



DEVELOPMENT OF AN EXPERT SYSTEM RULEBASE FOR IDENTIFYING CONTACT ALLERGENS

M. D. BARRATT, D. A. BASKETTER, M. CHAMBERLAIN, M. P. PAYNE*, G. D. ADMANS†
and J. J. LANGOWSKI†

Unilever Environmental Safety Laboratory, Sharnbrook, Bedford MK44 1LQ, *Occupational Medicine and Hygiene Laboratory, Health and Safety Executive, Sheffield S3 7HQ and †LHASA UK, School of Chemistry, University of Leeds, Leeds LS2 9JT, UK

Abstract—There are currently no *in vitro* methods for the identification of skin sensitizers (contact allergens). Knowledge relating chemical structure to toxicity can be programmed into expert systems. An historical database containing results of 294 defined single substances tested in the guinea pig maximization test to a single protocol has been used to derive a set of structural alerts for skin sensitization. Where possible, the approach used was to group the substances according to their most likely mechanism of reaction with skin proteins. Where no mechanism could be identified, structural alerts were derived empirically for groups of molecules with similar chemical functionality. This process has currently resulted in the production of 40 structure-activity rules, which have been incorporated into the expert system DEREK. Rulebases of this type have potential for use as a preliminary screen in toxicological hazard identification and may ultimately lead to a reduction in the use of laboratory animals.

Introduction

Knowledge relating chemical structure to toxicity can be programmed into expert systems. DEREK (an acronym for 'Deductive Estimation of Risk from Existing Knowledge') is one such expert system (Sanderson and Earnshaw, 1991), being developed by LHASA UK at the School of Chemistry, University of Leeds. DEREK embodies both a controlling programme and a chemical rulebase. The chemical rulebase consists of descriptions of molecular substructures previously found to correlate with specific toxicological endpoints. The user communicates with DEREK by means of interactive computer graphics, by drawing the two-dimensional chemical structure of the query molecule on the screen.

There are currently no *in vitro* methods for the identification of skin sensitizers (contact allergens). To fill this gap, a rulebase is under development based on the well-recognized correlation of the protein reactivity of chemicals with their sensitization potential (Dupuis and Benezra, 1982). An historical (Unilever) database containing the results of testing 294 defined single substances in the guinea pig maximization test to a single protocol (Magnusson and Kligman, 1970), has been used to derive a set of structural alerts for skin sensitization.

Rationale

Test chemicals were grouped into four categories on the basis of the guinea pig maximization data—strong sensitizers (70–100% positive); moderate sensitizers (30–70% positive); weak sensitizers (up to 30% positive), and non-sensitizers (negative). To

identify groups of chemicals with common reaction mechanisms, initially only chemicals classifying as strong or moderate sensitizers were considered. These were the chemicals which currently would be classified as skin sensitizers according to the EC criteria (EEC, 1983).

For the purposes of identifying structural alerts, the chemicals were divided into groups, either based on reaction mechanisms or derived empirically. The principal groups are acylating agents, alkylating/aryllating agents, 'Michael' electrophiles and precursors, aldehydes and precursors, free radical generators, 'thiol-exchange' agents, and others (empirical). Forty structural alerts were then identified from these groups of compounds and programmed into the DEREK system.

Results and Discussion

Typical screens showing (a) a chemical structure prior to processing for structural alerts, (b) the structure with the structural alert identified and (c) the reference screen containing information about the rule which has been triggered, are shown in Fig. 1 a–c.

Of the 294 molecules in the database, 135 had sensitization scores such that they would be classified as sensitizers according to Annex V of the EC Dangerous Substances Directive (i.e. strong and moderate, + 30% or more positive) and 159 would be classified as non-sensitizers (made up of 120 negative and 39 weak).

Of the 135 sensitizers, DEREK identified 132 as containing a structural alert likely to lead to skin

sensitization. Of the 120 negative chemicals, DEREK found no structural alert for 98 molecules, but identified 22 as false positives. Of the 39 weak sensitizers (chemicals showing some response but not classifying as sensitizers), DEREK found structural alerts for 16.

The three sensitizers not identified by DEREK were abietic acid, cinnamic alcohol and anethole. The sensitization potential of the former two molecules has been shown to arise from sensitizing impurities (Basketter, 1992; Karlberg, 1991); refinements to the rulebase are required to accommodate anethole and similar molecules.

