Abstract: The prediction of tissue-to-plasma water partition coefficients (Kpu) from in vitro and in silico data using the tissue-composition based model (Rodgers & Rowland, J Pharm Sci. 2005, 94(6):1237-48.) is well established. However, distribution of basic drugs, in particular into lysosome-rich lung tissue, tends to be under-predicted by this approach. The aim of this study was to develop an extended mechanistic model for the prediction of Kpu which accounts for lysosomal sequestration and the contribution of different cell types in the tissue of interest. The extended model is based on compound-specific physicochemical properties and tissue composition data to describe drug ionization, distribution into tissue water and drug binding to neutral lipids, neutral phospholipids and acidic phospholipids in tissues, including lysosomes. Physiological data on the types of cells contributing to lung, kidney and liver, their lysosomal content and lysosomal pH were collated from the literature. The predictive power of the extended mechanistic model was evaluated using a dataset of 28 basic drugs (pKa ≥ 7.8, 17 β-blockers, 11 structurally diverse drugs) for which experimentally determined Kpu data in rat tissue have been reported. Accounting for the lysosomal sequestration in the extended mechanistic model improved the accuracy of Kpu predictions in lung compared to the original Rodgers model (56% drugs within 2-fold or 88% within 3-fold of observed values). Reduction in the extent of Kpu under-prediction was also evident in liver and kidney. However, consideration of lysosomal sequestration increased the occurrence of over-predictions, yielding overall comparable model performances for kidney and liver, with 68% and 54% of Kpu values within 2-fold error, respectively. High lysosomal concentration ratios relative to cytosol (>1000-fold) were predicted for the drugs investigated; the extent differed depending on the lysosomal pH and concentration of acidic phospholipids among cell types. Despite this extensive lysosomal sequestration in the individual cells types, the maximal change in the overall predicted tissue Kpu was <3-fold for lysosome-rich tissues investigated here. Accounting for the variability in cellular physiological model input parameters, in particular lysosomal...
fraction of the cellular volume occupied by the lysosomes, only partially explained discrepancies between observed and predicted Kpu data in the lung. Improved understanding of the system properties, e.g., cell/organelle composition is required to support further development of mechanistic equations for the prediction of drug tissue distribution. Application of this revised mechanistic model is recommended for prediction of Kpu in lysosome-rich tissue to facilitate the advancement of physiologically-based prediction of volume of distribution and drug exposure in the tissues.
Incorporation of lysosomal sequestration in the mechanistic model for prediction of tissue distribution of basic drugs

Frauke Assmus, J. Brian Houston and Aleksandra Galetin

Centre for Applied Pharmacokinetic Research, School of Health Sciences, University of Manchester, Stopford Building, Oxford Road, M13 9PT Manchester, United Kingdom

*To whom correspondence should be addressed:

Dr Aleksandra Galetin
Centre for Applied Pharmacokinetic Research
School of Health Sciences
The University of Manchester, Stopford Building
Oxford Road, Manchester, M13 9PT, UK
Tel: (+) 44 161 275 6886
Email: Aleksandra.Galetin@manchester.ac.uk

Abstract: 402
Intro: 866
Discussion: 1838
References: 58

KEYWORDS

Prediction of tissue distribution; lysosomal sequestration; basic drugs; PBPK modelling
ABSTRACT

The prediction of tissue-to-plasma water partition coefficients ($K_{pu}$) from *in vitro* and *in silico* data using the tissue-composition based model (Rodgers & Rowland, J Pharm Sci. 2005, 94(6):1237-48.) is well established. However, distribution of basic drugs, in particular into lysosome-rich lung tissue, tends to be under-predicted by this approach. The aim of this study was to develop an extended mechanistic model for the prediction of $K_{pu}$ which accounts for lysosomal sequestration and the contribution of different cell types in the tissue of interest. The extended model is based on compound-specific physicochemical properties and tissue composition data to describe drug ionization, distribution into tissue water and drug binding to neutral lipids, neutral phospholipids and acidic phospholipids in tissues, including lysosomes. Physiological data on the types of cells contributing to lung, kidney and liver, their lysosomal content and lysosomal pH were collated from the literature. The predictive power of the extended mechanistic model was evaluated using a dataset of 28 basic drugs ($pK_a \geq 7.8$, 17 β-blockers, 11 structurally diverse drugs) for which experimentally determined $K_{pu}$ data in rat tissue have been reported. Accounting for the lysosomal sequestration in the extended mechanistic model improved the accuracy of $K_{pu}$ predictions in lung compared to the original Rodgers model (56% drugs within 2-fold or 88% within 3-fold of observed values). Reduction in the extent of $K_{pu}$ under-prediction was also evident in liver and kidney. However, consideration of lysosomal sequestration increased the occurrence of over-predictions, yielding overall comparable model performances for kidney and liver, with 68% and 54% of $K_{pu}$ values within 2-fold error, respectively. High lysosomal concentration ratios relative to cytosol (>1000-fold) were predicted for the drugs investigated; the extent differed depending on the lysosomal pH and concentration of acidic phospholipids among cell types. Despite this extensive lysosomal sequestration in the individual cells types, the maximal change in the overall predicted tissue $K_{pu}$ was <3-fold for lysosome-rich tissues investigated here. Accounting for the variability in cellular physiological model input parameters, in particular lysosomal pH and fraction of the cellular volume occupied by the lysosomes, only partially explained discrepancies between observed and predicted $K_{pu}$ data in the lung. Improved understanding of the system properties, e.g., cell/organelle composition is required to support further development of mechanistic equations for the prediction of drug tissue distribution. Application of this revised mechanistic model is recommended for prediction of $K_{pu}$ in lysosome-rich tissue to facilitate
the advancement of physiologically-based prediction of volume of distribution and drug exposure in
the tissues.
1 Introduction

Tissue-to-plasma water partition coefficients ($K_{pu}$) are important drug-related distribution parameters in physiologically-based pharmacokinetic (PBPK) models [1-4]. The experimental determination of $K_{pu}$ requires in vivo studies under steady-state conditions which are laborious and time-intensive. Therefore, several in silico models for the prediction of $K_{pu}$ from drug-specific physicochemical properties and tissue composition data have been developed [5-9]. An initial model by Poulin et al. [5] described tissue distribution based on passive processes and accounted for drug dissolution in tissue water in addition to drug binding to neutral lipids, neutral phospholipids and macromolecules. Subsequently, Rodgers et al. incorporated the contribution of electrostatic interactions of moderate-to-strong bases ($pK_a > 7$) with acidic phospholipids [6, 10]; a separate mechanistic model was developed for the prediction of tissue distribution of acids, very weak bases, neutrals and zwitterions [7]. In the case of moderate-to-strong bases, only the nonionized drug was assumed to bind to neutral lipids and neutral phospholipids, whereas electrostatic interaction with acidic phospholipids was assumed to be the predominant factor for the lipid binding of cationic drug species [6]. The Rodgers & Rowland model improved accuracy in $K_{pu}$ predictions of 28 moderate-to-strong bases across a range of tissues [6]. Despite improved model performances relative to Poulin et al., under-prediction in $K_{pu, lung}$ prevailed and both under- and over-predictions of $K_{pu, brain}$ were evident [6]. The same trend in under-prediction of $K_{pu, lung}$ for lipophilic bases was confirmed in a comprehensive analysis of six tissue distribution prediction models applied to a set of 81 acidic, basic, neutral and zwitter-ionic compounds [11].

Disregarding lysosomal sequestration of basic drugs in the Rodgers & Rowland model [6] is one possible explanation for the prevailing under-prediction of $K_{pu, lung}$. Lysosomes are acidic organelles found in almost all animal and human cells where they act as the main digestive system [12]. They are abundant in lungs, liver, kidney and spleen, and in certain cell types such as macrophages where they may contribute even >10% of the cellular volume [13, 14]. Extensive accumulation of cationic amphiphilic drugs (CAD) in alveolar macrophages and other cell types has been reported and attributed to large extent to subcellular sequestration of basic drugs in lysosomes [13, 15-20]. This ‘trapping’ phenomenon is attributed to pH differences between the cytosol (pH ~ 7.2) and acidic lysosomes (pH ~ 4.5 - 5), as protonation in the lysosomal environment shifts the equilibrium of the
basic drug towards the ionized species [13, 21]. This extensive drug ionization in lysosomes and reduced permeability of ionized drug species across the lysosomal membrane may lead to potentially very high drug concentrations in this organelle relative to the cytosol. In addition to the pH gradient, binding of cationic drugs to phospholipids in lysosomal membrane contributes to this subcellular accumulation [13, 21]. Prolonged accumulation of CAD in lysosomes has been associated with increased risk of drug-induced phospholipidosis, an important safety concern for potential drug candidates [22, 23]. CAD-induced phospholipidosis is characterized by extensive accumulation of phospholipids in cells, formation of lamellar inclusion bodies, drug accumulation in association with increased phospholipids and reversibility after discontinuation of drug treatment [22]. Several mechanisms may be involved in the accumulation of phospholipids including direct or indirect inhibition of lysosomal phospholipases [22]. Potential additional implications of drug accumulation in lysosomes on either drug efficacy, drug resistance and/or drug-drug interactions at the level of lysosomes have all been highlighted [14, 24].

