Natalie.Fey@Bristol.ac.uk

University of BRISTOL

Data-driven discovery and optimisation of organometallic catalysts

Dr Natalie Fey

Dial-a-Molecule Annual Meeting

3rd July 2019, York

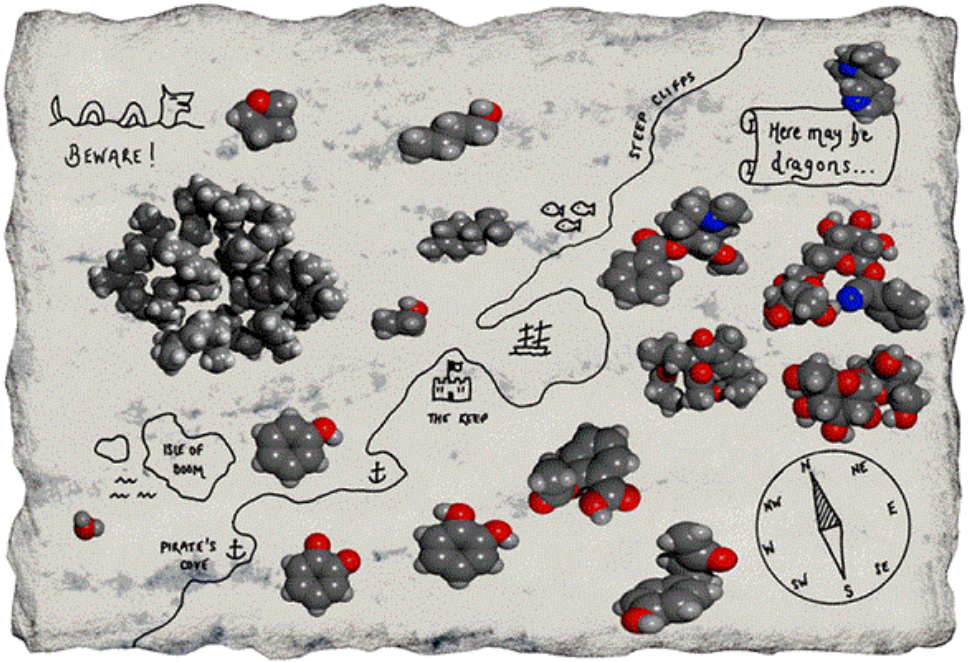
bristol.ac.uk

Opportunities from Data

- Select “best” synthetic route
  - Better use of resources/time.
  - Break “cultural attachments” to routes.
  - Find new approaches.
  - When to stop.
- Broaden our toolkit, explore new chemistry.
- Deepen our understanding (mechanism, catalyst/substrate matching).

- Combine experiment and computation to work towards prediction.

Property Databases

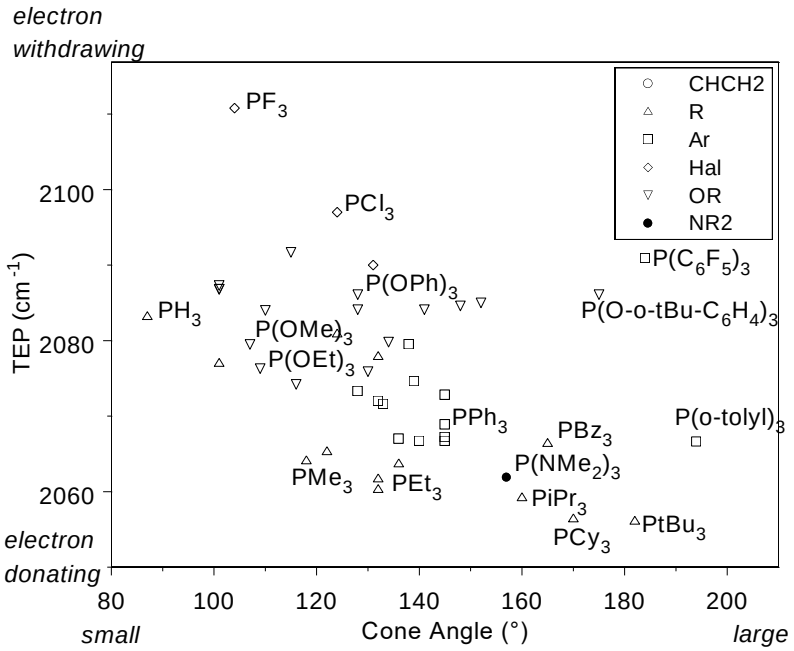


Chem. Central J. **2015**, *9*, 38

Ligand Knowledge Bases

- Ligands used to tune complex properties.
- Characterise ligands (size, electronics).
- Map ligand properties.
- Derive structure-property relationships.

- Are two dimensions sufficient?
- Transferability to other ligand types.

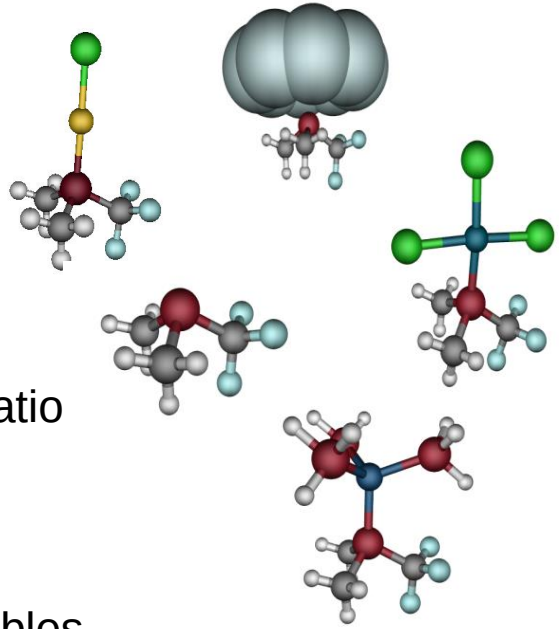


Tolman, *Chem. Rev.* **1977**, *77*, 313–348.

For other examples: QALE, Cundari, Rothenberg.

See *Coord. Chem. Rev.* **2009**, *253*, 704-722.

LKB-P



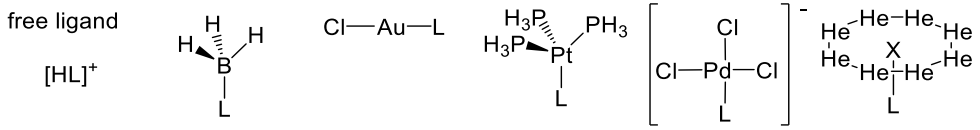
- 28 descriptors, quite highly correlated.

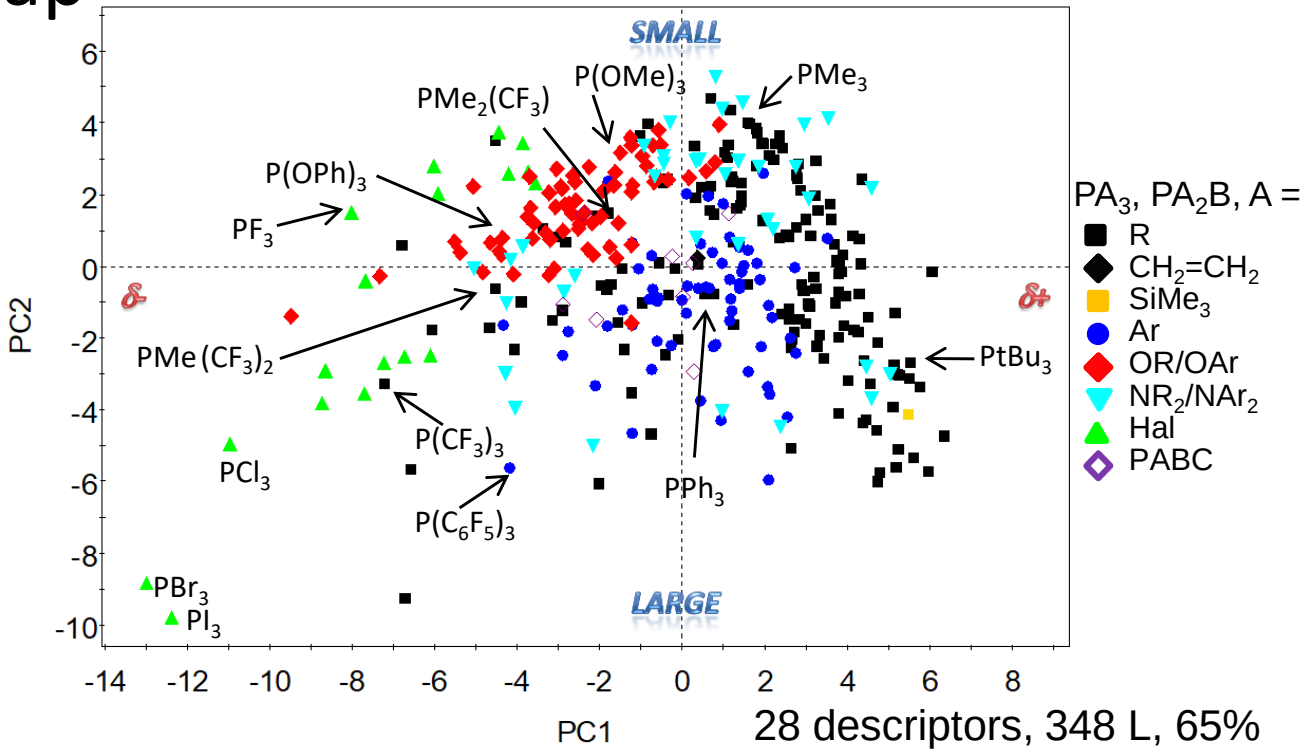
- Projection methods:
  - Reduce number of variables (dimensions) by transformatio variables.
  - Optimally represent distances between objects.
- Principal Component Analysis
  - Describe variation of data in terms of uncorrelated variables.
  - Derive from linear combinations of original variables.
  - Maximum variance in first PC, decreasing order of importance, orthogonal.

