A new approach to target-specific virtual library screening

KeyFinder libraries have been developed by Key Organics in partnership with Prosarix to provide a cutting-edge service to support early-phase drug discovery.

Focused library approaches aim to provide an enriched set of synthesised compounds (of the order of ~10,000) for screening against a specific target family to which they have been designed. The design process can include many different approaches, sometimes as simple as incorporating privileged scaffolds. Overall, focused libraries claim to offer a more target-relevant chemical space than is offered by a generic HTS screen and can also enable a faster, more developed route to identify hits than a fragment-centric approach. Despite the merits of using focused libraries to increase hit rates, the chemical space offered is often somewhat limited. In contrast, target-centric virtual libraries can be designed that encompass a much larger chemical space (ie millions) which can provide increased novelty and, in combination with virtual screening protocols, allows inclusion of selectivity determinants at the hit identification stage, potentially shortening the subsequent optimisation phase of the project. Key Organics and Prosarix have developed a new target-focused virtual library approach, 'KeyFinder', that combines structure-based virtual screening protocols with virtual libraries, constructed only from in-house reagents, to provide rapid exploration of tractable chemistry space and follow-on synthesis of candidate compounds.

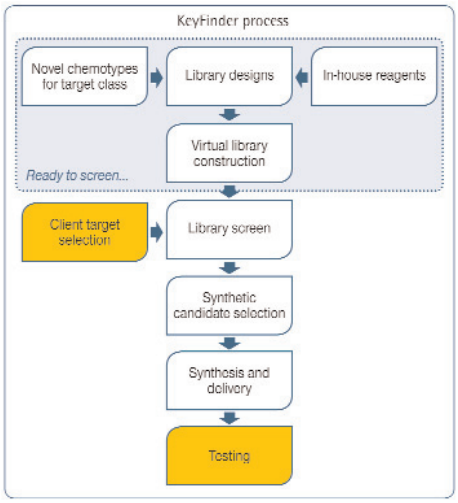

Fig 1. KeyFinder process. Virtual screening of pre-constructed virtual libraries is conducted to a target of interest leading to the identification of candidate compounds for synthesis.

Based in Camelford, UK, Key Organics is a leading provider of chemistry services, and Prosarix, based in Cambridge, UK, is a specialist provider of computational chemistry services. Prosarix has considerable experience in delivering hit-finding and lead optimisation capabilities to a wide range of targets/ indications. The two companies have worked together on a number of projects and have now formed a more formal partnership to provide KeyFinder libraries. These 'ready-to-run' virtual libraries with follow-on synthesis combine a number of advantageous features including a novel and large drug-like chemistry space designed for a specific target family; library designs based on in-house cores and templates, with tractable chemistry; virtual screening of the library against the desired pharmacology profile; design and synthesis of compounds; and a unique service capability for each target class with bespoke design and chemistry.

Library design is based on state-of-the-art modelling capabilities, using target structure and related family homology models, broad enough to be relevant for different receptor subtypes. The final chemotypeselection is based on predicted receptor affinity using a large panel of targets, resulting in many novel chemotypes in a final target class library.

KeyFinder provides considerable benefits when compared to conventional hit finding:

* Integrated design and synthesis workflows from inception
* Ready-to-run receptor models provide the basis for rapid in silico screening for the desired subtype pharmacology
* Tractable chemistry routes support rapid delivery of designed compounds
* The novel drug-like chemistry space enables expedient transition to the lead optimisation phase
* Patentable chemotypes enable IP generation
* The integrated in silico design and chemistry team enables the acceleration of projects
* A sound basis for rapid follow-on lead optimisation

The technology typically delivers 20 to 100+ compounds that are designed to have drug-like properties in quantities ranging from 1 to 50mg of material with purities of 95% or greater as determined by LCMS and NMR. The KeyFinder process is summaried in Fig 1.

Current libraries: S1PR library and Kinase library

To date, two KeyFinder virtual libraries have been designed, a smaller S1PR library and a larger Kinase library, both target classes being of current pharmaceutical interest, with each representing different design challenges. The S1PR is a highly targeted library to one set of closely related class A GPCRs. In contrast, the Kinase library required the inclusion of a much wider chemical space to reflect the structural diversity inherent in this large class of targets.

