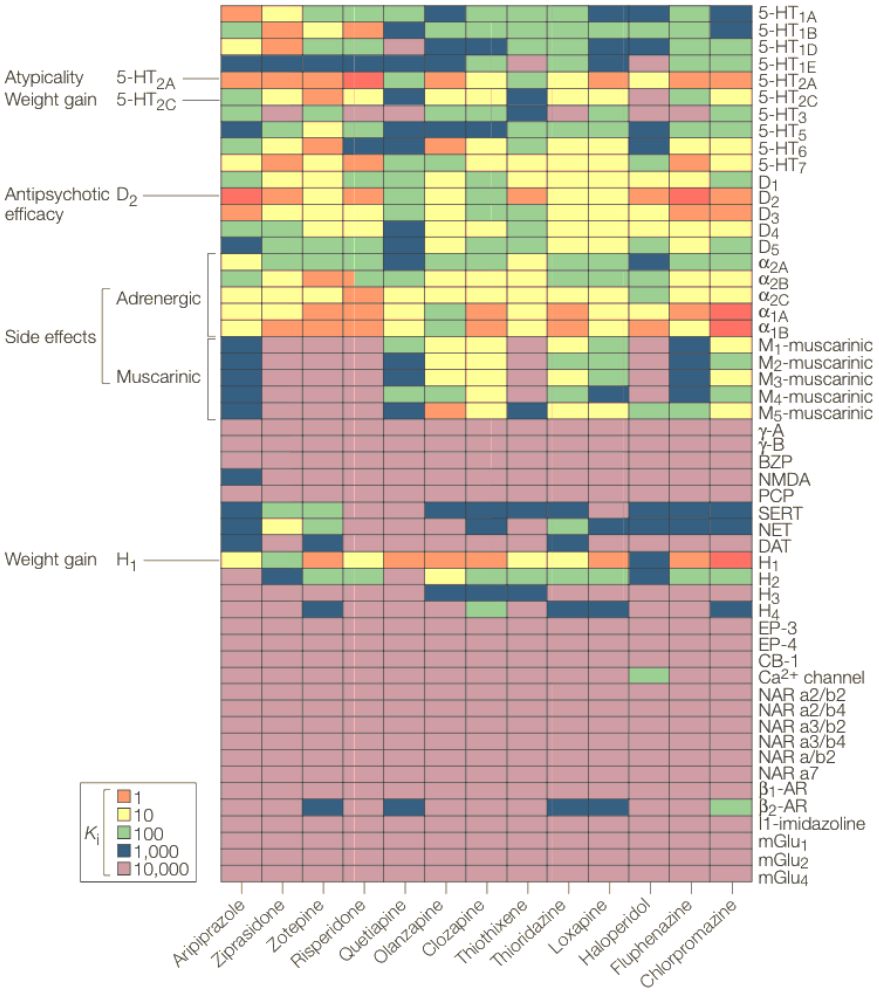


Figure 6: Polypharmacology profile of 13 antipsychotic drugs. Their affinity values for each one of the 53 targets is plotted in different colors. Extracted from (Roth *et al.*, 2004).[33]

efforts were initiated to complete and integrate some of the existing data in several databases. In parallel, new screening methods are still emerging and several attempts have been made in order to screen both new and known compounds across a large number of orphan targets.

Drug polypharmacology has revolutionized the entire pharmacology field and the drug discovery industry, introducing a more complex and holistic view and opening a new horizon in drug discovery.

In silico drug discovery

Nowadays, computational tools are an essential part of any scientific laboratory but this happened mainly during the last 70 years. ENIAC (Electronic Numerical Integrator and Computer) was the first electronic computer and was created in 1945 in the University of Pennsylvania (Fig. 7). Since then, the computing power grew exponentially, allowing to solve increasingly computation-intensive problems.[36]

Computational chemistry, later subdivided into chemoinformatics and molecular modelling, was born in the late 1950s with the work of Louis Ray and Russell Kirsch describing an algorithm for substructure searching in digital computers in 1957.[38] In the 1960s, several chemical databases

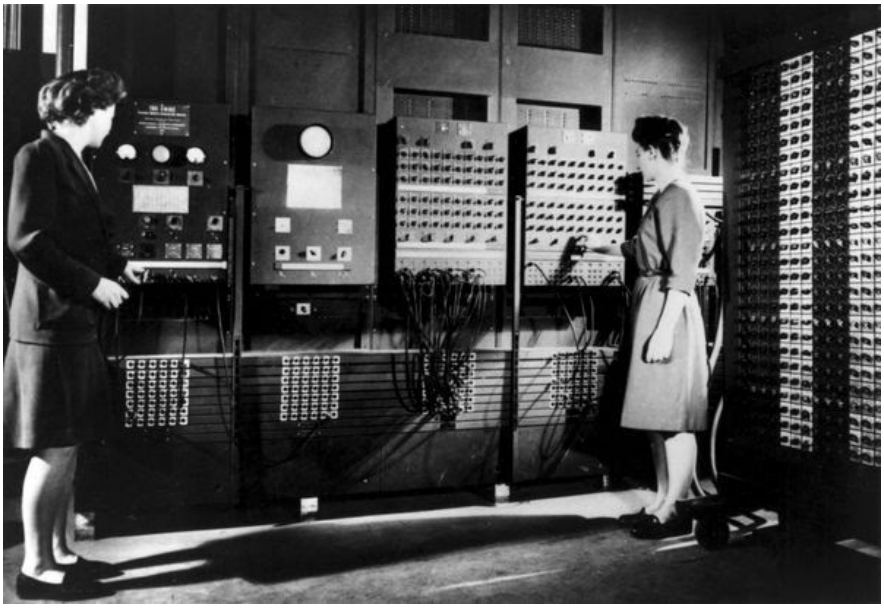


Figure 7: ENIAC control panels. The ENIAC programmers Frances Bilas and Betty J. Jennings can be observed standing next to the electronic computer. Photo extracted from CHM Revolution website.[37]

and computational tools started to emerge, such as the Chemical Abstracts Service (CAS) Registry System, which developed a unique index number to identify chemical entities.[39] In 1963, G. Vleduts proposed that computers could be used to suggest sequences of reactions that could result in a desired synthetic target.[39,40] This was the first step of computer-aided drug design (CADD). This discipline uses computational approaches to evaluate the potential activity of specific chemical entities against the target(s) of interest. In drug discovery, CADD is extremely useful to reduce costs and time, minimizing the number of molecules that should be tested

in preclinical *in vitro* and later *in vivo* assays.[41] Accordingly, CADD approaches are now widely used in modern drug discovery.

In 1962, Corwin H. Hansch published the first computational correlation study between the biological activity and the physicochemical parameters of a molecule.[42,43] This was the first time that a quantitative structure-activity relationship (QSAR) model was implemented.[44] With the appearance of advanced computational, statistical and mathematical tools during the following decades, QSAR rapidly became an essential part of any drug discovery process and pharmacological research. It was not until the 1980s that the use of X-ray crystallography to obtain three-dimensional (3D) structures of molecules allowed the introduction of 3D-QSAR models.[45] The first two 3D-QSAR methods to be presented were comparative molecular field analysis (CoMFA) and hypothetical active site lattice (HASL).[39] These new 3D-QSAR models allowed the introduction during the 1980s of structure-based drug design (SBDD), as mentioned in the previous section.[45]

A few years before, in 1971, the Protein Data Bank (PDB) was created by Walter Hamilton as a computational database in order to store information of experimental protein structures obtained by both X-ray crystallography and Nuclear Magnetic Resonance (NMR).[39,46] This database, that initially contained

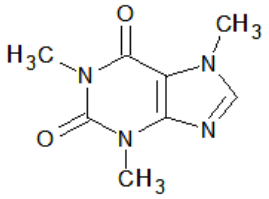
just seven protein structures, is nowadays one of the most used macromolecule structural databases, containing more than 170.000 entries.[47]

In 1982, Kuntz introduced a new computational approach in order to study protein-ligand interactions, named DOCK.[48] This was the first method that allowed to analyse the degree of complementarity between a protein binding site and a ligand.[39]

Just a few years later, Tim Berners-Lee created one of the most powerful inventions of computer science: the world wide web system (WWW).[49,50] It rapidly became a globally used system around the world that revolutionized almost all aspects of human life.[49] This technology allowed the creation and spread of lots of bioinformatic resources such as the GenBank database[51] (1992) or the EMBL Nucleotide Sequence Data Library[52] (1993), which was the first nucleotide sequence database.[49] Other important databases were created in the next decade, such as the well-known NCBI databases PubMed (launched publicly in 1997), Human Genome (1999), or other online platforms such as DrugBank[53] (2006) or ChEMBL[54] (2010) databases. Moreover, lots of bioinformatics tools became completely accessible and sometimes with a graphical interface to facilitate usage across the entire scientific community.[49] Internet quickly became not just a completely new source of resources and tools, but the most important library of scientific literature.

The consolidation of new compound databases and the exponential growth of them promoted the establishment of new molecule classification systems, such as the Universal Protein code (UniProt) for proteins, in 2002, or the International Chemical Identifier (InChI code) for small molecules, created by the IUPAC organization in 2005 (Fig. 8).

The increasing amount of data available nowadays and the computational power provided by the new modern supercomputers allow us to compute extremely complex problems in relatively short periods of time and to provide analytics tools for all these big heterogenous data. In this regard, in the last decade, several big data databases



InChI=1S/C8H10N4O2/c1-10-4-9-6-5(10)7(13)12(3)8(14)11(6)2/h4H,1-3H3 (caffeine)

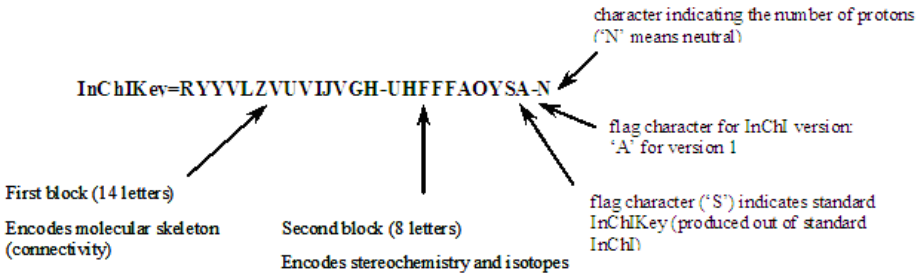


Figure 8: InChiKey nomenclature of caffeine molecule. Each part of the inChiKey nomenclature is shown. Figure extracted from Inchi-trust website.[55]

appeared, such as the FDA Adverse Event Reporting System (FAERS), in 2012, to collect spontaneous drug safety reports, or SureChembl, in 2016, to store chemical data extracted from patents.[56]

In the last few years, advances in machine learning are extremely promising and the potential uses of them in the drug discovery process and personalized healthcare are just being observed. Machine learning can be defined as the implementation of algorithms that can automatically improve through data, detect patterns on it and use them to predict future outcomes or make decisions.[57] The most widely used machine learning techniques in drug discovery and pharmacological studies are support vector machine (one of the most used classifier) (Fig. 9), random forest, or artificial neural network, which are algorithms that, given a set of parameters, can analyse the possible combination of them in order to predict the potential results.[58] Some examples of the utility of machine learning in drug discovery and development include prognosis of medical events based on patient records, new assessment methods in order to predict pharmacokinetic, ADMET and efficacy properties, prediction of target structure and inference of ligand-protein interactions and adverse drug reactions.[57,59]

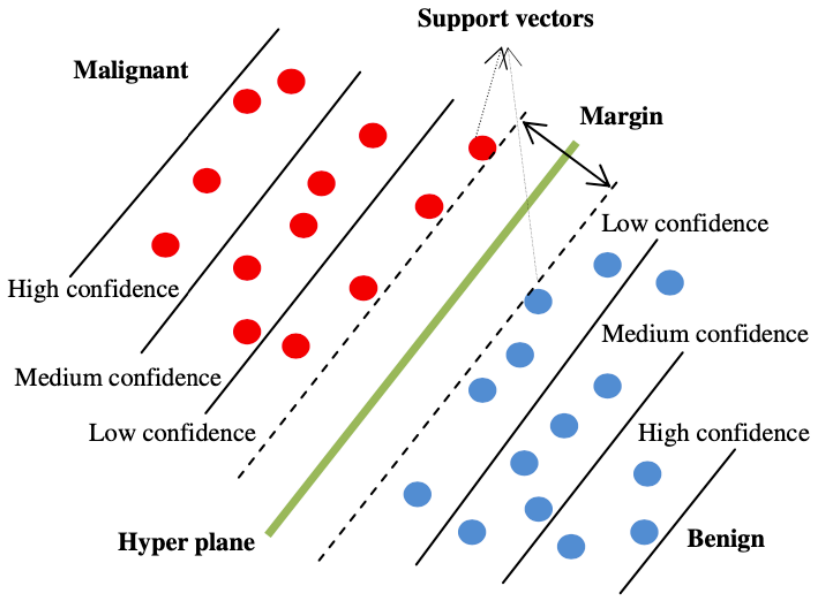


Figure 9: Illustrative example of a Support Vector Machine classifier used to differentiate between tumour and benign tumours. Extracted from (Khazendar *et al*., 2014).[60]

Pharmaceutical business

Since 1980, the pharmaceutical industry has experienced significant growth, both in revenues and investment, although the number of new drugs that are approved remained constant.[61,62] The exponential establishment of a capitalist economic model covering almost the entire world and a globalization process that expanded and universalized the markets of the late twentieth and early twenty-first century, promoted this extraordinary expansion of the pharmaceutical industry.[63]

The increasing cost and complexity of drug development contributed to transform the pharmaceutical companies, via fusions and acquisitions, into the big players that they are nowadays. As a result, 73% of the total pharmaceutical sales during 2015 were attributable to the 25 largest pharmaceutical companies.[64] Thus, a few big companies control more than 2/3 of the new drugs approved, and the same trend is being observed in the last twenty years.[65] There are certain risks associated with this market monopolization, such as the increasing cost of new therapeutics, the difficulties to find treatments for rare diseases or the increasing power and influence of the owners of these companies.

Big pharma not only achieved a huge influence in society but also managed to become one of the most profitable business models, raising concerns about its profits and prices.[66] A recent study compared 35 large pharmaceutical companies with 357 large non-pharmaceutical companies just to find out that the gross profit of the first was significantly higher than the latter.[64] Not surprisingly, and due to the huge profit expected, some reports found that the number of marketing organizations that control new molecules, although they have little to do with drug discovery, is increasing.[65]

Moreover, big pharmaceutical companies are benefiting from public investment, both directly and/or indirectly. For instance, NIH along with other public organizations was responsible for

33% of biomedical research funding from 2003 to 2008, assuming some of the so-called risk of failure of the pharma companies.[66] This partnership rarely contributes to lowering the prices of new therapeutics in the market. Hence, the pharmaceutical model is becoming a huge profitable business that probably will merge with other companies to develop massive capitalistic conglomerates.

Drug discovery and development process

Drug discovery and development process (Fig. 10) experimented several changes in the last decades. Complexity and risk involved in the development and marketing of a new drug make it a very cost-effective and extremely time

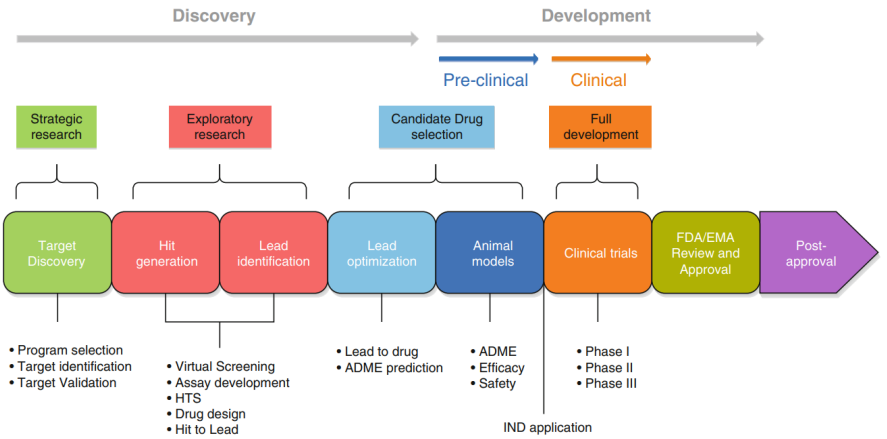


Figure 10: Drug discovery and development process representation. It shows the different subphases and some of the most important aspects of each one of them. Figure extracted from (Duelen *et al.*, 2019).[41]

consuming (15 years on average) process (Fig. 11). All the research and development needed for a drug to reach the market can cost up to $1 billion.[67] Although this process has undergone some changes in the last decades, several steps have been well established and remained quite invariable.

The first step in drug development is to choose what condition or disease is wanted to improve and identify which target is the one that requires activation, modulation or inhibition to revert or improve a certain pathology.[41] The pharmacological and biochemical background on the pathology, knowledge on marketed drugs and target profiling are essential to select an appropriate target candidate.[68] However, the step of target validation is absolutely crucial to ensure that the target is clearly associated with the disease and to confirm its druggability. This process is extremely important because a

	Target to Hit	Hit to Lead	Lead Optim	Non-Clinical	Phase 1	Phase 2	Phase 3	Sub to Launch
# per Launch	24.3	19.4	14.6	12.4	8.6	4.6	1.6	1.1
P(TS)	80%	75%	85%	69%	54%	34%	70%	91%
Cycle time (yrs)	1.0	1.5	2.0	1.0	1.5	2.5	2.5	1.5
Cost/launch ($mil)	$94	$166	$414	$150	$273	$319	$314	$48

Figure 11: Table of estimated resources expended in a drug discovery and development process until the launch of the drug. Figure adapted from (Mohs and Greig, 2017).[69]

bad decision could be translated into an enormous waste of time and money.[67]

The second major step in a drug development process is the hit generation. It is based on the identification of active molecules (hits) with a confirmed *in vitro* activity for the target.[68] A large variety of screening methods exist to identify these hits. Among them, high throughput screening (HTS) is based on a robotically testing a large chemical library against a specific target or phenotype. In order to reduce the cost and to take advantage of the existing knowledge, a more focused screening can be performed by preselecting molecules likely to interact with the wanted target. To aid in this preselection, other more specialized screening methods can be used, including structured-based drug discovery, that uses X-ray crystallography to design molecules based on their structure, physiological screening, a more holistic approach that evaluates drug effects at a tissue-level, and virtual screening, that exploits the concepts of molecular similarity and computational docking to predict what compounds may interact with a given target.[68] In the event that hits are identified, an exhaustive evaluation of their biochemical, pharmacokinetics, ADME and physiological properties is required.

The next crucial step of drug discovery is the hit optimization. To this aim, a hit-to-lead process is necessary to optimize all

properties of the hit that require improvement, including pharmacokinetics, ADME profile, and target efficacy and selectivity. Once the criteria for optimal properties are met, the molecule(s) will progress in the process to obtain *in vivo* proof of principle. At the end of this hit-to-lead process, the one, or few, lead(s) that met the optimal properties may be nominated to become a drug candidate.[5]

After hit optimization, a second optimization process of those leads identified is carried out in order to improve any remaining deficient properties without affecting the favourable ones.[68] Moreover, a thorough characterization of these leads is necessary before considering them as possible drug candidates. In general terms, this lead characterization is based on models of genotoxicity, *in vivo* models of general behaviour, high-dose pharmacology, PK/PD studies, dose linearity and repeat dosing PK.[68] Compounds that are successfully characterized are considered possible preclinical drug candidates. It is important to mention that the lead-to-drug process does not stop with the first drug candidate. It is needed to continue exploring all the lead compounds as possible back up candidates if the other ones fail in the next preclinical or clinical steps.

The preclinical phase starts directly when a specific drug candidate is nominated. The preclinical stage includes both *in vitro* assays as well as *in vivo* animal studies. In this last one,

animal models of the human body are used in order to evaluate not only ADME, but efficacy and safety properties and to ensure that there are no major toxicity or safety issues. Once the preclinical phase is passed, the company needs to submit an Investigational New Drug (IND) application explaining the preclinical results and their future clinical plan to the appropriate regulatory agency of the country the company wants to commercialize this future drug.

After that, clinical trials may start. They are divided in three different phases. It is during this stage that the drug is finally tested in humans. In Phase 1, the drug is tested in a small number of healthy people in order to assure that the drug is safe and no major complications are associated with its administration. In Phase 2, the drug is tested in patients suffering from the pathology in order to evaluate the efficacy of the drug. At last but not least, in phase 3 the number of patients increases considerably in order to evaluate in a more exhaustive and significant way both drug efficacy and safety. Once the third phase is successfully completed, it needs to be presented into the corresponding regulatory agency that will ultimately decide if the drug is suitable to be marketed.[70]

However, drug development does not stop here. A post-marketing phase (phase IV studies) is needed in order to evaluate the real drug effectiveness in a non-interventional observational surveillance study.[71,72]

A large number of professionals from many different fields are involved in the whole course of this large and complex process. Although a huge number of resources are invested to discover and develop a new drug, the pharmaceutical industry is one of the most (if not the most) profitable industries over the world above bank, oil, gas or media industries.[64]

2. METABOLOMICS

Metabolomics is the cellular, tissular and organismal metabolism study of small chemical compounds, also known as metabolites.[73] These small molecules are the substrates and products that participate in all the biological processes involved in a living system, and include both exogenous and endogenous molecules. The complete set of these small molecules and their interactions in a biological system are known as the metabolome.[73]

Metabolomics: origin and evolution of a new omics discipline

Metabolomics was introduced for the first time in 1998 by Stephen Oliver.[74] However, metabolomics studies started some decades before. In the 1950s, several studies stressed the importance of metabolic profiling systems to explain complex biological systems in a more precise way.[75] In 1955, Donald Nicholson collected and published all the known metabolic reactions by that time, which represented the first holistic review of an organism metabolome.[75,76] These metabolic systems promoted the idea of a precision medicine, where each metabolism described in a biological system could be treated differently.[77]

In the next decades, the paradigm “DNA makes RNA makes protein” gained prominence and caused a swift expansion of the biggest omics disciplines: genomics, proteomics and transcriptomics (Fig. 12). In 1992, the first chromosome sequence, belonging to S. cerevisiae, was completely sequenced.[78] This discovery showed that the protein-encoding genes which were known by then represented only about 20% of the entire set of genes, changing the perception of the functional genome. In parallel, with the arrival of new analytic methods, such as mass spectrometry or NMR, functional analysis of proteomics helped to gain information about the metabolite role in biological systems and the metabolite profile could be finally related to the biological function.[73]

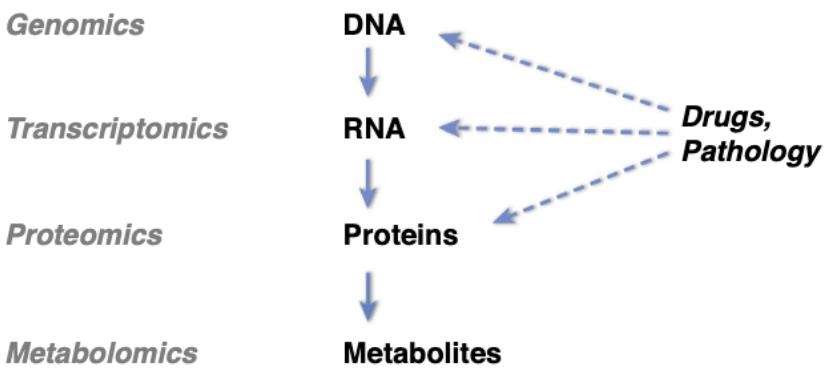


Figure 12: Conceptual flow of “DNA makes RNA makes Proteins” idea and the place of metabolomics amongst the other omics sciences. Figure extracted from (Fillet and Frédérich, 2015).[79]

In the 1990s, metabolic control analysis (firstly introduced by Kacser and Burns in 1973) was widely recognised and used.[80] This approach allowed to quantify the control distribution over concentrations and metabolic fluxes in biological systems.[81] Based on this analysis, Stephen Oliver suggested in 1998 the concept of metabolome in a functional analysis of the yeast genome.[74] He observed that concentrations of yeast endogenous metabolites vary constantly and he proposed that this variation allows to maintain fluxes through the metabolic networks in a constant way.[73]

At this point, metabolomics made a leap forward and became an important discipline in biological research. The incredible advances in mass spectrometry and chromatographic techniques, as long as the development of computational techniques, such as *de novo* approaches, allowed the identification and classification of metabolites and the determination of their interactions and roles in biological systems.[82] Some of these *de novo* techniques predict pathways based on the known information of reactions and hypothesize what intermediate metabolites are needed to complement metabolic pathways. Despite the continuous characterization of new metabolites, its coverage is still low compared to the estimated complete metabolite map (Fig. 13).[83]

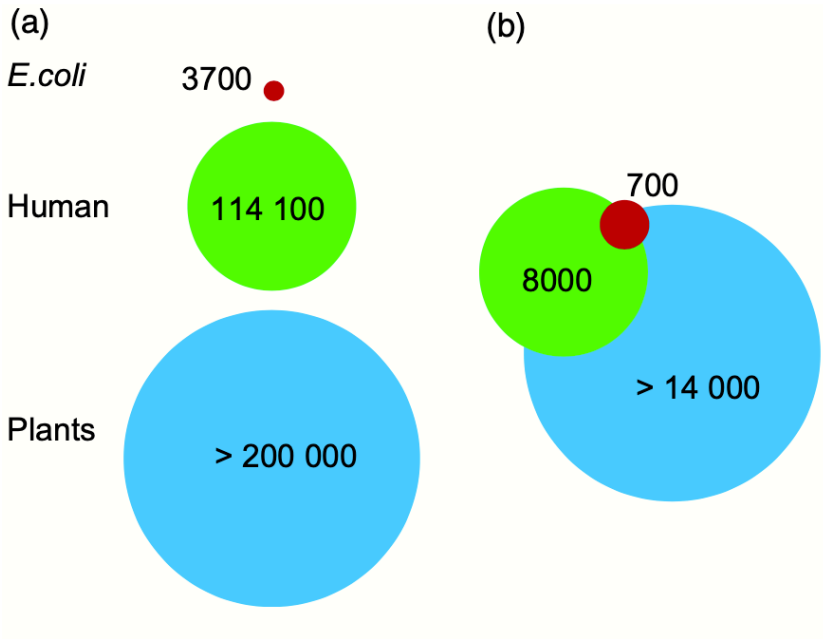


Figure 13: Representation of the a) estimated number of metabolites present in *E. cole*, Human and Plants and b) the approximate number of metabolites that can be measurable nowadays. Figure extracted from (Alseekh and Fernie, 2018).[83]

The substantial increase in the amount of metabolomic information (including metabolite compounds, metabolite pathways and metabolite profiling) led to the creation of several metabolomic databases in order to annotate and classify all the data. For example, in 2007 the Human Metabolome DataBase (HMDB) was created with information on more than 114,000 different metabolites that could be found in the human body.[84,85] At the same time, other databases started to include information about metabolites to complement their data, such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) PATHWAY database.[86] KEGG PATHWAY is a repository of

metabolic pathway maps representing the available knowledge of the molecular interaction, reaction and relation networks for both metabolic, genetic, environmental or drug information. An interesting effort has been made by KEGG developers to focus on representing disease pathway maps as perturbed states of the molecular network system. In these systems, genetic, environmental or metabolic factors, along with drugs, are considered as perturbants.[86] Another example of the growth of metabolic data online repositories is the public IUPHAR/BPS Guide to PHARMACOLOGY database, which is a curated source of ligand-activity-target relationships that includes information of more than 10,500 ligand, including 509 metabolites, 802 endogenous peptides or 326 natural products.[87,88]

In recent years, the consolidation of computational science and the emergence of new methodological approaches are inspiring new scientific fields to incorporate metabolomics in their research lines.

Metabolomics, drug discovery and personalized medicine

In the last decades, the generalized idea that almost all chronic and long-term diseases have an eminently genetic origin was widely extended. However, genome sequencing and single nucleotide polymorphism (SNP) characterization have

revealed that the association between disease and a specific genetic origin is less stronger than expected.[89] Therefore, a change of paradigm was needed in order to incorporate other important aspects of disease aetiology such as epigenetic, microbiomic or environmental factors. Metabolomics substantially helped to clarify the chemical causes of some important chronic diseases such as atherosclerosis, diabetes and some types of cancer. Therefore, metabolites have an important impact in the appearance and progress of diseases.[89]

For example, several studies suggested that type 2 diabetes is related not only with genetic and environmental factors, but also with the levels of certain types of amino acids. It is suggested that high levels of isoleucine, leucine, valine, phenylalanine and tyrosine could be associated with an increased risk of developing type 2 diabetes in the short-to-long term.[90–92] This example illustrates how metabolomics can lead to the identification of new disease biomarkers. The metabolomics ability to predict specific metabolic conditions that are related with the aetiology of specific diseases is leading to the idea of drug treatment into a more personalized medicine. Although the application of individual metabolomics is still limited, and its true impact is yet to be demonstrated, the usability of metabolomics in precision medicine seems to be clear and it is expected to increase in the next decades.[93]

Within this context, metabolite-based drug discovery (MBDD) emerged as a new possible approach to identify and test new therapies (Fig. 14).[89] In fact, a new concept known as pharmacometabolomics appeared in order to study both pharmaceutical drugs and endogenous metabolites with a metabolomic approach.[94] In drug metabolism, a huge number of enzymes, proteins and organs are involved, making the whole process extremely complex and variable depending on age, gender, diet, weight and even locations. Metabolomics allows to monitorize drug responses and, in line with the

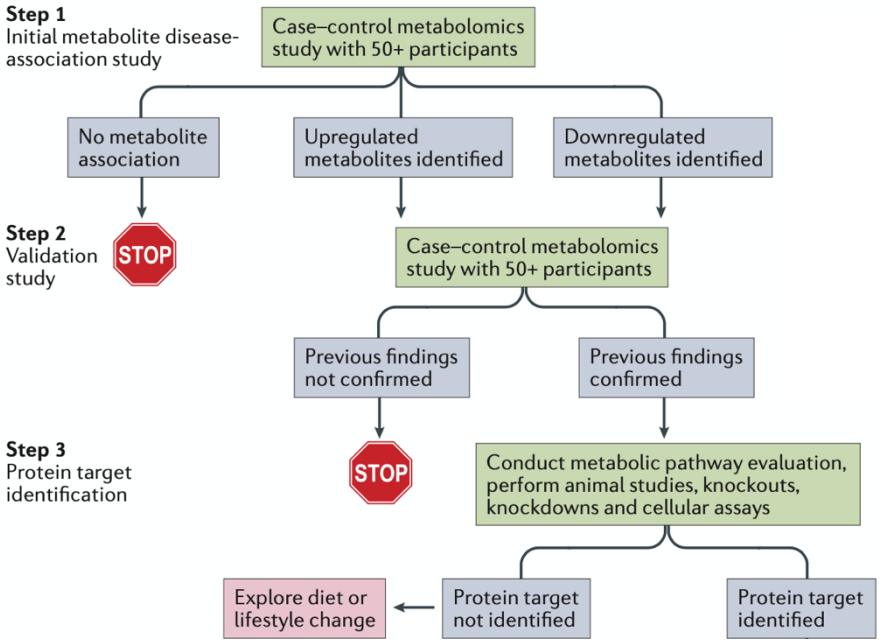


Figure 14: Example decision tree of the first steps of a metabolite-based drug discovery process. Figure adapted from (Wishart, 2016).[89]

previously mentioned personalized medicine, be able to customize drug dosing.[89]

In this regard, a recent study suggested that the metabolome could be an incredible basis for a more individualized medicine.[95] Applying machine learning algorithms, they tried to predict metabolite levels in individuals on the basis of gut microbiome, clinical parameters, diet, host genetics, lifestyle or anthropometric measurements.[95] They presented a significant prediction power of the profiled metabolites. This is a great advance towards a mechanistic understanding of metabolome and its possible alterations, leading to a possible use in personalized medicine that could design interventions in order to modulate the metabolite levels and to adjust these alterations towards improving the undesired conditions.[95]

In spite of the high hopes put on metabolomics over the last 30 years, its true level of impact on the entire drug discovery and development process has yet to be fully determined. However, there is no doubt on the positive implications that this omics technology is having over the entire scientific community.

