PHARMASEL-PROSPERITY NEWSLETTER

PharmaSEL Prosperity

A bimonthly newsletter

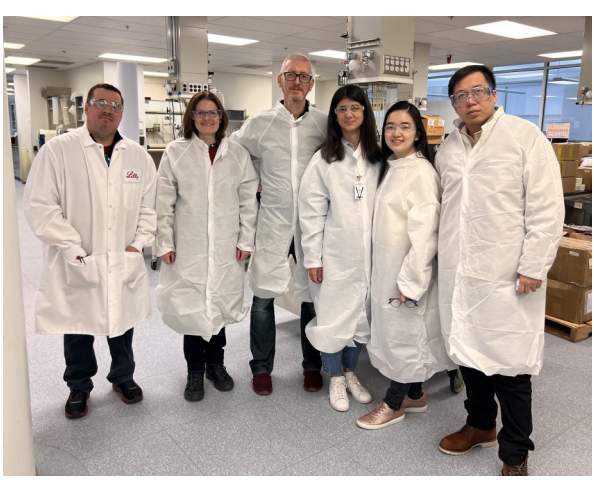
ISSUE 8 JANUARY 2023

In this issue:

- *The UCL research group - WP3 shares their experiences from their visit to Lilly.*
- *Research from UCL group - WP3 on machine learning techniques for HPLC system modelling enhancement.*
- *James Botwright, a PhD student in WP1, shares his background and perspective on the partnership.*
- *A glance at a recent publication on the most kinetically influential reaction pathway driving the generation of HCN from Oxyma/DIC adduct.*
- *A highlight of our recent awards, conferences/posters, and publications.*

Visit to Lilly: UCL group - WP3

Members of the UCL Work Package 3 team visited Lilly facilities in Indianapolis this month, after spending time at FOCAPO (Foundations of Computer-Aided Process Operations) 2023 in San Antonio, where the researchers were able to present their work.



"As part of the visit, we were taken around the laboratories to see the equipment which relates to our modelling efforts, a very useful experience for all of us and a big thank you to Colin Bent and Scott Hartzell for the tour" - Professor Eva Sorensen

"Visiting Lilly in Indianapolis was an incredible opportunity to share knowledge and ideas, and discuss future plans. During the visit, we also had the chance to enjoy a tour of the laboratories. A big thank you to those who made this visit possible" - Dr Monica Tirapelle

"The visit to the lab made us realise the complexity of implementing even only a few, for example, three, columns in series or parallel, which we should keep in mind during modelling and simulation" - Miss Dian Ning Chia & Mr Fanyi Duanmu

Research Feature: Machine learning techniques for HPLC system modelling enhancement (Alberto Marchetto, WP3)

Methodology

The methodology implemented so far allows for the computation of solute retention factors as a function of their molecular properties and the mobile phase composition in isocratic HPLC methods. The methodology is based on the following three steps. Firstly, quantitative-structure property relationships (QSPRs) based on ridge regression predict the solute parameters of linear solvation-energy relationships (LSERs) for small molecules of interest. The QSPR built use molecular descriptors as model inputs, and were calibrated and validated with experimental values of LSER solute parameters of small molecules taken from a public database. LSERs equipped with the predicted solute parameters are then used to predict the solvent strength coefficient (S_S), and the natural logarithm of the retention factor with pure water as the mobile phase ($logk_w$), for the small molecules of interest. Applying linear solvent strength (LSS) theory, the S_S and K_w yield the desired small molecule retention factors k. An applicability domain (AD) analysis based on the Mahalanobis distance between molecules in the LSER solute parameter space is also proposed, which allows to reduce the risk of extrapolating the two LSER.

Table 1. *Predictive performances of the QSPR based on a ridge regression used to predict the four LSER solute parameters E, S, A and B.*

Error metrics	*E*	*S*	*A*	*B*
RMSEP*	0.10	0.22	0.10	0.10
R^2 §	0.98	0.89	0.91	0.95

* RMSEP: root mean squared error of prediction
§ R^2: coefficient of determination

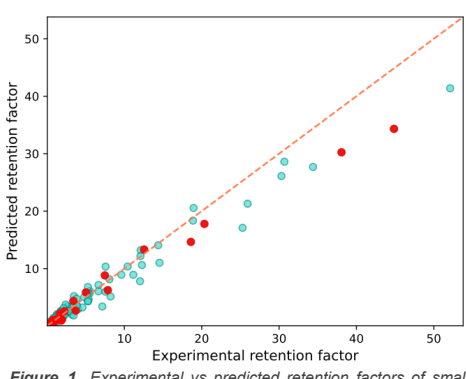


Figure 1. *Experimental vs predicted retention factors of small molecules not used for QSPR and LSER calibration, at different mobile phase compositions. In red, molecules found outside the applicability domain of the LSER.*

Case Study

Predictive performances of the QSPR for the 1,281 molecules in the QSPR validation data set (i.e., molecules not used for QSPR training) are reported in Table 1 for the four LSER solute parameters predicted, i.e. E, S, A and B (the fifth solute parameter usually appearing in LSER, V, can be easily calculated from solute molecular structure). Satisfactory performances are obtained for most LSER solute parameters. In particular, the root mean squared error of prediction (RMSEP) and the coefficient of determination R^2 obtained for E, A and B are not far away from what had been obtained with much more complex ML algorithms[2]. Figure 1 shows experimental *vs* predicted retention factors (k_{exp} vs k_{pred}) for small molecules not used for QSPR and LSER calibration, at different mobile phase compositions. Retention factor predictions were more accurate for low values of k_{exp} (high organic modifier fractions), whereas predictions tend to be less accurate for molecules with high k_{exp} values (low organic modifier fractions). Including all the molecules, a median relative error (MedRE) of 19% was found, while excluding the molecules found outside the AD (red dots in Figure 1), the MedRE decreased down to 16%. Considering that predictions were done starting from information about the molecular structures only, and that QSPR, LSER and the LSS theory had been applied, the initial results prove the potential of such a data driven approach.

Relevance to Lilly and Future Work

The methodology is of interest for the PharmaSEL-Prosperity partnership, for example, to provide a good starting point for HPLC method development. Indeed, it was proven capable of predicting small molecule retention factors in analytical HPLC systems without any prior experimentation. Overall, the methodology will allow for minimization of costs and solvent consumption, and significantly speed up analytical HPLC method development. Future work will advance the methodology by improving QSPR predictions, e.g., by considering more complex QSPR accounting for non-linear relationships between molecular descriptors and LSER solute parameters, and by applying alternative techniques to assess the applicability domain of LSER.

Alberto Marchetto
PhD Student
Department of Chemical Engineering / University College London
alberto.marchetto.21@ucl.ac.uk

Researcher Spotlight: James Botwright, WP1

James completed his masters at Durham University in 2016 with a project supervised by Dr Elizabeth Grayson investigating the potential use of fluorinated protecting groups as NMR markers in the synthesis of oligosaccharides. He then worked for topical formulation development company MedPharm as a formulation developer and analyst, developing HPLC methods for the analysis of APIs and determining their stability in various excipients. In 2018, James moved on to ATDBio, an oligonucleotide synthesis company, doing preparative HPLC and analysis of oligos. During the pandemic, James spent lockdowns in the lab as ATDBio produced primers and probes for the detection of COVID-19 by PCR.

Desiring more freedom and greater challenge than was to be found in preparative and analytical HPLC, James decided to return to academia. He started his PhD at Queen Mary University of London in September 2021 and is now investigating the use of organic solvent nanofiltration membranes in liquid phase peptide synthesis. This research has been ongoing within the Livingston group for a number of years, with both James and predecessor Jet Yeo of Imperial College London being involved in WP1 of the PharmaSEL Prosperity partnership. Liquid phase peptide synthesis has already demonstrated to be an attractive alternative to solid phase peptide synthesis on a large scale with companies such as Ajinomoto and Jitsubo offering industrial synthesis of peptides by their own liquid phase methods. The use of nanofiltration membranes has the potential to make the liquid phase synthesis more convenient due to its automatability.

James has worked on the synthesis of oligosaccharides, oligonucleotides and oligopeptides and upon completion of his PhD will be looking for somewhere to cash in his biological oligomer bingo card.

James Botwright
PhD Student
Queen Mary University of London
j.a.s.botwright@qmul.ac.uk

Featured Publication

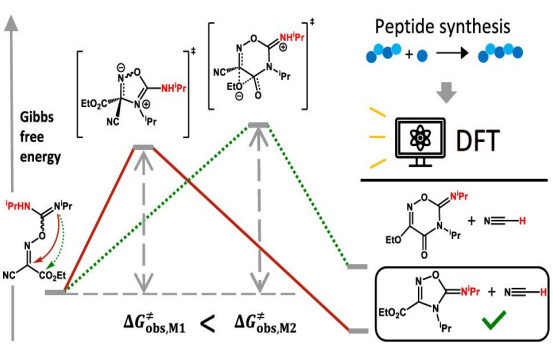
Uncovering the Most Kinetically Influential Reaction Pathway Driving the Generation of HCN from Oxyma/DIC Adduct: A Theoretical Study

Lingfeng Gui, Claire S. Adjiman, Amparo Galindo, Fareed Bhasha Sayyed, Stanley P. Kolis, and Alan Armstrong

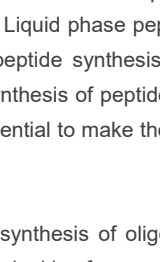
DOI: 10.1021/acs.iecr.2c03145

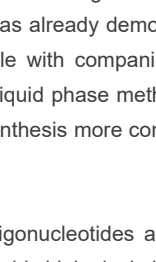
Abstract

The combination of ethyl (hydroxyimino) cyanoacetate (Oxyma) and diisopropylcarbodiimide (DIC) has demonstrated superior performance in amino acid activation for peptide synthesis. However, it was recently reported that Oxyma and DIC could react to generate undesired hydrogen cyanide (HCN) at 20 °C, raising safety concerns for the practical use of this activation strategy. To help minimize the risks, there is a need for a comprehensive investigation of the mechanism and kinetics of the generation of HCN. Here we show the results of the first systematic computational study of the underpinning mechanism, including comparisons with experimental data. Two pathways for the decomposition of the Oxyma/DIC adduct are derived to account for the generation of HCN and its accompanying cyclic product. These two mechanisms differ in the electrophilic carbon atom attacked by the nucleophilic sp2-nitrogen in the cyclization step and in the cyclic product generated. On the basis of computed "observed" activation energies, ΔG_{obs}, the mechanism that proceeds via the attack of the sp2-nitrogen at the oxime carbon is identified as the most kinetically favorable one, a conclusion that is supported by closer agreement between predicted and experimental 13C NMR data. These results can provide a theoretical basis to develop a design strategy for suppressing HCN generation when using Oxyma/DIC for amino acid activation.



Awards





Dr Othman Almusaimi **&** ***Prof Daryl R. Williams***

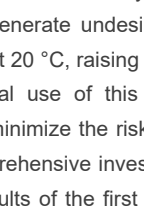
Secured funding from the Institute of Deep Tech Entrepreneurship (DT Prime) / Imperial College London, to develop new technology to deploy novel classes of peptides.

Upcoming Conferences / Posters

1. **O. Almusaimi, S. V. Morse, L. Lombardi, D. R. Williams**. Successful synthesis of glial-specific blood-brain barrier shuttle peptide following fragment condensation approach on solid-phase resin. ACS Spring 2023, Crossroads of Chemistry, Indianapolis, IN, 26th to 30th March 2023.
2. **L. Lombardi, Y. Shi, A. Falanga, H. Azevedo, S. Galdiero**. Enhancing the potency of antimicrobial peptides through molecular engineering and self-assembly. ACS Spring 2023, Crossroads of Chemistry, Indianapolis, IN, 26th to 30th March 2023.

Publications

1. Othman Al Musaimi, Varshitha Gavva, and D.R. Williams. *Greener Cleavage of Protected Peptide Fragments from Sieber Amide Resin*. *ChemistryOpen*. **2022**. 10.1002/open.202200236.
2. Lingfeng Gui, Claire S. Adjiman, Amparo Galindo, Fareed Bhasha Sayyed, Stanley P. Kolis, and Alan Armstrong. *Uncovering the Most Kinetically Influential Reaction Pathway Driving the Generation of HCN from Oxyma/DIC Adduct: A Theoretical Study*. *Ind. Eng. Chem. Res.* **2023**.10.1021/acs.iecr.2c03145.
3. Othman Al Musaimi, Oscar M. Mercado-Valenzo, and D.R. Williams. Factors Influencing the Prediction Accuracy of Model Peptides in Reversed-Phase Liquid Chromatography. *Separations*. **2023**. 10.3390/separations10020081



Co-Editor-in-Chief
Hamish Mitchell
PhD Student
Imperial College London
hamish.mitchell16@imperial.ac.uk

Co-Editor-in-Chief
Othman Almusaimi
Research Associate
Imperial College London
o.al-musaimi@imperial.ac.uk

Disclaimer: This research program is funded by Eli Lilly and Company, Imperial College of London, University College London and from EPSRC Grant Ref: EP/T518207/1. The contents of this newsletter are chosen and edited by researchers and students of Imperial College London. Their editorial work is independent of Eli Lilly and Company and EPSRC.

Imperial College London · UCL · Queen Mary University of London · Lilly · UKRI Engineering and Physical Sciences Research Council · Royal Academy of Engineering