

# Modellierung PBPK-relevanter Verteilungskoeffizienten organischer Stoffe

Von der Fakultät für Chemie und Physik  
der Technischen Universität Bergakademie Freiberg  
genehmigte

**DISSERTATION**

zur Erlangung des akademischen Grades

*Doctor rerum naturalium*  
(*Dr. rer. nat.*),

vorgelegt

von Dipl.-Chem. Stefanie Stöckl

geboren am 03.02.1984 in Parsberg.

Gutachter: Prof. Dr. Gerrit Schüürmann, Leipzig, Freiberg  
Prof. Dr. Andreas Schäffer, Aachen

Tag der Verleihung: 20.12.2013







Diese Arbeit wurde finanziert durch die Europäischen Forschungsprojekte 2-FUN (Full-chain and UNcertainty Approaches for Assessing Health Risks in Future ENvironmental Scenarios), Vertragsnummer 036976) und OSIRIS (Optimised Strategies for Risk Assessment of Industrial Chemicals through Integration of Non-Test and Test Information, Vertragsnummer GOCE-CT-2007-037017)



## Danksagung

In erster Linie möchte ich mich bei meinem Doktorvater Herrn Prof. Gerrit Schüürmann für die Bereitstellung des interessanten Themas bedanken und das große Vertrauen, welches er in mich und meine Arbeitskraft steckte. Für die Erstellung des Gutachtens danke ich ihm ebenfalls.

Herrn Prof. Dr. Andreas Schäffer des Lehrstuhls für Umweltbiologie und Chemodynamik der RWTH Aachen danke ich sehr für die Erstellung des Zweitgutachtens.

Bei Herrn Dr. Ralph Kühne des Departments Ökologische Chemie bedanke ich mich für seine engagierte und kritische Betreuung, die zum Gelingen dieser Arbeit beigetragen haben. Auch Ralf-Uwe Ebert hat mich nie im Stich gelassen, wenn es um fachliche Fragen zu ChemProp, Dampfdruck und weitere Themen ging. Dafür danke ich ihm.

Meiner Vertrauensperson Dr. Torsten Thalheim und seiner lieben Frau Diana danke ich für das stets offene Ohr bei Fragen und Unklarheiten sowie konstruktive Hinweise bei der Erstellung der Arbeit.

Ein großer Dank geht an alle, die diese Arbeit Korrektur gelesen haben Verbesserungsvorschläge brachten: Nadin Ulrich, Dr. Ralph Kühne, Rebecca Hiltrop, Julia Thürich, Dr. Torsten Thalheim und Thomas Putzmann.

Ganz besonderer Dank gilt meinen Arbeitskollegen am Helmholtz-Zentrum für Umweltforschung und neu gewonnenen Freunden, vor allem Franziska Schramm, Rebecca Hiltrop, Anja Laqua, Stefanie Finsterbusch-Kettner und Denis Mulliner, die mir auch in schwierigeren Zeiten hilfreich zur Seite standen. Der tolle Zusammenhalt der Arbeitsgruppe hat mir oft neue Motivation gegeben.

Ein riesen Dank geht an „die Mädels“. Danke für leckere Bohnensuppe, Sport im Turm, tolle Wanderungen, lustige Heimatbesuche und einen Mord's Spaß während der gesamten Zeit. Ohne euch hätte ich wohl nicht durchgehalten.

Meinen Eltern Maria und Bernhard sowie meiner kleinen Schwester Sarah danke ich auf herzlichste Weise. Ihr Ansporn, Zuspruch und ihr Beistand zu jedem Zeitpunkt haben zum erfolgreichen Abschluss dieser Arbeit beigetragen.

Zu guter Letzt möchte ich mich zutiefst bei meinem Freund Thomas Putzmann bedanken, der in Höhen und Tiefen zu mir steht und mich so liebt wie ich bin.

# Inhaltsverzeichnis

## Inhaltsverzeichnis

Abkürzungen, Akronyme und Symbole.....	I
Abbildungsverzeichnis .....	IV
Tabellenverzeichnis.....	VII
1 Einleitung und Motivation .....	1
2 Theorie zu Verteilungskoeffizienten und Dampfdruck.....	4
2.1 Definitionen von Verteilungskoeffizienten.....	4
2.2 Erforderliche Koeffizienten für PBPK-Modelle .....	4
2.2.1 Blut/Luft-Verteilung.....	9
2.2.2 Fett/Luft- und Fett/Blut-Verteilung.....	15
2.3 Literatur-Modelle zur Vorhersage von Verteilungskoeffizienten.....	16
2.3.1 Lineare und nicht-lineare Modelle .....	16
2.3.2 LSER-Modelle .....	21
2.4 Dampfdruck.....	24
2.4.1 Zusammenhang des Dampfdrucks mit der Henry-Konstante .....	24
2.4.2 Vorhersagemodelle für den Dampfdruck.....	27
3 Material und Methoden .....	32
3.1 Modelliermethoden .....	32
3.1.1 Fragmentmodell .....	32
3.1.2 Elektrotopologische Indices .....	34
3.2 Charakterisierungen der Datensätze.....	35
3.2.1 Datensatz von $K_{ba}$ .....	35
3.2.2 Datensatz von $K_{fa}$ .....	38
3.2.3 Datensatz von $P_v$ .....	41
3.3 Verteilungskoeffizienten und Parameter zur Modellierung.....	44
3.3.1 Abraham-Parameter .....	45
3.3.2 Wasser/Luft-Verteilungskoeffizient.....	45
3.3.3 Oktanol/Wasser-Verteilungskoeffizient.....	46
3.3.4 Oktanol/Luft-Verteilungskoeffizient.....	47
3.3.5 Membran/Wasser-Verteilungskoeffizient .....	47
3.3.6 Humanalbumin/Wasser-Verteilungskoeffizient.....	48
3.3.7 Olivenöl/Luft-Verteilungskoeffizient.....	48
3.3.8 Wasserlöslichkeit .....	49
3.4 Programme .....	50
3.5 Statistische Parameter .....	50
4 Modellierung von Verteilungskoeffizienten .....	54
4.1 Diskussion der Abraham-Modelle aus der Literatur .....	54
4.1.1 Blut/Luft-Verteilung.....	54
4.1.2 Fett/Luft-Verteilung .....	56
4.2 $K_{ba}$ -Modelle .....	58
4.2.1 Modellierung von $\log K_{ba}$ mit Linearen Solvatations-Energie-Beziehungen ..	58
<i>Ausreißer-Betrachtung: Tetracyanoethylenoxid (TCNEO)</i> .....	64
4.2.2 Modellierung von $\log K_{ba}$ mit anderen Verteilungskoeffizienten .....	66
4.2.2.1 Zusammengefasste LSER-Gleichungen.....	77
4.2.2.2 Einsatz von physiologischen Fraktionsanteilen .....	78
4.2.3 Fragmentmodell für $K_{ba}$ .....	81
4.2.4 $K_{ba}$ -Modellierung mit E-States .....	85
4.2.5 Konsensmodellierungsstrategie.....	86

## Inhaltsverzeichnis

4.3	$K_{fa}$ -Modelle .....	87
4.3.1	Modellierung von $\log K_{fa}$ mit Linearen Solvations-Energie-Beziehungen ...	87
4.3.2	Modellierung von $\log K_{fa}$ mit anderen Verteilungskoeffizienten .....	93
4.3.3	Konsensmodellierungsstrategie.....	102
4.4	Verteilungskoeffizient $K_{fb}$ .....	103
5	Dampfdruck-Modellierung.....	113
5.1	Literatur-Modelle .....	113
5.2	Modellierung von $\log P_v$ mit dem H+S-Modell.....	126
5.3	Modellierung von $\log P_v$ mit dem Mackay-Ansatz.....	128
5.4	Modellvergleiche für den Dampfdruck .....	135
6	Zusammenfassung und Ausblick .....	138
7	Literaturverzeichnis.....	141
8	Anhang .....	152
8.1	Blut/Luft-Verteilungskoeffizient $K_{ba}$ .....	152
8.1.1	$K_{ba}$ -Daten und -Werte.....	152
8.1.2	Werte aller zur $K_{ba}$ -Modellierung verwendeten Verteilungskoeffizienten ...	165
8.1.3	Fragmente und E-States .....	171
8.2	Fett/Luft-Verteilungskoeffizient $K_{fa}$ .....	173
8.2.1	$K_{fa}$ -Daten und -Werte .....	173
8.2.2	Werte aller zur $K_{fa}$ -Modellierung verwendeten Verteilungskoeffizienten.....	181

## Abkürzungen, Akronyme und Symbole

### Abkürzungen, Akronyme und Symbole

a	Luftphase (englisch: <i>air</i> )
AFC	Atom/Fragment-Beitrags-Methode (englisch: <i>atom/fragment contribution</i> )
ACF	atomzentrierte Fragmente (englisch: <i>atom centred fragments</i> )
ADME	Absorption, Distribution, Metabolisierung und Exkretion
atm	Physikalische Atmosphäre
b	Blutphase
ber.	berechnet
bzw.	beziehungsweise
<i>c</i>	Konzentration
$c_a$	Konzentration eines Stoffes in Luft
$c_{tot}$	Gesamtkonzentration (englisch: <i>total concentration</i> )
ca.	circa
CAS	Registrierungsnummer (englisch: <i>Chemical Abstract Service</i> )
CV	Kreuzvalidierung (englisch: <i>cross validation</i> )
$c_w$	Konzentration eines Stoffes in Wasser
Da	Dalton
Desk.	Deskriptor(en)
E-States	elektrotopologische Indices
<i>etc.</i>	<i>et cetera</i>
exp.	experimentell
f	Fettphase
GA	genetischer Algorithmus
<i>H</i>	Henry-Konstante
HBN	Anzahl an Wasserstoffbrücken (englisch: <i>hydrogen bond number</i> )
HIA	menschliche intestinale Resorption (englisch: <i>human intestinal absorption</i> )
HM	heuristische Methode
HSA	Humanalbumin (englisch: <i>human serum albumin</i> )
IVE	iterative uniformative Variableneliminierung
<i>K</i>	Verteilungskoeffizient
$K_{ba}$	Blut/Luft-Verteilungskoeffizient
$K_{fa}$	Fett/Luft-Verteilungskoeffizient
$K_{fb}$	Fett/Blut-Verteilungskoeffizient
$K_{HSA}$	Humanalbumin/Wasser-Verteilungskoeffizient

## Abkürzungen, Akronyme und Symbole

$K_{\text{HSA/a}}$	Humanalbumin/Luft-Verteilungskoeffizient
$K_{\text{mw}}$	Membran/Wasser-Verteilungskoeffizient
$K_{\text{oa}}$	Oktanol/Luft-Verteilungskoeffizient
$K_{\text{olive}}$	Olivenöl/Luft-Verteilungskoeffizient
$K_{\text{ow}}$	Oktanol/Wasser-Verteilungskoeffizient
$K_{\text{ta}}$	Gewebe/Luft-Verteilungskoeffizient
$K_{\text{wa}}$	Wasser/Luft-Verteilungskoeffizient
KNN (ANNs)	künstliche neuronale Netze (englisch: <i>artificial neural networks</i> )
kNN	k-nächste-Nachbarn-Algorithmus (englisch: <i>k-nearest-neighbours</i> )
log	Logarithmus, logarithmisch
LOO	Kreuzvalidierungsverfahren Leave-One-Out
LSER	Lineare Solvations-Energie-Beziehungen (englisch: <i>linear solvation energy relationship</i> )
m	Membranphase
$me$	mittlerer absoluter Fehler (englisch: <i>mean error</i> )
MLR	multiple lineare Regression
$mne$	maximaler negativer Fehler (englisch: <i>maximum negative error</i> )
$mpe$	maximaler positiver Fehler (englisch: <i>maximum positive error</i> )
MW	molekulare Masse (englisch: <i>molecular weight</i> )
$n$	Stoffmenge
NLR	nicht-lineare Regression
Nr.	Nummer
o	Oktanolphase
olive	Olivenölphase
OSIRIS	EU-Projekt mit dem Akronym: Optimised Strategies for Risk Assessment of Industrial Chemicals through Integration of Non-Test and Test Information
p	Proteinphase
Pa	Pascal
PBPK	physiologiebasierte Pharmakokinetik
PCA	Hauptkomponentenanalyse (englisch: <i>Principal Component Analysis</i> )
PCR	Hauptkomponentenregression (englisch: <i>Principal Component Regression</i> )
PLS	Pfadanalyse Partial Least Squares
$P$	Partialdruck
$P_v$	Dampfdruck
QSAR	Quantitative Struktur-Aktivitätsbeziehungen (englisch: <i>quantitative structure activity relationships</i> )
QSPR	Quantitative Struktur-Eigenschaftsbeziehung (englisch: <i>quantitative structure property relationship</i> )

## Abkürzungen, Akronyme und Symbole

$R$	universelle Gaskonstante
RBFNNs	radiale Basisfunktion neuronale Netzwerke
$rms$	Wurzel der Fehlerquadratsumme (englisch: <i>root-mean-square-error</i> )
$sd$	Standardfehler (englisch: <i>standard deviation</i> )
SMILES	chemischer Strukturcode (englisch: <i>Simplified Molecular Input Line Entry Specification</i> )
SPME	Festphasenmikroextraktion (englisch: <i>solid phase micro extraction</i> )
$S_w$	Wasserlöslichkeit
SVM	vektorgestützte Methode (englisch: <i>support vector machine</i> )
$T$	(System-)Temperatur
$t$	Gewebe (englisch: <i>tissue</i> )
$T_b$	Siedetemperatur (englisch: <i>boiling temperature</i> )
$T_m$	Schmelztemperatur (englisch: <i>melting temperature</i> )
UVE	uniformative Variableneliminierung
$V$	Volumen
$V_s$	Feststoffvolumen
$V_w$	Wasservolumen
$V_{tot}$	Gesamtvolumen (englisch: <i>total volume</i> )
vgl.	vergleiche
VOCs	flüchtige organische Substanzen (englisch: <i>volatile organic compounds</i> )
w	Wasserphase
z. B.	zum Beispiel

## Abbildungsverzeichnis

### Abbildungsverzeichnis

Abbildung 2.1:	Schematische Darstellung eines PBPK-Modells. (Einzelne Bilder aus [14]).....	5
Abbildung 2.2:	Schematische Darstellung des Blut/Luft-Verteilungskoeffizient: Wie Fremdstoffe in den Körper gelangen. (Einzelne Bilder aus [14]).....	6
Abbildung 2.3:	Schematische Darstellung der Fett/Blut-Verteilung, hier am Beispiel der Niere gezeigt. (Einzelne Bilder aus [14]) .....	7
Abbildung 2.4:	Schematische Darstellung der Blut/Luft-Verteilung in den beiden Kompartimenten Blut und Luft.....	9
Abbildung 2.5:	Schematische Darstellung der Blut/Luft-Verteilung mit Subkompartimenten. ....	10
Abbildung 2.6:	Detaillierte schematische Darstellung des Kompartiments Blut.....	10
Abbildung 2.7:	Einfache schematische Darstellung des Kompartiments Blut mit seinen drei Komponenten Wasser (w), Protein (p) und Lipid (L). ....	12
Abbildung 4.1:	Diagramm der neu angepassten LSER-Gleichung (4.4) mit experimentellen Deskriptoren. Auftragung von vorhergesagtem gegen experimentellen $\log K_{ba}$ . ....	60
Abbildung 4.2:	Strukturformeln der fünf Glykolether 2-Methoxyethanol, 2-Ethoxyethanol, 2-Butoxyethanol, 2-Isopropoxyethanol und 1-Methoxy-2-propanol. ....	60
Abbildung 4.3:	Strukturformeln der Ausreißer 1,4-Dioxan, Isophoron und Cyclohexanon. ....	61
Abbildung 4.4:	Auftragung von vorhergesagtem gegen experimentelles $\log K_{ba}$ über die neu angepasste LSER-Gleichung (4.5) mit berechneten Deskriptoren. ....	62
Abbildung 4.5:	Strukturformel von Tetracyanoethylenoxid.....	64
Abbildung 4.6:	Auftragung von vorhergesagtem gegen experimentelles $\log K_{ba}$ nach Gleichung (4.7). ....	67
Abbildung 4.7:	Diagramm von berechnetem $\log K_{ba}$ über Gleichung (4.8) gegen experimentelles $\log K_{ba}$ . ....	69
Abbildung 4.8:	Diagramm von berechnetem $\log K_{ba}$ über Gleichung (4.9) gegen experimentelles $\log K_{ba}$ . ....	70
Abbildung 4.9:	Auftragung von berechneten $\log K_{ba}$ -Werte über Gleichung (4.10) gegen experimentelle Werte. ....	71
Abbildung 4.10:	Auftragung von berechneten $\log K_{ba}$ -Werte über Gleichung (4.13) gegen experimentelle Werte. ....	73
Abbildung 4.11:	Auftragung von berechneten $\log K_{ba}$ -Werte über Gleichung (4.18) gegen experimentelle Werte. ....	79
Abbildung 4.12:	Auftragung von berechneten $\log K_{ba}$ -Werte über Gleichung (4.19) gegen experimentelle Werte. ....	80
Abbildung 4.13:	Strukturformel von Divinylether.....	83
Abbildung 4.14:	Vorhergesagte $\log K_{ba}$ -Werte mit dem Fragmentmodell aufgetragen gegen experimentelle $\log K_{ba}$ -Werte.....	84
Abbildung 4.15:	Diagramm der neu angepassten LSER-Gleichung (2.21) für $\log K_{fa}$ mit exp. Stoffparametern; der Pfeil zeigt den am stärksten überschätzten Stoff Hexachlorethan. ....	89
Abbildung 4.16:	Strukturformel von Hexachlorethan. ....	89

## Abbildungsverzeichnis

Abbildung 4.17: Diagramm der neu angepassten LSER-Gleichung (2.21) für $\log K_{fa}$ mit ber. Deskriptoren; die sechs gekennzeichneten Stoffe sind mit $> 0,5$ logarithmischen Einheiten überschätzt. ....	90
Abbildung 4.18: Strukturformel von Ethen. ....	91
Abbildung 4.19: Strukturformel von Tetrachlormethan. ....	91
Abbildung 4.20: Strukturformel von Oktan. ....	91
Abbildung 4.21: Strukturformel von Cyclopropan. ....	91
Abbildung 4.22: Strukturformel von Tricyclodekan. ....	92
Abbildung 4.23: Grafische Darstellung der berechneten gegen experimentelle Werte des Fett/Luft-Verteilungskoeffizienten. ....	93
Abbildung 4.24: Strukturformel von Dekan. ....	94
Abbildung 4.25: Grafische Darstellung des Modells (4.23) mit $K_{olive}$ ; die drei gekennzeichneten Datenpunkte sind die am stärksten überschätzten Stoffe im Datensatz. ....	95
Abbildung 4.26: Auftragung der Regression von berechnetem $\log K_{fa}$ gegen experimentelles $\log K_{fa}$ über Gleichung (4.24). ....	97
Abbildung 4.27: Auftragung der Regression von mit Gleichung (4.25) berechnetem $\log K_{fa}$ gegen experimentelles $\log K_{fa}$ . ....	98
Abbildung 4.28: Strukturformel von Tetrachlorethen. ....	103
Abbildung 4.29: Strukturformel von Styrol. ....	104
Abbildung 4.30: Auftragung der mit LSER berechneten $\log K_{fb}$ - Werte gegen experimentelle Werte. Eingezeichnet ist die xy-Gerade und der Ausreißer Tetrachlorethen (eingekreist). ....	106
Abbildung 4.31: Grafische Darstellung berechneten $\log K_{fb}$ -Werte mit den Modellen der anderen Verteilungskoeffizienten Nr. 3 aus Tabelle 4.5 und Nr. 3 aus Tabelle 4.12 gegen experimentelle Werte. ....	107
Abbildung 4.32: Auftragung von berechnetem $\log K_{fb}$ über die LSER-Gleichungen von $\log K_{fa}$ und $\log K_{ba}$ gegen experimentelle Werte. Die Werte streuen um die xy-Gerade. ....	109
Abbildung 4.33: Strukturformeln von Difluormethan, 1-Nitropropan und 2-Nitropropan. .	109
Abbildung 4.34: Grafische Darstellung der vorhergesagten $\log K_{fb}$ -Werte über andere Verteilungskoeffizienten gegen experimentelle Werte. Die Werte streuen um die xy-Gerade. ....	111
Abbildung 4.35: Auftragung von berechnetem $\log K_{fb}$ über LSER und Fragment gegen experimentelle Werte. Die Werte streuen um die xy-Gerade. ....	112
Abbildung 5.1: Strukturformel von Bioallethrin. ....	114
Abbildung 5.2: Auftragung von berechneten gegen experimentelle $\log P_v$ -Werte aus dem Programm EPI Suite [118]. ....	115
Abbildung 5.3: Strukturformel von 1,4,5,8-Tetraaminoanthrachinon (DB1). ....	116
Abbildung 5.4: Auftragung von mit Henry-Konstante und Wasserlöslichkeit berechneten gegen experimentelle aus EPI Suite $\log P_v$ -Werte. ....	117
Abbildung 5.5: Strukturformel von 1,4-Diamino-5-nitroanthracen-9,10-dion (C.I. Disperse Violet 8). ....	118
Abbildung 5.6: Auftragung berechneter gegen experimentelle Dampfdruck-Werte; über das H+S-Modell berechnet, wobei die Wasserlöslichkeit aus dem klassenbasierten Modell stammt. ....	118
Abbildung 5.7: Auftragung berechnete gegen experimentelle $\log P_v$ -Werte; berechnet über das Grain-Watson-Modell bei 25 °C. ....	120

## Abbildungsverzeichnis

Abbildung 5.8:	Auftragung der berechneten gegen experimentelle $\log P_v$ -Werte; berechnet über das Grain-Antoine-Modell bei 25 °C. ....	121
Abbildung 5.9:	Auftragung berechnete gegen experimentelle $\log P_v$ -Werte; berechnet über die Konstanten $A$ , $B$ und $C$ des Grain-Antoine-Modell. ....	122
Abbildung 5.10:	Strukturformeln von C.I. Disperse Red 1 (1), Tribenuronmethyl (3), 4-Hydroxy-4'-nitroazobenzen (4) und Methyl-7-chloro-2-[[[(4-trifluoromethoxyphenyl)(methoxycarbonyl)amino]carbonyl]-2,9-dihydroindeno[2,1-E][1,3,4]oxadiazin (2). ....	123
Abbildung 5.11:	Strukturformel von Amicarbazon. ....	124
Abbildung 5.12:	Auftragung der berechneten gegen die experimentellen Dampfdruck-Werte; berechnet über das Mackay-Modell. ....	125
Abbildung 5.13:	H+S-Modell mit den Teildatensätzen für flüssige ° und feste + Substanzen im $P_v$ -Datensatz. ....	128
Abbildung 5.14:	Mackay-Modell mit den Teildatensätzen für flüssige ° und feste + Substanzen im $P_v$ -Datensatz. ....	130
Abbildung 5.15:	Strukturformeln von Trimethylbleichlorid und Methansulfonylfluorid. ...	131
Abbildung 5.16:	Strukturformel von 1,4-Difluorbenzol. ....	131
Abbildung 5.17:	Auftragung der vorhergesagten gegen experimentelle $\log P_v$ -Werte für das Flüssigmodell (5.11). ....	133
Abbildung 5.18:	Kurzes Mackay-Modell ohne den zweiten Term mit den Teildatensätzen für flüssige ° und feste + Substanzen im $P_v$ -Datensatz....	134

## Tabellenverzeichnis

### Tabellenverzeichnis

Tabelle 2.1:	Kombinationen von Verteilungskoeffizienten zur Beschreibung des Wasserteils und des organischen Teils im Blut. ....	11
Tabelle 4.1:	Statistische Parameter der nachgerechneten LSER-Gleichung für $\log K_{ba}$ (2.22) mit experimentellen und berechneten Abraham-Deskriptoren.....	59
Tabelle 4.2:	Statistische Parameter der neu angepassten LSER-Gleichung (4.4) mit experimentellen Abraham-Deskriptoren im Vergleich zu zwei Literaturmodellen.....	61
Tabelle 4.3:	Statistische Parameter der neu angepassten LSER-Gleichung mit experimentellen und berechneten Abraham-Deskriptoren. ....	63
Tabelle 4.4:	LSER-Deskriptoren von TCNEO. ....	64
Tabelle 4.5:	Modelle, Regressions- und Statistikparameter für $\log K_{ba}$ -Modelle nach dem Schema $a \cdot \log A + b \cdot \log B + c \cdot \log C + d$ . ....	75
Tabelle 4.6:	Vergleich der Regressions- und ausgewählter Statistikparameter der Modelle (4.19) und (4.13) mit $n = 190$ . ....	81
Tabelle 4.7:	Fragmente mit Regressionsparametern, Fehlern und p-Werte für das Fragmentmodell für $\log K_{ba}$ . ....	82
Tabelle 4.8:	Korrelationen der Fehler von vier empfohlenen Modellen, wobei drei verschiedene Modelltypen vertreten sind. ....	86
Tabelle 4.9:	Statistische Parameter der nachgerechneten LSER-Gleichung für $\log K_{fa}$ (2.23) mit experimentellen und berechneten Abraham-Deskriptoren.....	88
Tabelle 4.10:	Statistische Parameter der neu angepassten LSER-Gleichungen (4.20) und (4.21) für $\log K_{fa}$ mit experimentellen und berechneten Abraham-Deskriptoren.....	92
Tabelle 4.11:	Statistische Parameter der Gleichung (4.23) mit $K_{olive}$ mit experimentellen und berechneten Abraham-Deskriptoren. ....	96
Tabelle 4.12:	Modelle, Regressions- und Statistikparameter für $\log K_{fa}$ -Modelle nach dem Schema $a \cdot \log A + b \cdot \log B + c \cdot \log C + d$ . ....	100
Tabelle 4.13:	Korrelationen der Fehler von drei empfohlenen Modellen, wobei zwei verschiedene Modelltypen vertreten sind. ....	102
Tabelle 4.14:	Substanzen, für die experimentelle Werte für $\log K_{fb}$ in der Literatur vorhanden sind. ....	104
Tabelle 5.1:	Statistische Parameter der Berechnung des Dampfdrucks mit dem internen Vorhersage-Programm und Vergleich mit dem internen experimentellen $P_v$ -Datensatz von EPI Suite.....	114
Tabelle 5.2:	Statistik-Ergebnisse des Grain-Watson-Modells zur Vorhersage von $\log P_v$ bei 25 °C und 37 °C ....	119
Tabelle 5.3:	Statistische Parameter bei Anwendung des Mackay-Modells auf den hier behandelten Datensatz. ....	124
Tabelle 5.4:	Vergleich der statistischen Parameter der hier vorgestellten Literatur-Modelle zur Dampfdruck-Vorhersage. ....	126
Tabelle 5.5:	$P_v$ -Modelle mit Regressions- und Statistikparameter.....	136

## Tabellenverzeichnis

Tabelle 8.1:	Verwendete Substanzen für die $K_{ba}$ -Modellierung mit Name, SMILES-Code, ChemProp ID und CAS Nr. und experimentelle aus der Literatur gesammelte Werte für Mensch und Ratte und Quellen. ....	152
Tabelle 8.2:	Experimentelle und berechnete LSER-Deskriptoren für den $K_{ba}$ -Datensatz.....	159
Tabelle 8.3:	Alle Verteilungskoeffizienten mit Werten, die zur Modellierung von $K_{ba}$ verwendet werden. ....	165
Tabelle 8.4:	Alle 27 Fragmente, die als Ausgangspunkt für die Erstellung des Fragmentmodells für den Blut/Luft-Verteilungskoeffizienten verwendet werden. ....	171
Tabelle 8.5:	Einige wichtige Atomgruppen der E-States mit den dazugehörigen Gruppensymbolen. ....	172
Tabelle 8.6:	Verwendete Substanzen für die $K_{fa}$ -Modellierung mit Name, SMILES-Code, ChemProp ID und CAS Nr. und experimentelle aus der Literatur gesammelte Werte für Mensch und Ratte und Quellen. ....	173
Tabelle 8.7:	Experimentelle und berechnete LSER-Deskriptoren für den $K_{ba}$ -Datensatz.....	177
Tabelle 8.8:	Alle Verteilungskoeffizienten mit Werten, die zur Modellierung von $K_{fa}$ verwendet werden. ....	181

# Einleitung

## 1 Einleitung und Motivation

Physiologie-basierte Pharmakokinetik (PBPK)-Modelle beschreiben den Verbleib von Chemikalien im Körper von Mensch oder Tieren. Prozesse die diesen Verbleib hauptsächlich regulieren sind Aufnahme eines Stoffes (*absorption*), die Verteilung im Körper (*distribution*), der (bio-)chemische Um- und Abbau (*metabolism*) sowie die Ausscheidung (*excretion*) und werden ADME-Parameter genannt [1-5]. PBPK-Modelle beruhen auf der anatomischen und physiologischen Struktur des Körpers und zu einem gewissen Maß auf der dazugehörigen Biochemie. Üblicherweise sind sie Mehr-Kompartimenten-Modelle, in denen die Kompartimente als Organe oder Gewebe definiert sind und die Verbindungen untereinander die Blut- oder Lymphflüsse darstellen. Für jedes Kompartiment kann ein System von Differentialgleichungen zur Bestimmung der Konzentration oder Menge der Substanz aufgestellt werden. Seine Parameter repräsentieren Blutfluss, Atemfrequenz, Organvolumina *etc.*, für die die nötigen Informationen entweder eigene experimentelle Daten oder in wissenschaftlichen Publikationen verfügbar sind. Die Beschreibung des Körpers mit einem Modell ist sehr vereinfacht, aber es muss die Balance zwischen realer Komplexität und modellhafter Einfachheit gefunden werden [6].

PBPK-Modelle sind durch die mathematische Beschreibung von anatomischen, physiologischen, physischen und chemischen Phänomenen, die im komplexen ADME-Prozess beteiligt sind, mechanistisch. Eine starke Vereinfachung und ein hoher Grad an Empirismus sind in diesen Modellen dennoch vorhanden. PBPK-Modelle können einen rein voraussagenden Nutzen haben, aber auch andere Anwendungen wie statistische Folgerungen mithilfe der Entwicklung von Bayesschen Statistikmethoden, die komplexe Modelle behandeln [7]. Das gilt sowohl für Toxizitäts-Risikobewertung als auch für therapeutische Arzneimittelentwicklung.

Interessanterweise war das erste pharmakokinetische Modell im Jahr 1937, das in der wissenschaftlichen Literatur beschrieben wurde [8], ein PBPK-Modell. Es brachte jedoch damalige Computerberechnungen an ihre Grenzen. Der Schwerpunkt verlagerte sich danach auf einfachere Modelle, für die eine analytische Lösung erhalten werden konnte. Die Verfügbarkeit von Computern und numerischen Integrationsalgorithmen kennzeichnete ein erneutes Interesse an physiologischen Modellen in den frühen 1970er Jahren. Viele wissenschaftliche Publikationen haben bis heute PBPK-Modelle beschrieben sowie angewandt und zusätzlich werden einige Software-Tools in diesem Bereich angeboten [9-11].

## Einleitung

Für PBPK-Modelle werden Verteilungskoeffizienten benötigt, die wiedergeben, in welcher Phase oder in welchem Kompartiment im Körper sich ein Stoff bevorzugt aufhält. Daraus kann auf eine mögliche toxikologische Gefahr einer einzelnen Substanz geschlossen oder der Wirkungsbereich von Pharmaka verstanden und gesteuert werden.

Ziel dieser Arbeit ist es, PBPK-relevante Verteilungskoeffizienten zu modellieren, wobei die folgenden Kompartimente für eine genaue Untersuchung des Verbleibs von Chemikalien im Körper benötigt werden: Blut/Luft, Fett/Luft, Fett/Blut, Leber/Luft, Leber/Blut, Muskel/Luft und Muskel/Blut. Da der Fett/Blut-Verteilungskoeffizient in einer einfachen Näherung über die Fett/Luft- und Blut/Luft-Koeffizienten berechnet werden kann, muss dieser nicht direkt modelliert werden. Alle anderen Verteilungskoeffizienten, welche die Organe Leber und Muskel betreffen, können über den Fettgehalt des jeweiligen Organs oder Gewebes berechnet werden. Somit reduziert sich der Modellierungsaufwand und beläuft sich nur auf die beiden Verteilungskoeffizienten: Blut/Luft und Fett/Luft.

Diese Verteilungskoeffizienten werden mit quantitativen Struktur-Aktivitätsbeziehungen, sogenannten QSAR-Modellen erstellt (QSAR, englisch: *quantitative structure activity relationship*), die die Eigenschaften von Chemikalien nur auf der chemischen Struktur basierend vorhersagen und somit keine experimentellen Eingabeparameter erlauben.

Für den Blut-Luft-Verteilungskoeffizienten werden drei Ansätze verfolgt: der LSER-Ansatz, der auf linearen Solvations-Energie-Beziehungen (LSER, englisch: *linear solvation energy relationship*) beruht, ein Ansatz mit anderen Verteilungskoeffizienten und ein Fragmentmodell, welches in Kapitel 3.1.1 auf Seite 32ff beschrieben ist. Für den Fett/Luft-Verteilungskoeffizienten werden nur die ersten beiden Modellieransätze verwendet, da damit bereits gute Ergebnisse erzielt werden können und der Datensatz für die Modellierung mit Fragmenten zu klein ist.

Eine wichtige Größe für die Modellierung dieser pharmakokinetischen Parameter ist die Henry-Konstante, die dem Luft/Wasser-Verteilungskoeffizienten entspricht. Dieser kommt bei der Modellierung mit anderen Verteilungskoeffizienten zum Einsatz und kann wiederum mit Hilfe des Dampfdrucks und der Wasserlöslichkeit bestimmt werden. Aus diesem Grund ist der Dampfdruck ein wichtiger Parameter für die PBPK-Modellierung, weshalb bereits bestehende Dampfdruck-Modelle miteinander verglichen und zwei Literatur-Ansätze genauer untersucht und modifiziert werden.

Ziel dieser Arbeit ist es, Vorhersagemodelle für PBPK-Zielgrößen, im konkreten Fall für Verteilungskoeffizienten und für den Dampfdruck, zu erstellen. Dabei werden Modelle aus

## Einleitung

der Literatur auf Verbesserungspotential geprüft, gegebenenfalls erweitert und neue sowie bessere entwickelt. Diese sollen nur auf der chemischen Struktur basieren und erlauben somit keine experimentellen Eingabeparameter. Mit unterschiedlichen und mehreren Modellalternativen steigt die Richtigkeit der Ergebnisse und durch die Vielfalt können die jeweils benötigten Aspekte der Vorhersage gezielter berücksichtigt werden. Ausgangspunkt für die Modellierung von Verteilungskoeffizienten ist eine LSER-Gleichung von Abraham *et al.* [12], welche mit experimentellen Stoffparametern brauchbare Vorhersagen ermöglicht, mit berechneten Deskriptoren jedoch nicht sehr befriedigend ist. Die Motivation, die sich daraus ergibt, ist methodisch komplementäre Modelle zu entwickeln, die zusammen eine größere Konfidenz ergeben, als getrennt betrachtet.

## 2 Theorie zu Verteilungskoeffizienten und Dampfdruck

### 2.1 Definitionen von Verteilungskoeffizienten

Die Definition von Verteilungskoeffizienten ist sehr einfach; sie besteht lediglich aus dem Quotienten der Konzentration eines Stoffes in einem Kompartiment und der Konzentration eines Stoffes in einem anderen, benachbarten Kompartiment. Dieses Verhältnis definiert den Verteilungskoeffizient  $K$ , der in Gleichung (2.1) für die Kompartimente A und B als  $K_{AB}$  gezeigt ist.

$$K_{AB} = \frac{\text{Konzentration der Chemikalie in Kompartiment A}}{\text{Konzentration der Chemikalie in Kompartiment B}} \text{ (im Gleichgewicht)} \quad (2.1)$$

### 2.2 Erforderliche Koeffizienten für PBPK-Modelle

PBPK-Modelle werden genutzt, um den Verbleib von Stoffen im Körper zu beschreiben. Diese Modelle nutzen biologisch relevante Parameter, z. B. Gewebe/Blut-Verteilungskoeffizienten, Organvolumina, Blutfluss-, Aufnahme-, Metabolisierungs- und Ausscheidungs-raten und leisten einen wichtigen Beitrag zur Umwelt-Risikobewertung [13]. Schematisch beschreibt ein PBPK-Modell den Körper des Menschen oder eines jeden beliebigen Säugetiers im Hinblick auf den Verbleib von chemischen Fremdstoffen (siehe Abbildung 2.1).

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

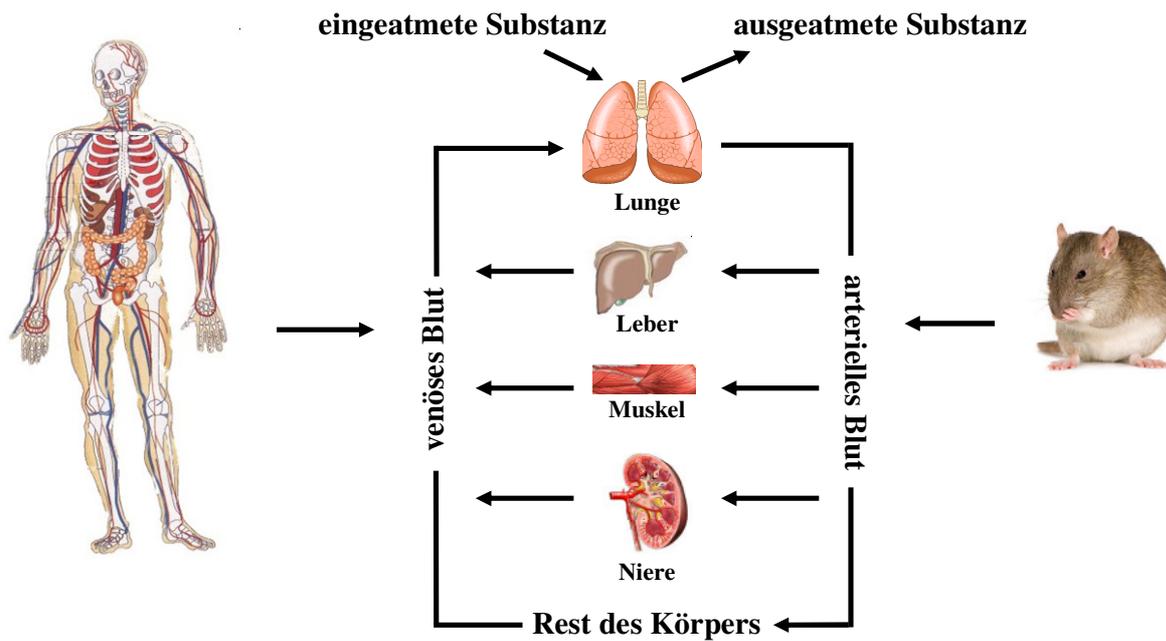


Abbildung 2.1: Schematische Darstellung eines PBPK-Modells. (Einzelne Bilder aus [14])

Somit dürfen wichtige Organe, wie die Lunge, die Leber, die Nieren und das Kompartiment Muskel bei einem PBPK-Modell nicht fehlen. Bedeutend ist natürlich auch der Blutfluss, der hauptsächlich für den Transport der eingedrungenen Chemikalie verantwortlich ist. Eine flüchtige Substanz wird eingeatmet und gelangt über die Lunge in einer bestimmten Konzentration in den Blutkreislauf (Abbildung 2.2). Dieser Prozess wird durch den Blut/Luft-Verteilungskoeffizienten bestimmt.

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

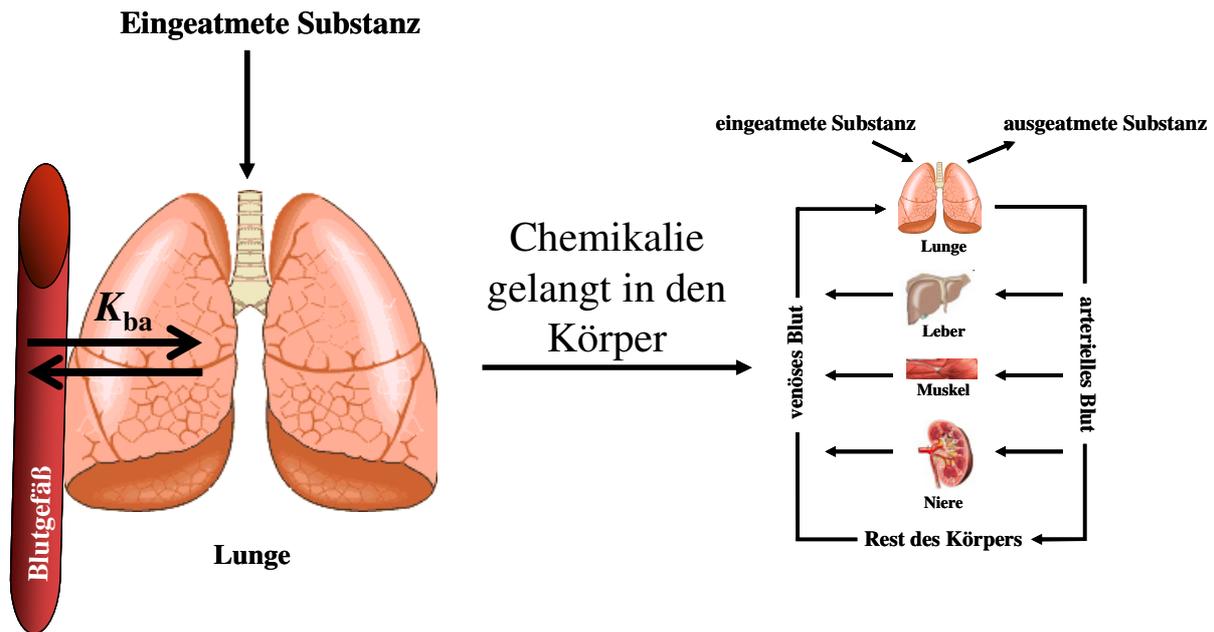


Abbildung 2.2: Schematische Darstellung des Blut/Luft-Verteilungskoeffizienten: Wie Fremdstoffe in den Körper gelangen. (Einzelne Bilder aus [14])

Über den Blutkreislauf kann der Fremdstoff in die oben genannten Organe gelangen, wobei er in den Körperkompartimenten akkumulieren kann. Diese sogenannte Bioakkumulation ist die Anreicherung einer Substanz im Organismus nach der Aufnahme über die Lunge, den Magen oder die Haut und tritt vorwiegend bei Substanzen auf, die weder rasch biochemisch ab- oder umgebaut noch zügig ausgeschieden werden. Die angereicherte Substanz kann in einem erhöhten Maße dem Körper Schaden zufügen. Sie kann aber auch weiter im Blutstrom fließen bis sie letztendlich über den Urin ausgeschieden wird, wobei die Niere die Schlüsselrolle spielt. Des Weiteren kann der Schadstoff in der Leber metabolisiert oder wieder über die Lunge ausgeatmet werden. Die hier behandelte Modellierung spezialisiert sich auf das Blut/Luft- und Fett/Luft-System, in denen sich flüchtige organische Substanzen (VOCs, englisch: *volatile organic compounds*) verteilen. Sind diese beiden Koeffizienten bekannt, kann auch der Fett/Blut-Verteilungskoeffizient errechnet werden. Mit einer groben Näherung kann über den Fettgehalt der Zielorgane und Gewebe der komplette Satz von bestimmten Verteilungskoeffizienten zwischen verschiedenen Geweben, Organen, Blut, Fett und Luft, wie sie als Eingangsgrößen für PBPK-Modelle benötigt werden, bestimmt werden. Bei dieser Näherung werden verschiedene Unterkompartimente, beispielsweise Proteine, vernachlässigt und es wird keine Unterscheidung zwischen Speicherfett und Membranfett vorgenommen. Deshalb wird als Vereinfachung für den Organ/Blut-Verteilungskoeffizienten zunächst der

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

Fett/Blut-Verteilungskoeffizient modelliert (Gleichung (2.2)), der über den jeweiligen Fettgehalt des Gewebes oder Organs mit einem Faktor (hier  $f_i$ ) umgerechnet werden kann. Hierfür wird ein Gleichungssystem erstellt, in dem jeweils anteilig der Fett- und Wassergehalt in dem Gewebe von Interesse eine Rolle spielt.

$$\frac{c_{\text{Organ}}}{c_{\text{Blut}}} = f_i \cdot \frac{c_{\text{Fett}}}{c_{\text{Blut}}} \quad (2.2)$$

Bildhaft ist dieser Sachverhalt am Beispiel der Niere in Abbildung 2.3 gezeigt. Ein Fremd- oder Wirkstoff befindet sich im Blutkreislauf und kann in das Zielorgan gelangen. Dieser Verteilungskoeffizient  $K$  ist definiert als das Verhältnis der Konzentrationen einer Chemikalie im Gewebe (in diesem Beispiel das anteilige Fettgewebe der Niere) und im Blut im Gleichgewicht [15].

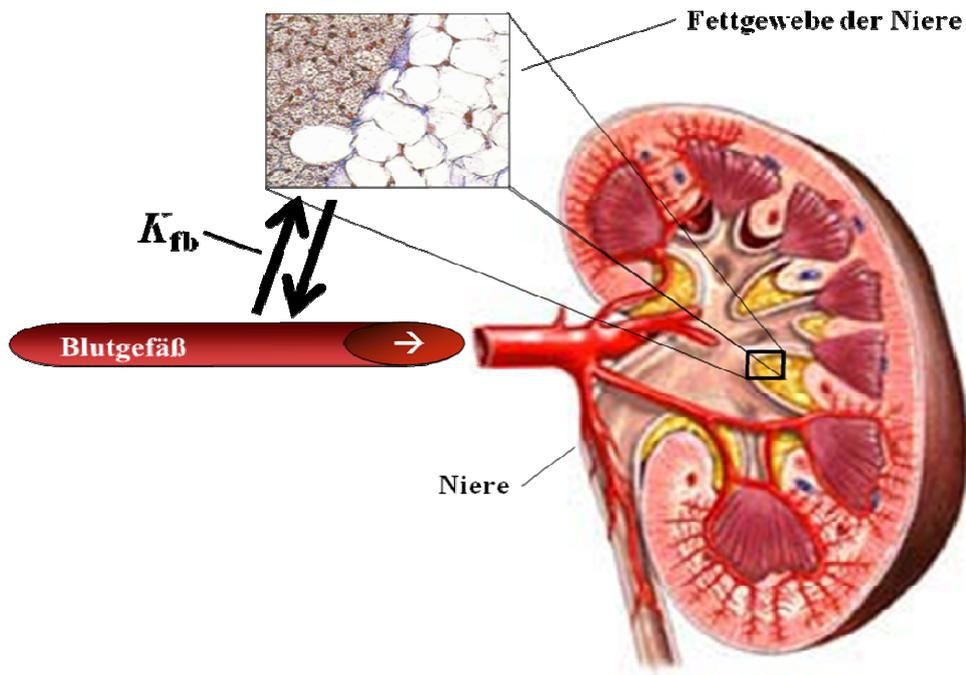


Abbildung 2.3: Schematische Darstellung der Fett/Blut-Verteilung, hier am Beispiel der Niere gezeigt. (Einzelne Bilder aus [14])

Die Umrechnung über den Fettanteil erfolgt durch Multiplikation mit einem Faktor, der den Fettanteil in dem Gewebe von Interesse wiedergibt. Dies ist am Beispiel des Leber/Blut-Koeffizient in Gleichung (2.3) gezeigt.

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

$$K_{lb} = f_1 \cdot K_{fb} \quad (2.3)$$

Um die relevanten Verteilungskoeffizienten für ein PBPK-Modell zu erhalten, dabei aber den Aufwand überschaubar zu halten, werden nur die Grundkoeffizienten Blut-zu-Luft und Fett-zu-Luft modelliert, aus denen die anderen abgeleitet werden.

Auch die Literatur belegt, dass diese Umrechnungen über den Lipidgehalt der Organe für PBPK-relevante Verteilungskoeffizienten sinnvoll sind.

Poulin und Krishnan [16] beschreiben den Gewebe/Luft-Verteilungskoeffizienten für ein PBPK-Modell mit dem Oktanol/Wasser-Verteilungskoeffizienten  $\log K_{ow}$ . Die Basis hierfür ist, dass die Löslichkeit einer Chemikalie in Gewebe oder Blut als Summe ihrer Löslichkeit in neutralen Lipiden, Phospholipiden und in Wasser, die in diesen Kompartimenten enthalten sind, ausgedrückt werden kann. Somit werden Gewebe und Organe über deren Lipidgehalt definiert. Die Autoren berechnen demnach den Gewebe/Blut-Koeffizienten für mehrere Stoffe basierend auf den Fettgehalt der Gewebe und den Oktanol/Wasser- oder Öl/Saline-Verteilungskoeffizienten der Verbindungen. Ihre Ergebnisse zeigen, wenn der Gewebe/Luft-Verteilungskoeffizient für ein Gewebe bekannt ist, können die Koeffizienten für die anderen Gewebearten auch über ihren Lipidgehalt berechnet werden. Daraus ergibt sich, dass der Blut/Luft-Koeffizient über den Fettgehalt im Blut berechnet wird und jeder Gewebe/Blut-Koeffizient aus den Gewebe/Luft- und Blut/Luft-Werten berechnet werden kann [13,17]. Für einige Verbindungen jedoch reicht der Fettgehalt im Blut allein nicht aus, um den Blut/Luft-Wert zu beschreiben. In diesen Fällen sollte die Bindung der Stoffe an Blutproteine zusätzlich berücksichtigt werden [17].

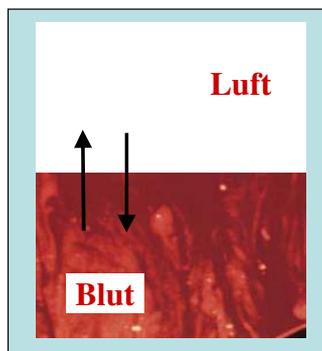
Eine weitere Arbeitsgruppe bestimmt die Werte für verschiedene Gewebe/Blut-Verteilungen über lineare Korrelationen mit dem Fett/Blut-Wert [18] für 35 hydrophile und hydrophobe Substanzen mit unterschiedlichen chemischen Strukturen.

Diese Literaturbeispiele zeigen, dass diese Näherung durchaus sinnvoll ist und auch anderswo Anwendung findet.

### 2.2.1 Blut/Luft-Verteilung

Der Blut/Luft-Verteilungskoeffizient gibt wieder, wie sich die Stoffkonzentrationen zwischen Blut und Luft verteilen. Anschaulich ist das in Abbildung 2.4 dargestellt und mathematisch wird es nach Gleichung (2.4) berechnet. Luft wird im Folgenden immer mit „a“ (Luft, englisch: *air*) abgekürzt.

$$K_{ba} = \frac{c_{\text{blood}}}{c_{\text{air}}} \quad (2.4)$$



**Abbildung 2.4:** Schematische Darstellung der Blut/Luft-Verteilung in den beiden Kompartimenten Blut und Luft.

Schematisch kann Blut in Wasser und organische Komponenten aufgeteilt werden, wie in Abbildung 2.5 dargestellt. Blut ist eine Körperflüssigkeit, die aus Blutzellen, suspendiert im Blutplasma besteht. Plasma, das einen ca. 55-prozentigen Anteil des Blutes ausmacht, besteht vor allem aus Wasser (90 % - 95 % Vol.) und beinhaltet gelöste Proteine, Glucose, Mineral-Ionen, Hormone, Kohlenstoffdioxid, Thrombozyten (Blutplättchen) und Blutzellen. Die Löslichkeit von VOCs im Blut ist in erster Linie abhängig von der Zusammensetzung, die auch die Fett- und Wasseranteile berücksichtigt sowie vom Grad der Bindung der Chemikalie an spezifische Komponenten im Blut, beispielsweise an Hämoglobin, Plasmaproteine, *etc.* [19]. Das Kompartiment Blut kann in eine wässrige und eine organische Fraktion aufgeteilt werden, die zelluläre Komponenten enthält. Auch Poulin und Krishnan beschreiben die Verteilung im Blutkompartiment als Summe der Verteilungen in Erythrozyten und Plasma [16], wobei Plasma hauptsächlich aus Wasser besteht und die Erythrozyten den Hauptbestandteil der zellulären Anteile ausmachen. Sowohl die wässrige Phase, als auch die

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

organische, zelluläre Phase beinhalten einen kleinen Anteil an Proteinen, von denen Humanalbumin (HSA, englisch: *human serum albumin*) den größten Teil ausmacht.

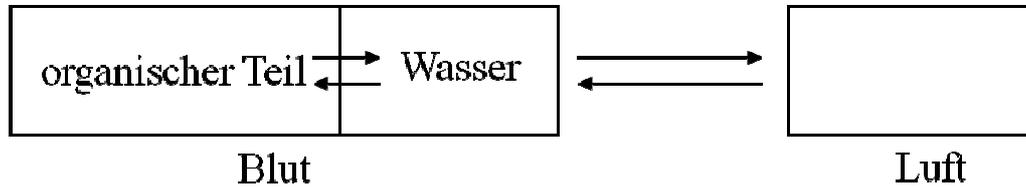


Abbildung 2.5: Schematische Darstellung der Blut/Luft-Verteilung mit Subkompartimenten.

Daher sollte es möglich sein, die Verteilung von Chemikalien zwischen Blut und Luft als eine Kombination der Verteilungen Wasser/Luft ( $\log K_{wa}$ ) und Oktanol/Wasser ( $\log K_{ow}$ ) oder Protein HSA/Wasser ( $\log K_{HSA}$ ) zu beschreiben. Der Verteilungskoeffizient Wasser/Luft ist die inverse dimensionslose Henry-Konstante, die über den Dampfdruck ermittelt werden kann. Der Beitrag für den organischen Teil kann zwischen einzelnen Koeffizienten für Oktanol, Membran, Protein HSA und Lipid oder als eine Kombination von mehreren variieren. Somit kann  $\log K_{ba}$  durch eine multilineare Regression von z. B.  $\log K_{wa}$  und  $\log K_{ow}$  vorhergesagt werden. Daraus ergeben sich entsprechende Regressionsparameter. Davon ausgehend ist das gesamte Modellkonzept ein mechanistischer Ansatz, wobei die Empirie darin steckt, dass kein richtiges Lipid modelliert wird, sondern dafür Hilfsgrößen verwendet werden. Da die Komponenten des organischen Teils nicht im direkten Kontakt mit der Luft stehen sondern von Wasser umgeben sind (siehe Abbildung 2.6), findet die Verteilung zwischen dem organischen Kompartiment und Wasser statt.

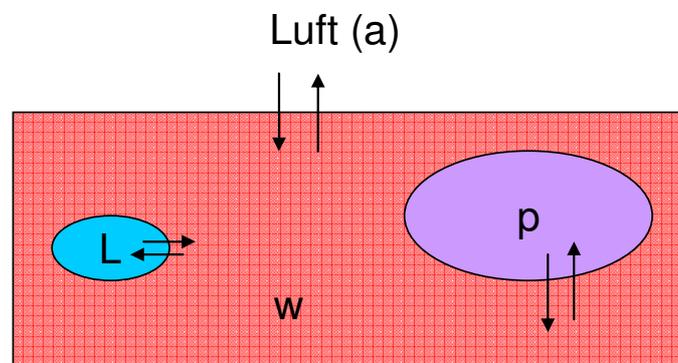


Abbildung 2.6: Detaillierte schematische Darstellung des Kompartiments Blut.

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

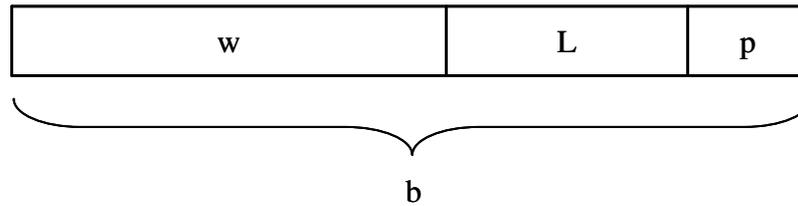
Beim 1-Parameter-Modell wird nur der Verteilungskoeffizient Wasser-zu-Luft herangezogen, um den Blut/Luft-Verteilungskoeffizienten für PBPK zu modellieren. Bei den 2-Parameter-Modellen kommt zusätzlich eine organische Komponente hinzu, die Oktanol, Protein, Membran oder Olivenöl sein kann und somit über die Verteilungskoeffizienten  $K_{ow}$ ,  $K_{HSA}$ ,  $K_{mw}$  oder  $K_{olive}$  beschrieben werden kann. Auch 3-Parameter-Modelle werden in dieser Arbeit untersucht, bei denen der organische Teil mit einer Kombination aus Oktanol und Protein oder Protein und Membran beschrieben wird. Diese Kombinationen sind in Tabelle 2.1 dargestellt. Die Modelle 1 bis 7 werden berechnet und die Ergebnisse diskutiert.

**Tabelle 2.1: Kombinationen von Verteilungskoeffizienten zur Beschreibung des Wasserteils und des organischen Teils im Blut.**

Kombinationen von Verteilungskoeffizienten zur Beschreibung von Blut	Wasser	organischer Teil (mit Hilfsgrößen Oktanol, Protein, Membran und Olivenöl)
Modell Nr. 1	$K_{wa}$	-
Modell Nr. 2	$K_{wa}$	$K_{ow}$
Modell Nr. 3	$K_{wa}$	$K_{HSA}$
Modell Nr. 4	$K_{wa}$	$K_{mw}$
Modell Nr. 5	$K_{wa}$	$K_{olive}$
Modell Nr. 6	$K_{wa}$	$K_{ow}, K_{HSA}$
Modell Nr. 7	$K_{wa}$	$K_{HSA}, K_{mw}$

In Abbildung 2.7 wird das Kompartiment Blut (b) dargestellt, welches in die drei Komponenten Wasser (w), Protein (p) und Lipid (L) aufgeteilt wird. Der organische Teil im Blut setzt sich somit aus Protein und Lipid zusammen. Insgesamt werden die Faktoren 50 % für Wasser, 40 % für Lipide und 10 % für Proteine angenommen und in die Vorhersage einberechnet. Lipide werden im Serum in Form von Lipoproteinen transportiert, da sie wasserunlöslich sind. In diesem Modell werden sie als eigene Phase angesehen.

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden



**Abbildung 2.7:** Einfache schematische Darstellung des Kompartiments Blut mit seinen drei Komponenten Wasser (w), Protein (p) und Lipid (L).

Für ein solches System lässt sich die Verteilung von Chemikalien zwischen Blut (b) und Luft (a) aus den einzelnen Bestandteilen des Blutes folgendermaßen (2.5) beschreiben:

$$K_{ba} = \Phi_w K_{wa} + \Phi_p K_{pa} + \Phi_L K_{La} \quad (2.5)$$

Die Fraktionen  $\Phi_w$ ,  $\Phi_p$  und  $\Phi_L$  sind jeweils die Volumenanteile in Wasser (siehe Gleichung (2.6)), Protein und Lipid.  $K_{wa}$ ,  $K_{pa}$  und  $K_{La}$  sind die Verteilungskoeffizienten zwischen den drei einzelnen Kompartimenten und Luft (a). In Gleichung (2.7) ist die Gesamtkonzentration  $c_{tot}$  entsprechend aus der Summe der Einzelkonzentrationen beschrieben.

$$\text{mit } \Phi_w = \frac{V_w}{V_w + V_p + V_L} \quad (2.6)$$

$$c_{tot} = \Phi_w c_w + \Phi_p c_p + \Phi_L c_L \quad (2.7)$$

Etwas genauer stellt Abbildung 2.6 (Seite 10) die Bestandteile und die Zusammensetzung des Blutes dar. Hier ist die größte Phase das Wasser (w), in dem die kleineren Anteile Lipid (L) und Protein (p) gelöst sind. Zusätzlich sind die Verteilungen zwischen den einzelnen Phasen mit Pfeilen dargestellt. Dadurch wird ersichtlich, dass zwischen Lipid und Wasser ein Konzentrationsgleichgewicht der Chemikalie vorliegt. Damit herrscht auch ein Gleichgewicht zwischen Wasser und Luft sowie Wasser und Protein.

Die Verteilung der Stoffe in Luft und Blut wird durch den Quotienten der Konzentrationen in der jeweiligen Phase bestimmt (Gleichung (2.8)). Die Phase Blut kann hier in beliebige Fraktionen  $i$  aufgeteilt werden, nach Abbildung 2.7 in Wasser, Lipid und Protein.

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

$$K_{ab} = \frac{c_a}{c_b} = \frac{c_a}{\sum_i \Phi_i \cdot c_i} \quad (2.8)$$

Aus dem Quotienten der Konzentrationen kann der Verteilungskoeffizient gewonnen werden, der sich wiederum aus den einzelnen Blutkompartimenten zusammensetzen lässt. Gleichung (2.9) zeigt die drei Fraktionen Wasser, Protein und Lipid analog zu Gleichung (2.5). Als Lipid kann beispielsweise Membran herangezogen werden oder Olivenöl, wie auch in [20] verwendet.

$$\frac{c_b}{c_a} = K_{ba} = \Phi_w \cdot \frac{c_w}{c_a} + \Phi_p \cdot \frac{c_p}{c_a} + \Phi_L \cdot \frac{c_L}{c_a} \quad (2.9)$$

$\underbrace{\hspace{1.5cm}} \quad \underbrace{\hspace{1.5cm}} \quad \underbrace{\hspace{1.5cm}}$ 
 $K_{wa} \quad K_{pa} \quad K_{La}$

$\downarrow \quad \downarrow$ 
 $K_{pw} \cdot K_{wa} \quad K_{mw} \cdot K_{wa}$

Die Koeffizienten  $K_{pa}$  und  $K_{La}$  können jeweils aus  $K_{pw}$  und  $K_{mw}$  (Membran/Wasser-Verteilungskoeffizient) durch Multiplikation mit  $K_{wa}$  gewonnen werden. Durch mathematisches Umformen entsteht Gleichung (2.10).

$$K_{ba} = K_{wa} (\Phi_w + \Phi_p K_{pw} + \Phi_L K_{mw}) \quad (2.10)$$

Nach Logarithmieren von Gleichung (2.10) ergibt sich:

$$\log K_{ba} = \log K_{wa} + \log(\Phi_w + \Phi_p K_{pw} + \Phi_L K_{mw}) \quad (2.11)$$

$$\underbrace{\hspace{10cm}}_{\approx a \log K_{pw} + b \log K_{mw}} \quad (2.12)$$

Gleichung (2.11) beschreibt die Verteilung chemischer Stoffe in Blut und entspricht damit der detaillierten Darstellung von Blut, die in Abbildung 2.6 (Seite 10) gezeigt wird. Der zweite

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

Teil der Gleichung (2.11) kann näherungsweise wie in Gleichung (2.12) beschrieben werden. Diese genäherte Gleichung (2.12) ist empirisch proportional zu Gleichung (2.11) und kann über eine multilineare Regression der Verteilungskoeffizienten  $\log K_{wa}$ ,  $\log K_{pw}$  und  $\log K_{mw}$  angepasst werden.

Gleichung (2.11) kann mit physiologischen Werten für die einzelnen Fraktionen für Wasser, Protein und Lipid berechnet werden. Als plausibelste Annahme wird  $\Phi_w$  gleich 0,5 gesetzt,  $\Phi_p$  ist 0,1 und  $\Phi_L$  erhält einen Wert von 0,4 (siehe Abbildung 2.7). Die Ergebnisse beider Ansätze ((2.11) und (2.12)) werden in dieser Arbeit miteinander verglichen.

Einige Autoren nutzen für die Abschätzung der Plasmaproteinbindung von Chemikalien im Blut die Bindung an das Protein Human Serum Albumin (HSA) [21-24]. Diese Bindung kann auch als Verteilungskoeffizient zwischen dem Protein und Wasser beschrieben werden, nämlich als  $K_{HSA}$ . Damit gibt es verschiedene Ansätze den  $\log K_{HSA}$ -Wert zu bestimmen, wie beispielsweise aus dem Prozentteil, der im Plasma gebunden ist (%PPB) (siehe Gleichung (2.13), [23]).

$$\log K_{HSA} = \log\left(\frac{\%PPB}{101 - \%PPB}\right) \quad (2.13)$$

Das Einsetzen des Wertes 101 basiert auf einem Trick, dass  $\log K_{HSA}$  auf den Wert 2 begrenzt wird und somit eine Kappung der Obergrenze gegeben ist, wenn eine Chemikalie zu 100 % an HSA gebunden ist. Mit der Annahme der Bildung eines binären Komplexes zwischen dem Liganden und dem HSA im Blut sowie einem Überschuss an Albumin, verglichen mit der Konzentration des Wirkstoffes, kann der  $\log K_{HSA}$ -Wert in die Affinitätskonstante  $\log K_A$ , wie in Gleichung (2.14) dargestellt [24], umgewandelt werden [23].

$$\log K_A = \log \frac{\%PPB}{101 - \%PPB} - \log c(\text{HSA}) \quad (2.14)$$

Die Affinitätskonstante  $K_A$  gibt die Neigung einer Chemikalie an HSA zu binden wieder. Damit ergibt sich für  $K_A$  ein Zusammenhang zur HSA-Konzentration.

### 2.2.2 Fett/Luft- und Fett/Blut-Verteilung

Die Fett/Luft-Verteilung  $K_{fa}$  ist definiert als die Konzentration des Stoffes in Fett geteilt durch die Konzentration des Stoffes in der Luft (siehe Gleichung (2.15)).

$$K_{fa} = \frac{c_{fat}}{c_{air}} \quad (2.15)$$

Für die Modellierung des  $K_{fa}$ -Verteilungskoeffizienten können verschiedene Methoden gewählt werden. Zum Einen kann die Modellierung mit einer linearen Gleichung nach Abraham *et al.* [12,25] erfolgen. Zum Anderen kann das Kompartiment Fett näherungsweise mit der hydrophoben Oktanolphase oder auch mit Olivenöl gleichgesetzt werden. Die Verteilung von Chemikalien zwischen Fett und Luft wird anhand der Anpassung an die Verteilung Oktanol/Luft ( $\log K_{oa}$ , siehe Kapitel 3.3.4) oder Olivenöl/Luft ( $\log K_{olive}$ , siehe Kapitel 3.3.7) vorhergesagt. Des Weiteren können die Verteilungskoeffizienten  $\log K_{mw}$  (siehe Kapitel 3.3.5) und  $\log K_{HSA}$  verwendet und mit den anderen kombiniert und als multilineare Regression an die Zielgröße  $\log K_{fa}$  angepasst werden. Für den HSA/Wasser-Verteilungskoeffizienten kann auch die Umrechnung mit  $K_{wa}$  durchgeführt werden, was in  $K_{HSA/a}$  resultiert. Auch dieser Koeffizient wurde zur Vorhersage herangezogen.

Der Fett-Blut-Verteilungskoeffizient ist im Eingang des Kapitels 2.2 (siehe Gleichung (2.2) und Abbildung 2.3, Seite 7) bereits erläutert worden und kann aus dem Fett/Luft- (2.15) und dem Blut/Luft-Verteilungskoeffizienten (2.4) errechnet werden. Dies ist in Gleichung (2.16) gezeigt.

$$K_{fb} = \frac{K_{fa}}{K_{ba}} \quad (2.16)$$

Diese Quotientenbildung ist eine Näherung, da sich die nicht reinen Phasen für Fett und Blut aus den Verteilungskoeffizienten  $K_{fa}$  und  $K_{ba}$  möglicherweise anders verhalten als die Reinphasen. Die Modellierung von  $\log K_{fa}$  und die Berechnung des Blut/Fett-Verteilungskoeffizienten werden in dieser Arbeit im späteren Verlauf gezeigt.

### 2.3 Literatur-Modelle zur Vorhersage von Verteilungskoeffizienten

Eine Vielzahl von Modellen wurde basierend auf QSAR zur Vorhersage von Verteilungskoeffizienten von VOCs entwickelt. Auch für den Blut/Luft- und Fett/Luft-Verteilungskoeffizienten für Mensch und Ratte existieren QSAR-Modelle. Da VOCs wie beispielsweise halogenierte Kohlenwasserstoffe ein hohes Risikopotenzial für Mensch und Umwelt besitzen, werden viele Studien mit diesen Stoffen vorgenommen.

#### 2.3.1 Lineare und nicht-lineare Modelle

Ein lineares Modell für den Blut/Luft-Verteilungskoeffizienten wurde von Luan *et al.* [19] genutzt, das heuristische Methode (HM) genannt wird. Die Autoren verwendeten einen Datensatz, der in ihrer Arbeit enthalten ist und aus 143 Stoffen besteht und aus Abraham *et al.* [26] übernommen wurde. Dieser Datensatz wurde in einen Trainingssatz von 114 Stoffen und einem Testsatz von 29 Substanzen geteilt.

Eine weitere lineare Methode ist die multiple linear Regression (MLR), die eine weitverbreitete statistische Technik in QSAR- und QSPR- Studien (quantitative Struktur-Eigenschaftsbeziehung, englisch: *quantitative structure property relationship*) ist. Konoz *et al.* [27] haben den Blut/Luft-Verteilungskoeffizienten von 143 VOCs vorhergesagt. Sie wählten die Deskriptoren mit Hilfe eines genetischen Algorithmus (GA) aus und spezifizierten sie mit Hilfe von einem erstellten GA-MLR Modell. GA diente hierbei als Optimierungswerkzeug und Technik als Anleitung zur Zufallssuche, beruhend auf dem Prinzip der Evolution und der natürlichen Genetik. Die von Abraham *et al.* [30] übernommenen experimentellen Daten, sind in ihrer Arbeit zugänglich. Der Datensatz wurde zufällig in drei Gruppen (Trainings-, Test- und Validierungssatz), die entsprechend 107, 18 und 18 Moleküle enthalten, aufgeteilt.

Hauptkomponentenanalyse-basierte Regressionsmethoden (PCA, englisch *Principal Component Analysis*) gehören ebenfalls zu den linearen statistischen Techniken, die häufig in QSAR-Studien genutzt werden. Diese beinhalten Hauptkomponentenregression (PCR, englisch: *Principal Component Regression*) und Partial Least Squares (PLS). Basak *et al.* [15] modellierten toxikokinetische Parameter von VOCs basierend auf berechneten molekularen Deskriptoren mit anschließender PLS-Analyse. Sie sagten Gewebe/Blut-Verteilungskoeffizienten mit Hilfe eines hierarchischen QSAR (HiQSAR)-Ansatz vorher, der topostrukturelle, topochemische, 3-dimensionale und quantenchemische Deskriptoren

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

verwendet. Hierbei ist die molekulare Struktur der einzige geforderte Eingabeparameter. Für die Vorhersage von Gewebe/Luft-Verteilungskoeffizienten nutzten sie einen zugänglichen Datensatz von 131 Chemikalien. Die Autoren verglichen die Effektivität des HiQSAR-Ansatzes mit physicochemisch-basierten Bestimmungsmethoden für Gewebe/Luft-Verteilungskoeffizienten, die experimentelle Saline/Luft- und Olivenöl/Luft-Koeffizienten nutzten [28-30].