3. DRUG SAFETY

Drug safety refers to the study of the safety of drug candidates and medicines throughout the drug discovery and development process that aims to increase the benefits and minimize the risks of drug administrations.[96] Once the drug reaches the market, the field of pharmacovigilance includes a number of different approaches and techniques to perform target safety assessments and postmarketing surveillance studies. Nowadays, drug safety and pharmacovigilance are essential pillars of the drug discovery and development process, and they represent an enormous effort on time and cost for the pharmaceutical industry. However, the study of the safety of drugs started even before the big pharmaceutical companies were born.

Origin of pharmacovigilance and safety pharmacology

Unfortunately, the history of pharmacovigilance is full of big scandals and deaths, that motivated and precipitated changes in legislations, lifestyle and marketing. In 1848, the death of a Hanna Greener, an English woman, after a chloroform administration motivated an investigation to understand the relation between the use of this substance and its death.[96] This study preceded a series of investigations regarding more deaths related to chloroform that The Lancet finally published

in 1893.[97] This was one of the first publications of a pharmacovigilance study relating a specific adverse effect associated with drug administration.

Few years later, in 1906, the US Food and Drug Administration (FDA) was created with the mandate to control the safety, efficacy and security of drugs, food compounds, veterinarian products and some other products.[98] Furthermore, in 1938 the FDA administration approved the Federal Food, Drug and Cosmetic act, which established that drug safety needs to be tested, validated and finally approved by the FDA before drug marketing can take place. This new law was approved after

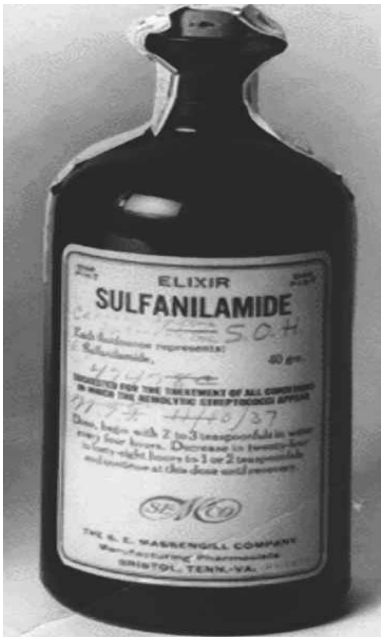


Figure 15: Original 1-gallon bottle of sulfanilamide elixir. Photo extracted from Drug Times web.[99]

what is known as ‘The Sulfanilamide Elixir disaster’. This event occurred in 1937, after some studies suggested that sulfanilamide, in powder or tablet forms, was able to revert streptococcal infections. However, in its liquid form, sulfanilamide was dissolved in diethylene glycol, an extremely toxic substance (Fig. 15). As sulfanilamide was already approved in other forms, the liquid solution was distributed without pharmacology studies. After a few months and more than 100 deaths, the real nature of this substance was detected and the product was withdrawn from the market.[98]

Another crucial change was conducted in the 1960s in response to another tragedy. In 1961, William G. McBride and Widukind Lenz published, in parallel, alarming studies in which they warned of a correlation between the administration of thalidomide and newborn congenital malformations (Fig. 16).[100,101] Thalidomide was approved in the late 1950s and was widely prescribed to pregnant women as an antiemetic until 1962, when it was withdrawn from the market. Although it is difficult to figure out how many babies were affected, it is estimated that about 100,000 children worldwide were born with some kind of malformations that in so many times led to a premature death.

The thalidomide tragedy brought different regulatory revisions and a global change in the entire drug discovery and development process was adopted worldwide. One of the most

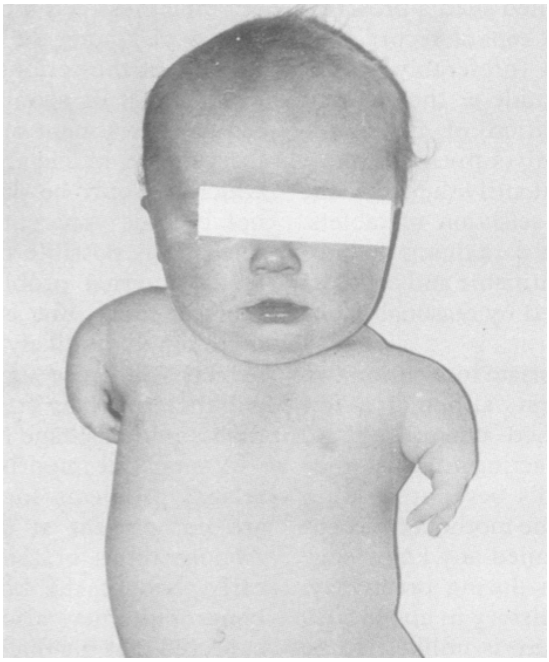


Figure 16: newborn deformities produced by Thalidomide effects. Figure adapted from (Smithells and Newman, 1992).[102]

important regulatory advances was the consolidation of the animal testing in drug safety, making it mandatory before the administration in humans.[96] During the next decade, several countries included rigorous toxicity testing approaches in the non-clinical research of new drugs and exhaustive postmarketing surveillance for new drug was also mandatory.

From that moment on, several epidemiologic studies were done to evaluate the possible adverse reactions of drugs. The first one to be done was guided by the Boston Collaborative Drug Surveillance Program, created in 1966.[96] As previously seen, in the 1980s, it was demonstrated that drugs could also

interact with other targets apart from those mainly responsible for its mechanism of action.[30] Those secondary targets may be often the cause of adverse drug reactions or drug side-effects. Accumulating evidences in this regard led to the acceptance that drugs were not as selective as once thought. With the concept of drug polypharmacology, safety pharmacology was implemented.

Polypharmacology and safety pharmacology: A new pharmacovigilance paradigm

The International Conference on Harmonization (ICH) defined this discipline as “those studies that investigate the potential undesirable pharmacodynamic effects of a substance on physiological functions in relation to exposure in the therapeutic range and above”.[103] Safety pharmacology was based on the idea that organ functions are also drug targets and the principal cause of adverse reactions in humans.[104]

One of the first countries to promote organ function assessments was Japan. These first evaluations included different organ categories in order to prioritize their relevance and to suggest what kind of studies were required. Rapidly, during the 1990s, the Japanese guidelines were adopted by the entire pharmaceutical community as the main organ function assessments.[104] In 1990, Europe, United States and

Japan joined efforts to create the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). This initiative was founded to promote the development of unified technical and registration guidelines for pharmaceutical drugs that could be used and accepted worldwide. However, it was not until the year 2000 that ICH launched for the first time a guideline for safety pharmacology, named ICH Topic S7A.[104]

ICH Topic S7B incorporated an *in vitro* pharmacological assay on Potassium ion channel subunit alpha (IKv11.1) also known as hERG, as an absolute requirement for a new drug to be approved.[105] This was the first time a target-based study was included as a mandatory requisite for a drug to be approved. In 1990, hERG was found to be related to the appearance of a deadly ventricular tachyarrhythmia, called Torsades de

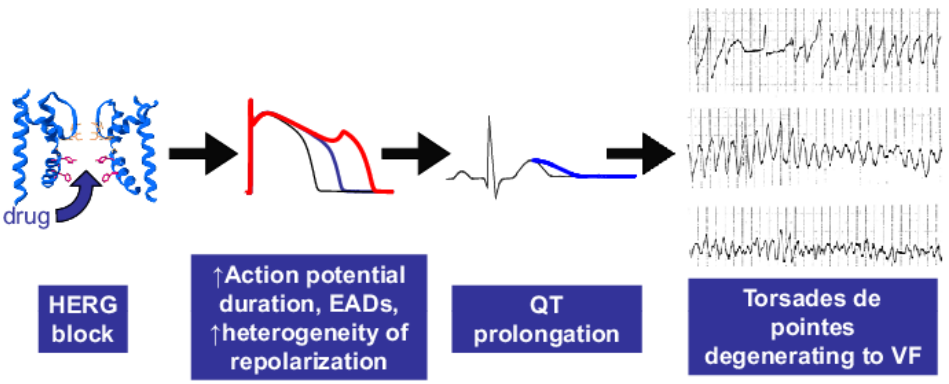


Figure 17: Illustration of a drug induced Torsade de Pointes through a drug induced HERG block like Terfenadine. Figure extracted from (Kannankeril, *et al.*, 2010).[109]

Pointes (TdP). This association was discovered due to the increasing number of reports of deadly patients that took terfenadine and the appearance of TdP during the 1980s.[106,107] Terfenadine was an antihistaminic drug approved for the treatment of allergies.[108] As a non-cardiac drug, pharmaceutical agencies did not study the possible cardiac effects of this drug. Therefore, in 1997 this drug was withdrawn from the market and it preceded several more non-cardiac drugs to be withdrawn due to cardiac side effects. It was determined that the inhibition of hERG cardiac receptors by this drug causes a prolonged QT interval and consequently, in some cases, a deadly TdP (Fig. 17).[108,109]

Off-target interactions as a cause of adverse drug reactions.

Since the incorporation of *in vitro* pharmacology assays to test for affinity on hERG, plenty of new studies have been performed to assess any other possible causal relationships between off-target interactions and drug side effects. In addition, as seen previously, technological advances in molecular biology and pharmacology, such as the improvement of X-ray crystallography, together with the revolutionizing advent of computational approaches, led to a substantial increase in the number of targets available for

testing and, with it, the number of new active chemical substances.

Moreover, polypharmacology studies have shown that, based on the specific drug therapeutic effect, a target can be understood as primary (on-target) or secondary (off-target) to define if the drug is expected to exert its therapeutic effect on it or not, respectively. This notwithstanding, toxic or adverse effects could be related with both primary and secondary targets. Primary adverse effects are described as an exaggerated pharmacologic effect at the target of interest while secondary adverse effects are events produced as a result of the interaction or modulation of other targets.[110] Nowadays in drug discovery, a general strategy during lead optimisation is to maximise *in vitro* binding affinity for the primary target(s) in order to try to improve target selectivity and to reduce secondary affinities and thus, adverse drug effects.[111] However, the final drug candidate is usually not the molecule with the strongest on-target affinity but the one with the best equilibrium between all properties.[112]

In vitro secondary pharmacology is a cost-effective approach used by the pharmaceutical industry to anticipate potential safety issues of a new drug candidate as early as possible during the drug discovery process.[113] In 2017, a study led by Thomas Papoian proposed that safety margins based on *in vitro* binding affinities are better predictors for determining the

risk of a drug to produce a specific side effect than measures involving only the maximum therapeutic free plasma drug concentration *in vivo*.[114] In this study, they analysed the well-known relation between valvular heart disease (VHD) and a huge agonism activity on 5-hydroxytryptamine receptor 2B (5HT2B). Interestingly, they observed that the agonist drugs that do not produce VHD have activity values significantly lower (more than two order of magnitude lower) than the activity of the endogenous metabolite of this target, serotonin, suggesting that some kind of relation may exist between the minimum level of affinity needed by a drug to be active on a specific target and the endogenous metabolome activity.[114]

This interesting approach not only showed that *in vitro* secondary pharmacology is useful in order to assess the possible relationship between side effects and secondary targets, but presented a promising new approach in drug safety that may be extremely relevant in order to understand why drugs need a certain level of affinity to be active on their targets and why these levels of affinity could be extremely different between drugs interacting with exactly the same target. This was one of the first studies that connected metabolomics and drug safety in order to explain the implications that the endogenous metabolome could have in the origin of drug adverse events and it was partially responsible for this Thesis.

4. SUDDEN CARDIAC ARREST

Epidemiology of Sudden cardiac arrest

A sudden cardiac arrest (SCA) is the abrupt cessation of the heart's ability to pump blood through the body within a short time period (less than an hour) from the onset of symptoms, leading in most of the cases to an unexpected death (also called, then, sudden cardiac death).[115,116] Although classifying based on clinical circumstances can be misleading, and understanding that almost 40% of sudden deaths can be unwitnessed,[117] SCA is generally caused by myocardial infarctions, cardiac arrhythmias and/or myocardial ischaemia.[118] However, other potential causes such as aortic rupture, pulmonary embolism or drug abuse are also extremely linked to its advent.[118] Accordingly, SCA is understood as a multicausal event, where an accumulation of various and often interacting causes promotes its appearance.[119–123] Furthermore, since a big number of SCAs occur in individuals without previously diagnosed heart conditions and out of the hospital, the prevention of SCA is even more difficult.

In the last decades, incredible improvements have been done regarding the identification of causative factors, prevention and new effective treatments of SCA. However, sudden cardiac arrest is still one of the major public health issues worldwide. It

is estimated that the average survival rate in the European countries is only 10%.[118] Moreover, in Europe, it is related to more than 50% of cardiac deaths and approximately 20% of all natural deaths.[118] These parameters are very similar in the United States (USA), being the 50% of the cardiac deaths and around 12% of all natural deaths.[115] Therefore, it is estimated that approximately between 300,000 and 450,000 cases of SCA occur annually in USA.[124]

A great improvement in cardiac resuscitation has been done in the last decades that has been translated into an increased survival rate from out-of-hospital events.[118] It is crucial to reduce the delay between its appearance and defibrillation. In order to do that, a lot of effort has been put in several European countries into distributing automated external defibrillators (AED) with a fast deployment all across the territory.[118]

Risk factors identification. First step for the detection and prevention of SCA

Age is one of the main risk factors associated with SCA, growing exponentially from the age of 45 (Fig. 18). Although a peak is also observed during the early infancy (< 5 years old), the majority of SCAs cases occurred in adults after the age of 45.[119] Moreover, the causative and risk factors of SCA seem to be extremely different in young people, being

cardiomyopathies and coronary genetic abnormalities the principal cause of SCA, alongside drug abuse.[125] In terms of gender, approximately 75% of SCA events occur in men.[115] This percentage means that, in general, men have between three and four times more risk of SCA than women. However, this difference becomes smaller with age.[119]

The detection of risk factors for SCA and the identification of SCA-associated diseases become an essential part of SCA prevention. Risk factors associated with SCA could have a variety of completely different origins. As a cardiovascular

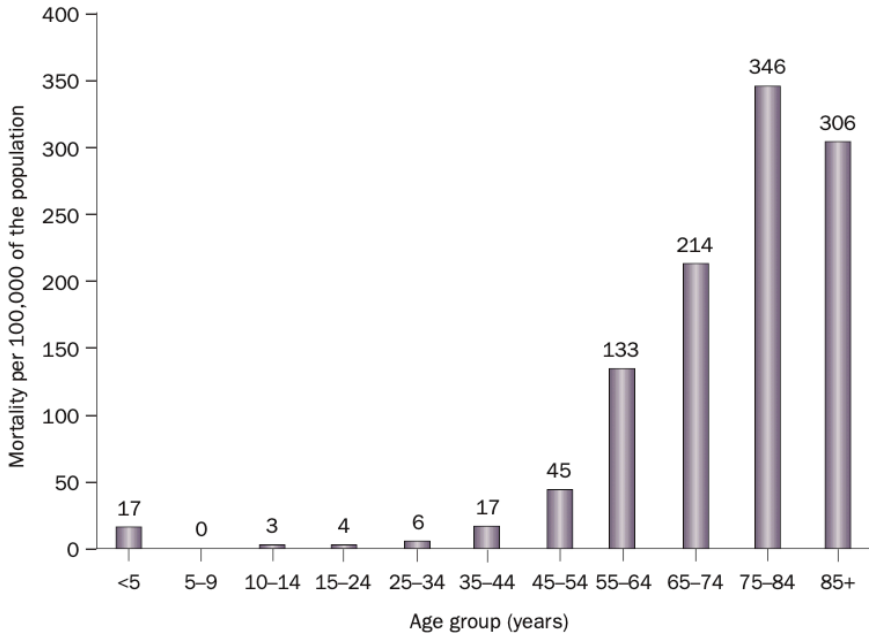


Figure 18: Sudden cardiac arrest mortality distribution by age group among residents of Multnomah County, in Oregon (USA), during an entire year between 2002 and 2003. Figure extracted from (Adabag *et al.*, 2010).[119]

disease likely being caused by some other cardiovascular conditions such as heart failure, myocardial infarction or other coronary artery diseases, it is suggested to adopt the cardiovascular risk factors as possible indicators for SCA.[126] In this regard, Framingham Risk Score (USA)[127] and HeartScore (EU)[128] are two platforms created in order to estimate the cardiovascular risk of an individual. The most important risk factors to be considered, apart from age and gender, are cigarette exposure, hypertension, obesity, type 2 diabetes or hyperlipidemia. These treatable cardiovascular factors are the basis of a conventional preventive therapy strategy (Fig. 19).[126]

However, other SCA risk factors have been suggested such as left ventricular dysfunction or hypertrophy, poor heart functional status and an abnormal electrocardiogram (such as prolonged QT interval, T-wave alternans or Tpeak-Tend conditions).[129] All these factors have been proposed as potential risk markers of SCA. As observed, almost all of these factors are measurable in anatomical screening or with clinical risk profiling (Fig. 19).[126] Their evaluation is extremely important to try to prevent the accumulation of them and reduce the risk of sudden cardiac arrest.

Furthermore, there is strong evidence of genetic predispositions that influence the risk of SCA.[130,131] The genome screening and genome wide-association studies are used in order to identify these variants that predispose SCA

individuals.[132,133] These variants can be used as genetic markers and could be effective as personalized risk predictors. One of the most relevant genetic markers associated with SCA is SCN5A, a gene that encodes the voltage-gated sodium channel type 5 subunit alpha ($Na_V1.5$). This transmembrane protein is mainly found in cardiac muscle regulating the inbound sodium flux, inducing a depolarization and the consequent action potential in the excitable cells.[134] A channel structural or functional defect, led by some SCN5A mutation, may lead to a reduction of the sodium current and a consequent arrhythmogenic syndromes, a sudden cardiac arrest and ultimately a fatal outcome. Several genetic variations in SCN5A have been related to SCA, such as the rs7626962 mutation (Ser1103Tyr), which is based on an amino acid substitution between domains II and III of the $Na_V1.5$, the rs11720524 mutation, which interrupts a specific transcription factor binding site of the gene, rs41312391, a mutation that alters the expression of another gene (MoG1) implicated in the histone deubiquitinating complexes.[129]

Another protein family implicated in SCA is the family of potassium channels. They play an important role in the repolarization process of the cardiac action potential and, consequently, in the maintenance of normal cardiac rhythm.[135] An abnormal rate of cardiac repolarization can lead to SCA.[136] Both KCNH2 (also known as hERG), which encodes the potassium channel $K_V11.1$, and KCNQ1 genes, which encode

the potassium channel $K_V7.1$, are extremely important for the pathogenesis of SCA. Several mutations of these two genes are associated with the appearance of SCA.[137,138]

The third most common gene associated with SCA is RyR2.[129] The cardiac ryanodine receptor (RyR2) is a calcium channel involved in the calcium release, facilitating the binding of this ion to contractile proteins of the heart's muscle and activating systolic contraction.[129] A misregulation of RyR2 gene function could lead to a bad release of calcium ions, leading to an intermittent contractile and electrical activity, that could lead to a cardiac arrhythmia and the consequent SCA.[139] Several known RYR2 mutations, such as rs3766871 or rs376687, have been related to an increase of the risk of SCA.[129]

There are several more mutational genes implicated in the appearance of SCA, such as the nitric oxide synthase 1 adaptor protein (NOS1AP), which is implicated in the

Strategy	Examples	Measures	Power
Conventional risk factors	Framingham risk index	Predict evolution of disease	High for population; low for individual
Anatomic disease screening	Computed tomography	Identify abnormal coronary arteries	High for anatomic identification; low for individual event prediction
Clinical risk profiling	EF, stress testing, imaging techniques	Extent of disease	High for small, high-risk subgroups; low for large, low-risk subgroups
Transient risk predictors	Inflammatory markers	Prediction of unstable plaques	Uncertain feasibility
Personalized risk predictors	Familial/genetic profiles	Individual SCD expression	Uncertain clinical applicability; in evolution

Figure 19: Different SCA prevention strategies based on the detection and evaluation of different risk factors. Figure extracted from (Myerburg and Goldberger, 2017).[126]

bioavailability of some QT-prolonging drugs.[140] Although they have a much lower frequency in the population, their implications on the SCA appearance could be extremely important to analyse and detect susceptibilities.[129]

The detection of novel potential genetic markers and the development of new techniques that can improve the prognosis of SCA are two of the principal scientific goals regarding SCA.

Drug-induced sudden cardiac arrest

A large number of studies have been done in order to analyse the causal association between drug consumption and sudden cardiac arrest. There is an increasing number of drugs prescribed for a wide range of indications (including non-cardiac ones) such as arrhythmia, infection, depression, allergy or sedation that are directly related to a prolongation of the QT interval and a drug-induced arrhythmia that can lead to a fatal sudden cardiac death.[141]

As seen in the last chapter, one of the main mechanisms of action involved in SCA is the inhibition of potassium channels (both hERG and KCNQ1) that in turn block the repolarization and produces a QT prolongation. This is the most common mechanism of action of the majority of drugs involved in both

Torsades de Pointes and sudden cardiac death.[141] However, alternative mechanisms of drug-induced arrhythmogenesis and sudden death are being demonstrated. For example the inhibition of $Na_V1.5$ by several blocker agents has also been related to be the causant of QT prolongation.[142]

Accordingly, there is a rather long list of drugs, both cardiac (e.g., antiarrhythmics) and non-cardiac (e.g., antipsychotics, antidepressants or anesthetics) that have been reported to be associated with SCA.[143–148]

Sicouri and Antzelevitch published in 2008 a list of antipsychotic and antidepressant drugs that are known to increase the risk of SCD. They included typical antipsychotics (chlorpromazine, pimozide, thioridazine, perphenazine, trifluoperazine, haloperidol, droperidol), atypical antipsychotics (clozapine, quetiapine, risperidone, sultopride, loxapine and ziprasidone), tricyclic antidepressants (amitriptyline, amoxapine, clomipramine, citalopram, desipramine, doxepin, imipramine, nortriptyline, trimipramine) as well as other antidepressants (fluoxetine, maprotiline, lithium, sertraline and venlafaxine). Almost all of these drugs prolong the QT interval, inducing TdP arrhythmias and the consequent SCD.[143]

The relation between the anesthetic use and the appearance of cardiac arrests has been well studied and documented.[149–155] Some peri- and post-operatory studies revealed a

significant increase of cardiac arrests when local or total anesthesia, such as bupivacaine, halothane or isoflurane, is administered.[153–155]

A remarkable warning note needs to be added regarding antiarrhythmic drugs and the implications on SCA. Antiarrhythmic agents, such as amiodarone, have been for many years the principal treatment for cardiac arrest resuscitation. However, nowadays there is a great uncertainty about the benefits of their use in cardiac arrest resuscitation.[156] In fact, some antiarrhythmic drugs have been shown to be responsible for an increased risk of SCA. It is the case of encainide and flecainide, two antiarrhythmic agents of Ic class that could produce proarrhythmic response and the consequent cardiac arrest.[157,158]

Nowadays, worldwide regulatory agencies include several preclinical safety assays to detect whether new drug candidates show arrhythmogenic signals, including potassium channels inhibition screening in cell lines, isolated cardiac tissue recordings and QT-interval monitoring.[141] The detection of SCA-related drugs at a preclinical level is becoming increasingly helpful in order to prevent drug-induced sudden cardiac arrest or to anticipate the associated risk of some marketed new drugs, avoiding their prescription to the most susceptible and vulnerable individuals.

OBJECTIVES

The main objectives pursued by this Thesis can be summarized as follows:

i. To analyse the relationship (if any) between the affinities of endogenous metabolites for their native proteins and drugs for their primary targets.

ii. To reassess drug polypharmacology in the context of the human endogenous metabolome.

iii. To redefine drug safety margins relative to *in vitro* affinities of endogenous metabolites.

iv. To compare *in vitro* affinities of natural compounds and drugs for primary targets.

v. To explore and analyse the association between drugs and sudden cardiac arrest using databases of spontaneous drug safety reports.

The first and primary objective is addressed in Chapter 1, where an analysis of a large number of curated drug-target affinities reveals that the optimised affinity of endogenous ligands for their native targets can serve as a reference affinity baseline for the primary pharmacology of drugs. Results in

Chapter 1 provide answers also for the second and third objectives. Subsequently, results from Chapter 2 provided additional evidence for estimating the safety risk associated with drugs based on the endogenous metabolome. In Chapter 3, all the knowledge acquired in the previous two Chapters is used to gain a deeper understanding on the marked differences between drugs and natural compounds. Finally, Chapter 4 focuses on the particular use case, namely, the study of sudden cardiac arrest using postmarketing spontaneous drug safety reports.

RESULTS

1. The human endogenous metabolome as a pharmacology baseline for drug discovery

Drug Discovery Today 24(9):1359-6446 (2019)

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.

[#] These authors contributed equally to this work.
* Corresponding autors: jmestres@imim.cat

Bofill A, Jalencas X, Oprea TI, Mestres J. The human endogenous metabolome as a pharmacology baseline for drug discovery. Drug Discov Today. 2019 Sep;24(9):1806-1820. doi: 10.1016/j.drudis.2019.06.007.

Abstract

We have limited understanding of the wide variation of *in vitro* affinities of drugs for their targets. An analysis of a highly curated set of 442 interactions between 293 drugs and 79 primary targets reveals that 67% 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 optimized 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. A new approach for estimating safety margins is proposed.

Teaser

Two thirds of drugs have more potent *in vitro* affinities for their primary targets than those of the corresponding endogenous ligands that can also serve as a baseline to estimate off-target safety margins.

Keywords: endogenous metabolites, drug design, drug efficacy, drug safety, 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 and pharmacokinetic 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 due to 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].

This may in part explain the substantial variation of binding affinities observed for drugs targeting the same protein but also across different proteins [7]. For example, the serotonin receptor subtype 1A (5HT1A) is one of the protein targets recognised to be involved in the mechanisms of action of brexpiprazole and vortioxetine [7]. However, while 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 may require achieving certain levels of affinity for their primary target(s) and why these levels of affinity may be considerably different due to efficacy and receptor reserve across primary target(s) is at the core of modern preclinical drug discovery.

In humans, most drug targets are proteins whose function 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 thousands of chemical species [12]. Each endogenous ligand binds to its native protein with a certain affinity that has been sensitively optimised by evolution and that may 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 (*vide supra*). 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 are consistent with the majority of drugs having *in vitro* affinities for their primary target(s) above those of the corresponding metabolite/substrate-target interaction. The implications for secondary pharmacology and off-target safety margins derived from it are discussed [15]. In particular, the case of drug-induced valvular heart disease when the small molecule drug is an agonist of the serotonin 5HT2B receptor is analysed in detail.

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}). 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 (Supplementary Table 1). The vast majority of triads (90%) involve K_i binding affinities.

The 442 pairs of drug and metabolite affinities for the same protein are plotted in Figure 1. On average, drug affinities are 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) are higher than the corresponding metabolite-target interactions (solid gray line in Figure 1). This percentage

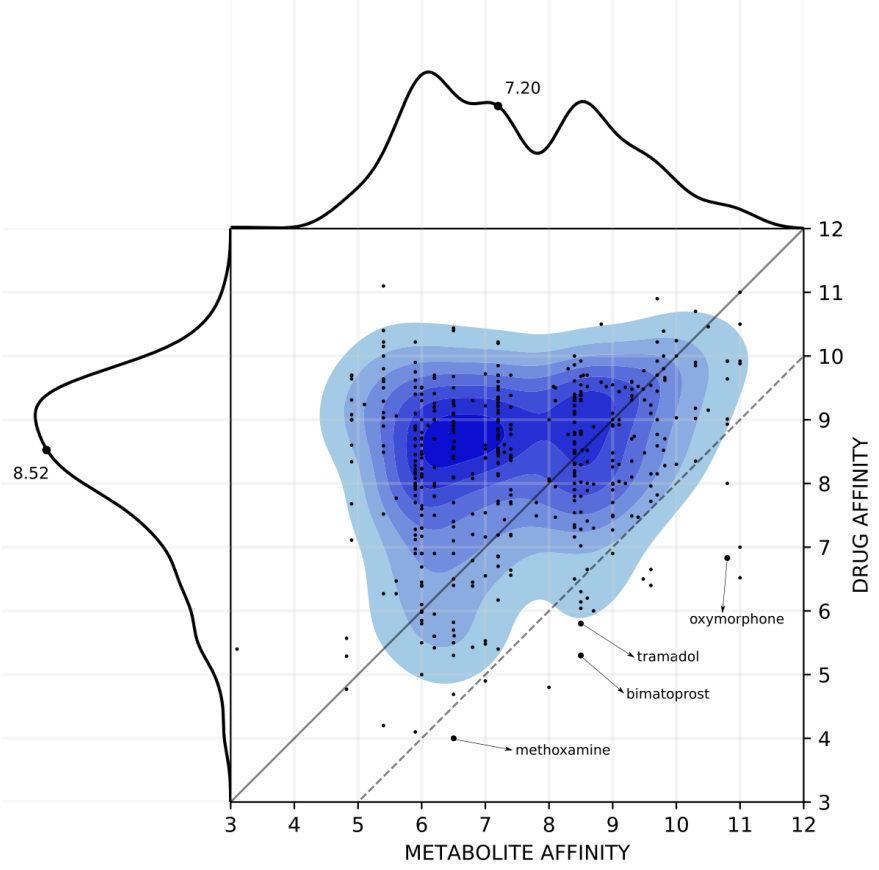


Figure 1 | Density plot of the 442 drug-target-metabolite triads. Drug affinities for their primary target(s) are plotted against the corresponding endogenous metabolite affinities. A kernel density estimation is shaded in blue tones, darkest blue corresponding to highest density regions. Solid and dashed lines correspond to drug affinities being, respectively, equal to and two orders of magnitude lower than the corresponding metabolite affinities for the same protein. 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 text are also indicated.

increases to 85% and 96% when considering drug affinities one order and two orders of magnitude (dashed gray line in Figure 1) below metabolite affinities, respectively.

Only 4% of drug affinities for primary target(s) are found to be over 100-fold lower than the corresponding metabolite affinities (Supplementary Table 2). A close examination of these 19 drug-target interactions suggests that the assignment of the interacting protein as primary target of the drug could benefit from a careful revision. For example, oxymorphone is a semi-synthetic opioid analgesic that has all three μ, δ, and κ opioid receptors assigned as mechanism of action targets with binding affinities (pK_i) of 9.44, 7.30, and 6.83, respectively [7]. However, oxymorphone's affinity 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 may suggest that the analgesic effect of oxymorphone is unlikely to be conducted through the activation of the κ opioid receptor but mainly due to 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 due solely to 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 could be 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 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 may not be 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 likely 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 play 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 Figure 1 into two sets of 160 and 260 interactions that involve drugs acting as agonists and antagonists on their targets, respectively (Figure 2). The 160 drug agonist interactions (Figure 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 (Figure 2b). The remaining 22 interactions correspond to drugs being annotated as partial agonists (14) or inverse agonists (8). 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 (Figure 2a) follows the diagonal line. Compared with Figure 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 gray line in Figure 2a), although this percentage

increases up to 91% if a drug-metabolite affinity window of minus two orders of magnitude is considered (dashed gray line in Figure 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 range of affinity values. In contrast, the 260 interactions from drug antagonists appear very much concentrated well above the diagonal line, with a clear shift towards targets with low affinity endogenous ligands when compared to drug agonists. In this case, over three quarters (79%) of the 260 drug antagonist affinities for their mechanism of action target(s) have already higher affinity values than the corresponding metabolite-target interactions (solid gray line in Figure 2b), with 98% of them fitting within a range of two orders of magnitude below the endogenous ligand affinities (dashed gray line in Figure 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. In contrast, antagonist drugs do need higher affinities

to completely block the action of endogenous ligands when released.