$$PC_1 = a_{1,1}v_1 + a_{1,2}v_2 + \ldots + a_{1,n}v_n$$
$$PC_2 = a_{2,1}v_2 + a_{2,2}v_2 + \ldots + a_{2,n}v_n$$

BP86/6-31G*, LACV3P on Au, Pd, Pt

LKB-P Map





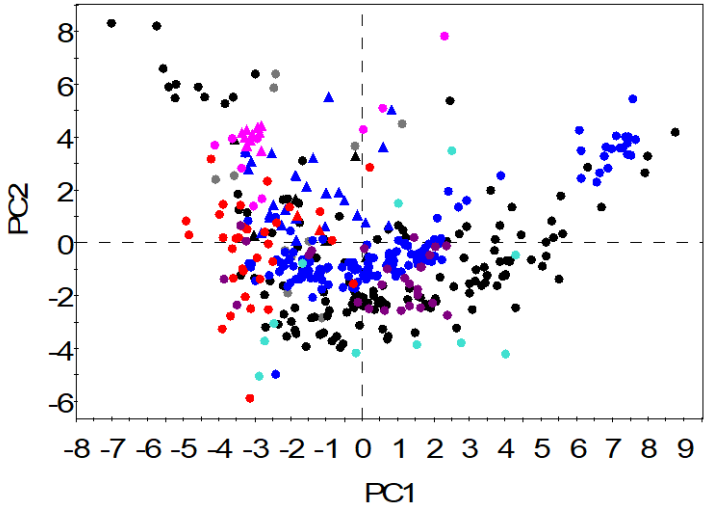
28 descriptors, 348 L, 65%

bristol.ac.uk

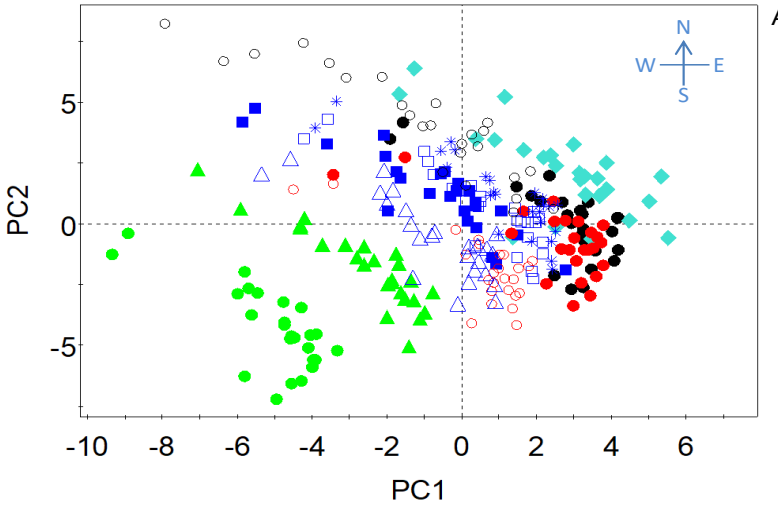
Dr Jesús Jover, with JNH, AGO, GCLJ

Organometallics, **2010**, *29*, 6245;
Chem. Eur. J. **2006**, *12*, 291

Other ligands

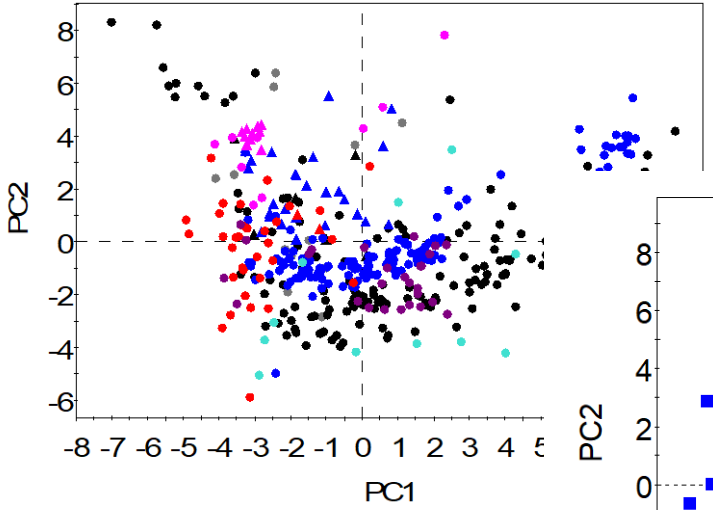


LKB-PP
Organometallics, **2012**, 31, 5302;
Organometallics **2008**, *27*, 1372

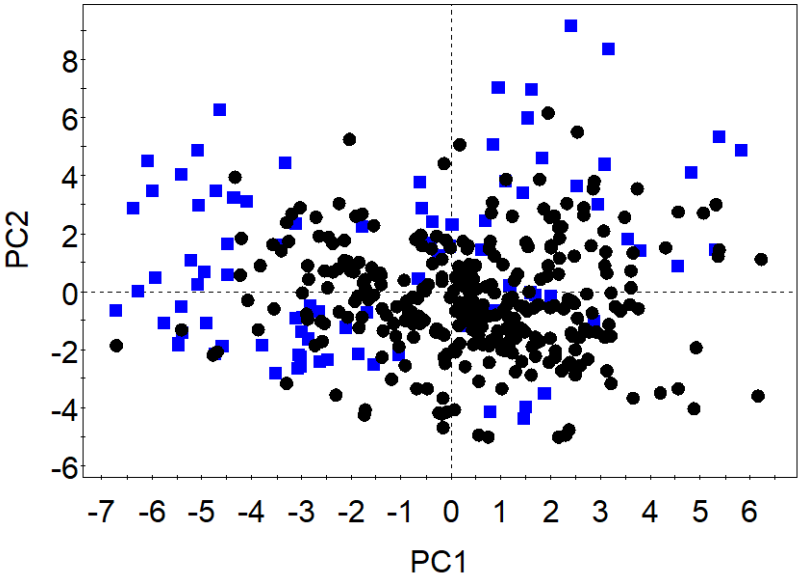


LKB-PP$_{screen}$
Dalton Trans. **2013**, 42, 172

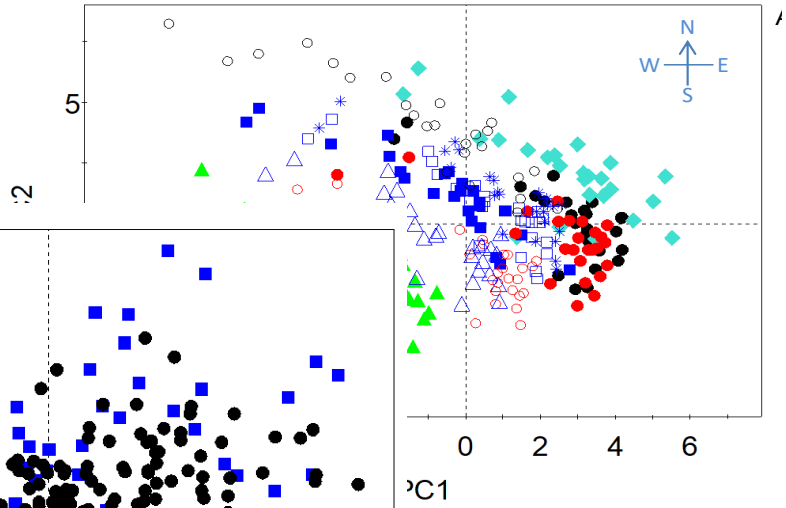
Other ligands



LKB-PP
Organometallics, **2012**, 31
