Symeres
Making Molecules Matter. Together.

DRUG DISCOVERY

Unlocking the Potential of High-Throughput Screening: Symegold Library Design and Expansion Insights

LEON VAN BERKOM

The results you achieve will be in direct proportion to the effort you apply

Denis Waitley

Abstract: *This paper emphasizes the importance of compound libraries and their quality in shaping the success of drug discovery campaigns. Specifically, it delves into the design principles and physicochemical attributes of Symeres' Symegold library, highlighting its role in hit-finding activities. The paper details the criteria for selecting compounds for the library, the synthesis process, and the management of the compound collection. The paper advocates for the strategic utilization of diverse and high-quality compound libraries, exemplified by the Symegold collection, as a crucial component of successful hit discovery programs.*

Introduction

In the pharmaceutical industry, high-throughput screening (HTS) remains the cornerstone of drug discovery. HTS involves swiftly and automatically testing large compound libraries against a biological target, enabling the identification of compounds with desired biological activity from thousands to sometimes millions of candidates. Over decades, the pharmaceutical industry has continued to invest in maintaining, expanding, and improving compound libraries for use in HTS campaigns. These endeavors are justified, as the quality of hits identified through HTS forms the foundation of a drug discovery campaign and can significantly influence its success or failure. While the size of a library is commonly perceived as a crucial factor determining its utility, there is growing recognition that factors such as chemical diversity, novelty, and sp^3 richness also play a pivotal role in determining a library's effectiveness.

A cornerstone of Symeres' hit-finding activities is our 75,000 compound Symegold library that was created as part of the European Lead Factory (ELF) consortium. The compounds in this collection are based on diverse leadlike properties and structural novelty. Due to the success of the library, Symeres continues to expand the Symegold collection and aims to add 10,000 novel molecules by the end of 2025.

Symegold
Your Exclusive Compound Library

This paper elaborates on the design principles and physicochemical attributes of the Symegold library. At Symeres, we recognize the paramount importance of clear project governance in achieving success. In the case of the Symegold collection, three pillars underpin its foundation: library design, library production, and compound management. Each of these pillars is further detailed in the subsequent sections below.

Library design

At Symeres, we uphold the principle that quality eclipses quantity. The establishment of the Symegold library was grounded in the conviction that its constituents must adhere to six fundamental criteria: innovation, druglikeness (avoiding undesirable reactive chemical groups, frequent hitters, toxicophores, aggregators, and similar undesirable features),

(calculated) properties, novelty, chemical tractability, and potential for diversification (Figure 1).

Harnessing the innovative spirit of Symeres' workforce, scientists are encouraged to propose concepts for scaffolds that are aligned with the aforementioned principles to the Symegold design committee. This committee, comprising of organic and medicinal chemists, evaluates each new idea against the Symegold design standards. Only after approval from this expert committee are the scaffolds and their associated libraries selected for synthesis. All new concepts are archived in a virtual idea repository known as ScafVault. Presently, ScafVault houses over 300 ideas that have been sanctioned for synthesis.

While the majority of these concepts yield druglike molecules that adhere to Lipinski's rule of five, the Symegold collection also contains molecules that do not strictly fall within that physicochemical space, such as scaffolds and pseudonatural compounds. While the entire collection can undergo comprehensive screening, there is also the flexibility to focus screening efforts on specific compound classes, such as macrocycles or pseudonatural products.

