Computational Characterisation of Chemicals and Datasets in Terms of Organic Functional Groups - a New Toxtree Rulebase

Romualdo Benigni\textsuperscript{a}, Olga Tcheremenskaia\textsuperscript{a} and Andrew Worth\textsuperscript{b}

\textsuperscript{a} Istituto Superiore di Sanità, Environment and Health Department, Rome, Italy

\textsuperscript{b} European Commission - Joint Research Centre, Institute for Health & Consumer Protection, Systems Toxicology Unit, Ispra, Italy
The mission of the JRC-IHCP is to protect the interests and health of the consumer in the framework of EU legislation on chemicals, food, and consumer products by providing scientific and technical support including risk-benefit assessment and analysis of traceability.
ABSTRACT

Toxtree is a freely available, user-friendly and extensible software application that is designed to make structure-based predictions for a number of toxicological endpoints and mechanisms of chemical action. The platform has been developed by the Joint Research Centre in collaboration with Ideaconsult Ltd (Sofia, Bulgaria) with a range of modules developed by various contributors. One of the modules developed as an extension to Toxtree is aimed at the identification of organic functional groups in query chemicals. The rulebase consists of 204 organic functional groups recognised by the “Checkmol” program, which was developed by Dr Norbert Haider, University of Vienna. A new Functional Group Profiler, has been coded as a Toxtree module by the Istituto Superiore di Sanità’ (Rome, Italy). The Toxtree profiler, called ISSFUNC, can be used to screen and characterise chemicals as a basis for read-across, category formation and (Q)SAR analysis. It can also be used for the global comparison of datasets, such as model training and test sets and chemical inventories.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISS</td>
<td>Istituto Superiore di Sanità</td>
</tr>
<tr>
<td>ISSCAN</td>
<td>Istituto Superiore di Sanità database on chemical carcinogens</td>
</tr>
<tr>
<td>EPA</td>
<td>Environmental Protection Agency (USA)</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>JRC</td>
<td>Joint Research Centre</td>
</tr>
<tr>
<td>QSAR</td>
<td>Quantitative Structure-Activity Relationships</td>
</tr>
<tr>
<td>REACH</td>
<td>Registration, Evaluation, Authorisation and Restriction of Chemicals</td>
</tr>
</tbody>
</table>
## CONTENTS

1. Introduction .......................................................................................................................... 4
2. The Functional Groups........................................................................................................... 5
3. Use cases................................................................................................................................ 5
   3.1 Basic applications ............................................................................................................. 5
   3.2 Comparing chemical databases ....................................................................................... 6
4. Summary and Conclusions .................................................................................................... 9
5. Acknowledgements and Disclaimer .................................................................................... 9
6. References .............................................................................................................................. 10
1. Introduction

In the context of the recent developments in chemicals regulations and regulatory needs worldwide, the progress in chemoinformatics technology is particularly timely in providing an essential tool to support the chemical assessment process. Until now, the assessment of chemical risks in the European Union (EU) has been largely based on traditional toxicology. However, legislative, societal and practical realities (too many chemicals, too few resources) have created new inducements and opportunities to encourage the use and acceptance of “alternative” approaches, which can reduce substantially the need for experimental toxicological testing.

In 2003, the European Commission (EC) adopted a legislative proposal for a new chemical assessment and management system called REACH (Registration, Evaluation and Authorisation of CHemicals). Article 13(1) of the REACH regulation (EC, 2006) states that:

“Information on intrinsic properties of substances may be generated by means other than tests, provided that the conditions set out in Annex XI are met. In particular for human toxicity, information shall be generated whenever possible by means other than vertebrate animal tests, through the use of alternative methods, for example, in vitro methods or qualitative or quantitative structure-activity relationship models or from information from structurally related substances (grouping or read-across).”

REACH has introduced a dramatic change in the EU regulatory framework - it explicitly provides the basis for the use of structure-activity relationship models, together with other “non-testing” approaches, for predicting the environmental and toxicological properties of chemicals, in the interests of time-effectiveness, cost-effectiveness and animal welfare. This change is increasingly being reflected in other pieces of EU legislation (Worth, 2010).

The science that aims to understand the relationships between chemical structure and the biological activity of molecules is evolving to support three distinct activities: category formation, read-across, and (Quantitative) Structure-Activity Relationship ((Q)SAR) analysis. A chemical category is a group of chemicals whose physicochemical and human health and/or environmental toxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity. If this similarity is recognised with sufficient evidence, all the chemicals in the category can be assessed (and regulated) in the same way. Another approach to fill data gaps is read-across. In the read-across approach, endpoint information for one chemical is used to predict the same endpoint for another chemical, which is considered to be “similar” in some way (usually on the basis of structural similarity, but increasingly also on the basis of mechanistic or biological similarity).

Regarding the third approach, the scientific foundation of (Q)SAR models lies in physical organic chemistry, where chemical behaviour and activity are estimated solely from a knowledge of chemical structure. (Q)SAR modeling has been widely used in pharmacology, toxicology and physical chemistry, and its capabilities and limitations are relatively well understood (Hansch & Leo, 1995; Franke & Gruska, 2003; Worth et al., 2007).

The extensive use of estimation techniques such as (Q)SARs, read-across and grouping of chemicals, where appropriate and in a suitably constrained context, has the potential to effect reductions in the use of animals for toxicity assessment. At the same time, all these approaches need to be supported by adequate technological tools. Fortunately, the recent years have witnessed a dramatic progress in the field of manipulation of the chemical structure with computers, ranging from chemical relational databases, to calculation of chemical descriptors and derivation of qualitative and quantitative structure-activity relationships (Chen, 2006; Muchmore et al., 2010).
Among the software tools specifically aimed at supporting the (Q)SAR analyses in the regulatory assessment of chemicals is the expert system Toxtree. Toxtree (http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE) is an open-source, freely available software application that places chemicals into categories and predicts various kinds of toxic effects by applying various decision tree approaches. All estimation methods are structurally-based. The Toxtree platform was developed by Ideaconsult Ltd. (Sofia, Bulgaria) under the terms of a JRC contract (Worth et al., 2007). A range of modules (plug-ins) have also been developed by various contributors. In its present version (2.1.0; June 2011), it contains modules for estimating: a) oral toxicity (Cramer scheme); b) aquatic modes of action (Verhaar scheme); c) skin and eye irritation and corrosion; d) carcinogenicity and in vitro and in vivo mutagenicity; e) ability to act as Michael acceptors; f) persistence / biodegradation potential (START rulebase); and sites of cytochrome P450-mediated oxidation.

All the above rulebases can be used independently to estimate endpoints or properties. Another use is to consider the various rulebases in combination for a wider profiling of the chemicals. In this report, we describe the implementation of a new rulebase for Toxtree, aimed at characterising chemicals in terms of organic functional groups. The rulebase consists of 204 organic functional groups recognized by the “Checkmol” program (http://merian.pch.univie.ac.at/~nhaider/cheminf/cmmm.html) which was developed by Dr Norbert Haider (University of Vienna, Austria).

