Department of Chemistry | Lancaster University

MSc Chemistry by Research

Research Project

Thesis

Project Title: *In silico* toxicity prediction (ISTP) as a means of screening species for materials science and engineering

Author: Sidra Shah, BSc (Hons)

Student Number: 32222192

Supervisor: John Hardy

Affiliation: Department of Chemistry & Materials Science Institute

Lancaster University

ABSTRACT

Sustainability is an issue of growing importance in academia, industry and society. Academic research in safe, sustainable chemistry is therefore the focus of increased attention and resource, and the use of computational tools (e.g., *in silico* toxicity prediction) can facilitate the development of green chemistry approaches to industrially/societally relevant issues. The focus of this thesis is the application of *in silico* toxicity screening to a variety of different substances, including low/high molecular weight organic species, organometallic species, and inorganic species, demonstrating their potential for developing polymer-based materials with technical and medical applications that are "safe by design".

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADME	Absorption, Distribution, Metabolism, and Excretion
ADR's	Adverse Drug Reactions
EC3	Skin sensitisation potency value
ECHA	European Chemicals Agency
EURL ECVAM	European Union Reference Laboratory for alternatives to animal testing
FITC	Fluorescein isothiocyanate
GMPT	Guinea Pig Maximisation Test
ISTP	*In Silico* Toxicity Prediction
LLNA	Local Lymph Node Assay
MSDS	Material Safety Data Sheets
QSAR	Quantitative Structure Activity Relationship
REACH	Registration Evaluation, Authorisation, and restrictions of Chemicals
ROP	Ring Opening Polymerisation
SOHN	Self – organising, Hypothesis Network
SVHC	Substance of Very High Concern

ACKNOWLEDGEMENTS

I would like to express my gratitude and appreciation for my Supervisor Dr John Hardy at Lancaster University for the continuous support, mentorship, and guidance in assisting this research project. The Hardy group, particularly Mark Ashton and James Fong for the technical inputs and advice, and Peter Carrington (Department of Engineering) for insights about optoelectronic materials. Externally Amanda Hughes (University of Liverpool) for insights about printable perovskites for solar cells. Jennifer Garden (University of Edinburgh) for insights about catalysts used to produce polymer-based materials. I am extremely thankful to my family and friends who have shown unconditional support this year, particularly my husband Zulfiqar for his patience, and endless encouragement that has kept me going and made this project possible. Lastly, my mum and babies Adam and Ameer for putting up with mummy over the past few years.

TABLE OF CONTENTS

AIMS & OBJECTIVES

- **Aim**: To assess the effectiveness of the Lhasa suite (specifically, Derek Nexus and Sarah Nexus) in screening compounds important for developing polymer-based materials, and critically assessing their toxicity profiles based upon Lhasa's data and the academic literature.
- **Objective 1:** Examine toxicity profiles of solvents (e.g., low molecular weight organic solvents and ionic liquids).
- **Objective 2:** Examine toxicity profiles of catalytic species (e.g., nanozymes, heterogenous catalysts, organometallic catalysts, photocatalysts functioning via 1 or 2 proton processes).
- **Objective 3:** Examine toxicity profiles of species for upconversion (e.g., bulk solids/crystalline materials, amorphous materials, nanoparticles, and species in solution)
- **Draw informed conclusions**: Collect the data to highlight the relatively untapped potential for *in silico* toxicity screening in polymer materials science and engineering (potentially helping assess issues of global importance such as plastic/electronic waste [via assessment of the environmental/health hazards of microplastics and/or additives for plastics/electronics], issues which the global population is increasingly aware of).

1. INTRODUCTION

1.1 TOXICOLOGY AND COMPUTATIONAL CHEMISTRY

Computational based technologies are fast-growing, predominantly within pharmaceutical and the medical sector, becoming a popular modelling tool in green chemistry and sustainable chemical production. The term '*In Silico*' refers to 'Experimentation performed on computer or via computer simulation'.[1]

Toxicology is the study of chemical safety, measurements and analysis of harmful substances that are contacted with the human body, animals, and the environment. It is a diverse field, importantly so for medical areas and the environmental sector.[1] Toxins could be physical, chemical, or biological. The studies of Paracelsus (1500 AD) raised an awareness into certain chemicals possessing toxic profiles and causing harm to plants and animals, he recognised the body's response to these chemicals was subject to the dose received. These early studies showed that a larger dosage form could potentially be toxic in larger quantities and forms may be proven harmless in smaller doses. Paracelsus was one of the founders of toxicology and the discoverer of dose-response relationship.[2]

1.2 *IN SILICO* TOXICITY MODELLING APPLICATIONS

In silico toxicity screening and artificial intelligence are becoming the driving forces for toxicological studies, these computational methods are revolutionising industries.[3] Earlier detection of high-risk chemicals, gives better insights into the elimination process of hazardous substances. This has given computational methods a proactive outlook into toxicology modelling, examples of which include: quantity-structure-activity relationship (QSAR), in-vitro and in-vivo extrapolation methods, next generation sequencing, high throughput screening, machine learning and endpoint predictions. [3]

In silico toxicity screening predictions (ISTP) is gaining popularity across a range of sectors. Pharmacological, physical, and toxicological properties of compounds are found to be determined through high-throughput screening. This method has proven to predict toxicity from chemical structures, proving its ability to be financially cost effective and time efficient as early assessment of mutagenic compounds can reduce the time in pre-clinical studies and *in vivo* analysis.[4]

In silico toxicology testing is becoming a popular form of study within pharmacology and environmental chemistry within the recent years, it uses computational modelling and screening

of large data sets. Database modelling includes pharmacophore, Quantity Structure Activity Relationship (QSAR). *In silico* toxicity screening is particularly popular within the pharmaceutical industry, for many years different tools have been evaluated to reduce the time and cost as well as ethical concerns around animal testing. These methods have shown an ability to identify toxic compounds and reactions at an earlier stage of drug development.[4]

1.3 REGULATORY

The safety around chemical production and management is a growing concern for the European Chemicals Agency (ECHA), they centralise the Registration Evaluation, Authorisation, and restrictions of Chemicals (REACH) system, the main purpose is to protect human health and the environment, by ensuring safe chemical production and management. Classification labelling and Packaging of substances and mixtures (CLP) and Regulation Organisation for economic cooperation and development (OCED) are included in the decision making bodies in toxicology that benefit from and are investing in *in silico* toxicity screening.[5] Centralised systems are operated for these chemicals to identify hazardous compounds with intrinsic properties, enhance and develop alternatives such as computational toxicity screening for the chemical industry as a regulatory requirement to report and eliminate toxic adverse reactions prematurely. REACH is promoting the use of alternatives to animal testing, assessing hazardous properties of substances with QSAR methods, REACH provides a guidance on chemical safety and the assessment of chemicals using computational methods by a non-testing tiered approach.[3]

1.4 GREEN CHEMISTRY

Green chemistry is a term used since the 1990s to describe the development of chemical products and processes in a way that reduces or eliminates the use of hazardous compounds. Green chemistry applies to a chemical product's entire life cycle from design, manufacture, usage, to disposal. Sustainable chemistry is a term that was introduced more recently, and which the Organisation for Economic Co-operation and Development (OECD) defines as "a scientific concept that seeks to improve the efficiency with which natural resources are used to meet human needs for chemical products and services.".[6]

Together the approaches encourage the development of chemical products and processes using safer practices that reduce hazards and waste production, achieved by the cost-effective design and manufacture of chemical products with minimal/no pollution. Producing chemicals from

feedstocks, reagents, and solvents that are less toxic to human health is key for this approach, lowering risks for environmental and health damage.[7]

One of the *'12 principles in Green Chemistry'* Is to '*use catalysts' not 'stoichiometric reagents'* so the risk of excess toxic waste to the environment is reduced. Catalysts speed up reactions and are effective in small amounts. Making processes high in atom economy and reduced toxicity for chemical synthesis.[7] An example of such catalytic processes involve organometallic catalysts used in Ring Opening Polymerisation (ROP), which have become an alternative to traditional metal catalysts, proving low-cost synthesis and ability to control the polymerisation reaction precisely.[8] This field has emerged rapidly, with a wide range of appropriate catalysts, some of which explored in this thesis.

1.5 ISTP IN DRUG DISCOVERY

In vivo testing is essential as a standard stage of drug discovery, and still compulsory for approval of drugs on the market. In the United States of America, federal regulations around drug candidates and approval of drugs were written and introduced by the Food and Drug Administration (FDA) body in 1906. The aim is to regulate safe medicine and food worldwide.[8] The concerns around animal testing procedures are still deep rooted and rising, the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) regulations are a European initiative that endorse use of computational alternatives where possible. Computational methods have become indispensable, eliminating toxic affects at the premature stage of drug discovery is essential, as these characteristics of target drug molecules are leading to late-stage failures at the clinical testing stage. *In silico* toxicity screening can be an essential tool from the discovery stage where potential new medicinal targets are screened through to characterising and optimising and ensuring safety of selected candidate through to clinical studies in humans.[8]

Characterisation of target pharmaceuticals involves assessing the pharmacokinetic profile, drug molecule absorption, distribution, metabolism, and elimination of target drug compounds. A drug candidate is screened for biological, physical, and toxicological mechanistic properties for therapeutic applications within various species and humans. For analysis of drug molecules efficacy, both *in vivo* and *in vitro* routes of study are significant.[9] *In vivo* studies allow understanding of the effects of drug release within living organisms, and the toxicological properties are examined via *In vitro* analysis.

Impurities are known to be found in synthesis, from the starting materials, reagents and intermediate stages, or by-products. The presence of unwanted impurities cannot be avoided,

however, limiting the toxicity of these substances can be done. The ICH has published a set of guidelines to minimise exposure to toxic drug substances and disclosing potential risks such as carcinogenicity, genotoxicity, and nephrotoxicity. These guidelines are set for definitive risks to human health, positive test data is collected from studies, evaluated, and updated.[10]

Drug failures have mounted in recent years, traditionally bioactivity and toxicity analysis were determined at a later stage of development. Innovative technologies have permitted early screening of toxic structures, allowing early detection and elimination, as this is adding to major costings for pharmaceutical companies. In 1997 Kennedy *et al*,[11] reported that out of the failed new drug entities 16% failed in animal toxicity testing and 14% failed due to adverse effects in humans. It is crucial to discover potential toxicities earlier on as 10-12 years it takes for a drug to reach the market, is associated with significantly higher costs.[11]

Toxicity screening covers a range of adverse drug reactions (ADRs) throughout the drug development process. The use of computational methods facilitates the elimination and structurisation of drug candidates' failure before investing in their synthesis.[4]

1.6 DEREK NEXUS AND SARAH NEXUS SOFTWARE

Computational software Derek Nexus and Sarah Nexus was used in this research project which are components of a knowledge-based software suite that provides instant results and predictions based on the chemical structures screened. Derek Nexus is a modelling software which can be used to predict the toxicity of chemicals in mammals and bacteria. This software uses a knowledge base to match parts of the toxicophores to alerts, apply reasoning to assess the likelihood for a typical prediction and give an overall prediction of toxicity for selected species or parts of the query structure. Currently, the certified Derek knowledge base has 890 alerts and 74 endpoints with 9 parent endpoints.[12]

Sarah Nexus is a statistical software which uses QSAR methodology to predict mutagenicity. Modelling is based on publicly available data and uses custom molecular fragmentation methods and a Self-Organising Hypothesis Network (SOHN) approach; predictions are easily interpretable, and a level of confidence is provided as a value. With a Sarah prediction query compound is entered into the software, the compound is then fragmented, and a hypothesis is identified and supporting examples are retrieved. A confidence value is calculated, considering Ames results and similarity of supporting examples.[12]

1.7 NANOTOXICOLOGY

The rapid demand for and exciting applications of nanotechnology has created a gap in the computational market for toxicity studies around nano-scale particles for nano-medicine and nano-biotechnologies.[13] Nano technologies have increased the regulatory demand for *in silico* studies and reduction of risk and time associated with animal studies, providing support for analysis of large data sets, QSAR, QSPR, molecular dynamics and ADME. These significant advances in non-animal tests are finding a better method to determine cytotoxic hazards. Analysis from predictions for non-mutagens is shown by high accuracy in data screened. [13]

With the advances in nanomedicine, and the recent development of new nano substances in nano systems have become available, nanomaterials for medical diagnosis, prevention, and disease control. Regulatory issues have also arisen, the European union are exploring the current applications and opportunities for nanotechnology to meet medical needs and standardizing the compliance around and nanotechnology.

The application of nanotechnology in medicine has introduced the ability of targeted drug release, increasing localisation of therapeutic dose to intended tissues.[14] Nanoparticle-based drug delivery systems have the effective ability to target specific cells, and organs, so healthy tissues remain clear of adverse toxic effects.

The pharmaceutical advancement of nanomedicine comes with its challenges, during the development process physicochemical characterisation and pharmacokinetics needs must be assessed. Nanotoxicological evaluation is required to eliminate *in vivo* and *vitro* physicochemical properties. Understanding the potential risks of nanomaterials is important for each life cycle of event and characterisation through to various stages of commercialisation.[15]

1.8 MUTAGENICITY

Mutagenicity testing is a crucial tool to predict long term risk to human health, and data acquired acts as a fast and reliable way to determine toxicity. A number of studies have shown a correlation between mutagenicity and carcinogenicity, mutagens are found naturally occurring environmentally or chemically produced both causing cancerous affects.[16] Mutagenicity testing is often a requirement for regulatory validation for chemicals, and mutagenicity is an important toxicological endpoint for medical and pharmaceutical companies.

The AMES test is a widely accepted biological procedure using bacteria to assess a chemical's potential to cause mutations in the DNA of a test organism. It is a biological assay determining the mutagenic ability of chemical compounds. Whilst this method allows identification of mutations present in strains, it is also used to detect the mutagenicity of drugs, reagents, solvents, and other soluble liquid suspensions. Interpretation of the results shows proportionality between mutagenicity and the number of colonies observed on plates.

Cytotoxins are substances that effect cell viability, and genotoxins are those which damage the genome within a cell, causing damage to the DNA. Genotoxic chemicals have the potential to interact with various proteins in mitosis, DNA damage could lead to carcinogenesis or lay the foundations for congenital disorders.[17] Safety is a great concern for these various chemicals causing mutagenic effects, cytotoxic behaviours are to be screened for various categories of compounds in flavours, preservatives, drugs, and solvents. *In vitro* analysis could be conducted as a regulatory requirement to detect adverse long-term effects.

2. EXPERIMENTAL

2.1 MATERIALS AND METHODOLOGY

2.1.1 LHASA LIMITED

The products used in this research are designed and created by Lhasa Limited, a computational software development company. Derek Nexus for managing toxic chemical information, and Sarah Nexus for managing mutagenicity information, both designed to meet Organisation for Economic Co-operation and Development (OECD) guidelines.

2.1.2 DEREK NEXUS

Derek Nexus is an expert knowledge-based toxicology software that gives accurate toxicity predictions for a range of endpoints. Derek Nexus allows the evaluation of potential toxic chemicals; it also allows properties of chemicals to be configured for R&D purposes to redefine molecular structures. Derek Nexus gives toxicity predictions of given chemical query structures inputted, alerts are then fired, and predictions are made, including quantitative EC3 predictions for skin sensitisation, negative predictions for bacterial mutagenicity and skin sensitisation. The likelihood of structures possessing toxic behaviour is measured against a probability scale (certain to impossible).

2.1.3 FEATURES AND BENEFITS

The categories of evidence considered for reasonings are listed below:

1. Endpoint alert
2. Alert
3. Endpoint
4. Species
5. Toxicity data
6. Physiochemical properties
7. Molecular weight

These affect the likelihood of predictive outcome.

List of properties characterised by Derek Nexus software:

- Transpirable predictions

- Customisable reports
- Reducing risk in R&D
- One interface, multiple predictions
- A Derek alert is a Lhasa alert
- Rapid toxicity assessment

An alert is a set of structural features and data attached to a molecule, allowing users to visualise a given effect. As it currently stands, Derek knowledge base has 890 alerts and 74 endpoints and 9 parents endpoints.[12]

2.1.3 ENDPOINTS

Endpoints are designed to give each alert a direct definition or outcome of result, these are toxicological.

- Carcinogenicity
- Mutagenicity
- Genotoxicity
- Skin Sensitisation
- Teratogenicity
- Irritation
- Respiratory Sensitisation
- Reproductive Toxicity

The most common endpoint prediction as a standard are skin sensitisation and mutagenicity.[12]

2.1.4 NEGATIVE PREDICTIONS

Derek Nexus's expert toxicity prediction software generates negative predictions with functionality for endpoints, this is for the Ames Mutagenicity test. Figure 1 shows a workflow model designed by Derek Nexus highlighting the process behind a negative prediction. If a query compound entered does not give any alerts a negative prediction is given with potential outcomes, being inactive, inactive with misclassified features and inactive with unclassified features. Table 1 defined each and explains how to associate this with query compounds.[18]

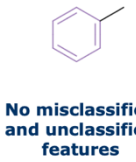
FIGURE 1: NEGATIVE PREDICTIONS IN DEREK NEXUS WORKFLOW [12]

Derek Nexus's workflow is shown above, this is used in the software autonomously to generate predictions. A query compound is entered in a chemically correct way, the Nexus knowledge-based system then allows the compound to be identified and matched to literature from many sources, a positive/negative prediction is then made. An expert review by the user is then conducted, putting the knowledge acquired to the test and making a prediction based on the information presented.[18]

In the absence of an alert, query structures are compared to an external database (Ames test and Skin sensitisation). This is to assess and eliminate other causes for concern. When an alert shows 'No misclassified or unclassified features' it is said to be a highly confident negative prediction. 'contains misclassified features' and 'contains unclassified features' means a slightly lower confidence and some features may be a cause for concern.[12] The presence of misclassified and unclassified features leads to increased variability and reduces the confidence in the negative prediction.

This method for making negative *in vitro* predictions for data sets is a challenging model, a positive endpoint result is given from a structural feature identified to cause a certain effect, thus identified from literature, as a negative prediction is relying on the absence of this affect and relying on features presented in the query which may not have been identified, or predicted incorrectly or may not belong the software.[19] Considering a negative prediction is ambiguous, as there is readily available information available for users with easy access . Negative predictions are taken from a range of textual data which is stored in Derek's knowledge-based system which has taken years to acquire.

TABLE 1: EXPLAINING TYPES OF NEGATIVE PREDICTIONS IN DEREK NEXUS[12]

No misclassified and unclassified features	Misclassified features	Contains unclassified features
		
▪ Query compound has a feature in common with a non-alerting positive compound in the data sets. ▪ A non-alerting positive compound is experimentally found to be positive a particular assay (e.g., Ames test).	▪ These features are those that have been found positive in data sets. ▪ This is not automatically considered as a red flag for toxicity. ▪ Feature may only be present in one positive compound. ▪ Even though feature may be present in one positive compound it may not be promoting toxicity.	▪ Unclassified features are those that have not been found in the data set. ▪ This prediction type query compound has no alerts fired, includes structural fragment that is not covered in respective data sets. ▪ Features not matched with public data set. ▪ Derek has no given reason for this compound to be positive, expert assessment may be required.

2.1.5 EC3 PREDICTIONS FOR SKIN SENSITISATION

Derek's predictions are based on EC3 (effective concentration for a stimulation index of 3) values which depict to skin potency level of alerting compounds. The value generated allows extrapolation of hazard and risk.

Skin sensitisation is a required endpoint in Derek Nexus, current version contains 90 skin sensitisation alerts. This prediction is an important requirement for various chemical assessments, REACH, CLP and EURL EVCVAM (European Union reference Laboratory for

alternatives to animal testing). To extend this current hazard identification to risk a potency prediction would be required.

According to the OECD, there are defined approaches which need to be followed for careful hazard identification and skin sensitisation. The specific guideline which relates to this subject is no. 497 Defined approaches on skin sensitisation. A defined approach (DA) has been identified and the mechanism associated with the study of skin sensitisation is a well-defined Adverse Outcome Pathway (AOP).[20]

For a skin sensitisation alert that is fired, a potency level EC3 prediction is also generated. This value is a quantitative measurement of potency. The Local Lymph Node Assay (LLNA) gives an EC3 value to give a quantitative potency prediction and EC3 values correlates with the human sensitisation induction thresholds. Skin sensitisation has been a common factor to predict hazard potential in chemicals, Derek Nexus uses this endpoint as a standard for structures ended in query section.[12]

In vivo methods used as comparison are Local lymph node assays (LLNA's) and Guinea pig maximisation tests. according to the OECD guidelines, a negative result for skin sensitisation is through *In Silico* Toxicity Screening would be favourable alongside animal testing methods. [21]

TABLE 2: THE EUROPEAN CENTRE FOR ECOTOXICOLOGY AND TOXICOLOGY OF CHEMICALS (ECOTOC) CLASSIFICATION SCHEME ASSIGNED TO CATEGORIES OF CHEMICALS BASED ON THEIR POTENCY (EC3) TO AID RISK ASSESSMENTS. [23]

ECETOC Classification	
EC3 Value	Potency Category
<0.1	Extreme
≥0.1 to <1	Strong
≥1 to <10	Moderate
≥10 to ≥100	Weak

EC3 negative predictions are based on a nearest neighbour model, in which the closest neighbours are selected from a reference group of compounds that trigger the same alert as the query compound. For the closest neighbours, a similarity score is determined and an EC3 prediction is created. Out of the more than 650 compounds in the Lhasa EC3 collection, the closest neighbours are chosen. EC3 values are taken from a broad literature range curated by Lhasa.[23]

FIGURE 2: LIKELIHOOD LEVELS USED BY DEREK NEXUS IN ORDER FROM IMPOSSIBLE TO CERTAIN.[23]

The figure above is demonstrating how Derek measures each alert and gives each a level of likelihood. This framework of likelihood levels is used throughout the reports generated from the software.

TABLE 3: DEREK NEXUS'S DEFINED LEVELS OF LIKELIHOOD WITH THE FOLLOWING DEFINITIONS.[12]

Level of likelihood	Definition
Certain	There is proof the proposition is true
Probable	There is at least one strong argument that the proposition is true. And there are no arguments against it.
Plausible	The weight of evidence supports the proposition
Equivocal	There is an equal weight of evidence for and against
Doubted	The weight of evidence opposes proposition
Improbable	There is at least one strong argument that the proposition is false and there are no arguments that it is true.
Impossible	There is proof that the proposition is false.

2.1.6 SARAH NEXUS

Sarah Nexus is a statistical based software built for the prediction of mutagenicity. ICH M7 guidelines propose that computational methods employed to predict toxicology in particular Ames mutagenicity by using QSAR methodologies, both expert rules based and statistical based. Sarah Nexus is a unique machine learning methodology, employing a custom molecular fragmentation method and SOHN approach. This model contains 11,774 structures which have been standardised and fragmented. This software combines information imputed for each fragment to scientific valid rules, a measure of confidence is given for each prediction, this is directly made available with supporting data and analysis. A level of confidence is given as a value with each prediction, a defined applicability domain.[24]

Data referenced in Sarah Nexus equates to 9882 unique structures (4716 positive/5166 negative). This is publicly available database alongside member donated data. Sarah creates structural units for individual atoms, functional groups, and rings. A reduced graph is created for a molecule and fragments the molecule between structural units according to the specified depth. If a structure contains a particular feature, then it is associated with toxicity for a given endpoint.

Sarah prediction is generated once a query compound is entered, it is then fragmented and matched against structural compounds within the internal database. A hypothesis is identified and generated for the query compound, supporting examples are then retrieved. A confidence value is calculated considering the Ames result and similarity between supporting examples.

Confidence per hypothesis is calculated by Ames test activity of supporting examples, overall confidence is the weighted average of the individual hypothesis. Confidence relates to the accuracy of the prediction by considering the likelihood and the reliability of the model.[24]

TABLE 4: DEFINING THE PREDICTIVE TERMINOLOGY USED IN SARAH NEXUS REPORT.[24]

⟹ Out of domain	Fragments not covered by the training set results with no prediction. Based on presence of unknown fragments. If part of the structure is not covered then the query structure is outside domain.
⟹ Equivocal	In the absence of strong overall signal, equivocal call is made. A strong argument cannot be made based on the training set of compounds, any hypothesis generated for the query compound for either activity or inactivity in bacterial reverse mutation assay (Ames test).
⟹ Positive	The query structure is predicted to be positive in bacterial reverse mutation assay (Ames test).
⟹ Negative	The query structure is predicted to be negative in bacterial revers mutation assay (Ames test).
⟹ Confidence	The overall confidence prediction is determined from the individual hypothesis activated by the inputted structure. These are, in turn, based on the signal and the Tanimoto similarity of the nearest neighbours to the query structure used to build the hypothesis. In the absence of any hypotheses being activated by the query compound, the signal and Tanimoto similarity of the nearest neighbours from the entire training set are used to generate the overall confidence.

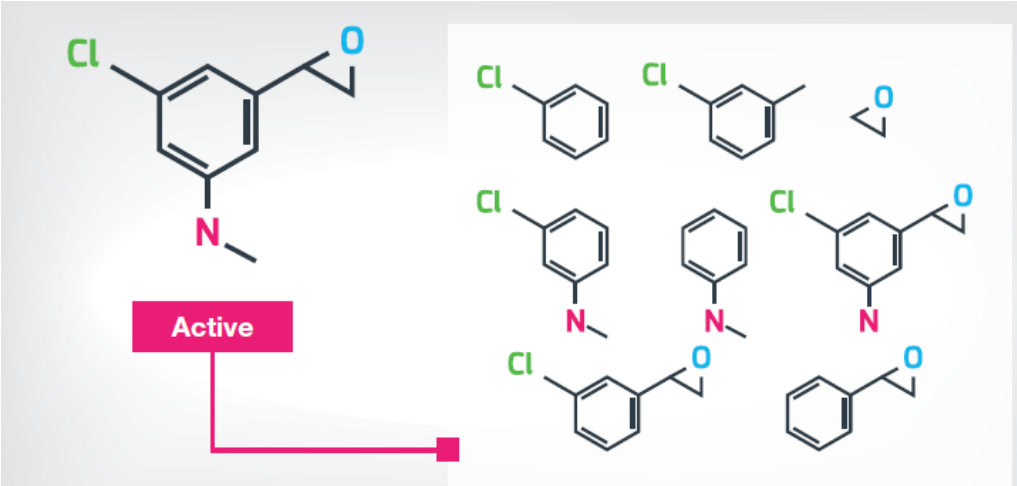


FIGURE 3: SHOWING QUERY COMPOUND SPLIT INTO RELEVANT FIGURES.[23]

Sarah Nexus uses a similar approach to Derek Nexus's knowledge bases system, from entry of query compound, prediction set up, batch processing and reporting. Sarah has the capacity to produce fragments that are found in molecules from the training set, breaking down the compound for effective comparison with pre-determine fragments. It can create a hierarchy of models, with some being more global and others being more local, providing users with the best of both worlds. Sarah will choose the best model for each component and examine the data that is available for each fragment and then combines it according to criteria that are supported by science. A transparent prediction is made by outlining the relative relevance of each local model's contribution and the supporting data. The level of confidence for each prediction is given per fragment, therefore proving all the information for experts to judge and draw conclusions.[25] ICH M7 guidelines for complementary expert rule based and statistical systems for mutagenicity prediction are used and Sarah Nexus is a transparent, well validated statistical system.

2.2 EXPERIMENTAL METHODS

2.2.1 SELECTION OF CHEMICALS

Molecular species were identified and sourced through an online search and selection using PubMed, Web of science for this experimental study. Organometallic catalysts sourced through a collaboration with Jennifer Garden at Edinburgh University; solvents for processing chemicals to generate energy materials were discussed with Amanda Hughes at the University of Liverpool. 7 data sets have been created which have been detailed below, careful selection of similar structures, common functional groups or bonding were chosen. This is to allow characteristic predictions and testing for the applications and possibilities using *in silico* toxicity screening to assess a range of compounds.

TABLE 5: CATEGORIES OF DATASETS AND LIST OF QUERY STRUCTURES WITHIN EACH

Categories of Datasets	**Name**
Solvents popular for polymer dissolution	▪ Acetonitrile ▪ Carbon dioxide (CO_2) ▪ Cyclohexanone ▪ Diethylcarbonate ▪ Dimethylcarbonate ▪ Butyl acrylate ▪ *N*-formylmorphine ▪ Pyridine ▪ Solketal ▪ Water
Ionic Liquids	▪ Bis-2-ethylhexyl-phosphate ▪ Butyl-3-methylimidazolium (BMIM) ▪ Bromide ▪ Hexafluorophosphate ▪ Phosphinate ▪ Tetrabutylammonium ▪ Tetrafluoroborate
Solvents for energy materials	▪ 2-Methoxyethanol (2ME) ▪ 2-aminoethanol (2MEA) ▪ 3-Methyl – 1,2-oxazol-5-ol (3Mox) ▪ Cyrene ▪ Diethylformamide (DEF) ▪ Dimethylformamide (DMF) ▪ Dimethylacetamide (DMAc) ▪ 1,3-Dimethyl-2-imidazolidinone (DMI) ▪ Dimethylpropyleneurea (DMPU) ▪ Dimethyl Sulfoxide (DMSO) ▪ Gamma-Butyrolactone (GBL)

	▪ Hexamethylphosphoramide HMPA) ▪ *N*-methylacetamide (NMAc) ▪ *N*-Methylpyrrolidone (NMP) ▪ Sulfolane ▪ Trimethyl Phosphate (TMP) ▪ Tetramethylurea (TMU)
Nanozyme Catalysts	▪ Copper (Cu) ▪ Copper Phosphate (Cu_3PO_4) ▪ Copper chloride ($CuCl_2$) ▪ Copper Oxide (CuO) ▪ Iron Oxide (Fe_3O_4) ▪ Gadolinium trichloride ($GdCl_3$) ▪ Manganese dioxide (MnO_2) ▪ Molybdenum Sulfide (MoS_2)
Heterogenous Catalysts	▪ Gold ▪ Indium Oxide (In_2O_3) ▪ Iridium Oxide ($Ir0_2$) ▪ Platinum (Pt) ▪ Platinum-Iridium alloy (Pt-Ir) ▪ Tin Oxide (SnO_2) ▪ Titanium Nitride (TiN) ▪ Zeigler Natta
Organometallic catalyst	▪ LZn2OBn ▪ $[LNaZn_2Et_2(THF)_2]$ ▪ $[LKZn_2Et_2(THF)_2]$ ▪ C1 ▪ C1b ▪ C1c ▪ C1d ▪ LZn_2OBn ▪ *rac*-(BDI-1)ZnOCH(Me) ▪ $[Zn_2L\ Et(HMDS)_2]$ ▪ $[Zn_2L\ Et(OiPr)_2]$
Photoinitiators	▪ 2-hydroxy-2-methylpropiophenone - 1PP ▪ 3-6-B3FL - fluorenone - 2PP ▪ 22-dimethoxy-2- phenylacetonephenone - 1PP ▪ AQN - anthraquinone derivatives ▪ BDAB -Bisdiethylaminobenzophenone - 2PP ▪ BSEA – Water Soluble 2PP ▪ Camphorquinone - teeth 3DP ▪ Diphenyl-246- trimethoxybenzoylphosphine oxide - 1PP ▪ E2CK - WaterSoluble 2PP ▪ Irgacure 369 - 2PP ▪ Irgacure 2959 - 2PP ▪ P2CK - WaterSoluble 2P ▪ UC species 1-6.

2.2.2 EXPERIMENTAL DESIGN

Nexus program v.2.2.2 (Build 282, Jul 2018), Derek Nexus v.6.0.1 software system was used as part of the knowledge suite to study various categories of molecular species in this research. Selected chemical structures from their categories were drawn in PerkinElmer ChemDraw 19.1.21 software. Files were saved as .mol and simultaneously imported to Derek for toxicity predictions, identification of toxicophores within query structures. Test reports exported and saved in .pdf format under chemical name. Table 6 demonstrates the parameters defined for the test of each selected compound in Derek Nexus.

TABLE 6: PROCESSING OPTIONS FOR DEREK NEXUS USED TO GENERATE PREDICTIONS IN DEREK NEXUS.[26]

Selected species	**Bacterium, Mammal**
Selected knowledge base	Derek knowledge base 6.0.1
Reasoning level?	At least equivocal
Perceive tautomers?	Yes
Perceive mixtures?	Yes
Match alerts without rules	No
Show open likelihood	No
Show negative prediction	Yes
Show rapid prototypes	Yes
Filter nearest	Yes

2.2.3 CLASSIFICATION OF STRUCTURES USING A TRAFFIC LIGHT SYSTEM

All molecular structures from the various categories have been screened using Derek Nexus and Sarah Nexus, the data acquired from the reports was summarised in tables, which can be found in the appendices. Data was further summarised into a traffic light system which was created to define levels of safety and determine the extent of toxicities for the purpose of this research. The tables with summarised data can be found in the appendix section.

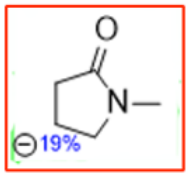
TABLE 7: SUMMARISING THE TRAFFIC LIGHT SYSTEM USED TO CLASSIFY STRUCTURES IN EACH DATA SET

COLOUR	No. OF ALERTS FIRED
Green	0
Amber	≥1
Red	≥2

The traffic light system has been employed to highlight the predicted toxicity of the structures examined:

- Predicted to be safe in green.
- Predicted to be potentially toxic in amber.
- Predicted to be probably toxic in red.

Molecular structures with no alerts fired have been categorised and highlighted in green; structures with 1 or more alerts are categorised and highlighted in amber, and structures with two or more alerts are categorised and highlighted in red. However, it is noteworthy and important to delve into the literature to critically assess the predicted outcomes when the structures are highlighted amber or red.



NMP

FIGURE 4: EXAMPLE OF STRUCTURE DRAWN FOR PRESENTATION PURPOSES IN THIS RESEARCH

As an exemplar structure, NMP has been computationally analysed by Derek Nexus and Sarah Nexus, results obtained from the experimental study have been summarised into tables and then transferred to a traffic light system. Structures have been evaluated and screened multiple times by the software, the results generated have been summarised in tables in the appendices. For simplicity depending on the frequency of alerts each structure has fired, categories have been made and each structure has been placed into a green, amber and red.

The highlighted box represents the level of safety, and -/+ circular sign indicates a positive or negative Sarah result, a prediction percentage will have been stated with the structure. Sarah nexus gives an overall percentage of the predictive outcome, this has been stated at the bottom left corner of the structures highlighted box.

3. Results & Discussion

Historically, the majority of chemicals have been derived from fossil fuel feedstocks, and as there are only finite amounts of fossil fuel reserves left, and as we prepare for a post-fossil fuel world, attention has begun to shift to methods to: reduce the amount of chemicals we use (e.g. via atom efficient chemistry) and concomitantly reduce the waste we create; to reuse chemicals/products, or to recycle or ensure they are safe to dispose of without damaging the environment. A selection of systems are analysed herein to understand the potential utility of ISTP to screen molecules/materials for various technical/medical applications.

3.1 SOLVENTS

Solvents are widely used in synthesis, processing, analysis and delivery of chemicals, and the class of chemicals is too wide to screen holistically, consequently a discrete selection have been analysed here as an example of the utility of ISTP to them. The most widely used "green solvents" in chemistry are water, supercritical carbon dioxide and ionic liquids, although biomass-derived bio-renewable solvents are under investigation.

3.1.1 LOW MOLECULAR WEIGHT SOLVENTS

A small selection of low molecular weight solvents potentially useful in solution phase polymer synthesis/dissolution, results of the analysis have been presented in Figure 5.[27] Noting *N*-butylpyrrolidone (NBP) is a non-toxic substitute for NMP.[28] The EU's REACH legislations require rigorous assessments of high-risk solvents for chemical processes, providing safety to industries, reducing hazardous waste, and considering the recycle of chemicals. Solvents such as *N*-methyl pyrrolidinone (NMP) and dimethylformamide (DMF) are known for causing hazardous effects to human health and the environment, they are now classed as substances of very high concern (SVHC) both are polar aprotic solvents used in dissolution of common polymers – polyvinylidene difluoride, polyurethane, and polyacrylonitrile.[29] Low molecular weight polar solvents are both a significant challenge and a huge potential for the polymer industry. Green replacement possibilities for the polar aprotic class of solvents are currently limited, and research and development of new green solvents tends to focus on solvents for chemical reactions and the pharmaceutical industry rather than polymer dissolution. There is a significant need to research bio-effective, cost-effective solvents, there are green polar aprotic solvent alternatives, this study has used Derek Nexus and Sarah Nexus to find toxicities of alternative examples for polymer dissolution.

3.1.1.1 SOLVENTS STRUCTURES

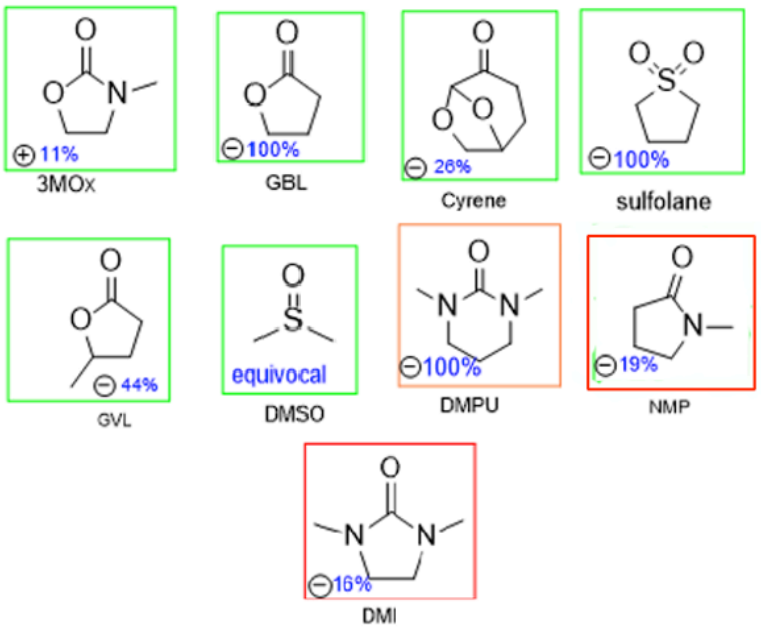


FIGURE 5: LOW MOLECULAR WEIGHT SOLVENTS SCREENED WITH DEREK NEXUS AND SARAH NEXUS

Low molecular weight solvents have been used in molecular polymer synthesis, 19 selected solvents were inputted and screened through Derek Nexus and Sarah Nexus program. The traffic light system has been employed to highlight structures that are predicted to be safe in green, potentially toxic in amber and probably toxic in red. A mixture of results can be observed, solvents that were classified as safe whereby did not show any alerts 3MOx, GBL, Cyrene and GVL.

DMPU is a versatile solvent used in polymerisation processes, alerts fired for testicular toxicity. NMP is a restricted chemical, according to REACH. Groupings of these structures also showing similarity within their chemical structures.

There could be some discretion in the data as sulfolane is suspected to be reprotoxic, causing damage to fertility and an unborn child, it can also cause oral toxicity these have not shown up in any alerts in Derek Nexus program. NMP and DMI are known toxic solvents, shown to be toxic in the Nexus program by multiple alerts being fired.

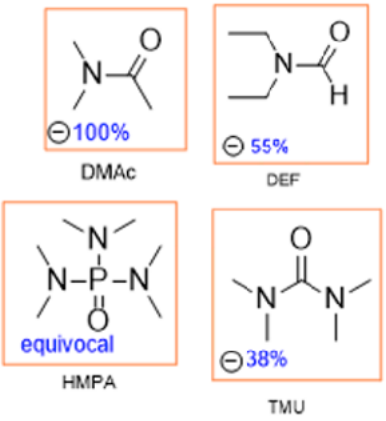


FIGURE 6: LOW MOLECULAR WEIGHT AMBER SOLVENTS SCREENED USING DEREK NEXUS AND SARAH NEXUS

These low molecular weight solvents were also screened using Derek Nexus and Sarah Nexus within this data set. In the amber region found were HMPA, DEF, TMU and DMAc reside, DMAc is a substance of Very High Concern (SVHC). These solvents showed alerts for skin sensitisation and acute toxicity.

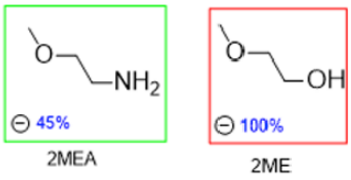


FIGURE 7: 2MEA AND 2ME SOLVENTS SCREENED USING DEREK NEXUS AND SARAH NEXUS

2-methoxyethanol (2ME) is a volatile protic solvent, used for many different processes, in dyes, resins and as an additive in plane de-icing solutions. It is known to cause bone marrow toxicity, developmental toxicity, nephrotoxicity, and testicular toxicity. Derek Nexus results have also highlighted skin sensitisation in mammals as non-sensitiser, denoting that 2ME has the potential to also cause skin irritation reactions in mammals.[30]

2ME and 2MEA consist of a similar chemical structure, the change in functional group hydroxyl to amine has significantly changed to toxicity of the solvent. Even though toxic alerts have not been shown for 2MEA, material safety data sheets (MSDS) show skin toxicity and harm to eyes.

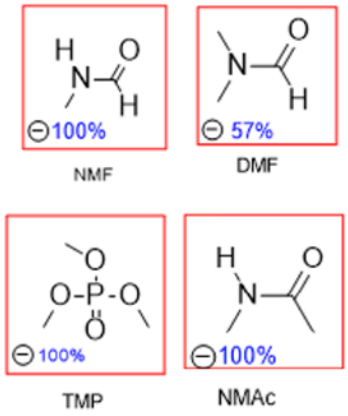


FIGURE 8: LOW MOLECULAR WEIGHT SOLVENTS SCREENED USING DEREK NEXUS AND SARAH NEXUS SHOWING HIGH TOXICITY ALERTS

The structures in this figure considered toxic are NMF, DMF, TMP and NMAc, with known toxicities such as skin and eye irritation as well as acute toxicity. These solvents have been used commonly in laboratory process, synthesis and polymer production. Toxic properties of these are known, Derek Nexus has flagged up these toxicities proving the ability of the software to generate data reliably.

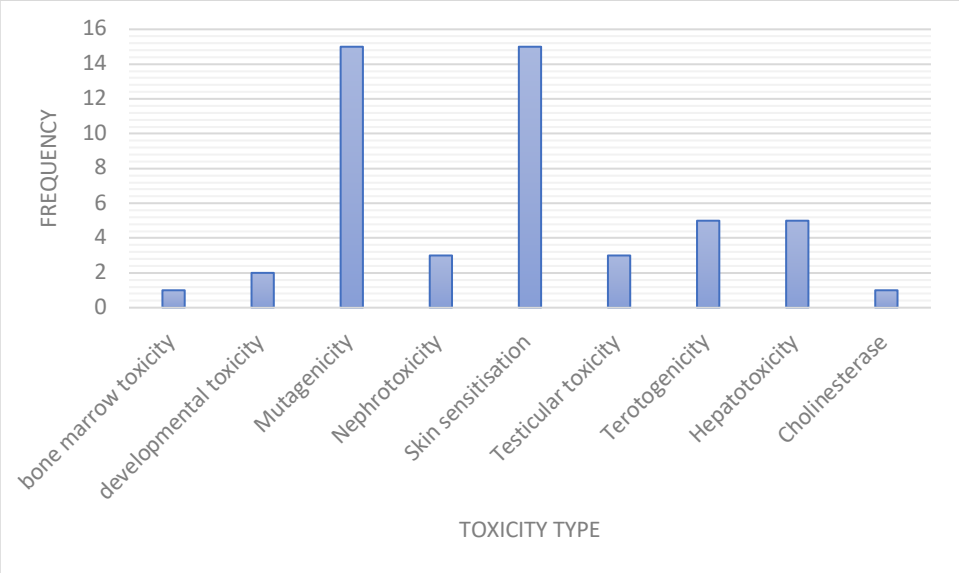


FIGURE 9: TOXICITY TYPES IDENTIFIED IN HUGHES SOLVENT STRUCTURES DATA SET AGAINST FREQUENCY OF ALERTS.

The results generated from the reports of this low molecular weight solvent data set, show various toxicities, as Figure 9 above has demonstrated. It can be deduced that although toxicity to human health is known to be a trait for these solvents there are distinct types of health-related issues caused by these chemicals.

Although misclassified and unclassified features have been shown, alerts have also been fired for many solvent structures. The commonality between these chosen solvents are the toxicity type and alerts fired.

Alert 034 monothioglycol or glycol monoalkyl ether, alkoxy- or alkylthio-carboxylic acid or precursors gives the following reasonings.

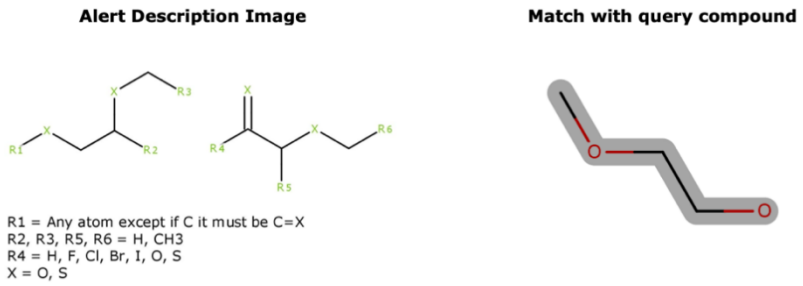


FIGURE 10: DEREK NEXUS REPORT GENERATED SHOWING ALERT DESCRIPTION IMAGE MATCHED WITH QUERY COMPOUND.

This report highlights those compounds which can release glycol monoalkyl ethers $ROCH_2CH_2OH$ where R is a methyl or ethyl group may cause teratogenicity/foetotoxicity.[31]

Figure 10 is taken from Derek Nexus report where monoethyl precursors are taken from the inputted structure and matched to referenced data within the software.

The toxicity of 1,2-ethylene glycols to bone marrow is discussed in this alert. This rapid prototype warning was created utilising a confidential data set of 1467 compounds that were categorised according to whether or not they caused bone marrow lesions in oral rat repeat dosage experiments, which were typically 28 days in length. This rapid prototype alarm was triggered by one molecule in the data set, and that compound proved hazardous to the bone marrow.

If a chemical contains a rapid prototype alert for bone marrow toxicity, then it is considered equivocal that the chemical will cause bone marrow toxicity in mammals and impossible in bacteria. The variation in rule outcome with species is achieved via use of the variable “Species dependent variable 8.” [30]

Sarah predictions have been negative, with few positive results shown in the table. Sarah prediction for 3Mox was given positive with 11% confidence, with the hypothesis being overruled by 781 training set examples.

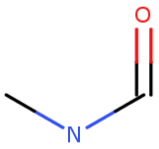
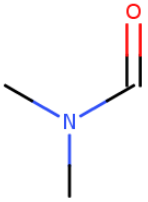
Hepatotoxicity and teratogenicity gave higher frequency of alerts in this data set, with skin sensitisation and mutagenicity being the highest toxicity endpoint.

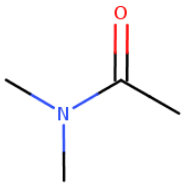
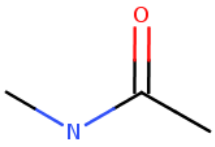
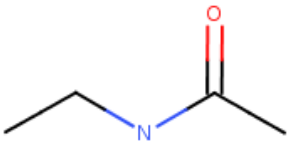
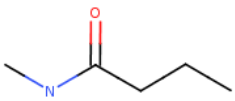
Most frequently alert 696, was present in the Derek Nexus prediction for this solvent dataset. This alert is defined by teratogenicity of short chain alkyl amides. These low molecular weight amides can potentially cause various skeletal, craniofacial, and central nervous system defects including limb and digit abnormalities.

However, there is no direct evidence to prove short chain amides are teratogenic in humans, chemical species with alert 696 have a plausible consideration that chemical will cause teratogenicity in mammals and impossible in bacteria.

Test data: Example 1 is *N*-methylformamide, the study was conducted in rats showing teratogenicity gives a positive result. [32]

TABLE 8: SUMMARISING ALERT 696 SHORT CHAIN ALKYL AMIDES EXAMPLES.[26]

Name	Structure	Test/Results
N-methylformamide CAS Number: 123-39-7		Species: Rat Assay: Teratogenicity study Endpoint(s): Teratogenicity Result: positive Species: Rabbit Assay: Teratogenicity study Endpoint(s): Teratogenicity Result: positive
N,N-dimethylformamide CAS Number: 68-12-2		Species: Rat Assay: Teratogenicity study Endpoint(s): Teratogenicity Result: positive Species: Rabbit Assay: Teratogenicity study Endpoint(s): Teratogenicity Result: positive
N,N-dimethylacetamide		Species: Rat Assay: Teratogenicity study Endpoint(s): Teratogenicity

CAS Number: 127-19-5		Result: positive Species: Rabbit Assay: Teratogenicity study Endpoint(s): Teratogenicity Result: positive
N-methyl acetamide CAS Number: 79-16-3		Species: Rat Assay: Teratogenicity study Endpoint(s): Teratogenicity Result: positive Species: Rabbit Assay: Teratogenicity study Endpoint(s): Teratogenicity Result: positive
N-ethylacetamide CAS Number: 625-50-3		Species: Rat Assay: Teratogenicity study Endpoint(s): Teratogenicity Result: positive
N-methylbutyramide CAS Number: 17794-44-4		Species: Rat Assay: Teratogenicity study Endpoint(s): Teratogenicity Result: positive

Hepatoxicity is a known toxicity endpoint with 5 alerts given in data set, low molecular weight solvents containing formamide derivatives is probable in alert 553. This alert is comprised in most structures. Hepatoxicity is defined by toxicity and damage to the liver function.

Low molecular weight solvents dimethylformamide (DMF), 1,3-Dimethyl-2-imidazolidinone (DMI), *N*-methylformamide (NMF) and TMP have given plausible alerts in Derek Nexus for Hepatotoxicity endpoint. Necrosis and/or micro vesicular steatosis are forms of hepatocellular injuries that have been reported as dose dependant by formamide derivatives.[32]

Formamide derivatives are known as occupational toxicant, these toxic effects have been reported by Zimmerman, Craig *et al* conducted a study in mice and rats with NMF, and DMF being administered into several species gave a positive response to repeated exposure of large doses.[33]

It is believed that *N*-alkyl carbamic acid thioester conjugates, which are produced in mammalian systems during the metabolism of *N*-alkylformamides, may play a significant role in mediating the hepatotoxicity of the parent formamides, possibly by releasing the highly reactive methyl isocyanate at cell membranes.[34]

Derek prediction 553 formamide derivative prediction reports a positive result in human species from a hepatotoxicity case study assay, also a positive result in rats. The prediction report for DMF shows an exact match with query compound was found dimethylformamide, a second prediction would confirm the positive outcome of the alert. Clinical test data in workers has shown a positive correlation between exposure to DMF and associated toxic liver injuries, liver biopsies taken showed hepatocellular necrosis and microvesicular steatosis across the smooth endoplasmic reticulum.[35] This study shows direct correlation, however results cannot be absolutely generalised as occupational history has not been considered.

Teratogenicity is the ability of a toxicant or chemical to cause damage to a developing foetus, substances can cause physical or functional defects in human embryo once a pregnant woman is exposed to different levels of certain substances. Teratogenicity is found as a toxicological endpoint in mammals for solvents DMF, NMAc and NMF. Teratogenicity was reported in mammals as probable, and the alert matched is 696 short chain alkyl amide and exact match from the database N-methylformamide. A study conducted by Von Kreybig T *et al*, showed teratogenicity with a positive test result in rats.[36]

A study in rabbits by Merkle J and Zeller H in 1980 also provided a positive result in the teratogenicity study. Rule 828 in Derek highlights that if teratogenicity study is positive in rabbits, it is certain that species dependant variable is probable. If these chemicals are known to give a positive response in a teratogenic study in rats, then is it considered certain that the chemical will cause teratogenicity in rats, probable in mammals other than the rat and impossible in bacteria.[37]

Nephrotoxicity (Renal toxicity) is a kidney related problem when the body is exposed to drugs or chemical toxicants causing damage overtime with exposure. When this occurs one's body is unable to get rid of excess waste and urine, blood electrolytes such as K^+ and Mg^+ become elevated, dehydration may occur because of this, and kidneys can fail. The serum creatine is present to make energy for muscles, kidneys can filter large amounts of this chemical out. When kidney problems are presented, less creatine will be filtered causing lower blood flow and in turn causing urinary infections.[38]

Nephrotoxicity endpoint has been fired as an alert for 2-methylhydroxyethanol, DMI, TMP. Equivocal for DMI, 2ME with alert matched RapidPrototype069 1,2-Ethyleneglycol or derivative, and TMP with alerts matched to RapidPrototype066 phosphate or phosphonate. The nephrotoxicity of phosphates and phosphonates is discussed, this rapid prototype warning was created utilising a data set of 731 compounds that were categorised according to whether kidney histopathologic lesions were present or absent in trials employing repeated oral doses in rats, most of which lasted 28 days. This rapid prototype warning was triggered by nine substances in this data collection, five of which were nephrotoxic.

Skin sensitisation in TMP is plausible, 462 Alkyl ester of phosphoric or phosphonic acid. Skin sensitisation in mammal is NON-SENSITISER, is the alert fired for this data set for low molecular weight solvents.

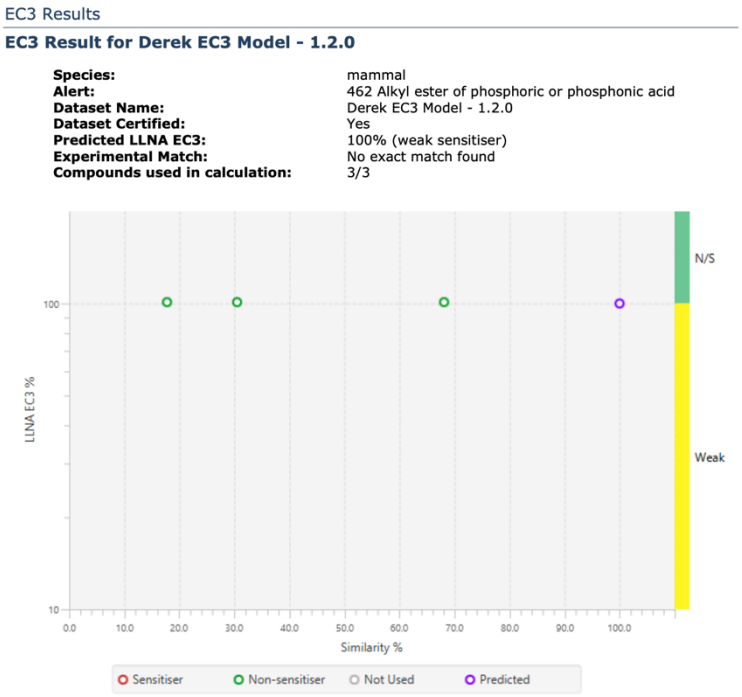


FIGURE 11: EC3 DATA TAKEN FROM DEREK NEXUS REPORT FOR TMP, SKIN SENSITISATION.[26]

Skin sensitisation: guinea pig maximisation test, local lymph node assay (LLNA). Figure 11 is showing the skin sensitisation LLNA EC3 value as predicted 100% in purple, showing alkyl ester of phosphoric acid to be a weak sensitiser.[39]

This endpoint was fired for TMP, alert 003 Organophosphorus, esters and organophosphorus, anhydride and their derivatives inhibit the family of enzymes known as cholinesterases, examples are sarin and ethyl pyrophosphate, ethyl bromophos, 4-nitrophenyl methyl phosphinate. *In vivo* cholinesterase activity study for Sarin and 4-nitrophenyl methyl (phenyl) phosphinate, provided a result of reduced activity for cholinesterase inhibition endpoint[40] in guinea pigs. Test data for ethyl bromophos and trichlorfon also showed reduced activity in the study of *in vivo* cholinesterase inhibition in rats.[38] The final chemical for alert 003 was ethyl pyrophosphate, this *in vivo* study was conducted in rabbits to assess cholinesterase inhibition, results also showed reduced activity.[41]

From Figure 11 the final endpoint to be discussed is testicular toxicity in 2ME and DMI, DMPU. In DMI and DMPU rapid prototype163 Ethylene or propylene urea describes the testicular toxicity of ethylene and propylene, this rapid prototype alert derived from a data set of 1515 chemicals, classified on the basis of histopathologic lesions in testis in oral rat repeat dose studies mostly on a 28 day duration.[39] Rapid prototype alert was activated by three chemicals with a toxic effect to testis. Rapid prototype alert 163 is a marker that indicates whether a substance will be harmful to the testicles in mammals and impossible in bacteria.

Testicular toxicity found to be plausible in mammals in Derek prediction for 2ME,[42] alert matched: 067 Monothioglycol or glycol monoalkyl ether, alkoxy- or alkylthio-carboxylic acid, or precursors. Testicular toxicity may be caused by classes of compounds which can release glycol monoalkyl ethers $ROCH_2CH_2OH$, where R is a methyl or ethyl group. Esters, thioesters, acid halides and the corresponding aldehydes are also included as possible precursors of the acid.

Overall, from the results obtained it is important for further investigations to take place, the toxicities outlined are dependent on the amount used, method of use and/or exposure to the organism, nevertheless, the desirability of using greener solvents for industrial processes is the driving force behind investigations to switch from conventional solvents.[28]

In a circular economy, chemistry will serve as the foundation for novel goods that are manufactured from renewable feedstocks and are intended to be reused, recycled, or have the feedstock refilled naturally. Increasingly, the materials used to make things will be valued as resources on a par with raw materials rather than being simply thrown away. In this viewpoint, the function of chemists in a society devoid of waste is being discussed in a variety of academic and non-academic settings.[43]

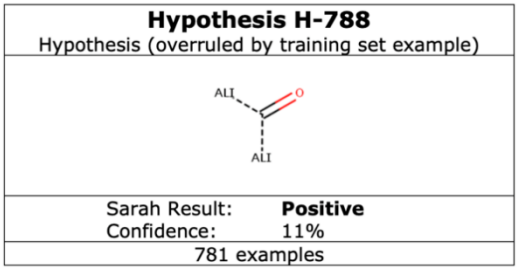


FIGURE 12: POSITIVE SARAH RESULT DISPLAYED IN THE HYPOTHESIS FOR 3MOX[26]

From this set of green structures screened for Sarah, negative results have been found for Ames test. This figure was taken from the 3MOx report highlighting the percentage of 11% confidence in positive result from 781 examples screened in the database.

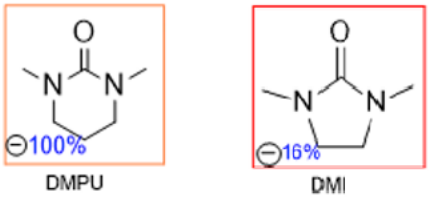


FIGURE 13: COMPARISON OF DMPU AND DMI SOLVENTS

1,3-Dimethyl-2-imidazolidinone (DMI) and N,N'-Dimethylpropyleneurea (DMPU) are shown above in figure 13, both used as solvents in polymer processing. DMPU is amber as alerts for skin sensitisation and testicular toxicity were shown, MSDS's for this solvent also show eye irritation and reproductive toxicity. DMI is toxic solvent, multiple alerts in Derek Nexus have also highlighted different toxicities such as hepatotoxicity, nephrotoxicity, skin irritation and testicular toxicity.

3.1.1.2 CHALLENGES FOR SOLVENT STRUCTURES

Solvent structures in this section have given a range of results, safer solvents showing fewer alerts in green and common hazardous solvents showing multiple alerts for toxicity, thus proving the working ability of Derek Nexus and Sarah Nexus.

The overall impact of these safer alternatives cannot be assumed, as the overall assessment over the chemical's life cycle must be considered before concluding sustainability. Industries are left with choices of solvents that do not meet the specifications and technical requirement for the desired product results or yield.

3.1.2 SOLVENTS FOR PROCESSING PERVOSKITES

A small selection of low molecular weight solvents potentially useful in solution processing perovskites for solar cell applications have been analysed in Figure 13.[44] Perovskite cells have shown high durability and already marketed as products for solar films in commercial technology, solar cells have the ability for greater efficiency.[45] The main advantage is the high-performance ability and cost effectiveness. Examples include, metal halide complexes, lead, or methylammonium lead halides. The most promising path toward stable perovskite solar cells is inorganic connections and inorganic perovskite compositions.[46]

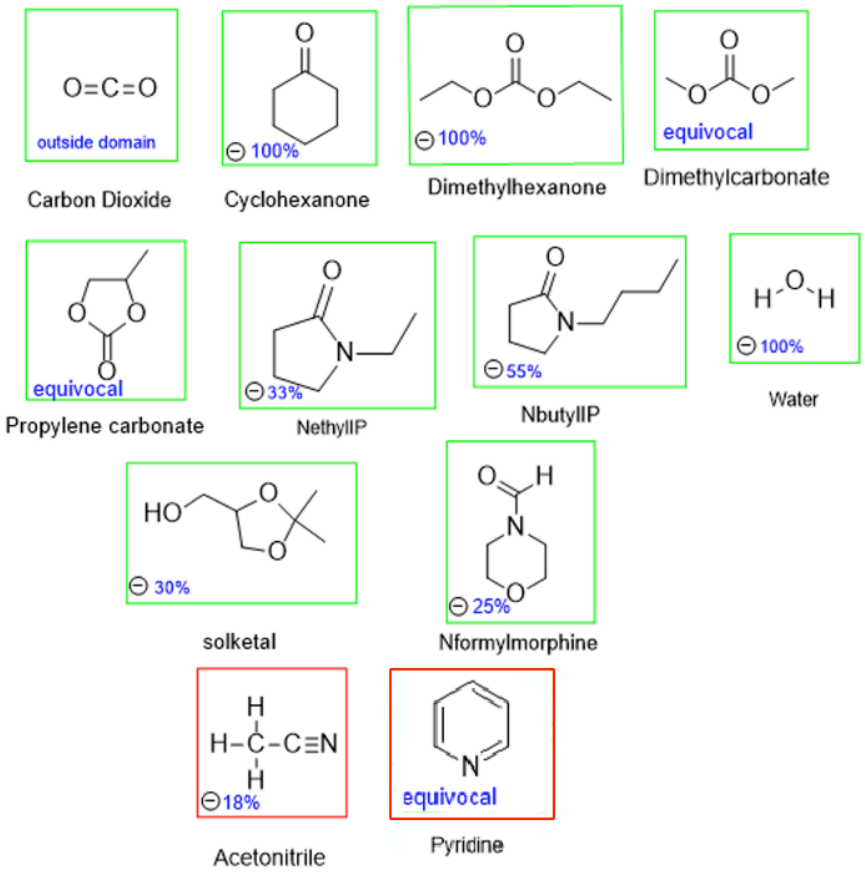


FIGURE 13: SOLVENTS USED IN POLYMER DISSOLUTION SCREENED BY DEREK NEXUS AND SARAH NEXUS AND SUMMARISED USING THE TRAFFIC LIGHT SYSTEM.

Due to the exceptional thermal stability, inorganic perovskite solar cells (PVSCs) have attracted great interest. The most distinctive photoelectric characteristics can be recognised and lead to

interest in three-dimensional organic-inorganic lead halide perovskites. Due to their distinctive photoelectric characteristics and straightforward production, three-dimensional (3D) organic–inorganic lead halide perovskites have become the subject of intense investigation. [47]

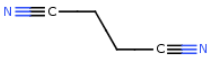
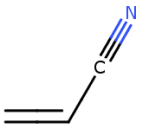
Low molecular weight solvents used in materials processing are displayed above, most these solvents are shown to be safe as they are green as no alerts were fired through Derek Nexus, despite toxicity shown and safety requirements when handling in the materials safety data sheets.

Acetonitrile as a solvent is popular choice, as it is highly polar and volatile. It is commonly used in the pharmaceuticals industry to cast plastic materials, extractive distillation in petrochemical industry and commonly in analytical laboratories for HPLC analysis and many more. It is distinctively known for its toxic characteristics and high flammability, it is metabolised into hydrogen cyanide and thiocyanate, both poisonous and hazardous to human health and the environment.[48]

Acetonitrile gave matching toxicity alerts for cyanide type effects in mammal as plausible, mutagenicity *in vitro* in bacterium as inactive, skin sensitisation in mammal is non-sensitiser and teratogenicity as plausible with exact compound matched as an alert. Acetonitrile belongs to the nitrile class of compounds, these can release cyanide upon metabolism, and produce typical cyanide toxicity. Cyanohydrins and cyanohydrin esters are likely to be released for most nitrile compounds.[49]

Alkyl, alkenyl and alkynyl nitriles with up to a C4 chain are also active, causing similar toxicity affects.[50] There must be at least one alpha hydrogen, and the chain may only carry non-polar substituents. A second nitrile group, separated from the first by two or more carbon atoms will also remove the activity.[51]

TABLE 9: SUMMARY OF NITRILES LINKED TO CAUSING TERATOGENICITY

 **Beta-aminopropionitrile** Species: rat Assay: Teratogenicity Study Result: Positive Species: Monkey Assay: Teratogenicity Study Result: Positive Species: Hamster Assay: Teratogenicity Study Result: Positive	 **Succinonitrile** Species: Hamster Assay: Teratogenicity Study Result: Positive	 **Acrylonitrile** Species: rat Assay: Teratogenicity Study Result: Positive Species: Hamster Assay: Teratogenicity Study Result: Positive Species: Hamster Assay: Teratogenicity Study Result: Positive

From the table summary above, alerts matched to nitrile toxicity have been highlighted and noted. Aliphatic nitriles have shown to cause malformations in rodents and foetal deaths which are characterised by the increased resorptions. One of the study's above by Steffek *et al* has shown limb and tail abnormalities, as vertebral column defects develop.[52]

Aliphatic nitrile can be seen in rats and hamsters in various studies, acrylonitrile, aminoacetonitrile and succinonitrile have been reported to cause vertebral defects, neonatal death, this is dose dependant. From the study into three baboons, foetal macerations and spinal bifida had been observed in one foetus after intravenous administration at 500mg/kg a day.[52] Dichloroacetonitrile was tested in rats it had shown malformations such as cardiovascular, digestive and urogenital. Acrylonitrile injections administered to Hamsters caused embryonic toxicities and teratogenic effects.[53] These *in vivo* studies have further confirmed the toxicities caused by acetonitrile, and the cyanide type effects caused.

Due to the increased open circuit voltage, improved film uniformity, and improved hydrophobicity of pyridine derivatives, which further increases the long-term durability, they are utilised as charge carriers, or hole transporting materials (HTMs), in perovskite solar cells (PSCs).[50]

The data also states Sarah results from the predictions, negative result, out of domain and equivocal.

3.1.2.2 COMPARISON OF CYCLOHEXANONE AND PYRIDINE

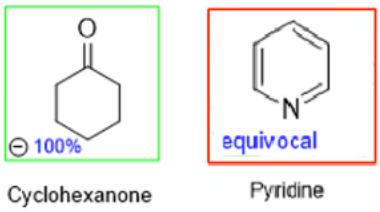


FIGURE 14: CYCLOHEXANONE AND PYRIDINE STRUCTURES

The safety and hazard identification for cyclohexanone solvent is known acute toxicity, skin irritation and serious eye damage. Skin sensitisation has been flagged by Derek nexus for cyclohexanone only. Pyridine is commonly used as a solvent to make many products such as medicines, flavourings and vitamins, it is a known solvent causing damage to skin and eyes. Derek Nexus reports have shown pyridine with skin sensitisation alerts.

3.1.2.3 CHALLENGES FOR PERVOSKITES

The results generated from Derek Nexus have confirmed known toxicities of acetonitrile and giving unclassified or misclassified features with no alerts for the solvents in this data set. There could be potential affects that have not been detected by the software.[54] Overall, it can be seen from this data set that green solvents are less toxic and may be better for industrial purposes and the environment. Derek Nexus has supported the results seen in MSDS's for these solvents.

3.1.3 IONIC LIQUIDS

Ionic liquids are a class of salts that have low melting points and can be potent solvents and electrolytes.[55] Small selection of ionic liquids potentially useful in polymer synthesis was analysed Figure 11.[56] With examples of the utility of such ionic liquids including (but not limited to): 1-butyl-3-methylimidazolium hexafluorophosphate (BMIM-PF_6) used in the free radical polymerization of methyl methacrylate,[57] atom transfer radical polymerization (ATRP) of methyl methacrylate,[58] catatonically polymerized styrene,[59] electrochemical synthesis of polypyrrole,[60] enzymatic polymerisation using *Candida antarctica* lipase-catalysed formation of polycaprolactone;[61] 1-butyl-3-methylimidazolium tetrafluoroborate (BMIM-BF_4) used in the anionic polymerization of methyl methacrylate,[62] cationic ring opening polymerization of 3-ethyl-3-hydroxymethyloxetane,[64] transition metal-catalysed polymerizations (e.g., Rhodium(I) catalysed polymerization of phenylacetylene),[65] enzymatic polymerisation using *Candida antarctica* lipase-catalysed formation of polycaprolactone;[61] tetrabutylammonium bromide (TBAB)[66] used in the group transfer polymerisation of methyl methacrylate;[67] 1-butyl-2,3-dimethylimidazolium hexafluorophosphate (BDMIM-PF_6) used in the ring-opening metathesis polymerization of norbornene;[68] phosphonium-based ILs containing different ions (e.g., trihexyl(tetradecyl)phosphonium bis(2-ethylhexyl)phosphate (Cyphos™ IL 349) used in the microwave-assisted ring opening polymerization of ε-caprolactone;[69] 1-butyl-3-methylimidazolium bromide (BMIM-Br) used in the step-growth polymerization of poly(ether ketones).[70]

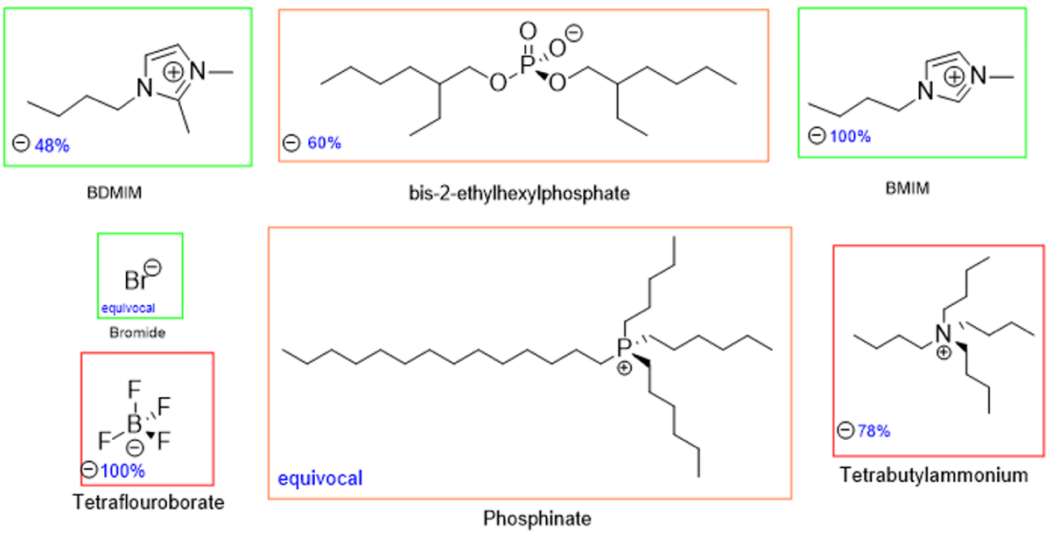


FIGURE 15: IONIC LIQUIDS STRUCTURES SCREENED IN DEREK NEXUS AND SARAH AND CATEGORISED USING THE TRAFFIC LIGHT SYSTEM

Ionic liquids used in polymerisation, are known as green solvents for their applications in industrial process due to the high ionic conductivity, chemical stability, and thermal stability in processes. Ionic liquids in processes are increasing, applications for functional polymers. [71]

The structures shown above in Figure 15 show BDMIM, bromide, BMIM, as safer ionic liquids. In the amber region phosphinate cation and bis-2-ethylphosphate have some characteristic properties of low volatility and low toxicity.

For this set of data common alerts fired are Organophosphorus-di- or tri- ester for Bis-2-ethylhexyl-phosphate, main alerts fired are as follows: 653 Organophosphorus-di- or tri- ester, RapidPrototype007 Phosphine, or phosphine oxide, 218 Quaternary ammonium salt, 436 quaternary ammonium salt and 216 Boron halide. Common endpoints found for these ionic liquids are hepatotoxicity, carcinogenicity, irritation to the eye, irritation to the skin, and thyroid toxicity. Chemicals containing organophosphorus, bis-2-ethylhexyl-phosphate fired an alert describing hepatoxicity, causing liver problems such as hepatocellular necrosis, hypertrophy. Derek Nexus literature database shows positive results for *in vivo* tests in studies in rats. Matches examples are tris(2-butoxyethyl) phosphate,[72] phosphamidon[73], diphenyl 2-ethylhexyl phosphate[73]. From these *in vivo* studies hepatotoxicity in mammals is considered plausible, however there is limited evidence to show that these compounds can cause damage in humans.

RapidPrototype007 Phosphine or phosphine oxide for phosphinate structure shown hepatotoxicity as equivocal. This rapid prototype alert was derived using data from 731 chemicals, studied histopathologic legions in rat liver repeated dose which took place over 28 days. From these five chemicals activated this alert as hepatotoxic. Therefore, this data cannot be further extrapolated to humans, phosphinate is for this reason in the amber region as its safety cannot be assured.

Tetrafluoroborate has been highlighted in red, as Boron halides are readily hydrolysed producing highly corrosive irritants. Known irritants are Boron trifluoride which is classified as corrosive-R35, causing severe burns, Boron trichloride and Boron tribromide - 34 and R35, respectively. For this structural alert, skin and irritation potential can be noted as these chemicals are highly reactive substances.[74] Alert 242 Fluoroborate salt has fired thyroid toxicity, this is caused by inorganic fluoroborate ions such as sodium or potassium fluoroborate, these have the potential to inhibit the iodide uptake pump of thyroid glands, this would increase the levels of thyroid stimulating hormones (TSH) production leading to follicular cell hyperplasia[75] Derek Nexus has provided the following reasonings: carcinogenicity in mammal as plausible, irritation to the eye I

mammal as plausible, irritation to the respiratory tract in mammal as plausible, irritation of the skin in mammal as plausible, and thyroid toxicity in mammal as plausible.

Quaternary ammonium salt is the alert matched to tetrabutylammonium, known irritants which fire the alert include: Stearylphenylethyldimethylammonium tosylate Benzyl-2-hydroxydodecyldimethyl ammonium benzoate (R34) Didecylmethyl alkoxyammonium chloride (R34) Hexadecyltrimethyl ammonium chloride (R34).[76] Quaternary ammonium compounds' surfactant qualities, which necessitate a lipophilic chain in the molecule, are what cause them to be irritating. Tri(*n*-butyl)benzylammonium 4-hydroxynaphthalene-1-sulphonate and dimethyldistearyl-ammonium bisulphate are two examples of quaternary ammonium compounds that may only moderately irritate skin or eyes.[77]

Cationic surfactants have been described using the QSAR model by Patlewicz *et al*, molecular volume, LogP and micelle concentrations have been used as parameters, these were to derive the model using back propagation neural network analysis.[78]

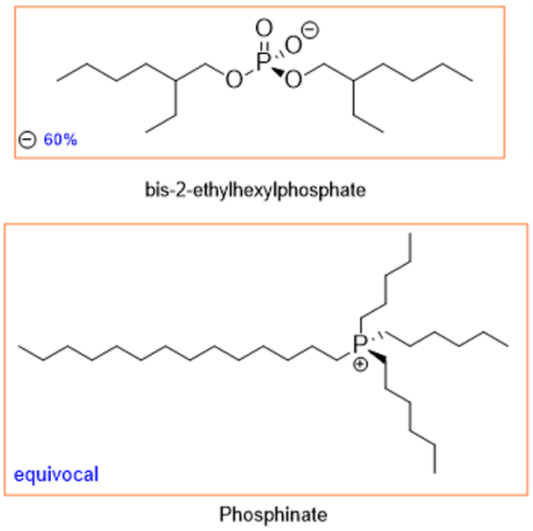


FIGURE 16: COMPARISON OF PHOSPHINATE AND BIS-2-ETHYLHEXYLPHOSPHATE

From figure 16 Phosphinate and bis-2-ethylhexylphosphate can be compared, with the common phosphine group. Phosphine gas on its own is a toxic gas, these ligands add to the enhanced toxic effects. Derek Nexus reports for both compounds show similarity in alerts, hepatotoxicity in mammal as equivocal in phosphinate and plausible in bis-2-ethylhexylphosphate, as well as

additional carcinogenicity in mammal as opened. Skin sensitisation is mammal as non-sensitiser for both.

3.1.3.3 CHALLENGES FOR IONIC LIQUIDS

Although red alerts have not been fired for many solvents in this data set, skin sensitisation is a potential for all. These solvents are described as weak sensitisers due insufficient supporting data to predict EC3 values. Skin sensitisation as equivocal as no sufficient data is present to make a strong EC3 prediction, Derek Nexus gives 30% as weak sensitiser. Sarah Nexus has provided negative predictions for these ionic liquid structures; no mutagenic activity has been shown in the Ames test.

New applications of ionic liquid solvents include creating new classes of materials. Advanced classes of stimuli responsive solvogels, Ionic liquids are more desirable over conventional solvents, they possess favourable physicochemical properties, ranging from non-volatile and non-flammable and water soluble.[71] This preference is leading to applications and use in industry.

It is feasible to adjust the Ionic liquid, the polymer, or both to produce a wide range of multifunctional composites and meet the unique needs of various applications. The goal of this effort is to produce smart IL/polymer-based materials for applications such as responsive and sensitive actuators, batteries, fuel cells, and biomedical applications. It also focuses on innovative materials and techniques relating to Ionic liquids and polymers.[79]

3.2 CATALYSTS FOR POLYMERISATION

Biological catalysts have evolved to catalyse various reactions; the most widely known biological catalysts are enzymes (soluble enzymes are homogeneous catalysts, whereas membrane-bound enzymes are heterogeneous catalysts), however, non-protein-based biomolecules can exhibit catalytic properties (e.g., ribozymes, and synthetic deoxyribozymes). While biological catalysts are a potent and exciting class of catalysts, the amount of computational resource required to evaluate such large structures was unavailable during the course of the project, and instead we focus on a variety of somewhat simpler structures that may be useful for catalysis.

3.2.1 HETEROGENEOUS CATALYSIS

3.2.1.1 INDUSTRIAL HETEROGENEOUS CATALYSIS FOR POLYMERISATION.

An historically and industrially relevant example of heterogeneous catalysis for polymerisation is olefin polymerization for which the 1963 Nobel Prize in Chemistry was awarded to Karl Ziegler, for his discovery of first titanium-based catalysts, and Giulio Natta, for using them to prepare stereoregular polymers from propylene.[80] Ziegler–Natta catalysts ($TiCl_3$ on $MgCl_2$) have been used in the commercial manufacture of various polyolefins since 1956, currently at a scale of millions of tonnes per annum.[81]

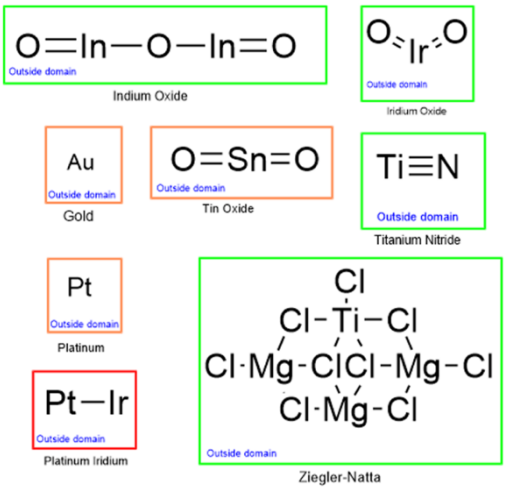


FIGURE 17: HETEROGENOUS CATALYSTS SCREENED BY DEREK NEXUS AND SARAH NEXUS AND SUMMARISED USING TRAFFIC LIGHT SYSTEM

This set of data has presented catalysts with toxicity alert 443 found in metal salts, including gold, platinum, Pt-Ir, for skin toxicity, many metals have the potential to cause allergic contact dermatitis.[82]

Other metals such as Indium oxide, iridium oxide, IrO_2, TiN and the Zeigler-Natta catalyst, did not fire any alerts for specific toxicities. Mutagenicity *in vitro* in bacterium is inactive and skin

sensitisation in mammal is predicted non-sensitiser, these outcomes signify substructures within the input structure as having the potential to cause mutagenicity in bacterium, and potential skin sensitisation in mammals. As the reasoning given in Derek states that unclassified features are present, Lhasa's skin sensitisation negative prediction dataset does not contain some features in the molecules, therefore verification would be required in this case.

Certain metals in contact with human skin for long periods of time can be absorbed through sweat and cause immunological sensitisation to soluble metal compounds in susceptible individuals, which can result in a delayed hypersensitivity reaction following future contact with that metal or its compounds.[83] Metals producing skin sensitisation reactions are nickel, chromium and cobalt, usually observed through jewellery, medical or dental implants. [84]

Platinum plays an important role in industrial processes, its presence and toxicological effects in a molecule must be accurately and quickly identified. Derek Nexus software successfully detected platinum toxicities for platinum compound and platinum-iridium, according to reports, molecules with charged and reactive ligand structures are more likely to cause allergies when exposed to platinum.

Both the guinea pig maximisation test and the local lymph node assay have given positive results; however, the derived potency does not match human clinical experience. Copper chloride, for example, elicited a robust reaction in the LLNA, even though copper salts are rarely skin sensitizers in humans.[75] Predicted LLNA EC3 value for 1.5% weak sensitiser.[74]

Certain studies in various species have shown a response for skin sensitisation, example compounds are potassium dichromate, nickel sulphate, sodium dichromate, ammonium dichromate.[86]

From this set of heterogenous catalysts, platinum-Iridium, a metal alloy is showing chromosomal damage as a toxicity alert for Platinum–iridium, Pt-Ir. Alerts for skin sensitisation as well as chromosomal damage *in vitro* with a matched alert 640 platinum compound are present. Test data from Nexus shows a positive result for an *in vitro* micronucleus test in humans with endpoint chromosome damage *in vitro* in platinum (II) chloride. A similar positive result was found in a study conducted by Migliore *et al* in 2002, an *in vitro* micronucleus test showing positive chromosome damage *in vitro*.[85]

3.2.1.2 CHALLENGES FOR INDUSTRIAL HETEROGENEOUS CATALYSTS

The performance Ziegler Natta catalyst in polymerisation is seen through sustainable processes, Derek Nexus fired no alerts giving this a safety score, however catalysts can reduce their effects by reacting with various monomers causing poisonous effects and differences in yields.[87] Various greener solvents and catalysed used in polymerisation are shown to be safer, than those with known toxic metallic properties.

Derek Nexus has outlined metal toxicities for this data set of catalysts, many alerts have been shown up with potential toxic effects. Safer catalysts have been highlighted also, the range of results gives this data set versatility, results can be applied to polymer applications, thus showing the positive benefits of Derek Nexus software. However, these catalysts have not been successfully tested with Sarah Nexus as these shown outside domain results.

3.2.2 NANOZYMES

Nanozymes are nanomaterials with enzyme-like activities; such nanomaterials tend to be producible on large scales, with high structural stability and tuneable catalytic activities.[88] Examples of nanozyme structures include: CuO nanozymes used for the polymerisation of *N*,*N*-dimethylacrylamide and methylene-bisacrylamide to yield hydrogels;[89] iron oxide (Fe_3O_4) nanozymes used for the polymerisation of peptides containing aromatic amino acids yielding nanogels via peroxidase-like activity;[90] MnO_2 nanozymes used for the polymerization of humic acid;[91] MoS_2 nanozymes used for the polymerisation of trithiocyanuric acid to polythiocyanuric acid via disulphide linkage formation;[92] copper phosphate ($Cu_3(PO_4)_2$) nanozymes used for the polymerization of poly(ethylene glycol) diacrylate;[93] Cu,Gd@BHb^{FITC} nanozymes (prepared from $GdCl_3$, $CuCl_2$, Bovine haemoglobin (BHb) and FITC; yielding 2-3 nm nanoparticles including Cu(0) and Cu(II)) catalyse the oxidative polymerization of dopamine (DA) to PDA-dots,[94] however the complex/unclear chemistry of the latter example highlights complications in analysis of such structures (likewise with dopamine-hybridized carbon quantum dots-supported Pd single atoms that catalyse the in situ free-radical polymerization of poly(ethylene glycol) diacrylate without heating or UV light irradiation).[95]

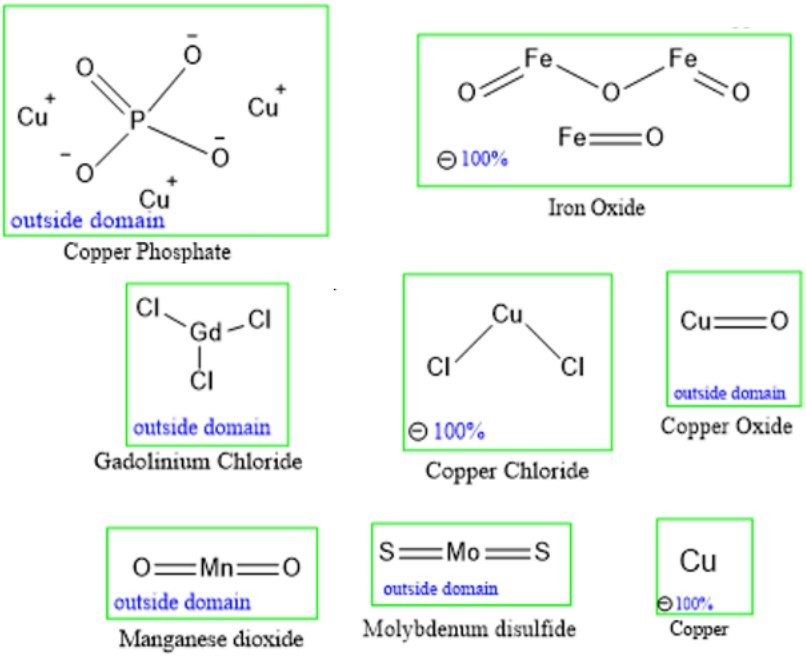


FIGURE 18: NANOZYME STRUCTURES SCREENED USING DEREK NEXUS AND SARAH NEXUS AND HIGHLIGHTED USING TRAFFIC LIGHT SYSTEM

Nanozymes have exceptional physical properties and are widely used and applied to many important fields including medicine. Above we can see a small selection of species that have been used in various bio-medical applications. Iron oxide has been used as a nanozyme to accelerate biomedicine, targeted drug delivery and customised diagnostic imaging due to special magnetic properties.[96] Iron oxide nanoparticles Fe_2O_3 and Fe_3O_4 have constituted of typical nanozymes, with physical characteristics of catalytic species. Fe_3O_4 nanoparticles participating in single electron transfer processes and boosting peroxidase-like activity by a factor of about ten.

Manganese dioxide (MnO_2) nanozyme has shown enzyme-like properties, high stability, and biocompatibility under physiological conditions with oxidase like activities.[97] Report for Derek Nexus showed no alerts, but 'unclassified features' which means there could be possibility of potential skin sensitisations.

As for Copper based nanozymes, copper phosphate, copper chloride, copper phosphate has shown no alerts fired in the Derek Nexus database, and these structures are highlighted in green. Copper based compounds are of great interest, due to their potential electrical conductivity and

cost-effective methods of producing nanoparticles, which are being investigated for nano devices.[98]

3.2.2.2 CHALLENGES NANOZYME CATALYSTS

From these nanozymes screened, it can be deduced that these catalysts are safe to use, and low toxicity profile. This would determine sustainability for biomedical processes, nanozymes can be combined with other nanomaterials, besides pure iron oxide, to create multifunctional hybrid nanocomplexes.

All have been highlighted in green, however this does not define them or mean they are completely safe. For the purpose of this research assumptions can be made from the data shown during the Derek Nexus test run.

Copper containing species cause metal toxicities, Copper phosphate causes serious eye damage and skin irritation, copper chloride causes eye and skin and respiratory irritation.

Domains for the Derek nexus software cause concerns of ambiguity in the results, this is as the software stated unclassified features for iron oxide, The reason for these nanozymes being placed in the green category is due to no alerts being fired. This does not mean these catalysts can be considered safe. Operating domains of the software need to be re-tested to validate these results.

3.2.3 HOMOGENEOUS CATALYSIS

3.2.3.1 ORGANIC CATALYSIS

Well defined macromolecules have become possible to design and prepare due to the ring-opening polymerization (ROP) of lactones and other heterocyclic monomers. Organo-metallic catalysts have created an interesting appeal and topics of interest within the polymer industry.[99] Aliphatic polyesters and polycarbonates are becoming popular as biodegradable polymers, also very favourable for biomedical applications as they are environmentally friendly.[100] A further alternatives to this are Ring Opening Polymerisation (ROP). The degradable and biocompatible properties of ROP's have led to recent advances into further research and applications.[101]

3.2.3.2 ORGANOMETALLIC CATALYSTS

Ligand design has been the subject of intense research interest in the development of organometallic catalysts, and heterometallic cooperativity offers a promising strategy to further tune organometallic catalyst function.[102]

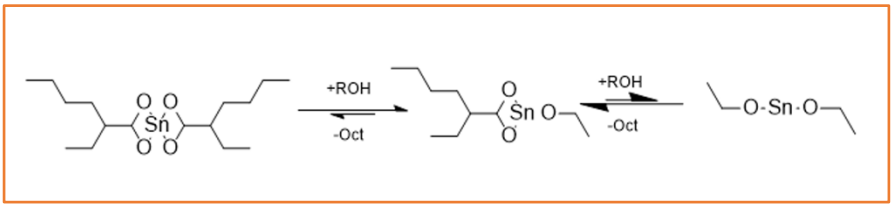


FIGURE 19: SCHEME SHOWING EQUILIBRIUM OF SN(OCT)2 AND AN ALCOHOL AS COINITIATOR. LEFT: SN (OCT)2 MIDDLE: THE MONO-ALKOXIDE SPECIES (HIGHLY ACTIVE) RIGHT: THE BISALKOXIDE.[103]

3.2.3.3 HETEROMETALLIC BIMETALLIC ZINC CATALYSTS

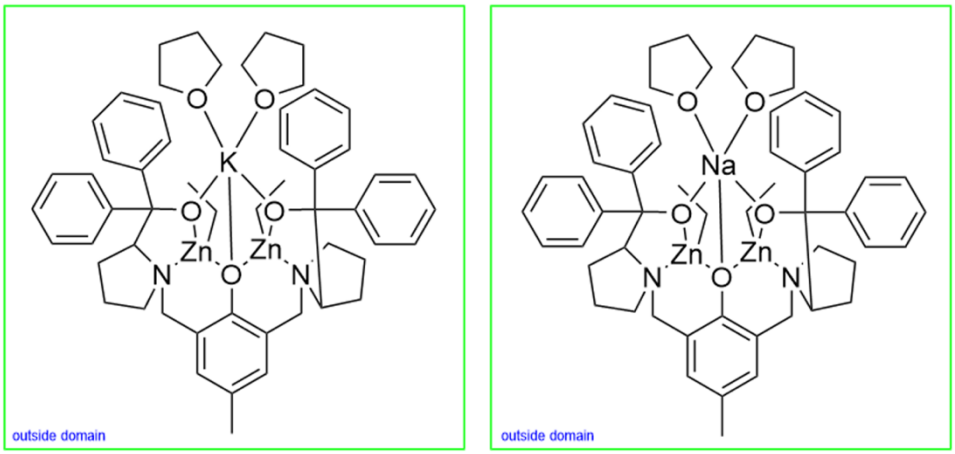


FIGURE 20: STRUCTURES TAKEN FROM GARDEN *ET AL*, SCREENED USING DEREK NEXUS AND SARAH NEXUS PREDICTION HIGHLIGHTED IN GREEN [104]

The production of catalysts is conducted in any forms, cyclic ester ring opening polymerisation has shown to be an effective method to produce polyesters, such as poly lactic acid (PLA) and poly(3-caprolactone) (PCL) and poly(d-valerolactone) (PVL).[105]

The structures drawn above have been taken from Garden *et al* literature paper from chemical science, $[LNaZn_2Et_2(THF)_2]$ and $[LKZn_2Et_2(THF)_2]$ through the synthesis of novel heterometallic complexes for cyclic ester ROP. Heterometallic species have been reported to show increased catalytic activity compared with monometallic complexes.[106]

Derek Nexus has reported these structures as non-mutagenic and non-sensitisers, these catalytic species are predicted to be inactive in the bacterial *in vitro* (Ames) mutagenicity test. Structures have also not been found in the skin sensitisation data set, therefore considered as non-sensitisers, no alerts have also been fired.

Fragments from these structures could not be found in the Sarah Nexus training data library, resulting in an out of domain prediction.

3.2.3.4 HETEROMETALLIC SALEN (AL) SPECIES

Cooperative heterometallic Salen (Al) species have experimentally explored, they have displayed excellent catalytic properties in lactide ring opening polymerisation (ROP). Heterometallic bimetallic species have been explored and experimentally combined to improve catalytic performance. Poly(Lactic acid) is a favourable degradable polymer, a great alternative to conventional polymers with biomedical applications.[107]

The significance of using heterometallic metals as catalyst over homometallic, is the significant enhancement of the C-H bonded metal halogen exchange, and olefin polymerisation, this combination is showing a greater promise for Ring Opening Polymerising.

Aluminium Salen catalysts have been reported to show excellent stereo control and high selectivity, very few catalysts have been studied for Lactide ROP. Heterometallic catalyst that are reported have asymmetric ligands, these offers simplicity for the complexes. [105-106]

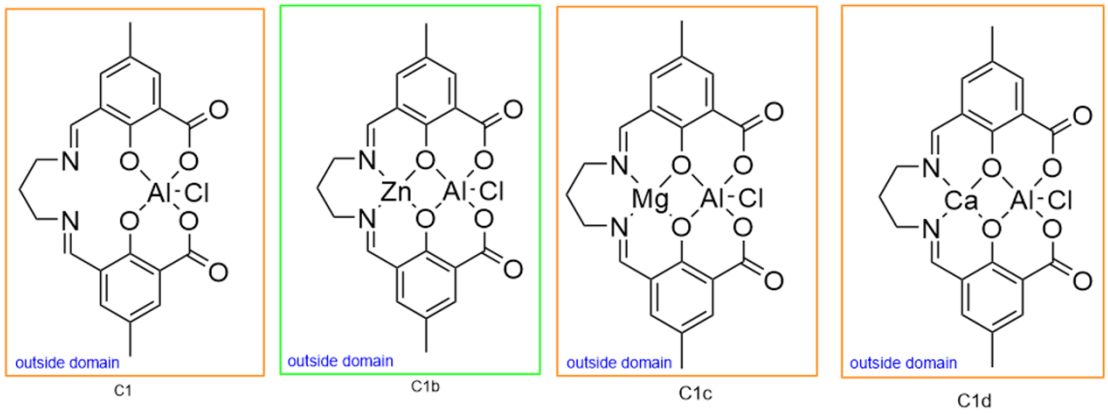


FIGURE 21: HETEROMETALLIC (SALEN) AL COMPLEXES - SHOWING STRUCTURES FROM GARDEN AT AL USED IN ROP, SCREENED USING DEREK NEXUS AND SARAH NEXUS

Heterometallic molecular species can improve catalytic activity compared to their homometallic analogues. Cooperative Al/Mg and Al/Zn combinations have been studied by Garden *et al*, this set of complexes has given an interesting set of results.

These Salen complexes have been synthesised at different metal sites, the optimised reaction conditions are as follows: 1) Et_2AlCl in tetrahydrofuran (THF) at -78°C, 2) 1 equiv. bBu_2Mg in THF at room-temperature (RT) and 3) 1 eqiuv. Et_2Zn in THF at RT.[106]

The computational study on these molecular structures using Derek Nexus gave a mixture of results, from Figure 18 amber and green highlighted structures are shown.

Derek Nexus has reported C1b complex is the Zinc/Aluminium complex, is classed as safe, with no alerts fired. This has been highlighted in green by the traffic light system. This complex provided an out of domain result in Sarah Nexus, this could have been due to the structural features not being recognised in Lhasa's programs Fragments from these structures could not be found in the Sarah Nexus training data library, resulting in an out of domain prediction. C1b would be considered as non-mutagenic and a non-sensitiser.

Matching structures have also not been found in the skin sensitisation data set, therefore considered as non-sensitisers. The Lhasa Ames test reference does not match any structural alerts for bacterial *in vitro* mutagenicity, it is predicted Inactive.

Molecular structure of C1 and C1c are highlighted in amber, showing alerts fired through Derek Nexus for skin sensitisation as equivocal. Skin sensitisation prediction has equivocal results, alerts matched; 444 Imine or alpha, beta-unsaturated imine for C1 and C1C.[107]

In GMPT positive results have also been reported for various substituted phenols, examples including 4-styrylphenol (4-hydroxystilbene), and chloroxylenol (4-chloro-2,3-dimethylphenol).[123] C1C structure has a similar chemistry to C1, however for this molecular structure 439 substituted phenol alert has become apparent in Derek Nexus.[74]

It is also important to note that negative results have also been reported for various substituted phenols in LLNA, GMPT and human maximization tests. The toxicity data for substituted phenols is varied, and they may cause skin sensitisation via numerous potentially overlapping routes. As a result, the alert's reach encompasses a wide spectrum of structurally distinct substances. On the phenol ring, hydrogen, alkyl, aryl, and halogen substituents are all allowed, as are further aromatic ring fusions. Compounds containing phenolic derivates have been associated with skin sanitation reactions *in vivo* and *in vitro*.[108]

Species:	mammal
Alert:	439 Substituted phenol
Dataset Name:	Derek EC3 Model - 1.2.0
Dataset Certified:	Yes
Predicted LLNA EC3:	0.23% (strong sensitiser)
Experimental Match:	No exact match found
Compounds used in calculation:	10/21

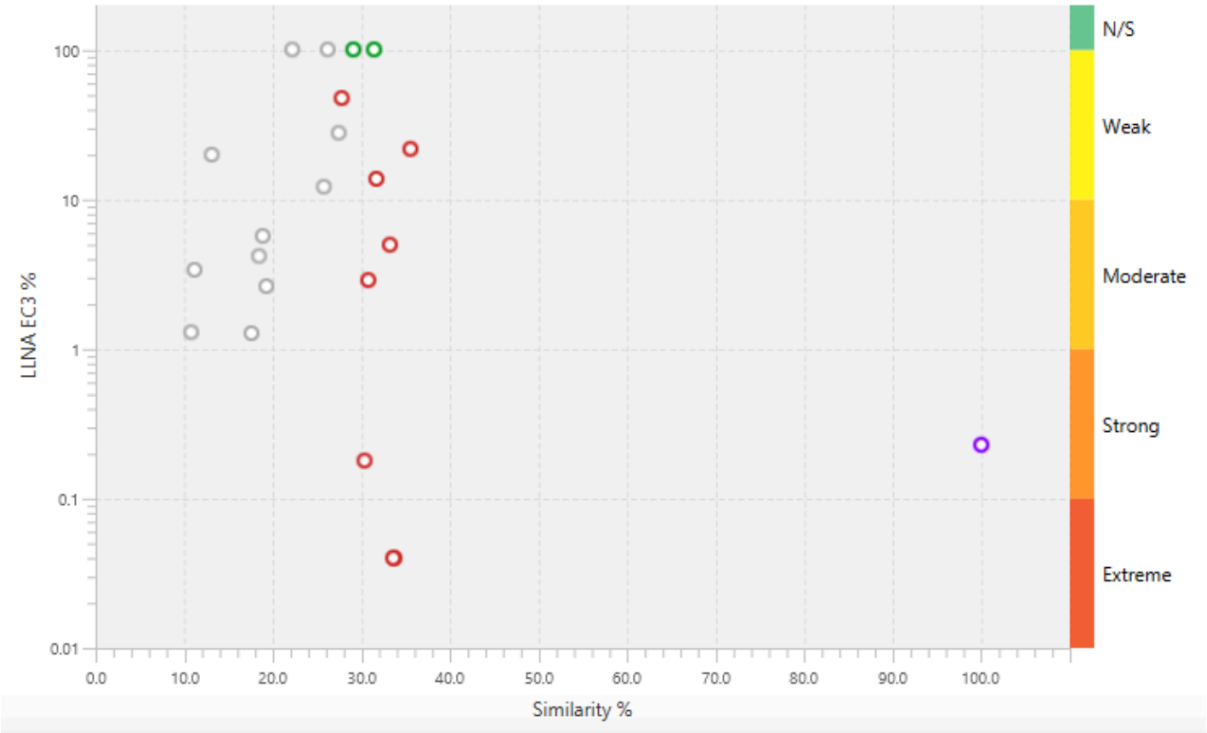


FIGURE 22: LLNA PREDICTION TAKEN FROM DEREK NEXUS REPORT FOR SUBSTITUTED PHENOLS [109]

This EC3 prediction above shows phenol as a strong sensitiser, with predicted value of 0.23% for skin sensitisation in mammals.

From the complexes above illustrated in Figure 21, C1C had been a challenge to screen, as the valency of atoms would not match the requirements of the software, causing re-occurring errors in Derek Nexus. After liaising with Lhasa's Scientists, this structure was separated and analysed with a Mg^{2+} as separate ion with the structure.

The reason for this molecular structure also being highlighted amber as it has shown alerts in Derek for skin sensitisation. Mutagenicity *in vitro* bacterium is Inactive, for structure 1 and 2, contains unclassified features for structure 1. And no misclassified features for structure 2.

Skin sensitisation has been reported as being equivocal by Derek Nexus, with the 439 substituted phenol being the matching alert. The literature data set in Lhasa has shown LLNA, GMPT, and human maximisation test to show skin sensitisation in mammals. Potential mechanisms reported

are for quinone; pre/prohapten producing an electrophilic Michael acceptor.[110] Phenolic radical potential mechanism is that of pre/prohapten producing a free radical.

Although Phenol itself has given a negative result for skin sensitisation in its LLNA test in mice,[74] Phenol derivatives and substituted phenols have been reported to show positive results, such as 4-styrylphenol has given a strong positive result in the Maximisation test with Guinea pigs, 2,4,5-trichlorophenol tested positive in LLNA.[111]

For the complex an EC3 prediction is found at 0.48% as a strong sensitiser, similar matches have been found for substituted phenols, as weak sensitisers.

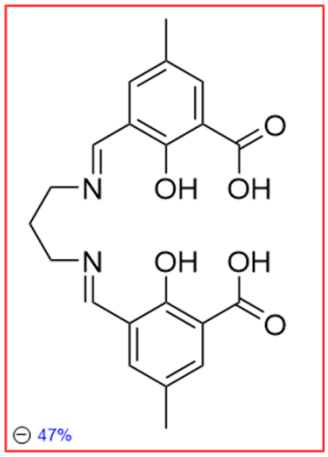


FIGURE 23: LIGAND TAKEN FROM SCHEME GARDEN *ET AL*, ROP IN LACTIDE POLYMERISATION STRUCTURES [106]

This structural ligand drawn above, has highlighted the potential toxic effects of heterometallic complex synthesis which can be incorporated into various metallic catalysts sites. This ligand has shown common red alert endpoints in Derek Nexus. Hepatotoxicity as plausible, and skin sensitisation as equivocal, matched to the results above for C1 and C1c.

The first alert describes skin sensitisation potential by substituted phenol analogues, although phenol itself gives a negative result for skin sensitisation in animals, derivatives of phenol have the potential to cause toxicity as positive results have been reported. In LLNA derivatives including 2,5-dimethyl phenol, 3,4-dimethyl phenol.[112]

Chemical found for this alert is *N,N*'-(2,2-dimethylpropylidene)hexamethylenediamine, species and assay have been unspecified by Derek, result showed GHS Category 1 for skin sensitisation.

For this ligand constructed, reasonings provided suggest that hepatoxicity is plausible for a chemical when alert 666- is presented in the results. Hepatotoxicity can be caused in mammals and is possible in bacteria.[113] Hepatoxicity is caused by toxicity to normal liver function, there is limited human data. EC3 result for skin sensitisation is 2.9%, moderate sensitiser.

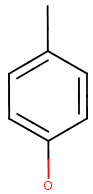
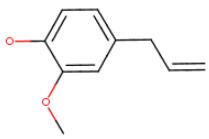
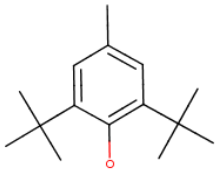
In experimentation with animals' low glutathione levels, and high dosages of para-alkylphenols were found to have a considerable impact on liver function enzyme levels, as well as produce centrilobular necrosis and cholestasis. Normal cell function can be altered by toxicological mechanisms, such as metabolic activation of quinone methide reactive intermediates.

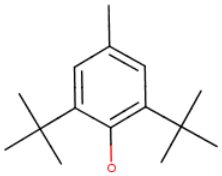
Reactive quinone methide metabolites are found to be involved in acute hepatotoxicity of p-cresol, eugenol and 4-hydroxyphenylacetone. These metabolites can form adducts with macromolecules altering the cellular functions leading to cell death.

Studies have shown that the stability of formed quinone methide should be considered, in the rat liver model, metabolites of para-alkyl-2—methoxyphenols with half-lives of 10 seconds to 10 minutes, those with shorter half-lives were less cytotoxic, this is due to them either reacting before reaching the critical cellular nucleophiles or being too stable to react.

.

TABLE 10: SUMMARY OF MATCHED EXAMPLES FOR ALERT 666-PARA-ALKYLPHENOL OR DERIVATIVE

Chemical name	Structure	Assay	Endpoint	Result /Reference
P-cresol		Hepatotoxicity case report Human	Hepatotoxicity	Positive Green MA. (1975) report)., *Medicine, Science, and the Law*, 15, 65-66 … Kamijo Y, Soma K, Kokuto M, Ohbu M, Fuke C and Ohwada T. (2003) Hepatocellular injury with hyperaminotransferasemia after cresol ingestion., Archives of Pathology and Laboratory Medicine, 127, 364-366
Eugenol		Hepatotoxicity study Mouse	Hepatotoxicity	Positive
butylated hydroxytoluene		Hepatotoxicity study Rat	Hepatotoxicity	Positive Hepatotoxicity study data reported by Mizutani *et al* relate to testing in animals with depleted glutathione. Mizutani T, Satoh K, Nomura H and Nakanishi K. (1991)

				Hepatotoxicity of eugenol in mice depleted of glutathione by treatment with DL-buthionine sulfoximine., Research Communications in Chemical Pathology and Pharmacology, 71, 219-230
butylated hydroxytoluene		Hepatotoxicity study Rat	Hepatotoxicity	Powell CJ, Connelly JC, Jones SM, Grasso P and Bridges JW. (1986) Hepatic responses to the administration of high doses of BHT to the rat: their relevance to hepatocarcinogenicity., Food and Chemical Toxicology, 24, 1131-1143 DOI: 10.1016/0278-6915(86)90299-1

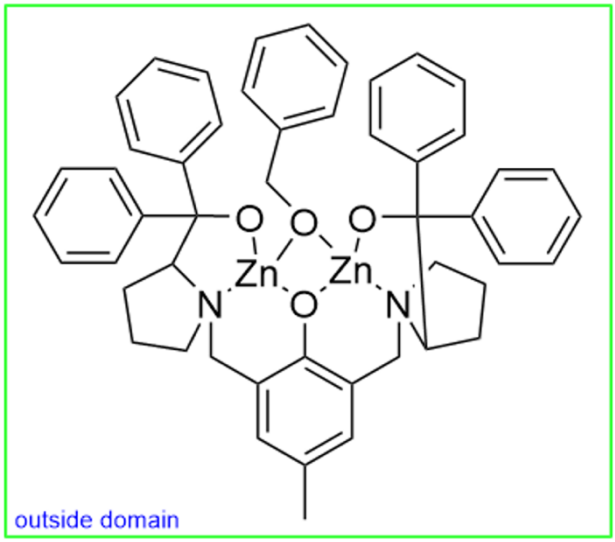


FIGURE 24: CATALYST FROM GARDEN *ET AL*, SCREENED USING DEREK NEXUS AND SARAH NEXUS AND HIGHLIGHTED GREEN USING TRAFFIC LIGHT SYSTEM [114]

This macromolecular structure shown LZn_2OBn catalyst, has proven to show excellent catalytic activity for a range cyclic esters, by the works of Garden *et al* it has been reported to be the fastest catalyst for poly(ε-caprolactone-block-lactic acid).[114] Also it is seen to be the first catalyst to selectively prepare poly(ε- caprolactone-block-lactic acid-block-β-butyrolactone) in a one- pot synthesis, this has further lead to production of a range of block polyesters.[115]

LZn_2OBn, however, has a single starting group. The high activity seen here may result from the metal cooperativity, where one Lewis acidic in Zn centre coordinates a cyclic ester and the other Zn carries the alkoxide group (OBn or the polymer chain) to carry out nucleophilic attack on the coordinated monomer. A possible way to take use of metal–metal cooperativity and achieve high activity and careful control over the polymer structure is through the creation of bimetallic, monoinitiator catalysts.[116]

Derek Nexus reports have not fired any alerts for this catalyst, however reasoning details show mutagenicity as inactive, as substructures within this macromolecule do not match any alerts in the Derek structural database therefore predicted as inactive in the bacterial *in vitro* (Ames) mutagenicity test. This structure also shows no matching structural examples for skin sensitisation.

Fragments from these structures could not be found in the Sarah Nexus training data library, this has resulted in an out of domain prediction.

3.2.3.7 ZINC AND MAGNESIUM ALKOXIDES

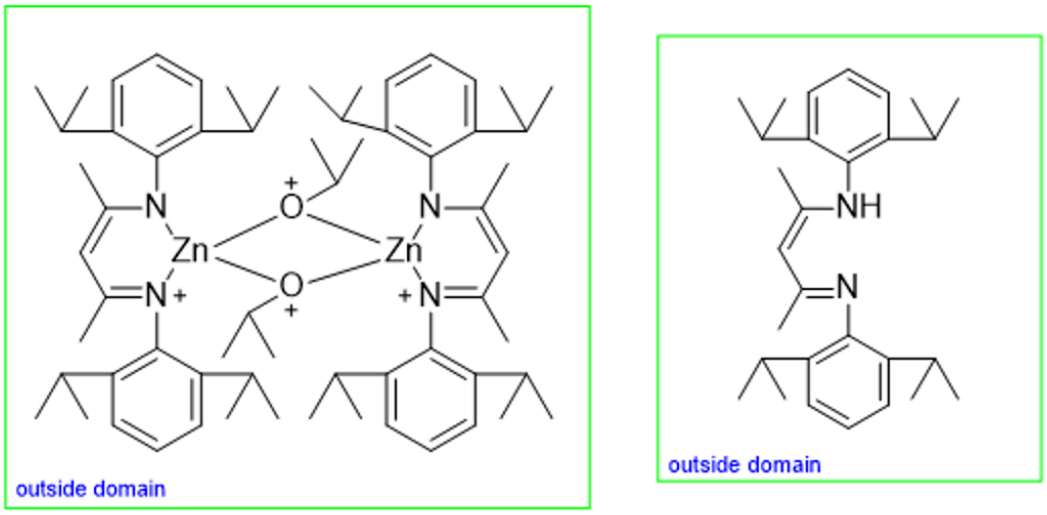


FIGURE 25: METAL ALKOXIDE CATALYSTS STRUCTURES SCREENED USING DEREK NEXUS AND SARAH NEXUS, HIGHLIGHTED GREEN USING TRAFFIC LIGHT SYSTEM[117]

Coates *et al* studied a series of Zinc and Magnesium alkoxides. Molecular structure of catalyst rac-(BDI-1)ZnOCH(Me)CO_2Me for lactide polymerisation applications was shown to be desirable through experimentation with these catalysts.[118]

Structures screened through Derek Nexus have indicated safe usage of this catalyst, no alerts have been fired for this set. The ligand is overall shown to be highlighted in the green area, however skin sensitisation in mammal is non-sensitiser as this molecular catalyst is a macromolecular structure and some features have not been identified by the Lhasa skin sensitisation negative prediction dataset. As unclassified features are present, further verification would be required for a valid assessment,

An outside domain outcome has been predicted by Sarah Nexus for the mutagenicity *in vitro* endpoint. This means at least one atom which is present in the fragment of compound submitted

Structures of the family of 4 zinc complexes studied as rac-LA ROP catalysts herein: amido complexes [Zn_2L Et$(HMDS)_2$] 1, [Zn_2L Pr$(HMDS)_2$] 2, and [ZnL Open(HMDS)] 3 and the corresponding alkoxides [Zn_2L Et$(OiPr)_2$] 4, [12] [Zn_2L Pr$(OiPr)_2$] 5, and [ZnL Open$(OiPr)_2$] 6.[123]

Williams *et al*, showing structure activity of bi-metallic complexes, 4 macromolecular structures taken from the works and screened using Derek Nexus and Sarah Nexus.

These complexes are desired in ROP polymerisations, as dizinc complexes are stronger catalysts, the ligands are also highly electron donating, shown to be more stable than monozinc catalysts.[122]

These macromolecular structures can be seen to be green; no alerts have been fired, mutagenicity is seen as inactive and overall, non-sensitisers for these showing safe usage and disposal. Reports have highlighted mutagenicity *in vitro* bacterium as inactive, and skin sensitisation in mammal as non-sensitiser. N.B. Organocatalysts are easier to screen than organometallic catalysts.

3.2.3.8 CHALLENGES FOR ORGANOMETALLIC CATALYSTS

Derek Nexus provided interesting results for this data set, however Sarah Nexus did not meet the expectation of giving reliable results, due to the nature of the structures. As these catalysts contained complexities for the software, as the structures were too large to be screened gave outside domain results. This is something to be flagged with Lhasa as an improvement for the future. This will give versatility and better projectability for catalytic polymer applications.

3.3 PHOTOINITIATORS

Photopolymerisation uses light (typically UV-visible) to initiate polymerization reactions which can be used to form linear and crosslinked polymer structures, which can be employed for additive manufacturing which is a key underpinning technology of the current industrial revolution (Industry 4.0)[124] this can take place either via single-photon absorption by an initiator (typically a photon with a short wavelength [e.g. UV] through linear absorption), or via a two-photon process, wherein the initiator absorbs two near infrared (NIR) photons with a long wavelength through nonlinear absorption. 1-photon polymerisation (1PP) underpins classic stereolithography, whereas 2-photon polymerisation (2PP) underpins direct laser writing (DLW). Absorption processes mean that 1PP typically occurs at surfaces/interfaces of samples and the resin/ink yields planar structures (which can be grown into 3D structures by an automated process with moving surfaces/interfaces); by comparison, NIR femtosecond laser pulses can be focused into the bulk of a sample loaded with resin/ink enabling 3D structuring.

3.3.1 PHOTOINIATOR STRUCTURES

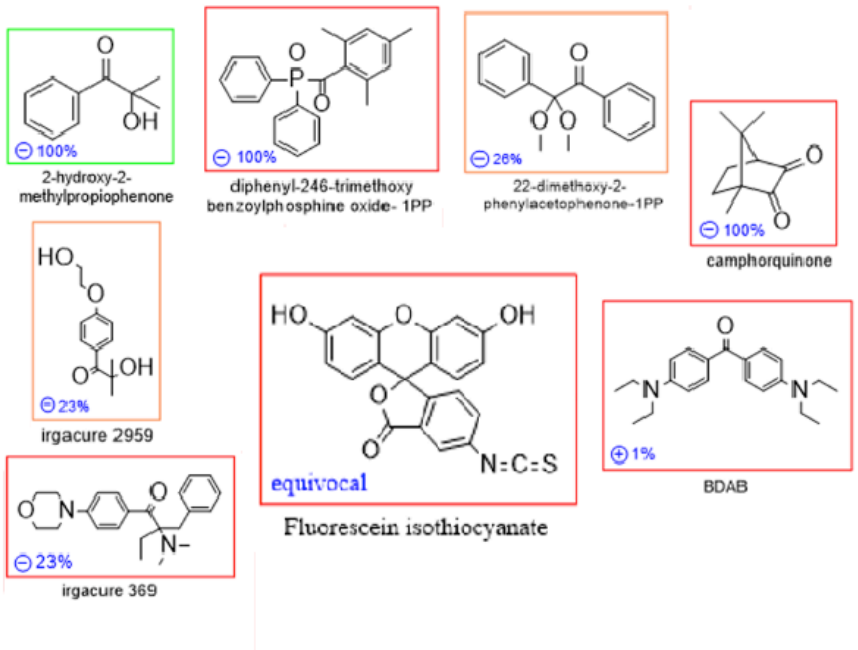


FIGURE 26: PHOTOINITIATOR STRUCTURES PRESENTED ABOVE USING TRAFFIC LIGHT SYSTEM, SCREENED WITH DEREK NEXUS AND SARAH NEXUS

There is a vast array of photoinitiators in the literature (some of which are commercially available), however, the range of structures precludes a comprehensive study of their toxicities, and a few examples of 1PP and 2PP initiators are highlighted. Examples of commercially available 1PP initiators include: Camphorquinone (used in dental composites and 3DP,[125] 2-hydroxy-2-methylpropiophenone, 2,2-dimethoxy-2-phenylacetonephenone, diphenyl-2,4,6-trimethoxybenzoylphosphineoxide. Examples of 2PP initiators include organic solvent soluble benzoquinones,[126] anthraquinones,[127] fluorenones – b3fl[128], indeed water soluble initiators (e.g. Irgacure 2959,[129] Irgacure 369,[130] which can be used to prepare biomaterials (e.g. hydrogels); in some cases (e.g. P2CK, G2CK, BSEA) in the absence/presence of presence of living cells.[130-133]

3.3.1.1 RED STRUCTURES

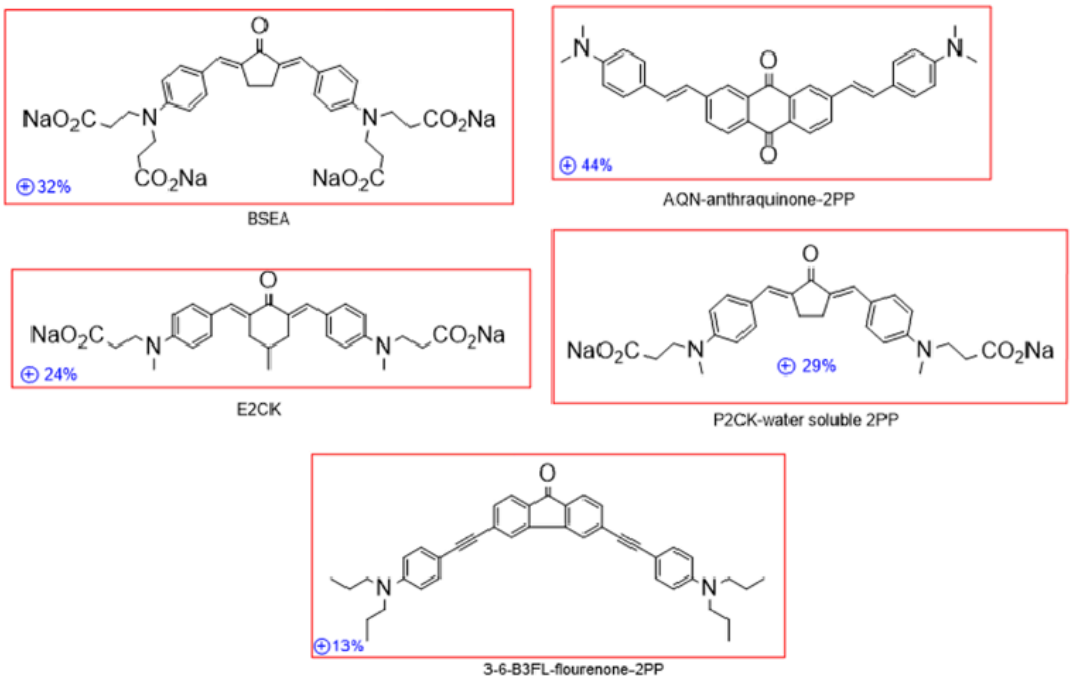


FIGURE 27: PHOTOINITIATORS SCREENED USING DEREK NEXUS AND SARAH NEXUS SHOWING RED ALERTS

3.3.1.2 CYCLIC PHOTOINITIATORS STRUCTURES

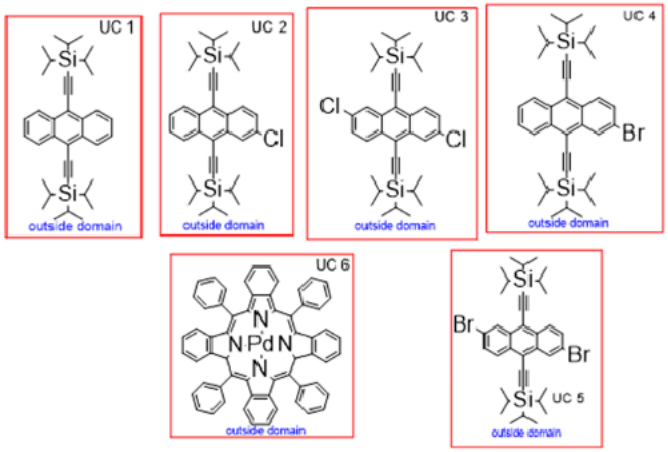


FIGURE 28: METAL CONTAINING CYCLIC PHOTOINITIATORS SCREENED USING DEREK NEXUS AND SARAH NEXUS

Photoinitiators studied above employ a range of toxic properties.[134] Multiple alerts have been found for each query structure. This data highlighted the highest number of alerts per query structures screened in both Derek Nexus ad Sarah Nexus.

Within this data set the highest numbers of positive alerts in Sarah Nexus have also been found. 2-hydroxy-2methylpropiophenenone is considered safe in comparison to the data set, highlighted in green with a negative Sarah Nexus result 100%.

Silicone based photoinitiators are well known, due to the excellent thermal and chemical properties displayed by this metal. Silicone is also used as an additive in polymer materials to improve the thermal characteristics.[135] Known toxicities screened by Derek Nexus are Hepatoxicity, Mutagenicity and Nephrotoxicity.

Derek Nexus reports for these structures have been summarised below.

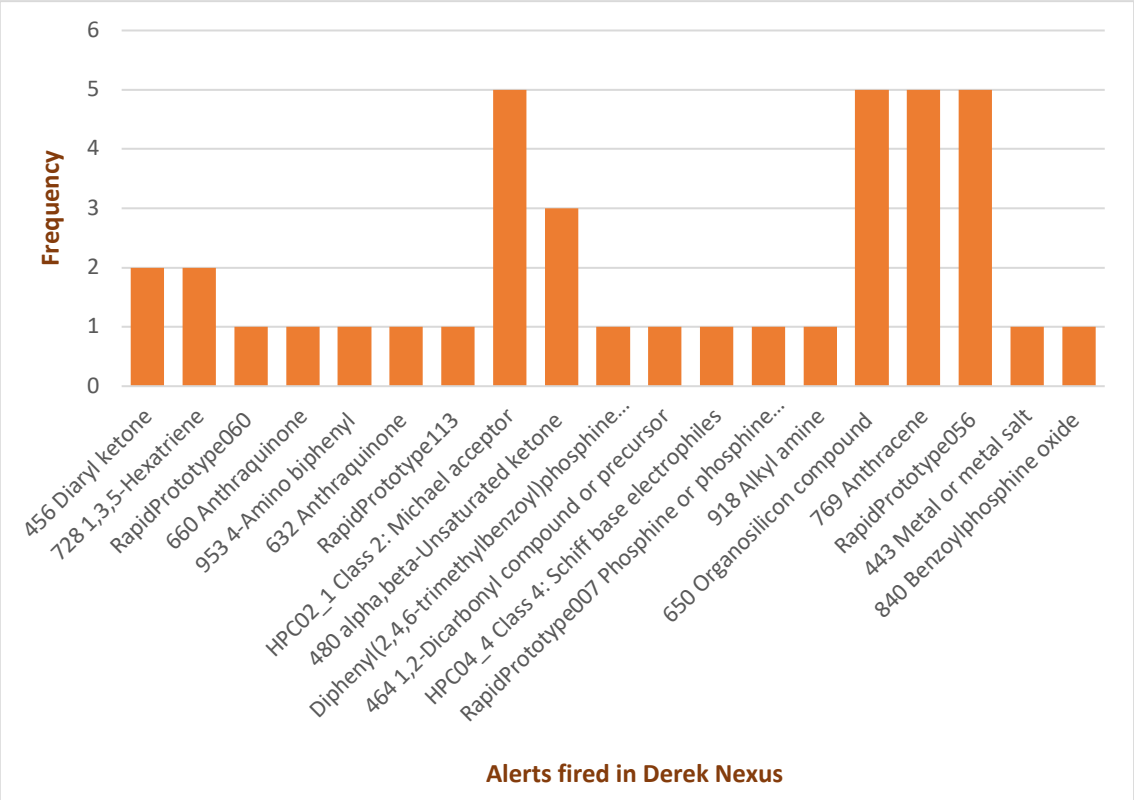


FIGURE 29: SUMMARY OF TYPES OF ALERTS FIRED THROUGH DEREK NEXUS AND THE FREQUENCY OF OCCURRENCE FOR PHOTOINITIATORS DATA SET

Various alerts have been fired by photoinitiators query structures, most frequent alerts are 650 Organosilicon compound, 769 Anthracene, Rapid prototype 056 and HPC02_1 Class 2: Michael acceptor, followed by 48- alpha, beta-unsaturated ketone. 456 Diaryl ketone and 728 1,3,5-Hexatriene alerts were fired twice across the dataset.[136]

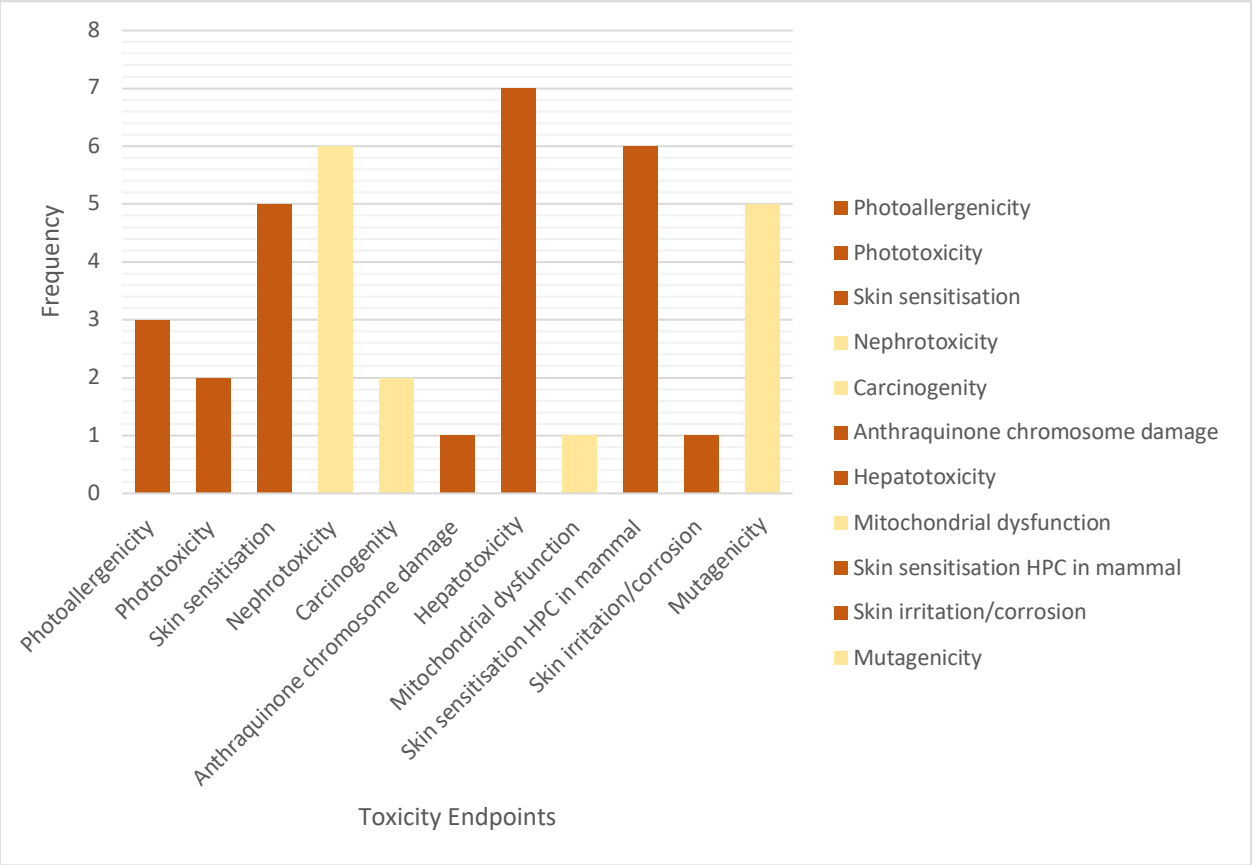


FIGURE 30: THE FREQUENCY OF EACH ENDPOINT FIRED IN PHOTOINITIATOR DATA SET

Most common endpoints are Hepatotoxicity, Skin sensitisation and nephrotoxicity. Many regulatory organisations are enforcing laws and limits on the chemical usage in industries, including EURL ECVAM), implementing a highly integrated process of safety and risk assessment techniques to ensure compliance. From the illustration above, it is clear that metal-salt derivatives are showing higher toxicity levels than respective metals.

Irgacure 369-2PP fired alerts in Derek Nexus for skin sensitisation/irritation as plausible, whilst irgacure 2959-2PP fired no alerts, however carcinogenicity in mammals is open, substructures in the query structure screened have potential toxic properties which can cause cancer.

For this set of data, a common toxicity endpoint is photoallergenicity, this has occurred for 3-6-B3FL - fluorenone - 2PP, AQN - anthraquinone derivatives, BDAB-bisdiethylaminobenzophenone - 2PP and UC species 6. Diaryl ketones and their cyanoacrylate precursors have the potential to cause photo allergenic reactions, this is based on human photo patch tests.[137] Photoallergenicity takes place via free radical reactions, this depends on the substituents present on the ring systems. The features of light absorption and photochemical behaviour of the chemicals that

contain the diaryl ketone chromophore are pH dependent since many of these compounds also contain ionising groups.[152]

Nephrotoxicity is another endpoint re-occurring for this set of Photoinitiators, 2,2-dimethoxy-2-pheynylacetonephenone-1PP, UC species 1-5. A rapid prototype alert described for methylene glycol and its derivatives, data taken from 731 chemicals and studied histopathological lesions over 28 days with rodents.[153]

Hepatotoxicity was presented in 7 photoinitiators species, being the most common endpoint for compounds containing the anthraquinone group, these compounds are known to cause hepatocellular membrane damage and liver damage, both seen in experimentation with animals and humans.[138]

Skin sensitisation and Mutagenicity are default endpoints studied by Derek Nexus, Mutagenicity was inactive for most species, except the UC species 1-5 where mutagenicity has been plausible in bacteria.

Positive Sarah results are displayed above, with higher levels of toxicities and positive results shown through mutagenic Ames test.

3.3.1.3 FLOURESCEIN ISOTHIOCYNATE (FITC)

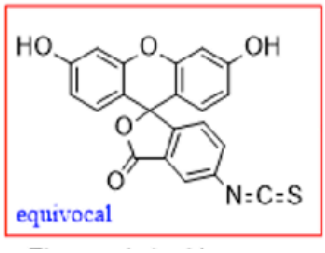


FIGURE 31: FLOURESCEIN ISOTHIOCYNATE(FITC)

The molecule fluorescein isothiocyanate (FITC) would be considered a red alert through the traffic light system. The alerts fired for this are 304 Isocyanate or isothiocyanate, which cause chromosomal damage, clastogenicity seen through Mutagenicity Ames test.

From the Derek Nexus screening report, several toxicities have been highlighted as plausible. Chromosomal damage, mutagenicity *in vitro* bacterium, phototoxicity and skin sensitisation. Alert matched to chromosomal damage *in vitro* as plausible is 304 Isocyanate or isothiocyanate.

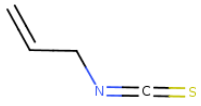
Chromosomal (clastogenicity) seen via *in vitro* chromosome aberration test, and Mutagenicity through the Ames test. There is a potential for DNA interaction, methyl isocyanate and phenyl isocyanate is demonstrated to reach with exocyclic amino group of DNA bases deoxyadenosine, deoxyguanosine and deoxycytidine.[139]

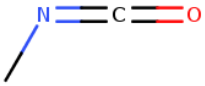
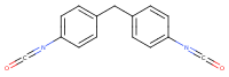
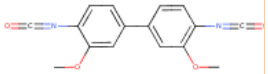
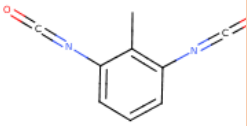
Yamaguchi tested a number of isocyanates and isothiocyanates, in bacteria *Salmonella typhimurium* and found a positive mutagenicity response for all isothiocyanates compounds, with a potency degree. In the Ames test Methyl isocyanate has also shown to be mutagenic.[140]

Rihova *et al*, has also reported mutagenicity of allyl isothiocyanate in *Escherichia coli* with activation only. It was found that allyl isothiocyanate is only mutagenic when activated in *E. coli*. However, in this instance, mixed function oxidases containing cytochrome P-450 were not believed to be the cause of the transformation of allyl isothiocyanate into mutagenic products.[141]

For alert 304 – Isocyanate or isothiocyanate, chemicals allyl Isothiocyanate[142], methyl isocyanate[143], 4,4'-diphenylmethane diisocyanate[144], 3,3'-dimethoxybenzidine-4,4'-diisocyanate and 2,6-diisocyanatotoluene all showed positive result during *in vivo* experimentation in species selected hamster, *in vitro* chromosome aberration test showing endpoint chromosome damage *in vitro*. Below in Table 11 is a summary and referenced sources matched to the alerting chemical.

TABLE 11: SUMMARY OF STRUCTURAL ALERTS MATCHING ALERT 304 – ISOCYANATE

Chemical name	Structure	Assay / species	Endpoint	Result /comments
Allyl Isothiocyanate		*In vitro* chromosome aberration test Hamster	Chromosome damage *in vitro*	Positive

Methyl isocyanate		*In vitro* chromosome aberration test Hamster	Chromosome damage *in vitro*	Positive
4,4'-diphenylmethane diisocyanate		*In vitro* chromosome aberration test Hamster	Chromosome damage *in vitro*	Positive
3,3'-dimethoxybenzidine-4,4'-diisocyanate		*In vitro* chromosome aberration test Hamster	Chromosome damage *in vitro*	Positive ,
2,6-diisocyanatotoluene		*In vitro* chromosome aberration test Hamster	Chromosome damage *in vitro*	Positive

Alert 410 has been matched with skin sensitisation for FICT, in mammals such as guinea pig Buehler test, guinea pig maximisation test (GPMT). The potential mechanism is Haptem acting as the electrophilic thiocarbamylating agent. Isocyanates are electrophilic, attaining the potential to react with skin proteins and DNA and depleting cellular and physiological functions, and forming dithiocarbamate derivatives.[145]

In numerous skin sensitization experiments, the isothiocyanates have shown skin sensitization action.[140-144] Examples include allyl isothiocyanate in the guinea pig Buehler assay, rhodamine B isothiocyanate in the mouse ear swelling test, and phenyl isothiocyanate in the GPMT. Thioglucoside of allyl isothiocyanate was integrated with enzyme myrosinase and mixed with petrolatum, this produced a positive reaction, immunochemical toxic reaction occurred for plants

of the Cruciferae, mustard plant family.[146] Although, shown to be allergens to horticultural plants, allergic contact dermatitis cannot be determined.

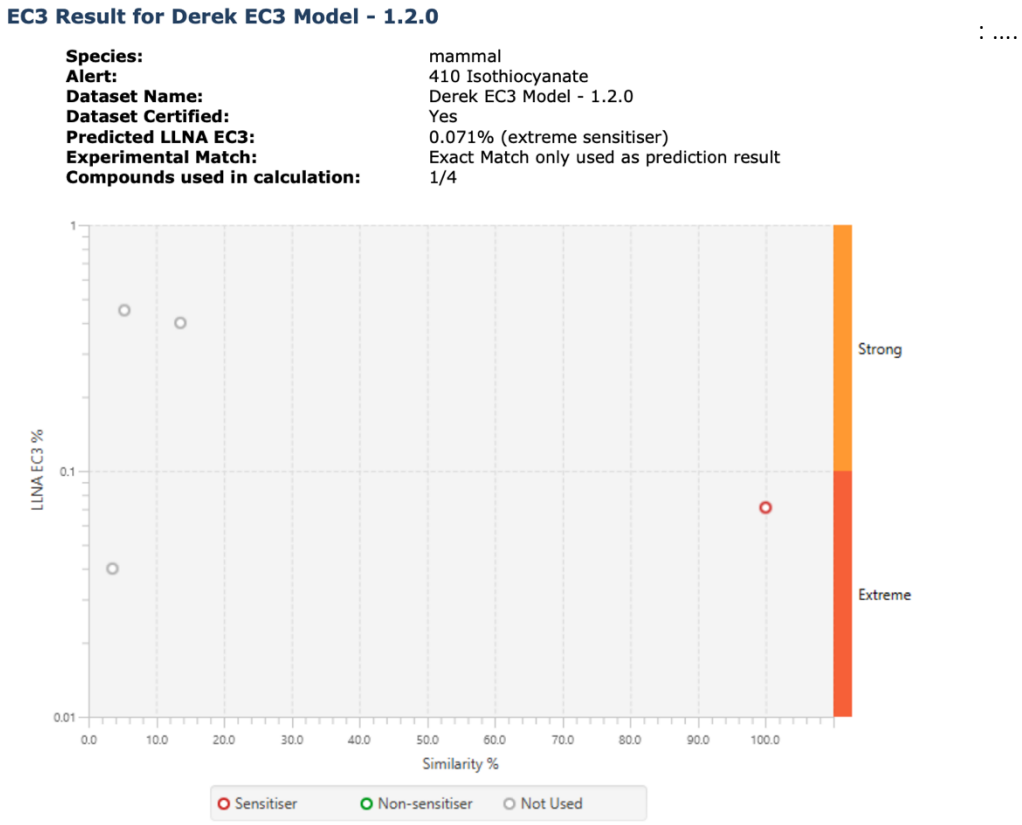


FIGURE 32: SHOWING EC3 PREDICTION FOR SKIN SENSITISATION 410 ISOTHIOCYANATE[26]

Above is an example of EC3 prediction showing 0.071% as an extreme sensitiser from the predicted LLNA value. Skin sensitisation has been reported positive through *In vivo* experimentation through the Guinea pig Buehler test, chemicals matched to this alert was Allyl Isothiocyanate.[147]

Phenyl isothiocyanate[148], also showed a positive result in the Guinea pig maximisation test. Rhodamine B isothiocyanate is an example that also matched, *in vivo* experimentation Ear swelling test in mice. FITC increased the number of dendritic cells of lymph nodes and initiated a proliferative *in vitro* response, dictated a positive result for skin sensitisation as the endpoint.[149]

Phototoxicity endpoint was highlighted for FITC, the alert matched was 763 Fluorescein or derivative.[149] This alert describes toxicity of fluorescein-like xanthene dyes, photochemical

studies of xanthene dyes such as fluorescein (FL), rose Bengal (RB) and erythrosine.[150] A chemical has the potential to generate systemic or localised phototoxic reactions if it has a phototoxic structural elements present. The molecule's photochemical characteristics will also determine whether it is phototoxic or not. Since short-wavelength UV radiation has a harder time penetrating skin, chemicals that absorb at wavelengths greater than 290 nm are more likely to be phototoxic.

Due to the limited prevalence of xanthene dyes as singlet oxygen photosensitisers in photodynamic therapy, there have been a few reports of photosensitisation reactions.[151]

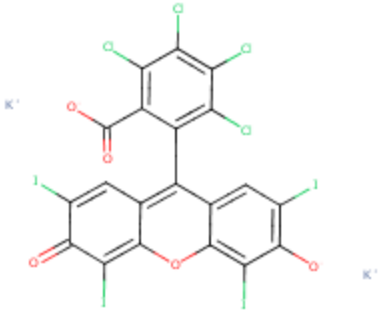


FIGURE 33: ROSE BENGAL CAS NUMBER: 11121-48-5 TAKEN FROM DEREK NEXUS REPORT FOR FITC [150]

The structure above is Rose Bengal, studied by Wiener *et al*., a photosensitisation case report in humans indicated a positive result for phototoxicity. A similar study also showed a positive result through an *in vitro* photohaemolysis test in humans.[152] The final study for Rose Bengal was undertaken in mice, a 3T3 NRU phototoxicity test which also gave a positive result.

These studies have shown that Rose Bengal creates a considerable amount of singlet oxygen and exists in its triplet state and can cause significant photo-damage under certain exposure to visible light radiation.

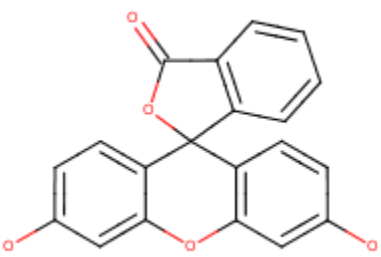


FIGURE 34: FLUORESCEIN TAKEN FROM SARAH NEXUS REPORT

Fluorescein tested positive for humans in a photosensitisation case report, determining phototoxicity. Therefore, from the data it can be concluded that phototoxicity in mammals is plausible. The Sarah Nexus prediction for FITC has given an equivocal for the prediction at 8% level, therefore a compelling argument cannot be made based on this assumption.

3.3.1.4 CHALLENENGES FOR PHOTOINITIATORS

This data set has presented the highest number of alerts for Derek Nexus, and positive results for mutagenicity in Sarah Nexus. Several toxicities have been found, these can be matched to known literature. Photoinitiators have shown interesting results, many endpoints proving the application of the software and ability to flag up toxicities successfully. Metal containing photoinitiators UC1-UC6 species showed multiple red alerts, hepatotoxicity, entered into the software and further expert review would be required. Contradiction in results is common and can be seen for FITC's alerts. For the same substrate negative results have also been reported, this could be due to instability and high bacterial toxicity associates with cyanates.[153]

4. CONCLUSION

From this research it can be concluded that *in silico* toxicity screening and the use of the Lhasa suite is beneficial and promising for medical advancement, pharmaceutical drug development and environmental sustainability. As discussed, as a cost-effective and time saving tool it is proven to function as an extremely useful software, with the ability to screen large sets of data rapidly.

This study has successfully demonstrated the use of toxicity prediction tools for environmental and green chemistry, the software is designed to work for pharmaceutical industries as well as chemical industries as the works are complimentary to both. An excellent benefit is for REACH, and health and safety regulatory bodies, defining the levels of safety for marketed chemicals, this tool is highly favourable in allowing foundations and regulations to be put in place for a safer environment and safer practices for chemical industries.

The experimental results for solvents for polymerisation in this research has proven the ability and application of Derek Nexus and Sarah Nexus. From the low molecular weight solvents section multiple toxicities have been flagged up by the software, such as DMI, NMP, NMF and TMP. Derek Nexus has shown skin sensitivity for common laboratory solvents such as HMPA, DMSO, DMPU, and 3MOx. Solvents for processing perovskites have presented a positive outcome for the applications of green solvents, with structures giving green alerts, shown to be less toxic and greener. Yet, there are known toxicities for common solvents here, for example cyclohexanone and dimehylhexanone which cause skin and eye irritation that have not been flagged up. Potential toxicities have not been detected by Derek Nexus software, which have been previously stated in MSDS's and handling procedures. Ionic liquids in this solvents section have shown fewer toxicities, proving the desirability over conventional solvents.

Catalysts screened for polymerisation have shown a comparable pattern within the data for tests using Derek Nexus, for heterogenous catalysts metal toxicities have been highlighted and alerts have shown hazardous effects. Safer catalysts have also been highlighted showing the potential application of the software. Nanozymes screened provided no alerts for toxicities, raising concerns of ambiguity and reliability of results. Homogenous catalysts showed interesting set of results, toxicities including hepatotoxicity and skin sensitisation. However, results for catalysts screened using Sarah Nexus have displayed out of domain outcome, due to the complexities of drawing and processing macromolecular structures, this would need further research and input from Lhasa Ltd.

The final section for photoinitiators species showed the highest number of alerts for toxicities in Derek Nexus, common endpoints for skin sensitisation, nephrotoxicity, photo-allergenicity and phototoxicity. Demonstrating the application and use of the software for simpler structures. This section also showed the highest number of positive results for mutagenicity in Sarah Nexus.

Both software Derek Nexus and Sarah Nexus have shown excellent features and benefits which have been illustrated throughout this research, however for this *in silico* toxicology screening method there is a need for a reliability and accuracy assessment. This should be conducted by experts in the field, these methods cannot be relied on completely without professional validation and quality control. Even though, these computational methods are facilitating gaps in the market and needs for animal testing there is still a clinical need to evaluate target compounds further.

For chemical compounds with features that cannot be matched by Derek Nexus, containing unclassified features, this would be considered a flaw as the software is not able to make confident predictions. Complex structures screened with Derek Nexus, allows the main feature to be assessed, a toxicity prediction can be made for each based on the substructures found, the level or intensity of toxicity is not highlighted, therefore can only be considered a guide. Another challenge is predicting repeated dosages and lower doses, as well-defined mechanism and adverse outcome pathway can vary between exposure limits.[156]

For Derek Nexus report results, the term 'equivocal' used to describe a prediction has various ambiguous meanings, it is used when the prediction cannot be resolved as positive or negative.[154] A misclassified feature in Derek Nexus given for a negative result, does to quantify the extent to which the prediction is valid, as misclassified compounds are found in Lhasa's Ames test reference set, and active in the Ames test, therefore reducing the confidence in the prediction.

From a clinical perspective, it would be beneficial if pharmaceutical companies collaborated with Lhasa, to provide them with confidential data for research purposes to re-design, implement and develop a methodology whereby structures are inputted into the databases for a higher clinical use. Guarding the confidentiality of data is empirical to research, it is to be ensured that similar features of these toxicants are imputed for easy retrieval in the software.[157]

In assessing the performance of Derek Nexus and Sarah Nexus, it should be noted that skin sensitisation predictions have not considered tautomeric forms, or individual components in the mixtures. No account is made for other toxic endpoint alerts, which could be present for the compounds, these would be considered weak sensitisers.

Sarah Nexus as a software that gives publicly available data, this is a great tool for scientists to find access to toxicological information within seconds of searches. As there is adequate information provided for a search to be conducted and initialised for research purposes. Sarah gave positive results for mutagenicity successfully for solvents, and photoinitiators, for heterometallic complexes Sarah Nexus has been unable to give reasonings or predictions, the inability to predict larger complex molecules causes a concern for further research and development into macromolecules. A set of complex substructures should be incorporated and developed into Sarah Nexus's software to improve the reliability of predictions and minimising out of domain results.

This research has successfully demonstrated the ability of *in silico* toxicity screening methods, i.e. the use of Derek Nexus and Sarah Nexus in identifying toxicities in solvents, catalytic species, organometallic and photoinitiators. Data from each category showed a similar toxicity profile, pattern in resulting endpoints and sensitisation predictions. The data has also highlighted the positive outcome for *in silico* methods for screening a range of compounds, which can further be applied to different industries and uses for various drugs and chemicals. Even though, It is clear that there is a need for further expert reviews, the reliability and credibility of this method must be critically assessed as required. *In silico* toxicity screening approach is dynamically growing, improving toxicity assessment methods, with the potential to reduce animal testing, and work as a beneficial computational tool for industries to create safer, environmentally greener products.

5. FUTURE WORK

The future of *in silico* toxicity screening as a tool for research is a fast-growing positive approach for multiple applications, there is a holistic need for further investigations into larger complex molecules. Identifying intricacies within sub structures submitted is also necessary, this would provide a precise toxicity prediction. Derek Nexus and Sarah Nexus have successfully shown the ability of computational methods to enhance toxicity screening. It would be interesting to explore more commercially available QSAR tools for enhanced toxicity predictions, such as degradation pathways using Zeneth, Vitic, QSAR toolbox.

There is a need for further research into Lhasa's software's Derek Nexus and Sarah nexus, to improve the reliability of negative predictions by combining the predictions from Sarah Nexus. Also, development of the software to access multiple structures simultaneously. Sarah Nexus to exhibit more complex structures within their literature databases, reducing misclassified and unclassified features presented in predications. Derek Nexus to define each alert to the targeted structure rather than generalising clinical use applied to cancer therapeutics.

Further research into green solvents used for ROP should be exhibited, organic solvents, ionic liquids and supercritical fluids are proven advantageous, but still a challenge a to purposefully develop and optimise non-toxic, thermally and chemically stable solvents for polymerisation. Further investigations into the design and synthesis of red-flagged solvents, explore ways to re-design and improve properties for safer applications. Also, using a greater database for a larger study to determine higher reliability of predicted results.

Heterometallic catalysts have shown excellent performance in lactide ROP, improving the catalyst design by incorporating smaller more selective ligand complexes could improve the efficacy of the polymerisation process. Salen catalysts have shown excellent stereo control, they require tailoring at higher temperatures and long reaction times, further investigations and careful control would lead to better yield of catalysts.

Research into re-designing photoinitiators as more favourable species for greener chemistry, improve chemical properties through experimentation. Improve work on perovskites solar cell applications highlighting the exciting potential in low-cost high production.

6. REFERENCES

1. N. Kleinstreuer, I. Tetko and W. Tong, *Chemical Research in Toxicology*, 2021, 34, 171-175.
2. P. Grandjean, *Basic & Clinical Pharmacology & Toxicology*, 2016, 119, 126-132.
3. A. Raies and V. Bajic, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2016, 6, 147-172.
4. J. Dearden, *Journal of Computer-Aided Molecular Design*, 2003, 17, 119-127.
5. M. Cronin and M. Yoon, *The History of Alternative Test Methods in Toxicology*, 2019, 287-300.
6. *Ec.europa.eu*, 2022, https://www.epa.gov/greenchemistry/basics-green-chemistry#definition
7. A. Ivanković, International Journal of Sustainable and Green Energy, 2017, 6, 39.
8. K. Brake, A. Gumireddy, A. Tiwari, H. Chauhan and D. Kumari, Pharmaceutica Analytica Acta, 2017, 08.
9. T. Bjornsson, J. Callaghan, H. Einolf and V. Fischer, *Drug Metabolism and Disposition*, 2003, 31, 815-832.
10. C. Hasselgren and J. Bercu, *Regulatory Toxicology and Pharmacology*, 2020, 118, 104807.
11. T. Kennedy, *Drug Discovery Today*, 1997, 2, 436-444.
12. D. Presa, Derek Nexus 6.0.1, Program Training, powerpoint, 2020
13. S. Shityakov, N. Roewer, J. Broscheit and C. Förster, *Computational Toxicology*, 2017, 2, 20-27.
14. Medical Device Network, Nanotechnology in Medicine, Technology Trends, 2020, https://www.medicaldevice-network.com/comment/nanotechnology-medicine-technology/
15. S. Soares, J. Sousa, A. Pais and C. Vitorino, *Frontiers in Chemistry*, 2018, 6.
16. D. Clive, *Regulatory Toxicology and Pharmacology*, 1985, 5, 79-100.
17. S. Sukumaran, R. Ranganatha and S. Chakravarthy, *International Journal of High Throughput Screening*, 2016, 1.
18. Williams, *Regulatory Toxicology and Pharmacology*, 2016, 76, 79-86.
19. M. Chilton and D. Macmillan, *Regulatory Toxicology and Pharmacology*, 2018, 95, 227-235.
20. *Oecd-ilibrary.org*, 2022, https://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals-section-4-health-effects_20745788
21. OECD iLibrary, Guideline no.497, 2021, https://www.oecd-ilibrary.org/docserver/b92879a4-en.pdf?expires=1656765305&id=id&accname=guest&checksum=8A852F79DE2ED1BBF85869DCACB6D9F2
22. T. Ohtake, Y. Maeda, T. Hayashi, H. Yamanaka, M. Nakai and M. Takeyoshi, *Toxicology*, 2018, 393, 9-14.
23. Lhasa Limited, https://www.lhasalimited.org/
24. Lhasa limited, Sarah Nexus 3.0, Program Training, 2020.
25. C. Hasselgren, J. Bercu, A. Cayley, K. Cross, S. Glowienke, N. Kruhlak, W. Muster, J. Nicolette, M. V. Reddy, R. Saiakhov and K. Dobo, *Regulatory Toxicology and Pharmacology*, 2020, 118, 104807.
26. Derek Nexus v.6.0.1 software system,
27. A. Zhenova, *Polymer International*, 2020, 69, 895-901.

28. E. Bisz, M. Koston and M. Szostak, *Green Chemistry*, 2021, 23, 7515-7521.
29. L. A. Malley, G. L. Kennedy, G. S. Elliott, T. W. Slone, W. Mellert, K. Deckardt, K. Kuttler, B. Hildebrand, M. I. Banton, R. J. Parod and J. C. Griffiths, *Drug and Chemical Toxicology*, 2001, 24, 315–338.
30. Lhasa limited, Derek Nexus, 2ME, Derek Nexus report, pdf. 2022.
31. European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC). (1995) The toxicology of glycol ethers and its relevance to man., *ECETOC Technical Report*, 6431.
32. W. Maddrey, *Gastroenterology*, Hepatotoxicity, 2000, 118, 984-985.
33. A. Borel and F. Abbott, *Chemical Research in Toxicology*, 1995, 8, 891-899.
34. A. Kalgutkar, I. Gardner and R. Obach, *Current Drug Metabolism*, 2005, 6, 161-225.
35. C. Redlich, A. West and L. Fleming, *Gastroenterology*, 1990, 99, 748-757.
36. T. von Kreybig, R. Preussman and W. Schmidt, (1968) Chemical composition and teratogenic activity in the rat. Part I. Carboxylic acid amides, carboxylic acid hydrazides, and hydroxamic acid., *Arzneimittel-Forschung*, 18 , 645-657
37. Merkle, J. and Zeller, H., 1980. Studies on acetamides and formamides for embryotoxic and teratogenic activities in the rabbit (author's transl). *Arzneimittel-forschung*, *30*(9), pp.1557-156
38. L. Sheets, *Fundamental and Applied Toxicology*, 1997, 35, 101-119.
39. M. Colovic, D. Krstic, T. Lazarevic-Pasti, A. Bondzic and V. Vasic, *Current Neuropharmacology*, 2013, 11, 315-335.
40. D. Anderson, L. Harris and C. Lieske, *Life Sciences*, 1991, 48, 2463-2468.
41. S. CALLAWAY and D. DAVIES, *British Journal of Pharmacology and Chemotherapy*, 1957, 12, 382-387.
42. C. Mebus, F. Welsch and P. Working, *Toxicology and Applied Pharmacology*, 1989, 99, 110-121.
43. A. Zhenova, *Polymer International*, 2020, 69, 895-901.
44. D. Di Girolamo and J. Pascual, *ACS Energy Letters*, 2021, 6, 959-968.
45. B. Miller-Chou and J. Koenig, *Progress in Polymer Science*, 2003, 28, 1223-1270.
46. J. Wang, J. Zhang and Y. Zhou, *Nature Communications*, 2020, 11.
47. J. Manser, J. Christians and P. Kamat, *Chemical Reviews*, 2016, 116, 12956-13008.
48. 1B. Xue, K. Dai, X. Zhang, S. Wang, C. Li, C. Zhao, X. Yang, Z. Xi, Z. Qiu, Z. Shen and J. Wang, *Chemosphere*, 2020, 258, 127067.
49. C. Goetz, *Encyclopedia of Movement Disorders*, 2010, 269-271.
50. E. Silver, S. Kuttab, T. Hasan, M. Hassan, Structural considerations in the metabolism of nitriles to cyanide in vivo. *Drug Metab Dispos*. 1982;10(5):495-498.
51. H. Tanii and K. Hashimoto, *Archives of Toxicology*, 1984, 55, 47-54.
52. A. Steffek, A. Verrusio and C. Watkins, *Teratology*, 1972, 5, 33-38.
53. C. Willhite, V. Ferm and R. Smith, *Teratology*, 1981, 23, 317-323.
54. S. Kwofie, K. Enninful, J. Yussif, L. Asante, M. Adjei, K. Kan-Dapaah, E. Tiburu, W. Mensah, W. Miller, L. Mosi and M. Wilson, *Molecules*, 2019, 24, 2299.
55. T. Erdmenger, C. Guerrero-Sanchez, J. Vitz, R. Hoogenboom and U. Schubert, *Chemical Society Reviews*, 2010, 39, 3317.
56. N. Singha, K. Hong and J. Mays, *Polymerized Ionic Liquids*, 2017, 1-22.
57. K. Hong, H. Zhang, J. Mays, A. Visser, C. Brazel, J. Holbrey, W. Reichert and R. Rogers, *Chemical Communications*, 2002, 1368-1369.
58. A. J. Carmicheal, D. M. Haddleton, S. A. F. Bon and K. R. Seddon, Chem. Commun., 2000, 1237.
59. T. Biedro? and P. Kubisa, *Journal of Polymer Science Part A: Polymer Chemistry*, 2004, 42, 3230-3235.
60. J. Pringle, J. Efthimiadis, P. Howlett, J. Efthimiadis, D. MacFarlane, A. Chaplin, S. Hall, D. Officer, G. Wallace and M. Forsyth, *Polymer*, 2004, 45, 1447-1453.

61. H. Uyama, T. Takamoto, & S. Kobayashi, Enzymatic Synthesis of Polyesters in Ionic Liquids. Polym J 34, 94–96 (2002).
62. T. Biedroń and P. Kubisa, *Journal of Polymer Science Part A: Polymer Chemistry*, 2007, 45, 4168-4172.
63. 63.T. Biedroń, M. Bednarek and P. Kubisa, *Macromolecular Rapid Communications*, 2004, 25, 878-881.
64. P. Mastrorilli, C. Nobile, V. Gallo, G. Suranna and G. Farinola, *Journal of Molecular Catalysis A: Chemical*, 2002, 184, 73-78.
65. K. Li, G. Wei, J. Darkwa and S. Pollack, *Macromolecules*, 2002, 35, 4573-4576.
66. National Center for Biotechnology Information (2022). PubChem Compound Summary for CID 74236, Tetrabutylammonium bromide. Retrieved August 19, 2022 from https://pubchem.ncbi.nlm.nih.gov/compound/Tetrabutylammonium-bromide.
67. R. Vijayaraghavan and D. MacFarlane, *Chemical Communications*, 2005, 1149.
68. S. Csihony, C. Fischmeister, C. Bruneau, I. Horváth and P. Dixneuf, *New J. Chem.*, 2002, 26, 1667-1670.
69. J. Kredatusová, H. Beneš, S. Livi, O. Pop-Georgievski, P. Ecorchard, S. Abbrent, E. Pavlova and D. Bogdał, *Polymer*, 2016, 100, 86-94.
70. Y. Kanetaka, S. Yamazaki and K. Kimura, *Macromolecules*, 2016, 49, 1252-1258.
71. J. Lu, F. Yan and J. Texter, *Progress in Polymer Science*, 2009, 34, 431-448.
72. Ministry of Health, Labour and Welfare, Japan (MHLW). (2001)_Tris (2-butoxyethyl) phosphate (78-51-3)., *Toxicity Testing Reports of Environmental Chemicals*,39
73. P. Bhatnagar and N. Jain, *Bulletin of Environmental Contamination and Toxicology*, 1986, 37, 767-773.
74. D. BASKETTER, *American Journal of Contact Dermatitis*, 1998, 9, 119-124.
75. Bowman WC and Rand MJ. (1980)_Textbook of Pharmacology., *Textbook of Pharmacology*.
76. K. SCHALLREUTER and K. SCHULZ, *Clinical and Experimental Dermatology*, 1986, 11, 460-466.
77. G. Lin and M. Hemming, *Food and Chemical Toxicology*, 1996, 34, 177-182.
78. G. Patlewicz, R. Rodford, G. Ellis and M. Barratt, *Toxicology in Vitro*, 2000, 14, 79-84.
79. D. Correia, L. Fernandes and P. Martins, *Advanced Functional Materials*, 2020, 30, 1909736.
80. J. Claverie and F. Schaper, *MRS Bulletin*, 2013, 38, 213-218.
81. G. Cecchin, G. Morini and F. Piemontesi, *Kirk-Othmer Encyclopedia of Chemical Technology*, 2003.
82. P. Bedello, M. Goitre, G. Roncarolo, S. Bundino and D. Cane, *Contact Dermatitis*, 1987, 17, 111-112.
83. J. Thyssen and T. Menné, *Chemical Research in Toxicology*, 2009, 23, 309-318.
84. A. Faurschou, T. Menné and J. Johansen, *Contact Dermatitis*, 2011, 64, 185-195.
85. L. Migliore, *Mutagenesis*, 2002, 17, 411-417.
86. Lhasa Limited, Derek Nexus, Pt, report, Leeds, 2020.
87. E. Participation, Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006 (Text with EEA relevance).
88. M. Liang and X. Yan, *Accounts of Chemical Research*, 2019, 52, 2190-2200.
89. Y. Ye, L. Xiao and B. Bin He, *Journal of Materials Chemistry B*, 2017, 5, 1518-1524.
90. P. Šálek, A. Golunova and J. Dvořáková, *Materials Letters*, 2020, 269, 127610.
91. K. Sun, Q. Liu and S. Li, *Journal of Hazardous Materials*, 2020, 393, 122393.
92. M. Sreeramareddygari, M. Somasundrum and W. Surareungchai, *New Journal of Chemistry*, 2020, 44, 5809-5818.
93. N. Nie, Y. Liu and B. Li, *Journal of Materials Chemistry B*, 2021, 9, 8094-8108.

94. S. Li, D. Liu and B. Wu, *Talanta*, 2022, 239, 123088.
95. H. He, Z. Fei and T. Guo, *Biomaterials*, 2022, 280, 121272.
96. L. Gao, K. Fan and X. Yan, *Theranostics*, 2017, 7, 3207-3227.
97. L. Liu, C. Wang and Y. Li, *Biomaterials Science*, 2021, 9, 5965-5976.
98. M. Samim, N. Kaushik and A. Maitra, *Bulletin of Materials Science*, 2007, 30, 535-540.
99. S. Penczek, J. Pretula and S. Slomkowski, *Chemistry Teacher International*, 2021, 3, 33–57.
100. N. Kamber, W. Jeong, R. Waymouth, R. Pratt, B. Lohmeijer and J. Hedrick, *Chemical Reviews*, 2007, 107, 5813-5840.
101. W. Gruszka and J. Garden, *Nature Communications*, 2021, 12.
102. P. Schäfer and S. Herres-Pawlis, *ChemPlusChem*, 2020, 85, 1044-1052.
103. W. Gruszka, A. Lykkeberg, G. Nichol, M. Shaver, A. Buchard and J. Garden, *Chemical Science*, 2020, 11, 11785-11790.
104. M. J. Stanford and A. P. Dove, Chem. Soc. Rev., 2010, 39, 486– 494.
105. C. Williams, *Chemical Society Reviews*, 2007, 36, 1573.
106. A. Gaston, Z. Greindl, C. Morrison and J. Garden, *Inorganic Chemistry*, 2021, 60, 2294-2303.
107. Lhasa Limited, Derek Nexus report C1C, Leeds, 2020.
108. C. L. Goh, Contact Dermatitis, 1989, 21, 166–171.
109. LLNA prediction taken from Derek Nexus report.
110. A. Aptula and D. Roberts, *Chemical Research in Toxicology*, 2006, 19, 1097-1105.
111. J. Ashby, *Toxicology*, 1995, 103, 177-194.
112. E. Participation, Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006 (Text with EEA relevance).
113. M. A Green, (1975) *Medicine, Science, and the Law*, 15 , 65-66
114. W. Gruszka, L. Walker, M. Shaver and J. Garden, *Macromolecules*, 2020, 53, 4294-4302.
115. L. Chen, W. Li, D. Yuan, Y. Zhang, Q. Shen and Y. Yao, *Inorganic Chemistry*, 2015, 54, 4699-4708.
116. A. Bustinza, M. Frías, Y. Liu and E. García-Bordejé, Catalysis Science & Technology, 2020, 10, 4061–4071.
117. B. Chamberlain, M. Cheng, D. Moore, T. Ovitt, E. Lobkovsky and G. Coates, *Journal of the American Chemical Society*, 2001, 123, 3229-3238.
118. M. Cheng, D. Moore, J. Reczek, B. Chamberlain, E. Lobkovsky and G. Coates, *Journal of the American Chemical Society*, 2001, 123, 8738-8749.
119. P. Schäfer and S. Herres-Pawlis, *ChemPlusChem*, 2020, 85, 1044-1052.
120. Y. Sarazin, J.-F. Carpentier, *Chem. Rev.* 2015, *115*, 3564–3614.
121. C. Stringer, R. Hickes and P. Botham, *Contact Dermatitis*, 1991, 24, 210-215.
122. C. Williams, L. Breyfogle, S. Choi, W. Nam, V. Young, M. Hillmyer and W. Tolman, *Journal of the American Chemical Society*, 2003, 125, 11350-11359.
123. A. Thevenon, C. Romain, M. Bennington, A. White, H. Davidson, S. Brooker and C. Williams, *Angewandte Chemie*, 2016, 128, 8822-8827.
124. U. Dilberoglu, B. Gharehpapagh, U. Yaman and M. Dolen, *Procedia Manufacturing*, 2017, 11, 545-554.
125. P. Delrot, D. Loterie, D. Psaltis and C. Moser, *Optics Express*, 2018, 26, 1766.
126. B. Haq, H. Khan, K. Alam, S. Attaullah, I. Zari and M. Mateenullah, *Applied Optics*, 2015, 54, 7020.
127. J. Xing, W. Chen, X. Dong, T. Tanaka, X. Fang, X. Duan and S. Kawata, *Journal of Photochemistry and Photobiology A: Chemistry*, 2007, 189, 398-404.
128. R. Liska, J. Polym. Sci. Part A 2011, 49, 3688–3699

129. E. Käpylä, T. Sedlačík, D. Aydogan, J. Viitanen, F. Rypáček and M. Kellomäki, *Materials Science and Engineering: C*, 2014, 43, 280-289.
130. A. Ovsianikov, M. Gruene, M. Pflaum, L. Koch, F. Maiorana, M. Wilhelmi, A. Haverich and B. Chichkov, *Biofabrication*, 2010, 2, 014104.
131. A. Ovsianikov, S. Mühleder, J. Torgersen, Z. Li, X. Qin, S. Van Vlierberghe, P. Dubruel, W. Holnthoner, H. Redl, R. Liska and J. Stampfl, *Langmuir*, 2013, 30, 3787-3794.
132. P. Sanjuan-Alberte, J. Vaithilingam, J. Moore, R. Wildman, C. Tuck, M. Alexander, R. Hague and F. Rawson, *Polymers*, 2021, 13, 1038.
133. X. Huang, Y. Zhang, M. Shi, L. Zhang, Y. Zhang and Y. Zhao, *European Polymer Journal*, 2021, 153, 110505.
134. K. Dietliker, T. Jung, J. Benkhoff, H. Kura, A. Matsumoto, H. Oka, D. Hristova, G. Gescheidt and G. Rist, Macromolecular Symposia, 2004, 217, 77–98.
135. I. D. Bartlett, J. M. Marshall and J. M. Maud, Journal of Non-Crystalline Solids, 1996, 198-200, 665–668.
136. Lhasa limited, Derek Nexus, report, UC1 species, 20222,pdf.
137. M. D. Barratt and K. R. Brown, Toxicology Letters, 1985, 24, 1–6.
138. A. C. de Groot and D. W. Roberts, *Contact Dermatitis*, 2014, 70, 193–204.
139. N. Tamura, K. Aoki and M. Lee, *Mutation Research Letters*, 1992, 283, 97-106.
140. T. Yamaguchi, *Agricultural and Biological Chemistry*, 1980, 44, 3017-3018.
141. E. Říhová, *Folia Microbiologica*, 1982, 27, 25-31.
142. A. Kasamaki, H. Takahashi and N. Tsumura, *Mutation Research Letters*, 1982, 105, 387-392.
143. B. Ganguly, *Environmental Science and Pollution Research*, 2019, 26, 18208-18229.
144. M. Ishidate, *Data book of chromosomal aberration test in vitro*, Elsevier, Amsterdam, 1988.
145. L. Mi and F. Chung, *Nutrition and Cancer*, 2008, 60, 12-20.
146. J. Mitchell and W. Jordan, *British Journal of Dermatology*, 1974, 91, 183-189.
147. E. BUEHLER, *Archives of Dermatology*, 1965, 91, 171.
148. S. Fregert, I. Dahlquist and L. Trulsson, *Contact Dermatitis*, 1983, 9, 87-88.
149. S. Macatonia and A. Edwards, *Immunology*, 1986, 59, 509-514.
150. D. Neckers, *Journal of Photochemistry and Photobiology A: Chemistry*, 1989, 47, 1-29.
151. M. DeRosa, *Coordination Chemistry Reviews*, 2002, 233-234, 351-371.
152. S. LOCK and J. FRIEND, *International Journal of Cosmetic Science*, 1983, 5, 39-49.
153. J. Mason, E. Zeiger and S. Haworth, *Environmental Mutagenesis*, 1987, 9, 19-28.
154. Derek Nexus report, 22-dimethoxy-2-pheynylacetonephenone-1PP, 20222, pdf.
155. P. D. King and M. C. Perry, *The Oncologist*, 2001, 6, 162–176.
156. R. S. Foster, A. Fowkes, A. Cayley, A. Thresher, A.-L. D. Werner, C. G. Barber, G. Kocks, R. E. Tennant, R. V. Williams, S. Kane and S. A. Stalford, *Genes and Environment*, 2020, 42.
157. M. T. D. Cronin, S. J. Enoch, C. L. Mellor, K. R. Przybylak, A. N. Richarz and J. C. Madden, *Toxicol. Res.*, 2017, 33, 173–182.

APPENDIX 1: DEREK NEXUS EXAMPLE REPORT

Lhasa Limited

Derek nexus

Derek Nexus Report

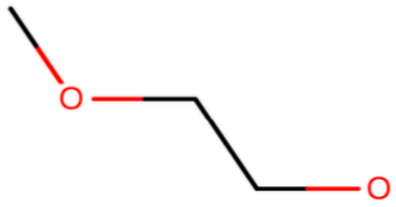
Report Information

Author	Report date	Prediction date	Program version
shahs5	12 February 2022 14:07:01	12 February 2022 14:00:20	Derek Nexus: 6.0.1, Nexus: 2.2.2

Processing Options

Selected Species bacterium, mammal	**Selected Knowledge Base(s)** Derek KB 2018 1.1	**Reasoning Level** At least EQUIVOCAL
Perceive tautomers Yes	**Perceive mixtures** Yes	**Match alerts without rules** No
Show Open likelihood Yes	**Show Negative Predictions** Yes	**Show Rapid Prototypes** Yes
Filter nearest neighbours on misclassified features Yes		

Submitted Compound - 2ME.mol



Smiles: O(CCO)C

Exact Mol Mass	76.0524
Average Mol Mass	76.09 (Source: Lhasa Limited, version 1.0)
Log Kp	-3.62 (Source: Potts & Guy, version 1.0)
Log P	-0.61 (Source: BioByte Corp., version 5.9)

Predictions

Knowledge Base: Derek KB 2018 1.1

Version	Last Modified Date	Certified by
1.1	23/11/2017 09:38:05	Lhasa Limited, Leeds, Yorkshire, UK

Reasoning Summary

◆ **Bone marrow toxicity in mammal is EQUIVOCAL**
- Alert matched: RapidPrototype156 1,2-Ethylene glycol

◆ **Developmental toxicity in mammal is PLAUSIBLE**
- Alert matched: 034 Monothioglycol or glycol monoalkyl ether, alkoxy- or alkylthio-carboxylic acid or precursors

◆ **Mutagenicity in vitro in bacterium is INACTIVE**
- No misclassified or unclassified features

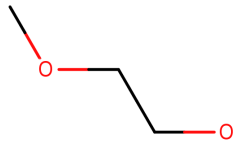
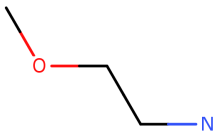
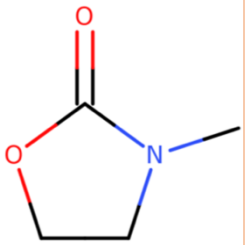
◆ **Nephrotoxicity in mammal is EQUIVOCAL**
- Alert matched: RapidPrototype069 1,2-Ethyleneglycol or derivative

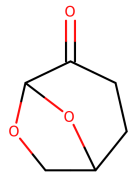
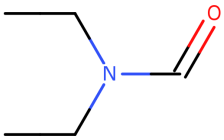
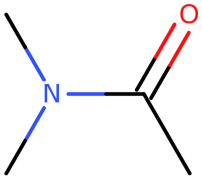
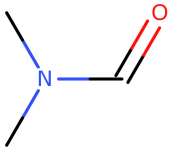
◆ **Skin sensitisation in mammal is NON-SENSITISER**
- No misclassified or unclassified features

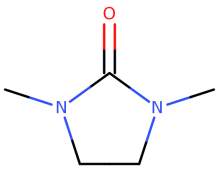
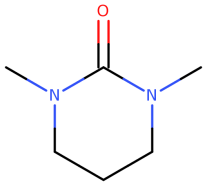
◆ **Testicular toxicity in mammal is PLAUSIBLE**
- Alert matched: 067 Monothioglycol or glycol monoalkyl ether, alkoxy- or alkylthio-carboxylic acid or precursors

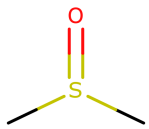
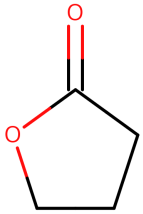
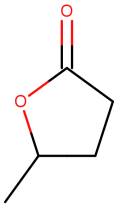
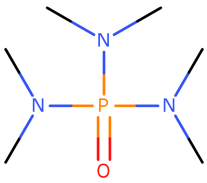
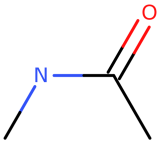
APPENDIX 2: SOLVENTS FOR ENERGY MATERIALS RESULTS TABLE

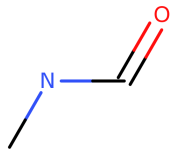
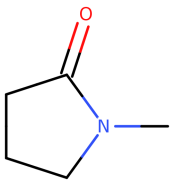
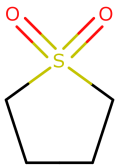
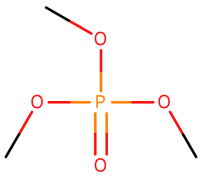
TABLE 12: DEREK NEXUS AND SARAH NEXUS SUMMARY OF RESULTS FOR HUGHES SOLVENT STRUCTURES

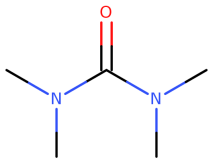
Name	Structure	Derek Nexus Alerts	Derek Nexus Reasoning	Sarah Nexus
2ME Smiles: O(CCO)C		[1] 034 Monothioglycol or glycol monoalkyl ether-alkoxy- or alkylthio-carboxylic acid or precursors [2] 067 Monothioglycol or glycol monoalkyl ether, alkoxy- or alkylthio-carboxylic acid or precursors [3] RapidPrototype069 1,2-Ethyleneglycol or derivative [4] RapidPrototype156 1,2-Ethylene glycol	▪ Bone marrow toxicity in mammal is EQUIVOCAL ▪ Developmental toxicity in mammal is PLAUSIBLE ▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Nephrotoxicity in mammal is EQUIVOCAL ▪ Skin sensitisation in mammal is NON-SENSITISER ▪ Testicular toxicity in mammal is PLAUSIBLE	Negative Confidence: 100%
2MEA Smiles: O(CCN)C		No alerts fired	No misclassified or unclassified features	Negative Confidence: 45%
3MOx Smiles: C1CN(C(O1)=O)C		No alerts fired	No misclassified or unclassified features	Positive Confidence: 11%

Cyrene **Smiles:** **C1C2CCC(C(O1)O2)=O**		No alerts fired	No misclassified or unclassified features	Negative Confidence: 26%
DEF **Smiles:** **N(C(=O)[H])(CC)CC**		**[1] 696** Short chain alkyl amide	▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER ▪ Teratogenicity in mammal is PLAUSIBLE	Negative Confidence: 55%
DMAc **Smiles:** **N(C(=O)C)(C)C**		**[1]** 696 Short chain alkyl amide	▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER ▪ Teratogenicity in mammal is PROBABLE	Negative Confidence: 100%
DMF **Dimethyl formamide** **Smiles:** **N(C(=O)[H])(C)C**		**[1] 553** Formamide derivative **[2] 696** Short chain alkyl amide	▪ Hepatotoxicity in mammal is PROBABLE - Alert matched: 553 Formamide derivative ▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER No misclassified or unclassified features	Negative Confidence: 57%

			▪ Teratogenicity in mammal is PROBABLE Alert matched: 696 Short chain alkyl amide Exact example match: dimethylformamide	
DMI Smiles: C1CN(C(N1C)=O)C		[1] RapidPrototype011 N,N'-Dialkyl ethylene urea [2] RapidPrototype041 N,N'-Dialkyl ethylene urea [3] Alert: RapidPrototype163 Ethylene- or propylene-urea	▪ Hepatotoxicity in mammal is EQUIVOCAL ▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Nephrotoxicity in mammal is EQUIVOCAL ▪ Skin sensitisation in mammal is NON-SENSITISER ▪ Testicular toxicity in mammal is EQUIVOCAL	Positive Confidence 16%
DMPU Smiles: O=C1N(C)CCCN1C		[1] RapidPrototype163 Ethylene- or propylene-urea	▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER ▪ Testicular toxicity in mammal is EQUIVOCAL	Negative Confidence: 100%

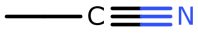
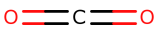
DMSO Dimethylsulfoxide Smiles: CS(C)=O		No alerts fired	▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER	Equivocal
GBL Smiles: C1CCC(O1)=O		No alerts fired	▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER	Negative Confidence: 100%
GVL Smiles: C1(CCC(O1) =O)C		No alerts fired	▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER	Negative Confidence: 44%
HMPA Smiles: P(=O)(N(C)C) (N(C)C)N(C)C		[1] 307 N-Methylol compound or precursor	▪ Mutagenicity *in vitro* in bacterium is EQUIVOCAL ▪ Skin sensitisation in mammal is NON-SENSITISER	Equivocal
NMAc Smiles: N(C(=O)C)(C)[H]		[1] 696 Short chain alkyl amide - N-methylacetamide (Exact match with Query Compound)	▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER ▪ Teratogenicity in mammal is PROBABLE	Negative Confidence: 100%

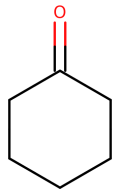
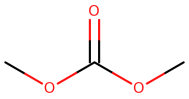
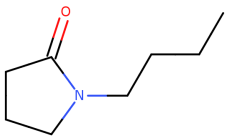
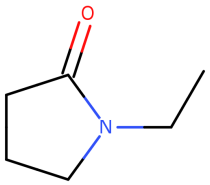
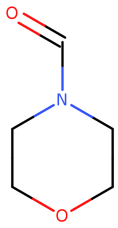
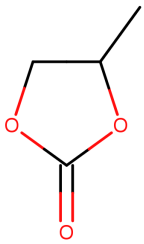
NMF Smiles: N(C(=O)[H])(C)[H]		**[1]** 553 Formamide derivative - N-methylformamide (Exact match with Query Compound) **[2]** 696 Short chain alkyl amide	▪ Hepatotoxicity in mammal is PROBABLE ▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER ▪ Teratogenicity in mammal is PROBABLE	Negative Confidence: 100%
NMP Smiles: C1CCN(C1=O)C		No alerts fired	▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER	Negative Confidence: 19%
Sulfolane Smiles: C1CCS(C1)(=O)=O		No alerts fired	No misclassified or unclassified features	Negative Confidence: 100%
TMP Smiles: P(=O)(OC)(OC)OC		**[1]** 003 Organophosphorus ester **[2]** 462 Alkyl ester of phosphoric or phosphonic acid **[3]** 653 Organophosphorus di- or tri-ester **[4]** RapidPrototype066 Phosphate or phosphonate	▪ Cholinesterase inhibition in mammal is PLAUSIBLE ▪ Hepatotoxicity in mammal is PLAUSIBLE ▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Nephrotoxicity in mammal is EQUIVOCAL	**Positive** Confidence: 100%

			▪ Skin sensitisation in mammal is PLAUSIBLE	
TMU Smiles: O=C(N(C)C)N(C)C		[1] 024 Polyalkyl urea	▪ Developmental toxicity in mammal is PLAUSIBLE ▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER	Negative Confidence: 38%

APPENDIX 3: SOLVENTS FOR POLYMER DISSOLUTION

TABLE 13: SUMMARY OF SARAH NEXUS AND DEREK NEXUS REASONING FOR SOLVENT POLYMER DISSOLUTION

Name	Structure	Derek Alerts	Derek Reasoning	Sarah Nexus
Acetonitrile Smiles: C([H])([H])([H])C#N		[1] 038 Nitrile [2] 604 Nitrile or derivative	Cyanide-type effects in mammal is PLAUSIBLE Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is NON-SENSITISER Teratogenicity in mammal is PLAUSIBLE	Negative Confidence: 18%
CO2 Smiles: O=C=O		No alerts fired	Contains unclassified features	Outside domain

Cyclohexanone Smiles: O=C1CCCCC1		No alerts fired	No misclassified or unclassified features	Negative Confidence: 100%
Dimethyl carbonate Smiles: O=C(OC)OC		No alerts fired	No misclassified or unclassified features	Equivocal
NButylP Smiles: C1CCN(C1=O)CCCC		No alerts fired	No misclassified or unclassified features	Negative Confidence: 55%
NethylP Smiles: C1CCN(C1=O)CC		No alerts fired	No misclassified or unclassified features	Negative Confidence: 33%
Nformylmorphine Smiles: C1COCCN1C(=O)[H]		No alerts fired	No misclassified or unclassified features	Negative Confidence: 25%
Propylene Carbonate Smiles: C1(COC(O1)=O)C		No alerts fired	No misclassified or unclassified features	Equivocal

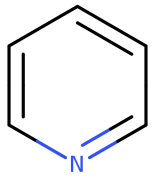
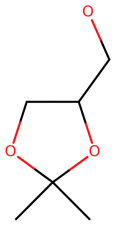
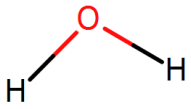
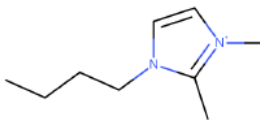
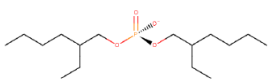
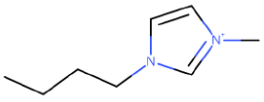
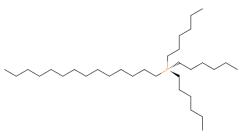
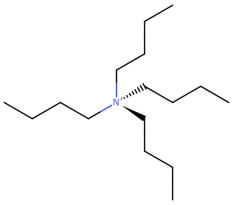
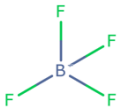
Pyridine **Smiles: C1=CN=CC=C1**		No alerts fired	No misclassified or unclassified features	Equivocal
Solketal **Smiles: C1(COC(O1)(C)C)CO**		No alerts fired	No misclassified or unclassified features	Negative Confidence: 30%
H_2O		No alerts fired	No misclassified or unclassified features	Negative Confidence: 100%

TABLE 14: SUMMARY IONIC LIQUIDS SCREENED WITH DEREK NEXUS AND SARAH NEXUS

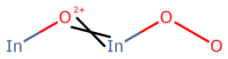
Name	Structure	Derek Alerts	Derek Reasoning	Sarah Nexus
BDMIM Smiles: C=1N(C(=[N+](C=1)C)C)CCCC		No alerts fired	No misclassified or unclassified features	Negative Confidence: 48%
Bis-2-ethylhexyl-phosphate Smiles: O=P(OCC(CCCC)CC)(OCC(CCCC)CC)[O-]		[1] 653 Organophosphorus-di- or tri- ester	▪ Carcinogenity in mammal is open ▪ Hepatotoxicity in mammal is PLAUSIBLE ▪ Mutagenicity *in vitro* in bacterium is INACTIVE ▪ Skin sensitisation in mammal is NON-SENSITISER	Negative Confidence: 60%
BMIM 1-butyl-3-methylimidazolium Smiles: C=1N(C=[N+](C=1)C)CCCC		No alerts fired	No misclassified or unclassified features	Negative Confidence: 100%
Bromide Smiles: [Br-]		No alerts fired	No misclassified or unclassified features	Equivocal
Phosphinate C([P+](CCCCCCCCCCCCCC)(CCCCCC)CCCCCC)CCCC		[1] RapidPrototype007 Phosphine or phosphine oxide	▪ Hepatotoxicity in mammal is EQUIVOCAL ▪ Skin sensitisation in mammal is NON-SENSITISER	Negative Confidence: 36%

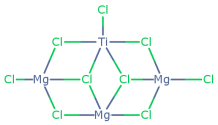
Tetrabutyl ammonium **Smiles: C([N+](CCCC)(CCCC)CCCC)CCC**		**[1]** 218 Quaternary ammonium salt **[2]** 436 Quaternary ammonium salt	▪ Irritation (of the eye) in mammal is PLAUSIBLE ▪ Irritation (of the skin) in mammal is PLAUSIBLE ▪ Mutagenicity *in vitro* in bacterium is INACTIVE	Negative Confidence: 78%
Tetrafluoroborate **Smiles: F[B-](F)(F)F**		**[1]** 216 Boron halide	▪ Carcinogenicity in mammal is PLAUSIBLE ▪ Irritation (of the eye) in mammal is PLAUSIBLE ▪ Irritation (of the respiratory tract) in mammal is PLAUSIBLE ▪ Irritation (of the skin) in mammal is PLAUSIBLE ▪ Thyroid toxicity in mammal is PLAUSIBLE	

APPENDIX 5: HETEROGENOUS CATALYSTS

TABLE 15: SUMMARY OF HETEROGENEOUS CATALYSTS SCREENED USING DEREK NEXUS AND SARAH NEXUS

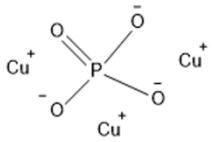
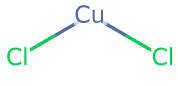
Name	Structure	Derek Alerts	Derek Reasoning	Sarah Nexus
Au (Gold) **Smiles: [Au]**	Au	**[1]** 443 Metal or metal salt	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is PLAUSIBLE	Outside domain

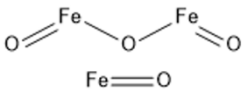
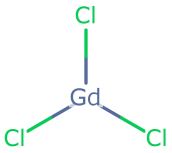
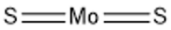
Indium Oxide In2O3 **Smiles:** **[InH2][OH+2]=[In]OO**		No alerts fired	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is non-sensitiser.	Outside domain
Indium Oxide In2O3 **Smiles: O=[In-]=[O+2]=[In-]=O**		No alerts fired	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is non-sensitiser.	Outside domain
IrO2 **Smiles: [Ir](=O)=O**		No alerts fired	Contains unclassified features	Outside domain
Platinum	Pt	**[1]** 443 Metal or metal salt	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is PLAUSIBLE	Outside domain
Pt-Ir **Smiles: [Pt][Ir]**		**[1]** 443 Metal or metal salt **[2]** 640 Platinum compound	Chromosome damage *in vitro* in mammal is PLAUSIBLE Skin sensitisation in mammal is PLAUSIBLE Mutagenicity *in vitro* in bacterium is INACTIVE	Outside domain
Tin oxide SnO2 **Smiles: [Sn](=O)=O**		**[1]** 463 Tin or tin compound	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is Non-sensitiser	Outside domain
Titanium Nitiride TiN **Smiles: [Ti]=N**		No alerts fired	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is non-sensitiser	Outside domain

Ziegler-Natta **Smiles:** **Cl[Mg]12Cl[Mg]35Cl4[Ti](Cl1)(Cl23)(Cl[Mg]4(Cl5)Cl)Cl**		No alerts fired	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is non-sensitiser	Outside domain

APPENDIX 6: NANOZYMES

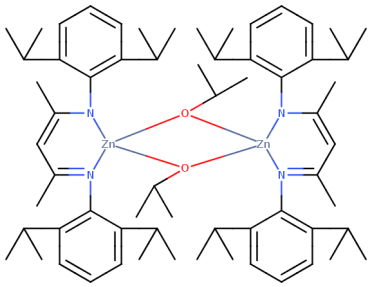
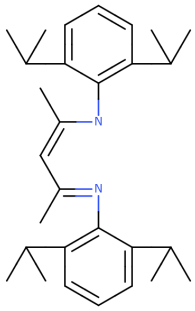
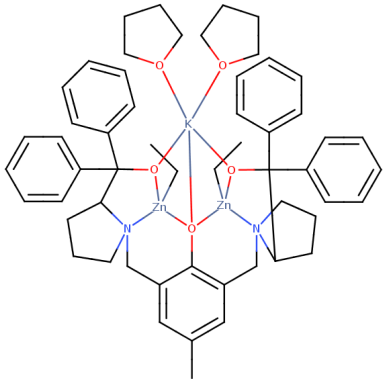
TABLE 16: NANOZYMES SCREENED USING DEREK NEXUS AND SARAH NEXUS

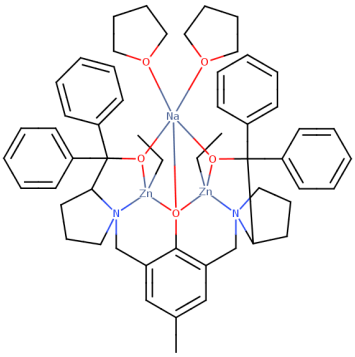
Name	Structure	Derek Alerts	Derek Reasoning	Sarah
Copper **Smiles:[Cu]**		No alerts fired	Contains misclassified features	Negative Confidence: 100 %
Cu_3PO_4 **Smiles:** **[Cu][Cu][Cu](P1(=O)OOO1)P2(=O)OOO2**		No alerts fired	Contains unclassified features	Outside domain
$CuCl_2$ **Smiles:** **[Cu](Cl)Cl**		No alerts fired	No misclassified or unclassified features	Negative Confidence: 100%
CuO **Smiles** **[Cu]=O**		No alerts fired	Contains unclassified features	Outside domain

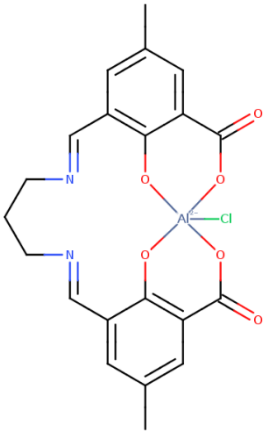
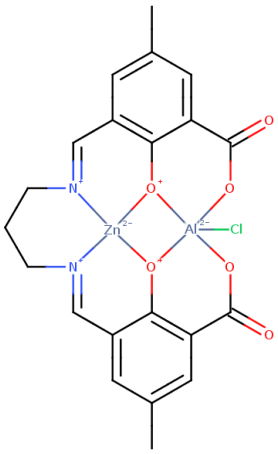
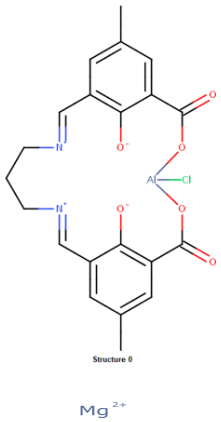
Fe_3O_4 Smiles: [Fe][Fe][Fe](=O)(=O)(=O)=O		No alerts fired	Contains unclassified features	Negative Confidence: 100%
$GdCl_3$ Smiles: [Gd](Cl)(Cl)Cl		No alerts fired	Contains unclassified features	Outside domain
MnO_2 Smiles: [Mn](=O)=O		No alerts fired	Contains unclassified features	Outside domain
MoS_2 Smiles: S=[Mo]=S		No alerts fired	Contains unclassified features	Outside domain

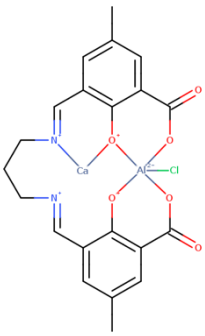
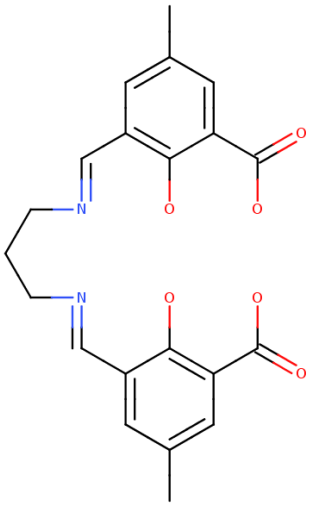
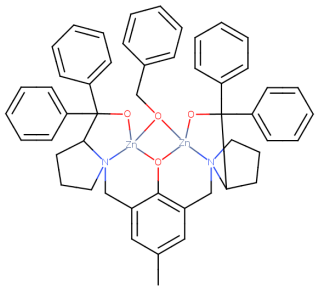
APPENDIX 7: ORGANOMETALLIC CATALYSTS

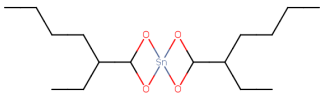
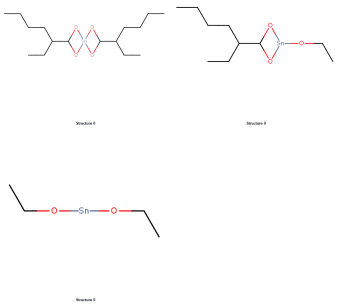
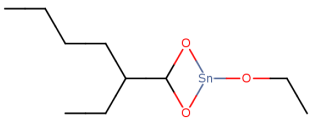
TABLE 17: SUMMARY OF RESULTS FOR ORGANOMETALLIC CATALYSTS SCREENED USING DEREK NEXUS AND SARAH NEXUS

Name	Structure	Derek Alerts	Derek Reasoning	Sarah Nexus
Coates-JACS-2001- complex		No alerts fired	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is NON-SENSITISER	Outside domain
Coates-JACS-2001-ligand Smiles: C(=NC1=C(C=CC=C1C(C)C)C(C)C)(C=C(NC2=C(C=CC=C2C(C)C)C(C)C)C)C		No alerts fired	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is NON-SENSITISER	Negative Confidence: 30%
File-reference: Garden-ChemSci-2020-JGH-K		No alerts fired	HERG channel inhibition *in vitro* in bacterium is OPEN HERG channel inhibition *in vitro* in mammal is OPEN Mutagenicity *in vitro* in bacterium is INACTIVE contains unclassified features	Outside domain

			Photoallergenicity in bacterium is OPEN Photoallergenicity in mammal is OPEN Skin sensitisation in mammal is NON-SENSITISER Contains unclassified features	
File reference: Garden-ChemSci-2020-JGH-Na.mol		No alerts fired	HERG channel inhibition *in vitro* in bacterium is OPEN HERG channel inhibition *in vitro* in mammal is OPEN Mutagenicity *in vitro* in bacterium is INACTIVE Contains unclassified features Photoallergenicity in bacterium is OPEN Photoallergenicity in mammal is OPEN Skin sensitisation in mammal is NON-SENSITISER Contains unclassified features	Outside domain

Garden-IC-2021-JGH-C1		[1] 444 Imine or alpha,beta-unsaturated imine	Skin sensitisation in mammal is EQUIVOCAL Alert matched: 444 Imine or alpha,beta-unsaturated imine Mutagenicity *in vitro* in bacterium in INACTIVE – contains unclassified features	Outside domain
Garden-IC-2021-JGH-C1B.mol		No alerts fired	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is NON-SENSITISER	Outside domain
Garden-IC-2021-JGH-C1C-coord-bond.mol		[1] 444 Imine or alpha,beta-unsaturated imine	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is EQUIVOCAL Skin sensitisation in mammal is NON-SENSITISER	Outside domain

Garden-IC-2021-JGH-C1D-		[1] 444 Imine or iminium salt	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is equivocal	Outside domain
File: Garden-IC-2021-JGH-ligand.mol Smiles: C1(=CC(=CC(=C1O)C(O)=O)C)C=NCCCN=CC2=C(O)C(=CC(=C2)C)C(O)=O		**[1]** 439 Substituted phenol **[2]** Alert: 444 Imine or alpha, beta-unsaturated imine **[3]** Alert: 666 para-Alkylphenol or derivative	Hepatotoxicity in mammal is PLAUSIBLE Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is EQUIVOCAL	Negative Confidence: 47%
Garden-Macromol-2020-JGH-all.m		No alerts fired	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is NON-SENSITISER	Outside domain

File: Herres-Pawlis-CPC-2020-JGHdioctanoate. Smiles: C(C1O[Sn]2(O1)OC(O2)C(CCCC)CC)(CC)CCCC		**[1]** 463 Tin or tin compound	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is PLAUSIBLE	Outside domain
Herres-Pawlis-CPC-2020-JGH-bisalkoxide.mol		**[1]** 463 Tin or tin compound	Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is PLAUSIBLE	Outside domain
Herres-Pawlis-CPC-2020-JGHmonoalkoxide Smiles: C(C1O[SnH](O1)OCC)(CC)CCCC		**[1]** 463 Tin or tin compound	Skin sensitisation in mammal is PLAUSIBLE	Outside domain

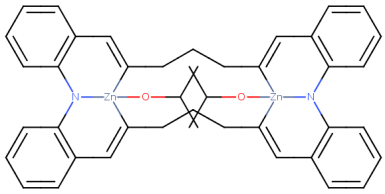
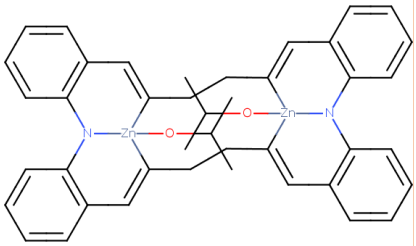
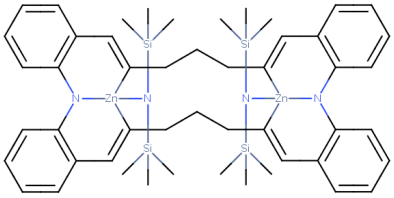
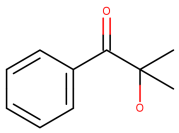
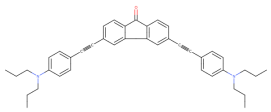
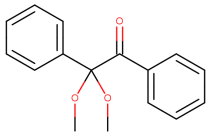
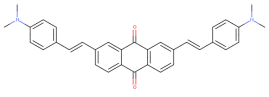
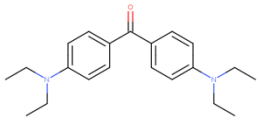
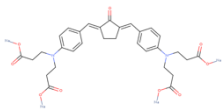
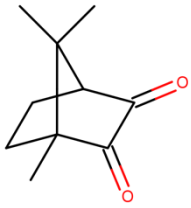
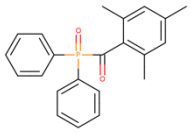
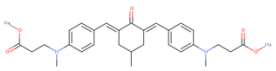
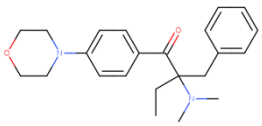
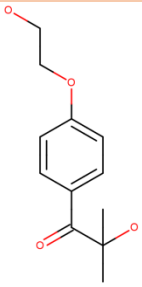
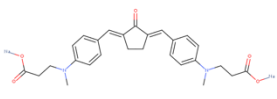
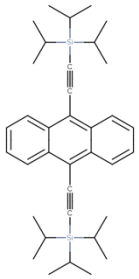
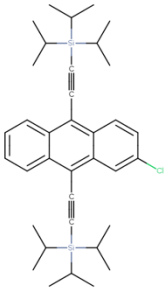
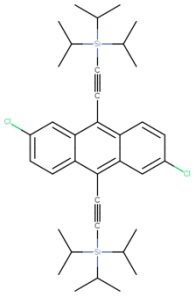
Williams-AIE-2016-JGH-A.mol		No alerts fired	Skin sensitisation in mammal is NON-SENSITISER Mutagenicity *in vitro* in bacterium is INACTIVE	Outside domain
Williams-AIE-2016-JGH-B.mol		No alerts fired	Skin sensitisation in mammal is NON-SENSITISER Mutagenicity *in vitro* in bacterium is INACTIVE	Outside domain
Williams-AIE-2016-JGH-C.mol		No alerts fired	Skin sensitisation in mammal is NON-SENSITISER Mutagenicity *in vitro* in bacterium is INACTIVE	Outside domain

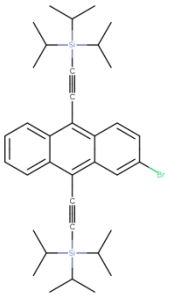
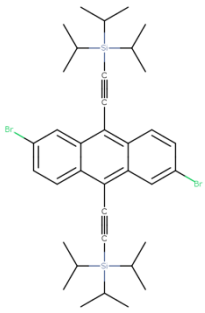
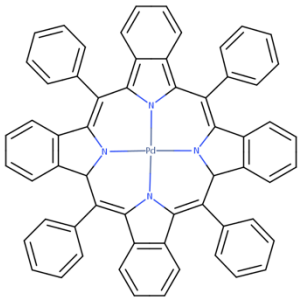
TABLE 18: SUMMARY OF RESULTS FOR PHOTOINITIATORS SCREENED USING DEREK NEXUS AND SARAH NEXUS

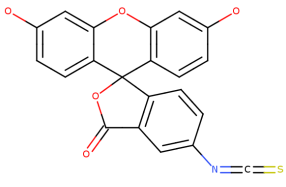
Name	Structure	Derek Alerts	Derek Reasoning	Sarah Nexus
2-hydroxy-2-methylpropiophenone - 1PP		No alerts fired	No misclassified or unclassified features	Negative Confidence: 100%
3-6-B3FL - fluorenone - 2PP		[1] 456 Diaryl ketone [2] 728 1,3,5-Hexatriene	Mutagenicity *in vitro* in bacterium is INACTIVE Photoallergenicity in mammal is PLAUSIBLE Phototoxicity in mammal is EQUIVOCAL Skin sensitisation in mammal is NON-SENSITISER	Positive Confidence: 13%
22-dimethoxy-2-phenylacetonephenone - 1PP		[1] RapidPrototype060 Methylene glycol or derivative	Mutagenicity *in vitro* in bacterium is INACTIVE Nephrotoxicity in mammal is EQUIVOCAL Skin sensitisation in mammal is NON-SENSITISER	Negative Confidence: 26%
AQN - anthraquinone derivatives		**[1]** 660 Anthraquinone **[2]** 953 4-Amino biphenyl, stilbene or derivative **[3]** 632 Anthraquinone derivative **[4]** RapidPrototype113	Carcinogenicity in mammal is PLAUSIBLE Anthraquinone Chromosome damage *in vivo* in mammal is PLAUSIBLE Hepatotoxicity in mammal is PLAUSIBLE Mitochondrial dysfunction in mammal is EQUIVOCAL Mutagenicity *in vitro* in bacterium is INACTIVE	Positive Confidence: 44%

		Anthraquinone derivative	Photoallergenicity in mammal is PLAUSIBLE	
BDAB - bisdiethylaminobenzophenone - 2PP		[1] 456 Diaryl ketone or precursor [2] HPC02_1 Class 2: Michael acceptor	Mutagenicity *in vitro* in bacterium is INACTIVE Photoallergenicity in mammal is PLAUSIBLE Skin sensitisation in mammal is NON-SENSITISER Skin sensitisation HPC in mammal is PLAUSIBLE	Positive Confidence: 1%
BSEA - WaterSoluble 2PP		**[1]** 480 alpha,beta-Unsaturated ketone or precursor **[2]** HPC02_1 Class 2: Michael acceptor	Skin sensitisation in mammal is PLAUSIBLE Skin sensitisation HPC in mammal is PLAUSIBLE	Positive Confidence: 32%
camphorquinone - teeth 3DP		[1] 464 1,2-Dicarbonyl compound or precursor [2] HPC04_4 Class 4: Schiff base electrophiles	Skin sensitisation in mammal is PLAUSIBLE Skin sensitisation HPC in mammal is OPEN	Negative Confidence: 100%
diphenyl-246-trimethoxybenzoylphosphine oxide - 1PP		**[1]** RapidPrototype007 Phosphine or phosphine oxide **[2]** 840 Benzoylphosphine oxide or analogue **[3]** Exact example	Hepatotoxicity in mammal is EQUIVOCAL Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is PLAUSIBLE	Negative Confidence: 100%

		match: diphenyl(2,4,6-trimethylbenzoyl)phosphine oxide		
E2CK - WaterSoluble 2PP		**[1]** 480 alpha,beta-Unsaturated ketone or precursor **[2]** HPC02_1 Class 2: Michael acceptor	Skin sensitisation in mammal is PLAUSIBLE Skin sensitisation HPC in mammal is PLAUSIBLE	Positive Confidence: 24%
irgacure 369 - 2PP		**[1]** 918 Alkyl amine **[2]** HPC02_1 Class 2: Michael acceptor	Skin irritation/corrosion in mammal is PLAUSIBLE Skin sensitisation in mammal is NON-SENSITISER Skin sensitisation HPC in mammal is PLAUSIBLE	Negative 23%
irgacure 2959 - 2PP		No alerts fired	Carcinogenicity in mammal is OPEN Mutagenicity *in vitro* in bacterium is INACTIVE Skin sensitisation in mammal is NON-SENSITISER	Negative 23%
P2CK - WaterSoluble 2P		[1] 480 alpha,beta-Unsaturated ketone or precursor [2] HPC02_1 Class 2: Michael acceptor	Skin sensitisation in mammal is PLAUSIBLE Skin sensitisation HPC in mammal is PLAUSIBLE	Positive 29%

UC species 1		[1] 650 Organosilicon compound [2] 769 Anthracene [3] RapidPrototype056 Organosilicon compound	Hepatotoxicity in mammal is PLAUSIBLE Mutagenicity *in vitro* in bacterium is PLAUSIBLE Nephrotoxicity in mammal is EQUIVOCAL Skin sensitisation in mammal is NON-SENSITISER	Outside domain
UC species 2		[1] 650 Organosilicon compound [2] 769 Anthracene [3] RapidPrototype056 Organosilicon compound	Hepatotoxicity in mammal is PLAUSIBLE Mutagenicity *in vitro* in bacterium is PLAUSIBLE Nephrotoxicity in mammal is EQUIVOCAL Skin sensitisation in mammal is NON-SENSITISER	Outside domain
UC species 3		[1] 650 Organosilicon compound [2] 769 Anthracene [3] RapidPrototype056 Organosilicon compound	Hepatotoxicity in mammal is PLAUSIBLE Mutagenicity *in vitro* in bacterium is PLAUSIBLE Nephrotoxicity in mammal is EQUIVOCAL Skin sensitisation in mammal is NON-SENSITISER	Outside domain

UC species 4		[1] 650 Organosilicon compound [2] 769 Anthracene [3] RapidPrototype056 Organosilicon compound	Hepatotoxicity in mammal is PLAUSIBLE Mutagenicity *in vitro* in bacterium is PLAUSIBLE Nephrotoxicity in mammal is EQUIVOCAL Skin sensitisation in mammal is NON-SENSITISER	Outside domain
UC species 5		[1] 650 Organosilicon compound [2] 769 Anthracene [3] RapidPrototype056 Organosilicon compound	Hepatotoxicity in mammal is PLAUSIBLE Mutagenicity *in vitro* in bacterium is PLAUSIBLE Nephrotoxicity in mammal is EQUIVOCAL Skin sensitisation in mammal is NON-SENSITISER	Outside domain
		[1] 728 1,3,5-Hexatriene [2] 443 Metal or metal salt	HERG channel inhibition *in vitro* in bacterium is OPEN HERG channel inhibition *in vitro* in mammal is OPEN Mutagenicity *in vitro* in bacterium is INACTIVE Photoallergenicity in bacterium is OPEN Photoallergenicity in mammal is OPEN Phototoxicity in mammal is EQUIVOCAL Skin sensitisation in mammal is PLAUSIBLE	Outside domain

FITC Smiles: C1(=CC=C2C (=C1)OC5=C (C23OC(C4=C 3 C=CC(=C4) N=C=S)=O) C=CC(=C5)O) O		[1] 304 Isocyanate or isothiocyanate - Chromosome damage (clastogenicity) [2] 410 Isothiocyanate - Skin sensitisation: guinea pig maximisation test, local lymph node assay [3] 763 Fluorescein or derivative – phototoxicity	Chromosome damage *in vitro* in mammal is PLAUSIBLE Mutagenicity *in vitro* in bacterium is PLAUSIBLE Skin sensitisation in mammal is PLAUSIBLE	Equivocal