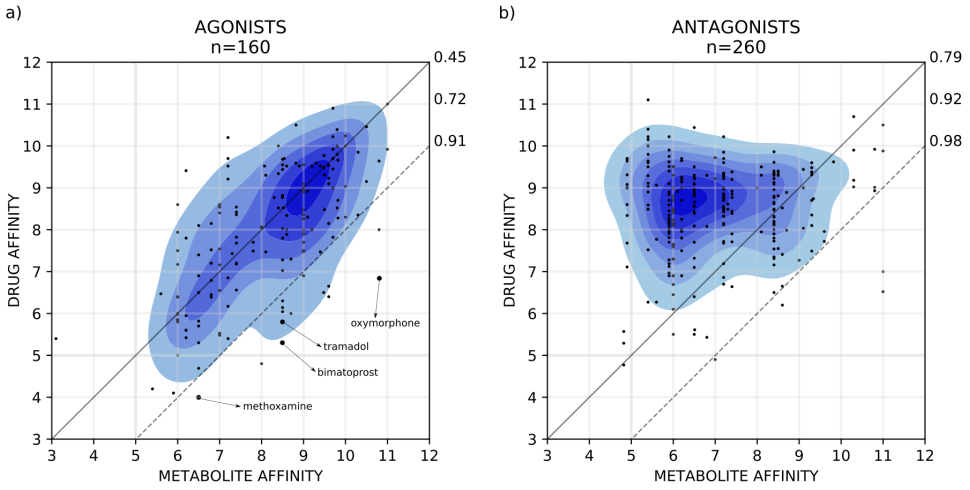


Figure 2 | Density plots of the 160/260 drug agonist/antagonist-target-metabolite triads**.** Affinities values of drug agonists (a) and antagonists (b) for their primary target(s) are plotted against the corresponding endogenous metabolite affinities. A kernel density estimation is shaded in blue tones, 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.

Comparing primary pharmacology across drug targets

A target-centered 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 and 9 nuclear hormone receptors. The height of each column in the circular plot of Figure 3 reflects the affinity between a drug target and its main endogenous ligand. As 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 serotonin, opioid, somatostatin, vasopressin and melatonin receptor families show *in vitro* binding affinities in the nanomolar range. In contrast, micromolar affinities seem to be sufficient for the endogenous ligands of acetylcholine (muscarinic), dopamine and adenosine receptor families. It seems thus clear that different proteins have evolved to interact with their native metabolites at different levels of affinity which may in turn translate into different lower-bound affinity criteria for any potential drug targeting them.

As illustrated in Figure 3, the endogenous ligand affinity for each protein sets a baseline above which drug affinities may spread. In principle, the lower the affinity baseline, the wider the affinity window allowed for drugs to optimise other pharmacodynamic and pharmacokinetic parameters. This is well reflected by the drug affinity distributions for targets having different endogenous ligand affinities. For example, serotonin and 17ß-estradiol have both *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 6 drugs that have 5HT1A as primary target and their pK_i values vary between 7.8 (right within the 100-fold difference from the serotonin affinity for 5HT1A) and 9.9. There are also 6 drugs having ESR1 assigned as mechanism of action target and, most interestingly, their affinities cover a similar range of pK_i values between 7.5 (close to the 100-fold difference from the 17□-estradiol affinity for ESR1) and 9.7. In 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 primary target covering a wide affinity window of pK_i values between 5.4 (within the 100-fold difference from the dopamine affinity for DRD2) and 10.2. Finally, (-)-noradrenaline has micromolar affinity (pK_i = 6.0) for the ß1 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 10-fold below the (-)-noradrenaline affinity for ADRB1) to 9.5.

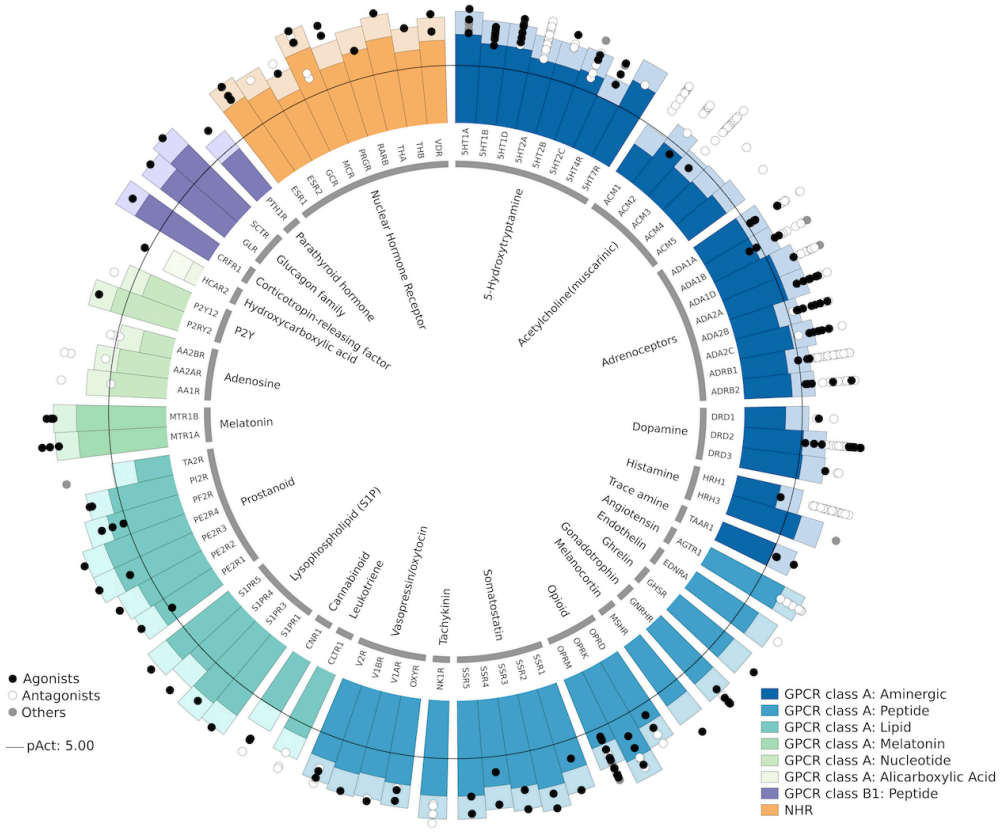


Figure 3 | Target-centered analysis of the 442 drug-receptor-metabolite triads. Each one of the 79 drug targets, 70 G protein-coupled receptors and 9 nuclear hormone receptors, is represented by a column in this circular plot. The height of the column reflects the affinity between a drug target and its main endogenous ligand. The tip of each column, in lighter color, 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.

Figure 4 illustrates these trends for the set of 26 targets for which pK_i data is available for more than 3 drugs (Supplementary Table 1). It is shown that there is a directly proportional relationship (Pearson correlation coefficient = 0.71; p-value = 4.52E-5; r^2 = 0.51) between the affinity of an endogenous ligand and the minimum affinity of a drug for its

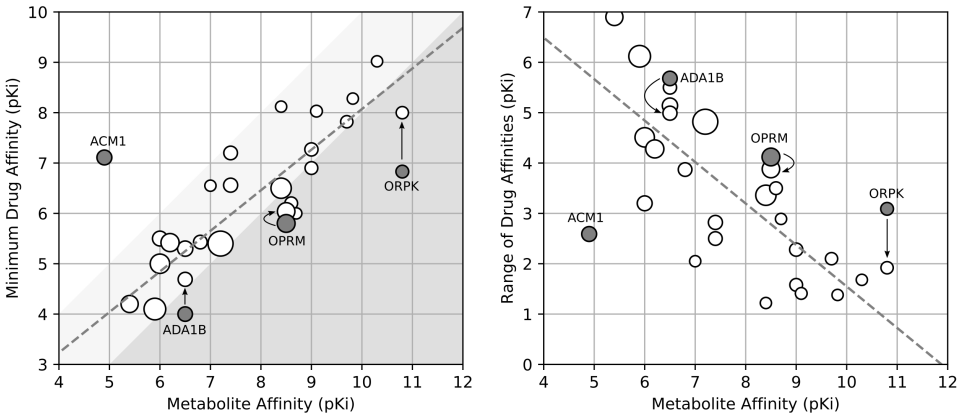


Figure 4 | Metabolite affinity as reference baseline for the primary pharmacology of drugs. Each circle represents one of the 26 targets for which pK_i data is available for more than 3 drugs. The size of the circle is proportional to the number of drugs associated with each target. Trends show that endogenous metabolite affinity is a) directly proportional to the minimum affinity of all drugs having the protein as primary target and b) inversely proportional to the range of drug affinities. Marked outliers (gray circles) are the muscarinic acetylcholine receptor M1 (ACM1), the κ opioid receptor (OPRK), the μ opioid receptor (OPRM) and the alpha-1B adrenergic receptor (ADRA1B). For OPRK, OPRM and ADRA1B their position 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 existent direct and inverse linear correlations, respectively.

primary target (Figure 4a). This suggests that the *in vitro*

binding affinity of the endogenous ligand for its native protein can be a good reference baseline for the objective primary pharmacology of drugs in preclinical research. It is shown also that there is an inversely proportional relationship (Pearson correlation coefficient = -0.64; p-value = 3.8E-4; r^2 = 0.41) between endogenous ligand affinities and the range of drug affinities for a given target (Figure 4b). This reflects the fact that, for every target, the upper-bound drug affinities reach always subnanomolar potencies, whereas the 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 ten micromolar 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 may 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

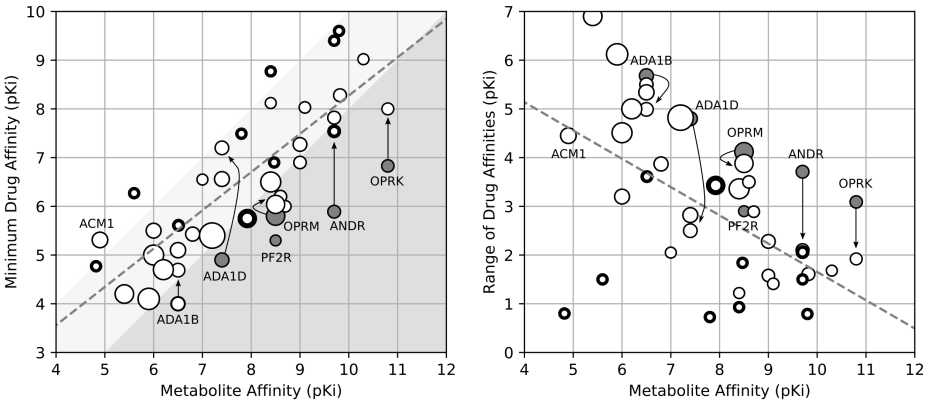


Figure 5 | Endogenous metabolite affinity as reference baseline for the primary pharmacology of drugs. Each circle represents one of the 37 targets for which pK_i data is available for more than 3 drugs. The size of the circle is proportional to the number of drugs associated with each target. Newly incorporated targets respect to Figure 4 appear as circles with tick black borders. The same trends observed in Figure 4 hold true: the endogenous ligand affinity for a protein is a) directly proportional to the minimum affinity of the drug for that protein and b) inversely proportional to the range of drug affinities for that protein. Targets with outlier drugs discussed in the text and removed from the correlations are shown in gray and their corrected position indicated by an arrow.

represented in the data we analyzed, drugs having ACM1 as primary target and with pK_i affinities around 4 might exist. In contrast, the case of OPRK is completely different. If one disregards oxymorphone (discussed above), the drug with minimum affinity is nalbuphine (pK_i = 8.0) that places OPRK right where one would expect. Similarly, discarding the discussed cases of tramadol and methoxamine for OPRM and ADA1B, respectively, pushes also 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 oxymorphone, tramadol and methoxamine results in stronger correlations between endogenous ligand affinities and minimum drug affinities (Pearson correlation coefficient = 0.87; p-value = 1.59E-8; r^2 = 0.76), on one hand, and range of drug affinities (Pearson correlation coefficient = -0.79; p-value =3.28E-6; r^2 = 0.62), on the other hand.

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. In order 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 (Supplementary Table 3) 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 148 drugs, 57 targets and 35 metabolites involved in those 202 new triads, 110 drugs, 9 targets and 9 metabolites were not represented in the original set.

The density plot of this extended set of 202 drug-receptor-metabolite triads (Supplementary Figure 1) is similar to the distribution displayed in Figure 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 3 drugs was available increased from 26 to 37. Accordingly, Figure 5 represents an expanded version of Figure 4, in which the 11 new targets are marked as white circles with thick black borders. As can be observed in Figure 5a, besides the special cases of ADA1B, OPRM and OPRK (drawn as gray circles) already discussed in Figure 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 above), 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], that has a significantly higher binding affinity for ANDR (pK_i=7.65) [28]. The removal of flutamide brings ANDR much closer to the region of accepted minimum drug affinity for this target. Another target marked as outlier in Figure 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 emphasize the importance of data completeness in this type of analyses [27].

After repositioning of those target outliers in Figure 5a, endogenous ligand affinities for a given protein were found to be directly proportional to the currently known minimum affinities of drugs for that protein, with a similar correlation to that observed already in Figure 4 (Pearson correlation coefficient = 0.83; p-value = 5.49E-10; r^2 = 0.68). On the other side, the inverse correlation between endogenous ligand affinities and the range of drug affinities for a given target (Figure 5b) is still found significant (Pearson correlation coefficient = -0.57; p-value = 3.01E-4; r^2 = 0.32) but some of the new targets with just 3 representative drugs show drug

affinity ranges clearly below the expected values for their respective targets. As already mentioned above, it is worth noting that this type of analyses is especially sensitive to data completeness and the range of drug affinities are susceptible

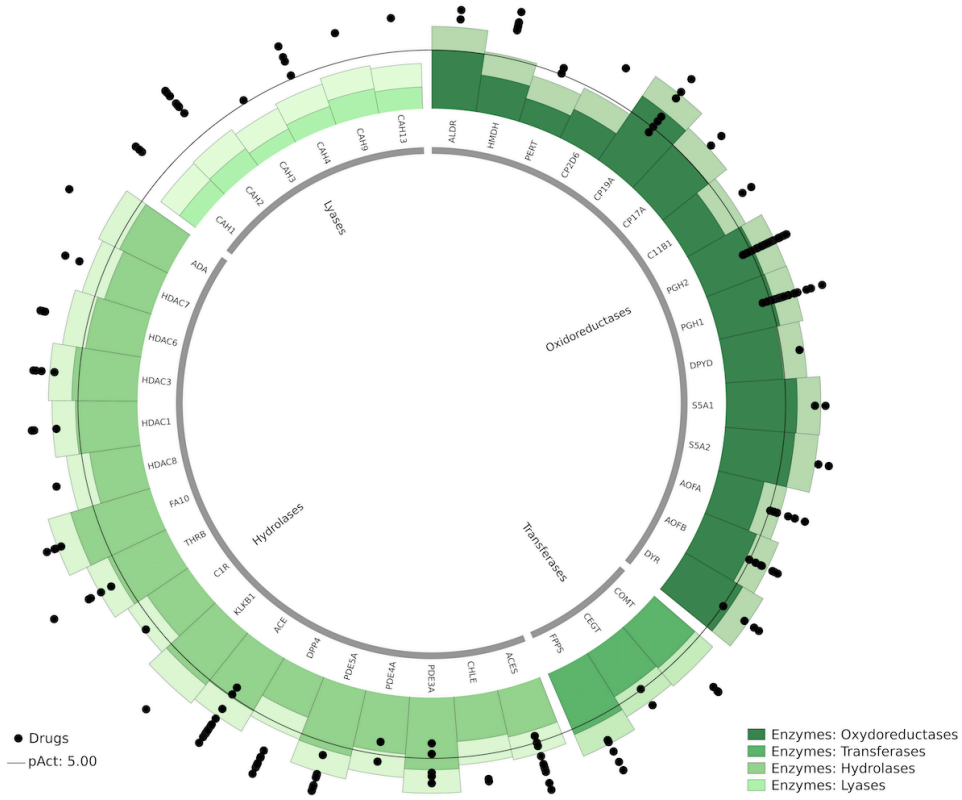


Figure 6 | Target-centered analysis of the 222 drug-enzyme-substrate triads. Each of the 41 enzymes is represented by a column in this circular plot. In contrast to the receptor families, the height of the column reflects the two order of magnitude window, in lighter color, above the affinity of the 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.

to variations as new drugs are included. As can be observed, the majority of the targets located well below the correlation line are actually targets with a small number of drugs (3), 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

In order 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, $K_M = (k_{-1} + k_2)/k_1$, being $(k_{-1} + k_2)$ the rate of breakdown and k_1 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 to 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 much higher than K_M values, on the negative log(molar) scale, in order 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 non-canonical 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 (Supplementary Table 4). The density plot of drug and endogenous substrate affinities for the same enzyme shows a clear accumulation of interactions above the diagonal solid line (Supplementary Figure 2). From a target-centered perspective, analogous to Figure 3 for receptors, Figure 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 Figure 6 reflects the affinity between the enzyme and its main endogenous substrate. Compared to 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 to drug-receptor affinities relative to their endogenous ligands.

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 G protein-coupled receptors [32], since it has been estimated that, on average, they may have biologically relevant binding affinities for up to ten members of this protein family [33]. Some of this secondary pharmacology may indeed be necessary for the efficacy of drugs addressing complex diseases [34] but binding to certain off-targets may 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 than 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 9 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 this work 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 in this work, we extended from 9 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 since no bibliographic evidence of links to VHD was found at present. All binding (pK_i) and VHD risk data for these 24 5HT2B agonist drugs are

provided as Supplementary Table 5. The distribution of *in vitro* binding affinities of the 24 5HT2B agonists is provided in Figure 7, which is analogous to Figure 1 but centered 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, 5 (42%) have pK_i values equal to or larger than the serotonin affinity, 7 (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, would have a pK_i value clearly below this two-order of magnitude window, although its racemic mixture, fenfluramine, would be right at the edge of it (pK_i = 6.4). In fact, it was the case of fenfluramine that alerted almost 20 years ago that its more potent 5HT2B agonist metabolite, norfenfluramine, could be responsible for its associated risk to 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. In contrast, of the 12 nonvalvulopathic drugs, 7 (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 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].

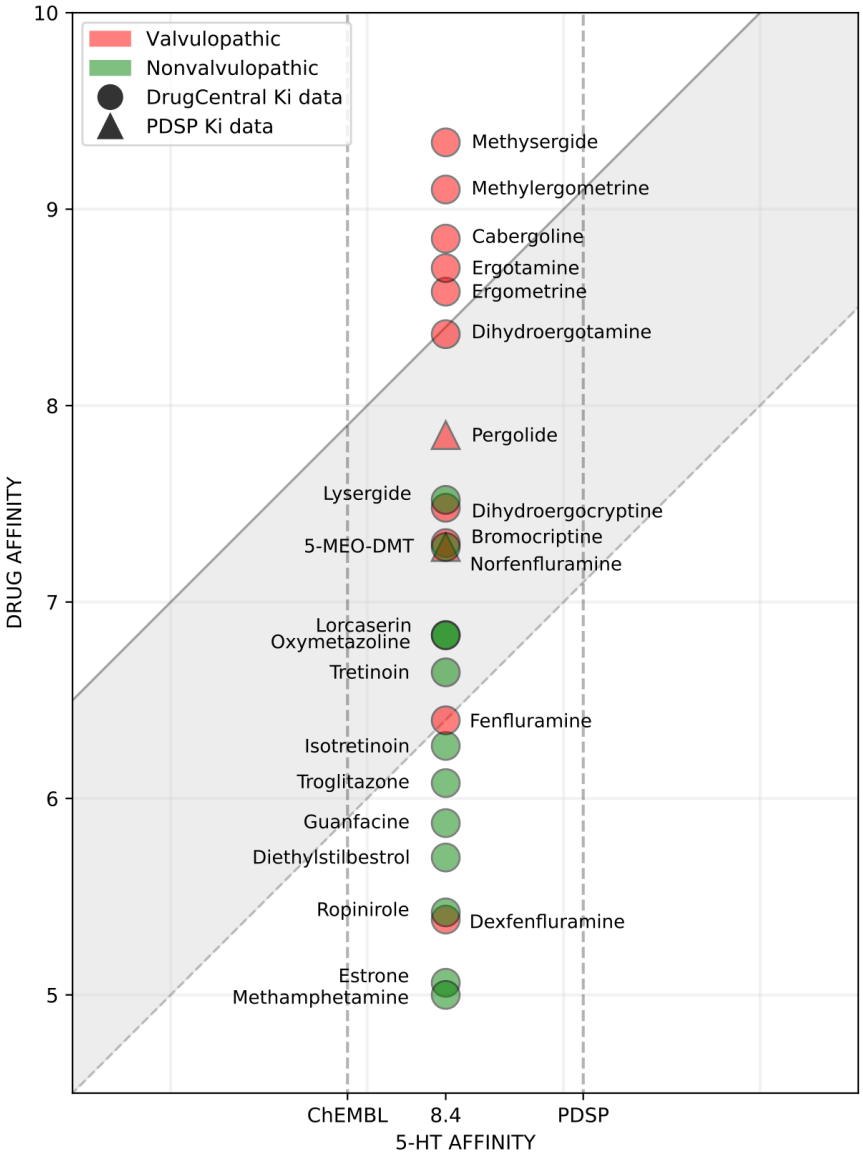


Figure 7 | Distribution of binding affinities (pK_i) for 24 5-HT_{2B} agonists. The set includes 12 valvulopathic drugs (in red) and 12 nonvalvulopathic drugs (in green). Drug affinities are aligned at the maximum binding affinity of serotonin for 5HT2B found in GtoPdb (8.4). For the sake of comparison, 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).

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 criteria 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 call of a drug to produce VHD. As an example, we took as reference value in Figure 7 the maximum affinity of serotonin for 5HT2B reported in GtoPdb (pK_i = 8.4). But the corresponding binding affinities found in other public sources may 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 leftwards, which would cause the nonvalvulopathic drugs isotretinoin and troglitazone to enter the gray zone. However, the serotonin affinity in PDSP (pK_i = 9.1) would move the entire drug affinity distribution rightwards, which would then cause that 10 out of the 12 nonvalvulopathic drugs would be safely placed below the two-order of magnitude window. Data robustness and dispersion is thus an important aspect to consider in this type of analyses.

Implications for initial assessments of safety margins

Many drug safety events are caused by the interaction of drugs with their own primary targets (pT) or with secondary (off-)targets (oT). 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 recognized [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 at the early stages of a drug discovery project and, in addition, it was recently shown that, at least for the case of 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 above.

Strategies to widen the off-target safety margins involve increasing the potency of the drug for the primary target and decreasing the drug's potency 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]. But the outcomes presented above may change the perspective from which off-target safety margins have been traditionally regarded. The concept is illustrated in Figure 8a for two hypothetical drugs (labelled as 1 and 2) representing two opposite case scenarios: on one hand, the affinity of drug 1 for its primary target (circle labelled

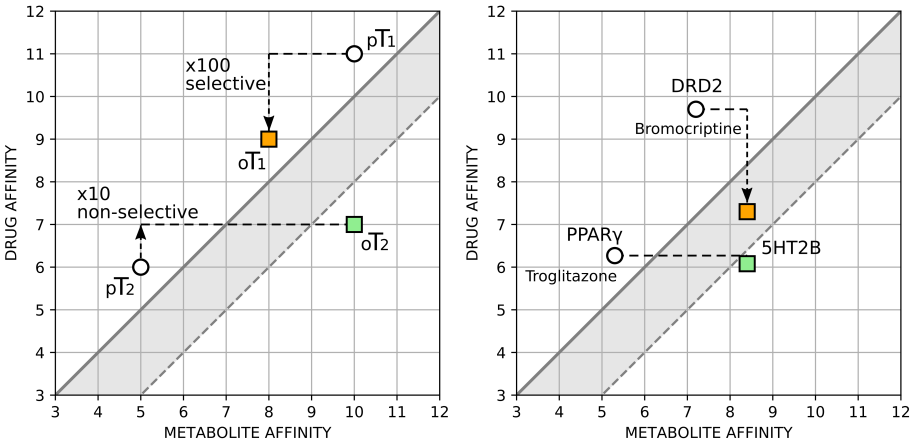


Figure 8 | A new perspective on safety margins considering endogenous ligand affinities. a) Schematic diagram plotting endogenous metabolite affinities *versus* drug affinities for a set of four exemplary proteins: two primary targets (pT_1 and pT_2, represented by circles) and two off-targets (oT_1 and oT_2, represented by squares). Despite its 100-fold selectivity, drug 1 may be at risk of showing the safety issue linked to off-target oT_1 (orange square), whereas drug 2 may likely be safe, even though its affinity for off-target oT_2 (green square) is 10-fold higher than the affinitity 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 (DRD2 and 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 5HT2B). See text for discussion.

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 likely to be biologically relevant; on the other hand, the affinity of drug 2 for its primary target (circle labelled as pT_2) is 10-fold 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 Figure 7). The result is that, in spite of its 100-fold selectivity, drug 1 may be at risk of producing the safety issue associated with affinity to oT_1, whereas the lack of selectivity may 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 above (Figure 7), bromocriptine and troglitazone are case studies resembling the hypothetical examples of drugs 1 and 2 in Figure 8a. The corresponding safety diagram for these two drugs is presented in Figure 8b. The binding affinity (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 just about 10-fold lower relative to the corresponding affinity for the endogenous ligand, placing this drug within the safety risk grey

zone. Bromocriptine exemplifies of the hypothetical drug 1 case: despite 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(serotonin) – pK_i(bromocriptine) = 1.10, foreshadow its potential risk for valvular heart disease. In contrast, the affinity of troglitazone for its primary target (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 to 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(serotonin) – pK_i(troglitazone) = 2.32, is a better predictor for its low risk in causing valvulopathy.

Overall, these results would favor 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):

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

Values of *SM* > *2* would be recommended.

Some practical limitations

There are however some aspects of the present study that merit further consideration. As already highlighted above, there are limitations associated with data completeness and bias, always present in this type of analyses [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). Also, even though there are 43 endogenous ligands assigned to those 442 interactions, a single one of them (serotonin) is involved in almost 15% of them. This very much 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 G protein-coupled receptors. A similar situation is encountered in the enzyme dataset 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, in order 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 characterize 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, as 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 critical as 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/substrates and drugs for the same protein was scarce, structural data of the complex between the protein with metabolites/substrates and drugs are rarer still. This notwithstanding, we searched the Protein Data Bank (PDB) [55] for entries of proteins that were co-crystallized with both

their endogenous metabolite/substrate and one of the drugs from our list. Two illustrative examples for G protein-coupled receptors were found, namely, the beta-2 adrenergic receptor (ADRB2) and the AA2AR. ADRB2 has been co-crystallized with its endogenous ligand, adrenaline (4ldo), and a drug antagonist, timolol (3d4s), whereas structures of AA2AR co-crystallized 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 Figure 9a. For the enzyme dataset, protein structures co-crystallized with both the endogenous substrate and a drug were found for 10 out of the 222 drug-enzyme-substrate triads collected. In total, 43 PDB entries were identified, 38 of them 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 Figure 9b. For both proteins, the alignment confirms that the endogenous metabolite/substrate and the drugs bind to the same site. Accordingly, along the same lines expressed above for pharmacological data, more efforts to resolve crystal structures of protein-metabolite/substrate/drug complexes for which *in vitro* affinity data is available would be an informative addition to a reference human endogenous ligand repository.

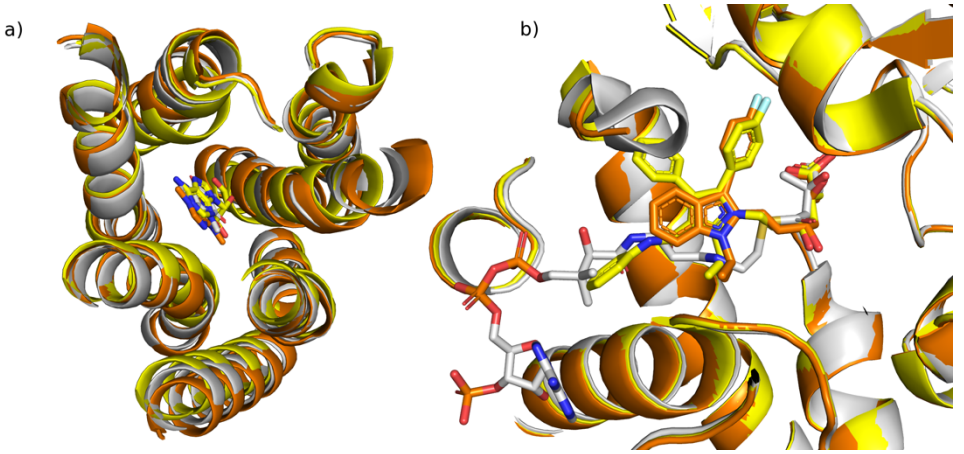


Figure 9 | Pairs of metabolite/substrate and drugs binding at the same protein site. Backbone superpositions of a) the adenosine 2A receptor co-crystallized with its endogenous ligand, adenosine (2ydo, in white), and two drug antagonists, theophylline (5mzj, in yellow) and caffeine (3rfm, in orange), and b) the 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase enzyme co-crystallized with its endogenous substrate, HMG-CoA (1dqn, in white), and two drug inhibitors, atorvastatin (1hwk, in yellow) and fluvastatin (1hwi, in orange).

Finally, we ought to emphasize 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 of the 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 (such as half-life, distribution and clearance) that are affected by interindividual variations (such as body mass, metabolism, drug-drug interactions, co-morbidities 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 co-existing 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 to the *in-vitro* to *in-vivo* translatability.

Concluding remarks

Understanding why a specific drug needs a certain level of affinity for their 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 148 drugs, 57 receptors, and 35 endogenous ligands, which included 110 drugs, 9 receptors, and 9 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 to 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.

Our findings that the human endogenous metabolome could serve as a pharmacology baseline for drug *in vitro* affinities on proteins was thus strengthened by the observations that endogenous ligand affinities for their native proteins are, on one hand, directly proportional to the minimum drug affinity among all drugs optimized 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 drug's primary pharmacology, results obtained in this work highlight also the impact that *in vitro* endogenous ligand affinities can have on assessing the risk of safety events linked to the drug's secondary pharmacology. 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, very much 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, as 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 drug's primary and secondary pharmacologies 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 in this work based on preclinical *in vitro* pharmacology data could provide additional metrics to assess the risk of clinical failure and contribute to reduce drug attrition.

Data and Methods

In vitro pharmacology of drugs and primary targets

We explored public pharmacology databases in search for interactions of drugs with their primary targets where quantitative affinity data was available for both the interactions between the drug and the target and between the target and its main endogenous metabolite. Accordingly, drug-target interactions labeled as being involved in the mechanism of

action of the drug with defined activity values were retrieved from DrugCentral [7]. Of the 4,486 drugs available in DrugCentral (downloaded on May 2017), 1,862 had defined affinity values against a given target, 936 of them having at least one interaction against a human target labeled as being involved in the drug's mechanism of action (primary targets of the drug). A total of 1,769 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 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 first the Guide to Pharmacology database (GtoPdb) [14] 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. There are however 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 favorable scenario when a range of affinities is provided, we aim at increasing the robustness of the 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}), taking the highest metabolite affinity available. We identified a total of 442 drug-receptor-metabolite triads involving 293 drugs, 79 proteins and 43 metabolites (Supplementary Table 1). Among them, 404 (90.4%) are pK_i values. In a second stage, we collected an additional set of 202 drug-receptor-metabolite triads involving 148 drugs, 57 targets

and 35 metabolites (Supplementary Table 3). Of them, 116 (57.4%) are 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 them, 2 drugs were annotated as 5HT2B agonists in DrugCentral [7]. An additional list of 8 drugs could be found annotated as 5HT2B agonists in ChEMBL [28]. A manual literature search of the remaining drugs allowed to confirm 11 other drugs as 5HT2B agonists and provide evidence supporting or rejecting their associated risk to produce 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 are recognised valvulopathic drugs. An additional set of 3 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 Supplementary Table 5.

Conflicts of interest

Dr. Oprea 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 Inc. Dr. Mestres was a former full time employee at Pharmacia&Upjohn (1996) and Organon (1997-2003). He is currently CEO of Chemotargets S.L.

