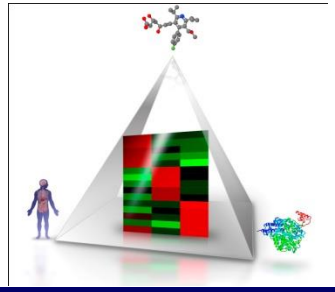
Using Chemical and Biological Data, In Particular Applied to Selecting Small Molecules to Increase Thermal Stability Of a Monoclonal Antibody

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

UNIVERSITY OF CAMBRIDGE



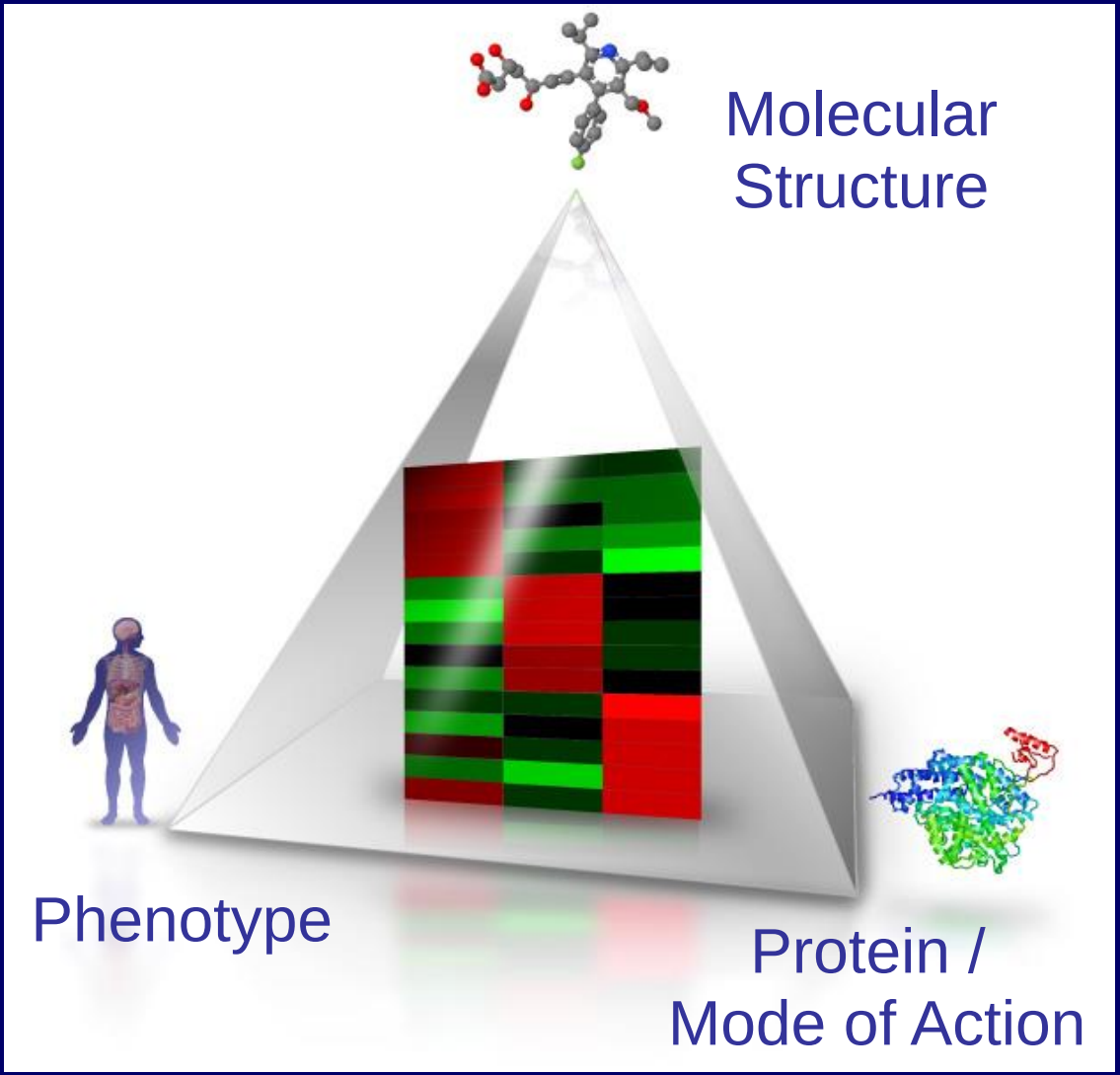
Outline

- 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 antibodides

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)

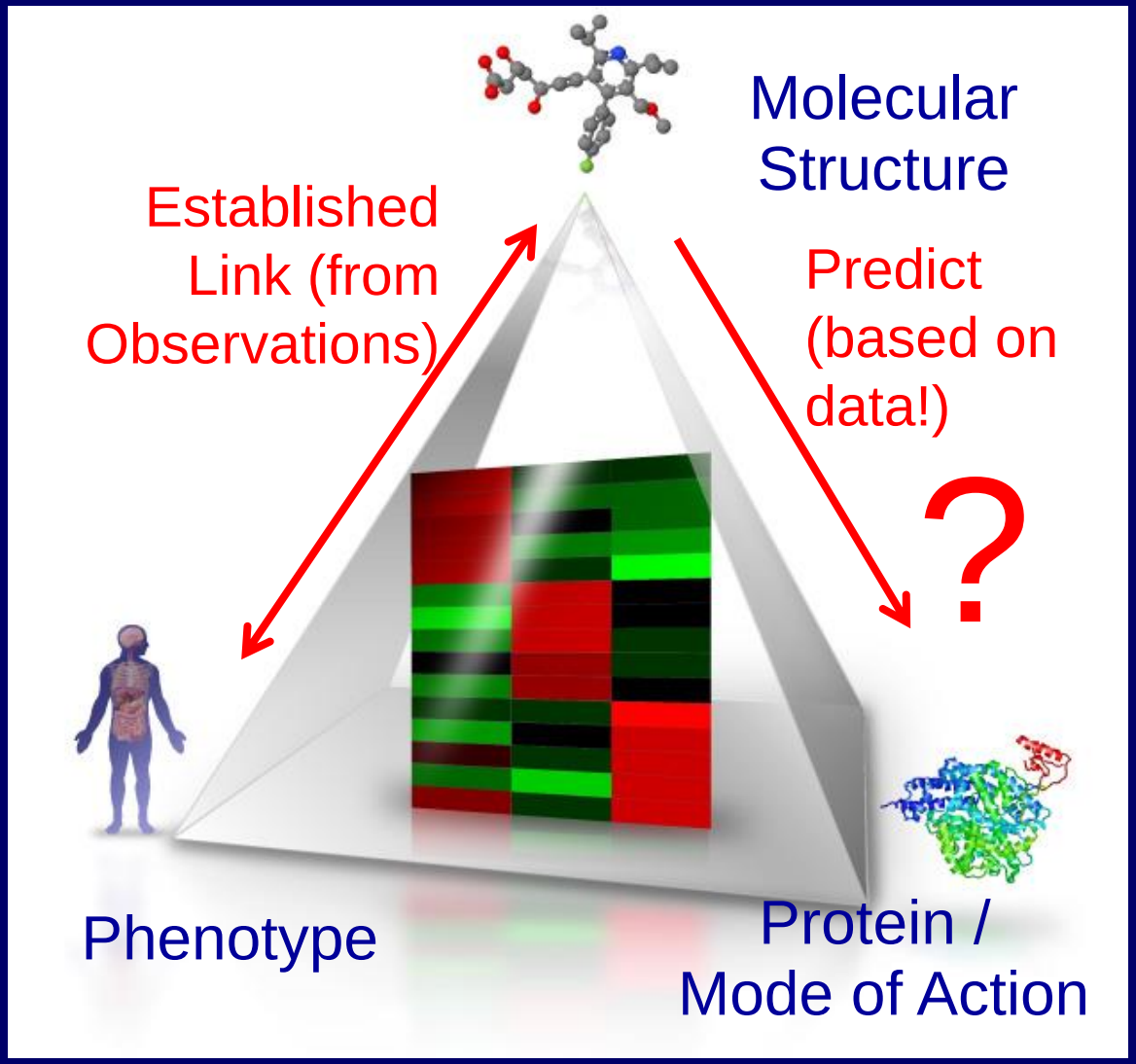
Modelling compound mixtures, traditional medicines