Methodological difficulties, e.g., the fragility of lysosomes during homogenization procedures, potential contamination of lysosomes and diffusion of drug from this organelle during the isolation procedure, have hampered direct measurements of lysosomal drug concentrations and determination of the relative importance of lysosomal sequestration in vitro [13]. Studies often rely on indirect assessment of the extent of lysosomal sequestration by i) using agents to abolish existing pH gradient between cytosol and lysosomes (e.g., ammonium chloride, monensin) [19, 20, 25] or ii) measuring cellular Kpu at low and high drug concentrations, assuming that the saturable component of the cellular uptake at a low drug concentration is attributed to the lysosomal sequestration [21]. Indirect methods neglect the effect of ionic strength and changes in membrane surface potential at higher drug concentrations due to insertion of cationic drugs into lipid bilayers (such as lysosomes), which may all affect drug binding to the membrane [26].

The aims of this study were to extend the existing Rodgers & Rowland mechanistic model [6] to account for lysosomal sequestration in the prediction of Kpu for moderate-to-strong bases, to investigate the impact of this process on Kpu predictions for lysosome-rich tissues and to assess whether consideration of lysosomal sequestration reduces the under-prediction trends previously
observed for basic drugs and $K_{pu,t_{long}}$ in particular. Physiological data on different cell types contributing to the overall tissue were collated for lung, kidney and liver, including also lysosomal volume fraction and lysosomal pH in individual cell types. Performances of the original and extended tissue-composition models were evaluated in a comparative manner using a dataset of 28 basic drugs employed for development of the original tissue distribution prediction tool [6]. The impact of variability in physiological model input parameters collated from the literature on $Kpu$ predictions was considered. The scope and limitations of the extended mechanistic model relative to the original Rodgers model and its application for the prediction of tissue distribution within physiologically-based framework are discussed.
2 Materials and Methods

2.1 Derivation of a mechanistic equation for the prediction of drug tissue distribution accounting for lysosomal sequestration

The tissue-to-plasma water partition coefficient, $K_{pu}$, is defined as the ratio of the total drug concentration in tissue, $C_{Tissue,Total}$, and the unbound drug concentration in plasma, $C_{u,p}$, at steady-state:

$$K_{pu} = \frac{C_{Tissue,Total}}{C_{u,p}}$$

(1)

For each tissue, $C_{Tissue,Total}$ is obtained by dividing the total amount of drug in tissue ($A_{Tissue,Total}$) by its volume ($V_{Tissue}$). $A_{Tissue,Total}$ can be expressed as the sum of amounts of drug in non-lysosomal compartments ($A_{Non-Lys}$) and in lysosomes ($A_{Lys}$), as shown in Equation 2.

$$C_{Tissue,Total} = \frac{A_{Tissue,Total}}{V_{Tissue}} = \frac{A_{Non-Lys} + A_{Lys}}{V_{Tissue}}$$

(2)

Expressing $A_{Lys}$ in terms of lysosomal volume ($V_{lys} = f_{lys} \cdot V_{Tissue}$) yields Equation 3, where $f_{lys}$ represents volume fraction of tissue occupied by lysosomes:

$$C_{Tissue,Total} = \frac{A_{Non-Lys}}{V_{Tissue}} + \frac{A_{Lys}}{V_{lys}} \cdot f_{lys}$$

(3)

or expressed in terms of concentrations:

$$C_{Tissue,Total} = C_{Non-Lys} + C_{Lys} \cdot f_{lys}$$

(4)

Rodgers et al. [6] have proposed a mechanistic equation for the prediction of the drug concentration in tissue and $K_{pu}$ for compounds with a basic $pK_a > 7$. Drug ionization and drug dissolution in intracellular and extracellular tissue water, in addition to drug binding to neutral lipids, neutral phospholipids and acidic phospholipids were taken into account. In the study by Rodgers et al., a distinction between the distribution of drug into cytosol and subcellular organelles, i.e., the contribution of lysosomal sequestration was not considered. Drug amounts and drug concentrations in tissue estimated by the original Rodgers model are referred to here as $A_{Tissue,R}$ and $C_{Tissue,R}$, respectively (Equations 5 and 6):
\[
C_{\text{Tissue,R}} = \frac{A_{\text{Tissue,R}}}{V_{\text{Tissue}}} = C_{\text{fR}} \cdot \left[ f_{\text{EW}} + \frac{1 + 10^{pK_a - p\text{pH}_{\text{EW}}}}{1 + 10^{pK_a - p\text{pH}_{\text{fR}}}} \cdot f_{\text{fR}} \right] \\
+ \left( P \cdot f_{\text{NL}} + \left( (0.3P + 0.7) \cdot f_{\text{NP}} \right) \right) \\
\]

\[
K_{\text{pu,R}} = \frac{C_{\text{Tissue,R}}}{C_{\text{fR}}} 
\]

where \( K_{\text{pu,R}} \) represents predicted \( K_{\text{pu}} \) values by the original Rodgers model. The terms \( f_{\text{EW}}, f_{\text{fW}}, f_{\text{NL}} \) and \( f_{\text{NP}} \) indicate the fractional tissue volumes of extracellular water, intracellular water (cytosol), neutral lipids and neutral phospholipids, respectively. \( P \) is the octanol/water partition coefficient and was used as a surrogate for binding of neutral drugs to neutral lipids whereas a weighted \( P \) was applied to reflect the affinity of neutral drugs to neutral phospholipids. More precisely, a hydrophilic-lipophilic balance of phospholipids similar to a mixture of 30% neutral lipids and 70% water was assumed. Electrostatic interaction of cationic drug with acidic phospholipids was described in terms of the association constant of basic compounds with acidic phospholipids (\( K_a \)) along with the concentration of acidic phospholipids in different tissues (\( \left| AP^+ \right|_f \)). Based on the assumption that \( K_a \) measured in red blood cells reflects \( K_a \) in all other tissues, \( K_a \) can be estimated from \( K_{\text{pu,BC}} \) (the blood cell-to-plasma water concentration ratio). Drug ionization and pH partitioning were described according to the \( pK_a \) and pH differences in extracellular water (\( p\text{H}_{\text{EW}} \)), plasma (\( p\text{H}_{\text{fR}} \), \( p\text{H}_{\text{fP}} \approx p\text{H}_{\text{EW}} \)) and intracellular water (cytosol, \( p\text{H}_{\text{fP}} \)). For further details about the model, the reader is referred to the literature [6] and the Supplementary Material.

The total amount of drug in tissue predicted by Rodgers et al. (\( A_{\text{Tissue,R}} \)) can be used as a surrogate for \( A_{\text{Non}-\text{Lys}} \) in Equation 3, resulting in the following relationship:

\[
C_{\text{Tissue,Total}} = \frac{A_{\text{Tissue,R}}}{V_{\text{Tissue}}} - \frac{A_{\text{Tissue,R}}}{V_{\text{Lys}}} \cdot f_{\text{fLys}} + \frac{A_{\text{Lys}}}{V_{\text{Lys}}} \cdot f_{\text{fLys}} 
\]

or in terms of drug concentrations:

\[
C_{\text{Tissue,Total}} = C_{\text{Tissue,R}} + f_{\text{fLys}} \cdot (C_{\text{Lys}} - C_{\text{Tissue,R}}) 
\]
Provided that $C_{\text{Lys}} \gg C_{\text{Tissue,R}}$, which is true for moderate-to-strong-bases accumulating in the acidic lysosomal environment, equation 8 simplifies to:

$$C_{\text{Tissue,Total}} = C_{\text{Tissue,R}} + f_{\text{Lys}} \cdot C_{\text{Lys}}$$  \hspace{1cm} (9)

Substitution of Equation 9 into Equation 1 yields:

$$Kpu^*_R = \frac{C_{\text{Tissue,R}} + f_{\text{Lys}} \cdot C_{\text{Lys}}}{C_{\text{U},P}} = Kpu_R + \frac{C_{\text{Lys}}}{C_{\text{U},P}} \cdot f_{\text{Lys}}$$  \hspace{1cm} (10)

Where $Kpu^*_R$ refers to the revised predicted tissue-to-plasma water partition coefficients that accounts for drug distribution into lysosomes.

