From Knowledge Generation to Knowledge Archive. A General Strategy Using TOPS-MODE with DEREK To Formulate New Alerts for Skin Sensitization

Ernesto Estrada,*,† Grace Patlewicz,‡ and Yaquelin Gutierrez†

BioComputing, X-ray Unit, RIAIDT, Edificio CACTUS, University of Santiago de Compostela, Santiago de Compostela 15782, Spain, and Safety and Environmental Assurance Centre (SEAC), Unilever Colworth, Colworth House, Sharnbrook, Bedford MK 44 1LQ, UK

Received October 29, 2003

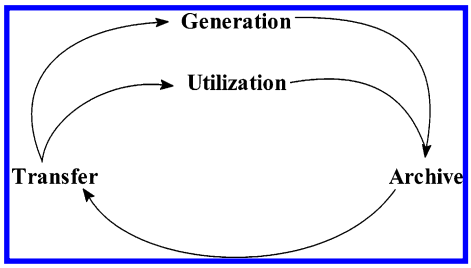
A general strategy for knowledge flow concerning skin sensitization based on the combined use of TOPS-MODE and DEREK expert system is proposed. TOPS-MODE is used as a knowledge generator, while DEREK represents the knowledge archive. A TOPS-MODE classification model allows the identification of structural fragments and groups responsible for strong/moderate skin sensitization. These structural contributions are sorted, analyzed, and graphically displayed in an appropriate way allowing the identification of several structural alerts for skin sensitization. Nine structural alerts already implemented in DEREK are identified using this strategy. They comprise, among others, alkyl halides, aldehydes, α,β-unsaturated compounds, aromatic amines, phenols, hydroquinone, isothiazolinone, and alkyl sulfonates. Four new hypotheses are generated using TOPS-MODE structural contributions to skin sensitization, which are not recognized as structural alerts by DEREK. They include the reduction of aromatic nitro groups and epoxidation reaction of double bonds as metabolic activation steps that can lead to reactive haptens which can trigger the skin sensitization mechanism. Another new alert is based on 1,2,5-thiadiazole-1,1-dioxide for which we have identified a possible mechanism explaining its strong skin sensitization profile. It is based on the existence of a tautomeric equilibrium and further reaction with nucleophiles, which are both supported by experimental evidence. Finally, we have identified a possible new mechanism for the skin sensitization of nonreactive compounds, which involves the formation of noncovalent complexes with proteins in a processing- and metabolism-independent way.

INTRODUCTION

The automatic prediction of molecular properties is a real necessity imposed by the huge amounts of data generated in both chemical and biological sciences. In these lines both pharmaceutical and consumer products industries are interested in evaluating thousands of chemicals in a fast and efficient manner. The prediction of toxicological properties of these chemicals is of particular importance due to many candidates failing or slowing down the development process as a consequence of toxicity problems or ambiguous toxicity findings. The importance of this aspect in molecular development is reflected in the existence of several *expert systems* for predicting toxicological properties of chemicals,[1] including DEREK[2] and TOPKAT.[3] Expert systems are artificial intelligence applications that use a *knowledge base* of human expertise to aid in solving problems,[4] such as deciding whether a chemical is toxic. In some respects this knowledge base is a *knowledge archive* where a collection of knowledge is expressed using some formal representation language, e.g., as rules. It preserves knowledge and allows it to remain intact once introduced into the system.

As new chemical entities are generated more rules need to be created to incorporate new knowledge about the toxicity profiles of novel structural features. In other words, a *knowledge generator* is needed that will provide new *structural alerts* to the knowledge archives in a cyclic manner that will keep the system updated for new challenges. This generator comprises activities associated with the entry of new knowledge into the system. Knowledge generation and knowledge archive are linked together in the general knowledge model that also includes *knowledge transfer* and *utilization* (see Scheme 1).[5] Transfer refers to activities associated with the flow of knowledge, while utilization includes activities and events connected with the application of this knowledge.

Scheme 1



At this point it is of great importance to know how to create this new knowledge from the data generated for new chemical entities. These data consist of discrete, objective facts about (toxicity) mechanisms or endpoint values. These data are sorted, analyzed, and displayed in a manner that enables their communication such that they are transformed

* Corresponding author fax: 34-981 547 077; e-mail: estrada66@yahoo.com.
† University of Santiago de Compostela.
‡ Safety and Environmental Assurance Centre (SEAC), Unilever Colworth.

10.1021/ci0342425 CCC: $27.50
Published on Web 01/31/2004

into *information*. However, knowledge still involves links between such information and its applications in such a way that it is closer to action than either information or data. We are accepting here that information is transformed into knowledge when it is in a useful form for making predictions about (toxicity) mechanisms and/or properties of new chemicals, which can be verified by experiment. Cheminformatics has been defined as the "mixing of information resources to transform data into information, and information into knowledge, for the intended purpose of making better decisions faster" in lead identification and optimization.[6] Consequently, the role of cheminformatics is to manage the data-information-knowledge pyramid as we propose here by combining TOPS-MODE as knowledge generator and DEREK as knowledge archive.

The topological substructural molecular design (TOPS-MODE) approach[7−13] has been applied to the study of skin sensitization of chemicals.[14] Using such a model we are able to generate data about groups and regions in skin sensitizing compounds that are responsible for such toxicity. These data, referred to here as *structural contributions* to skin sensitization, are transformed into information in this work by sorting, analyzing, and displaying it in an appropriate way. Using this information we formulate several hypotheses related to possible mechanisms of sensitization for such chemicals. These hypotheses can then be verified by experiment transforming this information into knowledge, which can then be converted into the form of rules for the expert system DEREK, which itself provides a useful form for transferring and utilizing knowledge. Several of the hypotheses generated here have been previously implemented into DEREK as rules, confirming the usefulness of TOPS-MODE in generating knowledge. Other new hypotheses are formulated in this work for the first time, and some of these still require experimental confirmation.

DEREK AND SKIN SENSITIZATION

The Deductive Estimation of Risk from Existing Knowledge (DEREK)[2] is an expert system for qualitative prediction of toxicity which has found great application in the chemical industry. Skin sensitization is among the endpoints predicted by DEREK.[15−18] The knowledge base for skin sensitization contains about 60 rules consisting of descriptions of molecular substructures that have been found in relation with this toxicity. This expert system has a user-friendly environment in which the user can introduce a chemical structure (query) using two-dimensional graphics or known chemical structure formats. The system identifies in the knowledge base those fragments which are present in the query and could be responsible for skin sensitization. This means that the user receives a prediction based on the knowledge contained in the knowledge base as well as an explanation and justification for the prediction made.

The rules implemented in DEREK for skin sensitization have the following appearance: **Rule 419** (Aldehydes): All compounds containing the structure R−CH(═O), where R is alkyl or aryl are skin sensitizers.

This rule is based on the knowledge generated from skin sensitization data for a significant amount of aldehydes. In fact, DEREK provides the user with several examples and explanations about skin sensitization data of these compounds. The rule can be rationalized from a chemical point of view on the basis of the knowledge that aldehydes can react with proteins via Schiff-base formation providing the hapten can trigger the sensitization mechanism. However, not all aldehydes can penetrate the skin to the same extent nor do they possess the same reactivity with amines to form Schiff-bases. Including more specific alerts will dramatically increase the number of rules in the knowledge base. Consequently, the nature of structural alerts and the number of rules in such an expert system are intimately related to each other. General structural alerts are formulated by using only a few simple rules, like rule 419 for aldehydes. However, more specific alerts need a greater number of rules. Thus a compromise needs to be reached between the generality of alerts and the number of rules.

TOPS-MODE AND SKIN SENSITIZERS CLASSIFICATION

The Topological Sub-Structural Molecular Design (TOPS-MODE) approach is based on the method of moments.[19] Our main modification consists of using the topological bond matrix (edge adjacency matrix) instead of the vertex adjacency matrix of the molecular graph. The second modification is the inclusion of bond weights in the main diagonal entries of the bond matrix in order to account for hydrophobic, electronic, and steric effects that could be involved in biological processes.[20−22] These bond weights are obtained from atomic contributions to partition coefficient,[23] polar surface,[24] and polarizability,[25] which are transformed into bond contributions to the same properties. The same approach is used to transform Gasteiger−Marsilli atomic charges,[26] van der Waals atomic radii,[27] and molar refraction[28] into bond weights. Bond weight for a bond (i,j), where i and j are the corresponding atoms linked together, is calculated as follows

$$w(i,j) = \frac{w_i}{\delta_i} + \frac{w_j}{\delta_j} \tag{1}$$

where w_i is the atomic weight and δ_i is the vertex degree of the corresponding atom. Using bond weighted matrices $\mathbf{B}(T)$, where T is the type of weight used (see below), we calculate spectral moments, which are defined as the trace of the kth power of the weighted bond matrix, i.e., the sum of the diagonal entries of the kth power of the weighted bond matrix

$$\mu_k^T = \mathbf{Tr}[\mathbf{B}(T)^k] = \sum_i b_{ii}(T)^k \tag{2}$$

where **Tr** is the trace of the matrix, k is the order of the spectral moment, $b_{ii}(T)$ are the diagonal entries of the weighted bond matrix, and T is the type of the bond weight: H (hydrophobic), PS (polar surface area), Pol (polarizability), vdW (van der Waals radii), Ch (Gasteiger−Marsilli charges), MR (molar refraction).

