

Dermato-allergologie

Développement de systèmes experts pour la prédiction du pouvoir sensibilisant

J.P. Lepoittevin

Laboratoire de dermatochimie associé au CNRS (UMR 7509), université Louis-Pasteur, clinique dermatologique, CHU, 67091 Strasbourg cedex, France

allergie de contact / haptènes / méthodes alternatives / relations structure-activité quantitatives / systèmes experts

allergic contact dermatitis / hapten / alternative methods / structure-activity relationships / expert systems

Un des grands objectifs de cette fin de siècle est la mise au point de tests dits « alternatifs » pour l'évaluation de l'activité pharmacologique ou toxicologique des nouvelles molécules mises sur le marché. L'allergie de contact ne fait pas exception à la règle et de nombreux programmes de recherche ont été lancés afin de développer des techniques *in vitro* de détection des substances allergisantes [1]. Jusqu'à présent, le caractère complexe du mécanisme biologique de l'allergie n'a pas permis, malgré les efforts considérables déployés ces dernières années, d'aboutir à un test fiable. Différentes voies prometteuses sont actuellement à l'étude mais n'aboutiront sans doute pas avant plusieurs années. Parallèlement à ces recherches à caractère biologique, une autre approche prend et va sans doute prendre beaucoup d'importance : l'étude des relations structure-activité. Il s'agit, à partir de caractéristiques moléculaires ou physicochimiques de prévoir et, dans certains cas, de quantifier le potentiel allergisant d'une nouvelle molécule sans qu'aucun test biologique ne soit pratiqué. Trois principaux axes sont actuellement à l'étude :

- la création de bases de données sur l'allergie ;
- la conception de systèmes informatiques experts ;

– le développement du *Quantitative Structure-Activity Relationship* (QSAR).

Ces trois approches, souvent complémentaires, en sont encore au stade du développement, mais l'on commence à en cerner le potentiel et les limitations. L'objectif de cet exposé n'est pas d'être exhaustif et de décrire tous les systèmes développés de par le monde, mais d'illustrer chaque approche, en prenant quelques exemples significatifs.

BASES MOLÉCULAIRES DE L'ALLERGIE DE CONTACT : SUPPORT DES RELATIONS STRUCTURE-ACTIVITÉ

La première étape dans le processus d'induction ou de révélation d'une allergie de contact est l'interaction de l'haptène, une petite molécule de faible poids moléculaire, avec une macromolécule biologique, dans ce cas une protéine, pour former un complexe antigénique. Il est couramment admis que cette interaction fait intervenir, dans la très grande majorité des cas, la formation d'une liaison covalente et donc irréversible entre l'haptène et la protéine [12]. Ceci implique que l'haptène possède des caractéristiques de réactivité chimique lui permettant de former des

liaisons avec les chaînes latérales des acides aminés. La ou les protéines ainsi modifiées vont ensuite être apprêtées par les cellules présentatrices d'antigène pour former des peptides antigéniques qui seront réexprimés à la surface de la cellule en association avec le complexe majeur d'histocompatibilité de classe I ou de classe II [11]. Ces peptides modifiés vont servir à la sélection et à l'activation des lymphocytes T portant un récepteur lymphocytaire de géométrie adéquate.

Quelles sont actuellement nos connaissances sur les relations existant entre la réactivité chimique et la structure de l'haptène, la nature et la sélectivité de la réponse immunitaire ?

Sélectivité chimique des haptènes vis-à-vis des acides aminés

Les études réalisées ces dernières années sur les mécanismes de modification des protéines par de petites molécules xénobiotiques ont permis de mettre en évidence toute une variété de mécanismes, en relation avec la structure de l'haptène. Une conséquence directe de cette diversité est l'existence d'une sélectivité dans la modification des acides aminés. Ainsi, nous avons pu montrer que les alcanesulfonates de méthyle, qui sont des agents méthylants lipophiles et des allergènes forts, modifient presque exclusivement les résidus histidine et méthionine des protéines [13] alors que les α -méthylène- γ -butyrolactones, principaux allergènes des plantes de la famille des *Asteraceae*, modifient principalement les résidus lysine [7]. Dans le même ordre d'idée, les dérivés dinitrohalogénés tels que le DNCB (dinitrochlorobenzène) ou DNFB (dinitrofluorobenzène), sont connus pour réagir facilement et rapidement avec les groupes lysine alors que le DNTB (dinitrothiocyanobenzène) réagit préférentiellement avec les résidus thiols. Il est à noter que, contrairement aux deux premiers dérivés qui sont des allergènes forts, le DNTB ne sensibilise que dans des conditions « dures » (en présence d'adjuvants immunostimulants) et se comporte plutôt comme un tolérogène lors d'une simple application épicutanée. Cette différence majeure dans l'activation du système immunitaire, liée, semble-t-il, à des différences de réactivité et de sélectivité, reste pour l'instant assez mal comprise.