2. The Functional Groups

The new rulebase identifies the classical organic functional groups (such as carbonyl, nitro or many others) present in the molecules, thus providing a tool to categorise and characterise the “chemical entities” under study (Feldman et al., 2005). The structural features are listed in Appendix 1.

In order to allow the users to discriminate between chemicals with higher and lower structure similarity, the functional groups have been divided into structural features with high specificity (HS) or low specificity (LS), nested hierarchically. For example, the low specificity feature FG75_LS (sulfonic acid derivatives) includes a broad range of compounds and has been further divided into a number of high specificity features: FG75_1_HS (sulfonic acid), FG75_2_HS (sulfonic acid ester), FG75_3_HS (sulfonamide). The high specificity features collect smaller sets of more closely related chemicals. A chemical can have simultaneously LS and HS features. By considering this information, the user can identify sub-categories of more mutually similar compounds and use these (qualitative) results for further analysis.

3. Use cases

3.1 Basic applications

The most immediate use of the Functional Group rulebase is the identification of similar chemicals. This forms the basis of the read-across procedure, which looks for a few chemicals with characteristics similar to those of a query chemical, whose missing data are to be extrapolated / interpolated from those of the “similar” chemicals. In a similar way, the category approach looks for chemicals to be grouped and assessed together. Another use of the Functional Group rulebase is to identify sets of congeneric chemicals that can be analysed together with a QSAR approach, such as the Hansch or extra-thermodynamic approach.

Whereas the Functional Group rulebase may provide the basic directions in finding similar chemicals, the simultaneous presence in Toxtree of a number of other estimation methods (e.g. Cramer, Verhaar, etc) permits a more sophisticated approach to such similarity searching for predictive toxicology. After a first categorisation through the Functional Group rulebase, it is possible to further sub-categorise by applying one or more of the other rulebases. This should be performed in a stepwise manner, where the functional groups provide the first clustering, and the hazard-based rulebase(s) refine the chemical
category by subdividing it into smaller clusters of chemicals with both chemical and toxicological similarity.

3.2 Comparing chemical databases

The task of comparing databases of chemicals is of utmost importance. In particular, it provides a means of putting into context and rationalising the performance of a toxicological assay or QSAR when applied to different sets of chemicals. Without anchoring performance statistics to a definition or description of the tested chemicals, it is not possible to give a rational explanation as to why a certain assay or model appears to have a different predictive performance when applied to different sets of chemicals. The chemical space of a QSAR model training set forms the basis of the model applicability domain (Netzeva et al., 2005).

As an example, we show the use of the Functional Group rulebase to compare two databases: a) the classical database of chemicals tested in the rodent carcinogenicity bioassay; and b) a database studied recently within the US EPA’s ToxCast Phase I exercise (Martin et al., 2010; Benigni et al., 2010).

The classical experimental carcinogenicity database is ISSCAN v3a (1141 unique chemicals) (Benigni et al., 2008). ISSCAN is available from the ISS website (http://www.iss.it/ampp/dati/cont.php?id=233&lang=1&tipo=7). The ToxCast data (309 unique chemicals) are contained in a dataset called: “ToxCast Phase I Data (AC50/LEC), downloadable from the ToxCast website (http://www.epa.gov/ncct/toxcast/data_sets.html).

The two sets of chemicals were combined together into an sdf file (with the exclusion of 46 chemicals common to the two databases), and the Functional Group rulebase was applied. A matrix, where each chemical was defined by the presence of the various functional groups, was obtained. The next step was the calculation of a distance matrix among all chemicals, based on the functional group profile. The Jaccard metric for similarity was used (Jaccard, 1912). The \( n \times n \) distance matrix was then reduced to 10 Principal Components (PC) (explained variance: 0.81). The last step was the application of Canonical Discriminant Analysis to the PCs, with the aim to separate Toxcast from the ISSCAN chemicals.

Figure 1 displays the distribution of the ISSCAN and ToxCast chemicals along the direction of maximum dissimilarity between the two databases (Canonical Component 1, Squared Canonical Correlation = 0.21). It appears that the functional group composition of the two databases is largely overlapping, with exceptions at the extremes of the X axis.
In particular, ISSCAN contains a large subset of simple aromatic amines that are not present in ToxCast (low values of the Canonical Component, representative examples in Figure 2a), whereas ToxCast contains a subset of complex structures only present in this database (high values of the Canonical Component, representative examples in Figure 2b).
It is interesting to link this observation with observations on the prevalent toxicity mechanisms of the two databases of chemicals. In fact, a recent analysis has estimated that in ISSCAN v3a around 70% of carcinogens have Structural Alerts for genotoxic carcinogenicity, whereas in ToxCast only 35% of carcinogens are thought to act through genotoxic carcinogenicity mechanisms (Benigni et al., 2010). As a matter of fact, the majority of aromatic amines present in ISSCAN, but not in ToxCast (Figure 2a), act by genotoxic mechanisms (Benigni & Bossa, 2011). This shows that the analysis of the functional group distribution in a set of chemicals can provide a powerful means of comparing datasets in terms of their mechanistic toxicology.
4. Summary and Conclusions

In this report, we have presented a new rulebase for Toxtree which provides a means of identifying a set of 204 organic functional groups in query chemicals. The organic functional groups profiler can be used to screen chemical databases and identify “similar” chemicals for the purposes of read-across, category formation and QSAR analysis. It can also be used for the global comparison of chemicals databases / inventories, as illustrated in this study with two chemical databases – ISSCAN and ToxCast. It is anticipated that the organic functional groups profiler will provide a useful means of categorizing chemicals, especially when used in combination with other hazard-based profilers within Toxtree.

5. Acknowledgements and Disclaimer

Any conclusions and opinions expressed in this document are those of the authors as individual scientists and do not constitute an official position by the JRC or the European Commission.
6. References