- 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

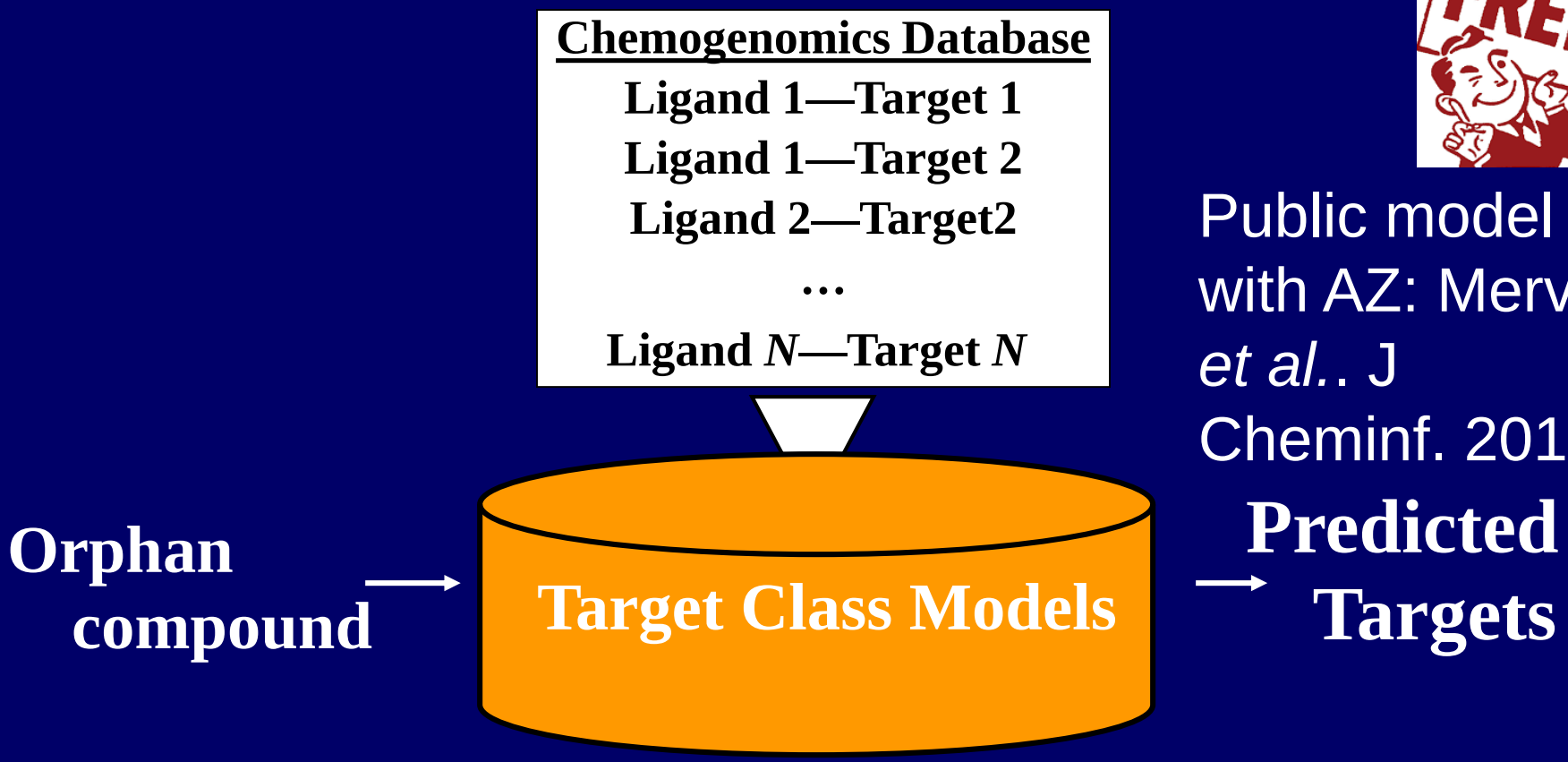


A. Koutsoukas *et* al., J Proteomics 2011 (74) 2554 – 2574.

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.