Lysosome-to-intracellular water partition coefficient ($Kpu_{\text{Lys}}$) is introduced in order to extend the original Rodgers model by a term reflecting lysosomal sequestration; it is defined as the ratio of the total lysosomal drug concentration ($C_{\text{Lys}}$) and the unbound concentration in the cytosol ($C_{\text{U},P}$) at steady state:

$$Kpu_{\text{Lys}} = \frac{C_{\text{Lys}}}{C_{\text{U},P}}$$  \hspace{1cm} (11)

Model assumptions adapted from the original Rodgers model and applied to the lysosomes are described below. Drug distribution between lysosome and cytosol is described in terms of drug ionization as well as drug dissolution in lysosomal and cytosolic intracellular water. In addition, drug binding to lysosomal lipids, i.e., neutral lipids, neutral phospholipids and acidic phospholipids, is considered. $P$ is used as a surrogate for the binding of neutral drugs to neutral lipids and a weighted $P$ (corresponding to the suggested hydrophilic-lipophilic balance of neutral phospholipids) is applied to reflect the affinity of neutral drugs to neutral phospholipids. Only the nonionized drug binds to neutral lipids and neutral phospholipids within the lysosome, whereas electrostatic interactions with acidic, lysosomal phospholipids govern the binding of cationic drug species. Only the nonionized species is assumed to cross the lysosomal membrane by passive diffusion, and equilibrium is assumed between the unbound nonionized species in the aqueous phase on each side of the lysosomal membrane. Drug ionization and pH partitioning are determined by the $pK_a$ of the drug and pH differences between cytosol containing intracellular water ($pH_{\text{U},P}$) and lysosomes ($pH_{\text{Lys}}$).
Relative water ($f_{IW}$) and lipid volume fractions ($f_{NL}, f_{NP}$) as well as the concentration of residual tissue components ($[AP^-]_{REM}$) in the lysosome were assumed to be equal to $f_{IW}, f_{NL}, f_{NP}$ and $[AP^-]_{REM}$ in non-lysosomal compartments of tissue (plasma and other membranes). Given the dominant electrostatic binding of moderate-to-strong bases to acidic phospholipids, $[AP^-]_{REM}$ was taken to be the total acidic phospholipid concentration (as in Rodgers et al.). Since lysosomes do not contain an extraorganelle compartment, the $f_{EW}$ was set to 0. Accordingly, $Kpu_{Lys}$ is given by:

$$Kpu_{Lys} = \frac{C_{Lys}}{C_{UW}} = \frac{1 + 10^{pK_a - pH_{Lys}}}{1 + 10^{pK_a - pH_{Lys}}} \cdot f_{IW} + \frac{K_a \cdot [AP^-]_T \cdot 10^{pK_a - pH_{Lys}}}{1 + 10^{pK_a - pH_{Lys}}} + \left( P \cdot f_{NL} + (0.3P + 0.7) \cdot f_{NP} \right)$$

Due to the acidic environment of lysosomal water, more pronounced pH differences between lysosome and cytosol (as compared to pH differences between cytosol and plasma, Equation 5) are driving the magnitude of $Kpu_{Lys}$. The concentration in cytosol ($C_{UW}$) is thereby the reference concentration for the calculation of $Kpu_{Lys}$. The concentration in cytosol ($C_{UW}$) can also be expressed in terms of $C_{U}$ according to:

$$C_{U,W} = C_U \cdot \left( \frac{1 + 10^{pK_a - pH_{Lys}}}{1 + 10^{pK_a - pH_{P}}} \right)$$

Equation 13 assumes that only the unbound unionised species crosses the plasma membrane, and that concentrations of unbound unionised drug in the aqueous phase on each side of the cell membrane are equal at equilibrium (for details see Supplementary Material). Substitution of Equation 13 into Equation 10 yields:

$$Kpu_R^* = Kpu_R + \left( \frac{1 + 10^{pK_a - pH_{Lys}}}{1 + 10^{pK_a - pH_{P}}} \right) \cdot \frac{C_{Lys}}{C_{UW}} \cdot f_{Lys}$$

or in terms of $Kpu_R$ and $Kpu_{Lys}$ (substitution of Equation 11 in Equation 14):

$$Kpu_R^* = Kpu_R + \left( \frac{1 + 10^{pK_a - pH_{Lys}}}{1 + 10^{pK_a - pH_{P}}} \right) \cdot Kpu_{Lys} \cdot f_{Lys}$$
Each tissue is comprised of various cell types that may differ in terms of \( f_{\text{Lys}} \) and \( pH_{\text{Lys}} \), both governing the extent of lysosomal sequestration. Hence, \( K_{\text{pu}}^* \) for a tissue comprised of different cell types can be expressed as:

\[
K_{\text{pu}}^* = K_{\text{pu}} + \sum_{\text{Cell Type}} \left( \frac{1 + 10^{p_{K_p} - pH_{pu}}}{1 + 10^{p_{K_p} - pH_{pu}}/10} \right) \cdot K_{\text{pu}} \cdot f_{\text{Lys}} \cdot f_{\text{Cell Type}}
\]

(16)

where \( K_{\text{pu}} \) is weighted by \( f_{\text{Lys}} \) in a particular cell type and by the fraction of that cell type in tissue (\( f_{\text{Cell Type}} \)). An illustration of the extended Rodgers model on the cell and tissue level using lung as an example is shown in Figure 1 and 2, respectively. Illustrations of the cell types considered for the extended Rodgers model for the kidney and liver are provided in the Supplementary Material (Figure S1).

2.2 Tissue-specific input parameters

Values for \( pH_p \) (\( pH_p = 7.4 \)) and \( pH_{pw} \) (\( pH_{pw} = 7.0 \)) as well as tissue-specific values for \( f_{\text{EW}} \), \( f_{\text{PW}} \), \( f_{\text{NL}} \), \( f_{\text{NP}} \) and \( [AP^-]_T \) in rat lung, liver and kidney were taken from [6]. The composition of the lysosomal membranes was assumed to be the same as in non-lysosomal compartments of the cell (plasma and membranes of other organelles). In addition, a critical literature review was conducted to collate physiological data for \( f_{\text{Cell Type}} \), \( f_{\text{Lys}} \) and \( pH_{\text{Lys}} \) in the tissues of interest, as summarized in Table 1.

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
</table>

In the case of lung tissue, cells with lysosomal origin (alveolar macrophages, type II cells) were considered separately from other residual cells (type I cells, interstitial cells, interstitial matrix, endothelial cells) [27]. Measured data for the lysosomal contribution to the cellular volume has been reported in rat alveolar macrophages \((3.7 - 13.9\%\), average \( f_{\text{Lys(AM)}} = 7.8\% \) [14], whereas only enzymatic studies of the acidic phosphatase activity (indirect marker for \( f_{\text{Lys}} \)) were available for type II cells [27-29]. Based on the activity of the lysosomal marker enzyme in type II cells relative to
alveolar macrophages [27], a lysosomal volume fraction \( f_{\text{Lys}(\text{Type II})} \) of 3% was assumed for the type II cells. With respect to the residual cells, \( f_{\text{Lys(Res.Cells)}} \) was set to 1% based on the consideration that a certain level of lysosomes in most cells is necessary for maintaining the digestive functions of a cell. Experimental data for lysosomal pH values were available for alveolar macrophages \( (pH_{\text{Lys(AM)}} = 4.75 \) (mouse; data in rat not available) [30]), and for type II cells \( (pH_{\text{Lys(Type II)}} = 5.1 \) (rat lamellar bodies), [31]), but not for residual cells. Therefore, the same lysosomal pH as in type II cells was assumed for residual cells.

Kidney tissue is extremely heterogeneous and data on lysosomal fractions in different kidney cells comprising various morphological units was limited; most data found on \( f_{\text{Lys(Kidney)}} \) in different cell types was reported as a sum of primary and secondary lysosomes, including hetero- and autophagosomes [32] which all have different properties (e.g., higher intra-vesicular pH) compared to primary lysosomes. The confidence in such \( f_{\text{Lys(Kidney)}} \) data was minimal and therefore the mean value of the lysosomal volume fraction in total lung (1.28 %) and liver (1.12%) was applied, resulting in the \( f_{\text{Lys(Kidney)}} \) of 1.2% (Supplementary Material, Section S2). For the purpose of this analysis, it was assumed that different cell types in the kidney show no difference in the lysosomal volume fraction. Due to the lack of experimentally determined lysosomal pH data in rat kidney, the average value \( (pH_{\text{Lys(Kidney)}} = 5.1) \) measured in different kidney-derived cell systems was used (e.g. MDCK cells [33], cultured human proximal tubular cells [34], for all references see Table 1). For consistency, only lysosomal pH values measured with the well-established FITC-Dextran method [30] were included.