### Appendix 1 Structural features incorporated into the Organic Functional Groups Profiler

<table>
<thead>
<tr>
<th>No</th>
<th>Alert ID</th>
<th>Alert Title</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FG1</td>
<td>cation</td>
<td>any positive charge</td>
</tr>
<tr>
<td>2</td>
<td>FG2</td>
<td>anion</td>
<td>any negative charge</td>
</tr>
<tr>
<td>3</td>
<td>FG3_LS</td>
<td>carbonyl compound: aldehyde or ketone</td>
<td><img src="" alt="structure" /> $R^1 = H, alkyl, aryl$ $R^2 = H, alkyl, aryl$</td>
</tr>
<tr>
<td>4</td>
<td>FG3_1_HS</td>
<td>aldehyde</td>
<td><img src="" alt="structure" /> $R = H, alkyl, aryl$</td>
</tr>
<tr>
<td>5</td>
<td>FG3_2_HS</td>
<td>ketone</td>
<td><img src="" alt="structure" /> $R^1 = alkyl, aryl$ $R^2 = alkyl, aryl$</td>
</tr>
<tr>
<td>6</td>
<td>FG4_LS</td>
<td>thiocarbonyl compound: aldehyde or ketone</td>
<td><img src="" alt="structure" /> $R^1 = H, alkyl, aryl$ $R^2 = H, alkyl, aryl$</td>
</tr>
<tr>
<td>7</td>
<td>FG4_1_HS</td>
<td>thioaldehyde</td>
<td><img src="" alt="structure" /> $R = H, alkyl, aryl$</td>
</tr>
<tr>
<td>8</td>
<td>FG4_2_HS</td>
<td>thioketone</td>
<td><img src="" alt="structure" /> $R^1 = alkyl, aryl$ $R^2 = alkyl, aryl$</td>
</tr>
<tr>
<td>9</td>
<td>FG5</td>
<td>imine</td>
<td><img src="" alt="structure" /> $R^1 = H, alkyl, aryl$ $R^2 = H, alkyl, aryl$ $R^3 = H, alkyl, aryl$</td>
</tr>
<tr>
<td>10</td>
<td>FG6</td>
<td>hydrazone</td>
<td><img src="" alt="structure" /> $R^1 = H, alkyl, aryl$ $R^2 = H, alkyl, aryl$ $R^3 = H, alkyl, aryl$ $R^4 = H, alkyl, aryl$</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>----</td>
<td>----------</td>
<td>-------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>11</td>
<td>FG7</td>
<td>semicarbazone</td>
<td><img src="image" alt="Structure for semicarbazone" /> $R^1 = H, \text{alkyl, aryl}$ $R^2 = H, \text{alkyl, aryl}$ $R^3 = H, \text{alkyl, aryl}$ $R^4 = H, \text{alkyl, aryl}$ $R^5 = H, \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>12</td>
<td>FG8</td>
<td>thiosemicarbazone</td>
<td><img src="image" alt="Structure for thiosemicarbazone" /> $R^1 = H, \text{alkyl, aryl}$ $R^2 = H, \text{alkyl, aryl}$ $R^3 = H, \text{alkyl, aryl}$ $R^4 = H, \text{alkyl, aryl}$ $R^5 = H, \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>13</td>
<td>FG9</td>
<td>oxime</td>
<td><img src="image" alt="Structure for oxime" /> $R^1 = H, \text{alkyl, aryl}$ $R^2 = H, \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>14</td>
<td>FG10</td>
<td>oxime ether</td>
<td><img src="image" alt="Structure for oxime ether" /> $R^1 = H, \text{alkyl, aryl}$ $R^2 = H, \text{alkyl, aryl}$ $R^3 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>15</td>
<td>FG11</td>
<td>ketene</td>
<td><img src="image" alt="Structure for ketene" /> $R^1 = H, \text{alkyl, aryl}$ $R^2 = H, \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>16</td>
<td>FG12</td>
<td>ketene acetal derivative</td>
<td><img src="image" alt="Structure for ketene acetal derivative" /> $R^1 = H, \text{alkyl, aryl}$ $R^2 = H, \text{alkyl, aryl}$ $X = \text{any hetero atom}$ $Y = \text{any hetero atom}$</td>
</tr>
<tr>
<td>17</td>
<td>FG13</td>
<td>carbonyl hydrate</td>
<td><img src="image" alt="Structure for carbonyl hydrate" /> $R^1 = H, \text{alkyl, aryl}$ $R^2 = H, \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>18</td>
<td>FG14</td>
<td>hemiacetal</td>
<td><img src="image" alt="Structure for hemiacetal" /> $R^1 = H, \text{alkyl, aryl}$ $R^2 = H, \text{alkyl, aryl}$ $R^3 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>19</td>
<td>FG15</td>
<td>acetal</td>
<td><img src="image" alt="Structure for acetal" /> $R^1 = H, \text{alkyl, aryl}$ $R^2 = H, \text{alkyl, aryl}$ $R^3 = \text{alkyl, aryl}$ $R^4 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>----</td>
<td>----------</td>
<td>--------------</td>
<td>-----------</td>
</tr>
</tbody>
</table>
| 20 | FG16     | hemiaminal   | ![Structure](image) R¹ = H, alkyl, aryl  
                                R² = H, alkyl, aryl  
                                R³ = H, alkyl, aryl  
                                R⁴ = H, alkyl, aryl  
                                R⁵ = H, alkyl, aryl |
| 21 | FG17     | aminal       | ![Structure](image) R¹ = H, alkyl, aryl  
                                R² = H, alkyl, aryl  
                                R³ = H, alkyl, aryl  
                                R⁴ = H, alkyl, aryl  
                                R⁵ = H, alkyl, aryl |
| 22 | FG18     | thiohemiaminal | ![Structure](image) R¹ = H, alkyl, aryl  
                                R² = H, alkyl, aryl  
                                R³ = H, alkyl, aryl  
                                R⁴ = H, alkyl, aryl  
                                R⁵ = H, alkyl, aryl |
| 23 | FG19     | thioacetal   | ![Structure](image) R¹ = H, alkyl, aryl  
                                R² = H, alkyl, aryl  
                                R³ = alkyl, aryl    
                                R⁴ = alkyl, aryl    |
| 24 | FG20     | enamine      | ![Structure](image) R¹ = H, acyl, alkyl, aryl  
                                R² = H, acyl, alkyl, aryl  
                                R³ = H, acyl, alkyl, aryl  
                                R⁴ = H, acyl, alkyl, aryl  
                                R⁵ = H, acyl, alkyl, aryl |
| 25 | FG21     | enol         | ![Structure](image) R¹ = H, acyl, alkyl, aryl  
                                R² = H, acyl, alkyl, aryl  
                                R³ = H, acyl, alkyl, aryl |
| 26 | FG22     | enolether    | ![Structure](image) R¹ = H, acyl, alkyl, aryl  
                                R² = H, acyl, alkyl, aryl  
                                R³ = H, acyl, alkyl, aryl  
                                R⁴ = alkyl, aryl    |
<p>| 27 | FG23_LS  | alcohol      | <img src="image" alt="Structure" /> R = alkyl, aryl    |
| 28 | FG23_1_HS| primary alcohol | <img src="image" alt="Structure" /> R = alkyl, aryl    |</p>
<table>
<thead>
<tr>
<th>No</th>
<th>Alert ID</th>
<th>Alert Title</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>FG23_2_HS</td>
<td>secondary alcohol</td>
<td>R¹ = alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = alkyl, aryl</td>
</tr>
<tr>
<td>30</td>
<td>FG23_3_HS</td>
<td>tertiary alcohol</td>
<td>R¹ = alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R³ = alkyl, aryl</td>
</tr>
<tr>
<td>31</td>
<td>FG23_4_HS</td>
<td>1,2-diol</td>
<td>R¹ = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R³ = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R⁴ = H, alkyl, aryl</td>
</tr>
<tr>
<td>32</td>
<td>FG23_5_HS</td>
<td>1,2-aminoalcohol</td>
