Summary
Structural GenomiX, Inc. (SGX) has combined high-throughput X-ray crystallography with a fragment-based approach to lead identification/lead optimization. FAST™ (Fragments of Active Structures) exploits crystallographic screening to detect, visualize, and identify small lead-like ligands (molecular weight 150-200, typical hit rates 1-5%) that bind productively to the target protein. Each member of the FAST™ fragment library is amenable to rapid chemical elaboration at two or three points to provide access to unrivaled chemical diversity using parallel organic synthesis. Initial lead optimization involves using knowledge of the structure of the target-fragment complex and advanced computational chemistry tools to guide synthesis of small focused linear (one-dimensional) libraries. These linearly elaborated fragments are then evaluated with in vitro biochemical assays and crystallography. Thereafter, optimal variations at each point of chemical diversity are combined to synthesize focused combinatorial (two- or three-dimensional) libraries that are again examined with assays and crystallography. (The theoretical chemical diversity of the fully elaborated FAST™ fragment library exceeds 160 million compounds.) These focused combinatorial libraries typically contain multiple novel compounds of low molecular weight (<350) that bind the target protein with low nM IC_{50} values and display considerable selectivity. Finally, lead series are prioritized for further medicinal chemistry and compound advancement efforts using the results of cellular and animal model assays, in vitro and in vivo ADME and in vitro toxicology studies in concert with structural information. Application of the FAST™ fragment-based lead generation process to various oncology targets will be presented.
Structure based approaches have attracted the interest of the pharmaceutical and biotechnology industries for more than two decades. Until the advent of high-brightness synchrotron sources and cryogenic sample preservation methods in the 1990s, however, protein structures were only attainable after lengthy and resource-intensive efforts, greatly limiting their contributions to drug discovery. More recently, Nienaber and colleagues at Abbott Laboratories demonstrated the feasibility of using X-ray crystallography to screen for drug discovery starting points (Nienaber et al., 2000). Despite its promise, crystallographic screening has not been adopted widely because the approach depends critically on (i) the availability of high quality crystals of the target protein early in a drug discovery campaign, (ii) the ability to produce large numbers of target-small molecule complex structures in a time frame compatible with the needs of the medicinal chemist, and (iii) a reliable, efficient strategy for exploiting structural information to discover and optimize drug candidates. Today, high-throughput X-ray crystallography, advanced computational tools, and parallel organic synthesis have been combined to meet these challenges and provide new opportunities for lead discovery and optimization.

By way of historical introduction, it is important to understand how efforts to introduce X-ray crystallography, computational design, and parallel organic synthesis combined with high-throughput screening into main-stream drug discovery fared. More than a decade ago X-ray crystallography was taken up by most of the large pharmaceutical companies, because of its promise to deliver “rationally-designed” drugs. Regrettably, this anticipated revolution in drug discovery failed to materialize. Earlier incarnations of the X-ray method usually proved to be too slow to keep up with the pace of medicinal chemistry. In addition, computational design strategies based on the structure of the protein target by itself did not in fact permit rational design of small molecule inhibitors of biochemical and biological function. Two other would be revolutionary methods came along later, both of which also failed to live up to expectations. The advent of combinatorial chemistry did allow chemists to make large numbers of molecules quickly, but these compounds rarely displayed the drug like properties needed to make it through pre-clinical testing and into patients. High-throughput screening, which was developed to take advantage of the chemical bounty provided by combinatorial approaches, also failed to deliver useful lead compounds for many targets.

In response to pressure to reduce discovery timelines for new pharmaceuticals and increase the rate at which new chemical entities are discovered and brought to market, Structural GenomiX, Inc. (SGX) has developed a novel technology called Fragments of Active Structures (FAST™) that serves as an engine for high-quality lead generation. With the benefit of technical advances that have reduced many of their historical liabilities, this new approach to drug discovery exploits the strengths of X-ray crystallographic screening, structure based lead optimization, computational design, and combinatorial chemistry. For crystallographic screening to deliver maximum value to the medicinal chemist, fragment hits must be suitable to rapid, iterative "hit-to-lead" optimization. The FAST™ process is based on a library of small, lead-like fragments that are fully enabled for chemical modification. With access to industry leading structure
determination capabilities, SGX quickly produces potent and selective small molecule inhibitors of drug targets by leveraging medicinal chemistry resources with structural information and insights from computational chemistry.

**FAST™ Fragment Hit Discovery**

The SGX fragment-based approach begins with crystallographic screening of a target protein against the FAST™ library of ~1,000 small lead-like fragments (molecular weight 150-200). Each fragment has been tailored to include multiple points of diversity, or chemical “handles,” which are amenable to rapid parallel synthesis. The entire library is then screened within one to two days using a dedicated, state-of-the-art synchrotron beamline, SGX-CAT (Figure 1). Crystallographic screening entails soaking of pre-formed protein crystals with mixtures of ten structurally dissimilar fragments, followed by X-ray analysis of the soaked crystals. This approach yields hit rates of 1-5%, which is significantly higher than the hit rates of 0.010-0.025% typical for traditional high-throughput screening (HTS). The resulting structures provide “direct looks” at each one of the 10-50 lead-like fragments that bind to the crystalline target protein. The presence of bromine atoms in more than half of the SGX fragments facilitates the crystallographic screening process, using anomalous scattering signals to definitively establish the orientation of the bound fragment. Crystallographic screening hits are evaluated in terms of mode of binding and measured IC\textsubscript{50} (typically 1mM to 500\textmu M) to identify the most promising subset of candidates for lead optimization against the target.

