AstraZeneca
IMED Biotech Unit

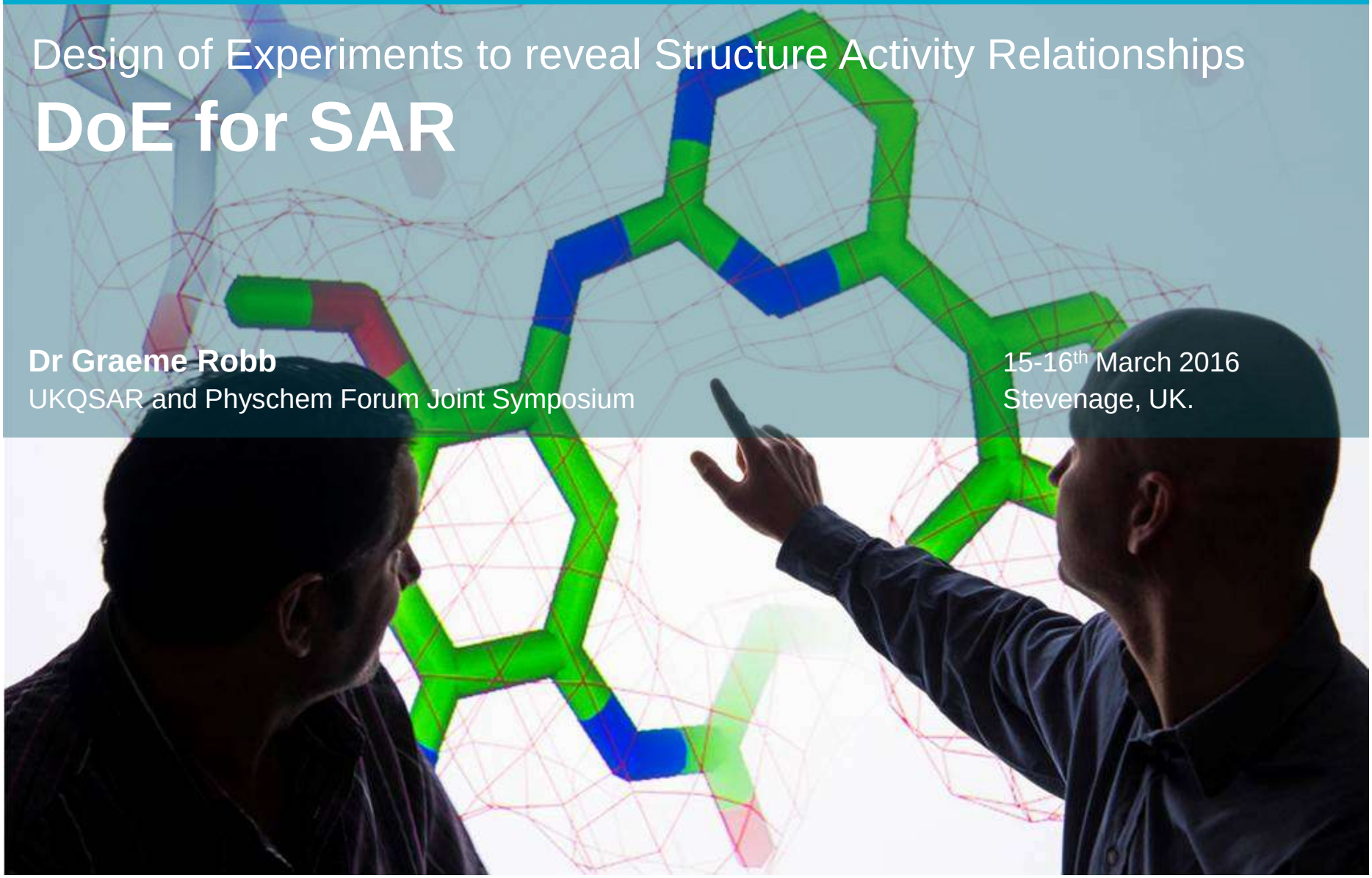
Design of Experiments to reveal Structure Activity Relationships

DoE for SAR

Dr Graeme Robb

UKQSAR and Physchem Forum Joint Symposium

15-16th March 2016
Stevenage, UK.

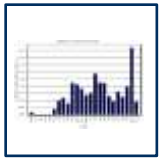


Topics

Design of Experiments: what and why?

Application to Medicinal Chemistry

Making DoE Chemocentric



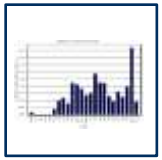
Results: does it work?

Topics

Design of Experiments: what and why?

Application to Medicinal Chemistry

Making DoE Chemocentric



Results: does it work?

Design of Experiments (DoE)

A statistical technique to identify which experiments to perform to tell you maximum information about a system.

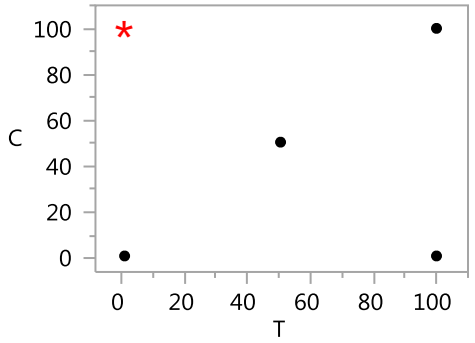
Particularly relevant when the cost (money, time, *etc.*) of an experiment is high.

