Structure-Activity Relationships for Skin Sensitization Potential: Development of Structural Alerts for Use in Knowledge-Based Toxicity Prediction Systems†

M. P. Payne and P. T. Walsh*

Occupational Medicine and Hygiene Laboratory, Health and Safety Executive, Broad Lane, Sheffield S3 7HQ, U.K.

Received June 26, 1993®

The development of qualitative structure-activity relationships for the prediction of skin sensitization potential, based on structural alerts (substructures associated with a toxicological mechanism), and suitable for incorporation as rules into a knowledge-based system is described. The structure dependence of the skin sensitization mechanism may be largely defined in terms of the presence or metabolic/nonmetabolic formation of protein reactive functional groups on the test compound and by the physicochemical requirements of significant skin penetration. The proposed structural alerts were tested on a data set of diverse chemicals. The results showed that the alerts have potential as preliminary indicators of skin sensitization potential for a wide range of low molecular weight chemicals.

INTRODUCTION

The use of well-validated structure-activity relationships (SAR) to predict toxicity has the potential to produce considerable efficiency gains for a regulatory authority by directing testing resources on those chemicals of greatest potential toxicological hazard and on their more likely areas of toxicological concern. An important requirement for the application of such SAR to the hazard assessment of chemicals, when there is a lack of experimental data, is that they should be readily defensible and transparent in their application. Toxicity prediction rules based on well-validated structural alerts (substructures associated with a particular toxicological end point), obtained from the requirements of accepted toxicological mechanisms or, more empirically, from the toxicological properties of a well-defined range of similar substances, offer promise in this context. Such SAR have been developed for mutagenicity.[1] SAR may also provide valuable insights into the mode of action within a class of chemicals.

Skin sensitization (allergic contact dermatitis) is an important industrial disease known to be caused by a very wide range of chemicals.[2] Animal tests for this end point are incorporated in the "base-set" battery of regulatory tests for many types of chemicals including new industrial chemical substances. The mechanism is a general one which is reasonably well understood and indicates a common set of physicochemical and chemical determinants for activity.[3] Thus a mechanistically-led SAR analysis of skin sensitization would appear to be feasible. The large number of experimental variables in conventional animal tests and the predominant use of a subjectively assessed end point limits the derivation of *quantitative* SAR for skin sensitization. However *qualitative* SAR using structural alerts suggested by mechanistic considerations are supported by human and experimental skin sensitization studies.[3,4]

Subject to their validation with high-quality data, these structural alerts may be incorporated as "rules" into a knowledge-based system (e.g. DEREK—deductive estimation of risk from existing knowledge[5,6]), providing a useful tool for regulatory toxicologists when the potential toxicological hazards of chemicals are assessed. This paper describes the development of some skin sensitization SAR rules suitable for inclusion into such a system.

SKIN SENSITIZATION

The term skin sensitization, in the context of this paper, refers to T-cell mediated sensitization *via* the dermal route. The process occurs in two phases: the induction of sensitization following initial exposure and the subsequent elicitation of the sensitization reaction following later exposure. The process may be broken down into the following steps:

Induction of Sensitization Phase

(I) penetration of the allergen into the epidermis through the outer horny layer of the skin (*stratum corneum*)
(II) uptake or surface (membrane) binding of the chemical (or its protein conjugate) by antigen-presenting cells
(III) metabolism to reactive hapten (if required)
(IV) reaction of hapten with cellular protein to form antigen
(V) processing of antigen by the cell, including its presentation on the outer surface of the cell in association with genetically determined MHC transcellular glycoproteins
(VI) passage of antigen bearing cells to lymph nodes
(VII) recognition of antigen by specific T-cells
(VIII) proliferation of specific T-cells and their dissemination around the body by the blood system

Elicitation of Sensitization Phase

(IX) recurrence of stages I–V
(X) presentation of processed antigen to enhanced population of specific T-cells in skin
(XI) production of local inflammatory response

Details of hapten processing by antigen presenting cells still remain to be elucidated. The structure dependence of the mechanism may be largely defined in terms of the presence or metabolic/nonmetabolic formation of protein reactive functional groups on the allergen[3,4] and by the physicochemical requirements of significant skin penetration which have been

† Presented at the 3rd International Conference on Chemical Structures, Noordwijkerhout, The Netherlands, June 6–10, 1993.
® Abstract published in *Advance ACS Abstracts*, January 15, 1994.

0095-2338/94/1634-0154$04.50/0 Published 1994 by the American Chemical Society

described in the literature;[7] i.e., stages I–IV. The influence of chemical structure on these stages is now addressed.

Protein Reactivity. All low molecular weight chemicals or their metabolic products, which function immunologically as haptens, must combine with protein or polypeptides to form an antigen capable of eliciting an allergic reaction.[8] The prediction of whether a reaction of a hapten with protein is likely *in vivo* may initially be based on knowledge of protein reactivity[9,10] and/or the experimental observation of protein conjugate formation *in vitro* (e.g., ref 11) or the reaction of hapten with a model nucleophile such as *n*-butylamine or aniline. Our examination of experimental skin sensitization data in the published literature (e.g., from ref 12) obtained using well-established protocols for simple, primarily low molecular weight (<500), monofunctional organic compounds showed that most compounds which possessed protein reactive groups were significant sensitizers. Exceptions were found for some corrosive compounds and/or those deactivated by hydrolysis (e.g., low molecular weight acid chlorides) or for ionic compounds with poor skin penetration characteristics (e.g., Cr^{3+} salts). Reactions with skin proteins under physiological conditions may occur by several mechanisms and involve a range of reactive groups on the allergen and protein; examples of structural alerts associated with such reactions are illustrated below in the Structural Alerts section.

Simple structural alerts of high sensitivity may in some instances be based purely on the presence of a particular type of functional group of high reactivity, e.g., aliphatic aldehydes. However, it is clear that for a high frequency of skin sensitization to be exhibited the necessary level of reactivity of a functional group may in many instances require appropriate activation by other groups or substituents in the molecule. As a result, accurate structural alerts for skin sensitization will frequently be complex. For example, 2,4-dinitrohalobenzenes and halotriazines are both sensitizers as a consequence of nitro group or heterocyclic nitrogen-based activation of the halogen to substitution by nucleophilic protein amino groups. Other substituents (e.g., CF_3) may also activate toward such nucleophilic attack or participate as leaving groups in such reactions (e.g., OMe) and may be defined in the structural alerts.