Acknowledgments

We are grateful to Rita Moita Santos and Anne Hersey for their kind assistance in the construction of the circular plots in Figures 3 and 6. K_i 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 data associated with this article can be found, in the online version, at https:doi.org/10.1016/j.drudis.2019.06.007

Biographies of all authors

Andreu Bofill is a PhD candidate at the Research Group on Systems Pharmacology under the supervision of Dr. Jordi Mestres. He studied Human Biology and later obtained a MSc in Bioinformatics for Health Science at University Pompeu Fabra. His principal research interests include drug interactions, drug-induced adverse event detection and personalized medicine.

Xavier Jalencas is a Research Scientist at Chemotargets. He has a PhD in Biomedicine from the University Pompeu Fabra for his research on structure-based approaches to polypharmacology performed at the Research Group on Systems Pharmacology under the supervision of Dr. Jordi Mestres. After a post-doctoral 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, MD PhD is Professor of Medicine and Pharmaceutical Sciences, and Chief, Translational Informatics Division, at the Department of Internal Medicine, University of New Mexico School of Medicine in Albuquerque, New Mexico (USA). He is also Guest Professor at the Institute of Medicine, Gothenburg University in 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 8 US patents. Since 2014, Dr. Oprea is PI for the Illuminating the Druggable Genome Knowledge Management Center, a NIH Common Fund initiative. Dr. Oprea's work 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 in 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. After a post-doctoral stay at Pharmacia&Upjohn in Kalamazoo (Michigan, USA), in 1997 he joined the Molecular Design &

Informatics department at N.V. Organon in Oss (The Netherlands) and in 2000 he was appointed Head of Computational Medicinal Chemistry at Organon Laboratories in Newhouse (Scotland, 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, 10 patents among them.

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, S337-345
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. Biotchnol.* 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, 10.1016/j.echo.2009.05.002
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. Editors (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

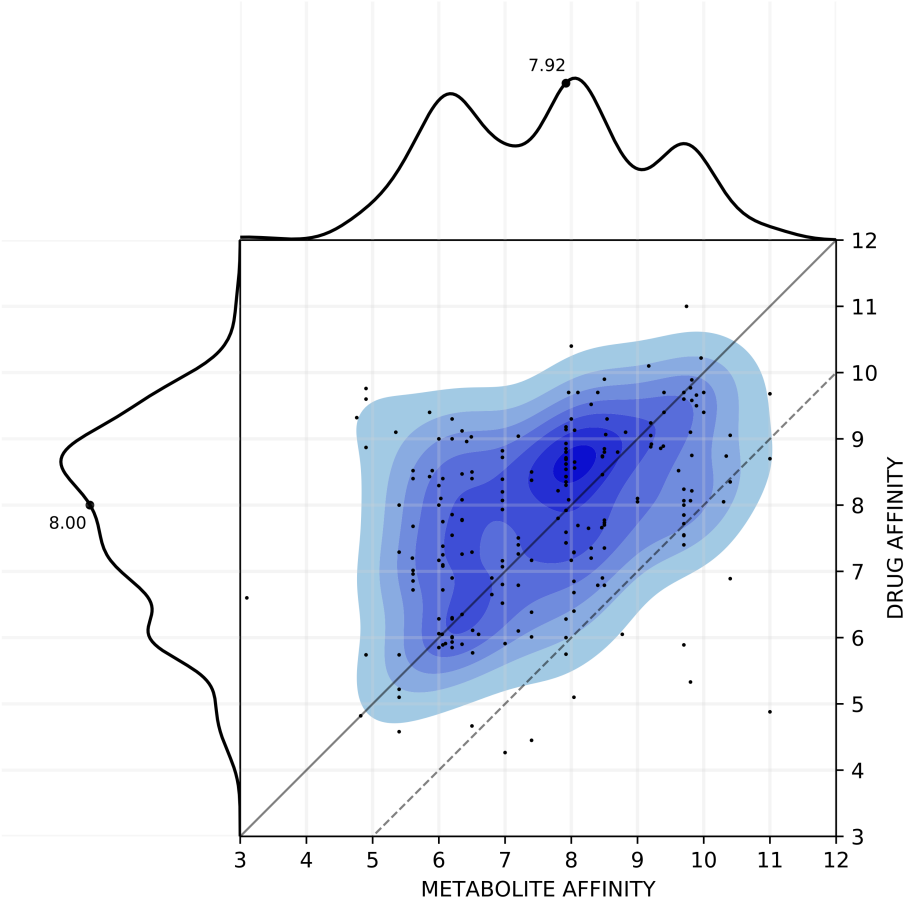
45. Huang, X.-P. *et al*. (2009) Parallel functional activity profiling reveals valvulopathogens are potent 5-hydroxytryptamine(2B) receptor agonists: implications for drug safety assessment. *Mol. Pharmacol*. 76, 710–722
46. Fitzgerald, L. W. *et al.* (2000) Possible role of valvular serotonin 5-HT(2B) 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? *The 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, Guidelines for ATC classification and DDD assignment. Oslo (2017)
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, (2007) J. 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.*, doi: 10.1038/d41573-019-00074-z

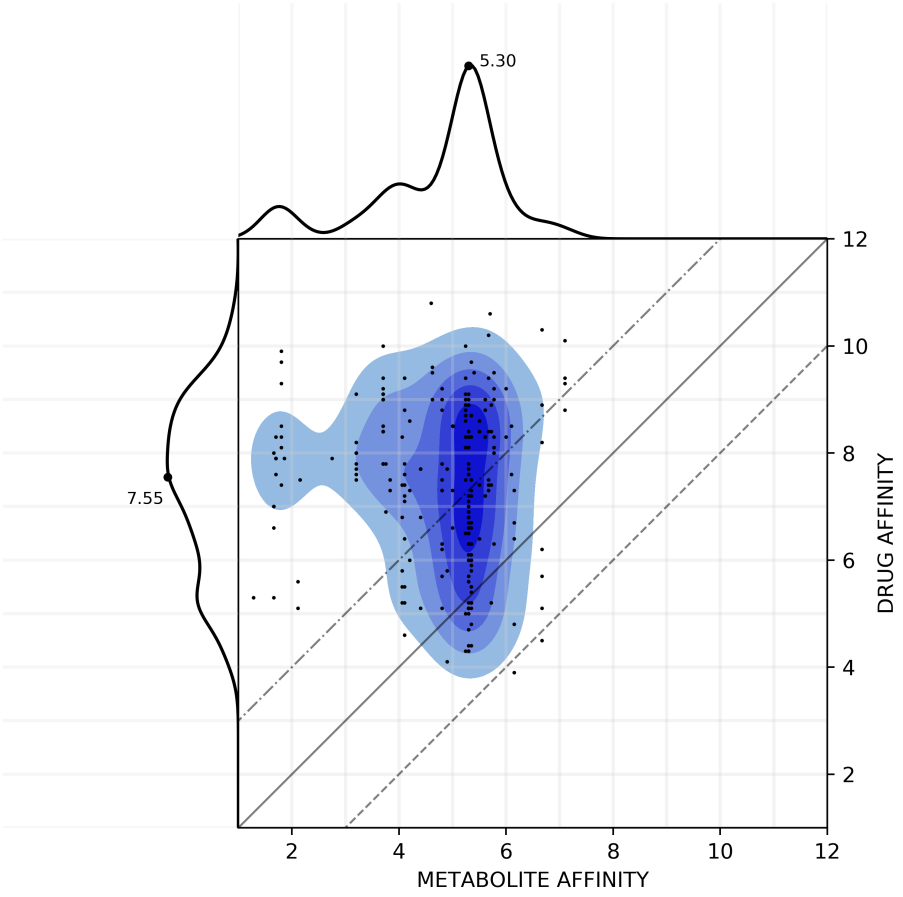
Supplementary Materials

Supplementary Figure 1:



Supplementary Figure 1. Density plot for the 202 drug-receptor-metabolite triads of the extended pharmacology dataset.

Supplementary Figure 2:



Supplementary Figure 2. Density plot for the 222 drug-enzyme-substrate triads of the extended biochemistry dataset.

Supplementary tables are enclosed in the next links except for the Supplementary Table 5 that is included in the next page:

Supplementary Table 1: Original list of 442 drug-receptor-metabolite triads, involving 293 drugs, 79 receptors, and 43 endogenous metabolites.
https://drive.google.com/file/d/1kCFlb1bmaGXR_-wBbMcpw_saTyPhK6x_

Supplementary Table 2: List of the 19 drug affinities for primary receptor targets found to be over 100 fold less potent than the corresponding endogenous metabolite affinities.
https://drive.google.com/file/d/1VRZMlOVtYqQUxRe6nnYw2KrHbS7Wte1S

Supplementary Table 3: Extended list of 202 drug-receptor-metabolite triads, involving 148 drugs, 57 receptors, and 35 endogenous metabolites.
https://drive.google.com/file/d/19H1V0fVFXyhAxgPpbXWhVC7HKqIQAwdk

Supplementary Table 4: Extended list of 222 drug-enzyme-substrate triads, involving 164 drugs, 41 enzymes, and 32 endogenous substrates.
https://drive.google.com/file/d/1kdGPx21OZ6gC3fwxSP-o9WDRtZe7zhQO

Supplementary Table 5:

Drug	Drug activity	Drug activity source	5-HT affinity for 5-HT2B	Activity type	VHD	VHD references
Methysergide	9,34	Drug Central	8,4	Ki	Yes	1, 2
Methylergometrine	9,10	Drug Central	8,4	Ki	Yes	3
Cabergoline	8,85	Drug Central	8,4	Ki	Yes	3, 4
Ergotamine	8,70	Drug Central	8,4	Ki	Yes	2, 3
Ergometrine	8,58	Drug Central	8,4	Ki	Yes	3
Dihydroergotamine	8,36	Drug Central	8,4	Ki	Yes	3, 5
Pergolide	7,85	PDSP	9,3	Ki	Yes	2, 3
Lysergide	7,52	Drug Central	8,4	Ki	No	
Dihydroergocryptine	7,48	Drug Central	8,4	Ki	Yes	
Bromocriptine	7,30	Drug Central	8,4	Ki	Yes	4, 7
Norfenfluramine	7,28	PDSP	9,3	Ki	Yes	1, 2, 6
5-MEO-DMT	7,28	Drug Central	8,4	Ki	No	
Lorcaserin	6,83	Drug Central	8,4	Ki	No	3
Oxymetazoline	6,83	Drug Central	8,4	Ki	No	2, 8, 9
Tretinoin	6,64	Drug Central	8,4	Ki	No	
Fenfluramine	6,40	Drug Central	8,4	Ki	Yes	1, 2, 6
Isotretinoin	6,27	Drug Central	8,4	Ki	No	
Troglitazone	6,08	Drug Central	8,4	Ki	No	
Guanfacine	5,87	Drug Central	8,4	Ki	No	2, 8, 9
Diethylstilbestrol	5,70	Drug Central	8,4	Ki	No	
Ropinirole	5,42	Drug Central	8,4	Ki	No	2, 3, 8
Dexfenfluramine	5,38	Drug Central	8,4	Ki	Yes	1, 6
Estrone	5,06	Drug Central	8,4	Ki	No	
Methamphetamine	5,00	Drug Central	8,4	Ki	No	

VHD references:

1. 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, 10.1016/j.echo.2009.05.002.
2. Hutcheson, J. D. et al. (2011) Serotonin receptors and heart valve disease—It was meant 2B. Pharmacol. Ther. 132, 146–157.
3. 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.
4. 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.
5. Roth, B. L. (2007) Drugs and valvular heart disease. N. Engl. J. Med. 356, 6–9.
6. I. verson, S. L. and Smith, D. A. Editors (2016) Metabolite safety in drug development. John Wiley & Sons.
7. Tan, L. C. S. et al. (2009) Bromocriptine use and the risk of valvular heart disease. Mov. Disord. 24, 344–349.
8. Elangbam, C.S. (2010) Drug-induced valvulopathy: an update. Toxicol Pathol. 38, 837–848.
9. 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.

Supplementary Table 5. List of 24 5-HT_{2B} agonist drugs included in the valvular heart disease analysis.

2. The impact of the human endogenous metabolome on the secondary pharmacology of drugs

Andreu Bofill[1], Xavier Jalencas[1] 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.

* Corresponding author: jmestres@imim.cat

Abstract

A better understanding of the pharmacological activity of drugs in a specific biological system would contribute to more effective and safer individualized drug treatments. The secondary pharmacology of drugs is directly related to their affinity to targets distinct from their intended therapeutic target. The present work explores the impact that the human endogenous metabolome has on drug polypharmacology. An analysis of 2845 interactions between 403 drugs and 169 human targets shows that more than 20% of secondary targets have affinities below two orders of magnitude of the affinity of the native endogenous metabolite for that target. These drug-target interactions can be considered as biologically non significant, as it is unlikely they produce any relevant effect in the organism. Accordingly, we propose a new approach to reevaluate drug bioactivity to better approximate the real drug polypharmacology landscape.

Introduction

Drug safety is one of the main pillars of the pharmaceutical industry. Producing safe drugs without reducing their efficacy is one of the major challenges of drug discovery. A general strategy used in order to achieve this goal is to improve the on-target selectivity maximizing the *in vitro* binding affinity for the therapeutic targets, expecting to minimize the secondary pharmacology of the drug at the same time.

However, it is almost impossible for a drug to only interact with its therapeutic targets with no secondary pharmacology at all. The polypharmacology of drugs, referred to their capacity to interact with multiple targets, leads to a much more complex scenario.[1] Estimating a more realistic drug polypharmacology map would be a significant step forward in associating drug effects to specific drug-target interactions. This knowledge would significantly improve drug treatments by anticipating drug adverse events or identifying new therapeutic targets for existing drugs.[2,3]

In the last decades, metabolomics has emerged as a promising discipline in the drug discovery process. It describes and quantifies all small molecules (endogenous or exogenous) involved in a biological system.[4–6] Genetic and environmental alterations are traduced to metabolite adjustments as an ultimate response of the organism to changes.[7] In

consequence, the metabolome reflects with precision the cellular and tissular behaviour, being the variations in metabolites level a good indicator of individual alterations.[8] In fact, the use of metabolites as disease biomarkers is already being applied in order to anticipate the appearance of certain chronic diseases such as cancer or diabetes, allowing a more individualized medicine.[9–11]

The evolutionary optimised affinity of endogenous ligands for their native proteins can serve as a baseline for the primary pharmacology of drugs. We demonstrated in a previous work that the safety risks of drugs arising from their secondary pharmacology depend on the affinities that endogenous ligands have for the proteins involved.[12] Accordingly, we suggested that the minimum level of affinity that is needed for a drug to be able to produce a significant effect by interacting with a specific protein is linked to the affinity of the endogenous ligand for this protein.[12]

The large-scale analysis presented in this work introduces a new approach for a more accurate estimation of drug polypharmacology based on the analysis of the endogenous human metabolome impact on drug-protein interactions. It also explores the potential use of this approach in precision medicine for the identification of therapeutic biomarkers.

Drugs pharmacology and endogenous ligands

A list of 403 drugs was extracted from a previous work.[12] The 403 drugs in this list include all drugs for which a quantitative interaction with a target involved in its mechanism of action was available in DrugCentral.[13] Additionally, all targets required to have an assigned principal endogenous metabolite with a known quantitative affinity value. This list was used to interrogate DrugCentral[13], Chembl[14] and Guide To Pharmacology (GtoPdb)[15] databases in order to extract all available drug-target interactions in humans. The targets implicated in the mechanism of action of these drugs were extracted from the DrugCentral repository[13]. A total of 3641 drug-protein interactions involving all 403 drugs and 406 proteins was finally retrieved. The list of human targets involved in those interactions was used to interrogate the Guide to Pharmacology database (GtoPdb)[15] to determine their native endogenous ligands and extract the maximum affinity value reported for their interaction. All drug-target-metabolite triads were considered if both drug-target and target-metabolite interactions had been measured in the same activity units (Ki, Kd, IC50 or EC50). In order to complement the metabolite-target interaction list, we used Chembl database[14] to search for metabolite-target affinities that are not found in GtoPdb with the same activity units than the drug-target one. A total of 2845 drug-target-metabolite triads were

ultimately identified, involving 403 drugs, 169 human targets and 71 endogenous metabolites (Table S1 in the supplementary information). A total of 602 of these triads corresponded to a mechanism of action target (primary) for the drug and 2243 to a secondary target.

For both primary and secondary drug pharmacology we inspected the number of drug-target interactions presenting an affinity at least two orders of magnitude below the corresponding metabolite affinity for the target. As suggested in previous studies, this 100-fold difference between the drug and the human endogenous metabolite could be used as a baseline in order to discriminate biologically non-relevant interactions from the ones that will produce a biological effect.[12] In the whole set of 2845 drug-target-metabolite triads, we determined that a 16.73% of the interactions are, in fact, non-significant. This percentage decreases to 4.49% for drug-target interactions involving primary targets, and increases to 20.02% for secondary targets (Figure 1A). This clear difference in the number of non-relevant interactions between primary and secondary drug interactions can be attributed to the fact that drugs are carefully optimised to interact specifically with their primary targets,[12] so the number of non-relevant interactions for primary pharmacology is expected to be insignificant compared with off-targets. The few cases of non-relevant primary interactions could be attributed to prodrugs or to possible data errors. On the other hand, we classified a 20%

of the secondary interactions as non-relevant. Off-target interactions constitute the majority of drug-target interactions (78.84%) and consequently the main culprit for the polypharmacology of these drugs. According to our results, the impact of the endogenous metabolome is relevant on secondary targets rather than on primary pharmacology. Consequently, a more realistic estimation of the significant drug-target interaction space reduces the number of interactions with off-targets, thus reducing the candidates to produce the safety events reported for drugs.

The Anatomical Therapeutic Chemical (ATC) classification system was used to stratify drugs and analyze the impact of the non-relevant interactions in each one of these drug groups. Under this classification, a drug can actually be part of more than one group depending on the tissue they are intended to affect or their therapeutic properties. Of the 403 drugs, 372 are classified in at least one ATC code. A total of 12 drug groups of the ATC anatomical-level were considered: (A) Alimentary tract and metabolism system (number of drugs (d) = 43, % of non-relevant interactions (%n) = 19.05%), (B) Blood and blood forming organs (d=7, %n=28.13%), (C) Cardiovascular system (d=74, %n=16.47%), (D) Dermatological drugs (d=29, %n=17.86), (G) Genitourinary system and reproductive hormones (d=44, %n=24.95%), (H) Systemic hormonal preparations, excluding sex hormones and insulins (d=25, %n=28.75%), (L) Antineoplastic and

immunomodulating agents (d=18, %n=29.63%), (M) Musculoskeletal system (d=4, %n=14.29%), (N) Nervous system (d=102, %n=13.91%), (R) Respiratory system (d=71, %n=14.92%), (S) Sensory organs (d=39, %n=19.92%) and (V) Various ATC structures (d=7, %n=24.14%). The number of drug-target interactions involved in each one of the 12 drug classes aforementioned has been analyzed individually and presented in Figure 1B. As can be observed, a substantial number of drug-target interactions are annotated to the nervous system ATC anatomical-level, while some other groups like B, M or V have a really low number of interactions. The right panel of Figure 1B represents a comparison of the percentage of drug-target interactions in all ATC groups, allowing us to visualize the percentage of non-relevant interactions (plotted in black) among the different groups. So, although the percentage of non-relevant interactions changes depending on the ATC group, there are no significant variations between them, being the average of non-relevant interactions across all the groups of 16.73%.

A finer grouping of drugs corresponding to the third level of the ATC system classification was also inspected. We only considered those groups with at least 20 drug-target interactions. Some variability can be observed over the different drug groups (Figure S1 in the supplementary information). We can recognize several groups presenting percentages of non-relevant interactions significantly higher

than the rest, like opioids (N02A, %n=42.42%), estrogens (G03C, %n=38.67%) or Hormones and related agents (L02A, %n=45.45%). The list of drug groups with the number of drug-target interactions and the percentage of non-relevant interactions for each one of them is provided in Table S2 in the supplementary information. Some pharmacological drug classes seem to have more predisposition to be affected by metabolites, and consequently present a higher number of non-relevant interactions. However, some metabolite targets have been well-studied for decades while other groups of targets may not be so explored and thus, the lower metabolomic impact we observe on these groups could also reflect this lack of data completion.

In order to better analyze the impact of the metabolome on different drug classes, we generated a tripartite network with drugs, targets and metabolites. Removing all non-relevant interactions from the network allows us to compare the full polypharmacology network of a certain group of drugs with the potential drug-protein interaction network considering the influence of the endogenous metabolome. The Cytoscape software platform[16] was used to construct, format and visualize the interaction networks between drugs, metabolites, and proteins.

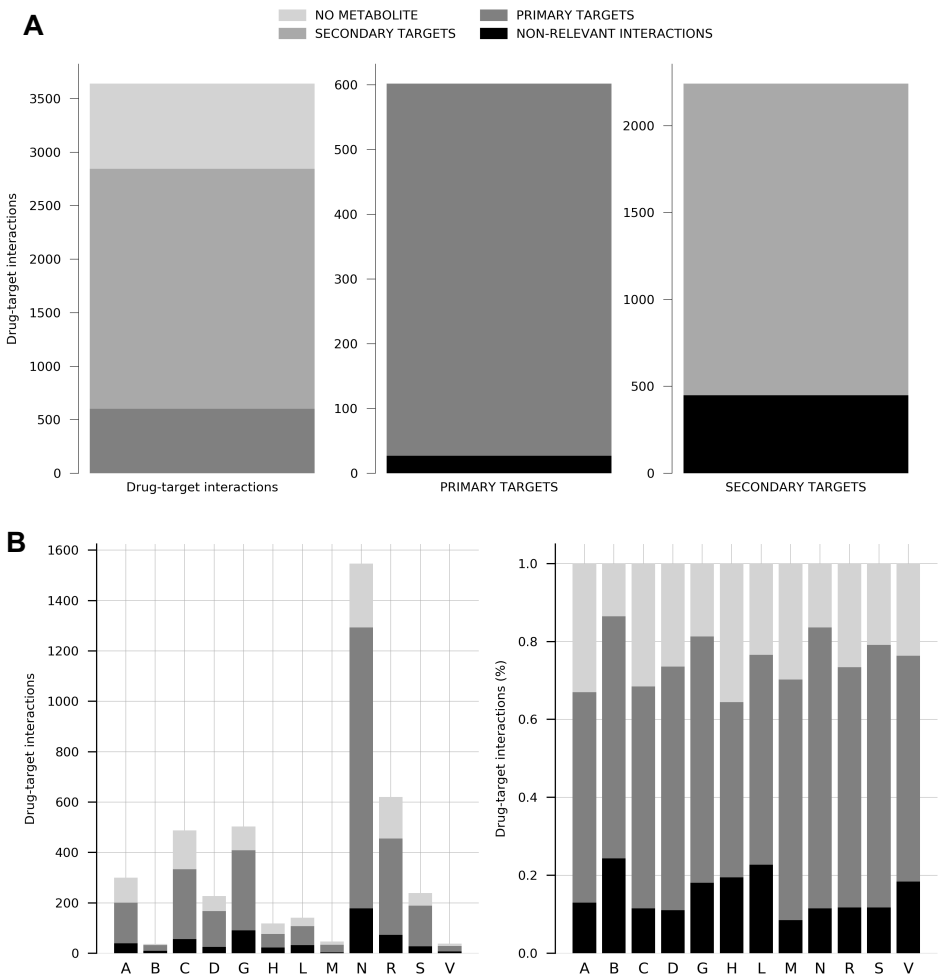


Figure 1: Drug-target distributions. **A)** Bar plots reflecting the total number of drug-target interactions and the number of interactions involving primary or secondary targets. The first bar plot shows the number of drug-target interactions without a known association of a human endogenous metabolite with this target (light gray), and the interactions with metabolite information, splitted between primary (dark gray) and the secondary (gray) targets. The second and third bar plots represent the primary and secondary drug-target interactions, divided between the biologically relevant and non-relevant (black) interactions. **B)** Absolute and relative drug-target distribution over each one of the 12 drug groups of the ATC anatomical-level. The number of non-metabolite associated drug-target interactions (light gray), biologically relevant (dark gray) and biologically non-relevant (black) interactions are represented in different colors in both bar plots.

Metabolomic impact on estrogens

Estrogens (G03C) are used mostly in menopausal hormone therapy and hormonal birth control, but also for many other indications like cancer treatment or feminizing hormone therapy.[17,18] The estrogen pharmacological group is formed by natural and semisynthetic estrogens (G03CA), synthetic estrogens (G03CB) and estrogens in combinations with other drugs (G03CC). Specifically, the 5 estrogen drugs under study are: Diethylstilbestrol (G03CB, G03CC, L02AA), estradiol (G03CA), estriol (G03CA, G03CC), estrone (G03CA) and ethinylestradiol (G03CA, L02AA).

The derived drug-target network is shown in Figure 2A and consists of 75 interactions between these five aforementioned drugs and 43 target proteins. The data regarding these interactions can be found in Table S3 in the supplementary information. The average number of target interactions per drug is 15.0, and the average of drug interactions per target is 1.74. The node size is proportional to the number of interactions in which the molecule participates. All five drugs are reported to exert their therapeutic effect through the same primary target, the estrogen receptor (solid edges in the network). Just 29 of the 75 drug-target interactions had drug affinity values at least two orders of magnitude below the metabolite affinity for the same target and are considered non-relevant (light-gray edges). These non-relevant drug-

target interactions represent the 38.67% of the drug-target-metabolite triads known. Figure 2B shows the same polypharmacology network without these 29 non-relevant interactions. As can be observed, not only the number of interactions has decreased, but the number of targets on the network is also lower. A total of 16 targets have been removed from the estrogen drugs polypharmacology network due to the absence of relevant interactions, representing 38.10% of the initial targets. So, the new estimated polypharmacology of estrogen drugs consists in 46 drug-target interactions between the five estrogen drugs and 24 targets. In this filtered polypharmacological network, the average number of target interactions per drug is 9.2 and the average of drug interactions per target is 0.85. This case exemplifies how comparing the affinity of endogenous ligands with the affinity of drugs can be a good method to reduce and narrow down drug polypharmacology by keeping only significant targets, which is key in drug safety risk evaluation.

The metabolite-drug network (Supplementary Figure S2) is formed by a total of 34 interactions between these five drugs and 17 human endogenous metabolites. The average number of target interactions between drugs and metabolites is 2.11 and is represented by the size of the edges. The node size reproduces the number of interactions in which the molecule participates, reflecting its relevance inside the group. As can be observed in this network, the most relevant metabolite is 17B-estradiol, which is the principal endogenous

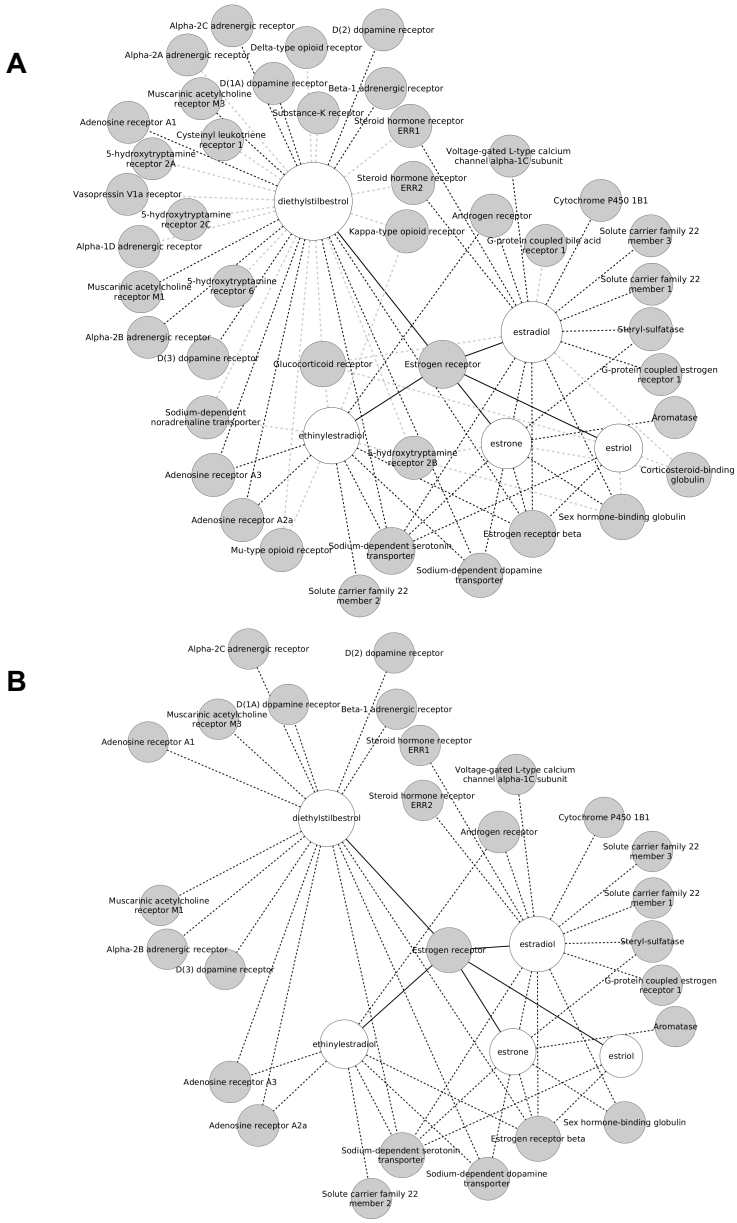


Figure 2: Full versus filtered polypharmacology of estrogen drugs (G03C). White nodes stand for drugs and gray nodes stand for human protein targets. Solid line edges correspond to primary/on-target interactions and dashed lines to secondary/off-target interactions. **A)** Full drug-target interaction network. Light gray edges correspond to non-relevant interactions. **B)** Filtered drug-protein interaction network, where non-relevant interactions have been removed from the network.

metabolite of the estrogen receptor. The edge color reflects the percentage of non-relevant drug-target interactions in which this metabolite is participating (being black 100% and light gray 0%), helping to identify metabolites with a key role in non-relevant interactions. In this regard, cortisol (7 non-relevant interactions (n)=7, percentage of non-relevant interactions (%n)=100%) and serotonin (n=5, %n=100%) are the two metabolites with a greater impact in the estrogens group. Other metabolites like dynorphin B, dynorphin A, adrenaline, neurokinin A, lithocholic acid or vasopressin are also important in order to discriminate and clean non-relevant interactions.

Metabolomic impact on opioid drugs

Opioids (N02A) are one of the most used analgesics drug classes. Although some of them also have antitussive and antidiarrheal action (codeine and loperamide), their main indication is to treat severe pain.[19] In this study we analyse 9 opioid drugs: oxycodone, nalbuphine, buprenorphine, morphine, hydromorphone, fentanyl, butorphanol, tapentadol and buprenorphine. We found 31 different interactions between these 9 opioids and 8 human targets (Table S4 in supplementary information). The resulting drug-target network is plotted in Figure 3A. The average number of interactions per drug is 3.44 and the average of drugs per target is 3.875. As can be observed, the principal mechanism of action target

(solid edges in Fig. 3A) is conducted via the mu-type opioid receptor, being annotated as primary target for all the 9 studied opioids. However, the other two opioid receptors, delta and kappa, are also assigned as primary targets for some of the drugs in the group.

When taking into account the metabolomic information, we determined that 12 of these drug-target interactions had affinities at least two orders of magnitude below the endogenous metabolite affinity for the same target (light-gray edge color in Fig. 3A). The new estimated opioid polypharmacology network is shown in Figure 3B and is formed by 19 interactions between the 9 aforementioned opioids and 6 human protein targets. There are two targets that are found biologically non-relevant regarding the polypharmacology of these 9 opioid drugs. The corrected average number of interactions per drug is 3.8 and the average of drug interactions per target is 3.17. Interestingly, we found a drug (tramadol) that apparently loses all its mechanism of action interactions. Tramadol is an analgesic drug and theoretically exerts its therapeutic function through the μ opioid receptor. However, previous studies have already suggested that its effect could be conducted by one of the products of its metabolism, which has a higher affinity for the μ opioid receptor. The interaction of the drug metabolite with the receptor would be considered significant, as its affinity is above the baseline of two log units.[20]

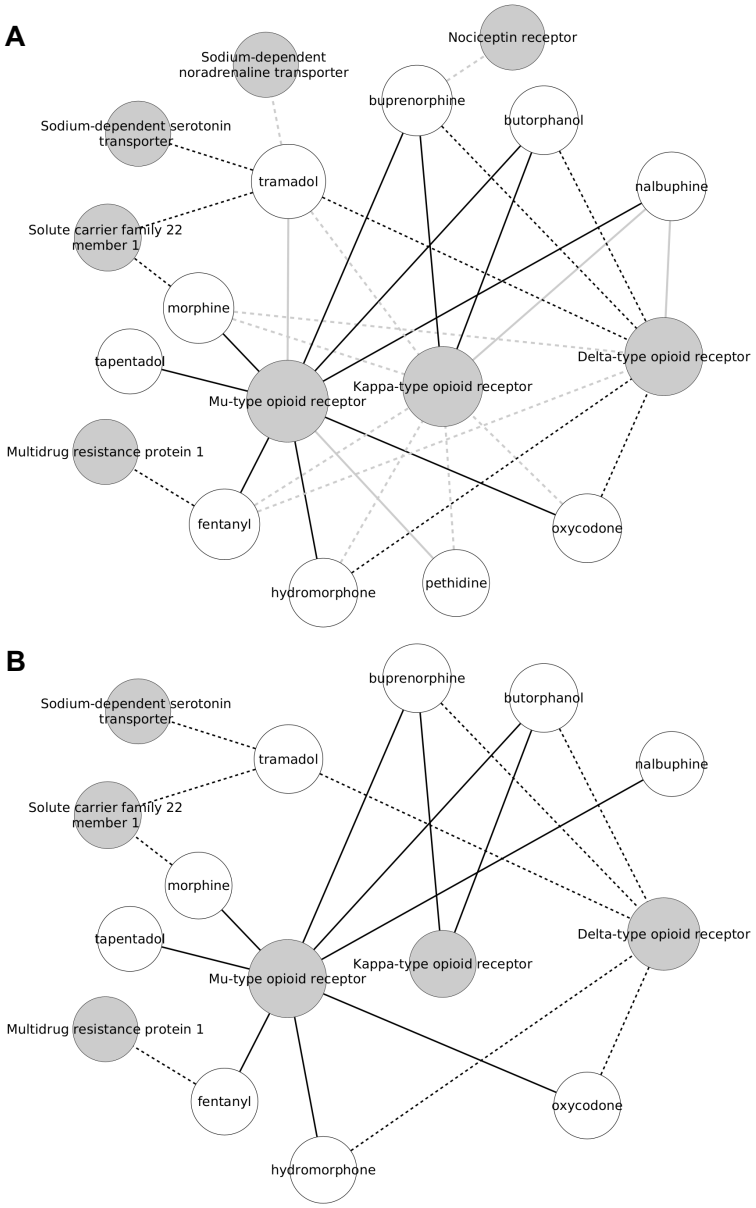


Figure 3: Full versus filtered polypharmacology of opioids drugs (N02A). White and gray nodes represent drugs and human protein targets, respectively. Solid line edges correspond to primary interactions and dashed lines to secondary interactions. **A)** Full drug-target interaction network. Non-relevant interactions are plotted in light gray color .**B)** Potential drug-protein interaction network filtered based on the endogenous metabolome. Non-relevant interactions have been removed from the network.

The metabolite-drug network can be found in Figure S3 in supplementary materials. A total of 31 interactions between these 9 opioids agents and 8 human endogenous metabolites are depicted in it. As can be seen, dynorphin B is the endogenous molecule participating in the largest number of target interactions as a principal endogenous metabolite, followed by dynorphin A and L-enkephalin. However the metabolites with a greatest impact over the opioids groups that help to determine the non-relevant interactions are dynorphin A (non-relevant interactions (n)=5, percentage of non-relevant interactions (%n)=71.43%, L-enkephalin (n=3, %n=42.86%) and nociceptin (n=2, %n=100%).

Conclusions and Discussion

The development of metabolomics is an important step forward in several scientific disciplines, including drug discovery. The determination and quantification of metabolites under specific conditions, such as genetic alterations, environmental factors (diet, pollution, smoking, age) or several diseases, can improve the understanding of drug effects in individual human bodies. This knowledge will allow us to create more effective and safer personalized drug treatments, as drug safety will be predicted in a more precise and personalized manner. However, metabolomics is still a young and relatively unexplored field. The real number of different human metabolites is still unknown to the present date, since a huge

part of them have still not been neither identified nor quantified nor target-profiled. Estimated approaches have reported ranges from 3,000 essential metabolites to approximations of 20,000 unidentified metabolites that are not essential for growth and development but could be of significant importance for prognosis, diagnosis and for the identification of surrogate markers for different disease conditions as well as for a better understanding of applied translational systems biology.[21,22] Despite just considering a small portion of the human endogenous metabolome, the presented approach already reveals an inherent role of human endogenous metabolites in drugs bioactivity.

This work explores the potential impact that the human endogenous metabolome has on drug pharmacology. Based on evidence, we propose a reference baseline to discriminate the interactions that will produce a biological effect from biologically non-relevant interactions. The cases of estrogen and opioids drugs reveal how some drug-target interactions are, in fact, non-relevant and not likely to cause any of the safety events associated with the administration of the drug. This reduction of the drug polypharmacology space allows us to better understand the real behaviour of drugs in biological systems and help to identify the main liabilities involved in the secondary pharmacology of drugs. The ultimate implication of the impact of the endogenous metabolome in drug polypharmacology is the implementation of a more

personalized medicine, where individualized treatments based on personal levels of endogenous metabolites are taken into account to minimize risk factors associated with drugs

References

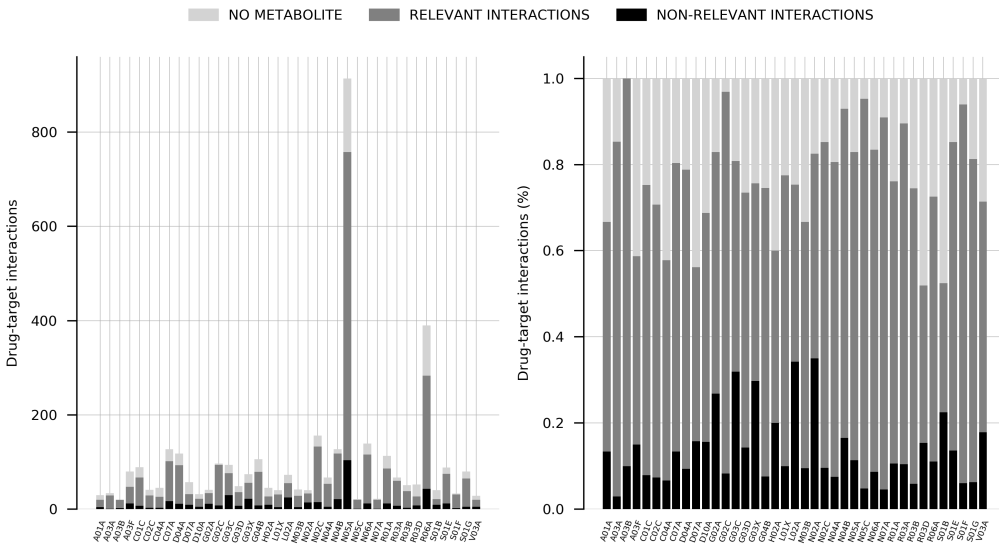
1. Jalencas, X. & Mestres, J. On the origins of drug polypharmacology. *MedChemComm* vol. **4** 80–87 (2013).
2. Mestres, J., Gregori-Puigjané, E., Valverde, S. & Solé, R. V. The topology of drug–target interaction networks: implicit dependence on drug properties and target families. *Molecular BioSystems* vol. **5** 1051 (2009).
3. O'Hagan, S. & Kell, D. B. Understanding the foundations of the structural similarities between marketed drugs and endogenous human metabolites. *Front. Pharmacol.* **6**, 105 (2015).
4. Liu, X. & Locasale, J. W. Metabolomics: A Primer. *Trends in Biochemical Sciences* vol. **42** 274–284 (2017).
5. Kell, D. B. & Oliver, S. G. The metabolome 18 years on: a concept comes of age. *Metabolomics* vol. **12** (2016).
6. Zamboni, N., Saghatelian, A. & Patti, G. J. Defining the metabolome: size, flux, and regulation. *Mol. Cell* **58**, 699–706 (2015).
7. Dettmer, K. & Hammock, B. D. Metabolomics--a new exciting field within the 'omics' sciences. *Environ. Health Perspect.* **112**, A396–7 (2004).
8. Nicholson, J. K., Everett, J. R. & Lindon, J. C. Longitudinal pharmacometabonomics for predicting patient responses to therapy: drug metabolism, toxicity and efficacy. *Expert Opin. Drug Metab. Toxicol.* **8**, 135–139 (2012).

9. Wang, T. J. *et al.* 2-Aminoadipic acid is a biomarker for diabetes risk. *J. Clin. Invest.* **123**, 4309–4317 (2013).
10. Palmer, N. D. *et al.* Metabolomic profile associated with insulin resistance and conversion to diabetes in the Insulin Resistance Atherosclerosis Study. *J. Clin. Endocrinol. Metab.* **100**, E463–8 (2015).
11. Shah, N. J., Sureshkumar, S. & Shewade, D. G. Metabolomics: A Tool Ahead for Understanding Molecular Mechanisms of Drugs and Diseases. *Indian J. Clin. Biochem.* **30**, 247–254 (2015).
12. Bofill, A., Jalencas, X., Oprea, T. I. & Mestres, J. The human endogenous metabolome as a pharmacology baseline for drug discovery. *Drug Discov. Today* **24**, 1806–1820 (2019).
13. Ursu, O. *et al.* DrugCentral 2018: an update. *Nucleic Acids Res.* **47**, D963–D970 (2019).
14. Gaulton, A. *et al.* The ChEMBL database in 2017. *Nucleic Acids Research* vol. **45** D945–D954 (2017).
15. Sharman, J. L. *et al.* Accessing Expert-Curated Pharmacological Data in the IUPHAR/BPS Guide to PHARMACOLOGY. *Curr. Protoc. Bioinformatics* **61**, 1.34.1–1.34.46 (2018).
16. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

17. Kuhl, H. Pharmacology of estrogens and progestogens: influence of different routes of administration. *Climacteric* **8 Suppl 1**, 3–63 (2005).
18. Wesp, L. M. & Deutsch, M. B. Hormonal and Surgical Treatment Options for Transgender Women and Transfeminine Spectrum Persons. *Psychiatr. Clin. North Am.* **40**, 99–111 (2017).
19. KuKanich, B. & Wiese, A. J. Opioids. *Veterinary Anesthesia and Analgesia* 207–226 (2017) doi:10.1002/9781119421375.ch11.
20. Lai, J., Ma, S. W., Porreca, F. & Raffa, R. B. Tramadol, M1 metabolite and enantiomer affinities for cloned human opioid receptors expressed in transfected HN9.10 neuroblastoma cells. *Eur. J. Pharmacol.* **316**, 369–372 (1996).
21. Kouskoumvekaki, I. & Panagiotou, G. Navigating the human metabolome for biomarker identification and design of pharmaceutical molecules. *J. Biomed. Biotechnol.* **2011**, (2011).
22. Schmidt, C. W. Metabolomics: what's happening downstream of DNA. *Environmental Health Perspectives* vol. **112** (2004).

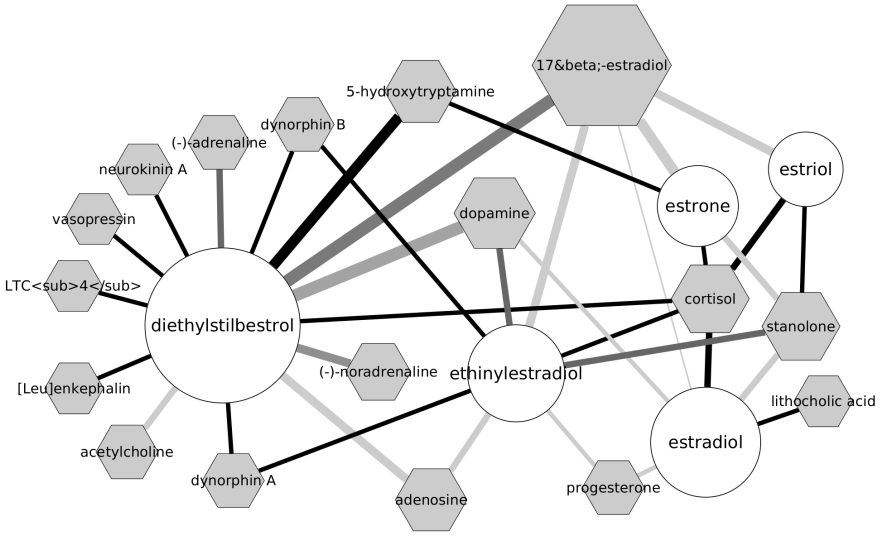
Supplementary Materials

Supplementary Figure 1:



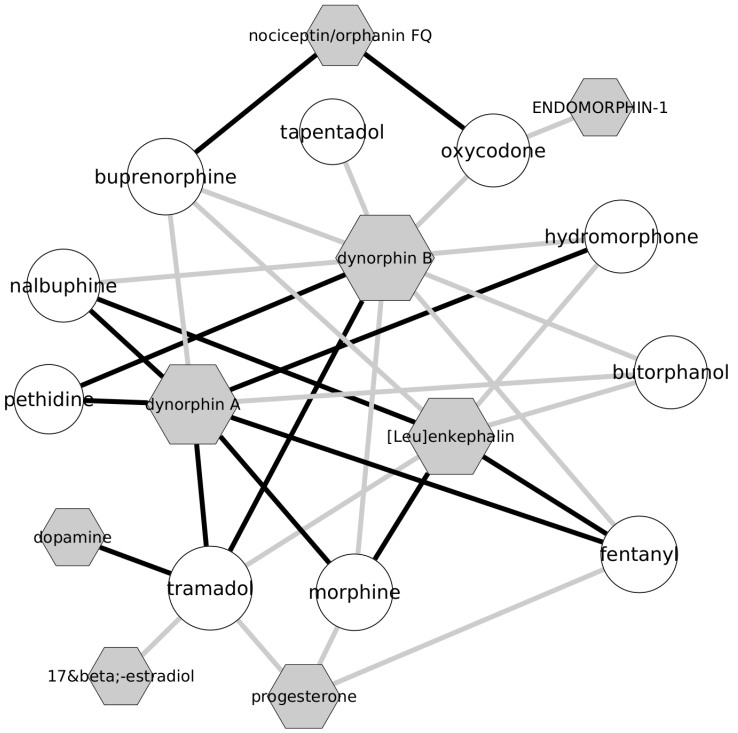
Supplementary Figure 1: Absolute and relative drug-target distribution over each one of the pharmacological drug subgroups corresponding to the third level of the ATC system classification. There are only represented these pharmacological groups with drugs participating in at least 20 drug-target interactions. The number of non-metabolite associated drug-target interactions (light gray) and relevant (dark gray) and non relevant (black) drug-target-metabolite interactions are represented in different colors in both bar plots.

Supplementary Figure 2:



Supplementary figure 2: Metabolite-drug network. It is formed by 34 interactions between 5 estrogen drugs and 17 human endogenous metabolites. Drugs and metabolites are represented as white circles and grey hexagons, respectively. The size of all the nodes (both drugs and metabolites) represents the number of interactions in which the molecule participates. The size of the edges represents the number of target interactions that relate these drug and metabolite edges. The black (1) to white (0) scale of colors of the edges represent the % of non-relevant drug-target interactions where this specific metabolite participates.

Supplementary Figure 3:



Supplementary figure 3: Metabolite-drug network. It is formed by 31 interactions between 9 opioid drugs and 8 human endogenous metabolites. Drugs and metabolites are represented as white circles and grey hexagons, respectively. The size of all the nodes (both drugs and metabolites) represents the number of interactions in which the molecule participates. The size of the edges represents the number of target interactions that relate these drug and metabolite edges. The black (1) to white (0) scale of colors of the edges represent the % of non-relevant drug-target interactions where this specific metabolite participates.

Supplementary tables are enclosed in the next links:

Supplementary Table 1: List of 2845 drug-target-metabolite triads involving 403 drugs, 169 human targets and 71 endogenous metabolites.
https://drive.google.com/file/d/1s7EEbgljcP6tTJtJJ0Z4ICBVoPJ8Htj5

Supplementary Table 2:
List of drug groups of the ATC pharmacological-level. The number of drug-target interactions and the percentage of non-relevant interactions for each one of these drug groups is also provided.
https://drive.google.com/file/d/15WANigLEfoHeSz8kgBhRtJXYsPuAmfvD

Supplementary Table 3:
List of drug-target-metabolite interactions related with estrogens.
https://drive.google.com/file/d/12NzGjKPf810IWJ8xJRINAVpM2NQsKUWp

Supplementary Table 4:
List of drug-target-metabolite interactions implicating opioid agents.
https://drive.google.com/file/d/1GX9cT5F90ItSmZ_9iwtCmNu6n4M2Bh0G

3. The impact of the human endogenous metabolome on natural compounds

Andreu Bofill[1]. Xavier Jalencas[1] 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.

* Corresponding authors: jmestres@imim.cat

Abstract

Natural compounds have been present in the environment of humans since their origin as a wide variety of chemical substances produced by other living organisms. Although some of them are known to produce therapeutic or toxic effects, the majority do not produce any effect in the human organism. Differently to drugs and endogenous metabolites, natural compounds have not been optimized neither by humans nor evolution to produce any biological effect by interacting with any human protein. It has been demonstrated in a previous work that the affinity needed by drugs to interact with their targets and produce an effect is related to the affinity of the human endogenous metabolites for those proteins. In the present work we demonstrate that in the same way as it happens with drugs, there is a relation between the affinity of the endogenous metabolites for their native proteins and the minimum affinity needed by natural compounds to interact with those proteins and produce a significant effect.

Introduction

In the last 30 years, omics technologies, such as genomics, transcriptomics and proteomics have gained prominence and have become important players in many scientific disciplines.[1] However, the gene expression assembly line from DNA to RNA to proteins did not seem to reflect the entire complexity of biological systems.[2] Therefore, metabolomics emerged in order to study small chemical compounds, known as metabolites, at a cellular, tissular and organismal level.[2] These metabolites, including both endogenous and exogenous molecules, are collectively referred as the metabolome and constitute the substrates and products of all chemical reactions that take place in an organism.[3] Accordingly, changes in the metabolome are the ultimate answer of an organism or biological system to environmental or genetic alterations.[4] Metabolomics represents a step forward towards a more holistic understanding of biological processes and its promising contributions to biomedical research and drug discovery have been coming out in the last years.[5–7]

The environmental contribution not only to disease appearance, but also to the pharmacology and bioactivity of drugs or to the chronification of several conditions is widely recognized.[8,9] For example, the ingestion of specific food compounds, traditional medical plants and other natural compounds have been associated in many occasions to

metabolite changes or even to several illnesses such as diabetes, obesity, cardiovascular diseases or cancer.[10–14] However, the majority of food compounds and natural products in our environment do not exert any effect in our organism, neither good nor bad. This is probably because these natural products are not biologically optimised to interact with human proteins to produce any effect.

We demonstrated in a previous work that the evolutionary optimised affinity of endogenous ligands for their native proteins can serve as a baseline for the primary pharmacology of drugs.[15] Moreover, we proposed that the minimum level of affinity that is needed for a drug to be able to modulate the activity of a specific protein is given by the optimised affinity of the native endogenous ligand for this protein. The presence of safety liabilities arising from the secondary pharmacology of drugs seems also to be clearly dependent on this metabolite affinity baseline.[15] Furthermore we suggested that this baseline could serve to differentiate non-relevant interactions from those that will produce a biological effect.

In the present work, we aimed to study the role of the human endogenous metabolome in the activity that natural compounds have in the human body. We introduce a new approach to better estimate the pharmacology of natural compounds based on the impact of the endogenous

metabolome. A comparison between drugs and natural compounds was done in order to contrast the different implications the endogenous metabolome has in both types of exogenous molecules.

Natural compounds and endogenous ligands

The FooDB database[16] was used to obtain an accurate and comprehensive list of 14,186 food constituents. FooDB is the most complete resource and the reference database on chemical food constituents associated with an extremely rich metadata, including a complete collection of biochemical, compositional and physiological data.[17] In parallel, we extended the list of natural compounds with 4,385 herbal compounds obtained through a meticulous and exhaustive research of literature. A manual curation of each one of these compounds was carried out to ensure its natural origin. In order to clearly differentiate between human endogenous metabolites and natural exogenous compounds, all molecules annotated in the Guide To Pharmacology (GtoPdb) database[18] as human endogenous metabolites were excluded from our list. The final list of natural compounds was used to interrogate ChEMBL[19] and GtoPdb[18] databases in order to extract all known interactions between these natural compounds and human targets. A total of 2,996 interactions involving 856 molecules and 656 human protein targets were obtained. The

list of human proteins involved in those interactions was also used to extract from GtoPdb[18] each of their native human endogenous ligands and extract their reported maximum affinity value. Only those compound-target-metabolite triads with both interactions measured with the same activity units (Ki, IC50 or EC50) were accepted. A total of 260 compound-target-metabolite triads were identified, involving 96 natural compounds, 67 human targets and 38 endogenous metabolites. 86.21% of these triads involved Ki binding affinities while 10.34% of them involved IC50 affinities and just the 3.45% corresponded to EC50 values. As much as 124 of these 260 triads involved 25 natural compounds that are approved as drugs in at least one country. On the contrary, 136 of these triads involved compounds that have never been approved as drugs.

In order to differentiate and compare the set of natural compounds from the set of drugs, we just took into account the natural compounds that have never been approved as drugs. So, our final natural compound set included 136 compound-target-metabolite triads involving 71 natural compounds, 47 human proteins and 28 endogenous metabolites (Table S1 in the supplementary information). Noteworthy, 16 of these 71 natural compounds are being investigated in different phases of clinical trials.

The 136 pairs of natural compounds and metabolite affinities for the same protein are plotted in Fig. 1, where the metabolite and the compound affinities for a specific target are represented in the x and y axis respectively. The median of natural compound affinities was over 30-fold lower than the corresponding metabolite affinities, with median negative logarithm affinities of 5.76 and 7.35, respectively. On a perfect correlation between compound and metabolite affinities, the data points would follow the solid black diagonal line. As can be observed, there is a clear tendency for natural products to present lower affinities for their targets than their principal native endogenous metabolites. In fact, this density plot reveals that most of the points (83.09%) are accumulated below the diagonal line (solid-grey line in Fig. 1), which corresponds to compound affinities for their targets being lower than the principal endogenous metabolite affinity for the same targets. Indeed, 44.12% of these compound affinities were even lower than the corresponding metabolite-target interactions minus two orders of magnitude.

On the contrary, just for three of these 136 natural compounds (2.21%) the affinities to their targets were over 100-fold higher than their corresponding metabolite affinities (dashed-gray top line in Fig. 1), increasing to 9 (6.62%) with a threshold at 10-fold. A close examination of these top 9 interactions suggests that although they have not been clearly associated with any therapeutic use and no drugs have been

developed based on any of them, they are involved in biologically relevant interactions with human proteins.

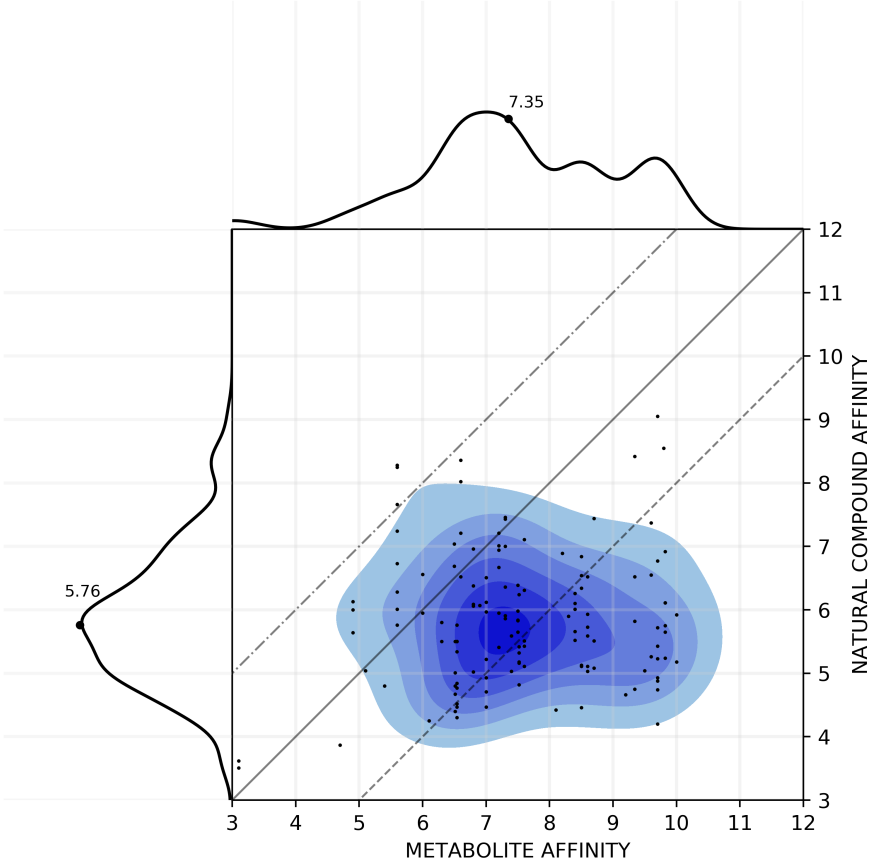


Figure 1: Density plot of the 136 compound-target-metabolite triads. The natural compound affinities for their target(s) are plotted against the corresponding endogenous metabolite affinities. A kernel density estimation is shaden in blue tones, darkest blue corresponding to highest density regions. Solid line corresponds to natural compound affinities being equal to the corresponding metabolite affinities for the same protein, while the two dashed lines correspond to the natural compound affinities being 2 orders of magnitude higher or lower than the corresponding metabolite affinities for the same protein, respectively. The distributions of natural compound affinities (left) and metabolite affinities (top) are included, with the median negative logarithm values.

For example, isocorypalmine is a mono-demetyhlated analog of tetrahydropalmatine and can be found (also with tetrahydropalmatine itself) in plants of the genus *Corydalis*, which include one of the most frequently used herbs to treat drug addiction in the traditional Chinese medicine.[20,21] The administration of isocorypalmine in mouse models is linked to a reduction of the cocaine-induced locomotor hyperactivities, being this compound a promising agent for treatment of cocaine addiction.[21,22] Isocorypalmine is annotated with binding affinities (pKi) of 8.28, 8.02, 7.43, 7.21 and 7.11 for the dopamine D1, D5, D3, D2 and D4 receptors respectively, being the difference between these interactions and the endogenous metabolite (dopamine) affinity for the same receptors of 2.68, 1.42, 0.13, 0.01 and -0.49. On the contrary, tetrahydropalmatine presents binding affinities (pKi) of 6.73, 6.52, 5.95 and 5.86 for Dopamine D1, D5, D2 and D3 receptors respectively, being its differential with the principal endogenous metabolite (dopamine) for each one of these targets of 1.13, -0.08, -1.25 and -1.44 orders of magnitude. These results suggest that the action of tetrahydropalmatine might be conducted mainly through the interaction with D1 and D5 receptors, but probably not through its interaction with D2 and D3. On the other hand, isocorypalmine presents binding affinities for D3 and D2 receptors similar to dopamine, suggesting that this compound could also interact with these two receptors. Along these lines, some studies reported that,

in mouse models, isocorypalmine acts as a high-affinity partial agonist of D1 and D5 dopamine receptors and a moderate-affinity antagonist of D2, D3 and D4 receptors, while tetrahydropalmatine just exert its actions through an activation of D1 and D5 dopamine receptors.[21]

Similarly, stepholidine, which is an active ingredient of *Stephania intermedia* plant, seems to have a strong binding affinity (pKi) for dopamine D1 and D5 receptors (8.25 and 8.36, respectively). In fact, the stepholidine affinity for dopamine D1 receptor is more than 2.5 orders of magnitude higher than the affinity of the principal endogenous metabolite for the same receptor, dopamine, while the difference for dopamine D5 receptor is more than 1.5 orders of magnitude. At the same time, we also found stepholidine binding affinity information for dopamine D2, D3 and D4 receptors (6.94, 7.00, 5.43, respectively). For dopamine D2 and D3 receptors, these binding affinities seem to be similar to the affinity of dopamine for each one of these targets (differences of -0.26 and -0.30 orders of magnitude). However, the affinity of stepholidine for dopamine D4 receptor is more than two orders of magnitude lower than the affinity of dopamine for the same target, suggesting that this natural compound is unlikely to conduct any effect through the activation of dopamine D4 receptor. A huge number of studies have been done regarding the possible role of stepholidine as a pan-dopamine receptor antagonist. It is proposed that stepholidine may possess dopamine D1-like

receptor (dopamine D1 and D5 receptors) agonistic and D2-like receptor (D2, D3, D4) antagonistic properties.[23–26] However, our results may suggest that stepholidine may not be active on dopamine D4 receptor. The observed dopamine-effect observed may be conducted through the activation of the other dopamine receptors, but not likely through the activation of dopamine D4 receptor, like some literature suggests.[24] Some studies in rodents proposed stepholidine as an attenuant of drug-induced reinforcement, making it a promising drug candidate to treat drug abuse.[25,27–29]

Another example involves muscarine, a well-known natural compound that can be found in certain types of mushrooms, such as *Clitocybe* or *Inocybe* species.[30] It is a parasympathomimetic substance that interacts with muscarinic acetylcholine receptors (actually giving the name to this group of acetylcholine receptors), causing a huge activation of the peripheral parasympathetic nervous system.[30] In this regard, we found binding affinity values of 6.00, 7.04 and 6.01 for muscarinic acetylcholine receptors M1, M2 and M4, respectively. All of these interactions have higher binding affinities than their principal endogenous metabolite, acetylcholine. The well-known effect of muscarine through the activation of muscarinic receptors aligns perfectly with the proposed idea that the affinity of the endogenous metabolite for their native targets can be used to determine if an exogenous molecule is capable of interacting with these

targets and produce a significant effect beyond its binding affinity value alone.

All these examples clearly indicate that these natural compounds, although not being used as drugs effectively produce an effect in the human organism. So, although lots of these natural compounds have activities at least two orders of magnitude lower than the metabolite affinity for the same target (biologically non-relevant interactions), several of these interactions seem to be biologically relevant and responsible for the effects that these compounds could produce. These results also validate the importance of taking into account the endogenous metabolome in order to assess whether a chemical compound has any biological effect through the interaction with a specific protein.[15]

Drug *vs.* natural compound interactions

In order to compare the impact of the human endogenous metabolites affinities between natural compounds and drugs, we retrieved a list of 403 drugs from a previous work.[15] These 403 drugs represent all the DrugCentral drugs interacting with at least one human target which is involved in its mechanism of action and is also associated to a known principal endogenous metabolite according to GtoPDB[18] or Chembl[19] databases. To be able to fairly compare the

interactions of these drugs with the interactions of the set of natural compounds, only the maximum affinity value was taken into account for each natural compound interaction if several of them were available. Our final set was composed of 71 different natural compounds.

So, with these two sets we generated a density plot comparison, where drug affinities for their primary targets and natural compounds affinities for their maximum affinity targets are plotted against their corresponding endogenous metabolite affinity for the same target (Fig. 2). As can be observed, these two sets of compounds generate considerably different density estimations. Although the median of metabolite affinities between drugs and natural compounds is almost the same (7.25 and 7.52, respectively) and both groups seem to follow a relatively similar distribution, the median of both drugs and natural compounds affinities are extremely different. The median of drug affinities was found to be over almost 3 orders of magnitude higher than the median for natural compounds with median negative logarithm affinities of 8.43 and 5.76, respectively. This result could have led us to conclude that the compounds that are known to produce a therapeutic effect just need a higher affinity than the other ones. However, we can observe how the tendencies of both density plots are also remarkably different. The drug density plot (red points in Fig. 2) reveals that 67% of all drug-protein interactions have affinity values above the corresponding metabolite-protein affinities

and that up to 96% of those drug affinities have values within two orders of magnitude of metabolite affinities for the same protein (dashed line in Fig. 2). So, there is a trend for low drug

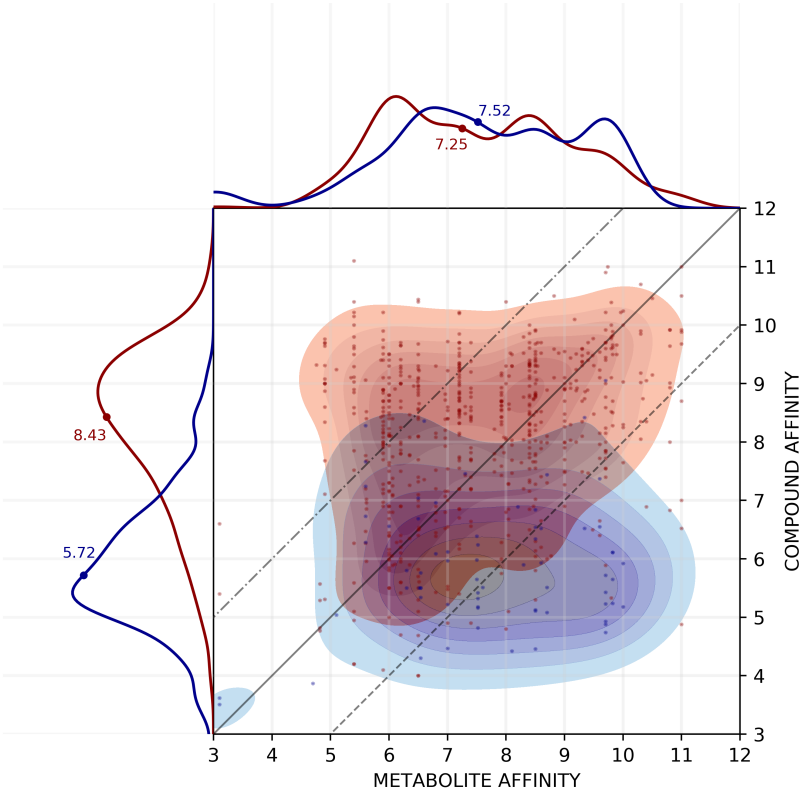


Figure 2: Density plot comparison between drugs (red) and natural compounds (blue) affinities. Drug affinities for their primary target(s) and natural compounds affinities for all their maximum affinity targets are plotted against the corresponding endogenous metabolite affinities. A kernel density estimation is shaded in red and blue tones, darkest corresponding to highest density regions. The solid line corresponds to natural compound affinities being equal to the corresponding metabolite affinities for the same protein, while the two dashed lines correspond to the natural compound affinities being 2 orders of magnitude higher or lower than the corresponding metabolite affinities for the same protein, respectively. Natural compound and drug (left) and metabolite (top) affinities distributions are included, with the median negative logarithm values, for both drug and natural compound groups in red and blue respectively.

affinities to correlate with low metabolite affinities and not with high ones, but with high drug affinities, the window of metabolite affinities is extremely much wider.[15] However, natural compounds do not follow this trend. It seems that natural compounds present a variable distribution of affinity values over the wide distribution of metabolite affinities for the same target. Indeed, the span of compound affinities does not differ between high and low metabolite affinities. As opposed to drugs, natural compound density plot (blue points in Fig. 2) reveals that 82% of natural compounds affinities have a maximum affinity value below the corresponding metabolite-target affinities and 48% of them have values below two orders of magnitude of the metabolite affinities for the same target.

As previously mentioned, a total of 25 compounds from the curated list of natural compounds collected in this study were actually approved drugs (Fig. S1 in the supplementary material). Analyzing the affinities of the maximum affinity target of those natural compounds, we can observe an extremely similar density distribution than the one given by all the 403 drugs. There is a trend for low such natural compounds affinities to correlate with low metabolite affinities and not with high ones. For high compounds affinities, the window of compound affinities is much wider, although the majority of compound affinities are correlated with high metabolite affinities too. There is a great correlation between these compounds and the metabolite affinities. So, for those natural

compounds that are expected to produce a therapeutic effect, their affinities for their maximum affinity targets are found to have an equal pattern to the entire set of drugs.

Conclusions and Discussion

Similarly to what happens with all drugs, natural compounds that produce specific known therapeutic effects through the interaction with a particular protein need to present a minimum level of affinity given by the optimised affinity of the endogenous ligand for this protein. However, the majority of natural compounds are not biologically optimised to produce any effect. As observed, almost half of these natural compounds have a maximum affinity interaction lower than the endogenous metabolite affinity for the same protein. Moreover a large number of them have an affinity value below the metabolite affinity minus two orders of magnitude. These results suggest that a considerable number of these interactions are non-relevant and, consequently, these compounds do not produce any effect on human bodies.

Furthermore, several of the natural compounds never approved as drugs have at least one target with an affinity higher than the corresponding metabolite affinity minus two orders of magnitude. This indicates that a biological effect could be expected when these exogenous natural compounds

would be introduced in the human organism, as we expose in the previous examples. Even though these products are not developed or approved yet as drugs, they are part of our environment as food or toxic products. In fact, some of these compounds could have been optimized by their original species to be toxic to their depredators as a defensive method. Such is the case of muscarine. Some other natural compounds could present, as previously seen, possible therapeutic effects not yet exploited or investigated.

In summary, the study of natural compounds can be a source of new chemical agents, either therapeutic or toxic, and the endogenous metabolome can contribute to better discern whether the interaction of a particular compound with a protein target is likely to produce an effect in the organism.

References

1. Fillet, M. & Frédérich, M. The emergence of metabolomics as a key discipline in the drug discovery process. *Drug Discov. Today Technol.* **13**, 19–24 (2015).
2. Kell, D. B. & Oliver, S. G. The metabolome 18 years on: a concept comes of age. *Metabolomics* vol. 12 (2016).
3. Zamboni, N., Saghatelian, A. & Patti, G. J. Defining the metabolome: size, flux, and regulation. *Mol. Cell* **58**, 699–706 (2015).
4. Dettmer, K. & Hammock, B. D. Metabolomics--a new exciting field within the 'omics' sciences. *Environ. Health Perspect.* **112**, A396–7 (2004).
5. Manach, C., Hubert, J., Llorach, R. & Scalbert, A. The complex links between dietary phytochemicals and human health deciphered by metabolomics. *Mol. Nutr. Food Res.* **53**, 1303–1315 (2009).
6. Trujillo, E., Davis, C. & Milner, J. Nutrigenomics, proteomics, metabolomics, and the practice of dietetics. *J. Am. Diet. Assoc.* **106**, 403–413 (2006).
7. Bar, N. *et al.* A reference map of potential determinants for the human serum metabolome. *Nature* (2020) doi:10.1038/s41586-020-2896-2.
8. Smith, S. E. & Rawlins, M. D. Drug Metabolism: Environmental Influences. *Variability in Human Drug Response* 84–109 (1976) doi:10.1016/b978-0-407-43301-4.50011-9.

9. Norman, R. E., Carpenter, D. O., Scott, J., Brune, M. N. & Sly, P. D. Environmental exposures: an underrecognized contribution to noncommunicable diseases. *Rev. Environ. Health* **28**, 59–65 (2013).
10. Brownell, K. D. Fast Food and Obesity in Children. *PEDIATRICS* vol. **113** 132–132 (2004).
11. Kozuka, C. *et al.* Natural food science based novel approach toward prevention and treatment of obesity and type 2 diabetes: Recent studies on brown rice and γ-oryzanol. *Obesity Research & Clinical Practice* vol. **7** e165–e172 (2013).
12. Fardet, A. & Boirie, Y. Associations between food and beverage groups and major diet-related chronic diseases: an exhaustive review of pooled/meta-analyses and systematic reviews. *Nutrition Reviews* vol. **72** 741–762 (2014).
13. Bingham, S. & Riboli, E. Diet and cancer — the European Prospective Investigation into Cancer and Nutrition. *Nature Reviews Cancer* vol. **4** 206–215 (2004).
14. Reedy, J. *et al.* Higher diet quality is associated with decreased risk of all-cause, cardiovascular disease, and cancer mortality among older adults. *J. Nutr.* **144**, 881–889 (2014).
15. Bofill, A., Jalencas, X., Oprea, T. I. & Mestres, J. The human endogenous metabolome as a pharmacology baseline for drug discovery. *Drug Discov. Today* **24**, 1806–1820 (2019).

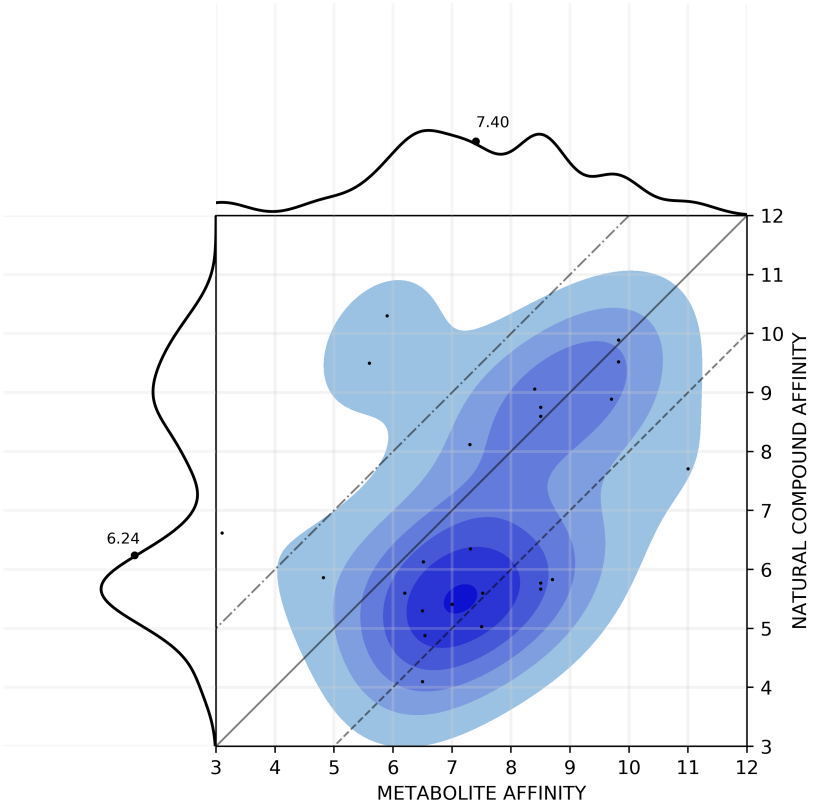
16. David Wishart, University of Alberta. FooDB version 1.0. *FooDB* https://foodb.ca.
17. Sorokina, M. & Steinbeck, C. Review on natural products databases: where to find data in 2020. Journal of Cheminformatics vol. **12** (2020).
18. Sharman, J. L. et al. Accessing Expert-Curated Pharmacological Data in the IUPHAR/BPS Guide to PHARMACOLOGY. Curr. Protoc. Bioinformatics **61**, 1.34.1–1.34.46 (2018).
19. Gaulton, A. et al. The ChEMBL database in 2017. Nucleic Acids Research vol. **45** D945–D954 (2017).
20. Min, X. et al. A database on treating drug addiction with traditional Chinese medicine. Addiction **102**, 282–288 (2007).
21. Xu, W. et al. L-isocorypalmine reduces behavioral sensitization and rewarding effects of cocaine in mice by acting on dopamine receptors. Drug Alcohol Depend. **133**, 693–703 (2013).
22. Ma, Z.-Z. et al. Isoquinoline alkaloids isolated from Corydalis yanhusuo and their binding affinities at the dopamine D1 receptor. Molecules **13**, 2303–2312 (2008).
23. Yang, K., Jin, G. & Wu, J. The neuropharmacology of (-)-stepholidine and its potential applications. Curr. Neuropharmacol. **5**, 289–294 (2007).
24. Meade, J. A. et al. (-)-Stepholidine is a potent pan-dopamine receptor antagonist of both G protein- and β-

arrestin-mediated signaling. Psychopharmacology vol. **232** 917–930 (2015).

25. Hicks, C. et al. Dopamine D1-Like Receptor Agonist and D2-Like Receptor Antagonist (−)-Stepholidine Reduces Reinstatement of Drug-Seeking Behavior for 3,4-Methylenedioxypyrovalerone (MDPV) in Rats. ACS Chemical Neuroscience vol. **9** 1327–1337 (2018).
26. Jin, G.-Z., Zhu, Z.-T. & Fu, Y. (−)-Stepholidine: a potential novel antipsychotic drug with dual D1 receptor agonist and D2 receptor antagonist actions. Trends in Pharmacological Sciences vol. **23** 4–7 (2002).
27. Yue, K. et al. L-Stepholidine, a naturally occurring dopamine D1 receptor agonist and D2 receptor antagonist, attenuates heroin self-administration and cue-induced reinstatement in rats. Neuroreport **25**, 7–11 (2014).
28. Wang, W. et al. The effect of L-stepholidine, a novel extract of Chinese herb, on the acquisition, expression, maintenance, and re-acquisition of morphine conditioned place preference in rats. Neuropharmacology **52**, 355–361 (2007).
29. Ma, B. et al. L-stepholidine, a natural dopamine receptor D1 agonist and D2 antagonist, inhibits heroin-induced reinstatement. Neurosci. Lett. **559**, 67–71 (2014).
30. George, P. & Hegde, N. Muscarinic toxicity among family members after consumption of mushrooms. Toxicol. Int. **20**, 113–115 (2013).

Supplementary Materials

Supplementary Figure 1:



Supplementary Figure 1: Density plot of the 25 drug-like natural compounds – target-metabolite triads. The natural compounds affinities for their 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 line corresponds to natural compound affinities being equal to the

corresponding metabolite affinities for the same protein, while the two dashed lines correspond to the natural compound affinities being 2 orders of magnitude higher or lower than the corresponding metabolite affinities for the same protein, respectively. The distributions of natural compound affinities (left) and metabolite affinities (top) are included, with the median negative logarithm values.

Supplementary Table 1:

136 natural compounds never approved as drugs. 'drug' column shows if a drug has been investigated to become a future drug (PD) or no (C).

https://drive.google.com/file/d/1wzvMfgxzd9Js_DSj3W8t2FqLBYtllraE

4. Drug-induced sudden cardiac arrest: An analysis of drug report databases

Andreu Bofill[1]. Xavier Jalencas[1] 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.

* Corresponding authors: jmestres@imim.cat

Abstract

Sudden cardiac arrest (SCA) represents a major health issue in our society and it is not expected to decrease in a short- and medium-term future. To better understand the risk factors and mechanisms involved in SCA susceptibility is of great relevance to alleviate this burden. In this work, an analysis of spontaneous drug safety reports identified 12 drugs likely associated with SCA. An exhaustive analysis of each one of them was done to find known evidences of their potential causal relation with SCA. We further explored the pharmacology of these SCA-related drugs in order to understand which are the most common mechanisms of action involved in the appearance of this event.

Introduction

Sudden cardiac arrest (SCA) is one of the main public health issues in Europe. The increase of the life expectancy during the last decades in all European countries has caused a considerable increment of SCA and it is expected to rise further. Nowadays, it is responsible for more than 50% of the cardiac deaths and approximately 20% of all natural deaths in Europe.[1] SCA is understood as the sudden cessation of the heart's ability to pump blood through the body and it is caused by acute myocardial infarction, cardiac arrhythmia and/or myocardial ischaemia, being the result of various interrelated causes.[2]

In the past decades, a large number of studies have explored the possible risk and causative factors of SCA.[3–6] Although great improvements in new effective treatments, identification and prevention of sudden cardiac arrest have resulted in a significant mortality reduction, numbers are still incredibly high, being the average survival rate in the European countries of only 10%. Risk factors associated with SCA have different origins, including genetic polymorphisms, ambiental or lifestyle factors (such as alcoholism, obesity, fibrosis or smoking), or drug use. In fact, in most of the cases it is a collection of coexisting risk factors that leads to the SCA outcome.

Therefore, there is still room for improvement regarding the prevention and recognition of risk factors associated with SCA appearance at an individual level. A considerable number of studies have already revealed the relationship between the administration of specific drugs and a consequent SCA event, and tried to clarify the mechanism of action leading to this outcome.[7–12] In fact, some drugs have been withdrawn from the market due to its association with sudden cardiac death. This is the case of sertindole, an atypical antipsychotic used in the treatment of schizophrenia that was withdrawn from the market in 1998 due to an increased risk of sudden cardiac death.[13] Although this clear relation between drug administration and SCA, there is a large need to further deepen and expand this knowledge. The detection of drugs and their mechanisms of action related with SCA would have immediate therapeutic implications, allowing to design better and individualized treatments and, along with the detection of risk factors, to prevent the administration of some of these drugs to vulnerable individuals susceptible to SCA.

In the present study, we aimed to obtain a consolidated and curated list of drugs linked to the SCA event. To do so, we analyzed postmarketing reports associated with SCA in order to extract significant signals related to its appearance. We analysed these drugs and explored their drug-target interactions in order to evaluate the possible mechanisms of action involved in the appearance and development of SCA.

Parallelly, with these consolidated and corroborated drugs, we studied which are the most commonly reported comorbidities alongside SCA.

Drug-induced sudden cardiac arrest

The FDA Adverse Event Reporting System (FAERS) is a public database that contains more than 15 million records of adverse drug cases reported during the last 50 years, between 1969 and 2020, mostly in the United States. Each report contains information on the list of drugs the patient was taking at the time, alongside with the disease and safety effects suffered, as well as some personal characteristics (i.e. age, weight and gender). Similarly, VigiBase is the WHO database of individual case safety reports that contains more than 20 million records of adverse events since 1968. Disproportionality analyses of this type of repositories allow for identifying drugs potentially linked to specific side effects and for further analyzing the contextual factors associated with drug-induced events like SCA.

The CLARITY platform[14], developed by Chemotargets, was used to perform the present study. Initially, a standardization process is applied to both adverse event terms and drug names to be able to navigate through the terms in an appropriate way. To standardize adverse events, it uses

MedDRA, which is a dictionary of medical terms for regulatory activities developed by the International council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) to promote the unification of regulatory terminologies. In consequence, adverse events are provided in a controlled vocabulary where entries are annotated with a unique medDRA code. In fact, several terms can be mapped to the same code, such as 'sudden cardiac arrest' and 'cardiac arrest', which are assigned to the same unique code and collapsed to the term 'cardiac arrest'. So, the term that was used to extract SCA reports is 'cardiac arrest'.

A total of 130,252 reports linked with the term 'cardiac arrest' were extracted. The average age of the patients reported was 53.38 years and the average weight 74.58 kg. There is almost no gender bias in the reports, being 52.0% and 48.0% the percentage of reports for men and women, respectively. A total of 2,502 drugs contained at least one report. In order to detect those reported drugs most related to SCA, we applied the following four statistical and disproportionality criteria. First of all, the frequency of drug-safety event reports over the total number of reports mentioning the drug needed to be greater than 1%. In that situation, the specific drug-safety association was considered **frequent**. Secondly, we computed the Proportional Reporting Ratio (PRR), which is the ratio between the frequency of reports of a specific drug that are related to a specific adverse

event and the proportion of reports of that advent reported alongside all drugs. We determined a threshold of at least 2 in order to consider the drug-safety association as **overreported**. Then, we analysed the number of reports where the drug is annotated as primary or secondary suspects for the appearance of the safety event. The drug-safety event association was considered **suspicious** and taken into account, if there are at least 10 reports considering the drug to be the primary or secondary suspect for the safety event. Finally, the drug-safety association needed to be **instantiated**, which means that there is at least one report with just this specific drug and the safety event. If a drug-safety event association is frequent, overreported, suspicious and instantiated for at least three consecutive years, then we considered that association as **consolidated**.

Applying these criteria to the whole set of drugs having at least one report linked to SCA left us with 16 drugs having SCA as a consolidated safety event. However, an exhaustive analysis of these 16 drugs revealed that four of them (alteplase, esmolol, nitric acid and reteplase) are in fact used in a cardiac arrest resuscitation context. Being SCA an actual indication of those drugs instead of an unwanted secondary effect, they were left out of the analysis. The final list of 12 drugs associated with SCA is shown in Table 1.

We used the Anatomical Therapeutic Chemical (ATC) classification system to stratify drugs and to analyze whether drugs related with SCA are most associated with some particular drug group. Under this classification, a drug may actually belong to more than one ATC group depending on the part of the organism they are intended to actuate and their therapeutic characteristics. A total of 6 drug groups of the ATC anatomical-level involving drugs related to SCA were found: (B) Blood and blood forming organs (d=1), (C) Cardiovascular system (d=4), (J) Antiinfectives for systemic use (d=1), (N) Nervous system (d=5), (S) Sensory organs (d=1) and (V) Various ATC structures (d=2). As can be observed, a predominant part of the SCA-related drugs are therapeutically relevant for the nervous (41%) and cardiovascular systems (33%). The specific ATC codes are provided in Table 1. An exhaustive analysis of these 12 drugs follows.

The neurological drugs that we found are part of three different therapeutic ATC levels. We found an antiepileptic (fosphenytoin), three different psycholeptics, including two hypnotics or sedatives (midazolam and dexmedetomidine) and one antipsychotic (droperidol), and finally a local anesthetic (bupivacaine). Fosphenytoin is an anticonvulsant drug used in the acute treatment of convulsive status epilepticus and to decrease seizure activity by increasing efflux of sodium ions across cell membranes in the motor cortex. Fosphenytoin is a phenytoin prodrug which can be efficiently administered

intravenously to deliver phenytoin. It is reported to produce several critical side effects, including a range of cardiac complications.[15] Some studies have directly related the administration of fosphenytoin to the induction of cardiac arrest.[16,17] Droperidol is another one of the psycholeptics found to be related to SCA. It is an antipsychotic drug with sedative and antiemetic properties. It is an antagonist of dopamine receptors in the central nervous system and several reports suggest a strong association between the administration of

DRUG	ATC CODE
adenosine	C01EB10, J05AB03, S01AD06
aprotinin	B02AB01
bupivacaine	N01BB01, N01BB10, N01BB51
dexmedetomidine	N05CM18
diatrizoic acid	V08AA01
droperidol	N05AD08
encainide	C01BC08
fosphenytoin	N03AB05
midazolam	N05CD08
regadenoson	C01EB2
sugammadex	V03AB35
verapamil	C08DA01,C08DA51

Table 1: The list of 12 SCD-consolidated drugs and their ATC classification code.

typical antipsychotic drugs, including droperidol, and an increased risk of sudden cardiac arrest.[7,18,19] Midazolam is a benzodiazepine with sedative and anxiolytic properties. It is generally used as palliative care in dying patients. Midazolam administration has been associated with cardiac arrest and respiratory depression, particularly in elderly patients.[20] In the same way, dexmedetomidine, an alpha-2-adrenoreceptor agonist with similar sedative, anxiolytic and analgesic properties than midazolam, has been also related with a significant elevated risk of SCA.[20–22] The last neurological drug that we found is bupivacaine. It is used as a local anesthetic in surgery operations. It blocks the generation of peripheral nerve impulses, preventing depolarization and minimizing the growth of the action potential. The causal relation between the bupivacaine administration and the cardiac arrest event is well and widely known. In fact, this drug is used to induce cardiac-arrest in several animal models.[23–25]

Four cardiovascular drugs were found to be strongly related to SCA: one is a selective calcium channel blocker with direct cardiac effects (verapamil) and the other three are cardiac drugs, one class Ic antiarrhythmic (encaine), and two other cardiac preparations (adenosine and regadenoson). Adenosine is an endogenous purine nucleoside, used to treat supraventricular tachycardia. In fact, it is among the most frequently used cardiac medications in the clinical practice of anesthesiology. It is a known potent vasodilator of coronary

and cerebral resistance vessels.[26] This drug is used to produce an adenosine-induced cardiac arrest in patients with prior heart surgery or patients undergoing thoracic aorta endovascular repair.[26,27] In fact, it was reported that the heartbeat stopped immediately for 18-58s just after a severe adenosine injection and the electroencephalogram power was reduced a 57%.[26] Regadenoson is an agonist of adenosine receptors that acts as a coronary vasodilator. It is a selective adenosine-2A receptor agonist and it is used to induce pharmacological stress in myocardial perfusion imaging procedures. It seems to produce less side-effects than the adenosine due to its higher selectivity.[28] However the FDA released a safety announcement a few years ago advising of the serious risk of heart attack and death associated with regadenoson.[29] Encainide was a class Ic antiarrhythmic agent. It has been withdrawn from the market after it was associated with an excess mortality due to arrhythmia and shock after acute recurrent myocardial infarction.[30] Several studies reported the relation between administration of encainide and the increased incidence of cardiac arrest.[31,32] Finally, verapamil is a selective calcium channel blocker that is used in order to reduce the blood pressure and to control the supraventricular tachycardia. Several reports associated the administration and the overdose of verapamil with the cardiac arrest appearance.[25,33,34]

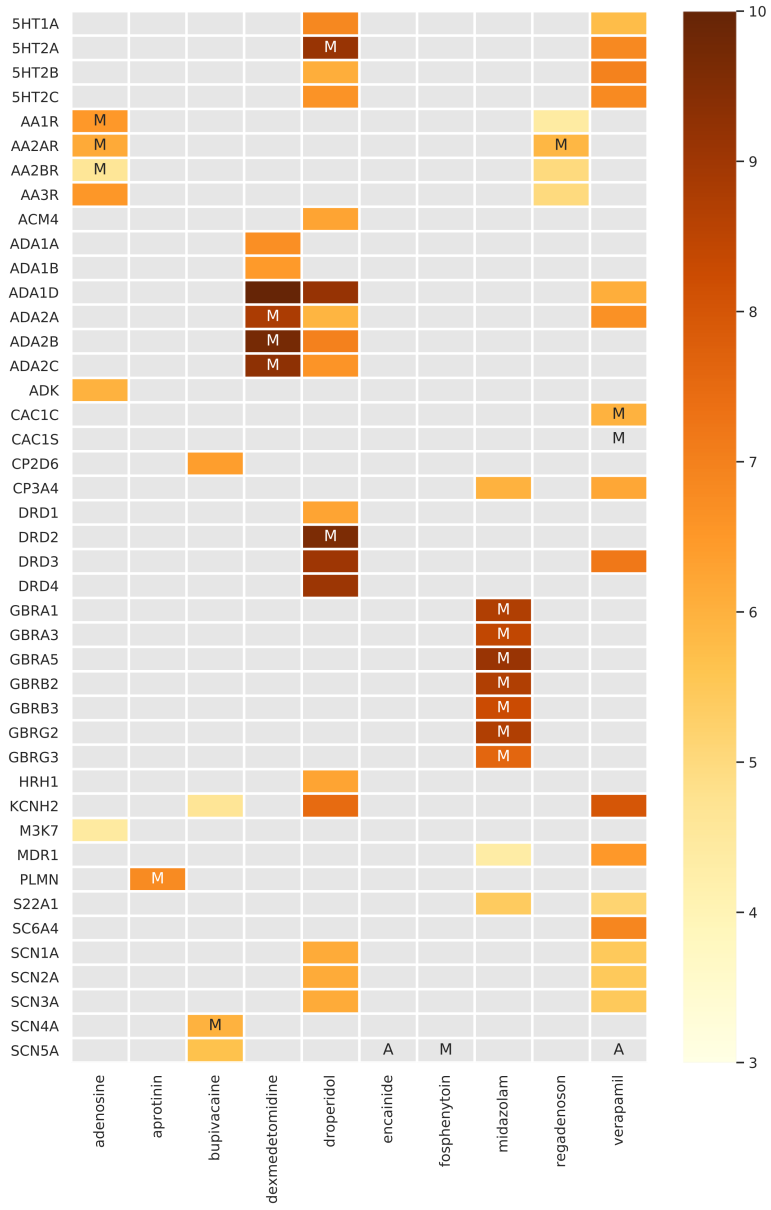


Figure 1: Interaction heatmap for the 10 drugs linked to cardiac arrest. Only targets with two known interactions or with a single strong interaction (> 1μM) are included. The complete list can be accessed at Supplementary Table S1. Each interaction is colored according to its quantitative affinity. Interactions with no quantitative affinity are labeled as “A” (active), while interactions involved in the mechanism of action of the drug are labeled as “M”. Both drugs and targets are sorted alphabetically.

Another drug found in this study strongly related with SCA is aprotinin. It is a proteinase inhibitor used as an antihemorrhagic agent. It is widely used in cardiac surgery in order to reduce the bleeding in operatory, post-operatory and general hospital stay. Several studies associated the administration of aprotinin with an increased risk of several outcome events such as myocardial infarction, renal failure or heart failure, converting this drug into a potential cause of cardiac arrest.[35,36]

Finally, the last two SCA-related drugs that we present in this work are sugammadex and diatrizoic acid. Sugammadex is used in general anaesthesia as the first selective relaxant binding agent. It is a modified gamma-cyclodextrin widely used as an antidote to reverse neuromuscular blockade caused by the administration of some neuromuscular-blocking agents such as rocuronium or vecuronium.[37] Some case reports have been published relating the appearance of a cardiac arrest after the sugammadex administration in healthy patients.[38,39] Finally, diatrizoic acid is a contrast agent used in medical imaging, specially as an x-ray contrast. Although some studies pointed that the administration of diatrizoic acid could be associated to pronounced effects in cardiac function and to an increased risk of some cardiac complications such as prolonged QT interval, no significant evidences has been reported relating the use of diatrizoic acid with the appearance of sudden cardiac

arrest.[40,41] In consequence, we suggest an accurate reexamination of this drug.

The meticulous review of these drugs confirmed their relation to SCA, thus supporting the methodology to retrieve meaningful signals from drug-reporting databases.

Characterisation of the main mechanisms of action linked to drugs associated with SCA

In order to further evaluate the drug-induced sudden cardiac arrest appearance and be able to explore its mechanisms of action, we extracted all the interactions of SCA-consolidated drugs obtained in the previous section with human targets. So, the list of 12 drugs was used to interrogate DrugCentral[42], Chembl[43] and Guide To Pharmacology (GtoPdb)[44] databases to extract their known interactions with human targets. From the 12 aforementioned drugs, we were able to find at least one target for 10 of them. A total of 107 interactions involving 10 drugs and 77 human targets were retrieved (Figure 1 and Table S1 from supplementary material). We did not find any interaction for two of the drugs (diatrizoic acid and sugammadex) in any of the databases used to extract the activity information. This is not surprising as neither of them produce their therapeutic effect by targeting a human protein. Sugammadex is used for the reversal of the neuromuscular

blockade induced by rocuronium and vecuronium (two usual anesthetic drugs) by forming complexes with them and reducing their concentration in the neuromuscular junction. Diatrizoic acid is a common contrast medium used in x-ray radiology.

In order to discriminate between primary and secondary targets, we retrieved from DrugCentral[42] all drug-target interactions labeled as being involved in the mechanism of action (MoA) of the drug. Consequently, all other interactions not labeled as MoA were classified as secondary targets. A total of 21 of the 107 found interactions were implicated in the MoA of the drug, involving 9 drugs and 20 targets. Although this valuable information truly helped in understanding our drug set, we did not discriminate between primary and secondary targets when assessing the possible implication of these targets on SCA appearance.