<td>R¹ = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R³ = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R⁴ = H, alkyl, aryl</td>
</tr>
<tr>
<td>33</td>
<td>FG23_6_HS</td>
<td>phenol</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>FG23_7_HS</td>
<td>1,2-diphenol</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>FG23_8_HS</td>
<td>enediol</td>
<td>R¹ = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = H, alkyl, aryl</td>
</tr>
<tr>
<td>36</td>
<td>FG24_LS</td>
<td>ether</td>
<td>R¹ = alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = alkyl, aryl</td>
</tr>
<tr>
<td>37</td>
<td>FG24_1_HS</td>
<td>dialkylether</td>
<td>R¹ = alkyl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = alkyl</td>
</tr>
<tr>
<td>38</td>
<td>FG24_2_HS</td>
<td>alkylarylether</td>
<td>R¹ = alkyl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = aryl</td>
</tr>
<tr>
<td>39</td>
<td>FG24_3_HS</td>
<td>diarylether</td>
<td>R¹ = aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = aryl</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>----</td>
<td>----------</td>
<td>---------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>40</td>
<td>FG25</td>
<td>thioether</td>
<td>$R^1 \text{--} S \text{--} R^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^1 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^2 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>41</td>
<td>FG26</td>
<td>disulfide</td>
<td>$R^1 \text{--} S \text{--} S \text{--} R^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^1 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^2 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>42</td>
<td>FG27</td>
<td>peroxide</td>
<td>$R^1 \text{--} O \text{--} O \text{--} R^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^1 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^2 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>43</td>
<td>FG28</td>
<td>hydroperoxide</td>
<td>$R \text{--} O \text{--} OH$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>44</td>
<td>FG29</td>
<td>hydrazine derivative</td>
<td>$R^1 \text{--} H \text{--} N \text{--} N \text{--} R^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^1 = \text{H, acyl, alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^2 = \text{H, acyl, alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^3 = \text{H, acyl, alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^4 = \text{H, acyl, alkyl, aryl}$</td>
</tr>
<tr>
<td>45</td>
<td>FG30</td>
<td>hydroxylamine</td>
<td>$R \text{--} N \text{--} O \text{--} R^3$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^1 = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^2 = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^3 = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td>46</td>
<td>FG31_LS</td>
<td>amine</td>
<td>$R^3 \text{--} N \text{--} R^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^1 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^2 = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^3 = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td>47</td>
<td>FG31_1_HS</td>
<td>primary aliphatic amine</td>
<td>$R\text{--}NH_2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R = \text{alkyl}$</td>
</tr>
<tr>
<td>48</td>
<td>FG31_2_HS</td>
<td>primary aromatic amine</td>
<td>$R\text{--}NH_2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R = \text{aryl}$</td>
</tr>
<tr>
<td>49</td>
<td>FG31_3_HS</td>
<td>secondary aliphatic amine</td>
<td>$R^1 \text{--} N \text{--} R^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^1 = \text{alkyl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^2 = \text{alkyl}$</td>
</tr>
<tr>
<td>50</td>
<td>FG31_4_HS</td>
<td>secondary mixed amine (aryl alkyl)</td>
<td>$R^1 \text{--} N \text{--} R^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^1 = \text{alkyl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$R^2 = \text{aryl}$</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>----</td>
<td>----------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>51</td>
<td>FG31_5_HS</td>
<td>secondary aromatic amine</td>
<td>![Structure Image] (R¹ = aryl, R² = aryl)</td>
</tr>
<tr>
<td>52</td>
<td>FG31_6_HS</td>
<td>tertiary aliphatic amine</td>
<td>![Structure Image] (R¹ = alkyl, R² = alkyl, R³ = alkyl)</td>
</tr>
<tr>
<td>53</td>
<td>FG31_7_HS</td>
<td>tertiary mixed amine</td>
<td>![Structure Image] (R¹ = alkyl, R² = aryl, R³ = alkyl)</td>
</tr>
<tr>
<td>54</td>
<td>FG31_8_HS</td>
<td>tertiary aromatic amine</td>
<td>![Structure Image] (R¹ = alkyl, aryl, R² = alkyl, aryl, R³ = alkyl, aryl)</td>
</tr>
<tr>
<td>55</td>
<td>FG31_9_HS</td>
<td>quaternary ammonium salt</td>
<td>![Structure Image] (R¹ = alkyl, aryl, R² = alkyl, aryl, R³ = alkyl, aryl, R⁴ = alkyl, aryl)</td>
</tr>
<tr>
<td>56</td>
<td>FG32</td>
<td>N-oxide</td>
<td>![Structure Image] (R¹ = alkyl, aryl, R² = alkyl, aryl, R³ = alkyl, aryl)</td>
</tr>
<tr>
<td>57</td>
<td>FG33_LS</td>
<td>halogen derivative (alkyl or aryl)</td>
<td>![Structure Image] (R—X, X = F, Cl, Br, I, R = alkyl, aryl)</td>
</tr>
<tr>
<td>58</td>
<td>FG33_1_HS</td>
<td>alkyl fluoride</td>
<td>![Structure Image] (R—F, R = alkyl)</td>
</tr>
<tr>
<td>59</td>
<td>FG33_2_HS</td>
<td>alkyl chloride</td>
<td>![Structure Image] (R—Cl, R = alkyl)</td>
</tr>
<tr>
<td>60</td>
<td>FG33_3_HS</td>
<td>alkyl bromide</td>
<td>![Structure Image] (R—Br, R = alkyl)</td>
</tr>
<tr>
<td>61</td>
<td>FG33_4_HS</td>
<td>alkyl iodide</td>
<td>![Structure Image] (R—I, R = alkyl)</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>----</td>
<td>-------------</td>
<td>---------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>62</td>
<td>FG33_5_Hs</td>
<td>aryl fluoride</td>
<td>$\text{R} = \text{aryl}$</td>
</tr>
<tr>
<td>63</td>
<td>FG33_6_Hs</td>
<td>aryl chloride</td>
<td>$\text{R} = \text{aryl}$</td>
</tr>
<tr>
<td>64</td>
<td>FG33_7_Hs</td>
<td>aryl bromide</td>
<td>$\text{R} = \text{aryl}$</td>
</tr>
<tr>
<td>65</td>
<td>FG33_8_Hs</td>
<td>aryl iodide</td>
<td>$\text{R} = \text{aryl}$</td>
</tr>
<tr>
<td>66</td>
<td>FG34_LS</td>
<td>organometallic</td>
<td>$\text{R} = \text{any metal}$</td>
</tr>
<tr>
<td></td>
<td>compound</td>
<td></td>
<td>$\text{M} = \text{any metal}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\text{R} = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>67</td>
<td>FG34_1_Hs</td>
<td>organolithium</td>
<td>$\text{R} = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>68</td>
<td>FG34_2_Hs</td>
<td>organomagnesium</td>
<td>$\text{R} = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td>compound</td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>FG35_LS</td>
<td>carboxylic acid</td>
<td>$\text{R} = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td>derivative</td>
