Field-based virtual screening: New trends to increase the chemical diversity of your leads

Alessandro Deplano*¹, Javier Vázquez¹, Albert Herrero¹, Enric Gibert¹, Enric Herrero¹, F. Javier Luque ²

¹ Pharmacelera, Plaça Pau Vila, 1, Sector 1, Edificio Palau de Mar, Barcelona 08039, Spain.
² Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Sciences, Institute of Biomedicine (IBUB), and Institute of Theoretical and Computational Chemistry (IQTC-UB), University of Barcelona, Av. Prat de la Riba 171, Santa Coloma de Gramenet E-08921, Spain.

* Corresponding author: alessandro.deplano@pharmacelera.com
Field-based virtual screening: New trends to increase the chemical diversity of your leads

Graphical abstract:

Virtual Screening: A Way To Reduce Experimental Costs

Chemical space filtering: only those compounds with chances to become a hit are tested experimentally
Abstract:

Computational chemistry methods can significantly reduce experimental costs in early stages of a drug development project by filtering out unsuitable candidates and discovering new chemical matter. Molecular alignment is a key pre-requisite for 3D similarity evaluation between compounds and pharmacophore elucidation. Relying on the hypothesis that the variation in maximal achievable binding affinity for an optimized drug-like molecule is largely due to desolvation, we explore herein a novel small molecule 3D alignment strategy that exploits the partitioning of molecular hydrophobicity into atomic contributions in conjunction with information about the distribution of hydrogen-bond donor/acceptor groups in each compound. A brief description of the method, as implemented in the software package PharmScreen, is presented. The computational procedure is calibrated by using a dataset of 402 molecules pertaining to 14 distinct targets taken from the literature and validated against the CCDC AstraZeneca test set of 121 experimentally derived molecular overlays. The results confirm the suitability of MST based-hydrophobic parameters for generating molecular overlays with correct predictions obtained for 100%, 93%, and 55% of the molecules classified into easy, moderate and hard sets, respectively. The potential of this tool in a drug discovery campaign is then evaluated in a retrospective study with the aim to evaluate the correlations between activities and similarity score of a series of sigma-1 receptor ligands. The results confirm the suitability of the tool for Drug Discovery purposes finding the 67% of the most active ligands (≤10 nM) in Q1 of the ranking and the most active compound in position five.

Keywords: Drug Discovery; Virtual Screening; Molecular Alignment; Ligand-based; Hydrophobicity
Speech Goals

• Present the virtual screening techniques and how they can help finding better leads with high chemical diversity respect the reference structure.
  – Hydrophobicity in CADD
  – The value of considering multiple fields (electrostatic, steric and hydrophobic) when performing molecular alignment and virtual screening
  – The importance of finding chemical diversity using in-silico technologies
  – Case study
Which Two Are More Similar?

Strawberry  Orange  Basketball

There is no single measure of similarity:

“What is the essence of a molecule? What is it made of? What will it do?”
Structurally similar molecules tend to have similar properties:

Problem: Subjective concept, with multiple ways of defining similarity
- 1D, 2D or 3D descriptors
- The weighting of these descriptors
- Mathematical expression of the similarity function.

3D-based similarity methods:

NONSUPERPOSITIONAL

The analysis of atomic distances to a set of reference positions

SUPERPOSITIONAL

Correct alignment is critical
Hydrophobicity vs Binding Affinity And Activity

A correlation emerges between the $\text{pIC}_{50}$/ $\text{pK}_{i}$ and the global hydrophobic similarity index


The defined druggability model assumes that favorable drug binding is largely driven by the hydrophobic effect
Can We Adopt Only Hydrophobic Descriptors?

Previous implementation based on **empirical hydrophobic descriptors**

- **Molecular Lipophilicity Potential (MLP)**
  Combines empirical fragmental contribution to lipophilicity with a distance-dependent function.