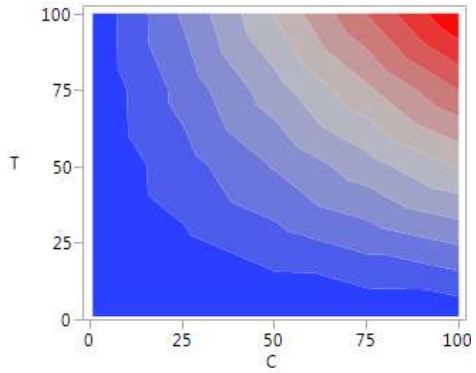
For example, it is used in pharmaceutical development in order to maximise the yield of a synthetic reaction.

How DoE works: Screening Design

- In simple terms, determine the minimum number of experiments you need to perform to cover the **expected variability** in your inputs.
- *e.g.* A reaction where you vary reagent concentration (C) and temperature (T)
  - Can explore the system with a **minimum of 4** experiments, assuming only linear effects
  - Include interaction term (C*T) with 5 experiments
  - Fit output yields to equation and pick optimal values of C and T

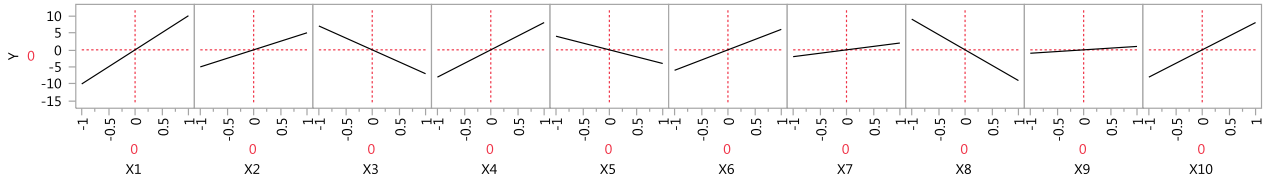
$$Y = X_1*C + X_2*T + X_3*C*T + X_4$$

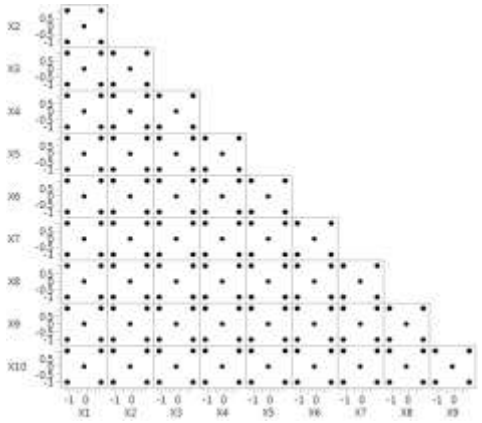




Multidimensional DoE

- Linear designs become **more efficient** at higher dimensionality
  - Can explore 10 variables with a **minimum of 16** experiments

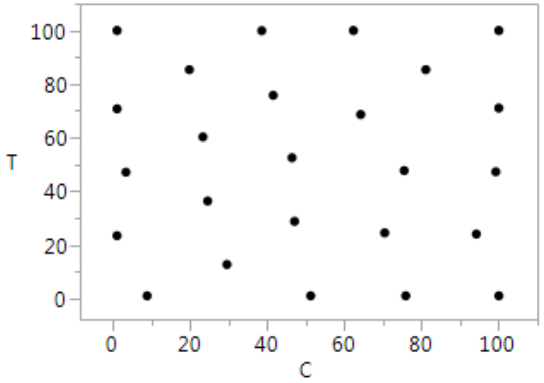




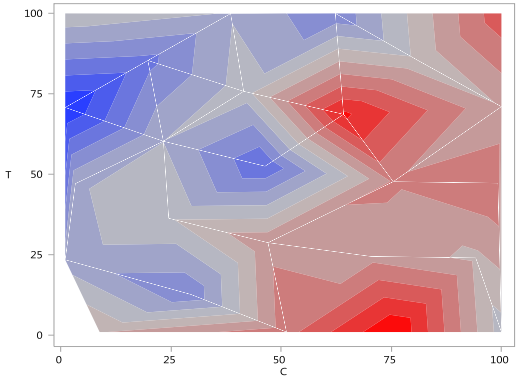
- If interactions are important it takes many more
  - **57 experiments** to include 2nd-order interactions
  - What about 3rd, 4th, *etc* order terms?
- Full modelling of the interactions becomes **prohibitively expensive**

Covering Design

- Rather than modelling as an equation of linear effects and interaction terms, create a design of **equally spaced points** in all variables
  - Resolution of design determines the number of experiments
- Thinking about only concentration and temperature again
  - A 25-point design **covers the space** very effectively



  - Has the advantage of being able to describe **complex landscape**, not just smooth functions



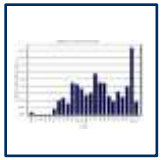
- Better than screening design for higher dimensions

Topics

Design of Experiments: what and why?

Application to Medicinal Chemistry

Making DoE Chemocentric



Results: does it work

So, what has this to do with drug design?

