

TOXICOLOGICAL SCIENCES

Critical evaluation of low-molecular weight respiratory sensitizers and their protein reactivity potential towards lysine residues

Nora L. Krutz^{*1}, Ian Kimber², Cindy A. Ryan³, Petra S. Kern¹, G. Frank Gerberick⁵

¹NV Procter & Gamble Services Company SA, Strombeek-Bever, Belgium.

² University of Manchester, Manchester, UK.

³The Procter & Gamble Company, Mason, OH, United States. Present address: ToxTech Solutions, LLC, Milan, IN, United States

⁵GF3 Consultancy LLC, West Chester, OH, United States.

***corresponding author:** Dr. Nora L. Krutz

NV Procter & Gamble Services Company SA
Temselaan 100, 1853 Strombeek-Bever, Belgium
Phone: +32 245 66040; Email: krutz.n@pg.com

Keywords: Respiratory sensitization, *in chemico*, DPRA, peptide reactivity

Running Head: Protein reactivity potential of respiratory sensitizers

Abstract

Interest in the development of methods to evaluate the respiratory sensitization potential of low molecular weight chemicals continues, but no method has yet been generally accepted or validated.

A lack of chemical reference standards, together with uncertainty regarding relevant immunological mechanisms, have hampered method development. The first key event in the development of either skin or respiratory sensitization is the formation of stable adducts of the chemical with host proteins. This event is measured in the Direct Peptide Reactivity Assay (DPRA) using cysteine- and lysine-containing model peptides. It is hypothesized that protein reactivity and subsequent adduct formation may represent the earliest point of divergence in the pathways leading to either skin or respiratory sensitization.

DPRA data for 200 chemicals were compiled and grouped into respiratory, skin and non-sensitizers. Chemicals grouping was based on extensive literature research and expert judgment.

To evaluate if chemical groups represent different peptide reactivity profiles, peptide reactivity data were clustered and compared with information on protein binding mechanisms and chemical categories available via the OECD Toolbox.

Respiratory sensitizers (n=15) showed a significant (3-fold) higher lysine reactivity than skin sensitizers (n=129). However, this difference was driven largely by the high representation of acid anhydrides among the respiratory sensitizers that showed clear lysine selectivity.

Collectively, these data suggest that preferential reactivity for either cysteine or lysine is associated primarily with chemical structure, and that lysine preference is not a unifying characteristic of chemical respiratory allergens.

Introduction

Many low molecular weight chemicals (900 Dalton) have the potential to induce skin sensitization resulting in allergic contact dermatitis (ACD). Others, far fewer in number, are able to induce sensitization of the respiratory tract associated with respiratory allergy and asthma (Kimber et al. 2011). Although the distinction between contact allergens and chemical respiratory allergens is not absolute, there are clearly differences between chemical allergens with respect to the characteristics of immune responses they preferentially elicit, and the form that sensitization will take. The factors that drive different pathways of sensitization remain unclear, and it is likely that there are a number of important influences (Kimber et al. 2011; Kimber et al. 2018).

There has been a continuing commitment to the development of methods for evaluation the respiratory sensitizing potential of chemicals, but no method has yet been formally validated or widely accepted (Arts 2020; Basketter et al. 2017; Kimber et al. 2018). Several factors have contributed to the lack of such a method, including uncertainty about the precise immunological mechanisms through which chemicals cause sensitization of the respiratory tract (Arts 2020; Isola et al. 2008; Kimber et al. 2010). Moreover, there has often been confusion about the discrimination between true chemical respiratory allergens that, by definition, induce sensitization via immunological processes, and chemicals that cause occupational asthma through other, non-immunological, mechanisms (most recently reviewed by Arts 2020). As a result of this confusion, it has been problematic to compile definitive lists of chemicals that provide reference standards for evaluating the ability of test methods to distinguish between chemical respiratory allergens, contact allergens and non-sensitizing chemicals.

Recently similarities and differences between contact allergens and chemical respiratory allergens have been mapped in Adverse Outcome Pathways (AOP) for skin and respiratory sensitization (Kimber et al. 2014; Kimber et al. 2018; Sullivan et al. 2017). One possibility is that mechanistic

aspects of the first Key Event (KE)/Molecular Initiating Event (MIE) in the pathways leading to acquisition of sensitization, that is the formation of stable adducts with host proteins, represents the earliest point of divergence between contact allergens and respiratory allergens. One potentially important aspect of the electrophilic interaction of chemical allergens with host proteins is selectivity for different amino acid residues, and in particular, selective association with either cysteine or lysine. A number of previous investigations have sought to determine whether the preferential association with cysteine or lysine provides a signature for contact or respiratory allergens (Dik et al. 2016; Lalko et al. 2011; Lalko et al. 2012). Results have been somewhat variable, and therefore there is a need now for a more systematic approach to address this question. We here describe a detailed examination of the amino acid reactivity profiles of respiratory and skin sensitizers. For this purpose, we have attempted to distinguish carefully between true chemical respiratory allergens, and contact allergens that lack the potential to induce sensitization of the respiratory tract. From these groups of chemical allergens (and from chemicals identified as lacking sensitizing activity) we have selected for further analysis those for which there are data available on their behavior in the OECD 442C compliant *in vitro* predictive test for sensitizing chemicals, the Direct Peptide Reactivity Assay (DPRA) (Gerberick et al. 2004). We here describe patterns of reactivity with cysteine and lysine displayed by skin and respiratory sensitizing chemicals, and by non-sensitizing chemicals.

Materials and Methods

DPRA data set

A set of chemicals with Direct Peptide Reactivity (DPRA) data - according to OECD 442c (OECD 2020b) - was derived from published sources (Avonto et al. 2018; Dik et al. 2016; Gerberick et al. 2004; Gerberick et al. 2007; Hirota et al. 2018; Hoffmann et al. 2018; Lalko et al. 2013; Lalko et

