Review Article

Overview of the advancement in the Drug Discovery and contribution in the drug development Milestones

ABSTRACT.

The development of a new chemical entity from original idea to the regulatory approval for launch into the market is a very long and complex journey. This complex process can take between 10-15 years, with a big financial cost of over $1 billion. The molecule can be derived from many sources including academic and clinical research and from the commercial sector. It takes many years to develop valid evidence-based supporting document before selecting a target for advancement of the drug discovery programme. The success of a chosen target can motivate the pharmaceutical industry or some academic institutions to embark on a number of early processes to identify molecules which possess suitable characteristics to make acceptable drugs. This review gives an overview of the drug discovery odyssey at the preclinical level,the identification of initial target and validation, through assay development, high throughput screening, hit identification, lead optimization and finally the selection of a candidate molecule for clinical development after the regulatory new drug approval.

Keywords new chemical entity, drug discovery; target identification, target validation, high throughput screening (HTS), assay development; screening cascade; lead optimization, IND.

UNDER PEER REVIEW

INTRODUCTION

A drug discovery process comes into action once the five-challenging question for drug discovery and development has been answered. These questions are; Is there a medical need? What is the prevalence of the disease? What is the market potential? Do we have a biochemical target? Can the compound be synthesized that are target selective, potent *in vivo*, and bio-available? Are the compounds efficacious in disease models, show dose response, and are not toxic? and finally, is there a disease or clinical condition without suitable medical products available and it has this unmet clinical need [1-3]. A drug or new chemical entity (NCE) is any substance, other than food, that is used to prevent, diagnose, treat or relieve symptoms of a disease or abnormal condition [3]. It can also be considered as a substance that alters mood or body function, or that can be habit-forming or addictive, especially a narcotic [3, 4].

The initial research, often occurring in academia, generates data to develop a hypothesis that the inhibition or activation of a protein or pathway will result in a therapeutic effect in a disease state. The outcome of this activity is the selection of a target which may require further validation prior to progression into the lead discovery phase in order to justify a drug discovery effort. During lead discovery, an intensive search enables identification of a drug-like small molecule or biological therapeutic agent, typically termed a

development candidate, that will progress into preclinical, and if successful, into clinical development and ultimately be a marketed medicine [5].

Types of Pharmaceutical Companies

There are different types of pharma companies involved in drug discovery and development. They include;

- ✓ **Pharmaceutical Drug Discovery/Development** They take drug discovery from hit, lead optimization to the drug development phase and marketed product. This category of company includes examples like Pfizer, GlaxoSmithKline, Merck Sharp, Novartis, AstraZeneca [5].
- ✓ **The Pharmaceutical Drug Delivery companies**; They take the lead compound from different sources to development and market. Most generic drugs are produced by these companies like Élan Corporation, Alza Corporation.
- ✓ **The Biotech-Pharmaceuticals**; They develop biotechnology tools, methods, devices and gene targets for the drug development integration. They usually collaborate with other pharma companies or go into merger. Examples include Genentech, Amgen.
- ✓ **Contract Research Organization (CRO).** They are not directly involved in the drug development process, but are in charge of running and conducting clinical trials for the pharma companies [5].

Major challenges of Pharmaceutical companies

The major challenges Pharma companies face are;
Time and money-It is a long process to develop a molecule and launch in the market (10-15 years). Competition, R&D spending, Patent life, Price controls, Government legislation, Regulatory requirements, Managed health care, Cost of new enabling technology, Management of alliances and biotech venture [5].
Medicines discovery dates back to the Neanderthals to the people of Mesopotamia, Egypt, Greece and China where man used herbal remedies to treat the sick people. In mediaeval times the quest for the *elixir of life* was pursued by alchemists, but it is the scientists of the past 100-150 years who have had success by translating laboratory-based discoveries into drugs that have significantly survived millions of lives [5, 6]. There was the collapse of the German stock market in 1873 and during the recovery period there was an economic boom that led to an expansion of chemical and electrical industries [3, 7]. The rising interest in the investment in the manufacture of synthetic dyes eventually placed Germany ahead of all its competitors in technology. Consequently, the German chemists did not only become very influential in the field of organic chemistry, but also led to the upsurge of the German pharmaceutical industry 8].

At the leading end of the industry were the leading manufacturers F. Bayer & Company and Farbenfabriken Hoechst who realized that their chemists researching and developing dyes also had the potential to produce new medicines [9]. One such scientist was Paul Ehrlich, who was fascinated by his research on colourful dyes and their capacity to interact with histological and cellular structures [4, 10]. Ehrlich from his research over many decades, later benefitted from chemical companies who provided hundreds to thousands of new dyes for him to continue his research. By establishing that these dyes were biologically active, evaluated a number of compounds probably exceeding thousands or even millions as a part of the high-throughput screening (HTS) now employed in academia and industry today [11]. Ehrlich elucidated that, the biological effect of a chemical compound such as a dye depends on its chemical composition and the cell on which it acts [12]. He was able to establish a connection between chemistry, biology and

UNDER PEER REVIEW

medicine in a very scientifically valid manner. Ehrlich postulated that chemical dyes were the catalyst for this *revolutionary association* [13]. In addition to Ehrlich fascination of dyes, he was also inspired by his colleagues conducting research in the field of immunology like, Louis Pasteur, Robert Koch, Emil von Behring and Shibasaburo Kitasato [2]. Towards the end of the 20th century, Ehrlich developed the receptor theory, which was the key instrument that led to the understanding of how the binding of drugs to various types of receptors could occur due to structural differences in chemical compositions. Ehrlich's research with the treatment of infectious diseases with drugs derived from the German dye industry motivated him to develop ways of using organic chemistry to modify certain starting dyes in various ways in finding new chemical structures with promising potential for biological activity [13, 15]. Erlich has been described as the founder of chemotherapy and his 'magic bullet concept' is still what today's scientists strive to aim for when developing small molecules that attack pathogens yet remain harmless to healthy tissues [9].

During the two World Wars, essential medicines normally supplied by Germany was scarce and led to gradual shift for exploitation of synthetic drugs. Synthetic organic chemistry became an exceptional important discipline and is still one of the backbones of drug discovery [16]. Synthetic organic chemistry has continually adapted to integrate innovative techniques and methodologies central to drug development. Much synthetic drug discovery stemmed from cancer drug development and was born from an observation that mustard gas, employed in chemical warfare during World Wars I and II, destroyed lymphatic tissue and bone marrow formation. The observations made by Drs. Gilman, Goodman and co-workers laid the foundation for conducting the first clinical trials with nitrogen mustards (βchloroethylamines) in 1942 at Yale-New Haven Hospital, although report of the clinical results was only made public four years later, due to the spirit of secrecy during World War II [17, 18]. An array of DNA alkylating agents was produced which triggered an increased understanding of DNA in the 1950s. A number of other agents were also discovered, such as the vinca alkaloids and purine/pyrimidine synthesis inhibitors [19], with the research mostly sponsored by the National Cancer Institute (NCI). Such studies led to the assessment of primarily cytotoxic agents and by the 1970s, the importance of natural product-based early drug discovery was achieved [20]. The new synthesis of many of these frequently promising, novel agents was often too complex and too expensive to allow progression into early stages of clinical trials. This situation caused a paradigm shift from natural product screening to a stage of discovery initiatives, providing an opportunity for the identification of natural products as potential lead compounds [17, 21]. These lead compounds were subsequently developed into bioactive molecules, which were more improved to synthesis.

Recent advances in organic chemistry have successfully led to the complete synthesis of many complex natural products, a milestone that has significantly improved the ease with which chemists can now deal with the complexity of many of these naturally-derived architectural structures [22]. Synthetic chemistry has also been instrumental in the development of drug delivery and prodrug strategies, which now focuses on the development of selective therapeutics with reduced side-effect profiles [23]. Although research in cancer medicines was the focus of many synthetic drug discovery, this was conducted in collaboration with research in other therapeutic disease areas as illustrated in figure 1, with an elaborate chronology of the drug innovation processes..

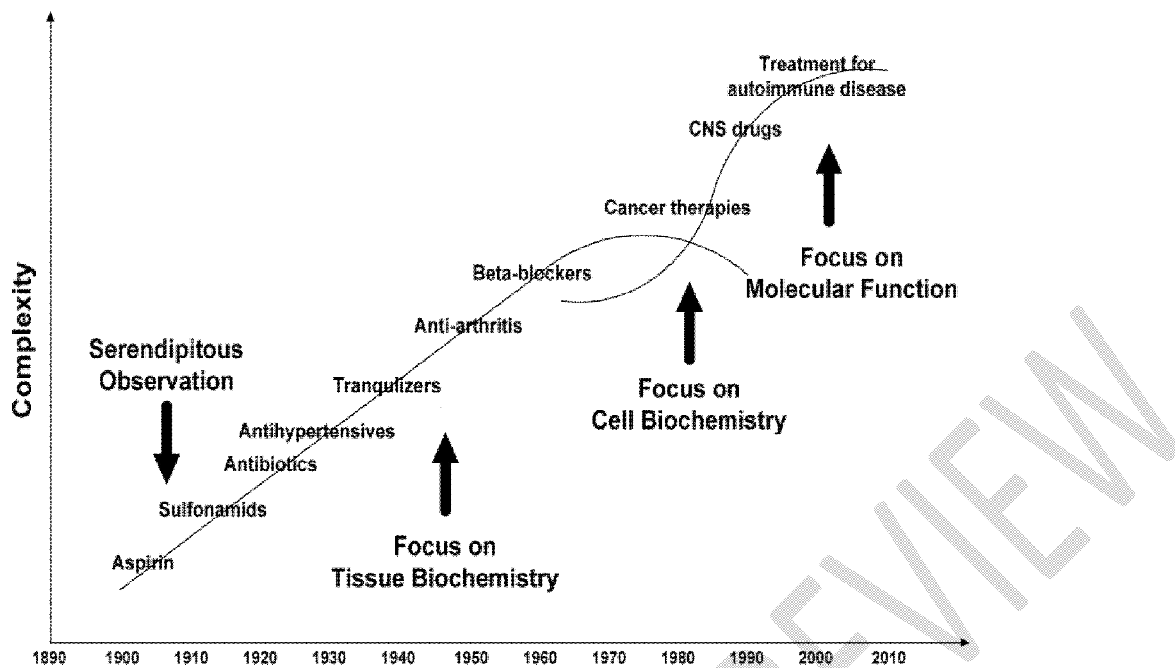

Fig. 1. Chronology of drug innovation process [22]

Today, the emergence of the genomics era and the focus of events at the molecular level is changing the landscape of drug discovery. Data mining reveals the generation of many data that has caused many not only to speculate on an expanding druggable genome, but also given new opportunities for taking drug discovery to the next level [23]. The number of gene products that are targets for existing drugs has been a topic for much debate and depends on the analysis performed. However, a valid estimate is in the region of 300-500 gene products potential drug targets [7, 24]. As the human genome is estimated to encode 20,000-25,000 human gene products, the number of drug targets is likely to increase. However, it will take some time to validate targets at the protein level, which has an added level of complexity. Both gene and protein expression profiling methodologies have been developed although with major challenges. In the last two decades these methods are used to monitor and catalogue changes in the expression of genes and their respective protein products [25]. There is more priority in the interest of understanding of human disease at the molecular level than to the elucidation of changes in biochemical processes associated with disease phenotypes.