- **Can we use the principles of DoE to pick a library of compounds with the minimum size, that will give as maximum information?**
  - Here an 'experiment' is an individual compound to make
  - Goal is not accurate prediction, but identification of promising SAR to focus on

- Rather than independent variables; our inputs are **chemical structures** and can only be altered in **discrete** amounts (can't add ½ an atom)

- Problem: "[Chemical] Space is big. Really big. You just won't believe how vastly, hugely, mind-bogglingly big it is. I mean, you may think it's a long way down the road to the chemist, but that's just peanuts to [chemical] space." *– Douglas Adams*

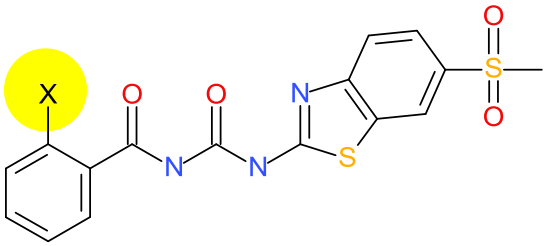
- The number of potential molecules is huge (JL Reymond, *Acc.Chem.Res.*, 2015)
  - Using up to 13 atoms = **977 million** real molecules
  - Using up to 17 atoms = **166 billion** real molecules
  - Drugs typically have up to 40 atoms, estimated to be **10^60** drug-like molecules !!!!

Framing the problem

- How to possibly represent the **vastness** of chemical space?
  - Don't try. Consider a limited scope.
    - *e.g*. Fix most of a molecule and vary a smaller part
- What **inputs** to use for the design?
  - Atoms and bonds are discrete
  - Describe high-dimensional space
- Instead use **properties** derived from the structure
  - *e.g.* LogD, ring count, molecular weight, H-bond acceptors, donors, *etc*
  - These are an incomplete representation of structure, but may suffice
  - Many properties are correlated and not independently variable
  - These are continuous and so suitable for DoE
- **Requires careful selection of properties for DOE**

Simple Screening Design

- Design of antagonists for **GHS-R1a**
- Focussed development at a single position



- Picking only small groups, there are >100 possibilities
- Use properties for DoE

Journal of Medicinal Chemistry

J. Med. Chem., **2014**, *57* (14), pp 6128–6140

Article

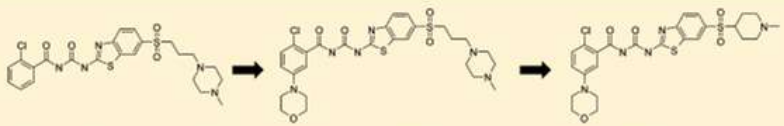
pubs.acs.org/jmc

Identification, Optimization, and Pharmacology of Acylurea GHS-R1a Inverse Agonists

William McCoull,* Peter Barton, Alastair J. H. Brown, Suzanne S. Bowker, Jennifer Cameron, David S. Clarke, Robert D. M. Davies, Alexander G. Dossetter, Anne Ertan, Mark Fenwick, Clive Green, Jane L. Holmes, Nathaniel Martin, David Masters, Jane E. Moore, Nicholas J. Newcombe, Claire Newton, Helen Pointon, Graeme R. Robb, Christopher Sheldon, Stephen Stokes, and David Morgan

AstraZeneca, Mereside, Alderley Park, Macclesfield SK10 4TG, U.K.

Supporting Information



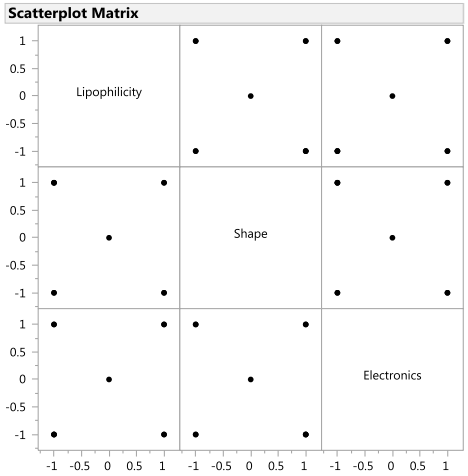
Compound	14 (AZ-GHS-14)	22 (AZ-GHS-22)	38 (AZ-GHS-38)
GHS-R1a binding IC_{50}	1.3 nM	0.77 nM	6.7 nM
GHS-R1a function	partial agonist	inverse agonist	inverse agonist
CNS exposure	-	non-CNS penetrant	CNS penetrant

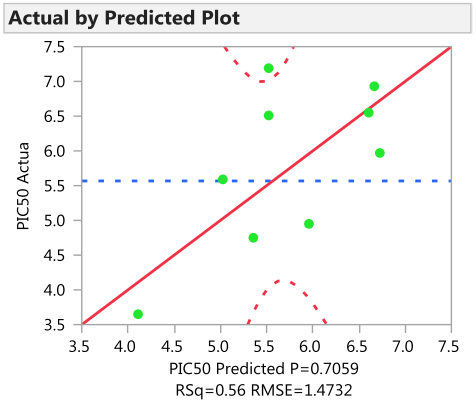
ABSTRACT: Ghrelin plays a major physiological role in the control of food intake, and inverse agonists of the ghrelin receptor (GHS-R1a) are widely considered to offer utility as antiobesity agents by lowering the set-point for hunger between meals. We identified an acylurea series of ghrelin modulators from high throughput screening and optimized binding affinity through structure−activity relationship studies. Furthermore, we identified specific substructural changes, which switched partial agonist activity to inverse agonist activity, and optimized physicochemical and DMPK properties to afford the non-CNS penetrant inverse agonist **22** (AZ-GHS-22) and the CNS penetrant inverse agonist **38** (AZ-GHS-38). Free feeding efficacy experiments showed that CNS exposure was necessary to obtain reduced food intake in mice, and it was demonstrated using GHS-R1a null and wild-type mice that this effect operates through a mechanism involving GHS-R1a.