**Elaborating FAST™ Hits**

With detailed insights from the structure of each high-priority fragment bound to the target protein, small linear (one-dimensional) libraries are designed through modifications at each site of chemical diversity. The presence of bromine atoms in more than half of the SGX fragments provides a second important process advantage, enabling carbon-carbon formation via Suzuki coupling or the Heck reaction. By applying an advanced computational filter, chemical modifications predicted to make undesirable steric clashes with the target are excluded, thereby eliminating application of expensive synthetic resources to chemical modifications that are unlikely to increase potency and selectivity. Each one-dimensional library is synthesized and then evaluated using \textit{in vitro} biochemical assays. Compounds with improved IC\textsubscript{50} values relative to the starting point or parent fragment (typically 10-100\textmu M) are then examined by X-ray crystallography to visualize the effect of each chemical modification on mode of binding and room to move in terms of further synthetic chemistry.

The most promising substitutions at each chemical “handle” are combined in small combinatorial (two- or three-dimensional) libraries of modifications at each site of chemical diversity. Again, advanced SGX computational tools are used to limit synthesis efforts to only those combinations of chemical modifications likely to improve potency and selectivity. Each combinatorial library is synthesized and then evaluated using \textit{in vitro} biochemical assays and X-ray crystallography. At this stage, the best of these multiply elaborated fragments have IC\textsubscript{50} values of 10-100nM and often display considerable selectivity against closely related proteins. Most targets yield at least two structurally distinct lead series that can be evaluated for activity in cells and animal
models of disease.

During the process of lead optimization via FAST™ structure-guided lead "engineering," particular attention is paid to keeping the mass of the inhibitor as low as possible. Typical molecular weights of doubly and triply elaborated fragments are <350, which represent an important factor in determining the bioavailability of a drug candidate. Similar attention is paid to other determinants of bioavailability, such as solubility, lipophilicity, the number of freely rotatable bonds, and the number of hydrogen bond donors and acceptors (Wenlock et al., 2002). Chemical moieties with known liabilities in terms of toxicity and undesirable ADME properties were excluded from both the FAST™ library and the reagents used for parallel synthesis of the focused linear and combinatorial libraries around each high-priority fragment hit.

Although the SGX fragment library consists of ~1,000 compounds, the combinatorial nature of the FAST™ process ensures that the library has unrivalled potential in terms of chemical diversity and typically yields novel compounds after only two rounds of chemical synthesis. The number of elaborations possible at each site of chemical diversity ranges from 400 to 40,000. Thus, the worst case scenario of only 400 possible modifications at each one of only two chemical "handles" gives a minimum chemical diversity of 400*400=160,000 distinct compounds/fragment or 160 million distinct compounds from the FAST™ library. Less conservative estimates of the potential chemical diversity of the SGX fragment library give the number of distinct compounds in the hundreds of billions. The combination of high-throughput X-ray crystallography and advanced computational tools allows SGX to synthesize only the small fraction of this enormous bounty of potential compounds that is likely to yield compounds with desirable target binding properties. Throughout the process, fragment modifications generating previously patented compounds are excluded and computational filters are used to ensure that expensive synthetic resources are used to explore only novel regions of chemical shape space.

The Changing Face of Lead Discovery

The FAST™ process is a highly efficient and adaptable lead discovery strategy for using large numbers of X-ray structures of lead-like fragments bound to important disease targets. These structures serve as the building blocks for novel drug candidates by guiding further modifications of initial screening hits. The method exploits high-throughput X-ray crystallography to identify and visualize synthetically enabled fragments bound productively to kinases and other enzymes. Thereafter, knowledge of how each fragment binds to its target is combined with advanced computational tools to maximally leverage the efforts of parallel organic synthesis to produce "engineered" leads that are potent, reasonably selective, and have drug-like physiochemical properties. The speed with which a FAST™ campaign can be prosecuted (3-6 months) and the ability of the process to access an enormous scope of diverse and novel compounds makes it both competitive with and complementary to HTS (Table 1). FAST™ may also provide a useful means of approaching targets, such as proteases, for which HTS often fails to yield acceptable medicinal chemistry starting points. None of this would be possible without
dedicated, full time access to a high-brightness synchrotron beamline, an optimally
designed fragment library, and proprietary computational chemistry software and
hardware developed and retained by SGX.

References
Nienaber, V.L., Richardson, P.L., Klighofer, V. Bouska, J.J., Giranda, V.L, and Greer, J.

comparison of physiochemical property profiles of development and marketed oral drugs.

Stephen K. Burley, M.D., D.Phil., F.R.S.C., is the Chief Scientific Officer and Senior Vice
President of Research at SGX (San Diego) and Adjunct Professor at the Rockefeller
University.
Table 1. Comparison of FAST™ and high-throughput screening (HTS) combined with lead optimization.

<table>
<thead>
<tr>
<th></th>
<th>FAST™</th>
<th>HTS/Lead Optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screening Library Size</td>
<td>&gt;160MM compounds</td>
<td>~2MM compounds</td>
</tr>
<tr>
<td>Hit Rate</td>
<td>1-5%</td>
<td>0.01-0.025%</td>
</tr>
<tr>
<td>Lead IC₅₀/Lead MW</td>
<td>&lt;100nM/~350 MW</td>
<td>&lt;100nM/~500 MW</td>
</tr>
<tr>
<td>Novelty Potential</td>
<td>High</td>
<td>Limited by HTS library</td>
</tr>
<tr>
<td>Time Required</td>
<td>3-6 months</td>
<td>8-15 months</td>
</tr>
<tr>
<td>Drug Like Properties</td>
<td>Built in at the outset</td>
<td>Addressed later</td>
</tr>
<tr>
<td>Target Limitations</td>
<td>Crystallographically enabled targets required (i.e., not applicable to GPCRs and ion channels, at present).</td>
<td>Many targets (e.g., proteases) fail to produce useful starting points for lead optimization.</td>
</tr>
</tbody>
</table>