The drug discovery objective is geared towards the generation of identifiable therapeutic targets that can reduce the drug development attrition [6]. The mapping of the human genome was a great breakthrough with scientists working at the interface of chemistry and biology in drug discovery, making use of the data available to them to discover new blockbuster drugs. The ever-increasing cost of drug discovery has been a major hindrance for Research and development, especially in times of recession. There is serious question if research and development (R&D) in the emerging markets can become an opportunity to advance to the next level of understanding of how to develop successful drugs. [7].

Why do we discover new drugs?

UNDER PEER REVIEW

There is a continuous search for new chemical entities due to the following factors; An increased discovery of large therapeutic compounds from natural products, emergence of many pharmaceutical companies in the last 50 years. Academic and research institutions have witnessed a global increase. Much progress has been made in the understanding of disease aetiology and processes, with increased knowledge on the mechanism that controls or eliminates. Most current disease treatment only lead to symptom relief with some undesirable side effects in some cases. The drug rresistance and tolerance, ineffective against pathogenic invasions in cases of (Tuberculosis, HIV malaria are well documented. The changing lifestyle and increase in life span have created life style disease, and changing social attitudes have created more market for "lifestyle drugs". The need to discover treatment for old and evolving diseases has increased due to inadequacies of current medicines for HIV, Diabetes, Cancer etc. There has also been a significant progress in molecular biology (Sequencing of Human genome). OMICs & protein engineering has given room for more understanding of precise disease mechanism, biochemical pathways and discovery of new targets [25-27].

2. The evolution of modern drug discovery

At the beginning of the 20th century, drug discovery was initiated mostly by individuals such as Paul Ehrlich and his associates. The approach now requires collaboration in a multidisciplinary setting from various disciplines including chemistry, computational modelling, structural biology and pharmacology [13]. The information available from the literature about a specific disease or target is now used by the research team to decide what intervention would be most likely to bring about the desired result. The exact nature of the project progression depends on the resources available: for example, an academic group may not be as expansive as a large pharmaceutical company in terms of how to tackle the problems of validating a novel target or developing 'hit' and 'lead' compounds that will be able to modulate that target [28]. The drug discovery process outlined in Figure 2 is, therefore, an approximate model which is employed by pharmaceutical companies, but one which a small biotech company or a university also can engage in through multiple collaborations. This discovery process can be instigated at several points and adapted to bring about the results needed to take a project to the next level in line with a recent in-depth review of the early drug discovery process.

UNDER PEER REVIEW

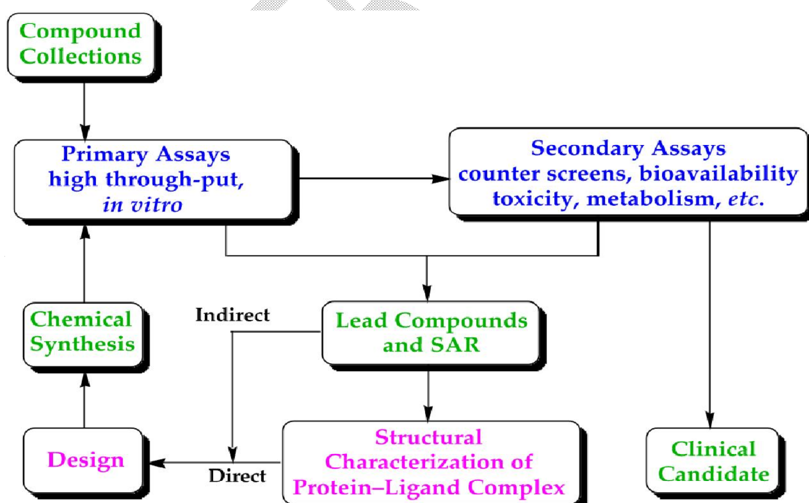

Fig. 2. The multi-faceted drug discovery process in the 21st Century.

Up until the mid-80s, drug discovery was focused on the isolation of natural products and medicinal chemistry was central for a research team to find more potent and selective compounds than the natural

product or synthetic compound themselves. After isolation and characterization of the natural products, structure-activity relationship (SAR) studies were and still are a vital tool in optimizing a pharmacophore [29]. Initially, a drug design process was an instrumental course of action between the synthesis of new compounds by a synthetic/medicinal chemist and the screening of these for biological activity by a pharmacologist. The drug discovery process was chemistry-focused rather than target driven. The discovery process of a drug now involves a multidisciplinary effort that is synergistic, which often encompasses high throughput screening (HTS) procedures. It is also one that often follows rules that are based on empirical findings from clinical investigations such as Lipinski's rule of 5 [9]. 'Hit' compounds are progressed into a 'lead' compound, which undergoes thorough pharmacological and toxicological testing. The results of these tests enable a research team to decide whether it is profitable to continue with the progression of a specific project.

The scenario with medicinal chemists is often to screen virtual or commercial libraries of compounds to identify hit molecules. The second stage is to prepare libraries of small molecules based around the hit molecule, measure their activity and correlate the results to determine the chemical structure with optimum activity [12]. This analysis may make use of SARs, computational chemistry, combinatorial chemistry and enzymatic and cellular assays, to help unravel biological activity derived from unique mechanism of action of a small molecule. The selection of a lead compound and the development of a synthetic pathway for its preparation on a large scale for preclinical and clinical investigations must also be considered at an early stage in the discovery process. If the lead molecule cannot be synthesized on a large scale, progression to clinical evaluation will not be possible. Similarly, researchers must also devise suitable *in vitro* and *in vivo* tests to assess the activity and toxicity of the compounds produced. If there is no suitable way of testing a hit or lead molecule *in vivo* the project may come to a halt unless it is decided to spend resources on developing appropriate models [3, 28-30].

Nowadays, hit and lead molecules with proven activity are assessed for susceptibility for phase I and II metabolism in the very early stages of the discovery process. For example, many HTS technologies are now available to detect cytochrome P450 (CYP) substrates or inhibitors, which should decrease the number of withdrawals of novel drugs from the market due to affinity for major CYP metabolizing isozymes. HTS CYP data can be used to guide medicinal chemist away from these interactions at an early stage and in certain cases might entirely solve the issue by targeted modification of the CYP interacting functionality [32].
HTS methodologies have been developed and have enabled research teams to generate vast numbers of compound variations of a desired pharmacophore.

Combinatorial chemistry

Combinatorial chemistry (combichem) was first applied to the generation of peptide arrays in 1984 and evolved rapidly into a new discipline that was recognized to have revolutionized drug discovery [33]. The early generations of combichem scientists captured the interest of the industry, and modified the common use of a number of terminologies, and abbreviations that became widespread in the literature including deconvolution, diversomer, split-and-mix, multipin, SPOC or SPOS (solid-phase organic chemistry or synthesis), submonomer synthesis, T-bag (Teflon bag) to name a few [17, 34]. The early scientists working in the combichem environment required different management solutions to classical synthetic chemists. For example, the chemists involved in planning a traditional synthesis to obtain a target compound or a natural product typically conducted a retrosynthetic analysis to determine the best, and perhaps cheapest way to obtain the target. On the other hand, combinatorial chemists primarily considered forward synthesis strategies that are founded and in which building blocks are commercially available or indeed worth

UNDER PEER REVIEW

synthesizing. In the same manner, chemical information systems that can be quickly accessed via updated databases of inventory and commercially available reagents are invaluable tools in reagent acquisition by the combinatorial chemists. As combichem matured from solid-phase synthesis to solid-supported synthesis, new synthetic strategies and techniques also evolved. Some of these are now well integrated into the drug design process including microwave synthesis [5, 11, 36], fluorous synthesis [21], click chemistry [37] and flow reactors [19]. As with traditional drug design, combichem relies on organic synthesis methodologies and exploits automation and miniaturization to synthesize large libraries of compounds, which can accelerate the drug discovery process. The combinatorial approach is often systematic and repetitive, using sets of commercially available chemical reagents to form a diverse set of molecular entities. The combinatorial approach is very powerful in early stage discovery and allows HTS to take place, combining rapid synthesis of chemical compounds to be screened using both enzymatic and cellular assays for evaluation. The quick turnaround of data allows a flow of information, which enables second and third generation of compounds to be generated in rapid fashion. Combichem mostly concerns "parallel" synthesis and "split and-mix" synthesis as illustrated in (Figure 3).