Simple Screening Design

- Pick three largely **orthogonal parameters**
  - Lipophilicity (pi)
  - Shape (volume)
  - Electronics (sigma-p)

- Could make a minimum of 5, but expect non-linearity so make 10
  - Select a representative compound for each design point (close in properties)



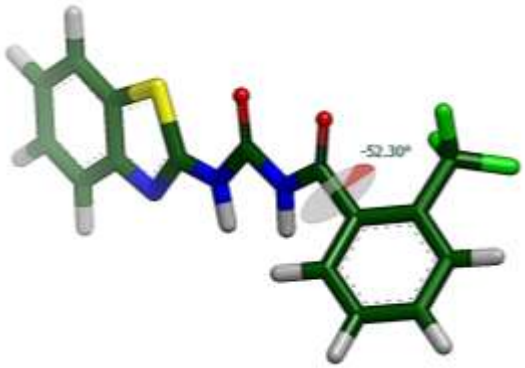


Summary of Fit	
RSquare	0.56201
RSquare Adj	-0.31397
Root Mean Square Error	1.473204
Mean of Response	5.567264
Observations (or Sum Wgts	10

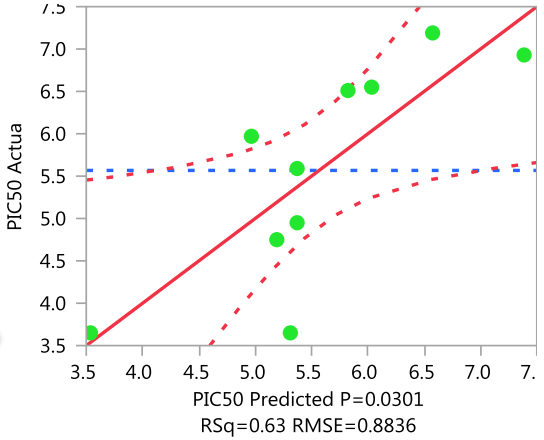
Problem: no model for activity!

Simple Screening Design

- Apply a bit of chemistry
  - Outliers on the model show different **3D shape** from those on the line.



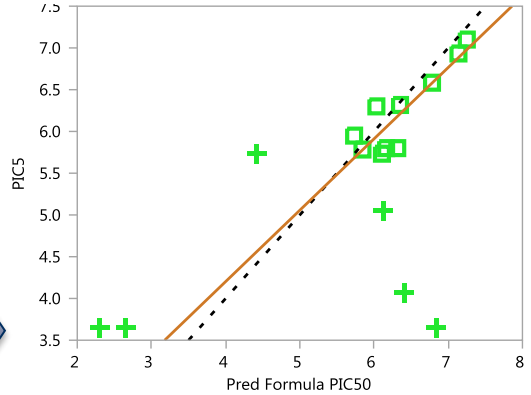
Include extra dihedral angle term



Summary of Fit	
RSquare	0.632318
RSquare Adj	0.527265
Root Mean Square Error	0.883647
Mean of Response	5.567264
Observations (or Sum Wgts	10

Gives a better model

pi and sigma-p drop out



Summary of Fit	
RSquare	0.775641
RSquare Adj	0.747596
Root Mean Square Error	0.253969
Mean of Response	6.2281
Observations (or Sum Wgts	10

Model identifies designs with activity!

11 (of 14) true positives

Screening Design Summary

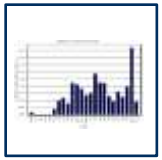
- **Good**: confident we have identified optimal group for affinity
  - Made only 27 compounds (from >100)
- **Challenge**: would we have reached this point without using DoE?
  - Probably, but with either more compounds or less confidence
- **Challenge**: wasted effort on inactive / unattractive compounds?
  - True. In order to populate the property space, sometimes we make compounds containing unattractive groups, like nitro or thioether
- **Bad**: Picked molecules were a compromise – closest to design points
  - In some cases there may be no close compound for a specific design point
- **Bad**: had to adapt model part way through
  - Original design failed
- **Bad**: Complex, noisy model for one small group
  - Possibly couldn't apply this to multiple or larger groups

Topics

Design of Experiments: what and why?

Application to Medicinal Chemistry

Making DoE Chemocentric



Results: does it work?

Chemocentric DoE

- The SAR landscape is generally too complex for a Screening Design
- **Covering Design** gives better SAR description
  - Property space presents the same challenges as before