Eine weitere Gruppe, die PLS nutzte, um HIA (englisch: *human intestinal absorption*) vorherzusagen, war Stenberg *et al.* [31]. Eine Gesamtanzahl von 674 Pharmaka, die einen großen Bereich an physiko-chemischen Eigenschaften und Permeabilität abdecken, wurde in dieser Arbeit untersucht. Der Datensatz wurde in drei Teilsätze geteilt, basierend auf dem Level des experimentellen und rechnerischen Aufwandes eines jeden Stoffes. In der Arbeit von Stenberg *et al.* [31] ist nur ein kleiner Teil der Verbindungen aufgeführt. Es gibt auch die Möglichkeit, die PLS-Technik mit einer unformativen Variableneliminierung (UVE-PLS) oder mit einer iterativen UVE-Prozedur (IVE-PLS) zu modifizieren. Dies kann zur Verdeutlichung von Bereichen, die zur Bindungsaktivität beisteuern genutzt werden, wie Polanski *et al.* [32] vorschlugen. Sie verwendeten die Applikation der CoMSA Methode, um 3D QSAR von 50 illustrierten Steroidaromatase-Inhibitoren zu analysieren. Hierzu separierten sie ihre Serie in zwei Gruppen, einen Trainingssatz bestehend aus 25 und einen Testsatz, der auch 25 Moleküle enthält.

Eine Extrapolation zwischen den verschiedenen Spezies von inhalativen Pharmakokinetiken wird von Beliveau *et al.* [33] angewandt. Das Modell bestimmt Fett-, Leber-, Muskel- und Blut-zu-Luft-Verteilungskoeffizienten in Ratte und Mensch. Werden diese einzelnen Verteilungskoeffizienten zusammengefasst zu einem allgemeinen, kann dieser als Gewebe/Luft-Verteilungskoeffizient geschrieben werden ( $K_{ta}$ ). Dieser wird berechnet mit Gleichung (2.17),

$$K_{ta} = F_{nL} \cdot K_{oil} + F_w \cdot K_{wa} + f_b \cdot F_p \cdot K_{pa} \quad (2.17)$$

wobei  $F_{nL}$  die Volumenfraktion von neutralen Lipiden ist,  $F_w$  den Gehalt der Wasseräquivalente,  $F_p$  den Proteingehalt und  $f_b$  die Fraktion des Gesamtproteins im Verteilungsprozess beschreibt. Zusätzlich werden die Verteilungskoeffizienten  $K_{oil}$  (pflanzliches Öl/Luft),  $K_{wa}$  und  $K_{pa}$  bei physiologischen Temperaturen verwendet. Die

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

benötigten logarithmischen Verteilungskoeffizienten werden mithilfe einfacher Atombeiträge über ein Fragmentmodell, das ebenfalls aus dieser Arbeitsgruppe stammt [34], berechnet. Die Veröffentlichung liefert vordefinierte Sätze von Parametern für Mensch und Ratte. Der Anwendungsbereich des Modells ist für VOCs definiert, wofür Fragmente für Alkane, Alkene und Halogene, außer Iod und Aromaten, zugänglich sind. Die Autoren verwendeten experimentelle Daten für 46 VOCs von Gargas *et al.* [35]. Nach Ersetzen des Zielparameters Gewebe/Luft-Verteilung durch Blut/Luft-Verteilung kann das Modell von Beliveau *et al.* [33] mit dem linearen physiologischen Ansatz aus Kapitel 2.2.1 (Seite 9ff), der ebenfalls Fraktionsanteile von Wasser, Lipiden und Proteinen enthält, und mit einem  $K_{fa}$ -Modell, das die gleichen Verteilungskoeffizienten in einer linearen Regression zur Modellierung verwendet, verglichen werden. Die Gleichung für den Blut/Luft-Verteilungskoeffizienten für Menschen mit konkreten Werten ist in Gleichung (2.18) gezeigt.

$$K_{ba} = 0,004 \cdot K_{oil} + 0,8217 \cdot K_{wa} + 0,1743 \cdot K_{pa} \quad (2.18)$$

Wird nun Fett als Gewebe betrachtet, kann dieses Literaturmodell mit den Modellen aus Kapitel 4.3.2 verglichen werden. Mit eingesetzten Parametern für Menschen und Vernachlässigung der Proteinbindung im Fett ergibt sich Gleichung (2.19).

$$K_{fa} = 0,7986 \cdot K_{oil} + 0,1514 \cdot K_{wa} \quad (2.19)$$

Beliveau *et al.* [33] bestimmten Verteilungskoeffizienten zusätzlich in vielen verschiedenen Geweben von Ratten (0,9 % Saline, Olivenöl, Blut, Leber, Muskeln und Fettgewebe) und beschrieben eine Methode, die eine modifizierte Form der Ampullen-Äquilibrierungs-Technik von Sato *et al.* [36] ist. Sie präsentieren die erhaltenen Ergebnisse für 55 niedermolekulare VOCs. Die Beziehungen zwischen Ratten Blut/Luft- und Menschen Blut/Luft-Verteilung sowie Olivenöl/Luft- und Fett/Luft-Verteilungskoeffizienten wurden mit linearer Regressionsanalyse untersucht. Die Modellierung der verschiedenen Gewebe/Luft- und Blut/Luft-Koeffizienten bezogen auf ihren Saline und Öl/Luft-Wert wurden mit linearer Regressionsanalyse untersucht. Diese Information über die Löslichkeit für eine große Vielzahl von flüchtigen Chemikalien ist notwendig, um PBPK-Modelle zu entwickeln.

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

Hansch *et al.* [37] schlagen vor, eine Prüfung der sogenannten Meyer-Overton-Hypothese vorzunehmen. Diese Hypothese korreliert die Fettlöslichkeit eines Narkosestoffs mit der Wirksamkeit ( $p^{-1}$ ,  $p$  ist der Narkosedruck) und sagt aus, dass eine narkotische Wirkung einsetzt, sobald genügend Moleküle des Narkosestoffs in den Lipidmembranen der Zelle gelöst sind. Die Autoren fanden eine ziemlich große Lücke in der bereits vorhandenen Datenbank eines einzelnen Systems, was sie dazu veranlasste, den Oktanol/Wasser-Verteilungskoeffizienten von 32 flüchtigen Substanzen zu messen. Die Narkosewirkung dieser Stoffe wurde untersucht.

Meulenberg *et al.* [38] legen nahe, dass die Verteilung von nichtreaktiven Chemikalien im Gewebe vollständig durch Lipophilie- und Hydrophiliewechselwirkungen der Verbindungen mit den Gewebestandteilen bestimmt wird. Sie haben Verteilungskoeffizienten für 137 VOCs in Olivenöl und Saline zusammengestellt sowie Verteilungskoeffizienten für Ratten- und Menschenblut, Fett, Gehirn, Leber und Nierengewebe veröffentlicht. Mit einer bilinearen Regressionsanalyse können die Verteilungskoeffizienten von VOCs in menschlichem Gewebe gut durch eine Linearkombination von  $K_{\text{Öl/Luft}}$  und  $K_{\text{Saline/Luft}}$  mit gewebespezifischen Regressionskoeffizienten beschrieben werden.

Paterson *et al.* [39] entwickelten einen einfachen, neuen Korrelationsansatz und wendeten ihn auf 24 VOCs an. Mit diesem können sie Gewebe/Blut-, Gewebe/Luft- und Blut/Luft-Verteilungskoeffizienten aus der Wasserlöslichkeit und dem Dampfdruck vorhersagen. Der Vorteil dieser Methode ist, dass sofort zu sehen ist, wie ein Verteilungskoeffizient durch jede beteiligte Phase beeinflusst wird. Mit diesem Ansatz ist es möglich, auf glaubhafte Verteilungskoeffizienten allein aus der Wasserlöslichkeit zu folgern.

Poulin *et al.* [16,40] wendeten einen Algorithmus zur Bestimmung der Gewebe/Luft-Verteilungskoeffizienten für VOCs an, der auf der Gesamtvolumenfraktion von Wasser und Phospholipiden in Gewebe und Plasma sowie auf der ungebundenen Fraktion in fetthaltigem Gewebe und Plasma basiert. Dieser Algorithmus wird verwendet, um den Gewebe/Luft-Verteilungskoeffizienten von 45 VOCs in Ratten zu bestimmen. Die Verbindungen sind vor allem Ketone, Alkohole, Acetatester, Alkane, Halogenalkane, aromatische Kohlenwasserstoffe und Diethylether. Die abhängige Variable  $K_{\text{ta}}$  wird mit Hilfe oben genannter Fraktionen beschrieben.

Varnek *et al.* [41] führten einen neuen theoretischen Ansatz ein, bei dem die unbekannte Zielgröße mit ähnlichen Größen oder Eigenschaften verglichen wird und der Wert für die

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

unbekannte Variable aus den bereits verfügbaren experimentellen Daten vorhergesagt werden. Dieser sogenannte induktive Ansatz ist eine Art Wissenstransfer und eine gute Alternative zur teuren und zeitintensiven Beschaffung von neuen experimentellen Daten. Die Arbeit von Varnek *et al.* [41] zeigt die Anwendung eines induktiven Ansatzes in QSAR-Studien, was in der Literatur auch als induktive Transfer-Methode bezeichnet wird. Durch Nutzen von multitask learning (MTL) und sequentially feature nets (FN) Techniken werden mit Struktur-Eigenschafts-Modellierungen sogar für kleine und komplexe Datensätze Ergebnisse erhalten. Die Autoren sammelten Mensch- und Rattendaten und verwendeten insgesamt 141 organische Verbindungen, um Vorhersagemodelle für Gewebe/Luft-Verteilungskoeffizienten zu erstellen.

Zhang *et al.* [38] nutzten ein nicht-lineares Regressionsmodell (NLR), um Gewebe/Blut-Verteilungskoeffizienten vorherzusagen. Der Datensatz besteht aus 80 in deren Arbeit zugänglichen Verbindungen und dazugehörigen Koeffizienten, die wahllos in zwei Datensets geteilt wurden: ein Trainingssatz mit 67 Verbindungen und ein Testsatz mit 13 Stoffen. Ihre Arbeit kann als eine Erweiterung der Hansch-Gleichung [43] angesehen werden, die zur Modellierung vieler Aspekte von chemisch-biologischen Wechselwirkungen sowohl in der Medikamenten- und Pestizidforschung als auch in vielen Gebieten der Toxikologie hilfreich ist.

Bekannte Ansätze für nicht-lineare Methoden zur Vorhersage von PBPK-relevanten Verteilungskoeffizienten sind künstliche neuronale Netze (KNN, englisch: *artificial neural networks* (ANNs)), die große Popularität in QSAR-Studien erlangten. Konoz *et al.* [27] nutzten ANNs zur Vorhersage von Blut/Luft-Verteilungskoeffizienten von 143 VOCs. Der Datensatz wurde von Abraham *et al.* [26] übernommen. Auch Wessel *et al.* [44] nutzte KNN zur HIA-Bestimmung von Medikamenten. Ihr Datensatz enthält 86 Verbindungen, die mit ihren Werten und Koeffizienten öffentlich zugänglich sind. Zur Vorhersage von menschlichen Blut/Luft-Verteilungskoeffizienten ( $K_{ba}$ ) von VOCs, wendeten Luan *et al.* [19] nicht-lineare radiale Basisfunktion neuronale Netzwerke an (RBFNNs). Der Datensatz wurde ebenfalls von Abraham *et al.* [26] übernommen. Er wurde in einen Trainingssatz von 114 und einen Testsatz von 29 Verbindungen geteilt.

Ein weiteres nicht-lineares Regressionsmodell (NLR) schlugen Deretey *et al.* [45] zur Bestimmung von HIA mit einer Datensatzgröße von 124 Molekülen basierend auf einen bereits erwähnten Datensatz von Zhao *et al.* [46] vor.

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

Einige weitere Ansätze für QSPR-Studien wurden veröffentlicht, beispielsweise der Sphärenausgrenzungsalgorithmus (SEA, englisch: *sphere exclusion algorithm*), der von Gunturi *et al.* [47] zur Modellierung von HIA genutzt wurde. Die zugänglichen HIA Daten von 175 Arzneimitteln in ihrer Arbeit wurden aus verschiedenen Literaturstellen zusammengestellt. Gunturi *et al.* [47] wendeten die Methodik der Nächste-Nachbarn-Klassifikation (kNN, englisch: *k-nearest neighbours*) an. Dies ist eine Methode zur Klassifizierung von Objekten basierend auf nächsten Trainingsbeispielen in einem Merkmalsraum. Sie nutzten kNN zusammen mit einem genetischen Algorithmus (kNN-QSAR-GA), um signifikante und unabhängige Deskriptoren auszuwählen. Raevsky *et al.* [52] sagten auch HIA mit dem kNN Ansatz vorher, wobei 38 neutrale Verbindungen, 29 Kationen, 20 Anionen und 13 Zwitterionen zum Einsatz kamen. Alle 100 Verbindungen sind in dieser Veröffentlichung zugänglich.

Eine weitere mathematische Methode zur Vorhersage von HIA ist die sigmoidale Kurvenanpassung nach Boltzmann, wie es in [51,53] beschrieben ist. Auch die Ridge Regression (RR), eine Form der MLR, ausgeführt von Basak *et al.* [32] ist eine geeignete Regressionsmethode, um menschliche Gewebe/Luft-Verteilungskoeffizienten vorherzusagen. In ihrer Studie nutzten die Autoren 131 Chemikalien, die in der Veröffentlichung gelistet sind und verglichen ihre Ergebnisse mit denen von Hoerl *et al.* [54].

Mit diesen Modellen werden unterschiedliche Zielgrößen, wie verschiedene Verteilungskoeffizienten oder HIA, vorhergesagt. Fast alle dieser gezeigten Modelle besitzen kleinere Datensätze oder sind auf deutlich komplexere Strukturen, die jedoch keine VOCs sind, anzuwenden als die in dieser Arbeit vorgestellten Modelle. Die Herangehensweisen an die Modellierung sind sehr unterschiedlich und es werden zum Teil neue Modellierungswege erarbeitet. Dadurch erweitert sich die Auswahl der vorhandenen Modelle, die je nach Anforderung an die Zielgröße und Datensatzgröße entsprechend gewählt werden können.

### 2.3.2 LSER-Modelle

Eines der bedeutendsten Modelle, welches einen wichtigen generellen Ansatz zur Vorhersage von Verteilungskoeffizienten  $K$  verwendet, ist gegenwärtig als Abraham-Gleichung (Gleichungen (2.20) und (2.21)) bekannt [12]. Es ist ein mechanistisch orientiertes Modell, das auf linearen Solvatations-Energie-Beziehungen basiert.

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

$$\log K = c + eE + sS + aA + bB + vV \quad (2.20)$$

$$\log K = c + eE + sS + aA + bB + lL \quad (2.21)$$

Die fünf Terme dieser Gleichungen beschreiben intermolekulare Wechselwirkungen in grundlegender Form. Dabei ist jeder Term das Produkt eines Stoffdeskriptors (Großbuchstabe) und eines Phasenparameters (Kleinbuchstabe).  $E$  ist die Exzess-Molrefraktion,  $S$  beschreibt die Polarität und Polarisierbarkeit,  $A$  und  $B$  repräsentieren die Wasserstoffbrücken-Donor- und Akzeptorstärke der Verbindung und  $V$  ist das charakteristische Volumen nach McGowan [51]. Der Parameter  $B$  kann in  $B^H$  und  $B^O$  unterschieden werden.  $B^H$  definiert die Wasserstoffbrücken-Basizität für Verteilungen in wasserfreien Systemen wie reines Oktanol im Falle Oktanol/Luft-Verteilung.  $B^O$  beschreibt die Wasserstoffbrücken-Basizität für Verteilungen in wässrigen Systemen, z. B. wasser-gesättigtes Oktanol im Falle der Oktanol/Wasser-Verteilung [52].  $B^O$  und  $B^H$  werden oft zusammengefasst und mit dem allgemeinen Symbol  $B$  angegeben. Die korrespondierenden Phasenparameter sind als  $e$ ,  $s$ ,  $a$ ,  $b$  und  $v$  gegeben. Der Term  $c$  steht für die Regressionskonstante. Der Parameter  $e$  beschreibt die Tendenz der Phase, durch  $\pi$ - und n-Elektronen zu wechselwirken und  $s$  gibt die Polarität des Lösungsmittels wieder. Phasenparameter  $a$  ist die Wasserstoffbrücken-Basizität der Phase, der wiedergibt, ob eine basische Phase mit einer aciden gelösten Substanz eine Wechselwirkung eingeht und  $b$  ist die Wasserstoffbrücken-Acidität der Phase. Der  $vV$  Term in Gleichung (2.20) enthält sowohl Unterschiede in Hohlraum-Effekten als auch Unterschiede in allgemeinen Dispersionswechselwirkungen. Der  $v$ -Parameter ist somit ein Maß für den Unterschied zwischen Hydrophobie oder Lipophilie der zwei Phasen [53]. Gleichung (2.20) wird normalerweise für die Verteilung in flüssig/flüssig-Systemen gebraucht. Für Gas/flüssig-Verteilungen sollte der Term  $vV$  durch  $lL$  (2.21) ausgetauscht werden, wobei  $L$  der logarithmische Hexadecan/Luft-Verteilungskoeffizient bei 20 °C ist und  $l$  der korrespondierende Phasenparameter. Ansätze vom Typ Abraham funktionieren normalerweise gut mit experimentellen Deskriptoren. Jedoch ist die Vorhersage der Koeffizienten (außer  $V$ ) aus der chemischen Struktur [54,55] immer noch ein Problem [56].

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

Ein LSER-Modell nach Gleichung (2.21) zur Bestimmung von Blut/Luft-Verteilungskoeffizienten für Mensch und Ratte wurde von Abraham *et al.* [26] eingeführt (Gleichung (2.22)).

$$\log K_{ba} = -1,069 + 0,456 E + 1,083 S + 3,738 A + 2,580 B^H + 0,376 L \quad (2.22)$$
$$n = 196, r^2 = 0,938, rms = 0,319$$

Die Statistikparameter  $n$ ,  $r^2$  und  $rms$  sind die Anzahl der Daten, das Bestimmtheitsmaß und der absolute mittlere Fehler  $rms$  (englisch: *root mean square-error*).

Auch für den Fett/Luft-Verteilungskoeffizienten wurde eine LSER-Gleichung von Abraham und Ibrahim [25] veröffentlicht (Gleichung (2.23)).

$$\log K_{fa} = -0,052 + 0,051 E + 0,728 S + 1,783 A + 0,332 B + 0,743 L \quad (2.23)$$
$$n = 129, r^2 = 0,958, sd = 0,194$$

Die Datensätze dieser beiden Modelle sind etwas größer als die in dieser Arbeit verwendeten, da neben VOCs auch anorganische Stoffe und Verbindungen enthalten sind, auf denen in dieser Arbeit nicht der Fokus liegt. Die Datensätze wurden bezüglich VOCs etwas erweitert, wodurch eine ähnliche Anzahl an Stoffen resultiert. Die LSER-Gleichungen wurden jeweils neu angepasst und anschließend mit denen aus der Literatur verglichen.

Sprunger *et al.* [57] erweiterten den Ansatz für Blut/Luft durch Hinzufügen eines linearen Indikatorterms, der den Einfluss verschiedener Spezies beschreibt. Die damit erweiterte Abraham-Gleichung gilt für Mensch, Ratte, Maus und andere Tiere, bei der insgesamt 358 experimentelle Werte für die verschiedenen Spezies eingehen (Gleichung (2.24)).

$$\log K_{ba} = -1,167 + 0,487 E + 1,093 S + 3,636 A + 2,771 B + 0,366 L + i I \quad (2.24)$$
$$n = 358, r^2 = 0,949, rms = 0,293$$

Die Konstante  $i$  wird als Korrektur für das Mensch-Modell ( $i = 0$ ) zu den anderen Spezies verwendet.  $I$  ist die Speziesindikatorvariable, die an das Basis-Abraham-Modell addiert wird und spezifisch für jede Spezies ist. Für Ratten beträgt  $i$  in diesem Gesamtmodell 0,165. Im Modell für Mensch und Ratte ist dieser Wert leicht höher, nämlich  $i = 0,201$  (siehe Gleichung

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

(2.25)). Dieser fassbare, aber dennoch kleine Unterschied bekräftigt den Ansatz, Mensch- und Ratten-Daten zu einem Modell zusammenzufassen.

$$\log K_{ba} = -1,164 + 0,434 E + 1,120 S + 3,070 A + 2,485 B + 0,368 L + 0,201 I \quad (2.25)$$

$n = 146, r^2 = 0,951, rms = 0,270$

Andere Abraham-Modelle machen es möglich, HIA [46], Hautpermeation und Haut/Wasser-Verteilung bei 37 °C vorausszusagen [58]. Das Modell von Zhao *et al.* [46] zur Bestimmung von HIA war auf 241 Arzneistoffe trainiert, wobei alle Chemikalien und zugehörigen Daten in deren Arbeit aufgelistet sind. Das Abraham und Martins-Modell [58] verwendet Daten für 119 gelöste Substanzen aus der Literatur.

Das bedeutende Abraham-Modell wird mit experimentellen Deskriptoren nachgerechnet und im Falle von Blut/Luft mit den Ergebnissen von Sprunger *et al.* [57] verglichen. Da aber ein Modell nur auf berechneten Größen basieren und ganz ohne experimentelle Werte auskommen soll, ist eine Überarbeitung dieser Modelle sinnvoll.

### 2.4 Dampfdruck

Der Dampfdruck beschreibt den Phasenübergang von einem festen oder flüssigen Reinstoff in den gasförmigen Aggregatzustand. Das kann in einem, von flüssig zu gasförmig, oder in zwei Schritten, von fest über flüssig zu gasförmig, erfolgen. Die Sublimation wird nicht weiter berücksichtigt, da sie nur einen kleinen Teil der verfügbaren Substanzen betrifft und die Modellierung durch die zusätzlich auftretende Schmelzwärme an Komplexität gewinnt.

Ein Modellierungsansatz zum Dampfdruck wird verwendet, um über ihn die Berechnung der Henry-Konstante und somit auch des Wasser/Luft-Verteilungskoeffizienten zu ermöglichen.

#### 2.4.1 Zusammenhang des Dampfdrucks mit der Henry-Konstante

Die Verteilung zwischen Luft und Wasser kann mit mehreren Größen beschrieben werden, die mathematisch miteinander verbunden sind. Die Henry-Konstante  $H$  charakterisiert das Gleichgewicht von Stoffen zwischen Luft und Wasser [52]. Eine fundamentale thermodynamische Eigenschaft ist der Henry-Koeffizient, der die Druckabhängigkeit der Löslichkeit von idealen Gasen in einer idealen Lösung beschreibt und in Druckeinheiten ausgedrückt wird [59]. In den Umweltwissenschaften sind auch andere Definitionen in

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

Gebrauch. In seiner dimensionslosen Form ( $K_{aw}$ ) kann dieser als das Verhältnis der Konzentration des Stoffes in der Luft  $c_a$  zur Konzentration des Stoffes im Wasser  $c_w$  definiert werden:

$$K_{aw} = \frac{c_a}{c_w} \quad (2.26)$$

Dieser Verteilungskoeffizient wird auch die dimensionslose Henry-Konstante genannt. Im ökologischen Zusammenhang werden Vorhersagemethoden in der Regel mit  $\log K_{aw}$ -Werten kalibriert [59]. Obwohl dieser Parameter nicht über eine formale Einheit verfügt, müssen die zugrunde liegenden Konzentrationseinheiten beachtet werden, da ein  $K_{aw}$ , der auf Stoffmengenanteilen basiert, sich sowohl relativ als auch absolut von einem  $K_{aw}$ , der auf molaren Konzentrationen basiert, unterscheidet. Für die Gasphase im Gleichgewicht mit einer flüssigen Lösung kann näherungsweise das ideale Gasgesetz für den jeweiligen Stoff von Interesse angewendet werden [52]:

$$P = \frac{n}{V} RT = c_a RT \quad (2.27)$$

Hier ist  $P$  der Druck,  $n$  die Stoffmenge,  $V$  das Volumen,  $R$  die universelle Gaskonstante und  $T$  die Temperatur. Wenn die Konzentration in Luft dem Partialdruck in Luft entspricht, führt dieser Ausdruck über die Zustandsgleichung für ideale Gase zu:

$$H = K_{aw} RT = \frac{P}{c_w} \quad (2.28)$$

Diese Definition basiert auf dem Verhältnis des Partialdruckes zur Konzentration des Stoffes in Wasser  $c_w$  und wird üblicherweise als Henry-Konstante angesehen. Aus dieser Gleichung (2.28) ist die Temperaturabhängigkeit des Verhältnisses zwischen dem Verteilungskoeffizient Luft/Wasser und der Henrykonstante bekannt. Zusätzlich existiert eine intrinsische Temperaturabhängigkeit für den Verteilungskoeffizienten selbst. Beide können in der gleichen Art modelliert werden [59].

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

Viele organische Chemikalien sind nicht vollkommen mit Wasser mischbar. Die Konzentration in Wasser am Sättigungspunkt ist die Wasserlöslichkeit  $S_w$  [59]. Wenn die Konzentration eines Stoffes in Wasser die Löslichkeitsgrenze erreicht, bildet sich durch weitere Zugabe der Substanz eine separate Phase, welche die Phase der reinen Substanz ist. Da der gelöste Zustand im Gleichgewicht mit der Gasphase und der reinen Substanzphase ist, befindet sich auch die letztere mit der Gasphase im Gleichgewicht. Damit ist bei der Konzentration, an der die Löslichkeitsgrenze erreicht wird, die sich bildende Substanzphase mit der gelösten und mit der Gasphase im Gleichgewicht. Infolgedessen ist der Partialdruck  $P$  bei Sättigung in der Gasphase gleich dem Dampfdruck  $P_v$  der reinen Verbindung [60]. Daraus folgt, dass die Henry-Konstante auch als der Quotient auf Dampfdruck und Wasserlöslichkeit (Gleichung (2.29)) definiert werden kann.

$$H = \frac{P_v}{S_w} \quad (2.29)$$

Der Dampfdruck beschreibt das Gleichgewicht eines Stoffes zwischen seiner Fest- oder Flüssigphase und seiner Gasphase. Im Gegensatz dazu definiert die Henry-Konstante das Gleichgewicht eines Stoffes im Kompartiment Wasser mit dem Anteil des Stoffes in der Gasphase.

Eine genaue Betrachtung zeigt, dass Gleichung (2.29) in der Tat eine angenäherte Beziehung für die Henry-Konstante ist. Die Näherung basiert auf der Annahme, dass sich die Gasphase nach dem idealen Gasgesetz verhält und die gelöste Konzentration so gering ist, dass die Lösung ideal verdünnt vorliegt. Dies kann im Konflikt mit der Gleichsetzung von  $c_w$  und  $S_w$  stehen. Die sich bildende separate Substanzphase ist tatsächlich eine reine Stoffphase, wenn die Stoffmenge in Lösung größer als  $S_w$  ist. Dennoch wird Gleichung (2.29) verwendet, um  $H$  zu bestimmen, sofern keine experimentellen Daten vorhanden sind [52]. In diesem Fall werden vorzugsweise experimentelle Daten für  $P_v$  und  $S_w$  verwendet.

Modelle zur Vorhersage der Henry-Konstante aus der molekularen Struktur nutzen normalerweise den dekadischen Logarithmus des dimensionslosen  $K_{aw}$  als Zielgröße. Experimentelle Werte für  $H$  können in  $K_{aw}$  über die Gleichung (2.28) umgerechnet werden. Für Realgase gibt es im Vergleich zum idealen Gas eine Restriktion. Hierbei kann der Partialdruck den Atmosphärendruck nicht übersteigen. Als Konsequenz daraus ist ihre

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

gemessene Wasserlöslichkeit unterhalb der Sättigung in Lösung bei einem Dampfdruck, der dem Atmosphärendruck entspricht [59]. Aus einer streng thermodynamischen Sicht, ist Gleichung (2.28) nur bei unendlicher Verdünnung gültig. Daraus ergibt sich, dass die Henry-Konstante keine Konstante ist, sondern auf den Grenzwert einer molaren Fraktion von 0 zugeht und dass die Gleichung (2.29) immer eine Näherung ist, da eine Sättigung nie einer unendlichen Verdünnung entspricht. Allerdings sind diese Überlegungen von geringerer Bedeutung, da die Genauigkeit der Vorhersage für die Umweltmodellierung nicht relevant bzw. vernachlässigbar ist. Der Dampfdruck besitzt eine hohe PBPK-Relevanz, da er direkt proportional zum Henry-Koeffizienten ist (Gleichung (2.29)), dieser wiederum leicht in den Verteilungskoeffizienten  $K_{wa}$  umgerechnet werden kann (Gleichung (2.28)).

### 2.4.2 Vorhersagemodelle für den Dampfdruck

Der Dampfdruck ist eine stark temperaturabhängige Größe. Dennoch wird der Dampfdruck bei 25 °C häufig als Deskriptor in Modellen genutzt. Für PBPK-Modelle sollte die angewendete Temperatur allerdings 37 °C betragen, da sie sich auf physiologische Bedingungen beziehen. Die Temperaturabhängigkeit wird durch die Clausius-Clapeyron-Gleichung (2.30) beschrieben, bei der der Dampfdruck mit der Temperatur im Zusammenhang steht.

$$\frac{d(\ln P_v)}{dT} = \frac{\Delta H}{RT^2} \quad (2.30)$$

Die Größe  $\Delta H$  entspricht entweder der Verdampfungsenthalpie bei Übergängen vom flüssigen in den gasförmigen Aggregatzustand oder der Sublimationsenthalpie bei Übergängen vom festen Zustand in den gasförmigen. Die Abweichung des betrachteten gasförmigen Stoffes vom Verhalten eines idealen Gases kann durch die Zuhilfenahme des Kompressibilitätsunterschieds der zwei Phasen verrechnet werden (Gleichung (2.31)). Dieser Parameter besitzt einen Wert zwischen 1,00 und 0,91 [61].

$$\Delta Z = \frac{P_v \Delta V}{RT} \quad (2.31)$$

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

Dies führt zu  $(\Delta Z \cdot RT^2)$  als Nenner in Gleichung (2.30) anstatt  $(RT^2)$ . Durch Integration der Clausius-Clapeyron-Gleichung von einer beliebigen Referenz-Temperatur (z. B. Siedepunkt  $T_b$ ) auf die Umgebungstemperatur von Interesse (z. B.  $T = 37 \text{ °C} = 310 \text{ K}$ ) ergibt sich die Basis für eine wichtige Klasse von Vorhersage-Methoden für  $P_v$  [60]. Eine große Anzahl an Dampfdruck-Vorhersagemodelle nutzt den Siedepunkt bei Atmosphärendruck, einige von ihnen verwenden zusätzlich den Schmelzpunkt.

Das Modell von Sepassi *et al.* [62] gilt als das Genaueste und wird somit für den allgemeinen Gebrauch empfohlen [60]. Allerdings hängt die Zuverlässigkeit von der richtigen Einschätzung der Rotationssymmetrie ab und damit in der Regel von einer hinreichend korrekten Geometrie. Für feste und flüssige Stoffe besteht die Beschränkung, dass der Siedepunkt größer sein muss als die Systemtemperatur zur Berechnung des Dampfdruckes (z. B.  $37 \text{ °C}$ ). Neben der Rotationssymmetrie werden als weitere Parameter die effektive Anzahl der Torsionsbindungen  $\tau$  und die Anzahl an Wasserstoffbrücken (HBN) verwendet, für die jeweils Modelle existieren.

Ein weiteres Modell ist eine von C. F. Grain modifizierte Gleichung, die der französische Forscher Ch. Antoine 1888 zuerst vorlegte. Dieses empirische Antoine-basierte Modell von Grain [63] ist temperaturabhängig und gibt als Zielgröße den natürlichen Logarithmus des Dampfdruckes aus (siehe Gleichung (2.32)).

$$\ln P_v = A - \frac{B}{T - C} \quad (2.32)$$

Die Konstanten  $A$ ,  $B$  und  $C$  sind in den Gleichungen (2.33), (2.34) und (2.35) dargestellt.

$$A = \frac{B}{T_b - C} + \text{korr} \quad (2.33)$$

(wenn  $P = 1 \text{ atm}$  bei  $T = T_b$ )

$$B = \frac{\Delta H}{\Delta Z} RT(T_b - C)^2 \quad (2.34)$$

(von der Clausius-Clapeyron Gleichung (2.30) und wenn  $P = 1 \text{ atm}$  bei  $T = T_b$ )

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

$$C = -18 + 0,19 T_b \quad (2.35)$$

Die Konstante  $C$  (2.35) wird über die Regel von Thomson [64] berechnet. Die Parameter zur Berechnung der Konstanten sind ein Korrekturfaktor ( $\text{korr} = 1,01 \cdot 10^5$ ), der Kompressibilitätsunterschied ( $\Delta Z = 0,97$  ([61,65], Definition siehe Gleichung (2.31)) und die Verdampfungsenthalpie  $\Delta H$ , die mit Gleichung (2.36) berechnet wird.

$$\Delta H = \frac{\Delta S_b}{R} = K_f (\text{konst} + R \ln T_b) \quad (2.36)$$

Die Größe  $\Delta S_b$  ist die Entropie der Verdampfung. Weitere Parameter sind die Konstante ( $\text{konst} = 8,75 \text{ kal/mol}\cdot\text{K} = 36,6345 \text{ J/mol}\cdot\text{K}$ ), die Temperatur ( $T$  in K), die ideale Gaskonstante ( $R = 1,987 \text{ kal/mol}\cdot\text{K} = 8,314 \text{ J/mol}\cdot\text{K}$ ) und die Fishtine-Konstante ( $K_f$ ), die das Dipolmoment von polaren und unpolaren Stoffen berücksichtigt [66].  $K_f$  ist ein diskreter Tabellenwert für aliphatische, alicyclische und polare wasserstoffbrückengebundene aromatische Stoffe. Für Verbindungen mit einem aromatischen Ring wie das polare Benzol, Pyridin, Furan und Thiofuran-Derivate wird  $K_f$  aus dem Dipolmoment nach Gleichung (2.37) berechnet. Bei Verbindungen mit zwei aromatischen Ringen, beispielsweise Naphthalin, wird Gleichung (2.38) verwendet. Das Dipolmoment wird durch Substituentenbeiträge errechnet [65,66].

$$K_f = 1 + \frac{2\mu}{100} \quad (2.37)$$

$$K_f = 1 + \frac{\mu}{100} \quad (2.38)$$

Das Antoine-basierte Modell von C. F. Grain (Gleichung (2.32)) ist nur für die Dampfdruckbestimmung von Stoffen anwendbar, die bei der ausgewählten Systemtemperatur in flüssigem oder gasförmigem Zustand vorliegen [63]. Im Falle von festen Stoffen wird eine Korrektur angewendet (siehe Gleichung (2.39)).

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

$$\ln(P_v(\text{fest})) = \ln(P_v(\text{unterkühlte Flüssigkeit})) + 6,81(1 - T_m) \quad (2.39)$$

Das sogenannte Grain-Antoine-Modell kann auch weiter modifiziert werden, wie es von Sage & Sage [67] beschrieben ist (siehe Gleichung (2.40)). Hierbei spielen die Effekte der van-der-Waals-, Dipol-Dipol- und Wasserstoffbrückenbindungswechselwirkungen eine Rolle.

$$S_v = K_f \cdot \ln(82,06 \cdot T_b) \quad (2.40)$$

Die Berechnung von  $K_f$  erfolgt auch hier nach Fishtine [65,66] wie in Gleichungen (2.37) und (2.38) gezeigt.

Ein weiteres Modell zur Berechnung des Dampfdruckes ist das Grain-Modell basierend auf Watson [63], welches für feste und flüssige organische Chemikalien anwendbar ist (Gleichungen (2.41), (2.42) und (2.43)).

$$\ln P_v = \frac{S_v}{\Delta Z \cdot R} \cdot f(T_p, m) + \ln(\text{korr}) \quad (2.41)$$

$$T_p = \frac{T}{T_b} \quad (2.42)$$

$$f = 1 - \frac{(3 - 2T_p)^m}{T_p} - 2m(3 - 2T_p)^{m-1} \ln T_p \quad (2.43)$$

Bei Flüssigkeiten ist  $m = 0,19$  und bei Feststoffen gilt folgende Regel. Ist  $T_p$  größer als 0,6, erhält  $m$  einen Wert von 0,36, liegt  $T_p$  unter 0,5, beträgt  $m = 1,19$ . Für  $T_p$  zwischen 0,5 und 0,6 hat  $m$  den Wert 0,8. Alternativ dazu kann  $m$  auch zwischen 0,38 und 1,19 interpoliert werden, was mit einer Modifikation durch Sage & Sage [67] geschieht, bei der  $m$  in Abhängigkeit von  $T_p$  unterschiedliche Werte einnimmt (Gleichung (2.44)):

$$m = 0,4133 - 0,2575 \cdot T_p \quad (2.44)$$

## Verteilungskoeffizienten, Dampfdruck und Modelliermethoden

Die Parameter  $k_{\text{orr}}$ ,  $\Delta Z$  und  $S_v$  in Gleichung (2.41) sind mit den Parametern des Grain-Antoine-Modells gleich ((2.33), (2.31) und (2.36)). Für Feststoffe wird die gleiche Funktion für  $f$  verwendet, wie für Flüssigkeiten (2.43). Da das Grain-Watson-Modell auch für Feststoffe anwendbar ist, benötigt es keine Fugazitätskorrektur durch den Schmelzpunkt. Trotzdem muss festgelegt werden, ob ein Stoff bei der Systemtemperatur fest oder flüssig vorliegt.

Obwohl mit dem Sepassi-Modell bessere Ergebnisse erzielt werden als mit den Grain-Methoden, ist seine Anwendungsbreite deutlich eingeschränkter und verlässliche Ergebnisse werden nur mit einer eindeutigen Molekülsymmetrie erreicht. Deshalb ist es nicht das geeignetste Modell zur Dampfdruckbestimmung.

Ein weiteres Dampfdruck-Modell ist von Mackay *et al.* [68], welches für feste und flüssige Chemikalien angewendet werden kann. Es wird durch Integration der Clausius-Clapeyron-Gleichung (2.30) mit einigen Näherungen für die Temperaturabhängigkeit von  $\Delta H$  erhalten (siehe Gleichung (2.45)).

$$\log P_v = -4,6 \cdot \left(\frac{T_b}{T} - 1\right) - 2,95 \cdot \left(\frac{T_m}{T} - 1\right) \quad (2.45)$$

In dieser Gleichung wurde die Regel von Trouton [69] (2.46) angewendet, um die Siedentropie zu bestimmen.

$$\Delta S_b = 88 \frac{\text{J}}{\text{mol} \cdot \text{K}} \Rightarrow \frac{\Delta S_b}{2,3R} = 4,6 \quad (2.46)$$

Der Übergang vom flüssigen in den gasförmigen Zustand wird mit dem Siedepunkt-Term beschrieben. Der Schmelzpunkt-Term wird nur für Feststoffe benötigt (bei  $T_m > T$ ) und wird somit als Indikatorterm verwendet. Von allen der hier vorliegenden Methoden ist das Mackay-Modell das einzige, das abgesehen von  $T_b$  und  $T_m$  keine Verwendung von strukturellen Informationen der Verbindungen macht. Dieses Modell wurde für Kohlenwasserstoffe und halogenierte Stoffen kalibriert. Daher ist zu berücksichtigen, dass die Methode nur bedingt auf polarere Verbindungen, die nicht Teil des Trainingsatzes waren, angewendet werden kann [68].

### 3 Material und Methoden

Nachfolgend werden Modellierungsmethoden und -ansätze, die zur Vorhersage des Blut/Luft-Koeffizienten verwendet werden, aufgeführt und erläutert. Hierzu zählen das Fragmentmodell und die Verwendung von elektrotopologischen Indices (E-States). Des Weiteren werden die Blut/Luft-, Fett/Luft- und Dampfdruck-Datensätze charakterisiert, deren Wertebereiche vorgestellt und eine detaillierte Betrachtung der verwendeten Verteilungskoeffizienten und eine Erklärung zu den Abraham-Parametern gegeben. Zusätzlich werden die verwendeten Programme angeführt sowie die verwendeten statistischen Parameter beschrieben und erörtert.

#### 3.1 Modelliermethoden

Die in dieser Arbeit angeführten Modelle werden mit multilinearer Regression angepasst, wobei einige als Parameter LSER-Deskriptoren nutzen, andere wiederum verschiedene Verteilungskoeffizienten, mit denen die Kompartimente Blut, Luft und Fett beschrieben werden. Für den Blut/Luft-Verteilungskoeffizienten werden mit der Beschreibung der Eigenschaft mit Fragmenten sowie mit elektrotopologischen Indices zwei weitere Ansätze verwendet, die im Folgenden näher erläutert werden.

##### 3.1.1 Fragmentmodell

Das Fragmentmodell basiert auf der Zerlegung eines chemischen Moleküls in Strukturfragmente, beispielsweise funktionelle Gruppen, denen festgelegte Beiträge zugeordnet werden. Diese Beiträge sind additiv. Es existiert damit eine bestimmte Anzahl an Fragmenten für ein Molekül, die genau dieses bestimmte Molekül wiedergeben. Hierzu werden keine Korrekturfaktoren genutzt. Diese Gruppenbeitragsmethode (englisch: *group contribution method*) nutzt das Prinzip der gleichen Teilstrukturen chemischer Stoffe in vielen verschiedenen Molekülen. Die kleinsten gemeinsamen Bestandteile sind die Atome und die Bindungen. Die betrachteten organischen Stoffe sind aus Kohlenstoff und Wasserstoff aufgebaut, optional enthalten sie zusätzlich Sauerstoff, Stickstoff, Halogene und Schwefel. Zusammen mit einer Einfach-, Doppel- und Dreifachbindung können nur aus diesen Atomtypen und drei Bindungsarten tausende von Molekülen gebaut werden. Die nächsten, leicht komplexeren Bausteine aus Einzelteilen sind funktionelle Gruppen, die selbst aus wenigen Atomen und Bindungen aufgebaut sind.

## Material und Methoden

Mit einer Gruppenbeitragsmethode können Eigenschaften von reinen Chemikalien und Mischungen durch die Nutzung von Gruppen- oder Atomeigenschaften vorhergesagt werden. Dies reduziert die Anzahl der benötigten Daten erheblich, da nicht die Kenntnis der Eigenschaften von Hunderten von Stoffen sondern Daten von nur ein paar Dutzend bekannt sein müssen.

Mit Fragmentmodellen können thermodynamische Eigenschaften aus der Molekülstruktur abgeschätzt und vorhergesagt werden, weshalb sie sich gut für die Erstellung von QSPR-Modellen für Verteilungskoeffizienten eignen.

In Publikationen aus den Jahren 2003 und 2005 nutzten Beliveau *et al.* [33,34] Gruppenbeitragsmethoden zur Vorhersage von verschiedenen Verteilungskoeffizienten, die für PBPK-Modelle benötigt werden. Chemische Fragmente oder Gruppen haben eindeutige und konstante Beiträge, die aufsummiert werden, um den Parameter von Interesse ( $P_i$ ) zu berechnen (Gleichung (3.1)).

$$\log P_i = \sum_{i=1}^n f_i \times C_i \quad (3.1)$$

In Gleichung (3.1) ist  $f_i$  die Häufigkeit des Auftretens einer Gruppe  $i$  in einem gegebenen Molekül und  $C_i$  der Beitrag des Fragmentes  $i$  zum logarithmierten Verteilungskoeffizienten  $P_i$ . Diese QSPR-Gleichung sagt aus, dass jedes Strukturfragment einen additiven und konstanten Beitrag zu den Verteilungskoeffizienten liefert [34]. Durch die Nutzung von experimentell ermittelten Verteilungskoeffizienten zusammen mit der Häufigkeit, mit der jedes Fragment der im Datensatz enthaltenen Stoffe auftritt, können die fragmentspezifischen  $C_i$ -Werte quantifiziert werden [33].

Als Erweiterung dieser „reinen“ Fragmentmethode können Korrekturfaktoren eingesetzt werden. Diese beschreiben beispielsweise Wechselwirkungen zwischen den einzelnen Fragmenten und werden über ihre Häufigkeit des Vorkommens eingerechnet. Zudem gibt es auch Indikatoren, bei denen ein bestimmtes Merkmal entweder in die Regression eingeht oder nicht. Sie arbeiten nach dem „Ja-Nein-Prinzip“.

Ein wichtiger Verteilungskoeffizient, der mit dieser erweiterten Fragmentmethode erfolgreich abgeschätzt werden kann, ist der Oktanol/Wasser-Verteilungskoeffizient, beschrieben von Meylan und Howard im Jahre 1994 [70]. Ein damals neuer Fragmentansatz, die

## Material und Methoden

Atom/Fragment-Beitrags-Methode AFC (englisch: *atom/fragment contribution*) ist ein "reduktionistischer" Ansatz und wurde mit multilinearer Regression von zuverlässigen, experimentellen  $\log K_{ow}$ -Werten entwickelt. Die Regression wurde in zwei separaten und aufeinanderfolgenden Schritten durchgeführt. Die erste Regression korrelierte Atom/Fragment-Werte mit  $\log K_{ow}$  und die zweite korrelierte Korrekturfaktoren mit  $\log K_{ow}$ . Der Oktanol/Wasser-Verteilungskoeffizient eines Stoffes wurde durch Aufsummieren aller Atom/Fragment-Werte und Korrekturfaktoren bestimmt, die in einer Struktur enthalten sind. Eine erste Überlegung war die sorgfältige Auswahl von einfachen Atomen/Fragmenten, die in beliebiger Kombination verbunden werden konnten, um möglicherweise fehlende Fragmente auszuschließen. Die AFC-Methode wurde nur für organische und siliciumorganische Verbindungen und ausgewählte organische Salze entwickelt. Anorganische Verbindungen und ihre experimentellen  $\log K_{ow}$ -Werte wurden aus der Betrachtung ausgeklammert.

### 3.1.2 Elektrotopologische Indices

Die elektrotopologischen Indices E-States (englisch: *electrotopological state index*) wurden in der QSAR-Analyse unter anderem durch Kier und Hall im Jahre 1990 [71,72] bekannt gemacht. E-States-Ansätze besitzen eine hohe Wichtigkeit in der QSAR-Modellierung. Mit Indices auf Atomebene, die direkt von Molekül-Verbindungstabellen berechnet werden, gibt es die Möglichkeit Atome und Regionen im Molekül, die für die Aktivität wichtig sind, zu identifizieren. Diese Indices vereinen den elektronischen Charakter mit der topologischen Umgebung eines jeden Skelettatoms in einem Molekül [73]. Der E-State eines Skelettatoms ist als intrinsischer Wert formuliert, der mit einem Störungsterm addiert wird (ähnlich dem Fragmentansatz). Letzterer stammt von der elektronischen Wechselwirkung innerhalb der topologischen Umgebung jedes Atoms im Molekül. Somit variiert der E-State für ein bestimmtes Atom (oder Atomtyp) von Molekül zu Molekül und ist abhängig von der genauen Struktur der Verbindung. Zum Beispiel ist der berechnete E-State-Wert für ein aromatisches Kohlenstoffatom kleiner, wenn es zu einem Kohlenstoff, der mit einer Hydroxygruppe substituiert ist, benachbart ist und sogar noch kleiner, wenn die OH-Gruppe direkt an das Kohlenstoffatom gebunden ist. Beide dieser Werte sind zudem kleiner als der E-State-Wert eines unsubstituierten Kohlenstoffatoms in Benzol [74]. Der intrinsische Zustand basiert auf der Kier-Hall-Elektronegativität [72] und ist abgeleitet von dem Verhältnis dieser Elektronegativität zur Anzahl der skelettartigen Sigmabindungen für dieses Atom. Die

## Material und Methoden

molekulare Konnektivität wird mit einem  $\delta$  dargestellt und gibt die Anzahl der Anbindungen (Ecken) im Skelett (Graph) wieder. Sie berechnet sich aus der Größe  $\sigma$ , der Anzahl der Elektronen in  $\sigma$ -Orbitalen, subtrahiert mit der Anzahl an Wasserstoffatome, die an das Atom gebunden sind [74].

Dieser E-States-Ansatz wurde für die chemometrische Modellierung des Blut/Luft-Verteilungskoeffizienten verwendet, zum Einen als reines E-States-Modell, als Kombination mit dem Fragmentmodell und letztendlich ein LSER-Modell kombiniert mit E-States (siehe Kapitel 4.2.4). Einige wichtige Atomgruppen der E-States mit den dazugehörigen Gruppensymbolen sind in Tabelle 8.4 im Anhang 8.1.3 gezeigt.

### 3.2 Charakterisierungen der Datensätze

Die Charakterisierung der Datensätze für die Verteilungskoeffizienten  $K_{ba}$  und  $K_{fa}$  und den Dampfdruck  $P_v$  beruht auf der strukturellen Konstitution der Verbindungen und geschieht im Hinblick auf die Atomtypen, die Komplexität und die Polarität.

#### 3.2.1 Datensatz von $K_{ba}$

Für den Blut/Luft-Verteilungskoeffizienten von Mensch und Ratte werden zunächst die Werte für 196 Stoffe von Abraham *et al.* [26] genommen. Edelgase, Verbindungen in elementarer Form und anorganischen Verbindungen werden entfernt, da diese für die PBPK-Modellierung sowie für die verschiedenen Methoden, die zur Modellierung verwendet werden, nicht relevant und passend sind. Mit Datenprüfung und Ergänzung aus verschiedenen Literaturquellen [12,38,57,75] konnten insgesamt 191  $\log K_{ba}$  Werte von organischen Verbindungen für Mensch und Ratte ermittelt werden (siehe Anhang 8.1.1, Tabelle 8.1). Mensch- und Ratten-Daten werden dabei vereint, das heißt, bei Verbindungen zu denen Mensch- und Ratten-Daten vorhanden sind, wird der Mittelwert gebildet wie es von Abraham *et al.* [26] gezeigt wird, da der experimentelle Fehler von  $\log K_{ba}$  (Mensch) und  $\log K_{ba}$  (Ratte) Daten im Allgemeinen größer ist, als der systematische Fehler zwischen beiden. Auch Peyret und Krishnan [1] behaupten, dass der Menschen-Blut/Luft-Verteilungskoeffizient am besten mit dem Ratten-Blut/Luft-Verteilungskoeffizienten vorherzusagen ist. Im hier betrachteten Datensatz sind 147 Werte für Mensch und 134 für Ratte enthalten, wovon sich 90 überschneiden. Für Stoffe, bei denen beide Werte zur Verfügung standen, wurden die Mittelwerte verwendet. Im Mittel weichen die Werte der Überschneidungen um 0,2

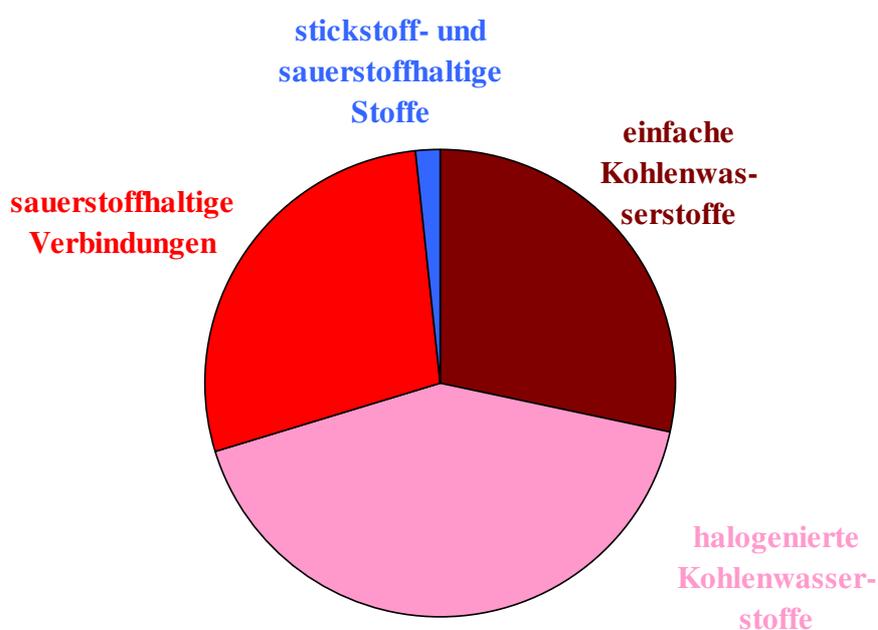
## Material und Methoden

voneinander ab. Die größte Abweichung zeigt einen Unterschied von 1,07 absolut, welcher bei dem Alkan Nonan zu finden ist.

Die 191 Stoffe des Datensatzes besitzen einen molekularen Massenbereich von 26,04 Da bis 252,75 Da und decken folgende Atomklassen ab: olefinische, alkyliche und aromatische Kohlenstoffe, Halogene (F, Cl, Br, I) gebunden an einen aliphatischen oder einen aromatischen Kohlenstoff, Hydroxyl-Sauerstoff gebunden an einen aliphatischen Kohlenstoff, Ether-Sauerstoff, Carbonyl-Sauerstoff (Aldehyde und Ketone), heteroaromatische Sauerstoff (Furan), aliphatische Verbindungen mit Nitrogruppen und ein Epoxid mit vier Cyanidgruppen (Tetracyanoethylenoxid).

Um die chemische Domäne des Datensatzes zu charakterisieren, wurden die Stoffe bezüglich ihrer Zusammensetzung aus unterschiedlichen Atomtypen, der Komplexität und ihrer Polarität untersucht.

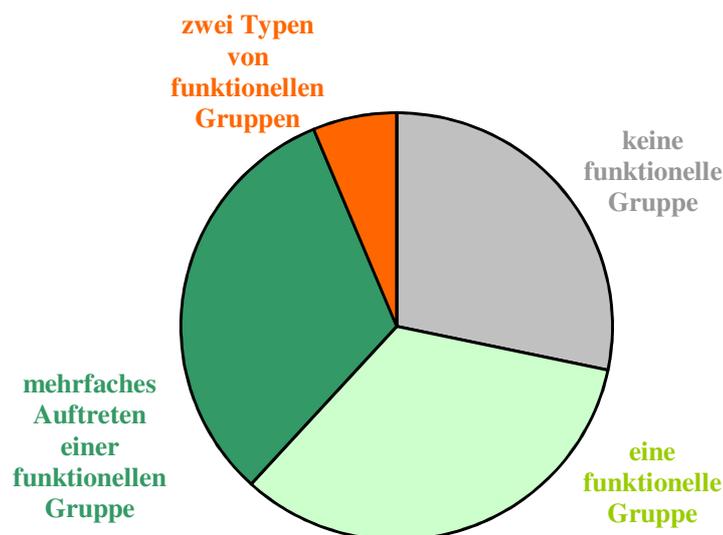
Die Klassifizierung des 191 Stoffe umfassenden Datensatzes bezüglich der verschiedenen Strukturgruppen auf der Basis von Atomtypen zeigt: 54 (28 %) Verbindungen sind einfache Kohlenwasserstoffe, 80 (42 %) sind halogenierte Kohlenwasserstoffe, 54 (28 %) sind sauerstoffhaltige Stoffe, mit oder ohne Halogene, und drei (2 %) beinhalten sowohl Stickstoff als auch Sauerstoff (siehe Schema 3.1). Diese letzten drei sind 1-Nitropropan, 2-Nitropropan und Tetracyanoethylenoxid, das komplizierter aufgebaut ist, als die anderen beiden.



Schema 3.1: Atomtypen des  $K_{ba}$ -Datensatzes.

## Material und Methoden

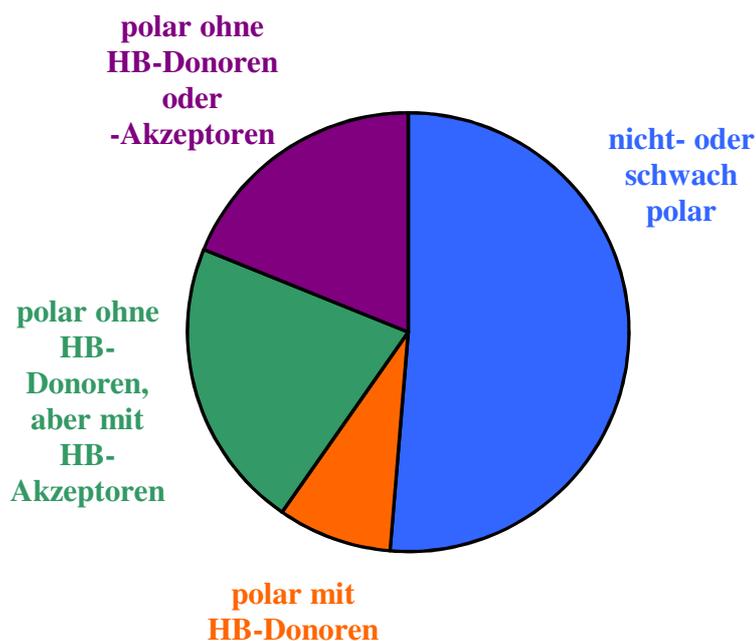
Im Hinblick auf die Komplexität des Datensatzes (siehe Schema 3.2) besitzen die 54 (28 %) einfachen Kohlenwasserstoffe keine funktionelle Gruppe. 64 (34 %) Verbindungen enthalten eine einzige funktionelle Gruppe. Zudem gibt es 61 (32 %) Stoffe mit häufigerem Auftreten der gleichen funktionellen Gruppe und zwölf (6 %) Verbindungen besitzen mindestens zwei verschiedene funktionelle Gruppen.



Schema 3.2: Komplexität des  $K_{ba}$ -Datensatzes.

Die Stoffe im Trainingssatz wurden weiterhin bezüglich ihrer unterschiedlichen Polaritäten, wie sie von Schüürmann *et al.* [76] gezeigt sind, analysiert. Der größte Bereich umfasst 98 (51 %) nicht oder schwach polare Verbindungen, 16 (8 %) sind polare Substanzen mit Wasserstoffbrücken-Donoren (HB-Donoren), 41 (22 %) sind polar auf Grund von Wasserstoffbrücken-Akzeptoren (HB-Akzeptoren) und 36 (19 %) Stoffe sind polare Chemikalien ohne Wasserstoffbrücken-Donoren und -Akzeptoren (siehe Schema 3.3).

## Material und Methoden



Schema 3.3: Polarität des  $K_{ba}$ -Datensatzes.

Diese Datensatz-Charakterisierung zeigt, dass die Verbindungen bezüglich ihrer Atomtypen hauptsächlich sehr einfach strukturiert sind und dass der Datensatz eine mäßige Komplexität aufweist. Im Vergleich dazu weist die Polarität mit ungefähr 50 % polaren und 50 % unpolaren Verbindungen ein ausgewogenes Bild auf.

Der Wertebereich der  $\log K_{ba}$ -Werte erstreckt sich von  $-1,520$  für 1,1,1,2,2,3,3,4,4-Nonafluorbutan bis  $4,508$  für 2-Methoxyethanol. Dies entspricht einem Bereich von sechs logarithmischen Einheiten.

### 3.2.2 Datensatz von $K_{fa}$

Menschen- und Rattendaten für den Fett/Luft-Verteilungskoeffizienten wurden hauptsächlich von Meulenberg *et al.* [38] und aus verschiedenen anderen Literaturquellen [18,35,75,77] entnommen. Insgesamt wurden  $\log K_{fa}$ -Werte für 96 organische Stoffe erhalten. Wiederum wurden Menschen- und Rattendaten vereint, da der mittlere Fehler zwischen beiden so gering ist, dass er vernachlässigt werden kann [25]. Es wurden auch hier die Mittelwerte für die Stoffe, für die beide Werte zur Verfügung standen, verwendet. Im  $\log K_{fa}$ -Datensatz sind 44 Werte für Mensch und 78 für Ratte enthalten, wovon sich 26 überschneiden. Die Werte der Überschneidungen weichen im Mittel um 0,08 voneinander ab. Die größte Abweichung zeigt der verzweigte Alkohol 2-Methyl-1-propanol mit einem absoluten Unterschied von 0,27.

## Material und Methoden

Die 96 Stoffe aus dem Datensatz besitzen eine Molekülmasse von 28,05 Da bis 236,74 Da und beinhalten Stoffklassen der einfachen Kohlenwasserstoffe, halogenierten Kohlenwasserstoffe, sauerstoffhaltigen Stoffe (optional mit Halogenen) sowie stickstoffhaltigen Verbindungen (optional sauerstoffhaltig).

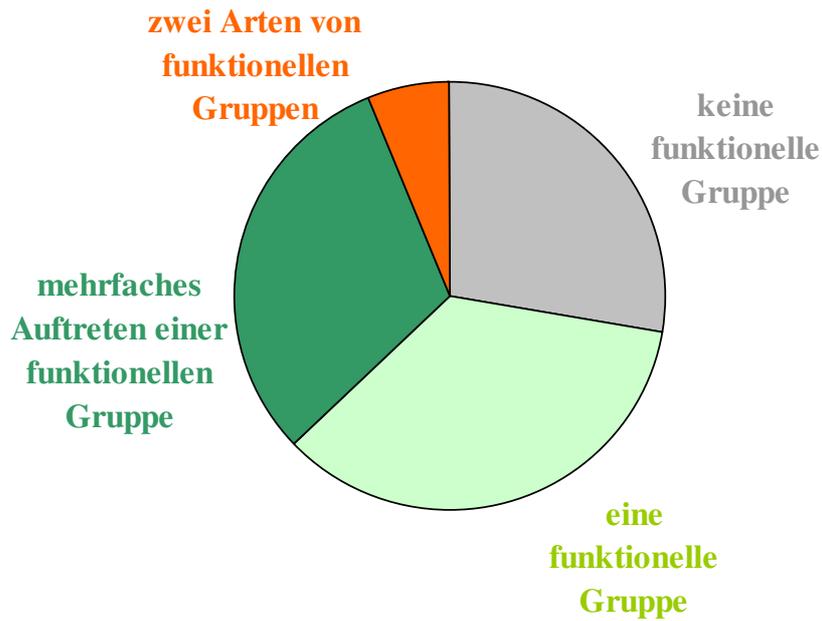
Eine Klassifizierung der Stoffe im Datensatz bezüglich der Existenz verschiedener Gruppen von Atomtypen sieht folgendermaßen aus (siehe Schema 3.4): Im Datensatz mit 96 Verbindungen haben 27 (28 %) Stoffe eine einfache Kohlenwasserstoff-Struktur, 38 (40 %) sind halogenierte Kohlenwasserstoffe, 29 (30 %) sind sauerstoffhaltige Verbindungen, die auch Halogene enthalten können sowie zwei (2 %) Stoffe besitzen Stickstoff und Sauerstoff.



Schema 3.4: Atomtypen des  $K_{fa}$ -Datensatzes.

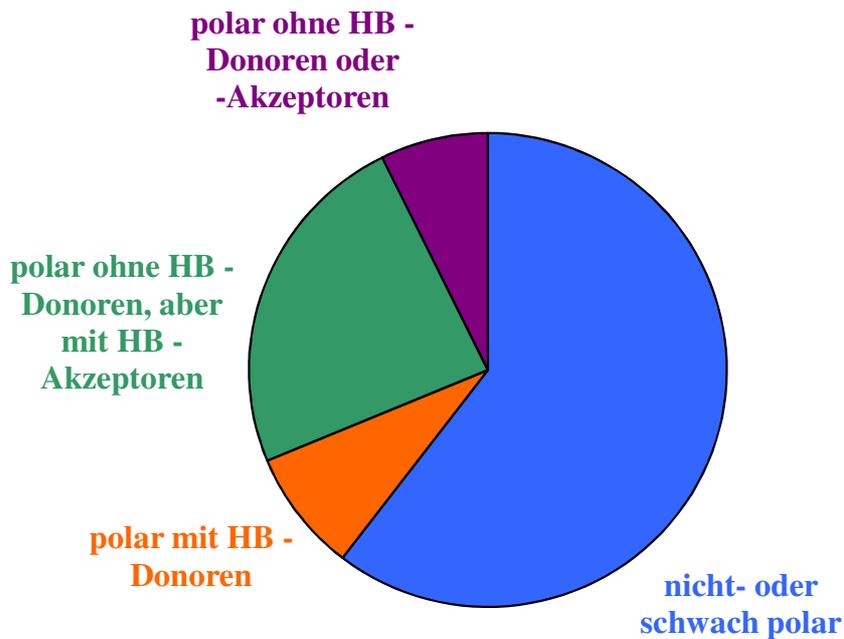
Bezüglich der Komplexität des  $K_{fa}$ -Datensatzes besitzen 27 (28 %) Verbindungen keine funktionelle Gruppe, 34 (36 %) haben eine einzige funktionelle Gruppe, 29 (30 %) Stoffe besitzen eine funktionelle Gruppe doppelt bzw. mehrfach und 6 (6 %) Stoffe besitzen zwei verschiedene funktionelle Gruppen (siehe Schema 3.5).

## Material und Methoden



Schema 3.5: Aufteilung der Komplexität des  $K_{fa}$ -Datensatzes.

Der Datensatz wurde weiterhin bezüglich der Polarität untersucht (siehe Schema 3.6). Der größte Anteil mit 58 (61 %) Stoffen ist nicht oder schwach polar, 8 (8 %) Stoffe sind polar mit HB-Donoren, 23 (24 %) besitzen HB-Akzeptor-Polarität und 7 (7 %) Stoffe sind polar ohne Wasserstoffbrücken-Donoren oder -Akzeptoren.



Schema 3.6: Polarität des  $K_{fa}$ -Datensatzes.

## Material und Methoden

Diese Datensatz-Charakterisierung zeigt, dass die Stoffe hauptsächlich sehr einfach bezüglich Atomtypen und Komplexität sind.

Für den logarithmierten Fett/Luft-Verteilungskoeffizienten  $\log K_{fa}$  liegen die Werte im Bereich von  $-0,022$  für Ethen bis  $4,077$  für *m*-Methylstyrol.

Bei einem Vergleich der Datensatzcharakterisierungen für den Blut/Luft- und Fett/Luft-Verteilungskoeffizienten stellt sich heraus, dass sich die Verteilungen der Atomtypen und Komplexität sehr stark ähneln. Lediglich die Polaritätsverteilung in den beiden Datensätzen unterscheidet sich. Während im Blut/Luft-Datensatz ca. 50 % der Verbindungen polar sind, trifft dies im Fall des Fett/Luft-Datensatzes nur zu knapp 40 % zu. Der größte Unterschied bei den polaren Stoffen findet sich bei denen, die keine Wasserstoffbrücken-Donoren oder HB-Akzeptoren sind.

### 3.2.3 Datensatz von $P_v$

Ein Datensatz von 1675 experimentellen  $\log P_v$ -Werten wurde von Ralf-Uwe Ebert aus unserem Arbeitskreis gesammelt [78-112] und für diese Arbeit zur Verfügung gestellt. Die Werte für den Dampfdruck erstrecken sich über einen logarithmischen Bereich von  $-13,9$  [Pa] für Nicosulfuron bis  $6,1$  [Pa] für Carbonylsulfid, was einem Bereich von 20 Größenordnungen entspricht.

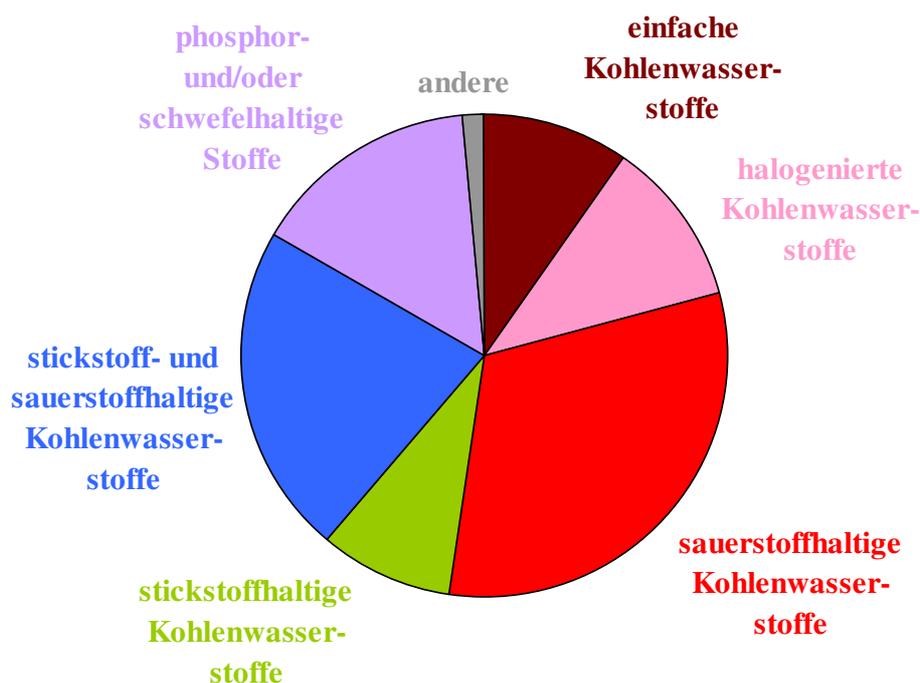
Der Dampfdruck  $P_v$  ist temperaturabhängig. Die meisten der experimentellen Werte sind bei  $25\text{ °C}$  gegeben. Die Modellierung erfolgt jedoch bei einer Temperatur von physiologischen  $37\text{ °C}$ . Eine Auftragung von berechneten  $P_v$  nach Grain-Antoine (siehe Kapitel 4.4) bei  $25\text{ °C}$  gegen die gleiche Berechnung bei  $37\text{ °C}$  ergibt eine Steigung von  $0,92$  und einen Achsenabschnitt von  $0,55$  mit einem Bestimmtheitsmaß  $r^2$  von  $0,999$ . Der Unterschied der Berechnungen bei beiden Temperaturen mit einem mittleren Fehler von  $0,55$  [Pa] logarithmischen Einheiten hat keinen erheblichen Einfluss verglichen mit dem Gesamtbereich der Daten. Selbst der maximale Unterschied von  $1,6$  [Pa] logarithmischen Einheiten ist noch relativ klein. Somit können ohne Bedenken die experimentellen Werte bei  $25\text{ °C}$  für die Modellierung von  $P_v$  bei  $37\text{ °C}$  hergenommen werden.

Im  $P_v$ -Datensatz sind viele verschiedene Stoffklassen enthalten wie nichtaromatische und aromatische Kohlenwasserstoffe, homogen und gemischt halogenierte Verbindungen, sauerstoffhaltige Stoffe, ohne und mit Halogenen, stickstoffhaltige Chemikalien ohne und mit Halogenen, Stoffe, die sowohl Sauerstoff, als auch Stickstoff enthalten und zusätzlich

## Material und Methoden

optional halogeniert sind. Des Weiteren finden sich Isocyanate und Stickoxide und schwefelhaltige Verbindungen, sowohl mit als auch ohne Halogene sowie sauerstoffhaltige und stickstoffhaltige Schwefelverbindungen und viele andere mehr. Der molekulare Massenbereich erstreckt sich von 18,0 Da bis 1052,7 Da.