1
2
3 93 al. 2012; Natsch et al. 2013; OECD 2020a; Urbisch et al. 2015; Wareing et al. 2017; Yamashita et
4
5 94 al. 2015) and in-house generated data for four chemicals (unpublished). Missing CAS numbers
6
7 95 were amended and a general test chemical name was chosen per CAS number. The DPRA data
8
9 96 was adapted as follows: negative peptide depletion values were set at 0% depletion, all peptide
10
11 97 depletion values were rounded to a whole number, and redundant data points were excluded from
12
13 98 further analyses. Data points were considered as redundant if both cysteine and lysine peptide
14
15 99 depletion values were identical after rounding up to whole numbers. Chemicals were excluded
16
17 100 from the data set for one of the following reasons: they were natural products without defined
18
19 101 structures (e.g., treemoss, jasmine, ylang ylang) or other mixtures (e.g., kathon CG, the mixture of
20
21 102 2,4-/2,6-toluene diisocyanate); co-elution occurred with either the cysteine or lysine peptide; there
22
23 103 was uncertainty about chemical identification; the evidence for clear grouping as skin or
24
25 104 respiratory sensitizer was insufficient.. In total, 200 chemicals with 845 DPRA data points and 337
26
27 105 non-redundant DPRA data points for each, cysteine and lysine reactivity, were used in analyses
28
29 106 (Suppl. Table 1).
30
31
32
33
34 107 The 200 chemicals were further annotated with chemical category information according to “US-
35
36 108 EPA New Chemical Categories” and protein binding mechanism according to Organisation for
37
38 109 Economic Co-operation and Development (OECD) and OASIS which were derived from the
39
40 110 Quantitative Structure Activity Relationship (QSAR) Toolbox 4.4 (OECD 2020a) and are
41
42 111 provided in Suppl. Table 1. “US-EPA New Chemical Categories” is a profiler to identify whether
43
44 112 a chemical belongs to one of the 66 chemical categories analyzed in the Environmental Protection
45
46 113 Agency (EPA) program and is used for hazard assessments of chemicals by the EPA (OECD
47
48 114 2020a; US EPA 2010). Protein binding by OECD (102 alerts) and by OASIS (112 alerts) were
49
50 115 used as profilers to identify protein binding mechanisms (OECD 2020a). Results of both profilers
51
52 116 were consolidated and in case of mixed results (e.g., Acetylation by OECD and Schiff base
53
54
55
56
57
58
59
60

formation by OASIS) were combined. In case of “No alert found” by only one of the profilers, the result was overruled by the alert of the other profiler.

Chemical Grouping

Prior to analysis, the 200 chemicals with DPRA data were grouped according to the following descriptors: respiratory sensitizer (n=15), skin sensitizer (n=129), or non-sensitizer (n=56). It is well known that the number of chemicals for which there is strong evidence to support a respiratory sensitizer grouping are far fewer than skin sensitizers (Arts 2020; Kimber et al. 2014). Much effort on developing QSAR alerts and databases has been done, but lesser effort on the identification of true respiratory sensitizers versus potential respiratory sensitizers and non-immunological asthmagens (Arts 2020). To assure that all chemicals selected as respiratory allergens have strong supportable evidence, a thorough expert review of the literature was conducted using a weight of evidence approach. Clinical evidence in humans that the substance can lead to sensitization of the respiratory tract associated with allergic symptoms required the availability of confirmed cases of allergic asthma or other related allergic respiratory conditions (rhinitis, extrinsic allergic alveolitis) along with an immunological assessment (e.g., IgE levels by ELISA, skin prick test data). It was also critical to have reliable data supporting the extent and type of exposure to the specific chemical being evaluated. Thus, there was a need for confirmation of direct exposure to the chemical. Chemicals where the evidence available suggested the asthma symptoms were caused by non-immunological mechanisms as well as agents causing anaphylaxis were not included in the respiratory sensitizer grouping.

For the skin sensitizer and non-sensitizer analyzed in this paper, the groupings were based on those proposed in previous publications (Api et al. 2017; Basketter et al. 2014; Hoffmann et al. 2018). Briefly, existing LLNA data were used along with human evidence. The human evidence consisted

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

141 of data from human maximization tests, human repeat insult patch tests, and/or diagnostic patch

142 tests. The LLNA data were collected from the literature, LLNA databases (e.g., NICEATM) and

143 from the proprietary database of the Research Institute for Fragrance Materials (RIFM). The

144 LLNA studies were conducted according to the OECD test guideline 429 ‘Skin sensitization: local

145 lymph node assay’ (OECD 2010) or similar protocols.

146

147 *Data analysis*

148 The arithmetic mean of the non-redundant lysine and cysteine peptide depletion per chemical was

149 calculated and used to generate boxplots to reflect the distribution of lysine and cysteine peptide

150 reactivity according to the sensitizing groups, chemical categories and protein binding

151 mechanisms.

152 Boxplots, the cluster analysis as well as graphs were generated using JMP® Pro 15 (SAS Institute

153 Inc., Cary, North Carolina, US). A Student’s t-test (with $\alpha = 0.01$; ns = $p > 0.05$, ** = $p \leq 0.01$,

154 *** = $p \leq 0.001$) was conducted to test whether the two groups of respiratory and skin sensitizers

155 are equal with regard to their cysteine, lysine and average peptide reactivity or not using JMP®

156 Pro 15. For the cluster analysis, one to ten clusters were examined using K-Means cluster analysis

157 and the most suitable number of clusters was identified based on the cubic clustering criterion

158 (CCC) to minimizing the within-cluster sum of squares. The distribution of chemical categories

159 per cluster was generated with Microsoft® Excel ® for Office 365 MSO 16 (ProPlus).

160

Page 7

Results

Peptide reactivity per sensitizer groups

The arithmetic means of the non-redundant cysteine and lysine DPRA depletion values were calculated for each chemical (n=200) and chemicals were grouped as respiratory sensitizers (n=15), skin sensitizers (n=129), or non-sensitizers (n=56). Respiratory sensitizers were determined based on weight of evidence and expert judgement and are listed in Table 1.

In order to display the distribution of the lysine and cysteine peptide depletion data and identify potential differences between the sensitizer/non-sensitizer groups, boxplots were used (Figure 1).

A Student's t-test (with $\alpha = 0.01$) was conducted to identify significant differences between respiratory and skin sensitizers. Significant differences were found with regard to lysine ($p < 0.0001$) and average of lysine and cysteine ($p = 0.0089$), but not cysteine ($p = 0.9933$) peptide reactivity. The results are aligned with the assumption that respiratory sensitizers are more lysine than cysteine reactive.

However, the variability (interquartile range box and whiskers, outliers) is remarkably high for all groups. The median of lysine peptide reactivity shows a nine times higher lysine depletion value for respiratory than skin sensitizers (64% versus 7%), whereas the arithmetic mean shows a three times higher lysine depletion value than skin sensitizers (52% versus 17%). This may be partly explained by the number of outliers (n=13, including 4 acrylates) in the group of skin sensitizers for lysine reactivity (highly reactive compounds towards both peptides with the exception of bromohexane (CAS# 111-25-1), but also by the high variability and the low number of respiratory sensitizers.

Non-sensitizers demonstrated minimal reactivity towards the cysteine and lysine peptides (in average below 6.4%) except 19 chemicals (including 7 acids) that showed peptide depletion despite not being considered as skin sensitizers (see comment section in Suppl. 1).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

185

186 *Clustering chemicals according to peptide reactivity*

187 To identify potential peptide reactivity patterns or clusters of all 200 chemicals, a principal
188 component (PC) analysis with 2 variables, cysteine and lysine peptide depletion, was conducted.
189 Using a k-Means analysis (one to ten clusters) and the cubic clustering criterion (CCC), eight
190 clusters were identified as the most suitable number of clusters and are shown as a Biplot (Figure
191 2) and as values in a table (Table 2).