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

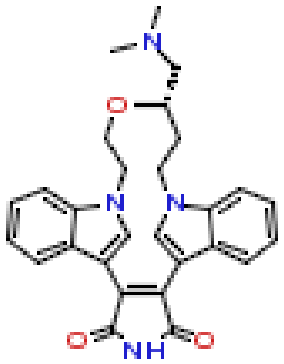


Prediction Examples: Gleevec, Ruboxistaurin

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

- Ruboxistaurin (Lilly/Takeda),Phase III
  - PKCb

Molecule	Targets	Scores
	ABL1	46.50
	PDGFRB	28.99
	KIT	22.02
	CDK9	21.30
	BRAF	16.13
	FLT1	13.09
	PLK1	8.05
	BTK	5.44

Molecule	Targets	Scores
Chiral 	PRKCB1	95.81
	CAMK2G	87.48
	PRKCG	66.35
	PRKCA	56.99
	PRKCD	52.44
	PRKCH	51.41
	PRKCE	50.42
	PRKCZ	42.48

Understanding rat sleep data

- Project with Eli Lilly
- Male Wistar rats

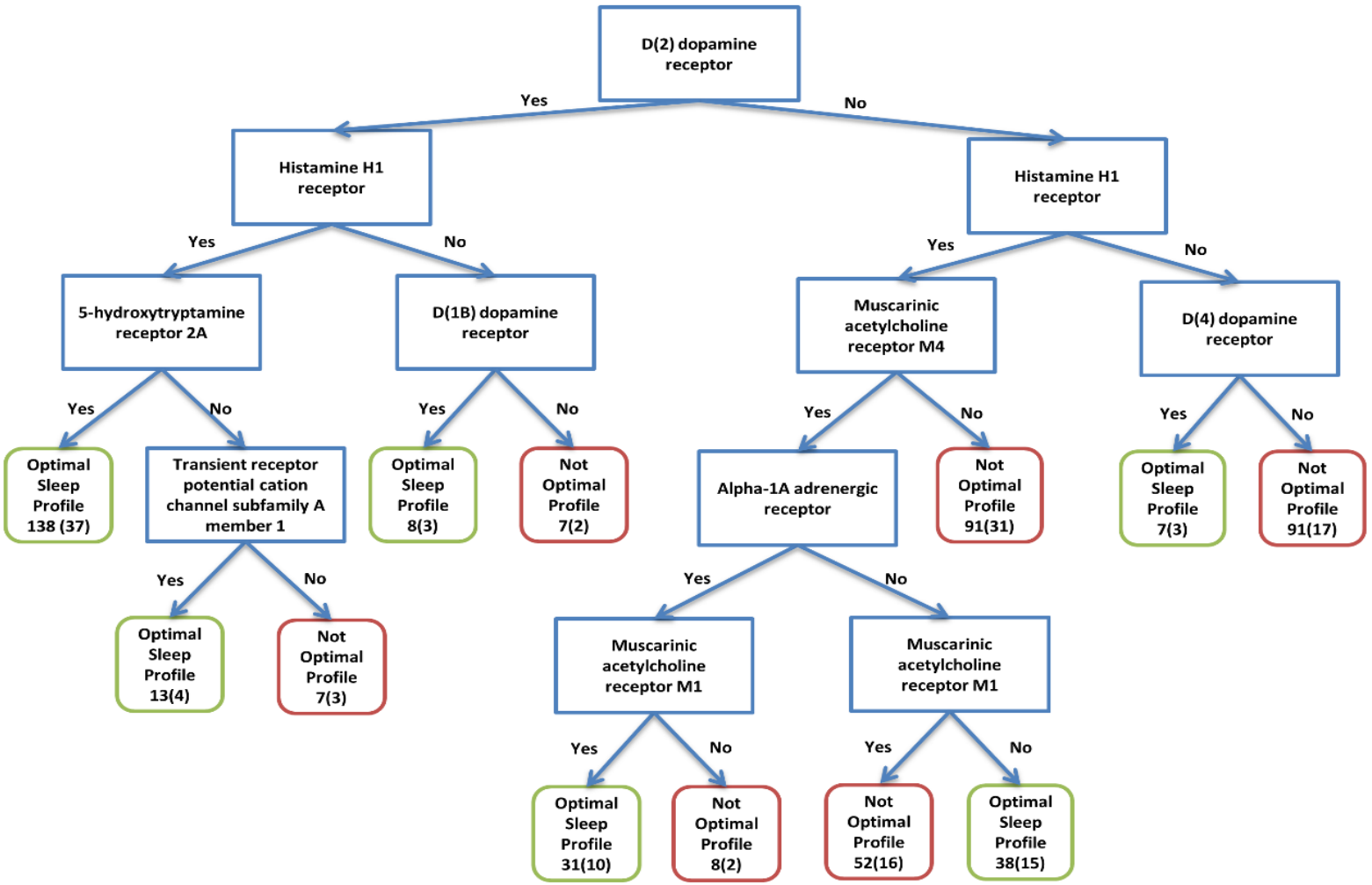
Work by Georgios Drakakis

- 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

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

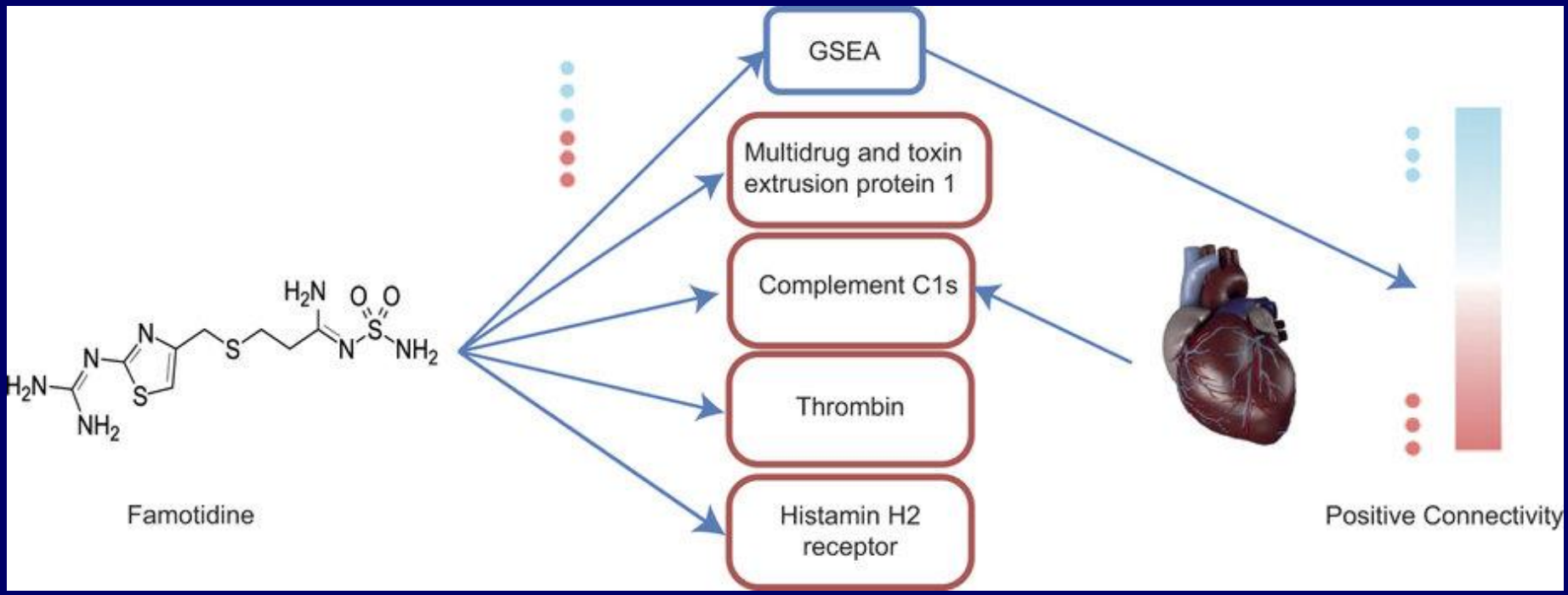
- *5 out of 7 marketed drugs (71%) tested increased sleep parameters (a sixth led to hyperactivity!)*