Skin sensitization potency was measured by performing the local lymph node assay (LLNA)[29−31] at the same experimental conditions for all the chemicals studied and then deriving the EC3 values. In this work the EC3 values are ranked qualitatively according to their potencies:[32] class 1 signified strong/moderate sensitizers (EC3 < 10%), class

2 nonstrong/moderate sensitizers, i.e., weak, extremely weak, and nonsensitizers (EC3 > 10%). A data set of 93 compounds was used to find a discriminant model applying TOPS-MODE approach. Spectral moments of the different types were used to discriminate between strong/moderate skin sensitizers from weak/extremely weak/nonsensitizers. The linear discriminant model obtained is given below[14]

$$\begin{aligned}\text{Class}(1) = {} & 1.331\mu_1^{\text{H}} - 0.00598\mu_4^{\text{H}} + 0.00781\mu_2^{\text{PS}} - \\ & 2.1366 \cdot 10^{-4}\mu_3^{\text{PS}} + 0.0755\mu_1^{\text{MR}} + 0.0319\mu_2^{\text{MR}} - \\ & 1.1133\mu_5^{\text{Pol}} - 2.3797\mu_1^{\text{Ch}} + 0.1547\mu_3^{\text{Ch}} + 0.00425\mu_6^{\text{Ch}} + \\ & 2.0932\mu_1^{\text{vdW}} - 0.8683\mu_2^{\text{vdW}} + 0.7954 \quad (3)\end{aligned}$$

$$\text{Wilks} - \lambda = 0.61 \quad F(12,63) = 3.39 \quad D^2 = 2.52 \quad p < 0.0007$$

where λ is the Wilks' statistics, D^2 is the squared Mahalanobis distance, and F is the Fisher ratio. This model was able to correctly classify 80% of the 75 compounds in the training set. Thirty-one of the 39 strong/moderate sensitizers were correctly predicted. Eight compounds were predicted to be "false negatives", i.e., compounds that were strong/moderate sensitizers but predicted to be weak/extremely weak/nonsensitizers (see Table 1). In a cross-validation experiment this model classified correctly 7 from 8 compounds for a 83.3% of good classification. This model was further tested by an external prediction set of 15 compounds, from which 13 were correctly classified and one was unclassified.

TOPS-MODE STRUCTURAL INTERPRETATION

The main advances of using the TOPS-MODE approach to study skin sensitization as compared with other approaches, such as those based on MOLCONN, CODESSA, TOPKAT, DEREK, etc., is 2-fold. On one hand, TOPS-MODE permits the development of quantitative structure−activity models in a similar way to those approaches using molecular descriptors, such as MOLCONN or CODESSA. On the other hand, it permits the interpretion of the results in terms of fragment contribution identifying structural alerts that can be responsible for the undesired activity in a similar way as DEREK does. The structural interpretation of the TOPS-MODE results is carried out by using the bond contributions to skin sensitization. Bond contributions to skin sensitization are calculated on the basis of the local spectral moments, which are defined as the diagonal entries of the different powers of the weighted bond matrix

$$\mu_k^T(i) = b_{ii}(T)^k \quad (4)$$

where $\mu_k^T(i)$ is the kth local moment of the bond i. It is straightforward to realize that total moments are the sum of the local moments. Consequently we can substitute total moments of the weighted bond matrix by their expressions in terms of local moments in the classification model for skin sensitization

$$\text{Class} = b_0 + \sum_k a_k\mu_k^T = b_0 + \sum_k\sum_i a_k\mu_k^T(i) \quad (5)$$

where b_0 is the intercept in the linear discriminant model. This expression can be rearranged into the following one:

$$\text{Class} = b_0 + \sum_k a_k\mu_k^T(1) + \sum_k a_k\mu_k^T(2) + \cdots + \sum_k a_k\mu_k^T(m) \quad (6)$$

Each one of the summations in expression (6) is the contribution of the corresponding bond to the skin sensitization of the studied compound. This means that a compound is classified as a skin sensitizer if the sum of its bond contributions to sensitization is greater than zero, i.e., Class > 0. These bond contributions can be represented as circles of radii proportional to the contribution. Two colors are used to differentiate positive and negative contributions. Here we use red circles for sensitizing contributions and blue circles for nonsensitizing ones. There is no a priori justification for the existence of bond additive schemes for this or any other property. However, it is well-known that several molecular descriptors and physicochemical parameters are currently expressed as a linear combination of bond contributions, such as polarizability, partition coefficients, molar refraction, etc. Thus, the use of any of these descriptors in a QSAR model permits represention of the studied property in a bond additive scheme in a similar way that the TOPS-MODE does.

STRUCTURAL ALERTS ALREADY EXISTING IN DEREK "FIRED" BY TOPS-MODE

Structural alerts consist of molecular fragments, which are known or believed to be responsible for the chemical reactivity component of the toxic action. The procedure that we will use for formulating the hypothesis of structural alerts with TOPS-MODE is as follows. First, we calculate bond contributions for all bonds in the skin sensitizers in our data set. Using positive contributions to skin sensitization we identified the groups, fragments, or regions that could be responsible for the toxic effect of such chemical. Then, we group these substructures into classes that allow some generalizations to be made about such fragments, e.g., functional groups. Finally, we check whether these groups could form haptens with proteins via known chemical reactions or after metabolism.

A. Alkyl Halides. This class of compounds is included in DEREK structural alert # 413. The carbon−halogen bond was identified in several alkyl halides present in our data set as contributing positively to strong/moderate skin sensitization. Most of the alkyl halides in this data set contain long carbon chains, which are identified by TOPS-MODE as contributing negatively to the skin sensitization. However, these long carbon chains could be of particular importance for the penetration of these compounds through the skin. This is observed by the fact that short chain alkyl halides are less sensitizing than longer chain ones. Nevertheless, the carbon−halogen bond is the only one in these compounds that can react with proteins forming the hapten via nucleophilic substitution of the halogen. In Figure 1 we provide the graphic representation of bond contributions in some alkyl halides included in the current study. As can be seen the contribution decreases in the order Br > Cl > I. Alkyl iodines included in the current data set are all weak sensitizers, which justifies the TOPS-MODE finding that the carbon−iodine bond has a negative contribution to strong/moderate sensitization.

Table 1. List of Compounds Used as Training, Cross-Validation, and External Prediction Sets and Skin Sensitization Classification According to TOPS-MODE