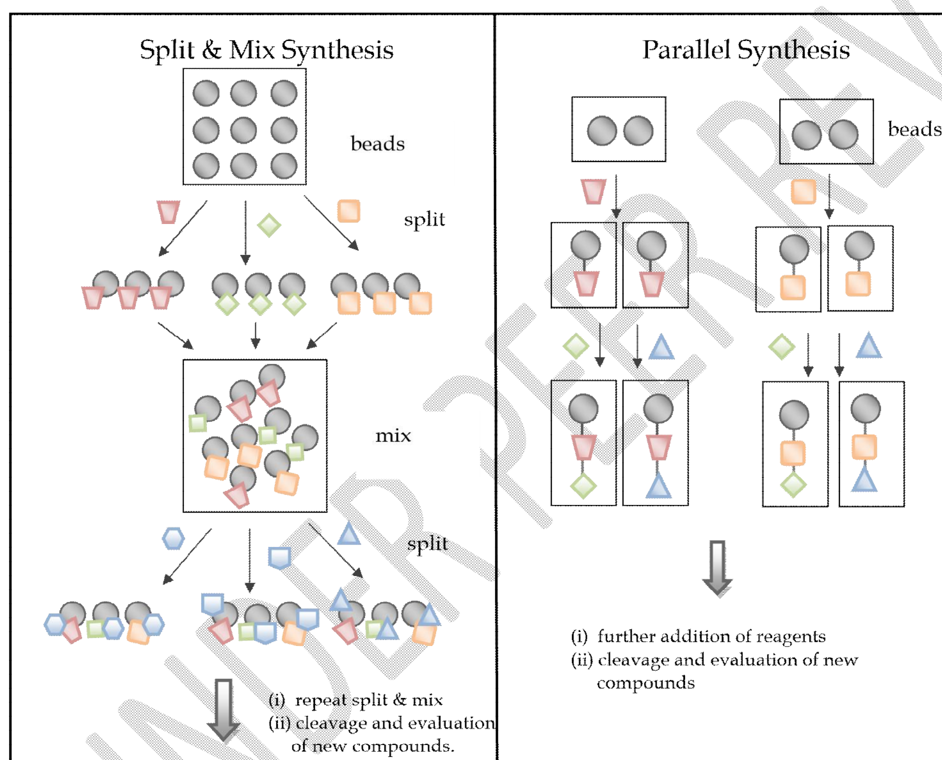

Fig. 3. Combinatorial chemistry approaches. The parallel synthesis is generally used to generate larger quantities of a small number of compounds and split and mix to generate smaller quantities of a larger number of molecules [15].

UNDER PEER REVIEW

Discovery of small molecules for testing of biological pathways and potential new targets

The human genome mapping developed as a better and improved understanding of both pathological causes and function of biological targets and the development of HTS technologies was intended to result in a higher number of new chemical entities (NCEs or for medicinal use. Unfortunately, for some reason this has not been the case due to a number of reasons. Computational molecular modelling has provided scientists with an insight into biochemical events at the molecular level [37]. An understanding of the binding process of small molecules to many macromolecules such as DNA is well understood, however the

same cannot be said about other targets. Many investigations are still to be done probably due to the lack of interest or belief that so-called "undruggable" proteins can be successfully targeted. It has been estimated that only 10–14% of the proteins encoded in the human genome are 'druggable' using existing 'drug-like' molecules [6, 38]. However, given that the *chemical space*, the complete set of all possible small molecules, has been calculated to comprise 10^{30}–10^{200} structures, depending on the parameters used [39], there are large number of yet uncovered chemical structures. Considering the limitations of chemical libraries in addressing challenging targets, it is important to recognize that the vast majority of accessible libraries of small molecules are based on existing drugs [40]. Drugging targets that are recognized as targets and exercising principles such as Lipinski's "rule of five" that have yielded success in the past is safe territory, so it is perfectly understandable that research can be continued in that line

Chemical genetics

Modern genetics began with the applied and theoretical work of the nature of inheritance in plants by German-Czech scientist Gregor Mendel in the mid-19th century. In comparison, the science in chemical genetics is only a couple of decades old, but has become popular in recent years. Chemical genetics has very much its origin in classical genetics and adapted most of the methods and terminology already established [41]. Genetic knockouts have been the key concept to illustrate biological pathways and causal agents of pathological diseases. Now, in addition to chemical genetics, the fields of chemical biology and related modern fields are enabling small molecules to be discovered and developed and used as chemical 'knockdowns' [12, 42]. Chemical genetic principle underlies most of the experimental sciences and explains why our depth of understanding of biological systems has been largely determined by the availability of tools that can be used to disrupt them [43]. In order to close the genotype-phenotype gap, biological research has to go beyond genomics, proteomics, and dissection of biological systems into their different constituents [44].

Protein function is regulated in complex networks with other biomacromolecules, small molecules and supramolecular structures like membranes whereas, genetic manipulation results in a permanent alteration of the native structure of the network. Chemical perturbations with small molecule modulators of protein function providing temporal control using dose-response explorations without fundamentally transforming the biological network [9, 45]. It is very attractive to use small molecules to perturb a biological system because of their dynamic nature, which offers many advantages such as: (i) ability to target a single domain of a multidomain protein, (ii) allows precise temporal control that is critical for rapidly acting processes, (iii) can target orthologous or paralogous proteins, enabling comparisons between species or redundant functions, and (iv) do not directly alter the concentrations of a targeted protein, thus avoiding indirect effects on multiprotein complexes [3, 34, 46]

To elevate the complexity of the test system to reflect for example upon a diseased state of a cell combination chemical genetics (CCG) can be employed. CCG can be defined as the systematic testing of multiple perturbations involving chemical probes and can include either chemical combinations or mixed chemical and genetic perturbations. [47]. Classical and chemical genetics (Figure 4) are generally divided into *forward* screens, in which uncharacterized perturbers are tested against a selected phenotype to detect genes associated with that phenotype, and *reverse* studies, in which a specific gene or protein is modulated and multiple phenotypes are monitored to determine the effects of that specific target [15, 4
28, 48]. Studies involving combined perturbations can be similarly classified with the mechanistic focus shifted from individual targets to interactions between them [49]

.

UNDER PREVIEW

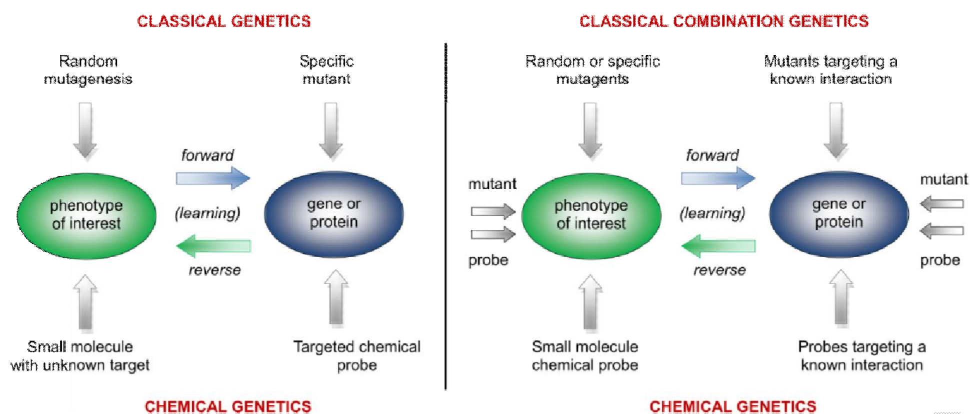

Fig. 4. Combined perturber studied in the context of forward and reverse genetics. The essence of classical and chemical genetics is to explore the function of individual genes or proteins. In combination chemical genetics, the focus of investigations shifts from individual targets to interactions between them or conditional target dependencies, and the perturbations are applied as combinations

Chemical Biology.

Chemical biology has made significant contribution in drug discovery and offering new technologies that can progress the understanding of human health. Given the temporal control offered by small molecules and the ability to use combinations of small molecule modulators, chemical biology tries to complement the use of pure gbiological analysis to study a wide range of biological systems [49]. Chemical biology aims to answer questions in complex test systems and may provide the field with commercial chemical probes that can be used to probe pathways and elucidate more about biological targets. The discovery of the potent and selective deacetylase inhibitors tubacin and histacin are examples of how powerful chemical genetics can be in combination with computational methods such as principal component analysis [50]. However, good chemical probes for *in vitro* and especially *in vivo* perturbation are not easy to come by as small molecules are generally pleiotropic and they have multiple dose-dependent molecular targets that are often not fully characterized, and may lead to unexpected activities. Obstacles and challenges are similar to those in drug development: small molecules often have inherent problems such as *in vitro* aggregation, poor solubility, difficulty in crossing biological membranes and reactive or toxic functionalities. At present, development of chemical probes for *in vivo* testing may be too ambitious a goal. As a result, evaluation of the effect of chemical 'knockdowns' in clinically relevant tissue should in the near future be in more complex assays that mimic for example malignant tissue.

Target identification

Drugs fail in the clinical development process for two main reasons; the first is that they do not work and the second is that they are not safe. As such, one of the most important steps in developing a new drug is target identification and validation. A target is a broad term which can be applied to a range of biological entities which may include for example proteins, genes and RNA [51]. A good target needs to be efficacious, safe, meet clinical and commercial needs and, above all, be 'druggable'. A 'druggable' target is accessible to the identified drug molecule, be it a small molecule or larger biologicals, and upon binding, elicit a biological response which may be measured both *in vitro* and *in vivo* [7, 52] It is now shown that certain target classes are more amenable to small molecule drug discovery, for example, G-protein-coupled

UNDER PEER REVIEW

receptors (GPCRs), whereas antibodies are good at blocking protein/protein interactions. There are also examples of phenotypes in humans where mutations can nullify or over activate the receptor, for example, the voltage gated sodium channel NaV1.7, both mutations incur a pain phenotype, insensitivity or oversensitivity respectively [53]. An alternative approach is to use phenotypic screening to identify disease relevant target. **Figure 5 illustrate the d**rug discovery process from target ID and validation through to filing of a compound and the approximate timescale for these processes. FDA, Food and Drug Administration; IND, Investigational New Drug; NDA, New Drug Application. The drug discovery time line also shown in figure 6.

