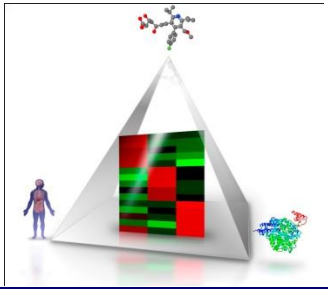
Cheminformatics and Bioinformatics Approaches Applicable to Neglected Diseases

Andreas Bender, PhD
Lecturer for Molecular Informatics
Unilever Centre for Molecular Science Informatics
University of Cambridge
Fellow of King's College, Cambridge

UNIVERSITY OF CAMBRIDGE



Outline

- Our take on chemical and biological information

- Cases where our approaches were applied / could be applied to neglected diseases
  - Mode-of-action analysis from phenotypic screens
  - Compound selection against protein mutants
  - Analysis of compound combination screens against Plasmodium
  - Using gene expression data for compound selection and assessment

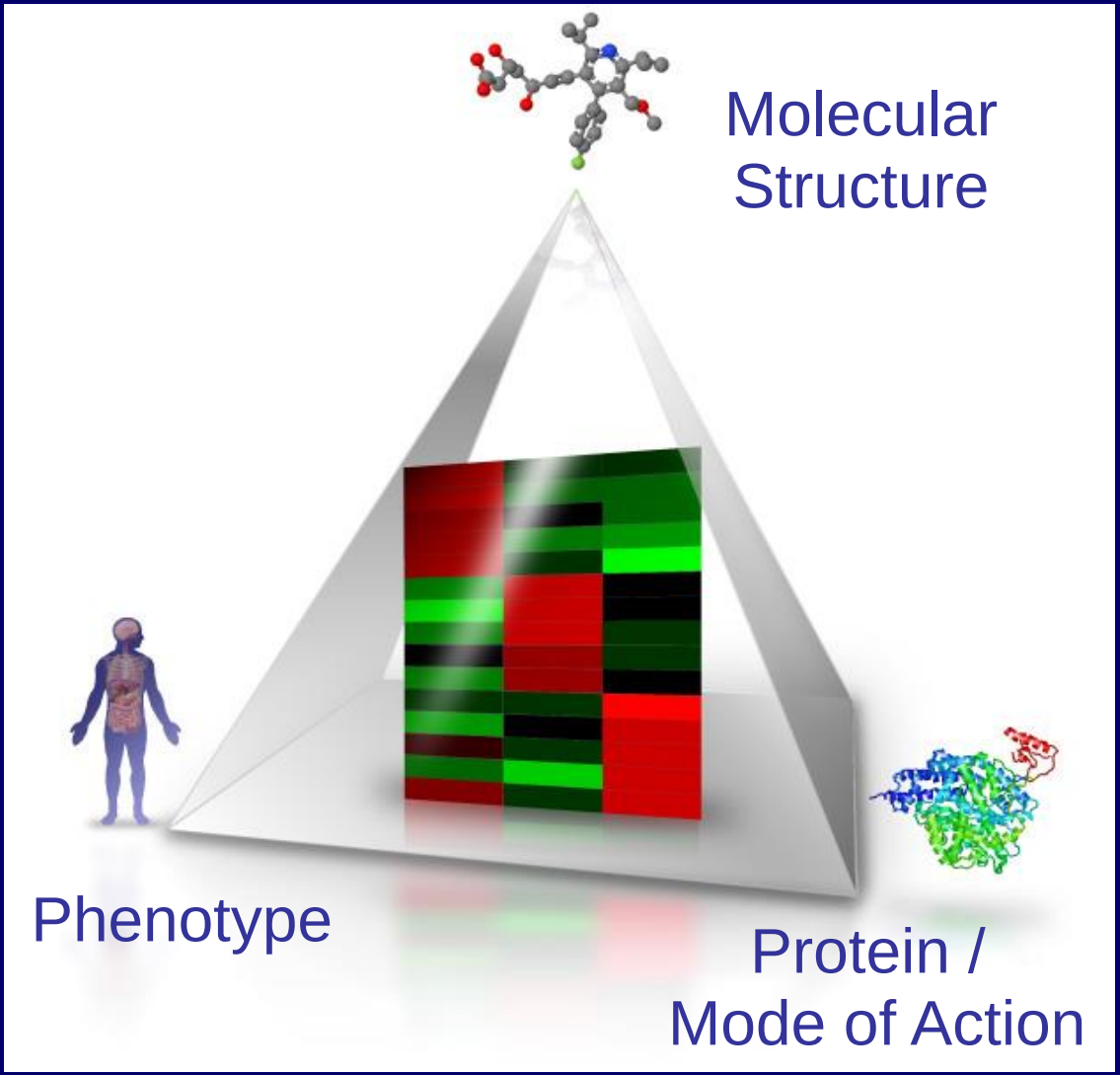
More and More Data is Available...