- **Hydropathic INTeractions (HINT) scoring function**
  Rank compounds according to hydrophobic complementarity

Our Strategy: Atomic-Level Contributions To Hydrophobicity

**MST Model**

*Derived from the Quantum Mechanical IEF/PCM-MST Solvation Models*

*Partitioning of the solvation free energy in the MST continuum models.*

\[
\Delta G_{\text{sol}} = \Delta G_{\text{ele}} + \Delta G_{\text{cav}} + \Delta G_{\text{vW}}
\]

\[
\Delta G_{\text{sol}} = \sum_{i=1}^{N} \Delta G_{\text{sol},i} = \sum_{i=1}^{N} (\Delta G_{\text{ele},i} + \Delta G_{\text{cav},i} + \Delta G_{\text{vW},i})
\]

**Atomic Contribution to Log P**

\[
\log P_X = \sum_{i=1}^{N} \log P_{X,i} = \sum_{i=1}^{N} -\frac{\Delta G_{X,i}^{\text{vW}}}{2.303RT} \quad (X: \text{ele, cav, vW})
\]

\[
\log P_{\text{total}} = \log P_{\text{ele}} + \log P_{\text{cav}} + \log P_{\text{vW}}
\]

*Electrostatic contributions*

*Non electrostatic contributions*

---

Why Use QM-Based Methods?

- The atomic contribution is influenced by the whole molecule
  - Take into account conformation impact
  - Model new chemical groups not present in empirical databases


\[
\begin{align*}
\text{HOOC} & \quad +1.1 \\
+2.2 & \\
\text{NO}_2 &
\end{align*}
\]
Hydrophobic Descriptors Validated for QSAR

  - Hydrophobic fields usage in QSAR studies

- T. Ginex¹, J. Muñoz-Muriedas², E. Herrero³, E. Gibert³, P. Cozzini⁴, F. J. Luque¹, “Application of the Quantum Mechanical IEF/PCM-MST Hydrophobic Descriptors to Selectivity in Ligand Binding”, Journal of Molecular Modelling (JMM), June 2016
  - Hydrophobic fields usage in selectivity evaluation
PharmScreen: MST-based Alignment

Parameter calculation

Expansion center and tensors calculation

Alignment pool

Similarity Function

Tanimoto

Tversky

Final Alignment

\[ Q = \sum_{i=1}^{N} \log P_{o/w, j} \left( 3 \bar{r}_i \cdot \bar{r}_j - |\bar{r}_i|^2 |\bar{r}_j|^2 \right) \]

\[ I = \sum_{i=1}^{N} \log P_{o/w, j}^{cav} \left( |\bar{r}_i|^2 1 - \bar{r}_i \cdot \bar{r}_j \right) \]

Molecular Fields are agnostic to chemotypes
Better Ligand-Receptor Interaction Model

Traditional fields (Shape – Electro)  Ref overlay  Crystal overlay  PIM-1 INHIBITORS ALIGNMENT

PharmScreen interaction fields

PharmScreen fields better represent ligand-protein interactions vs traditional fields
PharmScreen Provides Superior Alignment

AZ / CCDC Dataset:
1456 crystal structures from 121 receptors
PharmScreen Provides Superior Alignment

AZ / CCDC Dataset:
1456 crystal structures from 121 receptors

<table>
<thead>
<tr>
<th></th>
<th>Easy</th>
<th>Moderate</th>
<th>Hard</th>
<th>Unfeasible</th>
</tr>
</thead>
<tbody>
<tr>
<td>AstraZeneca</td>
<td>95%</td>
<td>73%</td>
<td>39%</td>
<td>0%</td>
</tr>
<tr>
<td>MolAlign</td>
<td>100%</td>
<td>76%</td>
<td>54%</td>
<td>0%</td>
</tr>
<tr>
<td><strong>PharmScreen</strong></td>
<td><strong>100%</strong></td>
<td><strong>96%</strong></td>
<td><strong>72%</strong></td>
<td><strong>12.5%</strong></td>
</tr>
</tbody>
</table>
Do These Descriptors Provide The Same Overlays?

Percentage of equal overlays between hydrophobic/HB and steric/electrostatic fields

Generated overlays differ significantly for complex cases highlighting the complementarity of both approaches
**Study**

**Project goal:** Virtual screening quality evaluation.
Explore correlations between activities and molecular similarity.