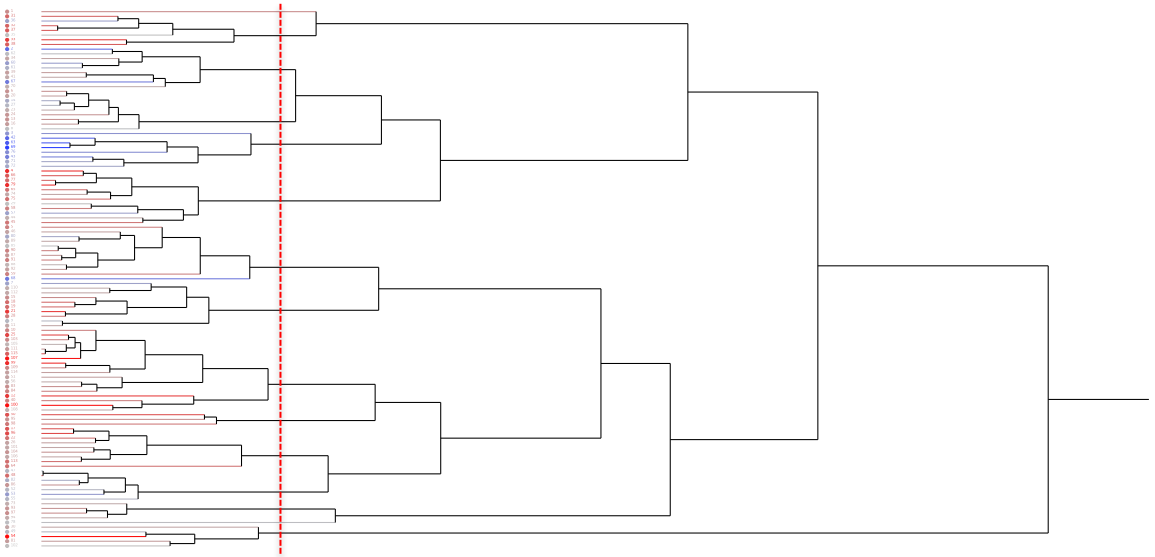
1. Challenge: **Do compounds represent design points?**

2. Challenge: **Do we have to make unattractive compounds?**

3. Challenge: **How many properties are required; and which ones?**

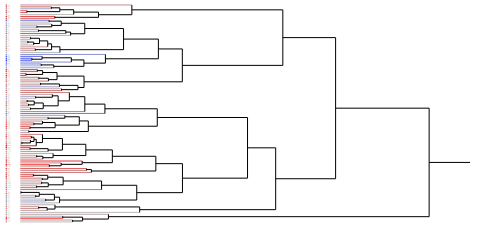
Compounds != design points ?

- Traditional covering design selection aims for evenly-spaced design points
  - But difficulties in converting design points into compounds
- Identify a ‘candidate’ set, *i.e.* a list of all potential compounds that could be made
- Appropriate design should best represent the compound set
  - Use clustering to define well-spaced, representative experimental design



Which property space?

- Clustering of properties is one option
  - Can use data compression, *e.g*. PCA



- Clustering works for chemical structure too
  - Using chemical fingerprints
  - Or try clustering by pharmacophore

- We have a chemistry-ready DoE technique, but...
  - **Which selection method performs best?**
  - **Does this method outperform regular DoE?**
  - **Does this method outperform a medicinal chemist?**

Test conditions

- A set of 115 compounds were made on GHS-R1a project, varying one group
  - This is our candidate set (Q: is this fair?)
  - What proportion must be made to identify the best of the rest?
  - Pick candidates by various methods
    1. Screening design, based on property data (first 3 principle components)
    2. Covering design, based on clustering of properties
    3. Covering design, based on clustering of pharmacophores (2D, path-based)
    4. Covering design, based on clustering of chemical fingerprints
    5. Completely random selection
    6. Temporal order, *i.e.* the order the chemist made them

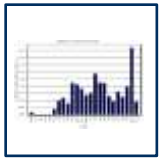
- In each case, build a linear-regression QSAR model from the testset
  - Measure of accuracy: testset prediction RMSE
  - Compound selection within clusters is random
  - For robustness, repeat each 20 times to get average result

Topics

Design of Experiments: what and why?