Die Klassifizierung des 1675 Verbindungen enthaltenen Datensatzes nach Substanzen mit verschiedenen Gruppen von Atomtypen ist in Schema 3.7 dargestellt und setzt sich aus 165 (10 %) einfachen Kohlenwasserstoffen, 186 (11 %) halogenierten Kohlenwasserstoffen, 523 (32 %) sauerstoffhaltigen Verbindungen, die auch Halogene enthalten können, 149 (9 %) stickstoffhaltigen (optional halogenierten) sowie 372 (22 %) Stoffen, die Stickstoff und Sauerstoff enthalten, zusammen. Darüber hinaus sind 258 (15 %) phosphor- und/oder schwefelhaltige Stoffe und 22 (1 %) sind andere Verbindungen, die nicht in eine der vorher genannten Kategorien einzuordnen sind.

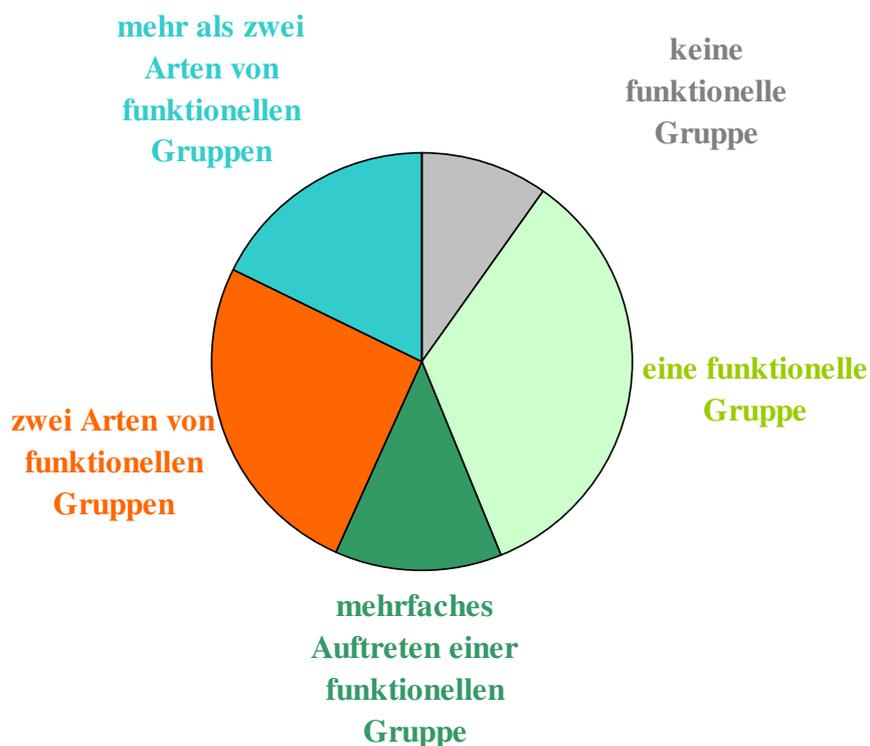


Schema 3.7: Atomtypen des  $P_v$ -Datensatzes.

Wird die Komplexität des  $P_v$ -Datensatzes betrachtet, kann an Schema 3.8 visuell die Verteilung an funktionellen Gruppen erkannt werden. Somit haben 165 (10 %) Stoffe keine funktionelle Gruppe und 569 (34 %) besitzen eine funktionelle Gruppe. Bei 215 (13 %) Verbindungen tritt eine funktionelle Gruppe mehrfach auf, bei 427 (25 %) Stoffen liegen zwei

## Material und Methoden

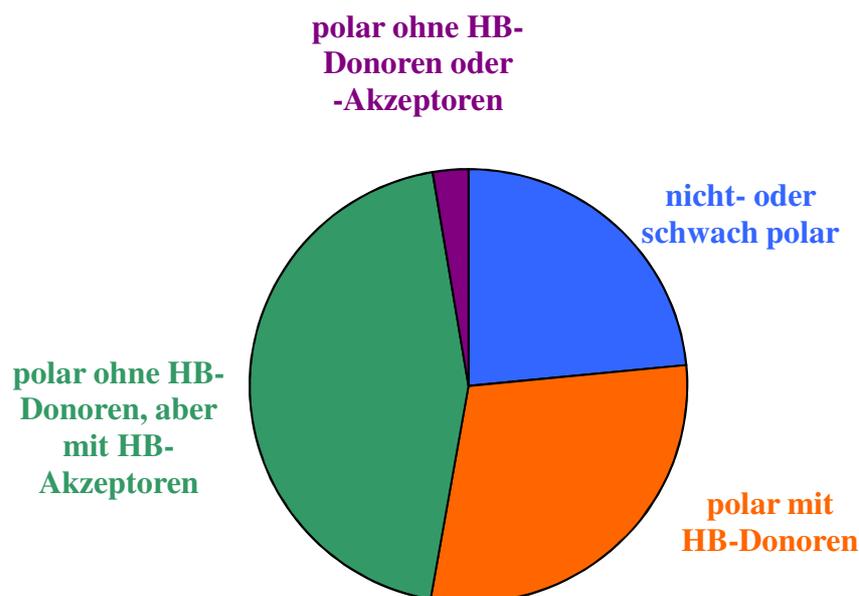
Arten von funktionellen Gruppen vor und bei 298 (18 %) sind mehr als zwei Arten von funktionellen Gruppen vorhanden. Eine Verbindung kann dieser Einteilung nicht zugeordnet werden und wird als eine nicht klassifizierte Struktur ausgewiesen. Hierbei handelt es sich um das anorganische Wasser.



**Schema 3.8:** Komplexität des  $P_v$ -Datensatzes.

In Anbetracht der Polarität zeigt der  $P_v$ -Datensatz, dass 393 (23 %) Substanzen nicht oder schwach polar sind. 491 (29 %) sind polar mit HB-Donoren, 747 (45 %) Stoffe sind polar mit HB-Akzeptoren und 44 Verbindungen, was 3 % entspricht, sind polar ohne HB-Donoren oder -Akzeptoren (siehe Schema 3.9).

## Material und Methoden



Schema 3.9: Polarität des  $P_v$ -Datensatzes.

Verglichen mit den Datensätzen der Verteilungskoeffizienten (Kapitel 3.2.1, Seite 35ff und 3.2.2, Seite 38ff), sind im  $P_v$ -Datensatz viel mehr verschiedene Atomtypen enthalten, was sich eindeutig in der Komplexität widerspiegelt. Auch die Polarität ist stärker auf polare Verbindungen verlagert, während nur noch weniger als 25 % nicht bzw. schwach polar sind.

### 3.3 Verteilungskoeffizienten und Parameter zur Modellierung

Für die Modellierung der Verteilungskoeffizienten Blut/Luft und Fett/Luft wurden verschiedene Ansätze verwendet. Einer davon nutzt andere Verteilungskoeffizienten, die die Zielgrößen mittels multilinearer Regression beschreibt. Die verwendeten Verteilungskoeffizienten sind der Wasser/Luft-Verteilungskoeffizient ( $\log K_{wa}$ ), der Oktanol/Wasser-Verteilungskoeffizient ( $\log K_{ow}$ ), der Oktanol/Luft-Verteilungskoeffizient ( $\log K_{oa}$ ), der Membran/Wasser-Verteilungskoeffizient ( $\log K_{mw}$ ), der Humanalbumin/Wasser-Verteilungskoeffizient ( $\log K_{HSA}$ ) und der Olivenöl/Luft-Verteilungskoeffizient ( $\log K_{olive}$ ). Für diese können entweder experimentelle Daten oder berechnete Werte aus verschiedenen Vorhersagemodellen wie dem von Abraham *et al.* [53,113] eingeführten LSER-Ansatz verwendet werden.

Die Modellierung des Dampfdrucks  $P_v$  erfolgt mit dem Verteilungskoeffizienten  $K_{wa}$ , bzw. der daraus resultierenden Henry-Konstante  $H$  und der Wasserlöslichkeit  $S_w$  (siehe Kapitel 2.4.1, Seite 24ff).

## Material und Methoden

### 3.3.1 Abraham-Parameter

Ansätze nach Abraham funktionieren mit experimentellen Deskriptoren generell gut [25,26]. Die Berechnung der Stoffdeskriptoren, ausgehend von der chemischen Struktur, ist prinzipiell möglich [54,114], aber sowohl die Güte als auch der Anwendungsbereich zur Beschreibung physiko-chemischer Eigenschaften von Chemikalien wie beispielsweise Verteilungskoeffizienten sind möglicherweise noch nicht ausreichend [56]. Mit dem Ziel, ein Vorhersagemodell zu entwickeln, das nur auf der chemischen Struktur basiert, können keine experimentellen Eingabeparameter erlaubt sein. Deswegen werden die neuen  $K_{ba}$ - und  $K_{fa}$ -Modelle mit berechneten LSER-Deskriptoren erstellt. Jedoch werden als Vergleich die Berechnungen mit experimentellen Deskriptoren, die zum Großteil aus der ChemProp-Datenbank [115,116] stammen, zusätzlich gezeigt.

### 3.3.2 Wasser/Luft-Verteilungskoeffizient

Zur Bestimmung des Blut/Luft-Koeffizienten wird unter anderem der Wasser/Luft-Verteilungskoeffizient  $\log K_{wa}$  verwendet, da Wasser mit rund 50 % ein wichtiger Bestandteil des Blutes ist. Dieser wird mit dem weit verbreiteten Modell HENRYWIN [117] berechnet, welches ein Teil der Software EPI Suite [118] ist. Hierbei wird die Henry-Konstante organischer Stoffe bei 25 °C aus den jeweiligen Strukturen bestimmt. Anschließend wird die Henry-Konstante in den Logarithmus des dimensionslosen Verteilungskoeffizienten  $K_{aw}$  durch Division mit der idealen Gaskonstante  $R$  und der Temperatur  $T$  umgerechnet. Nach Inversion ergibt sich der Koeffizient  $K_{wa}$ , welcher hier in logarithmierter Form verwendet wird ( $\log K_{wa}$ ).

Alternativ existiert eine große Anzahl an publizierten Modellen für  $K_{aw}$  sowie auch computergestützte Versionen [115,116]. Vornehmlich wird auch ein Ansatz nach Abraham *et al.* [119] angewendet, um  $\log K_{aw}$  (3.2) oder  $\log K_{wa}$  (3.3) zu beschreiben.

$$\log K_{aw} = 0,994 - 0,577 E - 2,549 S - 3,813 A - 4,841 B^H + 0,869 V \quad (3.2)$$

$$\log K_{wa} = -0,994 + 0,577 E + 2,549 S + 3,813 A + 4,841 B^H - 0,869 V \quad (3.3)$$

$$n = 408, r^2 = 0,99, rms = 0,151$$

Die Erklärung der Abraham-Deskriptoren ist in Kapitel 2.3.2 (Seite 22) sowie die der Statistikparameter in Kapitel 3.5 (Seite 50ff) erörtert. Die Diskussion von Schüürmann *et al.*

## Material und Methoden

[116] zeigt, dass dieses Modell dem HENRYWIN-Modell [117] sowohl in der Güte, als auch in der Allgemeingültigkeit unterlegen ist. Dennoch existieren im Zusammenhang mit dieser Arbeit zwei interessante Aspekte, diese Methode einzusetzen.

1) Wird dieses  $\log K_{wa}$ -Modell zusammen mit anderen Verteilungskoeffizienten, die ebenfalls mit LSER berechnet werden, verwendet, ergibt sich eine konsistente Modellierung, die ausschließlich auf Abraham-Deskriptoren beruht.

2) Die nähere Untersuchung der einzelnen Koeffizienten erlaubt eine mechanistische Betrachtung der zugrundeliegenden Wechselwirkungen (siehe Kapitel 4.1).

Der Wasser/Luft-Verteilungskoeffizient kann über einfache Umrechnungen über den Dampfdruck  $P_v$  gewonnen werden (siehe Kapitel 2.4.1, Seite 24ff), weshalb die Modellierung des Dampfdrucks in dieser Arbeit zusätzlich betrachtet wird. Somit spielt der Dampfdruck eine wichtige Rolle für PBPK-Modelle.

Der dimensionslose Henry-Koeffizient kann auch mit einem Fragment-Modell aus unserer Arbeitsgruppe berechnet werden. Dieses Modell ist noch nicht publiziert, aber in der OSIRIS-Version der hausinternen Software ChemProp [115,116] (siehe Kapitel 3.4, Seite 54) verfügbar.

### 3.3.3 Oktanol/Wasser-Verteilungskoeffizient

Zur Modellierung des Blut/Luft-Verteilungskoeffizienten wird auch der Oktanol/Wasser-Verteilungskoeffizient verwendet. Blut besteht neben Wasser aus organischen Anteilen, die mit dem organischen Oktanol beschrieben werden können. Poulin und Krishnan haben 1995 zur Vorhersage von Gewebe/Blut-Verteilungskoeffizienten experimentelle  $\log K_{ow}$ -Werte verwendet [16]. Zur Berechnung der  $\log K_{ow}$ -Werte existiert auch eine LSER-Gleichung (3.4) nach Abraham [120]. Die von den Autoren verwendete statistische Größe  $sd$  ist der Standardfehler (englisch: *standard deviation*).

$$\log K_{ow} = 0,088 + 0,562 E - 1,054 S + 0,032 A - 3,460 B^O + 3,814 V \quad (3.4)$$

$$n = 613, r^2 = 0,99, sd = 0,116$$

Der Verteilungskoeffizient  $\log K_{ow}$  kann auch mit dem Modell KOWWIN [121], das ein Teil von EPI Suite [118] ist, berechnet werden. Dieses nutzt eine Fragment-Methode zur Berechnung. Hierbei wird eine Struktur in Fragmente geteilt (Atome oder größere

## Material und Methoden

funktionelle Gruppen) und die Werte der Koeffizienten für jedes Fragment oder jede Gruppe aufsummiert. Diese Methodik wird von Meylan und Howard im Jahre 1995 [70] beschrieben (siehe Kapitel 3.1.1).

### 3.3.4 Oktanol/Luft-Verteilungskoeffizient

Als Pendant zum Oktanol/Wasser-Verteilungskoeffizienten für die Modellierung von  $\log K_{ba}$  wird der Oktanol/Luft-Verteilungskoeffizient zur  $\log K_{fa}$ -Modellierung verwendet. Dabei ist das zweite Kompartiment Luft identisch und die Bereiche Blut und Oktanol können einander gegenübergestellt werden. Der Verteilungskoeffizient  $\log K_{oa}$  wird aus  $\log K_{ow}$  und  $\log K_{wa}$  nach Gleichungen (3.4) (Kapitel 3.3.3, Seite 46) und (3.3) (Kapitel 3.3.2, Seite 45) entweder manuell oder mit EPI Suite [118] berechnet. Das entsprechende Modul heißt KOAWIN [122], bei dem die Berechnung sowohl mit experimentellen Parametern aus der Datenbank, als auch mit berechneten Eingangsgrößen aus KOWWIN [121] und HENRYWIN [117] durchgeführt werden kann.

$$\log K_{oa} = \log K_{ow} + \log K_{wa} \quad (3.5)$$

$$K_{ow} \cdot K_{wa} = \frac{c_o}{c_w} \frac{c_w}{c_a} = \frac{c_o}{c_a} = K_{oa} \quad (3.6)$$

Die  $\log K_{oa}$  Berechnung kann ebenfalls mit Hilfe einer LSER-Gleichung für wasserfreies Oktanol durchgeführt werden (siehe Gleichung (3.7), [123,124]).

$$\log K_{oa} = -0,120 + 0,203 E + 0,560 S + 3,576 A + 0,702 B^H + 0,939 L \quad (3.7)$$

$$n = 156, r^2 = 0,997, sd = 0,125$$

$\log K_{oa}$  kann über ein Fragment-Modell, das in unserer Gruppe entwickelt wurde, berechnet werden. Dieses Modell ist noch nicht publiziert, aber in der internen Version von ChemProp [115,116] implementiert.

### 3.3.5 Membran/Wasser-Verteilungskoeffizient

Für den Verteilungskoeffizienten zwischen Membran und Wasser wurde in unserer Arbeitsgruppe eine LSER-Gleichung erstellt, nachdem die Zielgröße experimentell mittels

## Material und Methoden

Gleichgewichtsdialyse, Ultrazentrifugation und Festphasenmikroextraktion (SPME, englisch: *solid phase micro extraction*) bestimmt wurde [125]. Die Membran ist in diesem Fall eine biologische, aus einer Lipiddoppelschicht aufgebaute Struktur, die in Zellen vorkommt. Die dazugehörige Gleichung zum Membran/Wasser-Verteilungskoeffizienten ist in [125] gezeigt. Die Regression der 197 Stoffe zeigt ein Bestimmtheitsmaß von  $r^2 = 0,93$  und eine Wurzel der Fehlerquadratsumme von 0,39 (siehe Kapitel 3.4). Alle Stoffparameter, mit Ausnahme von  $E$ , sind negativ, wobei die Wasserstoffbrücken-Akzeptorstärke  $B$  den größten Einfluss hat. Gefolgt wird er vom Volumen nach McGowan, das einen moderaten Einfluss auf die Membran/Wasser-Verteilung besitzt. Die Deskriptoren  $E$ ,  $S$  und  $A$  besitzen kleinere Phasenparameter, was darauf hindeutet, dass sie nur geringe Einflüsse auf die Verteilung ausüben. Die Fehler für die einzelnen Phasenparameter betragen minimal 0,20 ( $e$ ) und maximal 0,28 ( $b$ ).

Dieses Modell ist in der OSIRIS-Version von ChemProp [115,116] (siehe Kapitel 3.4) als implizierter Deskriptor verfügbar.

### 3.3.6 Humanalbumin/Wasser-Verteilungskoeffizient

Der HSA/Wasser-Verteilungskoeffizient  $K_{\text{HSA}}$  wird mit einer LSER-Gleichung (3.8) berechnet, die Valko *et al.* [23] 2003 veröffentlichten.

$$\log K_{\text{HSA}} = -1,28 + 0,82 E - 0,36 S + 0,18 A - 1,97 B^{\text{H}} + 1,62 V \quad (3.8)$$

$$n = 52, r^2 = 0,83, rms = 0,33$$

Der verwendete HSA/Luft-Koeffizient  $K_{\text{HSA}/a}$  wird aus  $\log K_{\text{HSA}}$  mit  $\log K_{\text{wa}}$  umgerechnet (siehe Gleichung (3.10)).

$$\log K_{\text{HSA}/a} = \log K_{\text{HSA}} + \log K_{\text{wa}} \quad (3.9)$$

### 3.3.7 Olivenöl/Luft-Verteilungskoeffizient

Für den Olivenöl/Luft-Verteilungskoeffizienten (oder kurz:  $K_{\text{olive}}$ ) gibt es eine veröffentlichte LSER-Gleichung [20], mit der dieser Verteilungskoeffizient berechnet werden kann. Der  $bB$ -Term mit einem ursprünglichen Koeffizienten von 0,09 war statistisch nicht signifikant und

## Material und Methoden

kann daher vernachlässigt werden, um damit zu Gleichung (3.10) zu kommen [20]. Das Modell ist auch in die Software ChemProp [115,116] (siehe Kapitel 3.4) implementiert.

$$\log K_{\text{olive}} = -0,156 - 0,254 E + 0,859 S + 1,656 A + 0,873 L \quad (3.10)$$

$$n = 215, r^2 = 0,981, sd = 0,196$$

Dieser Verteilungskoeffizient wird bei der  $\log K_{\text{ba}}$ -Modellierung für die Beschreibung der organischen Phase im Blut genutzt und bei der  $\log K_{\text{fa}}$ -Modellierung als Äquivalent für Fett eingesetzt.

### 3.3.8 Wasserlöslichkeit

Wasser ist aufgrund seiner Polarität ein gutes Lösungsmittel für viele Stoffe. Generell gilt, dass die Wasserlöslichkeit  $S_w$  mit abnehmender Polarität des Stoffes abnimmt. Substanzen, die sich gut in Wasser lösen (z. B. Salze) werden als hydrophil (wasserliebend) bezeichnet, während Stoffe, die sich nicht gut in Wasser lösen (z. B. Fette und Öle), als hydrophob (wassermeidend) bekannt sind. Die Fähigkeit einer Substanz, sich in Wasser zu lösen, ist stark temperaturabhängig und dadurch bestimmt, ob die Anziehungskräfte zwischen Substanz und Wasser mit den Wasserstoffbrückenbindungen gleichziehen oder sie sogar übertreffen oder nicht. Wenn ein Stoff Eigenschaften besitzt, die diese starken intermolekularen Kräfte nicht überwinden können, wird er aus dem Wasser verdrängt und löst sich somit nicht.

Feststoffe und Gase verhalten sich in Bezug auf die Wasserlöslichkeit unterschiedlich. Gase lösen sich ohne eine feste Grenze der löslichen Menge proportional zum Partialdruck des Gases in Wasser (Henry-Gesetz, siehe Kapitel 2.4.1, Seite 24ff). Die hierbei als Löslichkeit bezeichnete Gleichgewichtskonzentration je Druckeinheit nimmt mit abnehmender Temperatur zu. Dagegen lösen sich Feststoffe bei abnehmender Temperatur meistens schlechter in Wasser, wovon es aber auch viele Ausnahmen gibt, z. B. das Natriumsalz Thenardit [126].

Der Deskriptor  $S_w$  zur Modellierung von  $P_v$  kann über verschiedene Methoden berechnet werden, zum Einen aus dem Oktanol/Wasser-Verteilungskoeffizienten  $K_{\text{ow}}$  [127] oder zum Anderen aus der Struktur über verschiedenste Modellansätze [128-134]. Beruhend auf einem LSER-Ansatz ist das Modell von Abraham und Le [134] für flüssig/fest-Systeme (Gleichung (3.11)) eine Möglichkeit,  $\log S_w$  zu bestimmen.

## Material und Methoden

$$\log S_w = 0,518 - 1,004 E + 0,771 S + 2,168 A + 4,238 B^H - 3,362 A \cdot B^H - 3,987 V \quad (3.11)$$

Statt der üblichen fünf Terme  $E$ ,  $S$ ,  $A$ ,  $B$  und  $V$  (siehe Kapitel 2.3) tritt ein weiterer Term auf, der das Produkt aus Wasserstoffbrücken-Donor- und -Akzeptorstärke  $A$  und  $B$  beschreibt. Der dekadische Logarithmus von  $S_w$  wird in der Einheit mol/L verwendet. Über eine in ChemProp [115,116] implementierte Modellauswahl wird  $\log S_w$  für die hier gezeigten Modellierungen gewonnen [135].

### 3.4 Programme

Für den Großteil der Datenbearbeitung wurde Microsoft Office Excel 2007 [136] verwendet. Die multilinenen Regressionsanalysen wurden mit dem Softwarepaket Statistica 8.0 [137] durchgeführt, während die Grafiken mit OriginPro 8G [138] angefertigt wurden. Für die Ausgabe vieler Verteilungskoeffizienten wurden verschiedene Programme aus der Software EPI Suite [118] und das hausinterne Programm ChemProp [115,116] verwendet. Letzteres wurde auch zur Modellierung des  $\log K_{ba}$ -Fragmentmodells sowie zur Ausgabe von Statistiken herangezogen. Das Statistik-Programm R [139] diente zur Durchführung von Kreuzvalidierungen [140] sowie zur Ausgabe weiterer Statistik-Parameter.

Das System ChemProp [115,116] enthält eine große Anzahl an Berechnungsmethoden für umweltrelevante physikalisch-chemische Stoffeigenschaften, die entweder aus der Literatur stammen oder in unserer Arbeitsgruppe entwickelt wurden. Diese Parameter sind z. B. Verteilungseigenschaften, reine Stoffdeskriptoren und -eigenschaften, Umweltverhalten, Ökotoxizität und Toxizität. Bekannte Modelltypen sind Fragmentmodelle (siehe Kapitel 3.1.1, Seite 32ff), LSER-Gleichungen nach Abraham (siehe Kapitel 2.3.2, Seite 21ff), Read-Across-Ansätze, die zum Teil auf atomzentrierten Fragmenten (ACF, englisch: *atom centred fragments*) beruhen und Modelle, die topologische und elektrotopologische Indices verwenden (siehe Kapitel 3.1.2, Seite 34f). Die aktuelle ChemProp-Version ist 5.2.9 [141].

### 3.5 Statistische Parameter

Die Kalibrierungsgüte des Modells zur Berechnung der Verteilungskoeffizienten  $K_{ba}$  und  $K_{fa}$  und des Dampfdrucks  $P_v$  wird mit dem quadrierten Korrelationskoeffizienten  $r^2$  (Bestimmtheitsmaß) und dem dazugehörigen Fehler, die Wurzel der Fehlerquadratsumme  $rms$  bemessen. Das Bestimmtheitsmaß  $r^2$  ist durch folgende Gleichung definiert [59,140,142]:

## Material und Methoden

$$r^2 = 1 - \frac{\sum_i^n (y_i^{fit} - y_i^{exp})^2}{\sum_i^n (y_i^{exp} - y^{mean})^2} \quad (3.12)$$

In Gleichung (3.12) ist  $y_i^{exp}$  der experimentelle Zielwert eines Stoffes  $i$  (im vorliegenden Fall von  $\log K_{ba}$ ,  $\log K_{fa}$  oder  $\log P_v$ ),  $y_i^{fit}$  der regressionsgefittete Wert davon und  $y^{mean}$  der experimentell beobachtete Mittelwert (*mean*). Im vorliegenden Datensatz wurden  $n$  Werte vorhergesagt. Die Kalibrierungsgüte nimmt Werte von  $0 \leq r^2 \leq 1$  ein.

Die Vorhersagequalität des Modells wird mit dem quadrierten Vorhersagekorrelationskoeffizient  $q^2$  [50,142] quantifiziert. Der Unterschied zum  $r^2$  besteht darin, dass ein Modell nicht neu kalibriert, sondern auf einen anderen Datensatz angewendet wird. Deshalb wird statt  $y_i^{fit}$  die Variable  $y_i^{pred}$  verwendet. Dieses  $q^2$  kommt somit nur bei Anwendung von bereits bestehenden Modellen auf die hier gezeigten Datensätze von  $K_{ba}$ ,  $K_{fa}$  und  $P_v$  zum Einsatz.

$$q^2 = 1 - \frac{\sum_i^n (y_i^{pred} - y_i^{exp})^2}{\sum_i^n (y_i^{exp} - y^{mean})^2} \quad (3.13)$$

In Gleichung (3.13) gibt  $y_i^{pred}$  der vorhergesagte (nicht gefittete) Zielwert wieder und  $y^{mean}$  ist der Mittelwert der experimentellen  $\log K_{ba}$ -,  $\log K_{fa}$  oder  $\log P_v$ -Werte, der für die Bewertung der Vorhersagegüte genutzt wird. Ein perfektes Modell besitzt die Vorhersagegüte  $q^2 = 1$ . Im Gegensatz zu  $r^2$  kann  $q^2$  auch negativ werden.

Für ein neu angepasstes Modell, zu dem keine externen Validierungsdaten vorliegen, werden  $r^2$  und  $q^2$  aus einem Datensatz berechnet und eine Leave-One-Out-Kreuzvalidierung (LOO-CV, englisch: *cross validation*) angewendet. Zur Darstellung der entsprechenden Modellierungsgüte wird  $q_{CV}^2$  angegeben [140]. Bei diesem Verfahren wird jeweils ein experimenteller Wert aus dem Datensatz entfernt und durch ein modifiziertes Modell

## Material und Methoden

vorhergesagt. Dieses Modell wird mit dem verbleibenden N-1 großen Datensatz erzeugt. Daher wird diese Art von Modellvalidierung als Leave-One-Out-Kreuzvalidierung bezeichnet.

$$q_{CV}^2 = 1 - \frac{\sum_i^n (y_i^{\text{pred}(N-1)} - y_i^{\text{exp}})^2}{\sum_i^n (y_i^{\text{exp}} - y_{\text{mean}}^{N-1,i})^2} \quad (3.14)$$

In Gleichung (3.14) gibt  $y_i^{\text{pred}(N-1)}$  den vorhergesagten Zielwert an, die mit LOO ermittelt wurde und  $y_{\text{mean}}^{N-1,i}$  ist der Mittelwert der experimentellen  $\log K_{\text{ba}^-}$ ,  $\log K_{\text{fa}^-}$  oder  $\log P_v$ -Werte, jeweils ohne den einen weggelassenen Wert.

Zusätzlich zu  $r^2$  und  $q^2$  kann die Vorhersagequalität des Modells mit dem *rms*-Fehler charakterisiert werden (Gleichung (3.15)).

$$rms = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^{\text{pred}} - y_i^{\text{exp}})^2} \quad (3.15)$$

Je kleiner dieser Wert ist, desto besser ist ein Modell.

Für die Kreuzvalidierung wird zur Berechnung des *rms* ein Freiheitsgrad abgezogen und wieder  $y_i^{\text{pred}(N-1)}$  betrachtet (siehe Gleichung (3.16), [140]).

$$rms_{CV} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i^{\text{pred}(N-1)} - y_i^{\text{exp}})^2} \quad (3.16)$$

Zusätzlich zu  $r^2$ ,  $q^2$  und *rms* kann noch der systematische Fehler *bias* (siehe Gleichung (3.17)) betrachtet werden, der wiedergibt, ob sich die Fehler im Mittel aufheben.

$$bias = \frac{\sum_i^n (y_{i,1} - y_{i,2})}{n} \quad (3.17)$$

## Material und Methoden

Zur Vervollständigung der Statistikparameter werden der maximale negative Fehler *mne* (englisch: *maximum negative error*; größte Unterschätzung) und der maximale positive Fehler *mpe* (englisch: *maximum positive error*; größte Überschätzung) angegeben. Sie zeigen die Abweichung eines bestimmten Punktes von der Regressionsgeraden an.

Der mittlere absolute Fehler *me* (englisch: *mean error*) wird mit folgender Gleichung berechnet:

$$me = \frac{1}{n} \sum_{i=1}^n |y_i - \bar{y}| \quad (3.18)$$

Alle Regressionsparameter der multilinearen Regressionen werden auf statistische Signifikanz geprüft. Dies geschieht mittels des Signifikanzwerts  $p$  (p-Wert). Der p-Wert ist frei wählbar, wird jedoch häufig, wie auch in dieser Arbeit, mit einem Signifikanzniveau von 0,05 verwendet [137]. Dies bedeutet, dass mit einem p-Wert von größer als 5 % das Ergebnis als nicht signifikant angesehen wird.

### 4 Modellierung von Verteilungskoeffizienten

In diesem Kapitel werden alle Ergebnisse der Modellierung von Verteilungskoeffizienten detailliert vorgestellt und diskutiert. Den größten Teil machen dabei die Modelle für den Blut/Luft-Verteilungskoeffizienten aus. Dazu gehören alle Modelle basierend auf LSER und auf anderen Verteilungskoeffizienten, ein Fragmentmodell und ein auf elektrotopologischen Indices (E-States) gestütztes QSAR-Modell. Des Weiteren werden diese entwickelten Modelle sowohl untereinander als auch mit bereits bestehenden Blut/Luft-Modellen verglichen. Der Fett/Luft-Verteilungskoeffizient wird auch mit anderen Verteilungskoeffizienten beschrieben. Hier kommen  $\log K_{oa}$ ,  $\log K_{mw}$ ,  $\log K_{HSA}$ ,  $\log K_{HSA/a}$  und  $\log K_{olive}$  zum Einsatz. Auch diese Modelle werden miteinander verglichen. Für beide Verteilungskoeffizienten wird eine Konsensmodellstrategie angewendet und bewertet. Abschließend wird die Berechnung von  $\log K_{fb}$  aus den anderen beiden Verteilungskoeffizienten gezeigt.

#### 4.1 Diskussion der Abraham-Modelle aus der Literatur

Alle in dieser Arbeit verwendeten Verteilungskoeffizienten können über LSER-Gleichungen berechnet werden. Daher bietet es sich an, die einzelnen Abraham-Deskriptoren und somit die unterschiedlichen Einflüsse der intermolekularen Wechselwirkungen miteinander zu vergleichen. Diese Wechselwirkungen sind Moleküleigenschaften neutraler Verbindungen wie beispielsweise Van-der-Waals-Wechselwirkungen und Wasserstoffbrückenbindungen.

##### 4.1.1 Blut/Luft-Verteilung

Wird der bereits bestehende Abraham-Ansatz für die Blut/Luft-Verteilung [26] mit den LSER-Gleichungen der Deskriptoren  $\log K_{aw}$  und  $\log K_{HSA}$ ,  $\log K_{ow}$  und  $\log K_{mw}$  verglichen, wird eine bessere Einsicht in die relevanten Wechselwirkungen erlangt.

$$\log K_{ba} = -1,069 + 0,456 E + 1,083 S + 3,738 A + 2,580 B^H + 0,376 L \quad (4.1)$$

$$n = 196, r^2 = 0,938, rms = 0,319$$

Wie bei Gleichung (4.1) gesehen werden kann, sind alle Koeffizienten positiv und somit geben sie entweder die Präferenz der Blutphase oder die Abweisung der Luftphase wieder.

## Modellierung von Verteilungskoeffizienten

Die wichtigsten Deskriptoren sind die Wasserstoffbrücken-Donorstärke  $A$  und die H-Brücken-Akzeptorstärke  $B$  der Substanz mit hohen Werten der Phasenparameter von 3,738 für  $a$  und 2,580 für  $b$ . Wie eindeutig zu sehen ist, wird die Blutphase bevorzugt von Stoffen, die Wasserstoffatome für H-Brückenbindungen zur Verfügung stellen können und zu einem geringeren Anteil von Molekülen, die als Wasserstoffbrücken-Akzeptor fungieren können. Die Polarität und Polarisierbarkeit  $S$  ist mäßig mit einem Phasenparameterwert von 1,083 gewichtet. Dies bedeutet, dass auch die Polarität eine Verbindung in das Blutkompartiment drängt. Die Deskriptoren mit kleineren Beiträgen sind die Exzess-Molrefraktion  $E$  und der logarithmische Hexadecan/Luft Verteilungskoeffizient  $L$ , wobei beide eine leichte Präferenz der Blutphase wiedergeben.

Wird die Gleichung für die Wasser/Luft-Verteilung (Gleichung (4.2)) betrachtet, wird deutlich, dass die H-Brücken-Donorstärke  $A$  und die H-Brücken-Akzeptorstärke  $B$  die größten Beiträge mit Phasenparametern von 3,813 und 4,841 liefern.

$$\log K_{wa} = -0,994 + 0,577 E + 2,549 S + 3,813 A + 4,841 B^H - 0,869 V \quad (4.2)$$

$$n = 408, r^2 = 0,99, rms = 0,151$$

Im Vergleich zum Blut besteht fast die gleiche Präferenz für H-Brücken-Donoren. Jedoch ist sie für H-Brücken-Akzeptoren stärker signifikant und auch um einiges größer, als für die Donoren. Die Akzeptor- oder Basizitätsbevorzugung der Blutphase ist durch den Wasseranteil im Blut begründet, während die Acidität mit den organischen Komponenten in Zusammenhang steht. Das wird eindeutig bei den Modellen für  $\log K_{ow}$  (Gleichung (3.4)),  $\log K_{mw}$  (siehe Kapitel 3.3.5 und [125]) und  $\log K_{HSA}$  (Gleichung (3.8)) mit einem schwach positiven Koeffizienten  $a$  ( $< 0,6$ ) bestätigt. Bei allen drei betrachteten Modellen ist  $b$  deutlich negativ ( $< -1,9$ ), was darauf hindeutet, dass die organische Phase im Blut H-Brücken-Akzeptoren abweist. Mit der allgemeinen Bevorzugung der Blut- gegenüber der Luftphase von Wasserstoffbrücken-Donoren und -Akzeptoren, können Akzeptoren bevorzugt im Wasserkompartiment im Blut gefunden werden, während Donoren in beiden Teilkompartimenten erwartet werden. Sie bevorzugen mindestens jedoch die Wasserphase. Der im Allgemeinen mäßige Effekt von  $S$  (Polarität/Polarisierbarkeit) ist bei der Verteilung in Luft und Wasser (Gleichung (4.2)) betragsmäßig größer als bei der Blut/Luft-Verteilung, während  $S$  bei der Oktanol/Wasser-Verteilung (Gleichung (3.4)) einen ähnlichen Betrag

## Modellierung von Verteilungskoeffizienten

besitzt. Für die Serum- (Gleichung (3.8)) und Membran/Wasser-Verteilung (siehe Kapitel 3.3.5 und [125]) ist der Wert für die Polarität und Polarisierbarkeit deutlich kleiner. Aus den Vorzeichen geht hervor, dass Substanzen mit Wasserstoffbrücken-Akzeptoren aufgrund entsprechender Wechselwirkungen die Wasserphase, nicht aber die organische oder die Luftphase bevorzugen. Sich so verhaltende Verbindungen werden hauptsächlich im Wasser gefunden und weniger im Membran- und Proteinanteil des Blutes. Im Gegensatz dazu ist die allgemeine Präferenz für Wechselwirkungen der Substanzen durch  $n$  und  $\pi$  Elektronen ( $E$ ) mit dem Blutkompartiment zu einem gewissen Anteil unausgeglichen. Dies bedeutet, dass die Werte für den Koeffizienten  $e$  durchweg positiv sind, aber in ihrer Wertigkeit nicht genau übereinstimmen.

Aufgrund ihrer unterschiedlichen Natur sind die Terme  $L$  und  $V$  schwerer zu vergleichen, als die vorangegangenen Deskriptoren. Der negative Beitrag des Volumenterms  $V$  zur Wasser/Luft-Verteilung (Gleichung (4.2)) zeigt die Schwierigkeit für größere Moleküle in die Wasserphase zu gelangen. Sehr viel leichter ist es bei den Fällen mit organischen Anteilen, in denen der dominierende Effekt die abnehmende Flüchtigkeit mit zunehmender Größe ist, welcher sich in einem positiven  $V$ -Term für die Oktanol-, Membran- sowie HSA-Verteilung zeigt. Von großen Molekülen wird daher erwartet, dass sie die organische Phase bevorzugen. Der Einfluss der Molekülgröße auf die Verteilung zwischen Blut und Luft ist jedoch gering. Dadurch kann der kleine  $L$ -Term in der  $\log K_{ba}$ -Gleichung durch einen entsprechenden kleinen  $V$ -Term ersetzt werden (siehe (2.20) und (2.21)).

### 4.1.2 Fett/Luft-Verteilung

Auch für die Fett/Luft-Verteilung existiert bereits ein LSER-Modell von Abraham *et al.* [25], welche in Gleichung (4.3) dargestellt ist.

$$\log K_{fa} = 0,0511 E + 0,728 S + 1,783 A + 0,332 B^H + 0,743 L - 0,052 \quad (4.3)$$

$$n = 129, r^2 = 0,958, sd = 0,194$$

Hierbei ist auffällig, dass alle Phasenparameter positiv sind. Der wichtigste Deskriptor ist die Wasserstoffbrücken-Acidität  $A$  mit einem relativ großen Wert von 1,783. Daraus wird ersichtlich, dass vor allem Stoffe, die Wasserstoffatome für H-Brückenbindungen zur Verfügung stellen können, die Fettphase im Vergleich zur Luftphase bevorzugen, da Luft

## Modellierung von Verteilungskoeffizienten

inert ist. Wie im Vergleich mit Gleichung (4.2) zu sehen ist, ist Fett ein viel schwächerer Donor als Wasser. Bei Molekülen, die als Akzeptor bei Wasserstoffbrücken fungieren können, ist dies zu einem sehr viel geringeren Anteil gegeben. Die Polarität und Polarisierbarkeit  $S$  und der logarithmische Hexadecan/Luft-Verteilungskoeffizient  $L$  sind mäßig gewichtet mit Phasenparameterwerten von 0,728 und 0,743. Das bedeutet, dass sowohl ein hoher Wert für die Polarität, als auch für den logarithmische Hexadecan/Luft-Verteilungskoeffizienten eine Verbindung aus dem Luftkompartiment vertreibt. Die Deskriptoren mit kleineren Beiträgen sind die Exzess-Molrefraktion  $E$  und die H-Brücken-Akzeptorstärke, wobei beide eine leichte Präferenz der Fettphase wiedergeben.

Wird nun diese Gleichung mit den LSER-Gleichungen der zum Modellieren verwendeten Verteilungskoeffizienten  $\log K_{oa}$ ,  $\log K_{mw}$ ,  $\log K_{HSA}$  und  $\log K_{olive}$  verglichen, zeigt sich, dass nur die Oktanol/Luft-Gleichung (3.7) ausschließlich positive Vorfaktoren enthält. Die hier vorliegenden Zahlenwerte für die Polarität und Polarisierbarkeit  $S$  und den logarithmischen Hexadecan/Luft-Verteilungskoeffizienten  $L$  liegen in einer ähnlichen Größenordnung mit nur geringer Abweichung der Phasenparameterwerte. Bei der Wasserstoffbrücken-Acidität und -Basizität hingegen sind die Phasenparameter annähernd doppelt so groß wie im Vergleich zur Fett/Luft-Verteilung. Das begründet sich daraus, dass Moleküle mit H-Akzeptoren und H-Donoren eine doppelt so hohe Affinität zur Oktanolphase wie zur Fettphase besitzen. Bei der Exzess-Molrefraktion  $E$  ist diese Tendenz sogar viermal stärker. Die Werte für  $S$  von  $\log K_{mw}$  (siehe Kapitel 3.3.5 und [125]) und  $\log K_{HSA}$  (3.8) besitzen mit einem relativ kleinen Beitrag negative Vorzeichen, was darauf schließt, dass sich polare und polarisierbare Verbindungen bevorzugt in Wasser aufhalten. Auch  $B$  ist bei diesen beiden Gleichungen negativ, jedoch sind diese Werte sehr viel größer als bei der LSER-Gleichung der Zielgröße  $\log K_{fa}$ . Damit kann nachgewiesen werden, dass Stoffe mit HB-Akzeptoren die Wasserphase deutlich gegenüber der Membran- oder Proteinphase bevorzugen. Der positive Beitrag der Volumenterme  $V$  zur Membran/Wasser- und Protein/Wasser-Verteilung zeigt die Tendenz für größere Moleküle, sich in der wässrigen Phase zu befinden. Die Exzess-Molrefraktion  $E$  und die Wasserstoffbrücken-Acidität  $A$  besitzen kleine, positive Beiträge, was bedeutet, dass sich Moleküle mit HB-Donoren eher in der organischen Phase aufhalten. Bei der LSER-Gleichung von  $\log K_{olive}$  ist keine Wasserstoffbrückenbasizität  $B$  angegeben, da bei ihr der Phasenparameter  $b$  sehr klein sowie statistisch nicht signifikant ist (siehe Kapitel 3.3.7, Seite 48 und [20]). Den größten Einfluss besitzt hier die HB-Donorstärke  $A$  mit einem positiven

## Modellierung von Verteilungskoeffizienten

und ähnlich großen Wert wie bei der  $\log K_{fa}$ -Gleichung. Auch die Polarität und Polarisierbarkeit  $S$  und der logarithmische Hexadecan/Luft-Verteilungskoeffizient  $L$  besitzen sowohl das gleiche Vorzeichen, als auch einen ähnlichen Wert wie in der  $\log K_{fa}$ -Gleichung. Das verdeutlicht, dass sich die Verteilung zwischen Fett und Luft sehr gut mit der Olivenöl/Luft-Verteilung beschreiben lässt, welches sich auch im Modellierungserfolg in dieser Arbeit zeigt.

### 4.2 $K_{ba}$ -Modelle

Zur Beschreibung der Verteilung organischer Stoffe zwischen Blut und Luft werden verschiedene Ansätze verwendet. Diese sind ein auf intermolekularen Wechselwirkungen beruhender LSER-Ansatz, eine multilineare Regression mit anderen, bereits existierenden, Verteilungskoeffizienten und ein auf die chemische Struktur beruhender Fragmentansatz. Eine QSAR-Modellierung wird anhand von elektrotopologischen Indices durchgeführt. Es wird jeweils der Anwendungsbereich der Modelle untersucht, da dieser eine entscheidende Rolle für die Substanzauswahl spielt und die Eignung des Modells bestätigt, um verlässliche Vorhersagen zu treffen. Ein Vergleich der einzelnen Modelle untereinander und mit bereits existierenden Literaturmodellen wird angeführt.

#### 4.2.1 Modellierung von $\log K_{ba}$ mit Linearen Solvatations-Energie-Beziehungen

Wie in Kapitel 2.3 erläutert, existieren für die  $\log K_{ba}$  Modellierung einige LSER-Ansätze, die in der Literatur beschrieben sind. Bei dem Abraham-Modell (siehe Gleichungen (2.20) und (2.21), Kapitel 2.3.2, Seite 21) von Abraham *et al.* [26], können entweder experimentelle (exp.) oder berechnete (ber.) Deskriptoren (Desk.) verwendet werden. Die experimentellen Werte stammen aus der Datenbank von ChemProp [115,116]. Die berechneten Werte werden nach Platts *et al.* [143] bestimmt, wobei die Berechnung auch im Programm ChemProp [115,116] enthalten ist.

Wird das Modell (2.22) (Kapitel 2.3.2, Seite 23) mit den in der Literatur angegebenen, feststehenden Regressionsparametern nachgerechnet und beide Arten von Deskriptoren eingesetzt, ergibt sich das in Tabelle 4.1 dargestellte Ergebnis. Mit experimentellen Stoffparametern resultiert deutlich besseres Ergebnis als mit berechneten. So zeigt der Einsatz von experimentellen Deskriptoren eine Vorhersage- und Modellierungsgüte von  $r^2 = q^2 = 0,92$  und einen *rms*-Fehler von 0,31. Im Vergleich dazu wird mit berechneten Deskriptoren nur ein  $r^2$  von 0,68, ein  $q^2$  von 0,64 sowie ein *rms* von 0,68 erreicht. Auffallend ist auch der große

## Modellierung von Verteilungskoeffizienten

Wert der größten Überschätzung mit 5,90 bei einer Verwendung von berechneten Stoffparametern.

**Tabelle 4.1:** Statistische Parameter der nachgerechneten LSER-Gleichung für  $\log K_{ba}$  (2.22) mit experimentellen und berechneten Abraham-Deskriptoren.

Statistischer Parameter	exp. Desk.	ber. Desk.
$n$	191	191
$r^2$	0,92	0,68
$q^2$	0,92	0,64
$rms$	0,31	0,68
$bias$	-0,02	-0,01
$me$	0,24	0,42
$mne$	-0,91	-1,37
$mpe$	0,66	5,90

Wird das Modell (2.22) (Kapitel 2.3.2, Seite 23) in einer multiplen linearen Regression mit dem hier vorliegenden Datensatz (siehe Kapitel 3.2.1, Seite 35ff) und experimentellen LSER-Stoffparametern neu angepasst, ergibt sich Gleichung (4.4) mit entsprechenden Regressionsparametern.

$$\log K_{ba} = (-0,95 \pm 0,07) + (0,53 \pm 0,10) E + (1,03 \pm 0,12) S + (3,50 \pm 0,25) A + (2,61 \pm 0,13) B^H + (0,34 \pm 0,02) L \quad (4.4)$$

$$n = 191, r^2 = 0,92, q_{CV}^2 = 0,92, rms = 0,31, rms_{CV} = 0,32, bias = 0,00, me = 0,24, mne = -0,88, mpe = 0,68$$

Diese Regressionsparameter ähneln denen in der Literatur-Gleichung (2.22) (Kapitel 2.3.2, Seite 23). Die Auftragung der vorhergesagten (ber.)  $\log K_{ba}$ -Werte gegen die experimentellen (exp.  $\log K_{ba}$ ) ist in Abbildung 4.1 zu sehen und zeigt eine gute Vorhersage der experimentellen Werte durch die Anpassung der LSER-Gleichung mit experimentellen Deskriptoren. Dies belegen auch die statistischen Parameter (siehe Kapitel 3.5, Seite 50ff). Mit einem Bestimmtheitsmaß von 0,92 und einem sehr kleinen  $rms$  von 0,31 ist dieses Modell ein außerordentlich gutes Modell. Die Werte für das Bestimmtheitsmaß und den  $rms$  sind im Vergleich zu denen des nachgerechneten Abraham-Modells mit experimentellen Deskriptoren ähnlich (vgl. Tabelle 4.1). Dies ist jedoch nicht verwunderlich, da der Datensatz zu ca. 95 % dem aus der Literatur entspricht. Der  $bias$  von 0,00 ergibt sich durch die Anpassung und das  $q^2$  ist gleich dem  $r^2$ . Das ist ein Artefakt der Kleinste-Fehler-Anpassung, dabei wird

## Modellierung von Verteilungskoeffizienten

automatisch  $r^2 = q^2$ . Daher werden für alle folgenden, neu angepassten Modelle die Parameter  $q^2$  und  $bias$  nicht berücksichtigt.

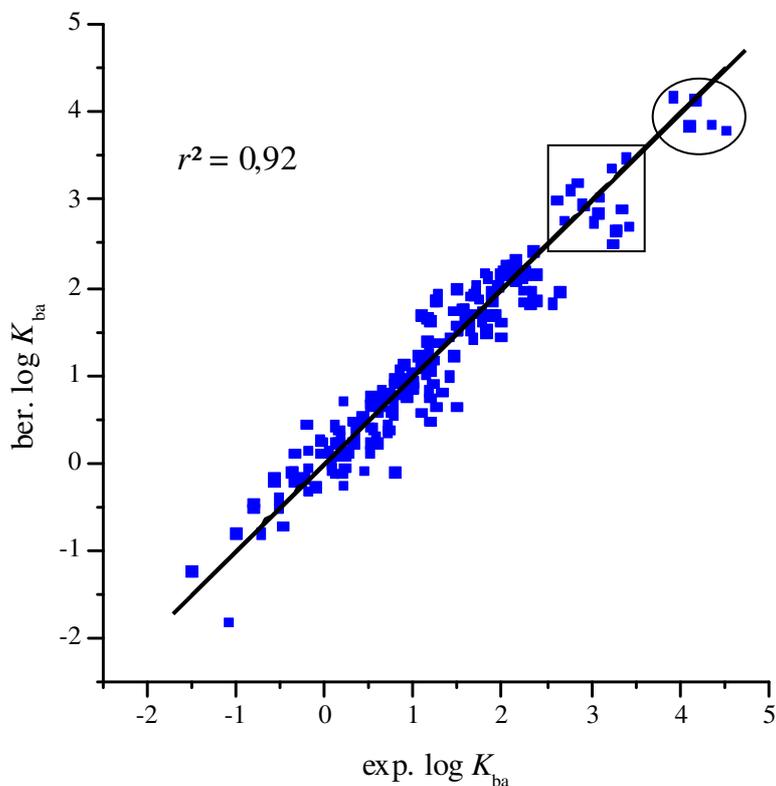


Abbildung 4.1: Diagramm der neu angepassten LSER-Gleichung (4.4) mit experimentellen Deskriptoren. Auftragung von vorhergesagtem gegen experimentellen  $\log K_{ba}$ .

Es ist zu erkennen, dass sich im oberen Bereich der Regressionsgeraden zwei Gruppen von Chemikalien (eingekreist und mit Viereck umgeben) abgrenzen, die (vor allem die eingekreiste) annähernd senkrecht zur  $xy$ -Geraden liegen. Es handelt sich hierbei um Glykolether (Abbildung 4.2) in der mit Kreis versehenen Gruppe und um Alkohole und die drei ringförmigen und sauerstoffhaltigen Verbindungen 1,4-Dioxan, Isophoron und Cyclohexanon (siehe Abbildung 4.3) in der mit Kästchen umgebenen Gruppe.

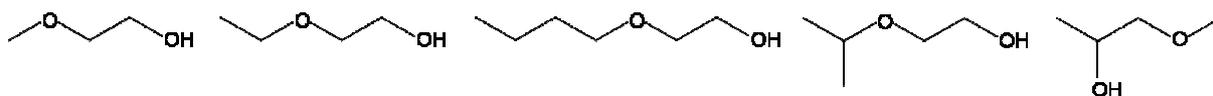


Abbildung 4.2: Strukturformeln der fünf Glykolether 2-Methoxyethanol, 2-Ethoxyethanol, 2-Butoxyethanol, 2-Isopropoxyethanol und 1-Methoxy-2-propanol.

## Modellierung von Verteilungskoeffizienten

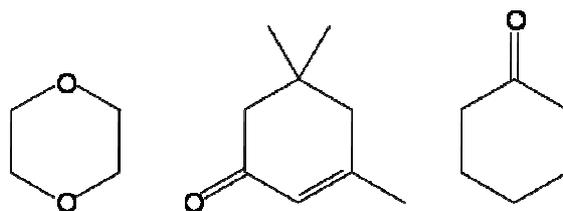


Abbildung 4.3: Strukturformeln der Ausreißer 1,4-Dioxan, Isophoron und Cyclohexanon.

Ohne die Gruppe der fünf Glykolether ergibt sich ein Bestimmtheitsmaß von 0,91 und ohne beide Gruppen ein Wert von 0,89.

Tabelle 4.2 zeigt den Vergleich zu den Literatur-Modellen von Abraham *et al.* [26] (Gleichung (4.1), Kapitel 4.1.1, Seite 54) und Sprunger *et al.* [57] (Gleichung (2.24), Kapitel 2.3.2, Seite 23). Der Datensatz für Mensch und Ratte von Abraham *et al.* [26] beinhaltet 5 Stoffe mehr. Dennoch ist die Güte des Modells fast gleich zu dem aus vorliegender Arbeit, während der *rms*-Fehler jedoch etwas schlechter ist. Dahingegen haben Sprunger *et al.* [57] einen kleineren Datensatz und zugleich sind sowohl das Bestimmtheitsmaß als auch der *rms* leicht besser.

**Tabelle 4.2:** Statistische Parameter der neu angepassten LSER-Gleichung (4.4) mit experimentellen Abraham-Deskriptoren im Vergleich zu zwei Literaturmodellen.

Statistischer Parameter	Vorliegende Arbeit	Abraham <i>et al.</i> [26]	Sprunger <i>et al.</i> [57]
$n$	191	196	146
$r^2$	0,92	0,94	0,95
$q_{CV}^2$	0,92	-	-
$rms$	0,31	0,32	0,27
$rms_{CV}$	0,32	-	-
$me$	0,24	-	-
$mne$	-0,88	-	-
$mpe$	0,68	-	-

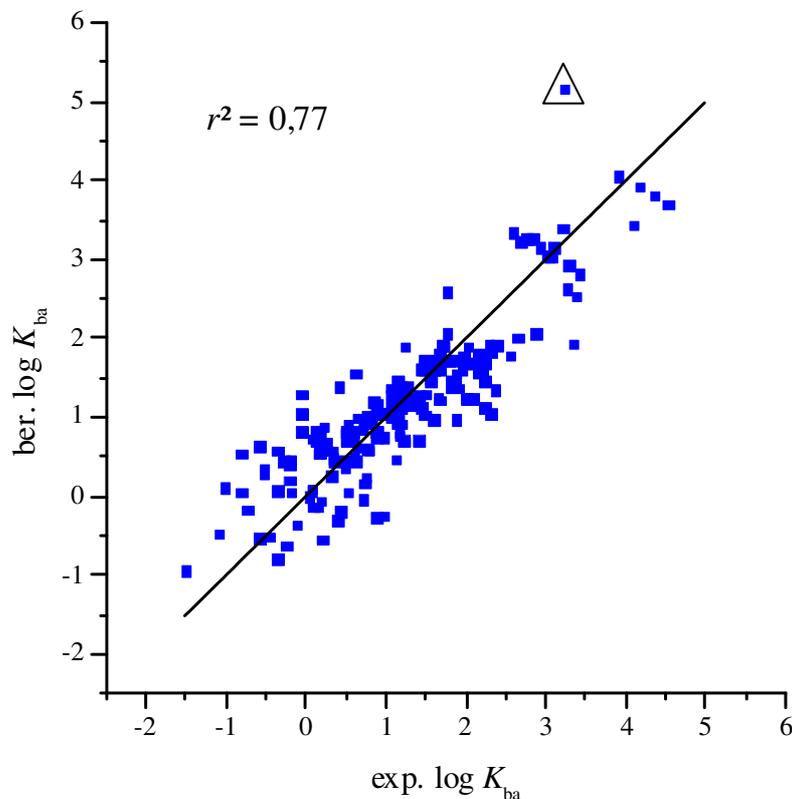
Da das Vorhersagemodell für den Blut/Luft-Verteilungskoeffizienten nicht mit Hilfe von experimentellen Deskriptoren angewendet werden, sondern allein mit vorhergesagten Hilfsgrößen ein gutes Ergebnis erzielen soll, werden nur berechnete Deskriptoren verwendet. Daraus ergibt sich folgende Regressionsgleichung (4.5) mit -parametern:

## Modellierung von Verteilungskoeffizienten

$$\log K_{ba} = (-0,33 \pm 0,12) + (1,30 \pm 0,20) E - (0,45 \pm 0,21) S + (4,77 \pm 0,47) A + (3,02 \pm 0,31) B^H + (0,23 \pm 0,05) L \quad (4.5)$$

$$n = 191, r^2 = 0,77, q_{CV}^2 = 0,68, rms = 0,54, rms_{CV} = 0,64, me = 0,41, mne = -1,38, mpe = 1,96$$

Die sich daraus ergebende Statistik ist deutlich schlechter als die mit experimentellen Deskriptoren. Die maximalen negativen und positiven Fehler haben Werte von über einer logarithmischen Einheit. Die dazugehörige Auftragung dieser Anpassung ist in Abbildung 4.4 gezeigt.



**Abbildung 4.4:** Auftragung von vorhergesagtem gegen experimentelles  $\log K_{ba}$  über die neu angepasste LSER-Gleichung (4.5) mit berechneten Deskriptoren.

Die erhaltenen Phasenparameter (außer  $l$ ) unterscheiden sich deutlich von den Phasenparametern aus der Anpassung mit experimentellen Deskriptoren. Die Polarität des Lösungsmittels  $s$  besitzt sogar ein anderes Vorzeichen. Die Statistikparameter zeigen deutlich, dass die Qualität des Modells im Vergleich zu dem mit experimentellen Deskriptoren abnimmt (siehe Tabelle 4.3).

## Modellierung von Verteilungskoeffizienten

**Tabelle 4.3:** Statistische Parameter der neu angepassten LSER-Gleichung mit experimentellen und berechneten Abraham-Deskriptoren.

Statistischer Parameter	exp. Desk.	ber. Desk.
$n$	191	191
$r^2$	0,92	0,77
$q_{CV}^2$	0,92	0,68
$rms$	0,31	0,54
$rms_{CV}$	0,32	0,64
$me$	0,24	0,41
$mne$	-0,88	-1,38
$mpe$	0,68	1,96

Zudem ist ein stark überschätzter Ausreißer mit dem  $mpe$ -Wert von 1,96 zu erkennen, der in Abbildung 4.4 mit einem Dreieck gekennzeichnet ist. Es handelt sich um den Tetracyanoethylenoxid (TCNEO), der in der nachfolgenden *Ausreißer-Betrachtung* näher erörtert wird. Ohne TCNEO ergibt sich mit berechneten LSER-Stoffparametern die Modellgleichung (4.6) und die dazugehörige untenstehende Statistik.

$$\log K_{ba} = (-0,64 \pm 0,12) + (0,81 \pm 0,20) E + (0,50 \pm 0,24) S + (4,30 \pm 0,44) A + (3,22 \pm 0,28) B^H + (0,24 \pm 0,04) L \quad (4.6)$$

$$n = 190, r^2 = 0,81, q_{CV}^2 = 0,79, rms = 0,50, rms_{CV} = 0,51, me = 0,37, mne = -1,68, mpe = 1,27$$

Der Anwendungsbereich des Modells für die Zielgröße  $\log K_{ba}$  ist bereits in Kapitel 3.2.1 (Seite 35ff) dargestellt. Die entsprechenden Werte der berechneten Abraham-Deskriptoren für dieses Modell erstrecken sich über die Bereiche für  $E$  von  $-0,66$  bis  $0,964$ ,  $S$  von  $-0,045$  bis  $1,45$ ,  $A$  von  $0,003$  bis  $0,348$ ,  $B$  von  $-0,099$  bis  $0,733$  sowie  $L$  von  $0,232$  bis  $5,762$ .

Um den Blut/Luft-Verteilungskoeffizienten kompetent vorherzusagen, ist der LSER-Ansatz somit ein annehmbares Mittel der Wahl. Mit experimentellen Deskriptoren können bereits gute Erfolge erzielt werden. Das ist jedoch weniger der Fall, wenn berechnete Deskriptoren verwendet werden. Zur Entwicklung eines Vorhersagemodells ist es entscheidend, dass nur berechnete Größen eingehen und keine experimentellen Werte gebraucht werden. Daher müssen weitere alternative Methoden gefunden werden.

## Modellierung von Verteilungskoeffizienten

### *Ausreißer-Betrachtung: Tetracyanoethylenoxid (TCNEO)*

Der Ausreißer Tetracyanoethylenoxid (TCNEO) (IUPAC Name: oxirane-2,2,3,3-tetracarbonitrile, CAS Nummer: 3189-43-3) ist der einzige Feststoff im  $K_{ba}$ -Datensatz und besitzt somit andere Eigenschaften als die anderen enthaltenen Substanzen.

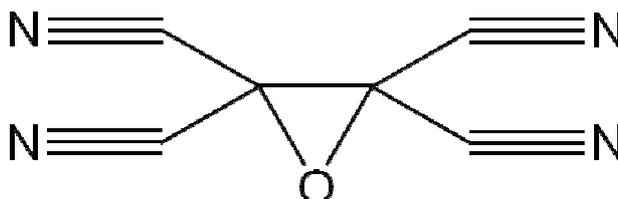


Abbildung 4.5: Strukturformel von Tetracyanoethylenoxid.

Dieser Ausreißer ist in der Anpassung mit experimentellen LSER-Deskriptoren (siehe Abbildung 4.1, Kapitel 4.2.1, Seite 60) deutlich besser vorhergesagt, da er im Vergleich zum Einsatz von berechneten Stoffparametern mit einem Fehler von  $-0,70$  nur leicht unterschätzt ist (Abbildung 4.4, Kapitel 4.2.1, Seite 62). Der Grund dafür ist, dass sich die experimentellen und berechneten Deskriptoren stark unterscheiden. Eine Ausnahme bildet  $A$ , das null oder annähernd null ist (siehe Tabelle 4.4). Es fällt auf, dass die experimentellen Deskriptoren immer kleiner sind als die berechneten. Der größte Unterschied tritt beim logarithmischen Hexadecan/Luft-Verteilungskoeffizienten  $L$  auf, bei dem jedoch die Phasenparameter sich nur leicht voneinander unterscheiden ( $0,34$  und  $0,23$ , Differenz  $0,11$ ).

Tabelle 4.4: LSER-Deskriptoren von TCNEO.

LSER Desk.	$E$	$S$	$A$	$B$	$L$
exp.	0,390	1,000	0,000	0,520	2,543
ber.	0,785	3,416	0,003	1,554	5,617
Unterschied	-0,395	-2,416	-0,003	-1,034	-3,074

Das lässt vermuten, dass die Vorhersage der Abraham-Deskriptoren für Feststoffe nicht geeignet ist. In der Arbeit von Abraham *et al.* [26], aus der die Daten genommen wurden, wird diese Substanz als flüchtiger Stoff eingestuft, was jedoch nicht korrekt ist. Die Autoren haben den Wert für den Blut/Luft-Verteilungskoeffizienten wiederum von Gargas *et al.* [144], die diesen gemessen haben.

## Modellierung von Verteilungskoeffizienten

Bei dieser Detailanalyse des Ausreißers TCNEO wird untersucht, welche weiteren Eigenschaften den Unterschied zu den anderen Stoffen machen. TCNEO ist ein Feststoff unter Normalbedingungen mit einem Schmelzpunkt von 176 °C [145]. In Abbildung 4.4 (Kapitel 4.2.1, Seite 62) ist z. B. deutlich zu sehen, dass sich der Stoff anders verhält als die anderen Substanzen. Wie in Tabelle 8.3 im Anhang 8.1.2 (Seite 165ff) zu sehen ist, besitzt TCNEO einen sehr niedrigen  $\log K_{mw}$ -Wert von  $-3,58$ , wohingegen der  $\log K_{mw}$ -Bereich aller anderen Stoffe von  $-0,22$  bis  $5,23$  geht. Möglicherweise können hier die gerechneten Abraham-Eingangsparemeter fehlerhaft sein. Dieser Fakt kann am besten überprüft werden, indem die berechneten mit experimentellen Abraham-Parametern von strukturähnlichen Substanzen oder zumindest von Substanzen mit entsprechenden oder ähnlichen funktionellen Gruppen verglichen werden. Wird die Gegenüberstellung der experimentellen LSER-Daten von 15 ausgewählten Substanzen, die Cyanidgruppen enthalten und in der Software ChemProp [115,116] implementiert sind, mit den berechneten und experimentellen Werte von TCNEO durchgeführt, ist ersichtlich, dass die experimentellen Deskriptoren übereinstimmen, jedoch die berechneten TCNEO-Deskriptoren bis auf  $A$  alle sehr viel größer als die experimentellen Stoffparameter sind. Somit kann angenommen werden, dass die berechneten Deskriptoren fehlerhaft sind und diese nur mit Vorsicht verwendet werden können. Die vorhandenen experimentellen Abraham-Parameter für TCNEO werden in Tabelle 4.4 in Kapitel 4.2.1 mit den berechneten Parametern verglichen. Zusätzlich besitzt TCNEO eine ungewöhnliche inverse Henry-Konstante mit einem hohen  $\log K_{wa}$ -Wert von  $8,20$  (berechnet mit EPI Suite [117,118], siehe Kapitel 3.3.2, Seite 45f), während der Bereich aller anderen Stoffe im Datensatz zwischen  $-3,54$  und  $5,77$  liegt. Ein experimenteller Wert ist dagegen nicht vorhanden. Wird dieser Verteilungskoeffizient für TCNEO mit der LSER-Gleichung (3.3) (Seite 45) mit berechneten Abraham-Deskriptoren ermittelt, ergibt sich der logarithmische Wert  $14,87$  und mit dem Fragment-Modell, welches in ChemProp [115,116] verfügbar ist, ein ähnlicher Wert von  $15,41$ . Das bedeutet, dass das LSER-Modell mit dem Fragment-Modell stärker übereinstimmt als mit EPI Suite. Auch für den Oktanol/Wasser-Verteilungskoeffizienten (Kapitel 3.3.3, Seite 46) ist kein experimenteller Wert verfügbar. Die Vorhersage mit der LSER-Gleichung (3.4) (Seite 46) mit berechneten Abraham-Deskriptoren ergibt einen Wert von  $-4,89$ , während die Berechnung mit EPI Suite  $0,52$  liefert. Die Werte unterscheiden sich somit um mehr als 5 Größenordnungen. Das könnte ein weiterer Hinweis auf schlechte oder gar falsche Abraham-Eingangsparemeter sein.

## Modellierung von Verteilungskoeffizienten

TCNEO wird nachfolgend aus dem Datensatz entfernt und alle weiteren Modellierungsansätze werden ohne den Stoff TCNEO durchgeführt. Die Datensatzgröße beläuft sich somit auf 190 Stoffe.

### 4.2.2 Modellierung von $\log K_{ba}$ mit anderen Verteilungskoeffizienten

Die Modellierung des Blut/Luft-Verteilungskoeffizienten mit anderen Verteilungskoeffizienten erfolgt nach Tabelle 2.1 aus Kapitel 2.2.1 (Seite 11). Da Blut definitiv aus Wasser und weiteren Komponenten, wie Lipide, Proteine *etc.*, besteht, wird zuerst eine Regression zwischen  $\log K_{ba}$  vs.  $\log K_{wa}$  betrachtet (Modell Nr. 1). Die zugehörige lineare Regression ist in Gleichung (4.7) ( $\log K_{wa}$ : EPI Suite) dargestellt.

$$\log K_{ba} = (0,53 \pm 0,02) \cdot \log K_{wa} + (0,96 \pm 0,04) \quad (4.7)$$

$$n = 190, r^2 = 0,72, q_{CV}^2 = 0,72, rms = 0,59, rms_{CV} = 0,59, me = 0,46, mne = -1,27, mpe = 1,79$$

Die inverse, dimensionslose Henry-Konstante  $K_{wa}$ , berechnet mit dem Modell HENRYWIN [117] aus der Software EPI Suite [118] (siehe Kapitel 3.3.2), wird mit der Zielgröße der Blut/Luft-Verteilung korreliert (Abbildung 4.6). Die Gerade durch die Datenpunkte ist die xy-Gerade.

## Modellierung von Verteilungskoeffizienten

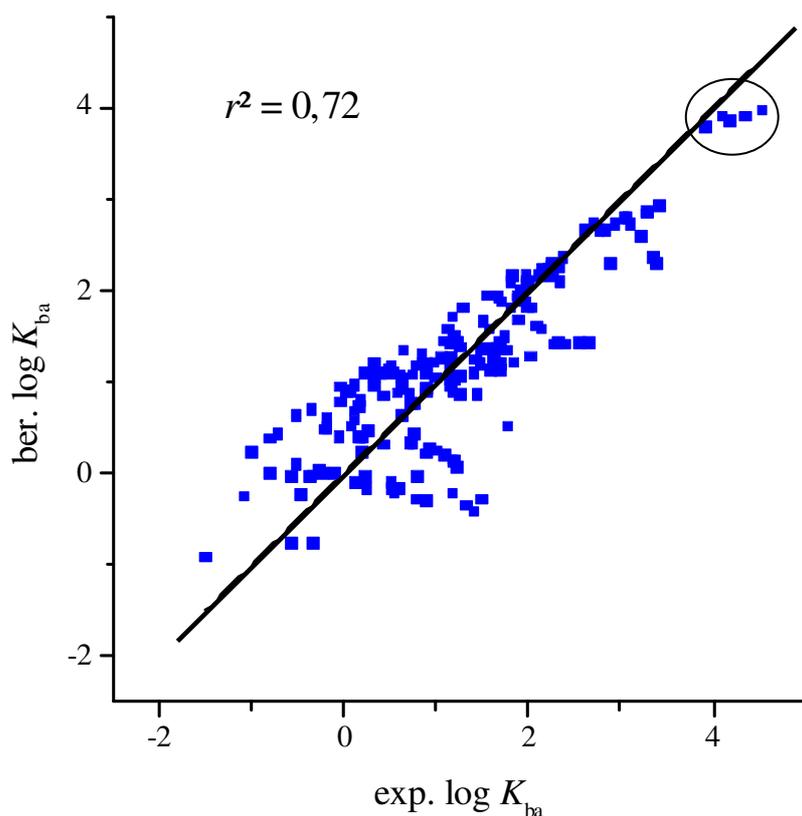


Abbildung 4.6: Auftragung von vorhergesagtem gegen experimentelles  $\log K_{ba}$  nach Gleichung (4.7).