192 The eight clusters can be partitioned into the top three clusters with the highest average lysine
193 depletion ($\geq 58\%$) with the number 1, 3 and 6 as well as the top three clusters with the highest
194 average cysteine depletion ($\geq 73\%$) with the number 1, 5, 6, and 8. Cluster 6 contains chemicals
195 with the highest average cysteine and lysine peptide depletion (both $\geq 87\%$), cluster 8 those with
196 only high average cysteine depletion ($\geq 95\%$) and cluster 3 with only high lysine depletion ($\geq 73\%$).
197 Cluster 4 represents minimally reactive chemicals (both $\leq 3\%$ depletion) and, clusters 2 and 7,
198 somewhat medium cysteine and lysine reactive chemicals, respectively. Clusters 1 and 2 have the
199 highest standard deviation, but all clusters show standard deviations (based on mean values from
200 different publications) that are below or at the max assay standard deviation for replicates for
201 cysteine ($< 14.9\%$) and lysine depletion ($< 11.6\%$) (OECD 2020b).

202 Six of eight clusters contain at least one of the 15 respiratory sensitizers including cluster 4, the
203 minimal reactivity cluster. Cluster 7 (somewhat medium cysteine and lysine reactivity) and 8 (only
204 high cysteine reactivity) contain none of the respiratory sensitizers. The majority of the respiratory
205 sensitizers (Table 1) represent anhydrides and diisocyanates, which separate in two distinct clusters
206 with the highest number of respiratory sensitizers: Cluster 3 (only high lysine reactive chemicals,
207 n=8) with five respiratory sensitizers and cluster 5 (high cysteine reactive chemicals, n=19) with
208 four respiratory sensitizers. Cluster 3 contains phthalic anhydride, cyclohexane dicarboxylic

Page 9

anhydride (or hexahydrophthalic anhydride), benzene-tricarboxylic anhydride (or trimellitic anhydride), tetrachlorophthalic anhydride, and glutaraldehyde as well as bromohexane (skin sensitizer) and two non-sensitizers. This cluster represents an average cysteine depletion of 12% and an average lysine depletion of 73% (Table 2). Almost all anhydrides of this data set are located in cluster 3, with the exception of maleic anhydride, which is in a cluster with higher cysteine depletion (cluster 1, label number 150) and chlorendic anhydride, which is – together with the respiratory sensitizer piperazine (label number 175) – in a cluster of minimal peptide depletion (cluster 4, label number 99). Cluster 5 contains hexamethylene diisocyanate, 2,6-toluene diisocyanate, methylenediphenyl diisocyanate, isophorone diisocyanate as well as 13 skin sensitizers and 1 non-sensitizer. This cluster contains almost all diisocyanates, except 2,4-toluene diisocyanate (cluster 2, label number 28). Cluster 5 represents 95% average cysteine depletion and 31% average lysine depletion (Table 2).

Peptide reactivity per chemical category and protein binding mechanism

In order to further elucidate, if the chemical categories play a role in the overall lysine and cysteine peptide reactivity of the sensitizer/non-sensitizer groups, chemicals were annotated with their US-EPA New Chemical Categories derived from the QSAR Toolbox 4.4 (OECD 2020a).

In total, 128 of the 200 chemicals were categorized according to the US-EPA New Chemical Categories, thus 36% were not categorized. Of the 128 categorized chemicals, 14 are respiratory sensitizers (missing: chloramine-T, label number 98), 84 are skin sensitizers and 30 are non-sensitizers with a total of 21 different chemical categories identified. In alignment with the cluster analysis results, anhydrides tend to be more lysine- versus cysteine-reactive with the exception of maleic anhydride (label number 150) and chlorendic anhydride (label number 99) and diisocyanates tend to be more cysteine- versus lysine-reactive (Figure 3). No other chemical

category shows a clear distinction between cysteine and lysine reactivity or the difference is based on an $n=1$ such as e.g. aldehydes grouped as respiratory sensitizer. Furthermore, mechanistic protein binding domains were analyzed to determine whether or not there was a preferential lysine or cysteine reactivity. Protein binding alerts generated using OECD and OASIS from the QSAR Toolbox 4.4 (OECD 2020a) found that for 103 of the 200 chemicals (13 respiratory sensitizers, 84 skin sensitizers and 6 non-sensitizers), none of the 13 distinct groups of protein binding mechanisms was particularly lysine- or cysteine-reactive (Figure 4). Mechanisms such as acetylation and SNAr, as well as partly Michael addition, only showed a tendency towards higher lysine reactivity compared to other binding mechanisms.

Chemical categories per cluster

The clusters identified by principal component analysis and the analysis of the peptide reactivity per chemical category showed that peptide-binding preference can be specific for chemical categories such as anhydrides and diisocyanates. To differentiate, if clusters represent specific chemical categories (as in the case of anhydrides and diisocyanates) or if chemical categories can be found in various clusters, a distribution of the 21 chemical categories in the eight different clusters is visualized (Figure 5).

Figure 5 shows that certain chemical categories (e.g. aldehydes, esters and phenols) can be found in almost all clusters and certain clusters (e.g. cluster 4) contain multiple chemical categories (12 of 21). As shown in Figure 2, anhydrides can be found predominantly in cluster 3, but also in cluster 1 and 4. However, cluster 3 also contains other chemical categories such as neutral organics, as well as the promiscuous aldehydes and esters. Diisocyanates are predominantly in cluster 5, which in turn also contains e.g. polynitroaromatics, acrylates and ester. Cluster 4, which contains the minimally peptide reactive chemicals is the most diverse (12 chemical categories) and with the

highest number of chemicals (n=90) of which 36 chemicals are not categorized. All other clusters revealed four (cluster 3) to eight different chemical categories per cluster.

Discussion

One of the most interesting, and seemingly intractable, questions in the study of allergic sensitization is why some chemical sensitizers are associated with respiratory allergy and asthma, whereas others are associated only with skin sensitization and allergic contact dermatitis (ACD). The question is of more than simply academic interest because the answer might hold the key to elucidating mechanistic differences between classes of chemical allergens, and provide a basis for predictive identification of chemical respiratory allergens; a goal that has so far remained elusive. It has been argued that there must exist a point of divergence on the pathway leading to sensitization where contact allergens and respiratory allergens diverge such that different qualities of immune response are elicited (Kimber et al., 2018). One theory is that the nature of initial interaction with proteins to form an immunogenic hapten-protein conjugates represents that point of divergence. This theory has been lent credence by results of preliminary studies that reported differences between contact and respiratory allergens with respect to amino acid selectivity (Hettick et al. 2012; Holsapple et al. 2006; Hopkins et al. 2005; Lalko et al. 2012). A theory has been that contact allergens favor cysteine-rich cellular proteins, whereas respiratory allergens favor interaction with lysine-rich human serum proteins (Hopkins et al. 2005). However, not all studies have observed the same patterns (Dik et al. 2016; Lalko et al. 2011; Lalko et al. 2012). The purpose of the investigations reported here was, therefore, to explore in a more systematic way whether there exist consistent differences between contact and respiratory allergens with respect to amino acid selectivity during hapten-protein formation. For this purpose it was necessary first to identify groups of chemicals, based upon literature review and expert judgement, that could be considered with confidence as being skin sensitizers only, respiratory sensitizers, or non-sensitizing chemicals. This was an important preliminary step because research on the characteristics of chemical respiratory allergens has been plagued by an inability to distinguish

between true respiratory sensitizers that act through immunological mechanisms, and other chemicals that cause asthma via other (non-immunological) mechanisms or agents causing anaphylaxis. Thus, one important objective was to distinguish true respiratory allergens from chemicals associated with non-allergic asthma using a weight of evidence approach.

