Estimació de propietats de biopartició mitjançant mesures cromatogràfiques

Marta Hidalgo Rodríguez

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) i a través del Dipòsit Digital de la UB (**diposit.ub.edu**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) y a través del Repositorio Digital de la UB (**diposit.ub.edu**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service and by the UB Digital Repository (**diposit.ub.edu**) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Universitat de Barcelona
Facultat de Química
Departament de Química Analítica
Programa de doctorat: Química Analítica del Medi Ambient i la Pol·lució

Estimació de propietats de biopartició mitjançant mesures cromatogràfiques

Marta Hidalgo Rodríguez

Memòria de tesi doctoral presentada per a optar
al grau de Doctor per la Universitat de Barcelona

Sota la direcció de:
Dr. Martí Rosés i Pascual, Dra. Elisabet Fuguet i Jordà
del Departament de Química Analítica de la Universitat de Barcelona

Barcelona, setembre de 2013

El Dr. Martí Rosés i Pascual i la Dra. Elisabet Fuguet i Jordà

CERTIFIQUEN

Que la present memòria de tesi doctoral, que porta per títol “Estimació de propietats de biopartició mitjançant mesures cromatogràfiques”, ha estat realitzada sota la seva direcció per la Sra. Marta Hidalgo Rodríguez al Departament de Química Analítica de la Universitat de Barcelona i que tots els resultats presentats són fruit de la recerca realitzada per l’esmentada doctoranda.

Barcelona, setembre de 2013

Dr. Martí Rosés i Pascual

Catedràtic del
Departament de Química Analítica
de la Universitat de Barcelona

Dra. Elisabet Fuguet i Jordà

Professora Agregada Interina
del Departament de Química Analítica
de la Universitat de Barcelona

Als meus pares,

AGRAÏMENTS

La realització d'aquesta tesi ha marcat una etapa de la meva vida, de la qual em sento satisfeta i afortunada. No només han estat anys d'esforç i dedicació els que m'han conduït a assolir aquest repte, sinó també el suport i l'empenta de molta gent a la que vull expressar la meva gratitud.

En primer lloc, vull donar les gràcies als meus directors de tesi, el Dr. Martí Rosés i la Dra. Elisabet Fuguet. Us agraeixo tots els coneixements científics que m'heu aportat, així com haver-me ensenyat a fer recerca. També la confiança que heu depositat en mi, els vostres consells i la bona relació que hem mantingut i espero seguir mantenint.

També vull agrair a la Dra. Clara Ràfols les seves aportacions científiques a la tesi, a més del seu recolzament al llarg d'aquests anys.

Gràcies també a la Dra. Elisabeth Bosch per haver-me admès en el grup de recerca, per interessar-se pels meus progressos i per les seves valuoses orientacions.

I en conjunt, vull agrair a tots quatre haver-me fet sentir tan ben acollida des del primer dia que vaig trepitjar el laboratori. A part, he tingut el plaer de compartir amb tots vosaltres l'assistència a congressos, dels que guardo records imborrables.

El fet que hagi estat tan a gust al grup de recerca també té molt a veure amb el bon ambient que s'ha respirat entre els companys de laboratori. Molts companys que han anat canviant al llarg dels anys i que si em posés a fer una llista potser em deixaria a algú i seria injust. No obstant, sí que m'agradaria fer una menció especial a aquells amb els que he mantingut una relació més estreta. Juanma, moltíssimes gràcies per tot, fer la tesi en paral·lel ens ha permès compartir molts moments i considero que ha estat un gran privilegi. Cristina, crec que recordarem durant molt temps algunes anècdotes que vam viure juntes al laboratori, moltes gràcies per les estones que hem passat, així com pels teus ànims. Joan Marc, gràcies per les teves xerrades i el teu caràcter animat que amenitzen la feina. Elham, gràcies a tu hi ha hagut moments de laboratori difícils d'oblidar. Rufielyn, I am very happy you came for some months to the lab so we had the opportunity to meet each other. Eli Lázaro, gràcies per la teva predisposició a ajudar-me sempre que t'ho he demanat i per les estones que hem compartit. Sara, gràcies per la teva ajuda durant els últims mesos.

Del departament, també vull fer una menció especial a la Sílvia. Ha estat una gran sort comptar amb algú de confiança fora del grup, així com compartir els moments de finalització de la tesi. Ànims, que ja queda menys per fer una celebració de doctores.

Agraeixo també l'ajuda del personal de secretaria en temes administratius, especialment a la Carme i a la Manoli.

I would also like to thank Dr. Holger Fischer and Dr. Manfred Kansy for giving me the opportunity to do a predoctoral stay in their research group of F. Hoffmann-La Roche (Basel). It was a priceless experience. On one hand, not only I learned new concepts and techniques, but also I was able to see how an important pharmaceutical company works. I want to thank my lab colleagues (Frauke, Severin, Björn, Isabelle, Virginie) for making me feel at home. On the other hand, I had the chance of living abroad, meeting lots of people and travelling around the wonderful Switzerland.

Agraeixo també al Dr. Enrique Ferrero i al Dr. Jorge Malfeito l'haver-me permès formar part del grup de I+D+i d'Acciona Agua durant aquest últim any. Compaginar el treball en aquesta empresa amb l'escriptura de la tesi ha fet que fos un any intens, no obstant ha merescut la pena. A més del gran aprenentatge que m'ha suposat, la investigació que realitzen m'ha semblat d'allò més interessant i he pogut treballar amb bellíssimes persones. Moltes gràcies per tot. Un agraïment especial a la Carme, amb qui ja vaig compartir laboratori a la facultat i amb qui ha estat un plaer tornar a compartir-lo a l'empresa, gràcies per la teva ajuda i els teus ànims.

Molts dels companys que he anomenat s'han convertit en bons amics, però deixant de banda el món científic també tinc grans amics a qui vull donar les gràcies. Mary, Alicia, Cristina, Oriol, Javi, Alex, Cristal, Marta, Núria, Juanca, Héctor, Edgar, Carol, Nat, Marti, moltes gràcies pel vostre suport i pels moments de desconnexió que m'heu ofert. Uns moments imprescindibles per agafar energies i tirar endavant, estic encantada de poder compartir vivències amb vosaltres.

I com no, un dels meus majors agraïments és per la meva família. Pares, germà, neboda, cunyada, tiets, cosins, avis, gràcies pel vostre interès i recolzament. En particular, dedico aquesta tesi als meus pares per haver-me proporcionat el camí per assolir aquest repte. Al meu pare li regalo el més gran dels meus somriures com a satisfacció d'haver aconseguit aquest objectiu, perquè sé que és el millor regal que et puc fer. La meva mare ha estat un pilar fonamental, moltíssimes gràcies per estar

sempre pendent de mi i de l'avanç de la tesi, pel coratge que m'encomanes dia rere dia i pel teu suport incondicional. Les teves paraules han estat i segueixen sent molt importants per mi.

L'altre dels meus majors agraïments és per l'Axel. Les estones que hem passat junts al laboratori em deixen grans records. Però sobretot, gràcies infinites per haver-me proporcionat el luxe de descobrir com ets durant aquests anys. Després de tot aquest temps, t'has convertit en una persona indispensable a la meva vida i sense les teves incansables dosis de tranquil·litat, el teu inestimable suport i la teva ajuda aquesta tesi hagués estat molt més difícil. Moltíssimes gràcies per estar sempre al meu costat.

ABREVIATURES I ACRÒNIMS

a	Basicitat per pont d'hidrogen del medi
A	Acidesa per pont d'hidrogen del solut
a_u	Coeficient *a* normalitzat
ADMET	Absorció, distribució, metabolisme, excreció i toxicitat
b	Acidesa per pont d'hidrogen del medi
B	Basicitat per pont d'hidrogen del solut
b_u	Coeficient *b* normalitzat
BB	Coeficient de distribució sang-cervell
bio	Biològic
Brij 35	Èter de dodecilpolioxietilè(23)
c	Constant de la correlació
C18	Fase estacionària d'octadecilsilà
C_{nar}	Concentració de narcosi en capgrossos
C_o	Concentració de solut a la fase octanòlica
C_w	Concentració de solut a la fase aquosa
cal	Calculat
CHAPS	3-[(3-colamidopropil)dimetilamoni]-1-propansulfonat
CHAPSO	3-[(3-colamidopropil)dimetilamoni]-2-hidroxi-1-propansulfonat
CHI	Índex d'hidrofobicitat cromatogràfica
CMC	Concentració micel·lar crítica
Col	Colesterol
cos θ	Cosinus de l'angle que formen dos vectors LSER
crom	Cromatogràfic

d	Distància entre dos vectors LSER normalitzats
D'	Distància entre dos vectors LSER
DAD	Detector de fila de díodes (*diode array*)
DPPC	Dipalmitoilfosfatidilcolina
DPPG	Dipalmitoilfosfatidilglicerol
DTAC	Clorur de dodeciltrimetilamoni
e	Capacitat del medi per interaccionar amb els electrons –n i –π del solut
E	Refractivitat molar en excés del solut
e_u	Coeficient *e* normalitzat
EKC	Cromatografia electrocinètica
EOF	Flux electroosmòtic
EPA	Agència de protecció mediambiental dels Estats Units d'Amèrica
exp	Experimental
F	Paràmetre estadístic F de Fisher
HIA	Absorció intestinal humana
HPLC	Cromatografia de líquids d'alta resolució
HTAB	Bromur d'hexadeciltrimetilamoni
IAM	Columna d'HPLC de membrana artificial immobilitzada
k	Factor de retenció d'una substància en un sistema cromatogràfic
K_{OC}	Coeficient de sorció sòl-aigua, normalitzat pel contingut de carboni orgànic del sòl
K_p	Coeficient de permeabilitat a través de la pell
k_w	Factor de retenció en cromatografia de líquids extrapolat a aigua pura
l	Longitud d'un vector LSER
LC_{50}	Concentració de compost que provoca la mort del 50% de la població

LDS	Dodecilsulfat de liti
LEKC	Cromatografia electrocinètica de liposomes
LFER	Relacions lineals de l'energia lliure
LPFOS	Perfluorooctanosulfonat de liti
LSER	Relacions lineals de l'energia lliure de solvatació
MeCN	Acetonitril
MEEKC	Cromatografia electrocinètica de microemulsions
MEKC	Cromatografia electrocinètica micel·lar
MELC	Cromatografia de líquids de microemulsions
MLC	Cromatografia de líquids micel·lar
MLR	Regressió lineal múltiple
MSC18	Columna d'HPLC XTerra MSC18
n	Número de punts de la correlació
p	Pendent de la correlació
P	Coeficient de partició d'una substància entre dues fases líquides immiscibles
$P_{o/w}$	Coeficient de partició octanol-aigua d'una substància
PAAU	poli(11-acrilamidaundecanoat de sodi)
PC	Component principal
PCA	Anàlisi per components principals
poli-EKC	Cromatografia electrocinètica de polímers
poli-(SDeS)	poli(9-decenilsulfat de sodi)
pred	Predit
q	Ordenada a l'origen de la correlació
QSAR	Relacions quantitatives estructura-activitat

r	Coeficient de la correlació
RP18	Columna d'HPLC XTerra RP18
RP-HPLC	Cromatografia de líquids d'alta resolució en fase invertida
s	Capacitat del medi per establir interaccions dipol-dipol i dipol-dipol induït
S	Dipolaritat/polaritzabilitat del solut
s_u	Coeficient *s* normalitzat
σ_{bio}^2	Variància deguda a la precisió de les dades biològiques
σ_{corr}^2	Variància total de la correlació biològico-cromatogràfica
σ_{crom}^2	Variància deguda a la precisió de les dades cromatogràfiques
σ_d^2	Variància deguda a la dissimilitud entre la parella de sistemes biològic i cromatogràfic comparats
SC	Colat de sodi
SD	Desviació estàndard
SD_{bio}^2	Estimació de la variància deguda a la precisió de les dades biològiques
$SD_{corr\ cal}^2$	Estimació de la variància total de la correlació biològico-cromatogràfica
$SD_{corr\ exp}^2$	Variància total experimental obtinguda a la correlació biològico-cromatogràfica
SD_{crom}^2	Estimació de la variància deguda a la precisió de les dades cromatogràfiques
SD_d^2	Estimació de la variància deguda a la dissimilitud entre la parella de sistemes biològic i cromatogràfic comparats
SDC	Deoxicolat de sodi
SDS	Dodecilsulfat de sodi
SP	Propietat del solut en un sistema de partició, relacionada amb l'energia lliure de solvatació

STC	Taurocolat de sodi
t_0	Temps mort en HPLC
t_{eof}	Temps d'elució del marcador de flux electroosmòtic en MEKC
t_{mc}	Temps d'elució del marcador micel·lar en MEKC
t_R	Temps de retenció d'una substància en un sistema cromatogràfic
Triton X-100	Èter d'octilfenilpolioxietilè(10)
TTAB	Bromur de tetradeciltrimetilamoni
UV	Ultraviolat
v	Hidrofobicitat del medi
V	Volum de McGowan del solut
v_u	Coeficient v normalitzat
$\vec{w}$	Vector de coeficients LSER
$\vec{w}_u$	Vector de coeficients LSER normalitzats
x_i	Fracció molar de i

ÍNDEX

OBJECTIU 1

ESTRUCTURA DE LA TESI 5

ABSTRACT 11

1. INTRODUCCIÓ 17

1.1. Sistemes de biopartició 19

1.2. Sistemes fisicoquímics per emular sistemes de biopartició 25

1.2.1. Sistema de partició octanol-aigua 25

1.2.2. Sistemes cromatogràfics 27

1.2.2.1. Sistemes de cromatografia de líquids d'alta resolució 27

1.2.2.2. Sistemes de cromatografia electrocinètica micel·lar 31

1.2.2.3. Altres sistemes cromatogràfics 35

1.3. Model de paràmetres de solvatació 37

1.3.1. Perspectiva general dins els estudis QSAR 37

1.3.2. Equació general de solvatació 41

1.3.2.1. Aplicació a sistemes de biopartició 45

1.3.2.2. Aplicació a sistemes fisicoquímics 47

1.4. Similitud entre sistemes biològics i fisicoquímics de partició 51

2. PUBLICACIONS 59

Article I. *Solute-solvent interactions in micellar electrokinetic chromatography: VII. Characterization of sodium cholate-sodium deoxycholate mixed-micellar systems* 63

Article II. *Estimation of biological properties by means of chromatographic systems: evaluation of the factors that contribute to the variance of biological-chromatographic correlations* 71

Article III. *Modeling nonspecific toxicity of organic compounds to the fathead minnow fish by means of chromatographic systems* 81

Article IV. *Performance of chromatographic systems to model soil-water sorption* 97

Article V. *Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds* 107

3. RESULTATS I DISCUSSIÓ 119

3.1. Caracterització de sistemes de MEKC mitjançant l'equació general de solvatació 121

3.1.1. Sistemes amb fases micel·lars mixtes de SC i SDC 122

3.1.2. Sistema de STC 131

3.2. Recopilació bibliogràfica de sistemes caracteritzats segons el model de paràmetres de solvatació 137

3.3. Estimació de propietats de biopartició a partir de mesures cromatogràfiques d'HPLC i MEKC 143

3.3.1. Desenvolupament d'una metodologia per avaluar la bondat de les correlacions biològico-cromatogràfiques 143

3.3.2. Modelació de sistemes de biopartició concrets: toxicitat en el peix *fathead minnow*, sorció sòl-aigua i permeabilitat a través de la pell 156

3.3.2.1. Selecció de possibles sistemes cromatogràfics per modelar-los 156

3.3.2.2. Correlacions experimentals 164

3.3.2.3. Aplicacions il·lustratives: estimació de valors de K_{OC} i K_p 180

CONCLUSIONS 187

REFERÈNCIES 193

SUPPORTING INFORMATION 211

OBJECTIU

L'objectiu principal d'aquesta tesi doctoral ha estat estudiar l'emulació de processos biològics de partició a partir de sistemes cromatogràfics. Per tal de disposar de sistemes cromatogràfics que puguin ser útils per a aquest propòsit, s'ha pretès caracteritzar nous sistemes cromatogràfics mitjançant el model de paràmetres de solvatació, un model multiparamètric que permet conèixer com afecta cada tipus d'interacció al procés que experimenten els soluts neutres en els sistemes de partició. També s'ha pretès recopilar bibliogràficament tant sistemes cromatogràfics com de biopartició caracteritzats mitjançant aquest model, fent-lo servir com a marc comú per comparar ambdós tipus de processos. La finalitat ha estat avaluar la capacitat dels sistemes cromatogràfics per estimar propietats de biopartició, tot examinant les correlacions biològico-cromatogràfiques i proposant una metodologia per identificar els sistemes cromatogràfics més adequats per modelar processos biològics. Amb l'aplicació d'aquesta metodologia a sistemes seleccionats com són la toxicitat en el peix *fathead minnow*, la sorció sòl-aigua i la permeabilitat a la pell de substàncies neutres, s'ha volgut demostrar el potencial dels sistemes cromatogràfics per estimar propietats biològiques a partir de mesures ràpides, senzilles i econòmiques, oferint una alternativa pràctica i sostenible a la seva determinació experimental en éssers vius.

ESTRUCTURA DE LA TESI

Aquesta memòria de tesi doctoral es presenta com a compendi de publicacions. La memòria es troba principalment dividida en tres parts: una introducció general, una còpia completa dels articles publicats i una discussió global dels resultats obtinguts.

A la introducció, es posa en context la línia de recerca seguida en el transcurs de la tesi. Inicialment, es parla dels sistemes biològics de partició. Es detallen les seves característiques fonamentals, se'n posen exemples i s'expliquen breument els tres sistemes que es voldran emular: la toxicitat en el peix *fathead minnow*, la sorció sòl-aigua i la permeabilitat en pell de substàncies neutres. S'exposa la dificultat de determinar experimentalment propietats de biopartició i la conseqüent necessitat de tenir a l'abast mètodes de predicció. Es fa referència a la possibilitat d'estimar dades biològiques a partir de sistemes fisicoquímics.

Seguidament, es presenten els tres tipus de sistemes fisicoquímics de partició en els que s'ha focalitzat l'estudi d'emulació de sistemes biològics: el sistema de partició octanol-aigua i els sistemes cromatogràfics d'HPLC i de MEKC. No només s'expliquen breument els seus fonaments sinó que també es fa referència a la seva aptitud i als seus avantatges per modelar processos de biopartició.

La resta de la introducció es dedica a explicar el model de paràmetres de solvatació i la seva utilitat per caracteritzar sistemes de partició tant de caràcter biològic com fisicoquímic. Es mostren múltiples exemples de caracteritzacions que es discuteixen vers la informació que proporciona el model. Finalment, es detalla com avaluar la semblança entre sistemes biològics i fisicoquímics a partir de diferents eines basades en aquest model. No obstant, es fa palesa la necessitat de disposar d'eines complementàries que permetin estimar la bondat amb què un sistema fisicoquímic pot emular un de biològic, el que permetrà identificar els sistemes més adients. La introducció acaba d'aquesta manera, havent presentat la temàtica de la tesi i enllaçant amb el treball de recerca efectuat.

A continuació, es presenten les cinc publicacions fruit de la investigació duta a terme durant el doctorat. Constitueixen el nucli principal de la memòria. Apareixen per ordre cronològic de publicació i és on es poden consultar els detalls experimentals i els resultats de cadascun dels treballs.

A la tercera part, la de resultats i discussió, es presenten i es raonen de manera conjunta els resultats més rellevants dels cinc articles publicats. També s'inclouen alguns aspectes no publicats que han contribuït a la recerca efectuada.

El primer apartat està dedicat a la caracterització de sistemes cromatogràfics, concretament de MEKC, mitjançant el model de paràmetres de solvatació. La caracterització de sistemes mixtos de colat i deoxicolat de sodi correspon a l'article I. La caracterització del sistema de taurocolat de sodi no es troba específicament publicada però es fa servir als articles III-V.

El segon apartat presenta la compilació de sistemes biològics i fisicoquímics de partició que es troben caracteritzats a la literatura mitjançant el model de paràmetres de solvatació. Els dos tipus de sistemes es comparen mitjançant aquest model i es justifica l'elecció dels tres sistemes biològics a estudiar: la toxicitat en el peix *fathead minnow*, la sorció sòl-aigua i la permeabilitat en pell de compostos neutres. Aquesta feina tampoc es troba explícitament publicada en tota la seva extensió, però part de la informació que s'extreu condueix i es fa servir a les publicacions II-V.

El tercer i últim apartat fa referència als estudis d'estimació de propietats biològiques mitjançant mesures de MEKC i HPLC, el que correspon als articles II-V i dóna lloc al títol de la tesi. D'una banda, es presenta la metodologia desenvolupada per estimar la capacitat predictiva de les correlacions biològico-cromatogràfiques, essent un factor clau per identificar quins sistemes cromatogràfics són adequats per modelar processos de biopartició (article II). D'altra banda, aquesta metodologia es fa servir per seleccionar sistemes cromatogràfics que emulin els tres processos biològics escollits a l'apartat previ. Posteriorment, es duen a terme comprovacions experimentals. A més, també es mostren aplicacions pràctiques d'estimació de valors de sorció sòl-aigua i permeabilitat en pell de soluts neutres a partir de mesures cromatogràfiques (articles III-V).

Les conclusions més importants que s'extreuen de tots els resultats obtinguts es poden consultar a continuació de les tres parts de la memòria mencionades.

Cal esmentar que aquesta tesi opta a la menció internacional de doctorat per haver realitzat una estada de recerca de sis mesos a la seu de l'empresa farmacèutica

F. Hoffmann-La Roche, a Basilea (Suïssa). En aquesta estada, es va posar a punt un nou assaig *in vitro* per estudiar la distribució de fàrmacs en teixits, utilitzant concretament teixits de rata. Un cop optimitzat, l'assaig es va fer servir per determinar la distribució d'un ampli ventall de substàncies farmacèutiques en el cervell, el fetge, el greix i els testicles. Els resultats es van comparar amb els d'altres assajos *in vitro* i *in vivo*.

La temàtica de l'estada efectuada en aquest centre europeu s'emmarca perfectament dins el projecte de la tesi, doncs complementa l'estudi de sistemes de partició i la cerca de metodologies per conèixer propietats biològiques sense necessitat de determinar-les directament en éssers vius. No obstant, tot i que es preveu una publicació al respecte, l'empresa encara necessitarà més temps per completar l'estudi amb altres teixits. Per això, donat que aquesta memòria es presenta com a compendi de publicacions, no s'ha pogut incloure la part de recerca referent a l'estada.

ABSTRACT

The title of the present thesis is *Estimation of biopartitioning properties by means of chromatographic measurements*.

The main goal of the thesis is showing the potential of chromatographic systems to model biopartitioning processes, which is a topic of great interest due to the following reasons. Experimental determination of biological properties, including toxicological and environmental ones, usually involves very expensive, time-consuming and difficult procedures because of the complexity of biological samples. Sometimes, these procedures are even ethically questionable. As a consequence, estimation methods from simple measurements are very valuable. In this sense, chromatographic systems can provide practical and sustainable approaches since their measurements are easy, fast, economic and reliable.

The key issue is finding which chromatographic systems emulate well biopartitioning processes. The solvation parameter model is very useful for this purpose since it allows the characterization and the comparison of both physicochemical and biological types of partitioning systems. Any type of system ruled by the passive transport of neutral solutes between two phases can be characterized by this model. The characterization provides insight into the types and the relative strength of the chemical interactions that control the partitioning process. Besides, it is possible to compare different systems by means of their characterizations. The interest lies on identifying similar biological and chromatographic systems. In this case, good correlations are expected between them, making then possible the estimation of the biopartitioning property by means of chromatographic measurements in the directly related system.

The research carried out in the thesis is based on this subject. As a result, five articles have been published in peer-reviewed journals of significant impact factor. These publications are the basis of the work exposed here.

Several chromatographic systems have been initially characterized using the solvation parameter model. They consist of MEKC systems of sodium cholate-sodium deoxycholate mixtures on one hand and pure sodium taurocholate on the other hand. The study of the mixed-micellar systems is reported in the first article. This work not only pays attention to the influence of the surfactant composition in the mixtures, but

also deals with the effect of the set of solutes, the descriptors and the pH of the buffer on the characterizations. Regarding the sodium taurocholate MEKC system, it has been characterized because of its similarity to numerous biological processes. The corresponding characterization has been used for the studies reported in the articles III-V.

The main part of the thesis has been dedicated to study the performance of chromatographic systems to emulate biological ones. First, it has required the compilation from literature of biological and physicochemical partitioning systems characterized by the solvation parameter model. The compiled biological systems have very different nature: blood-brain distribution, skin permeation and tadpole narcosis are only few examples. The compiled physicochemical systems, mainly chromatographic, can be divided into three groups: solvent-water partition systems, electrokinetic chromatography systems and high-performance liquid chromatography systems. The similarity between each biological and physicochemical system has been evaluated according to a distance parameter related to the characterization of the systems by the solvation parameter model. However, in order to identify conveniently which physicochemical systems can model biological ones, it has been observed the need to quantify the precision that could be achieved in the correlations of biological and physicochemical systems. These circumstances have led to the articles II-V.

Thus, a systematic procedure has been developed to evaluate the goodness that could be obtained in the emulation of a biological system by a chromatographic one. It is based on the examination of the three sources of error that determine the overall variance of the correlations between biological and chromatographic systems: the variance of the biological data, the variance of the chromatographic data and the variance attributed to the dissimilarity between the two correlated systems. These contributions of variance can be easily predicted through the characterization of the systems by the solvation parameter model. Their sum estimates the overall variance expected for the correlation, which indicates the precision that could be obtained in the estimation of the biological property by means of the chromatographic measurement.

This procedure has been presented in the second article, proving also its usefulness by some illustrative examples of correlations between biological and MEKC systems. After applying the mentioned methodology to the correlations, the calculated variances have been compared to the experimental ones.

The thesis has continued focusing on the estimation of three particular biopartitioning properties by means of chromatographic measurements: the nonspecific toxicity to the *fathead minnow* fish, the soil-water sorption and the human skin permeation of neutral organic compounds, which correspond to the articles III-V respectively. The systematic procedure mentioned above has been applied to evaluate the suitability of several chromatographic systems to estimate these biopartitioning properties. According to this procedure, the most promising systems of HPLC and MEKC, in addition to the reference octanol-water partition system, have been selected to finally test their performances by experimental correlations. The chosen chromatographic systems are: three HPLC systems based on two conventional C18 columns and one immobilized artificial membrane column; two MEKC systems of sodium dodecylsulfate on one hand and sodium taurocholate on the other hand.

For the modeling of human skin permeation, a correction using the volume of the solute has been required in the correlations. The introduction of this new variable (i.e. the volume of the solute in the biological-chromatographic correlations) has been properly discussed in the fifth article.

Last but not least, the best biological-chromatographic correlation models obtained have been used for the prediction of soil-water sorption values and skin permeability data of neutral compounds. The precision achieved has been compared to other approaches used for the estimation of biopartitioning properties, for example calculation from structure.

In order to allow English readers to understand the thesis, the main conclusions reached are also written in English, in addition to this abstract, the five published papers and the supporting information.

INTRODUCCIÓ

1.1. SISTEMES DE BIOPARTICIÓ

Els processos de biopartició es basen en la distribució de substàncies entre una fase aquosa i una fase orgànica de caràcter biològic, per exemple una membrana cel·lular o un sòl. Aquestes dues fases formen el que anomenem sistemes de biopartició.

La distribució que experimenta un compost en un sistema de biopartició és significativament complexa perquè la composició exacta de les fases sol estar poc definida i la disposició espacial és rarament homogènia. En conseqüència, els experiments en sistemes de biopartició són normalment difícils per raons tècniques, també solen requerir temps i costos elevats i, a més, sovint plantegen qüestions ètiques referents a l'experimentació amb éssers vius. D'altra banda, aquests experiments permeten conèixer propietats de biopartició de compostos que són claus en molts àmbits, per exemple, a l'hora de desenvolupar nous productes com fàrmacs o bé a l'hora d'avaluar la toxicitat i el risc mediambiental de compostos d'origen industrial o substàncies emergents que poden ser contaminants.

Existeixen processos de biopartició de naturalesa molt diversa. L'absorció intestinal [1], la distribució de fàrmacs entre la sang i òrgans com el cervell [2] o la narcosi en capgrossos [3] són alguns dels que han estat àmpliament estudiats. Independentment del seu caràcter farmacològic o ecotoxicològic, ens referim a ells com a sistemes de biopartició o biològics. Altres exemples són la toxicitat de compostos en organismes aquàtics, la sorció de substàncies en sòls i la permeabilitat de soluts a través de la pell, dels quals es parlarà més detalladament a continuació.

Els estudis de toxicitat aquàtica tenen com a objectiu determinar si compostos presents a les aigües particionen vers organismes aquàtics exhibint algun efecte tòxic. Es tracta d'estudis d'importància vital per avaluar el dany mediambiental que poden ocasionar els vessaments de productes químics a les aigües [4]. Existeixen nombrosos protocols experimentals i mètodes estàndard d'anàlisi per determinar la toxicitat aquàtica de substàncies químiques, els quals es basen principalment en l'experimentació amb organismes aquàtics. Entre ells, el peix *fathead minnow*

(*Pimephales promelas*) és una de les espècies considerades per l'Agència de protecció mediambiental dels Estats Units (EPA) com a model biològic en estudis de toxicologia aquàtica [5]. La Figura 1 mostra una imatge d'aquest peix. Es tracta d'una espècie petita, que presenta tolerància a diversos tipus d'aigües (i per tant a diferents paràmetres com el pH, la terbolesa o la temperatura), té un cicle reproductiu/evolutiu ben definit i produeix molta descendència. Precisament aquestes característiques fan que el *fathead minnow* sigui una espècie adequada per dur a terme experiments controlats i reproduïbles de toxicitat aquàtica de compostos químics. De fet, s'ha convertit en un model biològic primordial que s'utilitza àmpliament tant a nivell regulatori com a nivell d'investigació [6]. Un dels experiments més comuns consisteix en determinar la concentració letal del compost químic que produeix la mort del 50% de la població de peixos *fathead minnow*. Aquesta propietat es coneix de forma abreujada com a LC_{50} (de l'anglès *lethal concentration for median mortality*, normalment expressada en $mol \cdot l^{-1}$) i es tracta d'una propietat de biopartició molt comuna en estudis de toxicitat aquàtica.

Figura 1. Peix *fathead minnow* (*Pimephales promelas*) [7].

La sorció de compostos orgànics en sòls correspon a un altre sistema de biopartició molt important dins l'àmbit mediambiental. Es tracta d'un procés clau a l'hora de determinar el transport i destí de substàncies contaminants, podent avaluar així el seu impacte sobre la natura [8,9]. El procés de sorció fa referència a la captació del solut per part del sòl, independentment del mecanisme (absorció, adsorció, intercanvi iònic) [10]. En aquest procés, la propietat de biopartició que normalment s'estudia és el coeficient de sorció del compost, que descriu quantitativament el seu grau de

distribució entre un sòl i una fase aquosa en contacte. Per a un compost químic orgànic, el coeficient de sorció varia considerablement en funció del tipus de sòl, però s'ha observat que és directament proporcional a la quantitat de matèria orgànica que conté el sòl, especialment en el cas de substàncies orgàniques neutres. Per aquest motiu, la propietat més acceptada i utilitzada és el coeficient de sorció normalitzat en base al contingut de carboni orgànic del sòl, K_{OC} (amb unitats $ml \cdot g^{-1}$), ja que d'aquesta manera es considera minimitzada la variabilitat entre sòls [11,12]. La determinació de valors de K_{OC} és especialment rellevant dins l'àmbit agroquímic perquè l'ús estès d'adobs i pesticides (Figura 2) fa necessari controlar el risc mediambiental que suposen.

Figura 2. Fumigació d'un camp de conreu amb pesticides [7].

Altres sistemes de biopartició es relacionen amb el cos humà, com és el cas de l'absorció de substàncies químiques a través de la pell [13]. Aquest és un procés essencial per a les indústries farmacèutiques i cosmètiques, i també des del punt de vista toxicològic quan s'avalua l'exposició ocupacional a compostos químics. L'absorció dèrmica d'una substància en solució aquosa es mesura normalment mitjançant el coeficient de permeabilitat de la substància a través de la pell, K_p (expressat en $cm \cdot s^{-1}$). Aquesta propietat de biopartició indica la capacitat de penetració i difusió de la

substància a través de la pell, que és bàsicament un procés de difusió passiva. La capa còrnia (*stratum corneum*) correspon a la capa més externa de l'epidermis i és la principal barrera que ha de creuar una substància per ser absorbida. El seu pas a través de la resta de capes de l'epidermis i la dermis dóna lloc finalment a l'entrada de la substància dins el sistema circulatori [10,14]. La Figura 3 il·lustra de manera simplificada l'estructura de la pell humana.

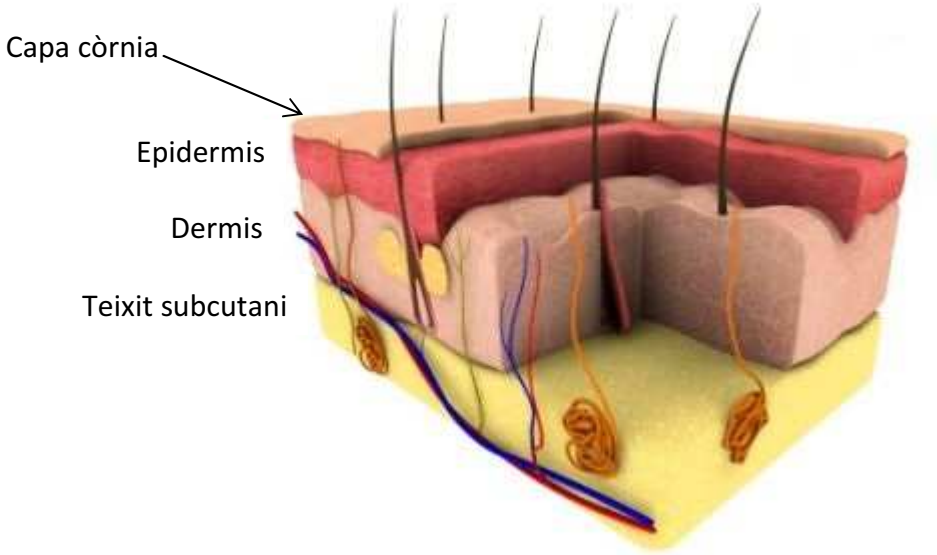


Figura 3. Secció transversal de la pell [15].

Tot i la importància indiscutible de conèixer propietats de biopartició com les esmentades, també s'han comentat prèviament les dificultats sovint associades a la seva determinació experimental (procediments experimentals complicats, llargs, costosos i èticament qüestionables). Per aquests motius interessa disposar de models que permetin predir propietats de biopartició de substàncies químiques més fàcilment, en menys temps, de manera més econòmica i oferint alternatives a l'experimentació amb éssers vius. De fet, els models de predicció s'han convertit en eines de cribratge indispensables degut al número creixent de compostos dels quals es necessita conèixer alguna de les seves propietats [10]. Els models de predicció contribueixen a la reducció del número d'assajos *in vivo* i *in vitro* necessaris, disminueixen la dificultat, el cost i el

temps requerits i, per tant, es tracta d'alternatives sostenibles a la determinació directa de propietats de biopartició.

Al llarg dels anys, molts models QSAR (relacions quantitatives estructura-activitat, de l'anglès *quantitative structure-activity relationships*) [16,17] han estat desenvolupats amb aquest propòsit. Relacionen l'activitat o propietat d'un compost amb una sèrie de descriptors moleculars del propi compost. Un cop establert un model QSAR, no només s'obté informació de l'activitat o propietat en estudi, sinó que el propi model pot servir per predir l'activitat o propietat de nous compostos si es coneixen els seus descriptors. El model de paràmetres de solvatació és un tipus de QSAR que s'ha fet servir amb molt d'èxit per descriure tant sistemes de partició biològics com fisicoquímics [18,19].

L'ús d'un mateix model per caracteritzar sistemes biològics i fisicoquímics ofereix la possibilitat de comparar-los. D'aquesta manera és possible identificar sistemes fisicoquímics que emulin processos biològics i en aquest cas s'espera obtenir una bona correlació entre les seves propietats. Aleshores, la propietat biològica d'un nou compost es podria predir de forma senzilla a partir de la seva mesura en el sistema fisicoquímic directament relacionat, tal com il·lustra la Figura 4. El principal avantatge respecte a l'estimació QSAR és que no cal conèixer valors de descriptors pels nous soluts, el que fa suscitar molt interès per l'emulació de sistemes biològics mitjançant sistemes fisicoquímics.

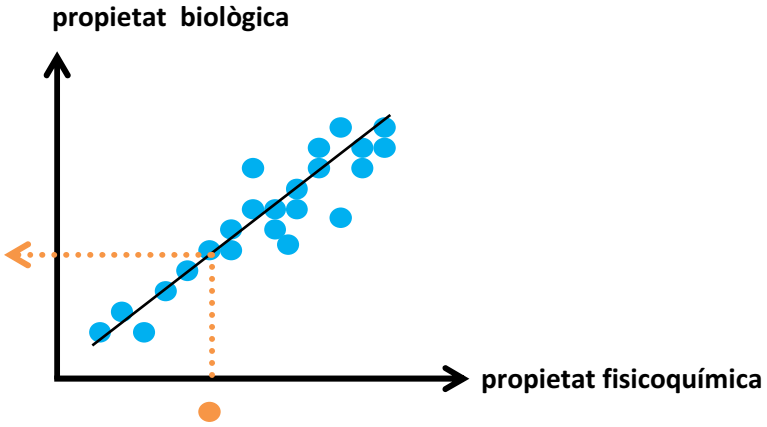


Figura 4. Predicció de la propietat biològica d'un compost (●) a partir d'un model de correlació biològico-fisicoquímic.

Entre els diferents sistemes fisicoquímics, el de partició octanol-aigua ha estat el clàssic sistema de referència per modelar sistemes de biopartició [20]. No obstant, els sistemes cromatogràfics han anat guanyant molt terreny en aquest àmbit en els darrers anys per la seva facilitat d'ús [21-23].

1.2. SISTEMES FISICOQUÍMICS PER EMULAR SISTEMES DE BIOPARTICIÓ

1.2.1. Sistema de partició octanol-aigua

El coeficient de partició d'una substància (P) es defineix com el quocient de les concentracions d'aquesta substància en la seva forma neutra entre dues fases líquides immiscibles que es troben en equilibri. Quan una de les fases és aigua i l'altra octanol, es parla del coeficient de partició octanol-aigua ($P_{o/w}$) [24]:

$$P_{o/w} = \frac{C_o}{C_w} \quad (1)$$

C_o correspon a la concentració de la substància a la fase octanòlica i C_w a la concentració de la substància a la fase aquosa. Sovint es fa servir la forma logarítmica del coeficient de partició octanol-aigua, és a dir, log $P_{o/w}$.

El log $P_{o/w}$ d'un compost indica el seu caràcter lipofílic [25] i, per tant, la seva afinitat per una matriu lipídica (representada per l'octanol atès al seu caràcter apolar i oliós). No cal confondre, però, el concepte de lipofilicitat amb el d'hidrofobicitat. En un sistema format per una fase lipídica i una aquosa, la hidrofobicitat considera la interacció entre el compost i l'aigua, mentre que la lipofilicitat fa referència a la interacció del compost amb el lípid. En realitat, la lipofilicitat engloba la hidrofobicitat i un terme de polaritat, i és la propietat que es mesura a través del log $P_{o/w}$.

La lipofilicitat de fàrmacs, productes agroquímics i substàncies químiques en general va guanyar molta importància arran dels treballs pioners de Meyer [26] i Overton [27]. També va ser fonamental la posterior contribució de Hansch i Fujita, ja que van ser els primers en proposar l'ús del log $P_{o/w}$ per modelar la partició de substàncies en sistemes biològics [28,29]. Des de llavors, nombrosos models fisicoquímics basats en el log $P_{o/w}$ s'han fet servir per descriure processos de

biopartició. Des del punt de vista farmacològic, el log $P_{o/w}$ ha donat lloc a models que permeten predir les propietats d'absorció, distribució, metabolisme, excreció i toxicitat (ADMET) de substàncies químiques [30,31], el que ha fet que es converteixi en un índex de lipofilicitat clau en el procés de desenvolupament i descobriment de nous fàrmacs. A la vegada, el log $P_{o/w}$ també ha guanyat molta importància dins l'àmbit mediambiental [10], ja que el caràcter lipofílic dels compostos sovint està relacionat amb la seva sorció en sòls i la seva concentració en peixos i animals en general.

Atès que la lipofilicitat s'ha convertit en una de les propietats fisicoquímiques més estudiades, els mètodes per a la mesura i predicció del log $P_{o/w}$ han estat dels més desenvolupats. El mètode d'equilibrament de fases, àmpliament conegut pel seu nom anglès *shake-flask* [24,32,33], és el mètode tradicional i de referència per a la determinació experimental de valors de log $P_{o/w}$. Aquest mètode permet obtenir resultats molt acurats si es controlen bé tots els detalls experimentals, no obstant és llarg, tediós i requereix quantitats significatives de compostos purs. Els mètodes potenciomètrics [34,35] són una bona alternativa en el cas de compostos àcids o bàsics, però també necessiten una quantitat relativament gran de la substància pura. D'altra banda, les tècniques cromatogràfiques esdevenen la millor opció quan cal determinar el log $P_{o/w}$ de molts compostos en poc temps, ja que les anàlisis són ràpides i es poden automatitzar. A més, es necessita poca quantitat de compost, aquest no cal que sigui pur i els valors de log $P_{o/w}$ que s'obtenen són força acurats (encara que no tant com els obtinguts mitjançant *shake-flask* o potenciometria). S'ha estès especialment l'ús de la cromatografia de líquids d'alta resolució en fase invertida (RP-HPLC) per a la determinació del log $P_{o/w}$ de substàncies. De fet, són nombrosos els mètodes de RP-HPLC desenvolupats per mesurar la lipofilicitat de compostos, per exemple fent servir el factor de retenció extrapolat a aigua pura (log k_w) [36], l'índex d'hidrofobicitat cromatogràfica (CHI) [37,38] o la combinació de la retenció cromatogràfica amb descriptors derivats de l'estructura dels soluts [39-41].

Donada la importància de la lipofilicitat, també s'han desenvolupat nombrosos mètodes de càlcul de log $P_{o/w}$ que han donat lloc a una gran varietat de programes. ClogP, ALOGPS i ACD/logP són alguns exemples dels softwares existents [42]. Tot i que el log $P_{o/w}$ d'una mateixa substància calculat mitjançant diversos programes pot ser

discordant perquè fan servir algoritmes diferents, els programes de càlcul de log $P_{o/w}$ es troben fàcilment a l'abast i es consideren una eina molt útil per a una primera avaluació de la lipofilicitat de manera simple, ràpida i barata.

Respecte a les bases de dades de valors de log $P_{o/w}$, la majoria de dades han estat compilades per Leo i col·laboradors [43]. De fet, la seva base de dades del Pomona College a través del programa Bio-Loom [44] és la més coneguda, encara que també n'hi ha d'altres disponibles [45,46].

1.2.2. Sistemes cromatogràfics

1.2.2.1. Sistemes de cromatografia de líquids d'alta resolució

La cromatografia de líquids d'alta resolució (HPLC, de l'anglès *high-performance liquid chromatography*) [47,48] és una tècnica analítica basada en la diferent distribució dels components d'una mescla entre una fase mòbil líquida i una fase estacionària que són immiscibles i de diferent polaritat. Les substàncies, en funció de la seva naturalesa, es distribueixen de forma diferent entre les fases de manera que adquireixen diferent mobilitat i elueixen a temps diferents, separant-se així en pics cromatogràfics que poden ser analitzats qualitativament i/o quantitativament. La diferent polaritat de les fases juga un paper fonamental en la separació. Quan la fase mòbil és més polar que la fase estacionària es parla de cromatografia de líquids en fase invertida (RP-HPLC).

Les fases mòbils més comuns en RP-HPLC són mescles hidroorgàniques d'aigua o tampó i un modificador orgànic com el metanol, l'acetonitril o el tetrahidrofurà. Existeixen dos modes d'elució: elució isocràtica, quan la composició de la fase mòbil es manté constant; elució en gradient, quan la fracció de fase aquosa i modificador orgànic varien de forma programada durant l'anàlisi cromatogràfica.

L'element separador en HPLC és una columna que conté la fase estacionària. Les fases estacionàries més populars en RP-HPLC contenen espècies orgàniques enllaçades a una superfície sòlida, normalment un suport sòlid de sílice. Entre les fases enllaçades més comuns es troben les cadenes alquíliques C18 (fases estacionàries d'octadecilsilà).

La Figura 5 mostra dues fases estacionàries formades per cadenes C18 enllaçades que corresponen a dues columnes convencionals de RP-HPLC (XTerra RP18 i XTerra MSC18).

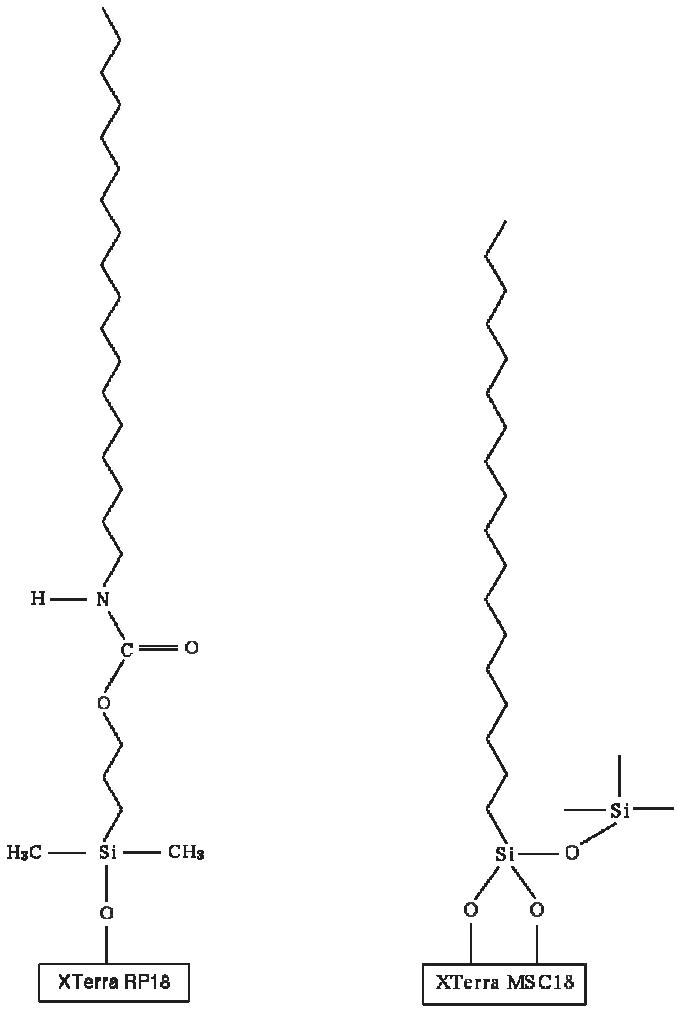


Figura 5. Estructura de dues columnes C18 [49].

Entre els diferents paràmetres cromatogràfics, el factor de retenció s'utilitza freqüentment per descriure la migració que experimenten les substàncies en un determinat sistema cromatogràfic. El factor de retenció d'una substància en HPLC (k_{HPLC}) es defineix com:

$$k_{HPLC} = \frac{t_R - t_0}{t_0} \qquad (2)$$

on t_R és el temps de retenció de la substància, és a dir, el temps que triga en arribar al detector des del moment que és injectada en el sistema; t_0 correspon al temps mort, és a dir, el temps que triga un solut que no es retingui en absolut en arribar al detector després de ser injectat.

La idea de fer servir sistemes de RP-HPLC per emular processos de biopartició es deu principalment a que la distribució que experimenta un solut en RP-HPLC es pot considerar molts cops anàloga al seu comportament en un sistema de biopartició. Es tracta, en ambdós casos, de processos dinàmics on les propietats hidrofòbiques, electròniques i estèriques tant del solut com del sistema són determinants. A més, es poden establir similituds entre ambdós tipus de processos. Per exemple, tant en un sòl com en un sistema de RP-HPLC, una fase aquosa percola a través d'una microestructura inert i porosa (la fracció mineral en un sòl i la sílice de la columna en RP-HPLC), la qual es troba enllaçada a una fase orgànica solvatada i immòbil (la matèria orgànica del sòl i la cadena alquílica de la columna en RP-HPLC) que determina la retenció selectiva dels compostos orgànics. Com a conseqüència, és raonable esperar que existeixi relació entre la retenció cromatogràfica en RP-HPLC i el coeficient de sorció de substàncies orgàniques en sòls, tal com han publicat diversos autors [50-52]. De la mateixa manera, també existeixen publicacions que relacionen la retenció obtinguda en RP-HPLC amb altres propietats de biopartició com la concentració de narcosi en capgrossos [21], el factor de bioconcentració en peixos [53] o el coeficient de partició i el coeficient de permeabilitat a la pell [54], entre d'altres.

Més enllà de l'analogia que pot haver entre sistemes de RP-HPLC i sistemes biològics, el creixent interès per estimar propietats biològiques a partir de mesures de RP-HPLC també es deu als avantatges que ofereix la tècnica cromatogràfica. Principalment, les anàlisis són simples, econòmiques, fiables, fàcils d'automatitzar, no requereixen molt temps, tampoc molta quantitat de mostra i aquesta no cal que sigui pura. Un avantatge addicional i molt important és la flexibilitat amb la que es poden canviar les propietats dels sistemes de RP-HPLC, la qual cosa s'aconsegueix fàcilment variant les característiques de la fase mòbil i/o estacionària; tan sols modificant els components o les composicions de la fase mòbil o bé fent servir una altra columna cromatogràfica es poden obtenir diferents sistemes de RP-HPLC. Aquesta flexibilitat és

clau a l'hora d'emular processos de biopartició ja que obre el ventall de possibles sistemes de RP-HPLC que s'assimilin a sistemes biològics.

El propòsit de disposar de sistemes de RP-HPLC més similars a processos de biopartició ha donat lloc, fins i tot, al desenvolupament de fases estacionàries. És el cas de les membranes artificials immobilitzades (IAM), introduïdes per Pidgeon i Venkataram [55] a finals dels anys 80 com a fases estacionàries que emulen a les biomembranes.

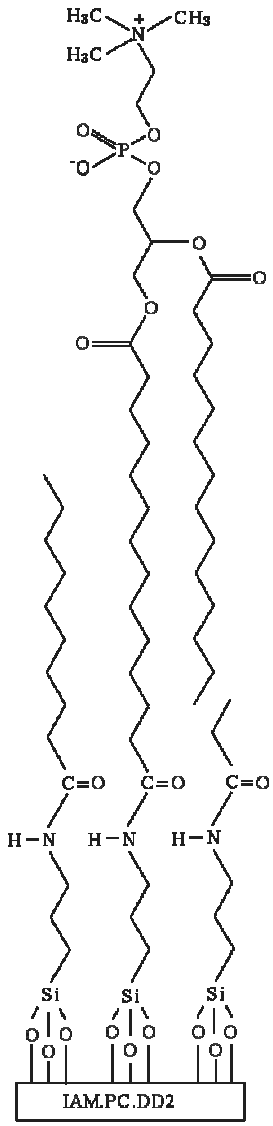


Figura 6. Fase estacionària d'una columna IAM [49].

Les columnes IAM es caracteritzen per emprar els principals constituents de la membrana cel·lular, els fosfolípids. La Figura 6 mostra la fase estacionària d'una columna IAM, preparada enllaçant covalentment molècules de fosfatidilcolina a la sílice i formant una monocapa de fosfolípids que mimetitza l'entorn lipídic de les membranes cel·lulars.

L'objectiu de les columnes IAM és obtenir prediccions ràpides i precises de la interacció de fàrmacs amb membranes biològiques mitjançant mesures cromatogràfiques. Per això en els últims anys s'ha correlacionat la retenció de substàncies en columnes IAM amb diferents paràmetres biològics com l'absorció intestinal en rates [56,57] o la penetració de compostos a través de la barrera hematoencefàlica [58,59]. S'ha demostrat que aquest tipus de columnes tenen un gran potencial per modelar la distribució de substàncies en sistemes biològics [60-62].

1.2.2.2. Sistemes de cromatografia electrocinètica micel·lar

La cromatografia electrocinètica (EKC, de l'anglès *electrokinetic chromatography*) és una altra tècnica de separació, híbrida de l'electroforesi capil·lar i la cromatografia de líquids [63,64]. Combina el moviment electroforètic i electroosmòtic dels soluts i els solvents amb el mecanisme de separació cromatogràfica, ja que en la separació intervé tant la pròpia migració electroforètica dels soluts com la partició d'aquests entre l'electròlit suport i una fase pseudoestacionària. La fase pseudoestacionària pot estar formada per diferents tipus d'additius (micel·les, polímers, liposomes, microemulsions) que donen lloc a les diferents variants de cromatografia electrocinètica. La modalitat més popular és la cromatografia electrocinètica micel·lar (MEKC, sigles angleses de *micellar electrokinetic chromatography*), desenvolupada als anys 80 per Terabe i col·laboradors [65] i caracteritzada per contenir dissolucions micel·lars com a fase pseudoestacionària.

Els sistemes de MEKC estan formats doncs per una solució d'electròlit que es fa servir com a medi conductor, normalment un tampó aquós, on també hi ha dissolt un tensioactiu iònic a una concentració superior a la seva concentració micel·lar crítica (CMC). Per sobre de la CMC els monòmers de tensioactiu s'agrupen formant

micel·les (Figura 7) i d'aquesta manera donen lloc a la fase micel·lar que actua com a pseudoestacionària [66].

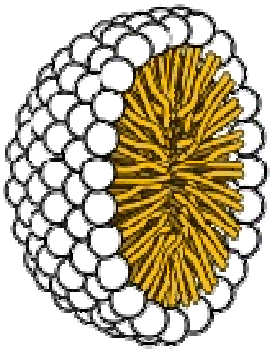


Figura 7. Estructura d'una micel·la esfèrica [7].

Els tensioactius, també anomenats surfactants o detergents, són molècules formades per caps hidrofílics i cues hidrofòbiques (representades a la Figura 7). En funció de la càrrega elèctrica de la part hidrofílica existeixen diferents tipus de tensioactius: aniònics (SDS, SC, LPFOS,...), catiònics (TTAB, HTAB, DTAC,...), zwitteriònics (CHAPS, CHAPSO,...) i no iònics (Brij 35, Triton X-100,...). Alguns d'ells es mostren a la Figura 8.

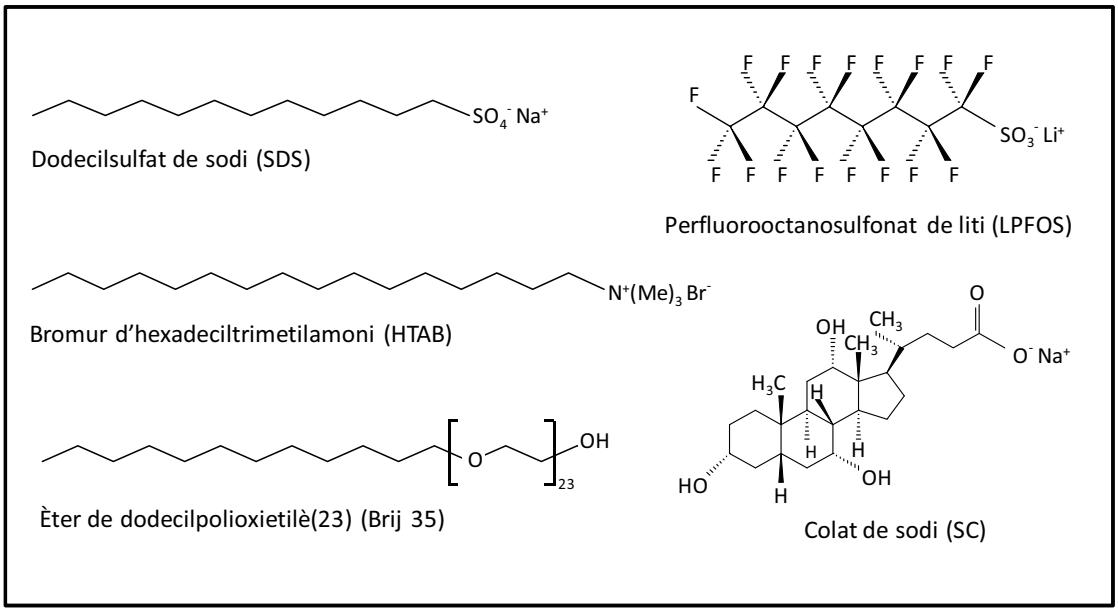


Figura 8. Estructura de diversos tensioactius fets servir en MEKC.

Els sistemes de MEKC poden estar formats per un tensioactiu o per mescles de tensioactius, però en qualsevol cas cal que continguin almenys un tensioactiu iònic que proporcioni càrrega a les micel·les perquè aquestes migrin. Al llarg del seu desplaçament, les micel·les interaccionen amb els anàlits i aquests experimenten un procés de partició entre la fase aquosa (electròlit) i la fase micel·lar (pseudoestacionària) que depèn tant de la naturalesa dels soluts com del sistema. Gràcies a que els soluts a més de migrar electroforèticament també experimenten partició, en MEKC no només és possible la separació de compostos carregats sinó també de compostos neutres [67]. La Figura 9 il·lustra el principi de separació de tres substàncies neutres de diferent hidrofobicitat (S1, S2 i S3). El flux electroosmòtic (EOF) correspon al moviment de la solució d'electròlit dins el capil·lar i es considera la força impulsora en electroforesi capil·lar.

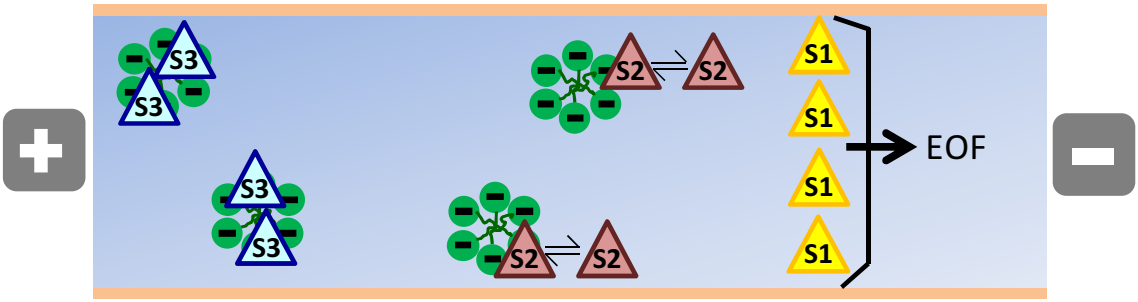


Figura 9. Separació de substàncies neutres en MEKC.

Per determinar el factor de retenció d'un compost neutre en MEKC (k_{MEKC}) cal fer servir un marcador de flux electroosmòtic i un marcador micel·lar [67,68], els quals determinaran la finestra d'elució del compost. El marcador de flux electroosmòtic ha de ser una substància neutra, petita i polar que no interaccioni gens amb les micel·les (S1 a la Figura 9). Per tant, es trobarà a la fase aquosa tot el temps que triga en travessar el capil·lar, eluint a un temps t_{eof}. El metanol, el dimetilsulfòxid, l'acetona i la formamida són exemples de marcadors de flux electroosmòtic. Contràriament, el marcador micel·lar ha de particionar totalment amb les micel·les (S3 a la Figura 9), pel que es necessita una substància neutra, gran i molt hidrofòbica, com pot ser

l'octilbenzè o la dodecanofenona. El marcador micel·lar coeluirà amb les micel·les a un temps t_{mc}. Qualsevol compost neutre parcialment soluble a la fase micel·lar (S2 a la Figura 9) eluirà a un temps de retenció t_R comprès entre la finestra d'elució que determinen t_{eof} i t_{mc}, de manera que el seu factor de retenció es pot calcular mitjançant la següent expressió:

$$k_{MEKC} = \frac{t_R - t_{eof}}{t_{eof}\left(1 - \frac{t_R}{t_{mc}}\right)} \quad (3)$$

En els darrers anys s'han obtingut bones correlacions entre factors de retenció en MEKC i diferents paràmetres biològics, demostrant així el potencial dels sistemes de MEKC per modelar processos de biopartició. Entre les propietats biològiques correlacionades es troben la biodisponibilitat de corticosteroides [69], la narcosi en capgrossos [70], diferents paràmetres d'ecotoxicitat de compostos orgànics [71,72] i la toxicitat de fàrmacs en humans [73]. L'aptitud dels sistemes de MEKC per emular sistemes biològics es deu principalment al seu caràcter amfipàtic i anisotròpic. Les parts hidrofíliques i hidrofòbiques dels tensioactius donen lloc a fases micel·lars que proporcionen unes interaccions hidrofòbiques i electròniques assimilables a les que regeixen alguns processos de biopartició. Per això els sistemes de MEKC esdevenen bons models biològics [74].

D'altra banda, les mesures en sistemes de MEKC són pràctiques i ofereixen una sèrie d'avantatges que les fan especialment atractives per predir propietats de biopartició. De fet, incorporen tant els avantatges de la cromatografia de líquids com els de l'electroforesi capil·lar. Es caracteritzen per ser acurades, per temps d'anàlisi curts, pel fet de que es poden automatitzar, per requerir petites quantitats de mostra i també de solucions, no cal que la mostra sigui pura i són anàlisis econòmiques i senzilles de realitzar. Addicionalment, un aspecte molt positiu és l'enorme facilitat per manipular la selectivitat dels sistemes de MEKC, la qual cosa es pot aconseguir simplement modificant el tipus de tensioactiu, fent servir mescles de tensioactius, afegint additius (ciclodextrines, urea, etc.) o afegint solvents orgànics al tampó de

separació [66,75-77]. Entre aquestes opcions, és ben sabut que l'elecció del tensioactiu és el factor que més determina la selectivitat d'un sistema de MEKC [78,79]. Per aquest motiu s'han estudiat i es continuen estudiant sistemes de MEKC amb diferents propietats segons el tensioactiu o les mescles de tensioactius que constitueixen la fase pseudoestacionària [80]. Aquesta gran flexibilitat és especialment útil des del punt de vista de la modelització biològica ja que permet disposar de sistemes de MEKC que resultin adequats per predir propietats de biopartició. De fet, la possibilitat d'ajustar les característiques d'un sistema de MEKC per tal de fer-les més semblants a les d'un sistema biològic ha convertit la cromatografia electrocinètica micel·lar en una tècnica potent i versàtil per emular sistemes de biopartició.

1.2.2.3. Altres sistemes cromatogràfics

Val a dir que existeixen altres sistemes cromatogràfics a part dels descrits que també poden ser útils per predir propietats biològiques de compostos químics. Destaquen els sistemes de cromatografia líquida micel·lar (MLC) i de microemulsions (MELC), així com els sistemes de cromatografia electrocinètica de liposomes (LEKC) i també de microemulsions (MEEKC).

La cromatografia líquida micel·lar [81] i la cromatografia líquida de microemulsions [82] són, de fet, modalitats de la cromatografia de líquids en fase invertida convencional (RP-HPLC) detallada a l'apartat 1.2.2.1. La principal diferència és que en MLC es fan servir fases mòbils micel·lars preparades a partir de dissolucions de tensioactius per sobre de la seva CMC, mentre que en MELC es fan servir fases mòbils formades per microemulsions preparades en solució aquosa a partir d'un oli, un surfactant i un cosolvent. Encara que s'han obtingut bones correlacions entre paràmetres biològics i factors de retenció en sistemes de MLC [83,84] i MELC [85], cal tenir en compte que emprar dissolucions micel·lars o de microemulsions en cromatografia de líquids implica treballar amb molta cura i fer servir protocols de neteja adequats per tal de no fer malbé ni escurçar la vida útil tant del cromatògraf com de la columna cromatogràfica.

També existeixen treballs [86-88] on s'han relacionat satisfactòriament propietats biològiques i factors de retenció en sistemes de cromatografia electrocinètica de

liposomes (LEKC) i de microemulsions (MEEKC). Ambdós tipus de sistemes són variants de la cromatografia electrocinètica, de la mateixa manera que ho són els sistemes de MEKC tractats a l'apartat 1.2.2.2. En LEKC [89] i MEEKC [90], però, els additius que formen la fase pseudoestacionària són liposomes i microemulsions respectivament. Els sistemes de LEKC i MEEKC requereixen més treball experimental que els de MEKC ja que la preparació de liposomes i microemulsions és un procés més llarg, tediós i delicat que la preparació de dissolucions micel·lars. A més, els sistemes de LEKC i MEEKC presenten normalment més problemes d'estabilitat que els de MEKC i també cal tenir més precaució pel que fa a la neteja del capil·lar i l'equip d'electroforesi.

En l'emulació de processos de biopartició mitjançant sistemes fisicoquímics, no només cal tenir en compte l'aptitud del sistema fisicoquímic, sinó que també interessa que sigui un sistema pràctic a l'hora de treballar. En aquest sentit, els sistemes de MLC, MELC, LEKC i MEEKC presenten més desavantatges que els de RP-HPLC convencional i els de MEKC. Per aquest motiu, aquesta tesi s'ha centrat principalment en avaluar l'habilitat de sistemes cromatogràfics de RP-HPLC convencional i de MEKC per predir propietats biològiques de substàncies.

1.3. MODEL DE PARÀMETRES DE SOLVATACIÓ

1.3.1. Perspectiva general dins els estudis QSAR

Tal com el seu nom indica, les relacions quantitatives estructura-activitat (QSAR) [16,17] relacionen de manera quantitativa l'activitat biològica de compostos químics amb els seus descriptors moleculars. Els models QSAR es basen en la teoria de les relacions lineals de l'energia lliure, LFER (sigles angleses de *linear free energy relationships*). Aquestes assumeixen que la variació d'energia lliure total d'un procés és una combinació lineal de les variacions d'energia lliure associades a cadascuna de les diferents interaccions moleculars implicades, les quals es poden descriure mitjançant descriptors moleculars.

Un model QSAR es pot expressar de manera simplificada mitjançant la següent equació:

$$\text{Activitat biològica} = f(\text{descriptors moleculars}) \quad (4)$$

D'aquesta equació es dedueix l'objectiu general dels estudis QSAR: establir els paràmetres moleculars independents que proporcionen una correlació estadísticament significativa respecte a una determinada activitat biològica per a un conjunt de soluts. La Figura 10 il·lustra el fonament d'un model QSAR i les seves dues aplicacions bàsiques [91], les quals es resumeixen a continuació:

- Aportar informació sobre el procés biològic mitjançant la identificació i interpretació dels factors clau que determinen la resposta biològica. Per tant, els models QSAR poden ser molt útils per conèixer els mecanismes d'acció que segueixen els compostos per exhibir una determinada activitat biològica.
- Predir l'activitat biològica de nous compostos. Un cop establert un model QSAR, aquest permet estimar l'activitat de substàncies no incloses en el model del sistema biològic considerat.

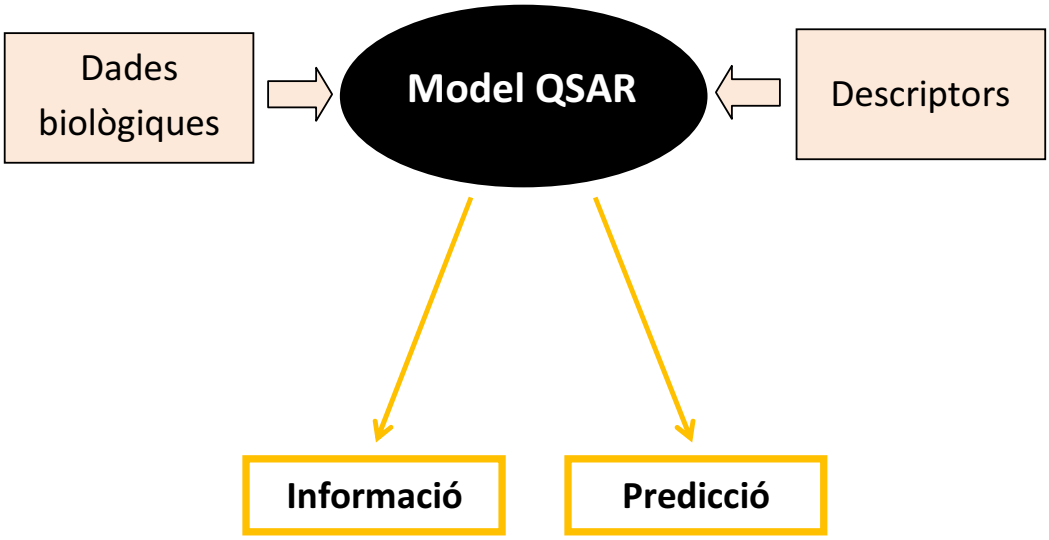


Figura 10. Fonament dels estudis QSAR.

A la vegada, aquestes dues aplicacions dels models QSAR contribueixen de manera molt significativa a la reducció d'assajos *in vivo* i *in vitro*. Donada la gran necessitat de conèixer l'activitat biològica d'un gran número de compostos en molts àmbits (indústries farmacèutiques, alimentàries, agroquímiques, estudis mediambientals, etc.), els models QSAR esdevenen eines fonamentals i necessàries de cribratge. D'aquesta manera permeten reduir costos, esforços i temps, així com minimitzar el número d'assajos experimentals amb éssers vius. Per tant, es tracta de models de gran utilitat [92-96].

Pel que fa al desenvolupament de models QSAR [97], la tècnica més utilitzada és la regressió lineal múltiple, MLR (de l'anglès *multiple linear regression*), entre l'activitat biològica d'una sèrie de compostos i els seus corresponents descriptors. Existeixen molts tipus de descriptors moleculars que es poden fer servir (estructurals, fisicoquímics, electroestàtics, etc.) i aquests poden ser experimentals o calculats (E-DRAGON [98] i CODESSA Pro [99] són alguns dels programes de càlcul comercials disponibles). La bondat del model dependrà tant de la qualitat com de la quantitat tant de les dades biològiques com dels descriptors. Convé que les dades biològiques siguin homogènies i representatives, és a dir, que presentin una uniformitat (no s'agrupin en clústers per exemple) i cobreixin l'espai químic d'interès (els compostos fets servir en

el model han cobrir un interval de variació pel que fa al valor de l'activitat biològica en estudi). A la vegada, també convé que els descriptors compleixin certs requisits: han de ser rellevants de manera que expliquin bé la variabilitat de l'activitat biològica pel conjunt de soluts analitzats, no han d'estar intercorrelacionats per tal d'evitar informació redundant i, a més, convé fer servir descriptors interpretables i el mínim nombre possible per facilitar la comprensió del model.

Avaluar la bondat del model un cop construït és clau per determinar la fiabilitat de la informació i les prediccions que se'n derivin [97]. Amb tal propòsit es poden fer servir paràmetres estadístics (coeficients de correlació, desviacions estàndard, etc.) i també validacions amb diferents conjunts de soluts que es coneixen pel seu nom anglès com a *training set* i *test set*. El *training set* és el conjunt de soluts amb el que s'estableix el model i s'avalua la seva robustesa (validació interna). El *test set* és independent del *training set* i correspon al conjunt de soluts que es fa servir per avaluar la capacitat predictiva del model (validació externa). Cal també destacar que un aspecte molt important en el desenvolupament i avaluació de models QSAR és la correcta detecció i interpretació de valors anòmals (*outliers* en anglès), normalment a partir de gràfics i anàlisi estadística.

Entre els diferents models QSAR, destaquen els que descriuen processos de solvatació. El terme solvatació fa referència a la situació en què una molècula de solut està envoltada d'un excés de molècules de solvent com a resultat de les forces intermoleculars solut-solvent que s'estableixen. Aquests processos tenen lloc gràcies a un canvi en l'energia lliure de solvatació i es basen en un tipus de LFER que s'anomenen relacions lineals de l'energia lliure de solvatació, LSER (de l'anglès *linear solvation energy relationships*) [100]. Aquesta energia es considera suma de cadascuna de les energies implicades en les diferents etapes del procés de solvatació [101,102], les quals es detallen a continuació i s'il·lustren a la Figura 11:

- Etapa 1. Formació d'una cavitat en el solvent per allotjar el solut. Implica el trencament del clúster de molècules de solvent, les quals són desplaçades al formar-se la cavitat. La cavitat ha de ser adequada a la mida i la forma del solut.
- Etapa 2. Inserció del solut en la cavitat.

- Etapa 3. Establiment d'interaccions solut-solvent. Les interaccions intermoleculars que s'estableixen entre el solut i el solvent són de diferent tipus:

  - interaccions de dispersió (també anomenades de London)
  - interaccions degudes a la presència de dipols
  - interaccions degudes a la formació de ponts d'hidrogen

  En aquesta etapa, les molècules de solvent sobrants es reorganitzen i s'afegeixen a l'estructura general del solvent.

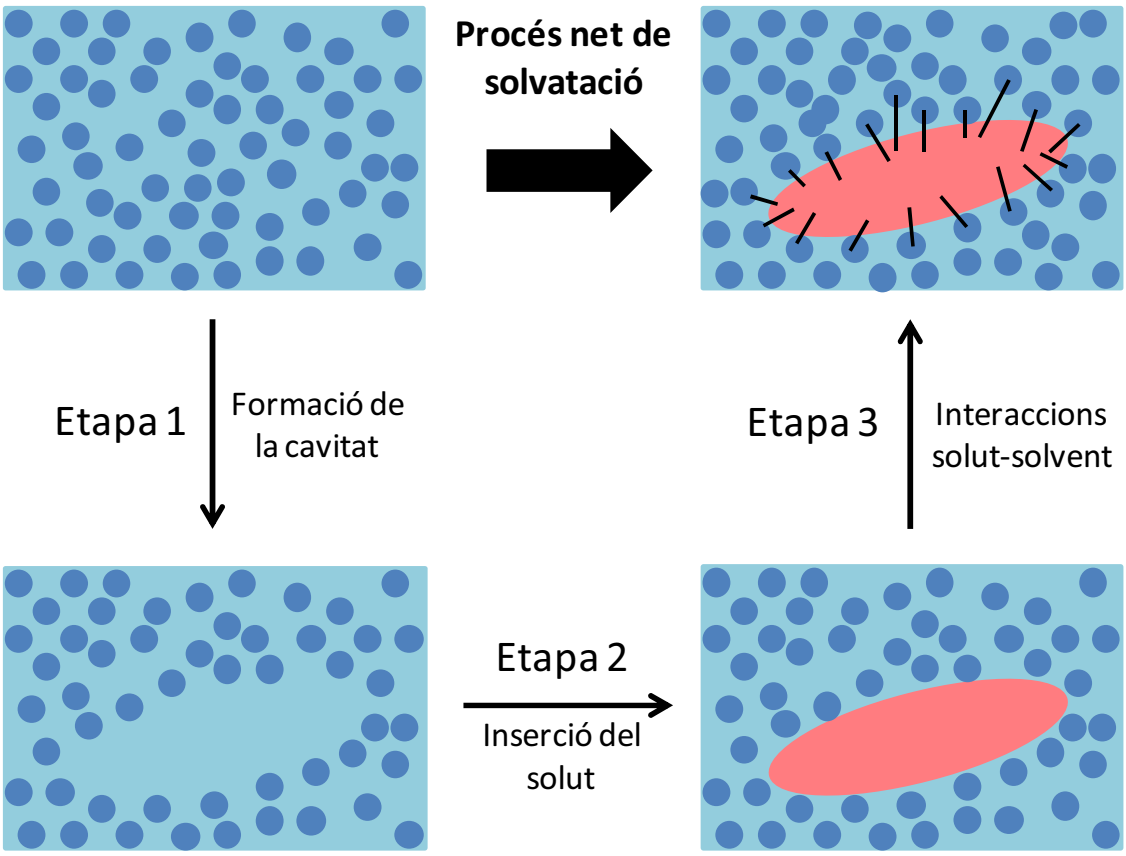


Figura 11. Etapes del procés de solvatació.

Els models basats en LSER han de reflectir els principals factors que governen aquestes etapes de solvatació i, per tant, han d'englobar els següents tres termes: terme de cavitat, terme de dipolaritat/polaritzabilitat i terme d'enllaç per pont d'hidrogen.

A principis de la dècada dels 90, Abraham i col·laboradors [18,103-106] van definir cadascun d'aquests termes en base a descriptors moleculars del solut, el que va donar lloc al model de paràmetres de solvatació, àmpliament acceptat i utilitzat des d'aleshores.

1.3.2. Equació general de solvatació

El model de paràmetres de solvatació es basa en la següent equació general de solvatació [18], aplicable a tots aquells processos que impliquin la transferència passiva de soluts neutres entre dues fases condensades:

$$\log SP = c + eE + sS + aA + bB + vV \quad (5)$$

La variable dependent (SP a l'equació 5) correspon a la propietat d'interès del solut en el sistema de partició considerat, que pot ser una constant d'equilibri o qualsevol altra propietat directament relacionada amb l'energia lliure de solvatació. Es pot tractar tant d'una propietat de partició biològica (coeficient de distribució sang-cervell (BB), coeficient de permeabilitat en pell (K_p), etc.) com fisicoquímica (coeficient de partició octanol-aigua ($P_{o/w}$), factor de retenció cromatogràfic (k), etc.) [19].

Tal com es pot observar, l'equació general de solvatació és multiparamètrica. *E*, *S*, *A*, *B* i *V* són les variables independents i corresponen als descriptors moleculars del solut proposats per Abraham i col·laboradors [18,103-106]. Aquests descriptors es troben definits a continuació, tot englobant-los dins els termes de cavitat, dipolaritat/polaritzabilitat i enllaç per pont d'hidrogen, esmentats a la secció anterior:

- terme de cavitat
  - *V*: *volum de McGowan*. Correspon al volum molecular del solut calculat mitjançant l'algoritme de McGowan [103]. Representa l'energia necessària per trencar les forces de cohesió del solvent i crear una cavitat adequada a la mida del solut per poder allotjar-lo. Està estretament relacionat amb les forces de dispersió solut-solvent.

- terme de dipolaritat/polaritzabilitat
  - *E*: *refractivitat molar en excés*. Mesura de la capacitat del solut per interaccionar amb el solvent mitjançant electrons –n i –π i es relaciona amb la polaritzabilitat.
  - *S*: *dipolaritat/polaritzabilitat*. Reflecteix la capacitat del solut per establir interaccions amb el solvent de tipus dipol-dipol i dipol-dipol induït, a més de mesurar també la polaritzabilitat.
- terme d'enllaç per pont d'hidrogen
  - *A*: *acidesa per pont d'hidrogen del solut*. Mesura la capacitat donadora de ponts d'hidrogen del solut. Es relaciona doncs amb els enllaços per pont d'hidrogen formats a l'interaccionar grups donadors del solut amb parells d'electrons lliures de grups acceptors del solvent.
  - *B*: *basicitat per pont d'hidrogen del solut*. Mesura la capacitat acceptora de ponts d'hidrogen del solut. Per tant, es relaciona amb els enllaços per pont d'hidrogen formats a l'interaccionar parells d'electrons lliures de grups acceptors del solut amb grups donadors del solvent.

D'aquests cinc descriptors, dos (*V* i *E*) es poden obtenir per càlcul i tres (*S*, *A*, *B*) cal determinar-los experimentalment [102,107]. El volum de McGowan del solut (*V*) i la seva refractivitat molar en excés (*E*) es poden calcular directament a partir de l'estructura molecular del propi compost. En canvi, la dipolaritat/polaritzabilitat (*S*) i l'acidesa i la basicitat del solut per pont d'hidrogen (*A* i *B*, respectivament) s'han de determinar experimentalment mitjançant sistemes cromatogràfics o de partició líquid-líquid i emprant l'equació general de solvatació (equació 5). A l'actualitat, els descriptors d'Abraham que s'han determinat i es troben disponibles corresponen a un ampli ventall de substàncies, uns 5000 compostos de naturalesa molt variada, des de gasos senzills com l'hidrogen fins a fàrmacs i pesticides d'estructura més complexa [108,109]. Cal dir que també és possible obtenir estimacions dels descriptors [109].