Wie vermutet, ist eine deutliche Korrelation zwischen Blut und Wasser zu Luft Verteilungen zu erkennen und das Modell Nr. 1 eine gute Grundlage, um die Modellierung mit anderen Verteilungskoeffizienten zu erweitern. Das Bestimmtheitsmaß beträgt jedoch nur 0,72 und der *rms* 0,59, was für ein zuverlässiges Modell von hoher Güte nicht ausreichend ist, zumal die Blut/Luft-Modelle auf LSER-Basis eine deutlich bessere Statistik aufweisen (siehe Kapitel 2.3, Gleichungen (2.22) und (2.24), Seite 23). Eine Gruppe von fünf Stoffen ist dabei auffällig, die am äußersten rechten und oberen Ende der xy-Geraden zu finden ist, jedoch gut vorhergesagt wird. Diese Stoffe sind in Abbildung 4.6 eingekreist. Hierbei handelt es sich um die Glykolether 2-Methoxyethanol, 2-Ethoxyethanol, 2-Butoxyethanol, 2-Isopropoxyethanol und 1-Methoxy-2-propanol (siehe Abbildung 4.2, Seite 60). Die Gruppierung folgt bereits aus den experimentellen Daten. Die Vorhersage der Alkohole mit  $\log K_{wa}$  ist für das Blut/Luft-Modell präzise.

Aus diesem Ergebnis und der Überlegung, dass Blut neben Wasser auch aus organischen Komponenten besteht (siehe Kapitel 2.2.1, Seite 9ff), wird die Größe des Oktanol/Wasser-Verteilungskoeffizienten (siehe Kapitel 3.3.3, Seite 48), hinzugezogen. Es wird demnach der

## Modellierung von Verteilungskoeffizienten

organische Teil des Blutes mit Oktanol beschrieben. Nach Tabelle 2.1 (Kapitel 2.2.1, Seite 11) ist die Modellierung des Blut/Luft-Verteilungskoeffizienten mit einer multilinenen Regression von  $K_{wa}$  und  $K_{ow}$  Modell Nr. 2. Der Verteilungskoeffizient  $K_{ow}$  dient hier zusätzlich zum Wasser als Fett- oder Proteinrepräsentant. Aus der Modellierung des Blut/Luft-Verteilungskoeffizienten mit Wasser/Luft- und Oktanol/Wasser-Verteilung resultiert in Gleichung (4.8) beschriebener multilinearer Zusammenhang.

$$\log K_{ba} = (0,71 \pm 0,02) \cdot \log K_{wa} + (0,38 \pm 0,03) \cdot \log K_{ow} + (0,04 \pm 0,09) \quad (4.8)$$

$$n = 190, r^2 = 0,83, q_{CV}^2 = 0,83, rms = 0,46, rms_{CV} = 0,47, me = 0,36, mne = -1,06, mpe = 1,79$$

Gleichung (4.8) bildet Modell Nr. 2 aus Tabelle 2.1 (Kapitel 2.2.1, Seite 11) mit  $\log K_{wa}$  (HENRYWIN [117] aus Epi Suite [118]) und  $\log K_{ow}$  (KOWWIN [121] aus EPI Suite [118]) als Deskriptoren ab. Die Konstante  $c$  ist mit einem p-Wert von 0,62 nicht signifikant (siehe Kapitel 3.4) und besitzt einen größeren Standardfehler als der Wert des Regressionsparameters selbst. Werden nun die aus Gleichung (4.8) berechneten Werte gegen die experimentellen  $\log K_{ba}$ -Werte aufgetragen, ergibt sich der Graph, der in Abbildung 4.7 dargestellt ist. Die Statistik ist im Vergleich zum 1-Parameter-Modell (Gleichung (4.7), Seite 66) deutlich besser. Das Bestimmtheitsmaß ist mit einem Wert von 0,83 deutlich verbessert. Auch hier ist die xy-Gerade aufgezeigt.

## Modellierung von Verteilungskoeffizienten

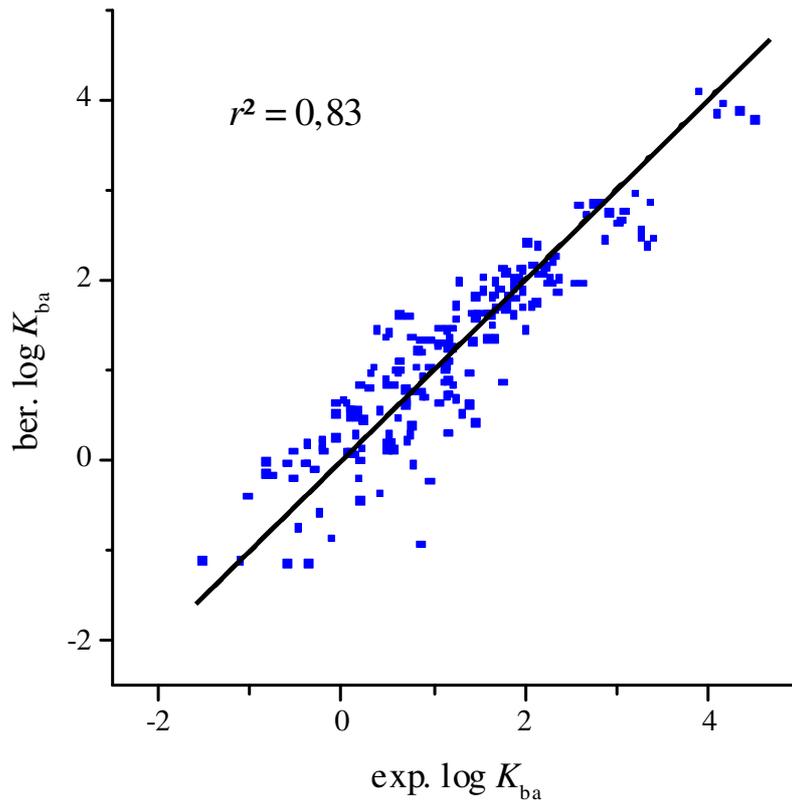


Abbildung 4.7: Diagramm von berechnetem  $\log K_{ba}$  über Gleichung (4.8) gegen experimentelles  $\log K_{ba}$ .

Die Gruppe der Alkohole wird jedoch mit dieser Regression schlechter beschrieben als im Modell davor (Gleichung (4.7), Seite 66). Somit kann gefolgert werden, dass Alkohole in diesem Datensatz mit Hinzunahme von  $\log K_{ow}$  schlechter modelliert werden als ohne die Oktanol/Wasser-Verteilung.

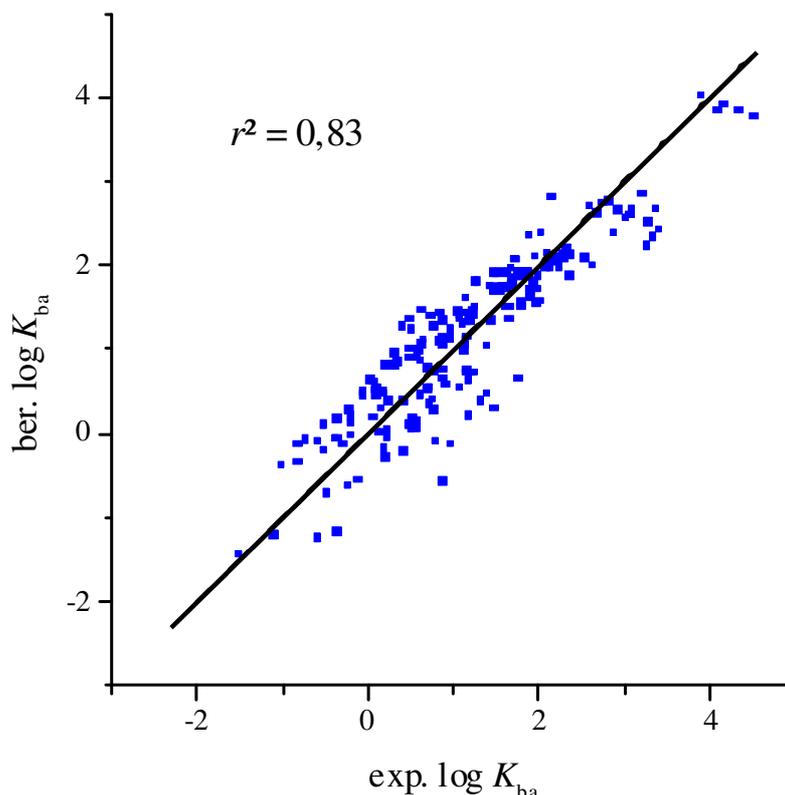
Insgesamt ist das Modell noch nicht zufriedenstellend, weshalb mit dem Humanalbumin/Wasser-Verteilungskoeffizient, der den Proteinanteil im Blut modellieren soll, ein weiterer Deskriptor betrachtet wird. In Tabelle 2.1 (Kapitel 2.2.1, Seite 11) ist die Modellierung des Blut/Luft-Verteilungskoeffizienten mit einer multilinearen Regression von  $\log K_{wa}$  (HENRYWIN [117] aus Epi Suite [118]) und  $\log K_{HSA}$  (LSER) Modell Nr. 3.

$$\log K_{ba} = (0,64 \pm 0,02) \cdot \log K_{wa} + (0,71 \pm 0,06) \cdot \log K_{HSA} + (1,09 \pm 0,04) \quad (4.9)$$

$$n = 190, r^2 = 0,83, q_{CV}^2 = 0,83, rms = 0,45, rms_{CV} = 0,46, me = 0,37, mne = -0,89, mpe = 1,44$$

## Modellierung von Verteilungskoeffizienten

Die mit Gleichung (4.8) vorhergesagten Werte werden gegen die experimentellen  $\log K_{ba}$ -Werte aufgetragen (Abbildung 4.8). Die Statistik ist im Vergleich zum vorherigen Modell (Gleichung (4.8)(4.7)) ähnlich, nur die *rms*-Fehler sind leicht besser. Erneut ist die xy-Gerade durch die Datenpunkte gezeichnet.



**Abbildung 4.8:** Diagramm von berechnetem  $\log K_{ba}$  über Gleichung (4.9) gegen experimentelles  $\log K_{ba}$ .

Der organische Teil im Blut zur Modellierung von  $\log K_{ba}$  kann auch durch ein Membranmodell beschrieben werden, da der Stoffaustausch von der Lunge in die Blutbahn durch eine sogenannte alveolokapilläre Membran erfolgt. Die Membran spielt damit bei der Verteilung eine wichtige Rolle. Sowohl die Blutgefäßzellen als auch die Blutkörperchen besitzen Zellmembranen, in die der Fremd- und Wirkstoff aus dem wässrigen Blutserum diffundieren kann. Dieser Diffusionsvorgang ist auch umgekehrt möglich. Daher gilt der Membran/Wasser-Verteilungskoeffizient als sinnvoller Deskriptor bei der Modellierung des Blut/Luft-Verteilungskoeffizienten, wie auch schon in Kapitel 2.2.1 in Tabelle 2.1 (Seite 11) dargestellt. Bei der linearen Regression von  $\log K_{wa}$  (HENRYWIN [117] aus Epi Suite [118]) und  $\log K_{mw}$  (LSER) (siehe Gleichung (4.10)) ergibt sich das Modell Nr. 4 aus Tabelle 2.1.

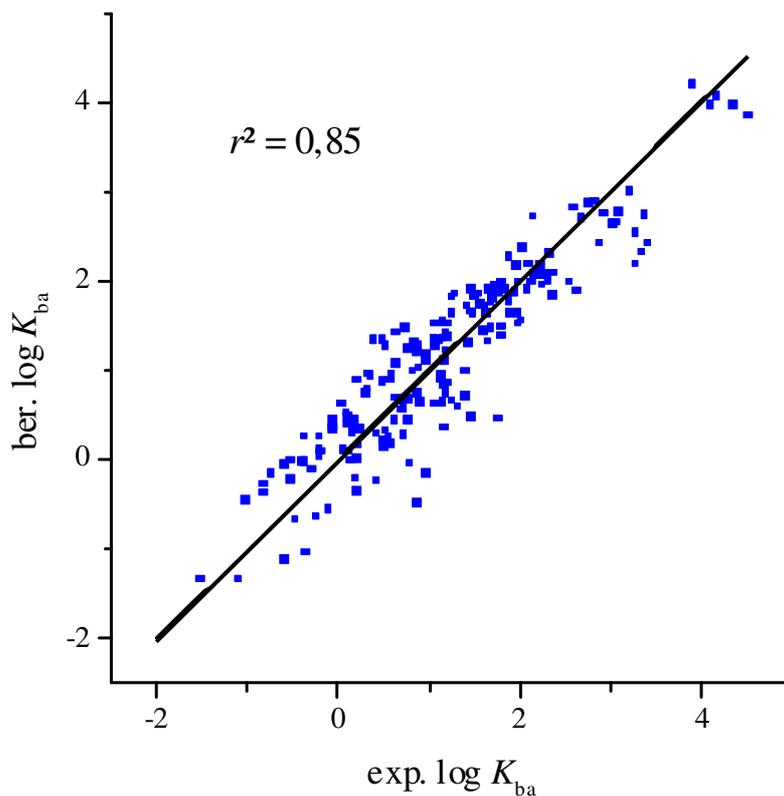
## Modellierung von Verteilungskoeffizienten

Dieses Modell zeigt leicht bessere statistische Parameter als die vorangegangenen Modellen die  $\log K_{ow}$  und  $\log K_{HSA}$  nutzten (Gleichung (4.8) und Gleichung (4.9)). Möglicherweise kann der organische Teil im Blut mit  $K_{mw}$  besser beschrieben werden als mit den anderen beiden Verteilungskoeffizienten. Die Regressionskonstante  $c$  ist nicht signifikant.

$$\log K_{ba} = (0,67 \pm 0,02) \cdot \log K_{wa} + (0,49 \pm 0,04) \cdot \log K_{mw} - (0,11 \pm 0,09) \quad (4.10)$$

$$n = 190, r^2 = 0,85, q_{CV}^2 = 0,84, rms = 0,44, rms_{CV} = 0,44, me = 0,36, mne = -0,97, mpe = 1,34$$

Die daraus resultierende Darstellung mit Gleichung (4.10) berechneten gegen experimentelle  $\log K_{ba}$ -Werte ist in Abbildung 4.9 zu sehen.



**Abbildung 4.9:** Auftragung von berechneten  $\log K_{ba}$ -Werte über Gleichung (4.10) gegen experimentelle Werte.

Als weitere Hilfsgröße zur Beschreibung des organischen Teils im Blut kann der Olivenöl/Luft-Verteilungskoeffizient eingesetzt werden. So wird der Blut/Luft-Verteilungskoeffizient mit einer multilinenen Regression von  $\log K_{wa}$  (HENRYWIN [117]) aus Epi Suite

## Modellierung von Verteilungskoeffizienten

[118]) und  $\log K_{\text{olive}}$  (LSER) modelliert. In Tabelle 2.1 (Kapitel 2.2.1, Seite 11) wird dieses Modell (Gleichung (4.11)) als Modell Nr. 5 bezeichnet.

$$\log K_{\text{ba}} = (0,71 \pm 0,02) \cdot \log K_{\text{wa}} + (0,34 \pm 0,03) \cdot \log K_{\text{olive}} + (0,10 \pm 0,08) \quad (4.11)$$

$$n = 190, r^2 = 0,84, q_{\text{CV}}^2 = 0,83, rms = 0,45, rms_{\text{CV}} = 0,46, me = 0,36, mne = -1,09, mpe = 1,38$$

Die Konstante ist mit einem p-Wert von 0,23 nicht signifikant. Im Statistikvergleich mit anderen Modellen schneidet das Modell Nr. 5 etwas schlechter ab.

Zusätzlich zu den 2-Parameter-Modellen können auch 3-Parameter-Modelle getestet werden, bei denen der Wasseranteil im Blut wieder mit  $K_{\text{wa}}$  vorhergesagt wird und der organische Teil mit zwei der Hilfsgrößen, beispielsweise mit  $K_{\text{ow}}$  und  $K_{\text{HSA}}$ . Dadurch wird der organische Teil durch Oktanol und das Protein Humanalbumin dargestellt. In Tabelle 2.1 (Kapitel 2.2.1, Seite 11) handelt es sich um Modell Nr. 6 ( $\log K_{\text{wa}}$ : HENRYWIN [117] aus Epi Suite [118],  $\log K_{\text{ow}}$ : KOWWIN [121] aus EPI Suite [118],  $\log K_{\text{HSA}}$ : LSER), das hier in Gleichung (4.12) gezeigt ist.

$$\log K_{\text{ba}} = (0,67 \pm 0,03) \cdot \log K_{\text{wa}} + (0,17 \pm 0,10) \cdot \log K_{\text{ow}} + (0,41 \pm 0,19) \cdot \log K_{\text{HSA}} + (0,62 \pm 0,27) \quad (4.12)$$

$$n = 190, r^2 = 0,84, q_{\text{CV}}^2 = 0,83, rms = 0,45, rms_{\text{CV}} = 0,47, me = 0,37, mne = -0,98, mpe = 1,61$$

Die Statistik ist bei den 3-Parameter-Modellen nicht besser als bei den 2-Parameter-Modellen und der Deskriptor  $\log K_{\text{ow}}$  ist mit einem p-Wert von 0,09 nicht signifikant. Daher ist keine Verbesserung zu beobachten und daher ist dieses Modell nicht zu empfehlen.

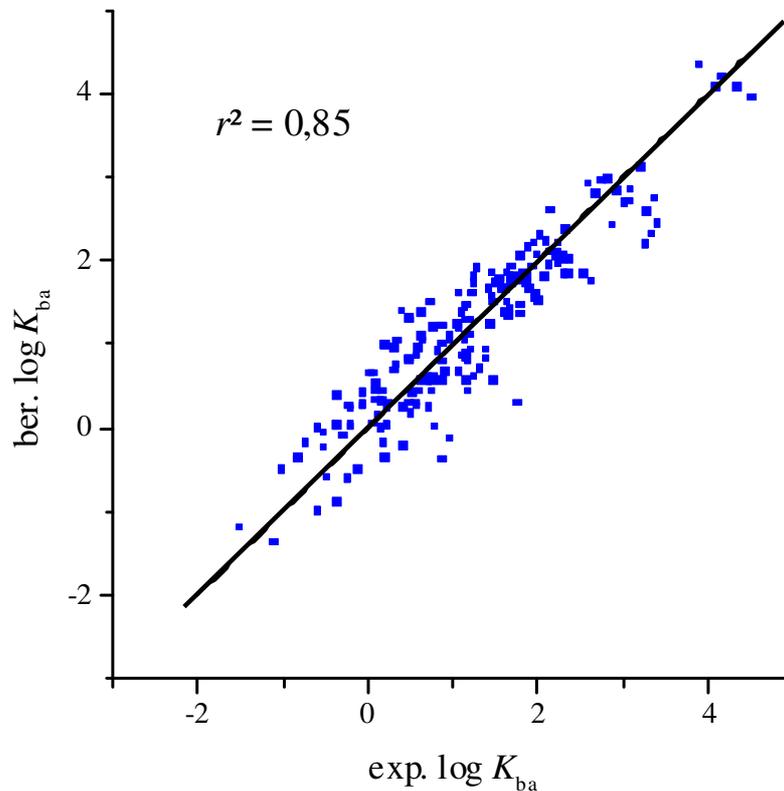
Modell Nr. 7 aus Tabelle 2.1 (Kapitel 2.2.1, Seite 11) ist ebenfalls ein 3-Parameter-Modell, bei dem die organischen Blutkomponenten mit Protein und Membran beschrieben werden. Die Modellierung findet daher mit dem Humanalbumin/Wasser-Verteilungskoeffizienten und dem Membran/Wasser-Verteilungskoeffizienten statt. Der Wasserteil wird wieder mit  $K_{\text{wa}}$  vorhergesagt.

## Modellierung von Verteilungskoeffizienten

$$\log K_{ba} = (0,68 \pm 0,02) \cdot \log K_{wa} - (0,66 \pm 0,29) \cdot \log K_{HSA} + (0,90 \pm 0,19) \cdot \log K_{mw} - (1,14 \pm 0,46) \quad (4.13)$$

$$n = 190, r^2 = 0,85, q_{CV}^2 = 0,85, rms = 0,43, rms_{CV} = 0,44, me = 0,34, mne = -1,01, mpe = 1,45$$

Gleichung (4.13) zeigt dieses Modell mit den Parametern  $\log K_{wa}$  (HENRYWIN [117] aus Epi Suite [118]),  $\log K_{HSA}$  (LSER) und  $\log K_{mw}$  (LSER). Die Fehler sind hierzu wieder kleiner, die Statistik allgemein sogar besser als die der 2-Parameter-Modelle und alle Parameter sind signifikant. Die Fehler  $rms$  und  $me$  sind im Vergleich zu den anderen Modellen am kleinsten. Dies deutet darauf, dass durch die Hinzunahme eines dritten Koeffizienten die Vorhersage weiter verbessert werden kann.



**Abbildung 4.10:** Auftragung von berechneten  $\log K_{ba}$ -Werte über Gleichung (4.13) gegen experimentelle Werte.

Die grafische Lösung der berechneten  $\log K_{ba}$ -Werte gegen die experimentellen Werte mit Gleichung (4.13) ist in Abbildung 4.10 gezeigt. Weiterhin bleibt die Gruppierung der Glykolether (siehe Abbildung 4.2, Seite 60) am rechten oberen Rand bestehen. Die Linie durch die aufgetragenen Werte ist die xy-Gerade, die die bestmögliche Vorhersage darstellt.

## **Modellierung von Verteilungskoeffizienten**

Eine tabellarische Zusammenfassung der Modelle zur Beschreibung des Blut/Luft-Verteilungskoeffizienten mit anderen Verteilungskoeffizienten ist in Tabelle 4.5 gezeigt.

## Modellierung von Verteilungskoeffizienten

**Tabelle 4.5: Modelle, Regressions- und Statistikparameter für log  $K_{ba}$ -Modelle nach dem Schema  $a \cdot \log A + b \cdot \log B + c \cdot \log C + d$ .**

Modell Nr.	Verteilungskoeffizienten	$a$	$b$	$c$	$d$	$r^2$	$q_{CV}^2$	$rms$	$rms_{CV}$	$me$	$mne$	$mpe$
1	$A = K_{wa}$	$0,53 \pm 0,02$	-	-	$0,96 \pm 0,04$	0,72	0,72	0,59	0,59	0,46	-1,27	1,79
2	$A = K_{wa}, B = K_{ow}$	$0,71 \pm 0,02$	$0,38 \pm 0,03$	-	$0,04^* \pm 0,09$	0,83	0,83	0,46	0,47	0,36	-1,06	1,79
3	$A = K_{wa}, B = K_{HSA}$	$0,64 \pm 0,02$	$0,71 \pm 0,06$	-	$1,09 \pm 0,04$	0,83	0,83	0,45	0,46	0,37	-0,89	1,44
4	$A = K_{wa}, B = K_{mw}$	$0,67 \pm 0,02$	$0,49 \pm 0,04$	-	$-0,11^* \pm 0,09$	0,85	0,84	0,44	0,44	0,36	-0,97	1,34
5	$A = K_{wa}, B = K_{olive}$	$0,71 \pm 0,02$	$0,34 \pm 0,03$	-	$0,10^* \pm 0,08$	0,84	0,83	0,45	0,46	0,36	-1,09	1,38
6	$A = K_{wa}, B = K_{ow}, C = K_{HSA}$	$0,67 \pm 0,03$	$0,17^* \pm 0,10$	$0,41 \pm 0,19$	$0,62 \pm 0,27$	0,84	0,83	0,45	0,47	0,37	-0,98	1,61
7	$A = K_{wa}, B = K_{HSA}, C = K_{mw}$	$0,68 \pm 0,02$	$-0,66 \pm 0,29$	$0,90 \pm 0,19$	$-1,14 \pm 0,46$	0,85	0,85	0,43	0,44	0,34	-1,01	1,45

\* nicht signifikant

## Modellierung von Verteilungskoeffizienten

Ausgehend von der Überlegung, dass die Hauptkomponente im Blut Wasser ist, wird dafür der Verteilungskoeffizient  $K_{wa}$  zur Modellierung eingesetzt, was in Tabelle 4.5 als Modell Nr. 1 bezeichnet wird. Dieses Modell wird durch Hinzunahme eines zweiten Deskriptors verbessert. Dieser beschreibt die organische Komponente im Blut. Analog der Tabelle 2.1 (Kapitel 2.2.1, Seite 11) werden dafür versuchsweise die Hilfsgrößen  $K_{ow}$ ,  $K_{HSA}$ ,  $K_{mw}$  und  $K_{olive}$  eingesetzt und stellen die Modelle Nr. 2 bis 5 dar. Diese Modelle zeichnen sich durch geringere Fehler und bessere Vorhersageparameter als das 1-Deskriptor-Modell aus, was durch die Verwendung eines zweiten Parameters erklärbar ist. Damit wird nun sowohl der Wasserteil als auch der organische Teil im Blut beschrieben. Untereinander sind die Statistiken dieser Modelle sehr ähnlich, wobei der Einsatz des Membran/Wasser-Verteilungskoeffizienten die geringsten Fehler ( $me = 0,36$  und  $rms = 0,44$ ) aufweist und Modell Nr. 4 statistisch ein wenig besser abschneidet als die anderen 2-Parameter-Modelle. Das bedeutet, dass der organische Teil im Blut am besten mit der Hilfsgröße „Membran“ dargestellt wird, was sich damit erklären lässt, dass z. B. die Blutkörperchen wie alle Zellen mit einer Zellmembran umgeben sind und darin Proteine in Form von Membranproteinen vorliegen. Dadurch ist der  $K_{mw}$  eine bedeutende und sinnvolle Ergänzung zu  $K_{wa}$  um zusätzlich zum Wasserteil auch den organischen Teil im Blut gut modellieren zu können. Durch Hinzunahme eines weiteren Deskriptors wird versucht, den organischen Teil im Blut genauer zu beschreiben und eine Modellverbesserung zu erzielen. Bei Modell Nr. 6 (Tabelle 4.5) werden dafür  $K_{ow}$  und  $K_{HSA}$  verwendet. Der Deskriptor  $K_{ow}$  ist in diesem 3-Parameter-Modell nicht signifikant und es ergibt sich keine Verbesserung. Das bedeutet, dass die Kombination aus Oktanol und dem Protein Humanalbumin den organischen Teil im Blut nicht besser beschreiben als mit den einzelnen Deskriptoren. Der Oktanol/Wasser-Verteilungskoeffizient besitzt einen ähnlichen Informationsgehalt wie  $K_{HSA}$  und deshalb kann keine Verbesserung erzielt werden. Wird hingegen der Humanalbumin/Wasser-Verteilungskoeffizient als zweite Hilfsgröße zu  $K_{mw}$  zur Beschreibung des organischen Teils im Blut genutzt (siehe Modell Nr. 7 in Tabelle 4.5, Seite 75), wird der organische Teil im Blut mit den Hilfsgrößen „Membran“ und „Protein“ dargestellt und das Modell minimal verbessert. Dieses Modell nutzt als einziges einen Verteilungskoeffizienten mit einem negativen Regressionsparameter, nämlich der für  $\log K_{HSA}$  beträgt mit  $-0,66$ , der zusätzlich einen relativ großen Fehler von  $0,29$  hat. Der Anwendungsbereich des Modells Nr. 7 aus Tabelle 4.5 (Seite 75) ist durch die Wertebereiche seiner drei Deskriptoren, deren

## Modellierung von Verteilungskoeffizienten

logarithmische Werte sich für  $\log K_{wa}$  von  $-3,537$  bis  $8,198$  und für  $\log K_{HSA}$  von  $-3,373$  bis  $1,455$  sowie für den Membran/Wasser-Verteilungskoeffizienten von  $-3,580$  bis  $5,226$  erstrecken, gegeben. Bei diesem Modell werden  $K_{HSA}$  und  $K_{mw}$  mit Abraham-Eingangsparameter berechnet (siehe Kapitel 3.3.6, Seite 48 und Kapitel 3.3.5, Seite 47) und  $K_{wa}$  mit EPI Suite [117,118]. Wird das gesamte Modell mit LSER-Parametern gerechnet und somit  $K_{wa}$  über die LSER-Gleichung (3.3) ermittelt, ergibt sich eine deutlich schlechtere Statistik ( $r^2 = 0,65$ ,  $rms = 0,67$ ). Das zeigt deutlich, dass die Güte des Modells durch die unterschiedliche Herkunft der Eingangsparameter beeinflusst werden kann. Umso wichtiger ist es, dass die Qualität der Deskriptorenberechnung hochwertig ist.

### 4.2.2.1 Zusammengefasste LSER-Gleichungen

Da die Verteilungskoeffizienten für diese Modelle, wie sie in Kapitel 4.2.2 (Seite 66ff) beschrieben sind, alle mit Abraham-Deskriptoren berechnet werden können, besteht die Möglichkeit diese als ein einzelnes, neues Abraham-Modell zu beschreiben. Als Beispiel für diese Zusammenfassung wird das Modell Nr. 3 aus Tabelle 4.5 ( $K_{wa}$ : HENRYWIN [117] aus Epi Suite [118],  $K_{HSA}$ : LSER) hergenommen (siehe Gleichung (4.14)).

$$\log K_{ba} = (0,64 \pm 0,02) \cdot \log K_{wa} + (0,71 \pm 0,06) \cdot \log K_{HSA} + (1,09 \pm 0,04) \quad (4.14)$$

$$n = 190, r^2 = 0,83, q_{CV}^2 = 0,83, rms = 0,45, rms_{CV} = 0,46, me = 0,37, mne = -0,89, mpe = 1,44$$

$K_{wa}$  wird für das Modell im vorangegangenen Kapitel und in Gleichung (4.14) mit HENRYWIN [117] aus der Software EPI Suite [118] erhalten, kann aber auch über eine LSER-Gleichung berechnet werden (siehe Kapitel 3.3.2). Danach ergibt sich Gleichung (4.15) ( $K_{wa}$ : LSER,  $K_{HSA}$ : LSER).

$$\log K_{ba} = (0,78 \pm 0,03) \cdot \log K_{wa} + (0,67 \pm 0,08) \cdot \log K_{HSA} + (0,94 \pm 0,05) \quad (4.15)$$

$$n = 190, r^2 = 0,74, q_{CV}^2 = 0,73, rms = 0,57, rms_{CV} = 0,58, me = 0,42, mne = -2,40, mpe = 1,33$$

Das Ergebnis des  $K_{ba}$ -Modells ändert sich dadurch erheblich. Es ist mit einem Bestimmtheitsmaß von  $0,74$  im Gegensatz zu  $0,83$  sowie einem  $rms$  von  $0,57$  im Gegensatz zu  $0,45$  deutlich schlechter. Dadurch zeigt sich, dass mit verschiedenen Quellen der

## Modellierung von Verteilungskoeffizienten

Eingangsgrößen eine unterschiedliche Vorhersagegüte resultiert und das Problem der Deskriptoren-Berechnung dieser Stoffe nicht entscheidend gelöst werden kann.

Werden nun für die logarithmierten Verteilungskoeffizienten  $\log K_{wa}$  und  $\log K_{HSA}$  die LSER-Gleichungen (3.3) und (3.8) eingesetzt, folgt Gleichung (4.16) für den Blut/Luft-Verteilungskoeffizienten.

$$\begin{aligned} \log K_{ba} = & 0,78 \cdot (-0,994 + 0,577 E + 2,549 S + 3,813 A + 4,841 B^H - 0,869 V) \\ & + 0,67 \cdot (-1,28 + 0,82 E - 0,36 S + 0,18 A - 1,97 B^H + 1,62 V) + 0,94 \end{aligned} \quad (4.16)$$

Nach Ausmultiplizieren mit den Regressionsparametern und Zusammenfassen der LSER-Deskriptoren ergibt sich die als Gleichung (4.17) beschriebene Endgleichung.

$$\log K_{ba} = 1,00 E + 1,75 S + 3,09 A + 2,46 B^H + 0,41 V - 0,65 \quad (4.17)$$

$$n = 190, r^2 = 0,74, rms = 0,57, me = 0,42, mne = -2,46, mpe = 1,29$$

Wird die Zielgröße mit dieser neu berechneten Gleichung (4.17) vorhergesagt, folgt das gleiche Ergebnis wie in Gleichung (4.15), aus dem das Bestimmtheitsmaß  $r^2 = 0,74$  und ein  $rms$  von 0,57 resultieren. Verglichen mit dem Literatur-Modell (LSER-Gleichung (4.5), Kapitel 4.2.1, Seite 61) ist die Statistik im Allgemeinen etwas schlechter. Auffällig ist, dass der Deskriptor für die Polarität und Polarisierbarkeit in die Literaturgleichung negativ eingeht, während sie ein positives Vorzeichen besitzt. Des Weiteren nutzt die Literatur-Gleichung den logarithmischen Hexadecan/Luft-Verteilungskoeffizienten  $L$ , wohingegen in der zusammengefassten LSER-Gleichung das charakteristische Volumen nach McGowan  $V$  verwendet wird. Diese Stoffparameter sind jedoch nach Abraham *et al.* [119] austauschbar. Alle weiteren Deskriptoren sind in ihren Phasenparametern vergleichbar.

Es können damit sämtliche Modelle, bei denen  $K_{ba}$  mit anderen Verteilungskoeffizienten beschrieben wird (Tabelle 4.5, Seite 75) und diese über eine LSER-Gleichung berechnet werden können, zusammengefasst werden.

### 4.2.2.2 Einsatz von physiologischen Fraktionsanteilen

Ausgehend von den Denkansätzen und Modellen (2.11) und (2.12) in Kapitel 2.2.1 (Seite 13) wird der Blut/Luft-Verteilungskoeffizient unter Einsatz von physiologischen Fraktions-

## Modellierung von Verteilungskoeffizienten

anteilen modelliert. Die Gleichung (2.11) entspricht annähernd der tatsächlichen Blutzusammensetzung, da die Fraktionsanteile von Wasser, Proteinen und (Membran-)Lipiden einbezogen werden. Für die Fraktionsanteile werden die Werte 0,5 für Wasser, 0,1 für Protein und 0,4 für Lipid verwendet. Für Protein wird wieder das Humanalbumin ( $K_{\text{HSA}}$ ) und für die Lipid/Wasser-Verteilung der  $K_{\text{mw}}$  verwendet, woraus Gleichung (4.18) entsteht. Nach Einsetzen der Verteilungskoeffizienten  $K_{\text{wa}}$  (HENRYWIN [117] aus Epi Suite [118]),  $K_{\text{HSA}}$  (LSER) und  $K_{\text{mw}}$  (LSER) und Ausrechnen der Prognosewerte ergeben sich die Statistikparameter wie sie unter der Gleichung stehen.

$$\log K_{\text{ba}} = \log K_{\text{wa}} + \log (0,5 + 0,1 \cdot K_{\text{HSA}} + 0,4 \cdot K_{\text{mw}}) \quad (4.18)$$

$$n = 190, r^2 = 0,84, rms = 0,45, me = 0,52, mne = -1,15, mpe = 2,61$$

Die Modellstatistik zeigt, dass das Bestimmtheitsmaß mit 0,84 deutlich verbessert ist, aber der *mpe*-Fehler mit über zweieinhalb logarithmischen Einheiten sehr groß ist. Dieses Ergebnis zeigt sich auch in der grafischen Darstellung (Abbildung 4.11), in der die xy-Gerade als schwarze Linie eingezeichnet ist.

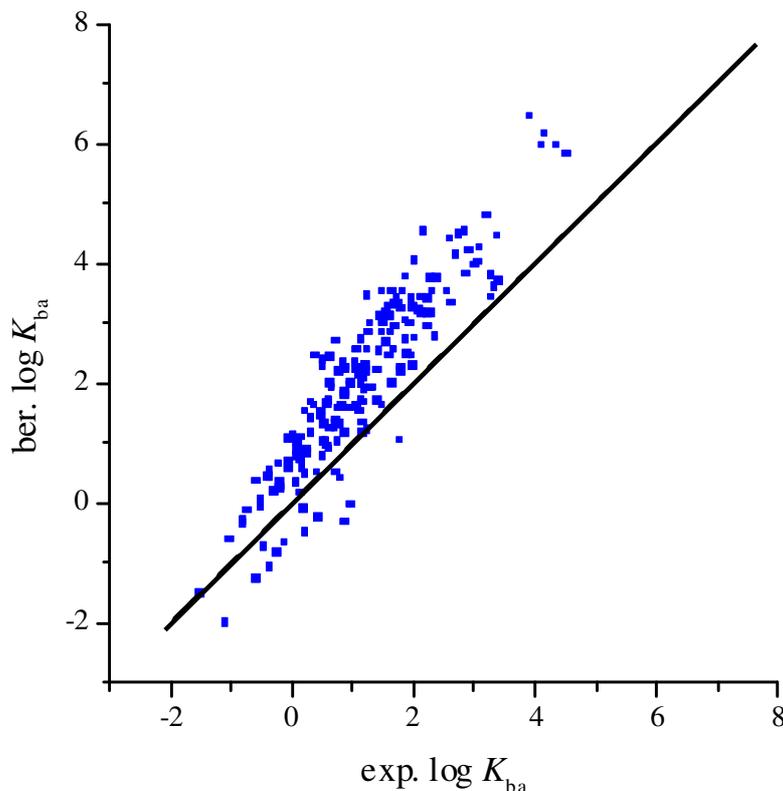


Abbildung 4.11: Auftragung von berechneten  $\log K_{\text{ba}}$ -Werte über Gleichung (4.18) gegen experimentelle Werte.

## Modellierung von Verteilungskoeffizienten

Mit steigenden  $\log K_{ba}$ -Werten weichen die berechneten Werte immer stärker von der xy-Gerade ab. Aus diesem Grund wird die Gleichung (4.18) dahingehend modifiziert, dass eine Regression durchgeführt und die Gleichung dadurch angepasst wird. Hierbei entsteht Gleichung (4.19) ( $K_{wa}$ : HENRYWIN [117] aus Epi Suite [118],  $K_{HSA}$ : LSER und  $K_{mw}$ : LSER) mit folgenden Regressions- und Statistikparametern.

$$\log K_{ba} = (0,67 \pm 0,02) \cdot \log K_{wa} + (0,51 \pm 0,04) \cdot \log (0,5 + 0,1 \cdot K_{HSA} + 0,4 \cdot K_{mw}) + (0,04 \pm 0,08) \quad (4.19)$$

$$n = 190, r^2 = 0,85, q_{CV}^2 = 0,85, rms = 0,43, rms_{CV} = 0,43, me = 0,35, mne = -0,97, mpe = 1,35$$

Bei Modell Gleichung (4.19) ist die Konstante nicht signifikant und muss somit nicht berücksichtigt werden. Grafisch ist das Modellergebnis in Abbildung 4.12 dargestellt.

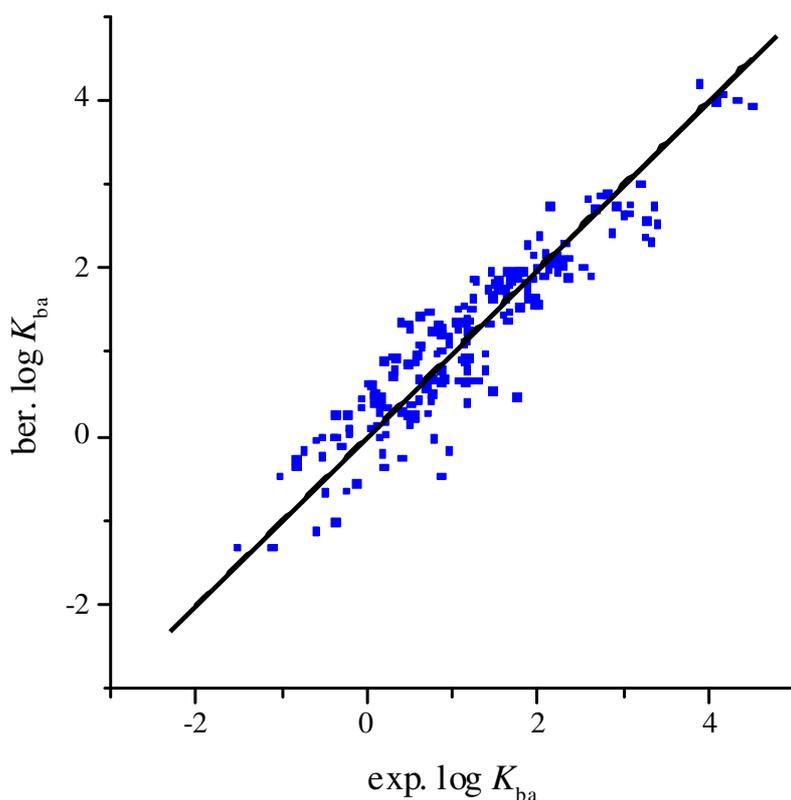


Abbildung 4.12: Auftragung von berechneten  $\log K_{ba}$ -Werte über Gleichung (4.19)(4.18) gegen experimentelle Werte.

## Modellierung von Verteilungskoeffizienten

Die Modelliererergebnisse dieser Gleichung werden mit denen der Näherung dieser Gleichung verglichen, bei der der mittlere Term  $b \cdot \log(0,5 + 0,1 \cdot K_{\text{HSA}} + 0,4 \cdot K_{\text{mw}})$  durch die Summe der logarithmierten Protein/Wasser- und Membran/Wasser-Verteilungskoeffizienten ersetzt wird (siehe Gleichung (2.12)). Dieses Modell entspricht genau dem Modell Nr. 7 (Gleichung 4.13) aus Kapitel 4.2.2 (Tabelle 4.5, Seite 75). Bei einer Korrelation der vorhergesagten Werte miteinander ergibt sich ein Korrelationskoeffizient von 0,97. Dies bedeutet, dass die Vorhersage der Zielgröße  $\log K_{\text{ba}}$  mit beiden Modellen sehr ähnlich ist. Ein Vergleich der Regressionen ist in Tabelle 4.6 gezeigt.

**Tabelle 4.6: Vergleich der Regressions- und ausgewählter Statistikparameter der Modelle (4.19) und (4.13) mit  $n = 190$ .**

Modell	$n$	$a$	$b$	$c$	$d$	$q_{\text{CV}}^2$	$rms_{\text{CV}}$
(4.19)	190	$0,67 \pm 0,02$	$0,51 \pm 0,04$	$0,02^* \pm 0,08$	-	0,85	0,43
(4.13)	190	$0,68 \pm 0,02$	$-0,66 \pm 0,29$	$0,90 \pm 0,19$	$-1,14 \pm 0,46$	0,85	0,44

\* nicht signifikant

Die Vorhersagegüte  $q_{\text{CV}}^2$ , der  $rms_{\text{CV}}$ -Fehler und der Regressionsparameter  $a$  sind bei beiden Modellen sehr ähnlich. Die Regressionsparameter für  $\log K_{\text{HSA}}$  und  $\log K_{\text{mw}}$  in Modell (4.13) (Seite 73) werden mit  $-0,66$  und  $0,90$  angegeben. Verglichen mit den Fraktionsanteilen in Modell (4.18) ( $0,1$  für  $\log K_{\text{HSA}}$  und  $0,4$  für  $\log K_{\text{mw}}$ ) multipliziert mit dem Regressionsparameter  $0,51$  (ergibt  $0,051$  für  $\log K_{\text{HSA}}$  und  $0,204$  für  $\log K_{\text{mw}}$ ) ist der Trend bezüglich der Größenordnungen prinzipiell der Gleiche.

Zusammenfassend kann gesagt werden, dass beide Modelle vergleichbare Ergebnisse liefern und es somit keinen großen Unterschied macht, ob der Verteilungskoeffizient  $\log K_{\text{ba}}$  mit Gleichung (4.19) oder mit der Näherung (4.13) gerechnet wird.

### 4.2.3 Fragmentmodell für $K_{\text{ba}}$

Die theoretischen Grundlagen des Fragmentmodells sind in Kapitel 3.1.1 (Seite 32ff) gezeigt. Für den Blut/Luft-Datensatz mit 190 Substanzen (Kapitel 3.2.1, Seite 35ff) werden ausgehend von einer Ursprungsliste aus unserem Arbeitskreis 16 Fragmente verwendet, die zur Beschreibung der Zielgröße geeignet sind. Die verwendeten Fragmente liefern einen wichtigen Beitrag zum Modell und sind in Tabelle 4.7 dargestellt. Zusätzlich listet Tabelle 4.7

## Modellierung von Verteilungskoeffizienten

auch die Parameter aus der linearen Regression und die Konstante. Zudem sind die zugehörigen Fehler und p-Werte (siehe Kapitel 3.5 Seite 53) angegeben. Den größten Beitrag liefert mit einem Wert von 2,91 Fragment Nr. 8, welches eine Hydroxygruppe (OH) ist. Dieses ist jedoch verhältnismäßig gering mit einer Häufigkeit von 16 vertreten. Den zweitgrößten Beitrag liefert die Carbonylgruppe (O=C, Fragment Nr. 10, Häufigkeit = 20 und Parameter = 2,30), gefolgt von der Nitrogruppe (NO<sub>2</sub>, Fragment Nr. 12, Parameter = 2,29), die nur zweimal im gesamten Datensatz vertreten ist und dadurch einen hohen Fehler von 0,24 besitzt. Am häufigsten treten Wasserstoffatome und Wasserstoffe gebunden an eine Carbonylgruppe (H + H-CO, Fragment Nr. 3, Häufigkeit = 1341) auf. Danach folgen das Kohlenstoffatom (C, Fragment Nr. 2) mit einer Häufigkeit von 674 und der aromatische Kohlenstoff (Car, Fragment Nr. 4), der eine Häufigkeit von 190 vorweist.

**Tabelle 4.7: Fragmente mit Regressionsparametern, Fehlern und p-Werte für das Fragmentmodell für  $\log K_{ba}$ .**

Nr.	Fragment	Häufigkeit	Parameter	Fehler	p-Wert
1	+C_tert	46	-0,51	0,08	0,000000
2	C	674	0,33	0,05	0,000000
3	H + H-CO	1341	-0,09	0,02	0,000020
4	Car	190	0,30	0,02	0,000000
5	Cl	99	0,57	0,03	0,000000
6	Br + I	20	0,91	0,07	0,000000
7	+XCar + XCd	31	-0,24	0,05	0,000009
8	OH	16	2,91	0,10	0,000000
9	O	32	1,27	0,08	0,000000
10	O=C	20	2,30	0,11	0,000000
11	+O-CO	9	-1,83	0,17	0,000000
12	NO <sub>2</sub>	2	2,29	0,24	0,000000
13	+FCar	27	-0,19	0,04	0,000001
14	+CX <sub>3</sub>	27	-0,41	0,10	0,000052
15	+CX <sub>4</sub>	4	-1,07	0,20	0,000000
16	+O-C-X	13	-0,37	0,06	0,000000
	Konstante	190	-0,36	0,08	0,000011

Im Vergleich zu Chlor (Cl, Fragment Nr. 5) sind die Halogene Brom und Iod zu einem Fragment zusammengefasst (Br + I, Fragment Nr. 6). Halogene werden im Allgemeinen mit X bezeichnet, was für die Fragmente Nr. 14 (+CX<sub>3</sub>) und Nr. 15 (+CX<sub>4</sub>) jeweils drei oder vier beliebige entweder gleiche oder unterschiedliche Halogene an ein Kohlenstoffatom gebunden, bedeutet. Ein Halogen (mit Ausnahme von Fluor) gebunden an ein aromatisches Kohlenstoffatom oder eines mit Doppelbindung, wird mit +XCar + XCd (Fragment Nr. 7)

## Modellierung von Verteilungskoeffizienten

bezeichnet. Fluor gebunden an einen aromatischen Kohlenstoff (+FCar, Fragment Nr. 13) wird hierbei jedoch gesondert betrachtet. Ein Halogen gebunden an ein Kohlenstoffatom, das wiederum an ein Sauerstoffatom gebunden ist, wird mit +O-C-X bezeichnet (Fragment Nr. 16). Ein sehr wenig vertretenes Fragment mit einer Häufigkeit von nur neun Fällen, ist ein Sauerstoffatom gebunden an eine Carbonylgruppe (+O-CO, Fragment Nr. 11). Weitere Gruppen bilden Sauerstoffatome (O, Fragment Nr. 9) und tertiäre Kohlenstoffatome (+C\_tert, Fragment Nr. 1). Die Konstante der linearen Regression ist in der letzten Zeile von Tabelle 4.7 aufgeführt. Den größten p-Wert besitzt Fragment Nr. 14, +CX<sub>3</sub>, welches jedoch immer noch mit einem Wert von 0,000052 verschwindend gering ist.

Die Statistik dieses angepassten Modells ist mit folgenden Parametern gezeigt.

$$n = 190, r^2 = 0,92, q_{CV}^2 = 0,90, rms = 0,32, rms_{CV} = 0,35, me = 0,25, mne = -1,26, mpe = 0,90$$

Die grafische Darstellung der vorhergesagten  $\log K_{ba}$ -Werte gegen die experimentellen Werte ist in Abbildung 4.14 dargestellt. Zentral gelegen ist eine Substanz zu finden, die mit diesem Fragmentmodell überschätzt wird. Hierbei handelt es sich um Divinylether (siehe Abbildung 4.13), mit einer Überschätzung von 1,26.

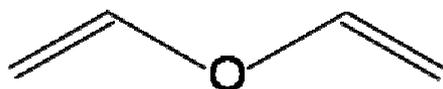
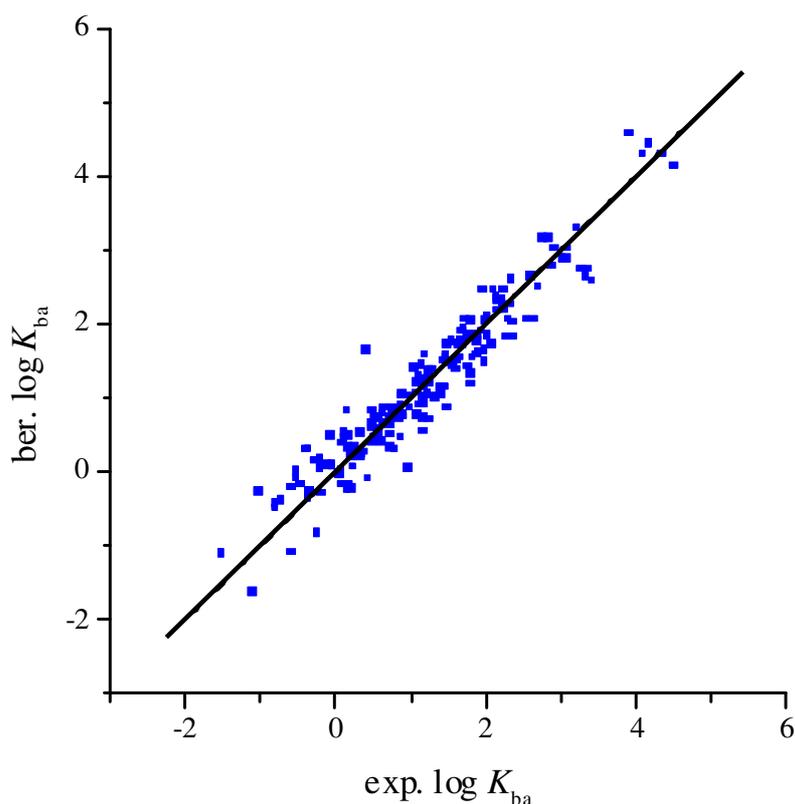


Abbildung 4.13: Strukturformel von Divinylether.

Auffällig ist auch, dass wiederum die Monoglykolether (siehe Abbildung 4.2, Seite 60) am rechten oberen Ende senkrecht zur xy-Geraden angeordnet sind.

## Modellierung von Verteilungskoeffizienten



**Abbildung 4.14:** Vorhergesagte  $\log K_{ba}$ -Werte mit dem Fragmentmodell aufgetragen gegen experimentelle  $\log K_{ba}$ -Werte.

Mit den betrachteten 16 Fragmenten wird ein  $K_{ba}$ -Modell für 190 Stoffe mit einem Bestimmtheitsmaß von 0,92 und einem  $rms$  von 0,32 erhalten. Damit ist dieses Modell als sehr zufriedenstellend bewertbar. Durch Hinzunahme weiterer Fragmente könnte das Modell leicht verbessert werden, doch dieses Ergebnis ist ein guter Kompromiss zwischen geringer Fragmentanzahl und zufriedenstellender Güte des Modells. Vergleichend dazu erreicht das LSER-Modell mit berechneten Eingangsparametern ein Bestimmtheitsmaß von 0,77 und einen  $rms$  von 0,54 (siehe Kapitel 4.2.1, Abbildung 4.4 und Tabelle 4.3, Seite 66) und die Modelle mit anderen Verteilungskoeffizienten (Kapitel 4.2.2, Tabelle 4.5, Seite 75) bestenfalls ein  $r^2$  von 0,85 sowie einen  $rms$  von 0,43.

Der Anwendungsbereich des hier dargestellten Fragmentmodells wird bestimmt durch die Eingangsdaten, welche in diesem Fall die experimentellen  $\log K_{ba}$ -Werte von 191 Stoffen (siehe Kapitel 3.2.1) sind. Diese erstrecken sich mit Werten von  $-1,52$  bis  $4,508$  über 8 logarithmische Einheiten.

## Modellierung von Verteilungskoeffizienten

Wird dieses Fragmentmodell mit den fünf LSER-Parametern aus Kapitel 2.3.2 (Seite 21) in einer linearen Regression erweitert, verbessert sich das  $r^2$  mit einem Wert von 0,92 nicht. Auch in Kombination mit E-States (siehe Kapitel 3.1.2, Seite 34) ergibt sich keine Verbesserung.

Zusammenfassend wird das Fragmentmodell als sehr positiv bewertet und kann definitiv als Vorhersagemodell für den Blut/Luft-Verteilungskoeffizienten verwendet und auch empfohlen werden.

### 4.2.4 $K_{ba}$ -Modellierung mit E-States

Die Modellierung des Blut/Luft-Verteilungskoeffizienten mit elektrotologischen Indices ist in Kapitel 3.1.2 (Seite 34) dargestellt. 79 wichtige Atomgruppen der E-States mit den dazugehörigen Gruppensymbolen sind in Tabelle 8.4 im Anhang 8.1.3 gezeigt. Zur Erläuterung der wichtigsten E-States kann gesagt werden, dass ein großes „S“ jeweils Summe bedeutet und ein kleines „s“ für eine Einfachbindung (englisch: *single bond*) steht. Dementsprechend deutet das kleine „d“ auf eine Doppelbindung hin und das kleine „a“ steht für aromatische Systeme. Am Ende eines jeden elektrotologischen Index steht das Atom oder die Gruppe zu der die jeweiligen Bindungen führen. Bei Index SssCH<sub>2</sub> führen zu einer CH<sub>2</sub>-Gruppe zwei Einfachbindungen, während bei Index SdsCH eine Einfach- und eine Doppelbindung zu einer Kohlenwasserstoffgruppe führen. Wichtig ist hierbei der Hybridisierungszustand der Kohlenstoffatome, wobei es sich hier um sp<sup>3</sup>- und sp<sup>2</sup>-Hybridisierung handelt. Dieser zeigt den elektronischen Charakter bezogen auf die topologische Umgebung eines jeden Skelettatoms, was das Charakteristikum für E-States ist. SH-Donor und SH-Akzeptor sind die Summen an Wasserstoffbrücken-Donoren und Wasserstoffbrücken-Akzeptoren (vgl. Abraham-Parameter *A* und *B* in Kapitel 2.3) im gesamten Datensatz. Die elektrotologischen Indices werden zuerst für jedes Skelettatom einzeln berechnet, wobei die Summe der Werte pro Atom ausgegeben wird. Diese sogenannte Summe des Moleküllevels wird für jeden Stoff angepasst.

Ein über schrittweise Regression erhaltenes Modell stellt sich als nicht besonders hilfreich heraus. Mit vier E-States (SsCH<sub>3</sub>, SdO, SsF und SaaCH) wird ein Bestimmtheitsmaß von 0,63 und ein *rms*-Fehler von 0,69 erreicht. Auch als Ergänzung zum Fragmentmodell und in Kopplung mit LSER sind keine befriedigenden Ergebnisse zu erreichen. Somit ist die

## Modellierung von Verteilungskoeffizienten

Modellierung mit E-States für den Verteilungskoeffizienten  $K_{ba}$  im Vergleich zu den anderen Methoden nicht konkurrenzfähig und wird nicht mehr weiter betrachtet.

### 4.2.5 Konsensmodellierungsstrategie

In den Kapiteln 4.2.1 (Seite 58ff), 4.2.2 (Seite 66ff) und 4.2.3 (Seite 81ff) werden neu entwickelte Modelle zur Vorhersage des Verteilungskoeffizienten  $K_{ba}$  vorgestellt. Dabei werden verschiedene Ansätze verfolgt, die auf unterschiedlichen QSAR-Eigenschaften der Chemikalien im Datensatz beruhen. Der LSER-Ansatz basiert auf intermolekularen Wechselwirkungen, das Fragmentmodell auf Strukturchemie und die Modelle mit anderen Verteilungskoeffizienten auf die Verteilung in den Kompartimenten, die für PBPK-Modelle relevant sind. Die beste Vorhersage ist mit dem Fragmentmodell möglich, gefolgt vom Einsatz anderer Verteilungskoeffizienten. Mit dem LSER-Ansatz ist die Vorhersagegüte die schwächste. Der Grund für die verschiedenen Typen von Modellen ist, dass eine Konsensmodellschlussfolgerung durchgeführt werden kann, bei der verschiedene Modelltypen zur Vorhersage von  $K_{ba}$  kombiniert werden. Stimmen die Vorhersagen der verschiedenen Modelle überein und sind die Methoden ausreichend unterschiedlich, so steigt insgesamt die Vertrauenswürdigkeit der vorhergesagten Werte. Dazu werden vier Modelle empfohlen und ihre Fehler in Korrelationen miteinander verglichen. Zu den empfohlenen Modellen gehören das LSER-Modell (4.6) (Seite 63) mit berechneten Stoffparametern, das Fragmentmodell aus Kapitel 4.2.3 (Seite 81ff) und die Modelle Nr. 3 ( $K_{wa}+K_{HSA}$ ) und 4 ( $K_{wa}+K_{mw}$ ) aus Tabelle 4.5 (Seite 75) (siehe Gleichungen ((4.9) (Seite 69) und (4.10) (Seite 71)), die in Kapitel 2.2.1, Seite 9ff) dargestellt sind. Der Modellierungstyp der Modelle Nr. 3 und 4 ist gleich, da beide mit anderen Verteilungskoeffizienten gerechnet werden. Die Fehler dieser vier Modelle werden miteinander korreliert und die Ergebnisse in Tabelle 4.8 gezeigt.

**Tabelle 4.8: Korrelationen der Fehler von vier empfohlenen Modellen, wobei drei verschiedene Modelltypen vertreten sind.**

Modell	LSER (4.6)	$K_{wa}+K_{HSA}$ (4.9)	$K_{wa}+K_{mw}$ (4.10)	Fragmentmodell
LSER (4.6)	1,00			
$K_{wa}+K_{HSA}$ (4.9)	0,25	1,00		
$K_{wa}+K_{mw}$ (4.10)	0,26	0,97	1,00	
Fragmentmodell	0,39	0,34	0,37	1,00

## Modellierung von Verteilungskoeffizienten

Die Fehlerkorrelationen sollten gering sein, um die Modelle statistisch unabhängig zu betrachten sowie unabhängige Informationen zu liefern. Die Modelle (4.9) und (4.10) weisen wie erwartet eine sehr gute Fehlerkorrelation von 0,97 auf, da sie auch den gleichen Modelltyp verwenden. Diese werden in Kombination nicht als Vorhersagemodell empfohlen. Eine schlechte Fehlerkorrelation ist bei den Modellen (4.9) ( $K_{wa}+K_{HSA}$ ) und (4.6) (LSER) mit einem Korrelationskoeffizienten von 0,25 eindeutig der Fall. Die Vorhersage mit diesen Modellen liegt somit eng beisammen, woraus sich eine große Konfidenz für eine kombinierte Anwendung ergibt. Die Modelle sind ausreichend unabhängig voneinander. Ein vergleichbar hohes Konfidenzmaß (0,26) zeigen die Modelle (4.10) ( $K_{wa}+K_{mw}$ ) und LSER (4.6). Es wird somit empfohlen den LSER-Ansatz und eines der Modelle mit anderen Verteilungskoeffizienten zu nutzen, um den Blut/Luft-Verteilungskoeffizienten mit der Konsensmodellierungsstrategie vorherzusagen.

### 4.3 $K_{fa}$ -Modelle

Auch für die Modellierung des Verteilungskoeffizienten für Fett/Luft werden verschiedene Ansätze verfolgt. Es wird sowohl der LSER-Ansatz nach Abraham und Ibrahim [25] angewendet als auch andere Verteilungskoeffizienten, die dafür geeignet sind, die Fett/Luft-Verteilung zu beschreiben. Die Modelle werden untereinander verglichen sowie die Qualität der Modelle diskutiert.

#### 4.3.1 Modellierung von $\log K_{fa}$ mit Linearen Solvatations-Energie-Beziehungen

Für die Modellierung des Fett/Luft-Verteilungskoeffizienten wird zunächst die LSER-Gleichung nach Abraham (siehe Gleichung (2.23) in Kapitel 2.3, Seite 23) nachgerechnet und neu angepasst. Dafür können jeweils experimentelle und berechnete Deskriptoren verwendet werden. Die statistischen Ergebnisse der Nachrechnungen sind in Tabelle 4.9 gezeigt.

## Modellierung von Verteilungskoeffizienten

**Tabelle 4.9:** Statistische Parameter der nachgerechneten LSER-Gleichung für  $\log K_{fa}$  (2.23) mit experimentellen und berechneten Abraham-Deskriptoren.

Statistischer Parameter	exp. Desk.	ber. Desk.
$n$	96	96
$r^2$	0,96	0,86
$q^2$	0,96	0,85
$rms$	0,17	0,32
$bias$	-0,04	-0,03
$me$	0,13	0,22
$mne$	-0,40	-0,59
$mpe$	0,46	1,19

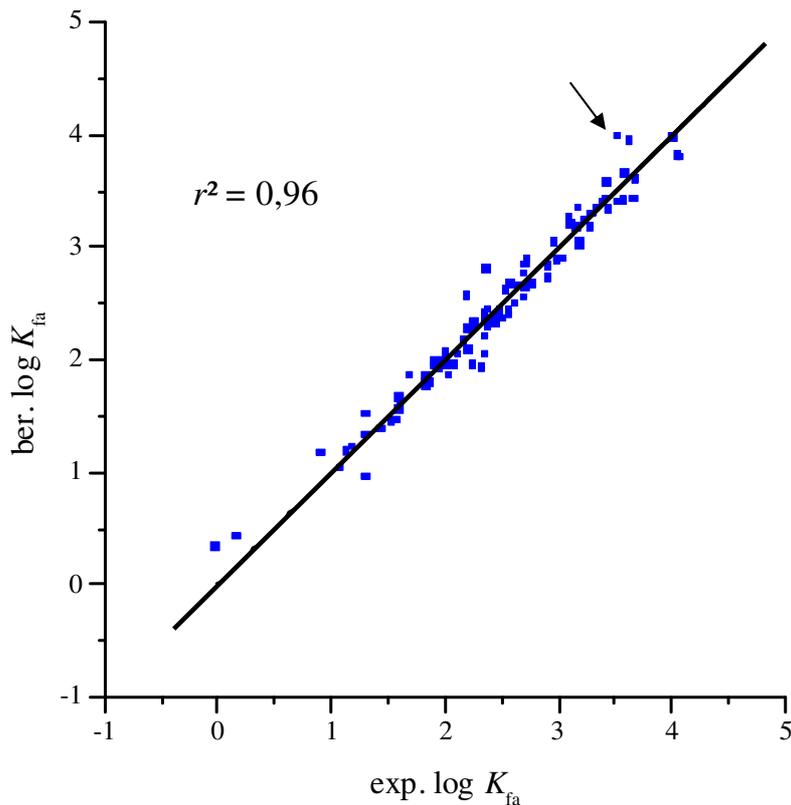
Das Bestimmtheitsmaß mit experimentellen Deskriptoren ( $r^2 = 0,96$ ) ist deutlich verbessert als mit berechneten ( $r^2 = 0,86$ ), wie es auch bei  $\log K_{ba}$  der Fall ist. Auch der  $rms$  ist mit experimentellen Stoffparametern kleiner und somit besser (siehe Tabelle 4.9). Ausgehend von der allgemeinen Gleichung (2.21) (Seite 22) mit experimentellen Deskriptoren ergeben sich bei der Neuanpassung ähnliche Werte für die Statistik und die Regressionsparameter (siehe Gleichung (4.20)) wie bei der nachgerechneten Literaturgleichung (Gleichung (2.23), Seite 23).

$$\log K_{fa} = (0,04 \pm 0,06) + (0,06 \pm 0,08) E + (0,51 \pm 0,09) S + (1,72 \pm 0,17) A + (0,39 \pm 0,12) B^H + (0,75 \pm 0,02) L \quad (4.20)$$

$$n = 96, r^2 = 0,96, q_{CV}^2 = 0,95, rms = 0,16, rms_{CV} = 0,17, me = 0,12, mne = -0,38, mpe = 0,47$$

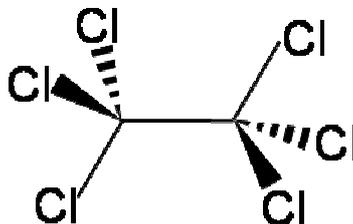
Die Werte für die Konstante und die Exzess-Molrefraktion sind dabei nicht signifikant (Erklärung siehe Kapitel 3.5, Seite 50ff). Die grafische Darstellung der berechneten gegen die experimentellen Werte ist in Abbildung 4.15 gezeigt. Die Gerade stellt die xy-Gerade dar.

## Modellierung von Verteilungskoeffizienten



**Abbildung 4.15:** Diagramm der neu angepassten LSER-Gleichung (2.21) für  $\log K_{fa}$  mit exp. Stoffparametern; der Pfeil zeigt den am stärksten überschätzten Stoff Hexachlorethan.

Die Vorhersage des Blut/Fett-Verteilungskoeffizienten mit dem LSER-Ansatz unter Einsatz von experimentellen Stoffparametern zeigt ein sehr gutes Modell mit kleinen Fehlern ( $rms_{CV} = 0,17$ ). Den größten Fehler besitzt Hexachlorethan (Abbildung 4.16) mit einem Wert von 0,47, welches in Abbildung 4.15 am oberen Ende der Grafik dargestellt ist (siehe Pfeil).



**Abbildung 4.16:** Strukturformel von Hexachlorethan.

Werden berechnete Stoffdeskriptoren für die Neuanpassung von Gleichung (2.21) (Seite 22) genutzt, ist das Modell insgesamt schlechter. Die Vorhersagegüte beträgt nur noch 0,85 und

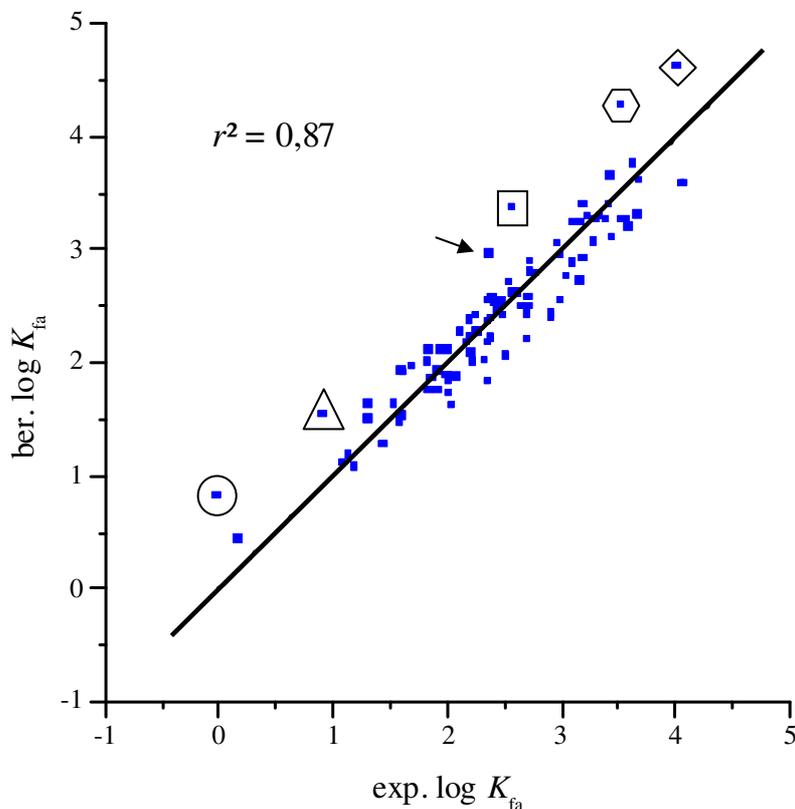
## Modellierung von Verteilungskoeffizienten

die Fehler sind deutlich größer ( $rms_{CV} = 0,31$ ). Die zugehörigen Regressionsparameter sind in Gleichung (4.21) gezeigt, wovon jedoch  $e$ ,  $s$  und  $b$  nicht signifikant sind.

$$\log K_{fa} = (0,30 \pm 0,10) + (0,22 \pm 0,19) E + (0,14 \pm 0,21) S + (1,88 \pm 0,36) A + (0,48 \pm 0,31) B^H + (0,69 \pm 0,04) L \quad (4.21)$$

$$n = 96, r^2 = 0,87, q_{CV}^2 = 0,85, rms = 0,29, rms_{CV} = 0,31, me = 0,22, mne = -0,59, mpe = 1,19$$

In Abbildung 4.17 sind die mit Gleichung (4.21) berechneten gegen experimentelle Werte von  $\log K_{fa}$  aufgetragen. Die xy-Gerade ist als schwarze Linie eingezeichnet.



**Abbildung 4.17:** Diagramm der neu angepassten LSER-Gleichung (2.21) für  $\log K_{fa}$  mit ber. Deskriptoren; die sechs gekennzeichneten Stoffe sind mit  $> 0,5$  logarithmischen Einheiten überschätzt.

Der größte Ausreißer in diesem LSER-Modell ist Ethen (Abbildung 4.18) mit einer Verfehlung des Zielwertes um 0,86. Ethen ist in Abbildung 4.17 eingekreist dargestellt.

## Modellierung von Verteilungskoeffizienten

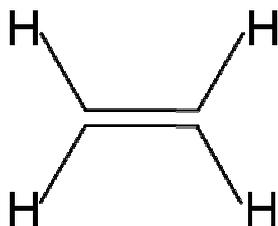


Abbildung 4.18: Strukturformel von Ethen.