With respect to physiological input data for liver, only liver parenchym (84.1% of total liver) was considered for the calculation of \( K_{\text{pu,Lys}} \), whereas intercellular spaces (Disse space, sinusoidal lumen, biliary canaliculi) were not taken into account [35]. Hence, predicted \( K_{\text{pu,Lys}} \) values were weighted by a factor of 0.841 (Supplementary Material, Section S2). Experimental data for lysosomal pH values in rat hepatocytes \( (pH_{\text{Lys(Hept)}} = 4.72) \) and rat Kupffer cells \( (pH_{\text{Lys(Kupffer)}} = 4.94) \) as well as lysosomal volume fractions in liver cells were obtained from literature (corresponding references listed in Table
1). Since no measured pH$_{Lys}$ values in endothelial and fat-storing cells were available, a value of pH$_{Lys}$ = 5.1 was assumed for these cells. Only lysosomal pH values measured in intact cells were included since pH$_{Lys}$ data obtained in isolated lysosomes tend to overestimate the lysosomal pH (data not shown).

2.3 Selection of the dataset and compound-specific input parameters

The measurement of Kpu requires extensive in vivo studies with tissue collection and published data for experimentally determined Kpu in different tissues are sparse. Rodgers et al. have collated experimental Kpu data in lung, kidney, liver and other tissues from 11 studies, along with $K_a$ data required for application of the model [6]. To validate the extended mechanistic model and to facilitate comparative evaluation with the original Rodgers model, we have used the same set of 28 basic compounds ($pK_a > 7$) previously published by Rodgers et al. It is worth noting that Graham et al. [11] have compared different Kpu prediction models using a 81 compound dataset which contained 8 additional moderate-to-strong bases ($pK_a \geq 7$). However, none of these additional 8 compounds had Kpu measurements in all of the three tissues investigated in our study (liver, lung, kidney) which made comparisons of model accuracies across different tissues difficult.

Compound-specific input parameters, i.e., measured log P and $pK_a$ data, were taken from Rodgers et al. [6], Table 2. For inaperisone, only predicted log P and $pK_a$ values were available and were taken as reported. $K_a$ values were derived from reported Kpu$_{BC}$ data and additional input parameters ($pH_{BC}, f_{PW,BC}, f_{NL,BC}, f_{NP,BC}, A_{P_{BC}}$) specific to red blood cells as described previously [6]. Where negative $K_a$ values were obtained (phenycyclidine and nicotine), the electrostatic interaction of these drugs with acidic phospholipids was assumed to be negligible and $K_a$ was set to zero.

2.4 Data analysis

The original (eq. 6) and extended (eq. 16) mechanistic models were used to predict Kpu$_R$ and Kpu*$_R$ values, respectively for the set of 28 drugs in lysosome-rich tissues in rat, i.e., lung, liver and kidney.
Model performance was assessed using a two-fold error margin as criterion; for consistency with previous analyses of $Kpu$ predictions [5, 6], percentage of under- and over-predictions within 3-fold was also assessed. In addition, the coefficient of determination ($r^2$, squared Pearson correlation coefficient), as derived from linear regression between observed and predicted $Kpu$ data, was calculated. Absolute deviations between observed and predicted data were calculated in terms of the geometric mean-fold error:

$$gmfe = 10^{\frac{1}{n} \sum_{i=1}^{n} \log_{10} \left( \frac{PRED_i}{OBS_i} \right)}$$

where $PRED_i$ and $OBS_i$ refer to the predicted and observed $Kpu$ value for the $i$th compound, respectively, and where $n$ is the total number of observations.

The impact of lysosomal sequestration on accumulation of basic drugs in lungs, liver and kidney was assessed by calculating the ratio between $Kpu_R$ and $Kpu_R^*$, thus providing a fold-change in predicted tissue distribution by accounting for this process. In the sensitivity analysis, fold changes in $Kpu_R^*$ were calculated from the ratio of $Kpu_R^*$ derived using the lowest and highest reported physiological $f_{Lys(AM)}$ estimate (3.7% and 13.9%, respectively). An analogous approach was applied to assess the impact of variability in $f_{Lys(Typ II)}$ and $pH_{Lys}$. Since experimental data for $f_{Lys(Typ II)}$ and $pH_{Lys(Typ II)}$ were missing, a range of $f_{Lys(Typ II)} = 3$- 6% and $pH_{Lys(Typ II)} = 4.7 – 5.4$ was assumed. The applied range for $pH_{Lys(AM)}$ (4.69–4.81) represents the standard deviation of the experimentally determined mean value for this parameter [30]. In addition, sensitivity analyses was performed to investigate the impact of individual physicochemical parameter on the predicted lung $Kpu_R^*$ by using the same $\log P$ and varying $pK_a$ ($pK_a = 7$-$11$), or vice versa, the same $pK_a$ and varying $\log P$ ($\log P = 1$-$5$), as detailed in the Supplementary Material Section 3.
3 Results

3.1 Characterization of the dataset

The 28 compound dataset used for model validation comprised of 17 beta-blockers and 11 additional, structurally diverse basic drugs ($pK_a = 7.8 - 10.0$) of intermediate to high lipophilicity ($\log P = 0.88 - 4.96$). Key physicochemical properties ($pK_a$, $\log P$) for all drugs in the dataset are listed in Table 2, along with estimated $Kpu_{BC}$ and $K_a$ data used as model input parameters. The dataset covered a broad range of experimental $Kpu$ values ($Kpu_{lung} = 2.5 - 1426$, $Kpu_{kidney} = 7 - 364$, $Kpu_{liver} = 3 - 327$) with the highest extent of drug tissue distribution observed in lung for R-propranolol.

Table 2

3.2 The prediction of drug distribution into lung

Lung tissue is comprised of various different cell types of which alveolar macrophages ($f_{AM} = 4.1\%$ of lung tissue) contain the highest lysosomal volume fraction (average $f_{Lys(AM)} = 7.8\%$ [14]). Similar to alveolar macrophages, type II cells are of lysosomal origin, but with lower contribution of lysosomes to the cellular volume compared to alveolar macrophages [27]. In the initial analysis, type II cells were combined with residual cells in the extended model resulting in $f_{Res.Cells} = 95.9\%$ with $f_{Lys(Res./TypeII Cells)} = 1\%$ (scenario ‘AM and residual cells’). Subsequently, $Kpu_R^*$ values were predicted by considering type II cells ($f_{TypeII} = 8.3\%$ with $f_{Lys(TypeII)}$ of 3%) separately to alveolar macrophages ($f_{AM} = 4.1\%$ with $f_{Lys(AM)} = 7.8\%$) and residual cells ($f_{Res.Cells} = 87.6\%$ with $f_{Lys(Res.Cells)} = 1\%$).

Predicted lung $Kpu_R$ and $Kpu_R^*$ values and summary of performances of different models are shown in Tables 2 and 3. Comparison of experimental and predicted $Kpu_{lung}$ obtained from the original and the extended Rodgers model (‘AM and residual cells only’) are shown in Figure 3A. The high tendency of all drugs in the dataset to accumulate in lysosomes was evident, as $Kpu_{Lys}$ in alveolar macrophages ranged from 69 (nicotine) to >4000 (propranolol), with $Kpu_{Lys} > 1000$ for 22/28 drugs. Similar trends were evident in $Kpu_{Lys}$ for residual and type II cells, although the overall extent of
accumulation relative to the cytosolic concentrations was less pronounced reflecting differences in the pH gradient. Predicted lysosomal $K_{pu,\text{lys}}$ for individual drugs and cell types are provided in the Supplementary Material Table S2.

Figure 3, Table 3

While the lung $K_{pu,k}$ values predicted by the original Rodgers model were strongly correlated with experimental data ($r^2 = 0.73$), the under-prediction was apparent, as 60% of the drugs were outside the 2-fold range of the observed values. The consideration of lysosomal sequestration resulted in drug-dependent increase in predicted lung $K_{pu,k}$; fold increase ranged between 1.78 - 2.33 (model based on AM and residual cells) and 1.85 - 2.46 (model based on AM, type II cells and residual cells, Supplementary Material Table S3). The only exception was phencyclidine, for which only a minor change in predicted $K_{pu}$ ($\leq 8\%$) was observed with incorporation of the lysosomal sequestration (Table 2). The increase in predicted lung $K_{pu}$ values was reflected in the reduced percentage of under-predictions (using 2-fold error margin) from 60% (original model) to 32% (model based on AM and residual cells) and to 28% (model based on AM, type II cells and residual cells). The following drugs remained under-predicted even with the extended model (phencyclidine, S-propranolol, verapamil, S-betaxolol, R-betaxolol, biperiden) with their predicted lung $K_{pu}$ values representing 31-43% of the observed (phencyclidine being the most pronounced outlier). In the case of imipramine and S-oxprenolol, under-prediction was less pronounced and the predicted $K_{pu}$ values were on the borderline of the 2-fold error. The most pronounced under-prediction was evident for compounds with the highest measured tissue accumulation ($K_{pu} > 33$, Table 2), whereas drugs showing low accumulation in the lung tissue were generally predicted well, with the exception of phencyclidine. Lidocaine $K_{pu,\text{lung}}$ was over-predicted (10-fold), consistent with the original Rodgers model. Use of the extended mechanistic model resulted in the slight over-prediction of $K_{pu,\text{lung}}$ in the case of S-acebutolol and inaperisone, in contrast to the original model.

