Using Chemical and Biological Data, In Particular Applied to Selecting Small Molecules to Increase Thermal Stability Of a Monoclonal Antibody

Andreas Bender, PhD
Natural Philosopher for Molecular Informatics, University of Cambridge
Fellow of King’s College, Cambridge
CTO, Healx
Outline

- Chemical and biological data – what is out there (and how can we use it?)

- Case studies from mode-of-action analysis and compound selection

- Application of informatics methods to select small molecules for the thermal stabilization of antibodies
Core Data Considered: Chemistry, Phenotype, Targets / Mode of Action
So what’s the point of it all? We would like to answer questions!

- “What is the reason upon treatment with A for phenotypic effect B?”
  -> Mode of Action

- “Which compound should I make to achieve effect C in a biological system?”
  -> Chemistry

- “Does patient D or patient E respond better to drug F?”
  -> Phenotype / Phenotype Change
More generally, where can we ‘model properties’?

- Where output property is determined eg by structure (input space)
- In principle *any* type of property that is a function of input
- Can be either data- or model-driven
- The more data, the better

- For us, the most interesting part is link between chemical structure, and biological effect (which descriptors, models, … to use)
Group Research Organized in Clusters
(Numbers = number of people working on project)

Mode-of-action analysis
- Mode-of-action analysis (‘target prediction’) (~7)
- Modelling bioactivities on target families (~2)

Modelling compound mixtures, traditional medicines
- Mixture modelling (~6; ERC Starting Grant)
- Traditional medicines/natural products (~3)

Integrating chemical and biological data
- Pharmacogenomics/toxicogenomics (~2)
- Gene expression/RNA-Seq data for compound selection (stem cell differentiation), mode of action analysis (~3)
Starting from *in vivo* efficacy we can predict the MoA, based on ligand chemistry.

Exploiting known bioactivity data for new decisions: Target predictions

- The models enable automated prediction of the targets or target families of orphan ligands given only their chemical structures.

Chemogenomics Database

- Ligand 1—Target 1
- Ligand 1—Target 2
- Ligand 2—Target 2
- ...
- Ligand N—Target N

Public model with AZ: Mervin et al. J Cheminf. 2015

Orphan compound ➔ Target Class Models ➔ Predicted Targets
Prediction Examples: Gleevec, Ruboxistaurin

- Gleevec (Novartis),
  - Launched
  - Targets Bcr-Abl, c-kit, PDGFRb

- Ruboxistaurin (Lilly/Takeda), Phase III
  - PKCb
Understanding rat sleep data

- Project with Eli Lilly
- Male Wistar rats

Treated with ~500 sleep-inducing compounds, dozens of readouts from EEG/EMG, Abdominal Minimitter, Cage that define ‘good sleep’

- Q: What are bioactivity profiles associated with compounds inducing good sleep?

- Going from single to multiple targets (polypharmacology), and from single to multiple simultaneous MoA hypotheses for given phenotype

Work by Georgios Drakakis
Decision trees learn receptor bioactivity profiles associated with ‘good’ and ‘bad’ sleep.
Prospective validation on both target and phenotypic level

- 7 marketed drugs/drug combinations were selected which are predicted to modulate sleep, are dissimilar to the training set, but were not annotated with this side effect

- 5 out of 7 marketed drugs (71%) tested increased sleep parameters (a sixth led to hyperactivity!)

