Université de Montréal

Développement d'approches quantitatives de type "structure-activité" pour la modélisation pharmacocinétique

par

Martin Béliveau

Département de santé environnementale et santé au travail

Faculté de médecine

Thèse présentée à la Faculté des études supérieures

en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)

en Santé Publique

option Toxicologie de l'environnement

Août, 2004

Université de Montréal

Faculté des études supérieures

Cette thèse intitulée:

Développement d'approches quantitatives de type "structure-activité" pour la modélisation pharmacocinétique

présentée par Martin Béliveau

a été évaluée par un jury composé des personnes suivantes:

Docteur Claude Viau, président-rapporteur

Docteur Kannan Krishnan, directeur de recherche

Docteur Robert Tardif, co-directeur de recherche

Docteur Michel Gérin, membre du jury

Docteur Donald Mackay, examinateur externe

Sommaire

Les modèles pharmacocinétiques à base physiologique (PCBP) permettent, entre autres, de prédire la concentration sanguine ou tissulaire d'un xénobiotique à partir de la dose d'exposition. Les équations algébriques et différentielles qui constituent un modèle PCBP contiennent plusieurs paramètres qui peuvent être classés en trois groupes: les paramètres physiologiques, les paramètres physicochimiques, et les paramètres biochimiques. Pour de nombreuses espèces on peut retrouver dans la littérature les valeurs des paramètres physiologiques. Par contre, les valeurs des paramètres physicochimiques et biochimiques sont souvent inexistantes. En effet, leur détermination peut être relativement complexe et exiger la réalisation d'études *in vivo* ou *in vitro*. Une alternative possible consiste à estimer pour une substance la valeur de ces paramètres à partir de l'information portant sur sa structure moléculaire à l'aide d'une approche "structure-activité" (QSAR). Ce projet de recherche visait à développer et à évaluer des approches QSAR pour simuler la dose interne des xénobiotiques dans une perspective d'analyse du risque toxicologique. Les objectifs spécifiques étaient: : i) d'utiliser les approches actuelles en QSAR pour estimer les paramètres pharmacocinétiques qui régissent la dose interne afin d'utiliser la structure moléculaire directement en analyse du risque toxicologique, ii) de développer une méthodologie QSAR selon les approches conventionnelles englobant plusieurs familles de substances chimiques, ce qui permettrait de prédire le profil pharmacocinétique d'une substance chez le rat pour un scénario d'exposition donné, iii) de développer une méthodologie QSAR selon les approches conventionnelles qui permettrait de prédire le profil pharmacocinétique d'une substance chez l'humain pour un scénario d'exposition donné, iv) d'extrapoler le profil pharmacocinétique d'une espèce à l'autre en utilisant une méthodologie QSAR qui relie la structure chimique aux paramètres régissant la cinétique

des substances, v) de développer une approche QSAR permettant de relier directement la structure moléculaire à un indice de concentration interne chez le rat, soit la concentration sanguine à l'état stationnaire, vi) d'extrapoler celle-ci à l'humain à l'aide d'une approche QSAR reliant la structure moléculaire et le mécanisme régissant la concentration sanguine à l'état stationnaire, soit la clairance systémique. Premièrement, nous avons appliqué une approche QSAR déjà existante (le Free-Wilson des chloroéthanes) à un contexte d'analyse du risque pour le 1,1,1-trichloroéthane. Pour la première fois, une approche QSAR a servi à estimer les paramètres pharmacocinétiques qui déterminent la concentration interne utilisée en analyse du risque toxicologique. Deuxièmement, une approche QSAR-PCBP originale décrivant la relation entre la structure moléculaire de composés organiques volatils appartenant à plusieurs familles (halométhanes, haloéthanes, haloéthylènes, hydrocarbures aromatiques) et les paramètres pharmacocinétiques, a été développée et validée chez le rat, ce qui a permis d'établir une relation "structure – profil pharmacocinétique" pour des scénarios d'exposition ne menant pas à un état stationnaire. Troisièmement, cette approche QSAR-PCBP a été appliquée à l'humain, ce qui a permis d'établir une relation "structure – profil pharmacocinétique" pour des scénarios d'exposition ne menant pas à un état stationnaire chez l'humain. Quatrièmement, un QSAR qui relie la structure aux paramètres régissant la cinétique des substances à été développé, ce qui a permis d'extrapoler le profil pharmacocinétique des concentrations tissulaires d'une espèce à l'autre en utilisant l'information sur la physiologie de l'organisme. Cinquièmement, une approche reliant la structure moléculaire et la concentration sanguine à l'état stationnaire a été développée chez le rat. Sixièmement, les concentrations stationnaires ont été extrapolées à l'humain en utilisant les QSAR spécifiques au mécanisme déterminant ces concentrations, soit la clairance systémique. Les différents modèles QSAR-PCBP chez le rat et l'humain ont été validés en comparant les résultats obtenus à des concentrations sanguines expérimentales

obtenues de la littérature. Ce projet a démontré pour la première fois que l'utilisation de la structure moléculaire en analyse de risque toxicologique est possible, et permet de déterminer des critères de priorisation lors du développement et de l'étude de nouvelles substances chimiques, en plus de nous aider ainsi à mieux comprendre l'impact de la structure moléculaire sur la toxicité des xénobiotiques.

Mots clés: Modélisation PBPK, QSAR, structure, pharmacocinétique, état stationnaire, paramètres pharmacocinétiques, composés volatils, approches *in silico*, clairance.

Abstract

Physiologically-based pharmacokinetic (PBPK) models are tools that are useful for predicting the blood and tissue concentrations of a chemical from information on the external dose. The algebraic and differential equations on which PBPK models are based are solved using physiological, physicochemical and biochemical parameters. Whereas information on the value of physiological parameters can be found in the literature for many species, physicochemical and biochemical parameter values for many chemicals are often lacking. This data gap is most often filled by obtaining experimental values using *in vivo* or *in vitro* methodologies. However, a more efficient alternative is to directly relate the structure of the compound of interest to the value of these chemical-specific parameters using quantitative structure-activity relationships (QSARs) and chemical structure information. The objective of this study was to develop QSAR methodologies that could be used as tools for simulating the internal concentration of chemicals within different species, so they can be integrated within a risk assessment framework. The specific objectives were to: i) use current QSAR methodologies within a risk assessment framework by relating chemical structure to the pharmacokinetic (PK) parameters that determine internal concentrations, ii) develop a novel QSAR methodology that can relate the structure of volatile organic chemicals belonging to multiple families to their corresponding PK profile in rat according to the exposure scenario, iii) apply this novel QSAR methodology to the prediction of the PK profile in humans, iv) extrapolate the PK profile from one species to another by developing a QSAR methodology that relates molecular structure to the chemical-specific component of PK parameters, v) establish the relationship between molecular structure and steady-state concentrations in rats, and vi) extrapolate those steady-state concentrations to humans by relating structure to the clearance parameters involved.

First, an established QSAR methodology (a Free-Wilson approach that describes chloroethanes) was used within a risk assessment framework. For the first time, a QSAR was used to estimate the internal concentrations corresponding to established safe levels by relating structure to pharmacokinetic parameters. Second, a novel QSAR-PBPK approach that related the structure of chemicals belonging to various families (halomethanes, haloethanes, haloethylenes and aromatic hydrocarbons) to chemical-specific pharmacokinetic parameters was developed and validated in rats. This QSAR-PBPK framework established a structure–PK profile relationship for non steady-state exposure scenarios. Third, this QSAR-PBPK framework was applied to humans by predicting the PK profile using the relationship between structure and human specific PK parameters. Fourth, internal concentrations and PK profiles were extrapolated from one species to another using QSARs specific to the chemical component of each model parameter. Fifth, an approach that directly related chemical structure and steady-state internal concentrations was developed in rats. Sixth, the steady-state concentrations were extrapolated to humans by using the relationships between structure and the underlying parameter involved: the systemic clearance. Therefore, the internal concentrations in multiple species, whether they resulted from scenarios leading to steady-state conditions or not, were related to molecular structure using this novel QSAR-PBPK methodology. The use of these molecular structure-internal concentration relationships within risk assessment should facilitate research priorities when dealing with the development of new compounds or the study of existing ones, as well as help to understand the potential impact of chemical structure on human health.

Keywords: PBPK models, QSAR, QSPR, structure, pharmacokinetics, steady-state, pharmacokinetic parameters, VOCs, *in silico* approaches, clearance.

Table des Matières

Liste des tableaux

Chapitre II

Chapitre V

Chapitre VI

Chapitre VII

Liste des figures

Chapitre II

Chapitre III

Chapitre IV

Chapitre V

Liste des sigles et des abréviations

1,1-DCE	1,1-dichloroéthane
1,2-DCE	1,2-dichloroéthane
ACSL®	Logiciel de simulation ("Advanced Continuous Simulation Language")
AUC	Surface sous la courbe
BEN	Benzène
BS	Structure de base
C_a	Concentration artérielle
C_{air}	Concentration dans l'air
Ca_{ss}, Cb_{ss}, SS-BC	Concentration artérielle (ou sanguine) à l'état stationnaire
C_b	Concentration de produit lié
CE	Chloroéthane
CFM	Chloroforme
C_{inh}, C_i	Concentration inhalée
CL	Clairance
CL_h	Clairance hépatique
CL_{int}	Clairance intrinsèque
CL_p	Clairance pulmonaire
CM	Chlorométhane
CMD	Fichier du programme ACSL®
COV, VOC	Composé organique volatil
C_p	Concentration de protéines
CP, PC	Coefficient de partage

C_s, C_f, C_i	Contribution du fragment ou concentration inhalée
CSL	Fichier du programme ACSL®
C_t	Concentration dans le tissu t
CT	Tétrachlorure de carbone
C_v	Concentration veineuse
C_{vt}	Concentration veineuse sortant du tissu t
CYP2E1	Cytochrome P-450 de la famille 2E1
CYP450, CYP	Cytochrome P-450
dA_{met}/dt, RAM	Taux de changement de la quantité métabolisée
dA_t/dt	Taux de changement de la quantité dans le tissu t
DCM	Dichlorométhane
DE_{50}	Dose effective causant une réponse donnée chez 50% des individus
D_l	Coefficient pour la diffusion dans la fraction lipidique (couche cornée)
D_p	Coefficient pour la diffusion dans la fraction protéinique (couche cornée)
E	Coefficient d'extraction hépatique
EBZ	Ethylbenzène
F	Fréquence
F_b	Capacité de liaison protéinique
F_{nelt}, F_{nl}	Fraction équivalente en lipide neutre dans le tissu
F_p	Fraction de protéine dans le sang
F_s, f_i	Fréquence de la présence du fragment dans la molécule
F_u	Fraction libre dans le sang
F_{wet}, F_w	Fraction équivalente en eau dans le tissu

K_a	Constante de liaison protéinique
$K_{a(abs)}$	Constante d'absorption orale
K_f, K_{fc}	Constante d'élimination de premier ordre
K_m	Constante d'affinité Michaëlis-Menten
K_p	Constante de perméabilité dermale
LFE	Linear Free-Energy
MW	Poids moléculaire
N	Nombre de sites de liaison
NOAEL, DSENO	Dose sans effet nocif observé
PAH	Hydrocarbure aromatique polycyclique
P_b, $P_{b:a}$, $P_{b:a(app)}$	Coefficient de partage sang:air (apparent)
PCBP, PBPK	Pharmacocinétique à base physiologique
PCBs	Biphényles polychlorés
$P_{he:a}$	Coefficient de partage hexadécane:air
PK	Pharmacocinétique
pK_a	Constante d'acidité (logarithme)
$P_{o:a}$	Coefficient de partage huile végétale:air
$P_{o:w}$	Coefficient de partage octanol:eau
$P_{p:a}$	Coefficient de partage protéine:air
$P_{p:w}$	Coefficient de partage protéine:eau (couche cornée)
P_{PBPK}	Paramètre d'un modèle PCBP
PRESS	Somme des carrés de l'erreur résiduelle de prédiction
PSA	Surface polaire
P, P_t, $P_{t:b}$	Coefficient de partage tissu:air ou tissu:sang (indice l = foie, s ou m = muscles, f = gras, et r = rapidement perfusés)

$P_{t:a}$	Coefficient de partage tissu:air
$P_{vo:w}$	Coefficient de partage huile végétale:eau
$P_{w:a}$	Coefficient de partage eau:air
Q_c	Débit cardiaque
Q_l	Perfusion sanguine du foie
Q_p	Ventilation pulmonaire
QSAR, QSPR	Relation "structure-activité" ou "structure-propriété"
Q_t	Perfusion sanguine du tissu t
R^2	coefficient d'ajustement
SPSS	Logiciel statistique ("Statistical Package for the Social Sciences")
SR	Indice de sensibilité
SS	État stationnaire
SSY	Somme du carré des réponses
STY	Styrène
TCE	1,1,1-trichloroéthane
TMB	Triméthylbenzène
TOL	Toluène
UNIFAC	Description de la structure chimique standardisée
V_m, V_x	Volume molaire
V_{max}, V_{maxc}	Vélocité maximum du métabolisme
V_{nb}	Fraction de lipides neutres dans le sang
V_{nt}	Fraction de lipides neutres dans le tissu
V_{pb}	Fraction de phospholipides dans le sang
V_{pt}	Fraction de phospholipides dans le tissu
V_t	Volume du tissu t

V_{wb}	Fraction d'eau dans le sang
V_{wt}	Fraction d'eau dans le tissu
XYL	m-Xylène
Δt	Pas d'intégration

Tous mes rêves ne sont possibles que grâce
à deux personnes particulières qui n'ont jamais
cessées d'être fières de leur fils.
Merci à mes parents.
Je vous aime.

Remerciements

Je tiens à remercier les docteurs Kannan Krishnan et Robert Tardif, pour m'avoir donné la chance de poursuivre mes travaux à leur cotés. Si séparément ils sont travaillants, attentifs, et compétents, vous ne pouvez imaginer l'effet de synergie sur un étudiant lorsqu'ils travaillent ensemble. Ils ne lâchent jamais prise, même dans les moments difficiles. Merci à vous.

Mes remerciements vont aussi à Mme Ginette Charest-Tardif et Mme Pierrette Gagnon, pour leur appui technique et leur aide précieuse.

De plus j'aimerai remercier Dr Sami Haddad, Dr Patrick Poulin, Dre Annie St-Pierre, M. Jérôme Lavoué, ainsi que les équipes des docteurs Krishnan et Tardif. Les amitiés formées avec ces personnes m'ont permis de continuer ma recherche de savoir.

Finalement, et non le moindre, j'aimerai remercier ma douce moitié, mon épouse Stéphanie Rossy. Elle reste, simplement, le soleil de ma vie.

CHAPITRE PREMIER:

1 Introduction générale

1.1 Introduction à la modélisation pharmacocinétique à base physiologique

La modélisation pharmacocinétique à base physiologique (PCBP) consiste à décrire mathématiquement l'absorption, la distribution, le métabolisme et l'élimination de substances dans l'organisme. Les équations différentielles qui y sont utilisées sont basées sur la physiologie de l'organisme, ainsi que sur les caractéristiques physico-chimiques des substances. Ces équations permettent la prédiction de la cinétique des substances dans un organe cible (Krishnan et Andersen 2001). Les relations entre les paramètres qui servent d'intrants dans ces équations, c'est-à-dire les paramètres physiologiques (débit alvéolaire, débit cardiaque, volumes tissulaires et débits sanguins aux tissus), physico-chimiques (coefficients de partage, coefficients de diffusion, et constantes d'absorption), et biochimiques (taux d'élimination, constantes de liaisons et métaboliques) déterminent le devenir d'une substance dans un organisme.

Ainsi, les modèles PCBP permettent de prédire la concentration sanguine ou tissulaire d'une substance à partir de la dose d'exposition. Puisque la toxicité d'une substance est considérée comme étant directement liée à sa concentration tissulaire (Andersen *et al.* 1987), ces modèles sont très utilisés en analyse du risque toxicologique. Ainsi, grâce à la modélisation PCBP, la détermination de la concentration d'une substance dans l'organe cible en fonction du temps et de la concentration d'exposition devient possible. De plus, le recours à une représentation mathématique de mécanismes biologiques permet de tenir compte de plusieurs phénomènes comme la saturation des processus de biotransformation et son influence sur la concentration tissulaire.

L'utilisation des modèles PCBP en analyse du risque toxicologique présente certains avantages. Premièrement, puisqu'ils incorporent des mécanismes biologiques qui facilitent le calcul de la concentration interne, leur utilisation permet de relier celle-ci plutôt que la

dose d'exposition à l'effet observé, ce qui réduit l'incertitude reliée aux méthodes traditionnelles d'extrapolation des doses d'expositions. Ainsi, le scénario d'exposition, la voie d'absorption, la dose et l'espèce peuvent être considérés lors d'extrapolations (Krishnan et Andersen 2001). Deuxièmement, la modélisation PCBP combinée à des études toxicocinétiques chez l'animal permet de mieux comprendre les mécanismes responsables de la toxicocinétique d'une substance donnée. En effet, dans certains cas le profil cinétique généré par le modèle PCBP ne correspond pas aux données déterminées expérimentalement. Cela peut être dû à d'autres processus biologiques, qui doivent être incorporés dans le modèle, ou alors cela peut indiquer que certains paramètres du modèle doivent être modifiés afin de mieux refléter la réalité expérimentale (Clewell et Andersen 1987; Haddad *et al.* 1998). En outre, les hypothèses qui sont ainsi générées peuvent être ensuite vérifiées expérimentalement.

Un modèle PCBP est constitué de groupements tissulaires, ou compartiments, représentant les tissus de l'organisme où la substance chimique est distribuée. Le degré de regroupement tissulaire (c.-à-d., le nombre de compartiments) dépend de la substance à l'étude et de l'objectif du modèle (Krishnan et Andersen 2001). Chaque compartiment tissulaire est décrit mathématiquement par une équation différentielle de bilan de masse, où le taux de changement de la quantité d'une substance dans le compartiment est fonction du taux d'entrée et du taux de sortie du compartiment. Selon les caractéristiques chimiques de la substance d'intérêt (p.ex., composé volatil ou non), la représentation fonctionelle (c.-à-d., les équations mathématiques) peut varier.

Puisque l'absorption des composés organiques volatils (COV) de faible poids moléculaire n'est pas limitée par leur diffusion, mais plutôt par la perfusion sanguine au tissu, le taux d'entrée dans le compartiment dépend de la concentration artérielle (C_a) et du débit sanguin au tissu (Q_t), alors que le taux de sortie dépend de la concentration veineuse

sortant du tissu (C_{vt}) et du débit sanguin au tissu. Le bilan de masse, dans ce cas, est calculé comme suit:

$$\frac{\delta A_t}{\delta t} = Q_t C_a - Q_t C_{vt} \quad (1)$$

où $\delta A_t/\delta t$ = taux de changement de la quantité de produit dans le tissu (mg/h)

L'équation 1 représente un tissu homogène et présume que la concentration veineuse tissulaire est en équilibre avec la concentration tissulaire, un équilibre qui est déterminée par le coefficient de partage tissu:sang (P_t).

Lorsque le compartiment représente un tissu pouvant métaboliser la substance, l'équation de bilan de masse utilisée comporte un terme additionnel qui décrit le taux de clairance métabolique:

$$\frac{\delta A_t}{\delta t} = Q_t C_a - Q_t C_{vt} - \frac{\delta A_{met}}{\delta t} \quad (2)$$

où $\delta A_{met}/\delta t$ = taux de métabolisme (mg/h).

Le terme $\delta A_{met}/\delta t$ dans l'équation 2 peut représenter un processus métabolique de premier ordre:

$$\frac{\delta A_{met}}{\delta t} = K_f C_{vt} V_t \quad (3)$$

où K_f = constante métabolique de premier ordre (h^{-1}), et

V_t = volume du tissu (L).

Pour une biotransformation de premier ordre dans le foie, l'équation suivante, basée sur le concept de clairance hépatique (CLh = Ql*E), peut être substituée à l'équation 3 (Poulin et Krishnan 1998):

$$\frac{\delta A_{met}}{\delta t} = Q_l \cdot E \cdot C_a \quad (4)$$

où Q_l = débit sanguin au foie (L/h), et

E = coefficient d'extraction hépatique $[CL_{int}/(Q_l+CL_{int})]$

CL_{int} = clairance intrinsèque (V_{max}/K_m)

V_{max} = vélocité maximale du métabolisme (mg/h), et

K_m = constante d'affinité (mg/L)

Par contre, si le processus métabolique est saturable, l'équation suivante est utilisée pour représenter le taux de métabolisme:

$$\frac{\delta A_{met}}{\delta t} = \frac{V_{max} C_{vt}}{K_m + C_{vt}} \tag{5}$$

Les différents compartiments tissulaires sont reliés par la circulation systémique. Le sang artériel amène la substance vers les tissus, alors que les concentrations veineuses tissulaires se regroupent pour former le sang veineux.

Certains modèles PCBP de COV n'ont pas de compartiment distinct qui représente le sang. La concentration artérielle est plutôt représentée comme la solution, à l'équilibre, de l'équation du bilan de masse représentant le compartiment pulmonaire, où les échanges gazeux ont lieu. On considère qu'il y a équilibre lorsque toute la quantité de substance qui quitte l'air inspiré passe vers le sang et que le tissu pulmonaire distribue la vapeur entre l'espace alvéolaire et le sang artériel en fonction du coefficient de partage sang:air. Dans ce cas l'équation suivante peut être dérivée:

$$C_a = \frac{Q_p C_{inh} + Q_c C_v}{Q_c + \frac{Q_c}{P_b}} \tag{6}$$

où P_b = coefficient de partage sang:air

Q_p = ventilation pulmonaire (L/h)

C_{inh} = concentration de substance inhalée (mg/L)

Q_c = débit cardiaque (L/h)

L'équation 6 reste valide tant que les trois conditions suivantes sont rencontrées: 1) il y a un équilibre rapide de la substance entre le sang et l'espace alvéolaire, 2) il n'y a aucune biotransformation significative dans le tissu pulmonaire, et 3) la capacité de stockage des poumons est négligeable. Le système sanguin veineux (C_v), quant à lui, est représenté comme la solution à l'équilibre de l'équation de bilan de masse du compartiment veineux:

$$C_v = \frac{\sum_t^n Q_t C_{vt}}{Q_c} \qquad (7)$$

Une fois la description des compartiments tissulaires et des liens entre les compartiments complétée, on doit procéder à l'estimation des paramètres qui sont inclus dans les équations mathématiques du modèle PCBP.

1.2 Estimation des paramètres

Les équations algébriques et différentielles qui constituent un modèle PCBP sont résolues à l'aide de plusieurs paramètres qui peuvent être classés selon trois groupes: les paramètres physiologiques (volumes tissulaires, taux de perfusion sanguin, débit cardiaque, taux de ventilation alvéolaire), les paramètres physicochimiques (coefficient de partage sang:air, coefficient de partage tissu:sang, constante d'absorption orale, coefficient de perméabilité tissulaire), et les paramètres biochimiques (vélocité maximale, constante d'affinité Michaëlis-Menten).

Les caractéristiques physiologiques d'un organisme sont représentées dans un modèle PCBP par les débits sanguins et pulmonaires, ainsi que les volumes tissulaires. Pour de nombreuses espèces (p. ex., poisson, rat, souris, ou humain) on peut retrouver la valeur expérimentale de ces paramètres dans la littérature (Caster *et al.* 1956; Domench *et al.*

1969; Arms et Travis 1988; Hetrick *et al.* 1991; Ross *et al.* 1991; Brown *et al.* 1997). Lorsque les données requises ne sont pas disponibles dans la littérature, on peut les estimer à l'aide de certaines équations allométriques validées (Krishnan et Andersen 1991). Il est aussi possible de déterminer les valeurs requises de façon expérimentale en laboratoire (Mauderly 1990; Delp *et al.* 1991). Les valeurs des paramètres physicochimiques et biochimiques sont, quant à elles, plus difficile à obtenir de la littérature et sont souvent inexistantes.

Les valeurs des paramètres physicochimiques et biochimiques, nécessaires à l'élaboration d'un modèle PCBP pour une substance particulière, peuvent être obtenues expérimentalement à partir d'approches *in vivo* ou *in vitro*.

1.2.1 APPROCHES *IN VIVO*

Les approches *in vivo* impliquent la collecte de données pharmacocinétiques chez un animal exposé à la substance, suivi de l'analyse de ces données avec un modèle PCBP. En ajustant les paramètres manquant du modèle de façon à ce que celui-ci décrive le mieux possible les données expérimentales, la valeur d'un paramètre peut ainsi être estimée. Par contre, une telle procédure n'est efficace que si au plus un ou deux paramètres doivent être estimés. L'estimation de la valeur de ces paramètres par une telle procédure *in vivo*, surtout en ce qui concerne les produits non-volatils, peut cependant s'avérer un processus lourd qui nécessite une grande quantité d'animaux (Krishnan et Andersen 2001). Néanmoins, cette procédure est fréquemment utilisée dans le cas des paramètres biochimiques (Andersen *et al.* 1980; Gargas *et al.* 1986; Gargas et Andersen 1989; Gargas *et al.* 1990).

Les méthodes expérimentales pour déterminer les P_t *in vivo* se basent sur le calcul du rapport des concentrations tissulaires et sanguines une fois à l'équilibre. Deux méthodes

sont couramment réalisées selon le type de substance étudiée. Des expositions par inhalation sont employées dans le cas des COV, puisque cette voie d'administration se prête bien aux produits volatils (Gargas *et al.* 1989). Par contre, dans le cas des composés non-volatils, l'administration intraveineuse en infusion constante ou en bolus est employée (Chen et Gross 1979; Lam *et al.* 1982; Lin *et al.* 1982; Gallo *et al.* 1987). Le P_t est un paramètre important qui détermine la distribution tissulaire d'une substance; il correspond au rapport entre les coefficients de partage tissu:air ($P_{t:a}$) et sang:air (P_b). Les valeurs des coefficients de partage tissu:air ($P_{t:a}$) ou du P_b sont souvent déterminées sous certaines conditions *in vitro*.

1.2.2 APPROCHES *IN VITRO*

Les approches *in vitro* qui permettent, entre autre, de limiter l'utilisation d'animaux, n'ont été utilisées jusqu'à présent que pour estimer les coefficients de partage. Les constantes métaboliques déterminées *in vitro* ne peuvent être incorporées directement dans les modèles PCBP, même si à cet égard l'utilisation d'hépatocytes isolés et de fraction post-mitochondriales semble prometteuse (Krishnan et Andersen, 2001).

Les $P_{t:a}$ ou P_b peuvent être déterminés *in vitro* selon plusieurs méthodes expérimentales. Dans le cas des COV, on utilise préférablement la méthode dite des "flacons à l'équilibre", qui consiste à mesurer la distribution "relative" d'une quantité connue de substance entre deux phases, la première constituée par un homogénat de tissu ou du sang, et la deuxième, l'air contenu dans le flacon maintenu à 37°C. Le rapport des quantités dans les deux phases correspond au coefficient de partage tissu:air ou sang:air. Le rapport entre ces deux coefficients permet de déterminer les coefficients tissu:sang (Sato et Nakajima 1979; Gargas *et al.* 1989; Kaneko *et al.* 1994; Tardif *et al.* 1997). Dans le cas des substances non

volatiles, la méthode de dialyse à l'équilibre (Lin *et al.* 1982; Law *et al.* 1991; Murphy *et al.* 1995; Haddad *et al.* 1998) ou la méthode d'ultrafiltration à l'aide des composés marqués (Lin *et al.* 1978; Jepson *et al.* 1994), sont utilisées. Ces deux méthodes sont basées sur l'équilibre entre un homogénat de tissu et une solution tampon, et les concentrations de substance marquée dans les deux phases sont mesurées par scintillation liquide.

Toutes les approches *in vitro* mentionnées ci-haut nécessitent le recours à du tissu animal.

Afin de réduire le nombre d'animaux utilisés, les recherches se sont tournées vers le développement de solutions de rechange, particulièrement les approches dites *in silico*, pour estimer la valeur des paramètres des modèles PCBP. Deux types d'approches *in silico* sont utiles dans ce contexte. La première concerne l'utilisation des données portant sur les paramètres disponibles dans la littérature afin de développer des équations mathématiques qui associent les caractéristiques des substances chimiques à la valeur des paramètres. L'approche QSAR (relation structure-activité quantitative) classique en est un exemple. La deuxième approche, *in silico*, concerne l'estimation de la valeur des paramètres des modèles PCBP à l'aide d'algorithmes mécanistes qui tiennent compte des interrelations qui existent entre les déterminants chimiques et biologiques. Avant d'aller plus loin, nous présenterons un bilan de l'état des connaissances concernant les approches *in silico* (QSARs et algorithmes mécanistes) qui ont été proposées pour estimer la valeur de certains des paramètres des modèles PCBP.

1.3 Méthodologie générale des approches in silico

La section suivante présente brièvement certains aspects méthodologiques sur lesquels se fondent les deux type d'approches *in silico*, soit les approches QSAR et les algorithmes mécanistes.

1.3.1 QSAR

Cette approche consiste à relier une activité biologique, ou de façon plus spécifique et en ce qui nous concerne une *propriété* (le coefficient sang:air, par exemple), à des caractères structuraux qui sont spécifiques à une substance à partir d'une fonction mathématique (*f*):

$$\text{Propriété biologique} = f(\text{caractère structurel}) \quad (8)$$

Puisque des données empiriques sont utilisées afin de dériver la fonction mathématique qui décrira correctement la relation décrite plus haut, les fonctions résultantes peuvent être linéaires, multilinéaires ou supralinéaires, suivant la nature des données expérimentales. On retrouve deux type de QSAR dans la littérature qui permettent d'estimer la valeur des paramètres des modèles PCBP: les modèles "Linear-free energy" (LFE) et les modèles "Free-Wilson".

1.3.1.1 Modèles de type LFE

Les modèles de type LFE sont en réalité des relations quantitatives qui décrivent une activité biologique en fonction de la structure moléculaire, sur la base de principes thermodynamiques (Hansch et Leo 1995). À la base de cette approche courante de Hansch figure le principe voulant que les différences observées dans l'ampleur d'une activité

biologique parmi une série de substances chimiques correspondent à un changement dans l'énergie libre (ΔG) du système impliqué. Puisque la différence dans l'activité biologique et le changement d'énergie libre sont souvent reliés de façon linéaire, la relation mathématique qui en découle est nommée "linear free energy relationship". Puisqu'il est très difficile de déterminer directement ΔG dans un système biologique, c'est plutôt ses composantes thermodynamiques d'énergie (ΔE), d'enthalpie (ΔH) et d'entropie (ΔS), qui sont utilisées, et elles sont représentées par une série de descripteurs structuraux qui peuvent être dérivés pour chaque substance et qui sont spécifiques (Seydel et Schaper 1982). Les coefficients de ces descripteurs (c.-à-d., pente et segment sur l'axe des ordonnées) sont ensuite analysés, par régression, en utilisant des techniques statistiques normales. Dans ce type d'approche, les descripteurs structuraux sont classés selon trois types de catégories: électrostatiques, stériques ou hydrophobes. Les modèles LFE peuvent incorporer des descripteurs appartenant à une ou plusieurs de ces catégories, et ce en fonction de la signification statistique (p.ex., la valeur de *p*) de chaque caractère structurel dans le model final.

1.3.1.1.1 CARACTÈRES ÉLECTROSTATIQUES DANS LES MODÈLES DE TYPE "LFE"

Les caractères structuraux de catégorie "électrostatique" sont typiquement ceux qui décrivent des tendances à accepter ou à donner des électrons, des charges atomiques partielles et la densité des champs électrostatiques. Ils sont définis par les constantes de Hammet (σ), les paramètres de résonance ("R values"), les paramètres inductifs ("F values") et les valeurs de Taft (ρ*, σ*, E_s) (Hansch et Leo 1995). Puisque les molécules ionisées peuvent traverser difficilement (par diffusion passive) les membranes biologiques et que les constantes d'effets électrostatiques sont dérivées à partir des caractéristiques d'ionisation d'une molécule, une relation entre le comportement pharmacocinétique de

toutes les substances chimiques et des caractères strictement électrostastiques n'apparait pas appropriée.

1.3.1.1.2 CARACTÈRES STÉRIQUES DANS LES MODÈLES DE TYPE "LFE"

Les effets stériques sont typiquement représentés par des valeurs calculées pour la réfractivité molaire et le paramètre stérique de Taft. Cependant, puisque les effets stériques décrivent "l'encombrement" d'une molécule, la description des caractères stériques peut inclure le volume moléculaire, la masse moléculaire, la surface moléculaire, le niveau d'arborescence de la chaîne de carbone, etc. (Hansch et Leo 1995). Les indices de connectivité moléculaires et les caractères dérivés d'une analyse QSAR "3D" peuvent aussi être considérés comme étant de nature stérique, même si dans certains cas la détermination de la relation entre la structure et ces caractères est plutôt intuitive et souvent obscure. De plus, l'obtention de la valeur de ces caractères pour une substance spécifique requiert souvent l'utilisation de logiciels spécialisés en modélisation chimique.

1.3.1.1.3 CARACTÈRES HYDROPHOBES DANS LES MODÈLES DE TYPE "LFE"

Les effets hydrophobes que l'on incorpore dans les équations des modèles de type LFE sont fréquemment représentés par le log (en base 10) du coefficient de partage octanol:eau (Po:w) ou le caractère hydrophobe π, qui est lui-même est dérivé du Po:w (Hansch et Leo 1995). Mais d'autres coefficients de partage (p.ex., eau:air [Pw:a]. huile vegétale:air [Po:a], huile végétale:eau, n-hexadécane:air [Phe:a]) et coefficients de solubilité ont aussi été utilisés. Le caractère hydrophobe (octanol:air, huile:air, ou eau:air) a souvent servi à relier la structure à certains paramètres des modèles PCBP. L'utilisation de données expérimentales regroupées chez plusieurs espèces (rat, humain, poisson) dans des régressions a mené à une grande variation dans les valeurs des coefficients Po:w, Po:a, ou

Pw:a. Puisque les coefficients de partage (CP) sang:air ou tissu:air représentent la distribution d'une substance entre une matrice biologique (qui est constituée principalement de lipide et d'eau) et l'air, il semble naturel que de bonnes corrélations soient obtenues en utilisant une mesure pour une autre matrice appropriée et l'air (c.à.d., Pw:a ou Po:a). De plus, puisque les tissus sont principalement composés d'eau, de lipides et de protéines, on a suggéré que les coefficients des caractères hydrophobes représentant ces diverses composantes reflétaient la composition tissulaire (Abraham et Weathersby 1994).

1.3.1.2 Les modèles de type "Free Wilson"

Même si la relation entre les paramètres pharmacocinétiques et les caractères hydrophobes a été fréquemment explorée, le développement de telles relations basées sur d'autres types de caractères dans une approche LFE n'est pas facile. La signification de ces caractères tels qu'ils sont utilisés dans les équations LFE est fréquemment obscure, surtout en ce qui a trait à l'élucidation des mécanismes qui gouvernent les processus pharmacocinétiques. Il faut aussi une base de données substantielle pour résoudre ces équations et explorer de façon satisfaisante toute les combinaisons structurelles possibles, ce qui s'avère souvent impossible à réaliser. Ainsi, certaines alternatives à l'approche LFE ont été explorées. Par exemple, une approche consiste à développer une série de constantes à partir de la relation entre l'activité biologique et la nature, ainsi que l'occurrence, de groupes fonctionnels dans une molécule mère (Free et Wilson 1964). Cette méthodologie peut être décrite par la relation mathématique suivante:

$$\text{Activité} = A + \Sigma_i \Sigma_j G_{ij} X_{ij} \qquad (9)$$

où A est défini comme la moyenne de l'activité biologique de la série de substances, G_{ij} comme la contribution à l'activité d'un groupe fonctionnel **i** à la position **j** et X_{ij} la présence (1.0) ou l'absence (0.0) du groupe fonctionnel **i** à la position **j** dans la molécule.

L'approche Free-Wilson suppose que la contribution des groupes fonctionnels est additive et qu'une base de donnée assez large soit disponible pour déterminer la contribution des groupes fonctionnels. Des QSAR de type Free-Wilson ont été développés dans le domaine de la pharmacocinétique pour relier la structure moléculaire des sulfamidés à leur taux d'absorption orale ($K_{a\,(oral)}$) (Seydel et Schaper 1982). En se basant sur la valeur des contributions des groupes fonctionnels ainsi dérivée, il est possible de prédire le composé ayant la valeur de $K_{a\,(oral)}$ la plus élevée, simplement par addition des fragments dont la contribution est la plus importante. Récemment, des algorithmes Free-Wilson qui relient structure et paramètres de modèles PCBP ont été développés chez le rat, l'humain et le poisson pour une série de chloroéthanes (Fouchécourt et Krishnan 2000; Fouchécourt *et al.* 2000). Les algorithmes développés ont ensuite été intégrés dans un modèle PCBP afin de simuler la cinétique de ces substances chez plusieurs espèces. Le facteur limitant d'une telle approche, par contre, est que le modèle Free-Wilson développé pour les chloroéthanes ne peut pas être utilisé pour prédire la valeur des paramètres de substances qui n'ont pas la structure de base et les groupes fonctionnels communs à la famille étudiée. En revanche, une telle limite peut être surmontée par l'utilisation d'algorithmes de type mécanistes.

1.3.2 ALGORITHMES MÉCANISTES

Contrairement aux approches *in silico* décrites plus haut, les algorithmes "mécanistes" sont basés sur des considérations biologiques et ne requierent pas de données expérimentales au départ. Des informations portant sur les processus biologiques

spécifiques qui déterminent la valeur d'un paramètre PCBP sont recueillies et des relations mathématiques entre la valeur du paramètre PCBP et le déterminant biologique sont établies. Ces relations doivent être capables de prédire la valeur d'un paramètre PCBP, et les valeurs prédites peuvent ensuite êtres comparées aux valeurs expérimentales afin de valider l'algorithme. L'incertitude concernant l'estimation de la valeur des paramètres pour de nouvelles substances est ainsi réduite, parce que l'algorithme est basé sur des mécanismes biologiques connus. Lorsque les valeurs estimées diffèrent des valeurs expérimentales, des hypothèses concernant la présence d'autres mécanismes impliqués peuvent être générées et ces mécanismes peuvent êtres subséquemment incorporés de façon mathématique dans l'algorithme afin d'être vérifiés. Théoriquement, ce type d'algorithme peut être développé pour n'importe quel paramètre de modèle PCBP et est indépendant de la famille chimique ou de la structure moléculaire. Le développement et l'applicabilité de tels algorithmes ne sont pas limités que par notre compréhension des processus biologiques et mécanismes qui déterminent la valeur d'un paramètre.

Présentement, plusieurs QSAR (à la fois de type LFE et Free-Wilson) et algorithmes mécanistes qui facilitent l'estimation des paramètres de modèles PCBP spécifiques à la substance sont disponibles dans la littérature. Toutes ces approches *in silico* seront abordées plus en détails dans les sections suivantes.

1.4 Problématique

Les valeurs des paramètres physicochimiques et biochimiques, nécessaires pour le développement des modèles PCBP, sont présentement obtenues à partir d'études *in vivo* ou *in vitro*. Une alternative possible consiste à estimer la valeur des paramètres spécifiques à une substance, comme les constantes physicochimiques et biochimiques, à partir de

l'information sur sa structure moléculaire. Les approches *in silico* permettant d'estimer la valeur des paramètres des modèles PCBP se sont centrées sur le développement de QSAR de type LFE et d'algorithmes mécanistes. Alors que les QSAR de type LFE ont l'avantage d'être facilement dérivés, ils sont par contre limités à la famille chimique dont ils sont dérivés. Une grande quantité d'information expérimentale est ainsi requise, ce qui réduit l'applicabilité de l'algorithme. En outre, les estimations des paramètres obtenues ne peuvent être extrapolées entre espèces différentes. Plusieurs doutes existent aussi concernant la pertinence, d'un point de vue mécaniste, de certains caractères structurels employés dans ce type d'équations. Les nouvelles approches mécanistes offrent l'avantage de ne pas être limitées à une famille chimique, et de plus, elles peuvent êtres extrapolées à d'autres espèces, en spécifiant les valeurs des contenus en lipides et en eau de l'espèce étudiée. Par contre, les valeurs de Pw:a et Po:w spécifiques à la structure, et nécessaires à ces algorithmes mécanistes, ne peuvent être dérivées que par des approches qui fragmentent la structure de façon différente. Différentes banques de fragments sont donc nécessaires, selon l'approche considérée (Baum 1998; Boethling et Mackay 2000; Sangster 1997). De plus, il n'existe pas d'approches mécanistes présentement pour l'estimation des paramètres biochimiques. Les QSAR de type Free-Wilson peuvent être utilisés pour établir des liens directs entre le nombre ou la nature des fragments moléculaires et la valeur d'un paramètre physicochimique et biochimique intégré dans un modèle PCBP. Cela a déjà été démontré avec les chloroéthanes. L'étude de cas présentée dans l'article 1 suggère que le développement de modèles de type PCBP-QSAR, dans lequel le nombre et/ou la nature des fragments moléculaires serait le seul élément variable qui permettrait une simulation de la cinétique de la série, est possible. Par contre, les QSAR Free Wilson ne peuvent, tout comme les QSAR de type LFE, que prédire la valeur des paramètres pour les substances d'une même famille (p.ex., les chloroéthanes) et sont donc limités dans leur application.

Si les paramètres spécifiques à une substance requis pour élaborer un modèle PCBP sont reliés à la structure, cela suggère que les QSAR peuvent êtres intégrés à un modèle PCBP et cela permet de relier la structure d'une substance à sa pharmacocinétique. Puisque le profil pharmacocinétique est généré en utilisant seulement l'information concernant la structure de la substance comme intrant au modèle, un modèle QSAR-PCBP peut également simuler le profil pharmacocinétique des substances pour différents scénarios d'exposition. En considérant les limites des approches existantes discutées précédemment, toute approche PCBP-QSAR ne pourra être appliquée avec un certain degré de succès en analyse de risque que si elle démontre: i) qu'elle peut être généralisable à plusieurs familles de substances chimiques, ii) qu'elle peut être extrapolée à diverses espèces, et iii) qu'elle permet d'extrapoler les concentrations tissulaires ou sanguines obtenues à diverses doses d'exposition. Le développement de ce type de QSAR, qui estime les paramètres d'accumulation d'une substance selon l'espèce, permettrait de simuler la cinétique selon des scénarios d'exposition anticipés pour de nouvelles substances. Par conséquent, la dose interne qui correspond à la DSENO ou au risque unitaire pourrait être prédite et extrapolée d'une espèce à l'autre, et ce, en utilisant seulement l'information sur la structure moléculaire d'une substance. Le processus d'extrapolation entre espèces s'en trouverait amélioré du fait qu'il considèrerait la dose interne.

1.5 Objectifs

Ce projet de recherche vise à évaluer les approches QSAR présentement disponibles et d'en proposer de nouvelles permettant de les utiliser en analyse du risque toxicologique. Une approche originale développée permettant d'utiliser les modèles existants et le développement d'approches QSAR permettant de simuler la dose interne seront proposées.

Cet objectif s'appuie sur i) l'utilisation de la modélisation pharmacocinétique comme outil fiable d'estimation permettant de prédire la cinétique des composés dans l'organisme et; ii) le développement d'approches permettant de relier la structure moléculaire à la dose interne, de façon directe ou par l'intermédiaire de ses paramètres pharmacocinétiques. Les objectifs spécifiques sont i) d'utiliser les approches actuelles en QSAR pour estimer les paramètres pharmacocinétiques qui régissent la dose interne afin d'utiliser la structure moléculaire directement en analyse du risque toxicologique, ii) de développer une méthodologie QSAR englobant plusieurs familles de substances chimiques, ce qui nous permettra de prédire le profil pharmacocinétique d'une substance chez le rat selon divers scénarios d'exposition, iii) de développer une méthodologie QSAR qui nous permettra de prédire le profil pharmacocinétique d'une substance chez l'humain selon divers scénarios d'exposition, iv) d'extrapoler le profil pharmacocinétique d'une espèce à l'autre en utilisant une méthodologie QSAR qui relie la structure aux paramètres régissant la cinétique des substances, v) de développer une approche QSAR permettant de relier directement la structure moléculaire à un indice de concentration interne chez le rat, soit la concentration sanguine à l'état stationnaire, vi) extrapoler celle-ci vers l'humain à l'aide d'une approche QSAR reliant la structure moléculaire et le mécanisme régissant la concentration sanguine à l'état stationnaire, soit la clairance systémique, tout ceci en vue d'utiliser plus efficacement la structure moléculaire en analyse du risque toxicologique.

1.6 Démarches expérimentales et organisation de la thèse

Premièrement, une approche QSAR que l'on peut retrouver dans la littérature (le Free-Wilson des chloroéthanes) sera utilisée dans un contexte d'analyse du risque (Article I). Pour la première fois, une approche QSAR Free-Wilson sera utilisée afin de prédire les

paramètres pharmacocinétiques d'un modèle PCBP et ce modèle sera utilisé en analyse du risque toxicologique (Objectif 1).

Deuxièment (Article II), une approche QSAR-PCBP originale concernant la relation structure moléculaire – paramètres pharmacocinétiques, qui tient compte des limites des approches actuelles, sera développée et validée chez le rat, ce qui permettra d'établir une relation structure – profil pharmacocinétique pour les scénarios d'exposition qui ne mènent pas à l'état stationnaire. Cette approche sera appliquée à plusieurs familles de substances chimiques (Objectif 2).

Troisièmement (Article III), l'approche QSAR-PCBP développée et validée chez le rat dans l'article I sera appliquée à l'humain, ce qui permettra d'établir une relation structure – profil pharmacocinétique pour les scénarios d'exposition qui ne mènent pas à l'état stationnaire chez une autre espèce que le rat (Objectif 3).

Quatrièmement (Article IV), la structure d'une substance sera reliée aux mécanismes derrière les paramètres régissant la cinétique des substances. Ceci permet d'extrapoler le profil pharmacocinétique d'une espèce à l'autre en utilisant seulement l'information sur la physiologie et en gardant l'information sur la chimie constante à travers les espèces (Objectif 4).

Cinquièmement (Article V), puisque souvent les valeurs limites d'exposition en analyse du risque toxicologique se basent des concentrations sanguines à l'état stationnaire, une approche reliant directement la structure moléculaire à la concentration sanguine à l'état stationnaire sera développée chez le rat (Objectif 5).

Sixièmement (Article VI), les concentrations sanguines à l'état stationnaire seront extrapolées vers l'humain à l'aide d'une approche QSAR reliant la structure moléculaire et le mécanisme régissant la concentration sanguine à l'état stationnaire, soit la clairance systémique (Objectif 6).

Par la réalisation de ce dernier objectif de la thèse, les doses internes chez plusieurs espèces, quelles soient à l'état stationnaire ou non, auront été reliées à la structure moléculaire par l'utilisation de cet outil novateur qu'est la modélisation QSAR-PCBP.

Finalement, le dernier chapitre de la thèse présente une discussion générale sur l'implication de cette recherche en analyse du risque toxicologique. Les limites d'utilisation de l'approche développée ainsi que les applications possibles seront discutées.

CHAPITRE DEUXIÈME:

2 Article I

In silico approaches for physiologically-based pharmacokinetic modeling

Beliveau, M., Krishnan, K. (2003). In silico approaches for physiologically-based pharmacokinetic modeling. Dans: *Animal replacement approaches in toxicology*, Ed. Salem and Katz.

In silico approaches for physiologically-based pharmacokinetic modeling

Martin Béliveau and Kannan Krishnan

Groupe de recherche en toxicologie humaine (TOXHUM)

Faculté de médecine, Université de Montréal,

Case Postale 6128, Succursale Centre-ville

Montréal, P.Q., Canada, H3C 3J7

2.1 Introduction

Physiologically-based pharmacokinetic (PBPK) models are increasingly being used for conducting dose-, route-, species- and exposure scenario-extrapolations required for risk assessments (Andersen *et al.*, 1987). PBPK models are basically mechanism-based mathematical descriptions of the processes of absorption, distribution, metabolism, and excretion in the intact organism (Krishnan and Andersen, 2001). The algebraic and differential equations constituting the PBPK models are solved with the knowledge of various input parameters, namely, physiological (tissue volumes, blood flow rates, cardiac output, alveolar ventilation rate), physicochemical (blood:air partition coefficients, tissue:blood partition coefficients, absorption rate constants, permeability coefficients), and biochemical (maximal velocity, Michaelis affinity constant) parameters. Whereas the information on physiological parameters can be obtained from the biomedical literature (Arms and Travis, 1988), this is frequently not the case for physicochemical (partition coefficients, absorption constants, and permeability coefficient) and biochemical parameters (hepatic or renal clearances, maximal velocity of metabolism, and Michaelis affinity constant).

The physicochemical and biochemical parameters needed for constructing chemical-specific PBPK models can be obtained using *in vivo* or *in vitro* approaches. *In vivo* approaches involve collection of pharmacokinetic data in exposed animals and analysis of such data using a PBPK model. By adjusting the model simulations to match the experimental data, the numerical values of the missing parameters can be estimated. Such a procedure is reliably applied for estimating one or two parameters at a time. Parameter estimation using *in vivo*

studies, particularly for non-volatiles, can be tedious and can require extensive use of animals (Krishnan and Andersen, 2001).

The *in vitro* methods, facilitating reduced animal use, have been proven to be useful only for estimating partition coefficients. The *in vitro* derived metabolism constants cannot be directly incorporated within PBPK models, even though freshly isolated hepatocytes and post-mitochondrial fractions appear to hold some promise (Krishnan and Andersen, 2001). The considerations of cost-effectiveness as well as reduction/replacement of animal use have led to the development of other alternative approaches, particularly *in silico* approaches, for estimating PBPK model parameters. Two kinds of *in silico* approaches are useful in this context. The first one involves the use of available data for various PBPK parameters in order to develop equations that associate characteristics of chemicals to the magnitude of the parameters. An example of this category is the classical QSAR (quantitative structure-activity relationship) approach. Another *in silico* approach involves the development of mechanistic algorithms based on an understanding of the interrelationships among certain biological and chemical determinants in order to predict the numerical value of PBPK model parameters. The objectives of this chapter are: (1) to review the state-of-the art of *in silico* approaches (QSARs, biologically-based algorithms) for estimating PBPK model parameters, and (2) to illustrate how the *in silico*-based PBPK models can be used in human health risk assessment applications.

2.2 *Methodological basis of* in silico *approaches*

The following paragraphs provide a brief description of the methodological basis of the two types of *in silico* approaches, namely QSARs and biologically-based algorithms.

2.2.1 QSARs

The QSARs typically relate a biological activity or more specifically, a *property* in this context, to structural features specific to chemicals through a mathematical function (*f*):

$$\text{Biological property} = f(\text{structural feature}) \qquad [1]$$

Since empirical data are used to derive the mathematical function fulfilling the above relationship, depending upon the nature of the data the functions can be linear, multilinear or supralinear. Two types of QSARs have been used to estimate the value of PBPK model parameters: linear-free energy (LFE) models and Free-Wilson models.

2.2.1.1 LFE-type models

LFE-type models are quantitative relationships that describe activity as a function of chemical structure, relying upon the principles of thermodynamics (Hansch and Fujita, 1964). The basis of the commonly used Hansch approach is that the differences in magnitude of a given biological activity within a series of chemicals correspond to changes in the free energy (ΔG) during the processes involved. As the difference in biological activity and the change in free energy are

likely to be linearly related, the resulting mathematical relationships are referred to as "linear free energy relationships". Since it is very difficult to directly determine Δ G in biological systems, it's thermodynamic components such as the energy (ΔE), enthalpy (ΔH) and entropy (ΔS) are used instead, and are represented by a series of structural descriptors which can be derived for any given molecule (Seydel and Schaper, 1982). The coefficients for these descriptors (i.e., slopes and intercepts) are then regressed using standard statistical techniques. In this type of approach, structural descriptors can be broadly classified into three general types: electrostatic, steric or hydrophobic. LFE models can incorporate one or many of these categories of structural descriptors, based on the statistical significance of each feature in the final model.

2.2.1.1.1 ELECTROSTATIC FEATURES IN LFE-TYPE MODELS

Electronic effects typically include electron donating and withdrawing tendencies, partial atomic charges and electrostatic field densities as defined by Hammett sigma (σ) values, resonance parameters (R values), inductive parameters (F values) and Taft substituent values (ρ^*, σ^*, E_s). Because ionized molecules cannot pass through biological membranes and electrostatic effects constants are derived from ionization characteristics of the molecule, relating pharmacokinetic behavior of all chemicals to only electrostatic features is not totally relevant. Abraham *et al.* (1994) used dipolarity/polarizability, among others, as electrostatic descriptors for relating structure to the numerical values of tissue:air and blood:air partition coefficients of a series of chemicals. They observed that water, plasma, blood, lung, kidney, muscle, brain, fat and olive oil, progressively

became less dipolar/polarizable due to increasing lipid content (0 to 100 %). Thus, the electrostatic descriptor is relevant only for functionally substituted compounds such as 1-propanol. Lewis and Dickins (2002) have shown the importance of electrostatic descriptors such as ionization potential and pK_a in relating structure of various drugs to metabolic rates and binding to CYP450. Correlations were improved with the incorporation of these descriptors, especially for ionizable or polar drugs. This is probably related to the binding site of the compound on the CYP protein, which contains polar amino acids.

2.2.1.1.2 STERIC FEATURES IN LFE-TYPE MODELS

Steric effects are conventionally represented by values calculated for molar refractivity and the Taft steric parameter. However, since steric effects describe the "bulkiness" of the molecule, they can include molecular volume, molecular weight, surface area, carbon chain branching, etc. Molecular connectivity indices and features derived from 3D QSAR can also be considered as being of a steric nature, although the relationship between structure and these features is often obscure and not as intuitive, in certain instances. Furthermore, obtaining the chemical-specific values for these features often requires the use of specialized chemical modeling software.

Gargas *et al.* (1988) used steric descriptors in order to relate structure to PBPK model parameter values, and they suggested that a LFE equation combining connectivity indices and *ad hoc* descriptors (such as the number of halogens in a compound) provided better descriptions for tissue:air partition coefficients than using either descriptors alone. Use of connectivity indices is limited because the

relationship between these descriptors and structure is not informative in a transparent and direct manner. For example, first order valence connectivity index represents both structural and electronic features of a compound, in a complex way. Order indices represent multiple substitution patterns of the halogens on the carbons in the compounds, flexibility in polymers, or halogen substitution patterns. More intuitive are the other descriptors used by Gargas *et al.* (1988), namely the number of halogen atoms present in the molecule, though these parameters may not always be present in all molecules of interest. Gargas *et al.* (1988) also used steric parameters to relate structure to maximal velocity of metabolism (V_{max}), but because of the high level of electronic information contained in the connectivity indices used in the descriptions, the authors suggested that more accurate modeling of V_{max} could be attempted by using both steric descriptors as well as more specific electronic information such as charge distribution.

2.2.1.1.3 HYDROPHOBIC FEATURES IN LFE-TYPE MODELS

Hydrophobic features in LFE-type equations are frequently represented by using log octanol:water partition coefficient (Log $P_{o:w}$) or the hydrophobic parameter, π, which is derived from $P_{o:w}$. However, other partition coefficients (e.g., water:air [$P_{w:a}$], oil:air [$P_{o:a}$], oil:water, n-hexadecane:air [$P_{he:a}$] partition coefficients) and solubility parameters have also been used. Hydrophobic parameters (namely, octanol:air, oil:air, or water:air) have been extensively used for relating structure and PBPK model parameters. The use of various datasets and experimental data in multiple species (rat, human or fish) in the regressions have led to varied coefficient values for $P_{o:w}$, $P_{o:a}$ or $P_{w:a}$. Because blood:air and

tissue:air partition coefficients (PC) represent distribution between a biological matrix (consisting mostly of lipid and water) and air, it is logical that reasonably good correlations are obtained using a partition measure for another relevant matrix and air (i.e., $P_{w:a}$ or $P_{o:a}$). Furthermore, since tissue is composed of water, lipids and protein, the coefficients of these descriptors have been suggested to reflect tissue composition (Abraham *et al.,* 1994). Recently, Meulenberg and Vijverberg (2000) after an extensive review of the literature, found that values of the coefficients obtained following regression analysis using $P_{w:a}$, $P_{o:a}$ and the experimental PC values for rats and humans were essentially the same as the tissue water and lipid content, highlighting the importance of tissue composition in the partitioning process.

2.2.1.2 Free Wilson-type models

Although the relationships between pharmacokinetic parameters and hydrophobic determinants have been explored frequently, the development of such relationships using other determinants in the LFE approach is not as straight forward. Furthermore, the mechanistic significance of many of these determinants as used in LFE equations is often unclear as to how they relate to PK processes, and in order to successfully explore all possible structural combinations, a substantial dataset is required and is often unavailable. Because of this, alternatives to the LFE approach have been explored. In this regard, Free and Wilson (1964) developed a series of substituent constants by relating biological activity with the nature and frequency of occurrence of specific functional groups in

the parent molecule. This methodological approach is reflected by the following equation:

$$\text{Activity} = A + \Sigma_i \Sigma_j G_{ij} X_{ij} \quad [2]$$

where A was defined as the average biological activity for the series, G_{ij} the contribution to activity of a functional group **i** in the jth position and X_{ij} the presence (1.0) or absence (0.0) of the functional group **i** in the jth position.

The Free Wilson approach requires that the contributions of substituents be additive and that a sufficiently large database be available to facilitate the determination of the contribution of various substituents. In the pharmacokinetic arena, the Free-Wilson type QSARs have been developed for the rate of oral absorption ($K_{a\ (oral)}$) of sulfonamides (Seydel and Schaper, 1982). Based on the fragment constants derived, it was possible to predict the compound with the highest $K_{a\ (oral)}$ value, by combining the fragments with the highest contributions. More recently, Free-Wilson algorithms for relating structure to PBPK model parameters for a series of chloroethanes in rats, humans and fish have been developed (Fouchécourt and Krishnan, 2000; Fouchécourt *et al.* 2000). These algorithms were then successfully integrated into a PBPK model in order to simulate the kinetics of these chemicals in the various species. The limiting factor of such an approach, however, is that the Free Wilson model developed for chloroethanes could not be used to predict the parameter values for chemicals lacking the common structure and the substituents. Such a limitation can be overcome with the development and use of biologically-based algorithms.

2.2.2 BIOLOGICALLY-BASED ALGORITHMS

Contrary to the *in silico* methods described above, biologically-based algorithms do not require *a priori* knowledge of experimental data. Here, information on specific biological processes that determine the magnitude of a PBPK parameter is gathered, and predictive mathematical relationship between the PBPK parameter and biological determinants is developed. The predictions of the algorithm are then compared with experimental data for validation purposes. Uncertainty regarding the prediction of parameter values for *de novo* compounds is somewhat reduced, because the algorithm is based on known biological mechanisms. In cases where predicted values differ from experimental data, hypotheses concerning other plausible mechanisms can be generated and incorporated within the algorithm for further verification. Theoretically, these types of algorithms can be developed for any PBPK parameter regardless of the chemical class or molecular structures. The development and application of such algorithms is only limited by the current level of understanding of the mechanistic basis and phenomena that determine the magnitude of the PBPK model parameters.

At the present time, several QSARs (LFE and Free-Wilson) and biologically-based algorithms are available to facilitate the prediction of chemical-specific PBPK model parameters. All of these *in silico* approaches, as detailed in the following section, have been uniquely applied to estimate the PBPK model parameters of organic substances.

2.3 In silico approaches for PBPK model parameters

2.3.1 IN SILICO APPROACHES FOR TISSUE:AIR PARTITION COEFFICIENTS

Tissue:air PCs describe the relative concentrations of volatile organic chemicals (VOCs) in tissues and air at steady-state. Both LFE-type QSAR models and biologically-based algorithms have been developed for predicting tissue:air partition coefficients of a variety of chemicals (Table 1).

In developing an LFE-type QSAR for human tissue:air PCs, Abraham *et al.* (1994) observed that non-polar solutes only needed hexadecane:air partitioning (a hydrophobic descriptor) whereas an electrostatic descriptor (π_2^H) was important for functionally substituted compounds such a 1-propanol. Abraham *et al.* (1994) correctly underscore the importance of limiting the use of these equations for interpolation, particularly for the kinds of chemicals considered during model development. Gargas *et al.* (1988) used connectivity indices in order to correlate structure with the rat tissue:air PCs of a series of haloalkanes. In this study, it was found that fluorine substituents reduced the tissue solubility, with the greatest effect being observed in biological matrices with greater proportion of water (e.g., blood). To the contrary, chlorine and bromine substituents increased solubility in all tissues. Because of the electronegativity of these atoms (F>Cl>Br) it was suggested that these atoms increased the solubility in the medium via dispersion interactions. In other words, these atoms increase lipophilicity characteristic of the molecule and it is therefore difficult to separate the purely "steric" from the purely "hydrophobic" influence because of the interdependency of these parameters as they relate to tissue solubility. As listed in Table 1, most of the published studies have established the quantitative relationship between tissue:air PCs and

descriptors such as $P_{w:a}$ and $P_{o:a}$, due to the importance of the solubility of chemicals in tissue water and tissue lipids. Currently, there do not exist any Free-Wilson type models for tissue:air PCs of VOCs.

Mechanistic algorithms for predicting tissue:air PCs have, however, been developed on the basis of the determinants of two components, namely, tissue water:air PC and tissue lipid:air PC. Whereas the tissue water:air PC is considered to be the same as the inverse of the Henry's law constant, the tissue lipid:air PC is assumed to be equivalent to $P_{o:w}$. Along these lines, Poulin and Krishnan (1996c) developed the following algorithm to predict tissue:air partition coefficients ($P_{t:a}$):

$$P_{t:a} = [P_{o:w} P_{w:a} (V_{nt} + 0.3V_{pt})] + [P_{w:a}(V_{wt} + 0.7V_{pt})] \quad [3]$$

where $P_{o:w}$ = n-octanol:water partition coefficient, $P_{w:a}$ = water:air partition coefficent, V_{nt} = volume fraction of neutral lipid in tissue, V_{pt} = volume fraction of phospholipid in tissue, and V_{wt} = volume fraction of water in tissue.

In the above equation, $P_{o:w}$ and $P_{w:a}$ can be directly estimated from knowledge of molecular structure (Hine and Mookerjee, 1975; Hansch and Leo, 1979) , whereas the volume fractions of tissue components can be found in the literature for a number of species (Poulin *et al.*, 1999) or established experimentally. The above equation has been used to predict rat and human $P_{t:a}$ (liver, muscle, fat) of several alkanes, haloalkanes and aromatic hydrocarbons. For oxygen-containing VOCs (alcohols, esters, ethers), using vegetable oil instead of n-octanol as the lipid surrogate provides better estimates of their tissue solubility (Poulin and Krishnan, 1996c).

2.3.2 IN SILICO APPROACHES FOR BLOOD:AIR PCS

The blood:air PC is an important parameter since it influences the extent and rate of absorption, distribution and elimination of VOCs. Table 2 presents the various LFE-type and Free-Wilson-type QSAR models, as well as mechanistically-based *in silico* approaches that have been developed so far for the prediction of blood:air PCs of VOCs.

The LFE-type QSARs have benefited from the numerous studies involving anaesthetic gases in humans (Eger and Larson, 1964; Cowles *et al.*, 1971; Steward *et al.* 1973; Saraiva *et al.*, 1977; Laass, 1987). Since anaesthetic-like compounds are relatively lipophilic, the best regression equations are mostly those that contain hydrophobic parameters such as $P_{o:a}$ and $P_{w:a}$ or measures of solubility in lipids and water. Batterman *et al.* (2002) related the human blood:air PCs of trihalomethanes to various descriptors including molecular weight and number of bromine atoms in the compound. Since these descriptors also tend to be correlated with lipophilicity (i.e., increases in molecular weight or number of bromine tend to increase $P_{o:w}$), these types of correlations, especially for such a reduced dataset, are to be expected. DeJongh *et al.* (1997) and Meulenberg and Vijverberg (2000) used the hydrophobic descriptors $P_{o:w}$, $P_{o:a}$, and $P_{w:a}$ to relate rat blood:air PC to the structure of VOCs. However, contrary to the regression with tissue:air PCs they could only derive adequate regressions when a significant intercept was included. Since partitioning into lipids and water, was taken into account by the hydrophobic descriptors, presence of an intercept was interpreted as being the result of significant binding to blood proteins. To-date, there have not been any attempts to correlate the magnitude of this binding intercept to LFE-type descriptors.

Free-Wilson-type QSARs have recently been developed for a series of chloroethanes. Each chemical in the chloroethane family was described with a common basic structure (BS) of two carbons (C-C) as well as a set of substituent groups. The working hypothesis was that each substituent group in the structure had an additive and constant contribution to the blood:air PC (P_b) as reflected by the following equation (Free and Wilson, 1964):

$$P_b = BS + \sum fs \times Cs \qquad [4]$$

where BS = contribution of the basic structure to P_b, fs = frequency of occurrence of the substituent S in the chemical and Cs = contribution of the substituent S to P_b.

The frequency and identity of each fragment in a given molecule was provided as input along with rat and human experimental values of P_b for each chloroethane, and multiple linear regression analyses on the experimental data were conducted to identify the contribution of the basic structure and the substituent groups (Fouchécourt and Krishnan, 2000). Group contribution to blood:air PC values, however, is different from one species to another, as can be seen in Table 2.

Biologically-based algorithms for predicting blood:air PCs should be able to account for chemical solubility in blood lipids (phospholipids and neutral lipids), solubility in blood water fraction as well as protein binding. Poulin and Krishnan (1996c) proposed the following algorithm for predicting P_b of VOCs:

$$P_b = [P_{o:w}\, P_{w:a}\, (V_{nb} + 0.3V_{pb})] + [P_{w:a}(V_{wb} + 0.7V_{pb})] \qquad [5]$$

where V_{nb} =volume fraction of neutral lipid in blood, V_{pb} = volume fraction of phospholipid in blood, and V_{wb} = volume fraction of water in blood.

The predictions of rat blood:air partition coefficient obtained using the above equation were found to be adequate for relatively hydrophilic organics (e.g., alcohols, ketones, acetate esters), but not for relatively lipophilic organic chemicals for which the predicted P_b was significantly lower than the experimental values. The blood:air partition coefficient of a chemical is a composite number that can represents two processes occurring in the blood, namely, solubility and binding. Whereas chemical solubility is likely to be determined by the neutral lipid, phospholipid and water contents in blood, the binding would appear to be associated with plasma proteins and/or hemoglobin. For alcohols, acetate esters and ketones, rat and human blood:air partition coefficients appear to be adequately predicted using solubility-based algorithms. For more lipophilic VOCs (e.g., alkanes, haloalkanes, aromatic hydrocarbons), however, the blood:air partition coefficients obtained using the solubility-based algorithm are lower than the experimental data. The fact that the rat blood:air partition coefficient of lipophilic VOCs are under-predicted has been explained by the potential binding of these substances to blood proteins (Poulin and Krishnan, 1996b). However, such a discrepancy between predicted and experimental data is less pronounced in humans. Given that differences in blood lipid composition between rat and human are minor, interspecies differences in binding to blood components (either affinity constants or number of binding sites) can possibly be the basis (Wiester *et al.* 2002). At the present time, there are no validated mechanistic algorithms for predicting association constants for blood protein binding of organic chemicals. However, a qualitative approach for identifying VOCs that can bind to blood

proteins has been developed on the basis of the consideration of structural features and lipophilic characteristics (Poulin *et al.* 1999).

2.3.3 IN SILICO APPROACHES FOR TISSUE:BLOOD PCS

Tissue:blood PCs are fundamental parameters required for the construction of PBPK models. These parameters not only determine the tissue concentration at steady-state but also influence the time taken to attain steady-state. In the case of VOCs, the tissue:blood PCs can be computed by dividing the tissue:air PCs with the blood:air PC. It is useful, however, to develop approaches for predicting tissue:blood PCs directly. The *in silico* approaches published to-date for the purpose of estimating tissue:blood PCs of organic chemicals are presented in Table 3. Among the available *in silico* approaches, the LFE-type QSARs have mainly focused on using steric or hydrophobic descriptors. Abraham *et al.* (1994) developed equations for many tissues using the McGowan volume (V_x), an indicator of compound bulk. The brain:blood PCs of a series of CNS-acting pharmaceutical agents (most notably H_2R antagonists) have been extensively studied and related to the steric descriptors V_x, molar volume (V_m), molecular weight (MW) and polar surface area (PSA). Parham *et al.* (1997) also developed QSARs using steric descriptors, for estimating adipose tissue:blood PCs of a series of polychlorinated biphenyls (PCBs) congeners. These descriptors included some that described planarity, position of chlorines, and the effect of the chlorines on the adjacent carbons. It was shown that the PC depended mostly on the presence or absence of adjacent non-chlorine-substituted *meta* and *para* carbons. Since PCB congeners with substituted *meta-para* pairs tend to be more slowly

eliminated than those with unsubstituted pairs, it is suggested that the reason for this slower elimination might be the higher adipose tissue:blood PC, which leads to a greater storage of PCBs in this tissue. Most of the work relating to hydrophobic descriptors have involved a series of basic drugs in rabbit tissues, for which tissue:blood (or plasma) PCs were related to $P_{o:w}$ using an exponential function (Yokogawa *et al.*, 1990, 2002). For rat tissue:blood PCs, incorporation of $P_{o:w}$ data in regression equations has necessitated significant intercepts (DeJongh *et al.*,1997), due to the blood protein binding of these chemicals.

Free-Wilson-type QSARs have also been developed for the tissue:blood PCs of chloroethanes in rats, humans, and fish (Table 3). Because experimental PCs for fish were limited to 3 chemicals (in stead of the possible 9 in the series, as for rats and humans), the number of fragments in rats and humans is not the same as in fish (4 vs 2, respectively). The statistical power of the fish model is therefore low, as compared to datasets used in LFE-type QSARs, which may consist of up to a hundred or so chemicals, typically the case when data on pharmaceuticals are included.

Regarding the biologically-based algorithms for the estimation of tissue:blood partition coefficients, the following published by Poulin and Krishnan (1995a,b) is of use with VOCs that do not exhibit significant binding to blood proteins:

$$P_{t:b} = \frac{P_{o:w}(V_{nt} + 0.3V_{pt}) + (V_{wt} + 0.7V_{pt})}{P_{o:w}(V_{nb} + 0.3V_{pb}) + (V_{wb} + 0.7V_{pb})} \qquad [6]$$

If blood protein binding is important (as is the case for lipophilic VOCs), then it should be additionally accounted for.

2.3.4 IN SILICO APPROACHES FOR PROTEIN BINDING

Some chemicals bind to specific transport proteins. For example, various drugs (e.g., warfarin) bind to serum albumin. Other chemicals bind to tissue proteins. In these cases, partitioning of a compound in that tissue will depend not only on solubility in tissue water and lipids, but also on the binding to protein. Since only unbound chemical can freely distribute across tissue membranes, determination of the fraction unbound (f_u) becomes important. For chemicals that reversibly bind to proteins, f_u is related to the binding affinity constant (K_a) as follows:

$$1/f_u = 1+K_a(nC_p-C_b) \qquad [7]$$

where n = number of binding sites on the protein

C_p = molar concentration of binding protein, and

C_b = molar concentration of bound chemical

The available *in silico* approaches for estimating f_u and K_a are listed in Table 4. These approaches have mainly relied upon the use of hydrophobic descriptors in LFE-type equations. For example, Nestorov *et al.* (1998) related the ratio of fraction bound to fraction unbound for various tissues in rats to the $P_{o:w}$ in LFE-type QSARs for a series of barbituric acids. The number of binding sites was assumed to be one for all tissues. The K_a or f_u in human blood for penicillins, organic acids, cephalosporins, and aromatic acids have also related to hydrophobic descriptors. These observations seem to suggest that non-specific reversible binding to tissue protein is mostly a lipophilic process, probably because of the intrinsic hydrophobic

nature of the binding site. To date there have been no Free-Wilson type QSARs developed to describe protein binding.

Even if there are no mechanistic algorithms for the prediction of K_a or f_u values, Poulin *et al.* (1999) proposed a semi-quantitative mechanistic algorithm that considers protein binding for facilitating the calculation of apparent blood:air PC. The apparent blood:air PC reflects the ratio of steady-state arterial blood concentration ($C_{a(total)}$) to the atmospheric concentration (C_{air}) of the chemical as follows:

$$P_{b:a,(app)} = \frac{C_{a(total)}}{C_{air}} \qquad [8]$$

Considering the components of $C_{a(total)}$, the above equation can be rewritten as follows:

$$P_{b:a(app)} = \frac{C_{a,free}}{C_{air}} + \frac{C_{a,bound}}{C_{air}} \qquad [9]$$

Since $C_{a,\,bound} = C_{a,\,free}K_aC_p/(1+K_aC_{a,\,free})$,

$$P_{b:a(app)} = \frac{C_{a,free}}{C_{air}} + \left[\frac{C_{a,free} * K_a * C_p}{C_{air} * \left(1 + K_a * C_{a,free}\right)}\right] \qquad [10]$$

where K_a = binding association constant and C_p = concentration of binding proteins.

The preceding equation can be rewritten as,

$$P_{b:a(app)} = \frac{C_{a,free}}{C_{air}} * \left[1 + \frac{K_a * C_p}{1 + K_a * C_{a,free}}\right] \quad [11]$$

The first term of the above equation corresponds to solubility-based $P_{b:a}$. Therefore, replacing $C_{a,free}/C_{air}$ with $P_{b:a}$, the above equation becomes:

$$P_{b:a(app)} = P_{b:a(pred)} * \left[1 + \frac{K_a * C_p}{1 + K_a * C_{a,free}}\right] \quad [12]$$

This equation can be incorporated within PBPK models to calculate the $P_{b,app}$ as a function of time and exposure concentration. Here, the only additional, chemical-specific parameter that is required relates to K_a. At the present time, there is no validated animal-replacement algorithm for predicting association constants for blood protein binding of organic chemicals. However, based on the analysis presented by Poulin and Krishnan (1996b), it would appear that the average K_a value for rat hemoglobin binding is 1930 M^{-1} for several VOCs (i.e., chemicals with a molecular volume of < 300 cubic Angstroms, log $P_{o:w}$ >1, and lacking oxygen in the molecule). This information may be used, at the present time, to provide a "first-cut" estimate of K_a and $P_{b:a(app)}$ for purposes of PBPK modeling of VOCs in the absence of experimental data.

2.3.5 IN SILICO APPROACHES FOR CLEARANCE CONSTANTS

Hepatic clearance (CL_h) in PBPK models is described as the product of the extraction ratio (E) and liver blood perfusion rate (Q_l). The extraction ratio depends upon the intrinsic clearance (CL_{int}), which is equal to the ratio of V_{max}/K_m for first order conditions. *In silico* approaches published so far, regarding the estimation of CL_h, V_{max} and K_m are listed in Tables 5 – 7. Most of the *in silico* models have been

developed with data on pharmaceutical products. LFE-type QSARs for benzodiazepines relating CL_h in humans with electrostatic descriptors such as the ionization potential (Lewis, 2000) and such QSARs for V_{max} of n-demethylation of ethylamines using molecular length have also been developed (Lewis, 2001). Hydrophobic descriptors have been used to develop QSARs of CL_h of a series of basic drugs in rabbits (Ishizaki *et al.,* 1997; Yokogawa *et al.*, 2002). LFE-type QSARs relating K_m to electrostatic descriptors such as pK_a have been developed for the acetylation of sufonamides (Table 7) (Seydel and Schaper, 1982). Additionally, the Hammet constant has been used to relate demethylation and sulfatation affinity to the structure of halo-nitro and phenol compounds (Hansch and Leo, 1995). K_m has also been related to purely hydrophobic descriptors such as $P_{o:w}$ in the case of affinity for demethylation, glucuronidation, hydrolysis, and sulfation of phenols, phenylhippurates, and morphines (Hansch and Leo, 1995). Recent studies have shown, however, that in most cases, a combination of all three types of descriptors – electrostatic, lipophilic and steric - best describe affinity to metabolizing enzymes in LFE-type models (Lewis and Dickins, 2002).

There have only been a few attempts to develop QSAR models of the hepatic clearance, intrinsic clearance and metabolism constants (V_{max}, K_m) for environmental pollutants. Waller *et al.* (1996) developed a 3-D QSAR model for estimating intrinsic clearance of a small series of VOCs that are substrates of CYP2E1. Gargas *et al.* (1988) related the V_{max} of a series of haloalkanes to connectivity indices. Parham and Portier (1998) predicted rates of metabolism of a series of PCB congeners using steric descriptors, although implications of the inclusion of such descriptors were not discussed. Since $k_{cat} = V_{max}/[E_t]$, where $[E_t]$ = total enzyme concentration, some authors explored quantitative relationships

between the catalytic rate (k_{cat}) and structural descriptors (Table 6). These efforts indicate the varying importance of steric, hydrophobic, and electrostatic processes, depending on the nature of the binding site and enzymes involved (Tables 6-7).

The development of Free-Wilson type QSARs for V_{max} and K_m of VOCs has been attempted for chloroethanes (Fouchécourt and Krishnan, 2000). No mechanistic algorithms, however, are available for predicting CL_h, CL_{int}, V_{max} or K_m *a piori* without any experimental work. An interim approach, at least for VOCs, would involve the use of physiological limits of clearance in order to estimate the range of blood (or tissue) concentration possible in an individual (or population). In other words, even without knowing the exact rate of metabolism, it should be possible to establish the theoretical limits of blood concentration curves (Poulin *et al.*, 1999). For example, in the case of chemicals metabolised in the liver, the rate of amount metabolised (RAM) would be equal to:

$$RAM = C_a * CL_h \quad [13]$$

where C_a = arterial concentration of chemical.

Because $CL_h = Q_l * E$ and E cannot be lower than 0 or higher than 1, the envelope of possible concentrations can be obtained by setting CL_h in the above equation to its physiological limits, i.e., Q_l or 0. Although in certain cases this range can be quite large, it can provide a "first-cut" estimate of the possible effect of metabolism, and is especially appropriate for human populations in which the metabolic rate is substantially variable.

There have been limited attempts to develop QSARs for renal and total body clearance of chemicals (e.g., xylidines) using the LFE and Free-Wilson approaches, but the predictive power was limited to the substituents in the dataset (Table 5) (Seydel and Schaper, 1982).

2.3.6 IN SILICO APPROACHES FOR SKIN PERMEABILITY CONSTANTS

Table 8 provides a list of *in silico* approaches available for predicting the skin permeability constant (K_p). K_p is essential to simulate the pharmacokinetics of chemicals following topical application or dermal contact. The LFE-type QSARs relating K_p to V_x, molecular weight and connectivity indices have been developed for alcohols, steroids, esters, and several drugs. Recent studies have suggested molecular volume to be the most effective steric descriptor (Patel *et al.*, 2002). However, the most significant overall descriptor is log $P_{o:w}$, as demonstrated by the equations directly relating $P_{o:w}$ and K_p (Table 8).

A mechanistic equation for the estimation of K_p in human skin was also developed by Poulin and Krishnan (2001):

$$K_p = \frac{P_{vo:w} * 0.028 * D_l}{0.0340} + \frac{P_{p:w} * 0.88 * D_p}{0.0018} \qquad [14]$$

where $P_{vo:w}$ = vegetable oil :water PC,

$P_{p:w}$ =protein:water PC for stratum corneum,

D_l = coefficient for diffusion into the lipid fraction of stratum corneum, and

D_p=coefficient for diffusion into the protein fraction of stratum corneum.

In the above algorithm, the coefficients 0.028, 0.034, 0.88, and 0.0018 are species-specific and refer to the fractional content of lipid in skin, the path length for a hypothetical tortuous diffusion pathway in stratum corneum, the sum of the fractional content of water and protein in skin, and the path length for a

hypothetical transcellular diffusion pathway across the corneocytes of stratum corneum, respectively. The remaining terms are chemical-specific and can be derived from molecular structure. $P_{p:w}$ in stratum corneum has been related to $P_{o:w}$ through a empirical relationship (Poulin and Krishnan, 2001). Validation of this algorithm was accomplished with human data for a series of structurally-unrelated acids, alcohols, and hydrocarbons. While Poulin and Krishnan (2001) suggest that additional processes (metabolism, binding) would probably have to be considered for certain compounds, all predictions using the above equation were within a factor of two of the experimental data. Since hydrogen-bonding effect was not considered in this equation, K_p predicted using the above equation will be limited to organic compounds with weak hydrogen bonding capabilities. Predictions of K_p for chemicals with high hydrogen bonding capabilities (containing at least two important hydrogen bonding groups) can be accomplished by appropriately accounting for this phenomenon.

2.3.7 IN SILICO APPROACHES FOR ORAL ABSORPTION CONSTANTS

The rate constant for oral absorption of chemicals may be a single number reflecting the sum total of all processes involved, or be a set of constants each one of which may describe the rate of absorption along the gastrointestinal tract. The simplest description of oral absorption in PBPK models uses a first order rate constant ($K_{a\ (oral)}$) to approximate this process. And the value of this parameter is frequently obtained by fitting the model to serial blood concentration data obtained following oral dosing.

In silico efforts in modeling absorption have mainly focused on developing software packages that are capable of predicting the faction absorbed in humans (Balimane *et al.,* 2000; Agoram *et al.,* 2001; Raevsky and Schaper, 1998; Raevsky *et al.,* 2002). Among these the ACAT and the iDEA models are commonly being used for pre-clinical analysis by the pharmaceutical industry (Agoram *et al.,* 2001; Grass and Sinko, 2002). These models do not provide separate estimates of absorption rate constants, since they also integrate first pass metabolism by the liver in their estimation of fraction absorbed. LFE-type QSARs that have been developed to relate $K_{a\ (oral)}$ to structural descriptors or $P_{o:w}$ are summarized in Table 9. Mostly, these equations show the critical importance of lipophilicity in estimating absorption of certain substances, even though this is not necessarily the case for other substances. There is also evidence that partial atomic charge and dipole moment (electrostatic descriptors), as well as surface area and volume (steric descriptors) are related to the percent absorbed in humans (Balimane *et al.*, 2000). However, the role of these parameters, if any, in determining the hepatic first pass effect in addition to the rate of absorption remains unclear (see Tables 5-7). Currently, no mechanistic algorithms are available to provide predictions of $K_{a\ (oral)}$ values of chemicals prior to any testing.

The *in silico* approaches described above can be incorporated within PBPK models to obtain simulations of blood and tissue concentrations for new chemicals, prior to testing, as described in the following section.

2.4 *Integrating* in silico *approaches into risk assessment*

PBPK models facilitate the incorporation of the estimates of the various chemical-specific parameters with those of physiological parameters to simulate the pharmacokinetics of chemicals in intact animals. The simulations of internal dose measures obtained with these models are increasingly used to replace external dose in risk assessment calculations to enhance the scientific basis of the methodology. For the risk assessment of systemic acting chemicals, PBPK modeling has been used for establishing the internal dose corresponding to the NOAEL in animal and subsequent extrapolation to human safe levels (Reitz *et al.*, 1988), while for the risk assessment of carcinogens, these models have been used to derive a unit risk based on internal dose measures and subsequently risk levels in exposed humans (Andersen *et al.*, 1987). The use of *in silico* approaches, presented above, could facilitate risk assessment for *de novo* compounds using only molecular structure as input. However, before applying this to new, untested chemicals, the robustness of such an approach would first have to be studied with known chemicals. The robustness can be studied using a leave-one-out procedure, where one chemical within a dataset is removed from the analysis and new coefficients for the QSAR model parameters are derived. The risk assessment can then be performed for the "left out" chemical, using the derived equations. Such a procedure would be useful in illustrating the usefulness of the *in sillico* approaches for estimating PBPK model parameters in risk assessment. A case study is presented here using the Free-Wilson QSARs for chloroethanes (Fouchécourt and Krishnan, 2000) for the PBPK modeling and health risk assessment of 1,1,1-trichloroethane (methyl chloroform).

2.4.1 FREE-WILSON QSARS FOR CHLOROETHANES

Quantitative relationships between molecular fragments of eight haloethanes (chloroethane, 1,1 dichloroethane, 1,2 dichloroethane, 1,1,2 trichloroethane, 1,1,1,2 tetrachloroethane, 1,1,2,2 tetrachloroethane, pentachloroethane, and hexachloroethane) and the chemical-specific parameters required for PBPK modeling [tissue:air PC (P_t, subscripts l = liver, f=fat, and s=slowly perfused), blood:air PC (P_b), V_{maxc} (μmol/hr/kg), and K_m (μM)] were established according to an additive model developed by Free and Wilson (1964). Table 10 presents the molecular fragments used to describe the chloroethanes. The common basic structure and the substitution positions for the structural fragments are presented in Figure 1. Fragments consisted of hydrogens ($-H_3$), chloride ($-ClH_2$), dichlorides ($-Cl_2H$), and/or trichlorides ($-Cl_3$). For simplicity, these will be referred to H_3, Cl, Cl_2, and Cl_3. Any investigated chloroethane can be reconstituted by assembling the number of appropriate fragments with the basic structure (Figure 1). The working hypothesis was that each substituent group in the structure had an additive and constant contribution to the parameter of interest (P_i), as reflected by the following equation (Free and Wilson, 1964):

$$P_i = BS_i + \sum f_s \times C_{si} \qquad [15]$$

where BS_i = contribution of the basic structure to P_i, f_s = frequency of occurrence of the substituent S in the chemical and C_{si} = contribution of the substituent S to P_i.

By providing experimentally-determined rat P_t, P_b, V_{maxc} and K_m (Gargas *et al.*, 1988;Gargas *et al.*, 1989) for the 8 chloroethanes, along with the frequency of

occurrence of each substituent in each chloroethane (f_s) (Table 10) as input to a commercially available software (QSAR®-PC, Biosoft®, Cambridge, UK) the parameters of eqn. 15 were quantified. The output of the multiple linear regression performed by the software provided the contribution of the basic structure and the substituent groups. Any negative value resulting from the application of the QSAR equation was replaced by a default value of 1 (e.g., P_b of chloroethane in rats and humans).

Molecular fragments and their contributions to PBPK model parameters in rats and humans, as developed using the Free Wilson approach, are presented in Tables 11 and 12. The correlation between the estimates obtained using the *in silico* approach (i.e., Free Wilson approach) and experimental data on partition coefficients and metabolic constants is presented in Figures 2 and 3. The mean (±SD) of the ratio of experimental values to predicted values in rats was 1.17±1.2 for P_b, 0.96±0.17 for P_l, 0.89±0.21 for P_s, 0.97±0.24 for P_f, 1.03±0.35 for V_{maxc}, and 0.99±0.25 for K_m. The mean (±SD) of the experimental to predicted ratios in humans was 1.08±0.73 for P_b, 0.96±0.19 for P_l, 0.98±0.14 for P_s, and 1.06±0.38 for P_f. These results suggest that both tissue:blood partition coefficients and metabolic constants of structurally-related chemicals (e.g., chloroethanes) can be adequately described by a Free-Wilson model. This model was then applied to predict the PBPK model parameters for methyl chloroform (a chemical that was not part of the initial dataset) in rats and humans. Results are presented in Tables 13 and 14. The ratio of predicted to experimental values in rats ranged from 0.41 (for P_b) to 1.56 (for K_m), whereas the ratio of predicted to experimental values in humans ranged from 0.46 (for P_b) to 2.1 (for P_f). These QSAR predictions were

then integrated within a PBPK model to generate simulations of the pharmacokinetics of methyl chloroform in rats and humans.

2.4.2 Integrating Free-Wilson QSARs into PBPK models

The QSARs for parameter estimation were incorporated within a PBPK model that consisted of four compartments (liver, fat, richly perfused and slowly perfused tissues) interrelated by the blood circulation (Ramsey and Andersen, 1984). Tissue uptake was blood flow limited and was determined by the chemical specific tissue:blood PCs:

$$\frac{\partial Ai}{\partial t} = Qi \cdot (Ca - Cvi) \qquad [16]$$

$$Ai = \int_0^t \frac{\partial Ai}{\partial t} \qquad [17]$$

$$Cvi = \frac{Ai}{Vi \cdot Pi} \qquad [18]$$

where $\partial Ai/\partial t$ = rate of change in the amount of chemical in tissue i (mg/hr),

Qi = blood flow to tissue i (L/hr),

Ca = arterial blood concentration of chemical (mg/L),

Cvi = venous blood concentration of chemical leaving tissue i (mg/L),

Ai = quantity of chemical in tissue i (mg),

Vi = volume of tissue i (L), and

Pi = tissue:blood partition coefficient for tissue i.

In the metabolizing organ (i.e., liver) tissue uptake was determined additionally by the rate of metabolism:

$$\frac{\partial Al}{\partial t} = Ql \cdot (Ca - Cvi) - \frac{Vmax^* Cvl}{Km + Cvl} \qquad [19]$$

The PBPK model was written in Advanced Continuous Simulation Language (ACSL®, Aegis Technologies, Huntsville, AL). ACSL® simulation requires two components: a continuous simulation language (CSL) file, which contains the program (i.e., constants, QSAR algorithms for the model parameters, differential equations, and integration algorithms) and a command (CMD) file, which contains simulation conditions (i.e., exposure frequency and duration) and other chemical specific information provided by user input. Since Free Wilson QSARs along with the values of C_s and BS (Tables 11 and 12), for the partition coefficients and metabolic constants in rats and humans were included in the CSL file, only the frequencies of the substituents (Table 10) needed to be entered as input in the CMD file in order to simulate the kinetics of any given chloroethane.

Validation of the integrated Free Wilson-PBPK model was accomplished by comparing the output of this model to the output of a PBPK model that contained experimentally-determined parameter values. This model validation strategy is depicted in Figure 4 for 1,1,1 trichloroethane. Figure 5 compares the simulations of blood concentration profiles in rats obtained with the integrated QSAR-PBPK model and the conventional PBPK model for all chloroethanes in the series. The steady-state arterial blood and tissue concentrations of 1,1,1-trichloroethanes in rats and humans obtained using the integrated QSAR-PBPK model and the

conventional PBPK models are compared in Table 15 for a 1 ppm exposure. Ratios were 1.79±0.67 (range: 1.03 in fat to 2.22 in slowly perfused tissues) in rat and 1.42±0.33 (range: 1.06 in fat to 1.85 in richly perfused tissues) in humans.

2.4.3 QSAR-BASED RISK ASSESSMENT OF METHYL CHLOROFORM

The integrated *in silico*-PBPK modeling approach was used to establish the internal dose corresponding to a lifetime exposure to the rat no-observed adverse effect level (NOAEL, 875 ppm) of methyl chloroform (Reitz et al., 1988). For VOCs like methyl chloroform, lifetime exposures are expected to result in steady-state conditions. Therefore steady-state arterial blood concentrations (Ca_{ss}) were used as internal dose surrogates. The Ca_{ss} in rats during a lifetime continuous exposure to the NOAEL was obtained using the conventional and QSAR-based PBPK models. The Ca_{ss} given by the QSAR-PBPK model corresponding to the rat NOAEL was 59.3 mg/L whereas the corresponding value obtained with the experimental data based PBPK model was 24.9 mg/L. For risk assessment purposes, the human exposure concentration yielding the equivalent internal dose (i.e., Ca_{ss}) was established using the QSAR-PBPK model or alternatively the conventional PBPK model. The human exposure concentrations corresponding to the rat Ca_{ss} of 59.3 mg/L (which is the internal dose during the lifetime exposure to NOAEL) were 6342 ppm and 4252 ppm as obtained with the QSAR-based approach and the experimental-data based PBPK models, respectively.

2.5 Conclusions and Future Directions

Currently, physicochemical and biochemical parameters required for PBPK modeling are obtained by conducting *in vivo* or *in vitro* studies. Alternatively, chemical-specific parameters such as physicochemical and biochemical constants can be estimated from information on molecular structure. *In silico* approaches for estimating PBPK model parameters have mainly centred on LFE-type QSARs and mechanistically-based equations. While LFE QSARs have the advantage of being easily derived, they are limited to the chemical class for which they are developed. Furthermore, resulting parameter estimates cannot be extrapolated across species. There are also growing concerns over the mechanistic relevance of some of the structural descriptors used in these types of equations. The emerging mechanistically-based approaches offer the advantage of being relevant regardless of the chemical family, and are amenable to interspecies extrapolations. The applicability of these approaches has largely been verified with inhaled VOCs. Even though these approaches are conceptually applicable to non-volatile organics as well, it becomes more challenging to predict the other PBPK model parameters required for modeling the kinetics of these chemicals (i.e. tissue diffusion coefficients, association constants for binding, oral absorption rates, and dermal permeability coefficients). As our level of understanding of the mechanistic determinants of each of these parameters improves, *in silico* approaches to provide *a priori* predictions of these parameters can be developed.

Direct relationships between number or nature of the molecular fragments and the values of the physicochemical and biochemical parameters used in PBPK modeling can be established using Free-Wilson-type QSARs. This has been done

with chloroethanes. The case study presented in this chapter suggests that the development of QSAR-type PBPK models, in which the number and/or nature of the molecular fragments alone could be varied to provide simulations of the kinetics of chemicals, is feasible. Further, this case study has demonstrated the manner in which QSARs-based PBPK models can be used in human health risk assessment.

If the chemical-specific parameters required for PBPK modeling are related to molecular structure, then the QSARs can be incorporated within a PBPK modeling framework to relate molecular structure to the pharmacokinetics of chemicals. The pharmacokinetic profiles can be generated using only the information on the number of each molecular fragment specific to the chemical as input in the model. Using this type of framework, the QSAR-PBPK model can be used to simulate the pharmacokinetic profiles of chemicals for varying exposure scenarios.

Incorporation of species-specific QSARs for parameter estimation allows the prediction of the kinetics and accumulation of chemicals in a variety of species (e.g., fish, rat, mouse, humans) for many different exposure scenarios. Because of this, the internal dose corresponding to a NOAEL or unit risk in one species can be easily extrapolated to another species, using only chemical structure information. The scientific basis of interspecies extrapolation is increased, because the resulting extrapolation is based on internal dose and not exposure dose. In this study, this was illustrated by using a QSAR-PBPK model to simulate the internal dose of methyl chloroform in rats and humans. Continued research in this area should facilitate the development and validation of more mechanism based *in silico*

approaches for predicting the PBPK model parameters and pharmacokinetic profiles of chemicals, thus reducing the need for animal experiments but still contributing to the cost-efficient conduct of scientifically-sound health risk assessments for chemicals of concern.

2.6 References

1. Abraham, M. H., Kamlet, M. J., Taft, R. W., Doherty, R. M., and Weathersby, P. K. (1985). Solubility properties in polymers and biological media. 2. The correlation and prediction of the solubilities of nonelectrolytes in biological tissues and fluids. J.Med.Chem. **28**, 865-870.

2. Abraham, M. H. and Weathersby, P. K. (1994). Hydrogen bonding. 30. Solubility of gases and vapors in biological liquids and tissues. J.Pharm.Sci. **83**, 1450-1456.

3. Agoram, B., Woltosz, W.S., Bolger, M.B. (2001). Predicting the impact of physiological and biochemical processes on oral drug bioavailability. Adv. Drug Deliv. Rev. **50**, s41-s67.

4. Andersen, M.E., Clewell, H.J., Gargas, M.L., Smith, F.A., and Reitz, R.H. (1987). Physiologically based pharmacokinetics and the risk assessment process for methylene chloride. Toxicol. Appl. Pharmacol. **87**, 185-205.

5. Arms, A. D. and Travis, C. C. Reference Physiological Parameters in Pharmacokinetic Modeling. Office of Health and Environmental Assessment. EPA. EPA/600/6-88/004, 1-1-7.16. 1988. Washington, DC.

6. Balimane, P.V., Chong, S., and Morrison, R.A. (2000). Current methodologies used for the evaluation of intestinal permeability and absorption. J. Pharmacol. Toxicol. Methods. **44**, 301-312.

7. Batterman, S., Zhang, L., Wang, S., and Franzblau, A. (2002). Partition coefficients for the trihalomethanes among blood, urine, water, milk and air. Sci.Total Environ. **284**, 237-247.

8. Bertelsen, S. L., Hoffman, A. D., Gallinat, C. A., Elonen, C. M., and Nichols, J. W. (1998). Evaluation of log Kow and tissue lipid content as predictors of chemical partitioning to fish tissues. Environ.Toxicol.Chem. **17**, 1447-1455.

9. Bird, A. E. and Marshall, A. C. (1967). Correlation of serum binding of penicillins with partition coefficients. Biochem.Pharmacol. **16**, 2275-2290.

10. Cabala, R., Svobodova, J., Feltl, L., and Tichy, M. (1992). Direct determination of partition coefficients of volatile liquids between oil and gas by gas chromatography and its use in QSAR analysis. Chromatographia **34**, 601-606.

11. Clark, D. E. (1999). Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 2. Prediction of blood-brain barrier penetration. J.Pharm.Sci. **88**, 815-821.

12. Connell, D. W., Braddock, R. D., and Mani, S. V. (1993). Prediction of the partition coefficient of lipophilic compounds in the air-mammal tissue system. Sci.Total Environ. **Suppl Pt 2**, 1383-1396.

13. Cowles, A. L., Borgstedt, H. H., and Gillies, A. J. (1971). Solubilities of ethylene, cyclopropane, halothane and diethyl ether in human and dog blood at low concentrations. Anesthesiol. **35**, 203-211.

14. Csanady, G. A. and Laib, R. J. (1990). Use of linear free energy relationships in toxicology: prediction of partition coefficients of volatile lipophilic compounds. Arch.Toxicol. **64**, 594-596.

15. Csanady, G. A., Laib, R. J., and Filser, J. G. (1995). Metabolic transformation of halogenenated and other alkenes - a theoretical approach. Estimation of metabolic reactivities for in vivo conditions. Toxicol. Lett. **75**, 217-223.

16. DeJongh, J., Verhaar, H. J., and Hermens, J. L. (1997). A quantitative property-property relationship (QPPR) approach to estimate in vitro tissue-blood partition coefficients of organic chemicals in rats and humans. Arch.Toxicol. **72**, 17-25.

17. Eger, E. I. and Larson, C. P. Jr. (1964). Anaesthetic solubility in blood and tissues; values and significance. Br. J. Anaesth. **36**, 140-149.

18. Ekins, S. and Obach, R. S. (2000). Three-dimensional quantitative structure activity relationship computational approaches for prediction of human in vitro intrinsic clearance. J.Pharmacol.Exp.Ther. **295**, 463-473.

19. Feher, M., Sourial, E., and Schmidt, J. M. (2000). A simple model for the prediction of blood-brain partitioning. Int.J.Pharm. **201**, 239-247.

20. Feingold, A. (1976). Estimation of anesthetic solubility in blood. Anesth.Analg. **55**, 593-595.

21. Fiserova-Bergerova, V., Tichy, M., and Di Carlo, F. J. (1984). Effects of biosolubility on pulmonary uptake and disposition of gases and vapors of lipophilic chemicals. Drug Metab. Rev. **15**, 1033-1070.

22. Fouchécourt, M.-O. and Krishnan, K. (2000). A QSAR-type PBPK model for inhaled chloroethanes. Toxicol.Sci. **54**, 88.

23. Fouchécourt,M.-O., Walker,J., and Krishnan,K. (2000). An integrated QSAR-PBPK model for conducting rat-fish extrapolation of the biokinetics of chloroethanes. In Handbook on QSARs for Predicting Effects of Chemicals on Environmental-Human Health Interactions (J.Walker, Ed.), SETAC Press, Pensecola, FL.

24. Free, S. M. and Wilson, J. W. (1964). A mathematical contribution to structure-activity studies. J.Med.Chem. **7** , 395-399.

25. Gargas, M. L., Seybold, P. G., and Andersen, M. E. (1988). Modeling the tissue solubilities and metabolic rate constant (Vmax) of halogenated methanes, ethanes, and ethylenes. Toxicol. Lett. **43**, 235-256.

26. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H., and Andersen, M. E. (1989). Partition coefficients of low-molecular-weight volatile chemicals in various liquids and tissues. Toxicol. Appl.Pharmacol. **98**, 87-99.

27. Ghafourian, T. and Fooladi, S. (2001). The effect of structural QSAR parameters on skin penetration. Int.J.Pharm. **217**, 1-11.

28. Grass, G.M. and Sinko, P.J. (2002). Physiologically-based pharmacokinetic simulation modeling. Adv. Drug Deliv. Rev. **54**, 433-451.

29. Hansch, C. and Fujita, T. (1964). ρ-σ-π analysis. A method for the correlation of biological activity and chemical structure. J. Am. Chem. Soc. **86**, 1616-1626.

30. Hansch,C. and Leo,A. (1979). The fragment method of calculating partition coefficients. In Substituent constants for correlation analysis in chemistry and biology (C.Hansch and A.Leo, Eds.), pp. 18-43. Wiley, New York.

31. Hansch,C. and Leo,A. (1995). QSAR in metabolism. In Exploring QSAR: Fundamentals and applications in chemistry and biology (C.Hansch and A.Leo, Eds.), pp. 299-343. ACS Professionnal Reference Book of American Chemical Society, Washington.D.C.

32. Hine, J. and Mookerjee, P. K. (1975). The intrinsic hydrophobic character of organic compounds: correlations in terms of structural contributions. J.Org.Chem. **40**, 511-522.

33. Ishizaki, J., Yokogawa, K., Nakashima, E., and Ichimura, F. (1997). Relationships between the hepatic intrinsic clearance or blood cell- plasma partition coefficient in the rabbit and the lipophilicity of basic drugs. J.Pharm.Pharmacol. **49**, 768-772.

34. Kalizan, R. and Markuszewski, M. (1996). Brain/blood distribution described by a combination of partition coefficients and molecular mass. Int. J. Pharmacol. **145**, 9-16.

35. Kaneko, T., Wang, P. Y., and Sato, A. (1994). Partition coefficients of some acetate esters and alcohols in water, blood, olive oil, and rat tissues. Occup.Environ.Med. **51**, 68-72.

36. Kim, K. H. (1993). 3D-Quantitative structure-activity relationships:describing hydrophobic interactions directly from 3D structures using a comparative molecular field analysis (CoMFA) approach. Quant. Struct.-Act. Relat. **12**, 232-238.

37. Krishnan,K. and Andersen,M.E. (2001). Physiologically based pharmacokinetic modeling in toxicology. In Principles and methods of toxicology (A.W.Hayes, Ed.), pp. 193-241. Taylor & Francis, Philadelphia.

38. Laass,W. (1987). Estimation of blood/air partition coefficients of organic solvents. In QSAR in drug design and toxicology (D.Hadzi and B.Jerman-Blazic, Eds.), pp. 131-134. Elsevier Science Publishers, Amsterdam.

39. Laznicek, M., Kvetina, J., Mazak, J., and Krch, V. (1987). Plasma protein binding-lipophilicity relationships: interspecies comparison of some organic acids. J.Pharm.Pharmacol. **39**, 79-83.

40. Lewis, D. F. (2000). On the recognition of mammalian microsomal cytochrome P450 substrates and their characteristics: towards the prediction of human p450 substrate specificity and metabolism. Biochem.Pharmacol. **60**, 293-306.

41. Lewis, D. F. and Dickins, M. (2002). Factors influencing rates and clearance in P450-mediated reactions: QSARs for substrates of the xenobiotic-metabolizing hepatic microsomal P450s. Toxicol. **170**, 45-53.

42. Lewis, D. F. V. (2001). COMPACT: a structural approach to the modelling of cytochromes P450 and their interactions with xenobiotics. J.Chem.Technol.Biotechnol. **76**, 237-244.

43. Lombardo, F., Blake, J. F., and Curatolo, W. J. (1996). Computation of brain-blood partitioning of organic solutes via free energy calculations. J.Med.Chem. **39**, 4750-4755.

44. Meulenberg, C. J. and Vijverberg, H. P. (2000). Empirical relations predicting human and rat tissue:air partition coefficients of volatile organic compounds. Toxicol. Appl. Pharmacol. **165**, 206-216.

45. Moss, G. P., Dearden, J. C., Patel, H., and Cronin, M. T. (2002). Quantitative structure-permeability relationships (QSPRs) for percutaneous absorption. Toxicol. In Vitro **16**, 299-317.

46. Nestorov, I., Aarons, L., and Rowland, M. (1998). Quantitative structure-pharmacokinetics relationships: II. A mechanistically based model to evaluate the relationship between tissue distribution parameters and compound lipophilicity. J.Pharmacokinet.Biopharm. **26**, 521-545.

47. Norinder, U. and Haeberlein, M. (2002). Computational approaches to the prediction of the blood-brain distribution. Adv.Drug Deliv.Rev. **54**, 291-313.

48. Parham, F. M., Kohn, M. C., Matthews, H. B., DeRosa, C., and Portier, C. J. (1997). Using structural information to create physiologically based pharmacokinetic models for all polychlorinated biphenyls. Toxicol. Appl.Pharmacol. **144**, 340-347.

49. Parham, F. M. and Portier, C. J. (1998). Using structural information to create physiologically based pharmacokinetic models for all polychlorinated biphenyls. II. Rates of metabolism. Toxicol. Appl.Pharmacol. **151**, 110-116.

50. Patel, H., ten Berge, W., and Cronin, M. T. D. (2002). Quantitative structure-activity relationships (QSARs) for the prediction of skin permeation of exogenous chemicals. Chemosphere **48**, 603-613.

51. Paterson, S. and Mackay, D. (1989). Correlation of tissue, blood, and air partition coefficients of volatile organic chemicals. Br.J.Ind.Med. **46**, 321-328.

52. Perbellini, L., Brugnone, F., Caretta, D., and Maranelli, G. (1985). Partition coefficients of some industrial aliphatic hydrocarbons (C5- C7) in blood and human tissues. Br.J.Ind.Med. **42**, 162-167.

53. Poulin, P. and Krishnan, K. (1995a). A biologically-based algorithm for predicting human tissue: blood partition coefficients of organic chemicals. Hum.Exp.Toxicol. **14**, 273-280.

54. Poulin, P. and Krishnan, K. (1995b). An algorithm for predicting tissue: blood partition coefficients of organic chemicals from n-octanol: water partition coefficient data. J.Toxicol. Environ.Health **46**, 117-129.

55. Poulin, P. and Krishnan, K. (1996a). A tissue composition-based algorithm for predicting tissue:air partition coefficients of organic chemicals. Toxicol. Appl.Pharmacol. **136**, 126-130.

56. Poulin, P. and Krishnan, K. (1996b). A mechanistic algorithm for predicting blood:air partition coefficients of organic chemicals with the consideration of reversible binding in hemoglobin. Toxicol. Appl.Pharmacol. **136**, 131-137.

57. Poulin, P. and Krishnan, K. (1996c). Molecular structure-based prediction of the partition coefficients of organic chemicals for physiological pharmacokinetic models. Toxicol. Methods **6**, 117-137.

58. Poulin,P., Beliveau,M., and Krishnan,K. (1999). Mechanistic animal-replacement approaches for predicting pharmacokinetics of organic chemicals. In Toxicity assessment alternatives: methods, issues,opportunities. (H.Salem and S.A.Katz, Eds.), pp. 115-139. Humana Press Inc., Totowa, NJ.

59. Poulin, P. and Krishnan, K. (2001). Molecular structure-based prediction of human abdominal skin permeability coefficients for several organic compounds. J.Toxicol. Environ.Health **62**, 143-159.

60. Raevsky, O. A. and Schaper, K.-J. (1998). Quantitative estimation of hydrogen bond contribution to permeability and absorption processes of some chemicals and drugs. Eur.J.Med.Chem. **33**, 799-807.

61. Raevsky, O. A., Schaper, K.-J., Artursson, P., and McFarland, J.W. (2002). A novel approach for prediction of intestinal absorption of drugs in humans

based on hydrogen bond descriptors and structural similarity. Quant. Struct.-Act. Relat. **20**, 402-413.

62. Ramsey, J. C. and Andersen, M. E. (1984). A physiologically-based description of the inhalation pharmacokinetics of styrene in rats and humans. Toxicol. Appl.Pharmacol. **73**, 159-175.

63. Reitz, R. H., McDougal, J. N., Himmelstein, M. W., Nolan, R. J., and Schumann, A. M. (1988). Physiologically-based pharmacokinetic modeling with methyl chloroform: Implications for interspecies, high-low dose and dose-route extrapolations. Toxicol. Appl. Pharmacol. **95**, 185-199.

64. Saraiva, R. A., Willis, B. A., Steward, A., Lunn, J. N., and Mapleson, W. W. (1977). Halothane solubility in human blood. Br.J.Anaesth. **49**, 115-119.

65. Sato, A. and Nakajima, T. (1979). Partition coefficients of some aromatic hydrocarbons and ketones in water, blood and oil. Br.J.Ind.Med. **36**, 231-234.

66. Seydel, J. K. and Schaper, K. J. (1982). Quantitative structure-pharmacokinetic relationships and drug design. Pharmacol.Ther. **15**, 131-182.

67. Smith, D. S., Jones, B. C., and Walker, D. K. (1996). Design of drugs involving the concepts and theories of drug metabolism and pharmacokinetics. Medicinal Res. Rev. **16**, 243-266.

68. Steward, A., Allott, P. R., Cowles, A. L., and Mapleson, W. W. (1973). Solubility coefficients for inhaled anaesthetics for water, oil and biological media. Br.J.Anaesth. **45**, 282-293.

69. Testa, B., Crivori, P., Reist, M., and Carrupt, P.-A. (2000). The influence of lipophillicity on the pharmacokinetic behavior of drugs: concepts and examples. Perspect.Drug Discov.Des. **19**, 179-211.

70. Tichy,M., Fiserova-Bergerova,V., and Di Carlo,F.J. (1984). Estimation of biosolubility of hydrophilic compounds - QSAR study. In QSAR in toxicology and xenobiochemistry (M.Tichy, Ed.), pp. 225-231. Elsevier Sciences Publishers, Amsterdam.

71. Tichy, M. (1991). QSAR approach to estimation of the distribution of xenobiotics and the target organ in the body. Drug Metabol.Drug Interact. **9**, 191-200.

72. Tichy, M. (1991). QSAR approach to target organ estimation. Sci.Total Environ. **109-110**, 407-410.

73. Waller, C. L., Evans, M. V., and McKinney, J. D. (1996). Modeling the cytochrome P450-mediated metabolism of chlorinated volatile organic compounds. Drug Metab. Dispos. **24**, 203-210.

74. Wiester, M. J., Winsett, D. W., Richards, J. H., Doerfler, D. L., and Costa, D. L. (2002). Partitioning of benzene in blood: influence of hemoglobin type in humans and animals. Environ.Health.Perspect. **110**, 255-261.

75. Yamaguchi, T., Yabuki, M., Saito, S., Watanabe, T., Nishimura, H., Isobe, N., Shono, F., and Matsuo, M. (1996). Research to develop a predicting system of mammalian subacute toxicity, (3). Construction of a predictive toxicokinetics model. Chemosphere **33**, 2441-2468.

76. Yokogawa, K., Nakashima, E., Ishizaki, J., Maeda, H., Nagano, T., and Ichimura, F. (1990). Relationships in the structure-tissue distribution of basic drugs in the rabbit. Pharm.Res. **7**, 691-696.

77. Yokogawa, K., Ishizaki, J., Ohkuma, S., and Miyamoto, K. (2002). Influence of lipophilicity and lysosomal accumulation on tissue distribution kinetics of basic drugs: a physiologically based pharmacokinetic model. Methods Find.Exp.Clin.Pharmacol. **24**, 81-93.

Table 1 : *In silico* approaches for estimating the tissue:air partition coefficients (P) of chemicals

Approach[a]	Species[b]	Chemical Class[c]	Reference
QSARs: LFE-type equations			
Electrostatic descriptors			
$\text{Log } P_{adipose:air} = -0.294 - 0.172R_2 + 0.729\pi_2^H + 1.7474\alpha_2^H + 0.219\beta_2^H + 0.895\text{Log}P_{he:a}$	H	Inert gases; LMWVOCs	(2)
$\text{Log } P_{brain:air} = -1.074 + 0.427R_2 + 0.286\pi_2^H + 2.781\alpha_2^H + 2.787\beta_2^H + 0.609\text{Log}P_{he:a}$	H	Inert gases; LMWVOCs	(2)
$\text{Log } P_{heart:air} = -1.208 + 0.128R_2 + 0.987\pi_2^H + 0.643\alpha_2^H + 1.783\beta_2^H + 0.597\text{Log}P_{he:a}$	H	Inert gases; LMWVOCs	(2)
$\text{Log } P_{kidney:air} = -1.084 + 0.417R_2 + 0.226\pi_2^H + 3.624\alpha_2^H + 2.926\beta_2^H + 0.534\text{Log}P_{he:a}$	H	Inert gases; LMWVOCs	(2)
$\text{Log } P_{liver:air} = -1.031 + 0.059R_2 + 0.774\pi_2^H + 0.593\alpha_2^H + 1.049\beta_2^H + 0.654\text{Log}P_{he:a}$	H	Inert gases; LMWVOCs	(2)
$\text{Log } P_{lung:air} = -1.300 + 0.667R_2 + 0.680\pi_2^H + 3.539\alpha_2^H + 3.35\beta_2^H + 0.458\text{Log}P_{he:a}$	H	Inert gases; LMWVOCs	(2)
$\text{Log } P_{muscle:air} = -1.14 + 0.544R_2 + 0.216\pi_2^H + 3.4714\alpha_2^H + 2.924\beta_2^H + 0.578\text{Log}P_{he:a}$	H	Inert gases; LMWVOCs	(2)
Steric descriptors			
$\text{Log } P_{adipose:air} = (0.734\,{}^1x^v) - (0.029x^v_s) - (1.57(1/{}^1x)) - (0.559(1/{}^1x^v)) - 0.098\,{}^3x^v_c + 2.213$	R	Haloalkanes	(25)
$\text{Log } P_{adipose:air} = 0.734\,{}^1X^v - 0.0291X_s^v - 1.570/{}^1X^v - 0.559\,{}^1X^v - 0.098\,{}^4X_c^v + 2.213$	R	Haloalkanes	(14)
$\text{Log } P_{adipose:air} = 0.563N_{Cl} + 1.028N_{Br} + 0.467N_C + 0.270Q_H - 0.199N_F - 0.097$	R	Haloalkanes	(7, 25)
$\text{Log } P_{adipose:air} = 1.037\,{}^1x^v - (0.007(1/x^v_s)) + 0.022Q_H - 0.177\,{}^3x^v_c - 0.199N_F - 0.0036$	R	Haloalkanes	(25)
$\text{Log } P_{liver:air} = (1.072\,{}^1x^v) - (0.021(1/x^v_s)) + (0.647(1/{}^1x^v)) - (0.304\,{}^4x^V_C) - 1.212$	R	Haloalkanes	(25)
$\text{Log } P_{liver:air} = 0.366N_{Cl} - 0.588N_{Br} + 0.345Q_H - 0.179N_F - 0.007$	R	Haloalkanes	(25)
$\text{Log } P_{liver:air} = -0.685\,{}^1x^v - (0.020(1/x^v_s)) + 0.232Q_H + (0.298(1/{}^1x^v)) + 0.104N_{Cl} - 0.726$	R	Haloalkanes	(25)
$\text{Log } P_{liver:air} = 1.072\,{}^1X^v - 0.021/X_s^v + 0.647/{}^1X^v - 0.304\,{}^4X_c^v - 1.212$	R	Haloalkanes	(14)
$\text{Log } P_{muscle:air} = 0.379Q_H - 0.278N_{Cl} + 0.536N_{Br} - 0.190N_F + 0.169N_{Cl} - 0.439$	R	Haloalkanes	(25)
$\text{Log } P_{muscle:air} = 0.399\,{}^1x^v - (0.007(1/x^v_s)) + 0.295Q_H + 0.259\,{}^4x^v_{pc} - 0.194N_F - 0.217$	R	Haloalkanes	(25)
$\text{Log } P_{muscle:air} = (0.995\,{}^1x^v) - (0.018(1/x^v_s)) - (0.424\,{}^4x^v_c) - (0.559(1/{}^1x^v) + (0.602(1/{}^1x^v)) - 1.334$	R	Haloalkanes	(25)
Hydrophobic descriptors			
$\text{Log}(P_{adipose:water} - V_{wt}) = 0.9P_{o:w} + 0.31$	F	Chloroethanes; Benzene	(8)
$\text{Log}(P_{kidney:water} - V_{wt}) = 0.72P_{o:w} - 0.56$	F	Chloroethanes; Benzene	(8)
$\text{Log}(P_{liver:water} - V_{wt}) = 1.06P_{o:w} - 1.43$	F	Chloroethanes; Benzene	(8)
$\text{Log}(P_{muscle:water} - V_{wt}) = 0.63P_{o:w} - 0.60$	F	Chloroethanes; Benzene	(8)
$\text{Ln } P_{adipose:air} = 0.032T_b - 5.456$	H	Haloalkanes	(14)

$\ln P_{liver:air}=0.022T_b-4.638$	H	Haloalkanes	(14)
$\log P_{adipose:air}=0.209+0.0628\log P_{w:a}+0.8868\log P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{adipose:air}=0.21\log P_{o:a}+0.24\log P_{w:a}$	H	Hydrophilic VOCs	(72)
$\log P_{adipose:air}=0.782\log P_{o:a}+0.201\log P_{w:a}+0.432$	H	Hydrophobic VOCs	(71)
$\log P_{adipose:air}=0.901\log P_{o:a}+0.150$	H	LMWVOCs	(21)
$\log P_{adipose:air}=0.174+0.910\log P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{brain:air}=-0.16\log P_{o:a}+0.82\log P_{w:a}+0.47$	H	Hydrophilic VOCs	(72)
$\log P_{brain:air}=0.274+0.537\log P_{w:a}+0.444\log P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{brain:air}=0.394+1.096\ \log P_{w:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{brain:air}=0.471\log P_{o:a}+0.630\log P_{w:a}-0.305$	H	Hydrophobic VOCs	(71)
$\log P_{brain:air}=0.844\log P_{o:a}-1.124$	H	LMWVOCs	(21)
$\log P_{brain:air}=-0.850+0.773\ \log P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{brain:air}=-3.692+1.253R_G$	H	Inert gases; LMWVOCs	(1)
$\log P_{kidney:air}=-0.18\log P_{o:a}+0.82\log P_{w:a}+0.53$	H	Hydrophilic VOCs	(72)
$\log P_{kidney:air}=0.277+1.111\log P_{w:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{kidney:air}=0.466\log P_{o:a}+0.379\log P_{w:a}-0.332$	H	Hydrophobic VOCs	(71)
$\log P_{kidney:air}=0.700\log P_{o:a}-0.877$	H	LMWVOCs	(21)
$\log P_{kidney:air}=-0.920+0.764\log P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{liver:air}=-0.388+0.502\log P_{w:a}+0.497\log P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{liver:air}=0.432+1.064\log P_{w:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{liver:air}=0.746\log P_{o:a}+0.178\log P_{w:a}-0.767$	H	Hydrophobic VOCs	(71)
$\log P_{liver:air}=0.871\log P_{o:a}-1.044$	H	LMWVOCs	(21)
$\log P_{liver:air}=-0.875+0.773\log P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{lung:air}=-0.21\log P_{o:a}+0.91\log P_{w:a}+0.41$	H	Hydrophilic VOCs	(72)
$\log P_{lung:air}=-0.057+0.870\log P_{w:a}+0.146\log P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{lung:air}=0.057+0.978\log P_{w:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{lung:air}=0.373\log P_{o:a}+0.416\log P_{w:a}-0.216$	H	Hydrophobic VOCs	(71)
$\log P_{lung:air}=0.644\log P_{o:a}-0.815$	H	LMWVOCs	(21)
$\log P_{lung:air}=-0.833+0.911\log P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{muscle:air}=-0.19\log P_{o:a}+0.82\log P_{w:a}+0.54$	H	Hydrophilic VOCs	(72)
$\log P_{muscle:air}=0.49\log P_{o:a}+0.39\log P_{w:a}-0.31$	H	Hydrophobic VOCs	(72)
$\log P_{muscle:air}=-0.263+0.575\log P_{w:a}+0.423\log P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\log P_{muscle:air}=0.351+1.108\log P_{w:a}$	H	Inert gases; LMWVOCs	(1)

$\text{Log } P_{muscle:air}=0.652\text{Log}P_{o:a}-0.702$	H	LMWVOCs	(21)
$\text{Log } P_{muscle:air}=-0.852+0.768\text{Log}P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\text{Log } P_{muscle:air}=-3.247+0.965R_G$	H	Inert gases; LMWVOCs	(1)
$P_{adipose:air}=0.447P_{o:a}+0.075P_{w:a}+6.59$	H	LMWVOCs; CFCs	(44)
$P_{brain:air}=(0.026S_o+0.51S_w)/S_a$	H	LMWVOCs	(51)
$P_{brain:air}=0.020P_{o:a}+0.380P_{w:a}+0.94$	H	LMWVOCs; CFCs	(44)
$P_{kidney:air}=(0.014S_o+0.51S_w)/S_a$	H	LMWVOCs	(51)
$P_{kidney:air}=0.011P_{o:a}+0.400P_{w:a}+0.69$	H	LMWVOCs; CFCs	(44)
$P_{kidney:air}=-0.391+0.550\text{Log}P_{w:a}+0.440\text{Log}P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$P_{liver:air}=(0.028S_o+0.51S_w)/S_a$	H	LMWVOCs	(51)
$P_{liver:air}=0.028P_{o:a}+0.79$	H	LMWVOCs; CFCs	(44)
$P_{muscle:air}=0.014P_{o:a}+0.384P_{w:a}+0.94$	H	LMWVOCs; CFCs	(44)
$\text{Ln } P_{adipose:air}=0.032T_b-5.456$	R	LMWVOCs	(14)
$\text{Ln } P_{liver:air}=0.022T_b-4.638$	R	LMWVOCs	(14)
$\text{Log } P_{adipose:air}=0.920\text{Log}P_{o:a}+0.136$	R	LMWVOCs	(26)
$\text{Log } P_{adipose:air}=0.927\text{Log}P_{o:a}-0.032\text{Log}P_{w:a}+0.120$	R	LMWVOCs	(26)
$\text{Log } P_{adipose:air}=1.027\text{Log}P_{o:a}-0.046\text{Log}P_{w:a}-0.119$	R	Haloalkanes	(25)
$\text{Log } P_{liver:air}=0.574\text{Log}P_{o:a}+0.302\text{Log}P_{w:a}-0.278$	R	Haloalkanes	(25)
$\text{Log } P_{liver:air}=0.730\text{Log}P_{o:a}+0.128\text{Log}P_{w:a}-0.550$	R	LMWVOCs	(26)
$\text{Log } P_{muscle:air}=0.477\text{Log}P_{o:a}+0.365\text{Log}P_{w:a}-0.374$	R	Haloalkanes	(25)
$\text{Log } P_{muscle:air}=0.644\text{Log}P_{o:a}+0.180\text{Log}P_{w:a}-0.725$	R	LMWVOCs	(26)
$P_{adipose:air}=0.594P_{o:a}+0.085P_{w:a}+9.40$	R	LMWVOCs; CFCs	(44)
$P_{brain:air}=0.054P_{o:a}+0.832P_{w:a}$	R	LMWVOCs; CFCs	(44)
$P_{kidney:air}=0.097P_{o:a}+0.826P_{w:a}$	R	LMWVOCs; CFCs	(44)
$P_{liver:air}=0.026P_{o:a}+0.878P_{w:a}+2.36$	R	LMWVOCs; CFCs	(44)
$P_{muscle:air}=0.010P_{o:a}+0.772P_{w:a}+0.29$	R	LMWVOCs; CFCs	(44)
Mechanistically-based equations			
$P_{tissue:air}=(S_sV_{wt}+S_vV_{nt}+0.7S_sV_{pt}+0.3S_vV_{pt})/S_a$	R, H	LMWVOCs	(55)
$P_{tissue:air}=P_{o:w}P_{w:a}(V_{nt}+0.3V_{pt})+P_{w:a}(V_{wt}+0.7V_{pt})$	R, H	LMWVOCs	(57)

[a] π_2^H=dipolarity/polarizability, α_2^H= overall hydrogen-bond acidity, β_2^H= overall hydrogen-bond basicity, $^1X^v$, X_s^v, ^1X, $^4X^v_c$, $^3X^v_c$, $^4X^v_{pc}$=connectivity indices, N_{Br}= number of bromines, N_C= number of carbons, N_{Cl}= number of chlorines, N_F= number of fluorines, $P_{he:a}$= hexadecane:air partition coefficient, $P_{o:a}$= n-octanol:air (or vegetable oil:air) partition coefficient, $P_{o:w}$= n-

octanol:water (or vegetable oil:water) partition coefficient, $P_{w:a}$= water:air partition coefficient, Q_H=polarity due to presence of hydrogen atoms in the molecule, R_2 = molar refractivity, R_g= mean solvent solubility, S_a= solubility in air, S_o= solubility in n-octanol (or vegetable oil), S_s= solubility in saline, S_v= solubility in vegetable oil, S_w=solubility in water, T_b= boiling point, V_{nt}= volume fraction of neutral lipids in tissues, V_{pt}= volume fraction of phospholipids in tissues, and V_{wt}= volume fraction of water in tissues.

[b] F=fish, H=human, and R=rat.

[c] CFCs = chlorofluorocarbons, LMWVOCs=low molecular weight volatile organic chemicals, and VOCs = volatile organic chemicals.

Table 2 : *In silico* approaches for estimating the blood:air partition coefficients (P) of chemicals

Approach[a]	Species[b]	Chemical Class[c]	Reference
QSARs: LFE-type equations			
Electrostatic descriptors			
$\text{Log } P_{blood:air} = -1.269+0.612R_2+0.916\pi_2^H+3.614\alpha_2^H+3.381\beta_2^H+0.362\text{Log}P_{he:a}$	H	Inert gases; LMWVOCs	(2)
$\text{Log } P_{plasma:air} = -1.48+0.490R_2+2.04\pi_2^H+3.5074\alpha_2^H+3.911\beta_2^H+0.157\text{Log}P_{he:a}$	H	Inert gases; LMWVOCs	(2)
Steric descriptors			
$\text{Log } P_{blood:air} = 0.0072MW+0.197$	H	Trihalomethanes	(7)
$\text{Log } P_{blood:air} = 0.321N_{Br}+1.06$	H	Trihalomethanes	(7)
$P_{blood:air} = 0.07MW+5.59$	H	Aliphatic hydrocarbons	(52)
$\text{Log } P_{blood:air} = 0.443Q_H-0.303N_F+0.225N_{Cl}+0.510N_{BR}+0.155N_C-0.104$	R	Haloalkanes	(25)
Hydrophobic descriptors			
$\text{Log } (P_{blood:water}- V_{wb})=0.7P_{o:w}-0.75$	F	Chloroethanes; Benzene	(8)
$\text{Ln } P_{blood:air} = 0.038T_b-13.3$	H	Aliphatic hydrocarbons	(14)
$\text{Log } P_{blood:air} = 0.0109T_b-2.584$	H	Trihalomethanes	(7)
$\text{Log } P_{blood:air} = -0.14\text{Log}P_{o:a}+0.86\text{Log}P_{w:a}+0.47$	H	Hydrophilic VOCs	(72)
$\text{Log } P_{blood:air} = 0.685\text{log}P_{o:a}-0.6565$	H	Trihalomethanes	(7)
$\text{Log } P_{blood:air} =0.45\text{Log}P_{w:a}+1.21$	H	VOCs	(38)
$\text{Log } P_{blood:air}= -0.003\text{Log}P_{w:a}+1.47$	H	VOCs	(38)
$\text{Log } P_{blood:air}= -0.074+0.802\text{Log}P_{w:a}+0.218\text{Log}P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\text{Log } P_{blood:air}= -0.07\text{Log}S_w+1.21$	H	VOCs	(38)
$\text{Log } P_{blood:air}= -0.09\text{Log}P_{o:a}+2.45$	H	VOCs	(38)
$\text{Log } P_{blood:air}= -0.102+0.675\text{Log}P_{w:a}+0.315\text{Log}P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\text{Log } P_{blood:air}= -0.295+0.588\text{Log}P_{w:a}+0.411\text{Log}P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\text{Log } P_{blood:air}= -0.338\text{Log}P_{o:a}+3.121$	H	Halogenated hydrocarbons	(70)
$\text{Log } P_{blood:air}= -0.6737+0.5319\text{Log}P_{o:a}\text{Log}P_{w:a}$	H	VOCs	(65)
$\text{Log } P_{blood:air}= 0.695\text{Log}P_{o:a}-1.076$	H	LMWVOCs	(21)
$\text{Log } P_{blood:air}= -0.820+0.754\text{Log}P_{o:a}$	H	Inert gases; LMWVOCs	(1)
$\text{Log } P_{blood:air}=0.09\text{Log}S_w+8.25\text{Log}V_o-11.09$	H	VOCs	(38)

Log $P_{blood:air}$=0.11LogS_w+1.91	H	VOCs	(38)
Log $P_{blood:air}$=0.180Log$P_{o:a}$+0.889Log$P_{w:a}$+0.054	H	Hydrophobic VOCs	(71)
Log $P_{blood:air}$=0.20LogS_w+1.29	H	VOCs	(38)
Log $P_{blood:air}$=0.22Log$P_{w:a}$+0.67Log$P_{o:a}$-0.98	H	VOCs	(38)
Log $P_{blood:air}$=0.22LogS_w+10.78LogV_w-40.99	H	VOCs	(38)
Log $P_{blood:air}$=0.262+0.996Log$P_{w:a}$	H	Inert gases; LMWVOCs	(1)
Log $P_{blood:air}$=0.27Log1000/P+5.10LogV_o-6.67	H	VOCs	(38)
Log $P_{blood:air}$=0.31LogS_w+3.90LogV_o-4.53	H	VOCs	(38)
Log $P_{blood:air}$=0.35Log1000/P+1.01	H	VOCs	(38)
Log $P_{blood:air}$=0.35LogS_w+0.79Log1000/P+1.34LogV_o-2.23	H	VOCs	(38)
Log $P_{blood:air}$=0.37LogS_w+10.09LogV_w-38.40	H	VOCs	(38)
Log $P_{blood:air}$=0.38LogS_w+0.91Log1000/P-0.45	H	VOCs	(38)
Log $P_{blood:air}$=0.45LogS_w+0.81Log1000/P-0.40	H	VOCs	(38)
Log $P_{blood:air}$=0.48LogS_w+0.75Log1000/P+1.67LogV_o-2.77	H	VOCs	(38)
Log $P_{blood:air}$=0.51Log1000/P+0.37	H	VOCs	(38)
Log $P_{blood:air}$=0.581Log$P_{o:a}$+0.332Log$P_{w:a}$-0.599	H	LMWVOCs	(26)
Log $P_{blood:air}$=0.63Log1000/P+0.38	H	VOCs	(38)
Log $P_{blood:air}$=0.65Log$P_{o:a}$-0.84	H	VOCs	(38)
Log $P_{blood:air}$=0.851LogS_w+1.78	H	VOCs	(38)
Log $P_{blood:air}$=0.984Log$P_{w:a}$+0.053	H	Ketones; ethers; gases	(70)
Log $P_{blood:air}$=1.07Log$P_{w:a}$+0.27Log$P_{o:a}$-0.79	H	VOCs	(38)
Log $P_{blood:air}$=1.21LogV_o-0.17	H	VOCs	(38)
Log $P_{blood:air}$=3.05-0.34$P_{o:n}$	H	Ketones	(10)
Log $P_{blood:air}$=-3.922+1.369R_G	H	Inert gases; LMWVOCs	(1)
Log $P_{blood:air}$=5.89LogV_w-21.43	H	VOCs	(38)
Log $P_{blood:air}$=7.86LogV_o-10.40	H	VOCs	(38)
/id Log $P_{blood:air}$=8.90LogV_w-33.40	H	VOCs	(38)
Log $P_{milk:air}$ = 0.900log$P_{o:a}$-1.095	H	Trihalomethanes	(7)
Log $P_{plasma:air}$= -0.079+0.896Log$P_{w:a}$+0.149Log$P_{o:a}$	H	Inert gases; LMWVOCs	(1)
Log $P_{plasma:air}$= -0.082+0.894Log$P_{w:a}$+0.152Log$P_{o:a}$	H	Inert gases; LMWVOCs	(1)
Log $P_{plasma:air}$= -0.848+0.890Log$P_{o:a}$	H	Inert gases; LMWVOCs	(1)
Log $P_{plasma:air}$= -3.696+1.208R_G	H	Inert gases; LMWVOCs	(1)

$\text{Log } P_{plasma:air}=0.038+1.019\text{Log}P_{w:a}$	H	Inert gases; LMWVOCs	(1)
$P_{blood:air} = 0.0072P_{o:a}+0.898P_{w:a}+0.03$	H	LMWVOCs; CFCs	(44)
$P_{blood:air} = 0.08e^{0.0308Tb}$	H	Aliphatic hydrocarbons	(52)
$P_{blood:air}=0.00442P_{o:a}$	H	Aliphatic hydrocarbons	(52)
$P_{blood:air}=0.88P_{w:a}+0.012$	H	VOCs	(20)
$P_{blood:air}=0.89P_{w:a}+0.011P_{o:a}$	H	LMWVOCs	(70)
$P_{blood:air}=0.90\text{Log}P_{w:a}-461$	H	Esters; Alcohols	(35)
$P_{blood:air}=P_{w:a}+(P_{o:a}/100)$	H	Anaesthetics	(17)
$P_{blood:air}=S_w(1+0.0035P_{o:w})/S_a$	H	LMWVOCs	(51)
$\text{Log } P_{blood:air}=P_{w:a}*\{V_{lb}P_{o:w}^{0.85}+V_{prb}(86.2/P_{o:w}+3.70)+V_{wb}$	H, R	LMWVOCs	(12)
$\text{Log } P_{blood:air}=0.426\text{Log}P_{o:a}+0.515\text{Log}P_{w:a}-0.070$	R	Haloalkanes	(25)
$\text{Log } P_{blood:air}=0.553\text{Log}P_{o:a}+0.351P_{w:a}-0.286$	R	LMWVOCs	(26)
$P_{blood:air}=0.0054P_{o:a}+0.931P_{w:a}+1.16$	R	LMWVOCs; CFC	(44)
QSARs: Free-Wilson-type equations			
$P_{blood:water} = BS_{(C-C)}\ (28.4) + nCL_2(-12.9) + nCL_3(12.9)$	F	Chloroethanes	(23)
$P_{blood:air} = BS_{(C-C)}\ (26.2) + nH_3(-34.9) + nCL(-4.51) + nCL_2(29.4) + nCL_3(11.5)$	H	Chloroethanes	(22)
$P_{blood:air} = BS_{(C-C)}\ (45.6) + nH_3(-51.5) + nCL(-8.86) + nCL_2(36.4) + nCL_3(11.1)$	R	Chloroethanes	(22)
Mechanistically-based equations			
$P_{blood:air} = P_{o:w}P_{w:a}(V_{nb}+0.3V_{pb})+P_{w:a}(V_{wb}+0.7V_{pb})$	R, H	LMWVOCs	(57)
$P_{blood:air} =[f_e(S_sV_{we}+S_vV_{ne}+0.7S_sV_{pe}+0.3S_vV_{pe})+f_p(S_sV_{wp}+S_vV_{np}+0.7S_sV_{pp}+0.3S_vV_{pp})]/S_a$	R, H	LMWVOCs	(56)

[a] π_2^H = dipolarity/polarizability, α_2^H= overall hydrogen-bond acidity, β_2^H= overall hydrogen-bond basicity, BS = basic structure, f_e= erythrocyte fraction in blood, f_p=plasma fraction in blood, MW=molecular weight, N_{Br}= number of bromines, N_C= number of carbons, N_{Cl}= number of chlorines, nCL= number of CL fragments, nCL_2= number of CL_2 fragments, nCL_3= number of CL_3 fragments, N_F= number of fluorines, nH_3= number of H_3 fragments, P=vapor pressure, $P_{he:a}$= hexadecane:air partition coefficient, $P_{o:a}$= n-octanol:air (or vegetable oil:air) partition coefficient, $P_{o:n}$= vegetable oil:nitrogen partition coefficient, $P_{o:w}$= n-octanol:water (or vegetable oil:water) partition coefficient, $P_{w:a}$= water:air partition coefficient, Q_H= polarity due to presence of hydrogen atoms in the molecule, R_2= molar refractivity, R_g= mean solvent solubility, S_a= solubility in air, S_s= solubility in saline, S_v= solubility in vegetable oil, S_w= solubility in water, T_b= boiling point, V_{lb}= volume fraction of lipids in blood, V_{nb}= volume fraction of neutral lipids in blood, V_{ne}= volume fraction of neutral lipids in tissues, V_{np}= volume fraction of neutral lipids in plasma, V_o=surface tension, V_{pb}= volume fraction of phospholipids in blood, V_{pe}= volume fraction of phospholipids in

erythrocytes, V_{pp}= volume fraction of phospholipids in plasma, V_{prb}= volume fraction of proteins in blood, V_w=heat released due to evaporation at the boiling point, V_{wb}= volume fraction of water in blood, V_{we}= volume fraction of water in erythrocytes, and V_{wp}= volume fraction of water in plasma.

[b] F=fish, H=human, and R=rats.

[c] CFCs = chlorofluorocarbons, LMWVOCs=low molecular weight volatile organic chemicals, and VOCs = volatile organic chemicals.

Table 3 : *In silico* approaches for estimating the tissue:blood partition coefficients (P) of chemicals

Approach[a]	Species[b]	Chemical Class[c]	Reference
QSARs: LFE-type equations			
Steric descriptors			
$\text{Log } P_{adipose:blood} = 0.168+0.198R_2+0.130\pi_2^H -1.211\alpha_2^H-3.267\beta_2^H+2.275V_x$	H	Inert gases; LMWVOCs; CFCs	(2)
$\text{Log } P_{brain:blood} = -0.166+0.239R_2-0.626\pi_2^H -0.368\alpha_2^H-0.615\beta_2^H+1.072V_x$	H	Inert gases; LMWVOCs; CFCs	(2)
$\text{Log } P_{brain:blood} =-0.0148PSA+0.152LogP_{o:w}+0.139$	H	Inert gases; HMWOCs; LMWVOCs	(11)
$\text{Log } P_{brain:blood}=1.359+0.338LogP_{cyh}-0.00618V_m$	H	H_2-R antagonists	(34)
$\text{Log } P_{heart:blood} = -0.346+0.204\pi_2^H -2.150\alpha_2^H-0.853\beta_2^H+0.931V_x$	H	Inert gases; LMWVOCs; CFCs	(2)
$\text{Log } P_{kidney:blood} = -0.188+0.226R_2-0.559\pi_2^H -0.433\beta_2^H+0.832V_x$	H	Inert gases; LMWVOCs; CFCs	(2)
$\text{Log } P_{liver:blood} = -0.270+0.233R_2-0.375\pi_2^H -1.004\alpha_2^H-1.118\beta_2^H+0.832V_x$	H	Inert gases; LMWVOCs; CFCs	(2)
$\text{Log } P_{lung:blood} = -0.150-0.195\pi_2^H +0.389V_x$	H	Inert gases; LMWVOCs; CFCs	(2)
$\text{Log } P_{muscle:blood} = -0.222-0.479\pi_2^H -0.517\beta_2^H+0.999V_x$	H	Inert gases; LMWVOCs; CFCs	(2)
$P_{adipose:plasma}=1.9988-0.5004UNS+0.1793NPL+0.05931DIFF^2$	H	PCBs	(48)
$\text{Log } P_{brain:blood} = 0.088+0.264R_2-0.966\pi_2^H-0.705\Sigma\alpha_2^H-0.756\Sigma\beta_2^H+1.189V_x$	R	H_2-R antagonists	(47)
$\text{Log } P_{brain:blood} = -0.088+0.272 \text{ Log } P_{o:w}-0.00116MW$	R	H_2-R antagonists	(34)
$\text{Log } P_{brain:blood}=-0.00116MW+0.272LogP_{o:w}-0.088$	R	Inert gases; volatile hydrocarbons	(47)
$\text{Log } P_{brain:blood}=-0.01V_m+0.35LogP_{o:w}+0.99I_3+1.25$	R	Drug-like molecules	(47)
$\text{Log } P_{brain:blood}=-0.021PSA-0.003MV+1.643$	R	Inert gases; HMWOCs; LMWVOCs	(11)
$\text{Log } P_{brain:blood}=-0.0322DPSA+1.33$	R	HMWOCs	(47)
$\text{Log } P_{brain:blood}=-0.038+0.198R_2-0.687\pi_2^H-0.715\Sigma\alpha_2^H-0.698\Sigma\beta_2^H+0.995V_x$	R	H_2-R antagonists; Inert gases; SOMs	(47)
$\text{Log } P_{brain:blood}=-0.218(N_N+N_O)+0.235logP_{o:w}-0.027$	R	HMWOCs	(47)
$\text{Log } P_{brain:blood}=0.476+0.541LogP_{o:w}-0.00794MW$	R	H_2-R antagonists	(34)
$\text{Log } P_{brain:blood}=1.296+0.309LogP_{cyh}-0.00570MW$	R	H_2-R antagonists	(34)
Hydrophobic descriptors			
$\text{Log } P_{brain:blood}=0.39LogP_{o:w}+0.68$	H	Drugs, Hormones	(66)
$\text{Log } P_{brain:blood}=0.054G^o+0.43$	H	H_2-R antagonists; LMWVOCs	(43)
$P_{adipose:blood}=[(V_{lt}P_{o:w}^{A1}+V_{wt})/(V_{lb}P_{o:w}^{A2}+V_{wb})]+B$	H, R	LMWVOCs	(16)
$P_{brain:blood}=[(V_{lt}P_{o:w}^{A1}+V_{wt})/(V_{lb}P_{o:w}^{A2}+V_{wb})]+B$	H, R	LMWVOCs	(16)
$P_{kidney:blood}=[(V_{lt}P_{o:w}^{A1}+V_{wt})/(V_{lb}P_{o:w}^{A2}+V_{wb})]+B$	H, R	LMWVOCs	(16)
$P_{liver:blood}=[(V_{lt}P_{o:w}^{A1}+V_{wt})/(V_{lb}P_{o:w}^{A2}+V_{wb})]+B$	H, R	LMWVOCs	(16)

$P_{muscle:blood}=[(V_{lt}P_{o:w}^{A1}+V_{wt})/(V_{lb}P_{o:w}^{A2}+V_{wb})]+B$	H, R	LMWVOCs	(16)
$Ln\ P_{kidney:blood}=0.0065\sum o$	R	HMWOCs	(75)
$Ln\ P_{liver:blood}=0.025\sum i$	R	HMWOCs	(75)
$Ln\ P_{muscle:blood}=0.0069\sum i$	R	HMWOCs	(75)
$Log\ P_{brain:blood}=0.035\Delta G_{solv}+0.259$	R	H_2-R antagonists; LMWVOCs	(47)
$Log\ P_{brain:blood}=0.4275-0.3873n_{acc,solv}+0.1092LogP_{o:w}-0.0017A_{pol}$	R	Drugs; LMWVOCs; Anaesthetics	(19)
$Log\ P_{brain:blood}=1.979+0.373LogP_{cyh}-0.00275V_{wav}$	R	H_2-R antagonists	(34)
$Log\ P_{brain:plasma} = -0.48*\Delta logP_{oct-cyc}+0.89$	R	H_2-R antagonists	(69)
$Ln\ P_{adipose:blood}=0.05\sum i+0.021$	R	HMWOCs	(75)
$P_{adipose:blood}=0.915P_{o:w}^{0.573}$	Rb	Basic drugs	(76)
$P_{adipose:plasma} = 0.016P_{o:w}^{1.255}$	Rb	Basic drugs	(77)
$P_{bone\ marrow:blood} =1.975P_{o:w}^{0.273}$	Rb	Basic drugs	(76)
$P_{bone:plasma} = 0.036P_{o:w}^{0.947}$	Rb	Basic drugs	(77)
$P_{brain:blood}=3.157P_{o:w}^{0.312}$	Rb	Basic drugs	(76)
$P_{brain:plasma} = 0.062P_{o:w}^{0.984}$	Rb	Basic drugs	(77)
$P_{gut:blood} = 3.002P_{o:w}^{0.346}$	Rb	Basic drugs	(76)
$P_{gut:plasma} = 0.058P_{o:w}^{1.02}$	Rb	Basic drugs	(77)
$P_{heart:blood}=1.678P_{o:w}^{0.422}$	Rb	Basic drugs	(76)
$P_{heart:plasma} = 0.032P_{o:w}^{1.098}$	Rb	Basic drugs	(77)
$P_{kidney:plasma} = 0.075P_{o:w}^{1.037}$	Rb	Basic drugs	(77)
$P_{liver:plasma} = 0.064P_{o:w}^{0.884}$	Rb	Basic drugs	(77)
$P_{lung:blood}=1.158P_{o:w}^{0.565}$	Rb	Basic drugs	(76)
$P_{lung:plasma} = 0.031P_{o:w}^{1.236}$	Rb	Basic drugs	(77)
$P_{muscle:blood}=4.928P_{o:w}^{0.221}$	Rb	Basic drugs	(76)
$P_{muscle:plasma} = 0.099P_{o:w}^{0.889}$	Rb	Basic drugs	(77)
$P_{skin:blood}=2.997P_{o:w}^{0.256}$	Rb	Basic drugs	(76)
$P_{skin:plasma} = 0.058P_{o:w}^{0.927}$	Rb	Basic drugs	(77)
$P_{spleen:blood}=3.002P_{o:w}^{0.346}$	Rb	Basic drugs	(76)
QSARs: Free-Wilson-type equations			
$P_{adipose:blood} = BS_{(C-C)}\ (94.5) + nCL_2(-29.2) + nCL_3(29.2)$	F	Chloroethanes	(23)
$P_{liver:blood} = BS_{(C-C)}\ (2.93) + nCL_2(-0.238) + nCL_3(0.238)$	F	Chloroethanes	(23)
$P_{muscle:blood}= BS_{(C-C)}\ (3.02) + nCL_2(-0.175) + nCL_3(0.175)$	F	Chloroethanes	(23)
$P_{adipose:blood} = BS_{(C-C)}\ (49.2) + nH_3(-0.440) + nCL(-14.54) + nCL_2(-6.65) + nCL_3(26.5)$	H	Chloroethanes	(22)

$P_{liver:blood} = BS_{(C-C)}$ (2.64) + nH_3(-0.61) + nCL(-0.66) + nCL_2(-0.18) + nCL_3(1.68)	H	Chloroethanes	(22)
$P_{muscle:blood} = BS_{(C-C)}$ (1.11) + nH_3(0.08) + nCL(-0.02) + nCL_2(-0.21) + nCL_3(0.15)	H	Chloroethanes	(22)
$P_{adipose:blood} = BS_{(C-C)}$ (30.1) + nH_3(-9.88) + nCL(-6.02) + nCL_2(-3.90) + nCL_3(17.3)	R	Chloroethanes	(22)
$P_{liver:blood} = BS_{(C-C)}$ (1.79) + nH_3(-0.9) + nCL(-0.38) + nCL_2(-0.21) + nCL_3(1.27)	R	Chloroethanes	(22)
$P_{muscle:blood} = BS_{(C-C)}$ (0.69) + nH_3(-0.12) + nCL(0.04) + nCL_2(-0.12) + nCL_3(0.17)	R	Chloroethanes	(22)
Mechanistically-based equations			
$P_{tissue:blood} = (S_oV_{nt}+S_w*0.7V_{pt}+S_o*0.3V_{pt}+S_wV_{wt})/(S_oV_{nb}+S_w*0.7V_{pb}+S_o*0.3V_{pb}+S_wV_{wb})$	H	LMWVOCs	(53)
$P_{tissue:blood} = (P_{o:w}V_{nt}+V_{wt}+P_{o:w}*0.3V_{pt}+0.7V_{pt})/[f_e(P_{o:w}V_{ne}+V_{we}+P_{o:w}*0.3V_{pe}+0.7V_{pe})$ $+f_p(P_{o:w}V_{np}+V_{wp}+P_{o:w}*0.3V_{pp}+0.7V_{pp})]$	R	Ketones; Alcohols; Esters	(54)
$P_{tissue:blood} = [P_{o:w}(V_{nt}+0.3V_{pt})+(V_{wt}+0.7V_{pt})]/[P_{o:w}(V_{nb}+0.3V_{pb})+(V_{wb}+0.7V_{pe})]$	R, H	LMWVOCs	(56)

[a] π_2^H=dipolarity/polarizability, α_2^H=overall hydrogen-bond acidity, β_2^H=overall hydrogen-bond basicity, ΔG_{solv}=free energy of solvation in hexadecane, Σi=molecular structure Fujita value, Σo=molecular structure Fujita value, A_1,A_2=Collander-type coefficient, A_{pol}=polar surface area, B=correction factor, BS=Basic structure, DIFF=variable dependent on the number of chlorine atoms in the aromatic cycle, DPSA=Dynamic polar surface area, f_e=fraction of erythrocytes in blood, f_p=fraction of plasma in blood, I_3=variable dependent on the presence of an amino nitrogen or carboxyl group, MV=molecular volume, MW=Molecular weight, $n_{acc,solv}$=number of solvated hydrogen-bond acceptors, nCL=number of CL fragments, nCL_2=number of CL_2 fragments, nCL_3=number of CL_3 fragments, nH_3=number of H_3 fragments, N_N=number of nitrogens, N_O=number of oxygens, NPL=variable dependant on the number of chlorine atoms in the molecule in ortho position, ^oG=Gibbs free energy related to the solvation of the substance in water, P_{cyh}=cyclohexane:water partition coefficient, $P_{o:w}$=n-octanol:water partition coefficient (or vegetable oil:water), $P_{oct-cyc}$=octanol-cyclohexane, PSA=polar surface area, R_2=Excess molar refraction, S_o=solubility in n-octanol (or vegetable oil), S_w=solubility in water, UNS=variable dependent on the number of atoms in the molecule that are not chlorines, V_{lb}=volume fraction of lipids in blood, V_{lt}=volume fraction of lipids in tissue, V_m=molar volume, V_{nb}=volume fraction of neutral lipids in blood, V_{ne}=volume fraction of neutral lipids in erythrocytes, V_{np}=volume fraction of neutral lipids in plasma, V_{nt}=volume fraction of neutral lipids in tissues, V_{pb}=volume fraction of phospholipids in blood, V_{pe}=volume fraction of phospholipids in erythrocytes, V_{pp}=volume fraction of phospholipids in plasma, V_{pt}=volume fraction of phospholipids in tissues, V_{wav}=volume of water needed in order to solubilize the substance, V_{wb}=volume fraction of water in blood, V_{we}=volume fraction of water in erythrocytes, V_{wp}=volume fraction of water in plasma, V_{wt}=volume fraction of water in tissue, and V_x=McGowan characteristic volume.

[b] F=fish, H=human, and R=rats.

[c] CFCs = chlorofluorocarbons, HMWOCs = high molecular weight organic chemicals, LMWVOCs=low molecular weight volatile organic chemicals, PCBs=polychlorobiphenyls, and VOCs = volatile organic chemicals.

Table 4: *In silico* approaches for estimating protein binding of chemicals[a]

Approach[b]	Species[c]	Chemical Class	Reference
QSARs: LFE-type equations			
Hydrophobic descriptors			
$Log\ (1/f_{u\ (plasma)}-1) = 0.994*LogP_{o:w}-1.10$	H	aromatic acids	(69)
$Log\ (1/f_{u\ (plasma)}-1) = 0.994LogP_{o:w}-1.10$	H	organic acids	(39)
$Log\ (1/K_{a\ (plasma)}) = -3.91LogP_{o:w}^{2}+13LogP_{o:w}-13.7$	H	cephalosporins	(69)
$Log\ (1-f_{u\ (brain)}) = 0.36\ Log\ P_{o:w}-1.07$	H	Barbiturates	(66)
$Log\ (1-f_{u\ (plasma)}) = 0.276\ Log\ P_{o:w}+1.2$	H	Penicillins	(66)
$Log\ (1-f_{u\ (plasma)}) = 0.30\ Log\ P_{o:w}-1.03$	H	Barbiturates	(66)
$Log\ (1-f_{u\ (plasma)}) = 0.33\ Log\ P_{o:w}+1.94$	H	Tetracyclines	(66)
$Log\ 1/K_{a\ (albumin\ binding)} = -0.85\ Log\ P_{o:w}+2.73$	H	Sulfapyrimidines; Sulfapyridines	(66)
$Log\ 1/K_{a\ (albumin\ binding)} = -0.97\ Log\ P_{o:w}+3.24$	H	Sulfapyridines	(66)
$Log\ 1/K_{a\ (albumin\ binding)} = -0.97\ Log\ P_{o:w}-0.70I+3.24$	H	Sulfapyrimidines; Sulfapyridines	(66)
$Log\ 1/K_{a\ (albumin\ binding)} = -0.99\ Log\ P_{o:w}+2.49$	H	Sulfapyrimidines	(66)
$Log\ K_{albumin\ binding} = 0.89\ Log\ P_{o:w}+1.47$	H	Sulfapyrimidines	(66)
$Log\ K_{albumin\ binding} = 1.15\ Log\ P_{o:w}+1.23$	H	Sulfonamides	(66)
$Log\ K_{albumin\ binding} = 1.23\ Log\ P_{o:w}-0.056$	H	Steroid bisguanylhydrazones	(66)
$Log\ K_{albumin\ binding} = 1.32Log\ P_{o:w}+0.37$	H	Penicillins	(66)
$Log\ K_{albumin\ binding} = 1.39\ Log\ P_{o:w}-1.19$	H	Cardenolides	(66)
$Log\ K_{albumin\ binding} = 1.65\ Log\ P_{o:w}-2.57$	H	Steroid hormones	(66)
$Log\ K_{plasma\ protein\ binding} = 0.73\Delta R_{mui}+1.46$	H	Sulfapyridines	(66)
$Log\ K_{a\ (blood\ protein\ binding)} = 0.504\Sigma\pi-0.665$	H	Penicillins	(9)
$Log\ (1/f_{u\ (plasma)}-1) = 1.011LogP_{o:w}-1.745$	R	organic acids	(39)
$Log\ (1-f_u)/f_{u\ (adipose)} = Log\ 0.750+0.936Log\ P_{o:w}$	R	Barbituric acids	(46)
$Log\ (1-f_u)/f_{u\ (brain)} = Log\ 0.073+0.860Log\ P_{o:w}$	R	Barbituric acids	(46)
$Log\ (1-f_u)/f_{u\ (gut)} = Log\ 0.099+0.824Log\ P_{o:w}$	R	Barbituric acids	(46)
$Log\ (1-f_u)/f_{u\ (heart)} = Log\ 0.135+0.780Log\ P_{o:w}$	R	Barbituric acids	(46)
$Log\ (1-f_u)/f_{u\ (kidney)} = Log\ 0.676+0.619Log\ P_{o:w}$	R	Barbituric acids	(46)
$Log\ (1-f_u)/f_{u\ (liver)} = Log\ 1.775+0.504Log\ P_{o:w}$	R	Barbituric acids	(46)

$\text{Log } (1\text{-}f_u)/f_{u\ (lung)} = \text{Log } 0.164+0.841\text{Log } P_{o:w}$	R	Barbituric acids	(46)
$\text{Log } (1\text{-}f_u)/f_{u\ (muscle)} = \text{Log } 0.080+0.835\text{Log } P_{o:w}$	R	Barbituric acids	(46)
$\text{Log } (1\text{-}f_u)/f_{u\ (pancreas)} = \text{Log } 0.022+1.095\text{Log } P_{o:w}$	R	Barbituric acids	(46)
$\text{Log } (1\text{-}f_u)/f_{u\ (plasma)} = \text{Log } 0.016+0.975\text{Log } P_{o:w}$	R	Barbituric acids	(46)
$\text{Log } (1\text{-}f_u)/f_{u\ (red\ blood\ cell)} = \text{Log } 0.178+0.677\text{Log } P_{o:w}$	R	Barbituric acids	(46)
$\text{Log } (1\text{-}f_u)/f_{u\ (skin)} = \text{Log } 0.271+0.736\text{Log } P_{o:w}$	R	Barbituric acids	(46)
$\text{Log } (1\text{-}f_u)/f_{u\ (spleen)} = \text{Log } 0.126+0.841\text{Log } P_{o:w}$	R	Barbituric acids	(46)
$\text{Log } (1\text{-}f_u)/f_{u\ (stomach)} = \text{Log } 0.058+0.939\text{Log } P_{o:w}$	R	Barbituric acids	(46)
$\text{Log } (1\text{-}f_u)/f_{u\ (testis)} = \text{Log } 0.120+0.747\text{Log } P_{o:w}$	R	Barbituric acids	(46)
$\text{Log } K_{plasma\ protein\ binding} = 0.33\Delta R_{mui}\text{-}0.53I+4.08$	R	Sulfapyridines	(66)
$\text{Log } (1/f_{u\ (plasma)}\text{-}1) = 1.016\text{Log}P_{o:w}\text{-}1.275$	Rb	organic acids	(39)

[a] K = protein affinity constant (Freundlich isotherm), K_a = protein affinity constant (Scatchard isotherm) and f_u = unbound fraction.

[b] $P_{o:w}$=octanol:water partition coefficient, π=molecular hydrophobicity constant, I=family indicator variable, ΔR_{mui}=variable dependant on the resistance constant due to diffusion of the non-ionized form in the lipid membrane

[c] H=humans, Rb=rabbit, and R=rat.

Table 5: *In silico* approaches for estimating clearances (CL) of chemicals

Approach[a]	Species[b]	Chemical Class[c]	Reference
QSARs: LFE-type equations			
Electrostatic descriptors			
$Log\ CL_{(hepatic)} = 0.64LogP_{o:w} - 0.98IP + 9.33$	H	Benzodiazepines	(40)
$Log\ CL_{(hepatic)} = 0.055Energy - 0.95IP - 0.53HBD + 10.63$	H	Benzodiazepines	(40)
$Log\ CL_{(hepatic)} = 0.067Energy - 1.01IP - 0.34HBD - 0.43\Delta E + 14.66$	H	Benzodiazepines	(40)
$Log\ CL_{(hepatic)} = 0.094Energy - 1.18IP - 0.74\Delta E + 18.65$	H	Benzodiazepines	(40)
$Log\ CL_{(hepatic)} = 0.65LogP_{o:w} - 0.40IP - 0.37HBD + 0.0025Hf + 3.63$	H	Benzodiazepines	(40)
$Metabolic\ ratio = 2.72Q_6 + 1.96E_H + 0.014S_N + 6.43$	R	Dichlorobiphenyls	(41)
Steric descriptors			
$1/Log\ CL_{(intrinsic;\ hepatic)} = 3.58 - 0.058S\_sCl - 0.57S\_aaO - 0.47Shadow\ Z\ length - 0.75CIC$	H	Commercially available drugs	(18)
$1/Log\ CL_{(intrinsic;\ hepatic)} = -3.11 - 0.10Dipole\text{-}mag + 13.25Jurs\text{-}RPCG + 0.57Jurs\text{-}RPCS + 0.00013A_{pol}$	H	Commercially available drugs	(18)
$CL_{(intrinsic;\ hepatic)} = 25Steric + 44Electrostatic + 20LUMO + 11HINT$	R	Haloalkanes	(73)
Hydrophobic descriptors			
$CL_{(intrinsic;\ hepatic)} = 0.0555P_{o:w}^{1.05}$	H	Basic drugs	(77)
$Log\ CL_{(renal)} = -0.24(Log\ P_{o:w})^2 - 0.04\ Log\ P_{o:w} + 0.58$	H	Probenecid analogs	(66)
$Log\ CL_{(renal)} = -Log(0.35 + 0.013\ P_{o:w}^{1.12})$	H	Probenecid analogs	(66)
$Log\ CL_{(renal)} = -0.5logP_{o:w} + 3$	H	NSAID	(67)
$Log\ CL_{(renal)} = -0.5LogP_{o:w} + 13$	H	ß-blockers	(67)
$Log\ CL_{(hepatic)} = -0.54\Delta R_{mui} - 0.51$	R	Sulfonamides	(66)
$Log\ CL_{(renal)} = -0.41\Delta R_{mui} - 0.80$	R	Sulfonamides	(66)
$Log\ CL_{(renal)} = -0.51LogP_{o:w} - 0.33$	R	Sulfapyridines	(75)
$Log\ CL_{(renal)} = -Log[0.048 + 6.98 \times 10^{-4}(10^{\Sigma\pi})^{1.394}$	R	Xylidines	(66)
$Log\ CL_{(total)} = -0.74\Delta R_{mui} + 0.22pK_a - 1.73$	R	Sulfonamides	(66)
$Log\ E_{(hepatic)} = 0.045LogP_{o:w} - 0.32$	R	HMWOCs	(75)
$CL_{(intrinsic;\ hepatic)} = 3.828P_{o:w}^{0.676}$	Rb	Basic drugs	(33)
$CL_{(intrinsic;\ hepatic)} = 0.0875P_{o:w}^{1.338}$	Rb	Basic drugs	(33)
$CL_{(intrinsic;\ hepatic)} = 0.248P_{o:wapp}^{1.289}$	Rb	Basic drugs	(33)

QSARS: Free-Wilson-type equations

Log $CL_{(renal)} = 0.417\ R_2(CH_3)-0.744\ R_1(OC_3H_7)+1.33$	R	Xylidines	(66)
Log $CL_{(total)} = 0.49\ R_2(CH_3)+0.57\ R_2(C_2H_5)+0.25\ R_1(OC_3H_7)+1.76$	R	Xylidines	(66)

[a] ΔE=variable related to molecular orbitals, ΔR_{mui}=variable dependant on the resistance constant due to diffusion of the non-ionized form in the lipid membrane, π = molecular hydrophobicity constant, A_{pol}=polar surface area, CIC=complementary information content, Dipole-mag=dipole moment, E_H=HOMO energy, Electrostatic=Coulombic interaction energy, Energy=minimum internal energy, HBD=potential hydrogen bond donor atoms in the molecule, Hf=enthalpy of formation, HINT=Hydrophobic field energy, IP=ionization potential, Jurs-RPCG=relative positive charge, Jurs-RPCS=relative positive charge surface area, LUMO=lowest unoccupied molecular orbital energy, $P_{o:w}$=n-octanol:water partition coefficient (or vegetable oil:water), $P_{o:w,\ app}$=apparent octanol:water partition coefficient, Q_6=net atomic charge on carbon atom at biphenyl ring position, $R_2(CH_3)$ = methyl fragment at R_2 position, $R_2(C_2H_5)$= ethyl fragment at R_2 position, $R_1(OC_3H_7)$= propyl ether fragment at R_1 position, S_aaO=E-state indice for oxygen atoms with two aromatic bonds, S_sCl=E-state indice for chlorine atoms with a single bond, Shadow Z length=length of the molecule in Z dimension, S_N=total nucleophillic superdelocalizability, and Steric=Van der Waals interaction energy.

[b] F=fish, H=human, and R=rats.

[c] CFCs = chlorofluorocarbons, HMWOCs = high molecular weight organic chemicals, LMWVOCs=low molecular weight volatile organic chemicals, and VOCs = volatile organic chemicals.

Table 6: *In silico* approaches for estimating reaction rates of chemicals[a]

Approach[b]	Species[c]	Chemical Class[d]	Reference
QSARs: LFE-type equations			
Electrostatic descriptors			
Log $V_{(oxidation)}$ =0.894-0.111diameter-0.007ΔE	H	Nitriles	(41)
Log $k_{cat\ (oxidation)}$ =19.97-0.024$\Delta H^{\neq}$-0.95IP	H	Toluenes	(41)
Log $V_{(oxidation)}$ =26.90-2.58IP	H	Halothanes	(41)
Log $k_{cat\ (oxidation)}$ =0.024Vol-0.23μ-1.14	H	Barbiturates	(41)
Log $k_{cat\ (oxidation)}$ =1.33-0.15μ	H	Anilines	(41)
Log $k_{cat\ (demethylation)}$ = -0.68 σ^{-}+ 1.06	R	$X\text{-}C_6H_4N(CH_3)_2$	(31)
Steric descriptors			
Log $k_{cat}/K_{m\ (Oxidation)}$ =0.0347SA-2.29ΔE+1.92	H	Toluenes	(41)
Log $V_{max\ (n\text{-}demethylation)}$ =0.18 Length – 1.94	H	Ethylamines	(42)
Log $V_{max\ (n\text{-}demethylation)}$ =3.50 Length – 0.13 $Length^2$-23.9	H	Ethylamines	(42)
Log $V_{(oxidation)}$ = 2.4861-0.1364NPL*NSIDE+0.5694UNS-0.2433NOM*NMC +0.001227MW*NUNSTOT+0.8242IND-1.1493MOD	R	PCBs	(49)
Log $V_{max\ (oxidation)}$=-1.676$^4x^v_c$+0.424$^3x^v_c$-0.134x^v_{pc}+1.622	R	Haloalkanes	(25)
$V_{(n\text{-}demethylation)}$ = 0.005SA-0.52	R	Amines	(41)
$V_{(n\text{-}demethylation)}$ = 0.038SA-0.00001SA^2-25.64	R	Amines	(41)
Hydrophobic descriptors			
Log $k_{cat}/K_{m\ (demethylation)}$ = 0.53 $LogP_{o.w}$ +3.47	R	$X\text{-}C_6H_4N(CH_3)_2$	(31)
Log V = 0.55 $LogP_{o.w}$	R	Barbiturates	(31)
QSARS: Free-Wilson-type equations			
V_{maxc} = $BS_{(C\text{-}C)}$ (51.6) + nH_3(14.6) + nCL(-4.84) + nCL_2(10.2) + nCL_3(-16.9)	R, H	Chloroethanes	(22)

[a] k_{cat} = catalytic rate, K_m = enzyme affinity constant, V=metabolic rate, V_{max}=maximal velocity of metabolism, and V_{maxc}=body weight normalized maximal velocity of metabolism,

[b] ΔE=LUMO energy – HOMO energy, $\Delta H^{\neq}$ = hydrogen abstraction energy, μ=dipolar moment of the molecule, $^4X^v_c$, $^3X^v_c$, X_{vpc}=connectivity indices, BS=Basic structure, diameter=diameter of the molecule, IND=variable dependant on experimental

data used, IP=ionization potential, Length=Length of the molecule, MOD=variable dependant on experimental data used, MW=molecular weight, nCL=number of CL fragments, nCL_2=number of CL_2 fragments, nCL_3=number of CL_3 fragments, nH_3=number of H_3 fragments, NMC=number of meta chlorines, NOM=number of adjacent unsubstituted ortho-meta carbon pairs, NPL=variable dependent on the number of chlorine atoms in the molecule in ortho position, NSIDE=variable dependant on the number of chlorine atoms in the molecule in meta position, NUNSTOT=variable dependant on the number of chlorine atoms in the molecule, $P_{o:w}$=n-octanol:water partition coefficient (or vegetable oil:water), SA=Surface area, UNS=variable dependant on the number of atoms in the molecule that are not chlorine, Vol=Volume of the molecule, and σ^-=Hammet constant.

[c] F=fish, H=human, and R=rats.

[d] PCBs = polychlorobiphenyls

Table 7: *In silico* approaches for estimating the Michaelis-Menten affinity constant (K_m) of chemicals

Approach[a]	Species[b]	Chemical Class[c]	Reference
QSARs: LFE-type equations			
Electrostatic descriptors			
Log $1/K_{m\ (demethylation)}$ =0.46Log$P_{o:w}$+0.63σ^-+2.62	H	X-$C_6H_4N(CH_3)_2$	(31)
Log $1/K_{m\ (sulfation)}$=0.92Log$P_{o:w}$-1.48MR_4-0.64MR_3+1.04MR_2+0.67σ^-+4.01	H	Phenols	(31)
Log $K_{m\ (acetylation)}$ =-0.42LogP_{ui}+0.14pK_a-2.89	H	Sulfonamides	(66)
$K_{m\ (oxidation)}$ = [(μ_i*a*d(π)$_i$)/|lpi-b|]+c	R	Alkenes	(15)
Log $K_{m\ (acetylation)}$ = 0.17pK_a-0.69	R	Sulfonamides	(66)
Log $K_{m\ (acetylation)}$ =-0.42$\Delta R_{mu:i}$+0.15pK_a-1.39	R	Sulfonamides	(66)
Log $K_{m\ (acetylation)}$ = 0.07pK_a+0.31Log$P_{o:w}$-0.33	Rb	Sulfonamides	(66)
Hydrophobic descriptors			
Log $1/K_{m\ (oxidation)}$ = 1.39Log $P_{o:w}$-0.22Log $P_{o:w}^2$-0.50	H	Barbiturates	(41)
Log $1/K_{m\ (sulfatation)}$ =2.93F_2+1.16π_2+0.91π_3+0.82MR_2-0.59$I_{\beta OH}$+1.29I_{ET}+2.59	H	Phenols	(31)
-Log $K_{m\ (oxidation)}$ = 43.27-4.03ΔE-0.60Log $P_{o:w}$	H	Toluenes	(41)
Log $1/K_{m\ (glucoronidation)}$=0.83Log$P_{o:w}$+1.37	R	Phenols	(31)
Log $1/K_{m\ (hydrolysis)}$ =0.056$Z1_{H2O}$+0.051$Z2_{H2O}$+0.026$Z3_{H2O}$+0.04$Z4_{H2O}$+4.616	R	Phenylhippurates	(36)
Log $1/K_{m\ (hydrolysis)}$ =0.066$Z1_{H2O}$+4.259	R	Phenylhippurates	(36)
Log $1/K_{m\ (hydrolysis)}$ =0.44π+4.08	R	Phenylhippurates	(36)
Log $1/K_{m\ (hydrolysis)}$=0.40π+4.40	R	Phenylhippurates	(36)
Log $1/K_{m\ (NADP\text{-}oxidation)}$ =0.69Log$P_{o:w}$+2.90	R	Drugs	(66)
Log $K_{m\ (n\text{-}demethylation)}$ =-0.55Log$P_{o:w}$+2.67	R	Morphines	(31)
Log $K_{m\ (oxidation)}$ = 0.61Log $P_{o:w}$ + 2.23	R	Carbamates	(31)
Log $K_{m\ (oxidation)}$ = 1.02Log $P_{o:w}$ + 2.98	R	Pyrazoles	(31)
Log $K_{m\ (oxidation)}$ = 1.05Log $P_{o:w}$ + 1.22	R	4-nitrophenyl alkyl ethers	(31)
Log $K_{m\ (oxidation)}$ =0.79Log $P_{o:w}$ + 1.46	R	Alkylbenzenes	(31)
Log $K_{m\ (oxidation)}$ = 1.04Log $P_{o:w}$ + 1.10	Rb	Toluenes	(31)
QSARs: Free-Wilson-type equations			

$Km = BS_{(C-C)}(3.8) + nH_3(-2.59) + nCL(-0.37) + nCL_2(0.79) + nCL_3(0.19)$	R, H	Chloroethanes	(22)

[a] ΔE=LUMO energy – HOMO energy, σ -=Hammet constant, μ = dipolar moment of the molecule, π,π_2,π_3=molecular hydrophobicity constants, ΔR_{mui}=variable dependant on the resistance constant due to diffusion of the non-ionized form in the lipid membrane, a = orbital availability, b = LUMO energy, BS=Basic structure, c = variable dependant on the molecular size, $d(\pi)$ = normalized electron density, F_2=variable dependant on the electrical field induced by ortho positioned atoms, $I\beta_{OH}$=variable dependant on the number of β OH groups in the molecule, I_{ET}=variable dependant on the family of the substance, IP = ionization potential, $MR_{2,3,4}$=Molar refractivity indices, nCL=number of CL fragments, nCL_2=number of CL_2 fragments, nCL_3=number of CL_3 fragments, nH_3=number of H_3 fragments, pK_a=log dissociation constant of an acid in water, $P_{o:w}$=n-octanol:water partition coefficient (or vegetable oil:water), P_{ui}=n-octanol:water partition coefficient for the non-ionized form, and $Z1,2,3,4_{H2O}$=variables corresponding to the potential energy for the interaction between the molecule and water.

[b] F=fish, H=human, and R=rats.

[c] PCBs = polychlorobiphenyls

Table 8: *In silico* approaches for estimating the skin permeability coefficient (K_p) of chemicals

Approach[a]	Species[b]	Chemical Class[c]	Reference
QSARs: LFE-type equations			
Electrostatic descriptors			
$\text{Log } K_p = -0.626\Sigma Ca - 23.8\Sigma(Q+)/\alpha - 0.289SsssCH - 0.0357SsOH - 0.482I_B + 0.405B_R + 0.834$	H	LMWVOCs; HMWOCs	(45)
$\text{Log } K_p = 0.44R_2 - 0.49\pi_2^H - 1.48\Sigma\alpha_2^H - 3.44\Sigma\beta_2^H + 1.94V_x - 5.13$	H	LMWVOCs; HMWOCs	(45)
$\text{Log } K_p = -0.59\pi_2^H - 0.63\Sigma\alpha_2^H - 3.48\Sigma\beta_2^H + 1.79V_x - 5.05$	H	LMWVOCs; HMWOCs	(45)
$\text{Log } K_p = -5.33 - 0.62\pi_2^H - 0.38\Sigma\alpha_2^H - 3.34\Sigma\beta_2^H + 1.85V_x$	H	Alcohols, Steroids	(27)
Steric descriptors			
$K_p = (b_1 + 0.0025/(b_2 + b_3 + P_{o:w}^{b4}))^{-1} MW^{b5}$	H	LMWVOCs; HMWOCs	(45)
$K_p = (b_1 + b_2 P_{o:w}) e^{(b3MW)}$	H	LMWVOCs; HMWOCs	(45)
$\text{Log } K_p = -5.14 - 0.47\Sigma Ca + 0.23\Sigma Cd + 0.038Pol$	H	Alcohols, Steroids	(60)
$\text{Log } K_p = -6.14 - 0.42\Sigma Ca + 0.23\Sigma Cd + 0.21L - 0.11W$	H	Alcohols, Steroids	(60)
$\text{Log } K_p = -7.29 + 0.15Pol$	H	Alcohols	(60)
$\text{Log } K_p = b_1 + b_2 \text{Log} P_{o:w} + b_3 MW^{0.5}$	H	LMWVOCs; HMWOCs	(45)
$\text{Log } K_p = -0.428\delta - 4.80^4 X_{PC}^V + 28.06$	H	Hydrocorticone esters	(27)
$\text{Log } K_p = 0.652\text{Log} P_{o:w} - 0.00603MW - 0.623ABSQon - 0.313SsssCH - 2.3$	H	Dermal drugs; LMWVOCs; HMWOCs	(50)
$\text{Log } K_p = 0.77\text{Log} P_{o:w} - 0.0103MW - 2.33$	H	LMWVOCs; HMWOCs	(45)
$\text{Log } K_p = -0.786OT + 0.252^2\kappa - 1.617q_s^+ - 5.767$	H	Alcohols, Steroids	(27)
$\text{Log } K_p = 0.82\text{Log} P_{o:w} - 0.0093V_m - 0.039MP_t - 2.36$	H	Steroids	(45)
$\text{Log } K_p = 0.84\text{Log} P_{o:w} - 0.07(\text{log} P_{o:w})^2 - 0.27Hb - 1.84\text{LogMW} + 4.39$	H	LMWVOCs; HMWOCs	(45)
$\text{Log } K_p = 28.4q^- + 0.018V_m + 2.824$	H	Barbiturates; Isoquinoline; Salicyclic acid	(27)
$\text{Log } K_p = 3.99\text{logTA} + 4.53q_s^- - 0.762OT - 11.364$	H	Alcohols, Steroids	(27)
Hydrophobic descriptors			
$K_p = 1.17*10^{-7} P_{o:w}^{0.751} + 2.73*10^{-8}$	H	Pharmaceuticals	(45)
$K_p = b_1(P_{o:w}^{b2}/(b_3 + P_{o:w}^{b2}))$	H	HMWOCs	(45)
$\text{Log } K_p = -0.207\ \text{Log} P_{o:w}^2 + 1.49\ \text{Log} P_{o:w} - 5.42$	H	Steroids	(66)
$\text{Log } K_p = -0.37\text{Log} P_{o:w}^2 + 2.39\text{Log} P_{o:w} - 8.71$	H	Phenols	(69)
$\text{Log } K_p = 0.544\text{log} P_{o:w} - 2.88$	H	Aliphatic alcohols	(66)
$\text{Log } K_p = 0.80\text{log} P_{o:w} - 8.883$	H	Hydrocorticone esters	(27)

Log K_p=-1.46ΔLog$P_{o:w}$+0.29Log$P_{o:w}$-3.75	H	Alcohols, steroids	(69)
Log K_p=-4.36-0.38ΣCa+0.24ΣCd	H	Steroids	(60)
Mechanistically-based equations			
K_p =($P_{vo:w}$*0.028D_l/0.0340)+($P_{p:w}$*0.88D_p/0.0018)	H	Acids; Alcohols; Hydrocarbons	(59)

[a] δ=solubility parameter, π_2^H=dipolarity/polarizability, α_2^H=overall hydrogen-bond acidity, β_2^H=overall hydrogen-bond basicity, ΣCa=Hydrogen bond acceptor free energy in the molecule, ΣCd=Hydrogen bond donor in the molecule, $^2\kappa$=molecular shape index, $^4X_{pc}^v$=connectivity indices, ABSQon=sum of absolute charges on oxygen and nitrogen atoms, b_1,b_2,b_3,b_4,b_5=regression coefficients without any assigned role, B_R=number of rotatable bonds, D_l=coefficient for diffusion into the lipid fraction of stratum corneum, D_p=coefficient for diffusion into the protein fraction of stratum corneum, Hb=number of hydrogen bonds formed by the substance, I_B=Balaban index, L=molecular length, MP_t=melting point, MW=molecular weight, OT=number of hydrogen bonding heteroatoms, $P_{o:w}$=n-octanol:water partition coefficient (or vegetable oil:water), Pol=describes bulk or volume related effects, $P_{p:w}$=Protein:water partition coefficient for stratum corneum, $P_{vo:w}$=vegetable oil:water partition coefficient, q^-=the most negative charge on the hydrogen bond accepting heteroatoms, Q^+/α=positive charge per unit volume, q_s^-=sum of atomic charges on hydrogen bonding heteroatoms, q_s^+=sum of atomic charges on hydrogen bonding hydrogens, R_2=Excess molar refraction, SsOH=sum of E-state indices for all hydroxy groups, SsssCH=sum of E-state indices for all methyl groups, TA=total solvent accessible surface, V_m=molar volume, V_x=McGowan characteristic volume, and W=molecular width.

[b] F=fish, H=human, and R=rats.

[c] CFCs = chlorofluorocarbons, HMWOCs = high molecular weight organic chemicals, LMWVOCs=low molecular weight volatile organic chemicals, and VOCs = volatile organic chemicals.

Table 9: *In silico* approaches for estimating the oral absorption constant (K_a) of chemicals

Approach[a]	Species[b]	Chemical Class	Reference
QSARs: LFE-type equations			
Electrostatic descriptors			
$Log\ K_{a\ (absorption)} = -0.58 Log\ P_{o:w} + 0.35 pK_a - 1.77$	F	Barbiturates	(66)
Hydrophobic descriptors			
$K_{a\ (absorption)} = k_m[(1/R_f)-1]^{n'}/(Q+[(1/R_f)-1]^{n'})$	R	Sulfonamides	(69)
$Log\ K_{a\ (absorption)} = -0.04\ (Log\ P_{o:w})^2 + 0.22 Log\ P_{o:w} + 0.04$	R	Sulfonamides	(66)
$Log\ K_{a\ (absorption)} = 0.067 + LogP_{o:w} - Log(1.4 + P_{o:w})$	R	Pharmaceuticals	(75)
$Log\ K_{a\ (absorption)} = -0.082(Log\ P_{o:w})^2 + 0.268 Log\ P_{o:w} + 3.96$	R	Organic acids	(66)
$Log\ K_{a\ (absorption)} = -0.09\ (Log\ P_{o:w})^2 + 0.44 Log\ P_{o:w} - 0.396$	R	Sulfonamides	(66)
$Log\ K_{a\ (absorption)} = 0.09 Log\ P_{o:w} + 0.83$	R	Xanthenes	(66)
$Log\ K_{a\ (absorption)} = 0.18\ Log\ P_{o:w} + 0.23$	R	Carbamates	(66)
$Log\ K_{a\ (absorption)} = 0.24 Log\ P_{o:w} - 1.37$	R	Antihistamines	(66)
$Log\ K_{a\ (absorption)} = 0.30\ Log\ P_{o:w} - 0.07(Log\ P_{o:w})^2 - 2.38$	R	Sulfonylureas	(66)
$Log\ K_{a\ (absorption)} = 0.3 Log\ P_{o:w} - 0.57 Log(0.34\ P_{o:w} + 1) - 0.15I - 0.74$	R	Carbamates	(66)
$Log\ K_{a\ (absorption)} = 0.46 Log\ P_{o:w} - 0.36 Log(0.60 P_{o:w} + 1) - 0.23$	R	Sulfonylureas	(66)
$Log\ K_{a\ (absorption)} = 0.5\ Log\ P_{o:w} - 0.61 Log(0.07\ P_{o:w} + 1) - 0.39$	R	Sulfonamides	(66)
$Log\ K_{a\ (absorption)} = 0.502\ Log\ P_{o:w} - Log(0.053\ P_{o:w}^{0.0862} + 1) - 0.384$	R	Sulfonamides	(66)
$Log\ K_{a\ (absorption)} = 0.56 Log\ P_{o:w} - (0.04\ P_{o:w}^{0.84} + 1) - 0.63$	R	Phenols; Anilines; Esters	(66)
$Log\ K_{a\ (absorption)} = 1.36 Log\ P_{o:w} + 0.36$	R	Organic anions	(66)
QSARs: Free-Wilson-type equations			
$Log\ K_{a\ (absorption)} = BS_{(BEN\text{-}SO2\text{-}NHCONH)}\ (-2.272) + n{*}H\ (0) + n{*}2\text{-}CH_3\ (0.088) + n{*}4\text{-}CH_3\ (0.074) + n{*}4\text{-}C_2H_5\ (0.163) + n{*}4\text{-}OCH_3\ (-0.229) + n{*}2\text{-}NO_2\ (-0.324) + n{*}3\text{-}NO_2\ (-0.207) + n{*}4\text{-}NO_2\ (-0.323) + n{*}4\text{-}Cl\ (0.198) + n{*}4\text{-}Br\ (0.122) + n{*}n\text{-}C_4H_9\ (0) + n{*}CH_3\ (-0.638) + n{*}C_2H_5\ (-0.361) + n{*}n\text{-}C_3H_7\ (-0.145) + n{*}i\text{-}C_3H_7\ (-0.244) + n{*}i\text{-}C_4H_9\ (-0.035) + n{*}t\text{-}C_4H_9\ (0.149) + n{*}cy\text{-}C_6H_{11}\ (0.135) + n{*}allyl\ (-0.419) + n{*}C_6H_5\ (-0.088)$	R	Sulfonylureas	(66)

[a] allyl=allyl group, Br=Bromine, BS=Basic structure, C_2H_5=Ethyl group, C_6H_5=aromatic ring group, CH_3=Methyl group, Cl=Chlorine group, cy-C_6H_{11}=cyclohexyl group, H=Hydrogen group, I=family indicator variable, i-C_3H_7=iso-propyl group, i-C_4H_9=iso-butyl group, k_m=rate constant for transfer out of the membrane, n=group occurrence in molecule, n'=constant specific to the equation without any given role, n-C_3H_7=n-propyl group, n-C_4H_9=n-butyl group, NO_2=Nitrooxide group, OCH_3=Methyl

ether group, pK_a=log dissociation constant of an acid in water, $P_{o:w}$=n-octanol:water partition coefficient (or vegetable oil:water), Q=constant specific to the equation without any given role, R_f=reverse-phase TLC lipophilicity parameter, and t-C_4H_9=tert-butyl group.

[b] F=fish and R=rats.

Table 10 : Frequency of occurrence of molecular fragments for each chloroethane of the series.

Chemical	**BS[a]**	**H_3**	**Cl**	**Cl_2**	**Cl_3**
Chloroethane	1	1	1	0	0
1,1-dichloroethane	1	1	0	1	0
1,2-dichloroethane	1	0	2	0	0
1,1,1-trichloroethane	1	1	0	0	1
1,1,2-trichloroethane	1	0	1	1	0
1,1,1,2-tetrachloroethane	1	0	1	0	1
1,1,2,2-tetrachloroethane	1	0	0	2	0
Pentachloroethane	1	0	0	1	1
Hexachloroethane	1	0	0	0	2

[a]BS = basic structure (C-C)

Table 11 : Contributions[a] of chloroethane structural features to rat partition coefficients[b] and metabolic constants[c]

Fragments	P_b	P_l	P_s	P_f	V_{maxc}	K_m
BS	56.8	2.02	0.746	28.9	52.7	3.75
Cl_2	42.7	-0.319	-0.0181	-1.16	9.40	0.863
Cl_3	7.00	1.60	0.233	14.1	-15.3	-0.0932
Cl	-9.60	-0.506	0.00710	-7.22	-7.22	-0.234
H_3	-50.1	-0.653	-0.0770	-8.56	12.9	-1.65
r^2	0.98	0.91	0.99	0.96	0.82	0.88

[a] Contributions were obtained by multiple linear regression from experimental data on chloroethane, 1,1- dichloroethane, 1,2 –dichloroethane, 1,1,2- trichloroethane, 1,1,1,2- tetrachloroethane, 1,1,2,2- tetrachloroethane, pentachloroethane, and hexachloroethane. BS = basic structure (C-C).

[b] P_b, P_l, P_s and P_f refer to blood:air, liver:blood, slowly perfused tissue:blood and fat:blood partition coefficients, respectively.

[c] V_{maxc} (µmol/hr/kg) and K_m (µM) refer to maximal velocity of metabolism and affinity constant, respectively.

Table 12 : Contributions[a] of chloroethane structural features to human partition coefficients[b]

Fragments	P_b	P_l	P_s	P_f
BS	37.4	2.72	1.099	38.9
Cl_2	29.6	-0.365	-0.163	0.105
Cl_3	7.53	2.16	0.166	12.2
Cl	-8.92	-0.446	0.0510	-10.6
H_3	-39.3	-0.699	0.0450	-1.05
r^2	0.83	0.98	0.91	0.94

[a] Contributions were obtained by multiple linear regression from experimental data on chloroethane, 1,1- dichloroethane, 1,2 –dichloroethane, 1,1,2- trichloroethane, 1,1,1,2- tetrachloroethane, 1,1,2,2- tetrachloroethane, and hexachloroethane. BS = basic structure (C-C).

[b] P_b, P_l, P_s and P_f refer to blood:air, liver:blood, slowly perfused tissue:blood and fat:blood partition coefficients, respectively.

Table 13: Comparison of experimental[a] (Exp) and QSAR-estimated (Est) values of rat partition coefficients[b] and metabolic constants[c] for 1,1,1-trichloroethane.

Parameter	Exp	Est
P_b	5.67	13.7
P_l	1.52	2.97
P_s	0.56	0.90
P_f	46.4	34.5
V_{maxc}	43.1	50.3
K_m	3.14	2.01

[a] Experimental data from Fouchécourt and Krishnan (2000).

[b] P_b, P_l, P_s and P_f refer to blood:air, liver:air, slowly perfused tissue:air and fat:air partition coefficients, respectively.

[c] V_{maxc} (µmol/hr/kg) and K_m (µM) refer to maximal velocity of metabolism and affinity constant, respectively.

Table 14 Comparison of experimental[a] (Exp) and QSAR-estimated (Est) values of human partition coefficients[b] for 1,1,1-trichloroethane.

Parameter	Exp	Est
P_b	2.53	5.56
P_l	3.40	4.18
P_s	1.25	1.31
P_f	104	50.1

[a] Experimental data from Fouchécourt and Krishnan (2000).

[b] P_b, P_l, P_s and P_f refer to blood:air, liver:air, slowly perfused tissue:air and fat:air partition coefficients, respectively.

Table 15 : Steady-state tissue concentrations (µg/L) of 1,1,1-trichloroethane in rat and humans estimated using the conventional (PBPK) and QSAR-based (QSAR) physiologic model following a continuous exposure to 1 ppm.

Tissue	Rat		Human	
	QSAR	PBPK	QSAR	PBPK
Blood	22.9	16.6	12.8	8.5
Liver	6.77	4.22	6.23	5.61
Slowly perfused	20.7	9.31	16.7	10.7
Fat	790	770	502	472
Richly perfused	68.1	25.2	53.4	28.9

Table 16: Steady-state arterial blood concentration (Ca_{ss}) obtained using the conventional (PBPK) and QSAR-based (QSAR) physiological model in rats exposed to the NOAEL of 1,1,1-trichloroethane (875 ppm) and the corresponding environmental concentration (C_i) in humans derived using the human conventional (PBPK) and QSAR-based (QSAR) physiological models

Endpoint	QSAR	PBPK
Rat Ca_{ss} (mg/L)	59.3	24.9
Human C_i (ppm)[a]	6342	4252

[a] Calculated using the QSAR-derived Ca_{ss} (59.3 mg/L).

FIGURE LEGENDS

Figure 1: Chemical description methodology used in this study. The chemicals are represented as a basic structure (C-C) with substituents on the two carbons. Examples of the description of 1,1,1 trichloroethane and 1,1,2,2 tetrachloroethane are presented.

Figure 2: Comparison of rat experimental and predicted parameter values. Experimental values from Fouchécourt and Krishnan (2000).

Figure 3: Comparison of human experimental and predicted parameter values. Experimental values from Fouchécourt and Krishnan (2000).

Figure 4: Quantitative structure-activity relationship (QSAR) physiologically-based pharmacokinetic (PBPK) modeling framework. User input consists of the exposure scenario and chemical structure information such as the number of fragments constituting the molecule. This information is fed to the program which contains the model constants, the Free-Wilson type SPR, the contribution values of each molecular fragment (C_s) and of the basic structure (BS) to the model parameters (P), and the simulation algorithms. The model can then simulate the pharmacokinetics of the chemical in biota and provide its profile as output. The example of 1,1,1 trichloroethane is shown.

Figure 5: Quantitative structure-activity relationship (QSAR) based physiologically based pharmacokinetic (PBPK) model tissue concentrations versus conventional PBPK model tissue concentrations in rats as estimated in this study.

Figure 1

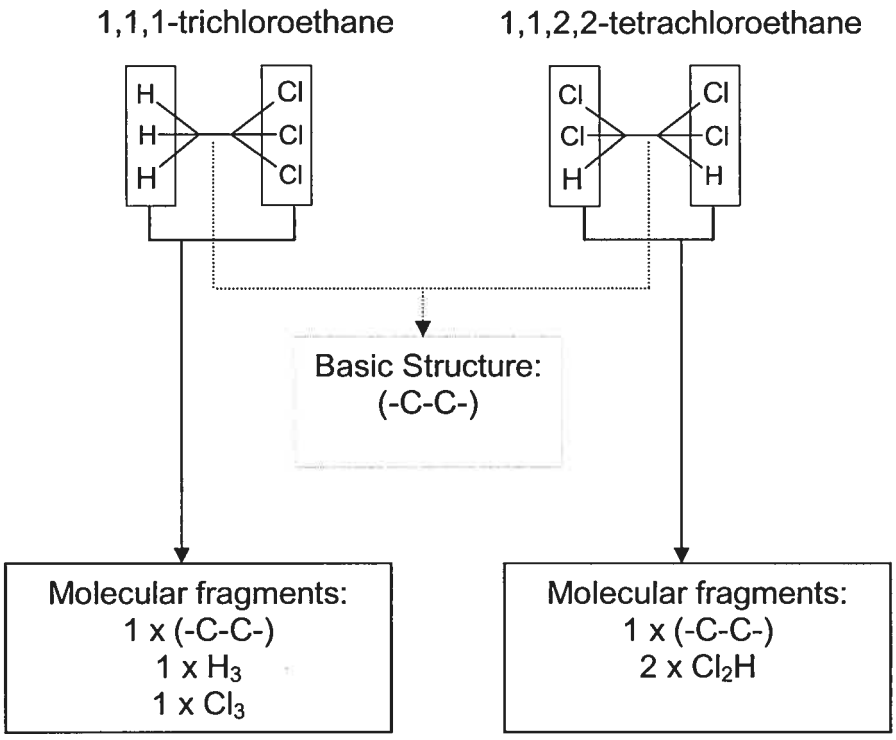

1,1,1-trichloroethane
1,1,2,2-tetrachloroethane
H
H
H
Cl
Cl
Cl
Cl
Cl
H
Cl
Cl
H
Basic Structure:
(-C-C-)
Molecular fragments:
1 x (-C-C-)
1 x H_3
1 x Cl_3
Molecular fragments:
1 x (-C-C-)
2 x Cl_2H

Figure 2

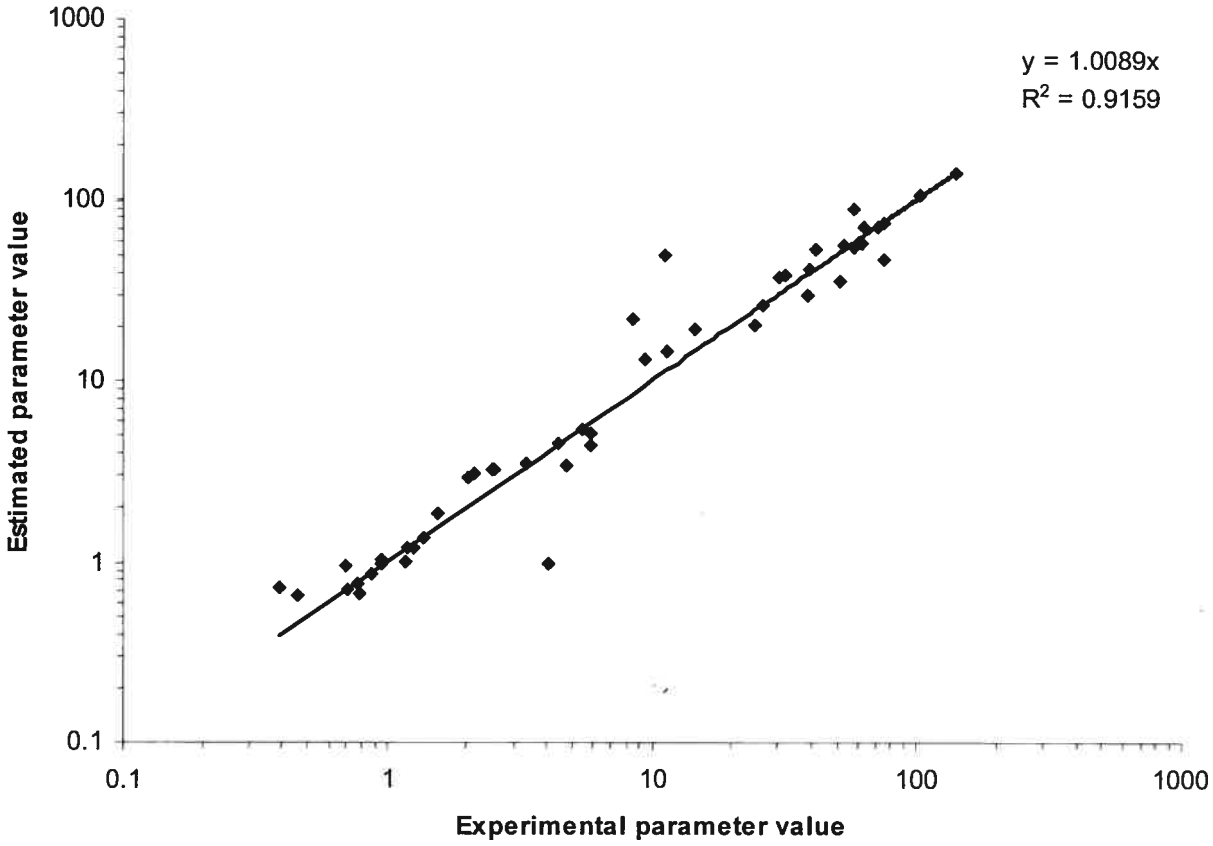

y = 1.0089x
R2 = 0.9159
Estimated parameter value
Experimental parameter value
1000
100
10
1
0.1
0.1
1
10
100
1000

Figure 3

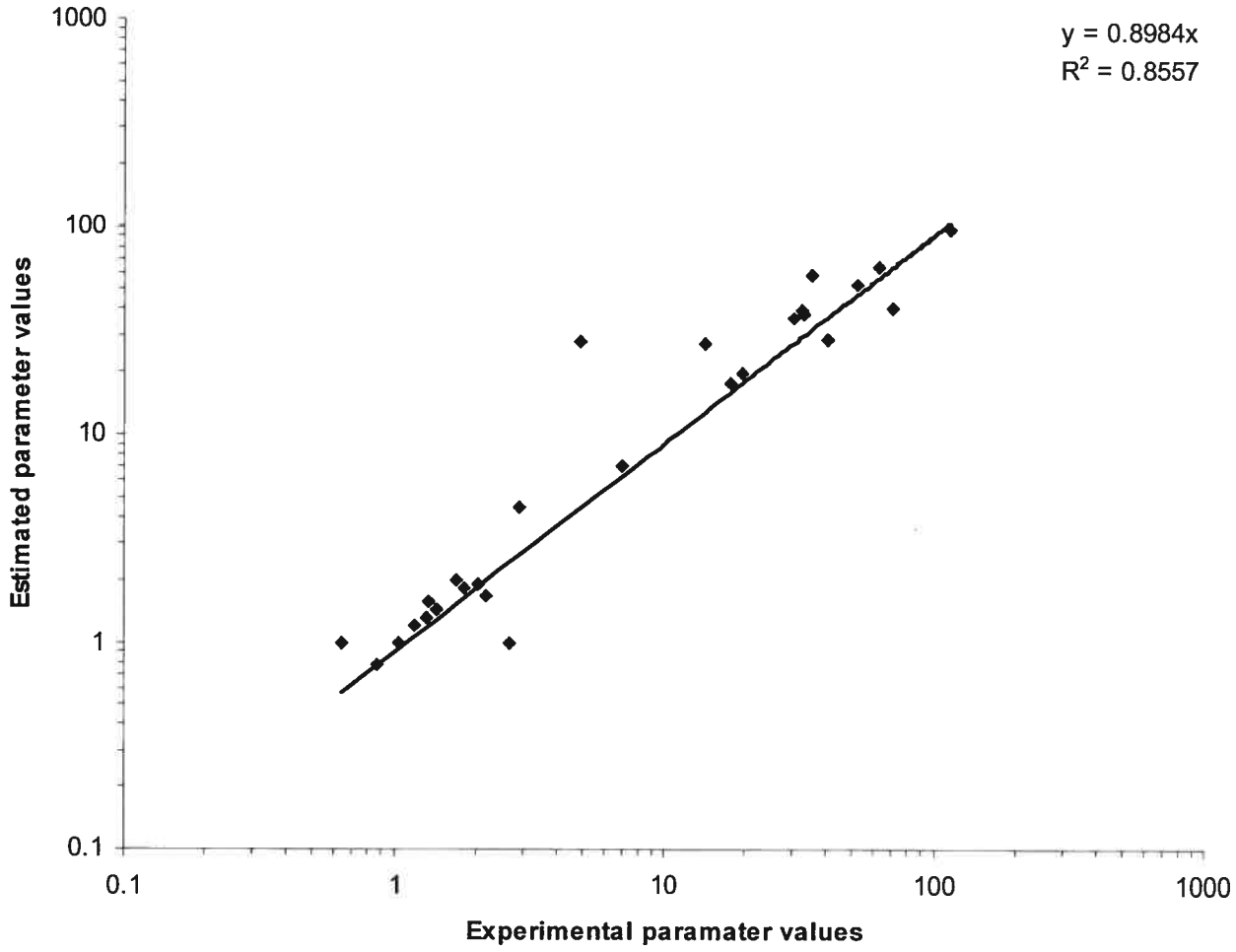

y = 0.8984x
R2 = 0.8557
Estimated parameter values
Experimental paramater values
1000
100
10
1
0.1
0.1
1
10
100
1000

Figure 4

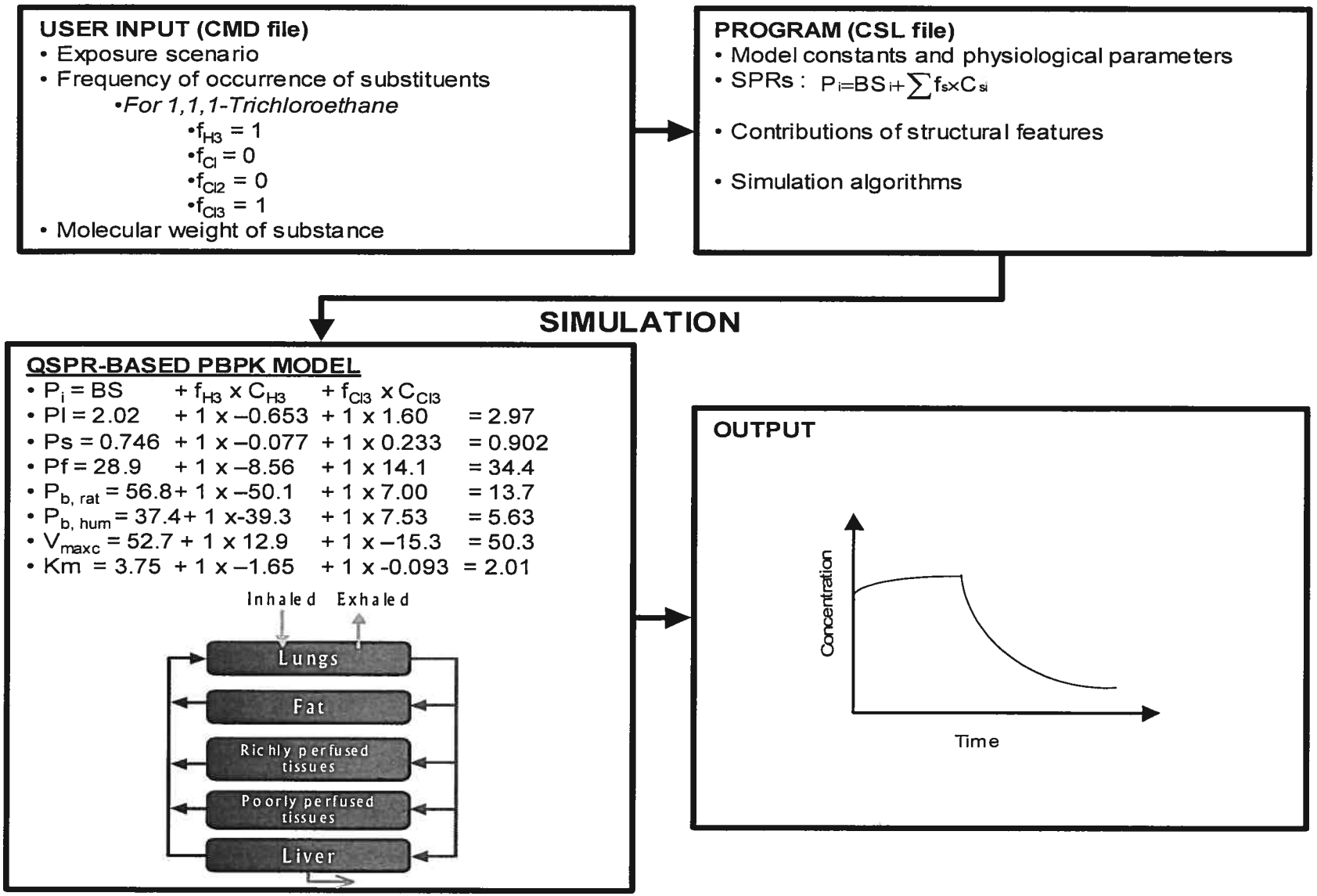
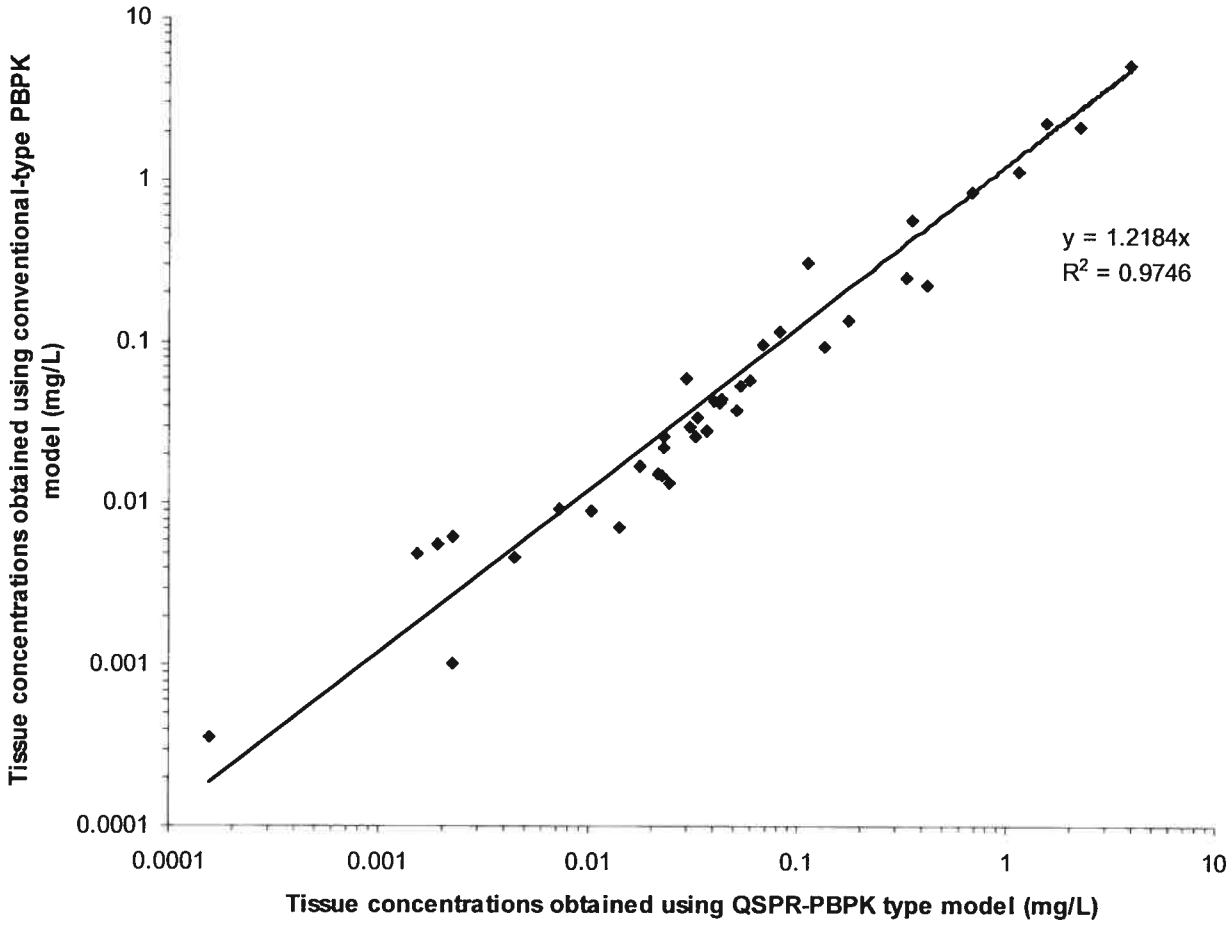

USER INPUT (CMD file)
• Exposure scenario
• Frequency of occurrence of substituents
•For 1,1,1-Trichloroethane
•f_{H3} = 1
•f_{Cl} = 0
•f_{Cl2} = 0
•f_{Cl3} = 1
• Molecular weight of substance
PROGRAM (CSL file)
• Model constants and physiological parameters
• SPRs : $P_i = BS_i + \sum f_s \times C_{si}$
• Contributions of structural features
• Simulation algorithms
SIMULATION
QSPR-BASED PBPK MODEL
• P_i = BS + f_{H3} x C_{H3} + f_{Cl3} x C_{Cl3}
• PI = 2.02 + 1 x –0.653 + 1 x 1.60 = 2.97
• Ps = 0.746 + 1 x –0.077 + 1 x 0.233 = 0.902
• Pf = 28.9 + 1 x –8.56 + 1 x 14.1 = 34.4
• $P_{b, rat}$ = 56.8+ 1 x –50.1 + 1 x 7.00 = 13.7
• $P_{b, hum}$ = 37.4+ 1 x-39.3 + 1 x 7.53 = 5.63
• V_{maxc} = 52.7 + 1 x 12.9 + 1 x –15.3 = 50.3
• Km = 3.75 + 1 x –1.65 + 1 x -0.093 = 2.01
Inhaled
Exhaled
Lungs
Fat
Richly perfused tissues
Poorly perfused tissues
Liver
OUTPUT
Concentration
Time

Figure 5


Tissue concentrations obtained using conventional-type PBPK model (mg/L)
Tissue concentrations obtained using QSPR-PBPK type model (mg/L)
y = 1.2184x
R2 = 0.9746
10
1
0.1
0.01
0.001
0.0001
0.0001
0.001
0.01
0.1
1
10

CHAPITRE TROISIEME:

3 Article II

Quantitative structure-property relationships for physiologically-based pharmacokinetic modeling of volatile organic chemicals in rats

Quantitative structure-property relationships for physiologically-based pharmacokinetic modeling of volatile organic chemicals in rats

Martin Béliveau, Robert Tardif and Kannan Krishnan

Groupe de recherche en toxicologie humaine (TOXHUM), Université de Montréal, Case Postale 6128, Succ. Centre-Ville, Montréal, PQ, H3C 3J7, Canada.

Running title: QSPRs for PBPK modeling

Address for correspondence: Dr. Kannan Krishnan, Département de santé environnementale et santé au travail, Faculté de médecine, Université de Montréal, 2375 Chemin de la Côte Ste-Catherine, Bureau 4105, Montréal, Québec, Canada, H3T 1A8. Tel: (514) 343-6581, Fax: (514) 343-2200. email: [redacted]

Quantitative structure-property relationships for physiologically-based pharmacokinetic modeling of volatile organic chemicals in rats. Béliveau, M., Tardif, R., and Krishnan, K. (2003). *Toxicol. Appl. Pharmacol.* **189** 221-232.

The objective of present study was to develop quantitative structure-property relationships (QSPRs) for the chemical-specific input parameters of rat PBPK models (i.e., blood:air partition coefficient (P_b), liver:air partition coefficient (P_l), muscle:air partition coefficient (P_m), fat:air partition coefficient (P_f), and hepatic clearance (CL_h)), for simulating the inhalation pharmacokinetics of volatile organic chemicals (VOCs). The literature data on P_b, P_l, P_f, and P_m for 46 low molecular weight VOCs as well as CL_h for 25 such VOCs primarily metabolized by CYP2E1 (alkanes, haloalkanes, haloethylenes, and aromatic hydrocarbons) were analysed to develop QSPRs. The QSPRs developed in this study were essentially multilinear additive models, which imply that each fragment in the molecular structure has an additive and constant contribution to partition coefficients and hepatic clearance. Most of the values in the calibration set could be reproduced adequately with the QSPR approach which involved the calculation of the sum of the frequency of occurrence of fragments (CH_3, CH_2, CH, C, C=C, H, Cl, Br, F, benzene ring, and H in benzene ring structure) times the fragment-specific contributions determined in this study. The QSPRs for P_b, P_l, P_m, P_f and CL_h were then included within a PBPK model which only required the specification of the frequency of occurrence of fragments in a molecule along with exposure concentration and duration as input for conducting pharmacokinetic simulations. This QSPR-PBPK model framework facilitated the prediction of the inhalation pharmacokinetics of four VOCs present in the calibration dataset (toluene,

dichloromethane, trichloroethylene, and 1,1,1-trichloroethane), and four VOCs which were not part of the calibration set (1,2,4-trimethyl benzene, ethyl benzene, 1,3-dichloropropene and 2,2-dichloro-1,1,1-trifluoroethane) but which could be described using the molecular fragments investigated in the present study. The QSPRs developed in this study should be potentially useful for providing a "first-cut" evaluation of the inhalation pharmacokinetics of VOCs prior to experimentation, as long as the number and nature of the fragments do not exceed the ones in the calibration dataset used in this study.

Keywords: QSAR, PBPK models, pharmacokinetics, VOCs.

3.1 Introduction

Internal dose, rather than the exposure concentration, is increasingly being used in health risk assessment to enhance the scientific basis of the process (Andersen *et al.*, 1987; Clewell, III *et al.*, 2002; Krishnan and Andersen, 2001). The use of internal dose in risk calculations facilitates the consideration of the appropriate toxic moiety as well as exposure scenario and interspecies differences in pharmacokinetics. Direct measurements of internal dose, however, are not always possible. In such circumstances, physiologically-based pharmacokinetic (PBPK) models are of use in simulating the internal dose of chemicals in exposed biota. PBPK models simulate pharmacokinetics and internal dose of chemicals on the basis of the interrelationships among certain critical mechanistic determinants such as physiological (cardiac output, alveolar ventilation rate, tissue blood flow rates, and tissue volumes), physicochemical (blood:air and tissue:blood partition coefficients), and biochemical parameters (maximal velocity of metabolism [V_{max}] and Michaelis affinity constant [K_m], or alternatively hepatic clearance) (Krishnan and Andersen, 2001).

The construction of PBPK models for new chemicals requires the knowledge of physiological, physicochemical and biochemical parameters. Although the numerical values of the species-specific physiological parameters are available in the literature (Arms and Travis, 1988), this is not necessarily the case for the chemical-specific partition coefficients and biochemical parameters. These parameters are usually determined by *in vitro* or *in vivo* experimentation (Krishnan and Andersen, 2001). Partition coefficients required for PBPK modeling can as well be obtained using animal-replacement methods (reviewed in Payne and

Kenny, 2002). A method proposed by Poulin and Krishnan (1996) facilitates the prediction of tissue:air, blood:air and tissue:blood partition coefficients from knowledge of species-specific tissue and blood composition (i.e., water, neutral lipid and phospholipid content), as well as n-octanol:water partition coefficient ($P_{o:w}$) and water:air partition coefficient ($P_{w:a}$) of chemicals. $P_{w:a}$ and $P_{o:w}$, in turn, can be predicted from knowledge of the molecular structure of organic chemicals using algorithms developed by Hine and Mookerjee (Hine and Mookerjee, 1975) and Hansch and Leo (Hansch and Leo, 1979). However, the molecular fragments required to predict $P_{w:a}$ are different from the molecular fragments required to predict $P_{o:w}$ and as such tissue:air and blood:air partition coefficients required for PBPK modeling cannot be predicted using a single set of fragments as per the existing methods. Conventional QSAR-type approach represents a feasible alternative. Accordingly, specific structural fragments can be hypothesized to consistently contribute to PBPK model parameter values, specifically the partition coefficients and biochemical parameters. It should then be possible to compute the numerical values of these chemical-specific PBPK model parameters by summing up the contribution of each molecular fragment towards each parameter.

The objectives of this study were therefore: (i) to develop quantitative structure-property relationships (QSPRs) for the various input parameters of rat PBPK models (i.e., blood:air partition coefficient (P_b), liver:air partition coefficient (P_l), muscle:air partition coefficient (P_m), fat:air partition coefficient (P_f), and hepatic clearance (CL_h)), and (ii) to verify the usefulness of such QSPRs by incorporating them within a PBPK model to simulate the inhalation pharmacokinetics of volatile organic chemicals (VOCs).

3.2 Methods

QSPRs for PBPK Model Parameters

Quantitative relationships between molecular fragments of VOCs belonging to various classes (alkanes, haloalkanes, haloethylenes, and aromatic hydrocarbons) and the chemical-specific parameters required for PBPK modeling (P_b, P_l, P_m, P_f and CL_h) were developed according to an additive model by Gao *et al.* (Gao *et al.*, 1992). In this modeling approach, chemical fragments or groups are used to describe a molecule, facilitating the description of all atoms in a molecule. As such, each molecular fragment is assumed to have a unique and finite contribution to the PBPK model parameter. Table 1 presents the molecular fragments used to describe the VOCs investigated in this study. The alkanes were described using the basic UNIFAC (UNIversal Function group Activity Coefficient) groups such as CH_3, CH_2, CH, and C (Ochsner and Sokoloski, 1985; Hansen *et al.*, 1991). Whereas the alkenes were described using two molecular fragments, namely, C=C and H, substituent groups in the alkanes and alkenes included Cl, Br, and F. The benzene ring skeleton was described as a single fragment (i.e., AC), and the hydrogen atoms attached to the ring were represented separately (H_AC). The characterization and representation of molecular fragments in VOCs investigated in the present study, as described above, were done according to Martin and Young (Martin and Young, 2001) who successfully developed a QSAR model of acute toxicity of organic chemicals to aquatic species using the additive model of Gao *et al.* (Gao, Govind, and Tabak, 1992). The VOCs investigated in the present study can be reconstituted by adding the frequency of occurrence of each fragment shown in Table 1.

The experimentally-determined mean values of P_b, P_l, P_f, and P_m for the VOCs listed in Table 1 were obtained from Gargas *et al.* (1989). The partition coefficients had been determined using blood and tissues of Fischer-344 rats as per the vial equilibration technique of Sato and Nakajima (1979). The partition coefficients reported by Gargas *et al.* (1989) correspond to experimental measures and not derived with the knowledge of lipid and water content of tissue and blood. The metabolism constants (V_{maxc} (mg/kg/hr), K_m (mg/L) or Kf_c ($hr^{-1}\cdot kg^{-1}$) for several VOCs with partition coefficient data in Gargas *et al.* (1989) were obtained from Gargas *et al.* (1988) (i.e., chloromethanes, chloroethanes, and chloroethylenes). Since the metabolic constants for some VOCs in Gargas *et al.* (1989) were not determined by Gargas *et al.* (1988), they were obtained from other literature sources (Ramsey and Andersen (1984), Lily *et al.* (1998), Ali and Tardif (1999), Bogaards *et al.* (2001) and Haddad *et al.* (2000)).

The V_{maxc} values were used to calculate V_{max} (mg/hr) using the equation $V_{maxc}*BW^{0.74}$ where BW = 0.25 kg. V_{max} and K_m values were then used to calculate CL_h as follows (Renwick, 2001):

$$CL_h = \frac{Q_l * V_{max}/K_m}{V_{max}/K_m + Q_l} \qquad [1]$$

where Q_l = liver blood flow rate (1.34 L/hr, Haddad *et al.*, 2000). In case of chemicals reported to be eliminated via a first order process rather than (or in addition to) a saturable pathway (dichloromethane, vinyl chloride, tetrachloroethylene, chloroethane, and 1,1,1-trichloroethane) (Gargas *et al.* 1988), Kf_c ($hr^{-1}\cdot kg^{-1}$) values reported by Gargas *et al.* (1988) were used to calculate Kf (hr^{-1}) using the equation $Kf_c*BW^{-1/3}$ where BW = 0.25 kg. V_{max}, K_m, and Kf values were then used to calculate CL_h as follows:

$$CL_h = \frac{Q_l * [(V_{max}/K_m) + Kf * V_l]}{[(V_{max}/K_m) + Kf * V_l] + Q_l} \qquad [2]$$

where V_l = volume of liver (0.01 L, Haddad *et al.*, 2000) and Kf = first order elimination rate (hr^{-1}). The CL_h value obtained as above is usually applicable for relatively low exposure concentrations when the K_m is greater than the venous blood concentration leaving the liver (Renwick, 2001).

The numerical values of P_b, P_l, P_f, P_m, and CL_h (i.e., P_{PBPK}) were used along with the frequency of occurrence (f) of each fragment (i) in the various VOCs in a multilinear additive model (Martin and Young, 2001):

$$\log P_{PBPK} = \sum_{i=1}^{ng} f_i \cdot C_i \qquad [3]$$

to characterize the numerical value representing the contribution of each fragment i (C_i). The above QSPR implies that each fragment in the structure has an additive and constant contribution to the parameter of interest (P_{PBPK}).

Multilinear regressions were performed using a commercially-available statistical software package (SPSS® for Windows® v10.0.7, SPSS Inc., Chicago, IL), and the results obtained were essentially the values of C_i for each of the eleven structural fragments (CH_3, CH_2, CH, C, C=C, H, Cl, Br, F, AC, H_AC) specific to the PBPK parameter of interest. The statistical significance ($p < 0.05$) of the regression coefficients was assessed by computing *t*-statistic (> 1.8) (Martin and Young, 2001).

Cross-validation was performed to evaluate all QSPRs. The cross validation corresponds to a "leave-one-out" procedure, in which a chemical in the series is removed from the analysis and coefficients are recalculated (Wold, 1991). The newly obtained C_i values are then used to estimate the parameter value of the

"left-out" chemical. This step is then repeated for each chemical in the series. Performance of the model is assessed on the basis of the difference between the predicted value of the parameter when a chemical is "left-out" and the experimental value (the deleted residual) for that chemical. The sum of the squares of these residuals (i.e., the predictive residual sum of squares (PRESS) statistic) over the sum of squares of the response values (SSY) is used as a general indication of the validity of a model (Wold, 1991). To be a reasonable QSPR model, the ratio must be <0.4. A value of PRESS/SSY of <0.1 indicates an excellent model (Wold, 1991).

QSPR-PBPK Modeling

The PBPK model used in the present study consisted of four tissue compartments, namely liver, fat, slowly perfused tissues, and rapidly perfused tissues, interconnected by systemic circulation through the pulmonary exchange compartment (Ramsey and Andersen, 1984). The algebraic and differential equations constituting the model were identical to those of Ramsey and Andersen (1984), with the exception of hepatic metabolism which was calculated as hepatic clearance (CL_h) times arterial blood concentration (C_a) (Poulin and Krishnan, 1999). The QSPR-PBPK model was written as a program in Advanced Continuous Simulation Language (ACSL®, Aegis Technologies, Huntsville, AL). For conducting simulations using ACSL®, two files are required: (i) a continuous simulation language code file (.CSL), which contains the program (i.e., constants, QSPRs for model parameters, differential equations, and integration algorithms) and (ii) a command (.CMD) file, which contains simulation conditions (i.e., exposure frequency and duration) and other chemical-specific input data (i.e.,

fragments present in the molecule being simulated). Since QSPRs for P_b, P_l, P_m, P_f and CL_h were included in the .CSL file, only the frequency of occurrence of the fragments in a molecule needed to be entered along with exposure concentration and duration as input in the .CMD file, to obtain pharmacokinetic simulations. A schematic representation of the modeling methodology used in the present study is depicted in Figure 1. The QSPR-PBPK model codes written in ACSL® can be obtained by contacting the authors.

The QSPR-PBPK modeling framework was initially used to predict the inhalation pharmacokinetics of four VOCs present in the calibration dataset (toluene, dichloromethane, trichloroethylene, and 1,1,1-trichloroethane). For each of these chemicals, the exposure concentration and duration were provided as input along with the nature and number of fragments (Table 2) to the QSPR-PBPK model to obtain simulations of their inhalation pharmacokinetics. The simulations were compared with previously published experimental data obtained in rats exposed to these chemicals individually. Additionally, the QSPR-PBPK model was used to predict the inhalation pharmacokinetics of VOCs which were not part of the calibration set but which could be described using the molecular fragments of chemicals investigated in the present study (Table 1). This set consisted of 1,2,4-trimethylbenzene, ethyl benzene, 1,3-dichloropropene and 2,2-dichloro-1,1,1-trifluoroethane. The pharmacokinetic simulations for these VOCs were obtained solely from knowledge of their molecular structure and exposure conditions (Table 3). The QSPR-PBPK model simulations were then compared with experimental data on the inhalation pharmacokinetics of these chemicals in the rat obtained from the literature.

3.3 Results

QSPRs for PBPK model parameters

The contributions of each of the 11 molecular fragments to rat P_b, P_l, P_m, and P_f, as obtained from analysis of data for 46 VOCs, are presented in Table 4. This table also presents the results of such analysis for rat CL_h based on data for 25 VOCs (last two colums of Table 4). For all five PBPK model parameters, the overall regression was highly positive and significant (R^2 for P_b, P_l, P_m, P_f, and CL_h were 0.958, 0.965, 0.941, 0.993, and 0.893, respectively). Furthermore, the PRESS/SSY statistic of the QSPR model for P_b, P_l, P_m, P_f and CL_h was 0.06, 0.05, 0.09, 0.01, and 0.20 (0.10 when excluding 1,1,1-trichloroethane and tetrachloroethylene), respectively. The PRESS/SSY statistic values for QSPRs of partition coefficients are within the suggested benchmark value (0.1) for excellent models. It should be noted, however, that for each of the five PBPK parameters at least three C_i are not statistically significant (*t* statistic < 1.8).

The correlation between experimental and QSPR-estimated values of partition coefficients and hepatic clearance is presented in Figure 2. The mean (±SD) experimental/estimated ratio was 1.12±0.61 for P_b, 1.08±0.41 for P_l, 1.05±0.46 for P_m, 1.06±0.42 for P_f, and 0.99±0.26 for CL_h (excluding CL_h for 1,1,1-trichloroethane and tetrachloroethylene). Overall, of the 184 experimental partition coefficients, 170 were within a factor of two of the QSPR-estimated values, whereas six estimates were within a factor of 2-3 of the experimental values. For the remaining chemicals, the experimental to predicted ratio was greater than three (P_b: cyclohexane (3.8), 2,2,4-trimethylpentane (0.24), and vinyl chloride (3.09); P_f: difluoromethane (3.68); P_l: difluoromethane (0.01) and vinyl chloride (3.14); P_m:

methylchloride (30), cyclohexane (61)). Regarding hepatic clearance, the QSPR-estimates were within a factor of two for all 25 chemicals except for two very poorly metabolized VOCs. For these two VOCs, the predicted vs. experimental CL_h values were 0.44 vs. 0.004 (tetrachloroethylene) and 0.56 vs. 0.07 (1,1,1-trichloroethane). Overall, these results suggest that blood:air partition coefficients, tissue:air partition coefficients, and hepatic clearance of a majority of structurally-unrelated VOCs investigated in the present study can be adequately described with a group contribution model.

QSPR-PBPK modeling

The QSPRs developed in this study were incorporated within PBPK models such that predictions of pharmacokinetics of VOCs could be obtained solely with knowledge of molecular structure and exposure conditions. Figure 3 presents the QSPR-PBPK model simulations of the inhalation pharmacokinetics of four VOCs from the calibration set, namely, toluene, dichloromethane, trichloroethylene and 1,1,1-trichloroethane, and compares them with the experimental data collected in rats. In order to obtain these simulations, only the number of the fragments along with the exposure concentration and duration (Table 2) were provided as input to the QSPR-PBPK model.

The integrated QSPR-PBPK model was also used to predict the inhalation pharmacokinetics of VOCs that were not part of the calibration set. Again, only by providing molecular fragment information along with exposure characteristics (Table 3), predictions of the pharmacokinetics of 1,2,4-trimethylbenzene, ethyl benzene, 1,3-dichloropropene and 2,2-dichloro-1,1,1-trifluoroethane were obtained using the QSPR-PBPK model. Figure 4 compares the model predictions with the inhalation pharmacokinetic data obtained from the literature for these chemicals.

The QSPR-PBPK model, in general, predicts the blood kinetics of inhaled VOCs with only knowledge of their molecular structure, exposure concentration and exposure duration. The level of correspondence between the simulations and experimental data was even greater when the CL_h values were derived from QSPRs that excluded 1,1,1-trichloroethane and tetrachloroethylene (not shown).

3.4 Discussion

QSAR models have been developed for various chemical properties (e.g., vapor pressure, n-octanol:water partition coefficient) and toxicological responses (e.g., mutagenicity, carcinogenicity, cytotoxicity) (Akers *et al.*, 1999; Barratt, 2000; Benigni and Richard, 1998; Blaha *et al.*, 1998; Cronin and Dearden, 1995; Enslein and Borgstedt, 1989; Franke *et al.*, 2001; Greene, 2002; Hansch and Leo, 1979; Hansch and Leo, 1995; Hine and Mookerjee, 1975; Mumtaz *et al.*, 1995; Rosenkranz and Klopman, 1993; Sabljic, 2001; Steward *et al.*, 1973). The feasibility of developing such models for simulating pharmacokinetics of untested chemicals has not been undertaken previously. The present study was based on the premise that, if QSPRs for PBPK model parameters are feasible, then it should be possible to predict inhalation pharmacokinetics of VOCs based on molecular structure information. There are published linear free-energy methodologies that permit the estimation of one or more of the PBPK model parameters of organic chemicals using information on molecular connectivity, Log $P_{o:w}$ and Henry's law constants, or alternatively using 3-D QSAR (Abraham *et al.*, 1985; Abraham and Weathersby, 1994; Csanady *et al.*, 1995; Csanady and Laib, 1990; Gargas, Seybold, and Andersen, 1988; Parham *et al.*, 1997; Parham and Portier, 1998; Waller *et al.*, 1996). None of the published approaches facilitate the estimation of all PBPK model parameters for chemicals belonging to more than one family. The present study for the first time developed QSPRs for estimating all chemical-specific parameters (partition coefficients and hepatic clearance) and validated them by successful incorporation within PBPK models to simulate the inhalation pharmacokinetics of several VOCs.

A previous effort attempted to relate molecular structure information and PBPK model parameters using the Free-Wilson approach (Fouchécourt and Krishnan, 2000). The resulting Free-Wilson algorithms were limited by their applicability to one family of substances (i.e., chloroethanes). An alternative is to use the group contribution approach (Gao, Govind, and Tabak, 1992; Martin and Young, 2001) which accounts for the contribution of each atom in the molecular structure. This approach is different from the Free-Wilson approach in that it does not require a common structure among the chemicals investigated. The common "basic" structure is easily identifiable while studying chemicals belonging to the same family (e.g., the two carbon backbone in the chloroethanes (Fouchécourt and Krishnan, 2000)) but becomes a hurdle while considering multiple groups of chemicals such as chloromethanes, ethanes, ethylenes, and aromatic hydrocarbons. In such cases, the establishment of a chemical class-specific QSPR is certainly possible, but the tradeoff becomes one of statistical power, since the number of chemicals in each family having experimental PBPK parameter values is very limited, at least at the present time. For example, the chloroethylene family contains only five chemicals for which PBPK model parameters have been determined experimentally. The alternative is to use Gao's approach, as has been done in the present study, since it does not require the existence of a common basic structure and facilitates the identification of independent variables (i.e., number of fragments or groups in a molecule) based on visual inspection of a compound's structural formula (Martin and Young, 2001).

The partition coefficients and metabolic parameters required for developing chemical-specific PBPK models are most often obtained *in vivo* or *in vitro* (reviewed by Krishnan and Andersen (Krishnan and Andersen, 2001)). Algorithms

and QSPRs for estimating blood:air, tissue:air and tissue:blood partition coefficients on the basis of $P_{o:w}$ and $P_{w:a}$ have become available (DeJongh *et al.*, 1997; Meulenberg and Vijverberg, 2000; Poulin and Krishnan, 1995; Poulin and Krishnan, 1996). According to these methods, however, the estimation of $P_{w:a}$ and $P_{o:w}$ from molecular structure information requires the development and use of two different sets of molecular fragments for each chemical. Alternatively, direct relationships between the number or nature of molecular fragments and the values of chemical-specific parameters (i.e., partition coefficients and hepatic clearance) used in PBPK modeling may be established. This has been attempted recently with chloroethanes (Fouchécourt and Krishnan, 2000) using a Free-Wilson approach, suggesting the potential applicability of additive approaches based on molecular fragment contributions for QSAR modeling of PBPK parameters. The present study has demonstrated the feasibility of the development of QSPRs for VOCs belonging to multiple families and its integration within PBPK modeling framework to simulate the pharmacokinetics of new chemicals.

The ability of the QSPR-PBPK model to simulate adequately the kinetics of new chemicals is not necessarily a result of accurate prediction of all individual chemical-specific parameters, but it could very well indicate that the net impact of critical determinants of blood concentration, as computed from QSPR-derived parameters is similar to that of experimentally-derived PBPK model parameters. However, for most of the chemicals in the calibration set, the experimental and estimated values of individual PBPK model parameters compared favorably. The most sensitive chemical-specific parameter will vary according to the chemical being simulated, the exposure scenario, and the endpoint examined. QSPRs should strive towards accurately predicting the most sensitive parameter(s) for the

chemical, endpoint and exposure scenario of interest. Following the conventional sensitivity analysis for the chemical-specific parameters, (e.g., Tardif *et al.*, 2002) more detailed analysis can be performed to evaluate the sensitivity of fragment-specific contributions with respect to the endpoint of interest. Molecular fragment(s), critical to the kinetic profiles, can subsequently be identified.

Results of the present study show that it is possible to describe PBPK model parameters (e.g., blood:air partition coefficients, tissue:air partition coefficients and hepatic clearance) as the cumulative addition of the contributions of various molecular fragments for a series of chemicals. Not all fragment contribution values in Table 4 are statistically significant (*t* value > 1.8). This may be a consequence of either the fact that some of these fragments do not make a statistically-significant contribution to the parameter value, or that there is not enough data to facilitate the determination of a contribution value of statistical significance, for some fragments. It should also be noted that no judgment of data quality was done, even though most data came from a single laboratory and were generated using the same experimental protocol. Generally, a high degree of correlation existed between the experimental and QSPR-estimated values of PBPK model parameters. The correlation was strongest for P_f, which is not surprising given the fact, a similar parameter, namely $P_{o:w}$, has been shown to vary predictably among chemicals based on the number and nature of fragments (Hansch and Leo, 1979).

Of all the PBPK model parameters, the metabolic constants represented the most difficult in terms of QSPR development. In the present study, initial regression studies were conducted using the maximal velocity for metabolism (V_{max}) and Michaelis affinity constant (K_m) of the VOCs. Although both V_{max} and K_m could be modeled adequately (R^2 > 0.80), the QSPR for K_m failed the cross

validation process (PRESS/SSY = 0.7). Further efforts were undertaken to develop QSPR models of intrinsic clearance (CL_{int}), which represents the ratio of V_{max}/K_m. Even though the QSPR for CL_{int} represented a significant improvement over the V_{max} and K_m models, the adequacy of this model fit was influenced by the CL_{int} value of isoprene that was well outside the range of the experimental values of the other chemicals. Since some VOCs are also metabolized via a first order process such as GSH conjugation, modeling of total hepatic clearance, CL_h, was undertaken. The rate of metabolism in liver can be computed either using CL_h or using V_{max}/K_m for first order conditions. Even though both approaches give identical results (Poulin and Krishnan, 1999), the use of CL_h permits a more direct evaluation of the impact of hepatic blood flow limitation on hepatic metabolic processes considered together (Johanson and Naslund, 1988; Kedderis, 1997). The present study successfully developed a QSPR for CL_h and integrated it within PBPK models to provide predictions of the kinetics of VOCs.

Of the VOCs investigated in the present study, carbon tetrachloride, 1,1,1-trichloroethane, hexachloroethane, and tetrachloroethylene exhibited poor hepatic extraction ($E<0.5$). Of these, only 1,1,1-trichloroethane and tetrachloroethylene were not modeled adequately by the QSPR. Gargas *et al.* (1988) reported similar observations during the modeling of V_{max} of halogenated alkanes and alkylenes. The QSPR developed in this study did actually predict 1,1,1-trichloroethane and tetrachloroethylene as being poorly metabolized ($E = 0.42$ and 0.30, respectively), but not to the extent the experimental data suggest ($CL_h = 0.02$ and 0.001, respectively). Note that these are the only two chemicals in the dataset exclusively metabolized by a first order process at experimental exposure concentrations ($K_f = 5$ and 0.3, respectively). However, this discrepancy between experimental and

predicted CL_h did not necessarily have a significant impact on the kinetic profile for these chemicals, because of the lower sensitivity of CL_h. This is clearly illustrated in the case of 1,1,1-trichloroethane, for which the QSPR-PBPK model simulations describe the experimental data fairly well (Figure 3C).

The CL_h, as derived and modeled in the present study is only appropriate for conducting simulations of the pharmacokinetics of chemicals at relatively low exposure concentrations. At high exposure concentrations when saturable mechanisms come into play, knowledge of V_{max} and K_m is essential to be able to adequately simulate the metabolism rate and pharmacokinetics of chemicals in biota. Therefore, more complex QSPR approaches may have to be explored for modeling V_{max} and K_m. The use of the fragment-based additive approach, developed in the present study, in our opinion, should be limited to low molecular weight VOCs (i) metabolized by P-450 2E1 and (ii) that possess fragments (number and nature) that are similar to those investigated in the present study. Since the isoenzyme involved in the metabolism of other chemicals may be different (e.g., PAHs, PCBs) (Chan and Hellebone, 1995; Lewis, 2000), additional mechanistic data at the qualitative and quantitative levels should be considered (e.g., presence or absence of the role of an isoenzyme, its concentration in tissues) such that the resulting QSPRs can be of use in predicting metabolism of untested chemicals.

For the time being, extrapolations using QSPRs to predict PBPK model parameters for other chemicals can be made as long as they contain fragments corresponding to the number and nature, investigated in this study. Extrapolation to predict the kinetics of chemicals with fragments other than those investigated in this study, or larger number of certain fragments (e.g., 4 F, 5 CH_3) than those

investigated in this study are not appropriate. With increasing chain length and molecular weight, it is likely that more complex approaches may be necessary to account for non-linearity and qualitative changes in mechanisms (presence vs absence of protein binding, saturation of liposolubility characteristics, differences in isoenzymes involved in metabolism).

In conclusion, the present study has demonstrated that it is possible to develop QSPRs for PBPK model parameters and use them for conducting *a priori* prediction of the inhalation pharmacokinetics of VOCs in rats. The QSPRs developed in this study should be potentially useful for providing a "first-cut" evaluation of the pharmacokinetics of chemicals, prior to experimentation, as long as the number and nature of the fragments do not exceed the ones in the calibration dataset used in this study.

Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council and the Canadian Network of Toxicology Centres. M.B. is recipient of a scholarship from the Natural Sciences and Engineering Research Council and Fonds de Recherche sur la Nature et les Technologies. K.K. is recipient of a research scholarship from Fonds de la Recherche en Santé du Québec (1992-2004).

3.5 References

Abraham, M. H., Kamlet, M. J., Taft, R. W., Doherty, R. M., and Weathersby, P. K. (1985). Solubility properties in polymers and biological media. 2. The correlation and prediction of the solubilities of nonelectrolytes in biological tissues and fluids. *J. Med. Chem.* **28**, 865-870.

Abraham, M. H., and Weathersby, P. K. (1994). Hydrogen bonding. 30. Solubility of gases and vapors in biological liquids and tissues. *J. Pharm. Sci.* **83**, 1450-1456.

Akers, K. S., Sinks, G. D., and Schultz, T. W. (1999). Structure-toxicity relationships for selected halogenated aliphatic chemicals. *Environ. Toxicol. Pharmacol.* **7**, 33-39.

Ali, N., and Tardif, R. (1999). Toxicokinetic modeling of the combined exposure to toluene and n-hexane in rats and humans. *J. Occup. Health* **41**, 95-103.

Andersen, M. E., Clewell, H. J. I., Gargas, M. L., Smith, F. A., and Reitz, R. H. (1987). Physiologically-based pharmacokinetics and risk assessment process for methylene chloride. *Toxicol. Appl. Pharmacol.* **87**, 185-205.

Arms, A. D. and Travis, C. C. (1988). *Reference Physiological Parameters in Pharmacokinetic Modeling. Office of Health and Environmental Assessment.* EPA/600/6-88/004, US EPA, Washington, DC.

Barratt, M. D. (2000). Prediction of toxicity from chemical structure. *Cell Biol. Toxicol.* **16**, 1-13.

Benigni, R., and Richard, A. M. (1998). Quantitative structure-based modeling applied to characterization and prediction of chemical toxicity. *Methods Enzymol.* **14**, 284-276.

Blaha, L., Damborsky, J., and Nemec, M. (1998). QSAR for acute toxicity of saturated and unsaturated halogeneted aliphatic compounds. *Chemosphere* **36**, 1345-1365.

Bogaards, J. J., Freidig, A. P., and van Bladeren, P. J. (2001) Prediction of isoprene diepoxide levels in vivo in mouse, rat and man using enzyme kinetic data in vitro and physiologically-based pharmacokinetic modelling. *Chem. Biol. Interact.* **138**, 247-265.

Brown, R. P., Delp, M. D., Lindstedt, S. L., Rhomberg, L. R., and Beliles, R. P. (1997). Physiological parameter values for physiologically based pharmacokinetic models. *Toxicol. Ind. Health* **13**, 407-484.

Chan, Z., and Hellebone, B. R. (1995). A QSAR for steroidal compound interaction with cytochrome P4501A1. *Environ. Toxicol. Chem.* **14**, 597-603.

Clewell, H. J., III, Andersen, M. E., and Barton, H. A. (2002). A consistent approach for the application of pharmacokinetic modeling in cancer and noncancer risk assessment. *Environ. Health Perspect.* **110**, 85-93.

Cronin, M. T. D., and Dearden, J. C. (1995). QSAR in toxicology. 3. Prediction of chronic toxicities. *Quant. Struct. -Act. Relat.* **14**, 329-334.

Csanady, G. A., and Laib, R. J. (1990). Use of linear free energy relationships in toxicology: prediction of partition coefficients of volatile lipophilic compounds. *Arch. Toxicol.* **64**, 594-596.

Csanady, G. A., Laib, R. J., and Filser, J. G. (1995). Metabolic transformation of halogenenated and other alkenes - a theoretical approach. Estimation of metabolic reactivities for in vivo conditions. *Toxicol. Lett.* **75**, 217-223.

Dallas, C. E., Gallo, J. M., Ramanthan, R., Muralidhara, S., and Bruckner, J. V. (1991). Physiological pharmacokinetic modeling of inhaled trichloroethylene in rats. *Toxicol. Appl. Pharmacol.* **110**, 303-314.

Dallas, C. E., Ramanthan, R., Muralidhara, S., Gallo, G. M., and Bruckner, J. V. (1989). The uptake and elimination of 1,1,1-trichloroethane during the following inhalation exposures in rats. *Toxicol. Appl. Pharmacol.* **98**, 385-397.

DeJongh, J., Verhaar, H. J., and Hermens, J. L. (1997). A quantitative property-property relationship (QPPR) approach to estimate in vitro tissue-blood partition coefficients of organic chemicals in rats and humans. *Arch. Toxicol.* **72**, 17-25.

Enslein, K., and Borgstedt, H. H. (1989). A QSAR model for the estimation of carcinogenicity: example application to an azo-dye. *Toxicol. Lett.* **49**, 107-121.

Fouchécourt, M.-O., and Krishnan, K. (2000). A QSAR-type PBPK model for inhaled chloroethanes. *Toxicol. Sci.* **54**, 88.

Franke, R., Gruska, A., Giuliani, A., and Benigni, R. (2001). Prediction of rodent carcinogenicity of aromatic amines: a quantitative structure-activity relationship model. *Carcinogenesis* **22**, 1561-1571.

Gao, C., Govind, R., and Tabak, H. H. (1992). Application of the group contribution method for predicting the toxicity of organic chemicals. *Environ. Toxicol. Chem.* **11**, 631-636.

Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H., and Andersen, M. E. (1989). Partition coefficients of low-molecular-weight volatile chemicals in various liquids and tissues. *Toxicol. Appl. Pharmacol.* **98**, 87-99.

Gargas, M. L., Seybold, P. G., and Andersen, M. E. (1988). Modeling the tissue solubilities and metabolic rate constant (Vmax) of halogenated methanes, ethanes, and ethylenes. *Toxicol. Lett.* **43**, 235-256.

Greene, N. (2002). Computer systems for the prediction of toxicity: an update. *Adv. Drug Deliv. Rev.* **54**, 417-431.

Haddad, S., Charest-Tardif, G., Tardif, R., and Krishnan, K. (2000). Validation of a physiological modeling framework for simulating the toxicokinetics of chemicals in mixtures. *Toxicol. Appl. Pharmacol.* **167**, 199-209.

Hansch, C., and Leo, A. (1979). The fragment method of calculating partition coefficients. In *Substituent constants for correlation analysis in chemistry and biology* (Hansch, C., and Leo, A., Eds.), pp. 18-43 Wiley, New York, NY.

Hansch, C., and Leo, A. (1995). *Exploring QSAR.* American Chemical Society, Washington, D.C.

Hansen, H.K., Rasmussen, P., Fredenslund, A., Schiller, M. and Gmehling, J. (1991). Vapor liquid equilibria by UNIFAC group contibution, 5. Revision and extension. Ind. Eng. Chem. Res. **30**, 2355-2358.

Hine, J., and Mookerjee, P. K. (1975). The intrinsic hydrophobic character of organic compounds: correlations in terms of structural contributions. *J. Org. Chem.* **40**, 511-522.

Johanson, G., and Naslund, P. H. (1988). Spreadsheet programming: A new approach in physiologically based modeling of solvent toxicokinetics. *Toxicol. Lett.* **41**, 115-127.

Kedderis, G. L. (1997). Extrapolation of in vitro enzyme induction data to humans in vivo. *Chem. Biol. Interact.* **107**, 109-121.

Krishnan, K., and Andersen, M. E. (2001). Physiologically based pharmacokinetic modeling in toxicology. In *Principles and methods of toxicology* (Hayes, A. W., Ed.), pp. 193-241 Taylor & Francis, Philadelphia, PA.

Lewis, D. F. (2000). On the recognition of mammalian microsomal cytochrome P450 substrates and their characteristics: towards the prediction of human p450 substrate specificity and metabolism. *Biochem. Pharmacol.* **60**, 293-306.

Lilly, P. D., Andersen, M. E., Ross, T. M., and Pegram, R. A. (1998) A physiologically based pharmacokinetic description of the oral uptake, tissue dosimetry, and rates of metabolism of bromodichloromethane in the male rat. *Toxicol. Appl. Pharmacol.* **150**, 205-217.

Martin, T. M., and Young, D. M. (2001). Prediction of the acute toxicity (96-h LC_{50}) or organic compounds to the fathead minnow (*Pimephales promelas*) using a group contribution method. *Chem. Res. Toxicol.* **14**, 1378-1385.

Meulenberg, C. J., and Vijverberg, H. P. (2000). Empirical relations predicting human and rat tissue:air partition coefficients of volatile organic compounds. *Toxicol. Appl. Pharmacol.* **165**, 206-216.

Mumtaz, M. M., Knauf, L. A., Reisman, D. J., Peirano, W. B., DeRosa, C. T., Gombar, V. K., Enslein, K., Carter, J. R., Blake, B. W., Huque, K. I., and Ramanujam, V. M. S. (1995). Assessment of effect levels of chemicals from quantitative structure-activity relationship (QSAR) models. I. Chronic lowest-observed-adverse-effect level (LOAEL). *Toxicol. Lett.* **79**, 131-143.

Ochsner, A.B., and Sokoloski, T.D. (1985). Prediction of solubility in nonideal multicomponent systems using the UNIFAC group contribution model. *J. Pharm. Sci.* **74**, 634-637.

Parham, F. M., Kohn, M. C., Matthews, H. B., DeRosa, C., and Portier, C. J. (1997). Using structural information to create physiologically based pharmacokinetic models for all polychlorinated biphenyls. *Toxicol. Appl. Pharmacol.* **144**, 340-347.

Parham, F. M., and Portier, C. J. (1998). Using structural information to create physiologically based pharmacokinetic models for all polychlorinated biphenyls. II. Rates of metabolism. *Toxicol. Appl. Pharmacol.* **151**, 110-116.

Payne, M. P., and Kenny, L. C. (2002). Comparison of models for the estimation of biological partition coefficients. *J. Toxicol. Environ. Health A* **65**, 897-931.

Poulin, P., and Krishnan, K. (1995). An algorithm for predicting tissue: blood partition coefficients of organic chemicals from n-octanol: water partition coefficient data. *J. Toxicol. Environ. Health* **46**, 117-129.

Poulin, P., and Krishnan, K. (1996). Molecular structure-based prediction of the partition coefficients of organic chemicals for physiological pharmacokinetic models. *Toxicol. Methods* **6**, 117-137.

Poulin, P., and Krishnan, K. (1999). Molecular structure-based prediction of the toxicokinetics of inhaled vapors in humans. *Int. J. Toxicol.* **18**, 7-18.

Ramsey, J. C., and Andersen, M. E. (1984). A physiologically-based description of the inhalation pharmacokinetics of styrene in rats and humans. *Toxicol. Appl. Pharmacol.* **73**, 159-175.

Renwick, A. G. (2001). Toxicokinetics: pharmacokinetics in toxicology. In: *Principles and methods of toxicology* (Hayes, A. W., Ed.), pp. 137-191 Taylor & Francis, Philadelphia, PA.

Rosenkranz, H. S., and Klopman, G. (1993). Structural relationships between mutagenicity, maximum tolerated dose, and carcinogenicity in rodents. *Environ. Mol. Mutagen.* **21**, 193-206.

Sabljic, A. (2001). QSAR models for estimating properties of persistant organic pollutants required in evaluation of their environmental fate and risk. *Chemosphere* **43**, 363-375.

Sato, A., and Nakajima, T. (1979) Partition coefficients of some aromatic hydrocarbons and ketones in water, blood and oil. *Br. J. Ind. Med.* **36**, 231-234

Steward, A., Allott, P. R., Cowles, A. L., and Mapleson, W. W. (1973). Solubility coefficients for inhaled anaesthetics for water, oil and biological media. *Br. J. Anaesth.* **45**, 282-293.

Stott, W. T., and Kastl, P. E. (1986). Inhalation pharmacokinetics of techical grade 1,3-dichloropropene in rats. *Toxicol. Appl. Pharmacol.* **85**, 332-341.

Swiercz, R., Rydzynski, K., Wasowicz, W., Majcherek, W., and Wesolowski, W. (2002). Toxicokinetics and metabolism of pseudocumene (1,2,4-trimethylbenzene) after inhalation exposure in rats. *Int. J. Occup. Med. Environ. Health* **15**, 37-42.

Tardif, R., Droz, P.-O., Charest-Tardif, G., Pierrehumbert, G., and Truchon, G. (2002). Impact of human variability on the biological monitoring of exposure to toluene: I. Physiologically based toxicokinetic modeling. *Toxicol. Lett.* **134**, 155-163.

Vinegar, A., Williams, R. J., Fisher, J. W., and McDougal, J. N. (1994). Dose-dependant metabolism of 2,2-dichloro-1,1,1-trifluoroethane: a physiologically based pharmacokinetic model in the male Fisher 344 rat. *Toxicol. Appl. Pharmacol.* **129**, 103-113.

Waller, C. L., Evans, M. V., and McKinney, J. D. (1996). Modeling the cytochrome P450-mediated metabolism of chlorinated volatile organic compounds. *Drug Metab. Dispos.* **24**, 203-210.

Wold, S. (1991). Validation of QSAR's. *Quant. Struct. -Act. Relat.* **10**, 191-193.

Figure legends

FIGURE 1: Schematic representation of the QSPR-PBPK modeling framework used in the present study.

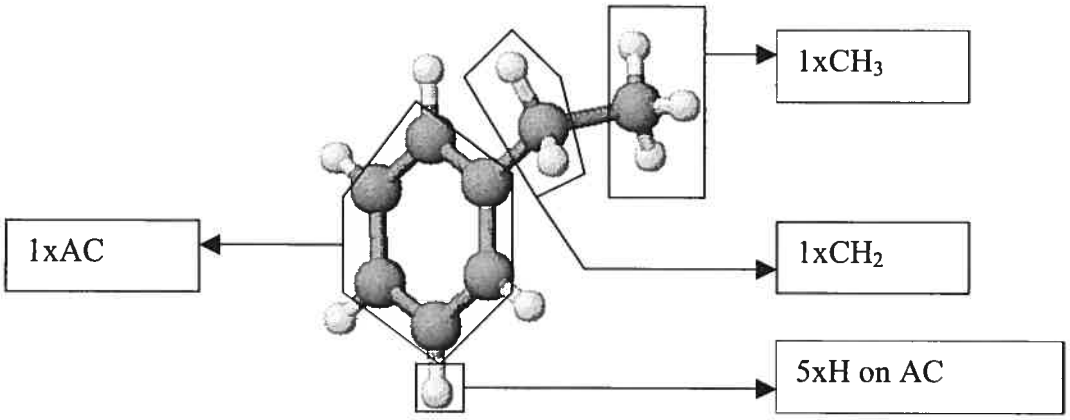
FIGURE 2: Comparison of experimental values (Exp.) with the QSPR-derived estimates (Est.) of (A) log blood:air partition coefficients (R^2 = 0.967, y = 0.968x), (B) log liver:air partition coefficients (R^2 = 0.973, y = 0.973x), (C) log fat:air partition coefficients (R^2 = 0.995, y = 0.995x), (D) log muscle:air partition coefficients (R^2 = 0.954, y = 0.955x) of 46 volatile organic chemicals and (E) hepatic clearance (R^2 = 0.936, y = 0.936x) for 25 chemcials. Experimental data from Ramsey and Andersen (1984), Gargas *et al.* ((Gargas *et al.*, 1988; Gargas *et al.*, 1989)), Lily *et al.* (1998), Ali and Tardif (1999), Haddad *et al.* (2000), and Bogaards *et al.* (2001).

FIGURE 3: Comparison of the QSPR-PBPK model simulations (solid lines) with the experimental data (symbols) on venous blood concentrations in rats following inhalation exposure to (A) toluene (50 ppm, 4hr), (B) dichloromethane (100 ppm, 4 hr), (C) 1,1,1-trichloroethane (50 ppm, 2 hr), and (D) trichloroethylene (50 ppm, 2 hr). Data were obtained from the literature (Dallas *et al.*, 1989; 1991; (Haddad *et al.*, 2000)).

FIGURE 4: Comparison of the QSPR-PBPK model simulations (solid lines) with the experimental data (symbols) on venous blood concentrations in rats following inhalation exposure to (A) 1,2,4-trimethylbenzene (100 ppm, 6 hr), (B) 2,2-dichloro-1,1,1-trifluoroethane (1000 ppm, 4 hr), (C) 1,3-dichloropropene (300 ppm, 3 hr), and (D) ethylbenzene (50 ppm, 4 hr). Data were obtained from the literature (Stott

and Kastl, 1986; Vinegar *et al.*, 1994; (Haddad, Charest-Tardif, Tardif, and Krishnan, 2000); (Swiercz *et al.*, 2002).

Figure 1



USER INPUT (.CMD file)
- Exposure conc. = 50 ppm
- Exposure duration = 4 hr
- Frequency of occurrence of substituents
  - *For Ethylbenzene*
    - $f_{CH3} = 1$
    - $f_{CH2} = 1$
    - $f_{AC} = 1$
    - $f_{H\_AC} = 5$
    - All other $f_i = 0$

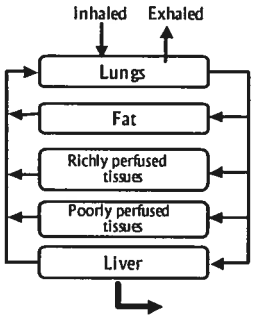
PROGRAM (.CSL file)
- Model constants and physiological parameters
- QSPRs for model parameters :

$$P_{PBPK} = \sum f_i \times C_i$$

- Contributions of structural features
- Simulation algorithms

QSPR-BASED PBPK MODEL

- Log P_t = $f_{CH3} \times C_{CH3}$ + $f_{CH2} \times C_{CH2}$ + $f_{AC} \times C_{AC}$ + $f_{H(AC)} \times C_{H\_AC}$
- CL_h = 1 x 0.398 + 1 x 5.76E-2 + 1 x –6.29E-2 + 5 x 7.25E-2= 1.03
- Log P_b = 1 x 7.18E-2 + 1 x 0.109 + 1 x 2.85 + 5 x –0.297 = 1.55
- Log P_l = 1 x 1.55E-2 + 1 x 0.234 + 1 x 3.76 + 5 x –0.408 = 1.97
- Log P_f = 1 x 0.366 + 1 x 0.435 + 1 x 2.92 + 5 x -5.58E-2 = 3.44
- Log P_s = 1 x –1.97E-2+ 1 x 0.122 + 1 x 3.65 + 5 x -0.466 = 1.52



OUTPUT

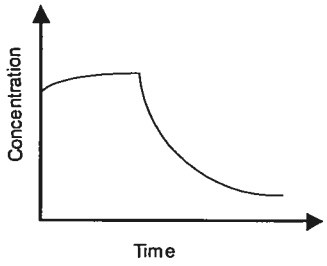


Figure 2

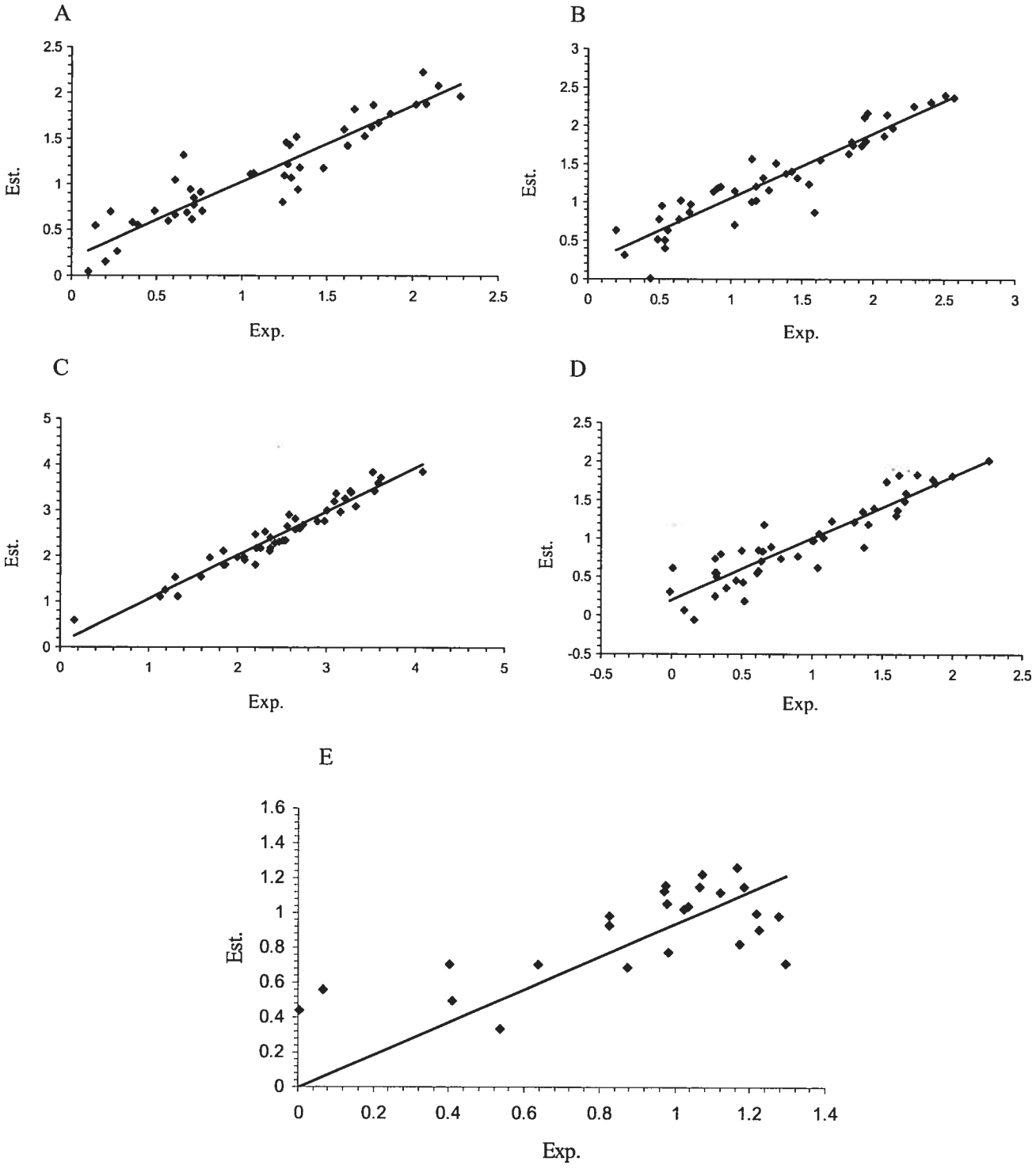


Figure 3

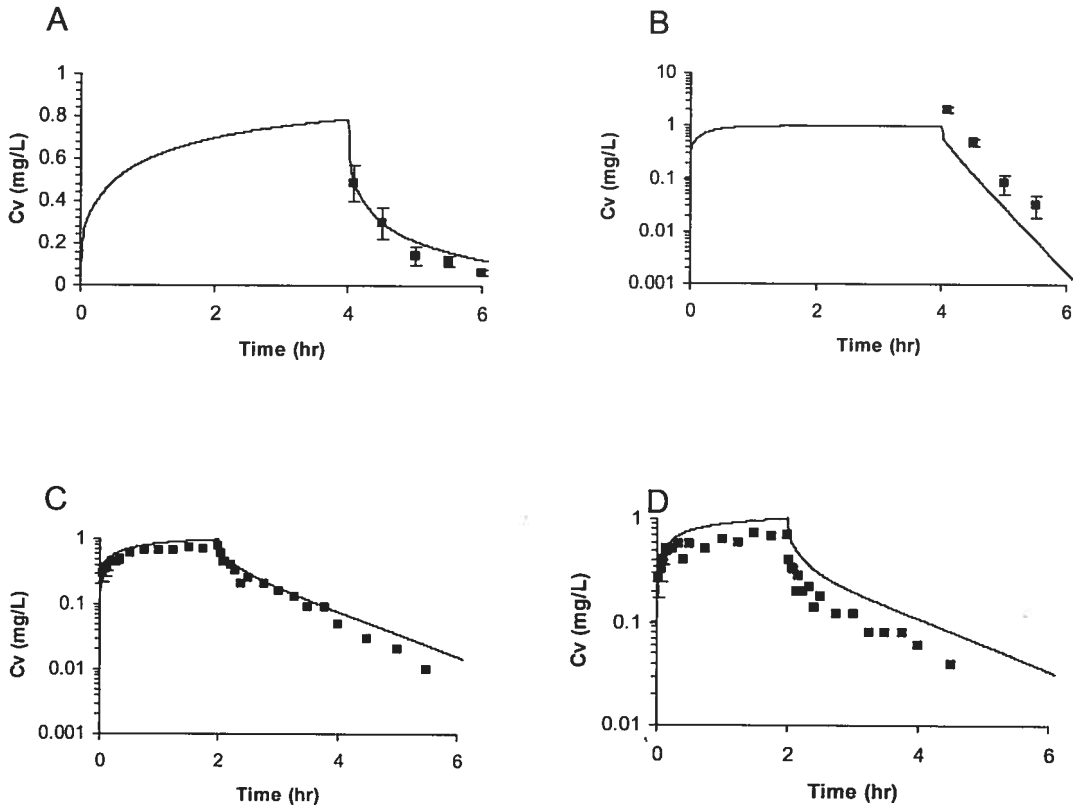


Figure 4

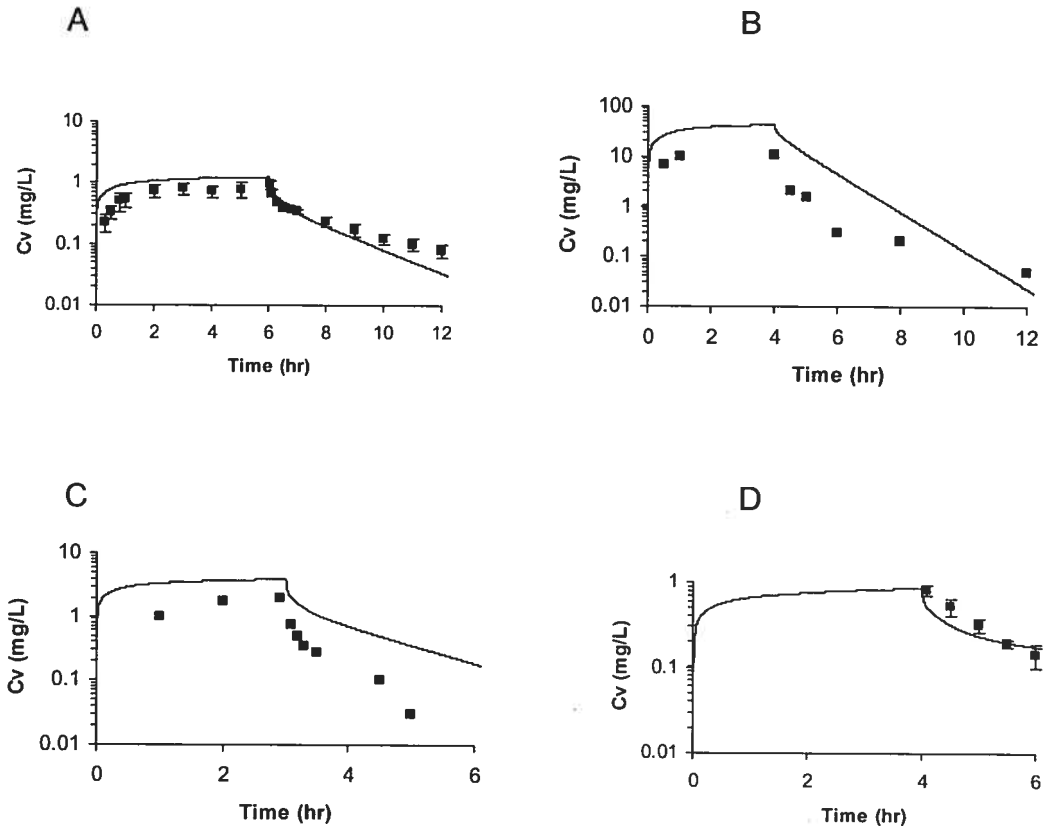


Table 1: Frequency of occurrence of molecular fragments in the VOCs investigated in the present study.

Chemicals	CH_3	CH_2	CH	C	C=C	H	Br	Cl	F	AC	H_AC
Chloromethane	1	0	0	0	0	0	0	1	0	0	0
Dichloromethane	0	1	0	0	0	0	0	2	0	0	0
Chloroform	0	0	1	0	0	0	0	3	0	0	0
Carbon tetrachloride	0	0	0	1	0	0	0	4	0	0	0
Difluoromethane	0	1	0	0	0	0	0	0	2	0	0
Fluorochloromethane	0	1	0	0	0	0	0	1	1	0	0
Bromochloromethane	0	1	0	0	0	0	1	1	0	0	0
Dibromomethane	0	1	0	0	0	0	2	0	0	0	0
Chlorodibromomethane	0	0	1	0	0	0	2	1	0	0	0
Chloroethane	1	1	0	0	0	0	0	1	0	0	0
1,1-Dichloroethane	1	0	1	0	0	0	0	2	0	0	0
1,2-Dichloroethane	0	2	0	0	0	0	0	2	0	0	0
1,1,1-Trichloroethane	1	0	0	1	0	0	0	3	0	0	0
1,1,2-Trichloroethane	0	1	1	0	0	0	0	3	0	0	0
1,1,1,2-Tetrachloroethane	0	1	0	1	0	0	0	4	0	0	0
1,1,2,2-Tetrachloroethane	0	0	2	0	0	0	0	4	0	0	0
Pentachloroethane	0	0	1	1	0	0	0	5	0	0	0
Hexachloroethane	0	0	0	2	0	0	0	6	0	0	0
1,2-Dibromoethane	0	2	0	0	0	0	2	0	0	0	0
1-Bromo-2-chloroethane	0	2	0	0	0	0	1	1	0	0	0

1,1,1-Trifluoro-2-chloroethane	0	1	0	1	0	0	0	1	3	0	0
1,1,1-Trifluoro-2-bromo-2-chloroethane	0	0	1	1	0	0	1	1	3	0	0
1-Chloropropane	1	2	0	0	0	0	0	1	0	0	0
2-Chloropropane	2	0	1	0	0	0	0	1	0	0	0
1,2-Dichloropropane	1	1	1	0	0	0	0	2	0	0	0
n-Propyl bromide	1	2	0	0	0	0	1	0	0	0	0
isopropylbromide	2	0	1	0	0	0	1	0	0	0	0
n-Hexane	2	4	0	0	0	0	0	0	0	0	0
n-Heptane	2	5	0	0	0	0	0	0	0	0	0
Cyclohexane	0	5	0	0	0	0	0	0	0	0	0
2,3,4-Trimethyl pentane	5	0	3	0	0	0	0	0	0	0	0
2,2,4-Trimethyl pentane	5	1	1	1	0	0	0	0	0	0	0
Vinyl chloride	0	0	0	0	1	3	0	1	0	0	0
1,1-Dichloroethylene	0	0	0	0	1	2	0	2	0	0	0
cis-1,2,-Dichloroethylene	0	0	0	0	1	2	0	2	0	0	0
Trichloroethylene	0	0	0	0	1	1	0	3	0	0	0
Tetrachoroethylene	0	0	0	0	1	0	0	4	0	0	0
Vinyl bromide	0	0	0	0	1	3	1	0	0	0	0
Benzene	0	0	0	0	0	0	0	0	0	1	6
Chlorobenzene	0	0	0	0	0	0	0	1	0	1	5
Toluene	1	0	0	0	0	0	0	0	0	1	5
Styrene	0	0	0	0	1	3	0	0	0	1	5
m-Methylstyrene	1	0	0	0	1	3	0	0	0	1	4

m-Xylene	2	0	0	0	0	0	0	0	0	1	4
Allyl chloride	0	1	0	0	1	3	0	1	0	0	0
Isoprene	1	0	0	0	2	5	0	0	0	0	0
n	30	36	15	9	11	25	11	65	9	6	29

Table 2: Molecular fragments and exposure characteristics of chemicals used for the internal validation of the QSPR-PBPK model.

Chemical	Fragment	Frequency of occurence	Exposure conc. (ppm)	Exposure duration (hr)	Reference
Dichloromethane	CH_2	1	100	4	(Haddad *et al.*, 2000)
	Cl	2			
Toluene	CH_3	1	50	4	(Haddad *et al.*, 2000)
	AC	1			
	H_AC	5			
1,1,1-Trichloroethane	CH_3	1	50	2	(Dallas *et al.*, 1989)
	C	1			
	Cl	3			
Trichloroethylene	C=C	1	50	2	(Dallas *et al.*, 1991)
	H	1			
	Cl	3			

Table 3: Molecular fragments and exposure characteristics of chemicals used for the external validation of the QSPR-PBPK model.

Chemical	Fragment	Frequency of occurence	Exposure conc. (ppm)	Exposure duration (hr)	Reference
1,2,4-Trimethylbenzene	AC	1	100	6	(Swiercz *et al.*, 2002)
	H_AC	3			
	CH_3	3			
1,3-Dichloropropene	CH_2	1	300	3	(Stott and Kastl, 1986)
	C=C	1			
	H	2			
	Cl	2			
Ethyl benzene	CH_3	1	50	4	(Haddad *et al.* 2000)
	CH_2	1			
	AC	1			
	H_AC	5			
2,2-Dichloro-1,1,1-trifluoroethane	CH	1	1000	4	(Vinegar *et al.*, 1994)
	C	1			
	Cl	2			
	F	3			

Table 4: Fragment-specific contributions to PBPK model parameters[a]

Fragments	Log P_b	t[b]	Log P_l	t	Log P_m	t	Log P_f	t	CL_h	t
CH_3	0.0718	1.21	0.0155	0.252	-0.0197	0.319	0.366	7.47	0.388	2.89
CH_2	0.109	2.98	0.234	6.20	0.122	3.23	0.435	14.5	-0.186	0.200
CH	0.0789	0.700	0.359	3.08	0.266	2.27	0.330	3.56	-0.464	2.03
C	-0.606	3.73	0.0318	0.189	-0.105	0.621	-0.285	2.13	-1.44	4.67
C=C	-0.494	1.76	0.257	0.881	-0.707	0.243	0.327	1.41	-1.71	3.75
H	0.236	2.24	-0.0305	-0.0280	0.0813	0.743	0.155	1.79	0.813	4.79
Br	0.834	12.0	0.700	9.68	0.622	8.60	1.17	20.3	0.523	2.98
Cl	0.481	10.7	0.384	8.23	0.322	6.90	0.735	19.8	0.537	6.00
F	0.0203	0.287	-0.113	1.54	-0.911	1.24	0.0752	1.29	-	-
AC	2.85	3.50	3.76	4.46	3.65	4.32	2.92	4.35	0.128	0.094
H_AC	-0.292	1.79	-0.408	2.41	-0.446	2.64	-0.558	0.415	0.611	0.235

[a] P_b, P_l, P_m, P_f, and CL_h refer to blood:air, liver:air, muscle:air, fat:air partition coefficients, and hepatic clearance, respectively.

[b] t-statistic > 1.8 indicates significance at $p < 0.05$.

CHAPITRE QUATRIÈME:

4 Article III

A spreadsheet program for modeling quantitative structure-pharmacokinetic relationships for inhaled volatile organics in humans

Submitted to: SAR and QSAR in Environmental Research

A SPREADSHEET PROGRAM FOR MODELING QUANTITATIVE STRUCTURE-PHARMACOKINETIC RELATIONSHIPS FOR INHALED VOLATILE ORGANICS IN HUMANS

M. Béliveau, R.Tardif, K.Krishnan*

Groupe de Recherche en Toxicologie Humaine (TOXHUM), Université de Montréal, Case Postale 6128, Succ. Centre-Ville, Montréal, PQ, H3C 3J7, Canada.

Address for correspondence: Dr. Kannan Krishnan, Département de santé environnementale et santé au travail, Faculté de médecine, Université de Montréal, 2375 Chemin de la Côte Ste-Catherine, Bureau 4105, Montréal, Québec, Canada, H3T 1A8. Tel: (514) 343-6581, Fax: (514) 343-2200. email: [redacted]

Abstract

The extent and profile of target tissue exposure to toxicants depends upon the pharmacokinetic processes, namely, absorption, distribution, metabolism and excretion. The present study developed a spreadsheet program to simulate the pharmacokinetics of inhaled volatile organic chemicals in humans based on molecular structure information. The approach involved the construction of a human physiologically-based pharmacokinetic (PBPK) model, and the estimation of its parameters based on quantitative structure-property relationships (QSPRs) in an Excel® spreadsheet. The compartments of the PBPK model consisted of liver, adipose tissue, poorly perfused tissues and richly perfused tissues connected by circulating blood. The parameters required were: human physiological parameters such as cardiac output, breathing rate, tissue volumes and tissue blood flow rates (obtained from the biomedical literature), tissue:air partition coefficients (obtained using QSPRs developed with rat data), blood:air partition coefficients (Pb) and hepatic clearance (CL). Using literature data on human Pb and CL for several VOCs (alkanes, alkenes, haloalkanes and aromatic hydrocarbons), multilinear additive QSPR models were developed. The numerical contributions to human Pb and CL were obtained for eleven structural fragments (CH3, CH2, CH, C, C=C, H, Cl, Br, F, Benzene ring, and H in benzene ring structure). Using these data as input, the PBPK model written in Excel® spreadsheet simulated the inhalation pharmacokinetics of ethylbenzene (33 ppm, 7 hr) and dichloromethane (100 ppm, 6 hr) in humans exposed to these chemicals. The QSPRs developed in this study should be useful for predicting the inhalation pharmacokinetics of VOCs in humans, prior to testing and experimentation.

Keywords: QSAR; PBPK models; pharmacokinetics; structure; internal dose; VOCs.

4.1 Introduction

Quantitative structure-activity relationship (QSAR), quantitative structure-property relationship (QSPR) and physiologically-based pharmacokinetic (PBPK) models are increasingly finding a place in the tool box of risk assessors [1]. When these models are developed using specialized software, the user can obtain standardized results easily but cannot visualize the modeling processes per se, i.e., how the software numerically resolves each equation in a model, as well as how the numerical output from each equation is provided as input to another. The knowledge of these processes is essential for an understanding of the interrelationships between the parameters and the model equations, as well as the equations themselves.

Spreadsheet programming offers a unique environment for visualizing the real-time solution and working of the model. Spreadsheet programs are commonly used and easily understood by biologists and scientists in general. The feasibility of solving PBPK model equations using a spreadsheet program (such as Microsoft Excel®) has been previously demonstrated [2,3]. In this approach, the influence of physiological, physicochemical and biochemical parameters on internal dose measures (e.g., blood concentration) can be visualized in real time and depicted using the spreadsheet's native graphing function. Since recent work has demonstrated the usefulness of QSAR/QSPR approaches in relating input parameters for PBPK models (partition coefficients, metabolic clearance) to chemical structure information [4], it should be possible to construct a QSPR-PBPK

model in spreadsheets which would facilitate the simulation of pharmacokinetics of chemicals in humans as a function of chemical structure information. Such a framework, if developed, can uniquely permit real-time visualization of the impact of structural fragments on the pharmacokinetic outcome in humans.

The objective of the present study was therefore to develop a methodology for integrating QSPRs within PBPK models using a spreadsheet program in order to simulate the inhalation kinetics of volatile organic chemicals (VOCs) in humans.

4.2 Methods

PBPK Model and Equations

The PBPK model used in this study describes the human body as a set of four tissue compartments (liver, slowly perfused tissues, richly perfused tissues and fat) interconnected by systemic blood circulation [2,3]. The algebraic and differential equations constituting the model and the manner in which they were entered in the spreadsheet are given in Table 1. Three sets of parameters (physiological, physicochemical and biochemical) were necessary In order to solve the PBPK model equations (Table 2). Physiological parameters include flows such as cardiac output (Qch), alveolar ventilation (Qph), tissue blood flow rates (Qlh, Qfh, Qrh and Qsh), as well as tissue volumes (Vlh, Vfh, Vrh, and Vsh). Physicochemical parameters include the blood:air partition coefficient (Pbh) and tissue:blood partition coefficients (Plh, Prh, Psh, and Pfh). Tissue:blood partition coefficients for volatile organic chemicals (VOCs) are obtained by dividing the tissue air partition coefficients by the blood:air partition coefficient. The principal biochemical parameter included in the PBPK model is the hepatic clearance (CL_h). The partition coefficients and hepatic clearance were estimated using QSPRs.

The overall approach consisted of linking structure-PBPK parameter relationships with a human PBPK model written in Microsoft Excel® to simulate the pharmacokinetic profile as a function of the molecular structure. A workbook was created in Microsoft Excel® and it consisted in 3 sheets: i) a parameter sheet, ii) a PBPK model sheet, and iii) a simulation sheet.

Parameter Sheet

All user input (shaded in the spreadsheet) necessary to run the PBPK model are found in the parameter sheet (Figure 1). These input parameters include chemical name (B2), molecular weight (D2), and exposure descriptions such as inhaled concentration (D5) and exposure duration (D6), as well as time functions for the simulations such as the integration interval (B17). Also present in this sheet are the human physiological parameters (G3:J8). The chemical-specific parameters were calculated using fragment-specific contributions and frequency of occurrence of fragments according to a linear additive model [4]. This QSPR model assumes that each molecular fragment has a unique, finite and additive contribution to the PBPK model parameter. In this study, the molecular fragments and their contributions to tissue:air partition coefficients of VOCs in rats, as determined previously, were used to calculate the tissue:air partition coefficients for human PBPK models, since the rat-human difference in tissue composition is not marked [3]. The numerical values of fragment contributions to the liver:air (Pl), muscle:air (Pm) and fat:air (Pf) partition coefficients, obtained from a previous rat modeling effort [4], are provided in Table 3 and their entry into spreadsheets is depicted in Figure 1.

The remaining chemical-specific parameters for the human PBPK models were estimated following the development of QSPRs. A multilinear model (as developed for the rat tissue:air partition coefficients) was used to relate the structure of VOCs to human blood:air partition coefficient and hepatic clearance. The experimentally-determined mean values of human Pb (blood:air partition coefficient) as well as rodent K_m (Michaelis affinity constant) and V_{max} (maximal

velocity for metabolism) for several VOCs were obtained from the literature [5-7]. The human Pb values were available for the following VOCs: alkanes (1-chloropropane, 2-chloropropane, a,2-dichloropropane, n-propyl bromide, cyclohexane, 2,2,4-trimethyl pentane), halomethanes (chloromethane, dichloromethane, chloroform, carbon tetrachloride, chlorodibromomethane), haloethanes (chloroethane, 1,1-dichloroethane, 1,2-dichloroethane, 1,1,1-trichloroethane, 1,1,2-trichloroethane, 1,1,1,2-tetrachloroethane, hexachloroethane, 1-bromo-2-chloroethane), alkenes (1,2-dichloroethylene, cis-1,2-dichloroethylene, trichloroethylene, terachloroethylene, vinyl chloride, vinyl bromide) and aromatic hydrocarbons (benzene, chlorobenzene, toluene, styrene, m-xylene). The V_{max} and K_m values were only available for the halomethanes, haloethanes, alkenes and aromatic hydrocarbons [5-7].

While V_{max} in humans was calculated using the following allometric equation: V_{max} (mg/hr/kg) x $BW^{0.74}$ where BW = body weight (kg), K_m was assumed to be species-invariant and is consistent with current state of knowledge when the same isoenzyme is known to be involved in the metabolism of VOCs in both the rodents and humans. The K_m and V_{max} values, obtained from rat studies, were used to calculate CL_h as follows [8]:

$$CL_h = \frac{Q_l * V_{max}/K_m}{V_{max}/K_m + Q_l} \qquad [1]$$

where Q_l = liver blood flow rate (109 L/hr) [9]. The CL_h value obtained as above is usually applicable for relatively low exposure situations when the K_m is greater than the venous blood concentration leaving the liver [8]. The numerical values of human Pb and CL_h (i.e., P_{PBPK}) were used along with the frequency of occurrence of each fragment (f) in the various VOCs in a multilinear additive model [4]:

$$\text{Log}P_{PBPK} = \sum_{i=1} f_i \cdot C_i \qquad [2]$$

to characterize the numerical value representing the contribution of each fragment i (C_i).

Multilinear regressions were performed to solve eqn. 2 for human parameters using a commercially-available statistical software package (SPSS® for Windows® v10.0.7, SPSS Inc., Chicago, IL), and the results obtained were essentially the values of C_i for each of the eleven structural fragments (CH_3, CH_2, CH, C, C=C, H, Cl, Br, F, AC, H_AC) specific to the parameter of interest. The statistical significance ($p < 0.05$) of the regression coefficients was assessed by computing *t*-statistic (> 1.8) [4].

Cross-validation was performed to evaluate all QSPRs. The sum of the squares of the residuals (i.e., the predictive residual sum of squares (PRESS) statistic) over the sum of squares of the response values (SSY) was used for validation [4]. The PRESS statistic measures how well the regression equation fits the data. It is computed by removing the ith datapoint from the dataset, computing the regression equation without this datapoint, predicting that point based on the regression equation, then computing the residual ("leave-one-out" procedure). This process is repeated for each datapoint, followed by the summing up of the sqares of each calculated residual (i.e., the difference between the predicted value of the parameter when a chemical is "left-out" and the experimental value for that chemical). To be a reasonable QSPR model, the ratio must be <0.4. A value of PRESS/SSY of <0.1 indicates an excellent model [10].

These validated QSPRs (i.e., the previously validated QSPRs for the tissue:air PC and the newly developed QSPRs for human blood:air PC and CL_h)

were written in Excel® within the lower half of the parameter sheet, reserved for calculated parameters (Figure 1). The QSPRs for each chemical-specific PBPK model parameter as well as their description in the spreadsheet are presented in Table 4. The fragment-specific contribution to the parameter values (B23:L27) were used along with the frequency of occurrence of the appropriate fragment in the molecule (B13:L13) as specified by the user, for estimating the chemical-specific PBPK model parameters (M23:N27) in humans. The lower half of this sheet also contains the calculated inhaled concentration in mg/L (B29) and as well as the tissue:blood partition coefficients calculated as tissue:air over blood:air (C33:C37).

PBPK Model Sheet

The PBPK model sheet contains the differential and algebraic equations necessary to simulate the pharmacokinetics of chemicals in humans. All equations entered in this sheet are presented in Table 1. These equations were copied onto each line of the spreadsheet as described previously [2]. Simulation over time was accomplished by numerically solving the differential equations of the PBPK model according to Euler algorithm with an integration interval (Δt) of 0.005 h [2].

Simulation Sheet

The simulation sheet consists essentially of the graphical output of the simulation in progress. It also contains any experimental pharmacokinetic data available for the chemical of interest which can be compared with the simulation output of the

model. Once these three sheets are set up, the QSPR-PBPK model can be used to provide simulations of the pharmacokinetics of chemicals. A workbook consisting of these three sheets for QSAR-PBPK modeling, developed in this study, can be obtained by writing the authors.

Spreadsheet Simulation

Simulation was accomplished by select-copying the last line in the PBPK model sheet using the fill handle function to the desired simulation time (row B), as described previously [4]. Spreadsheet windows could be arranged to visually accommodate all three sheets.

Model Validation

To validate the QSPR-PBPK model for humans, simulations for ethyl benzene (EBZ) (1*CH3+1*CH2+1*AC+5*H_AC) kinetics following exposure to 33 ppm (7 hr exposure) and dichloromethane (DCM) (2*CL+1*CH2) at 100 ppm (6 hr exposure) were performed as described and the simulated blood concentration profiles were compared to experimental values found in the literature. Of these two chemicals, ethyl benzene was not part of the calibration set whereas dichloromethane was. It should be noted that the pharmacokinetic data of dichloromethane was not used in developing the QSPR model, rather its blood:air and hepatic clearance values were. So the pharmacokinetic data for dichloromethane used for validation were not part of the original calibration data sets used for QSPR development.

Impact of the Fragment Constants on the Internal Dose

The sensitivity of the simulated internal dose (the area under the plasma concentration vs time curve, AUC) to the fragments of the two chemicals (EBZ and DCM) was evaluated. In this case, the impact on the internal dose (AUC) was computed as a sensitivity ratio (SR):

$$SR = \frac{\left(\frac{AUC_{+1\%} - AUC_{0\%}}{AUC_{0\%}}\right) * 100}{\left(\frac{Cf_{+1\%} - Cf_{0\%}}{Cf_{0\%}}\right) * 100}$$

where $AUC_{+1\%}$ = simulated AUC when fragment contribution is increased by 1%,

$AUC_{0\%}$ = simulated AUC with original fragment contribution,

$Cf_{+1\%}$ = fragment contribution (increased by 1%), and

$Cf_{0\%}$ = original fragment contribution.

Cf was increased by 1%, for each fragment contributing to a single PBPK parameter or with respect to all PBPK parameters simultaneously, and the resulting AUCs computed.

4.3 Results

Validation of Rat QSPRs With Human Tissue:air PC Data

Figure 2 presents a comparison of the rat QSPR model predictions of tissue:air PCs and the human experimental values for the same chemicals taken from the literature. Overall, the regression was positive and significant (R^2 = 0.96), demonstrating the validity of rat QSARs for predicting human tissue:air partition coefficients of the VOCs investigated in the present study.

QSPRs for Human Blood:air PC and CL_h

The fragment-specific contributions to human blood:air partition coefficients and hepatic clearance as determined with the QSPRs developed in this study are presented in Table 5. The overall regression was highly positive and significant (R^2 for Pb and CL_h were 0.92 and 0.96, respectively). Furthermore, the PRESS/SSY statistic of for these parameters was 0.16 and 0.06, respectively. These values are within the suggested PRESS/SSY benchmark value (0.4) for an adequate QSPR model. It should be noted, however, that not all fragment contributions were statistically significant (*t* statistic < 1.8). This may either be due to the fact that the available data were not sufficient to yield statistical confidence or that not all fragments contribute significantly to P_{PBPK}.

The mean (±SD) experimental/QSPR-estimated ratio was 1.13±0.61 for Pb and 1.00±0.2 for CL_h. Overall, of the 54 experimental CL_h and Pb parameter values, 47 were within a factor of two of the QSPR-estimated values, whereas 5

Pb estimates were within a factor of 2-3 of the experimental values. The remaining cases had an experimental to predicted ratio smaller than 0.33 (carbon tetrachloride (0.23) and cyclohexane (0.30)).

The contributions of each of the 11 molecular fragments to human Pb and CL_h as obtained from QSPR modeling, as well as the fragment contributions to the tissue:air partition coefficients previously obtained in rats [4], were entered into the lower half of the parameter sheet of the workbook as shown in Figure 1.

QSPR-PBPK Modeling

The workbook consisting of three spreadsheets was then used to simulate the inhalation pharmacokinetics of VOCs based solely on the knowledge of molecular structure and exposure information. The solution to the set of QSPRs and differential equations constituting the integrated model for EBZ and DCM resulted in temporal evolution of blood and tissue concentrations of these chemicals in humans. Portions of these simulated data calculated and displayed for each integration interval, for human exposure to 33 ppm EBZ and 100 ppm DCM are depicted in Figures 3 and 4. These figures also present a comparison of the QSPR-PBPK model simulations of the inhalation pharmacokinetics for EBZ and DCM with experimental data collected in humans. In order to obtain these simulations, only the number of the fragments along with the exposure concentration and duration were provided as input to the QSPR-PBPK model contained in the workbook. Because of the nature of linkages among data contained in various cells of spreadsheets, every time an input parameter is changed its impact on the pharmacokinetic endpoint can be visualized on a real-

time basis. To illustrate this, the sensitivity of simulated internal dose (the area under the blood concentration vs time curve, AUC) of EBZ and DCM associated with change in fragment contribution values was evaluated. These analyses indicated that the aromatic ring in EBZ and chlorine in DCM were the most sensitive parameters affecting the global result of internal dose simulations (Table 6). An examination of the sensitivity of internal dose to changes in fragment-specific contributions of individual input parameters, indicated that the aromatic hydrogen in CL_h and chlorine contribution in Pb were the most sensitive parameters for EBZ and DCM, respectively.

4.4 Discussion

Risk assessment approaches are increasingly being focused on the use of tissue or blood concentrations of chemicals rather than their concentrations in the environment. The rate and magnitude of uptake and accumulation in biota are determined by pharmacokinetic processes. The PBPK models are scientifically-sound tools that allow the simulation of blood and tissue concentrations of chemicals as a function of the rates of their absorption, distribution, metabolism and excretion. Simulation of the pharmacokinetics of chemicals in humans from molecular structure information integrated within PBPK models, as done in the present study, provides a novel approach of practical use to risk assessors.

Even though a variety of simulation software has been used to develop PBPK models [3], the use of spreadsheets is preferable since the user can visualize the mechanics of the simulation and see the results when the simulation is conducted. The spreadsheet programming also provides a unique framework to visualize the impact of structural fragments on the pharmacokinetic profile of chemicals in humans. Towards this goal, the present study constructed an Excel ® workbook consisting of three spreadsheets. This contrasts with the previous spreadsheet developed by Haddad (1996) in which all of the model parameters, graphs, and equations were contained on one spreadsheet. However, due to need to integrate QSPR equations within the PBPK model framework, the one-spreadsheet-type model proved impractical, though that does not in any way reduce the validity. In order to improve the practicality of the spreadsheet model, we split the components of the QSPR-PBPK model into three work sheets. The model parameters, QSPRs and exposure scenario data were entered into a

parameter sheet. All of the model equations and simulations over time were contained in the model sheet. Finally, the real time graphical output of blood and time concentration over time and the various experimental data needed for comparison purposes were located within the simulation sheet. These three sheets were subsequently rearranged such that user input and real time graphical simulation can be accommodated.

For modeling the pharmacokinetics of VOCs in humans, QSPRs relating chemical structure and parameters need to be developed, as has previously been done for VOCs in rats [4]. Ideally, a compilation of tissue:air partition coefficients experimentally determined in human tissues would be used for the derivation of human tissue:air QSPRs. However, human tissue:air partition coefficient data is limited. Because of the similarity in the tissue composition between rats and humans, already available rat tissue:air QSPRs [4] were used for predicting human tissue:air PCs. For blood:air PCs and CL_h values, however, due to interspecies differences in blood protein binding capacity and metabolic enzyme activity, we sought to the development of QSPRs for these two parameters in humans.

Following the development of human-specific QSPRs for CL_h and Pb, these were incorporated, along with the rat QSPRs for tissue:air partition coefficients, into the parameter section of the workbook. This enabled the simulation of the kinetics of VOCs belonging to different classes in humans using only the frequency of occurrence of specific fragments in the chemical simulated. By simply varying the number of fragments in a molecule (therefore simulating a different chemical each time the user input is changed), different blood concentration profiles could be obtained, and, for a given exposure scenario, the chemical possessing a

desired pharmacokinetic profile within a series (e.g., the chemical displaying the lowest C_{max}) can be identified.

C_{max}, or any other pharmacokinetic endpoint (i.e., AUC of the parent chemical), predicted with a model, is only as accurate as the model's most sensitive parameter(s). The most sensitive of these parameters will depend on the endpoint of interest and the chemical being studied. Because of this, QSPRs, as developed for pharmacokinetic models, should strive towards predicting the most sensitive parameter(s).

There are therefore two levels of sensitivities that need to be accounted for when assessing the impact of any given fragment on the PK, the first relating to the innate sensitivity of the model to its chemical-specific parameters, and the second relating to the sensitivity of QSPRs for a parameter to its fragment contributions. In other words, we need to distinguish between the impact of varying the contribution value for a fragment in general, and the impact of varying the contribution value for a fragment in the model's most sensitive parameter.

Because sensitivity ratios of blood concentration profiles is the result of both intrinsic sensitivity of the model to the parameters and sensitivity of the parameters to the group contributions within the QSPRs, these two cannot be considered separately. For example, a large variation in the contribution of a fragment to the richly perfused tissue:air PC is not expected to have a significant effect on the internal dose if the internal dose is not sensitive to changes in the value of this model parameter.

It is also expected that the overall value of the fragment contribution has an impact on the sensitivity of the internal dose. For example, the aromatic hydrogen in EBZ contributes the most to the CL_h observed. Therefore, any error on that

value is expected to have the greatest impact on the outcome. Similarly, the chlorine in DCM contributes the most to the blood:air PC, and the internal dose is sensitive to that variation as illustrated in Table 6.

This sensitivity analysis can be used as a tool for developing better QSPRs. Critical fragments can be identified and further experimental work with chemicals containing these fragments can be undertaken. Increasing the available data on chemicals containing fragments that have been identified as having a significant impact on the internal dose can improve the overall confidence in the QSPR generated and the resulting PBPK parameter value for a *de novo* chemical. In conclusion, the QSPRs developed in this study should be useful in predicting – prior to any experimentation - the inhalation pharmacokinetics in humans of low molecular weight VOCs that contain one or more of the 11 fragments evaluated in this study and metabolized by hepatic P-450 2E1.

Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council and the Canadian Network of Toxicology Centres. M.B. is recipient of scholarships from the Natural Sciences and Engineering Research Council of Canada and Fonds de Recherche sur la Nature et les Technologies.

4.5 References

[1] Beliveau, M. and Krishnan, K. (2003) "In silico approaches for physiologically based pharmacokinetic modeling". In: Salem, H. and Katz, S.A. eds, Alternative Toxicological Methods, (CRC Press LLC, Boca Raton), pp. 479-532.

[2] Haddad, S., Pelekis, M. and Krishnan, K. (1996) "A methodology for solving physiologically based pharmacokinetic models without the use of simulation softwares." *Toxicol. Letters* **85**, 113-126.

[3] Krishnan, K. and Andersen, M.E. (2001) "Physiologically based pharmacokinetic modeling in toxicology." In: Hayes, A.W. ed, Principles and methods of toxicology, (Taylor & Francis, Philadelphia), pp. 193-241.

[4] Beliveau, M., Tardif, R. and Krishnan, K. (2003) "Quantitative structure-property relationships for physiologically based pharmacokinetic modeling of volatile organic chemicals in rats", *Toxicol. Appl. Pharmacol.* **189**, 221-32.

[5] Gargas, M.L., Burgess, R.J., Voisard, D.E., Cason, G.H. and Andersen, M.E. (1989) "Partition coefficients of low-molecular-weight volatile chemicals in various liquids and tissues", *Toxicol. Appl. Pharmacol.* **98**, 87-99.

[6] Gargas, M.L., Seybold, P.G. and Andersen, M.E. (1988) "Modeling the tissue solubilities and metabolic rate constant (Vmax) of halogenated methanes, ethanes, and ethylenes", *Toxicol. Letters* **43**, 235-256.

[7] Haddad, S., Charest-Tardif, G., Tardif, R. and Krishnan, K. (2000) "Validation of a physiological modeling framework for simulating the toxicokinetics of chemicals in mixtures." *Toxicol. Appl. Pharmacol.* **167**, 199-209.

[8] Poulin, P. and Krishnan, K. (1999) "Molecular structure-based prediction of the toxicokinetics of inhaled vapors in humans." *Int. J. Toxicol.* **18**, 7-18.

[9] Tardif, R., Charest-Tardif, G., Brodeur, J. and Krishnan, K. (1997) "Physiologically based pharmacokinetic modeling of a ternary mixture of alkyl benzenes in rats and humans." *Toxicol. Appl. Pharmacol.* **144**, 120-134.

[10] Wold, S. (1991) "Validation of QSAR's", *QSAR* **10**, 191-193.

[11] Andersen, M.E., Clewell, H.J.I. and Gargas, M.L. (1991) "Physiologically-based pharmacokinetic modeling with dichloromethane, its metabolite carbon monoxide and blood carboxyhemoglobin in rats and humans." *Toxicol. Appl. Pharmacol.* **108**, 14-27.

[12] Ward, R.C., Travis, C.C., Hetrick, D.M., Andersen, M.E. and Gargas, M.L. (1988) "Pharmacokinetics of tetrachloroethylene." *Toxicol. Appl. Pharmacol.* **93**, 108-117.

[13] Reitz, R.H., Mandrela, A.L., Corley, R.A., Quast, J.F., Gargas, M.L., Andersen, M.E., Staats, D.A. and Conolly, R.B. (1990) "Estimating the risk of liver cancer associated with human exposures to chloroform using physiologically-based pharmacokinetic modeling." *Toxicol. Appl. Pharmacol.* **105**, 443-459.

[14] Tardif, R., Lapare, S., Charest-Tardif, G., Brodeur, J. and Krishnan, K. (1995) "Physiologically-based pharmacokinetic modeling of a mixture of toluene and xylene in humans", *Risk Anal.* **15**, 335-342.

[15] Ramsey, J.C. and Andersen, M.E. (1984) "A physiologically-based description of the inhalation pharmacokinetics of styrene in rats and humans." *Toxicol. Appl. Pharmacol.* **73**, 159-175.

[16] Paustenbach, D., Andersen, M.E., Clewell, H.J. and Gargas, M.L. (1988) "A physiologically-based pharmacokinetic model for inhaled carbon tetrachloride in the rat." *Toxicol. Appl. Pharmacol.* **96**, 191-211.

[17] Tardif, R., Lapare, S., Krishnan, K. and Brodeur, J. (1993) "Physiologically based modeling of the toxicokinetic interaction between toluene and m-xylene in the rat", *Toxicol. Appl. Pharmacol.* **120**, 266-273.

[18] Reitz, R.H., Gargas, M.L., Andersen, M.E., Provan, W.M. and Green, T.L. (1996) "Predicting cancer risk from vinyl chloride exposure with a physiologically based pharmacokinetic model." *Toxicol. Appl. Pharmacol.* **137**, 253-267.

[19] Allen, B.C. and Fisher, J.W. (1993) "Pharmacokinetic modeling of trichloroethylene and trichloroacetic acid in humans." *Risk Anal.* **13**, 71-86.

[20] Fisher, J.W., Gargas, M.L., Jepson, G.W., Allen, B. and Andersen, M.E. (1991) "Physiologically-based pharmacokinetic modeling with trichloroethylene and its metabolite, trichloroacetic acid in the rat and mouse." *Toxicol. Appl. Pharmacol.* **109**, 183-195.

[21] Reitz, R.H., Gargas, M.L., Mendrala, A.L. and Schumann, A.M. (1996) "In vivo and in vitro studies of perchloroethylene metabolism for physiologically based pharmacokinetic modeling in rats, mice, and humans." *Toxicol. Appl. Pharmacol.* **136**, 289-306.

[22] Andersen, M.E., Clewell, H.J.I., Gargas, M.L., Smith, F.A. and Reitz, R.H. (1987) "Physiologically-based pharmacokinetics and risk assessment process for methylene chloride." *Toxicol. Appl. Pharmacol.* **87**, 185-205.

Table 1 PBPK model equations, as entered in the spreadsheet[a].

Tissue	*Equations*	*Cell*	*Spreadsheet*
Blood	$Ca = \frac{QpCi + QcCv}{Qc + Qp/Pb}$	D4	=((Qch*U3)+(Qph*C4))/(Qch+(Qph/Pbh))
	$Cv = \frac{QlCvl + QsCvs + QrCvr + QfCvf}{Qc}$	U4	=(((Qlh*H4)+(Qfh*L4)+(Qrh*P4)+(Qsh*T4))/Qch)
Liver	$dAl/dt = Ql(Ca - Cvl) - CLh * Ca$	E4	=Qlh*($D4-H3)-CLhqsar*$D4
	$Al = dAl/dt * \Delta t + Al_{t-1}$	F4	=(E4*t)+F3
	$Cl = Al/Vl$	G4	=F4/Vlh
	$Cvl = Cl/Pl$	H4	=G4/Plh
Fat	$dAf/dt = Qf(Ca - Cvf)$	I4	=(Qfh*($D4-L3))
	$Af = dAf/dt * \Delta t + Af_{t-1}$	J4	=I4*t+J3
	$Cf = Af/Vf$	K4	=J4/Vfh
	$Cvf = Cf/Pf$	L4	=K4/Pfh
Richly	$dAr/dt = Qr(Ca - Cvr)$	M4	=(Qrh*($D4-P3))
	$Ar = dAr/dt * \Delta t + Ar_{t-1}$	N4	=M4*t+N3
	$Cr = Ar/Vr$	O4	=N4/Vrh
	$Cvr = Cr/Pr$		=O4/Prh

		P4	
Slowly	$dAs/dt = Qs(Ca - Cvs)$ $As = dAs/dt * \Delta t + As_{t-1}$ $Cs = As/Vs$ $Cvs = Cs/Ps$	Q4 R4 S4 T4	=(Qsh*($D4-T3)) =Q4*t+R3 =R4/Vsh =S4/Psh

[a] adapted from Haddal *et al.* [2].

Table 2 Model parameters for the PBPK model presented in this study.

Parameter	*Value*	*Cell*	*Cell Name*
Physiological			
Flows (L/hr)			
Cardiac output[a]	417	J3	Qch
Alveolar ventilation[b]	417	J4	Qph
Liver[c]	109	J5	Qlh
Fat[c]	20.9	J6	Qfh
Richly Perfused Tissues[c]	183.7	J7	Qrh
Slowly Perfused Tissues[c]	104.4	J8	Qsh
Volumes (L)[d]			
Liver	1.82	I5	Vlh
Fat	13.3	I6	Vfh
Richly Perfused Tissues	3.5	I7	Vrh
Slowly Perfused Tissues	43.4	I8	Vsh
Physicochemical			
Blood/air	N23	B33	Pbh
Liver/blood	$\frac{N24}{N23}$	C34	Plh
Richly/blood	$\frac{N24}{N23}$	C36	Prh
Slowly/blood	$\frac{N25}{N23}$	C37	Psh
Fat/blood	$\frac{N26}{N23}$	C35	Pfh

Biochemical			
CLh	M27	N27	CLhqsar

[a] Based on an allometric constant of 18 (H3) [9].

[b] Based on an allometric constant of 18 (H4) [9].

[c] Calculated using data on fraction of cardiac output (H5 to H8) [9].

[d] Calculated using data on fractional volume (G5 to G8) [9].

Table 3 Fragment contributions to the rat tissue:air partition coefficients[a].

Fragments	*Log P_l*	*Log P_m*	*Log P_f*
CH_3	0.0155	-0.0197	0.366
CH_2	0.234	0.122	0.435
CH	0.359	0.266	0.330
C	0.0318	-0.105	-0.285
C=C	0.257	-0.707	0.327
H	-0.0305	0.0813	0.155
Br	0.700	0.622	1.17
Cl	0.384	0.322	0.735
F	-0.113	-0.911	0.0752
AC	3.76	3.65	2.92
H_AC	-0.408	-0.446	-0.558

[a] from Beliveau *et al.* [4].

Table 4 Quantitative-structure property relationships (QSPR) as entered in the spreadsheet model.

QSPR		
Algebraic[a]	*Spreadsheet*	*Cell*
$LogPb = \sum_{i=1}^{n} Cf_{Pb,i} * nf_i$	=Parameters!$B4*B23+Parameters!$B5*C23+Parameters!$B6*D23+Parameters!$B7*E23+ Parameters!$B8*F23+Parameters!$B9*G23+Parameters!$B10*H23+Parameters!$B11*I23+ Parameters!$B12*J23+Parameters!$B13*K23+Parameters!$B14*L23	M23
$LogPl = \sum_{i=1}^{n} Cf_{Pl,i} * nf_i$	=Parameters!$B4*B24+Parameters!$B5*C24+Parameters!$B6*D24+Parameters!$B7*E24+ Parameters!$B8*F24+Parameters!$B9*G24+Parameters!$B10*H24+Parameters!$B11*I24+ Parameters!$B12*J24+Parameters!$B13*K24+Parameters!$B14*L24	M24
$LogPm = \sum_{i=1}^{n} Cf_{Pm,i} * nf_i$	=Parameters!$B4*B25+Parameters!$B5*C25+Parameters!$B6*D25+Parameters!$B7*E25+ Parameters!$B8*F25+Parameters!$B9*G25+Parameters!$B10*H25+Parameters!$B11*I25+ Parameters!$B12*J25+Parameters!$B13*K25+Parameters!$B14*L25	M25
$LogPf = \sum_{i=1}^{n} Cf_{Pf,i} * nf_i$	=Parameters!$B4*B26+Parameters!$B5*C26+Parameters!$B6*D26+Parameters!$B7*E26+ Parameters!$B8*F26+Parameters!$B9*G26+Parameters!$B10*H26+Parameters!$B11*I26+	M26

	Parameters!$B12*J26+Parameters!$B13*K26+Parameters!$B14*L26	
$CLh = \sum_{i=1}^{n} Cf_{CLh,i} * nf_i$	=Parameters!$B4*B27+Parameters!$B5*C27+Parameters!$B6*D27+Parameters!$B7*E27+ Parameters!$B8*F27+Parameters!$B9*G27+Parameters!$B10*H27+Parameters!$B11*I27+ Parameters!$B12*J27+Parameters!$B13*K27+Parameters!$B14*L27	M27

[a] where Pb, Pl, Pm, Pf, CLh, Cf, and nf are defined as blood:air, liver:air, muscle:air, and fat:air PC, hepatic clearance, contribution of fragment to parameter and number of fragment in molecule, respectively.

Table 5 Fragment contributions (Cf) to the human blood:air partition coefficient (Pb) and hepatic clearance (CLh).

Fragment	*Cf*			
	Log Pb	*t*	*CLh*	*t*
CH3	-4.82E-2	0.607	45.5	5.79
CH2	0.136	2.912	10.1	1.12
CH	0.262	1.473	-14.3	1.02
C	-0.193	0.824	-73.7	3.35
C=C	-0.116	0.310	-16.0	0.445
H (on C=C)	5.73E-2	0.407	19.5	1.45
Br	0.585	4.57	42.4	4.95
Cl	0.316	4.84	29.2	4.85
AC	3.05	2.55	-211	2.86
H (on AC)	-0.347	1.48	49.3	3.54

Table 6 Sensitivity ratios ($x10^{-3}$) reflecting the extent of change in human blood:air (P_b), liver:air (P_l), slowly perfused:air (P_m), fat:air (P_f) and hepatic clearance (CL_h) due to 1% change in the value of the fragments (CH2, CH2, AC, H_AC, Cl, CH2).

Group contribution change / *In PBPK model param.*	*in ethyl benzene*				*in dichloromethane*	
	*CH3*1.01*	*CH2*1.01*	*AC*1.01*	*H_AC*1.01*	*CH2*1.01*	*Cl*1.01*
P_b	-17.0	48.0	1052	-622	172	800
P_l	-0.000003	-0.00004	-0.0008	0.0004	-0.143	-0.474
P_m	0.000043	-0.00027	-0.0142	0.0035	-0.485	-2.58
P_f	-0.000029	-0.00003	-0.0002	0.0002	-50.1	-171
CL_h	-560	-125	2659	-2978	-99.7	-574
Global	-577	-76.8	3761	-3567	21.6	45.6

FIGURE LEGENDS

Figure 1 Input parameters sheet for the QSPR-PBPK workbook (ethyl benzene example shown).

Figure 2 Comparison of the rat QSPR estimated tissue:air partition coefficients and human experimental tissue:air partition coefficients. Experimental data from the literature [9, 11-21].

Figure 3 Ethyl benzene QSPR-PBPK workbook and simulation for a 33 ppm exposure (7 hrs) in humans. Experimental data from the literature [9].

Figure 4 Dichloromethane QSPR-PBPK workbook and simulation for a 100 ppm exposure (6 hrs) in humans. Experimental data from the literature [22].

Figure 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Chemical specific information					Physiological Parameters								
2	Chemical Name =	Ethyl Benzene	MW =	106.17		Tissue	fraction V (L	fraction Q (L	V (L)	Q (L/hr)				
3	Fragment	frequency in molecule				Body (c)		18	70	417.49				
4	CH_3	1	Exposure conditions			Lung (p)		18		417.49				
5	CH_2	1	Cinh (ppm)	33		Liver (l)	0.026	0.26	1.82	108.55				
6	CH	0	Length (hr)	7		Fat (f)	0.19	0.05	13.3	20.87				
7	C	0				Richly (r)	0.05	0.44	3.5	183.70				
8	C=C	0				Slowly (s)	0.62	0.25	43.4	104.37				
9	H	0				Total	0.896	1	62.02	417.49				
10	Br	0												
11	Cl	0												
12	F	0												
13	AC	1												
14	H_AC	5												
15														
16	Time Functions													
17	integration interval (t, h	0.005												
18														
19														
20	Calculated Parameters													
21								Fragments						
22	Contributions to	CH_3	CH_2	CH	C	C=C	H	Br	Cl	F	AC	H_AC	Sum of con	Value
23	Log P_b	-4.82E-02	0.136	2.62E-01	-0.193	-0.116	5.73E-02	0.585	0.316	0	3.05	-0.347	1.403	25
24	Log P_l	1.55E-02	0.234	0.359	3.18E-02	0.257	-3.05E-02	0.7	0.384	-0.113	3.76	-0.408	1.970	93
25	Log P_m	-1.97E-02	0.122	0.266	-0.105	-7.07E-01	8.13E-02	0.622	0.322	-9.11E-01	3.65	-0.446	1.522	33
26	Log P_f	0.366	0.435	0.33	-0.285	0.327	0.155	1.17	0.735	7.52E-02	2.92	-5.58E-01	0.931	9
27	CL_r	45.5	10.1	-14.3	-73.7	-16	19.5	42.4	29.2	0	-211	49.3	91.100	91.100
28														
29	Cinh (mg/L)	0.143296933												
30														
31	Physicochemical Parameters													
32	Tissue	Pt:a	Pt:b											
33	Blood	25.3												
34	Liver	93.2	3.687											
35	Fat	8.5	0.337											
36	Richly	93.2	3.687											
37	Slowly	33.3	1.317											

Figure 2

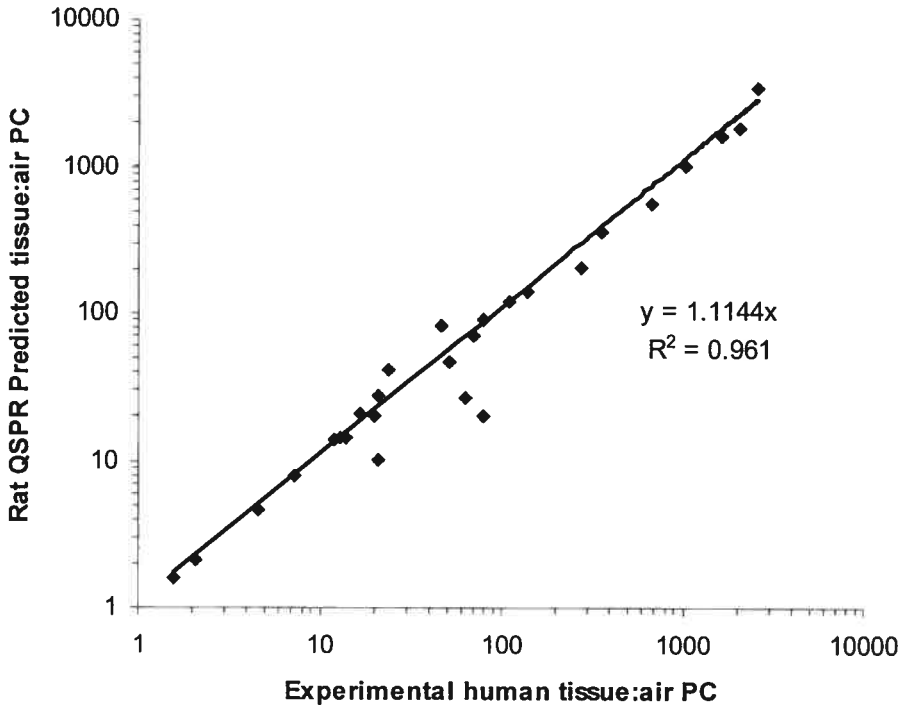

10000
1000
100
10
1
Rat QSPR Predicted tissue:air PC
y = 1.1144x
R2 = 0.961
1
10
100
1000
10000
Experimental human tissue:air PC

Figure 3

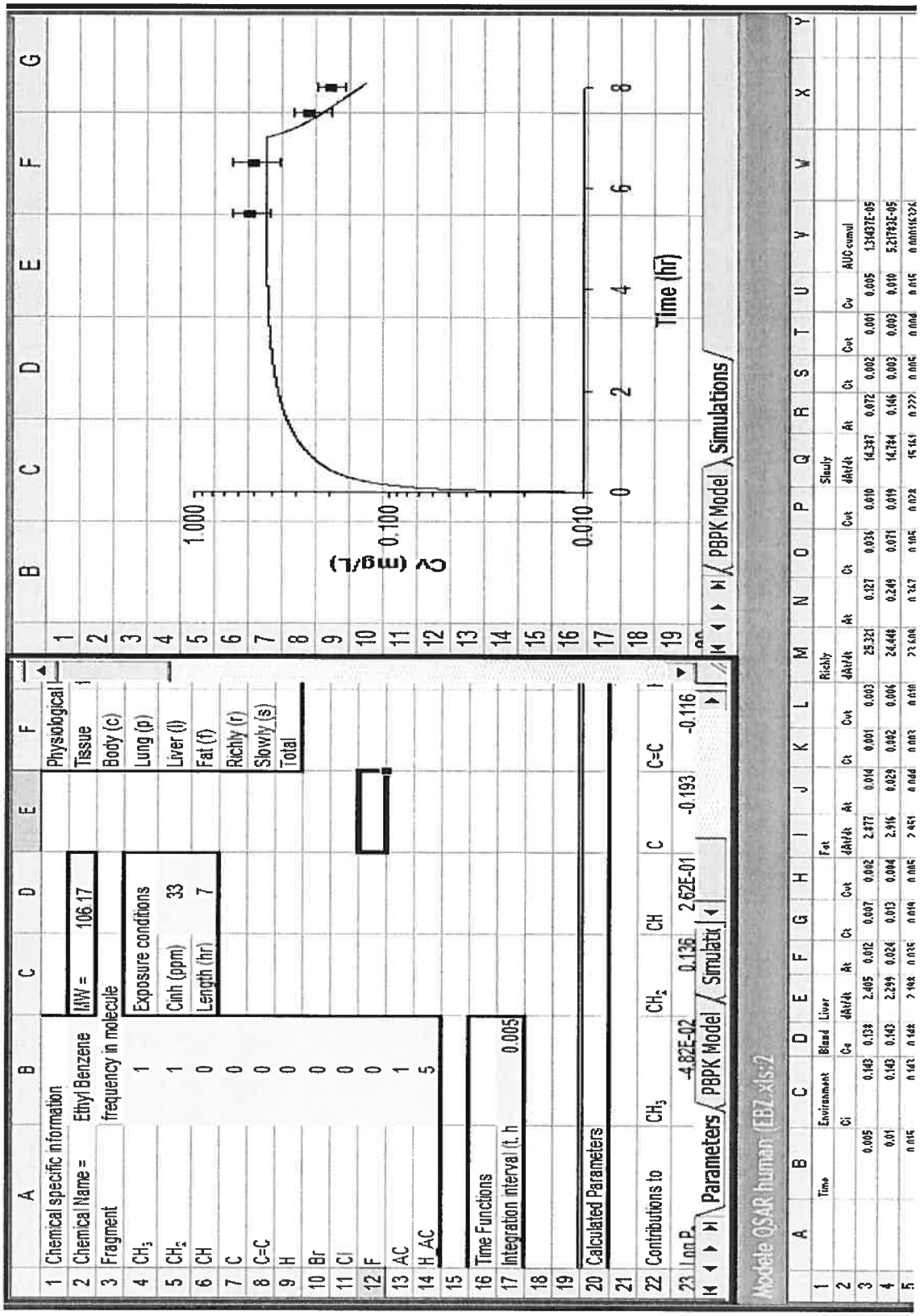
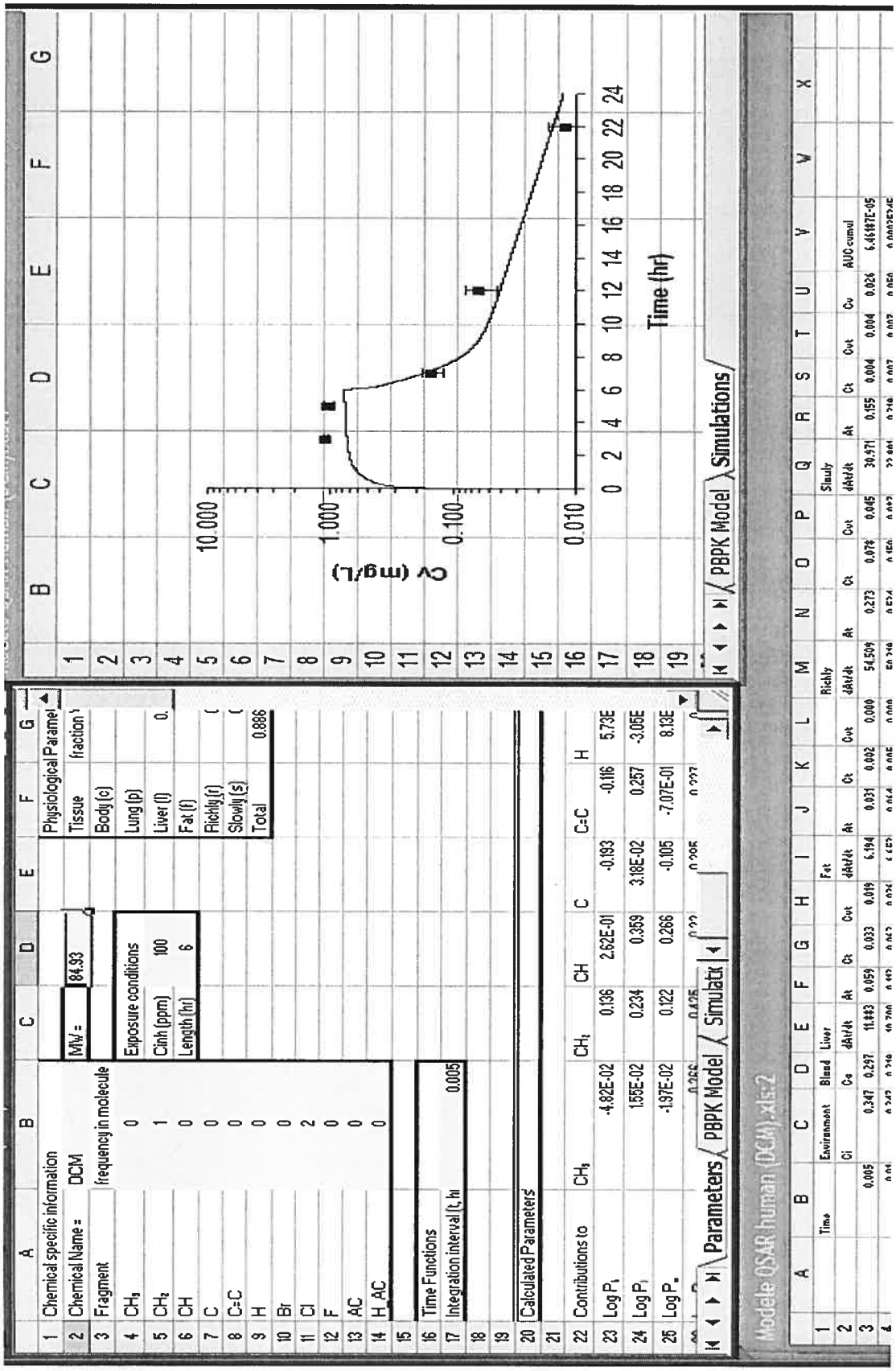

Chemical specific information
Chemical Name = Ethyl Benzene
MW = 106.17
Fragment frequency in molecule
Exposure conditions
Cinh (ppm) 33
Length (hr) 7
Physiological
Tissue
Body (c)
Lung (p)
Liver (l)
Fat (f)
Richly (r)
Slowly (s)
Total
Time Functions
Integration interval (t, h 0.005
Calculated Parameters
Contributions to
Parameters
PBPK Model
Simulations
Time (hr)
Modele QSAR human EBZ.xls:2

Figure 4


Chemical specific information
Chemical Name = DCM
MW = 84.93
Physiological Parame
Tissue fraction
Body (c)
Lung (p)
Liver (l) 0.
Fat (f)
Richly (r)
Slowly (s)
Total 0.886
Exposure conditions
Cinh (ppm) 100
Length (hr) 6
Time Functions
Integration interval (t_hr 0.005
Calculated Parameters
Contributions to
Log P
Parameters
PBPK Model
Simulatn
Simulations
Modele QSAR human (DCM).xls:2
Cv (mg/L)
Time (hr)

CHAPITRE CINQUIEME:

5 Article IV

Quantitative structure-property relationships for interspecies extrapolation of the pharmacokinetics of organic chemicals

For submission to: Chemical Research in Toxicology

Quantitative structure-property relationships for interspecies extrapolation of the inhalation pharmacokinetics of organic chemicals

Martin Béliveau[†], John Lipscomb[‡], Robert Tardif[†] and Kannan Krishnan[†*]

[†]Groupe de Recherche en Toxicologie Humaine (TOXHUM), Université de Montréal, Case Postale 6128, Succ. Centre-Ville, Montréal, PQ, H3C 3J7, Canada.

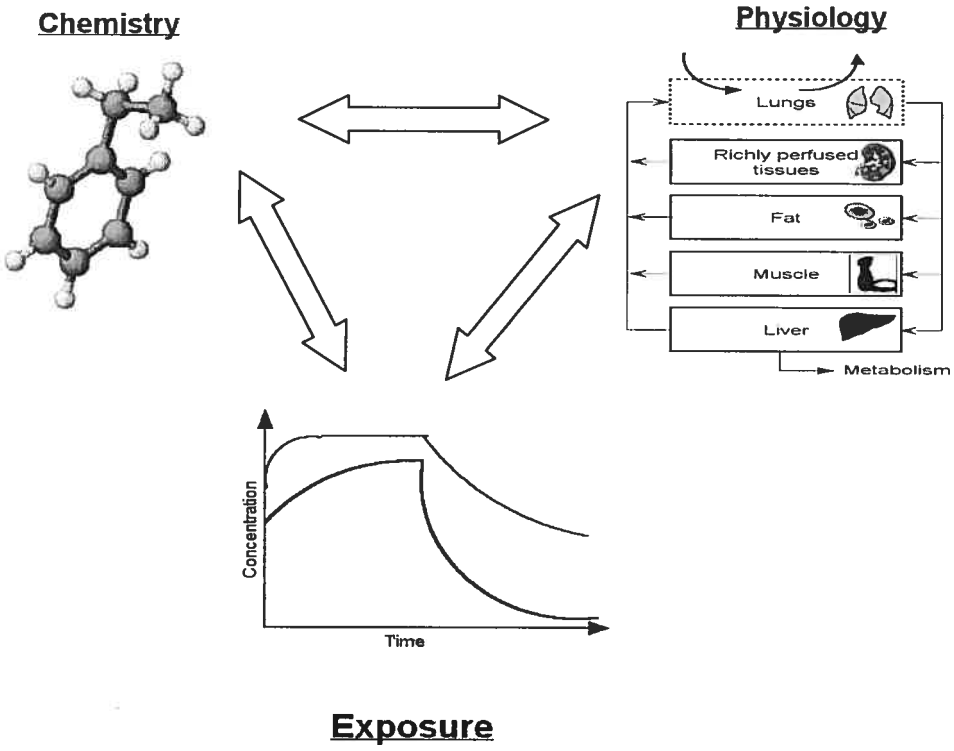
[‡] US Environmental Protection Agency, National Center for Environmental Assessment, 26 W Martin Luther King Drive, MD 190, Cincinnati, OH, 45268, USA.

*Corresponding author

Running title: QSPRs for the interspecies extrapolation of VOCs

Keywords: PBPK modeling, VOCs, pharmacokinetics, QSPR

Figure for the cover page



Abstract

The objectives of this study were to (i) develop quantitative structure-property relationships (QSPRs) for tissue:air partition coefficients (P_t) and hepatic clearance (CL_h), and (ii) conduct interspecies extrapolations of the pharmacokinetics of volatile organic chemicals (VOCs) by incorporating the above QSPRs within a physiologically-based pharmacokinetic (PBPK) modeling framework. P_t was predicted using the following algorithm: $F_{nlet}*P_{o:a} + F_{wet}*P_{w:a}+ f_b*F_p*P_{p:a}$, where F_{nlet} = fraction of neutral lipid equivalents in tissue t, F_{wet} = fraction of water equivalents in tissue t, $P_{o:a}$ = vegetable oil:air partition coefficient, $P_{w:a}$ = water:air partition coefficient, f_b = binding capacity of the protein fraction, F_p = fraction of binding protein in tissue, and $P_{p:a}$ = protein:air partition coefficient. CL_h was estimated as follows: $Q_l*[(CL_{int}*C_{2E1}*V_l)/(Q_l+CL_{int}*C_{2E1}*V_l)]$ where CL_{int} = normalized intrinsic clearance, Q_l = blood flow to the liver, C_{2E1} = concentration of CYP2E1 enzyme in the liver of the species of interest, and V_l = volume of the liver. QSPRs relating molecular fragments of 46 VOCs and parameters required for estimating P_t or CL_h ($P_{o:a}$, $P_{w:a}$, $P_{p:a}$, and CL_{int}) were established using a group contribution method (Σ fi*Ci where f = frequency of occurrence of the group i in a given molecule, and Ci = contribution of the group i to $P_{o:a}$, $P_{w:a}$, $P_{p:a}$, or CL_{int}). Values of group contributions were determined by multilinear regression of experimental data. The species-specific parameters required for solving the above algorithms were obtained from the literature. These algorithms, once incorporated into a multispecies PBPK modeling framework, enabled extrapolation of the kinetics of chemicals across species. The pharmacokinetics of dichloromethane, toluene, as well as two *de novo* compounds (1,2,4-trimethyl benzene and ethyl benzene) were extrapolated from rat to human and used to validate the methodology. The present

study has demonstrated that it is possible to extrapolate pharmacokinetic behavior of chemicals from rats to humans on the basis of QSPRs and species-specific physiological information.

Footnotes

[1] Abbreviations: PBPK : Physiologically-based pharmacokinetic, QSPR : Quantitative-structure property relationship, P_b : blood:air partition coefficient , P_l : liver:blood partition coefficient, P_m : muscle:blood partition coefficient, P_f : fat:blood partition coefficient, CL_h : hepatic clearance, VOC : volatile organic chemical, P_t : tissue:air partition coefficients, P_b : blood:air partition coefficients, $P_{o:a}$: oil:air partition coefficient, $P_{w:a}$: water:air partition coefficient, and $P_{p:a}$: protein:air partition coefficient, CL_{int} : enzyme content-normalized intrinsic clearance, Ql : blood flow to the liver, C_{2E1} : concentration of CYP2E1 enzyme in the liver of the species of interest, and Vl : volume of the liver, PRESS : predicted residual sum of squares statistic, SSY : the sum of squares of the response values, PC : partition coefficient, ACSL : Advanced Continuous Simulation Language, PK : pharmacokinetics, V_{max} : maximal velocity for metabolism, K_m : Michaelis affinity constant. MSP : microsomal protein, $P_{o:w}$: octanol:water partition coefficient.

5.1 Introduction

Pharmacokinetic models are useful for simulating the blood and tissue concentrations of chemicals in biota as a function of time and exposure conditions. Physiologically-based pharmacokinetic (PBPK) models are mathematical descriptions of uptake and disposition of chemicals, finding use in the conduct of interspecies and inter-chemical extrapolations. For conducting inter-chemical extrapolation of the pharmacokinetics in a given species (e.g., rat), the chemical-specific PBPK model parameters need to be characterized. Beliveau *et al.* (*1*) recently used quantitative structure-property relationships (QSPRs) to elucidate the contribution of each molecular fragment towards the numerical value of PBPK model parameters (blood:air partition coefficient (P_b), liver:blood partition coefficient (P_l), muscle:blood partition coefficient (P_m), fat:blood partition coefficient (P_f) and hepatic clearance (CL_h)) for the rat.

Although the fragment-specific contributions towards PBPK model parameters are expected to remain the same in a given biological species (i.e., rat, in this case), the QSPRs are not directly applicable to another species of interest (humans). The interspecies differences in fragment contributions and PBPK model parameter values result from interspecies differences in mechanistic determinants (e.g., lipid content, water content, enzyme concentrations). In fact, the magnitude of PBPK model parameters such as partition coefficients and metabolic constants depends on the chemical characteristics (e.g., lipophilicity, vapour pressure, molecular volume) as well as biological characteristics (e.g., enzyme levels, binding site, lipid and water contents of tissues and blood). The chemical-specific characteristics will remain the same regardless of the species, whereas the

biological parameters are likely to change from one species to another. Hypothetically then, any QSPR useful for interspecies extrapolation should be based on the consideration of chemical-specific characteristics and species-dependent biological characteristics. Such QSPRs for PBPK model parameters, which can be used in interspecies extrapolation of pharmacokinetics of chemicals, have not yet been developed.

The overall objective of this study was therefore to develop QSPRs for conducting interspecies extrapolations of PBPK model parameters. The specific objectives were: (i) to develop QSPRs that can account for chemical-specific and species-specific characteristics of blood:air partition coefficients, tissue:air partition coefficients and hepatic clearance of volatile organic chemicals (VOCs), and (ii) to conduct rat-to-human extrapolations of the inhalation pharmacokinetics of some VOCs (toluene, dichloromethane, ethylbenzene and 1,2,4-trimethyl benzene) by incorporating the above QSPRs within a PBPK modeling framework.

5.2 Methods

Chemicals

In the present study, QSPR development was undertaken using experimental data for 46 VOCs (monochloromethane, dichloromethane, trichloromethane, tetrachloromethane, difluoromethane, fluorochloromethane, bromochloromethane, dibromomethane, chlorodibromomethane, monochloroethane, 1,1-dichloroethane, 1,2-dichloroethane, 1,1,1-trichloroethane, 1,1,2-trichloroethane, 1,1,1,2-tetrachloroethane, 1,1,2,2-tetrachloroethane, pentachloroethane, hexachloroethane, 1,2-dibromoethane, 1-bromo-2-chloroethane, 1,1,1-trifluoro-2-chloroethane, 1,1,1-trifluoro-2-bromo-2-chloroethane, 1-chloropropane, 2-chloropropane, 1,2-dichloropropane, n-propylbromide, Isopropyl bromide, n-hexane, n-heptane, cyclohexane, 2,3,4-trimethylpentane, 2,2,4-trimethylpentane, monochloroethylene, 1,1-dichloroethylene, 1,2-dichloroethylene, trichloroethylene, tetrachloroethylene, vinyl bromide, benzene, chlorobenzene, toluene, styrene, methylstyrene as well as *m*-, *o*- and *p*-xylenes) (*2, 3*). All chemicals were described using various combinations of eleven fragments (CH_3, CH_2, CH, C, C=C, H, Cl, Br, F, benzene ring, and H in the benzene ring structure) as done previously (*4, 5*) (Table 1).

QSPR for partition coefficients

Tissue:air or blood:air partition coefficients (Pt or Pb) representing the equilibrium ratio of chemical concentration in the biological matrix and air, were predicted using the following equation based on previous work (*6*):

$$Pt \text{ or } Pb = Fnl*Po{:}a + Fw*Pw{:}a + fb*Fp*Pp{:}a \quad [1]$$

where Fnl = fraction of neutral lipid equivalents in the biological matrix (i.e., tissue or blood), Fw = fraction of water equivalents in the biological matrix, fb= binding capacity of proteins in the biological matrix, Fp = fraction of binding protein in the biological matrix, Po:a = vegetable oil:air partition coefficient, Pw:a = water:air partition coefficient, Pp:a = protein:air partition coefficient, and * indicates multiplication.

In order to predict Pb and Pt using the above equation, four species-specific parameters (Fnl, Fw, fb and Fp) and three chemical-specific parameters (Po:a, Pw:a and Pp:a, representing solubility in vegetable oil, water, and proteins) were required. For the VOCs investigated in the present study, protein binding is negligible in all biological matrices except blood (i.e., fb*Fp = 0), such that the third term of Eqn. [1] equals zero while calculating Pt values. However, the fb*Fp term is critical to correctly compute Pb values of VOCs, and neglecting this term leads to significant underestimation of the true partition coefficient (data not shown) between blood and air. The species-specific data on Fnl and Fw were obtained from the literature (Table 2) whereas the chemical-specific input parameters (Po:a, Pw:a and Pp:a) were determined using QSPRs.

QSPRs for the partition coefficients (Po:a, Pw:a and Pp:a) were established using a group contribution method (*5, 7*). Accordingly, chemical fragments or groups are considered to have unique and constant contributions, which can be summed to calculate the parameter of interest (P_i), as follows (*5*):

$$\log P_i = \sum_{i=1}^{ng} f_i \cdot C_i \qquad [2]$$

where f = frequency of occurrence of the group i in a given molecule, and C_i = contribution of the fragment i to P_i.

By providing experimentally-determined Pw:a and Po:a, as well as Pp:a derived from experimental data (*2*), along with the frequency of occurrence of each fragment in the 46 VOCs, the fragment-specific Ci values were quantified in this study.

QSPR for hepatic clearance

Hepatic clearance (CL_h) was estimated using the following mechanistic algorithm:

$$CL_h = Ql \cdot \left(\frac{CL_{int} \cdot C_{2E1} \cdot Vl}{Ql + CL_{int} \cdot C_{2E1} \cdot Vl} \right) \qquad [3]$$

where CL_{int} = enzyme content-normalized intrinsic clearance (L/h/μmol CYP2E1), Ql = blood flow to the liver (L/h), C_{2E1} = concentration of CYP2E1 protein in the liver of the species of interest (μmol/L), and Vl = volume of the liver (L).

In order to predict CL_h, the numerical values of species-specific parameters such as Ql, Vl and C_{2E1} as well as chemical-specific CL_{int} were needed. Whereas the Vl and C_{2E1} for rats and humans were obtained from the literature (*8, 9*), a

group contribution QSPR (Eqn 2) was developed for relating CL_{int} (which is essentially the ratio of the enzyme turnover number divided by the substrate affinity) to molecular fragments in 46 VOCs (*3*).

QSPRs development and validation

The group contributions (Ci in Eqn 2) towards Po:a, Pw:a, Pp:a and CL_{int} were determined following multilinear regression analysis of the experimental data with a commercially-available statistical software package (SPSS® for Windows® v10.0.7, SPSS Inc., Chicago, IL).

Statistical significance ($p < 0.05$) was achieved for a fragment contribution when the computed t-statistic was greater than 1.8. All QSPR results were cross-validated by computing the predicted residual sum of squares (PRESS) statistic and the sum of squares of the response values (SSY) using a statistical software (SPSS® for Windows® v10.0.7, SPSS Inc., Chicago, IL). The ratio PRESS/SSY was used as an indication of the validity of the QSPRs. To be a reasonable QSAR or QSPR model, the PRESS/SSY ratio must be smaller than 0.4. A PRESS/SSY value < 0.1 is considered to be indicative of an excellent model (*10*).

For developing QSPRs of Po:a and Pw:a, the experimental data for all 46 VOCs were obtained from Gargas *et al.* (*2*). For QSPR modeling of Pp:a, however, these values were obtained as the difference between experimental Pb values (*2*) and those predicted based on the consideration of lipid and water solubility of individual VOCs (*11-14*). The resulting QSPR results were input to Eqn 1 for calculating species-specific and tissue-specific partition coefficients. For this purpose, knowledge of Fnl, Fw, fb and Fp were required. The tissue-specific and species-specific data on Fnl, Fw and Fp were obtained from published

literature (Table 2). The fb value for blood was set to 1 for the rat and 0.5 for humans, except in case of VOCs possessing a geometric volume > 300 Å. In such cases, fb was considered to be zero (*12*).

For QSPR modeling of $CL_{int,}$ chemical-specific values were obtained by initially dividing experimental data on maximal velocity of metabolism (V_{max}) by the Michaelis-Menten affinity constant (K_m) obtained in rats (*3*), and then dividing the resulting values by the rat liver content of CYP2E1 (rat: 4800 pmol/g liver). The rat liver CYP2E1 level was derived from data on microsomal content of the isozyme-specific protein (approximately 80 pmol CYP2E1/mg microsomal protein) and the data on microsomal protein content of rat liver (approximately 60 mg/g liver) (9, 15). For conducting interspecies extrapolation of hepatic clearance, the species-specific values of Vl, Ql and CYP2E1 levels were specified. Ql in rat and humans was set to 2.11 and 108 L/h, respectively (*6*). Vl was set to 0.01 L and 1.82 L in rats and humans, respectively (*6*). CYP2E1 level was set to 4800 pmol/ml in rat and 2482 pmol/ml in humans (*8, 9, 15*).

The adequacy of QSPR model-derived values of rat and human blood:air partition coefficients, tissue:air partition coefficients and hepatic clearance was accomplished by comparing them with experimental values obtained from the literature (*2, 3, 16-23*).

QSAR-PBPK modeling in rats and humans

The PBPK model used in the present study consisted of four tissue compartments interconnected by systemic circulation through the pulmonary exchange compartment (Figure 1), as described by Ramsey and Andersen (*24*). The algebraic and differential equations constituting the model were identical to

those of Ramsey and Andersen (*24*), with the exception of hepatic metabolism which was calculated as CL_h times arterial blood concentration (C_a) (*25*) (Appendix). The QSPR-PBPK model was written as a program in Advanced Continuous Simulation Language (ACSL®, Aegis Technologies, Huntsville, AL). For conducting simulations using ACSL®, two files are required: (i) a continuous simulation language file (.CSL), which contains the program (i.e., constants, QSPRs for model parameters, differential equations, and integration algorithms) and (ii) a command (.CMD) file, which contains simulation conditions (i.e., exposure concentration, frequency and duration) and other chemical-specific input data (i.e., fragments present in the molecule being simulated). Since the QSPRs were included in the .CSL file, only the frequency of occurrence of the fragments in a molecule needed to be specified along with exposure concentration and duration, to obtain pharmacokinetic simulations. The QSPR-PBPK model codes written in ACSL® can be obtained by contacting the senior author.

The QSPR-PBPK model framework was initially validated by comparing the predicted inhalation pharmacokinetics of two VOCs present in the calibration dataset (toluene and dichloromethane). For each of these chemicals, the exposure concentration and duration were provided as input along with the nature and number of fragments (Table 3) to the QSPR-PBPK model to obtain simulations of their inhalation pharmacokinetics. When extrapolating the kinetics to another species (i.e., from rat to human), the species-specific physiological parameters were replaced with appropriate values (*26, 27*) and then the simulations were conducted. The resulting simulations were compared with previously published experimental data for toluene and dichloromethane obtained in rats and humans (*26, 28, 29*).

In a further validation effort, the QSPR-PBPK model was used to predict the inhalation pharmacokinetics of VOCs which were not part of the calibration set but which could be described using the molecular fragments of chemicals investigated in the present study (Table 1). This set consisted of 1,2,4-trimethylbenzene and ethyl benzene. The pharmacokinetic simulations for these chemicals were obtained solely from knowledge of their molecular structure, exposure conditions (Table 3), and species-specific physiological information. The QSPR-PBPK model simulations for 1,2,4-trimethylbenzene and ethyl benzene were then compared with experimental data on the inhalation pharmacokinetics of these chemicals in rats and humans obtained from the literature (*26, 28, 30, 31*).

5.3 Results

QSPRs for Po:a, Pw:a, Pp:a and CL_{int}

The contributions of each of the 11 molecular fragments to chemical-specific parameters, namely Po:a, Pw:a, Pp:a, and CL_{int}, as obtained from analysis of data for 46 VOCs, are presented in Table 4. For all four parameters, the overall regression was highly positive and significant (R^2 for Po:a, Pw:a, Pp:a, and CL_{int} were 0.994, 0.727, 0.947, and 0.969, respectively). Moreover, the PRESS/SSY statistic of the QSPR model for Po:a, Pw:a, Pp:a, and CL_{int} was 0.01, 0.41, 0.02, and 0.06, respectively. The values of PRESS/SSY for Po:a, Pp:a, and CL_{int} were within the suggested benchmark value (0.1) for excellent QSPR models, whereas the value of PRESS/SSY for Pw:a was close to the suggested benchmark value (0.4) for an adequate QSPR model. It should be noted, however, that for each of the four parameters at least 3 fragment contributions are not statistically significant (t statistic < 1.8), indicating that not all fragments contribute significantly to Pi.

Tissue:air partition coefficients: QSPR-based prediction for the rat and extrapolation to humans

The QSPRs for Po:a, Pw:a and Pp:a were then combined with the tissue composition data for computing rat tissue:air partition coefficients. The correlation between rat experimental values and QSPR predictions of tissue:air partition coefficients is depicted in Figure 2. The mean (±SD) [range] of the experimental to predicted ratios in rats was 1.4±1.6 [0.4-11] for liver:air partition coefficient (Pl), 1.6±1.3 [0.4-7.5] for muscle:air partition coefficient (Pm), and 1.19±0.6 [0.3-3.2] for fat:air partition coefficient (Pf). Overall, of the 138 experimental Pl, Pm or Pf

values, 112 were within a factor of two of the QSPR-based predictions, whereas 22 predictions were within a factor of 2-3 of the experimental values. The remaining cases were: Pf of isoprene (3.2), Pl of allyl chloride (11) and Pm of both isoprene (5.6) and allyl chloride (7.5).

Using the QSPRs for Po:a, Pw:a, and Pp:a, and by replacing the tissue composition for the rat with those of humans in Eqn 1, rat-to-human extrapolations of tissue:air partition coefficients were accomplished. The correlation between experimental values and QSPR-predictions of human tissue:air partition coefficients is presented in Figure 3. The mean (±SD) [range] of the predicted to experimental ratios in humans was 1.8±1.3 [0.3-6] for Pl, 2.1±1.3 [0.6-6] for Pm, and 1.2±0.6 [0.21-2.3] for Pf. Overall, 42 predicted values of Pl, Ps, or Pf were within a factor of two of the experimental values, whereas 18 predictions were within a factor of 2-3 of the experimental values. The outliers were: 1,3-butadiene (4.4), chlorobenzene (4), and n-pentane (6) for Pm, carbon tetrachloride (0.2) for Pf, as well as 1,3-butadiene (6) and n-heptane (3.6) for Pl.

Blood:air partition coefficient: QSPR-based estimation for the rat and extrapolation to humans

The correlation between rat and human experimental and estimated values of blood:air partition coefficients are presented in Figure 4. The mean (±SD) [range] of the experimental to predicted ratios was 0.87±0.44 [0.21-1.88] in humans and 1.10±0.53 [0.24-2.69] in rats. For Pb values in rats and humans, of the 78 experimental values, 68 were within a factor of two of the biologically-based-estimations, whereas 10 estimates were within a factor of 2-3 of the experimental values. The remaining cases were: carbon tetrachloride in both rat

(0.24) and humans (0.21), as well as chloroform (0.31) and vinyl bromide (0.32) in humans.

Hepatic clearance: QSPR-based estimation for the rat and extrapolation to humans

For CL_h, the mean (±SD) [range] of the experimental to predicted ratios was 1.01±0.22 [0.61-1.67] and 1.00±0.14 [0.73-1.39] in rats and humans, respectively, and all experimental values were within a factor of 1.5 of the QSPR estimations except for chloroform (1.67) and *cis*-1,2-dichloroethylene (0.61) in rats. These results suggest that both partition coefficients and hepatic clearance of VOCs belonging to several chemical classes can be adequately described and extrapolated from one species to another using mechanistic QSPRs.

QSPR-PBPK modeling

The mechanistic QSPRs described above were incorporated into a multispecies PBPK modeling framework that enabled rat-to-human extrapolation of the inhalation pharmacokinetics of VOCs. By interchanging the species-specific parameters and keeping all other parameters constant, the PBPK model was used to simulate the kinetics of toluene and dichloromethane initially in the rat and then in humans. These VOCs were in the calibration set for QSPR development for partition coefficients and hepatic clearance. The integrated QSPR-PBPK model adequately simulated the kinetics in rats exposed for 4 hr to 100 ppm toluene and 100 ppm dichloromethane, as well as in human volunteers exposed for 6 hr to 50 ppm dichloromethane and for 7 hr to 17 ppm toluene (Figure 5).

The QSPR-PBPK model was then used to simulate the inhalation pharmacokinetics of two other VOCs that were not part of the original calibration set but which consisted of fragments elucidated in the present study (1,2,4-trimethyl benzene [1*AC+3*H(on AC)+3*CH3] and ethyl benzene [1*AC+5*H(on AC)+CH3+CH2]). Accordingly, using chemical structure information for these VOCs, their partition coefficients and hepatic clearance were calculated with QSPRs the results of which were then used along with physiological parameters in PBPK equations to provide simulations of pharmacokinetics. The resulting simulations of the kinetics in rats exposed to 100 ppm trimethylbenzene for 4 hr and to 50 ppm ethylbenzene for 4 hr were generally within a factor of 2 of the corresponding experimental data obtained from the literature (Figure 6). The same QSPR-PBPK was then used to conduct rat-to-human extrapolation of the inhalation pharmacokinetics of trimethylbenzene and ethylbenzene. The resulting simulations compared well with experimental data obtained in volunteers exposed for 2 hr to 2 ppm trimethylbenzene and for 7 hr to 33 ppm ethylbenzene (Figure 6).

5.4 Discussion

PBPK models are increasingly being used by risk assessors for conducting interspecies extrapolation of target tissue dose. The parameters required for constructing these models in a single species as well as for using them for extrapolation purposes are not always available. This hurdle can be overcome by combining QSAR approaches with physiological data to derive input parameters for PBPK models. With such an approach, two areas of data need can be addressed: (1) inter-chemical differences in input parameters for a single species, and (2) interspecies differences in input parameter values for a given chemical. If extrapolation of pharmacokinetics across species and across chemicals is to be accomplished, these two types of differences must be simultaneously accounted for. PBPK models provide a unique framework that can account for part of the interspecies differences in physiology, particularly tissue blood flows and volumes. However, inter-chemical differences must still be accounted for in the form of physicochemical and biochemical parameters. There are several ways of generating these data. First, *in vitro* or *in vivo* studies can be conducted to generate the various partition coefficients and metabolic rate of chemicals of interest in both the test species and humans. While such data are desirable, they may not always be available in the literature and often significant effort and resources would be required to generate these data for a series of new chemicals. Second, *in silico* QSARs for each input parameter for PBPK modeling in each species of interest can be developed. While the development of such models with available data sets is less labor intensive, they are specific to the species in which such data were collected. The feasibility of developing species-specific QSPRs

depends largely on the availability of such data in a particular species. Furthermore, to move from one species to another, new sets of data and analyses would be required so that species-specific QSPRs can be developed anew. Another promising approach, developed in this study, involves the development of mechanistic algorithms that are based on chemical-specific QSPRs and species-specific characteristics. In this approach, extrapolation across chemicals can be accomplished simply by varying chemical-specific information, while extrapolation across species is accomplished by varying the relevant biological information contained in the QSPRs.

The tissue:air and blood:air partition coefficients are key input parameters for PBPK models. Existing algorithms require data on Po:w and Pw:a derived from two different QSPR approaches. For example, to estimate Pow of dichloromethane according to the program KOWWIN®, the structural fragments CH2 (n=1), CL (n=2) as well as an equation constant (n=1) are used. However, for predicting Pw:a of chemicals according to HENRYWIN®, dichloromethane is described as an intact fragment (CH2-CL2). This example emphasizes the fact that the currently available QSPRs do not allow the use of a single set of descriptors to predict all partition coefficients required for PBPK modeling. Furthermore, the HENRYWIN® program provides estimates of Pw:a for 25°C, which should then be extrapolated to physiological temperature (37°C), before it can be used for calculating tissue:air and blood:air partition coefficients. In this study, these obstacles were overcome by developing QSPRs that use a single set of fragments to model Po:a and Pw:a at 37°C.

An alternative approach for relating molecular structure information and PBPK model parameters would be to the Free-Wilson approach (*33*). Even though

the Free-Wilson algorithms could have been explored for modeling Pw:a and Po:w in the present study, their applicability is limited to one family of substances (i.e., chloroethanes) and species. The group contribution approach (*5, 7*) used in the present study is different from the Free-Wilson approach in that it does not require a common structure among the chemicals investigated. Whereas the common "basic" structure is easily identifiable while studying chemicals belonging to the same family (e.g., the two carbon backbone in the chloroethanes (*33*)), it does become a hurdle while modeling multiple groups of chemicals such as chloromethanes, ethanes, ethylenes, and aromatics. In such cases, the establishment of a chemical class-specific algorithm is certainly possible, but the tradeoff becomes one of statistical power, since the number of chemicals in each family for which PBPK model parameters are available is considerably low. For example, the chloroethylene family consists of a maximum of 6 members. The Gao's group contribution approach, used in the present study, does not require the existence of a common basic structure among the chemicals and facilitates the identification of independent variables (i.e., number of fragments or groups in a molecule) based on visual inspection of a compound's structural formula (*5*). Isomers or diastereomers, that cannot be adequately modeled using structural fragments, are described with an independent variable taking a value of 0 or 1 depending on the spatial positioning of the group. In order to have statistical meaning, however, many isomers of the same type (e.g., cis-trans or meta-ortho-para isomers) must be present in the dataset. The chemical set used in this study contained only one *cis-trans* isomer (1,2-dichloroethylene) and position isomer (xylene), preventing the use of position corrections in the regression equations. While the Po:a of *o*-, *m*-, and *p*-xylene are comparable (3534, 3245, and 3319,

respectively), this is not the case for *cis*- and *trans*-1,2-dichloroethylene (278 and 178, respectively), highlighting the importance of incorporating such position data into the QSPR, when data are available.

The present study has demonstrated the feasibility of the development of QSPRs for pharmacokinetic determinants of VOCs belonging to multiple families and their integration within PBPK modeling framework to simulate the pharmacokinetics of new chemicals in rats and humans. Reasonably good agreement between simulated and experimental pharmacokinetics of VOCs was observed in the present study. Even though the overall success of the QSPR-PBPK model does not necessarily imply that individual chemical-specific parameters are accurately predicted by the QSPR equations or that species-specific PBPK model parameters are accurately predicted by the biologically-based equations, it does indicate that the net impact of these critical determinants of blood concentration (as computed from QSPR-derived Po:a, Pw:a, Pp:a and CL_{int}) are similar to that of experimentally-derived PBPK model parameters. It is clear from the results that, for most of the chemicals in the calibration set, the experimental and estimated values of individual PBPK model parameters compared favorably.

Results of the present study show that it is possible to describe several PBPK model parameters (e.g., blood:air partition coefficients, tissue:air partition coefficients and hepatic clearance) in two species using four QSPRs for chemical-specific parameters (Pw:a, Po:a, Pp:a and CL_{int}). Not all fragment contribution values in Table 4 are statistically significant (t value > 1.8). This may be a consequence of either the fact that some of these fragments do not make a statistically significant contribution to the chemical-specific parameter value, or that

there is just not enough data to facilitate the determination of a contribution value of statistical significance, for some fragments. It should also be noted that no judgment of data quality was done, even though all data used in generating the QSPRs were obtained using the same experimental protocol. Generally, a high degree of correlation existed between the experimental and QSPR-estimated values of chemical parameters. As information regarding partition coefficients of diverse chemicals becomes available in the literature, these can be integrated to the existing database.

The tissue:air partition coefficients in general could be adequately predicted with knowledge of VOC solubility in tissue lipids and water (*34*). This, however, is not adequate in the case of blood for which additional partitioning of VOCs into proteins should be accounted for (*11, 12*). Recent in depth studies on blood protein binding of benzene in rats and humans have suggested the role of a species-specific component, even though it has not yet been characterized (*35*). This process is represented in Eqn. 2 with parameters reflecting the capacity of binding proteins (fb) as well as the volume fraction of blood proteins present in blood (Fp). The species-specific fraction of binding (fb*Fp) was derived semi-empirically as a combination of the fraction of blood protein present in blood as well as the relative binding capacities in rats and humans. While inter-chemical differences in binding would be expected to result from difference in binding affinities (reflected in this study by fb), interspecies differences would be expected to result from difference in binding protein concentration (Fp) or number of binding sites on the protein (fb). Recent studies have demonstrated that interspecies difference in blood protein concentrations cannot alone account for interspecies differences in Pb of some VOCs (*35*). The number of binding sites on rat

hemoglobin has been experimentally determined as 2 for chloroform (a representative chemical of the VOCs studied) (*13*). No such evaluation of the number of binding sites in human blood has been reported.

Of all the PBPK parameters, the metabolic constants represent the most difficult to model. Initial regression studies were conducted using maximal velocity for metabolism (V_{max}) and the Michaelis affinity constant (K_m) of the VOCs independently. Although both V_{max} and K_m could be modeled adequately (R^2 > 0.80), the QSPR for K_m failed the cross-validation process (PRESS/SSY = 0.7). Further efforts were then undertaken to develop QSPRs of intrinsic clearance ($CL_{int} = V_{max}/K_m$), normalized for CYP2E1 content. Even though the QSPR for CL_{int} represented a significant improvement over the V_{max} and K_m models, the adequacy of this model fit was significantly influenced by the CL_{int} value of isoprene that was outside the range of experimental values of the other chemicals. Modeling of the CL_h was then attempted. The use of CL_h in PBPK models allows the evaluation of the relative roles of blood flow and CL_{int} in the hepatic metabolism of chemicals. Even though both approaches give identical results (*25*) for first order conditions, the use of CL_h permits a more direct evaluation of the impact of hepatic blood flow limitation on the amount metabolized (*36, 37*). The present study successfully implemented a biologically-based algorithm for CL_h that uses a QSPR for CL_{int} along with species-specific information on enzyme content (e.g., CYP2E1). The CL_h, as derived and modeled in the present study, is only appropriate for conducting simulations of the pharmacokinetics of chemicals at low exposure concentrations. At high exposure concentrations when saturable mechanisms come into play, knowledge of V_{max} and K_m is essential to be able to adequately simulate the metabolism rate and pharmacokinetics of chemicals in biota.

Therefore, more complex QSPR approaches may have to be explored for modeling V_{max} and K_m. Overall, the use of the fragment-based QSPR approach, developed in the present study, should be limited to low molecular weight VOCs (i) metabolized by P-450 2E1 and (ii) that possess fragments (number and nature) that are similar to those investigated in the present study. Extrapolation to predict the kinetics of chemicals with fragments other than those investigated in this study, or larger number of certain fragments (e.g., 4 F, 5 CH_3) than those investigated in this study is not appropriate. With increasing chain length and molecular weight, it is likely that more complex approaches may be necessary to account for non-linearity and qualitative changes in mechanisms (presence vs absence of protein binding, saturation of liposolubility characteristics, differences in isoenzymes involved in metabolism). In theory, the only limit to the development of such algorithms is the knowledge of the mechanistic processes behind the endpoint of interest and the relative importance of chemical- and species- specific information forming the basis of QSPRs.

In conclusion, the present study has demonstrated that it is possible to extrapolate PBPK model parameters from rats to humans using QSPRs in mechanistic algorithms. These parameter estimates, once integrated with species-specific physiological parameters, allow *a priori* prediction of the inhalation pharmacokinetics of VOCs, at exposure concentrations below saturation. The QSPRs developed in this study should be potentially useful for providing "first-cut" evaluation of the inhalation pharmacokinetics of VOCs in multiple species, prior to experimentation, as long as the number and nature of the fragments do not exceed the ones in the calibration dataset used in this study.

5.5 Appendix

The following algebraic and mass balance equations based on Ramsey and Andersen (*24*), were used to calculate the tissue and blood concentrations of chemicals. The PBPK model was written as a computer program containing the equations for arterial blood, venous blood, liver, muscle, fat, and richly perfused tissues. In these equations, * denotes multiplication.

Arterial blood

The concentration of chemical in arterial blood (C_a) was calculated using the following expression, based on the mass balance equation for lungs which specifies that loss of chemical from the air is balanced by an identical gain in arterial blood:

$$C_a = \frac{Q_p * C_i + Q_c * C_v}{Q_c + Q_p / P_b} \qquad [4]$$

where Q_c = cardiac output, C_v = venous blood concentration, Q_p = alveolar ventilation rate, C_i = inhaled chemical concentration, and P_b = blood:air partition coefficient.

Muscle, fat, and richly perfused tissues

The rate of change in the amount of chemical in non-metabolizing tissues (dA_t/dt) (t = adipose, muscle, and richly perfused tissues) was described with a mass balance equation based on tissue blood flow (Q_t) and the arterio-venous concentration difference (C_a-C_{vt}):

$$\frac{\partial A_t}{\partial t} = Q_t * (C_a - C_{vt}) \qquad [5]$$

The amount of chemical in tissue (A_t), concentration in tissue (C_t) and concentration in venous blood leaving the tissue (C_{vt}) were calculated as follows:

$$A_t = \int_0^t \frac{\partial A_t}{\partial t} \quad [6]$$

$$C_t = \frac{A_t}{V_t} \quad [7]$$

$$C_{vt} = \frac{C_t}{P_t} \quad [8]$$

where P_t = tissue:blood partition coefficient, and V_t = volume of tissue t.

Liver

The rate of change in the amount of chemical in liver (dA_l/dt) was described using a mass balance equation based on liver blood flow rate (Q_l), the arterio-venous concentration difference (C_a-C_{vl}) and the rate of metabolism (dA_{met}/dt) calculated as hepatic clearance (CL_h) times arterial blood concentration:

$$\frac{\partial A_l}{\partial t} = Q_l * (C_a - C_{vl}) - \frac{\partial A_{met}}{\partial t} \quad [9]$$

$$\frac{\partial A_{met}}{\partial t} = CL_h * C_a \quad [10]$$

The amount of chemical in liver (A_l), concentration in liver (C_l), and the concentration in venous blood leaving the liver (C_{vl}) were calculated per Eqns. [4]-[6] using liver:blood partition coefficient (P_l) and liver volume (V_l).

Venous blood

The concentration of a chemical in mixed venous blood was calculated by accounting for the venous contributions of all tissue compartments:

$$C_v = \frac{\sum_{i=1}^{3} C_{vti} * Q_{ti} + C_{vl} * Q_l}{Q_c} \qquad [11]$$

where t = adipose, muscle, and richly perfused tissues, respectively.

5.6 References

1. Beliveau, M., Tardif, R., and Krishnan, K. (2003) Quantitative structure-property relationships for physiologically-based pharmacokinetic modeling of volatile organic chemicals in rats. *Toxicol. Appl. Pharmacol. 189*, 221-232.

2. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H., and Andersen, M. E. (1989) Partition coefficients of low-molecular-weight volatile chemicals in various liquids and tissues. *Toxicol. Appl. Pharmacol. 98*, 87-99.

3. Gargas, M. L., Seybold, P. G., and Andersen, M. E. (1988) Modeling the tissue solubilities and metabolic rate constant (Vmax) of halogenated methanes, ethanes, and ethylenes. *Toxicol. Lett.. 43*, 235-256.

4. Beliveau, M., Tardif, R., and Krishnan, K. (2003) Quantitative structure-property relationships for physiologically based pharmacokinetic modeling of volatile organic chemicals in rats. *Toxicol Appl Pharmacol. 189*, 221-32.

5. Martin, T. M. and Young, D. M. (2001) Prediction of the acute toxicity (96-h LC_{50}) or organic compounds to the fathead minnow (*Pimephales promelas*) using a group contribution method. *Chem. Res. Toxicol. 14*, 1378-1385.

6. Krishnan, K. and Andersen, M. E. (2001) Physiologically based pharmacokinetic modeling in toxicology. In *Principles and methods of toxicology* (Hayes, A. W., ed) pp. 193-241, Taylor & Francis, Philadelphia.

7. Gao, C., Govind, R., and Tabak, H. H. (1992) Application of the group contribution method for predicting the toxicity of organic chemicals. *Environ. Toxicol. Chem.. 11*, 631-636.

8. Lipscomb, J. C., Teuschler, L. K., Swartout, J. C., Striley, C. A. F., and Snawder, J. E. (2003) Variance of Microsomal Protein and Cytochrome P450 2E1 and 3A Forms in Adult Human Liver. *Toxicol. Mech. Methods. 13*, 45-51.

9. Davis, R. R., Murphy, W. J., Snawder, J. E., Striley, C. A. F., Henderson, D., Khan, A., and Krieg, E. F. (2002) Susceptibility to the ototoxic properties of toluene is species specific. *Hearing Res. 166*, 24-32.

10. Wold, S. (1991) Validation of QSAR's. *QSAR. 10*, 191-193.

11. Poulin, P. and Krishnan, K. (1996) A mechanistic algorithm for predicting blood:air partition coefficients of organic chemicals with the consideration of reversible binding in hemoglobin. *Toxicol. Appl. Pharmacol. 136*, 131-137.

12. Poulin, P., Beliveau, M., and Krishnan, K. (1999) Mechanistic animal-replacement approaches for predicting pharmacokinetics of organic chemicals. In *Toxicity assessment alternatives: methods, issues,opportunities.* (Salem, H.andKatz, S. A., eds) pp. 115-139, Humana Press Inc., Totowa, NJ.

13. Beliveau, M. and Krishnan, K. (2000) Concentration dependency of rat blood: air partition coefficients of some volatile organic chemicals. *J.Toxicol Environ.Health A. 60*, 377-389.

14. Beliveau, M. and Krishnan, K. (2000) Estimation of rat blood:air partition coefficients of volatile organic chemicals using reconstituted mixtures of blood components. *Toxicol. Lett. 116*, 183-188.

15. Carlile, D. J., Zomorodi, K., and Houston, J. B. (1997) Scaling factors to relate drug metabolic clearance in hepatic microsomes, isolated hepatocytes, and the intact liver: studies with induced livers involving diazepam. *Drug Metab. Dispos. 25*, 903-911.

16. Paterson, S. and Mackay, D. (1989) Correlation of tissue, blood, and air partition coefficients of volatile organic chemicals. *Br. J. Ind. Med. 46*, 321-328.

17. Fiserova-Bergerova, V. and Diaz, M. L. (1986) Determination and prediction of tissue-gas partition coefficients. *Int. Arch. Occ. Environ. Health. 58*, 75-87.

18. Sato, A. and Nakajima, T. (1979) Partition coefficients of some aromatic hydrocarbons and ketones in water, blood and oil. *Br. J. Ind. Med. 36*, 231-234.

19. Meulenberg, C. J. and Vijverberg, H. P. (2000) Empirical relations predicting human and rat tissue:air partition coefficients of volatile organic compounds. *Toxicol. Appl. Pharmacol. 165*, 206-216.

20. Filser, J. G., Johanson, G., Kessler, W., Kreuzer, P. E., Stei, P., Baur, C., and Csanady, G. A. (1993) A pharmacokinetic model to describe toxicokinetic interactions between 1,3-butadiene and styrene in rats: predictions for human exposure. *IARC Scientific Publications. 127*, 65-78.

21. Laparé, S., Tardif, R., and Brodeur, J. (1993) Effect of various exposure scenarios on the biological monitoring of organic solvents. I. Toluene and xylene. *Int. Arch. Occ. Environ. Health. 64*, 569-580.

22. Gearhart, J. M., Mahle, D. A., Greene, R. J., Seckel, C. S., Flemming, C. D., Fisher, J. W., and Clewell, H. J. I. (1993) Variability of physiologically based pharmacokinetic (PBPK) model parameters and their effects on PBPK model predictions in risk assessment for perchloroethylene (PCE). *Toxicol. Lett. 68*, 131-144.

23. Fisher, J. W., Mahle, D. A., Bankston, L., Greene, R. J., and Gearhart, J. M. (1997) Lactational transfer of volatile chemicals in breast milk. *Am. Ind. Hyg. Assoc. J. 58*, 425-431.

24. Ramsey, J. C. and Andersen, M. E. (1984) A physiologically-based description of the inhalation pharmacokinetics of styrene in rats and humans. *Toxicol. Appl. Pharmacol. 73*, 159-175.

25. Poulin, P. and Krishnan, K. (1999) Molecular structure-based prediction of the toxicokinetics of inhaled vapors in humans. *Int. J. Toxicol. 18*, 7-18.

26. Tardif, R., Charest-Tardif, G., Brodeur, J., and Krishnan, K. (1997) Physiologically based pharmacokinetic modeling of a ternary mixture of alkyl benzenes in rats and humans. *Toxicol. Appl. Pharmacol. 144*, 120-134.

27. Astrand, I. (1983) Effect of physical exercise on uptake, distribution and elimination of vapors in man. In *Modeling of inhalation exposure to vapors: uptake,*

distribution, and elimination (Fiserova-Bergerova, V., ed) pp. 107-130, CRC Press Inc., Boca Raton, FL.

28. Haddad, S., Charest-Tardif, G., Tardif, R., and Krishnan, K. (2000) Validation of a physiological modeling framework for simulating the toxicokinetics of chemicals in mixtures. *Toxicol. Appl. Pharmacol. 167*, 199-209.

29. Andersen, M. E., Clewell, H. J. I., and Gargas, M. L. (1991) Physiologically-based pharmacokinetic modeling with dichloromethane, its metabolite carbon monoxide and blood carboxyhemoglobin in rats and humans. *Toxicol. Appl. Pharmacol. 108*, 14-27.

30. Swiercz, R., Rydzynski, K., Wasowicz, W., Majcherek, W., and Wesolowski, W. (2002) Toxicokinetics and metabolism of pseudocumene (1,2,4-trimethylbenzene) after inhalation exposure in rats. *Int. J. Occ. Med. Environ. Health. 15*, 37-42.

31. Jarnberg, J., Johanson, G., Lof, A., and Stahlbom, B. (1997) Inhalation toxicokinetics of 1,2,4-trimethylbenzene in volunteers: comparison between exposure to white spirit and 1,2,4-trimethylbenzene alone. *Sci. Total Environ. 199*, 65-71.

32. Poulin, P. and Krishnan, K. (1996) Molecular structure-based prediction of the partition coefficients of organic chemicals for physiological pharmacokinetic models. *Toxicol. Methods. 6*, 117-137.

33. Fouchécourt, M.-O. and Krishnan, K. (2000) A QSAR-type PBPK model for inhaled chloroethanes. *Toxicol. Sci. 54*, 88.

34. Poulin, P. and Krishnan, K. (1996) A tissue composition-based algorithm for predicting tissue:air partition coefficients of organic chemicals. *Toxicol. Appl. Pharmacol. 136*, 126-130.

35. Wiester, M. J., Winsett, D. W., Richards, J. H., Doerfler, D. L., and Costa, D. L. (2002) Partitioning of benzene in blood: influence of hemoglobin type in humans and animals. *Environ. Health Persp. 110*, 255-261.

36. Johanson, G. and Naslund, P. H. (1988) Spreadsheet programming: A new approach in physiologically based modeling of solvent toxicokinetics. *Toxicol. Lett. 41*, 115-127.

37. Kedderis, G. L. (1997) Extrapolation of in vitro enzyme induction data to humans in vivo. *Chem.-Biol. Inter. 107*, 109-121.

Table 1 : Frequency of occurrence of molecular fragments in the VOCs investigated in the present study.

Chemicals	CH_3	CH_2	CH	C	C=C	H	Br	Cl	F	AC	H_AC
Halomethanes											
Chloromethane	1	0	0	0	0	0	0	1	0	0	0
Dichloromethane	0	1	0	0	0	0	0	2	0	0	0
Chloroform	0	0	1	0	0	0	0	3	0	0	0
Carbon tetrachloride	0	0	0	1	0	0	0	4	0	0	0
Difluoromethane	0	1	0	0	0	0	0	0	2	0	0
Fluorochloromethane	0	1	0	0	0	0	0	1	1	0	0
Bromochloromethane	0	1	0	0	0	0	1	1	0	0	0
Dibromomethane	0	1	0	0	0	0	2	0	0	0	0
Chlorodibromomethane	0	0	1	0	0	0	2	1	0	0	0
Haloethanes											
Chloroethane	1	1	0	0	0	0	0	1	0	0	0
1,1-Dichloroethane	1	0	1	0	0	0	0	2	0	0	0
1,2-Dichloroethane	0	2	0	0	0	0	0	2	0	0	0
1,1,1-Trichloroethane	1	0	0	1	0	0	0	3	0	0	0
1,1,2-Trichloroethane	0	1	1	0	0	0	0	3	0	0	0
1,1,1,2-Tetrachloroethane	0	1	0	1	0	0	0	4	0	0	0
1,1,2,2-Tetrachloroethane	0	0	2	0	0	0	0	4	0	0	0
Pentachloroethane	0	0	1	1	0	0	0	5	0	0	0
Hexachloroethane	0	0	0	2	0	0	0	6	0	0	0

1,2-Dibromoethane	0	2	0	0	0	0	2	0	0	0	0
1-Bromo-2-chloroethane	0	2	0	0	0	0	1	1	0	0	0
1,1,1-Trifluoro-2-chloroethane	0	1	0	1	0	0	0	1	3	0	0
1,1,1-Trifluoro-2-bromo-2-chloroethane	0	0	1	1	0	0	1	1	3	0	0
Alkanes											
1-Chloropropane	1	2	0	0	0	0	0	1	0	0	0
2-Chloropropane	2	0	1	0	0	0	0	1	0	0	0
1,2-Dichloropropane	1	1	1	0	0	0	0	2	0	0	0
n-Propyl bromide	1	2	0	0	0	0	1	0	0	0	0
isopropylbromide	2	0	1	0	0	0	1	0	0	0	0
n-Hexane	2	4	0	0	0	0	0	0	0	0	0
n-Heptane	2	5	0	0	0	0	0	0	0	0	0
Cyclohexane	0	6	0	0	0	0	0	0	0	0	0
2,3,4-Trimethyl pentane	5	0	3	0	0	0	0	0	0	0	0
2,2,4-Trimethyl pentane	5	1	1	1	0	0	0	0	0	0	0
Alkenes											
Vinyl chloride	0	0	0	0	1	3	0	1	0	0	0
1,1-Dichloroethylene	0	0	0	0	1	2	0	2	0	0	0
cis-1,2,-Dichloroethylene	0	0	0	0	1	2	0	2	0	0	0
Trichloroethylene	0	0	0	0	1	1	0	3	0	0	0
Tetrachoroethylene	0	0	0	0	1	0	0	4	0	0	0

Vinyl bromide	0	0	0	0	1	3	1	0	0	0	0
Allyl chloride	0	1	0	0	1	3	0	1	0	0	0
Isoprene	1	0	0	0	2	5	0	0	0	0	0
Aromatic hydrocarbons											
Benzene	0	0	0	0	0	0	0	0	0	1	6
Chlorobenzene	0	0	0	0	0	0	0	1	0	1	5
Toluene	1	0	0	0	0	0	0	0	0	1	5
Styrene	0	0	0	0	1	3	0	0	0	1	5
m-Methylstyrene	1	0	0	0	1	3	0	0	0	1	4
m-Xylene	2	0	0	0	0	0	0	0	0	1	4
n	30	36	15	9	11	25	12	64	9	6	29

Table 2: Composition of rat and human tissues and blood (32).

Tissue and species	Fnl[a]	Fw[a]	Fp[a]
Fat			
Rat	0.8536	0.1215	-
Human	0.7986	0.1514	-
Liver			
Rat	0.0425	0.7176	-
Human0	0.0473	0.7400	-
Muscle			
Rat	0.0117	0.7471	-
Human	0.0378	0.7573	-
Blood			
Rat	0.002	0.8423	0.156
Human	0.004	0.8217	0.1743

[a]. Fnl = neutral lipid equivalent contained in tissues and blood (expressed as volume fraction of biological matrix), and Fw = water equivalent content of tissues and blood (as volume fraction of the biological matrix), Fp = protein content blood.

Table 3: Molecular fragments and exposure characteristics of chemicals used for the validation of the QSPR-PBPK model.

Chemical	Fragment	Frequency of occurence	Rats		Humans		Reference
			Exposure conc. (ppm)	Exposure duration (hr)	Exposure conc. (ppm)	Exposure duration (hr)	
Dichloromethane	CH_2	1	100	4	100	6	(*28, 29*)
	Cl	2					
Toluene	CH_3	1	50	4	17	7	(*26, 28*)
	AC	1					
	H_AC	5					
Ethyl benzene	CH_3	1	50	4	33	7	(*26, 28*)
	CH_2	1					
	AC	1					
	H_AC	5					
1,2,4-Trimethylbenzene	AC	1	100	6	2	2 (at 50 W)	(*30, 31*)
	H_AC	3					
	CH_3	3					

Table 4: Fragment-specific contributions[a] towards oil:air partition coefficients (Po:a), water:air partition coefficient (Pw:a), protein:air partition coefficient (Pp:a), as well as intrinsic clearance (CL_{int} [L/hr/µmol CYP2E1]).

Fragments	Log Po:a	t	Log Pw:a	t	Log Pp:a	t	Log CL_{int}	t
CH3	0.354	7.33	-3.76E-2	0.396	0.306	5.51	1.552	7.83
CH2	0.441	14.9	-.223	3.83	0.182	5.40	0.514	2.28
CH	0.377	4.13	-0.477	2.65	-0.111	1.06	7.83E-2	0.224
C	-0.354	2.68	-1.49	5.74	-1.06	6.94	-0.871	1.57
C=C	0.197	0.867	-1.94	4.32	-0.877	3.36	0.591	0.654
H (on =)	0.134	1.57	0.555	3.30	0.492	5.03	0.383	1.13
Br	1.174	20.8	0.622	5.58	1.15	18.5	1.00	4.64
Cl	0.776	21.3	0.468	6.51	0.764	18.2	0.522	3.43
F	0.136	2.37	0.229	2.03	0.241	3.67	-	-
AC	3.729	5.65	0.650	0.500	1.97	2.61	-7.646	4.12
H (on AC)	-0.190	1.44	-6.24E-2	0.240	-2.81E-2	0.186	1.535	4.65
r^2	0.994		0.727		0.984		0.969	
PRESS/SSY	0.01		0.41		0.02		0.06	

[a] Contributions were obtained by multiple linear regression of experimental data . t-statistic > 1.8 indicates statistical significance ($p < 0.05$).

Figure legends

Figure 1: QSPR-PBPK modeling framework presented in this study.

Figure 2: Comparison of experimental versus QSPR-predicted tissue:air partition coefficients (PC) in rats. A) Log fat:air PC, B) Log liver:air PC, C) Log muscle:air PC.

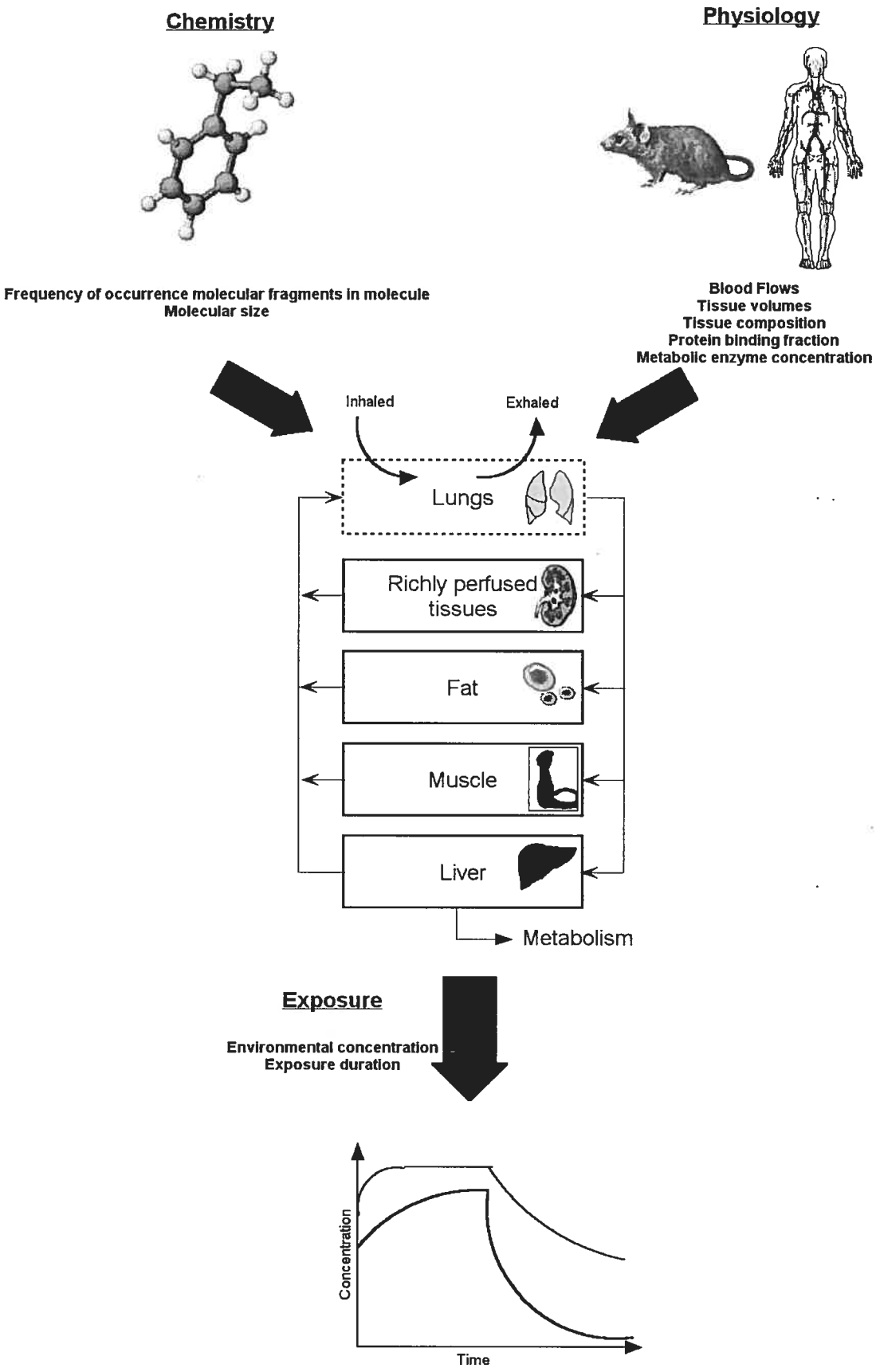
Figure 3: Comparison of experimental versus QSPR-predicted tissue:air partition coefficients (PC) in humans. A) Log fat:air PC, B) Log liver:air PC, C) Log muscle:air PC.

Figure 4: Comparison of experimental versus QSPR-derived blood:air partition coefficients (PC). A) in rats, B) in humans.

Figure 5: Comparison of QSPR-PBPK model simulations (lines) with experimental data (*26, 28, 29*) (symbols) on blood concentrations following inhalation exposure to (A) 100 ppm dichloromethane (4 hr) in rats, (B) 50 ppm toluene (4 hr) in rats, (C) 100 ppm dichloromethane (6 hr) in humans, and (D) 17 ppm toluene (7 hr) in humans..

Figure 6: Comparison of QSPR-PBPK model simulations (lines) with experimental data (*26, 28, 30, 31*) (symbols) on blood concentrations following inhalation exposure to (A) 100 ppm 1,2,4-trimethyl benzene (6 hr) in rats, (B) 50 ppm ethylbenzene (4 hr) in rats, (C) 2 ppm 1,2,4-trimethyl benzene (2 hr, 50 W workload) in humans and (D) 33 ppm ethylbenzene (7 hr) in humans.

Figure 1


Chemistry
Physiology
Frequency of occurrence molecular fragments in molecule
Molecular size
Blood Flows
Tissue volumes
Tissue composition
Protein binding fraction
Metabolic enzyme concentration
Inhaled
Exhaled
Lungs
Richly perfused tissues
Fat
Muscle
Liver
Metabolism
Exposure
Environmental concentration
Exposure duration
Concentration
Time

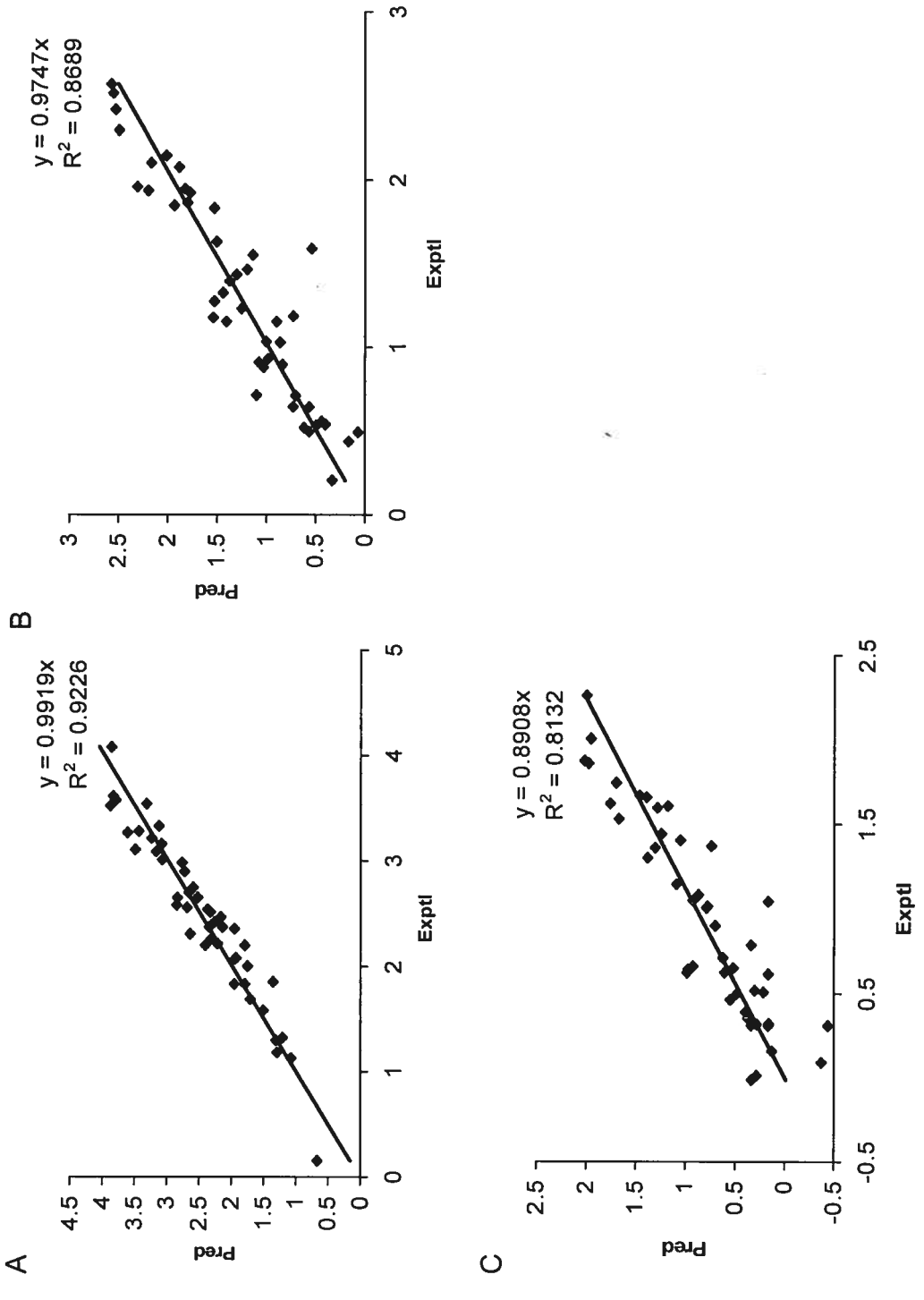

A
y = 0.9919x
R2 = 0.9226
Pred
Exptl
B
y = 0.9747x
R2 = 0.8689
Pred
Exptl
C
y = 0.8908x
R2 = 0.8132
Pred
Exptl

Figure 2

Figure 3

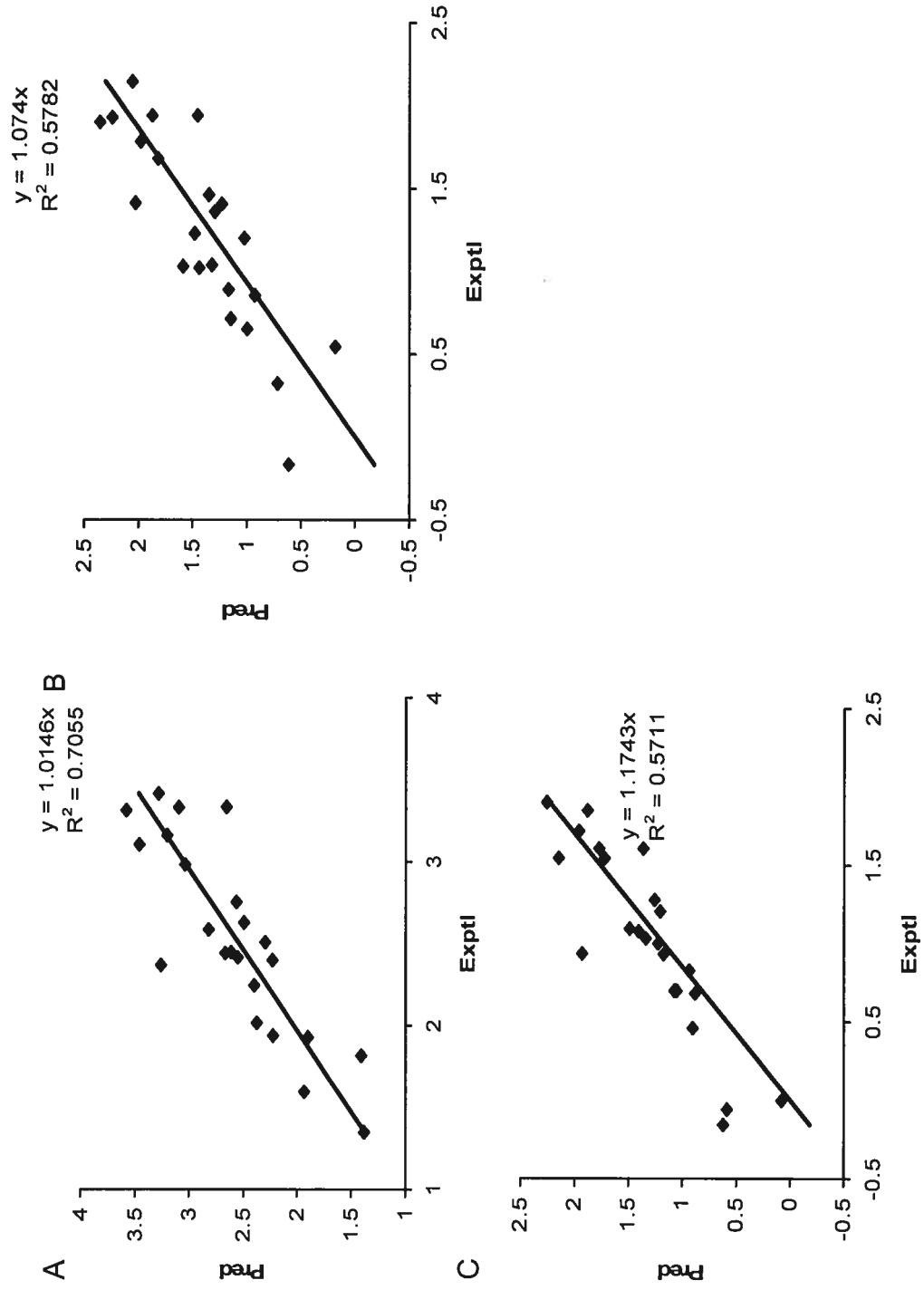
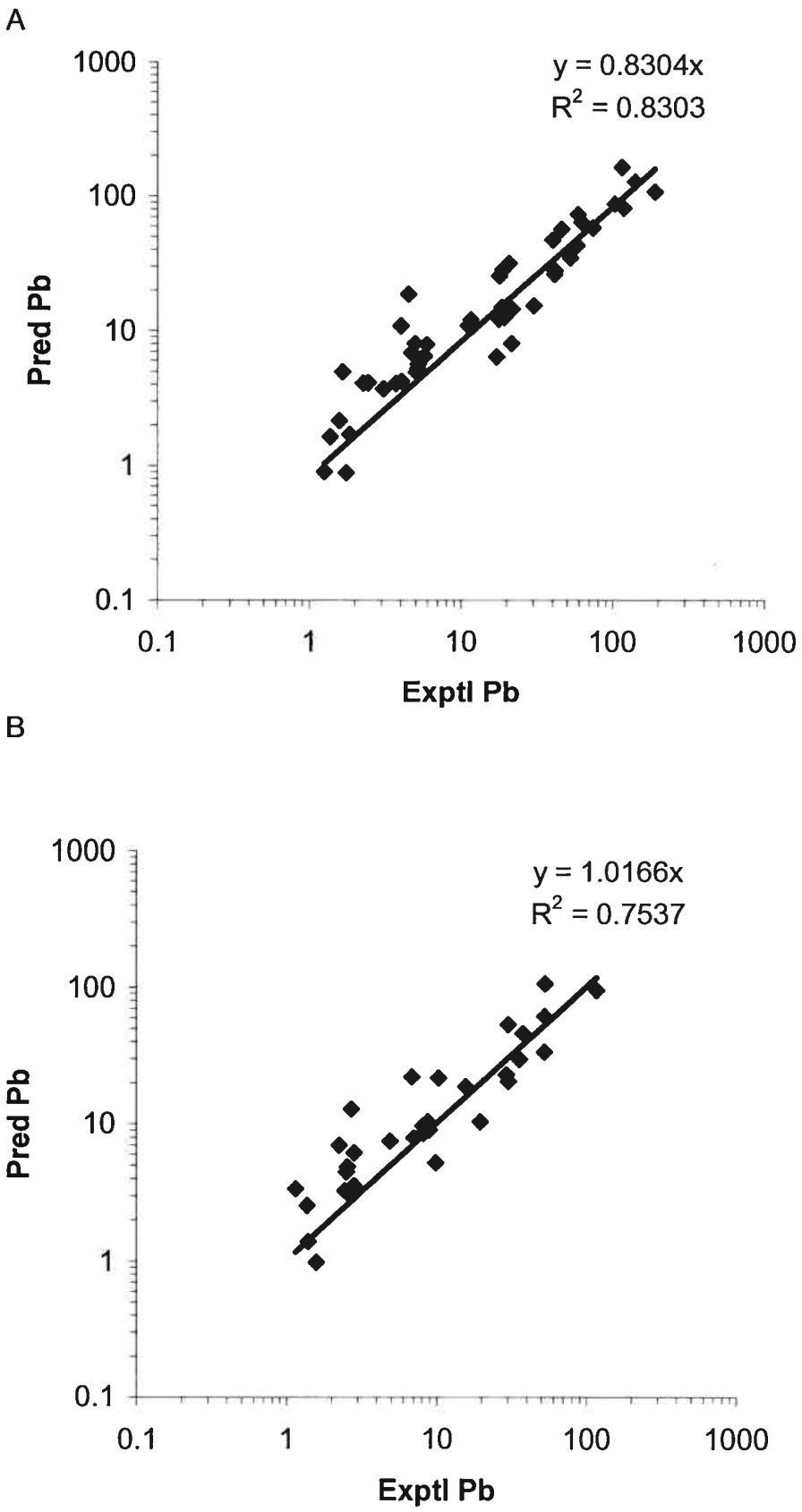

A
y = 1.0146x
R² = 0.7055
Pred
Exptl
B
y = 1.074x
R² = 0.5782
Pred
Exptl
C
y = 1.1743x
R² = 0.5711
Pred
Exptl

Figure 4


A
y = 0.8304x
R2 = 0.8303
Pred Pb
Exptl Pb
1000
100
10
1
0.1
0.1
1
10
100
1000
B
y = 1.0166x
R2 = 0.7537
Pred Pb
Exptl Pb
1000
100
10
1
0.1
0.1
1
10
100
1000

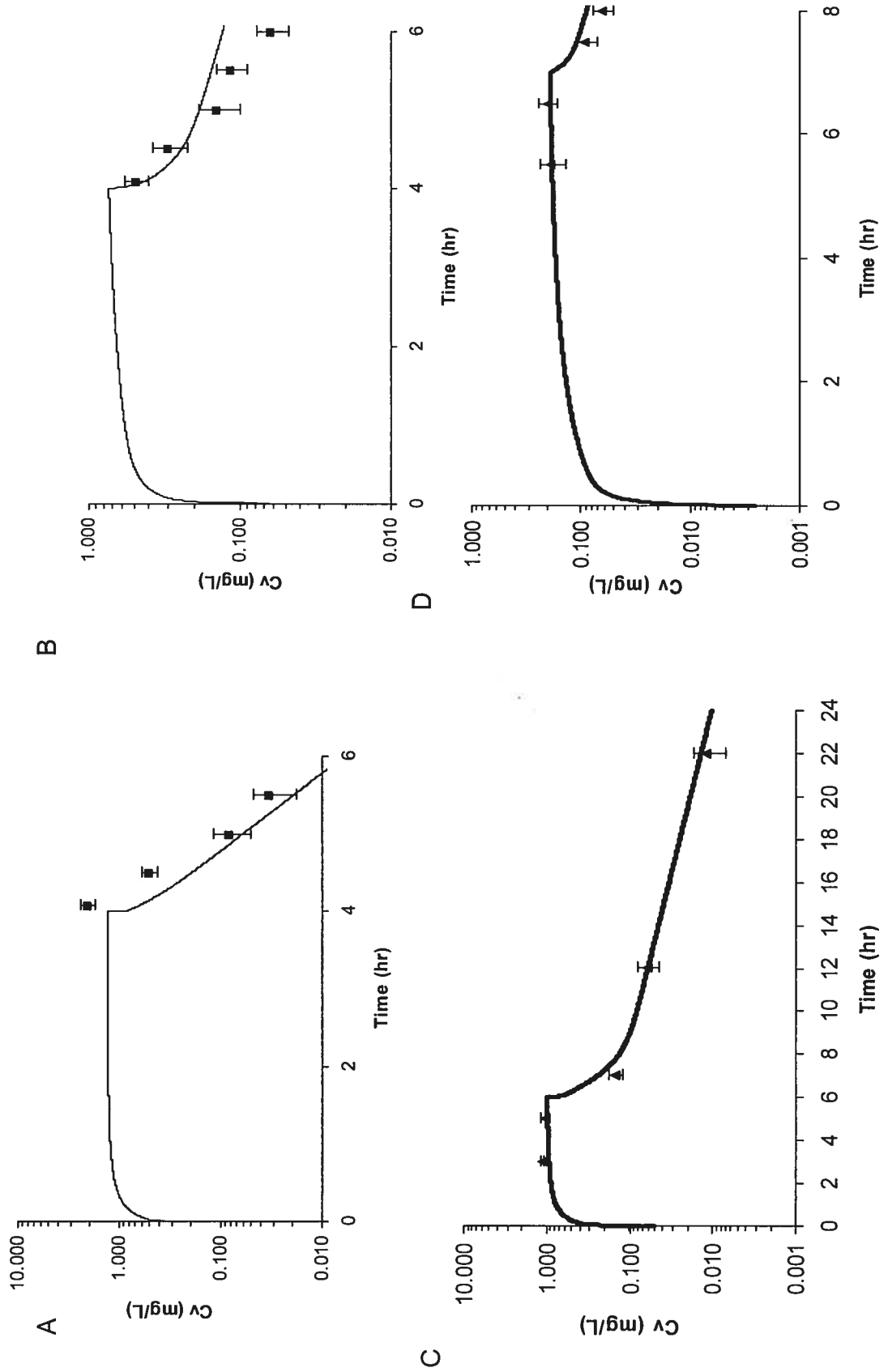

A
B
C
D
Cv (mg/L)
Time (hr)
10.000
1.000
0.100
0.010
0.001
0
2
4
6
8
10
12
14
16
18
20
22
24

Figure 5

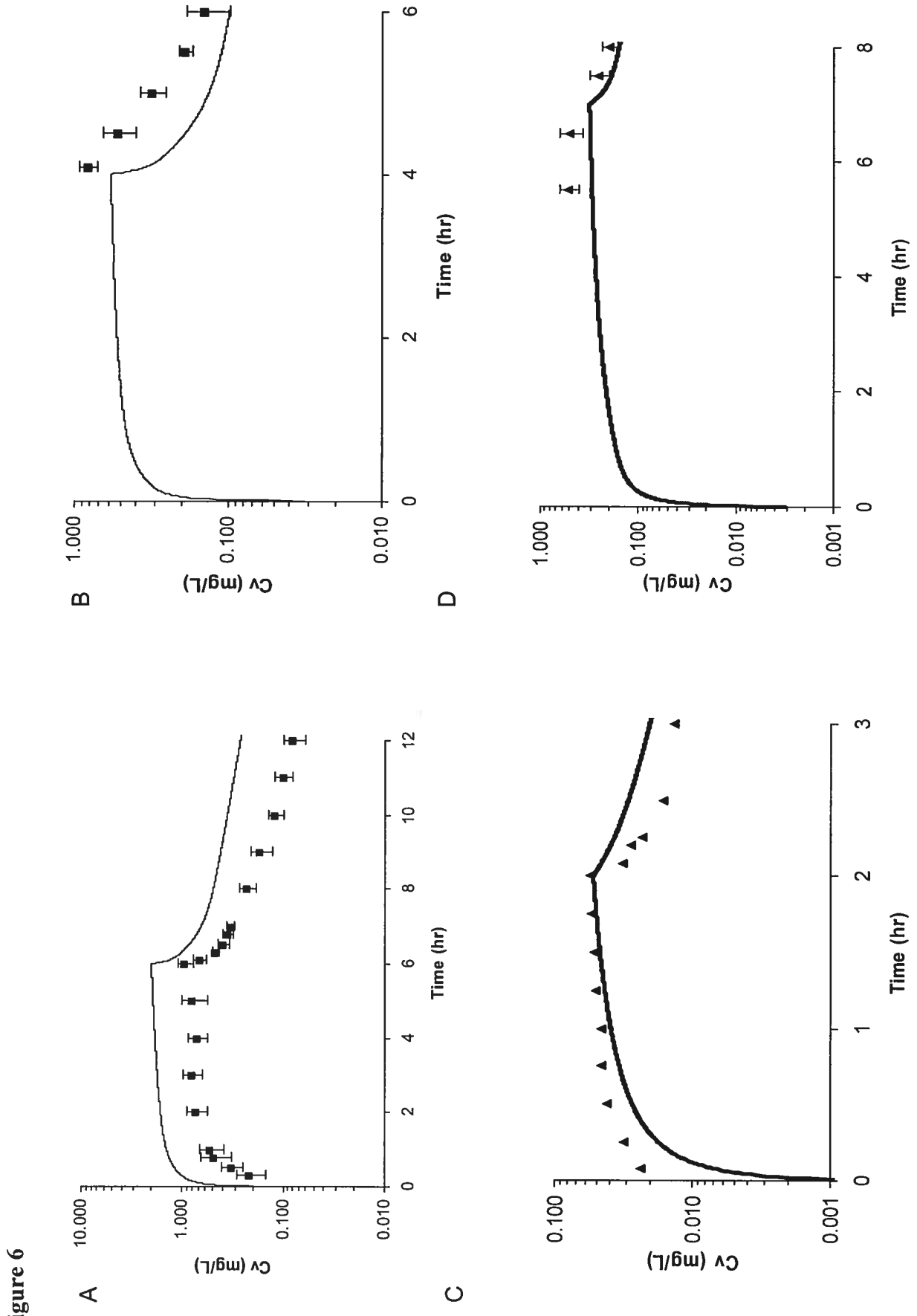


Figure 6

CHAPITRE SIXIEME:

6 Article V

Molecular structure-based prediction of the steady-state blood concentrations of inhaled organics in rats

For submission to: Toxicology Mechanisms and Methods

Molecular structure-based prediction of the steady-state blood concentrations of inhaled organics in rats

Martin Béliveau and Kannan Krishnan

Groupe de recherche en toxicologie humaine (TOXHUM), Université de Montréal, Case Postale 6128, Succ. Centre-Ville, Montréal, PQ, H3C 3J7, Canada.

Running title: Structure-blood concentration relationships

Address for correspondence: Dr. Kannan Krishnan, Département de santé environnementale et santé au travail, Faculté de médecine, Université de Montréal, 2375 Chemin de la Côte Ste-Catherine, Bureau 4105, Montréal, Québec, Canada, H3T 1A8. Tel: (514) 343-6581, Fax: (514) 343-2200. email: [redacted]

Abstract

Chronic exposure to volatile organic chemicals (VOCs) in the environment leads to steady-state conditions. The establishment of quantitative relationships between steady-state blood concentrations and molecular structures of VOCs can be potentially useful. The objective of this study was therefore to investigate the relationship between the steady-state arterial blood concentration (Ca_{ss}) in rat and the molecular structures of 19 VOCs belonging to multiple chemical families (alkanes, haloalkanes, haloalkenes and aromatics). The overall approach consisted of developing quantitative relationships between molecular fragments (CH_3, CH_2, CH, C, C=C, H, Cl, benzene ring and H in benzene ring structure) in alkanes, haloalkanes, haloethylenes, and aromatic hydrocarbons, as well as their Ca_{ss} (associated with 1 µmol/L inhalation exposure) according to an additive fragment model. This modeling approach implies that each fragment in the structure of a chemical has an additive and constant contribution to its Ca_{ss}. A multi-linear regression was performed using a commercially-available statistical software package, and the results obtained were essentially the contributions associated with each of the nine structural fragments towards Ca_{ss} in the rat continuously exposed to 1 µmol/L VOC in the air. The resulting model estimated adequately the Ca_{ss} of VOCs initially used in the calibration (estimated/experimental ratio: 1.04±0.30, mean±SD). This molecular structure *vs* Ca_{ss} relationship was then validated using an external dataset on Ca_{ss} for three aliphatic hydrocarbons (octane, 2-methyl octane and 1-nonene; 100 ppm exposures). The ratio of predicted to experimental Ca_{ss} for these chemicals ranged from 0.6 to 1.2. The results of this study suggest that steady-state blood

concentrations of inhaled VOCs can be predicted using structure-activity type models.

6.1 Introduction

Quantitative structure-activity relationships (QSARs) are useful for screening chemicals exhibiting potentially beneficial or undesirable activities (Cronin and Dearden 1995; Benigni and Richard 1998; Akers *et al.* 1999; Barratt 2000). Whereas numerous QSAR models for biological activity are available (e.g., Blaha *et al.* 1998; Franke *et al.* 2001; Greene 2002), such models relating molecular structures to internal doses of chemicals are only beginning to be developed. Beliveau *et al.* (2003) recently developed an approach to predict pharmacokinetics of inhaled volatile organic chemicals (VOCs) in the rat from molecular structure and exposure data. Initially, these authors developed quantitative relationships between the molecular structures of chemicals and their pharmacokinetic determinants (i.e., partition coefficients and hepatic clearance), which were then integrated with physiological parameters to simulate the internal dose of inhaled VOCs in rats. By integrating quantitative structure-property relationships (QSPRs) into physiologically-based pharmacokinetic (PBPK) models, structure-pharmacokinetic relationships could thus be simulated for any exposure concentration and duration (Beliveau *et al.* 2003).

Chronic, lifetime exposure is the most frequently considered exposure scenario in human health risk assessment applications (e.g., Andersen *et al.* 1987, 2000). Such continuous exposures lead to steady-state (SS) conditions within the organism (Andersen 1981; Pelekis *et al.* 1997; Csanady and Filser 2001). Therefore, the development of quantitative relationships between molecular structure and internal dose measures is relevant. There has been one previous attempt to develop QSAR models of steady-state blood concentrations (SS-BCs) in

rats (Fouchecourt and Krishnan 1999). In this study, the authors developed quantitative relationships between structural fragments of a series of aliphatic hydrocarbons and their SS-BCs using Free-Wilson approach. Since this particular approach requires that the chemicals have a common "back-bone" in their structure, it cannot be applied to model chemicals belonging to multiple families. This limitation can be overcome with the use of group contribution approach (Gao *et al.* 1992, Martin and Young 2001, Beliveau *et al.* 2003), which accounts for the contribution of all structural fragments in a molecule. This approach is different from the Free-Wilson approach in that it does not require a common "basic" structure among the chemicals investigated, which is a hurdle when considering multiple chemicals belonging to different families. The Gao's additive model can potentially be used to relate the SS-BC to structural fragments of VOCs belonging to multiple families, but this aspect has never been investigated.

The objective of this study was therefore to investigate the quantitative relationship between the steady-state blood concentrations in rat and the molecular structures of VOCs belonging to multiple chemical families (alkanes, haloalkanes, haloalkenes and aromatics).

6.2 Methods

Model development

Quantitative models of the relationship between molecular fragments of VOCs belonging to various classes (alkanes, haloalkanes, haloethylenes, and aromatic hydrocarbons) and their steady-state arterial blood concentration (Ca_{ss}, associated with continuous inhalation exposure to 1 µmol/L in air) were developed according to Gao *et al.* (1992). In this modeling approach, chemical fragments or groups are used to describe a molecule, accounting for all atoms in the molecule. As such, each molecular fragment is assumed to have a unique and finite contribution to Ca_{ss}. Table 1 presents the molecular fragments used to describe the VOCs investigated in this study. The characterization and representation of molecular fragments in VOCs investigated in the present study were done according to Beliveau *et al.* (2003). The alkanes were described using the basic UNIFAC (UNIversal Function group Activity Coefficient) groups such as CH_3, CH_2, CH, and C (Ochsner and Sokoloski 1985; Hansen *et al.* 1991). Whereas the alkenes were described using two molecular fragments, namely, C=C and H, substituent groups in the alkanes and alkenes also included Cl. The benzene ring was described as a single fragment (i.e., AC), and the hydrogen atoms attached to the ring were represented separately (H_AC). The VOCs investigated in the present study can be reconstituted by adding the frequency of occurrence of each fragment shown in Table 1.

The Ca_{ss} was described using a multilinear additive QSPR model (Martin and Young 2001; Beliveau *et al.* 2003) in which the frequency of occurrence (f) of

each fragment (i) in the various VOCs (Table 1) was related to their respective Ca_{ss} according to the following equation:

$$LogCa_{ss} = \sum_{i=1} f_i \cdot Cf_i \qquad [1]$$

where Cf_i represents the numerical contribution of each fragment i to the Ca_{ss}. The above QSPR implies that each fragment in the structure has an additive and constant contribution to Ca_{ss}. It should be noted that all QSPR model predictions presented in this paper represent the antilog of the Ca_{ss} values obtained with the above equation.

QSPR modeling is appropriately done using Ca_{ss} data expressed in terms of molar concentration, in order to account for different molecular weights. Therefore, Ca_{ss} (μM) required for QSPR modeling were obtained for dichloromethane (Andersen *et al.* 1991), tetrachloroethylene (Ward *et al.* 1988), toluene (Haddad *et al.* 2000), *m*-xylene (Haddad *et al.* 2000), styrene (Ramsey and Andersen 1984), carbon tetrachloride (Paustenbach *et al.* 1988), ethyl benzene (Haddad *et al.* 2000), chloroform (Reitz *et al.* 1990), trichloroethylene (Fisher *et al.* 1991), vinyl chloride (Reitz *et al.* 1996), and benzene (Haddad *et al.* 2000) using PBPK models. All these models describe the processes of absorption, distribution, metabolism and elimination and have been validated with blood concentration data collected near steady-state in rats of various strains. These PBPK models were written in Advanced Continuous Simulation Language (ACSL®, AEgis technologies, Huntsville, AL) and compiled to simulate the Ca_{ss} associated with continuous exposure to 1 μmol/L of each chemical in air (or 24.45 ppm).

This dataset was supplemented with seven other long chain hydrocarbons (hexane, heptane, nonane, decane, 2-methyl heptane, 2-methyl nonane, 1-octene

and 1-decene) for which experimental values on steady-state blood concentration were available (Zahlsen *et al.* 1992, 1993; Fouchecourt and Krishnan 1999). These experimental values were divided by the inhalation exposure concentration (100 ppm, 4.1 µmol/L air), considering that this exposure concentration is likely to be within the first order range for hepatic metabolism of these hydrocarbons (Fouchecourt and Krishnan 1999). The remaining three experimental steady-state blood concentrations available from the Zahlsen *et al.* studies (octane, 2-methyl octane, 1-nonene) were used for validation purposes.

A multilinear regression was performed on the available Ca_{ss} values and their corresponding frequency of occurrence using a commercially-available statistical software package (SPSS® for Windows® v10.0.7, SPSS Inc., Chicago, IL), and the results obtained were essentially the values of Cf_i in Eqn. 1 for each of the nine structural fragments (CH_3, CH_2, CH, C, C=C, H, Cl, AC, H_AC) specific to Ca_{ss} in rats exposed to 1 µmol/L air. The statistical significance ($p < 0.05$) of the regression coefficients was assessed by computing *t*-statistic (> 1.8) (Beliveau *et al.*, 2003).

Model validation

A cross-validation, as described in Beliveau *et al.* (2003), was performed to evaluate this QSPR. The predictive residual sum of squares (PRESS) statistic over the sum of squares of the response values (SSY) was used as a general indication of the validity of a model (Wold 1991). The PRESS statistic measures how well the regression equation fits the data. It is computed by removing the ith datapoint from the dataset, computing the regression equation without this datapoint, predicting

that point based on the regression equation, then computing the residual ("leave-one-out" procedure). This process is repeated for each datapoint, followed by the summing up of the sqares of each calculated residual (i.e., the difference between the predicted value of the parameter when a chemical is "left-out" and the experimental value for that chemical). A ratio <0.4 was considered to be a reasonable acceptance criteria to establish the validity of the model (Wold 1991). The Ca_{ss} from this cross-validated QSPR were then obtained and compared with experimental SS-BC for three aliphatic hydrocarbons (octane, 2-methyl octane, 1-nonene) that were not part of the original calibration data set (Zahlsen *et al.* 1992, 1993).

6.3 Results

The unit contributions of the nine molecular fragments to rat Ca_{ss}, as obtained from analysis of data for 19 VOCs, are presented in Table 2. The overall regression was highly positive and significant (R^2 = 0.896). Furthermore, the PRESS/SSY statistic of the QSPR model was 0.19. The PRESS/SSY statistic values for QSPRs of Ca_{ss} are within the suggested benchmark value (0.4) for adequate models. It should be noted, however, that four Cf_i are not statistically significant (*t* statistic < 1.8).

Tables 3 and 4 compare the QSPR-derived values with the Ca_{ss} values used for constructing the model (at 1 µmol/L air exposure). The QSPR-derived Ca_{ss} varied on average by a factor of 1.04±0.30 (smallest: 0.6, 1-octene; highest: 1.67, styrene) from the Ca_{ss} used for calibrating the model. These results suggest that Ca_{ss} of VOCs belonging to several families investigated in the present study can be described with a group contribution model.

The QSPR model was then used to predict the steady-state blood concentration of three aliphatic hydrocarbons that were not part of the calibration data set (Table 5). The estimated/experimental Ca_{ss} ratio was 1.2 for octane, 1.1 for 2-methyl octane, and 0.6 for 1-nonene.

6.4 Discussion

Quantitative structure-property models allow the estimation of input parameters for pharmacokinetic models, which in turn can be used to generate the blood concentration profiles associated with defined exposure scenarios (e.g., Beliveau *et al.* 2003). For the health risk assessment of environmental chemicals, continuous lifetime exposure is the scenario considered and it would likely lead to steady-state conditions within the organism, specifically in the case of volatile organic chemicals. In such instances, the calculation of steady-state blood concentrations (Ca_{ss}) does not have to rely on "full-blown" PBPK models but would only require the knowledge of inhaled dose, pulmonary clearance and hepatic clearance. For a given exposure concentration then, the inter-chemical difference in Ca_{ss} is a result of the difference in total clearance (pulmonary plus hepatic), which can possibly be linked to structure. If such quantitative models of the relationship between structural fragments and Ca_{ss} are developed for a series of chemicals, then predictions for other chemicals can possibly be made. Direct relationships between structure and experimental Ca_{ss} in rats have been established for a limited number of aliphatic hydrocarbons having a common "skeletal" group in their structures using a Free-Wilson approach (Fouchecourt and Krishnan 1999). However, the establishment of Free-Wilson type QSPRs is hampered by the limited amount of experimental data on Ca_{ss} available for each specific sub-set or class of chemicals such as halomethanes, haloethanes or aromatics. An alternative approach, as implemented in this study, would be to evaluate these diverse chemicals simultaneously and apply a group contribution model that would also increase the overall power of the resulting QSPRs. The

resulting group contributions can also be applied without restriction to any sub-group of chemical families (like aliphatic hydrocarbons, in this instance), because the contributions are derived from the whole dataset. Any intrinsic error in the values estimated would therefore be expected to be the same across the sub-groups.

The group contributions obtained in this study are potentially useful in assessing the impact of varying chemical structures on steady-state blood concentration. For example, in the case of a series of aromatic hydrocarbons, following continuous exposure to 1 µmol/L in air, Ca_{ss} would range from 3.35 to 7.17 µM (Table 6). Considering benzene, toluene, xylene and trimethylbenzene, as the number of CH_3 increases from 0 to 1 to 2 to 3, there is not a linear and proportional increase in Ca_{ss} (Table 6). This is essentially due to the log-linear relationship between Ca_{ss} and Cf, as well as the fact that increase in contribution by CH_3 is compensated by a decrease in the contribution by H on the aromatic ring. Similarly, in a series of halogenated methanes, Ca_{ss} values for 1 µmol/L inhalation exposure would increase from 0.69 to 5.24 (Table 7) but it essentially is a net result of the difference between the increase in Ca_{ss} due to increase in the number of CL and corresponding decrease due to change in the number of hydrogens. In comparing 1,1-dichloroethane and 1,1,1-trichloroethane values, for example, it can be seen that the increase in the number of chlorines (with corresponding decrease in the number of hydrogen atoms) has a net effect of decreasing Ca_{ss} (Table 7).

The present study used Ca_{ss} values associated with 1 µmol/L inhalation exposure, for calibration of the model. Therefore, the resulting QSPR predictions should be multiplied by exposure concentration of interest to compute the

appropriate Ca_{ss} values. Such a calculation is only valid as long the exposure levels are within the linear kinetic range. QSPR-based extrapolation of Ca_{ss} to exposure levels beyond the linear range of metabolic clearance would lead to an underestimation of the true values and is therefore not recommended. The present methodology remains useful however, as a tool for providing a "first cut" estimation of steady-state blood concentrations for *de novo* compounds, given that environmental exposures to VOCs lead to steady-state condition and that the anticipated inhalation exposure concentrations are low, respecting the linear kinetics assumption. For higher exposure concentrations, calculations of blood concentrations can be conducted using molecular structure information in PBPK models (Beliveau *et al.*, 2003).

Overall, the present study developed quantitative relationships between molecular structure of VOCs belonging to more than one family and their steady-state blood concentrations in rats. The QSPR model developed in this study is a potentially useful tool for providing a "first-cut" estimate of the Ca_{ss} for VOCs as long as the nature and number of molecular fragments contained in such compounds do not exceed what was used in the calibration set in this study.

Acknowledgement

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). M.B. is recipient of a scholarship from the Natural Sciences and Engineering Research Council and Fonds de Recherche sur la Nature et les Technologies.

6.5 References

Akers, K. S., Sinks, G. D., and Schultz, T. W. (1999). Structure-toxicity relationships for selected halogenated aliphatic chemicals. *Environ. Toxicol. Pharmacol.* **7**, 33-39.

Andersen, M. E. (1981). A physiologically based toxicokinetic description of the metabolism of inhaled gases and vapors: analysis at steady state. *Toxicol. Appl. Pharmacol.* **60**, 509-26.

Andersen, M. E., Clewell, H. J. I., and Gargas, M. L. (1991). Physiologically-based pharmacokinetic modeling with dichloromethane, its metabolite carbon monoxide and blood carboxyhemoglobin in rats and humans. *Toxicol. Appl. Pharmacol.* **108**, 14-27.

Andersen, M. E., Clewell, H. J. I., Gargas, M. L., Smith, F. A., and Reitz, R. H. (1987). Physiologically-based pharmacokinetics and risk assessment process for methylene chloride. *Toxicol. Appl. Pharmacol.* **87**, 185-205.

Andersen, M. E., Sarangapani, R., Frederick, C. B., and Kimbell, J. S. (2000). Dosimetric adjustment factors for methyl methacrylate derived from a steady-state analysis of a physiologically based clearance-extrapolation model. *Inh. Toxicol.* **11**, 899-926.

Barratt, M. D. (2000). Prediction of toxicity from chemical structure. *Cell Biol. Toxicol.* **16**, 1-13.

Beliveau, M., Tardif, R., and Krishnan, K. (2003). Quantitative structure-property relationships for physiologically based pharmacokinetic modeling of volatile organic chemicals in rats. *Toxicol Appl Pharmacol* **189**, 221-32.

Benigni, R., and Richard, A. M. (1998). Quantitative structure-based modeling applied to characterization and prediction of chemical toxicity. *Methods Enzymol.* **14**, 284-76.

Blaha, L., Damborsky, J., and Nemec, M. (1998). QSAR for acute toxicity of saturated and unsaturated halogeneted aliphatic compounds. *Chemosphere* **36**, 1345-65.

Cronin, M. T. D., and Dearden, J. C. (1995). QSAR in toxicology. 3. Prediction of chronic toxicities. *Quant.-Struct.-Act.-Rel.* **14**, 329-34.

Csanady, G. A., and Filser, J. G. (2001). The relevance of physical activity for the kinetics of inhaled gaseous substances. *Arch Toxicol* **74**, 663-72.

Fisher, J. W., Gargas, M. L., Jepson, G. W., Allen, B., and Andersen, M. E. (1991). Physiologically-based pharmacokinetic modeling with trichloroethylene and its metabolite, trichloroacetic acid in the rat and mouse. *Toxicol. Appl. Pharmacol.* **109**, 183-95.

Fouchecourt, M.-O., and Krishnan, K. (1999). Quantitative relationship between steady-state blood concnetrations and structural features of aliphatic hydrocarbons. *Toxicol. Lett.* **110**, 177-82.

Franke, R., Gruska, A., Giuliani, A., and Benigni, R. (2001). Prediction of rodent carcinogenicity of aromatic amines: a quantitative structure-activity relationship model. *Carcinogenesis* **22**, 1561-71.

Gao, C., Govind, R., and Tabak, H. H. (1992). Application of the group contribution method for predicting the toxicity of organic chemicals. *Environ. Toxicol. Chem.* **11**, 631-36.

Greene, N. (2002). Computer systems for the prediction of toxicity: an update. *Adv. Drug Deliv. Rev.* **54**, 417-31.

Haddad, S., Charest-Tardif, G., Tardif, R., and Krishnan, K. (2000). Validation of a physiological modeling framework for simulating the toxicokinetics of chemicals in mixtures. *Toxicol. Appl. Pharmacol.* **167**, 199-209.

Hansen, H. K., Rasmussen, P., Fredenslund, A., Schiller, M., and Gmehling, J. (1991). Vapor liquid equilibria by UNIFAC group contibution, 5. Revision and extension. *Ind. Eng. Chem. Res.* **30**, 2355-58.

Martin, T. M., and Young, D. M. (2001). Prediction of the acute toxicity (96-h LC_{50}) or organic compounds to the fathead minnow (*Pimephales promelas*) using a group contribution method. *Chem. Res. Toxicol.* **14**, 1378-85.

Ochsner, A. B., and Sokoloski, T. D. (1985). Prediction of solubility in nonideal multicomponent systems using the UNIFAC group contribution model. *J Pharm Sci* **74**, 634-7.

Paustenbach, D., Andersen, M. E., Clewell, H. J., and Gargas, M. L. (1988). A physiologically-based pharmacokinetic model for inhaled carbon tetrachloride in the rat. *Toxicol. Appl. Pharmacol.* **96**, 191-211.

Pelekis, M., Krewski, D., and Krishnan, K. (1997). Physiologically based algebraic expressions for predicting steady-state toxicokinetics of inhaled vapors. *Toxicol. Methods* **7**, 205-25.

Ramsey, J. C., and Andersen, M. E. (1984). A physiologically-based description of the inhalation pharmacokinetics of styrene in rats and humans. *Toxicol. Appl. Pharmacol.* **73**, 159-75.

Reitz, R. H., Gargas, M. L., Andersen, M. E., Provan, W. M., and Green, T. L. (1996). Predicting cancer risk from vinyl chloride exposure with a physiologically based pharmacokinetic model. *Toxicol. Appl. Pharmacol.* **137**, 253-67.

Reitz, R. H., Mandrela, A. L., Corley, R. A., Quast, J. F., Gargas, M. L., Andersen, M. E., Staats, D. A., and Conolly, R. B. (1990). Estimating the risk of liver cancer associated with human exposures to chloroform using physiologically-based pharmacokinetic modeling. *Toxicol. Appl. Pharmacol.* **105**, 443-59.

Ward, R. C., Travis, C. C., Hetrick, D. M., Andersen, M. E., and Gargas, M. L. (1988). Pharmacokinetics of tetrachloroethylene. *Toxicol. Appl. Pharmacol.* **93**, 108-17.

Wold, S. (1991). Validation of QSAR's. *Quant.-Struct.-Act.-Rel.* **10**, 191-93.

Zahlsen, K., Eide, I., Nilsen, A. M., and Nilsen, O. G. (1992). Inhalation kinetics of C6 to C10 aliphatic, aromatic and naphthenic hydrocarbons in rat after repeated exposures. *Pharmacol Toxicol* **71**, 144-9.

Zahlsen, K., Eide, I., Nilsen, A. M., and Nilsen, O. G. (1993). Inhalation kinetics of C8 to C10 1-alkenes and iso-alkanes in the rat after repeated exposures. *Pharmacol Toxicol* **73**, 163-8.

Table 1: Frequency of occurrence of molecular fragments in the VOCs investigated in the present study.

Fragment / Chemical	CH_3	CH_2	CH	C	C=C	$H_{(on =)}$	CL	AC	$H_{(on AC)}$
1-Decene	1	7	0	0	1	3	0	0	0
1-Octene	1	5	0	0	1	3	0	0	0
2-Methyl heptane	3	4	1	0	0	0	0	0	0
2-Methyl nonane	3	6	1	0	0	0	0	0	0
Benzene	0	0	0	0	0	0	0	1	6
Carbon tetrachloride	0	0	0	1	0	0	4	0	0
Chloroform	0	0	1	0	0	0	3	0	0
Decane	2	8	0	0	0	0	0	0	0
Dichloromethane	0	1	0	0	0	0	2	0	0
Ethyl benzene	1	1	0	0	0	0	0	1	5
Heptane	2	5	0	0	0	0	0	0	0
Hexane	2	4	0	0	0	0	0	0	0
m-Xylene	2	0	0	0	0	0	0	1	4
Nonane	2	7	0	0	0	0	0	0	0
Styrene	0	0	0	0	1	3	0	1	5
Tetrachloroethylene	0	0	0	0	1	0	4	0	0
Toluene	1	0	0	0	0	0	0	1	5
Trichloroethylene	0	0	0	0	1	1	3	0	0
Vinyl chloride	0	0	0	0	1	3	1	0	0

n	26	65	4	1	7	16	17	5	25

Table 2: Fragment specific contributions (C_f, [µmol/L air]$^{-1}$) to the steady-state arterial blood concentration (Ca_{ss}) of volatile organic chemicals in rats.

Fragment	$C_{f\,(rat)}$	t^a
CH_3	-0.296	2.99
CH_2	0.109	3.42
CH	0.304	1.74
C	-0.0666	0.231
C=C	0.387	1.37
$H_{(on =)}$	-0.127	1.22
CL	0.138	2.34
AC	2.46	2.97
$H_{(on\ AC)}$	-0.323	2.10

[a] *t*-statistic > 1.8 indicates significance at $p<0.05$.

[b] A PRESS/SSY statistic < 0.4 indicates an adequate model.

Table 3: Comparison of QSPR model estimations (P) with the values used in the calibration of the model (E) for the steady-state arterial blood concentration (Ca_{ss}, µM) in rats for a series of volatile organic chemicals.

Chemical	$Ca_{ss\,(P)}$[a]	$Ca_{ss\,(E)}$[b]	P/E Ratio
Dichloromethane	2.43	3.94	0.62
Tetrachloroethylene	8.72	10.5	0.83
Toluene	3.56	4.29	0.83
m-Xylene	3.79	3.97	0.95
Styrene	7.17	4.30	1.67
Carbon tetrachloride	3.07	3.07	1.00
Ethyl benzene	4.58	5.77	0.79
Chloroform	5.24	4.11	1.27
Trichloroethylene	4.74	3.60	1.32
Vinyl chloride	1.40	1.20	1.16
Benzene	3.35	3.52	0.95

[a] Calculated using Equation 1 for an exposure concentration of 1 µmol/L air.

[b] Estimated at steady-state following continuous inhalation exposure to 1 µmol/L air, using validated PBPK models.

Table 4: QSPR predicted (P) vs experimental (E) steady-state arterial blood concentration (Ca_{ss}, µM) in rats for a series of aliphatic hydrocarbons.

Chemical	$Ca_{ss\,(P)}$ [a]	$Ca_{ss\,(E)}$ [b]	P/E Ratio
1-Decene	12.18	16.27	0.7
1-Octene	7.37	11.7	0.6
2-Methyl heptane	2.90	3.07	0.9
2-Methyl nonane	4.79	5.77	0.8
Decane	7.78	6.37	1.2
Heptane	3.66	2.47	1.5
Hexane	2.85	2.87	1.0
Nonane	6.05	4.03	1.5

[a] Predictions obtained using Equation [1] for 1 µmol/L air exposure were multiplied with the corresponding exposure concentration (i.e., 4.1 µmol/Lair or 100 ppm).

[b] Experimental values from Zahlsen *et al.* (1992, 1993).

Table 5: QSPR predicted (P) vs experimental (E) steady-state arterial blood concentration (Ca_{ss}, µM) in rats for three aliphatic hydrocarbons.

Chemical	Fragments					$Ca_{ss\ (P)}$[a]	$Ca_{ss\ (E)}$[b]	P/E Ratio
	CH_3	CH_2	CH	C=C	$H_{(on\ =)}$			
1-Nonene[c]	1	6	0	1	3	9.47	15.03	0.6
2-Methyl octane	3	5	1	0	0	3.73	3.33	1.1
Octane	2	6	0	0	0	4.71	3.9	1.2

[a] Predicted using Equation [1] for 1 µmol/L air exposure and multiplied by the actual exposure concentration (4.1 µmol/L air, 100 ppm).

[b] Experimental values from Zahlsen *et al.* (1992, 1993).

[c] For 1-nonene: Log $Ca_{ss\ (1\ \mu mol/L\ air)}$ = [1*CH_3 (-0.296)] + [6*CH_2 (0.109)] + [1*C=C (0.387)] + [3*$H_{(on=)}$ (-0.127)] = 0.364 => $Ca_{ss\ (1\ \mu mol/L\ air)}$ = $10^{0.364}$ = 2.31 µM (per µmol/L air). For a 100 ppm (4.1 µmol/L air) exposure: Ca_{ss} = 2.31 * 4.1 = 9.47 µM.

Table 6: Comparison of the arterial blood concentration at steady-state (Ca_{ss}) predicted according to the molecular structure of a series of benzenes in rats exposed by inhalation to 1 µmol/L air.

Fragment	C_f[a]	Frequency of occurrence of fragment in chemical[b]					
		BEN	TOL	EBZ	XYL	TMB	STY
CH_3	-0.296	0	1	1	2	3	0
CH_2	0.109	0	0	1	0	0	0
C=C	0.387	0	0	0	0	0	1
$H_{(on =)}$	-0.127	0	0	0	0	0	3
AC	2.46	1	1	1	1	1	1
$H_{(on AC)}$	-0.323	6	5	5	4	3	5
Ca_{ss} (µM)		3.35	3.56	4.58	3.79	4.02	7.17

[a] C_f = contribution of the molecular fragment to Ca_{ss}.

[b] BEN, TOL, EBZ, XYL, TMB, and STY represent benzene, toluene, ethylbenzene, m-xylene, trimethylbenzene, and styrene, respectively.

Table 7: Comparison of the arterial blood concentration at steady-state (Ca_{ss}) predicted according to the molecular structure of a series of chlorinated hydrocarbons in rats exposed by inhalation to 1 µmol/L.

Fragment	C_f[a]	Frequency of occurrence of fragment in chemical[b]							
		CM	DCM	CFM	CT	CE	1,1-DCE	TCE	1,2-DCE
CH_3	-0.296	1	0	0	0	1	1	1	0
CH_2	0.109	0	1	0	0	1	0	0	2
CH	0.304	0	0	1	0	0	1	0	0
C	-0.067	0	0	0	1	0	0	1	0
CL	0.138	1	2	3	4	1	2	3	2
Ca_{ss} (µM)		0.69	2.43	5.24	3.07	0.89	1.92	1.13	3.12

[a] C_f = contribution of the molecular fragment to Ca_{ss} in rats and humans, respectively.

[b] CM, DCM, CFM, CT, CE, 1,1-DCE, TCE and 1,2-DCE represent chloromethane, dichloromethane, chloroform, carbon tetrachloride, chloroethane, 1,1-dichloroethane, 1,1,1-trichloroethane, and 1,2-dichloroethane, respectively.

CHAPITRE SEPTIEME:

7 Article VI

Molecular-structure-based interspecies extrapolation of the steady-state concentrations of inhaled organics

For submission to: Inhalation Toxicology

Molecular-structure-based interspecies extrapolation of the steady-state concentrations of inhaled organics

Martin Béliveau, Robert Tardif, and Kannan Krishnan

Groupe de recherche en toxicologie humaine (TOXHUM), Université de Montréal, Case Postale 6128, Succ. Centre-Ville, Montréal, PQ, H3C 3J7, Canada.

Running title: QSPRs for interspecies extrapolation of steady-state concentrations

Address for correspondence: Dr. Kannan Krishnan, Département de santé environnementale et santé au travail, Faculté de médecine, Université de Montréal, 2375 Chemin de la Côte Ste-Catherine, Bureau 4105, Montréal, Québec, Canada, H3T 1A8. Tel: (514) 343-6581, Fax: (514) 343-2200. email: [redacted]

Abstract

The objective of this study was to develop a methodology for extrapolating steady-state (SS) blood concentrations from rats to humans on the basis of information on chemical structure and physiological determinants of the clearance processes involved. The SS blood concentrations were extrapolated from rats to humans for a series of 11 structurally-unrelated volatile organic chemicals (VOCs) using the hepatic and pulmonary clearance (CL) predicted in each species with QSPRs obtained from the literature. Accordingly, only the frequency of occurrence (f) of each fragment in the various VOCs was provided as input. Each CL was then summed to determine the species-specific systemic CL and used to predict the corresponding Ca_{ss} in each species. The Ca_{ss} estimated using total CL in rats was compared to the values obtained using the previously validated rat PBPK models. The Ca_{ss} values, in the rat, obtained using the two approaches varied by a factor of 1.28±0.37. The ratio of predicted to experimental values in humans averaged 1.19±0.58. This study has demonstrated that QSPRs relating chemical structure to the mechanistic determinants of clearance processes can be used to assess interspecies difference in steady-state blood concentrations.

7.1 Introduction

Environmental exposure to organic chemicals most often results in steady-state conditions within the organism. Under such exposure conditions, steady-state (SS) blood concentrations (Cb_{ss}) are essentially the result of systemic clearance mechanisms within the body, and the use of more complex physiologically-based pharmacokinetic (PBPK) modeling becomes unnecessary. On a unit exposure-basis, Cb_{ss} levels are expected to be inversely proportional to systemic clearance, such that a direct relationship between Cb_{ss} and molecular structure can be established. Such relationships have been developed for a specific chemical class or multiple chemical families in rats using a quantitative structure-property relationship (QSPR) based on a fragment contribution approach (Fouchecourt and Krishnan 1999; Beliveau *et al.* 2004). However, the fragment contributions derived are expected to vary according to the species in which the steady-state data were obtained. The most straightforward approach would involve deriving fragment contributions in both species in order to perform interspecies extrapolations. The magnitude of the species component within each contribution values could therefore be easily quantified for each extrapolation combination (e.g., rat-human, mouse-human, etc.). However, steady-state data are not always readily available, in which case the resulting statistical power of the QSPR is diminished. Since Cb_{ss} is essentially driven by systemic clearance, an alternative approach is to relate structure to the underlying clearance mechanisms, for which data are more readily available. Furthermore, this provides a scientifically-sound way of extrapolating Cb_{ss} across species, as the chemical-specific parameters can be fixed and the biological parameters can be varied according to the species of interest. QSPRs

relating chemical structure to determinants of SS blood concentration, i.e., clearances in both rats and humans have already been developed (Beliveau *et al.* 2003, 2004). Building on these data, the present study developed a methodology for extrapolating steady-state blood concentrations from rats to humans on the basis of information on chemical structure and physiological determinants of the clearance processes involved.

7.2 Methods

Steady-state concentration – structure relationship

The Cb_{ss} (µmol/L) of volatile organics can be described using the following mechanistic algorithm (Andersen 1981; Pelekis *et al.* 1997; Csanady and Filser 2001):

$$Cb_{ss} = \frac{Q_p \cdot C_i}{CL_p + CL_h} \qquad [1]$$

where Q_p = pulmonary ventilation (L/h)

C_i = inhaled concentration (µmol/L air)

CL_p = pulmonary clearance (L/h), and

CL_h = hepatic clearance (L/h)

While QSPRs relating structure to hepatic clearance are available in the literature, there are no known such algorithms for CL_p. Since CL_p is essentially the ratio of pulmonary ventilation over the blood:air partition coefficient (Q_p/P_b), human and rat-specific QSPRs relating structure to P_b can be used.

Beliveau *et al.* (2003; 2004) developed QSPRs relating molecular fragments of a structurally diverse set of volatile organic chemicals (VOCs) to CL_h and P_b of rats as well as humans. In those studies, chemicals were described using chemical fragments (Table 1), each with a constant and additive contribution to the pharmacokinetic parameter of interest (CL_h and P_b) as follows:

$$Cb_{ss} = \frac{Q_p \cdot C_i}{\frac{Q_p}{10^{\sum f_i \cdot Cf(P_b)_i}} + \sum_{i=1} f_i \cdot Cf(CL_h)_i} \qquad [2]$$

where f = frequency of occurrence of the fragment i in the chemical of interest

Cf = contribution of the fragment to the parameter of interest (P_b and CL_h, respectively).

Using the above equation, the frequencies of occurrence of fragments were specified to determine the Cb_{ss} of a series of chemicals in a given species (i.e., rat). To extrapolate from rats to humans, the species-specific parameters (Q_p and the specific contribution of each fragment) were replaced. The species-specific contributions of each fragment, obtained from the literature, are listed in Table 2.

Parametrization

Q_p was set at 5.38 L/hr and 417 L/hr in rats and humans, respectively (Ramsey and Andersen 1984; Tardif et al., 1997). The frequency of occurrence of fragments (Table 1) along with the contribution values to Log P_b and CL_h in rat and humans (Beliveau *et al.* 2003, 2004) (Table 2) were used to compute the corresponding chemical-specific parameters of clearance. The CL values were then used to estimate the Cb_{ss} of the VOCs presented in Table 1 in each species for an exposure of 1 µmol/L air (equivalent to 24.45 ppm) according to Eqn. 2.

Validation

The Cb_{ss} obtained using the QSPR-based algorithm was compared to SS concentration data obtained from validated rat and human PBPK models for dichloromethane (Andersen *et al.* 1991), tetrachloroethylene (Ward *et al.* 1988; Loizou 2001), toluene (Tardif *et al.* 1997; Haddad *et al.* 2000), *m*-xylene (Tardif *et al.* 1997; Haddad *et al.* 2000), styrene (Ramsey and Andersen 1984), carbon

tetrachloride (Paustenbach *et al.* 1988; El-Masri *et al.* 1996), ethyl benzene (Tardif *et al.* 1997; Haddad *et al.* 2000), chloroform (Corley *et al.* 1990; Reitz *et al.* 1990), trichloroethylene (Fisher *et al.* 1991; Allen and Fisher 1993), vinyl chloride (Reitz *et al.* 1996), and benzene (Medinsky *et al.* 1996; Sherwood and Sinclair 1999; Haddad *et al.* 2000). The PBPK models written in Advanced Continuous Simulation Language (ACSL®, AEGIS Technologies, Huntsville, AL) were compiled to simulate the Cb_{ss} associated with continuous exposure to 1 µmol/L air.

7.3 Results

The fragment contributions for the P_b and CL_h in rat and humans were used to determine the Cb_{ss} in both species according to Eqn. 2 for 11 structurally-unrelated VOCs presented in Table 1. The frequency of occurrence of each fragment in a given molecule was used as a chemical-specific, species-independent, parameter. Table 3 presents the Cb_{ss} obtained with the QSPR-based algorithm for the 11 structurally-unrelated VOCs. These Cb_{ss} are compared with SS concentrations derived using validated PBPK models in Table 3. The ratio of predicted to experimental values in rats averaged 1.28±0.37 with a range of 0.80 [tetrachloroethylene] to 1.91 [vinyl chloride]. Following vinyl chloride, the largest discrepancies were observed for carbon tetrachloride [1.81] and benzene [1.65]. The ratios for the remaining chemicals in the series averaged 1.09.

The steady-state concentrations in humans were estimated (Table 3) by substituting the species-specific parameters in Eqn. 2 (pulmonary ventilation and species-specific contributions). Table 3 also compares these concentrations to those obtained using validated PBPK models. The ratio of predicted to experimental values in humans averaged 1.19±0.58 with a range of 0.43 [carbon tetrachloride] to 2.64 [tetrachloroethylene]. Following tetrachloroethylene, the largest discrepancies were observed for carbon tetrachloride and ethyl benzene [1.59]. The ratios for the remaining chemicals averaged 1.05.

7.4 Discussion

Because environmental exposure most often leads to steady-state conditions within the organism, internal blood concentrations of compounds can be estimated using simple algebraic equations that relate exposure to concentration through clearance of the compound from the blood. Although steady-state levels can be directly related to chemical structure (Fouchecourt and Krishnan 1999), this necessitates some *a priori* knowledge of the concentrations expected in the species of interest, in the form of experimental data that is not always available, either at all or in sufficient quantity to generate a suitable QSPR. Experimental data on the clearance or the mechanistic determinants of clearance can be more readily obtained, however. Using available QSPRs for clearance processes allows for the extrapolation from one species to another, by keeping chemical specific data constant and adapting species-specific parameters to the species of interest. This offers the advantage of not necessitating prior knowledge of the PK endpoint itself in any species.

In this study, QSPRs for CL in animals and humans were used to predict the Cb_{ss} because the interspecies differences in physiology were taken into account separately from the interchemical differences in Eqn. 1. To extrapolate across species (across columns in Table 3), only the physiological parameters were changed. To extrapolate across chemicals (across rows in Table 3) only the chemical-specific parameters (i.e., the frequency of occurrence of each fragment within a molecule) were varied. Table 4 presents the rat-to-human SS blood concentration ratios for the chemicals investigated in this study. Given that the ratio of steady-state concentrations is essentially the result of the ratio of different

clearance processes in rat and humans, this ratio can be estimated based solely on the available QSPRs relating structure to clearance processes within each species. This allows for the establishment of a direct link between structure and the pharmacokinetic uncertainty factor used in current risk assessment methodologies.

Saturation of metabolic pathways can be taken into consideration in Eqn. 2 by replacing chemical-specific information by the physiological limits of hepatic clearance (i.e., induction/inhibition of metabolic pathways will result in hepatic extraction ratios ranging between 0 and 1). This can then be used as a "worst case" scenario to establish a range of possible concentrations, without requiring any additional chemical-specific information. This provides an advantage over the empirical relationship between steady-state levels and molecular structure (Fouchecourt and Krishnan 1999).

Because hepatic and pulmonary clearances are the main components of the systemic clearance of VOCs (Andersen 1981; Pelekis *et al.* 1997; Krishnan and Andersen 2001), these components were individually represented in Eqn. 2. For chemicals suspected of additional clearances within the organism (e.g., chemicals with important renal clearance), additional clearance terms would need to be included. As with hepatic clearance, the physiological limits of the clearing organ of interest can provide a best guess toward assessing the impact of such clearance on the potential blood concentrations.

Further analyses on the sensitivity of the rat to human blood concentration ratio with regard to information on the CL processes and, indirectly, the fragment specific contributions can help evaluate the potential impact of chemical structure on these processes. For certain chemical structures, uncertainty with regard to the contribution of each chemical fragment (e.g., CH_3, CH_2, CH, C, C=C, $H_{(on =)}$, Br, Cl,

F, AC, $H_{(on\ AC)}$, Table 2) to certain pharmacokinetic characteristics would not have a significant impact on the concentration ratio, given the relative importance of the competing physiological processes. Such sensitivity analyses are of potential use in determining which chemical features are critical to the determination of the blood concentrations.

Overall, the method implemented in this study is useful for evaluating the steady-state blood concentrations of VOCs in multiple species, on the basis information on chemical structure and physiological determinants of clearance processes.

Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council and the Canadian Network of Toxicology Centres. M.B. is recipient of a scholarship from the Natural Sciences and Engineering Research Council and Fonds de Recherche sur la Nature et les Technologies. .

7.5 References

Allen, B. C., and Fisher, J. W. (1993). Pharmacokinetic modeling of trichloroethylene and trichloroacetic acid in humans. *Risk Analysis* **13**, 71-86.

Andersen, M. E. (1981). A physiologically based toxicokinetic description of the metabolism of inhaled gases and vapors: analysis at steady state. *Toxicology and Applied Pharmacology* **60**, 509-26.

Andersen, M. E., Clewell, H. J. I., and Gargas, M. L. (1991). Physiologically-based pharmacokinetic modeling with dichloromethane, its metabolite carbon monoxide and blood carboxyhemoglobin in rats and humans. *Toxicol. Appl. Pharmacol.* **108**, 14-27.

Beliveau, M., Tardif, R., and Krishnan, K. (2003). Quantitative structure-property relationships for physiologically based pharmacokinetic modeling of volatile organic chemicals in rats. *Toxicol Appl Pharmacol* **189**, 221-32.

Beliveau, M., Tardif, R., and Krishnan, K. (2004). A spreadsheet program for conducting molecular structure-based simulations of the inhalation pharmacokinetics of volatile organic chemicals. *SAR and QSAR in Environmental Research*, In press.

Corley, R. A., Mandrela, A. L., and Smith, F. A. (1990). Development of a physiologically-based pharmacokinetic model for chloroform. *Toxicol. Appl Pharmacol.* **103**, 512-27.

Csanady, G. A., and Filser, J. G. (2001). The relevance of physical activity for the kinetics of inhaled gaseous substances. *Arch Toxicol* **74**, 663-72.

El-Masri, H. A., Thomas, R. S., Sabados, G. R., Phillips, J. K., Constan, A. A., Benjamin, S. A., Andersen, M. E., Mehendale, H. M., and Yang, R. S. H. (1996). Physiologically based pharmacokinetic/pharmacodynamic modeling of the toxicologic interaction between carbon tetrachloride and kepone. *Arch. Toxicol.* **70**, 704-13.

Fisher, J. W., Gargas, M. L., Jepson, G. W., Allen, B., and Andersen, M. E. (1991). Physiologically-based pharmacokinetic modeling with trichloroethylene and its metabolite, trichloroacetic acid in the rat and mouse. *Toxicol. Appl. Pharmacol.* **109**, 183-95.

Fouchecourt, M.-O., and Krishnan, K. (1999). Quantitative relationship between steady-state blood concnetrations and structural features of aliphatic hydrocarbons. *Toxicol. Lett.* **110**, 177-82.

Haddad, S., Charest-Tardif, G., Tardif, R., and Krishnan, K. (2000). Validation of a physiological modeling framework for simulating the toxicokinetics of chemicals in mixtures. *Toxicol. Appl. Pharmacol.* **167**, 199-209.

Krishnan, K., and Andersen, M. E. (2001). Physiologically based pharmacokinetic modeling in toxicology. In Principles and methods of toxicology (A. W. Hayes, ed., pp. 193-241. Taylor & Francis, Philadelphia.

Loizou, G. D. (2001). The application of physiologically based pharmacokinetic modelling in the analysis of occupational exposure to perchloroethylene. *Toxicol Lett.* **124**, 59-69.

Medinsky, M. A., Kenyon, E. M., Seaton, M. J., and Schlosser, P. M. (1996). Mechanistic considerations in benzene physiological model development. *Environ. Health Perspect.* **104**, 1399-404.

Paustenbach, D., Andersen, M. E., Clewell, H. J., and Gargas, M. L. (1988). A physiologically-based pharmacokinetic model for inhaled carbon tetrachloride in the rat. *Toxicol. Appl. Pharmacol.* **96**, 191-211.

Pelekis, M., Krewski, D., and Krishnan, K. (1997). Physiologically based algebraic expressions for predicting steady-state toxicokinetics of inhaled vapors. *Toxicol. Methods* **7**, 205-25.

Ramsey, J. C., and Andersen, M. E. (1984). A physiologically-based description of the inhalation pharmacokinetics of styrene in rats and humans. *Toxicol. Appl. Pharmacol.* **73**, 159-75.

Reitz, R. H., Gargas, M. L., Andersen, M. E., Provan, W. M., and Green, T. L. (1996). Predicting cancer risk from vinyl chloride exposure with a physiologically based pharmacokinetic model. *Toxicol. Appl. Pharmacol.* **137**, 253-67.

Reitz, R. H., Mandrela, A. L., Corley, R. A., Quast, J. F., Gargas, M. L., Andersen, M. E., Staats, D. A., and Conolly, R. B. (1990). Estimating the risk of liver cancer associated with human exposures to chloroform using physiologically-based pharmacokinetic modeling. *Toxicol. Appl. Pharmacol.* **105**, 443-59.

Sherwood, R. J., and Sinclair, G. C. (1999). New PBPK model applied to old occupational exposure to benzene. *AIHAJ.* **60**, 259-65.

Tardif, R., Charest-Tardif, G., Brodeur, J., and Krishnan, K. (1997). Physiologically based pharmacokinetic modeling of a ternary mixture of alkyl benzenes in rats and humans. *Toxicol. Appl.Pharmacol.* **144**, 120-34.

Ward, R. C., Travis, C. C., Hetrick, D. M., Andersen, M. E., and Gargas, M. L. (1988). Pharmacokinetics of tetrachloroethylene. *Toxicol. Appl. Pharmacol.* **93**, 108-17.

Table 1 Frequency of occurrence of molecular fragments in the volatile organic chemicals investigated in the present study.

Fragment / Chemical	CH_3	CH_2	CH	C	C=C	$H_{(on =)}$	CL	AC	$H_{(on AC)}$
Dichloromethane	0	1	0	0	0	0	2	0	0
Tetrachloroethylene	0	0	0	0	1	0	4	0	0
Toluene	1	0	0	0	0	0	0	1	5
m-Xylene	2	0	0	0	0	0	0	1	4
Styrene	0	0	0	0	1	3	0	1	5
Carbon tetrachloride	0	0	0	1	0	0	4	0	0
Ethyl benzene	1	1	0	0	0	0	0	1	5
Chloroform	0	0	1	0	0	0	3	0	0
Trichloroethylene	0	0	0	0	1	1	3	0	0
Vinyl chloride	0	0	0	0	1	3	1	0	0
Benzene	0	0	0	0	0	0	0	1	6
n	26	65	4	1	7	16	17	5	25

Table 2 Fragment specific unit contribution (C_f) to the blood:air partition coefficient (P_b) and hepatic clearance (CL_h) of volatile organic chemicals in rats and humans[a].

Fragments	Rat		Human	
	Log P_b	CL_h	Log P_b	CL_h
CH_3	0.0718	0.388	-4.82E-2	45.5
CH_2	0.109	-0.186	0.136	10.1
CH	0.0789	-0.464	2.62E-1	-14.3
C	-0.606	-1.44	-0.193	-73.7
C=C	-0.494	-1.71	-0.116	-16
$H_{(on =)}$	0.236	0.813	5.73E-2	19.5
Br	0.834	0.523	0.585	42.4
Cl	0.481	0.537	0.316	29.2
F	0.0203	-	-	-
AC	2.85	0.128	3.05	-211
$H_{(on AC)}$	-0.292	0.611	-0.347	49.3

[a] Data from Beliveau *et al.* (2003, 2004)

Table 3 Steady-state arterial blood concentrations (Cb_{ss}, µmol/L) in rats and humans exposed via inhalation (1 µmol/L air) to a series of volatile organic chemicals as estimated from validated PBPK models and the QSPR-based algorithm.

Chemical	Rat Cb_{ss}			Human Cb_{ss}		
	PBPK	QSPR	PBPK/ QSPR	PBPK	QSPR	PBPK/ QSPR
Dichloromethane	3.94	4.00	0.99	3.72	2.99	1.24
Tetrachloroethylene	10.48	8.43	1.24	8.44	3.20	2.64
Toluene	4.29	5.34	0.80	4.48	4.03	1.11
m-Xylene	3.97	4.38	0.91	3.73	4.71	0.79
Styrene	4.30	4.15	1.04	3.79	4.36	0.87
Carbon tetrachloride	3.07	5.57	0.55	2.28	5.31	0.43
Ethyl benzene	5.77	6.90	0.84	6.16	3.88	1.59
Chloroform	4.11	4.11	1.00	6.56	4.21	1.56
Trichloroethylene	3.60	5.05	0.71	3.92	2.88	1.36
Vinyl chloride	1.20	2.29	0.52	0.95	1.68	0.56
Benzene	3.52	5.82	0.60	2.94	3.22	0.91

Table 4 Rat to human steady-state arterial blood concentrations (Cb_{ss}, μmol/L) ratios for a series of volatile organic chemicals derived from the QSPR algorithm.

Chemical	QSPR-based algorithm rat Cb_{ss}[a]	QSPR-based algorithm human Cb_{ss}[a]	Rat/Human Cb_{ss} Ratios
Dichloromethane	4.00	2.99	1.34
Tetrachloroethylene	8.43	3.20	2.63
Toluene	5.34	4.03	1.33
m-Xylene	4.38	4.71	0.93
Styrene	4.15	4.36	0.95
Carbon tetrachloride	5.57	5.31	1.05
Ethyl benzene	6.90	3.88	1.78
Chloroform	4.11	4.21	0.98
Trichloroethylene	5.05	2.88	1.75
Vinyl chloride	2.29	1.68	1.36
Benzene	5.82	3.22	1.81

[a] Data from Table 3 for a 1 μmol/L air exposure.

CHAPITRE HUITIEME:

8 Discussion Générale

Présentement, la majorité des approches QSAR visent à établir la relation entre la structure moléculaire d'une série de substances analogues et une réponse toxique à ces substances (Hansch and Leo, 1995). L'indice de toxicité utilisé est souvent exprimé en terme d'exposition à une dose externe de la substance, c.-à-d., une dose effective causant une réponse donnée chez 50% des individus (DE_{50}) ou une dose sans effet nocif observé [DSENO (NOAEL)]. Puisque l'effet toxique n'est pas nécessairement relié à la dose externe administrée, mais plutôt à la concentration interne dans le tissu-cible, le fait d'être capable de relier la structure moléculaire à un indicateur de concentration tissulaire constituerait une amélioration par rapport aux approches existantes. Les modèles PCBP se fondent sur les mécanismes biologiques responsables de la toxicocinétique d'une substance dans l'organisme pour traduire une expostion externe en terme de concentration tissulaire interne, en tenant compte de l'espèce et du scénario d'exposition. C'est pourquoi leur utilisation en analyse du risque toxicologique est de plus en plus répandue. Les paramètres pharmacocinétiques nécessaires à l'élaboration de modèles PCBP pour de nouvelles substances sont à caractère physiologique, physicochimique, ou biochimique. Alors que les paramètres physiologiques sont souvent disponibles dans la littérature, les valeurs des paramètres physicochimiques et biochimiques, lesquels sont spécifiques à la substance étudiée, sont parfois plus difficiles à identifier. Puisqu'il n'est pas toujours possible ou efficace de determiner de façon expérimentale la valeur de ces paramètres pour chacune des substances, le dévelopment de modèles pour de nouvelles substances est présentement limité par la disponibilité de données sur ces paramètres.

Afin de pallier cette situation, une démarche visant à prédire les paramètres pharmacocinétiques d'une subtstance à partir de sa structure moléculaire peut s'avérer utile. Selon la quantité d'information requise ou disponible, l'une ou l'autre des démarches présentées dans les Chapitre 2 à 7 peut s'avérer applicable. Par exemple, lorsque la

concentration sanguine à l'état stationnaire est recherchée, le processus peut être simplifié à celui présenté dans les chapitres 6 et 7, puisqu'un modèle pharmacocinétique complet s'avère inutile dans ce contexte, le nombre de paramètres necessaires pour déterminer l'état stationnaire étant limité essentiellement au processus de clairance. L'exercice se limitera donc à associer la structure à la clairance systémique pour des substances similaires ou alors, à determiner un éventail de valeurs de clairances possibles en se basant sur les limites physiologiques.

Lorsque l'exposition ne conduit pas à l'état stationaire, l'utilisation d'un modèle QSAR-PCBP est nécessaire. Dans ce cas, l'approche utilisée dépendra du niveau d'information sur les relations structures-paramètres pharmacocinétiques disponibles dans la littérature. Trois situations peuvent se présenter. Premièrement, si la structure s'apparente à une famille générale de produits chimiques pour laquelle il existe déjà des QSARs pour les paramètres pharmacocinétiques établis dans une espèce précise, ceux-ci peuvent être directement utilisés avec un modèle PCBP afin de déterminer la concentration interne peu importe le scénario d'exposition. Par contre, puisque ces QSAR sont spécifiques à l'espèce chez laquelle ils ont étés déterminés, la concentration tissulaire qui en résulte ne peut être appliquée directement à d'autres espèces. Deuxièmement, pour extrapoler la concentration tissulaire à d'autres espèces, ou lorsqu'il n'est pas possible d'obtenir des QSAR spécifiques à chaque paramètre des modèles PCBP pour chaque espèce, la structure moléculaire peut être reliée à la composante qui est spécifique à la structure chimique pour un paramètre pharmacocinétique. Puisque cette composante, qui est spécifique à la structure (p.ex., Po:w ou Pw:a), reste inchangée peu importe l'espèce, la quantité d'information (sous forme de QSAR disponible) requise s'en trouve réduite. La composante physiologique de chaque paramètre, quant à elle, peut être adaptée à l'espèce que l'on désire étudier, et il est ainsi possible de relier et d'extrapoler la concentration interne à plusieurs espèces. En utilisant

cette approche, il est possible de simuler la concentration interne de façon simultanée chez plusieurs espèces représentatives d'un écosystème donné (par exemple, après le déversement accidentel d'une nouvelle substance dans l'environnement), et ce, même si l'on ne dispose comme information que de la structure moléculaire de la substance. De plus, grâce à la flexibilité de la modélisation PCBP, il est possible de tester un scénario hypothétique permettant de vérifier si la substance risque de s'accumuler dans la chaîne alimentaire, le seul point limitant étant la disponibilité de la valeur des paramètres physiologiques.

Troisièmement, si la disponibilité des données concernant les processus mécanistes des paramètres pharmacocinétiques permettant les approches mentionnées ci-dessus est limitée ou inexistante, il demeure possible d'estimer de façon préliminaire la concentration interne à l'état stationnaire. Cette information peut se présenter sous la forme d'une valeur précise ou d'une plage de valeurs (concernant la clairance pulmonaire ou hépatique, par exemple), selon les informations disponibles. Puisque les expositions de type environnementales mènent très souvent à des concentrations internes à l'état stationnaire, ce type d'information, même sommaire, peut-être d'une grande utilité en analyse du risque toxicologique même en absence de données cinétiques sur une substance.

La majorité des QSAR disponibles présentement dans la littérature se limitent à des familles de substances spécifiques (v. les Free-Wilson, par exemple), principalement parce que les propriétés (lipophiles, stériques ou électrostatiques) utilisées dans le développement qui en sont la base sont différentes d'une famille chimique à l'autre. Classer les différentes substances selon leur famille chimique limite le nombre de substances qui peuvent être modélisées et diminue ainsi la puissance du QSAR obtenu. Par exemple, dans le cas des COV, développer un QSAR pour chaque famille chimique signifie que les chloroéthylènes compteraient seulement six composés. Il est donc avantageux de regrouper

le plus de substances possibles sous un même QSAR, dans la mesure où l'indice modélisé le justifie. Pour ce qui est de l'approche développée dans l'article 1, en augmentant graduellement le nombre de carbones et le poids moléculaire (c.-à-d. en incluant de plus en plus de familles chimiques), certaines approches plus complexes seront nécessaires afin de tenir compte du manque de linéarité et du changement graduel des mécanismes impliqués, comme par exemple, la présence ou l'absence de liaison protéinique, la saturation de la liposolubilité ou un changement des isoenzymes responsables du métabolisme (les COV sont métabolisés principalement par le CYP2E1, alors que les hydrocarbures plus complexes auraient plusieurs voies métaboliques possibles). Par contre, lorsque les déterminants d'un paramètre pharmacocinétique sont suffisament connus, ils peuvent être facilement intégrés à l'algorithme bio-chimique approprié et à sa relation à la structure moléculaire tel que démontré précédemment (articles 2 à 6).

Ici, il importe de clarifier l'utilisation du terme mécanisme dans le contexte du présent travail. Globalement, ce terme réfère à l'ensemble des composantes biologiques reliées ensemble (sous la forme d'une équation ou d'une série d'équations mathématiques) déterminant par exemple la concentration sanguine. D'un point de vue pharmacocinétique, l'utilisation d'un modèle PCBP est mécaniste, puisque la concentration interne est établie à partir des composantes biologiques qui déterminent les processus d'absorption (p.ex., coefficient de partage sang:air), de distribution (p.ex., coefficient de partage tissu:sang), de métabolisme (p.ex., processus de clairance hépatique) et d'élimination (p.ex., constante d'élimination rénale de premier ordre). De même, l'utilisation de la solubilité dans l'eau et les lipides pour déterminer la valeur des coefficients de partage est un excercice mécaniste. Puisque le résultat obtenu est basé sur la valeur de ces composantes, une approche mécaniste n'implique pas une connaissance préalable de la valeur d'un paramètre modélisé. Toute relation qui établit un lien entre une fonction (sous forme d'observations

expérimentales) et des variables indépendantes est dite "empirique", peu importe si ces variables dépendent ou non de mécanismes biologiques/chimiques. Les modèles QSAR et pharmacocinétiques classiques sont par définition "empiriques", puisque qu'une connaisance préalable de la fonction cible (ou du paramètre modélisé) est requise afin d'établir la valeur des variables de l'équation. Par ailleurs, les modèles QSAR-PCBP développés dans cette étude (v. articles 1, 2, 3 et 4), sont un amalgame de relations empiriques (structure-paramètres, structure-solubilité) et mécanistes (coefficients de partages, concentrations internes). Par exemple, dans l'article 5, les concentrations internes chez le rat sont reliées de façon empirique à la structure moléculaire, alors que dans l'article 6, la concentration interne est prédite de façon mécaniste à l'aide de relations empiriques existant entre la structure et les composantes pharmacocinétiques qui définissent la clairance.

Les variables indépendantes qui sont généralement incluses dans les équations de régression QSAR de type LFE représentent souvent des mécanismes chimiques comme la distribution électronique ou le potentiel de liaison chimique (acide/base). Ainsi dans le cas où la valeur des coefficients dérivés est importante, les équations empiriques peuvent fournir d'importants indices sur la nature des mécanismes chimiques impliqués. Par contre, les études portant sur l'élaboration de "vrais" modèles chimiques qui font la relation entre la structure moléculaire et la solubilité dans l'eau ou l'octanol sans données expérimentales n'existent pas présentement (Sangster 1997; Baum 1998; Boethling et Mackay 2000).

Alors que les isomères de position sont relativement faciles à distinguer selon l'approche développée dans l'article 2, certains autres type d'isomères (cis/trans, énantiomères ou diastéréoisomères) le sont moins. L'approche utilisée le plus souvent est d'insérer une variable indépendante qui prend la valeur de 0 ou 1 selon la position spatiale d'un groupement (R ou S, par exemple). Afin d'avoir une certaine puissance statistique, des

données sur plusieurs isomères du même type (c.-à-d., cis/trans, *m-o-p-*, ou R/S) sont nécessaires. Au cours de cette étude, seuls un isomère cis/trans (1,2-dichloroéthylène), un isomère de position *m-o-p-* (xylène) et deux énantiomères (1,1,1-trifluoro-2-chloroéthane et 2,3,4-triméthylpentane) faisaient partie de la base de donnée utilisée, ce qui a limité l'insertion de variables indépendantes pouvant corriger ce genre de phénomène, la puissance statistique étant faible. L'importance potentielle de l'agencement spatial d'une molécule peut être illustrée en comparant les diverses valeurs disponibles. Alors que les $P_{o:a}$ des isomères o-, m- et p-xylene sont comparables (3534, 3245 et 3319), cela n'est pas le cas pour les *cis* et *trans*-1,2-dichloroéthylène (278 et 178). Ceci illustre bien l'importance d'intégrer ce genre d'information lorsque cela est possible. Certaines études devront sûrement porter sur l'approfondissement des connaissances concernant la relation entre l'agencement spatial et les diférents paramètres pharmacocinétiques.

Parmi les COV étudiés, le tétrachlorure de carbone, le 1,1,1-trichloroéthane, l'hexachloroéthane et le tétrachloroéthylène ont une faible extraction hépatique chez le rat ($E<0.5$). De ceux-ci, seuls le 1,1,1-trichloroéthane et le tétrachloroéthylène ne sont pas modélisés de façon adéquate par le QASR developpé dans l'article 2. Ce genre de difficulté a déjà été rapporté dans la littérature (Gargas et al., 1988), dans le cadre d'études portant sur des QSAR reliant structure et la valeur de V_{max} pour certains haloalcanes. Il importe de noter que le QSAR (chez le rat) présenté dans l'article 2 prédit correctement le faible coefficient d'extraction hépatique de ces deux substances ($E = 0.42$ et 0.30), même si les valeurs sont plus élevées que celle de la littérature ($CL_h = 0.02$ et 0.001). Il faut aussi noter que ce sont les deux seules substances parmi toutes celles étudiées qui sont métabolisées exclusivement par un processus d'ordre premier à des concentrations "environmentales". Malgré tout, cette divergence entre les clairances estimées et expérimentales n'a pas nécessairement d'impact sur le profil cinétique de ces substances (voir Figure 3C dans

l'article 2, par exemple). Ce genre d'observation est valable pour d'autres divergences au niveau des paramètres: l'impact net dépendra de l'importance relative du paramètre pharmacocinétique dans le modèle (c.-à-d., la sensibilité). Même si une étude de sensibilité a été réalisée (article 4), il n'en demeure pas moins que la relation entre la variation des fragments et la variation de la concentration interne dans le temps gagnerait à être approfondie, puisqu'il s'agit d'une relation plutôt complexe.

L'applicabilité des approches développées a été vérifiée pour les COV. Conceptuellement, les approches peuvent êtres appliquées au non-volatils. Par contre, pour ce faire, il faudrait prédire de façon adéquate certains paramètres pharmacocinétiques additionnels, tels que les coefficients de diffusion tissulaire, les constantes de liaison protéinique, les constantes d'absoption orales ou percutanées. Un algorithme mécaniste existe déjà pour la constante d'absorption percutanée (v. article 1), ce qui facilite l'établissement de la relation structure-paramètre. Plusieurs QSAR portant sur la constante d'absorption orale existent. Il serait interessant d'approfondir ce champ d'étude, puisque 1) les mécanismes qui contribuent directement à la valeur de la constante d'absorption orale sont encore méconnus et 2) les valeurs de plusieurs constantes d'absorption orale sont disponibles dans la littérature, particulièrement pour des produits pharmaceutiques.

Par ailleurs, les modèles QSAR utilisés dans les articles 2 à 6 permettent de réduire l'incertitude qui est associée à la mesure de dose interne déterminant la toxicité d'une substance. Par contre, peu importe l'approche utilisée, le résultat sera une valeur à laquelle un certain degré d'erreur est associé. En l'absence de valeurs expérimentales, des structures plus ou moins similaires à la substance étudiée sont utilisées pour estimer les valeurs des paramètres pharmacocinétiques. Quantifier cette erreur devient donc un exercise essentiel permettant d'estimer le niveau de confiance qui caractérise la valeur prédite d'un paramètre ou d'une variable. Les résultats des travaux présentés dans le Chapitre 4 sont un pas vers

l'élaboration d'une méthodologie permettant de quantifier l'erreur associée à l'estimation de la concentration interne appropriée. Lorsqu'on utilise un QSAR de fiabilité limitée pour estimer un paramètre sensible du modèle, l'incertitude concernant la valeur obtenue est élevée. En revanche, lorsqu'un QSAR robuste est utilisé pour estimer un paramètre sensible du modèle, l'incertitude diminue. De plus, en quantifiant la concentration interne appropriée en fonction des fragments moléculaires présents, l'incertitude concernant le QSAR lui-même peut être évaluée. Ainsi l'incertitude serait plus grande dans le cas où le fragment démontrant une grande sensibilité vis-à-vis la concentration tissulaire (comme par exemple, le fragment AC de l'éthyl benzène de l'Article IV) serait aussi celui générant la plus grande erreur dans le QSAR lui même. Une petite erreur dans l'estimation de la valeur du fragment en tant que tel aurait ainsi le potentiel de causer une erreur importante dans l'estimation de la concentration interne. L'évaluation de ces deux types de sensibilités est ainsi primordiale dans toute estimation QSAR-PCBP afin de quantifier l'erreur, et ainsi l'incertitude, associée à l'estimation de la concentration interne générée.

Les algorithmes bio-chimiques présents dans les articles 1 à 6 permettent de prédire la cinétique et l'accumulation de substances chez une gamme d'espèces (p.ex., poisson, rat, souris, humain), et ce, pour plusieurs scénarios d'exposition différents. La concentration interne qui correspond au risque unitaire ou à la DSENO chez une espèce peut donc être facilement extrapolée à une autre espèce en utilisant seulement l'information sur la structure moléculaire. Le caractère scientifique du processus d'analyse du risque en santé publique est ainsi rehaussé, surtout lorsque l'on compare cette nouvelle approche à l'utilisation actuelle des QSAR dans ce contexte (c.-à-d., estimations basées sur la dose externe et non interne). Toutes avancées dans le domaine des approches *in silico* en pharmacocinétique réduit le besoin d'expérimentations animales et accroit la fiabilité du processus d'analyse du risque en santé publique.

Globalement, cette étude a pour la première fois proposé une méthodogie qui permettra d'utiliser la structure moléculaire en analyse du risque toxicologique en utilisant des outils QSAR. Elle a démontré 1) comment les outils QSARs disponibles actuellement pourraient être utilisés en analyse du risque, 2) que les lacunes des méthodes QSAR actuelles pouvaient être corrigées en développant de nouvelles méthodologies mieux adaptées, 3) que ces nouvelles méthodologies s'adaptent selon l'espèce étudiée et l'entité interne appropriée, et finalement 4) que l'incertitude associée à l'estimation de cette entité interne doit être caractérisée afin de poser un jugement éclairé sur sa valeur scientifique.

CHAPITRE NEUVIEME:

9 Bibliographie

Abraham, M. H., et Weathersby, P. K. (1994). Hydrogen bonding. 30. Solubility of gases and vapors in biological liquids and tissues. *J. Pharm. Sci.* **83**, 1450-56.

Andersen, M. E., Clewell, H. J. I., Gargas, M. L., Smith, F. A., et Reitz, R. H. (1987). Physiologically-based pharmacokinetics and risk assessment process for methylene chloride. *Toxicol. Appl. Pharmacol.* **87**, 185-205.

Andersen, M. E., Gargas, M. L., Jones, R. A., and Jenkins, L. J. (1980). Determination of the kinetic constants for metabolism of inhaled toxicants in vivo by gas uptake measurements. *Toxicol. Appl. Pharmacol.* **54**, 116.

Arms, A. D., et Travis, C. C. (1988). Reference Physiological Parameters in Pharmacokinetic Modeling. Office of Health and Environmental Assessment. US EPA, Washington, DC.

Baum, E. J. (1998). *Chemical property estimations: Theory and Practice*. CRC Press LLC, Boca Raton, FL.

Boethling, R. S., and Mackay, D., eds. (2000). *Handbook of Property Estimation Methods for Chemicals: Environmental and Health Sciences*. CRC Press LLC, Boca Raton, FL.

Brown, R. P., Delp, M. D., Lindstedt, S. L., Rhomberg, L. R., et Beliles, R. P. (1997). Physiological parameter values for physiologically based pharmacokinetic models. *Toxicol. Ind. Health* **13**, 407-84.

Caster, W. O., Poncelet, J., Simon, A. B., et Armstrong, W. B. (1956). Tissue weights of the rat. I. Normal values determined by dissection and chemical methods. *Proc. Soc. Exper. Biol. Med.* **91**, 122-26.

Chen, H. S. G., et Gross, J. F. (1979). Estimation of tissue to plasma partition coefficients used in physiological pharmacokinetic models. *J. Pharmacokin. Biopharm.* **7**, 117-25.

Clewell, H. J. I., et Andersen, M. E. (1987). Dose, species and route extrapolation using physiologically-based pharmacokinetic models. *Drinking Water and Health* **8**, 159-82.

Delp, M. D., Manning, R. O., Bruckner, J. V., et Armstrong, R. B. (1991). Distribution of cardiac output during diurnal changes of activity in rats. *Am. J. Physiol.* **261**, H1487-93.

Domench, R. J., J.E., H., Noble, M. M., Saunder, K. B., Hensen, J. R., et Subijanto, S. (1969). Total and regional coronary blood flow measured by radioactive microsphere in conscious and anesthetized dogs. *Circ. Res.* **25**, 581-96.

Fouchécourt, M.-O., et Krishnan, K. (2000). A QSAR-type PBPK model for inhaled chloroethanes. *Toxicol. Sci.* **54**, 88.

Fouchécourt, M.-O., Walker, J., et Krishnan, K. (2000). An integrated QSAR-PBPK model for conducting rat-fish extrapolation of the biokinetics of chloroethanes. In Handbook on QSARs for Predicting Effects of Chemicals on Environmental-Human Health Interactions (J. Walker, ed. SETAC Press, Pensecola, FL.

Free, S. M., et Wilson, J. W. (1964). A mathematical contribution to structure-activity studies. *J. Med. Chem.* **7**, 395-99.

Gallo, J. M., Lam, F. C., et Perrier, D. G. (1987). Area method for the estimation of partition coefficients for physiological pharmacokinetic models. *J. Pharmacokin. Biopharm.* **15**, 271-80.

Gargas, M. L., Seybold, P. G., and Andersen, M. E. (1988). Modeling the tissue solubilities and metabolic rate constant (Vmax) of halogenated methanes, ethanes, and ethylenes. *Toxicol. Lett.* **43**, 235-56.

Gargas, M. L., et Andersen, M. E. (1989). Determinating the kinetic constants of chlorinated ethane metabolism in the rat from rates of exhalation. *Toxicol. Appl. Pharmacol.* **99**, 344-53.

Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H., et Andersen, M. E. (1989). Partition coefficients of low-molecular-weight volatile chemicals in various liquids and tissues. *Toxicol. Appl. Pharmacol.* **98**, 87-99.

Gargas, M. L., Clewell, H. J., et Etersen, M. E. (1986). Metabolism of inhaled dihalomethanes in vivo: Differentiation of kinetic constants for two independent pathways. *Toxicol. Appl. Pharmacol.* **82**, 211-23.

Gargas, M. L., Clewell, H. J., et Andersen, M. E. (1990). Gas uptake inhalation techniques and the rates of metabolism of chloromethanes, chloroethanes and chloroethylenes in the rat. *Inh. Toxicol.* **2**, 319.

Haddad, S., Withey, J. R., Laparé, S., Law, F. C. P., et Krishnan, K. (1998). Physiologically-based pharmacokinetic modeling of pyrene in the rat. *Environ. Toxicol. Pharmacol.* **5**, 245-55.

Hansch, C., et Leo, A. (1995). *Exploring QSAR*. American Chemical Society, Washington, D.C.

Hetrick, D. M., Jarabek, A. M., et Travis, C. C. (1991). Sensitivity analysis for physiologically-based pharmacokinetic models. *J. Pharmacokin. Biopharm.* **19**, 1-20.

Jepson, G. W., Hoover, D. K., Black, R. K., McCafferty, J. D., Mahle, D. A., et Gearhart, J. M. (1994). A partition coefficient determination method for nonvolatile chemicals in biological tissues. *Fund. Appl. Toxicol.* **22**, 519-24.

Kaneko, T., Wang, P. Y., et Sato, A. (1994). Partition coefficients of some acetate esters and alcohols in water, blood, olive oil, and rat tissues. *Occup. Environ. Med.* **51**, 68-72.

Krishnan, K., et Andersen, M. E. (1991). Interspecies scaling in pharmacokinetics. In New Trends in Pharmacokinetics (A. Rescigno et A. K. Thakkur, eds.), pp. 203-26. Plenum Press, New York.

Krishnan, K., et Andersen, M. E. (2001). Physiologically based pharmacokinetic modeling in toxicology. In Principles and methods of toxicology (A. W. Hayes, ed., pp. 193-241. Taylor & Francis, Philadelphia.

Lam, G., Chen, M. L., et Chiou, W. L. (1982). Determination of tissue: blood partition coefficients in physiologically-based pharmacokinetic models. *J. Pharma. Sci.* **71**, 454-56.

Law, F. C. P., Abeini, S., et Kennedy, C. J. (1991). A biologically-based toxicokinetic model for pyrene in rainbow trout. *Toxicol. Appl. Pharmacol.* **110**, 390-402.

Lin, J. H., Hayashi, M., Awazu, S., et Hanano, M. (1978). Correlation between in vitro and in vivo drug metabolism rate: Oxidation of ethoxybenzamide in rat. *J. Pharmacokin. Biopharm.* **6**, 327-37.

Lin, J. H., Sugiyama, Y., Awazu, S., et Hanano, M. (1982). In vitro and in vivo evaluation of the tissue to blood partition coefficients for physiological pharmacokinetic models. *J. Pharmacokin. Biopharm.* **10**, 637-47.

Mauderly, J. L. (1990). Measurement of respiration and respiratory responses during inhalation exposures. *J. Am. Coll. Toxicol.* **9**, 397-406.

Murphy, J. E., Janszen, D. B., et Gargas, M. L. (1995). An in vitro method for determination of tissue partition coefficients of non-volatile chemicals such as 2,3,7,8-tetrachlorodibenzo-p-dioxin and estradiol. *J. Appl. Toxicol.* **15**, 147-52.

Poulin, P., et Krishnan, K. (1998). A quantitative structure-toxicokinetic relationship model for highly metabolised chemicals. *ATLA* **26**, 45-59.

Ross, R., Leger, L., Guardo, R., de Guise, J., et Pike, B. G. (1991). Adipose tissue volumes measured by magnetic resonance imaging and computerized tomography in rats. *J. Appl. Physiol.* **70**, 2164-72.

Sangster, J. (1997). *Octanol-Water Partition Coefficients: Fundamentals and Physical Chemistry*. John Wiley &Sons, Chichester, England.

Sato, A., et Nakajima, T. (1979). Partition coefficients of some aromatic hydrocarbons and ketones in water, blood and oil. *Br. J. Ind. Med.* **36**, 231-34.

Seydel, J. K., et Schaper, K. J. (1982). Quantitative structure-activity relationship of some chlorinated hydrocarbons. *Pharmacol. Ther.* **15**, 131-82.

Tardif, R., Charest-Tardif, G., Brodeur, J., et Krishnan, K. (1997). Physiologically based pharmacokinetic modeling of a ternary mixture of alkyl benzenes in rats and humans. *Toxicol. Appl. Pharmacol.* **144**, 120-34.