HTS and SAR Analysis of Chemicals from the Elk River Spill

Scott S Auerbach[1], Neepa Choksi[2], , Stephen Ferguson[1], Jui-Hua Hsieh[1], Daniel Svoboda[3], Glenn Myatt[4], Matthew Wampole[5], Aleck Sedykh[6], Josep Prous, Jr.[7], Scott Masten[1], Nigel Walker[1], Mary Wolfe[1], Raymond R. Tice[1], John R. Bucher[1]
[1]NIH/NIEHS/DNTP, RTP, NC, USA; [2]ILS, RTP, NC, USA; [3]Social & Scientific Systems, Durham, NC, USA; [4]Leadscope, Inc., Columbus, OH, USA; [5]Thompson Reuters, New York, NY, USA; [6]MultiCASE Inc., Beachwood, OH, USA;
[7]Prous Institute for Biomedical Research, Barcelona, Spain

Abstract

In January 2014, a chemical mixture used for cleaning coal was accidently released into the Elk River in West Virginia, resulting in contamination of the water supply for nearly 300,000 people. The mixture included 4-methylcyclohexanemethanol (MCHM), 1,4-cyclohexanedimethanol (CHDM), 4-methoxymethylcyclohexanemethanol (MMCHM), dimethyl 1,4-cyclohexanedicarboxylate (DMCHDC), and methyl 4-methylcyclohexanecarboxylate (MMCHC) a proprietary mixture of propylene glycol ethers containing propylene glycol phenyl ether (PPH) and dipropylene glycol phenyl ether (DiPPH). Limited toxicological data were available for these chemicals; therefore, a review of the available Tox21 high throughput screening (HTS) assay data was performed along with a structure activity relationship (SAR) analysis to predict potential toxicological effects. Four chemicals (MCHM, CHDM, DMCHDC, PPH) were evaluated in 27 stress response pathway and nuclear receptor assays. All four chemicals were inactive at concentrations up to ~100 µM. For SAR evaluations, 7 platforms were used (ADMET Predictor™, CASE Ultra, Leadscope®, MetaDrug™, Prous Institute SymmetrySM, Toxtree, VEGA). The results were organized based on the category of predicted endpoints and an integrated analysis was performed. Consistent with observed effects in West Virginia residents following the spill, the SAR analyses predicted that a number of the chemicals would be irritating to the skin, eyes, and lung. The chemicals also were predicted to produce effects in the liver and kidney, which is largely consistent with the limited, available toxicity data. A focused assessment of dose-response for effects on these and other target organs/systems will be performed as part of ongoing NTP toxicological studies.

Introduction

In January 2014, approximately 10,000 gallons of a liquid used to wash coal and remove impurities that contribute to pollution during combustion were spilled from a leaking tank into the Elk River. The spill led to the contamination of the water supply of nearly 300,000 people within nine counties in the Charleston, West Virginia metropolitan area. Reports of licorice odors at homeowner taps and hospital admittances were signs that the population was exposed to contaminated tap water. A recent CDC report found that one-fifth of households that received contaminated water reported health effects that were believed to be related to the chemical spill. The report indicates that most of the health effects involved rashes and skin irritation; however, respiratory illnesses, nausea, and diarrhea were reported also.

The information available to date indicates that the major constituent of the spilled liquid was a technical product (crude MCHM) containing 4-methyl-1-cyclohexanemethanol (MCHM, CASRN 34885-03-5) as the major constituent. A proprietary mixture containing predominantly dipropylene glycol phenyl ether (DiPPH, CASRN 51730-94-0) and propylene glycol phenyl ether (PPH, CAS 770-35-4) was also reported to be present in the leaking tank at <10 % by weight. Based on Material Safety Data Sheets, several additional chemicals were likely present in the spilled liquid at lower levels and these are noted in Table 1.

The chemicals involved in the West Virginia Elk River Spill were nominated by the Centers for Disease Control and Prevention (CDC)/Agency for Toxic Substances and Disease Registry (ATSDR) for toxicological characterization. To respond to the request for additional toxicology data by the CDC/ATSDR, the NTP plans to perform a number of studies of relatively short duration (Table 1) to provide information relevant to the potential exposures of the Charleston residents. The chemicals of greatest concern (e.g., the chemicals of higher abundance in the spilled material) will be studied in rodent toxicology models, in lower organisms, and using predictive modeling approaches. Chemicals of more limited concern (e.g., minor constituents) will be evaluated using lower organisms and *in silico* approaches. A major focus of the toxicological characterization will be the use of chemical and bioinformatics-based predictive models. The models make it possible to: (1) query for effects on nearly all biological processes; (2) suggest a need for longer-term, more comprehensive toxicology studies; and (3) provide a conservative estimate of the dose levels for longer-term studies where health effects would be anticipated, should they occur. Another significant consideration in selecting the types of studies to be performed is the need to assess the potential for acute exposures to result in irreversible effects. For this reason, several of the assessments will evaluate effects on fetal and early post-natal development, effects that are often irreversible.

This presentation reviews the results of the quantitative high throughput screening (qHTS) and the structure activity response (SAR) analysis of the chemicals spilled into the Elk River.

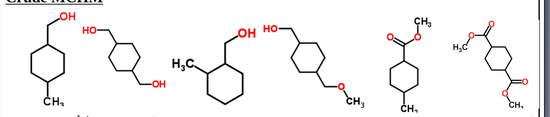
The Intramural Research Program of the National Institute of Environmental Health Sciences (NIEHS) supported this poster. Technical support was provided by ILS under NIEHS contract HHSN27320140003C. The views expressed above do not necessarily represent the official positions of any Federal agency. Since the poster was written as part of the official duties of the authors, it can be freely copied. The findings and conclusions in this presentation are those of the authors and do not represent the views of the NTP or any final NTP determination or policy.

Table 1. Proposed NTP Studies

	Studies							
Test Article [Abbreviation, CAS Number]	Rat Prenatal Toxicity	Mouse Dermal Irritation and Hypersensitivity	5-Day Rat Toxicogenomic	Bacterial Mutagenicity	Zebrafish Developmental	Nematode Toxicity	High Throughput Screening	Structure Activity Relationship (SAR) Analysis
4-Methylcyclohexanemethanol [MCHM, 34885-03-5]	X	X	X	X	X	X	X	X
Dipropylene glycol phenyl ether [DiPPH, 51730-94-0]			X	X	X	X		X
Propylene glycol phenyl ether [PPH, 770-35-4]			X	X	X	X	X	X
1,4-Cyclohexanedimethanol (CHDM; 105-08-8)				X	X	X	X	X
2-Methylcyclohexanemethanol [2MCHM, 2105-40-0]				X	X	X		X
4-(Methoxymethyl)cyclohexanemethanol [MMCHM, 98955-27-2]				X	X	X		X
Dimethyl 1,4-cyclohexanedicarboxylate [DMCHDC, 94-60-0]				X	X	X	X	X
Methyl 4-methylcyclohexanecarboxylate [MMCHC, 51181-40-9]				X	X	X		X
Technical product ["crude MCHM"]		X	X	X	X	X		

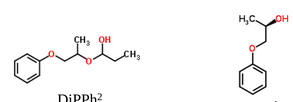
Figure 1. Spilled Chemicals

Crude MCHM



MCHM[1,*] (34885-03-5) | CHDM* (105-08-8) | 2MCHM (2105-40-0) | MMCHM (98955-27-2) | MMCHC (51181-40-9) | DMCHDC* (94-60-0)

Glycol Ethers



DiPPh[2] (25961-90-4) | PPh* (770-35-4)

[1]Major constituent of the spilled material; [2]Four Isomers; *Screened in Tox21 Phase 2

qHTS Results

- Results from 21 Tox21 assays were reviewed
  - Nuclear receptors, Cell Stress, and Cell Viability
  - 15 point dose response up to 92 µM
  - Each assay was run 3 separate times to determine reproducibility
  - A chemical was deemed to be active if the exhibit a curve classification of 1.1, 1.2, 2.1, or 2.2 in 2 out 3 assay runs
- **None of the 6 chemicals were active in any of the assays**
- **Important**: Analytic evaluation of the library is ongoing
  - Verified chemicals: 1,4-Cyclohexanedimethanol (CHDM); Dimethyl 1,4-cyclohexanedicarboxylate (DMCHDC); Phenoxyisopropanol
  - Unverified chemicals: **4-methylcyclohexanemethanol (MCHM)**; Propylene glycol phenyl ether (PPH); Cyclohexanemethanol, 4-((ethenyloxy)methyl)-

Smiles for SAR

Name	CASRN	SMILES
1,4-Cyclohexanedimethanol	105-08-8	C1(CCC(CO)CC1)CO
4-Methoxymethylcyclohexanemethanol	98955-27-2	C1CC(CCC1CO)COC
4-Methylcyclohexanecarboxylic acid	4331-54-8	C1(C(O)=O)CCC(C)CC1
Propylene glycol phenyl ether	770-35-4	O(c1ccccc1)C[C@@H](O)C
4-Methylcyclohexanemethanol	34885-03-5	OCC1CCC(CC1)C
1,4-Cyclohexanedicarboxylic acid, dimethyl ester	94-60-0	C1(C(OC)=O)CCC(C(OC)=O)CC1
Methyl 4-methylcyclohexanecarboxylate	51181-40-9	C1(CCC(CC1)C)C(=O)OC
2-Methylcyclohexanemethanol	2105-40-0	C1[C@@H]([C@@H](CCC1)C)CO
Dipropylene glycol phenyl ether (isomer 1)	NOCAS	CC(CO)OC(C)COC1=CC=CC=C1
Dipropylene glycol phenyl ether (isomer 2)	344884-62-4	CC(O)COC(C)COC1=CC=CC=C1
Dipropylene glycol phenyl ether (isomer 3)	NOCAS	CC(O)COCC(C)OC1=CC=CC=C1
Dipropylene glycol phenyl ether (isomer 4)	25961-90-4	CC(CO)OCC(C)OC1=CC=CC=C1

Figure 2. SAR Software Activity Calls

- Chemical structures were evaluated across 6 software packages (Software: Leadscope, CASE Ultra, VEGA, Toxtree, MetaDrug, ADMETPredictor)
  - Activity calls (positive, negative or out of domain) as reported by the different software packages are summarized in the figure below

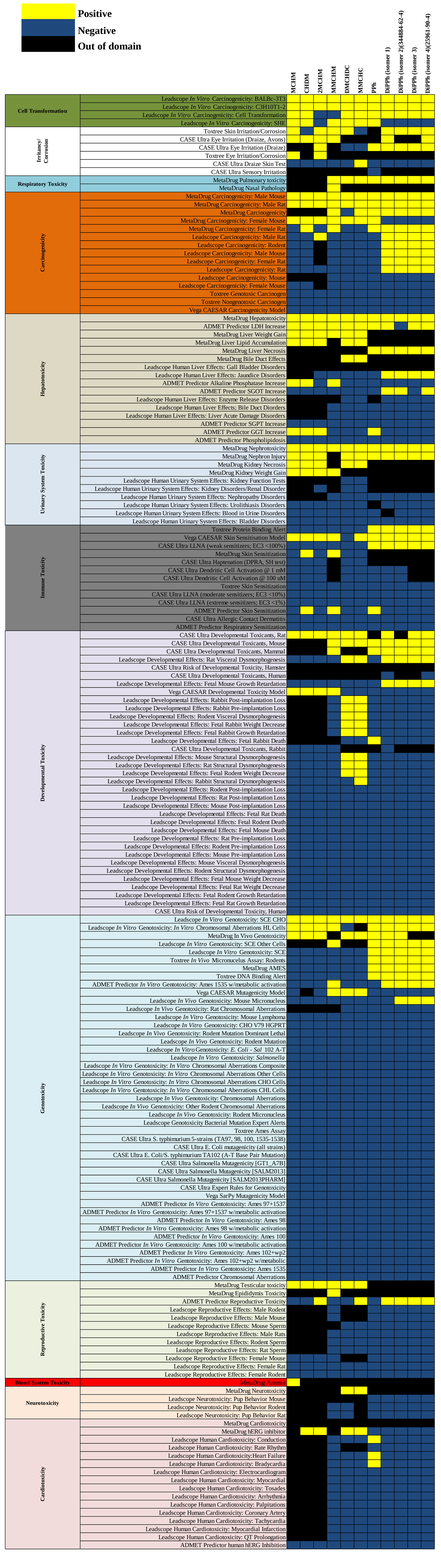


Figure 3. Summary of Overall Results

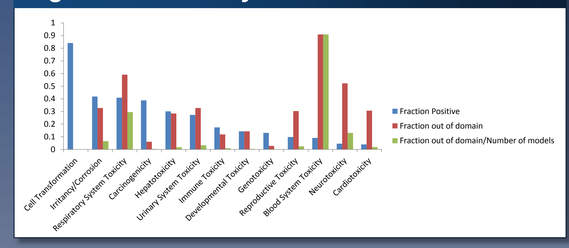
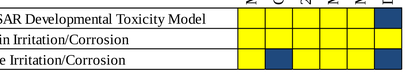


Figure 4. Review of Positive SAR Calls

- SAR model results called "**positive**" by the software were reviewed by a scientist with experience evaluating toxicity-related SAR
  - Considerations of reviewing scientist
    - Model probability score or confidence in call
    - Biological plausibility of the features driving the positive call
    - Domain of the model (structural similarity of the test chemical to model training data)

MCHM Class

Moderate to high confidence postive call based on expert review
Low confidence postive or negative by model call



Explanation of MCHM Scientific Review

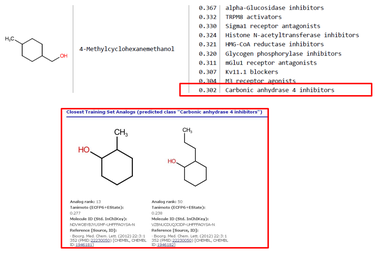
Endpoint	Model Outcome	Summary of Results	Scientific Interpretation of SAR Results
Vega CAESAR Developmental Toxicity Model	Positive, good reliability	Structural similarity of six compounds within the training set of the model ranged from 0.675 to 0.889. Model descriptors for the test compound were stated to be within the range of descriptors within the model training set. Additionally, "strongly similar" compounds with known experimental values were identified in the training set.	While the platform identified several chemicals as structurally similar, closer inspection of the structures indicates that only a single compound contains the cyclohexane alcohol feature present in the query compound. The remainder of the chemicals either contain several cyclic moieties, a carboxylic acid or ketone moiety, or are straight chain alcohols. The chemical and electronic nature of these compounds are likely different from the query compound. There is some caution that should be taken when using the activities of these compounds to predict the activity of the query compound. Therefore, there is a lower confidence in the prediction that the compound is a developmental toxicant.
Toxtree Skin Irritation/Corrosion	Positive (Irritating to skin)	The presence of an ethylene glycol ether moiety was identified as a structural feature associated with skin irritation potential.	The model evaluates chemicals based on physicochemical properties (e.g., Log P, lipid solubility, and surface tension) and the presence of structural alerts associated with skin irritation and/or corrosion. For the current evaluation, predicted physicochemical properties were included for most of the endpoints. The one endpoint that was excluded was lipid solubility. There is lower confidence in the prediction.
Toxtree Eye Irritation/Corrosion	Positive (Serious lesions to the eye)	The presence of an aliphatic alcohol moiety was identified as a structural feature associated with eye irritation potential.	The model evaluates chemicals based on physicochemical properties (e.g., Log P, lipid solubility, and surface tension) and the presence of structural alerts associated with eye irritation and/or corrosion. For the current evaluation, predicted physicochemical properties were included for most of the endpoints. The one endpoint that was excluded was lipid solubility. There is lower confidence in the prediction.

PPh Class

None of the models that reported positive calls for PPh of DiPPh were deemed to be of moderate or high confidence following expert review

Figure 5: Mode of Action Prediction with Symmetry

Prous Institute Symmetry chemical structure analysis software was used to predict the potential modes of chemical modes of action. The results of the MCHM analysis are shown below



Conclusions

- qHTS on a subset of the spilled chemicals indicated they were not active at doses up to ~100 µM
- Caution should be taken when accepting model predictions at face value
  - Limited structural complexity of chemicals under consideration combined with chemical space the models are trained on significantly limits confidence in the predictions
- Scientific review of the positive model predictions related to developmental toxicity and irritancy/corrosion to be of at least moderate confidence for a number of the chemicals contained in crude MCHM
  - Irritancy findings for chemicals in crude MCHM are consistent with the available data
- The large fraction of "out of domain" calls in certain "icitites" such as neurotoxicity, blood, and respiratory system indicates significant blind spots when predicting the toxicity of the spilled chemicals that would need to be addressed through testing
- Agglomerative analysis of model predictions provides a means to perform high level weight of signal evaluations of the overall SAR characterization of the spilled chemicals