

Research Article

The kinetic Direct Peptide Reactivity Assay (kDPRA): Intra- and Inter-Laboratory Reproducibility in a Seven-Laboratory Ring Trial

Britta Wareing¹, Susanne N. Kolle¹, Barbara Birk¹, Nathalie Alépée², Tina Haupt³, Rishil Kathawala⁴, Petra S. Kern⁵, Laurent Nardelli², Hans Raabe⁴, Marián Rucki⁶, Cindy A. Ryan⁷, Sjoerd Verkaart⁸, Walter M. A. Westerink⁸, Robert Landsiedel¹ and Andreas Natsch³

¹BASF SE Experimental Toxicology and Ecology, Ludwigshafen, Germany; ²L'Oréal Research & Innovation, Aulnay-sous-Bois, France; ³Fragrances S&T, Ingredients Research, Givaudan Schweiz AG, Kemptthal, Switzerland; ⁴Institute for In Vitro Sciences, Inc., Gaithersburg, MD, USA; ⁵Procter & Gamble Services NV/SA, Brussels, Belgium; ⁶National Institute of Public Health, Centre of Toxicology and Health Safety, Prague, Czech Republic; ⁷Procter & Gamble, Mason, OH, USA; ⁸Charles River Laboratories, Den Bosch BV, The Netherlands

Abstract

While the skin sensitization hazard of substances can be identified using non-animal methods, the classification of potency into UN GHS sub-categories 1A and 1B remains challenging. The kinetic direct peptide reactivity assay (kDPRA) is a modification of the DPRA wherein the reaction kinetics of a test substance towards a synthetic cysteine-containing peptide are evaluated. For this purpose, several concentrations of the test substance are incubated with the synthetic peptide for several incubation times. The reaction is stopped by addition of monobromobimane, which forms a fluorescent complex with the free cysteine of the model peptide. The relative remaining non-depleted amount of peptide is determined. Kinetic rate constants are derived from the depletion vs concentration and time matrix and used to distinguish between UN GHS sub-category 1A sensitizers and test substances in sub-category 1B/not classified test substances. In this study, we present a ring trial of the kDPRA with 24 blind-coded test substances in seven laboratories. The intra- and inter-laboratory reproducibility were 96% and 88%, respectively (both for differentiating GHS Cat 1A sensitizers from GHS Cat 1B/not classified). Following an independent peer review, the kDPRA was considered to be acceptable for the identification of GHS Cat 1A skin sensitizers. Besides GHS Cat 1A identification, the kDPRA can be used as part of a defined approach(es) with a quantitative data integration procedure for skin sensitization potency assessment. For this aim, next to reproducibility of classification, the quantitative reproducibility and variability of the rate constants were quantified in this study.

1 Introduction

The mechanism underlying skin sensitization is complex, but it is well understood and described as an adverse outcome pathway by OECD (OECD, 2014). During the last decade, significant progress has been made in the field of non-animal tests, and several test methods to address skin sensitization have meanwhile been adopted as OECD test guidelines (TGs). To date however, the regulatory accepted test methods were validated to address skin sensitization hazard but not sensitizer potency, although potency information is essential for risk assessment purposes.

The molecular initiating event in the adverse outcome pathway for skin sensitization is the covalent binding of the sensitizer to skin proteins, and a first method to address this key event was adopted as OECD TG 442C in 2015. This method, the direct peptide reactivity assay (DPRA), uses two synthetic peptides (one containing a cysteine, one containing a lysine residue) that are incubated with a single concentration of the test substance. After 24 hours incubation, the concentrations of remaining, non-depleted peptide are determined using HPLC.

All currently adopted non-animal OECD TGs to assess skin sensitization, including the DPRA, provide information on the

Received April 29, 2020; Accepted June 4, 2020;
Epub June 10, 2020; © The Authors, 2020.

ALTEX 37(4), 639–651. doi:10.14573/altex.2004291

Correspondence: Andreas Natsch, PhD
Fragrances S&T, Ingredients Research,
Givaudan Schweiz AG
Kemptpark 50, 8310 Kemptthal, Switzerland
(andreas.natsch@givaudan.com)

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is appropriately cited.



skin sensitizing hazard but cannot be used as a stand-alone test to address skin sensitization (OECD, 2020) or sufficiently address skin sensitizing potency (Wareing et al., 2017). A score-based approach for identification of GHS 1A chemicals has been proposed, however, reproducibility of the underlying scores derived from quantitative data in the DPRA and the hCLAT assay has not yet been documented (Nukada et al., 2013).

The kinetic direct peptide reactivity assay (kDPRA) is a modification of the OECD-adopted *in chemico* DPRA (described in Appendix I of OECD TG 442C, (OECD, 2020)). The kDPRA uses the cysteine-containing test peptide (Ac-RFAACAA-COOH; Cys), also used in the DPRA, while it does not use a lysine-containing peptide. The final concentration of the Cys peptide (0.5 mM) and the reaction medium (25% acetonitrile in phosphate buffer) are identical in the kDPRA and in the DPRA. While the DPRA measures only at one concentration of the test substance (5 mM in the Cys peptide reaction mixture) and at one, not exactly defined, time point (≥ 24 h), the kDPRA performs parallel reactions at five test substance concentrations (5, 2.5, 1.25, 0.625 and 0.3125 mM) and at six defined time points (10, 30, 90, 150, 210 and 1440 min) at $25 \pm 2.5^\circ\text{C}$. The residual concentration in % relative to the VC of the Cys peptide after the respective reaction time is measured by stopping the reaction by the addition of monobromobimane (mBrB). Highly reactive and non-fluorescent mBrB rapidly reacts with unbound cysteine moieties of the model peptide to form a fluorescent complex. The residual concentration in % relative to the VC of the Cys thus can be determined, and the depletion vs time and concentration matrix is used to calculate rate constants.

It has been shown that kDPRA differentiates GHS Cat 1A sensitizers from GHS 1B/not classified substances with a balanced accuracy of 85% (based on 180 test substances), with a sensitivity of 84% (38/45), and a specificity of 86% (116/135) relative to LLNA results (Natsch et al., 2020). In addition, the prediction of human skin sensitization for 123 test substances that fall within the kDPRA's applicability domain has a balanced accuracy of 76%, a sensitivity of 64% (21/33), and a specificity of 89% (80/90) (Natsch et al., 2020). On the basis of the overall data available ($n = 180$), the kDPRA's applicability domain was shown to include a variety of organic functional groups, the full range of skin sensitization potencies (as determined in *in vivo* studies), and diverse physicochemical properties.

After setting up a standard operating procedure (SOP) to conduct the kDPRA and to perform rate constant calculations, the primary goal of this study was to assess the transferability of the method (using six test substances) to five naïve labs and then assess intra- and inter-laboratory reproducibility of the method using 24 blind-coded substances. The reproducibility of the assay based on the log-transformed rate constants as well as the clas-

sification reproducibility to differentiate Cat 1A sensitizers from Cat 1B sensitizers/not classified substances according to UN GHS were evaluated.

2 Materials and methods

2.1 Test substance selection

Test substances for the ring trial for evaluation of transferability and reproducibility were selected based upon their published characterization for potency in mice and humans (Basketter et al., 2014; ICCVAM, 2011).

Transfer phase

During the transfer phase, 6 DPRA-positive sensitizers were tested twice: 2,4-dichloronitrobenzene (CAS RN 97-00-7), oxazolone (CAS RN 15646-46-5), formaldehyde (CAS RN 50-00-0), ethylene glycol dimethacrylate (CAS RN 97-90-5), benzylidene acetone (CAS RN 122-57-6), and 2,3-butanedione (CAS RN.431-03-8). While ethylene glycol dimethacrylate was the positive control originally reported for the kDPRA (Wareing et al., 2017), the remaining 5 are proficiency chemicals for the DPRA listed in OECD TG 442C (OECD, 2020). The transfer phase test substances were purchased from Sigma Aldrich and distributed by BASF non-blinded.