For the purposes of this exercise classification as a true respiratory allergen required a sound clinical evidence of exposure resulting in symptoms in the respiratory tract (related to e.g., asthma, rhinitis, alveolitis) along with an immunological assessment (e.g., IgE levels by ELISA, skin prick test data). It must be acknowledged that using these criteria some generally recognized respiratory sensitizers may not have been included in the respiratory allergen category, either through lack of evidence or peptide reactivity data from DPRA studies, conducted according to OECD 442c.

However, for the purpose of determining whether respiratory allergens display clear amino acid preferences it was important to be certain of allergenic properties. For instance, chemicals such as azodicarbonamide and methyl methacrylate that are sometimes cited as being respiratory sensitizers were excluded from the respiratory allergen group based on a holistic review of the available data (Arts and Kimber 2018; Borak et al. 2011). For the skin sensitizers and non-sensitizer chemicals considered in this paper, the groupings were based on those proposed previously (Api et al. 2017; Basketter et al. 2014; Hoffmann et al. 2018). The supporting data were based on LLNA studies and on human data (human maximization tests, human repeated insult patch tests, and diagnostic patch tests). However, it must be appreciated that known chemical respiratory allergens are able to induce positive responses in the LLNA (Basketter et al. 2017), and that many (but not all) chemical respiratory allergens have some potential to cause ACD as well as sensitization of the respiratory tract (Kimber et al. 2011).

1
2
3 307 The important point is, however, that in addressing the objective of these investigations only
4
5 308 chemicals known to cause sensitization of the respiratory tract were included in the respiratory
6
7 309 allergen set. It is acknowledged that applying these criteria some true respiratory allergens may
8
9 310 have been excluded based on either a lack of standard DPRA data, and/or the absence of
10
11 311 convincing evidence to support this classification.
12
13
14 312 Data from DPRA studies, conducted according to OECD 442c, were used to map the reactivity
15
16 313 profiles of respiratory, skin and non-sensitizers towards cysteine- and lysine-containing peptides.
17
18 314 Cysteine and lysine reflect hard and soft nucleophiles, respectively, and are prone to react with
19
20 315 either hard or soft electrophiles (LoPachin et al. 2019; Mekenyan et al. 2014). The DPRA is the
21
22 316 most widely used peptide reactivity assay with a large and consistent dataset available. Although
23
24 317 the applicability domain of the DPRA can be extended by modifications to the protocol, for
25
26 318 instance to measure the peptide reactivity of inorganic compounds (e.g., platinum compounds)
27
28 319 (Hemming et al. 2019), results from modified protocols were not included here to ensure
29
30 320 consistency and comparability of DPRA data.
31
32
33
34 321 Respiratory sensitizers displayed a significantly (3-fold) higher reactivity with lysine than did skin
35
36 322 sensitizers. However, this difference was driven largely by the high representation among the
37
38 323 respiratory sensitizer set of acid anhydrides that displayed a clear preference for lysine. This is
39
40 324 consistent with the report of Dik et al. (2016) which found that amino acid binding preference is
41
42 325 chemical class specific, and not necessarily indicative of respiratory sensitization potential.
43
44
45 326 Anhydrides (other than chlorendic anhydride) showed high reactivity with lysine usually paired
46
47 327 with low cysteine reactivity (except in the case of maleic anhydride). This is a profile different
48
49 328 from that observed with diisocyanates that displayed stronger reactivity with cysteine. The
50
51 329 reactivity clusters of anhydrides and diisocyanates suggest that the former (shared with the skin
52
53
54
55
56
57
58
59
60

sensitizer bromohexane and two non-sensitizers) is less common for skin sensitizers than the diisocyanate-cluster, which is grouped together with 13 skin sensitizers.

Dik et al. (2016) have shown that anhydrides tend to react well with cysteine and lysine following even short exposure times, but interestingly lose reactivity towards cysteine after more prolonged (24 h) incubation. In general, anhydrides are readily hydrolyzed to their corresponding acids which are usually not protein reactive (Kreiling et al. 2017; Yamashita et al. 2015). Their hydrolysis could be pH-dependent and thus, lysine-reactivity may differ depending on the buffer system used, which has been shown recently for isocyanates and anhydrides in the Peroxidase Peptide Reactivity Assay (Ryan et al. 2020). In aqueous culture systems, anhydrides may undergo rapid hydrolysis prior to reaching cellular molecular targets which could lead to false-negative results in *in vitro* tests (Narita et al. 2017). Therefore, the instability of such anhydrides can affect the overall reactivity assessment. Nevertheless, the high lysine reactivity and low cysteine reactivity pattern of anhydrides is consistent with previous findings of other studies and seem to be less common for skin sensitizers. Phthalic anhydride exclusively reacts with lysine residues (Gerberick et al. 2007; Natsch 2010), and it is a respiratory allergen that is believed not to cause skin sensitization or ACD (Basketter and Kimber 2016).

The exceptions to the anhydride-reactivity pattern are maleic and chlorendic anhydride. Maleic anhydride hydrolyzes to acid similar to other anhydrides (Environment Agency Austria 2015) but shows high cysteine reactivity, and therefore may not lose reactivity towards cysteine after 24 h incubation in the DPRA. Chlorendic anhydride, which is also thought to hydrolyze to acid immediately (ECHA 2011), behaves slightly differently in the DPRA. It shows low lysine reactivity combined with low cysteine reactivity, albeit being reactive to cysteine (>72% depletion) after 20 min in the DPRA (Dik et al. 2016). Other aliphatic amines (skin sensitizers and non-

sensitizers) also show rather low peptide reactivity, except 2,5-diaminotoluene sulfate (a skin sensitizer).

Apart from diisocyanates and anhydrides, no other chemical category was found to have a distinct preferential amino acid reactivity profile. However, many (36%) of the skin sensitizers and non-sensitizers and one respiratory sensitizer (chloramine-T) could not be categorized according to US-EPA New Chemical Categories (structure categorization of chemical groups), which is a similar deficit as reported by Golden et al. (2020). Less discriminatory than chemical categories were protein binding mechanisms since none of those appeared to be associated with a preference lysine reactivity.

In conclusion, there is no clear discrimination between confirmed chemical respiratory allergens and skin sensitizing chemicals with respect to cysteine/lysine binding preferences. The interpretation is that if the pathways leading to skin and respiratory sensitization diverge at the stage of hapten-protein adduct formation the distinction is not as simple as a preferential association with cysteine or lysine. It may be that the kinetics of conjugate formation, and/or the promiscuity of protein interactions play more decisive roles.