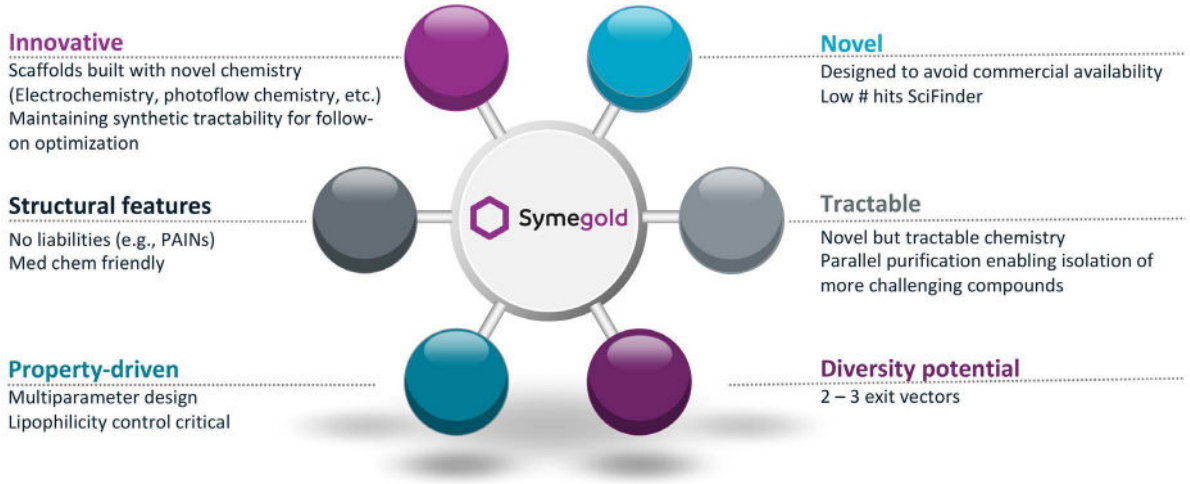

Figure 1. Symegold design principles

The emergence of cross-coupling methodologies in the 1980s has provided chemists with an expanded synthetic toolbox, but this has also resulted in a noticeable shift in the molecular landscape, which is characterized by molecules with reduced sp^3-like character and increased planarity.[1] To counteract this tendency towards a 'molecular flatland,' the excessive use of sp^2-hybridized atoms in drug candidates must be minimized in favor of sp^3-hybridized atoms. In that sense, spirocyclic molecules present a dual advantage: not only do these molecules inherently possess three-dimensional characteristics, but they also offer a high degree of rigidity. Thus, for these reasons, the Symegold collection is notably enriched with spirocyclic scaffolds. Figure 2 illustrates examples of spirocyclic scaffolds in the Symegold collection.

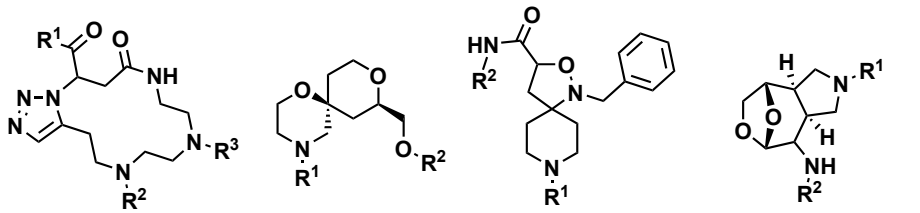

Figure 2. Examples of Symegold scaffolds

The profound influence of natural products on treating a diverse array of ailments has been extensively documented. For centuries, both plant- and marine-derived remedies have played crucial roles in addressing a wide spectrum of diseases, notably in fields such as oncology and antibiotics. While the potential of natural compounds to inspire new drug candidates is commonly acknowledged, there has only been limited discourse on their ability to serve as powerful starting materials for generating novel therapeutic agents unrelated to their biological origins. This notion of utilizing natural-compound-derived scaffolds has found successful application in several Symegold scaffolds. The Symegold library is currently bolstered by approximately 4,000 of these semisynthetic compounds. Figure 2 illustrates an example of this concept and highlights the innovative and imaginative essence inherent in the Symegold library.

Macrocyclic compounds represent a compelling frontier in drug discovery, offering a myriad of advantages and utilities that render them invaluable in pharmaceutical research and development.[3] Macrocycles tend to bind to biological targets with high affinity, and they often exhibit remarkable resilience against metabolic degradation. For this reason, there has been a surge of interest in macrocyclic compound libraries. The Symegold library was therefore designed to contain 2,000 macrocyclic molecules.

Library production

Utilizing in-house-developed algorithms, nominated scaffolds are first enumerated in silico into a structurally diverse set of final compounds. This virtual library is constructed from Symeres' extensive collection of almost 3,000 reagents (amines, carboxylic acids, halides, sulfonyl halides, etc.). Drawing on past project experience, problematic building blocks are identified and either removed or substituted with alternatives to ensure a high synthetic success rate. To enhance the structural diversity of the library, Symeres opts not to exhaustively explore all possible variations of building blocks. Instead, an enhanced secondary selection based on both diversity and physicochemical properties increases coverage of chemical space by including the most dissimilar compounds.

After the design of the virtual library has been completed, the virtual concepts are put into practice through the synthesis of key intermediates or scaffolds on the multigram scale. Those scaffolds are then used in the synthesis of the corresponding compound libraries. To allow for efficient follow-up upon identifying hit compounds in the HTS campaign, key intermediates and scaffolds are stored, enabling the rapid preparation of analogues from advanced building blocks. A more comprehensive overview of Symeres' parallel chemistry platform is provided in a separate white paper that is accessible on our website.

Properties

The sets undergo enumeration to capture highly desirable attributes that are crucial for medicinal chemistry research. Informed by input from internal experts and external partners, our selection criteria yield a wide array of diverse compounds within specified limits for hit-finding libraries. Numerous factors are considered for the final selection, including cLogP, polar surface area, molecular weight, and number of aromatic rings (Figure 3).

