Scottish Computational Chemistry Symposium 2023

20 June 2023
University of Glasgow

http://www.chem.gla.ac.uk/compscot2023/

University of Glasgow | School of Chemistry

ARCHIE-WeSt

ROYAL SOCIETY OF CHEMISTRY | THEORETICAL CHEMISTRY GROUP

SCM Software for Chemistry & Materials

ROYAL SOCIETY OF CHEMISTRY Glasgow & West of Scotland Local Section

Programme

Tuesday 20 June 2023
University of Glasgow, School of Chemistry, Joseph Black Building, Glasgow G12 8QQ

from 9:30	Registration Tea, coffee & pastries	Conference Rm (A4-41a)
	Session I (Chair: Dr Hans Senn)	Main LT (B4-19)
10:30	Dr Hans Senn: *Welcome and opening remarks*	
10:35	**K1** Prof Carles Bo (ICIQ, Tarragona, Spain) *Taming computational chemistry data: ioChem-BD and beyond*	
11:25	**C1** Dr Suryoday Prodhan (University of Liverpool) *From monomer sequence to charge mobility in semiconductor polymers via model reduction*	
11:45	**C2** Hamish W. A. Swanson (University of Strathclyde) *Computational insights of sequence dependence in chemically fueled self-assemblies*	
12:05	**C3** Jake A. Thompson (University of Glasgow) *Calculation of redox potentials in aqueous transition-metal-substituted polyoxotungstates*	
12:25	**C4** Dr Karina Kubiak-Ossowska (ARCHIE-WeSt/University of Strathclyde) *ARCHIE-WeSt – High-performance computing*	
12:35	Lunch	Conference Rm (A4-41a)
	Session II (Chair: Dr Laia Vilà-Nadal)	Carnegie LT (C3-05)
13:40	**C5** Dr Ephrath Solel (University of Edinburgh) *Dispersion vs. steric hindrance: reinvestigating classic steric factors*	
14:00	**C6** Dr M. Arif Sajja (Heriot-Watt University) *A Comparison of non-covalent interactions in σ-alkane complexes of Rh in the solid state*	
14:20	**C7** Luca Craciunescu (Heriot-Watt University) *Theoretical investigation of bimolecular collisions: NO + CO_2*	
14:40	**C8** Dr Maria Jose Aliaga (SCM, Amsterdam, Netherlands) *SCM, AMS2023 & opportunities for collaborations*	
15:00	Tea & coffee	Conference Rm (A4-41a)
	Session III (Chair: Prof Stuart Macgregor)	Main LT (B4-19)
15:30	**C9** Alister S Goodfellow (University of St. Andrews) *Exploring ligand effects in Mn-catalysis with DFT and machine learning*	
15:50	**C10** J. Jasmin Güven (University of Edinburgh) *Exploiting the path-independence of free energies to study potential inhibitors for β-lactamases*	
16:10	**C11** Rohan Gorantla (University of Edinburgh) *From proteins to ligands: decoding deep learning methods for binding affinity prediction*	
	Poster Session	Conference Rm (A4-41a)
16:30	**P1** – **P26** Posters, drinks & cheese	
17:50	Announcement of poster prizes, closing remarks	

Abstracts of Talks

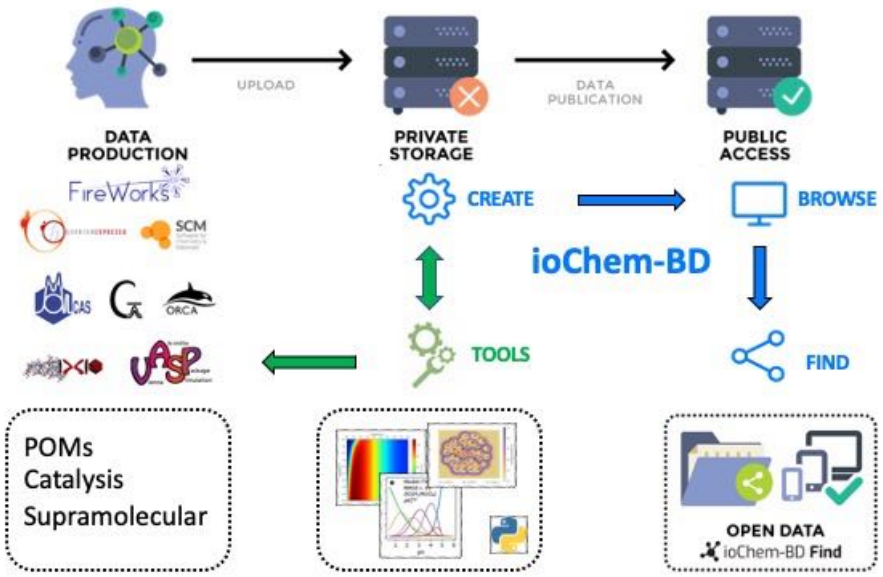
K1 Taming computational chemistry data: ioChem-BD and beyond

Carles Bo

Institute of Chemical Research of Catalonia (ICIQ), 43007 Tarragona, Spain. E-mail: cbo@iciq.cat.

This talk will introduce the basics and illustrate the main features of *ioChem-BD*.[1] This web platform has a modular architecture and a user-friendly interface designed to keep input and output files from computational chemistry studies safe, to extract meaningful data from the raw results, and to publish the results from computational projects in open-data form. Since we launched the platform in 2015, it has evolved significantly.[2] Since 2019, the code for some modules is distributed as open-source and it is being used by researchers worldwide.

ioChem-BD enables complex tasks automation through data workflows (see Figure 1). Some examples of new tools developed by our group to tackle different aspects of reactivity and catalysis will be showcased. *POMSimulator*[3] generates reaction networks automatically from a given molecular dataset and solves the corresponding simultaneous multi-equilibria equations to obtain speciation-phase diagrams; *amk-tools*[4] provides a framework to parse large chemical reaction networks and visualize them interactively, facilitating the chemical interpretation of the obtained results; *gTOFfee*[5] is a generalization of the energy span model that can be applied to catalytic cycles of arbitrary complexity, and *OntoRXN*[6] is an ontology that allows to express reaction networks in a semantic manner. All these pieces of software are aimed at analyzing and organizing computational catalysis results, and being naturally interconnected with *ioChem-BD*, are good add-ons to the Catalysis Informatics toolbox.



[1] https://www.iochem-bd.org.
[2] Álvarez-Moreno, M., de Graaf, C., López, N., Maseras, F., Poblet, J. M., Bo, C., *J. Chem. Inf. Model*. **2015**, *55*, 95–103. Bo, C., Maseras, F., López, N., *Nat. Catal.* **2018**, *1*, 809–810.
[3] Petrus, E., Segado, M., Bo, C. *Chem. Sci.* **2020**, *11*, 8448–8456. Petrus, E., Bo, C. *J. Phys. Chem. A* **2021**, *125*, 23, 5212–5219. Petrus, E., Segado, M., Bo, C. Inorg. Chem. **2022**, *61*, 35, 13708–13718.
[4] Garay-Ruiz, D., Álvarez-Moreno, M., Bo, C., Martínez-Núñez, E., *ACS Phys. Chem. Au* **2022**, *2*, 225–236.
[5] Garay-Ruiz, D., Bo, C., *ACS Catal.* **2020**, *10*, 12627−12635.
[6] Garay-Ruiz, D., Bo, C., *J. Cheminf.* **2022**, *14*, 29.

C1 From monomer sequence to charge mobility in semiconductor polymers via model reduction

Suryoday Prodhan, Rex Manurung, Alessandro Troisi

Department of Chemistry, University of Liverpool, Liverpool L69 3BX, UK. E-mail: suryoday.prodhan@liverpool.ac.uk.

In search of high charge-carrier mobility polymeric systems, it is essential to establish the connection between chemical structures and transport properties of conjugated polymers. In this talk, we will present an efficient model reduction scheme for polymer semiconductors that can be utilized to compute intra-chain charge-carrier mobility from the monomer sequence. The reduced model can be used in conjunction with any quantum dynamics approach, but at the present context it is explored assuming that transport takes place through incoherent hopping events between states of different degrees of delocalization. The procedure is developed by considering 28 realistic polymers, for which a quantitative correlation is established between charge localization characteristics and charge mobility. The polymer dataset helps in establishing plausible ranges for all the microscopic parameters of the model and therefore it has been used to determine the maximum plausible improvement in mobility. The reduced model is also used to provide some insight on the observation that the highest mobility polymers do not have very broad valence bands: there is indeed a range of the inter-monomer coupling for which this parameter has little effect on the mobility.

S. Prodhan, R. Manurung, A. Troisi, *Adv. Funct. Mater*. **2023**, DOI: 10.1002/adfm.202303234.

C2 Computational insights of sequence dependence in chemically fueled self-assemblies

Christine M. E. Kriebisch[b], Brigitte A. K. Kriebisch[b], Hamish W. A. Swanson[a], Alexander van Teijlingen[a], Zoe MacPherson[a], Alexander M. Bergmann[b], Tell Tuttle[a], Job Boekhoven[b]

[a]Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow G1 1XL, UK. E-mail: hamish.swanson@strath.ac.uk. [b]School of Natural Sciences, Department of Chemistry, Technical University of Munich, 85748 Garching, Germany.

The equilibrium assembly of short peptides and their derivatives is contingent on peptide sequence in a manner that has significant implications for resultant structure of nanoassemblies.[1] For both modified and unmodified di- and tripeptides a wealth of data has been generated such that it has been possible to produce design rules and rationales to characterize this relationship.[2 3] In the context of non-equilibrium chemically fueled self-assembly this relationship is more nuanced; being a process with both a reaction coordinate and an *assembly coordinate*, along which morphological transitions can occur. Sequence dependencies can dictate the occupancy of assembled states and through their alteration it is possible to impact directly the exploration of the assembly coordinate and cause kinetic trapping of assembled states.[4]

With a view to the development of adaptive self-assembly for nanotechnology applications, it is necessary to generate design rules such that kinetic trapping can be controlled, and short peptide sequences can be selected *de novo* to yield systems with desired properties. Currently these design rules are emerging,[5] though to deepen our understanding of the critical features of sequencing a theoretical rationale is required. Thus, we have been motivated to adapt a coarse-grained constant pH molecular dynamics model, with stochastic charge neutralisation,[6] to enable the simulation of non-equilibrium self-assemblies in a manner which is sensitive to system reactivity. Consequently, it has been possible to rationalize the role of amino acids at given sequence positions in dictating the evolution of assembly morphology and thus strengthen our understanding of emerging design rules.

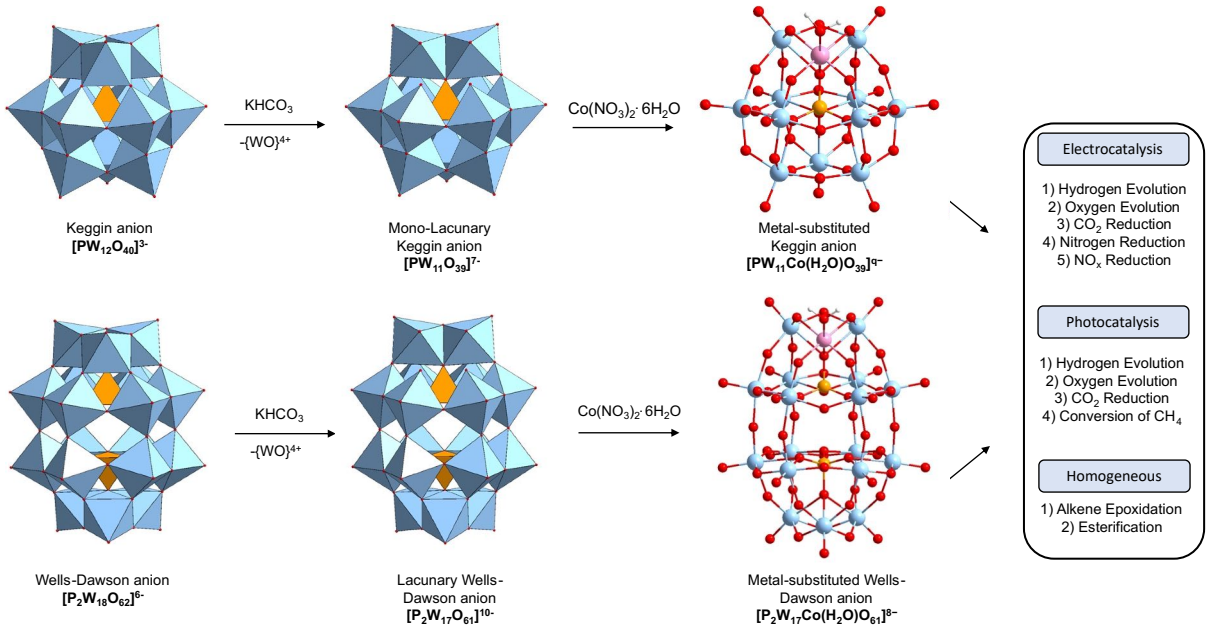
[1] P. W. J. M. Frederix, G. G. Scott, Y. M. Abul-Haija, D. Kalafatovic, C. G. Pappas, N. Javid, N. T. Hunt, R. V. Ulijn, T. Tuttle, *Nat. Chem.* **2015**, 7, 30–37.
[2] A. Lampel, R. V. Ulijn, T. Tuttle, *Chem. Soc. Rev.* **2018**, 47, 3737–3758.
[3] S. Fleming, R. V. Ulijn, *Chem. Soc. Rev.* **2014**, 43, 8150–8177.
[4] M. Tena-Solsona, J. Boekhoven, *Isr. J. Chem.* **2019**, 59, 898–905.
[5] B. A. K. Kriebisch, C. M. E. Kriebisch, A. M. Bergmann, C. Wanzke, M. Tena-Solsona, J. Boekhoven, *Chemsystemschem* **2022**.
[6] A. van Teijlingen, H. W. A. Swanson, K. H. A. Lau, T. Tuttle, *J. Phys. Chem. Lett.* **2022**, 13, 4046–4051.

C3 Calculation of redox potentials in aqueous transition-metal-substituted polyoxotungstates

Jake A. Thompson[a], Rebeca González-Cabaleiro[b], Laia Vilà-Nadal[a]

[a]School of Chemistry, University of Glasgow, Glasgow G12 8QQ, UK. [b]Faculty of Civil Engineering and Geosciences, Delft University of Technology, 2628 CD Delft, Netherlands.

Polyoxometalates (POMs) have received significant interest owing to their diversity of topologies, structural flexibility, and functionality at the nanoscale.[1-3] These compounds exhibit excellent redox stability and are integral precursors for numerous electrochemical reactions. Density Functional Theory (DFT) calculations have been employed to systematically study the accuracy of various exchange–correlation functionals in describing the redox properties of mono-substituted polyoxotungstates. We propose a new model by incorporating K^+ counter-ions and observe their effect on structural parameters and electrochemical properties. The influence of the self-interaction error (SIE) was significantly reduced by incorporating counter-ions to our approximations. However, overestimations across in $K_x[PW_{11}M(H_2O)O_{39}]^{q-x}$ salts were prominent attributed to the over-stabilization of the ion-pairs. Functionals containing large contributions to Hartree–Fock produced large uncertainties in redox potentials attributed to exaggerated proximity of the ion-pair. Our results emphasize that understanding the electrolyte environment is essential to obtaining reasonable agreement between theoretical and experimental results.[4]



Scheme. General synthetic strategy for obtaining metal–substituted Keggin, $[PW_{11}M(H_2O)O_{39}]^{q-}$, and Wells–Dawson, $[P_2W_{17}M(H_2O)O_{61}]^{q-}$, derivatives.

[1] A. V. Anyushin, A. Kondinski, T. N. Parac-Vogt, *Chem. Soc. Rev.* **2020**, *49*, 382–432.
[2] N. I. Gumerova, A. Rompel, *Chem. Soc. Rev.* **2020**, *49*, 7568–7601.
[3] N. v. Izarova, M. T. Pope, U. Kortz, *Angew. Chem., Int. Ed. Engl*. **2012**, *51*, 9492–9510.
[4] J. Thompson, R. Gonzalez-Cabaleiro, L. Vilà-Nadal, *ChemRxiv* **2022**, DOI: 10.26434/chemrxiv-2022-mr2mr.