Els coeficients de l'equació general de solvatació (*e*, *s*, *a*, *b*, *v*) s'obtenen per regressió lineal múltiple entre els valors de la propietat SP, determinats per a un conjunt de soluts neutres en el sistema de dues fases considerat, i els descriptors d'aquests soluts. Conèixer els coeficients d'un sistema significa tenir el sistema caracteritzat segons el model de paràmetres de solvatació. Els coeficients aporten

informació sobre el medi, concretament reflecteixen les diferències entre les dues fases en les que es distribueix el solut pel que fa a la propietat del descriptor que acompanyen:

- *e*: representa la diferent capacitat de les fases per interaccionar amb els electrons –n i –π del solut.
- *s*: indica la diferent habilitat de les fases per establir interaccions tipus dipol-dipol i dipol-dipol induït amb el solut.
- *a*: mesura la diferència de les fases en basicitat per pont d'hidrogen.
- *b*: mesura la diferència de les fases en acidesa per pont d'hidrogen.
- *v*: informa de la diferent facilitat de les fases per allotjar el solut i, per tant, de la hidrofobicitat del sistema.

Pel que fa a la constant *c* de l'equació, està relacionada amb la proporció de fases. Aquesta constant modifica el valor de la propietat d'interès SP però no afecta a la selectivitat del sistema.

Per caracteritzar correctament un sistema mitjançant l'equació general de solvatació i, per tant, obtenir uns coeficients fiables i representatius del procés de partició descrit, cal complir una sèrie de requisits pel que fa al conjunt de soluts emprats a la caracterització [79,102]:

- Els soluts han de ser de naturalesa química variada (incloent compostos aromàtics i alifàtics) i cobrir un ampli interval de valors pel que fa a cadascun dels descriptors i també a la propietat SP.
- Els descriptors dels soluts no han de presentar colinealitat. Altes covariàncies entre descriptors fan difícil determinar la contribució de cada tipus d'interacció a la propietat d'interès SP i en conseqüència es poden obtenir incerteses altes en els coeficients.
- Matemàticament calen com a mínim 7 soluts per poder fer la regressió multilineal i obtenir els coeficients i estadístics de la caracterització. A la pràctica, però, s'ha observat que calen com a mínim uns 20 soluts per obtenir resultats acceptables.

- Els soluts han de ser neutres al pH de treball, doncs l'equació general de solvatació no contempla interaccions entre ions.
- Els soluts han de ser adequats al sistema de detecció fet servir per determinar el seu valor SP, és a dir, el valor de la propietat en estudi. Per exemple, si es vol caracteritzar un sistema cromatogràfic i les dades es determinen mitjançant detecció ultraviolada (UV) cal seleccionar un conjunt de soluts que absorbeixin a l'ultraviolat.
- Cal evitar soluts amb molt d'error en la mesura de la propietat d'interès SP. Per exemple, en la caracterització de sistemes de MEKC és millor no incloure soluts amb retenció molt propera al marcador de flux electroosmòtic o al marcador micel·lar ja que les seves mesures solen anar associades a un error considerable. De la mateixa manera, en HPLC cal evitar soluts amb temps de retenció molt llargs.

La correcta caracterització d'un sistema mitjançant l'equació 5 permet entendre les interaccions químiques que governen el procés estudiat. Per obtenir informació del comportament de les dues fases del sistema respecte a les interaccions que estableixen amb el solut, cal interpretar els coeficients de la caracterització en base al seu significat químic, magnitud i signe [102]:

- Químicament, els coeficients són específics i complementaris del descriptor molecular que acompanyen. Així doncs, el coeficient *b* acompanya al descriptor molecular *B* però *B* es refereix a la basicitat per pont d'hidrogen del solut i *b* a l'acidesa per pont d'hidrogen del medi.
- La magnitud dels coeficients reflecteix la diferència entre la capacitat d'interacció de cadascuna de les dues fases. Per exemple, un coeficient amb elevat valor absolut indica que el grau d'interacció de cadascuna de les fases amb el solut és molt diferent pel tipus d'interacció en qüestió (el solut interacciona molt amb una de les fases i gens amb l'altra). Tanmateix, un coeficient elevat indica que aquest tipus d'interacció té molta importància en el valor de la propietat en estudi. Per tant, comparar la magnitud dels coeficients per a un mateix sistema també permet conèixer quin tipus d'interacció té més influència en el valor de la propietat d'interès.

- El signe dels coeficients indica quina de les dues fases té una capacitat d'interacció major amb el solut, és a dir, cap a quina de les fases està afavorida la partició en el tipus d'interacció considerat. Mentre el signe positiu indica que la partició està afavorida cap a la fase 1, el signe negatiu indica que ho està cap a la fase 2, entenent la partició del solut entre dues fases com a: $\frac{[\text{solut}]_{\text{fase 1}}}{[\text{solut}]_{\text{fase 2}}}$. Normalment, la fase 1 correspon a una fase de caràcter orgànic mentre que la fase 2 es refereix a una fase aquosa. Així, un coeficient amb signe positiu indica que la fase orgànica interacciona més amb el solut que l'aquosa segons la propietat del solut considerada, de manera que té lloc la transferència del solut cap a la fase orgànica.

A continuació es detallaran i s'interpretaran algunes caracteritzacions de processos tant biològics com fisicoquímics de partició que han estat descrits amb èxit mitjançant l'equació general de solvatació.

1.3.2.1. Aplicació a sistemes de biopartició

La Taula 1 mostra alguns exemples de sistemes biològics de diferent naturalesa que es troben caracteritzats a la literatura amb el model de paràmetres de solvatació.

Taula 1. Caracteritzacions de sistemes biològics de partició.

Sistema biològic	**Propietat SP**	**Coeficients**					**Ref.**
		e	***s***	***a***	***b***	***v***	
Narcosi en capgrossos	log ($1/C_{nar}$)	0.770	-0.696	0.243	-2.592	3.343	[3]
Sorció sòl-aigua	log K_{OC}	0.74	-*	-0.31	-2.27	2.09	[110]
Distribució sang-cervell	log BB	0.511	-0.886	-0.724	-0.666	0.861	[111]
Absorció intestinal	log HIA	-0.025	0.141	-0.409	-0.514	0.204	[1]
Permeabilitat a la pell	log K_p	-0.106	-0.473	-0.473	-3.000	2.296	[112]
Toxicitat en el peix *fathead minnow*	-log LC_{50}	0.418	-0.182	0.417	-3.574	3.377	[113]

* Coeficient no significatiu al 95% de nivell de confiança

Els coeficients, obtinguts a l'aplicar l'equació 5, aporten informació sobre els processos biològics descrits. A totes les caracteritzacions de la Taula 1 s'observa que el coeficient *b* és negatiu i el coeficient *v* positiu. El fet de que *b* sigui negatiu significa que com més bàsic sigui el solut per pont d'hidrogen (més elevat sigui el descriptor *B*), més desafavorida estarà la seva transferència des de la fase aquosa fins la fase de caràcter biològic, el que succeeix en tots els sistemes contribuint així a reduir el valor de la propietat SP en tots els casos. D'altra banda, el coeficient *v* és positiu a totes les caracteritzacions i, per tant, el volum del solut (*V*) contribueix sempre positivament a la partició dels soluts cap a la fase biològica, de manera que com més gran sigui *V* més augmentarà la propietat SP de qualsevol dels sistemes considerats. No s'observa, però, cap tendència comuna pel que fa a la magnitud dels coeficients *b* i *v*, ni tampoc pel que fa al signe i la magnitud de la resta de coeficients.

A les caracteritzacions de la narcosi en capgrossos i de la toxicitat en el peix *fathead minnow*, s'observa que *b* i *v* són els coeficients amb major valor absolut, el que significa que la basicitat per pont d'hidrogen (*B*) i el volum (*V*) són les propietats del solut que més determinen el valor de log $(1/C_{nar})$ i -log LC_{50}. La principal diferència d'aquests dos sistemes respecte a la resta és que el coeficient *a* és positiu. En conseqüència, l'acidesa del solut per pont d'hidrogen (*A*) afavoreix tant la narcosi en capgrossos com la toxicitat en el peix *fathead minnow*, encara que no de forma determinant perquè el coeficient no té un valor absolut gaire gran.

La sorció sòl-aigua de compostos orgànics i la permeabilitat en pell de substàncies químiques també són processos de biopartició condicionats per la basicitat del solut per pont d'hidrogen (*B*) i el seu volum (*V*), ja que *b* i *v* també són els coeficients amb major valor absolut a les seves caracteritzacions. En conseqüència, l'acidesa per pont d'hidrogen (*b*) i la hidrofobicitat (*v*) són les propietats que més diferencien les dues fases que formen cadascun d'aquests dos sistemes. El signe negatiu del coeficient *a* indica que l'acidesa del solut per pont d'hidrogen (*A*) contribueix a la seva permanència a la fase aquosa i, per tant, s'oposa a la transferència del solut cap al sòl en un cas i cap a la pell en l'altre cas. El procés de sorció sòl-aigua també es caracteritza per tenir un coeficient *s* no significatiu, el que reflecteix que el sòl i l'aigua tenen una dipolaritat/polaritzabilitat equivalent i, per tant, el descriptor *S* no té

influència en el valor de log K_{OC}. En canvi, a la caracterització de permeabilitat en pell el coeficient *s* té la mateixa magnitud i signe que el coeficient *a*, el que significa que la pell i la fase aquosa en què es troba dissolt el solut presenten les mateixes diferències quant a dipolaritat/polaritzabilitat i basicitat per pont d'hidrogen.

La distribució sang-cervell i l'absorció intestinal són els sistemes de la Taula 1 amb caracteritzacions més diferents respecte a la resta. En el procés de distribució sang-cervell s'observa que la dipolaritat/polaritzabilitat (*S*) i el volum (*V*) són les propietats del solut que més contribueixen al valor de log BB, ja que *s* i *v* són els coeficients amb valor absolut més gran. No obstant, no hi ha molta diferència entre la magnitud dels cinc coeficients d'aquest sistema (en comparació amb les notables diferències que es poden observar entre la magnitud dels coeficients de caracteritzacions ja comentades). A diferència de la resta de sistemes, el coeficient *a* té un valor absolut gran a la caracterització de la distribució sang-cervell i, per tant, l'acidesa del solut per pont d'hidrogen (*A*) juga un paper fonamental en aquest procés. Aquesta propietat del solut s'oposa a la seva partició cap al cervell, de la mateixa manera que ho fan la seva basicitat (*B*) i dipolaritat/polaritzabilitat (*S*) (tant *a*, com *b* com *s* són negatius i contribueixen a disminuir el valor de log BB). Contràriament, *E* i *V* sí que afavoreixen la partició del solut cap al cervell (*e* i *v* són positius) i conseqüentment són responsables d'augmentar el paràmetre log BB. Pel que fa a l'absorció intestinal, els trets més característics són que l'acidesa i la basicitat del solut per pont d'hidrogen, *A* i *B* respectivament, són les propietats del solut més determinants en aquest procés. Tant un augment de *A* com de *B* contribueix a disminuir l'absorció intestinal (log HIA). Es tracta de l'únic sistema de la Taula 1 en què *v* no és dels coeficients amb major valor absolut i, per tant, el volum del solut (*V*) no és la propietat que més influeix en aquest procés de biopartició.

1.3.2.2. Aplicació a sistemes fisicoquímics

Són molts els sistemes fisicoquímics de partició que han estat caracteritzats a partir de l'equació general de solvatació [80,102,114-116]. A la Taula 2 es troben alguns d'ells, incloent sistemes de partició aigua-solvent i sistemes cromatogràfics d'HPLC i de MEKC (formats per tensioactius purs i mescles de tensioactius).

Taula 2. Caracteritzacions de sistemes de partició aigua-solvent i cromatogràfics.

Sistema fisicoquímic	Propietat SP	Coeficients					Ref.
		e	***s***	***a***	***b***	***v***	
Octanol-aigua	log $P_{o/w}$	0.562	-1.054	0.034	-3.460	3.814	[117]
Hexà-aigua	log P	0.579	-1.723	-3.599	-4.764	4.344	[107]
Columna RP18 - MeCN/tampó aquós	log k_{HPLC}	0.233	-0.460	-0.260	-1.574	1.582	[54]
Columna MSC18 - MeCN/tampó aquós	log k_{HPLC}	0.151	-0.463	-0.424	-1.719	1.759	[54]
Columna IAM - MeCN/tampó aquós	log k_{HPLC}	0.292	-0.344	0.141	-1.193	1.161	[54]
SDS-tampó aquós	log k_{MEKC}	0.558	-0.596	-0.266	-1.674	2.717	[118,119]
SC-tampó aquós	log k_{MEKC}	0.691	-0.693	0.117	-1.937	2.274	[118,119]
TTAB-tampó aquós	log k_{MEKC}	0.902	-0.617	0.766	-2.410	2.634	[118,119]
SDS/Brij 35-tampó aquós	log k_{MEKC}	0.667	-0.515	0.335	-2.723	2.767	[70]
LPFOS/LDS-tampó aquós	log k_{MEKC}	0.16	-0.31	-0.58	-0.85	2.64	[120]

Es poden observar tendències comuns entre els coeficients de les caracteritzacions presentades a la Taula 2. Per a tots els sistemes fisicoquímics considerats, el coeficient *b* és negatiu i el coeficient *v* positiu (el que coincideix amb el comportament observat també en tots els sistemes de biopartició de la Taula 1). A part, les caracteritzacions fisicoquímiques de la Taula 2 també tenen en comú que el coeficient *e* és positiu i el coeficient *s* negatiu, independentment que es tracti d'un sistema de partició aigua-solvent o un sistema d'HPLC o de MEKC.

Tenint en compte el signe positiu del coeficient *v*, la fase de caràcter orgànic dels sistemes considerats (octanol, hexà, columnes d'HPLC, tensioactius en MEKC) sempre té més facilitat que la fase aquosa per crear una cavitat de la grandària de la molècula del solut i les forces de dispersió són més favorables per allotjar-lo. Es tracta d'un fet raonable, doncs l'aigua és un solvent altament organitzat i cohesiu, de manera que és

més fàcil trencar les interaccions entre molècules de solvent orgànic que entre molècules d'aigua. En conseqüència, com més gran sigui el volum del solut (*V*) més tendència tindrà a anar cap a la fase orgànica. El signe també positiu del coeficient *e* indica que com més elevat sigui el descriptor *E*, més afinitat tindrà el solut per la fase orgànica, augmentant així el valor de la propietat SP en tots els sistemes considerats. Contràriament, el fet de que *s* i *b* siguin sempre negatius a la Taula 2 significa que la fase aquosa d'aquests sistemes presenta una major dipolaritat/polaritzabilitat i acidesa per pont d'hidrogen que la fase de caràcter orgànic. En conseqüència, un augment de *S* i *B* afavoreix la permanència dels soluts a l'aigua i contribueix a disminuir el valor de SP, és a dir, el coeficient de partició o la retenció cromatogràfica segons el sistema fisicoquímic que es tracti.

Pel que fa a la magnitud dels coeficients, un altre tret en comú dels sistemes de la Taula 2 és que els coeficients *b* i *v* són els de major valor absolut a totes les caracteritzacions (no s'observa el mateix en els sistemes biològics de la Taula 1). Aquest fet implica que les dues fases que formen cada sistema es diferencien majoritàriament en l'acidesa per pont d'hidrogen (*b*) i la hidrofobicitat (*v*). Com a conseqüència, la basicitat del solut per pont d'hidrogen (*B*) i el seu volum (*V*) són els descriptors que més influeixen en el valor de les propietats fisicoquímiques d'aquests sistemes.

El coeficient *a* és l'únic que no presenta cap comportament general, ni pel que fa a signe ni pel que fa a magnitud, entre els sistemes de la Taula 2. Per tant, l'acidesa del solut per pont d'hidrogen (*A*) afectarà de manera diferent a cadascun dels processos fisicoquímics considerats. Per exemple, l'efecte de *A* és pràcticament nul en la partició octanol-aigua ja que el coeficient *a* és pràcticament zero (el que significa que l'octanol i l'aigua tenen una basicitat per pont d'hidrogen equivalent). En canvi, el descriptor *A* és determinant en el sistema de partició hexà-aigua degut a l'elevat valor absolut del coeficient *a*, el qual és negatiu i contribueix a disminuir el coeficient de partició. En canvi, el valor positiu de *a* en el sistema TTAB-tampó aquós, entre d'altres, indica que l'acidesa del solut per pont d'hidrogen (*A*) contribueix a la seva transferència cap a la fase pseudoestacionària i, per tant, fa augmentar la retenció cromatogràfica en aquest sistema de MEKC.

1.4. SIMILITUD ENTRE SISTEMES BIOLÒGICS I FISICOQUÍMICS DE PARTICIÓ

L'aplicabilitat del model de paràmetres de solvatació tant a sistemes biològics com fisicoquímics de partició té un avantatge fonamental i és que ofereix un marc comú per comparar-los. Els sistemes caracteritzats segons l'equació general de solvatació (equació 5) es poden comparar en base a les seves caracteritzacions, és a dir, estudiant la relació entre els seus coeficients LSER per tal d'avaluar la seva similitud. Donat que la mateixa equació 5 es pot aplicar satisfactòriament a ambdós tipus de processos, ha de ser possible trobar sistemes biològics i fisicoquímics similars. En aquest cas, les propietats d'ambdós processos estaran correlacionades i es podrà estimar la propietat de biopartició d'una substància a partir de la seva mesura fisicoquímica.

El model de paràmetres de solvatació, per tant, no només permet conèixer les interaccions solut-solvent que governen un procés de partició, així com predir la propietat en estudi per a nous compostos si es coneixen els seus descriptors, sinó que també és de gran utilitat per comparar sistemes caracteritzats i poder trobar sistemes fisicoquímics que modelin processos de biopartició.

Tal com s'explicarà a continuació, existeixen diferents tipus d'eines, tant de caràcter visual com matemàtic, per comparar les caracteritzacions de sistemes i avaluar així la seva similitud.

L'anàlisi per components principals (PCA) és una tècnica quimiomètrica que ha estat emprada per molts autors [1,80,85,121-123] per obtenir una distribució visual de sistemes en funció dels seus coeficients LSER. Aquesta tècnica [124] transforma la matriu original de dades (coeficients dels sistemes a comparar) en dues noves matrius, *scores* i *loadings*, i també s'obtenen els components principals (PCs). Els PCs són les noves variables i expliquen la variància de la matriu de dades originals amb menys paràmetres que els inicials. La matriu de *loadings* explica com els paràmetres inicials (coeficients dels sistemes) contribueixen a cada PC. La matriu de *scores* proporciona el valor de cada PC pels sistemes considerats a la matriu original. Per tant, la

representació dels valors de *scores* dels dos primers PCs, PC1 i PC2 (ja que són els que expliquen la major variància de les dades originals), permet obtenir una distribució gràfica de fàcil visualització dels sistemes en funció de les similituds i diferències entre els seus coeficients.

Cal tenir en compte que si la matriu original no ha estat pretractada, l'anàlisi per components principals posarà de relleu les diferències entre sistemes amb valors extrems de coeficients. Per exemple, distribuirà de forma molt allunyada dos sistemes que tinguin un valor molt elevat del coeficient *v*, en un cas, i un valor molt baix de *v*, en l'altre cas. No obstant, pot ser que aquests dos sistemes presentin el mateix comportament pel que fa a les interaccions solut-solvent (selectivitat) si els seus coeficients són proporcionals. Per aquest motiu és recomanable pretractar la matriu original dels coeficients. El pretractament de normalització per files (cada fila correspon a la caracterització d'un sistema) assumeix que els cinc coeficients de cada sistema (*e*, *s*, *a*, *b*, *v*) són els components d'un vector ($\vec{w}$) que té la següent longitud *l* :

$$l = \sqrt{e^2 + s^2 + a^2 + b^2 + v^2} \tag{6}$$

Per a cada sistema, els coeficients es normalitzen tal com es detalla a continuació:

$$e_u = \frac{e}{l} \tag{7}$$

$$s_u = \frac{s}{l} \tag{8}$$

$$a_u = \frac{a}{l} \tag{9}$$

$$b_u = \frac{b}{l} \tag{10}$$

$$v_u = \frac{v}{l} \tag{11}$$

on *l* és la longitud del vector i (e_u, s_u, a_u, b_u, v_u) esdevenen els components del vector normalitzat ($\vec{w}_u$).

Aquest pretractament permet que tots els sistemes amb coeficients proporcionals tinguin els mateixos coeficients normalitzats. Per tant, quan la matriu de coeficients normalitzats s'analitza per components principals és quan s'aconsegueix una distribució dels sistemes segons la seva selectivitat, és a dir, s'agrupen tenint en compte el pes relatiu dels seus coeficients [80]. Els sistemes que es trobin més pròxims en la distribució de PCA seran els que presentin major similitud, mentre que sistemes molt separats espacialment tindran característiques molt diferents respecte a les interaccions solut-solvent que estableixen amb els soluts.

La Figura 12 mostra un exemple en què es comparen diversos sistemes de MEKC mitjançant PCA després d'haver normalitzat els seus coeficients (equacions 7-11). Al gràfic de *scores* representat a la figura, el sistema de LPFOS apareix molt separat espacialment de la resta de sistemes, el que indica que té una selectivitat molt diferent. Aquest fet s'atribueix a que el LPFOS, a diferència dels tensioactius que formen els altres sistemes, conté múltiples àtoms de fluor a la seva cadena alquílica que li atorguen propietats molt diferents. A la figura també s'observa que la resta de sistemes apareixen força pròxims entre ells, el que significa que presenten un comportament similar respecte a les interaccions que estableixen amb els soluts. Tots ells són sistemes de MEKC formats per tensioactius amb estructures semblants, bàsicament cadenes hidrocarbonades o anells. Destaca la proximitat dels sistemes de LDS i SDS, els de SDC i SC i els de TTAB i HTAB. Els sistemes que formen aquestes parelles tenen estructures especialment semblants, el que fa que tinguin una selectivitat molt similar que queda reflectida al gràfic mitjançant una gran proximitat espacial.

El principal avantatge que presenta l'anàlisi per components principals és que permet manipular de forma conjunta un gran nombre de dades i, per tant, és possible comparar a la vegada la similitud/dissimilitud de molts sistemes caracteritzats mitjançant l'equació de solvatació. No obstant, les distribucions visuals obtingudes per PCA canvien quan s'inclouen o es treuen dades de la matriu original, és a dir, quan s'afegeixen o es treuen sistemes per comparar. La raó és que aleshores canvia el

conjunt de valors de coeficients a representar i, en conseqüència, varia el pes de cada coeficient en els PCs, tot comportant una distorsió de la representació visual dels sistemes. Per aquest motiu molts cops es prefereix fer servir mètodes de comparació de sistemes de caire més matemàtic, com els que es detallaran seguidament.

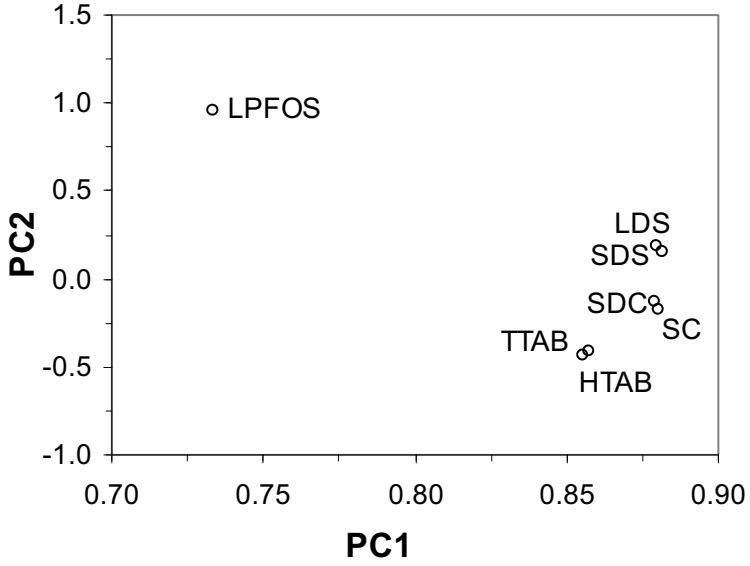


Figura 12. Gràfic de *scores* de diversos sistemes de MEKC respecte als dos primers components principals després de normalitzar els seus coeficients [118].

Un mètode proposat per Abraham i Rosés [114] i utilitzat en nombrosos treballs [21,22,70,79] consisteix en estudiar la proporcionalitat dels coeficients dividint-los entre el més gran (normalment *v*) i comparant les relacions *e*/*v*, *s*/*v*, *a*/*v* i *b*/*v* de diferents sistemes. Si aquestes relacions són semblants significa que els sistemes comparats tenen coeficients proporcionals i es poden considerar similars.

Ishihama i Asakawa [125] van proposar un altre mètode, el qual ha estat emprat també per altres autors [126,127]. Es basa en considerar els coeficients com a components d'un vector en cinc dimensions i calcular el cosinus de l'angle (cos θ) que formen els vectors de dos sistemes com a mesura de la seva similitud (Figura 13). Si els dos sistemes són matemàticament similars, el cosinus de l'angle ha de tenir un valor molt proper a 1.

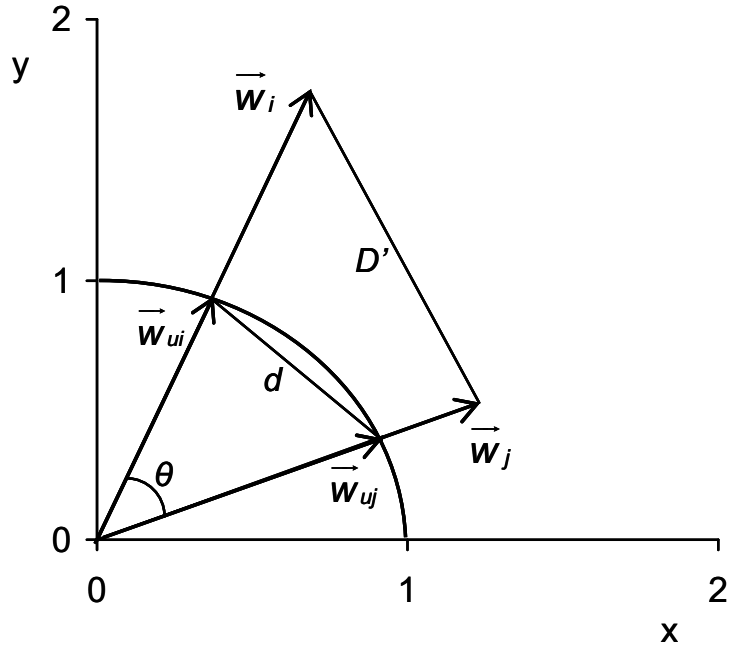


Figura 13. Representació esquemàtica en dues dimensions dels vectors de coeficients LSER de dos sistemes ($\vec{w}_i$, $\vec{w}_j$), els seus corresponents vectors normalitzats ($\vec{w}_{ui}$, $\vec{w}_{uj}$), i els paràmetres θ, *D'* i *d*.

Abraham i Martins [112] van proposar una altre paràmetre de comparació de sistemes, la distància *D'*, utilitzat també en altres treballs [126,127]. En aquest cas, els coeficients del sistema es consideren un punt en un espai de cinc dimensions, de manera que el càlcul de la distància entre dos punts (*D'*) mesura la similitud entre els dos sistemes comparats (Figura 13). Es considera que un valor de *D'* comprès o inferior a 0.5-0.8 és indicador de que els dos sistemes són químicament similars.

El nostre grup de recerca va desenvolupar un nou paràmetre de distància [54], *d*, que es considera molt útil per comparar sistemes caracteritzats mitjançant l'equació general de solvatació [85,116]. Es tracta d'una eina matemàtica que també considera un espai de cinc dimensions i mesura la distància entre els vectors de dos sistemes, però en aquest cas després d'haver normalitzat els coeficients (equacions 7-11). Per tant, el paràmetre *d* correspon a la distància entre dos vectors unitaris normalitzats

(Figura 13). La distància *d* informa de la similitud entre sistemes, de manera que com més petita sigui la *d* entre dos sistemes més matemàticament semblants seran. Aquests seran idèntics o proporcionals en el cas de *d* = 0. La distància *d* es pot calcular mitjançant la següent expressió:

$$d=\sqrt{(e_{ui}-e_{uj})^2+(s_{ui}-s_{uj})^2+(a_{ui}-a_{uj})^2+(b_{ui}-b_{uj})^2+(v_{ui}-v_{uj})^2} \quad (12)$$

on el subíndex *u* fa referència a la prèvia normalització dels coeficients i els subíndexs *i*, *j* representen els dos sistemes comparats.

Cal mencionar que els paràmetres de distància *D'* i *d* no donen la mateixa informació [54,116]. La distància *D'* serà gran quan les longituds dels vectors siguin molt diferents entre si, però en canvi els vectors unitaris d'aquests sistemes poden tenir una distància *d* petita.

D'altra banda, el paràmetre *d* [54] sí que està directament relacionat amb el cos θ proposat per Ishihama i Asakawa. Tal com il·lustra la Figura 13, els dos vectors normalitzats i la distància *d* formen un triangle isòsceles, de manera que el cos θ i la *d* es poden relacionar per trigonometria:

$$\cos\theta = \cos(2\arcsin(d/2)) \quad (13)$$

Per tant, es tracta de dues eines matemàtiques que mesuren el mateix concepte. No obstant, l'escala del cosinus no és lineal i només valors de cosinus entre 0.97 i 1 indicaran sistemes similars. En canvi, la distància *d* sí que té una escala lineal i indica la semblança entre sistemes mitjançant un interval més ampli, entre 0-0.25 aproximadament. Conseqüentment, la distància *d* es considera més fàcil i intuïtiva per mesurar la similitud entre parells de sistemes caracteritzats mitjançant l'equació general de solvatació.

La Taula 3 mostra un exemple de la utilització dels paràmetres cos θ, *D'* i *d* per comparar parelles de sistemes. Concretament, a la taula es comparen tres sistemes de biopartició i un sistema cromatogràfic ja esmentats al llarg d'aquesta introducció. Al

comparar l'absorció intestinal amb el sistema cromatogràfic RP18, els valors dels tres paràmetres (cos θ << 0.97, *D'* >> 0.8, *d* >> 0.25) condueixen a la mateixa conclusió: es tracta de dos processos molt diferents. En canvi, aquesta concordança entre els tres paràmetres ja no s'observa al comparar les altres dues parelles de sistemes. Quan es compara la permeabilitat en pell i la narcosi en capgrossos amb el sistema cromatogràfic, tant el cos θ (amb un valor comprès entre 0.97 i 1) com la *d* (amb un valor inferior a 0.25) indiquen que ambdós sistemes biològics són semblants al cromatogràfic. No obstant, el valor obtingut pel paràmetre *D'* és força gran (superior a 0.5-0.8) i, per tant, no proporciona la mateixa informació. Aquests dos exemples il·lustren que és possible que dos sistemes similars puguin estar separats per una distància *d* petita i una distància *D'* gran.

Taula 3. Valors dels paràmetres cos θ, *D'* i *d* obtinguts entre tres sistemes de biopartició i el sistema cromatogràfic format per la columna RP18-MeCN/tampó aquós.

Sistema biològic	**Columna RP18-MeCN/tampó aquós** [54]		
	cos θ	***D'***	***d***
Absorció intestinal [1]	0.721	1.863	0.747
Permeabilitat a la pell [112]	0.980	1.644	0.200
Narcosi en capgrossos [3]	0.975	2.176	0.223

Determinar la similitud entre sistemes biològics i fisicoquímics és útil per identificar possibles sistemes fisicoquímics que emulin processos de biopartició. No obstant, la bondat amb què un sistema biològic és modelat mitjançant un sistema fisicoquímic ve donada per la precisió que s'obté al correlacionar les seves propietats. Aquesta precisió està òbviament relacionada amb la similitud dels sistemes, però també ve determinada per la incertesa de les dades que es correlacionen. Donat que cap dels paràmetres descrits per comparar sistemes permet quantificar aquesta precisió, es posa de manifest la necessitat de disposar d'una metodologia complementària. Una metodologia que permeti estimar la precisió que s'espera obtenir a les correlacions biològico-fisicoquímiques i que, en funció de la bondat de les correlacions, estableixi

uns criteris sistemàtics per seleccionar sistemes fisicoquímics adients per emular processos de biopartició. Aquesta metodologia ha estat objecte de desenvolupament i aplicació al llarg de la present tesi.

PUBLICACIONS

A continuació es presenta el llistat de publicacions derivades de la recerca desenvolupada durant la realització de la tesi doctoral, així com una còpia de cadascuna d'elles:

ARTICLE I: *Solute-solvent interactions in micellar electrokinetic chromatography: VII. Characterization of sodium cholate-sodium deoxycholate mixed-micellar systems*
Marta Hidalgo-Rodríguez, Elisabet Fuguet, Clara Ràfols, Martí Rosés
J. Chromatogr. A 2010, *1217*, 1701-1708

ARTICLE II: *Estimation of biological properties by means of chromatographic systems: evaluation of the factors that contribute to the variance of biological-chromatographic correlations*
Marta Hidalgo-Rodríguez, Elisabet Fuguet, Clara Ràfols, Martí Rosés
Anal. Chem. 2010, *82*, 10236-10245

ARTICLE III: *Modeling nonspecific toxicity of organic compounds to the fathead minnow fish by means of chromatographic systems*
Marta Hidalgo-Rodríguez, Elisabet Fuguet, Clara Ràfols, Martí Rosés
Anal. Chem. 2012, *84*, 3446-3452

ARTICLE IV: *Performance of chromatographic systems to model soil-water sorption*
Marta Hidalgo-Rodríguez, Elisabet Fuguet, Clara Ràfols, Martí Rosés
J. Chromatogr. A 2012, *1252*, 136-145

ARTICLE V: *Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds*
Marta Hidalgo-Rodríguez, Sara Soriano-Meseguer, Elisabet Fuguet, Clara Ràfols, Martí Rosés
Eur. J. Pharm. Sci. 2013,
DOI: http://dx.doi.org/10.1016/j.ejps.2013.04.005

Journal of Chromatography A, 1217 (2010) 1701–1708

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Solute–solvent interactions in micellar electrokinetic chromatography: VII. Characterization of sodium cholate–sodium deoxycholate mixed-micellar systems

Marta Hidalgo-Rodríguez, Elisabet Fuguet, Clara Ràfols, Martí Rosés*

Departament de Química Analítica and Institut de Biomedicina de la Universitat de Barcelona, Universitat de Barcelona, Martí i Franquès, 1-11, E-08028 Barcelona, Spain

ARTICLE INFO

Article history:
Received 8 September 2009
Received in revised form
21 December 2009
Accepted 4 January 2010
Available online 11 January 2010

Keywords:
Micellar electrokinetic chromatography
Solvation parameter model
Solute descriptors
Sodium cholate
Sodium deoxycholate
Surfactants

ABSTRACT

Micellar electrokinetic chromatography (MEKC) systems with mixed pseudostationary phases of the bile surfactants sodium cholate (SC) and sodium deoxycholate (SDC) have been characterized by means of the solvation parameter model. The importance of characterizing systems with an appropriate set of solutes that embrace a wide range of descriptor values has been proven as they can significantly influence the value of the system constants. The fit of the solvation parameter model to the experimental log *k* data has been compared for each SC–SDC system when the Abraham descriptors and the Poole optimized descriptors, recently proposed, are used. In both cases, the variation in MEKC surfactant composition results in similar changes in the coefficients of the correlation equations, which in turn leads to similar information on solute–solvent and solute–micelle interactions. It is demonstrated that SDC is more hydrogen-bond acidic and hydrophobic but slightly less polarizable than SC. Systems with intermediate selectivity are obtained through mixtures of both surfactants.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Micellar electrokinetic chromatography (MEKC) is a well-established separation technique based on electrophoretic and electroosmotic principles whose distinguishing feature is the addition of a surfactant above its critical micelle concentration to the separation buffer. As a consequence, solutes are separated not only by migration but also by distribution between the aqueous phase (bulk electrolyte) and the pseudostationary phase (charged micelles), which allows the separation of complex mixtures of both neutral and ionized solutes [1–3]. The uncharged solutes are separated according to their distribution constants between the aqueous and micellar phases, whereas the charged solutes are separated as a result of their distribution between phases and their electrophoretic mobility. The main advantage of MEKC is the possibility of modifying the migration behaviour and separation in a very easy and flexible way, only by changing the nature of the pseudostationary phase. This can be achieved by proper selection of the surfactant type or by the addition of complexing agents (e.g., cyclodextrins, urea, etc.) or organic solvents to the separation solution [3–8]. However, it is generally accepted that the choice of surfactant is the most important consideration for varying the chemical nature of the medium and optimizing selectivity [8,9].

The solvation parameter model is strongly recommended to characterize MEKC systems [10] because it allows a better understanding of the types and the relative strength of the chemical interactions that control retention. This model, developed by Abraham [11], provides information not only about how different the characterized systems are, but also about the magnitude of the different interactions between the phases (aqueous phase and micellar phase in MEKC) and neutral solutes. It is based on linear free energy relationships (LFERs) established with solute descriptors. The suitable form for MEKC can be written as:

$$\log k = c + eE + sS + aA + bB + vV \quad (1)$$

where k is the MEKC retention factor and E, S, A, B and V are the solute descriptors proposed by Abraham [11]. E is an excess molar refraction, S is the solute dipolarity/polarizability, A and B are the solute's effective hydrogen-bond acidity and hydrogen-bond basicity, respectively, and V is McGowan's solute characteristic volume. The correlation coefficients of Eq. (1) are characteristic of the system (micellar phase + aqueous buffer) and reflect the system properties that are complementary to the corresponding solute property, i.e. they are system constants.

Since the solute property correlated to the solvation descriptors (log k for MEKC) can be any property related to free energy,

* Corresponding author. Tel.: +34 93 403 92 75; fax: +34 93 402 12 33.
E-mail address: marti@apolo.qui.ub.es (M. Rosés).

0021-9673/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2010.01.001

the solvation parameter model has also been applied to the characterization of many other physicochemical systems apart from MEKC systems [10,12,13], as well as it has been demonstrated to be extremely useful for the characterization of many biological processes [14–16].

As the MEKC retention factor is directly related to the partition of the solute between the micellar phase and the aqueous phase, *e* refers to the difference in capacity of each phase to interact with solute π- and n-bonding electrons; *s* is a measure of the difference in capacity of the micellar phase and the aqueous phase to take part in dipole–dipole and dipole-induced dipole interactions; *a* and *b* represent the differences in hydrogen-bond basicity and acidity, respectively, between the micellar and the aqueous phases; *v* is a measure of the relative ease of forming a cavity for the solute in the buffer and the micelles. The intercept of the correlation, *c*, is related to the pseudostationary/aqueous buffer phase ratio and its value influences the retention time but not the selectivity. For any MEKC system the coefficients of the correlation equation can be obtained by multiple linear regression analysis of the experimental log *k* values acquired for a group of varied solutes with known descriptor values. As the proper selection of an adequate collection of solutes is extremely important to accurately determine the coefficients of Eq. (1), the literature gives some recommendations [9,10,17]. In general terms, this set of solutes must have properties sufficiently varied to define all interactions in Eq. (1) and be of sufficient number to establish the statistical validity of the model. In order to obtain a good fit of the solvation parameter model, well-characterized descriptor values with the minimal uncertainty for the selected set of compounds are also required. Abraham et al. [18] constructed the available database of descriptor values, which collects 4000 solutes characterized by all or some of their descriptors. Recently, Poole et al. [19] have proposed a collection of new descriptor values which were optimized for chromatographic methods using well-characterized systems of gas chromatography, reversed-phase chromatography, micellar electrokinetic chromatography, and liquid–liquid partitioning. These Poole optimized descriptors are expected to afford a better fit to the experimental data and smaller standard deviations for the system constants than the Abraham descriptors.

With the wide range of surfactants commercially available, several individual MEKC surfactants have been characterized through the solvation parameter model [3,4,8,9,13,17,20]. Such model has also been applied to the characterization of mixed-micellar systems [5,21–23]. The attention that has been paid to the use of mixed micelles is due to the fact that the properties of the pseudostationary phase, and therefore the coefficients of Eq. (1) for the MEKC system, can be continuously varied by changing the proportion of the surfactants in the mixture. This is especially interesting as it can facilitate the fine-tuning of selectivity, as has been reported in some studies carried out to classify the chemical selectivity of electrokinetic chromatography systems [13,23].

In this work, we characterize mixed-micellar systems of sodium cholate (SC) and sodium deoxycholate (SDC) (Fig. 1). The manuscript is part of a series devoted to the characterization and selectivity of different MEKC systems, composed of individual or mixed micelles, and to the study of several factors that influence the results. Some individual surfactant systems included in this series are the anionic sodium dodecyl sulfate (SDS), lithium dodecyl sulfate (LDS), lithium perfluorooctanesulfonate (LPFOS), SC, SDC and the cationic tetradecyltrimethylammonium bromide (TTAB) and hexadecyltrimethylammonium bromide (HTAB). Regarding mixed-micellar systems, this series has only included the characterization of the following binary mixtures: SDS with the neutral surfactant Brij 35, and two surfactants with very different properties, LDS and LPFOS. In the present work, we characterize another mixture composed of two individual bile salt surfactants that are very similar, SC and SDC. We study how their composition in the mixtures affects the coefficients of the solvation parameter model. The effect of the set of solutes and descriptors used and the pH of the buffer on the characterization coefficients are also discussed.

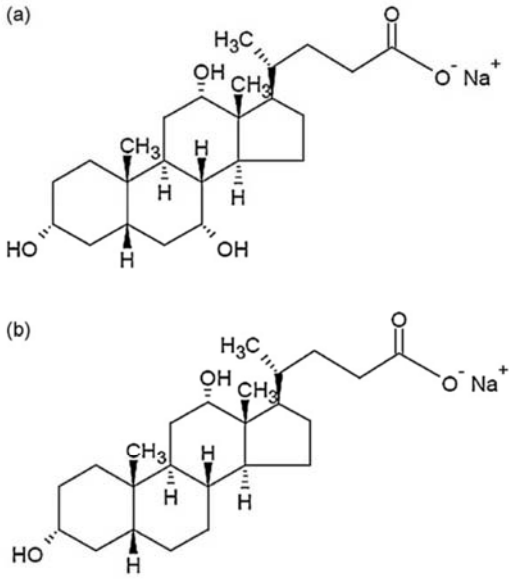


Fig. 1. Chemical structure of the monomers of the bile surfactants: (a) sodium cholate (SC) and (b) sodium deoxycholate (SDC).

2. Experimental

2.1. Apparatus and conditions

All separations were performed with a Beckman P/ACE System 5500 Capillary Electrophoresis with a UV diode array detector. The fused-silica separation capillaries were 40 cm effective length × 50 μm I.D. and were obtained from Composite Metal Services Ltd (Shipley, West Yorkshire, UK). The capillary was activated by the following washing sequence: water (10 min), 1 *M* NaOH (10 min), water (5 min), 0.1 *M* NaOH (10 min), water (5 min) and separation solution (20 min). This sequence was also used for conditioning the capillary when the pseudostationary phase was changed. As daily conditioning when using the same pseudostationary phase, the capillary was flushed with water for 5 min, followed by 0.1 *M* NaOH for 5 min, water for 5 min and separation solution for 20 min. Prior to each separation when using the same pseudostationary phase the capillary was flushed with separation solution for 3 min. Retention measurements were made at 25 °C and +15 kV. Detection was at 214 nm.

All systems (separation solutions) were 20 m*M* in aqueous buffer. The systems at pH 8 were prepared by dissolving the surfactants in NaH_2PO_4:$Na_2B_4O_7$ (65:35) buffer. For 80 m*M* SC at pH 7 the system was prepared by dissolving the surfactant in NaH_2PO_4:Na_2HPO_4 (50:50) buffer. For 80 m*M* SC at pH 6 the system was prepared by dissolving the surfactant in NaH_2PO_4:Na_2HPO_4 (50:50) buffer and neutralizing with HCl. Surfactant concentrations were similar to the ones used by other authors [4,8,17], and they were chosen to be well above the critical micelle concentration (CMC) in order to obtain a reasonable volume of pseudostationary phase and an acceptable elution window. The proportion of each surfactant in the mixtures was chosen in such a way that the total surfactant concentration ranged from 40 m*M* to 80 m*M*, with the total surfactant concentration varied in increments of 10 m*M*.

Solutes were dissolved in methanol (used as electroosmotic flow marker) and contained 2 mg mL^{-1} of dodecanophenone as micellar marker [24]. The injection of methanol produces a local disruption of the micellar phase. For low micellar concentrations, the disruption causes peak splitting [25], which can be avoided by working at higher concentrations such as those used in this work. The concentration of the solutes was 2 mg mL^{-1}, except for the alcohols which were 40% (v/v) in order to obtain measurable absorbance. All solutions were filtered through 0.45-μm nylon syringe filters obtained from Albet (Dassel, Germany). Samples were introduced into the capillary by applying a pressure of 0.5 p.s.i. for 1 s (1 p.s.i. = 6894.76 Pa). All measurements were taken in triplicate.

2.2. Reagents and materials

Hydrochloric acid (25% in water), sodium hydroxide (>99%), sodium dihydrogenphosphate monohydrate (>99%), disodium hydrogenphosphate (>99%) and methanol (HPLC grade) were from Merck (Darmstadt, Germany). Disodium tetraborate decahydrate (>99.5%) was from Sigma (Steinheim, Germany). SC (>98%) was from Fluka (Steinheim, Germany). SDC (98%) and dodecanophenone (98%) were from Aldrich (Steinheim, Germany). Water was purified by a Milli-Q plus system from Millipore (Bedford, MA, USA), with a resistivity of 18.2 MΩ cm. The test solutes employed were reagent grade or better and obtained from several manufacturers (Merck (Darmstadt, Germany), Sigma (Steinheim, Germany), Fluka (Steinheim, Germany), Aldrich (Steinheim, Germany), Carlo Erba (Milano, Italy), Baker (Deventer, Netherlands)).

2.3. Calculation

The MEKC retention factor, k, was calculated according to Eq. (2) with the migration time of methanol used to determine the electroosmotic flow (t_0), and the migration time of dodecanophenone used to determine the migration time of the micelles (t_m). t_R is the solute migration time:

$$k = \frac{t_R - t_0}{(1 - (t_R/t_m))t_0} \quad (2)$$

Microsoft Excel XP was used to perform data calculations and multiple linear regression analysis.

3. Results and discussion

3.1. Influence of the set of solutes on the characterization of systems

Although most MEKC systems are usually characterized at pH 7, it was decided to characterize all the systems studied here at pH 8 because SDC presents low solubility at pH 7.

The 80 m*M* SC pure surfactant system at pH 8 was the first one to be characterized by means of the solvation parameter model through Eq. (1) by analysis of the log k data of a series of 69 solutes with known E, S, A, B and V descriptor values. The studied solutes and their descriptors are given in Table 1, where both the Abraham descriptor values [18] and the optimized descriptor values recently proposed by Poole et al. [19] for some of them are presented. In this section, only the Abraham descriptors have been employed for the system characterizations. The collection of solutes has been selected according to a study carried out in a previous work [17] in which they were considered an appropriate group of compounds for MEKC characterizations covering a wide range of solute descriptor values. Initially, the compounds 4-chlorophenol, catechol, resorcinol, hydroquinone, 2-naphthol, and 1,2,3-trihydroxybenzene were not included in the characterization set because they have aqueous pK_a values between 9 and 10 and therefore they are susceptible to be partially ionized at the working pH (pH 8). The k values obtained were calculated by Eq. (2). The system constants and the statistics for the fit of the solvation parameter model to the experimental log k data for the 80 m*M* SC system at pH 8 are summarized in Table 2 (correlation a). The compounds of the set that were not included in the final correlation (because they coeluted with the electroosmotic flow marker or were outliers) are also detailed in the table.

Noticeable differences were observed between the coefficients in Eq. (1) (system constants) for the 80 m*M* SC system characterized at pH 8 in this work, and a 80 m*M* SC system characterized at pH 7 in a previous work [17]. Since literature reports that pH does not have effect on the system constants [8], we decided to characterize 80 m*M* SC at pH 7, as well as, at pH 6 in order to prove if the change of pH was or not responsible for the differences found between the coefficients.

The characterizations of the 80 m*M* SC system at pH 7 and 6, unlike the 80 m*M* SC system at pH 8, were carried out including the substances 4-chlorophenol, catechol, resorcinol, hydroquinone, 2-naphthol, and 1,2,3-trihydroxybenzene because these phenols are totally neutral at these values of working pH. The system constants and the statistics for the fit of the solvation parameter model to the experimental log k data for each correlation are also shown in Table 2 (correlations e and f).

Comparing the coefficients of the correlations a, e, and f detailed in Table 2, it is observed that they are nearly identical for 80 m*M* SC at pH 7 and 6, whereas those of the system at pH 8 are different. Since the only difference between the characterization at pH 8 and at pH 7 and 6 is the fact that the substances 4-chlorophenol, catechol, resorcinol, hydroquinone, 2-naphthol, and 1,2,3-trihydroxybenzene were not included at pH 8, it was decided to calculate again the correlations at pH 7 and 6 without these phenols in order to investigate if they were responsible for the discrepancies between the coefficients. The results obtained are also summarized in Table 2 (correlations b and c).

After removing these six phenols, coefficients became very similar, independently of the pH of the separation solution. According to this fact, it is confirmed that changing the pH (after excluding ionized solutes) has little influence on the selectivity, as it is reported in the literature [8]. Taking everything into account, what is highly remarkable is the role that 4-chlorophenol, catechol, resorcinol, hydroquinone, 2-naphthol, and 1,2,3-trihydroxybenzene play in these characterizations. These phenolic compounds are especially important because they are of the few ones of the set of solutes with a high value for the A descriptor. As a consequence, if 4-chlorophenol, catechol, resorcinol, hydroquinone, 2-naphthol, and 1,2,3-trihydroxybenzene are not included in the characterizations, a very narrow range of values for the A descriptor is used and the accuracy in the determination of system constants diminishes. In fact, it can be observed that when these phenolic compounds are not included in the correlations, the a coefficient is approximately 50% higher than the value obtained when the more varied set of solutes that includes them is used. Therefore, this reveals how important is to employ a varied set of solutes that span a wide range of descriptor values when systems are characterized by means of the solvation parameter model.

As the compounds 4-chlorophenol, catechol, resorcinol, hydroquinone, 2-naphthol, and 1,2,3-trihydroxybenzene have a great effect on the system constants, it was studied if their ionisation degree was significant enough to include them in the correlation at pH 8. These phenolic compounds were injected in 80 m*M* SC at pH 8. Afterwards, the log k values obtained for the 80 m*M* SC system at pH 8 were plotted against those obtained before at pH 6 and 7 for all the solutes of the set (Fig. 2a). According to the proved little effect of the pH on the selectivity, good correlations should be obtained

Table 1
Solute descriptors used in the solvation parameter model.

Solute	V	Abraham descriptors [18]				Poole optimized descriptors [19]			
		E	S	A	B	E	S	A	B
Propan-1-ol	0.5900	0.236	0.42	0.37	0.48				
Propan-2-ol	0.5900	0.212	0.36	0.33	0.56				
Butan-1-ol	0.7309	0.224	0.42	0.37	0.48	0.224	0.44	0.34	0.52
Pentan-1-ol	0.8718	0.219	0.42	0.37	0.48	0.219	0.44	0.34	0.52
Pentan-3-ol	0.8718	0.218	0.36	0.33	0.56	0.218	0.40	0.27	0.57
Propan-1,3-diol	0.6487	0.397	0.91	0.77	0.85				
Butan-1,4-diol	0.7860	0.395	0.93	0.72	0.90				
Pentan-1,5-diol	0.9305	0.388	0.95	0.72	0.91				
Thiourea	0.5696	0.840	0.82	0.77	0.87				
Benzene	0.7164	0.610	0.52	0.00	0.14	0.608	0.51	0.00	0.14
Toluene	0.8573	0.601	0.52	0.00	0.14	0.606	0.50	0.00	0.14
Ethylbenzene	0.9982	0.613	0.51	0.00	0.15	0.613	0.50	0.00	0.14
Propylbenzene	1.1391	0.604	0.50	0.00	0.15	0.610	0.50	0.00	0.14
Butylbenzene	1.2800	0.600	0.51	0.00	0.15	0.595	0.50	0.00	0.14
p-Xylene	0.9982	0.613	0.52	0.00	0.16	0.615	0.49	0.00	0.16
Naphthalene	1.0854	1.340	0.92	0.00	0.20	1.240	0.91	0.00	0.19
Chlorobenzene	0.8388	0.718	0.65	0.00	0.07	0.718	0.66	0.00	0.06
Bromobenzene	0.8914	0.882	0.73	0.00	0.09	0.882	0.72	0.00	0.09
Anisole	0.9160	0.708	0.75	0.00	0.29	0.712	0.77	0.00	0.31
Benzaldehyde	0.8730	0.820	1.00	0.00	0.39	0.813	1.03	0.00	0.39
Acetophenone	1.0139	0.818	1.01	0.00	0.48	0.806	1.03	0.00	0.50
Propiophenone	1.1548	0.804	0.95	0.00	0.51	0.804	1.03	0.00	0.50
Butyrophenone	1.2957	0.797	0.95	0.00	0.51	0.798	1.03	0.00	0.50
Valerophenone	1.4366	0.795	0.95	0.00	0.50	0.795	1.03	0.00	0.50
Heptanophenone	1.7184	0.720	0.95	0.00	0.50				
Benzophenone	1.4808	1.447	1.50	0.00	0.50	1.224	1.33	0.00	0.58
Methyl benzoate	1.0726	0.733	0.85	0.00	0.46	0.738	0.92	0.00	0.44
Benzyl benzoate	1.6804	1.264	1.42	0.00	0.51	1.264	1.30	0.00	0.59
Benzonitrile	0.8711	0.742	1.11	0.00	0.33	0.742	1.14	0.00	0.33
Aniline	0.8162	0.955	0.96	0.26	0.50	0.955	1.00	0.25	0.43
o-Toluidine	0.9751	0.970	0.90	0.23	0.59	0.966	1.05	0.19	0.49
3-Chloroaniline	0.9390	1.050	1.10	0.30	0.36				
4-Chloroaniline	0.9390	1.060	1.10	0.30	0.35	1.017	1.13	0.37	0.31
2-Nitroaniline	0.9904	1.180	1.37	0.30	0.36	1.182	1.44	0.39	0.35
3-Nitroaniline	0.9904	1.200	1.71	0.40	0.35	1.248	1.60	0.47	0.42
4-Nitroaniline	0.9904	1.220	1.91	0.42	0.38	1.236	1.83	0.60	0.34
Nitrobenzene	0.8906	0.871	1.11	0.00	0.28	0.846	1.14	0.00	0.27
2-Nitroanisole	1.0902	0.965	1.34	0.00	0.38				
Benzamide	0.9728	0.990	1.50	0.49	0.67	1.258	1.34	0.65	0.66
4-Aminobenzamide	1.0726	1.340	1.94	0.80	0.94				
Acetanilide	1.1137	0.870	1.36	0.46	0.69	0.960	1.12	0.53	0.71
4-Chloroacetanilide	1.2357	0.980	1.50	0.64	0.51				
Phenol	0.7751	0.805	0.89	0.60	0.30	0.769	0.76	0.72	0.32
3-Methylphenol	0.9160	0.822	0.88	0.57	0.34	0.810	0.78	0.67	0.35
2,3-Dimethylphenol	1.0569	0.850	0.90	0.52	0.36	0.866	0.77	0.59	0.40
2,4-Dimethylphenol	1.0569	0.840	0.80	0.53	0.39				
Thymol	1.3387	0.822	0.79	0.52	0.44				
Furan	0.5363	0.369	0.53	0.00	0.13				
2,3-Benzofuran	0.9053	0.888	0.83	0.00	0.15	0.921	0.77	0.00	0.19
Quinoline	1.0443	1.268	0.97	0.00	0.51	1.268	1.09	0.00	0.60
Pyrrole	0.5774	0.613	0.73	0.41	0.29				
Pyrimidine	0.6342	0.606	1.00	0.00	0.65				
Antipyrine	1.5502	1.320	1.50	0.00	1.48				
Caffeine	1.3632	1.500	1.60	0.00	1.33	1.557	1.62	0.00	1.27
Corticosterone	2.7389	1.860	3.43	0.40	1.63				
Cortisone	2.7546	1.960	3.50	0.36	1.87				
Hydrocortisone	2.7975	2.030	3.49	0.71	1.90				
Estradiol	2.1988	1.800	3.30	0.88	0.95				
Estriol	2.2575	2.000	3.36	1.40	1.22				
Monuron	1.4768	1.140	1.50	0.47	0.78				
Myrcene	1.3886	0.483	0.29	0.00	0.21				
α-Pinene	1.2574	0.446	0.14	0.00	0.12				
Geraniol	1.4903	0.513	0.63	0.39	0.66				
4-Chlorophenol[a]	0.8975	0.915	1.08	0.67	0.20	1.016	0.79	0.89	0.21
Catechol[a]	0.8338	0.970	1.10	0.88	0.47				
Resorcinol[a]	0.8338	0.980	1.00	1.10	0.58	0.968	0.91	1.37	0.51
Hydroquinone[a]	0.8338	1.000	1.00	1.16	0.60				
2-Naphthol[a]	1.1441	1.520	1.08	0.61	0.40	1.457	1.18	0.81	0.35
1,2,3-Trihydroxybenzene[a]	0.8925	1.165	1.35	1.35	0.62				

[a] Phenols with pK_a values between 9 and 10 and therefore susceptible to be partially ionized at pH 8.

Table 2
Coefficients in Eq. (1) for the 80 m*M* SC micellar separation system at pH 8, 7 and 6.

System	Coefficients						Statistics				Solutes excluded from the correlations
	c	*e*	*s*	*a*	*b*	*v*	*n*	*r*	SD	*F*	
Phenols with pK_a 9–10 not included											
(a) 80 m*M* SC at pH 8	−1.544 (0.110)	1.022 (0.185)	−0.849 (0.138)	0.152 (0.134)	−1.955 (0.132)	2.266 (0.116)	60	0.958	0.241	120	Propan-1,3-diol[a], Butan-1,4-diol[a] Furan[a]
(b) 80 m*M* SC at pH 7	−1.543 (0.110)	1.022 (0.185)	−0.819 (0.140)	0.144 (0.135)	−1.959 (0.133)	2.245 (0.122)	59	0.956	0.242	112	Propan-1,3-diol[a], Butan-1,4-diol[a] Furan[a], α-Pinene[b]
(c) 80 m*M* SC at pH 6	−1.570 (0.110)	1.018 (0.184)	−0.819 (0.140)	0.141 (0.134)	−1.954 (0.132)	2.246 (0.121)	59	0.956	0.241	113	Propan-1,3-diol[a], Butan-1,4-diol[a] Furan[a], α-Pinene[b]
Phenols with pK_a 9–10 included											
(d) 80 m*M* SC at pH 8	−1.501 (0.101)	0.910 (0.152)	−0.776 (0.116)	0.109 (0.096)	−1.961 (0.126)	2.255 (0.112)	66	0.958	0.234	133	Propan-1,3-diol[a], Butan-1,4-diol[a] Furan[a]
(e) 80 m*M* SC at pH 7	−1.502 (0.101)	0.915 (0.152)	−0.748 (0.118)	0.095 (0.096)	−1.968 (0.127)	2.235 (0.117)	65	0.956	0.234	126	Propan-1,3-diol[a], Butan-1,4-diol[a] Furan[a], α-Pinene[b]
(f) 80 m*M* SC at pH 6	−1.531 (0.100)	0.916 (0.151)	−0.751 (0.117)	0.097 (0.095)	−1.963 (0.126)	2.237 (0.116)	65	0.957	0.232	127	Propan-1,3-diol[a], Butan-1,4-diol[a] Furan[a], α-Pinene[b]

Standard deviations are in parentheses.
[a] Solutes that coelute with methanol.
[b] Solute considered outlier (with a standard residual higher than |2.5|).

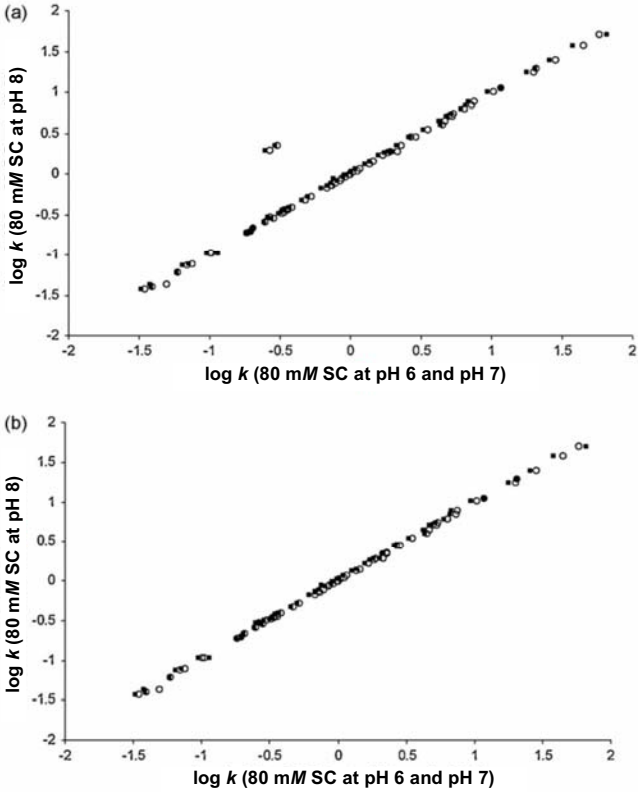


Fig. 2. (a) Log *k* values for the 80 m*M* SC system at pH 6 (■) and pH 7 (○) against the log *k* values at pH 8. (b) Log *k* values for the 80 m*M* SC system at pH 6 (■) and pH 7 (○) against the log *k* values at pH 8, with the values for catechol and 1,2,3-trihydroxybenzene determined at pH 8 in phosphate buffer.

if these compounds are not significantly ionized. The correlations illustrated in Fig. 2a show a good fit for all the compounds except for the phenols catechol and 1,2,3-trihydroxybenzene, which deviate from both correlations. However, the most probably reason for the deviations of these two compounds is the fact that they have vicinal diol groups in their chemical structure and the borate buffer, which was used mixed with phosphate buffer, has a well-known ability to complex vicinal diol groups [26,27]. In order to prove this, catechol and 1,2,3-trihydroxybenzene were injected again in 80 m*M* SC at pH 8 but in this case the aqueous buffer was prepared only with phosphate. Plots of the log *k* values of the 80 m*M* SC system at different pHs were represented again including these new log *k* values acquired for catechol and 1,2,3-trihydroxybenzene at pH 8. Plots are shown in Fig. 2b. The good correlations observed in this figure for all the compounds of the set confirm that the previously noticed deviations of catechol and 1,2,3-trihydroxybenzene were due to the modification of the effective mobility that both compounds experiment when borate buffer interacts with their consecutive diols. The good correlations shown in Fig. 2b also prove that 4-chlorophenol, catechol, resorcinol, hydroquinone, 2-naphthol, and 1,2,3-trihydroxybenzene are not significantly ionized at pH 8, since none deviates from the values acquired at pH 7 and 6 (pH in which they are clearly neutral). In fact, a simple calculation shows that the ionisation degree of these compounds in water at pH 8 is less than 5% (except for the most acidic compound 1,2,3-trihydroxybenzene). This degree is decreased by partition of the neutral forms of the compounds to the pseudostationary phase. Thus, it is possible to include them in any characterization at pH 8. For the 80 m*M* SC at pH 8, system constants were calculated again including these six phenols and the results are shown in Table 2 (correlation d).

3.2. Characterization of SC–SDC mixtures

Separation systems composed by mixtures of SC–SDC at different concentrations were also characterized by means of the solvation parameter model through Eq. (1) by analysis of the log k data of the 69 solutes selected (Table 1). Correlation equations for each system were calculated employing both the Abraham descriptors and the Poole optimized descriptors, whose values are detailed in Table 1, in order to check if the use of these optimized descriptors improves the correlation equations of the SC–SDC systems studied in this work.

The coefficients obtained for each system and the statistics for the fit of the solvation parameter model to the experimental log k data when the Abraham descriptors were employed in the regressions are summarized in Table 3. The model provides a good statistical fit and its interpretation is chemically reasonable. The solutes that were not included in the correlations are detailed in the table (all them coeluted with methanol, except for heptanophenone that coeluted with dodecanophenone in the 40 mM SDC system). The coefficients obtained for the 80 mM SC and 40 mM SDC systems (only ones that are reported in the literature) are similar to the ones obtained for other authors in similar conditions [8,17]. The small differences observed in some constants may come from the different solutes employed in the characterizations.

The constants for each system were also calculated employing the Poole optimized descriptors that are available for some solutes of the characterization set (for the rest of solutes of the set, the Abraham descriptors were used). The results are summarized in Table 4.

As the correlation equation of each system was calculated with the same group of solutes in Tables 3 and 4, the coefficients and statistics presented in these tables can be directly compared. For all the characterized systems the fit of the experimental data to the solvation parameter model was only slightly better when the proposed optimized descriptors were used. Regarding the coefficients, it is observed that for all the systems, the values of e, b, and v are a bit higher when the Poole optimized descriptors are used in the correlation equations instead of the Abraham descriptors. On the contrary, it is observed that the s and a coefficients are a bit lower if the characterizations are performed using the Poole optimized descriptors. Anyway, the coefficients summarized in both tables lead to the same chemical interpretation.

In order to illustrate more clearly that the experimental data better fit the model when the Poole optimized descriptors are used, we also calculated the system constants for all the studied SC–SDC systems just including the solutes with available Poole optimized descriptors. The F statistic ranged from 71 to 83 when the Abraham descriptors were used, whereas it ranged from 144 to 197 with the Poole optimized descriptors. Regarding the standard deviation, it ranged from 0.190 to 0.212 using the Abraham descriptors, whereas it ranged from 0.127 to 0.151 using the Poole optimized descriptors. Taking into account these statistics, a greater improvement in the fit is observed when the Poole optimized descriptors are used instead of the Abraham descriptors. However, we consider that interpreting the system constants according to these regressions is not rigorous enough, since the solutes with available Poole optimized descriptors are not sufficiently representative to define all interactions (40 solutes that do not include large polar hydrogen-bonding solutes such as corticosteroids). Therefore, the chemical interpretation of the characterizations will be discussed according to the regressions calculated with the 69 solutes, i.e. using Abraham descriptors for the 69 solutes in one hand (Table 3) and the available Poole optimized descriptors for 40 solutes and Abraham descriptors for the remaining 29 solutes on the other hand (Table 4).

As shown in Tables 3 and 4, the largest coefficients in absolute value are v and b independently of the composition of SC and

Table 3
Coefficients in Eq. (1) for the SC–SDC mixed-micellar systems at 25 °C and pH 8 calculated employing the Abraham descriptors.

System	x_{SDC}	Coefficients						Statistics				Solutes excluded from the correlations
		c	e	s	a	b	v	n	r	SD	F	
80 mM SC	0.000	−1.501 (0.101)	0.910 (0.152)	−0.776 (0.116)	0.109 (0.096)	−1.961 (0.126)	2.255 (0.112)	66	0.958	0.234	133	Propan-1,3-diol, Butan-1,4-diol Furan
60 mM SC–10 mM SDC	0.143	−1.578 (0.095)	0.897 (0.144)	−0.768 (0.110)	0.125 (0.090)	−1.903 (0.119)	2.241 (0.106)	66	0.961	0.221	145	Propan-1,3-diol, Butan-1,4-diol Furan
40 mM SC–20 mM SDC	0.333	−1.658 (0.091)	0.858 (0.142)	−0.795 (0.108)	0.145 (0.090)	−1.847 (0.119)	2.300 (0.105)	67	0.961	0.221	147	Propan-1,3-diol Butan-1,4-diol
20 mM SC–30 mM SDC	0.600	−1.753 (0.102)	0.865 (0.155)	−0.809 (0.116)	0.117 (0.097)	−1.827 (0.130)	2.354 (0.114)	66	0.955	0.239	125	Propan-2-ol, Propan-1,3-diol Butan-1,4-diol
40 mM SDC	1.000	−1.786 (0.112)	0.813 (0.170)	−0.803 (0.125)	0.061 (0.103)	−1.771 (0.140)	2.400 (0.127)	64	0.947	0.255	102	Propan-1-ol, Propan-2-ol, Propan-1,3-diol Butan-1,4-diol, Heptanophenone

Standard deviations are in parentheses.

Table 4
Coefficients in Eq. (1) for the SC–SDC mixed-micellar systems at 25 °C and pH 8 calculated employing the Poole optimized descriptors (Abraham descriptors were used for solutes with no available optimized descriptors).

System	x_{SDC}	Coefficients						Statistics				Solutes excluded from the correlations
		c	*e*	*s*	*a*	*b*	*v*	*n*	*r*	SD	*F*	
80 m*M* SC	0.000	−1.536 (0.094)	0.956 (0.146)	−0.825 (0.110)	0.029 (0.081)	−1.944 (0.119)	2.303 (0.104)	66	0.964	0.216	158	Propan-1,3-diol, Butan-1,4-diol Furan
60 m*M* SC–10 m*M* SDC	0.143	−1.610 (0.090)	0.936 (0.138)	−0.814 (0.105)	0.045 (0.077)	−1.883 (0.113)	2.287 (0.099)	66	0.966	0.205	169	Propan-1,3-diol, Butan-1,4-diol Furan
40 m*M* SC–20 m*M* SDC	0.333	−1.686 (0.088)	0.892 (0.138)	−0.836 (0.104)	0.059 (0.078)	−1.821 (0.115)	2.340 (0.099)	67	0.965	0.209	166	Propan-1,3-diol, Butan-1,4-diol
20 m*M* SC–30 m*M* SDC	0.600	−1.785 (0.098)	0.909 (0.152)	−0.854 (0.113)	0.033 (0.084)	−1.804 (0.126)	2.396 (0.108)	66	0.960	0.226	141	Propan-2-ol, Propan-1,3-diol Butan-1,4-ciol
40 m*M* SDC	1.000	−1.822 (0.109)	0.865 (0.169)	−0.852 (0.123)	−0.022 (0.091)	−1.750 (0.137)	2.444 (0.122)	64	0.952	0.244	112	Propan-1-ol, Propan-2-ol, Propan-1,3-diol Butan-1,4-ciol, Heptanophenone

Standard deviations are in parentheses.

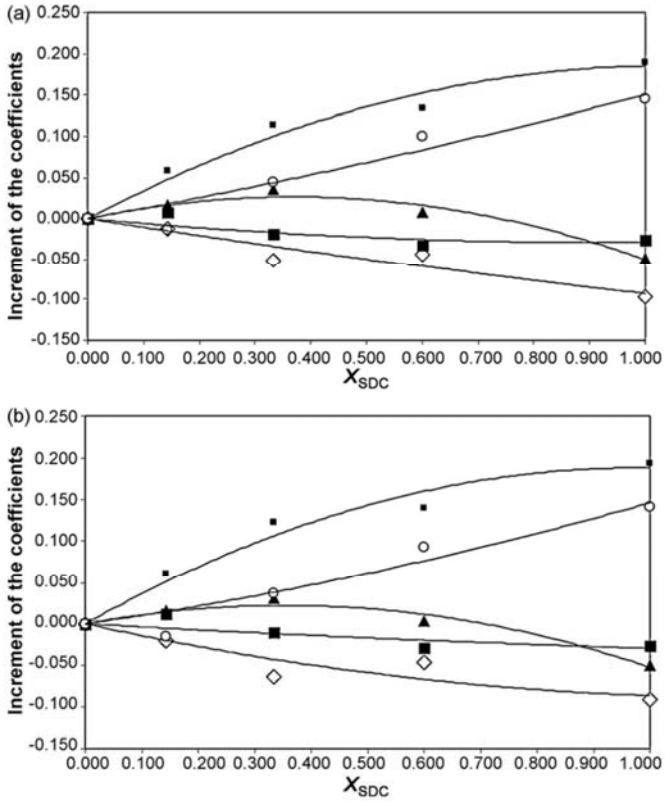


Fig. 3. Plot of the increment of each coefficient depending on the molar fraction of SDC in the mixed-micellar systems when the coefficients in Eq. (1) have been calculated with: (a) the Abraham descriptors, and (b) the Poole optimized descriptors: (◊) Δ*e*; (■) Δ*s*; (▲) Δ*a*; (■) Δ*b*; (○) Δ*v*.

SDC. This means that the McGowan's characteristic volume (*V*) and the hydrogen-bond basicity (*B*) are the solute descriptors that contribute most to retention. On the other hand, the *a* coefficient is nearly zero in all the systems, which indicates that the hydrogen-bond acceptor ability of micelles is similar to that of water and the retention of solutes is then scarcely influenced by its hydrogen-bond acidity (*A* descriptor plays a minor role). Regarding the sign of the coefficients, in all the SC–SDC systems the *e*, *a* and *v* coefficients are positive whereas the *s* and *b* coefficients are negative. According to the positive sign, the partition of the solutes towards the micellar phase is favoured by their excess molar refraction (*E*), their hydrogen-bond acidity (*A*) and their molar volume (*V*). On the contrary, since the *s* and *b* coefficients are negative, the more dipolar and hydrogen-bond basic is a solute (high values of *S* and *B*, respectively), the more favoured is its partition into the aqueous phase and it becomes less retained.

In order to study the variation of the coefficients depending on the composition of SC–SDC in the separation buffer, the increment of each coefficient has been plotted against the composition of SDC in the mixtures, using the coefficients of the 80 m*M* SC system as a zero point. Fig. 3a shows the plot corresponding to the variation of the system constants calculated with the Abraham descriptors whereas Fig. 3b represents the variation of the coefficients that were obtained employing the Poole optimized descriptors. It can be observed that Fig. 3a and b is nearly identical, which means that both collection of descriptors lead to the same conclusions about the variation of the coefficients. The hydrogen-bond acidity is the property that makes the two surfactants more different since the *b* coefficient has the greatest increment, followed by the *v* coefficient, which represents the hydrophobicity of the system. The increment of the rest of coefficients is modest according to the similar nature

of SC and SDC bile salts. From Fig. 3a and b we come to the conclusion that SDC is more hydrogen-bond acidic than SC ($b_{SDC} > b_{SC}$) and it is also a bit more hydrophobic ($v_{SDC} > v_{SC}$). However, SDC is moderately less polarizable than SC since it is observed a smooth decrease of the e coefficient ($e_{SDC} < e_{SC}$). On the other hand, there is hardly any difference in dipolarity and hydrogen-bond basicity between both surfactants, since a meaningful variation of its coefficients is not observed ($s_{SDC} \approx s_{SC}$, $a_{SDC} \approx a_{SC}$). This means that at any concentration of SC and SDC, the micellar pseudophase has the same capacity to take part in dipole–dipole interactions with the solute, as well as the same hydrogen-bond basicity.

4. Conclusions

It has been shown how significant is employing a suitable set of solutes for characterizing systems by means of the solvation parameter model in order to obtain accurate coefficients, as well as it has been proved that the pH of the aqueous phase does not have effect on the studied MEKC system characterizations. Initial comparison of the results obtained here for 80 m*M* SC at pH 8 and literature results for 80 m*M* SC at pH 7 showed significant differences, but we have demonstrated that these differences are caused by differences in the set of solutes used, not by differences in pH. No significant differences were found when the same set of non-ionized solutes were used for the two characterizations.

Both the Abraham descriptor values and the optimized descriptor values recently proposed by Poole et al. have been employed to calculate the correlation equations of all the studied SC–SDC mixtures. It has been proved that both group of descriptors lead to the same variation of coefficients depending on the composition of SC–SDC in the systems. Although the statistics obtained are a bit better when the optimized descriptors are used, there are not optimized descriptor values for all the solutes of the set. For this reason, enlarging the optimized descriptors database would be very interesting.

The characterizations of SC–SDC mixed-micellar phases demonstrate that the addition of SDC to SC micelles increases the hydrogen-bond acidity of the micellar phase and its hydrophobicity whereas the rest of interactions are only slightly changed. Therefore, the more rich in SDC is the micellar phase, the more appropriate the system is to separate mixtures of compounds of different hydrogen-bond basicity and size.

Acknowledgements

We thank the Ministerio de Ciencia e Innovación of the Spanish Government and the Fondo Europeo de Desarrollo Regional of the European Union (project CTQ2007-61608/BQU) for financial support. M.H.-R. also thanks the Ministerio de Educación of the Spanish Government for being supported by a grant (AP2007-01109).

References

[1] S. Terabe, N. Chen, K. Otsuka, Adv. Electrophoresis 7 (1994) 87.
[2] D.R. Baker, Capillary Electrophoresis, Wiley-Interscience, New York, 1995.
[3] C.F. Poole, S.K. Poole, J. Chromatogr. A 792 (1997) 89.
[4] S. Yang, M.G. Khaledi, Anal. Chem. 67 (1995) 499.
[5] S.K. Poole, C.F. Poole, Anal. Commun. 34 (1997) 57.
[6] Z. Liu, H. Zou, M. Ye, J. Ni, Y. Zhang, J. Chromatogr. A 863 (1999) 69.
[7] C. Fu, M.G. Khaledi, J. Chromatogr. A 1216 (2009) 1901.
[8] S.K. Poole, C.F. Poole, Analyst 122 (1997) 267.
[9] C.F. Poole, S.K. Poole, M.H. Abraham, J. Chromatogr. A 798 (1998) 207.
[10] M. Vitha, P.W. Carr, J. Chromatogr. A 1126 (2006) 143.
[11] M.H. Abraham, Chem. Soc. Rev. 22 (1993) 73.
[12] M.H. Abraham, M. Rosés, C.F. Poole, S.K. Poole, J. Phys. Org. Chem. 10 (1997) 358.
[13] E. Fuguet, C. Ràfols, E. Bosch, M.H. Abraham, M. Rosés, Electrophoresis 27 (2006) 1900.
[14] M.H. Abraham, H.S. Chadha, F. Martins, R.C. Mitchell, M.W. Bradbury, J.A. Gratton, Pestic. Sci. 55 (1999) 78.
[15] M.H. Abraham, Y.H. Zhao, J. Le, A. Hersey, C.N. Luscombe, D.P. Reynolds, G. Beck, B. Sherborne, I. Cooper, Eur. J. Med. Chem. 37 (2002) 595.
[16] L.M. Sprunger, J. Gibbs, W.E. Acree Jr., M.H. Abraham, QSAR Comb. Sci. 28 (2009) 72.
[17] E. Fuguet, C. Ràfols, E. Bosch, M.H. Abraham, M. Rosés, J. Chromatogr. A 942 (2002) 237.
[18] M.H. Abraham, A. Ibrahim, A.M. Zissimos, J. Chromatogr. A 1037 (2004) 29.
[19] C.F. Poole, S.N. Atapattu, S.K. Poole, A.K. Bell, Anal. Chim. Acta 652 (2009) 32.
[20] M.D. Trone, M.G. Khaledi, Anal. Chem. 71 (1999) 1270.
[21] M. Rosés, C. Ràfols, E. Bosch, A.M. Martínez, M.H. Abraham, J. Chromatogr. A 845 (1999) 217.
[22] E. Fuguet, C. Ràfols, E. Bosch, M. Rosés, M.H. Abraham, J. Chromatogr. A 907 (2001) 257.
[23] C. Fu, M.G. Khaledi, J. Chromatogr. A 1216 (2009) 1891.
[24] E. Fuguet, C. Ràfols, E. Bosch, M. Rosés, Electrophoresis 23 (2002) 56.
[25] C. Ràfols, A. Poza, E. Fuguet, M. Rosés, E. Bosch, Electrophoresis 23 (2002) 2408.
[26] E. Fuguet, M. Reta, C. Gibert, M. Rosés, E. Bosch, C. Ràfols, Electrophoresis 29 (2008) 2841.
[27] J.P. Landers, R.P. Oda, M.D. Schuchard, Anal. Chem. 64 (1992) 2846.

Anal. Chem. 2010, 82, 10236–10245

Estimation of Biological Properties by Means of Chromatographic Systems: Evaluation of the Factors That Contribute to the Variance of Biological−Chromatographic Correlations

Marta Hidalgo-Rodríguez, Elisabet Fuguet, Clara Ràfols, and Martí Rosés*

Departament de Química Analítica and Institut de Biomedicina de la Universitat de Barcelona, Universitat de Barcelona, Martí i Franquès, 1-11, E-08028 Barcelona, Spain

The performance of chromatographic systems to emulate biological systems is evaluated in terms of the precision that can be achieved. The variance obtained when biological parameters are correlated against physicochemical ones can be decomposed in three terms: the variance of the biological data, the variance of the physicochemical data, and the variance caused by the dissimilarity between the two correlated systems (biological and physicochemical). The three terms contribute to the overall variance observed when measurements in chromatographic systems are correlated with experimental biological properties. The Abraham linear free energy relationships (LFERs) provide a very good approach to characterize biological and physicochemical systems and thus the variance of the analyzed data and the similarity/dissimilarity between them. The contribution of the three variances to the precision of the biological parameter estimated in this way is evaluated from the characterization of the biological and chromatographic systems by means of the Abraham model. The proposed method is able to estimate the goodness of chromatographic systems to predict particular biological properties. In particular, this method is illustrated by comparison of toxicity data (−log LC_{50}) for the fish fathead minnow with retention data (log *k*) in several micellar electrokinetic chromatography (MEKC) systems and also by correlations between retention data (log *k*) in the sodium taurocholate (STC) MEKC system and data of several biological systems.

Experimental determination of biological parameters, including toxicological and environmental ones, usually involves very expensive, time-consuming, and difficult procedures because of the complexity of biological samples. Sometimes, these procedures are even ethically questionable. Because of these experimental complications, predictive models for such processes are highly desirable. In this respect, models of predictive capacity should be general and fast, in order to be suitable for giving an approximate value for the biological property of interest.

Since experimental determination of physicochemical properties is normally cheaper, faster, and easier than it is for biological properties, physicochemical systems are thought to be a good approach to estimate biological parameters. Thus, there is a lot of interest in developing physicochemical systems that emulate biological processes in order to be able to predict biological properties of chemical compounds through the measurement of physicochemical properties. Among the existing physicochemical systems, chromatographic systems are very popular because of their facility of measurement.

In particular, MEKC offers suitable chromatographic systems to mimic biological ones since the solvation and distribution properties of MEKC systems can be easily modified by appropriate selection of the surfactant type, by mixture of different surfactants, or by addition of complexing agents or organic solvents to the separation buffer.[1−5] This means that there is a wide range of MEKC systems that can be compared with biological ones, and in addition, the characteristics of a given MEKC system can be manipulated in a simple way to try to make it more similar to a biological system of interest.

The solvation parameter model, which was developed by Abraham,[6] has successfully described many interesting biological processes (e.g., blood−brain distribution, chemical toxicity, soil−water sorption)[7−9] as well as physicochemical ones (e.g., solvent−water partition, reversed-phase liquid chromatography (RPLC), micellar electrokinetic chromatography (MEKC))[10] ruled by the partition of solutes between two phases. Since this model has been applied to characterize both biological and physicochemical systems, it is possible to compare them by means of their characterizations.

So far, several authors have compared their characterized chromatographic systems with biological ones in order to test their

* To whom correspondence should be addressed. Phone: 34 93 403 92 75. Fax: 34 93 402 12 33. E-mail: marti@apolo.qui.ub.es.

(1) Poole, C. F.; Poole, S. K. *J. Chromatogr., A* **1997**, *792*, 89–104.
(2) Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. *J. Chromatogr., A* **2002**, *942*, 237–248.
(3) Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. *J. Chromatogr., A* **2009**, *1216*, 6877–6879.
(4) Fuguet, E.; Ràfols, C.; Bosch, E.; Rosés, M.; Abraham, M. H. *J. Chromatogr., A* **2001**, *907*, 257–265.
(5) Liu, Z.; Zou, H.; Ye, M.; Ni, J.; Zhang, Y. *J. Chromatogr., A* **1999**, *863*, 69–79.
(6) Abraham, M. H. *Chem. Soc. Rev.* **1993**, *22*, 73–83.
(7) Platts, J. A.; Abraham, M. H.; Zhao, Y. H.; Hersey, A.; Ijaz, L.; Butina, D. *Eur. J. Med. Chem.* **2001**, *36*, 719–730.
(8) Hoover, K. R.; Acree, W. E., Jr.; Abraham, M. H. *Chem. Res. Toxicol.* **2005**, *18*, 1497–1505.
(9) Poole, S. K.; Poole, C. F. *Anal. Commun.* **1996**, *33*, 417–419.
(10) Vitha, M.; Carr, P. W. *J. Chromatogr., A* **2006**, *1126*, 143–194.

10.1021/ac102626u © 2010 American Chemical Society
Published on Web 11/24/2010

suitability to model different biological processes.[11−14] Nevertheless, a study about the criterion for considering if a chromatographic system emulates well enough a biological one had not been carried out yet from this approach, and this is the scope of the present article.

The performance of a chromatographic system to mimic a biological one is assessed according to the overall variance obtained when the biological and the chromatographic data are correlated, since the precision with which the biological property can be estimated by means of the chromatographic measurement depends on this variance. The precision of the biological data, the precision of the chromatographic data, and the dissimilarity between the biological and the chromatographic correlated systems are the three factors that mainly determine the overall variance of its correlation. This work is focused on evaluating the contribution of each of these factors to the overall variance of several correlations between biopartitioning and MEKC systems, previously characterized by means of the solvation parameter model. According to these contributions, the modeling of the toxicity for the fish fathead minnow (−log LC_{50}) by means of various MEKC systems is assessed as well as the goodness of the sodium taurocholate (STC) MEKC system to predict some biological properties of interest.

THEORETICAL BASIS

Many properties of solutes can be related to structural descriptors by means of quantitative structure−activity relationships (QSARs). In particular, linear free energy relationships (LFERs) assume that the free energy change associated with the solute−solvent property of interest is linearly related to solute and solvent appropriate molecular descriptors. Among the variety of models based on these principles, the solvation parameter model developed by Abraham[6] is one of the most widely accepted in order to achieve a better understanding of the types and the relative strength of the chemical interactions that control any solvation process of a neutral compound.[10] Since the solute property correlated to the solvation descriptors can be any property related to free energy, the solvation parameter model has been successfully applied to characterize many biological processes, including some toxicological and environmental ones, as well as a wide range of physicochemical processes ruled in any case by the passive transport of solutes between two phases. Some examples of biological processes described by means of the solvation parameter model are the drug transport across the blood−brain barrier (log PS, log BB),[7,15,16] human intestinal absorption (log{ln[100/(100 − % Abs.)]}),[17] skin permeation and partition (log K_p and log K_{sc}),[18] tadpole narcosis (log $(1/C_{nar})$),[19] chemical toxicity for several aquatic organisms (−log LC_{50}, −log IGC_{50}),[8,20] and soil−water sorption (log K_{OC}).[9] Regarding the physicochemical systems, many solvent−water partition processes,[21−23] EKC separation systems,[24] and reversed-phase liquid chromatography (RPLC) systems[12,25−27] have been also characterized through the solvation parameter model.

The solvation parameter model is based on the following equation:[6]

$$\log \mathrm{SP} = c + eE + sS + aA + bB + vV \qquad (1)$$

where SP is the dependent solute property in a given system and can be described by an equilibrium constant or some other free energy related property. The E, S, A, B, and V independent variables are the solute descriptors proposed by Abraham. E represents the excess molar refraction, S is the solute dipolarity/polarizability, A and B are the solute's effective hydrogen-bond acidity and hydrogen-bond basicity, respectively, and V is McGowan's solute volume. Equation 1 does not include any ionic interaction, and thus it is not applicable to charged compounds, only to neutral ones. The coefficients of the equation are characteristic of the system and reflect its complementary properties to the corresponding solute property. As they represent the difference in solvation properties between the two phases that compose the system, e refers to the difference in capacity of each phase to interact with solute π- and n-electrons; s is a measure of the difference of the two phases in capacity to take part in dipole−dipole and dipole−induced dipole interactions; a and b represent the differences in hydrogen-bond basicity and acidity, respectively, between both phases; v is a measure of the relative ease of forming a cavity for the solute in the two phases. For any system, the coefficients of the correlation equation can be obtained by multiple linear regression analysis between the log SP values acquired for an appropriate group of solutes and their descriptor values.

The solvation parameter model not only provides information about the magnitude of the different interactions between the phases and neutral solutes for a characterized system but also allows one to compare how different various characterized systems are regarding these interactions. In fact, several approaches can be used to evaluate the similarity among systems by studying the relationship between their LFER coefficients. For instance, principal component analysis (PCA) is a chemometric tool applied by several authors to obtain a visual distribution of systems

(11) Rosés, M.; Ràfols, C.; Bosch, E.; Martinez, A. M.; Abraham, M. H. *J. Chromatogr., A* **1999**, *845*, 217–226.
(12) Lázaro, E.; Ràfols, C.; Abraham, M. H.; Rosés, M. *J. Med. Chem.* **2006**, *49*, 4861–4870.
(13) Liu, J.; Sun, J.; Wang, Y.; Liu, X.; Sun, Y.; Xu, H.; He, Z. *J. Chromatogr., A* **2007**, *1164*, 129–138.
(14) Lu, R.; Sun, J.; Wang, Y.; Li, H.; Liu, J.; Fang, L.; He, Z. *J. Chromatogr., A* **2009**, *1216*, 5190–5198.
(15) Gratton, J. A.; Abraham, M. H.; Bradbury, M. W.; Chadha, H. S. *J. Pharm. Pharmacol.* **1997**, *49*, 1211–1216.
(16) Abraham, M. H. *Eur. J. Med. Chem.* **2004**, *39*, 235–240.
(17) Abraham, M. H.; Zhao, Y. H.; Le, J.; Hersey, A.; Luscombe, C. N.; Reynolds, D. P.; Beck, G.; Sherborne, B.; Cooper, I. *Eur. J. Med. Chem.* **2002**, *37*, 595–605.
(18) Abraham, M. H.; Martins, F. *J. Pharm. Sci.* **2004**, *93*, 1508–1523.
(19) Abraham, M. H.; Ràfols, C. *J. Chem. Soc., Perkin Trans. 2* **1995**, 1843–1851.
(20) Hoover, K. R.; Flanagan, K. B.; Acree, W. E., Jr.; Abraham, M. H. *J. Environ. Eng. Sci.* **2007**, *6*, 165–174.
(21) Abraham, M. H.; Chadha, H. S.; Whiting, G. S.; Mitchell, R. C. *J. Pharm. Sci.* **1994**, *83*, 1085–1100.
(22) Abraham, M. H.; Chadha, H. S.; Dixon, J. P.; Leo, A. J. *J. Phys. Org. Chem.* **1994**, *7*, 712–716.
(23) Abraham, M. H.; Zissimos, A. M.; Acree, W. E., Jr. *New J. Chem.* **2003**, *27*, 1041–1044.
(24) Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. *Electrophoresis* **2006**, *27*, 1900–1914.
(25) Abraham, M. H.; Rosés, M.; Poole, C. F.; Poole, S. K. *J. Phys. Org. Chem.* **1997**, *10*, 358–368.
(26) Bolliet, D.; Poole, C. F.; Rosés, M. *Anal. Chim. Acta* **1998**, *368*, 129–140.
(27) Lepont, C.; Poole, C. F. *J. Chromatogr., A* **2002**, *946*, 107–124.

according to the similarity of their coefficients.[17,20,28] The great advantage that PCA offers is the joint manipulation of large sets of data, which allows the comparison of many systems by means of very few visual distributions. However, these visual distributions become different just by including or removing some systems from the initial group. This is because the set of coefficient values to be represented changes and the weight of each coefficient in the principal components also does, which lead to the distortion of the visual distribution. For this reason, many authors prefer other approaches that compare the similarities among systems in a more mathematical way. For two systems, a common method consists of comparing the ratios of their coefficients such as e/v, s/v, a/v, b/v.[11,29] Another approach, developed by Ishihama and Asakawa, is based on considering the LFER coefficients as components of a vector in five-dimensional space and taking the angle (θ) between two LFER vectors as a measure of their mathematical similarity.[30] Later, Abraham and Martins considered the LFER coefficients of any system as a point in five-dimensional space and suggested the distance (D') between two points as the measure scale to know the chemical similarity of the pair of systems.[18] In a previous work,[12] we proposed another mathematical procedure for comparing systems previously characterized by means of the solvation parameter model. It is based on a distance parameter, d, which measures the similarity between two systems by means of their normalized unitary vectors, also considering them in a five-dimensional space. The normalized vector of a system ($\vec{w}_u$) has the following components (normalized coefficients):

$$e_u = \frac{e}{l} \quad (2)$$

$$s_u = \frac{s}{l} \quad (3)$$

$$a_u = \frac{a}{l} \quad (4)$$

$$b_u = \frac{b}{l} \quad (5)$$

$$v_u = \frac{v}{l} \quad (6)$$

where l is the length of the coefficients vector calculated as follows

$$l = \sqrt{e^2 + s^2 + a^2 + b^2 + v^2} \quad (7)$$

Considering two systems, e.g., one biological (i) and one chromatographic (j), their LFERs can be written as

$$\log SP_i = c_i + l_i(e_{ui}E + s_{ui}S + a_{ui}A + b_{ui}B + v_{ui}V) \quad (8)$$

$$\log SP_j = c_j + l_j(e_{uj}E + s_{uj}S + a_{uj}A + b_{uj}B + v_{uj}V) \quad (9)$$

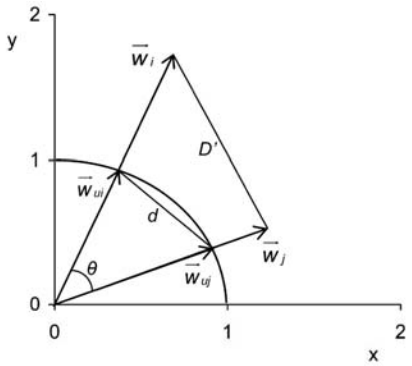


Figure 1. Plot in two dimensions of two LFER vectors ($\vec{w}_i$,$\vec{w}_j$) and their corresponding normalized coefficients vectors ($\vec{w}_{ui}$,$\vec{w}_{uj}$) including the d, D', and θ parameters between the two vectors.

and the d distance between their $\vec{w}_{ui}$ and $\vec{w}_{uj}$ normalized vectors provides a measure of the mathematical similarity between the two systems, as it is represented in Figure 1. The d distance can be calculated according to the following equation, in which all the coefficients have been previously normalized:

$$d = \sqrt{(e_{ui} - e_{uj})^2 + (s_{ui} - s_{uj})^2 + (a_{ui} - a_{uj})^2 + (b_{ui} - b_{uj})^2 + (v_{ui} - v_{uj})^2} \quad (10)$$

The smaller d is, the more mathematically similar are the two compared systems. As this parameter was already demonstrated[12] to be a very easy, intuitive, and useful tool to compare systems characterized by means of the solvation parameter model, we decided to employ it in this work for estimating the similarity between biological and chromatographic systems.

EXPERIMENTAL SECTION

All the experimental retention factors in the MEKC systems considered in this work were determined in our research group. Most of them correspond to previous works, and experimental details are specified in the corresponding references. The only retention factors included in this paper that do not belong to any publication are those of the STC MEKC system. For this physicochemical system, the characterization published by Poole et al.[31] was taken into account although the experimental retention factors required for the correlations against biological properties were determined in our laboratory.

Apparatus. All separations were performed with a Beckman P/ACE System 5500 Capillary Electrophoresis with a UV diode array detector. The fused-silica separation capillary was 40 cm effective length × 50 μm i.d. and was obtained from Composite Metal Services Ltd. (Shipley, West Yorkshire, U.K.). The capillary was activated by the following washing sequence: water (10 min), 1 M NaOH (10 min), water (5 min), 0.1 M NaOH (10 min), and separation solution (20 min). As daily conditioning, the capillary was flushed with water for 5 min, followed by 0.1 M NaOH for 5 min, water for 5 min, and separation solution for 20 min. Prior to each separation, the capillary was flushed with separation solution for 3 min. Retention measurements were made at 25 °C and +15 kV. Detection was at 214 nm.

(28) Sprunger, L. M.; Gibbs, J.; Acree, W. E., Jr.; Abraham, M. H. *QSAR Comb. Sci.* **2009**, *28*, 72–88.
(29) Abraham, M. H.; Treiner, C.; Rosés, M.; Ràfols, C.; Ishihama, Y. *J. Chromatogr., A* **1996**, *752*, 243–249.
(30) Ishihama, Y.; Asakawa, N. *J. Pharm. Sci.* **1999**, *88*, 1305–1312.
(31) Poole, S. K.; Poole, C. F. *Analyst* **1997**, *122*, 267–274.

Chemicals. Hydrochloric acid (25% in water), sodium hydroxide (>99%), sodium dihydrogenphosphate monohydrate (>99%), and methanol (HPLC grade) were from Merck (Darmstadt, Germany). Disodium tetraborate decahydrate (>99.5%) was from Sigma (Steinheim, Germany). Taurocholic acid sodium salt hydrate (STC) (98%) was from Acros Organics (Geel, Belgium). Dodecanophenone (98%) was from Aldrich (Steinheim, Germany). Water was purified by a Milli-Q plus system from Millipore (Bedford, MA), with a resistivity of 18.2 MΩ cm. The test solutes employed were reagent grade or better and obtained from several manufacturers (Merck (Darmstadt, Germany), Sigma (Steinheim, Germany), Fluka (Steinheim, Germany), Aldrich (Steinheim, Germany), Carlo Erba (Milano, Italy), and Baker (Deventer, The Netherlands)).

Procedure. The separation solution was 50 mM in STC and 10 mM in aqueous buffer, pH 8. It was prepared dissolving the surfactant in 10 mM NaH_2PO_4−10 mM $Na_2B_4O_7$ (50:50) and neutralizing with HCl.

Solutes were dissolved in methanol (used as electroosmotic flow marker) and contained 2 mg mL^{-1} of dodecanophenone as the micellar marker.[32] The concentration of the solutes was 2 mg mL^{-1}. All solutions were filtered through 0.45 μm nylon syringe filters obtained from Albet (Dassel, Germany). Samples were introduced into the capillary by applying a pressure of 0.5 psi for 1 s (1 psi = 6894.76 Pa). All measurements were taken in triplicate.

The MEKC retention factor, k, was calculated according to eq 11 with the migration time of methanol used to determine the electroosmotic flow (t_0) and the migration time of dodecanophenone used to determine the migration time of the micelles (t_m). t_R is the solute migration time:

$$k = \frac{(t_R - t_0)}{\left(1 - \frac{t_R}{t_m}\right)t_0} \tag{11}$$

Microsoft Excel XP was used to perform data calculations.

RESULTS AND DISCUSSION

In order to know if a chromatographic system emulates well a biological one, the corresponding biological and chromatographic data can be correlated, obtaining an equation of the following type:

$$\log SP_{bio} = q + p \log SP_{chrom} \tag{12}$$

where SP_{bio} and SP_{chrom} are the correlated biological and chromatographic properties. In this work, the correlated chromatographic property is k, the retention factor in MEKC.

The overall variance of the biological−chromatographic systems' correlation is a measure of the precision of the biological property estimation. Thus, the evaluation of the overall variance is the key issue to know the performance of the chromatographic system to mimic the biological one.

(32) Fuguet, E.; Ràfols, C.; Bosch, E.; Rosés, M. *Electrophoresis* **2002**, *23*, 56–66.
(33) http://www.epa.gov/ncct/dsstox/sdf_epafhm.html.

The main sources of the variance of any biological−chromatographic systems' correlation come from three different factors: the precision of the correlated biological data, the precision of the correlated chromatographic data, and the error coming from the dissimilarity between the two correlated systems. The precision of the correlation can be expressed in terms of variance by the following expression:

$$\sigma_{corr}^2 = \sigma_{bio}^2 + \sigma_{chrom}^2 + \sigma_d^2 \tag{13}$$

where σ_{corr}^2 is the overall variance that results from the correlation between the data of the biological and the chromatographic systems, and σ_{bio}^2, σ_{chrom}^2, and σ_d^2 are the contributions to the overall variance of the factors mentioned above: the precision of the biological data, the precision of the chromatographic data, and the error coming from the dissimilarity between the two correlated systems. In order to assess the goodness of any biological−chromatographic systems' correlation, the contribution of each factor to the overall variance must be carefully examined.

Contribution of the Precision of the Biological Data. The precision of the original biological and chromatographic data has a meaningful effect on the standard deviation of the correlation. The contribution of the original biological data is especially important because the complexity of biological processes makes that their experimental values are usually determined with a significant uncertainty. In addition, the great diversity of biological mechanisms causes that some biological properties can be measured more precisely than others, and as a consequence, more or less error is transferred to the correlation depending on the considered biological system. Therefore, the contribution of this source of error can be quite different according to the biological system under consideration, but in any case it must be taken into account to discuss the overall variance of the correlation. It is evident that in principle we cannot expect a precision in the biological−chromatographic correlation better than the one of the original biological data, although it is often obtained when the correlated data belongs to a series of compounds of the same family.

For biological systems well characterized by means of the solvation parameter model, i.e., assuming that the model error can be negligible, the standard deviation of their characterization (SD_{bio}) can be taken as an estimation of the precision of the original biological data. Thus, the value of σ_{bio}^2 (eq 13) can be directly estimated through the statistics (SD_{bio}^2) of the biological system characterization:

$$\sigma_{bio}^2 \approx SD_{bio}^2 \tag{14}$$

Contribution of the Precision of the Chromatographic Data. Although the experimental chromatographic data is normally measured more easily, precisely, and accurately than the experimental biological data, its precision also contributes to the standard deviation obtained in the biological−chromatographic correlation and, therefore, it is another source of random error. Unlike what happens with biological systems, the precision of the chromatographic data does not change too much depending on the considered system because the mechanisms that rule chromatographic processes are not so varied as the biological ones.

Considering chromatographic systems characterized through the solvation parameter model, the precision of the original chromatographic data can be estimated by means of the standard deviation of their characterization (SD_{chrom}), as it happens with the biological data, assuming again that the model error is insignificant. Nevertheless, the contribution of this source of error to the overall precision of the biological−chromatographic correlation is in this case affected by its slope (represented as p in eq 12): the larger is the slope of the correlation, the larger is the contribution of the precision of the original chromatographic data to the overall standard deviation. Consequently, σ_{chrom}^2 (eq 13) can be estimated as follows:

$$\sigma_{chrom}^2 \approx (p\ SD_{chrom})^2 \quad (15)$$

Contribution of the Dissimilarity between the Correlated Systems. When biological data is correlated against chromatographic data, the dissimilarity between the considered biological and chromatographic systems is a factor that greatly determines its correlation degree. Thus, the contribution of this factor (expressed as σ_d^2 in eq 13) to the overall variance must be taken into account to interpret the biological−chromatographic correlation.

As it has been mentioned before, the d distance (eq 10) is a good parameter to measure the mathematical similarity between two systems. According to this parameter, the smaller is d between a biological and a chromatographic system, the more similar are these two systems, and the correlation of their data is expected to be less influenced by σ_d^2. In fact, two systems, biological and chromatographic ones, with $d = 0$ would have the same unitary vector and thus they would be, by definition, linearly related ($\sigma_d^2 = 0$).

σ_d^2 can be estimated by means of a correlation between calculated values for both biological and chromatographic systems. These calculated values (for log SP_{bio} and log k) are obtained just multiplying the LFER equation of the corresponding system by the descriptors of each considered solute. When values calculated in this way for both biological and chromatographic systems are correlated, the overall variance is not affected by the precision of the correlated data, since the values calculated through the solvation parameter model make 0 of these contributions:

$$SD_{corr\ cal}^2 = \underbrace{SD_{bio\ cal}^2}_{0} + \underbrace{(p\ SD_{chrom\ cal})^2}_{0} + SD_d^2 \quad (16)$$

Then, as it is observed in eq 16, the overall variance of the correlation between calculated values can be directly attributed to the dissimilarity between the two correlated systems. Therefore, the contribution of the dissimilarity between a pair of biological−chromatographic systems (σ_d^2) can be estimated by means of the variance obtained when calculated values are correlated ($SD_{corr\ cal}^2$), such as the following expression shows

$$\sigma_d^2 \approx SD_{corr\ cal}^2 \quad (17)$$

Examples for Testing the Contributions. Biological and chromatographic systems chosen to carry out this study are presented in Table 1. The characterization by means of the solvation parameter model is detailed for each system together with the statistics, as well as their corresponding normalized coefficients calculated by eqs 2−7. The selected systems presented in Table 1 were compiled from a wide range of biological and MEKC systems characterized in the literature through the solvation parameter model followed by the calculation of the d distance (eq 10) between each pair of biological−MEKC systems.

In order to demonstrate that the procedure mentioned above properly estimates the contributions of the three factors to the overall variance of a biological−chromatographic correlation, first of all, only two factors were considered (σ_{chrom}^2 and σ_d^2) by means of correlating calculated biological data with experimental chromatographic data. Using calculated biological data implies that the precision of the experimental biological data does not affect the overall variance of the correlation. Thus, when calculated biological data is correlated against experimental chromatographic data, the obtained variance is only contribution of the other two factors, which are the precision of the experimental chromatographic data and the dissimilarity between the two correlated systems:

$$SD_{corr\ cal}^2 = \underbrace{SD_{bio\ cal}^2}_{0} + (p\ SD_{chrom})^2 + SD_d^2 \quad (18)$$

Experimental retention factors of the MEKC system composed of SDS[2,3] were correlated against calculated values of biological properties of very different nature. For each biological system, calculated values were obtained multiplying its corresponding system constants (Table 1) by the descriptors (Table 2) of each solute with known retention factor in the SDS system. As it can be observed in Table 3, pairs of biological-SDS systems with different similarity (see the range of d values) were selected to examine the contribution of the precision of the chromatographic data and the contribution of the dissimilarity between the two systems. If these contributions are well estimated according to the method detailed in the previous sections, their sum ($SD_{corr\ cal}^2$) must match with the experimental variance obtained in the correlation between experimental retention factors and calculated biological properties ($SD_{corr\ exp}^2$). Table 3 shows the estimation of the error attributed to the precision of the correlated chromatographic data ($(p\ SD_{MEKC})^2$), and the estimation of the error attributed to the dissimilarity between the systems (SD_d^2), for each pair of biological-SDS systems. When the sum of these two contributions ($SD_{corr\ cal}^2$) is compared with the variance obtained in the correlation ($SD_{corr\ exp}^2$), values nearly identical are observed for all the cases, independent of the considered biological system and its similarity with regard to the MEKC system of SDS. As a consequence, it is demonstrated that these two contributions are well estimated according to the proposed method. For the correlations that lead to $SD_{corr\ exp}^2$, all solutes with known retention factor in the SDS system characterization were considered ($n_{corr\ exp}$). The same 63 solutes were considered to calculate the estimations that lead to $SD_{corr\ cal}^2$.

Another MEKC system was also chosen to repeat the procedure just described, in which only two contributions are taken into account. It was the MEKC system composed of LPFOS as surfactant,[2,3] whose properties are quite different to those of the

Table 1. System Constants and Normalized Coefficients for All the Studied Biological and MEKC Systems

		coefficients							statistics				normalized coefficients			
	SP property	c	e	s	a	b	v	n	r	SD	F	e_u	s_u	a_u	b_u	v_u
Biological Systems																
human intestinal absorption[17]	log[ln[100/(100 − % Abs.)]]	0.544	−0.025	0.141	−0.409	−0.514	0.204	127	0.894	0.290	94	−0.036	0.201	−0.582	−0.732	0.290
human skin partition[18]	log K_{sc}	0.341	0.341	−0.206	−0.024	−2.178	1.850	45	0.962	0.216	97	0.118	−0.071	−0.008	−0.755	0.641
tadpole narcosis[19]	log $(1/C_{nar})$	0.582	0.770	−0.696	0.243	−2.592	3.343	114	0.954	0.337	217	0.177	−0.160	0.056	−0.594	0.766
chemical toxicity for the fathead minnow fish (*Pimephales promelas*)[8]	−log LC_{50}	0.996	0.418	−0.182	0.417	−3.574	3.377	196	0.976	0.276	779.5	0.084	−0.037	0.084	−0.721	0.681
chemical toxicity for the guppy fish (*Poecilia reticulata*)[8]	−log LC_{50}	0.811	0.782	−0.230	0.341	−3.050	3.250	148	0.973	0.280	493.1	0.172	−0.051	0.075	−0.671	0.715
chemical toxicity for the bluegill fish (*Lepomis macrochirus*)[8]	−log LC_{50}	0.903	0.583	−0.127	1.238	−3.918	3.306	66	0.984	0.272	360	0.110	−0.024	0.233	−0.738	0.623
chemical toxicity for the goldfish fish (*Carassius auratus*)[8]	−log LC_{50}	0.922	−0.653	1.872	−0.329	−4.516	3.078	51	0.983	0.277	254	−0.112	0.321	−0.057	−0.776	0.529
chemical toxicity for the *Tetrahymena Pyriformis* protozoa[20]	−log IGC_{50}	0.616	0.413	−0.048	0.348	−2.707	2.944	192	0.982	0.205	1002	0.102	−0.012	0.086	−0.671	0.729
soil−water sorption[9]	log K_{OC}	−0.070	0.750	−0.460	0.000	−1.950	2.590	49	0.982	0.170	292	0.223	−0.137	0.000	−0.580	0.771
MEKC Systems																
40 mM SDS, 20 mM phosphate, pH 7[2,3]	log k	−1.680	0.558	−0.596	−0.266	−1.674	2.717	63	0.991	0.120	618	0.169	−0.180	−0.080	−0.507	0.822
40 mM LPFOS, 20 mM phosphate, pH 7[2,3]	log k	−1.410	−0.113	−0.243	−0.876	−0.455	1.966	62	0.970	0.190	180	−0.051	−0.110	−0.395	−0.205	0.887
20 mM TTAB, 20 mM phosphate, pH 7[2,3]	log k	−1.851	0.902	−0.617	0.766	−2.410	2.634	53	0.994	0.089	746	0.237	−0.162	0.201	−0.632	0.691
50 mM SDS−15 mM Brij 35, 20 mM phosphate, pH 7[11]	log k	−1.285	0.667	−0.515	0.335	−2.723	2.767	27	0.985	0.109	140	0.167	−0.129	0.084	−0.683	0.694
50 mM STC, 20 mM phosphate-tetraborate, pH 8[31]	log k	−2.100	0.600	−0.340	0.000	−2.060	2.430	40	0.989	0.090	377	0.184	−0.104	0.000	−0.632	0.746

Table 2. Solute Descriptors Employed in the Present Study

solute	E	S	A	B	V
propan-1-ol	0.236	0.42	0.37	0.48	0.5900
propan-2-ol	0.212	0.36	0.33	0.56	0.5900
butan-1-ol	0.224	0.42	0.37	0.48	0.7309
pentan-1-ol	0.219	0.42	0.37	0.48	0.8718
pentan-3-ol	0.218	0.36	0.33	0.56	0.8718
propan-1,3-diol	0.397	0.91	0.77	0.85	0.6487
butan-1,4-diol	0.395	0.93	0.72	0.90	0.7860
pentan-1,5-diol	0.388	0.95	0.72	0.91	0.9305
thiourea	0.840	0.82	0.77	0.87	0.5696
benzene	0.610	0.52	0.00	0.14	0.7164
toluene	0.601	0.52	0.00	0.14	0.8573
ethylbenzene	0.613	0.51	0.00	0.15	0.9982
propylbenzene	0.604	0.50	0.00	0.15	1.1391
butylbenzene	0.600	0.51	0.00	0.15	1.2800
p-xylene	0.613	0.52	0.00	0.16	0.9982
naphthalene	1.340	0.92	0.00	0.20	1.0854
chlorobenzene	0.718	0.65	0.00	0.07	0.8388
bromobenzene	0.882	0.73	0.00	0.09	0.8914
anisole	0.708	0.75	0.00	0.29	0.9160
benzaldehyde	0.820	1.00	0.00	0.39	0.8730
acetophenone	0.818	1.01	0.00	0.48	1.0139
propiophenone	0.804	0.95	0.00	0.51	1.1548
butyrophenone	0.797	0.95	0.00	0.51	1.2957
valerophenone	0.795	0.95	0.00	0.50	1.4366
heptanophenone	0.720	0.95	0.00	0.50	1.7184
benzophenone	1.447	1.50	0.00	0.50	1.4808
methyl benzoate	0.733	0.85	0.00	0.46	1.0726
benzyl benzoate	1.264	1.42	0.00	0.51	1.6804
benzonitrile	0.742	1.11	0.00	0.33	0.8711
aniline	0.955	0.96	0.26	0.50	0.8162
o-toluidine	0.970	0.90	0.23	0.59	0.9751
3-chloroaniline	1.050	1.10	0.30	0.36	0.9390
4-chloroaniline	1.060	1.10	0.30	0.35	0.9390
2-nitroaniline	1.180	1.37	0.30	0.36	0.9904
3-nitroaniline	1.200	1.71	0.40	0.35	0.9904
4-nitroaniline	1.220	1.91	0.42	0.38	0.9904
nitrobenzene	0.871	1.11	0.00	0.28	0.8906
2-nitroanisole	0.965	1.34	0.00	0.38	1.0902
benzamide	0.990	1.50	0.49	0.67	0.9728
4-aminobenzamide	1.340	1.94	0.80	0.94	1.0726
acetanilide	0.870	1.36	0.46	0.69	1.1137
4-chloroacetanilide	0.980	1.50	0.64	0.51	1.2357
phenol	0.805	0.89	0.60	0.30	0.7751
3-methylphenol	0.822	0.88	0.57	0.34	0.9160
2,3-dimethylphenol	0.850	0.90	0.52	0.36	1.0569
2,4-dimethylphenol	0.840	0.80	0.53	0.39	1.0569
thymol	0.822	0.79	0.52	0.44	1.3387
furan	0.369	0.53	0.00	0.13	0.5363
2,3-benzofuran	0.888	0.83	0.00	0.15	0.9053
quinoline	1.268	0.97	0.00	0.51	1.0443
pyrrole	0.613	0.73	0.41	0.29	0.5774
pyrimidine	0.606	1.00	0.00	0.65	0.6342
antipyrine	1.320	1.50	0.00	1.48	1.5502
caffeine	1.500	1.60	0.00	1.33	1.3632
corticosterone	1.860	3.43	0.40	1.63	2.7389
cortisone	1.960	3.50	0.36	1.87	2.7546
hydrocortisone	2.030	3.49	0.71	1.90	2.7975
estradiol	1.800	3.30	0.88	0.95	2.1988
estriol	2.000	3.36	1.40	1.22	2.2575
monuron	1.140	1.50	0.47	0.78	1.4768
myrcene	0.483	0.29	0.00	0.21	1.3886
α-pinene	0.446	0.14	0.00	0.12	1.2574
geraniol	0.513	0.63	0.39	0.66	1.4903
4-chlorophenol	0.915	1.08	0.67	0.20	0.8975
catechol	0.970	1.10	0.88	0.47	0.8338
resorcinol	0.980	1.00	1.10	0.58	0.8338
hydroquinone	1.000	1.00	1.16	0.60	0.8338
2-naphthol	1.520	1.08	0.61	0.40	1.1441
1,2,3-trihydroxybenzene	1.165	1.35	1.35	0.62	0.8925

SDS system. Table 4 shows the biological systems whose calculated values were correlated against the experimental retention factors of the LPFOS system. As it is observed, all the pairs of biological-LPFOS systems present a high d value because the LPFOS system has very different properties with regard to any

Table 3. Evaluation of the Correlations between Experimental Retention Factors (log *k*) in the 40 mM SDS System and Calculated Properties for the Detailed Biological Systems

	MEKC system: 40 mM SDS, pH 7					
biological systems	d	$(p\ SD_{MEKC})^2$	SD_d^2	$SD_{corr\ cal}^2$	$SD_{corr\ exp}^2$	$n_{corr\ exp}$
soil−water sorption	0.139	0.013	0.022	0.035	0.034	63
tadpole narcosis	0.173	0.018	0.048	0.067	0.066	63
chemical toxicity for the guppy fish (*Poecilia reticulata*)	0.282	0.019	0.126	0.145	0.144	63
chemical toxicity for the *Tetrahymena Pyriformis* protozoa	0.310	0.015	0.108	0.124	0.122	63
human skin partition	0.338	0.006	0.097	0.103	0.103	63
chemical toxicity for the fathead minnow fish (*Pimephales promelas*)	0.347	0.018	0.177	0.196	0.194	63
chemical toxicity for the bluegill fish (*Lepomis macrochirus*)	0.469	0.015	0.427	0.442	0.440	63
chemical toxicity for the goldfish fish (*Carassius auratus*)	0.700	0.026	0.593	0.619	0.615	63
human intestinal absorption	0.879	0.000	0.029	0.029	0.029	63

Table 4. Evaluation of the Correlations between Experimental Retention Factors (log *k*) in the 40 mM LPFOS System and Calculated Properties for the Detailed Biological Systems

	MEKC system: 40 mM LPFOS, pH 7					
biological systems	d	$(p\ SD_{MEKC})^2$	SD_d^2	$SD_{corr\ cal}^2$	$SD_{corr\ exp}^2$	$n_{corr\ exp}$
soil−water sorption	0.622	0.028	0.210	0.237	0.237	62
tadpole narcosis	0.651	0.037	0.374	0.411	0.411	62
chemical toxicity for the *Tetrahymena Pyriformis* protozoa	0.712	0.030	0.418	0.448	0.447	62
chemical toxicity for the guppy fish (*Poecilia reticulata*)	0.722	0.036	0.528	0.564	0.563	62
human skin partition	0.736	0.011	0.227	0.237	0.237	62
chemical toxicity for the fathead minnow fish (*Pimephales promelas*)	0.750	0.032	0.616	0.648	0.647	62
chemical toxicity for the goldfish fish (*Carassius auratus*)	0.871	0.055	1.147	1.203	1.200	62
human intestinal absorption	0.874	0.001	0.027	0.028	0.028	62
chemical toxicity for the bluegill fish (*Lepomis macrochirus*)	0.884	0.018	0.952	0.970	0.970	62

Table 5. Evaluation of Correlations between Experimental Aquatic Toxicity of Organic Compounds (−log LC_{50}) to the Fathead Minnow Fish and Experimental Retention Factors (log *k*) for the Detailed MEKC Systems

	biological system: chemical toxicity for the fathead minnow fish *(Pimephales promelas)*											
MEKC systems	d	SD_{bio}^2	n_{bio}	$(p\ SD_{MEKC})^2$	n_{MEKC}	F_{MEKC}	SD_d^2	n_d	F_d	$SD_{corr\ cal}^2$	$SD_{corr\ exp}^2$	$n_{corr\ exp}$
50 mM SDS−15 mM Brij 35, pH 7	0.131	0.076	196	0.009	27	0.12	0.020	69	0.26	0.105	0.100	19
20 mM TTAB, pH 7	0.246	0.076	196	0.008	53	0.11	0.066	69	0.87	0.150	0.078	23
40 mM SDS, pH 7	0.347	0.076	196	0.017	63	0.22	0.178	69	2.34	0.272	0.208	27
40 mM LPFOS, pH 7	0.750	0.076	196	0.019	62	0.25	0.671	69	8.83	0.767	0.525	26

biological system. The $(p\ SD_{MEKC})^2$ and SD_d^2 estimated values were calculated for each pair of biological-LPFOS systems and are detailed in Table 4. Their sum ($SD_{corr\ cal}^2$) is also presented in the table as well as the variance directly obtained from the correlation between biological calculated data and LPFOS experimental data ($SD_{corr\ exp}^2$). When $SD_{corr\ cal}^2$ and $SD_{corr\ exp}^2$ values are compared for each pair of systems, they match again perfectly. For the examples shown in the table, 62 solutes ($n_{corr\ exp}$) were considered in all correlations since this is the number of solutes with known retention factor in the LPFOS system and the same solutes were considered for the calculated values of the biological property. The next stage of this work is based on the correlation of both experimental biological and physicochemical data in order to evaluate the three factors that contribute to any biological−chromatographic systems' correlation.

Experimental toxicity data (−log LC_{50}) for the fish fathead minnow[8] was correlated against experimental retention factors in several MEKC systems. Since aquatic organisms play an essential role in the food chain, the evaluation of the acute and chronic toxicity of chemicals in them is useful to know the environmental damage caused by these chemicals. Particularly, the fathead minnow is one of the species outlined by EPA guidelines[33] as a biological model in aquatic toxicology studies. Table 5 shows the MEKC systems that were chosen to be correlated with the fathead minnow system and, as it can be observed, they cover different d distances. According to the proposed method, SD_{bio}^2, $(p\ SD_{MEKC})^2$, and SD_d^2 were calculated for each pair of systems. The corresponding values are detailed in the table together with the number of solutes considered in each case. n_{bio} and n_{MEKC} correspond to the number of solutes of the biological and MEKC systems characterization, respectively. With regard to n_d, the same 69 solutes were used to calculate SD_d^2 for all pairs of systems (eq 16). These 69 solutes are those selected in a previous work[2,3] as an appropriate group of compounds that cover a wide range of solute descriptor values. The sum of the estimated contributions ($SD_{corr\ cal}^2$) as well as the variance directly obtained from the correlation of both experimental biological and MEKC data ($SD_{corr\ exp}^2$) are shown in Table 5. For each pair of fathead minnow-MEKC systems, $SD_{corr\ cal}^2$ gives a fair enough approximation to $SD_{corr\ exp}^2$. Obviously, $SD_{corr\ cal}^2$ and $SD_{corr\ exp}^2$ values presented in Table 5 for each pair of systems are not so identical as the ones presented in Tables 3 and 4, where only two factors were taken into account. However, when both experimental physicochemical and biological data are correlated and the three factors are examined, it is already expected not to obtain exactly the same values for $SD_{corr\ cal}^2$ and $SD_{corr\ exp}^2$. This is because of the different number and type

Table 6. Evaluation of Correlations between Experimental Retention Factors (log *k*) in 50 mM STC and Experimental Properties Values for the Detailed Biological Systems

	MEKC system: 50 mM STC, pH 8											
biological systems	d	SD_{bio}^2	n_{bio}	$(p\ SD_{MEKC})^2$	n_{MEKC}	F_{MEKC}	SD_d^2	n_d	F_d	$SD_{corr\ cal}^2$	$SD_{corr\ exp}^2$	$n_{corr\ exp}$
soil−water sorption	0.077	0.029	49	0.007	40	0.24	0.005	69	0.17	0.041	0.022	14
chemical toxicity for the guppy fish (*Poecilia reticulata*)	0.105	0.078	148	0.004	40	0.05	0.030	69	0.38	0.113	0.086	22
chemical toxicity for the *Tetrahymena Pyriformis* protozoa	0.156	0.042	192	0.008	40	0.19	0.039	69	0.93	0.090	0.055	29
chemical toxicity for the fathead minnow fish (*Pimephales promelas*)	0.183	0.076	196	0.010	40	0.13	0.056	69	0.74	0.143	0.087	24

of solutes considered in the calculation of the contribution of each factor (n_{bio}, n_{MEKC}, n_d) and in the correlation of the experimental values ($n_{corr\ exp}$), where only those solutes with a known value for both the biological and physicochemical properties can be included. Nevertheless, as it can be observed in Table 5, $SD_{corr\ cal}^2$ gives a good estimation of $SD_{corr\ exp}^2$, which means that the proposed method is suitable to know how much each factor contributes to the overall variance. Thus, it is possible to evaluate the goodness of a biological−chromatographic correlation according to the contribution of each factor. Once these contributions have been calculated, it is necessary to pay attention to the contribution of the precision of the correlated original data, especially the biological data that usually has a greater uncertainty associated. Since in the correlation of experimental data it is not expected to obtain a better precision than the one of the original data, the SD_{bio}^2 value must be compared with the SD_d^2 value for the considered pair of systems. If we look at the examples presented in Table 5, the contribution of the dissimilarity between the correlated systems is lower than the contribution of the variance of the biological data ($SD_d^2 < SD_{bio}^2$) for the SDS-Brij 35 and fathead minnow pair of systems as well as for the TTAB and fathead minnow systems. This means that not much more error is introduced in these correlations with regard to the error that the original data contains. For this reason, the correlations between SDS-Brij 35 and fathead minnow, on one hand, and TTAB and fathead minnow, on the other hand, are considered good enough and both MEKC systems are thought to emulate well the toxicity in the fathead minnow fish. On the contrary, the SDS and LPFOS systems are not considered to simulate well this biological system. As it is shown in Table 5, in both cases the error attributed to the dissimilarity of the biological-MEKC systems exceeds the error attributed to the precision of the biological data ($SD_d^2 > SD_{bio}^2$). The high SD_d^2 values are in agreement with the high d values for these pair of systems. This means that the properties of the correlated

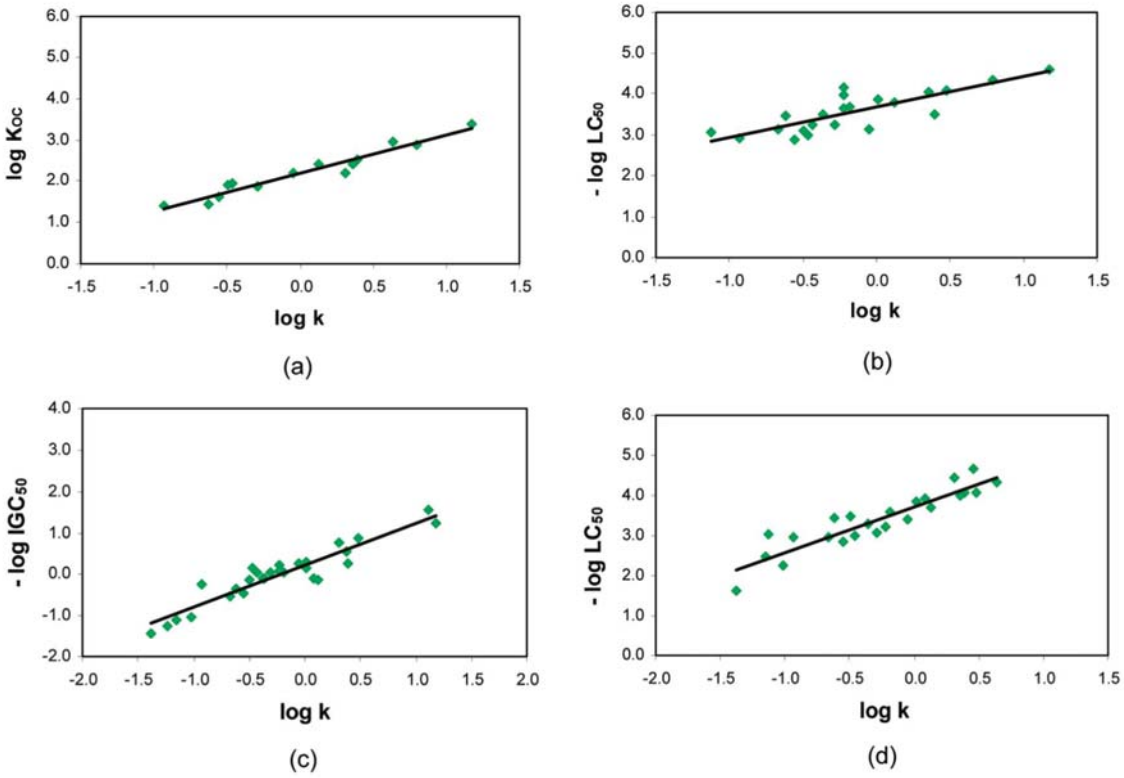


Figure 2. Plots of experimental (a) log K_{OC} (soil−water partition logarithm), (b) −log LC_{50} (median mortality lethal concentration logarithm) for the guppy fish, (c) −log IGC_{50} (median inhibitory growth cell concentration logarithm) for the *Tetrahymena Pyriformis* protozoa, and (d) −log LC_{50} (median mortality lethal concentration logarithm) for the fathead minnow fish vs experimental log *k* (retention factor logarithm) in 50 mM STC.

systems are too different to obtain correlations in which biological values can be correctly predicted through measurements in these MEKC systems.

This conclusion can be also reached by testing with the Fisher *F*-test the significance of the variances introduced by the chromatographic system and the dissimilarity between correlated systems in reference to the biological one. In Table 5, we have included the calculated *F* values for this two factors ($F_{MEKC} = \{(p\ SD_{MEKC})^2/SD_{bio}^2\}$ and $F_d = (SD_d^2/SD_{bio}^2)$). F_{MEKC} values are always lower than 1, showing that the variance of the chromatographic system does not significantly increase the variance of the biological system. The F_d values for 50 mM SDS−15 mM Brij35 and 20 mM TTAB are also lower than 1. Only the 40 mM SDS and 40 mM LPFOS systems have F_d values higher than 1 and even larger than 1.37, which is the critical *F* value for a significant contribution of SD_d^2 for a significance level of 0.05 and 68 ($n_d - 1$) and 195 ($n_{bio} - 1$) degrees of freedom. Thus, it is clear that these two systems are not appropriate to emulate aquatic toxicity for the fathead minnow fish because they significantly increase the variance of the original biological data.

In the previous work where the *d* parameter was proposed,[12] it was considered that distances between 0 and 0.25 indicate that the two compared systems are similar enough to be well correlated. Under this assumption, we decided to examine correlations between MEKC-biological systems with $d < 0.25$. As a MEKC system, we chose the one composed of STC as a surfactant[31] because this one had distances lower than 0.25 regarding several biological systems. Table 6 shows the systems whose experimental biological properties were correlated against the experimental retention factors in the STC system and also the detailed distances between each pair of systems. According to the method proposed in this work, the contributions of the three factors to the overall variance were estimated (SD_{bio}^2, $(p\ SD_{MEKC})^2$, and SD_d^2), and their values are presented in Table 6. In the same way that in Table 5, their sum ($SD_{corr\ cal}^2$) results a good approach to the variance obtained when experimental values for both systems are correlated ($SD_{corr\ exp}^2$). Such correlations are illustrated in Figure 2. If the goodness of each correlation is examined according to the error attributed to each factor, the conclusion is that all four biological systems are well emulated by the STC system. This conclusion comes out from the fact that for all the STC-biological systems, the SD_d^2 value is lower than the SD_{bio}^2 value. In agreement with distances, the higher is *d*, the higher is the contribution of the dissimilarity between systems, but in any case the error due to this factor increases the overall variance much more than it does the error of the original data. Therefore, these biological properties are considered to be well predicted by means of the retention factor in the STC system.

The same conclusion is also reached by *F*-test because all calculated F_{MEKC} and F_d values are lower than 1, as can be observed in Table 6. Moreover, it is worth to point out that the examples in Table 6 demonstrate how important is considering the contributions of each factor to interpret the correlations. Thus, in spite of the greater distance and SD_d^2 value for the toxicity in the *Tetrahymena Pyriformis*-STC systems, the overall variance obtained in its correlation ($SD_{corr\ exp}^2 = 0.055$) is lower than the one obtained for the toxicity in guppy-STC systems ($SD_{corr\ exp}^2 = 0.086$), and this is due to the greater error associated with the original data of this last biological system ($SD_{bio}^2 = 0.078$). Figure 3 summarizes the procedure proposed in this work to evaluate the modeling of biological processes by means of chromatographic systems.

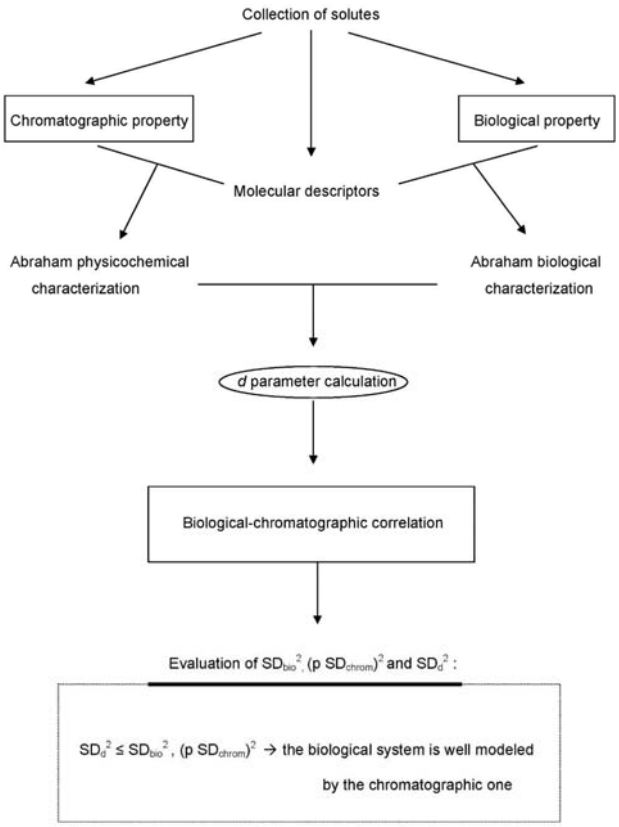


Figure 3. Diagram of the procedure for evaluating the performance of chromatographic systems to emulate biological ones.

CONCLUSIONS

A new method is proposed to evaluate the performance of chromatographic systems to model biological ones. This method is based on the examination of the three factors that contribute to the variance obtained when biological systems are correlated against chromatographic ones: the precision of the biological data, the precision of the chromatographic data, and the dissimilarity between the two correlated systems. It has been shown that the contribution of these factors can be easily estimated through the statistics of the Abraham characterization of the systems and the values of the properties calculated from these characterizations.

By means of this method, we can conclude that a biological property is well predicted by means of a chromatographic measurement if the error attributed to the dissimilarity between the two systems is lower or similar to the error attributed to the precision of the correlated original data, especially the biological data that usually is more uncertain than the chromatographic data. The proposed method can complement standard statistical procedures such as the use of training and test sets. Moreover, it can be used as a preliminary procedure to establish what chromatographic systems are worth being tested to emulate a

particular biological system, provided that they have been well characterized by the solvation parameter model. The correlations between several biological and MEKC systems have been assessed according to this method, observing that some of the considered biological systems are well emulated by MEKC systems.

ACKNOWLEDGMENT

We thank the Ministerio de Educación of the Spanish Government and the Fondo Europeo de Desarrollo Regional of the European Union (Project CTQ2007-61608/BQU) for financial support. M.H.-R. also thanks the Ministerio de Educación of the Spanish Government for being supported by a scholarship (Grant AP2007-01109).

Received for review October 5, 2010. Accepted November 4, 2010.

AC102626U

analytical chemistry

Article

pubs.acs.org/ac

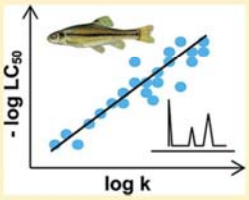
Modeling Nonspecific Toxicity of Organic Compounds to the Fathead Minnow Fish by Means of Chromatographic Systems

Marta Hidalgo-Rodríguez, Elisabet Fuguet, Clara Ràfols, and Martí Rosés*

Departament de Química Analítica and Institut de Biomedicina de la Universitat de Barcelona, Universitat de Barcelona, Martí i Franquès, 1-11, E-08028 Barcelona, Spain

Supporting Information

ABSTRACT: The performance of chromatographic systems to mimic aquatic toxicity to the fathead minnow fish is evaluated taking into account the factors that contribute to the variance of biological–chromatographic correlations. These factors are the precision to measure the fathead minnow toxicity, the precision of the surrogate chromatographic system, and the error from the dissimilarity between the fathead minnow and chromatographic systems. The precisions are estimated through the characterization of the systems by the solvation parameter model. Several chromatographic systems as well as the common reference octanol–water partition system have been selected to test their ability to model the nonspecific toxicity to the fathead minnow by means of the proposed approach. Predictions and experimental tests show that the micellar electrokinetic chromatography system of sodium taurocholate and chromatographic measurements using an immobilized artificial membrane column provide the most precise estimations of this biopartitioning property. The octanol–water partition system, the conventional C18 high-performance liquid chromatography systems, and the micellar electrokinetic chromatography system of sodium dodecylsulfate show worse performances.



The environment is regularly exposed to chemical substances derived from industrial and natural biological processes. Aquatic toxicity studies are required in order to assess the environmental hazard and risk of chemicals that are released into our waterways.[1] Standard test methods and experimental protocols based on testing with aquatic organisms have been established for evaluating the environmental effects of chemicals. Particularly, the fathead minnow fish (*Pimephales promelas*) is one of the species outlined by EPA guidelines[2] as a biological model in aquatic toxicology studies. The fathead minnow model plays an essential role in this field because it has been widely used both for regulatory testing and research, for example to screen the toxicity of new compounds or to examine life-stages/end points for chemicals with different modes of action.[3] However, the experimental investigation through animal tests is usually a lengthy, costly, and technically difficult process, even ethically questionable. For these reasons, surrogate physicochemical systems capable of estimating biological properties in a faster, more economic and easier manner have become of potential interest. Undoubtedly, implementation of physicochemical testing in common practices of toxicity estimation will contribute to more sustainable development of chemicals.

Physicochemical modeling is based on the premise that the structure of a compound determines its properties. Thus, many properties of solutes can be related to structural descriptors by means of quantitative structure–activity relationships (QSARs). Among the QSAR models, the solvation parameter model proposed by Abraham[4,5] provides a robust approach for characterizing the properties of partitioning processes. Both biological[6−12] and physicochemical[13−17] systems ruled by the passive transport of solutes between two phases have been successfully described by means of the solvation parameter model equation:

$$\log SP = c + eE + sS + aA + bB + vV \quad (1)$$

where SP is the dependent solute property in a given partitioning system, i.e. an equilibrium constant or some other free energy related property. The E, S, A, B, and V independent variables are the solute descriptors proposed by Abraham. E represents the excess molar refraction, S is the solute dipolarity/polarizability, A and B are the solute's effective hydrogen-bond acidity and hydrogen-bond basicity, respectively, and V is McGowan's solute volume. The coefficients of the equation are characteristic of the system and reflect its complementary properties to the corresponding solute property. As they represent the difference in solvation properties between the two phases of the partitioning system, e refers to the difference in capacity of each phase to interact with solute π- and n-electrons; s is a measure of the difference of the two phases in capacity to take part in dipole–dipole and dipole–induced dipole interactions; a and b represent the differences in hydrogen-bond basicity and acidity, respectively, between both phases; v is a measure of the relative ease of forming a cavity for the solute in the two phases. For any system, the coefficients of this correlation equation can be obtained by multiple linear regression analysis between the log SP values acquired for an appropriate group of solutes and their descriptors values.

Received: February 13, 2012
Accepted: March 12, 2012
Published: March 12, 2012

ACS Publications © 2012 American Chemical Society

The solvation parameter model has been used for describing nonspecific aquatic toxicity of neutral organic compounds to several aquatic organisms.[10,11,18,19] Thus, the observed toxicity data of compounds that act through nonspecific modes of action (lacking of covalent interactions between toxicant and organism) have been correlated through the Abraham equation (eq 1). For fathead minnow, Hoover et al.[10] obtained the following correlation equation:

$$-\log \mathrm{LC}_{50} = 0.996 + 0.418E - 0.182S + 0.417A - 3.574B + 3.377V \quad (2)$$

with $n = 198$, $r = 0.976$, SD = 0.276, and $F = 779$. LC_{50} is the median mortality lethal molar concentration of the solute (96 h end point), n denotes the number of solutes, r corresponds to the multiple correlation coefficient, SD is the standard deviation, and F is the Fischer statistic. One of the advantages of the Abraham equation is the information derived from its coefficients, which are specific of the studied system. Thus, coefficients of eq 2 provide a general understanding of the solute properties that lead to nonspecific aquatic toxicity on fathead minnow. Size is the solute property that contributes the most to increase toxicity (v is the largest positive coefficient), whereas hydrogen-bond basicity is the solute property most opposed to the transfer of the solute from water to the biomembrane (b is the lowest negative coefficient), which means that it reduces toxicity.

Once the coefficients of the solvation parameter model are known for a particular system, two approaches can be used in order to estimate the dependent solute property for other compounds not considered in the Abraham equation. The nonspecific aquatic toxicity to fathead minnow can be estimated for new compounds through eq 2, just multiplying its coefficients by the descriptors of the new compounds, if they are known or can be determined. Alternatively, eq 2 can be compared with the correlation equation of other systems characterized by means of the solvation parameter model in order to identify surrogate systems, i.e. systems with coefficients proportionally close to these of eq 2. Thus, good correlations are expected between the fathead minnow data and the data from these surrogate systems, being able then to estimate the toxicity to fathead minnow.

This last approach is the aim of the present work, which is focused on the performance of chromatographic systems, basically micellar electrokinetic chromatography (MEKC) and high-performance liquid chromatography (HPLC) systems, to emulate nonspecific toxicity of neutral organic compounds to the fathead minnow fish. As well, the octanol–water partition system has been studied because it is widely accepted as the reference model for membrane biopartitioning and thus it is the most common system in biological–physicochemical correlations.[20,21] Despite this fact, the measurement of octanol–water partition coefficients is lengthy and tedious. On the contrary, chromatographic systems are very popular because they are easy to use and for this reason several authors have chosen them to test their suitability to model biological processes.[22−27]

In this study, possible chromatographic systems surrogated to fathead minnow toxicity have been identified through calculation of the d distance parameter.[23] The precision achieved in the toxicity prediction by the surrogate systems has been evaluated from the precision of the correlated biological data, the precision of the correlated chromatographic data, and the error coming from the dissimilarity between the two correlated systems (measured by the d distance).[28] Finally, the prediction of the performance of the chromatographic surrogate systems has been checked by direct correlation of experimental data from fathead minnow toxicity and chromatographic log k values.

THEORY

Similarity between Systems. The similarity between systems characterized by means of the solvation parameter model (eq 1) can be measured through the d distance parameter,[23] which is calculated according to the following expression:

$$d = [(e_{ui} - e_{uj})^2 + (s_{ui} - s_{uj})^2 + (a_{ui} - a_{uj})^2 + (b_{ui} - b_{uj})^2 + (v_{ui} - v_{uj})^2]^{1/2} \quad (3)$$

where the subscript u means that all their coefficients have been previously normalized and the subscripts i, j represent the two compared systems. The coefficients are normalized as follows:

$$e_u = \frac{e}{l} \quad (4)$$

$$s_u = \frac{s}{l} \quad (5)$$

$$a_u = \frac{a}{l} \quad (6)$$

$$b_u = \frac{b}{l} \quad (7)$$

$$v_u = \frac{v}{l} \quad (8)$$

where l is the length of the coefficients' vector:

$$l = (e^2 + s^2 + a^2 + b^2 + v^2)^{1/2} \quad (9)$$

Considering the coefficients of any system as a vector in a five-dimensional space, the d parameter measures the distance between the normalized unitary vectors of a pair of systems. Thus, the d distance provides a measure of the similarity between the two considered systems: the smaller the d, the closer the two systems. For instance, two systems, one biological (bio) and one chromatographic (chrom)

$$\log \mathrm{SP}_{\mathrm{bio}} = c_{\mathrm{bio}} + l_{\mathrm{bio}}(e_{u\mathrm{bio}}E + s_{u\mathrm{bio}}S + a_{u\mathrm{bio}}A + b_{u\mathrm{bio}}B + v_{u\mathrm{bio}}V) \quad (10)$$

$$\log \mathrm{SP}_{\mathrm{chrom}} = c_{\mathrm{chrom}} + l_{\mathrm{chrom}}(e_{u\mathrm{chrom}}E + s_{u\mathrm{chrom}}S + a_{u\mathrm{chrom}}A + b_{u\mathrm{chrom}}B + v_{u\mathrm{chrom}}V) \quad (11)$$

will show $d = 0$ if they have the same normalized vector, and thus, they must show a linear correlation

$$\log \mathrm{SP}_{\mathrm{bio}} = c_{\mathrm{bio}} - \frac{l_{\mathrm{bio}}}{l_{\mathrm{chrom}}}c_{\mathrm{chrom}} + \frac{l_{\mathrm{bio}}}{l_{\mathrm{chrom}}}\log \mathrm{SP}_{\mathrm{chrom}} \quad (12)$$

In practice, the correlation is tested by plotting the experimental biological data against the experimental chromatographic data and thus:

$$\log \mathrm{SP}_{\mathrm{bio}} = q + p\log \mathrm{SP}_{\mathrm{chrom}} \quad (13)$$

dx.doi.org/10.1021/ac2034453 | Anal. Chem. 2012, 84, 3446−3452

Table 1. Physicochemical Systems Considered in This Work

systems	type of system[a]	experimental conditions	ref
1 DPPG:DPPC	LEKC	12 mM dipalmitoylphosphatidyl glycerol (DPPG):dipalmitoylphosphatidyl choline (DPPC) (30:70), 25 mM HEPES pH 7.5	30
2 STC	MEKC	50 mM sodium taurocholate (STC), 20 mM phosphate pH 7	31
3 pentanol–water	solvent–water partition	diverse	32
4 SDS microemulsion	MEEKC	1.4 wt.% sodium dodecyl sulfate (SDS), 6.49 wt.% butan-1-ol, 0.82 wt.% heptane, 500 mM phosphate+100 mM borate pH 7	33
5 IAM	HPLC	immobilized artificial membrane prepared with phosphatidyl choline (IAM.PC.DD2) column, 10 mM phosphate buffer pH 7:acetonitrile (60:40)	23
6 octanol–water	solvent–water partition	diverse	13
7 RP18	HPLC	Xterra RP18 column, 10 mM phosphate buffer pH 7:acetonitrile (60:40)	23
8 MSC18	HPLC	Xterra MSC18 column, 10 mM phosphate buffer pH 7:acetonitrile (60:40)	23
9 poly-(SDeS)	polymeric EKC	10 g/L poly(sodium 9-decenyl sulfate) (SDeS), 100 mM phosphate pH 7	34
10 SDS	MEKC	40 mM sodium dodecylsulfate (SDS), 20 mM phosphate pH 7	35, 36
11 Brij 35 C18	MELC	AT. chrom C18 column, 6.0 w/w % polyoxyethylene(23) dodecyl ether (Brij 35), 6.6 w/w % butanol, 0.8 w/w % heptane, 86.6 w/w % 50 mM phosphate buffer pH 7	24
12 LPFOS	MEKC	40 mM lithium perfluorooctanesulfonate (LPFOS), 20 mM phosphate pH 7	35, 36

[a]LEKC = liposome electrokinetic chromatograpy; MEKC = micellar electrokinetic chromatography; MEEKC = microemulsion electrokinetic chromatography; HPLC = high-performance liquid chromatography; MELC = microemulsion liquid chromatography.

The closer the distance between the two systems to zero, the closer q and p are to the intercept and slope predicted by eq 12.

Precision of the Correlation of the Biological–Chromatographic Systems. The key issue to know the performance that could by achieved by a physicochemical system in the emulation of a biological one falls on the overall variance obtained in the biological–physicochemical correlation. Three factors contribute to the overall variance obtained in such correlations: the precision of the biological data (σ_{bio}), the precision of the chromatographic data (σ_{chrom}), and the dissimilarity between the two correlated systems (σ_d). It was demonstrated[28] that these three contributions can be easily estimated through the statistics of the Abraham characterizations and the calculated values from these characterizations, according to the following expressions:

$$\sigma_{bio}^2 \approx SD_{bio}^2 \quad (14)$$

$$\sigma_{chrom}^2 \approx (pSD_{chrom})^2 \quad (15)$$

$$\sigma_d^2 \approx SD_d^2 \quad (16)$$

where SD_{bio} is the standard deviation of the Abraham characterization of the biological system, SD_{chrom} corresponds to the standard deviation of the characterization of the chromatographic system, p is the slope of the correlation between the two considered systems, and SD_d is the standard deviation of the correlation between calculated values for both biological and chromatographic properties. Such calculated values are obtained applying the corresponding Abraham equation to the descriptors of the solutes of the test set (usually the biological set).

The overall variance of the biological–chromatographic correlation (SD_{corr}^2) is estimated by the addition of these three terms:

$$SD_{corr\ cal}^2 = SD_{bio}^2 + (pSD_{chrom})^2 + SD_d^2 \quad (17)$$

The lower the $SD_{corr\ cal}^2$, the better the surrogate system estimates the biological property.

EXPERIMENTAL SECTION

Apparatus. HPLC and MEKC measurements were done in the same instruments previously described.[23,28]

Chemicals. Methanol (HPLC grade), hydrochloric acid (25% in water), sodium hydroxide (>99%), sodium dihydrogenphosphate monohydrate (>99%), and sodium dodecylsulfate (SDS) (>99%) were from Merck (Darmstadt, Germany). Acetonitrile (HPLC grade) was from VWR International (West Chester, Pennsylvania, USA). Taurocholic acid sodium salt hydrate (98%) was from Acros Organics (Geel, Belgium). Dodecanophenone (98%) was from Aldrich (Steinheim, Germany). Water was purified by a Milli-Q plus system from Millipore (Bedford, MA, USA), with a resistivity of 18.2 MΩ cm. The test solutes employed were reagent grade or better and obtained from several manufacturers (Merck, Sigma (Steinheim, Germany), Fluka (Steinheim, Germany), Aldrich, Carlo Erba (Milano, Italy), Baker (Deventer, Netherlands)).

Procedure. *MEKC Systems.* The sodium taurocholate (STC) system was a separation solution 50 mM in taurocholic acid sodium salt hydrate and 20 mM in phosphate aqueous buffer adjusted to pH 7. For the SDS system, the separation solution was 40 mM in SDS and 20 mM in phosphate aqueous buffer adjusted to pH 7. All solutes were dissolved in methanol (used as electroosmotic flow marker) and contained 2 mg mL^{-1} of dodecanophenone as micellar marker.[29] The concentration of the solutes was 2 mg mL^{-1}. All solutions were filtered through 0.45-μm nylon syringe filters obtained from Albet (Dassel, Germany). Samples were introduced into the capillary by applying a pressure of 0.5 psi for 1 s. All measurements were taken in triplicate.

HPLC Systems. The mobile phase was a mixture of acetonitrile:10 mM phosphate aqueous buffer adjusted to pH 7 (40:60). All solutes were dissolved in methanol, and their concentrations were 0.1 mg mL^{-1}. The injection volume was 10 μL. Isocratic conditions were used at a flow rate of 1 mL min^{-1}. The column hold-up time was determined by using an aqueous solution of potassium bromide (0.1 mg mL^{-1}) as an unretained solute, whose detection was performed at 200 nm. All measurements were taken in triplicate.

Table 2. Estimation of the Contributions that Determine the Overall Variance in the Correlations between Data for Fathead Minnow and Chromatographic and Partitioning Data of the Considered Physicochemical Systems

	nonspecific toxicity to the fathead minnow fish, $SD_{bio}^2 = 0.076$, $n_{bio} = 198$						
systems	d	q_{cal}	p_{cal}	$(p_{cal}\ SD_{chrom})^2$	n_{chrom}	SD_d^2	$SD_{corr\ cal}^2$
1 DPPG:DPPC LEKC	0.095	3.48	1.08	0.006	27	0.011	0.093
2 STC MEKC	0.147	3.66	1.35	0.023	56	0.014	0.113
3 pentanol–water	0.153	1.14	1.01	0.033	106	0.064	0.174
4 SDS MEEKC	0.176	2.72	1.11	0.010	53	0.093	0.179
5 IAM HPLC	0.186	3.41	2.75	0.017	51	0.039	0.132
6 octanol–water	0.198	1.39	0.88	0.010	613	0.083	0.170
7 RP18 HPLC	0.259	2.20	1.95	0.012	55	0.149	0.237
8 MSC18 HPLC	0.295	2.24	1.72	0.015	55	0.209	0.300
9 poly-(SDeS) EKC	0.335	3.63	1.14	0.006	35	0.154	0.236
10 SDS MEKC	0.347	2.96	1.26	0.023	63	0.111	0.210
11 Brij 35 C18 MELC	0.475	1.35	2.44	0.024	26	0.591	0.692
12 LPFOS MEKC	0.750	3.44	1.33	0.064	62	0.661	0.801

Table 3. Experimental Evaluation of the Performance of the Studied Chromatographic Systems to Emulate Nonspecific Toxicity of Organic Compounds to the Fathead Minnow Fish[a]

	nonspecific toxicity to the fathead minnow fish					
systems	q	p	$SD_{corr\ exp}^2$	n	r	F
2 STC MEKC	3.68 (0.05)	1.12 (0.07)	0.118	50	0.917	254
5 IAM HPLC	3.67 (0.04)	2.20 (0.14)	0.084	48	0.915	235
6 Octanol–water	1.48 (0.06)	0.85 (0.02)	0.152	183	0.946	1530
7 RP18 HPLC	2.79 (0.11)	1.43 (0.13)	0.139	48	0.854	124
8 MSC18 HPLC	2.96 (0.12)	1.13 (0.13)	0.191	48	0.792	78
10 SDS MEKC	3.22 (0.07)	1.02 (0.08)	0.147	49	0.868	143

[a]Standard deviations in brackets.

RESULTS AND DISCUSSION

Selection of Systems to Emulate Nonspecific Toxicity of Neutral Organic Compounds to Fathead Minnow. A large number of systems characterized through the Abraham equation (eq 1) have been examined in order to identify surrogate systems for the toxicity to fathead minnow. Some illustrative chromatographic and partitioning systems of different nature are presented in Table 1. They are divided into three groups: solvent–water partition systems, systems of electrokinetic chromatography (EKC), and HPLC systems. All of them are based on chemical interactions that control the distribution of solutes between the two phases that compose the system, which is the same principle that rules the nonspecific toxicity to fathead minnow.

The Abraham correlation equations of the considered chromatographic and partitioning systems, and a careful comparison of them are given in the Supporting Information. The calculated d values of the compared systems are presented in Table 2.

We applied the approach described in the Theory section for estimating the precision of biological–chromatographic correlations. Since p in eq 15 is the experimental value of eq 13, which is not initially known, we used the slope of the correlation between calculated values (p_{cal}) as an estimation of the slope of the experimental correlation (p). In order to obtain p_{cal} and also SD_d^2, calculated values from the Abraham characterizations were required, and we used for this purpose the set of solutes employed in the Abraham characterization of the toxicity to fathead minnow (eq 2),[10] i.e. the biological system to be emulated (198 solutes).

Table 2 shows the contributions of variance that would be obtained for each pair of fathead minnow-physicochemical systems according to our prediction method, as well as the resulting overall variance from the sum of such contributions ($SD_{corr\ cal}^2$) (eq 17). The intercept (q_{cal}) and slope (p_{cal}) of the calculated correlations are also presented.

Since $(p_{cal}\ SD_{chrom})^2$ is usually much lower than SD_{bio}^2, a low overall variance and therefore a good performance will be obtained if SD_d^2 is lower than or similar to SD_{bio}^2. Taking this criterion into account, systems 1, 2, 3, and 5 should emulate well nonspecific toxicity of neutral organic compounds to fathead minnow ($SD_d^2 < SD_{bio}^2$, which will give an overall variance less than twice the variance of the original biological data). Among these systems, we chose the STC MEKC system (system 2) and the IAM HPLC system (system 5) to determine and correlate their experimental values against the ones of the fathead minnow system because measurements in the LEKC system of DPPG:DPPC (system 1) and the pentanol–water partition system (system 3) are more time-consuming and costly. Liposome systems are very troublesome regarding vesicle preparation and stability and chromatographic systems (both HPLC and EKC) are in general preferable to water–partition systems because they are usually more economic, easier to automate, and faster. However, they have some limitations concerning the high retention time of some substances in HPLC and the significant error associated to MEKC measurements in extreme regions of the elution window.

The Brij 35 C18 MELC system (system 11) and the MEKC system of LPFOS (system 12) seem to be very weak models for the estimation of the toxicity to fathead minnow. They have SD_d^2 much higher than SD_{bio}^2, meaning that a lot of error

dx.doi.org/10.1021/ac2034453 | Anal. Chem. 2012, 84, 3446–3452

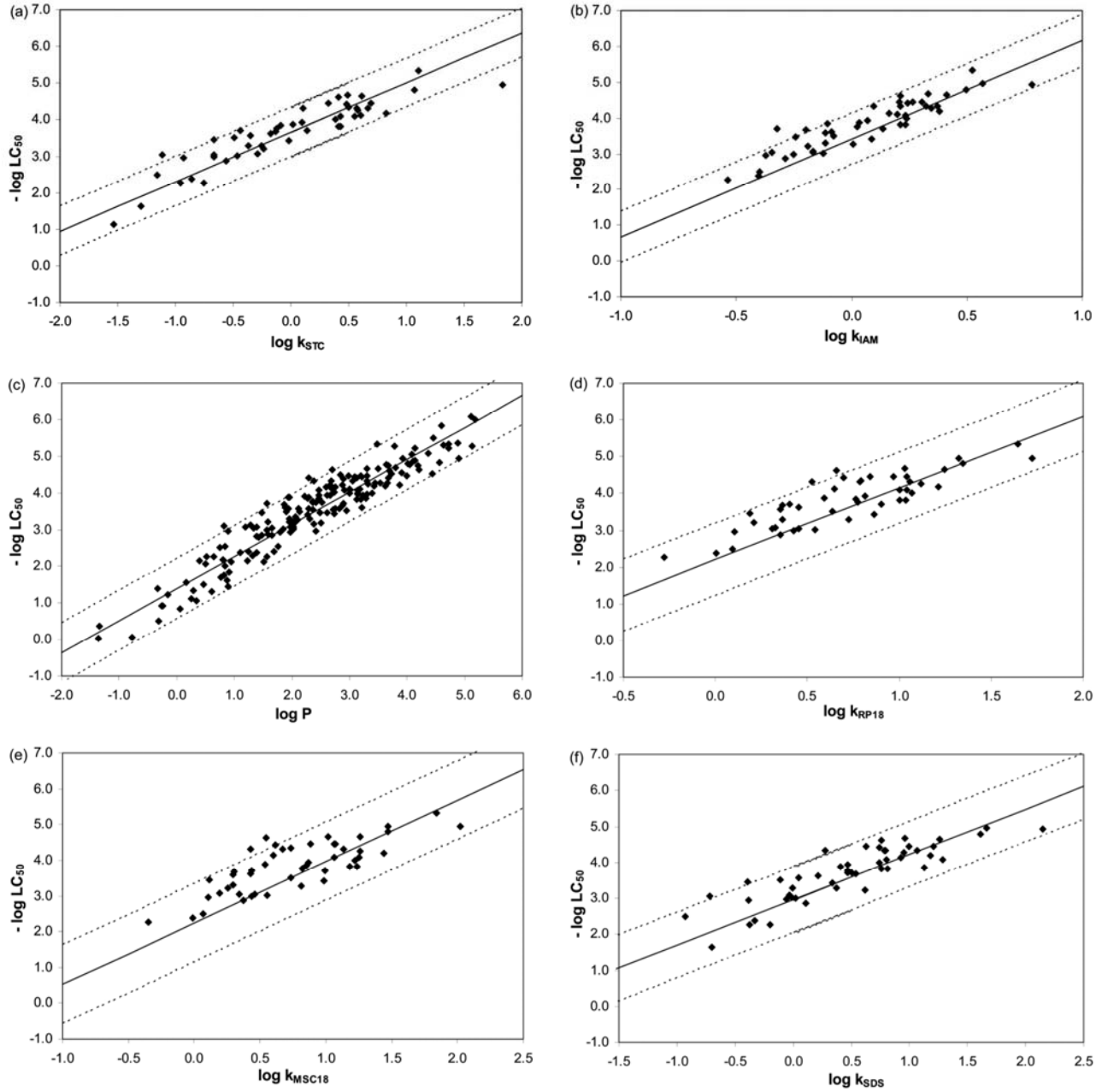


Figure 1. Plots of experimental $-\log LC_{50}$ (median mortality lethal concentration logarithm) for fathead minnow vs experimental log k (retention factor logarithm) or log P (octanol–water partition coefficient logarithm) in the studied physicochemical systems: (a) STC MEKC, (b) IAM HPLC, (c) octanol–water, (d) RP18 HPLC, (e) MSC18 HPLC, and (f) SDS MEKC. The lines correspond to the correlation between values calculated for both properties through their Abraham model characterizations (—) and the corresponding range for 95% confidence level ($\pm 2\ SD_{corr\ cal}$) (- - -).

would be introduced in their correlations due to the dissimilarity between these chromatographic systems and the fathead minnow. For the rest of systems of Table 2, their SD_d^2 values are similar or a bit higher than their SD_{bio}^2 values, what means that they are not expected to model toxicity to fathead minnow so well as systems 1, 2, 3, and 5, but fairly good. Among them, we selected systems 6, 7, 8, and 10 to study their experimental correlations since they are quite popular as physicochemical systems. It was interesting to select the octanol–water partition (system 6) because it is a reference system to estimate hydrophobicity and correlate it to many biological properties, its characterization through the solvation parameter model includes a large set of solutes,[13] and we had experimental values available for a lot of solutes from a reliable database.[37] We also decided to consider systems 7 and 8 because they correspond to the common C18 chromatographic columns and there is a lot of interest in knowing their ability to predict aquatic toxicities. In the same vein, system 10 was selected due to the popularity of SDS as surfactant in MEKC.

Regarding the relationship between the d distance parameter and SD_d^2, the larger d is, the higher SD_d^2 should apparently be (the more different two systems are, the more error associated to their dissimilarity should be). However, this tendency is not exactly what we observe in Table 2. Systems with close d values

dx.doi.org/10.1021/ac2034453 | Anal. Chem. 2012, 84, 3446–3452

not always have increasing SD_d^2 values as d increases. This fact is because the distance only gives a rough idea of the similarity between pair of systems, whereas the calculation of SD_d^2 is a much more accurate estimation of how much the similarity between systems affects its correlation.

Evaluation of the Performance of Chromatographic Systems to Estimate Nonspecific Toxicity to the Fathead Minnow Fish. The predictions of the precision performance of the surrogate systems have been tested by direct correlations of experimental $-\log LC_{50}$ values[10] vs log k data through eq 13. The number of solutes with known values for both properties has been enlarged in order to obtain representative correlations. Therefore, we checked the solutes with known $-\log LC_{50}$,[10] and we selected a large set of representative solutes (see the Supporting Information) to determine their log k in the chosen chromatographic systems.

The results of the experimental correlations are shown in Table 3 and Figure 1. It is noticeable that almost all the experimental points fall into our predictions, for a confidence level of 95% ($\pm 2\ SD_{corr\ cal}$), even for the octanol−water partition system (Figure 1c) with many more experimental data.

According to the statistics, the IAM HPLC and the STC MEKC systems (Figure 1a and b) are the best choice to model the nonspecific toxicity to fathead minnow because their correlations have associated the lowest experimental variances, which agrees with our predictions ($SD_{corr\ cal}^2$ of Table 2). The similarity of these systems to the fathead minnow one leads to a low SD_d^2 value (in accordance with the low d value) and thus to a low overall error ($SD_{corr\ exp}^2$).

The rest of tested systems provide not so good estimations because they are already more different to the fathead minnow system, which is reflected in their SD_d^2 values that are higher than their SD_{bio}^2 values. The octanol−water partition, RP18 HPLC, and SDS MEKC systems (Figure 1c, d, and f) give fair experimental correlations since their SD_d^2 are only slightly higher than their SD_{bio}^2. Although the SDS MEKC system presents the highest d value, the contribution of error due to its dissimilarity toward the fathead minnow is not the highest one. Its SD_d^2 value, and also the corresponding $SD_{corr\ cal}^2$ value, lay between the ones of octanol−water and RP18 HPLC systems, which agrees with the obtained $SD_{corr\ exp}^2$. The chromatographic system of MSC18 (Figure 1e) has the highest experimental variance and shows the worst performance to predict nonspecific toxicity to fathead minnow because of their very poor similarity, indicated by the high SD_d^2 value.

■ CONCLUSIONS

Models based on chromatographic systems are very useful for the emulation of biopartitioning processes. The precision of the estimation of a biological property by means of chromatographic measurements depends on three factors: the variance of the biological data, the variance of the chromatographic data, and the variance because of the dissimilarity between the two correlated systems. The proposed method for prediction of these contributions through the solvation parameter model has been proved to calculate well the resulting overall variance. Several chromatographic systems to mimic the nonspecific toxicity of neutral organic compounds to the fathead minnow fish have been identified according to this prediction method. Experimental tests agree with our predictions and show that the best surrogate systems for the nonspecific toxicity to fathead minnow are the IAM HPLC and the STC MEKC systems. The octanol−water partition, the SDS MEKC, and the common C18 HPLC systems show a worse performance, although their correlations are still quite acceptable.

■ ASSOCIATED CONTENT

Supporting Information

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: 34 93 403 92 75. Fax: 34 93 402 12 33. E-mail: marti.roses@ub.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the Ministerio de Ciencia e Innovación of the Spanish Government and the Fondo Europeo de Desarrollo Regional (FEDER) of the European Union (project CTQ2010-19217/BQU) for financial support. M.H-R. also thanks the Ministerio de Educación of the Spanish Government for a supporting scholarship (AP2007-01109).

■ REFERENCES

(1) Van Leeuwen, C. J.; Hermens, J. L. M. *Risk assessment of chemicals: an introduction*; Kluwer Academic: Dordrecht, 1995.
(2) http://www.epa.gov/ncct/dsstox/sdf_epafhm.html.
(3) Ankley, G. T.; Villeneuve, D. L. *Aquat Toxicol* **2006**, *78*, 91−102.
(4) Abraham, M. H. *Chem. Soc. Rev.* **1993**, *22*, 73−83.
(5) Abraham, M. H.; Chadha, H. S.; Martins, F.; Mitchell, R. C.; Bradbury, M. W.; Gratton, J. A. *Pestic. Sci.* **1999**, *55*, 78−88.
(6) Platts, J. A.; Abraham, M. H.; Zhao, Y. H.; Hersey, A.; Ijaz, L.; Butina, D. *Eur. J. Med. Chem.* **2001**, *36*, 719−730.
(7) Abraham, M. H.; Zhao, Y. H.; Le, J.; Hersey, A.; Luscombe, C. N.; Reynolds, D. P.; Beck, G.; Sherborne, B.; Cooper, I. *Eur. J. Med. Chem.* **2002**, *37*, 595−605.
(8) Abraham, M. H.; Martins, F. *J. Pharm. Sci.* **2004**, *93*, 1508−1523.
(9) Abraham, M. H.; Ràfols, C. *J. Chem. Soc. Perkin Trans. 2* **1995**, 1843−1851.
(10) Hoover, K. R.; Acree, W. E. Jr.; Abraham, M. H. *Chem. Res. Toxicol.* **2005**, *18*, 1497−1505.
(11) Hoover, K. R.; Flanagan, K. B.; Acree, W. E. Jr.; Abraham, M. H. *J. Environ. Eng. Sci.* **2007**, *6*, 165−174.
(12) Poole, S. K.; Poole, C. F. *J. Chromatogr. A* **1999**, *845*, 381−400.
(13) Abraham, M. H.; Chadha, H. S.; Whiting, G. S.; Mitchell, R. C. *J. Pharm. Sci.* **1994**, *83*, 1085−1100.
(14) Abraham, M. H.; Rosés, M.; Poole, C. F.; Poole, S. K. *J. Phys. Org. Chem.* **1997**, *10*, 358−368.
(15) Lepont, C.; Poole, C. F. *J. Chromatogr. A* **2002**, *946*, 107−124.
(16) Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. *Electrophoresis* **2006**, *27*, 1900−1914.
(17) Vitha, M.; Carr, P. W. *J. Chromatogr. A* **2006**, *1126*, 143−194.
(18) Gunatilleka, A. D.; Poole, C. F. *Anal. Commun.* **1999**, *36*, 235−242.
(19) Gunatilleka, A. D.; Poole, C. F. *Analyst* **2000**, *125*, 127−132.
(20) Sangster, J. *Octanol-water partition coefficients: fundamentals and physical chemistry*; Wiley: Chichester, 1997.
(21) Wezel, A. P. v; Opperhuizen, A. *Crit. Rev. Toxicol.* **1995**, *25*, 255−279.
(22) Rosés, M.; Ràfols, C.; Bosch, E.; Martinez, A. M.; Abraham, M. H. *J. Chromatogr. A* **1999**, *845*, 217−226.
(23) Lázaro, E.; Ràfols, C.; Abraham, M. H.; Rosés, M. *J. Med. Chem.* **2006**, *49*, 4861−4870.
(24) Liu, J.; Sun, J.; Wang, Y.; Liu, X.; Sun, Y.; Xu, H.; He, Z. *J. Chromatogr. A* **2007**, *1164*, 129−138.

(25) Lu, R.; Sun, J.; Wang, Y.; Li, H.; Liu, J.; Fang, L.; He, Z. *J. Chromatogr. A* **2009**, *1216*, 5190−5198.
(26) Cimpean, D. M.; Poole, C. F. *Analyst* **2002**, *127*, 724−729.
(27) Poole, C. F.; Gunatilleka, A. D.; Poole, S. K. *Adv. Chromatogr.* **2000**, *40*, 159−230.
(28) Hidalgo-Rodríguez, M.; Fuguet, E.; Ràfols, C.; Rosés, M. *Anal. Chem.* **2010**, *82*, 10236−10245.
(29) Fuguet, E.; Ràfols, C.; Bosch, E.; Rosés, M. *Electrophoresis* **2002**, *23*, 56−66.
(30) Burns, S. T.; Agbodjan, A. A.; Khaledi, M. G. *J. Chromatogr. A* **2002**, *973*, 167−176.
(31) Hidalgo-Rodríguez, M.; Rivas, P.; Fuguet, E.; Ràfols, C.; Rosés, M. manuscript in preparation.
(32) Abraham, M. H.; Nasezadeh, A.; Acree, W. E. *Ind. Eng. Chem. Res.* **2008**, *47*, 3990−3995.
(33) Abraham, M. H.; Treiner, C.; Rosés, M.; Ràfols, C.; Ishihama, Y. *J. Chromatogr. A* **1996**, *752*, 243−249.
(34) Akbay, C.; Shamsi, S. A. *Electrophoresis* **2004**, *25*, 635−644.
(35) Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. *J. Chromatogr. A* **2002**, *942*, 237−248.
(36) Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. *J. Chromatogr. A* **2009**, *1216*, 6877−6879.
(37) Bio-Loom online database. http://www.biobyte.com.

Modeling nonspecific toxicity of organic compounds to the fathead minnow fish by means of chromatographic systems

Supporting information

Marta Hidalgo-Rodríguez, Elisabet Fuguet, Clara Ràfols, and Martí Rosés*

Departament de Química Analítica and Institut de Biomedicina de la Universitat de Barcelona, Universitat de Barcelona, Martí i Franquès, 1-11, E-08028 Barcelona, Spain

* Corresponding author. E-mail: marti.roses@ub.edu; Telephone: 34 93 403 92 75

TABLE OF CONTENTS

	PAGE
Comparison of systems according to their Abraham characterization equations	S-2
Selection of representative solutes for the experimental correlations	S-4
FIGURES	
Figure S-1: PCA for solute descriptors	S-5
TABLES	
Table S-1: Characterization of the studied systems	S-7
Table S-2: Experimental log k data	S-8

Comparison of systems according to their Abraham characterization equations

Table S-1 here

The liposome electrokinetic chromatography system of DPPG:DPPC (system 1) shows very close normalized coefficients to the fathead minnow system, which is also reflected in the very low value of the *d* distance. This means that this system is expected to emulate very well the nonspecific toxicity to fathead minnow. Regarding systems 2-6, the corresponding *d* values are quite low, in accordance with the fact that their normalized coefficients are quite similar to the ones of fathead minnow. For example, the values for a_u are in general slightly positive as the one of fathead minnow. The main difference between the coefficients falls on the s_u values, which are noticeably more negative for all the mentioned physicochemical systems.

Systems 7-12 show considerably differences in the normalized coefficients compared to the ones of fathead minnow and in consequence large *d* values. The main differences are the s_u and a_u values which are in general quite more negative (in fact, a_u is negative for these systems whereas it is positive for fathead minnow). It is worth to point out that the poly-SDeS EKC system (system 9) presents a clearly different selectivity with regard to the mentioned physicochemical and fathead minnow systems, according to the relevant differences in its normalized coefficients. a_u and s_u values for system 9 are not very negative, what differs from the mentioned physicochemical systems and in this sense it becomes more similar to fathead minnow. However, system 9 has a b_u value far less negative than the one of fathead minnow. Regarding system 11, the values for s_u and a_u are very different from the ones of fathead minnow, which leads to a high distance value between the two systems. Nevertheless, the most different normalized coefficients with regard to the ones of fathead minnow correspond to the LPFOS MEKC system (system 12). Especially, the values for a_u and b_u are very different, and also the v_u coefficient differs considerably. In agreement with this fact, the distance between system 12 and fathead minnow is huge (d=0.750), being the highest

one of the cases illustrated on the Table and meaning that the LPFOS system is a very poor system to surrogate the nonspecific toxicity to the fathead minnow fish.

The I_{bio}/I parameter gives an approximate value for the sensitivity that could be obtained if the fathead minnow toxicity is estimated by means of measurements in the considered physicochemical systems. In general, the I_{bio}/I values shown in the Table are close to 1, except for some specific cases. These cases are the HPLC systems (RP18, MSC18, and especially IAM), the Brij 35 C18 MELC system, and the LPFOS MEKC system with high I_{bio}/I values. Therefore, these systems are expected to have a bigger sensitivity for the estimation of toxicity to fathead minnow (a certain variation in the physicochemical data causes a higher variation in the estimated biological property).

Table S-2 here

Selection of representative solutes for the experimental correlations

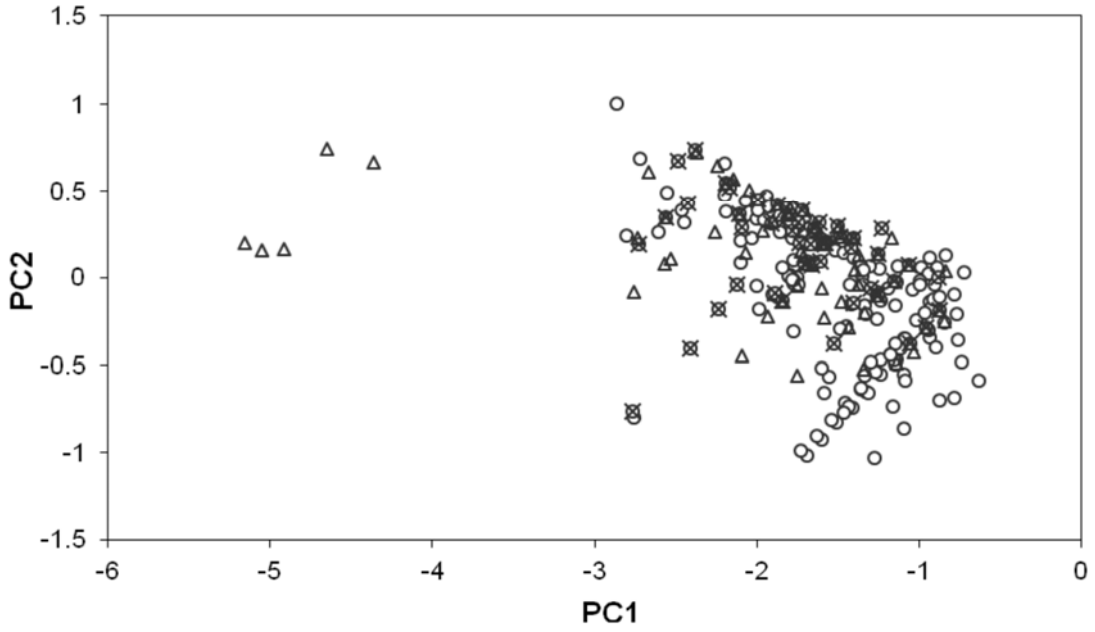
Figure S-1 here

Figure S-1 shows how the solutes of the selected MEKC systems (systems 2, 10), the selected HPLC systems (systems 5, 7, 8), and the octanol-water partition system (system 6) are distributed according to the scores values of the fathead minnow solutes.[1] The compounds chosen for the experimental correlations cover a wide range of the chemical space considered in the fathead minnow characterization. In Figures S-1a and S-1b, the unique region of the plots covered by solutes considered in the fathead minnow but not in the experimental correlations corresponds to the right-bottom quadrant, which includes solutes that we could not measure because they are not suitable for the UV detection (alcohols, ethers, hydrocarbons). In the left-upper quadrant of Figures S-1a and S-1b there are also five corticosteroids used in the chromatographic systems whose huge descriptor values make them far away from the rest of clustered substances. However, no values for toxicity to fathead minnow were measured for these solutes, and therefore their chemical space is not covered by the solutes of the biological system to be correlated.

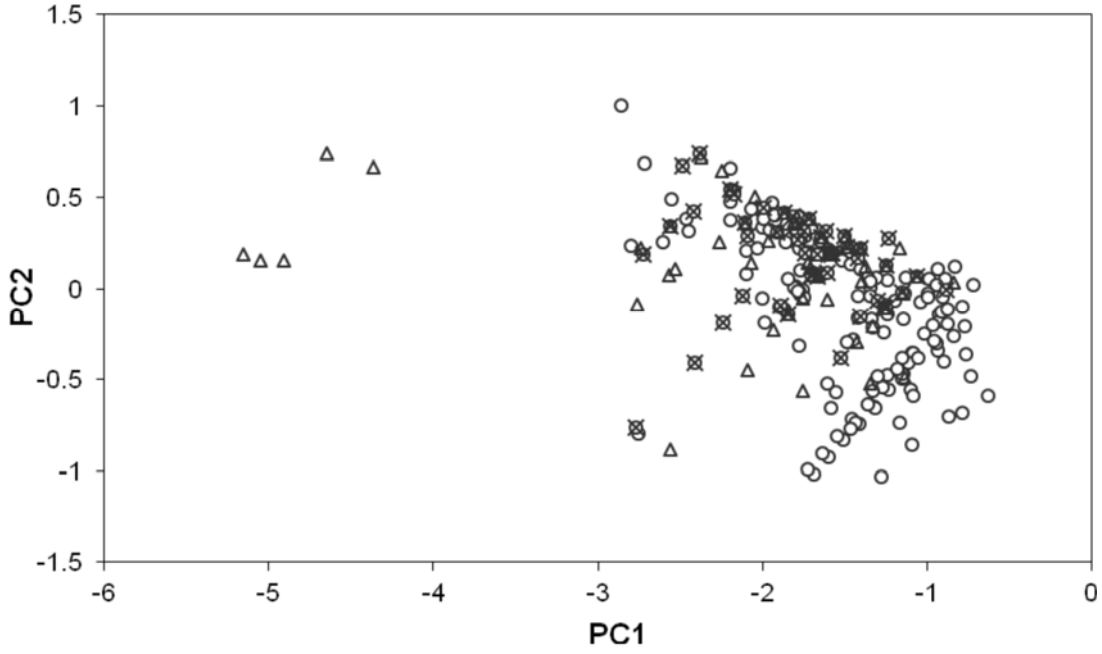
References

(1) PCA was performed with *Matlab package version R2007a from MathWorks (Natick, MA, USA).*

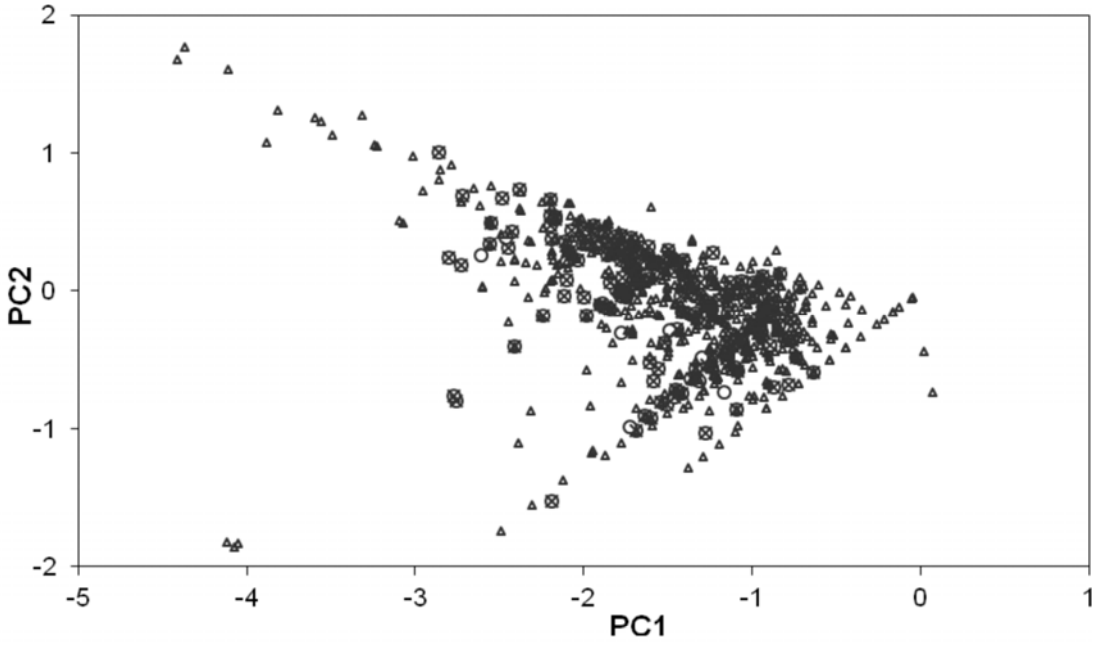
FIGURES



(a)



(b)



(c)

Figure S-1: Scores plot of the principal component analysis (PCA) of the solute descriptors included in the fathead minnow Abraham characterization (O), and the projection of the solute descriptors considered in the chromatographic Abraham characterization (Δ) and in the experimental correlation (X) for: (a) MEKC systems, (b) HPLC systems, (c) octanol-water partition system.

TABLES

Table S-1

Systems coefficients and statistics of the solvation parameter model, normalized coefficients, and values of the *I*, *d*, and I_{bio}/I parameters for the systems studied

System	Coefficients						Statistics				Normalized coefficients					*I*	*d*	I_{bio}/I
	c	*e*	*s*	*a*	*b*	*v*	n	r	SD	F	e_u	s_u	a_u	b_u	v_u			
Biological system																		
Fathead minnow	0.996	0.418	-0.182	0.417	-3.574	3.377	198	0.976	0.276	779	0.084	-0.037	0.084	-0.721	0.681	4.956	0.000	1.00
Chromatographic and partitioning systems																		
1 DPPG:DPPC LEKC	-2.21	0.45	-0.44	0.71	-3.23	3.13	27	0.995	0.070	-	0.098	-0.096	0.154	-0.703	0.681	4.597	0.095	1.08
2 STC MEKC	-1.838	0.613	-0.513	0.319	-2.396	2.390	56	0.984	0.112	305	0.175	-0.147	0.091	-0.686	0.684	3.492	0.147	1.42
3 Pentanol-water	0.185	0.367	-0.732	0.105	-3.100	3.395	106	0.988	0.180	823	0.079	-0.157	0.022	-0.664	0.727	4.671	0.154	1.06
4 SDS MEEKC	-1.133	0.279	-0.692	-0.060	-2.805	3.048	53	0.994	0.090	791	0.066	-0.164	-0.014	-0.667	0.724	4.214	0.176	1.18
5 IAM HPLC	-0.710	0.292	-0.344	0.141	-1.193	1.161	51	0.991	0.047	480	0.169	-0.199	0.081	-0.689	0.671	1.731	0.186	2.86
6 Octanol-water	0.088	0.562	-1.054	0.034	-3.460	3.814	613	0.997	0.116	23162	0.106	-0.199	0.006	-0.655	0.721	5.286	0.198	0.94
7 RP18 HPLC	-0.172	0.233	-0.460	-0.260	-1.574	1.582	55	0.996	0.056	1121	0.101	-0.200	-0.113	-0.683	0.686	2.305	0.259	2.15
8 MSC18 HPLC	-0.152	0.151	-0.463	-0.424	-1.719	1.759	55	0.994	0.072	764	0.059	-0.182	-0.167	-0.676	0.692	2.543	0.295	1.95
9 poly-(SDeS) EKC	-2.930	0.522	-0.040	-0.141	-1.637	2.946	35	0.991	0.067	-	0.153	-0.012	-0.041	-0.480	0.863	3.414	0.335	1.45
10 SDS MEKC	-1.680	0.558	-0.596	-0.266	-1.674	2.717	63	0.991	0.120	618	0.169	-0.180	-0.080	-0.507	0.822	3.305	0.347	1.50
11 Brij 35 C18 MELC	0.698	-0.003	-0.412	-0.415	-0.912	0.887	26	0.986	0.064	141	-0.002	-0.294	-0.296	-0.651	0.633	1.400	0.475	3.54
12 LPFOS MEKC	-1.410	-0.113	-0.243	-0.876	-0.455	1.966	62	0.970	0.190	180	-0.051	-0.110	-0.395	-0.205	0.887	2.216	0.750	2.24

Table S-2

Experimental data (log k) for the selected chromatographic systems for correlation to the nonspecific toxicity to the fathead minnow fish

Solute	MEKC systems		HPLC systems		
	STC	SDS	IAM	RP18	MSC18
1,2,4-trimethylbenzene	0.821	1.185	0.381	1.214	1.440
1,2-dimethylbenzene	0.411	0.758	0.211	1.000	1.184
1,3-dimethylbenzene	0.427	0.816	0.235	1.038	1.233
1,4-dichlorobenzene	0.581	0.952	0.346	1.121	1.260
1,4-dimethylbenzene	0.427	0.801	0.235	1.040	1.254
1-butanol	-1.296	-0.703	-	-	-
1-nitronaphthalene	0.694	1.207	0.304	1.042	1.069
1-pentanol	-0.956	-0.199	-	-	-
1-propanol	-1.533	-	-	-	-
2,3-benzofuran	0.099	0.468	0.070	0.815	0.868
2,4-dichlorophenol	0.572	0.793	0.373	0.788	0.670
2,4-dimethylphenol	0.019	0.405	0.032	0.593	0.538
2-naphthol	0.412	0.763	0.210	0.657	0.548
3,4-dichloroaniline	0.498	0.784	0.212	0.793	0.738
3-methylphenol	-0.372	-0.005	-0.115	0.364	0.295
4-chloro-3-methylphenol	0.486	0.737	0.242	0.698	0.616
4-chloroaniline	-0.174	0.214	-0.088	0.454	0.426
4-chlorophenol	0.101	0.274	0.096	0.527	0.429
4-methylphenol	-0.350	0.050	-0.111	0.356	0.298
4-nitroaniline	-0.287	-0.035	-0.167	0.325	0.193
4-nitrotoluene	-0.121	0.465	0.022	0.776	0.822
4-tert-butylphenol	0.692	1.000	0.263	0.842	0.887
acenaphthene	-	1.666	0.569	1.325	1.473
acetophenone	-0.557	0.111	-0.287	0.354	0.376
aniline	-0.930	-0.384	-0.372	0.104	0.107
benzamide	-0.753	-0.372	-0.535	-0.276	-0.346
benzene	-0.489	-0.110	-0.080	0.636	0.737
benzonitrile	-0.667	-0.060	-0.250	0.426	0.435
benzophenone	0.557	1.284	0.198	1.000	1.062
benzyl alcohol	-0.861	-0.331	-0.402	0.004	-0.015
biphenyl	1.066	1.608	0.495	1.348	1.469
bromobenzene	0.327	0.631	0.209	0.971	1.063
chlorobenzene	0.144	0.468	0.137	0.904	0.995
dibutyl phthalate	1.107	-	0.525	1.647	1.844
diethyl phthalate	-0.084	1.130	-0.105	0.764	0.851
dimethyl phthalate	-0.441	0.501	-0.323	0.405	0.434
diphenylamine	0.615	1.259	0.410	1.247	1.256
diuron	0.606	0.922	0.161	0.646	0.603
ethyl 4-aminobenzoate	-0.130	0.540	-0.198	0.366	0.301
ethylbenzene	0.387	0.744	0.238	1.070	1.224
furan	-1.113	-0.715	-0.344	0.307	0.338
naphthalene	0.665	1.064	0.324	1.059	1.137
nitrobenzene	-0.463	0.020	-0.121	0.544	0.551
N-methylaniline	-0.667	-0.011	-0.169	0.451	0.457

N,N-dimethylaniline	-0.258	0.372	0.007	0.725	0.815
pentylbenzene	1.832	2.147	0.781	1.722	2.021
phenol	-0.663	-0.390	-0.241	0.189	0.117
pyrrole	-1.162	-0.927	-0.401	0.095	0.071
quinoline	-0.235	0.615	-0.189	0.208	0.252
thymol	0.493	0.962	0.333	1.030	1.014
toluene	-0.019	0.342	0.086	0.861	0.983

Journal of Chromatography A, 1252 (2012) 136–145

Contents lists available at SciVerse ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Performance of chromatographic systems to model soil–water sorption

Marta Hidalgo-Rodríguez, Elisabet Fuguet, Clara Ràfols, Martí Rosés*

Departament de Química Analítica and Institut de Biomedicina de la Universitat de Barcelona, Universitat de Barcelona, Martí i Franquès, 1-11, E-08028 Barcelona, Spain

ARTICLE INFO

Article history:
Received 5 April 2012
Received in revised form 11 June 2012
Accepted 15 June 2012
Available online 25 June 2012

Keywords:
Soil
Solvation parameter model
Micellar electrokinetic chromatography
High-performance liquid chromatography
Surrogate systems

ABSTRACT

A systematic approach for evaluating the goodness of chromatographic systems to model the sorption of neutral organic compounds by soil from water is presented in this work. It is based on the examination of the three sources of error that determine the overall variance obtained when soil–water partition coefficients are correlated against chromatographic retention factors: the variance of the soil–water sorption data, the variance of the chromatographic data, and the variance attributed to the dissimilarity between the two systems. These contributions of variance are easily predicted through the characterization of the systems by the solvation parameter model. According to this method, several chromatographic systems besides the reference octanol–water partition system have been selected to test their performance in the emulation of soil–water sorption. The results from the experimental correlations agree with the predicted variances. The high-performance liquid chromatography system based on an immobilized artificial membrane and the micellar electrokinetic chromatography systems of sodium dodecylsulfate and sodium taurocholate provide the most precise correlation models. They have shown to predict well soil–water sorption coefficients of several tested herbicides. Octanol–water partitions and high-performance liquid chromatography measurements using C18 columns are less suited for the estimation of soil–water partition coefficients.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The sorption of organic chemicals by soil from water plays an essential role in determining the transport and fate of contaminants in the environment. Therefore, it becomes a process of great importance in the field of environmental protection and risk assessment [1–3].

Reliable soil sorption data is required to evaluate the ecological impact that may follow the distribution of chemicals in the environment. The distribution equilibrium of a compound between a soil and a contact aqueous phase is mainly dominated by its sorption to the soil organic matter, which varies widely with the soil type. This variability is usually minimized by normalizing the soil–water partition coefficient to the organic carbon content of the soil. Thus, the organic carbon normalized sorption coefficient, K_{OC}, is the most accepted constant for describing the soil–water sorption of a solute independently of the type of soil considered [4–6]. There are several techniques for the experimental determination of K_{OC} such as the classical batch methods [7–10], the packed column techniques [11,12], and the continuous stirred tank reactor methods [13]. Although direct determination of K_{OC} is preferred, tests involving soil are very expensive, time-consuming, and tedious, and they are also subject to a variety of experimental difficulties and artifacts. In addition, the large number of existing chemicals makes impossible to test all of them and screening methods are consequently required. Therefore, there is a potential interest in finding practical and sustainable alternatives for the estimation of K_{OC} values.

Several models based on quantitative structure–activity relationships (QSARs) have been developed with such purpose. QSARs assume that a well defined process such as the soil sorption of organic compounds is quantitatively described by means of several physicochemical properties and structural descriptors. Thus, many of the reported QSARs for modeling soil–water sorption have been established by correlation of K_{OC} against experimental molecular properties: octanol–water partition coefficients and aqueous solubilities have been widely used as descriptors [5,9,14], as well as retention measurements in reversed-phase chromatography [5,15–17]. Also QSAR models based on theoretical descriptors from the molecular structure of chemicals have been developed to find correlations with K_{OC}. For example, many authors have used molecular connectivity indices [5,9,14,18]. Models that combine a variety of calculated structural parameters and are applied to a large and diverse set of chemicals are also reported in literature [19,20].

A particular approach within QSARs, the solvation parameter model [21], has been proved to be specially useful for characterizing the uptake of neutral organic compounds by soil from water

* Corresponding author at: Departament de Química Analítica, Universitat de Barcelona, Martí i Franquès, 1-11, E-08028 Barcelona, Spain.
Tel.: +34 93 403 92 75; fax: +34 93 402 12 33.
E-mail address: marti.roses@ub.edu (M. Rosés).

0021-9673/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.chroma.2012.06.058

[22,23]. Whereas other QSARs based on a single compound descriptor are only applicable to a limited range of solute structures, the five descriptors of the solvation parameter model make it robust with respect to structure diversity [24]. In this sense, Goss and coworkers [25,26] highlighted the suitability of using polyparameter linear free energy relationships (LFERs) (e.g. the solvation parameter model) instead of LFERs of one parameter (i.e. correlations of K_{OC} against one solely descriptor).

The characterization of soil–water sorption by the solvation parameter model allows the estimation of K_{OC} values for untested chemicals as long as their solute descriptors are known. Another possibility for estimating K_{OC} data, which does not require solute descriptors, consists of identifying surrogate physicochemical systems of the soil–water sorption process and measuring a physicochemical property directly related to K_{OC} in the surrogate system. Identification of appropriate surrogate systems can be achieved within the framework of the solvation parameter model just comparing the characterizations of the soil–water sorption and different physicochemical systems [27,28]. The correlation model provided by a surrogate system should give a precise estimation of K_{OC}.

The aim of this work is finding surrogate chromatographic systems to estimate soil–water partition coefficients of unionized organic compounds using the solvation parameter model approach. We have focused on the study of chromatographic systems, basically high-performance liquid chromatography (HPLC) and micellar electrokinetic chromatography (MEKC) systems, because of two reasons. On one hand, there is analogy between the behaviour of a neutral organic solute in a reversed-phase chromatography system and the soil–water sorption process: the solute is transported in both cases by an aqueous phase in contact with a solvated organic layer and it distributes between the two phases according to its nature. On the other hand, chromatographic analyses are easy to perform, as well as cheap and fast, and therefore they can represent suitable alternatives for the direct experimental determination of K_{OC} values. We have also decided to evaluate the performance of the octanol–water partition system since it is a classical model of biopartitioning processes [29].

A systematic procedure to evaluate the performance of surrogate systems have been already established [30,31] and applied to the emulation of nonspecific toxicity of organic compounds to the fathead minnow fish [32]. According to this systematic procedure, the most promising surrogate chromatographic systems have been selected to finally test them by experimental correlations. The aim of this work is to show its usefulness with regard to soil–water sorption modeling.

2. Theory

2.1. The solvation parameter model

The solvation parameter model was proposed by Abraham [21,33] to describe processes ruled by the passive transport of neutral solutes between two phases. It is based on the following equation:

$$\log SP = c + eE + sS + aA + bB + vV \quad (1)$$

where SP is the dependent solute property in a given partitioning system, i.e. an equilibrium constant or some other free energy related property. The E, S, A, B and V independent variables are the solute descriptors proposed by Abraham. E represents the excess molar refraction, S is the solute dipolarity/polarizability, A and B are the solute's effective hydrogen-bond acidity and hydrogen-bond basicity, respectively, and V is McGowan's solute volume. The coefficients of the equation (e, s, a, b, v) are characteristic of the system and reflect its complementary properties to the corresponding solute property. For any system, the coefficients of this correlation equation can be obtained by multiple linear regression analysis between the log SP values acquired for an appropriate group of solutes and their descriptor values.

Many partitioning systems, both biological [23,34–39] and physicochemical [40–44] ones, have been successfully described by means of Eq. (1). Thus, Poole and Poole [23] characterized the sorption of neutral organic compounds by soil from water:

$$\log K_{OC} = 0.21 + 0.74E - 0.31A - 2.27B + 2.09V \quad (2)$$

The statistics of this correlation equation were $n = 129$, $r = 0.977$, SD = 0.248, and $F = 655$, where n indicates the number of solutes included in the characterization, r is the multiple correlation coefficient, SD corresponds to the standard deviation, and F denotes the Fischer statistic. Eq. (2) gives insight into the intermolecular interactions that control the soil–water sorption mechanism. Size is the solute property that contributes the most to the partition of compounds towards soil since v is the largest positive coefficient. On the contrary, solute's hydrogen-bond basicity is the property that plays a major role in reducing soil sorption because b is the most negative coefficient. The fact that s is 0 indicates that soil and water present the same dipolarity/polarizability ability. Therefore, the S descriptor has no influence in the K_{OC} value.

2.2. The d distance as a measure of similarity of systems

The d parameter [30] is a useful tool to measure the similarity between systems characterized through the solvation parameter model (Eq. (1)). It considers the five coefficients of a system as components of a vector in a five-dimensional space, whose length corresponds to

$$l = \sqrt{e^2 + s^2 + a^2 + b^2 + v^2} \quad (3)$$

The coefficients of the solvation parameter equation can be normalized as follows

$$e_u = \frac{e}{l} \quad (4)$$

$$s_u = \frac{s}{l} \quad (5)$$

$$a_u = \frac{a}{l} \quad (6)$$

$$b_u = \frac{b}{l} \quad (7)$$

$$v_u = \frac{v}{l} \quad (8)$$

where the u subscript indicates the normalization of the coefficients.

d measures the distance between the normalized unitary vectors of two systems. It can be calculated according to the following expression, which has been applied to the soil–water sorption system (*soil*) and a chromatographic one (*chrom*):

$$d = \sqrt{(e_{u\,soil} - e_{u\,chrom})^2 + (s_{u\,soil} - s_{u\,chrom})^2 + (a_{u\,soil} - a_{u\,chrom})^2 + (b_{u\,soil} - b_{u\,chrom})^2 + (v_{u\,soil} - v_{u\,chrom})^2} \quad (9)$$

The lower d is, the more similar the soil–water sorption and the chromatographic systems are.

According to their normalized coefficients, the soil–water sorption and the chromatographic systems can be expressed as

$$\log K_{OC} = c_{soil} + l_{soil}(e_{u\,soil}E + s_{u\,soil}S + a_{u\,soil}A + b_{u\,soil}B + v_{u\,soil}V) \quad (10)$$

$$\log k = c_{chrom} + l_{chrom}(e_{u\,chrom}E + s_{u\,chrom}S + a_{u\,chrom}A + b_{u\,chrom}B + v_{u\,chrom}V) \quad (11)$$

where $\log K_{OC}$ and $\log k$ correspond to the logarithm of the dependent solute property of each system, i.e. the soil–water partition coefficient normalized to the organic carbon content and the retention factor in the chromatographic system respectively. The d distance between both systems will be 0 if they have the same normalized vector. In this case, the chromatographic system may model the soil–water sorption process and the following linear correlation must hold:

$$\log K_{OC} = c_{soil} - \frac{l_{soil}}{l_{chrom}}c_{chrom} + \frac{l_{soil}}{l_{chrom}} \log k \quad (12)$$

The d distance is closely related to the angle θ_{ij} between the two normalized vectors, which was proposed by Ishihama and Asakawa [45] as a measure of the similarity between the two systems. Their relation [30] can be written as

$$\cos \theta_{ij} = \cos\left(2 \arcsin\left(\frac{d}{2}\right)\right) \quad (13)$$

2.3. Evaluation of the predictive ability of $\log K_{OC}$–$\log k$ correlation models

The performance of chromatographic systems to mimic soil–water sorption is in practice tested by correlating experimental data for a representative group of solutes:

$$\log K_{OC} = q + p \log k \quad (14)$$

where q is the experimental intercept and p is the experimental slope of the correlation. As the d distance (Eq. (9)) between the two systems comes close to zero, q and p should be closer to the intercept and slope predicted by Eq. (12), respectively.

The overall variance of the correlation determines the precision with which K_{OC} values can be estimated by means of chromatographic measurements. Therefore, the sources of variance should be examined in order to evaluate the goodness of the correlation. There are three sources of variance: the uncertainty of the soil sorption data (σ_{soil}), the uncertainty of the chromatographic data (σ_{chrom}), and the error due to the dissimilarity between the two systems (σ_d). These three contributions of variance were proven [31] to be easily estimated through the characterization of the two systems by means of the solvation parameter model, according to the following expressions:

$$\sigma^2_{soil} \approx SD^2_{soil} \quad (15)$$

$$\sigma^2_{chrom} \approx (p\,SD_{chrom})^2 \quad (16)$$

$$\sigma^2_{d} \approx SD^2_{d} \quad (17)$$

where SD_{soil} and SD_{chrom} correspond to the standard deviation of the soil–water sorption and chromatographic systems' characterization respectively, p is again the slope of the correlation, and SD_d is the standard deviation from the correlation between calculated values of $\log K_{OC}$ and $\log k$. The calculated values of a system can be obtained applying the corresponding Abraham equation to the descriptors of a set of solutes (usually the biological set).

Thus, the overall variance of the correlation (SD^2_{corr}) can be predicted as sum of these three contributions:

$$SD^2_{corr\,cal} = SD^2_{soil} + (p\,SD_{chrom})^2 + SD^2_d \quad (18)$$

The lower $SD^2_{corr\ cal}$ is, the better the performance of the chromatographic system to emulate the sorption of neutral organic compounds by soil from water is expected to be.

3. Experimental

3.1. Apparatus and conditions

3.1.1. HPLC systems

The retention data was measured in an IAM.PC.DD2 column (100 mm × 4.6 mm I.D., 12 μm particle size, Regis Technologies Inc., Morton Grove, IL, USA), XTerra MSC18 and XTerra RP18 columns (150 mm × 4.6 mm I.D., 5 μm particle size, Waters Corporation, Milford, MA, USA). All measurements were performed with a Shimadzu liquid chromatograph (Kyoto, Japan) equipped with two Shimadzu LC-20AD pumps and a Shimadzu SPD-M20A diode array detector. The temperature was controlled at 25.0 ± 0.1 °C with a Shimadzu CTO-10AS column oven.

The mobile phase was a mixture of 40% acetonitrile and 60% 10 mM phosphate aqueous buffer adjusted to pH 7. All solutes were dissolved in methanol and their concentrations were 0.1 mg mL^{-1}. The injection volume was 10 μL. Isocratic conditions were used at a flow rate of 1 mL min^{-1}. The column hold-up time was determined by using an aqueous solution of potassium bromide (0.1 mg mL^{-1}) as an unretained solute, whose detection was performed at 200 nm. All measurements were taken in triplicate.

3.1.2. MEKC systems

MEKC separations were performed with a Beckman P/ACE System 5500 Capillary Electrophoresis with a UV diode array detector. The fused-silica capillary was 40 cm effective length × 50 μm ID and was obtained from Composite Metal Services Ltd. (Shipley, West Yorkshire, UK). The capillary was activated by the following washing sequence: water (10 min), 1 M NaOH (10 min), water (5 min), 0.1 M NaOH (10 min), and separation solution (20 min). As daily conditioning, the capillary was flushed with water for 5 min, followed by 0.1 M NaOH for 5 min, water for 5 min and separation solution for 20 min. Prior to each separation the capillary was flushed with 0.1 M NaOH for 1 min and separation solution for 3 min. Retention measurements were made at 25 °C and +15 kV. Detection was at 214 nm.

The sodium taurocholate (STC) system was a separation solution 50 mM in taurocholic acid sodium salt hydrate and 20 mM in phosphate aqueous buffer adjusted to pH 7. The same buffer conditions were used to prepare the separation solution of the sodium dodecylsulfate (SDS) system, which was 40 mM in SDS. All solutes were dissolved in methanol (used as electroosmotic flow marker) and contained 2 mg mL^{-1} of dodecanophenone as micellar marker [46]. The concentration of the solutes was 2 mg mL^{-1}. All solutions were filtered through 0.45-μm nylon syringe filters obtained from Albet (Dassel, Germany). Samples were introduced into the capillary by applying a pressure of 0.5 p.s.i. for 1 s. All measurements were taken in triplicate.

3.2. Reagents and materials

Methanol (HPLC grade), hydrochloric acid (25% in water), sodium hydroxide (>99%), sodium dihydrogenphosphate monohydrate (>99%), and SDS (>99%) were from Merck (Darmstadt, Germany). Acetonitrile (HPLC grade) was from VWR International (West Chester, Pennsylvania, USA). Taurocholic acid sodium salt

Table 1
Description of several physicochemical systems discussed in this work.

Systems	Type of system[a]	Experimental conditions	Ref.
1 DPPG:DPPC:Chol	LEKC	10 mM dipalmitoylphosphatidyl glycerol (DPPG):dipalmitoylphosphatidyl choline (DPPC):cholesterol (Chol) (24:46:30), 25 mM HEPES pH 7.5	[47]
2 RP18	HPLC	Xterra RP18 column, 10 mM phosphate buffer pH 7:acetonitrile (60:40)	[30]
3 STC	MEKC	50 mM sodium taurocholate (STC), 20 mM phosphate pH 7	[48]
4 SDS microemulsion	MEEKC	1.4 wt.% sodium dodecyl sulfate (SDS), 6.49 wt.% butan-1-ol, 0.82 wt.% heptane, 500 mM phosphate + 100 mM borate pH 7	[49]
5 Pentanol–water	Solvent–water partition	Diverse	[50]
6 MSC18	HPLC	Xterra MSC18 column, 10 mM phosphate buffer pH 7:acetonitrile (60:40)	[30]
7 Octanol–water	Solvent–water partition	Diverse	[40]
8 IAM	HPLC	Immobilized artificial membrane prepared with phosphatidyl choline (IAM.PC.DD2) column, 10 mM phosphate buffer pH 7:acetonitrile (60:40)	[30]
9 SDS	MEKC	40 mM sodium dodecylsulfate (SDS), 20 mM phosphate pH 7	[51,52]
10 PAAU	Polymeric EKC	58 mM poly(sodium 11-acrylamidoundecanoate) (PAAU), 12.5 mM phosphate-tetraborate pH 9.2	[53]
11 SDS C18	MELC	AT. Chrom C18 column, 3.3% (w/w) sodium dodecylsulfate (SDS), 6.6% (w/w) butanol, 1.6% (w/w) heptane, 88.5% (w/w) 50 mM phosphate buffer pH 7	[54]
12 LPFOS	MEKC	40 mM lithium perfluorooctanesulfonate (LPFOS), 20 mM phosphate pH 7	[51,52]

[a]LEKC: liposome electrokinetic chromatography; HPLC: high-performance liquid chromatography; MEKC: micellar electrokinetic chromatography; MEEKC: microemulsion electrokinetic chromatography; MELC: microemulsion liquid chromatography.

hydrate (98%) was from Acros Organics (Geel, Belgium). Dodecanophenone (98%) was from Aldrich (Steinheim, Germany). Water was purified by a Milli-Q plus system from Millipore (Bedford, MA, USA), with a resistivity of 18.2 MΩ cm. The test solutes employed were reagent grade or better and obtained from several manufacturers (Merck, Sigma (Steinheim, Germany), Fluka (Steinheim, Germany), Aldrich, Carlo Erba (Milano, Italy), Baker (Deventer, Netherlands), Riedel-deHaën (Seelze, Germany)).

3.3. Calculation

Principal component analysis (PCA) was performed with Matlab package version R2007a from MathWorks (Natick, MA, USA). Microsoft Excel was used for data calculations and multiple linear regression analysis.

3.3.1. HPLC systems

The HPLC retention data was expressed by means of the retention factor, k:

$$k = \frac{(t_R - t_0)}{t_0} \quad (19)$$

where t_R corresponds to the solute retention time and t_0 is the column hold-up time determined by an aqueous potassium bromide solution.

3.3.2. MEKC systems

The MEKC retention factor, k, was calculated as follows:

$$k = \frac{(t_R - t_0)}{(1 - (t_R/t_m))t_0} \quad (20)$$

t_R is the solute migration time, t_0 corresponds to the migration time of methanol (used to determine the electroosmotic flow), and t_m is the migration time of dodecanophenone (used to determine the migration time of the micelles).

4. Results and discussion

4.1. Surrogate chromatographic systems to emulate soil–water sorption

Several physicochemical systems considered in this work are presented in Table 1, which includes some possible surrogate systems for the soil–water sorption process as well as other illustrative physicochemical systems. They were selected from our database, where many physicochemical systems characterized through the solvation parameter model are compiled. In fact, most systems of Table 1 were also studied in a previous publication about modeling nonspecific toxicity of organic compounds to the fathead minnow fish [32]. They consist of physicochemical systems of different nature, which basically correspond to systems of electrokinetic chromatography, high-performance liquid chromatography, and solvent–water partition. Table 2 shows the Abraham characterizations, normalized coefficients (Eqs. (4)–(8)), and l (Eq. (3)), d (Eq. (9)), and l_{soil}/l (Eq. (12)) parameters of the systems presented in Table 1.

The systems can be compared according to their normalized coefficients. Thus, we can easily observe that a general difference between the normalized coefficients of the soil–water sorption and the physicochemical systems falls on s_u, which is 0 for the soil–water whereas it is negative for the rest of systems. Especially, s_u is very negative for the SDS C18 MELC system (system 11), which also presents a huge negative a_u value in comparison to soil–water. In the same vein, the normalized coefficients of the LPFOS MEKC system (system 12) differ enormously from the ones of soil–water sorption. The most noticeable difference is found in b_u (far less negative), and there are also big differences in e_u (the only negative value of the table), a_u (quite more negative), and v_u (considerably more positive). The normalized coefficients of the other physicochemical systems compared to the ones of soil–water sorption do not present so pronounced differences. Even they are not too different among them. These observations completely agree with the d distance values shown in the Table, which are quite close for systems 1–10 but they become much larger for systems 11 and 12.

The l_{soil}/l parameter corresponds to the slope of Eq. (12). Therefore, it indicates the sensitivity that could be obtained in the soil–water sorption–physicochemical data correlations. The highest l_{soil}/l values belong to the systems of HPLC (systems 2, 6, 8), the PAAU poly-EKC (system 10), the SDS C18 MELC (system 11), and the LPFOS MEKC (system 12), meaning that they are expected to provide the highest sensitivity. On the contrary, the octanol–water partition system (system 7) presents the lowest l_{soil}/l and consequently the lowest expected sensitivity.

Nevertheless, neither the comparison of normalized coefficients nor the d parameter allow direct computation of the precision that could be achieved in the estimation of K_{OC} values by means of measurements in these physicochemical systems. For this reason, we predicted the contributions of variance that determine the overall error in soil–water–physicochemical systems' correlations (Section

Table 2
Characterizations by the solvation parameter model of the systems presented in Table 1 as well as the corresponding normalized coefficients, and l, d, and l_{soil}/l parameters.

Systems	Coefficients						Statistics				Normalized coefficients					l	d	l_{soil}/l
	c	e	s	a	b	v	n	r	SD	F	e_u	s_u	a_u	b_u	v_u			
Soil–water sorption	0.21	0.74	0.00	−0.31	−2.27	2.09	129	0.977	0.248	655	0.232	0.000	−0.097	−0.712	0.656	3.188	0.000	1.00
Physicochemical systems																		
1 DPPG::DPPC:Chol LEKC	−2.30	0.54	−0.55	0.32	−3.12	3.01	29	0.990	0.080	–	0.122	−0.125	0.072	−0.707	0.682	4.415	0.239	0.72
2 RP18 HPLC	−0.172	0.233	−0.460	−0.260	−1.574	1.582	55	0.996	0.056	1121	0.101	−0.200	−0.113	−0.683	0.686	2.305	0.243	1.38
3 STC MEKC	−1.838	0.613	−0.513	0.319	−2.396	2.390	56	0.984	0.112	305	0.175	−0.147	0.091	−0.686	0.684	3.492	0.249	0.91
4 SDS MEEKC	−1.133	0.279	−0.692	−0.060	−2.805	3.048	53	0.994	0.090	791	0.066	−0.164	−0.014	−0.667	0.724	4.214	0.260	0.76
5 Pentanol–water	0.185	0.367	−0.732	0.105	−3.100	3.395	106	0.988	0.180	823	0.079	−0.157	0.022	−0.664	0.727	4.671	0.264	0.68
6 MSC18 HPLC	−0.152	0.151	−0.463	−0.424	−1.719	1.759	55	0.994	0.072	764	0.059	−0.182	−0.167	−0.676	0.692	2.543	0.265	1.25
7 Octanol–water	0.088	0.562	−1.054	0.034	−3.460	3.814	613	0.997	0.116	23162	0.106	−0.199	0.006	−0.655	0.721	5.286	0.272	0.60
8 IAM HPLC	−0.710	0.292	−0.344	0.141	−1.193	1.161	51	0.991	0.047	480	0.169	−0.199	0.081	−0.689	0.671	1.731	0.276	1.84
9 SDS MEKC	−1.680	0.558	−0.596	−0.266	−1.674	2.717	63	0.991	0.120	618	0.169	−0.180	−0.080	−0.507	0.822	3.305	0.327	0.96
10 PAAU poly-EKC	−1.86	0.26	−0.16	−0.27	−1.05	2.11	18	0.997	0.051	403	0.109	−0.067	−0.113	−0.439	0.882	2.392	0.382	1.33
11 SDS C18 MELC	0.727	0.101	−0.645	−0.922	−0.951	0.901	26	0.990	0.078	205	0.058	−0.373	−0.533	−0.550	0.521	1.730	0.635	1.84
12 LPFOS MEKC	−1.410	−0.113	−0.243	−0.876	−0.455	1.966	62	0.970	0.190	180	−0.051	−0.110	−0.395	−0.205	0.887	2.216	0.701	1.44

2.3), since it is the key issue for the identification of possible surrogate systems. Table 3 shows all the predicted variances.

The error attributed to the precision of the chromatographic data (Eq. (16)) was predicted using the slope of the correlation between calculated values (p_{cal}) because the experimental one (p) is initially not known. Both the intercept and slope of the calculated correlations (q_{cal} and p_{cal}, respectively) are detailed in Table 3 too. The set of solutes used to obtain the calculated correlations corresponds to 124 solutes from the total of 129 considered in the soil–water sorption characterization (n_{soil}) [23]. The remaining 5 solutes (tetracene, benz[a]anthracene, 1,2,5,6-dibenzanthracene, benzo[a]pyrene, hexachlorobenzene) were excluded from these correlations because they had very high calculated values (both $\log K_{OC}$ and $\log k$) which led to a big deviation. Calculated $\log k$ values were especially high (mainly above 2, which cannot be considered very reliable) because of their large descriptors. Thus, one of the limitations of the presented approach may be that compounds with extreme solvation parameters can present significant deviations.

In Table 3, SD^2_{soil} is observed to be higher than any value of $(p_{cal}\ SD_{chrom})^2$, which reflects the larger experimental variability usually associated with biological properties because of their complex matrix. Since $(p_{cal}\ SD_{chrom})^2$ is small, we consider SD^2_d should be on the order of or lower than SD^2_{soil} to obtain correlation models with an acceptable predictive ability. Although the STC MEKC system (system 3) presents the only SD^2_d lower than SD^2_{soil}, systems 1, 8, and 9 have SD^2_d values very close to the variance of the original soil–water sorption data. Thus, the calculated overall variances for these systems ($SD^2_{corr\ cal}$) (Eq. (18)) are about twice the variance of the biological data and they are expected to mimic well the soil–water sorption process. Among these systems, we decided to test experimentally systems 3, 8, and 9, but not system 1 because working with liposomes is very costly and experimentally tricky. On the contrary, chromatographic systems are easy of use, although they present some limitations regarding the error associated to MEKC measurements in extreme regions of the elution window and the high retention time of some substances in HPLC.

The SD^2_d values of the other physicochemical systems already surpass the SD^2_{soil}. Undoubtedly, the huge SD^2_d and consequently $SD^2_{corr\ cal}$ of the SDS C18 MELC and LPFOS MEKC systems (systems 11 and 12, respectively) mean that they are too much different from soil–water sorption to provide good estimations of K_{OC} values. Nevertheless, the remaining systems (systems 2, 4, 5, 6, 7, 10) present a SD^2_d contribution not too greater than SD^2_{soil} and consequently they could achieve a fair performance in modeling soil–water sorption. Thus, we decided to test experimentally the HPLC systems of RP18 and MSC18 (systems 2, 6) due to the great interest in proving the feasibility of conventional C18 columns for the estimation of K_{OC} values. System 7 was also selected because the octanol–water partition system has been classically taken as reference in biological emulations [29]. Furthermore, it is a well-studied system [55] whose Abraham characterization includes many solutes [40] and experimental log P values can be obtained for most of them from a reliable database [56], i.e. we did not have to measure octanol–water partitions. In fact, we were reluctant to test experimentally the pentanol–water partition (system 5) because partition experiments are very tedious and time-consuming. Systems 4 and 10 were not selected because other systems of electrokinetic chromatography with better predicted performances (lower $SD^2_{corr\ cal}$) were already chosen.

The comparison of the d distances with the SD^2_d values of Table 3 highlights the usefulness of the variance predictions. Since both d and SD^2_d indicate the similarity between the soil–water sorption and each physicochemical system, d and SD^2_d should increase in parallel as the pairs of systems become more dissimilar. However,

Table 3
Prediction of the variance contributions that determine the overall error in soil–water sorption vs. chromatographic data correlations for the considered systems.

Systems	Soil–water sorption				$SD^2_{soil} = 0.062$	$n_{soil} = 129$
	d	q_{cal}	p_{cal}	$(p_{cal}\ SD_{chrom})^2$	SD^2_d	$SD^2_{corr\ cal}$
1 DPPG:DPPC:Chol LEKC	0.239	2.32	0.87	0.005	0.073	0.139
2 RP18 HPLC	0.243	1.20	1.47	0.007	0.151	0.220
3 STC MEKC	0.249	2.19	1.06	0.014	0.052	0.128
4 SDS MEEKC	0.260	1.56	0.88	0.006	0.140	0.208
5 Pentanol–water	0.264	0.27	0.81	0.021	0.120	0.203
6 MSC18 HPLC	0.265	1.26	1.29	0.009	0.195	0.265
7 Octanol–water	0.272	0.48	0.69	0.006	0.123	0.191
8 IAM HPLC	0.276	2.02	2.14	0.010	0.084	0.156
9 SDS MEKC	0.327	1.70	1.01	0.015	0.067	0.143
10 PAAU poly-EKC	0.382	2.18	1.35	0.005	0.101	0.167
11 SDS C18 MELC	0.635	1.53	1.11	0.008	0.524	0.594
12 LPFOS MEKC	0.701	2.26	1.18	0.050	0.383	0.494

this behaviour is not strictly observed in the table mainly because most d values are too close to reflect the differences among the systems. Especially in these cases, the predicted SD^2_d values allow us to fine-tune the selection of possible surrogate systems.

4.2. Suitability of the selected chromatographic systems to estimate K_{OC} values

Experimental $\log K_{OC}$ values [23] have been correlated against experimental $\log k$ data from the chosen chromatographic systems (Eq. (14)) in order to prove their feasibility for emulating soil–water sorption.

The solutes to test and correlate experimentally were previously examined by PCA. The solutes with known $\log K_{OC}$ [23] distributed according to the scores plot of the two main PCs (PC1 and PC2, which capture the 96% of the total data variance) are shown in Fig. 1. The solutes used in the characterizations of the chromatographic systems are also represented in the scores plots of Fig. 1. Thus, the solutes to test and correlate experimentally

Table 4
Experimental $\log k$ data determined for the selected chromatographic systems and correlated with $\log K_{OC}$ values.

Solute	HPLC systems			MEKC systems	
	RP18	MSC18	IAM	STC	SDS
1,2-Dichlorobenzene	1.072	1.206	0.380	0.694	0.988
1,2-Dimethylbenzene	1.000	1.184	0.211	0.411	0.758
1,3-Dimethylbenzene	1.038	1.233	0.235	0.427	0.816
1,4-Dichlorobenzene	1.121	1.260	0.346	0.581	0.952
1,4-Dimethylbenzene	1.040	1.254	0.235	0.427	0.801
1-Butanol	–	–	–	−1.296	−0.703
1-Pentanol	–	–	–	−0.956	−0.199
1-Propanol	–	–	–	−1.533	–
2,4-Dichlorophenol	0.788	0.670	0.373	0.572	0.793
2-Chlorophenol	0.459	0.362	0.002	−0.167	0.103
3-Chlorophenol	0.555	0.417	0.106	0.095	0.252
4-Bromophenol	0.602	0.492	0.180	0.266	0.442
4-Chloroaniline	0.454	0.426	−0.088	−0.174	0.214
4-Methylaniline	0.234	0.260	−0.245	−0.639	0.176
Acetanilide	−0.008	−0.085	−0.400	−0.597	−0.191
Acetophenone	0.354	0.376	−0.287	−0.557	0.111
Acridine	0.682[a]	0.718[a]	0.225[a]	0.526[a]	1.834
Aniline	0.104	0.107	−0.372	−0.930	−0.384
Anisole	0.611	0.691	−0.105	−0.352	0.128
Anthracene	1.536	1.655	0.823	1.350	2.399
Benzamide	−0.276	−0.346	−0.535	−0.753	−0.372
Benzene	0.636	0.737	−0.080	−0.489	−0.110
Benzophenone	1.000	1.062	0.198	0.557	1.284
Biphenyl	1.348	1.469	0.495	1.066	1.608
Bromobenzene	0.971	1.063	0.209	0.327	0.631
Butylbenzene	1.490	1.742	0.603	1.444	1.726
Chlorobenzene	0.904	0.995	0.137	0.144	0.468
Diethylphthalate	0.764	0.851	−0.105	−0.084	1.130[a]
Dimethylphthalate	0.405	0.434	−0.323	−0.441	0.501
Ethyl 4-hydroxybenzoate	0.413	0.285	−0.048	0.089	0.595
Ethylbenzene	1.070	1.224	0.238	0.387	0.744
Methyl benzoate	0.682	0.610	−0.135	−0.256	0.430
Naphthalene	1.059	1.137	0.324	0.665	1.064
Nitrobenzene	0.544	0.551	−0.121	−0.463	0.020
N-methylaniline	0.451	0.457	−0.169	−0.667	−0.011
N,N-dimethylaniline	0.725	0.815	0.007	−0.258	0.372
Phenantrene	1.486	1.595	0.777	1.529	2.263
Phenol	0.189	0.117	−0.241	−0.663	−0.390
Phenylurea	−0.245	−0.309	−0.418	−0.648	−0.202
Propylbenzene	1.264	1.496	0.410	0.810	1.231
Pyrene	1.710	1.820	1.053	1.712	–
Toluene	0.861	0.983	0.086	−0.019	0.342

[a] Solutes excluded from the correlations because they were considered outliers (with a standard residual higher than |2.5|).

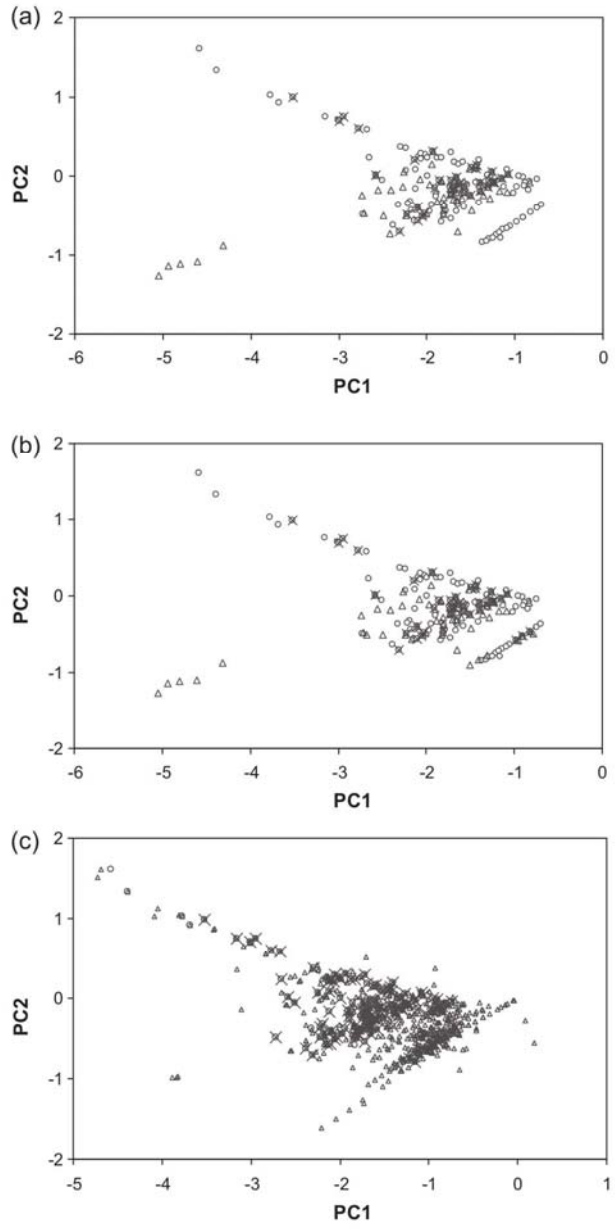


Fig. 1. Scores plot after performing PCA to the solute descriptors considered in the soil–water sorption characterization (○), and projection of the solute descriptors included in the characterization of the selected chromatographic systems (Δ) and in the experimental correlations (×) for: (a) HPLC systems, (b) MEKC systems, (c) octanol–water partition system.

are observed to cover well the chemical space where most compounds from the soil–water and the chromatographic systems are clustered. The log K_{OC} solutes in the left-upper quadrant are those excluded from the calculated correlations in the previous section (tetracene, benz[a]anthracene, 1,2,5,6-dibenzanthracene, benzo[a]pyrene, hexachlorobenzene). Their large descriptor values make them move away from the rest of substances in the scores plot, being also responsible for the big deviation of their calculated values from the calculated values of the rest of solutes. Thus, when the large descriptors of a solute make the studied property too high (e.g. log $k > 2$, which would correspond to a very retained substance and therefore measured with a considerable error), it is advisable not to take the solute into account. Accordingly, the exclusion of tetracene, benz[a]anthracene, 1,2,5,6-dibenzanthracene, benzo[a]pyrene and hexachlorobenzene from the set of solutes to test and correlate experimentally is justified. The right-bottom quadrant of Fig. 1a also shows substances from the soil–water sorption characterization that have not been experimentally tested and correlated because they were not appropriate for the UV detection in the chromatographic systems. Last but not least, the five solutes shown in the left-bottom quadrant of Fig. 1a and b are corticosteroids that were used in the chromatographic characterizations. They are also isolated from the rest of solutes because of their huge descriptor values ($E \approx 2$, $S > 3$, $V > 2$, whereas these descriptors are around 1 for the rest of solutes of the chromatographic characterizations; $B > 1$, whereas it is around 0.5 for most of the solutes that are not corticosteroids). No values of log K_{OC} are known for these corticosteroids. All in all, the solutes chosen to test and correlate experimentally provide a representative set.

The experimental log k data of the solutes tested in the selected chromatographic systems is presented in Table 4. The experimental log P values were obtained from a reliable database [56], as it was previously mentioned.

Table 5 collects the results of the experimental correlations, which are illustrated in Fig. 2. Apart from the experimental points, Fig. 2 also shows our predictions by means of a 95% of confidence range ($\pm 2\ SD_{corr\ cal}$).

The IAM HPLC system provides the best correlation model according to its lowest experimental variance ($SD^2_{corr\ exp}$) and the little dispersion observed in the experimental points of Fig. 2e. The soil–water sorption data is also well modeled by means of chromatographic measurements in the MEKC systems of STC and SDS (Fig. 2b and f) since the $SD^2_{corr\ exp}$ of their correlations is still quite low. In fact, the experimental results (Table 5) prove that all three systems (IAM HPLC, STC MEKC, SDS MEKC) are the best choice to model soil–water sorption, which is in agreeement with our predictions ($SD^2_{corr\ cal}$ of Table 3). Thus, the fact that the SD^2_d predicted for these systems are similar to the log K_{OC} uncertainty ($SD^2_d \sim SD^2_{soil}$) implies that their chromatographic measurements are expected to estimate well log K_{OC} values, which is demonstrated by the low overall variance obtained experimentally ($SD^2_{corr\ exp}$).

The performance achieved by the other tested systems is poorer. Our predictions already show that the RP18 HPLC, MSC18 HPLC, and octanol–water partition systems are more different to the soil–water sorption process. Their dissimilarity (estimated by SD^2_d in Table 3) causes the increase of the overall error of their correlations, as we predicted by $SD^2_{corr\ cal}$ and we prove experimentally by their $SD^2_{corr\ exp}$. Thus, the experimental points appear more dispersed in Fig. 2a, c, and d. The HPLC systems based on C18 columns are shown to be the least suited for surrogating soil–water sorption, especially the MSC18 HPLC system.

It is worth noting that almost all the experimental points of the tested systems are within the dotted lines that cover the confidence range of our predictions. This behaviour is observed regardless both the number of data (there are many more points in Fig. 2d) and the different ability of these systems to mimic soil–water sorption.

4.3. Herbicides test

The more precise correlation models obtained in the previous section were used to estimate the log K_{OC} values of several herbicides, which are agrochemical compounds of potential environmental interest. Their log K_{OC} values were also estimated using the method proposed by Poole and Poole [23], i.e. calculating the log K_{OC} by the solvation parameter model equation (Eq. (2)) for the compounds with known descriptors. Thus, we could compare the

Table 5
Results from the experimental correlations between log K_{OC} and log k values for the selected chromatographic systems (standard deviations in brackets).

Systems	Soil–water sorption					
	q	p	$SD^2_{corr\ exp}$	n	r	F
2 RP18 HPLC	1.30 (0.13)	1.48 (0.15)	0.188	38	0.856	99
3 STC MEKC	2.23 (0.05)	1.14 (0.07)	0.113	41	0.935	270
6 MSC18 HPLC	1.46 (0.14)	1.19 (0.15)	0.245	38	0.808	68
7 Octanol–water	0.46 (0.08)	0.71 (0.03)	0.150	114	0.922	633
8 IAM HPLC	2.20 (0.05)	2.17 (0.12)	0.075	38	0.945	303
9 SDS MEKC	1.72 (0.07)	1.08 (0.07)	0.111	39	0.923	213

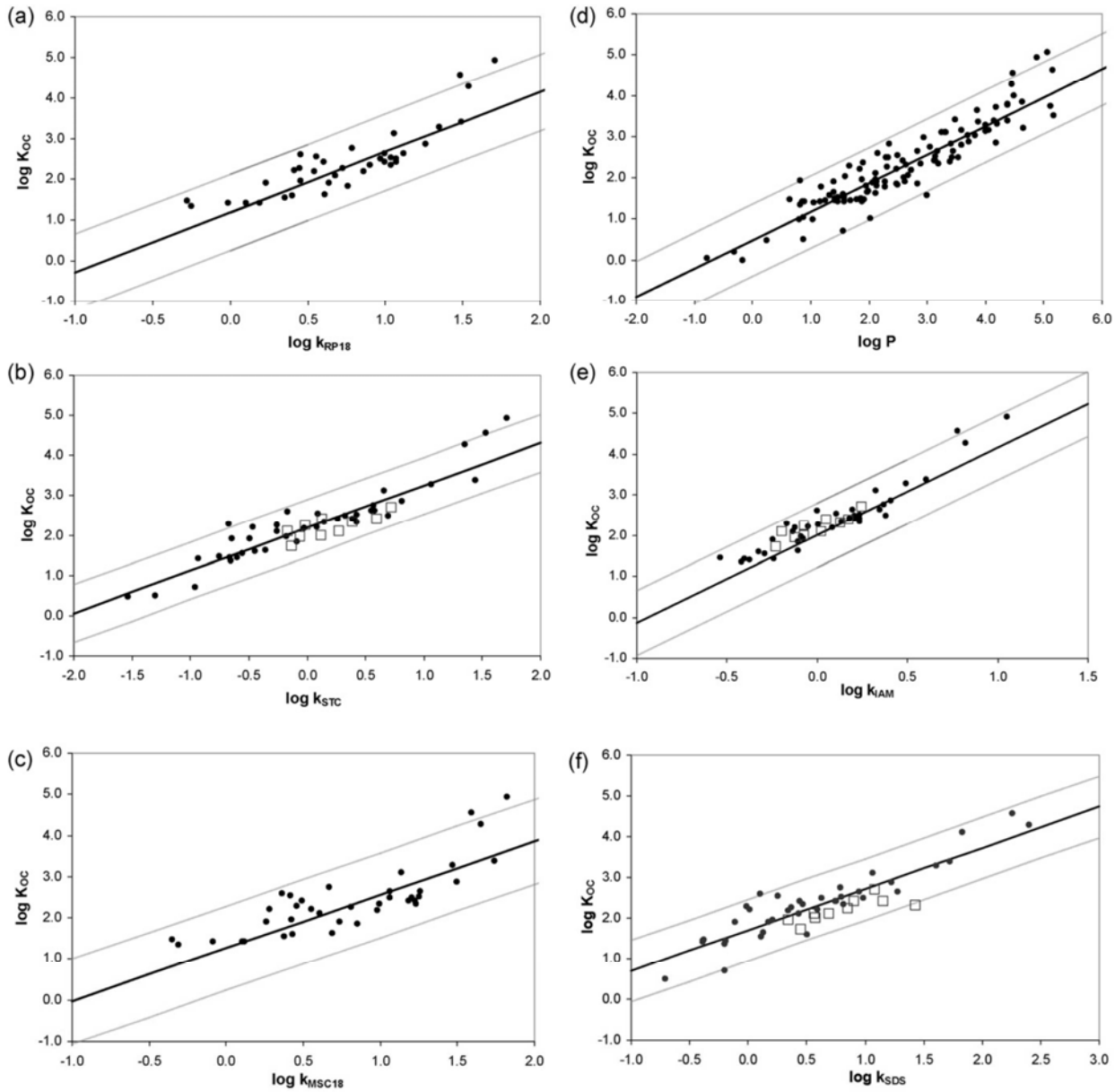


Fig. 2. Plots of the experimental correlations between log K_{OC} and log k data for the tested chromatographic systems: (a) RP18 HPLC, (b) STC MEKC, (c) MSC18 HPLC, (d) octanol–water, (e) IAM HPLC, and (f) SDS MEKC. The correlations between values calculated through the Abraham characterizations are also shown (—), as well as the corresponding confidence range for 95% confidence level (± 2 $SD_{corr\ cal}$) (.....). The experimental points for several tested pesticides (□) are included in (b) STC MEKC, (e) IAM HPLC, and (f) SDS MEKC.

Table 6
Experimental and predicted log K_{OC} values for the tested herbicides.

Tested herbicides	log K_{OC} exp	log K_{OC} pred			
		IAM HPLC	STC MEKC	SDS MEKC	Solvation parameter model
Monuron	1.95	1.93	2.17	2.09	2.22
Diuron	2.40	2.57	2.92	2.70	2.73
Linuron	2.70	2.73	3.06	2.89	–
Fluometuron	2.00	2.04	2.37	2.34	2.01
Metobromuron	2.10	2.25	2.54	2.47	–
Metoxuron	1.72	1.71	2.08	2.21	2.10
Simazine	2.10	1.78	2.04	2.33	1.95
Propazine	2.40	2.31	2.38	2.96	2.35
Atrazine	2.24	2.05	2.22	2.64	2.15
Terbuthylazine	2.32	2.48	2.68	3.26	2.66

K_{OC} data estimated by the different approaches with the experimental K_{OC} data of these herbicides.

The tested herbicides were monuron, diuron, linuron, fluometuron, metobromuron, metoxuron, simazine, propazine, atrazine, and terbuthylazine. They belong to two different families: phenyureas and triazines. Their experimental log K_{OC} values are well known and were taken from literature [20]. They are collected in Table 6. The log k values of these herbicides were determined experimentally in the chromatographic systems which present the best performances in modeling soil–water sorption: the IAM HPLC, the STC MEKC, and the SDS MEKC.

Both experimental log K_{OC} and log k values have been included in Fig. 2 according to the chromatographic system considered, i.e. Fig. 2b, e, f. As it is shown, the data points of the tested herbicides fall into the 95% confidence range of our predictions of precision.

The log K_{OC} values of these herbicides were afterwards predicted by means of the experimental chromatographic data determined in each system, just applying the corresponding correlation model of the system (Table 5). The results are presented in Table 6. This Table also includes the log K_{OC} values estimated by means of the solvation parameter model equation, i.e. multiplying the coefficients of Eq. (2) by the descriptors of the herbicides (we were not able to estimate the log K_{OC} of linuron and metobromuron by this way because their solvation descriptors are not known). Table 6 shows that the closest estimations of log K_{OC} to the experimental values are provided by the IAM HPLC system, which agrees with the best precision of this system (Table 5) as it was shown in the previous section.

The predicted against the experimental log K_{OC} values of the tested herbicides result in the following correlation equations (for values predicted by means of the IAM HPLC, the STC MEKC, and the SDS MEKC systems, as well as the solvation parameter model, respectively):

$$\log K_{OC}\,\text{pred}_{IAM} = -0.20(\pm 0.44) + 1.09(\pm 0.20)\log K_{OC}\,\text{exp} \quad (SD = 0.167) \tag{21}$$

$$\log K_{OC}\,\text{pred}_{STC} = 0.26(\pm 0.59) + 0.99(\pm 0.27)\log K_{OC}\,\text{exp} \quad (SD = 0.224) \tag{22}$$

$$\log K_{OC}\,\text{pred}_{SDS} = 0.31\,(\pm 0.65) + 1.04(\pm 0.30)\log K_{OC}\,\text{exp} \quad (SD = 0.247) \tag{23}$$

$$\log K_{OC}\,\text{pred}_{\text{solvation model}} = 0.57(\pm 0.79) + 0.80(\pm 0.37)\log K_{OC}\,\text{exp} \quad (SD = 0.235) \tag{24}$$

The error associated to each intercept and slope is presented in brackets, as well as the standard deviation of each correlation equation. The slopes are not significantly different from 1 and the intercepts from 0 for a 95% confidence level, meaning that the log K_{OC} values are well predicted. They are especially well predicted by the correlation models of the surrogated chromatographic systems, since the errors of the intercepts and slopes of Eqs. ((21)–(23)) are lower than the ones for prediction by the solvation parameter model (Eq. (24)). Therefore, correlation with log k for the proposed surrogate systems of the IAM HPLC, STC MEKC and SDS MEKC systems provide even better log K_{OC} predictions than calculation from the solute descriptors by the solvation parameter model for soil–water sorption. Last but not least, a further advantage is that solvation descriptors are not required, making our proposed method applicable to solutes with unknown solute descriptors.

5. Conclusions

Chromatographic systems provide a fast, economic, and reliable alternative to the direct determination of K_{OC} values. The predictive ability of models based on the correlation between soil–water sorption and chromatographic data should be assessed to find surrogate chromatographic systems of the soil–water sorption process. The overall variance of the correlation is shown to be well predicted by addition of the error attributed to the soil–water sorption data, the error of the chromatographic data, and the error due to the dissimilarity between the two correlated systems, which are estimated by means of the solvation parameter model approach. Several chromatographic systems chosen by this procedure have been finally tested experimentally. Both predictions and experimental correlations show that the IAM HPLC system and the MEKC systems of STC and SDS are the most suitable for the estimation of K_{OC} values. Several herbicides have been tested in these systems calculating their K_{OC} values from their log k retention with the expected precision. The precision is better than that obtained by estimating K_{OC} from the solvation parameter descriptors of the solutes and allows determination of K_{OC} data from simple HPLC or MEKC measurements when the descriptors are unknown. Less precise K_{OC} estimations are obtained through measurements in the octanol–water partition system and above all the conventional HPLC systems based on C18 columns, which present the weakest performance.

Acknowledgements

The authors thank the Ministerio de Ciencia e Innovación of the Spanish Government and the Fondo Europeo de Desarrollo Regional (FEDER) of the European Union (project CTQ2010-19217/BQU) for financial support. Furthermore, M. H-R. thanks the Ministerio de Educación of the Spanish Government for a supporting scholarship (AP2007-01109).

References

[1] B.L. Sawhney, K.K. Brown, Reactions and Movement of Organic Chemicals in Soils, Soil Science Society of America, Madison, 1989.
[2] R.J. Wagenet, P. Baveye, B.A. Stewart, Interacting Processes in Soil Science, Lewis Publishers CRC Press Inc., Boca Raton, 1992.
[3] R.J. Watts, Hazardous Wastes: Sources, Pathways, Receptors, John Wiley & Sons, New York, 1998.
[4] C.T. Chiou, Partition and Adsorption of Organic Contaminants in Environmental Systems, John Wiley & Sons, New York, 2002.
[5] W.J. Doucette, Environ. Toxicol. Chem. 22 (2003) 1771.

[6] G. Bronner, K. Goss, Environ. Sci. Technol. 45 (2011) 1307.
[7] American Society for Testing and Materials, Standard Test Method for Determining a Sorption Constant (K_{OC}) for an Organic Chemical in Soil and Sediments, 11.05, Annual Book of ASTM Standards, Philadelphia, 1993, p. 731.
[8] Organization for Economic Cooperation and Development, OECD Guidelines for the Testing of Chemicals, Paris, 1993, p. 992.
[9] B.M. Gawlik, N. Sotiriou, E.A. Feicht, S. Schulte-Hostede, A. Kettrup, Chemosphere 34 (1997) 2525.
[10] American Society for Testing and Materials, Standard Test Method for 24-h Batch-Type Measurement of Containment Sorption by Soils and Sediment, 11.04, Annual Book of Standards, Philadelphia, 2001, p. 120.
[11] M.L. Brusseau, R.E. Jessup, P.S.C. Rao, Environ. Sci. Technol. 24 (1990) 727.
[12] L.S. Lee, P.S.C. Rao, M.L. Brusseau, Environ. Sci. Technol. 25 (1991) 722.
[13] H. De Jonge, T.J. Heimovaara, J.M. Verstraten, Soil Sci. Soc. Am. J. 63 (1999) 297.
[14] Z. Gerstl, J. Contam. Hydrol. 6 (1990) 357.
[15] B.M. Gawlik, E.A. Feicht, W. Karcher, A. Kettrup, H. Muntau, Chemosphere 36 (1998) 2903.
[16] G. Szabó, J. Guczi, W. Kördel, A. Zsolnay, V. Major, P. Keresztes, Chemosphere 39 (1999) 431.
[17] J.M. Bermúdez-Saldaña, L. Escuder-Gilabert, M.J. Medina-Hernández, R.M. Villanueva-Camañas, S. Sagrado, Trends Anal. Chem. 25 (2006) 122.
[18] J.R. Baker, J.R. Mihelcic, A. Sabljic, Chemosphere 45 (2001) 213.
[19] J. Huuskonen, J. Chem. Inf. Comput. Sci. 43 (2003) 1457.
[20] I. Kahn, D. Fara, M. Karelson, U. Maran, P.L. Andersson, J. Chem. Inf. Comput. Sci. 45 (2005) 94.
[21] M.H. Abraham, Chem. Soc. Rev. 22 (1993) 73.
[22] S.K. Poole, C.F. Poole, Anal. Commun. 33 (1996) 417.
[23] S.K. Poole, C.F. Poole, J. Chromatogr. A 845 (1999) 381.
[24] M.H. Abraham, J.M.R. Gola, R. Kumarsingh, J.E. Cometto-Muniz, W.S. Cain, J. Chromatogr. B 745 (2000) 103.
[25] K. Goss, R.P. Schwarzenbach, Environ. Sci. Technol. 35 (2001) 1.
[26] T.H. Nguyen, K. Goss, W.P. Ball, Environ. Sci. Technol. 39 (2005) 913.
[27] C.F. Poole, A.D. Gunatilleka, S.K. Poole, Adv. Chromatogr. 40 (2000) 159.
[28] D.M. Cimpean, C.F. Poole, Analyst 127 (2002) 724.
[29] R.N. Smith, C. Hansch, M.M. Ames, J. Pharm. Sci. 64 (1975) 599.
[30] E. Lázaro, C. Ràfols, M.H. Abraham, M. Rosés, J. Med. Chem. 49 (2006) 4861.
[31] M. Hidalgo-Rodríguez, E. Fuguet, C. Ràfols, M. Rosés, Anal. Chem. 82 (2010) 10236.
[32] M. Hidalgo-Rodríguez, E. Fuguet, C. Ràfols, M. Rosés, Anal. Chem. 84 (2012) 3446.
[33] M.H. Abraham, H.S. Chadha, F. Martins, R.C. Mitchell, M.W. Bradbury, J.A. Gratton, Pestic. Sci. 55 (1999) 78.
[34] J.A. Platts, M.H. Abraham, Y.H. Zhao, A. Hersey, L. Ijaz, D. Butina, Eur. J. Med. Chem. 36 (2001) 719.
[35] M.H. Abraham, Y.H. Zhao, J. Le, A. Hersey, C.N. Luscombe, D.P. Reynolds, G. Beck, B. Sherborne, I. Cooper, Eur. J. Med. Chem. 37 (2002) 595.
[36] M.H. Abraham, F. Martins, J. Pharm. Sci. 93 (2004) 1508.
[37] M.H. Abraham, C. Ràfols, J. Chem. Soc. Perkin Trans. 2 (1995) 1843.
[38] K.R. Hoover, W.E. Acree Jr., M.H. Abraham, Chem. Res. Toxicol. 18 (2005) 1497.
[39] K.R. Hoover, K.B. Flanagan, W.E. Acree Jr., M.H. Abraham, J. Environ. Eng. Sci. 6 (2007) 165.
[40] M.H. Abraham, H.S. Chadha, G.S. Whiting, R.C. Mitchell, J. Pharm. Sci. 83 (1994) 1085.
[41] M.H. Abraham, M. Rosés, C.F. Poole, S.K. Poole, J. Phys. Org. Chem. 10 (1997) 358.
[42] C. Lepont, C.F. Poole, J. Chromatogr. A 946 (2002) 107.
[43] E. Fuguet, C. Ràfols, E. Bosch, M.H. Abraham, M. Rosés, Electrophoresis 27 (2006) 1900.
[44] M. Vitha, P.W. Carr, J. Chromatogr. A 1126 (2006) 143.
[45] Y. Ishihama, N. Asakawa, J. Pharm. Sci. 88 (1999) 1305.
[46] E. Fuguet, C. Ràfols, E. Bosch, M. Rosés, Electrophoresis 23 (2002) 56.
[47] S.T. Burns, A.A. Agbodjan, M.G. Khaledi, J. Chromatogr. A 973 (2002) 167.
[48] M. Hidalgo-Rodríguez, P. Rivas, E. Fuguet, C. Ràfols, M. Rosés, Manuscript in preparation.
[49] M.H. Abraham, C. Treiner, M. Rosés, C. Ràfols, Y. Ishihama, J. Chromatogr. A 752 (1996) 243.
[50] M.H. Abraham, A. Nasezadeh, W.E. Acree, Ind. Eng. Chem. Res. 47 (2008) 3990.
[51] E. Fuguet, C. Ràfols, E. Bosch, M.H. Abraham, M. Rosés, J. Chromatogr. A 942 (2002) 237.
[52] E. Fuguet, C. Ràfols, E. Bosch, M.H. Abraham, M. Rosés, J. Chromatogr. A 1216 (2009) 6877.
[53] C. Fujimoto, Electrophoresis 22 (2001) 1322.
[54] J. Liu, J. Sun, Y. Wang, X. Liu, Y. Sun, H. Xu, Z. He, J. Chromatogr. A 1164 (2007) 129.
[55] J. Sangster, Octanol–Water Partition Coefficients: Fundamentals and Physical Chemistry, Wiley, Chichester, 1997.
[56] Bio-Loom online database: <http://www.biobyte.com>.

ARTICLE IN PRESS

European Journal of Pharmaceutical Sciences xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

European Journal of Pharmaceutical Sciences

journal homepage: www.elsevier.com/locate/ejps

Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds

Marta Hidalgo-Rodríguez, Sara Soriano-Meseguer, Elisabet Fuguet, Clara Ràfols, Martí Rosés *

Departament de Química Analítica and Institut de Biomedicina de la Universitat de Barcelona, Universitat de Barcelona, Martí i Franquès, 1-11, E-08028 Barcelona, Spain

ARTICLE INFO

Article history:
Received 16 October 2012
Received in revised form 22 March 2013
Accepted 9 April 2013
Available online xxxx

Keywords:
Skin permeation prediction
High-performance liquid chromatography
Micellar electrokinetic chromatography
Octanol–water partition
Quantitative structure-activity relationships

ABSTRACT

Several chromatographic systems (three systems of high-performance liquid chromatography and two micellar electrokinetic chromatography systems) besides the reference octanol–water partition system are evaluated by a systematic procedure previously proposed in order to know their ability to model human skin permeation. The precision achieved when skin-water permeability coefficients are correlated against chromatographic retention factors is predicted within the framework of the solvation parameter model. It consists in estimating the contribution of error due to the biological and chromatographic data, as well as the error coming from the dissimilarity between the human skin permeation and the chromatographic systems. Both predictions and experimental tests show that all correlations are greatly affected by the considerable uncertainty of the skin permeability data and the error associated to the dissimilarity between the systems. Correlations with much better predictive abilities are achieved when the volume of the solute is used as additional variable, which illustrates the main roles of both lipophilicity and size of the solute to penetrate through the skin. In this way, the considered systems are able to give precise estimations of human skin permeability coefficients. In particular, the HPLC systems with common C18 columns provide the best performances in emulating the permeation of neutral compounds from aqueous solution through the human skin. As a result, a methodology based on easy, fast, and economical HPLC measurements in a common C18 column has been developed. After a validation based on training and test sets, the method has been applied with good results to the estimation of skin permeation of several hormones and pesticides.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Dermal absorption (Hoang, 1992; Kielhorn et al., 2006) is a global term that refers to the transport of chemicals from the outer surface of the skin into the skin and the systemic circulation. It is a key process in drug delivery studies of pharmaceutical and cosmetic interest, as well as in the fields of dermal toxicology, risk assessment and occupational exposure of environmental pollutants.

Commonly, the outermost layer of the skin, the stratum corneum, controls dermal absorption. Substances must initially permeate through the stratum corneum to be absorbed into the body following dermal exposure. Hence it is called the rate-determining barrier. This process can only occur by passive diffusion. The passage through the remaining sublayers of the epidermis and dermis ends up with the entrance of the substances into the blood capillaries.

Dermal absorption from exposure to chemicals in aqueous solution is typically described by skin-water permeability coefficients (K_p). The K_p of a solute is actually the linear velocity (cm s^{-1}) that describes its movement through the skin, which is related to the partition coefficient of the solute in the membrane (K_m), its membrane diffusion coefficient (D_m), and also the thickness of the membrane (h) (Lieb and Stein, 1986):

$$K_p = \frac{K_m D_m}{h} \quad (1)$$

Skin-water permeability coefficients, K_p, can be determined using a wide range of experimental techniques, both *in vivo* and *in vitro* experiments (Kielhorn et al., 2006). Over the years, a large number of K_p data has been generated and several compilations are available (Flynn, 1990; Johnson et al., 1995; Vecchia and Bunge, 2003; EDETOX). However, the measurement of dermal absorption of chemicals is laborious, costly and time-consuming, besides the existence of conflicting experimental protocols, the ethical difficulties with regard to human and animal testing, and the increasing legislation. Hence, there is a need for methods capable of predicting K_p values in a simple, economic, fast and ethical way.

Two general types of models, theoretical and empirical, can be developed to predict skin-water permeability coefficients of chemicals. Both types of models are complementary. They not only

* Corresponding author. Tel.: +34 93 403 92 75; fax: +34 93 402 12 33.
E-mail address: marti.roses@ub.edu (M. Rosés).

0928-0987/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejps.2013.04.005

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

allow the estimation of K_p values for untested chemicals, reducing in this way the need for *in vivo* and *in vitro* testing, but also provide insights into the permeation mechanism.

Theoretical models are mechanistic approaches that attempt to simulate the dynamics involved in the skin permeation process. They consist in mathematical transport models that relate dermal absorption to parameters like diffusion coefficients and partition coefficients, taking into account factors such as possible routes of penetration and interactions between the permeating chemicals and the skin constituents (Mitragotri et al., 2011; Tezel and Mitragotri, 2003; Wilschut et al., 1995).

Empirical models are based on physicochemical parameters and they relate K_p values to physicochemical descriptors of chemicals, for example descriptors of hydrophobicity, polarity, solubility, hydrogen bonding or size (Geinoz et al., 2004). In general, empirical models for skin permeation usually reveal the distinct important role of lipophilicity and molecular size (Patel et al., 2002). Many scientists have constructed quantitative structure-activity relationships (QSARs) (Hansch and Leo, 1995) as empirical models for describing human skin permeation. Several overviews of the main developed QSARs can be found in the literature (Cronin, 2006; Fitzpatrick et al., 2004, 2008; Geinoz et al., 2004; Maibach, 2007; Moss et al., 2002; Walker et al., 2003). Although most QSARs employ the multiple linear regression analysis, also nonlinear algorithms have been used to establish skin permeability models (Neely et al., 2009), especially artificial neural networks (Chen et al., 2007; Degim et al., 2003). In the last years, the developed computational models based on calculated descriptors (Buchwald and Bodor, 2001; Neely et al., 2009; Neumann et al., 2006) have become very helpful and powerful approaches to provide approximate estimations of K_p values. However, skin permeation models based on experimental descriptors are still generally preferred when more accuracy and reliability are required in the prediction of skin-water permeabilities of chemicals.

The solvation parameter model proposed by Abraham (1993) and Abraham et al. (1999) has been one of the QSARs applied to the human skin permeation process (Abraham et al., 1995b; Abraham et al., 1997, 1999; Abraham and Martins, 2004; Zhang et al., 2012). This model is very useful to characterize any partitioning system ruled by the passive transport of solutes between two phases. Thus, Abraham and Martins (2004) proposed the following equation for characterizing the permeation of neutral solutes from aqueous solution through the human skin after analyzing an extensive and carefully examined data set:

$$\log K_p = -5.246 - 0.106E - 0.473S - 0.473A - 3.000B + 2.296V \quad (2)$$

whose statistics were $n = 119$, $r = 0.912$, $SD = 0.461$, $F = 112$. This equation was obtained through the analysis of 119 compounds of different chemical nature, which covered a wide range of $\log K_p$ values. In that work (Abraham and Martins, 2004), the reliability of Eq. (2) was demonstrated by dividing the 119 compounds in random training and test sets, which were used to establish the equation and predict log K_p values successfully. In Eq. (2), the dependent solute property, $\log K_p$ (with K_p in cm s^{-1} and referred to the permeation of neutral solutes at 37 °C), is expressed as linear combination of several solute descriptors. The E, S, A, B, and V independent variables are the solute descriptors proposed by Abraham: E represents the excess molar refraction, S is the solute dipolarity/polarizability, A and B are the solute's effective hydrogen-bond acidity and hydrogen-bond basicity, respectively, and V is McGowan's solute volume. Abraham et al. (2004) constructed the available database of descriptor values, which collects 4000 solutes characterized by all or some of their descriptors. E and V can be calculated directly from the structure of the compound, whereas S, A, and B need to be determined experimentally using chromatographic or liquid–liquid distribution systems (Abraham et al., 1999). The coefficients of the equation are obtained by multiple linear regression analysis between the values of the studied property ($\log K_p$ in Eq. (2)) acquired for an appropriate group of solutes and their descriptors. The coefficients are characteristic of the partitioning system and reflect its complementary properties to the corresponding solute property.

Therefore, the characterization of a system by the solvation parameter model provides insight into the types and relative strength of the solute–solvent interactions that control the described process (Vitha and Carr, 2006). Thus, Eq. (2) shows that size is the only solute property that contributes to increase skin permeation (the only positive coefficient is the coefficient of V) whereas the rest of solute properties are opposed to the transfer of the solute from aqueous solution to human skin (the rest of coefficients are negative). Also, Eq. (2) points out that that size and hydrogen bonding (mostly hydrogen bond acceptor basicity) are the most determinant solute properties in human dermal absorption (the coefficients of V and B are the highest in absolute values).

The characterization of human skin permeability by Eq. (2) allows the estimation of K_p data for untested solutes as long as their solute descriptors are known. It consists in multiplying the coefficients of the equation by the descriptors of the new compounds. Another possibility is the comparison of Eq. (2) with the characterization of other systems, given that many biological (Abraham and Ràfols, 1995; Hoover et al., 2005; Platts et al., 2001; Poole and Poole, 1999) and physicochemical (Abraham et al., 1994; Lázaro et al., 2006; Vitha and Carr, 2006) processes have been successfully described by the solvation parameter model. By comparison of the characterizations, surrogate physicochemical systems can be identified for human skin permeation. K_p values for untested solutes can be then estimated by measuring a directly related physicochemical property in the surrogate system. In this instance, solute descriptors for the test compounds are not required.

Among the existing physicochemical systems, the octanol–water partition system has been the classical model for biopartitioning processes (Sangster, 1997; Smith et al., 1975). Nevertheless, chromatographic systems have gained popularity in this field (Cimpean and Poole, 2002; Escuder-Gilabert et al., 2003; Poole et al., 2000) because they encompass the main interactions between solutes and biological membranes (hydrophobic, steric, electronic contributions) and they also preserve the advantages of chromatographic techniques (reproducibility, speed, automatization). For these reasons, several authors have studied the feasibility of predicting K_p values of chemicals from chromatographic measurements (Barbato et al., 1998; Martínez-Pla et al., 2003; Nasal et al., 1995; Turowskil and Kaliszan, 1997; Wang et al., 2009).

In this work, the modeling of human skin permeation has been studied by means of both the reference octanol–water partition system and several possible surrogate chromatographic systems, basically of high-performance liquid chromatography (HPLC) and micellar electrokinetic chromatography (MEKC). For the octanol–water system, experimental log$P$ values were obtained from a reliable database (Bio-Loom). For chromatographic systems, experimental log$k$ values were determined in this work. The performances of the studied systems have been evaluated according to an established systematic procedure (Hidalgo-Rodríguez et al., 2010; Lázaro et al., 2006) which was successfully applied to assess the ability of the same systems to mimic the nonspecific toxicity of neutral organic compounds to the fathead minnow fish (Hidalgo-Rodríguez et al., 2012a) and the sorption of neutral organic chemicals by soil from water (Hidalgo-Rodríguez et al.,

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

Table 1
Chromatographic systems considered in this work.

Systems	Type of system	Experimental conditions	References
MSC18	HPLC	Xterra MSC18 column, 10 mM phosphate buffer pH 7:acetonitrile (60:40)	Lázaro et al., 2006
RP18	HPLC	Xterra RP18 column, 10 mM phosphate buffer pH 7:acetonitrile (60:40)	Lázaro et al., 2006
Octanol–water	Solvent-water partition	Diverse	Abraham et al., 1994
IAM	HPLC	Immobilized artificial membrane prepared with phosphatidyl choline (IAM.PC.DD2) column, 10 mM phosphate buffer pH 7:acetonitrile (60:40)	Lázaro et al., 2006
STC	MEKC	50 mM sodium taurocholate (STC), 20 mM phosphate pH 7	Hidalgo-Rodríguez et al., 2012a
SDS	MEKC	40 mM sodium dodecylsulfate (SDS), 20 mM phosphate pH 7	Fuguet et al., 2002a

2012b). The aim of this paper is showing the usefulness of this procedure to identify good surrogate physicochemical systems for the human skin permeation process. Thus the precision achieved in the estimation of human skin permeabilities by correlations with the studied systems has been predicted within the framework of the solvation parameter model, taking into account the error from the original data and the dissimilarity between the correlated systems. The predicted performances have been finally tested by experimental correlations. This work also presents the significant effect observed in the predictive abilities of the correlation models after adding the volume of the solute as additional variable in the human skin permeability-chromatographic data correlations.

2. Theory

2.1. Estimation of the precision achieved in biological-chromatographic correlations

The ability of a chromatographic system to model a biological one, such as the permeation of aqueous neutral solutes through the human skin, is in practice tested by correlating experimental data for a representative set of solutes:

$$\log K_p = q + p \log k \tag{3}$$

where q and p are the experimental intercept and slope of the correlation respectively, K_p corresponds to the skin-water permeability coefficient, and k to the chromatographic retention factor (it would be a partition coefficient, P, in the case of a water-solvent partition system like octanol–water).

The overall variance of the correlation determines the precision that could be achieved if K_p data for untested solutes is predicted by measurements in the chromatographic system. Therefore, the examination of the overall variance is crucial to evaluate the predictive ability of the correlation model. The overall variance (σ^2_{corr}) is contribution of three terms:

$$\sigma^2_{corr} = \sigma^2_{bio} + \sigma^2_{chrom} + \sigma^2_d \tag{4}$$

σ_{bio} and σ_{chrom} correspond to the intrinsic error of the biological and chromatographic data respectively, and σ_d is the error coming from the dissimilarity between the two correlated systems.

A particular method was developed (Hidalgo-Rodríguez et al., 2010) and successfully applied (Hidalgo-Rodríguez et al., 2012a,b) to estimate each contribution of variance by means of the solvation parameter model approach, just as follows:

$$\sigma^2_{bio} \approx SD^2_{bio} \tag{5}$$

$$\sigma^2_{chrom} \approx (p\ SD_{chrom})^2 \tag{6}$$

$$\sigma^2_d \approx SD^2_d \tag{7}$$

where SD_{bio} and SD_{chrom} are, respectively, the standard deviation of the biological and the chromatographic systems' characterization by the solvation parameter model; p corresponds to the slope of the correlation; SD_d is the standard deviation from the correlation between calculated values for both the biological and the chromatographic systems. The calculated values of a system can be obtained by means of its characterization, just multiplying the Abraham coefficients of the system by the descriptors of a set of solutes (usually the biological set). The overall variance can be then calculated by addition of these three estimated contributions:

$$SD^2_{corrcal} = SD^2_{bio} + (pSD_{chrom})^2 + SD^2_d \tag{8}$$

The lower $SD^2_{corrcal}$, the better predictive ability of the biological-chromatographic correlation is expected.

2.2. The d distance as a tool for measuring the dissimilarity of systems

The d distance (Lázaro et al., 2006) is a useful parameter to indicate the dissimilarity between two systems characterized by the solvation parameter model. It is closely related to the angle θ_{ij} proposed by Ishihama and Asakawa (Ishihama and Asakawa, 1999) also as a measure of dissimilarity between pair of systems. Both parameters can be related (Lázaro et al., 2006) by the following expression:

$$\theta_{ij} = 2\ arcsin(d/2) \tag{9}$$

The d distance can be easily calculated as follows:

$$d = \sqrt{(e_{ui} - e_{uj})^2 + (s_{ui} - s_{uj})^2 + (a_{ui} - a_{uj})^2 + (b_{ui} - b_{uj})^2 + (v_{ui} - v_{uj})^2} \tag{10}$$

where e, s, a, b, and v are the five coefficients of the characterization of a system by the solvation equation, the subscripts i and j represent the two compared systems, and u indicates that all the coefficients have been previously normalized. The coefficients are normalized according to

$$e_u = \frac{e}{l} \tag{11}$$

$$s_u = \frac{s}{l} \tag{12}$$

$$a_u = \frac{a}{l} \tag{13}$$

$$b_u = \frac{b}{l} \tag{14}$$

$$v_u = \frac{v}{l} \tag{15}$$

where l is the length of the coefficients' vector:

$$l = \sqrt{e^2 + s^2 + a^2 + b^2 + v^2} \tag{16}$$

Since the d parameter measures the distance between the normalized vectors of a pair of systems, the higher d is, the less similar the two systems are. They would be perfectly correlated in the case of $d = 0$.

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

Table 2
Characterization of the systems by the solvation parameter model and values for the *d* distance.

System	Coefficients						Statistics				Normalized coefficients					*d*
	c	*e*	*s*	*a*	*b*	*v*	*n*	*r*	*SD*	*F*	e_u	s_u	a_u	b_u	v_u	
Biological																
Human skin permeation	−5.426	−0.106	−0.473	−0.473	−3.000	2.296	119	0.912	0.461	112	−0.028	−0.123	−0.123	−0.782	0.598	0.000
Chromatographic																
MSC18 HPLC	−0.152	0.151	−0.463	−0.424	−1.719	1.759	55	0.994	0.072	764	0.059	−0.182	−0.167	−0.676	0.692	0.181
RP18 HPLC	−0.172	0.233	−0.460	−0.260	−1.574	1.582	55	0.996	0.056	1121	0.101	−0.200	−0.113	−0.683	0.686	0.200
Octanol–water	0.088	0.562	−1.054	0.034	−3.460	3.814	613	0.997	0.116	23162	0.106	−0.199	0.006	−0.655	0.721	0.268
IAM HPLC	−0.710	0.292	−0.344	0.141	−1.193	1.161	51	0.991	0.047	480	0.169	−0.199	0.081	−0.689	0.671	0.316
STC MEKC	−1.838	0.613	−0.513	0.319	−2.396	2.390	56	0.984	0.112	305	0.175	−0.147	0.091	−0.686	0.684	0.323
SDS MEKC	−1.680	0.558	−0.596	−0.266	−1.674	2.717	63	0.991	0.120	618	0.169	−0.180	−0.080	−0.507	0.822	0.412

Table 3
Predicted variances for the human skin permeation-chromatographic correlations.

Systems	Human skin permeation			SD^2_{bio} = 0.213 n_{bio} = 119	
	q_{cal}	p_{cal}	$(p_{cal}\ SD_{chrom})^2$	SD^2_d	$SD^2_{corrcal}$
MSC18 HPLC	−6.08	1.31	0.009	0.400	0.621
RP18 HPLC	−6.12	1.45	0.007	0.407	0.626
Octanol–water partition	−6.54	0.44	0.003	0.691	0.906
IAM HPLC	−5.48	1.58	0.006	0.598	0.816
STC MEKC	−5.57	0.47	0.003	0.830	1.045
SDS MEKC	−5.72	0.01	0.000	1.016	1.229

Table 4
Experimental log*k* values for the tested solutes in the studied chromatographic systems.

Solute	HPLC systems			MEKC systems	
	MSC18	RP18	IAM	STC	SDS
1-Butanol	–	–	–	−1.296	−0.703
1-Pentanol	–	–	–	−0.956	−0.199
1-Propanol	–	–	–	−1.533	–
2,4-Dichlorophenol	0.670	0.788	0.373	0.572	0.793
2-Chlorophenol	0.362	0.459	0.002	−0.167	0.103
2-Naphthol	0.548	0.657	0.210	0.412	0.763
2-Nitro-p-phenylenediamine	−0.160	−0.064	−0.425	−0.562	−0.335
3-Methylphenol	0.295	0.364	−0.115	−0.372	−0.005
3-Nitrophenol	0.255	0.403	−0.014	−0.177	−0.032
4-Bromophenol	0.492	0.602	0.180	0.266	0.442
4-Chloro-3-methylphenol	0.616	0.698	0.242	0.486	0.737
4-Chlorophenol	0.429	0.527	0.096	0.101	0.274
4-Hydroxybenzyl alcohol	−0.433	−0.388	−0.571	−0.831	−0.710
4-Methylphenol	0.298	0.356	−0.111	−0.350	0.050
8-Methoxypsoralen	0.491	0.507	−0.066	0.103	1.098
Aniline	0.107	0.104	−0.372	−0.930	−0.384
Anisole	0.691	0.611	−0.105	−0.352	0.128
Benzaldehyde	0.356	0.324	−0.292	−0.694	−0.070
Benzene	0.737	0.636	−0.080	−0.489	−0.110
Benzyl alcohol	−0.015	0.004	−0.402	−0.861	−0.331
Benzyl nicotinate	0.691	0.626	−0.021	0.262	1.295
Caffeine	−0.457	−0.485	−0.723	−0.721	−0.304
Corticosterone	0.324	0.296	−0.090	0.096	2.016
Dexamethasone	0.246	0.258	−0.140	0.021	1.697
Estradiol	0.635	0.827	0.575	0.901	2.111
Ethylbenzene	1.224	1.070	0.238	0.387	0.744
Hydrocortisone	0.014	0.052	−0.242	−0.029	1.467
Methyl 4-hydroxybenzoate	0.113	0.230	−0.184	−0.192	0.270
o-Phenylenediamine	−0.286	−0.285	−0.526	−1.054	−0.554
Phenol	0.117	0.189	−0.241	−0.663	−0.390
Progesterone	1.311	1.147	0.430	0.792	–
Resorcinol	−0.298	−0.204	−0.403	−0.661	−0.705
Testosterone	0.685	0.672	0.233	0.468	2.141
Thymol	1.014	1.030	0.333	0.493	0.962
Toluene	0.983	0.861	0.086	−0.019	0.342

3. Materials and methods

3.1. Apparatus and conditions

3.1.1. HPLC systems

The retention data was measured in an IAM.PC.DD2 column (100 × 4.6 mm I.D., 12 μm particle size, Regis Technologies Inc., Morton Grove, IL, USA), XTerra MSC18, XTerra RP18, and XBridge C18 columns (150 × 4.6 mm I.D., 5 μm particle size, Waters Corporation, Milford, MA, USA). All measurements were performed with a Shimadzu liquid chromatograph (Kyoto, Japan) equipped with two Shimadzu LC-20AD pumps and a Shimadzu SPD-M20A diode array detector. The temperature was controlled at 25.0 ± 0.1 °C with a Shimadzu CTO-10AS column oven.

The mobile phase was a mixture of 40% acetonitrile and 60% 10 mM phosphate aqueous buffer adjusted to pH 7. All solutes were dissolved in methanol and their concentrations were 0.1 mg ml^{-1}. The injection volume was 10 μl. Isocratic conditions were used at a flow rate of 1 ml min^{-1}. The column hold-up time was determined by using an aqueous solution of potassium bromide (0.1 mg ml^{-1}) as an unretained solute, whose detection was performed at 200 nm. All measurements were taken in triplicate.

3.1.2. MEKC systems

MEKC separations were performed with a Beckman P/ACE System 5500 Capillary Electrophoresis with a UV diode array detector. The fused-silica capillary was 40 cm effective length × 50 μm I.D. and was obtained from Composite Metal Services Ltd. (Shipley, West Yorkshire, UK). The capillary was activated by the following washing sequence: water (10 min), 1 M NaOH (10 min), water (5 min), 0.1 M NaOH (10 min), and separation solution (20 min). As daily conditioning, the capillary was flushed with water for 5 min, followed by 0.1 M NaOH for 5 min, water for 5 min and separation solution for 20 min. Prior to each separation the capillary was flushed with 0.1 M NaOH for 1 min and separation solution for 3 min. Retention measurements were made at 25 °C and +15 kV. Detection was at 214 nm.

The sodium taurocholate (STC) system was a separation solution 50 mM in taurocholic acid sodium salt hydrate and 20 mM in phosphate aqueous buffer adjusted to pH 7. The same buffer conditions were used to prepare the separation solution of the sodium dodecylsulfate (SDS) system, which was 40 mM in SDS. All solutes were dissolved in methanol (used as electroosmotic flow marker) and contained 2 mg ml^{-1} of dodecanophenone as micellar marker (Fuguet et al., 2002b). The concentration of the solutes was 2 mg ml^{-1}. All solutions were filtered through 0.45-μm nylon syringe filters obtained from Albet (Dassel, Germany). Samples were introduced into the capillary by applying a pressure of 0.5 p.s.i. for 1 s. All measurements were taken in triplicate.

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

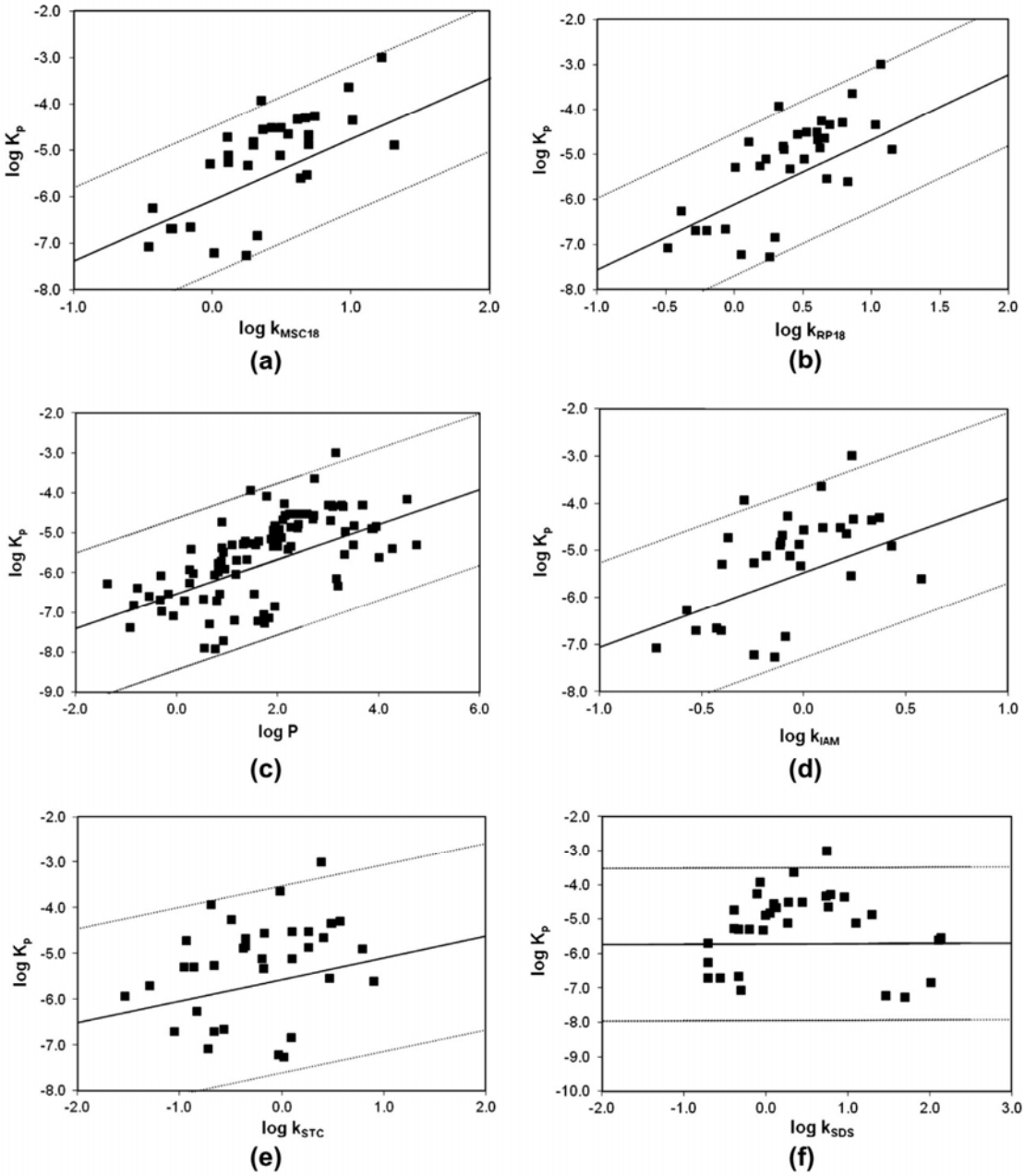


Fig. 1. Plots of experimental logK_p (skin-water permeability coefficient logarithm) against log$k$ (retention factor logarithm) or log$P$ (octanol–water partition coefficient logarithm) in the studied systems: (a) MSC18 HPLC, (b) RP18 HPLC, (c) octanol–water partition, (d) IAM HPLC, (e) STC MEKC, (f) SDS MEKC. The correlations between calculated values for both properties by means of their Abraham characterizations are also shown (—), as well as the ranges for 95% of confidence level (±2$SD_{corr\ cal}$) (·····).

3.2. Reagents

Methanol (HPLC grade), hydrochloric acid (25% in water), sodium hydroxide (>99%), sodium dihydrogenphosphate monohydrate (>99%), and SDS (>99%) were from Merck (Darmstadt, Germany). Acetonitrile (HPLC grade) was from VWR International (West Chester, Pennsylvania, USA). Taurocholic acid sodium salt hydrate (98%) was from Acros Organics (Geel, Belgium). Dodecanophenone (98%) was from Aldrich (Steinheim, Germany). Water was purified by a Milli-Q plus system from Millipore (Bedford, MA, USA), with a resistivity of 18.2 MΩ cm. The test solutes employed were reagent grade or better and obtained from several manufacturers (Merck, Sigma (Steinheim, Germany), Fluka (Steinheim, Germany), Aldrich, Carlo Erba (Milano, Italy), Baker (Deventer, Netherlands)).

3.3. Calculation

Microsoft Excel was used for data calculations and multiple linear regression analysis.

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

ARTICLE IN PRESS

Table 5
Results of the experimental $\log K_p$ against $\log k$ correlations (standard deviations in brackets).

Systems	Human skin permeation					
	q	p	$SD^2_{correxp}$	n	r	F
MSC18 HPLC	−5.89 (0.18)	1.77 (0.31)	0.602	32	0.718	32
RP18 HPLC	−5.99 (0.19)	1.91 (0.33)	0.596	32	0.721	33
Octanol–water partition	−6.39 (0.14)	0.49 (0.06)	0.666	96	0.623	60
IAM HPLC	−5.09 (0.17)	1.94 (0.54)	0.870	32	0.547	13
STC MEKC	−5.13 (0.18)	0.57 (0.29)	1.024	35	0.331	4
SDS MEKC	−5.23 (0.21)	−0.05 (0.23)	1.203	33	0.038	1

Table 6
Predicted variances when the correction of volume is applied to the human skin permeation-chromatographic correlations.

Systems	Human skin permeation				$SD^2_{bio} = 0.213$ $n_{bio} = 119$	
	q_{cal}	p_{cal}	t_{cal}/p_{cal}	$(p_{cal}\ SD_{chrom})^2$	SD^2_d	$SD^2_{corrcal}$
MSC18 HPLC	−5.15	1.58	−0.39	0.013	0.032	0.258
RP18 HPLC	−5.21	1.75	−0.35	0.010	0.041	0.263
Octanol–water partition	−5.76	0.81	−1.13	0.009	0.063	0.284
IAM HPLC	−4.19	2.29	−0.32	0.012	0.131	0.355
STC MEKC	−3.72	1.15	−0.88	0.017	0.169	0.398
SDS MEKC	−3.06	1.50	−1.50	0.033	0.123	0.368

Table 7
Results of the experimental correlations of $\log K_p$ against $\log k$ and V (standard deviations in brackets).

Systems	Human skin permeation						
	q	p	t	$SD^2_{correxp}$	n	r	F
MSC18 HPLC	−4.62 (0.14)	2.02 (0.14)	−1.04 (0.09)	0.118	32	0.953	144
RP18 HPLC	−4.76 (0.15)	2.14 (0.16)	−1.01 (0.10)	0.139	32	0.945	120
Octanol–water partition	−5.44 (0.11)	0.79 (0.04)	−1.15 (0.09)	0.239	96	0.885	168
IAM HPLC	−3.58 (0.24)	2.56 (0.35)	−1.12 (0.16)	0.335	32	0.860	41
STC MEKC	−3.20 (0.29)	1.38 (0.21)	−1.39 (0.20)	0.410	35	0.809	30
SDS MEKC	−2.37 (0.25)	1.73 (0.17)	−2.85 (0.23)	0.208	33	0.913	75

3.3.1. HPLC systems

The HPLC retention data was expressed by means of the retention factor, k:

$$k = \frac{(t_R - t_0)}{t_0} \tag{17}$$

where t_R corresponds to the solute retention time and t_0 is the column hold-up time determined by an aqueous potassium bromide solution.

3.3.2. MEKC systems

The MEKC retention factor, k, was calculated as follows:

$$k = \frac{(t_R - t_0)}{\left(1 - \frac{t_R}{t_m}\right) t_0} \tag{18}$$

t_R is the solute migration time, t_0 corresponds to the migration time of methanol (used to determine the electroosmotic flow), and t_m is the migration time of dodecanophenone (used to determine the migration time of the micelles).

4. Results and discussion

4.1. Ability of chromatographic systems to model human skin permeation

Table 1 presents the chromatographic systems considered in this work for evaluation of their suitability to estimate skin-water permeability coefficients of neutral compounds (Eq. (3)). They are three HPLC systems based on two conventional C18 columns (MSC18 and RP18) and one immobilized artificial membrane (IAM) column, two MEKC systems composed of sodium taurocholate (STC) on one hand and sodium dodecylsufalte (SDS) on the other hand, and also the classical octanol–water partition system.

Systems of Table 1 were selected mainly because of two reasons: the similarity of the systems to the biological process to mimic, measured from the d distance parameter (see Section 2.2), and the advantages of chromatographic techniques in terms of cost, time, and popularity.

The d distances between human skin permeation and each system of Table 1 were calculated according to Eq. (10). Table 2 shows the characterization of the systems by the solvation parameter model, their normalized coefficients, and the calculated d distances, which are low or at least not too high to dismiss the possibility of obtaining good correlation models for human skin permeation. According to the d parameter, the HPLC systems with C18 columns are the closest to the human skin permeation system, whereas the MEKC systems are the most different to this biopartitioning process. However, the d parameter alone does not quantify the precision that can be achieved in the estimation of K_p by means of measurements in the chromatographic systems.

Hence the methodology described in the Section 2.1 was applied to predict the variances that will affect the correlations between human skin permeation and each chromatographic system. The results are presented in Table 3 (see that the systems are ordered according to increasing d distances).

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

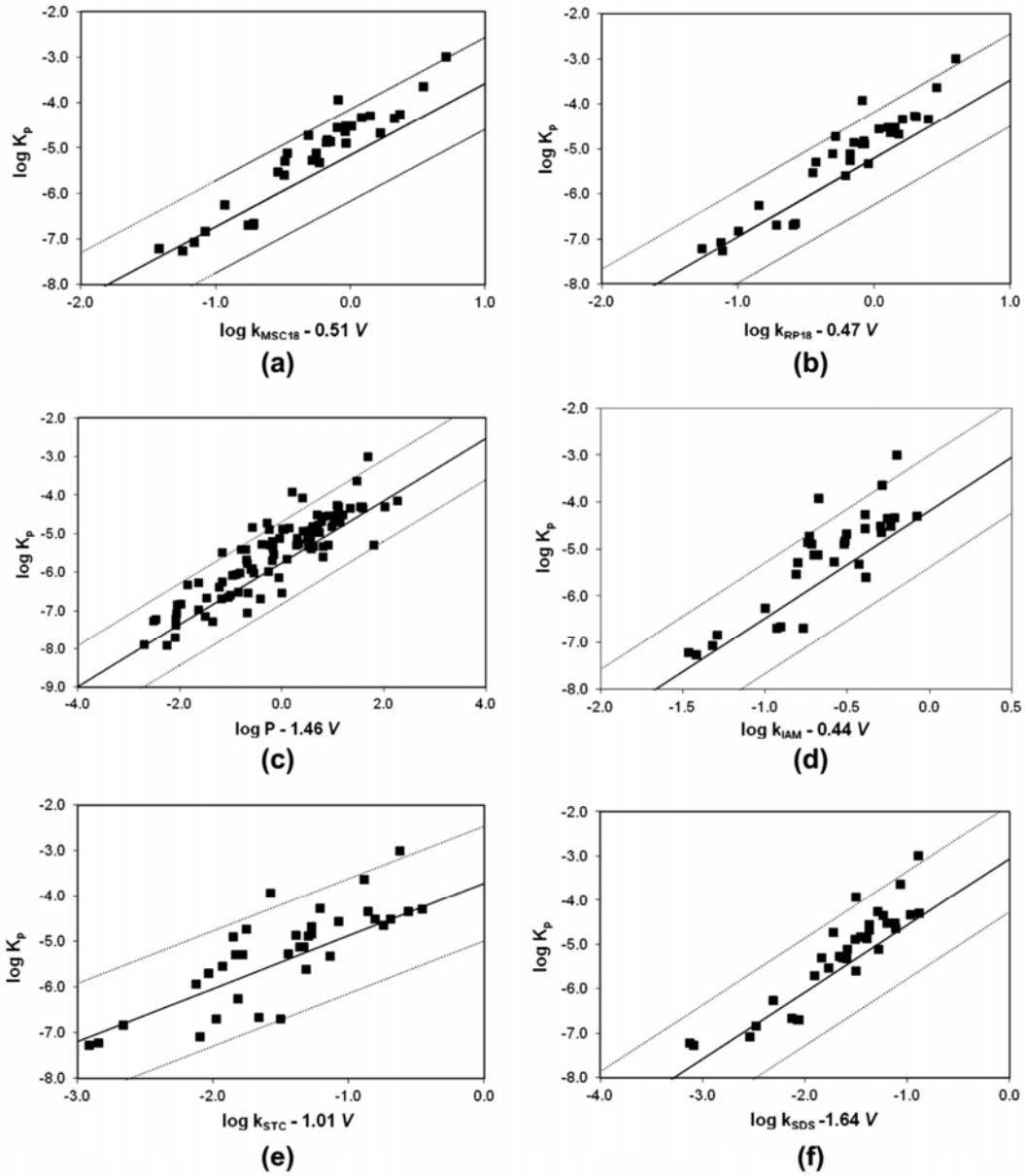


Fig. 2. Experimental correlations corrected by the volume of the solute, i.e. $\log K_p$ vs. $\log k$ and V or $\log K_p$ vs. $\log P$ and V, for the studied systems: (a) MSC18 HPLC, (b) RP18 HPLC, (c) octanol–water partition, (d) IAM HPLC, (e) STC MEKC, (f) SDS MEKC. The lines correspond to the correlations between calculated values and V (—) and their corresponding ranges for 95% of confidence level ($\pm 2SD_{corr\ cal}$) (·····).

The contribution of variance due to the intrinsic error of the chromatographic data (Eq. (6)) was predicted using the slope of the correlation between calculated values (p_{cal}) because the experimental slope (p) is initially not known. Both the intercept and slope of the calculated correlations (q_{cal} and p_{cal} respectively) are included in Table 3. The same 119 solutes of the human skin permeation characterization (Eq. (2)) (Abraham and Martins, 2004) were used to obtain the calculated correlations.

The predicted variances of Table 3 for K_p estimation from chromatographic measurements show that the error coming from the original skin permeability data (SD^2_{bio}) is quite significant. Biological data should be ideally measured in a single laboratory using designated protocols and with uniform samples taken from the same source. However, this hardly ever occurs and the uncertainty of the experimental data can contribute considerably to get less precise correlations. Nevertheless, the major contribution of variance comes in all cases from the dissimilarity between the human skin permeation and the chromatographic systems (SD^2_d), whose increase agrees with the increase of the d distance. In fact, SD^2_d is twice SD^2_{bio} at least, which means that a lot of error will be introduced in the correlations with regard to the intrinsic error of the original data. This fact is reflected in the high values obtained for

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

Table 8
Experimental logK_p values (average), log$k$ values in the column XBridge C18, and volume of the selected solutes.

Solute	Average logK_p	Log$k$	V
Cortisone[a]	−8.26	−0.029	2.7546
Estriol[b]	−7.96	−0.219	2.2575
Prednisolone[a]	−7.91	−0.113	2.7546
Antipyrine[b]	−7.74	−0.407	1.4846
Cortexolone[a]	−7.68	0.292	2.7389
Dexamethasone[b]	−7.51	0.184	2.9131
Caffeine[a]	−7.30	−0.543	1.3632
Resorcinol[b]	−6.79	−0.368	0.8338
Hydroxyprogesterone[a]	−6.78	0.787	2.6802
Aminopyrine[b]	−6.55	−0.161	1.8662
Testosterone[a]	−6.46	0.582	2.3827
Pregnenolone[b]	−6.38	1.256	2.6645
Pyridine[a]	−6.30	−0.364	0.6753
Catechol[b]	−6.20	−0.245	0.8338
Estrone[a]	−6.00	0.827	2.1558
Progesterone[b]	−5.89	1.274	2.6215
Atrazine[a]	−5.56	0.509	1.6196
Benzyl alcohol[b]	−5.54	−0.006	0.9160
Methyl-4-hydroxybenzoate[a]	−5.48	0.065	1.1313
Aniline[b]	−5.38	0.057	0.8162
N,N-dimethylaniline[a]	−5.26	0.856	1.0980
2-Toluidine[b]	−5.00	0.251	0.9571
4-Bromophenol[a]	−4.88	0.506	0.9501
3-Xylene[b]	−4.66	1.262	0.9982
Cumene[a]	−4.41	1.471	1.1391
Ethylbenzene[b]	−3.85	1.261	0.9982
Styrene[a]	−3.75	1.082	0.9552

[a] Training set.
[b] Test set.

the predicted overall variances ($SD^2_{corrcal}$). According to the predictions, poor performances are therefore expected in the modeling of human skin permeability by these systems alone.

The performances were checked by direct correlations of experimental logK_p (Abraham and Martins, 2004) against experimental log$k$ (log$P$ for the octanol–water partition system). Some compounds used in the human skin permeation characterization (Eq. (2)) (Abraham and Martins, 2004) were injected in each chromatographic system to determine their experimental retention factors, k. The corresponding experimental log$k$ values are detailed in Table 4. As far as the octanol–water partition system is concerned, experimental log$P$ values were obtained from a reliable database (Bio-Loom). Fig. 1 shows the plots and Table 5 collects the details of the correlations obtained.

Fig. 1 clearly illustrates the high dispersion of the data points in all plots, which is also reflected in the results presented in Table 5. We can see that the dispersion of the points increases from Fig. 1a to Fig. 1f, which agrees with the decrease of the correlation coefficient r and the Fischer statistic F detailed in Table 5. Especially, there is no significant correlation at all between experimental logK_p and log$k$ data for the SDS MEKC system, as it is clearly observed in Fig. 1f and in the poor correlation statitstics of Table 5 ($r = 0.038$, $F = 1$).

Table 5 also shows that the overall variances of the experimental correlations ($SD^2_{correxp}$) are certainly high, which agree with the previous predictions. In fact, the $SD^2_{correxp}$ values are nearly identical to the predicted ones of Table 3 ($SD^2_{corrcal}$), which highlights the goodness of the method used for predicting the variances. This systematic methodology is proven therefore to be very useful for

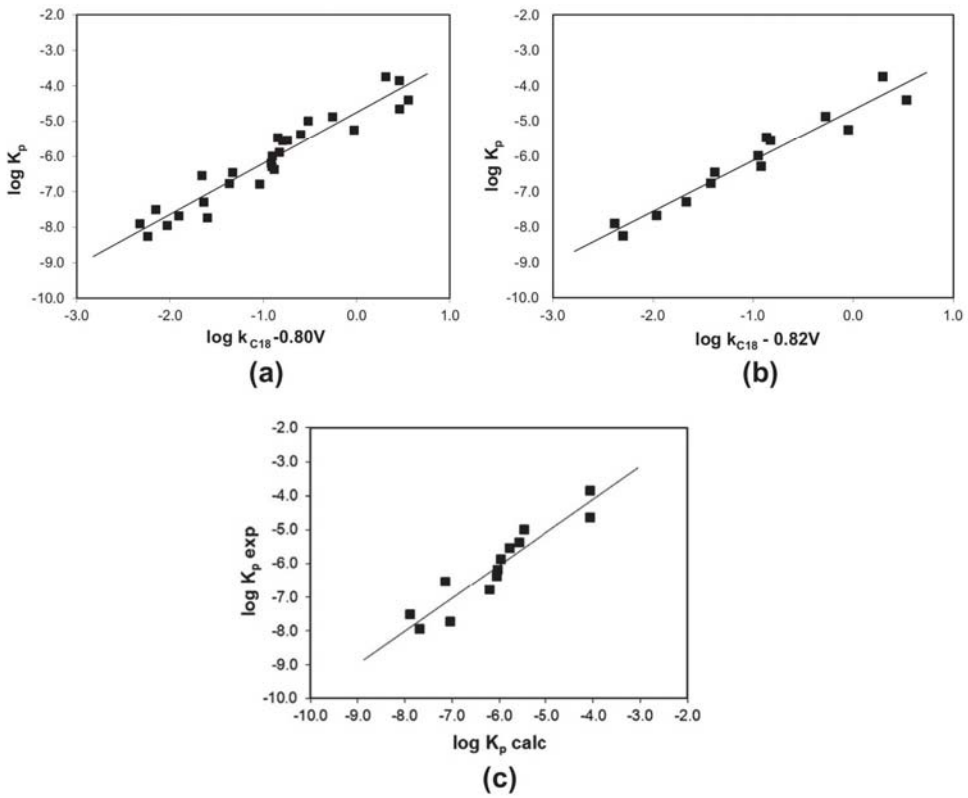


Fig. 3. Experimental correlations corrected by the volume of the solute, i.e. logK_p vs. log k and V for the system XBridge C18: (a) global correlation, (b) calibration set correlation. Fig 3c is the plot of the average experimental logK_p values against the logK_p values estimated in this work for the compounds of the test set.

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

Table 9
Comparison of experimental and estimated $\log K_p$ values.

Compound	Experimental $\log K_p$				Estimated $\log K_p$		
	Abraham and Martins, 2004	Patel et al., 2002	Neumann et al., 2006	Neely et al., 2009	Eq. (2)	Patel et al., 2002	This work
Pregnenolone	–	−6.38	−6.38	−6.38	−4.70	−6.25	−6.04
Estriol	–	−7.96	−7.96	−7.96	−6.67	−6.86	−7.68
Progesterone	−4.90	−6.38	−6.38	−5.92	−4.54	−6.51	−5.96
Dexamethasone	−7.27	–	–	−7.75	−6.71	–	−7.88
Propazine	–	–	–	–	−5.13	–	−5.65
Simazin	–	–	–	–	−5.77	–	−6.08
Terbutylazin	–	–	–	–	−5.07	–	−5.55
Fluomethuron	–	–	–	–	–	–	−5.81
Linuron	–	–	–	–	−4.85	–	−5.32
Isoproturon	–	–	–	–	−5.25	–	−6.02
Monuron	–	–	–	–	−5.43	–	−6.11

predicting the ability of chromatographic systems to emulate the human skin permeation process, which is a key issue to select appropriate surrogate systems to estimate K_p data.

According to both predicted and experimental variances of Tables 3 and 5, as well as Fig. 1, human skin permeation is not well modeled by any of the considered physicochemical systems. Although the two conventional HPLC systems of C18 columns are the best suited because their correlations have the lowest overall variances, these variances are already too high to provide precise estimations of human skin permeability coefficients. The poor performances can be attributed to the fact that the chromatographic systems do not entirely encompass the main factors that control dermal absorption. For this reason the next section will deal with the correction of the correlations by an additional parameter in order to improve the predictive abilities.

4.2. Correction of solute volume in human skin permeation against chromatographic data correlations

Biological-chromatographic correlation models with limited predictive abilities, such as the ones of the previous section, are unfortunately very common because good surrogate chromatographic systems of biological processes are difficult to find.

Under these circumstances, the addition of another independent variable to the correlation can help to mimic better the biological property under study. A suitable variable for this purpose can be the volume of the solute, which provided good results in other previous (Abraham and Ràfols, 1995; Abraham et al., 1995a; Rosés et al., 1999). On one hand, McGowan's characteristic volume can be easily calculated for any solute from the molecular formula and the number of rings, using the algorithm of Abraham and McGowan for the number of bonds in a molecule (Abraham and McGowan, 1987; Abraham, 1993). Thus, it is an additive property of quite simple calculation. On the other hand, the volume of the solute is usually a determinant property in any biopartitioning process. In particular, Eq. (2) shows its large influence in the permeation of neutral solutes through the human skin. The suitability of molecular volume as descriptor for predicting skin permeation was already highlighted by Potts and Guy (Potts and Guy, 1992, 1995) and Barratt (Barratt, 1995). Potts and Guy presented the use of molecular volume in combination with $\log P$, as well as in combination with descriptors for solute hydrogen bond acidity and basicity. Barratt used molecular volume together with $\log P$ and melting point. In fact, many scientists have developed QSARs for skin permeation with a generic form consisting of a term of lipophilicity and also a term of molecular size (Cronin et al., 1999; Fitzpatrick et al., 2004; Geinoz et al., 2004; Patel et al., 2002). Both factors are crucial in the penetration of solutes through the skin, as it was mentioned in the introduction. However, the use of volume alone does not result in satisfactory predictions. Volume should be combined with another term related to lipophilicity in order to improve the correlations. In this work chromatographic retention factors represent the lipophilicity term and the addition of the solute's volume to the correlations attempt to account for the molecular size influence in the skin permeation process.

Hence, we decided to test the inclusion of the solute's volume (i.e. the V descriptor of Eq. (2)) in the $\log K_p$ against $\log k$ correlations. When the correction of volume is applied, Eq. (3) can be converted into Eq. (19):

$$\log K_p = q + p(\log k + t/pV) \tag{19}$$

Eq. (19) consists of a correlation with two independent variables ($\log k$ and V). V corresponds to the McGowan's volume of the solute and t is its corresponding coefficient in the correlation.

Table 6 shows the predictions of variances after applying the correction of volume (Eq. (19)). Thus, calculated $\log K_p$ from Eq. (2) were correlated against calculated log k and McGowan's volumes for the 119 solutes of the human skin permeation characterization. A huge decrease in SD_d^2 and, consequently also in $SD_{corrcal}^2$, is observed for all systems of Table 6 in comparison to the predicted values of Table 3. In fact, the SD_d^2 contributions of Table 6 are all below the SD_{bio}^2, which means that the correlations corrected by the volume are expected to be mainly affected by the inherent error of the K_p values rather than by the dissimilarity between the human skin permeation-chromatographic systems. Therefore, the addition of volume to the correlations is expected to improve them significantly and provide good correlation models allowing precise estimations of K_p values for untested solutes.

The same experimental data of the previous section, experimental $\log K_p$ and log k values, were then correlated including the solute's volume as additional descriptor (Eq. (19)). The results are collected in Table 7. In agreement with the predicions (Table 6), much lower $SD_{correxp}^2$ were obtained compared to those of Table 5 (without the correction of volume). Also the increase of the r and F statistics (compare Table 7 with Table 5) shows the improvement of the correlations after the correction of volume, which is highly noticeable for the SDS MEKC system (r goes from 0.038 to 0.913 and F goes from 1 to 75). Thus, the selected human skin permeation-chromatographic correlations improve significantly when the solute's volume is used as additional variable, which confirms that retention factors alone are not enough for a good modeling of dermal absorption. These results agree with a number of QSARs which show that skin permeation is mediated both by the lipophilicity and the molecular size of the permeant (Cronin et al., 1999; Fitzpatrick et al., 2004; Geinoz et al., 2004; Patel et al., 2002). Thus, the passage of a molecule through the stratum corneum and its diffusion are highly influenced by both lipophilicity and molecular size of the compound. However, whereas lipophilicity contributes positively, molecular size contributes negatively in the permeation process.

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

Fig. 2 clearly illustrates much improved correlations in comparison to Fig. 1. As it has been mentioned, the most astonishing improvement is observed in the correlation of the SDS MEKC system (compare Fig. 1f with Fig. 2f). This fact can be explained by looking at the normalized coefficients of the systems, which are detailed in Table 2. The SDS MEKC system presents the most different normalized coefficient of volume (v_u) with regard to the v_u value of human skin permeation. Therefore, the correlation of both systems is coherently the most corrected by the inclusion of volume as additional variable.

Apart from the experimental points, Fig. 2 also includes our predictions for a 95% of confidence level ($\pm 2SD_{corr\,cal}$). As it is observed, almost all the points fall into our predictions, regardless of the chromatographic system considered.

According to the experimental overall variances of Table 7 and the plots shown in Fig. 2, the conventional HPLC systems based on C18 columns provide very precise estimations of skin-water permeability coefficients when the correlations are corrected by the solute's volume. The octanol–water partition system and the SDS MEKC system are the following best suited systems to model human skin permeation. The systems of IAM HPLC and particularly STC MEKC present the worst performances as surrogates of human skin permeation (their correlations have the highest dispersion). However, they still mimic well this biopartitioning process after the correction of volume.

Last but not least, we can observe both in Table 7 and in Fig. 2 that the term of volume (t/p in Eq. (19)) is negative in all the studied correlations, which agrees with QSARs proposed by other authors based on lipophilicity and solute size parameters (Cronin et al., 1999; Fitzpatrick et al., 2004; Geinoz et al., 2004; Patel et al., 2002). This fact means that the $\log K_p$ values predicted from chromatographic measurements and the solute's volume are lower than those predicted only from chromatographic data. Besides, the coefficient of V shows the importance of the volume correction. For the HPLC systems, the coefficient of $V(t)$ is lower in absolute value than the retention factor term, which means that the chromatographic term has more influence in the correlations than the solute's volume. Thus, the two HPLC systems with C18 columns are the best suited to model human skin permeation mainly because of the contribution of the chromatographic data. Table 7 shows that both chromatographic data and solute's volume have the same influence on the correlation of the STC MEKC system. For the rest of systems, octanol–water and SDS MEKC, the solute's volume contributes more to the correlations than the chromatographic data does (term t is higher than p in absolute value). In fact, the biggest coefficient of V corresponds to the correlation of SDS MEKC, which has been shown to improve the most after the volume correction.

4.3. Current estimation of skin permeation values through HPLC measurements

Once identified the best surrogate systems to estimate skin permeation, they can be used to predict $\log K_p$ values for new compounds. It has been proved that the two HPLC systems based on C18 columns (XTerra MSC18 and XTerra RP18) are the ones that provided better results. Since C18 columns are in general very similar in partition properties, it seems reasonable that the previous conclusion can be extended to other C18 columns. Column properties may change slightly because of modifications in the manufacturing process or column use, or even a particular column may no longer be available. Thus, lab calibration of the particular column to be used will facilitate the practical implementation of the proposed model in different physchem labs, which may use C18 columns of different makers, qualities, specifications and performances. Therefore, we have selected a common C18 column (XBridge C18 from Waters), different from the two used in the study of previous sections, to establish a method for the current estimation of skin permeation.

To establish the experimental methodology we have selected a set of 27 substances that belong to different families, have different structural complexity, and have known and varied experimental $\log K_p$ values. These values have been compiled from different literature sources (Abraham and Martins, 2004; Neely et al., 2009; Neumann et al., 2006; Patel et al., 2002). In case that more than one experimental $\log K_p$ value is published for a same compound, we have used the average of the published values. After injecting the selected substances into the HPLC system, a multiple linear regression analysis has been done between the experimental $\log K_p$ and the parameters $\log k$ and volume. Table 8 shows the selected compounds, its experimental $\log K_p$ values, the retention factor logarithm in the C18 column and its volume. Fig. 3a shows the obtained correlation, which has the following correlation equation and statistics:

$$\log K_p = -4.76 \pm 0.18 + 1.44 \pm 0.12 \log k - 1.16 \pm 0.10V$$
$$SD^2 = 0.152 \quad n = 27 \quad r = 0.913 \quad F = 126 \tag{20}$$

It can be observed that the standard deviation, the correlation coefficient, and the coefficients of the equation are of the same order as the ones obtained for the MSC18 and RP18 systems with a different set of compounds.

To prove the validity of the surrogate system to estimate $\log K_p$ values, the set of 27 compounds has been divided into two subsets: a training set and a test set. The aim of the training set is to establish a group of compounds that can be used to calibrate the HPLC system. Then, the test set is used to validate the goodness of the calibration set to predict new $\log K_p$ values. To construct the two substets, compounds have been distributed according to its $\log K_p$ value, and then they have been divided alternatively into the training set and the test set. Table 8 shows the subset to which belongs each compound. Again, a multiple linear regression has been performed with the 14 compounds of the training set. Fig. 3b shows the plot of the experimental $\log K_p$ values versus the $\log k$ and the volume of the compounds, and the regression equation and statistics are the following:

$$\log K_p = -4.69 \pm 0.26 + 1.44 \pm 0.18 \log k - 1.19 \pm 0.13V$$
$$SD^2 = 0.129 \quad n = 14 \quad r = 0.970 \quad F = 86 \tag{21}$$

The obtained statistics allow the internal validation of the training set so that the robustness of the system is guaranteed: the coefficients of the two parameters of the correlation ($\log k$ and V) have a good significance level, the variance (SD^2) agrees with the expected one for C18 HPLC systems, the correlation coefficient (r) is quite above 0.8, and Fisher's F parameter is high.

The test set has been used to perform the external validation of the system, i. e. ensure its predictive ability. To this purpose, the experimental $\log K_p$ values have been plotted against the values predicted through the correlation of the training set (Fig. 3c), obtaining the following equation:

$$Log\ K_p\ exp = -0.21 \pm 0.68 + 0.97 \pm 0.11 \log K_p calc$$
$$SD^2 = 0.201 \quad n = 13 \quad r = 0.937 \quad F = 79 \tag{22}$$

A very good concordance is observed between the experimental and the predicted values. The slope of the trend line is close to the unity, r is also above 0.8, and Fisher's F parameter is also high. The variance of the regression (0.201) is only slightly higher than the one of the training set (0.129). Therefore, it can be concluded that the selected training set is able to predict $\log K_p$ values.

Next step has been the prediction of the skin permeation for a series of hormones and pesticides, and when possible, the comparison of the obtained values against experimental or calculated

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

ARTICLE IN PRESS

values from literature. Results are shown in Table 9. The first four columns show experimental $\log K_p$ values taken from different sources (Abraham and Martins, 2004; Neely et al., 2009; Neumann et al., 2006; Patel et al., 2002). The last three columns show the values obtained through three different approaches: the values estimated through the solvation parameter model (Eq. (2)) (Abraham and Martins, 2004; ACD/Percepta, 2013), the values calculated by Patel (Patel et al., 2002), and finally the values obtained through the proposed method. It can be observed that there is very good agreement between the experimental values and the ones proposed in this work for the series of hormones (around 5% error). However, higher differences are observed when our data is compared to the ones of Patel (around 10%) and the highest difference is when compared to data predicted by Eq. (2) (15–30% error). Regarding the series of pesticides, there are no experimental values to compare with. However, as Abraham experimental descriptors for almost all pesticides are known (ACD/Percepta, 2013) we could calculate the skin permeation values of these compounds through Eq. (2). When they are compared to our results a difference of 10–15% is observed. However, the good agreement observed between the experimental $\log K_p$ of the series of hormones and our results, suggests that the proposed values for the pesticide series should not be very different from the experimental ones.

5. Conclusions

Human skin permeability coefficients can be conveniently estimated by chromatographic measurements as long as suitable surrogate chromatographic systems are identified, which has been studied in this work. The ability of chromatographic systems to model human skin permeation has been shown to be well predicted by addition of the error of the skin permeation data, the error of the chromatographic data, and the error due to the dissimilarity between the human skin permeation and the chromatographic systems. These three contributions of error are easily estimated by the solvation parameter model approach. The performances of several chromatographic systems to mimic human skin permeation have been predicted by this procedure and tested by experimental correlations. Poor human skin permeation-chromatographic correlation models have been both predicted and obtained experimentally. Nevertheless, the correlations have been very much improved after introducing the solute's volume as an additional variable, resulting then in correlation models with good predictive abilities to estimate K_p data for untested solutes. These results are in agreement with other author's studies and illustrate the key processes of human skin permeation: partition and diffusion, which are modeled as a combination of chromatographic retention and solute's volume descriptors. With volume correction, the conventional HPLC systems based on C18 columns show very good performances in modeling human skin permeation, the SDS MEKC and the octanol–water partition systems are also quite good, and the IAM HPLC and the STC MEKC, among the studied systems, provide the least precise estimations of K_p data but still acceptable precisions.

The best surrogate system, an HPLC system with a C18 column, has been chosen to establish a methodology for the prediction of human skin permeability of untested compounds by chromatographic measurements. This methodology, based on easy, fast, and cheap experimental measurements, has been validated by training and tests sets and has been finally used to predict the skin permeation of several hormones and pesticides. The K_p values have been predicted with the expected precision using descriptors with fully chemical sense, and they are in agreement with the K_p data reported by other authors. In addition, K_p values estimated in this way have been proposed for some pesticides.

Acknowledgements

The authors thank the Ministerio de Ciencia e Innovación of the Spanish Government and the Fondo Europeo de Desarrollo Regional (FEDER) of the European Union (Project CTQ2010-19217/BQU) for financial support. Furthermore, M.H.-R. thanks the Ministerio de Educación of the Spanish Government for a supporting scholarship (AP2007-01109).

References

Abraham, M.H., 1993. Scales of solute hydrogen-bonding: their construction and application to physicochemical and biochemical processes. Chem. Soc. Rev. 22, 73–83.
Abraham, M.H., Martins, F., 2004. Human skin permeation and partition: general linear free-energy relationship analyses. J. Pharm. Sci. 93, 1508–1523.
Abraham, M.H., McGowan, J.C., 1987. The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography. Chromatographia 23, 243–246.
Abraham, M.H., Råfols, C., 1995. Factors that influence tadpole narcosis. An LFER analysis. J. Chem. Soc. Perkin Trans. 2, 1843–1851.
Abraham, M.H., Chadha, H.S., Whiting, G.S., Mitchell, R.C., 1994. Hydrogen bonding. 32. An analysis of water-octanol and water-alkane partitioning and the delta log P parameter of seiler. J. Pharm. Sci. 83, 1085–1100.
Abraham, M.H., Chadha, H.S., Dixon, J.P., Råfols, C., Treiner, C., 1995a. Hydrogen bonding. Part 40. Factors that influence the distribution of solutes between water and sodium dodecylsulfate micelles. J. Chem. Soc. Perkin Trans. 2, 887–894.
Abraham, M.H., Chadha, H.S., Mitchell, R.C., 1995b. The factors that influence skin penetration of solutes. J. Pharm. Pharmacol. 47, 8–16.
Abraham, M.H., Martins, F., Mitchell, R.C., 1997. Algorithms for skin permeability using hydrogen bond descriptors: the problem of steroids. J. Pharm. Pharmacol. 49, 858–865.
Abraham, M.H., Chadha, H.S., Martins, F., Mitchell, R.C., Bradbury, M.W., Gratton, J.A., 1999. Hydrogen bonding part 46: a review of the correlation and prediction of transport properties by an LFER method: physicochemical properties, brain penetration and skin permeability. Pestic. Sci. 55, 78–88.
Abraham, M.H., Ibrahim, A., Zissimos, A.M., 2004. Determination of sets of solute descriptors from chromatographic measurements. J. Chromatogr. A 1037, 29–47.
ACD/Percepta. <http://www.acdlabs.com> Advanced Chemistry Development, Inc., Release (Build 2203, 29 January 2013).
Barbato, F., Cappello, B., Miro, A., La Rotonda, M., Quaglia, F., 1998. Chromatographic indexes on immobilized artificial membranes for the prediction of transdermal transport of drugs. Il Farmaco 53, 655–661.
Barratt, M.D., 1995. Quantitative structure–activity relationships for skin permeability. Toxicol. In Vitro 9, 27–37.
Bio-Loom Online Database. <http://www.biobyte.com>.
Buchwald, P., Bodor, N., 2001. A simple, predictive, structure-based skin permeability model. J. Pharm. Pharmacol. 53, 1087–1098.
Chen, L., Lian, G., Han, L., 2007. Prediction of human skin permeability using artificial neural network (ANN) modeling. Acta Pharmacol. Sin. 28, 591–600.
Cimpean, D.M., Poole, C.F., 2002. Systematic search for surrogate chromatographic models of biopartitioning processes. Analyst 127, 724–729.
Cronin, M.T.D., 2006. The prediction of skin permeability using quantitative structure–activity relationship methods. In: Riviere, J.E. (Ed.), Dermal Absorption Models in Toxicology and Pharmacology. CRC Press, Boca Raton, pp. 113–134.
Cronin, M.T.D., Dearden, J.C., Moss, G.P., Murray-Dickson, G., 1999. Investigation of the mechanism of flux across human skin in vitro by quantitative structure–permeability relationships. Eur. J. Pharm. Sci. 7, 325–330.
Degim, T., Hadgraft, J., Ilbasmis, S., Ozkan, Y., 2003. Prediction of skin penetration using artificial neural network (ANN) modeling. J. Pharm. Sci. 92, 656–664.
EDETOX. <http://www.ncl.ac.uk/edetox>. Evaluations and Predictions of Dermal Absorption of Toxic Chemicals Website. European Union Framework V: Quality of Life, Environment and Health Key Action Funding (Project No. QLKA-2000-00196).
Escuder-Gilabert, L., Martínez-Pla, J.J., Sagrado, S., Villanueva-Camañas, R.M., Medina-Hernández, M.J., 2003. Biopartitioning micellar separation methods: modelling drug absorption. J. Chromatogr. B 797, 21–35.
Fitzpatrick, D., Corish, J., Hayes, B., 2004. Modelling skin permeability in risk assessment––the future. Chemosphere 55, 1309–1314.
Fitzpatrick, D., Golden, D., Corish, J., 2008. Modeling skin permeability in risk assessment. Drugs Pharm. Sci. 177, 287–297.
Flynn, G.L., 1990. Physicochemical determinants of skin absorption. In: Gerrity, T.R., Henry, C.J. (Eds.), Principles of Route-to-Route Extrapolation for Risk Assessment. Elsevier, New York, pp. 93–127.
Fuguet, E., Råfols, C., Bosch, E., Abraham, M.H., Rosés, M., 2002a. Solute-solvent interactions in micellar electrokinetic chromatography. III. Characterization of the selectivity of micellar electrokinetic chromatography systems. J. Chromatogr. A 942, 237–248.

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

ARTICLE IN PRESS

Fuguet, E., Ràfols, C., Bosch, E., Rosés, M., 2002b. Solute-solvent interactions in micellar electrokinetic chromatography: IV. Characterization of electroosmotic flow and micellar markers. Electrophoresis 23, 56–66.
Geinoz, S., Guy, R.H., Testa, B., Carrupt, P., 2004. Quantitative structure-permeation relationships (QSPeRs) to predict skin permeation: a critical evaluation. Pharm. Res. 21, 83–92.
Hansch, C., Leo, A., 1995. Exploring QSAR, Fundamentals and Applications in Chemistry and Biology. ACS Professional Reference Book. American Chemical Society, Washington.
Hidalgo-Rodríguez, M., Fuguet, E., Ràfols, C., Rosés, M., 2010. Estimation of biological properties by means of chromatographic systems: evaluation of the factors that contribute to the variance of biological-chromatographic correlations. Anal. Chem. 82, 10236–10245.
Hidalgo-Rodríguez, M., Fuguet, E., Ràfols, C., Rosés, M., 2012a. Modeling nonspecific toxicity of organic compounds to the fathead minnow fish by means of chromatographic systems. Anal. Chem. 84, 3446–3452.
Hidalgo-Rodríguez, M., Fuguet, E., Ràfols, C., Rosés, M., 2012b. Performance of chromatographic systems to model soil–water sorption. J. Chromatogr. A 1252, 136–145.
Hoang, K.T., 1992. Dermal Exposure Assessment: Principles and Applications. U.S. Environmental Protection Agency, Office of Health and Environmental Assessment, Washington, DC, EPA/600/8-91/011B.
Hoover, K.R., Acree Jr., W.E., Abraham, M.H., 2005. Chemical toxicity correlations for several fish species based on the Abraham solvation parameter model. Chem. Res. Toxicol. 18, 1497–1505.
Ishihama, Y., Asakawa, N., 1999. Characterization of lipophilicity scales using vectors from solvation energy descriptors. J. Pharm. Sci. 88, 1305–1312.
Johnson, M.E., Blankschtein, D., Langer, R., 1995. Permeation of steroids through human skin. J. Pharm. Sci. 84, 1144–1146.
Kielhorn, J., Melching-Kollmuss, S., Mangelsdorf, I., 2006. Environmental Health Criteria 235: Dermal Absorption. World Health Organization Press, Geneva.
Lázaro, E., Ràfols, C., Abraham, M.H., Rosés, M., 2006. Chromatographic estimation of drug disposition properties by means of immobilized artificial membranes (IAM) and C18 columns. J. Med. Chem. 49, 4861–4870.
Lieb, W.R., Stein, W.D., 1986. Transport and Diffusion Across Cell Membranes. Academic Press, Orlando.
Maibach, H.I., 2007. Dermal absorption models in toxicology and pharmacology. Clin. Toxicol. 45, 736.
Martínez-Pla, J.J., Martín-Biosca, Y., Sagrado, S., Villanueva-Camañas, R.M., Medina-Hernández, M.J., 2003. Biopartitioning micellar chromatography to predict skin permeability. Biomed. Chromatogr. 17, 530–537.
Mitragotri, S., Anissimov, Y.G., Bunge, A.L., Frasch, H.F., Guy, R.H., Hadgraft, J., Kasting, G.B., Lane, M.E., Roberts, M.S., 2011. Mathematical models of skin permeability: an overview. Int. J. Pharm. 418, 115–129.
Moss, G.P., Dearden, J.C., Patel, H., Cronin, M.T.D., 2002. Quantitative structure–permeability relationships (QSPRs) for percutaneous absorption. Toxicol. In Vitro 16, 299–317.
Nasal, A., Sznitowska, M., Buciński, A., Kaliszan, R., 1995. Hydrophobicity parameter from high-performance liquid chromatography on an immobilized artificial membrane column and its relationship to bioactivity. J. Chromatogr. A 692, 83–89.
Neely, B.J., Madihally, S.V., Robinson Jr., R.L., Gasem, K.A.M., 2009. Nonlinear quantitative structure–property relationship modeling of skin permeation coefficient. J. Pharm. Sci. 98, 4069–4084.
Neumann, D., Kohlbacher, O., Merkwirth, C., Lengauer, T., 2006. A fully computational model for predicting percutaneous drug absorption. J. Chem. Inf. Model. 46, 424–429.
Patel, H., Berge, W.T., Cronin, M.T.D., 2002. Quantitative structure–activity relationships (QSARs) for the prediction of skin permeation of exogenous chemicals. Chemosphere 48, 603–613.
Platts, J.A., Abraham, M.H., Zhao, Y.H., Hersey, A., Ijaz, L., Butina, D., 2001. Correlation and prediction of a large blood–brain distribution data set-an LFER study. Eur. J. Med. Chem. 36, 719–730.
Poole, S.K., Poole, C.F., 1999. Chromatographic models for the sorption of neutral organic compounds by soil from water and air. J. Chromatogr. A 845, 381–400.
Poole, C.F., Gunatilleka, A.D., Poole, S.K., 2000. In search of a chromatographic model for biopartitioning. Adv. Chromatogr. 40, 159–230.
Potts, R.O., Guy, R.H., 1992. Predicting skin permeability. Pharm. Res. 9, 663–669.
Potts, R.O., Guy, R.H., 1995. A predictive algorithm for skin permeability: the effects of molecular size and hydrogen bond activity. Pharm. Res. 12, 1628–1633.
Rosés, M., Ràfols, C., Bosch, E., Martínez, A.M., Abraham, M.H., 1999. Solute-solvent interactions in micellar electrokinetic chromatography. Characterization of sodium dodecyl sulfate-Brij 35 micellar systems for quantitative structure-activity relationship modelling. J. Chromatogr. A 845, 217–226.
Sangster, J., 1997. Octanol-Water Partition Coefficients: Fundamentals and Physical Chemistry. Wiley, Chichester.
Smith, R.N., Hansch, C., Ames, M.M., 1975. Selection of a reference partitioning system for drug design work. J. Pharm. Sci. 64, 599–606.
Tezel, A., Mitragotri, S., 2003. On the origin of size-dependent tortuosity for permeation of hydrophilic solutes across the stratum corneum. J. Control. Release 86, 183–186.
Turowskil, M., Kaliszan, R., 1997. Keratin immobilized on silica as a new stationary phase for chromatographic modelling of skin permeation. J. Pharm. Biomed. Anal. 15, 1325–1333.
Vecchia, B.E., Bunge, A.L., 2003. Skin absorption databases and predictive equations. In: Guy, R., Hadgraft, J. (Eds.), Transdermal Drug Delivery. Marcel Dekker, New York, pp. 57–141.
Vitha, M., Carr, P.W., 2006. The chemical interpretation and practice of linear solvation energy relationships in chromatography. J. Chromatogr. A 1126, 143–194.
Walker, J.D., Rodford, R., Patlewicz, G., 2003. Quantitative structure–activity relationships for predicting percutaneous absorption rates. Environ. Toxicol. Chem. 22, 1870–1884.
Wang, Y., Sun, J., Liu, H., Liu, J., Zhang, L., Liu, K., He, Z., 2009. Predicting skin permeability using liposome electrokinetic chromatography. Analyst 134, 267–272.
Wilschut, A., ten Berge, W.F., Robinson, P.J., McKone, T.E., 1995. Estimating skin permeation. The validation of five mathematical skin permeation models. Chemosphere 30, 1275–1296.
Zhang, K., Chen, M., Scriba, G.K.E., Abraham, M.H., Fahr, A., Liu, X., 2012. Human skin permeation of neutral species and ionic species: extended linear free-energy relationship analyses. J. Pharm. Sci. 101, 2034–2044.

Please cite this article in press as: Hidalgo-Rodríguez, M., et al. Evaluation of the suitability of chromatographic systems to predict human skin permeation of neutral compounds. Eur. J. Pharm. Sci. (2013), http://dx.doi.org/10.1016/j.ejps.2013.04.005

RESULTATS I DISCUSSIÓ

3.1. CARACTERITZACIÓ DE SISTEMES DE MEKC MITJANÇANT L'EQUACIÓ GENERAL DE SOLVATACIÓ

Des del seu desenvolupament [18], el model de paràmetres de solvatació ha estat àmpliament utilitzat per caracteritzar sistemes cromatogràfics i obtenir així informació de les interaccions solut-solvent que s'estableixen [102]. Disposar de sistemes cromatogràfics ben caracteritzats permet conèixer el tipus i la força de les interaccions que controlen la retenció i la selectivitat i, per tant, resulta molt útil per millorar separacions cromatogràfiques de compostos. A la vegada, gràcies a que aquest model també permet descriure processos de biopartició, una altra gran utilitat és que ofereix un marc comú per comparar els sistemes cromatogràfics amb els de biopartició. D'aquesta manera es converteix en una eina clau per trobar sistemes cromatogràfics amb característiques similars a sistemes de biopartició, de manera que sigui possible estimar propietats biològiques a partir de mesures cromatogràfiques en sistemes anàlegs.

En aquest context, part d'aquesta tesi s'ha dedicat a ampliar el coneixement de sistemes de cromatografia electrocinètica micel·lar a partir de la seva caracterització mitjançant el model de paràmetres de solvatació. Concretament, s'han estudiat els següents sistemes: d'una banda sistemes formats per fases pseudoestacionàries mixtes dels tensioactius colat i deoxicolat de sodi (SC i SDC respectivament); d'altra banda el sistema micel·lar format pel tensioactiu taurocolat de sodi (STC). En els tres casos es tracta de tensioactius aniònics que pertanyen al grup de les sals biliars. Un dels motius per triar-los és que les sals biliars probablement constitueixen un dels grups de tensioactius més utilitzats en MEKC després del dodecilsulfat de sodi (SDS), de manera que interessa aportar noves dades sobre ells.

3.1.1. Sistemes amb fases micel·lars mixtes de SC i SDC

L'ús de mescles de tensioactius en MEKC és molt interessant de cara a obtenir sistemes amb diferent selectivitat. Les propietats de la fase pseudoestacionària no tan sols es poden modificar pel simple fet de mesclar diversos tensioactius, sinó que també la diferent proporció d'aquests tensioactius a la mescla pot donar lloc a sistemes amb selectivitats variades. Per tant, l'ús de les fases micel·lars mixtes obre el ventall de selectivitats que es poden obtenir en MEKC, que és un aspecte fonamental per tal de trobar sistemes adequats per a separacions cromatogràfiques o per a l'emulació de sistemes de biopartició.

Són diversos els autors que s'han interessat en aquest tema i tenen publicacions al respecte [70,75,120,128-130]. Seguint aquesta línia de recerca, i també amb la idea de complementar una sèrie de treballs publicats prèviament sobre la caracterització i selectivitat de sistemes de MEKC [70,80,118-120,131], en aquesta tesi s'han decidit caracteritzar sistemes mixtos de SC i SDC en MEKC a partir del model de paràmetres de solvatació (article I).

En aquest estudi, d'una banda s'ha avaluat la variació de les propietats de solvatació en sistemes de MEKC amb diferent contingut de SC i SDC. D'altra banda, també s'ha avaluat la influència de diversos factors en la caracterització dels sistemes, concretament el conjunt de soluts, el pH de la fase aquosa i els descriptors. A continuació s'exposaran els trets més rellevants d'aquest treball.

Els tensioactius estudiats, el SC i SDC, són les sals de dos dels principals àcids biliars, l'àcid còlic i deoxicòlic respectivament. Ambdós tensioactius tenen un carboxilat com a grup polar, fet que els converteix en aniònics. Cal remarcar que la seva estructura deriva del colesterol i són per tant esteroides. Cadascun d'ells conté 24 àtoms de carboni, trihidroxilats en el cas del SC i dihidroxilats en el cas del SDC, de manera que es diferencien en un grup hidroxil (Figura 1, article I).

La caracterització de sistemes de SC i SDC ha comprès l'estudi tant dels tensioactius purs com de diverses mescles de tots dos. La majoria de sistemes han estat caracteritzats a pH 8, donada la baixa solubilitat del SDC a pH 7 (pH al qual es caracteritzen normalment els sistemes de MEKC). El sistema de SC pur ha estat

caracteritzat a pH 6 i pH 7, a part de també a pH 8. En total, s'han caracteritzat 7 sistemes (Taula 4), 3 dels quals es diferencien en el pH de la fase aquosa (sistemes 1-3) i 5 dels quals es diferencien en la composició i concentració dels surfactants (sistemes 3-7). El grup de substàncies per caracteritzar els sistemes (Taula 1, article I) s'ha escollit d'acord amb un treball previ [118] en el qual es va comprovar que era un conjunt adequat perquè complia els requisits detallats a l'apartat 1.3.2 de la introducció.

Taula 4. Sistemes de SC i SDC caracteritzats.

Sistema	SC	SDC	x_{SDC}	pH aquós
1	80 mM	-	0.00	pH 6
2	80 mM	-	0.00	pH 7
3	80 mM	-	0.00	pH 8
4	60 mM	10 mM	0.14	pH 8
5	40 mM	20 mM	0.33	pH 8
6	20 mM	30 mM	0.60	pH 8
7	-	40 mM	1.00	pH 8

La caracterització del tensioactiu pur de SC variant el pH de la fase aquosa (sistemes 1-3 de la Taula 4) ha posat de manifest tant la influència del pH com el paper que juguen els soluts fets servir a la caracterització. A la Taula 2 de l'article I es mostren els coeficients LSER obtinguts per a cadascun d'aquests tres sistemes.

Inicialment, el SC es va caracteritzar a pH 8 sense incloure 6 fenols del conjunt de soluts de caracterització, pensant en la seva susceptibilitat d'estar ionitzats a aquest pH de treball. Es tracta dels fenols detallats a la Taula 5. En canvi, aquestes substàncies sí que van ser incloses a les caracteritzacions de pH 6 i pH 7, donat que són clarament neutres a aquests valors de pH. Comparant les caracteritzacions obtingudes (a, e i f de la Taula 2 de l'article I), s'observa que les de pH 6 i pH 7 són pràcticament idèntiques, però en canvi hi ha diferències respecte a la de pH 8. La discrepància més notable es troba en el coeficient *a*, que té un valor aproximadament un 50% superior a la caracterització de pH 8 (caracterització a), en comparació amb les de pH 6 i pH 7 (caracteritzacions e i f).

Taula 5. Fenols amb pK_a 9-10, susceptibles d'estar ionitzats a pH 8, amb els seus corresponents descriptors d'Abraham.

Compost	Descriptors d'Abraham				
	E	***S***	***A***	***B***	***V***
4-Clorofenol	0.915	1.08	0.67	0.20	0.8975
Catecol	0.970	1.10	0.88	0.47	0.8338
Resorcinol	0.980	1.00	1.10	0.58	0.8338
Hidroquinona	1.000	1.00	1.16	0.60	0.8338
2-Naftol	1.520	1.08	0.61	0.40	1.1441
1,2,3-Trihidroxibenzè	1.165	1.35	1.35	0.62	0.8925

Per tal d'avaluar només l'efecte del pH, es van calcular novament les caracteritzacions de pH 6 i pH 7 excloent els sis fenols esmentats. Els resultats llavors mostren que les caracteritzacions que no inclouen aquests fenols (correlacions a, b i c de la Taula 2 de l'article I) tenen els mateixos coeficients LSER independentment del pH de treball (pH 8, 7 i 6 respectivament). En conseqüència, queda així reflectit que el pH aquós per si sol no té pràcticament influència en la selectivitat dels sistemes de MEKC, la qual cosa corrobora el que reporten altres autors [78].

D'altra banda, aquests resultats posen de manifest la importància del conjunt de soluts amb els que es caracteritza un sistema mitjançant el model de paràmetres de solvatació. Un cop comprovat que el pH no influeix, es pot concloure que els 6 fenols de la Taula 5 són aleshores els responsables de les diferències que s'observaven inicialment entre els coeficients de SC a pH 8 i els coeficients de SC a pH 7 i pH 6 (correlacions a, d i e de la Taula 2 de l'article I, respectivament). Fent un cop d'ull al conjunt de soluts de caracterització i als seus descriptors (Taula 1, article I), es pot observar que aquests fenols són dels compostos amb més acidesa per pont d'hidrogen, és a dir, amb un descriptor *A* més elevat. Concretament, el valor de *A* segons la base de dades d'Abraham [107-109] es troba entre 0.61 i 1.35 per a aquests fenols (veure Taula 5), mentre que és inferior o directament zero per a quasi tota la resta de soluts. Això implica que si els 6 fenols mencionats no s'inclouen a les

caracteritzacions, el descriptor *A* no comprèn un interval prou representatiu de valors. En conseqüència, disminueix l'exactitud amb què es determinen els coeficients de solvatació. Per tant, queda il·lustrada la importància de caracteritzar els sistemes amb un grup de substàncies que cobreixin un ampli interval de valors pel que fa als descriptors.

Aquests fets han portat a investigar si els fenols de la Taula 5 es troben significativament ionitzats a pH 8, ja que en cas de trobar-se neutres interessa incloure'ls a les caracteritzacions per tal d'obtenir coeficients els més acurats possible. Per tal de comprovar-ho, aquests fenols van ser injectats al sistema de SC a pH 8. Seguidament, els seus valors de log k (juntament amb els de la resta de substàncies del conjunt de caracterització) van ser representats respecte als obtinguts en els sistemes de SC a pH 6 i pH 7 (valors de pH als quals són clarament neutres). A la Figura 2 de l'article I es pot observar que només el catecol i el 1,2,3-trihidroxibenzè es desvien inicialment de la correlació. En tots dos casos, però, aquesta desviació desapareix quan es deixa d'emprar tampó de borat per preparar el sistema de SC a pH 8. Es comprova així que la desviació inicial d'ambdós compostos es devia al complex que formen els seus hidroxils veïnals amb el tampó de borat, un fenomen ja reportat en vàries ocasions a la literatura [132,133]. En definitiva, la bona correlació entre la retenció dels fenols a pH 8 i la seva retenció a pH 7 i pH 6 indica que no es troben significativament ionitzats a pH 8. Conseqüentment, aquests fenols seran inclosos a les següents caracteritzacions.

Un altre dels aspectes estudiats ha estat la influència del valor dels descriptors sobre les caracteritzacions. Per això, els 5 sistemes de SC i SDC corresponents als números 3-7 de la Taula 4 s'han caracteritzat emprant dos conjunts de descriptors. El primer conjunt prové de la base de dades construïda per Abraham i col·laboradors [107-109], la qual disposa de valors coneguts de tots o alguns descriptors de solvatació per a milers de substàncies. El segon conjunt de descriptors correspon a valors de descriptors optimitzats per Poole i col·laboradors [134] a través de mètodes cromatogràfics (mitjançant sistemes ben caracteritzats de cromatografia de gasos, cromatografia de líquids en fase invertida, cromatografia electrocinètica micel·lar i partició líquid-líquid), per tal d'obtenir descriptors amb menor incertesa. A la Taula 1

de l'article I es detallen els valors disponibles de cada conjunt de descriptors pels soluts de caracterització.

Els coeficients de solvatació obtinguts per a cada sistema de SC-SDC després de correlacionar els valors de log k i els descriptors d'Abraham de les substàncies es mostren a la Taula 3 de l'article I. A la Taula 4 del mateix article es mostren els coeficients obtinguts emprant els descriptors optimitzats de Poole, disponibles només per a 40 soluts del conjunt de caracterització (pels 29 soluts restants cal emprar els descriptors d'Abraham). Si es comparen les Taules 3 i 4, els coeficients obtinguts per a cada sistema són semblants i condueixen a la mateixa interpretació química, la qual serà comentada més endavant. Pel que fa als valors estadístics, els ajustos són lleugerament millors quan es fan servir els descriptors optimitzats. De fet, si es tornen a calcular les correlacions però incloent només els 40 soluts amb descriptors optimitzats disponibles, s'observa més clarament que les estadístiques són millors (valors considerablement inferiors de SD i superiors de *F*) quan es fan servir els descriptors optimitzats (Taula 6).

Taula 6. Comparació dels intervals obtinguts per a SD i *F* a les caracteritzacions dels cinc sistemes estudiats de SC-SDC, incloent només les substàncies amb descriptors optimitzats disponibles (*n* = 40) i emprant ambdós conjunts de descriptors.

Paràmetre estadístic	**Conjunt de descriptors utilitzats**	
	Descriptors d'Abraham	**Descriptors optimitzats de Poole**
SD	0.190 - 0.212	0.127 - 0.151
F	71 - 83	144 - 197

Aquests resultats són totalment coherents amb el treball efectuat per Poole i col·laboradors, que van decidir optimitzar valors de descriptors per tal de poder obtenir millors ajustos de les dades a l'equació general de solvatació. No obstant, es posa de manifest la necessitat de tenir a l'abast valors de descriptors optimitzats per a més soluts, ja que en aquest estudi els 40 compostos amb descriptors optimitzats disponibles no formen un grup prou representatiu de substàncies. Per això, a

continuació es discutiran les caracteritzacions dels sistemes de SC i SDC en base a les correlacions que comprenen tot el conjunt de soluts (Taules 3 i 4 de l'article I).

Quan s'interpreten químicament els coeficients LSER obtinguts a cada sistema, es poden observar les següents característiques comunes de solvatació independentment de la composició de SC i SDC:

- La fase aquosa i pseudoestacionària es diferencien principalment en la seva hidrofobicitat i acidesa per pont d'hidrogen (*v* i *b* són els coeficients de magnitud més alta), mentre que ambdues fases presenten aproximadament la mateixa basicitat per pont d'hidrogen (*a* és el coeficient de menor valor absolut, molt proper a 0). Això significa que les propietats del solut que més influeixen en la seva retenció són el volum (*V*) i la basicitat per pont d'hidrogen (*B*), mentre que l'acidesa per pont d'hidrogen (*A*) quasi no hi afecta.

- Cada coeficient té el mateix signe en tots els sistemes estudiats, de manera que cada descriptor influencia sempre la retenció en el mateix sentit independentment del contingut en SC i SDC.

- D'acord amb el signe dels coeficients, la retenció d'un compost augmentarà com més gran siguin la seva refractivitat molar en excés (*E*), acidesa per pont d'hidrogen (*A*) i volum (*V*) (els coeficients *e*, *a* i *v* són positius i indiquen que es veurà afavorida la partició cap a la fase micel·lar, augmentant així la retenció del solut en el sistema). Contràriament, els soluts menys retinguts seran aquells amb valors alts dels descriptors *B* i *S* (*b* i *s* són negatius i per tant els soluts tendiran a anar cap a la fase aquosa).

En general, aquestes observacions sobre el comportament dels soluts en tots els sistemes de SC i SDC estudiats concorden amb la tendència observada a la majoria de processos fisicoquímics de partició (apartat 1.3.2.2 de la introducció).

A la Taula 7 es poden comparar les caracteritzacions obtingudes en aquesta tesi (calculades amb els descriptors d'Abraham, veure article I) pels dos sistemes purs de SC i SDC (sistemes 3 i 7 de la Taula 4) amb les caracteritzacions d'altres autors. Cal remarcar que encara que s'hagin dut a terme en diferents condicions experimentals, les caracteritzacions són comparables perquè la selectivitat dels sistemes de MEKC ve

principalment determinada per l'elecció del tensioactiu (condicions experimentals com el pH, la composició del tampó o la concentració del tensioactiu tenen poca influència) [78].

Taula 7. Comparació de les caracteritzacions obtingudes per diferents autors en sistemes de MEKC de SC i de SDC purs (desviacions estàndard entre parèntesi).

Sistema	Coeficients						Estadístics				Ref.
	c	***e***	***s***	***a***	***b***	***v***	***n***	***r***	**SD**	***F***	
SC	-1.501 (0.101)	0.910 (0.152)	-0.776 (0.116)	0.109 (0.096)	-1.961 (0.126)	2.255 (0.112)	66	0.958	0.234	133	Article I
	-1.408 (0.085)	0.691 (0.127)	-0.693 (0.097)	0.117 (0.078)	-1.937 (0.111)	2.274 (0.104)	62	0.962	0.191	139	[118, 119]
	-1.229 (0.074)	0.324 (0.064)	-0.392 (0.080)	-*	-2.052 (0.103)	2.173 (0.125)	45	0.978	0.17	220	[135]
	-1.82	0.55 (0.08)	-0.65 (0.08)	0.08 (0.05)	-2.51 (0.10)	2.75 (0.12)	36	0.992	0.07		[130]
	-1.48 (0.11)	0.50 (0.08)	-0.76 (0.10)	0.12 (0.06)	-2.52 (0.11)	2.66 (0.13)	59	0.986	0.10	294	[79]
	-1.71 (0.12)	0.63 (0.08)	-0.47 (0.09)	-*	-2.29 (0.13)	2.45 (0.12)	40	0.983	0.110	241	[78]
SDC	-1.786 (0.112)	0.813 (0.170)	-0.803 (0.125)	0.061 (0.103)	-1.771 (0.140)	2.400 (0.127)	64	0.947	0.255	102	Article I
	-1.833 (0.120)	0.926 (0.202)	-0.867 (0.146)	0.070 (0.143)	-1.785 (0.146)	2.422 (0.130)	58	0.949	0.260	94	[118, 119]
	-1.69 (0.12)	0.66 (0.08)	-0.47 (0.09)	-*	-2.47 (0.13)	2.67 (0.11)	40	0.986	0.11	286	[78]

* Coeficient no significatiu al 95% de nivell de confiança

Pel que fa al sistema de SC pur, el coeficient *e* obtingut en aquest treball és el més elevat dels de la Taula 7, mentre que la resta de coeficients tenen valors força semblants a alguns dels publicats per altres autors (els coeficients *s* i *a* són similars als

de les referències [79,118,130] i els coeficients *b* i *v* són similars als de les referències [118,135]).

Els coeficients obtinguts pel sistema de SDC pur guarden molta semblança amb els publicats a la referència [118] (aquesta mateixa referència també conté la caracterització de SC més similar a la de l'article I). Aquest fet es pot atribuir a que aquests sistemes s'han caracteritzat pràcticament amb el mateix conjunt de substàncies, que és un conjunt ampli i variat. La resta de caracteritzacions d'altres autors inclouen un número més reduït i/o menys representatiu de soluts i d'aquí que hi hagi més diferències en el valor dels coeficients. Per tant, es posa novament de manifest la rellevància de les substàncies fetes servir per caracteritzar sistemes mitjançant l'equació general de solvatació.

En general, però, es pot concloure que els coeficients de solvatació obtinguts pel SC i SDC purs (article I) descriuen el mateix comportament (igual signe) que les caracteritzacions extretes de la bibliografia i mostrades a la Taula 7, de manera que són resultats concordants amb els d'altres autors.

Finalment, per obtenir informació de les diferències i similituds entre les propietats dels dos tensioactius, cal comparar els coeficients obtinguts pels dos sistemes purs i els tres sistemes mixtos de SC i SDC (sistemes 3-7 de la Taula 4). A la Figura 14 es poden apreciar visualment les variacions dels coeficients en funció de la composició de la fase micel·lar en els 5 sistemes estudiats.

D'acord amb la figura, el coeficient *b* és el que presenta un major increment a mesura que augmenta la composició de SDC, el que indica que el SDC és més àcid per pont d'hidrogen que el SC. També destaca el increment del coeficient *v*, el que reflecteix que els sistemes més rics en SDC són els que tenen més facilitat per allotjar el solut, tot indicant que el tensioactiu SDC és més hidrofòbic que el SC. Pel que fa al coeficient *e*, s'observa que experimenta una lleugera disminució, de manera que la fase micel·lar es torna menys polaritzable quan augmenta la fracció molar de SDC. Aquest fet suggereix que el SDC és lleugerament menys polaritzable que el SC. D'altra banda, quasi no s'observa variació dels coeficients *s* i *a* en funció de la composició de la fase micel·lar, del que es pot deduir que el SC i el SDC tenen una dipolaritat i una basicitat per pont d'hidrogen molt similars.

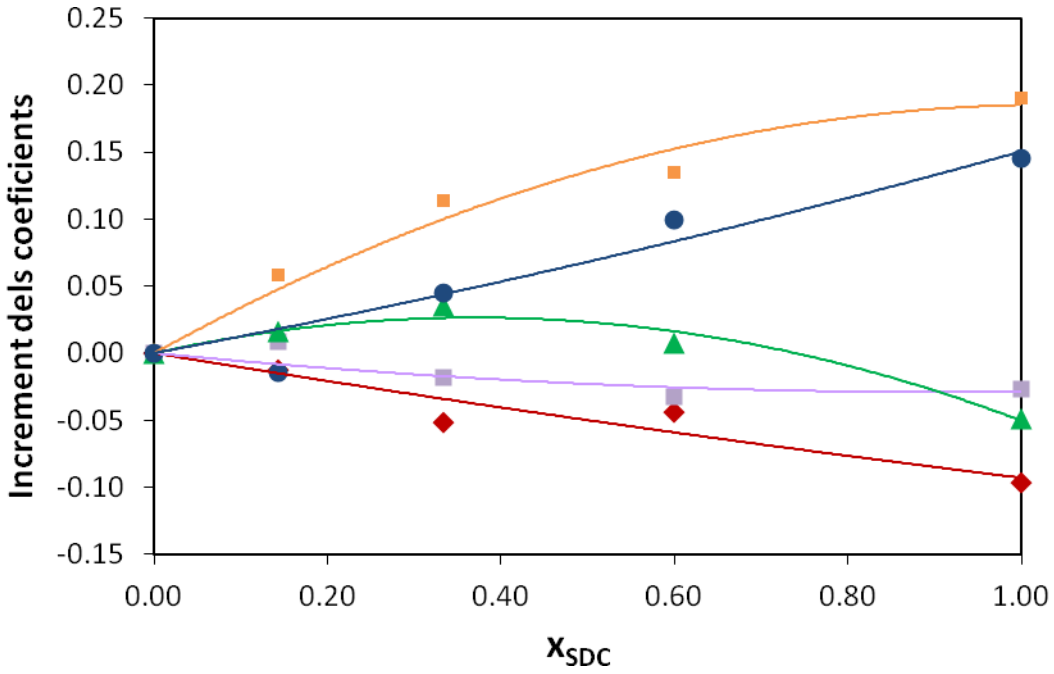


Figura 14. Increment de cada coeficient en funció de la fracció molar de SDC en els sistemes estudiats (coeficients calculats amb els descriptors d'Abraham): Δ*e* (◆), Δ*s* (■), Δ*a* (▲), Δ*b* (■) i Δ*v* (●).

Tot i haver obtingut variacions modestes dels coeficients en els sistemes estudiats (la qual cosa ja concorda amb la similar naturalesa dels dos tensioactius emprats), aquestes variacions reflecteixen que la diferent composició de SC i SDC en fases micel·lars mixtes afecta a la selectivitat. Segons els resultats, per separar mescles de compostos amb distinta basicitat per pont d'hidrogen i mida (diferent valor dels descriptors *B* i *V* respectivament), seran més adequats els sistemes que contenen més SDC a la fase micel·lar. Aquesta conclusió prové d'observar que les interaccions que es veuen més afectades a l'afegir SDC a les micel·les de SC són l'acidesa per pont d'hidrogen i la hidrofobicitat de la fase micel·lar (*b* i *v* respectivament), les quals augmenten amb el contingut de SDC.

3.1.2. Sistema de STC

El taurocolat de sodi, STC, és un tensioactiu que també pertany al grup de les sals biliars. Prové de l'àcid taurocòlic, que a la seva vegada prové de l'àcid còlic conjugat amb l'aminoàcid taurina. Es tracta d'un tensioactiu aniònic que té un grup sulfonat com a grup polar (Figura 15).

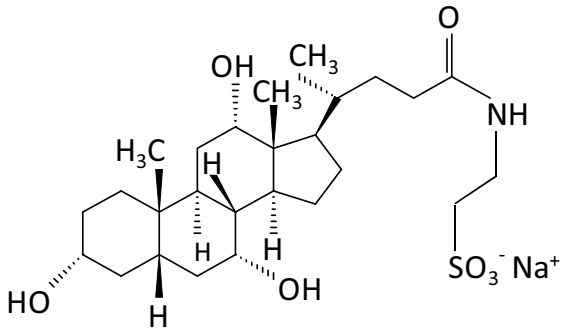


Figura 15. Estructura del taurocolat de sodi (STC).

El sistema de MEKC format per STC va ser caracteritzat per Poole i col·laboradors [78] anys enrere, obtenint els coeficients detallats a la Taula 8.

La caracterització de Poole i col·laboradors s'ha fet servir inicialment per comparar la similitud d'aquest sistema cromatogràfic amb sistemes de biopartició de naturalesa diversa a partir del càlcul de la distància *d* (veure apartat 1.4 de la introducció). Alguns exemples dels valors de *d* obtinguts es mostren a la Taula 6 de l'article II. Es pot observar que es tracta de distàncies petites ($d < 0.25$) que suggereixen que aquest sistema de MEKC guarda similitud amb sistemes de biopartició i, per tant, sembla tenir potencial per emular-los.

Aquestes circumstàncies han conduït a caracteritzar de nou el sistema de STC, però amb un grup de soluts més gran i variat que el que van utilitzar Poole i col·laboradors. Concretament, amb el conjunt de substàncies ja emprades a les caracteritzacions dels sistemes de SC i SDC de l'apartat anterior (donat que es va publicar que aquest conjunt era adequat per caracteritzar sistemes de MEKC mitjançant l'equació general de solvatació [118]). La intenció és obtenir uns coeficients LSER més representatius

d'aquest sistema de MEKC, per poder així avaluar amb més fiabilitat la seva capacitat per modelar sistemes de biopartició.

Inicialment, el sistema de STC ha estat caracteritzat reproduint les condicions experimentals de concentració de tensioactiu (STC 50 mM) i pH aquós (pH 8) emprades per Poole i col·laboradors [78]. A la Taula 8 es mostren els coeficients de solvatació obtinguts (STC pH 8). La caracterització consta de 55 soluts, després d'haver exclòs aquells que coeluïen (1-propanol, 2-propanol, 1,3-propandiol, 1,4-butandiol, 1,5-pentandiol, pirimidina, mircè, α-pinè) amb el marcador de flux electroosmòtic (metanol) o marcador micel·lar (dodecanofenona) i també aquells considerats com a aberrants (4-aminobenzamida, tiourea, cafeïna, antipirina, benzamida, acetanilida) per tenir residus estàndard superiors a l2.5l a la regressió. De fet, aquests són els criteris que es van fer servir també per excloure compostos a les caracteritzacions de SC i SDC de l'apartat anterior.

Addicionalment, el sistema de MEKC de STC també ha estat caracteritzat a pH 7 (amb la mateixa concentració de tensioactiu, 50 mM). Tot i haver comprovat a l'apartat anterior que el pH no té gaire influència a les caracteritzacions, s'ha volgut determinar qualsevol petita variació en els coeficients. El primer i principal motiu pel qual s'ha cregut convenient disposar de la caracterització a pH 7 és perquè a aquest pH és possible incloure compostos àcids per pont d'hidrogen sense problemes d'ionització (per exemple, fenols amb pK_a 8-9, que seran neutres a pH 7 però estaran ionitzats o parcialment ionitzats a valors de pH superiors). Tal com s'ha vist a l'apartat anterior, convé incloure aquest tipus de substàncies (que aporten valors del descriptor *A*) per obtenir resultats representatius. El segon motiu és que pH 7 és el pH més emprat per reproduir el pH fisiològic (pH 7.4) al qual tenen lloc alguns processos de biopartició. Per tant, sembla raonable treballar a pH 7 de cara a estudiar l'emulació d'aquest tipus de processos.

La caracterització de STC a pH 7 també es troba detallada a la Taula 8. En aquest cas, també s'han exclòs substàncies degut a coelució (2-propanol, 1,3-propandiol, 1,4-butandiol, pirimidina) i per resultar punts aberrants a la regressió (tiourea, α-pinè, 4-aminobenzamida, mircè, butilbenzè, cafeïna, antipirina, heptanofenona, benzamida).

Taula 8. Comparació de la caracterització del sistema de MEKC de STC obtinguda per Poole i col·laboradors amb les caracteritzacions obtingudes en aquesta tesi (desviacions estàndard entre parèntesi).

Sistema	Coeficients						Estadístics				Ref.
	c	***e***	***s***	***a***	***b***	***v***	***n***	***r***	**SD**	***F***	
STC pH 8 (Poole)	-2.10 (0.09)	0.60 (0.07)	-0.34 (0.07)	-*	-2.06 (0.10)	2.43 (0.09)	40	0.989	0.09	377	[78]
STC pH 8	-1.939 (0.054)	0.700 (0.081)	-0.561 (0.059)	0.356 (0.048)	-2.534 (0.094)	2.493 (0.071)	55	0.986	0.104	346	No publicat
STC pH 7	-1.838 (0.053)	0.613 (0.085)	-0.513 (0.064)	0.319 (0.053)	-2.396 (0.100)	2.390 (0.084)	56	0.984	0.112	305	Article III

* Coeficient no significatiu al 95% de nivell de confiança

Comparant les tres caracteritzacions de la Taula 8, totes presenten el mateix comportament pel que fa al signe dels coeficients. El signe positiu dels coeficients *e*, *a* (excepte a la caracterització de Poole en que *a* no és significatiu) i *v* indiquen, respectivament, que la fase micel·lar té més capacitat per interaccionar amb els electrons –n i –π del solut, és més bàsica per pont d'hidrogen i més hidrofòbica que la fase aquosa. El signe negatiu dels coeficients *s* i *b* reflecteix, respectivament, que la fase micel·lar té menys habilitat per establir interaccions tipus dipol-dipol i dipol-dipol induït amb el solut i que és menys àcida per pont d'hidrogen que la fase aquosa. En concordança amb els signes dels coeficients, la retenció d'un compost augmentarà com més grans siguin els seus descriptors *E*, *A* i *V*, ja que es veurà afavorida la seva partició cap a la fase micel·lar. En canvi, soluts amb valors alts dels descriptors *S* i *B* veuran afavorida la seva partició cap a la fase aquosa i, per tant, la seva retenció disminuirà.

A part del signe, un altre tret en comú de les tres caracteritzacions està relacionat amb la magnitud dels coeficients. En totes les caracteritzacions, els coeficients *b* i *v* són els de major valor absolut i el coeficient *a* el de menor. Per tant, hi ha concordança en què la basicitat del solut per pont d'hidrogen i el seu volum són les propietats que més

condicionen la seva retenció en el sistema de MEKC, mentre que la seva acidesa per pont d'hidrogen és la que menys.

Sí que s'aprecien diferències entre les caracteritzacions si s'observen en detall els valors dels coeficients. Les diferències són mínimes entre els sistemes STC pH 8 i STC pH 7 caracteritzats en aquesta tesi (tal com s'esperava donada la poca influència del pH ja comentada prèviament). Tot i així, es pot dir que s'han obtingut coeficients en general lleugerament inferiors en valor absolut pel STC a pH 7.

En canvi, sí que s'observen diferències més significatives entre el valor dels coeficients obtinguts per Poole i col·laboradors i els obtinguts en aquesta tesi. La principal diferència es troba en el coeficient *a*, el qual no és significatiu a un 95% de nivell de confiança per a Poole i col·laboradors, tot indicant que la fase micel·lar i aquosa tenen aproximadament la mateixa basicitat per pont d'hidrogen. Per tant, segons els seus resultats, l'acidesa per pont d'hidrogen del solut (*A*) no té influència en la retenció. Ara bé, el valor del coeficient *a* obtingut en aquesta tesi pel sistema de STC és aproximadament 0.3. Això reflecteix que la fase micel·lar és més bàsica per pont d'hidrogen que l'aquosa, de manera que com més gran sigui el descriptor *A* d'un solut més tendirà a anar cap a la fase micel·lar i a retenir-se en el sistema.

A l'hora de justificar aquesta diferència en el valor del coeficient *a* cal revisar els compostos fets servir com a conjunt de caracterització. Tal com s'observa als estadístics, Poole i col·laboradors van obtenir els coeficients de solvatació emprant 40 soluts, mentre que les altres dues caracteritzacions contenen 55 i 56 soluts. L'interval de valors del descriptor *A* que cobreixen aquests soluts correspon a 0.30-0.69 en el cas de Poole i col·laboradors i 0.23-1.40 a les altres dues caracteritzacions. Per tant, a les caracteritzacions d'aquesta tesi s'ha cobert un major interval del descriptor *A*. Conseqüentment, s'ha obtingut un valor significativament diferent del coeficient *a* que és més exacte i representatiu del sistema cromatogràfic.

Pel que fa a la resta de coeficients, els valors de *e* i *v* obtinguts en aquest tesi són força semblants als de Poole i col·laboradors, mentre que els valors de *s* i *b* són més negatius.

En línies generals, les caracteritzacions obtingudes pel STC concorden amb la de Poole i col·laboradors, atribuint-se les diferències al conjunt de soluts de caracterització. El fet d'haver emprat un conjunt més ampli i variat ha permès caracteritzar millor el sistema, complint així amb el propòsit inicial.

Tal com es veurà als següents apartats, la caracterització de Poole i col·laboradors es farà servir per il·lustrar una metodologia que té com a objectiu avaluar la bondat amb què es poden estimar propietats de biopartició a partir de mesures cromatogràfiques (article II). Després, es farà servir la caracterització de STC a pH 7 obtinguda en aquesta tesi per tal d'estudiar l'emulació de diversos sistemes biològics mitjançant sistemes cromatogràfics (articles III, IV i V).

3.2. RECOPILACIÓ BIBLIOGRÀFICA DE SISTEMES CARACTERITZATS SEGONS EL MODEL DE PARÀMETRES DE SOLVATACIÓ

Amb el propòsit de comparar sistemes biològics i fisicoquímics de partició a partir del marc comú que ofereix el model de paràmetres de solvatació, s'ha fet una recopilació bibliogràfica dels sistemes que es troben caracteritzats a la literatura mitjançant aquest model. També s'han inclòs en aquesta recopilació els sistemes cromatogràfics caracteritzats en aquesta tesi i explicats a l'apartat 3.1. El llistat de sistemes biològics i fisicoquímics recopilats i les seves caracteritzacions es poden consultar a l'apartat d'informació suplementària (Taules SI-1 i SI-2). En total, s'han recopilat 28 sistemes biològics i 92 sistemes fisicoquímics.

Els sistemes de biopartició recopilats són de naturalesa molt diversa. Es poden distingir sistemes com la distribució de substàncies entre la sang i el cervell, processos de permeabilitat a través de barreres com la pell, sistemes basats en processos d'absorció i/o adsorció com l'absorció intestinal humana i la sorció de compostos en sòls, i també molts sistemes que descriuen la toxicitat no específica (sense interaccions covalents) de compostos en diferents tipus d'organismes (bàsicament peixos, amfibis, crustacis, protozous i bacteris).

Pel que fa als sistemes fisicoquímics recopilats, també se'n poden distingir de diferents tipus. Fonamentalment, es poden dividir en tres grups: sistemes de partició aigua-solvent, sistemes de cromatografia de líquids i sistemes de cromatografia electrocinètica. Entre els sistemes de partició aigua-solvent es troba, per exemple, el clàssic i popular sistema de partició octanol-aigua, encara que també n'hi ha d'altres com el sistema de partició pentanol-aigua o hexà-aigua. Respecte als sistemes de cromatografia de líquids, hi ha sistemes convencionals de RP-HPLC i també sistemes de MLC i MELC. Finalment, entre els sistemes de cromatografia electrocinètica recopilats, es poden diferenciar les següents modalitats: sistemes de MEKC, MEEKC, LEKC i

poli-EKC. En general, s'ha parlat de la majoria d'aquests sistemes fisicoquímics de partició a l'apartat 1.2 de la introducció.

A partir d'aquesta recopilació bibliogràfica, s'han comparat els sistemes biològics amb els fisicoquímics. Per obtenir una idea de la seva similitud/dissimilitud, s'ha fet servir el paràmetre de distàncies *d* explicat a l'apartat 1.4 de la introducció. Aquest paràmetre es calcula a partir dels coeficients de solvatació normalitzats, per la qual cosa ha estat necessari normalitzar prèviament els coeficients de cada sistema. Els coeficients normalitzats també es troben detallats a l'apartat d'informació suplementària (Taules SI-1 i SI-2). El càlcul de la distància *d* entre cada sistema biològic i fisicoquímic recopilat ha donat lloc a una matriu de distàncies de dimensions 92x28. Aquesta matriu es pot consultar al CD-ROM adjunt (Taula SI-3).

Els valors de *d* calculats han estat molt útils per a una primera avaluació de la semblança entre els sistemes. Per exemple, es pot observar que el sistema de distribució sang-cervell presenta *d* altes (la majoria entre 0.6-0.8) respecte als sistemes fisicoquímics considerats. Aquest fet indica que el procés que experimenta un compost al distribuir-se entre la sang i el cervell és molt diferent a qualsevol dels processos que tenen lloc en els sistemes fisicoquímics. Per tant, els sistemes fisicoquímics considerats no semblen adequats per emular aquest procés biològic. Passa el mateix amb l'absorció intestinal humana. Contràriament, a la matriu de distàncies es pot observar que altres sistemes de biopartició com la narcosi en capgrossos o la sorció en sòls tenen valors de *d* petits (inferiors o al voltant de 0.25) amb força sistemes fisicoquímics. Això indica que es tracta de sistemes amb característiques similars, essent aleshores probable que les propietats de biopartició i fisicoquímiques d'aquests processos estiguin correlacionades.

Gràcies a la recopilació bibliogràfica i al càlcul de la distància *d*, ha estat possible seleccionar amb criteri diversos sistemes de biopartició per estudiar la seva emulació mitjançant sistemes fisicoquímics. Els sistemes escollits han estat els tres següents: la toxicitat no específica de substàncies en el peix *fathead minnow* [113], la sorció de compostos en sòls [110] i la permeabilitat de soluts a través de la pell [112] (tots tres s'han tractat breument a l'apartat 1.1 de la introducció). Els motius que han portat a escollir-los s'exposen a continuació:

- Aquests sistemes de biopartició tenen distàncies *d* inferiors o de l'ordre de 0.25 respecte a força sistemes fisicoquímics. Això indica semblança entre ells i, per tant, possibilitat de trobar sistemes fisicoquímics que modelin bé aquests sistemes de biopartició. Donat que aquest és l'objectiu, ha estat un dels factors primordials que s'han tingut en compte per escollir-los.

- Un altre aspecte important és que tots tres sistemes de biopartició es troben ben caracteritzats mitjançant l'equació general de solvatació. Si es consulta l'apartat d'informació suplementària (Taula SI-1), es pot veure que els seus coeficients LSER s'han obtingut amb bons ajustos estadístics i incloent un gran número de substàncies. El fet de disposar de sistemes ben caracteritzats és indispensable de cara a estudiar la seva emulació.

- També ha estat necessari fixar-se en el número i tipus de substàncies per a les quals es coneixen i es troben disponibles valors de les propietats de biopartició d'aquests sistemes (LC_{50}, K_{OC} i K_p). Si es vol emular un determinat sistema biològic, cal tenir a l'abast valors de la seva propietat biològica per a un número elevat de soluts. Addicionalment, cal que aquests soluts siguin adequats per ser mesurats en els sistemes fisicoquímics (per exemple, hauran de ser compostos que absorbeixen a l'UV si es fa servir un detector DAD al sistema fisicoquímic). Només així es disposarà de suficients valors, tant de la propietat biològica com fisicoquímica, per a uns mateixos soluts. Això permetrà obtenir correlacions representatives d'ambdues propietats i avaluar la capacitat del sistemes fisicoquímics per modelar sistemes biològics. Els sistemes de biopartició escollits compleixen aquests requisits respecte als soluts.

- Es tracta de sistemes de biopartició d'interès científic i social, donat que les seves corresponents propietats (LC_{50}, K_{OC} i K_p) donen informació de l'efecte que poden tenir les substàncies sobre el medi ambient o el propi cos humà.

- Per últim, s'han volgut escollir processos de biopartició de caire diferent, amb una fase orgànica biològica molt distinta (el peix *fathead minnow*, el sòl i la pell humana). D'aquesta manera es pretén estudiar l'emulació de tres casos concrets, però de naturalesa molt variada.

Un cop escollits els sistemes de biopartició a estudiar, s'ha avaluat la seva similitud respecte a sistemes fisicoquímics a partir dels coeficients normalitzats i la distància *d*. A continuació se'n fan algunes discussions.

A la Taula S-1 de l'article III es pot comparar el sistema de toxicitat en el peix *fathead minnow* amb 12 dels sistemes fisicoquímics recopilats bibliogràficament. El sistema cromatogràfic número 1, que es tracta d'un sistema de LEKC, presenta uns coeficients normalitzats molt semblants al del *fathead minnow* i, conseqüentment, una distància molt petita. D'aquí es dedueix que les interaccions solut-solvent que regeixen ambdós processos són molt similars i, per tant, s'espera que aquest sistema cromatogràfic emuli molt bé la toxicitat en el peix *fathead minnow*. Els valors de *d* pels sistemes 2-6 també són força baixos, d'acord amb que els coeficients normalitzats d'aquests sistemes fisicoquímics també s'assemblen als del *fathead minnow* (per exemple a_u és lleugerament positiu en quasi tots els casos). Per tant, sembla que també podrien ser bons candidats per emular-lo.

En canvi, els sistemes cromatogràfics 7-12 presenten coeficients normalitzats que difereixen cada cop més dels del *fathead minnow* (s_u i a_u són en general bastant més negatius) i, conseqüentment, valors creixents de *d*. Les diferències són molt pronunciades pels sistemes 11 i 12, especialment pel 12. Aquest sistema de MEKC format pel tensioactiu LPFOS té uns valors de a_u i b_u, també de v_u, molt diferents respecte als del *fathead minnow*, fet que queda reflectit en un valor de distància molt gran (0.750, el valor més alt de la taula). D'aquí es dedueix que és un sistema fisicoquímic amb característiques completament distintes a les del *fathead minnow*, el que fa pensar que no tindrà aptitud per modelar-lo.

Una discussió del mateix tipus, però en aquest cas considerant la sorció de compostos en sòls, es pot fer a partir de la Taula 2 de l'article IV. A part d'aquest sistema biològic, a la taula es mostren novament 12 dels sistemes cromatogràfics i de partició compilats bibliogràficament (la majoria d'ells coincideixen amb els de la Taula S-1 de l'article III). El sistema de partició sòl-aigua també es pot comparar amb els sistemes fisicoquímics de la taula a partir dels coeficients normalitzats.

La principal diferència i la més generalitzada es troba en el coeficient s_u, que es relaciona amb la polaritat de les fases. s_u és 0 pel sistema de sorció en sòls, però és en

canvi negatiu per a tots els sistemes fisicoquímics considerats. s_u és especialment negatiu pel sistema de MELC número 11, que també té un valor molt negatiu de a_u en comparació amb el sistema biològic. No obstant, el sistema de MEKC que conté LPFOS (sistema 12) és el que presenta uns coeficients normalitzats més diferents respecte a la sorció en sòls (e_u negatiu, a_u força més negatiu, b_u molt menys negatiu i v_u considerablement més positiu en el sistema 12). Les grans diferències que s'observen en els coeficients normalitzats dels sistemes 11 i 12 queden també reflectides per valors de distància elevats (0.635 i 0.701 respectivament). Es tracta, per tant, de dos sistemes fisicoquímics que no guarden gens de semblança amb el procés de sorció en sòls, de manera que no s'espera que les seves propietats estiguin correlacionades.

Pel que fa a la resta de sistemes (sistemes 1-10), no s'observen diferències molt acusades entre els seus coeficients normalitzats i els del sistema biològic. En conseqüència, els valors de *d* no són molt grans i fan pensar que alguns d'ells podrien ser bons emuladors de la sorció en sòls. És destacable que aquests 10 sistemes cromatogràfics i de partició tenen distàncies similars (compreses entre 0.243 i 0.382), tot indicant que aleshores el seu grau de semblança respecte a la sorció en sòls és també similar.

Tal com s'ha vist, la comparació dels coeficients normalitzats i l'ús de la distància *d* permeten avaluar la semblança entre sistemes, obtenint així una idea aproximada de quins sistemes fisicoquímics podrien o no podrien ser adequats per modelar processos de biopartició. Ara bé, l'èxit de l'emulació es basarà en la precisió que s'obtindrà al correlacionar les dades biològiques i fisicoquímiques. Per tant, convé poder quantificar aquesta precisió de cara a identificar amb un criteri més sòlid quins sistemes fisicoquímics tenen capacitat per emular un determinat procés biològic. Per aquest motiu, en aquesta tesi s'ha desenvolupat una metodologia per conèixer la precisió que es pot obtenir a les correlacions biològico-fisicoquímiques. Aquesta metodologia s'explicarà a continuació (apartat 3.3.1) i s'aplicarà posteriorment a l'emulació dels sistemes de biopartició escollits (apartat 3.3.2).

3.3. ESTIMACIÓ DE PROPIETATS DE BIOPARTICIÓ A PARTIR DE MESURES CROMATOGRÀFIQUES D'HPLC I MEKC

La capacitat d'un sistema fisicoquímic per emular un sistema biològic ve determinada per la variància que s'obté al correlacionar les dades biològiques amb les fisicoquímiques, ja que la precisió amb què la propietat biològica pot ser estimada a partir de la mesura fisicoquímica depèn d'aquesta variància. Per tant, si es vol conèixer la bondat amb la qual sistemes fisicoquímics poden modelar sistemes biològics, cal examinar la variància de les correlacions biològico-fisicoquímiques. Amb tal propòsit, s'ha desenvolupat una metodologia que ha estat posteriorment utilitzada per estudiar casos concrets d'estimació de propietats de biopartició a partir de mesures cromatogràfiques.

3.3.1. Desenvolupament d'una metodologia per avaluar la bondat de les correlacions biològico-cromatogràfiques

Un sistema biològic i un sistema fisicoquímic, concretament cromatogràfic, que hagin estat caracteritzats mitjançant el model de paràmetres de solvatació, es poden expressar d'acord amb els seus coeficients normalitzats de la següent manera:

$$\log SP_{bio} = c_{bio} + l_{bio} (e_{u\,bio} E + s_{u\,bio} S + a_{u\,bio} A + b_{u\,bio} B + v_{u\,bio} V) \quad (14)$$

$$\log SP_{crom} = c_{crom} + l_{crom} (e_{u\,crom} E + s_{u\,crom} S + a_{u\,crom} A + b_{u\,crom} B + v_{u\,crom} V) \quad (15)$$

En realitat, les equacions 14 i 15 corresponen a l'equació general de solvatació (equació 5) expressada en funció de la longitud del vector de coeficients LSER (*l*,

equació 6) i els coeficients LSER normalitzats (e_u , s_u , a_u , b_u i v_u, equacions 7-11). Els subíndex *bio* i *crom* serveixen per distingir el sistema biològic del cromatogràfic.

En el cas ideal de que el sistema biològic i cromatogràfic tinguin el mateix vector normalitzat, la distància *d* entre ells (equació 12) serà 0 i, per tant, es tractarà de sistemes matemàticament idèntics. En aquest cas, el sistema cromatogràfic serà adequat per modelar el procés biològic i es complirà la següent correlació lineal:

$$\log SP_{bio} = c_{bio} - \frac{l_{bio}}{l_{crom}} c_{crom} + \frac{l_{bio}}{l_{crom}} \log SP_{crom} \qquad (16)$$

Tal com s'observa a l'equació 16, l_{bio}/l_{crom} correspondrà al pendent de la correlació. Conseqüentment, informarà de la sensibilitat a l'hora d'estimar la propietat biològica a partir del sistema cromatogràfic.

Quan es compara un sistema biològic amb un sistema cromatogràfic, el valor de l_{bio}/l_{crom} es pot calcular fàcilment a partir de les seves caracteritzacions. Tot i que aquests sistemes tinguin una *d* diferent de 0, el càlcul de l_{bio}/l_{crom} pot donar una idea inicial de la sensibilitat aproximada que es podria obtenir si s'intentés predir la propietat biològica a partir del sistema cromatogràfic.

Amb aquesta intenció, s'ha calculat el valor de l_{bio}/l_{crom} entre la toxicitat en el peix *fathead minnow* i diferents sistemes fisicoquímics de partició, principalment cromatogràfics. Els resultats es mostren a la Taula S-1 de l'article III, on es veu que s'obtenen valors propers a 1 en la majoria de casos. No obstant, l_{bio}/l_{crom} és força més gran que 1 pels següents sistemes: els tres sistemes d'HPLC basats en les columnes IAM, RP18 i MSC18; el sistema de MELC basat en la columna C18 i Brij 35; el sistema de MEKC que conté LPFOS. De manera conseqüent, s'espera que aquests 5 sistemes cromatogràfics tinguin major sensibilitat a l'hora d'estimar la toxicitat en el peix *fathead minnow*.

A la Taula 2 de l'article IV, es mostra el paràmetre l_{bio}/l_{crom} calculat entre el sistema de sorció sòl-aigua i diversos sistemes majoritàriament cromatogràfics. També en aquest cas, els tres sistemes d'HPLC amb columnes RP18, MSC18 i IAM, així com el

sistema de MEKC compost per LPFOS, donen lloc a valors de I_{bio}/I_{crom} superiors a 1 (juntament amb els sistemes PAAU poli-EKC i SDS C18 MELC). En conseqüència, les variacions en les dades cromatogràfiques d'aquests 6 sistemes són les que ocasionaran una major variació en el coeficient de sorció sòl-aigua.

A la pràctica, la bondat amb la qual un sistema biològic és emulat per un sistema cromatogràfic s'avalua correlacionant les dades experimentals d'ambdós sistemes per a un grup representatiu de substàncies, obtenint així una equació del següent tipus:

$$\log SP_{bio} = q + p \log SP_{crom} \quad (17)$$

on q i p són respectivament l'ordenada a l'origen i el pendent de la correlació. Com més proper a 0 sigui el valor de la distància d entre els dos sistemes, més s'assemblaran els valors de q i p a l'ordenada a l'origen i el pendent predits per l'equació 16, respectivament.

La variància total d'aquesta correlació (equació 17) determina la precisió amb què la propietat biològica d'un compost pot ser estimada a partir de seva mesura cromatogràfica. Depenent d'aquesta precisió, el sistema cromatogràfic es considerarà més o menys adequat per emular el sistema biològic. Per tant, examinar la variància total d'una correlació biològico-cromatogràfica resulta clau per avaluar la bondat amb què un sistema biològic és modelat per un sistema cromatogràfic.

Tal com s'ha exposat a l'article II, la variància total que s'obté al correlacionar dades experimentals biològiques i cromatogràfiques (σ_{corr}^2) prové dels tres factors següents: la precisió de les dades biològiques, la precisió de les dades cromatogràfiques i l'error degut a la dissimilitud entre els dos sistemes correlacionats. σ_{corr}^2 es pot expressar llavors com a suma de les contribucions d'aquests tres factors:

$$\sigma_{corr}^2 = \sigma_{bio}^2 + \sigma_{crom}^2 + \sigma_d^2 \quad (18)$$

σ_{bio}^2 i σ_{crom}^2 són, respectivament, les contribucions de variància degudes a la precisió de les dades biològiques i cromatogràfiques correlacionades. σ_d^2 correspon a la

contribució de variància deguda al grau de dissimilitud entre els sistemes biològic i cromatogràfic comparats.

Després de descompondre la variància total de les correlacions biològico-cromatogràfiques en les contribucions dels tres factors que la determinen (equació 18), s'ha examinat cadascuna d'aquestes contribucions i s'ha proposat com estimar-les.

En primer lloc, s'ha examinat la contribució de variància deguda a la precisió de les dades biològiques (σ_{bio}^2). Donat que qualsevol correlació biològico-cromatogràfica es veurà afectada per la incertesa dels valors biològics, es tracta d'un factor que cal avaluar. De fet, no s'esperen obtenir correlacions amb millor precisió que la de les dades de partida (això només pot ocórrer si les dades correlacionades pertanyen a conjunts de compostos homòlegs, els quals cal evitar a fi d'obtenir resultats representatius).

Degut a la complexitat dels processos biològics, les dades biològiques solen tenir força error intrínsec perquè són difícils de determinar experimentalment. A més, si es volen compilar suficients dades d'una determinada propietat biològica, moltes vegades cal recórrer a dades obtingudes per diversos autors i mesurades en diferents laboratoris (fent servir distints protocols i a partir de mostres que no tenen la mateixa uniformitat ni procedència). Aquest fet també contribueix a que les dades biològiques que es correlacionen tinguin un error considerable. Un altre aspecte a comentar és que la precisió de les dades biològiques pot variar molt depenent del sistema biològic considerat. Donat que existeix una gran diversitat de processos biològics, les seves corresponents propietats es poden mesurar amb més precisió en alguns sistemes que en altres. Per tant, a l'hora d'interpretar el resultat d'una correlació biològico-cromatogràfica, així com de comparar-la amb altres correlacions, és indispensable fixar-se en el sistema biològic correlacionat i en la precisió de les seves dades.

S'ha proposat estimar la contribució de variància que prové de les dades biològiques (σ_{bio}^2) a partir de la caracterització del sistema biològic mitjançant el model de paràmetres de solvatació. Assumint que l'error del model és negligible, es pot considerar que la desviació estàndard de la seva caracterització (SD_{bio}) és una

estimació de la precisió dels valors biològics. Conseqüentment, es proposa estimar σ_{bio}^2 de la següent manera:

$$\sigma_{bio}^2 \sim SD_{bio}^2 \quad (19)$$

De fet, si es consulten els sistemes biològics de l'apartat d'informació suplementària (Taula SI-1), s'observa que els valors SD concorden perfectament amb el que s'ha comentat sobre la precisió de les dades biològiques. Es tracta de valors de SD en general força elevats (al voltant de 0.3) i que varien bastant entre alguns sistemes (per exemple, SD és 0.461 pel sistema de permeabilitat en pell mentre que és 0.205 pel sistema de toxicitat en el protozou T*etrahymena pyriformis*).

En segon lloc, s'ha examinat la contribució de variància deguda a la precisió de les dades cromatogràfiques (σ_{crom}^2). Òbviament, en una correlació biològico-cromatogràfica, l'error que contenen els valors cromatogràfics també es traslladarà a la correlació i serà per tant un altre factor a avaluar.

Les dades cromatogràfiques (i fisicoquímiques en general) són més fàcils de determinar experimentalment que les biològiques, de manera que se solen mesurar amb més exactitud i precisió. Això significa que s'espera que σ_{crom}^2 contribueixi menys que σ_{bio}^2 a la variància total de la correlació (σ_{corr}^2). Una altra característica de la precisió dels valors cromatogràfics és que no sol variar gaire entre sistemes.

Per estimar σ_{crom}^2 també s'ha proposat fer servir la caracterització del corresponent sistema mitjançant el model de paràmetres de solvatació. Assumint novament que el model té un error negligible, la desviació estàndard de la caracterització del sistema cromatogràfic (SD_{crom}) es pot atribuir a la precisió de les seves dades. No obstant, a part de SD_{crom}, també és necessari considerar el pendent de la correlació biològico-cromatogràfica (p) per estimar σ_{crom}^2. El motiu és que al correlacionar les dades d'un sistema biològic i un sistema cromatogràfic, com més gran sigui p, més afectarà l'error de les dades cromatogràfiques a la variància total. Per tant, es proposa estimar σ_{crom}^2 de la següent manera:

$$\sigma_{crom}^2 \sim (p\, SD_{crom})^2 \quad (20)$$

A l'apartat d'informació suplementària, Taula SI-2, apareixen detallats els valors de SD obtinguts a la caracterització d'un ampli ventall de sistemes cromatogràfics. En concordança amb el que s'ha comentat sobre la precisió de les dades cromatogràfiques, es pot veure que els valors de SD no són gaire alts (sobre 0.1) i que no hi ha moltes diferències entre ells.

En tercer i últim lloc, s'ha examinat la contribució de variància que prové de la dissimilitud entre els dos sistemes correlacionats (σ_d^2). El grau de semblança entre un sistema biològic i un sistema cromatogràfic condicionarà molt la bondat de la seva correlació, per la qual cosa es tracta d'un factor que és fonamental avaluar.

Donat que la distància *d* també es relaciona amb la dissimilitud entre parelles de sistemes, en principi s'espera que els valors de *d* i σ_d^2 siguin coherents entre si. Per exemple, dos sistemes idèntics tindrien tant *d* com σ_d^2 iguals a 0, ja que el fet que no hi haguessin diferències entre els sistemes faria que la seva correlació no es veiés afectada per aquest factor. En canvi, per a parelles de sistemes cada cop més diferents entre si s'hauria d'observar tant un augment de *d* com de σ_d^2. Dos sistemes amb propietats molt diferents (*d* elevada) no es correlacionaran bé degut a la seva dissimilitud, de manera que la contribució d'error degut a aquest factor hauria de ser també alta (valor elevat de σ_d^2).

Considerant un sistema biològic i un cromatogràfic, σ_d^2 es pot estimar a partir de la correlació entre valors calculats per a ambdós sistemes. Els valors calculats de cada sistema es poden obtenir a través de la seva caracterització mitjançant el model de paràmetres de solvatació, multiplicant els coeficients LSER de la caracterització pels descriptors moleculars dels compostos considerats. Assumint que l'error del model és negligible, es pot considerar que els valors calculats d'aquesta manera no tenen error associat. Per tant, la variància de la correlació entre valors calculats no es veurà afectada per la precisió de les dades ($SD_{bio\ cal}^2$ i $(p\ SD_{crom\ cal})^2$ seran 0) i es pot atribuir directament a la dissimilitud entre els sistemes (SD_d^2):

$$SD_{corr\ cal}^2 = \underbrace{SD_{bio\ cal}^2}_{0} + \underbrace{(p\ SD_{crom\ cal})^2}_{0} + SD_d^2 \qquad (21)$$

Així, s'ha proposat estimar la contribució de variància deguda a la dissimilitud de dos sistemes (σ_d^2) a partir de la variància que s'obté al correlacionar els seus valors calculats:

$$\sigma_d^2 \sim SD_d^2 \qquad (22)$$

En resum, la bondat d'una correlació biològico-cromatogràfica es pot avaluar mitjançant la metodologia proposada. Aquesta metodologia consisteix en examinar els tres factors que determinen la variància total (precisió de les dades biològiques, precisió de les dades cromatogràfiques i dissimilitud entre els dos sistemes correlacionats), tot estimant la contribució de cada factor a partir del model de paràmetres de solvatació. La variància total (equació 18) es pot predir llavors com a suma de les contribucions estimades (equació 23):

$$SD_{corr\ cal}^2 = SD_{bio}^2 + (p\ SD_{crom})^2 + SD_d^2 \qquad (23)$$

D'aquesta manera, $SD_{corr\ cal}^2$ permet quantificar la precisió que s'espera obtenir a la correlació biològico-cromatogràfica.

Aquest procediment queda també il·lustrat a la Figura 16:

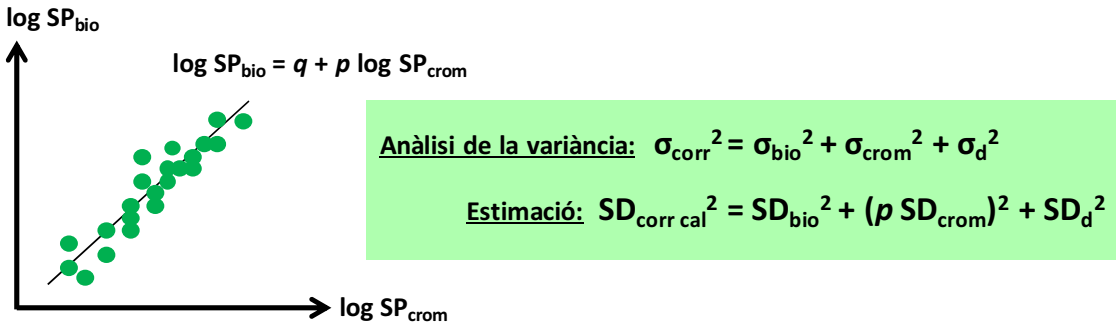


Figura 16. Anàlisi de la variància total d'una correlació biològico-cromatogràfica en funció dels factors que la determinen. Estimació de la contribució d'aquests factors a la variància total.

A continuació, la metodologia proposada s'ha provat amb diferents exemples de sistemes biològics i cromatogràfics, concretament de MEKC (article II).

Per tal d'avaluar l'estimació de la contribució de les diferents variàncies, inicialment la metodologia s'ha provat considerant només dos dels factors que contribueixen a la variància total: la precisió de les dades cromatogràfiques (σ_{crom}^2) i la dissimilitud entre els sistemes (σ_d^2). Per tal de no considerar el tercer factor (la precisió de les dades biològiques, σ_{bio}^2), s'han correlacionat dades biològiques calculades (amb la qual cosa s'assumeix que σ_{bio}^2 és 0) respecte a dades cromatogràfiques experimentals. Conseqüentment, la variància total d'aquestes correlacions (σ_{corr}^2) depèn només de σ_{crom}^2 i σ_d^2 (les dades biològiques no aporten error a les correlacions perquè són calculades).

Taula 9. Avaluació de correlacions entre valors biològics calculats i cromatogràfics experimentals pels sistemes detallats.

Sistema biològic	**Sistema de MEKC**					
	SDS [118,119]			**LPFOS** [118,119]		
	d	**$SD_{corr\ cal}^2$**	**$SD_{corr\ exp}^2$**	***d***	**$SD_{corr\ cal}^2$**	**$SD_{corr\ exp}^2$**
Sorció sòl-aigua [136]	0.139	0.035	0.034	0.622	0.237	0.237
Narcosi en capgrossos [3]	0.173	0.067	0.066	0.651	0.411	0.411
Toxicitat en el peix *guppy* [113]	0.282	0.145	0.144	0.722	0.564	0.563
Toxicitat en el protozou *Tetrahymena pyriformis* [122]	0.310	0.124	0.122	0.712	0.448	0.447
Partició a la pell [112]	0.338	0.103	0.103	0.736	0.237	0.237
Toxicitat en el peix *fathead minnow* [113]	0.347	0.196	0.194	0.750	0.648	0.647
Toxicitat en el peix *bluegill* [113]	0.469	0.442	0.440	0.884	0.970	0.970
Toxicitat en el peix *goldfish* [113]	0.700	0.619	0.615	0.871	1.203	1.200
Absorció intestinal [1]	0.879	0.029	0.029	0.874	0.028	0.028

S'han comparat 9 sistemes biològics de naturalesa diversa amb 2 sistemes de MEKC, un format per SDS i l'altre per LPFOS. Per a cada parella de sistemes biològic-MEKC, s'han estimat les contribucions de variància $(p\ SD_{crom})^2$ i SD_d^2. La seva suma ha permès estimar la variància total, $SD_{corr\ cal}^2$ (equació 23 amb $SD_{bio}^2 = 0$ en aquest cas). Els valors de $SD_{corr\ cal}^2$ s'han comparat amb els de la variància total experimental, $SD_{corr\ exp}^2$, tal com es mostra a la Taula 9.

Es pot observar que s'han obtingut valors pràcticament idèntics de $SD_{corr\ cal}^2$ i $SD_{corr\ exp}^2$ per a tots els casos, independentment dels sistemes biològics i de MEKC correlacionats i de la seva semblança (veure *d*). Per tant, es demostra que el mètode proposat aconsegueix predir bé la variància total quan només es consideren com a fonts d'error la incertesa de les dades cromatogràfiques i la dissimilitud entre els sistemes.

Seguidament, s'ha prosseguit considerant ja els tres factors que determinen la variància total de les correlacions biològico-cromatogràfiques: σ_{bio}^2, σ_{crom}^2 i σ_d^2. Amb aquest propòsit s'han correlacionat dades experimentals tant biològiques com cromatogràfiques: d'una banda, valors de toxicitat en el peix *fathead minnow* respecte a factors de retenció en diversos sistemes de MEKC (Taula 10); d'altra banda, diferents propietats de biopartició respecte a factors de retenció en el sistema de MEKC de STC (Taula 11).

Per a cada parella de sistemes, s'ha estimat la contribució de variància de cada factor (SD_{bio}^2, $(p\ SD_{crom})^2$ i SD_d^2) i la variància total que s'espera obtenir a la correlació ($SD_{corr\ cal}^2$, equació 23). Novament, els valors de $SD_{corr\ cal}^2$ s'han comparat amb els de la variància obtinguda experimentalment, $SD_{corr\ exp}^2$.

A totes dues taules (Taules 10 i 11), es pot veure que $SD_{corr\ cal}^2$ dóna una força bona aproximació de $SD_{corr\ exp}^2$. El fet que $SD_{corr\ cal}^2$ i $SD_{corr\ exp}^2$ s'acostessin més al considerar només dos factors (Taula 9), té a veure amb les substàncies: quan s'han considerat dos factors s'han emprat valors biològics calculats, podent fer servir aleshores els mateixos soluts per obtenir $SD_{corr\ cal}^2$ i $SD_{corr\ exp}^2$; en canvi, al considerar els tres factors, s'han emprat dades experimentals tant biològiques com cromatogràfiques, de manera que cada contribució ha hagut de ser estimada amb un conjunt diferent de soluts i només alguns d'ells (els que coincidien) s'han fet servir a les correlacions experimentals. En

aquest últim cas (que serà el cas habitual) és normal que la utilització de diferents grups de substàncies faci que $SD_{corr\ cal}^2$ no predigui exactament $SD_{corr\ exp}^2$, sinó de manera aproximada. No obstant, els valors obtinguts de $SD_{corr\ cal}^2$ s'aproximen prou bé als de $SD_{corr\ exp}^2$ i condueixen a les mateixes conclusions, tal com es comentarà a continuació. Per tant, es demostra que la metodologia de predicció de variàncies funciona bé quan es consideren els tres factors que la determinen, ja que permet predir correctament la variància que s'obtindrà a correlacions biològico-cromatogràfiques.

Taula 10. Avaluació de correlacions entre valors experimentals de -log LC_{50} pel peix *fathead minnow* i log k pels sistemes de MEKC detallats.

Sistema de MEKC	**Toxicitat en el peix *fathead minnow*** [113]**,** $\mathbf{SD_{bio}^2 = 0.076}$				
	d	$\mathbf{(p\ SD_{crom})^2}$	$\mathbf{SD_d^2}$	$\mathbf{SD_{corr\ cal}^2}$	$\mathbf{SD_{corr\ exp}^2}$
SDS/Brij 35 [70]	0.131	0.009	0.020	0.105	0.100
TTAB [118,119]	0.246	0.008	0.066	0.150	0.078
SDS [118,119]	0.347	0.017	0.178	0.272	0.208
LPFOS [118,119]	0.750	0.019	0.671	0.767	0.525

Taula 11. Avaluació de correlacions entre valors experimentals de diverses propietats de biopartició i log k pel sistema de MEKC de STC.

Sistema biològic	**Sistema de MEKC: STC** [78]					
	d	$\mathbf{SD_{bio}^2}$	$\mathbf{(p\ SD_{crom})^2}$	$\mathbf{SD_d^2}$	$\mathbf{SD_{corr\ cal}^2}$	$\mathbf{SD_{corr\ exp}^2}$
Sorció sòl-aigua [136]	0.077	0.029	0.007	0.005	0.041	0.022
Toxicitat en el peix *guppy* [113]	0.105	0.078	0.004	0.030	0.113	0.086
Toxicitat en el protozou *Tetrahymena pyriformis* [122]	0.156	0.042	0.008	0.039	0.090	0.055
Toxicitat en el peix *fathead minnow* [113]	0.183	0.076	0.010	0.056	0.143	0.087

Aquest mètode de contribució de variàncies resulta molt útil per avaluar la bondat amb què un sistema biològic és emulat per un sistema cromatogràfic i, per tant, de cara a escollir quins sistemes cromatogràfics seran més adequats per a l'emulació. Donat que l'error total d'una correlació biològico-cromatogràfica s'espera que sigui superior al de les dades de partida, s'ha proposat comparar la contribució de variància deguda a la dissimilitud entre els sistemes (SD_d^2) amb la contribució de variància que prové de les dades originals (en general SD_{bio}^2, ja que les dades cromatogràfiques solen tenir un error força inferior). Si $SD_d^2 \sim SD_{bio}^2$ significa que no s'introduirà gaire més error a la correlació del que ja tenen les dades originals, de manera que es podrà considerar que el sistema cromatogràfic emula bé el biològic. Contràriament, si $SD_d^2 >> SD_{bio}^2$ significa que la dissimilitud entre els sistemes aportarà força més error a la correlació que les seves pròpies dades, de manera que la variància total serà massa alta i el sistema biològic no serà ben modelat pel cromatogràfic.

A la Taula 10, s'observa que $SD_d^2 < SD_{bio}^2$ quan es compara la toxicitat en el peix *fathead minnow* amb els dos sistemes de MEKC formats per SDS/Brij 35 i TTAB. Aquesta mateixa situació es dóna per a totes les parelles de sistemes de la Taula 11. En tots els casos, aquest fet concorda amb valors baixos de *d* ($d < 0.25$). D'aquí es dedueix que els sistemes biològics i de MEKC comparats són força semblants i que s'espera obtenir correlacions precises (ja que s'estimen valors baixos de $SD_{corr\ cal}^2$). Els valors obtinguts de $SD_{corr\ exp}^2$ confirmen que efectivament aquests sistemes biològics són ben emulats pels corresponents sistemes de MEKC.

Tot el contrari és el que s'observa a la Taula 10 quan es compara la toxicitat en el peix *fathead minnow* amb els sistemes de MEKC de SDS i de LPFOS: $d > 0.25$ i $SD_d^2 >> SD_{bio}^2$, especialment en el cas del LPFOS. Conseqüentment, degut a la dissimilitud entre els sistemes, es pronostica una variància total alta ($SD_{corr\ cal}^2$) que queda confirmada experimentalment ($SD_{corr\ exp}^2$). Per tant, aquests sistemes de MEKC no són adequats per emular la toxicitat en el peix *fathead minnow*.

Aquesta metodologia, a part de ser útil per conèixer quins sistemes biològics i cromatogràfics es correlacionaran bé, també permet interpretar les correlacions. A la Taula 11, segons el paràmetre *d*, s'esperaria obtenir una millor correlació entre la parella de sistemes toxicitat en el peix *guppy*-STC que entre la parella de sistemes

toxicitat en el protozou *Tetrahymena pyriformis*-STC (la *d* és inferior en el primer cas). En canvi, a les correlacions experimentals s'obté el contrari ($SD_{corr\ exp}^2$ és superior en el primer cas). Gràcies a l'anàlisi de les contribucions de variància, podem deduir que aquest fet es pot atribuir a una major incertesa de les dades originals per a la toxicitat en el peix *guppy* (SD_{bio}^2 és 0.078, mentre que per a la toxicitat en el protozou T*etrahymena pyriformis* és 0.042).

A mode de resum, a la Figura 17 s'esquematitza el procediment proposat per avaluar la bondat de les correlacions biològico-cromatogràfiques. En el següent apartat, s'estudiarà l'emulació de diversos sistemes biològics mitjançant sistemes cromatogràfics a partir d'aquest procediment.

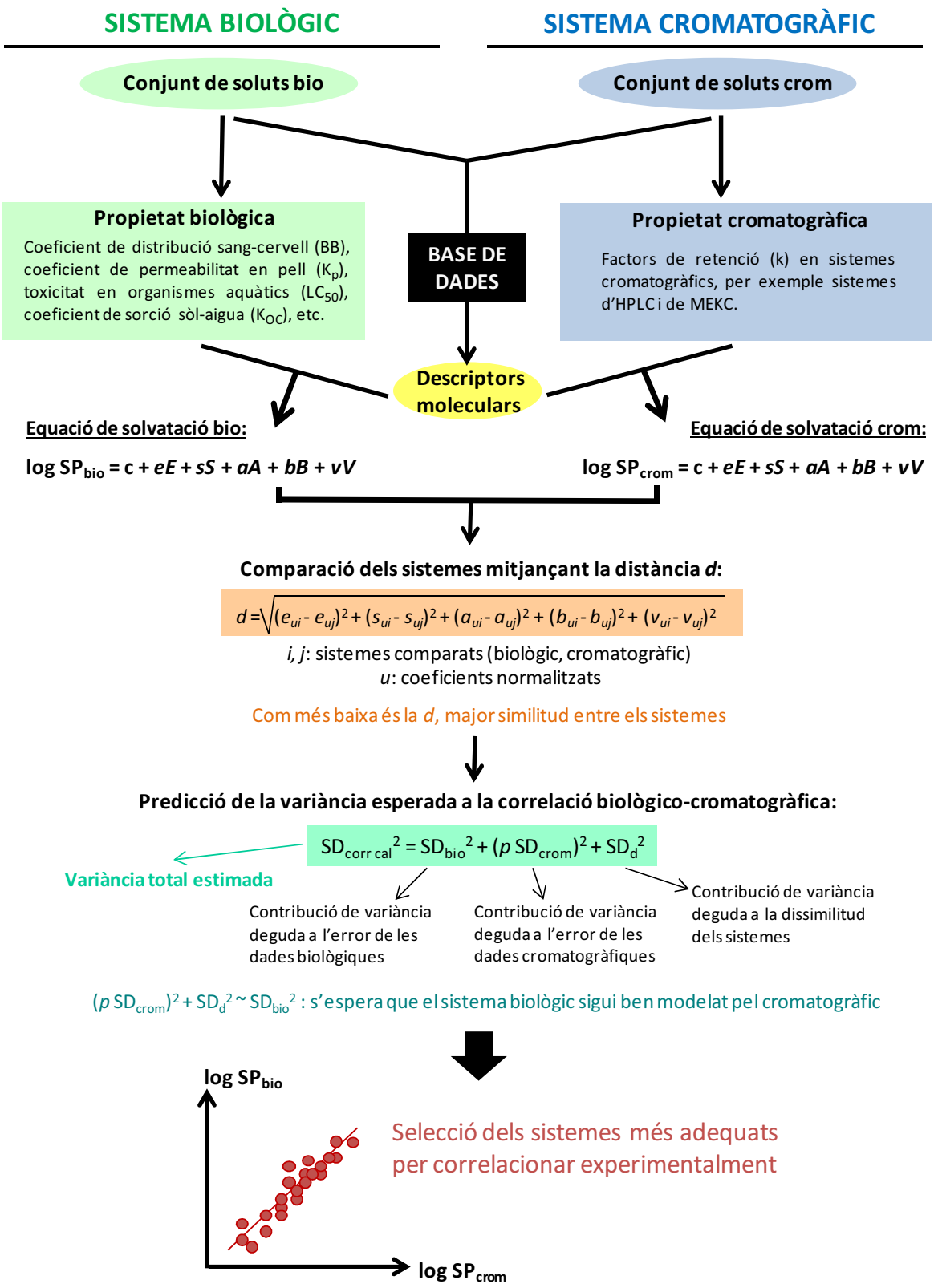


Figura 17. Diagrama del procediment proposat per estudiar l'emulació de sistemes biològics mitjançant sistemes cromatogràfics.

3.3.2. Modelació de sistemes de biopartició concrets: toxicitat en el peix *fathead minnow*, sorció sòl-aigua i permeabilitat a través de la pell

Tal com s'ha explicat a l'apartat 3.2, s'han escollit tres sistemes de biopartició per estudiar la seva emulació mitjançant sistemes cromatogràfics: la toxicitat aquàtica no específica de compostos orgànics neutres en el peix *fathead minnow*, la sorció en sòls de substàncies neutres en fase aquosa i la permeabilitat de soluts neutres en solució aquosa a través de la pell humana.

Amb el propòsit de trobar sistemes cromatogràfics anàlegs que permetin estimar les seves propietats, el primer pas ha estat seleccionar sistemes cromatogràfics amb possibilitats d'acomplir aquest objectiu. Això s'ha dut a terme a partir de la distància *d* i el mètode de predicció de variàncies exposat a l'apartat anterior. A continuació, s'ha comprovat la bondat amb què els sistemes cromatogràfics escollits permeten estimar les propietats de biopartició a partir de correlacions experimentals. Finalment, els millors models de correlació s'han aplicat de manera pràctica a la predicció de valors de K_{OC} i K_p de substàncies d'interès. Tot aquest estudi s'explicarà en detall a continuació i correspon als articles III, IV i V de la present tesi doctoral.

3.3.2.1. Selecció de possibles sistemes cromatogràfics per modelar-los

A la Taula 12 es presenten alguns exemples dels sistemes cromatogràfics i de partició que han estat comparats amb els processos de biopartició que es volen emular. Tots els sistemes de la Taula 12 provenen de la compilació bibliogràfica presentada a l'apartat 3.2, la qual es pot consultar a l'apartat d'informació suplementària (Taula SI-2). Es pot observar que els sistemes de la taula són força diversos. A part de sistemes de partició aigua-solvent, s'han contemplat diferents tipus de sistemes de cromatografia electrocinètica (MEKC, MEEKC, LEKC, poli-EKC) i de cromatografia de líquids (HPLC, MELC).

Taula 12. Sistemes fisicoquímics comparats amb els sistemes de biopartició que es volen modelar.

Nom del sistema		Condicions experimentals	Ref.
Sistema	**Tipus**		
DPPG:DPPC	LEKC	12 mM dipalmitoilfosfatidilglicerol:dipalmitoilfosfatidilcolina (30:70), tampó 25 mM HEPES pH 7.5	[137]
STC	MEKC	50 mM taurocolat de sodi, tampó 20 mM fosfat pH 7	Article III
Pentanol-aigua	Partició solvent-aigua	Diverses	[138]
Microemulsió de SDS	MEEKC	1.44% (w/w) dodecilsulfat de sodi, 6.49% (w/w) 1-butanol, 0.82% (w/w) heptà, 91.25% (w/w) tampó 500 mM fosfat+100 mM borat pH 7	[139]
IAM	HPLC	Columna IAM.PC.DD2, tampó 10 mM fosfat pH 7:acetonitril (60:40)	[54]
Octanol-aigua	Partició solvent-aigua	Diverses	[117]
RP18	HPLC	Columna Xterra RP18, tampó 10 mM fosfat pH 7:acetonitril (60:40)	[54]
MSC18	HPLC	Columna Xterra MSC18, tampó 10 mM fosfat pH 7:acetonitril (60:40)	[54]
poli-(SDeS)	Poli-EKC	10 $g \cdot l^{-1}$ poli(9-decenilsulfat de sodi), tampó 100 mM fosfat pH 7	[140]
SDS	MEKC	40 mM dodecilsulfat de sodi, tampó 20 mM fosfat pH 7	[118,119]
Brij 35 C18	MELC	Columna AT. Chrom C18, 6.0% (w/w) èter de dodecilpolioxietilè(23), 6.6% (w/w) butanol, 0.8% (w/w) heptà, 86.6% (w/w) tampó 50 mM fosfat pH 7	[85]
LPFOS	MEKC	40 mM perfluorooctanosulfonat de liti, tampó 20 mM fosfat pH 7	[118,119]
DPPG:DPPC:Col	LEKC	10 mM dipalmitoilfosfatidilglicerol:dipalmitoilfosfatidilcolina:colesterol (24:46:30), tampó 25 mM HEPES pH 7.5	[137]
PAAU	Poli-EKC	58 mM poli(11-acrilamidaundecanoat de sodi), tampó 12.5 mM fosfat-tetraborat pH 9.2	[141]
SDS C18	MELC	Columna AT. Chrom C18, 3.3% (w/w) dodecilsulfat de sodi, 6.6% (w/w) butanol, 1.6% (w/w) heptà, 88.5% (w/w) tampó 50 mM fosfat pH 7	[85]

La distància *d* és una de les eines que permet comparar aquests sistemes fisicoquímics amb els de biopartició. De fet, a l'apartat 3.2, aquests sistemes ja es van comparar amb la toxicitat en el peix *fathead minnow* i amb la sorció sòl-aigua a partir dels coeficients normalitzats i del paràmetre *d* (discussions corresponents a la Taula S-1 de l'article III i a la Taula 2 de l'article IV, respectivament). No obstant, a partir d'aquestes discussions va quedar palès que cal estimar la precisió que s'espera obtenir a les correlacions biològico-cromatogràfiques si es volen identificar quins són els sistemes cromatogràfics més adients per emular processos de biopartició. Per això, ara s'ha procedit a aplicar la metodologia de predicció de variàncies exposada a l'apartat 3.3.1.

Els sistemes de la Taula 12 s'han comparat amb la toxicitat en el peix *fathead minnow* (Taula 13) i la sorció sòl-aigua (Taula 14) a partir de la variància que s'ha predit per a les seves correlacions (les Taules 13 i 14 són una simplificació de la Taula 2 de l'article III i de la Taula 3 de l'article IV, respectivament). Per a cada parella de sistemes biològic i cromatogràfic, s'ha estimat cada contribució de variància (SD_{bio}^2, $(p_{cal}\ SD_{crom})^2$ i SD_d^2) i la variància total ($SD_{corr\ cal}^2$) que s'espera obtenir a la correlació. En ambdues taules, els sistemes cromatogràfics apareixen classificats segons si SD_d^2 és inferior a SD_{bio}^2, de l'ordre, o molt superior, ja que tal com s'ha comentat prèviament és un aspecte que condiciona molt la bondat de les correlacions. Dins de cada classificació apareixen ordenats per valors creixents de $SD_{corr\ cal}^2$.

Els sistemes cromatogràfics amb $SD_d^2 < SD_{bio}^2$ respecte al sistema biològic, es poden considerar que seran bons emuladors. En aquests casos, no s'introduirà gaire més error a les correlacions biològico-cromatogràfiques respecte a l'error que contenen les dades originals. Per això, les variàncies totals que es prediuen, $SD_{corr\ cal}^2$, són baixes.

No obstant, dels 4 sistemes cromatogràfics i de partició que es troben en aquesta situació respecte a la toxicitat en el peix *fathead minnow* (veure Taula 13), només s'han escollit els sistemes STC MEKC i IAM HPLC per comprovar experimentalment la seva capacitat de modelació. El motiu per no escollir el sistema DPPG:DPPC LEKC és que la preparació de dissolucions de liposomes és un procés que requereix temps i és delicat, a part que els liposomes tendeixen a presentar problemes d'estabilitat. Tampoc s'ha escollit el sistema de partició pentanol-aigua perquè en general els

sistemes de partició aigua-solvent requereixen un treball experimental llarg i tediós de preparació de les particions i equilibrament de fases. En canvi, els sistemes escollits són d'HPLC i MEKC, dues tècniques que ofereixen rapidesa, simplicitat i menor cost econòmic. Pel que fa a la sorció sòl-aigua, també s'ha escollit el sistema STC MEKC per comprovar la bondat de la seva correlació experimental (tal com s'observa a la Taula 14, $SD_d^2 < SD_{bio}^2$, el que dóna lloc a un valor de $SD_{corr\ cal}^2$ baix).

Taula 13. Comparació de sistemes cromatogràfics i de partició amb la toxicitat en el peix *fathead minnow* a partir de la predicció de variàncies i la distància *d*.

Toxicitat en el peix *fathead minnow* ($SD_{bio}^2 = 0.076$)			
Comparació SD_d^2 i SD_{bio}^2	**Sistema**	***d***	**$SD_{corr\ cal}^2$**
$SD_d^2 < SD_{bio}^2$	DPPG:DPPC LEKC	0.095	0.093
	STC MEKC	0.147	0.113
	IAM HPLC	0.186	0.132
	Pentanol-aigua	0.153	0.174
$SD_d^2 \sim SD_{bio}^2$	Octanol-aigua	0.198	0.170
	SDS MEEKC	0.176	0.179
	SDS MEKC	0.347	0.210
	poli-(SDeS) poli-EKC	0.335	0.236
	RP18 HPLC	0.259	0.237
	MSC18 HPLC	0.295	0.300
$SD_d^2 >> SD_{bio}^2$	Brij 35 C18 MELC	0.475	0.692
	LPFOS MEKC	0.750	0.801

Els sistemes de les Taules 13 i 14 pels quals s'ha estimat que SD_d^2 és lleugerament superior a SD_{bio}^2, però no molt superior sinó de l'ordre ($SD_d^2 \sim SD_{bio}^2$), s'espera que encara puguin emular de manera acceptable els processos de biopartició. Òbviament, però, s'espera obtenir correlacions biològico-cromatogràfiques menys precises (s'han estimat valors superiors de $SD_{corr\ cal}^2$) que en els casos en què $SD_d^2 < SD_{bio}^2$. No obstant,

la variància total estimada encara no és prou gran com per desestimar la possibilitat d'obtenir bones correlacions.

Taula 14. Comparació de sistemes cromatogràfics i de partició amb la sorció sòl-aigua a partir de la predicció de variàncies i la distància *d*.

Sorció sòl-aigua (SD_{bio}^2 = 0.062)			
Comparació SD_d^2 i SD_{bio}^2	**Sistema**	***d***	**$SD_{corr\ cal}^2$**
$SD_d^2 < SD_{bio}^2$	STC MEKC	0.249	0.128
$SD_d^2 \sim SD_{bio}^2$	DPPG:DPPC:Col LEKC	0.239	0.139
	SDS MEKC	0.327	0.143
	IAM HPLC	0.276	0.156
	PAAU poli-EKC	0.382	0.167
	Octanol-aigua	0.272	0.191
	Pentanol-aigua	0.264	0.203
	SDS MEEKC	0.260	0.208
	RP18 HPLC	0.243	0.220
	MSC18 HPLC	0.265	0.265
$SD_d^2 >> SD_{bio}^2$	LPFOS MEKC	0.701	0.494
	SDS C18 MELC	0.635	0.594

Entre els sistemes amb $SD_d^2 \sim SD_{bio}^2$, s'han seleccionat el d'octanol-aigua, SDS MEKC, RP18 HPLC i MSC18 HPLC per fer les comprovacions experimentals tant de l'emulació de la toxicitat en el peix *fathead minnow* com de la sorció sòl-aigua. S'han escollit concretament aquests sistemes per les següents raons: el sistema de partició octanol-aigua s'ha fet servir clàssicament com a model de referència de processos biològics [20,24] i, a més, es poden obtenir valors experimentals de log $P_{o/w}$ a través d'una base de dades fiable i a l'abast com és el Bio-Loom [44] (sense necessitat de fer experiments de partició); el SDS és el tensioactiu més popular i emprat en aplicacions analítiques i, conseqüentment, resulta interessant conèixer la capacitat del sistema SDS MEKC per emular processos biològics; els sistemes RP18 HPLC i MSC18 HPLC es

basen en columnes C18, l'ús de les quals està molt estès i és molt comú en laboratoris, de manera que existeix molt interès en conèixer la seva aptitud per predir propietats de biopartició. Entre els sistemes amb SD_d^2 ~ SD_{bio}^2 respecte a la sorció sòl-aigua, també s'ha escollit el sistema IAM HPLC donat que les columnes IAM van ser especialment dissenyades amb la finalitat de mimetitzar entorns biològics.

A partir de les prediccions de variància també es poden identificar quins sistemes cromatogràfics no seran adients per modelar processos biològics. Es tracta d'aquells pels que s'estima un valor de SD_d^2 molt superior a SD_{bio}^2, de manera que són sistemes amb característiques molt diferents als de biopartició. Per això, a les Taules 13 i 14 es pot veure que en aquests casos es preveu obtenir correlacions biològico-cromatogràfiques molt afectades per la variància (valors alts de $SD_{corr\ cal}^2$).

Pel que fa al paràmetre *d* també detallat a les Taules 13 i 14, es pot comprovar que els seus valors són en general coherents amb la classificació dels sistemes segons SD_d^2 i segons el valor predit de $SD_{corr\ cal}^2$: les parelles de sistemes biològic i cromatogràfic amb distància *d* baixa es correlacionaran sense gaire error associat a la dissimilitud entre els sistemes. Per tant, s'espera una correlació amb poca variància (valor baix de $SD_{corr\ cal}^2$). S'observa el contrari per a les parelles de sistemes amb *d* alta. No obstant, el fet que *d* i $SD_{corr\ cal}^2$ no augmentin estrictament en paral·lel es deu a que la distància només dóna una idea aproximada de la dissimilitud entre parelles de sistemes. Quan hi ha parelles amb *d* properes, la distància no és capaç de diferenciar quina serà la influència de la dissimilitud dels sistemes sobre la variància de la correlació.

En definitiva, s'han escollit els mateixos 6 sistemes (5 cromatogràfics i el de partició octanol-aigua) per comprovar amb dades experimentals la seva capacitat per modelar la toxicitat en el peix *fathead minnow* i la sorció sòl-aigua. També s'ha decidit que es comprovarà experimentalment la seva bondat per emular el sistema de permeabilitat a través de la pell. Aquests sistemes es detallen a la Taula 15, juntament amb els valors calculats de *d* i $SD_{corr\ cal}^2$ respecte a cadascun dels tres processos de biopartició.

Taula 15. Valors de *d* i $SD_{corr\ cal}^2$ entre els sistemes cromatogràfics seleccionats i els tres processos de biopartició estudiats.

Sistema	Toxicitat en el peix *fathead minnow*		Sorció sòl-aigua		Permeabilitat a la pell		
	d	**$SD_{corr\ cal}^2$**	***d***	**$SD_{corr\ cal}^2$**	***d***	**$SD_{corr\ cal}^2$** sense corrección del volum del solut	**$SD_{corr\ cal}^2$** amb corrección del volum del solut
RP18 HPLC	0.259	0.237	0.243	0.220	0.200	0.626	0.263
MSC18 HPLC	0.295	0.300	0.265	0.265	0.181	0.621	0.258
IAM HPLC	0.186	0.132	0.276	0.156	0.316	0.816	0.355
SDS MEKC	0.347	0.210	0.327	0.143	0.412	1.229	0.368
STC MEKC	0.147	0.113	0.249	0.128	0.323	1.045	0.398
Octanol-aigua	0.198	0.170	0.272	0.191	0.268	0.906	0.284

Després d'aplicar el mètode de predicció de variàncies entre el sistema de permeabilitat a la pell i cadascun dels sistemes cromatogràfics seleccionats, a la Taula 15 es pot observar que la variància total que es prediu és molt alta per a totes les correlacions (veure $SD_{corr\ cal}^2$ sense correcció del volum del solut).

Si es revisen les contribucions de variància (SD_{bio}^2, $(p_{cal}\ SD_{crom})^2$ i SD_d^2) que han donat lloc a aquests valors de $SD_{corr\ cal}^2$ (Taula 3, article V), per una banda es veu que SD_{bio}^2 és força alt per a la permeabilitat en pell (concretament 0.213, mentre que per a la toxicitat en el peix *fathead minnow* era 0.076 i per a la sorció sòl-aigua 0.062). Per altra banda, els valors de SD_d^2 són molt superiors als de SD_{bio}^2 per a totes les correlacions entre la permeabilitat en pell i els sistemes cromatogràfics, essent aquest el principal factor pel qual es prediuen valors tan alts de $SD_{corr\ cal}^2$. Per tant, no s'espera obtenir bons models de correlació entre dades de log K_p i dades de log k dels sistemes cromatogràfics seleccionats.

En aquesta situació, per a la permeabilitat en pell s'ha decidit incloure una variable addicional a les correlacions per tal d'intentar millorar-les. S'ha cregut que el volum de McGowan del solut (*V*) podia ser una variable adequada per a aquest propòsit pels següents motius:

- *V* es pot calcular fàcilment per a qualsevol solut a partir de la seva estructura [18,103].

- El volum de McGowan del solut és un descriptor molecular que en general influencia molt els processos de partició i, conseqüentment, pot ajudar a descriure'ls. És el cas de la permeabilitat en pell, ja que si es consulta la seva caracterització d'Abraham (veure apartat d'informació suplementària, Taula SI-1) es pot veure que el terme de volum té molt pes.

- Existeixen treballs en els que s'ha aconseguit millorar la correlació entre dues propietats fent servir *V* com a variable addicional [3,70,142].

- Concretament, alguns autors han demostrat que *V* és un descriptor adient per descriure la permeabilitat a través de la pell [143-145]. De fet, són molts els models QSAR [146-149] que han estat desenvolupats per a la permeabilitat en pell basant-se en dos termes que es consideren clau en aquest procés: un terme de lipofilicitat i un altre de mida molecular. Per això, sembla coherent correlacionar les dades de log K_p respecte a les de log k i *V*, amb la intenció que els factors de retenció representin el terme de lipofilicitat i el volum representi el terme de mida molecular del compost.

En incloure *V* com a variable addicional a les correlacions, l'equació 17 passa a tenir la següent expressió:

$$\log SP_{bio} = q + p\,(\log SP_{crom} + t/p\ V) \qquad (24)$$

Per tant, les correlacions amb correcció de volum tindran dues variables independents: log SP_{crom} i *V*. Pel que fa a *t*, correspon al coeficient de *V* a la correlació (*t/p* és la correcció de volum al terme cromatogràfic).

Les contribucions de variància per a les correlacions entre la permeabilitat en pell i els sistemes cromatogràfics de la Taula 15 s'han calculat novament considerant *V* com a variable independent addicional (equació 24). La mateixa Taula 15 mostra que les variàncies totals que es prediuen per a les correlacions amb correcció de volum són molt inferiors a les que s'havien obtingut sense fer servir el volum. De fet, amb correcció de volum, SD_d^2 és inferior a SD_{bio}^2 per a tots els casos (Taula 6, article V) i per

això els valors de $SD_{corr\ cal}^2$ que es mostren a la Taula 15 són baixos. Per tant, es preveu que l'addició del volum a les correlacions permeti estimar correctament valors de log K_p a partir de mesures en els sistemes cromatogràfics seleccionats.

En resum, d'acord amb les prediccions de $SD_{corr\ cal}^2$ que es detallen a la Taula 15, els sistemes IAM HPLC i STC MEKC sembla que seran els que millor prediguin valors de -log LC_{50}. També es preveu que aquests sistemes, juntament amb el de SDS MEKC, siguin els que proporcionin models de correlació més precisos per estimar dades de log K_{OC}. Pel que fa a l'estimació de log K_p mitjançant mesures cromatogràfiques, s'espera que els sistemes més adients siguin els d'HPLC basats en columnes C18.

3.3.2.2. Correlacions experimentals

La bondat dels 6 sistemes seleccionats per emular els 3 processos de biopartició de toxicitat en el peix *fathead minnow*, sorció sòl-aigua i permeabilitat en pell de substàncies orgàniques neutres s'ha comprovat mitjançant correlacions experimentals. S'han correlacionat dades experimentals de cada propietat de biopartició (-log LC_{50}, log K_{OC} i log K_p respectivament) respecte a dades experimentals de cada sistema seleccionat (log k en tots els casos, excepte log $P_{o/w}$ pel sistema de partició octanol-aigua), el que correspon a l'equació 17. Pel sistema de permeabilitat a través de la pell, també s'han fet correlacions incloent el volum del solut (*V*) com a variable independent addicional, el que correspon a l'equació 24.

Els valors experimentals de les propietats de biopartició s'han obtingut de la literatura, concretament dels articles en què es troben publicades les caracteritzacions d'Abraham que s'han considerat pel sistema de toxicitat en el peix *fathead minnow* [113], la sorció sòl-aigua [110] i la permeabilitat a través de la pell [112]. Pel que fa a les dades experimentals dels sistemes cromatogràfics, tots els valors de log k s'han determinat al laboratori injectant les substàncies en les condicions característiques de cada sistema (veure Taula 12). Els valors de log $P_{o/w}$ són els únics que s'han obtingut sense treball experimental gràcies a la base de dades Bio-Loom [44].

De cara a obtenir correlacions experimentals representatives, ha calgut examinar prèviament els soluts a correlacionar. Això s'ha fet a través de l'anàlisi per components principals. A la Figura 18 es mostra un exemple de selecció de substàncies per

correlacionar la sorció sòl-aigua amb mesures cromatogràfiques en els sistemes d'HPLC.

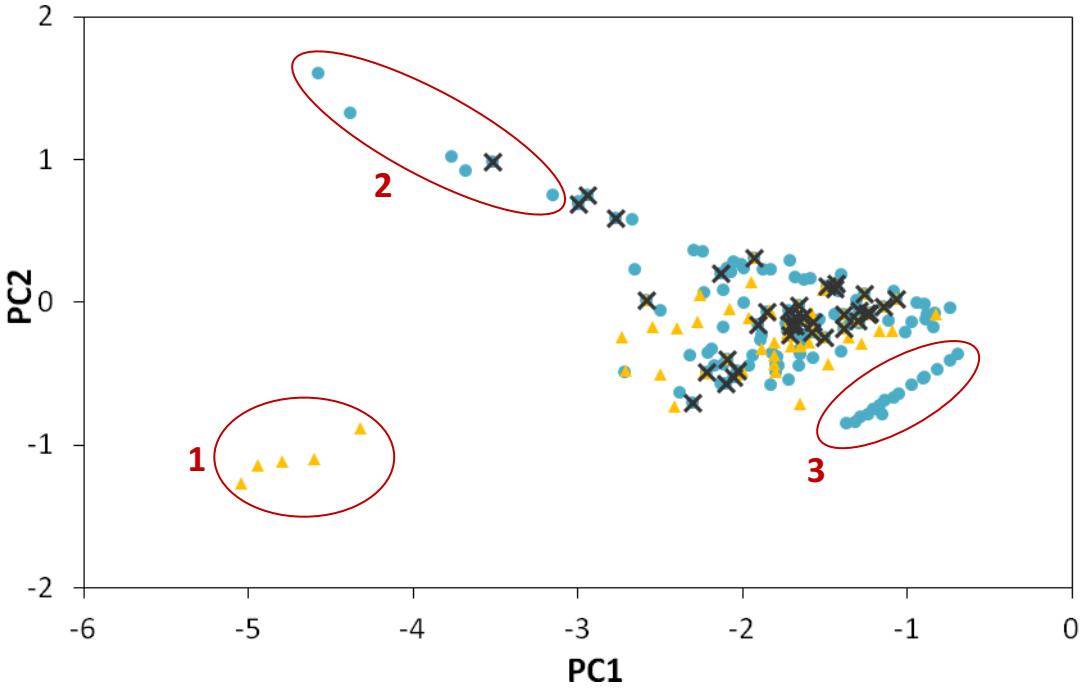


Figura 18. Gràfic de *scores* corresponent al PCA dels descriptors dels soluts inclosos a la caracterització del sistema de sorció sòl-aigua (●), projecció dels descriptors dels soluts que formen part de la caracterització dels sistemes d'HPLC (RP18, MSC18 i IAM) (▲) i selecció dels soluts a correlacionar experimentalment (**X**).

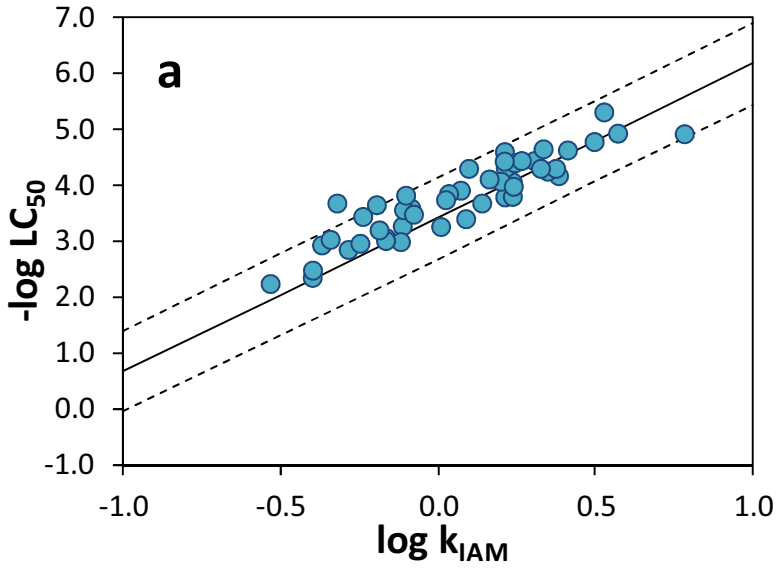
Cal assegurar-se que els compostos seleccionats cobreixen majoritàriament l'espai químic tant del sistema de biopartició com dels sistemes cromatogràfics, el que es pot veure que es compleix a la Figura 18. Només hi ha 3 grups aïllats de soluts que no seran correlacionats pels següents motius:

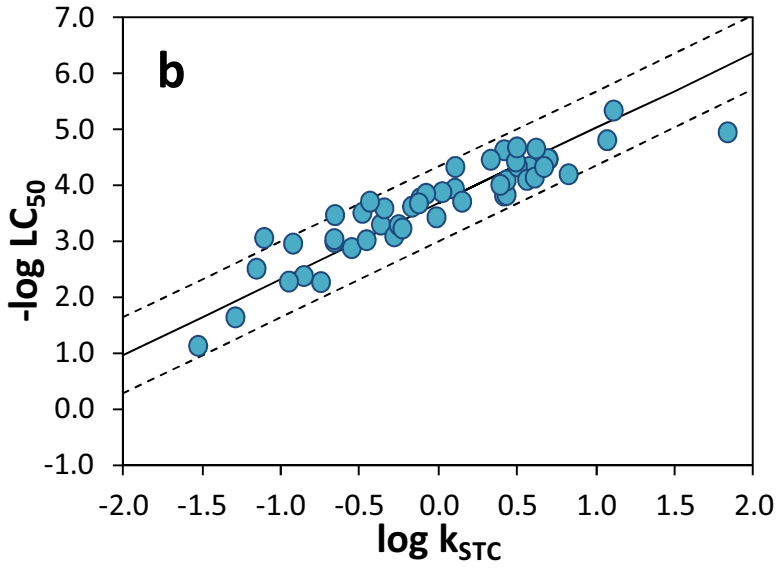
- El grup 1 correspon a 5 corticosteroides inclosos a les caracteritzacions d'Abraham dels sistemes cromatogràfics. Es troben allunyats de la resta de compostos perquè tenen descriptors moleculars molt grans en comparació amb els dels altres soluts. Aquests 5 corticosteroides no es troben a la caracterització del sistema de sorció sòl-aigua i no es coneixen els seus valors de log K_{OC}.

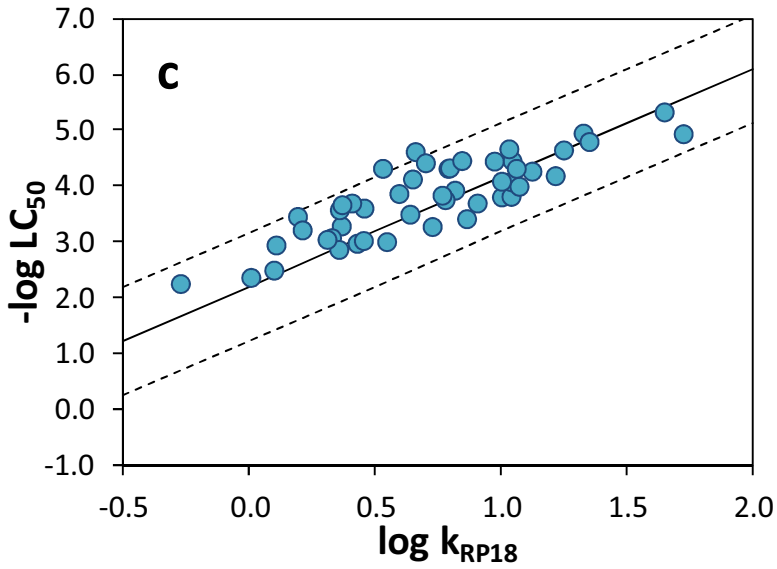
- Les substàncies no seleccionades del grup 2 són el tetracè, benzantracè, 1,2,5,6-dibenzantracè, benzoapirè i hexaclorobenzè. Es tracta de soluts que també tenen descriptors molt grans que fan que s'allunyin de la zona on es troben la majoria de compostos. Aquestes 5 substàncies no seran correlacionades experimentalment perquè els seus valors calculats tant de log K_{OC} com de log k (valors calculats a partir de les corresponents caracteritzacions d'Abraham) ja presenten molta dispersió respecte als de la resta de soluts. A més, els valors de log k calculats són majors a 2, el que correspondria a substàncies molt retingudes i llavors mesurades amb força error. Això reflecteix que els compostos amb descriptors extrems que es desvien molt poden ser una limitació de la metodologia presentada.

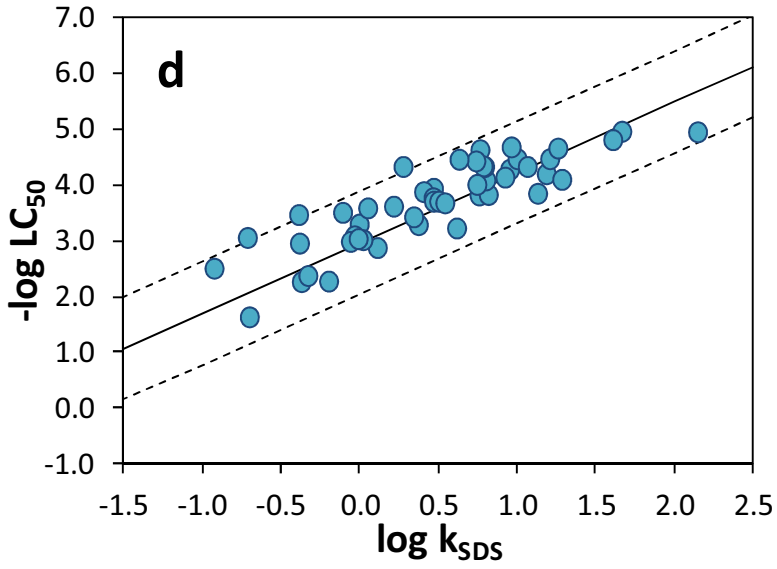
- Per últim, el grup 3 correspon a substàncies amb valors coneguts de log K_{OC} però que no es podran detectar per UV (mode de detecció utilitzat en els sistemes d'HPLC), de manera que no es podrà mesurar el seu log k.

A continuació es mostren els gràfics i les equacions que s'han obtingut a les correlacions experimentals de la toxicitat en el peix *fathead minnow* (Figura 19 i Taula 16), la sorció sòl-aigua (Figura 20 i Taula 17) i la permeabilitat a la pell (Figura 21 i Taula 18). Tant els gràfics com les equacions de les taules apareixen ordenats de més a menys precisió obtinguda. En el cas de la permeabilitat en pell, es mostren els gràfics amb i sense correcció de volum. Aquests es troben ordenats de més a menys precisió obtinguda segons les correlacions amb correcció de volum, que són les que es detallen a la Taula 18.


a
-log LC50
log kIAM
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-1.0
-0.5
0.0
0.5
1.0


b
-log LC50
log kSTC
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0


c
-log LC50
log kRP18
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0


d
-log LC50
log kSDS
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5

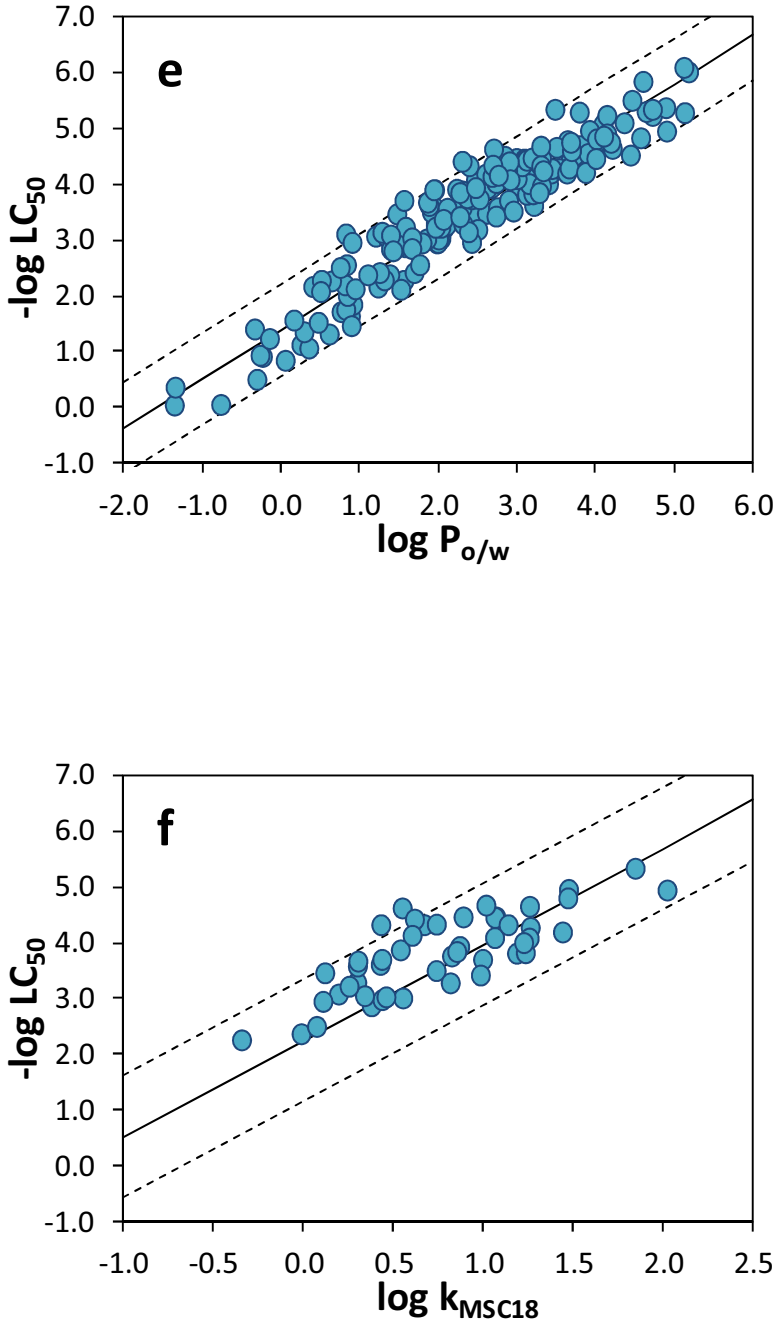
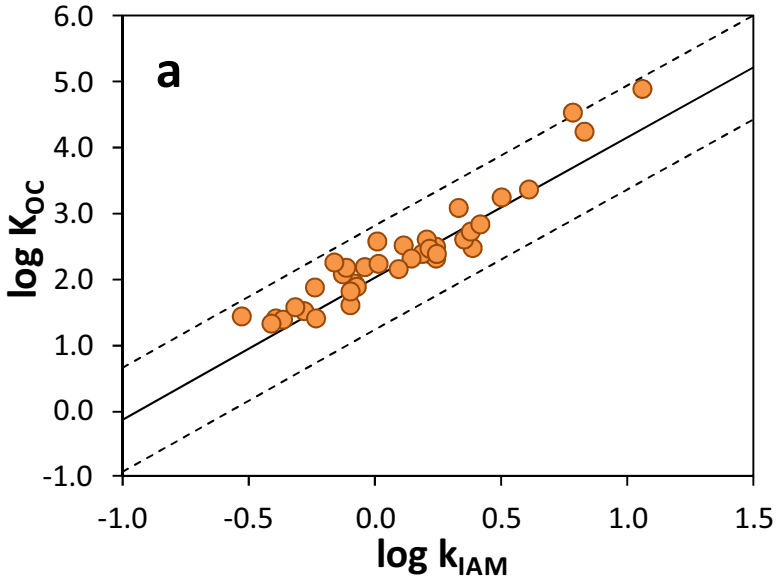
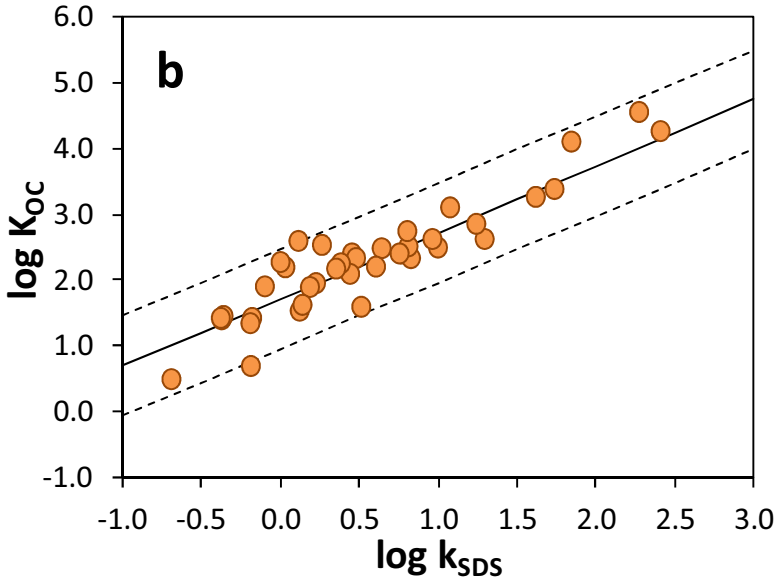


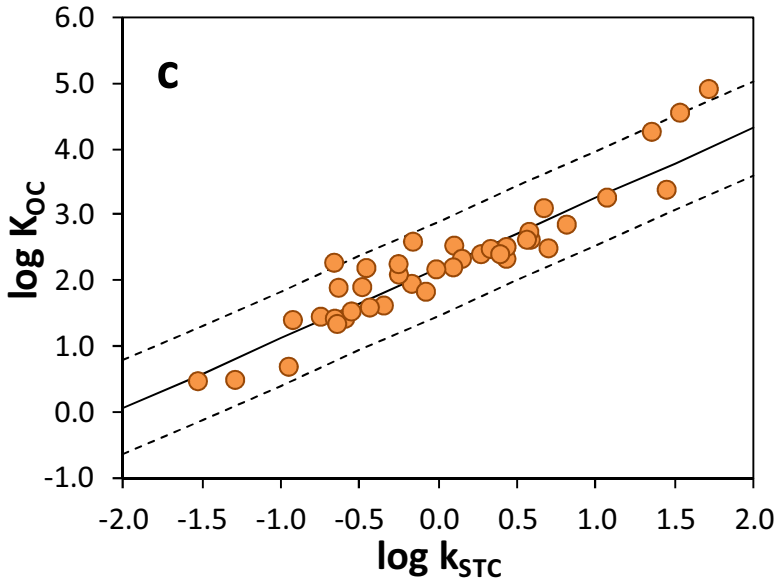
Figura 19. Correlacions entre dades experimentals de -log LC_{50} i log k (o log $P_{o/w}$) pels sistemes seleccionats: (a) IAM HPLC, (b) STC MEKC, (c) RP18 HPLC, (d) SDS MEKC, (e) octanol-aigua, (f) MSC18 HPLC. També es mostren les correlacions entre valors calculats a partir de les caracteritzacions d'Abraham (—) i els corresponents intervals per a un 95% de confiança (± 2 $SD_{corr\ cal}$) (----).

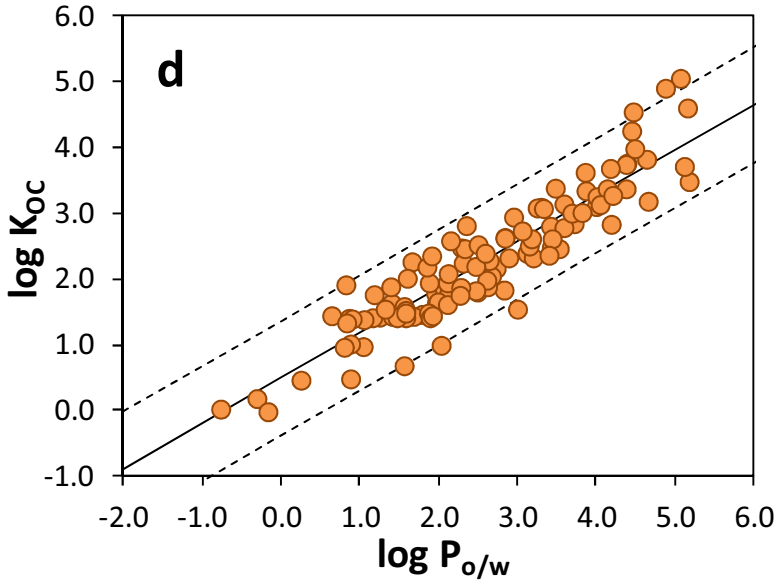
Taula 16. Resultats de les correlacions experimentals entre la toxicitat en el peix *fathead minnow* i els sistemes cromatogràfics seleccionats (desviacions estàndard entre parèntesi).

Sistema	**Toxicitat en el peix *fathead minnow***					
	q	***p***	$SD_{corr\ exp}^{2}$	***n***	***r***	***F***
IAM HPLC	3.67 (0.04)	2.20 (0.14)	0.084	48	0.915	235
STC MEKC	3.68 (0.05)	1.12 (0.07)	0.118	50	0.917	254
RP18 HPLC	2.79 (0.11)	1.43 (0.13)	0.139	48	0.854	124
SDS MEKC	3.22 (0.07)	1.02 (0.08)	0.147	49	0.868	143
Octanol-aigua	1.48 (0.06)	0.85 (0.02)	0.152	183	0.946	1530
MSC18 HPLC	2.96 (0.12)	1.13 (0.13)	0.191	48	0.792	78


a
log K_{OC}
log k_{IAM}
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-1.0
-0.5
0.0
0.5
1.0
1.5


b
log K_{OC}
log k_{SDS}
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0


c
log K_{OC}
log k_{STC}
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0


d
log K_{OC}
log $P_{o/w}$
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
-2.0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
6.0

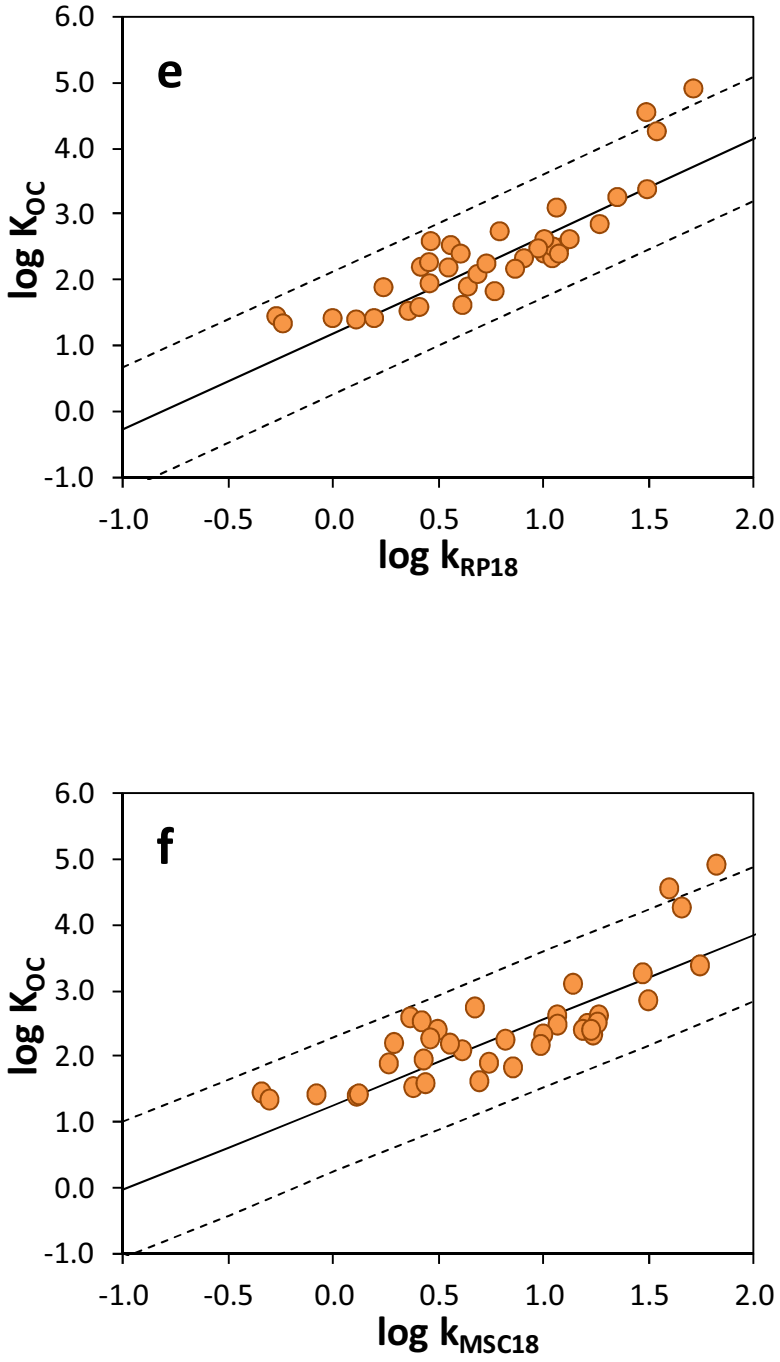
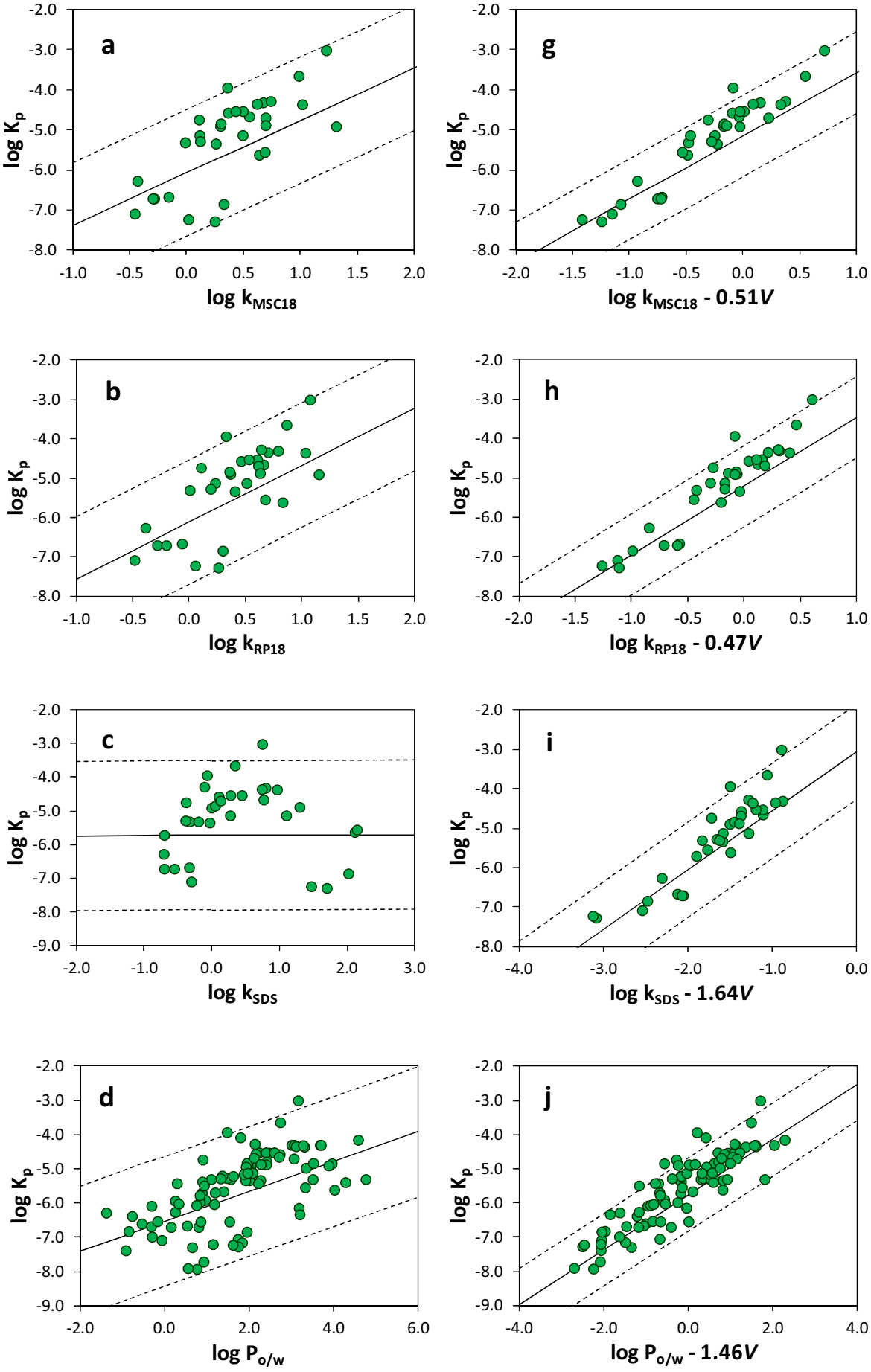


Figura 20. Correlacions entre dades experimentals de log K_{OC} i log k (o log $P_{o/w}$) pels sistemes seleccionats: (a) IAM HPLC, (b) SDS MEKC, (c) STC MEKC, (d) octanol-aigua, (e) RP18 HPLC, (f) MSC18 HPLC. També es mostren les correlacions entre valors calculats a partir de les caracteritzacions d'Abraham (—) i els corresponents intervals per a un 95% de confiança (± 2 $SD_{corr\ cal}$) (----).

Taula 17. Resultats de les correlacions experimentals entre la sorció sòl-aigua i els sistemes cromatogràfics seleccionats (desviacions estàndard entre parèntesi).

Sistema	**Sorció sòl-aigua**					
	q	***p***	$\mathbf{SD_{corr\ exp}}^{2}$	***n***	***r***	***F***
IAM HPLC	2.20 (0.05)	2.17 (0.12)	0.075	38	0.945	303
SDS MEKC	1.72 (0.07)	1.08 (0.07)	0.111	39	0.923	213
STC MEKC	2.23 (0.05)	1.14 (0.07)	0.113	41	0.935	270
Octanol-aigua	0.46 (0.08)	0.71 (0.03)	0.150	114	0.922	633
RP18 HPLC	1.30 (0.13)	1.48 (0.15)	0.188	38	0.856	99
MSC18 HPLC	1.46 (0.14)	1.19 (0.15)	0.245	38	0.808	68


a
log K_p
log k_MSC18
g
log K_p
log k_MSC18 - 0.51V
b
log K_p
log k_RP18
h
log K_p
log k_RP18 - 0.47V
c
log K_p
log k_SDS
i
log K_p
log k_SDS - 1.64V
d
log K_p
log P_o/w
j
log K_p
log P_o/w - 1.46V

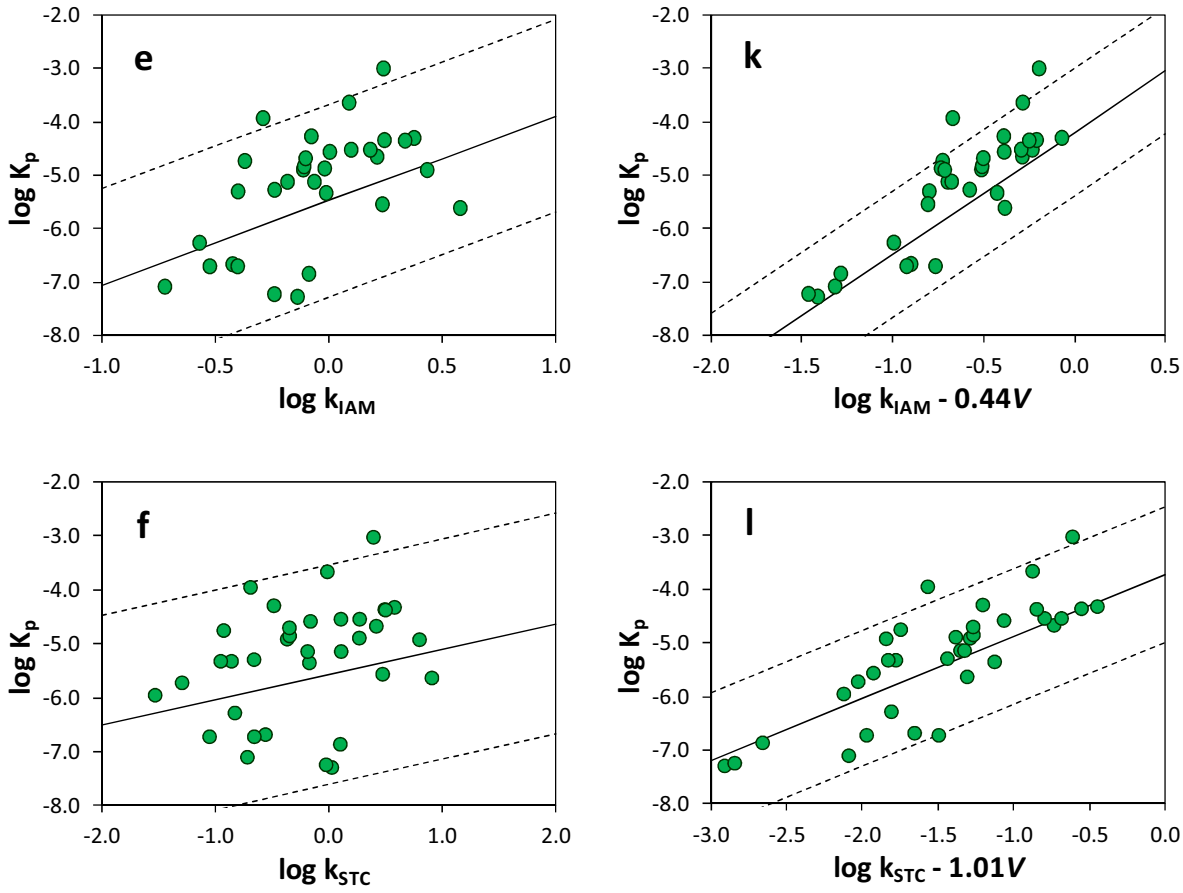


Figura 21. Correlacions entre dades experimentals de log K_p i log k (o log $P_{o/w}$) pels sistemes seleccionats: (a) MSC18 HPLC, (b) RP18 HPLC, (c) SDS MEKC, (d) octanol-aigua, (e) IAM HPLC, (f) STC MEKC. Correlacions amb correcció de volum, log K_p respecte a log k (o log $P_{o/w}$) i *V* pels mateixos sistemes: (g) MSC18 HPLC, (h) RP18 HPLC, (i) SDS MEKC, (j) octanol-aigua, (k) IAM HPLC, (l) STC MEKC. També es mostren les correlacions entre valors calculats a partir de les caracteritzacions d'Abraham (—) i els corresponents intervals per a un 95% de confiança (± 2 $SD_{corr\ cal}$) (----).

Taula 18. Resultats de les correlacions experimentals entre la permeabilitat en pell i els sistemes cromatogràfics seleccionats, aplicant la correcció de volum (desviacions estàndard entre parèntesi).

Sistema	**Permeabilitat a la pell**						
	q	***p***	***t***	$SD_{corr\ exp}^2$	***n***	***r***	***F***
MSC18 HPLC	-4.62 (0.14)	2.02 (0.14)	-1.04 (0.09)	0.118	32	0.953	144
RP18 HPLC	-4.76 (0.15)	2.14 (0.16)	-1.01 (0.10)	0.139	32	0.945	120
SDS MEKC	-2.37 (0.25)	1.73 (0.17)	-2.85 (0.23)	0.208	33	0.913	75
Octanol-aigua	-5.44 (0.11)	0.79 (0.04)	-1.15 (0.09)	0.239	96	0.885	168
IAM HPLC	-3.58 (0.24)	2.56 (0.35)	-1.12 (0.16)	0.335	32	0.860	41
STC MEKC	-3.20 (0.29)	1.38 (0.21)	-1.39 (0.20)	0.410	35	0.809	30

Es pot observar que pràcticament tots els punts experimentals de les figures 19-21 es troben dins de l'interval de confiança determinat per les prediccions de variància (± 2 $SD_{corr\ cal}$ per a un 95% de confiança). Aquest fet és independent del sistema de biopartició i cromatogràfic correlacionats, de la seva dissimilitud i del número de punts experimentals. Per tant, posa de manifest que les prediccions de variància s'avenen amb els resultats experimentals.

Donat que els gràfics de cada figura estan ordenats d'acord amb valors creixents de la variància obtinguda a les correlacions ($SD_{corr\ exp}^2$), també s'observa com cada cop els punts apareixen més dispersos.

A la Figura 21 es poden comparar les correlacions de permeabilitat en pell amb i sense correcció de volum. La comparació dels gràfics de cada sistema cromatogràfic és una demostració visual clara que les correlacions milloren de manera molt significativa quan s'afegeix *V* com a variable independent. L'ajust millora de manera especialment espectacular per a la correlació de SDS MEKC. Es tracta d'un fet coherent si es consulten les caracteritzacions d'Abraham (apartat d'informació suplementària): el sistema SDS MEKC és el sistema cromatogràfic amb un coeficient normalitzat de volum (v_u) més diferent respecte al sistema de permeabilitat en pell, de manera que és lògic que la correlació d'ambdós sistemes és la que quedi més corregida a l'afegir el volum del solut com a variable independent.

La Taula 19 recull de manera conjunta les variàncies experimentals que s'han obtingut a les correlacions mostrades a les figures 19-21.

Si es compara la Taula 19 amb la Taula 15, es pot observar que les variàncies experimentals concorden amb el que s'havia predit, de manera que els valors de $SD_{corr\ cal}^2$ i $SD_{corr\ exp}^2$ condueixen a les mateixes conclusions.

Els sistemes cromatogràfics IAM HPLC i STC MEKC són els que emulen millor la toxicitat en el peix *fathead minnow*, tal com demostren els baixos valors de $SD_{corr\ exp}^2$ i la poca dispersió dels punts a les Figures 19a i 19b. Per tant, és possible predir de manera precisa valors de -log LC_{50} de compostos orgànics neutres a partir de mesures cromatogràfiques en aquests sistemes, el que concorda amb el que es preveia a partir de la predicció de variàncies. La resta de sistemes no modelen tan bé aquest procés de

biopartició però encara ho fan de manera correcta, ja que les variàncies experimentals no són gaire altes. El sistema MSC18 HPLC és el que té menys habilitat per modelar-lo, tal com s'havia previst (la correlació d'aquest sistema té el valor més elevat tant de $SD_{corr\ exp}^2$ com de $SD_{corr\ cal}^2$).

Taula 19. Variàncies obtingudes a les correlacions experimentals dels sistemes de biopartició i cromatogràfics.

Sistema	Toxicitat en el peix *fathead minnow*	Sorció sòl-aigua	Permeabilitat a la pell	
	$SD_{corr\ exp}^2$	$SD_{corr\ exp}^2$	$SD_{corr\ exp}^2$ sense correcció del volum del solut	$SD_{corr\ exp}^2$ amb correcció del volum del solut
RP18 HPLC	0.139	0.188	0.596	0.139
MSC18 HPLC	0.191	0.245	0.602	0.118
IAM HPLC	0.084	0.075	0.870	0.335
SDS MEKC	0.147	0.111	1.203	0.208
STC MEKC	0.118	0.113	1.024	0.410
Octanol-aigua	0.152	0.150	0.666	0.239

Pel que fa a la sorció en sòls, els models de correlació més precisos els proporcionen els sistemes IAM HPLC, SDS MEKC i STC MEKC. Les prediccions de variància ja ho auguraven i ha quedat comprovat experimentalment (les seves correlacions són les que tenen menys $SD_{corr\ exp}^2$). També en concordança amb les prediccions, els sistemes d'HPLC basats en columnes C18 són els menys adequats per emular aquest procés de biopartició (a les Figures 20e i 20f s'observa que les seves correlacions són les que presenten més dispersió).

Quant a l'emulació de la permeabilitat a través de la pell, queda demostrat experimentalment que la retenció cromatogràfica per si sola no és suficient per descriure bé aquest procés. En canvi, sí que s'obtenen bones correlacions respecte a valors de log K_p si es fan servir conjuntament com a variables independents la retenció cromatogràfica i el volum del solut. Aquest fet reflecteix els fonaments del procés de

permeabilitat d'una substància aquosa a través de la pell: depèn de la lipofilicitat del compost, que és un terme que contribueix de manera positiva en el procés i que queda representat per la retenció cromatogràfica, i també depèn de la mida molecular del compost, que contribueix de manera negativa i que queda representat pel volum de McGowan del solut. Experimentalment es comprova que quan les correlacions es corregeixen amb el volum del solut, els sistemes d'HPLC RP18 i MSC18 es correlacionen de manera precisa amb la permeabilitat en pell (les seves correlacions tenen poc error, s'ha obtingut un valor baix de $SD_{corr\ exp}^2$). La resta de sistemes, especialment el de IAM HPLC i sobre tot el de STC MEKC, donen lloc a models de correlació amb una dispersió considerablement major i, per tant, no són gaire adequats per modelar aquest procés de biopartició.

3.3.2.3. Aplicacions il·lustratives: estimació de valors de K_{OC} i K_p

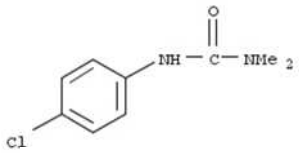
Els models de correlació més precisos obtinguts per a la sorció sòl-aigua s'han emprat per predir valors de log K_{OC} de diversos herbicides. Es tracta dels models de correlació corresponents als sistemes cromatogràfics IAM HPLC, STC MEKC i SDS MEKC (veure Taula 17).

Els herbicides són compostos agroquímics d'ús molt estès degut a la seva utilitat per eliminar les males herbes i les plantes no desitjades. No obstant, també poden tenir efectes negatius sobre el medi ambient i la salut, essent llavors contraproduents. És fonamental conèixer el seu valor de log K_{OC} i per això han estat escollits per avaluar la capacitat predictiva dels models de correlació cromatogràfics.

Els herbicides estudiats es mostren a la Figura 22. Pertanyen a dues famílies: les fenilurees i les triazines. A la figura també es troba detallat el seu valor de log K_{OC} experimental, que s'ha extret de la literatura [150].

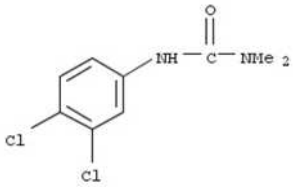
Aquests herbicides han estat injectats en els sistemes IAM HPLC, STC MEKC i SDS MEKC. Seguidament, el seu log K_{OC} ha estat predit a partir del log k determinat a cada sistema, tot emprant el corresponent model de correlació (equació obtinguda experimentalment per a cada sistema a l'apartat anterior, Taula 17).

FENILUREES



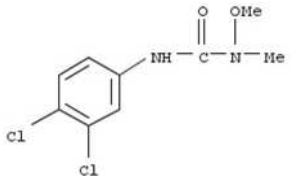
Monuró

log K_{OC} exp = 1.95
log K_{OC} $pred_{IAM\ HPLC}$ = 1.93
log K_{OC} $pred_{STC\ MEKC}$ = 2.17
log K_{OC} $pred_{SDS\ MEKC}$ = 2.09
log K_{OC} $pred_{model\ solvatació}$ = 2.22



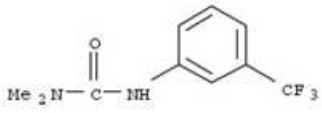
Diuró

log K_{OC} exp = 2.40
log K_{OC} $pred_{IAM\ HPLC}$ = 2.57
log K_{OC} $pred_{STC\ MEKC}$ = 2.92
log K_{OC} $pred_{SDS\ MEKC}$ = 2.70
log K_{OC} $pred_{model\ solvatació}$ = 2.73



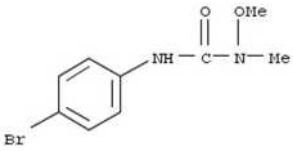
Linuró

log K_{OC} exp = 2.70
log K_{OC} $pred_{IAM\ HPLC}$ = 2.73
log K_{OC} $pred_{STC\ MEKC}$ = 3.06
log K_{OC} $pred_{SDS\ MEKC}$ = 2.89
log K_{OC} $pred_{model\ solvatació}$ = ---



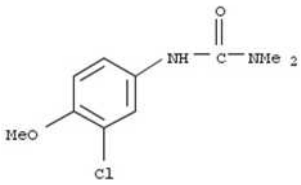
Fluometuró

log K_{OC} exp = 2.00
log K_{OC} $pred_{IAM\ HPLC}$ = 2.04
log K_{OC} $pred_{STC\ MEKC}$ = 2.37
log K_{OC} $pred_{SDS\ MEKC}$ = 2.34
log K_{OC} $pred_{model\ solvatació}$ = 2.01



Metobromuró

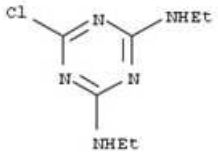
log K_{OC} exp = 2.10
log K_{OC} $pred_{IAM\ HPLC}$ = 2.25
log K_{OC} $pred_{STC\ MEKC}$ = 2.54
log K_{OC} $pred_{SDS\ MEKC}$ = 2.47
log K_{OC} $pred_{model\ solvatació}$ = ---



Metoxuró

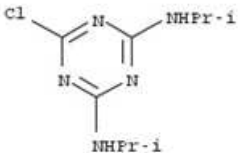
log K_{OC} exp = 1.72
log K_{OC} $pred_{IAM\ HPLC}$ = 1.71
log K_{OC} $pred_{STC\ MEKC}$ = 2.08
log K_{OC} $pred_{SDS\ MEKC}$ = 2.21
log K_{OC} $pred_{model\ solvatació}$ = 2.10

TRIAZINES



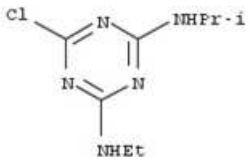
Simazina

log K_{OC} exp = 2.10
log K_{OC} $pred_{IAM\ HPLC}$ = 1.78
log K_{OC} $pred_{STC\ MEKC}$ = 2.04
log K_{OC} $pred_{SDS\ MEKC}$ = 2.33
log K_{OC} $pred_{model\ solvatació}$ = 1.95



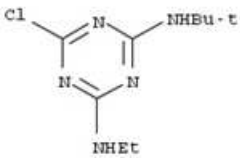
Propazina

log K_{OC} exp = 2.40
log K_{OC} $pred_{IAM\ HPLC}$ = 2.31
log K_{OC} $pred_{STC\ MEKC}$ = 2.38
log K_{OC} $pred_{SDS\ MEKC}$ = 2.96
log K_{OC} $pred_{model\ solvatació}$ = 2.35



Atrazina

log K_{OC} exp = 2.24
log K_{OC} $pred_{IAM\ HPLC}$ = 2.05
log K_{OC} $pred_{STC\ MEKC}$ = 2.22
log K_{OC} $pred_{SDS\ MEKC}$ = 2.64
log K_{OC} $pred_{model\ solvatació}$ = 2.15



Terbutilazina

log K_{OC} exp = 2.32
log K_{OC} $pred_{IAM\ HPLC}$ = 2.48
log K_{OC} $pred_{STC\ MEKC}$ = 2.68
log K_{OC} $pred_{SDS\ MEKC}$ = 3.26
log K_{OC} $pred_{model\ solvatació}$ = 2.66

Figura 22. Herbicides estudiats amb els corresponents valors de log K_{OC} experimentals i predits.

D'altra banda, el log K_{OC} dels herbicides estudiats també s'ha estimat mitjançant la caracterització de la sorció sòl-aigua a partir del model de paràmetres de solvatació [110] (la caracterització es pot consultar a l'apartat d'informació suplementària, Taula SI-1; per obtenir els valors de log K_{OC} s'han multiplicat els coeficients LSER pels descriptors dels herbicides, encara que pel linuró i metobromuró no ha estat possible perquè no es coneixen els seus descriptors).

Els valors de log K_{OC} estimats tant a partir de cadascun dels sistemes cromatogràfics com a partir del model de paràmetres de solvatació es detallen també a la Figura 22.

Es pot observar que els log K_{OC} predits més pròxims als experimentals són els que s'obtenen a partir de mesures cromatogràfiques en el sistema IAM HPLC. Això concorda amb els resultats experimentals de l'apartat anterior, en els que es va veure que aquest sistema era el que millor emulava la sorció sòl-aigua.

Quan es representen les dades de log K_{OC} estimades en cada cas respecte a les experimentals, s'obtenen les ordenades a l'origen, els pendents i les desviacions estàndard que es mostren a la Taula 20.

Taula 20. Valors de *q*, *p* i SD corresponents a les correlacions de log K_{OC} estimats a partir dels sistemes cromatogràfics IAM STC, STC MEKC i SDS MEKC i a partir del model de paràmetres de solvatació respecte a log K_{OC} experimentals pels herbicides estudiats.

log K_{OC} pred mitjançant:	**log K_{OC} pred vs. log K_{OC} exp**		
	q	***p***	**SD**
IAM STC	-0.20 (0.44)	1.09 (0.20)	0.167
STC MEKC	0.26 (0.59)	0.99 (0.27)	0.224
SDS MEKC	0.31 (0.65)	1.04 (0.30)	0.247
Model de paràmetres de solvatació	0.57 (0.79)	0.80 (0.37)	0.235

El fet que *q* no sigui significativament diferent de 0 ni *p* significativament diferent de 1 per a un 95% de confiança indica que els valors de log K_{OC} es prediuen bé en tots els

casos. A partir dels sistemes cromatogràfics s'obté una major precisió que a partir del model de paràmetres de solvatació, donat que l'error de *p* i *q* és menor. A més, l'estimació de log K_{OC} mitjançant mesures cromatogràfiques no requereix l'ús de descriptors (evitant, per exemple, fets com que el K_{OC} del linuró i metobromuró no es poden predir a partir de l'equació de solvatació perquè no es coneixen descriptors moleculars per a aquests soluts).

També s'han fet servir els models de correlació més precisos obtinguts per a la permeabilitat en pell, és a dir, els dels sistemes d'HPLC RP18 i MSC18 (veure Taula 18), per establir un mètode de predicció de log K_p mitjançant mesures cromatogràfiques. Donat que ambdós sistemes cromatogràfics es basen en columnes C18, les quals solen ser força similars quant a propietats de partició, s'ha comprovat que en general els sistemes d'HPLC amb fases estacionàries C18 són igual d'adequats per modelar la permeabilitat en pell. Per això, de cara a establir un mètode de predicció de valors de K_p s'ha escollit una columna C18 diferent de les emprades anteriorment (XBridge C18, en comptes de XTerra RP18 i XTerra MSC18). En aquest cas, al ser una columna diferent, s'ha fet un calibratge previ.

El conjunt de substàncies escollides per establir aquest mètode cromatogràfic ha constat de 27 compostos amb característiques variades i dades experimentals de K_p conegudes. El llistat de compostos, els seus valors experimentals de log K_p i els corresponents valors del descriptor de volum es poden consultar a la Taula 8 de l'article V. Aquests valors experimentals s'han obtingut a través de diverses fonts bibliogràfiques [112,149,151,152], fent la mitjana en el cas de disposar de valors diferents segons la font.

Després d'injectar totes les substàncies en el sistema d'HPLC format per la columna XBridge C18, s'han correlacionat les dades experimentals de log K_p respecte a les de log k i el volum. La correlació global obtinguda correspon a la següent equació:

$$\log K_p = -4.76\ (\pm 0.18) + 1.44\ (\pm 0.12)\log k - 1.16(\pm 0.10)V \qquad (25)$$

$SD^2 = 0.152$, $n = 27$, $r = 0.913$, $F = 126$

L'equació 25 mostra que l'ajust de la correlació és molt semblant al que s'havia obtingut prèviament amb les columnes XTerra C18, verificant així que el sistema d'HPLC amb la columna XBridge C18 és igualment adequat per predir valors de log K_p.

La predicció de coeficients de permeabilitat s'ha validat dividint els soluts estudiats en un *training set* de 14 soluts i un *test set* de 13 soluts (ordenant els soluts per log K_p creixent i seleccionant de manera alterna els d'un *set* i els de l'altre, segons es mostra a la Taula 8 de l'article V).

Amb el *training set* s'ha establert un model de correlació entre dades de log K_p i dades de log k i *V*. L'equació 26 mostra el resultat d'aquesta correlació, on s'observa que s'ha obtingut un bon ajust i que la precisió concorda amb l'esperada pels sistemes d'HPLC basats en columnes C18:

$$\log K_p = -4.69\ (\pm 0.26) + 1.44\ (\pm 0.18)\log k - 1.19(\pm 0.13)V \quad (26)$$

$SD^2 = 0.129$, $n = 14$, $r = 0.970$, $F = 86$

El model de correlació establert amb el *training set* s'ha fet servir per predir valors de log K_p dels soluts del *test set*, avaluant així la capacitat predictiva del mètode. L'equació 27 correspon als log K_p predits respecte als experimentals pel *test set*:

$$\log K_p \text{ pred} = -0.56\ (\pm 0.63) + 0.90\ (\pm 0.10)\log K_p \text{ exp} \quad (27)$$

$SD^2 = 0.187$, $n = 13$, $r = 0.937$, $F = 79$

S'observa una bona concordança entre els valors predits i experimentals, amb una variància baixa i un pendent molt proper a 1 i ordenada a 0, el que també es pot apreciar a la Figura 23.

Finalment, el model de correlació establert amb el *training set* s'ha fet servir per predir valors de log K_p d'un conjunt d'hormones i pesticides (concretament herbicides, pràcticament els mateixos pels quals s'ha estimat el log K_{OC} prèviament). Els valors de log K_p predits segons el mètode proposat s'han comparat amb els valors experimentals disponibles per a alguns d'aquests soluts en diferents fonts bibliogràfiques [112,149,151,152]. També s'han comparat amb els valors estimats mitjançant altres

procediments: a partir de la pròpia caracterització del sistema de permeabilitat en pell mitjançant el model de paràmetres de solvatació [109,112] i a partir de l'equació proposada per Patel i col·laboradors [149]. La Figura 24 mostra els compostos i les dades de log K_p experimentals i estimades.

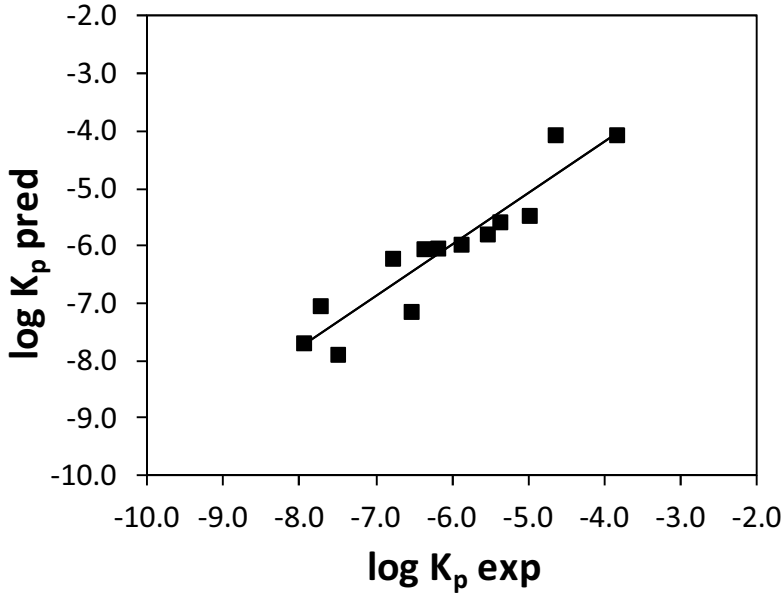
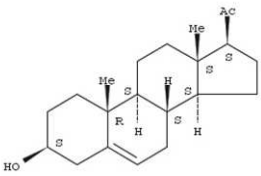


Figura 23. Correlació entre els valors de log K_p predits i experimentals pel *test set*.

Per a les hormones, els log K_p predits segons mètode proposat concorden bé amb els valors experimentals (amb un error del 5%), mentre que disten més dels valors estimats a partir de l'equació de Patel i col·laboradors (error del 10%) i sobre tot dels valors estimats amb l'equació general de solvatació (error del 15-30%).

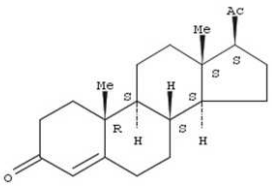
Pel que fa als herbicides, no hi ha valors experimentals disponibles ni tampoc predits a partir de l'equació de Patel i col·laboradors. Per això, els log K_p estimats segons mètode proposat només s'han comparat amb els predits a partir del model de paràmetres de solvatació, apreciant-se una diferència del 10-15%. Donat que s'ha demostrat que el model de correlació utilitzat prediu amb precisió els valors de K_p del *test set* i de les hormones, cal pensar que els valors de K_p predits pels herbicides deuen ser molt semblants als experimentals.

HORMONES



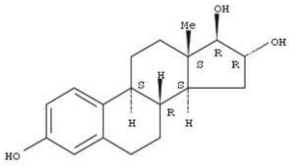
Pregnenolona

log K_p $exp_{Ref\ 149}$= -6.38
log K_p $exp_{Ref\ 152}$ = -6.38
log K_p $exp_{Ref\ 151}$ = -6.38
log K_p $pred_{model\ solvatació}$= -4.70
log K_p $pred_{equació\ Patel}$ = -6.25
log K_p $pred_{XBridge\ C18\ HPLC}$ = -6.04



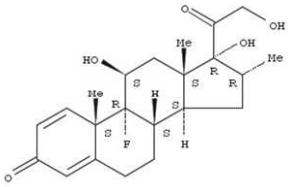
Progesterona

log K_p $exp_{Ref\ 112}$ = -4.90
log K_p $exp_{Ref\ 149}$= -6.38
log K_p $exp_{Ref\ 152}$ = -6.38
log K_p $exp_{Ref\ 151}$ = -5.92
log K_p $pred_{model\ solvatació}$= -4.54
log K_p $pred_{equació\ Patel}$ = -6.51
log K_p $pred_{XBridge\ C18\ HPLC}$ = -5.96



Estriol

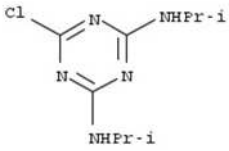
log K_p $exp_{Ref\ 149}$ = -7.96
log K_p $exp_{Ref\ 152}$ = -7.96
log K_p $exp_{Ref\ 151}$ = -7.96
log K_p $pred_{model\ solvatació}$= -6.67
log K_p $pred_{equació\ Patel}$ = -6.86
log K_p $pred_{XBridge\ C18\ HPLC}$ = -7.68



Dexametasona

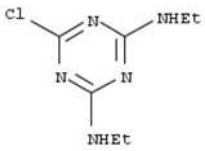
log K_p $exp_{Ref\ 112}$= -7.27
log K_p $exp_{Ref\ 151}$ = -7.75
log K_p $pred_{model\ solvatació}$= -6.71
log K_p $pred_{equació\ Patel}$ = ---
log K_p $pred_{XBridge\ C18\ HPLC}$ = -7.88

HERBICIDES



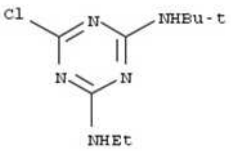
Propazina

log K_p exp = ---
log K_p $pred_{model\ solvatació}$= -5.13
log K_p $pred_{equació\ Patel}$ = ---
log K_p $pred_{XBridge\ C18\ HPLC}$ = -5.65



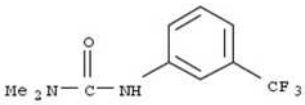
Simazina

log K_p exp = ---
log K_p $pred_{model\ solvatació}$= -5.77
log K_p $pred_{equació\ Patel}$ = ---
log K_p $pred_{XBridge\ C18\ HPLC}$ = -6.08



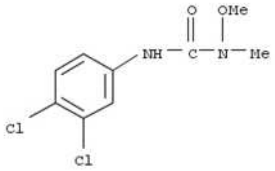
Terbutilazina

log K_p exp = ---
log K_p $pred_{model\ solvatació}$= -5.07
log K_p $pred_{equació\ Patel}$ = ---
log K_p $pred_{XBridge\ C18\ HPLC}$ = -5.55



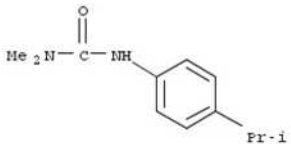
Fluometuró

log K_p exp = ---
log K_p $pred_{model\ solvatació}$= ---
log K_p $pred_{equació\ Patel}$ = ---
log K_p $pred_{XBridge\ C18\ HPLC}$ = -5.81



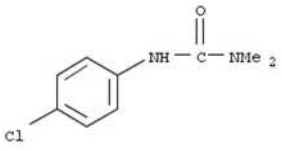
Linuró

log K_p exp = ---
log K_p $pred_{model\ solvatació}$= -4.85
log K_p $pred_{equació\ Patel}$ = ---
log K_p $pred_{XBridge\ C18\ HPLC}$ = -5.32



Isoproturó

log K_p exp = ---
log K_p $pred_{model\ solvatació}$= -5.25
log K_p $pred_{equació\ Patel}$ = ---
log K_p $pred_{XBridge\ C18\ HPLC}$ = -6.02



Monuró

log K_p exp = ---
log K_p $pred_{model\ solvatació}$= -5.43
log K_p $pred_{equació\ Patel}$ = ---
log K_p $pred_{XBridge\ C18\ HPLC}$ = -6.11

Figura 24. Comparació dels valors de log K_p experimentals i predits per a les hormones i pesticides estudiats.

CONCLUSIONS

The results obtained throughout the thesis have led to the following conclusions:

1. The characterization of mixed SC-SDC MEKC systems by the solvation parameter model has shown that the micellar phase has more capacity than the aqueous phase to interact with solute –n and –π bonding electrons, as well as it is more hydrophobic, independently of the surfactant composition. Regardless of the content in SC-SDC, the micellar phase is less dipolar/polarizable and hydrogen-bond acid than the aqueous phase. Both phases are similar in hydrogen-bond basicity.

2. Slight changes in selectivity have been observed in the characterizations of the mixed-micellar systems, which agree with the similar nature of SC and SDC. The most remarkable changes point out that the more rich in SDC is a system, the more suitable is to separate substances according to their different hydrogen-bond basicity and size.

3. The use of both Abraham descriptors and Poole optimized descriptors in the characterizations of the SC-SDC mixtures have led to the same conclusions, although better statistics have been obtained with the optimized descriptors. However, optimized descriptors are only available for some solutes. Hence, it is advisable to enlarge the database of such descriptors.

4. The solvation coefficients obtained in the characterization of the STC MEKC system have indicated that the partition of the solutes towards the micellar phase is favoured by their excess molar refraction, hydrogen-bond acidity and molar volume. On the contrary, the more dipolar and hydrogen-bond basic is a solute, the more favoured is its partition into the aqueous phase and, consequently, it becomes less retained.

5. The characterizations of both SC-SDC systems and the STC system at different pH values of the aqueous phase have shown the little influence of the pH on the solvation coefficients, which agrees with other authors' works. The factor that is really relevant and has a major effect on the accuracy of the characterizations is the

set of solutes. It has been proved that the set of solutes should be enough varied and large to cover a wide range of descriptor values. The studied MEKC systems have been characterized meeting this requirement about solutes, obtaining therefore more accurate characterizations than other authors.

6. The solvation parameter model proposed by Abraham has been proved to be very useful not only to characterize partitioning systems but also to compare biological and physicochemical processes once characterized. It provides a good approach to understand the solute-solvent interactions that control a process, as well as to examine the variance of the analyzed data and the similarity/dissimilarity between the systems.

7. The calculation of the *d* distance parameter between pairs of systems characterized through the solvation model has been observed to be very functional and suitable for an approximate evaluation of their similarity. Nevertheless, in order to study the emulation of a biological system by a physicochemical one, it has arisen the need to quantify the precision that could be obtained in the correlation of the biological and physicochemical data.

8. A new methodology has been developed to evaluate the performance of chromatographic systems to model biological ones in terms of the precision that can be achieved. It consists in analyzing the three factors that contribute to the overall variance obtained in biological-chromatographic correlations: the precision of the biological data, the precision of the chromatographic data and the dissimilarity between the two correlated systems. The contribution of these factors has been demonstrated to be easily estimated through the characterization of the systems by the solvation parameter model (using the statistics of the characterizations and the values of the properties calculated from these characterizations).