<td></td>
<td>$\text{X} = \text{any hetero atom}$</td>
</tr>
<tr>
<td>70</td>
<td>FG35_1_Hs</td>
<td>carboxylic acid</td>
<td>$\text{R} = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td>71</td>
<td>FG35_2_Hs</td>
<td>carboxylic acid</td>
<td>$\text{R} = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td>salt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>FG35_3_Hs</td>
<td>carboxylic acid</td>
<td>$\text{R}^{1} = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td>ester</td>
<td></td>
<td>$\text{R}^{2} = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>----</td>
<td>--------------</td>
<td>----------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>73</td>
<td>FG35_4_HS</td>
<td>lactone</td>
<td><img src="image" alt="lactone structure" /></td>
</tr>
<tr>
<td>74</td>
<td>FG35_5_HS</td>
<td>carboxylic acid</td>
<td><img src="image" alt="carboxylic acid primary amide" /></td>
</tr>
<tr>
<td></td>
<td></td>
<td>primary amide</td>
<td>$R = H$, alkyl, aryl</td>
</tr>
<tr>
<td>75</td>
<td>FG35_6_HS</td>
<td>carboxylic acid</td>
<td><img src="image" alt="carboxylic acid secondary amide" /></td>
</tr>
</tbody>
</table>
|    |              | secondary amide      | $R^1 = H$, alkyl, aryl  
|    |              |                      | $R^2 = \text{alkyl, aryl}$    |
| 76 | FG35_7_HS    | carboxylic acid      | ![carboxylic acid tertiary amide](image) |
|    |              | tertiary amide       | $R^1 = H$, alkyl, aryl  
|    |              |                      | $R^2 = \text{alkyl, aryl}$  
|    |              |                      | $R^3 = \text{alkyl, aryl}$  |
| 77 | FG35_8_HS    | lactam               | ![lactam structure](image)      |
|    |              |                      | $R = H$, alkyl, aryl            |
| 78 | FG35_9_HS    | carboxylic acid      | ![carboxylic acid hydrazide](image) |
|    |              | hydrazide            | $R^1 = H$, alkyl, aryl  
|    |              |                      | $R^2 = \text{alkyl, aryl}$  
|    |              |                      | $R^3 = \text{alkyl, aryl}$  
|    |              |                      | $R^4 = H$, alkyl, aryl        |
| 79 | FG35_10_HS   | carboxylic acid      | ![carboxylic acid azide](image) |
|    |              | azide                | $R = H$, alkyl, aryl            |
| 80 | FG35_11_HS   | hydroxamic acid      | ![hydroxamic acid](image)       |
|    |              |                      | $R = H$, alkyl, aryl            |
| 81 | FG35_12_HS   | carboxylic acid      | ![carboxylic acid amidine](image) |
|    |              | amidine              | $R^1 = H$, alkyl, aryl  
|    |              |                      | $R^2 = \text{alkyl, aryl}$  
|    |              |                      | $R^3 = \text{alkyl, aryl}$  
|    |              |                      | $R^4 = H$, alkyl, aryl        |
| 82 | FG35_13_HS   | carboxylic acid      | ![carboxylic acid amidrazone](image) |
|    |              | amidrazone           | $R^1 = H$, alkyl, aryl  
|    |              |                      | $R^2 = \text{alkyl, aryl}$  
|    |              |                      | $R^3 = \text{alkyl, aryl}$  
<p>|    |              |                      | $R^4 = H$, alkyl, aryl        |
| 83 | FG36         | nitrile              | <img src="image" alt="nitrile structure" />     |
|    |              |                      | $R = H$, alkyl, aryl            |</p>
<table>
<thead>
<tr>
<th>No</th>
<th>Alert ID</th>
<th>Alert Title</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>84</td>
<td>FG37_LS</td>
<td>acyl halide</td>
<td>R = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>X = F, Cl, Br, I</td>
</tr>
<tr>
<td>85</td>
<td>FG37_1_HS</td>
<td>acyl fluoride</td>
<td>R = H, alkyl, aryl</td>
</tr>
<tr>
<td>86</td>
<td>FG37_2_HS</td>
<td>acyl chloride</td>
<td>R = H, alkyl, aryl</td>
</tr>
<tr>
<td>87</td>
<td>FG37_3_HS</td>
<td>acyl bromide</td>
<td>R = H, alkyl, aryl</td>
</tr>
<tr>
<td>88</td>
<td>FG37_4_HS</td>
<td>acyl iodide</td>
<td>R = H, alkyl, aryl</td>
</tr>
<tr>
<td>89</td>
<td>FG38</td>
<td>acyl cyanide</td>
<td>R = H, alkyl, aryl</td>
</tr>
<tr>
<td>90</td>
<td>FG39</td>
<td>imido ester</td>
<td>R¹ = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R³ = H, alkyl, aryl</td>
</tr>
<tr>
<td>91</td>
<td>FG40</td>
<td>imidoyl halide</td>
<td>R¹ = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>R² = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>X = F, Cl, Br, I</td>
</tr>
<tr>
<td>92</td>
<td>FG41_LS</td>
<td>thioacarboxylic acid</td>
<td>R = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td>acid derivative</td>
<td>X = any hetero atom</td>
</tr>
<tr>
<td>93</td>
<td>FG41_1_HS</td>
<td>thioacarboxylic acid</td>
<td>R = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>X = OH, SH</td>
</tr>
<tr>
<td>94</td>
<td>FG41_2_HS</td>
<td>thioacarboxylic acid</td>
<td>R¹ = H, alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td>acid ester</td>
<td>R² = alkyl, aryl</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>X = O, S</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>----</td>
<td>---------------</td>
<td>---------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>95</td>
<td>FG41_3_HS</td>
<td>thiolactone</td>
<td><img src="image" alt="Thiolactone Structure" /></td>
</tr>
<tr>
<td>96</td>
<td>FG41_4_HS</td>
<td>thiocarboxylic acid amide</td>
<td><img src="image" alt="Thiocarboxylic Acid Amide Structure" /></td>
</tr>
<tr>
<td>97</td>
<td>FG41_5_HS</td>
<td>thiolactam</td>
<td><img src="image" alt="Thiolactam Structure" /></td>
</tr>
<tr>
<td>98</td>
<td>FG42</td>
<td>imidothioester</td>
<td><img src="image" alt="Imidothioester Structure" /></td>
</tr>
<tr>
<td>99</td>
<td>FG43</td>
<td>oxohetarene</td>
<td><img src="image" alt="Oxohetarene Structure" /></td>
</tr>
<tr>
<td>100</td>
<td>FG44</td>
<td>thioxohetarene</td>
<td><img src="image" alt="Thioxohetarene Structure" /></td>
</tr>
<tr>
<td>101</td>
<td>FG45</td>
<td>iminohetarene</td>
<td><img src="image" alt="Iminohetarene Structure" /></td>
</tr>
<tr>
<td>102</td>
<td>FG46_LS</td>
<td>orthocarboxylic acid derivative</td>
<td><img src="image" alt="Orthocarboxylic Acid Derivative Structure" /></td>
</tr>
<tr>
<td>103</td>
<td>FG46_1_HS</td>
<td>carboxylic acid orthoester</td>
<td><img src="image" alt="Carboxylic Acid Orthoester Structure" /></td>
</tr>
<tr>
<td>104</td>
<td>FG46_2_HS</td>
<td>carboxylic acid amide acetal</td>
<td><img src="image" alt="Carboxylic Acid Amide Acetal Structure" /></td>
</tr>
<tr>
<td>105</td>
<td>FG47</td>
<td>carboxylic acid anhydride</td>
<td><img src="image" alt="Carboxylic Acid Anhydride Structure" /></td>
</tr>
</tbody>
</table>