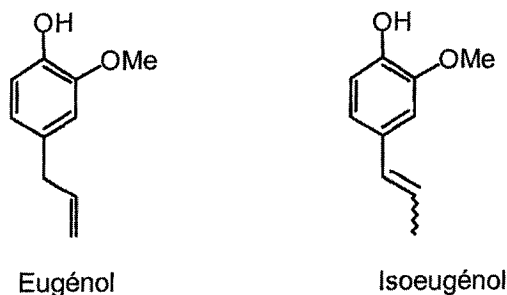


Figure 1. Structure de l'eugénol et de l'isoeugénol.

Importance des mécanismes de réaction

La sélectivité des haptènes pour certains acides aminés est directement liée à la densité électronique sur le site de réaction (déterminée par la structure de la molécule) et par la nature de la réaction chimique qui va se dérouler. Les réactions chimiques sont en effet classées en plusieurs groupes en fonction de leurs caractéristiques et des mécanismes de rupture et de formation des liaisons. Ainsi des molécules de structures très voisines peuvent donner lieu, par exemple lors de la métabolisation, à la formation d'intermédiaires différents qui conduiront à des mécanismes de réaction différents. Le cas de l'eugénol et de l'isoeugénol (figure 1) est à ce titre très intéressant.

Ces deux molécules, présentes dans de nombreux extraits naturels, ont en effet une structure très voisine, mais un potentiel allergisant très différent. Alors que l'eugénol est un sensibilisant faible, l'isoeugénol est un sensibilisant relativement fort. Des études récentes réalisées chez la souris semblent indiquer que ces deux molécules très semblables ont en fait un métabolisme et des mécanismes de réaction très différents [4], qui pourraient expliquer les différences observées dans le potentiel de sensibilisation. Ainsi, l'eugénol semble métabolisé en ortho-quinone électrophile après une étape de déméthylation (figure 2) alors que l'isoeugénol semble oxydé directement en paraméthylènequinone (figure 3).

Conséquences sur le pouvoir sensibilisant

Une conséquence directe de la sélectivité des haptènes pour les acides aminés va être l'existence de différences dans les sites de modification des protéines et dans la densité de ces modifications. Ces

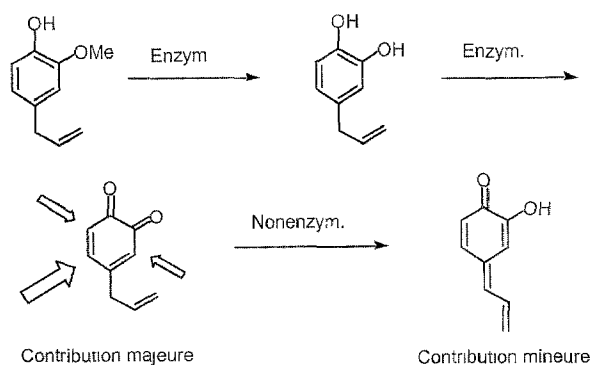


Figure 2. Mécanisme proposé pour la sensibilisation à l'eugénol chez la souris. Les flèches indiquent les centres réactifs potentiels.

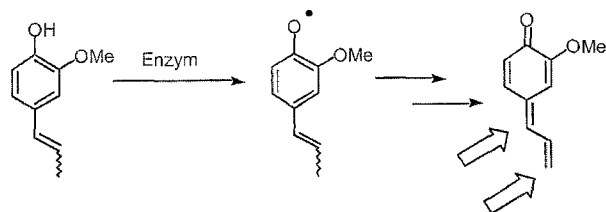


Figure 3. Mécanisme proposé pour la sensibilisation à l'isoeugénol chez la souris. Les flèches indiquent les centres réactifs potentiels.

différences, au départ purement chimiques, semblent de plus en plus avoir un impact majeur sur la réponse du système immunitaire. Ainsi, la sélectivité des sites d'hapténisation va participer directement à la sélection des fragments peptidiques qui vont être présentés par les cellules présentatrices d'antigènes aux lymphocytes T, et donc à la sélection des récepteurs lymphocytaires. Cette sélectivité va également contrôler indirectement le taux d'hapténisation de la ou des protéine(s). Or il semble que la densité des épitopes à la surface des cellules présentatrices d'antigènes influe, directement ou indirectement, sur l'orientation de la réponse immunitaire vers un profil Th1 ou Th2. Ainsi, de fortes densités d'épitope semblent orienter la réponse vers un profil Th1 alors que de faibles densités semblent orienter la réponse vers un profil Th2 [10].

BASES DE DONNÉES SUR L'ALLERGIE DE CONTACT

On peut considérer plusieurs types de bases de données en allergie de contact. Le cas le plus souvent

rencontré est celui des bases de type « personnel » qui contiennent toutes les informations allergologiques d'une consultation privée ou d'un service hospitalier. L'un des meilleurs exemples est sans doute celui du service de dermato-allergologie de Leuven [5] qui a développé un logiciel d'aide au diagnostic très complet. Dans le cas qui nous intéresse, l'objectif est un peu différent. Il s'agit de répertorier, dans une base de donnée aussi complète que possible, tous les cas d'allergie de contact rapportés dans la littérature en introduisant pour chacun d'eux le maximum d'informations. L'orientation d'une telle base n'est plus clinique mais « chimique ».

Développement

La base de donnée actuellement la plus complète est sans doute celle développée à Palo-Alto par CCS Associate en collaboration avec Maibach de l'université de San Francisco et Benezra [3]. Cette base de donnée est alimentée principalement par les cas d'allergie publiés dans *Contact Dermatitis* depuis 1975. Pour chaque allergène, les informations compilées et introduites dans la base sont :

- l'identification : nom, structure chimique, numéro de *Chemical Abstract* ;
- les données allergologiques : sensibilisant primaire ou éliciteur, nombre de cas rapportés, état du patient (sujet avec eczéma, autre pathologie, sujet sain) ;
- les données techniques : type de test (Open Epicutaneous Test, patch test), concentration, véhicule, intensité de la réaction, tests sur l'humain ou sur un modèle animal, nombre de sujets contrôles, référence de *Contact Dermatitis* ;
- l'évaluation : degré de confiance.

Ce dernier point mérite sans doute quelques explications. Le degré de confiance est chiffré de 0 à 5 et représente une évaluation de la pertinence de l'information. Il ne s'agit en aucun cas de porter un jugement sur le travail de l'expérimentateur mais plutôt d'évaluer le poids que doit avoir cette information dans la base de donnée. Il s'agit d'un point important si l'on veut utiliser de telles bases dans un sens prédictif. Les critères d'appréciation du degré de confiance sont par exemple la présence de contrôles, la pureté des composés testés, le nombre de cas, etc.

La recherche dans la base de donnée peut ensuite s'effectuer par nom, numéro de *Chemical Abstract*, structure chimique et par tous les paramètres biologiques définis précédemment. Ils peuvent bien entendu être utilisés seuls ou en combinaison. Cette exploitation des données est assurée par logiciel Prophet. Il s'agit d'une arborescence hiérarchique basée sur des structures et des sous-structures chimiques choisies pour leur importance en allergie de contact. Ce type d'approche, développé dans les années 1970 pour la création de bases de données sur les produits carcinogènes, a été étendu à l'allergie de contact. L'arborescence de ce système est actuellement basée sur 250 structures chimiques.

L'utilité principale d'un tel système est de répondre à des questions du type :

- Est-ce qu'un composé appartient à un groupe allergisant ?
- Quels sont les composés susceptibles de donner des réactions en allergie croisée ?
- Quelle concentration et quel véhicule utiliser pour tester ce produit ?
- Quelle est la pertinence de ces réponses ?

Limitations

Les principales limitations de ce système proviennent de son mode de constitution. En effet, les données sont extraites de la littérature et donc soumises au « biais » inhérents aux publications scientifiques. Par exemple, une réaction sévère sur quelques individus sera souvent plus largement documentée que des réactions modérées chez un plus grand nombre de patients.

SYSTÈMES EXPERTS

Les systèmes dits « experts » sont des programmes d'aide à la décision. Ils reposent sur l'exploitation de larges bases de données et doivent être capables, à partir des données structurales, physicochimiques, électroniques, etc., d'un nouveau composé et par comparaison relative avec les données historiques contenues dans sa mémoire, de faire une recommandation de décision concernant le statut allergisant d'une molécule. Les nouveaux paramètres ainsi définis doivent aller alimenter en retour la base de donnée et améliorer le système expert.

Les différentes approches

Dans le domaine des systèmes experts, de très nombreux paramètres peuvent être considérés.

Structures d'alerte

Il s'agit de structures chimiques totales ou partielles, connues pour présenter des risques allergisants. Elles sont principalement définies par l'analyse de la littérature et les bases de données sur l'allergie de contact. Elles permettent au système expert la détection d'un « risque ». L'un des systèmes basés sur ce principe, actuellement sur le marché, est nommé DEREK (pour *Deductive Estimation of Risk from Existing Knowledge*).

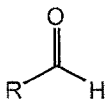
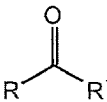
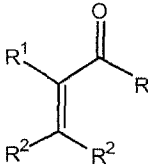
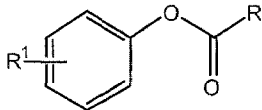
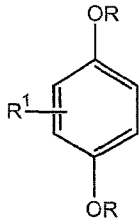
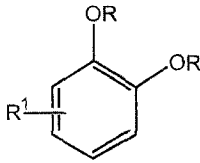
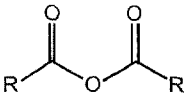
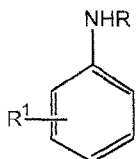
Ce système expert comprend un logiciel de « contrôle » chargé d'analyser la structure des molécules et une base de donnée comprenant des « règles » sous formes de sous-structures connues pour être associées à une activité allergisante [2]. Cette approche peut être combinée à des techniques analytiques pour identifier les molécules potentiellement sensibilisantes présentes dans des mélanges complexes tels que les parfums ou les déodorants. Ainsi, lors de l'analyse de 71 déodorants, 84 molécules sur les 226 identifiées comportaient au moins une structure d'alerte et 70 pouvaient être métabolisées en molécules comportant au moins une structure d'alerte [14]. Les principales structures d'alerte utilisées lors de cette étude sont représentées dans le *tableau I*.

Indicateurs de risque

Ces indicateurs sont des paramètres physicochimiques qui sont en faveur d'un potentiel allergisant. Chaque paramètre positif est considéré comme un facteur aggravant sans que l'intensité de cette aggravation ne soit prise en compte. On peut citer par exemple comme indicateurs de risque :

- présence d'une ou de plusieurs fonctions chimiques réactives ;
- indice de réfraction molaire ;
- nombre de liaisons hydrogènes ;
- nombre de paires électroniques sur oxygène ou azote ;
- nombre de NH ou OH ;
- nombre de groupes ionisables ;
- nombre de structures aromatiques ;

Tableau I. Structures d'alerte utilisées pour l'évaluation du potentiel sensibilisant des molécules identifiées dans les déodorants.

<i>Groupe chimique</i>	<i>Substituants</i>	<i>Structure</i>
Aldéhydes	R = alkyle, aryle	
Cétones	R, R' = alkyle, aryle	
Aldéhydes, cétones, esters et amides α,β insaturés	R = H, C, N, O (sauf OH) R' = sauf hétéroatomes, esters, cétones R ² = pas aryl sauf si R = H	
Esters de phényle	R = alkyle, aryle R' = tous	
Hydroquinones et précurseurs O-alkylés	R = H, alkyle R' = tous	
Catéchols et précurseurs O-alkylés	R = H, alkyle R' = tous	
Anhydrides	R = tous	
Amines aromatiques primaires et secondaires	R = alkyle, aryle R' = tous	

- nombre d'hétérocycles ;
- nombre de groupes métabolisables, etc.

Chaque paramètre peut être affecté d'un coefficient plus ou moins important selon son influence dans le mécanisme de l'allergie. La prise en compte de tous ces facteurs doit permettre d'évaluer le risque potentiel. Le développement de tels systèmes experts en est à son tout début et va nécessiter une longue phase de validation pour distinguer les paramètres réellement importants de ceux qui sont anecdotiques.

Limitations

Une des principales limitations de ce système est que, pour le moment, le processus d'expertise s'effectue à partir de la structure de la molécule testée qui n'est pas forcément celle qui va réagir avec les protéines cutanées. Les processus de métabolisation ne sont pas pris en compte.

RELATIONS STRUCTURE-ACTIVITÉ QUANTITATIVES (QSAR)

L'objectif principal de l'approche dite QSAR est de définir un certain nombre de « quantificateurs » pour une réaction biologique donnée puis de les combiner pour faire une prévision quantitative de l'activité biologique. Il ne s'agit plus uniquement d'obtenir une réponse du type « molécule potentiellement allergisante » mais d'avoir une idée assez précise de l'intensité attendue pour la réponse biologique. Le quantificateur actuellement le plus développé est sans aucun doute le *Relative Alkylation Index* (RAI) [6].

Définition du RAI

Ce quantificateur représente une estimation du taux de modification des protéines par une molécule potentiellement allergisante et suppose que l'intensité de la réponse biologique est directement liée à ce taux de modification. Il s'agit d'une hypothèse simple prenant en compte trois paramètres : la dose du produit appliqué sur la peau, sa lipophilicité et sa réactivité chimique. Ces paramètres ont été choisis car il est connu expérimentalement depuis longtemps que la réponse allergisante augmente avec chacun de

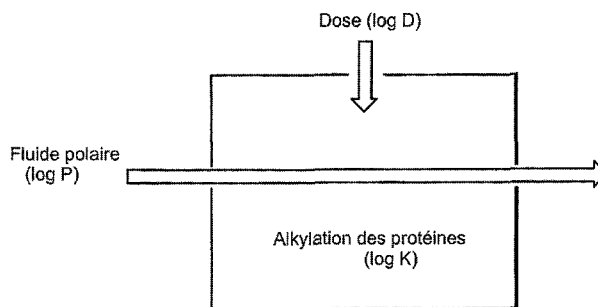


Figure 4. Schéma général du *Relative Alkylation Index* (RAI).

ces paramètres. Dans le modèle du RAI, l'épiderme est divisé en une suite d'unités de volume fini. Ces volumes, de nature principalement lipophile, sont parcourus par un fluide polaire et contiennent un grand nombre de protéines (figure 4).

La quantité de produit entrant dans une unité de volume est supposée proportionnelle à la dose du produit appliqué sur la peau. Elle est exprimée par le logarithme de la dose D . Le temps de séjour dans l'unité de volume des molécules à étudier est proportionnel à leur caractère lipophile quantifié par le logarithme du coefficient de partage P . Celui-ci représente l'affinité relative de l'allergène dans un mélange octanol/eau. Plus le produit est lipophile et donc son $\log P$ important, plus il séjournera longtemps dans l'unité de volume. La réactivité chimique de l'allergène vis-à-vis des protéines est quant à elle exprimée par le logarithme de sa constante cinétique k vis-à-vis d'un nucléophile modèle : la butylamine. Il est supposé que cette réaction est du premier ordre, c'est-à-dire que les protéines sont toujours en excès par rapport à l'allergène.

Il est alors possible d'exprimer le taux relatif de modification des protéines par l'expression :

$$\text{RAI} = a \log k + b \log P + c \log D$$

où a , b et c représentent le poids relatif de chaque paramètre. Pour l'instant, ces constantes sont déterminées expérimentalement et ne sont valables que pour une série de molécules ayant le même mécanisme de réaction. C'est pour cela que le RAI est défini comme un index relatif d'alkylation.

L'intensité de la réaction d'allergie augmente avec chacun de ces paramètres et donc avec le RAI

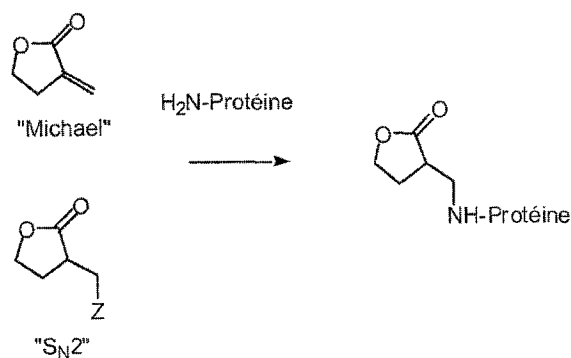


Figure 5. Réaction des lactones α - β insaturées et substituées avec les protéines.

jusqu'à un certain point appelé point de surcharge (*overload*). Ce point, qui est une constatation expérimentale, est encore assez mal expliqué mais correspond vraisemblablement à une saturation des mécanismes biologiques ainsi qu'à la mise en place des mécanismes de régulation.

Application

Une série de molécules dont la structure est voisine de celles des α -méthylène- γ -butyrolactones ont été synthétisées et étudiées [8, 9]. Ces molécules forment par réaction avec les protéines le même type d'épitope que celui obtenu par réaction avec l' α -méthylène- γ -butyrolactone correspondante (*figure 5*).

Les différents substituants ont été choisis pour couvrir une gamme de réactivité significative. Ces molécules permettent donc d'une part d'observer l'influence de la réactivité chimique sur l'activité biologique et d'autre part de développer le RAI en faisant varier simultanément la constante cinétique et les facteurs physicochimiques.

Avant d'évaluer l'activité biologique de ces composés, nous avons évalué leur réactivité relative vis-à-vis de la butylamine (constante cinétique k) choisie comme modèle de nucléophile. Les coefficients de partage P sont quant à eux déterminés par le calcul.

L'activité allergisante a été testée chez le cobaye et ces composés se sont révélés être de bons allergènes. Les premières modélisations en utilisant les cons-

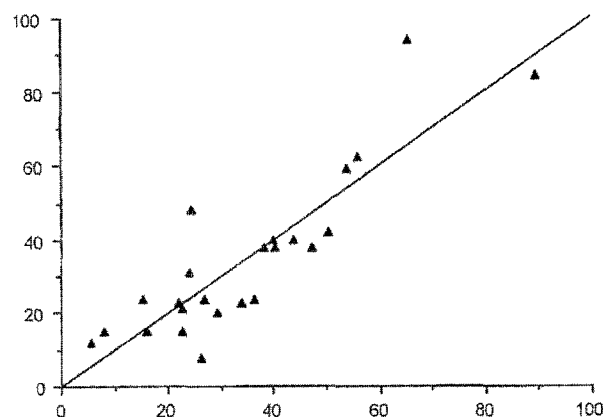


Figure 6. Corrélation entre réponses observées et mesurées dans le cas des lactones substituées.

tantes cinétiques déterminées précédemment et les $\log P$ calculés indiquent une bonne corrélation pour un RAI égal à :

$$\log D + 2 \log k + 0,5 \log P$$

À partir de cette définition du RAI, il est ensuite possible de définir l'expression de la réaction biologique :

$$\text{Réponse biologique} = a \text{ RAI}_i - b (\text{RAI}_i)^2 + c \text{ RAI}_c + d$$

où a , b et c sont des coefficients positifs et d une constante correspondant à une série donnée de composés. Dans notre cas, si la réponse biologique est exprimée en pourcentage de la réponse maximum observable, et en partant des résultats biologiques obtenus lors de cette étude, nous arrivons à une expression du type :

$$\% \text{ de réponse} = 73,0 \text{ RAI}_i - 7,2 (\text{RAI}_i)^2 + 17,5 \text{ RAI}_c - 208$$

Comme on peut le constater sur le graphe de la *figure 6*, il existe une bonne corrélation entre les réponses biologiques observées et celles calculées à partir de l'équation ci-dessus.