Inter- and intra-laboratory trial (reproducibility assessment)

The 24 test substances used during the blind-coded testing included two correct negatives and 21 correct positives in the Cys-only DPRA according to literature data (ICCVAM, 2011). The test set intentionally also included one sensitizer known to be negative/minimally reactive in the DPRA (i.e., dihydrocoumarin). The 24 test substances are listed in Table 1 including their protein reaction mechanistic domain assignment. This set of 24 test substances is intentionally strongly biased for positive test substances (sensitizers), as the kDPRA is intended for potency discrimination within test substances rated reactive. Two negative test substances were nevertheless included to also assess reproducibility for non-reactive test substances.

2.2 Participating laboratories

This ring trial validation study was conducted by a total of 7 laboratories (in alphabetical order): BASF SE Experimental Toxicology and Ecology (Germany), Charles River Laboratories Den Bosch BV (The Netherlands), Givaudan Schweiz AG (Switzerland), Institute for In Vitro Sciences, Inc. (USA), L'Oréal Research & Innovation (France), National Institute of Public Health (Czech Republic), Procter & Gamble (USA). The lead labs, BASF SE

Abbreviations

A, alanine; ACN, acetonitrile; C, cysteine; CAS RN, Chemical Abstracts Service Registry Number; Cat, category; Cys, cysteine; CN, correct negative; CP, correct positive; DNCB, 2,4-dinitrochlorobenzene; dp, depletion; DPRA, direct peptide reactivity assay; DSA05, dose per skin area that produces a positive response in 5% of the tested humans; EC3, estimated concentration leading to a stimulation index of 3 in the local lymph node assay; EGDMA, ethylene glycol dimethacrylate; F, phenylalanine; FN, false negative; FP, false positive; GHS, Globally Harmonized System of Classification, Labeling and Packaging of Chemicals; HPLC, high performance liquid chromatography; IC-CVAM, Interagency Coordinating Committee on the Validation of Alternative Methods; k, kinetic rate constant; kDPRA, kinetic DPRA; l, liquid; LLNA, local lymph node assay; Lys, lysine; MA, Michael acceptor; mBrB, monobromobimane; n, number of chemicals; NC, negative control; OECD, Organisation for Economic Co-operation and Development; OECD TG, OECD test guideline; PC, positive control; QP, quinone precursor; R, arginine; RC, reactive carbonyl; s, solid; SB, Schiff base former; SC, substance control; SD, standard deviation; SOP, standard operating procedure; SNAr, aromatics reacting by nucleophilic substitutions

Tab. 1: Test substances for the reproducibility (inter- and intra-laboratory) assessment

Test substance name	Abbreviation	Physical form ^a	CAS RN	LLNA EC3 (%)	ICCVAM ^b human potency (DSA ₀₅ , µg/cm ²)	GHS Cat LLNA	GHS Cat Human	Cys-depletion (%) ^c	Mechanistic domain
(Chloro)methylisothiazolinone	CMI	l	26172-55-4	0.009	5	1A	1A	96.3	SN ₂ -reaction
Glyoxal	GLY	l	107-22-2	0.6	345	1A	1A	56.5	Schiff base
Methylisothiazolinone	MI	s	2682-20-4	0.4	223.5	1A	1A	97.9	SN ₂ -reaction at the S-atom proposed
Methyl-2-octynoate	MOY	l	111-12-6	0.45	388	1A	1A	97.2	Michael acceptor
4-Phenylenediamine	PPD	s	106-50-3	0.15	30	1A	1A	95.3	Quinone methide(s)/imines
Tetrachlorsalicylanilide	TCS	s	1154-59-2	0.04	27	1A	1A	36.8	Acyl transfer
Isoeugenol	IE	l	97-54-1	1.35	1016	1A	1A	92.6	Quinone methide(s)/imines
Bourgeonal	BOU	l	18127-01-0	4.3	1541	1B	1B	17.7	Schiff base
Carvone	CAR	l	6485-40-1	12.9	19284	1B	1B	25.7	Michael acceptor
Dihydrocoumarin	DHC	l	119-84-6	5.6	759	1B	1B	0.0	Acyl transfer
Hydroxycitronellal	HC	l	107-75-5	22.2	5237	1B	1B	32.3	Schiff base
Imidazolidinyl urea	IU	s	39236-46-9	24	3846	1B	1B	38.4	Acyl transfer
Methylhexanedione	MHD	l	13706-86-0	26.0	3595	1B	1B	25.8	Schiff base
Perillaaldehyde ^d	PA	l	2111-75-3	4.04	1484	1B	1B	31.9	Michael acceptor / Schiff base
Phenyl benzoate	PB	s	93-99-2	18.33	52489	1B	1B	50.9	Acyl transfer
Phenylpropion aldehyde	PPA	l	93-53-8	6.3	692	1B	1B	37.4	Schiff base
Tetramethyldiuram disulfide	TMD	s	137-26-8	2.93	4544	1B	1B	99.5	SN ₂ -reaction at the S-atom proposed
Benzisothiazolinone	BI	s	2634-33-5	4.8	50	1B	1A	97.7	SN ₂ -reaction at the S-atom proposed
Benzylidene acetone	BA	s	122-57-6	3.7	299	1B	1A	93.5	Michael acceptor
δ-Damascone	DAM	l	57378-68-4	3.55	193	1B	1A	No data	Michael acceptor
Diethylmaleate	DEM	l	141-05-9	4.7	400	1B	1A	99.9	Michael acceptor
trans-2-Hexenal	HEX	l	6728-26-3	4.05	49	1B	1A	97.9	Michael acceptor / Schiff base
4-Methoxy-acetophenone	MAP	s	100-06-1	> 50	Non-sensitizer	NC	unknown	2.4	No alert
Chlorobenzene	CB	l	108-90-7	> 25	Non-sensitizer	NC	unknown	0.4	No alert

^a Abbreviations: l, liquid; s, solid; ^b ICCVAM (2011); ^c According to the Cys-only prediction model described in OECD TG 442C (OECD, 2020), test substances with a Cys-depletion of 13.89% or above are considered positive. ^d Substance is called p-Mentha-1,8-dien-7-al in the OECD DA SS database.

Experimental Toxicology and Ecology (Germany) and Givaudan Schweiz AG (Switzerland), were responsible for protocol authorship, organization of test-substance selection and procurement, and statistical evaluations. The other 5 laboratories were naïve to the kDPRA (thereof 2 were also naïve to the standard DPRA according to OECD TG 442C, (OECD, 2020)).

2.3 Kinetic direct peptide reactivity assay (kDPRA)

Procedure

Test substances were dissolved in acetonitrile (ACN) or in pH 7.5 phosphate buffer, if not soluble in ACN, to yield stock solutions of 20 mM. Thereafter, dilution series of 20, 10, 5, 2.5 and 1.25 mM were prepared.



The kDPRA consisted of the following steps: In case of ACN-soluble substances, 120 μL of 0.667 mM Cys-peptide solution in pH 7.5 phosphate buffer was added to each well of a black 96-well plate. Next, 40 μL of the respective substance solution was added to each well. This yielded 0.5 mM peptide concentration and substance concentrations of 5, 2.5, 1.25, 0.625 and 0.3125 mM (final ratios of peptide: test substance = 1:10, 1:5, 2:5, 4:5, 8:5). All substances were tested in triplicate within the same run.

In case of pH 7.5 phosphate buffer-soluble substances, 80 μL of 1.0 mM Cys-peptide solution in phosphate buffer (pH 7.5) was added to each well of a black 96-well plate. Next, 40 μL ACN was added and, finally, 40 μL of the respective substance solution. This yielded the same composition of samples as for the ACN-soluble substances described above.

Each 96-well plate comprised control samples as follows: 12 wells of a negative control (NC) containing the peptide and vehicle; 12 wells of blank control (BC) containing pH 7.5 phosphate buffer (without peptide) and the vehicle; 1 sample per concentration of the positive control (PC) cinnamic aldehyde, and 1 sample per substance and concentration of a substance control (SC) containing the respective test substance and the buffer but no peptide. The SC served for identification of interference of the test substance with the fluorescence measurement and as a background measurement.

The plates were sealed with impermeable foil directly after application of the substance, shaken for 5 min on a plate shaker and thereafter incubated in the dark at $25^\circ\text{C} \pm 2.5^\circ\text{C}$. The substances were incubated for 10, 30, 90, 150, 210 and 1440 min. After the respective reaction time, each test run was stopped by the addition of 40 μL of 3 mM mBrB solution (diluted in acetonitrile). Highly reactive non-fluorescent mBrB rapidly reacts with unbound cysteine moieties of the model peptide to form a fluorescent complex. The higher the intensity in fluorescence, the more cysteine moieties remained unbound after the respective reaction time and the less peptide-reactive is the test substance. The substance solutions containing mBrB were then further incubated for 5 min in the dark on a plate shaker. Then fluorescence was detected using an excitation filter of 390 nm and an emission filter of 480 nm.

Fluorescence intensities were normalized relative to the substance without the peptide (SC) as well as the phosphate buffer and acetonitrile (BC; background fluorescence). Relative peptide depletion was expressed as percent decrease in relation to the mean of the NC wells. Further, a pair-wise comparison of each substance concentration group with the VC was performed using the Welch t-test (two-sided) for the hypothesis of equal means.

Data evaluation

Depletion values were further evaluated if the criteria for positivity was reached at the highest concentration of one reaction time point (13.89% Cys-peptide depletion, based on the Cys-only prediction model described in OECD TG 442C (OECD, 2020)) and if the depletion was statistically significantly different ($p < 0.05$) from NC.

For each incubation time, the remaining (non-reacted) amount of Cys-peptide was determined and its natural logarithm plot-

ted against the respective substance concentration. For each time point for which the regression line gave a correlation > 0.9 , the calculated slope was divided by the incubation time to determine the second order reaction rate constant k ($\text{min}^{-1}\text{mM}^{-1}$) (Roberts and Natsch, 2009). This value was transformed to the rate constant in ($\text{s}^{-1}\text{M}^{-1}$) and the logarithm was taken. The maximum value observed at any time point was taken as the $\log k_{\text{max}}$ and used for the further evaluation.

Acceptance criteria

The results of a 96-well test plate were considered valid if the following conditions were met:

- Positive control (PC): The $\log k$ of the PC at 90 min is within the following range: $-1.75 \text{ M}^{-1}\text{s}^{-1}$ to $-1.40 \text{ M}^{-1}\text{s}^{-1}$. If no $\log k$ is obtained at 90 min, the value at 150 min is used instead and lies in the following range: $-1.90 \text{ M}^{-1}\text{s}^{-1}$ to $-1.45 \text{ M}^{-1}\text{s}^{-1}$
- Vehicle control (VC): The coefficient of variation of the 12 VC values of a plate is $< 12.5\%$ for at least 5 of the 6 time points.

If one or more of these criteria were not met, a repetition of the run was considered. Further, the runs for substances were repeated if non-linear behavior of results was obtained in order to exclude data bias due to artifacts, e.g., pipetting errors.

Prediction model

The maximum rate constant observed, $\log k_{\text{max}}$, is used in the kDPRA to distinguish between two levels of skin sensitization potency, i.e., to discriminate GHS subcategory 1A from GHS subcategory 1B/not classified. The prediction model was developed from a dataset on 180 substances with LLNA reference data (Natsch et al., 2020):

Reaction rate	kDPRA prediction
$\log k_{\text{max}} \geq -2.0$	GHS subcategory 1A
$\log k_{\text{max}} < -2.0$	GHS subcategory 1B or not classified

2.4 Transfer phase

For the transfer phase, the naïve laboratories received the protocol, an evaluation sheet and run validity criteria, and a preliminary proficiency range based on previous results of the lead labs with the same test substances. Two telephone conferences were held to clarify questions related to the conduct of the assay, but no hands-on training was conducted.

2.5 Blind-coded testing

In total, 24 different test substances were assessed under GLP-like conditions during the blind-coded testing in seven different laboratories. Test-substance procurement, blind-coding and distribution were conducted by an external service (BioTeSys GmbH, Esslingen, Germany). All seven participating labs tested all 24 test substances in one repetition (inter-laboratory reproducibility). Further, out of the 24 test substances, a random subset of 12 test substances was tested in two additional repetitions (with a different code for each run) in three or four labs (additional intra-laboratory reproducibility). Thus, for 12 test substances the intra-laboratory comparison was conducted in 3 labs and for the re-

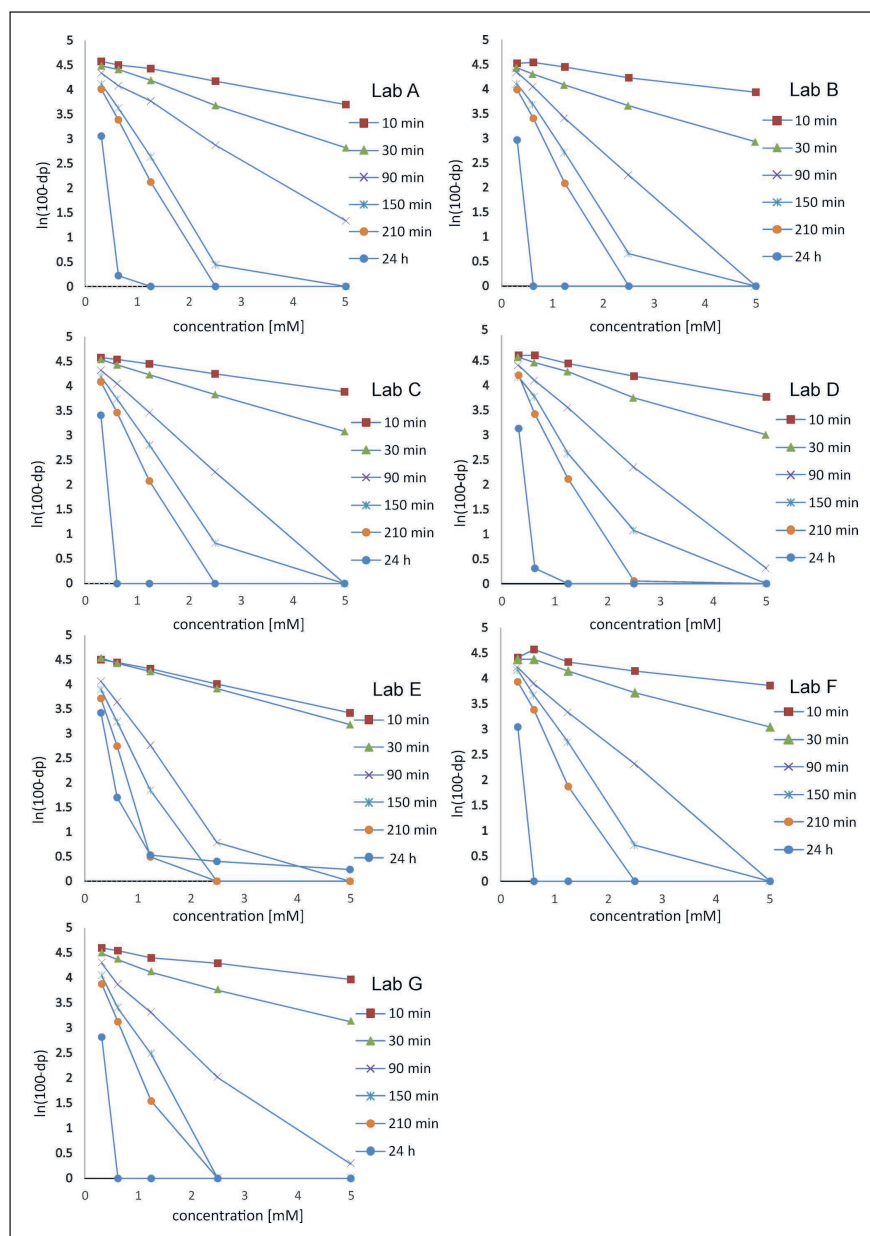


Fig. 1: Depletion matrices for DNCB in one run from all seven laboratories

The natural logarithm of the non-depleted peptide concentrations is plotted vs the concentration of the test substance at each time point for laboratories A-G.

maintaining 12 test substances in 4 labs (in total $12 \times 3 + 12 \times 4 = 84$ intra-laboratory comparisons). This was fully randomized, so no laboratory received the same test substances for intra-laboratory reproducibility. The codes Lab1 - Lab8 are thus attributed to different labs for each chemical, Lab 5 - Lab 8 always referring to the labs performing intra-laboratory testing for a particular chemical.

3 Results and discussion

To illustrate a typical result of the kDPRA assay, the depletion matrices for DNCB determined in the seven participating laboratories during the transfer phase are shown in Figure 1. For each exposure time point, the natural logarithm of the remaining

Cys-peptide (in % relative to the VC) is plotted against test substance concentration.

3.1 Transferability

The five naïve labs tested six (non-coded) known sensitizers and the positive control (PC) cinnamic aldehyde to establish the assay within their labs and to familiarize themselves with the study protocol. Overall, the results reported were very similar to those obtained by the two lead labs (Fig. 2) with the exception of formaldehyde, which was much less reactive in all three repetitions performed at Lab E.

The kDPRA was easily transferable to five naïve labs without hands-on training, and the SOP was found sufficiently detailed to perform the test in all participating laboratories. There were no

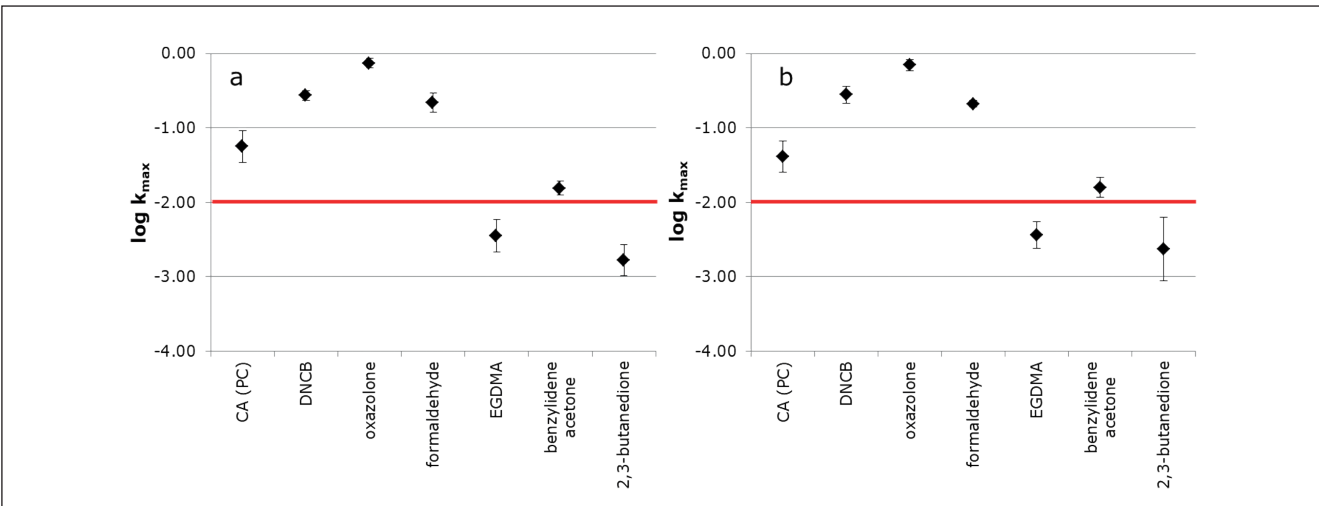


Fig. 2: Reproducibility of $\log k_{\max}$ values for the PC and six non-coded test substances for kDPRA establishment
Mean values and SDs of 4 runs (in two lead labs; panel A) and a total of 11 runs (in five naïve labs, panel B) with two or three valid runs conducted per naïve lab are shown. Note that the outlier Lab E was removed for the formaldehyde values. The solid red line indicates the cut-off $\log k_{\max} = -2.0$.

significant technical obstacles specific to the method. Figure 2 indicates that the six test substances first tested in the two lead laboratories resulted in very similar $\log k_{\max}$ values when tested in the five naïve labs. Moreover, the standard deviation (SD) from the four experiments in the two lead labs is similar to the SD in the eleven experiments in the five naïve labs, and thus the variability was not significantly increased by moving from the lead labs to the naïve labs.

3.2 Blind-coded testing

Once the laboratories had successfully tested the six substances of the transfer phase, they progressed to testing the 24 blind-coded

substances. After the test of the blind-coded substances was completed at the seven participating laboratories, all spreadsheets containing the blind-coded data were collected by the lead labs and provided to BioTeSys as an external and independent data repository site before the code for unblinding was provided by the latter. Analysis and biostatistics on the decoded data were then conducted at Givaudan.

Reproducibility of positive control (PC)

For the PC cinnamic aldehyde, the $\log k_{\max}$ values and the rate constants at a fixed time (90 min and 150 min) were reported for each experiment (summarized in Tab. 2).

Tab. 2: Reproducibility of positive control in the blind-coded phase: $\log k_{\max}$, $\log k_{90 \text{ min}}^b$ and $\log k_{150 \text{ min}}$ values ($\text{M}^{-1}\text{s}^{-1}$)

	$\log k_{\max}$ values ($\text{M}^{-1}\text{s}^{-1}$)				$\log k_{90 \text{ min}}$ values ($\text{M}^{-1}\text{s}^{-1}$) ^b				$\log k_{150 \text{ min}}$ values ($\text{M}^{-1}\text{s}^{-1}$)			
	AVG	SD	Min	Max	AVG	SD	Min	Max	AVG	SD	Min	Max
All labs (blind coded testing)	-1.35	0.13	-1.51	-1.15	-1.58	0.04	-1.64	-1.53	-1.66	0.04	-1.75	-1.62
Lab A	-1.37	0.16	-1.59	-0.87	-1.53	0.07	-1.66	-1.43	-1.62	0.05	-1.71	-1.50
Lab B	-1.44	0.14	-1.63	-1.16	-1.60	0.04	-1.68	-1.53	-1.67	0.03	-1.71	-1.57
Lab C	-1.32	0.19	-1.59	-0.99	-1.54	0.07	-1.66	-1.37	-1.63	0.06	-1.74	-1.51
Lab D	-1.46	0.23	-1.88	-0.81	-1.60	0.08	-1.75	-1.48	-1.75	0.07	-1.88	-1.62
Lab E	-1.15	0.23	-1.52	-0.73	-1.62	0.08	-1.75	-1.48	-1.65	0.06	-1.73	-1.53
Lab F	-1.51	0.18	-1.70	-1.09	-1.64	0.07	-1.75	-1.52	-1.69	0.04	-1.77	-1.65
Lab G	-1.22	0.15	-1.54	-1.01	-1.56	0.04	-1.64	-1.47	-1.64	0.09	-1.71	-1.26
All labs (transfer phase)^a	-1.35	0.22	-1.76	-0.94	-1.60	0.08	-1.73	-1.38	-1.68	0.09	-1.92	-1.45

^a Shown for comparison; ^b The $k_{90 \text{ min}}$ value is the rate used to decide on acceptability of an experiment. In case no rate is calculated at 90 min (reaction not linear or not statistically significant), then $k_{150 \text{ min}}$ can be considered. Labs had to report the 150 min value instead of the 90 min value in only 4 of 148 runs during the blind-coded testing. AVG, average; SD, standard deviation; Min, minimum; Max, maximum

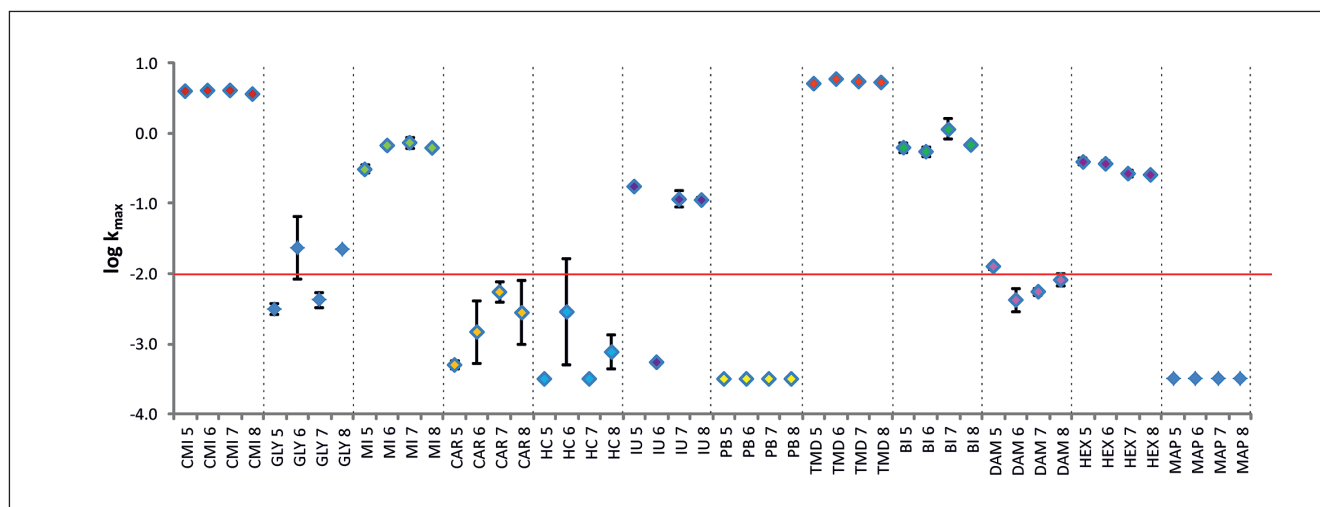


Fig. 3: Intra-laboratory testing: Variability expressed as average values and SD in repeated intra-laboratory testing (3 times each) in four labs

For non-reactive test substances ($\log k_{\max} < -3.46$ corresponding to Cys-depletion of $<13.89\%$ at 5 mM after 24 h), a default value of -3.5 was assigned to allow plotting the results. Abbreviated test substance names (see Tab. 1) and the number attributed to the test laboratory for testing that particular test substance are indicated on the x-axis. The solid red line indicates the cut-off $\log k_{\max} = -2.0$.

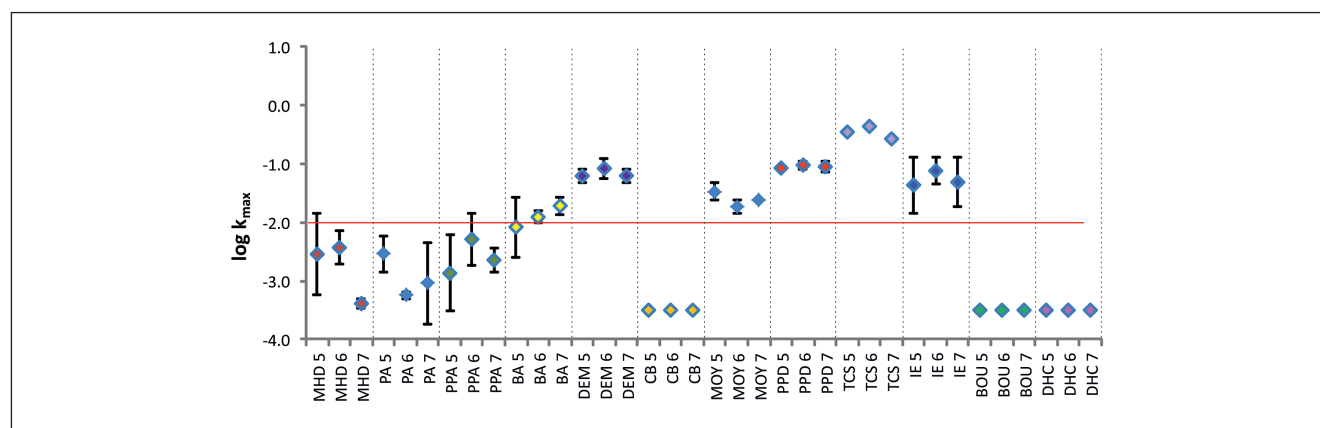


Fig. 4: Intra-laboratory testing: Variability expressed as average values and SD in repeated intra-laboratory testing (3 times each) in three labs

For non-reactive test substances ($\log k_{\max} < -3.46$, corresponding to Cys-depletion of $<13.89\%$ at 5 mM after 24 h) a default value of -3.5 was assigned to allow plotting the results. Abbreviated test substance names (see Tab. 1) and the number attributed to the test laboratory for testing that particular test substance are indicated on the x-axis. The solid red line indicates the cut-off $\log k_{\max} = -2.0$.

The average log k_{\max} of all valid runs during the blind-coded testing phase was -1.35, and this value is identical to the value obtained as average value of all labs in the transfer phase. Thus, very comparable results were obtained in both phases of this study. The average log k_{\max} value varied between -1.15 and -1.51 for the seven labs. The SD for intra-laboratory reproducibility of the log k_{\max} was between 0.14 and 0.23, similar to the average SD obtained for all test substances in the intra-laboratory reproducibility (0.158; see subsection on intra-laboratory reproducibility).

The inter-laboratory variability of the PC was even lower for the rate constants derived at 90 and 150 min, which are used as accep-

tance criteria, and log $k_{90 \min}$ varied between -1.53 and -1.64 for the seven labs (Tab. 2), while the intra-laboratory SD was between 0.04 and 0.08. The overall SD for all runs was 0.04. Log $k_{150 \min}$ varied between -1.62 and -1.75 for the seven labs (Tab. 2), while the intra-laboratory SD was between 0.03 and 0.09. The overall SD for all runs was 0.04.

Intra-laboratory reproducibility

The average and the SDs of log k_{\max} are shown in Figure 3 for test substances tested in four labs and in Figure 4 for test substances tested in three labs.

Tab. 3: Intra-laboratory reproducibility of the classification of test substances^a

Substance	Lab 5 rep1	Lab 5 rep2	Lab 5 rep3	Lab 6 rep1	Lab 6 rep2	Lab 6 rep3	Lab 7 rep1	Lab 7 rep2	Lab 7 rep3	Lab 8 rep1	Lab 8 rep2	Lab 8 rep3	Lab 5 AVG ^b	Lab 6 AVG	Lab 7 AVG	Lab 8 AVG	Labs with consistent repetitions
(Chloro) methylisothiazolinone	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	4 of 4
Glyoxal	1B	1B	1B	1B	1A	1A	1B	1B	1B	1A	1A	1A	1B	1A	1B	1A	3 of 4
Methylisothiazolinone	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	4 of 4
Methyl-2-octynoate	1A	1A	1A	1A	1A	1A	1A	1A	1A				1A	1A	1A		3 of 3
4-Phenylenediamine	1A	1A	1A	1A	1A	1A	1A	1A	1A				1A	1A	1A		3 of 3
Tetrachlorsalicylanilide	1A	1A	1A	1A	1A	1A	1A	1A	1A				1A	1A	1A		3 of 3
Isoeugenol	1A	1A	1A	1A	1A	1A	1A	1A	1A				1A	1A	1A		3 of 3
Bourgeonal	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r				n-r	n-r	n-r		3 of 3
Carvone	1B	1B	1B	1B	1B	1B	n-r	1B	1B	1B	1B	1B	1B	1B	1B	1B	4 of 4
Dihydrocoumarin	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r				n-r	n-r	n-r		3 of 3
Hydroxycitronellal	n-r	n-r	n-r	1B	1B	1B	n-r	n-r	n-r	1B	1B	n-r	n-r	1B	n-r	1B	4 of 4
Imidazolidinyl urea	1A	1A	1A	1B	n-r	n-r	1A	1A	1A	1A	1A	1A	1A	1B	1A	1A	4 of 4
Methylhexanedione	1B	1B	1B	n-r	1B	1B	1B	1B	1B				1B	1B	1B		3 of 3
Perillaaldehyde	1B	1B	1B	1B	1B	1B	1B	1B	1B				1B	1B	1B		3 of 3
Phenyl benzoate	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	4 of 4
Phenylpropionaldehyde	1B	1B	n-r	1B	1A	1B	1B	1B	n-r				1B	1B	1B		2 of 3
Tetramethyldiuram disulfide	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	4 of 4
Benzosiothiazolinone	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	4 of 4
Benzylidene acetone	1A	1A	1B	1A	1A	1A	1A	1A	1A				1B	1A	1A		2 of 3
δ-Damascone	1A	1A	1A	1B	1B	1B	1B	1B	1B	1B	1B	1B	1A	1B	1B	1B	4 of 4
Diethylmaleate	1A	1A	1A	1A	1A	1A	1A	1A	1A				1A	1A	1A		3 of 3
trans-2-Hexenal	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	4 of 4
4-Methoxy-acetophenone	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r	4 of 4
Chlorobenzene	1B	n-r	n-r	n-r	n-r	n-r	n-r	n-r	n-r				n-r	n-r	n-r		3 of 3
Total																	81 of 84

^a Abbreviations: n-r, non-reactive; ^b AVG indicates the rating of the test substance by the average log k_{\max} determined from all repetitions in a particular lab, this value is used for inter-laboratory reproducibility of class prediction below.

For most test substances, the intra-laboratory variability was low and the SDs between individual runs were very small. For those test substances with very low SD (below 0.3 on the logarithmic scale corresponding to a two-fold difference in the kinetic rate), the values reported by the different laboratories were also very close to each other. The average SD of the 24 test substances in intra-laboratory testing was 0.158, and the average was <0.1 for 10 test substances, with a further five test substances being non-reactive in all laboratories. These included the two non-sensitizers (4-methoxy-acetophenone and chlorobenzene) and 3,4-dihydro-

coumarin, which are also non-Cys-reactive in DPRA (Natsch et al., 2013). In addition, phenylbenzoate and bourgeonal were reproducibly non-reactive in kDPRA despite the fact that Cys-depletion had been reported in the DPRA (Bauch et al., 2011 and unpublished data). This low intra-laboratory variability of the rate constants and quantification of this variability should be considered an important aspect for uncertainty analysis when these data are later used in defined approaches for potency assessment on a continuous scale. For most validated *in vitro* assays, only reproducibility of hazard classification has been fully document-

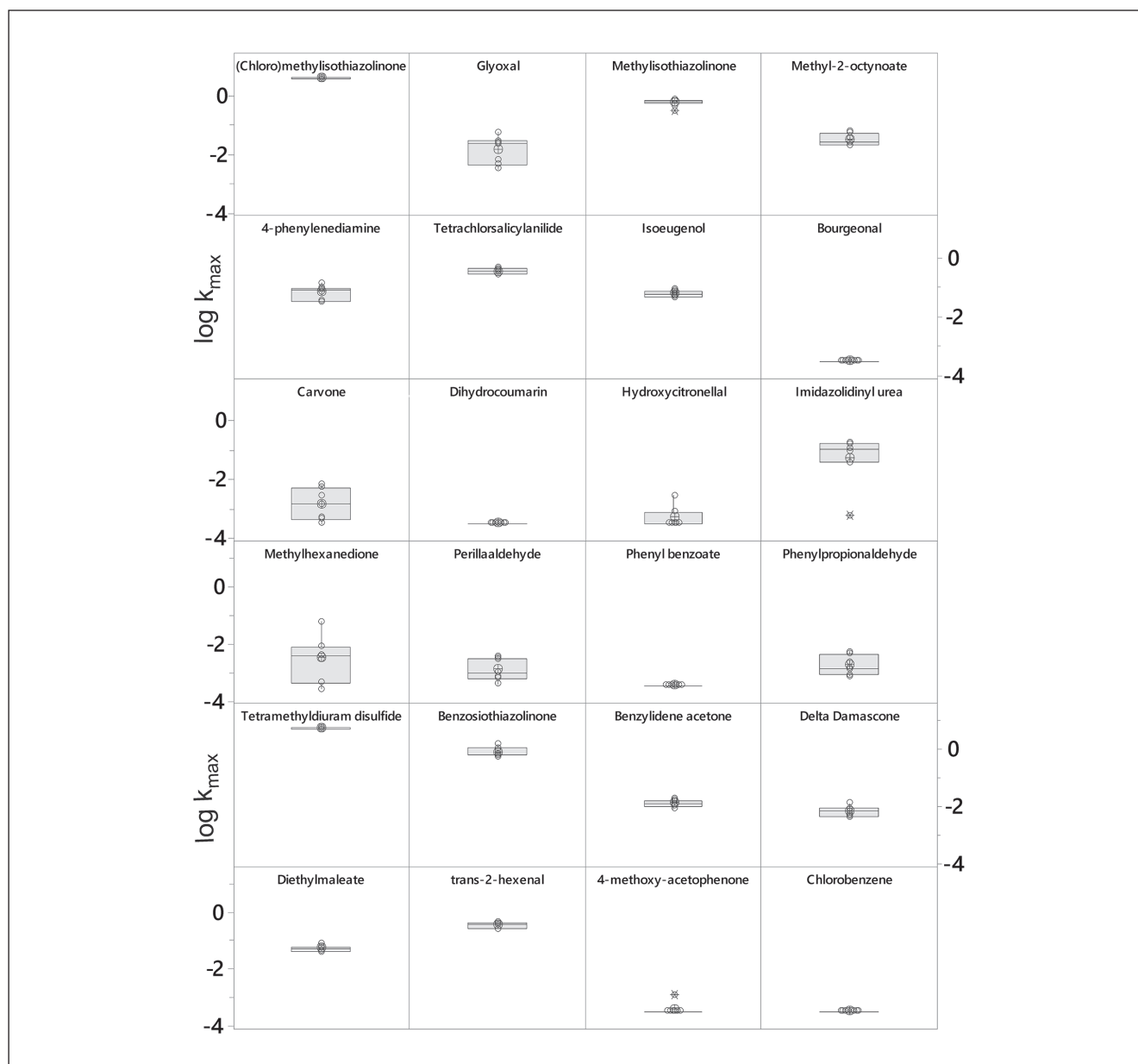


Fig. 5: Log k_{\max} values from inter-laboratory testing

The 7 individual lab results (circles; the average from repeated testing for labs that tested a particular test substance several times), the interquartile range box, and the average (horizontal line) are shown.

ed in validation studies but not reproducibility of the continuous outputs (peptide depletion, concentration-response for biological marker induction or cytotoxicity).

For four of the 24 test substances (CAR, MHD, PPA, IE, see Tab. 1 for full chemical names), a higher intra-laboratory variability was observed (logarithmic SD between 0.35 and 0.45), but this increased variability was observed in multiple labs (see below), and it appeared to be intrinsic to certain test substances. While the exact reason for enhanced variability for specific test substances is not known, some explanations can be proposed. Very slowly react-

ing test substances like CAR and MHD may have higher variation, as variation over a prolonged incubation time may be cumulative. Test substances triggering peptide oxidation like PPA are known to be subject to higher variability. For PPA, this had already been observed in the DPRA pre-validation study (reported in Dimitrov et al., 2016). Finally, pre-haptens that spontaneously oxidize, like IE, may be more variable, as autooxidation is known to be a self-catalyzed process and hence prone to more stochastic effects.

Intra-laboratory data for prediction of UN GHS sensitizer classes according to the prediction cut-off log k_{\max} -2 are provid-

**Tab. 4: log k_{\max} values ($M^{-1}s^{-1}$) from inter-laboratory testing**For laboratories 5-8, the average from repeated intra-laboratory testing is shown.^a

Substance	Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7	Lab 8	AVG	SD
(Chloro)methylisothiazolinone	0.58	0.64	0.57		0.60	0.61	0.61	0.56	0.60	0.029
Glyoxal	-1.57	-2.23	-1.25		-2.51	-1.64	-2.37	-1.65	-1.89	0.478
Methylisothiazolinone	-0.26	-0.15	-0.19		-0.51	-0.17	-0.13	-0.21	-0.23	0.130
Methyl-2-octynoate	-1.28	-1.60	-1.71	-1.22	-1.47	-1.73	-1.61		-1.52	0.202
4-phenylenediamine	-1.50	-1.17	-0.86	-1.47	-1.07	-1.02	-1.05		-1.16	0.238
Tetrachlorsalicylanilide	-0.55	-0.35	-0.44	-0.47	-0.46	-0.36	-0.57		-0.46	0.086
Isoeugenol	-1.07	-1.22	-1.23	-1.16	-1.36	-1.12	-1.31		-1.21	0.103
Bourgeonal	n-r	n-r	n-r	n-r	n-r	n-r	n-r		n-r	n-r
Carvone	-2.16	n-r	-3.37		-3.30	-2.83	-2.26	-2.56	-2.75	0.512
Dihydrocoumarin	n-r	n-r	n-r	n-r	n-r	n-r	n-r		n-r	n-r
Hydroxycitronellal	n-r	n-r	n-r		n-r	-2.55	n-r	-3.12	-2.83	0.383
Imidazolidinyl urea	-1.02	-0.73	-1.41		-0.76	n-r	-0.94	-0.95	-1.30	0.245
Methylhexanedione	-2.44	-3.61	-1.22	-2.12	-2.54	-2.43	-3.39		-2.54	0.795
Perillaaldehyde	-3.44	-3.19	-2.56	-2.44	-2.54	-3.25	-3.04		-2.92	0.401
Phenyl benzoate	n-r	n-r	n-r		n-r	n-r	n-r	n-r	n-r	n-r
Phenylpropionaldehyde	-3.17	-3.10	-2.37	-2.90	-2.87	-2.29	-2.64		-2.76	0.343
Tetramethyldiuram disulfide	0.75	0.74	0.73		0.71	0.77	0.74	0.73	0.74	0.020
Benzosiothiazolinone	-0.07	0.18	-0.21		-0.20	-0.26	0.06	-0.17	-0.10	0.161
Benzylidene acetone	-1.94	-1.98	-1.78	-1.82	-2.08	-1.91	-1.72		-1.89	0.125
δ-Damascone	-2.06	-2.34	-2.17		-1.90	-2.38	-2.26	-2.09	-2.17	0.169
Diethylmaleate	-1.38	-1.25	-1.32	-1.38	-1.21	-1.08	-1.20		-1.26	0.109
trans-2-Hexenal	-0.38	-0.37	-0.34		-0.41	-0.43	-0.57	-0.59	-0.44	0.100
4-Methoxy-acetophenone	n-r	-2.91	n-r		n-r	n-r	n-r	n-r	n-r	n-r
Chlorobenzene	n-r	n-r	n-r	n-r	n-r	n-r	n-r		n-r	n-r

^a Abbreviations: AVG, indicates the rating of the test substance by the average log k_{\max} determined from all repetitions in all labs; n-r, non-reactive; SD, standard deviation

ed in Table 3. In summary, using the kDPRA to differentiate GHS Cat 1A vs GHS Cat 1B/not classified, attribution to GHS Cat 1A was consistent in 81 of 84 instances, hence intra-laboratory reproducibility with the classification prediction model for identifying 1A test substances was 96%. Variable predictions were mostly observed for test substances with a log k_{\max} very close to the classification cut-off.

Based on data in Figures 3 and 4 and in Table 3, the kDPRA has proven to be very reproducible in intra-laboratory testing, both when predicting binary classification as well as when considering the numerical log k_{\max} values on a continuous scale.

Inter-laboratory reproducibility

Log k_{\max} values obtained during the inter-laboratory reproducibility assessment are summarized in Table 4 and Figure 5.

These data demonstrate the inter-laboratory variability of the 7 laboratories for all 24 test substances. The inter-laboratory reproducibility was high for most test substances with an average SD for inter-laboratory comparison of 0.244, which was slightly higher than in the intra-laboratory comparisons and which corresponds to a variation around the geometric mean of 1.75 fold. Again, quantification of the uncertainty of the rate constant determinations on a continuous scale is an important attribute for uncertainty analysis of subsequent models and risk assessments relying on log k_{\max} values. In general, similar test substances that had higher variability in intra-laboratory testing also exhibited higher inter-laboratory variability (GLY, CAR, MHD, PPA), indicating that this is an intrinsic property of the test substances and not due to experimental variability.

There is one significant outlier in the whole dataset: Lab E re-

Tab. 5: Inter-laboratory reproducibility for GHS category determination based on the cut-off log k_{\max} of -2.0
Data for the laboratories testing the test substances once.

Substance	Lab 1	Lab 2	Lab 3	Lab 4	Consistent 1A vs. 1B/NC?
(Chloro)methylisothiazolinone	1A	1A	1A		YES
Glyoxal	1A	1B	1A		NO
Methylisothiazolinone	1A	1A	1A		YES
Methyl-2-octynoate	1A	1A	1A	1A	YES
4-phenylenediamine	1A	1A	1A	1A	YES
Tetrachlorsalicylanilide	1A	1A	1A	1A	YES
Isoeugenol	1A	1A	1A	1A	YES
Bourgeonal	n-r	n-r	n-r	n-r	YES
Carvone	1B	n-r	1B		YES
Dihydrocoumarin	n-r	n-r	n-r	n-r	YES
Hydroxycitronellal	n-r	n-r	n-r		YES
Imidazolidinyl urea	1A	1A	1A		YES
Methylhexanedione	1B	1B	1A	1B	NO
Perillaaldehyde	1B	1B	1B	1B	YES
Phenyl benzoate	n-r	n-r	n-r		YES
Phenylpropionaldehyde	1B	1B	1B	1B	YES
Tetramethyldiuram disulfide	1A	1A	1A		YES
Benzosiothiazolinone	1A	1A	1A		YES
Benzylidene acetone	1A	1A	1A	1A	YES
δ -Damascone	1B	1B	1B		YES
Diethylmaleate	1A	1A	1A	1A	YES
trans-2-Hexenal	1A	1A	1A		YES
4-Methoxy-acetophenone	n-r	1B	n-r		YES
Chlorobenzene	n-r	n-r	n-r	n-r	YES
n consistent	22				
Reproducibility (%)	(22/24) 92%				

ported very low / no reactivity (in repeated intra-laboratory testing) for IU; thus, for an unknown reason, this lab obtained different results for test substances related to formaldehyde. This outlier is highly consistent for IU and formaldehyde in that particular lab, and it appears to be linked to the chemistry. Formaldehyde does form a reversible peptide-adduct (data not shown), and for unknown reasons the reaction must have been reversed prior to reaction or during the reaction with mBrB in that particular laboratory.

Benzylidene acetone was tested both in the transfer phase and in the blind-coded phase. The log k_{\max} value from the seven labs in the blind-coded phase was -1.89 ± 0.13 , while

it was -1.80 ± 0.12 in the transfer phase, hence a very similar result and similar variability was observed in both open and blinded testing.

When applying the prediction cut-off (log $k_{\max} = -2.0$) to differentiate GHS Cat 1A vs GHS Cat 1B/not classified, a consistent result was obtained for 22 test substances (when testing each test substance once, Tab. 5) and for 20 test substances (when testing each test substance three times, Tab. 6). This resulted in an inter-laboratory reproducibility for identifying 1A test substances of 92% (for laboratories testing the test substance once) and 83% (laboratories testing the test substances three times), respectively, with an average for the two indepen-



Tab. 6: Inter-laboratory reproducibility for GHS category determination based on the log k_{\max} cut-off of -2.0
Data for laboratories testing the test substances three times.^a

Substance	Lab 5	Lab 6	Lab 7	Lab 8	Consistent 1A vs. 1B/NC ?
(Chloro)methylisothiazolinone	1A	1A	1A	1A	YES
Glyoxal	1B	1A	1B	1A	NO
Methylisothiazolinone	1A	1A	1A	1A	YES
Methyl-2-octynoate	1A	1A	1A		YES
4-phenylenediamine	1A	1A	1A		YES
Tetrachlorsalicylanilide	1A	1A	1A		YES
Isoeugenol	1A	1A	1A		YES
Bourgeonal	n-r	n-r	n-r		YES
Carvone	1B	1B	1B	1B	YES
Dihydrocoumarin	n-r	n-r	n-r		YES
Hydroxycitronellal	n-r	1B	n-r	1B	YES
Imidazolidinyl urea	1A	1B	1A	1A	NO
Methylhexanedione	1B	1B	1B		YES
Perillaaldehyde	1B	1B	1B		YES
Phenyl benzoate	n-r	n-r	n-r	n-r	YES
Phenylpropionaldehyde	1B	1B	1B		YES
Tetramethyldiuram disulfide	1A	1A	1A	1A	YES
Benzosiothiazolinone	1A	1A	1A	1A	YES
Benzylidene acetone	1B	1A	1A		NO
δ-Damascone	1A	1B	1B	1B	NO
Diethylmaleate	1A	1A	1A		YES
trans-2-Hexenal	1A	1A	1A	1A	YES
4-Methoxy-acetophenone	n-r	n-r	n-r	n-r	YES
Chlorobenzene	n-r	n-r	n-r		YES
n consistent	20				
Reproducibility (%)	(20/24) 83%				

^a The average log k_{\max} from three intra-laboratory experiments was taken to make the prediction.

dent evaluations of 88%.

When evaluating these values against other published validation studies where each test substance was typically tested three times, this analysis is a bit more stringent as 50% of the comparisons are made in 4 labs, and the random chance of congruent results falls from 25% to 13% with testing in 4 labs instead of 3 labs (i.e., the more labs, the higher the chance of producing one deviating result).

4 Conclusion

The kDPRA proved to be transferable to laboratories without hands-on training, and highly reproducible results for the positive control were obtained. The within-laboratory reproducibility of the kDPRA for assigning GHS Cat 1A was 96%, and the between-laboratory reproducibility for 24 test substances was 88%.

The average SD of the logarithmic rate in intra-laboratory testing was 0.158, which corresponds to a variation around the geometric mean of k_{\max} of 1.44-fold, while the average SD for inter-laboratory comparison was 0.244, which corresponds to a variation of 1.75-fold. This quantification of variability on a continuous

scale will be beneficial for uncertainty analysis in risk assessment.

Following an independent peer review, the kDPRA validation study was considered to demonstrate that the method is acceptable for the predictive identification of skin sensitization potency categories. The kDPRA has been added to the OECD Work Plan for the Test Guidelines Program for inclusion as an additional method in OECD TG 442C as project 4.317. Details and the draft updated OECD TG 442C will be discussed with the OECD expert group on skin sensitization during 2020.

References

- Basketter, D. A., Alepee, N., Ashikaga, T. et al. (2014). Categorization of chemicals according to their relative human skin sensitizing potency. *Dermatitis* 25, 11-21. doi:10.1097/DER.0000000000000003
- Bauch, C., Kolle, S. N., Fabian, E. et al. (2011). Intralaboratory validation of four in vitro assays for the prediction of the skin sensitizing potential of chemicals. *Toxicol In Vitro* 25, 1162-1168. doi:10.1016/j.tiv.2011.05.030
- Dimitrov, S., Detroyer, A., Piroird, C. et al. (2016). Accounting for data variability, a key factor in in vivo/in vitro relationships: Application to the skin sensitization potency (in vivo LLNA versus in vitro DPRA) example. *J Appl Toxicol* 36, 1568-1578. doi:10.1002/jat.3318
- ICCVAM (2011). ICCVAM Test Method Evaluation Report: Usefulness and Limitations of the Murine Local Lymph Node Assay for Potency Categorization of Chemicals Causing Allergic Contact Dermatitis in Humans. *NIH Publication No. 11-7709*.
- Natsch, A., Ryan, C. A., Foertsch, L. et al. (2013). A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation. *J Appl Toxicol* 33, 1337-1352. doi:10.1002/jat.2868
- Natsch, A., Haupt, T., Wareing, B. et al. (2020). Predictivity of the kinetic direct peptide reactivity assay (kDPRA) for sensitizer potency assessment and GHS subclassification. *ALTEX* 37, 652-664. doi:10.14573/altex.2004292
- Nukada, Y., Miyazawa, M., Kazutoshi, S. et al. (2013). Data integration of non-animal tests for the development of a test battery to predict the skin sensitizing potential and potency of chemicals. *Toxicol In Vitro* 27, 609-618. doi:10.1016/j.tiv.2012.11.006
- OECD (2014). The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. *OECD Series on Testing and Assessment, No. 168*. OECD Publishing, Paris. doi:10.1787/9789264221444-en
- OECD (2020). Test No. 442C: In Chemico Skin Sensitisation: Assays addressing the Adverse Outcome Pathway key event on covalent binding to proteins. *OECD Guidelines for the Testing of Chemicals, Section 4*. OECD Publishing, Paris. doi:10.1787/9789264229709-en
- Roberts, D. W. and Natsch, A. (2009). High throughput kinetic profiling approach for covalent binding to peptides: Application to skin sensitization potency of Michael acceptor electrophiles. *Chem Res Toxicol* 22, 592-603. doi:10.1021/tx800431x
- Wareing, B., Urbisch, D., Kolle, S. N. et al. (2017) Prediction of skin sensitization potency sub-categories using peptide reactivity data. *Toxicol In Vitro* 45, 134-145. doi:10.1016/j.tiv.2017.08.015

Conflict of interest

The authors of this publication are employees of the participating companies or organizations. The contribution to this publication was part of their work and is paid by their salaries.

Acknowledgments

The National Institute of Public Health (Czech Republic) was supported by ERDF/ESF project "International competitiveness of NIPH in research, development and education in alternative toxicological methods" (No. CZ.02.1.01/0.0/0.0/16_019/0000860).