Fragment screening: a comparison with other hit ID methods and challenges

Mike Hann
Director of Bio-Molecular Structure
Computational and Structural Chemistry
GSK Medicines Research Centre
Stevenage, UK

Mike.M.Hann@gsk.com
Five fortes of fragments

- The combinatorial explosion of chemistry space means that fragments can sample more of the available chemistry space at that level of complexity than is possible with more complex molecules.

- At lower complexity there is a higher probability of compounds matching the receptor even though they may be harder to detect. More complex molecules are more likely to have more “clashes” and thus do not fit.

- Medicinal chemists like to build molecules and so fragments are a great boot strap for structure based design. This plays to the strength of computational chemistry design

- By starting small and selecting the most Ligand Efficient compounds (eg DGbinding/number of heavy atoms), more Lead-like starting points are found which enhance the chances of successful Lead Optimisation campaigns.

- By reducing the number of pharmacophores in initial lead, only necessary interactions are built in to the compound as it is optimised. This should help ensure good developability properties of the resulting candidates.
The divergence of sampling rates of real compounds compared to the size of virtual chemistry space from the GDB* database at increasing levels of ligand complexity (as measured by the number of heavy atoms) – note the log scale.

The probabilities that ligands of different complexity (i.e., length) can match, be detected and the resultant “useful event”.

![Graph showing probability against length of ligand.](image)
# Fragment Screening

## Strengths
- Utilises the reduced complexity approach to increasing hit rate
- Focus on ligand efficiency
- Efficient sampling of chemical diversity.
- Enables structure based design and biophysics at outset
- Builds what you need

## Weaknesses
- Specialised methods needed to detect weak binding.
- Chemistry follow up needed to establish a lead quality molecule.
- Primarily limited to structure enabled targets
- Reductionist approach may over simplify complexity of interactions – i.e cooperativity lost
- Easy to squander a good hit
- Cost

## Opportunities
- Target tractability assessment
- Integration with other methods
- Increasing sensitivity of biophysics

## Threats
- More potent initial compounds from other methods.

---

Molecular complexity and fragment-based drug discovery: ten years on.
Curr Opin Chem Biol. 2011, 489-96. Leach AR, Hann MM.
### High Throughput Screening (HTS)

#### Strengths
- Diversity
- Proven track record
- Robustness
- Automation and miniaturisation
- Broadly applicable to biochemical and cellular assays

#### Weaknesses
- Compound collection maintenance
- False positives
- Cost – particularly infrastructure

#### Opportunities
- Synergy with other hit ID methods
- Spare capacity
- Label free detection methods (e.g., MS)
- Phenotypic screens

#### Threats
- Seen as expensive/slow but not so once infrastructure is in place.
- Perception as old technology!

---


### Encoded Libraries Technology – ELT*

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Huge numbers of compounds can be screened &gt; $10^9$</td>
<td>• Affinity only hence non-functional in absence of other assays</td>
</tr>
<tr>
<td>• Affinity selection and thus can give very potent compounds</td>
<td>• Aqueous chemistry only</td>
</tr>
<tr>
<td>• Minimal infrastructure cf HTS</td>
<td>• Complexity and size of molecules tends to be high</td>
</tr>
<tr>
<td>• Initial screening very quick</td>
<td>• Diversity is focussed around certain chemistries.</td>
</tr>
<tr>
<td></td>
<td>• Cost – chemistry making libraries and follow up off DNA to confirm hits.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Opportunities</th>
<th>Threats</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Target tractability assessment</td>
<td></td>
</tr>
</tbody>
</table>

*Design, synthesis and selection of DNA-encoded small-molecule libraries
Nature Chemical Biology 5, 647 - 654 (2009), Matt Clark, Barry Morgan et al
# Knowledge based screening

## Strengths
- In silico selects from widest diversity of tangible compounds in 2D or 3D
- Acoustic dispensing makes cherry picking easy from in house collections
- Good availability of compounds from suppliers

## Weaknesses
- Knowledge may be wrong or limiting!
- Docking and scoring poor
- Still need good assays!
- Cost of compounds if many purchased

## Opportunities
- Improvements in force fields and methods

## Threats
Challenges for Fragments

- With ever more sensitive detection methods how small should we go with the fragments we screen?
  - Re-exploring chemistry space is daunting if you get too elemental!
  - Non-additivity requires serendipidity to overcome so don’t go too small!

