

Development of structure–activity relationships (SARs) in allergic contact dermatitis

J.-P. Lepoittevin

Laboratoire de Dermatochimie associé au CNRS, UMR 7509, Université Louis Pasteur, Clinique Dermatologique, CHU, Strasbourg, France

Accepted 19 November 1998

Keywords: allergic contact dermatitis, alternative methods, expert systems, hapten, quantitative structure–activity relationships

Abstract

One of the major objectives at the end of this century is the development of “alternative” tests for the evaluation of the pharmacological and/or toxicological activity of newly developed molecules. Contact allergy is no exception to the rule and many research programs have been started to develop *in vitro* techniques for the detection of allergizing compounds. In parallel with these biological studies, another approach is becoming important and will no doubt become more so, namely, the study of structure–activity relationships (SARs). This consists of using molecular or physicochemical properties to predict and, in certain cases, quantify the allergizing potential of a new molecule without using any biological test. Three main approaches are currently under study: the creation of allergy databases, the design of “expert” computerized systems, and the development of quantitative SARs. These three often complementary approaches are still at the development stage, but we can begin to see their potential and limitations. The aim of this article is not to give an exhaustive description of all the systems developed worldwide, but to illustrate each approach by giving some important examples.

Abbreviations: (Q)SAR, (quantitative) structure–activity relationships; RAI, reactive alkylation index; RAI_c, reactive alkylation index calculated with the dose used at challenge; RAI_i, reactive alkylation index calculated with the dose used at incubation

Introduction

One of the major objectives at the end of this century is the development of “alternative” tests for the evaluation of the pharmacological and/or toxicological activity of newly developed molecules. Contact allergy is no exception to the rule and many research programs have been started to develop *in vitro*

techniques for the detection of allergizing compounds (Barbier et al., 1994). Until now, despite the considerable effort expended in the last few years, the complex character of the biological mechanism of allergy has prevented the development of a reliable test. Various promising routes are currently being explored, but are not expected to yield a result for several years.

In parallel with these biological studies, another approach is becoming important and will no doubt become more so, namely, the study of structure–activity relationships (SARs). This consists of using molecular or physicochemical properties to predict and, in certain cases, quantify the allergizing potential of a new molecule without using any biological test.

Three main approaches are currently under study:

1. The creation of allergy databases.
2. The design of “expert” computerized systems.
3. The development of QSARs (quantitative structure–activity relationships).

These three often complementary approaches are still at the development stage, but we can begin to see their potential and limitations. The aim of this paper is not to give an exhaustive description of all the systems developed worldwide but to illustrate each approach by giving some important examples.

Molecular basis of contact allergy and support for structure–activity relationships

The first stage in the induction or elicitation of a contact allergy is the interaction of the hapten – a small molecule of low molecular weight – with a protein, to form an antigenic complex. It is now accepted that, in the vast majority of cases, this interaction involves the formation of a covalent, and therefore irreversible, bond between the hapten and the protein (Lepoittevin et al., 1997). This implies that the hapten has chemical reactivity characteristics that allow it to form bonds with the side-chains of amino acids. The modified protein(s) is then processed by antigen-presenting cells, yielding antigenic peptides that are reexpressed on the surface of the cell in association with class I or

class II major histocompatibility complex molecules (Kalish, 1995). These modified peptides are then involved in the selection and activation of T lymphocytes that bear a lymphocyte receptor with a suitable geometry.

What do we currently know about the relationship between the chemical reactivity and structure of the hapten and the nature and selectivity of the immune response?