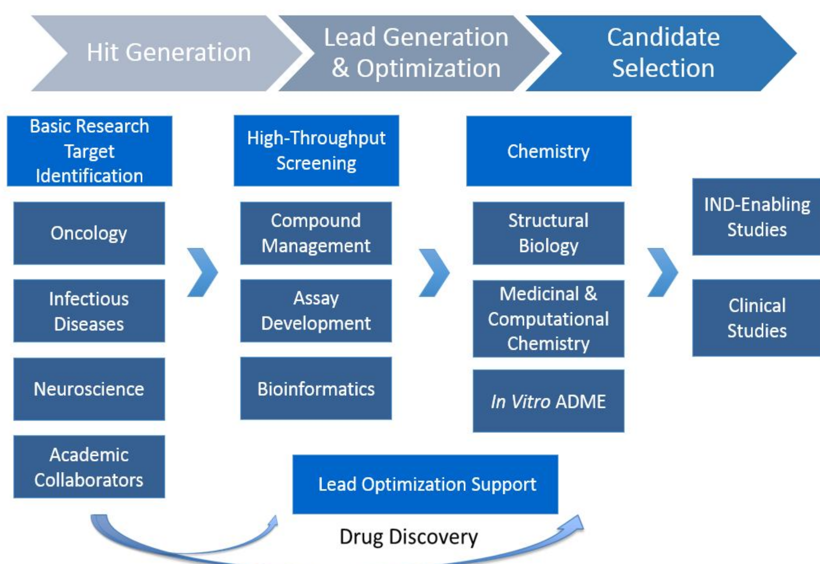

Figure 5: Drug discovery process from hit generation to lead optimization

UNDER PEER REVIEW

Data mining of available biomedical data has led to a significant increase in target identification. In this context, data mining refers to the use of a bioinformatics approach not only to help in identifying but also in selecting and prioritizing potential disease targets [54]. The data can be derived from a variety of sources but include publications and patent information, gene expression data, proteomics data, transgenic phenotyping and compound profiling data. Identification approaches also include examining mRNA/protein levels to determine whether they are expressed in disease and if they are correlated with disease exacerbation or progression. Another powerful approach in target identification is to look for genetic associations, for example, is there a link between a genetic polymorphism and the risk of disease or disease progression or is the polymorphism functional. Clones are individually screened by immunostaining and those that preferentially and strongly stained the malignant cells were chosen. The antigens recognized by those clones were isolated by immunoprecipitation and identified by mass spectroscopy [55].

Target validation

Once drug targets are identified, the target then needs to be fully prosecuted. Validation techniques range from *in vitro* tools through the use of whole animal models, to modulation of a desired target in disease model. Figure 6 gives al illustration of target identification and validation of multifunctional process.

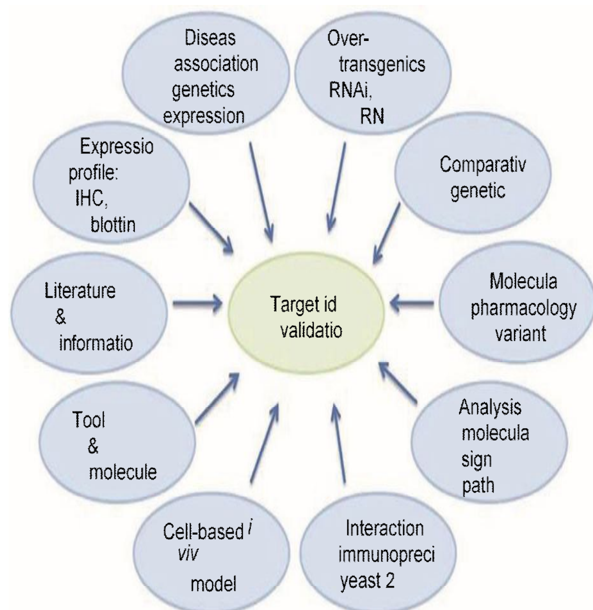

Figure 6: Target identification and validation of a multifunctional process

Antisense technology is a potentially powerful technique which utilizes RNA-like chemically modified oligonucleotides which are designed to be complimentary to a region of a target mRNA molecule [[56]. Binding of the antisense oligonucleotide to the target mRNA prevents binding of the translational machinery thereby blocking synthesis of the encoded protein. A prime example of the power of antisense technology was demonstrated by researchers at Abbott Laboratories who developed antisense probes to the rat P2X3 receptor [14]. When given by intrathecal minipump, to avoid toxicities associated with bolus injection, the phosphorothioate antisense P2X3 oligonucleonucleotides had marked anti-hyperanalgesic activity in the Complete Freund's Adjuvant model, demonstrating an unambiguous role for this receptor in chronic inflammatory states. Interestingly, after administration of the antisense oligonucleonucleotides was discontinued, receptor function and analgesic responses returned. Therefore, in contrast to the gene knockout approach, antisense oligonucleotide effects are reversible and a continued presence of the antisense is required for target protein inhibition [57]. The chemistry associated with creating oligonucleotides has resulted in some cases in molecules with limited bioavailability and pronounced toxicity, making their *in vivo* use problematic [58]. This has been compounded by non-specific actions, problems with controls for these tools and a lack of diversity and variety in selecting appropriate nucleotide probes [57].

UNDER PEER REVIEW

An alternative to gene knockouts is gene knock-ins, where a non-enzymatically functioning protein replaces the endogenous protein. These animals can have a different phenotype to a knockout, for example when the protein has structural as well as enzymatic functions [58], and these mice should predominantly mimic more closely what happens during treatment with drugs, that is, the protein is there but functionally inhibited [45]. More recently, the desire to be able to make tissue restricted and/or inducible knockouts has grown. Although these approaches are technically challenging, the most obvious reason for this is the need to overcome embryonic lethality of the homozygous null animals. Other reasons include avoidance of compensatory mechanisms due to chronic absence of a gene-encoded function and avoidance of developmental phenotypes. However, the use of transgenic animals is expensive and time-consuming. So, in order to circumvent some of these issues, the use of small interfering RNA (siRNA) s become increasingly popular for target validation.

Double-stranded RNA (dsRNA) specific to the gene to be silenced is introduced into a cell or organism, where it is recognized as exogenous genetic material and activates the RNAi pathway. The ribonuclease protein Dicer is activated which binds and cleaves dsRNAs to produce double-stranded fragments of 21–25 base pairs with a few unpaired overhang bases on each end [59]. These short double-stranded fragments are called siRNAs. These siRNAs are then separated into single strands and integrated into an active RNA-induced silencing complex (RISC). After integration into the RISC, siRNAs base pair to their target mRNA and induce cleavage of the mRNA, thereby preventing it from being used as a translation template [60]. The RNAi technology still has the major problem of delivery to the target cell, but many viral and non-viral delivery systems are currently under investigation [61].

Monoclonal antibodies are also an excellent target validation tool as they interact with a larger region of the target molecule surface, allowing for better discrimination between even closely related targets and often providing higher affinity. In contrast, small molecules are disadvantaged by the need to interact with the often more conserved active site of a target, while antibodies can be selected to bind to unique epitopes [62]. This exquisite specificity is the basis for their lack of non-mechanistic (or 'off-target') toxicity, which is a major advantage over small-molecule drugs. However, antibodies cannot cross cell membranes restricting the target class mainly to cell surface and secreted proteins. One impressive example of the efficacy of a mAb *in vivo* is that of the function neutralizing anti-TrkA antibody MNAC13, which has been shown to reduce both neuropathic pain and inflammatory hypersensitivity [63], thereby implicating NGF in the initiation and maintenance of chronic pain. Finally, the classic target validation tool is the small bioactive molecule that interacts with and functionally modulates effector proteins.

The hit discovery process

Following the process of target validation, it is during the hit identification and lead discovery phase of the drug discovery process that compound screening assays are developed. A 'hit' molecule can vary in meaning to different researchers but a hit is a compound which has the desired activity in a compound screen and whose activity is confirmed upon retesting [64]. A variety of screening paradigms exist to identify hit molecules as shown in figure 7. High throughput screening (HTS) involves the screening of the entire compound library directly against the drug target or in a more complex assay system, such as a cell-

UNDER PEER REVIEW

based assay, whose activity is dependent upon the target but which would also require secondary assays to confirm the site of action of compounds [17, 65]. This screening paradigm involves the use of complex laboratory automation but assumes no prior knowledge of the nature of the chemotype likely to have activity at the target protein. Focused or knowledge-based screening involves selecting from the chemical library smaller subsets of molecules that are likely to have activity at the target protein based on knowledge of the target protein and literature or patent precedents for the chemical classes likely to have activity at the drug target [66]. This type of knowledge has given rise, more recently, to early discovery paradigms using pharmacophores and molecular modelling to conduct virtual screens of compound databases [67]. Fragment screening involves the generation of very small molecular weight compound libraries which are screened at high concentrations and is typically accompanied by the generation of protein structures to enable compound progression [68].

Physiological screening. This is a tissue-based approach and looks for a method to identify molecules of interest. The output of a compound screen is typically termed a hit molecule, which has been demonstrated to have specific activity at the target protein [69]. Screening hits form the basis of a lead optimization chemistry programme to increase potency of the chemical series at the primary drug target protein. During the lead discovery, phase molecules are also screened in cell-based assays predictive of the disease state and in animal models of disease to characterize both the efficacy of the compound and its likely safety profile

Assay development

In the recombinant era the majority of assays in use within the industry rely upon the creation of stable mammalian cell lines over-expressing the target of interest or upon the overexpression and purification of recombinant protein to establish the biochemical assays, although in recent years there has been an increase in the number of reports describing the use of primary cell systems for compound screening [70]. Generally, cell-based assays have been applied to target classes such as membrane receptors, ion channels and nuclear receptors and typically generate a functional read-out as a consequence of compound activity [71]. In contrast, biochemical assays, which have been applied to both receptor and enzyme targets, often simply measure the affinity of the test compound for the target protein [11]. A plethora of assay formats have been enabled to support compound screening and the choice of the assay format is dependent upon the biology of the drug target protein, the equipment infrastructure in the host laboratory, the experience of the scientists in that laboratory, whether an inhibitor or activator molecule is sought and the scale of the compound screen [72].