Acknowledgements

This work was partly funded by The Federal Ministry of Education and Research (BMBF) grant 031L0128D.

Declaration of interests

This article was prepared during the normal course of the authors' affiliations or employment shown on the first page of the paper. None of these authors has participated in legal or regulatory

proceedings on the subject of this paper during the last 5 years. The authors have sole responsibility
for the preparation and content of this manuscript.

References

Api AM, Parakhia R, O'Brien D, Basketter DA. (2017). Fragrances Categorized According to Relative Human Skin Sensitization Potency. *Dermatitis*. **28**(5):299-307.

Arts J. (2020). How to assess respiratory sensitization of low molecular weight chemicals? *Int J Hyg Environ Health*. **225**:113469.

Arts J, Kimber I. (2018). Azodicarbonamide (ADCA): A reconsideration of classification as a respiratory sensitiser. *Regul Toxicol Pharmacol*. **94**:332-333.

Avonto C, Wang YH, Chittiboyina AG, Vukmanovic S, Khan IA. (2018). In chemico assessment of potential sensitizers: Stability and direct peptide reactivity of 24 fragrance ingredients. *J Appl Toxicol*.

Basketter D, Poole A, Kimber I. (2017). Behaviour of chemical respiratory allergens in novel predictive methods for skin sensitisation. *Regul Toxicol Pharmacol*. **86**:101-106.

Basketter DA, Alepee N, Ashikaga T, Barroso J, Gilmour N, Goebel C, Hibatallah J, Hoffmann S, Kern P, Martinozzi-Teissier S et al. (2014). Categorization of chemicals according to their relative human skin sensitizing potency. *Dermatitis*. **25**(1):11-21.

Basketter DA, Kimber I. (2016). Phthalic anhydride: Illustrating a conundrum in chemical allergy. *J Immunotoxicol*. **13**(6):767-769.

Borak J, Fields C, Andrews LS, Pemberton MA. (2011). Methyl methacrylate and respiratory sensitization: a critical review. *Crit Rev Toxicol*. **41**(3):230-268.

Dik S, Rorije E, Schwillens P, van Loveren H, Ezendam J. (2016). Can the Direct Peptide Reactivity Assay Be Used for the Identification of Respiratory Sensitization Potential of Chemicals? *Toxicol Sci*. **153**(2):361-371.

ECHA, European Chemicals Agency. (2011). Hydrolysis. CAS: 115-27-5. 1,4,5,6,7,7-hexachloro-8,9,10-trinorborn-5-ene-2,3-dicarboxylic anhydride (Chlorendic anhydride). <https://echa.europa.eu/de/registration-dossier/-/registered-dossier/2215/5/2/3/?documentUUID=3ddf62f3-023f-4761-ae01-ef3c6031dd58>.

Environment Agency Austria. 2015. CLH Report for Maleic anhydride. Proposal for Harmonised Classification and Labelling. Based on Regulation (EC) No 1272/2008 (CLP Regulation), Annex VI, Part 2. Version number 3.

Gerberick GF, Vassallo JD, Bailey RE, Chaney JG, Morrall SW, Lepoittevin JP. (2004). Development of a peptide reactivity assay for screening contact allergens. *Toxicol Sci*. **81**(2):332-343.

Gerberick GF, Vassallo JD, Foertsch LM, Price BB, Chaney JG, Lepoittevin JP. (2007). Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. *Toxicol Sci*. **97**(2):417-427.

Golden E, Maertens M, Hartung T, Maertens A. (2020). Mapping Chemical Respiratory Sensitization: How Useful Are Our Current Computational Tools? *Chem Res Toxicol*.

Hemming JDC, Hosford M, Shafer MM. (2019). Application of the direct peptide reactivity assay (DPRA) to inorganic compounds: a case study of platinum species. *Toxicology Research*. **8**(6):802-814.

Hettick JM, Siegel PD, Green BJ, Liu J, Wisnewski AV. (2012). Vapor conjugation of toluene diisocyanate to specific lysines of human albumin. *Anal Biochem*. **421**(2):706-711.

Hirota M, Ashikaga T, Kouzuki H. (2018). Development of an artificial neural network model for risk assessment of skin sensitization using human cell line activation test, direct peptide reactivity assay, KeratinoSens and in silico structure alert parameter. *J Appl Toxicol*. **38**(4):514-526.

- Hoffmann S, Kleinstreuer N, Alepee N, Allen D, Api AM, Ashikaga T, Clouet E, Cluzel M, Desprez B, Gellatly N et al. (2018). Non-animal methods to predict skin sensitization (I): the Cosmetics Europe database<sup/>. *Crit Rev Toxicol*. **48**(5):344-358.
- Holsapple MP, Jones D, Kawabata TT, Kimber I, Sarlo K, Selgrade MK, Shah J, Woolhiser MR. (2006). Assessing the potential to induce respiratory hypersensitivity. *Toxicol Sci*. **91**(1):4-13.
- Hopkins JE, Naisbitt DJ, Kitteringham NR, Dearman RJ, Kimber I, Park BK. (2005). Selective haptenation of cellular or extracellular protein by chemical allergens: association with cytokine polarization. *Chem Res Toxicol*. **18**(2):375-381.
- Isola D, Kimber I, Sarlo K, Lalko J, Sipes IG. (2008). Chemical respiratory allergy and occupational asthma: what are the key areas of uncertainty? *J Appl Toxicol*. **28**(3):249-253.
- Kimber I, Basketter DA, Dearman RJ. (2010). Chemical allergens - what are the issues? *Toxicology*. **268**(3):139-142.
- Kimber I, Basketter DA, Gerberick GF, Ryan CA, Dearman RJ. (2011). Chemical allergy: translating biology into hazard characterization. *Toxicol Sci*. **120** Suppl 1:S238-268.
- Kimber I, Dearman RJ, Basketter DA, Boverhof DR. (2014). Chemical respiratory allergy: reverse engineering an adverse outcome pathway. *Toxicology*. **318**:32-39.
- Kimber I, Poole A, Basketter DA. (2018). Skin and respiratory chemical allergy: confluence and divergence in a hybrid adverse outcome pathway. *Toxicol Res (Camb)*. **7**(4):586-605.
- Kreiling R, Gehrke H, Broschard TH, Dreessen B, Eigler D, Hart D, Hopflinger V, Kleber M, Kupny J, Li Q et al. (2017). In chemico, in vitro and in vivo comparison of the skin sensitizing potential of eight unsaturated and one saturated lipid compounds. *Regul Toxicol Pharmacol*. **90**:262-276.
- Lalko JF, Kimber I, Dearman RJ, Api AM, Gerberick GF. (2013). The selective peptide reactivity of chemical respiratory allergens under competitive and non-competitive conditions. *J Immunotoxicol*. **10**(3):292-301.
- Lalko JF, Kimber I, Dearman RJ, Gerberick GF, Sarlo K, Api AM. (2011). Chemical reactivity measurements: potential for characterization of respiratory chemical allergens. *Toxicol In Vitro*. **25**(2):433-445.
- Lalko JF, Kimber I, Gerberick GF, Foertsch LM, Api AM, Dearman RJ. (2012). The direct peptide reactivity assay: selectivity of chemical respiratory allergens. *Toxicol Sci*. **129**(2):421-431.
- LoPachin RM, Geohagen BC, Nordstroem LU. (2019). Mechanisms of soft and hard electrophile toxicities. *Toxicology*. **418**:62-69.
- Mekenyan O, Patlewicz G, Kuseva C, Popova I, Mehmed A, Kotov S, Zhechev T, Pavlov T, Temelkov S, Roberts DW. (2014). A mechanistic approach to modeling respiratory sensitization. *Chem Res Toxicol*. **27**(2):219-239.
- Narita K, Vo PTH, Yamamoto K, Kojima H, Itagaki H. (2017). Preventing false-negatives in the in vitro skin sensitization testing of acid anhydrides using interleukin-8 release assays. *Toxicol In Vitro*. **42**:69-75.
- Natsch A. (2010). The Nrf2-Keap1-ARE toxicity pathway as a cellular sensor for skin sensitizers--functional relevance and a hypothesis on innate reactions to skin sensitizers. *Toxicol Sci*. **113**(2):284-292.
- Natsch A, Ryan CA, Foertsch L, Emter R, Jaworska J, Gerberick F, Kern P. (2013). A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation. *J Appl Toxicol*. **33**(11):1337-1352.
- OECD, Organisation for Economic Co-operation and Development. (2010). Test No. 429: Skin Sensitisation - Local Lymph Node Assay. 23 July 2010. <https://doi.org/10.1787/9789264071100-en>.