Chemical selectivity of haptens toward amino acids

Studies performed in the last few years on the mechanisms involved in the modification of proteins by small xenobiotic molecules have made it possible to demonstrate the involvement of a wide variety of mechanisms, depending on the structure of the hapten. One direct consequence of this diversity is the existence of selectivity in the modification of amino acids. Thus, we have been able to show that methyl alkanesulfonates, which are lipophilic methylating agents and strong allergens, modify, almost exclusively, histidine and methionine residues in proteins (Lepoittevin and Benezra, 1992), while the α -methylene- γ -butyrolactones, the main allergens in plants of the Asteraceae family, mainly modify lysine residues (Franot et al., 1993). Similarly, dinitrohalogenated derivatives, such as DNCB (dinitrochlorobenzene) or DNFB (dinitrofluorobenzene), are known to react readily and rapidly with lysine groups, while DNTB (dinitrothiocyanobenzene) reacts preferentially with thiol residues. It should be noted that, in contrast to the first two derivatives, which are strong allergens, DTNB only sensitizes under extreme conditions (in the presence of immunostimulating adjuvants) and behaves more as a tolerogen on simple application to the skin. This major difference in the activation of the immune system, apparently linked to differences in reactivity and selectivity, is, for the moment, poorly understood.

The importance of reaction mechanisms

The selectivity of haptens for certain amino acids is directly related to the electron density at the reaction site (determined by the structure of the molecule) and to the type of chemical reaction that takes place. The chemical reactions can be divided into several groups on the basis of their characteristics and the mechanisms involved in the breaking and forming of bonds. Thus, molecules with very similar structures can give rise to the formation of different intermediates, for example during metabolization, which lead to different reaction mechanisms. The case of eugenol and isoeugenol is interesting in this respect.

These two molecules, present in many natural extracts, have very similar structures, but their allergizing potentials are very different, eugenol being a weak sensitizer, and isoeugenol a relatively strong one. Recent studies in the mouse seem to indicate that these two very similar molecules are metabolized in different ways and have very different reaction mechanisms (Bertrand et al., 1997), which could explain the observed differences in sensitization potential. Eugenol seems to be metabolized to an electrophilic orthoquinone after a demethylation step (Figure 1), whereas isoeugenol seems to be directly oxidized to a quinonemethide (Figure 2).

Consequences

One direct consequence of the selectivity of haptens for amino acids is differences in the site of modification of proteins and in the density of these modifications. These differences, initially purely chemical, appear more and more to have a major impact on the response of the immune system. The selectivity of haptenization sites is directly involved in the selection of the peptides that are presented by antigen-presenting cells to T lymphocytes and

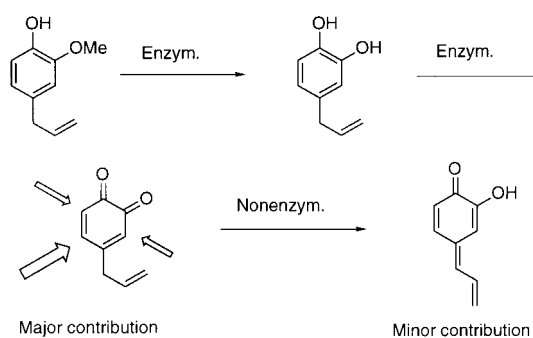


Figure 1. Proposed mechanism to explain the sensitizing potential of eugenol in mice. Arrows indicate major reactive sites.

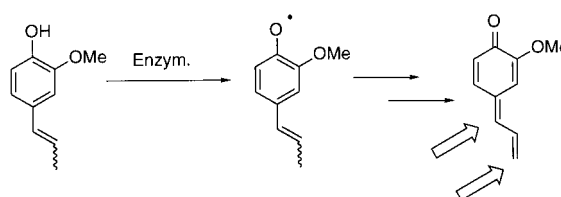


Figure 2. Proposed mechanism to explain the sensitizing potential of isoeugenol in mice. Arrows indicate major reactive sites.

thus in the selection of lymphocyte receptors. This selectivity also indirectly controls the haptenization level of the protein(s), since it appears that epitope density on the surface of the antigen-presenting cell could directly or indirectly influence the orientation of the immune response towards a Th1 or Th2 profile, a high epitope density favoring Th1 and a low epitope density Th2 (Murray, 1998).