UNDER PEER REVIEW

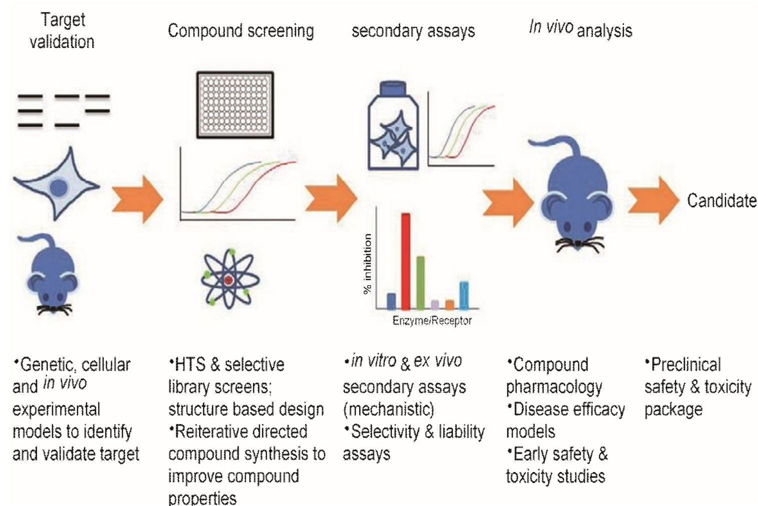

Figure 7: Summary overview of drug discovery High throughput screening (HTS) Assays

UNDER PEER REVIEW

High through put screening in drug discovery needs to demonstrate the following characteristics;

1. Pharmacological relevance of the assay. If available, studies should be performed using known ligands with activity at the target under study, to determine if the assay pharmacology is predictive of the disease state and to show that the assay is capable of identifying compounds with the desired potency and mechanism of action.
2. Reproducibility of the assay. Within a compound screening environment, it is a requirement that the assay is reproducible across assay plates, across screen days and, within a programme that may run for several years, across the duration of the entire drug discovery programme.
3. Assay costs. Compound screening assays are typically performed in microtitre plates. Within academia or for relatively small numbers of compounds assays are typically formatted in 96-well or 384-well microtitre plates whereas in industry or in HTS applications assays are formatted in 384-well or 1536-well microtire plates in assay volumes as small as a few microlitires [73]. In each case assay reagents and assay volumes are selected to minimize the costs of the assay.
4. Assay quality. Assay quality is typically determined according to the Z′ factor [74]. This is a statistical parameter that in addition to considering the signal window in the assay also considers the variance around both the high and low signals in the assay. The Z factor has become the industry standard means of measuring assay quality on a plate basis. The Z factor has a range of 0 to 1; an assay with a Z factor of greater than 0.4 is considered appropriately robust for compound screening although many groups prefer to work with assays with a Z factor of greater than 0.6. In addition to the Z factor assay quality is also monitored through the inclusion of pharmacological controls within each assay. Assays are deemed acceptable if the pharmacology of the standard compound(s) falls within predefined limits [8, 75].

Assay quality is affected by many factors. Generally, high-quality assays are created through implementing simple assay protocols with few steps, minimizing wash steps or plate to plate reagent transfers within the assay, through the use of stable reagents and biologicals, and through ensuring that all the instrumentation used to perform the assay is performing optimally. This is typically achieved through developing quality control practices for all items of laboratory [76].

5. Effects of compounds in the assay. Chemical libraries are typically stored in organic solvents such as ethanol or dimethyl sulphoxide (DMSO). Thus, assays need to be configured in a way that are not sensitive to the concentrations of solvents used in the assay. Typically, cell-based assays are intolerant to solvent concentrations of greater than 1% DMSO whereas biochemical assays can be performed in solvent concentrations of upto10%DMSO [77]. Studies are also performed to establish the false negative and false positive hit rates in the assay. When developing any HTS assay, which can involve the screening of several million molecules over several weeks, it is best practice to screen training sets of compounds to verify that the assay is performing acceptably.

Defining a hit series

Compound libraries have been assembled to contain small molecular weight molecules that obey chemical parameters such as the Lipinski Rule of Five [77], and more often have molecular weights of less than 400 and clogP (a measure of lipophilicity which affects absorption into the body) of less than 4. Molecules with these features have been termed 'drug-like', in recognition of the fact that the majority of clinically marketed drugs have a molecular weight of less than 350 and a cLogP of less than 3. It is critically important to initiate a drug discovery programme with a small simple molecule as lead optimization, to improve potency and selectivity, typically involves an increase in molecular weight which in turn can lead to safety and tolerability issues [78]. Once a number of hits have been obtained from virtual screening or HTS, the first role for the drug discovery team is to try to define which compounds are the best to work on. This triaging process is essential as, from a large library, a team will likely be left with many possible hits which they will need to reduce, confirm and cluster into series. There are several steps to achieving this. First, although this is less of a problem as the quality of libraries have improved, compounds that are known by the library curators to be frequent hitters in HTS campaigns need to be removed from further consideration [79]. Secondly, a number of computational chemistry algorithms have been developed to group hits based on structural similarity to ensure that a broad spectrum of chemical classes are represented on the list of compounds taken forward. Analysis of the compound hit list using these algorithms allows the selection of hits for progression based on chemical cluster, potency and factors such as ligand efficiency which gives an idea of how well a compound binds for its size (log potency divided by number of 'heavy atoms' i.e. non-hydrogen atoms, in a molecule) [80].

The next phase in the initial refinement process is to generate dose–response curves in the primary assay for each hit, preferably with a fresh sample of the compound. Showing normal competitive behaviour in hits is important. Compounds which give an all or nothing response are not acting in a reversible manner and indeed may not be binding to the target protein at all, with the activity at high concentrations arising from an interaction between the sample and another component of the assay system [81]. Reversible compounds are favoured because their effects can be more easily 'washed-out' following drug withdrawal, an important consideration when using in patients. Obtaining a dose–response curve allows the generation of a half

UNDER PEER REVIEW

maximal inhibitory concentration which is used to compare of the potencies of candidate compounds [45, 82]. Sourcing and using fresh samples of compounds for this exercise is highly desirable. Nearly all HTS libraries are stored as frozen DMSO solutions with the result that, after some time, the compound can become degraded or modified. Virtually anyone who has worked with libraries of this type has got anecdotes about how potent activity has disappeared when the compound was resynthesized and used in re-testing, although occasionally identification of potent impurities has allowed progress to be made.

With reliable dose–response curves generated in the primary assay for the target, the stage is set to examine the surviving hits in a secondary assay, if one is available, for the target of choice. This need not be an assay in a high throughput format but will involve looking at the effect of the compounds in a functional response, for example in a second messenger assay or in a tissue-or cell-based bioassay. Activity in this setting will give reassurance that compounds are able to modulate more intact systems rather than simply interacting with the isolated and often engineered protein used in the primary assay [83]. Throughout the confirmation process, medicinal chemists would be looking to cluster compounds into groups which could form the basis of lead series. As part of this process, consideration will be given to the properties of each cluster such as whether there is an identifiable structure–activity relationship (SAR) evolving over a number of compounds, that is, identification of a group of compounds which have some section or chemical motif in common and the addition of different chemical groups to this core structure results in different potencies. Issues of chemical synthesis would also be examined. Thus, ease of preparation, potential amenability to parallel synthesis and the ability to generate diversity from late-stage intermediates would be assessed [84].

Hit-to-lead phase

The aim of this stage of the work is to refine each hit series to try to produce more potent and selective compounds which possess PK properties adequate to examine their efficacy in any *in vivo* models that are available [85]. Typically, the work now consists of intensive SAR investigations around each core compound structure, with measurements being made to establish the magnitude of activity and selectivity of each compound. This needs to be carried out systematically and, where structural information about the target is known, structure-based drug design techniques using molecular modelling and methodologies such as X-ray crystallography and NMR can be applied to develop the SAR faster and in a more focused way. This type of activity will also often give rise to the discovery of new binding sites on the target proteins [13, 86].

A screening cascade at this time would generally consist of a relatively high throughput assay establishing the activity of each molecule on the molecular target, together with assays in the same format for sites where selectivity might be known, or expected to be, an issue. A compound meeting basic criterion at this stage would be escalated into a further bank of assays. These should include higher order functional investigations against the molecular target and also whether the compounds were active in primary assays in different species [87]. The key in vitro assays in early drug discovery is illustrated in table 1.

Table 1: Key *in vitro* assays in early drug discovery [9]

UNDER PEER REVIEW

Assays	Target value	Outcomes
CYP450-inhibition	➢ 10 mM	Main drug metabolic enzyme whose inhibition can cause toxicity
Caco-2 permeability P_{app}	➢ 1.¥ $10^{6}cm^{-1}$ (asymmetry <2)	Caco-2 colon carcinoma cell line-used for estimation of permeability across intestinal epithelium. Important for drug absorption from gut
Aqueous solubility	➢ 100 mM	Important for *in vitro* assays and *in vivo* drug delivery
Hep G2 hepatotoxicity	No effect at 50 ¥ IC_{50} or EC_{50}	Human HepG2 cells can act as a surrogate for effects of toxicity on human liver, an important cause of drug failure in the clinic
Log $D_{7.4}$	0-3 (for BBB penetration ca.2)	A measure of lipophilicity and movements across membranes
Microsomal stability Cl_{int}	< 30 $mLmin^{-1}$ mg^{-1} protein	Liver microsomes contain membrane-bound drug metablizing enzyme. This assay measures compound clearance and give an idea of how fast it will be cleared out *in vivo*
Cytotoxicity in suitable cell line	No effect at 50 ¥ IC_{50} or EC_{50}	Reduce the likelyhood of cellular toxicity *in vivo*
MDR1-MDCK permeability Papp	>1¥ 106 cm-1 (as	MDCK cells transfected with the MDR1 gene, which encodes the efflux protein P glycoprotein (P-gp). An important efflux transporter in many tissues including intestine, kidney and brain, P-gp can be used to predict intestinal and brain permeability

IC_{50}, half maximal inhibitory concentration

Solubility and permeability assessments are crucial in ruling in or out the potential of a compound to be a drug, that is, drug substance often needs access to a patient's circulation and therefore may be injected or more generally has to be adsorbed in the digestive system [88]. Deficiency in one or another parameter in a molecule can sometimes be put right. For example, formulation strategies can be used to design a tablet such that it dissolves in a particular region of the gut at a pH in which the compound is more soluble. A compound that lacks both these properties is very unlikely to become a drug no matter how potent it is in the primary screening assay [33, 89]. Microsomal stability is a useful measure of the ability of *in vivo* metabolizing enzymes to modify and then remove a compound. Hepatocytes are sometimes used in this sort of study instead and these will give more extensive results but are not used routinely as they need to be prepared freshly on a regular basis. CYP450 inhibition is examined as, among other things, it is an important predictor of whether a new compound might have an influence on the metabolism of an existing drug with which it may be co-administered [90]. If one or more of these properties is less than ideal, then it

UNDER PEER REVIEW

might be necessary to screen many more compounds specifically for those properties. Each programme will end up subtly different in this regard.