R¹ = H, alkyl, aryl
R² = H, alkyl, aryl
R³ = H, alkyl, aryl
R² = alkyl, aryl
R³ = alkyl, aryl

X = O, S
R = H, alkyl, aryl
R = H, alkyl, aryl
R = H, alkyl, aryl
R = H, alkyl, aryl
R = H, alkyl, aryl

X = OH, alkoxy, aryloxy, (substituted) amino, etc.
R¹ = H, alkyl, aryl
R² = H, alkyl, aryl
R² = alkyl, aryl
R³ = alkyl, aryl
R³ = alkyl, aryl
R³ = alkyl, aryl
R³ = H, alkyl, aryl
R² = H, alkyl, aryl
R² = H, alkyl, aryl
R² = H, alkyl, aryl
R¹, R² = H, alkyl, aryl
<table>
<thead>
<tr>
<th>No</th>
<th>Alert ID</th>
<th>Alert Title</th>
<th>Structure</th>
</tr>
</thead>
</table>
| 106| FG48_LS       | carboxylic acid imide                | ![Structural formula for carboxylic acid imide](image1) | $R^1, R^2 = H, alkyl, aryl$ $R^3 = H, alkyl, aryl$ ...
<p>| 107| FG48_1_HS     | carboxylic acid unsubstituted imide  | <img src="image2" alt="Structural formula for carboxylic acid unsubstituted imide" /> | $R^1, R^2 = H, alkyl, aryl$ |
| 108| FG48_2_HS     | carboxylic acid substituted imide    | <img src="image3" alt="Structural formula for carboxylic acid substituted imide" /> | $R^1, R^2 = H, alkyl, aryl$ $R^3 = anything but H$ |
| 109| FG49          | CO$_2$ derivative (general)          | any carbon with 4 valences to hetero atoms     |
| 110| FG50_LS       | carbonic acid derivative             | <img src="image4" alt="Structural formula for carbonic acid derivative" /> | $X, Y = any hetero atom$ |
| 111| FG50_1_HS     | carbonic acid monoester              | <img src="image5" alt="Structural formula for carbonic acid monoester" /> | $R = alkyl, aryl$ |
| 112| FG50_2_HS     | carbonic acid diester                | <img src="image6" alt="Structural formula for carbonic acid diester" /> | $R^1, R^2 = alkyl, aryl$ |
| 113| FG50_3_HS     | carbonic acid ester halide           | <img src="image7" alt="Structural formula for carbonic acid ester halide" /> | $R = alkyl, aryl$ $X = F, Cl, Br, I$ |
| 114| FG51_LS       | thiocarbonic acid derivative         | <img src="image8" alt="Structural formula for thiocarbonic acid derivative" /> | $X, Y = any hetero atom$ |
| 115| FG51_1_HS     | thiocarbonic acid monoester          | <img src="image9" alt="Structural formula for thiocarbonic acid monoester" /> | $R = alkyl, aryl$ |
| 116| FG51_2_HS     | thiocarbonic acid diester            | <img src="image10" alt="Structural formula for thiocarbonic acid diester" /> | $R^1, R^2 = alkyl, aryl$ |</p>
<table>
<thead>
<tr>
<th>No</th>
<th>Alert ID</th>
<th>Alert Title</th>
<th>Structure</th>
</tr>
</thead>
</table>
| 117| FG51_3_HS  | thiocarbonic acid ester halide                   | R = alkyl, aryl
X = F, Cl, Br, I                                                             |
| 118| FG52_LS    | carbamic acid derivative                         | R^1, R^2 = H, alkyl, aryl
X = OH, alkoxy, aryloxy, halogen                                             |
| 119| FG52_1_HS  | carbamic acid                                     | R^1, R^2 = H, alkyl, aryl                                                |
| 120| FG52_2_HS  | carbamic acid ester (urethane)                   | R^1, R^2 = H, alkyl, aryl
R^3 = alkyl, aryl                                                             |
| 121| FG52_3_HS  | carbamic acid halide                              | R^1, R^2 = H, alkyl, aryl
X = F, Cl, Br, I                                                             |
| 122| FG53_LS    | thiocarbamic acid derivative                     | R^1, R^2 = H, alkyl, aryl
X = OH, alkoxy, aryloxy, halogen                                             |
| 123| FG53_1_HS  | thiocarbamic acid                                 | R^1, R^2 = H, alkyl, aryl                                                |
| 124| FG53_2_HS  | thiocarbamic acid ester                          | R^1, R^2 = H, alkyl, aryl
R^3 = alkyl, aryl                                                             |
| 125| FG53_3_HS  | thiocarbamic acid halide                          | R^1, R^2 = H, alkyl, aryl
X = F, Cl, Br, I                                                             |
<p>| 126| FG54       | urea                                             | R^1, R^2, R^3, R^4 = H, alkyl, aryl                                      |
| 127| FG55       | isourea                                           | R^1, R^2, R^3, R^4 = H, alkyl, aryl                                      |</p>
<table>
<thead>
<tr>
<th>No</th>
<th>Alert ID</th>
<th>Alert Title</th>
<th>Structure</th>
</tr>
</thead>
</table>
| 128| FG56     | thiourea          | \[
R^1, R^2, R^3, R^4 = H, alkyl, aryl
\]                           |
| 129| FG57     | isothiourea       | \[
R^1, R^2, R^3, R^4 = H, alkyl, aryl
\]                           |
| 130| FG58     | guanidine         | \[
R^1, R^2, R^3, R^4 = H, alkyl, aryl
\]                           |
| 131| FG59     | semicarbazide     | \[
R^1, R^2, R^3, R^4, R^5 = H, alkyl, aryl
\]                           |
| 132| FG60     | thiosemicarbazide | \[
R^1, R^2, R^3, R^4 = H, alkyl, aryl
\]                           |
| 133| FG61     | azide             | \[
R ≡ N≡N≡N^−
\] R = alkyl, aryl |
| 134| FG62     | azo compound      | \[
R^1, R^2 = alkyl, aryl
\]                           |
| 135| FG63     | diazonium salt    | \[
R ≡ N≡N^+
\] R = alkyl, aryl |
| 136| FG64     | isonitrile        | \[
R ≡ N≡C^−
\] R = alkyl, aryl |
| 137| FG65     | cyanate           | \[
RO−C≡N
\] R = alkyl, aryl |
<table>
<thead>
<tr>
<th>No</th>