S1PR library: The sphingosine-1-phosphate receptor subfamily comprises five members (S1PR1-5). S1PR agonism has demonstrated therapeutic relevance for inflammatory conditions such as multiple sclerosis and Fingolimod is the first S1PR drug of this class to be approved. Significant research is now focusing on exploring the therapeutic potential of the different S1PR subtypes, with both selective or partially selective agonist and antagonists. S1P receptors are class A GPCRs and recently, in 2012, a crystal structure of S1P1 bound to an antagonist was solved.[1] This template enables the accurate modelling of antagonist states of S1P2-5 subtypes as well as predictions of the activated receptor conformations.

For the S1PR KeyFinder library, S1P2-5 receptor models were constructed and

several of these were validated by the docking performance of known modulators before being used to screen prospective chemotypes as fragments across the subtypes. Figure 2 shows the mapping of selectivity determinants in inactive S1P2-5 subtypes vs S1P1. High-scoring novel chemotypes have then been decorated using virtual library schemes with Key Organics' in-house reagents.

The S1PR library: consists of more than 800,000 enumerated compounds representing twelve different chemotypes. These compounds have been filtered to have molecular weights of less than 500 and PSAs of less than 140Å^2. The compounds are 100% drug-like by Lipinski criteria and 68% lead-like by Oprea criteria.

Kinase library: The structural kinome now consists of more than 2,500 deposited kinase structures, more than 1,200 of which are co-crystallised with ligands.[2] Despite this wealth of data, the design of highly selective inhibitors is still considered a major obstacle in kinase inhibitor design. Often, however, specificity aspects are not considered at the hit-finding stage. Therefore new in silico methodologies that incorporate predictions of selectivity to identify potential hits are advantageous.

The KeyFinder library construction applied to kinase inhibitor discovery combines a novel function (the Kscore) for scoring hinge-binding moieties with the generation of a synthetically tractable large virtual library (more than 14 million compounds) using available in-house reagents that include both known and novel cores. The KeyFinder process also includes structure-based virtual screening protocols to select compounds with the desired selectivity and potential for optimisation. Following selection, compounds can be rapidly synthesised for assay. The library has been validated in silico against observed kinase structural variation, e.g. DFG-in, DFG-out states and selectivity pockets.

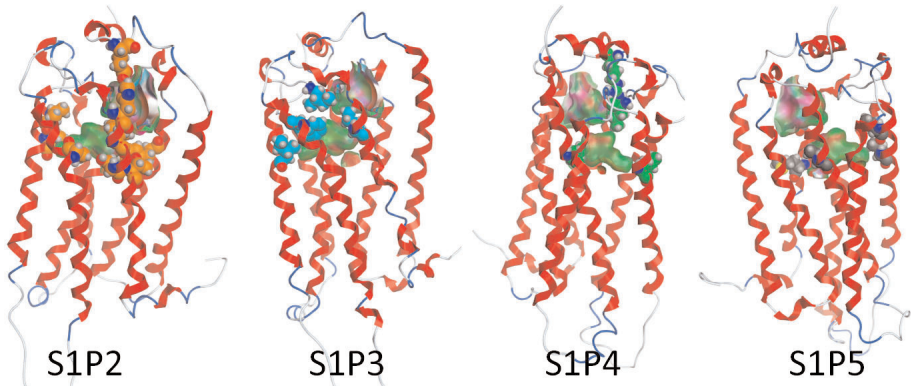

Fig 2. Specificity determinant mapping between S1P subtypes. Variations in residue identified between S1P1 and S1P2-5 are mapped onto homology models to determine areas of the binding pocket that may be utilised to select ligands with a desired selectivity profile.

The Kscore is a specific score for hinge-binding potential. Kinase inhibitors predominantly bind at the hinge region via one, two or three hydrogen bonds. Being able to determine the strength of these hydrogen bonds thus provides the basis for ranking the viability of hinge-binding fragments. A scoring function was developed based on Extended Huckel force-field in MOE (CCG Inc)[3] to score hinge-binding hydrogen bond strengths, mediated by conventional donor-acceptor pairs or by non-conventional hydrogen bonds, e.g. CH- or F-. This provides the basis for profiling cores against a diverse panel of kinases. Kscores for a substituted pyridine docked against a diverse panel of 118 kinases are shown in Fig 3.

Average Kscores across a kinase panel may be used to rank pan-kinase hinge-binding potential and scores for 89 different cyclic amides are shown in Fig 4.

Building the virtual library – core selection

Kscore was used to select reagents with high hinge binding potentials from a set of >9,000 available intermediates at Key Organics. Of the top ranking cores many represent privileged motifs observed within kinase patent literature. A recent Pfizer analysis of hinge binding cores determined their prevalence in compounds across different kinase targets.[4] Of the 20 most commonly observed cores, 17 are represented in the Kinase KeyFinder library. This suggests that some cores are privileged based on their high binding potentials. Importantly, many novel cores are also identified by the Kscore approach and form part of the library. A total of 137 hinge-binding cores were used as the basis for the current library.