- The challenge of fragment evolution without structures to guide?

- Enabling selective Polypharmacology

- Thinking its easy and thus not applying sufficient rigour and discipline to evolving towards a candidate.

- Integration not isolation and competition – the real opportunity for all these methods.
Medicinal chemistry guidelines

- Consider the chemical tractability (ligandability) of the target, and if it is poor then investigate different mechanisms of action or different pathways.
- Select multiple, low-complexity polar starting points with high binding enthalpy, and optimize enthalpically towards the lead compound.
- Select appropriate metrics for multidimensional optimization; use ligand efficiency and lipophilic efficiency metrics in hit-to-lead optimization and change to more complex metrics emphasizing dosage to support lead optimization.
- Evaluate available chemistries when entering extensive optimization; prepare what you designed and really want rather than what you can readily synthesize; design, synthesize and use proprietary building blocks rather than depend on chemistry catalogues.
- Do not be afraid to retrench to a series of lower potency if it has better physicochemical properties, particularly solubility; leave suboptimal scaffolds early; extensive optimization of a scaffold that is not amenable to achieving a desirable balance of potency and ADME (absorption, distribution, metabolism and excretion) properties is likely to be a waste of time and resources.
- Stay focused on the ‘sweet spot’ and committed to deliver high-quality compounds, but remain open-minded to the many ways this can be achieved.
- Resist timelines that compromise compound quality.

The challenge of medicinal chemistry – the role for nature and nurture in lead discovery and optimization
Acknowledgements

- Many at GSK present and past: 
  **Darren Green & Andy Pope**
  Andrew Leach, Paul Leeson, 
  Andy Brewster, Ian Churcher, 
  James Butler, Rob Young, 
  Alan Hill, Stephen Pickett, et al
- Anne Hersey, John Bradshaw, 
  Gavin Harper, Paul Gleeson, 
  Malcolm Weir, Harren Jhoti, 
  Barry Morgan
Back ups
A final thought on Fragment Based Drug Discovery

...and we can reduce our carbon footprint too, if we make them smaller and less lipophilic.
Finding the sweet spot in medicinal chemistry

Michael M. Hann and György M. Keserü

Given its position at the heart of small-molecule drug discovery, medicinal chemistry can have a key role in tackling the well-known productivity challenges in pharmaceutical research and development. In recent years, extensive analyses of successful and failed drug compounds have improved our understanding of the role of physicochemical properties in drug attrition, and clarified the difficulties in finding the ‘sweet spot’ in lead discovery and optimization. This figure summarizes suggested guidelines for improving compound quality.

Cultural changes are difficult and need supporting throughout an organization and at all phases. Exchange of views and access to data-driven knowledge are important and will help reduce ‘behavioural’ ‘addictions’. In-depth analysis of drug discovery case studies and success stories contributes substantially to those efforts, and towards the understanding of true compound quality.

Select multiple, low-complexity polar starting points with high binding affinities, and optimize methylylation towards the lead compound.

Select appropriate metrics for multidimensional optimization: use ligand efficiency and lipophilicity efficiency metrics in hit-to-lead optimization and change to more complex metrics emphasizing damage to support lead optimization.

On can be lead to revert to a series of lesser potency if it has better physicochemical properties. Extreme optimization of a scaffold that is not amenable to achieving a reasonable balance of potency and ADME properties is likely to be a waste of time and resources.

Consider the chemical tractability (lipophilicity) of the target, and, if possible, investigate different mechanisms of action or different pathways.

Evaluate available chemistries when targeting specific optimization: prepare what you designed and really want rather than what you can readily synthesize; design, synthesize and use proprietary building blocks rather than depend on chemistry catalogues.

Stay focused on the “sweet spot” and committed to deliver high-quality compounds, but remain open-minded to the many ways this can be achieved!

It’s the dose, stupid!!

- Concepts such as Drug Efficiency tells us how much of the dose actually is available in the biophase of interest.

\[ DRUGeff = \text{Biophase Concentration} \times \frac{100}{\text{Dose}} \]


- And more recently the use of Drug Efficiency Index as a strategy towards low therapeutic dose

\[ \text{DEI} = \log[\text{DRUGeff(%)}] + \text{pKd} \]

DEI is a correction of the *in vitro* affinity by the *in vivo* pharmacokinetic potential.