Key compounds which are beginning to meet the target potency and selectivity, as well as most of the physicochemical and ADME targets, should be assessed for PK in rats. Here one would normally be aiming for a half-life of >60 min when the compound is administered intravenously and a fraction in excess of 20% absorbed following oral dosing although sometimes, different targets require very different PK profiles [16, 91]. In large pharma with inhouse drug metabolism pharmacokinetics (DMPK) departments numerous compounds might be profiled while in academic environments there may be funds for only a predefined number of these expensive investigations [92] As the receptor antagonist programme, advanced through the hit-to-lead phase, a number of compounds were prepared which had potency in the nanomolar range and a benign selectivity profile except for some potency at the hERG channel, a potassium voltage-gated ion channel important for cardiac function and inhibition at which can cause cardiac liability. Ideally for hERG study the aim is for an activity over 30 uM or at least a 1000-fold selectivity for the target [93]. A number of these compounds were examined in PK studies and were found to have a reasonable half-life following intravenous dosing but poor plasma levels were noted when the compound was given orally to rats. It was felt that some of these compounds, representing the end of the hit-to lead phase of the project were, although not likely themselves to be progressed, capable of answering questions in disease models.

UNDER PEER REVIEW

Where do potential lead come from.

From various HTS assay leads can be selected. However, there are many sources of lead compounds as illustrated in figure 8, from acquisition compound, natural sources, newly synthesized or endogenous ligand [94].

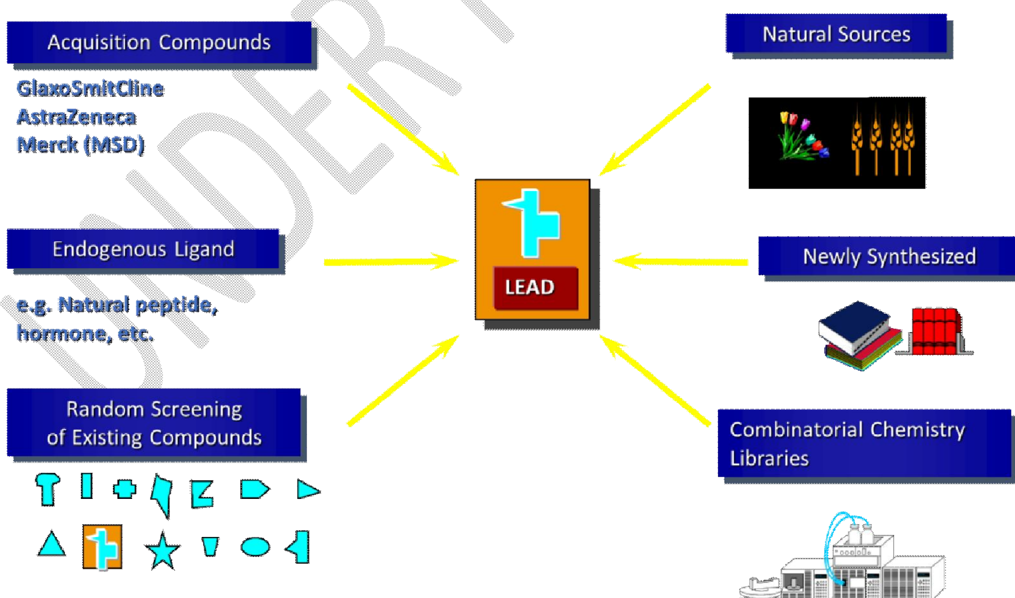

Figure 8: Sources of potential lead compounds [14]

Pre-screening of lead compounds using DEREK software.

Early prescreening of lead compound to detect early attrition of compound is possible through the use of DEREK software, (**D**eductive **E**stimation of **R**isk based on **E**xisting **K**nowledge). DEREK is an In-Silico screen which indicates whether a specific toxic response may occur. It does not provide a quantitative estimation of the prediction [42, 95]. DEREK has several rule bases, consisting of descriptions of molecular substructures (structural alerts) that have been associated with toxic end points (mutagenicity, carcinogenicity, skin irritation. Since substructures can exist in a variety of molecular contexts, the rules are not chemical specific but serve as broad generalizations with regard to the chemical structure (acid or halogen containing molecule, alkylating agent, Chemical substructure implicated in some toxic effect (Pharmacophore). For example, substructures like epoxides known to be mutagenic. Linked to some scientific justification: Derived from a knowledge base extracted from experimental data.

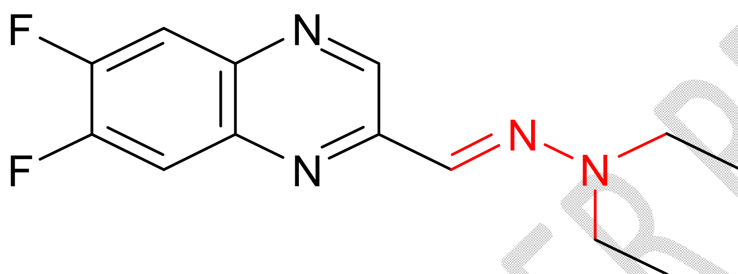

Figure 9. Chemical substructure implicated in some toxic effect- PHARMACOPHORE

Lead optimization phase

The aim of this final drug discovery phase is to maintain favourable properties in lead compounds while improving on deficiencies in the lead structure. Continuing with example above, the aim of the programme was now to modify the structure to minimize hERG liability and to improve the absorption of the compound. Thus, more regular checks of hERG affinity and CACO2 permeation were undertaken and compounds were soon available which maintained their potency and selectivity at the principal target but which had a much reduced hERG affinity and a better apparent permeation than initial lead compounds [96]. When examined for PK properties in rat one of these compounds, with 8 nM affinity at the receptor of interest, had an oral bioavailability of over 40% in rats and about 80% in dogs.

Compounds at this stage may be deemed to have met the initial goals of the lead optimization phase and are ready for final characterization before being declared as preclinical candidates. Discovery work does not cease at this stage. The team has to continue to explore synthetically in order to produce potential back up molecules, in case the compound undergoing further preclinical or clinical characterization fails and, more strategically, to look for follow-up series [97]. The stage at which the various elements that constitute further characterization are carried out will vary from company to company and parts of this process may be incorporated into the lead optimization phase. However, in general molecules need to be examined in models of genotoxicity such as the Ames test and in *in vivo* models of general behaviour such

UNDER PEER REVIEW

as the Irwin's test. High-dose pharmacology, PK/PD studies, dose linearity and repeat dosing PK looking for drug-induced metabolism and metabolic profiling all need to be carried out by the end of this stage [98, 99]. Consideration also needs to be given to chemical stability issues and salt selection for the putative drug substance. All the information gathered about the molecule at this stage will allow for the preparation of a target candidate profile which with together with toxicological and chemical manufacture and control considerations will form the basis of a regulatory submission to allow human administration to begin.

The process from hit generation to preclinical candidate selection often takes a long time and cannot in any way be considered a routine activity. There are rarely any short cuts and significant, intellectual input is required from scientists from a variety of disciplines and backgrounds. The quality of the hit-to-lead starting point and the expertise of the available team are the key determinants of a successful outcome of this phase of work. Typically, within industry for each project 200 000 to >10^6 compounds might be screened initially and during the following hit-to-lead and lead optimization programmes 100's of compounds are screened to hone down to one or two candidate molecules, usually from different chemical series [25, 52]. In academia screens are more likely to be of a focused nature due to the high cost of an extensive HTS or compounds are derived from a structure-based approach. Only 10% of small molecule projects within industry might make the transition to candidate, failing at multiple stages [16]. These can include the (i) inability to configure a reliable assay; (ii) no developable hits obtained from the HTS; (iii) compounds do not behave as desired in secondary or native tissue assays; (iv) compounds are toxic *in vitro* or *in vivo*; (v) compounds have undesirable side effects which cannot be easily screened out or separated from the mode of action of the target; (vi) inability to obtain a good PK or PD profile in line with the dosing regimen required in man, for example, if require a once a day tablet then need the compound to have a half-life *in vivo* suitable to achieve this; and (vii) inability to cross the blood brain barrier for compounds whose target lies within the central nervous system [65, 78]. The attrition rate for protein therapeutics, once the target has been identified, is much lower due to less off target selectivity and prior experience of PK of some proteins, for example, antibodies.

UNDER PEER REVIEW

Conclusion

Although relatively less costly than many processes carried out later on in the drug development and clinical phases, preclinical activity is sufficiently high risk and remote from financial return to often make funding it a problem. Ensuring transparency of the cost of each stage/assay within large pharma may help reduce some of their costs and there are some movements towards this as companies instigate a 'biotech' mentality and accountability for costs.

Once a candidate is selected, the attrition rate of compounds entering the clinical phase is also high, again only one in 10 candidates reaching the market but at this stage the financial consequences of failure are much higher. There has been considerable debate in industry as to how to improve the success rate, by 'failing fast and cheap'. Once a candidate reaches the clinical stage, it can become increasingly difficult to kill the project, as at this stage the project has become public knowledge and thus termination can influence confidence in the company and shareholder value. Carrying out more studies prior to clinical development such as improved toxicology screens (using failed drugs to inform these assays), establishing predictive translational models based on a thorough disease understanding and identifying biomarkers may help in

this endeavour. It is particularly in these later two areas where academic-industry partnerships could really add value preclinically and eventually help bring more effective drugs to patients.

REFERENCES

1. Sehgal A. Drug discovery and development using chemical genomics. .Curr Opin Drug Discov Devel. 2002 Jul;5(4):526-31. PMID: 12197310.