The hinge-binding cores were used to construct a virtual library based on 30 different tractable chemistry schemes consisting of one, two or three reactions. The majority of the library (more than 90%) is formed by only two reactions (ignoring protecting/de-protecting steps). The computed library has the following characteristics: of the computed compounds, more than two million have a molecular weight lower than 400, representing a more lead-like set. The library straddles lead-like and drug-like space where on one hand the lead-like compounds are better starting points for optimisation, but the larger drug-like compounds are more likely to offer features required for selectivity, which was a major driver in the library design. Scoring using the Multi Parameter Optimisation (MPO) function[5] has identified a subset of the library with high CNS potential for targeting CNS kinases. Conversely, MPO scoring can be used to select a library subset with limited predicted CNS penetration.

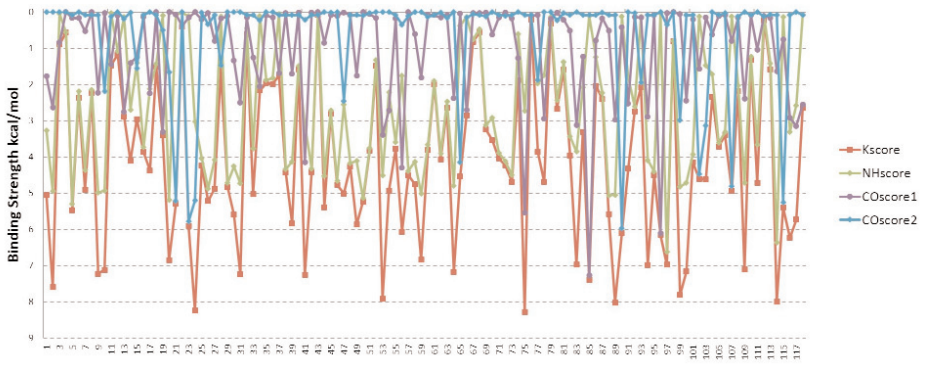

Fig 3. Profiling a substituted pyridine for kinase hinge-binding potential across a panel of 118 kinases.

Library validation

In order to assess the validity of the generated chemical space towards the observed diversity of kinase binding sites, a methodology was developed based on the recently described KLIF (kinase ligand interaction fingerprints) approach.[6] This approach describes kinase binding site variation according to the absence or

presence of features (i.e. the fingerprint) based on aligned kinase structures, and thus allows for rapid comparison of site features. An example of kinase binding site features is shown in Fig 5.

Structural examples of kinase site variation from the KLIF approach were determined as the basis for assessing the Kinase KeyFinder library (Table 1). Stringent pharmacophores were generated for each example, including features for selectivity pockets, in accordance with observed ligand-binding modes. Between 1% and 5% of the library was screened against each.

For each kinase structure tested, the structure-based pharmacophore identified hits. As such, this confirms the suitability of the library to yield hits throughout the kinome.

In summary

Large, synthetically tractable virtual libraries have been constructed for two target classes, S1PR and kinases. Both libraries have been subjected to in silico validation and for the kinase library this involved testing for its suitability to identify hits against structurally diverse kinase binding sites.

KeyFinder offers a rapid in silico screen of the libraries for the desired target profile of interest, leading to the identification of and synthesis of a focused library of hit candidates. Patentable chemotypes enable IP generation, and libraries, or particular chemotypes within the library, are accessible on a non-exclusive or exclusive basis. Follow-on lead optimisation support is available through Prosarix and Key Organics.

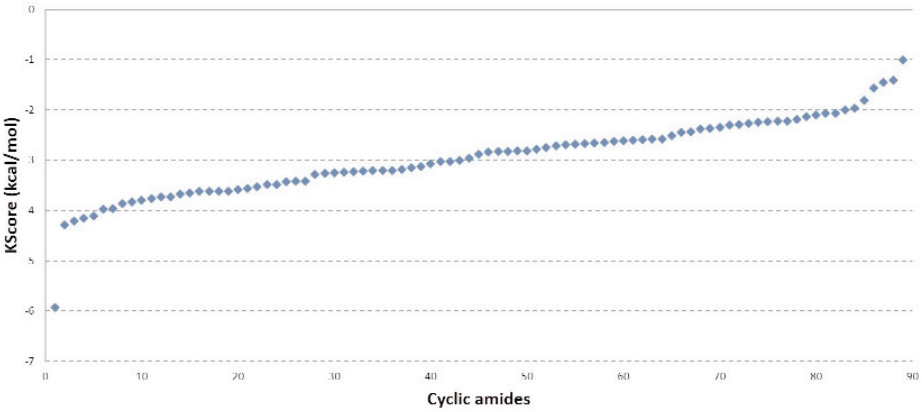

Fig 4. Average Kscores across a kinase panel may be used to rank pan-kinase hinge-binding potential.