Application to Medicinal Chemistry

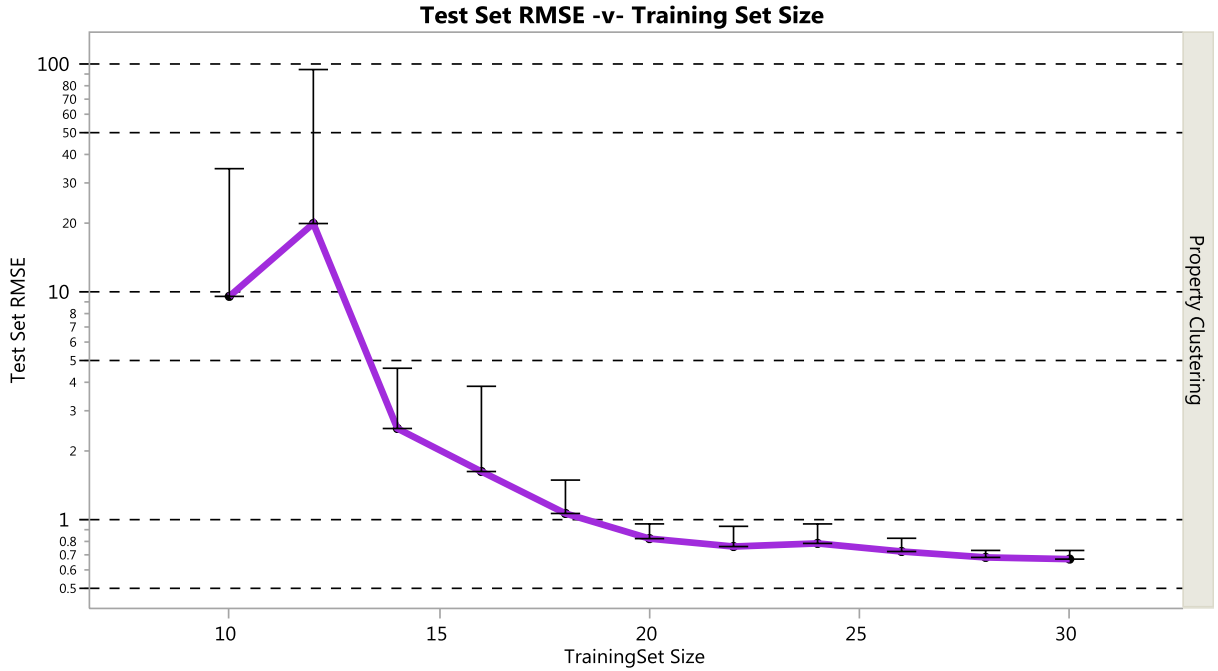
Making DoE Chemocentric



Results: does it work?

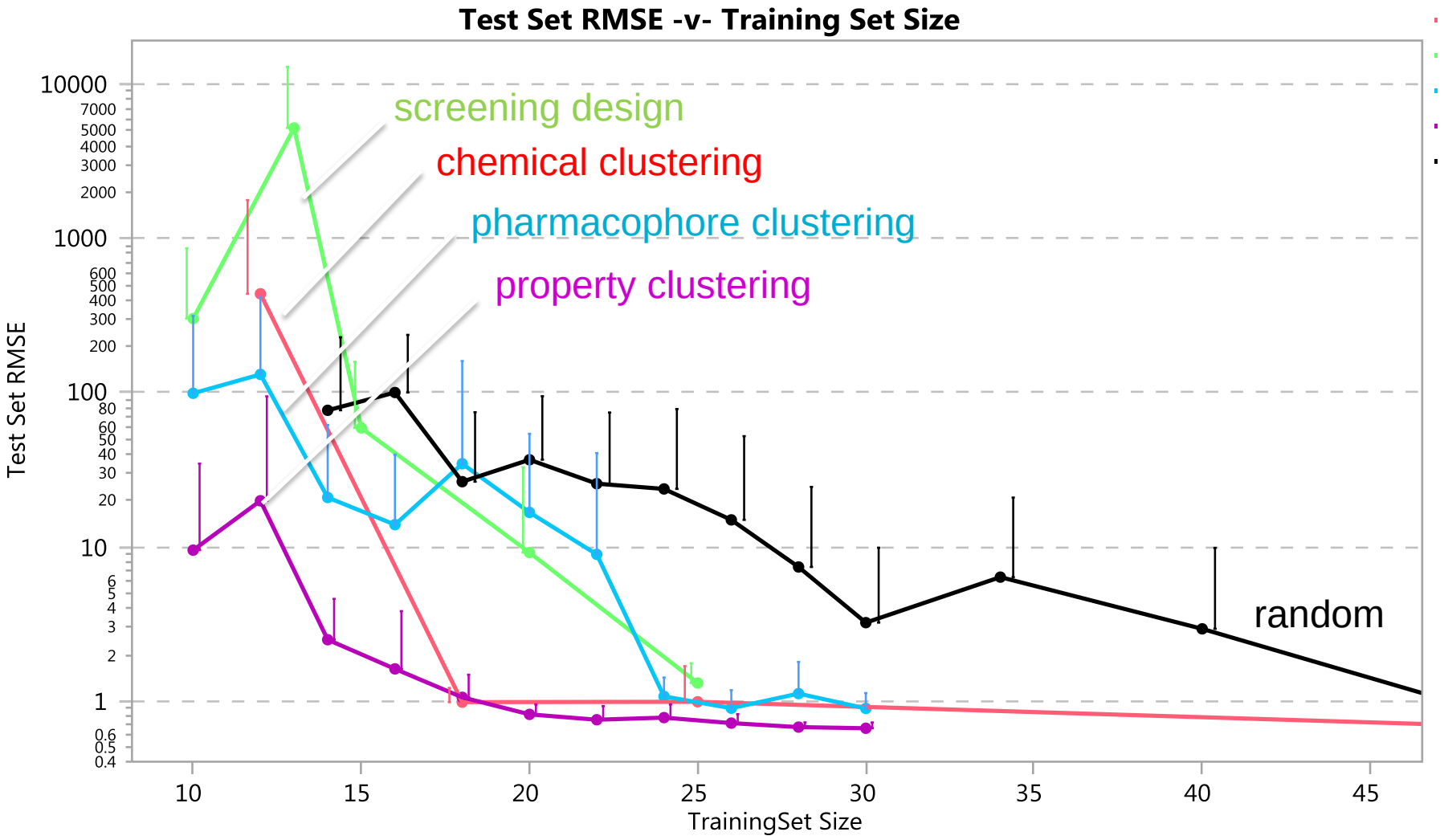
Covering Design by Property

- RMSE reduces then plateaus with increased design test set
  - QSAR model of pIC50 can't do better than RMSE ~ 0.7



- A training set of ~20 compounds is reasonably good at identifying the best of the remaining 95