OECD, Organisation for Economic Co-operation and Development. (2020a). The OECD QSAR Toolbox version 4.4. <https://qsartoolbox.org/>.

OECD, Organisation for Economic Co-operation and Development. (2020b). Test Guideline No. 442C: In Chemico Skin Sensitisation. Assays addressing the Adverse Outcome Pathway key event on covalent binding to proteins. Appendix I: In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA). <https://doi.org/10.1787/9789264229709-en>.

Ryan CA, Troutman JA, Kern PS, Quijano M, Dobson RLM, Jian Dai H, Burt TM, Gerberick GF. (2020). Refinement of the Peroxidase Peptide Reactivity Assay and Prediction Model for Assessing Skin Sensitization Potential. *Toxicol Sci.* **178**(1):88-103.

Sullivan KM, Enoch SJ, Ezendam J, Sewald K, Roggen EL, Cochrane S. (2017). An Adverse Outcome Pathway for Sensitization of the Respiratory Tract by Low-Molecular-Weight Chemicals: Building Evidence to Support the Utility of In Vitro and In Silico Methods in a Regulatory Context. *Applied In Vitro Toxicology.* **3**(3):213-226.

Urbisch D, Mehling A, Guth K, Ramirez T, Honarvar N, Kolle S, Landsiedel R, Jaworska J, Kern PS, Gerberick F et al. (2015). Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul Toxicol Pharmacol.* **71**(2):337-351.

US EPA, United States Environmental Protection Agency. 2010. United States Environmental Protection Agency (U.S. EPA) (2010) TSCA New Chemicals Program (NCP) Chemical Categories. Available: https://www.epa.gov/sites/production/files/2014-10/documents/ncp_chemical_categories_august_2010_version_0.pdf.

Wareing B, Urbisch D, Kolle SN, Honarvar N, Sauer UG, Mehling A, Landsiedel R. (2017). Prediction of skin sensitization potency sub-categories using peptide reactivity data. *Toxicol In Vitro.* **45**(Pt 1):134-145.

Yamashita K, Shinoda S, Hagiwara S, Miyazaki H, Itagaki H. (2015). Unsaturated fatty acids show clear elicitation responses in a modified local lymph node assay with an elicitation phase, and test positive in the direct peptide reactivity assay. *J Toxicol Sci.* **40**(6):843-853.

Figures

Figure 1: Overview of lysine and cysteine and average of lysine and cysteine peptide reactivity as % depletion of respiratory (n=15), skin (n=129, solid dots are outliers) and non-sensitizers (n=56, open circles are outliers) depicted as boxplots. Asterisks (*) and horizontal bars indicate intergroup differences between respiratory and skin sensitizers by Student's t-test (ns = $p > 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$). Results and graph are generated by JMP® Pro 15.

Figure 2: Biplot of principal component analysis of cysteine and lysine peptide reactivity of all reactive chemicals. Chemicals are clustered by k-Means Cluster (n=8) for cysteine (PC1) and lysine (PC2) peptide depletion (%). Data points represent arithmetic means of non-redundant DPRA cysteine and lysine depletion values of 200 chemicals, including 15 respiratory sensitizers (asterisks and label), 120 skin sensitizers (solid dots) and 56 non-sensitizers (open circles). Graph is generated by JMP® Pro 15.

Figure 3: Peptide reactivity of chemicals by chemical categories (US-EPA) and groups of respiratory (n=14, asterisks), skin (n=84, solid dots) and non-sensitizers (n=30, open circles). OECD QSAR toolbox version 4.4 categorized 128 chemicals according to US-EPA. Results are shown as boxplots by JMP® Pro 15.

Figure 4: Peptide reactivity of chemicals according to their protein binding mechanisms by OECD and OASIS and groups of respiratory (n=13), skin (n=84, solid dots) and non-sensitizers (n=6). OECD QSAR toolbox version 4.4 identified protein binding mechanisms for 103 chemicals. Results are shown as boxplots by JMP® Pro 15.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

524 Figure 5: Distribution of chemical categories per cluster. Chemical categories are according to US-
525 EPA and cluster are defined by principal component analysis (Figure 2). Shown are the number of
526 chemicals (n=200) per chemical category and cluster and asterisks represents the presence of at
527 least on respiratory sensitizer per chemical category and cluster. Graph is generated with Excel.

529 Tables

530 Table 1: List of chemicals with available DPRA data that are grouped as respiratory sensitizers
 531 (n=15) based on expert judgment using a weight-of-evidence approach.