no.	compound	EC3	class	pred.[a]	prob.[b]	no.	compound	EC3	class	pred.[a]	prob.[b]
1	trimellitic anhydride	9.2	1	1	91.08	55	*cis*-6-nonenal	23.1	2	1	66.59
2	1-bromoeicosane	6.1	1	1	56.68	56	clotrimazole	4.8	1	1	97.11
3	1-bromododecane	15.47	2	1	75.92	57	cyclamen aldehyde	20.5	2	2	2.24
4	1-bromohexadecane	2.3	1	1	67.01	58	dibromodicyanobutane	2.3	1	1	99.67
5	1-bromooctadecane	16.6	2	1	61.98	59	diethyl maleate	4.7	1	U[c]	50.16
6	1-bromopentadecane	5.15	1	1	69.39	60	dihydroeugenol	12.45	2	2	13.20
7	1-bromotetradecane	9.2	1	1	71.68	61	dimethyl sulfoxide	71.9	2	2	18.37
8	1-bromotridecane	10.2	1	1[c]	73.85	62	dodecyl methane sulfonate	8.8	1	U	47.50
9	1-chloro-2,4-dinitrobenzene	0.1	1	1	97.33	63	ethylene glycol dimethacrylate	36.5	2	2	23.02
10	1-chlorohexadecane	9.1	1	2	34.88	64	eugenol	13.95	2	2	46.50
11	1-chlorooctadecane	16.3	2	2	30.07	65	farnesal	11.7	2	2[c]	0.75
12	1-chlorotetradecane	20.2	2	2	40.03	66	formaldehyde	1.2	1	1	76.74
13	1-iododoecane	13.1	2	2	4.12	67	glutaraldehyde	0.1	1	1	89.01
14	1-iodohexadecane	19.1	2	2	2.69	68	glyoxal	0.6	1	1	92.38
15	1-iodononane	24.2	2	2[c]	5.63	69	hc red no.3	2.2	1	1	94.81
16	1-naphthol	1.3	1	1	80.03	70	hexyl cinnamic aldehyde	12.54	2	2	31.87
17	2,3-butanedione	11.3	2	2	26.33	71	hydroquinone	0.1	1	1[c]	90.84
18	2,4-heptadienal	4	1	1	87.86	72	hydroxycitronellal	25.25	2	2	8.25
19	2-amino-6-chloro-4-nitrophenol	2.2	1	1	95.57	73	1-bromodocosane	8.3	1	U	51.22
20	2-ethyl butraldehyde	68.15	2	2	18.10	74	isoeugenol	3.5	1	2	40.91
21	2-methoxy-4-methyl-phenol	5.8	1	2	24.83	75	l(−)perillaldehyde	7.95	1	2	6.84
22	2-methyl-5-hydroxyethylaminophenol	0.4	1	2[c]	10.14	76	lactic acid	14.3	2	1	77.73
23	2-methyl-2H-isothiazol-3-one	1.9	1	1	90.37	77	linalool	30.4	2	2[c]	19.22
24	2-methylundecanal	10	1	2[c]	14.83	78	lyral	17.1	2	2	8.00
25	2-nitro-*p*-phenylenediamine	0.4	1	1	95.71	79	MPT	1.4	1	1	70.48
26	2-phenylpropionaldehyde (±)	6.3	1	U	50.19	80	o-aminophenol	0.5	1	1	90.72
27	3-bromomethyl-5,5-dimethyldihydro-2(3H)-furanone	3.6	1	1	77.21	81	oleyl methane sulfonate	25	2	U	49.87
						82	palmitoyl chloride	8.8	1	2	17.92
28	3-methyl eugenol	31.6	2	2	21.09	83	paraphenylenediamine	0.29	1	1[c]	85.58
29	3-methyl isoeugenol	3.5	1	2	17.68	84	phenyl benzoate	17.05	2	1	74.13
30	4-allylanisole	20.15	2	2	32.52	85	phenylacetaldehyde	4.7	1	1	80.92
31	4-chloroaniline	6.5	1	1[c]	79.08	86	pyridine	71.9	2	1	77.19
32	4-nitrobenzyl bromide	0	1	1	97.35	87	qrm 2113	36.8	2	2	9.35
33	5-methyl eugenol	13.2	2	2	16.75	88	safranal	7.5	1	2	27.86
34	6-methyl eugenol	16.9	2	2	17.31	89	tetradecyl iodide	13.8	2	2[c]	3.33
35	7- bromotetradecane	21.3	2	2	24.48	90	tetramethylthiuram disulfide	5.2	1	1	94.26
36	12-bromo-1-dodecanol	6.9	1	1	94.27	91	*trans*-2-decenal	2.5	1	1	67.22
37	12-bromododecanoic acid	7.6	1	1[c]	95.10	92	*trans*-2-hexenal	5.5	1	1	76.11
38	abietic acid	11.27	2	2	11.30	93	xylene	95.8	2	2	16.55
39	alpha amyl cinnamaldehyde	12.05	2	2	34.30	94	diethylamine	39.78	2	2[d]	25.93
40	alpha-butyl cinnamaldehyde	13.7	2	2	36.82	95	2-mercaptobenzothiazole	9.669	1	1[d]	73.82
41	alpha-methyl cinnamaldehyde	4.5	1	1	58.75	96	oxazolone	0.013	1	U[d]	52.26
42	bromohexane	10.3	1	1[c]	85.91	97	toluene diisocyanate	0.109	1	1[d]	97.07
43	bromoundecane	19.6	2	1	77.87	98	phthalic anhydride	0.357	1	1[d]	81.33
44	butyl glycidyl ether	30.9	2	2[c]	35.85	99	benzocaine	22.026	1/2[e]	1[d]	64.60
45	c4-azlactone	2.1	1	outlier	19.64	100	benzylidene acetone	<10	1	1[d]	71.94
46	c6-azlactone	1.3	1	outlier	16.40	101	diethylphthalate		2	2[d]	46.98
47	c9-azlactone	2.8	1	outlier	12.37	102	isopropyl myristate	25−50	2	2[d]	2.67
48	c11-azlactone	16.1	2	2	10.17	103	4-methoxyacetophenone	>50	2	2[d]	23.60
49	c15 azlactone	17.8	2	2	6.80	104	1-(p-methoxyphenyl)-1-penten-3-one	<10	1	2[d]	16.79
50	c17-azlactone	19	2	2	5.53						
51	c19-azlactone	26.4	2	2	4.49	105	5-methyl-2,3-hexanedione	25−50	2	2[d]	12.26
52	camphorquinone	10	1	1	95.47	106	3-propylidenephthalide	<5	1	1[d]	56.63
53	cinnamic alcohol	20.6	2	1[c]	89.87	107	melatonin	>10	2	2[d]	3.30
54	cinnamic aldehyde	2.2	1	1	92.05	108	nimesulide	>10	2	1[d]	97.62

[a] Predicted classification according to model (6). [b] A posteriori probability of being in group 1, i.e., a strong/moderate sensitizer. [c] Compound in the cross-validation set. [d] Compound in an external prediction set. [e] Classified previously as a moderate sensitizer (Basketter, D. A.; et al. *Contact Dermatitis* **1995**, *33*, 28−32).

B. Aldehydes. This class of compounds is included in DEREK structural alert # 419. There were several strong/moderate saturated aldehydes in our data set of skin sensitization. TOPS-MODE identifies the carbonyl group of such compounds as contributing positively to skin sensitization. Saturated aldehydes can react with proteins via Schiff-base formation in which some basic amino acids such as lysine can attack the carbonyl group with subsequent loss of a water molecule.[33] In Figure 2 we illustrate the bond contributions obtained by using TOPS-MODE for some of these compounds.

Despite DEREK identifying all carbonyl groups in saturated aldehydes as alerts for skin sensitization there are some aldehydes such as 2-ethyl butyraldehyde, **20**, that are not sensitizers. TOPS-MODE classifies this compound correctly as a nonstrong/moderate sensitizer, but it still recognizes the carbonyl group as an alert for skin sensitization. A similar situation is observed for compounds **57** and **87**, the first is

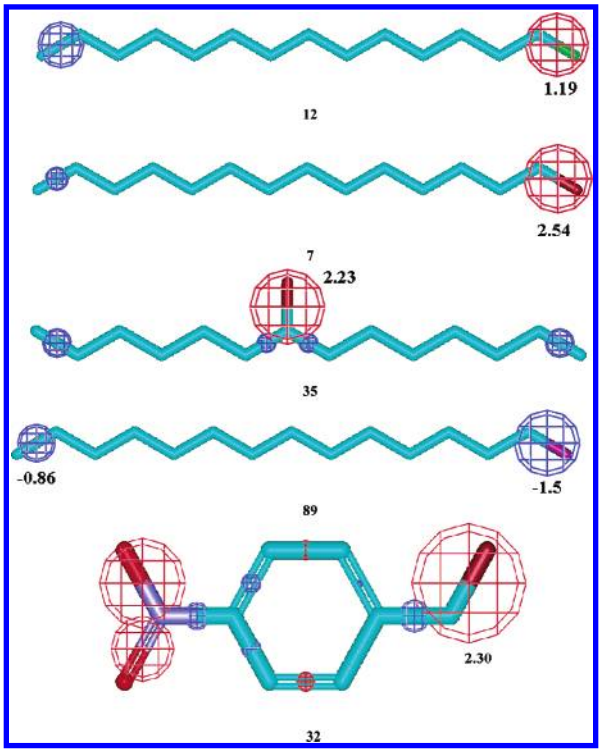


Figure 1. Structural contributions obtained from the TOPS-MODE classification model for some alkyl halides included in the current work.

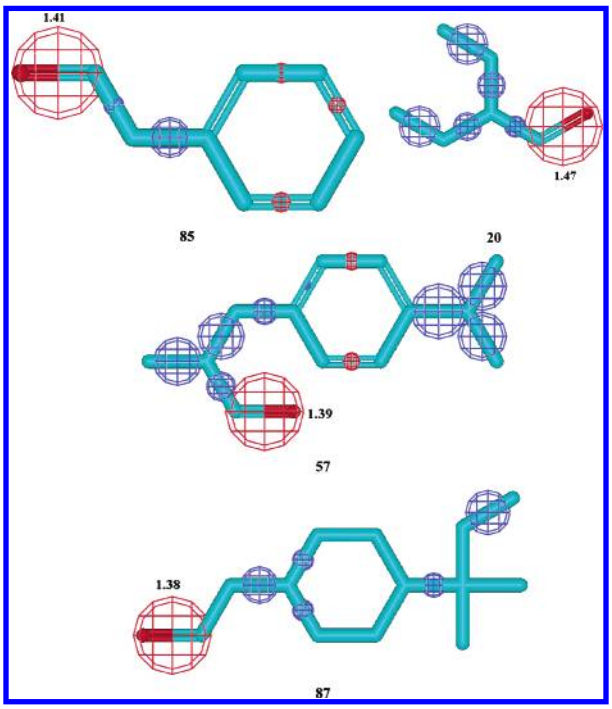


Figure 2. Structural contributions obtained from the TOPS-MODE classification model for some saturated aldehydes included in the current work.

a weak sensitizer and the second an extremely weak sensitizer, both posessing bulky groups at position 4 with respect to the chain having the carbonyl group. These bulky groups, which are not close to the reactive centers, are identified by TOPS-MODE as contributing negatively to skin sensitization. It is possible that in these cases the aldehydes form a noncovalent complex with proteins prior to forming the hapten, thus these groups may produce some steric