2. Nassar AE, Lopez-Anaya A. Strategies for dealing with reactive intermediates in drug discovery and development. .Curr Opin Drug Discov Devel. 2004 Jan;7(1):126-36. PMID: 14982156.
3. Lowe JA, Jones P, Wilson DM. The importance of target validation in drug discovery and development..Curr Opin Drug Discov Devel. 2009 Sep;12(5):581-4.

4. Sumner-Smith M Beginning to manage drug discovery and development knowledge. .Curr Opin Drug Discov Devel. 2001 May;4(3):319-24.

5. Emerging applications of network and pathway analysis in drug discovery and development.
6. Ganter B, Giroux CN.Curr Opin Drug Discov Devel. 2008 Jan;11(1):86-94. PMID: 18175271
7. .
8. Hutzler M, Messing DM, Wienkers LC. .Predicting drug-drug interactions in drug discovery: where are we now and where are we going? Curr Opin Drug Discov Devel. 2005 Jan;8(1):51-8.PMID: 15679172 .

9. Roberts SA. Drug metabolism and pharmacokinetics in drug discovery.; Curr Opin Drug Discov Devel. 2003 Jan;6(1):66-80.PMID:

10. Boran AD, Iyengar R. Systems approaches to polypharmacology and drug discovery. .Curr Opin Drug Discov Devel. 2010 May;13(3):297-309.PMID: 20443163

11. Watkins SM Lipomic profiling in drug discovery, development and clinical trial evaluation. .Curr Opin Drug Discov Devel. 2004 Jan;7(1):112-7.PMID: 14982154

12. Gasparini F, Bilbe G, Gomez-Mancilla B, Spooren W mGluR5 antagonists: discovery, characterization and drug development. Curr Opin Drug Discov Devel. 2008 Sep;11(5):655-65.PMID: 18729017 .

13. Hughes JP, Rees S, Kalindjian SB, Philpott KL.Principles of early drug discovery. Br J Pharmacol. 2011 Mar;162(6):1239-49. doi: 10.1111/j.1476-5381.2010.01127.x.PMID: 21091654.

UNDER PEER REVIEW

14. Fokunang CN, Tembe-Fokunang EA, Galega S, Forsuh FF, et al. In: Advances in Medicines and Biology, Leon V. Berhardt eds. Ethical implications and regulatory perspectives in the conduct of research in herbal medicine in cameroon. Advances in Medicines and Biology, 2019; vol 138.Nova Science publishers Inc, USA, 1-39p. ISBN:978-1-53615-066-7 EBOOKS, 270p..

15. Fokunang CN, Tembe-Fokunang EA, Awah P,, Djuidje Ngounoue M et al. The role of ethics in public health clinical research.:in alfonso j rodriguez-morsles (eds): current topics in public health. intech publishers: isnb.978-953-1121-14 intech may 2013 chapter 27. pp662-683.

16. Fokunang CN, Fonmboh DJ, Tembe EF, Mahop B, et al. Doping in Competitive Sports: stakes and challenges for resource limited countries. 2020. European Journal of pharmaceutical and medical research, Impact factor 6.22; ejpmr, 2020,7(12), 29-40; ISSN 2394-3211 EJPMR; www.ejpmr.com.

17. Tembe EF, Fonmboh DJ, Ngono MR, Banin AN, Fokunang LB, et al. Pharmacovigilance of Natural Herbal Medicines Research for Efficacy, Safety and Quality Assurance of Phytomedicine Products; Journal of Complementary and Alternative Medical Research 12(1): 21-37, 2020; ISSN: 2456-6276; DOI: 10.9734/JOCAMR/2020/v12i130198.

18. Dobgima JF, Ejoh RA Tembe EF, Bayaga H, Teke NG, et al. An Overview of Methods of Extraction, Isolation and Characterization of Natural Medicinal Plant Products in Improved Traditional Medicine Research Asian Journal of Research in Medical and Pharmaceutical Sciences 2020, Accepted Article no.AJRIMPS.61961 ISSN: 2457-0745

19. Mogue I, Njinkio LBN, Tembe Fokunang EA, Mbacham WF, Fokunang CN. Évaluation de la Toxicité Aigüe des Extraits Aqueux de Combretum micranthum (Combrétacée) et Essai de Mise en Forme Galénique d'un Phytomédicament pour le Traitement de la Fièvre Typhoïde Health Sciences and Disease; 2020; 21 (8): 1-15

20. Kamgain T, Tcha FG, Nnanga NE, Fokunang CN. From plant to drug: Galenic preformulation based on Phyllanthus amarus ;PKFokam Journal of Applied Science & Technology 2020,; . 2,(1,) : 1 – 14/ ISSN: 2707-2843 (Print) 2707-2851(online) www.pkfokam-jast.org

21. Glassman PM, Balthasar JP. Physiologically-based modeling of monoclonal antibody pharmacokinetics in drug discovery and development. Drug Metab Pharmacokinet. 2019 Feb;34(1):3-13. doi: 10.1016/j.dmpk.2018.11.002.

22. Mohs RC, Greig NH.Alzheimers Dement (N Y). Drug discovery and development: Role of basic biological research.2017 Nov 11;3(4):651-657. doi: 10.1016/j.trci.2017.10.005
23. Kaur T, Madgulkar A, Bhalekar M, Asgaonkar K. Molecular Docking in Formulation and Development. .Curr Drug Discov Technol. 2019;16(1):30-39. doi: 10.2174/1570163815666180219112421.PMID: 29468973
24. Renaud JP, Chung CW, Danielson UH, Egner U, Hennig M, Hubbard RE, Nar H opportunities. Nat Rev Drug Discov. 2016 Oct;15(10):679-98. doi: 10.1038/nrd.2016.123.
25. Kim H, Kim E, Lee I, Bae B, Park M, Nam H. Artificial Intelligence in Drug Discovery: A Comprehensive Review of Data-driven and Machine Learning Approaches.

UNDER PEER REVIEW

26. Biotechnol Bioprocess Eng. 2020;25(6):895-930. doi: 10.1007/s12257-020-0049-y.
27. Fryszkowska A, Devine PN. Biocatalysis in drug discovery and development. .Curr Opin Chem Biol. 2020 Apr;55:151-160. doi: 10.1016/j.cbpa.2020.01.012.

28. Pedreira JGB, Franco LS, Barreiro EJ. Chemical Intuition in Drug Design and Discovery. .Curr Top Med Chem. 2019;19(19):1679-1693. doi: 10.2174/1568026619666190620144142.PMID: 31258088

29. Hasselgren C, Myatt GJ. Computational Toxicology and Drug Discovery. .Methods Mol Biol. 2018;1800:233-244. doi: 10.1007/978-1-4939-7899-1_11.PMID: 29934896 .

30. Taylor, David. The Pharmaceutical Industry and the Future of Drug Development. Issues in Environmental Science and Technology. Royal Society of Chemistry 2015; 1–33. doi:10.1039/9781782622345-00001. ISBN 978-1-78262-189-8.
31. Paul, Steven M.; Mytelka, Daniel S.; Dunwiddie, Christopher T.; Persinger, Charles C.; Munos, Bernard H.; Lindborg, Stacy R.; Schacht, Aaron L. "How to improve R&D productivity: The pharmaceutical industry's grand challenge". Nature Reviews Drug Discovery, 2010' . 9 (3): 203–14. doi:10.1038/nrd3078. PMID 20168317. S2CID 1299234.
32. Prasad, Vinay; Mailankody, Sham. Research and development spending to bring a single cancer drug to market and revenues after approval JAMA Internal Medicine, 2017;. 177 (11): 1569–1575. doi:10.1001/jamainternmed.2017.3601. ISSN 2168-6106. PMC 5710275. PMID 28892524.
33. Moore, Thomas J.; Zhang, Hanzhe; Anderson, Gerard; Alexander, G. Caleb. Estimated costs of pivotal trials for novel therapeutic agents approved by the US Food and Drug Administration, 2015-2016". JAMA Internal Medicine 2018;. 178 (11): 1451–1457. doi:10.1001/jamainternmed.2018.3931. ISSN 2168-6106. PMC 6248200. PMID 30264133.
34. Sertkaya, A; Wong, HH.; Jessup, A; Beleche, T. Key cost drivers of pharmaceutical clinical trials in the United States". Clinical Trials; 2016; 13 (2): 117–26. doi:10.1177/1740774515625964. PMID 26908540. S2CID 24308679.
35. Maxmen A. .Busting the billion-dollar myth: how to slash the cost of drug development. Nature, 2016;. 536 (7617): 388–90. Bibcode:2016Natur.536..388M. doi:10.1038/536388a. PMID 27558048.
36. Boris Bogdan and Ralph Villiger, "Valuation in Life Sciences. A Practical Guide", 2008, 2nd edition, Springer Verlag.
37. Nielsen, Nicolaj Hoejer "Financial valuation methods for biotechnology", 2010. "Archived copy" (PDF). Archived from the original (PDF) 3010,on 2012-03-05. Retrieved 2014-11-25.
38. Wang Y. .Extracting Knowledge from Failed Development Programmes". Pharm Med, 2012;. 26 (2): 91–96. doi:10.1007/BF03256897. S2CID 17171991.
39. Herschel, M. (2012). "Portfolio Decisions in Early Development: Don't Throw Out the Baby with the Bathwater". Pharm Med 2012. 26 (2): 77–84. doi:10.1007/BF03256895. S2CID 15782597. Archived from the original on 2012-06-16. Retrieved 2012-06-12.
40. European Innovative Medicines Initiative. 2020. Retrieved 24 January 2020.
41. US Food and Drug Administration. 23 April 2018. Retrieved 24 January 2020.
42. Breakthrough Therapy". US Food and Drug Administration. 4 January 2018. Retrieved 24 January 2020.
UNDER PEER REVIEW

