A Chemogenomic Approach to Kinase Drug Discovery

Kevin P. Williams, *KPWbio Research, Chapel Hill, USA*

EuroConference Protein Kinase Inhibitors Full Abstract

Abstract

Targeted approaches to providing therapies for diseases such as cancer have become a major focus for drug discovery programs. In practice, the “targeted profiles” of the majority of kinase drugs have evolved more by serendipity than by design. While a number of technologies have demonstrated the usefulness of retrospectively identifying the inhibitory profile of drugs, the ability to prospectively mine for and define the selectivity of a compound at the onset of a program has lacked a database with a comprehensive and precise dataset. Historically, it has been difficult to use high throughput screening (HTS) data for selectivity or structure-activity relationship (SAR) analysis due to its heterogeneity in content and quality. By utilizing an industrialized systems process that integrates broad biology, high precision microfluidics technologies and robust chemistry a chemogenomic database has been generated. Progress on using the database as the initiation point for drug discovery programs is discussed.

Chemical Genomics

Chemical genomics differs from traditional drug discovery primarily in breadth and quality. A chemical genomics program should test not only validated targets, but also broad families of targets in order to identify patterns of activity as well as ensure selectivity against individual targets. For a successful chemical genomics program, a large, high quality, diverse collection of chemical agents must be assembled and characterized. Biological assays must be created for each target and these must be of high precision and accuracy to allow correlation of the data from all assays across the chemical library. Ideally, this is aided by utilizing a single technology platform. These two functions must then be merged in a specialized high throughput screening (HTS) operation. Finally, the data must be collected and mined to produce usable information.

Building the Platform – Chemistry, Targets and Assay Technology

One such chemical genomics approach to drug discovery, pioneered by Amphora, uses an industrialized system, encompassing high precision microfluidic-based assays, broad chemistry and biology information (reviewed in Janzen et al. 2004).

The key steps to building this process and database were to assemble a diverse small molecule library of druggable, highly pure and quantitated compounds (Popa-Burke et al. 2004). Purifying and quantifying each sample in compound libraries is a substantial scientific undertaking and is not a typical aspect of traditional HTS. However, advances in assay technology now enable accurate, reproducible data production from primary compound library screenings. The Amphora small molecule library currently consists of 135,000 compounds; all compounds were purified to >95% by liquid chromatography/mass spectrometry. Compound concentration in DMSO and aqueous is

determined primarily by chemi-luminescence nitrogen detection (CLND) (Popa-Burke et al. 2004). This high quality concentration data has aided predictive SAR analysis.

For kinases, target selection encompassed not only therapeutic targets and critical off-targets but also spanned multiple pathways. All known proteins participating in a targeted disease pathway represent potential drug targets, and all non-target proteins represent potential undesired interactions. Ideally, the chemical genomics approach would provide a small molecule for every kinase encoded by the genome as potential target validation tools or drug leads.

The ability to develop and transfer the large number of assays necessary to make the chemical genomics concept meaningful requires the implementation of defined and consistent processes at all steps of assay development and was greatly aided by using a single microfluidic-based technology platform. In contrast, assays that have multiple steps, including incubation times, numerous reagent additions, or washing steps are inherently noisier and are less precise. The more direct the assembly and readout of the assay, the easier it is to fulfill these requirements. Microfluidics is one of a limited number of functional assay formats that allows a direct readout of "inhibition". For enzymatic kinase assays, the assay reaction consists simply of kinase plus fluoroscein-labeled substrate peptide in the presence or absence of an inhibitor. Substrate is separated from phosphorylated product peptide electrophoretically on the microfluidic chip. The ability of these assays to directly measure both product and substrate peaks in an intervention-free manner provides the precision and comparability of inhibition data necessary for meaningful analysis. Greater than 80 kinase assays were developed with comparable biochemical and kinetic conditions. An advantage of broad initial screening of a purified compound library is the early identification of promiscuous compounds. Such behavior is frequent enough that enzymatic behavior filters are now included for inter-family inhibition so that such compounds are not selected for optimization. This enables the correct chemistry decisions to be made early in a program and allows series that have undesirable properties to be filtered out.

Mining the Database - Drug Discovery Programs

The Amphora chemogenomic database presently consists of >30 million data points based on the screening of >130,000 compounds versus >100 targets. The accuracy of the data has allowed an active compound to be defined at a six sigma inhibition threshold; typically the 99.9% confidence interval for designating a compound as active is in the 10-20% inhibition range. These threshold limits, in contrast to those typical of traditional HTS, have allowed the definition of "credible negative" values. This has allowed the screens to detect even weak inhibitors such as those identified for MAPKAPK-2. For the microfluidic-based high-throughput screen (HTS) of p38α MAP kinase the six sigma threshold was 16.4% and the precision radius was 6.2% (Norris et al. 2005). A tight precision radius allows for discrimination between close values, enabling the extraction of credible data points for SAR analyses.

We now have the ability to mine the dataset ("actives analysis") based on predefined target profiles and, in conjunction with scaffold clustering algorithms, can rapidly define chemical series. This has also proved to be a useful tool to guide target selection and analysis of multi-dimensional data obtained from compound profiling.

Examples of such analyses include "multi-targeted (e.g. Flt3/Kit/PDGFR) versus selective kinase (Akt) inhibitors", targeting mutation versus wild type, targeting parallel or downstream signaling pathways and profiling known inhibitors will be shown. Furthermore, in contrast to the disparate and limited datasets in the literature, we have begun using our high precision and comparable database to build sequence-independent dendrograms as tools to predict chemogenomic space.

The focus on precision and quality at all steps of the Amphora lead discovery process including library construction and handling, implementation of a high-precision microfluidic screening platform, statistical analysis of the HTS output, IT solutions and industrialized enzymology, has resulted in a chemogenomic database of compound data with comparable selectivity information against multiple target types and classes. Using this database as the starting point for multiple programs has allowed rapid progress to be made towards identifying lead scaffolds with potent analogs while maintaining the desired selectivity profiles. This becomes critical in the prioritization of which projects or scaffolds are advanced or are de-emphasized, leading to enhanced efficiency and productivity in drug discovery.

References

1. Janzen, W.; Bernasconi, P.; Cheatham, L.; Mansky, P.; Popa-Burke, I.; Williams, K.; Worley, J.; Hodge, N. (2004) Optimizing the Chemical Genomics Process. In: Chemical Genomics: Advances in Drug Discovery and Functional Genomics Applications. Ed. Darvas, Guttman, Dorman. Marcel Dekker, NY.

2. Popa-Burke, I.G.; Issakova, O.; Arroway, J.D.; Bernasconi, P.; Chen, M.; Coudurier, L.; Galasinski, S.; Janzen, W.P.; Lagasca, D.; Liu, D.; Lewis, R.S.; Mohney, R.P.; Sepetov, N.; Sparkman, D.A.; Hodge, C.N. (2004). Streamlined System for Purifying and Quantifying a Diverse Library of Compounds and the Effect of Compound Concentration Measurements on the Accurate Interpretation of Biological Assay Results. Analytical Chemistry, 76:7278-7287.

3. Norris, J.N., Williams K.P., Janzen W.P., Hodge C.N., Blackwell L. and Popa-Burke I. (2005). Selectivity of SB203580, SB202190 and Other Commonly Used p38 Inhibitors: Profiling against a Multi-Enzyme Panel. Lett Drug Design and Disc 2: 451-455.