Auch der vorhergesagte Wert für Tetrachlormethan (Abbildung 4.19) weicht um 0,82 vom experimentellen Wert ab und ist in Abbildung 4.17 von einem Viereck umgeben.

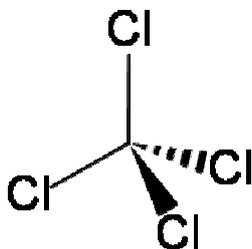


Abbildung 4.19: Strukturformel von Tetrachlormethan.

Vier weitere Substanzen in Abbildung 4.17 weisen Fehlerwerte von 0,61 bis 0,76 auf, wobei es sich um Oktan (Abbildung 4.20) (0,61) (mit Pfeil), Cyclopropan (Abbildung 4.21) (0,65) (mit Dreieck umgeben), Tricyclodekan (0,62) (mit Raute umgeben) und Hexachlorethan (Abbildung 4.16) (0,76) (mit einem Sechseck umgeben) handelt.



Abbildung 4.20: Strukturformel von Oktan.

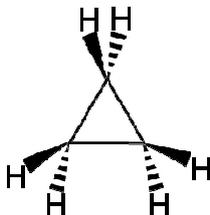


Abbildung 4.21: Strukturformel von Cyclopropan.

## Modellierung von Verteilungskoeffizienten

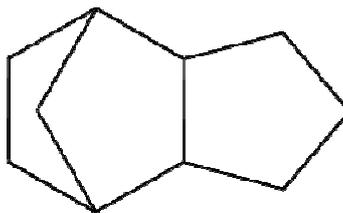


Abbildung 4.22: Strukturformel von Tricyclodekan.

Es fällt auf, dass alle diese sechs Ausreißer überschätzt werden. Dabei handelt es sich um vier Alkane und zwei chlorierte Alkane. Diese mittelgroßen Fehler resultieren in einer stärkeren Streuung um die  $xy$ -Gerade, als bei der Nutzung von experimentellen Eingangsgrößen. Ein Vergleich der statistischen Parameter ist in Tabelle 4.10 gezeigt.

**Tabelle 4.10:** Statistische Parameter der neu angepassten LSER-Gleichungen (4.20) und (4.21) für  $\log K_{fa}$  mit experimentellen und berechneten Abraham-Deskriptoren.

Statistischer Parameter	exp. Desk.	ber. Desk.
$n$	96	96
$r^2$	0,96	0,87
$q_{CV}^2$	0,95	0,85
$rms$	0,16	0,29
$rms_{CV}$	0,17	0,31
$me$	0,12	0,22
$mne$	-0,38	-0,51
$mpe$	0,47	0,86

Wie auch schon beim Blut/Luft-Verteilungskoeffizienten ist auch beim Fett/Luft-Koeffizienten das Modell mit experimentellen Deskriptoren besser. Dies liegt möglicherweise daran, dass die Berechnung der Deskriptoren nach Platts *et al.* [143] nicht exakt ist. Da ein Modell aber nur berechnete Eingangsgrößen verwenden soll, wird als Vergleich nur noch das Modell mit berechneten Deskriptoren genutzt. Die Anwendungsdomäne des Modells mit berechneten Abraham-Eingangsgrößen für  $\log K_{fa}$  ist in Kapitel 3.2.2 (Seite 38ff) dargestellt. Die Werte für die berechneten Deskriptoren erstrecken sich für  $E$  von -0,48 bis 0,889, für  $S$  von 0,009 bis 1,45, für  $A$  von 0,003 bis 0,348, für  $B$  von -0,015 bis 0,442 und für  $L$  von 0,232 bis 5,408. Mit dem Ziel ein besseres Modell zu entwickeln, als in der Literatur bereits vorhanden, werden andere Verteilungskoeffizienten für die Modellierung herangezogen.

## Modellierung von Verteilungskoeffizienten

### 4.3.2 Modellierung von $\log K_{fa}$ mit anderen Verteilungskoeffizienten

Es wird vermutet, dass der Verteilungskoeffizient  $\log K_{fa}$  mit dem Oktanol/Luft-Koeffizienten  $\log K_{oa}$  beschrieben werden kann (Kapitel 2.2.2, Seite 15). Eine Substanz ist lipophil, wenn sie sich gut in Fetten und Ölen löst. Der  $K_{ow}$ -Wert ist ein Maß für diese Lipophilie. Daher bietet sich der verwandte Koeffizient  $K_{oa}$  an, die Fett/Luft-Verteilung zu beschreiben.  $K_{oa}$  ist für die Modellierung geeignet, da gute Ergebnisse der linearen Regression gezeigt werden können (Gleichung (4.22)). Der Koeffizient  $\log K_{oa}$  wird hierbei mit dem Modul KOAWIN [122] aus EPI Suite [118] genommen (siehe Kapitel 3.3.4, Seite 47 und Kapitel 3.4, Seite 50). Die Auftragung ist in Abbildung 4.23 gezeigt, welche ein Bestimmtheitsmaß von 0,79 aufweist. Der dazugehörige  $rms$  beträgt 0,37.

$$\log K_{fa} = (0,65 \pm 0,03) \cdot \log K_{oa} + (0,66 \pm 0,10) \quad (4.22)$$

$$n = 96, r^2 = 0,79, q_{CV}^2 = 0,79, rms = 0,36, rms_{CV} = 0,37, me = 0,28, mne = -1,52, mpe = 0,79$$

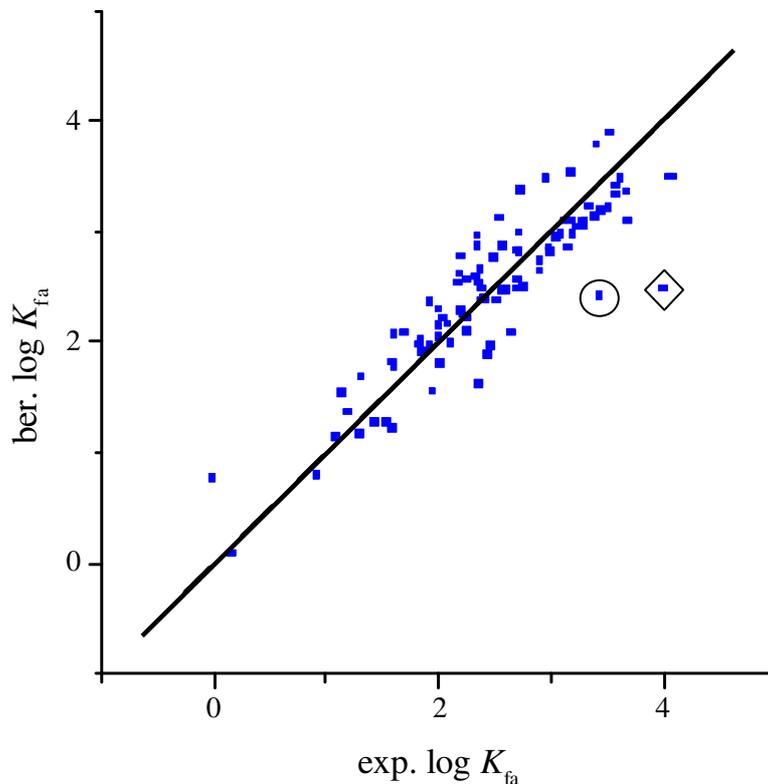


Abbildung 4.23: Grafische Darstellung der berechneten gegen experimentelle Werte des Fett/Luft-Verteilungskoeffizienten.

## Modellierung von Verteilungskoeffizienten

Die Substanzen Tricyclodekan (siehe Abbildung 4.22, Seite 92) (mit Raute umgeben) und Dekan (Abbildung 4.24) (eingekreist) sind am stärksten unterschätzt mit Fehlern von  $-1,52$  und  $-1,00$  und in Abbildung 4.23 ganz rechts zu sehen. Auch ist die Streuung um die  $xy$ -Gerade relativ stark.



Abbildung 4.24: Strukturformel von Dekan.

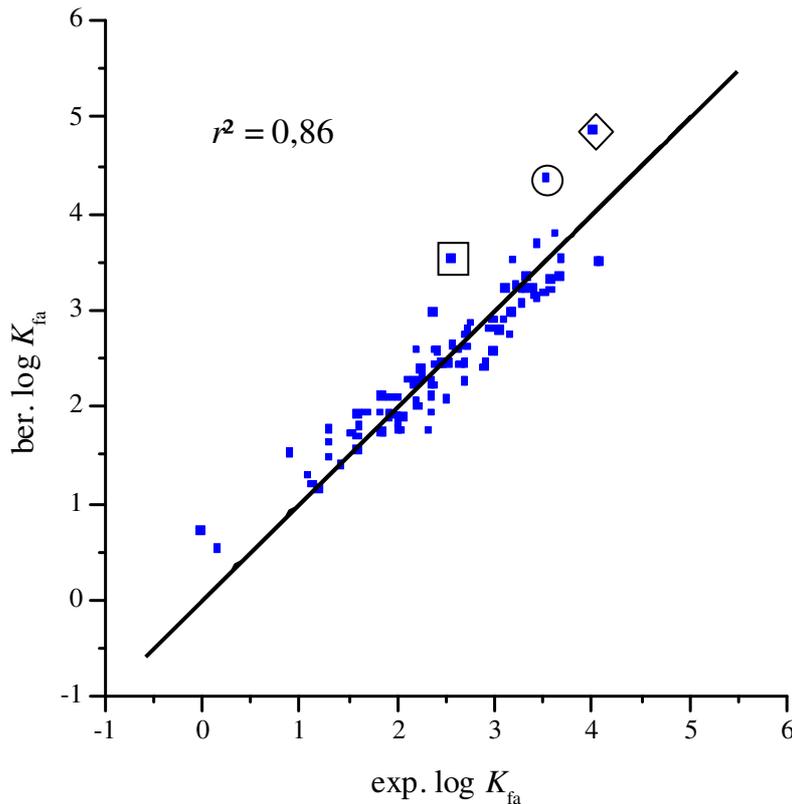
Wird mit  $K_{\text{olive}}$  (Kapitel 3.3.7, Seite 48f) ein weiterer Koeffizient zur Beschreibung der Fett/Luft-Verteilung untersucht und eine lineare Regression durchgeführt, ergeben sich die Statistikparameter und Gleichung (4.23) ( $\log K_{\text{olive}}$ : ber. LSER) wie unten gezeigt.

$$\log K_{\text{fa}} = (-0,81 \pm 0,03) \cdot \log K_{\text{olive}} + (0,37 \pm 0,09) \quad (4.23)$$

$$n = 96, r^2 = 0,86, q_{\text{CV}}^2 = 0,85, rms = 0,30, rms_{\text{CV}} = 0,32, me = 0,23, mne = -1,00, mpe = 0,58$$

Es ist deutlich zu erkennen, dass die Verwendung von  $K_{\text{olive}}$  ein noch besseres 1-Parameter-Modell als  $K_{\text{oa}}$  mit einem  $r^2$  von 0,86, einem  $q_{\text{CV}}^2$  von 0,85 und einem niedrigeren  $rms_{\text{CV}}$  von 0,32 liefert. Daraus ist ersichtlich, dass ein wichtiger Zusammenhang zwischen  $K_{\text{fa}}$  und  $K_{\text{olive}}$  vorliegt. Nach Anpassung und linearer Regression nach Gleichung (4.23) können die berechneten Werte gegen die experimentellen Werte wie es in Abbildung 4.25 gezeigt ist aufgetragen werden.

## Modellierung von Verteilungskoeffizienten



**Abbildung 4.25:** Grafische Darstellung des Modells (4.23) mit  $K_{olive}$ ; die drei gekennzeichneten Datenpunkte sind die am stärksten überschätzten Stoffe im Datensatz.

Hier lassen sich eindeutig drei überschätzte Ausreißer erkennen. Der größte davon befindet sich in Abbildung 4.25 relativ mittig bei den Werten  $\text{exp. log } K_{fa} = 2,55$  und  $\text{ber. log } K_{fa} = 3,55$  oberhalb der xy-Geraden. Dabei handelt es sich erneut um das vierfach halogenierte Tetrachlormethan (Abbildung 4.19, Seite 91) (mit einem Viereck umgeben) mit einem Fehler von einer logarithmischen Einheit. In Abbildung 4.25 liegt am oberen Ende der Grafik mit einem Fehler von 0,87 Tricyclodekan (Abbildung 4.22, Seite 92) (mit einer Raute gekennzeichnet) und links darunter Hexachlorethan (siehe Abbildung 4.16, Seite 89) (eingekreist), das um 0,85 logarithmischen Einheiten überschätzt ist.

Anscheinend lassen sich Verbindungen, deren Halogenanteil im Vergleich zum Kohlenstoffanteil sehr hoch ist, nicht gut mit diesem Modell vorhersagen. Hexachlorethan ist jedoch bereits bei dem LSER-Modell mit experimentellen Deskriptoren als Ausreißer aufgefallen. Somit ist die mäßige Vorhersage auf die LSER-Eingangsparameter gegründet. Dagegen können die cyclische Verbindung Tricyclodekan und das halogenierte Tetrachlormethan mit experimentellen LSER-Deskriptoren gut vorhergesagt werden, was möglicherweise auf ein Verteilungskoeffizientenproblem von  $K_{olive}$  hinweisen würde.

## Modellierung von Verteilungskoeffizienten

Der Parameter  $K_{\text{olive}}$  wird mit einer LSER-Gleichung nach Abraham berechnet, wie es in Kapitel 3.3.7 (Seite 48f) beschrieben ist. Bei der vergleichenden Betrachtung der Statistikparameter von Modellgleichung (4.23) unter Verwendung von experimentellen und berechneten Abraham-Deskriptoren für die  $K_{\text{olive}}$ -Berechnung fällt erneut auf, dass mit experimentellen Deskriptoren (exp.) deutlich bessere Ergebnisse erzielt werden (siehe Tabelle 4.11). Dieses lineare Modell ist jedoch nicht so gut, wie die reine LSER-Gleichung für  $\log K_{\text{fa}}$  von Abraham und Ibrahim [25] (siehe Tabelle 4.9, Seite 88 und Abbildung 4.15, Seite 89).

**Tabelle 4.11: Statistische Parameter der Gleichung (4.23) mit  $K_{\text{olive}}$  mit experimentellen und berechneten Abraham-Deskriptoren.**

Statistischer Parameter	ber.	exp.
$n$	96	96
$r^2$	0,86	0,93
$q_{\text{CV}}^2$	0,85	0,92
$rms$	0,30	0,21
$rms_{\text{CV}}$	0,32	0,23
$me$	0,23	0,15
$mne$	-1,00	-0,63
$mpe$	0,58	1,05

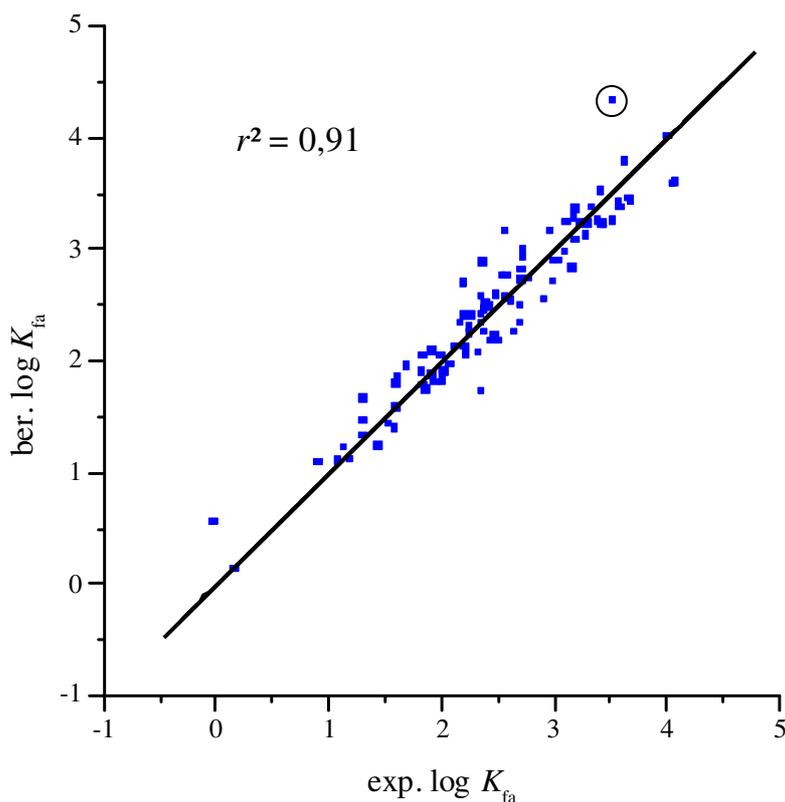
Da die Modelle mit berechneten Eingangsgrößen erstellt und nach Möglichkeit verbessert werden sollen, wird das 1-Parameter-Modell mit  $K_{\text{olive}}$  (4.23) zu einem 2-Parameter-Modell erweitert, indem der vorher betrachtete Verteilungskoeffizient  $K_{\text{oa}}$  mit  $K_{\text{olive}}$  gekoppelt wird. Die Gleichung für die  $\log K_{\text{fa}}$ -Vorhersage mit einer multilinearen Regression dieser beiden Verteilungskoeffizienten ist als Gleichung (4.24) ( $\log K_{\text{oa}}$ : KOAWIN [122] aus EPI Suite [118],  $\log K_{\text{olive}}$ : ber. LSER) dargestellt.

$$\log K_{\text{fa}} = (0,30 \pm 0,04) \cdot \log K_{\text{oa}} - (0,52 \pm 0,05) \cdot \log K_{\text{olive}} + (0,30 \pm 0,07) \quad (4.24)$$

$$n = 96, r^2 = 0,91, q_{\text{CV}}^2 = 0,91, rms = 0,24, rms_{\text{CV}} = 0,25, me = 0,18, mne = -0,82, mpe = 0,61$$

Das Bestimmtheitsmaß und die Vorhersagegüte betragen 0,91. Die Fehler sind sehr gering ( $rms = 0,24$  und  $me = 0,18$ ) und das Modell ist allgemein als sehr hochwertig zu betrachten, was die Darstellung in Abbildung 4.26 zeigt. Dieses Ergebnis ist mit einem  $rms_{\text{CV}}$  von 0,25 bei 96 Stoffen sehr zufriedenstellend.

## Modellierung von Verteilungskoeffizienten



**Abbildung 4.26:** Auftragung der Regression von berechnetem  $\log K_{fa}$  gegen experimentelles  $\log K_{fa}$  über Gleichung (4.24).

In Modell (4.24) und Abbildung 4.26 ist mit einem Fehler von 0,82 der größte Ausreißer Hexachlorethan (Abbildung 4.16, Seite 89) (eingekreist). Der Datenpunkt dieser Verbindung befindet sich am rechten oberen Ende von der Grafik (Abbildung 4.26). Hier ist  $\log K_{fa}$  mit zwei Verteilungskoeffizienten beschrieben, die einzeln über LSER-Parameter berechnet werden können.  $\log K_{olive}$  wird dabei über vier Abraham-Deskriptoren (siehe Gleichung (3.10) in Kapitel 3.3.7, Seite 48) und  $\log K_{oa}$  mit den üblichen fünf Deskriptoren (siehe Gleichung (3.7) in Kapitel 3.3.4, Seite 47) ermittelt. Der Oktanol/Luft-Verteilungskoeffizient kann auch aus  $K_{ow}$  und  $K_{aw}$  berechnet werden, welche wiederum mit LSER-Gleichungen berechnet werden können (siehe Kapitel 3.3.4, Seite 47). Diese  $K_{fa}$ -Modelle mit anderen Verteilungskoeffizienten können genau wie bei den  $K_{ba}$ -Modellen zu einer LSER-Gleichung zusammengefasst werden (siehe Kapitel 4.2.2.1, Seite 77f).

Als weitere mögliche Verbesserung des Fett/Luft-Modells kann noch ein weiterer, dritter Verteilungskoeffizient zum 2-Parameter-Modell hinzugenommen werden. Zu dem Modell mit  $K_{oa}$  und  $K_{olive}$  (Abbildung 4.26 und Gleichung (4.24)) wird der Membran/Wasser-Verteilungskoeffizient (Kapitel 3.3.5, Seite 47) hinzugenommen, woraus sich die multilineare Regression

## Modellierung von Verteilungskoeffizienten

(4.25) ( $\log K_{\text{oa}}$ : KOAWIN [122] aus EPI Suite [118],  $\log K_{\text{olive}}$ : ber. LSER,  $\log K_{\text{mw}}$ : ber. LSER) und unten stehende Statistikparameter ergeben.

$$\log K_{\text{fa}} = (0,37 \pm 0,04) \cdot \log K_{\text{oa}} - (0,37 \pm 0,05) \cdot \log K_{\text{olive}} + (0,16 \pm 0,04) \cdot \log K_{\text{mw}} + (0,16 \pm 0,08) \quad (4.25)$$

$n = 96$ ,  $r^2 = 0,93$ ,  $q_{\text{CV}}^2 = 0,91$ ,  $rms = 0,22$ ,  $rms_{\text{CV}} = 0,24$ ,  $me = 0,16$ ,  $mne = -0,54$ ,  $mpe = 0,81$

Grafisch ist dieses Modell in Abbildung 4.27 dargestellt. Die größten Ausreißer sind Hexachlorethan (Abbildung 4.16, Seite 89) (eingekreist), welcher mit einem Fehlerwert von 0,81 überschätzt ist und Oktan (Abbildung 4.20, Seite 91) (mit einem Viereck umgeben), der mit einem Fehlerwert von 0,70 überschätzt ist.

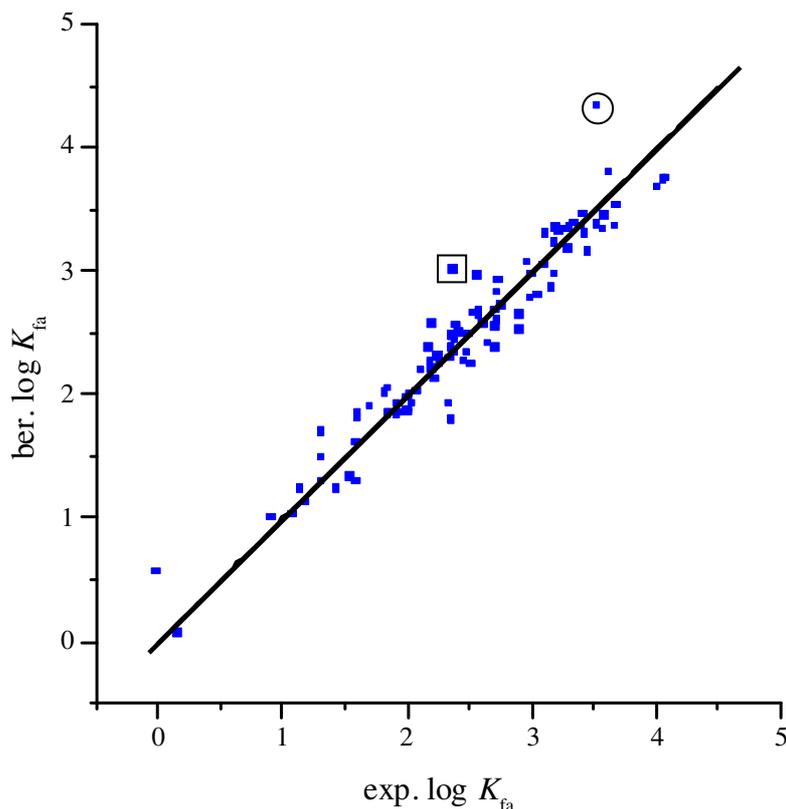


Abbildung 4.27: Auftragung der Regression von mit Gleichung (4.25) berechnetem  $\log K_{\text{fa}}$  gegen experimentelles  $\log K_{\text{fa}}$ .

Das 3-Parameter-Modell zeigt keine erhebliche Verbesserung, da sich die Vorhersagegüte von 0,91 nicht vergrößert. Die genutzten unabhängigen Parameter beschreiben ähnliche

## Modellierung von Verteilungskoeffizienten

Eigenschaften und somit bewirkt ein weiterer Verteilungskoeffizient keine Verbesserung für das Modell.

Die hier vorgestellten vier Kombinationen von Verteilungskoeffizienten als Deskriptoren für den Fett/Luft-Verteilungskoeffizienten, die in 1-Parameter-, 2-Parameter- und 3-Parameter-Modellen resultieren, sind in der Übersicht in Tabelle 4.12 gezeigt. Da die gesamten Koeffizienten über die Abraham-Stoffparameter berechnet werden können, ist es möglich die LSER-Gleichungen, wie es in Kapitel 4.2.2.1 (Seite 77f) gezeigt ist, zu einer Gleichung zusammenzufassen.

## Modellierung von Verteilungskoeffizienten

**Tabelle 4.12: Modelle, Regressions- und Statistikparameter für  $\log K_{fa}$ -Modelle nach dem Schema  $a \cdot \log A + b \cdot \log B + c \cdot \log C + d$ .**

Modell-Nr.	Verteilungskoeffizienten	$a$	$b$	$c$	$d$	$r^2$	$q_{CV}^2$	$rms$	$rms_{CV}$	$me$	$mne$	$mpe$
1	$A = K_{oa}$	$0,65 \pm 0,03$	-	-	$0,66 \pm 0,10$	0,79	0,79	0,36	0,37	0,28	-1,52	0,79
2	$A = K_{olive}$	$-0,81 \pm 0,03$	-	-	$0,37 \pm 0,09$	0,86	0,85	0,30	0,32	0,23	-1,00	0,58
3	$A = K_{oa}, B = K_{olive}$	$0,30 \pm 0,04$	$-0,52 \pm 0,05$	-	$0,30 \pm 0,07$	0,91	0,91	0,24	0,25	0,18	-0,82	0,61
4	$A = K_{olive}, B = K_{wa}, C = K_{HSA/a}$	$0,50 \pm 0,05$	$-0,44 \pm 0,08$	$0,68 \pm 0,10$	$1,23 \pm 0,14$	0,92	0,91	0,23	0,24	0,18	-0,76	0,44
5	$A = K_{oa}, B = K_{olive}, C = K_{mw}$	$0,37 \pm 0,04$	$-0,37 \pm 0,05$	$0,16 \pm 0,04$	$0,16 \pm 0,08$	0,93	0,91	0,22	0,24	0,16	-0,54	0,81

## Modellierung von Verteilungskoeffizienten

Das 1-Parameter-Modell, das  $K_{\text{olive}}$  zur Beschreibung der Zielgröße nutzt, ist das bessere Modell, welches nur einen Deskriptor nutzt und in Tabelle 4.12 als Modell Nr. 2 bezeichnet wird. Der Zusammenhang der Zielgrößen  $\log K_{\text{fa}}$  mit  $\log K_{\text{olive}}$  lässt sich auch in den LSER-Parametern der beiden einzelnen Gleichungen (2.23) (Kapitel 2.3.2, Seite 23) und (3.10) (Kapitel 3.3.7, Seite 48) erkennen. Somit besitzen die Phasenparameter für  $S$ ,  $A$  und  $L$  ähnliche Werte. Mit zwei Verteilungskoeffizienten ( $K_{\text{oa}}$  und  $K_{\text{olive}}$ ) ist die Fett/Luft-Verteilung besser beschrieben (Modell Nr. 3 in Tabelle 4.12, siehe Modell (4.24) und Abbildung 4.26). Der Anwendungsbereich dieses Modells ist durch die Wertebereiche der Deskriptoren  $\log K_{\text{oa}}$  und  $\log K_{\text{olive}}$  gegeben, deren logarithmische Werte sich für  $\log K_{\text{oa}}$  von  $-0,877$  bis  $4,939$  und für  $\log K_{\text{olive}}$  von  $0,207$  bis  $5,590$  erstrecken. Das 3-Parameter-Modell Nr. 5 erreicht ein  $r^2$  von  $0,93$  und einen  $rms$  von  $0,22$  und nutzt die Verteilungskoeffizienten  $K_{\text{olive}}$ ,  $K_{\text{oa}}$  und  $K_{\text{mw}}$ . Der Anwendungsbereich dieses Modells ist zusätzlich definiert durch den dritten Deskriptor  $\log K_{\text{mw}}$ , dessen Bereich sich von  $0,063$  bis  $4,408$  erstreckt.

Wird das Modell von Beliveau *et al.* [33] (Gleichung (2.17), Kapitel 2.3, Seite 17) auf den  $K_{\text{fa}}$ -Datensatz von 96 Stoffen (charakterisiert in Kapitel 3.2.2, Seite 38) angewandt, lässt es sich mit Modell Nr. 4 aus Tabelle 4.12 vergleichen, da diese beiden Modelle mit  $K_{\text{wa}}$ ,  $K_{\text{HSA/a}}$  und  $K_{\text{olive}}$  die gleichen Verteilungskoeffizienten nutzen. Das Literaturmodell ist in ChemProp [115,116] implementiert und auf 65 Stoffe anwendbar. Für die 31 sauerstoffhaltigen Substanzen wird kein Ergebnis geliefert. Neben der kleineren Anwendbarkeit des Modells von Beliveau *et al.* [33] liefert es zudem eine etwas schlechtere Statistik. Während das hier entwickelte Modell Nr. 24 ein Bestimmtheitsmaß von  $r^2 = 0,92$  und einen  $rms$  von  $0,23$  erreicht, sind die Parameter für das Vergleichsmodell  $r^2 = 0,90$  sowie  $rms = 0,28$ . Die Wertebereiche der einzelnen Deskriptoren sind für das Literaturmodell nur wenig begrenzter. Anstelle von  $5,4$  logarithmischen Einheiten für  $\log K_{\text{wa}}$  besitzt das Beliveau-Modell nur  $4,5$  logarithmische Einheiten und für  $\log K_{\text{HSA/a}}$  erstrecken sich die Werte statt über  $4,5$  logarithmische Einheiten gegenüber  $4$  logarithmischen Einheiten. Der Anwendungsbereich in Bezug auf den Olivenöl/Luft-Verteilungskoeffizienten ist bei dem hier entwickelten Modell und dem Literaturmodell gleich. Beide gehen über ca.  $6$  logarithmische Einheiten. Für die Vorhersage des Verteilungskoeffizienten Fett/Luft ist das Literaturmodell annähernd genauso gut wie das hier entwickelte Modell. Beide nutzen die gleichen drei Verteilungskoeffizienten als Deskriptoren, jedoch verwenden Beliveau *et al.* [33] Fraktionsanteile als Faktoren, während im hier entwickelten Ansatz die multilineare Regression als Anpassung verwendet wird.

## Modellierung von Verteilungskoeffizienten

Abschließend kann gesagt werden, dass die Fett/Luft-Verteilung mit anderen Verteilungskoeffizienten sehr zufriedenstellend zu beschreiben ist und mit zwei Deskriptoren (Modell Nr. 5,  $K_{oa}$  und  $K_{olive}$ ) besser ist, als das Literatur-Modell von Abraham und Ibrahim [25] (siehe Abbildung 4.17, Seite 90, Tabelle 4.10, Seite 92), welches berechnete Deskriptoren verwendet.

### 4.3.3 Konsensmodellierungsstrategie

Für die in den Kapiteln 4.3.1 (Seite 87) und 4.3.2 (Seite 93ff) entwickelten Modelle für den Verteilungskoeffizienten  $K_{fa}$  wird eine Konsensmodellschlussfolgerung durchgeführt. Die zwei Modelltypen, die hier betrachtet werden, beruhen auf dem LSER-Ansatz, der intermolekulare Wechselwirkungen als Basis nimmt und auf anderen Verteilungskoeffizienten, die die Verteilung in den PBPK-Kompartimenten beschreiben. Die bessere Vorhersage der Zielgröße  $K_{fa}$  wird mit der Verwendung von anderen Verteilungskoeffizienten erhalten. Werden diese beiden verschiedenen Modelltypen zur Vorhersage von  $K_{fa}$  kombiniert und die Methoden weisen einen ausreichenden Unterschied auf, so steigt die Vertrauenswürdigkeit der vorhergesagten Werte. Die Fehler von drei ausgewählten Modellen werden in Korrelationen miteinander verglichen. Zu diesen empfohlenen Modellen gehören das LSER-Modell (4.21) mit berechneten Stoffparametern (Seite 90), das Modell Nr. 3 ( $K_{oa}+K_{olive}$ ) aus Tabelle 4.12 (Seite 100) und das 3-Parameter-Modell Nr. 5 ( $K_{oa}+K_{olive}+K_{mw}$ ) ebenfalls aus Tabelle 4.12 (Seite 100) (siehe Gleichungen ((4.24) (Seite 96) und (4.25) (Seite 98)), Kapitel 4.3.2, Seite 93ff). Der Modellierungstyp der Modelle Nr. 3 und 5 ist der gleiche, da beide mit anderen Verteilungskoeffizienten gerechnet werden. Bei Modell Nr. 3 handelt es sich jedoch um ein 2-Parameter-Modell, während Modell Nr. 5 ein 3-Parameter-Modell ist. Die Fehler dieser drei Modelle werden miteinander korreliert und die Ergebnisse sind in Tabelle 4.13 gezeigt.

**Tabelle 4.13: Korrelationen der Fehler von drei empfohlenen Modellen, wobei zwei verschiedene Modelltypen vertreten sind.**

Modell	LSER (4.21)	$K_{oa}+K_{olive}$ (4.24)	$K_{oa}+K_{olive}+K_{mw}$ (4.25)
LSER (4.21)	1,00		
$K_{oa}+K_{olive}$ (4.24)	0,68	1,00	
$K_{oa}+K_{olive}+K_{mw}$ (4.25)	0,63	0,83	1,00

## Modellierung von Verteilungskoeffizienten

Die Korrelationen der Fehler sind verhältnismäßig hoch. Das bedeutet, dass die Modelle statistisch abhängig sind und somit keine unabhängigen Informationen liefern. Die Modelle, basierend auf anderen Verteilungskoeffizienten (4.24) ( $K_{oa}+K_{olive}$ ) und (4.25) ( $K_{oa}+K_{olive}+K_{mw}$ ), zeigen untereinander eine bessere Fehlerkorrelation (0,83) als jeweils einzeln mit dem LSER-Modell (4.21). Eine Kombination der Modelle (4.24) ( $K_{oa}+K_{olive}$ ) und (4.25) ( $K_{oa}+K_{olive}+K_{mw}$ ) wird daher nicht zur Vorhersage von  $K_{fa}$  empfohlen. Auch zu einer Kombination des LSER-Ansatzes (4.21) mit einem Modell des Typs multilineare Regression mit anderen Verteilungskoeffizienten wird nicht geraten, da eine Fehlerkorrelation von 0,68 bzw. 0,63 keine besonders große Konfidenz für die Anwendung aufweist.

### 4.4 Verteilungskoeffizient $K_{fb}$

Der dritte Verteilungskoeffizient  $\log K_{fb}$ , wie in den Kapiteln 2.2 (Seite 4ff) und 2.2.2 (Seite 15) erläutert und definiert, kann aus den beiden Verteilungskoeffizienten Blut/Luft und Fett/Luft errechnet werden. Dafür wird das Verhältnis der Koeffizienten, wie es in Gleichung (2.16) (Kapitel 2.2.2, Seite 15) gezeigt ist, gebildet oder die Logarithmen werden voneinander subtrahiert ( $\log K_{fa} - \log K_{ba}$ ). Es existieren experimentelle Werte aus der Literatur [38,75,77,146-148] für 12 Substanzen, die in Tabelle 4.14 mit ChemProp ID, CAS-Nummer, experimentellem  $\log K_{fb}$ -Wert und Quelle angegeben sind. Auffällig hierbei ist, dass es sich hauptsächlich um Kohlenwasserstoffe und nur eine Halogenverbindung (Tetrachlorethen, siehe Abbildung 4.28) handelt.

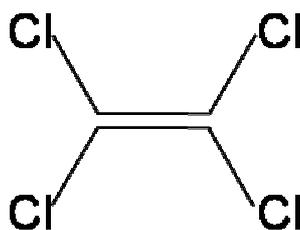


Abbildung 4.28: Strukturformel von Tetrachlorethen.

Diese Kohlenwasserstoffe sind entweder unverzweigt oder verzweigt oder sind, wie Styrol, aromatischer Struktur (siehe Abbildung 4.29).

## Modellierung von Verteilungskoeffizienten

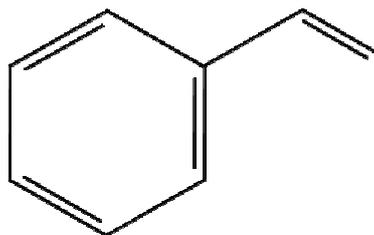


Abbildung 4.29: Strukturformel von Styrol.

Werden diese Stoffe mit den Datensätzen von  $K_{ba}$  und  $K_{fa}$  verglichen, fällt auf, dass nur drei Substanzen sowohl im  $K_{ba}$ -, als auch im  $K_{fa}$ -Datensatz vorhanden sind. Dennoch decken die beiden Koeffizienten zur Berechnung die  $K_{fb}$ -Daten bezüglich ihrer Komplexität und ihres Wertebereichs (1,699 – 3,092) ab. Somit können  $K_{ba}$  und  $K_{fa}$  zur Berechnung des dritten Verteilungskoeffizienten verwendet werden.

**Tabelle 4.14:** Substanzen, für die experimentelle Werte für  $\log K_{fb}$  in der Literatur vorhanden sind.

Nr.	Substanz	ChemProp ID	CAS-Nr.	exp $\log K_{fb}$	Quelle
1	Oktan	12	111-65-9	2,393	[75]
2	<i>n</i> -Nonan	971	111-84-2	2,380	[148], [146]
3	<i>n</i> -Dekane	1011	124-18-5	2,516	[38]
4	Undekane	1102	1120-21-4	2,723	[148]
5	Dodekane	1112	112-40-3	2,827	[148]
6	Styrol	58	100-42-5	1,699	[147]
7	Tetrachloroethen	159	127-18-4	3,092	[77]
8	3-Methyloktan	989	2216-33-3	2,274	[146]
9	4-Ethylheptan	987	2216-32-2	2,156	[146]
10	2,3-Dimethylheptan	978	3074-71-3	2,260	[146]
11	2,2,4-Trimethylhexan	995	16747-26-5	2,135	[146]
12	2,2,4,4-Tetramethylpentan	976	1070-87-7	2,217	[146]

Für diese zwölf in der Literatur mit experimentellen Werten gefundenen Substanzen kann eine Vorhersage begründet auf dem LSER-Prinzip (Gleichung (2.21), Kapitel 2.3.2, Seite 21) getroffen werden. Hierfür werden die mit berechneten Deskriptoren angepassten Gleichungen für  $\log K_{ba}$  aus Kapitel 4.2.1 (siehe Gleichung (4.5), Seite 61, hier Gleichung (4.26)) und für  $\log K_{fa}$  aus Kapitel 4.3.1 (siehe Gleichung (4.21), Seite 91, hier Gleichung (4.27)) hergenommen.

$$\log K_{ba} = (-0,33 \pm 0,12) + (1,30 \pm 0,20) E - (0,45 \pm 0,21) S + (4,77 \pm 0,47) A + (3,02 \pm 0,31) B^H + (0,23 \pm 0,05) L \quad (4.26)$$

$$n = 191, r^2 = 0,77, q_{CV}^2 = 0,68, rms = 0,54, rms_{CV} = 0,64, me = 0,41, mne = -1,38, mpe = 1,96$$

## Modellierung von Verteilungskoeffizienten

$$\begin{aligned} \log K_{fa} = & (0,30 \pm 0,10) + (0,22 \pm 0,19) E - (0,14 \pm 0,21) S + (1,88 \pm 0,36) A \\ & + (0,48 \pm 0,31) B^H + (0,69 \pm 0,04) L \end{aligned} \quad (4.27)$$

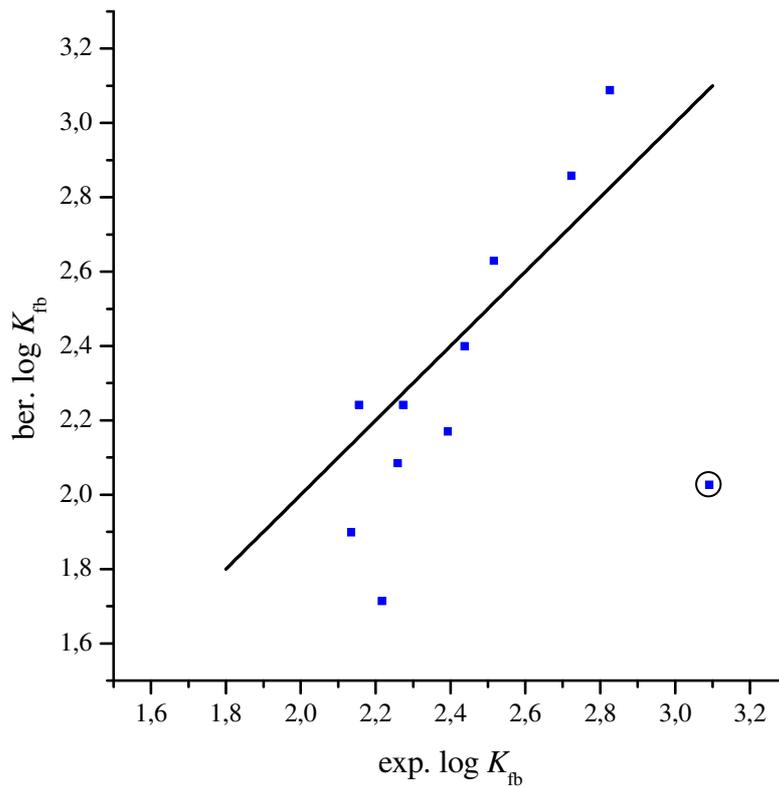
$$n = 96, r^2 = 0,87, q_{CV}^2 = 0,85, rms = 0,29, rms_{CV} = 0,31, me = 0,22, mne = -0,59, mpe = 1,19$$

Diese Vorhersage für die zwölf Stoffe des Fett/Blut-Verteilungskoeffizienten, basierend auf den LSER-Modellen von  $\log K_{ba}$  (Gleichung (4.26)) und  $\log K_{fa}$  (Gleichung (4.27)), ist durch die ermittelte Statistik nicht überzeugend.

$$n = 12, r^2 = 0,43, rms = 0,37, me = 0,25, mne = -1,07, mpe = 0,26$$

Die mit diesen Modellen errechneten Datenpunkte von  $\log K_{fb}$  sind in Abbildung 4.30 grafisch als Auftragung der berechneten gegen die experimentellen Werte dargestellt. Ein Ausreißer mit einer Unterschätzung von 1,07 logarithmischen Fehlereinheiten wurde ermittelt. Hierbei handelt es sich um die einzige Halogenverbindung Tetrachloroethen (siehe Abbildung 4.28, Seite 103) (eingekreist). Schon bei den Modellen für den Blut/Luft- und Fett/Luft-Verteilungskoeffizienten erscheinen mehrfach halogenierte Verbindungen als Ausreißer, was vermuten lässt, dass die hier entwickelten Modelle für diese Stoffklassen nicht geeignet sind und dafür andere Vorhersagemodelle benötigt werden. Ohne den gekennzeichneten Ausreißer erreicht das LSER-Modell für  $K_{fb}$  und elf Stoffen eine bessere Güte, was deutlich an der in Abbildung 4.30 eingezeichneten xy-Gerade erkennbar ist.

## Modellierung von Verteilungskoeffizienten



**Abbildung 4.30:** Auftragung der mit LSER berechneten  $\log K_{fb}$ -Werte gegen experimentelle Werte. Eingezeichnet ist die xy-Gerade und der Ausreißer Tetrachlorethen (eingekreist).

Ohne diesen Ausreißer ergibt die Statistik für elf Kohlenwasserstoffe folgende Parameter.

$$n = 11, r^2 = 0,87, rms = 0,22, me = 0,18, mne = -0,50, mpe = 0,26$$

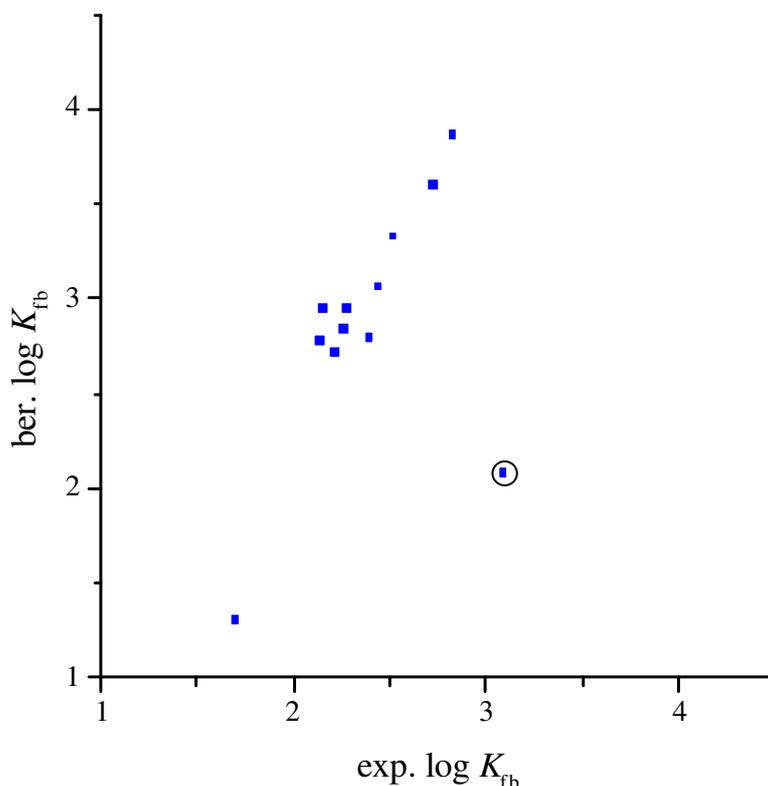
Das Bestimmtheitsmaß steigt deutlich an auf einen Wert von 0,87. Auch die Fehler sind erheblich kleiner und zeigen somit eine zufriedenstellende Vorhersage des Fett/Blut-Verteilungskoeffizienten. Eine Ausgleichsgerade durch die Datenpunkte besitzt ohne Tetrachlorethen eine Steigung  $> 1$ . Das bedeutet, dass Stoffe mit  $\log K_{fb}$ -Werten über 2,5 über- und Stoffe mit Werten unter 2,5 unterschätzt werden. Möglicherweise besteht hier ein Parametervorhersageproblem.

Die LSER-Modelle der Verteilungskoeffizienten  $K_{ba}$  und  $K_{fa}$  können für die Berechnung des dritten Koeffizienten verwendet werden, da aber nur eine geringe Anzahl an Daten verfügbar ist und zudem die Stoffe nur Kohlenwasserstoffe beinhalten, ist es kein sehr relevantes Modell.

## Modellierung von Verteilungskoeffizienten

Werden zur Berechnung des Fett/Blut-Verteilungskoeffizienten  $K_{fb}$  die Modelle mit anderen Verteilungskoeffizienten aus den Kapiteln 4.2.2 (Seite 66ff) und 4.3.2 (Seite 93ff) genutzt, kann für  $\log K_{ba}$  z. B. Modell Nr. 3 aus Tabelle 4.5 und für  $\log K_{fa}$  Modell Nr. 3 aus Tabelle 4.12 verwendet werden. In diesem  $K_{fb}$ -Modell sind insgesamt die vier Deskriptoren  $\log K_{wa}$ ,  $\log K_{oa}$ ,  $\log K_{olive}$  und  $\log K_{HSA}$  enthalten und für die zwölf Verbindungen wird eine mangelhafte Güte von einem Bestimmtheitsgrad von 0,23 und einem *rms*-Fehler von 0,56 erreicht. Die zugehörige Auftragung ist in Abbildung 4.31 zu sehen. Als Statistikparameter konnten folgende Werte ermittelt werden.

$$n = 12, r^2 = 0,23, rms = 0,56, me = 0,40, mne = -1,01, mpe = 0,88$$



**Abbildung 4.31:** Grafische Darstellung berechneten  $\log K_{fb}$ -Werte mit den Modellen der anderen Verteilungskoeffizienten Nr. 3 aus Tabelle 4.5 und Nr. 3 aus Tabelle 4.12 gegen experimentelle Werte.

Eine Ausgleichsgerade erzielt ohne den Ausreißerstoff Tetrachlorethen ein  $r^2$  von 0,90. Die gesamten Statistikparameter sind unten gezeigt.

## Modellierung von Verteilungskoeffizienten

$$n = 11, r^2 = 0,90, rms = 0,20, me = 0,16, mne = -0,17, mpe = 1,00$$

Die Steigung der Ausgleichsgerade ist für die elf Substanzen sehr steil mit einem Wert von 2, was keine perfekte Vorhersage der Werte für den Blut/Fett-Koeffizienten widerspiegelt, aber zumindest Näherungswerte bestimmbar macht.

Der Verteilungskoeffizient  $K_{fb}$  kann für diese zwölf Literaturstoffe über  $K_{ba}$  und  $K_{fa}$  mit dem LSER-Ansatz wesentlich besser als mit anderen Verteilungskoeffizienten errechnet werden.

Zur Erweiterung der Datenbasis für die Berechnung des dritten Koeffizienten kann zur Erlangung von mehr experimentellen Daten zunächst eine Schnittmenge der Stoffe aus den Datensätzen für  $K_{fa}$  und  $K_{ba}$  gebildet werden, wobei sich ergibt, dass alle 96 Substanzen aus dem  $K_{fa}$ -Datensatz auch im  $K_{ba}$ -Datensatz enthalten sind. Für diese Stoffe kann folglich ein quasi-experimenteller  $K_{fb}$ -Wert bestimmt werden (durch Subtraktion des logarithmierten Blut/Luft-Koeffizienten vom logarithmierten Fett/Luft-Koeffizienten) und dieser als Grundlage für die vorangegangenen Rechnungen genommen werden. Auch hier wird zunächst die Berechnung auf Grund der Abraham-Gleichung (Kapitel 2.3.2, Seite 21ff) durchgeführt. Für 96 Substanzen ergibt sich eine relativ breite Streuung um die xy-Gerade, was in Abbildung 4.32 gezeigt ist. Dennoch sind die Werte stark um die xy-Gerade verteilt und in Anbetracht der quasi-experimentellen Daten und der bereits fehlerbehafteten Ausgangsmodelle ist die Vorhersage zufriedenstellend.

## Modellierung von Verteilungskoeffizienten

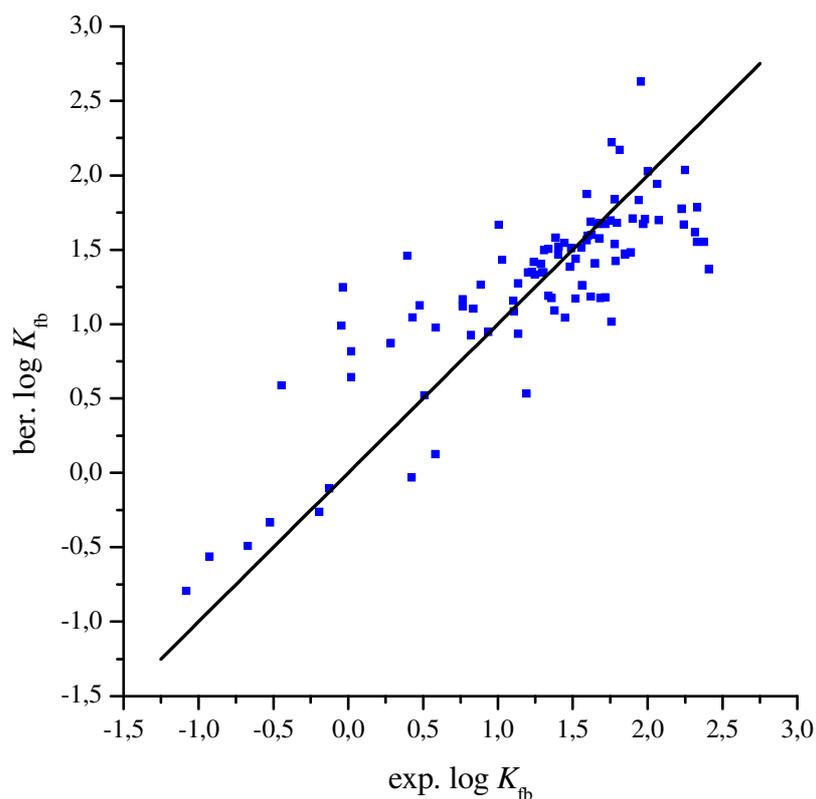


Abbildung 4.32: Auftragung von berechnetem  $\log K_{fb}$  über die LSER-Gleichungen von  $\log K_{fa}$  und  $\log K_{ba}$  gegen experimentelle Werte. Die Werte streuen um die xy-Gerade.

Die Stoffe mit Fehlerwerten über einer logarithmischen Einheit sind Difluormethan, 1-Nitropropan und 2-Nitropropan (siehe Abbildung 4.33).

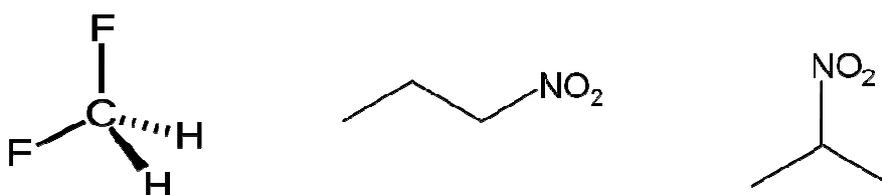


Abbildung 4.33: Strukturformeln von Difluormethan, 1-Nitropropan und 2-Nitropropan.

Die dazugehörige Statistik ist untenstehend gezeigt.

$$n = 96, r^2 = 0,69, rms = 0,33, me = 0,32, mne = -1,04, mpe = 1,28$$

Bei Hinzunahme der acht Stoffe aus der Literatur (siehe Tabelle 4.14), die im kombinierten Datensatz nicht vorhanden sind, wird nachfolgende Statistik erreicht.

## Modellierung von Verteilungskoeffizienten

$$n = 104, r^2 = 0,73, rms = 0,34, me = 0,31, mne = -1,04, mpe = 1,28$$

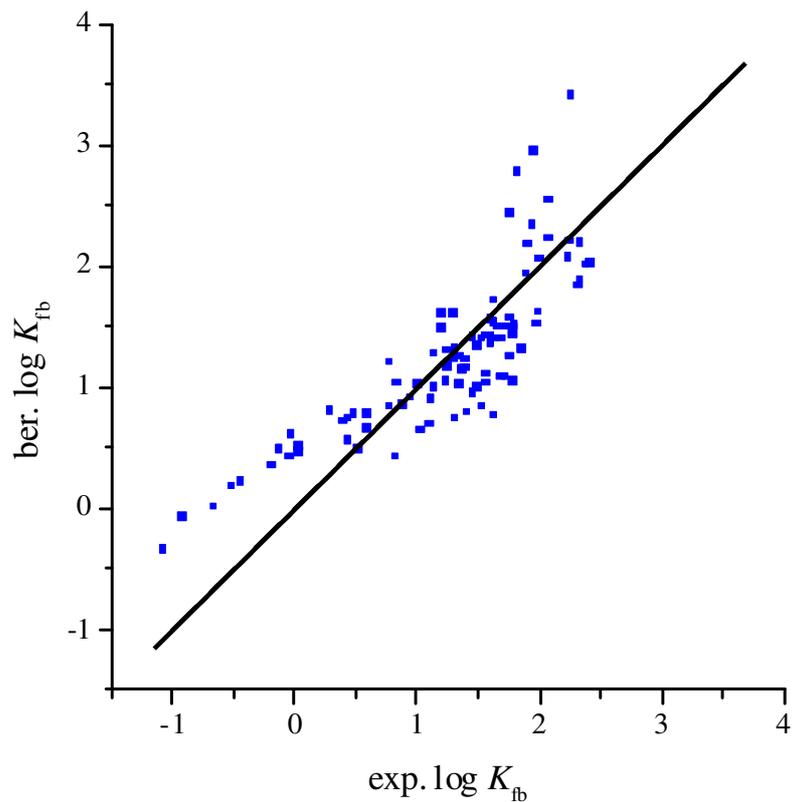
Das Bestimmtheitsmaß wird besser, aber die Fehler bleiben gleich bzw. verschlechtern sich im Falle von *rms*.

Wird der Fett/Blut-Verteilungskoeffizient über die Modelle mit anderen Verteilungskoeffizienten berechnet, kann z. B. wieder für  $\log K_{ba}$  z. B. Modell Nr. 3 aus Tabelle 4.5 und für  $\log K_{fa}$  Modell Nr. 3 aus Tabelle 4.12 verwendet werden. Die Deskriptoren des  $K_{fb}$ -Modells sind  $\log K_{wa}$ ,  $\log K_{oa}$ ,  $\log K_{olive}$  und  $\log K_{HSA}$ . Die Darstellung von berechneten gegen experimentelle  $\log K_{fb}$ -Werte ist in Abbildung 4.34 gezeigt. Die Statistik ist ähnlich der mit den LSER-Modellen berechneten.

$$n = 96, r^2 = 0,69, rms = 0,36, me = 0,34, mne = -0,84, mpe = 1,16$$

Die Streuung um die xy-Gerade ist ähnlich dem LSER-Modell (Abbildung 4.32). Auch hier können die quasi-experimentellen Werte als Grund dafür angesehen werden, dass die Güte nicht perfekt ist. Jedoch sieht dieses Modellergebnis viel besser aus, als das der zwölf experimentellen Werte aus der Literatur (Abbildung 4.31), was an der stark unterschiedlichen Datenmenge liegt. Es kommt jedoch deutlich zum Ausdruck, dass der dritte Koeffizient über die anderen beiden berechnet werden kann.

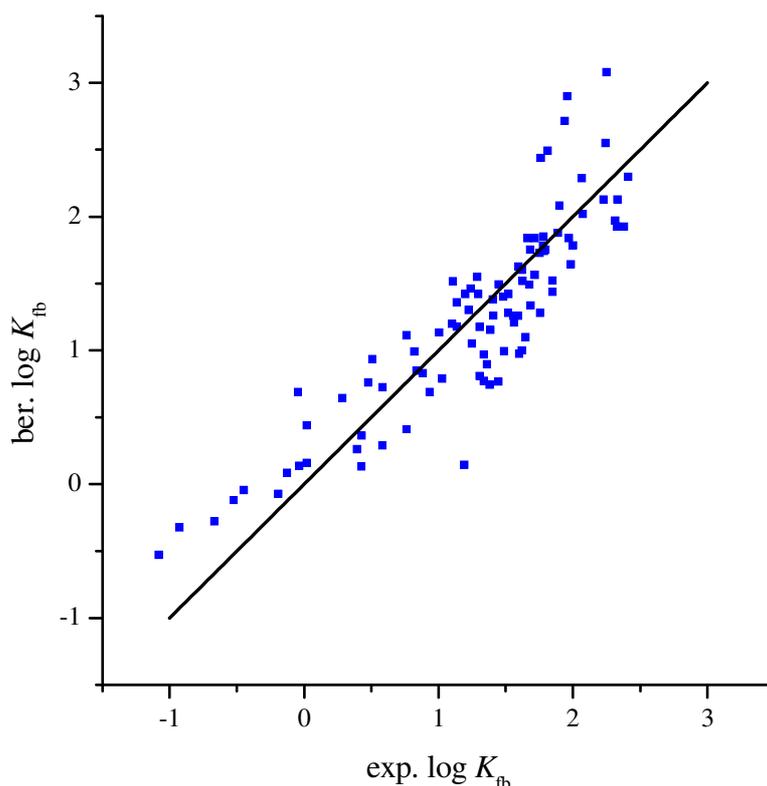
## Modellierung von Verteilungskoeffizienten



**Abbildung 4.34:** Grafische Darstellung der vorhergesagten  $\log K_{fb}$ -Werte über andere Verteilungskoeffizienten gegen experimentelle Werte. Die Werte streuen um die xy-Gerade.

Das Fragmentmodell für den Blut/Luft-Verteilungskoeffizienten (siehe Kapitel 4.2.3, Seite 81ff) kann auch zum Einsatz kommen und zusammen mit dem LSER-Modell für  $\log K_{fa}$  erreicht die Güte des  $K_{fb}$ -Modells ein  $r^2$  von 0,77, was als befriedigend angesehen wird. Die Auftragung von berechneten gegen experimentelle Werte ist in Abbildung 4.35 gezeigt.

## Modellierung von Verteilungskoeffizienten



**Abbildung 4.35:** Auftragung von berechnetem  $\log K_{fb}$  über LSER und Fragment gegen experimentelle Werte. Die Werte streuen um die xy-Gerade.

Abschließend kann gesagt werden, dass die Vorhersage des Fett/Blut-Verteilungskoeffizienten über den Blut/Luft- und Fett/Luft-Koeffizienten möglich ist, wenn die Modelle ausgehend von LSER, anderen Verteilungskoeffizienten oder Fragmenten verwendet werden. Mit der Datenbasis von zwölf experimentellen Werten aus der Literatur ist es nicht sinnvoll, eine Vorhersage zu treffen. Mit der erweiterten Datenbasis von 96 Stoffen lassen sich jedoch wichtige Erkenntnisse schließen. Der LSER-Ansatz ist prinzipiell ein gutes Modell, um Verteilungskoeffizienten zu beschreiben. Kombiniert mit dem Fragmentmodell für  $K_{ba}$  (Kapitel 4.2.3, Seite 81ff) verbessert sich das Modell deutlich. Ausgehend von anderen Verteilungskoeffizienten (die auch über LSER vorhergesagt werden können) ist ein zufriedenstellendes Modell zu erlangen. Deshalb wird empfohlen, den LSER-Ansatz bei der Vorhersage von Verteilungskoeffizienten in Betracht zu ziehen und speziell für den Fett/Luft-Koeffizienten eine Kombination von LSER und Fragment zu nutzen. Weiterführend kann für  $K_{fa}$  ebenfalls ein Fragmentmodell entwickelt werden. Demzufolge kann ein reines Fragmentmodell für die  $K_{fb}$ -Vorhersage verwendet werden und dieses mit dem hier gekoppelten Ansatz verglichen werden.

### 5 Dampfdruck-Modellierung

Da der Dampfdruck  $P_v$  in direktem Zusammenhang mit dem Verteilungskoeffizienten  $K_{wa}$  steht (siehe Kapitel 2.4.1, Seite 24ff), ist er ein bedeutender Parameter für die Modellierung von PBPK-relevanten Größen. Der Wasser/Luft-Verteilungskoeffizient findet eine wichtige Anwendung in der Beschreibung des Blut/Luft-Verteilungskoeffizienten in Kapitel 4.2.2 (Seite 66ff). Daher wird in diesem Kapitel die Modellierung des Dampfdrucks genauer betrachtet.

Einige Literaturmodelle für den Dampfdruck, beschrieben in Kapitel 2.4.2 (Seite 27ff), werden mit dem hier vorhandenen Datensatz (siehe Kapitel 3.2.3, Seite 41ff) nachgerechnet und die Ergebnisse untereinander verglichen. Ausgehend vom Zusammenhang des Dampfdruckes mit der Henry-Konstante und der Wasserlöslichkeit (Kapitel 2.4.1, Seite 24ff) sowie mit Hilfe des Mackay-Modells mit Modifizierungen werden  $P_v$ -Modellierungsansätze verfolgt. Die Modellierung des Dampfdrucks erfolgt demnach mit Hilfe der Parameter Henry-Konstante  $H$ , Schmelz- und Siedepunkt  $T_m$  und  $T_b$ , molekulare Masse  $MW$  und der Wasserlöslichkeit  $S_w$  (siehe Kapitel 3.3.8, Seite 49).

#### 5.1 Literatur-Modelle

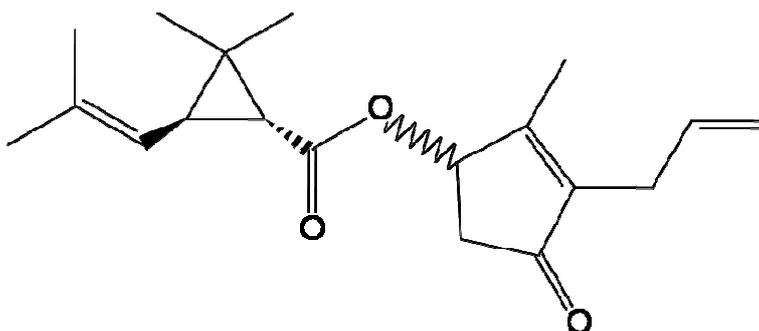
Eingangs wird der Datensatz, der in der Software EPI Suite [118] enthalten ist, mit  $n = 1482$  experimentellen Werten für den Dampfdruck herangezogen. Diese Stoffe werden gegen das interne Modell von EPI Suite [118] verglichen (MPBPWIN [90]). Dieses MPBPWIN-Modell sagt den Dampfdruck  $P_v$  mit drei einzelnen Methoden, der Antoine-Methode, der modifizierten Grain-Methode und der Mackay-Methode vorher (siehe Kapitel 2.4.2, Seite 27ff). Das bedeutet, dass der Dampfdruck für 1482 Stoffe mit dem internen Vorhersage-Programm berechnet und mit den programminternen experimentellen Daten verglichen wird. Dabei kann erwartet werden, dass die daraus resultierende Übereinstimmung besonders hoch ist. Das Ergebnis wird mit den bekannten Statistikparametern erfasst und ist in Tabelle 5.1 dargestellt. Eine Erklärung der Parameter ist in Kapitel 3.5 (Seite 50ff) zu finden.

## Dampfdruck-Modellierung

**Tabelle 5.1:** Statistische Parameter der Berechnung des Dampfdrucks mit dem internen Vorhersage-Programm und Vergleich mit dem internen experimentellen  $P_v$ -Datensatz von EPI Suite.

Statistischer Parameter	Wert
$n$	1482
$r^2$	0,92
$q^2$	0,92
$rms$	1,05
$bias$	0,21
$me$	0,60
$mne$	-9,97
$mpe$	5,83

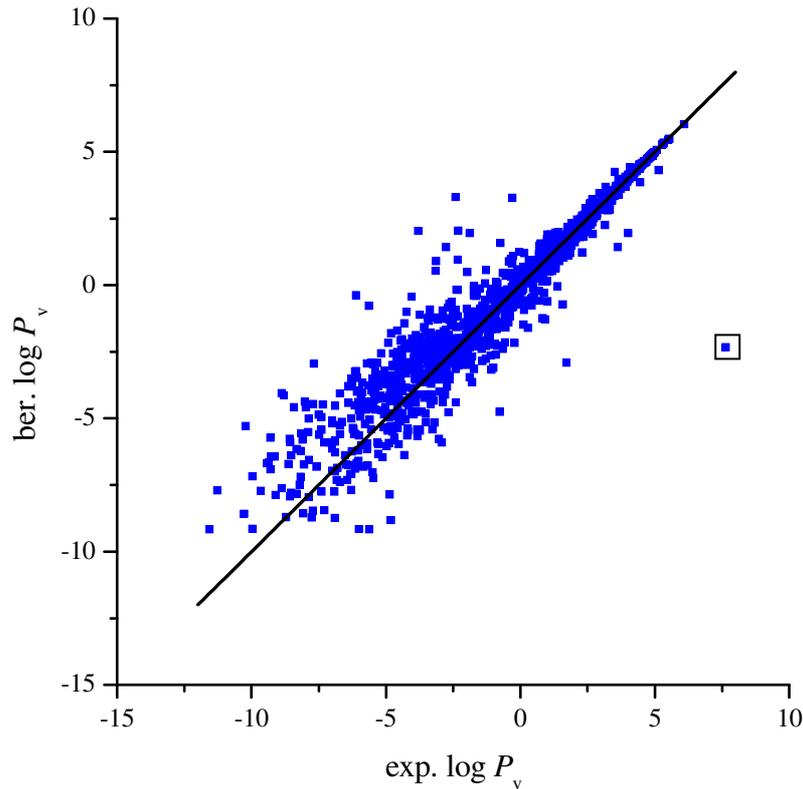
Das Modell ist mit einem Bestimmtheitsmaß von 0,92 global ein sehr gutes Modell. Der  $bias$  liegt bei 0,21 und der  $rms$  besitzt einen Wert von 1,05. Auffallend ist der größte negative Fehler  $mne$  von -9,97, der bei der Vorhersage von Bioallethrin, einem synthetischen Insektizid (siehe Abbildung 5.1), auftritt.



**Abbildung 5.1:** Strukturformel von Bioallethrin.

Bioallethrin besitzt einen experimentellen Dampfdruckwert von 7,64 logarithmischen Einheiten [Pa], wohingegen der vorhergesagte Wert bei -2,33 liegt. In Abbildung 5.2 sind die vorhergesagten Werte gegen die experimentellen Werte aufgetragen und die xy-Gerade durch die Wertepunkte eingezeichnet. Deutlich zu erkennen ist eine relativ breite Streuung um die xy-Gerade, vor allem im Bereich der kleineren  $\log P_v$ -Werte. Bioallethrin ist in Abbildung 5.2 mit einem Kästchen gekennzeichnet.

## Dampfdruck-Modellierung



**Abbildung 5.2:** Auftragung von berechneten gegen experimentelle log  $P_v$ -Werte aus dem Programm EPI Suite [118].

Von diesen 1482 Stoffen besitzen 99 Stoffe einen Fehler, der größer  $|2|$  ist. Das entspricht einem Anteil von 6,7 %.

Wird die Gleichung (2.29) in Kapitel 2.4.1 (Seite 26) nach dem Dampfdruck  $P_v$  aufgelöst, lässt sich dieser über Multiplikation der Henry-Konstante  $H$  mit der Wasserlöslichkeit  $S_w$  berechnen. Die logarithmischen Werte werden dementsprechend addiert (siehe Gleichung 5.1).

$$P_v = H \cdot S_w \rightarrow \log P_v = \log H + \log S_w \quad (5.1)$$

Dieses Modell wird zur Vereinfachung „H+S-Modell“ genannt. Bei dieser sehr einfachen Vorgehensweise wird die Henry-Konstante aus dem Programm EPI Suite [118] mit dem Modell HENRYWIN [117] gewonnen und die Wasserlöslichkeit über eine im Programm ChemProp [115,116] implementierte Modellauswahl [135] berechnet. Hierbei kann das Modell auf nur 1466 Stoffe angewendet werden, da nicht für alle die Wasserlöslichkeit berechnet werden kann. Die erhaltene Statistik ist unten angeführt.

$$n = 1466, r^2 = 0,88, q^2 = 0,88, rms = 1,44, me = 0,95, mne = -11,41, mpe = 9,01$$

## Dampfdruck-Modellierung

Mit diesem Modell wird ein schlechteres Ergebnis mit einem  $r^2$  von 0,88 erzielt und die Streuung um die xy-Gerade ist vor allem bei den höheren Werten deutlich stärker, als bei der programminternen Vorhersage (Abbildung 5.4). Es tritt der gleiche Ausreißer auf, Bioallethrin (Abbildung 5.1, Seite 114), der wieder stark unterschätzt ist (im Viereck). Diese Substanz besitzt einen hohen experimentellen Wert (7,64) und einen stark negativen Wert für die Wasserlöslichkeit (-5,56). Der vorhergesagte Wert ist somit auch negativ, wodurch eine starke Unterschätzung von über 10 logarithmischen Einheiten resultiert. Ein weiterer deutlicher Ausreißer ist 1,4,5,8-Tetraaminoanthrachinon, auch bekannt als C.I. Disperse Blue 1 (DB1) (siehe Abbildung 5.3), der auch stark unterschätzt ist.