C4 ARCHIE-WeSt – High-performance computing

Karina Kubiak-Ossowska

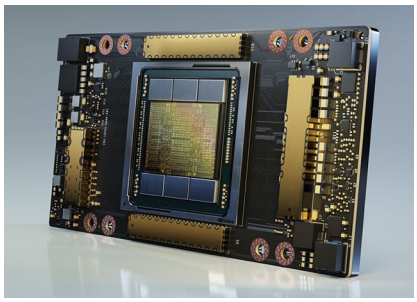
ARCHIE-WeSt, Department of Physics, University of Strathclyde, Glasgow G4 0NG. E-mail: karina.kubiak@strath.ac.uk or support@archie-west.ac.uk.

ARCHIE-WeSt
Academic and Research Computer Hosting Industry and Enterprise in the West of Scotland

ARCHIE-WeSt is a High Performance Computing (HPC) centre designed to support research computing in the West of Scotland. The aim of ARCHIE-WeSt is to provide computational resources for academic and industrial users*. ARCHIE-WeSt is a not-for-profit shared facility funded, hosted and managed by the University of Strathclyde.

The computational resources provided by ARCHIE-WeSt have been successfully used in many computational chemistry related projects which require to run calculations using GAMES, GAUSSIAN, ORCA, VASP, CP2K and many other software packages. The usage of ARCHIE-WeSt increases research output because researchers can run several jobs simultaneously on one or multiple cores.

Together with the recent addition of 24 NVidia A100 GPU cards, ARCHIE-WeSt offers computational power and capacity which is much greater than that available on a standard workstation or desktop computer. Therefore, computational problems at greater scale can be undertaken as well as significantly enhancing throughput.



NVidia A100 GPU card

To ensure the facility meets user needs the ARCHIE-WeSt team provides initial training and responsive personalized support: most support issues are addressed within few working hours. The application procedure is lightweight with access often granted within a few days, thereby minimising access barriers.

Our mission is to support your research. If you want to join our satisfied user community please visit www.archie-west.ac.uk or e-mail to support@archie-west.ac.uk.

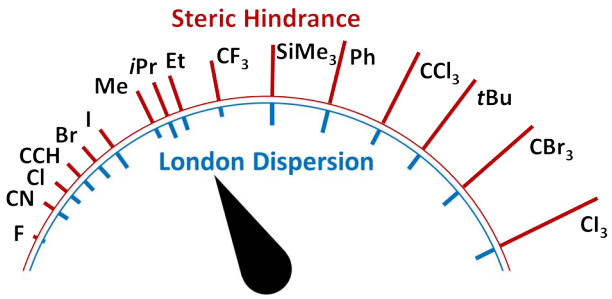
* Access charges apply for both academic and industry users

C5 Dispersion vs. steric hindrance: reinvestigating classic steric factors

Ephrath Solel

University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh EH9 3FJ, UK.

London dispersion (LD) interactions, the attractive part of the van-der-Waals interaction[1,2] hold somewhat of a unique position in the chemical world. Although their role in influencing macroscopic phenomena (such as the higher boiling points of larger alkanes) is well recognized, they are usually overlooked when discussing molecular phenomena. Substituents in reactions are generally considered as a source of "steric hindrance" and not as "steric attractors", better termed dispersion energy donors (DEDs). As such, their influence on reaction outcomes was quantified and presented by classic steric factors such as the A-value. We have shown, using computational quantum mechanical tools, that these well recognized steric factors have also an attractive LD component that balance part of the steric repulsion. By focusing on the LD component, we can explain various non-intuitive trends between substituents, such as the inconsistency between the size of the halogens and their A-values.[3] In addition, a systematic analysis of both the steric and dispersion interactions of the same molecules allows us to quantify the relative weights of the two effects and form a new DED scale.[4] Such corrected steric and LD factors could later be applied to explore the role of LD interactions also in other reactions. Our computations show that LD interactions have a significant influence on the overall relative stabilities and energetics in cyclohexane chair conformers, and also in related concerted reactions, and must not be ignored in reaction design.



[1] R. Eisenschitz, F. London, *Z. Phys.* **1930**, *60*, 491–527.
[2] F. London, *Trans. Faraday Soc.* **1937**, *33*, 8–26.
[3] E. Solel, M. Ruth, P. R. Schreiner, *J. Org. Chem.* **2021**, *86*, 7701–7713.
[4] E. Solel, M. Ruth, P. R. Schreiner, *J. Am. Chem. Soc.* **2021**, *143*, 20837–20848.

C6 A Comparison of Non-Covalent Interactions in σ-Alkane Complexes of Rh in the Solid State

M. Arif Sajjad, Stuart A. Macgregor

Institute of Chemical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS

σ-alkane complexes obtained by solid-state molecular organometallic chemistry (SMOM) are important models of key intermediates in C–H activation reactions. These σ-complexes can also act as heterogeneous catalysts for alkene isomerisation under solid/gas or solid/liquid conditions, processes where a parent alkane is replaced by another alkene or alkane.[1]

It has been suggested that the stability and reactivity of these σ-complexes are linked with the non-covalent interactions present between the microenvironment of anions and the cationic Rh centre.[2] For instance, both $[Rh(Cy_2P(CH_2)_2PCy_2)(NBA)][BAr^F_4]$, **[1-NBA][BAr$^F_4$]**, [NBA = norbornane, BAr^F_4 = $[B(3,5\text{-}(CF_3)_2C_6H_3)_4]$ and the analogue, $[Rh(Cy_2P(CH_2)_2PCy_2)(Propane)][BAr^F_4]$, **[1-Propane][BAr$^F_4$]**, feature an octahedral array of anions around a central Rh cation (see Figure 1), but differences in their properties were attributed to the greater contribution from C–H···F–C non-covalent interactions in the former.[2,3] Thus, to design robust, stable and reactive σ-alkane complexes, proper understanding of intermolecular interactions between the cation and its surrounding anions is crucial.

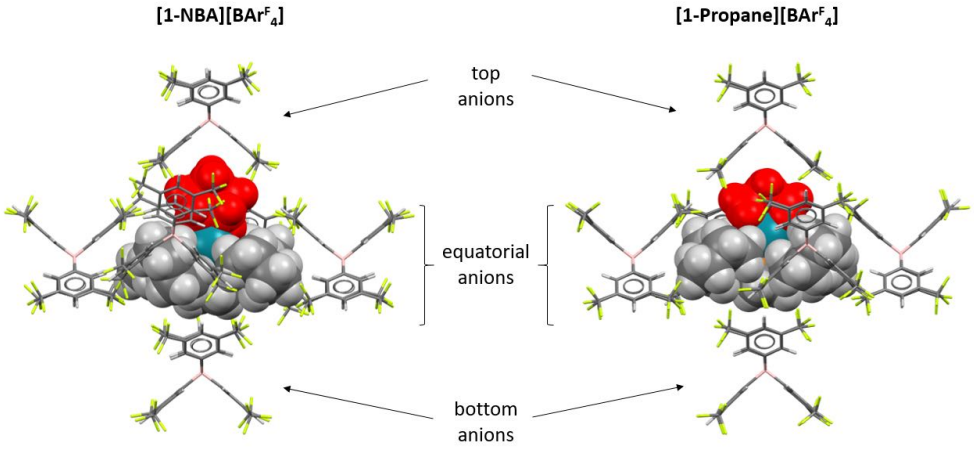


Figure 1. Octahedral arrangement of $[BAr^F_4]^-$ anions around the $[Rh(Cy_2P(CH_2)_2PCy_2)(\eta^2\eta^2\text{-alkane})]^+$ cations in **[1-NBA][1-BAr$^F_4$]** and **[1-Propane][BAr$^F_4$].** σ-alkane ligands are highlighted in red.

In this work, we employ computational modelling to identify the nature and strength of intermolecular non-covalent interactions in **[1-NBA][1-BAr$^F_4$]** and **[1-Propane][BAr$^F_4$]**.[4,5] The solid-state geometries have been optimised by periodic DFT calculations by relaxing only H and F atoms. All the cation-anion ($[1\text{-NBA}]^+$/$[1\text{-Propane}]^+$ – $[BAr^F_4]^-$) pairs with respect to their arrangement in the octahedron have been considered i.e., top, bottom and equatorial anions. Bonding analyses of ion-pairs are based on quantum theory of atoms in molecules (QTAIM) and independent gradient model based on Hirshfeld partition (IGMH) approaches.

[1] A. S. Weller, F. M. Chadwick, A. I. McKay, in *Adv. Organomet. Chem.*, Ed. P. J. Pérez, Academic Press, 2016, vol. 66, pp. 223–276.
[2] A. J. Bukvic, A. L. Burnage, G. J. Tizzard, A. J. Martínez-Martínez, A. I. McKay, N. H. Rees, B. E. Tegner, T. Krämer, H. Fish, M. R. Warren, S. J. Coles, S. A. Macgregor, A. S. Weller, *J. Am. Chem. Soc.* **2021**, *143*, 5106-5120.
[3] L. R. Doyle, E. A. Thompson, A. L. Burnage, A. C. Whitwood, H. T. Jenkins, S. A. Macgregor, A. S. Weller, *Dalton Trans.* **2022**, *51*, 3661-3665.
[4] A. G. Algarra, A. L. Burnage, M. Iannuzzi, T. Krämer, S. A. Macgregor, R. E. M. Pirie, B. Tegner, A. S. Weller, in *21st Century Challenges in Chemical Crystallography II: Structural Correlations, Data Interpretation*, Eds. D. M. P. Mingos, P. R. Raithby, Springer, Cham, 2020, pp. 183–228.
[5] M. A. Sajjad, S. A. Macgregor, A. S. Weller, *Faraday Discuss.*, **2023**, DOI: 10.1039/D3FD00009E.

C7 Theoretical investigation of bimolecular collisions: NO + CO_2

Luca Craciunescu, Martin J. Paterson

Institute of Chemical Sciences, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK

NO both in its $X^2\Pi$ ground and $A^2\Sigma^+$ first excited states is one of the most important radicals in atmospheric and combustion processes. Its collision and subsequent reactive or non-reactive quenching with a multitude of different molecules such as H_2, O_2, N_2 and H_2O has therefore been the subject of numerous experimental[1] and theoretical[2] studies through the years.

This work focuses on a comprehensive theoretical investigation of the NO $A^2\Sigma^+$ and ground state CO_2 system, as CO_2 is a relevant quencher because of its atmospheric significance. These technically challenging van-der-Waals (vdW) excited state potential energy surfaces (PESs) are explored using mainly coupled cluster methods and very diffuse basis sets, as the unpaired electron of the NO $A^2\Sigma^+$ sits in a Rydberg-type orbital. This cannot be described sufficiently by typically used augmented basis sets. The basis set superposition error is taken into account via counterpoise correction. By carrying out scans over relevant intermolecular coordinates, vdW-minima are identified and characterised. Those surfaces show deep vdW wells to up to 800 cm^{-1} and strong rotational anisotropy, which infers significant rotational energy transfer upon collision.

Finally, preliminary multi-state CASPT2 results for the interacting region are shown with the goal to elucidate electronic quenching mechanisms for the system, as done previously by Soulié and Paterson.[3]

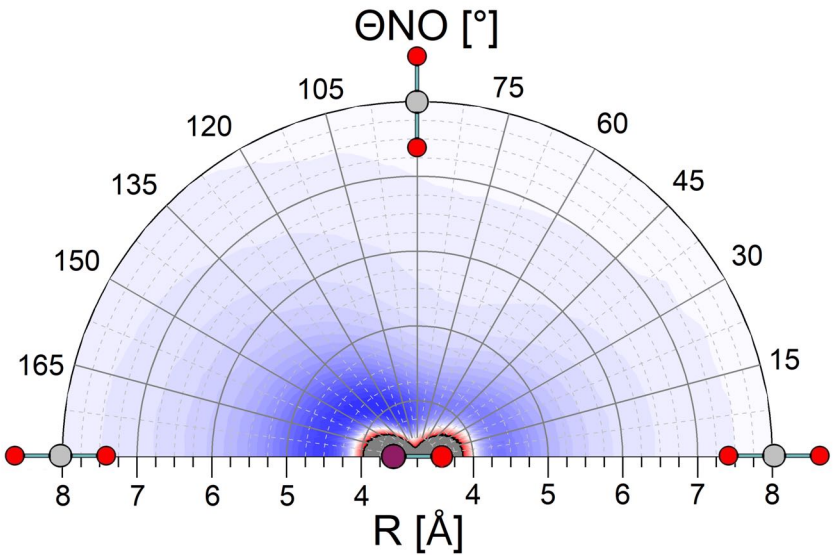


Figure 1: Exemplary calculated vdW-PES for the NO $A^2\Sigma^+$ and CO2 system.

[1] T. B. Settersten, B. D. Patterson, J. A. Gray, *J. Chem. Phys.* **2006**, *124*, 234308.
[2] C. Soulié, M. J. Paterson, *Phys. Chem. Chem. Phys.* **2022**, *24*, 7983-7993.
[3] C. Soulié, M. J. Paterson, *J. Chem. Phys* **2022**, *157*, 164304.

C8 SCM, AMS2023 and opportunities for collaborations

Maria Jose Aliaga, Sergio López, Fedor Goumans, Stan van Gisbergen

Software for Chemistry & Materials BV (SCM), De Boelelaan 1083, 1081 HV Amsterdam, Netherlands. E-mail: aliaga@scm.com

Originating from an academic density functional development group in the 1970s, SCM is a well-established and steadily growing scientific software company, both in terms of (academic) developers and functionality. Academic connections have always been fostered by SCM, also in the frame of public-funded collaborations, and the company has ample experience applying to and participating in different programs which have provided valuable funding for collaborations with academic (and industrial) partners. In this talk potential funding channels will be presented as well as what our friendly policies are for academic developers who are looking to contribute to our software.

The central framework in the Amsterdam Modeling Suite (AMS) enables the exploration of potential energy surfaces (PESs), mechanical, and electronic properties at several levels of theory. The unified AMS driver supports advanced PES explorations, molecular dynamics (MD) and Grand Canonical Monte Carlo (GCMC). Machine learned graph neural network potentials such as NEQUIP[1] and the universal model M3Gnet[2,3] can immediately be used for simulations of processes. The ParAMS module furthermore provides a comprehensive framework to build training data and optimize machine learned potentials (MLP), as well as ReaxFF and DFTB parameters. With different levels of electronic structure methods available in AMS, we are exploring ML methods to predict properties more efficiently for molecular materials. We welcome modelers and prospective developers from this computational chemistry community to discuss ideas and (funding) opportunities with us.

[1] J. Vandermause *et al*., *npj Comput. Mater.* **2020**, *6*, 20.
[2] C. Chen, S.P. A Ong, *Nat. Comput. Sci.* **2022**, *2*, 718–728.
[3] https://www.scm.com/m3gnet

C9 Exploring ligand effects in Mn-catalysis with DFT and machine learning

Alister S Goodfellow, Sarah Jane C. Sutcliffe, John B. O. Mitchell, Michael Bühl

School of Chemistry, University of St Andrews, St Andrews, Fife, KY16 9ST, UK.

Catalytic reactions involving 3d metals are widespread in the literature and present a move away from the traditional use of unsustainable, heavier metals. Manganese is active in the field of homogeneous catalysis for transfer hydrogenation reactions involving a range of substrates, including a variety of aromatic ketones.[1] Further experimental work in this field has shown that these reactions can be performed with enantioselectivity using a chiral ligand system.[2]

In this work, we explore ligand effects across a variety of different systems, from a detailed DFT perspective and using Machine Learning algorithms such as *k*-nearest neighbours. With DFT we rationalise and predict enantiocontrol through either stabilising a favoured or destabilising a disfavoured diastereomeric transition state and examine the effect of electronics on reactivity. On a larger scale, by using semi-empirical GFN2-xTB derived descriptors, we can predict DFT barriers to a mean absolute error of <1 kcal/mol, allowing comparison of relative activity across a range of different ligand systems at a far lower cost than with DFT.