Application of the extended model resulted in an improved overall prediction accuracy (GMFE <1.9) compared to the original model (GMFE = 2.653). The consideration of type II cells in addition to
alveolar macrophages and residual cells did not yield a significant improvement of the overall model performance (Table 3); hence, all the $K_{pu,lung}$ results above are based on the extended model based on AM and residual cells.

3.3 The prediction of drug distribution into kidney

Comparison of experimental and predicted $K_{pu}$ in rat kidney using the original and extended Rodgers model, are shown in Figure 3B, along with corresponding correlation statistics (Table 3). Despite the complexity of kidney tissue, the extended mechanistic model assumed equal contribution of lysosomes to the cellular volumes of different cell types and $f_{lyso(Kidney)}$ was set to 1.2%.

A linear relationship between experimental and predicted $K_{pu,K}$ and $K_{pu,K}'$ was evident, yielding an $r^2$ of 0.67-0.68. Similar to the lung, consideration of lysosomal sequestration in the model resulted in the overall increase in predicted $K_{pu}$ values (1.61 – 1.95 fold, with the exception of phencyclidine) and reduced extent of under-predictions (from 21% to 7% using 2-fold error or 4% when 3-fold error was considered). Compounds which were under-predicted by the extended model were S-acebutolol (42% of the observed $K_{pu}$) and nicotine (13% of the observed). However, the extent of over-predictions increased (25% of the dataset relative to 11% in the case of the original Rodgers model), with S-carvedilol (3.7-fold), R-propranolol (4-fold) and R-carvedilol (5.6-fold) being the most pronounced outliers. Apart from lidocaine and S-acebutolol, there was no overlap in drugs being over- or under-predicted in lung and kidney tissue. Application of the extended Rodgers model slightly decreased the overall model performance (GMFE of 1.73 vs. 1.91); however, the total number of over- and under-predictions remained unchanged (32%).

3.4 The prediction of drug distribution into liver

Liver is considered a lysosome-rich tissue, mainly due to the presence of Kupffer cells. Based on the analysis of collated literature data, these cells contain the highest lysosomal volume fraction among all liver parenchymal cell types ($f_{lyso(Kupffer)} = 13.6\%$), followed by endothelial cells ($f_{lyso(EndothCells)} = 6.86\%$). This is in contrast to hepatocytes ($f_{lyso(Hep)} = 0.82\%$) and fat-storing cells ($f_{lyso(Fat-Storing)} = 0.23\%$; details are shown in Table 1 and Figure S1B in the Supplementary Material. Despite their abundance of lysosomes, Kupffer and endothelial cells constitute only approximately 5% of total liver parenchyma.
whereas the largest proportion of liver is composed of hepatocytes (77.8% of total liver parenchym and 84.1% of total liver) [35]. Therefore, the extended mechanistic distribution model was evaluated following two approaches, namely i) applying a complex model with all abovementioned cell types (Kupffer, endothelial, fat-storing cells and hepatocytes) and ii) assuming a simplified model in which liver parenchym was considered to be composed of hepatocytes only with $f_{\text{Lys(Hep)}} = f_{\text{Lys(Parenchym)}} = 0.82\%$.

Figure 3C shows the comparison of experimental and predicted $Kpu_k$ and $Kpu_u$ data in rat liver using the original and the extended Rodgers model (hepatocytes-only model). As seen in the case of kidney, the consideration of lysosomal sequestration reduced the extent of under-prediction, and increased the number of over-predicted $Kpu$ values compared to the original analysis. A drug-dependent increase in predicted $Kpu$ was observed when lysosomal sequestration was considered in the predictive model, with a fold increase ranging from 1.94 - 2.31 (model based solely on hepatocytes) and 2.21 - 2.7 (model based on all cell types; phencyclidine was exception in both cases with a maximal 10% increase in predicted liver $Kpu$). Consideration of either hepatocytes alone or all cell types in the extended mechanistic model resulted in no under-predictions of liver $Kpu$ in contrast to 27% in the original analysis. However, pronounced over-prediction of liver $Kpu$ was evident with the extended model (46% outside 2-fold); this trend was apparent regardless of the cell types considered in the model for liver. For certain drugs, the fold over-prediction in liver $Kpu$ was comparable to kidney (R-carvedilol - 3.7, R-propranolol - 4.7, S-carvedilol - 2.4), whereas in some cases the extent of over-prediction was more pronounced relative to either kidney or lung (e.g., pentazocine – 26-fold). Lidocaine $Kpu$ was over-predicted for all tissues investigated. Increased number of over-predictions reduced predictive ability of the extended model when comparing performance using 3-fold limits (88% vs. 58% for the original and extended Rodgers model with all cell types, respectively). This was also reflected in the GMFE values (Table 2). Consideration of different cell types in the extended model for liver resulted in marginal differences in the total percentage of drugs with liver $Kpu$ predicted within the 2-fold (54-58%).
3.5 The prediction of drug distribution into lung - sensitivity analysis

All model predictions of $K_{puR}$ in lung shown here are based on the average values for the cellular physiological input parameters ($f_{lyso}$ and $pH_{lyso}$), as summarized in Table 1. In order to assess the impact of parameter variability on $K_{puR}$ predictions, a sensitivity analysis was performed by changing one parameter at a time according to the reported physiological range (Table 4). The fraction of cells was kept constant ($f_{AM} = 4.1\%$, $f_{Type-II} = 8.3\%$ and $f_{Res,Cells} = 87.6\%$) and no shift in the lysosomal pH in presence of drug was assumed.

Table 4

In the first instance, the sensitivity of $K_{puR}$ to $f_{Lys}$ in alveolar macrophages was investigated using five representative drugs selected from the 28-compound dataset. Rationale for their selection was either the occurrence of under-prediction (S-propranolol, imipramine, biperiden, S-oxprenolol) or over-prediction (lidocaine) by the original Rodgers & Rowland model. Three drugs had the same $pK_a$ (9.5) but covered a range of lipophilicity ($\log P = 2.18-4.8$), whereas the remaining two drugs differed in both their $pK_a$ and $\log P$ (biperiden: $pK_a = 8.8$, $\log P = 4.25$; lidocaine: $pK_a = 8$; $\log P = 2.44$).

Figure 4A shows experimental versus predicted $K_{puR}$ and $K_{puL}$ data for the five model drugs; the extended Rodgers model was parameterized with a range of $f_{Lys(AM)}$ (3.7-13.9%, increments of 0.1%); changes in predicted $K_{puR}$ as a function of $f_{Lys(AM)}$ are shown in Figure 4B.

The maximum 1.33-fold change in predicted $K_{puR}$ was observed over the range in $f_{Lys(AM)} = 3.7--13.9\%$ regardless of differences in $\log P$ and $pK_a$ of the five drugs. Increase in the $f_{Lys(AM)}$ resolved the under-prediction trend for all of the selected drugs with the exception of propranolol-S.
In order to investigate the impact of $f_{Lys(AM)}$ on $Kpu_R^*$ for a more diverse set of compounds, the sensitivity analysis was performed for the entire 28-compound dataset. Predicted $Kpu_R^*$ values for the individual drugs and for settings of $f_{Lys(AM)} = 0\%, 3.7\%$ and $13.9\%$ are shown in the Supplementary Material (Supplementary Material Table S4), along with comparisons of model performances (Supplementary Material Table S5). The analysis showed a 1.26 to 1.34-fold increase in predicted $Kpu_R^*$ across the physiological $f_{Lys(AM)}$ range (3.7-13.9%); as previously observed, the effect on predicted phencyclidine lung $Kpu$ was minimal (3%). The maximum fold-difference only marginally increased to 1.53 when comparing $Kpu_R^*$ predictions based on $f_{Lys(AM)} = 13.9\%$ and 0\% (assuming no lysosomal contribution in these cells).

Further analysis of the relationship between fold-changes in predicted $Kpu_R^*$ values and physicochemical properties of the 28-compound dataset showed the following trends (Figure S3, Supplementary Material): the absolute changes in predicted $Kpu$ across $f_{Lys(AM)}$ range (3.7-13.9%) increased with increasing log $P$ and increasing experimental $Kpu$, whereas no clear trend was evident for $pK_a$ based on the current dataset. However, relative changes in predicted $Kpu_R^*$ (i.e. fold-changes) using a range in $f_{Lys(AM)}$ was comparable for all drugs, independent of their log $P$, $pK_a$ or experimental $Kpu$. Sensitivity analyses was also performed by assessing the impact of individual physicochemical parameters on the predicted lung $Kpu_R^*$ ($pK_a = 7-11$ and log $P = 1-7$). Drugs with a basic $pK_a \geq \sim 8$ showed a plateau with almost constant fold-changes in predicted lung $Kpu_R^*$ over a wide log $P$ range (details are in the Supplementary Material, Figure S4).