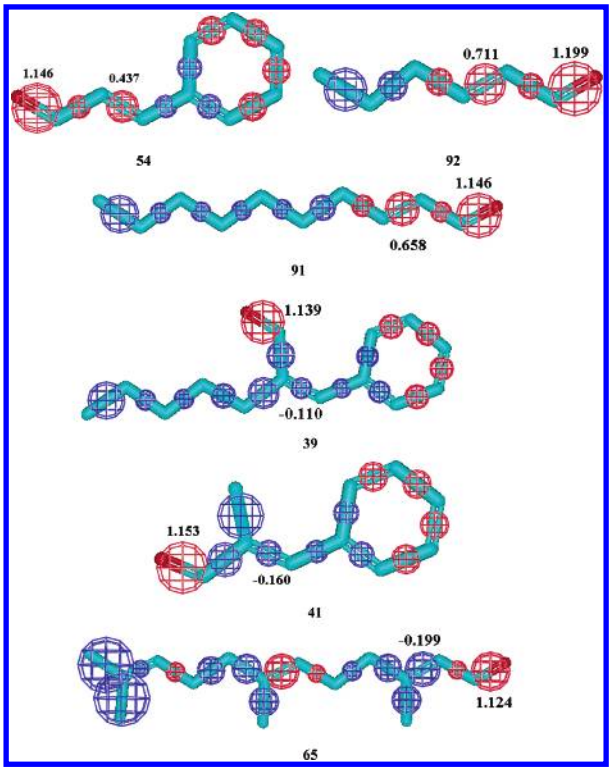


Figure 3. Structural contributions obtained from the TOPS-MODE classification model for some α,β-unsaturated aldehydes included in the current work.

hindrance for the reaction between amino groups of basic amino acids such as lysine and the carbonyl group.

C. α,β-Unsaturated Aldehydes, Amides, Esters, Ketones, Nitriles, and Nitro Compounds. This class of compounds is included in DEREK structural alert # 421. α,β-unsaturated aldehydes, amide, esters, ketones, nitrile, and nitro compounds can react with proteins by Michael addition of nucleophiles to the activated double bond.[34,35] It has been demonstrated in the case of α,β-unsaturated aldehydes, after the Michael addition a Schiff-base can also be formed by a new attack of the amine to the carbonyl group including some cyclization reactions. Our data set is well populated with α,β-unsaturated aldehydes. TOPS-MODE identifies both the double bond and the carbonyl group of these compounds as contributing to strong/moderate sensitization as seen in Figure 3 (see structures **54**, **91**, and **92**). Conjugated double bonds in structures **39**, **41**, and **65** are recognized by TOPS-MODE as contributing negatively to strong/moderate sensitization. The main difference between these compounds and other α,β-unsaturated aldehydes studied here is the presence of some steric hindrance on the conjugated double bond. This steric hindrance could be responsible for the lesser sensitizing potential of these compounds, e.g., compounds **39** and **65** are only weak sensitizers.

There are also two α,β-unsaturated esters in our data set: diethyl maleate and ethylene glycol dimethacrylate. The first is unclassified by TOPS-MODE as the percentage of classification as a strong/moderate sensitizer (50.16%) does not differ by more than 5% from that as a weak/extremely weak nonsensitizer (49.84%). TOPS-MODE is able to identify in both cases the double bond and the ester group as having positive contributions to the strong/moderate sensitization. Ethylene glycol dimethacrylate is, however, an extremely weak sensitizer with an EC3 value of 36.5%,

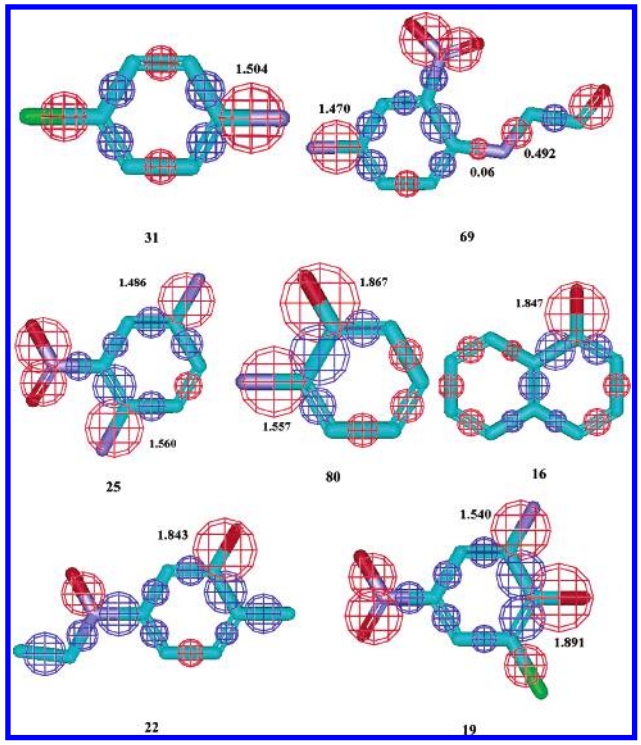


Figure 4. Structural contributions obtained from the TOPS-MODE classification model for some aromatic amines and phenols included in the current work.

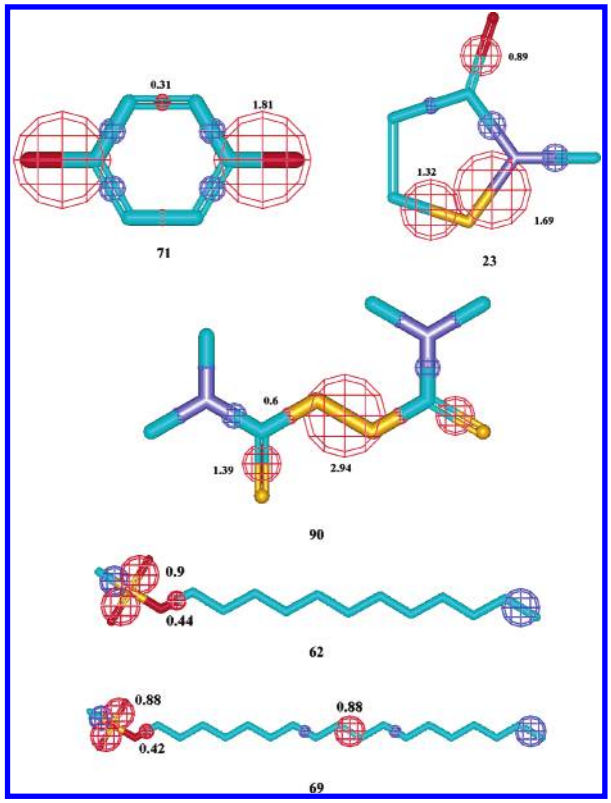


Figure 5. Structural contributions obtained from the TOPS-MODE classification model for some compounds of variable structures included in the current work, such as hydroquinone, isothiazolinone, thiuram disulfide, and alkyl sulfonates.

which is correctly classified by TOPS-MODE with a probability of 76.98% of being in the class of weak/extremely weak/nonsensitizers. Analyzing the bond contributions for this compound we can see a high negative contribution to sensitization for the methyl group directly bonded to the double bond (graphics not shown). This can be interpreted as an important steric effect hindering the double bond to a possible attack from a nucleophile in a Michael addition reaction.

It appears that the sensitization potential of some saturated aldehydes as well as α,β-unsaturated aldehydes and esters can be reduced by steric hindrance as observed by TOPS-MODE contributions and experimental observations. We do not consider it necessary to include a particular rule accounting for this specific fact to avoid an unnecessary proliferation of rules in expert systems such as DEREK. However, we do think that it could be convenient to include a comment accounting for the possible reduction of sensitization potential of chemicals due to the steric hindrance close to the reactive center like in carbonyl or conjugated double bonds of α,β-unsaturated aldehydes observed here.

D. Aromatic Primary or Secondary Amines. This class of compounds is included in DEREK structural alert # 427. It is believed that the sensitizing potential of aromatic amines depends on their transformation into reactive species which differ considerably in molecular size and chemical functionality from their precursors.[36] These metabolites can be formed via enzymatic transformation in the skin. One of these possible routes is the oxidation of amines to N-hydroxylamines which are then oxidized to nitroso compounds that can react with proteins through nucleophilic addition reaction.[36] TOPS-MODE identifies the amine group as contributing to strong/moderate skin sensitization and assigns them with high positive contributions as can be seen in Figure 4.

E. Phenols or Precursors. This class of compounds is included in DEREK structural alert # 439. There are several phenols reported in the literature as skin sensitizers. The DEREK system has a sensitization alert for phenols or phenol precursors. The alert description underlying the rule states that "skin sensitization for simple phenols is thought to arise as a result of formation of a phenolic radical which subsequently reacts with skin proteins" and that "whether the molecule will be a skin sensitizer will also depend on its percutaneous absorption". This last point is of particular interest. Although phenol itself is not a sensitizer,[37] other more lipophilic phenols such as 1-naphthol are strong sensitizers. There are several phenols in our data set that are all correctly classified by TOPS-MODE. In all cases TOPS-MODE recognizes the hydroxyl group as having a positive contribution to the skin sensitization as illustrated in Figure 4.