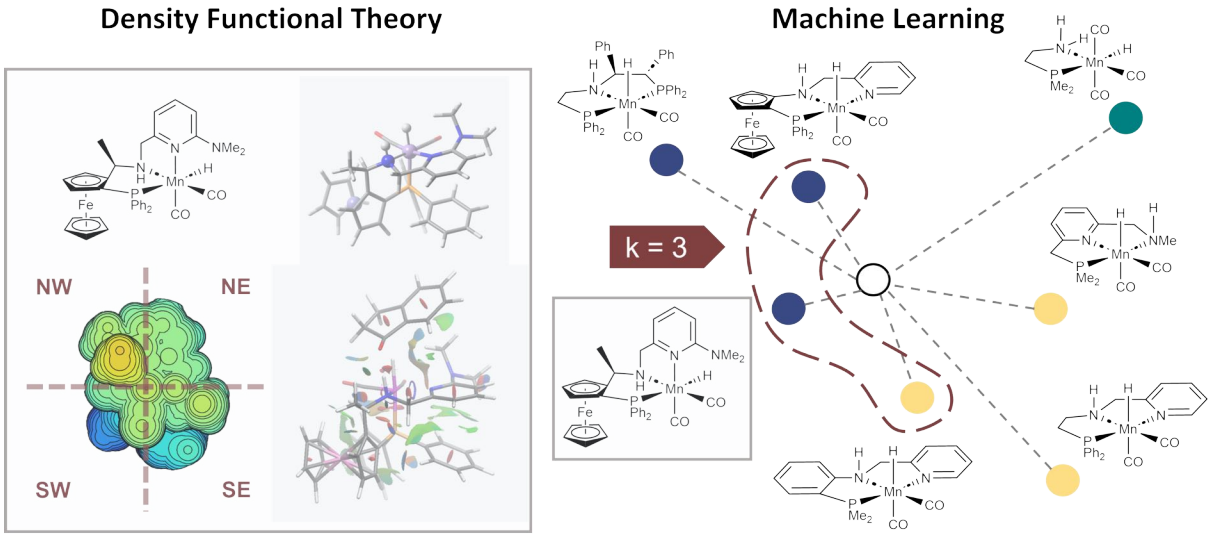


Figure 1. DFT analysis (left) and a schematic of a *k*-nearest neighbour ML algorithm (right).

[1] M. B. Widegren, G. J. Harkness, A. M. Slawin, D. B. Cordes, M. L. Clarke, *Angew. Chem. Int. Ed.* **2017**, *56*, 5825–5828.
[2] C. L. Oates, A. S. Goodfellow, M. Bühl, M. L. Clarke, *Angew. Chem. Int. Ed.* **2023**, *62*, e202212479.

C10 Exploiting the path-independence of free energies to study potential inhibitors for β-lactamases

J. Jasmin Güven, Antonia S. J. S. Mey

EaStCHEM School of Chemistry, University of Edinburgh, Edinburgh, UK.

Alchemical free energy (AFE) calculations offer a computationally manageable way of obtaining free energies of binding between a protein and a small drug-like molecule.[1] AFE calculations exploit the path-independence of free energies to obtain binding free energy estimates with accuracies of around 1 kcal/mol.[1] However, ensuring reliability for pharmaceutically-relevant protein targets and ligand series remains a challenge. An example of this is metallo-β-lactamase (MBLs) enzymes, where active-site zinc ions directly interact with the ligands and thus influence their binding affinities.[2] While bacteria have evolved many resistance mechanisms, the production of β-lactamases is the most common mechanism against antibiotics, such as penicillin.[3] Thus, building reliable and easy-to-use AFE methods is essential for using MBLs as drug targets.

Here, we present a set of baseline results of existing AFE methods calculated for a series of known inhibitors across serine- (KPC-2) and metallo-β-lactamases (VIM-2).[2] Conventional methods can be used for KPC-2 to achieve accuracies of around 0.8 – 1.2 kcal/mol, depending on the forcefield, simulation engine and perturbation network used. For VIM-2, we use methods to build a non-bonded zinc model based on the work by Li and Merz.[4] For KPC-2, we obtain an MUE of $0.80^{0.81}_{0.78}$ kcal/mol and a Spearman's rank correlation coefficient of $0.60^{0.61}_{0.59}$. Preliminary results for both KPC-2 and VIM-2 suggest that improvements to the methodology are needed in multiple areas. For AFE calculations to become computationally competitive for metalloproteins, streamlining the setup and calculations is essential. Overall, we highlight the need for new methodologies and optimised AFE methods to make them more practical for metalloproteins.

[1] A. S. J. S. Mey *et al. LiveCoMS* **2020**, *2*.
[2] O. A. Pemberton *et al., J. Med. Chem.* **2019**, *62*.
[3] T. Palzkill, *Ann. N. Y. Acad. Sci.* **2013**, *65.*
[4] P. Li, K. M. J. Merz, *J. Chem. Inf. Model.* **2016**, *56*.

C11 From proteins to ligands: decoding deep learning methods for binding affinity prediction

Rohan Gorantla[a, c], Andrea Weiße[b], Antonia S. J. S. Mey[c]

[a]School of Informatics, University of Edinburgh,Edinburgh EH8 9AB. [b]School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FF. [c]EaStCHEM School of Chemistry, University of Edinburgh, Edinburgh EH9 3FJ.

Accurate in silico predictions of protein-ligand binding affinity can significantly accelerate the early stages of drug discovery. Deep learning-based methods have shown promise recently, but their robustness for the virtual screening of large compound libraries on various targets needs improvement. [1] Understanding what these models learn from input protein and ligand data is essential to address this problem. We systematically investigated a sequence-based deep learning framework to assess the impact of protein and ligand encodings on two commonly used kinase datasets, Davis and KIBA. [2,3] The role of proteins is studied in two ways, using convolutional neural network-based encodings obtained from sequences and graph neural network-based encodings enriched with structural information. We study the convolutional neural network-based encodings using the protein representations given by a protein language model, ESM[4] and a handcrafted sequence with binding site residues from KLIFS[5]. The role of structural information is analysed through contact maps obtained from Alphafold2[6], ESM[4], and Pconsc4[7]. By introducing perturbations to the ligand graph representation of the SMILES string, we assess the effect of ligand encodings on the given graph neural network.

We find that protein encodings using contact maps do not significantly impact the binding affinity predictions. The sequence-based language model encodings perform better than handcrafted sequence encoding but give comparable results to the contact map-based encodings. In contrast, we find that testing variations of random ligand encodings led to a significant drop in performance on both datasets. As a result, ligand graph perturbation experiments show that the deep learning model relies heavily on ligand encodings for accurately predicting the binding affinity. This poses a challenge for generalisability and looking at off-target effects.

[1] M. Jiang *et al.*, *RSC Adv*. **2020**, *10*, 20701–20712.
[2] J. Tang *et al.*, *J. Chem. Inf. Model*. **2014**, *54*, 735–743.
[3] M. I. Davis *et al.*, *Nat. Biotechnol*. **2011**, *29,* 1046–1051.
[4] R. Rao *et al.*, *ICLR* **2020**.
[5] G. K. Kanev *et al.*, *Nucleic Acids Res.* **2021**, *49*, D562–D569
[6] J. Jumper *et al.*, *Nature* **2021**, 583–589.
[7] M. Michel, M. H. David, E. Arne, *Bioinformatics* **2019**, 2677–2679.

Abstracts of Posters

P1 Augmentation of FTIR spectral datasets using WGANs for pancreatic cancer liquid biopsy

Rose G. McHardy[a,b], Georgios Antoniou[b], Justin J. A. Conn[b], Matthew J. Baker[b,c], David S. Palmer[a,b]

[a]Department of Pure and Applied Chemistry, Thomas Graham Building, 295 Cathedral Street, University of Strathclyde, G1 1XL, UK. [b]Dxcover Ltd, Royal College Building, 204 George Street, Glasgow, G1 1XW, UK. [c]School of Medicine, Faculty of Clinical and Biomedical Sciences, University of Central Lancashire, Preston, PR1 2HE, UK

Over recent years, deep learning (DL) has become more widely used within the field of cancer diagnostics.[1] However, DL often requires large training datasets to prevent overfitting, which can be difficult and expensive to acquire. Data augmentation is a method that can be used to generate new data points to train DL models.[2]

In this study, we use attenuated total reflectance Fourier transform infrared (ATR-FTIR) spectra of patient serum samples and compare non-generative data augmentation methods to Wasserstein generative adversarial networks (WGANs)[3] in their ability to improve the performance of a convolutional neural network (CNN) to differentiate between pancreatic cancer and non-cancer samples in a total cohort of 625 patients.

The results show that WGAN augmented spectra improve CNN performance more than non-generative augmented spectra. When compared with a model that utilised no augmented spectra, adding WGAN augmented spectra to a CNN with the same architecture and same parameters, increased the area under the receiver operating characteristic curve (AUC) from 0.661 to 0.757, presenting a 15% increase in diagnostic performance. In a separate test on a colorectal cancer dataset, data augmentation using a WGAN led to an increase in AUC from 0.905 to 0.955. This demonstrates the impact data augmentation can have on DL performance for cancer diagnosis when the amount of real data available for model training is limited.

[1] G. Antoniou, J. J. A. Conn, B. R. Smith, P. M. Brennan, M. J. Baker, and D. S. Palmer, *Analyst,* **2023**, *148*, 1770-1776.
[2] S. I. Nikolenko, in *Synthetic Data for Deep Learning*, Springer Chem, 2021.
[3] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, *Advanced Neural Information Processing Systems,* **2017**, 5767-5777

P2 Halogen bonding with a flying molecule: the halogenabenzene bird

Tanja van Mourik

[a]*School of Chemistry, University of St Andrews, Scotland, UK.*

Halogen-substituted benzenes have received ample attention in the literature. Conversely, benzene structures where a carbon is replaced by a halogen have received much less attention so far. Such structures, labelled halogenabenzenes, were first introduced by Glukhovtsev in 1991.[1] Based on semiempirical calculations, he proposed a planar 8π-electron system. However with 8 π-electrons, this system is antiaromatic. Based on higher-level DFT and MP2 calculations, Rawashdeh et al. showed in 2017 that the planar iodabenzene structure is a transition state; the minima it is connected to are both an identical C_s-symmetric non-planar structure.[2] Rawashdeh et al. dubbed this structure "bird" because of the similarity with a flying bird (see Figure 1). The bird structure is unusual in the sense that the iodine is bonded to two carbon atoms. With the current interest in halogen bonds (X-bonds) in our group, we wondered how this topology would affect the halogen's ability to form X-bonds. To this aim, we investigated the ability of the bird-like halogenabenzene molecule, referred to as X-bird (X = Cl to At), to form halogen-bonded complexes with the nucleophiles H2O and NH3 using double-hybrid density functional theory and the aug-cc-pVTZ/aug-cc-pVTZ-PP basis set. The unusual structure of the X-bird results in two distinct σ-holes, roughly at the extension of the C–X bonds. Based on the behaviour of the interaction energy (which increases for heavier halogens) and van der Waals (vdW) ratio (which decreases for heavier halogens), it is concluded that the X-bird forms proper halogen bonds with H_2O and NH_3.[3]

Figure 1: The bird-like halogenabenzene molecule

[1] M. N. Glukhovtsev, *Russ. J. Org. Chem.* **1991**, *108*, 5299-5358. [2] A. M. Rawashdeh, P. Chakkingal Parambil, T. Zeng, R. Hoffmann, *J. Am. Chem. Soc.* **2017**, *139*, 7124-7129. [3] T. van Mourik, E.L. Cates, *J. Comput. Chem.* **2019**, *40*, 2111-2118.

P3 Electronic structure and reactivity of novel Ru–Zn heterobimetallic complexes

Lia Sotorrios,[a] Fedor M. Miloserdov,[b] Anne-Frédérique Pécharman,[b] John P. Lowey, [b] Stuart A. Macgregor,[a] Mary F. Mahon,[b] Michael K. Whittlesey[b]

[a]*Institute of Chemical Sciences, Heriot-Watt University, Edinburgh EH14 4AS.* [b]*Department of Chemistry, University of Bath, Bath BA2 3QD.*

Transition metal-main group metal heterobimetallic complexes can display cooperative and synergetic effects and have proven exceptionally useful in the activation of small molecules.[1] Nevertheless, the description of these systems in terms of electronic structure and bonding remains challenging. Our approach consists of the use of a wide range of computational techniques to study both the nature and reactivity of heterobimetallic complexes.

Recently, we reported the synthesis and reactivity of novel heterobimetallic complexes containing direct Ru–Zn bonds.[2] One of these systems was the unusual 8-coordinate complex $[Ru(PPh_3)_2(ZnMe)_4H_2]$ (**1**, Figure 2) in which the peripheral ZnMe ligands are the centre of reactivity.

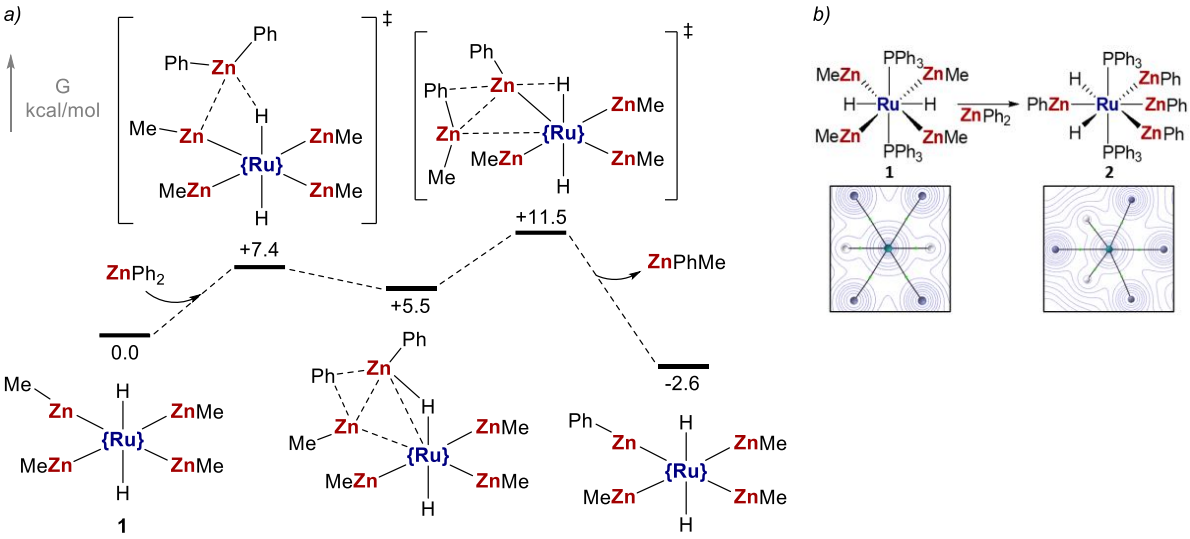


Figure 2. a) Free energy profile (kcal/mol) for ZnMe/ZnPh exchange in **1**. Theory level: (ωB97X-D(toluene)/def2TZVP/ /BP86/SDD(Ru, Zn, P, with pol. on P), 6-31G**. {Ru} = $[Ru(PPh_3)_2]$. b) QTAIM molecular graphs for **1** and **2** (equatorial planes).

The addition of $ZnPh_2$ to **1** resulted in the complete ZnMe/ZnPh exchange to form $[Ru(PPh_3)_2(ZnPh)_4H_2]$, **2**. DFT calculations show the reaction proceeds by the approach of $ZnPh_2$ aided by one of the hydride ligands. The Ph group is transferred into the adjacent zinc atom through a second transition state providing the mono-substituted complex (Figure 1*a*, $\Delta G^{\ddagger}$ = +11.5 kcal/mol; ΔG = –2.6). Subsequent ZnMe/ZnPh exchanges and a final isomerisation lead to complex **2**.