Label	Test chemical	CAS#
10	1,2,4-Benzene-tricarboxylic anhydride	552-30-7
12	1,2-cyclohexane dicarboxylic anhydride	85-42-7
16	1,6-hexamethylene diisocyanate	822-06-0
25	2,4,6-Trichloro-1,3,5-triazine	108-77-0
28	2,4-Toluene diisocyanate	584-84-9
30	2,6-Toluene diisocyanate	91-08-7
57	4,4'-Methylenediphenyl diisocyanate	101-68-8
98	Chloramine-T	127-65-1
99	Chlorendic anhydride	115-27-5
135	Glutaraldehyde	111-30-8
143	Isophorone diisocyanate	4098-71-9
150	Maleic anhydride	108-31-6
174	Phthalic Anhydride	85-44-9
175	Piperazine	110-85-0
190	Tetrachlorophthalic anhydride	117-08-8

532
 533 Table 2: Lysine and cysteine peptide reactivity as % depletion of the protein-reactive chemicals
 534 according to the 8 clusters (Figure 2). Results are shown as arithmetic means and standard
 535 deviation (STD) of all analyzed chemicals (n=200) and are generated by JMP® Pro 15.

Cluster	N	Cysteine reactivity (% depletion)		Lysine reactivity (% depletion)	
		Mean	STD	Mean	STD
1	8	73	±15	58	±9
2	29	44	±13	9	±6
3	8	12	±8	73	±6
4	90	4	±5	3	±4
5	18	95	±5	31	±8
6	10	95	±9	87	±6
7	11	9	±8	29	±9
8	26	95	±6	5	±6
Sum	200				

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supplement

Suppl. 1: List of analyzed chemicals (n=200) with according grouping into respiratory (n=15), skin (n=129) and non-sensitizers (n=56), their cysteine and lysine depletion (%) data and references, protein binding mechanisms (OECD and OASIS), chemical categories (US-EPA) and cluster information. DPRA results are shown as arithmetic mean for each CAS# and are based on non-redundant DPRA values.



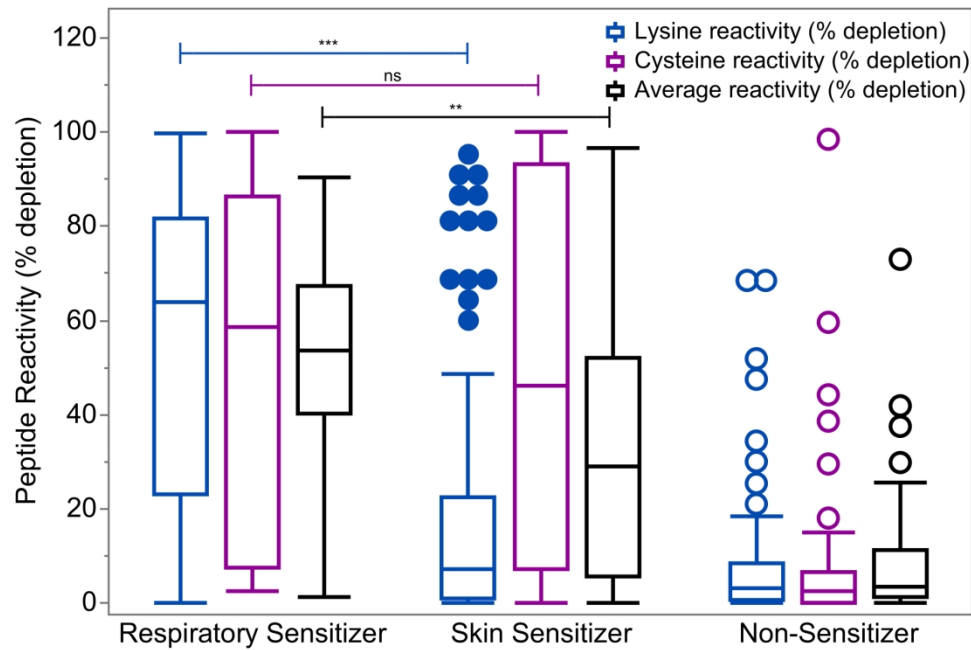


Figure 1: Overview of lysine and cysteine and average of lysine and cysteine peptide reactivity as % depletion of respiratory (n=15), skin (n=129, solid dots are outliers) and non-sensitizers (n=56, open circles are outliers) depicted as boxplots. Asterisks (*) and horizontal bars indicate intergroup differences between respiratory and skin sensitizers by Student's t-test (ns = $p > 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$). Results and graph are generated by JMP® Pro 15.

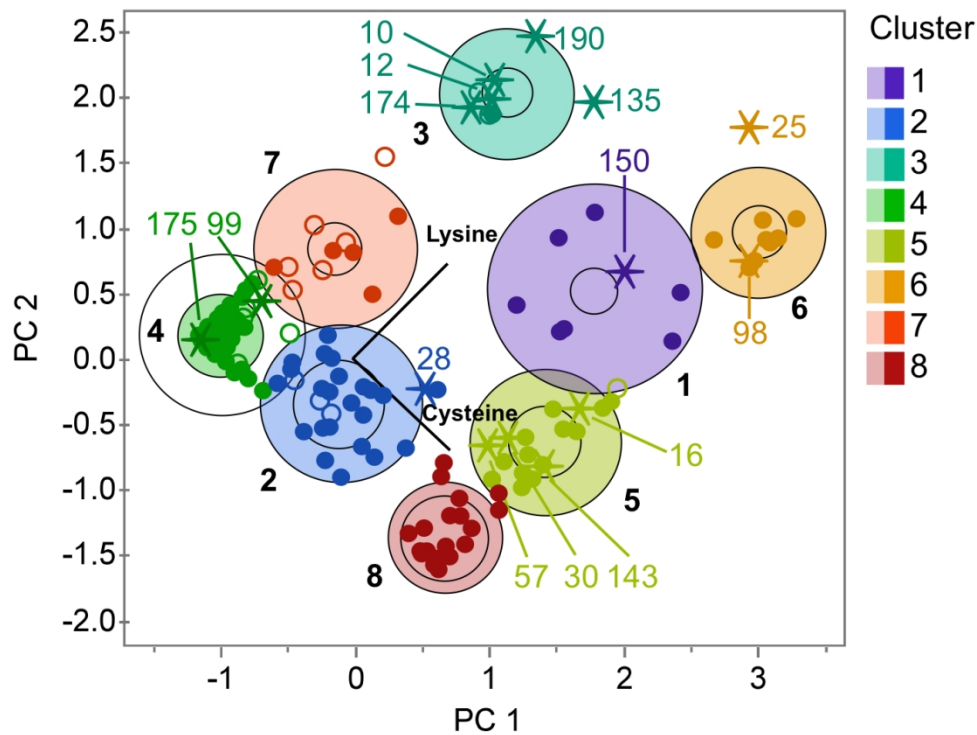


Figure 2: Biplot of principal component analysis of cysteine and lysine peptide reactivity of all reactive chemicals. Chemicals are clustered by k-Means Cluster (n=8) for cysteine (PC1) and lysine (PC2) peptide depletion (%). Data points represent arithmetic means of non-redundant DPRA cysteine and lysine depletion values of 200 chemicals, including 15 respiratory sensitizers (asterisks and label), 120 skin sensitizers (solid dots) and 56 non-sensitizers (open circles). Graph is generated by JMP® Pro 15.