Chemical Activation of Allergens. Some chemicals are not in themselves protein-reactive but become so following metabolism in the skin. The allowance for well-recognized routes of metabolism and to some extent nonmetabolic chemical transformations (e.g., hydrolysis, autooxidation) has been achieved[13] through a critical examination of the skin metabolism literature and an analysis of experimental skin sensitization data on groups of such compounds. For example, *o*- or *p*-dihydroxybenzenes are activated toward protein reactivity by the formation of quinones/semiquinones or related radical species;[14] 2- or 4-methoxy-substituted phenols require demethylation prior to such oxidation, which reduces their sensitization potential relative to the dihydroxy compounds,[15] and the dimethyl ether analogues are of very low sensitizing potential. A further example is the azo dyes which require, for their metabolic reduction to sensitizing anilines, strongly electron donating groups (e.g. hydroxy, amino) in the same aromatic ring.[16]

Skin Penetration. Some classes of compounds, particularly those of low protein reactivity, e.g., alkyl halides, additionally require high lipophilicity for activity. Such effects may be attributed to the need for good skin penetration and/or enhanced association by hydrophobic interactions with skin proteins. Theoretical treatments predict that skin penetration is governed primarily by lipophilicity and molecular size.[17] No general, accurate quantitative relationship has been derived that applies *in vivo* (or *in vitro*) for all chemical classes of polyfunctional small- or moderate-sized (molecular weight <1000) molecules.[18] However, for individual classes of chemicals (e.g., alcohols, phenols, and steroids) plots of skin penetration (as approximated by *in vitro* permeability coefficient) normally exhibit a rise in skin penetration with increasing octanol–water partition coefficient (log *P*) that allows quantitative structure–permeability relationships to be derived. There are however exceptions to this rule for certain classes of compounds, and a reduction in skin penetration at high values of log *P* is frequently observed. Consideration of calculated skin penetration rates with protein reactivity may therefore allow the refinement of structural alerts.

For ionic compounds skin penetration is usually low, e.g., Cr^{3+} ions are reported to be the active antigen in chromate sensitization[19] but Cr^{3+} salts and complexes are in fact weak or insignificant sensitizers as a result of poor skin penetration.[20] Ion pair formation however may provide a means of increasing effective lipophilicity, thereby enhancing skin penetration.

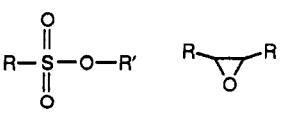
STRUCTURAL ALERTS

The following list of structural alerts is based on the structural requirements for reactivity with skin proteins, assessed where possible from existing published data on experimental skin sensitization. The alerts are classified, where possible, according to the anticipated reaction mechanism of the test substance with skin protein.

A. Skin Protein Alkylating Agents. Alkylation of proteins under physiological conditions commonly occurs at lysine and cysteine thiol groups and, when available, protonated nitrogen on histidine rings.[3] Alkylating agents, active under mild conditions, have an electropositive or conjugated alkyl center and a good leaving group (X). Such compounds include haloalkanes activated toward reactions with nucleophiles by conjugation with unsaturated groups: RCH_2X where R represents C=C, C≡C, Ar–, R′C=O, $-CO_2R'$, –CONR′, $R'R''NCH_2$–, $R'OCH_2$, and X = Cl, Br, and I, for example. (Ar denotes an aromatic ring).

Skin sensitization data are published for examples of most of these classes of compounds and confirm their activity. The order of reactivity and lipophilicity influencing skin penetration increases in the same order, X = Cl < Br < I hence skin sensitization potential is expected to increase in this order. Experimental data for 4-nitrobenzal halides supports this prediction.[21] Electron withdrawing groups in the aromatic ring increase reactivity and sensitization potential. Unactivated alkyl halides are not significantly sensitizing unless markedly lipophilic (i.e., log $P_{oct/water} > 4$).[22]

Other examples of sensitizing reactive alkylating agents include dialkyl sulfonates, dialkyl sulfates, and epoxides:



R, R′ = alkyl

B. Skin Protein Arylating Agents. Arylation of skin protein may result from nucleophilic substitution at an appropriately activated aromatic centre, e.g., ortho, para dinitrosubstituted halogenated aromatic compounds, by protein amino of sulfhydryl groups. The presence of heterocyclic nitrogen atoms

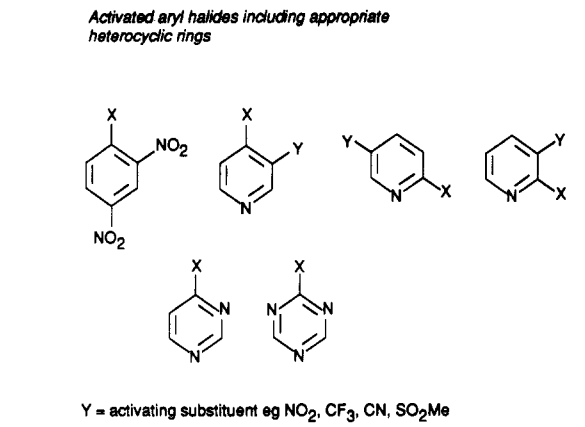


Figure 1. Skin sensitization structural alerts—skin protein arylating agents.

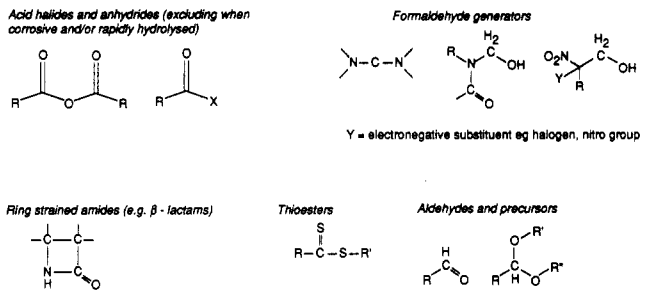


Figure 2. Skin sensitization structural alerts—skin protein acylating/sulfonating agents.

also strongly activates toward nucleophilic attack good leaving groups such as halogens in ortho or para positions.[23] The presence of two or more such activating groups allows reaction to occur under mild, e.g., physiological, conditions. Consequently 2,4-dinitrochlorobenzene and chlorotriazine fungicides such as Simazine are strong skin sensitizers.[24] Examples of structural alerts based on the potential for protein arylation are shown in Figure 1.

C. Skin Protein Acylating/Sulfonating Agents. Reactive carbonyl groups can acylate lysine amino groups in proteins thus.[9,10]

$$RCOX + \text{protein-}NH_2 \rightarrow \text{protein-}RNHCOR$$

Figure 2 shows structural alerts for various skin protein acylating or sulfonating agents which are known to have significant skin sensitization potential. In aqueous media such reactions compete with the hydrolysis of the reagent. As a consequence many highly reactive acyl halides or anhydrides are not significant skin sensitizers, although they are generally markedly corrosive or irritant. Esters with good leaving groups, such as *N*-hydroxysuccinamide (NHS), have been used to acylate proteins; however their sensitizing properties are not well established. Phenyl esters are similarly reactive and are generally marked skin sensitizers.[25] Alkyl esters are less reactive and are generally inactive in sensitization assays.