PDB	Kinase	Site Features
2OJG	Erk2	DFG-in, Hinge/AP
2WGJ	c-Met	DFG-in, Hinge/AP/FP-I
3UE4	Abl1	DFG-out-like, Hinge/AP/BP-I-A&B
3B8R	VEGFR2	DFG-in, Hinge/AP/BP-I-B/BP-II-in
4DCE	ALK	DFG-in, Hinge/AP/BP-I-B/BP-II-in/BP-II-A-in
3U6H	c-Met	DFG-in, Hinge/AP/BP-I-A&B/BP-II-in/BP-II-B-in
3B8Q	VEGFR2	DFG-out, Hinge/AP/BP-I-A&B/BP-II-out
1T46	c-Kit	DFG-out, Hinge/AP/BP-I-A&B/BP-II-out/BP-IV
4ASD	VEGFR2	DFG-out, Hinge/AP/BP-I-B/BP-II-out/BP-III

Table 1. Kinase structures used to create pharmacophores for library generation to explore available pockets in kinases.

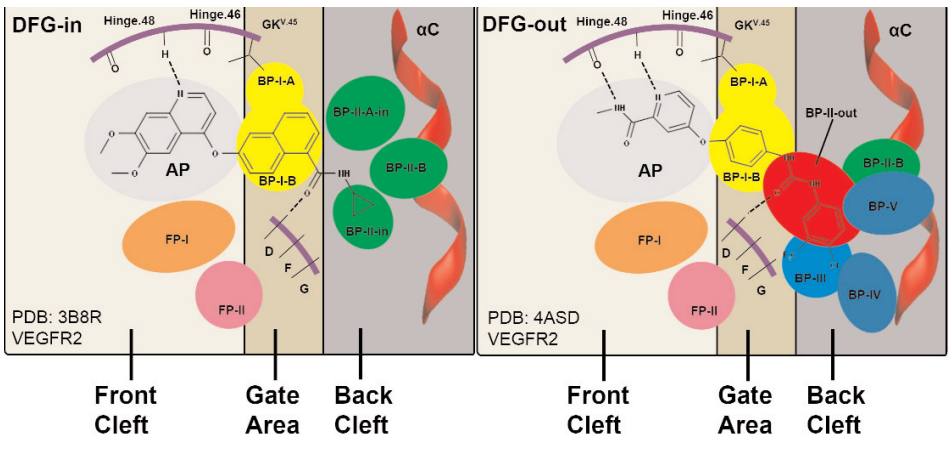

Fig 5. Schematic overview of available pockets in the kinase catalytic cleft exemplified with a DFG-in and DFG-out form of VEGF (adapted from ref 5).

REFERENCES

1. Hanson MA. et al (2012) Crystal structure of a lipid G protein-coupled receptor. Science 335:851-855.
2. http://www.rcsb.org/pdb/home/home.do
3. Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC H3A 2R7, Canada , 2013.
4. Xing et al. (2014) Scaffold mining of kinase hinge binders in crystal structure database, J. Comput. Aided Mol. Des. 28:13-23.
5. Wager et al. (2010) Moving beyond Rules: The Development of a Central Nervous System Multiparameter Optimization (CNS MPO) Approach To Enable Alignment of Druglike Properties, ACS Chem. Neurosci. 1:435-449.
6. Van Linden et al. (2014) KLIFS: A Knowledge-Based Structural Database To Navigate Kinase-Ligand Interaction Space, J. Med. Chem. 57(2):249-277.

Further information

Jon Heal
Prosarix Ltd
Newton Hall
Town Street, Newton
Cambridge CB22 7ZE
United Kingdom
Tel : +44 1223 870400
Email: jon.heal@prosarix.com
Web: www.prosarix.com

Steve Brough
Key Organics Ltd
Highfield Road Industrial Estate
Camelford, Cornwall PL32 9RA
United Kingdom
Tel: +44 1840 212137
Email: steveb@keyorganics.net
Web: www.keyorganics.net