Fragment-based structure-guided drug discovery: strategy, process, and lessons from human protein kinases

Stephen K. Burley, Gavin Hirst, Paul Sprengeler, and Siegfried Reich

INTRODUCTION

The experimental roots of fragment-based drug discovery can be found in the work of Petsko, Ringe, and coworkers, who were the first to report flooding of protein crystals with small organic solutes (e.g., compounds such as benzene with ten or fewer nonhydrogen atoms) to identify bound functional groups that might ultimately be transformed into targeted ligands.[1] The concept of linking fragments together to increase binding affinity was described as early as 1992 by Verlinde et al.[2] Computational screening of fragments, using tools such as DOCK[3,4] or MCSS,[5] was also described in the early 1990s. Pharmaceutical industry application of fragment screening began at Abbott Laboratories, where Fesik and coworkers pioneered "SAR by NMR" (structure/activity relationship by nuclear magnetic resonance).[6] In this spectroscopic approach, bound fragments are detected by NMR screening and subsequently linked together to increase affinity, as envisaged by Verlinde and coworkers.[2] Application of x-ray crystallography to detect and identify fragment hits was also pursued at Abbott.[7]

Fragment-based drug discovery has now been under way for more than a decade. Although Fesik and coworkers popularized the notion of linking fragments (as in their highly successful BCL-2 program), tactical emphasis appears to have largely shifted from fragment condensation to fragment engineering (or growing the fragment) to increase binding affinity and selectivity. Various biotechnology companies, including SGX Pharmaceuticals, Astex, and Plexxikon, have recently demonstrated that fragment-based approaches can indeed produce development candidates suitable for Phase I studies of safety and tolerability in patients (www.clinicaltrials.gov). Within many larger pharmaceutical companies, detection and optimization of fragments as a path to discovering new chemical entities appears to be gaining acceptance.

Before describing the SGX fragment-based drug discovery strategy, process, and lessons from human protein kinases, we review our current understanding of the nexus between chemical diversity of screening libraries and the challenges of compound screening that explain both the utility of starting the search for clinical development candidates with small fragments and the pharmaceutical industry's failure to realize the much-vaunted potential of combinatorial chemistry/high-throughput screening. Traditional drug discovery usually begins with a search for small molecule "hits" that demonstrate modest inhibition ($IC_{50} \sim 10\ \mu M$) of the molecular target in an in vitro biochemical assay. Promising hits are subsequently optimized into development candidates using iterative, trial-and-error methods and/or structure-guided design. The most commonly used approaches for finding hits have involved either high-throughput screening (HTS) of large compound libraries [typically 100,000–2,000,000 compounds with molecular weights (MW) of about 350–550] or opportunistic modification of substrate analogs and/or published active compounds. Although these methods have yielded a number of successfully marketed drugs, annual rates of new drug approval over the past decade have remained essentially unchanged despite dramatically increased research budgets at large pharmaceutical companies.

The fundamental shortcoming of any conventional HTS campaign derives from the reality that even the largest screening library exhibits only limited chemical diversity. For reference, the number of potential druglike molecules is predicted to be $\sim 10^{60}$,[8] which is comparable to accepted estimates for the total number of atoms comprising the universe. Such limitations effectively bias sampling of potential starting points for drug discovery and, therefore, may not yield the best lead series. Proprietary screening libraries are often biased toward certain structural classes, because these collections are composed of molecules synthesized for targets of historical importance rather than molecules chosen to sample leadlike chemical space.

Typical HTS libraries also consist of molecules (MW $\sim$350–550) larger than fragments (MW $< \sim$250), deliberately chosen because they yield more potent starting points for synthetic chemistry than do fragments (i.e., $IC_{50} < \sim 10\ \mu M$ versus $IC_{50} < \sim 10$ mM). Regrettably, subsequent optimization of these larger molecules is often complicated by the need to identify and remove functional groups to minimize molecular weight and hydrophobicity, while other functional groups must be added or modified to increase binding affinity. Thus, optimization of larger HTS hits into development candidates may require retrospective