Review Article

The chemistry of contact allergy: why is a molecule allergenic?

David Basketter1, An Dooms-Gossens2, Ann-Therese Karlberg3 and Jean-Pierre Lepoittevin4

1Unilever Environmental Safety Laboratory, Colworth House, Sharnbrook, Bedford MK44 1LQ, UK
2Department of Dermatology, University Hospital, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium
3Department of Occupational Dermatology, National Institute of Occupational Health, S-17184 Solna, Sweden
4Laboratoire de Dermatochimie associé au CNRS, Université Louis Pasteur, Clinique Dermatologique, CHU, F-67091 Strasbourg, France

This review concentrates on some specific aspects of the chemistry of allergic contact dermatitis. The way low molecular weight chemicals react with skin proteins to form complete allergens will be discussed and the development of molecular modelling techniques to analyse molecular recognition presented. Subsequently, how knowledge of the chemical structure can be used to estimate the allergenic activity of a molecule will be considered. This aspect includes work with qualitative and quantitative structure-activity relationships (SAR) in the field of contact allergy.

Key words: allergic contact dermatitis; chemical basis; structure-activity relationships; hapten; prohapten; nucleophile; electrophile; skin metabolism; cross-sensitivity; molecular modelling.


Accepted for publication 19 July 1995.

Contact allergy is, without doubt, among the pathological conditions in which chemistry plays an especially important rôle. Chemical reactions and interactions are involved throughout the biological process that will result in the patient developing delayed hypersensitivity, whether it be during the crossing of the cutaneous barrier (mainly controlled by the physicochemical properties of the allergen), during the formation of the hapten-protein complex (in which chemical bonds are involved) or during the phenomena of recognition between the antigen and the receptors on T lymphocytes (involving a discipline undergoing rapid development, i.e., that of supramolecular chemistry). To help avoid the inappropriate use of new allergens, the understanding of what makes a molecule a hapten has to be expanded. In order to promote work in the field within the European Society of Contact Dermatitis, a working party on the chemical bases of allergic contact dermatitis has been formed. Its main tasks include the following:

(i) chemical mechanisms in allergic contact dermatitis;
(ii) prohapten and skin metabolism of contact allergens;
(iii) structure-activity relationship for contact allergens.

Recently, there has been a major step forward in our understanding of the molecular basis of hapten recognition by T cells. Nevertheless this does not eliminate the need to understand the characteristics of the preceding processes. Indeed, it is true to say that the properties of a chemical are implicit in its molecular structure. To cause sensitization, a compound has to penetrate the skin, where it may be metabolized, and react with (Langerhans) cell surface proteins to form new chemi-
cal structures that are recognized as foreign. This review concentrates on some specific aspects of the chemistry of this process.

We discuss the way low molecular weight chemicals react with skin proteins to form complete antigens and the molecular recognition. Subsequently, we consider how knowledge of the chemical structure can be used to estimate the allergenic activity of a molecule. This aspect includes work with qualitative and quantitative (Q) structure-activity relationships (SAR) in the field of contact allergy. In our view, 2 other issues are of importance. It is not sufficient to understand the potential reactivity of a chemical; its ability of partition in the appropriate skin compartment is also crucial to the elaboration of contact allergy. However, physicochemical principles governing skin penetration have been well reviewed elsewhere (1). The issue of skin metabolism is also important, but again this has been reviewed elsewhere (2).

Some Chemical Reminders

Chemical bonds represent electronic interactions between atoms. They can be of various strengths and are characterized by the energy that they bring into play, a reflection of their stability. The energy is that which must be provided to break the bond between the 2 atoms. In general, a distinction is made between weak interaction, involving energy levels from a few calories to around 12 kilocalories per mole of complex, and strong interactions, covalent or coordinate bonds, with bond energies ranging from 50 to 100 kilocalories per mole. Weak interactions are normally grouped into 3 main categories, hydrophobic bonds, dipolar bonds and certain ionic bonds. Although these weak interactions involve modest energy levels and produce complexes of low stability, they are nonetheless of great biological importance, as they control virtually all the phenomena of recognition between receptors and substrates.

Strong interactions, mainly covalent bonds, result when 2 atoms share a pair of electrons and these are classically represented in chemical formulae by dashes. They involve high energies (50 to 100 kilocalories per mole) and are therefore very stable compared with the weak interactions. The 2 electrons required for bond formation can be contributed by both partners, in which case it is called a radical reaction, or can be provided by one of the atoms, which is especially electron-rich, and shared with the electron-poor atom; in this case, it is referred to as a reaction between a nucleophile (electron-rich) and an electrophile (electron-poor). These 2 terms, nucleophile and electrophile, represent the ability of a molecule, or rather an atom of this molecule, to donate or accept electrons to form a bond. Nucleophilic centres are rich in electrons and therefore partially negatively charged, while electrophilic centres, deficient in electrons, are partially positively charged.

There is also another type of relatively strong bond, comparable to covalent bonds, which occurs between metals or metal salts and electron-rich atoms (mainly hetero-atoms, such as nitrogen or oxygen). These interactions, known as coordinate bonds, permit these electron-rich groups (the ligands) to transfer part of their electron density to the metal and increase its stability. Coordinate bonds are characterized by the number of ligands and by a geometry characteristic both of the metal and of its degree of oxidation (Fig. 1). For example, Cobalt II (Co^{2+}) is characterized by a tetrahedral arrangement, nickel II (Ni^{2+}) by a square planar tetra-coordinated arrangement, and chromium III (Cr^{3+}) by a 6 ligand octahedral arrangement. It is the number of ligands and the geometry of these coordination complexes that determine whether the metals are allergic (3, 4).

The Principal Electrophilic Chemical Groups Present in Contact Allergens and their Interaction with the Skin

Many chemical groups have electrophilic properties and are able to react with various nucleophiles to form covalent bonds. Table 1 shows those chemical groups most frequently found in contact allergens and the mechanism by which they react with nucleophilic groups. There are 3 main types of mechanism (5), nucleophilic substitution on a saturated centre (e.g., alkyl halides and epoxides), nucleophilic substitution on an unsaturated centre (aromatic halides or esters) and nucleophilic addition (carbonyl derivatives and α, β-unsaturated systems).

If we consider the human body in its entirety and from a chemical viewpoint, it becomes apparent that a very large proportion of biological structures, especially nucleic acids and proteins, contain many electron-rich groups (nitrogen, phosphorous, oxygen or sulphur). We can thus consider the human body overall as being nucleophilic and thus able to react with electrophilic chemical sub-

![Fig. 1. Examples of coordination geometry.](image-url)
Table 1. Principal electrophilic groups seen in contact allergy

<table>
<thead>
<tr>
<th>Group</th>
<th>Name</th>
<th>Reaction mechanism</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-CH₂-X X</td>
<td>Allyl halide</td>
<td>Nucleophilic substitution on a saturated centre</td>
<td>Na-CH₂-X</td>
</tr>
<tr>
<td>X = Cl, Br, I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X = NO₂</td>
<td>Aryl halide</td>
<td>Nucleophilic substitution on an unsaturated centre</td>
<td></td>
</tr>
<tr>
<td>X = F, Cl, Br, I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-H</td>
<td>Aldehyde; R = H</td>
<td>Nucleophilic addition</td>
<td></td>
</tr>
<tr>
<td>Ketone; R = alkyl or aryl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R = OR</td>
<td>Ester; R = OR</td>
<td>Nucleophilic substitution on an unsaturated centre</td>
<td></td>
</tr>
<tr>
<td>R = NH₂</td>
<td>Amid; R = NH₂</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R = O₂</td>
<td>Epoxide</td>
<td>Nucleophilic substitution on a saturated centre</td>
<td></td>
</tr>
<tr>
<td>Lactone; X = O</td>
<td>Lactone; X = NH</td>
<td>Nucleophilic substitution on an unsaturated centre</td>
<td></td>
</tr>
<tr>
<td>R = H, R, OR</td>
<td>Unsaturated aldehydes and ketones</td>
<td>Nucleophilic addition</td>
<td></td>
</tr>
<tr>
<td>R-H</td>
<td>Para-quinone</td>
<td>Nucleophilic addition</td>
<td></td>
</tr>
<tr>
<td>N₃²⁺, Ca²⁺, Cr³⁺</td>
<td>Metal salts</td>
<td>Coordination bond</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Examples of allergizing molecules. The electrophilic centre is indicated by an arrow.

The Hapten-Protein Bond: Covalent or Non-covalent?

In biology, few phenomena are irreversible, the majority being controlled by equilibria. It is easy to understand that the more stable the hapten-protein complex, the greater the possibility of the immune system being able to process the immunological information, resulting in allergy. Given this, we can understand why the very strong (and thus difficult to reverse) covalent bond produces a maximal biological efficacy. It is therefore natural that it is this type of bond that is found in the majority of cases of allergy. However, it would be incorrect to think that only covalent bonds result in allergy. In the case of metal salts, the formation of covalent bonds is impossible and, in this case, a sufficiently stable coordination complex must be formed between the metal salt and the electron-rich residues of proteins. These coordination complexes are sufficiently stable, and the protein modification sufficiently great, to lead to allergy (12).

Metabolism and Prohaptens

Far from being an inert tissue, the skin is the site of many metabolic processes, which can result in structural modification of xenobiotics which penetrate into it. These metabolic processes, primarily
intended for the elimination of foreign molecules during detoxification, can, in certain cases, convert harmless molecules into derivatives with electrophilic, and therefore allergenic, properties. The metabolic processes are mainly based on oxidation reactions via extremely powerful enzymatic hydroxylation systems, such as the cytochrome P450 enzymes (13), but monoamine oxidases, which convert amines to aldehydes, and peroxidases are also found in the skin. The latter, when activated by the production of hydrogen peroxide during the oxidative stress following the introduction of a xenobiotic into the skin, can convert the electron-rich aromatic derivatives (aminated or hydroxylated) into quinones, which are powerful electrophiles. In this way, the long-chain catechols, responsible for the severe allergies to poison ivy (Rhus radicans L.) and poison oak (Rhus diversiloba T.), are oxidised in vivo to the highly-reactive orthoquinones (14) (Fig. 3). The same applies to para-phenylenediamine or hydroquinone derivatives, e.g., the allergens from Phacelia crenulata Torr., which can be converted into electrophilic paraquinones (15). Enzymatic hydrolyses can also occur with tulinosides A and B, found in the bulb of the tulip (Tulipa gesneriana L.) releasing the actual allergens, tulinalains A and B (16).

Non-enzymatic processes, such as atmospheric oxygen or ultra-violet irradiation, can also induce changes in the chemical structure of molecules. Many terpenes spontaneously auto-oxidize in air, producing allergenic derivatives. In the 50s, it was found that allergenic activity of turpentine was mainly due to hydroperoxides of one of the monoterpenes, Δ^{2}-carene (17). This is also the case for another monoterpenene, d-limonene found in citrus fruits. d-limonene itself is not allergenic but at air exposure hydroperoxides, epoxides and ketones are created which are strong allergens (18, 19). The diterpenoid resin acids in colophony are also activated by autoxidation. The main constituents abietic acid and dehydroabiestic acid are converted in highly reactive hydroperoxides and epoxides by contact with air (10, 20–22). These studies show that from the same prohapten different haptenes can be created which don’t necessarily cross-react.

All these molecules, which themselves have no electrophilic properties and cannot therefore be haptenes, but which can be converted to haptenes either by metabolic or non-enzymatic processes, are referred to as prohaptenes (5, 23). They play an important rôle in contact allergy because of their number and their highly reactive nature. Moreover, as the structure of the metabolized molecule can be very different from the structure of the initial molecule, the allergologic investigations become even more difficult.

**Haptens and Cross-allergy**

The factors that control molecular recognition during the elicitation stage are primarily the nature of the chemical group and the compatibility of the spatial geometry. Although the nature of the chemical group is very important and serves to define what are commonly called the group allergies, it cannot account for all structure-activity relationship. Receptor molecules are highly sensitive to volume and shape, and molecules must have a similar size and spatial geometry to be recognized by the same receptor. Thus, even though the molecules tulinalain A or B and alantolactone (the allergen of Imula heinemia L.) bear the same chemical group, α-methylene-γ- butyrolactone, they cannot give rise to cross-allergic reactions, as their spatial volumes are too different (Fig. 4). In contrast, isoalantolactone and alantolactone produce a cross-allergic reaction (24), since they share a homologous chemical group and spatial volume. The term cross-allergy is often misused and should be restricted to well-defined cases that can be called true cross-allergies (25, 26).

True cross-allergy between a sensitiser A and a
Fig. 4. Chemical structure and spatial representation of tulipalin A, alantolactone and isoalantolactone.

triggering agent B can be interpreted in various ways:

A and B are chemically and structurally similar hapten,
A is metabolized to a hapten that is similar to B,
B is metabolized to a hapten that is similar to A,
A and B are metabolized to similar hapten.

The identification of cross-allergic responses can be especially difficult, particularly in man, in whom the possibility of co- or poly-sensitization should not be ruled out. In addition, the metabolism of molecules can be very complex and 2 molecules with a priori little in common can be converted to derivatives that have a similar structure. It is therefore dangerous to draw conclusions from tests without knowing how the substances used are liable to be metabolized. Many reactions described as demonstrating cross-allergy are, without doubt, due to co-sensitization (26). Experimental studies in animals are often the only means of being really certain of what happens during recognition. The concept of the prohapten is very important in the interpretation of results in allergy. As the structure of the metabolized molecule can sometimes be very different from that of the initial molecule, it can be difficult to establish similarities of chemical groups and structure.

Molecular Modelling

In the last few years, molecular modelling has been shown to be a powerful tool in studies of conformational-dependent drug-receptor interactions and structure-activity relationship analysis (24–30). Despite the great potential of this technique, few attempts to analyze cross-reaction patterns in the field of allergic contact dermatitis have yet been reported. One reason may be the heterogeneous population of patients with heterogenous clinical histories, in which it is somewhat difficult to distinguish between actual cross-reaction and concomitant sensitization. A second reason is that, to be effective, structure-activity relationship studies need data for a wide range of molecules. The clinical investigation of contact dermatitis from corticosteroids, in which a large number of related substances have been tested on a large number of patients, represents a good opportunity to carry out such a structure-activity study. From statistical analysis of the clinical data, it was possible to advance an experimentally-supported hypothesis for cross-reaction patterns, as Coopman et al. (31) hypothesized that cross-reactions occur primarily within certain groups of corticosteroids. They distinguished 4 groups, group A consisting of hydrocortisone, tixocortol pivalate, and related compounds, group B consisting of triamcinolone acetonide, aminonide, and related compounds, group C consisting of betamethasone, dexamethasone, and related compounds, and group D consisting of esters such as hydrocortisone-17-butyrate and clobetasone-17-butyrate. Lepoittevin et al. (32) have correlated this with conformational characteristics, in order to establish a molecular bases for cross-reaction patterns in patients sensitized to corticosteroids.

The conformation of corticosteroids from groups A, B, C and D was analyzed. This study was based on 2 hypotheses. The 1st was that all corticosteroids should interact with proteins in a very similar way. All corticosteroid molecules were designed to interact with the same type of receptors, and thus should be more or less metabolized in similar ways. The 2nd hypothesis based on chemical observations, was that esters at position 21 are readily hydrolyzed to the free alcohol, while esters at position 17 are more resistant to hydrolysis, due to strong steric hindrance. Thus, for example, tixocortol pivalate was considered as tixocortol with a free thiol group at position 21, and alclometasone-17,21-dipropionate considered as alclometasone-17-propionate.

All molecules were drawn from energy-minimized building blocks and were then submitted to a multi-conformational analysis, in order to achieve the most energetically stable conformation. These conformations were then compared for analogies or differences in the Van der Waals' volumes that define the electronic shape of the molecule. As expected from the hypothesis, significant group-specific characteristics of volume and shape were found for molecules of group A, B
and D, but not for molecules of group C. The existence of groups A, B and D, as defined by the analysis of cross-reaction patterns in patients sensitized to corticosteroids, was fully supported by the conformational analysis of these molecules. Molecules of the same group have very similar spatial structures, explaining the cross-reactions observed. In addition, molecules from one group were sufficiently different from molecules of another group to explain the lack of cross-reactions observed between groups A, B and D. The volume occupied by specific groups on the α face of ring D seemed to be critical for the molecular recognition of corticosteroids by receptors of immuno-competent cells, while modifications of other parts of the molecule seemed to have little effects on the recognition patterns. As shown in Fig. 5, each group represents a well-defined, characteristic shape, which can be very useful for the prediction of potential cross-reactions of new corticosteroid molecules.

**Structure-activity Relationships for Contact Allergens**

Whilst it is true that some limited consideration was given in earlier decades to the relationship between chemical structure and the ability to cause contact allergy, see, e.g., (23), it is only in fairly recent times that a sustained effort has been made in this direction. An important turning point was early in the 1980s when Dupuis & Benezra (5) described, from an organic chemistry viewpoint, the key features that rendered a hapten protein reactive. This was closely followed by the description by Roberts & Williams (33) of an approach to mathematical modelling of skin sensitization potential, the relative alkylation index (RAI). In essence, this model proposed that the degree of sensitization was closely correlated with the extent to which skin protein became haptenated. This concept has been developed and extended in several pieces of work that will be discussed later in this section. However, an important area for attention is in the development of new computer-based systems for the identification of skin sensitizers. These are “expert systems” that use rules defined by experts to make predictions about novel chemicals. Since the body’s immune system is normally only reactive when stimulated, it is reasonable to suppose that the important factors that determine whether, an to what extent, a chemical can behave as a contact allergen, lie within the structure of that chemical. Consequently, one expects to see a

![Figure 5](image-url)

*Fig. 5. Characteristic shape of corticosteroids of class A, B and D.*
relationship between chemical structure and biological activity. This may be most apparent in families of organic molecules with similar characteristics, such as the nitrohalobenzenes (34) or the para-substituted benzenes (35). However, of greater importance may be families of chemicals where the mechanism of reaction with skin protein can be predicted to be the same. It is this aspect that was discussed in detail in the seminal work of Dupuis & Benezra (5). Nevertheless, our understanding of this area of science is far from complete. Apparently straightforward predictions of reaction mechanisms have proven difficult to substantiate in practice (36, 37). Complications arise from the need to understand the role of skin metabolism (38, 39) and the presence of strongly sensitizing contaminants, e.g., diethyl fumarate in malathion (40) and oxidation products in colophony (20).

QSAR models of skin sensitization have been based on the concept that the important factors governing the "quantity" of sensitization are the reactivity of the chemical, its ability to locate in the crucial epidermal environment and, of course, the dose applied. It was this philosophy that led to the development of the RAI model (33). This model was used initially to provide equations describing the skin sensitizing activities of sulfones and p-nitrobenzyl halides (34), but has since been adapted to cover alkyl alkane sulfonates (41), acrylates (42), haloalkanes (43), phenyl benzoates (44) and, most recently, a family of furanone derivatives (45, 46).

In general, RAI models use molar dose and log $P$ (where $P$ is the octanol water partition coefficient) as 2 parameters. When skin penetration is limited by high hydrophobicity, a $(\log P)^2$ term can be valuable (42, 44). However, the chemical reactivity term can be rather difficult to obtain. In some cases, it has proven possible to make estimates (42) and in others to eliminate it as an important variable (45), but these represent very restricted datasets. The best option, which was proposed in the original paper (33), is to measure the reactivity in vitro with a defined nucleophile, and this has been done with a recent series (45, 46). Unfortunately, this is not always an easy process, and the choice of acceptor nucleophile may be more important than we realise. Certainly, it is exceptional when we have any clue as to the nature of the important in vivo acceptor sites (47).

Another important problem of QSAR interpretation is the nature and quality of the biological data available for modelling. A more user-friendly approach has been described (34, 38), but it may be that QSARs will not prove generally applicable in skin sensitization. Their real value may be as indicators of when our mechanistic understanding is right, i.e., the QSAR predicts well within a homologous series of chemicals, or conversely when we are provided with the opportunity for novel thinking, i.e., a QSAR that seemed satisfactory, but fails to account for new results (49).

**Expert Systems**

This is an area of work that is still in its infancy, and should be distinguished from computer database type systems (50). Expert systems are rule-based, where rules are predefined by experts, not the computer. For skin sensitization, the best described and most advanced system is DEREK; DERivation of Risk from Existing Knowledge (51). A UK collaborative effort has led to the building of an expert system for skin sensitization based on the DEREK architecture (52). The original system was built on a rulebase of 40 rules, derived by expert analysis of guinea pig maximization test results (classified according to EEC criteria) on almost 300 chemicals. This rulebase has since been further refined and enhanced by analysis of data on another 40 chemicals (53). The rules essentially describe structural alerts for chemical reaction mechanisms and factors that enhance and/or limit them. So far, the expert system only considers the 2-dimensional structure of a potential allergen. However, chemical reactivity is best estimated from the 3-dimensional structure of a compound. Importantly, the current version of DEREK does not yet take account of the ability of the (pro)hapten to partition into the epidermis.

Others are known to be following related approaches, e.g., COMFA and neural network analysis of databases of chemicals, grouping chemicals into families with similar chemistries, searching for common structural elements, etc. (54), but generally, this work is as yet unpublished. It seems likely that, out of these activities, a consensus will emerge as to the spectrum of common reaction mechanisms and structural alerts for skin sensitization. What may then be left is the need to account for differences in skin penetration, and a measure of the extent to which skin enzymes may activate or inactivate different chemicals.

**Other Activities**

Given a large database, it is possible to carry out a statistical evaluation of the information. This has been done in one case by grouping guinea pig maximization test results on almost 300 chemicals into structurally similar type and then carrying out multivariate analysis (55). The outcome was only
partially successful, the predictive power of the method being rather poor.

Conclusion
The principles that we have discussed permit a reasoned approach to the chemical phenomena induced in contact allergy, however, we often have available only indirect evidence suggestive of one mechanism or another. Although the chemical bases for hapten-protein interactions can be investigated in the laboratory by the use of nucelophilic amino acids, small peptides and model proteins, and although a certain number of steps can be examined, at the present time, no method is available to follow a hapten step by step during the entire immunological process that leads to contact allergy. Many points await investigation, but in may cases a "chemical" analysis of the problem does allow us to understand and to foresee cross-allergies, and thus to predict better the sensitizing potency of chemicals and eventually to warn the patient about structurally-related products.

References
33. Roberts D W, Williams D L. The derivation of quantitative correlations between skin sensitization and physicochem-

Address:
J.-P. Lepoittevin
Laboratoire de Dermatochimie
Clinique Dermatologique
CHU, F-67091 Strasbourg Cedex
France