Limitations

Le RAI présente actuellement deux principales limitations. Il s'agit d'un système relatif qui donne de bons résultats à l'intérieur d'une série de composés

analogues, réagissant avec les protéines en suivant un même mécanisme, mais il ne permet pas encore d'évaluer le potentiel allergisant de molécules appartenant à deux familles très différentes. La seconde limitation est que ce modèle mathématique ne tient pas compte pour le moment des facteurs de pénétration cutanée dans l'évaluation de la dose de produit arrivant effectivement sur le site de réaction.

CONCLUSION

Nous sommes encore au tout début du développement des relations structure-activité en allergie de contact alors qu'elles ont déjà fait leurs preuves dans d'autres domaines. Du fait du mécanisme biologique complexe mis en jeu, le nombre des paramètres à prendre en compte est beaucoup plus important. Néanmoins, les quelques approches « pilotes » développées sont très prometteuses et devraient concurrencer ou compléter les tests alternatifs « biologiques ».

RÉFÉRENCES

- 1 Barbier A, Rizova E, Stampf JL, Lacheretz F, Pistor FHM, Bos JD, et al. Development of a predictive in vitro test for detection of sensitizing compounds (European BRIDGE project). In: *In Vitro Skin Toxicology (Irritation, Phototoxicity, Sensitization)*. New York : Liebert Inc. 1994.
- 2 Barratt MD, Basketter DA, Chamberlain M, Admans GD, Langowski JJ. An expert system rulebase for identifying contact allergens. *Toxic In Vitro* 1994 ; 8 : 1053-60.
- 3 Benezra C, Sigman CC, Perry IR, Helmes TC, Maibach HI. A systematic search for structure-activity relationships of skin contact sensitizers. *J Invest Dermatol* 1985 ; 85 : 351-6.
- 4 Bertrand F, Basketter DA, Roberts DW, Lepoittevin JP. 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.
- 5 Doms-Goossens A, Doms M, Drieghe J. Computers and patients information systems. In: Tycroft RJG, Menne T, Frosch PJ, eds. *Textbook of contact dermatitis*. Berlin: Springer Verlag; 1995. p. 769-84.
- 6 Fragnals R, Roberts DW, Lepoittevin JP, 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.
- 7 Franot C, Benezra C, Lepoittevin JP. Synthesis and interaction studies of ¹³C labeled lactone derivatives with a model protein using ¹³C NMR. *Bioorg Med Chem* 1993 ; 1 : 389-97.
- 8 Franot C, Roberts DW, Smith RG, Basketter DA, Benezra C, Lepoittevin JP. Structure-activity relationships for contact allergenic potential of γ,γ -dimethyl- γ butyrolactone derivatives. *Chem Res Toxicol* 1994 ; 7 : 297-306.
- 9 Franot C, Roberts DW, Basketter DA, Benezra C, Lepoittevin JP. Structure-activity relationships for contact allergenic potential of γ,γ -dimethyl- γ -butyrolactone derivatives. *Chem Res Toxicol* 1994 ; 7 : 307-12.
- 10 Hosken NA, Shibuya K, Heath AW, Murphy KM, O'Garra A. The effect of antigen dose on CD4+ T helper cell phenotype development in a T cell receptor- α β -transgenic model. *J Exp Med* 1995 ; 182 : 1579-84.
- 11 Kalish RS. Antigen processing: the gateway to the immune response. *J Am Acad Dermatol* 1995 ; 32 : 640-52.
- 12 Lepoittevin JP, Basketter DA, Goossens A, Kalberg AT, eds. *Allergic contact dermatitis: the molecular basis*. Berlin, Heidelberg: Springer-Verlag; 1997.
- 13 Lepoittevin JP, Benezra C. ¹³C-Enriched methylalkanesulfonates: new lipophilic methylating agents for the identification of nucleophilic amino acids of protein by NMR. *Tetrahedron Letters* 1992 ; 33 : 3875-8.
- 14 Rastogi SC, Lepoittevin JP, Johansen JD, Frosch PJ, Menné T, Bruze M, et al. Fragrances and other materials in deodorants: search for potentially sensitizing molecules using combined GC-MS and structure-activity relationships (SAR) analysis. *Contact Dermatitis* 1998 ; 39 : 293-303.