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

- Overall 78% correct on target level, ~71% on phenotypic level (across 4 MoA classes)

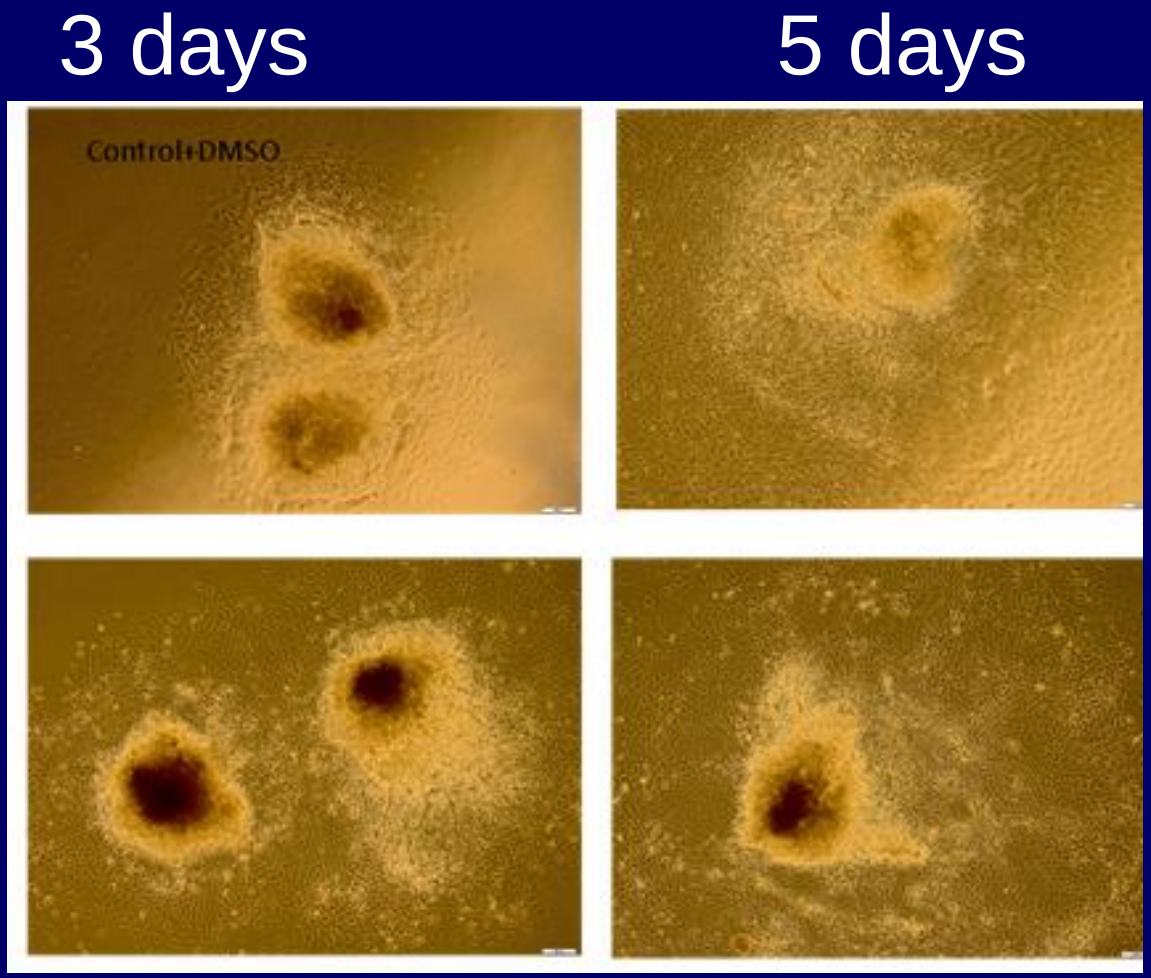
Combined gene expression / target prediction analysis for MoA analysis and compound selection

- 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)



KalantarMotamedi *et al. Cell Death Discovery* **2016**

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

- 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”
- Oyetayo *et al. Eur. J. Pharm. Sci.* (in revision)

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

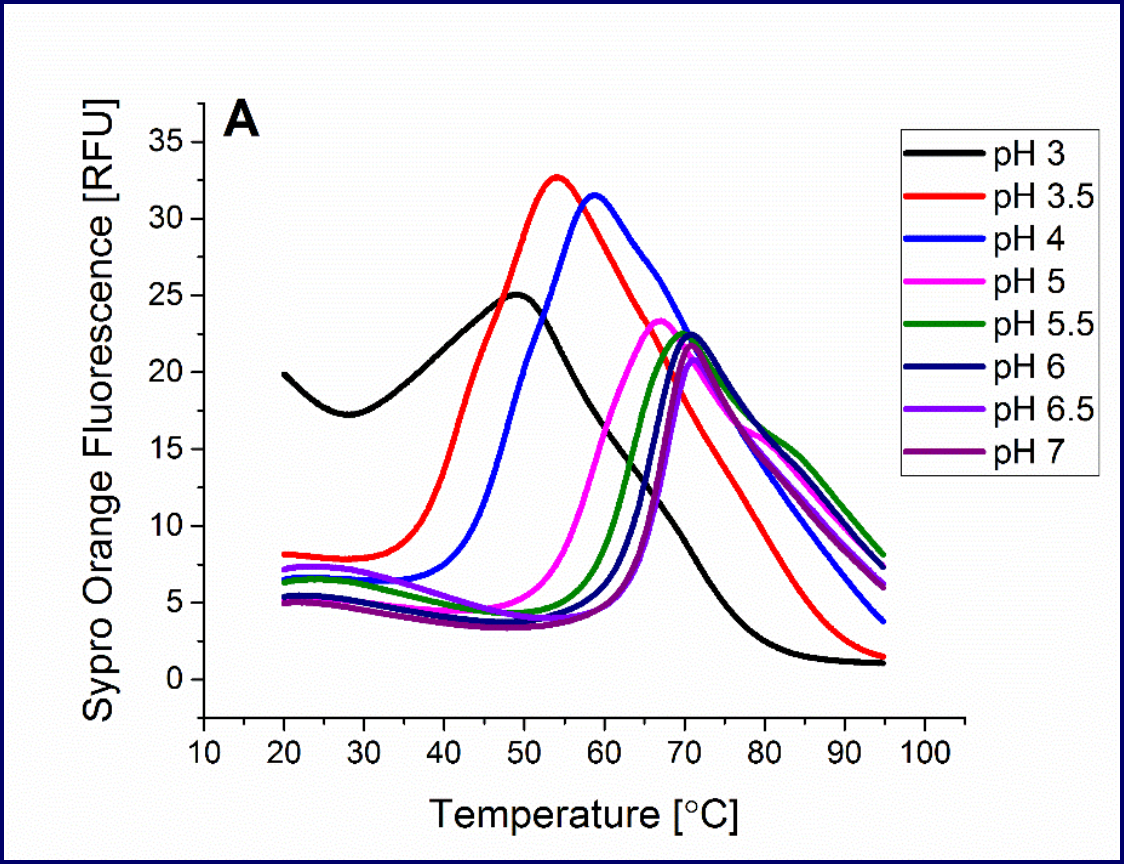
Selection of diverse compound library to determine Ab stabilizing properties