Organometallics **2008**, *27*



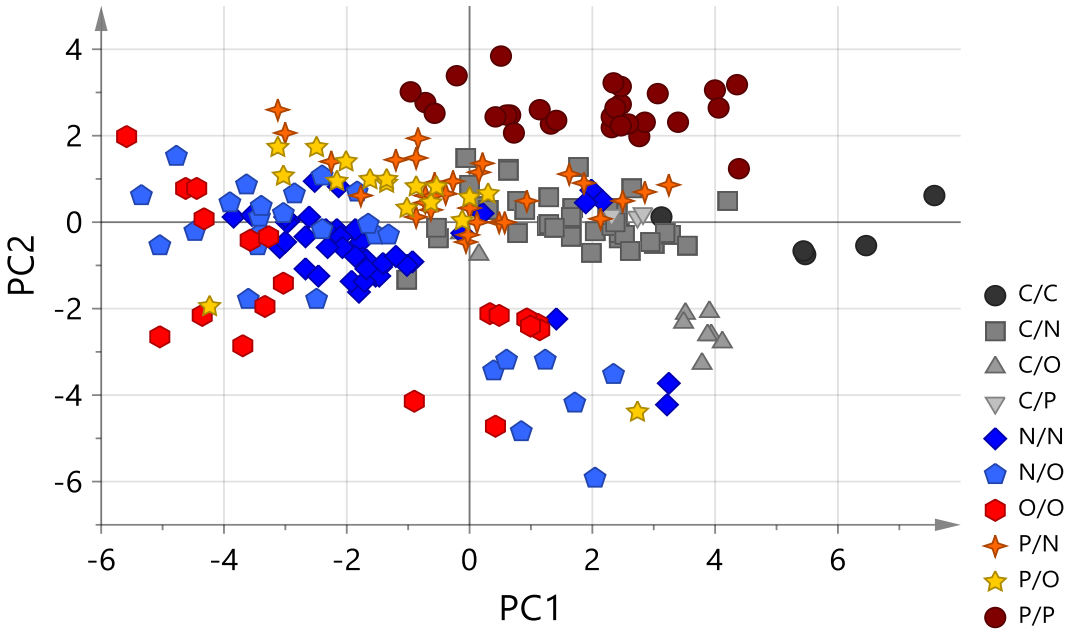
LKB-C
Dalton Trans. **2009**, 8183

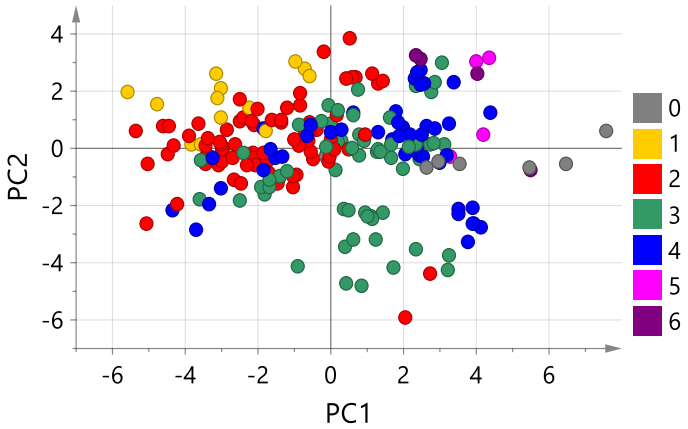


2013, 42, 172

bristol.ac.uk

Bidentate Ligand Space





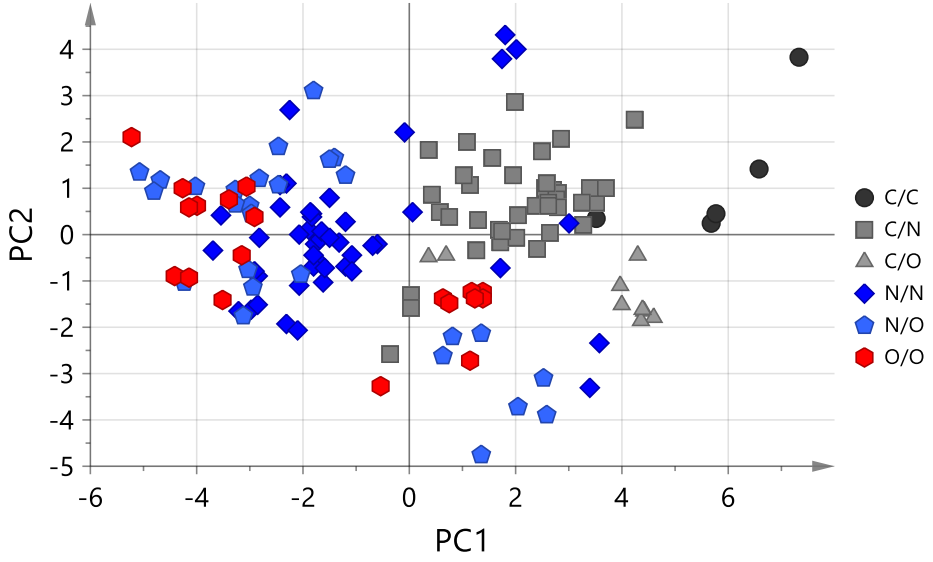
224 ligands, 52%

bristol.ac.uk

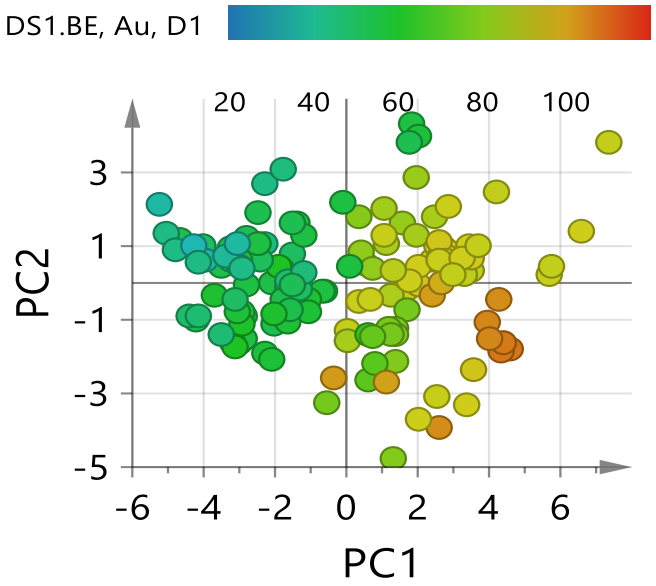
with Dr Charlotte Willans, Dr Bao Nguyen, Prof. Andrei Malkov, Dr Jonathan Moseley, Dr Simon Tyler

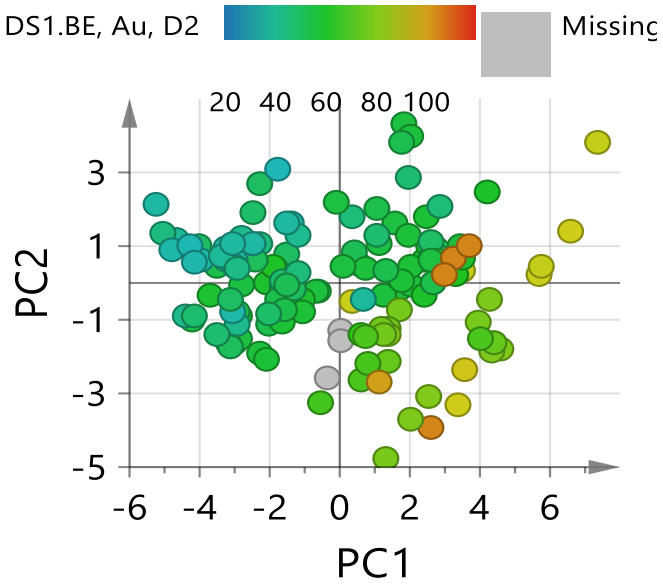
DIAL-A-MOLECULE
An EPSRC Grand Challenge Network

C-, N-, O-donor ligands only



146 ligands, 57%





Publication in preparation

bristol.ac.uk

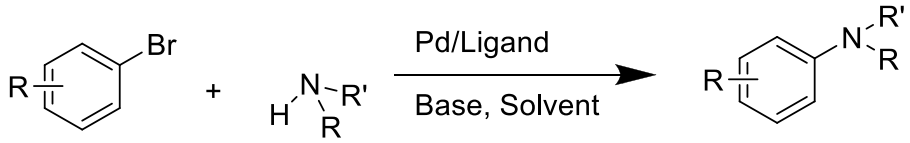
DIAL-A-MOLECULE
An EPSRC Grand Challenge Network