Methods comparison

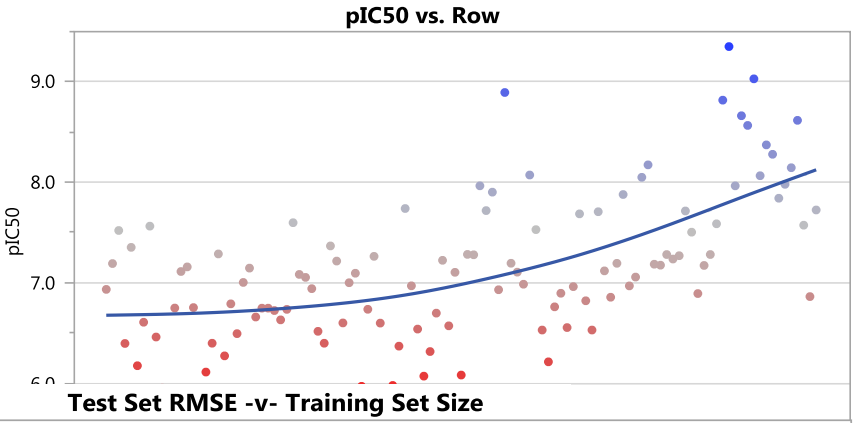


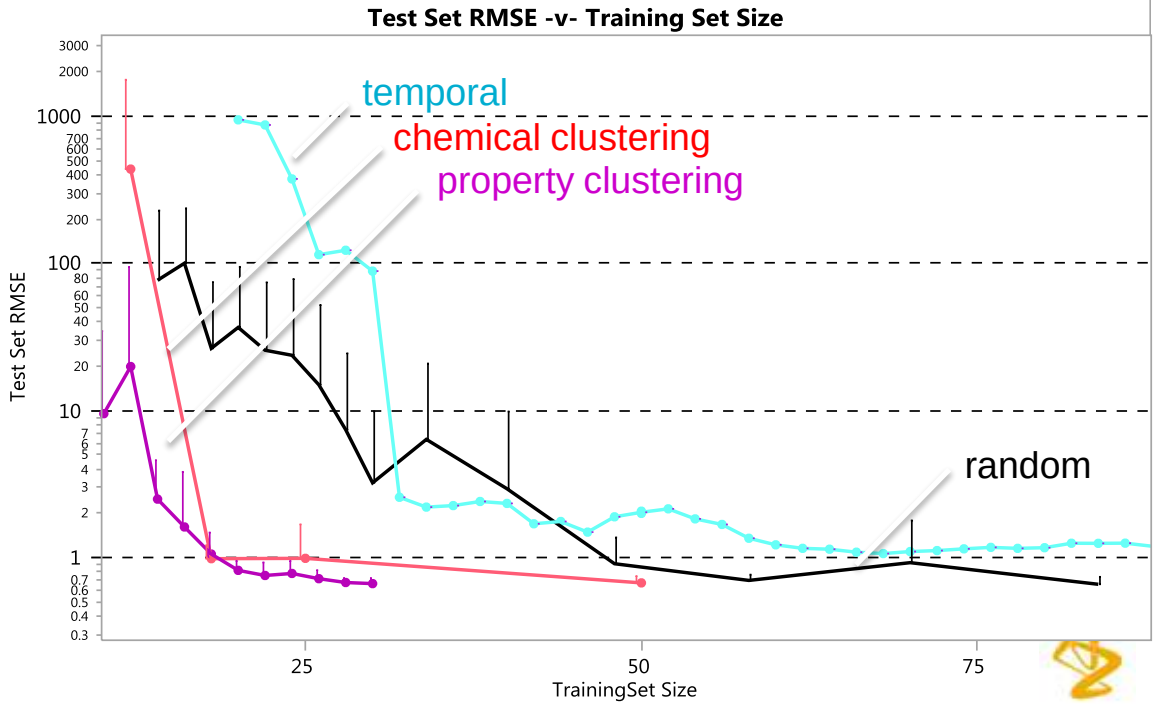
Methods comparison

- Any selection method is much better than random
  - Nice to confirm
- Covering designs outperform Screening design
  - Due to the non-linear, complex SAR landscape
- Chemical clustering and property clustering perform similarly
  - Property clustering seems superior at lower test-set size
- Pharmacophore clustering is more variable
  - 2D pharmacophore may be too crude
- **Now for the big test...**

Does DoE predict better than a chemist?

- Potency increased with time over the life of the project.
  - Of course, potency was not the only goal
- Analysis shows early compound give poor models for identifying the best of the later ones
  - No better than random
- DoE does much better at selecting which compounds to make first to identify the best of the rest
- *Caveat*
  - *No way to determine the false positive rate for unmade compounds*





A harder test

- 715 compounds, from three sub-series, with three points of diversity, made to inhibit 11β-HSD1.
- Can the DoE method handle a much more diverse set?

Journal of Medicinal Chemistry

J. Med. Chem., **2012**, *55* (23), pp 10652–10661

Article

pubs.acs.org/jmc

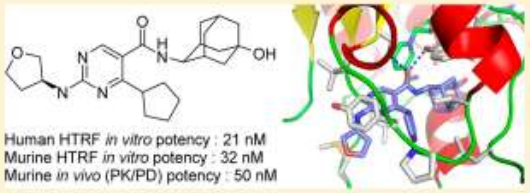
Free-Wilson and Structural Approaches to Co-optimizing Human and Rodent Isoform Potency for 11β-Hydroxysteroid Dehydrogenase Type 1 (11β-HSD1) Inhibitors

Frederick W. Goldberg,* Andrew G. Leach, James S. Scott, Wendy L. Snelson, Sam D. Groombridge, Craig S. Donald, Stuart N. L. Bennett, Cristian Bodin, Pablo Morentin Gutierrez, and Amy C. Gyte

