Compound library design and utilization trends of the past decade

Chemical industry has undergone several changes and has come a long way in the past eras. Early societies were already able to distinguish the biological effects from plants and could even extract compounds from them. Knowledge was passed on through tradition. This eventually led to the vast library of biological active substances we know today. Alchemists, often disguised as fools searching for gold, in fact experimented with natural chemical compounds and discovered that nature was makeable, in a sense that their experiments lead to the knowledge how under certain conditions compounds react to form new ones. From that moment on, existing compound libraries, partly written and partly passed on through tradition, mainly describing biological effects from plants and minerals, were extended with descriptions of synthesized compounds with biological effects.

When a layman directs his interest into today's state-of-the-art pharmaceutical science and visits the websites of leading pharmaceutical companies, he'll immediately notice the giant leap which has been made since the early days of compound design. All sites state that good health and finding solutions to the most pressing healthcare issues are vital to all of us. Information provided on these sites seldom reveals more than a product name and a safety sheet; hardly no information can be found on chemical structure and biological effect. Further reading will reveal that the search for compounds is no longer the exclusive competence of pharmaceutical industries and scientific research groups.
Pharmaceutics today is the search for collaboration and this is driven by the need to increase efficiency, reduce costs and find ways to manage the vast growing amount of data. Newly formed and highly specialized industries perform many of the tasks, i.e. compound synthesis and target screening, formerly done by the pharmaceutical industries themselves.

How are new medicines discovered?
Historically, new drugs were discovered through identifying the active ingredient from traditional remedies or by serendipitous discovery. Nowadays, chemical compounds are industrially synthesized and stored into chemical libraries. Modern drug discovery involves high-throughput screening of large compound libraries against isolated biological targets with hypothetical disease-modifying activity. The results of these experiments provide a starting point for drug design and for understanding the interaction or role of a particular biochemical process in (human) biology. Hits from these screens are tested in cultured cells and subsequently in animals and later on in human for sensitivity, selectivity, efficacy, and safety. This is a long lasting, inefficient, and expensive process with low rates of new therapeutic discovery.
Over the last decade a striking different approach in designing new compounds has taken place. Until approximately 10 years ago, molecules were designed as 'flat' molecules and synthesis was relatively easy by choosing from a set on the shelf of organic building blocks and coupling them together in the desired order with the target molecule. Standard reactions, such as Palladium coupling reactions, were

used. Altering or adding some reactive groups was sometimes necessary to complete the synthesis.

This is a good way to synthesize large amounts of certain organic chemical compounds in a short period of time. However, these 'flat' molecules have some undesirable properties, such as low water solubility and hence poor bioavailability. In a later stage of the process of drug development, complicated and time consuming procedures have to be carried out in order to enhance the solubility of the compound. The assembly of three-dimensional molecules, with different characteristics from 'flat' molecules, is more common nowadays. This kind of chemistry is more complicated and more challenging, but will eventually lead to compounds with better bioavailability. An example of a three-dimensional molecule is the macrocycle, i.e. cyclic macromolecule or macromolecular cyclic portion of a molecule. Natural product macrocycle drugs and their synthetic derivatives have been used for many years in different therapeutic areas. Despite their profound pharmacological properties and proven success, macrocycles have been poorly explored within drug discovery. Research will now focus on the wider use of macrocyclic scaffolds in medicinal chemistry. Specific characteristics of macrocycles, such as diverse functionality and stereochemical complexity, may result in high affinity and selectivity for protein targets, while maintaining bioavailability.

Library design

The process of drug design starts by searching large compound libraries to identify promising leads for potentially marketable drugs. High-throughput screening is a technique widely used to rapidly scan and analyze these libraries. A key feature for a successful library search is, however, the strategy used to design the library itself, and whether the design increases the probability of retrieving promising 'hits' and potential leads. Therefore, great emphasis in research is on compound library design strategies.
Over the last decade, the design of screening libraries has undergone several changes. Early attempts to design libraries often consisted of adding the largest set of compounds possible.

The first libraries typically included huge amounts of small-molecule structures. Essentially, the basic approach was to screen as many compounds against a particular (therapeutic) target in a given period of time as possible. Library design today is more sophisticated and centers around the methods used for choosing compound membership. The choice of compounds is often based on two design strategies: diversity-based design and target-based design.
Target-based compound libraries are collections of compounds specifically tailored to modulate an individual protein target or a family of related targets, such as kinases, G protein-coupled receptors, voltage-gated ion channels, or serine/cysteine proteases. These libraries may be selected from larger, more diverse collections using literature search and computational techniques such as in silico docking to the target or ligand similarity calculations using molecular fingerprints. Diversity-based or phenotypic screening libraries have been used for many years to identify potential drug candidates. An analysis of Swinney and Anthony published in 2011 showed that the contribution of phenotypic screening to the discovery of first-in-class drugs with new molecular mechanisms of action actually

exceeded that of target-based screening methods, in an era in which the major focus was on target-based approaches.

Collaborations

The search for collaboration between industry, small and medium-sized enterprises (SMEs), and academia has resulted in ambitious consortia such as the European Lead Factory and EU-OPEN-SCREEN. The European Lead Factory is an unique pan-European initiative for the promotion of new drug discoveries via open innovation and crowdsourcing. Launched at the beginning of 2013, the consortium is funded by the Innovative Medicines Initiative (IMI) and in-kind contribution from various partners for a period of 5 years. The initiative will be working toward the development of a library of 500,000 compounds, of which 300,000 previously safeguarded, high quality corporate compounds are contributed by European Federation of Pharmaceutical Industries and Associations (EFPIA) members. In addition, a substantial new Public Compound Collection of around 200,000 compounds is being constructed, based on syntheses by SMEs and academic institutions and exploiting the expertise available across all consortium partners. Together, the EFPIA and Public Compound Collections will form the Joint European Compound Collection, that will be used to screen against biological targets.

This Joint European Compound Collection gives the entire European research community access to a unique, high quality and extensive lead-like compound library.

The screening of compounds to assess their biological activity is also funded and performed by the European Lead Factory. Compounds in the EFPIA Compound Collection will be screened against both commercially contributed targets and targets from public sources. The European Lead Factory offers researchers the infrastructure and expertise needed to translate their innovative ideas and targets to ultrahigh-throughput screening facilities, and to provide them with quality hits for starting points for drug discovery. This means substantial cost savings for contributors as well as very productive exchange of ideas. State-of-the-art facilities are provided by the newly-established European Screening Centre, based in Scotland and the Netherlands.

The objective of the EU-OPENSCREEN initiative is to accelerate the discovery of biologically active substances in all areas of life sciences, including early drug discovery, microbiology, of 200,000 commercial molecules. This library is a collaborative selection effort of five European research groups each contributing 50,000 compounds to the selection.

Learn more about this topic at the 11th International Conference Intelligent Compound Libraries, 13 - 15 October, 2015- Berlin, Germany!