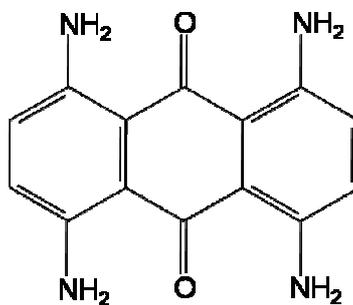
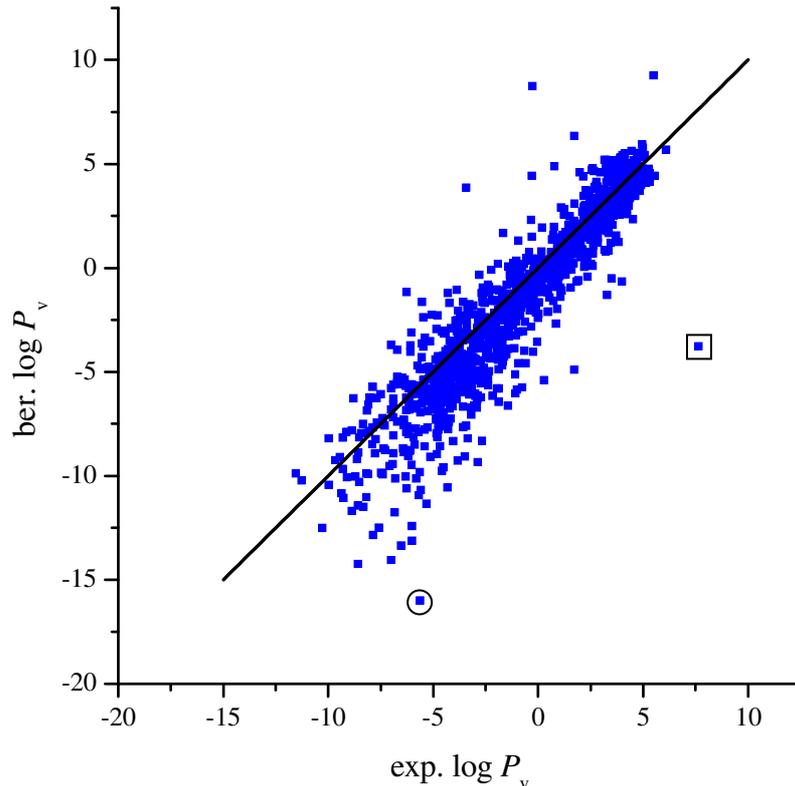


Abbildung 5.3: Strukturformel von 1,4,5,8-Tetraaminoanthrachinon (DB1).

Dieser Datenpunkt ist eingekreist. Der experimentelle Wert für DB1 liegt bei -5,62 und durch den sehr hohen negativen Wert für die Henry-Konstante (-14,30) resultiert ebenfalls eine Unterschätzung von über 10 logarithmischen Einheiten.

## Dampfdruck-Modellierung



**Abbildung 5.4:** Auftragung von mit Henry-Konstante und Wasserlöslichkeit berechneten gegen experimentelle aus EPI Suite  $\log P_v$ -Werte.

Dieses Modellergebnis mit dem Datensatz von EPI Suite wird mit dem gleichen Modell, das auf den hier vorliegenden Datensatz angewendet wird, verglichen. Dieser Datensatz beinhaltet 1675 Stoffe und wird in Kapitel 3.2.3 (Seite 41ff) charakterisiert.

Das Ergebnis aus der einfachen Addition von  $\log H$  und  $\log S_w$  (Gleichung 5.1) ist in Abbildung 5.6 gezeigt. Da die Wasserlöslichkeit über eine in ChemProp [115,116] implementierte Modellauswahl [135] nur für 1657 Stoffe berechnet werden kann, ist auch das Gesamtmodell nur auf diesen leicht eingeschränkten Datensatz anwendbar. Die fehlenden 18 Stoffe sind organische Stoffe, die Arsen, Selen, Blei, Silizium und Zinn beinhalten. Die Statistik ist unten stehend gezeigt.

$$n = 1657, r^2 = 0,88, q^2 = 0,88, rms = 1,48, me = 0,99, mne = -11,37, mpe = 9,01$$

Das H+S-Modell erreicht ein Bestimmtheitsmaß von 0,88 und besitzt somit die gleiche Güte, wie das Modell mit dem EPI Suite-Datensatz. Es ist jedoch auf deutlich mehr Stoffe anwendbar. Am stärksten unterschätzt ist 1,4-Diamino-5-nitroanthracen-9,10-dion, oder C.I. Disperse Violet 8 (siehe Abbildung 5.5), mit einem Fehlerwert von  $-11,37$  (mit Viereck umgeben).

## Dampfdruck-Modellierung

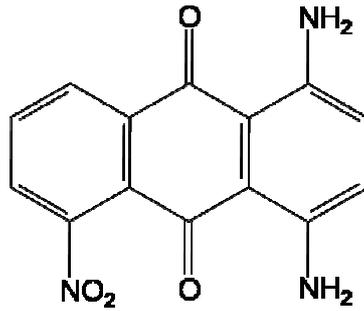


Abbildung 5.5: Strukturformel von 1,4-Diamino-5-nitroanthracen-9,10-dion (C.I. Disperse Violet 8).

Danach folgt 1,4,5,8-Tetraaminoanthrachinon (C.I. Disperse Blue 1, Abbildung 5.3, Seite 116) mit einer Unterschätzung von  $-10,44$ , der auch im Modell zuvor mit dem Datensatz von EPI Suite ein deutlicher Ausreißer ist (eingekreist).

Die Abbildung 5.6 zeigt berechnete gegen experimentelle  $\log P_v$ -Werte und als Idealvorhersage die xy-Gerade als Linie durch die Datenpunkte.

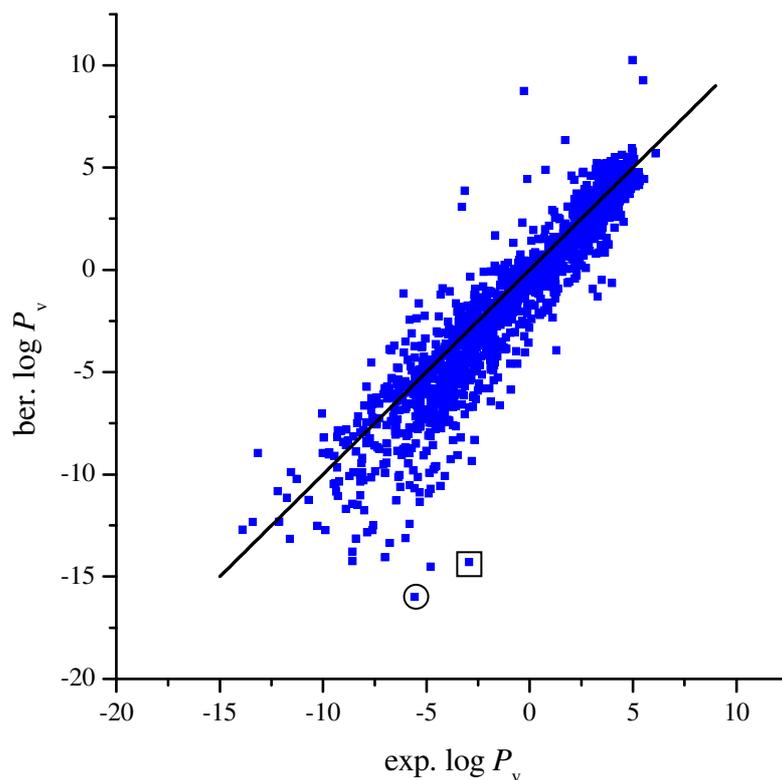


Abbildung 5.6: Auftragung berechneter gegen experimentelle Dampfdruck-Werte; über das H+S-Modell berechnet, wobei die Wasserlöslichkeit aus dem klassenbasierten Modell stammt.

Der Dampfdruck kann auch über andere Modelle vorhergesagt werden wie beispielsweise über das Grain-Watson-Modell. Dieses wird in Kapitel 2.4.2 (Seite 27ff) vorgestellt und ist, wie in

## Dampfdruck-Modellierung

Gleichung (2.41) dargestellt, bekannt. Als Systemtemperatur kann entweder 25 °C für die Raumtemperatur eingesetzt werden oder 37 °C, welche die physiologischen Bedingungen im Körper realitätsgetreu wiedergeben. Deshalb wird das Modell mit beiden Temperaturen gerechnet und die Ergebnisse miteinander verglichen. In Tabelle 5.2 sind die statistischen Parameter bei beiden Temperaturen angegeben. Das Modell ist in der Software ChemProp [115,116] implementiert und kann direkt auf den Datensatz von 1675 Stoffen angewendet werden. Es kann jedoch nicht für alle Substanzen ein Ergebniswert erhalten werden, weshalb sich die Anzahl der Daten  $n$  leicht unterscheidet. Die Güte der beiden Modell-Varianten ist vergleichbar mit Bestimmtheitsmaßen von 0,87 bei 25 °C und 0,86 bei 37 °C. Auch der *rms*-Wert weist nur eine Diskrepanz von 0,17 auf. Der maximale positive Fehler *mpe* ist für 25 °C (7,38) deutlich kleiner als bei 37 °C (12,00). Das leicht bessere Grain-Watson-Modell bei 25 °C ist grafisch in Abbildung 5.7 gezeigt.

**Tabelle 5.2:** Statistik-Ergebnisse des Grain-Watson-Modells zur Vorhersage von  $\log P_v$  bei 25 °C und 37 °C

Statistischer Parameter	Wert bei $T = 25\text{ °C}$	Wert bei $T = 37\text{ °C}$
$n$	1447	1429
$r^2$	0,87	0,86
$q^2$	0,81	0,84
<i>rms</i>	1,65	1,48
<i>bias</i>	-0,46	0,18
<i>me</i>	1,00	0,92
<i>mne</i>	-8,32	-7,63
<i>mpe</i>	7,38	12,00

Bei der grafischen Darstellung der berechneten gegen die experimentellen Dampfdruck-Werte ist deutlich zu erkennen, dass die Streuung um die xy-Gerade bei kleineren  $\log P_v$ -Werten stärker ist, als bei größeren Werten. Das deutet darauf hin, dass die Vorhersage mit dem Grain-Watson-Modell für kleinere (negative) Dampfdruck-Werte ungenauer ist, als für Werte größer Null.

## Dampfdruck-Modellierung

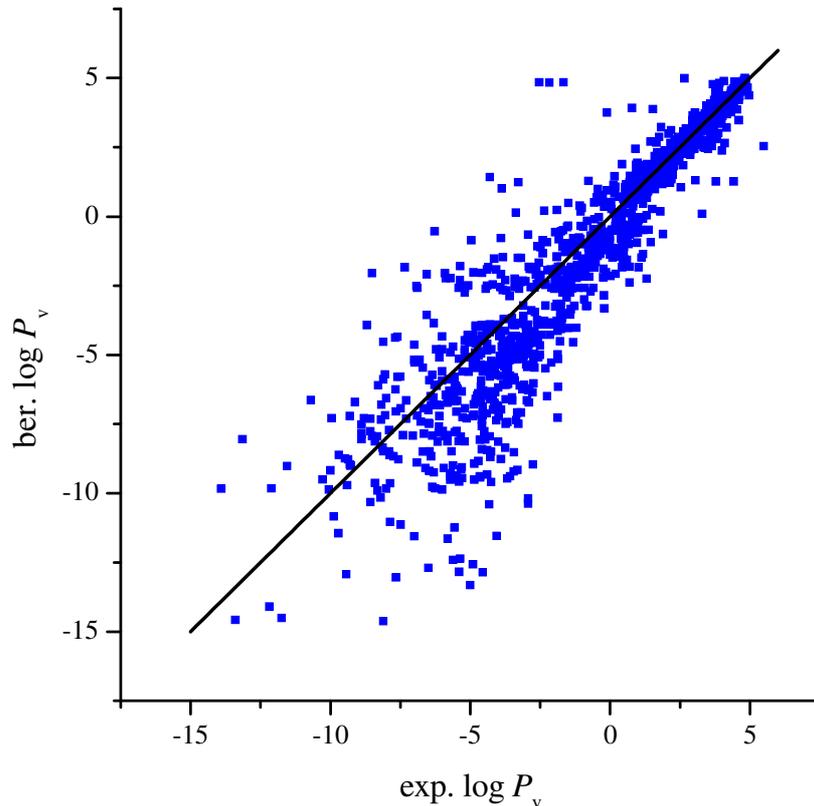


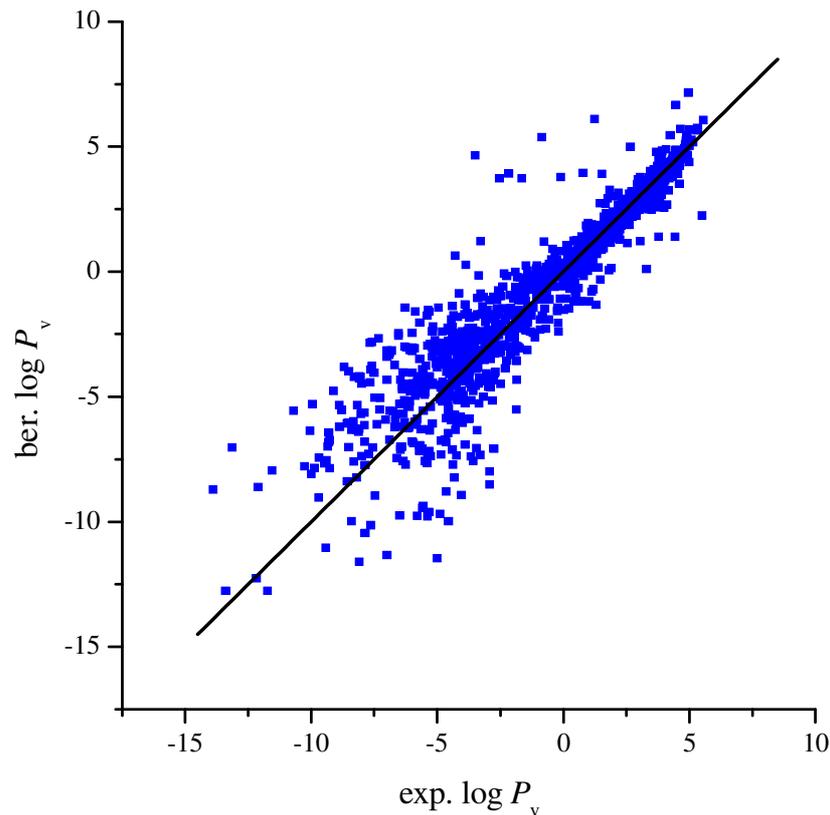
Abbildung 5.7: Auftragung berechnete gegen experimentelle  $\log P_v$ -Werte; berechnet über das Grain-Watson-Modell bei 25 °C.

Als weiteres Modell zur Vorhersage des Dampfdrucks wird das Grain-Antoine-Modell betrachtet. Auch dieses ist in ChemProp [115,116] implementiert und wird in Kapitel 2.4.2 (Seite 27ff) in den Gleichungen (2.32) bis (2.35) erläutert. Angewendet auf den Datensatz, der in Kapitel 3.2.3 (Seite 41ff) charakterisiert ist, kann für 1473 Stoffe eine Vorhersage bei 25 °C mit dem Grain-Antoine-Modell getroffen werden. Die betreffende Statistik ist hier gezeigt.

$$n = 1473, r^2 = 0,88, q^2 = 0,88, rms = 1,32, me = 0,78, mne = -6,48, mpe = 8,15$$

Das Bestimmtheitsmaß ist mit  $r^2 = 0,88$  etwas besser als beim Grain-Watson-Modell und auch der  $rms$  liegt mit 1,32 deutlich unter dem des vorher betrachteten Literatur-Modells. In Abbildung 5.8 ist zu sehen, dass die Streuung um die xy-Gerade etwas weniger stark ist, als in Abbildung 5.7.

## Dampfdruck-Modellierung

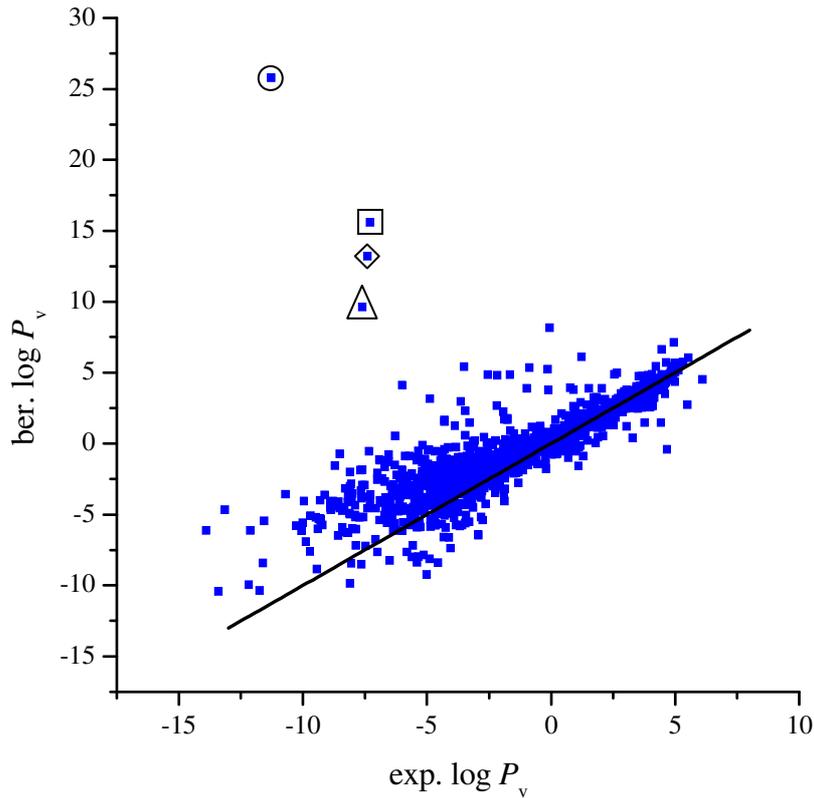


**Abbildung 5.8:** Auftragung der berechneten gegen experimentelle  $\log P_v$ -Werte; berechnet über das Grain-Antoine-Modell bei 25 °C.

Für dieses Modell wurden ebenfalls die zwei Temperaturen 25 °C und 37 °C eingesetzt, wobei als Fazit hervorgeht, dass sich die Ergebnisse der Vorhersage kaum unterscheiden. Aus einem Vergleich der Ergebnisse (grafisch nicht dargestellt) geht hervor, dass keine Parallelverschiebung vorliegt und die Vorhersagewerte mit einer Steigung von 0,92 und einem  $r^2$  von 0,99 sehr stark korrelieren. Daher muss nicht mehr zwischen den zwei Temperaturen unterschieden werden und die weiteren Modelle werden nur noch bei 25 °C gerechnet.

Die Konstanten  $A$ ,  $B$  und  $C$ , die das Grain-Antoine-Modell nutzt, können auch einzeln aus ChemProp [115,116] ausgegeben und das Modell nachgerechnet werden. Trotz der großen Anwendbarkeit auf 1619 Stoffe ist die Vorhersage um einiges schlechter als der vorangegangene Berechnungsansatz.

## Dampfdruck-Modellierung



**Abbildung 5.9:** Auftragung berechnete gegen experimentelle  $\log P_v$ -Werte; berechnet über die Konstanten  $A$ ,  $B$  und  $C$  des Grain-Antoine-Modell.

Deutlich auffällig in Abbildung 5.9 sind die vier Ausreißer, die mit Fehlerwerten von bis zu 37 logarithmischen Einheiten stark überschätzt sind. Hierbei handelt es sich um 2-[N-ethyl-4-(4-nitrophenyl)diazenylanilino]ethanol, auch bekannt als C.I. Disperse Red 1 (**1**), mit einem Fehler von 37,06 (eingekreist), gefolgt von Tribenuronmethyl (**3**) mit einer Überschätzung von 22,89 logarithmischen Einheiten (im Quadrat) sowie 4-Hydroxy-4'-nitroazobenzen (**4**) mit 20,63 logarithmischen Fehler-Einheiten (in der Raute) und mit einem Fehler von 17,24 ist Methyl-7-chloro-2-[[4-trifluoromethoxyphenyl](methoxycarbonyl)amino]carbonyl}-2,9-dihydroindeno-[2,1-E][1,3,4]oxadiazin (**2**) (im Dreieck) zu erwähnen (siehe Abbildung 5.10).

## Dampfdruck-Modellierung

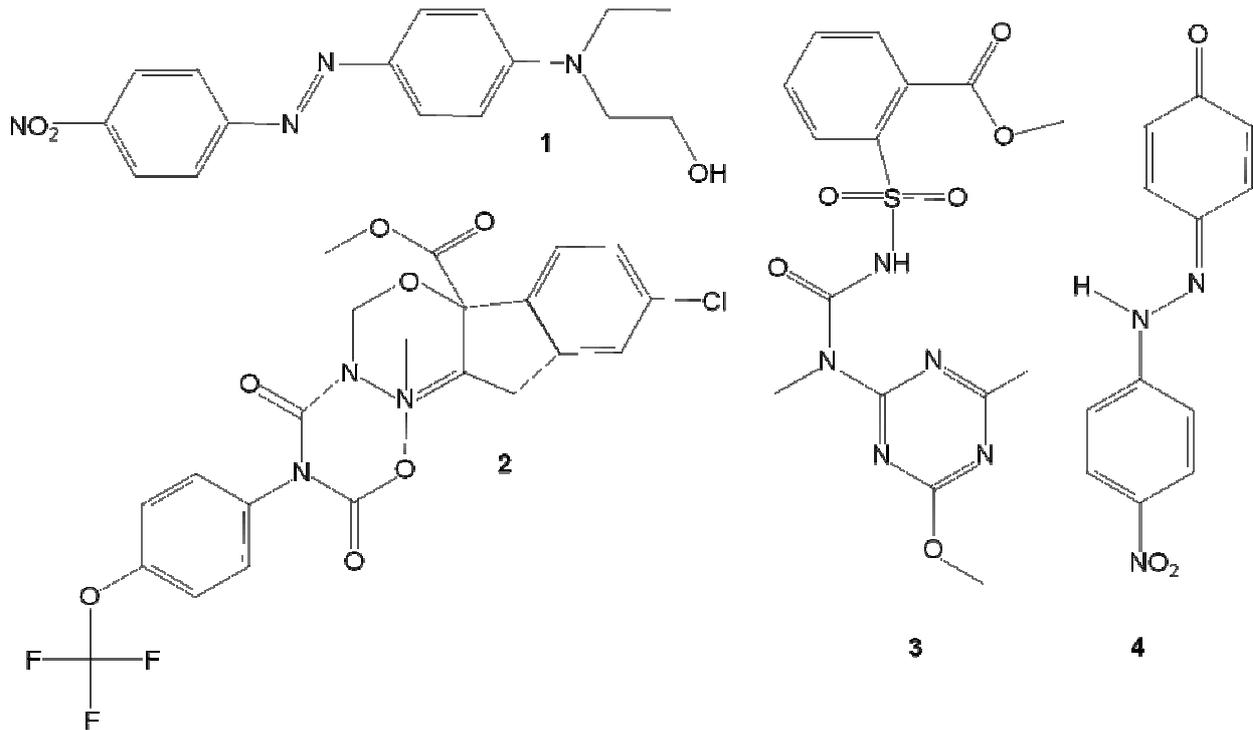


Abbildung 5.10. Strukturformeln von C.I. Disperse Red 1 (1), Tribenuronmethyl (3), 4-Hydroxy-4'-nitroazobenzene (4) und Methyl-7-chloro-2-[[4-(trifluoromethoxyphenyl)(methoxycarbonyl)amino]carbonyl]-2,9-dihydroindeno[2,1-E][1,3,4]oxadiazin (2).

Der größte Unterschied von diesen Stoffen zu den anderen ist der sehr viel kleinere Wert für die Konstante  $C$ , der negativ ist. Da die Konstante  $C$  nur von der Siedetemperatur abhängt und diese bei den besagten Stoffen auch ungewöhnlich klein ist, ist die deutliche Überschätzung darauf zurückzuführen. Das Modell erreicht ein Bestimmtheitsmaß von 0,72. Werden diese vier Ausreißer weggelassen, ergibt sich ein bei weitem besseres Bestimmtheitsmaß von 0,84.

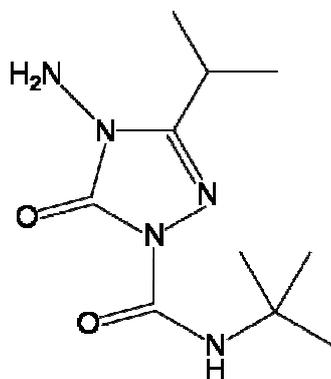
Das letzte Literaturmodell, das hier betrachtet wird, ist von Mackay *et al.* [68] und verwendet nur den Siede- und Schmelzpunkt des untersuchten Stoffes. Es ist in Kapitel 2.4.2 (Seite 27ff) vorgestellt (siehe Gleichung (2.45)). Werden die Werte für den Dampfdruck mit diesem, in ChemProp [115,116] implementierten Modell, vorhergesagt, können folgende Statistikparameter erhalten werden.

## Dampfdruck-Modellierung

**Tabelle 5.3:** Statistische Parameter bei Anwendung des Mackay-Modells auf den hier behandelten Datensatz.

Statistischer Parameter	Wert
$n$	1488
$r^2$	0,88
$q^2$	0,63
$rms$	2,29
$bias$	1,49
$me$	1,55
$mne$	-3,03
$mpe$	13,92

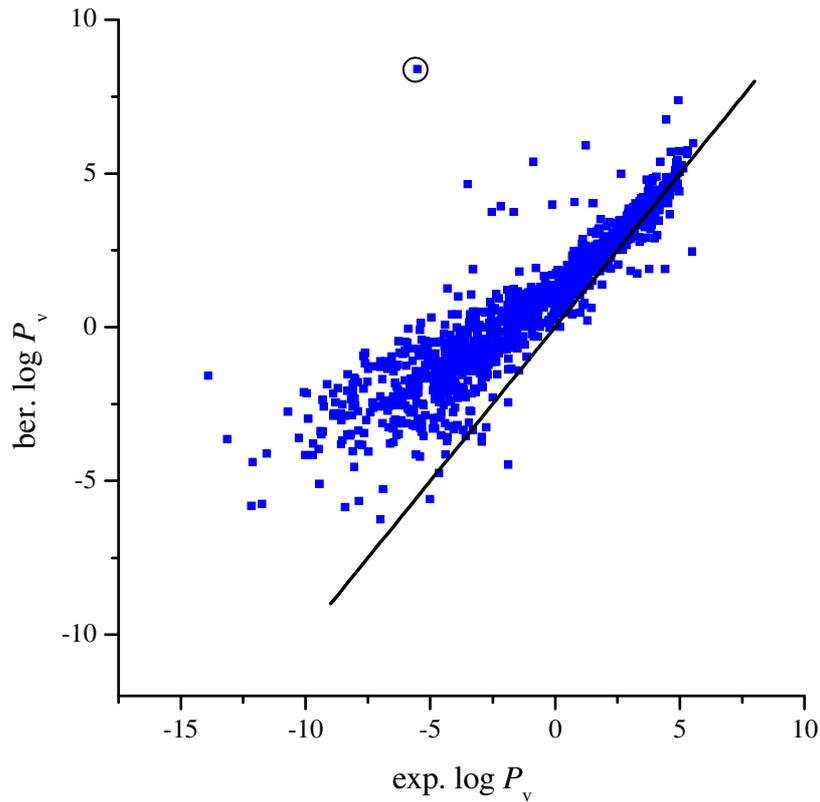
Das Modell ist somit auf 1488 Daten anwendbar und erreicht eine Güte von 0,88. Der größte positive Fehler  $mpe$  liegt bei 13,92, wobei es sich um Amicarbazon (Abbildung 5.11) handelt (eingekreist).



**Abbildung 5.11:** Strukturformel von Amicarbazon.

Der  $rms$  ist trotz der hohen Kalibrierungsgüte des Modells mit 2,29 verhältnismäßig groß. In der Darstellung Abbildung 5.12 sind die berechneten gegen die experimentellen  $\log P_v$ -Werte aufgetragen. Es ist zu erkennen, wie die Datenpunkte um die hier eingezeichnete  $xy$ -Gerade verteilt sind. Substanzen mit hohen  $\log P_v$ -Werten werden besser vorhergesagt, während Stoffe mit negativen Dampfdruck-Werten konstant zu hoch vorhergesagt werden. Auch nimmt die Streuung der Datenpunkte im unteren Wertebereich zu. Eine Neuanpassung und Modifikationen des Mackay-Modells werden in nachfolgendem Kapitel 5.2 genauer untersucht und vorgestellt.

## Dampfdruck-Modellierung



**Abbildung 5.12:** Auftragung der berechneten gegen die experimentellen Dampfdruck-Werte; berechnet über das Mackay-Modell.

Der Vergleich dieser vier vorgestellten Literaturmodelle miteinander, das H+S-Modell, die Ansätze von Grain-Watson, Grain-Antoine und Mackay zeigt, dass die Modell-Kalibrierungsgüte für alle Ansätze ähnlich ist (siehe Tabelle 5.4). Die Vorhersagegüte ist gleich oder annähernd gleich dem Bestimmtheitsmaß. Geringe Unterschiede liegen beim Mackay-Modell vor, bei dem das  $q^2$  nur 0,63 beträgt. Das H+S-Modell besitzt den größten Anwendungsbereich mit 1657 Stoffen im Vergleich zu den anderen, die für 1447 bis 1488 Stoffe anwendbar sind. Dieses Modell weist jedoch den größten negativen Fehler  $mne$  auf mit einem Wert von  $-11,37$ , während das Mackay-Modell mit  $13,92$  den größten positiven Fehler  $mpe$  besitzt.

## Dampfdruck-Modellierung

**Tabelle 5.4:** Vergleich der statistischen Parameter der hier vorgestellten Literatur-Modelle zur Dampfdruck-Vorhersage.

Statistischer Parameter	H+S	Grain-Watson	Grain-Antoine	Mackay
$n$	1657	1447	1473	1488
$r^2$	0,88	0,87	0,88	0,88
$q^2$	0,88	0,81	0,88	0,63
$rms$	1,48	1,65	1,32	2,29
$bias$	-0,65	-0,46	0,20	1,49
$me$	0,99	1,00	0,78	1,55
$mne$	-11,37	-8,32	-6,48	-3,03
$mpe$	9,01	7,38	8,15	13,92

Da die in diesem Kapitel gezeigten Literatur-Modelle nicht sehr aktuell sind, wird geklärt, ob mit dem hier verwendeten Datensatz (siehe Kapitel 3.2.3, Seite 41ff) durch Neukalibrierung und gegebenenfalls leichte Modifizierungen eine statistische Verbesserung möglich ist. Die Versuche einer Modelloptimierung werden mit dem H+S-Modell und dem Mackay-Modell durchgeführt, wobei versucht wird, eine bessere Vorhersage für den Dampfdruck zu erhalten.

### 5.2 Modellierung von $\log P_v$ mit dem H+S-Modell

Das in Kapitel 5.1 (Seite 113ff) erwähnte H+S-Modell wird hier in einer Anpassung der Deskriptoren  $\log H$  und  $\log S_w$  über eine multilineare Regression zur Modellierung des Dampfdrucks angewendet. Es resultiert eine leicht bessere Statistik als bei der Verwendung des Modells (5.1) (Seite 116), bei dem die Parameter lediglich addiert werden. Das Bestimmtheitsmaß bei 1657 Stoffen beträgt 0,89 und der  $rms$  sowie der mittlere Fehler  $me$  belaufen sich auf 1,29 bzw. 0,89. Der maximale negative Fehler liegt bei -8,04 und der maximale positive Fehler bei 8,20 (siehe Modell (5.2),  $\log H$ : EPI Suite [118], HENRYWIN [117],  $\log S_w$ : Modellauswahl [135] aus ChemProp [115,116]).

$$\log P_v = (0,78 \pm 0,01) \cdot \log H + (0,93 \pm 0,01) \cdot \log S_w + (0,90 \pm 0,06) \quad (5.2)$$

$$n = 1657, r^2 = 0,89, q_{CV}^2 = 0,88, rms = 1,29, rms_{CV} = 1,30, me = 0,89, mne = -8,04, mpe = 8,20$$

Dieser Gesamtdatensatz von 1657 Stoffen wird in zwei Teildatensätze aufgeteilt, in Feststoffe und in flüssige Stoffe bei 25 °C. Die Datensätze beinhalten ähnlich viele Stoffe (809 flüssige Substanzen und 848 Feststoffe). Für diese Teildatensätze wird das H+S-Modell neu angepasst. Das sogenannte Flüssigmodell ist in Gleichung (5.3) dargestellt.

## Dampfdruck-Modellierung

$$\log P_v = (0,77 \pm 0,01) \cdot \log H + (0,92 \pm 0,02) \cdot \log S_w + (1,07 \pm 0,08) \quad (5.3)$$

$$n = 809, r^2 = 0,84, q_{CV}^2 = 0,84, rms = 1,30, rms_{CV} = 1,31, me = 0,66, mne = -6,90, mpe = 8,31$$

Die Regressionsparameter sind zur Gesamtanpassung (5.2) fast gleich, jedoch sinkt die Vorhersagegüte von 0,88 auf 0,84. Für die Feststoffe bei 25 °C wird der Ansatz Feststoffmodell genannt und kann auf 848 Stoffe angewendet werden. Das Modell und die Statistik sind in Gleichung (5.4) gezeigt.

$$\log P_v = (0,75 \pm 0,02) \cdot \log H + (0,90 \pm 0,02) \cdot \log S_w + (0,65 \pm 0,10) \quad (5.4)$$

$$n = 848, r^2 = 0,81, q_{CV}^2 = 0,81, rms = 1,47, rms_{CV} = 1,48, me = 1,09, mne = -7,82, mpe = 6,46$$

Die Vorhersagegüte sinkt hierbei weiter auf 0,81 im Vergleich zu den beiden vorangegangenen Modellen. Die Regressionsparameter für  $\log H$  und  $\log S_w$  verkleinern sich leicht, während der  $rms$ -Fehler um 0,17 ansteigt. Somit ist das Feststoffmodell das schlechteste Modell und das Gesamtmodell für 1657 Stoffe das beste Modell, das auch in der Konsensmodellierung in Kapitel 5.4 (Seite 135ff) betrachtet wird. Die beiden Teilmodelle für flüssige und feste Substanzen sind in Abbildung 5.13 grafisch dargestellt. Die Feststoffe sind mit Kreuzen und die flüssigen Stoffe mit Kreisen eingezeichnet.

## Dampfdruck-Modellierung

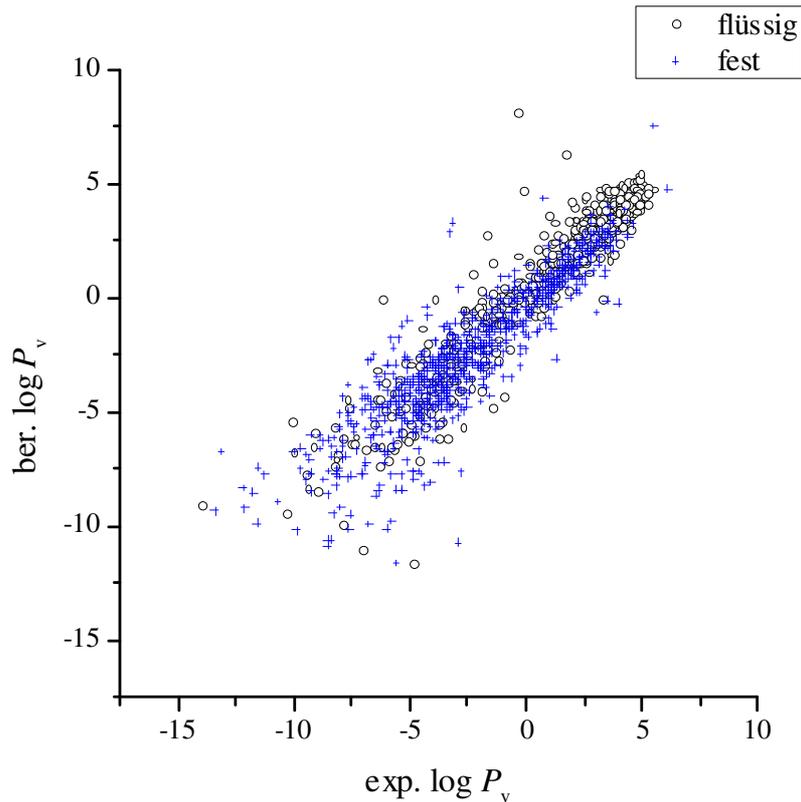


Abbildung 5.13: H+S-Modell mit den Teildatensätzen für flüssige  $\circ$  und feste  $+$  Substanzen im  $P_v$ -Datensatz.

Die festen und flüssigen Substanzen sind relativ über den gesamten Wertebereich verteilt, wobei ein Großteil der Flüssigkeiten im oberen Bereich ( $\log P_v \approx 5$ ) liegt. Zusätzlich sind ein paar wenige Ausreißer zu erkennen.

Die Aufteilung des Datensatzes in flüssige und feste Substanzen ist ein sehr interessanter Ansatz.  $S_w$  ist über die verschiedensten Modelle berechenbar, wodurch die Anwendbarkeit des H+S-Modells hoch ist.

### 5.3 Modellierung von $\log P_v$ mit dem Mackay-Ansatz

Der Mackay-Ansatz beinhaltet die Größen Siedepunkt  $T_b$  und Schmelzpunkt  $T_m$  zur Vorhersage des Dampfdrucks. Da es sich um den Übergang eines Stoffes vom Aggregatzustand flüssig in gasförmig handelt, werden flüssige Substanzen allein mit dem Siedepunkt beschrieben, während für das Modell für feste Stoffe der Schmelz- und Siedepunkt benötigt werden, um  $\log P_v$  vorherzusagen. Dieses Modell ist in Kapitel 2.4.2 (Seite 27ff) vorgestellt und in Gleichung (2.45) (Kapitel 2.4.2, Seite 31) mathematisch dargestellt. Das Ergebnis dieser Modellanwendung wird in Kapitel 5.1 (Seite 113ff) diskutiert und mit anderen Ansätzen verglichen.

Ausgehend von dieser Mackay-Gleichung wird zunächst eine Vereinfachung vorgenommen indem die Entropieterme entfernt werden. Dafür werden die Regressionsparameter  $a$  und  $b$

## Dampfdruck-Modellierung

eingeführt, die über eine multilineare Regression an die verschiedenen Modifikationen jeweils angepasst werden, wobei zusätzlich die Konstante  $c$  hinzukommt (Gleichung (5.5)).

$$\log P_v = a \cdot \left(\frac{T_b}{T} - 1\right) + b \cdot \left(\frac{T_m}{T} - 1\right) + c \quad (5.5)$$

Da  $T_m$  eine schwer zu berechnende Größe ist, spielt die Güte dieses Parameters für das Feststoff-Modell eine entscheidende Rolle. Wird in der Ausgangsgleichung (5.5) der Deskriptor  $T_m$  aus dem Programm MPBPWIN [90] von EPI Suite [118] gewonnen, kann für alle 1675 Substanzen ein experimenteller Wert erhalten bzw. die fehlenden mit berechneten ergänzt werden. Mit den Modellen aus ChemProp [115,116] kann nicht für alle Substanzen ein Wert erhalten werden, weshalb dieses Programm hier keine Anwendung findet. Mit dieser Anpassung (Gleichung (5.6)) wird ein großer Datensatz von 1675 Stoffen kalibriert und das Modell erreicht eine Güte von  $r^2 = 0,91$  und einen *rms*-Fehler von 1,16.

$$\log P_v = (-6,85 \pm 0,11) \cdot \left(\frac{T_b}{T} - 1\right) - (2,49 \pm 0,15) \cdot \left(\frac{T_m}{T} - 1\right) - (3,85 \pm 0,12) \quad (5.6)$$

$$n = 1675, r^2 = 0,91, q_{CV}^2 = 0,91, rms = 1,16, rms_{CV} = 1,16, me = 0,79, mne = -7,86, mpe = 8,90$$

Die Modell-Qualität ist für diesen großen Datensatz durchaus zufriedenstellend. Durch die Anpassung werden mehr Freiheitsgrade sowie Flexibilität erhalten und zugleich verbessert sich das Mackay-Modell (vgl. Tabelle 5.3, Seite 124) deutlich. Der *rms*-Fehler sinkt von 2,29 auf 1,16 und auch der mittlere absolute Fehler *me* sinkt von 1,55 auf 0,79. Zusätzlich steigt die Vorhersagegüte von 0,63 auf 0,91, was eine erhebliche Verbesserung darstellt. Findet ein Vergleich der theoretischen Entropieterme aus der Literatur [68] ( $-4,6$  für  $a$  und  $-2,95$  für  $b$ ) mit den Regressionsparametern in der neu angepassten Gleichung (5.6) statt, lässt sich feststellen, dass diese sehr ähnlich zueinander sind. Vor allem der Vorfaktor des Schmelzpunkt-Terms besitzt fast den gleichen Wert.

Der Gesamtdatensatz von 1675 Substanzen wird ebenfalls in flüssige und feste Stoffe aufgeteilt und für die Teildatensätze der Mackay-Ansatz neu angepasst. Dabei enthält der Datensatz für die flüssigen Stoffe 814 Substanzen und der zweite Teildatensatz besteht aus 861 Feststoffen. Somit ist die Aufteilung fast eins zu eins. Gleichung (5.7) zeigt das Modell für die 814 flüssigen Stoffe.

## Dampfdruck-Modellierung

$$\log P_v = (-6,80 \pm 0,10) \cdot \left(\frac{T_b}{T} - 1\right) - (0,25 \pm 0,18) \cdot \left(\frac{T_m}{T} - 1\right) - (1,24 \pm 0,18) \quad (5.7)$$

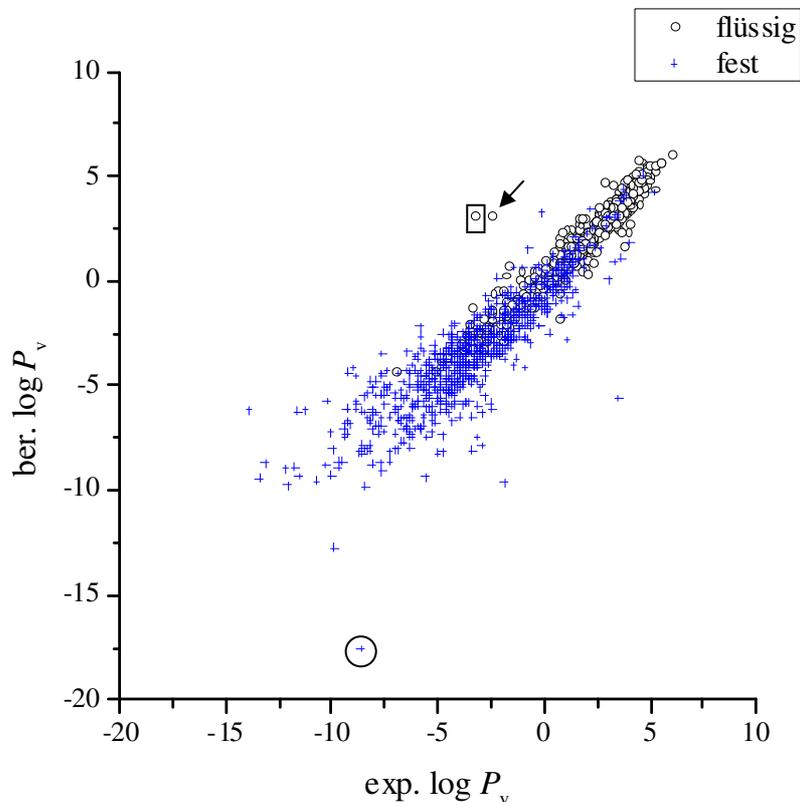
$$n = 814, r^2 = 0,91, q_{CV}^2 = 0,91, rms = 0,56, rms_{CV} = 0,57, me = 0,35, mne = -6,40, mpe = 2,57$$

Der Regressionskoeffizient für den  $T_m$ -Term ist nicht signifikant und die Vorhersagegüte bleibt gleich zu dem in vorherigem Modell. Die Fehler sind alle kleiner als in Modell (5.6). Der Regressionskoeffizient für den  $T_b$ -Term ist ähnlich zu dem in dem Modell davor. Findet diese Neukalibrierung mit den 861 Feststoffen aus dem  $P_v$ -Datensatz statt, ergibt sich Gleichung (5.8).

$$\log P_v = (-7,25 \pm 0,18) \cdot \left(\frac{T_b}{T} - 1\right) - (3,54 \pm 0,28) \cdot \left(\frac{T_m}{T} - 1\right) - (4,42 \pm 0,19) \quad (5.8)$$

$$n = 861, r^2 = 0,76, q_{CV}^2 = 0,76, rms = 1,45, rms_{CV} = 1,46, me = 1,05, mne = -9,09, mpe = 7,71$$

Die Statistik verschlechtert sich deutlich, was zeigt, dass die Anpassung des Mackay-Modells für Feststoffe so nicht optimal ist. Werden die berechneten gegen die experimentellen  $\log P_v$ -Daten für die Teildatensätze aufgetragen, ergibt sich Abbildung 5.14, in der die Feststoffe mit Kreuze und die flüssigen Verbindungen mit Kreisen dargestellt sind.



**Abbildung 5.14:** Mackay-Modell mit den Teildatensätzen für flüssige  $\circ$  und feste  $+$  Substanzen im  $P_v$ -Datensatz.

## Dampfdruck-Modellierung

Hierbei ist ersichtlich, dass die flüssigen Stoffe größtenteils im oberen Bereich liegen, während die Feststoffe weit ins Negative reichen. Es zeigen sich zwei überschätzte Ausreißer für die flüssigen Stoffe, bei denen es sich um Trimethylbleichlorid (Abbildung 5.15) (mit einem Pfeil markiert) und Methansulfonylfluorid (Abbildung 5.15) (mit Rechteck umgeben) handelt.

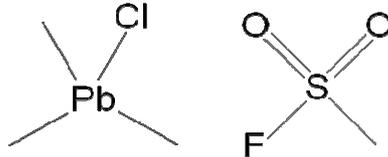


Abbildung 5.15: Strukturformeln von Trimethylbleichlorid und Methansulfonylfluorid.

Bei den Feststoffen sind drei Ausreißer zu erkennen, die deutlich unterschätzt werden. Der größte ist 1,4-Difluorobenzene (Abbildung 5.16) mit einem Fehler vom 9,09 (eingekreist in Abbildung 5.14).

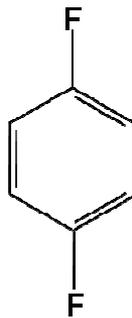


Abbildung 5.16: Strukturformel von 1,4-Difluorbenzol.

Als weitere Variation des Mackay-Ansatzes wird der zweite Term aus Gleichung (5.5) durch die molekulare Masse MW ersetzt, da der Schmelzpunkt stark von dieser abhängt und deshalb möglicherweise gut durch diese Hilfsgröße beschrieben werden kann. Diese Idee stammt aus dem Programm EPI Suite [118]. Die daraus resultierende Gleichung ist in (5.9) gezeigt.

$$\log P_v = (-8,52 \pm 0,14) \cdot \left(\frac{T_b}{T} - 1\right) - (0,0009 \pm 0,0006) \cdot MW - (2,08 \pm 0,14) \quad (5.9)$$

$$n = 1675, r^2 = 0,89, q_{CV}^2 = 0,89, rms = 1,24, rms_{CV} = 1,25, me = 0,84, mne = -7,57, mpe = 9,13$$

Das Modell ist auf alle 1675 Stoffe anwendbar und erreicht ein  $r^2$  und  $q_{CV}^2$  von 0,89. Der kleine Regressionsparameter für die molekulare Masse ist nicht signifikant, was darauf hindeutet, dass

## Dampfdruck-Modellierung

MW eine äußerst geringe Rolle im Modell (5.9) spielt und theoretisch auch weggelassen werden kann. Wird diese Annahme in die Praxis umgesetzt, ergibt sich für den Gesamtdatensatz die Modellgleichung (5.10) ohne den MW-Term. Dieses kurze Mackay-Modell enthält somit nur den Term der Siedetemperatur, was dem ersten Term der Mackay-Gleichung entspricht.

$$\log P_v = (-8,34 \pm 0,07) \cdot \left(\frac{T_b}{T} - 1\right) - (1,87 \pm 0,03) \quad (5.10)$$

$$n = 1675, r^2 = 0,89, q_{CV}^2 = 0,89, rms = 1,25, rms_{CV} = 1,25, me = 0,84, mne = -7,52, mpe = 8,87$$

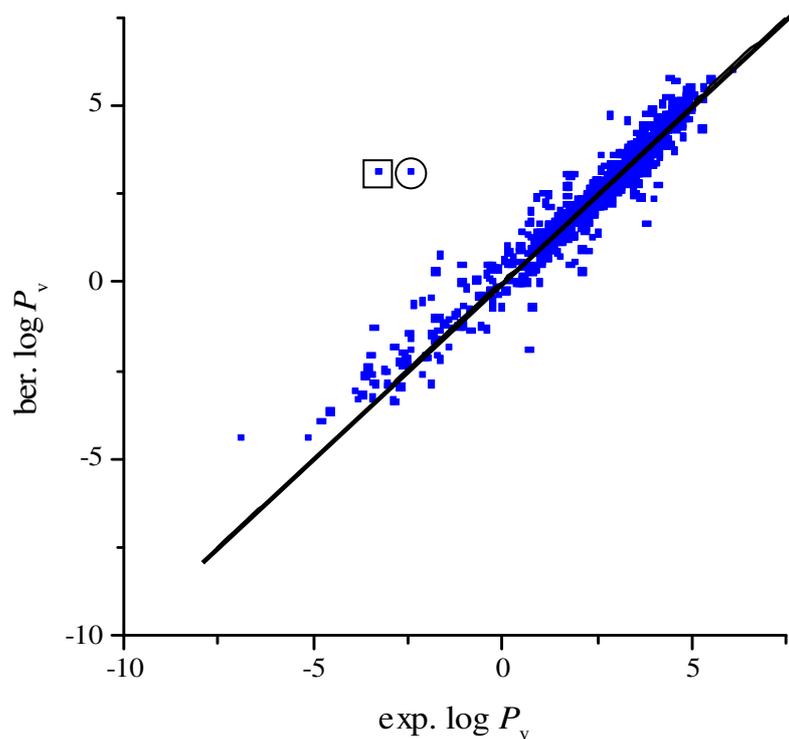
Die Statistik ist annähernd gleich zum Modell mit MW-Term (5.9) und für die große Anzahl an Stoffen sehr zufriedenstellend. Das zeigt, dass der zweite Term, der Feststoffterm, der Mackay-Gleichung nicht notwendig ist und der Dampfdruck nur mit dem Siedepunkt  $T_b$  beschrieben werden kann. Die Modellierung des Dampfdrucks mit dem Mackay-Ansatz könnte beendet werden. Es werden jedoch erneut die Teildatensätze der Stoffe für feste und flüssige Substanzen auf das Modell (5.10) angewendet. Die vereinfachte Gleichung, in der nur der erste Term der Ursprungsgleichung vorhanden ist, beschreibt den Phasenübergang von flüssig zu gasförmig und ist für Flüssigkeiten geeignet. Das Flüssigmodell ist für die 814 flüssigen Stoffe in Gleichung (5.11) dargestellt.

$$\log P_v = (-6,88 \pm 0,07) \cdot \left(\frac{T_b}{T} - 1\right) - (0,99 \pm 0,04) \quad (5.11)$$

$$n = 814, r^2 = 0,91, q_{CV}^2 = 0,91, rms = 0,56, rms_{CV} = 0,57, me = 0,35, mne = -6,42, mpe = 2,61$$

Bei einer Systemtemperatur von 25 °C erreicht es eine hohe Vorhersagegüte von 0,91 und einen sehr kleinen mittleren absoluten Fehler von 0,35. Die Fehler sind deutlich kleiner, als im Gesamtmodell (5.10). Auch in der grafischen Darstellung, gezeigt in Abbildung 5.17, lässt sich die hohe Güte dieses Vorhersagemodells für den Dampfdruck erkennen.

## Dampfdruck-Modellierung



**Abbildung 5.17:** Auftragung der vorhergesagten gegen experimentelle  $\log P_v$ -Werte für das Flüssigmodell (5.11).

Lediglich zwei Ausreißer liegen oberhalb der xy-Geraden, bei denen es sich wie im vorherigen Mackay-Modell auch um Trimethylbleichlorid mit einem Fehler von 5,53 logarithmischen Einheiten (eingekreist) und Methansulfonylfluorid (siehe Abbildung 5.15, Seite 131) mit einer Überschätzung von 6,42 logarithmischen Einheiten (mit Rechteck umgeben) handelt.

Außerdem ist in Abbildung 5.17 zu beobachten, dass im Datensatz mehr  $\log P_v$ -Werte größer Null vorhanden sind, diese enger zusammen liegen und weniger um die xy-Gerade streuen als die Werte kleiner Null.

Das vereinfachte und gekürzte Mackay-Modell, das nur die Siedetemperatur  $T_b$  nutzt und für die Anwendung von flüssigen Stoffen geeignet ist, wird aus exploratorischen Gründen ebenfalls für die 861 Feststoffe des  $P_v$ -Datensatzes gerechnet. Es liefert ein deutlich schlechteres Ergebnis als das Gesamt- (5.10) und Flüssigmodell (5.11) wie in Gleichung (5.12) zu sehen ist.

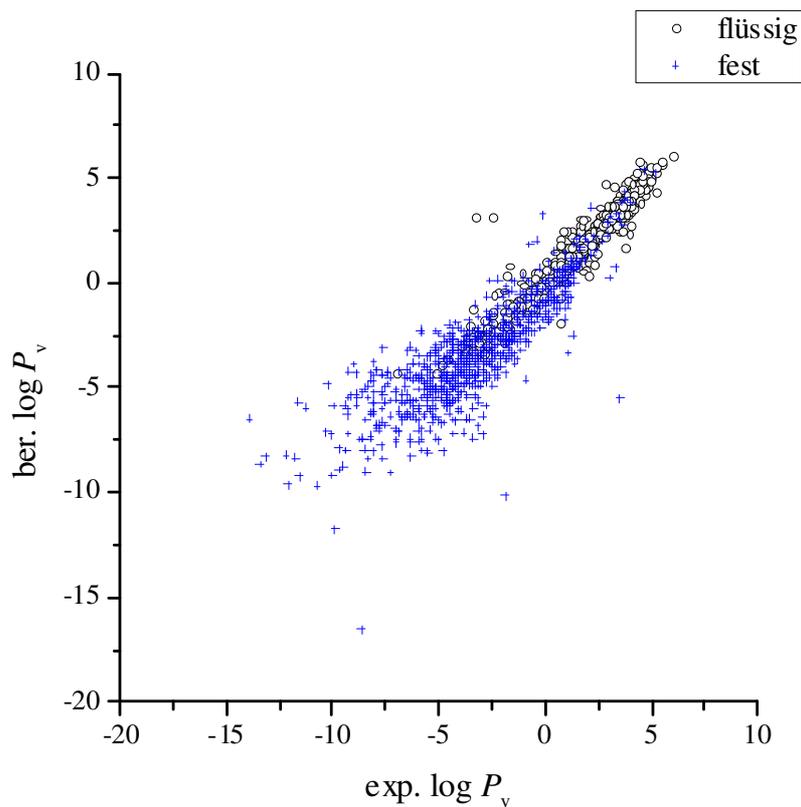
$$\log P_v = (-8,28 \pm 0,18) \cdot \left(\frac{T_b}{T} - 1\right) - (2,04 \pm 0,06) \quad (5.12)$$

$$n = 861, r^2 = 0,71, q_{CV}^2 = 0,71, rms = 1,59, rms_{CV} = 1,60, me = 1,17, mne = -7,38, mpe = 9,01$$

Für die Verbindungen, die bei Raumtemperatur im festen Aggregatzustand vorliegen, ist im Modell normalerweise sowohl der Anteil des Schmelzens, als auch der des Siedens enthalten

## Dampfdruck-Modellierung

(siehe Gleichung (5.5)). Beim kurzen Mackay-Modell sind nur der Siedepunkt und die Systemtemperatur enthalten und es zeigt für die Feststoffe eine schlechtere Statistik auf als für die flüssigen Substanzen oder auf den Gesamtdatensatz angewendet. Das könnte bedeuten, dass die Vorhersage des Dampfdrucks für Feststoffe mit diesem Modell nicht durchgeführt werden kann. Auch für das kurze Mackay-Modell werden für die Teildatensätze von festen (Kreuze) und flüssigen (Kreise) Verbindungen die berechneten gegen die experimentellen  $\log P_v$ -Werte aufgetragen (siehe Abbildung 5.18).



**Abbildung 5.18: Kurzes Mackay-Modell ohne den zweiten Term mit den Teildatensätzen für flüssige  $\circ$  und feste  $+$  Substanzen im  $P_v$ -Datensatz.**

Die Abbildung 5.14 zeigt ein sehr ähnliches Bild zum normalen Mackay-Modell. Zudem sind auch die Fehler der fünf deutlichen Ausreißer ähnlich. Auffallend ist, dass die Feststoffe hier mit einer polynomialen Anpassung besser beschrieben werden könnten als mit einer linearen Regression.

Die Dampfdruck-Vorhersage mit den beiden hier vorgestellten Mackay-Modellen ist für flüssige Substanzen deutlich besser als für Feststoffe. Da  $P_v$  direkt für die Vorhersage des PBPK-relevanten Verteilungskoeffizienten  $K_{ba}$  verwendet werden kann und dafür vorwiegend VOCs betrachtet werden, spielen die Feststoffe demzufolge eine untergeordnete Rolle. Die

## **Dampfdruck-Modellierung**

Flüssigmodelle zur Vorhersage des Dampfdrucks können somit direkt auf die Vorhersage von Verteilungskoeffizienten angewendet werden.

### **5.4 Modellvergleiche für den Dampfdruck**

Die entwickelten Dampfdruck-Modelle für den Gesamt- und die Teildatensätze aus den Kapiteln 5.2 (Seite 126ff) und 5.3 (Seite 128ff) werden in Tabelle 5.5 mit Regressions- und Statistikparameter gegenübergestellt.

## Dampfdruck-Modellierung

**Tabelle 5.5:**  $P_v$ -Modelle mit Regressions- und Statistikparameter.

Modell	Stoffe	$a$	$b$	$c$	$n$	$r^2$	$q_{CV}^2$	$rms$	$rms_{CV}$	$me$	$mne$	$mpe$
H+S	flüssig und fest	$0,78 \pm 0,01$	$0,93 \pm 0,01$	$0,90 \pm 0,06$	1657	0,89	0,88	1,29	1,30	0,89	-8,04	8,20
H+S	flüssig	$0,77 \pm 0,01$	$0,92 \pm 0,02$	$1,07 \pm 0,08$	809	0,84	0,84	1,30	1,31	0,66	-8,31	6,90
H+S	fest	$0,75 \pm 0,02$	$0,90 \pm 0,02$	$0,65 \pm 0,10$	848	0,81	0,81	1,47	1,48	1,09	-6,46	7,82
Mackay	flüssig und fest	$-6,85 \pm 0,11$	$-2,49 \pm 0,15$	$-3,85 \pm 0,12$	1675	0,91	0,91	1,16	1,16	0,79	-7,86	8,90
Mackay	flüssig	$-6,80 \pm 0,10$	$-0,25 \pm 0,18$	$-1,24 \pm 0,18$	814	0,91	0,91	0,56	0,57	0,35	-6,40	2,57
Mackay	fest	$-7,25 \pm 0,18$	$-3,54 \pm 0,28$	$-4,42 \pm 0,19$	861	0,76	0,76	1,45	1,46	1,05	-7,71	9,09
Mackay, kurz	flüssig und fest	$-8,34 \pm 0,07$	$-1,87 \pm 0,03$	-	1675	0,89	0,89	1,25	1,25	0,84	-7,52	8,87
Mackay, kurz	flüssig	$-6,88 \pm 0,07$	$-0,99 \pm 0,04$	-	814	0,91	0,91	0,56	0,57	0,35	-6,42	2,61
Mackay, kurz	fest	$-8,28 \pm 0,18$	$-2,04 \pm 0,06$	-	861	0,71	0,71	1,59	1,60	1,17	-7,38	9,01

## Dampfdruck-Modellierung

Das Mackay-Modell mit Einsatz von Siedepunkt und Schmelzpunkt ist das Beste dieser hier untersuchten Modelle. Die Flüssigmodelle des Mackay- und des kurzen Mackay-Ansatzes zeigen die beste Vorhersagegüte und die geringsten *rms*-Fehler. Bei diesen beiden Modellen ist es nicht relevant, ob eine Raumtemperatur von 25 °C oder eine physiologische Temperatur von 37 °C angenommen wird. Für die Systemtemperatur wurde Raumtemperatur gewählt, die Modelle können jedoch für PBPK-Modelle eingesetzt werden. Über den Dampfdruck können die Verteilungskoeffizienten  $K_{wa}$  und  $K_{ba}$  gewonnen werden. Somit ist die Modellierung von  $P_v$  ein Nebenweg und für die PBPK-Modelle konkret brauchbar.

Die Modelle der Feststoffe zeigen die schlechteste Statistik mit den größten Fehlern. Diese hier untersuchten multilinearen Modelle sind vermutlich nicht der richtige Ansatz, um den Dampfdruck von Feststoffen zu modellieren, was eine wichtige Aussage zur aktuellen  $P_v$ -Modellierbarkeit von Feststoffen ist. In Abbildung 5.18 (Kapitel 5.3, Seite 134) deutet sich an, dass eventuell eine polynomiale Herangehensweise zu einem Erfolg führen könnte.

Der Dampfdruck  $P_v$  findet Verwendung für den wichtigen PBPK-Verteilungskoeffizienten  $K_{wa}$  (siehe Kapitel 3.3.2, Seite 45f). In Gleichung (2.28) (Kapitel 2.4.1, Seite 24ff) ist die Umrechnung vom Dampfdruck zum Henry-Koeffizienten und  $K_{aw}$  gezeigt. Die Wasserlöslichkeit  $S_w$  ist über viele Modelle berechenbar und wird mit  $P_v$  für die Berechnung von  $K_{aw}$  eingesetzt (siehe Kapitel 2.4.1, Gleichung (2.28) und (2.29) , Seite 24ff). Diese Berechnung entspricht dem H+S-Modell.

### 6 Zusammenfassung und Ausblick

Ziel der Arbeit ist es, Modelle für die Verteilungskoeffizienten Blut/Luft und Fett/Luft zu erstellen, mit denen der für PBPK-Modelle relevanten Fett/Blut-Verteilungskoeffizienten zu berechnen. Da der inverse dimensionslose Henry-Koeffizient dabei zum Einsatz kommt und dieser aus dem Dampfdruck und der Wasserlöslichkeit gewonnen werden kann, wird die Modellierung des Dampfdrucks ebenfalls ins Auge gefasst. Der Dampfdruck ist damit ein wichtiger PBPK-relevanter Parameter, der in direktem Zusammenhang mit dem Wasser-/Luft- und zugleich dem Blut/Luft-Verteilungskoeffizienten steht.

Für alle Verteilungskoeffizienten und für den Dampfdruck wurden experimentelle Daten aus der Literatur gesammelt und die Datensätze in bezüglich ihrer Atomtypen, Komplexität und Polarität charakterisiert. Ausgehend von diesen Daten wurden Literaturansätze angepasst und neue multilineare Modelle erstellt.

Ein wichtiger Ansatz für die Modellierung von Verteilungskoeffizienten ist das LSER-Modell von Abraham *et al.* [26], welches für  $\log K_{ba}$  und  $\log K_{fa}$  bereits in der Literatur vorliegt. Diese Ansätze werden miteinander verglichen, wobei die Beiträge der einzelnen Deskriptoren diskutiert werden. Zur Blut/Luft-Modellierung wird das Literaturmodell angewendet und neu angepasst und es wird jeweils die Verwendung von experimentellen und berechneten Deskriptoren miteinander verglichen. Die Modelle mit experimentellen Deskriptoren erzielen die beste Güte mit einem Bestimmtheitsmaß von 0,92. Da Vorhersagemodelle aber ohne experimentelle Werte auskommen sollen, werden nur die näher betrachtet und diskutiert, die berechnete Deskriptoren verwenden. Hier beträgt die Modellgüte des LSER-Modells für  $\log K_{ba}$   $r^2 = 0,77$ . Auf den Fett/Luft-Verteilungskoeffizienten wird das Literaturmodell ebenfalls angewendet und eine Neuanpassung durchgeführt. Hierbei wird eine Modellgüte von  $r^2 = 0,96$  mit experimentellen Deskriptoren und von  $r^2 = 0,87$  mit berechneten Deskriptoren erreicht. Diese Ergebnisse zeigen, dass in der Modellierung von Verteilungskoeffizienten durchaus Potential zu Verbesserung vorhanden ist, weshalb weitere Ansätze der Modellierung verfolgt werden.

Es werden andere Verteilungskoeffizienten zur Modellierung von  $K_{ba}$  herangezogen, bei denen nach einer Analyse der Relevanz in Bezug auf die Zusammensetzung des Blutes die geeignetsten Verteilungskoeffizienten eingesetzt werden. Diese sind für den Wasserteil im Blut der Wasser/Luft-Verteilungskoeffizient und für den organischen Teil  $K_{ow}$ ,  $K_{HSA}$ ,  $K_{mw}$  sowie  $K_{olive}$ . Der organische Teil im Blut kann mit einem oder zwei der vier Verteilungs-

## Zusammenfassung und Ausblick

koeffizienten beschrieben werden. Ein Vergleich der Ergebnisse miteinander zeigt, dass die  $K_{ba}$ -Modelle mit anderen Verteilungskoeffizienten im besten Fall ein Bestimmtheitsmaß von 0,85 ergeben. Die Modellgüte übertrifft damit die des LSER-Ansatzes. Für den Blut/Luft-Verteilungskoeffizienten ermöglichen zusätzliche Betrachtungsweisen und Gedankenansätze ein tiefergehendes Verständnis für die Möglichkeiten und Anwendungen dieser Modelle.

Das Fragmentmodell mit 16 Fragmenten erzielt für den Blut/Luft-Verteilungskoeffizienten ein noch besseres Bestimmtheitsmaß ( $r^2 = 0,91$ ) als der Abraham-Ansatz und die Verwendung von anderen Verteilungskoeffizienten.

Das über schrittweise Regression erhaltene  $K_{ba}$ -Modell mit elektrotologischen Indices nutzt vier E-States und weist mit einem Bestimmtheitsmaß von 0,63 eine deutlich schlechtere Güte als alle anderen Modelle auf.

Im Falle der  $\log K_{fa}$ -Modellierung wird die Komponente Fett durch Oktanol und Olivenöl dargestellt. Das 2-Parameter-Modell mit den Verteilungskoeffizienten Oktanol/Luft und Olivenöl/Luft erreicht ein Bestimmtheitsmaß und Vorhersagequalität von 0,91 für 96 Verbindungen.

Zusammenfassend ist die Modellierung der Verteilungskoeffizienten  $K_{ba}$  und  $K_{fa}$  mit dieser Arbeit erfolgreich abgeschlossen. Es können jedoch weitere interessante Vorgehensweisen verfolgt werden, die auch in dieser Arbeit vorgestellt werden. Zum einen wird die Anwendung des Analogiekonzeptes Read-across, beispielsweise mit atomzentrierten Fragmenten erörtert, und zum anderen der vektorgestützten Methode SVM. Zudem ist hier deutlich geworden, dass die Güte der Eingangsparameter von ausschlaggebender Wichtigkeit ist und sich somit eine Modellverbesserung durch Optimierung der Eingangsparameter ergeben würde.

Der Fett/Blut-Verteilungskoeffizient wird aus den anderen beiden errechnet und die Ergebnisse der Berechnung vorgestellt. In der Literatur sind nur für sehr wenige Stoffe experimentelle Daten verfügbar und daher ist die Berechnung nicht sehr aussagekräftig. Deshalb wird aus den Datensätzen für  $\log K_{ba}$  und  $\log K_{fa}$  ein Datensatz an quasi-experimentellen Daten für  $\log K_{fb}$  erzeugt. Mit dem neu erstellten Datensatz von 96 Substanzen kann keine überzeugende Statistik erreicht werden. Im besten Fall, unter Einsatz des Fragmentmodells für  $K_{ba}$  und des LSER-Ansatzes für  $K_{fa}$ , wird ein Bestimmtheitsgrad von 0,77 erlangt.

## Zusammenfassung und Ausblick

Diese hier behandelten Verteilungskoeffizienten können in die eingangs vorgestellten PBPK-Modelle eingesetzt werden und damit wichtige physiko-chemische Prozesse und der Verbleib von Chemikalien im Körper vorhergesagt werden.

Auch der Dampfdruck spielt eine entscheidende Rolle für PBPK-Modelle, da er in direktem Zusammenhang zur Henry-Konstante und damit zum Wasser/Luft-Verteilungskoeffizienten steht. Für die Modellierung des Dampfdrucks wird zunächst ein Überblick und Vergleich der Literaturmethoden dargestellt. Anschließend werden das H+S-Modell und die Optimierung des Mackay-Ansatzes betrachtet. Hierbei wird die Erkenntnis gewonnen, dass der Dampfdruck für die großen Datensätze von über 1650 Verbindungen zufriedenstellend vorhergesagt werden kann. Die Ergebnisse mit dem Mackay-Ansatz sind generell besser als mit dem H+S-Modell. Bei einer Aufteilung der Datensätze in flüssige und feste Stoffe wird deutlich, dass die Modelle für die flüssigen Substanzen eine bessere Vorhersagequalität aufweisen als die für Feststoffe. Möglicherweise sind Dampfdruckmodelle nicht auf Feststoffe anwendbar. Zur Verifizierung dieser These müssten weitere Dampfdruck-Modelle wie die in der Theorie angeführten Ansätze nach Grain-Watson und Grain-Antoine oder mit dem in der Literatur erwähnten Modell von Sepassi *et al.* [62] untersucht werden.