- 21 out of the 27 predicted targets (78%) were validated

- Overall 78% correct on target level, ~71% on phenotypic level (across 4 MoA classes)
Combined gene expression / target prediction analysis for MoA analysis and compound selection

- Select compounds based *both* on gene expression and target prediction profiles
- Eg for stem cell differentiation

KalantarMotamedi *et al.* *Cell Death Discovery* 2016
Selected compound induces differentiation of stem cells into cardiac myocytes (by RT-PCR; work with Dr Nasr, Royan Institute, Isfahan)

Application of informatics methods to select small molecules for the thermal stabilization of antibodies

- Experimental work of Olubukayo-Opeyemi Oyetayo and Hans Kiefer, Biberach University of Applied Sciences; modelling performed by Oscar Mendez-Lucio (Cambridge)

- “Diversity selection, screening and quantitative structure-activity relationships of osmolyte-like additive effects on the thermal stability of a monoclonal antibody”

Aim

- Additives can contribute to the thermal stability of an antibody
- However, systematic relationships between structure and effect are usually unknown
  - Unspecific vs covalent interactions
  - Direct interactions vs altering water structure
  - Interaction with peptide backbone vs interactions with side chains (general vs protein-specific effects)
  - ...
Informatics contribution

- Hence, we used informatics methods to
  - Select a chemically diverse library (from given compound classes) before experiments
  - Generated structure-activity relationships after experiments to
    • Correlate/explain/understand stabilization effects observed
    • Select next round of stabilizing compounds with improved properties
Selection of diverse compound library to determine Ab stabilizing properties

- Amino acids, methylamamines and polyols
- Molecular weight < 300 (< 500 for polyols); sarcosine and mannitol used as queries for the methylamine and polyol class to identify similar compounds (>0.5, MACCS keys)
- Jarvis-Patrick clustering; diverse cluster centres selected
- Removed reactive/toxic compounds (according to MSDS)
- Solubility > 0.1M
- 84 compounds (29 amino acids, 18 methylamines, 37 polyols)
Methods: Antibody, readouts

- Recombinant human monoclonal antibody of the IgG1 subclass (mAb1) was produced in-house in CHO cells
- To determine unfolding differential scanning fluorimetry (DSF) was used
  - High throughput method
  - Non-equilibrium method though
  - Hence impact of extrinsic fluorescent probe on $T_m$, inability to measure reversibility of unfolding transitions
- Lowest observed thermal melting transition measured
Determining impact of pH: mAb1 in buffer at different pH

- Mostly Unfolded at pH 3; large pH impact
- pH 3.5 two melting transitions, at 6.5 one
- Hence osmolytes tested at pH 3.5 and 6.5
25% of methylamines, 50% of amino acids, 75% of polyols act as stabilizers (at both pH 3.5 and 6.5)

Charged compounds (amino acids, methylamines) show pH effect, much less so polyols!
Data used for QSAR model generation

- Measurements at pH 6.5 showed less error than at pH 3.5
- pH 6.5 also more relevant for practical processing steps, hence data obtained at this pH was used for QSAR model generation
QSAR model: Partial Least Squares (PLS)

- For 84 compounds 2D descriptors were calculated using MOE software.
- Removal of descriptors with low variance, normalization.
- TS potency at pH 6.5 used as output variable.
- Models were fit to all data points, model consistency and variable importance determined in leave-one-out cross-validation.
- Variable importance determined using ‘Variable Importance Projection’ (VIP).
Tm model fit across the amino acid, methylamine and polyol classes (‘global model’)

- RMSE = 4.77
- RMSE (LOO) = 6.07
Local models give much better correlations: Methylamines

R-Square = 0.997
Adj. R-Square = 0.997
RMSE = 0.398
Local models give (somewhat) better correlations: Amino acids
Local models give (somewhat) better correlations: Polyols
## Model statistics: Leave-one-out validation

<table>
<thead>
<tr>
<th>Model Statistic</th>
<th>Amino acids</th>
<th>Methylamin</th>
<th>Polyols</th>
<th>Global</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>29</td>
<td>18</td>
<td>37</td>
<td>84</td>
</tr>
<tr>
<td>Descriptors</td>
<td>63</td>
<td>56</td>
<td>51</td>
<td>60</td>
</tr>
<tr>
<td>R-Squared value</td>
<td>0.864</td>
<td>0.97</td>
<td>0.906</td>
<td>0.679</td>
</tr>
<tr>
<td>Adj. R-Squared value</td>
<td>0.848</td>
<td>0.967</td>
<td>0.891</td>
<td>0.645</td>
</tr>
<tr>
<td>RMSE</td>
<td>5.34</td>
<td>1.15</td>
<td>2.04</td>
<td>6.07</td>
</tr>
<tr>
<td>Components</td>
<td>9</td>
<td>10</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Variable Importance Projection: Polarity/hydrophobicity, accessible surface area are crucial