D. Michael Addition Electrophiles and Precursors. These are defined as compounds with an alkenic double bond conjugated (i.e., adjacent to) a multiply bonded electron withdrawing group which may react with nucleophilic groups of proteins, primarily cysteine sulfhydryl groups through a Michael type addition of the nucleophile to the terminal alkene carbon atom. Their reactivity with proteins has been well documented. Some examples are shown in Figure 3.

E. Thiol Exchange Compounds. Thiol exchange occurs as a consequence of nucleophilic attack of the thiol group on a

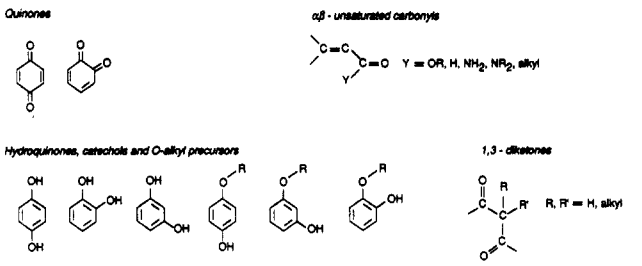


Figure 3. Skin sensitization structural alerts—Michael addition electrophiles and precursors.

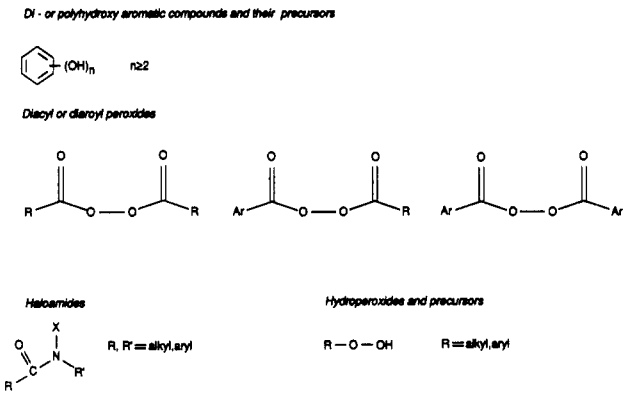


Figure 4. Skin sensitization structural alerts—free radical generators.

disulfide bridge, e.g.,

$$RSH + \text{protein-S-S-protein} \rightarrow \text{protein-S-S-R} + HS\text{-protein}$$

Structural types that may react in this way include the following thiols and disulfides:[9]

R–SH Ar–SH R–S–S–R Ar–S–S–R Ar–S–S–Ar

An example of a skin sensitizer of this class is benzothiazole.[26]

F. Free Radical Generators. Skin proteins can react with free radicals generated by enzymatic oxidation or photochemical reaction of several groups of compounds. Reaction can either occur by direct addition, e.g., of an alkyl or aryl group, or involve hydrogen abstraction by the radical species which is followed by further intermolecular reaction such as cross-linking. Such modified proteins may be sufficiently antigenic to result in sensitization. Radical reactions have been implicated in the reactions of several allergens such as eugenol.[15] Substructures associated with the generation of reactive radicals are shown in Figure 4.

G. Metabolism to Reactive Electrophiles. Reactive species which differ considerably in molecular size and chemical functionality from their precursors may be formed by means of enzymic transformations in the skin. The sensitizing properties of several groups of compounds, e.g., aromatic amines and dihydroxybenzenes shown in Figure 5, are generally considered to be dependent on such reactions.

H. Miscellaneous. Some structural alerts which cannot, at present, be assigned to a particular reaction mechanism but are nevertheless associated with significant skin sensitization potential are shown in Figure 6.

ASSESSMENT OF STRUCTURAL ALERTS

The predictive success of the above proposed structural alerts was tested using published European Community (EC) labeling information on substances notified under the Notification of New Substances Regulations.[27] The substances are labeled as skin sensitizers (using the R43 risk phrase)

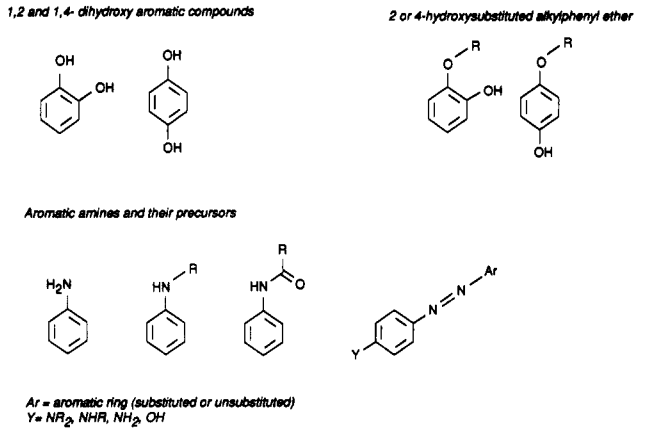


Figure 5. Skin sensitization structural alerts—metabolized to reactive electrophiles.

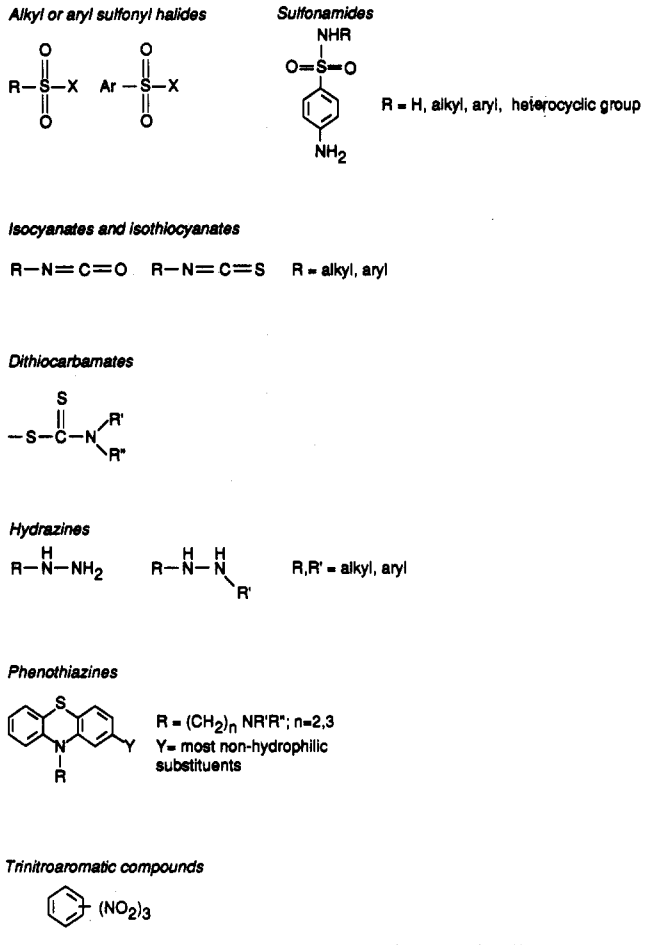


Figure 6. Skin sensitization structural alerts—miscellaneous.

primarily on the basis of the production of a minimum specified frequency of sensitization in animal tests. The substances are of diverse and frequently complex (polyfunctional) structural type and are considered to represent a good test of SAR methods. A list of the 93 compounds examined with their registry numbers and sensitization and structural alert data is shown in the table in Appendix 1. The results of the assessment of the proposed structural alerts on this sample of compounds are shown in Table 1.

The following discussion is concerned with the 43 substances which are sensitizers (25 true positives and 18 "false negatives") and the 4 nonsensitizers posssessing structural alerts (false positives). These substances are partitioned into those which do possess alerts and those which do not.

Table 1. Summary of Prediction Results

		actual		
		positives	negatives	total
predicted	positives	25	4	29
	negatives	18	46	64
	total	43	50	93

Table 2. Chemical Classes of Sensitizers with Structural Alerts

chemical class	ELINCS no.	structural alert type[a]	example structure no.
fluoro/chlorotriazine	400-010-9	B	I
	400-380-1		
	400-430-2		
	400-790-0		
	401-560-2		
	401-650-1		
	402-150-6		
	402-170-5		
$\alpha\beta$-unsaturated carbonyl	400-650-9	D	
	401-970-1		II
	402-650-4		
	402-970-4		
activated 2-/4-halopyridine/ halodiazine	400-290-2	B	III
N-chloroamide	401-570-7	C	IV
acid halide	401-800-6	C	V
sulfonamide	401-190-1	H	
	402-050-2		VI
alkyl isocyanate	402-290-8	H	
	402-440-2		VII
aromatic amine	402-190-4	G	VIII
disulfide	400-930-0	E	IX
p-NR_2, OR-substituted azo	400-460-6	G	X
benzothiazole	401-450-4	E	XI
	402-540-6		
benzyl halide	402-210-1	B	XII

[a] Refer to main text under Structural Alerts.

Substances Possessing Recognized Structural Alerts. The 25 substances that were correctly predicted to be sensitizing are classified in Table 2 according to their chemical class and then to their alerting type based on their protein reaction mechanism. The chemical structures of compounds **I–XII** in Table 2 are shown in Figure 7. Similar details for the four nonsensitizers which possess structural alerts (false positives; structures **XIII–XVI**) are shown in Table 3 and their structures are shown in Figure 8. All four compounds contain one or more primary or secondary aromatic amine groups, and one is an anthraquinone and another a chlorotriazine. The anthraquinone compound is an ionic salt which is expected to result in limited skin absorption. From studies of aromatic amines,[13] the fluorinated aromatic amine (**XIII**) is probably insufficiently sensitizing to be classified due to the reduction of the ease of oxidation to reactive intermediates (or their precursors) relative to aniline, by the electron withdrawing nature of the substituents.

Sensitizers Not Possessing Recognized Structural Alerts. The structures of the 18 skin sensitizers not covered by the above structural alerts (false negatives), suggest several potentially haptenic or prohaptenic chemical classes for further consideration; moreover, some of these compounds are highly lipophilic. A description of these substances is shown in Table 4. However, it should be noted that low levels of strongly sensitizing impurities in the tested preparations may, in some cases, make a significant contribution to the overall sensitization potential. Clearly, further information on the skin sensitization properties and/or potential protein reactivity of the classes of chemicals listed in Table 4 is required to develop reliable structural alerts from these initial results.

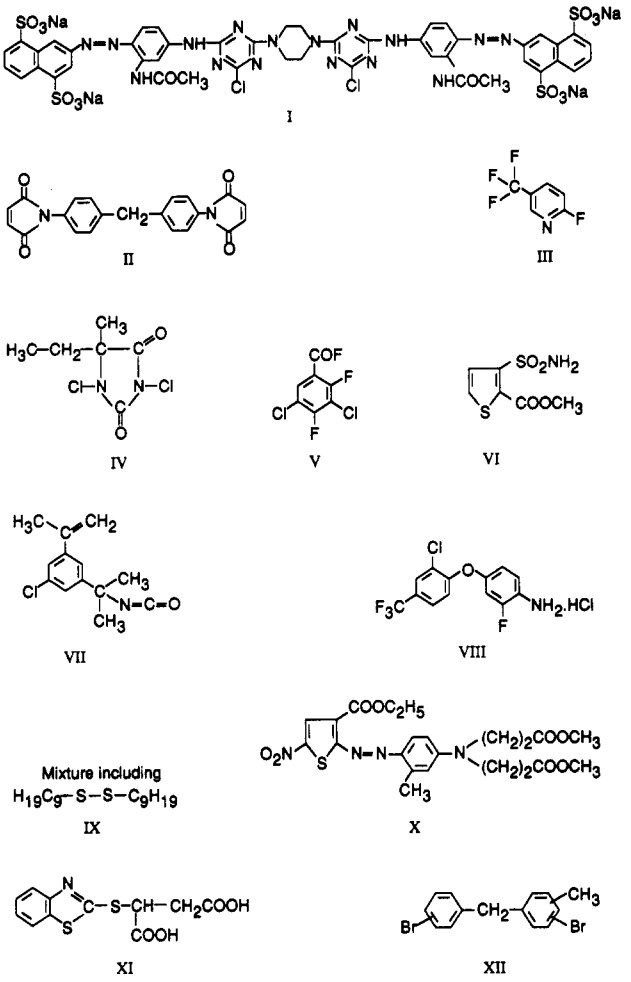


Figure 7. Structures of some sensitizers possessing structural alerts.

Table 3. Chemical Classes of Nonsensitizers Possessing Structural Alerts

chemical class	ELINCS no.	structural alert type[a]	example structure no.
primary/secondary (alkyl)	401-790-3	G	XIII
aromatic amine	400-680-2		XIV
halotriazine	400-120-7	B	XV
quinone/primary aromatic amine	400-350-8	D/G	XVI

[a] Refer to main text under Structural Alerts.

CONCLUSIONS

The combination of structural alerts based on protein reactivity with considerations of skin metabolism, skin penetration (governed predominantly by lipophilicity), and reactivity variation induced by substituent effects allows the qualitative prediction of skin sensitization potential for many simple, monofunctional chemicals and a substantial proportion of more complex, polyfunctional molecules. The proposed

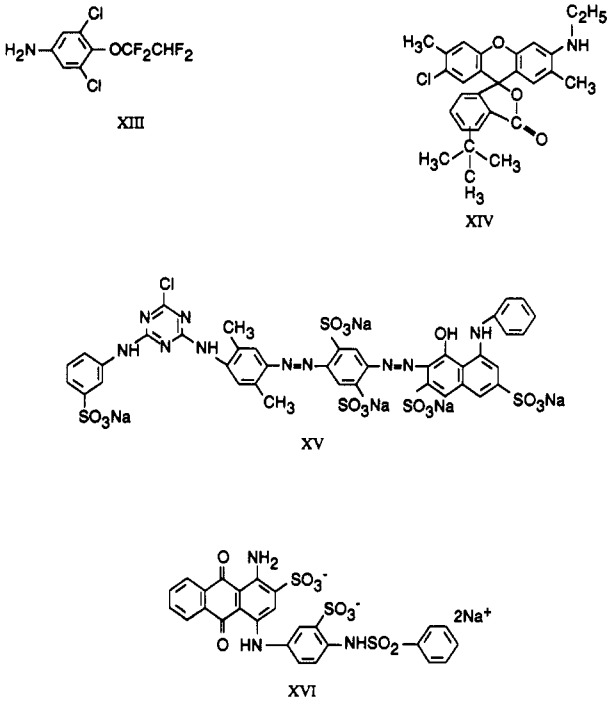


Figure 8. Structures of nonsensitizers possessing structural alerts.

structural alerts therefore have potential as preliminary indicators of skin sensitization potential for a wide range of chemicals of low molecular weight (less than approximately 600).

The overall predictive success of the alerts was 76% (71/93), with 58% of the compounds labeled as skin sensitizers and 92% of those not labeled as sensitizers being correctly predicted. Given that a considerable proportion (42%) of the compounds labeled as sensitizers were not identified by the proposed structural alerts, further validation, extension, and refinement of existing and new alerts using more experimental data are desirable for their confident use in regulatory matters. The application of toxicokinetic considerations, especially the modeling of skin penetration by means of lipophilicity and molecular size descriptors, and the identification, where possible, of skin metabolites may improve the predictive success of the system. Factors such as complex chemical group interactions and the present poor definition of skin metabolism may limit the ultimate sensitivity and accuracy of the approach. However, the guidance provided by the SAR developed to data should prove useful to regulatory toxicologists.

The structural alerts may be represented as rules in a knowledge-based system capable of automatic recognition of chemical structure, as in the DEREK system.[5,6] For use in DEREK, the structural alerts are coded into rules using the chemical structure languages PATRAN and CHMTRAN.[28] The structural alert is represented firstly in a PATRAN statement which can encode any substructure. Qualifying statements (e.g., restrictions on log P and interactions between

Table 4. Description of Sensitizers Not Possessing Structural Alerts

description	ELINCS no.	description	ELINCS no.
non-halo-substituted triazines/diazines	401-340-6	alkyl sulfonate–metal salt	401-640-7
	401-990-0	dialkyl imine	401-660-6
metal alkyl/hydrogen borate	401-040-5	*O*-alkyl hydroxlamine	402-030-3
benzotriazole/dialkyl phenol/alkyl ether	400-830-7	dialkyl benzamides	401-980-6
o-hydroxy naphthyl azo/aryl sultone/aryl amide	401-010-1		402-460-1
N-alkyl triazole/*t*-alkyl amine	401-280-0	α-amino–alkyl amide	402-840-7
alkyl thiosulfate salt	401-320-7	trinorbornenylpyridine	402-520-7
dihydropyrazole/aryl sultone/*t*-alkyl amine	401-410-6	alcohol	402-770-7
cyclic alkene/dialkyl carbonate	401-620-8	3,7-dichloroquinoline	402-780-1

Table 5. Substances Used in Assessment of Structural Alerts

name	ELINCS no.	CAS no.	skin sensitizer	alert
tetrasodium 3,3′-(piperazine-1,4-diylbis((6-chloro-1,3,5-triazine-4,2-diyl)imino(2-acetamido)-1,4-phenyleneazo))bis(naphthalene-1,5-disulfonate)	400-010-9	81898-60-4	+	+
2-fluoro-5-trifluoromethylpyridine	400-290-2	69045-82-5	+	+
disodium 6-((4-chloro-6-(*N*-methyl)-2-toluidino)-1,3,5-triazin-2-ylamino)-1-hydroxy-2-(4-methoxy-2-sulfonatophenylazo)naphthalene-3-sulfonate	400-380-1	86393-35-3	+	+
tetrasodium 2-(6-chloro-4-(2,5-dimethyl-4-(2,5-disulfonatophenylazo)phenylazo)-3-ureidoanilino)-1,3,5-triazin-2-ylamin)benzene-1,4-disulfonate	400-430-2		+	+
dimethyl(3-methyl-4-(5-nitro-3-ethoxycarbonyl-2-thienyl)azo)phenylnitrilodipropionate	400-460-6		+	+
methyl 2-(2-nitrobenzylidene)acetoacetate	400-650-9	39562-27-1	+	+
tetrasodium 5-benzamido-3-(5-(4-fluoro-6-(1-sulfonato-2-naphthylamino)-1,3,5-triazin-2-ylamino)-2-sulfonatophenylazo)-4-hydroxynaphthalene-2,7-disulfonate	400-790-0	85665-97-0	+	+
α-3-(3-(2*H*-benzotriazol-2-yl)-5-*tert*-butyl-4-hydroxyphenyl)propionyl-ω-hydroxypoly(oxyethylene) AND α-3-(3-(2*H*-benzotriazol-2-yl)-5-*tert*-butyl-4-hydroxyphenyl)propionyl-ω-3-(3-(2*H*-benzotriazol-2-yl)-5-*tert*-butyl-4-hydroxyphenyl)propionyloxypoly(oxyethylene)	400-830-7		+	–
C_{12}-14-*tert*-alkylammonium diphenyl phosphorothioate AND dinonyl sulfide (or disulfide)	400-930-0		+	+
dilithium 6-acetamido-4-hydroxy-3-(4-((2-sulfonatooxy)ethylsulfonyl)phenylazo)naphthalene-2-sulfonate	401-010-1		+	–
dibutyltin hydrogen borate	401-040-5	75113-37-0	+	–
methyl 2-(3-(4-methoxy-6-methyl-1,3,5-triazin-2-yl)-3-methylureidosulfonyl)benzoate	401-190-1	101200-48-0	+	+
N,*N*-bis(2-ethylhexyl)((1,2,4-triazol-1-yl)methyl)amine	401-280-0	91273-04-0	+	–
disodium *S*,*S*′-hexane-1,6-diyldi(thiosulfate) dihydrate	401-320-7		+	–
methyl α-((4,6-dimethoxypyrimidin-2-yl)ureidosulfonyl)-*o*-toluate	401-340-6	83055-99-6	+	–
(benzothiazol-2-ylthio)succinic acid	401-450-4	95154-01-1	+	+
dimethyl-2-(3-(4-chlorophenyl)-4,5-dihydropyrazol-1-ylphenylsulfonyl)ethylamine	401-410-6	10357-99-0	+	–
lithium sodium hydrogen 4-amino-6-(5-(5-chloro-2,6-difluoropyrimidin-4-ylamino)-2-sulfonatophenylazo)-5-hydroxy-3-(4-(2-(sulfonatooxy)ethylsulfonyl)phenylazo)naphthalene-2,7-disulfonate	401-560-2	108624-00-6	+	+
1,3-dichloro-5-ethyl-5-methylimidazolidine-2,4-dione	401-570-7	89415-87-2	+	+
cyclooct-4-en-1-yl methyl carbonate	401-620-8	87731-18-8	+	–
tin(II) methanesulfonate	401-640-7	53408-94-9	+	–
hexasodium 7-(4-(4-(4-(2,5-disulfonatoanilino)-6-fluoro-1,3,5-triazin-2-ylamino)-2-methylphenylazo)-7-sulfonatonaphthylazo)naphthalene-1,3,5-trisulfonate	401-650-1	85665-96-9	+	+
N,*N*′-(2,2-dimethylpropylidene)hexamethylenediamine	401-660-6	1000-78-8	+	–
3,5-dichloro-2,4-difluorobenzoyl fluoride	401-800-6	101513-70-6	+	+
1,1′-(methylenebis(4,1-phenylene)dipyrrole-2,5-dione AND *N*-(4-(4-(2,5-dioxopyrrol-1-yl)benzyl)-phenyl)acetamide AND 1-(4-(4-(5-oxo-2*H*-2-furylidenamino)benzylphenyl)pyrrole-2,5-dione	401-970-1		+	+
N-hexadecyl(or octadecyl)-*N*-hexadecyl(or octadecyl)benzamide	401-980-6		+	–
N,*N*′,*N*″,*N*‴-tetrakis(4,6-bis(butyl(*N*-methyl-2,2,6,6-tetramethylpiperidin-4-yl)amino)triazin-2-yl)-4,7-diazadecane-1,10-diamine	401-990-0	106990-43-6	+	–
O-ethylhydroxylamine	402-030-3	624-86-2	+	–
ethyl 3-sulfamoyl-2-thenoate	402-050-2		+	+
potassium sodium 5-(4-chloro-6-(*N*-(4-(4-chloro-6-(5-hydroxyl-2,7-disulfonato-6-(2-sulfonatophenylazo)-4-naphthylamino)-1,3,5-triazin-2-ylamino)phenyl-*N*-methyl)amino-1,3,5-triazin-2-ylamino)-4-hydroxy-3-(2-sulfonatophenylazo)naphthalene-2,7-disulfonate	402-150-6		+	+
trisodium 7-(4-(6-fluoro-4-(2-(2-vinylsulfonylethoxy)ethylamino)-1,3,5-triazin-2-ylamino)-2-ureidophenylazo)-naphthalene-1,3,6-trisuphonate	402-170-5	106359-91-5	+	+
4-(2-chloro-4-trifluoromethyl)phenoxy-2-fluoroaniline hydrochloride	402-190-4		+	+
(bromobenzyl)bromotoluene	402-210-1	99688-47-8	+	+
S-(3-trimethoxysilyl)propyl 19-isocyanato-11-(6-isocyanatohexyl)-10,12-dioxo-2,9,11,13-tetraazanonadecanethioate	402-290-8	85702-90-5	+	+
2-(3-(prop-1-en-2-yl)phenyl)prop-2-yl isocyanate	402-440-2	2094-99-7	+	+
(C_{16} or C_{18}-*n*-alkyl)(C_{16} or C_{18}-*n*-alkyl)ammonium 2-((C_{16} or C_{18}-*n*-alkyl)(C_{16} or C_{18}-*n*-alkyl)-carbamoyl)-benzenesulfonate	402-460-1		+	–
4-(1(or 4 or 5 or 6)-methyl-8,9,10-trinorborn-5-en-2-yl)pyridine, mixture of isomers	402-520-7		+	–
3-(bis(2-ethylhexyl)aminomethyl)benzothiazole-2(3*H*)-thione	402-540-6	105254-85-1	+	+
ethyl *trans*-3-dimethylaminoacrylate	402-650-4	924-99-2	+	+
methyl-4-phenylpentanol	402-770-7	92585-24-5	+	–
3,7-dichloroquinoline-8-carboxylic acid	402-780-1	84087-01-4	+	–
valinamide	402-840-7	20108-78-5	+	–
hydrogen sodium *N*-carboxylatoethyl-*N*-octadec-9-enylmaleamate	402-970-4		+	+
potassium μ-fluorobis(triethylaluminium)	400-040-2	12091-08-6	–	–
ammonium bis(1-(3,5-dinitro-2-oxiphenylazo)-3-(*N*-phenylcarbamoyl-2-naphtholato)chromate(1–)	400-110-2		–	–
pentasodium 5-anilino-3-(4-(4-(6-chloro-4-(3-sulfonatoanilino)-1,3,5-triazin-2-ylamino)-2,5-dimethylphenylazo)-2,5-disulfonatophenylazo)-4-hydroxynaphthalene-2,7-disulfonate	400-120-7		–	+
tetrasodium 5′(4,6-dichloro-5-cyanopyrimidin-2-ylamino)-4′-hydroxy-2,3′-azodinaphthalene-1,2′,5,7′-disulfonate	400-130-1		–	–
fatty acids, tall-oil, reaction products with iminodiethanol and boric acid	400-160-5		–	–
disodium 1-amino-4-(4-benzenesulfonamido-3-sulfonatoanilino)anthraquinone-2-sulfonate	400-350-8	85153-93-1	–	+
7,7-dimethyl-3-oxa-6-azaoctan-1-ol	400-390-6		–	–
ethylenediammonium *O*,*O*-bis(octyl)phosphorodithioate, mixed isomers	400-520-1		–	–
disodium 6-(2,4-dihydroxyphenylazo)-3-(4-(4-(2,4-dihydroxyphenylazo)anilino)-3-sulfonatophenylazo)-4-hydroxynaphthalene-2-sulfonate AND disodium 6-(2,4-diaminophenylazo)-3-(4-(4-(2,4-diaminophenylazo)anilino)-3-sulfonaphenylazo)-4-hydroxynaphthalene-2-sulfonate AND trisodium 6-(2,4-dihydroxyphenylazo)-3-(4-(4-(7-(2,4-dihydroxyphenylazo)-1-hydroxy-3-sulfonato-2-naphthylazo)anilino)-3-sulfonatophenylazo)-4-hydroxynaphthalene-2-sulfonate AND trisodium 6-(2,4-dihydroxyphenylazo)-3-(4-(4-(7-(2,4-dihydroxyphenylazo)-1-hydroxy-3-sulfonato-2-naphthylazo)anilino)-3-sulfonatophenylazo)-4-hydroxynaphthalene-2-sulfonate	400-570-4		–	–
dodecyl 3-(2,2,4,4-tetramethyl-21-oxo-7-oxa-3,20-diazadispiro(5.1.11.2)henicosan-20-yl)propionate AND tetradecyl.3-(2,2,4,4-tetramethyl-21-7-oxo-7-oxa-3,20-diazadispiro(5.1.11.2)henicosan-20-yl) propionate	400-580-9	85099-51-0	–	–
calcium 2,5-dichloro-4-(4-((5-chloro-4-methyl-2-sulfonatophenyl)azo)-5-hydroxy-3-methylpyrazol-1-yl) benzenesulfonate	400-710-4		–	–
5(or 6)-*tert*-butyl-2′-chloro-6′-ethylamino-3′,7′-dimethylspiro(isobenzofuran-1(1*H*),9′-xanthene)-3-one	400-680-2		–	+

Table 5 (Continued)

name	ELINCS no.	CAS no.	skin sensitizer	alert
trisodium bis(7-acetamido-2-(4-nitro-2-oxidophenylazo)-3-sulfonato-1-naphtholato)chromate(1-)	400-810-8		-	-
2,2-dimethyl-1,3-benzodioxol-4-ol	400-900-7	22961-82-6	-	-
2,4-dichloro-3-ethylphenol	401-060-4		-	-
butyl (dialkyloxy)dibutoxyphosphoryloxy))titanium(trialkyloxy) titanium phosphate	401-100-0		-	-
diethyl(ethyldimethylsilanolato)aluminium	401-160-8		-	-
sodium (1-(5-(4-(4-anilino-3-sulfophenylazo)-2-methyl-5-methylsulfonamidophenylazo)-4-hydroxy-2-oxido-3-(phenylazo)phenylazo)-5-nitro-4-sulfonato-2-naphtholato)iron(II)	401-220-3		-	-
(tris(chloromethyl)phthalocyaninato)copper(II), reaction products with *N*-methylpiperazine and methoxyacetic acid	401-260-1		-	-
bis(4-fluorophenyl)methyl(1,2,4-triazol-4-ylmethyl)silane hydrochloride	401-380-4		-	-
3-(3-methylpent-3-yl)isoxazol-5-ylamine	401-460-9	82560-06-3	-	-
potassium 2-hydroxycarbazole-1-carboxylate	401-630-2	96566-70-0	-	-
3-chloro-5-trifluoromethyl-2-pyridylamine	401-670-0	79456-26-1	-	-
4,4′-isobutylethylidenediphenol	401-720-1	6807-17-6	-	-
S-benzyl *N*,*N*-dipropylthiocarbamate	401-730-6	5288-80-9	-	-
2-chloroethyl chloropropyl 2-chloroethylphosphonate, mixture of isomers, AND 2-chloroethyl chloropropyl 2-chloropropylphosphonate, mixture of isomers	401-740-0		-	-
lead(II) methanesulfonate	401-750-5	17570-76-2	-	-
diethyl 2,4-dihydroxycyclodisiloxane-2,4-diylbis(trimethylene) diphosphonate, tetrasodium salt, reaction products with disodium metasilicate	401-770-4		-	-
3,5-dichloro-4-(1,1,2,2-tetrafluoroethoxy)aniline	401-790-3	104147-32-2	-	+
(*N*-benzyl-*N*-ethyl)amino-3′-hydroxyacetophenone hydrochloride	401-840-4	55845-90-4	-	-
sodium 3,5-dichloro-2-(5-cyano-2,6-bis(3-hydroxypropylamino)-4-methylpyridin-3-ylazo)-benzenesulfonate	401-870-8		-	-
3-chloro-4,5-α,α,α-pentafluorotoluene	401-930-3	77277-99-7	-	-
5-heptyl-1,2,4-triazol-3-ylamine AND 5-nonyl-1,2,4-triazol-3-ylamine	401-940-8		-	-
3-(dimethylamino)propylurea	401-950-2	31506-43-1	-	-
2-(2-bromoethoxy)anisole	402-010-4	4463-59-6	-	-
calcium octadecylxylenesulfonate	402-040-8		-	-
pentyl methylphosphinate AND 2-methylbutyl methylphosphinate	402-090-0	87025-52-3	-	-
benzyltributylammonium 4-hydroxynaphthalene-1-sulfonate	402-240-5	102561-46-6	-	-
2-(4,4-dimethyl-2,5-dioxooxazolidin-1-yl)-2′-chloro-5′-(2-(2,4-di-*tert*-pentylphenoxy)butyramido)-4,4-dimethyl-3-oxovaleranilide	402-260-4		-	-
zinc 2-hydroxy-5-C_{13-18}-alkylbenzoate	402-280-3		-	-
(ethyl-3-oxobutanoato-O_1,O'_3)(2-dimethylaminoethanolato)(1-methoxypropan-2-olato)-aluminium(III), dimerized	402-370-2		-	-
oxo-4-isopropyl-1-methyl-1,4-epoxycyclohexan-2-ol	402-470-6	107133-87-9	-	-
2-(4-(3-(4-chlorophenyl)-4,5-dihydropyrazolyl)phenylsulfonyl)ethyldimethylammonium hydrogen phosphonate	402-490-5	106359-93-7	-	-
trisodium (6-anilino-2-(5-nitro-2-oxidophenylazo)-3-sulfonato-1-naphtholato)(4-sulfonato-1,1′-azodi-2,2′-naphtholato)chromate(1-)	402-500-8		-	-
2-methyl-1-pentylpyridinium bromide	402-690-2		-	-
benzyl-2-hydroxydodecyldimethylammonium benzoate	402-610-6	113694-52-3	-	-
1-butyl-2-methylpyridinium bromide	402-680-8	26576-84-1	-	-
triethoxyisobutylsilane	402-810-3	17980-47-1	-	-
trisodium bis(2-(5-chloro-4-nitro-2-oxidophenylazo)-5-sulfonato-1-naphtholato)chromate(1-)	402-870-0	93952-24-0	-	-
bis(2,2,6,6-tetramethyl-4-piperidyl)succinate	402-940-0	62782-03-0	-	-

functional groups which can be represented by predefined fragments) are then written in CHMTRAN. The rules also contain comments such as literature references, rule date, etc., which can be accessed, if required, by the user. The system would be enhanced by the incorporation of knowledge of skin metabolism and nonmetabolic transformations and experimental or calculated physicochemical information for the prediction of skin penetration. The system could then perform a useful role in the initial stages of hazard assessment of industrial chemicals.

Additional information may be provided by substructure or similarity searching of large, toxicological skin sensitization databases using a substructure representing a structural alert or a substance containing the structural alert. Examination of similar compounds retrieved in the search, by an expert toxicologist and/or use of a set of rules for the application of structure-based analogy derived from existing knowledge, may generate additional knowledge. For example, the effect of interaction between other functional groups and the chosen structural alert on the skin sensitization potential may be examined. Alternatively, similarity searching can be performed using fragment descriptors developed for structure–activity relationships.[29] These descriptors are derived from the computation of links between atoms comprised of a pair of terminators and the distance between them expressed as the number of bonds.

ACKNOWLEDGMENT

We wish to thank Dr. Martin Barratt, Mr. David Basketter, and Dr. Mark Chamberlain of Unilever Research, Colworth, U.K., Dr. Jean-Pierre Lepoittevin of Université Louis Pasteur, Strasbourg, France; and Dr. Mark Cronin of Liverpool John Moores University, U.K., for useful discussions, encouragement and the provision of toxicological data.

APPENDIX I

Substances used in the assessment of structural alerts are given in Table 5.

REFERENCES AND NOTES

(1) Ashby, J.; Tennant, R. W. Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. *Mutat. Res.* **1988**, *204*, 17–115.

(2) Rycroft, R. J. G In *Textbook of Contact Dermatitis*; Rycroft, R. J. G., Menne, T., Frosch, P. J., Benezra, C., Eds.; Springer Verlag: Berlin, 1992; Chapter 11, pp 314–399.

(3) Leopittevin, J.-P.; Benezra, C. Chimie de l'allergie de contact. Pourquoi une molecule est-elle allergisante? (Chemistry of Contact Allergy. Why Is a Molecule Allergenic?) *Rev. Fr. Allergol.* **1991**, *31* (4), 235–241.

(4) Dupuis, G.; Benezra, C. *Allergic Contact Dermatitis to Simple Chemicals. A Molecular Approach*; Dekker: New York, 1982.
(5) Sanderson, D. M.; Earnshaw, C. G. Computer Prediction of Possible Toxic Action from Chemical Structure; The DEREK System. *Hum. Exp. Toxicol.* **1991**, *10*, 261–273.
(6) *DEREK Collaborative Group*, c/o LHASA UK, University of Leeds, Leeds LS2 9JT, U.K.
(7) Berner, G.; Cooper, E. R. In *Transdermal Delivery of Drugs*; Kydonieus, A. F., Berner, B., Eds.; CRC Press: Boca Raton, FL, 1987; Vol. II, pp 45–55.
(8) Roitt, I. M. *Essential Immunology*; Blackwell: Oxford, U.K. 1991; p 65.
(9) Wong, S. S. *Chemistry of protein conjugation and cross-linking*; CRC Press: Boca Raton, FL, 1991.
(10) Brinkley, M. *Teaching Editorial. A brief survey of methods for preparing conjugates with dyes, haptens and cross-linking reagents*. Bioconjugate Chem. **1992**, *3*, 2–13.
(11) Lepoittevin, J.-P.; Benezra, C. ^{13}C-enriched methyl alkanesulfonates: new lipophilic methylating agents for the identification of nucleophilic amino acids by NMR. *Tetrahedron Lett.* **1992**, *33*, 3875–3878.
(12) Wahlberg, J. E.; Boman, A. Guinea pig maximisation test. *Curr. Probl. Dermatol.* **1985**, *14*, 59–106.
(13) Payne, M. P.; Walsh, P. T.; Langowski, J. J. Skin Sensitisation structure-activity relationships for phenols and anilines and application of a qualitative rule-based system. In *Trends in QSAR and Molecular Modelling 92*; Wermuth, C. G., Ed.; ESCOM: Leiden, The Netherlands, 1993; pp 504–506.
(14) Schmidt, R. J.; Khan, L.; Chung, L. Y. Are free radicals and not quinones the haptenic species derived from urushiols and other contact allergenic mono- and dihydric alkylbenzenes? The significance of NADH, glutathione and redox cycling in the skin. *Arch. Dermatol. Res.* **1990**, *282*, 56–64.
(15) Barratt, M. D.; Basketter, D. A. Possible origin of the skin sensitisation potential of isoeugenol and related compounds. *Contact Dermatitis* **1992**, *27*, 98–104.
(16) Zbaida, S.; Stoddart, A. M.; Levine, W. G. Studies on the mechanism of reduction of azo dye carcinogen by rat liver microsomal cytochrome P-450. *Chem.-Biol. Interact.* **1989**, *69*, 61–71.
(17) Kasting, G. B.; Smith, R. L.; Cooper, E. R. Effect of lipid solubility and molecular size on percutaneous absorption. *Pharmacol. Skin.*, **1987**, *1*, 138–153.
(18) Ridout, G.; Houk, J.; Guy, R. H.; Santus, G. C.; Hadgraft, J.; Hall, L. L. An evaluation of structure-penetration relationships in percutaneous absorption. *Farmaco* **1992**, *47* (6), 869–892.
(19) Siegenthaler, U.; Laine, A.; Polak, L. Studies on contact sensitivity to chromium in the guinea pig. The role of valence in the formation of the antigenic determinant. *J. Invest. Dermatol.* **1983**, *80*, 44–47.
(20) Paustenbach, D. P.; Sheehan, P. J.; Paull, J. M.; Wisser, L. M.; Finley, B. L. Review of the allergic contact dermatitis hazard posed by chromium-contaminated soil: Identifying a safe concentration. *J. Toxicol. Environ. Health* **1982**, *37*, 177–207.
(21) Goodwin, B. F. J.; Johnson, A. W.; Roberts, D. W.; Williams, D. L. Relationships between chemical structure and skin sensitisation potential. In *Immunotoxicology*; Gibson, G. G., Hubbard, R., Parke, D. V., Eds.; Academic Press: London, 1983; pp 449–456.
(22) Basketter, D. A.; Roberts, D. W.; Cronin, M.; Scholes, E. W. The value of the local lymph node assay in quantitative structure-activity investigations. *Contact Dermatitis* **1992**, *27*, 137–142.
(23) March, J. *Advanced Organic Chemistry:Reactions, Mechanisms and Structure*; McGraw-Hill Kogakusha: Tokyo, 1977.
(24) Matsushita, T.; Aoyama, K. Cross reactions between some pesticides and the fungicide benomyl in contact allergy. *Ind. Health* **1981**, *19*, 77–83.
(25) Barratt, M. Personal communication, 1993.
(26) Basketter, D. A.; Scholes, E. W. Comparison of the local lymph node assay with the guinea-pig maximisation tests for the detection of a range of contact allergens. *Food Chem. Toxicol.* **1992**, *30*, 65–69.
(27) *Official Journal of the European Communities (1991)*; Annex 1 to EEC Commission Directive 91/325/EEC (March 1, 1991); Publications Office of EC:Luxembourg.
(28) Hopkinson, G. A. Computer-assisted organic synthesis design. Ph.D. Thesis, University of Leeds, U.K. 1985.
(29) Judson, P. N. Structural similarity searching using descriptors developed for structure-activity relationship studies. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 657–663.