F. Hydroquinone or Precursor. This class of compounds is included in DEREK structural alert # 417. Hydroquinone is one of the compounds in our data set which is a strong/moderate sensitizer. TOPS-MODE recognizes both hydroxyl groups as well as both double bonds at positions 2,3 and 5,6 in the phenyl ring as having positive contributions to strong/moderate skin sensitization. This compound is believed to be easily transformed into 1,4-quinone by oxidation reaction. 1,4-Quinone can react with proteins via the Michael addition reaction in which nucleophile groups in amino acids, mainly cysteine sulfydryl group, attack the activated double bonds.[38] This kind of mechanism is consistent with the bond contributions obtained by TOPS-MODE for this compound which are illustrated in Figure 5, compound **71**.

G. Isothiazolinone. This class of compounds is included in DEREK structural alert # 434. Isothiazolinones form one particular structural alert in DEREK. Only recently the possible mechanisms that are involved in the skin sensitization of these compounds have been studied experimentally.[39] These compounds can react with nucleophile groups in proteins by a different set of reactions. The first step is apparently an attack of the nucleophile to the carbon in the double bond which is closer to the sulfur atom. Then the bond N−S is broken down to open the ring. TOPS-MODE classifies 2-methyl-2H-isothiazol-3-one as a strong/moderate sensitizer in full agreement with experimental evidence. It recognizes most of the ring bonds as possibly responsible for the skin sensitization with the N−S and S−C bonds having the highest contributions (see Figure 5, structure **23**).

H. Thiuram Mono or Disulfide. This class of compounds is included in DEREK structural alert # 442. Tetramethylthiuram disulfide is classified by TOPS-MODE as a strong/moderate sensitizer where the disulfide bond is identified as having a very high contribution to skin sensitization. The carbon−sulfur double bond is also identified as contributing to skin sensitization. Disulfides are believed to react with proteins through thiol exchange as a consequence of nucleophilic attack of the thiol or disulfide group on a protein disulfide bridge involving redox cycling between the thiol and the corresponding disulfide.[40] However, DEREK contains a specific rule for these compounds apart from the structural alert for thiol or thiol exchange agents. DEREK also contains a rule for thioesters, which is a fragment present in thiuram disulfide. As a consequence this compound cannot only react as a thiol exchange agent but additionally as a thioester by breaking the bond between the thiocarbonyl and the sulfur. This bond is also recognized by TOPS-MODE as contributing to skin sensitization (see Figure 5, structure **90**).

I. Alkyl Sulfate or Sulfonate. This class of compounds is included in DEREK structural alert # 414. There are two dialkyl sulfonates in our data set: oleyl methane sulfonate and dodecyl methane sulfonate. The first is only a weak sensitizer, while the second is a strong/moderate sensitizer. Since there was insufficient information in our data set about the sensitization potential of this class of compounds, TOPS-MODE is unable to classify them. In fact, TOPS-MODE leaves both compounds as unclassified. However, TOPS-MODE was able to identify both sulfonyl groups as having positive contributions to skin sensitization. The highest contributions are those of the S═O bonds as well as that for the carbon−oxygen bond that joins together the sulfonyl group with the carbon chain (see Figure 5, structures **62** and **69**). TOPS-MODE also highlights bond S−CH_3 which has a high negative contribution to skin sensitization. These contributions, especially the latter, are in agreement with the belief that sulfonyl compounds are sensitizers on account of their alkylating effect on proteins.[41,42] It is well-known that the ion metanosulfonate (mesilate), which is present in both compounds, is a good leaving group in nucleophilic reactions.[43] This ion can be formed by breaking the bond C−O, which has a high positive contribution to skin sensitization (see Figure 5). Finally, it is noted that dodecyl methane sulfonate, which is a strong/moderate sensitizer, contains an isolated double bond which is recognized by TOPS-MODE as having an important positive contribution to sensitization. The possible contribution of double bonds to sensitization will be considered in a further section of this work.

GENERATION OF NEW HYPOTHESES USING TOPS-MODE

The new hypotheses obtained with the help of TOPS-MODE bond contributions to skin sensitization are generated following an identical procedure as before. These new hypotheses can form the basis for new structural alerts after experimental confirmation. They are generated by considering those groups or fragments recognized by TOPS-MODE which contribute positively to skin sensitization. However, the TOPS-MODE identification of such groups alone is insufficient for formulating a new hypothesis, but the existence of some chemical basis justifying possible mechanisms for the formation of haptens is also required. The final requisite for the formulation of a new hypothesis is that it has not been previously contained as a structural alert in the DEREK expert system.

J. Reduction of Aromatic Nitro Compounds. It is well-known that nitro groups in aromatic compounds can be reduced in biological media to amines. The intermediates of this reduction reaction are nitroso compounds and N-hydroxylamines. The first is known as an important reactive intermediate that can react with proteins by the addition of nucleophilic groups in amino acids to the N═O bond. This reduction mechanism is well-known to produce biological activation for nitro aromatic compounds and is indeed related to both mutagenic and carcinogenic activity.

In our data set there are five structures containing aromatic nitro groups: 1-chloro-2,4-dinitrobenzene, 2-amino-6-chloro-4-nitrophenol, 2-nitro-*p*-phenylenediamine, 4-nitrobenzyl bromide, and HC Red No3. They are strong sensitizers with values of EC3 not greater than 2.2%. TOPS-MODE recognizes in all cases that the NO bonds of the nitro group have high contributions to skin sensitization (see for examples **32** in Figure 1 or **19**, **25**, and **69** in Figure 4). There is no simple rule in DEREK concerning aromatic nitro compounds besides those for activated benzene rings and musk ambrette. However, Payne and Walsh have proposed a particular rule recognizing the skin sensitization of trinitroaromatic compounds as a miscellaneous structural alert.[17]

The main point concerning this possible alert including nitroaromatic compounds is whether the skin has the potential to metabolize these compounds via reduction of nitro groups. However, it has been observed experimentally that 2-nitro-*p*-phenylendiamine[44,45] (**25** in Figure 4), known as "coal-tar dye" and used in semipermanent (nonoxidative) and permanent (oxidative) hair dye formulations, is reduced into the skin. 17% of triaminobenzene, the reduction final product, was detected in a recent in vitro study using fuzzy rat skin.[46] In human skin it was also observed the formation of triaminobenzene with quantities depending on the formulation used. Thus, the fact is that nitrobenzenes can be reduced into the skin, and as a consequence the skin sensitization potential of these chemicals, which has received great attention due to their mutagenic and carcinogenic properties, should be investigated.

K. Epoxidation of Double Bonds. Epoxidation is a well-known metabolic activation route for chemicals. Epoxides are also known to develop skin sensitization, and DEREK

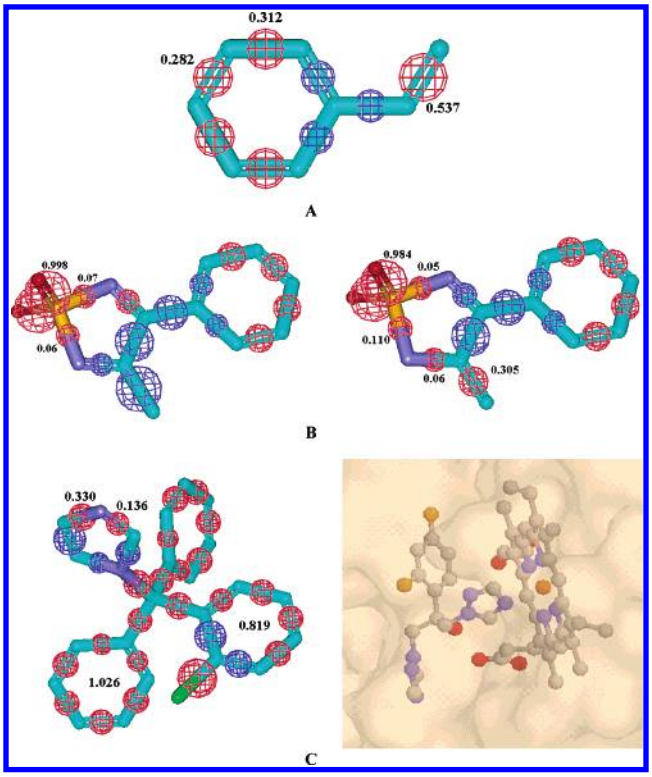
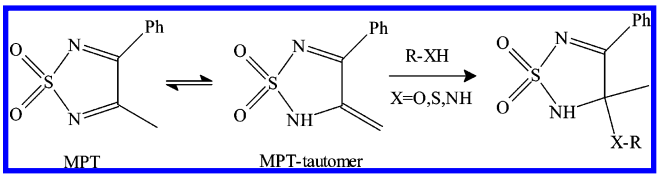


Figure 6. Structural contributions obtained from the TOPS-MODE classification model for (A) styrene; (B) 3-methyl-1,2,5-thiadiazole-1,1-dioxide (left) and its tautomer (right); (C) clotrimazole (left) and (right) a noncovalent complex between an azole antifungal compound and a protein (see text for explanation).

contains a particular rule for these chemicals. It is believed that epoxides are alkylating agents due to their reactivity with nucleophiles that can be added by breaking one of the C−O bonds of the epoxide. Epoxides can be formed from ethylenic structures in biological media.[47] TOPS-MODE has identified several double bonds not directly bonded to other groups that can activate them to Michael addition, as having a positive contribution to skin sensitization, e.g., **65** in Figure 3 and **69** in Figure 5. Unfortunately, all these compounds have groups different from the ethylenic one that can be responsible for the skin sensitization potentials of such compounds. To test whether an ethylenic group can be associated with skin sensitization we use TOPS-MODE to classify styrene as a skin sensitizer. This compound has no other functional group present apart from the ethylenic one that could be responsible for skin sensitization. Our model predicted styrene to be a strong/moderate sensitizer. The bond contributions for styrene show a high contribution to skin sensitization for the ethylenic double bond as shown in Figure 6A. This compound has been found to produce skin sensitization in humans using the patch test.[48] In this study it was concluded that styrene epoxide was the hapten due to the stronger reaction compared with styrene when tested in equimolar concentrations. According to these findings it is expected that styrene is the pro-hapten which has to undergo epoxidation in the skin before becoming a true hapten, i.e., the epoxide. There are other reports concerning the epoxidation of chemicals in the skin. For instance, aldrin, a pesticide having an isolated double bond in a cyclic system, has been reported to be epoxidated during skin absorption. According to this report more than 99% of the epoxidation product, dieldrin, was probably formed locally by dermal

Scheme 2



metabolism of percutaneously absorbed aldrin.[49−51] In a recent report styrene was found to produce skin sensitization in one Japanese worker of a fiber glass-reinforced plastic factory.[52]

Taking into account all these findings we propose to study the possible activation of double bonds via epoxidation in the skin. The structural features that this group requires in order to elicit skin sensitization need be explored to allow the formulation of the structural alert(s) concerning its skin sensitization.

L. 3-Methyl-1,2,5-thiadiazole-1,1-dioxide. There is another compound that has been reported as a strong/moderate sensitizer for which DEREK has no particular structural alert and flags nothing to report. 3-Methyl-4-phenyl-1,2,5-thiadiazole-1,1-dioxide (MPT) has an EC3 value of 1.4%, and it is well classified by TOPS-MODE with a probability of 70%. Figure 6B illustrates the bond contributions obtained from TOPS-MODE that identify several possible regions as responsible for the skin sensitization of this compound. It is known that this type of compound exists in a tautomeric equilibrium with the form given in Scheme 2. This tautomer can react via nucleophilic addition of nucleophiles such as alcohols and thiols to the C═N bond of the thiadiazole ring (see Scheme 2). For MPT, Caram et al. have found that the ethanol molecule does indeed add to the C═N bond positioned on the Me-substituted side of the substrate.[53−57]

TOPS-MODE predicts the tautomeric isomer of MPT to be a strong/moderate sensitizer with a probability of 90%. The contributions to the skin sensitization show that the SO_2 group has the largest positive values and the next highest contribution is for the $C{=}CH_2$ group (see Figure 6B), which indeed is expected to be attacked by nucleophiles (see Scheme 2). These findings indicate that compounds of the type of MPT can react with proteins by nucleophilic addition of thiol, amino, or hydroxyl groups to the $C{=}CH_2$ bond formed in one of the tautomers. This reactivity with nucleophiles is the necessary condition for the development of skin sensitization of these compounds. We propose to further investigate these types of compounds in order to confirm the hypothesis formulated here for the skin sensitization potential of MPT analogues.

M. Noncovalent Bond Hapten Formation. The generally accepted mechanism for skin sensitization consists of several stages in two main phases: induction and elicitation.[58] In the induction phase the allergen penetrates into the epidermis, where it is metabolized (if required) and then reacts with cellular proteins to form the antigen. This antigen is then recognized by specific T-cells, which finally proliferate and disseminate around the body through the blood. In the elicitation phase the antigen is presented to an enhanced population of specific T-cells in skin and then the production of local inflammatory response. The recognition of the hapten by T-cells in this mechanism necessarily implies covalent binding of the chemical or chemical metabolite to carrier proteins, which are presented as hapten-modified immuno-

genic peptides.[58] All the examples given in this work support this idea of the covalent binding between the chemical and a protein.

Clotrimazole is one compound with a strong/moderate sensitization potential which has been included in our data set. TOPS-MODE classifies this compound correctly. However, bond contributions to skin sensitization of this chemical as predicted by TOPS-MODE reveal some peculiar characteristics. As can be seen in Figure 6C most of the chemical bonds in clotrimazole are predicted to contribute to skin sensitization. Three main regions are clearly differentiated which are identified as contributing to skin sensitization. The first region is that formed by any (or both) of the nonsubstituted phenyl rings directly bonded to a quaternary carbon atom. The second region is that formed by some parts of the phenyl ring containing the chlorine atom, which is also recognized as contributing to the skin sensitization. Finally, the region around the nitrogen atom of the imidazole ring is also identified as contributing to the strong/moderate sensitization of clotrimazole.

The identification of almost all molecular regions in this compound as contributing to skin sensitization can be interpreted as the interaction of the whole molecule more than some specific parts of it with a biological system. This could be typical of noncovalent complexes between drugs and proteins or to the influence of general physicochemical properties on the biological response, such as adsorption or distribution. We advance here the hypothesis that clotrimazole could be acting by a noncovalent complex with proteins on the basis of some finding reported in the literature for this and analogous compounds. For instance, Penzotti et al. have developed a model for the pharmacophore of substrates interacting with P-glycoprotein.[59] According to this model two lipophilic regions separated by 7−9 Å, a hydrogen bond acceptor and a hydrogen bond donor, are necessary for the interaction with the protein in a similar way to the regions found in clotrimazole, which in fact it is a substrate and an inhibitor of P-glycoprotein. Another example is provided by the interaction of imidazole antifungal compounds with cytochrome P450, from which clotrimazole is a well-known inhibitor. In Figure 6C we illustrate the complex formed by the antifungal compound fluconazole and cytochrome P450 14 α-sterol demethylase (Cyp51) deposited in the Protein Data Bank (PDB) (PDB code 1EA1). The molecular graphic shows a region of the protein surface containing the pocket in which the ligand interacts with the prosthetic group of CYP, the ferric protoporphyrin IX. As can be seen in this figure, one of the nitrogens in the imidazole ring interacts with the ferric group of the protoporphyrin IX. This is the nitrogen atom identified by TOPS-MODE as one of those possibly responsible for skin sensitization. The other groups identified by TOPS-MODE are the phenyl rings, which in this figure are situated in hydrophobic cavities of the protein.

It is also known that most drugs, like clotrimazole, are not chemically reactive per se, and the metabolic activation was always postulated as the way to gain chemical reactivity before being presented to T-cells. However, Zanni et al.[60] demonstrated experimentally that this concept can be extended by showing that nonreactive drugs are recognized very rapidly in a processing- and metabolism-independent but still major histocompatibility complex (MHC) restricted way. These authors demonstrated that nonreactive drugs directly bind via noncovalent complexes to the MHC-peptide complex and the T cell receptor, resulting in T cell activation. This, of course, results in a new and revolutionary view of the mechanism of skin sensitization as it is not necessary a covalent bond between the chemical and a protein to be recognized by the T cells. This pharmacologic interaction of chemicals with immune receptors is called the "p-i concept".[61] The acceptance of this new view does not mean that the classical hapten model is not relevant as it remains the main way of presenting chemicals to T cells, but it does add the possibility of explaining why nonreactive chemicals such as clotrimazole can become immunogenic without the the formation of covalent complexes.

We propose to extend this study of skin sensitization using LLNA to other nonreactive chemicals that have been shown to activate T-cells via noncovalent complexes. It would also be interesting to demonstrate that clotrimazole can activate T cells and become immunogenic by forming a noncovalent complex with MHC-peptide complex. This demonstration will suppose new structural alert rules for these kinds of compounds, which necessarily should be further studied longer term by experiment and through using TOPS-MODE or other theoretical approaches.

DISCUSSION

In the first part of this work we have found several bond contributions for groups or fragments already contained in the expert system DEREK as structural alerts. This serves as a confirmation that these bond contributions generated by TOPS-MODE can be used to formulate structural alerts which can be conveniently deposited in knowledge archives such as DEREK. The quality of the predictions made by TOPS-MODE for bond contributions depends on the quality of the data set used to develop the model. There are several cases in which these bond contributions were generated from the data of several compounds, such as alkyl halides or aldehydes. In other cases, the bond contributions were obtained using very few examples in the data set. However, as TOPS-MODE makes a global description of the molecular structure it is expected that it is less dependent on the quantity of data than traditional methods used to generate a hypothesis.

As a matter of fact we give the example of how to generate the knowledge concerning the skin sensitization potential of thiuram mono or disulfide. Using the traditional method we have to collect a significant amount of data for these kinds of compounds. We can then extract the information to determine the structural pattern present in many skin sensitizers, and hence the structural alert may be formulated. TOPS-MODE, however, has been able to recognize this structural pattern from only one compound present in the data set. This is possible due to the nature of these descriptors. They describe the molecular structure as a whole in terms of hydrophobic, steric, and electronic characteristics of the molecules that can be transformed into local contributions. In this particular example TOPS-MODE is able to recognize the core of the structure which is responsible for skin sensitization leaving aside, for instance, the tertiary amine groups that are known to be inactive for skin sensitization.

There are some groups that have been recognized by TOPS-MODE as contributing positively to skin sensitization

and that are known to not have such activity. This is the case for instance of recognizing the carboxylic group as contributing positively to skin sensitization. It is known, however, that carboxylic acids are skin irritants but not sensitizers. Other examples are those of recognizing an aromatic carbon−chlorine bond in several compounds and cyano groups in dibromodicyanobutane as contributing positively to skin sensitization. There is no information available about the role of chlorine in aromatic rings concerning skin sensitization apart from that concerning activated phenyl rings. There is also no information concerning the role of cyano groups in skin sensitization, but some reports give this group as nonsensitizing. However, it is well-known that these and other groups can modify the sensitization potential of other groups. For instance, dibromodicyanobutane is a strong/moderate sensitizer alkyl halide even when it has a short chain. It is possible that the electron-withdrawing effect, as accounted for instance by Gasteiger−Marsilli charges, of the cyano groups adjacent to the bromines facilitate the nucleophilic substitution of them when reacting with proteins. The hydrophobic effect of halides such as chlorine in aromatic rings may also modify the sensitization potential of some compounds.

CONCLUSIONS

Decisions about whether a specific fragment should or should not be considered as a structural alert for skin sensitization or any other toxicological endpoint is a matter of human competence. However, identification of such fragments among the huge amount of data available today is a superhuman task. Theoretical and computational approaches can play a decisive role in transforming these data into valuable information as a first step to generate new knowledge. We have shown here that TOPS-MODE is a useful theoretical tool for identifying structural fragments and groups, which can be considered as structural alerts for skin sensitization. In the first part of this work we were able to identify several structural alerts that are well recognized in the literature and implemented in DEREK. In the second part we formulated four new hypotheses for skin sensitization structural alerts. One of these new hypotheses involves the formation of noncovalent complexes able to elicit sensitization. The confirmation of this hypothesis is now under study in one of our laboratories by testing on LLNA some nonreactive drugs known to bind noncovalently to the MHC-peptide complex. If this hypothesis is confirmed it implies a modification of the "classical" mechanism for skin sensitization of chemicals by introducing a new step involving the formation of noncovalent complexes with proteins.

Two other new hypotheses generated here with the help of TOPS-MODE are based on metabolic activation of chemicals previously known to interact with proteins. They are the reduction of nitro groups and the epoxidation of double bonds. For both cases there is wide experimental support confirming the metabolic activation of such groups in the skin. The fourth example analyzed in this work is a very interesting one. It is concerned with the skin sensitization of MPT, which is a strong sensitizer. This compound was the only one of this type present in the data set. TOPS-MODE has been able to classify it correctly and provide insights about the possible mechanism involved in the sensitization process. There is experimental support based on chemical facts that the tautomer of this compound can react with nucleophiles, which support our hypothesis for the mechanism of sensitization of this compound. This is an example of how the use of a theoretical approach like TOPS-MODE can save efforts and resources in the investigation of a toxicological endpoint. This is possible not only since TOPS-MODE allows the development of classification models but also it provides structural insights about the possible mechanisms involved. Charles A. Coulson has said, "Give us insights, not numbers". However, we have to recognize that it is possible to "Give insights through numbers" as shown in this work.

ACKNOWLEDGMENT

E.E. thanks Ministerio de Ciencia y Tecnología, Spain for a Ramón y Cajal contract as researcher at the University of Santiago de Compostela. E.E. also thanks Dr. D. Basketter, Dr. C. K. Smith Pease, and Dr. M. Chamberlain from SEAC, Unilever for useful discussions and encouragement.

REFERENCES AND NOTES

(1) Combes, R. D.; Judson, P. The use of artificial intelligence systems for predicting toxicity. *Pesticide Sci*. **1995**, *45*, 179−94.
(2) Sanderson, D. M.; Earnshaw, C. G. Computer prediction of possible toxic action from chemical structure; the DEREK system. *Human Exp. Toxicol*. **1991**, *10*, 261−73.
(3) Enslein, K.; Gombar, V. K.; Blake, B. W. Use of SAR in computer-assisted prediction of carcinogenicity and mutagenicity of chemicals by the TOPKAT program. *Mutat. Res*. **1994**, *305*, 47−61.
(4) Ignizio, J. P. *Introduction to Expert Systems: The Development and Implementation of Rule-Based Expert Systems*; McGraw-Hill: 1991.
(5) Newman, B. D.; Conrad, K. W. A framework for characterizing knowledge management methods, practices, and technologies. In *Proc. Third Int. Conf. On Practical Aspects of Knowledge Management* (*PAKM2000*); Reimer, U., Ed.; Basel, Switzerland, 2000.
(6) Hann, M.; Green, R. Chemoinformatics-a new name for an old problem? *Curr. Op. Chem. Biol*. **1999**, *3*, 379−383.
(7) Estrada, E.; Peña, A.; García-Domenech, R. Designing sedative/hypnotic compounds from a novel substructural graph-theoretical approach. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 583−595.
(8) Estrada, E.; Uriarte, E.; Montero, A.; Teijeira, M.; Santana, L.; De Clercq, E. A novel approach for the virtual screening and rational design of anticancer compounds. *J. Med. Chem*. **2000**, *43*, 1975−1985.
(9) Estrada, E.; Peña, A. In silico studies for the rational discovery of anticonvulsant compounds. *Bioorg. Med. Chem*. **2000**, *8*, 2755−2770.
(10) Estrada, E.; Uriarte, E. Quantitative structure−toxicity relationships using TOPS-MODE. 1. Nitrobenzene toxicity to *Tetrahymena pyriformis*. *SAR QSAR Environ. Res.* **2001**, *12*, 309−324.
(11) Estrada, E.; Molina, E.; Uriarte, E. Quantitative structure-toxicity relationships using TOPS-MODE. 2. Neurotoxicity of a noncongeneric series of solvents. *SAR QSAR Environ. Res.* **2001**, *12*, 445−459.
(12) Estrada, E.; Uriarte, E.; Gutierrez, Y.; Gonzalez, H. Quantitative structure-toxicity relationships using TOPS-MODE. 3. Structural factors influencing the permeability of commercial solvents through living human skin. *SAR QSAR Environ. Res.* **2003**, *14*, 145−163.
(13) Estrada, E.; Vilar, S.; Uriarte, E.; Gutierrez, Y. In silico studies toward the discovery of new anti-hiv nucleoside compounds with the use of TOPS-MODE and 2D/3D connectivity indices. 1. Pyrimidyl derivatives. *J. Chem. Inf. Comput. Sci*. **2002**, *42*, 1194−1203.
(14) Estrada, E.; Patlewicz, G.; Chamberlain, M.; Basketter, D.; Larbey, S. Computer-aided knowledge generation for understanding skin sensitization mechanisms: The TOPS-MODE approach. *Chem. Res. Tox*. **2003**, *16*, 1226−1235.
(15) Zinke, S.; Gerner, I.; Schlede, E. Evaluation of a rule base for identifying contact allergens by using a regulatory database: comparison of data on chemicals notified in the European Union with "Structural Alerts" used in the DEREK Expert System. *ATLA, Altern. Lab. Animals* **2002**, *30*, 285−298.
(16) Barratt, M. D.; Langowski, J. J. validation and subsequent development of the DEREK skin sensitization rulebase by analysis of the BgVV list of contact allergens. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 294−298.

(17) Payne, M. P.; Walsh, P. T. Structure−activity relationships for skin sensitization potential: Development of structural alerts for use in knowledge-based toxicity prediction systems. *J. Chem. Inf. Comput. Sci*. **1994**, *34*, 154−161.
(18) Barratt, M. D.; Basketter, D. A.; Chamberlain, M.; Payne, M. P.; Admans, G. D.; Langowski, J. J. Development of an expert system rulebase for identifying contact allergens. *Toxicol. Vitro* **1994**, *8*, 837−9.
(19) Burdett, J. K. Some structural problems examined using the method of moments. *Structure Bonding* **1987**, *65*, 29−90.
(20) Estrada, E. Spectral moments of the edge adjacency matrix in molecular graphs. 1. Definition and applications to the prediction of physical properties of alkanes. *J. Chem. Inf. Comput. Sci*. **1996**, *36*, 844−849.
(21) Estrada, E. Spectral moments of the edge-adjacency matrix of molecular graphs. 2. Molecules containing heteroatoms and QSAR applications. *J. Chem. Inf. Comput. Sci*. **1997**, *37*, 320−328.
(22) Estrada, E. Spectral moments of the edge-adjacency matrix of molecular graphs. 3. Molecules containing cycles. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 23−27.
(23) Wang, R.; Fu, Y.; Lai, L. A new atom-additive method for calculating partition coefficients. *J. Chem. Inf. Comput. Sci*. **1997**, *37*, 615−621.
(24) Ertl, P.; Rohde, B.; Selzer, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* **2000**, *43*, 3714−3717.
(25) Miller, K. J. Additivity methods in molecular polarizability. *J. Am. Chem. Soc*. **1990**, *112*, 8533−8542.
(26) Gasteiger, J.; Marsilli, M. A new model for calculating atomic charges in molcules. *Tetrahedron Lett*. **1978**, *34*, 3181−3184.
(27) Bondi, A. van der Waals volumes and radii. *J. Phys. Chem*. **1964**, *68*, 441−451.
(28) Ghose, A. K.; Crippen, G. M. Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure−activity relationships. 2. Modeling dispersive and hydrophobic interactions. *J. Chem. Inf. Comput. Sci*. **1987**, *27*, 21−35.
(29) Kimber, I.; Basketter, D. A.; Berthold, K.; Butler, M.; Garrigue, J.-L.; Lea, L.; Newsome, C.; Roggeband, R.; Steiling, W.; Stropp, G.; Waterman, S.; Wiemann, C. Skin sensitization testing in potency and risk assessment. *Toxicol. Sci.* **2001**, *59*, 198−208.
(30) Steiling, W.; Basketter, D.; Berthold, K.; Butler, M.; Garrigue, J. L.; Kimber, I.; Lea, L.; Newsome, C.; Roggeband, R.; Stropp, G.; Waterman, S.; Wiemann, C. Skin sensitization testing−new perspectives and recommendations. *Food Chem. Toxicol.* **2001**, *39*, 293−301
(31) Basketter, D. A.; Lea, L. J.; Dickens, A.; Briggs, D.; Pate, I.; Dearman, R. J.; Kimber, I. A comparison of statistical approaches to the derivation of EC3 values from local lymph node assay dose responses. *J. Appl. Toxicol.* **1999**, *19*, 261−266.
(32) Basketter, D. A.; Evans, P.; Gerberick, G. F.; Kimber, I. A. N. Factors affecting thresholds in allergic contact dermatitis: safety and regulatory considerations. *Contact Dermatitis* **2002**, *47*, 1−6.
(33) Patlewicz, G.; Roberts, D. W.; Walker, J. D. QSARs for the skin sensitization potential of aldehydes and related compounds. *QSAR Comb. Sci.* **2003**, *22*, 196−203.
(34) Meschkat, E.; Barratt, M. D.; Lepoittevin, J.-P. Studies of the chemical selectivity of hapten, reactivity, and skin sensitization potency. 2. NMR studies of the covalent binding of the ^{13}C-labeled skin sensitizers 2-[^{13}C]- and 3-[^{13}C]Hex-1-ene- and 3-[^{13}C]Hexane-1,3-sultones to human serum albumin. *Chem. Res. Tox*. **2001**, *14*, 118−126.
(35) Meschkat, E.; Barratt, M. D.; Lepoittevin, J.-P. Studies of the chemical selectivity of hapten, reactivity, and skin sensitization potency. 1. Synthesis and studies on the reactivity toward model nucleophiles of the ^{13}C-labeled skin sensitizers hex-1-ene- and hexane-1,3-sultones. *Chem. Res. Tox*. **2001**, *14*, 110−117.
(36) Boelsterli, U. A. *Mechanistic Toxicology*; Taylor & Francis: London, 2003; p 222.
(37) Basketter, D. A.; Lea, L.; Cooper, K.; Stocks, J.; Dickens, A.; Pate, I.; Dearman, R. J.; Kimber, I. Threshold for classification as a skin sensitiser in the local lymph node assay: a statistical evaluation. *Food Chem. Tox.* **1999**, *37*, 1167−1174.
(38) Bolton, J. L.; Trush, M. A.; Penning, T. M.; Dryhurst, G.; Monks, T. J. Role of quinines in toxicology. *Chem. Res. Tox*. **2000**, *13*, 135−160.
(39) Álvarez-Sánchez, R.; Basketter, D.; Pease, C.; Lepoittevin, J.-P. Studies of chemical selectivity of hapten, reactivity, and skin sensitisation potency. 3. Synthesis and studies on the reactivity toward model nucleophiles of the ^{13}C-labeled skin sensitizers, 5-chloro-2-methyl-isothiazol-3-one (MCI) and 2-methylisothiazol-3-one (MI). *Chem Res. Tox*. **2003**, *16*, 627−636.
(40) Munday, R. Toxicity of thiols and disulfides: involvement of free-radical species. *Free Radical Biol. Med.* **1989**, *7*, 659−673.
(41) Roberts, D. W.; Basketter, D. A. Quantitative structure−activity relationships: sulfonate esters in the local lymph node assay. *Contact Dermatitis* **2000**, 42, 154−161.
(42) Roberts, D. W.; Basketter, D. A. Further evaluation of the quantitative structure−activity relationships for skin-senstizing alkyl transfer agents. *Contact Dermatitis* **1997**, *37*, 107−112.
(43) Vollhardt, K. P. C. *Organic Chemistry*; W. H. Freeman and Co.: New York, 1987; pp 208−209.
(44) Elder, R. L. Final report on the safety assessment of 2-nitro-*p*-phenylenediamine and 4-nitro-*o*-phenylenediamine. *J. Am. Coll. Toxicol*. **1985**, *4*, 161−197.
(45) IARC. In *IARC Monographs on the Evaluation of the carcinogenic risk of chemicals to humans. 1,4-dimaino-2-nitro benzene (2-nitro-para-phenylenediamine)*; IARC, 57; Lyon, France, 1993.
(46) Yournick, J. J.; Bronaugh, R. L. Percutaneous penetration and metabolism of 2-nitro-*p*-phenylenediamine in human and fuzzy rat skin. *Toxicol. Appl. Pharmacol*. **2000**, *166*, 13−23.
(47) Smith Pease, C. K.; Hotchkiss, S. A. M. *Allergic Contact Dermatitis*; Taylor and Francis: London, 2001; p 132.
(48) Sjöborg, S.; Trulsson, F. Contact allergy to styrene and related chemicals *Contact Dermatitis* **1984**, 10, 94−96.
(49) Graham, M. J.; Williams, F. M.; Rawlins, M. D. Metabolism of aldrin to dieldrin by rat skin following topical application. *Food Chem. Toxicol*. **1991**, *29*, 707−711.
(50) MacPherson, S. E.; Scott, R. C.; Williams, F. M. Percutaneous absorption andmetabolism of aldrin by rat skin in diffusion cells. *Arch Toxicol*. **1991**, *65*, 599−602.
(51) Graham, M. J.; Williams, F. M.; Rawlins, M. D. Aldrin epoxidation by rat skin during percutaneous absorption in vivo. *Br. J. Clin. Pharmacol*. **1986**, *21*, 111.
(52) Minamoto, K.; Nagano, M.; Inaoka, T.; Futatsuka, M. Occupational dermatoses among fiberglass-reinforced plastic factory workers. *Contact Dermatitis* **2002**, *46*, 339−347.
(53) Aimone, S. L.; Caram, J. A.; Mirifico, M. V.; Vasini, E. J. A new type of imine alpha anion derived from 3-methyl-4-phenyl-1,2,5-thiadiazole 1,1-dioxide. *J. Phys. Org. Chem*. **2001**, *14*, 217−223.
(54) Caram, J. A.; Mirifico, M. V.; Vasini, E. J. Electrochemistry of 3,4-diphenyl-1,2,5-thiazole-1,1-dioxide (1) and its derivatives in ethanol-acteonitrile solutions and interactions of the dianion of (1) with metal cations. *Electroch. Acta* **1994**, *39*, 939−945.
(55) Caram, J. A.; Mirifico, M. V.; Aimone, S. L.; Vasini, E. J. 3,4-disubstituted derivatives of 1,2,5-thiadiazole-1,1-dioxide. Ethanol addition reactions and electro-reduction of 3-methyl-4-phenyl and 3,4-dimethyl derivatives in acetonitrile and ethanol solvents. *Can. J. Chem*. **1996**, *74*, 1564−1571.
(56) Rozas, M. F.; Piro, O. E.; Castellano, E. E.; Mirifico, M. V.; Vasini, E. J. Addition of aromatic nucleophiles to a C═N double bond of 1,2,5-thiazole-1,1-dioxide. *Molecules* **2000**, *5*, 503.
(57) Aimone, S. L.; Caram, J. A.; Mirifico, M. V.; Vasini, E. J. Electrochemical and spectroscopic study of the addition of several nucleophiles to 1,2,5-thiadiazole-1,1-dioxide derivatives. *J. Phys. Org. Chem*. **2000**, *13*, 272−282.
(58) Scheynus, A. 1997 Immunological apects. In:*Allergic Contact Dermatitis. The molecular basis*; Lepoittevin, J.-P., Basketter, D. A., Goossens, A., Karlberg, A.-T., Eds.; Springer: Berlin, pp 4−18.
(59) Penzotti, J. E.; Lamb, M. L.; Evensen, E.; Grootenhuis, D. J. A computational ensemble pharmacophore model for identifying substrates of P-glycoprotein. *J. Med. Chem*. **2002**, *45*, 1737−1740.
(60) Zanni, M. P.; Von Greyerz, S.; Brander, K. A.; Frutig, K.; Hari, Y.; Valitutti, S.; Pichler, W. J. HLA-restricted, processing- and metabolism-independent pathway of drug recognition by human $\alpha\beta$ lymphocytes. *J. Clin. Invest*. **1998**, *102*, 1591−1598.
(61) Pichler, W. J. Delayed drug hypersensitivity reactions. *Ann. Intern. Med*. **2003**, *139*, 683−693.

CI0342425