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
<th>VIP</th>
<th>Regression Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>SlogP_VSA7</td>
<td>Sum of the accessible surface area (in Å²) over all atoms $i$ such that SlogP of atom $i$ is in (0.25, 0.30]</td>
<td>2.027</td>
<td>-1.127</td>
</tr>
<tr>
<td>PEOE_VSA-1</td>
<td>Sum of the accessible surface area (in Å²) over all atoms $i$ such that the partial charge of atom $i$ is in [-0.10, -0.05)</td>
<td>1.929</td>
<td>0.678</td>
</tr>
<tr>
<td>logS</td>
<td>Log of aqueous solubility (mol/L)</td>
<td>1.798</td>
<td>-0.522</td>
</tr>
<tr>
<td>a_aro</td>
<td>Number of aromatic atoms</td>
<td>1.726</td>
<td>-0.438</td>
</tr>
<tr>
<td>b_ar</td>
<td>Number of aromatic bonds</td>
<td>1.726</td>
<td>-0.351</td>
</tr>
<tr>
<td>Q_VSA_FPOL</td>
<td>Fractional polar Van der Waals surface area</td>
<td>1.709</td>
<td>-0.351</td>
</tr>
<tr>
<td>Q_VSA_FHYD</td>
<td>Fractional hydrophobic Van der Waals surface area</td>
<td>1.709</td>
<td>-0.345</td>
</tr>
</tbody>
</table>
Also direct correlation of thermal shift with hydrophobicity/polarity parameters
Summary of antibody stabilization work

- Informatics methods were able to help us select diverse compounds.
- We were able to generate a model, which could be used two-fold:
  - To gain insight into parameters relevant for Ab stabilization (however, be aware of causality vs correlation, also multiple parallel effects are difficult to discriminate).
- For the selection of new compounds with improved properties.
Application of informatics methods to select small molecules for the thermal stabilization of antibodies

- Experimental work of Olubukayo-Opeyemi Oyetayo and Hans Kiefer, Biberach University of Applied Sciences; modelling performed by Oscar Mendez-Lucio (Cambridge)

- “Diversity selection, screening and quantitative structure-activity relationships of osmolyte-like additive effects on the thermal stability of a monoclonal antibody”

Acknowledgments

Avid Afzal
Azedine Zoufir
Ben Alexander-Dann
Bobby Glen
Chad Allen
Christoph Schlaffner
Dan Mason
Dezso Modos
Erin Oerton
Fatima Baldo
Fredrik Svensson
Fynn Krause
Georgios Drakakis
Kathryn Giblin
Keerthy Kumar
Krishna Bulusu
Leen Kalash
Lewis Mervin
Nitin Sharma
Ranjoo Choi
Richard Lewis
Samar Mahmoud
Salundi Basappa
Sharudul Paricharak
Sharif Siam
Siti Zuraidah Sobir
Stephanie Ashenden
Xianjun Fu
Yasaman Motamedi
Yue Kong

Sebastian Rohrer
Stefan Tresch
Antje Wolf
Klaus-Juergen Schleifer

AstraZeneca

Ola Engkvist
Thierry Kogej
Jay Mettetal
Samantha Hughes

Ian Stott

Hinrich Goehlmann
Herman van Vlijmen
Joerg K. Wegener

Lilly

Mike Bodkin
David Evans
Keith Wafford
Suzanne Brewerton

Institut Pasteur

Therese Malliavin
Isidro Cortes

Massimo Mercati
Anna Maidecchi
Jacopo Lucci