AstraZeneca, Mereside, Alderley Park, Macclesfield, SK10 4TG, United Kingdom

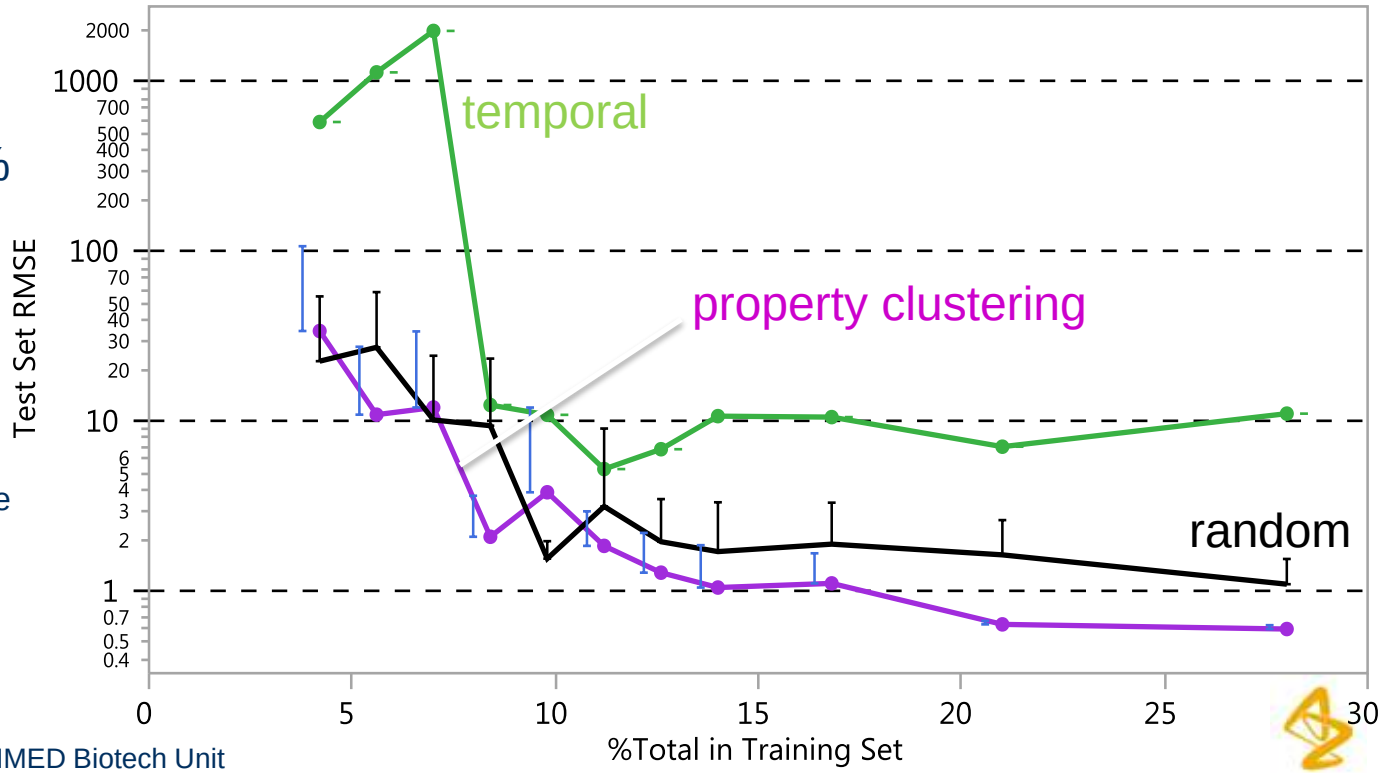
Supporting Information

ABSTRACT: 11β-Hydroxysteroid dehydrogenase 1 (11β-HSD1) has been a target of intensive research efforts across the pharmaceutical industry, due to its potential for the treatment of type II diabetes and other elements of the metabolic syndrome. To demonstrate the value of 11β-HSD1 in preclinical models, we required inhibitors with good potency against both human and rodent isoforms. Herein, we describe our efforts to understand how to co-optimize human and murine potency within the (5-hydroxy-2-adamantyl)-pyrimidine-5-carboxamide series. Two approaches are described—a data-driven (Free-Wilson) analysis and a structure-based design approach. The conclusions from these approaches were used to inform an efficient campaign to design compounds with consistently good human/murine potency within a $\log D_{7.4}$ range of 1−3. Compounds **20** and **26** demonstrated good rodent PK, which allowed us to demonstrate a PK/PD relationship in rat and mouse. We then evaluated **26** against glycemic and body weight end points in murine disease models, where it demonstrated glucose and body weight efficacy at 300 mg/kg/day but only body weight efficacy at 50 mg/kg/day, despite providing >90% target engagement in the liver.



A harder test

- DoE selection (and random selection) do a better job of identifying the better future potential compounds than the order they were actually made
- DoE selection gives similar result to random selection for low training set size
- DoE overtakes random when training set >10%



Note, all three series were made simultaneously for the temporal analysis

Summary

Design of Experiments can be applied to the problem of chemistry design

Adapting DoE to be more chemocentric vastly improves the predictive performance

Chemocentric DoE outperforms prediction based on the order that compounds were actually made

Could Chemocentric DoE be used to accelerate initial scoping in early chemistry?

Confidentiality Notice