Über den Dampfdruck kann der Verteilungskoeffizient  $K_{wa}$  bestimmt werden, der zur Beschreibung des Blutkompartiments in PBPK-Modellen eine wichtige Rolle spielt. Daher ist als Ausblick zu empfehlen, dass der  $P_v$ -Einsatz zur Berechnung von  $K_{wa}$  untersucht wird und nachfolgend erneut die Konsensmodellstrategie angewendet wird.

### 7 Literaturverzeichnis

- [1] Peyret T, Krishnan K 2011. QSARs for PBPK modelling of environmental contaminants. *Sar Qsar Environ. Res.* 22: 1-2, 129-169.
- [2] Miwa GT 1995. Goals, Design, Timing, and Future Opportunities for Nonclinical Drug-Metabolism Studies. *Toxicol. Pathol.* 23: 2, 131-135.
- [3] Frantz SW, Beatty PW, English JC, Hundley SG, Wilson AGE 1994. The Use of Pharmacokinetics As An Interpretive and Predictive Tool in Chemical Toxicology Testing and Risk Assessment - A Position Paper on the Appropriate Use of Pharmacokinetics in Chemical Toxicology. *Regul. Toxicol. Pharm.* 19: 3, 317-337.
- [4] Wilson AGE, Frantz SW, Keifer LC 1994. A Tiered Approach to Pharmacokinetic Studies. *Environ. Health Persp.* 102: 5-11.
- [5] Caldwell J, Gardner I, Swales N 1995. An Introduction to Drug Disposition - the Basic Principles of Absorption, Distribution, Metabolism, and Excretion. *Toxicol. Pathol.* 23: 2, 102-114.
- [6] 2-FUN, 2011, Full-chain and UNcertainty Approaches for Assessing Health Risks in Future ENvironmental Scenarios. EU Project, contract no. FP6-2005-GLOBAL-4-036976, 2007-2011. 2-FUN project website: <http://www.2-fun.org>.
- [7] Gelman A, Bois F, Jiang JM 1996. Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *J. Am. Stat. Assoc.* 91: 436, 1400-1412.
- [8] Teorell T 1937. Kinetics of distribution of substances administered to the body. *Arch. Int. Pharmacod. T.* 57: 205-240.
- [9] Bois FY, Jamei M, Clewell HJ 2010. PBPK modelling of inter-individual variability in the pharmacokinetics of environmental chemicals. *Toxicology* 278: 3, 256-267.
- [10] Bois FY 2010. Physiologically Based Modelling and Prediction of Drug Interactions. *Basic Clin. Pharmacol.* 106: 3, 154-161.
- [11] d'Yvoire MB, Prieto P, Blaauboer BJ, Bois FY, Boobis A, Brochet C, Coecke S, Freidig A, Gundert-Remy U, Hartung T, Jacobs MN, Lave T, Leahy DE, Lennernaes H, Loizou GD, Meek B, Pease C, Rowland M, Spendiff M, Yang J, Zeilmaker M 2007. Physiologically-based kinetic modelling (PBK modelling): Meeting the 3Rs agenda. *Atla-Altern. Lab. Anim.* 35: 6, 661-671.
- [12] Abraham MH, Weathersby PK 1994. Hydrogen-Bonding .30. Solubility of Gases and Vapors in Biological Liquids and Tissues. *J. Pharm. Sci.* 83: 10, 1450-1456.

## Literaturverzeichnis

- [13] Parham FM, Kohn MC, Matthews HB, DeRosa C, Portier CJ 1997. Using structural information to create physiologically based pharmacokinetic models for all polychlorinated biphenyls .1. Tissue:blood partition coefficients. *Toxicol. Appl. Pharm.* 144: 2, 340-347.
- [14] 2012. Bilder aus Webquellen:  
<http://www.allmystery.de/dateien/uf53569,1239648527,Mensch%20.jpg>,  
[http://www.selbsthilfe-organtransplantierter-nrw.de/files/leber\\_01.jpg](http://www.selbsthilfe-organtransplantierter-nrw.de/files/leber_01.jpg),  
<http://www.nlm.nih.gov/medlineplus/ency/images/ency/fullsize/19477.jpg>,  
[http://www.adpic.de/data/picture/detail/Ratte\\_142747.jpg](http://www.adpic.de/data/picture/detail/Ratte_142747.jpg),  
<http://reidhosp.adam.com/graphics/images/en/8674.jpg>.  
[http://www.chemgapedia.de/vsengine/media/vsc/de/ch/4/cm/carbonyl/bilder/niere\\_histonet.jpg](http://www.chemgapedia.de/vsengine/media/vsc/de/ch/4/cm/carbonyl/bilder/niere_histonet.jpg)
- [15] Basak SC, Mills D, Gute BD 2006. Prediction of tissue: air partition coefficients - theoretical vs. experimental methods. *Sar Qsar Environ. Res.* 17: 5, 515-5U3.
- [16] Poulin P, Krishnan K 1995. An Algorithm for Predicting Tissue-Blood Partition-Coefficients of Organic-Chemicals from N-Octanol-Water Partition-Coefficient Data. *J. Toxicol. Env. Health* 46: 1, 117-129.
- [17] Poulin P, Krishnan K 1996. A mechanistic algorithm for predicting blood:air partition coefficients of organic chemicals with the consideration of reversible binding in hemoglobin. *Toxicol. Appl. Pharm.* 136: 1, 131-137.
- [18] Fiserova-Bergerova V, Diaz ML 1986. Determination and Prediction of Tissue-Gas Partition-Coefficients. *Int. Arch. Occ. Env. Hea.* 58: 1, 75-87.
- [19] Luan F, Liu HT, Ma WP, Fan BT 2008. QSPR analysis of air-to-blood distribution of volatile organic compounds. *Ecotox. Environ. Safe.* 71: 3, 731-739.
- [20] Abraham MH, Ibrahim A 2006. Gas to olive oil partition coefficients: A linear free energy analysis. *J. Chem. Inf. Model.* 46: 4, 1735-1741.
- [21] Beaudry F, Coutu M, Brown NK 1999. Determination of drug-plasma protein binding using human serum albumin chromatographic column and multiple linear regression model. *Biomed. Chromatogr.* 13: 6, 401-406.
- [22] Hollosy F, Valko K, Hersey A, Nunhuck S, Keri G, Bevan C 2006. Estimation of volume of distribution in humans from high throughput HPLC-based measurements of human serum albumin binding and immobilized artificial membrane partitioning. *J. Med. Chem.* 49: 24, 6958-6971.
- [23] Valko K, Nunhuck S, Bevan C, Abraham MH, Reyncilds DP 2003. Fast gradient HPLC method to determine compounds binding to human serum albumin. Relationships with octanol/water and immobilized artificial membrane lipophilicity. *J. Pharm. Sci.* 92: 11, 2236-2248.
- [24] Kratochwil NA, Huber W, Müller F, Kansy M, Gerber PR 2002. Predicting plasma protein binding of drugs: a new approach. *Biochem. Pharmacol.* 64: 1355-1375.

## Literaturverzeichnis

- [25] Abraham MH, Ibrahim A 2006. Air to fat and blood to fat distribution of volatile organic compounds and drugs: Linear free energy analyses. *Eur. J. Med. Chem.* 41: 12, 1430-1438.
- [26] Abraham MH, Ibrahim A, Acree WE 2005. Air to blood distribution of volatile organic compounds: A linear free energy analysis. *Chem. Res. Toxicol.* 18: 5, 904-911.
- [27] Konož E, Golmohammadi H 2008. Prediction of air-to-blood partition coefficients of volatile organic compounds using genetic algorithm and artificial neural network. *Anal. Chim. Acta* 619: 2, 157-164.
- [28] Basak SC, Mills D, El-Masri HA, Mumtaz MM, Hawkins DM 2004. Predicting blood: air partition coefficients using theoretical molecular descriptors. *Environ. Toxicol. Phar.* 16: 1-2, 45-55.
- [29] Basak SC, Mills D, Hawkins DM, El-Masri HA 2003. Prediction of human blood: Air partition coefficient: A comparison of structure-based and property-based methods. *Risk Anal.* 23: 6, 1173-1184.
- [30] Basak SC, Mills D, Hawkins DM, El-Masri HA 2002. Prediction of tissue-air partition coefficients: A comparison of structure-based and property-based methods. *Sar Qsar Environ. Res.* 13: 7-8, 649-665.
- [31] Stenberg P, Norinder U, Luthman K, Artursson P 2001. Experimental and computational screening models for the prediction of intestinal drug absorption. *J. Med. Chem.* 44: 12, 1927-1937.
- [32] Polanski J, Gieleciak R 2003. The comparative molecular surface analysis (CoMSA) with modified uniformity variable elimination-PLS (UVE-PLS) method: Application to the steroids binding the aromatase enzyme. *J. Chem. Inf. Comput. Sci.* 43: 2, 656-666.
- [33] Beliveau M, Lipscomb J, Tardif R, Krishnan K 2005. Quantitative structure-property relationships for interspecies extrapolation of the inhalation pharmacokinetics of organic chemicals. *Chem. Res. Toxicol.* 18: 3, 475-485.
- [34] Beliveau M, Tardif R, Krishnan K 2003. Quantitative structure-property relationships for physiologically based pharmacokinetic modeling of volatile organic chemicals in rats. *Toxicol. Appl. Pharm.* 189: 3, 221-232.
- [35] Gargas ML, Burgess RJ, Voisard DE, Cason GH, Andersen ME 1989. Partition-Coefficients of Low-Molecular-Weight Volatile Chemicals in Various Liquids and Tissues. *Toxicol. Appl. Pharm.* 98: 1, 87-99.
- [36] Sato A, Nakajima T 1979. Partition coefficients of some aromatic hydrocarbons and ketones in water, blood and oil. *Brit. J. Ind. Med.* 36: 3, 231-234.
- [37] Hansch C, Vittoria A, Silipo C, Jow PYC 1975. Partition-Coefficients and Structure-Activity Relationship of Anesthetic-Gases. *J. Med. Chem.* 18: 6, 546-548.
- [38] Meulenberg CJW, Vijverberg HPM 2000. Empirical relations predicting human and rat tissue : air partition coefficients of volatile organic compounds. *Toxicol. Appl. Pharm.* 165: 3, 206-216.

## Literaturverzeichnis

- [39] Paterson S, Mackay D 1989. Correlation of tissue, blood, and air partition coefficients of volatile organic chemicals. *Brit. J. Ind. Med.* 46: 5, 321-328.
- [40] Poulin P, Theil FP 2000. A Priori prediction of tissue : plasma partition coefficients of drugs to facilitate the use of physiologically-based pharmacokinetic models in drug discovery. *J. Pharm. Sci.* 89: 1, 16-35.
- [41] Varnek A, Gaudin C, Marcou G, Baskin I, Pandey AK, Tetko IV 2009. Inductive Transfer of Knowledge: Application of Multi-Task Learning and Feature Net Approaches to Model Tissue-Air Partition Coefficients. *J. Chem. Inf. Model.* 49: 1, 133-144.
- [38] Zhang HB, Zhang YL 2006. Convenient nonlinear model for predicting the tissue/blood partition coefficients of seven human tissues of neutral, acidic, and basic structurally diverse compounds. *J. Med. Chem.* 49: 19, 5815-5829.
- [43] Hansch C, Maloney PP, Fujita T 1962. Correlation of Biological Activity of Phenoxyacetic Acids with Hammett Substituent Constants and Partition Coefficients. *Nature* 194: 4824, 178-&.
- [44] Wessel MD, Jurs PC, Tolan JW, Muskal SM 1998. Prediction of human intestinal absorption of drug compounds from molecular structure. *J. Chem. Inf. Comput. Sci.* 38: 4, 726-735.
- [45] Derety E, Feher M, Schmidt JM 2002. Rapid prediction of human intestinal absorption. *Quant. Struct. -Act. Rel.* 21: 5, 493-506.
- [46] Zhao YH, Le J, Abraham MH, Hersey A, Eddershaw PJ, Luscombe CN, Boutina D, Beck G, Sherborne B, Cooper I, Platts JA 2001. Evaluation of human intestinal absorption data and subsequent derivation of a quantitative structure-activity relationship (QSAR) with the Abraham descriptors. *J. Pharm. Sci.* 90: 6, 749-784.
- [47] Gunturi SB, Narayanan R, Khandelwal A 2006. In silico ADME modelling 2: Computational models to predict human serum albumin binding affinity using ant colony systems. *Bioorgan. Med. Chem.* 14: 12, 4118-4129.
- [48] Raevsky OA, Schaper KJ, Artursson P, McFarland JW 2002. A novel approach for prediction of intestinal absorption of drugs in humans based on hydrogen bond descriptors and structural similarity. *Quant. Struct. -Act. Rel.* 20: 5-6, 402-413.
- [49] Clark DE 1999. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 1. Prediction of intestinal absorption. *J. Pharm. Sci.* 88: 8, 807-814.
- [50] Hoerl AE, Kennard RW 1970. Ridge Regression - Biased Estimation for Nonorthogonal Problems. *Technometrics* 12: 1, 55-&.
- [51] Abraham MH, McGowan JC 1987. The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography. *Chromatographia* 23: 243-246.

## Literaturverzeichnis

- [52] Schüürmann G, Ebert RU, Kühne R 2006. Prediction of physicochemical properties of organic compounds from 2D molecular structure - Fragment methods vs. LFER models. *Chimia* 60: 10, 691-698.
- [53] Abraham MH 1993. Scales of Solute Hydrogen-Bonding - Their Construction and Application to Physicochemical and Biochemical Processes. *Chem. Soc. Rev.* 22: 2, 73-83.
- [54] Platts JA, Butina D, Abraham MH, Hersey A 1999. Estimation of molecular linear free energy relation descriptors using a group contribution approach. *J. Chem. Inform. Comput. Sci.* 39: 835-845.
- [55] Advanced Pharma Algorithms, Inc. 2004. ADME Boxes 2.5.7 Build 12. Toronto (ON, CAN).
- [56] Schüürmann G, Ebert R-U, Kühne R 2006. Prediction of physicochemical properties of organic compounds from 2D molecular structure - Fragment methods vs. LFER models. *Chimia*. 60: 691-698.
- [57] Sprunger LM, Gibbs J, Acree WE, Abraham MH 2008. Correlation of human and animal air-to-blood partition coefficients with a single linear free energy relationship model. *Qsar Comb. Sci.* 27: 9, 1130-1139.
- [58] Abraham MH, Martins F 2004. Human skin permeation and partition: General linear free-energy relationship analyses. *J. Pharm. Sci.* 93: 1508-1523.
- [59] Kühne R, Ebert RU, Schüürmann G 2005. Prediction of the temperature dependency of Henry's law constant from chemical structure. *Environ. Sci. Technol.* 39: 17, 6705-6711.
- [60] Schüürmann G., Ebert R.-U., Nendza M., Dearden J.C., Paschke A., Kühne R. 2007. Prediction of fate-related compound properties. In: C.J.van Leeuwen TGV (ed) Risk assessment of chemicals: an introduction 9. Springer, Dordrecht, pp. 375-455.
- [61] Miller DG 1964. Estimating Vapor Pressures-Comparison of Equations. *Ind. Eng. Chem.* 56: 3, 46-&.
- [62] Sepassi K, Myrdal PB, Yalkowsky SH 2006. Estimating pure-component vapor pressures of complex organic molecules: Part II. *Ind. Eng. Chem. Res.* 45: 25, 8744-8747.
- [63] Grain CF 1990. 14. Vapor pressure. In: Lyman WJ, Reehl WF, Rosenblatt DH (eds) Handbook of Chemical Property Estimation Methods. American Chemical Society, Washington (DC, USA), pp. 14-1-14-20.
- [64] Thomson G.W. 1959. Techniques of Organic Chemistry. 1, Part 1. Interscience, New York, p. 473.
- [65] Lyman WJ, Reehl WF, Rosenblatt DH 1982. Handbook of chemical property estimation methods : environmental behavior of organic compounds. McGraw-Hill, New York.
- [66] Fishtine SH 1963. Reliable Latent Heats of Vaporization Lv. *Ind. Eng. Chem.* 55: 6, 47-60.

## Literaturverzeichnis

- [67] Sage ML, Sage GW 2000. Vapor pressure. In: Boethling RS, Mackay D (eds) Handbook of Property Estimation Methods for Chemicals: Environmental and Health Sciences Boca Raton (FL, USA), pp. 53-65.
- [68] Mackay D, Bobra A, Chan DW, Shiu WY 1982. Vapor-Pressure Correlations for Low-Volatility Environmental Chemicals. *Environ. Sci. Technol.* 16: 10, 645-649.
- [69] Trouton F 1884. On Molecular Latent Heat.
- [70] Meylan WM, Howard PH 1995. Atom Fragment Contribution Method for Estimating Octanol-Water Partition-Coefficients. *J. Pharm. Sci.* 84: 1, 83-92.
- [71] Hall LH 1990. New Electrotopological Index for Atoms in Molecules - Use in Qsar Analysis. *Abstr. Pap. Am. Chem. S.* 199: 1-COMP.
- [72] Kier LB, Hall LH 1990. An Electrotopological-State Index for Atoms in Molecules. *Pharm. Res.* 7: 8, 801-807.
- [73] Hall LH, Mohny B, Kier LB 1991. The Electrotopological State - An Atom Index for Qsar. *Quant. Struct. -Act. Rel.* 10: 1, 43-51.
- [74] Hall LH, Kier LB 1995. Electrotopological State Indices for Atom Types: A Novel Combination of Electronic, Topological, and Valence State Information. *J. Chem. Inf. Comput. Sci.* 35: 6, 1039-1045.
- [75] Kamgang E, Peyret T, Krishnan K 2008. An integrated QSPR-PBPK modelling approach for in vitro-in vivo extrapolation of pharmacokinetics in rats. *Sar Qsar Environ. Res.* 19: 7-8, 669-680.
- [76] Schüürmann G, Ebert RU, Kühne R 2006. Prediction of the sorption of organic compounds into soil organic matter from molecular structure. *Environ. Sci. Technol.* 40: 22, 7005-7011.
- [77] Gearhart JM, Mahle DA, Greene RJ, Seckel CS, Flemming CD, Fisher JW, Clewell HJ 1993. Variability of Physiologically-Based Pharmacokinetic (Pbpb) Model Parameters and Their Effects on Pbpb Model Predictions in A Risk Assessment for Perchloroethylene (Pce). *Toxicol. Lett.* 68: 1-2, 131-144.
- [78] 1989. Handbuch Umweltchemikalien. Ecomed, Landsberg/Lech (D).
- [79] Industrieverband Agrar (ed) 1990. Wirkstoffe in Pflanzenschutz- und Schädlingsbekämpfungsmitteln. Physikalisch-chemische und toxikologische Daten. BLV Verl., München (D).
- [80] Worthing CR, Hance RJ (eds) 1991. The Pesticide Manual. The British Crop Protection Council, Farnham (UK).
- [81] Richardson ML, Gangolli S (eds) 1992. The Dictionary of Substances and their Effects. Royal Society of Chemistry, Cambridge (UK).

## Literaturverzeichnis

- [82] Lide DR (ed) 1992. Handbook of Chemistry and Physics. CRC Press, Boca Raton (FL, USA).
- [83] Tomlin C (ed) 1994. The Pesticide Manual. Incorporating the Agrochemicals Handbook. A world compendium. Royal Society of Chemistry / British Crop Protection Council, Cambridge (UK).
- [84] Hester RE, Harrison RM (eds) 1996. Chlorinated Organic Micropollutants. Issues in Environmental Science and Technology 6. Royal Society of Chemistry, Cambridge (UK).
- [85] Umweltbundesamt (UBA), Institut für Wasser-, Boden- und Lufthygiene (ed) 1996ff. Dokumentation wassergefährdender Stoffe. Datenblattsammlung zur Verwaltungsvorschrift wassergefährdender Stoffe (VwVwS) nach § 19g WHG. S. Hirzel Wissenschaftliche Verlagsgesellschaft, Stuttgart.
- [86] Tomlin CDS (ed) 1997. The Pesticide Manual. British Crop Protection Council, Farnham (UK).
- [87] British Crop Protection Council, Tomlin CDS (eds) 2000. The Pesticide Manual. A World Compendium. British Crop Protection Council, Farnham (UK).
- [88] Fachinformationszentrum Chemie (FIZ) Berlin Infotherm - thermophysical properties database: <http://www.chemistry.de/infotherm/servlet/infothermSearch>.
- [89] National Toxicology Program Chemical Health & Safety Data: [http://ntp-server.niehs.nih.gov/Main\\_Pages/Chem-HS.html](http://ntp-server.niehs.nih.gov/Main_Pages/Chem-HS.html) link broken.
- [90] U.S. Environmental Protection Agency 2008. MPBPVP 1.43.
- [91] Acros Organics <http://www.acros.be>.
- [92] Bamford HA, Baker JE, Poster DL 1998. Review of methods and measurements of selected hydrophobic organic contaminant aqueous solubilities, vapor pressures, and air-water partition coefficients. NIST Special Publication 928, USA.
- [93] Baughman GL, Perenich TA 1988. Fate of Dyes in Aquatic Systems .1. Solubility and Partitioning of Some Hydrophobic Dyes and Related-Compounds. *Environ. Toxicol. Chem.* 7: 3, 183-199.
- [94] Beauman JA, Howard PH . Physprop database. Syracuse Research Corporation, Syracuse (NY, USA). accessed at 29/07/2010.
- [95] Covarrubias-Cervantes M, Mokbel H, Champion D, Jose J, Voilley A 2004. Saturated vapour pressure of aroma compounds at various temperatures. *Food Chem.* 85: 2, 221-229.
- [96] European Chemicals Bureau International Uniform Chemical Information Database (IUCLID): <http://ecb.jrc.ec.europa.eu/iuclid-datasheet/75445.pdf>.
- [97] Goll ES, Jurs PC 1999. Prediction of vapor pressures of hydrocarbons and halohydrocarbons from molecular structure with a computational neural network model. *J. Chem. Inf. Comput. Sci.* 39: 6, 1081-1089.

## Literaturverzeichnis

- [98] Hornsby AG, Wauchope RD, Herner AE 1996. Pesticide properties in the environment. Springer, Berlin / New York (D / NY, USA).
- [99] Hovorka S, Dohnal V, Roux AH, Roux-Desgranges G 2002. Determination of temperature dependence of limiting activity coefficients for a group of moderately hydrophobic organic solutes in water. *Fluid Phase Equilibr.* 201: 1, 135-164.
- [100] Li J, Perdue EM, Pavlostathis SG, Araujo R 1998. Physicochemical properties of selected monoterpenes. *Environ. Int.* 24: 353-358.
- [101] Miller ME, Stuart JD 2000. Measurement of aqueous Henry's law constants for oxygenates and aromatics found in gasolines by the static headspace method. *Anal. Chem.* 72: 3, 622-625.
- [102] N'Guimbi J, Berro C, Mokbel I, Rauzy E, Jose J 1999. Experimental vapour pressures of 13 secondary and tertiary alcohols--correlation and prediction by a group contribution method. *Fluid Phase Equilibr.* 162: 1-2, 143-158.
- [103] Nass K, Lenoir D, Kettrup A 1995. Calculation of the Thermodynamic Properties of Polycyclic Aromatic-Hydrocarbons by An Incremental Procedure. *Angew. Chem. Int. Edit.* 34: 16, 1735-1736.
- [104] Oonk HAJ, van der Linde PR, Huinink J, Blok JG 1998. Representation and assessment of vapour pressure data; a novel approach applied to crystalline 1-bromo-4-chlorobenzene, 1-chloro-4-iodobenzene, and 1-bromo-4-iodobenzene. *J. Chem. Thermodyn.* 30: 897-907.
- [105] Sagebiel JC, Seiber JN 1993. Studies on the Occurrence and Distribution of Wood Smoke Marker Compounds in Foggy Atmospheres. *Environ. Toxicol. Chem.* 12: 5, 813-822.
- [106] Schmuckler ME, Barefoot AC, Kleier DA, Cobranchi DP 2000. Vapor pressures of sulfonyleurea herbicides. *Pest Manag. Sci.* 56: 6, 521-532.
- [107] Shiu WY, Ma KC 2000. Temperature dependence of physical-chemical properties of selected chemicals of environmental interest. I. Mononuclear and polynuclear aromatic hydrocarbons. *J. Phys. Chem. Ref. Data* 29: 41-130.
- [108] Sorbe G 1990. Sicherheitstechnische Kenndaten. Ecomed, Landsberg/Lech (D).
- [109] Stales CA, Peterson DR, Parkerton TF, Adams WJ 1997. The environmental fate of phthalate esters: A literature review. *Chemosphere* 35: 4, 667-749.
- [110] Tsuzuki M 2000. Thermodynamic estimation of vapor pressure for organophosphorus pesticides. *Environ. Toxicol. Chem.* 19: 1717-1726.
- [111] Verevkin SP 2002. Improved Benson increments for the estimation of standard enthalpies of formation and enthalpies of vaporization of alkyl ethers, acetals, ketals, and ortho esters. *J. Chem. Eng. Data* 47: 5, 1071-1097.
- [112] The University of Akron 2011. <http://www.uakron.edu>.

## Literaturverzeichnis

- [113] Abraham MH, Whiting GS, Doherty RM, Shuely WJ 1991. Hydrogen-Bonding .16. A New Solute Solvation Parameter,  $\Pi$ -2(H), from Gas-Chromatographic Data. *J. Chromatogr.* 587: 2, 213-228.
- [114] Advanced Pharma Algorithms, Inc. 2004. ADME Boxes 2.5.7 Build 12. Toronto (ON, CAN).
- [115] Schüürmann G, Kühne R, Kleint F, Ebert R-U, Rothenbacher C, Herth P 1997. A software system for automatic chemical property estimation from molecular structure. In: Chen F, Schüürmann G (eds) Quantitative Structure-Activity Relationships in Environmental Sciences - VII SETAC Press, Pensacola (FL, USA), pp. 93-114.
- [116] Schüürmann G, Ebert R-U, Nendza M, Dearden JC, Paschke A, Kühne R 2007. 9;9. Predicting fate-related physicochemical properties. In: van Leeuwen K, Vermeire T (eds) Risk Assessment of Chemicals. An Introduction. Springer Science, Dordrecht (NL), pp. 375-426.
- [117] U.S. Environmental Protection Agency 2008. HENRYWIN 3.20.
- [118] U.S. Environmental Protection Agency 2008. EPI Suite 4.0.
- [119] Abraham MH, Andonianhaftvan J, Whiting GS, Leo A, Taft RS 1994. Hydrogen-Bonding .34. the Factors That Influence the Solubility of Gases and Vapors in Water at 298-K, and A New Method for Its Determination. *J. Chem. Soc. Perk. T. 2*: 8, 1777-1791.
- [120] Abraham MH, Chadha HS, Whiting GS, Mitchell RC 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: 8, 1085-1100.
- [121] U.S. Environmental Protection Agency 2008. KOWWIN 1.67a.
- [122] U.S. Environmental Protection Agency 2008. KOAWIN v.1.10.
- [123] Platts JA, Abraham MH 2000. Partition of volatile organic compounds from air and from water into plant cuticular matrix: An LFER analysis. *Environ. Sci. Technol.* 34: 2, 318-323.
- [124] Abraham MH, Le J, Acree WE, Carr PW, Dallas AJ 2001. The solubility of gases and vapours in dry octan-1-ol at 298 K. *Chemosphere* 44: 4, 855-863.
- [125] Dabitz-Gutsche N 2010. Experimentelle Bestimmung der Verteilungskoeffizienten ausgewählter organischer Umweltchemikalien zwischen Membranlipiden und Wasser.
- [126] Shvedov D, Tremaine PR 2000. The solubility of sodium sulfate and the reduction of aqueous sulfate by magnetite under near-critical conditions. *J. Solution Chem.* 29: 10, 889-904.
- [127] Meylan WM, Howard PH, Boethling RS 1996. Improved method for estimating water solubility from octanol water partition coefficient. *Environ. Toxicol. Chem.* 15: 2, 100-106.

## Literaturverzeichnis

- [128] Hou TJ, Xia K, Zhang W, Xu XJ 2004. ADME evaluation in drug discovery. 4. Prediction of aqueous solubility based on atom contribution approach. *J. Chem. Inf. Comput. Sci.* 44: 1, 266-275.
- [129] Tetko IV, Tanchuk VY, Kasheva TN, Villa AEP 2001. Estimation of aqueous solubility of chemical compounds using E-state indices. *J. Chem. Inf. Comput. Sci.* 41: 6, 1488-1493.
- [130] Marrero J, Gani R 2002. Group-contribution-based estimation of octanol/water partition coefficient and aqueous solubility. *Ind. Eng. Chem. Res.* 41: 25, 6623-6633.
- [131] Kühne R, Ebert RU, Kleint F, Schmidt G, Schüürmann G 1995. Group-Contribution Methods to Estimate Water Solubility of Organic-Chemicals. *Chemosphere* 30: 11, 2061-2077.
- [132] Klopman G, Wang S, Balthasar DM 1992. Estimation of Aqueous Solubility of Organic-Molecules by the Group Contribution Approach - Application to the Study of Biodegradation. *J. Chem. Inf. Comput. Sci.* 32: 5, 474-482.
- [133] Huuskonen J 2001. Estimation of water solubility from atom-type electrotopological state indices. *Environ. Toxicol. Chem.* 20: 3, 491-497.
- [134] Abraham MH, Le J 1999. The correlation and prediction of the solubility of compounds in water using an amended solvation energy relationship. *J. Pharm. Sci.* 88: 9, 868-880.
- [135] Kühne R, Ebert R-U, Schüürmann G 2006. Model selection based on structural similarity – Method description and application to water solubility prediction. *J. Chem. Inf. Model.* 46: 636-641.
- [136] Microsoft Corporation 2007. Microsoft Office Excel 2007.
- [137] StatSoft Inc. 2008. STATISTICA 8.0.
- [138] OriginLab Corporation 2009. OriginPro 8G.0.63.988 SR6 8.0988.
- [139] Kurt Hornik 2009. R FAQ Version 2.9.2009-04-15.
- [140] Torsten Thalheim 2013. cvq2: Calculate the predictive squared correlation coefficient. R package version 1.1.0.
- [141] Kühne R 2011. Current version: <http://www.ufz.de/index.php?en=7160>.
- [142] Schüürmann G, Ebert RU, Chen JW, Wang B, Kühne R 2008. External Validation and Prediction Employing the Predictive Squared Correlation Coefficient - Test Set Activity Mean vs Training Set Activity Mean. *J. Chem. Inf. Model.* 48: 11, 2140-2145.
- [143] Platts JA, Butina D, Abraham MH, Hersey A 1999. Estimation of molecular linear free energy relation descriptors using a group contribution approach. *J. Chem. Inf. Comput. Sci.* 39: 5, 835-845.

## Literaturverzeichnis

[144] Gargas ML, Andersen ME, Teo SKO, Batra R, Fennell TR, Kedderis GL 1995. A Physiologically-Based Dosimetry Description of Acrylonitrile and Cyanoethylene Oxide in the Rat. *Toxicol. Appl. Pharm.* 134: 2, 185-194.

[145] 2013. SIGMA-ALDRICH. Sicherheitsdatenblatt Tetracyanoethylene oxide.

[146] Joshi G, Tremblay RT, Martin SA, Fisher JW 2010. Partition coefficients for nonane and its isomers in the rat. *Toxicology Mechanisms and Methods* 20: 9, 594-599.

[147] Ramsey JC, Andersen ME 1984. A Physiologically Based Description of the Inhalation Pharmacokinetics of Styrene in Rats and Humans. *Toxicol. Appl. Pharm.* 73: 1, 159-175.

[148] Smith AQ, Campbell JL, Keys DA, Fisher JW 2005. Rat tissue and blood partition coefficients for n-alkanes (C-8 to C-12). *International Journal of Toxicology* 24: 1, 35-41.

## Anhang

### 8 Anhang

Im Anhang werden alle Daten, die in dieser Arbeit verwendet werden, gezeigt.

#### 8.1 Blut/Luft-Verteilungskoeffizient $K_{ba}$

Für den Blut/Luft-Verteilungskoeffizienten werden die Stoffe aus dem Datensatz mit ihren experimentellen Daten gezeigt, alle Daten von Verteilungskoeffizienten, die zur Modellierung verwendet wurden und auch die Beiträge der Fragmente und E-States angeführt.

##### 8.1.1 $K_{ba}$ -Daten und -Werte

**Tabelle 8.1: Verwendete Substanzen für die  $K_{ba}$ -Modellierung mit Name, SMILES-Code, ChemProp ID und CAS Nr. und experimentelle aus der Literatur gesammelte Werte für Mensch und Ratte und Quellen.**

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
1	Ethan	CC	695	74-84-0	-1,070	-0,970	-1,020	[26]
2	<i>n</i> -Butan	CCCC	697	106-97-8	-	-0,530	-0,530	[26]
3	Pentan	CCCCC	1	109-66-0	-0,290	-	-0,290	[26]
4	Hexan	CCCCCC	3	110-54-3	0,070	0,350	0,210	[26]
5	Heptan	CCCCCCC	9	142-82-5	0,420	0,580	0,500	[26]
6	Oktan	CCCCCCCC	12	111-65-9	0,610	0,496	0,553	[26,57]
7	Nonan	CCCCCCCCC	971	111-84-2	1,700	0,630	1,165	[26]
8	Dekan	CCCCCCCCC	1011	124-18-5	1,920	1,020	1,470	[26]
9	Undekan	CCCCCCCCCCC	1102	1120-21-4	-	1,310	1,310	[57]
10	Dodekan	CCCCCCCCCCCC	1112	112-40-3	-	1,390	1,390	[57]
11	2-Methylpentan	CC(CCC)C	4	107-83-5	-0,390	-	-0,390	[26]
12	3-Methylpentan	CCC(CC)C	5	96-14-0	-0,370	-	-0,370	[26]
13	2,3-Dimethylbutan	CC(C(C)C)C	916	79-29-8	0,780	-	0,780	[26]
14	3-Methylhexan	CCC(CCC)C	921	589-34-4	0,110	-	0,110	[26]
15	2-Methylheptan	CC(C)CCCC	944	592-27-8	-	0,490	0,490	[26]
16	2,3,4-Trimethylpentan	CC(C(C(C)C)C)C	936	565-75-3	-	0,570	0,570	[26]

## Anhang

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
17	2-Methyloktan	C(C)(C)CCCCC	991	3221-61-2	-	0,520	0,520	[26]
18	2-Methylnonan	C(C)(C)CCCCCC	1012	871-83-0	-	0,760	0,760	[26]
19	2,2-Dimethylbutan	CCC(C)(C)C	6	75-83-2	-0,590	-	-0,590	[26]
20	2,2,4-Trimethylpentan	CC(CC(C)(C)C)C	13	540-84-1	0,200	0,250	0,225	[26]
21	Cyclopropan	C1CC1	912	75-19-4	-0,290	-0,120	-0,205	[26]
22	Cyclopentan	C1CCCC1	2	287-92-3	-	0,240	0,240	[26]
23	Cyclohexan	C1CCCCC1	7	110-82-7	0,190	0,140	0,165	[26]
24	Cycloheptan	C1CCCCC1	925	291-64-5	-	0,720	0,720	[26]
25	Methylcyclopentan	CC1CCCC1	8	96-37-7	-0,070	-	-0,070	[26]
26	Methylcyclohexan	CC1CCCCC1	11	108-87-2	0,610	0,790	0,700	[26]
27	1,2-Dimethylcyclohexan	C1C(C)C(C)CCC1	14	583-57-3	-	0,910	0,910	[26]
28	1,2,4-Trimethylcyclohexan	CC1CCC(C(C1)C)C	11587	2234-75-5	-	0,870	0,870	[26]
29	JP-10	C1CC2CC1C3C2CCC3	10708	2825-82-3	1,720	1,790	1,755	[26]
30	<i>tert</i> -Butylcyclohexan	C(C)(C)(C)C1CCCCC1	1101	3178-22-1	-	1,160	1,160	[26]
31	Ethen	C=C	700	74-85-1	-0,750	-0,310	-0,530	[26]
32	Propen	CC=C	701	115-07-1	-0,360	-0,060	-0,210	[26]
33	1-Okten	CCCCCCC=C	23	111-66-0	-	1,070	1,070	[26]
34	1-Nonen	CCCCCCCC=C	24	124-11-8	-	1,180	1,180	[26]
35	1-Deken	CCCCCCCCC=C	1396	872-05-9	-	1,210	1,210	[26]
36	1,3-Butadien	C(C=C)=C	705	106-99-0	0,090	-	0,090	[26]
37	2-Methyl-1,3-butadien	C(C(=C)C)=C	27	78-79-5	-0,120	0,320	0,100	[26]
38	Acetylen	C#C	706	74-86-2	-0,060	-	-0,060	[26]
39	Benzol	c1ccccc1	33	71-43-2	0,870	1,220	1,045	[26]
40	Toluol	Cc1ccccc1	34	108-88-3	1,120	1,160	1,140	[26]
41	Ethylbenzen	CCc1ccccc1	36	100-41-4	1,450	1,480	1,465	[26]
42	1,3-Dimethylbenzen ( <i>m</i> -Xylol)	Cc1cc(ccc1)C	38	108-38-3	1,510	1,660	1,585	[26]
43	<i>o</i> -Xylol	Cc1c(cccc1)C	35	95-47-6	1,530	1,300	1,415	[26]
44	<i>p</i> -Xylol	Cc1ccc(cc1)C	37	106-42-3	1,600	1,620	1,610	[26]
45	1,2,4-Trimethylbenzen	Cc1c(cc(cc1)C)C	40	95-63-6	1,770	1,160	1,465	[26]
46	Propylbenzen	CCCc1ccccc1	39	103-65-1	1,670	-	1,670	[26]
47	1,2,3-Trimethylbenzen	Cc1c(c(C)ccc1)C	42	526-73-8	1,820	-	1,820	[26]
48	1,3,5-Trimethylbenzen	Cc1cc(cc(c1)C)C	41	108-67-8	1,640	-	1,640	[26]

## Anhang

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
49	Isopropylbenzen	CC(c1ccccc1)C	45	98-82-8	1,570	-	1,570	[26]
50	<i>tert</i> -Butylbenzen	CC(c1ccccc1)(C)C	51	98-06-6	-	1,240	1,240	[26]
51	Styrol	C=Cc1ccccc1	58	100-42-5	1,730	1,600	1,665	[26]
52	<i>m</i> -Methylstyrol	c1c(C=C)cc(cc1)C	1726	100-80-1	-	2,280	2,280	[26]
53	<i>p</i> -Methylstyrol	c1c(ccc(c1)C)C=C	1729	622-97-9	-	2,370	2,370	[26]
54	Allylbenzen	C=CCc1ccccc1	1727	300-57-2	1,710	1,680	1,695	[26,75]
55	Fluorethan	FCC	1870	353-36-6	0,090	-	0,090	[26]
56	1-Fluorpropan	FCCC	6878	460-13-9	0,020	-	0,020	[26]
57	2-Fluorpropan	CC(F)C	9077	420-26-8	0,060	-	0,060	[26]
58	Difluormethan	FCF	1852	75-10-5	-	0,200	0,200	[26]
59	1,1-Difluorethan	FC(F)C	721	75-37-6	0,420	-	0,420	[26]
60	1,3-Difluorpropan	FCCCF	6879	462-39-5	0,969	-	0,969	[38]
61	Tetrafluormethan	FC(F)(F)F	718	75-73-0	-1,100	-	-1,100	[26]
62	1,1,2,2-Tetrafluorethan	C(F)(F)C(F)F	9098	359-35-3	-0,120	-	-0,120	[26]
63	1,1,1,2-Tetrafluorethan	FCC(F)(F)F	6630	811-97-2	-0,250	-	-0,250	[26]
64	1,1,2,2,3-Pentafluorpropan	C(F)(F)C(F)(F)CF	9089	679-86-7	-0,480	-	-0,480	[26]
65	1,1,2,4,4-Pentafluorbutan	FC(C(CC(F)F)F)F	12821		0,870	-	0,870	[26]
66	1,1,2,2,3,3,4,4-Oktafluorbutan	FC(C(C(C(F)F)(F)F)(F)F)F	12498	377-36-6	-0,360	-	-0,360	[26]
67	1,1,1,2,3,4,4,4-Oktafluorbutan	FC(C(C(C(F)(F)F)F)F)(F)F	12172	75995-72-1	-0,590	-	-0,590	[26]
68	1,1,1,2,2,3,3,4,4-Nonafluorbutan	FC(C(C(C(F)F)(F)F)(F)F)(F)F	12495	375-17-7	-1,520	-	-1,520	[26]
69	Chlormethan	CIC	710	74-87-3	0,230	0,390	0,310	[26]
70	Chlorethan	CICC	720	75-00-3	0,360	0,610	0,485	[26]
71	1-Chlorpropan	CICCC	123	540-54-5	0,460	0,720	0,590	[26]
72	2-Chlorpropan	CC(C)Cl	124	75-29-6	0,140	0,490	0,315	[26]
73	1-Chlorbutan	CICCCC	136	109-69-3	0,630	-	0,630	[26]
74	1-Chlorpentan	CICCCCC	144	543-59-9	0,870	-	0,870	[26]
75	Dichlormethan	CICCl	100	75-09-2	0,950	1,290	1,120	[26]
76	1,2-Dichlorethan	CICCCl	112	107-06-2	1,300	1,480	1,390	[26]
77	1,1-Dichlorethan	CIC(C)Cl	111	75-34-3	0,700	1,050	0,875	[26]
78	1,2-Dichlorpropan	CICCC(C)Cl	130	78-87-5	1,010	1,270	1,140	[26]
79	Trichlormethan	CIC(Cl)Cl	103	67-66-3	0,980	1,320	1,150	[26]
80	1,1,2-Trichlorethan	CIC(CCl)Cl	116	79-00-5	1,580	1,760	1,670	[26]

## Anhang

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
81	1,1,1-Trichlorethan	CIC(C)(Cl)Cl	115	71-55-6	0,500	0,760	0,630	[26]
82	1,2,3-Trichlorpropan	CICCCl	135	96-18-4	2,010	-	2,010	[26]
83	Tetrachlormethan	CIC(Cl)(Cl)Cl	107	56-23-5	0,570	0,660	0,615	[26]
84	1,1,1,2-Tetrachlorethan	CIC(CCl)(Cl)Cl	118	630-20-6	1,480	1,620	1,550	[26]
85	1,1,2,2-Tetrachlorethan	CIC(C(Cl)Cl)Cl	117	79-34-5	2,110	2,150	2,130	[26]
86	Pentachlorethan	CIC(C(Cl)Cl)(Cl)Cl	120	76-01-7	-	2,020	2,020	[26]
87	Hexachlorethan	CIC(C(Cl)(Cl)Cl)(Cl)Cl	122	67-72-1	1,720	1,800	1,760	[26]
88	1-Brompropan	BrCCC	125	106-94-5	0,850	1,090	0,970	[26]
89	2-Brompropan	CC(C)Br	126	75-26-3	0,410	0,860	0,635	[26]
90	Dibrommethan	BrCBr	102	74-95-3	-	1,870	1,870	[26]
91	1,2-Dibromethan	BrCCBr	114	106-93-4	-	2,080	2,080	[26]
92	Tribrommethan	BrC(Br)Br	106	75-25-2	2,020	2,270	2,145	[26]
93	Iodethan	ICC	110	75-03-6	0,830	-	0,830	[26]
94	Fluorchlormethan	FCCl	712	593-70-4	-	0,710	0,710	[26]
95	1-Chlor-2,2,2-trifluorethan	FC(F)(F)CCl	878	75-88-7	0,180	0,100	0,140	[26]
96	1-Chlor-1,2,2,2-tetrafluorethan	CIC(F)C(F)(F)F	6635	2837-89-0	-	0,181	0,181	[57]
97	1,1-Dichlor-1-fluorethan	C(F)(Cl)(Cl)C	1884	1717-00-6	-	0,320	0,320	[26]
98	Difluordichlormethan	FC(F)(Cl)Cl	715	75-71-8	-0,820	-	-0,820	[26]
99	1,1,1-Trifluor-2,2-dichlorethan	C(C(Cl)Cl)(F)(F)F	6686	306-83-2	-	0,610	0,610	[26]
100	1,2-Dichlortetrafluorethan	FC(C(F)(F)Cl)(Cl)F	722	76-14-2	-0,820	-	-0,820	[26]
101	Fluortrichlormethan	FC(Cl)(Cl)Cl	716	75-69-4	-0,060	-	-0,060	[26]
102	Tefluran	FC(C(F)Br)(F)F	6687	124-72-1	-0,220	-	-0,220	[26]
103	Halopropan	C(C(C(F)F)(F)F)Br	12570	679-84-5	0,750	-	0,750	[26]
104	Bromchlormethan	CICBr	101	74-97-5	0,790	1,620	1,205	[26]
105	1-Brom-2-chlorethan	CICCCl	113	107-04-0	1,470	1,720	1,595	[26]
106	Halothan	FC(C(Cl)Br)(F)F	1902	151-67-7	0,400	0,730	0,565	[26]
107	Bromdichlormethan	BrC(Cl)Cl	104	75-27-4	1,420	1,560	1,490	[26]
108	Chlordibrommethan	CIC(Br)Br	105	124-48-1	1,710	2,040	1,875	[26]
109	1,1-Difluorethen	FC(=C)F	724	75-38-7	-	-0,740	-0,740	[26]
110	Vinylchlorid	CIC=C	879	75-01-4	0,060	0,270	0,165	[26]
111	Allylchlorid	CICCC=C	160	107-05-1	-	1,240	1,240	[26]
112	cis-1,2-Dichlorethen	CIC=CCl	4438	156-59-2	0,980	1,330	1,155	[26]

## Anhang

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
113	<i>trans</i> -1,2-Dichlorethen	ClC=CCl	155	156-60-5	0,770	0,980	0,875	[26]
114	1,1-Dichlorethen	ClC(=C)Cl	154	75-35-4	-	0,700	0,700	[26]
115	Trichlorethen	ClC(=CCl)Cl	158	79-01-6	0,940	1,330	1,135	[26]
116	Tetrachlorethen	ClC(=C(Cl)Cl)Cl	159	127-18-4	1,090	1,280	1,185	[26]
117	Vinylbromid	C(Br)=C	723	593-60-2	0,360	0,610	0,485	[26]
118	1-Chlor-2,2-difluorethen	FC(=CCl)F	2108	359-10-4	0,060	-	0,060	[26]
119	$\beta$ -Chloropren	ClC(C=C)=C	2187	126-99-8	0,653	0,881	0,767	[57]
120	Fluorbenzen	Fc1ccccc1	167	462-06-6	-	1,060	1,060	[26]
121	1,2-Difluorbenzen	c1(F)c(cccc1)F	2206	367-11-3	-	0,960	0,960	[26]
122	1,4-Difluorbenzen	c1c(F)ccc(F)c1	2208	540-36-3	-	0,870	0,870	[26]
123	1,2,4-Trifluorbenzen	c1(F)c(F)cc(F)cc1	2212	367-23-7	-	0,760	0,760	[26]
124	1,3,5-Trifluorbenzen	c1(F)cc(F)cc(F)c1	2211	372-38-3	-	0,490	0,490	[26]
125	Pentafluorbenzen	c1(F)c(F)c(F)cc(F)c1F	2223	363-72-4	-	0,510	0,510	[26]
126	Methylpentafluorbenzen	c1(F)c(F)c(F)c(c(F)c1F)C	2276	771-56-2	-	0,730	0,730	[26]
127	Hexafluorbenzen	Fc1c(c(c(c(c1F)F)F)F)F	2228	392-56-3	-	0,390	0,390	[26]
128	Chlorbenzen	Clc1ccccc1	168	108-90-7	1,480	1,770	1,480	[26]
129	Chlorbenzen	Clc1ccccc1	168	108-90-7	1,480	1,770	1,625	[26]
130	1,2-Dichlorbenzen	Clc1c(cccc1)Cl	172	95-50-1	2,630	-	2,630	[26]
131	1,3-Dichlorbenzen	Clc1cc(ccc1)Cl	173	541-73-1	2,300	-	2,300	[26]
132	1,4-Dichlorbenzen	Clc1ccc(cc1)Cl	171	106-46-7	2,540	-	2,540	[12]
133	Brombenzen	Brc1ccccc1	169	108-86-1	-	1,720	1,720	[26]
134	4-Chlorbenzotrifluorid	C(F)(F)(F)c1ccc(cc1)Cl	2268	98-56-6	1,220	1,640	1,430	[26]
135	Methanol	CO	737	67-56-1	3,290	3,520	3,405	[26]
136	Ethanol	CCO	738	64-17-5	3,170	3,370	3,270	[26]
137	1-Propanol	CCCO	741	71-23-8	2,990	3,130	3,060	[26]
138	1-Butanol	CCCCO	262	71-36-3	2,970	3,190	3,080	[26]
139	1-Pentanol	CCCCCO	265	71-41-0	2,730	2,920	2,825	[26]
140	1-Hexanol	CCCCCCO	273	111-27-3	-	3,210	3,210	[26]
141	2-Methyl-1-propanol	CC(CO)C	264	78-83-1	2,890	2,940	2,915	[26]
142	3-Methyl-1-butanol	CC(CCO)C	270	123-51-3	2,580	2,920	2,750	[26]
143	2-Propanol	CC(O)C	740	67-63-0	2,920	3,110	3,015	[26]
144	2-Methyl-2-propanol	CC(C)(C)O	744	75-65-0	2,660	2,700	2,680	[26]

## Anhang

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
145	<i>tert</i> -Pentylalkohol	CCC(C)(C)O	268	75-85-4	-	2,590	2,590	[26]
146	Dimethylether	COC	763	115-10-6	1,160	-	1,160	[26]
147	Diethylether	CCOCC	422	60-29-7	1,090	1,120	1,105	[26]
148	Methyl- <i>tert</i> -butylether	COC(C)(C)C	427	1634-04-4	1,250	1,110	1,180	[26]
149	<i>tert</i> -Amylmethylether	C(OC)(CC)(C)C	6636	994-05-8	1,250	1,190	1,220	[26]
150	Ethyl- <i>tert</i> -butylether	C(OCC)(C)(C)C	6697	637-92-3	1,070	-	1,070	[26]
151	Ethyl- <i>tert</i> -pentylether	CC(OCC)(CC)C	9979	919-94-8	1,250	-	1,250	[26]
152	Divinylether	C(OC=C)=C	2603	109-93-3	0,410	-	0,410	[26]
153	Ethylenoxid	C1OC1	767	75-21-8	1,790	1,810	1,800	[26]
154	Propylenoxid	CC1CO1	442	75-56-9	1,812	1,778	1,795	[57]
155	1,2-Epoxy-3-buten	C1C(C=C)O1	5490	930-22-3	-	1,970	1,970	[26]
156	1,4-Dioxan	C1OCCOC1	765	123-91-1	-	3,267	3,267	[57]
157	Furan	c1ccco1	446	110-00-9	-	0,820	0,820	[26]
158	Propanon	CC(=O)C	754	67-64-1	2,350	2,370	2,360	[26]
159	Butan-2-on	CCC(=O)C	343	78-93-3	2,190	2,280	2,235	[26]
160	2-Pentanon	CCCC(=O)C	344	107-87-9	2,180	2,100	2,140	[26]
161	3-Pentanon	C(C)C(=O)CC	2513	96-22-0	2,210	-	2,210	[26]
162	2-Hexanon	CCCCC(=O)C	345	591-78-6	2,100	-	2,100	[26]
163	2-Heptanon	CCCCCC(=O)C	348	110-43-0	2,300	2,350	2,325	[26]
164	4-Methyl-2-pentanon	CC(CC(=O)C)C	346	108-10-1	2,010	1,900	1,955	[26]
165	3-Methylpentan-2-on	CCC(C(=O)C)C	6710	565-61-7	2,230	-	2,230	[26]
166	Cylohexanon	C1CCCCC(=O)C1	347	108-94-1	3,330	-	3,330	[26]
167	2-Methylcyclohexanon	O=C1C(CCCC1)C	5162	563-60-8	2,870	-	2,870	[26]
168	Isophoron	C1(=O)CC(C)(C)CC(=C1)C	7619	78-59-1	3,370	-	3,370	[26]
169	Ethylformiat	CCOC=O	381	109-94-4	1,650	-	1,650	[26]
170	Methylacetat	COC(C)=O	380	79-20-9	1,950	2,000	1,975	[26]
171	Ethylacetat	CC(=O)OCC	382	141-78-6	1,910	1,890	1,900	[26]
172	Propylacetat	CCCOC(C)=O	385	109-60-4	1,870	1,880	1,875	[26]
173	Butylacetat	CCCCOC(C)=O	390	123-86-4	1,920	1,950	1,935	[26]
174	Pentylacetat	CCCCCOC(C)=O	395	628-63-7	1,970	1,990	1,980	[26]
175	Isopropylacetat	CC(OC(=O)C)C	389	108-21-4	1,540	1,550	1,545	[26]
176	Isobutylacetat	CC(COC(=O)C)C	393	110-19-0	1,650	1,720	1,685	[26]

## Anhang

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
177	Isopentylacetat	<chem>O=C(C)OCCC(C)C</chem>	399	123-92-2	1,770	1,810	1,790	[26]
178	2-Methoxyethanol	<chem>COCCO</chem>	770	109-86-4	4,520	4,496	4,508	[26,57]
179	2-Ethoxyethanol	<chem>CCOCCO</chem>	771	110-80-5	4,340	-	4,340	[26]
180	2-Butoxyethanol	<chem>CCCCOCCO</chem>	448	111-76-2	3,900	-	3,900	[26]
181	2-Isopropoxyethanol	<chem>OCCOC(C)C</chem>	3579	109-59-1	4,160	-	4,160	[26]
182	1-Methoxy-2-propanol	<chem>OC(COC)C</chem>	6728	107-98-2	4,090	-	4,090	[26]
183	Desfluran	<chem>C(OC(C(F)(F)F)F)(F)F</chem>	9981	57041-67-5	-0,370	-	-0,370	[26]
184	Enfluran	<chem>FC(OC(C(Cl)F)(F)F)F</chem>	5207	13838-16-9	0,250	0,450	0,350	[26]
185	Isofluran	<chem>FC(C(OC(F)F)Cl)(F)F</chem>	4821	26675-46-7	0,150	0,250	0,200	[26]
186	Methoxyfluran	<chem>COC(C(Cl)Cl)(F)F</chem>	6326	76-38-0	1,16	1,400	1,280	[26]
187	Sevofluran	<chem>FCOC(C(F)(F)F)C(F)(F)F</chem>	6738	28523-86-6	-0,200	-	-0,200	[26]
188	Fluroxen	<chem>C=COCC(F)(F)F</chem>	5208	406-90-6	0,150	-	0,150	[26]
189	Tetracyanoethylenoxid	<chem>N#CC1(C#N)OC1(C#N)C#N</chem>		3189-43-3	-	3,220	3,220	[26]
190	1-Nitropropan	<chem>CCC[N+](=O)[O-]</chem>	576	108-03-2	2,270	2,350	2,310	[26]
191	2-Nitropropan	<chem>CC([N+](=O)[O-])C</chem>	577	79-46-9	2,190	2,260	2,225	[26]

## Anhang

**Tabelle 8.2: Experimentelle und berechnete LSER-Deskriptoren für den  $K_{ba}$ -Datensatz.**

Nr.	Name	exp.					ber.					
		<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>V</i>
1	Ethan	0	0	0	0	0,492	0,04	0,127	0,003	0,085	0,772	0,3904
2	<i>n</i> -Butan	0	0	0	0	1,615	0,04	0,127	0,003	0,085	1,77	0,6722
3	Pentan	0	0	0	0	2,162	0,04	0,127	0,003	0,085	2,269	0,8131
4	Hexan	0	0	0	0	2,668	0,04	0,127	0,003	0,085	2,768	0,954
5	Heptan	0	0	0	0	3,173	0,04	0,127	0,003	0,085	3,267	1,0949
6	Oktan	0	0	0	0	3,677	0,04	0,127	0,003	0,085	3,766	1,2358
7	Nonan	0	0	0	0	4,182	0,04	0,127	0,003	0,085	4,265	1,3767
8	Dekan	0	0	0	0	4,686	0,04	0,127	0,003	0,085	4,764	1,5176
9	Undekan	0	0	0	0	5,191	0,04	0,127	0,003	0,085	5,263	1,6585
10	Dodekan	0	0	0	0	5,696	0,04	0,127	0,003	0,085	5,762	1,7994
11	2-Methylpentan	0	0	0	0	2,503	0,025	0,088	0,003	0,103	2,54	0,954
12	3-Methylpentan	0	0	0	0	2,581	0,025	0,088	0,003	0,103	2,54	0,954
13	2,3-Dimethylbutan	0	0	0	0	2,495	0,01	0,049	0,003	0,121	2,312	0,954
14	3-Methylhexan	0	0	0	0	3,044	0,025	0,088	0,003	0,103	3,039	1,0949
15	2-Methylheptan	0	0	0	0	3,48	0,025	0,088	0,003	0,103	3,538	1,2358
16	2,3,4-Trimethylpentan	0	0	0	0	3,481	-0,005	0,01	0,003	0,139	3,082	1,2358
17	2-Methyloktan	0	0	0	0	3,966	0,025	0,088	0,003	0,103	4,037	1,3767
18	2-Methylnonan	0	0	0	0	4,453	0,025	0,088	0,003	0,103	4,536	1,5176
19	2,2-Dimethylbutan	0	0	0	0	2,352	0,019	0,048	0,003	0,136	2,356	0,954
20	2,2,4-Trimethylpentan	0	0	0	0	3,106	0,004	0,009	0,003	0,154	3,126	1,2358
21	Cyclopropan	0,41	0,23	0	0	1,314	0,248	0,277	0,003	0,071	1,627	0,4227
22	Cyclopentan	0,26	0,1	0	0	2,477	0,248	0,277	0,003	0,071	2,625	0,7045
23	Cyclohexan	0,31	0,1	0	0	2,964	0,248	0,277	0,003	0,071	3,124	0,8454
24	Cycloheptan	0,35	0,1	0	0	3,704	0,248	0,277	0,003	0,071	3,623	0,9863
25	Methylcyclopentan	0,23	0,1	0	0	2,907	0,233	0,238	0,003	0,089	2,896	0,8454
26	Methylcyclohexan	0,24	0,06	0	0	3,319	0,233	0,238	0,003	0,089	3,395	0,9863
27	1,2-Dimethylcyclohexan	0,28	0,24	0	0	3,847	0,218	0,199	0,003	0,107	3,666	1,1272
28	1,2,4-Trimethylcyclohexan	0,36	0,21	0	0	4,1	0,203	0,16	0,003	0,125	3,937	1,2681
29	JP-10	0,59	0,45	0	0,06	4,84	0,889	1,45	0,003	0,382	5,408	1,1918
30	<i>tert</i> -Butylcyclohexan	0,3	0,1	0	0,1	4,603	0,212	0,159	0,003	0,14	4,48	1,409
31	Ethen	0,11	0,1	0	0,07	0,289	0,158	0,107	0,003	0,109	0,618	0,3474

## Anhang

Nr.	Name	exp.					ber.					
		<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>V</i>
32	Propen	0,1	0,08	0	0,07	0,946	0,167	0,167	0,003	0,108	1,164	0,4883
33	1-Okten	0,09	0,08	0	0,07	3,568	0,167	0,167	0,003	0,108	3,659	1,1928
34	1-Nonen	0,09	0,08	0	0,07	4,073	0,167	0,167	0,003	0,108	4,158	1,3337
35	1-Deken	0,09	0,08	0	0,07	4,533	0,167	0,167	0,003	0,108	4,657	1,4746
36	1,3-Butadien	0,32	0,23	0	0,1	1,543	0,294	0,207	0,003	0,131	1,556	0,5862
37	2-Methyl-1,3-butadien	0,31	0,23	0	0,1	2,101	0,302	0,183	0,003	0,127	2,032	0,7271
38	Acetylen	0,19	0,6	0,06	0,04	0,14	0,328	0,345	0,167	0,127	0,794	0,3044
39	Benzen	0,61	0,52	0	0,14	2,786	0,656	0,577	0,003	0,137	2,944	0,7164
40	Toluen	0,6	0,52	0	0,14	3,325	0,664	0,553	0,003	0,133	3,42	0,8573
41	Ethylbenzen	0,61	0,51	0	0,15	3,778	0,664	0,553	0,003	0,133	3,919	0,9982
42	1,3-Dimethylbenzen (m-Xylol)	0,62	0,52	0	0,16	3,839	0,672	0,529	0,003	0,129	3,896	0,9982
43	<i>o</i> -Xylol	0,66	0,56	0	0,16	3,939	0,672	0,529	0,003	0,129	3,896	0,9982
44	<i>p</i> -Xylol	0,61	0,52	0	0,16	3,839	0,672	0,529	0,003	0,129	3,896	0,9982
45	1,2,4-Trimethylbenzen	0,68	0,56	0	0,19	4,441	0,68	0,505	0,003	0,125	4,372	1,1391
46	Propylbenzen	0,6	0,5	0	0,15	4,23	0,664	0,553	0,003	0,133	4,418	1,1391
47	1,2,3-Trimethylbenzen	0,73	0,61	0	0,19	4,565	0,68	0,505	0,003	0,125	4,372	1,1391
48	1,3,5-Trimethylbenzen	0,65	0,52	0	0,19	4,344	0,68	0,505	0,003	0,125	4,372	1,1391
49	Isopropylbenzen	0,6	0,49	0	0,16	4,084	0,649	0,514	0,003	0,151	4,19	1,1391
50	<i>tert</i> -Butylbenzen	0,62	0,49	0	0,18	4,413	0,643	0,474	0,003	0,184	4,505	1,28
51	Styrol	0,85	0,65	0	0,16	3,856	0,791	0,593	0,003	0,156	3,812	0,9552
52	<i>m</i> -Methylstyrol	0,87	0,65	0	0,18	4,375	0,799	0,569	0,003	0,152	4,288	1,0961
53	<i>p</i> -Methylstyrol	0,87	0,65	0	0,18	4,399	0,799	0,569	0,003	0,152	4,288	1,0961
54	Allylbenzen	0,72	0,6	0	0,22	4,136	0,791	0,593	0,003	0,156	4,311	1,0961
55	Fluorethan	0,05	0,35	0	0,1	0,576	-0,036	0,16	0,003	0,042	0,807	0,4081
56	1-Fluorpropan	0,03	0,35	0	0,13	1,103	-0,036	0,16	0,003	0,042	1,306	0,549
57	2-Fluorpropan	0	0,32	0	0,1	1,07	-0,051	0,121	0,022	0,06	1,078	0,549
58	Difluormethan	-0,32	0,487	0,065	0,05	0,042	-0,129	0,143	0,003	-0,015	0,232	0,2849
59	1,1-Difluorethan	-0,25	0,47	0,04	0,07	0,57	-0,144	0,104	0,053	0,003	0,503	0,4258
60	1,3-Difluorpropan	-0,2	0,55	0,12	0,21	1,35	-0,112	0,193	0,003	-0,001	1,341	0,5667
61	Tetrafluormethan	-0,58	-0,26	0	0	-0,817	-0,183	0,48	0,003	0,048	0,667	0,3203
62	1,1,2,2-Tetrafluorethan	-0,28	-0,3	0,3	0	0,289	-0,304	0,081	0,103	-0,053	0,342	0,4612
63	1,1,1,2-Tetrafluorethan	-0,39	0,16	0,16	0,05	0,403	-0,266	0,13	0,003	-0,011	0,551	0,4612

## Anhang

Nr.	Name	exp.					ber.					
		<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>V</i>
64	1,1,2,2,3-Pentafluorpropan	-0,45	0,17	0	0,03	0,68	-0,362	0,074	0,053	-0,037	0,8	0,6198
65	1,1,2,4,4-Pentafluorbutan	-0,49	0,48	0,23	0,14	1,2	-0,395	0,075	0,122	-0,078	1,147	0,7607
66	1,1,2,2,3,3,4,4-Oktafluorbutan	-0,79	0,08	0,15	0,15	1,456	-0,636	-0,045	0,103	-0,099	0,65	0,8138
67	1,1,1,2,3,4,4,4-Oktafluorbutan	-0,79	0,2	0,13	0,05	1,096	-0,578	0,055	0,103	-0,045	0,98	0,8138
68	1,1,1,2,2,3,3,4,4-Nonafluorbutan	-0,89	-0,43	0,04	0,1	0,7	-0,66	0,048	0,053	-0,055	0,831	0,8315
69	Chlormethan	0,25	0,43	0	0,08	1,163	0,167	0,284	0,003	0,078	1,12	0,3719
70	Chlorethan	0,23	0,4	0	0,1	1,678	0,167	0,284	0,003	0,078	1,619	0,5128
71	1-Chlorpropan	0,22	0,4	0	0,1	2,202	0,167	0,284	0,003	0,078	2,118	0,6537
72	2-Chlorpropan	0,18	0,35	0	0,12	1,97	0,152	0,245	0,022	0,096	1,89	0,6537
73	1-Chlorbutan	0,21	0,4	0	0,1	2,722	0,167	0,284	0,003	0,078	2,617	0,7946
74	1-Chlorpentan	0,21	0,38	0	0,09	3,223	0,167	0,284	0,003	0,078	3,116	0,9355
75	Dichlormethan	0,39	0,57	0,1	0,05	2,019	0,277	0,391	0,003	0,057	1,856	0,4943
76	1,2-Dichlorethan	0,42	0,64	0,1	0,11	2,573	0,306	0,441	0,003	0,084	2,52	0,6352
77	1,1-Dichlorethan	0,32	0,49	0,1	0,1	2,316	0,262	0,352	0,053	0,075	2,127	0,6352
78	1,2-Dichlorpropan	0,37	0,63	0,1	0,17	2,836	0,291	0,402	0,022	0,102	2,791	0,7761
79	Trichlormethan	0,43	0,49	0,15	0,02	2,48	0,406	0,559	0,053	0,082	2,586	0,6167
80	1,1,2-Trichlorethan	0,5	0,68	0,13	0,13	3,29	0,413	0,509	0,053	0,094	3,082	0,7576
81	1,1,1-Trichlorethan	0,37	0,41	0	0,09	2,733	0,4	0,519	0,003	0,115	2,901	0,7576
82	1,2,3-Trichlorpropan	0,55	0,65	0,03	0,31	3,582	0,43	0,559	0,022	0,108	3,692	0,8985
83	Tetrachlormethan	0,46	0,38	0	0	2,823	0,629	0,976	0,003	0,192	3,915	0,7391
84	1,1,1,2-Tetrachlorethan	0,54	0,63	0,1	0,08	3,641	0,563	0,676	0,003	0,147	3,91	0,88
85	1,1,2,2-Tetrachlorethan	0,6	0,76	0,16	0,12	3,803	0,508	0,577	0,103	0,091	3,59	0,88
86	Pentachlorethan	0,65	0,66	0,17	0,06	4,267	0,658	0,744	0,053	0,144	4,418	1,0024
87	Hexachlorethan	0,68	0,68	0	0	4,718	0,796	0,911	0,003	0,184	5,192	1,1248
88	1-Brompropan	0,37	0,4	0	0,12	2,62	0,34	0,363	0,003	0,067	2,546	0,7063
89	2-Brompropan	0,33	0,35	0	0,14	2,39	0,325	0,324	0,022	0,085	2,318	0,7063
90	Dibrommethan	0,71	0,69	0,11	0,07	2,886	0,623	0,549	0,003	0,035	2,712	0,5995
91	1,2-Dibromethan	0,75	0,76	0,1	0,17	3,382	0,652	0,599	0,003	0,062	3,376	0,7404
92	Tribrommethan	0,97	0,68	0,15	0,06	3,784	0,925	0,796	0,053	0,049	3,87	0,7745
93	Iodethan	0,64	0,4	0	0,15	2,573	0,677	0,4	0,003	0,078	2,54	0,6486
94	Fluorchlormethan	0,04	0,61	0,07	0,04	0,982	0,074	0,267	0,003	0,021	1,044	0,3896
95	1-Chlor-2,2,2-trifluorethan	0,01	0,4	0,15	0	1,168	-0,063	0,254	0,003	0,025	1,363	0,5659

## Anhang

Nr.	Name	exp.					ber.					
		<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>V</i>
96	1-Chlor-1,2,2,2-tetrafluorethan	-0,3	0,17	0,097	0,07	0,904	-0,137	0,298	0,053	0,014	1,281	0,5836
97	1,1-Dichlor-1-fluorethan	0,08	0,43	0,005	0,05	1,92	0,197	0,395	0,003	0,079	2,089	0,6529
98	Difluordichlormethan	0,04	0,04	0	0,04	0,998	0,223	0,728	0,003	0,12	2,291	0,5297
99	1,1,1-Trifluor-2,2-dichlorethan	-0,16	0,4	0,22	0	1,746	0,066	0,422	0,053	0,05	2,093	0,6883
100	1,2-Dichlortetrafluorethan	-0,19	0,05	0	0	1,427	-0,016	0,415	0,003	0,04	1,944	0,706
101	Fluortrichlormethan	0,21	0,24	0	0,07	1,95	0,426	0,852	0,003	0,156	3,103	0,6344
102	Tefluran	-0,07	0,21	0,2	0,02	1,37	0,036	0,377	0,053	0,003	1,709	0,6362
103	Halopropan	-0,07	0,28	0,2	0	2,03	0,014	0,277	0,053	-0,012	2,04	0,7771
104	Bromchlormethan	0,54	0,8	0,01	0,06	2,445	0,45	0,47	0,003	0,046	2,284	0,5469
105	1-Brom-2-chlorethan	0,57	0,7	0,1	0,09	2,982	0,479	0,52	0,003	0,073	2,948	0,6878
106	Halothan	0,1	0,39	0,13	0,05	1,953	0,239	0,501	0,053	0,039	2,521	0,7409
107	Bromdichlormethan	0,59	0,69	0,1	0,04	2,891	0,579	0,638	0,053	0,071	3,014	0,6693
108	Chlordibrommethan	0,78	0,68	0,12	0,1	3,304	0,752	0,717	0,053	0,06	3,442	0,7219
109	1,1-Difluorethen	-0,1	0	0	0,05	0,24	0,023	0,209	0,003	0,018	0,712	0,3828
110	Vinylchlorid	0,26	0,38	0	0,05	1,404	0,294	0,324	0,003	0,101	1,512	0,4698
111	Allylchlorid	0,33	0,56	0	0,05	2,109	0,294	0,324	0,003	0,101	2,011	0,6107
112	<i>cis</i> -1,2-Dichlorethen	0,44	0,61	0,11	0,05	2,439	0,43	0,541	0,003	0,093	2,406	0,5922
113	<i>trans</i> -1,2-Dichlorethen	0,43	0,41	0,09	0,05	2,278	0,43	0,541	0,003	0,093	2,406	0,5922
114	1,1-Dichlorethen	0,36	0,34	0	0,05	2,11	0,429	0,457	0,003	0,09	2,336	0,5922
115	Trichlorethen	0,52	0,37	0,08	0,03	2,997	0,565	0,674	0,003	0,082	3,23	0,7146
116	Tetrachlorethen	0,64	0,44	0	0	3,584	0,7	0,807	0,003	0,071	4,054	0,837
117	Vinylbromid	0,56	0,5	0	0,07	1,846	0,467	0,403	0,003	0,09	1,94	0,5224
118	1-Chlor-2,2-difluorethen	-0,34	0,285	0,15	0	0,723	0,159	0,426	0,003	0,01	1,606	0,5052
119	$\beta$ -Chloropren	0,45	0,36	0	0,15	2,513	0,429	0,34	0,003	0,12	2,38	0,7086
120	Fluorbenzen	0,48	0,57	0	0,1	2,788	0,538	0,628	0,003	0,126	2,952	0,7341
121	1,2-Difluorbenzen	0,39	0,63	0	0,06	2,843	0,42	0,636	0,003	0,146	2,879	0,7518
122	1,4-Difluorbenzen	0,38	0,6	0	0,06	2,766	0,42	0,679	0,003	0,069	2,96	0,7518
123	1,2,4-Trifluorbenzen	0,31	0,65	0	0,05	2,801	0,302	0,649	0,003	0,054	2,861	0,7695
124	1,3,5-Trifluorbenzen	0,27	0,37	0	0,08	2,642	0,302	0,616	0,003	-0,001	2,89	0,7695
125	Pentafluorbenzen	0,15	0,68	0	0,02	2,578	0,066	0,508	0,003	-0,026	2,556	0,8049
126	Methylpentafluorbenzen	0,06	0,59	0	0,01	3,244	0,074	0,484	0,003	-0,03	3,032	0,9458
127	Hexafluorbenzen	0,09	0,56	0	0,01	2,345	-0,052	0,397	0,003	-0,091	2,35	0,8226

## Anhang

Nr.	Name	exp.					ber.					
		<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>V</i>
128	Chlorbenzen	0,72	0,65	0	0,07	3,657	0,791	0,71	0,003	0,126	3,768	0,8388
129	Chlorbenzen	0,72	0,65	0	0,07	3,657	0,791	0,71	0,003	0,126	3,768	0,8388
130	1,2-Dichlorbenzen	0,87	0,78	0	0,04	4,518	0,926	0,8	0,003	0,146	4,511	0,9612
131	1,3-Dichlorbenzen	0,85	0,73	0	0,02	4,41	0,926	0,805	0,003	0,08	4,566	0,9612
132	1,4-Dichlorbenzen	0,83	0,75	0	0,02	4,435	0,926	0,843	0,003	0,069	4,592	0,9612
133	Brombenzen	0,88	0,73	0	0,09	4,041	0,964	0,789	0,003	0,115	4,196	0,8914
134	4-Chlorbenzotrifluorid	0,53	0,58	0	0,01	3,73	0,55	0,706	0,003	0,044	3,937	1,0328
135	Methanol	0,28	0,44	0,43	0,47	0,97	0,205	0,449	0,348	0,385	1,123	0,3082
136	Ethanol	0,25	0,42	0,37	0,48	1,485	0,205	0,449	0,348	0,385	1,622	0,4491
137	1-Propanol	0,24	0,42	0,37	0,48	2,031	0,205	0,449	0,348	0,385	2,121	0,59
138	1-Butanol	0,22	0,42	0,37	0,48	2,601	0,205	0,449	0,348	0,385	2,62	0,7309
139	1-Pentanol	0,22	0,42	0,37	0,48	3,106	0,205	0,449	0,348	0,385	3,119	0,8718
140	1-Hexanol	0,21	0,42	0,37	0,48	3,61	0,205	0,449	0,348	0,385	3,618	1,0127
141	2-Methyl-1-propanol	0,22	0,39	0,37	0,48	2,413	0,19	0,41	0,348	0,403	2,392	0,7309
142	3-Methyl-1-butanol	0,19	0,39	0,37	0,48	3,011	0,19	0,41	0,348	0,403	2,891	0,8718
143	2-Propanol	0,21	0,36	0,33	0,56	1,764	0,19	0,41	0,348	0,403	1,893	0,59
144	2-Methyl-2-propanol	0,18	0,3	0,31	0,6	1,963	0,184	0,37	0,348	0,436	2,208	0,7309
145	<i>tert</i> -Pentylalkohol	0,19	0,3	0,31	0,6	2,63	0,184	0,37	0,348	0,436	2,707	0,8718
146	Dimethylether	0	0,27	0	0,41	1,285	0,054	0,312	0,003	0,296	1,132	0,4491
147	Diethylether	0,04	0,25	0	0,45	2,015	0,054	0,312	0,003	0,296	2,13	0,7309
148	Methyl- <i>tert</i> -butylether	0,02	0,21	0	0,59	2,38	0,033	0,233	0,003	0,347	2,217	0,8718
149	<i>tert</i> -Amylmethylether	0,05	0,21	0	0,6	2,916	0,033	0,233	0,003	0,347	2,716	1,0127
150	Ethyl- <i>tert</i> -butylether	-0,02	0,16	0	0,6	2,72	0,033	0,233	0,003	0,347	2,716	1,0127
151	Ethyl- <i>tert</i> -pentylether	0	0,16	0	0,61	3,256	0,033	0,233	0,003	0,347	3,215	1,1536
152	Divinylether	0,26	0,39	0	0,13	1,76	0,308	0,392	0,003	0,342	1,916	0,6449
153	Ethylenoxid	0,25	0,74	0,07	0,32	1,371	0,261	0,462	0,003	0,402	1,487	0,3405
154	Propylenoxid	0,24	0,57	0	0,45	1,775	0,246	0,423	0,003	0,42	1,758	0,4814
155	1,2-Epoxy-3-buten	0,37	0,47	0	0,36	2,257	0,373	0,463	0,003	0,443	2,15	0,5793
156	1,4-Dioxan	0,33	0,75	0	0,64	2,87	0,274	0,647	0,003	0,733	2,844	0,681
157	Furan	0,37	0,53	0	0,13	1,913	0,395	0,477	0,003	0,162	2,063	0,5363
158	Propanon	0,18	0,7	0,04	0,49	1,696	0,179	0,598	0,003	0,419	1,891	0,547
159	Butan-2-on	0,17	0,7	0	0,51	2,287	0,179	0,598	0,003	0,419	2,39	0,6879

## Anhang

Nr.	Name	exp.					ber.					
		<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>V</i>
160	2-Pentanon	0,14	0,68	0	0,51	2,755	0,179	0,598	0,003	0,419	2,889	0,8288
161	3-Pentanon	0,15	0,66	0	0,51	2,811	0,179	0,598	0,003	0,419	2,889	0,8288
162	2-Hexanon	0,14	0,68	0	0,51	3,286	0,179	0,598	0,003	0,419	3,388	0,9697
163	2-Heptanon	0,12	0,68	0	0,51	3,76	0,179	0,598	0,003	0,419	3,887	1,1106
164	4-Methyl-2-pentanon	0,11	0,65	0	0,51	3,089	0,164	0,559	0,003	0,437	3,16	0,9697
165	3-Methylpentan-2-on	0,11	0,65	0	0,51	3,163	0,164	0,559	0,003	0,437	3,16	0,9697
166	Cylohexanon	0,4	0,86	0	0,56	3,792	0,387	0,748	0,003	0,405	3,744	0,8611
167	2-Methylcyclohexanon	0,37	0,83	0	0,56	4,055	0,372	0,709	0,003	0,423	4,015	1,002
168	Isophoron	0,51	1,12	0	0,53	4,74	0,51	0,745	0,003	0,474	4,746	1,2408
169	Ethylformiat	0,15	0,66	0	0,38	1,845	0,072	0,582	0,003	0,428	1,884	0,6057
170	Methylacetat	0,14	0,64	0	0,45	1,911	0,08	0,558	0,003	0,424	1,861	0,6057
171	Ethylacetat	0,11	0,62	0	0,45	2,314	0,08	0,558	0,003	0,424	2,36	0,7466
172	Propylacetat	0,09	0,6	0	0,45	2,819	0,08	0,558	0,003	0,424	2,859	0,8875
173	Butylacetat	0,07	0,6	0	0,45	3,353	0,08	0,558	0,003	0,424	3,358	1,0284
174	Pentylacetat	0,07	0,6	0	0,45	3,844	0,08	0,558	0,003	0,424	3,857	1,1693
175	Isopropylacetat	0,06	0,57	0	0,47	2,546	0,065	0,519	0,003	0,442	2,631	0,8875
176	Isobutylacetat	0,05	0,57	0	0,47	3,161	0,065	0,519	0,003	0,442	3,13	1,0284
177	Isopentylacetat	0,05	0,57	0	0,47	3,74	0,065	0,519	0,003	0,442	3,629	1,1693
178	2-Methoxyethanol	0,27	0,5	0,3	0,84	2,49	0,219	0,634	0,348	0,596	2,481	0,6487
179	2-Ethoxyethanol	0,24	0,52	0,31	0,81	2,792	0,219	0,634	0,348	0,596	2,98	0,7896
180	2-Butoxyethanol	0,2	0,5	0,3	0,83	3,806	0,219	0,634	0,348	0,596	3,978	1,0714
181	2-Isopropoxyethanol	0,2	0,47	0,3	0,91	3,17	0,204	0,595	0,348	0,614	3,251	0,9305
182	1-Methoxy-2-propanol	0,22	0,54	0,31	0,82	2,655	0,233	0,528	0,259	0,614	2,68	0,7896
183	Desfluran	-0,47	0,38	0,05	0,04	0,99	-0,451	0,253	0,103	0,136	0,913	0,6962
184	Enfluran	-0,24	0,54	0,01	0,1	2,009	-0,277	0,327	0,103	0,145	1,56	0,8009
185	Isofluran	-0,24	0,56	0,095	0,08	1,969	-0,248	0,377	0,103	0,172	1,725	0,8009
186	Methoxyfluran	0,11	0,67	0,07	0,14	2,864	0,11	0,474	0,053	0,263	2,641	0,8702
187	Sevofluran	-0,47	0,56	0	0,1	1,502	-0,48	0,296	0,053	0,179	1,536	0,8548
188	Fluroxen	0,18	0,3	0	0,1	1,6	-0,068	0,372	0,003	0,241	1,716	0,741
189	Tetracyanoethylenoxid	0,390	1,000	0,000	0,520	2,543	0,785	3,416	0,003	1,554	5,617	0,9593
190	1-Nitropropan	0,24	0,95	0	0,31	2,894	0,262	0,942	0,003	0,27	2,717	0,7055
191	2-Nitropropan	0,22	0,92	0	0,33	2,55	0,247	0,903	0,022	0,288	2,489	0,7055

## Anhang

### 8.1.2 Werte aller zur $K_{ba}$ -Modellierung verwendeten Verteilungskoeffizienten

**Tabelle 8.3: Alle Verteilungskoeffizienten mit Werten, die zur Modellierung von  $K_{ba}$  verwendet werden.**

Nr.	Name	$\log K_{wa}$	$\log K_{ow}$	$\log K_{oa}$	$\log K_{mw}$	$\log K_{HSA}$	$\log K_{HSA/a}$	$\log K_{olive}$
1	Ethan	-1,352	1,320	-0,032	1,140	-0,827	-2,216	1,137
2	<i>n</i> -Butan	-1,598	2,310	0,712	1,957	-0,371	-2,005	2,312
3	Pentan	-1,722	2,800	1,078	2,365	-0,143	-1,901	2,900
4	Hexan	-1,844	3,290	1,445	2,774	0,086	-1,794	3,487
5	Heptan	-1,968	3,780	1,812	3,183	0,314	-1,688	4,075
6	Oktan	-2,090	4,270	2,180	3,591	0,542	-1,581	4,662
7	Nonan	-2,214	4,760	2,546	4,000	0,770	-1,475	5,250
8	Dekan	-2,336	5,250	2,914	4,408	0,999	-1,367	5,837
9	Undekan	-2,459	5,740	3,281	4,817	1,227	-1,259	6,425
10	Dodekan	-2,582	6,230	3,648	5,226	1,455	-1,151	7,013
11	2-Methylpentan	-1,844	3,210	1,365	2,734	0,052	-1,828	3,421
12	3-Methylpentan	-1,844	3,210	1,365	2,734	0,052	-1,828	3,421
13	2,3-Dimethylbutan	-1,844	3,140	1,295	2,694	0,018	-1,862	3,355
14	3-Methylhexan	-1,968	3,710	1,742	3,143	0,280	-1,722	4,009
15	2-Methylheptan	-2,090	4,200	2,110	3,551	0,508	-1,615	4,596
16	2,3,4-Trimethylpentan	-2,090	4,050	1,960	3,472	0,441	-1,682	4,464
17	2-Methyloktan	-2,214	4,690	2,476	3,960	0,737	-1,509	5,184
18	2-Methylnonan	-2,336	5,180	2,844	4,369	0,965	-1,400	5,771
19	2,2-Dimethylbutan	-1,844	3,180	1,335	2,656	-0,004	-1,883	3,287
20	2,2,4-Trimethylpentan	-2,090	4,090	2,000	3,433	0,419	-1,704	4,396
21	Cyclopropan	-0,649	1,700	1,051	1,343	-0,631	-1,318	1,341
22	Cyclopentan	-0,895	2,680	1,785	2,160	-0,174	-1,107	2,516
23	Cyclohexan	-1,018	3,180	2,162	2,568	0,054	-1,002	3,104
24	Cycloheptan	-1,142	3,670	2,528	2,977	0,282	-0,897	3,691
25	Methylcyclopentan	-1,018	3,100	2,082	2,529	0,020	-1,035	3,038
26	Methylcyclohexan	-1,142	3,590	2,448	2,937	0,248	-0,931	3,625
27	1,2-Dimethylcyclohexan	-1,265	4,010	2,745	3,306	0,443	-0,859	4,147
28	1,2,4-Trimethylcyclohexan	-1,387	4,430	3,042	3,675	0,637	-0,787	4,668
29	JP-10	-0,799	3,590	2,791	2,283	0,106	-0,731	2,452
30	<i>tert</i> -Butylcyclohexan	-1,510	4,960	3,450	4,045	0,844	-0,703	5,188

## Anhang

Nr.	Name	$\log K_{wa}$	$\log K_{ow}$	$\log K_{oa}$	$\log K_{mw}$	$\log K_{HSA}$	$\log K_{HSA/a}$	$\log K_{olive}$
31	Ethen	-0,602	1,270	0,668	1,055	-0,840	-1,480	0,924
32	Propen	-0,796	1,680	0,884	1,430	-0,624	-1,458	1,474
33	1-Okten	-1,412	4,130	2,718	3,473	0,517	-0,932	4,412
34	1-Nonen	-1,535	4,620	3,085	3,882	0,745	-0,827	4,999
35	1-Deken	-1,657	5,120	3,463	4,291	0,973	-0,720	5,587
36	1,3-Butadien	-0,503	2,030	1,527	1,721	-0,421	-0,962	1,810
37	2-Methyl-1,3-butadien	-0,698	2,580	1,882	2,167	-0,170	-0,906	2,442
38	Acetylen	0,008	0,500	0,508	0,935	-0,862	-0,892	0,323
39	Benzen	0,657	1,990	2,647	2,114	-0,059	0,560	2,238
40	Toluen	0,614	2,540	3,154	2,559	0,193	0,769	2,869
41	Ethylbenzen	0,491	3,030	3,521	2,968	0,421	0,874	3,456
42	1,3-Dimethylbenzen (m-Xylol)	0,572	3,090	3,662	3,005	0,444	0,978	3,500
43	<i>o</i> -Xylol	0,572	3,090	3,662	3,005	0,444	0,978	3,500
44	<i>p</i> -Xylol	0,572	3,090	3,662	3,005	0,444	0,978	3,500
45	1,2,4-Trimethylbenzen	0,529	3,630	4,159	3,451	0,695	1,186	4,131
46	Propylbenzen	0,367	3,520	3,887	3,377	0,649	0,979	4,044
47	1,2,3-Trimethylbenzen	0,529	3,630	4,159	3,451	0,695	1,186	4,131
48	1,3,5-Trimethylbenzen	0,529	3,630	4,159	3,451	0,695	1,186	4,131
49	Isopropylbenzen	0,367	3,450	3,817	3,337	0,616	0,945	3,978
50	<i>tert</i> -Butylbenzen	0,246	3,900	4,145	3,667	0,788	0,996	4,432
51	Styrol	0,948	2,890	3,838	2,850	0,396	1,305	3,206
52	<i>m</i> -Methylstyrol	0,904	3,440	4,344	3,296	0,647	1,513	3,837
53	<i>p</i> -Methylstyrol	0,904	3,440	4,344	3,296	0,647	1,513	3,837
54	Allylbenzen	0,496	3,390	3,886	3,259	0,624	1,082	3,793
55	Fluorethan	0,050	1,260	1,310	1,236	-0,788	-0,776	1,352
56	1-Fluorpropan	-0,074	1,760	1,686	1,644	-0,560	-0,672	1,939
57	2-Fluorpropan	-0,074	1,680	1,606	1,615	-0,590	-0,702	1,845
58	Difluormethan	-1,077	0,710	-0,367	0,989	-0,946	-2,060	1,078
59	1,1-Difluorethan	-1,199	1,130	-0,069	1,386	-0,742	-1,979	1,526
60	1,3-Difluorpropan	-1,322	1,700	0,377	1,740	-0,521	-1,880	2,154
61	Tetrafluormethan	-2,273	1,190	-1,083	0,601	-1,178	-3,482	0,615
62	1,1,2,2-Tetrafluorethan	-1,796	0,940	-0,856	1,572	-0,688	-2,520	1,801
63	1,1,1,2-Tetrafluorethan	-1,796	1,680	-0,116	1,381	-0,776	-2,607	1,725

## Anhang

Nr.	Name	log $K_{wa}$	log $K_{ow}$	log $K_{oa}$	log $K_{mw}$	log $K_{HSA}$	log $K_{HSA/a}$	log $K_{olive}$
64	1,1,2,2,3-Pentafluorpropan	-2,219	2,040	-0,179	1,910	-0,517	-2,767	2,430
65	1,1,2,4,4-Pentafluorbutan	-2,341	1,790	-0,552	2,456	-0,223	-2,594	3,097
66	1,1,2,2,3,3,4,4-Oktafluorbutan	-3,238	2,870	-0,368	2,548	-0,253	-3,467	3,406
67	1,1,1,2,3,4,4,4-Oktafluorbutan	-3,238	2,870	-0,368	2,356	-0,348	-3,562	3,094
68	1,1,1,2,2,3,3,4,4-Nonafluorbutan	-3,537	3,480	-0,057	2,346	-0,374	-3,832	3,249
69	Chlormethan	0,475	1,090	1,565	1,099	-0,796	-0,359	1,042
70	Chlorethan	0,351	1,580	1,931	1,508	-0,568	-0,255	1,630
71	1-Chlorpropan	0,227	2,070	2,297	1,917	-0,339	-0,150	2,217
72	2-Chlorpropan	0,227	2,000	2,227	1,887	-0,370	-0,181	2,123
73	1-Chlorbutan	0,105	2,560	2,665	2,325	-0,111	-0,044	2,805
74	1-Chlorpentan	-0,018	3,050	3,032	2,734	0,117	0,061	3,393
75	Dichlormethan	0,428	1,340	1,768	1,534	-0,505	-0,115	1,634
76	1,2-Dichlorethan	0,306	1,830	2,136	1,846	-0,324	-0,056	2,066
77	1,1-Dichlorethan	0,306	1,760	2,066	1,931	-0,301	-0,033	2,082
78	1,2-Dichlorpropan	0,182	2,250	2,432	2,226	-0,126	0,018	2,559
79	Trichlormethan	0,881	1,520	2,401	1,825	-0,301	0,541	1,889
80	1,1,2-Trichlorethan	0,758	2,010	2,768	2,239	-0,073	0,647	2,461
81	1,1,1-Trichlorethan	0,758	2,680	3,438	2,127	-0,138	0,583	2,416
82	1,2,3-Trichlorpropan	0,635	2,500	3,135	2,564	0,118	0,715	2,995
83	Tetrachlormethan	-0,016	2,440	2,424	1,690	-0,296	-0,350	1,727
84	1,1,1,2-Tetrachlorethan	1,210	2,930	4,139	2,405	0,075	1,246	2,738
85	1,1,2,2-Tetrachlorethan	1,210	2,190	3,399	2,662	0,194	1,365	2,914
86	Pentachlorethan	1,664	3,110	4,774	2,828	0,341	1,968	3,190
87	Hexachlorethan	0,767	4,030	4,797	3,024	0,505	1,234	3,524
88	1-Brompropan	0,212	2,160	2,372	2,192	-0,119	0,055	2,528
89	2-Brompropan	0,212	2,080	2,292	2,163	-0,149	0,025	2,434
90	Dibrommethan	1,398	1,520	2,918	2,085	-0,064	1,296	2,255
91	1,2-Dibromethan	1,275	2,010	3,285	2,397	0,117	1,353	2,687
92	Tribrommethan	2,336	1,790	4,125	2,651	0,360	2,657	2,820
93	Iodethan	0,525	2,080	2,605	2,250	0,029	0,516	2,399
94	Fluorchlormethan	-0,324	1,030	0,706	1,262	-0,725	-1,087	1,356
95	1-Chlor-2,2,2-trifluorethan	-1,044	1,990	0,945	1,654	-0,555	-1,637	2,003
96	1-Chlor-1,2,2,2-tetrafluorethan	-1,344	1,860	0,516	1,672	-0,572	-1,953	1,979

## Anhang

Nr.	Name	log $K_{wa}$	log $K_{ow}$	log $K_{oa}$	log $K_{mw}$	log $K_{HSA}$	log $K_{HSA/a}$	log $K_{olive}$
97	1,1-Dichlor-1-fluorethan	0,007	2,370	2,376	1,855	-0,358	-0,390	2,138
98	Difluordichlormethan	-1,069	1,820	0,751	1,145	-0,737	-1,844	1,171
99	1,1,1-Trifluor-2,2-dichlorethan	-0,591	2,170	1,578	1,944	-0,352	-0,981	2,257
100	1,2-Dichlortetrafluorethan	-1,790	2,780	0,989	1,934	-0,377	-2,203	2,412
101	Fluortrichlormethan	-0,318	2,130	1,812	1,418	-0,516	-0,873	1,449
102	Tefluran	-0,859	1,950	1,090	1,947	-0,352	-1,249	2,290
103	Halopropan	-0,981	2,440	1,459	2,458	-0,076	-1,094	3,018
104	Bromchlormethan	0,913	1,430	2,343	1,809	-0,284	0,591	1,945
105	1-Brom-2-chlorethan	0,790	1,920	2,710	2,122	-0,103	0,648	2,376
106	Halothan	-0,107	2,260	2,153	2,219	-0,131	-0,276	2,568
107	Bromdichlormethan	1,367	1,610	2,977	2,100	-0,081	1,248	2,199
108	Chlordibrommethan	1,851	1,700	3,551	2,375	0,139	1,952	2,510
109	1,1-Difluorethen	-0,966	1,240	0,274	1,251	-0,751	-1,754	1,359
110	Vinylchlorid	-0,259	1,620	1,361	1,390	-0,593	-0,890	1,379
111	Allylchlorid	-0,144	1,930	1,786	1,799	-0,365	-0,547	1,967
112	<i>cis</i> -1,2-Dichlorethen	-0,115	1,980	1,865	1,725	-0,345	-0,499	1,834
113	<i>trans</i> -1,2-Dichlorethen	-0,115	1,980	1,865	1,725	-0,345	-0,499	1,834
114	1,1-Dichlorethen	-0,115	2,120	2,005	1,796	-0,310	-0,463	1,916
115	Trichlorethen	0,027	2,470	2,497	2,131	-0,063	-0,074	2,371
116	Tetrachlorethen	0,171	2,970	3,141	2,536	0,220	0,353	2,907
117	Vinylbromid	0,299	1,520	1,819	1,666	-0,373	-0,112	1,690
118	1-Chlor-2,2-difluorethen	-0,824	1,600	0,776	1,586	-0,504	-1,365	1,815
119	$\beta$ -Chloropren	-0,360	2,530	2,169	2,127	-0,139	-0,537	2,347
120	Fluorbenzen	0,590	2,190	2,780	2,061	-0,123	0,429	2,257
121	1,2-Difluorbenzen	0,523	2,390	2,913	1,944	-0,234	0,251	2,157
122	1,4-Difluorbenzen	0,523	2,390	2,913	2,152	-0,098	0,387	2,502
123	1,2,4-Trifluorbenzen	0,456	2,590	3,046	2,172	-0,125	0,293	2,605
124	1,3,5-Trifluorbenzen	0,456	2,590	3,046	2,367	-0,005	0,413	2,902
125	Pentafluorbenzen	0,320	2,990	3,310	2,427	-0,053	0,229	3,122
126	Methylpentafluorbenzen	0,278	3,540	3,818	2,873	0,198	0,438	3,753
127	Hexafluorbenzen	0,255	3,200	3,455	2,662	0,047	0,264	3,536
128	Chlorbenzen	0,788	2,640	3,427	2,519	0,224	0,974	2,774
129	Chlorbenzen	0,788	2,640	3,427	2,519	0,224	0,974	2,774

## Anhang

Nr.	Name	$\log K_{wa}$	$\log K_{ow}$	$\log K_{oa}$	$\log K_{mw}$	$\log K_{HSA}$	$\log K_{HSA/a}$	$\log K_{olive}$
130	1,2-Dichlorbenzen	0,917	3,280	4,197	2,860	0,461	1,341	3,193
131	1,3-Dichlorobenzen	0,917	3,280	4,197	3,062	0,590	1,469	3,513
132	1,4-Dichlorbenzen	0,917	3,280	4,197	3,068	0,598	1,477	3,537
133	Brombenzen	1,056	2,880	3,936	2,795	0,445	1,463	3,085
134	4-Chlorbenzotrifluorid	-0,152	3,600	3,448	3,135	0,504	0,314	3,850
135	Methanol	3,758	-0,630	3,128	0,063	-1,470	2,250	-1,351
136	Ethanol	3,635	-0,140	3,495	0,472	-1,242	2,355	-0,763
137	1-Propanol	3,512	0,350	3,862	0,881	-1,014	2,461	-0,176
138	1-Butanol	3,389	0,840	4,229	1,289	-0,785	2,566	0,412
139	1-Pentanol	3,265	1,330	4,595	1,698	-0,557	2,670	0,999
140	1-Hexanol	3,143	1,820	4,963	2,106	-0,329	2,776	1,587
141	2-Methyl-1-propanol	3,389	0,770	4,159	1,249	-0,819	2,532	0,346
142	3-Methyl-1-butanol	3,265	1,260	4,525	1,658	-0,591	2,636	0,933
143	2-Propanol	3,512	0,280	3,792	0,841	-1,047	2,427	-0,242
144	2-Methyl-2-propanol	3,389	0,730	4,119	1,171	-0,875	2,476	0,212
145	<i>tert</i> -Pentylalkohol	3,265	1,220	4,485	1,580	-0,646	2,580	0,799
146	Dimethylether	1,453	0,070	1,523	0,529	-1,203	0,212	0,204
147	Diethylether	1,207	1,050	2,257	1,346	-0,747	0,422	1,379
148	Methyl- <i>tert</i> -butylether	1,083	1,430	2,513	1,636	-0,608	0,438	1,767
149	<i>tert</i> -Amylmethylether	0,960	1,920	2,880	2,045	-0,379	0,543	2,354
150	Ethyl- <i>tert</i> -butylether	0,960	1,920	2,880	2,045	-0,379	0,543	2,354
151	Ethyl- <i>tert</i> -pentylether	0,838	2,410	3,248	2,454	-0,151	0,649	2,942
152	Divinylether	-0,192	1,680	1,488	1,110	-0,797	-1,027	0,877
153	Ethylenoxid	2,309	-0,050	2,259	-0,051	-1,472	0,799	-0,771
154	Propylenoxid	2,184	0,370	2,554	0,318	-1,278	0,869	-0,249
155	1,2-Epoxy-3-buten	2,190	0,730	2,920	0,608	-1,075	1,077	0,088
156	1,4-Dioxan	3,617	-0,320	3,297	-0,219	-1,629	1,950	-1,120
157	Furan	0,658	1,360	2,018	1,366	-0,578	0,042	1,292
158	Propanon	2,693	-0,240	2,453	0,325	-1,287	1,368	-0,149
159	Butan-2-on	2,570	0,260	2,830	0,733	-1,059	1,473	0,438
160	2-Pentanon	2,448	0,750	3,197	1,142	-0,831	1,579	1,026
161	3-Pentanon	2,448	0,750	3,197	1,142	-0,831	1,579	1,026
162	2-Hexanon	2,324	1,240	3,564	1,550	-0,602	1,684	1,613

## Anhang

Nr.	Name	$\log K_{wa}$	$\log K_{ow}$	$\log K_{oa}$	$\log K_{mw}$	$\log K_{HSA}$	$\log K_{HSA/a}$	$\log K_{olive}$
163	2-Heptanon	2,201	1,730	3,931	1,959	-0,374	1,789	2,201
164	4-Methyl-2-pentanon	2,324	1,160	3,484	1,511	-0,636	1,650	1,547
165	3-Methylpentan-2-on	2,324	1,160	3,484	1,511	-0,636	1,650	1,547
166	Cylohexanon	2,680	1,130	3,810	1,345	-0,634	2,008	1,230
167	2-Methylcyclohexanon	2,557	1,540	4,097	1,714	-0,440	2,079	1,751
168	Isophoron	2,568	2,620	5,188	2,338	-0,053	2,477	2,547
169	Ethylformiat	1,885	0,320	2,205	0,388	-1,292	0,555	0,002
170	Methylacetat	2,146	0,370	2,515	0,425	-1,269	0,839	0,045
171	Ethylacetat	2,021	0,860	2,881	0,833	-1,041	0,943	0,633
172	Propylacetat	1,899	1,360	3,258	1,242	-0,812	1,048	1,220
173	Butylacetat	1,776	1,850	3,626	1,651	-0,584	1,154	1,808
174	Pentylacetat	1,652	2,340	3,992	2,059	-0,356	1,258	2,396
175	Isopropylacetat	1,899	1,280	3,178	1,202	-0,846	1,015	1,154
176	Isobutylacetat	1,776	1,770	3,546	1,611	-0,618	1,120	1,742
177	Isopentylacetat	1,652	2,260	3,912	2,019	-0,389	1,225	2,329
178	2-Methoxyethanol	5,766	-0,910	4,856	0,270	-1,389	4,339	-1,109
179	2-Ethoxyethanol	5,644	-0,420	5,223	0,678	-1,161	4,445	-0,521
180	2-Butoxyethanol	5,398	0,570	5,968	1,495	-0,704	4,655	0,654
181	2-Isopropoxyethanol	5,521	0,000	5,520	1,047	-0,966	4,516	0,000
182	1-Methoxy-2-propanol	5,644	-0,490	5,153	0,663	-1,163	4,443	-0,386
183	Desfluran	-0,463	1,200	0,737	1,413	-0,862	-1,363	1,629
184	Enfluran	0,288	2,060	2,348	1,782	-0,595	-0,344	2,063
185	Isofluran	0,288	1,510	1,798	1,686	-0,642	-0,392	1,907
186	Methoxyfluran	1,639	2,020	3,659	1,808	-0,459	1,142	1,952
187	Sevofluran	-0,883	1,750	0,867	1,655	-0,738	-1,659	2,101
188	Fluroxen	-0,389	1,330	0,941	1,398	-0,743	-1,170	1,573
189	Tetracyanoethylenoxid	8,198	0,520	8,718	-3,580	-3,373	4,787	-5,917
190	1-Nitropropan	2,518	0,950	3,468	1,064	-0,793	1,687	1,018
191	2-Nitropropan	2,518	0,870	3,388	1,034	-0,823	1,657	0,923

## Anhang

### 8.1.3 Fragmente und E-States

Tabelle 8.4: Alle 27 Fragmente, die als Ausgangspunkt für die Erstellung des Fragmentmodells für den Blut/Luft-Verteilungskoeffizienten verwendet werden.

Nr.	Fragment
1	+C#C
2	+C_tert
3	+C_sec
4	+Csec-Csec
5	C
6	H + H-CO
7	Car
8	H-Car
9	F
10	Cl
11	Br +I
12	+XCar +XCd
13	OH
14	O
15	+O_ring
16	O=C
17	+O-CO
18	C#N
19	NO2
20	+FCar
21	+CX2
22	+CX3
23	+CX4
24	+CX-CX
25	+O-C-X
26	+O-CC-X
27	+hb:oh4v +u

## Anhang

**Tabelle 8.5:** Einige wichtige Atomgruppen der E-States mit den dazugehörigen Gruppensymbolen.

Nr.	Atomgruppe	Gruppensymbol	Nr.	Atomgruppe	Gruppensymbol
1	-Li	sLi	41	>SiH-	sssSiH
2	-Be-	ssBe	42	>Si<	ssssSi
3	>Be< [-2]	ssssBe	43	-PH2	sPH2
4	-BH-	ssBH	44	-PH-	ssPH
5	>B-	sssB	45	>P-	sssP
6	>B< [-1]	ssssB	46	->P=	dsssP
7	-CH3	sCH3	41	->P<	sssssP
8	=CH2	dCH2	48	-SH	sSH
9	-CH2-	ssCH2	49	=S	dS
10	=-CH	tCH	50	-S	ssS
11	=CH-	dsCH	51	aSa	aaS
12	aCHa	aaCH	52	>S=	dssS (sulfone)
13	>CH-	sssCH	53	>-S-<	ddssS (sulfate)
14	=C=	ddC	54	-Cl	sCl
15	>C-	tsC	55	-GeH3	sGeH3
16	=C<	dssC	56	-GeH2-	ssGeH2
11	aCa-	aasC	57	>GeH-	sssGeH
18	aaCa	aaaC	58	>Ge<	ssssGe
19	>C<	ssssC	59	-ASH2	ssAsH2
20	-NH3 [+1]	sNH3	60	-ASH-	ssAsH
21	-NH2	sNH2	61	>AS-	sssAs
22	-NH2- [+1]	ssNH2	62	->AS<	sssdAs
23	=-NH	dNH	63	->As-	sssssAs
24	-NH-	ssNH	64	-SeH	sSeH
25	aNHa	aaNH	65	=Se	dSe
26	=-N	tN	66	-Se-	ssSe
27	>NH- [+1]	sssNH	61	aSea	aaSe
28	=N-	dsN	68	>Se=	dssSe
29	aNa	aaN	69	>-Se=	ddssSe
30	>N-	sssN	70	-Br	sBr
31	-N<<	ddsN (nitro)	71	-SnH3	sSnH3
32	aaNs	aasN (N oxide)	72	-SnH2-	ssSnH2

## Anhang

Nr.	Atomgruppe	Gruppensymbol	Nr.	Atomgruppe	Gruppensymbol
33	>N< [+1]	ssssN (onium)	73	>SnH-	ssssSnH
34	-OH	SOH	74	>Sn<	ssssSn
35	=O	dO	75	-I	sI
36	-O-	ssO	76	-PbH3	sPbH3
31	aOa	aaO	77	-PbH2	ssPbH2
38	-F-	sF	78	>PbH-	ssssPbH
39	-SiH3	sSiH3	79	>Pb<	ssssPb
40	-SiH2-	ssSiH2	-	-	-

## 8.2 Fett/Luft-Verteilungskoeffizient $K_{fa}$

Für den Fett/Luft-Verteilungskoeffizienten werden die Stoffe aus dem Datensatz mit ihren experimentellen Daten und alle Daten von Verteilungskoeffizienten, die zur Modellierung verwendet wurden, gezeigt.

### 8.2.1 $K_{fa}$ -Daten und -Werte

**Tabelle 8.6: Verwendete Substanzen für die  $K_{fa}$ -Modellierung mit Name, SMILES-Code, ChemProp ID und CAS Nr. und experimentelle aus der Literatur gesammelte Werte für Mensch und Ratte und Quellen.**

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
1	Pentan	CCCCC	1	109-66-0	1.598		1.598	[38]
2	Hexan	CCCCCC	3	110-54-3	2.017	2.201	2.109	[38]
3	Heptan	CCCCCCC	9	142-82-5	2.554	2.579	2.567	[38]
4	Oktan	CCCCCCCC	12	111-65-9	2.367		2.367	[38]
5	<i>n</i> -Dekan	CCCCCCCCC	1011	124-18-5		3.426	3.426	[89]
6	2-Methylpentan	CC(CCC)C	4	107-83-5	1.940		1.940	[38]
7	3-Methylpentan	CCC(CC)C	5	96-14-0	2,009	-	2,009	[38]
8	3-Methylhexan	CCC(CCC)C	921	589-34-4	2,442	-	2,442	[38]
9	2,3,4-Trimethylpentan	CC(C(C(C)C)C)C	936	565-75-3	-	2,646	2,646	[38]
10	2,2-Dimethylbutan	CCC(C)(C)C	6	75-83-2	1,820	-	1,820	[38]

## Anhang

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
11	2,2,4-Trimethylpentan	CC(CC(C)(C)C)C	13	540-84-1	-	2,467	2,467	[38]
12	Cyclopropan	C1CC1	912	75-19-4	0,903	-	0,903	[38]
13	Cyclohexan	C1CCCCC1	7	110-82-7	2,415	2,371	2,393	[38]
14	Methylcyclopentan	CC1CCCC1	8	96-37-7	2,246	-	2,246	[38]
15	JP-10	C1CC2CC1C3C2CCC3	10708	2825-82-3	-	4,006	4,006	[39]
16	Ethen	C=C	700	74-85-1	-0,022	-	-0,022	[38]
17	2-Methyl-1,3-butadien	C(C(=C)C)=C	27	78-79-5	-	1,857	1,857	[38]
18	Benzol	c1ccccc1	33	71-43-2	2,579	2,636	2,608	[38]
19	Toluol	Cc1ccccc1	34	108-88-3	2,983	2,996	2,990	[38]
20	Ethylbenzen	CCc1ccccc1	36	100-41-4	3,246	3,192	3,219	[38]
21	1,3-Dimethylbenzen ( <i>m</i> -Xylol)	Cc1cc(ccc1)C	38	108-38-3	3,283	3,321	3,302	[38]
22	<i>o</i> -Xylen	Cc1c(cccc1)C	35	95-47-6	3,391	3,381	3,386	[38]
23	<i>p</i> -Xylen	Cc1ccc(cc1)C	37	106-42-3	3,305	3,243	3,274	[38]
24	Styrol	C=Cc1ccccc1	58	100-42-5	3,503	3,528	3,516	[38]
25	<i>m</i> -Methylstyrol	c1c(C=C)cc(cc1)C	1726	100-80-1	-	4,077	4,077	[38]
26	<i>p</i> -Methylstyrol	c1c(ccc(c1)C)C=C	1729	622-97-9	-	4,052	4,052	[38]
27	Allylbenzen	C=CCc1ccccc1	1727	300-57-2	-	3,680	3,680	[75]
28	Difluormethan	FCF	1852	75-10-5	-	0,155	0,155	[38]
29	Methylchlorid	ClC	710	74-87-3	-	1,130	1,130	[38]
30	Chlorethan	ClCC	720	75-00-3	-	1,587	1,587	[38]
31	1-Chlorpropan	ClCCC	123	540-54-5	-	2,072	2,072	[38]
32	2-Chlorpropan	CC(C)Cl	124	75-29-6	-	1,835	1,835	[38]
33	Dichlormethan	ClCCl	100	75-09-2	1,929	2,079	2,004	[38]
34	1,2-Dichlorethan	ClCCCl	112	107-06-2	-	2,698	2,698	[38]
35	1,1-Dichlorethan	ClC(C)Cl	111	75-34-3	-	2,215	2,215	[38]
36	1,2-Dichlorpropan	ClCC(C)Cl	130	78-87-5	-	2,698	2,698	[38]
37	Trichlormethan	ClC(Cl)Cl	103	67-66-3	2,447	2,307	2,377	[38]
38	1,1,2-Trichlorethan	ClC(CCl)Cl	116	79-00-5	-	3,158	3,158	[38]
39	1,1,1-Trichlorethan	ClC(C)(Cl)Cl	115	71-55-6	2,400	2,420	2,410	[38,12]
40	Tetrachlormethan	ClC(Cl)(Cl)Cl	107	56-23-5	-	2,555	2,555	[38]
41	1,1,1,2-Tetrachlorethan	ClC(CCl)(Cl)Cl	118	630-20-6	-	3,332	3,332	[38]
42	1,1,2,2-Tetrachlorethan	ClC(C(Cl)Cl)Cl	117	79-34-5	-	3,576	3,576	[38]

## Anhang

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
43	Pentachlorethan	<chem>C1C(Cl)Cl(Cl)Cl</chem>	120	76-01-7	-	3,615	3,615	[38]
44	Hexachlorethan	<chem>C1C(Cl)Cl(Cl)Cl(Cl)Cl</chem>	122	67-72-1	-	3,521	3,521	[38]
45	1-Brompropan	<chem>BrCCC</chem>	125	106-94-5	-	2,373	2,373	[38]
46	2-Brompropan	<chem>CC(C)Br</chem>	126	75-26-3	-	2,199	2,199	[38]
47	Dibrommethan	<chem>BrCBr</chem>	102	74-95-3	-	2,899	2,899	[38]
48	1,2-Dibromethan	<chem>BrCCBr</chem>	114	106-93-4	-	3,086	3,086	[38]
49	Fluorchlormethan	<chem>FCCl</chem>	712	593-70-4	-	1,188	1,188	[38]
50	1-Chlor-2,2,2-trifluoethan	<chem>FC(F)(F)CCl</chem>	878	75-88-7	1,531	1,326	1,429	[38]
51	Tefluran	<chem>FC(C(F)Br)(F)F</chem>	6687	124-72-1	1,301	-	1,301	[38]
52	Halopropan	<chem>C(C(C(F)F)(F)F)Br</chem>	12570	679-84-5	2,350	-	2,350	[12]
53	Bromochloromethan	<chem>C1CBr</chem>	101	74-97-5	-	2,512	2,512	[38]
54	1-Bromo-2-chloroethan	<chem>C1CCBr</chem>	113	107-04-0	-	2,982	2,982	[38]
55	Halothan	<chem>FC(C(Cl)Br)(F)F</chem>	1902	151-67-7	2,226	2,260	2,243	[38]
56	Chlordibrommethan	<chem>C1C(Br)Br</chem>	105	124-48-1	-	3,283	3,283	[38]
57	Vinylchlorid	<chem>C1C=C</chem>	879	75-01-4	-	1,301	1,301	[38]
58	Allylchlorid	<chem>C1CC=C</chem>	160	107-05-1	-	2,004	2,004	[38]
59	<i>cis</i> -1,2-Dichlorethylen	<chem>C1C=CCl</chem>	4438	156-59-2	-	2,356	2,356	[38]
60	<i>trans</i> -1,2-Dichlorethylen	<chem>C1C=CCl</chem>	155	156-60-5	-	2,170	2,170	[38]
61	1,1-Dichlorethylen	<chem>C1C(=C)Cl</chem>	154	75-35-4	-	1,836	1,836	[38]
62	Trichlorethen	<chem>C1C(=CCl)Cl</chem>	158	79-01-6	2,767	2,744	2,756	[38]
63	Tetrachlorethen	<chem>C1C(=C(Cl)Cl)Cl</chem>	159	127-18-4	3,161	3,214	3,188	[38,77]
64	1-Chlor-2,2-difluoethen	<chem>FC(=CCl)F</chem>	2108	359-10-4	1,301	-	1,301	[18]
65	Chlorbenzol	<chem>Clc1ccccc1</chem>	168	108-90-7	-	3,106	3,106	[38]
66	Methanol	<chem>CO</chem>	737	67-56-1	2,364	2,286	2,325	[38]
67	Ethanol	<chem>CCO</chem>	738	64-17-5	2,332	2,354	2,343	[38]
68	1-Propanol	<chem>CCCO</chem>	741	71-23-8	2,471	2,604	2,538	[38]
69	1-Butanol	<chem>CCCCO</chem>	262	71-36-3	-	2,954	2,954	[38]
70	1-Pentanol	<chem>CCCCCO</chem>	265	71-41-0	-	3,408	3,408	[38]
71	2-Methyl-1-propanol	<chem>CC(CO)C</chem>	264	78-83-1	2,589	2,857	2,723	[38]
72	3-Methyl-1-butanol (Isopentanol)	<chem>CC(CCO)C</chem>	270	123-51-3	-	3,176	3,176	[38]
73	2-Propanol	<chem>CC(O)C</chem>	740	67-63-0	2,255	2,438	2,347	[38]
74	Diethylether	<chem>CCOCC</chem>	422	60-29-7	1,699	1,679	1,689	[38]

## Anhang

Nr.	Name	SMILES-Code	ChemProp ID	CAS Nr.	exp. log $K_{ba}$ (Mensch)	exp. log $K_{ba}$ (Ratte)	exp. log $K_{ba}$ (Mittelwert)	Quelle
75	Divinylether	<chem>C(OC=C)=C</chem>	2603	109-93-3	1,602	-	1,602	[38]
76	Propanon	<chem>CC(=O)C</chem>	754	67-64-1	1,934	1,892	1,913	[38]
77	Butan-2-on	<chem>CCC(=O)C</chem>	343	78-93-3	2,210	2,301	2,256	[38]
78	2-Pentanon	<chem>CCCC(=O)C</chem>	344	107-87-9	-	2,571	2,571	[38]
79	2-Heptanon	<chem>CCCCCC(=O)C</chem>	348	110-43-0	-	3,663	3,663	[38]
80	4-Methyl-2-pentanon	<chem>CC(CC(=O)C)C</chem>	346	108-10-1	-	2,719	2,719	[38]
81	Methylacetat	<chem>COC(C)=O</chem>	380	79-20-9	-	1,996	1,996	[38]
82	Ethylacetat	<chem>CC(=O)OCC</chem>	382	141-78-6	-	2,185	2,185	[38]
83	<i>n</i> -Propylacetat	<chem>CCCOC(C)=O</chem>	385	109-60-4	-	2,711	2,711	[38]
84	<i>n</i> -Butylacetat	<chem>CCCCOC(C)=O</chem>	390	123-86-4	-	3,182	3,182	[38]
85	<i>n</i> -Pentylacetat	<chem>CCCCCOC(C)=O</chem>	395	628-63-7	-	3,572	3,572	[38]
86	Isopropylacetat	<chem>CC(OC(=O)C)C</chem>	389	108-21-4	-	2,481	2,481	[38]
87	Isobutylacetat	<chem>CC(COC(=O)C)C</chem>	393	110-19-0	-	3,045	3,045	[38]
88	Isopentylacetat	<chem>O=C(C)OCCC(C)C</chem>	399	123- 92-2	-	3,439	3,439	[38]
89	Desfluran	<chem>C(OC(C(F)(F)F)F)(F)F</chem>	9981	57041-67-5	1,079	-	1,079	[38]
90	Enfluran	<chem>FC(OC(C(Cl)F)(F)F)F</chem>	5207	13838-16-9	2,035	-	2,035	[38]
91	Isofluran	<chem>FC(C(OC(F)F)Cl)(F)F</chem>	4821	26675-46-7	1,843	1,992	1,918	[38]
92	Methoxyfluran	<chem>COC(C(Cl)Cl)(F)F</chem>	6326	76-38-0	2,902	-	2,902	[38]
93	Sevofluran	<chem>FCOC(C(F)(F)F)C(F)(F)F</chem>	6738	28523-86-6	1,585	-	1,585	[38]
94	Fluroxen	<chem>C=COCC(F)(F)F</chem>	5208	406-90-6	1,531	-	1,531	[38]
95	1-Nitropropan	<chem>CCC[N+](=O)[O-]</chem>	576	108-03-2	-	2,704	2,704	[38]
96	2-Nitropropan	<chem>CC([N+](=O)[O-])C</chem>	577	79-46-9	-	2,190	2,190	[38]

## Anhang

**Tabelle 8.7:** Experimentelle und berechnete LSER-Deskriptoren für den  $K_{ba}$ -Datensatz.

Nr.	Name	exp.					ber.					
		$E$	$S$	$A$	$B^H$	$L$	$E$	$S$	$A$	$B^H$	$L$	$V$
1	Pentan	0	0	0	0	2,162	0,04	0,127	0,003	0,085	2,269	0,8131
2	Hexan	0	0	0	0	2,668	0,04	0,127	0,003	0,085	2,768	0,954
3	Heptan	0	0	0	0	3,173	0,04	0,127	0,003	0,085	3,267	1,0949
4	Oktan	0	0	0	0	3,677	0,04	0,127	0,003	0,085	3,766	1,2358
5	<i>n</i> -Dekan	0	0	0	0	4,686	0,04	0,127	0,003	0,085	4,764	1,5176
6	2-Methylpentan	0	0	0	0	2,503	0,025	0,088	0,003	0,103	2,54	0,954
7	3-Methylpentan	0	0	0	0	2,581	0,025	0,088	0,003	0,103	2,54	0,954
8	3-Methylhexan	0	0	0	0	3,044	0,025	0,088	0,003	0,103	3,039	1,0949
9	2,3,4-Trimethylpentan	0	0	0	0	3,481	-0,005	0,01	0,003	0,139	3,082	1,2358
10	2,2-Dimethylbutan	0	0	0	0	2,352	0,019	0,048	0,003	0,136	2,356	0,954
11	2,2,4-Trimethylpentan	0	0	0	0	3,106	0,004	0,009	0,003	0,154	3,126	1,2358
12	Cyclopropan	0,41	0,23	0	0	1,314	0,248	0,277	0,003	0,071	1,627	0,4227
13	Cyclohexan	0,31	0,1	0	0	2,964	0,248	0,277	0,003	0,071	3,124	0,8454
14	Methylcyclopentan	0,23	0,1	0	0	2,907	0,233	0,238	0,003	0,089	2,896	0,8454
15	JP-10	0,59	0,45	0	0,06	4,84	0,889	1,45	0,003	0,382	5,408	1,1918
16	Ethen	0,11	0,1	0	0,07	0,289	0,158	0,107	0,003	0,109	0,618	0,3474
17	2-Methyl-1,3-butadien	0,31	0,23	0	0,1	2,101	0,302	0,183	0,003	0,127	2,032	0,7271
18	Benzol	0,61	0,52	0	0,14	2,786	0,656	0,577	0,003	0,137	2,944	0,7164
19	Toluol	0,6	0,52	0	0,14	3,325	0,664	0,553	0,003	0,133	3,42	0,8573
20	Ethylbenzen	0,61	0,51	0	0,15	3,778	0,664	0,553	0,003	0,133	3,919	0,9982
21	1,3-Dimethylbenzen ( <i>m</i> -Xylol)	0,62	0,52	0	0,16	3,839	0,672	0,529	0,003	0,129	3,896	0,9982
22	<i>o</i> -Xylen	0,66	0,56	0	0,16	3,939	0,672	0,529	0,003	0,129	3,896	0,9982
23	<i>p</i> -Xylen	0,61	0,52	0	0,16	3,839	0,672	0,529	0,003	0,129	3,896	0,9982
24	Styrol	0,85	0,65	0	0,16	3,856	0,791	0,593	0,003	0,156	3,812	0,9552
25	<i>m</i> -Methylstyrol	0,87	0,65	0	0,18	4,375	0,799	0,569	0,003	0,152	4,288	1,0961
26	<i>p</i> -Methylstyrol	0,87	0,65	0	0,18	4,399	0,799	0,569	0,003	0,152	4,288	1,0961
27	Allylbenzen	0,72	0,6	0	0,22	4,136	0,791	0,593	0,003	0,156	4,311	1,0961
28	Difluormethan	-0,32	0,487	0,065	0,05	0,042	-0,129	0,143	0,003	-0,015	0,232	0,2849
29	Methylchlorid	0,25	0,43	0	0,08	1,163	0,167	0,284	0,003	0,078	1,12	0,3719
30	Chlorethan	0,23	0,4	0	0,1	1,678	0,167	0,284	0,003	0,078	1,619	0,5128
31	1-Chlorpropan	0,22	0,4	0	0,1	2,202	0,167	0,284	0,003	0,078	2,118	0,6537

## Anhang

Nr.	Name	exp.					ber.					
		<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>V</i>
32	2-Chlorpropan	0,18	0,35	0	0,12	1,97	0,152	0,245	0,022	0,096	1,89	0,6537
33	Dichlormethan	0,39	0,57	0,1	0,05	2,019	0,277	0,391	0,003	0,057	1,856	0,4943
34	1,2-Dichlorethan	0,42	0,64	0,1	0,11	2,573	0,306	0,441	0,003	0,084	2,52	0,6352
35	1,1-Dichlorethan	0,32	0,49	0,1	0,1	2,316	0,262	0,352	0,053	0,075	2,127	0,6352
36	1,2-Dichlorpropan	0,37	0,63	0,1	0,17	2,836	0,291	0,402	0,022	0,102	2,791	0,7761
37	Trichlormethan	0,43	0,49	0,15	0,02	2,48	0,406	0,559	0,053	0,082	2,586	0,6167
38	1,1,2-Trichlorethan	0,5	0,68	0,13	0,13	3,29	0,413	0,509	0,053	0,094	3,082	0,7576
39	1,1,1-Trichlorethan	0,37	0,41	0	0,09	2,733	0,4	0,519	0,003	0,115	2,901	0,7576
40	Tetrachlormethan	0,46	0,38	0	0	2,823	0,629	0,976	0,003	0,192	3,915	0,7391
41	1,1,1,2-Tetrachlorethan	0,54	0,63	0,1	0,08	3,641	0,563	0,676	0,003	0,147	3,91	0,88
42	1,1,2,2-Tetrachlorethan	0,6	0,76	0,16	0,12	3,803	0,508	0,577	0,103	0,091	3,59	0,88
43	Pentachlorethan	0,65	0,66	0,17	0,06	4,267	0,658	0,744	0,053	0,144	4,418	1,0024
44	Hexachlorethan	0,68	0,68	0	0	4,718	0,796	0,911	0,003	0,184	5,192	1,1248
45	1-Brompropan	0,37	0,4	0	0,12	2,62	0,34	0,363	0,003	0,067	2,546	0,7063
46	2-Brompropan	0,33	0,35	0	0,14	2,39	0,325	0,324	0,022	0,085	2,318	0,7063
47	Dibrommethan	0,71	0,69	0,11	0,07	2,886	0,623	0,549	0,003	0,035	2,712	0,5995
48	1,2-Dibromethan	0,75	0,76	0,1	0,17	3,382	0,652	0,599	0,003	0,062	3,376	0,7404
49	Fluorchlormethan	0,04	0,61	0,07	0,04	0,982	0,074	0,267	0,003	0,021	1,044	0,3896
50	1-Chlor-2,2,2-trifluorethan	0,01	0,4	0,15	0	1,168	-0,063	0,254	0,003	0,025	1,363	0,5659
51	Tefluran	-0,07	0,21	0,2	0,02	1,37	0,036	0,377	0,053	0,003	1,709	0,6362
52	Halopropan	-0,07	0,28	0,2	0	2,03	0,014	0,277	0,053	-0,012	2,04	0,7771
53	Bromochloromethan	0,54	0,8	0,01	0,06	2,445	0,45	0,47	0,003	0,046	2,284	0,5469
54	1-Bromo-2-chloroethan	0,57	0,7	0,1	0,09	2,982	0,479	0,52	0,003	0,073	2,948	0,6878
55	Halothan	0,1	0,39	0,13	0,05	1,953	0,239	0,501	0,053	0,039	2,521	0,7409
56	Chlordibrommethan	0,78	0,68	0,12	0,1	3,304	0,752	0,717	0,053	0,06	3,442	0,7219
57	Vinylchlorid	0,26	0,38	0	0,05	1,404	0,294	0,324	0,003	0,101	1,512	0,4698
58	Allylchlorid	0,33	0,56	0	0,05	2,109	0,294	0,324	0,003	0,101	2,011	0,6107
59	<i>cis</i> -1,2-Dichlorethylen	0,44	0,61	0,11	0,05	2,439	0,43	0,541	0,003	0,093	2,406	0,5922
60	<i>trans</i> -1,2-Dichlorethylen	0,43	0,41	0,09	0,05	2,278	0,43	0,541	0,003	0,093	2,406	0,5922
61	1,1-Dichlorethylen	0,36	0,34	0	0,05	2,11	0,429	0,457	0,003	0,09	2,336	0,5922
62	Trichlorethen	0,52	0,37	0,08	0,03	2,997	0,565	0,674	0,003	0,082	3,23	0,7146
63	Tetrachlorethen	0,64	0,44	0	0	3,584	0,7	0,807	0,003	0,071	4,054	0,837

## Anhang

Nr.	Name	exp.					ber.					
		<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>V</i>
64	1-Chlor-2,2-difluorethen	-0,34	0,285	0,15	0	0,723	0,159	0,426	0,003	0,01	1,606	0,5052
65	Chlorbenzol	0,72	0,65	0	0,07	3,657	0,791	0,71	0,003	0,126	3,768	0,8388
66	Methanol	0,28	0,44	0,43	0,47	0,97	0,205	0,449	0,348	0,385	1,123	0,3082
67	Ethanol	0,25	0,42	0,37	0,48	1,485	0,205	0,449	0,348	0,385	1,622	0,4491
68	1-Propanol	0,24	0,42	0,37	0,48	2,031	0,205	0,449	0,348	0,385	2,121	0,59
69	1-Butanol	0,22	0,42	0,37	0,48	2,601	0,205	0,449	0,348	0,385	2,62	0,7309
70	1-Pentanol	0,22	0,42	0,37	0,48	3,106	0,205	0,449	0,348	0,385	3,119	0,8718
71	2-Methyl-1-propanol	0,22	0,39	0,37	0,48	2,413	0,19	0,41	0,348	0,403	2,392	0,7309
72	3-Methyl-1-butanol (Isopentanol)	0,19	0,39	0,37	0,48	3,011	0,19	0,41	0,348	0,403	2,891	0,8718
73	2-Propanol	0,21	0,36	0,33	0,56	1,764	0,19	0,41	0,348	0,403	1,893	0,59
74	Diethylether	0,04	0,25	0	0,45	2,015	0,054	0,312	0,003	0,296	2,13	0,7309
75	Divinylether	0,26	0,39	0	0,13	1,76	0,308	0,392	0,003	0,342	1,916	0,6449
76	Propanon	0,18	0,7	0,04	0,49	1,696	0,179	0,598	0,003	0,419	1,891	0,547
77	Butan-2-on	0,17	0,7	0	0,51	2,287	0,179	0,598	0,003	0,419	2,39	0,6879
78	2-Pentanon	0,14	0,68	0	0,51	2,755	0,179	0,598	0,003	0,419	2,889	0,8288
79	2-Heptanon	0,12	0,68	0	0,51	3,76	0,179	0,598	0,003	0,419	3,887	1,1106
80	4-Methyl-2-pentanon	0,11	0,65	0	0,51	3,089	0,164	0,559	0,003	0,437	3,16	0,9697
81	Methylacetat	0,14	0,64	0	0,45	1,911	0,08	0,558	0,003	0,424	1,861	0,6057
82	Ethylacetat	0,11	0,62	0	0,45	2,314	0,08	0,558	0,003	0,424	2,36	0,7466
83	<i>n</i> -Propylacetat	0,09	0,6	0	0,45	2,819	0,08	0,558	0,003	0,424	2,859	0,8875
84	<i>n</i> -Butylacetat	0,07	0,6	0	0,45	3,353	0,08	0,558	0,003	0,424	3,358	1,0284
85	<i>n</i> -Pentylacetat	0,07	0,6	0	0,45	3,844	0,08	0,558	0,003	0,424	3,857	1,1693
86	Isopropylacetat	0,06	0,57	0	0,47	2,546	0,065	0,519	0,003	0,442	2,631	0,8875
87	Isobutylacetat	0,05	0,57	0	0,47	3,161	0,065	0,519	0,003	0,442	3,13	1,0284
88	Isopentylacetat	0,05	0,57	0	0,47	3,74	0,065	0,519	0,003	0,442	3,629	1,1693
89	Desfluran	-0,47	0,38	0,05	0,04	0,99	-0,451	0,253	0,103	0,136	0,913	0,6962
90	Enfluran	-0,24	0,54	0,01	0,1	2,009	-0,277	0,327	0,103	0,145	1,56	0,8009
91	Isofluran	-0,24	0,56	0,095	0,08	1,969	-0,248	0,377	0,103	0,172	1,725	0,8009
92	Methoxyfluran	0,11	0,67	0,07	0,14	2,864	0,11	0,474	0,053	0,263	2,641	0,8702
93	Sevofluran	-0,47	0,56	0	0,1	1,502	-0,48	0,296	0,053	0,179	1,536	0,8548
94	Fluroxen	0,18	0,3	0	0,1	1,6	-0,068	0,372	0,003	0,241	1,716	0,741
95	1-Nitropropan	0,24	0,95	0	0,31	2,894	0,262	0,942	0,003	0,27	2,717	0,7055

## Anhang

Nr.	Name	exp.					ber.					
		<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>E</i>	<i>S</i>	<i>A</i>	<i>B<sup>H</sup></i>	<i>L</i>	<i>V</i>
96	2-Nitropropan	0,22	0,92	0	0,33	2,55	0,247	0,903	0,022	0,288	2,489	0,7055

## Anhang

### 8.2.2 Werte aller zur $K_{fa}$ -Modellierung verwendeten Verteilungskoeffizienten

**Tabelle 8.8: Alle Verteilungskoeffizienten mit Werten, die zur Modellierung von  $K_{fa}$  verwendet werden.**

Nr.	Name	$\log K_{wa}$	$\log K_{oa}$	$\log K_{mw}$	$\log K_{olive}$	$\log K_{HSA}$	$\log K_{HSA/a}$
1	Pentan	-1,758	1,682	2,365	-1,929	-0,143	-1,901
2	Hexan	-1,880	2,033	2,774	-2,364	0,086	-1,794
3	Heptan	-2,002	2,747	3,183	-2,800	0,314	-1,688
4	Oktan	-2,123	3,062	3,591	-3,236	0,542	-1,581
5	<i>n</i> -Dekan	-2,365	2,687	4,408	-4,107	0,999	-1,367
6	2-Methylpentan	-1,880	1,365	2,734	-2,136	0,052	-1,828
7	3-Methylpentan	-1,880	1,763	2,734	-2,136	0,052	-1,828
8	3-Methylhexan	-2,002	1,884	3,143	-2,571	0,280	-1,722
9	2,3,4-Trimethylpentan	-2,123	2,190	3,472	-2,549	0,441	-1,682
10	2,2-Dimethylbutan	-1,880	2,027	2,656	-1,942	-0,004	-1,883
11	2,2,4-Trimethylpentan	-2,123	1,996	3,433	-2,585	0,419	-1,704
12	Cyclopropan	-0,687	0,211	1,343	-1,444	-0,631	-1,318
13	Cyclohexan	-1,056	2,652	2,568	-2,751	0,054	-1,002
14	Methylcyclopentan	-1,056	2,199	2,529	-2,522	0,020	-1,035
15	JP-10	-0,837	2,791	2,283	-5,590	0,106	-0,731
16	Ethen	-0,640	0,161	1,055	-0,440	-0,840	-1,480
17	2-Methyl-1,3-butadien	-0,736	1,924	2,167	-1,703	-0,170	-0,906
18	Benzol	0,619	2,774	2,114	-2,748	-0,059	0,560
19	Toluol	0,576	3,296	2,559	-3,141	0,193	0,769
20	Ethylbenzen	0,453	3,642	2,968	-3,577	0,421	0,874
21	1,3-Dimethylbenzen ( <i>m</i> -Xylol)	0,534	3,732	3,005	-3,534	0,444	0,978
22	<i>o</i> -Xylen	0,534	3,794	3,005	-3,534	0,444	0,978
23	<i>p</i> -Xylen	0,534	3,700	3,005	-3,534	0,444	0,978
24	Styrol	0,910	3,899	2,850	-3,485	0,396	1,305
25	<i>m</i> -Methylstyrol	0,866	4,344	3,296	-3,878	0,647	1,513
26	<i>p</i> -Methylstyrol	0,866	4,329	3,296	-3,878	0,647	1,513
27	Allylbenzen	0,458	3,726	3,259	-3,921	0,624	1,082
28	Difluormethan	-1,114	-0,877	0,989	-0,207	-0,946	-2,060
29	Methylchlorid	0,437	1,353	1,099	-1,028	-0,796	-0,359
30	Chlorethan	0,313	1,773	1,508	-1,464	-0,568	-0,255

## Anhang

Nr.	Name	$\log K_{wa}$	$\log K_{oa}$	$\log K_{mw}$	$\log K_{olive}$	$\log K_{HSA}$	$\log K_{HSA/a}$
31	1-Chlorpropan	0,189	2,311	1,917	-1,900	-0,339	-0,150
32	2-Chlorpropan	0,189	2,045	1,887	-1,702	-0,370	-0,181
33	Dichlormethan	0,390	2,127	1,534	-1,735	-0,505	-0,115
34	1,2-Dichlorethan	0,268	2,797	1,846	-2,350	-0,324	-0,056
35	1,1-Dichlorethan	0,268	2,429	1,931	-2,025	-0,301	-0,033
36	1,2-Dichlorpropan	0,144	2,918	2,226	-2,588	-0,126	0,018
37	Trichlormethan	0,843	2,794	1,825	-2,566	-0,301	0,541
38	1,1,2-Trichlorethan	0,720	3,363	2,239	-2,955	-0,073	0,647
39	1,1,1-Trichlorethan	0,720	2,643	2,127	-2,726	-0,138	0,583
40	Tetrachlormethan	-0,054	2,778	1,690	-3,945	-0,296	-0,350
41	1,1,1,2-Tetrachlorethan	1,172	3,921	2,405	-3,700	0,075	1,246
42	1,1,2,2-Tetrachlorethan	1,172	4,214	2,662	-3,515	0,194	1,365
43	Pentachlorethan	1,626	4,321	2,828	-4,261	0,341	1,968
44	Hexachlorethan	0,729	4,939	3,024	-4,962	0,505	1,234
45	1-Brompropan	0,174	2,624	2,192	-2,297	-0,119	0,055
46	2-Brompropan	0,174	2,487	2,163	-2,100	-0,149	0,025
47	Dibrommethan	1,360	3,174	2,085	-2,530	-0,064	1,296
48	1,2-Dibromethan	1,237	3,536	2,397	-3,145	0,117	1,353
49	Fluorchlormethan	-0,362	1,080	1,262	-0,971	-0,725	-1,087
50	1-Chlor-2,2,2-trifluorethan	-1,082	0,945	1,654	-1,273	-0,555	-1,637
51	Tefluran	-0,897	1,580	1,947	-1,738	-0,352	-1,249
52	Halopropan	-1,018	1,459	2,458	-1,947	-0,076	-1,094
53	Bromochloromethan	0,875	2,634	1,809	-2,132	-0,284	0,591
54	1-Bromo-2-chloroethan	0,752	3,350	2,122	-2,748	-0,103	0,648
55	Haloethan	-0,145	2,381	2,219	-2,502	-0,131	-0,276
56	Chlordibrommethan	1,813	3,655	2,375	-3,362	0,139	1,952
57	Vinylchlorid	-0,297	1,564	1,390	-1,373	-0,593	-0,890
58	Allylchlorid	-0,182	2,277	1,799	-1,808	-0,365	-0,547
59	<i>cis</i> -1,2-Dichlorethylen	-0,153	2,868	1,725	-2,305	-0,345	-0,499
60	<i>trans</i> -1,2-Dichlorethylen	-0,153	2,868	1,725	-2,305	-0,345	-0,499
61	1,1-Dichlorethylen	-0,153	2,102	1,796	-2,172	-0,310	-0,463
62	Trichlorethen	-0,011	2,815	2,131	-3,104	-0,063	-0,074
63	Tetrachlorethen	0,133	3,540	2,536	-3,904	0,220	0,353

## Anhang

Nr.	Name	$\log K_{wa}$	$\log K_{oa}$	$\log K_{mw}$	$\log K_{olive}$	$\log K_{HSA}$	$\log K_{HSA/a}$
64	1-Chlor-2,2-difluorethen	-0,861	0,776	1,586	-1,577	-0,504	-1,365
65	Chlorbenzol	0,750	3,736	2,519	-3,547	0,224	0,974
66	Methanol	3,720	2,960	0,063	-1,734	-1,470	2,250
67	Ethanol	3,597	3,379	0,472	-2,170	-1,242	2,355
68	1-Propanol	3,474	3,769	0,881	-2,606	-1,014	2,461
69	1-Butanol	3,351	4,323	1,289	-3,041	-0,785	2,566
70	1-Pentanol	3,227	4,785	1,698	-3,477	-0,557	2,670
71	2-Methyl-1-propanol	3,351	4,158	1,249	-2,812	-0,819	2,532
72	3-Methyl-1-butanol (Isopentanol)	3,227	4,399	1,658	-3,248	-0,591	2,636
73	2-Propanol	3,474	3,530	0,841	-2,377	-1,047	2,427
74	Diethylether	1,169	2,189	1,346	-1,963	-0,747	0,422
75	Divinylether	-0,230	2,163	1,110	-1,780	-0,797	-1,027
76	Propanon	2,655	2,604	0,325	-1,968	-1,287	1,368
77	Butan-2-on	2,532	2,923	0,733	-2,404	-1,059	1,473
78	2-Pentanon	2,410	3,376	1,142	-2,839	-0,831	1,579
79	2-Heptanon	2,163	4,141	1,959	-3,711	-0,374	1,789
80	4-Methyl-2-pentanon	2,286	3,559	1,511	-3,046	-0,636	1,650
81	Methylacetat	2,108	2,508	0,425	-1,933	-1,269	0,839
82	Ethylacetat	1,983	2,991	0,833	-2,368	-1,041	0,943
83	<i>n</i> -Propylacetat	1,861	3,290	1,242	-2,804	-0,812	1,048
84	<i>n</i> -Butylacetat	1,738	3,720	1,651	-3,240	-0,584	1,154
85	<i>n</i> -Pentylacetat	1,614	4,100	2,059	-3,675	-0,356	1,258
86	Isopropylacetat	1,861	3,224	1,202	-2,575	-0,846	1,015
87	Isobutylacetat	1,738	3,511	1,611	-3,011	-0,618	1,120
88	Isopentylacetat	1,614	3,870	2,019	-3,446	-0,389	1,225
89	Desfluran	-0,501	0,737	1,413	-1,144	-0,862	-1,363
90	Enfluran	0,250	2,388	1,782	-1,728	-0,595	-0,344
91	Isofluran	0,250	1,991	1,686	-1,907	-0,642	-0,392
92	Methoxyfluran	1,601	3,030	1,808	-2,617	-0,459	1,142
93	Sevofluran	-0,921	0,867	1,655	-1,649	-0,738	-1,659
94	Fluroxen	-0,427	0,941	1,398	-1,684	-0,743	-1,170
95	1-Nitropropan	2,480	3,319	1,064	-2,964	-0,793	1,687
96	2-Nitropropan	2,480	3,243	1,034	-2,766	-0,823	1,657