43. Abad-Zapatero, C. . A Sorcerer's apprentice and the rule of five: from rule-of-thumb to commandment and beyond. Drug Discovery Today, 2007; Vol. 12, pp. 995-997
44. Angell, M. The Truth About the Drug Companies: How They Deceive Us and What to Do About It pp. Random House Publishing; 2004;, ISBN 0-37576094-37576096, New York, U.S.
45. Bauer, RA.; Wurst J.M. & Tan DS. Expanding the range of 'druggable' targets with natural product-based libraries: an academic perspective. Current Opinion in Chemical Biology,2010; Vol. 14, pp. 308-314
46. Bender A.; Fergus S, Galloway WRJD et al. Diversity oriented synthesis: A challenge for synthetic chemists. Chemical Genomics: Small Molecule Probes to Study Cellular Function, 2006; Vol. 58, pp. 47-60
47. Billingsley, ML. Druggable targets and targeted drugs: enhancing the development of new therapeutics. Pharmacology, 2008; Vol. 82, pp. 239-244
48. Boehringer M, Boehm HJ, Bur D et al. Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening. Journal of Medicinal Chemistry, 2000; Vol. 43, pp. 2664-2674
49. Bon, R. S. & Waldmann H. Bioactivity-guided navigation of chemical space. Account of Chemical Research, 2010; Vol. 43, pp. 1103-1114
50. Booth, B. & Zemmel R. Opinion - Quest for best. Nature Reviews Drug Discovery, 2003; Vol. 2, pp. 838-841
51. Burke, M. Innovation: Europe must do better. Chemistry World, 2011; Vol. 8, pp. 12 Burke, M. D. & Schreiber SL. A planning strategy for diversity-oriented synthesis. Angewandte Chemie-International Edition, 2004; Vol. 43, pp. 46-58
52. Chessari, G. & Woodhead AJ. (2009b). From fragment to clinical candidate--a historical perspective. Drug Discovery Today,2009; Vol. 14, pp. 668-675
53. Denny, WA. The contribution of synthetic organic chemistry to anticancer drug development. In Anticancer Drug Development, (Baguely B.C and Kerr D.J., eds), 2002; pp. 187-202, Academic Press
54. Dobson, CM. Chemical space and biology. Nature, 2004; Vol. 432, pp. 824-828
55. Edwards, PD; Albert JS, Sylvester M et al. Application of fragment-based lead generation to the discovery of novel, cyclic amidine beta-secretase inhibitors with nanomolar potency, cellular activity, and high ligand efficiency. Journal of Medicinal Chemistry, 2007; Vol. 50, pp. 5912-5925
56. Extance, A. (2011). EPSRC plans represent 'huge change'. Chemistry World, 2011; Vol. 8, p. 8
57. Feher, M. & Schmidt JM. Property distributions: Differences between drugs, natural products, and molecules from combinatorial chemistry. Journal of Chemical Information and Computer Sciences,2019; Vol. 43, pp. 218-227
58. Frantz, S. Creating the right chemistry. Nat Rev Drug Discov, 2003; Vol. 2, p. 163
59. Friedman, Y. Location of pharmaceutical innovation: 2000-2009. Nature Reviews Drug Discovery, 2010; Vol. 9, pp. 835-836
60. Fuller JC, Burgoyne NJ. & Jackson RM. Predicting druggable binding sites at the protein-protein interface. Drug Discovery Today, 2009; Vol. 14, pp. 155-161
61. Haggarty SJ, Koeller KM, Wong JC, et al. (2003). Multidimensional chemical genetic
62. analysis of diversity-oriented synthesis-derived deacetylase inhibitors using cell-based assays. Chemical Biology, 2003;Vol. 10, pp. 383-396
63. Hanson PR, Rolfe A. & Lushington GH. Reagent based DOS: A "Click, Click, Cyclize" strategy to probe chemical space. Organic & Biomolecular Chemistry, 2010; Vol. 8, pp. 2198-2203.
64. Hirsch, J. An anniversary for cancer chemotherapy. JAMA, 2006; Vol. 296, pp. 1518-1520

UNDER PEER REVIEW

65. Hopkins AL. & Groom CR. The druggable genome. Nature Reviews Drug Discovery, 2002; Vol. 1, pp. 727-730
66. Hughes, B. Evolving R&D for emerging markets. Nature Reviews Drug Discovery, 2010; Vol. 9, pp. 417-420
67. Justice BA, Badr NA. & Felder RA. 3D cell culture opens new dimensions in cell-based assays. Drug Discov Today, 2009; Vol. 14, pp. 102-107
68. Lombardino JG. & Lowe JA. The role of the medicinal chemist in drug discovery--then and now. Nat Rev Drug Discov, 2004; Vol. 3, pp. 853-862
69. Ma P, Gudiksen M; Fleming E; et al. Outlook - What drives success for specialty pharmaceuticals? Nature Reviews Drug Discovery, 2008; Vol. 7, pp. 563-567
70. Moos WH; Hurt CR. & Morales GA. Combinatorial chemistry: oh what a decade or two can do. Molecular Diversity, 2009; Vol. 13, pp. 241-245
71. Moura-Letts G, Diblasi CM; Bauer RA; et al. (2011). Solid-phase synthesis and chemical space analysis of a 190-membered alkaloid/terpenoid-like library. Proc Natl Acad Sci U S A, 2011; Vol. 108, pp. 6745-6750
72. Mullin, R. Battening the Hatches. Chemical & Engineering News, 2010; Vol. 88, p. 14
73. Munos, B. Lessons from 60 years of pharmaceutical innovation. Nature Reviews Drug Discovery, Vol. 2009; 8, pp. 959-968
74. Nghiem P. & Kawasumi M. Chemical genetics: Elucidating biological systems with small-molecule compounds. Journal of Investigative Dermatology, 2007; Vol. 127, pp. 15771584
75. Nielsen TE. & Schreiber SL. Towards the optimal screening collection: a synthesis strategy. Angew Chem Int Ed Engl,2008; Vol. 47, pp. 48-56
76. Overington JP, Al-Lazikani B. & Hopkins AL. Opinion - How many drug targets are there? Nature Reviews Drug Discovery, 2008; Vol. 5, pp. 993-996
77. Philpott KL, Hughes JP, Rees S, et al. Principles of early drug discovery. British Journal of Pharmacology, 2011; Vol. 162, pp. 1239-1249
78. Pittman D. Changing Their Ways. Chemical & Engineering News, 2010; Vol. 88, pp. 47-49
79. Pors, K.; Goldberg F. W.; Leamon C. P.; et al. The changing landscape of cancer drug discovery: a challenge to the medicinal chemist of tomorrow. Drug Discovery Today,2010 Vol. 14, pp. 1045-1050
80. Rooseboom M, Commandeur JN. & Vermeulen NP. Enzyme-catalyzed activation of anticancer prodrugs. Pharmacol Rev, 2004; Vol. 56, pp. 53-102
81. Rydzewski RM. The Drug Discovery Busines to Date. In: Real World Drug Discovery: A Chemist's Guide to Biotech and Pharmaceutical Research, 2008; pp. 1-52, Elsevier, ISBN: 978-970-908-046617-046610, Amsterdam, The Netherlands.
82. Shoichet BK, Hert J, Irwin JJ, et al. Quantifying biogenic bias in screening libraries. Nature Chemical Biology, 2009; Vol. 5, pp. 479-483
83. Sneader WE. Legacy of the past. Drug discovery: a history,2005; pp. John Wiley & Sons Ltd, ISBN-10 10-471-89979-89978, West Sussex (U.K.)
84. Spandl RJ, Bender A. & Spring DR. Diversity-oriented synthesis; a spectrum of approaches and results. Organic & Biomolecular Chemistry, 2008; vol. 6, pp. 11491158
85. Spring DR; Spandl RJ. & Bender A. Diversity-oriented synthesis; a spectrum of approaches and results. Organic & Biomolecular Chemistry, 2008; Vol. 6, pp. 11491158
86. Stockwell BR. Exploring biology with small organic molecules. Nature, 2004; Vol. 432, pp. 846-854
87. Strebhardt K. & Ullrich A. Paul Ehrlich's magic bullet concept: 100 years of progress. Nature Reviews Cancer, 2008; Vol. 8, pp. 473-480
UNDER PEER REVIEW

88. Swinney DC. & Anthony J. How were new medicines discovered? Nature Reviews Drug Discovery, 2011; Vol. 10, pp. 507-519
89. Tralau-Stewart CJ, Wyatt CA, Kleyn DE; et al. Drug discovery: new models for industry-academic partnerships. Drug Discov Today, 2009; Vol. 14, pp. 95-101
90. Tremblay JF. The Grand Experiment. Chemical & Engineering News, 2010; Vol. 88, p. 20
91. Ward S. Demographic factors in the Chinese health-care market. Nature Reviews Drug Discovery, 2008; Vol. 7, pp. 383-384
92. Wilhelm S, Carter C, Lynch M et al. Discovery and development of sorafenib: a multikinase inhibitor for treating cancer. Nature Reviews Drug Discovery, 2006; Vol. 5, pp. 835-844
93. Yang SH, Dou KF. & Song WJ. (2010). Prevalence of diabetes among men and women in China. New England Journal of Medicine,2010; Vol. 362, pp. 2425-2426; author reply 2426
94. Zartler, ER. & Shapiro MJ. Fragonomics: fragment-based drug discovery. *Current Opinion in Chemical Biology*, 2005; Vol. 9, pp. 366-370
95. Zhao, H. & Guo Z. Medicinal chemistry strategies in follow-on drug discovery. *Drug Discov Today*, 2009; Vol. 14, pp. 516-522
96. Zlokarnik G, Grootenhuis PD. & Watson JB. High throughput P450 inhibition screens in early drug discovery. *Drug Discov Today*, 2005; Vol. 10, pp. 1443-1450
97. JP Hughes[1], S Rees[2], SB Kalindjian[3] and KL Philpott[3.] Principles of early drug discovery British Journal of Pharmacology 2011;
98. Fox S, Farr-Jones S, Sopchak L, Boggs A, Nicely AW, Khoury R *et al*. (2006). High-throughput screening; Update on practices and success. J Biol Screen 11: 864–869.
99. Law R, Barker O, Barker JJ, Hesterkamp T, Godemann R, Andersen O *et al*. (2009). The multiple roles of computational chemistry in fragment-based drug design. J Comput Aided Mol Des 23: 459–473.

UNDER PEER REVIEW