- Amino acids, methylamines 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

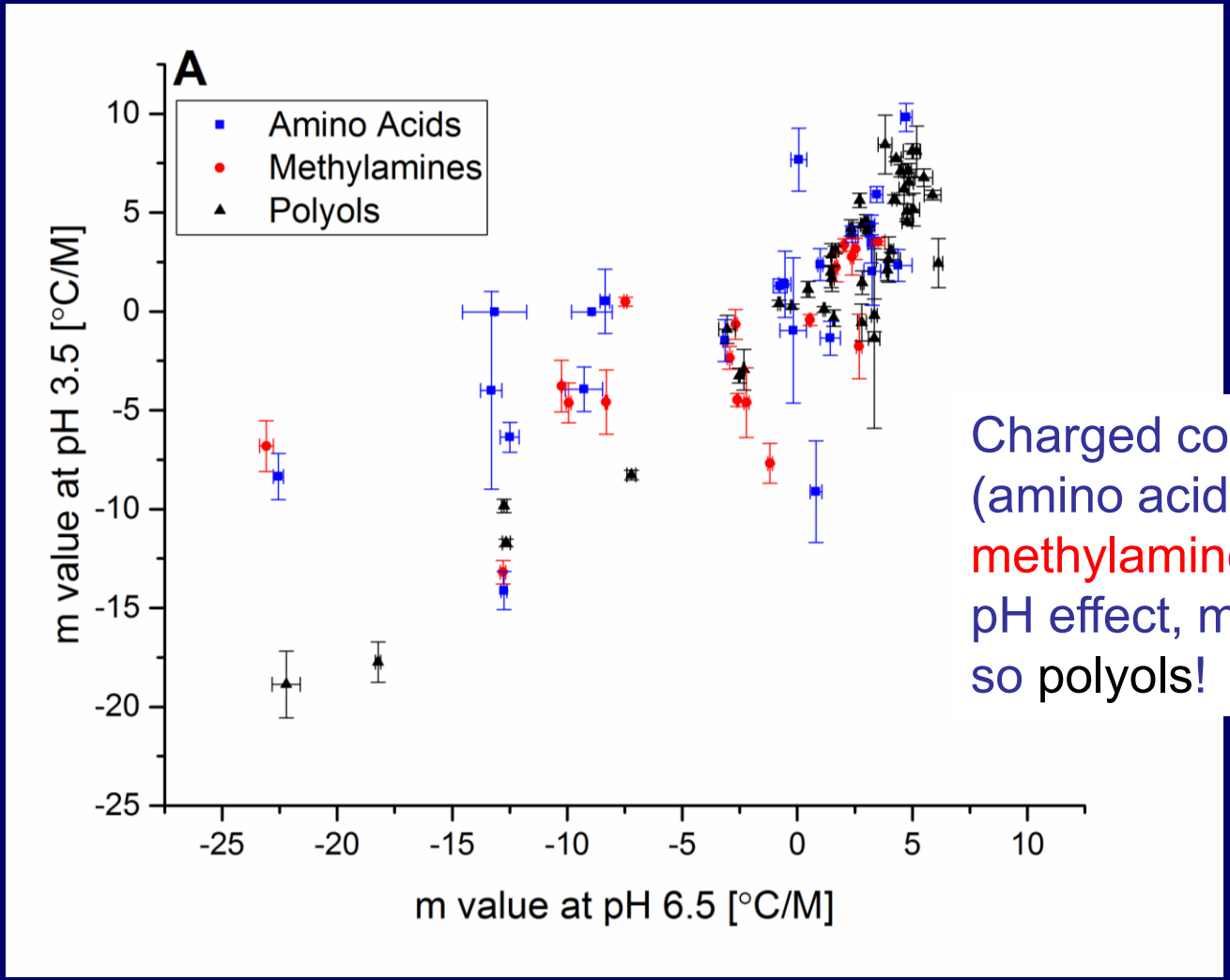
- Recombinant human monoclonal antibody of the IgG1 subclass (mAb1) was produced in-house in CHO cells
- To determine unfolding differential scanning fluorimetry (DSF) was used
  - High throughput method
  - Non-equilibrium method though
  - Hence impact of extrinsic fluorescent probe on T_m, inability 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