9. It has been shown that the proposed procedure provides a systematic criterion to know if a biopartitioning process is well emulated by a chromatographic system: the

contribution of variance attributed to the dissimilarity between the systems should be lower or similar to the contribution of variance attributed to the correlated data, especially the biological one, which usually has more uncertainty. In this case, the overall variance of the biological-chromatographic correlation will be low enough to allow the precise estimation of the biopartitioning property by means of chromatographic measurements. Therefore, this methodology can be used to understand the outcome of correlations between biological and chromatographic experimental data. Moreover, it is especially useful as a preliminary procedure to identify which chromatographic systems are worth being tested to model a particular biological process.

10. According to the *d* distance parameter and the procedure based on the prediction of the precision, the most practical and promising chromatographic systems for emulating the nonspecific toxicity to the *fathead minnow* fish, the soil-water sorption and the human skin permeability of neutral organic compounds have been successfully selected to be tested. They are three HPLC systems that consist of two conventional C18 columns (MSC18 and RP18) and one immobilized artificial membrane column (IAM), and two MEKC systems of SDS on one hand and STC on the other hand. The classical octanol-water partition system has also been chosen. After checking their performances by direct correlation of both biological and chromatographic experimental data, the good agreement between the predictions and the experimental results has proved the usefulness of the methodology employed.

11. It has been shown that the IAM HPLC and the STC MEKC systems are the most appropriate to provide precise estimations of nonspecific toxicity to the fathead minnow fish (-log LC_{50}). The same chromatographic systems, together with the SDS MEKC system, are also the best suited to estimate with precision soil-water sorption values (log K_{OC}). On the contrary, none of the tested systems has been seen to emulate well human skin permeation unless the volume of the solute, *V*, is included as an additional independent variable in the correlations. According to other authors' studies, it has been demonstrated that the modeling of human skin

permeation requires the combination of a lipophilicity term and a size term, which have been represented by chromatographic retention factors and volumes of the solutes in this work. With the volume correction, the HPLC systems with C18 columns, i.e. MSC18 and RP18, have shown to provide the most precise correlations to estimate human skin permeation data (log K_p values).

12. The potential of chromatographic systems to estimate biopartitioning properties has been finally illustrated by means of the prediction of log K_{OC} and log K_p values of several tested neutral compounds. The best correlation models obtained for the soil-water sorption have shown to predict log K_{OC} data of some herbicides from their log k retention with the expected precision. Regarding human skin permeation, a method has been established to estimate log K_p values from measurements in a HPLC system with a C18 column, since it has been proved to be the most suited chromatographic system for this purpose. This method has been validated by training and test sets and also applied with good results to the prediction of log K_p data of several hormones and herbicides. The successful predictions of both log K_{OC} and log K_p data obtained by this procedure have been compared to other approaches, for example predictions from the solvation parameter model or other methods based on calculations from the structure of the compounds. The precision achieved with the current methodology is comparable and even better than the approaches compared. Besides, the current methodology allows the estimation of biopartitioning properties from simple chromatographic measurements, without the requirement of descriptors, which are not always available and reduce the applicability of other approaches.

REFERÈNCIES

[1] Abraham, M. H.; Zhao, Y. H.; Le, J.; Hersey, A.; Luscombe, C. N.; Reynolds, D. P.; Beck, G.; Sherborne, B.; Cooper, I. On the mechanism of human intestinal absorption. *Eur. J. Med. Chem.* 2002, *37,* 595-605.

[2] Gratton, J. A.; Abraham, M. H.; Bradbury, M. W.; Chadha, H. S. Molecular factors influencing drug transfer across the blood-brain barrier. *J. Pharm. Pharmacol.* 1997, *49,* 1211-1216.

[3] Abraham, M. H.; Ràfols, C. Factors that influence tadpole narcosis. An LFER analysis. *J. Chem. Soc. Perkin Trans. 2,* 1995, 1843-1851.

[4] Van Leeuwen, C. J.; Hermens, J. L. M. Risk assessment of chemicals: an introduction. Kluwer Academic Publishers, 1995.

[5] http://www.epa.gov/ncct/dsstox/sdf_epafhm.html.

[6] Ankley, G. T.; Villeneuve, D. L. The fathead minnow in aquatic toxicology: past, present and future. *Aquat. Toxicol.* 2006, *78,* 91-102.

[7] http://en.wikipedia.org.

[8] Sawhney, B. L.; Brown, K. K. Reactions and movement of organic chemicals in soils. Soil Science Society of America, 1989.

[9] Wagenet, R. J.; Baveye, P.; Stewart, B. A. Interacting processes in soil science. Lewis Publishers, 1992.

[10] Boethling, R. S.; Mackay, D. Handbook of property estimation methods for chemicals. Environmental and health sciences. Lewis Publishers, 2000.

[11] Chiou, C. T. Partition and adsorption of organic contaminants in environmental systems. John Wiley & Sons, 2002.

[12] Bronner, G.; Goss, K. Sorption of organic chemicals to soil organic matter: influence of soil variability and pH dependence. *Environ. Sci. Technol.* 2011, *45,* 1307-1312.

[13] Kielhorn, J.; Melching-Kollmuss, S.; Mangelsdorf, I. Environmental health criteria 235: dermal absorption. World Health Organization, 2006.

[14] Scheuplein, R. J.; Blank, I. H. Permeability of the skin. *Physiol. Rev.* 1971, *51,* 702-747.

[15] http://www.omnimedicalsearch.com.

[16] Hansch, C.; Leo, A. Exploring QSAR. Fundamentals and applications in chemistry and biology. American Chemical Society, 1995.

[17] Hansch, C.; Leo, A.; Hoekman, D. Exploring QSAR. Hydrophobic, electronic and steric constants. American Chemical Society, 1995.

[18] Abraham, M. H. Scales of solute hydrogen-bonding: their construction and application to physicochemical and biochemical processes. *Chem. Soc. Rev.* 1993, *22,* 73-83.

[19] Abraham, M. H.; Chadha, H. S.; Martins, F.; Mitchell, R. C.; Bradbury, M. W.; Gratton, J. A. Hydrogen bonding part 46: a review of the correlation and prediction of transport properties by an LFER method: physicochemical properties, brain penetration and skin permeability. *Pestic. Sci.* 1999, *55,* 78-88.

[20] Smith, R. N.; Hansch, C.; Ames, M. M. Selection of a reference partitioning system for drug design work. *J. Pharm. Sci.* 1975, *64,* 599-606.

[21] Poole, C. F.; Gunatilleka, A. D.; Poole, S. K. In search of a chromatographic model for biopartitioning. *Adv. Chromatogr.* 2000, *40,* 159-230.

[22] Cimpean, D. M.; Poole, C. F. Systematic search for surrogate chromatographic models of biopartitioning processes. *Analyst* 2002, *127,* 724-729.

[23] Giaginis, C.; Tsantili-Kakoulidou, A. Quantitative structure-retention relationships as useful tool to characterize chromatographic systems and their potential to simulate biological processes. *Chromatographia* 2013, *76,* 211-226.

[24] Sangster, J. Octanol-water partition coefficients: fundamentals and physical chemistry. John Wiley & Sons, 1997.

[25] Caron, G.; Ermondi, G.; Scherrer, R. A. 5.18-Lipophilicity, polarity, and hydrophobicity. Comprehensive Medicinal Chemistry II. Elsevier, 2007.

[26] Meyer, H. To the theory of the alcohol arose: first report about which characteristics of the anaesthetic causes its narcotic effect. *Arch. Exp. Pathol. Pharmakol.* 1899, *42,* 109-118.

[27] Overton, E. Studien uber die Narkose. Fischer, Jena, 1901.

[28] Hansch, C.; Fujita, T. Analysis method for the correlation of biological activity and chemical structure. *J. Am. Chem. Soc.* 1964, *86,* 1616-1626.

[29] Fujita, T.; Iwasa, J.; Hansch, C. A new substituent constant, π, derived from partition coefficients. *J. Am. Chem. Soc.* 1964, *86,* 5175-5180.

[30] Winiwarter, S.; Ridderström, M.; Ungell, A.; Andersson, T. B.; Zamora, I. 5.22-Use of molecular descriptors for absorption, distribution, metabolism, and excretion predictions. Comprehensive Medicinal Chemistry II. Elsevier, 2007.

[31] Kerns, E. H.; Di, L. Drug-like properties: concepts, structure design and methods from ADME to toxicity optimization. *Eur. J. Drug Metab. Pharmacokinet.* 2008, *33,* 61-62.

[32] Danielsson, L.; Zhang, Y. Methods for determining n-octanol-water partition constants. *Trends Anal. Chem.* 1996, *15,* 188-196.

[33] EPA product properties test guidelines. OPPTS 830.7550, partition coefficient (n-octanol/water), shake flask method. 1996.

[34] Takács-Novák, K.; Avdeef, A. Interlaboratory study of log P determination by shake-flask and potentiometric methods. *J. Pharm. Biomed. Anal.* 1996, *14,* 1405-1413.

[35] Avdeef, A. pH-metric log P. Part 1. Difference plots for determining ion-pair octanol-water partition coefficients of multiprotic substances. *Quant. Struct. Act. Relat.* 1992, *11,* 510-517.

[36] Kaliszan, R.; Haber, P.; Bączek, T.; Siluk, D.; Valkó, K. Lipophilicity and pK_a estimates from gradient high-performance liquid chromatography. *J. Chromatogr. A* 2002, *965,* 117-127.

[37] Valkó, K.; Bevan, C.; Reynolds, D. Chromatographic hydrophobicity index by fast-gradient RP-HPLC: a high-throughput alternative to Log P/log D. *Anal. Chem.* 1997, *69,* 2022-2029.

[38] Rosés, M.; Bosch, E.; Ràfols, C.; Fuguet, E. Chromatographic hydrophobicity index (CHI). *Adv. Chromatogr.* 2012, *50,* 377-414.

[39] Valkó, K.; Du, C. M.; Bevan, C.; Reynolds, D. P.; Abraham, M. H. Rapid method for the estimation of octanol/water partition coefficient (Log P_{oct}) from gradient RP-HPLC retention and a hydrogen bond acidity term ($\Sigma\alpha_2^H$). *Curr. Med. Chem.* 2001, *8,* 1137-1146.

[40] Fikri, K.; Debord, J.; Bollinger, J.; Cledat, D.; Penicaut, B.; Robert, J. H. RP-HPLC lipophilicity studies for some (hetero)arylamides derived from 2-amino 4,6-dimethyl pyridine: introduction of an hydrogen-bond descriptor. *J. Liq. Chromatogr. Relat. Technol.* 2011, *34,* 1356-1366.

[41] Pallicer, J. M.; Calvet, C.; Port, A.; Rosés, M.; Ràfols, C.; Bosch, E. Extension of the liquid chromatography/quantitative structure–property relationship method to assess the lipophilicity of neutral, acidic, basic and amphotheric drugs. *J. Chromatogr. A* 2012, *1240,* 113-122.

[42] Mannhold, R.; Poda, G. I.; Ostermann, C.; Tetko, I. V. Calculation of molecular lipophilicity: state-of-the-art and comparison of log P methods on more than 96.000 compounds. *J. Pharm. Sci.* 2009, *98,* 861-893.

[43] Leo, A.; Hansch, C.; Elkins, D. Partition coefficients and their uses. *Chem. Rev.* 1971, *71,* 525-616.

[44] Bio-Loom online database: http://www.biobyte.com

[45] Sangster, J. Log K_{OW} databank. Sangster Research Laboratories, 1994.

[46] Howard, P. H.; Meylan, W. PHYSPROP database. Syracuse Research Corporation, 2000.

[47] Skoog, D. A.; Holler, F. J.; Nieman, T. A. Principios de análisis instrumental. McGraw-Hill, 2001.

[48] Miller, M. J. Chromatography. Concepts and contrasts. John Wiley & Sons, 2005.

[49] Lázaro, E.; Ràfols, C.; Rosés, M. Characterization of immobilized artificial membrane (IAM) and XTerra columns by means of chromatographic models. *J. Chromatogr. A* 2005, *1081,* 163-173.

[50] Gawlik, B. M.; Feicht, E. A.; Karcher, W.; Kettrup, A.; Muntau, H. Application of the European reference soil set (EUROSOILS) to a HPLC-screening method for the estimation of soil adsorption coefficients of organic compounds. *Chemosphere* 1998, *36,* 2903-2919.

[51] Szabó, G.; Guczi, J.; Kördel, W.; Zsolnay, A.; Major, V.; Keresztes, P. Comparison of different HPLC stationary phases for determination of soil-water distribution coefficient, K_{OC} values of organic chemicals in RP-HPLC system. *Chemosphere* 1999, *39,* 431-442.

[52] Bermúdez-Saldaña, J. M.; Escuder-Gilabert, L.; Medina-Hernández, M. J.; Villanueva-Camañas, R. M.; Sagrado, S. Chromatographic estimation of the soil-sorption coefficients of organic compounds. *Trends Anal. Chem.* 2006, *25,* 122-132.

[53] de Voogt, P.; Wegener, J. W. M.; Brinkman, U. A. T.; Govers, H. Retention of neutral and basic heteroaromatic hydrocarbons in RPLC systems and its use in predictive studies I. Concentration of the organic modifier. *Sci. Total Environ.* 1991, *109–110,* 69-87.

[54] Lázaro, E.; Ràfols, C.; Abraham, M. H.; Rosés, M. Chromatographic estimation of drug disposition properties by means of immobilized artificial membranes (IAM) and C18 Columns. *J. Med. Chem.* 2006, *49,* 4861-4870.

[55] Pidgeon, C.; Venkataram, U. V. Immobilized artificial membrane chromatography: supports composed of membrane lipids. *Anal. Biochem.* 1989, *176,* 36-47.

[56] Pidgeon, C.; Ong, S.; Liu, H.; Qiu, X.; Pidgeon, M.; Dantzig, A. H.; Munroe, J.; Hornback, W. J.; Kasher, J. S. IAM chromatography: an in vitro screen for predicting drug membrane permeability. *J. Med. Chem.* 1995, *38,* 590-594.

[57] Genty, M.; González, G.; Clere, C.; Desangle-Gouty, V.; Legendre, J. Determination of the passive absorption through the rat intestine using chromatographic indices and molar volume. *Eur. J. Pharm. Sci.* 2001, *12,* 223-229.

[58] Salminen, T.; Pulli, A.; Taskinen, J. Relationship between immobilized artificial membrane chromatographic retention and the brain penetration of structurally diverse drugs. *J. Pharm. Biomed. Anal.* 1997, *15,* 469-477.

[59] Reichel, A.; Begley, D. J. Potential of immobilized artificial membranes for predicting drug penetration across the blood-brain barrier. *Pharm. Res.* 1998, *15,* 1270-1274.

[60] Stewart, B. H.; Chan, O. H. Use of immobilized artificial membrane chromatography for drug transport applications. *J. Pharm. Sci.* 1998, *87,* 1471-1478.

[61] Taillardat-Bertschinger, A.; Carrupt, P.; Barbato, F.; Testa, B. Immobilized artificial membrane HPLC in drug research. *J. Med. Chem.* 2003, *46,* 655-665.

[62] Valkó, K. Application of high-performance liquid chromatography based measurements of lipophilicity to model biological distribution. *J. Chromatogr. A* 2004, *1037,* 299-310.

[63] Terabe, S. Electrokinetic chromatography: an interface between electrophoresis and chromatography. *Trends Anal. Chem.* 1989, *8,* 129-134.

[64] Pyell, U. Electrokinetic chromatography: theory, instrumentation and applications. John Wiley & Sons, 2006.

[65] Terabe, S.; Otsuka, K.; Ichikawa, K.; Tsuchiya, A.; Ando, T. Electrokinetic separations with micellar solutions and open-tubular capillaries. *Anal. Chem.* 1984, *56,* 111-113.

[66] Poole, C. F.; Poole, S. K. Interphase model for retention and selectivity in micellar electrokinetic chromatography. *J. Chromatogr. A* 1997, *792,* 89-104.

[67] Baker, D. R. Capillary Electrophoresis. John Wiley & Sons, 1995.

[68] Fuguet, E.; Ràfols, C.; Bosch, E.; Rosés, M. Solute-solvent interactions in micellar electrokinetic chromatography: IV. Characterization of electroosmotic flow and micellar markers. *Electrophoresis* 2002, *23,* 56-66.

[69] Yang, S.; Bumgarner, J. G.; Kruk, L. F. R.; Khaledi, M. G. Quantitative structure-activity relationships studies with micellar electrokinetic chromatography influence of surfactant type and mixed micelles on estimation of hydrophobicity and bioavailability. *J. Chromatogr. A* 1996, *721,* 323-335.

[70] Rosés, M.; Ràfols, C.; Bosch, E.; Martínez, A. M.; Abraham, M. H. Solute-solvent interactions in micellar electrokinetic chromatography. Characterization of sodium dodecyl sulfate-Brij 35 micellar systems for quantitative structure-activity relationship modelling. *J. Chromatogr. A* 1999, *845,* 217-226.

[71] Martín-Biosca, Y.; Escuder-Gilabert, L.; Marina, M. L.; Sagrado, S.; Villanueva-Camañas, R. M.; Medina-Hernández, M. J. Quantitative retention- and migration-toxicity relationships of phenoxy acid herbicides in micellar liquid chromatography and micellar electrokinetic chromatography. *Anal. Chim. Acta* 2001, *443,* 191-203.

[72] Bermúdez-Saldaña, J. M.; García, M. A.; Medina-Hernández, M. J.; Marina, M. L. Micellar electrokinetic chromatography with bile salts for predicting ecotoxicity of aromatic compounds. *J. Chromatogr. A* 2004, *1052,* 171-180.

[73] Yang, G.; Wang, Q.; Wang, S. Prediction of the pharmacodynamic properties of antihypertensive drugs by micellar electrokinetic capillary chromatography. *Asian J. Pharm. Sci.* 2007, *2,* 197-203.

[74] Woodrow, B. N.; Dorsey, J. G. Thermodynamics of micelle-water partitioning in micellar electrokinetic chromatography: comparisons with 1-octanol-water partitioning and biopartitioning. *Environ. Sci. Technol.* 1997, *31,* 2812-2820.

[75] Poole, S. K.; Poole, C. F. Influence of composition on the selectivity of a mixed-micellar buffer in micellar electrokinetic chromatography. *Anal. Commun.* 1997, *34,* 57-62.

[76] Liu, Z.; Zou, H.; Ye, M.; Ni, J.; Zhang, Y. Effects of organic modifiers on retention mechanism and selectivity in micellar electrokinetic capillary chromatography studied by linear solvation energy relationships. *J. Chromatogr. A* 1999, *863,* 69-79.

[77] Fu, C.; Khaledi, M. G. Selectivity patterns in micellar electrokinetic chromatography. *J. Chromatogr. A* 2009, *1216,* 1901-1907.

[78] Poole, S. K.; Poole, C. F. Characterization of surfactant selectivity in micellar electrokinetic chromatography. *Analyst* 1997, *122,* 267-274.

[79] Poole, C. F.; Poole, S. K.; Abraham, M. H. Recommendations for the determination of selectivity in micellar electrokinetic chromatography. *J. Chromatogr. A* 1998, *798,* 207-222.

[80] Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. Selectivity of single, mixed, and modified pseudostationary phases in electrokinetic chromatography. *Electrophoresis* 2006, *27,* 1900-1914.

[81] Dorsey, J. G. Micellar liquid chromatography. *Adv. Chromatogr.* 1987, *27,* 167-214.

[82] Marsh, A.; Clark, B. J.; Altria, K. D. A review of the background, operating parameters and applications of microemulsion liquid chromatography (MELC). *J. Sep. Sci.* 2005, *28,* 2023-2032.

[83] Breyer, E. D.; Strasters, J. K.; Khaledi, M. G. Quantitative retention-biological activity relationship study by micellar liquid chromatography. *Anal. Chem.* 1991, *63,* 828-833.

[84] Medina-Hernández, M. J.; Bonet-Domingo, E.; Ramis-Ramos, G.; García-Álvarez-Coque, M. C. On the retention of diuretics in micellar liquid chromatography and their site of action within the nephron. *Anal. Lett.* 1993, *26,* 1881-1889.

[85] Liu, J.; Sun, J.; Wang, Y.; Liu, X.; Sun, Y.; Xu, H.; He, Z. Characterization of microemulsion liquid chromatography systems by solvation parameter model and comparison with other physicochemical and biological processes. *J. Chromatogr. A* 2007, *1164,* 129-138.

[86] Ishihama, Y.; Oda, Y.; Asakawa, N. A Hydrophobicity scale based on the migration index from microemulsion electrokinetic chromatography of anionic solutes. *Anal. Chem.* 1996, *68,* 1028-1032.

[87] Xian, D.; Huang, K.; Liu, S.; Xiao, J. Liposome electrokinetic chromatography: an in vitro approach for predicting ecotoxicity. *Chromatographia* 2008, *67,* 407-412.

[88] Wang, Y.; Sun, J.; Liu, H.; Liu, J.; Zhang, L.; Liu, K.; He, Z. Predicting skin permeability using liposome electrokinetic chromatography. *Analyst* 2009, *134,* 267-272.

[89] Burns, S. T.; Khaledi, M. G. Rapid determination of liposome-water partition coefficients (K_{lw}) using liposome electrokinetic chromatography (LEKC). *J. Pharm. Sci.* 2002, *91,* 1601-1612.

[90] Watarai, H. Microemulsion capillary electrophoresis. *Chem. Lett.* 1991, 391-394.

[91] Walker, J. D.; Carlsen, L.; Hulzebos, E.; Simon-Hettich, B. Global government applications of analogues, SARs and QSARs to predict aquatic toxicity, chemical or physical properties, environmental fate parameters and health effects of organic chemicals. *SAR QSAR Environ. Res.* 2002, *13,* 607-616.

[92] Walker, J. D. Handbook on quantitative structure activity relationships (QSARs) for pollution prevention, toxicity screening, risk assessment and world wide web applications. SETAC Press, 2002.

[93] Walker, J. D. Handbook on quantitative structure activity relationships (QSARs) for predicting physical properties, bioconcentration potential and environmental fate of chemicals. SETAC Press, 2002.

[94] Walker, J. D. Handbook on quantitative structure activity relationships (QSARs) for predicting endocrine disruption potential of chemicals. SETAC Press: 2002.

[95] Walker, J. D. Handbook on quantitative structure activity relationships (QSARs) for predicting effects of chemicals on environmental-human health interactions. SETAC Press, 2002.

[96] Walker, J. D. Handbook on quantitative structure activity relationships (QSARs) for predicting ecological effects of chemicals. SETAC Press, 2002.

[97] Eriksson, L.; Jaworska, J.; Worth, A. P.; Cronin, M. T. D.; McDowell, R. M.; Gramatica, P. Methods for reliability and uncertainty assessment and for applicability evaluations of classification and regression based QSARs. *Environ. Health Perspect.* 2003, *111,* 1361-1375.

[98] E-DRAGON: http://www.vcclab.org.

[99] CODESSA Pro: http://www.codessa-pro.com.

[100] Politzer, P. S.; Murray, J. S. Quantitative treatments of solute/solvent interactions. Elsevier, 1994.

[101] Reichardt, C. Solvents and solvents effects in organic chemistry. Wiley-VCH, 2002.

[102] Vitha, M.; Carr, P. W. The chemical interpretation and practice of linear solvation energy relationships in chromatography. *J. Chromatogr. A* 2006, *1126,* 143-194.

[103] Abraham, M. H.; McGowan, J. C. The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography. *Chromatographia* 1987, *23,* 243-246.

[104] Abraham, M. H.; Grellier, P. L.; Prior, D. V.; Morris, J. J.; Taylor, P. J. Hydrogen bonding. Part 10. A scale of solute hydrogen-bond basicity using log K values for complexation in tetrachloromethane. *J. Chem. Soc. Perkin Trans. 2,* 1990, 521-529.

[105] Abraham, M. H.; Whiting, G. S.; Doherty, R. M.; Shuely, W. J. Hydrogen bonding. Part 13. A new method for the characterization of GLC stationary phases-the Laffort data set. *J. Chem. Soc. Perkin Trans. 2,* 1990, 1451-1460.

[106] Abraham, M. H.; Whiting, G. S.; Doherty, R. M.; Shuely, W. J. Hydrogen bonding. XVI. A new solute solvation parameter, π_2^H, from gas chromatographic data. *J. Chromatogr. A* 1991, *587,* 213-228.

[107] Abraham, M. H.; Ibrahim, A.; Zissimos, A. M. Determination of sets of solute descriptors from chromatographic measurements. *J. Chromatogr. A* 2004, *1037,* 29-47.

[108] Abraham, M.H. The UCL Database. University College London, 1997.

[109] ACD/Percepta, mòdul Absolv: http://www.acdlabs.com/products/percepta/predictors/absolv/

[110] Poole, S. K.; Poole, C. F. Chromatographic models for the sorption of neutral organic compounds by soil from water and air. *J. Chromatogr. A* 1999, *845,* 381-400.

[111] Platts, J. A.; Abraham, M. H.; Zhao, Y. H.; Hersey, A.; Ijaz, L.; Butina, D. Correlation and prediction of a large blood-brain distribution data set-an LFER study. *Eur. J. Med. Chem.* 2001, *36,* 719-730.

[112] Abraham, M. H.; Martins, F. Human skin permeation and partition: general linear free-energy relationship analyses. *J. Pharm. Sci.* 2004, *93,* 1508-1523.

[113] Hoover, K. R.; Acree, W. E., Jr.; Abraham, M. H. Chemical toxicity correlations for several fish species based on the Abraham solvation parameter model. *Chem. Res. Toxicol.* 2005, *18,* 1497-1505.

[114] Abraham, M. H.; Rosés, M.; Poole, C. F.; Poole, S. K. Hydrogen bonding. 42. characterization of reversed-phase high-performance liquid chromatographic C18 stationary phases. *J. Phys. Org. Chem.* 1997, *10,* 358-368.

[115] Lepont, C.; Poole, C. F. Retention characteristics of an immobilized artificial membrane column in reversed-phase liquid chromatography. *J. Chromatogr. A* 2002, *946,* 107-124.

[116] Poole, S. K.; Poole, C. F. Quantitative structure–retention (property) relationships in micellar electrokinetic chromatography. *J. Chromatogr. A* 2008, *1182,* 1-24.

[117] Abraham, M. H.; Chadha, H. S.; Whiting, G. S.; Mitchell, R. C. Hydrogen bonding. 32. An analysis of water-octanol and water-alkane partitioning and the Δlog P parameter of seiler. *J. Pharm. Sci.* 1994, *83,* 1085-1100.

[118] Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. Solute-solvent interactions in micellar electrokinetic chromatography. III. Characterization of the selectivity of micellar electrokinetic chromatography systems. *J. Chromatogr. A* 2002, *942,* 237-248.

[119] Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. Solute-solvent interactions in micellar electrokinetic chromatography III. Characterization of the selectivity of micellar electrokinetic chromatography systems. [Erratum to document cited in CA136:272381]. *J. Chromatogr. A* 2009, *1216,* 6877-6879.

[120] Fuguet, E.; Ràfols, C.; Bosch, E.; Rosés, M.; Abraham, M. H. Solute-solvent interactions in micellar electrokinetic chromatography: Selectivity of lithium dodecyl sulfate-lithium perfluorooctanesulfonate mixed-micellar buffers. *J. Chromatogr. A* 2001, *907,* 257-265.

[121] Mintz, C.; Acree, W. E., Jr.; Abraham, M. H. Correlation of minimum inhibitory concentrations toward oral bacterial growth based on the Abraham model. *QSAR Comb. Sci.* 2006, *25,* 912-920.

[122] Hoover, K. R.; Flanagan, K. B.; Acree, W. E., Jr.; Abraham, M. H. Chemical toxicity correlations for several protozoas, bacteria, and water fleas based on the Abraham solvation parameter model. *J. Environ. Eng. Sci.* 2007, *6,* 165-174.

[123] Sprunger, L. M.; Gibbs, J.; Acree, W. E., Jr.; Abraham, M. H. Linear free energy relationship correlation of the distribution of solutes between water and cetytrimethylammonium bromide (CTAB) micelles. *QSAR Comb. Sci.* 2009, *28,* 72-88.

[124] Malinowski, E. R. Factor Analysis in Chemistry. John Wiley & Sons, 1991.

[125] Ishihama, Y.; Asakawa, N. Characterization of lipophilicity scales using vectors from solvation energy descriptors. *J. Pharm. Sci.* 1999, *88,* 1305-1312.

[126] Abraham, M. H.; Zhao, Y. H. Characterisation of the water/o-nitrophenyl octyl ether system in terms of the partition of nonelectrolytes and of ions. *Phys. Chem. Chem. Phys.* 2005, *7,* 2418-2422.

[127] Abraham, M. H.; Acree, W. E.,Jr Comparative analysis of solvation and selectivity in room temperature ionic liquids using the Abraham linear free energy relationship. *Green Chem.* 2006, *8,* 906-915.

[128] Muijselaar, P. G.; Claessens, H. A.; Cramers, C. A. Characterization of pseudostationary phases in micellar electrokinetic chromatography by applying linear solvation energy relationships and retention indexes. *Anal. Chem.* 1997, *69,* 1184-1191.

[129] Khaledi, M. G.; Bumgarner, J. G.; Hadjmohammadi, M. Characterization of mixed micellar pseudostationary phases in electrokinetic chromatography using linear solvation energy relationships. *J. Chromatogr. A* 1998, *802,* 35-47.

[130] Trone, M. D.; Leonard, M. S.; Khaledi, M. G. Congeneric behavior in estimations of octanol-water partition coefficients by micellar electrokinetic chromatography. *Anal. Chem.* 2000, *72,* 1228-1235.

[131] Fuguet, E.; Ràfols, C.; Torres-Lapasió, J. R.; García-Álvarez-Coque, M. C.; Bosch, E.; Rosés, M. Solute-solvent interactions in micellar electrokinetic chromatography. 6. Optimization of the selectivity of lithium dodecyl sulfate-lithium perfluorooctanesulfonate mixed micellar buffers. *Anal. Chem.* 2002, *74,* 4447-4455.

[132] Landers, J. P.; Oda, R. P.; Schuchard, M. D. Separation of boron-complexed diol compounds using high-performance capillary electrophoresis. *Anal. Chem.* 1992, *64,* 2846-2851.

[133] Fuguet, E.; Reta, M.; Gilbert, C.; Rosés, M.; Bosch, E.; Ràfols, C. Critical evaluation of buffering solutions for pKa determination by capillary electrophoresis. *Electrophoresis* 2008, *29,* 2841-2851.

[134] Poole, C. F.; Atapattu, S. N.; Poole, S. K.; Bell, A. K. Determination of solute descriptors by chromatographic methods. *Anal. Chim. Acta* 2009, *652,* 32-53.

[135] Poole, S. K.; Durham, D.; Kibbey, C. Rapid method for estimating the octanol–water partition coefficient (log P_{ow}) by microemulsion electrokinetic chromatography. *J. Chromatogr. B* 2000, *745,* 117-126.

[136] Poole, S. K.; Poole, C. F. Model for the sorption of organic compounds by soil from water. *Anal. Commun.* 1996, *33,* 417-419.

[137] Burns, S. T.; Agbodjan, A. A.; Khaledi, M. G. Characterization of solvation properties of lipid bilayer membranes in liposome electrokinetic chromatography. *J. Chromatogr. A* 2002, *973,* 167-176.

[138] Abraham, M. H.; Nasezadeh, A.; Acree, W. E. Correlation and prediction of partition coefficients from the gas phase and from water to alkan-1-ols. *Ind. Eng. Chem. Res.* 2008, *47,* 3990-3995.

[139] Abraham, M. H.; Treiner, C.; Rosés, M.; Ràfols, C.; Ishihama, Y. Linear free energy relationship analysis of microemulsion electrokinetic chromatographic determination of lipophilicity. *J. Chromatogr. A* 1996, *752,* 243-249.

[140] Akbay, C.; Shamsi, S. A. Polymeric sulfated surfactants with varied hydrocarbon tail: II. Chemical selectivity in micellar electrokinetic chromatography using linear solvation energy relationships study. *Electrophoresis* 2004, *25,* 635-644.

[141] Fujimoto, C. Application of linear solvation energy relationships to polymeric pseudostationary phases in micellar electrokinetic chromatography. *Electrophoresis* 2001, *22,* 1322-1329.

[142] Abraham, M. H.; Chadha, H. S.; Dixon, J. P.; Ràfols, C.; Treiner, C. Hydrogen bonding. Part 40. Factors that influence the distribution of solutes between water and sodium dodecylsulfate micelles. *J. Chem. Soc. Perkin Trans. 2,* 1995, 887-894.

[143] Barratt, M. D. Quantitative structure-activity relationships for skin permeability. *Toxicol. In Vitro* 1995, *9,* 27-37.

[144] Potts, R. O.; Guy, R. H. Predicting skin permeability. *Pharm. Res.* 1992, *9,* 663-669.

[145] Potts, R. O.; Guy, R. H. A predictive algorithm for skin permeability: the effects of molecular size and hydrogen bond activity. *Pharm. Res.* 1995, *12,* 1628-1633.

[146] Cronin, M. T. D. The prediction of skin permeability using quantitative structure-activity relationship methods. Dermal absorption models in toxicology and pharmacology. Taylor and Francis CRC Press, 2006.

[147] Fitzpatrick, D.; Corish, J.; Hayes, B. Modelling skin permeability in risk assessment–the future. *Chemosphere* 2004, *55,* 1309-1314.

[148] Geinoz, S.; Guy, R. H.; Testa, B.; Carrupt, P. Quantitative structure-permeation relationships (QSPeRs) to predict skin permeation: a critical evaluation. *Pharm. Res.* 2004, *21,* 83-92.

[149] Patel, H.; Berge, W. t.; Cronin, M. T. D. Quantitative structure–activity relationships (QSARs) for the prediction of skin permeation of exogenous chemicals. *Chemosphere* 2002, *48,* 603-613.

[150] Kahn, I.; Fara, D.; Karelson, M.; Maran, U.; Andersson, P. L. QSPR treatment of the soil sorption coefficients of organic pollutants. *J. Chem. Inf. Comput. Sci.* 2005, *45,* 94-105.

[151] Neely, B. J.; Madihally, S. V.; Robinson, R. L., Jr.; Gasem, K. A. M. Nonlinear quantitative structure-property relationship modeling of skin permeation coefficient. *J. Pharm. Sci.* 2009, *98,* 4069-4084.

[152] Neumann, D.; Kohlbacher, O.; Merkwirth, C.; Lengauer, T. A fully computational model for predicting percutaneous drug absorption. *J. Chem. Inf. Model.* 2006, *46,* 424-429.

SUPPORTING INFORMATION

The supporting information is based on three tables: Tables SI-1, SI-2 and SI-3.

Tables SI-1 and SI-2 show, respectively, the compilation of biological and physicochemical partitioning systems characterized by the solvation parameter model. Both tables include the characterizations of the systems and their statistics, as well as the normalized coefficients that have been calculated for each system. They also include the corresponding references from literature and a list of abbreviations.

Table SI-3 is attached on the CD-ROM, being a matrix of dimensions 92x28. It details the *d* distance calculated between each pair of biological and physicochemical systems that have been compiled.

Table SI-1. Compilation of biopartitioning systems characterized by the solvation parameter model: system constants, statistics and calculated normalized coefficients.

nº	Biopartitioning system	SP property	Coefficients						Statistics				Normalized coefficients					Ref.
			c	*e*	*s*	*a*	*b*	*v*	*n*	*r*	SD	*F*	e_u	s_u	a_u	b_u	v_u	
1	Blood-brain distribution	log BB	0.044	0.511	-0.886	-0.724	-0.666	0.861	148	0.843	0.367	71	0.308	-0.534	-0.436	-0.401	0.519	[1]
2	Blood-brain permeation	log PS	-0.639	0.312	-1.009	-1.895	-1.636	1.709	30	0.933	0.520	32	0.097	-0.314	-0.590	-0.510	0.532	[2]
3	Intestinal absorption	log HIA	0.544	-0.025	0.141	-0.409	-0.514	0.204	127	0.894	0.290	94	-0.036	0.201	-0.582	-0.732	0.290	[3]
4	Skin partition	log K_{sc}	0.341	0.341	-0.206	-0.024	-2.178	1.850	45	0.962	0.216	97	0.118	-0.071	-0.008	-0.755	0.641	[4]
5	Skin permeation	log K_p	-5.426	-0.106	-0.473	-0.473	-3.000	2.296	119	0.912	0.461	112	-0.028	-0.123	-0.123	-0.782	0.598	[4]
6	Tadpole narcosis	log ($1/C_{nar}$)	0.582	0.770	-0.696	0.243	-2.592	3.343	114	0.954	0.337	217	0.177	-0.160	0.056	-0.594	0.766	[5]
	Nonspecific aquatic toxicity towards:																	
7	fish *fathead minnow* (*Pimephales promelas*)	-log LC_{50}	0.996	0.418	-0.182	0.417	-3.574	3.377	196	0.976	0.276	779.5	0.084	-0.037	0.084	-0.721	0.681	[6]
8	fish *guppy* (*Poecilia reticulata*)	-log LC_{50}	0.811	0.782	-0.230	0.341	-3.050	3.250	148	0.973	0.280	493.1	0.172	-0.051	0.075	-0.671	0.715	[6]
9	fish *bluegill* (*Lepomis macrochirus*)	-log LC_{50}	0.903	0.583	-0.127	1.238	-3.918	3.306	66	0.984	0.272	360	0.110	-0.024	0.233	-0.738	0.623	[6]
10	fish *golden orfe* (*Leuciscus idus melanotus*)	-log LC_{50}	-0.137	0.931	0.379	0.951	-2.392	3.244	49	0.967	0.269	127	0.218	0.089	0.223	-0.561	0.761	[6]
11	fish *goldfish* (*Carassius auratus*)	-log LC_{50}	0.922	-0.653	1.872	-0.329	-4.516	3.078	51	0.983	0.277	254	-0.112	0.321	-0.057	-0.776	0.529	[6]
12	fish *medaka high-eyes* (*Oryzias latipes*), 48 h	-log LC_{50}	0.834	1.047	-0.380	0.806	-2.182	2.667	50	0.969	0.292	133	0.282	-0.102	0.217	-0.588	0.719	[6]
13	fish *medaka high-eyes* (*Oryzias latipes*), 96 h	-log LC_{50}	-0.176	1.046	0.272	0.931	-2.178	3.155	44	0.980	0.277	182	0.256	0.066	0.228	-0.532	0.771	[6]
14	water flea *Daphnia magna,* 24 h	-log LC_{50}	0.915	0.354	0.173	0.420	-3.935	3.521	107	0.976	0.274	410	0.067	0.033	0.079	-0.741	0.663	[7]
15	water flea *Daphnia magna,* 48 h	-log LC_{50}	0.841	0.528	-0.025	0.219	-3.703	3.591	97	0.982	0.289	475	0.102	-0.005	0.042	-0.714	0.692	[7]
16	water flea *Ceriodaphnia dubia*	-log LC_{50}	2.234	0.373	-0.040	-0.437	-3.276	2.763	44	0.967	0.256	111	0.086	-0.009	-0.101	-0.758	0.639	[7]
17	water flea *Daphnia pulex*	-log LC_{50}	0.502	0.396	0.309	0.542	-3.457	3.527	45	0.981	0.311	233	0.079	0.062	0.109	-0.692	0.706	[7]
18	protozoa *Tetrahymena pyriformis*	-log IGC_{50}	0.616	0.413	-0.048	0.348	-2.707	2.944	192	0.982	0.205	1002	0.102	-0.012	0.086	-0.671	0.729	[7]
19	protozoa *Spirostomum ambiguum*	-log LC_{50}	0.148	0.111	0.288	0.687	-3.301	3.140	60	0.979	0.246	252	0.024	0.062	0.149	-0.715	0.680	[7]

(Table SI-1. Continued.)

nº	Biopartitioning system	SP property	Coefficients						Statistics				Normalized coefficients					Ref.
			c	*e*	*s*	*a*	*b*	*v*	*n*	*r*	SD	*F*	e_u	s_u	a_u	b_u	v_u	
20	protozoa *Entosiphon sulcantum*	-log IGC	0.489	0.894	0.355	1.108	-2.504	2.852	51	0.959	0.333	103	0.220	0.087	0.272	-0.615	0.701	[8]
21	protozoa *Uronema parduczi*	-log IGC	2.706	1.426	0.433	0.938	-1.025	2.599	59	0.961	0.326	127	0.432	0.131	0.284	-0.310	0.787	[8]
22	protozoa *Chilomonas paramecium*	-log IGC	0.440	1.129	0.160	0.442	-1.826	2.446	55	0.942	0.351	77	0.343	0.049	0.134	-0.555	0.744	[8]
23	bacterium *Pseudomonas putida*	-log IGC	0.752	0.955	0.092	-0.081	-2.088	2.947	87	0.963	0.269	209	0.255	0.025	-0.022	-0.559	0.788	[7]
	Bacterial growth inhibition towards:																	
24	bacterium *Porphyromonas gingivalis*	-log MIC	-3.320	1.111	-0.605	0.727	-1.904	2.423	126	0.952	0.313	232	0.326	-0.177	0.213	-0.558	0.711	[9]
25	bacterium *Selenomonas artemidis*	-log MIC	-3.008	0.982	-0.496	0.972	-2.643	2.312	116	0.933	0.268	149	0.258	-0.130	0.255	-0.694	0.607	[9]
26	bacterium *Streptococcus sobrinus*	-log MIC	-3.465	0.855	-0.465	0.735	-1.671	2.330	112	0.950	0.309	196	0.274	-0.149	0.236	-0.536	0.748	[9]
27	Alga cell permeation	$\log K_{cel}$	-2.235	0.000	-0.867	-3.143	-1.664	0.731	100	0.941	0.437	183	0.000	-0.232	-0.842	-0.446	0.196	[10]
28	Soil-water sorption	$\log K_{OC}$	0.21	0.74	0.00	-0.31	-2.27	2.09	129	0.977	0.25	655	0.232	0.000	-0.097	-0.712	0.656	[11]

Abbreviations:

BB: blood-brain distribution; **C_{nar}**: tadpole narcosis; **HIA**: human intestinal absorption; **IGC**: inhibitory growth concentration; **IGC_{50}**: median inhibitory growth concentration (for 50% of the population); **K_{cel}**: cell permeation; **K_{OC}**: soil water-sorption; **K_p**: skin permeation; **K_{sc}**: skin partition; **LC_{50}:** median mortality lethal concentration (for 50% of the population); **MIC**: minimum inhibitory concentration towards bacterial growth; **PS**: permeation across the blood-brain barrier

References:

[1] Platts, J. A.; Abraham, M. H.; Zhao, Y. H.; Hersey, A.; Ijaz, L.; Butina, D. Correlation and prediction of a large blood-brain distribution data set-an LFER study. *Eur. J. Med. Chem.* 2001, *36,* 719-730.

[2] Abraham, M. H. The factors that influence permeation across the blood-brain barrier. *Eur. J. Med. Chem.* 2004, *39,* 235-240.

[3] Abraham, M. H.; Zhao, Y. H.; Le, J.; Hersey, A.; Luscombe, C. N.; Reynolds, D. P.; Beck, G.; Sherborne, B.; Cooper, I. On the mechanism of human intestinal absorption. *Eur. J. Med. Chem.* 2002, *37*, 595-605.

[4] Abraham, M. H.; Martins, F. Human skin permeation and partition: General linear free-energy relationship analyses. *J. Pharm. Sci.* 2004, *93*, 1508-1523.

[5] Abraham, M. H.; Ràfols, C. Factors that influence tadpole narcosis. An LFER analysis. *J. Chem. Soc. Perkin Trans. 2* 1995, 1843-1851.

[6] Hoover, K. R.; Acree, W. E., Jr.; Abraham, M. H. Chemical toxicity correlations for several fish species based on the Abraham solvation parameter model. *Chem. Res. Toxicol.* 2005, *18*, 1497-1505.

[7] Hoover, K. R.; Flanagan, K. B.; Acree, W. E., Jr.; Abraham, M. H. Chemical toxicity correlations for several protozoas, bacteria, and water fleas based on the Abraham solvation parameter model. *J. Environ. Eng. Sci.* 2007, *6*, 165-174.

[8] Bowen, K. R.; Flanagan, K. B.; Acree Jr., W. E.; Abraham, M. H. Correlating toxicities of organic compounds to select protozoa using the Abraham model. *Sci. Total Environ.* 2006, *369*, 109-118.

[9] Mintz, C.; Acree, W. E., Jr.; Abraham, M. H. Correlation of minimum inhibitory concentrations toward oral bacterial growth based on the Abraham model. *QSAR Comb. Sci.* 2006, *25*, 912-920.

[10] Platts, J. A.; Abraham, M. H.; Hersey, A.; Butina, D. Estimation of molecular linear free energy relationship descriptors. 4. Correlation and prediction of cell permeation. *Pharm. Res.* 2000, *17*, 1013-1018.

[11] Poole, S. K.; Poole, C. F. Chromatographic models for the sorption of neutral organic compounds by soil from water and air. *J. Chromatogr. A* 1999, *845*, 381-400.

Table SI-2. Compilation of physicochemical partitioning systems characterized by the solvation parameter model: system constants, statistics and calculated normalized coefficients. The systems characterized throughout the thesis are also included.

nº	Physicochemical partitioning system	SP property	Coefficients						Statistics				Normalized coefficients					Ref.
			c	*e*	*s*	*a*	*b*	*v*	*n*	*r*	SD	*F*	e_u	s_u	a_u	b_u	v_u	
	Solvent-water partition:																	
1	Octanol	log $P_{o/w}$	0.088	0.562	-1.054	0.034	-3.460	3.814	613	0.997	0.116	23162	0.106	-0.199	0.006	-0.655	0.721	[1]
2	Isobutanol	log P	*	0.514	-0.693	0.020	-2.258	2.776	*	*	*	*	0.140	-0.188	0.005	-0.613	0.754	[2]
3	Pentanol	log P	0.185	0.367	-0.732	0.105	-3.100	3.395	106	0.988	0.180	823	0.079	-0.157	0.022	-0.664	0.727	[3]
4	Oleyl alcohol	log P	-0.359	-0.270	-0.528	-0.035	-4.042	4.204	*	*	*	*	-0.046	-0.090	-0.006	-0.690	0.717	[2]
5	Diethyl ether	log P	0.248	0.561	-1.016	-0.226	-4.553	4.075	239	0.969	0.352	722	0.090	-0.163	-0.036	-0.732	0.655	[4]
6	Dibutyl ether	log P	0.252	0.677	-1.506	-0.807	-5.249	4.815	87	0.976	0.310	329	0.092	-0.205	-0.110	-0.714	0.655	[5]
7	NPOE	log P	0.121	0.600	-0.459	-2.246	-3.879	3.574	55	0.967	0.308	142	0.104	-0.079	-0.388	-0.671	0.618	[6]
8	Ethyl acetate	log P	*	1.157	-1.397	-0.054	-4.261	3.726	*	*	*	*	0.195	-0.235	-0.009	-0.717	0.627	[7]
9	PGDP	log P	*	0.501	-0.828	-1.022	-4.640	4.033	*	*	*	*	0.079	-0.131	-0.162	-0.736	0.639	[7]
10	Olive oil	log P	*	0.574	-0.798	-1.422	-4.984	4.210	*	*	*	*	0.085	-0.118	-0.211	-0.738	0.624	[7]
11	Benzene	log P	0.142	0.464	-0.588	-3.099	-4.625	4.491	*	*	*	*	0.065	-0.082	-0.431	-0.643	0.624	[8]
12	Nitrobenzene	log P	*	0.576	0.003	-2.356	-4.420	4.263	*	*	*	*	0.087	0.000	-0.357	-0.669	0.646	[9]
13	Hexane	log P	0.361	0.579	-1.723	-3.599	-4.764	4.344	*	*	*	*	0.076	-0.227	-0.473	-0.627	0.571	[8]
14	Hexadecane	log P	0.087	0.667	-1.617	-3.587	-4.869	4.433	370	0.998	0.124	20236	0.087	-0.210	-0.466	-0.632	0.576	[1]
15	Cyclohexane	log P	0.159	0.784	-1.678	-3.740	-4.929	4.577	*	*	*	*	0.099	-0.212	-0.472	-0.623	0.578	[8]
16	Carbon disulfide	log P	0.047	0.686	-0.943	-3.603	-5.818	4.921	*	*	*	*	0.081	-0.111	-0.423	-0.684	0.578	[8]

(Table SI-2. Continued.)

nº	Physicochemical partitioning system	SP property	Coefficients						Statistics				Normalized coefficients					Ref.
			c	*e*	*s*	*a*	*b*	*v*	*n*	*r*	SD	*F*	e_u	s_u	a_u	b_u	v_u	
	Liquid chromatography:																	
17	RP18 HPLC	$\log k_{HPLC}$	-0.172	0.233	-0.460	-0.260	-1.574	1.582	55	0.996	0.056	1121	0.101	-0.200	-0.113	-0.683	0.686	[10]
18	MSC18 HPLC	$\log k_{HPLC}$	-0.152	0.151	-0.463	-0.424	-1.719	1.759	55	0.994	0.072	764	0.059	-0.182	-0.167	-0.676	0.692	[10]
19	IAM HPLC	$\log k_{HPLC}$	-0.710	0.292	-0.344	0.141	-1.193	1.161	51	0.991	0.047	480	0.169	-0.199	0.081	-0.689	0.671	[10]
20	SDS C18 MLC	$\log k_{MLC}$	0.572	0.157	-0.447	-0.387	-0.975	0.853	26	0.993	0.046	283	0.110	-0.312	-0.270	-0.681	0.595	[11]
21	Brij 35 C18 MLC	$\log k_{MLC}$	0.645	0.016	-0.344	-0.281	-1.026	0.949	26	0.990	0.051	191	0.011	-0.235	-0.192	-0.700	0.647	[11]
22	SDS C18, 0.8% heptane, MELC	$\log k_{MELC}$	0.665	0.117	-0.454	-0.495	-0.877	0.751	26	0.992	0.050	243	0.087	-0.339	-0.369	-0.654	0.560	[11]
23	SDS C18, 1.6% heptane, MELC	$\log k_{MELC}$	0.727	0.101	-0.645	-0.922	-0.951	0.901	26	0.990	0.078	205	0.058	-0.373	-0.533	-0.550	0.521	[11]
24	Brij 35 C18 MELC	$\log k_{MELC}$	0.698	-0.003	-0.412	-0.415	-0.912	0.887	26	0.986	0.064	141	-0.002	-0.294	-0.296	-0.651	0.633	[11]
25	Brij 35-SDS C18 MELC	$\log k_{MELC}$	0.706	0.027	-0.399	-0.474	-0.842	0.773	26	0.990	0.054	189	0.021	-0.307	-0.364	-0.647	0.594	[11]
	Electrokinetic chromatography:																	
26	SDS MEKC	$\log k_{MEKC}$	-1.680	0.558	-0.596	-0.266	-1.674	2.717	63	0.991	0.120	618	0.169	-0.180	-0.080	-0.507	0.822	[12,13]
27	SDecS MEKC	$\log k_{MEKC}$	-2.43	0.32	-0.24	0.00	-1.60	2.69	36	0.989	0.07	*	0.101	-0.076	0.000	-0.507	0.853	[14]
28	SOS MEKC	$\log k_{MEKC}$	-1.97	0.45	-0.31	-0.12	-1.87	2.85	23	0.994	0.06	*	0.130	-0.090	-0.035	-0.541	0.825	[14]
29	SDSu MEKC	$\log k_{MEKC}$	-1.92	0.33	-0.42	-0.02	-1.78	2.84	36	0.992	*	*	0.097	-0.124	-0.006	-0.524	0.837	[15]
30	SDCar MEKC	$\log k_{MEKC}$	-1.95	0.15	-0.39	0.23	-1.77	2.96	36	0.989	*	*	0.043	-0.112	0.066	-0.508	0.850	[15]
31	SDP MEKC	$\log k_{MEKC}$	-1.92	0.24	-0.55	0.15	-2.00	3.01	36	0.992	*	*	0.065	-0.150	0.041	-0.545	0.821	[15]
32	SDCV MEKC	$\log k_{MEKC}$	-1.65	0.42	-0.61	0.11	-2.38	2.94	36	0.992	*	*	0.109	-0.158	0.029	-0.617	0.762	[15]
33	SLSA MEKC	$\log k_{MEKC}$	-1.82	0.41	-0.37	0.10	-2.39	2.96	36	0.992	*	*	0.107	-0.096	0.026	-0.621	0.770	[15]
34	THADS MEKC	$\log k_{MEKC}$	-1.43	0.57	-0.66	-0.33	-1.56	2.56	24	0.997	0.06	532	0.182	-0.210	-0.105	-0.497	0.815	[16]
35	LDS MEKC	$\log k_{MEKC}$	-1.575	0.586	-0.595	-0.317	-1.565	2.609	63	0.989	0.128	526	0.185	-0.188	-0.100	-0.494	0.823	[12,13]

(Table SI-2. Continued.)

nº	Physicochemical partitioning system	SP property	Coefficients						Statistics				Normalized coefficients					Ref.
			c	*e*	*s*	*a*	*b*	*v*	*n*	*r*	SD	*F*	e_u	s_u	a_u	b_u	v_u	
36	LMT MEKC	$\log k_{MEKC}$	-1.90	0.51	-0.35	0.39	-2.37	2.88	36	0.99	0.06	*	0.134	-0.092	0.103	-0.624	0.758	[17]
37	ALE MEKC	$\log k_{MEKC}$	-1.89	0.44	-0.37	0.49	-2.41	2.92	36	0.99	0.06	*	0.114	-0.096	0.127	-0.624	0.756	[17]
38	SLN MEKC	$\log k_{MEKC}$	-1.99	0.44	-0.39	0.45	-2.32	2.92	36	0.99	0.06	*	0.116	-0.103	0.118	-0.610	0.768	[17]
39	$Mg(DS)_2$ MEKC	$\log k_{MEKC}$	-1.55	0.27	-0.42	-0.27	-1.88	3.02	35	0.990	*	*	0.075	-0.117	-0.075	-0.522	0.838	[18]
40	$Cu(DS)_2$ MEKC	$\log k_{MEKC}$	-1.510	0.350	-0.510	-0.260	-1.920	3.050	35	0.994	*	*	0.095	-0.139	-0.071	-0.524	0.832	[18]
41	SPN MEKC	$\log k_{MEKC}$	-1.72	0.42	-0.45	0.48	-2.58	3.11	36	0.99	0.07	*	0.102	-0.109	0.117	-0.627	0.756	[19]
42	AOT MEKC	$\log k_{MEKC}$	-1.82	0.34	-0.43	0.02	-3.02	3.09	43	0.986	*	*	0.078	-0.099	0.005	-0.693	0.709	[20]
43	LPFOS MEKC	$\log k_{MEKC}$	-1.410	-0.113	-0.243	-0.876	-0.455	1.966	62	0.970	0.190	180	-0.051	-0.110	-0.395	-0.205	0.887	[12,13]
44	SC MEKC	$\log k_{MEKC}$	-1.408	0.691	-0.693	0.117	-1.937	2.274	62	0.962	0.191	139	0.220	-0.220	0.037	-0.616	0.723	[12,13]
45	SDC MEKC	$\log k_{MEKC}$	-1.833	0.926	-0.867	0.070	-1.785	2.422	58	0.949	0.260	94	0.284	-0.265	0.021	-0.547	0.742	[12,13]
46	KDC MEKC	$\log k_{MEKC}$	-1.97	-0.53	-0.92	0.00	-2.50	3.10	18	0.991	0.10	139	-0.129	-0.223	0.000	-0.607	0.752	[16,21]
47	STC MEKC	$\log k_{MEKC}$	-2.10	0.60	-0.34	0.00	-2.06	2.43	40	0.989	0.09	377	0.184	-0.104	0.000	-0.632	0.746	[22]
48	STDC MEKC	$\log k_{MEKC}$	-1.99	0.67	-0.45	0.00	-2.17	2.62	40	0.991	0.09	430	0.192	-0.129	0.000	-0.621	0.749	[22]
49	KGDC MEKC	$\log k_{MEKC}$	-1.83	-0.60	-1.03	0.00	-1.99	2.78	16	0.970	0.12	89	-0.166	-0.284	0.000	-0.550	0.768	[16,21]
50	HTAB MEKC	$\log k_{MEKC}$	-1.830	1.112	-0.755	0.824	-2.437	2.710	49	0.994	0.081	690	0.280	-0.190	0.208	-0.614	0.682	[12,13]
51	TTAB MEKC	$\log k_{MEKC}$	-1.851	0.902	-0.617	0.766	-2.410	2.634	53	0.994	0.089	746	0.237	-0.162	0.201	-0.632	0.691	[12,13]
52	DTAC MEKC	$\log k_{MEKC}$	*	0.749	-0.430	0.871	-2.667	2.823	49	0.988	*	*	0.184	-0.106	0.214	-0.655	0.693	[23]
53	DHAB MEKC	$\log k_{MEKC}$	-2.96	1.46	-0.59	1.34	-4.38	4.01	41	0.96	*	*	0.232	-0.094	0.213	-0.697	0.638	[24]
54	50 mM SDS-5 mM Brij 35 MEKC	$\log k_{MEKC}$	-1.559	0.518	-0.444	0.311	-2.815	2.905	27	0.992	0.082	252	0.126	-0.108	0.076	-0.684	0.706	[25]
55	50 mM SDS-10 mM Brij 35 MEKC	$\log k_{MEKC}$	-1.412	0.622	-0.516	0.308	-2.699	2.829	27	0.987	0.105	154	0.155	-0.129	0.077	-0.674	0.706	[25]
56	50 mM SDS-15 mM Brij 35 MEKC	$\log k_{MEKC}$	-1.285	0.667	-0.515	0.335	-2.723	2.767	27	0.985	0.109	140	0.167	-0.129	0.084	-0.683	0.694	[25]

(Table SI-2. Continued.)

nº	Physicochemical partitioning system	SP property	Coefficients						Statistics				Normalized coefficients					Ref.
			c	*e*	*s*	*a*	*b*	*v*	*n*	*r*	SD	*F*	e_u	s_u	a_u	b_u	v_u	
57	10 mM LPFOS-30 mM LDS MEKC	log k_{MEKC}	-1.79	0.27	-0.41	-0.37	-1.20	2.74	40	0.991	0.081	446	0.088	-0.134	-0.121	-0.393	0.897	[26]
58	20 mM LPFOS-20 mM LDS MEKC	log k_{MEKC}	-1.85	0.16	-0.31	-0.58	-0.85	2.64	40	0.993	0.075	471	0.056	-0.109	-0.203	-0.298	0.925	[26]
59	30 mM LPFOS-10 mM LDS MEKC	log k_{MEKC}	-1.90	-0.02	-0.16	-0.76	-0.45	2.45	40	0.992	0.078	371	-0.008	-0.061	-0.291	-0.172	0.939	[26]
60	SC MEKC	log k_{MEKC}	-1.501	0.910	-0.776	0.109	-1.961	2.255	66	0.958	0.234	133	0.283	-0.241	0.034	-0.609	0.700	Article I
61	60 mM SC-10 mM SDC MEKC	log k_{MEKC}	-1.578	0.897	-0.768	0.125	-1.903	2.241	66	0.961	0.221	145	0.283	-0.242	0.039	-0.600	0.707	Article I
62	40 mM SC-20 mM SDC MEKC	log k_{MEKC}	-1.658	0.858	-0.795	0.145	-1.847	2.300	67	0.961	0.221	147	0.270	-0.250	0.046	-0.581	0.724	Article I
63	20 mM SC-30 mM SDC MEKC	log k_{MEKC}	-1.753	0.865	-0.809	0.117	-1.827	2.354	66	0.955	0.239	125	0.270	-0.252	0.036	-0.569	0.734	Article I
64	SDC MEKC	log k_{MEKC}	-1.786	0.813	-0.803	0.061	-1.771	2.400	64	0.947	0.255	102	0.254	-0.251	0.019	-0.554	0.751	Article I
65	STC MEKC	log k_{MEKC}	-1.838	0.613	-0.513	0.319	-2.396	2.390	56	0.984	0.112	305	0.175	-0.147	0.091	-0.686	0.684	Article III
66	SDS MEEKC	log k_{MEEKC}	-1.133	0.279	-0.692	-0.060	-2.805	3.048	53	0.994	0.090	791	0.066	-0.164	-0.014	-0.667	0.724	[27]
67	DPPG:DPPC:Chol LEKC	log k_{LEKC}	-2.30	0.54	-0.55	0.32	-3.12	3.01	27	0.990	0.08	*	0.122	-0.125	0.072	-0.707	0.682	[28]
68	DPPG:DPPC LEKC	log k_{LEKC}	-2.21	0.45	-0.44	0.71	-3.23	3.13	27	0.995	0.07	*	0.098	-0.096	0.154	-0.703	0.681	[28]
69	DHP LEKC	log k_{LEKC}	-2.68	0.42	-0.65	0.47	-3.27	3.59	41	0.99	*	*	0.085	-0.132	0.095	-0.662	0.727	[29]
70	DHP+Chol LEKC	log k_{LEKC}	-2.28	0.53	-0.77	0.43	-3.29	3.35	41	0.99	*	*	0.110	-0.160	0.089	-0.684	0.697	[29]
71	PAAU LEKC	log $k_{poly\text{-}EKC}$	-1.86	0.26	-0.16	-0.27	-1.05	2.11	18	0.997	0.051	403	0.109	-0.067	-0.113	-0.439	0.882	[30]
72	PSUA LEKC	log $k_{poly\text{-}EKC}$	-2.28	0.18	0.45	-0.15	-1.18	1.64	18	0.993	0.063	174	0.086	0.216	-0.072	-0.566	0.787	[30]
73	poly-SocS poly-EKC	log $k_{poly\text{-}EKC}$	-2.681	0.218	0.260	-0.144	-1.145	2.253	35	0.991	0.049	*	0.085	0.102	-0.056	-0.448	0.882	[31]
74	poly-SnoS poly-EKC	log $k_{poly\text{-}EKC}$	-3.023	0.482	0.078	-0.152	-1.495	2.907	35	0.988	0.078	*	0.146	0.024	-0.046	-0.452	0.879	[31]
75	poly-SDeS poly-EKC	log $k_{poly\text{-}EKC}$	-2.929	0.522	-0.049	-0.141	-1.637	2.946	35	0.991	0.067	*	0.153	-0.012	-0.041	-0.480	0.863	[31]
76	poly-SUS poly-EKC	log $k_{poly\text{-}EKC}$	-3.013	0.685	-0.190	-0.103	-1.772	3.182	35	0.990	0.080	*	0.185	-0.051	-0.028	-0.477	0.857	[31]
77	AGENT poly-EKC	log $k_{poly\text{-}EKC}$	-2.81	0.76	-0.07	0.45	-1.93	2.07	40	0.963	*	*	0.256	-0.024	0.151	-0.650	0.699	[32]

(Table SI-2. Continued.)

nº	Physicochemical partitioning system	SP property	Coefficients						Statistics				Normalized coefficients					Ref.
			c	*e*	*s*	*a*	*b*	*v*	*n*	*r*	SD	*F*	e_u	s_u	a_u	b_u	v_u	
78	OAGENT poly-EKC	$\log k_{poly\text{-}EKC}$	-1.65	0.71	-1.08	0.11	-2.29	2.06	40	0.962	*	*	0.212	-0.323	0.033	-0.685	0.616	[32]
79	DAGENT poly-EKC	$\log k_{poly\text{-}EKC}$	-1.98	0.59	-0.78	0.23	-2.42	2.39	40	0.979	*	*	0.166	-0.220	0.064	-0.683	0.674	[32]
80	SAGENT poly-EKC	$\log k_{poly\text{-}EKC}$	-1.75	0.63	-1.14	0.33	-2.64	2.51	40	0.941	*	*	0.161	-0.295	0.084	-0.680	0.646	[32]
81	AGESS poly-EKC	$\log k_{poly\text{-}EKC}$	-2.40	0.46	-0.43	0.27	-2.46	2.72	38	0.991	*	*	0.123	-0.115	0.072	-0.659	0.729	[33]
82	Elvacite 2669 poly-EKC	$\log k_{poly\text{-}EKC}$	-1.67	0.36	-0.19	0.07	-1.88	2.05	36	0.973	*	*	0.128	-0.068	0.025	-0.669	0.729	[34]
83	pOMAt-21-Na poly-EKC	$\log k_{poly\text{-}EKC}$	-2.662	0.470	-0.595	-0.409	-3.747	3.559	20	0.996	0.069	396.4	0.090	-0.114	-0.078	-0.715	0.679	[35]
84	pLMAt-15-Na poly-EKC	$\log k_{poly\text{-}EKC}$	-2.842	0.435	-0.672	-0.274	-3.704	3.654	20	0.995	0.078	303.1	0.084	-0.127	-0.051	-0.703	0.693	[35]
85	pSMAt-16-Na poly-EKC	$\log k_{poly\text{-}EKC}$	-2.729	0.653	-0.846	-0.495	-3.831	3.782	20	0.995	0.092	273.8	0.118	-0.154	-0.091	-0.695	0.686	[35]
86	pAt-13-Na poly-EKC	$\log k_{poly\text{-}EKC}$	-2.962	0.390	-0.395	-0.024	-3.524	3.582	20	0.998	0.112	117	0.077	-0.079	-0.004	-0.697	0.709	[35]
87	pLMAm-19-Na poly-EKC	$\log k_{poly\text{-}EKC}$	-2.690	0.374	-0.322	0.254	-2.453	2.880	20	0.995	0.055	266	0.097	-0.084	0.065	-0.641	0.754	[35]
88	pSAm-28-Na poly-EKC	$\log k_{poly\text{-}EKC}$	-2.570	0.421	-0.534	-0.187	-3.053	3.385	20	0.996	0.056	387	0.091	-0.115	-0.041	-0.661	0.735	[35]
89	pDHCHAt-2-Na poly-EKC	$\log k_{poly\text{-}EKC}$	-3.110	0.606	-0.598	-0.040	-2.579	2.906	20	0.988	0.096	120	0.153	-0.151	-0.010	-0.648	0.731	[36]
90	pDHCHAt-33-TEA poly-EKC	$\log k_{poly\text{-}EKC}$	-2.678	0.646	-0.462	0.241	-3.205	3.402	20	0.992	0.087	181	0.137	-0.097	0.051	-0.676	0.716	[36]
91	pLAt-9.2-TEA poly-EKC	$\log k_{poly\text{-}EKC}$	-3.146	0.496	-0.399	0.228	-3.188	3.153	20	0.993	0.074	210	0.110	-0.088	0.051	-0.703	0.695	[36]
92	ptOAm-49-Na poly-EKC	$\log k_{poly\text{-}EKC}$	-2.864	0.333	-0.438	0.434	-3.224	3.357	20	0.994	0.071	225	0.070	-0.093	0.091	-0.684	0.714	[36]

* Data not available in literature

Abbreviations:

AGENT: allyl glycidyl ether *N*-methyltaurine siloxane; **AGESS:** dodecane allyl glycidyl ether sulfite-modified siloxane; **ALE:** sodium *N*-lauroyl-*N*-methyl-β-alaninate; **AMPS:** 2-acrylamide-2-methyl-1-propane sulfonic acid; **AOT:** bis(2-ethylhexyl)sodium sulfosuccinate; **Brij 35:** polyoxyethylene(23) dodecyl ether**; C18**: HPLC column with octadecyl bonded silica stationary phase; **Chol:** cholesterol; **$Cu(DS)_2$:** copper dodecyl sulfate; **DAGENT:** dodecane allyl glycidyl ether *N*-methyl taurine siloxane; **DHAB:** dihexadecyldimethylammonium bromide; **DHP:** dihexadecylphosphate; **DPPC:** dipalmitoylphosphatidyl choline; **DPPG:** dipalmitoylphosphatidyl glycerol; **DTAC:** dodecyltrimethylammonium chloride; **Elvacite 2669:** poly(methyl methacrylate-ethyl acrylate-methacrylic acid); **HPLC:** high-performance liquid chromatography; **HTAB:** hexadecyltrimethylammonium bromide; **IAM:** immobilized artificial membrane column; **KDC:** potassium deoxycholate; **KGDC:** 3-β-glucopyranosyl-5β-cholan-12α-hydroxy-24-oic acid; **LDS:** lithium dodecyl sulfate; **LEKC:** liposome electrokinetic chromatography; **LMT:** sodium *N*-lauroyl-*N*-methyltaurate; **LPFOS:** lithium perfluorooctane sulfonate; **MEKC:** micellar electrokinetic chromatography; **MEEKC:** microemulsion electrokinetic chromatography; **MELC:** microemulsion liquid chromatography; **$Mg(DS)_2$:** magnesium dodecyl sulfate; **MLC:** micellar liquid chromatography; **MSC18**: XTerra MSC18 HPLC column; **NPOE**: *o*-nitrophenyl octyl ether; **OAGENT:** octane AGENT; **P:** solvent-water partition; **PAAU:** poly(sodium 11-acrylamidoundecanoate); **pAt-13-Na:** poly(AMPS-sodium lauryl acrylate-13); **pDHCHAt-2-Na:** poly(AMPS-sodium dihydrocholesteryl acrylate-2); **pDHCHAt-33-TEA:** poly(AMPS-triethylamine dihydrocholesteryl acrylate-33); **PGDP**: propylene glycol dipelargonate; **pLAt-9.2-TEA:** poly(AMPS-triethylamine lauryl acrylate-9.2); **pLMAt-15-Na:** poly(AMPS-sodium lauryl methacrylate-15); **pLMAm-19-Na:** poly(AMPS-sodium lauryl methacrylamide-19); **$P_{o/w}$:** octanol-water partition; **poly-EKC:** polymeric electrokinetic chromatography; **poly-SDeS:** poly(sodium 9-decenyl sulfate); **poly-SNoS:** poly(sodium 8-nonenyl sulfate); **poly-SOcS:** poly(sodium 7-octenyl sulfate); **poly-SUS:** poly(sodium 10-undecenyl sulfate); **pOMAt-21-Na:** poly(AMPS-sodium octyl methacrylate-21); **pSAm-28-Na:** poly(AMPS-sodium stearyl acrylamide-28); **pSMAt-16-Na:** poly(AMPS-sodium stearyl methacrylate-16); **PSUA:** poly(sodium 10-undecylenate); **ptOAm-49-Na:** poly(AMPS-sodium *tert*-octyl acrylamide-49); **RP18**: XTerra RP18 HPLC column; **SAGENT:** steryl AGENT; **SC:** sodium cholate; **SDC:** sodium deoxycholate; **SDCar:** sodium dodecyl carboxylate; **SDCV:** sodium dodecyl carbonyl valine; **SDecS:** sodium decyl sulfate; **SDP:** sodium dodecyl phosphate; **SDS**: sodium dodecyl sulfate; **SDSu:** sodium dodecyl sulfonate; **SLN:** sodium *N*-lauroylsarcosinate; **SLSA:** sodium dodecyl sulfoacetate; **SOS:** sodium octyl sulfate; **SPN:** sodium *N*-parmitoyl sarcosinate; **STC:** sodium taurocholate; **STDC:** sodium taurodeoxycholate; **THADS:** tris(hydroxymethyl)aminomethane dodecyl sulfate; **TTAB:** tetradecyltrimethylammonium bromide.

References:

[1] Abraham, M. H.; Chadha, H. S.; Whiting, G. S.; Mitchell, R. C. Hydrogen bonding. 32. An analysis of water-octanol and water-alkane partitioning and the Δlog P parameter of seiler. *J. Pharm. Sci.* 1994, *83,* 1085-1100.

[2] Abraham, M. M.; Chadha, H. S.; Dixon, J. P.; Leo, A. J. Hydrogen bonding. 39. The partition of solutes between water and various alcohols. *J. Phys. Org. Chem.* 1994, *7,* 712-716.

[3] Abraham, M. H.; Nasezadeh, A.; Acree, W. E. Correlation and prediction of partition coefficients from the gas phase and from water to alkan-1-ols. *Ind. Eng. Chem. Res.* 2008, *47,* 3990-3995.

[4] Abraham, M. H.; Zissimos, A. M.; Acree, W. E., Jr. Partition of solutes into wet and dry ethers; an LFER analysis. *New J. Chem.* 2003, *27,* 1041-1044.

[5] Abraham, M. H.; Zissimos, A. M.; Acree, W. E., Jr. Partition of solutes from the gas phase and from water to wet and dry di-n-butyl ether: a linear free energy relationship analysis. *Phys. Chem. Chem. Phys.* 2001, *3,* 3732-3736.

[6] Abraham, M. H.; Zhao, Y. H. Characterisation of the water/o-nitrophenyl octyl ether system in terms of the partition of nonelectrolytes and of ions. *Phys. Chem. Chem. Phys.* 2005, *7,* 2418-2422.

[7] Abraham, M. H.; Acree Jr., W. E. Characterisation of the water–isopropyl myristate system. *Int. J. Pharm.* 2005, *294,* 121-128.

[8] Abraham, M. H.; Ibrahim, A.; Zissimos, A. M. Determination of sets of solute descriptors from chromatographic measurements. *J. Chromatogr. A* 2004, *1037,* 29-47.

[9] Abraham, M. H. Hydrogen bonding. 31. Construction of a scale of solute effective or summation hydrogen-bond basicity. *J. Phys. Org. Chem.* 1993, *6,* 660-684.

[10] Lázaro, E.; Ràfols, C.; Abraham, M. H.; Rosés, M. Chromatographic estimation of drug disposition properties by means of immobilized artificial membranes (IAM) and C18 Columns. *J. Med. Chem.* 2006, *49,* 4861-4870.

[11] Liu, J.; Sun, J.; Wang, Y.; Liu, X.; Sun, Y.; Xu, H.; He, Z. Characterization of microemulsion liquid chromatography systems by solvation parameter model and comparison with other physicochemical and biological processes. *J. Chromatogr. A* 2007, *1164,* 129-138.

[12] Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. Solute-solvent interactions in micellar electrokinetic chromatography. III. Characterization of the selectivity of micellar electrokinetic chromatography systems. *J. Chromatogr. A* 2002, *942,* 237-248.

[13] Fuguet, E.; Ràfols, C.; Bosch, E.; Abraham, M. H.; Rosés, M. Solute-solvent interactions in micellar electrokinetic chromatography III. Characterization of the selectivity of micellar electrokinetic chromatography systems. [Erratum to document cited in CA136:272381]. *J. Chromatogr. A* 2009, *1216,* 6877-6879.

[14] Vitha, M. F.; Carr, P. W. A linear solvation energy relationship study of the effects of surfactant chain length on the chemical interactions governing retention and selectivity in micellar electrokinetic capillary chromatography using sodium alkyl sulfate sulfate buffers. *Sep. Sci. Technol.* 1998, *33,* 2075-2100.

[15] Trone, M. D.; Khaledi, M. G. Characterization of chemical selectivity in micellar electrokinetic chromatography. 4. Effect of surfactant headgroup. *Anal. Chem.* 1999, *71,* 1270-1277.

[16] Poole, C. F.; Poole, S. K.; Abraham, M. H. Recommendations for the determination of selectivity in micellar electrokinetic chromatography. *J. Chromatogr. A* 1998, *798,* 207-222.

[17] Trone, M. D.; Khaledi, M. G. Influence of ester and amide-containing surfactant headgroups on selectivity in micellar electrokinetic chromatography. *Electrophoresis* 2000, *21,* 2390-2396.

[18] Trone, M. D.; Mack, J. P.; Goodell, H. P.; Khaledi, M. G. Characterization of chemical selectivity in micellar electrokinetic chromatography: VI. Effects of surfactant counter-ion. *J. Chromatogr. A* 2000, *888,* 229-240.

[19] Trone, M. D.; Khaledi, M. G. Statistical evaluation of linear solvation energy relationship models used to characterize chemical selectivity in micellar electrokinetic chromatography. *J. Chromatogr. A* 2000, *886,* 245-257.

[20] Klotz, W. L.; Schure, M. R.; Foley, J. P. Rapid estimation of octanol–water partition coefficients using synthesized vesicles in electrokinetic chromatography. *J. Chromatogr. A* 2002, *962,* 207-219.

[21] Adlard, M.; Okafo, G.; Meenan, E.; Camilleri, P. Rapid estimation of octanol-water partition coefficients using deoxycholate micelles in capillary electrophoresis. *J. Chem. Soc., Chem. Commun.,* 1995, 2241-2243.

[22] Poole, S. K.; Poole, C. F. Characterization of surfactant selectivity in micellar electrokinetic chromatography. *Analyst* 1997, *122,* 267-274.

[23] Ishihama, Y.; Asakawa, N. Characterization of lipophilicity scales using vectors from solvation energy descriptors. *J. Pharm. Sci.* 1999, *88,* 1305-1312.

[24] Agbodjan, A. A.; Khaledi, M. G. Study of solute partitioning into cationic vesicles of dihexadecyldimethylammonium bromide using electrokinetic chromatography. *J. Chromatogr. A* 2003, *1004,* 145-153.

[25] Rosés, M.; Ràfols, C.; Bosch, E.; Martínez, A. M.; Abraham, M. H. Solute-solvent interactions in micellar electrokinetic chromatography. Characterization of sodium dodecyl sulfate-Brij 35 micellar systems for quantitative structure-activity relationship modelling. *J. Chromatogr. A* 1999, *845,* 217-226.

[26] Fuguet, E.; Ràfols, C.; Bosch, E.; Rosés, M.; Abraham, M. H. Solute-solvent interactions in micellar electrokinetic chromatography: Selectivity of lithium dodecyl sulfate-lithium perfluorooctanesulfonate mixed-micellar buffers. *J. Chromatogr. A* 2001, *907,* 257-265.

[27] Abraham, M. H.; Treiner, C.; Rosés, M.; Ràfols, C.; Ishihama, Y. Linear free energy relationship analysis of microemulsion electrokinetic chromatographic determination of lipophilicity. *J. Chromatogr. A* 1996, *752,* 243-249.

[28] Burns, S. T.; Agbodjan, A. A.; Khaledi, M. G. Characterization of solvation properties of lipid bilayer membranes in liposome electrokinetic chromatography. *J. Chromatogr. A* 2002, *973,* 167-176.

[29] Agbodjan, A. A.; Bui, H.; Khaledi, M. G. Study of solute partitioning in biomembrane-mimetic pseudophases by electrokinetic chromatography: Dihexadecyl phosphate small unilamellar vesicles. *Langmuir* 2001, *17,* 2893-2899.

[30] Fujimoto, C. Application of linear solvation energy relationships to polymeric pseudostationary phases in micellar electrokinetic chromatography. *Electrophoresis* 2001, *22,* 1322-1329.

[31] Akbay, C.; Shamsi, S. A. Polymeric sulfated surfactants with varied hydrocarbon tail: II. Chemical selectivity in micellar electrokinetic chromatography using linear solvation energy relationships study. *Electrophoresis* 2004, *25,* 635-644.

[32] Peterson, D. S.; Palmer, C. P. Novel alkyl-modified anionic siloxanes as pseudostationary phases for electrokinetic chromatography: II. Selectivity studied by linear solvation energy relationships. *Electrophoresis* 2001, *22,* 3562-3566.

[33] Schulte, S.; Palmer, C. P. Alkyl-modified siloxanes as pseudostationary phases for electrokinetic chromatography. *Electrophoresis* 2003, *24,* 978-983.

[34] Trone, M. D.; Leonard, M. S.; Khaledi, M. G. Congeneric behavior in estimations of octanol-water partition coefficients by micellar electrokinetic chromatography. *Anal. Chem.* 2000, *72,* 1228-1235.

[35] Shi, W.; Peterson, D. S.; Palmer, C. P. Effect of pendant chain lengths and backbone functionalities on the chemical selectivity of sulfonated amphiphilic copolymers as pseudo-stationary phases in electrokinetic chromatography. *J. Chromatogr. A* 2001, *924,* 123-135.

[36] Shi, W.; Palmer, C. P. Effect of pendent group structures on the chemical selectivity and performance of sulfonated copolymers as novel pseudophases in electrokinetic chromatography. *Electrophoresis* 2002, *23,* 1285-1295.

Table SI-3. Values of the *d* distance parameter calculated between each pair of biological and physicochemical partitioning systems.