Contact allergy databases

Introduction

Several types of contact allergy databases can be considered. That most often encountered is the "personal" database that contains all the allergology information from a private consultancy or a hospital department, one of the best

examples being that of the Dermato-Allergology Department at Leuven, which has developed a highly comprehensive diagnostic aid program (Dooms-Goossens et al., 1992). In our case, the objective is slightly different, being to list all the cases of contact allergy reported in the literature in a database that is as complete as possible, listing the maximum information for each case, with the orientation being not clinical, but “chemical”.

Development

Currently, the most complete database is that developed in Palo-Alto by CCS Associates in collaboration with H.I. Maibach of the University of San Francisco and Professor C. Benezra (Benezra et al., 1985). The main sources of data used are the cases of allergy published in *Contact Dermatitis* since 1975. The information compiled and entered in the database for each allergen is:

Identification: Name, chemical structure and *Chemical Abstracts* number.

Allergologic data: Primary sensitizer or eliciting agent, number of reported cases and condition of the patient (with eczema or another pathology, or healthy).

Technical data: Type of test (open epicutaneous test or patch test), concentration, vehicle, intensity of reaction, tests on man or on an animal model, number of subjects tested and *Contact Dermatitis* reference.

Evaluation: Confidence level.

The last point deserves some explanation. The confidence level is graded from 0 to 5 and represents an assessment of the relevance of the information. This in no way involves a judgement of the investigator's work, but rather assesses the weight to be given to this

information in the database, an important point if these databases are to be used predictively. Examples of the criteria used to assess the confidence level are the inclusion of controls, the purity of the compounds tested and the number of cases.

The database can be searched by name, by *Chemical Abstracts* number, by chemical structure, and by all the previously defined biological parameters, either independently or in combination. Searching is performed using the program PROPHET, which consists of a hierarchical tree based on chemical structures and substructures (Table 1) selected for their importance in contact allergy. This type of approach was developed in the 1970s to create carcinogen databases and was then extended to contact allergy. The tree is currently based on 250 chemical structures.

Table 1. Examples of groups and subgroups used to classify chemical substances

Group	Subgroup
Alcohol/thiols	Benzyl alcohols Enols
Aldehydes	Conjugated aldehydes Aromatic aldehydes Hydrazones
Terpenes	Monoterpenes Sesquiterpenes Diterpenes Triterpenes

The chief use of a system of this type is to answer questions such as:

- Does this compound belong to an allergizing group?
- Which compounds are liable to give allergic cross-reactions?
- What concentration and what vehicle should be used to test this product?
- What is the relevance of these responses?

Limitations: The main limitations of this system result from the way in which it is compiled, with the data being extracted from the literature and therefore subject to the inherent bias of scientific publications; for example, a severe reaction in a few individuals will often be better documented than moderate reactions in a larger number of patients.

Expert systems

Introduction

“Expert” systems are decision-aiding programs, based on the use of large databases, that can use, for example, the structural, physicochemical, and electronic data for a new compound and compare them with the data stored in memory to give an opinion as to the allergizing status of a molecule. The new parameters thus defined are then fed into the database to improve the expert system.

Various approaches

In the expert system field, many parameters can be considered.

Structural alerts: These consist of total or partial chemical structures known to present allergizing risks. They are mainly defined by analysis of the literature and contact allergy databases and allow the expert system to detect a “risk.” One system based on this approach is currently marketed under the name of DEREK (Deductive Estimation of Risk from Existing Knowledge).

This expert system consists of a “control” program that analyses the structure of the molecules and a database consisting of “rules” in the form of substructures known to be associated with allergizing activity (Barratt et al., 1994). Examples of “rules” are shown in Table 2.