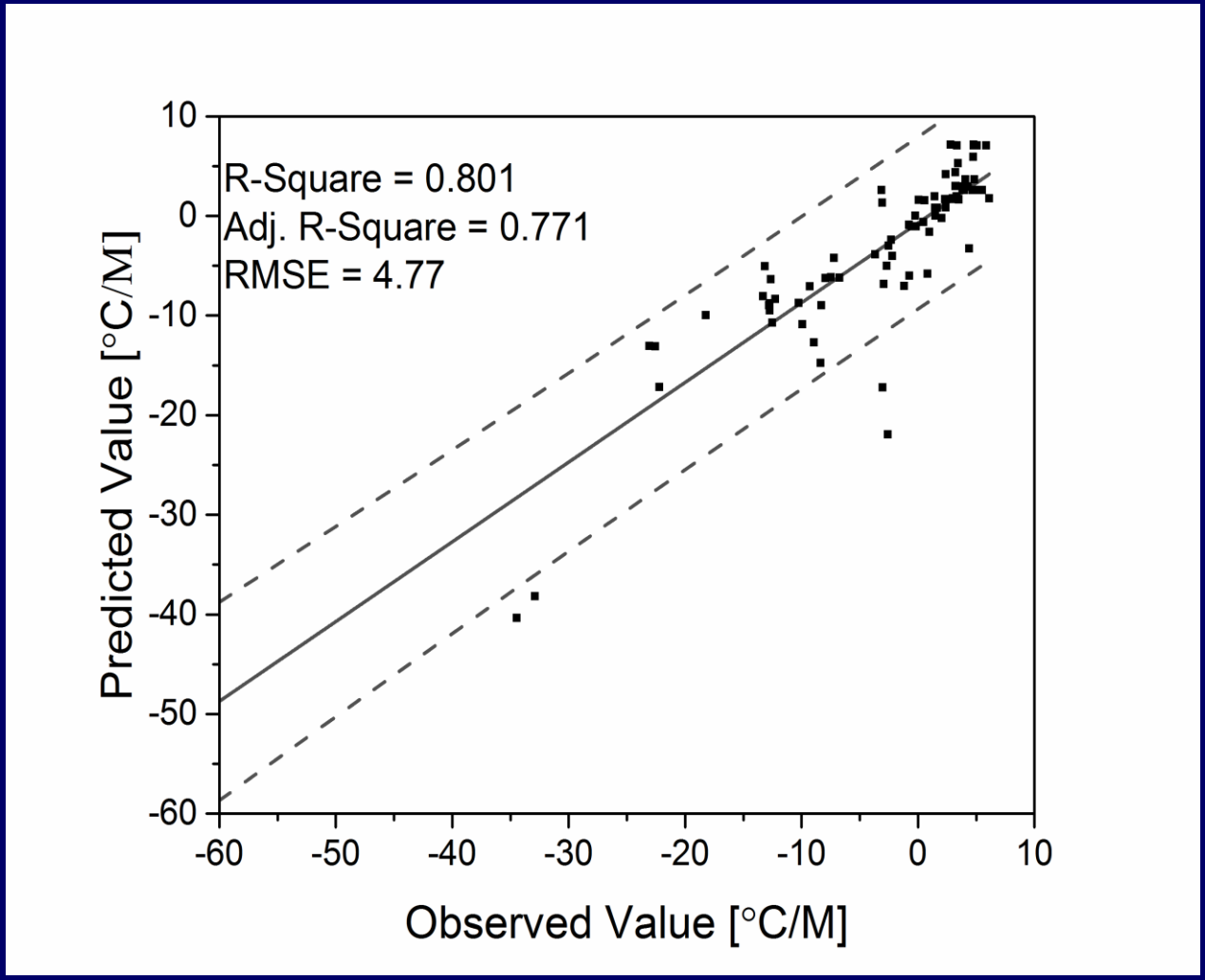
- Measurements at pH 6.5 showed less error than at pH 3.5
- pH 6.5 also more relevant for practical processing steps, hence data obtained at this pH was used for QSAR model generation

QSAR model: Partial Least Squares (PLS)

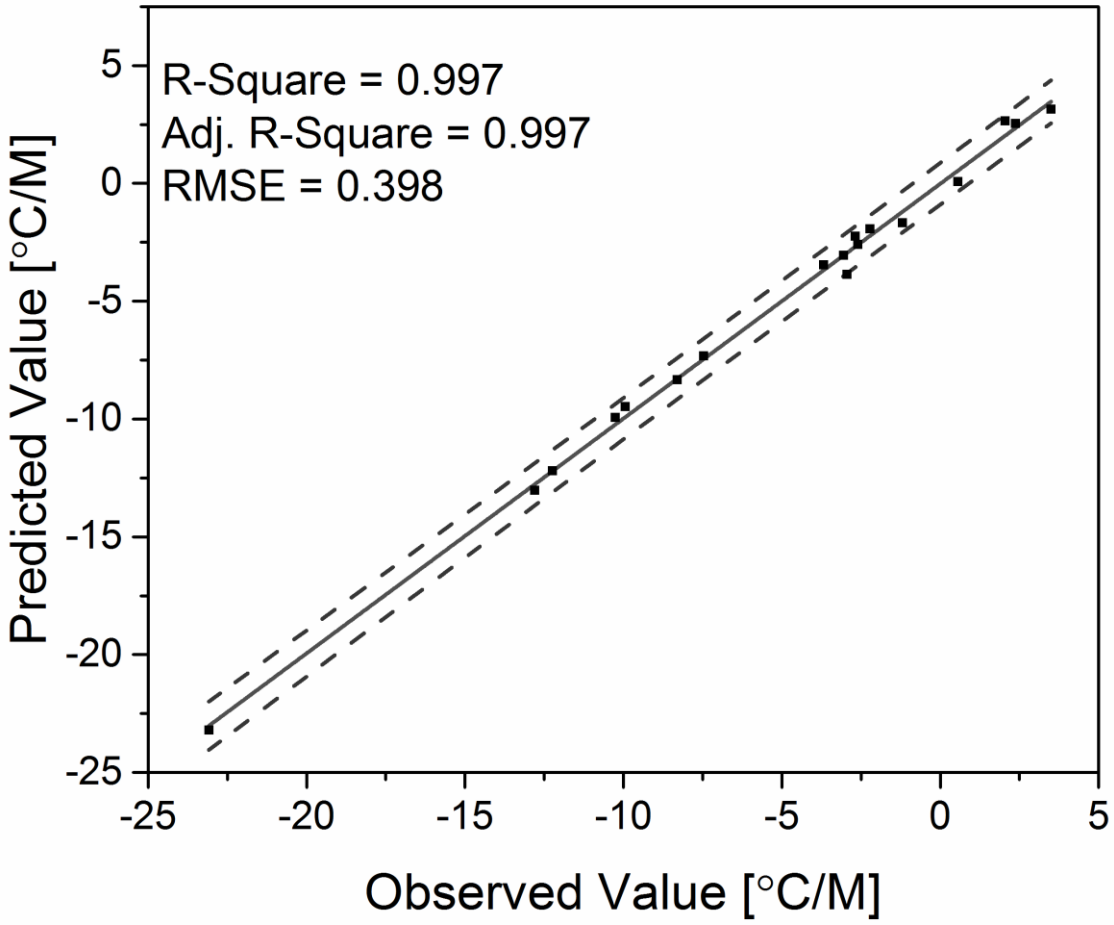
- For 84 compounds195 2D descriptors were calculated using MOE software
- Removal of descriptors with low variance, normalization
- TS potency at pH 6.5 used as output variable
- Models were fit to all data points, model consistency and variable importance determined in leave-one-out cross-validation
- Variable importance determined using ‘Variable Importance Projection’ (VIP)

Tm model fit across the amino acid, methylamine and polyol classes ('global model')