It is now widely accepted that the ability to react with a nucleophile either directly or after appropriate metabolism is a prerequisite for the majority of skin sensitizers. It should not be overlooked, however that

skin permeability is another key factor in determining whether or not a chemical is in practice a sensitizer. The octanol-water partition coefficient (P) of a molecule has been shown to be an important factor determining its skin permeability (Flynn, 1990), higher $\log P$ values (i.e. greater lipophilicity) broadly leading to greater permeability. Quantitative structure-activity relationships of skin sensitization invariably require the inclusion of $\log P$ as an independent variable (e.g. Roberts and Basketter, 1990).

It seems reasonable, therefore, that some molecules may contain a structural alert (i.e. be reactive, presumably by the same mechanism), but may not be skin sensitizers because their skin permeability is too low. For example, 4-aminobenzoic acid (a zwitterionic compound) and sodium 4-benzoyloxybenzoate are weak or non-sensitizers, whereas

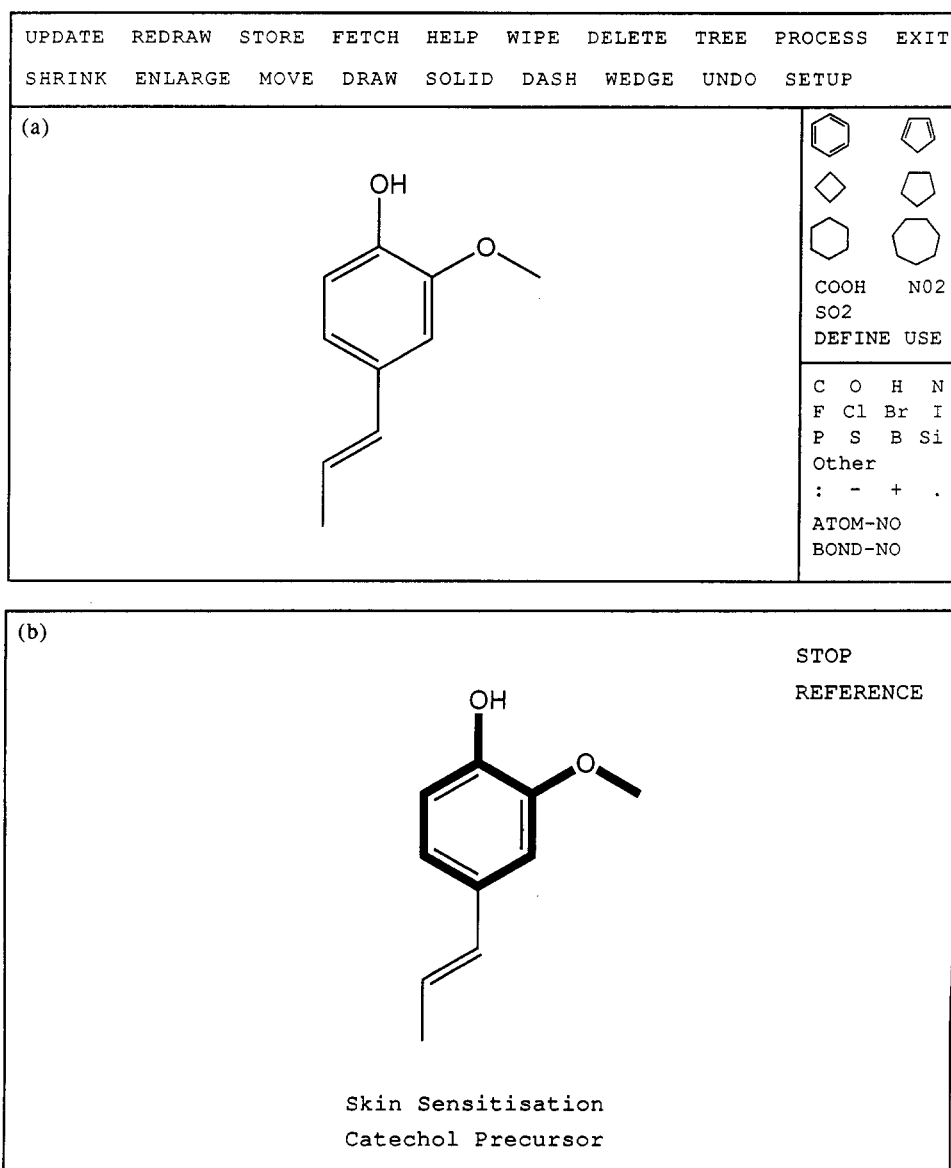


Fig. 1(a)-(b).