The impact of the remaining parameters ($f_{Lys(TpeII)}$, $pH_{Lys(AM)}$ and $pH_{Lys(TpeII)}$) on predicted lung $Kpu_R^*$ values for 28 compounds are shown in the Supplementary Material Table S4, with summary statistics in Table 4. Overall, predicted $Kpu_R^*$ values increased with an increase in the lysosomal volume fraction and a decrease in the lysosomal pH value. Similar to the analysis of the effect of $f_{Lys(AM)}$, fold changes in predicted lung $Kpu_R^*$ over a range in these physiological parameters were
relatively consistent for the drugs investigated and resulted in maximal 1.14- \((f_{LyS(TypeII)})\), 1.26- \((pH_{LyS(AM)})\) and 1.89-fold increases \((pH_{LyS(TypeII)})\) (Table 4). The analysis showed that fold changes in predicted lung \(Kpu^*_R\) were relatively low when considering the variability of the individual parameters in isolation. A combination of parameters is required in order to explain up to 2.46-fold change in lung \(Kpu^*_R\) evident when comparing the original and the extended Rodgers model.
4 Discussion

4.1 Extended mechanistic model for the prediction of $K_{pu}$

In this study, a previously published tissue composition-based model for the prediction of $K_{pu}$ [6] was extended to account for the contribution of lysosomal sequestration to drug tissue distribution. Analogous to the original model, drug ionization, drug distribution in tissue water and drug binding to neutral lipids, neutral phospholipids and acidic phospholipids were taken into account. In addition, drug binding to lysosomal lipids and pH differences between lysosome and cytosol were incorporated in the model. The extended mechanistic model also considered differences in cytosolic and lysosomal pH, lysosomal volume fractions and cell volume fractions among individual tissues/cell types; this was implemented using literature collated and critically assessed cellular physiological parameters. The aim of the work was to evaluate the predictive performance of the extended model for lysosome-rich tissues and tissues associated with under-prediction of $K_{pu}$ in the previous analyses (lung, liver and kidney). Due to limited availability of experimental $K_{pu}$ data and to allow direct evaluation against the existing model, the dataset consisted of the same 28 moderate-to-strong bases and physicochemical input data (i.e., $pK_a$, log $P$, $K_a$ (inferred from $K_{pu_{bc}}$)), as in Rodgers & Rowland work [6].

4.2 Model performance and the impact of lysosomal sequestration on predicted $K_{pu}$

The extended distribution model predicted extensive accumulation of all 28 drugs in lysosomes, as illustrated by the predicted $K_{pu_{lyso}} >1000$ for 79% of the drugs in the dataset in for example alveolar macrophages. The extent of lysosomal accumulation was less pronounced in hepatocytes, but still prominent ($K_{pu_{lyso}}$ ranging from 94 to 935 for nicotine and imipramine, respectively, Supplementary Material, Table S2), reflecting differences in lysosomal pH gradient and concentration of acidic phospholipids among cells. Despite high predicted accumulation in the lysosomes of individual cell types, maximal predicted increase in tissue $K_{pu}$ as a result of lysosomal sequestration did not exceed 3-fold for the dataset investigated (1.95 – 2.7 for kidney and liver, respectively); the fold change varied across tissues and assumptions made on their cellular composition (Table 2, Supplementary Material Section 2 and Table S3).
These results are not entirely surprising considering that lysosomes account for a relatively minor fraction of the entire tissue volume. For example, despite lysosomal volume fractions of 13.6% in Kupffer cells (liver) and up to 13.9% in alveolar macrophages (lung), the fraction of these lysosome-rich cell types in the individual tissues is generally low (<5%). In contrast, the largest proportion of liver and lung is composed of hepatocytes and residual cells (type I cells, interstitial cells, interstitial matrix, endothelial cells [27]), respectively, which are not particularly lysosome-rich (e.g., $f_{Lys(Rep)} = 0.82\%$). The same is true for kidney where limited data indicate that proximal tubule cells contain higher lysosomal volume fractions [32, 36, 37].

Application of the extended mechanistic model resulted in no under-prediction of liver $Kpu$ and significantly reduced the extent of under-prediction in the case of the lung (32%) and the kidney (<10%). However, increase in over-predictions of $Kpu$ relative to the original Rodgers model was evident across all tissues investigated. Overall, consideration of lysosomal sequestration improved $Kpu$ prediction accuracy for lungs, but resulted in comparable performance for liver and kidney (Table 3) relative to the model with no consideration of lysosomal trapping. Since the individual consideration of type II cells in addition to alveolar macrophages and residual cells did not yield a significant improvement in the overall lung $Kpu$, the extended model based on alveolar macrophages and combined residual cells represents the best compromise between predictive performance and model complexity in the case of lung.

Tissue distribution of basic drugs is driven by a combination of phospholipid binding (plasma and lysosomal membrane) and pH-driven lysosomal sequestration [14, 21, 25]. The extent of lysosomal accumulation has been suggested to be more prominent for dibasic compared to monobasic drugs, as illustrated by chloroquine [38]. Both the original and extended mechanistic tissue distribution models consider only the first $pK_a$ for $Kpu$ predictions. While this simplification is justified for the majority of compounds in the dataset with only one center of ionization, quinidine and timolol have two basic $pK_a$ values and ideally the $Kpu$ prediction model should account for this. Consideration of the highest basic $pK_a$ in the current study predicted $Kpu^*_r$ (lung) for both drugs within 2-fold of the experimental data. Over-prediction in the case of liver was seen for both drugs and for quinidine in kidney,
suggesting other contributing factors rather than consideration of a second $pK_a$ to the model performance.

4.3 Challenges and future perspectives for the mechanistic prediction of $K_{pu}$ accounting for lysosomal processes

Whereas lysosomal sequestration is linear at low drug concentrations, the capacity of the lysosomes to accumulate basic drugs is reduced at higher drug concentrations as a result of increasing lysosomal $pH$. Hence theoretical accumulation ratios in lysosomes suggested for certain drugs (e.g., for quinidine: 160,000, [13]) are most likely not achieved under *in vivo* conditions. Neither of the mechanistic distribution models account for these nonlinear effects. Although lysosomal $pH$ values are easily modified in the extended Rodgers model, a dynamic change of $pH_{Lys}$ as a function of drug concentration is not accommodated.

The uncertainty in some physiological input parameters such as $pH_{Lys}$, $f_{Lys}$ and $f_{Cell\ Type}$ (for tissues like kidney) is a general challenge, as the accuracy of $K_{pu}$ predictions critically depends on these values. An extensive literature research has been conducted to collate data for all required cellular input parameters. The composition of lysosomal membranes was not available for all cell types and therefore it was assumed that the fraction of lipids, phospholipids and water is the same as in the cell membrane; however, studies report some differences in mitochondrial membrane and high percentage of sphingolipids and lysophosphoglycerides in lysosomal relative to plasma membrane [39]. Physiological input parameters used in the extended model represent average values from different sources, animal strains and methodologies applied. For a number of parameters (e.g., $pH_{Lys}$ in residual cells in lung), no experimental data were reported and therefore estimates based on studies in other cell types had to be used. Despite these limitations, the sensitivity analysis showed that relatively minor changes in predicted $K_{puR}$ are expected when varying $pH_{Lys}$ and $f_{Lys}$ within the physiologically reported range. The disparities seen between predicted and observed $K_{pu}$ represent a combination of the uncertainty associated with some cellular parameters/assumptions (e.g., setting $f_{Lys(Kidney)}$ to 1.2%) and true biological variability in physiological parameters (e.g., lysosomal
It is important to note that tissue composition/ phospholipid content and cell phenotype may change under certain conditions (e.g., smoking [40, 41]) which may result in increased relative importance of lysosomal sequestration for drug tissue distribution and higher inter-individual variability. Significant deviations from average physiological input parameters are also anticipated in the case of drug-induced pulmonary phospholipidosis which is associated with excessive accumulation of phospholipids in lysosomes and the formation of multilamellar bodies. However, information about the direct impact of lysosomal phospholipid accumulation in the case of phospholipidosis on tissue distribution is scarce [42]. Extension of the current model to any other tissues (e.g., heart) requires collation of physiological data for that particular tissue together with cell composition data, all of which may be limited.

In addition to lysosomes, endosomes and the Golgi complex are acidic organelles [43, 44]) and accumulation of some basic drugs in mitochondria (basic pH environment) has been reported in the literature, but these organelles were not considered in the proposed model. Lamellar bodies in type II cells have been assumed to be more abundant than lysosomes due to their function in storing pulmonary surfactant [45]. Potential implications of pulmonary surfactants on the uptake of basic drugs have not been considered in the extended mechanistic model and may explain the remaining cases of under-prediction of $Kpu_{\text{lung}}$.

The accuracy of physicochemical parameters ($\log P$, $pK_a$) and inconsistencies between measured and calculated parameters (e.g., verapamil measured $\log P$ of 3.79 vs. predicted $\log P$ value of 4.80 (KOWWIN, v1.67) is a usual caveat when using predictive in silico tools. An additional source of uncertainty relates to the $K_a$ parameter derived from $\log P$, $pK_a$ and $Kpu_{\text{BC}}$. For highly lipophilic drugs and compounds with low-to-intermediate lipophilicity in combination with low $pK_a$, predicted $K_a$ values can become negative, as observed for phencyclidine and nicotine, respectively. For these drugs electrostatic interaction with acidic phospholipids in the cell (including the lysosome) was assumed to be negligible and $K_a$ was set to zero. Drug binding to acidic phospholipids in lysosomes is considered to amplify lysosomal drug accumulation driven by pH partitioning [13, 21]. Neglecting drug binding to acidic phospholipids in the lysosome leads to an underestimation of $Kpu_{\text{BC}}$ and in turn underestimates the fold changes in predicted $Kpu$ values when comparing the models. This effect has
been reflected in lower fold-increase in $K_{pu}$ and lower lysosomal accumulation ratios seen for phencyclidine and nicotine compared to all other drugs.

Potential impact of transporter mediated uptake/efflux to tissue distribution cannot be excluded entirely for the dataset investigated. Verapamil and quinidine are known P-glycoprotein substrates and predicted $K_{pu}$ values are expected to be over-estimated by the current mechanistic models. This trend was apparent for both drugs in the case of liver $K_{pu}$, but not in the case of lung or kidney (Table 2). Some of the over-or under-prediction of kidney $K_{pu}$ may be attributed to potential contributing role of renal uptake (e.g., OCT2) and efflux transporters (MATE-2K). However, no clear trends could be established between the extent of liver and kidney $K_{pu}$ mis-prediction and limited kinetic in vitro / in vivo data on the relevance of transporters for individual compounds in the dataset (collated data for individual drugs shown in the Supplementary Material Table S6). For example, in the case of the most pronounced outlier (carvedilol) only evidence for transporter (OCT2) inhibition was reported [46]. Further validation on an extended and structurally more diverse dataset is desirable and highly encouraged, as more experimental $K_{pu}$ data in lysosome-rich tissue become available.
5 Conclusions

In summary, the proposed mechanistic model for the prediction of $K_{pu}$ for moderate-to-strong bases ($pK_a > 7$) accounts for lysosomal sequestration by both pH-driven gradient and binding to acidic phospholipids. In addition, important drug-specific and tissue-specific cellular physiological input parameters (e.g., cell volume fractions, lysosomal volume fractions and lysosomal pH) are considered. Application of the proposed model reduced significantly the extent of $K_{pu}$ under-prediction in lysosome-rich tissues (lung in particular). Consideration of lysosomal sequestration resulted in the improved prediction accuracy (lung), whereas predictive performance of the model was comparable to the original Rodgers& Rowland model in the case of liver and kidney, despite an increase in over-predictions for these two tissues. The current analysis highlighted high subcellular accumulation for the drugs investigated which may have significant implications for drug safety and efficacy; however, the extent of the increase in the overall tissue distribution as a result of lysosomal sequestration did not exceed 3-fold. A sensitivity analysis confirmed the importance of lysosomal trapping for drug tissue distribution but highlighted that additional processes such as feedback mechanisms and active uptake may contribute to the remaining discrepancy between predicted and observed $K_{pu}$ values. To further validate the mechanistic model developed, direct measurements of drug lysosomal concentrations and the extent of drug accumulation are required. The application of the extended mechanistic distribution model is recommended for lysosome-rich tissues to support the advancement of PBPK modeling of volume of distribution and drug tissue exposure.
Acknowledgements

The authors would like to thank Prof. Malcolm Rowland and Dr David Hallifax for valuable discussions. Moreover, we thank Drs Ayse Ufuk, Daniel Scotcher and Laura Francis for their input on physiological input parameters.

Author contributions

FA developed and implemented the extended model, collated literature data, performed predictions, analyzed the results and wrote the manuscript. JBH and AG contributed to the design and analysis of the study and the writing of the manuscript.

Conflict of Interest

The authors declare no conflict of interest
FIGURE CAPTIONS

**Figure 1.** Illustration of the extended Rodgers model on the cellular level. Processes involved in drug distribution and considered by the extended Rodgers model are shown, including lysosomal sequestration of basic compounds.

**Figure 2.** Illustration of the extended Rodgers model on the tissue level using lung as an example. The fraction of various cell types in tissue ($f_{cell \_type}$), lysosomal volume fraction ($f_{lys}$) and lysosomal pH ($pH_{lys}$) are taken into account. Corresponding references listed in Table 1.

**Figure 3.** Predicted versus experimentally determined $K_{pu}$ values in A: rat lung (n=25), B: rat kidney (n=28) and C: rat liver (n=26) for basic compounds with $pK_a > 7$. Prediction based on either the original Rodgers model (black triangles) or the extended Rodgers model (grey triangles). The extended model was based on the following assumptions - A: Lung is comprised of alveolar macrophages ($f_{lys} = 7.8\%$) and residual cells ($f_{lys} = 1\%$), B: Kidney is comprised of various cell types with an average lysosomal volume fraction of $f_{lys} = 1.2\%$; and C: Liver is comprised of hepatocytes only ($f_{lys} = 0.82\%$). The dashed lines indicate the two fold error lines.

**Figure 4.** A: Predicted versus experimental $K_{pu}$ values for five model drugs, namely 1: S-propranolol; 2: imipramine; 3: S-oxprenolol; 4: biperiden; 5: lidocaine. Predicted $K_{pu}$ values were derived from the original (black triangles) and the extended Rodgers model parameterized with a range of lysosomal volume fractions in alveolar macrophages ($f_{lys} = 3.7 – 13.9\%$, grey lines). The remaining input parameters were kept constant at their average values listed in Table 1. The presence of AM, type II cells and residual cells was considered. The dashed lines indicate the two fold error lines. B: Predicted $K_{pu}^*$ values as a function of $f_{lys}$ for the 5 model drugs using the extended Rodgers model as described in A.
Figure 1.
Lung tissue

**Alveolar macrophages**
- $f_{cell type} = 4.1\%$
- $f_{lys} = 7.8\%$
- $pH_{lys} = 4.75$

**Type II cells**
- $f_{cell type} = 8.3\%$
- $f_{lys} = 3\%$
- $pH_{lys} = 5.1$

**Residual cells**
- $f_{cell type} = 87.6\%$
- $f_{lys} = 1\%$
- $pH_{lys} = 5.1$
Figure 3.

(A) Predicted Kpu (lung) vs. Experimental Kpu (lung)

(B) Predicted Kpu (kidney) vs. Experimental Kpu (kidney)

(C) Predicted Kpu (liver) vs. Experimental Kpu (liver)
Figure 4.
### Table 1. Tissue-specific input parameters for the extended Rodgers model used to predict \( K_{pu} \) values in lung, kidney and liver in rat.

<table>
<thead>
<tr>
<th></th>
<th>Lung</th>
<th>Kidney</th>
<th>Liver</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alveolar ( f_{cell_type} ) [%]</td>
<td>Type II ( f_{cell_type} ) [%]</td>
<td>Residual cells ( f_{cell_type} ) [%]</td>
</tr>
<tr>
<td></td>
<td>4.1(^1)</td>
<td>8.3(^1)</td>
<td>87.6(^1)</td>
</tr>
<tr>
<td>( f_{lys} ) [%]</td>
<td>7.8(^3)</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>( pH_{lys} )</td>
<td>4.75(^6)</td>
<td>5.1(^7)</td>
<td>5.1</td>
</tr>
</tbody>
</table>

**Fractional tissue volume (total tissue)** \(^\text{11}\)

| \( f_{NL} \)          | 0.022 | 0.012 | 0.014 |
| \( f_{NP} \)          | 0.0128 | 0.0242 | 0.0240 |
| \( f_{EW} \)          | 0.336 | 0.273 | 0.161 |
| \( f_{IW} \)          | 0.446 | 0.483 | 0.573 |

**Tissue concentration of acidic phospholipids (total tissue)** \(^\text{11}\)

| \( AP^- \) [mg/g] | 3.91 | 5.03 | 4.56 |
Average values for the volume fraction of different cell types in tissue reported for Fischer 344 rat and Sprague-Dawley rat [47]. Air and capillary volume was not considered.

Volume fraction of different cell types reported for Sprague-Dawley rat [35]. Intercellular spaces (15.9% of total liver volume) were not considered.

Mean value of lysosomal volume fraction in rat collated from literature as reported in [20].

Lysosomal volume fraction in total rat kidney, derived from the sum of absolute volumes for lysosomes in cortex, outer and inner stripe of the outer medulla and inner medulla [32]. The fractional volume of lysosomes was obtained by relating the total volume of lysosomes in kidney (34μL) to the total volume of a standardized kidney (1320μL).

Values taken as reported in [35].

Value measured in mouse alveolar macrophages [30].

Lysosomal pH measured in lamellar bodies of type II cells in rat [31].

Mean value from [34, 48-53], [33]. Lysosomal pH was measured in intact cells of different cell types and in different species.

Mean value of lysosomal pH in intact rat hepatocytes [38, 54-57].

Measured in rat Kupffer cells as reported in [58].

Values taken as reported by [6].
Table 2. Physicochemical properties and additional compound-specific input parameters used for the prediction of $K_{pu}$ in rat lung (n=25), kidney (n=28) and liver (n=26). Experimental and predicted $K_{pu}$ data, as obtained from the original and the extended Rodgers model, are provided.

<table>
<thead>
<tr>
<th>ID</th>
<th>Compound</th>
<th>Physicochemical properties</th>
<th>Tissue binding data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$pK_a$</td>
<td>$\log P$</td>
</tr>
<tr>
<td>1</td>
<td>R-Acebutolol</td>
<td>9.7</td>
<td>1.87</td>
</tr>
<tr>
<td>2</td>
<td>S-Acebutolol</td>
<td>9.7</td>
<td>1.87</td>
</tr>
<tr>
<td>3</td>
<td>R-Betaxolol</td>
<td>9.4</td>
<td>2.59</td>
</tr>
<tr>
<td>4</td>
<td>S-Betaxolol</td>
<td>9.4</td>
<td>2.59</td>
</tr>
<tr>
<td>5</td>
<td>Biperiden</td>
<td>8.8</td>
<td>4.25</td>
</tr>
<tr>
<td>6</td>
<td>R-Bisoprolol</td>
<td>9.4</td>
<td>1.87</td>
</tr>
<tr>
<td>7</td>
<td>S-Bisoprolol</td>
<td>9.4</td>
<td>1.87</td>
</tr>
<tr>
<td>8</td>
<td>R-Carvedilol</td>
<td>8.1</td>
<td>4.19</td>
</tr>
<tr>
<td>9</td>
<td>S-Carvedilol</td>
<td>8.1</td>
<td>4.19</td>
</tr>
<tr>
<td>10</td>
<td>Fentanyl</td>
<td>9.0</td>
<td>4.05</td>
</tr>
<tr>
<td>11</td>
<td>Imipramine</td>
<td>9.5</td>
<td>4.80</td>
</tr>
<tr>
<td>12</td>
<td>Inaperisone</td>
<td>9.1</td>
<td>3.72</td>
</tr>
<tr>
<td>13</td>
<td>Lidocaine</td>
<td>8.0</td>
<td>2.44</td>
</tr>
<tr>
<td>14</td>
<td>R-Metoprolol</td>
<td>9.7</td>
<td>2.15</td>
</tr>
<tr>
<td>15</td>
<td>S-Metoprolol</td>
<td>9.7</td>
<td>2.15</td>
</tr>
<tr>
<td>16</td>
<td>Nicotine</td>
<td>7.8 / 3.0</td>
<td>1.17</td>
</tr>
<tr>
<td>17</td>
<td>R-Oxyprenolol</td>
<td>9.5</td>
<td>2.18</td>
</tr>
<tr>
<td>18</td>
<td>S-Oxyprenolol</td>
<td>9.5</td>
<td>2.18</td>
</tr>
<tr>
<td></td>
<td>Drug</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>----------------------</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td></td>
<td>Pentazocine</td>
<td>8.5</td>
<td>3.31</td>
</tr>
<tr>
<td></td>
<td>Phencyclidine</td>
<td>9.4</td>
<td>4.96</td>
</tr>
<tr>
<td></td>
<td>R-Pindolol</td>
<td>8.8</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>S-Pindolol</td>
<td>8.8</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>Procaainamide</td>
<td>9.2</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>R-Propranolol</td>
<td>9.5</td>
<td>3.65</td>
</tr>
<tr>
<td></td>
<td>S-Propranolol</td>
<td>9.5</td>
<td>3.65</td>
</tr>
<tr>
<td></td>
<td>Quinidine</td>
<td>10.0/5.4</td>
<td>3.44</td>
</tr>
<tr>
<td></td>
<td>S-Timolol</td>
<td>9.2/8.8</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>Verapamil</td>
<td>8.5</td>
<td>3.79</td>
</tr>
</tbody>
</table>

a Experimental values as stated in Rodgers et al. [6].
b Mean values predicted using the online packages KOWWIN (erc.syrres.com), Interactive Analysis (LogP.com) and SPARC (ibmlc2.chem.uga.edu/sparc), as stated in Rodgers et al. [6].
c Blood cell-to-plasma water concentration ratio, as stated in Rodgers et al. [6].
d Association constant of basic compounds with acidic phospholipids, derived from $K_{pu}$, $pK_a$ and $log P$, as described in Rodgers et al. [6].
d1 Fold changes in $K_{pu}$ following consideration of lysosomal trapping. $d1.78 – 2.33$ (phencyclidine: 1.07); $d2.1.85 – 2.46$ (phencyclidine: 1.08); $d3.1.61 – 1.95$ (phencyclidine: 1.07); $d4.1.94 – 2.31$ (phencyclidine: 1.10); $d5.2.21 – 2.70$ (phencyclidine: 1.13).
Table 3. Comparison of the accuracy of $K_{pu}$ predictions in rat lung (n=25), kidney (n=28) and liver (n=26) using the original and the extended Rodgers model. Statistics are given based on the two-fold and three-fold (in brackets) error margin.

<table>
<thead>
<tr>
<th></th>
<th>Lung</th>
<th>Kidney</th>
<th>Liver</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rodgers, literature</td>
<td>Rodgers, extended</td>
<td>Rodgers, literature</td>
</tr>
<tr>
<td></td>
<td>AM$^a$, residual cells</td>
<td>AM, type II cells</td>
<td></td>
</tr>
<tr>
<td>Number of drug under-predicted</td>
<td>15 (9)</td>
<td>8 (2)</td>
<td>7 (1)</td>
</tr>
<tr>
<td>Number of drugs over-predicted</td>
<td>1 (1)</td>
<td>3 (1)</td>
<td>4 (1)</td>
</tr>
<tr>
<td>% of drugs under-predicted</td>
<td>60 (36)</td>
<td>32 (8)</td>
<td>28 (4)</td>
</tr>
<tr>
<td>% of drugs over-predicted</td>
<td>4 (4)</td>
<td>12 (4)</td>
<td>16 (4)</td>
</tr>
<tr>
<td>% of drugs within 2-fold (and 3-fold)</td>
<td>36 (60)</td>
<td>56 (88)</td>
<td>56 (92)</td>
</tr>
<tr>
<td>GMFE</td>
<td>2.653</td>
<td>1.898</td>
<td>1.890</td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.73</td>
<td>0.72</td>
<td>0.72</td>
</tr>
</tbody>
</table>

$^a$ Alveolar macrophages
$^b$ Diverse cells: Hepatocytes, Kupffer cells, Endothelial cells, fat-storing cells
Table 4. Fold-changes in predicted $Kpu^*_R$, derived from parameterization of the extended Rodgers model with a range of physiological input parameters.

<table>
<thead>
<tr>
<th></th>
<th>Lysosomal volume fraction</th>
<th>Lysosomal pH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alveolar macrophages</td>
<td>Type II cells</td>
</tr>
<tr>
<td>Range</td>
<td>3.7 – 13.9%</td>
<td>1 – 6%</td>
</tr>
<tr>
<td>Average value</td>
<td>7.8%</td>
<td>3%</td>
</tr>
<tr>
<td>Fold change in $Kpu$</td>
<td>1.26 – 1.34</td>
<td>1.11 – 1.14</td>
</tr>
<tr>
<td></td>
<td>(1.03)$^b$</td>
<td>(1.02)$^b$</td>
</tr>
<tr>
<td>Setting yielding lowest GMFE</td>
<td>13.9%</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alveolar macrophages</td>
<td>Type II / residual cells</td>
</tr>
<tr>
<td>Range</td>
<td>4.69 – 4.81</td>
<td>4.7 – 5.4</td>
</tr>
<tr>
<td>Average value</td>
<td>4.75</td>
<td>5.1</td>
</tr>
<tr>
<td>Fold change in $Kpu$</td>
<td>1.19 – 1.26</td>
<td>1.66 – 1.89</td>
</tr>
<tr>
<td></td>
<td>(1.02)$^b$</td>
<td>(1.08)$^b$</td>
</tr>
</tbody>
</table>

$^a$The fold changes were calculated for the 28-compound dataset and refer to changes of $Kpu^*_R$ using minimum and maximum values for either $f_{Lys}$ or $pH_{Lys}$ (in alveolar macrophages and type II cells) while keeping the other parameters constant (average values were used for the other parameters).

$^b$Number in brackets refer to fold-changes for phencyclidine (outlier).
References:


