disassembly into smaller pieces. Another significant weakness of the traditional drug discovery approach is the poor compliance of most screening hits with what we now recognize as being highly advantageous leadlike properties.[9–12] Such poor compliance frequently complicates and prolongs the lead optimization process and contributes, at least in part, to the relative paucity of new chemical entities plaguing the pharmaceutical industry. Finally, HTS hit rates tend to be very low (i.e., $<0.01\%$), which we now appreciate to be a direct consequence of screening libraries of compounds with MW ~350–550.[5]

Screening of small fragments (i.e., simple one- or two-ring heterocycles with MW $<$ ~250) instead of larger HTS library compounds overcomes all three of the shortcomings enumerated above. First, well-designed fragment libraries embody significantly more potential for chemical diversity than even the largest HTS libraries. For example, a 1,000-compound fragment library (each compound bearing two or more sites of chemical modification) can be readily elaborated into more than 10^8 accessible analogs (MW $<$ ~500). This extremely conservative estimate dwarfs even the largest HTS screening collections assembled within the pharmaceutical industry. Second, fragment libraries can be assembled with leadlike compounds exclusively, thereby maximizing the likelihood of successful optimization. Finally, small fragments exhibit an increased probability of binding to a given target (hit rates for many targets ~ 1–5%) versus larger, more complex molecules.[9] For a 1,000-compound fragment library such hit rates yield ten to fifty starting points for medicinal chemistry elaboration.

Such bounty does come at the expense of initial potency. Small fragment hits identified using various screening methods typically exhibit binding affinities of ~10 μM to 10 mM or even greater and are in some cases not measurable. The relatively weak character of fragment hits is deemed so unpalatable in some circles that adoption of fragment-based drug discovery approaches has been effectively inhibited within some organizations. Many medicinal chemists believe it more reasonable to attempt optimization from ~10 μM than an apparently 1,000-fold weaker starting point. Experience at both SGX and other fragment-based drug discovery companies has shown that 10 mM screening hits can in fact be rapidly optimized to better than 10 nM (see below).

The relationship between molecular weight and binding affinity has been explored intensively in various quarters. Astex has popularized the concept of ligand efficiency (LE),[13] which represents an indirect measure of the number of constituent atoms that participate in interactions with the target protein:

$$\mathrm{LE} = -\Delta G/(\text{no. of nonhydrogen atoms})$$
$$\sim -\mathrm{RT}\ln(\mathrm{IC}_{50})/(\text{no. of nonhydrogen atoms}).$$

In general, even weakly bound fragment hits are ligand efficient (LE $>$ 0.3), whereas most HTS hits are not (LE $<$ 0.3). During optimization of a fragment hit to generate nanomolar potency lead compounds, ligand efficiency can be monitored with the goal of maintaining/improving the balance between binding affinity and molecular weight (i.e., adding only atoms that contribute substantially to improved binding affinity).

As heralded by the MIT group's successful studies of protein crystals flooded with small organic solutes,[1] x-ray crystallographic screening has proven ideally suited to fragment-based drug discovery. The three-dimensional structure of the hit interacting with the target protein is the product of its detection. (Unlike NMR spectroscopy, x-ray crystallography provides a "direct look" at the protein/ligand complex.) A promising hit can be immediately qualified for further effort by establishing that it binds to the protein target in a well-defined orientation that is compatible with synthetic optimization. Such hits have been optimized by various structure-guided medicinal chemistry approaches.[13–16] Without three-dimensional structure validation and subsequent structural guidance, optimization of weakly bound fragments is extremely challenging because of the high propensity for nonspecific binding and false positives detected by biochemical assays. Application of x-ray crystallographic screening and/or cocrystallization with fragment hits has made fragment-based approaches to drug discovery both practical and successful.

SGX *FAST* FRAGMENT-BASED STRUCTURE-GUIDED DRUG DISCOVERY STRATEGY

Design of the SGX core fragment library

Recent studies of hit-to-lead optimization proposed a general definition of "leadlike" properties that increase the probability of successful optimization of hits to clinical candidates and successful prosecution of clinical development. These concepts have their origin in Lipinski's "rules,"[17] which describe properties of approved orally administered drugs: MW $<$ 500, calculated log P or Clog $P<5$, hydrogen-bond donors <5, and nitrogens+oxygens <10. Powerful though they are, Lipinski's rules are not appropriate for either HTS hits or evolving leads.[11,12] Hits usually increase in molecular weight, Clog P, and in number of rings and freely rotatable bonds during initial lead optimization and during subsequent development candidate optimization. Screening hits and leads should, therefore, be smaller than the molecular weight range embodied within Lipinski's rules. Teague et al.[12] initially proposed that leads should satisfy the following criteria: MW $<$ 350 and Clog $P<3.0$. Hann and Oprea[10] more recently proposed that leadlike molecules should have the following properties: MW $\leq$ 460, Clog $P \leq$ 4.2, freely rotatable bonds $\leq$ 10, $\leq$ 4 rings, hydrogen-bond donors $\leq$ 5, and hydrogen-bond acceptors $\leq$ 9.

"Leadlike" properties were originally proposed for molecules with binding affinities in the low micromolar range derived from HTS or combinatorial chemistry approaches. Fragment hits have binding affinities in the low micromolar to low millimolar range, thereby requiring