- But: How should we deal with it?

- Databases contain tens of millions of bioactivity data points, gene expression data, organ tox endpoint data, clinical trial data, ...
- *However*, integration – and utilization – of data is often not ideal

- This is what we aim to do in our group; integrate and analyze *heterogeneous* life science data; provide testable hypotheses; test those hypotheses

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*

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

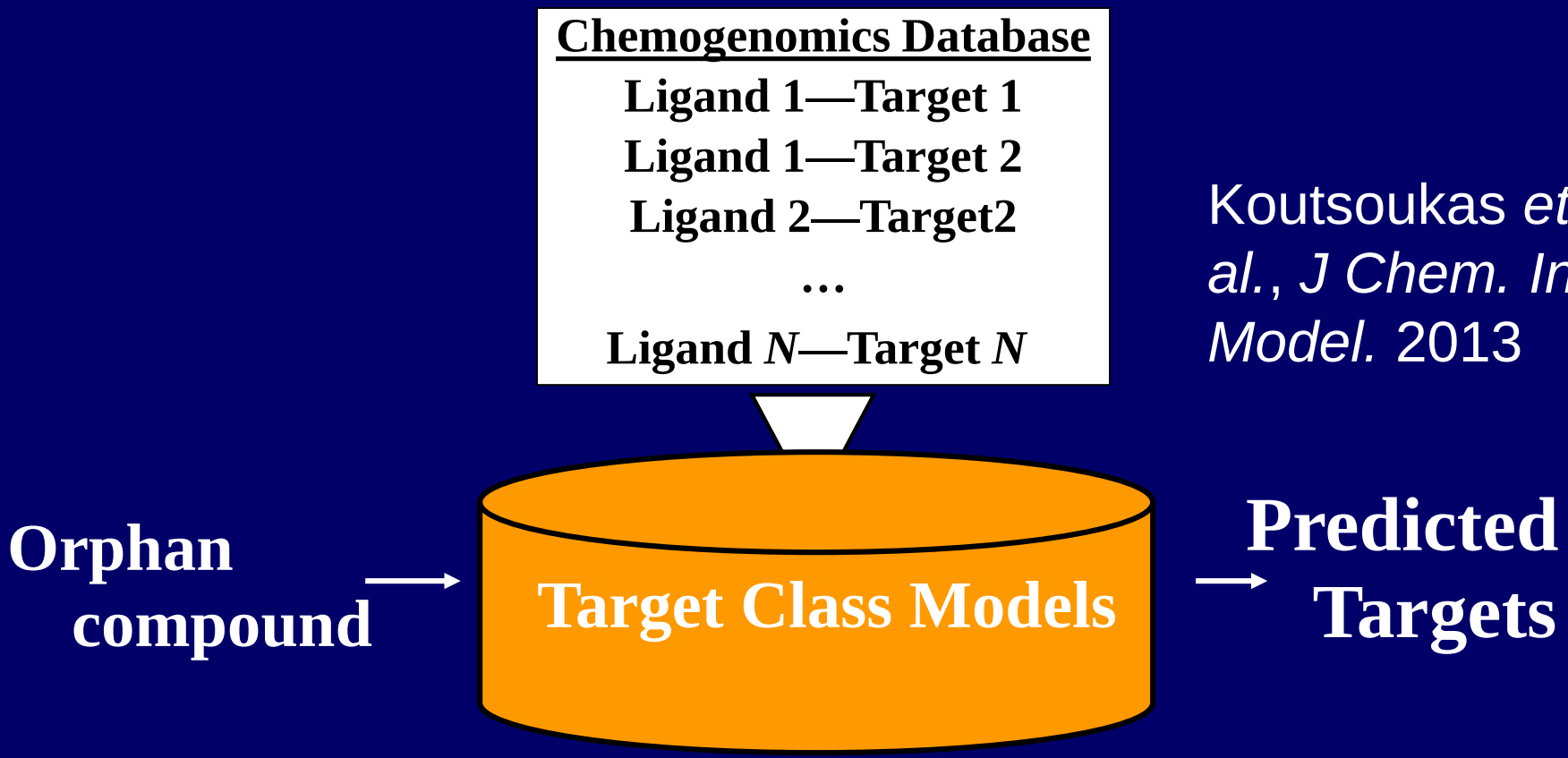
- Mixture modelling (~6; ERC Starting Grant)
- Traditional medicines/natural products (~3)