<th>Alert ID</th>
<th>Alert Title</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>138</td>
<td>FG66</td>
<td>isocyanate</td>
<td>$\ce{R-N=\equiv C=O}$ $\ce{R = alkyl, aryl}$</td>
</tr>
<tr>
<td>139</td>
<td>FG67</td>
<td>thiocyanate</td>
<td>$\ce{RS-\equiv C\equiv N}$ $\ce{R = alkyl, aryl}$</td>
</tr>
<tr>
<td>140</td>
<td>FG68</td>
<td>isothiocyanate</td>
<td>$\ce{R-N=\equiv C=\equiv S}$ $\ce{R = alkyl, aryl}$</td>
</tr>
<tr>
<td>141</td>
<td>FG69</td>
<td>carbodiimide</td>
<td>$^1\ce{R-N=\equiv C\equiv N=\equiv R^2}$ $\ce{R^1, R^2 = H, alkyl, aryl}$</td>
</tr>
<tr>
<td>142</td>
<td>FG70</td>
<td>nitroso compound</td>
<td>$\ce{R-N=O}$ $\ce{R = alkyl, aryl}$</td>
</tr>
<tr>
<td>143</td>
<td>FG71</td>
<td>nitro compound</td>
<td>$\ce{R-N=O}$ $\ce{R = alkyl, aryl}$</td>
</tr>
<tr>
<td>144</td>
<td>FG72</td>
<td>nitrite</td>
<td>$\ce{RO-N=O}$ $\ce{R = alkyl, aryl}$</td>
</tr>
<tr>
<td>145</td>
<td>FG73</td>
<td>nitrate</td>
<td>$\ce{RO-N=O}$ $\ce{R = alkyl, aryl}$</td>
</tr>
<tr>
<td>146</td>
<td>FG74_LS</td>
<td>sulfuric acid</td>
<td>$\ce{X-Y}$ $\ce{X, Y = any hetero atom}$</td>
</tr>
<tr>
<td>147</td>
<td>FG74_1_HS</td>
<td>sulfuric acid</td>
<td>$\ce{HO-SO-\equiv O}$</td>
</tr>
<tr>
<td>148</td>
<td>FG74_2_HS</td>
<td>sulfuric acid</td>
<td>$\ce{RO-SO-\equiv O}$ $\ce{R = alkyl, aryl}$</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>-----</td>
<td>-------------</td>
<td>------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>149</td>
<td>FG74_3_HS</td>
<td>sulfuric acid diester</td>
<td><img src="image" alt="diester" /> R¹, R² = alkyl, aryl</td>
</tr>
<tr>
<td>150</td>
<td>FG74_4_HS</td>
<td>sulfuric acid amide ester</td>
<td><img src="image" alt="ester" /> R¹ = alkyl, aryl R², R³ = H, alkyl, aryl</td>
</tr>
<tr>
<td>151</td>
<td>FG74_5_HS</td>
<td>sulfuric acid amide</td>
<td><img src="image" alt="amide" /> R¹, R² = H, alkyl, aryl</td>
</tr>
<tr>
<td>152</td>
<td>FG74_6_HS</td>
<td>sulfuric acid diamide</td>
<td><img src="image" alt="diamide" /> R¹, R², R³, R⁴ = H, alkyl, aryl</td>
</tr>
<tr>
<td>153</td>
<td>FG74_7_HS</td>
<td>sulfuryl halide</td>
<td><img src="image" alt="halide" /> X = F, Cl, Br, I Y = any hetero atom</td>
</tr>
<tr>
<td>154</td>
<td>FG75_LS</td>
<td>sulfonic acid derivative</td>
<td><img src="image" alt="derivative" /> R = alkyl, aryl X = any hetero atom</td>
</tr>
<tr>
<td>155</td>
<td>FG75_1_HS</td>
<td>sulfonic acid</td>
<td><img src="image" alt="acid" /> R = alkyl, aryl</td>
</tr>
<tr>
<td>156</td>
<td>FG75_2_HS</td>
<td>sulfonic acid ester</td>
<td><img src="image" alt="ester" /> R¹, R² = alkyl, aryl</td>
</tr>
<tr>
<td>157</td>
<td>FG75_3_HS</td>
<td>sulfonamide</td>
<td><img src="image" alt="amide" /> R¹ = alkyl, aryl R², R³ = H, alkyl, aryl</td>
</tr>
<tr>
<td>158</td>
<td>FG75_4_HS</td>
<td>sulfonyl halide</td>
<td><img src="image" alt="halide" /> R = alkyl, aryl X = F, Cl, Br, I</td>
</tr>
<tr>
<td>159</td>
<td>FG76</td>
<td>sulfone</td>
<td><img src="image" alt="sulfone" /> R¹, R² = alkyl, aryl</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>----</td>
<td>----------</td>
<td>-------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>160</td>
<td>FG77</td>
<td>sulfoxide</td>
<td><img src="image" alt="sulfoxide structure" /> ( R^1, R^2 = \text{alkyl, aryl} )</td>
</tr>
<tr>
<td>161</td>
<td>FG78_LS</td>
<td>sulfonic acid derivative</td>
<td><img src="image" alt="sulfonic acid derivative structure" /> ( R = \text{alkyl, aryl} ) ( X = \text{any hetero atom} )</td>
</tr>
<tr>
<td>162</td>
<td>FG78_1_HS</td>
<td>sulfonic acid</td>
<td><img src="image" alt="sulfonic acid structure" /> ( R = \text{alkyl, aryl} )</td>
</tr>
<tr>
<td>163</td>
<td>FG78_2_HS</td>
<td>sulfonic acid ester</td>
<td><img src="image" alt="sulfonic acid ester structure" /> ( R^1, R^2 = \text{alkyl, aryl} )</td>
</tr>
<tr>
<td>164</td>
<td>FG78_3_HS</td>
<td>sulfonic acid halide</td>
<td><img src="image" alt="sulfonic acid halide structure" /> ( R = \text{alkyl, aryl} ) ( X = \text{F, Cl, Br, I} )</td>
</tr>
<tr>
<td>165</td>
<td>FG78_4_HS</td>
<td>sulfonic acid amide</td>
<td><img src="image" alt="sulfonic acid amide structure" /> ( R^1 = \text{alkyl, aryl} ) ( R^2, R^3 = \text{H, alkyl, aryl} )</td>
</tr>
<tr>
<td>166</td>
<td>FG79_LS</td>
<td>sulfenic acid derivative</td>
<td><img src="image" alt="sulfenic acid derivative structure" /> ( R = \text{alkyl, aryl} ) ( X = \text{any hetero atom} )</td>
</tr>
<tr>
<td>167</td>
<td>FG79_1_HS</td>
<td>sulfenic acid</td>
<td><img src="image" alt="sulfenic acid structure" /> ( R = \text{alkyl, aryl} )</td>
</tr>
<tr>
<td>168</td>
<td>FG79_2_HS</td>
<td>sulfenic acid ester</td>
<td><img src="image" alt="sulfenic acid ester structure" /> ( R^1, R^2 = \text{alkyl, aryl} )</td>
</tr>
<tr>
<td>169</td>
<td>FG79_3_HS</td>
<td>sulfenic acid halide</td>
<td><img src="image" alt="sulfenic acid halide structure" /> ( R = \text{alkyl, aryl} ) ( X = \text{F, Cl, Br, I} )</td>
</tr>
<tr>
<td>170</td>
<td>FG79_4_HS</td>
<td>sulfenic acid amide</td>
<td><img src="image" alt="sulfenic acid amide structure" /> ( R^1 = \text{alkyl, aryl} ) ( R^2, R^3 = \text{H, alkyl, aryl} )</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>----</td>
<td>-----------</td>
<td>-----------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>171</td>
<td>FG80_LS</td>
<td>thiol</td>
<td>( R\rightarrow\text{SH} )  ( R = \text{alkyl, aryl} )</td>
</tr>
<tr>
<td>172</td>
<td>FG80_1_HS</td>
<td>alkylthiol</td>
<td>( R\rightarrow\text{SH} )  ( R = \text{alkyl} )</td>
</tr>
<tr>
<td>173</td>
<td>FG80_2_HS</td>
<td>arylthiol</td>
<td>( R\rightarrow\text{SH} )  ( R = \text{aryl} )</td>
</tr>
<tr>
<td>174</td>
<td>FG81_LS</td>
<td>phosphoric acid derivative</td>
<td>[ X, Y, Z = \text{O, N, Hal residue} ]</td>
</tr>
<tr>
<td>175</td>
<td>FG81_1_HS</td>
<td>phosphoric acid</td>
<td>[ \text{HO-P-OH} ]</td>
</tr>
<tr>
<td>176</td>
<td>FG81_2_HS</td>
<td>phosphoric acid ester</td>
<td>[ R = \text{alkyl, aryl} \  X, Y = \text{any O, N, Hal residue} ]</td>
</tr>
<tr>
<td>177</td>
<td>FG81_3_HS</td>
<td>phosphoric acid halide</td>
<td>[ X = \text{F, Cl, Br, I} \  Y, Z = \text{any O, N, Hal residue} ]</td>
</tr>
<tr>
<td>178</td>
<td>FG81_4_HS</td>
<td>phosphoric acid amide</td>
<td>[ R^1, R^2 = \text{H, alkyl, aryl} \  X, Y = \text{any O, N, Hal residue} ]</td>
</tr>
<tr>
<td>179</td>
<td>FG82_LS</td>
<td>thiophosphoric acid derivative</td>
<td>[ X, Y, Z = \text{any O, N, Hal residue} ]</td>
</tr>
<tr>
<td>180</td>
<td>FG82_1_HS</td>
<td>thiophosphoric acid</td>
<td>[ \text{HO-P-OH} ]</td>
</tr>
<tr>
<td>181</td>
<td>FG82_2_HS</td>
<td>thiophosphoric acid ester</td>
<td>[ R = \text{alkyl, aryl} \  X, Y = \text{any O, N, Hal residue} ]</td>
</tr>
<tr>
<td>No.</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>-----</td>
<td>----------</td>
<td>--------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>182</td>
<td>FG82_3_HS</td>
<td>thiophosphoric acid halide</td>
<td>$X = \text{F, Cl, Br, I}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$Y, Z = \text{any O, N, Hal residue}$</td>
</tr>
<tr>
<td>183</td>
<td>FG82_4_HS</td>
<td>thiophosphoric acid amide</td>
<td>$R^1, R^2 = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$X, Y = \text{any O, N, Hal residue}$</td>
</tr>
<tr>
<td>184</td>
<td>FG83_LS</td>
<td>phosphonic acid derivative</td>
<td>$R = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$X, Y = \text{any O, N, Hal residue}$</td>
</tr>
<tr>
<td>185</td>
<td>FG83_1_HS</td>
<td>phosphonic acid</td>
<td>$R = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>186</td>
<td>FG83_2_HS</td>
<td>phosphonic acid ester</td>
<td>$R^1, R^2 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$X = \text{any O, N, Hal residue}$</td>
</tr>
<tr>
<td>187</td>
<td>FG83_3_HS</td>
<td>phosphine</td>
<td>$R^1, R^2, R^3 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>188</td>
<td>FG83_4_HS</td>
<td>phosphinoxide</td>
<td>$R^1, R^2, R^3 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>189</td>
<td>FG84_LS</td>
<td>boronic acid derivative</td>
<td>$R = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$X, Y = \text{any O, N, Hal residue}$</td>
</tr>
<tr>
<td>190</td>
<td>FG84_1_HS</td>
<td>boronic acid</td>
<td>$R = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>191</td>
<td>FG84_2_HS</td>
<td>boronic acid ester</td>
<td>$R^1, R^2 = \text{alkyl, aryl}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$X = \text{any O, N, Hal residue}$</td>
</tr>
<tr>
<td>192</td>
<td>FG85</td>
<td>alkene</td>
<td>$R^1, R^2, R^3, R^4 = \text{H, alkyl, aryl}$</td>
</tr>
<tr>
<td>No</td>
<td>Alert ID</td>
<td>Alert Title</td>
<td>Structure</td>
</tr>
<tr>
<td>----</td>
<td>----------</td>
<td>-------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>193</td>
<td>FG86</td>
<td>alkyne</td>
<td>$R^1 \equiv \equiv R^2$ \hspace{1cm} $R^1, R^2 = H, \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>194</td>
<td>FG87</td>
<td>aromatic compound</td>
<td>any aromatic carbocyclic or heterocyclic structure, including cyclopentadienyl anion, tropylium cation, fulvene, tropone, pyridone-type lactams, etc.</td>
</tr>
<tr>
<td>195</td>
<td>FG88</td>
<td>heterocyclic compound</td>
<td>any cyclic structure with at least one non-carbon atom incorporated</td>
</tr>
<tr>
<td>196</td>
<td>FG89</td>
<td>alpha-aminoacid</td>
<td>$R^1, R^2 = H, \text{alkyl, aryl}$</td>
</tr>
<tr>
<td>197</td>
<td>FG90</td>
<td>alpha-hydroxyacid</td>
<td>$R = H, \text{alkyl, aryl}$</td>
</tr>
</tbody>
</table>
Abstract

Toxtree is a freely available, user-friendly and extensible software application that is designed to make structure-based predictions for a number of toxicological endpoints and mechanisms of chemical action. The platform has been developed by the Joint Research Centre in collaboration with Ideaconsult Ltd (Sofia, Bulgaria) with a range of modules developed by various contributors. One of the modules developed as an extension to Toxtree is aimed at the identification of organic functional groups in query chemicals. The rulebase consists of 204 organic functional groups recognised by the “Checkmol” program, which was developed by Dr Norbert Haider, University of Vienna. A new Functional Group Profiler has been coded as a Toxtree module by the Istituto Superiore di Sanita’ (Rome, Italy). The Toxtree profiler, called ISSFUNC, can be used to screen and characterise chemicals as a basis for read-across, category formation and (Q)SAR analysis. It can also be used for the global comparison of datasets, such as model training and test sets and chemical inventories.
How to obtain EU publications

Our priced publications are available from EU Bookshop (http://bookshop.europa.eu), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents. You can obtain their contact details by sending a fax to (352) 29 29-42758.
The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.