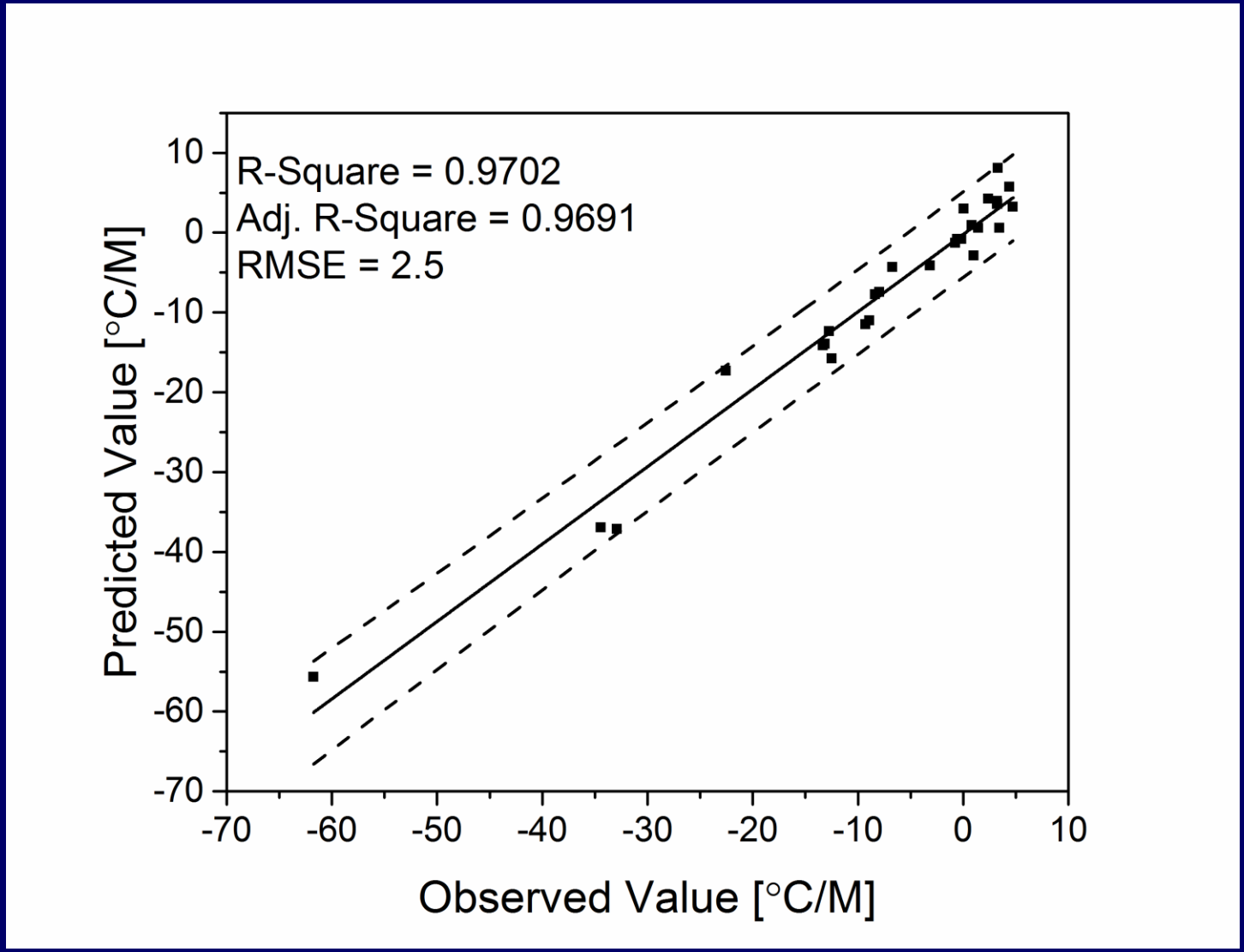
- RMSE = 4.77
- RMSE (LOO) = 6.07



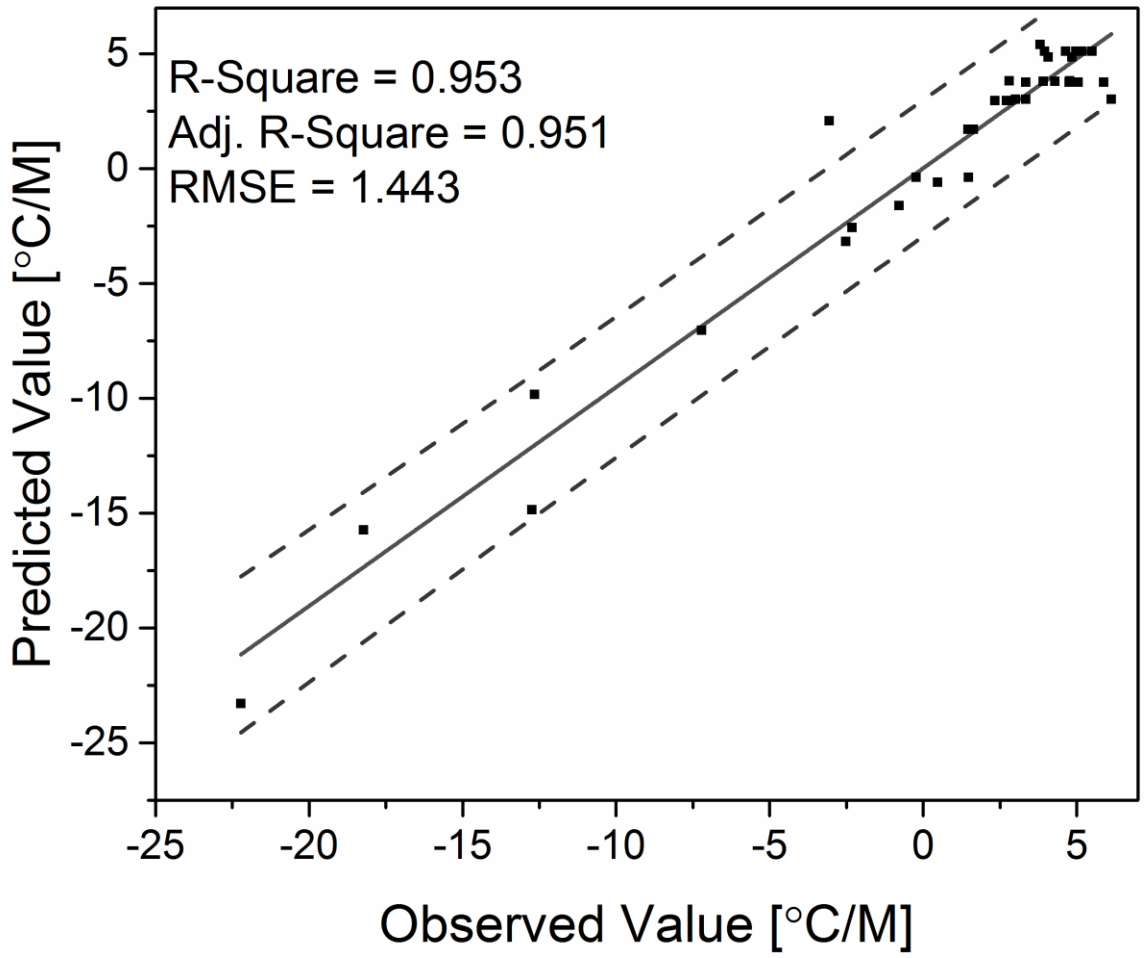
Local models give much better correlations: Methylamines



Local models give (somewhat) better correlations: Amino acids



Local models give (somewhat) better correlations: Polyols



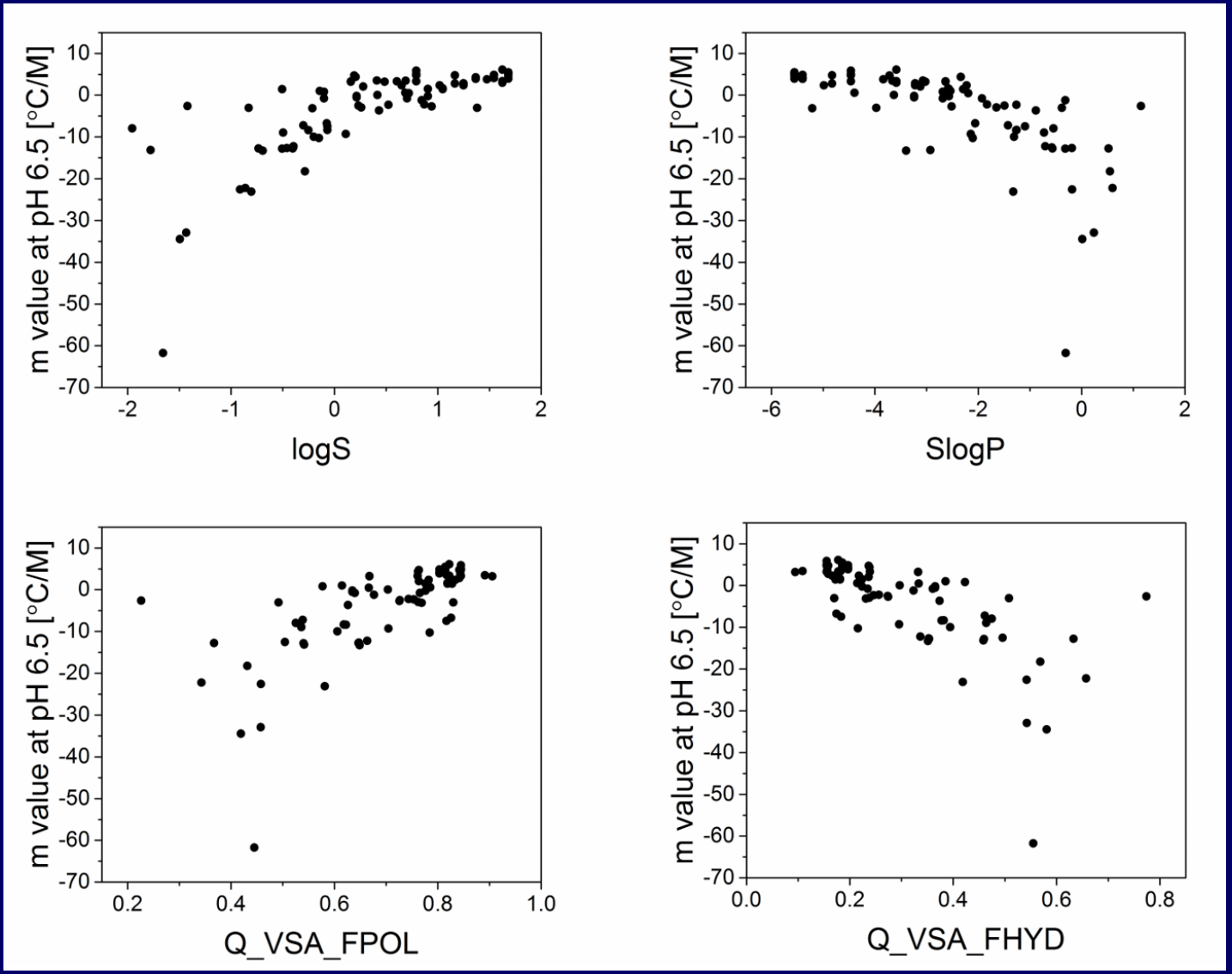
Model statistics: Leave-one-out validation

Model Statistic	Amino acids	Methylamin	Polyols	Global
Sample size	29	18	37	84
Descriptors	63	56	51	60
R-Squared value	0.864	0.97	0.906	0.679
Adj. R-Squared value	0.848	0.967	0.891	0.645
RMSE	5.34	1.15	2.04	6.07
Components	9	10	3	4

Variable Importance Projection: Polarity/hydrophobicity, accessible surface area are crucial

Variable Name	Description	VIP	Regression Coefficient
SlogP_VSA7	Sum of the accessible surface area (in Å^2) over all atoms *i* such that SlogP of atom *i* is in (0.25, 0.30]	2.027	-1.127
PEOE_VSA-1	Sum of the accessible surface area (in Å^2) over all atoms *i* such that the partial charge of atom *i* is in [-0.10, -0.05)	1.929	0.678
logS	Log of aqueous solubility (mol/L)	1.798	-0.522
a_aro	Number of aromatic atoms	1.726	-0.438
b_ar	Number of aromatic bonds	1.726	-0.351
Q_VSA_FPOL	Fractional polar Van der Waals surface area	1.709	-0.351
Q_VSA_FHYD	Fractional hydrophobic Van der Waals surface area	1.709	-0.345

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 antibodides

- 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”
- Oyetayo *et al. Eur. J. Pharm. Sci.* (in revision)

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
Shardul Paricharak
Sharif Siam
Siti Zuraidah Sobir
Stephanie Ashenden
Xianjun Fu
Yasaman Motamedi
Yue Kong

Sebastian Rohrer
Stefan Tresch
Antje Wolf
Klaus-Juergen Schleifer

Ola Engkvist
Thierry Kogej
Jay Mettetal
Samantha Hughes

Martin Augustin
Tom Klenka

Therese Malliavin
Isidro Cortes

Ian Stott

Hinrich Goehlmann
Herman van Vlijmen
Joerg K. Wegener

Mike Bodkin
David Evans
Keith Wafford
Suzanne Brewerton

Massimo Mercati
Anna Maidecchi
Jacopo Lucci