Risk indicators: These are physicochemical parameters that favor allergizing potential. Each positive parameter is considered as an aggravating factor without taking into account the intensity of aggravation. Examples of risk indicators are:

- Presence of one or more reactive chemical groups
- Molar refractive index
- Number of hydrogen bonds
- Number of electron pairs on oxygen or nitrogen
- Number of NH or OH groups
- Number of ionizable groups
- Number of aromatic structures
- Number of heterocycles
- Number of metabolizable groups

Each parameter can be allocated a numerical coefficient of varying size, depending on its effect on the mechanism of the allergy. The potential risk is evaluated by taking all these factors into account. The development of such systems is still in its infancy and will require a long validation phase to distinguish those factors that are really important from those that are anecdotal.

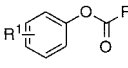
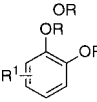
Limitations: One of the main limitations of this system is that, at the present time, the effects of metabolism are not considered and the process is based simply on the structure of the tested molecule, which is not necessarily that which reacts with the skin proteins.

Quantitative structure–activity relationships (QSAR)

Introduction

The main objective of the QSAR approach is to define a certain number of “quantifiers” for

Table 2. Examples of rules and chemical structures used by DEREK to predict the sensitizing potential

Chemical function	Substituents	Structure
Aldehydes	R = alkyl, aryl	
Ketones	R, R ¹ = alkyl, aryl	
Aldehydes, amides, esters and α,β unsaturated ketones	R = H, C, N, O (not OH) R ¹ = not heteroatoms, esters, ketones R ² = not aryl except if R = H	
Phenyl esters	R = alkyl, aryl R ¹ = any	
Hydroquinones and their O-alkyl precursors	R = H, alkyl R ¹ = any	
Catechols and their O-alkyl precursors	R = H, alkyl R ¹ = any	
Anhydrides	R = any	
Aromatic primary and secondary amines	R = alkyl, aryl R ¹ = any	
Alkyl halides	R = alkyl X = Cl, Br, I	

a given biological reaction, then combine these to give a quantitative estimate of the biological activity; it therefore no longer simply answers the question whether the molecule is potentially allergizing, but also gives a quite accurate indication of the expected intensity of the biological response. The currently most highly developed quantifier is the relative alkylation index (RAI) (Fraginals et al., 1991).

Definition of the RAI

This quantifier gives an estimate of the level of modification of proteins by a potentially allergizing molecule and assumes that the intensity of the biological response is directly related to the level of modification. This is a

simple hypothesis taking into account three parameters, the dose of the product applied to the skin, its lipophilicity and its chemical reactivity. These parameters were chosen because it has been known experimentally for a long time that the allergizing response increases with an increase in each of these parameters. In the RAI model, the epidermis is divided into a series of units of finite volume. These volumes, mainly lipophilic in nature, are traversed by a polar liquid and contain a large number of proteins (Figure 3).

The amount of product entering one unit of volume is assumed to be proportional to the dose of product applied to the skin and is expressed as the logarithm of the dose D . The residence time of the test molecules in the unit

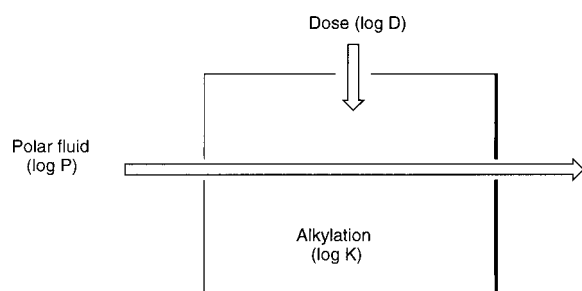


Figure 3. Basis of the Relative Alkylation Index (RAI).

volume is proportional to their lipophilic character measured by the logarithm of the partition coefficient P , which represents the relative affinities of the allergen in an octanol–water mixture; the more lipophilic the product and the greater its $\log P$, the longer it remains in the unit volume. The chemical reactivity of the allergen toward proteins is itself expressed as the logarithm of its rate constant k for a model nucleophile, butylamine. It is assumed that this reaction is first order, i.e., that the proteins are always in excess relative to the allergen.

