# AIMMS publication report for: 2022-05-18

1. Efficient Computation of Sequence Mappability (2022-5)
2. A Pathway-Based Genetic Score for Oxidative Stress: An Indic (2022-4)
3. Metabolic Conversions by Lactic Acid Bacteria during Plant P (2022-4)
4. Structural Properties of Phenylalanine-Based Dimers Revealed (2022-4)
5. Experiences From Developing Software for Large X-Ray Crystal (2022-4)
6. New potential candidates for astronomical searches discovere (2022-4)
7. Improving Results by Improving Densities: Density-Corrected (2022-4)
8. Using a *Caenorhabditis elegans* Parkinson’s Dise (2022-4)
9. C(spn)−X (n=1–3) Bond Activation by Palladium (2022-5)
10. *Hydrophilic interaction chromatography – mass spectrometry f (2022-4)*
11. *Optical control of Class A G protein-coupled receptors with (2022-4)*
12. *Workflows for optimization of enzyme cascades and whole cell (2022-4)*
13. *Chronic Release of Tailless Phage Particles from Lactococcus (2022-1)*
14. *Polycyclic Aromatic Hydrocarbons (PAHs) in Interstellar Ices (2022-3)*
15. *Puckering the Planar Landscape of Fragments: Design and Synt (2022-3)*
16. *Single-cell imaging of ERK and Akt activation dynamics and h (2022-3)*
17. *Exploring the statically screened G3W2 correction to the GW (2022-3)*

### 1) Efficient Computation of Sequence Mappability

* Charalampopoulos, P., Iliopoulos, C. S., Kociumaka, T., Pissis, S. P., Radoszewski, J., Straszyński, J.
* Bioinformatics, AIMMS, Bio Informatics (IBIVU), Reichman University, King's College London, University of California at Berkeley, University of Warsaw
* Algorithmica
* https://doi.org/10.1007/s00453-022-00934-y
* Corresponding author: Straszyński, J.
* Published May 2022 (early online 2 Feb 2022)
* Processed: 2022-5

Sequence mappability is an important task in genome resequencing. In the (k,m)-mappability problem, for a given sequence T of length n, the goal is to compute a table whose ith entry is the number of indices j≠ i such that the length-m substrings of T starting at positions i and j have at most k mismatches. Previous works on this problem focused on heuristics computing a rough approximation of the result or on the case of k= 1. We present several efficient algorithms for the general case of the problem. Our main result is an algorithm that, for k= O(1) , works in O(n) space and, with high probability, in O(n· min { mk, log kn}) time. Our algorithm requires a careful adaptation of the k-errata trees of Cole et al.[STOC 2004] to avoid multiple counting of pairs of substrings. Our technique can also be applied to solve the all-pairs Hamming distance problem introduced by Crochemore et al.[WABI 2017]. We further develop O(n2) -time algorithms to compute all (k,m)-mappability tables for a fixed m and all k∈ { 0 , … , m} or a fixed k and all m∈ { k, … , n}. Finally, we show that, for k, m= Θ (log n) , the (k,m)-mappability problem cannot be solved in strongly subquadratic time unless the Strong Exponential TimeHypothesisfails. This is an improved and extended version of a paper presented at SPIRE 2018.

### 2) A Pathway-Based Genetic Score for Oxidative Stress: An Indicator of Host Vulnerability to Phthalate-Associated Adverse Neurodevelopment

* Tanner, S., Thomson, S., Drummond, K., O’hely, M., Symeonides, C., Mansell, T., Saffery, R., Sly, P. D., Collier, F., Burgner, D., Sugeng, E. J., Dwyer, T., Vuillermin, P., Ponsonby, A. L.
* E&H: Environmental Health and Toxicology, AIMMS, Florey Institute of Neuroscience and Mental Health, University of Melbourne, Deakin University, The Minderoo Foundation, University of Queensland, WHO Collaborating Centre for Children's Health and Environment, Geelong Hospital, University of Oxford
* Antioxidants
* https://doi.org/10.3390/antiox11040659
* Corresponding author: Ponsonby, A. L.
* Published Apr 2022 (early online 29 Mar 2022)
* Processed: 2022-4

The developing brain is highly sensitive to environmental disturbances, and adverse exposures can act through oxidative stress. Given that oxidative stress susceptibility is determined partly by genetics, multiple studies have employed genetic scores to explore the role of oxidative stress in human disease. However, traditional approaches to genetic score construction face a range of challenges, including a lack of interpretability, bias towards the disease outcome, and often overfitting to the study they were derived on. Here, we develop an alternative strategy by first generating a genetic pathway function score for oxidative stress (gPFSox ) based on the transcriptional activity levels of the oxidative stress response pathway in brain and other tissue types. Then, in the Barwon Infant Study (BIS), a population-based birth cohort (n = 1074), we show that a high gPFSox, indicating reduced ability to counter oxidative stress, is linked to higher autism spectrum disorder risk and higher parent-reported autistic traits at age 4 years, with AOR values (per 2 additional pro-oxidant alleles) of 2.10 (95% CI (1.12, 4.11); p = 0.024) and 1.42 (95% CI (1.02, 2.01); p = 0.041), respectively. Past work in BIS has reported higher prenatal phthalate exposure at 36 weeks of gestation associated with offspring autism spectrum disorder. In this study, we examine combined effects and show a consistent pattern of increased neurodevelopmental problems for individuals with both a high gPFSox and high prenatal phthalate exposure across a range of outcomes, including high gPFSox and high DEHP levels against autism spectrum disorder (attributable proportion due to interaction 0.89; 95% CI (0.62, 1.16); p < 0.0001). The results highlight the utility of this novel functional genetic score and add to the growing evidence implicating gestational phthalate exposure in adverse neurodevelopment.

### 3) Metabolic Conversions by Lactic Acid Bacteria during Plant Protein Fermentations

* Engels, W., Siu, J., van Schalkwijk, S., Wesselink, W., Jacobs, S., Bachmann, H.
* Systems Bioinformatics, AIMMS, NIZO food research, Daiya Foods Inc.
* Foods
* https://doi.org/10.3390/foods11071005
* Corresponding author: Engels, W.
* Published 1 Apr 2022 (early online 29 Mar 2022)
* Processed: 2022-4

To secure a sustainable food supply for the rapidly growing global population, great efforts towards a plant-based diet are underway. However, the use of plant proteins comes with several challenges, such as improvement or removal of undesired flavours, and generation of desired texture properties. Fermentation holds large potential to alter these properties, but compared to dairy fermentations, our knowledge on strain properties in different plant-based substrates is still limited. Here, we explored different lactic acid bacteria for their ability to grow, produce flavour compounds, or remove off-flavour compounds from different plant proteins. For this, 151 LAB strains from dairy and non-dairy origins were cultured in plant protein plus coconut oil emulsions supplemented with glucose. Pea, chickpea, mung, fava, and soybean proteins were used in the study and bacterial strains for screening included the genera Streptococcus, Lactococcus, Lactobacillus, and Leuconostoc. Efficient, high throughput, screening on plant proteins was developed and strains were assessed for their ability to (i) acidify and decrease the pH; (ii) express key enzymes involved in the formation of amino acid derived flavours, which included PepN (aminopeptidase N), PepXP (X-prolyl dipeptidyl peptidase), EstA (esterase), BcAT (branched chain aminotransferase), CBL (cystathione beta lyase), and ArAT (aromatic aminotransferase); and (iii) improve the overall aroma profile by generating dairy/cheesy notes and decreasing off flavours. Suitable screening conditions were determined, and highlighted the importance that a sufficient heat treatment must be applied to samples containing plant proteins, prior to fermentation, as an outgrowth of spore forming Bacillus cereus was observed if the material was only pasteurised. Enzyme activities for strains measured in rich broth vs. a buffered protein solution showed little-to-no correlation, which illustrated the importance of screening conditions to obtain predictive enzyme measurements. Aroma formation analysis allowed to identify strains that were able to increase key aromas such as diacetyl, acetoin, 2-and 3-methyl butanol, and 2,3pentanedione, as well as decrease the off-flavours hexanal, pentanal, and nonanal. Our findings illustrate the importance of strain specific differences in the assessed functionalities and how a methodical approach to screening LAB can be applied to select suitable microorganisms that show promise in fermentation of plant proteins when applied in non-dairy cheese applications.

### 4) Structural Properties of Phenylalanine-Based Dimers Revealed Using IR Action Spectroscopy

* Stroganova, I., Bakels, S., Rijs, A. M.
* BioAnalytical Chemistry, AIMMS
* Molecules (Basel, Switzerland)
* https://doi.org/10.3390/molecules27072367
* Corresponding author: None
* Published 6 Apr 2022 (early online None)
* Processed: 2022-4

Peptide segments with phenylalanine residues are commonly found in proteins that are related to neurodegenerative diseases. However, the self-assembly of phenylalanine-based peptides can be also functional. Peptides containing phenylalanine residues with different side caps, composition, and chemical alteration can form different types of nanostructures that find many applications in technology and medicine. Various studies have been performed in order to explain the remarkable stability of the resulting nanostructures. Here, we study the early stages of self-assembly of two phenylalanine derived peptides in the gas phase using IR action spectroscopy. Our focus lies on the identification of the key intra- and intermolecular interactions that govern the formation of the dimers. The far-IR region allowed us to distinguish between structural families and to assign the 2-(2-amino-2-phenylacetamido)-2-phenylacetic acid (PhgPhg) dimer to a very symmetric structure with two intermolecular hydrogen bonds and its aromatic rings folded away from the backbone. By comparison with the phenylalanine-based peptide cyclic L-phenylalanyl-L-phenylalanine (cyclo-FF), we found that the linear FF dimer likely adopts a less ordered structure. However, when one more phenylalanine residue is added (FFF), a more structurally organized dimer is formed with several intermolecular hydrogen bonds.

### 5) Experiences From Developing Software for Large X-Ray Crystallography-Driven Protein-Ligand Studies

* Pearce, N. M., Skyner, R., Krojer, T.
* Organic Chemistry, AIMMS, The Oxford Science Park, Lund University
* Frontiers in Molecular Biosciences
* https://doi.org/10.3389/fmolb.2022.861491
* Corresponding author: Pearce, N. M.
* Published 11 Apr 2022 (early online None)
* Processed: 2022-4

The throughput of macromolecular X-ray crystallography experiments has surged over the last decade. This remarkable gain in efficiency has been facilitated by increases in the availability of high-intensity X-ray beams, (ultra)fast detectors and high degrees of automation. These developments have in turn spurred the development of several dedicated centers for crystal-based fragment screening which enable the preparation and collection of hundreds of single-crystal diffraction datasets per day. Crystal structures of target proteins in complex with small-molecule ligands are of immense importance for structure-based drug design (SBDD) and their rapid turnover is a prerequisite for accelerated development cycles. While the experimental part of the process is well defined and has by now been established at several synchrotron sites, it is noticeable that software and algorithmic aspects have received far less attention, as well as the implications of new methodologies on established paradigms for structure determination, analysis, and visualization. We will review three key areas of development of large-scale protein-ligand studies. First, we will look into new software developments for batch data processing, followed by a discussion of the methodological changes in the analysis, modeling, refinement and deposition of structures for SBDD, and the changes in mindset that these new methods require, both on the side of depositors and users of macromolecular models. Finally, we will highlight key new developments for the presentation and analysis of the collections of structures that these experiments produce, and provide an outlook for future developments.

### 6) New potential candidates for astronomical searches discovered in the electrical discharge of the PAH naphthalene and acetonitrile

* Loru, D., Steber, A. L., Thunnissen, J. M., Rap, D. B., Lemmens, A. K., Rijs, A. M., Schnell, M.
* BioAnalytical Chemistry, AIMMS, German Electron Synchrotron, Radboud University Nijmegen, University of Amsterdam, Kiel University
* Journal of Molecular Spectroscopy
* https://doi.org/10.1016/j.jms.2022.111629
* Corresponding author: Schnell, M.
* Published 19 Apr 2022 (early online None)
* Processed: 2022-4

The formation and dissociation mechanisms of polycyclic aromatic hydrocarbons (PAHs) as well as their reactivity with other interstellar molecules are elusive. In this work, we have investigated the electrical discharge chemistry of the PAH naphthalene and acetonitrile, a molecule known to be present in interstellar environments, using a combination of mass-selective IR-UV ion dip spectroscopy with the free electron laser FELIX in the mid-IR frequency region (550 – 1800 cm−1) and quantum chemical calculations. In addition to the species known to be produced in the electrical discharge of pure naphthalene, –CH3 and –CN substituted unsaturated hydrocarbons have been identified. Most of them, in particular those containing a nitrogen atom in the molecular framework, such as 7H-benzo[7]annulenecarbonitrile, have a substantial dipole moment and, therefore, can be considered as potential candidates for radio astronomical searches. Among the species observed, the two isomers 1- and 2-cyanonaphthalene, which have been recently detected in the TMC-1, have been identified in our experiment, thus continuing to highlight the use of electrical discharge sources as a valuable tool to produce astronomically relevant species.

### 7) Improving Results by Improving Densities: Density-Corrected Density Functional Theory

* Sim, E., Song, S., Vuckovic, S., Burke, K.
* Theoretical Chemistry, AIMMS, Yonsei University, University of California at Irvine
* Journal of the American Chemical Society
* https://doi.org/10.1021/jacs.1c11506
* Corresponding author: Sim, E.
* Published 20 Apr 2022 (early online 5 Apr 2022)
* Processed: 2022-4

Density functional theory (DFT) calculations have become widespread in both chemistry and materials, because they usually provide useful accuracy at much lower computational cost than wavefunction-based methods. All practical DFT calculations require an approximation to the unknown exchange-correlation energy, which is then used self-consistently in the Kohn-Sham scheme to produce an approximate energy from an approximate density. Density-corrected DFT is simply the study of the relative contributions to the total energy error. In the vast majority of DFT calculations, the error due to the approximate density is negligible. But with certain classes of functionals applied to certain classes of problems, the density error is sufficiently large as to contribute to the energy noticeably, and its removal leads to much better results. These problems include reaction barriers, torsional barriers involving I -conjugation, halogen bonds, radicals and anions, most stretched bonds, etc. In all such cases, use of a more accurate density significantly improves performance, and often the simple expedient of using the Hartree-Fock density is enough. This Perspective explains what DC-DFT is, where it is likely to improve results, and how DC-DFT can produce more accurate functionals. We also outline challenges and prospects for the field.

### 8) Using a <i>Caenorhabditis elegans</i> Parkinson&rsquo;s Disease Model to Assess Disease Progression and Therapy Efficiency

* Hughes, S., Dop, M. V., Kolsters, N., Klashorst, D. V. D., Pogosova, A., Rijs, A.
* E&H: Environmental Health and Toxicology, AIMMS, AMS - Ageing & Vitality, AMS - Musculoskeletal Health, BioAnalytical Chemistry
* Pharmaceuticals
* https://doi.org/10.3390/ph15050512
* Corresponding author: None
* Published 22 Apr 2022 (early online None)
* Processed: 2022-4

### 9) C(spn)−X (n=1–3) Bond Activation by Palladium

* Hansen, T., Sun, X., Dalla Tiezza, M., van Zeist, W. J., Poater, J., Hamlin, T. A., Bickelhaupt, F. M.
* Chemistry and Pharmaceutical Sciences, Theoretical Chemistry, AIMMS, Vrije Universiteit Amsterdam
* Chemistry - A European Journal
* https://doi.org/10.1002/chem.202103953
* Corresponding author: Hamlin, T. A.
* Published 6 May 2022 (early online 27 Dec 2021)
* Processed: 2022-5

We have studied the palladium-mediated activation of C(spn)−X bonds (n = 1–3 and X = H, CH3, Cl) in archetypal model substrates H3C−CH2−X, H2C=CH−X and HC≡C−X by catalysts PdLn with Ln = no ligand, Cl−, and (PH3)2, using relativistic density functional theory at ZORA-BLYP/TZ2P. The oxidative addition barrier decreases along this series, even though the strength of the bonds increases going from C(sp3)−X, to C(sp2)−X, to C(sp)−X. Activation strain and matching energy decomposition analyses reveal that the decreased oxidative addition barrier going from sp3, to sp2, to sp, originates from a reduction in the destabilizing steric (Pauli) repulsion between catalyst and substrate. This is the direct consequence of the decreasing coordination number of the carbon atom in C(spn)−X, which goes from four, to three, to two along this series. The associated net stabilization of the catalyst–substrate interaction dominates the trend in strain energy which indeed becomes more destabilizing along this same series as the bond becomes stronger from C(sp3)−X to C(sp)−X.

### *10) Hydrophilic interaction chromatography – mass spectrometry for metabolomics and proteomics: state-of-the-art and current trends*

* Kohler, I., Verhoeven, M., Haselberg, R., Gargano, A. F.
* BioAnalytical Chemistry, AIMMS, Vrije Universiteit Amsterdam, University of Amsterdam, Center for Analytical Sciences Amsterdam, Vrije Universiteit Amsterdam, Amsterdam
* Microchemical Journal
* https://doi.org/10.1016/j.microc.2021.106986
* Corresponding author: Kohler, I.
* Published Apr 2022 (early online 15 Nov 2021)
* Processed: 2022-4

Among all the –omics approaches, proteomics and metabolomics have received increased attention over the last decade. Both approaches have reached a certain level of maturity, showing their relevance i ...

### *11) Optical control of Class A G protein-coupled receptors with photoswitchable ligands*

* Wijtmans, M., Josimovic, I., Vischer, H. F., Leurs, R.
* Medicinal chemistry, AIMMS
* Current Opinion in Pharmacology
* https://doi.org/10.1016/j.coph.2022.102192
* Corresponding author: Leurs, R.
* Published Apr 2022 (early online None)
* Processed: 2022-4

The field of photopharmacology of Class A GPCR ligands has recently attracted attention. In this review we analyze 31 papers on currently available photoswitchable ligands for Class A GPCRs. Using the ...

### *12) Workflows for optimization of enzyme cascades and whole cell catalysis based on enzyme kinetic characterization and pathway modelling*

* Kuschmierz, L., Shen, L., Bräsen, C., Snoep, J., Siebers, B.
* Molecular Cell Physiology, AIMMS, University of Duisburg-Essen
* Current Opinion in Biotechnology
* https://doi.org/10.1016/j.copbio.2021.10.020
* Corresponding author: None
* Published Apr 2022 (early online 15 Nov 2021)
* Processed: 2022-4

To move towards a circular bioeconomy, sustainable strategies for the utilization of renewable, non-food biomass wastes such as lignocellulose, are needed. To this end, an efficient bioconversion of D ...

### *13) Chronic Release of Tailless Phage Particles from Lactococcus lactis*

* Liu, Y., Alexeeva, S., Bachmann, H., Martínez, J. A. G., Yeremenko, N., Abee, T., Smida, E. J.
* Systems Bioinformatics, AIMMS, Wageningen University & Research, TI Food and Nutrition, University of Amsterdam
* Applied and Environmental Microbiology
* https://doi.org/10.1128/AEM.01483-21
* Corresponding author: Smida, E. J.
* Published Jan 2022 (early online 11 Jan 2022)
* Processed: 2022-1

Lactococcus lactis strains residing in the microbial community of a complex dairy starter culture named "Ur"are hosts to prophages belonging to the family Siphoviridae. L. lactis strains (TIFN1 to TIF ...

### *14) Polycyclic Aromatic Hydrocarbons (PAHs) in Interstellar Ices: A Computational Study into How the Ice Matrix Influences the Ionic State of PAH Photoproducts*

* Ten Brinck, S., Nieuwland, C., Van Der Werf, A., Veenboer, R. M., Linnartz, H., Bickelhaupt, F. M., Fonseca Guerra, C.
* Theoretical Chemistry, AIMMS, Vrije Universiteit Amsterdam, Leiden University, Radboud University Nijmegen
* ACS Earth and Space Chemistry
* https://doi.org/10.1021/acsearthspacechem.1c00433
* Corresponding author: Bickelhaupt, F. M.
* Published 17 Mar 2022 (early online 21 Feb 2022)
* Processed: 2022-3

Polycyclic Aromatic Hydrocarbons (PAHs) in Interstellar Ices:A Computational Study into How the Ice Matrix Influences the Ionic State of PAH PhotoproductsIt has been experimentally observed that water ...

### *15) Puckering the Planar Landscape of Fragments: Design and Synthesis of a 3D Cyclobutane Fragment Library*

* Hamilton, D. J., Beemsterboer, M., Carter, C. M., Elsayed, J., Huiberts, R. E., Klein, H. F., O'Brien, P., de Esch, I. J., Wijtmans, M.
* Medicinal chemistry, AIMMS, Chemistry and Pharmaceutical Sciences, Vrije Universiteit Amsterdam, University of York
* ChemMedChem
* https://doi.org/10.1002/cmdc.202200113
* Corresponding author: Wijtmans, M.
* Published 11 Mar 2022 (early online None)
* Processed: 2022-3

Fragment-based drug discovery (FBDD) has a growing need for unique screening libraries. The cyclobutane moiety was identified as an underrepresented yet attractive three-dimensional (3D) scaffold. Syn ...

### *16) Single-cell imaging of ERK and Akt activation dynamics and heterogeneity induced by G-protein-coupled receptors*

* Chavez-Abiega, S., Gönloh, M. L., Gadella, T. W., Bruggeman, F. J., Goedhart, J.
* Systems Bioinformatics, AIMMS, University of Amsterdam
* Journal of cell science
* https://doi.org/10.1242/jcs.259685
* Corresponding author: Chavez-Abiega, S.
* Published 11 Mar 2022 (early online None)
* Processed: 2022-3

Kinases play key roles in signaling networks that are activated by G-protein-coupled receptors (GPCRs). Kinase activities are generally inferred from cell lysates, hiding cell-to-cell variability. To ...

### *17) Exploring the statically screened G3W2 correction to the GW self-energy: Charged excitations and total energies of finite systems*

* Förster, A., Visscher, L.
* Theoretical Chemistry, AIMMS
* Physical Review B
* https://doi.org/10.1103/PhysRevB.105.125121
* Corresponding author: Förster, A.
* Published 16 Mar 2022 (early online None)
* Processed: 2022-3

Electron correlation in finite and extended systems is often described in an effective single-particle framework within the GW approximation. Here, we use the statically screened second-order exchange ...