Integrating chemical and biological data

- Pharmacogenomics/toxicogenomics (~2)
- Gene expression/RNA-Seq data for compound selection and mode of action analysis etc. (~3)

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.



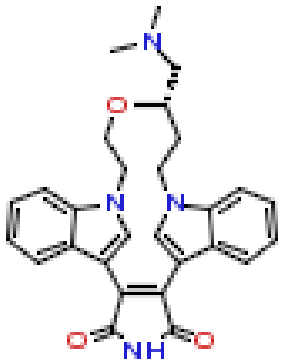
Koutsoukas *et al.*, *J Chem. Inf Model.* 2013

In silico target prediction, based on large databases:

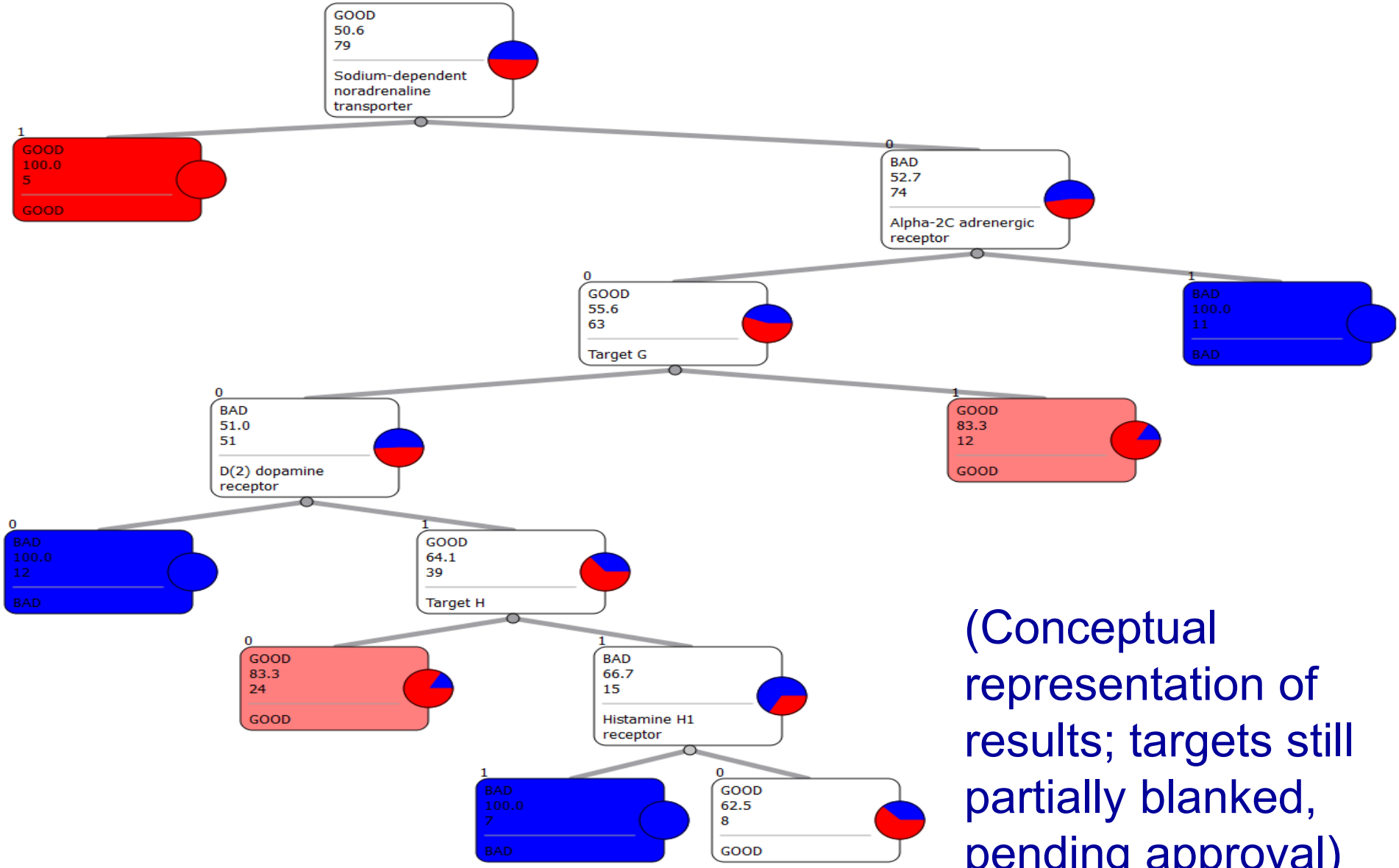
- Gleevec (Novartis),
  - Launched
  - Targets Bcr-Abl, c-kit, PDGFRb

- 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 polypharmacology of CNS-active compounds (with Eli Lilly)



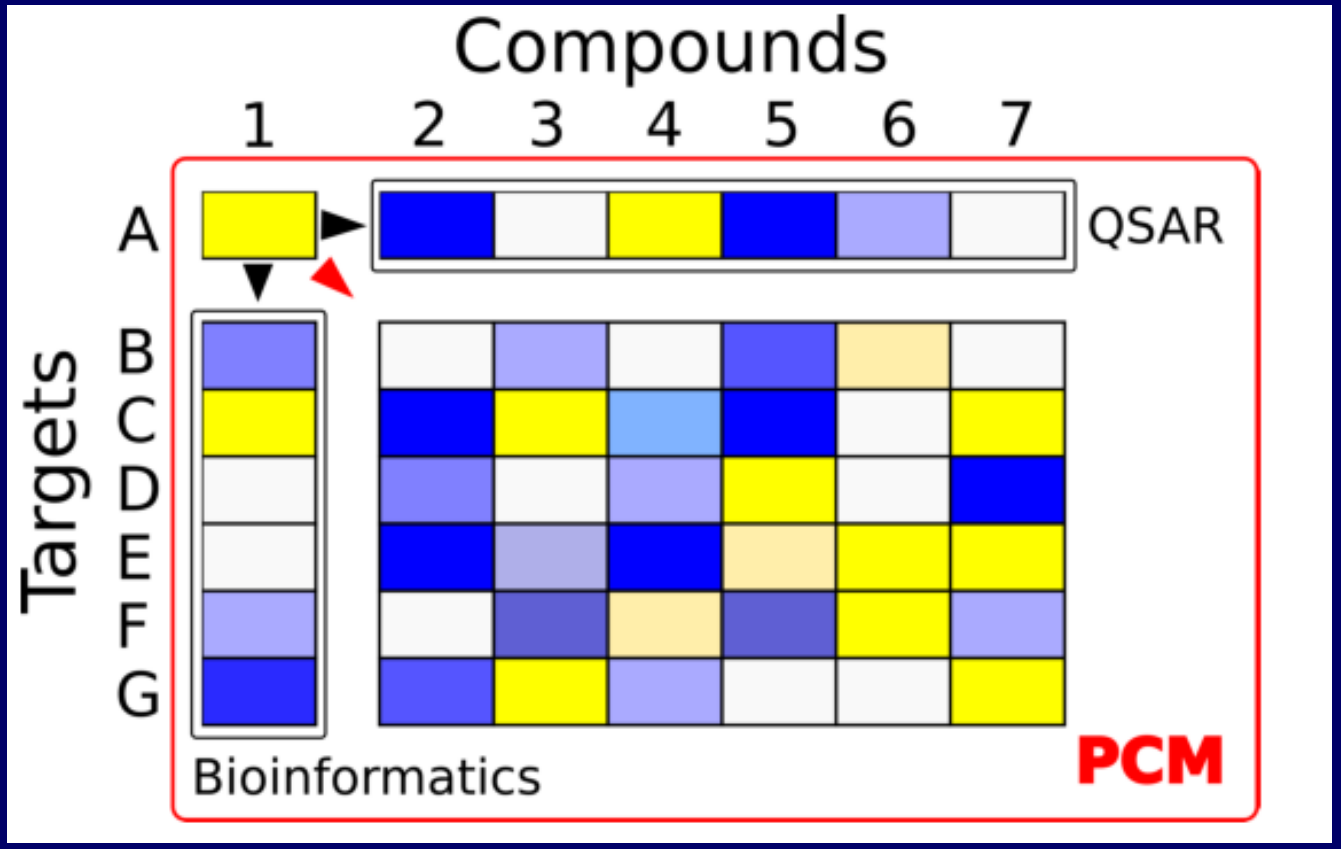
(Conceptual representation of results; targets still partially blanked, pending approval)

Items to consider for neglected diseases

- Targets are usually non-human proteins (= for which relatively little ligand data exists)
- If sufficient assay data *against isolated proteins* exist methods is applicable

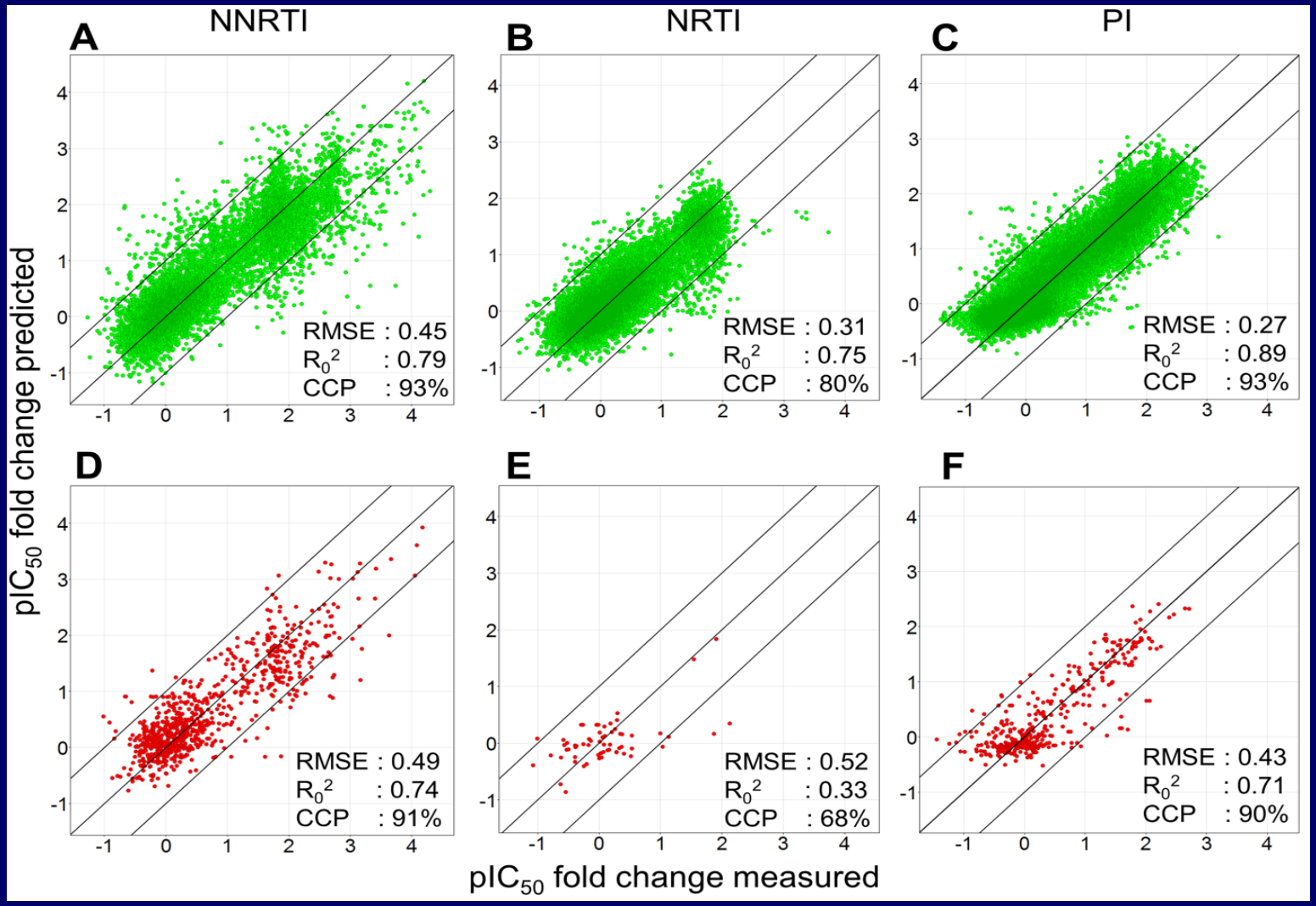
- Our experience with *P. falciparum*:
  - Extrapolation from human to parasite possible *in case of closely related proteins*
  - Otherwise *extrapolation remains difficult*
  - At least we can anticipate which cases likely work, which ones don’t

Integrated modelling of chemical and biological Data (eg for sets of protein mutants)



Isidro Cortes-Ciriano, Qurrat Ul Ain,*et al.* *MedChemComm* 2015, *Advance Article*

Drug efficacy prediction in HIV ‘works’ for >10k patients *across modes of action*



Project with J&J

PLoS Comp Biol 2013

Analysis of compound combination screens against *P falciparum*

- Compound combination screen against *P falciparum* was performed by the NIH
- About 600 6x6 dose response matrices available

- We aim to understand mechanistically which compound combinations are synergistic (and which ones antagonistic)
- Aim is to then propose novel combinations

- Currently still ongoing; happy to share results/methods when properly validated

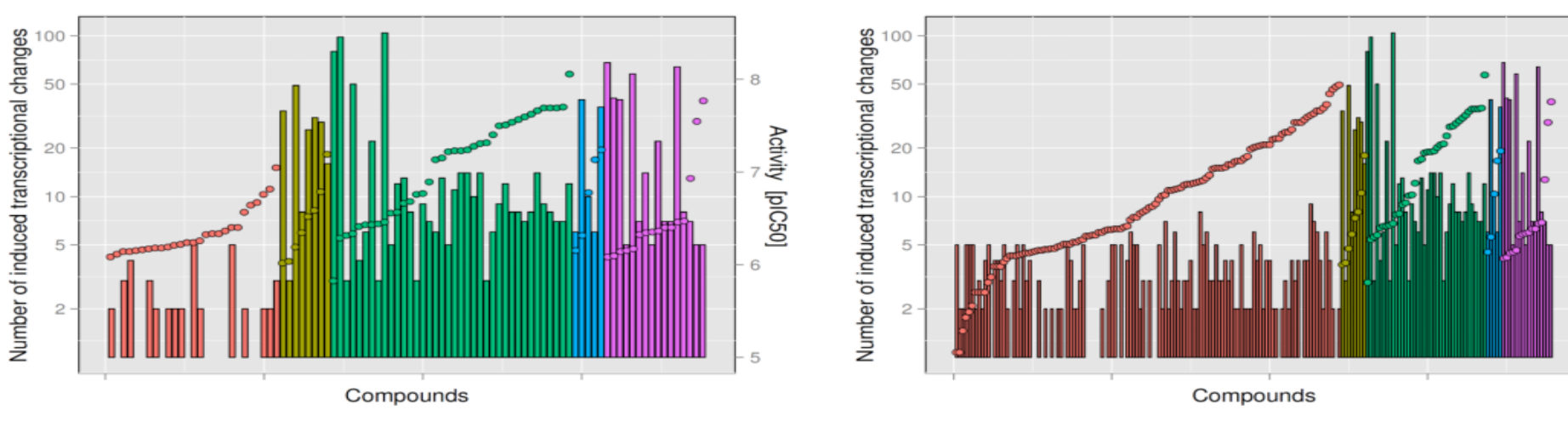
Evaluation of gene expression data for safety decision making

- Conducted by Johnson&Johnson/Janssen, with University of Cambridge (UK), Ghent University (BE), University of Rochester (US), Hasselt University (BE), KU Leuven (BE), Johannes Kepler University Linz (AT), and Durham University (UK)
- Funded by IWT

- Eight drug discovery projects in ‘big pharma’
- To what extent can gene expression data help with decision-making in lead optimization across disease areas, targets and scaffolds

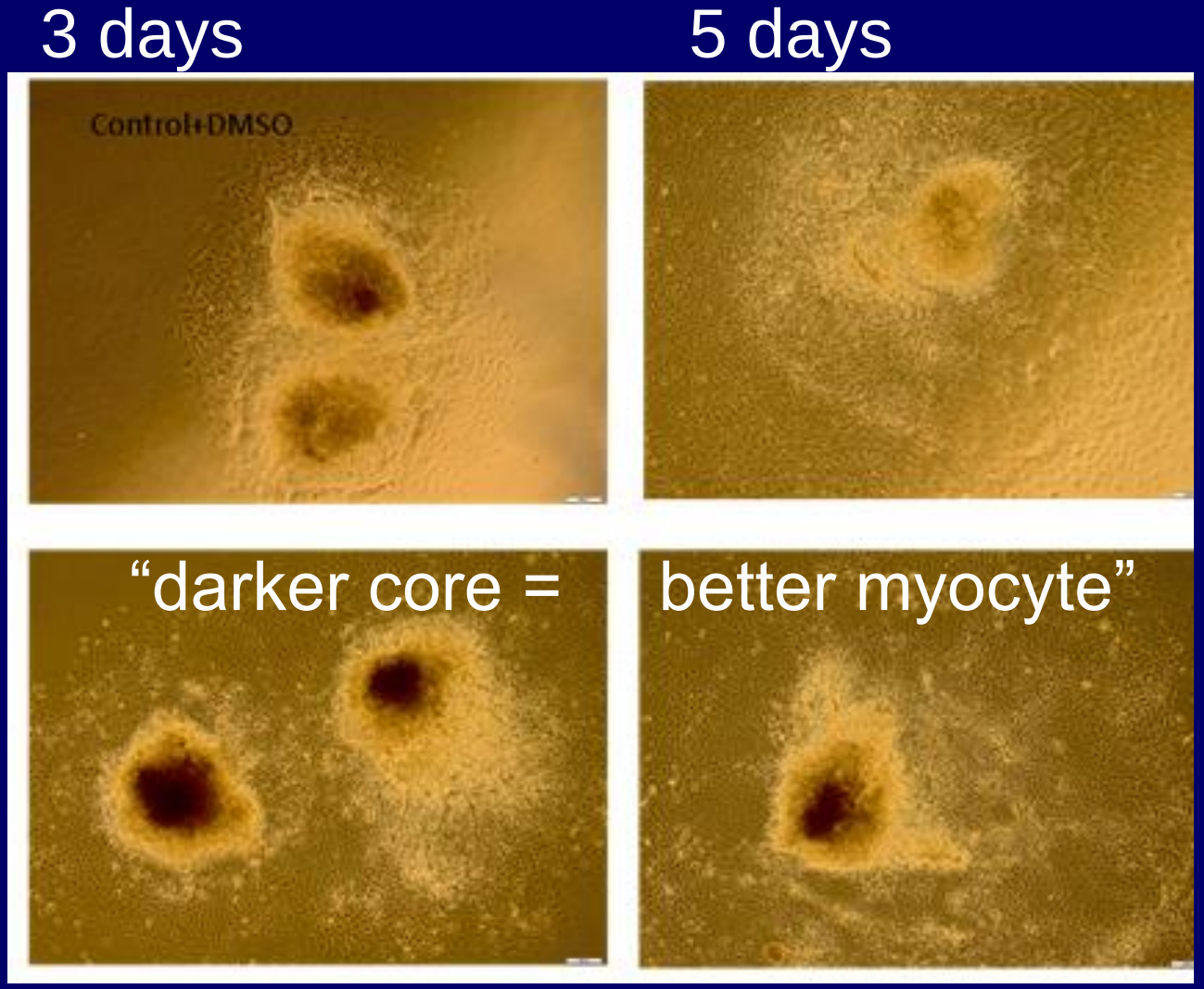
Prioritization of selective scaffold series for ROS1 project

- Bars: number of transcriptional changes; color: chemotype (A-E); dots: Activity on target



- Scaffold A was most selective, but not most potent (left)
- Follow up compounds *retained selectivity, increased potency*

Using gene expression data to select compounds, eg for stem cell differentiation (work with Dr Nasr, Royan Institute, Isfahan)

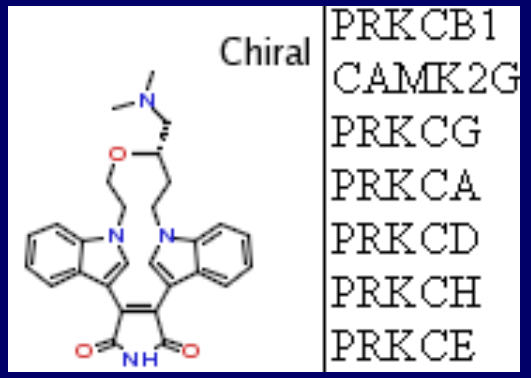


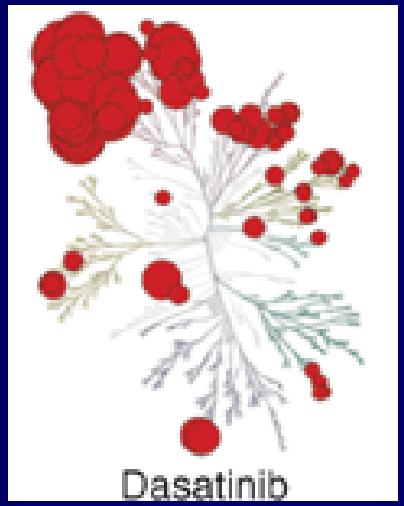
So what can our group contribute to research on neglected diseases?

- Data analysis methods to rationalize/understand phenotypic screening data

- Compound selection/prioritization methods, also for mutant/multiple protein targets

- Expertise in analyzing gene expression data related to both efficacy and toxicity of compounds





Acknowledgments

Aakash Ravindranath
Ain Qurrat
Alexios Koutsoukas
Avid Afzal
Bobby Glen
Chad Allen
Daniel Murrell
David Marcus
Fatima Baldo
Fazlin Mohd Fauzi
Georgios Drakakis
Krishna Bulusu
Lewis Mervin
Oscar Mendez Lucio
Richard Lewis
Rucha Chiddarwar
Shardul Paricharak
Salundi Basappa
Sharif Siam
Siti Zuraidah Sobir
Sonia Liggi
Sudeshna Guha Neogi
Yasaman Motamedi

Ad P. IJzerman
Bart Leidselink
Gerard van Westen

Sebastian Rohrer
Stefan Tresch
Antje Wolf
Klaus-Juergen Schleifer

Ola Engkvist
Thierry Kogej

Martin Augustin
Tom Klenka

Therese Malliavin
Isidro Cortes

Ian Stott

Hinrich Goehlmann
Herman van Vlijmen
Joerg K. Wegener

Mike Bodkin
David Evans
Suzanne Brewerton

Massimo Mercato
Anna Maidecchi