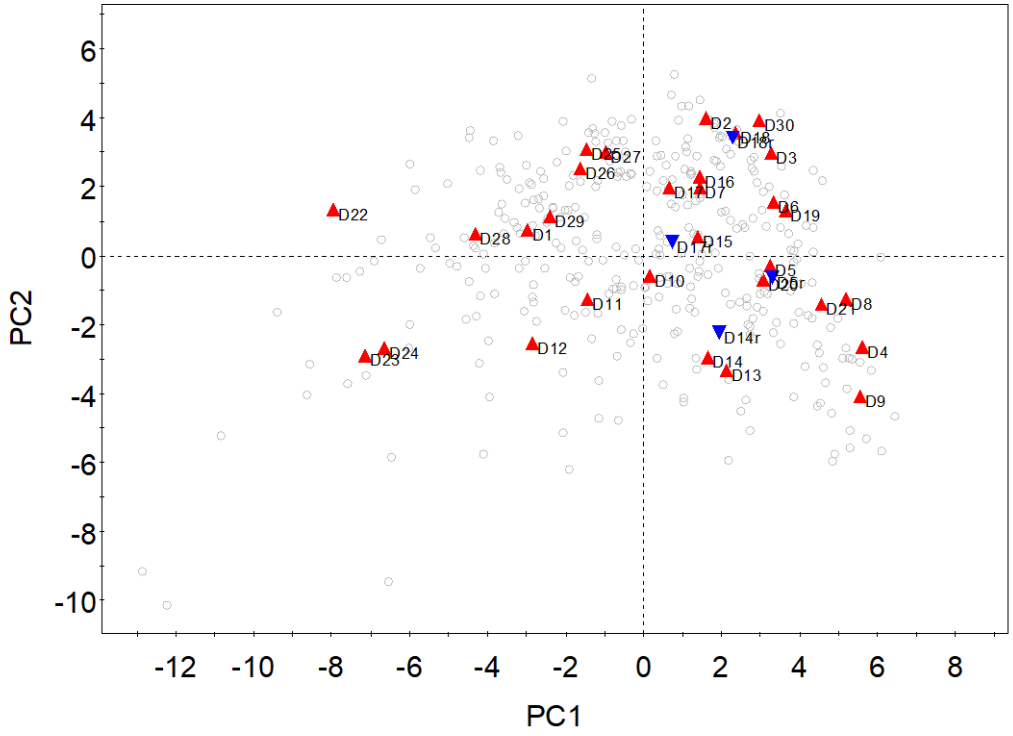
CatSci
Innovation in Catalysis

UNIVERSITY OF LEEDS

Loughborough University

Experiment

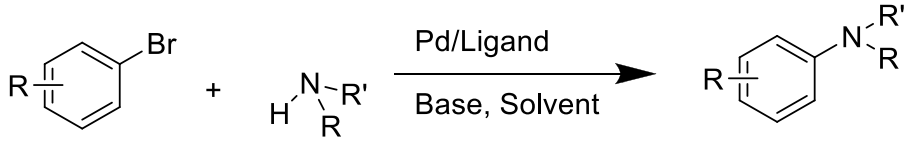


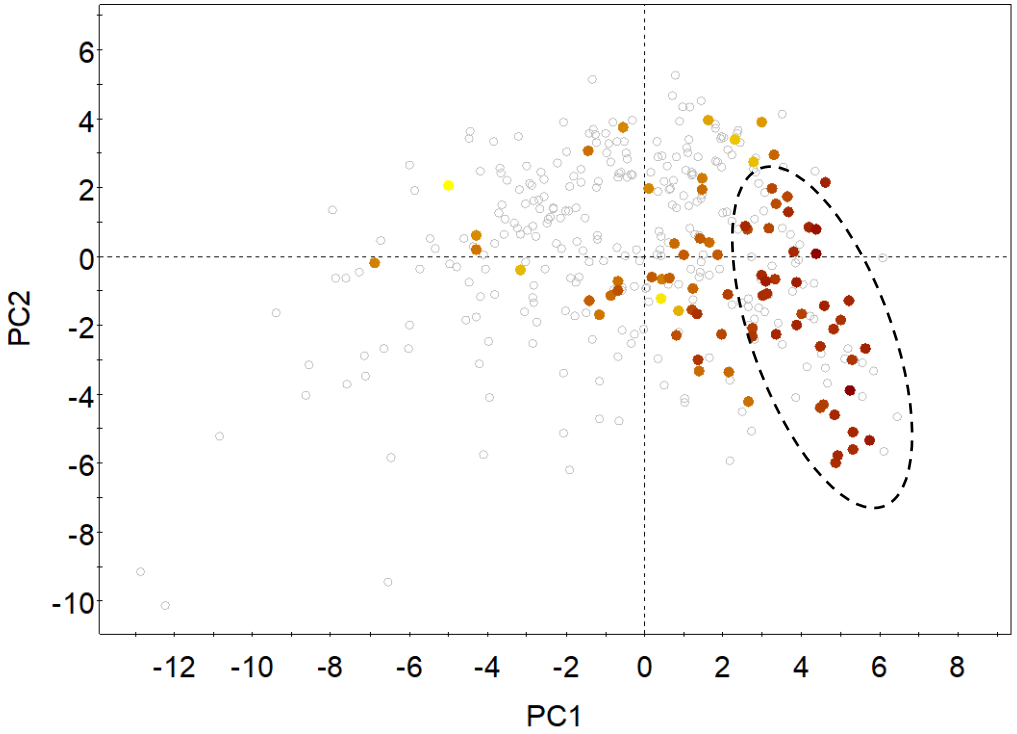


bristol.ac.uk

Dr Gareth Owen-Smith, Dr Jesús Jover, with Jeremy Harvey, Guy Orpen, Guy Lloyd-Jones

Experiment

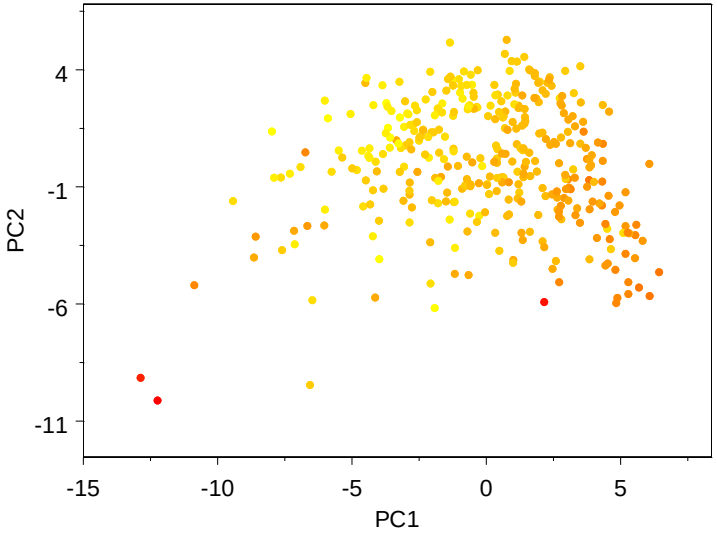


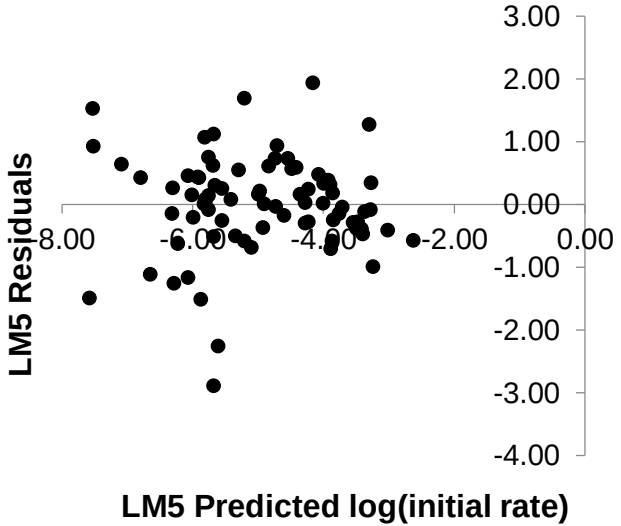


bristol.ac.uk

Dr Gareth Owen-Smith, Dr Jesús Jover, with
Jeremy Harvey, Guy Orpen, Guy Lloyd-Jones

Global Models:





LM5: Q(B fragm.), Q(Pd fragm.), Q(Pt fragm.), BE(Au), BE(Pd), P-Pt, ΔP-A(B), ΔP-A(Pt) [8], R^2 = 0.687, CV PE = 0.83

bristol.ac.uk

Dr Gareth Owen-Smith, Dr Jesús Jover, with Jeremy Harvey, Guy Orpen, Guy Lloyd-Jones

LASSO

(1000 random samples)

In statistics and machine learning, **lasso** (least absolute shrinkage and selection operator) (also **Lasso** or **LASSO**) is a **regression** analysis method that performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability of the statistical model it produces.

Lasso (statistics) - Wikipedia

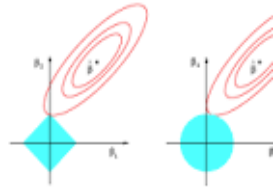
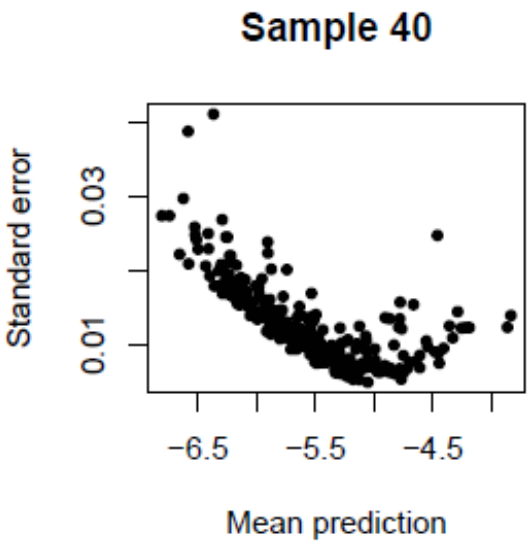


Figure 3.12: *Estimation picture for the lasso and ridge regression (right). Shown are contours error and constraint functions. The solid blue ar the constraint regions* $|\beta_1| + |\beta_2| \le t$ *and* $\beta_1^2 + \beta$

onlinecourses.science.psu.e

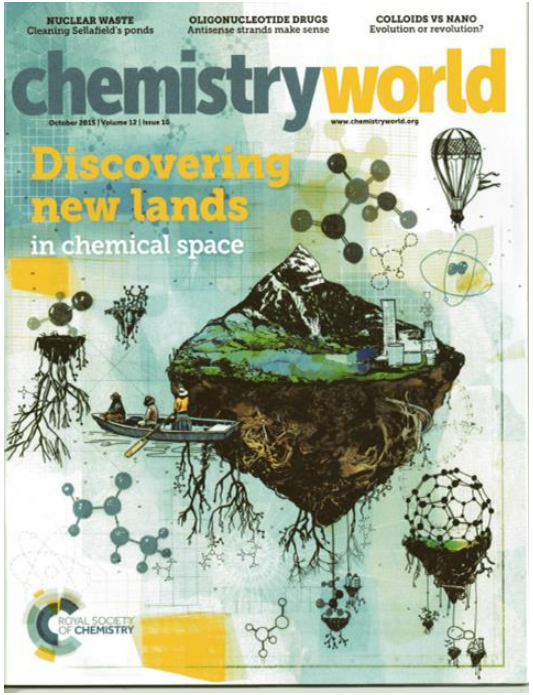


bristol.ac.uk

Dr Ben Swallow (JGI Seedcorn Funding)

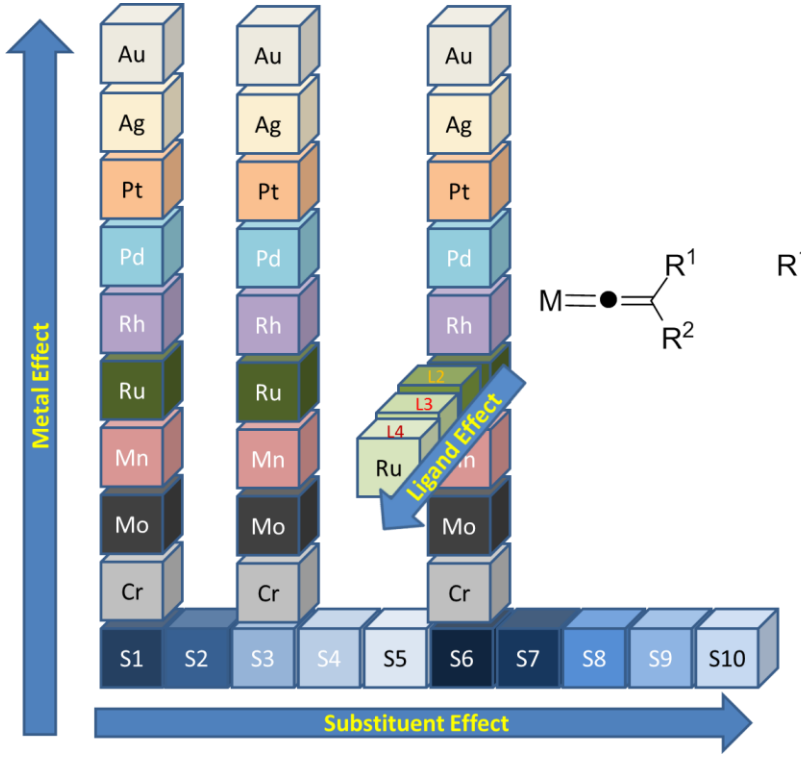
University of BRISTOL

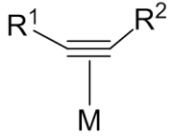
Beyond Ligands…



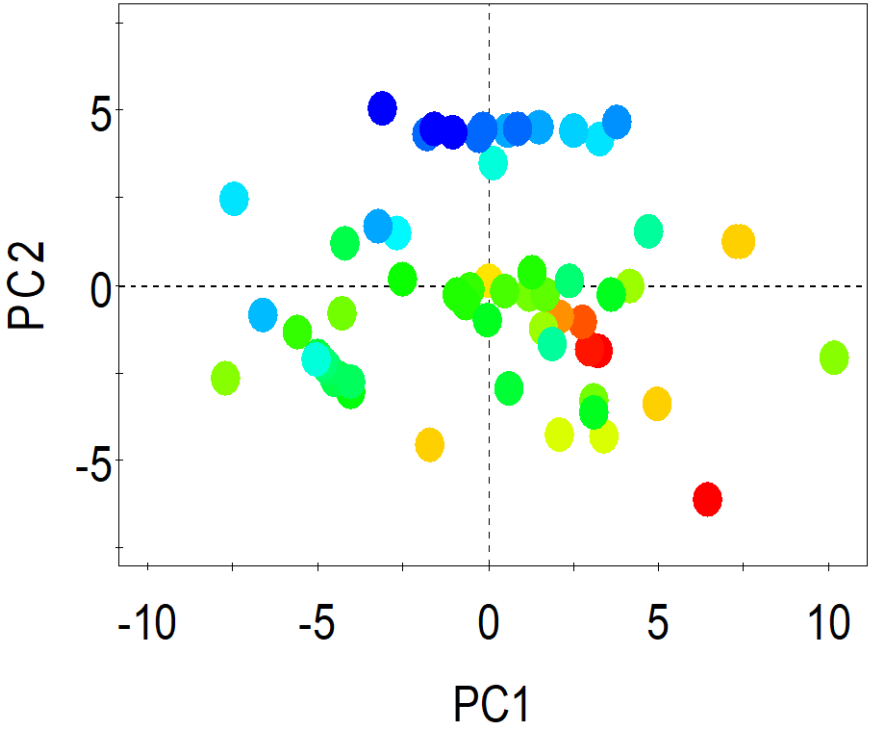
bristol.ac.uk

Alkynes & Vinylidenes





Tautomer Preference $\Delta E_{v\text{-}a}$(Ru)



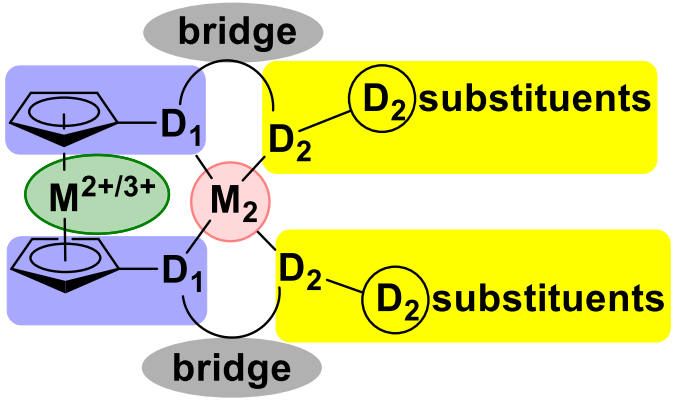
Chem. Commun. **2015**, *51*, 9702

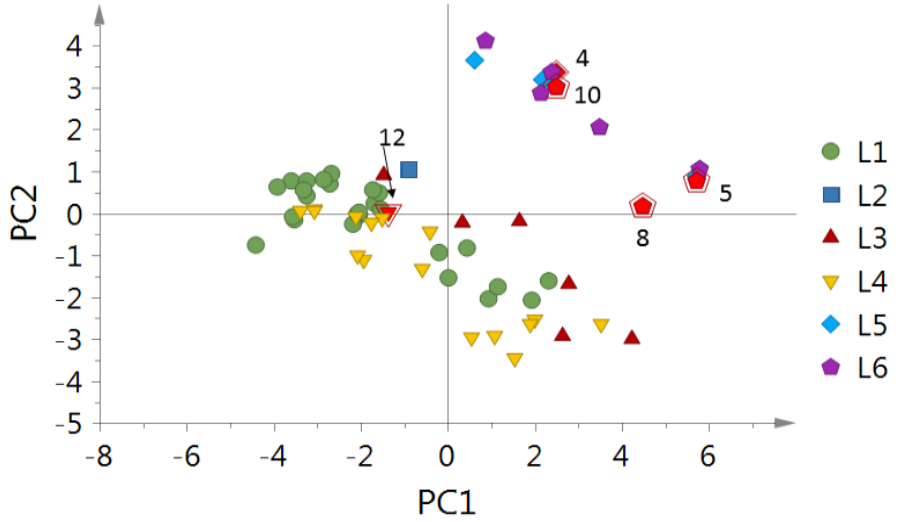
bristol.ac.uk

with Dr Jason Lynam,
Organometallics, **2014**, *33*, 1751

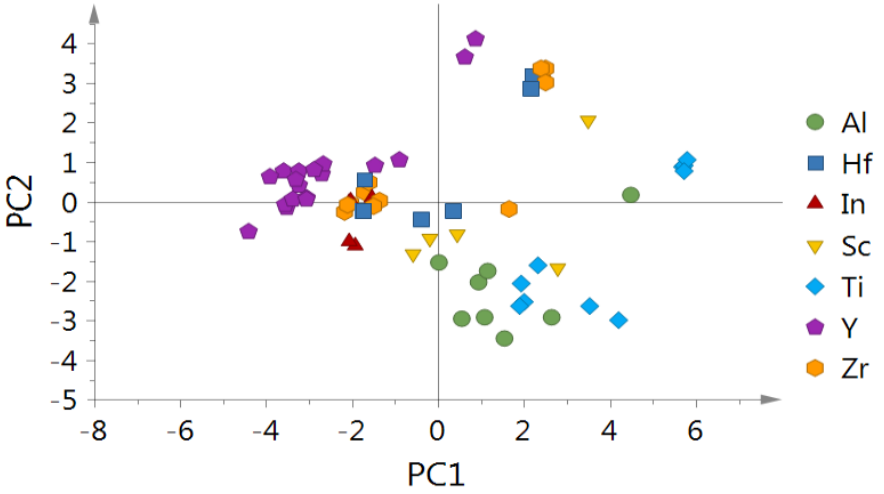
UNIVERSITY of York

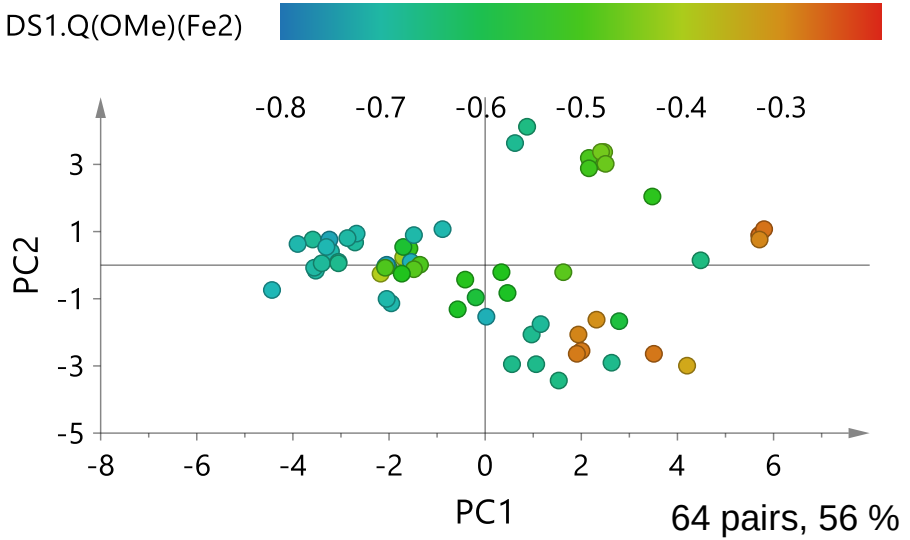
Redox-Switchable Catalysts



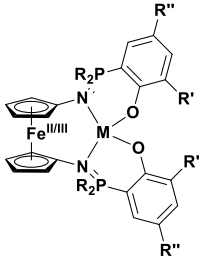


64 pairs, 56 %

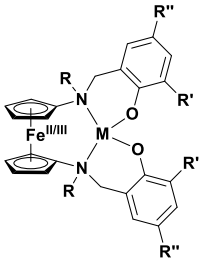




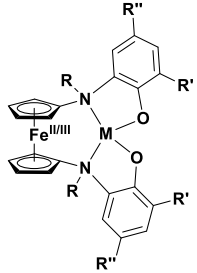
64 pairs, 56 %



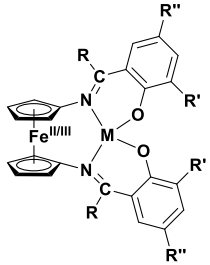
ML1 (phosphen)
M = Sc(OMe), Y(OMe), $Ti(OMe)_2$, $Zr(OMe)_2$, $Hf(OMe)_2$, Al(OMe), In(OMe)
R = H, Me, Ph, CF_3, OMe, OPh, *t*Bu
R' = H, Me, *t*Bu; R'' = H, Me, *t*Bu



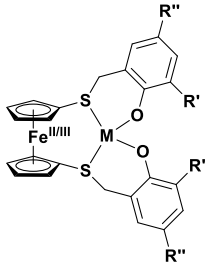
ML2 (salfan)
M = Y(OMe)
R = R' = R'' = H



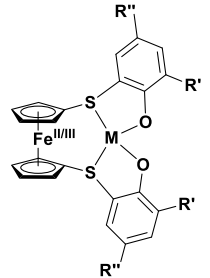
ML3 (salfan*)
M = Sc(OMe), Y(OMe), $Ti(OMe)_2$, $Zr(OMe)_2$, $Hf(OMe)_2$, Al(OMe)
R = H, R' = *t*Bu, R'' = H



ML4 (salfen)
M = Y(OMe), $Ti(OMe)_2$, $Zr(OMe)_2$, $Hf(OMe)_2$, Al(OMe), In(OMe)
R = H; R' = Me, CF_3, *t*Bu; R'' = H, Me



ML5 (thiolfan)
M = Y(OMe), $Ti(OMe)_2$, $Zr(OMe)_2$, $Hf(OMe)_2$
R' = R'' = H, *t*Bu



ML6 (thiolfan*)
M = Sc(OMe), Y(OMe), $Ti(OMe)_2$, $Zr(OMe)_2$, $Hf(OMe)_2$, Al(OMe)
R' = R'' = H, *t*Bu

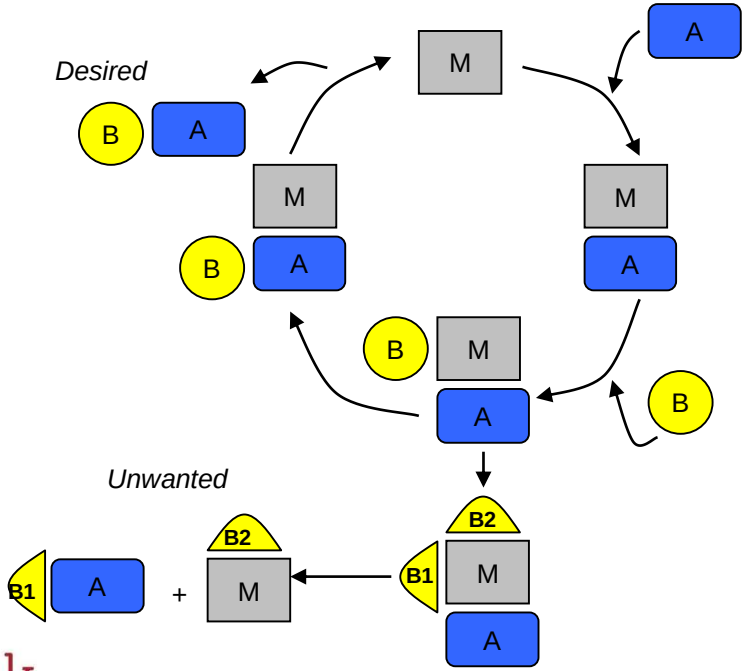
bristol.ac.uk

Chem. Commun., 2019, **55**, 7021-7024

UCLA College
Chemistry & Biochemistry

University of BRISTOL

Mechanism



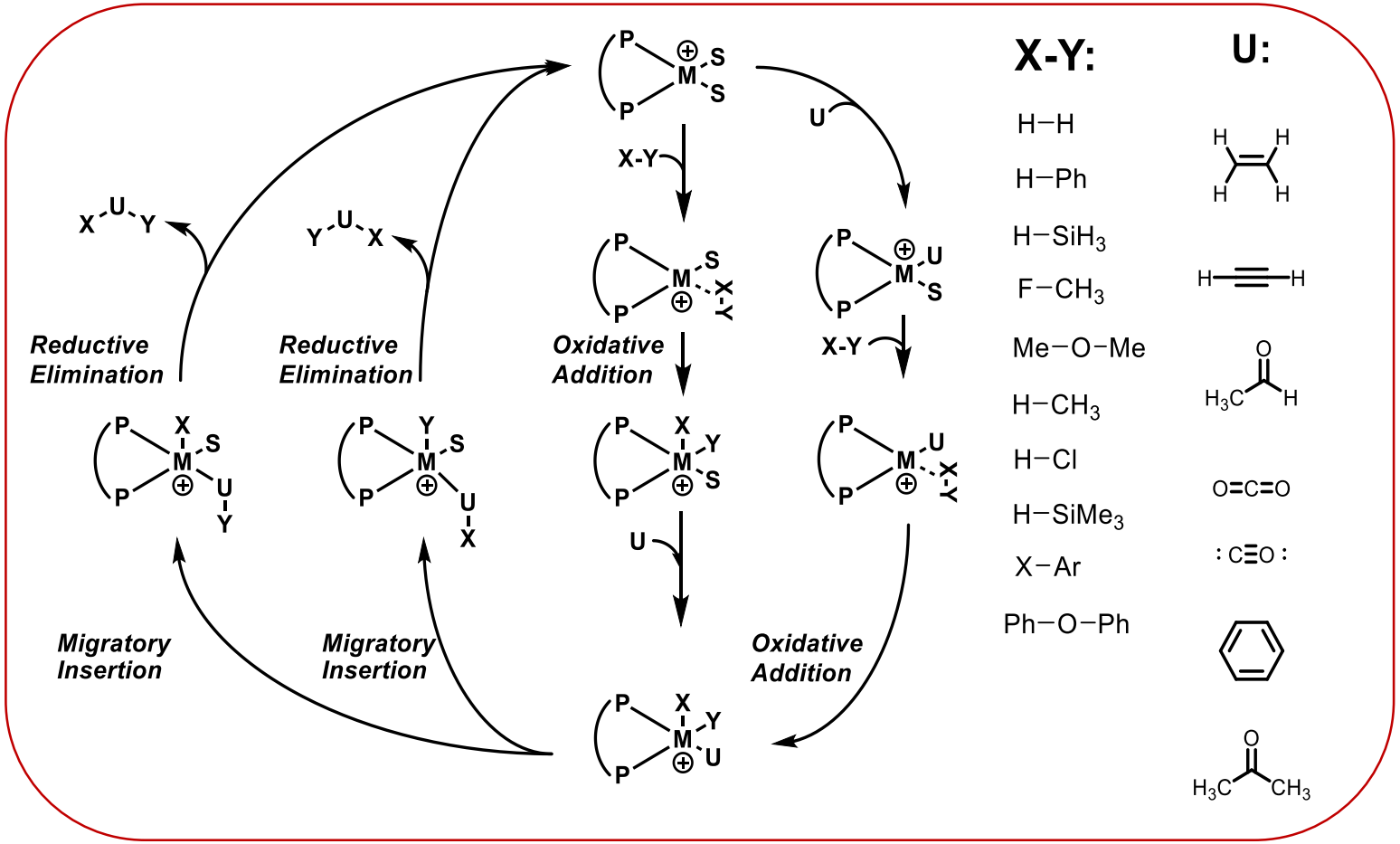
bristol.ac.uk

Bristol Reactivity in Catalysis Knowledgebase (BRiCK)

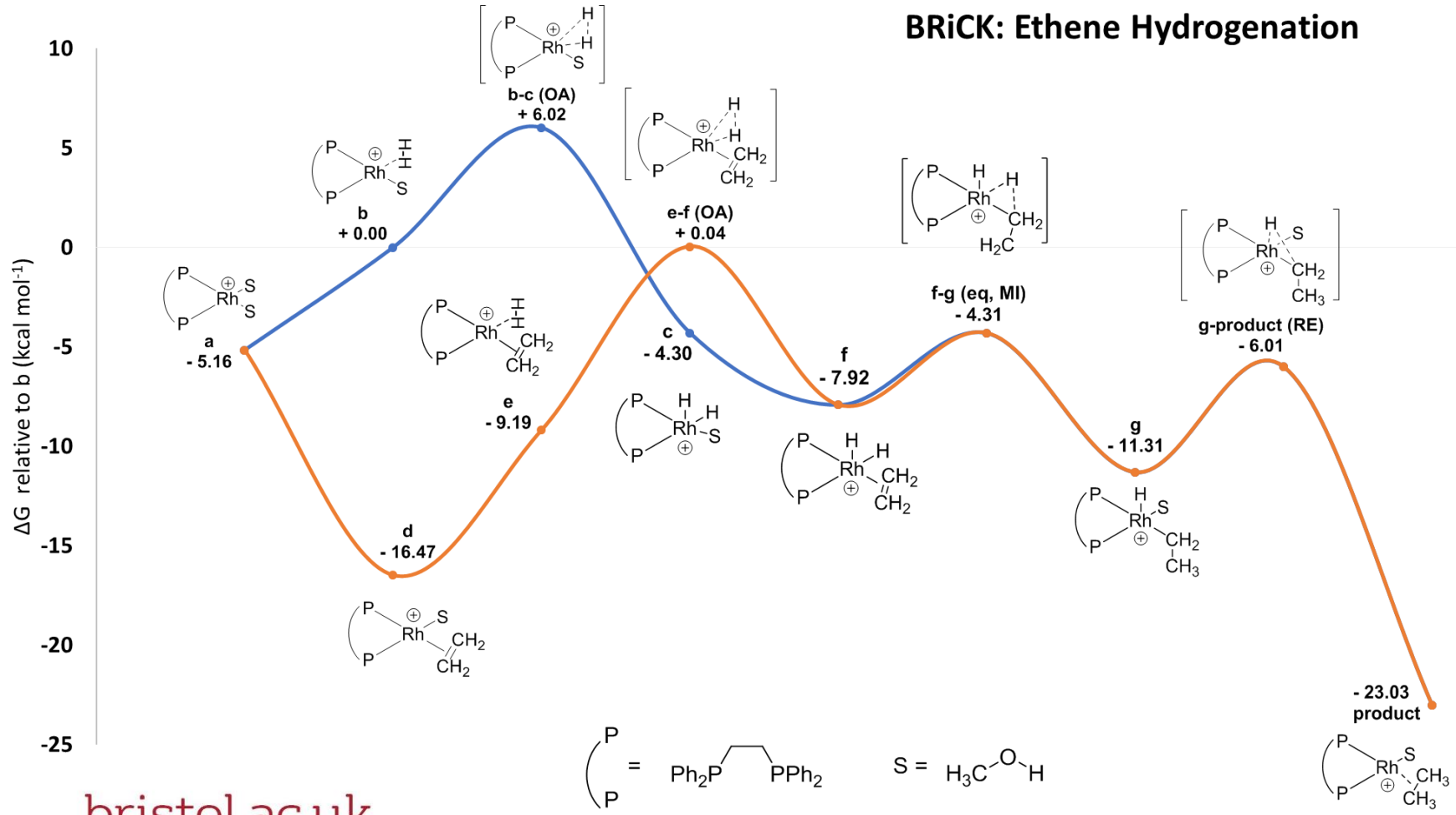
- Target: New applications for known catalysts.

- Systematic study of catalytic cycles.
  - Phase 1: Substrate effects on barriers.
  - Phase 2: Catalysts modification (metal, ligands).
  - Phase 3: Test experimentally.
  - Repeat…

Derek Durand, with Prof. Paul Pringle, Dr Jason Lynam,
Dr Ruth Webster, Dr Alex Cresswell,
coming soon: Dr Thibault Cantat, Dr Jordi Carbo


Reductive Elimination
Reductive Elimination
Oxidative Addition
Oxidative Addition
Migratory Insertion
Migratory Insertion
X-Y:
H–H
H–Ph
H–SiH3
F–CH3
Me–O–Me
H–CH3
H–Cl
H–SiMe3
X–Ar
Ph–O–Ph
U:

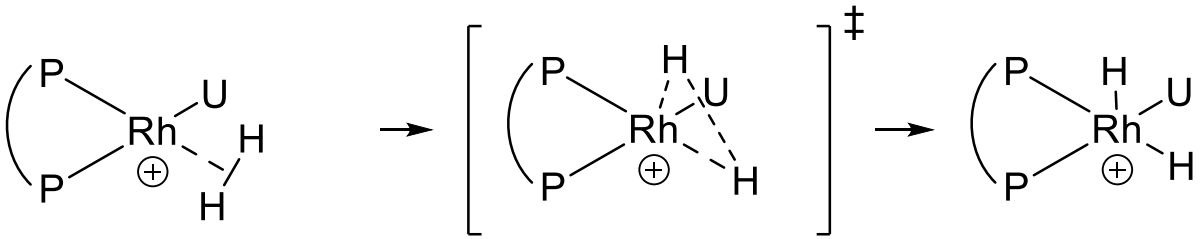
BRiCK: Ethene Hydrogenation



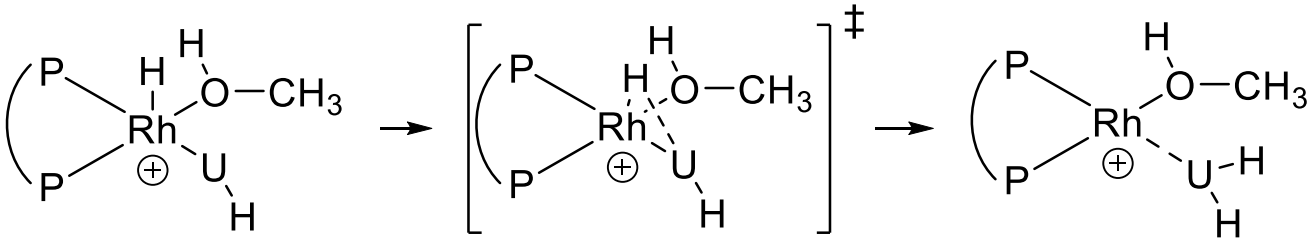
bristol.ac.uk

Derek Durand

Jaguar 8.5, B3PW91-D3, 6-31G(d), 298 K, PCM (methanol)



Substrate (U)	ΔΔG Oxidative Addition (kcal mol^{-1})
Ethene	9.23
Acetylene	4.90
Carbon Monoxide	8.71



Substrate (U)	ΔΔG Reductive Elimination (kcal mol^{-1})
Ethene	5.29
Acetylene	5.12
Carbon Monoxide	3.93

Derek Durand

Conclusions

- Databases of descriptors (ligands, substrates, catalysts).
  - Maps.
  - Models (experimental and computational responses).
  - Screening – design and optimisation.

- Mechanism.
  - Reasonable accuracy.
  - Multiple catalysts.
  - Selectivity and change in mechanism remain challenging.

We like data!

- Now moving on to mapping reactivity – BRiCK, with experimental testing.
- Extreme value prediction.

Acknowledgements

- Derek Durand, Claire McMullin, Jesús Jover, Gareth Owen-Smith, Ben Swallow, Mairi Haddow, Keshan Bolaky, Ben Rawe, Tim King
- Paul Pringle, Jeremy Harvey, Guy Orpen, Guy Lloyd-Jones
- Jason Lynam, Charlotte Willans, Bao Nguyen, Andrei Malkov, Jonathan Moseley, Simon Tyler, Paula Diaconescu, Ruth Webster, Alex Cresswell
- Bristol Centre for Computational Chemistry

EPSRC
Engineering and Physical Sciences Research Council

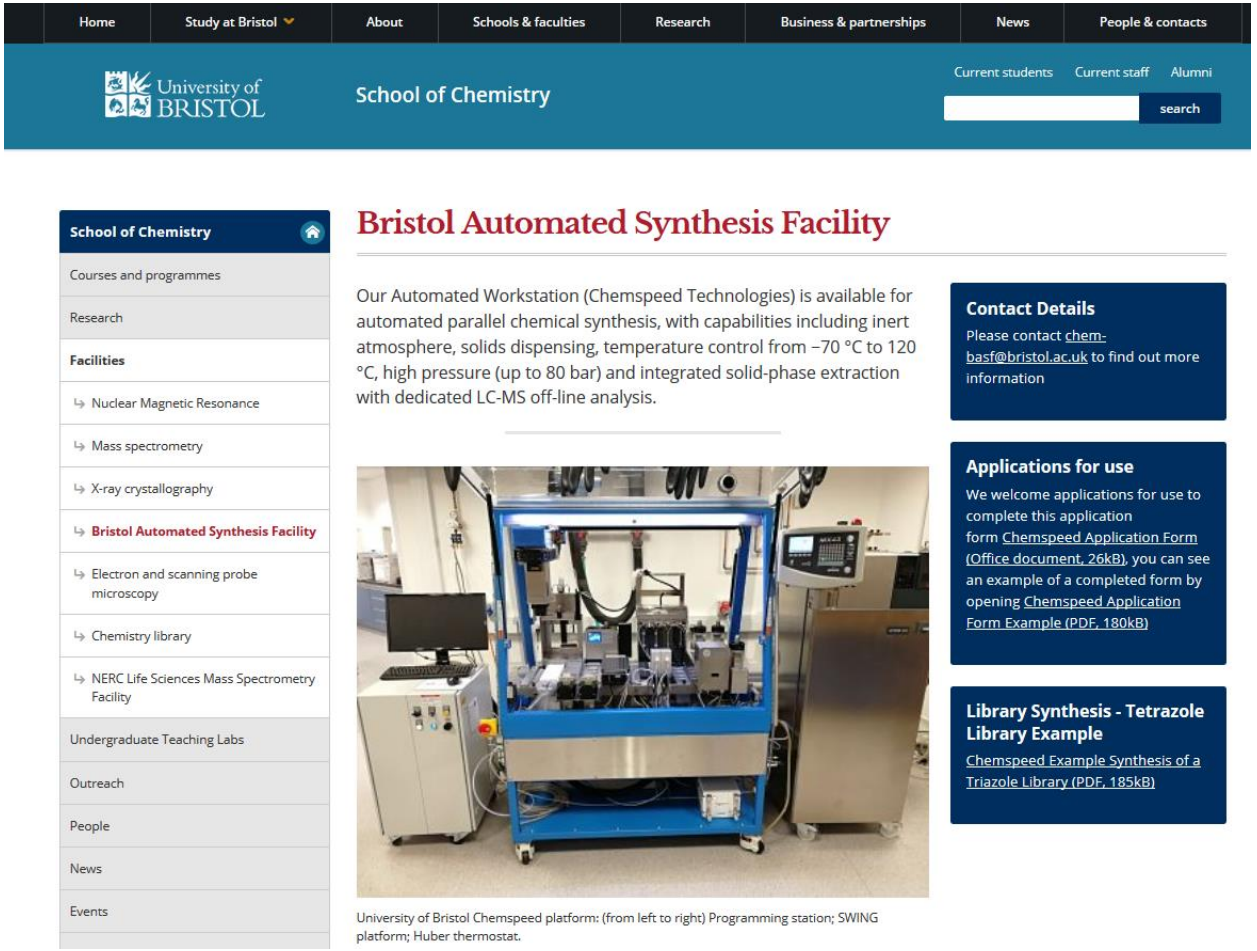
DIAL-A-MOLECULE
An EPSRC Grand Challenge Network

(CLM)

CatSci
Innovation in Catalysis

AstraZeneca (JJM)

University of York pump-priming fund, JGI Seedcorn Funding



bris.ac.uk/chemistry/facilities/basf

Physical Insights into Mechanistic Processes in Organometallic Chemistry

Faraday Discussion

2– 4 September 2019
York, UK

Upcoming deadlines

Early bird registration – 15 July 2019

Standard registration – 05 August 2019

http://rsc.li/organometallic-fd2019