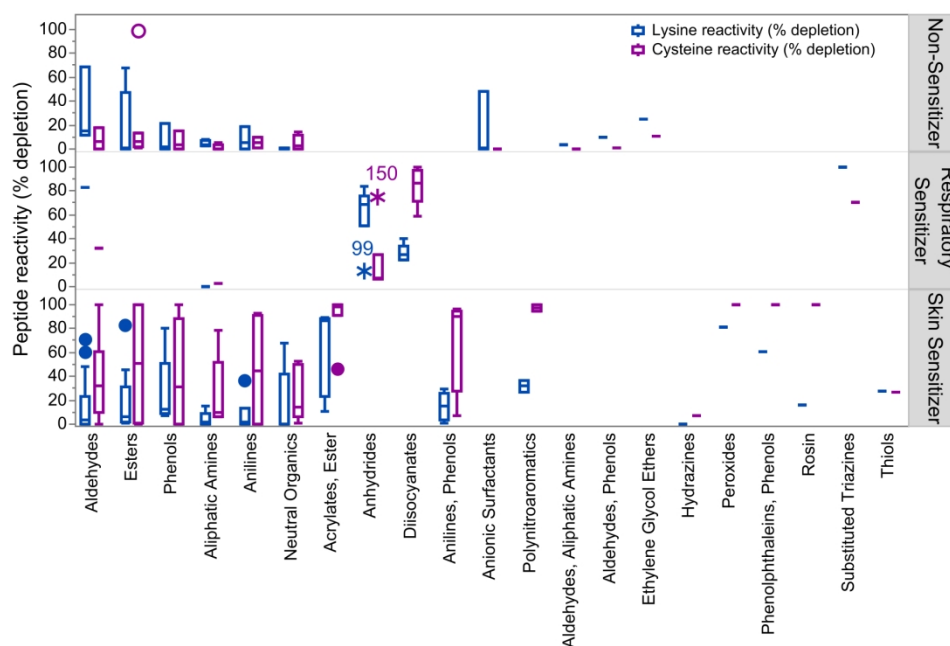


Figure 3: Peptide reactivity of chemicals by chemical categories (US-EPA) and groups of respiratory (n=14, asterisks), skin (n=84, solid dots) and non-sensitizers (n=30, open circles). OECD QSAR toolbox version 4.4 categorized 128 chemicals according to US-EPA. Results are shown as boxplots by JMP® Pro 15.

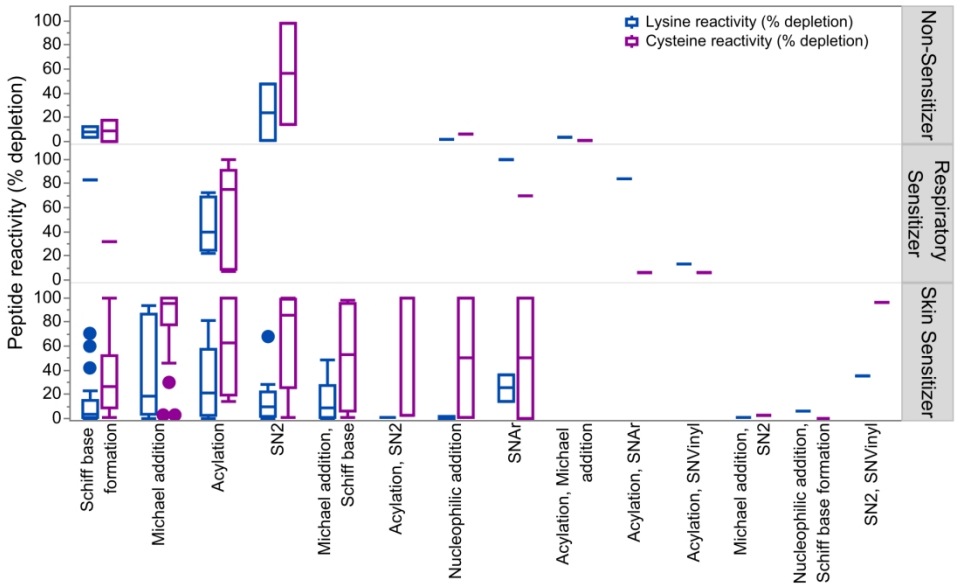


Figure 4: Peptide reactivity of chemicals according to their protein binding mechanisms by OECD and OASIS and groups of respiratory (n=13), skin (n=84, solid dots) and non-sensitizers (n=6). OECD QSAR toolbox version 4.4 identified protein binding mechanisms for 103 chemicals. Results are shown as boxplots by JMP® Pro 15.

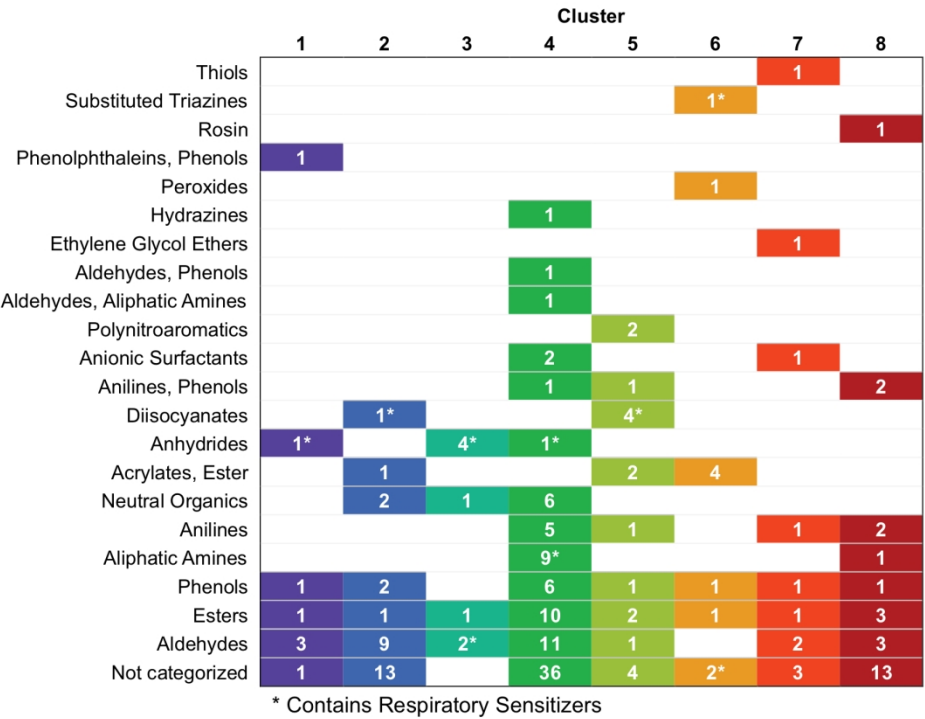


Figure 5: Distribution of chemical categories per cluster. Chemical categories are according to US-EPA and cluster are defined by principal component analysis (Figure 2). Shown are the number of chemicals (n=200) per chemical category and cluster and asterisks represents the presence of at least on respiratory sensitizer per chemical category and cluster. Graph is generated with Excel.