Quantum Theory of Atoms In Molecules (QTAIM) was used to analyse the electron density of **1** and **2**, focusing on Ru–Zn, Zn–H and Zn–Zn interactions (*b*, Figure 1). This suggested complexes **1** and **2** are best described as featuring 8-coordinate Ru centres. This was confirmed by the Extended Transition State (ETS) method combined with the Natural Orbitals for Chemical Valence (NOCV) theory and Non-Covalent Interaction (NCI) plots that indicated that any peripheral Zn···H and Zn···Zn interactions are at best weak.

[1] J. Campos, *Nat. Rev. Chem.*, **2020**, *4*, 696-702. [2] (a) F.M. Miloserdov, A. F. Pécharman, L. Sotorrios, N. A. Rajabi, J. P. Lowe, S. A. Macgregor, M. F. Mahon M. K. Whittlesey, *Inorg, Chem*, **2021**, *60*, 16256-16265. (b) L. Sotorrios, F. M. Miloserdov, A. F. Pécharman, J. P. Lowe, S. A. Macgregor, M. F. Mahon, M. K. Whittlesey, *Angew. Chem. Int. Ed.*, **2022**, *61*, e202117495.

P4 Computational modelling of a Chan–Lam aryl iodination reaction

Ambre Carpentier[a], Stuart A. Macgregor[a], Matthew Andrews[b], Allan J.B. Watson[b]

[a]Institute of Chemical Sciences, Heriot-Watt University, Edinburgh. E-mail: ac2070@hw.ac.uk. [b]EaStCHEM, School of Chemistry, University of St Andrews, St Andrews.

Amination reactions are traditionally performed by nucleophile-electrophile coupling, such as the Ullmann and Hartwig-Buchwald reactions, requiring harsh reaction conditions and extensive ligand development. The more recent nucleophile-nucleophile Chan-Lam reaction offers good to excellent yields under ambient conditions and allows 'ligandless' chemistry, using readily available boronic acid and inexpensive copper catalysts.[1,2] Despite the great potential of such a reaction, it remains poorly understood and few mechanistic studies are available.[3] The Chan-Lam reaction can be extended to the formation of various C–X bonds (where X = C, O, S, halogen).

Here we present a joint computational/experimental study of a Chan-Lam aryl iodination reaction (Figure 1). Starting from a $[Cu(OAc)_2]_2{\cdot}2H_2O$ pre-catalyst the active species (**Int1**) is generated by addition of phenanthroline and NaI. Density functional theory calculations then identified the most stable forms of **Int1** and **Int2**. This allowed us to define the mechanism for the transmetalation of boronic acid and the thermodynamics of the subsequent disproportionation step. Reductive elimination from Cu(III) was computed to be extremely facile and the transmetalation was therefore proposed to be the rate-limiting step.

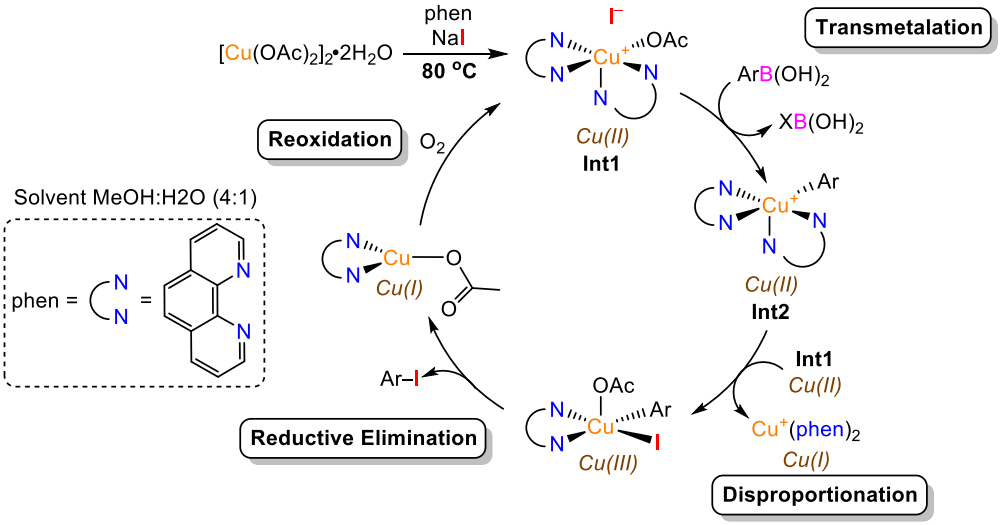


Figure 1. Proposed Catalytic Cycle for the Iodination Chan-Lam Reaction

[1] D. M. T. Chan, K. L. Monaco, R.-P. Wang and M. P. Winters, *Tetrahedron Lett.*, 1998, **39**, 2933-2936. [2] P. Y. S. Lam, C. G. Clark, S. Saubern, J. Adams, M. P. Winters, D. M. T. Chan and A. Combs, *Tetrahedron Lett.*, 1998, **39**, 2941-2944. [3] M. J. West, J. W. B. Fyfe, J. C. Vantourout and A. J. B. Watson, *Chem. Rev.*, 2019, **119**, 12491-12523.

P5 Using BioSimSpace to create an interoperable relative binding free energy pipeline

Anna M. Herz[a], Tahsin Kellici[b], Inaki Morao[b], Julien Michel[a]

[a]School of Chemistry, University of Edinburgh, Edinburgh EH9 3FJ, UK. [b]EVOTEC, Milton, Abingdon, Oxfordshire OX14 4RZ, UK.

AF (Alchemical Free Energy) methods are used to computationally predict binding affinities, with RBFE (Relative Binding Free Energy) specifically being used to rank compounds during preclinical drug discovery stages. Previous work has shown that relative free energies of hydration are reproducible to within roughly 0.2 kcal/mol between the molecular dynamics engines AMBER, SOMD, GROMACS, and CHARMM, however no universally applicable protocol was established and considerable attention to detail and package specific simulation protocols needed to be observed.[1] For RBFE, recent work has shown a reproducibility of roughly 0.5 kcal/mol between OpenMM, NAMD2, and NAMD3 but careful considerations of setup and analysis steps are crucial.[2]

This work aims to establish an interoperable FEP pipeline to enable the rapid testing of RBFE calculations and evaluate the reproducibility between three different molecular dynamics engines – SOMD, GROMACS, and AMBER. To this end the BioSimSpace python framework has been extended to support AMBER RBFE calculations. Efforts have focussed on capturing programmatically best-practices protocols for evaluating RBFEs with each simulation engine. The influence of setup or analysis protocols on the overall calculated RBFE has also been investigated by processing congeneric ligand series for the proteins tyrosine kinase 2 (TYK2), myeloid leukemia cell differentiation protein (MCL1) and mitogen-activated protein kinase 14 (p38α). We comment on the key factors that must be taken into consideration in order to achieve reproducibility of RBFEs to under 0.5 kcal/mol.

[1] H. H. Loeffler, S. Bosisio, G. Duarte Ramos Matos, D. Suh, B. Roux, D. L. Mobley and J. Michel, *J. Chem. Theory Comput.,* **2018**, *14,* 5567–5582. [2] A. D. Wade, A. P. Bhati, S. Wan and P. V. Coveney, *J. Chem. Theory Comput.*, **2022**, *18,* 3972–3987.

P6 Prediction of coiled coil oligomerisation with molecular dynamics simulations

Evangelia Notari[a], Christopher W. Wood[b], Julien Michel[a]

[a]EaStCHEM School of Chemistry, University of Edinburgh, Edinburgh EH9 3FJ, UK. [b]School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3BF, UK.

De novo protein design has mostly focused on algorithms that maximise the free energy difference between a single conformational protein state and other plausible macrostates, resulting in rigid and hyper-stable, non-functional folds. Since proteins often interconvert between different conformations to execute their functions, future progress in *de novo* protein design depends on advances in designing multi-state proteins, i.e. proteins that can switch between different states upon external perturbations. Coiled coils are an attractive design scaffold, given that they are one of few folds that can be accurately described with parametric equations and clear sequence-to-structure rules, which assist in the selection of sequences that can assemble into a desired oligomerisation state, i.e. dimer, trimer etc. However, unexpected oligomerisation states can often be adopted upon changes in the environment, such as pH shifts.[1]

We aim to develop a robust computational pipeline for predicting the free energy of oligomerisation of coiled coils, with the ultimate goal of designing multi-state coiled coils (Figure 1). We use metadynamics, an enhanced sampling Molecular Dynamics method that enables the estimation of the free energy of oligomerisation of the coiled coils, which can indicate a) whether specific oligomerisation states (e.g. dimer vs. trimer) and/or topologies (e.g. parallel vs. antiparallel) are more thermodynamically favourable over others, and b) whether two structurally distinct states are close energetically, hinting to a multi-state system.

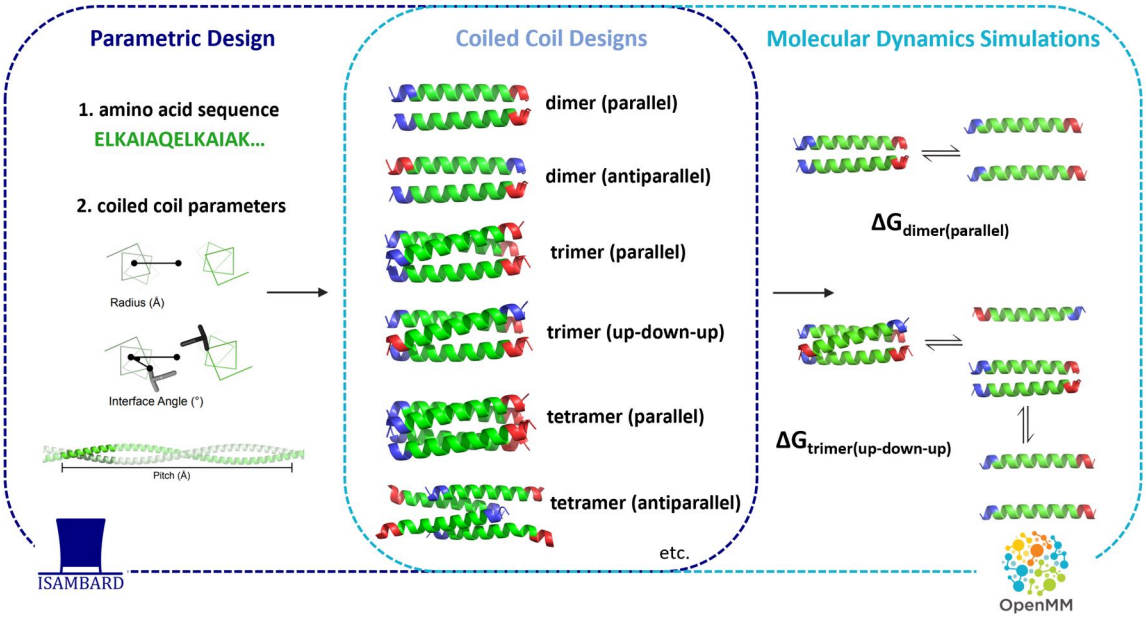


Figure 1 Proposed computational pipeline for the prediction of coiled coil oligomerisation.

[1] G. G. Rhys, C. W. Wood, J. L. Beesley, N. R. Zaccai, A. J. Burton, R. L. Brady, A. R. Thomson, D. N. Woolfson, *JACS* **2019**, *141(22)*, 8787–8797.

P7 PyRISM: statistical mechanical theory for modelling solvation

Abdullah Ahmad, David S. Palmer

Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow G1 1XL, UK.

The integral equation theory (IET) of liquids is an approach to modelling liquids and solvation that sits between explicit and implicit solvation models in terms of computational expense. The IET is a statistical-mechanical approach to understanding liquids, driven by density distribution functions to describe liquid structure and properties[1]. An open-source code (named pyRISM[2]) is reported that implements the 1-Dimensional Reference Interaction Site Model (1D-RISM) and its variations and also calculates solvation thermodynamics. Unlike existing open-source codes, which only solve the solvent-solvent 1D-RISM equations, or are only applicable to aqueous solutions at 298 K, pyRISM can be used for a wide-range of solvents and environmental conditions. We show that 1D-RISM can be used to generate descriptors for cheminformatics and machine learning purposes in the form of solvation free energy densities.

[1] Hansen, J.-P.; McDonald, I. R. *Theory of Simple Liquids*; Elsevier: Amsterdam, 1990.
[2] Ahmad, A. PyRISM v0.2.0 DOI: 10.5281/Zenodo.7107645, 2023.

P8 Towards a lead-free high sensitivity material to replace lead azide

Jack M. Hemingway[a], Peter Portius[b], Colin R. Pulham[a], Adam A. L. Michalchuk[c], Carole A. Morrison[a]

[a]EaStChem School of Chemistry and Centre for Science at Extreme Conditions (CSEC), University of Edinburgh, Edinburgh, UK. [b]Department of Chemistry, University of Sheffield, Sheffield S10 2TN, UK. [c]School of Chemistry, University of Birmingham, Birmingham B15 2TT, UK.

Energetic materials (EMs) are renowned for their ability to release a significant amount of energy upon exposure to a number of stimuli, which must be weighed against their sensitivity (the likelihood of initiation) [1]. Lead azide is a primary energetic which has both high energetic performance and is exceptionally sensitive to impact, which has resulted in it becoming a widely used detonator for more powerful secondary EMs [2]. However, lead-based compounds have many adverse environmental and public health effects, and as such its use has been limited. For these reasons, more sustainable replacements that can match lead azide for energetic performance are sought. Due to their inherent material properties, synthesis and characterisation of novel EMs is an inherently dangerous process, making the computational prediction of sensitivity an attractive alternative. Previous work within the group has resulted in the development of a vibrational up-pumping methodology for the prediction of impact sensitivity of molecular EMs computationally [3, 4]. In this work, the suitability of the vibrational up-pumping procedure when applied to metal containing coordinated energetics (such as lead azide) is explored. The impact sensitivity of lead azide and a number of copper containing potential replacements were predicted using vibrational up-pumping and rationalised through exploration of the vibrational landscape, in an attempt to reveal the origin of their sensitivity.

[1] T. M. Klapötke, *Chemistry of High-Energy Materials*, De Gruyter, Berlin, 5th edn., 2019.
[2] J. Akhavan, *The Chemistry of Explosives*, Royal Society of Chemistry, Cambridge, UK, 4th edn., 2022, pp. 35–37.
[3] A. A. Michalchuk, M. Trestman, S. Rudic, P. Portius, P. T. Fincham, C. R. Pulham and C. Morrison, Predicting the reactivity of energetic materials: an ab initio multi-phonon approach, *J. Mat. Chem. A*, **2019**, *7*, 19539–19553.
[4] A. A. L. Michalchuk, J. Hemingway and C. A. Morrison, Predicting the impact sensitivities of energetic materials through zone-center phonon up-pumping, *J. Chem. Phys.*, **2021**, *154*, 064105.

P9 Developing a general QSPR framework for property classification in KNIME

Connor MacDonald[a], Neil Berry[b], Emily Draper[a]

[a]*School of Chemistry, University of Glasgow, Glasgow G12 8QQ, UK.* [b]*Department of Chemistry, University of Liverpool, Liverpool L69 7ZD, UK.*

This project centers around the creation of machine learning models which will ultimately allow for the prediction of self-assembled aggregates. Due to the delicate balance of properties which lead to molecular self-assembly being poorly understood, the design of new materials for devices is often unachievable.[1] t is our hope that by being able to confidently predict the self-assembly behaviour of a molecule, the development of organic electronics will become much more accessible. This could pave the way for the design of responsive organic material with the capability of replacing metals in high-value mechanoresponsive devices.

To create an interpretable and freely available model, the open-source Konstanz Information Miner (KNIME) application with a library of chemistry and machine learning integrations could be utilised to develop a QSPR (Figure 1). KNIME is a visual programming platform making machine learning more accessible to those with limited programming knowledge (Figure 2).