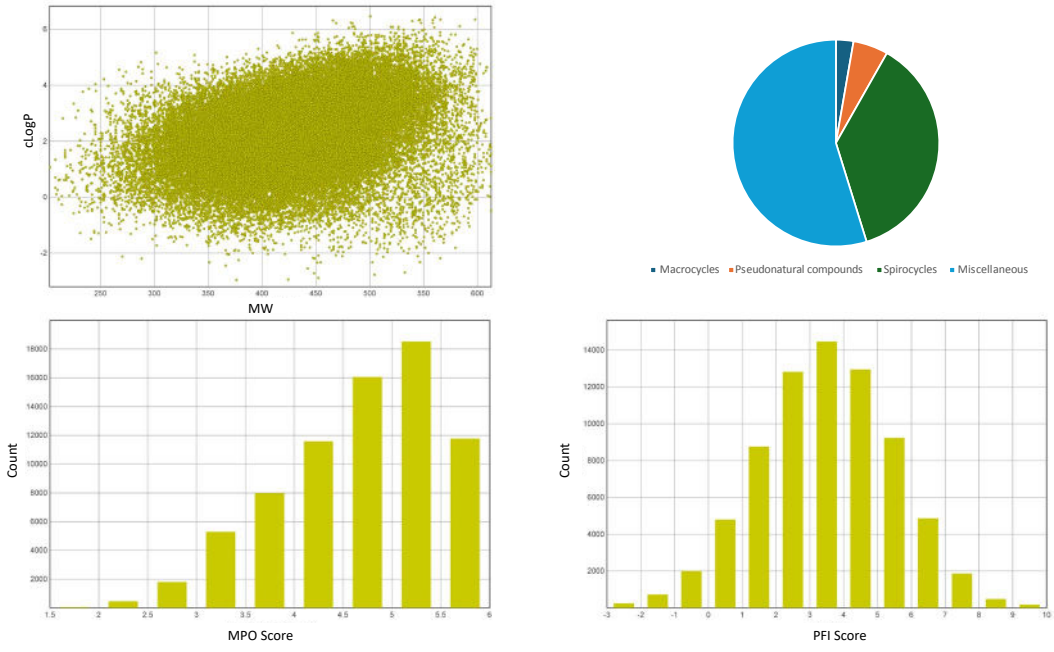

Figure 3. Selected properties of Symegold library

As anticipated, the majority of our compounds adhere to Lipinski's rule of five, while also striving to enhance the multiparameter optimization (MPO) score and to optimize the property forecast index (PFI), among other parameters.[4,5], Through our selection protocol and enumeration methodology, we even explore compounds at the periphery of these criteria, culminating in a set well-suited for diverse protein target types.

Compound management and HTS

Before compounds are added to the Symegold library, they undergo rigorous analysis and are subjected to the QA/QC team's scrutiny. Only compounds with a purity exceeding 90% are included in the library. Typically, the purity of compounds far exceeds this threshold, and the average purity of the compounds in the Symegold collection approaches 97%.

Solid aliquots of the final compounds are stored in barcoded vials at −20°C at Symeres' research site in Nijmegen, the Netherlands. For most compounds in the Symegold collection, at least 2 mg is available as neat stock, and aliquots of these solids can be provided with a turnaround time of less than 2 weeks. Additionally, DMSO aliquots can be accessed by collaborators through Axxam S.p.A in Milan, Italy.

The cornerstone elements for a thriving hit discovery program encompass top-tier compound libraries, purposeful assays, and meticulously crafted screening strategies. Although Symeres does not directly provide HTS services, it extends this capability through its strategic collaborator, Axxam S.p.A.[6] Clients intending to conduct an HTS campaign can leverage Axxam's assay development expertise, choose from a range of pre-existing assays, or transfer their validated assays to Axxam's platform. The Symegold library is offered by Axxam as part of their discovery services, alongside their AXXDiversity library and their RNA-targeting library.

Concluding remarks

At Symeres, we firmly believe that one of the fundamental ingredients for a successful hit discovery program is a highly diverse and high-quality compound library. While the number of compounds available in screening collections is ever increasing, it is important to look beyond the standard commercially available libraries, especially for difficult or uncommon targets. The Symegold collection presents a unique set of compounds that provide an excellent starting point for hit-finding programs. Because of our experience with the chemistry of these library compounds, we are also able to follow up with fast hit optimization, providing additional benefits for our customers.

Current and prospective clients can access high-quality chemistry from Symeres alongside state-of-the-art assay development and the HTS expertise of our strategic partner Axxam S.p.A.

1 *Lovering, J. Med. Chem.* ***2009****, 52, 6752–6756*
2 *DeCorte, J. Med. Chem.* ***2016****, 59, 9295–9304*
3 *Driggers, Nat. Rev. Drug Discovery* ***2008****, 7, 608–624*
4 *Wager, ACS Chem. Neurosci.* ***2016****, 7, 767–775*
5 *Young, Drug Discovery Today* ***2011****, 822–830*
6 *www.Axxam.com*

For further inquiries or collaboration opportunities, please contact us via: science@symeres.com
www.symeres.com

Symeres