**Data:**
- 174 sigma-1 receptor ligands from existing publications analyzed
- Public external references from RCSB Protein Data Bank: 5HK1 and 5HK2

**Workflow:**
- Library preparation
  - Generation 3D structure, isomers, tautomers and conformers of the molecules (~20,000 total molecules).
- As reference was used a ligand from a crystal structure external to the papers.
- Virtual screening with PharmScreen using hydrophobic and hydrogen bonds fields.

High Correlation PharmScreen Ranking And Active Hits

- Ligands with higher activity found in the initial results
  - Molecule with highest activity in position 5 of the VS ranking
- Molecule from the existing patent in position 15 of the VS ranking

67% of the active ligands (activity ≤10 nM) are in Q1

42% of the molecules with an activity between 10 nM and 100 nM are in Q1
**Project goal:** Virtual screening quality evaluation.

Explore how much chemical diversity can be retrieved

**Data:**
- 11 sets from Directory of Useful Decoys$^{1,2}$

**Workflow:**
- Use the reference structure provided in the dataset
- Virtual screening with PharmScreen using hydrophobic and hydrogen bonds fields.
- Compute weighted ROC curves and ROC enrichment$^3$

---

<table>
<thead>
<tr>
<th>Set</th>
<th>Actives</th>
<th>Decoys</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACE</td>
<td>46</td>
<td>1796</td>
</tr>
<tr>
<td>AChE</td>
<td>99</td>
<td>3859</td>
</tr>
<tr>
<td>CDK2</td>
<td>47</td>
<td>2070</td>
</tr>
<tr>
<td>COX-2</td>
<td>212</td>
<td>12606</td>
</tr>
<tr>
<td>EGFr</td>
<td>365</td>
<td>15560</td>
</tr>
<tr>
<td>Fxa</td>
<td>64</td>
<td>2092</td>
</tr>
<tr>
<td>HIVRT</td>
<td>34</td>
<td>1494</td>
</tr>
<tr>
<td>InhA</td>
<td>57</td>
<td>2707</td>
</tr>
<tr>
<td>P38</td>
<td>137</td>
<td>6779</td>
</tr>
<tr>
<td>PDGFrb</td>
<td>124</td>
<td>5603</td>
</tr>
<tr>
<td>VEGFr2</td>
<td>74</td>
<td>2647</td>
</tr>
</tbody>
</table>

---


PharmScreen Finds More Chemical Diversity

- **Virtual Screening** for 11 DUD sets (active hits clustered in families)

PharmScreen Finds More Chemical Diversity

- COX-2 (PDB: 1cx2), Cyclooxygenase-2 (prostaglandin synthase-2) study
  - 12818 compounds – 212 actives in 44 families

### Families found

<table>
<thead>
<tr>
<th>Method</th>
<th>Families found</th>
</tr>
</thead>
<tbody>
<tr>
<td>PharmScreen</td>
<td>9</td>
</tr>
<tr>
<td>FieldScreen</td>
<td>5</td>
</tr>
<tr>
<td>FieldScreen+P</td>
<td>6</td>
</tr>
<tr>
<td>2SHA</td>
<td>3</td>
</tr>
<tr>
<td>DOCK</td>
<td>2</td>
</tr>
<tr>
<td>OAAP</td>
<td>6</td>
</tr>
<tr>
<td>OAK</td>
<td>3</td>
</tr>
<tr>
<td>OAK_Flex</td>
<td>3</td>
</tr>
<tr>
<td>MACCS</td>
<td>4</td>
</tr>
</tbody>
</table>

Families found among first 50 structures

Active Structures found only by PharmScreen

3 more families found
Summary

- **Virtual Screening:**
  - Reduces the search space in initial drug discovery stages
  - Can provide **significant savings** in a drug discovery project

- **Pharmaceulera**'s field-based virtual screening technology:
  - **Full 3D representation of all relevant fields of interaction (shape, electrostatic and hydrophobic)** for molecular alignment AND similarity
  - Atomic-level LogP partitioning with semi-empirical quantum mechanical solvation models

Interaction fields are chemotype agnostic → more chemical diversity found
Thank you very much!