It is thus possible to express the relative level of modification of the proteins by the equation:

$$\text{RAI} = a \log k + b \log P + c \log D \quad (1)$$

where a , b and c represent the relative weight of each parameter. At the present time, these constants are determined experimentally and are only valid for a series of molecules with the same reaction mechanism, and it is for this reason that the RAI is defined as a relative index of alkylation.

The intensity of the allergic reaction increases with increase in each of these parameters, and thus with an increase in RAI, up to a certain limit called the overload point. This point, an experimental observation, is still poorly understood, but probably corresponds to the saturation of the biological mechanisms

and the setting up of regulatory mechanisms.

Application

A series of molecules with structures similar to those of the α -methylene- γ -butyrolactones has been synthesized and studied (Fraginals et al., 1991; Franot et al., 1994a,b). On reaction with proteins, these molecules form the same type of epitope as that produced by reaction with

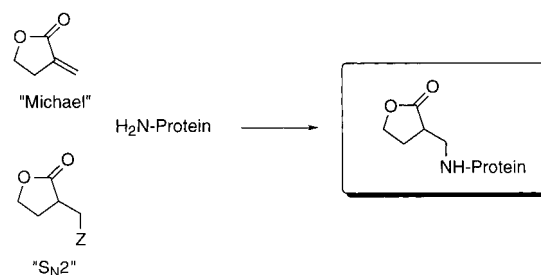


Figure 4. Reaction of α,β -unsaturated and substituted lactones with proteins.

the corresponding α -methylene- γ -butyrolactone (Figure 4).

The various substituents were chosen to cover a significant range of reactivity. These molecules make it possible both to observe the effect of the chemical reactivity on the biological activity and also to develop the RAI by simultaneously varying the kinetic constant and physicochemical factors.

Before assessing the biological activity of these compounds, we evaluated their relative reactivity for butylamine (rate constant k), chosen as a nucleophile model; the partition coefficients P were determined by calculation.

When the allergizing activity was tested in the guinea-pig, these compounds proved to be good allergens. Initial modeling using the previously determined kinetic constants and the calculated $\log P$ values indicates a good correlation for

$$\text{RAI} = \log D + 2 \log k + 0.5 \log P \quad (2)$$

Using this definition of the RAI, it is then possible to define the equation for the biological reactivity as

$$\text{Biological response} = a(\text{RAI}_i) - b(\text{RAI}_i)^2 + c(\text{RAI}_c) + d \quad (3)$$

where a , b and c are positive numerical coefficients and d is a constant for a given series of compounds. In our case, if the biological response is expressed as a percentage of the maximum observed response, using the biological results obtained during this study, we generate an equation of the type:

$$\% \text{ Response} = 73.0(\text{RAI}_i) - 7.2(\text{RAI}_i)^2 + 17.5(\text{RAI}_c) - 208 \quad (4)$$

As can be seen in Figure 5, there is a good correlation between the observed biological responses and those calculated using the above equation.

Limitations: Currently, RAI has two main limitations. First, it is a relative system that gives good results within a series of analogous

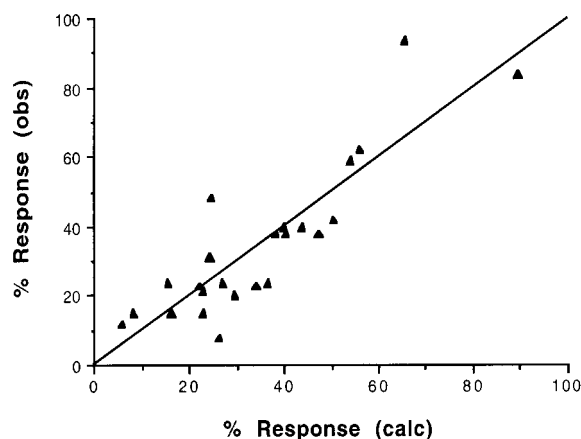


Figure 5. Correlation between observed and calculated response to substituted lactones according to equation (4).

compounds that react with proteins by the same mechanism, but cannot yet be used to assess the allergizing potential of molecules belonging to two very different families. Second, at present, this mathematical model does not take skin penetration factors into account in the evaluation of the dose of product actually reaching the reaction site.