It is a simple descriptor directly connected to efficacy and therapeutic dose with the potential to probe the balance between *in vitro* affinity and ADME properties.

*Application of drug efficiency index in drug discovery: a strategy towards low therapeutic dose. Montanari, Dino; Chiarparin, Elisabetta; Gleeson, Matthew Paul; Braggio, Simone; Longhi, Raffaele; Valko, Klara; Rossi, Tino. Expert Opinion on Drug Discovery*, Volume 6, Number 9, September 2011, pp. 913-920(8)
What we have come to know (or rediscover!)

- Large and lipophilic molecules are increasingly seen once more as bad!
- Lipinski’s 500/5 for oral bioavailability is increasingly seen as too lenient when it comes to the wider ADMET issues.
- We should be thinking 400/4 as a ceiling for better developability of drugs.

The expanding “sciences” of Medicinal Chemistry

- In hindsight, the rush to numbers as a solution to productivity obscured our collective memory and experience!
### Table 2
Percentages of compounds achieving defined target values in the various developability assays categorised by PFI or iPFI bins

<table>
<thead>
<tr>
<th>Assay / target value</th>
<th>&lt;3</th>
<th>3-4</th>
<th>4-5</th>
<th>5-6</th>
<th>6-7</th>
<th>7-8</th>
<th>8-9</th>
<th>9-10</th>
<th>&gt;10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solubility &gt;200 μM</td>
<td>89</td>
<td>83</td>
<td>72</td>
<td>58</td>
<td>33</td>
<td>13</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>%HSA &lt;95%</td>
<td>88</td>
<td>80</td>
<td>74</td>
<td>64</td>
<td>50</td>
<td>30</td>
<td>17</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>2C9 pIC₅₀ &lt;5</td>
<td>97</td>
<td>90</td>
<td>83</td>
<td>68</td>
<td>48</td>
<td>32</td>
<td>23</td>
<td>22</td>
<td>38</td>
</tr>
<tr>
<td>2C19 pIC₅₀ &lt;5</td>
<td>97</td>
<td>95</td>
<td>91</td>
<td>82</td>
<td>67</td>
<td>52</td>
<td>42</td>
<td>42</td>
<td>56</td>
</tr>
<tr>
<td>3A4 pIC₅₀ &lt;5</td>
<td>92</td>
<td>83</td>
<td>80</td>
<td>75</td>
<td>67</td>
<td>60</td>
<td>58</td>
<td>61</td>
<td>66</td>
</tr>
<tr>
<td>Clint &lt;3 ml/min/kg</td>
<td>79</td>
<td>76</td>
<td>68</td>
<td>61</td>
<td>54</td>
<td>42</td>
<td>41</td>
<td>39</td>
<td>52</td>
</tr>
<tr>
<td>Papp &gt;200 nm/s</td>
<td>20</td>
<td>30</td>
<td>46</td>
<td>65</td>
<td>74</td>
<td>77</td>
<td>65</td>
<td>50</td>
<td>33</td>
</tr>
<tr>
<td>hERG pIC₅₀ &lt;5 (+1 charge)</td>
<td>86</td>
<td>93</td>
<td>88</td>
<td>70</td>
<td>54</td>
<td>36</td>
<td>29</td>
<td>21</td>
<td>11</td>
</tr>
<tr>
<td>Promiscuity &lt;5 hits with pIC₅₀ &gt;5</td>
<td>85</td>
<td>78</td>
<td>74</td>
<td>65</td>
<td>49</td>
<td>30</td>
<td>20</td>
<td>13</td>
<td>7</td>
</tr>
</tbody>
</table>

PFI = mChrom log Dₐₜₗ₄ + #Ar

iPFI = mChrom log Pₐₜₗ₄ + #Ar

Sweet spots for absorption and potency are in conflict with sweet spots for other desirable properties?
The Knowledge Plot and how it influences Screening Strategies

The need for diversity is inversely proportional to the knowledge that you have on targets.

DIVERSITY NEEDED in screening sets.
The properties of drug molecules are ultimately a compromise between many different and often competing characteristics.