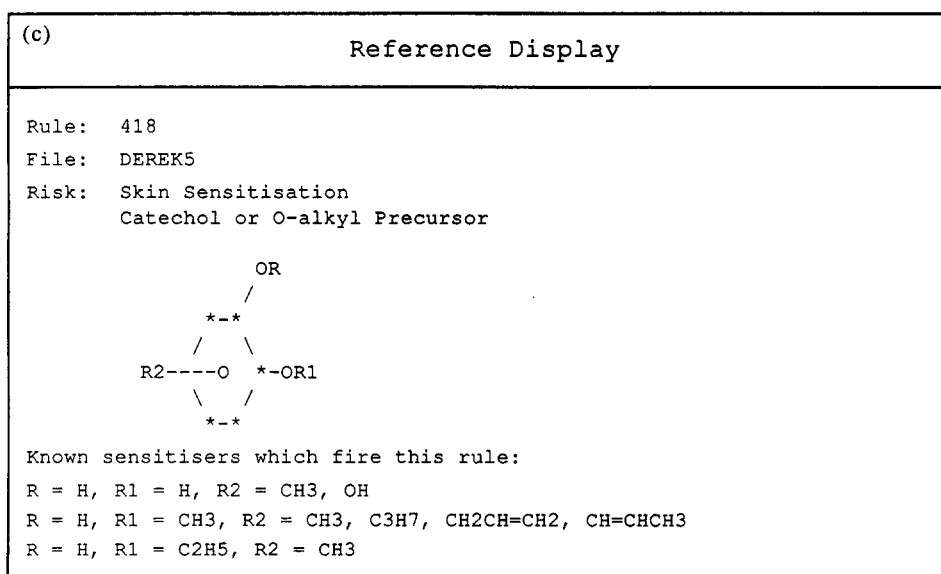


Fig. 1(c).

Fig. 1. Typical DEREK screens showing (a) a molecular structure (isoeugenol) prior to processing for structural alerts, (b) the structure with the structural alert identified and (c) the reference screen containing information about the rule which has been fired.

4-chloroaniline and phenyl benzoate are moderate or strong sensitizers. The former two molecules are both quite polar by virtue of possessing charged groups and will thus penetrate the skin much less readily than the latter two molecules. Several more of the 22 non-sensitizers identified by DEREK as containing structural alerts appear to fall into this category.

It is our intention to continue to develop the sensitization rulebase within DEREK in two ways: (a) to add additional rules [e.g. derived from sources such as classification data from the European List of New Chemical Substances (ELINCS)] and (b) to incorporate an estimate of the skin permeability in order to refine the identification of potential sensitizers. The potential of rulebases of this type for use in toxicological hazard screening can only increase and may ultimately lead to a reduction in the use of laboratory animals.

REFERENCES

- Basketter D. (1992) Skin sensitization to cinnamic alcohol: the role of skin metabolism. *Acta dermato-venereologica* **72**, 264–265.
- Dupuis G. and Benezra C. (1982) *Contact Dermatitis to Simple Chemicals: A Molecular Approach*. Marcel Dekker, New York.
- EEC (1983) EEC Commission Directive of 29 July 1983 adapting to technical progress for the fifth time Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. (Annex V). *Official Journal of the European Communities* **L257**, 1.
- Flynn G. L. (1990) Physicochemical determinants of skin absorption. In *Principals of Route-to-Route Extrapolation for Risk Assessment*. Edited by T. R. Gerrity and C. J. Henry. pp. 93–127. Elsevier, New York.
- Karlberg A.-T. (1991) Air oxidation increases the allergic potential of Tall Oil Rosin. Colophony contact allergens also identified in Tall Oil Rosin. *American Journal of Contact Dermatitis* **2**, 43–49.
- Magnusson B. and Kligman A. M. (1970) *Allergic Contact Dermatitis in the Guinea-Pig. Identification of Contact Allergens*. Charles A. Thomas, Springfield, IL.
- Roberts D. W. and Basketter D. A. (1990) A quantitative structure activity/dose response relationship for contact allergic potential of alkyl group transfer agents. *Contact Dermatitis* **23**, 331–335.
- Sanderson D. M. and Earnshaw C. G. (1991) Computer prediction of possible toxic action from chemical structure: the DEREK system. *Human and Experimental Toxicology* **10**, 261–273.