Conclusion

Although they have proved themselves in other domains, in contact allergy, we are still at the very beginning of the development of structure–activity relationships. Owing to the complex biological mechanism involved, the number of parameters that must be taken into account is much greater. Nevertheless, the few pilot approaches that have been developed are very promising and should rival or complement alternative biological tests.

References

- Barbier A, Rizova E, Stampf J-L et al. Development of a predictive *in vitro* test for detection of sensitizing compounds (European BRIDGE project). In: Rougier A, Goldberg AM, Maibach HI, eds. *In vitro* skin toxicology (irritation, phototoxicity, sensitization). New York: Liebert; 1994:341–50.
- Barratt MD, Basketter DA, Chamberlain M, Admans GD, Langowski JJ. An expert system rulebase for identifying contact allergens. *Toxicol in Vitro*. 1994;8:1053–60.
- Benezra C, Sigman CC, Perry IR, Helmes T C, Maibach HI. A systematic search for structure–activity relationships of skin contact sensitizers. *J Invest Dermatol*. 1985;85:351–6.
- Bertrand F, Basketter DA, Roberts DW, Lepoittevin J-P. Skin sensitization to eugenol and isoeugenol in mice: possible metabolic pathways involving ortho-quinone and quinone methide intermediates. *Chem Res Toxicol*. 1997;10:335–43.
- Dooms-Gossens A, Dooms M, Drieghe J. Computers and patients' information systems. In: Rycroft RJG, Menne T, Frosch PJ, Benezra C, eds. *Textbook of contact dermatitis*. Berlin: Springer-Verlag; 1992:771–84.
- Fraginals R, Roberts DW, Lepoittevin J-P, Benezra C. Refinement of the relative alkylation index (RAI) model for skin sensitization and application to mouse and guinea-pig test data for alkylsulfonates. *Arch Dermatol Res*. 1991;283:387–94.

- Franot C, Benezra C, Lepoittevin J-P. Synthesis and interaction studies of ^{13}C labeled lactone derivatives with a model protein using ^{13}C NMR. *Bioorg Med Chem*. 1993;1:389–97.
- Franot C, Roberts DW, Smith RG, Basketter DA, Benezra C, Lepoittevin J-P. Structure–activity relationships for contact allergenic potential of γ,γ -dimethyl- γ -butyrolactone derivatives Part I. *Chem Res Toxicol*. 1994a;7:297–306.
- Franot C, Roberts DW, Basketter DA, Benezra C, Lepoittevin J-P. Structure–activity relationships for contact allergenic potential of γ,γ -dimethyl- γ -butyrolactone derivatives Part II. *Chem Res Toxicol*. 1994b;7:307–12.
- Kalish RS. Antigen processing: the gateway to the immune response. *J Am Acad Dermatol*. 1995;32:640–52.
- Lepoittevin J-P, Benezra C. ^{13}C -Enriched methylalkane-sulfonates: new lipophilic methylating agents for the identification of nucleophilic amino acids of protein by NMR. *Tetrahedron Lett*. 1992;33:3875–8.
- Lepoittevin J-P, Basketter DA, Goossens A, Kalberg A-T (eds). *Allergic contact dermatitis: the molecular basis*. Berlin: Springer-Verlag; 1997.
- Murray JS. How the MHC selects Th1/Th2 immunity. *Immunol Today*. 1998;19:157–63.
- Address for correspondence:* Jean-Pierre Lepoittevin, Laboratoire de Dermatochimie, Clinique Dermatologique, CHU, F-67091 Strasbourg Cedex, France.
E-mail: jplepoit@chimie.u-strasbg.fr