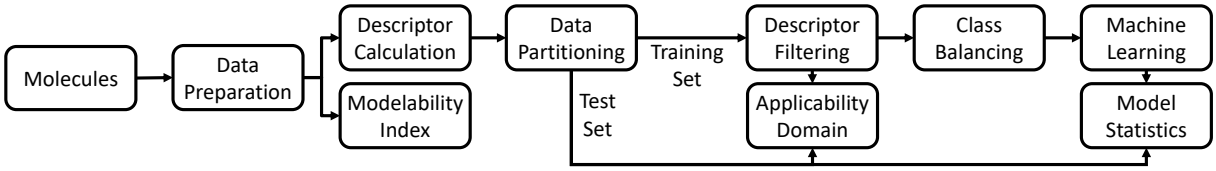


Figure 3: Flowchart representing QSPR modelling procedure.

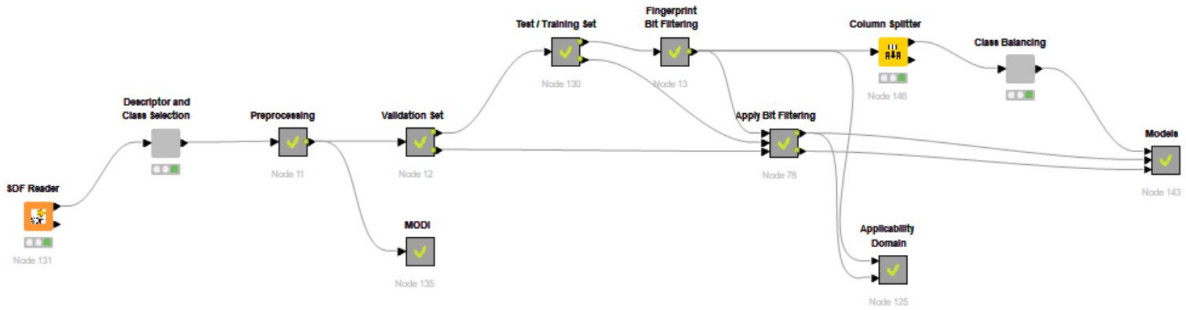


Figure 4: KNIME QSPR workflow

[1] J. Raeburn, A. Zamith Cardodo, D. J. Adams, *Chem. Soc Rev.* **2013**, *42*, 5142–5156.

P10 Prediction of pK_a *via* neural network potential

Ross J Urquhart, Alexander van Teijlingen, Tell Tuttle

Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow G1 1XL.

The logarithmic acid dissociation constant, pK_a, is an important parameter in fields such as drug discovery, analytical and environmental science among others. pK_a provides a quantitative measure of the acidity of a solution as well as a description of the ionization of a chemical. Experimentally, pK_a can be determined through titrations, calorimetry or NMR, but can also be determined theoretically through methods such as calculation through thermodynamic cycles, often relying on the accuracy of Density Functional Theory (DFT) calculations and more recently Machine Learning (ML).

Experimental determination of pK_a values are not always possible and theoretical pathways provide an alternative route to their determination. In 2004, Yates *et al.*[1] reported the determination of pK_a values for carbene molecules *via* a mixture of DFT and *ab initio* calculations. Since then, theoretical determination of pK_a has continued with studies reporting for molecules such as phenols,[3] and thiols.[4]

ML methods have also been employed in the prediction of pK_a values with toolkits such as PropKa[5] and Epik[6] readily available and parameterised towards proteins and small, drug-like molecules respectively. Often boasting accuracy near that of quantum methods and with increased speed, ML methods are increasingly attractive across many areas of chemistry including that of pK_a prediction. The ANAKIN-ME (ANI) suite[7] of neural network potentials provide near chemical accuracy at a fraction of the cost compared to DFT methods. One of the latest iterations of the suite, ANI-2x[8] boasts a 10^6 speed up over DFT whilst retaining similar chemical accuracy to the quantum method.

In this work, ANI-2x is used to predict the energies of gas and aqueous phase systems for use in the direct calculation of absolute pK_a values. Employing DFT reference calculations at differing levels of theory as validation we have explored the ability of ANI-2x for pK_a prediction on a subset of molecules.

[1] A. M. Magill, K. J. Cavell and B. F. Yates, *J. Am. Chem. Soc.*, 2004, **126,** 8717–8724.
[2] M. D. Liptak and G. C. Shields, *J. Am. Chem. Soc*, 2001, **123,** 7314–7319.
[3] S. Pezzola, S. Tarallo, A. Iannini, M. Venanzi, P. Galloni, V. Conte and F. Sabuzi, *Molecules*, 2022 **27** 8590.
[4] B. Thapa and H. B. Schlegel, *J. Phys. Chem. A*, 2016 **120** 5726–5735.
[5] M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7,** 525–537.
[6] J. C. Shelley, A. Cholleti, L. L. Frye, J. R. Greenwood, M. R. Timlin and M. Uchimaya, *J. Comput. Aided Mol. Des.*, 2007, **21,** 681–691.
[7] J. S. Smith, O. Isayev and A. E. Roitberg, *Chem. Sci.*, 2017, **8,** 3192–3203.
[8] C. Devereux, J. S. Smith, K. K. Huddleston, K. Barros, R. Zubatyuk, O. Isayev and A. E. Roitberg, *J. Chem. Theory Comput.*, 2020, **16,** 4192–4202.

P11 Blinded predictions and post hoc analysis of the second solubility challenge data: exploring training data and feature set selection for machine and deep learning models

Jonathan G. M. Conn[a], James W. Carter[a], Justin J. A. Conn[a], Vigneshwari Subramanian[b], Andrew Baxter[c], Ola Engkvist[d,e], Antonio Llinas[b], Ekaterina L. Ratkova[d], Stephen D. Pickett[f], James L. McDonagh[g], David S. Palmer[a]

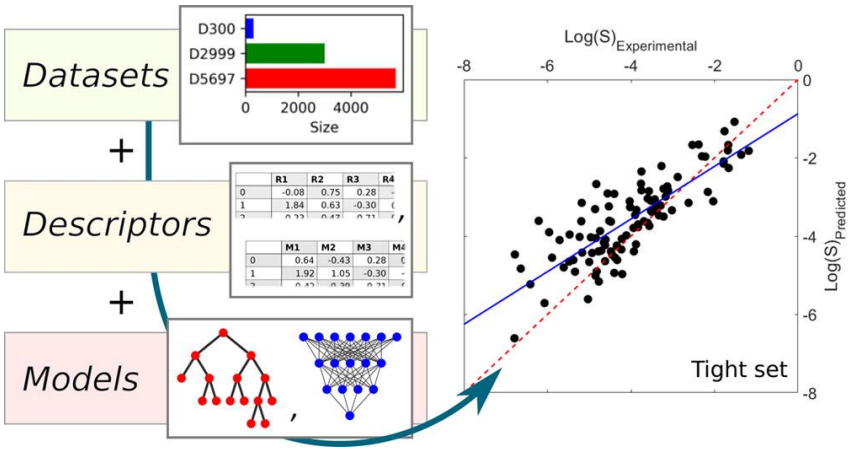
[a]Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow G1 1XL, UK. [b]Drug Metabolism and Pharmacokinetics, Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, SE-431 83 Göteborg, Sweden. [c]GSK Medicines Research Centre, Stevenage SG1 2NU, UK. [d]Medicinal Chemistry, Research and Early Development, Cardiovascular, Renal and Metabolism (CVRM), BioPharmaceuticals R&D, AstraZeneca, SE-431 50 Göteborg, Sweden. [e]Department of Computer Science and Engineering, Chalmers University of Technology, SE-312 96 Göteborg, Sweden. [f]Computational Sciences, GSK R&D Pharmaceuticals, Stevenage SG1 2NY, UK. [g]IBM Research Europe, Hartree Centre, SciTech Daresbury, Warrington, Cheshire WA4 4AD, UK.

Accurate methods to predict solubility from molecular structure are highly sought after in the chemical sciences. To assess the state of the art, the American Chemical Society organised a "Second Solubility Challenge" in 2019 [1], in which competitors were invited to submit blinded predictions of the solubilities of 132 drug-like molecules.

In the first part of this work [2], we describe the development of two models that were submitted to the Blind Challenge in 2019 but which have no previously been reported. These models were based on computationally inexpensive molecular descriptors and traditional machine learning algorithms and were trained on a relatively small data set of 300 molecules.

In the second part of the work, to test the hypothesis that predictions would improve with more advanced algorithms and higher volumes of training data, we compare these original predictions with those made after the deadline using deep learning models trained on larger solubility data sets consisting of 2999 and 5697 molecules.

The results show that there are several algorithms that are able to obtain state-of-the-art performance on the solubility challenge data sets, with the best model, a graph convolutional neural network, resulting in an RMSE of 0.86 log units. Critical analysis of the models reveals systematic differences between the performance of models using certain feature sets and training data sets. The results suggest that careful selection of high quality training data from relevant regions of chemical space is critical for prediction accuracy but that other methodological issues remain problematic for machine learning solubility models, such as the difficulty in modelling complex chemical spaces from sparse training data sets.



[1] A. Llinas, R. C. Glen, J. Goodman, *J. Chem. Inf. Model.*, **2019**, *48*, 1289-1303.
[2] J. G. M. Conn *et al.*, *J. Chem. Inf. Model.*, **2023**, *63*, 1099-1113.

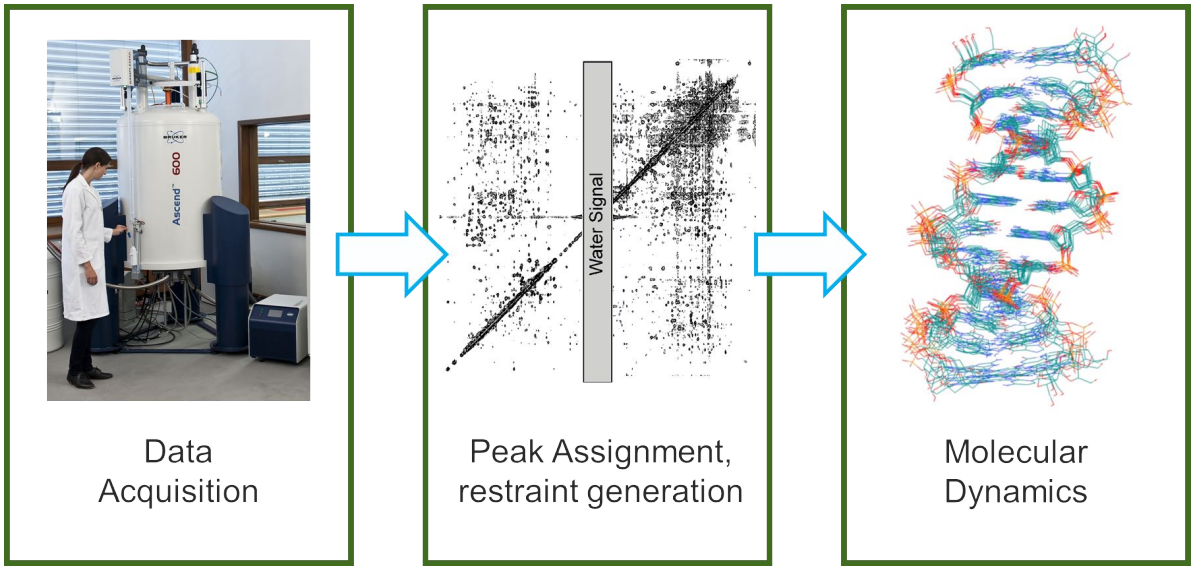
P12 Software to aid in DNA-ligand structure determination using solution NMR and restrained molecular dynamics simulations

Ivan Yankov, Jamie Whiters , David Palmer, Glenn Burley

Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow.

Introduction. Understanding the structure and dynamics of DNA-ligand complexes in solution is valuable in designing novel drug-like scaffolds. For example, structural information of minor groove binders is key to optimising the DNA binding domain of artificial transcription factors used in treating triplet expansion diseases. Coupling solution NMR with molecular dynamics simulations provides a powerful method for structural characterisation, but the complexity of the calculations and their dependence on model parameters can be a barrier to adoption.

Aims. Our aim is to transform the manual step-by-step NMR analysis process into a seamless start-to-finish uninterrupted calculation, in order to expedite the elucidation of DNA-ligand atomic coordinate ensembles.



Methods. We present a software package that combines readily available tools to automate NOESY intensity-to-distance analysis and couple it with the AMBER molecular dynamics program. The software package reduces unnecessary user interactions such that key inputs can be defined at the start of the workflow. CPU parallelisation reduces the NOESY-to-distance calculation time enabling users to receive live data during the assignment process. A template with default parameters and a tutorial on how to incorporate the workflow in your NMR analysis is presented.

Results. The application of the software is demonstrated for a palindromic dodecamer DNA duplex in complex with a novel polyamide ligand. The impact of NOESY restraints on the simulated structure of the DNA-polyamide complex is illustrated by contrasting a restrained and an unrestrained molecular dynamics simulation. The other factors that impact reproducibility of the NOESY intensity to distance conversion workflow are discussed.

Conclusion. The automated NOESY-to-distance analysis improves reproducibility and enables scalability of the process. Precious user time is reallocated from terminal manual interaction to expert analysis of the results. The tool is provided as open-source software and can be incorporated into NMR assignment software (e.g. Sparky) to aid the assignment process.

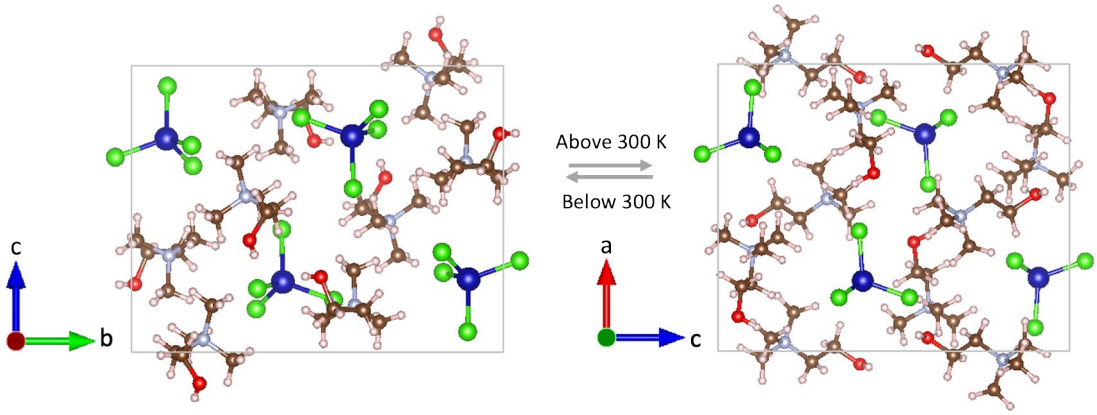
P13 First-principles modelling of choline-based plastic crystals

Shivani Grover, Joshua Levinsky, Claire Hobday

EaStCHEM School of Chemistry and Centre for Science at Extreme Conditions, University of Edinburgh, Edinburgh, UK.

Materials with large solid-state caloric effects induced by external field (mechanical, electric or magnetic field) are needed for the development of eco-friendly solid-state refrigeration technologies.[1] The refrigeration capacity is associated with a large isothermal entropy change or with a large temperature change induced by external stimuli, such as mechanical pressure (barocaloric effect) or electric field (electrocaloric effect) or magnetic field (magneto-caloric effect), effects that are enhanced near to phase transitions. To this end, organic-inorganic plastic crystals offer new interesting opportunities for the solid-state cooling applications due to promising barocaloric effects.[2] In these systems, huge entropy and/or enthalpy changes are associated with a symmetry breaking order-disorder transition, making them favourable for energy storage applications.

In the present work, we explore the potential of choline-based plastic crystals as promising barocaloric materials. Using methods of ab initio molecular dynamics as implemented in CP2K, we study the pressure- and temperature-induced structural phase transitions and properties of $(C_5H_{14}NO)CoCl_3$, and its potential as a barocaloric material.