We then ranked the obtained targets according to their number of drug interactions (Table S2 in supplementary material). The most representative target is the sodium voltage-gated channel protein type 5 subunit alpha ($Na_v1.5$), which is associated with four drugs. This protein, encoded by the gene SCN5A, is a transmembrane protein found mainly in cardiac muscle and is responsible for the inbound sodium flux, inducing a depolarization and the consequent action potential.[45,46] Several mutations of that gene have been

identified and associated with critical implications in the function of the channel, leading to a possible reduction of the sodium current.[46,47] Several of these mutations are largely associated with the Brugada Syndrome, Lev-Lenègre disease, long QT interval, atrial fibrillation or the sick sinus syndrome.[46,48–50] Moreover, rare variant mutations of this gene, which result in significantly shorter recovery times from inactivation, are related with a significant increase of sudden cardiac death cases.[51] The reduction of sodium levels could also be induced by the action of a drug. Some studies found that the blockade of the cardiac sodium channel by non-cardiac blocking drugs (i.e. bupivacaine) may be directly related to an increase of SCA risk, especially if some genetic alterations such as a mutation in the SCN5A gene are already present.[52,53]

Another channel was also found in the top most reported SCA-related proteins, the potassium voltage-gated channel subfamily H member 2 (KCNH2), encoded by the well-known gene hERG. This protein is part of the ion channel that conducts potassium out of the cardiac muscle and plays an important role in the repolarization process and the maintenance of normal cardiac rhythm.[54] In the same way as SCN5A, both hERG channel mutations or drug-blockade of hERG can be associated with several cardiac diseases such as an excessive prolongation of the QT interval, Torsades de Pointes, ventricular arrhythmias or sudden cardiac death.[55–58] Moreover, the administration of some drug groups, like

antipsychotics such as droperidol and verapamil has been associated with a significant increase of SCD risk by interacting with SCN5A.[12,18,59] These results clearly show that the blockade of both the sodium channel protein type 5 subunit alpha and the hERG channel are relevant contributors to drug-induced SCA.

Among the remaining targets we also found the four subtypes of adenosine receptors (AA1R, A2AR, AA2BR and AA3R). Adenosine receptors are G protein-coupled receptors with an important role in cardiac regulation. It has been reported that adenosine receptors are a key player in the development of heart failure and other cardiac complications, deriving in a sudden cardiac death.[60] More specifically, in a 2011 study on the implications of the adenosine receptor in the phenotype of heart failure, the authors proposed that a balance between A1/A3 and A2A receptors is needed in order to maintain a normal cardiac homeostasis, suggesting that an imbalance of these receptors could lead to possible cardiac complications.[60]

The list of targets also includes several alpha adrenergic receptors (ADA2A, ADA2B, ADA2C and ADA1D). Alpha adrenergic receptors are G-protein coupled receptors (GPCR) found in several parts of both peripheral and central neural systems. Although general functions have been described for the whole family of receptors, such as vasoconstriction or

smooth muscle contraction, each receptor has its particular functions. For example, alpha-1D adrenergic receptors are found mainly in epicardial coronary arteries mediating the coronary vasoconstriction.[61] On the contrary, alpha-2 adrenergic receptors, among other functions, participate in the inhibition norepinephrine release.[62] Several studies have proposed that the blockade of alpha-1 adrenergic receptors is related to an increased risk of cardiac complications such as heart failure or cardiac death in patients with hypertension, suggesting that an activation of alpha 1 adrenergic receptors could produce a cardioprotective effect.[63–65] However, we did not find evidences of the implication of alpha-2 receptors in cardiac arrest events. Our results suggest that the main mechanism of action implicated in SCA does not involve the alpha-2 adrenergic receptors but other of the already discussed targets linked to SCA.

Almost all the drugs described in this paper have a known interaction with some of the mentioned targets. However, aprotinin does not have a known activity value for any of them. In fact, its only known activity is for the plasminogen receptor, suggesting there might be a lack of information in the consulted databases that does not allow a correct detection of all drug-target interactions implicated with this drug reported side effect.

Identification of risk factors for drug-induced sudden cardiac arrest

We also analyzed in detail which are the safety-related events that are most linked with drug-induced SCA appearance. We based our study on the list of 12 drugs that have SCA as a consolidated safety event (positive set). In order to be able to compare this list of drugs with other non SCA-related drugs, we used CLARITY again to extract a list of drugs with no reports involving SCA. The top 45 most reported drugs with no reported SCA events were used as a negative set. These 45 drugs are reported associated with more than 3,900 safety events, none of them being SCA. A complete list of these 45 drugs can be found in the Supplementary Table S3.

For each drug in the positive and negative sets, we computed the frequencies of reports for which each safety event is reported alongside the drug. In the positive set, a total of 1,479 different safety events have been reported at least one time together with cardiac arrest and one of the 12 SCA-related drugs. A comparison analysis of the average frequencies of each safety event between the positive and the negative sets, is shown in Figure 2. The events most reported alongside SCA-consolidate drugs are plotted in red, while the events more frequently reported for drugs of the negative set are shown in yellow.

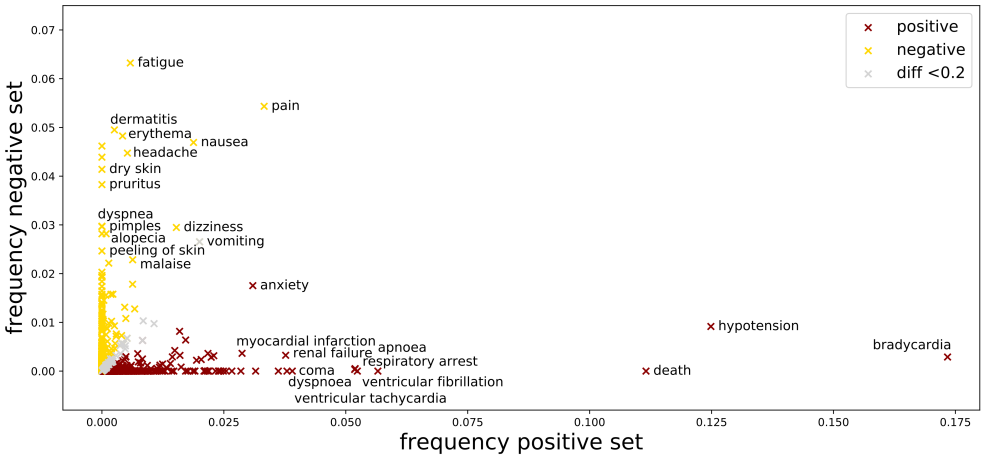


Fig. 2: Comparison between safety event frequencies reported alongside SCA-related drugs and these events frequencies reported with non-related SCA drugs. The most reported events with SCA-consolidated drugs are shown in red, while the ones most reported with non-related-SCA drugs are shown in yellow. Only the events with a higher frequency are labelled.

As can be observed, the safety events most frequently associated with SCA are cardiotoxicity, hemotoxicity and pneumotoxicity endpoints. A specially strong signal is shown for bradycardia, hypotension and death. Both bradycardia and hypotension are two well-known comorbidities of cardiac arrest.[66–68] Death is obviously a possible outcome of a sudden cardiac arrest. Although we found that death is present in just 11% of the reports alongside cardiac arrest and one of the SCA-related drugs, the real frequency is likely much larger, as lots of reporters do not register death as a comorbidity but as an outcome. However, some terms unrelated to cardiotoxicity or pneumotoxicity such as seizure, coma, loss of

consciousness or renal failure are also more frequently related with SCA consolidated drugs than with drugs in the negative set. Considering that some of those drugs are non-cardiac agents, as detailed in the previous section, the appearance of safety terms not directly associated with cardiotoxicity is expected. On the other hand, events such as fatigue, pain or nausea are clearly more common in the negative set of drugs and seem to hold no relationship with cardiac arrest. Another group of dermatologic events like dermatitis, erythema, dry skin and pruritus also appear to be unrelated to drugs with risk of SCA. For two events with a similar frequency of reports together with SCA such as pain and myocardial infarction (around 3 and 4% in both cases), the latter is a more specific alert for SCA as its frequency in the set of negative drugs is much lower. In summary, events with a high frequency difference between the positive and the negative sets can be used as an alert signal for SCA, as drugs reported for those events will likely be associated also with this outcome.

Conclusions

Sudden cardiac arrest represents a major health problem and one of the principal causes of death worldwide. To improve the knowledge of the relation between the administration of some drugs and an increased risk of sudden cardiac arrest entails an important step forward into a better prevention of this event. In

this work we used CLARITY to comprehensively analyse the relation between drugs and SCA in spontaneous safety reports. In this regard, 12 drugs were found to be disproportionally related to the SCA event. To confirm these findings, we exhaustively searched on previous literature the possible known relation of each one of these drugs with SCA. In 11 of these drugs we find evidence supporting the causal relation between the drug administration and the increased risk of sudden cardiac arrest. However, in the case of diatrizoic acid, although some other cardiac complications have been related to its administration, no significant evidence has been reported regarding its association with the appearance of sudden cardiac arrest. Accordingly, we suggest a reevaluation of the safety margins of this drug.

We also analysed the polypharmacology of these drugs to determine mechanisms of action leading to SCA. Some of the found targets, such as the sodium voltage-gated channel protein type 5 subunit alpha (SCN5A) or the potassium voltage-gated channel subfamily H member 2 (hERG) were associated with an increased risk of sudden cardiac arrest, which is well-established in the literature.

In summary, the study of postmarketing spontaneous drug safety reports is a helpful tool to evaluate and control drug safety. The criteria used in this paper has been demonstrated to be a suitable method to detect consolidated associations

between safety events and drug administration. The extension of this analysis to other side effects and drugs would allow for the identification of drug risks that might not be already common knowledge and could improve drug prescription safety.

References

1. Gillum, R. F. Geographic variation in sudden coronary death. *Am. Heart J.* **119**, 380–389 (1990).
2. Tan, H. L., Dagres, N., Böttiger, B. W., Schwartz, P. J. & ESCAPE-NET Investigators. European Sudden Cardiac Arrest network: towards Prevention, Education and New Effective Treatments (ESCAPE-NET). *European Heart Journal* vol. **39** 86–88 (2018).
3. Kaltman, J. R. *et al.* Screening for Sudden Cardiac Death in the Young. *Circulation* vol. 123 1911–1918 (2011).
4. Chugh, S. S. Sudden cardiac death in 2017: Spotlight on prediction and prevention. *International Journal of Cardiology* vol. **237** 2–5 (2017).
5. Fishman, G. I. *et al.* Sudden cardiac death prediction and prevention: report from a National Heart, Lung, and Blood Institute and Heart Rhythm Society Workshop. *Circulation* **122**, 2335–2348 (2010).
6. Risgaard, B. *et al.* Sudden Cardiac Death: Pharmacotherapy and Proarrhythmic Drugs: A Nationwide Cohort Study in Denmark. *JACC Clin Electrophysiol* **3**, 473–481 (2017).
7. Sicouri, S. & Antzelevitch, C. Sudden cardiac death secondary to antidepressant and antipsychotic drugs. *Expert Opin. Drug Saf.* **7**, 181–194 (2008).
8. Morentin, B., Callado, L. F., García-Hernández, S., Bodegas, A. & Lucena, J. The role of toxic substances in

sudden cardiac death. *Spanish Journal of Legal Medicine* vol. 44 13–21 (2018).

9. Haverkamp, W. & Breithardt, G. Drug-induced sudden cardiac death. *European heart journal* vol. **26** 1808–9; author reply 1809 (2005).
10. Daher, M. *et al.* Sudden cardiac arrest in general anesthesia as the first manifestation of anomalous origin of the left coronary artery. *Rev. Bras. Anestesiol.* **62**, 878–884 (2012).
11. Tabib, A. *et al.* [Sudden death during anesthesia: human error, drug related or cardiac death?]. *Therapie* **56**, 735–738 (2001).
12. Ray, W. A. *et al.* Antipsychotics and the risk of sudden cardiac death. *Arch. Gen. Psychiatry* **58**, 1161–1167 (2001).
13. Lewis, R., Bagnall, A.-M. & Leitner, M. Sertindole for schizophrenia. *Cochrane Database of Systematic Reviews* (2005) doi:10.1002/14651858.cd001715.pub2.
14. Garcia-Serna, R., Vidal, D., Remez, N. & Mestres, J. Large-Scale Predictive Drug Safety: From Structural Alerts to Biological Mechanisms. *Chem. Res. Toxicol.* **28**, 1875–1887 (2015).
15. Boucher, B. A. Fosphenytoin: a novel phenytoin prodrug. *Pharmacotherapy* **16**, 777–791 (1996).
16. DeToledo, J. C., Lowe, M. R., Rabinstein, A. & Villaviza, N. Cardiac arrest after fast intravenous infusion of phenytoin mistaken for fosphenytoin. *Epilepsia* **42**, 288 (2001).

17. Yumoto, T., Naito, H. & Nakao, A. Cardiac arrest after intravenous infusion of fosphenytoin. *International Journal of Case Reports and Images* vol. **9** 1 (2018).
18. Zhu, J. *et al.* Antipsychotic drugs and sudden cardiac death: A literature review of the challenges in the prediction, management, and future steps. *Psychiatry Res.* **281**, 112598 (2019).
19. Sicouri, S. & Antzelevitch, C. Mechanisms Underlying the Actions of Antidepressant and Antipsychotic Drugs That Cause Sudden Cardiac Arrest. *Arrhythm Electrophysiol Rev* **7**, 199–209 (2018).
20. Nordt, S. P. & Clark, R. F. Midazolam: A review of therapeutic uses and toxicity. *The Journal of Emergency Medicine* vol. **15** 357–365 (1997).
21. Ingersoll-Weng, E., Manecke, G. R. & Thistlethwaite, P. A. Dexmedetomidine and Cardiac Arrest. *Anesthesiology* vol. **100** 738–739 (2004).
22. Bharati, S., Pal, A., Biswas, C. & Biswas, R. Incidence of cardiac arrest increases with the indiscriminate use of dexmedetomidine: a case series and review of published case reports. *Acta Anaesthesiol. Taiwan* **49**, 165–167 (2011).
23. Hiller, D. B. *et al.* Epinephrine impairs lipid resuscitation from bupivacaine overdose: a threshold effect. *Anesthesiology* **111**, 498–505 (2009).
24. Mauch, J., Jurado, O. M., Spielmann, N., Bettschart-Wolfensberger, R. & Weiss, M. Resuscitation strategies

from bupivacaine-induced cardiac arrest. *Paediatr. Anaesth.* **22**, 124–129 (2012).

25. Vijay, B. S., Mitra, S. & Jamil, S. N. Refractory cardiac arrest due to inadvertent intravenous injection of 0.25% bupivacaine used for local infiltration anesthesia. *Anesth Essays Res* **7**, 130–132 (2013).
26. Plaschke, K., Böckler, D., Schumacher, H., Martin, E. & Bardenheuer, H. J. Adenosine-induced cardiac arrest and EEG changes in patients with thoracic aorta endovascular repair. *Br. J. Anaesth.* **96**, 310–316 (2006).
27. El-Menyar, A. & Gehani, A. Adenosine-induced tachyarrhythmia and cardiac arrest. *Future Cardiol.* **6**, 433–436 (2010).
28. Jager, P. L. *et al.* Regadenoson as a new stress agent in myocardial perfusion imaging. Initial experience in The Netherlands. *Revista Española de Medicina Nuclear e Imagen Molecular (English Edition)* vol. **33** 346–351 (2014).
29. Rosenblatt, J., Mooney, D., Dunn, T. & Cohen, M. Asystole following regadenoson infusion in stable outpatients. *J. Nucl. Cardiol.* **21**, 862–868 (2014).
30. Echt, D. S. *et al.* Mortality and Morbidity in Patients Receiving Encainide, Flecainide, or Placebo. *New England Journal of Medicine* vol. **324** 781–788 (1991).
31. Fish, F. A., Gillette, P. C. & Benson, D. W., Jr. Proarrhythmia, cardiac arrest and death in young patients receiving encainide and flecainide. The Pediatric

Electrophysiology Group. *J. Am. Coll. Cardiol.* **18**, 356–365 (1991).

32. Akiyama, T. *et al.* Increased risk of death and cardiac arrest from encainide and flecainide in patients after non-Q-wave acute myocardial infarction in the Cardiac Arrhythmia Suppression Trial. *The American Journal of Cardiology* vol. **68** 1551–1555 (1991).
33. McGOVERN, B. Precipitation of Cardiac Arrest by Verapamil in Patients with Wolff-ParkinsonWhite Syndrome. *Annals of Internal Medicine* vol. **104** 791 (1986).
34. Déniel, A., Fedrizzi, S., Lelong-Boulouard, V., Coquerel, A. & Alexandre, J. Fatal Cardiac Arrest Associated with Concomitant Bisoprolol and Verapamil Overdose. *J. Am. Geriatr. Soc.* **64**, 451–452 (2016).
35. Mangano, D. T., Tudor, I. C. & Dietzel, C. The Risk Associated with Aprotinin in Cardiac Surgery. *New England Journal of Medicine* vol. **354** 353–365 (2006).
36. McMullan, V. & Alston, R. P. III. Aprotinin and cardiac surgery: a sorry tale of evidence misused. *Br. J. Anaesth.* **110**, 675–678 (2013).
37. Bom, A., Hope, F., Rutherford, S. & Thomson, K. Preclinical pharmacology of sugammadex. *J. Crit. Care* **24**, 29–35 (2009).
38. Sanoja, I. A. & Toth, K. S. Profound Bradycardia and Cardiac Arrest After Sugammadex Administration in a

Previously Healthy Patient. *A & A Practice* vol. 12 22–24 (2019).

39. Ko, M. J. *et al.* Cardiac arrest after sugammadex administration in a patient with variant angina: a case report. *Korean Journal of Anesthesiology* vol. 69 514 (2016).
40. Mitsumori, M. *et al.* Effects of contrast media on the RR and QT interval during coronary arteriography. *Acta radiol.* **32**, 120–123 (1991).
41. Dunkel, J. A., Bøkenes, J., Karlsson, J. O. G. & Refsum, H. Cardiac Effects of Iodixanol Compared to Those of other Nonionic and Ionic Contrast Media in the Isolated Rat Heart. *Acta Radiologica* vol. 36 142–154 (1995).
42. Ursu, O. *et al.* DrugCentral: online drug compendium. *Nucleic Acids Res.* **45**, D932–D939 (2017).
43. Gaulton, A. *et al.* The ChEMBL database in 2017. *Nucleic Acids Res.* **45**, D945–D954 (2017).
44. Sharman, J. L. *et al.* Accessing Expert-Curated Pharmacological Data in the IUPHAR/BPS Guide to PHARMACOLOGY. *Curr. Protoc. Bioinformatics* **61**, 1.34.1–1.34.46 (2018).
45. Weidmann, S. The effect of the cardiac membrane potential on the rapid availability of the sodium-carrying system. *J. Physiol.* **127**, 213–224 (1955).
46. Veerman, C. C., Wilde, A. A. M. & Lodder, E. M. The cardiac sodium channel gene SCN5A and its gene product

NaV1.5: Role in physiology and pathophysiology. *Gene* **573**, 177–187 (2015).

47. Osman, J., Tan, S. C., Lee, P. Y., Low, T. Y. & Jamal, R. Sudden Cardiac Death (SCD) - risk stratification and prediction with molecular biomarkers. *J. Biomed. Sci.* **26**, 39 (2019).
48. Remme, C. A. *et al.* Overlap syndrome of cardiac sodium channel disease in mice carrying the equivalent mutation of human SCN5A-1795insD. *Circulation* **114**, 2584–2594 (2006).
49. Tan, H. L. *et al.* A sodium-channel mutation causes isolated cardiac conduction disease. *Nature* vol. **409** 1043–1047 (2001).
50. Wang, Q. *et al.* SCN5A mutations associated with an inherited cardiac arrhythmia, long QT syndrome. *Cell* vol. **80** 805–811 (1995).
51. Albert, C. M. *et al.* Cardiac sodium channel gene variants and sudden cardiac death in women. *Circulation* **117**, 16–23 (2008).
52. Weeke, P. *et al.* Antidepressant use and risk of out-of-hospital cardiac arrest: a nationwide case-time-control study. *Clin. Pharmacol. Ther.* **92**, 72–79 (2012).
53. Bardai, A. *et al.* Sudden cardiac arrest associated with use of a non-cardiac drug that reduces cardiac excitability: evidence from bench, bedside, and community. *Eur. Heart J.* **34**, 1506–1516 (2013).

54. Perrin, M. J., Subbiah, R. N., Vandenberg, J. I. & Hill, A. P. Human ether-a-go-go related gene (hERG) K channels: Function and dysfunction. *Progress in Biophysics and Molecular Biology* vol. 98 137–148 (2008).
55. van Noord, C., Sturkenboom, M. C. J. M., Straus, S. M. J. M., Witteman, J. C. M. & Stricker, B. H. C. Non-cardiovascular drugs that inhibit hERG-encoded potassium channels and risk of sudden cardiac death. *Heart* **97**, 215–220 (2011).
56. Brown, A. M. HERG block, QT liability and sudden cardiac death. *Novartis Found. Symp.* **266**, 118–31; discussion 131–5, 155–8 (2005).
57. Calderone, V., Testai, L., Martinotti, E., Del Tacca, M. & Breschi, M. C. Drug-induced block of cardiac HERG potassium channels and development of torsade de pointes arrhythmias: the case of antipsychotics. *J. Pharm. Pharmacol.* **57**, 151–161 (2005).
58. Wang, S. *et al.* Adrenergic regulation of the rapid component of delayed rectifier K+ currents in guinea pig cardiomyocytes. *J. Thorac. Dis.* **6**, 1778–1784 (2014).
59. Wu, C., Tsai, Y. & Tsai, H. Antipsychotic Drugs and the Risk of Ventricular Arrhythmia and/or Sudden Cardiac Death: A Nation-wide Case-Crossover Study. *Journal of the American Heart Association* vol. 4 (2015).
60. Feldman, A. M., Cheksis-Feiner, E., Hamad, E. & Chan, T. Adenosine receptor subtypes and the heart failure

phenotype: translating lessons from mice to man. *Trans. Am. Clin. Climatol. Assoc.* **122**, 198–214 (2011).

61. Jensen, B. C. *et al.* The alpha-1D Is the predominant alpha-1-adrenergic receptor subtype in human epicardial coronary arteries. *J. Am. Coll. Cardiol.* **54**, 1137–1145 (2009).
62. Giovannitti, J. A., Jr, Thoms, S. M. & Crawford, J. J. Alpha-2 adrenergic receptor agonists: a review of current clinical applications. *Anesth. Prog.* **62**, 31–39 (2015).
63. Jensen, B. C., O'Connell, T. D. & Simpson, P. C. Alpha-1-adrenergic receptors in heart failure: the adaptive arm of the cardiac response to chronic catecholamine stimulation. *J. Cardiovasc. Pharmacol.* **63**, 291–301 (2014).
64. O'Connell, T. D., Jensen, B. C., Baker, A. J. & Simpson, P. C. Cardiac alpha1-adrenergic receptors: novel aspects of expression, signaling mechanisms, physiologic function, and clinical importance. *Pharmacol. Rev.* **66**, 308–333 (2014).
65. Vanoli, E., Hull, S. S., Jr, Foreman, R. D., Ferrari, A. & Schwartz, P. J. Alpha 1-adrenergic blockade and sudden cardiac death. *J. Cardiovasc. Electrophysiol.* **5**, 76–89 (1994).
66. Yong, J., Hibbert, P., Runciman, W. B. & Coventry, B. J. Bradycardia as an early warning sign for cardiac arrest during routine laparoscopic surgery. *Int. J. Qual. Health Care* **27**, 473–478 (2015).

67. Kilgannon, J. H. *et al.* Early arterial hypotension is common in the post-cardiac arrest syndrome and associated with increased in-hospital mortality. *Resuscitation* **79**, 410–416 (2008).
68. Balasanmugam, C., Henriquez Felipe, C., Rodriguez, D. & Kulbak, G. Bradycardia, Hypotension, and Cardiac Arrest: A Complication of Local Anesthetics. *Cureus* **11**, e4033 (2019).

Supplementary Materials

Supplementary Table 1: List of 107 interactions involving 10 SCA-related drugs and 77 human proteins.
https://drive.google.com/file/d/1tFG3GuMpIFjQtkm5bh8sRhzUh5u4Ba79

Supplementary Table 2: List of 77 human targets sorted by number of SCA-related drugs:
https://drive.google.com/file/d/1zTmN4Yl06LcEnxqsB8YLzSAwlOplXis3

Supplementary Table 3: List of top 45 most-reported drugs without any SCA-report.
https://drive.google.com/file/d/1KHL7dDsJNV4wfuBXBA7Wrg5NAoJjq3N9

Discussion

The human endogenous metabolome as a baseline for primary pharmacology of drugs

The pharmaceutical industry puts a lot of effort into evaluating and optimising the bioactivity and pharmacology of drug candidates during the drug discovery process. A general strategy used in this process is to maximise the *in vitro* binding affinity for the targets involved in the mechanism of action of drug candidates to improve their target selectivity and, consequently, to minimize the drug secondary pharmacology that could result in unwanted secondary effects. However, different drugs have different levels of affinity with their mechanism-of-action targets and it is not known exactly why certain levels of affinity are needed by drugs to interact with their primary protein targets and exert a therapeutic action. This Thesis demonstrates that the minimum level of affinity needed by drugs to have a biologically-relevant interaction with a specific protein is related with the affinity of the endogenous metabolite for the same protein. Therefore, we propose that the metabolite affinity minus two orders of magnitude could be a useful affinity baseline for the primary pharmacology of drugs.

Each endogenous metabolite has been biologically selected by evolution in order to bind appropriately to its native protein with a specific level of affinity.[159] The results obtained in this Thesis suggest that this optimised endogenous

metabolite affinity for its native protein needs to be taken into account in order to explain the primary pharmacological activity of drugs. Our studies show that 96% of the drug-target primary interactions have affinity values above the principal endogenous metabolite affinity for the same target minus two orders of magnitude. Only 4% of these drug-target primary interactions are found below this reference baseline. Several cases are presented where interactions between drugs and targets with affinities below two orders of magnitude of metabolite affinities should not be assigned as primary target interactions. For example, there are several drugs that do not produce their effects through binding with targets currently assigned as their mechanism of action but their primary target is most likely yet to be uncovered. Another example of this wrong assignment found during this research was that several therapeutic effects related to the specific drug-target interactions are, indeed, produced by the action of a particular metabolite of the drug that interacts, usually, with the same target but with a significantly higher affinity. After an exhaustive analysis of these cases, we suggest that all relationships between a drug and its alleged mechanism of action targets with an affinity lower than the principal endogenous metabolite affinity for the target minus two orders of magnitude should be critically revised.

In parallel, we observed that the minimum drug affinity among all drugs optimised for a specific primary target is

directly proportional to the native endogenous ligand affinity for the same target. Conversely, the range of drug affinities for a given primary target is inversely proportional to the endogenous metabolite affinity. This clearly indicates that the endogenous metabolome contributes to determine which is the minimum level of affinity needed by drugs to exert its effect on their primary targets.

Impact of the human endogenous metabolome into secondary pharmacology of drugs

This thesis also highlights the implications that human endogenous metabolome could have on the secondary pharmacology of drugs. In the same way as for the drug primary pharmacology, we also propose that endogenous metabolite affinities should be used to establish safety margins for determining the risk of a drug to produce a specific side effect.

In order to exemplify this, we analysed the well-known relationship between valvular heart disease (VHD) and a strong agonist activity on the 5-hydroxytryptamine receptor 2B (5HT2B). Based on the study of 24 5HT2B agonist agents, we demonstrated that the minimum affinity needed by drugs to produce VHD is determined by the affinity of serotonin, the principal endogenous metabolite of 5HT2B receptor. In fact, it

could be established that a difference of two orders of magnitude below the *in vitro* affinity of serotonin for 5HT2B successfully discriminates the valvulopathic drugs from the ones devoid of risk to produce VHD. The evidence presented in this Thesis is in line with the study conducted by Papoian *et al.* in which an analysis of the affinity of 9 5HT2B agonists was conducted in order to evaluate the minimum level of affinity needed to produce VHD.[114] They suggested a 100-fold difference in affinity values between the endogenous ligand and drugs in order to explain the presence or absence of a valvulopathic effect in 8 5HT2B agonist drugs.[114]

In the same line, this thesis proposes a new method to assess off-target safety margins. A usual strategy to reduce off-target unwanted effects involve increasing the potency of the drug-target primary interaction to decrease the chances of secondary interactions and, consequently, reduce the adverse reactions associated with drug polypharmacology.[160] These safety margins are defined by the difference in binding affinities of the drug between the target and the off-target. However, considering the crucial implication of the human endogenous metabolome on secondary pharmacology of drugs we suggest that the off-target safety margins need to be redefined in order to incorporate this information. We propose that safety margin of drugs should also consider the difference in off-target binding affinities between the endogenous metabolite and the drug. In this regard, this Thesis suggests that a 100-fold

difference between the human endogenous metabolite affinities and the corresponding drug affinities for the same protein could be used as a safety margin criterion in order to discriminate the biologically-relevant from the biologically non-relevant drug-target interactions. Important implications for translational safety can be also inferred. Some experimental testing of endogenous metabolites and drugs is underway to provide additional confirmatory cases of these findings.

According to this definition, in the Chapter 2 of the results section, a new approach to estimate a more realistic drug polypharmacology landscape is introduced. Polypharmacology describes the ability of drugs to interact with multiple targets.[32] Nowadays, the drug-target interaction network of almost all current drugs is quite extensive and it is almost impossible to conceive a drug that only interacts with its therapeutic target. However, not all drug-target interactions are strong enough to produce a significant biological effect in the organism. Therefore, taking into consideration the reference baseline of endogenous metabolites interactions, we revised the polypharmacology of more than 400 drugs. This reassessment resulted in a significant reduction of the length of the pharmacological profile of drugs. This reduction of the drug polypharmacology could have several implications. First of all, it could allow us to better understand the real behaviour of drugs in biological systems. Consequently, it could also help to identify the main liabilities involved in the secondary

pharmacology of drugs and to detect specific pathways involved in these risk factors. The ultimate implication of the impact of the endogenous metabolome in drug polypharmacology is the implementation of a more personalized medicine, where individualized treatments based on personal levels of endogenous metabolites are taken into account to minimize risk factors associated with drugs.

Pharmacological view of natural compounds. The implications of human endogenous metabolome

This Thesis also compares the pharmacological activity of drugs and natural compounds and examines the impact of the endogenous metabolome on both sets. These results are presented in Chapter 3, where we show that the majority of natural compounds have affinity values below the metabolite affinity minus two orders of magnitude. As these exogenous compounds are not biologically optimized to interact with human targets, it seems plausible that their affinities are lower than the activity value of principal endogenous metabolites of these human targets, which are evolutionarily optimized to produce the required actions. However, we found that some of these natural compounds have affinities above the metabolite affinity baseline indicating that, as seen with drugs, it is plausible that they exert an effect in human organisms. These drug-like natural compounds could be toxic substances or new

drug candidates with both known or unknown therapeutic effects. We observed how the interactions between drug-like natural compounds with known therapeutic effects and their specific mechanism of action targets, need to present a minimum level of affinity given by the optimised affinity of the endogenous ligand for those targets.

Drugs and non-drug-like natural compounds have clear opposite behaviours. As previously mentioned, drugs have been optimized to produce at least a therapeutic effect interacting with their primary targets and, as we demonstrated in the first Chapter, being its affinity at least higher than the metabolite affinity baseline. On the contrary, non-drug-like natural compounds are not optimised, either biologically or by human intervention, to interact with any human target. For that reason, we found that these natural compounds follow an almost opposite trend than human drugs, being the majority of these affinity values below the metabolite affinity baseline.

Analysis of postmarketing adverse event reports. A review of drugs with a risk of inducing sudden cardiac arrest

Finally, an exhaustive study of postmarketing spontaneous adverse event reports was done and an accurate list of drugs significantly related to sudden cardiac arrest (SCA) is provided. This safety event is one of the major public health issues in

Europe causing, today, approximately 20% of all natural deaths.[161] After an intense curation of data, we found 12 drugs significantly consolidated to produce SCA and a comprehensive bibliographic analysis of these 12 SCA-related drugs was done, substantiating a possible causal link of these drugs with the appearance of sudden cardiac arrest.

Moreover, a study of the most relevant mechanisms of action of these 12 SCA-related drugs has also been performed in order to assess the biological pathways most related to sudden cardiac arrest. We concluded that the use of databases of adverse event reports is a good approach in order to study drug-induced safety events. Although this work remarks the role of several drugs in the induction of SCA, additional studies to understand more completely the key factors involved in the drug-induced sudden cardiac arrest are required.

Implications of metabolomics in drug discovery: insights and limitations

The studies presented in this Thesis highlight the important, yet traditionally highly neglected, role that the human endogenous metabolome has in biological systems and its implications for the effects of exogenous molecules in human bodies. Therefore, it represents a small but firm step in the consolidation of metabolomics as a key discipline with a

significant impact in the drug discovery process. As previous research suggested, a deeper understanding of the human endogenous metabolome could offer a whole new avenue to precision medicine.[93,95] In this respect, we suggest that adding this metabolomics perspective could help advancing the entire drug discovery process in multiple aspects such as, for example, refining the definition of safety margins and secondary effects, improving the assignment of primary targets, monitoring drug responses to customize drug dosing, or participating into the exploration of new therapeutic treatments.[89,93,95]

Nonetheless, there are also several limitations that merit further consideration. First of all, it is important to emphasise that biological systems are extremely complex and there are multiple components involved in the biological action of endogenous metabolites or drugs. In this regard, the metabolite concentration at the site of action fluctuates depending on the molecule type, tissue, cell type, environment or time and it could be certainly distinct in different individuals.[160,162] In parallel, drug exposure at the site of action also depends on several pharmacokinetics factors and individual particularities such as drug interference, comorbidities or environmental factors. Both drug exposure and metabolite abundance were not considered in any of the studies presented in this Thesis and it would be of interest to study in future works.

Another important limitation that we found during our studies was the recurrent issue of data completeness and the consequent bias in the available data, particularly for metabolites. Although several attempts were made in order to compile and present data for metabolites, there is a clear lack of information regarding the endogenous metabolites of human proteins and their pharmacology. We consider that it is extremely relevant to exhaustively identify the principal human endogenous metabolites, measure their *in vitro* binding affinities and finally identify their pharmacological profile. A reference public repository offering all this information would be a substantial improvement in order to better investigate the implications of human endogenous metabolome in drug discovery and other scientific fields.

Conclusions

The main conclusions that can be extracted from the results obtained in this Thesis can be summarized as follows:

a. There is a clear relationship between the minimum level of affinity needed by drugs for their primary targets and the affinity of the principal endogenous metabolite of these targets.

b. Endogenous metabolite affinities for their native proteins are directly proportional to the minimum affinity among all drugs optimized for a given primary target and inversely proportional to the range of drug affinities for that target.

c. The *in vitro* binding affinity of the endogenous metabolite for its native protein minus two orders of magnitude is a good reference baseline for the minimum affinity of a drug for its primary target.

d. *In vitro* binding affinities of endogenous metabolites for their native proteins offer a simple cost-effective means to alert on the risk of serious safety events linked to the drug's secondary pharmacology.

e. A difference of two orders of magnitude below the *in vitro* affinity of serotonin for 5HT2B has been determined and

demonstrated to successfully separate valvulopathic drugs from drugs devoid of risk to produce valvular heart disease.

f. Drug safety margins are recommended to be reviewed to include the difference in off-target binding affinities between the endogenous ligand and the drug.

g. Unlike drugs, natural compounds have not been synthetically optimized to produce a therapeutic effect in human organisms and thus, they are shown to have affinity values below the reference baseline set by the affinities of endogenous metabolites for their native proteins.

h. Those drug-like natural compounds that are capable to produce a therapeutic effect have affinity values for the human target above the reference baseline set by the affinities of metabolites for their native proteins.

i. Overall, the human endogenous metabolome may be acting, in some circumstances, as a defence mechanism against exogenous chemicals and its variability across individuals opens the door for new strategies in precision medicine.

References

1. Ng, R. Drugs: From Discovery to Approval. *John Wiley & Sons* (2011).
2. Jones, A. W. Early drug discovery and the rise of pharmaceutical chemistry. *Drug Testing and Analysis* vol. **3** 337–344 (2011).
3. Loriaux, D. L. & Lynn Loriaux, D. Diabetes and The Ebers Papyrus. *The Endocrinologist* vol. **16** 55–56 (2006).
4. Chast, F. A History of Drug Discovery. *The Practice of Medicinal Chemistry* 1–62 (2008) doi:10.1016/b978-0-12-374194-3.00001-9.
5. Drews, J. Drug Discovery: A Historical Perspective. *Science* vol. **287** 1960–1964 (2000).
6. Drews, J. *In Quest of Tomorrow's Medicines*. (Springer Science & Business Media, 2003).
7. Kibler, M. R. From the Mendeleev periodic table to particle physics and back to the periodic table. *Foundations of Chemistry* vol. **9** 221–234 (2007).
8. Pina, A. S., Hussain, A. & Roque, A. C. A. An historical overview of drug discovery. *Methods Mol. Biol.* **572**, 3–12 (2009).
9. Sertuerner & Sertuerner. Ueber das Morphium, eine neue salzfähige Grundlage, und die Mekonsäure, als Hauptbestandtheile des Opiums. *Annalen der Physik* vol. **55** 56–89 (1817).

10. Beswick, P. & Naylor, A. The role of medicinal chemistry in the drug discovery process. *Drug Discovery and Development* 119–134 (2013) doi:10.1016/b978-0-7020-4299-7.00009-3.
11. Lombardino, J. G. & Lowe, J. A. The role of the medicinal chemist in drug discovery — then and now. *Nature Reviews Drug Discovery* vol. **3** 853–862 (2004).
12. Langley, J. N. On the reaction of cells and of nerve-endings to certain poisons, chiefly as regards the reaction of striated muscle to nicotine and to curari. *J. Physiol.* **33**, 374–413 (1905).
13. Clark, A. J. *The Mode of Action of Drugs on Cells*. (1933).
14. Kenakin, T. Receptor theory. *Curr. Protoc. Pharmacol.* Chapter **1**, Unit1.2 (2008).
15. Ahlquist, R. P. A STUDY OF THE ADRENOTROPIC RECEPTORS. *American Journal of Physiology-Legacy Content* vol. **153** 586–600 (1948).
16. Moffat, J. G., Vincent, F., Lee, J. A., Eder, J. & Prunotto, M. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat. Rev. Drug Discov.* **16**, 531–543 (2017).
17. Takenaka, T. Classical vs reverse pharmacology in drug discovery. *BJU Int.* **88** Suppl **2**, 7–10; discussion 49–50 (2001).
18. Raut, A. A., Chorghade, M. S. & Vaidya, A. D. B. Reverse Pharmacology. *Innovative Approaches in Drug Discovery*

89–126 (2017) doi:10.1016/b978-0-12-801814-9.00004-0.

19. Carter, M. & Shieh, J. C. Molecular Cloning and Recombinant DNA Technology. *Guide to Research Techniques in Neuroscience* 207–227 (2010) doi:10.1016/b978-0-12-374849-2.00009-4.
20. Dagen, M. M. History of Insulin. *eLS* 1–3 (2016) doi:10.1002/9780470015902.a0003623.pub2.
21. Mullis, K. B. The Unusual Origin of the Polymerase Chain Reaction. *Scientific American* vol. **262** 56–65 (1990).
22. Goodwill, K. High-throughput x-ray crystallography for structure-based drug design. *Drug Discovery Today* vol. **6** 113–118 (2001).
23. Maveyraud, L. & Mourey, L. Protein X-ray Crystallography and Drug Discovery. *Molecules* vol. **25** 1030 (2020).
24. Bragg, W. H. The Reflection of X-Rays by Crystals. *Nature* vol. **91** 477–477 (1913).
25. Thomas, J. M. Centenary: The birth of X-ray crystallography. *Nature* **491**, 186–187 (2012).
26. Anderson, A. C. The Process of Structure-Based Drug Design. *Chemistry & Biology* vol. **10** 787–797 (2003).
27. Andricopulo, A. D. & Ferreira, L. L. G. *Chemoinformatics Approaches to Structure- and Ligand-Based Drug Design*. (Frontiers Media SA, 2019).
28. Liu, R., Li, X. & Lam, K. S. Combinatorial chemistry in drug discovery. *Curr. Opin. Chem. Biol.* **38**, 117–126 (2017).

29. Dibó, G. Combinatorial chemistry has matured in the last three decades: dedicated to Professor Árpád Furka on the occasion of his 80th birthday. *Molecular Diversity* vol. **16** 1–3 (2012).
30. Hopkins, A. L. Predicting promiscuity. *Nature* vol. **462** 167–168 (2009).
31. Antolín, A. A. & Mestres, J. The Impact of Distant Polypharmacology in the Chemical Biology of PARPs. *Concepts and Case Studies in Chemical Biology* 309–322 (2014) doi:10.1002/9783527687503.ch2.
32. Jalencas, X. & Mestres, J. On the origins of drug polypharmacology. *MedChemComm* vol. **4** 80–87 (2013).
33. Roth, B. L., Sheffler, D. J. & Kroeze, W. K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nat. Rev. Drug Discov.* **3**, 353–359 (2004).
34. Mestres, J., Gregori-Puigjané, E., Valverde, S. & Solé, R. V. Data completeness—the Achilles heel of drug-target networks. *Nature Biotechnology* vol. **26** 983–984 (2008).
35. Mestres, J., Gregori-Puigjané, E., Valverde, S. & Solé, R. V. The topology of drug-target interaction networks: implicit dependence on drug properties and target families. *Mol. Biosyst.* **5**, 1051–1057 (2009).
36. Strey, A. High Performance Computing. *International Encyclopedia of the Social & Behavioral Sciences* 6693–6697 (2001) doi:10.1016/b0-08-043076-7/00574-x.

37. ENIAC - CHM Revolution. https://www.computerhistory.org/revolution/birth-of-the-computer/4/78.
38. Ray, L. C. & Kirsch, R. A. Finding Chemical Records by Digital Computers. *Science* **126**, 814–819 (1957).
39. Willett, P. Chemoinformatics: a history. *WIREs Computational Molecular Science* vol. **1** 46–56 (2011).
40. Stepanets, G. F. & Vleduts, G. E. A method for the algorithmic description of the cyclical structure of a graph. *USSR Computational Mathematics and Mathematical Physics* vol. **3** 781–786 (1963).
41. Duelen, R. *et al.* Medicinal Biotechnology for Disease Modeling, Clinical Therapy, and Drug Discovery and Development. *Introduction to Biotech Entrepreneurship: From Idea to Business* 89–128 (2019) doi:10.1007/978-3-030-22141-6_5.
42. Hansch, C., Maloney, P. P., Fujita, T. & Muir, R. M. Correlation of Biological Activity of Phenoxyacetic Acids with Hammett Substituent Constants and Partition Coefficients. *Nature* vol. **194** 178–180 (1962).
43. Hansch, C. *et al.* The Correlation of Biological Activity of Plant Growth Regulators and Chloromycetin Derivatives with Hammett Constants and Partition Coefficients. *Journal of the American Chemical Society* vol. **85** 2817–2824 (1963).
44. Dearden, J. C. The History and Development of Quantitative Structure-Activity Relationships (QSARs).

Oncology 67–117 (2017) doi:10.4018/978-1-5225-0549-5.ch003.

45. Verma, J., Khedkar, V. M. & Coutinho, E. C. 3D-QSAR in drug design--a review. *Curr. Top. Med. Chem.* **10**, 95–115 (2010).
46. Goodsell, D. S. *et al.* RCSB Protein Data Bank: Enabling biomedical research and drug discovery. *Protein Sci.* **29**, 52–65 (2020).
47. wwPDB consortium. Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* **47**, D520–D528 (2019).
48. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R. & Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **161**, 269–288 (1982).
49. Gauthier, J., Vincent, A. T., Charette, S. J. & Derome, N. A brief history of bioinformatics. *Briefings in Bioinformatics* vol. **20** 1981–1996 (2019).
50. Berners-Lee, T., Cailliau, R., Groff, J. & Pollermann, B. World-wide web: the information universe. *Internet Research* vol. **20** 461–471 (2010).
51. Rice, C. M., Fuchs, R., Higgins, D. G., Stoehr, P. J. & Cameron, G. N. The EMBL data library. *Nucleic Acids Res.* **21**, 2967–2971 (1993).
52. Burks, C. *et al.* GenBank. *Nucleic Acids Res.* **20** Suppl, 2065–2069 (1992).

53. Wishart, D. S. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Research* vol. **34** D668–D672 (2006).
54. Gaulton, A. *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100–7 (2012).
55. Inchikey technical FAQ. https://www.inchi-trust.org/technical-faq-2/.
56. Papadatos, G. *et al.* SureChEMBL: a large-scale, chemically annotated patent document database. *Nucleic Acids Res.* **44**, D1220–8 (2016).
57. Hutchinson, L. *et al.* Models and Machines: How Deep Learning Will Take Clinical Pharmacology to the Next Level. *CPT Pharmacometrics Syst Pharmacol* **8**, 131–134 (2019).
58. Krogh, A. What are artificial neural networks? *Nat. Biotechnol.* **26**, 195–197 (2008).
59. Lima, A. N. *et al.* Use of machine learning approaches for novel drug discovery. *Expert Opinion on Drug Discovery* vol. **11** 225–239 (2016).
60. Khazendar, S. *et al.* Automated classification of static ultrasound images of ovarian tumours based on decision level fusion. *2014 6th Computer Science and Electronic Engineering Conference (CEEC)* (2014) doi:10.1109/ceec.2014.6958571.
61. Mullard, A. 2012 FDA drug approvals. *Nat. Rev. Drug Discov.* **12**, 87–90 (2013).

62. Munos, B. Lessons from 60 years of pharmaceutical innovation. *Nat. Rev. Drug Discov.* **8**, 959–968 (2009).
63. Lexchin, J. The Pharmaceutical Industry in Contemporary Capitalism. *Monthly Review* vol. **69** 37 (2018).
64. Ledley, F. D., McCoy, S. S., Vaughan, G. & Cleary, E. G. Profitability of Large Pharmaceutical Companies Compared With Other Large Public Companies. *JAMA* **323**, 834–843 (2020).
65. Kinch, M. S., Haynesworth, A., Kinch, S. L. & Hoyer, D. An overview of FDA-approved new molecular entities: 1827-2013. *Drug Discov. Today* **19**, 1033–1039 (2014).
66. Patterson, J. A. & Carroll, N. V. Should the United States government regulate prescription prices? A critical review. *Res. Social Adm. Pharm.* **16**, 717–723 (2020).
67. Smith, C. Drug target validation: Hitting the target. *Nature* vol. **422** 342–345 (2003).
68. Hughes, J. P., Rees, S., Kalindjian, S. B. & Philpott, K. L. Principles of early drug discovery. *British Journal of Pharmacology* vol. **162** 1239–1249 (2011).
69. Mohs, R. C. & Greig, N. H. Drug discovery and development: Role of basic biological research. *Alzheimers. Dement.* **3**, 651–657 (2017).
70. Friedman, L. M., Furberg, C. D. & DeMets, D. L. Introduction to Clinical Trials. *Fundamentals of Clinical Trials* 1–18 (2010) doi:10.1007/978-1-4419-1586-3_1.
71. Suvarna, V. Phase IV of Drug Development. *Perspect. Clin. Res.* **1**, 57–60 (2010).

72. Johnson-Pratt, L. R. Phase IV Drug Development: Post-Marketing Studies. *Principles and Practice of Pharmaceutical Medicine* 133–139 doi:10.1002/0470846291.ch12.
73. Kell, D. B. & Oliver, S. G. The metabolome 18 years on: a concept comes of age. *Metabolomics* vol. 12 (2016).
74. Oliver, S. Systematic functional analysis of the yeast genome. *Trends in Biotechnology* vol. **16** 373–378 (1998).
75. Zamboni, N., Saghatelian, A. & Patti, G. J. Defining the metabolome: size, flux, and regulation. *Mol. Cell* **58**, 699–706 (2015).
76. Pogson, C. An introduction to metabolic pathways By S. Dagley and D. E. Nicholson. Pp xi 343. Blackwell Scientific Publications, Oxford and Edinburgh. 1970. E3.75. *Endeavour* vol. **30** 154 (1971).
77. Williams, R. J. Biochemical Individuality. The Basis for the Genetotrophic Concept. *AIBS Bulletin* vol. **7** 43 (1957).
78. Oliver, S. G. *et al.* The complete DNA sequence of yeast chromosome III. *Nature* **357**, 38–46 (1992).
79. Fillet, M. & Frédérich, M. The emergence of metabolomics as a key discipline in the drug discovery process. *Drug Discov. Today Technol.* **13**, 19–24 (2015).
80. Kacser, H. & Burns, J. A. The control of flux. *Symp. Soc. Exp. Biol.* **27**, 65–104 (1973).
81. Curtis, R. K., Keira Curtis, R., Bishop, T. & Brand, M. D. Control Analysis of Metabolic Depression. *Cell and*

Molecular Response to Stress 283–296 (2002) doi:10.1016/s1568-1254(02)80022-5.

82. Brown, M. *et al.* Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. *The Analyst* vol. **134** 1322 (2009).
83. Alseekh, S. & Fernie, A. R. Metabolomics 20 years on: what have we learned and what hurdles remain? *Plant J.* **94**, 933–942 (2018).
84. Wishart, D. S. *et al.* HMDB: the Human Metabolome Database. *Nucleic Acids Research* vol. **35** D521–D526 (2007).
85. Wishart, D. S. *et al.* HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Research* vol. **46** D608–D617 (2018).
86. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
87. Alexander, S. P. H. *et al.* THE CONCIS GUIDE TO PHARMACOLOGY 2019/20: Introduction and Other Protein Targets. *Br. J. Pharmacol.* **176** Suppl 1, S1–S20 (2019).
88. Alexander, S. P. H., Mathie, A. & Peters, J. A. Guide to Receptors and Channels (GRAC), 5th edition. *Br. J. Pharmacol.* **164** Suppl 1, S1–324 (2011).
89. Wishart, D. S. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat. Rev. Drug Discov.* **15**, 473–484 (2016).

90. Wang, T. J. *et al.* Metabolite profiles and the risk of developing diabetes. *Nat. Med.* **17**, 448–453 (2011).
91. Wang, T. J. *et al.* 2-Aminoadipic acid is a biomarker for diabetes risk. *J. Clin. Invest.* **123**, 4309–4317 (2013).
92. Palmer, N. D. *et al.* Metabolomic profile associated with insulin resistance and conversion to diabetes in the Insulin Resistance Atherosclerosis Study. *J. Clin. Endocrinol. Metab.* **100**, E463–8 (2015).
93. Clish, C. B. Metabolomics: an emerging but powerful tool for precision medicine. *Cold Spring Harb Mol Case Stud* **1**, a000588 (2015).
94. Kaddurah-Daouk, R., Weinshilboum, R. M. & Pharmacometabolomics Research Network. Pharmacometabolomics: implications for clinical pharmacology and systems pharmacology. *Clin. Pharmacol. Ther.* **95**, 154–167 (2014).
95. Bar, N. *et al.* A reference map of potential determinants for the human serum metabolome. *Nature* (2020) doi:10.1038/s41586-020-2896-2.
96. Fornasier, G., Francescon, S., Leone, R. & Baldo, P. An historical overview over Pharmacovigilance. *Int. J. Clin. Pharm.* **40**, 744–747 (2018).
97. Report of the Lancet commission appointed to investigate the subject of the administration of chloroform and other anesthetic from a clinical standpoint. *The Lancet* vol. **141** 899–914 (1893).

98. Eaglstein, W. H. Brief History of the FDA. *The FDA for Doctors* 89–93 (2014) doi:10.1007/978-3-319-08362-9_24.
99. Gram, H. The Birth of the FDA. https://www.drugtimes.org/everyday/the-birth-of-the-fda.html (2018).
100. McBride, W. G. Thalidomide and Congenital Abnormalities. *Problems of Birth Defects* 199–199 (1961) doi:10.1007/978-94-011-6621-8_27.
101. Lenz, W. Thalidomide and Congenital Abnormalities. *Problems of Birth Defects* 199–199 (1962) doi:10.1007/978-94-011-6621-8_28.
102. Smithells, R. W. & Newman, C. G. Recognition of thalidomide defects. *Journal of Medical Genetics* vol. **29** 716–723 (1992).
103. EMA. ICH Topic S 7 A Safety Pharmacology Studies for Human Pharmaceuticals. *Euroepan Medicines Agency* https://www.ema.europa.eu/en/documents/scientific-guideline/ich-s-7-safety-pharmacology-studies-human-pharmaceuticals-step-5_en.pdf (2001).
104. Bass, A., Kinter, L. & Williams, P. Origins, practices and future of safety pharmacology. *J. Pharmacol. Toxicol. Methods* **49**, 145–151 (2004).
105. EMA. ICH Topic S 7 B The nonclinical Evaluation of the Potential for delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals. *EMA* https://www.ema.europa.eu/en/documents/scientific-

guideline/ich-s-7-b-nonclinical-evaluation-potential-delayed-ventricular-repolarization-qt-interval_en.pdf (2005).

106. Monahan, B. P. *et al.* Torsades de pointes occurring in association with terfenadine use. *JAMA* **264**, 2788–2790 (1990).
107. Mathews, D. R., McNutt, B., Okerholm, R., Flicker, M. & McBride, G. Torsades de pointes occurring in association with terfenadine use. *JAMA: the journal of the American Medical Association* vol. **266** 2375–2376 (1991).
108. Roy, M., Dumaine, R. & Brown, A. M. HERG, a primary human ventricular target of the nonsedating antihistamine terfenadine. *Circulation* **94**, 817–823 (1996).
109. Kannankeril, P., Roden, D. M. & Darbar, D. Drug-induced long QT syndrome. *Pharmacol. Rev.* **62**, 760–781 (2010).
110. Rudmann, D. G. On-target and off-target-based toxicologic effects. *Toxicol. Pathol.* **41**, 310–314 (2013).
111. Whitebread, S., Hamon, J., Bojanic, D. & Urban, L. Keynote review: In vitro safety pharmacology profiling: an essential tool for successful drug development. *Drug Discovery Today* vol. **10** 1421–1433 (2005).
112. Dambach, D. M. *et al.* Safety Lead Optimization and Candidate Identification: Integrating New Technologies into Decision-Making. *Chemical Research in Toxicology* vol. **29** 452–472 (2016).
113. Papoian, T. *et al.* Secondary pharmacology data to assess potential off-target activity of new drugs: a

regulatory perspective. *Nature Reviews Drug Discovery* vol. **14** 294–294 (2015).

114. Papoian, T. *et al.* Regulatory Forum Review*: 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 (2017).
115. Zipes, D. P. & Wellens, H. J. J. Sudden cardiac death. *Professor Hein J.J. Wellens: 33 Years of Cardiology and Arrhythmology* 621–645 (2000) doi:10.1007/978-94-011-4110-9_60.
116. Wellens, H. J. J. & Brugada, P. Sudden cardiac death: A multifactorial problem. *Sudden Cardiac Death* 285–296 (1991) doi:10.1007/978-94-009-0573-3_24.
117. Vreede-Swagemakers, J. J. M. de *et al.* Out-of-Hospital Cardiac Arrest in the 1990s: A Population-Based Study in the Maastricht Area on Incidence, Characteristics and Survival. *Journal of the American College of Cardiology* vol. **30** 1500–1505 (1997).
118. Tan, H. L., Dagres, N., Böttiger, B. W., Schwartz, P. J. & ESCAPE-NET Investigators. European Sudden Cardiac Arrest network: towards Prevention, Education and New Effective Treatments (ESCAPE-NET). *European Heart Journal* vol. **39** 86–88 (2018).
119. Adabag, A. S., Selcuk Adabag, A., Luepker, R. V., Roger, V. L. & Gersh, B. J. Sudden cardiac death: epidemiology

and risk factors. *Nature Reviews Cardiology* vol. **7** 216–225 (2010).

120. Chugh, S. S. *et al.* Epidemiology of Sudden Cardiac Death: Clinical and Research Implications. *Progress in Cardiovascular Diseases* vol. **51** 213–228 (2008).
121. Kaltman, J. R. *et al.* Screening for Sudden Cardiac Death in the Young. *Circulation* vol. **123** 1911–1918 (2011).
122. Fishman, G. I. *et al.* Sudden Cardiac Death Prediction and Prevention. *Circulation* vol. **122** 2335–2348 (2010).
123. Chugh, S. S. Sudden cardiac death in 2017: Spotlight on prediction and prevention. *Int. J. Cardiol.* **237**, 2–5 (2017).
124. Zheng, Z. J., Croft, J. B., Giles, W. H. & Mensah, G. A. Sudden cardiac death in the United States, 1989 to 1998. *Circulation* **104**, 2158–2163 (2001).
125. Jayaraman, R. *et al.* Risk Factors of Sudden Cardiac Death in the Young. *Circulation* vol. **137** 1561–1570 (2018).
126. Myerburg, R. J. & Goldberger, J. J. Sudden Cardiac Arrest Risk Assessment: Population Science and the Individual Risk Mandate. *JAMA Cardiol* **2**, 689–694 (2017).
127. D'Agostino, R. B. *et al.* General Cardiovascular Risk Profile for Use in Primary Care. *Circulation* vol. **117** 743–753 (2008).
128. Thomsen, T. HeartScore: a new web-based approach to European cardiovascular disease risk management. *Eur. J. Cardiovasc. Prev. Rehabil.* **12**, 424–426 (2005).

129. Osman, J., Tan, S. C., Lee, P. Y., Low, T. Y. & Jamal, R. Sudden Cardiac Death (SCD) - risk stratification and prediction with molecular biomarkers. *J. Biomed. Sci.* **26**, 39 (2019).
130. Albert, C. M. *et al.* Cardiac Sodium Channel Gene Variants and Sudden Cardiac Death in Women. *Circulation* vol. **117** 16–23 (2008).
131. Marcsa, B. *et al.* A Common Polymorphism of the Human Cardiac Sodium Channel Alpha Subunit (SCN5A) Gene Is Associated with Sudden Cardiac Death in Chronic Ischemic Heart Disease. *PLoS One* **10**, e0132137 (2015).
132. Bezzina, C. R. *et al.* Genome-wide association study identifies a susceptibility locus at 21q21 for ventricular fibrillation in acute myocardial infarction. *Nat. Genet.* **42**, 688–691 (2010).
133. Arking, D. E. *et al.* Genome-wide association study identifies GPC5 as a novel genetic locus protective against sudden cardiac arrest. *PLoS One* **5**, e9879 (2010).
134. Veerman, C. C., Wilde, A. A. M. & Lodder, E. M. The cardiac sodium channel gene SCN5A and its gene product NaV1.5: Role in physiology and pathophysiology. *Gene* vol. **573** 177–187 (2015).
135. Perrin, M. J., Subbiah, R. N., Vandenberg, J. I. & Hill, A. P. Human ether-a-go-go related gene (hERG) K channels: Function and dysfunction. *Progress in Biophysics and Molecular Biology* vol. **98** 137–148 (2008).

136. Ali, A., Butt, N. & Sheikh, A. S. Early repolarization syndrome: A cause of sudden cardiac death. *World J. Cardiol.* **7**, 466–475 (2015).
137. Winkel, B. O. G. *et al.* The Prevalence of Mutations inKCNQ1, KCNH2,andSCN5Ain an Unselected National Cohort of Young Sudden Unexplained Death Cases. *Journal of Cardiovascular Electrophysiology* vol. **23** 1092–1098 (2012).
138. Albert, C. M. *et al.* Cardiac sodium channel gene variants and sudden cardiac death in women. *Circulation* **117**, 16–23 (2008).
139. Belevych, A. E., Radwański, P. B., Carnes, C. A. & Györke, S. 'Ryanopathy': causes and manifestations of RyR2 dysfunction in heart failure. *Cardiovasc. Res.* **98**, 240–247 (2013).
140. Ramalho, D. & Freitas, J. Drug-induced life-threatening arrhythmias and sudden cardiac death: A clinical perspective of long QT, short QT and Brugada syndromes. *Rev. Port. Cardiol.* **37**, 435–446 (2018).
141. Killeen, M. J. Drug-induced arrhythmias and sudden cardiac death: implications for the pharmaceutical industry. *Drug Discov. Today* **14**, 589–597 (2009).
142. Shah, R. R. & Hondeghem, L. M. Refining detection of drug-induced proarrhythmia: QT interval and TRIaD. *Heart Rhythm* **2**, 758–772 (2005).

143. Sicouri, S. & Antzelevitch, C. Sudden cardiac death secondary to antidepressant and antipsychotic drugs. *Expert Opin. Drug Saf.* **7**, 181–194 (2008).
144. Morentin, B., Callado, L. F., García-Hernández, S., Bodegas, A. & Lucena, J. The role of toxic substances in sudden cardiac death. *Spanish Journal of Legal Medicine* vol. **44** 13–21 (2018).
145. Haverkamp, W. & Breithardt, G. Drug-induced sudden cardiac death. *European heart journal* vol. **26** 1808–9; author reply 1809 (2005).
146. Daher, M. *et al.* Sudden cardiac arrest in general anesthesia as the first manifestation of anomalous origin of the left coronary artery. *Rev. Bras. Anestesiol.* **62**, 878–884 (2012).
147. Tabib, A. *et al.* [Sudden death during anesthesia: human error, drug related or cardiac death?]. *Therapie* **56**, 735–738 (2001).
148. Ray, W. A. *et al.* Antipsychotics and the risk of sudden cardiac death. *Arch. Gen. Psychiatry* **58**, 1161–1167 (2001).
149. Black, S. Anesthesia-Related Cardiac Arrest in Children: Update from the Pediatric Perioperative Cardiac Arrest Registry. *Yearbook of Anesthesiology and Pain Management* vol. **2008** 91–92 (2008).
150. Braghiroli, K. S. *et al.* Perioperative and anesthesia-related cardiac arrests in geriatric patients: a systematic

review using meta-regression analysis. *Sci. Rep.* **7**, 2622 (2017).

151. Ellis, S. J. *et al.* Anesthesia-related cardiac arrest. *Anesthesiology* **120**, 829–838 (2014).
152. Gong, C.-L. *et al.* A study of anaesthesia-related cardiac arrest from a Chinese tertiary hospital. *BMC Anesthesiol.* **18**, 127 (2018).
153. Chalkiadis, G. A. & Branch, K. G. Cardiac arrest after isoflurane anaesthesia in a patient with Duchenne's muscular dystrophy. *Anaesthesia* vol. **45** 22–25 (1990).
154. Wyant, G. M. Halothane and cardiac arrest. *British Journal of Anaesthesia* vol. **32** 563 (1960).
155. Adekola, O. O., Asiyanbi, G. K., Desalu, I., Olatosi, J. O. & Kushimo, O. T. The outcome of anaesthesia related cardiac arrest in a Sub-Saharan tertiary hospital. *Egyptian Journal of Anaesthesia* vol. **32** 315–321 (2016).
156. Karlis, G. & Afantenou, S. Drugs in cardiac arrest: the rise and fall of antiarrhythmics. *Acta Cardiol.* **73**, 325–327 (2018).
157. Fish, F. A., Gillette, P. C. & Benson, D. W., Jr. Proarrhythmia, cardiac arrest and death in young patients receiving encainide and flecainide. The Pediatric Electrophysiology Group. *J. Am. Coll. Cardiol.* **18**, 356–365 (1991).
158. Akiyama, T., Pawitan, Y., Greenberg, H., Kuo, C. S. & Reynolds-Haertle, R. A. Increased risk of death and cardiac arrest from encainide and flecainide in patients

after non-Q-wave acute myocardial infarction in the Cardiac Arrhythmia Suppression Trial. CAST Investigators. *Am. J. Cardiol.* **68**, 1551–1555 (1991).

159. Suhre, K. *et al.* Human metabolic individuality in biomedical and pharmaceutical research. *Nature* **477**, 54–60 (2011).
160. Muller, P. Y. & Milton, M. N. The determination and interpretation of the therapeutic index in drug development. *Nature Reviews Drug Discovery* vol. **11** 751–761 (2012).
161. Gillum, R. F. Geographic variation in sudden coronary death. *Am. Heart J.* **119**, 380–389 (1990).
162. Bennett, B. D. *et al.* Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. Biol.* **5**, 593–599 (2009).