[1] X. Moya, S. Kar-Narayan, N. D. Mathur, *Nature Materials.* **2014,** *13*, 439-450. [2] B. Li, Y. Kawakita, S. Ohira-Kawamura, T. Sugahara, H. Wang, J. Wang, Y. Chen, S. I. Kawaguchi, S. Kawaguchi, K. Ohara, K. Li, D. Yu, R. Mole, T. Hattori, T. Kikuchi, S.-i. Yano, Z. Zhang, Z. Zhang, W. Ren, S. Lin, O. Sakata, K. Nakajima, Z. Zhang, *Nature* **2019,** *567*, 506-510.

P14 Towards a less expensive method of predicting impact sensitivity

Heather M. Quayle[a], Adam A. L. Michalchuk[b], Colin R. Pulham[a], Carole A. Morrison[a]

[a]EaStChem School of Chemistry and Centre for Science at Extreme Conditions (CSEC), University of Edinburgh, Edinburgh, UK. [b]School of Chemistry, University of Birmingham, Birmingham B15 2TT, UK.

The experimental development of new energetic materials (EMs) is a process often hindered by strict safety and performance criteria.[1,2] The outcome of testing of properties related to safety, such as impact sensitivity (IS; how hard a material must be hit to cause initiation), is greatly affected by changes in variables such as temperature, sample purity and grain size.[3,4] To that end, to develop new materials it has become important to be able to predict properties of EMs before manufacture, for example with computational modelling. A predictive model for IS, based on the vibrational up-pumping method of energetic impact initiation, has been produced and tested on a wide range of EMs.[5] Since sampling of the full Brillouin zone via density functional theory (DFT) is too computationally demanding, the model was adapted[6] to use only the gamma point to calculate vibrational spectra. Consequently, this model can be used well in a comparative nature, i.e., to predict if a material is a primary energetic (more sensitive to impact) or a secondary energetic (less sensitive), but works less well in producing a ranking of sensitivities, because under-sampling of the Brillouin zone leads to some loss of accuracy. The aim of this work is to develop a method to be able to sample the full Brillouin zone in a less computationally demanding way to speed up this process and allow for high throughput calculations. Initial developments have been made using the DFTB+ implementation[7] of the density functional tight binding (DFTB) model. This work will show the extent of progress made in this method development and probable next steps to improve the method.

[1] J. J. Sabatini and K. D. Oyler, *Crystals*, **2016,** *6*, 5.
[2] P. A. Davies, *J. Hazard. Mater*., **1994**, *38*, 75.
[3] V. J. Bellitto and M. I. Melnik, *Applied Surface Science*, **2010**, *256*, 3478–3481.
[4] R. W. Armstrong, C. S. Coffey, V. F. DeVost and W. L. Elban, *Journal of Applied Physics*, **1990**, *68*, 979–984.
[5] A. A. L. Michalchuk, M. Trestman, S. Rudic, P. Portius, P. T. Fincham, C. R. Pulham and C. A. Morrison, *J. Mat. Chem. A,* **2019**, *7*, 19539–19553.
[6] A. A. L. Michalchuk, J. Hemingway and C. A. Morrison, *J. Chem. Phys.,* **2021**, *154*, 064105.
[7] B. Hourahine et al*., J. Chem. Phys*. **2020**, *152*, 124101

P15 Development of molecular descriptors for quantitative structure-retention relationships

Madeleine Taylor[a], Roman Szucs[b], Lucy Morgan[c], Roland Brown[c], David Palmer[a]

[a]Pure and Applied Chemistry, University of Strathclyde, G1 1XL, UK. [b]Department of Analytical Chemistry, Faculty of Natural Sciences, Comenius University, Bratislava, Slovakia; [c]Pfizer Global R&D, Sandwich, UK.

High performance liquid chromatography (HPLC) is used at every stage in the development of new pharmaceutical compounds and processes. Quantitative structure-retention (QSRR) models are widely used to identify compounds in HPLC screening experiments, and to guide chromatographic method development[1]. The accuracy of QSRR models relies on high quality, appropriate molecular descriptors. The one-dimensional reference interaction site model (1D RISM) is a model of solvation based on spherically symmetrical density correlation functions. In this project, new molecular descriptors, tailored for retention time predictions, are developed from 1D RISM theory and tested in QSRR models. The usefulness of these physics-based descriptors has been proven previously for predictions of solvation free energy[2], entropy of solvation, and enthalpy of solvation[3].

HPLC retention time datasets for a diverse set of druglike molecules in 6 different chromatographic systems have been provide by Pfizer. 1D RISM equations were solved for the analyte molecules in various solvents with pyRISM solver software[4]. The calculations have been carried out for a selection of solvents that are common for reversed phase HPLC mobile phases, to tailor the descriptors to the purpose of QSRR. Molecular descriptors were extracted from the solvation free energy density function for each solute-solvent pair. Figure 1 shows the solvation free energy density for a small selection of druglike molecules in water, methanol, acetonitrile and coarse-grained acetonitrile to demonstrate the differences in solvation structure, and the molecular descriptors derived from them.

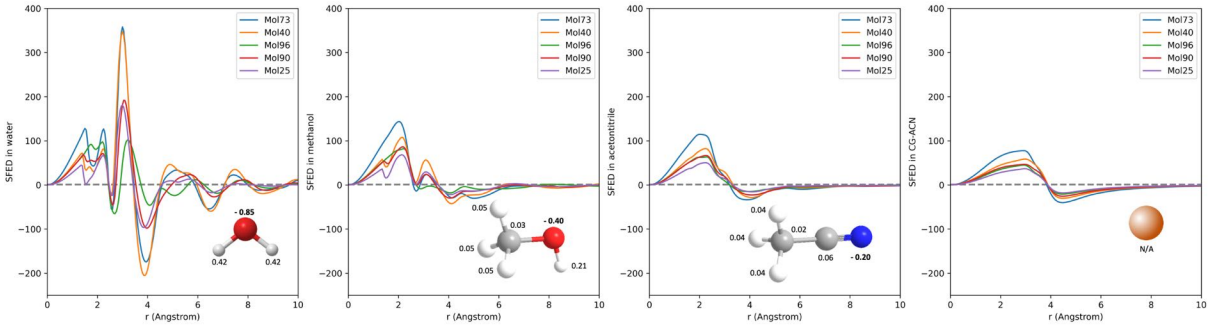


Figure 1. Solvation free energy density functionals for 5 molecules in water, methanol, acetonitrile, and coarse-grained acetonitrile.

QSRR models have been built using these descriptors, and they have been compared to other standard descriptors, including Mordred and VolSurf descriptor sets. The results demonstrate the applicability and usefulness of 1D RISM descriptors in QSRR models.

In future work, further methods of extracting descriptors from 1D RISM outputs will be tested. The chromatographic stationary phase will be modelled, to be used in combination with the mobile phase descriptors discussed above, to allow the dynamic equilibrium to be represented with these bespoke descriptors.

[1] P. R. Haddad, M. Taraji and R. Szücs, *Anal. Chem.*, **2021**, *93*, 228–256. [2] D. S. Palmer, M. Mišin, M. V. Fedorov and A. Llinas, *Mol. Pharm.*, **2015**, *12*, 3420–3432. [3] D. J. Fowles and D. S. Palmer, *Phys. Chem. Chem. Phys.*, **2023**, *25*, 6944–6954. [4] A. Ahmad, 2AUK/pyRISM, DOI: 10.5281/zenodo.7783600, 2023.

P16 A Multifunctional Drug Delivery System Based on Switchable Peptide-Stabilized Emulsions

Daniel Boas[a, b], Alexander van Teijlingen[c], Zohar Shpilt[a], Deborah E. Shalev[d, e], Edit Y. Tshuva[a], Tell Tuttle[c], Meital Reches[a, b]

[a]Institute of Chemistry, Hebrew University of Jerusalem, Jerusalem, 9190401, Israel. Center for Nanoscience and Nanotechnology, Hebrew University of Jerusalem, Jerusalem, 9190401, Israel. [c]WestCHEM, Department of Pure and Applied Chemistry, University of Strathclyde, Glasgow G1 1XL, UK. [d]Department of Pharmaceutical Engineering, Azrieli College of Engineering, Jerusalem, Israel. [e]Wolfson Centre for Applied Structural Biology, Hebrew University of Jerusalem, Jerusalem, Israel.

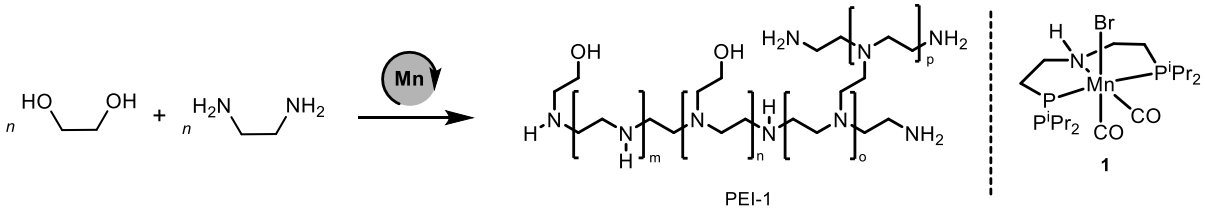
Emulsions are commonly used for drug delivery, yet they are usually limited to exclusively delivering either lipophilic compounds or hydrophilic compounds. This separation negates possible synergetic therapeutic roles between those compounds. Here, we introduce a novel design for a short peptide that can stabilize emulsions. Upon binding certain metal ions, the peptide acts as a molecular switch, changes conformation, and becomes amphiphilic. Spectroscopic methods, NMR, and molecular dynamics provide information on the mechanism of this complexation-triggered amphiphilicity. The stability of this unique emulsion is based on pH triggered histidine-metal bonds, which break under acidic conditions, granting the emulsion selectivity towards the extracellular pH of tumors.

P17 DFT study of the synthesis of polyethyleneimines from manganese catalysed coupling of ethylene glycol and ethylenediamine

Aniekan E. Owen, Claire N. Brodie, Julian S. Kolb, Michael Bühl, Amit Kumar

School of Chemistry, University of St Andrews, Scotland, UK. E-mail: aeo1@st-andrews.ac.uk.

Mn catalysed coupling of ethylene glycol and ethylenediamine provides a safer route for the synthesis of polyethyleneimines (PEI) from the drawback of aziridine which is known to be highly reactive, toxic, corrosive and volatile. PEI has applications in a range of industrial processes and products. The mechanistic studies of the catalytic pathway as supported by DFT calculations at the PBE0-D3/def2-TZVP/PCM//RI-BP86/def2-SVP/PCM level of theory show the reaction would favourably proceed by the formation and subsequent hydrogenation of imine intermediates [1]. Thermodynamic driving forces are studied for the most stable conformer of each intermediate encountered through various possible pathways. The key catalytic steps involve a number of (de)hydrogenation and (de)hydration processes using a Mn-PNP pincer catalyst **1** (with the so-called macho ligand). Regeneration of the active catalyst is computed to be the rate-limiting step, and the overall barrier for the process is influenced by the existence of an off-cycle adduct formed between the active catalyst and ethylene glycol, which has also been observed experimentally.



Scheme 1: Synthesis of PEI from the catalytic coupling of ethylene glycol and ethylenediamine using Mn-PNP macho complex.

[1] A. E. Owen, C. N. Brodie, J. S. Kolb, M. Bühl, A. Kumar, *Angew. Chem. Int. Ed.* **2023**, DOI: 10.1002/anie.202306655.

P18 First-principles study of electronic and optical properties in 1-dimensional oligomeric derivatives of telomestatin

Joelle Mergola-Greef[a], Bruce F. Milne[a, b]

[a]Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Meston Building, Old Aberdeen AB24 3UE, UK. [b]CFisUC, Departamento de Fisica, Universidade de Coimbra, 3004-516 Coimbra, Portugal.

The ground-state electronic structure and optical absorption profiles of a series of linear oligomers (n = 1–10) inspired by the macrocyclic natural product telomestatin (Figure 1) were investigated using real-space self-interaction corrected (time-dependent) density functional theory. The neutral species exhibit length-dependent formation of plasmonic excitations in the UV region, which is augmented by polaron-type absorption in the IR when the chains are doped with an extra electron/hole. When combined with a lack of absorbance in the visible region, these oligomers appear to be promising candidates for applications such as transparent antennas in dye-sensitised solar energy collection materials.

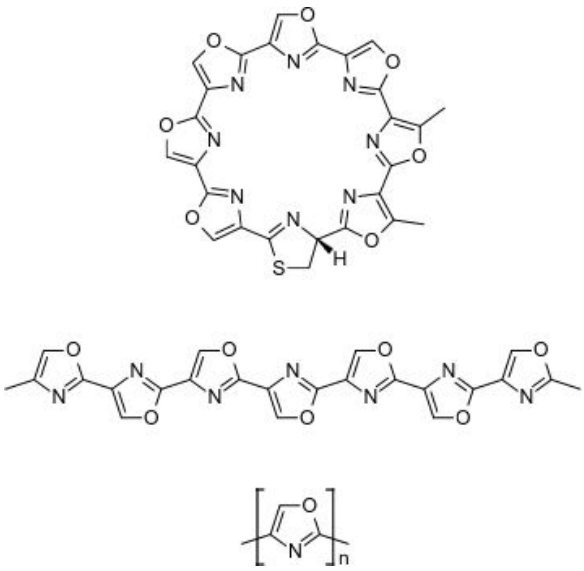


Figure 1. The all-cis natural product *R*-telomestatin (top), telomestatin-based linear all-trans model (middle), and overall schematic of the methyl-capped linear 1,4-polyoxazole oligomers explored in this study (bottom).

P19 Rationalising selective gold recovery through solid state computational modelling

Susanna Vance, Mateusz Mojsak, Carole A. Morrison, Jason B. Love

School of Chemistry, University of Edinburgh, Edinburgh EH9 3FJ.

Waste electronics and electrical equipment (e-waste) has become the fastest growing hazardous solid waste stream in the world, with approximately only one fifth being recycled.[1] Achieving its sustainable management is therefore crucial to moving towards a circular economy, and given the high content of base and precious metals in e-waste (often exceeding those in natural mineral ores), recycling is of great economic benefit.[2] Many thermo- and biochemical processes designed to recover metals from their ores have been adapted for metal recovery from e-waste, with hydro-, bio- and pyro-metallurgical processes the most common.[3] Recyclable precipitation processes have become increasingly popular as they provide significant advantages over solvent extraction and single-use precipitation methods.[4]

We have employed a simple tertiary diamide in the targeted precipitation of gold and other metals from acidic leach solutions from a variety of sources,[5] with excellent selectivity for gold easily achieved through tuning experimental conditions (see Figure 1). However, the factors that dictate the selectivity for gold over other metals still remain elusive. Therefore we have undertaken an extensive computational study to investigate if thermodynamics play a significant role in selectivity. To achieve this, a bespoke computational approach was developed, which facilitated the calculation of exchange energies for species in the solid state (Equation 1).

$$[MCl_4][Na]_{(s)} + [AuCl_4][LH]_{(s)} \rightleftharpoons [AuCl_4][Na]_{(s)} + [MCl_4][LH]_{(s)} \quad (1)$$

In addition, ligand strain was also quantified using this method and key intermolecular interactions were studied using quantum theory of atoms in molecules (QTAIM) and non-covalent interaction (NCI) plots, which rationalised the thermodynamic results.

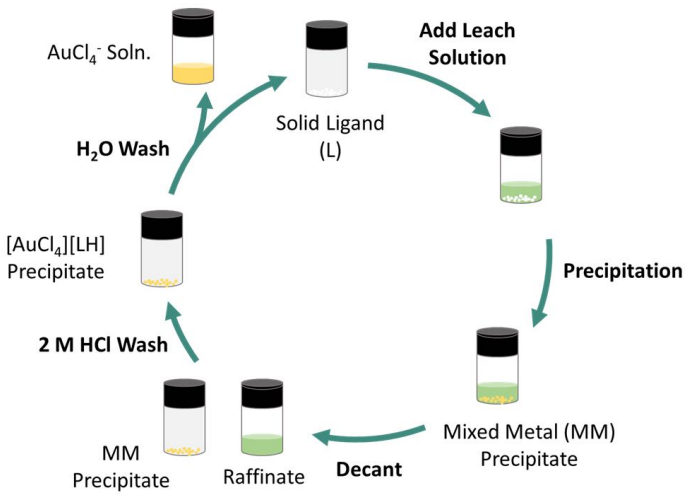


Figure 1. Overall schematic of precipitation process

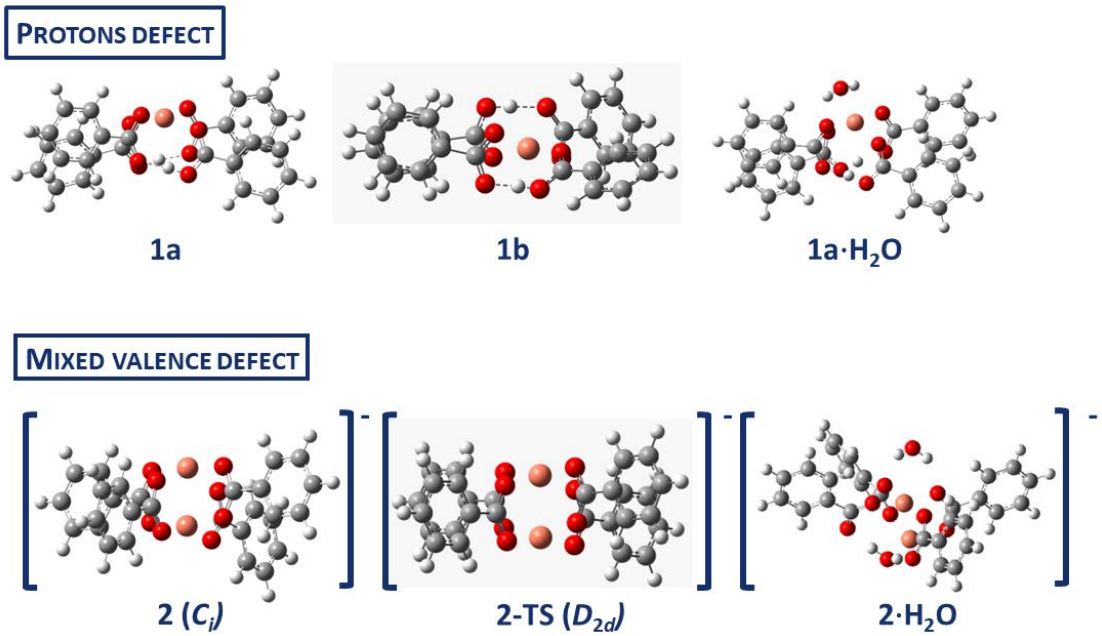
[1] A. K. Awasthi, J. Li, L. Koh and O. A. Ogunseitan, *Nat. Electron.*, **2019**, *2*, 86-89. [2] P. Tanskanen, *Acta Mater.*, **2013**, *61*, 1001-1011. [3] B. Debnath, R. Chowdhury and S. K. Ghosh, *Front. Environ. Sci. Engin.*, **2018**, *12.* [4] A. E. Lewis, *Hydrometallurgy*, **2010**, *104*, 222-234. [5] L. M. M. Kinsman, B. T. Ngwenya, C. A. Morrison and J. B. Love, *Nat. Commun.*, 2021, **12**, 6258.

P20 ^{13}C pNMR shifts of MOFs based on Cu(II)-dimers – DFT predictions for spin-½ defects

Edoardo Fusco, Michael Bühl

University of St Andrews

The HKUST and STAM families of MOFs are based on dinuclear Cu(II) paddlewheel dimer moieties (with spin 0 or 1). EPR experiments have furnished evidence for the presence of spin-½ centres.[1] These are presumed to arise from mononuclear Cu(II) defects in the normal paddlewheel moiety or from a mixed valence coordination site. So far, such spin-½ defect species were not detected in the solid-state NMR experiments,[2] which could be due to occurrence of the signals in a spectral area that has not been screened so far. As not even the order of magnitude of NMR chemical shifts of paramagnetic species (pNMR) can be reliably estimated a priori, *in silico* predictions are needed. We now report predictions of pNMR shifts of suitable defect models, namely a neutral model, where one of the two paddlewheel copper atoms is replaced either by two bridging protons [Cu(H)$_2$ defect] and a negatively charged, mixed-valence defect where one copper atom adopts the +1 oxidation state [Cu(I)Cu(II) defect] (Figure below).



The ^{13}C pNMR shifts and their temperature dependence were computed at the CAM-B3LYP/IGLO-II//PBE0-D3 level of theory, similar to the methodology employed by Ke, et al.[2] The defect centres showed a clear difference in the predicted pNMR spectra compared to the copper paddlewheel dimers, both in terms of spectral positions of the peaks and, even more importantly, their temperature dependence. Spin density analysis on the mixed valence complex revealed how its localised vs. delocalised nature impacts on structural details and on the spin delocalisation into the aromatic rings.

Our predictions should be useful to aid in the characterisation of spin-½ centres in MOFs based on Cu paddle wheel dimer motifs, suggesting that variable temperature experiments would be particularly suitable to distinguish between the defective and non-defective sites. Because the predicted pNMR shifts are quite sensitive to the nature of the defect, NMR should be an ideal analytical tool characterise these defects.

[1] H. EL Mkami, M.I.H. Mohideen, C. Pal, A. McKinlay, O. Scheimann, R.E. Morris, *Chem. Phys. Lett*. **2012**, 544, 17. [2] Z. Ke, D.M. Dawson, S.E. Ashbrook, M. Bühl, *Chem. Sci.* **2022**, 13, 2674.

P21 Barocaloric cooling: a combined experimental and computational study

Phillippa Partridge[a], Jenny Pringle[b], Charles J McMonagle[c], Anthony E. Phillips[d], Richard J. C. Dixey[d], Joshua Levinsky[a], Claire L. Hobday[a]

[a]EaStCHEM School of Chemistry and Centre for Science at Extreme Conditions, University of Edinburgh, Edinburgh, UK. [b]Institute for Frontier Materials, Deakin University, Melbourne, Australia. [c]SNBL at ESRF, Grenoble, France. [d]School of Physical and Chemical Sciences, Queen Mary University of London, London, UK.

Refrigeration accounts for 17% of the global electricity usage. The current vapour-compression technology releases greenhouse gases which has led to the demand for solid-state replacements.[1] Refrigerant effects can be induced in materials named barocaloric solids through the application of hydrostatic pressure which causes a change in isothermal entropy and adiabatic temperature of the system.

Organic ionic plastic crystals (OIPCs) are a class of materials which exhibit at least one solid-solid phase transition upon varying the temperature of the system. Upon cooling these structures, the mobility of the ions becomes limited, increasing the order within the structures. It is thought that OIPCs have the ability to undergo disorder-order phase transitions which would result in large entropy changes.[2] In this work we investigate the phase behaviour of the OIPC (cyanomethyl)trimethylammonium hexafluorophosphate, via experiment and simulation. Variable temperature and high-pressure crystallography and thermoanalytical methods were performed to understand if it has the desired characteristics of a barocaloric material. At room temperature, (cyanomethyl)trimethylammonium hexafluorophosphate is in a high symmetry, disordered, trigonal phase, upon cooling to 253 K the structure transitions to an ordered monoclinic phase. The isothermal entropy change associated with this solid-solid phase transition has been experimentally determined to be 138 J K^{-1} kg^{-1} and molecular dynamic simulations will be used in order to understand the configurational, rotational and vibrational contributions towards this value.

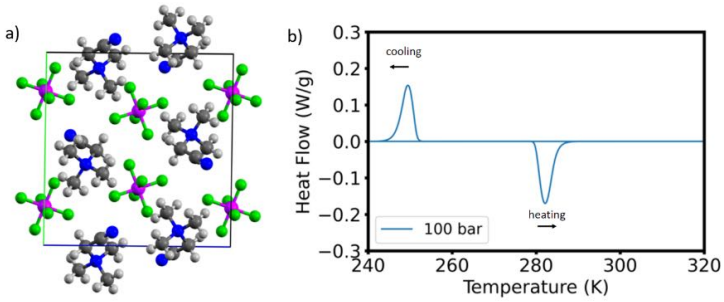


Figure 1. a) The low temperature structure of (cyanomethyl)trimethylammonium hexafluorophosphate at 120 K in P21/c. b) Differential Scanning Calorimetry plot upon heating and cooling at 100 bar showing the reversible phase transition.

[1] X. Moya and N. D. Mathur, *Science*, **2020**, *370*, 797–803. [2] J. M. Pringle, P. C. Howlett, D. R. MacFarlane and M. Forsyth, *J. Mater. Chem.*, **2010**, *20*, 2056.

P22 Quantifying intermolecular dispersion effects in σ-alkane complexes in the solid state

Carlos Martín-Fernández, M. Arif Sajjad, Stuart Macgregor

Institute of Chemical Sciences, Heriot-Watt University, Edinburgh EH14 4AS.

Amongst the many tools that have been developed for the analysis of non-covalent interactions,[1] Energy Decomposition Analyses (EDAs) have proven to be of great use. Of the many different methods and partitioning schemes available,[2] those based on local correlation methods can provide a well-founded definition of the different terms by taking into account the excitations of the localised orbitals on different fragments (see Fig 1a).[3]

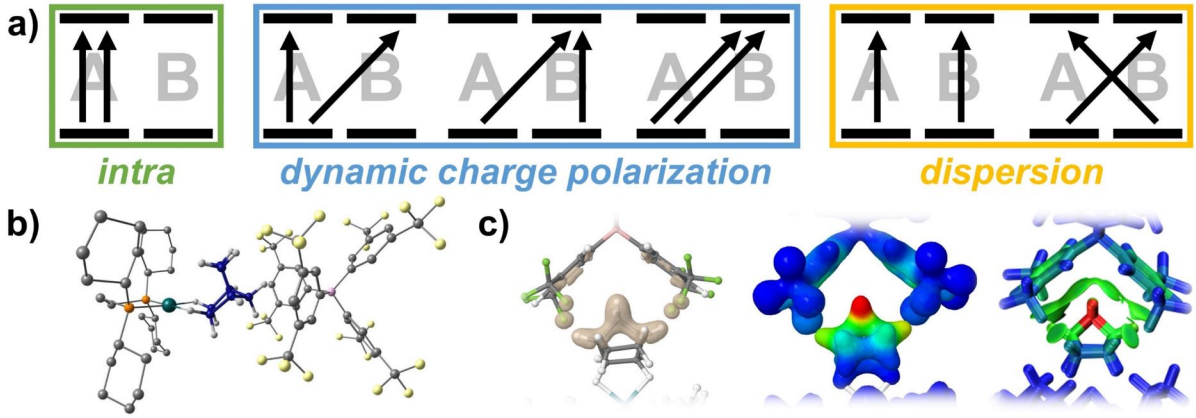


Figure 1. a) LED partition of the double excitations. b) [1-NBA]+ complex highlighting the interaction with the top [BArF4]− anion c) DID and mapped DID plots compared to IGM results for the top [1-NBA][BArF4] ion pair.

In this work, we have applied the recently developed local energy decomposition analysis (LED) based on DLPNO-CCSD(T)[4,5] calculations to compare a σ-norbornane complex (NBA, C_7H_{12}, see Fig. 1b) with a σ-propane analogue, that have both been thoroughly characterized both experimentally and with the aid of computational tools.[6-8] While *intra*molecular interactions between the NBA and the {RhP_2} fragment in **[1-NBA]**$^+$ have been previously assessed with LED methods,[9] here we consider *inter*molecular non-covalent interactions between the cations and the surrounding anions, thus accounting for the effect of the crystal environment. Due to the large size of the systems, together with the large computational cost associated with the DLPNO-CCSD(T) calculations, extensive benchmarking has been carried out in model systems to ensure a robust level of theory.

We show that around 65-70% of the interaction energy for the studied ion pairs is already described at the Hartree-Fock level, since it accounts for most of the electrostatic interaction arising from the charge separation. Interestingly, over 90% of the (remaining) correlation interaction energy corresponds to London dispersion, pointing to the importance of dispersion as a stabilizing factor for these systems. Additional insight into dispersive interactions is obtained from Dispersion Interaction Density (DID) plots,[10] which we can compare to more qualitative Independent Gradient Model (IGM) results (Fig. 1c).[8]

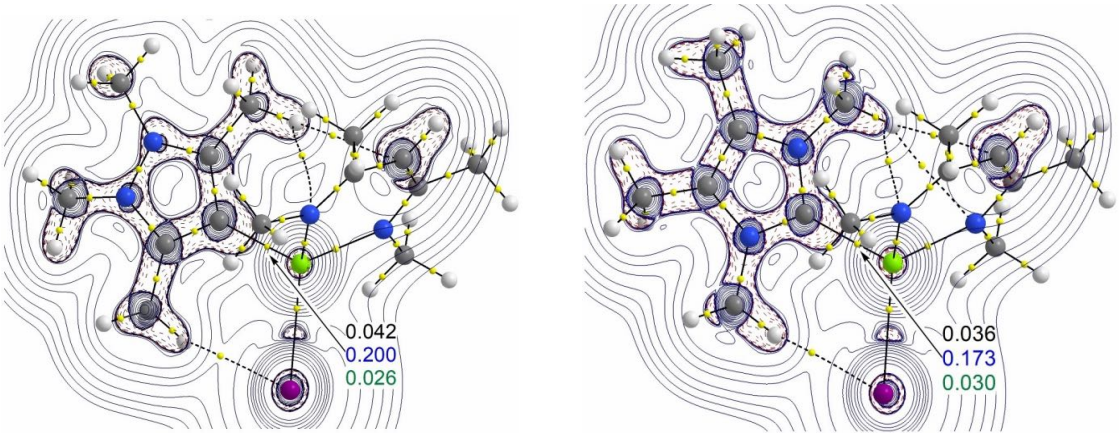
[1] E. Pastorczak and C. Corminboeuf, *J. Chem. Phys.* **2017**, *146*, 120901. [2] J. Andrés, P. W. Ayers, R. A. Boto et al., *J. Comput. Chem*. **2019**, *40*, 2248–2283. [3] M. Schütz, G. Rauhut and H.-J. Werner, *J. Phys. Chem. A* **1998**, *102*, 5997–6003. [4] G. Bistoni, *WIREs Comput Mol Sci*. **2020**, *10*, e1442. [5] W. B. Schneider, G. Bistoni, M. Sparta, et al., *J. Chem. Theory Comput.,* **2016**, *12*, 4778–4792. [6] A. G. Algarra, A. L. Burnage, M. Iannuzzi, et al., *Struct. Bond.*, **2020**, *186*, 183–228. [7] A. J. Bukvic, A. L. Burnage, G. J. Tizzard, et al., *J. Am. Chem. Soc.* **2021**, *143*, 5106–5120. [8] M. A. Sajjad, S. A. Macgregor and A. S. Weller, *Faraday Discuss.*, **2023**, *ASAP* DOI:10.1039/D3FD00009E. [9] Q. Lu, F. Neese and G. Bistoni, *Phys. Chem. Chem. Phys.*, **2019**, *21*, 11569–11577. [10] A. Wuttke and R. A. Mata, *J. Comput. Chem.* **2017**, *38*, 15–23.

P23 Exploring the bonding in β-diketiminate magnesium carbene and magnesium chalcogenide complexes: a DFT and QTAIM approach

Rochelle Ferns, Andreas Stasch, Tanja Van Mourik

School of Chemistry, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST.

Density Functional Theory and Quantum Theory of Atoms in Molecules were used to study the bonding in isomeric N-heterocyclic ligands viz. 1,2,3,5-tetramethylpyrazol-2-ylidene and 1,3,4,5-tetramethylimidazol-2-ylidene and their magnesium compounds. Cut back models of the compounds pyrazol-4-ylidene complex [(^{Dip}nacnac)MgI(^{Me}PZ)] and the isomeric imidazol-2-ylidene complex [(^{Dip}nacnac)MgI(^{Me}NHC)], (where ^{Me}PZ is 1,2,3,5-tetramethylpyrazol-2-ylidene, ^{Me}NHC is 1,3,4,5-tetramethylimidazol-2-ylidene, Dip = 2,6-*i*Pr$_2$-C$_6$H$_5$, nacnac = HC(MeCNAr)$_2$) were studied with Dip substituents replaced by methyl groups. DFT calculations were performed using Gaussian 16 at the M06/def2-TZVP and M06L/def2-SVP level of theory.[1,2] Natural Population Analysis (NPA) charges and Wiberg bond orders were calculated using Natural Bond Orbital (NBO) analysis. The optimised geometries of the molecules agree well with the available data from X-ray diffraction. Additionally, Quantum Theory of Atoms in Molecules was used to gain further insight into bonding in these compounds.[3] Bond critical points, bond ellipticity and Laplacian of the molecules were analysed to study the differences in the two compounds. It was observed that ^{Me}PZ is significantly less stable (ΔG= +45.4 kcal/mol, ΔH= +46.1 kcal/mol) than ^{Me}NHC. This difference is reduced by around 10 kcal/mol when these nucleophilic molecules are coordinated to the (^{Me}nacnac)MgI fragment (ΔG= +36.2 kcal/mol, ΔH= +36.5 kcal/mol). The results also confirm the higher nucleophilic nature of the ^{Me}PZ ligand as compared to the more common ^{Me}NHC carbene.[4] Furthermore, the molecules [{(^{iPrDip}NacNac)Mg}$_2$(μ-E)], E = O, S, Se and their cut back model systems [{(^{MeMe}NacNac)Mg}$_2$(μ-E)] (^{MeMe}NacNac = HC(MeNMe)$_2$), E = O, S, Se were studied to gain insight into the influence of the ligand bulk on the systems. DFT calculations performed at the M06/def2-TZVP level of theory showed that the full ligand model species [{(^{iPrDip}NacNac)Mg}$_2$(μ-E)] reproduced the overall structures found by X-ray diffraction. However, the cut back models [{(^{MeMe}NacNac)Mg}$_2$(μ-E)] reproduce only the linear Mg–O–Mg geometry with co-planar metal-ligand arrangements, but show significant bent Mg–E–Mg units for E = S and Se resulting in sub-90° bond angles.



QTAIM contour plots of the Laplacian of the electron density (solid lines positive, dashed lines negative) of [(^{Me}nacnac)MgI(^{Me}PZ)] (left) and [(^{Me}nacnac)MgI(^{Me}NHC)] (right). Values for the electron density, ∇ρ(r) (black) in e/bohr^3, Laplacian, ∇^{2}ρ(r) (blue) in e/bohr^5, and bond ellipticity, *e* (green), are given for the bond critical point (yellow) on the Mg–C bond path (black).

[1] Zhao, Y., Truhlar, D. G., *Theor. Chem. Acc.*, **2008**, *120*, 215-241.
[2] Weigend, F., Ahlrichs, R., *Phys. Chem. Chem. Phys.*, **2005**, 3297-3305.
[3] Platts, J. A., Overgaard, J., Jones, C., Iversen, B. B., Stasch, A., *J. Phys. Chem.*, **2009**, *115*, 194-200.
[4] Burnett S., de Vere-Tucker M., Davitt M., Cordes, D.B., Slawin, A.M., Ferns, R., van Mourik, T., Stasch, A., *Z. Anorg. Allg. Chem.,* **2022**, 648, e2022002.

P24 AI and AutoDock to Study a Novel Halogenase JDH3

Olena Holodaieva, Rebecca. J. M. Goss, Ying Zhan

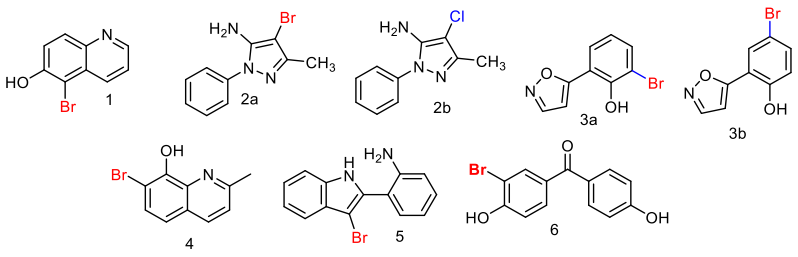
School of Chemistry, University of St Andrews

Many organic compounds with pharmaceutical applications or intermediates for their synthesis are halogenated with chlorine gas. In this case, a mixture of halogenation products is usually formed. VirX1 and other FAD-dependent Halogenases catalyse regioselective halogenation reactions under more favourable conditions compared to chlorine gas, which has a positive effect not only on the directionality and yield reaction, but also on the pricing policy of the final product and the environmental safety of the modern synthesis.

The structure of the VirX1 enzyme and its mechanism of action was studied by members of the Goss Lab.[1–3] However, there is a large variety of halogenases of which the spatial structure, their substrate and regioselectivity have not been discovered yet. That opens up a lot of opportunities for studying possible biotransformations.

JDH3 structure was predicted with AlphaFold and Phyre2. Comparison of the two models revealed differences in the structure of the internal pocket. The Phyre2 model demonstrated a wider and shorter pocket, then AlphaFold (size 62 × 71 × 85; volume 374170 Å^3 and size 60 × 72 × 94; volume 4060808 Å^3; relatively). The last model was chosen for the next research.

Active site with catalytic LYS was recognized, a map of the possible binding site was created using AutoSite v1.0. Presumably the binding site is located in the depth of the cavity containing the second WxWxIP and third FxxPxxSxG motifs. Flexible docking with grid box with size 22 × 22 × 24 was carried out with AutoDock (total number of torsions: 28).



Flexible docking allowed more accurate predict of the regioselectivity. So, for compounds **1**–**2** and **4**–**6** the only possible product of halogenation was predicted, while for structure 3 two possible regioisomers were predicted. In **3a** and **3b** cases, binding energy was –5.5 kcal/mol with distance to catalytic LYS 8.3 Å. This result was confirmed experimentally.

In all cases, the substrates were located into the depth of the inner cavity, relative potential binding site and interacted with it mainly hydrophobically with amino acid residues TYR and TRP (second motif), as well as PHE, PRO (third motif). In most cases, ILE (second motif) also interacted hydrophobically with the substrates. Hydroxyl group of the compounds **1**,**3**,**4** and **6** was stabilized by hydrogen bonds with SER of the third motif, which played a key role in compound orientation relative catalytic residue. Overall, second and third motifs promoted the specific orientation of the substrate relative to the catalytic residue. It can be stated with some certainty that the JDH3 binding site consists of the above listed residues.

To sum up, all docking results were confirmed experimentally, which indicates the correctness of the working model and the possibility of in-depth study of the binding site structure, the substrate and regioselectivity.

[1] D. S. Gkotsi, H. Ludewig, S. V. Sharma, J. A. Connolly, J. Dhaliwal, Y. Wang, W. P. Unsworth, R. J. K. Taylor, M. M. W. McLachlan, S. Shanahan, J. H. Naismith and R. J. M. Goss, *Nat. Chem.*, **2019**, *11*, 1091–1097.
[2] H. Ludewig, S. Molyneux, S. Ferrinho, K. Guo, R. Lynch, D. S. Gkotsi and R. J. M. Goss, *Curr. Opin. Struct. Biol.*, **2020**, *65*, 51–60.
[3] C. Crowe, S. Molyneux, S. V. Sharma, Y. Zhang, D. S. Gkotsi, H. Connaris and R. J. M. Goss, *Chem. Soc. Rev.*, **2021**, *50*, 9443–9481.

P25 Experimental and computational study of veratric acid synthesis starting from vanillin

Michael Nicolaou, Emma Gibson, Laia Vilà-Nadal

School of Chemistry, University of Glasgow, Glasgow, G12 8QQ, UK

Lignocellulosic biomass refers to the biomass obtained from plant matter and mainly consists of cellulose, hemicelluloses and lignin[1]. These biopolymers can be cleaved through acid catalysis or oxidation and used as feedstock for the production of value-added products (valorisation), such as platform chemicals, intermediates and fuel alternatives or additives. The use of acids as catalysts, however, impedes the green agenda of such reactions, as the use of homogeneous acids results in the production of aqueous waste, which need to be disposed of, as well as the need for separation processes.

Polyoxometalates (POMs) are transition metal (W, Mo, V)-oxo cluster species that can act as Lewis acids[2], as well as Bronsted acids in their protonated forms. POMs have also been demonstrated to anchor onto catalyst supports, forming heterogeneous catalysts[3]. Their ability to act as solid acid catalysts and application to biomass valorisation is already a field that is being explored and their have shown viable catalytic activity for reactions such as fructose dehydration to hydroxymethylfurfural (HMF)[4-5], a valuable product recognised to be amongst the "top 12" biorefinery-derived platform chemicals[6] and vanillin.

Here, we present preliminary computational work done to characterise vanillin, as well as a proposed synthesis of veratric acid starting from vanillin, using various density functional theory (DFT) methodologies.

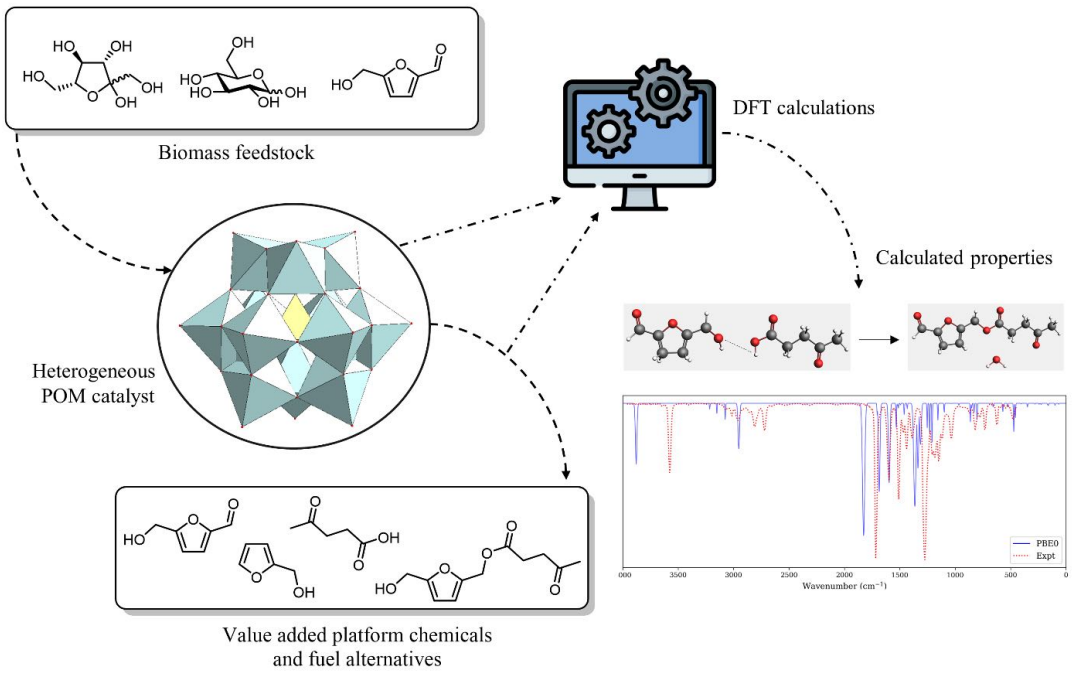


Figure 1. Polyoxometalate catalysts can be used for the conversion of biomass-derived compounds to value-added products. DFT computational methods can help elucidate the theoretical and mechanistic details of such reactions.

[1] Mathew, A. K. *et al.*, in: *Waste Biorefinery: Potential and perspectives*, **2018**, *ch. 9*, 267-297. [2] Izumi, Y. *et al.*, in: *Zeolite, clay and heteropoly acid in organic reactions*, **1992**, *ch. 3*, 99-157. [3] Chimienti, M. E. *et al.*, *Appl. Catal. A.*, **2001**, *208*, 7-19. [4] Qiu, G. *et al.*, *RSC Adv.*, **2018**, *8*, 32423-32433. [5] Da Costa, N. L. *et al.*, *Mol. Catal.*, **2021**, *500*, 111334. [6] Bozell, J. J. and Petersen, G. R., *Green Chem.*, **2010**, *12*, 539-554.

P26 Computational study into the effects of counter-cations on the $[P_8W_{48}O_{184}]^{40-}$ polyoxometalate wheel

Daniel Malcolm and Laia Vilà-Nadal

University of Glasgow, School of Chemistry, Glasgow G12 8QQ, UK.

We have recently been benchmarking the $[P_8W_{48}O_{184}]^{40-}$ POM structure in order to determine how to run computational calculations for these massive frameworks in the most efficient manner possible. Through a rigorous benchmarking process, we discovered that 8 potassium cations, located within the pore, provided us with the most accurate model in terms of mimicking empirical properties to a sufficient degree of accuracy whilst also requiring a relatively small number of computer cores and hours to achieve a successfully converged calculation. Additionally, we analysed two other similar POMs from the literature, $[As_8W_{48}O_{184}]^{40-}$ and $[Se_8W_{48}O_{176}]^{32-}$ respectively, in the hopes of determining whether they could be similarly incorporated into a POMzite network; given their close semblance in terms of local electron densities and interaction with potassium cations, we judge these POMs to be theoretically suitable as POMzite building blocks. We also experimented with substituting different cations into the $[P_8W_{48}O_{184}]^{40-}$ pore to observe the effect on pore dimensions and HOMO-LUMO gap; we found that we observed that the monocationic structures, particularly the $Li_8[P_8W_{48}O_{184}]^{32-}$ framework, yielded the least polarised structures. This correlates with the literature[1,2], validating our methodology for determining general POM characteristics and properties moving forwards.[3]

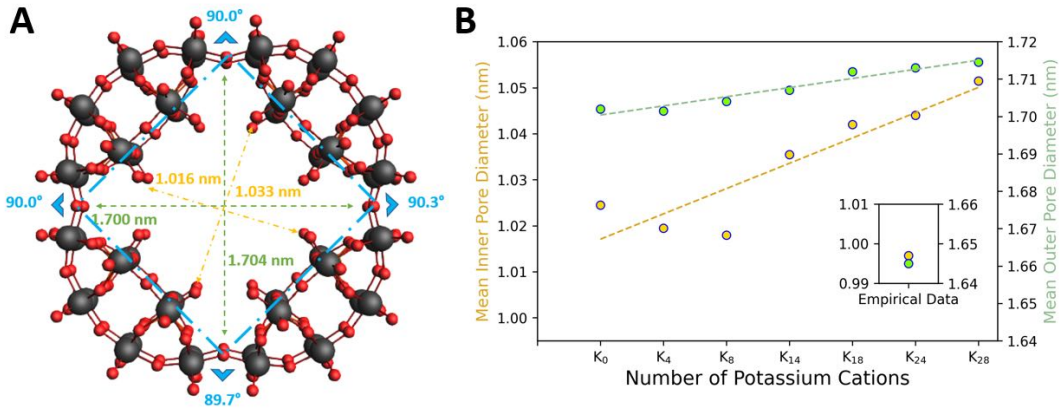


Figure 1. Measurement of angles and diameters within the $[P_8W_{48}O_{184}]^{40-}$(**A**) wheel, with angles (blue), the smaller, 'inner' pore diameter (yellow), and the larger, 'outer', pore diameter (green). (**B**) portrays the trend of the pore diameter to increase as more cations are included in the $\{P_8W_{48}\}$ structure. A subplot is included which shows the experimentally obtained values in relation to those measured from calculations.

[1] Yi, X., Izarova, N.V. and Kögerler, P., *Inorg. Chem.* **2017**, *56*, 13822-13828.
[2] Jiao, Y., Qin, C., Wang, X., Wang, C., Sun, C., Wang, H., Shao, K. and Su, Z., *Chem. – Asian J.* **2014**, *9*, 470-478.
[3] Malcolm D, Vilà-Nadal L. First steps in the inverse design of porous metal oxide materials: lacunary Wells–Dawson anion. *ChemRxiv* **2022**, DOI: