

Ionisation and Lipophilicity

Key parameters for the prediction of human intestinal absorption

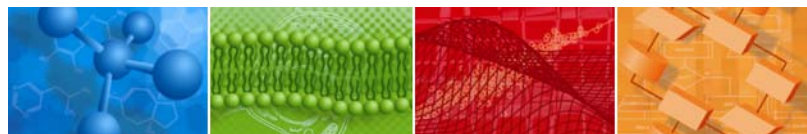
R. Didziapetris^{1,,} P. Japertas^{1,2,} A. Petrauskas³ and D. P. Reynolds⁴

¹Pharma Algorithms, Inc., A.Mickeviciaus 29, Vilnius 08117, Lithuania

²Vilnius University, Faculty of Chemistry, Naugarduko 24, Vilnius 03225, Lithuania

³Advanced Pharma Algorithms, Inc., 591 Indian Rd, Toronto, ON, Canada, M6P 2C4

⁴REYTEK Ltd., 11C Rothsay Road, Bedford, UK MK40 3PP



Pharma Algorithms

- Models that use simple descriptors are often qualitative e.g. rule of five
- Attempts to derive accurate quantitative models often use large numbers of descriptors or deploy non-linear statistical methods (e.g. Neural Networks)
- Very few approaches model the pH dependence of properties or incorporate pKa as a descriptor

Modelling Partition



- Partition: For molecules in their neutral form e.g. Hansch-Leo

$$cLogP = \sum X_i + \sum C_{ij} + const$$

X_i = group contributions

C_{ij} = correction factors

- Distribution:

- If more than one ionised form contributes to the overall property then the above equation cannot be modified directly to predict logD and pH dependence.
- Fraction ionised of a molecule or functional group is not a linear function of pH and pKa

Modelling Partition



Pharma Algorithms, Inc.

- Partition: For molecules in their neutral form e.g. Hansch-Leo

$$cLogP = \sum X_i + \sum C_{ij} + const$$

X_i = group contributions

C_{ij} = correction factors

- More generally:

$$cLogP = LFER$$

Where *LFER* = sum of terms comprising molecular descriptors and system coefficients

Modelling pH dependent Distribution



- The overall distribution is a summation of the partition of the differently charged species:

$$D = \sum F^i \times P^i$$

Where F^i = Fraction of species i and P^i = partition of species i

- Partition of an individual ionised species can be estimated from the partition of the neutral form using a Linear Free Energy approximation:

$$\text{Log}P^i = \text{LFER}$$

- The simplest form is: $\text{Log}P^i = \text{Log}P^0 + \Delta^i$

Appropriate values for octanol/water partition

are:	Δ^+ - mono-cation	= -3
	Δ^\pm - zwitterion	= -2.5
	Δ^- - mono-anion	= -4.1
	Δ^{2+} - bi-cation	= -5
	Δ^{2-} - bi-anion	= -5.0
	Δ^{+q} - covalent cation	= -4

Modelling a pH dependent property (X^{pH})



Pharma Algorithms, Inc.

$$X^{pH} = \sum F^i X^i$$

Where F^i = fraction of species i

If property X^i for species i is modelled by an *LFER* then:

$$\log X^{pH} = \log \left(\sum F^i 10^{LFER^i} \right) \quad (1)$$

Where F^i = fraction of species i (calculated from pH and pKa)

- The expression for logX is a non-linear function of molecular descriptors (including pH and pKa) and system coefficients
- Estimation of the system coefficients by fitting to experimental data requires application of non-linear optimisation techniques

Ionisation equilibria



– Ionisation can be calculated using the Henderson Hasselbalch relationships:

– Acid: $pH = pKa + \log_{10} [A^-]/[HA]$

$$[A^-]/[HA] = 10^{(pH-pKa)}$$

– Base: $pH = pKa - \log_{10} [BH^+]/[B]$

$$[BH^+]/[B] = 10^{(pKa-pH)}$$

Ionisation equilibria



For molecules with one acidic and one basic group
denote $x_a = 10^{[\text{pH}-\text{pKa}(\text{Acid})]}$ and $x_b = 10^{[\text{pKa}(\text{Base})-\text{pH}]}$

Neutral fraction $F^0 = 1 / [x_a + x_a x_b + x_b + 1]$

Zwitterion fraction $F^\pm = x_a x_b / [x_a + x_a x_b + x_b + 1]$

Cation fraction $F^+ = x_b / [x_a + x_a x_b + x_b + 1]$

Anion fraction $F^- = x_a / [x_a + x_a x_b + x_b + 1]$

Assumes that there is no-interaction between
ionisation centres

i.e. that micro-pKa's are the same as measured
macro-pKa

Ionisation equilibria

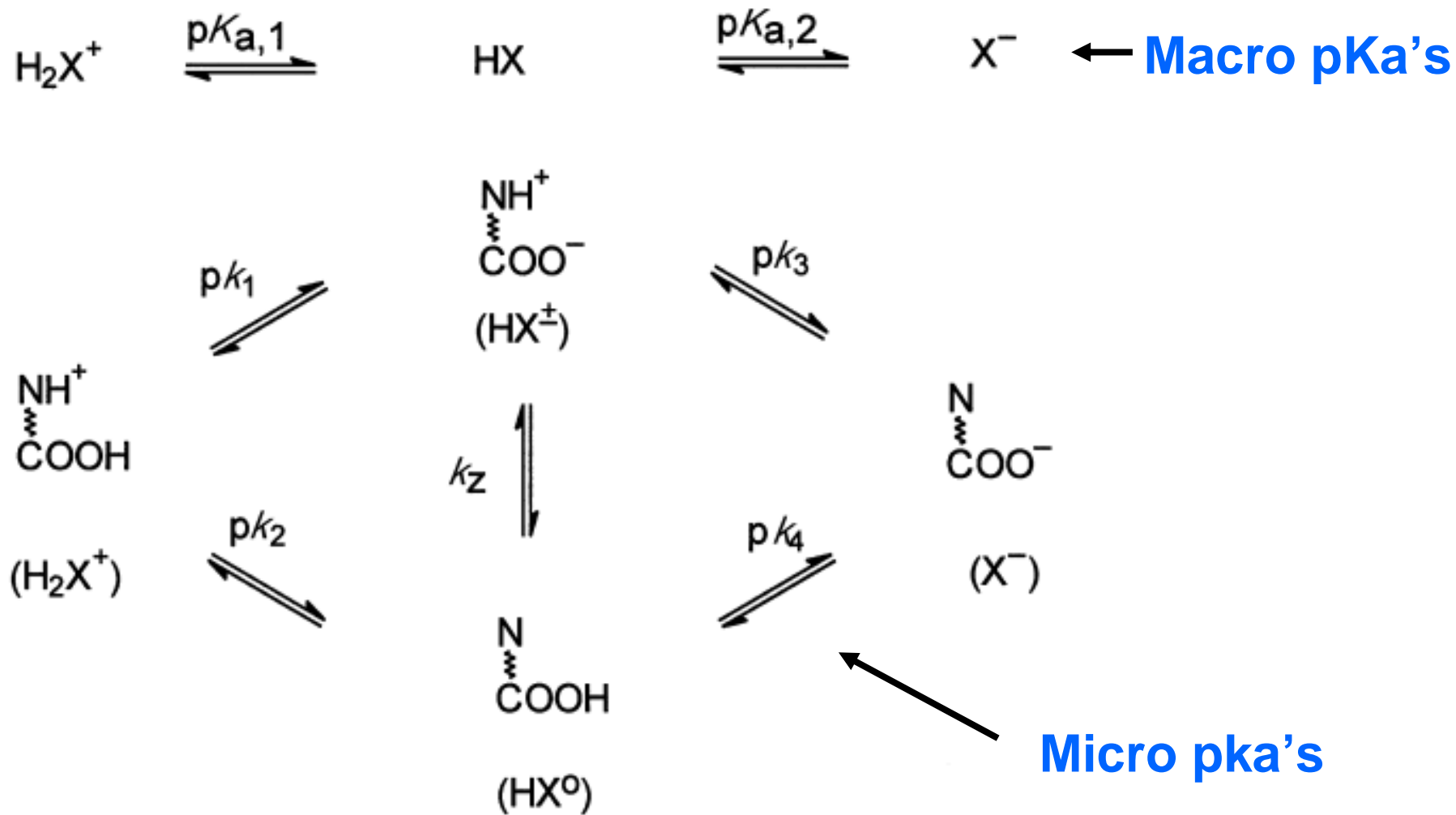


- Various software packages are available for the calculation of pKa from chemical structure
- Fractions of individual ionised forms can be rapidly calculated from either macro- or micro-pKa's
- QSPR's to predict pH dependent properties can be parameterised using non-linear optimisation techniques.

Ionisation scheme of a diprotic ampholyte



Pharma Algorithms, Inc.



pKa Prediction: Labetolol



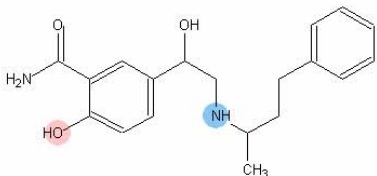
Pharma Algorithms, Inc.

ADME Boxes



- Databases
- Absolv
- Absorption
- Bioavailability
- Batch Calculations
- Distribution
- Ionization
- LogD
- pKa
- PhysChem
- Solubility
- MRDD
- P-gp Specificity
- PK Explorer

Predicted Values - pKa and Ion Fractions



Labetolol

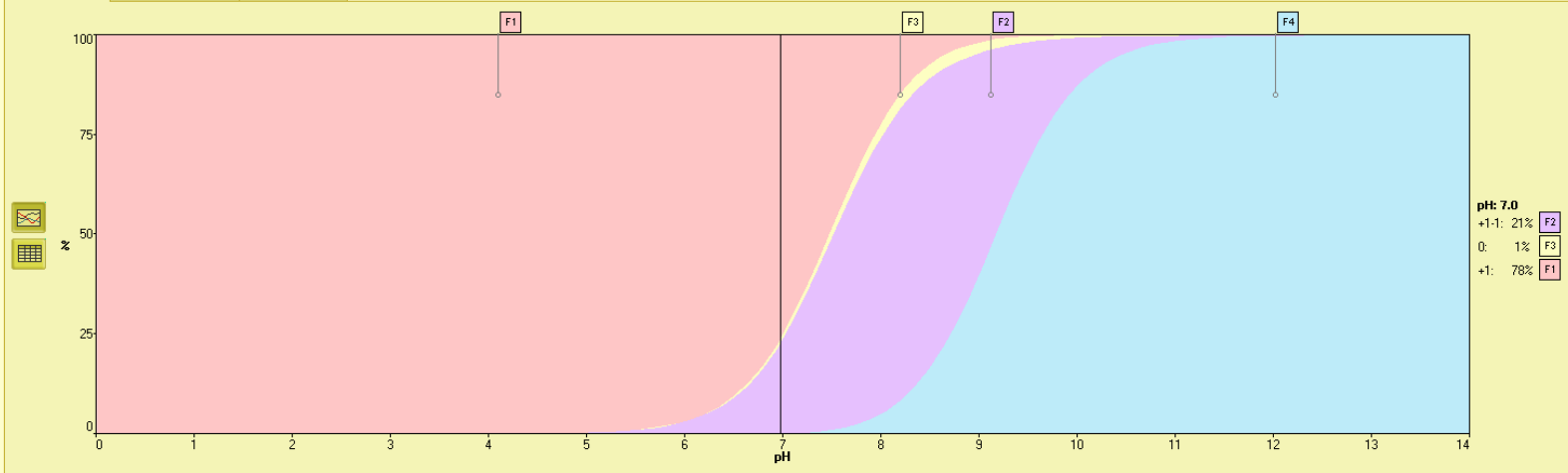


Strongest pKa(Acid): 7.5 ± 0.4
Strongest pKa(Base): 9.2 ± 0.4

pKa values: Contributions of microstages:

7.5 ± 0.4	95%	MS1	5%	MS2
9.2 ± 0.4	95%	MS3	5%	MS4

Net Charge vs pH Form Distribution vs pH Group State vs pH



pKa Prediction: Labetolol



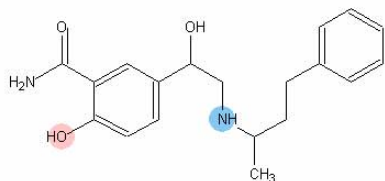
Pharma Algorithms, Inc.

ADME Boxes



- Databases
- Absolv
- Absorption
- Bioavailability
- Batch Calculations
- Distribution
- Ionization
- LogD
- pKa
- PhysChem
- Solubility
- MRDD
- P-gp Specificity
- PK Explorer

Predicted Values - pKa and Ion Fractions



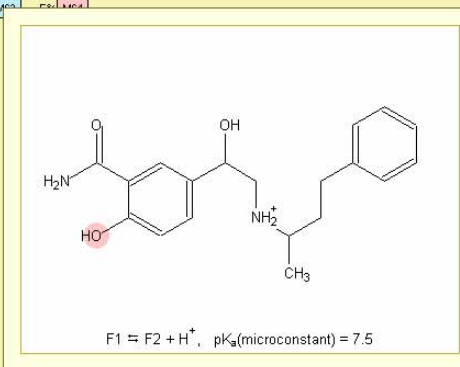
Labetolol



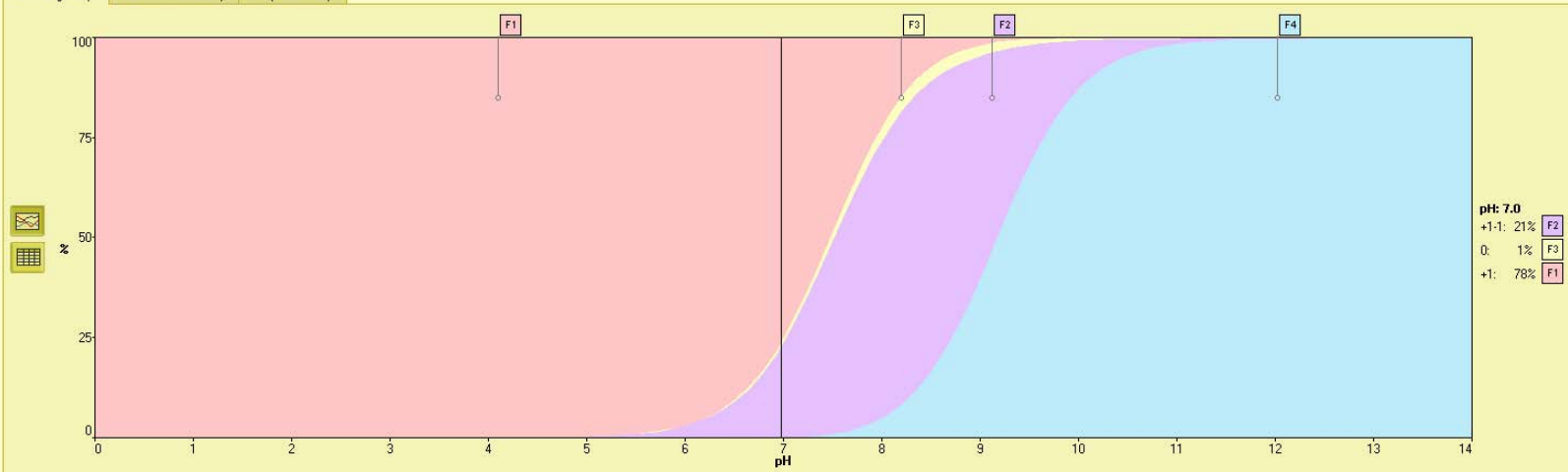
Strongest pKa(Acid): 7.5±0.4
Strongest pKa(Base): 9.2±0.4

pKa values: Contributions of microstages:

7.5±0.4 95% MS1 5% MS2
9.2±0.4 95% MS2 5% MS1



Net Charge vs pH | Form Distribution vs pH | Group State vs pH



pH: 7.0

+1-: 21% F2
0: 1% F3
+1: 78% F1

Labetolol



Pharma Algorithms, Inc.

		Calculated Fraction Ionised at pH7			
	Calculated pKa's	F^+	F_{\pm}	F^0	F^-
Macro pKa's	pKa1=7.5 pKa2=9.2	0.75	0.24	0.005	0.00
Micro pka's	pka ₁ =7.5 pka ₂ =8.8 pka ₃ =9.1 pka ₄ =7.8	0.72	0.27	0.01	0.00

QSAR example: HIA



- Many prediction models for Human Intestinal Absorption (HIA) have been reported
- A recent review [1] lists 23 models based on descriptors calculated from structure (published 1997-2005)
- None incorporate pKa as a descriptor and most use lipophilicity descriptors applicable to the neutral form
- A model published in 2006 [2] combined logD at pH 6 with HBD (or PSA) to predict absorption rate constants that correlated with experimental HIA values.
- In only three studies was the combined training + test set larger than 200 compounds
- The number of molecular descriptors used ranged from 1 up to 3387

[1] Hou et. al. (2006). "Recent advances in computational prediction of drug absorption and permeability in drug discovery." Curr Med Chem **13**(22): 2653-67.

[2] Linnankoski et. al. (2006). "Computational prediction of oral drug absorption based on absorption rate constants in humans." J Med Chem **49**(12): 3674-81.

Mechanistic Modelling



- A small number of mechanistic or physiologically based HIA models have been described.
- A quote from ref [1]:

‘Mechanistic (or rational) ADME models combine physicochemical properties of the compound and known physiological and anatomical properties, such as tissue lipid–water composition, with mathematical descriptions of fundamental biophysical processes that drive ADME behaviour. Physical laws with broad applicability guarantee the validity of the models within a wide range of chemical structure classes, and constitute the major advantage of mechanistic approaches against empirical models’

[1] Willmann, S., J. Lippert and C. Schmitt (2005). "From physicochemistry to absorption and distribution: predictive mechanistic modelling and computational tools." Expert Opin. Drug Metab. Toxicol. **1**(1): 159-168

HIA Modelling: Objectives



- Develop a model of passive intestinal absorption with broad applicability based on mechanistic insight
- Use simple and chemically meaningful molecular descriptors (i.e. logP, H-bond counts, molecular size, pKa); all of which can be predicted; some of which can be corrected by measurement
- Incorporate pH as an explicit variable in either a single compartment or multi-compartment model
- Use the largest available dataset to parameterise the model

HIA Data



- Our objective is a passive absorption model
- Ideal absorption data will be free of effects due to metabolism, active transport, poor solubility or slow dissolution.
- Compounds with appropriately evaluated human absorption data were obtained from:
 - Zhao and Abraham et. al. [1]
 - Willmann et. al. [2]
 - Remigius Didziapetris [3]
- The set of 169 molecules with data considered reliable by Abraham [1] was adopted as the test set.
- Data was available on a further 705 molecules which were used as the training set.
- Uneven distribution. The majority of compounds have HIA>90%

[1] Zhao, Y. H., J. Le, M. H. Abraham, A. Hersey et. al. (2001). J Pharm Sci **90**(6): 749-84.

[2] Willmann, S., W. Schmitt, J. Keldenich et. al. (2004). J Med Chem **47**(16): 4022-31.

[3] Selected from Pharma Algorithms Database. For an earlier publication and HIA classification model see: Zmuidinavicius, D., R. Didziapetris, P. Japertas, A. Avdeef and A. Petrauskas (2003). "Classification structure-activity relations (C-SAR) in prediction of human intestinal absorption." J Pharm Sci **92**(3): 621-33.

Equation from the CAT model

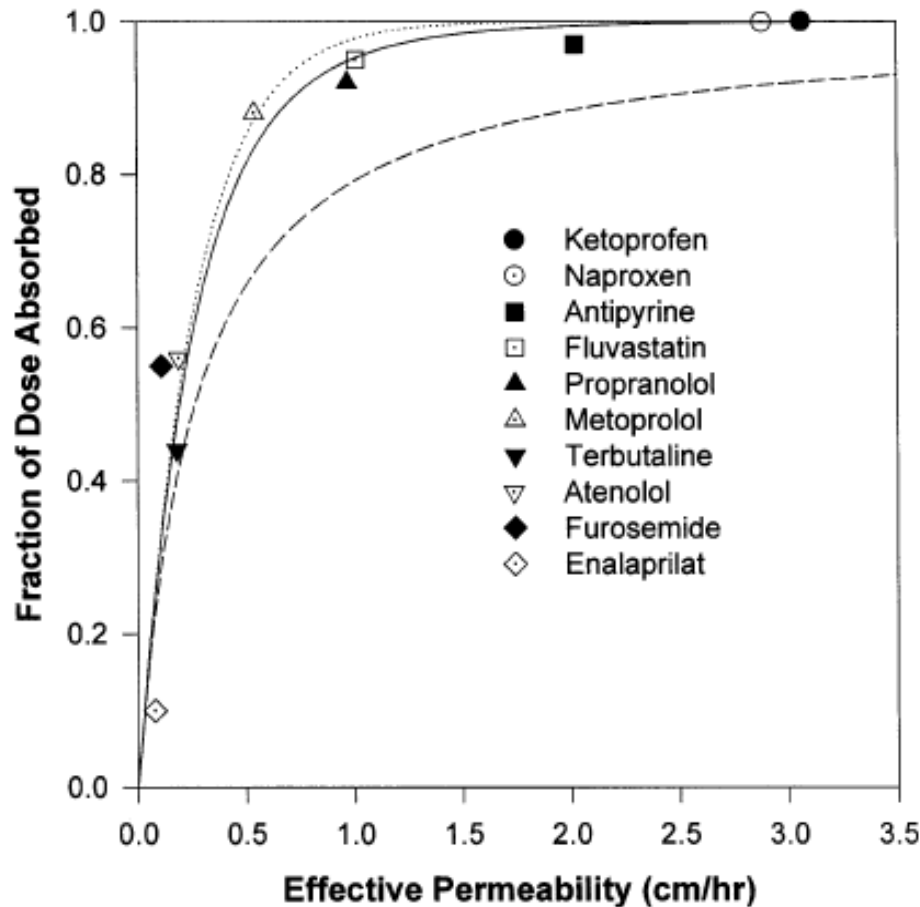


Fig. 2. The fraction of dose absorbed as a function of the human effective permeability, where (—) represents the prediction of the compartmental absorption and transit model, (---) represents the single compartment model, and (···), the plug flow model.

$$\text{Fraction Absorbed} = 1 - (1 + 0.54k_a)^{-7}$$

Yu, L. X. and G. L. Amidon (1999).

"A compartmental absorption and transit model for estimating oral drug absorption."

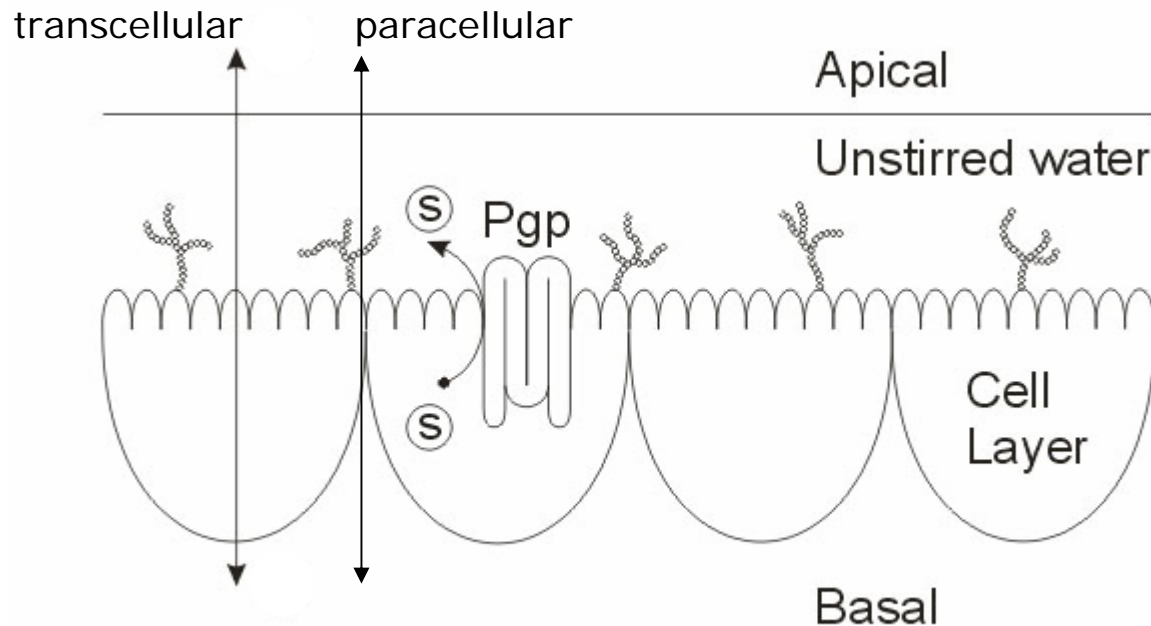
Int J Pharm **186**(2): 119-25

Passive Permeation



Overall absorption rate constant (k_a) is related to passive diffusion through an unstirred water layer (k_{Diff}), through pores between cells (k_{Para}) and through membrane bilayers (k_{Trans}).

$$1/k_a = 1/k_{Diff} + 1/(k_{Para} + k_{Trans})$$



Paracellular Permeability

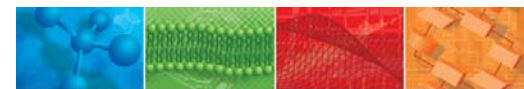


- Paracellular permeation is diffusion through the negatively charged tight junctions. It has been modelled using hydrodynamic theory as a size-restricted diffusion through a cylindrical pore within a negative electrostatic field-of-force [1].
- In Caco-2 monolayers it was observed that protonated amines permeate the pores faster than their neutral forms while organic cations were slower [1]
- A study [2] of a series of D-oligopeptides found a pronounced dependence of paracellular absorption on molecular weight, but in this series net charge had little effect (N.B. all molecules had 2 or 3 ionisable groups)
- In our model we used the Renkin function to model the size dependence and assigned an adjustable parameter i.e. relative paracellular permeability (p_{para}^i) to each type of ionic species. The final ratios were determined by fitting to experimental absorption data.

[1] Adson et. al. J Pharm Sci **83**(11): 1529-36

[2] He and Rowland et. al.(1996) Pharm Res **13**(11): 1673-8.

Paracellular Model Equations



Pharma Algorithms, Inc.

Paracellular permeability was estimated using equation (1):

$$P_{para} = C_{para} \frac{1}{r} f\left(\frac{r}{R}\right) \left(\sum^i F^i P_{para}^i \right) \quad (1)$$

Where the Renkin sieving function $0 < F\left(\frac{r}{R}\right) < 1$ is calculated by equation (2)

$$f\left(\frac{r}{R}\right) = \left[1 - \left(\frac{r}{R}\right) \right]^2 \left[1 - 2.104 \left(\frac{r}{R}\right) + 2.09 \left(\frac{r}{R}\right)^3 - 0.95 \left(\frac{r}{R}\right)^5 \right] \quad (2)$$

Molecular radius r is calculated by equation (3) from values of McGowan's molecular volume V_x estimated in Algorithm Builder:

$$r = \sqrt[3]{\frac{3V_x}{4\pi}} \times 166 \quad (3)$$

The apparent radius R of the pore was assigned as 5.65 Å based on precedent from previous work.

Transcellular Model Equations



- Transcellular permeability was estimated with simple LFER's
 - Permeability of the neutral molecule (k_{Trans}°) can be estimated from the lipophilicity of the neutral molecule

$$\text{Log } k_{\text{Trans}}^{\circ} = a_o + \log P_{\text{octanol}} (+ N_{\text{HD}} + \dots)$$

- Permeability of an ionised species can be estimated from the permeability of the neutral form:

$$\text{Log } k_{\text{Trans}}^i = \text{Log } k_{\text{Trans}}^{\circ} + \Delta_i$$

- Overall permeability is a summation of the permeability of all the contributing ionic species:

$$k_{\text{Trans}}^{\text{overall}} = \sum F_i \times (k_{\text{Trans}}^i)$$

- The fractions F^i are calculated from pH and pKa's using the Henderson-Haselbalch relationships
- Preliminary modelling suggested an optimum pH of 6.5 or 7 for a single pH model

Complete Model



Equation for %HIA

$$\%HIA = 1 - \left(1 + \frac{0.54}{\frac{1}{P_{uwl}} + \frac{1}{\sum_i F^i (10^{a_0 + a_1 \log P_{ow} + a_2 N_{HD} + \sum_j a_j I_j + \Delta_i}) + C_{para} \frac{1}{r} F\left(\frac{r}{R}\right) \left(\sum_i F^i P_{para}^i\right)}}} \right)^{-7}$$

The relationship between % HIA and the molecular descriptors is highly non-linear

This is a simple model i.e. the equation can be made much more complicated!

Complete Model



The equation contains two types of variables:

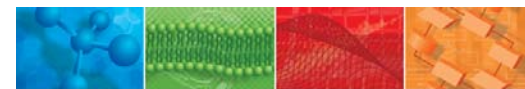
- **Calculated molecular descriptors**
(from Algorithm Builder or ADME Box software)
i.e. pK_a and F_i , $\log P_{\text{octanol}}$, N_{HD} and V_x

- **Descriptors assigned during the fitting process**
 1. Related to transcellular permeability
 2. Related to paracellular permeability
 3. The overall pH and constants that determine the relative contributions of the two pathways and the effect of the unstirred water layer

Data Fitting



- The RMS error between experimental and predicted %HIA was calculated and then minimised by variation of the model parameters
- Non-linear regression (e.g. using Excel ‘Solver’) cannot find a global minimum solution for a problem of this complexity
- By systematic variation of the descriptors the model can start simple and be made progressively more complex
- Increased complexity is only introduced if it improves the model
- A decision on the final parameterisation has still not been made
- The following slides show the development of a single compartment model at pH7 with negligible resistance from an unstirred water layer.



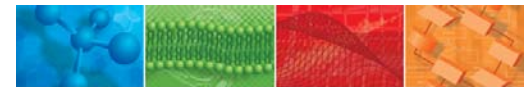
- With appropriate choice of constants the equation for %HIA simplifies from:

$$\%HIA = 1 - \left(1 + \frac{0.54}{\frac{1}{P_{uwl}} + \frac{1}{\sum_i F^i (10^{a_0 + a_1 \log P_{ow} + a_2 N_{HD} + \sum_j a_j I_j + \Delta_i}) + C_{para} \frac{1}{r} F\left(\frac{r}{R}\right) \left(\sum_i F^i P_{para}^i\right)}}} \right)^{-7}$$

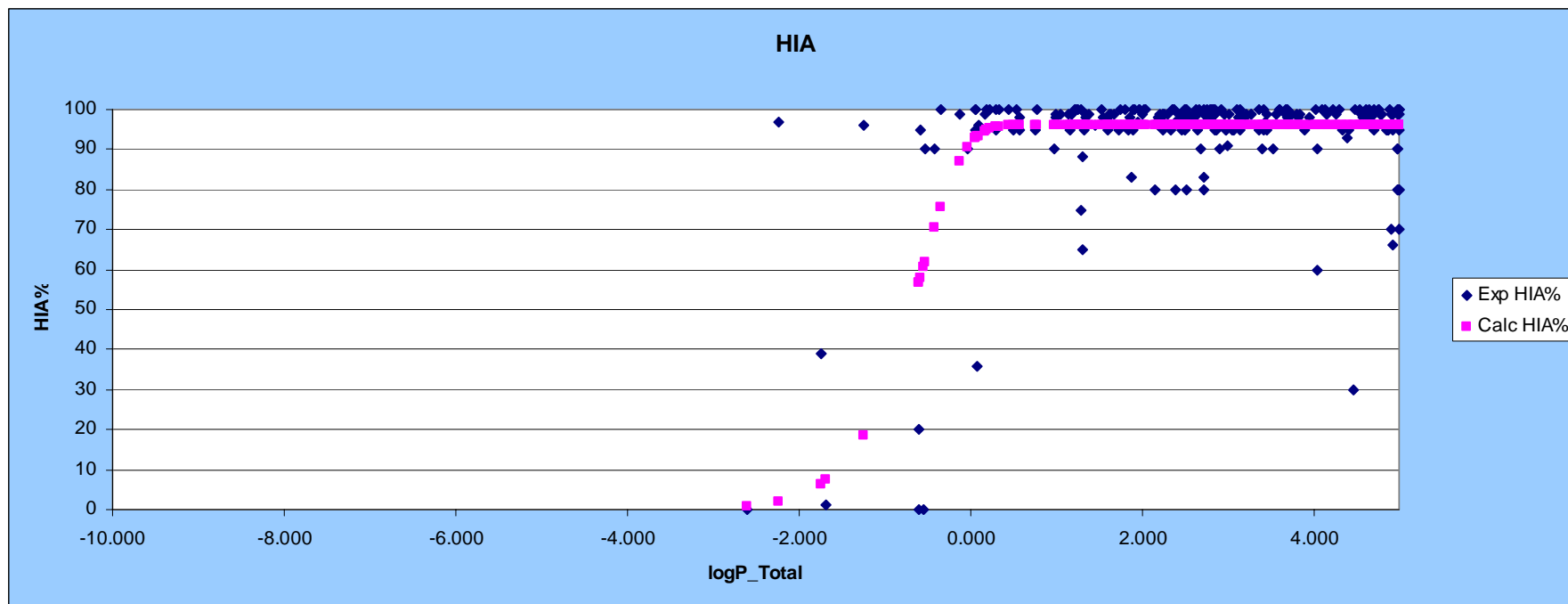
To:

$$\%HIA = 1 - \left(1 + 0.54 \left[10^{a_0 + a_1 \log P_{ow}} \right] \right)^{-7}$$

Data Fitting 1: Neutrals



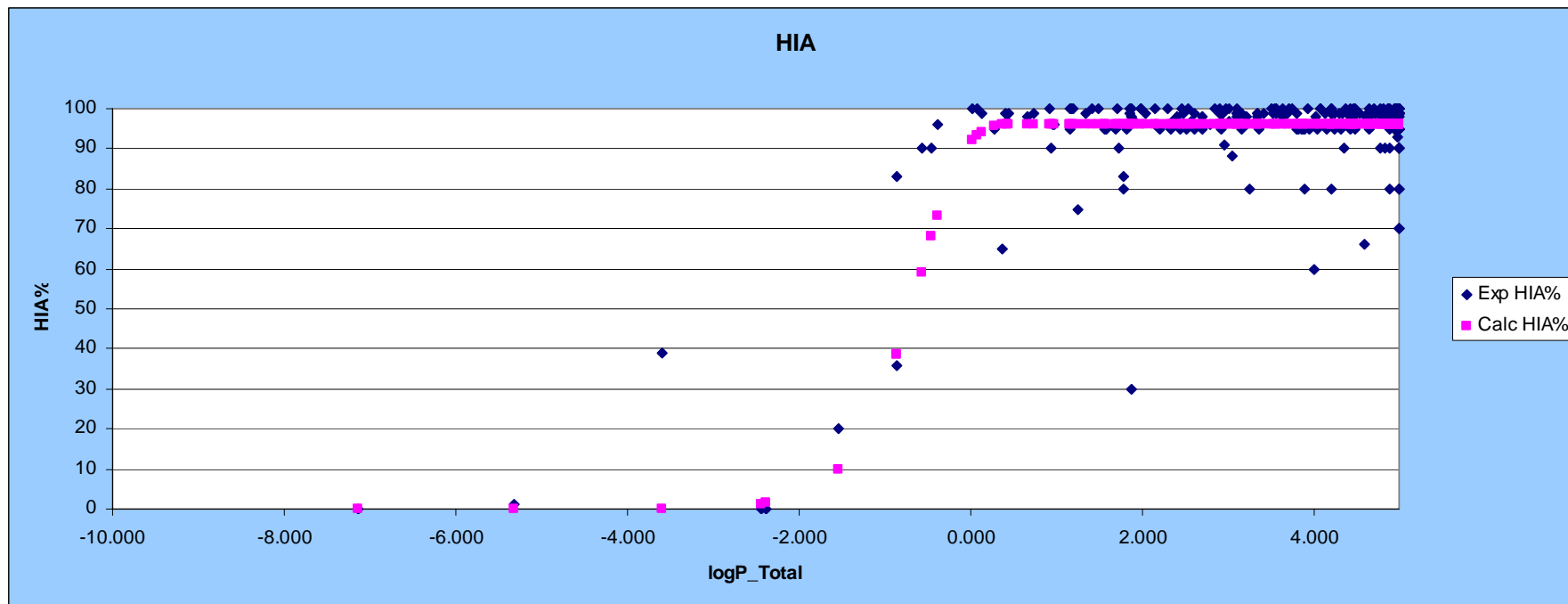
Pharma Algorithms, Inc.



– **Transcellular only:** $\text{Log } k_{\text{Trans}} = \log P_{\text{octanol}} + 0.85$

– 236 subset with $F^0 > 0.5$ at pH7 RMSE = 13.8 AAE = 6.4

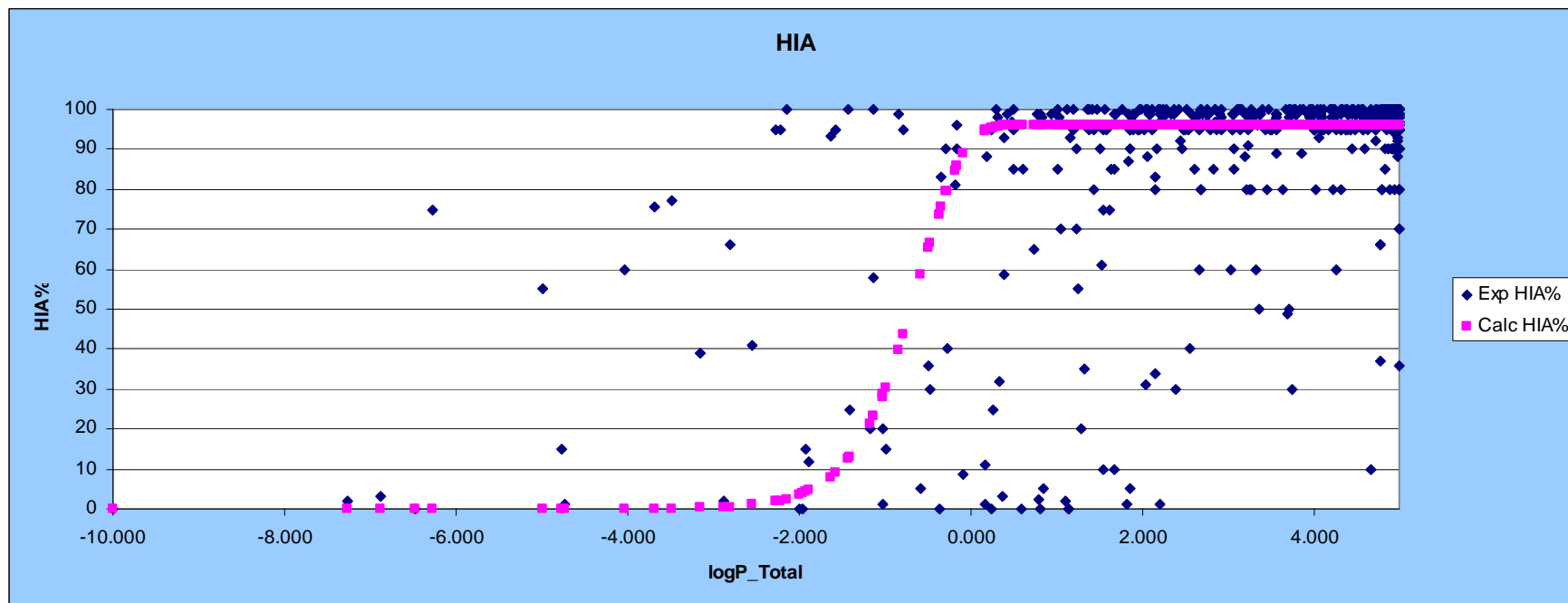
Data Fitting 2: Neutrals



– **Transcellular only:** $\log k_{\text{Trans}}^0 = \log P_{\text{octanol}} - 0.9N_{\text{HD}} + 3.51$

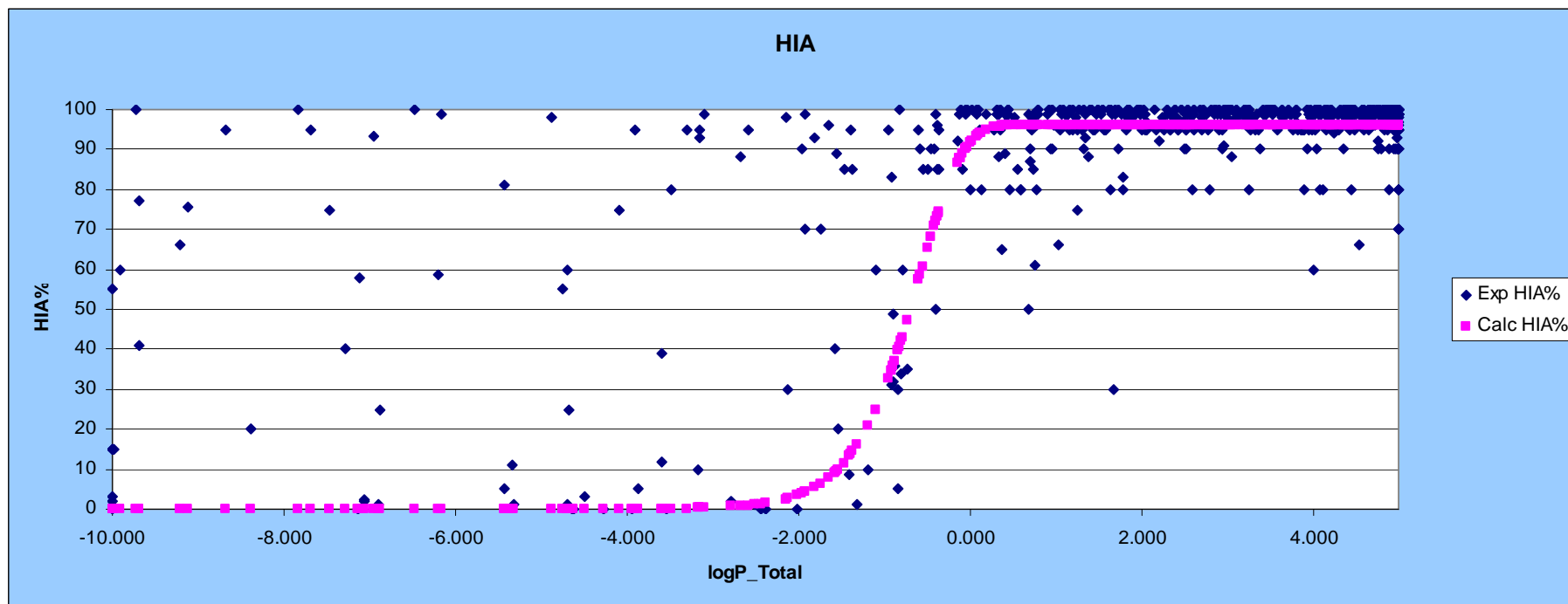
– N=235 subset with $F^0 > 0.5$ at pH7 RMSE = 9.0 AAE = 4.8

Data Fitting 3a: Acids, Bases and Neutrals



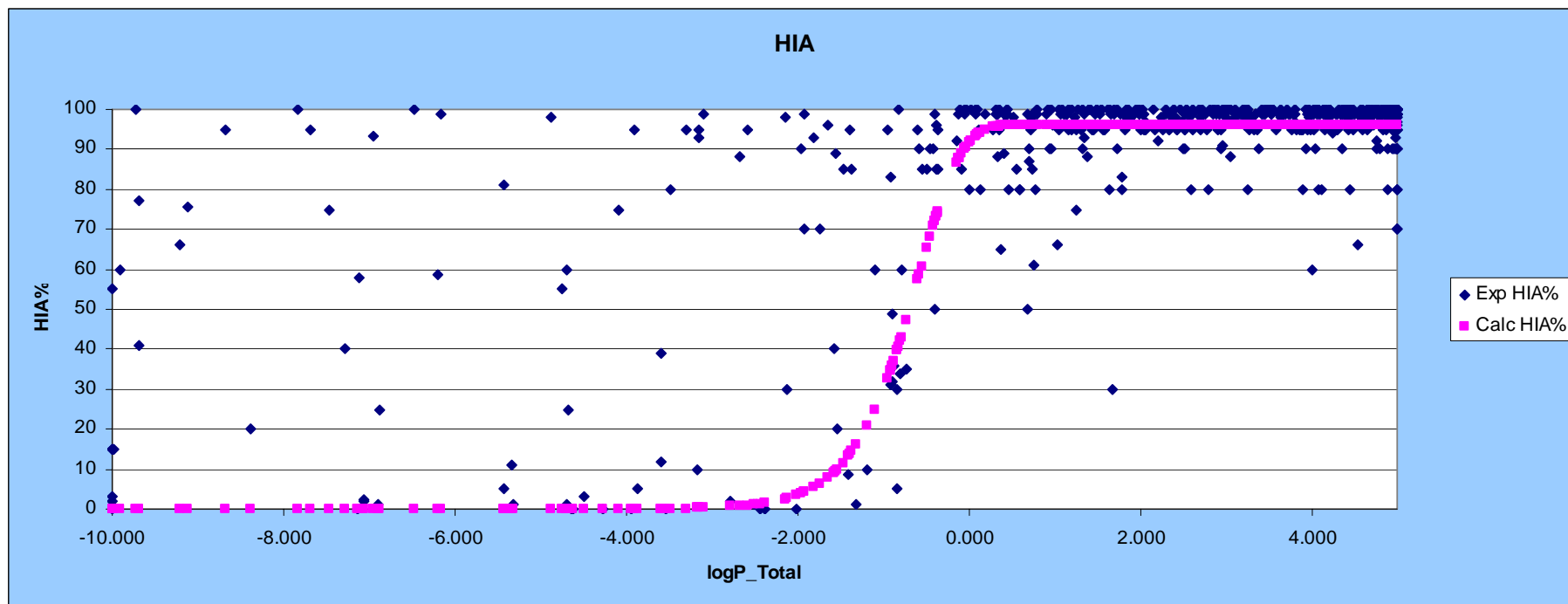
- **Transcellular only:** All species assumed to behave like neutrals
i.e. All $\Delta_i = 0$
$$\log k_{\text{Trans}}^i = \log k_{\text{Trans}}^0 + \Delta_i$$
$$\log k_{\text{Trans}}^0 = \log P_{\text{octanol}} - 0.9N_{\text{HD}} + 3.51$$
- N=630 i.e training set minus polyionic and quaternary ammonium compounds
- RMSE = 22.6 AAE = 10.3

Data Fitting 3b: Acids, Bases and Neutrals



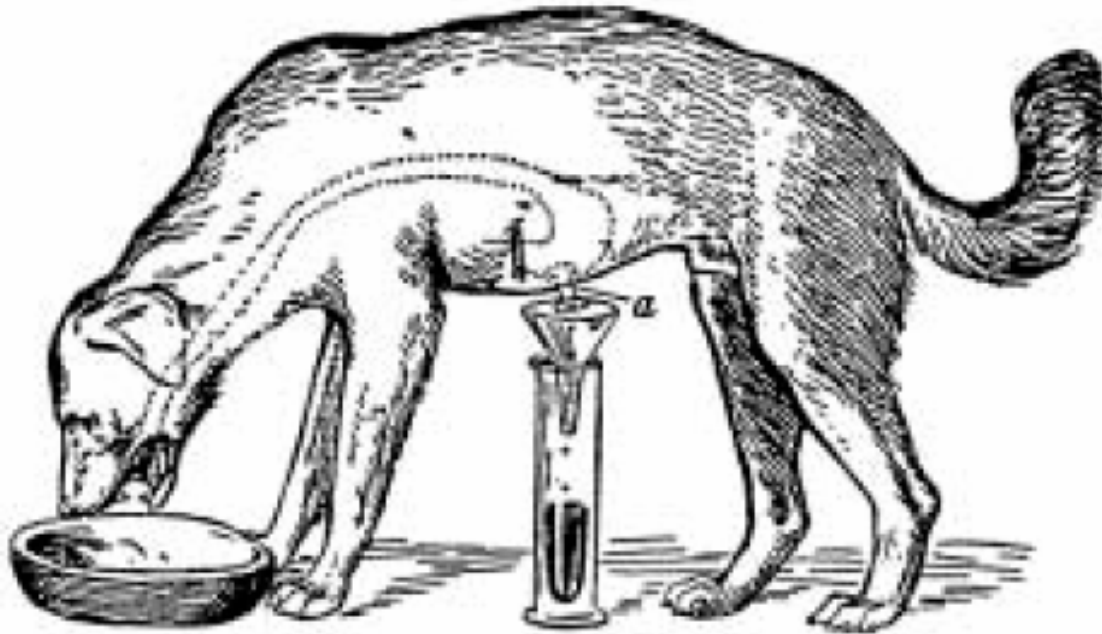
- **Transcellular only:** Ionic species assigned negligible permeability
i.e. All $\Delta_i = -10$
$$\log k_{\text{Trans}}^i = \log k_{\text{Trans}}^0 + \Delta_i$$
$$\log k_{\text{Trans}}^0 = \log P_{\text{octanol}} - 0.9N_{\text{HD}} + 3.51$$
- N=630 i.e. Training set minus polyionic and quaternary ammonium compounds
RMSE = 22.7 AAE = 10.3

Data Fitting 3b: Acids, Bases and Neutrals



- At first sight this result seems to imply that ionisation is not important
- What are appropriate values for Δ_i in the equation: $\log k_{\text{Trans}}^i = \log k_{\text{Trans}}^o + \Delta_i$
- Is there any independent evidence?

pH Dependence of the Absorption of Acids and Bases

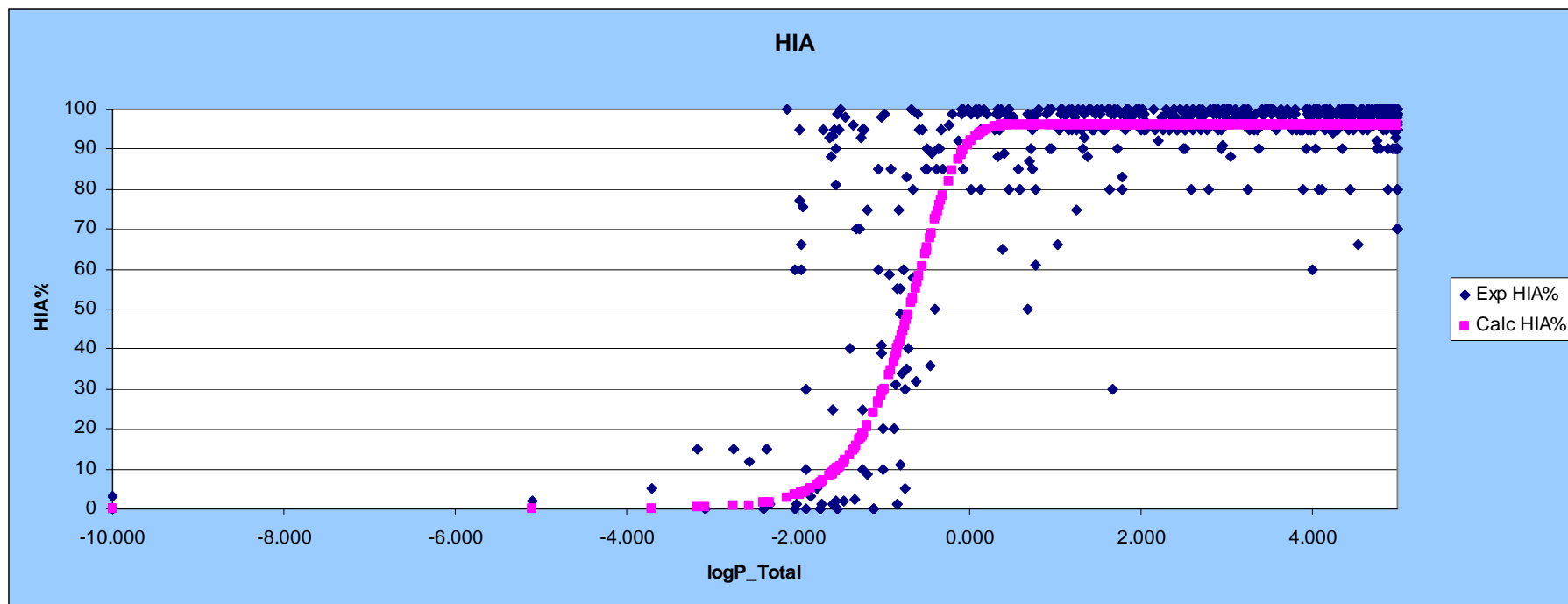


A strong base like strychnine, $pK_a = 8.3$, is not toxic for a dog with a pylorus ligation.

Data Fitting 4a: Introduction of Paracellular

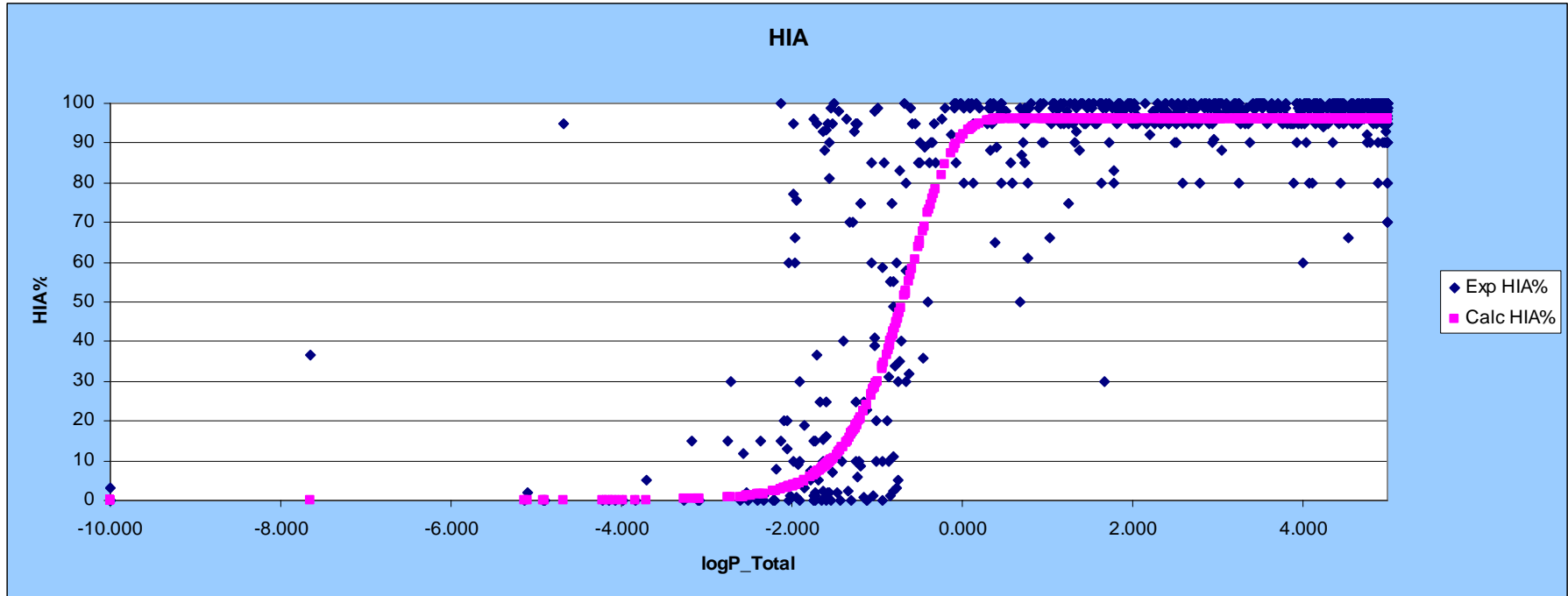


Pharma Algorithms, Inc.



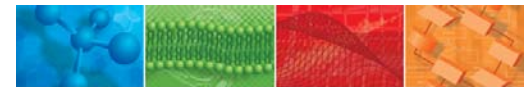
- **Transcellular:** Ionic species assigned negligible permeability
i.e. All $\Delta_i = -10$
$$\log k_{\text{Trans}}^i = \log k_{\text{Trans}}^0 + \Delta_i$$
$$\log k_{\text{Trans}}^0 = \log P_{\text{octanol}} - 0.9N_{\text{HD}} + 3.51$$
- **Paracellular:** $C_{\text{para}} = 50$ -All species assumed to have equal paracellular permeability
- N=630 Training set minus polyionic and quaternary ammonium compounds
RMSE = 18.8 AAE = 9.0

Data Fitting 4b: Full Training Set

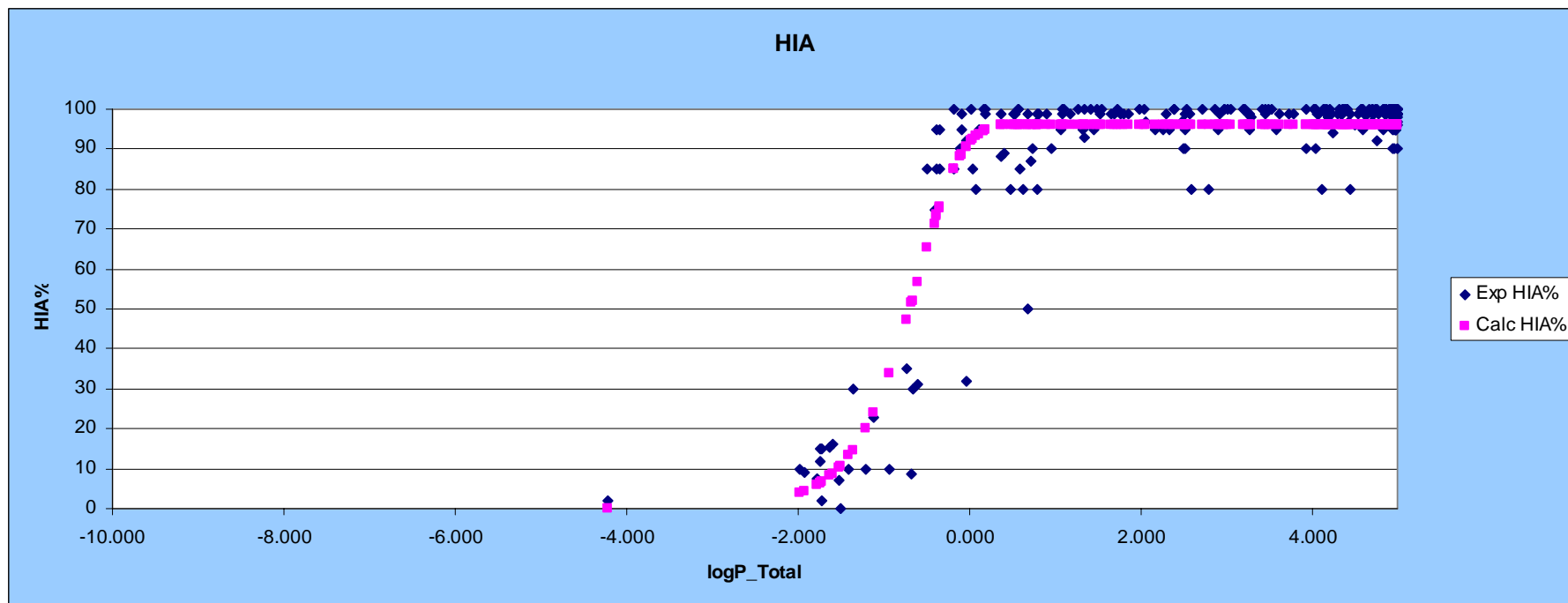


- **Transcellular:** Ionic species assigned negligible permeability
i.e. All $\Delta_i = -10$
$$\log k_{\text{Trans}}^i = \log k_{\text{Trans}}^0 + \Delta_i$$
$$\log k_{\text{Trans}}^0 = \log P_{\text{octanol}} - 0.9N_{\text{HD}} + 3.51$$
- **Paracellular:** $C_{\text{para}} = 46$ -All species assumed to have equal paracellular permeability
- N=702 Full training set. RMSE = 19 AAE = 9.3

Data Fitting 5b: Monoprotic-cations



Pharma Algorithms, Inc.



- **Transcellular:** Ionic species assigned negligible permeability

i.e. All $\Delta_i = -10$

$$\log k_{\text{Trans}}^i = \log k_{\text{Trans}}^0 + \Delta_i$$

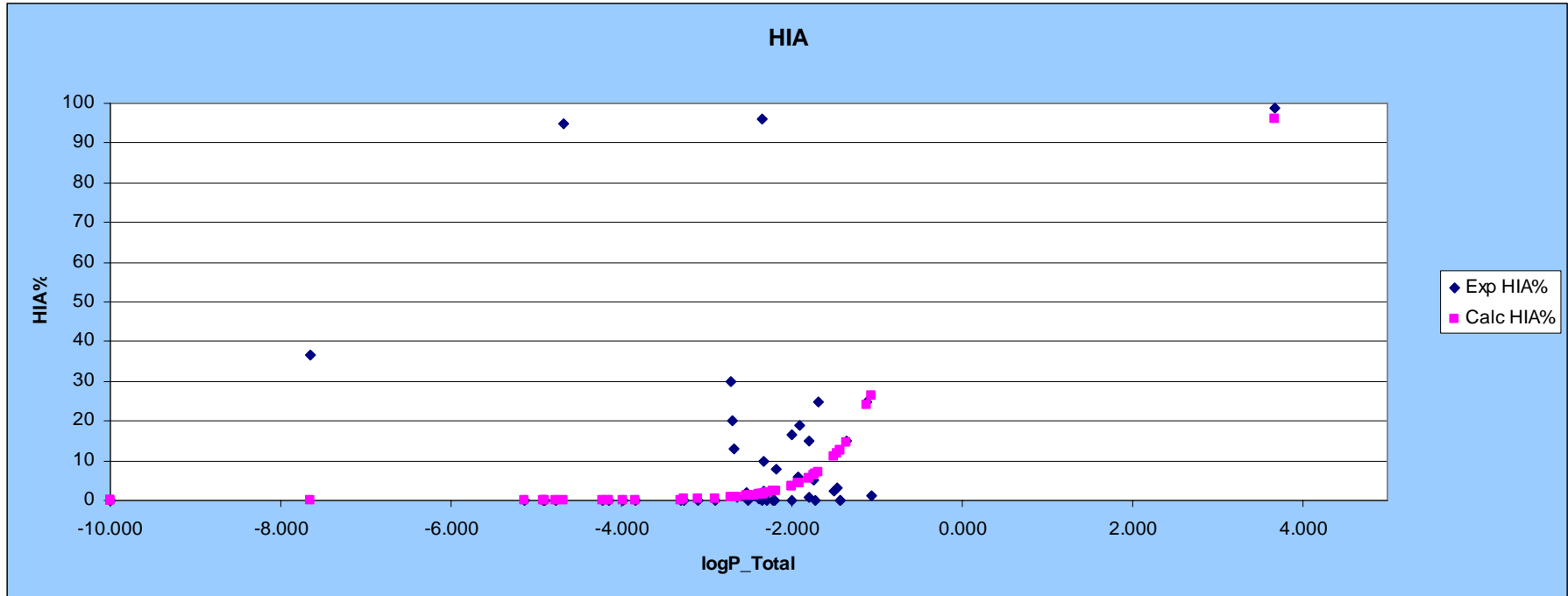
$$\log k_{\text{Trans}}^0 = \log P_{\text{octanol}} - 0.9N_{\text{HD}} + 3.51$$

- **Paracellular:** $C_{\text{para}} = 50$

Paracellular permeability of monoprotic-cations increased to 8 times the value of other species

- N=253 subset with $F^+ > 0.5$ RMSE = 8.1 AAE = 5.0

Data Fitting 5d: Poly-ions



- **Transcellular:** Ionic species assigned negligible permeability i.e. All $\Delta_i = -10$

$$\log k_{\text{Trans}}^0 = \log P_{\text{octanol}} - 0.9N_{\text{HD}} + 3.51$$

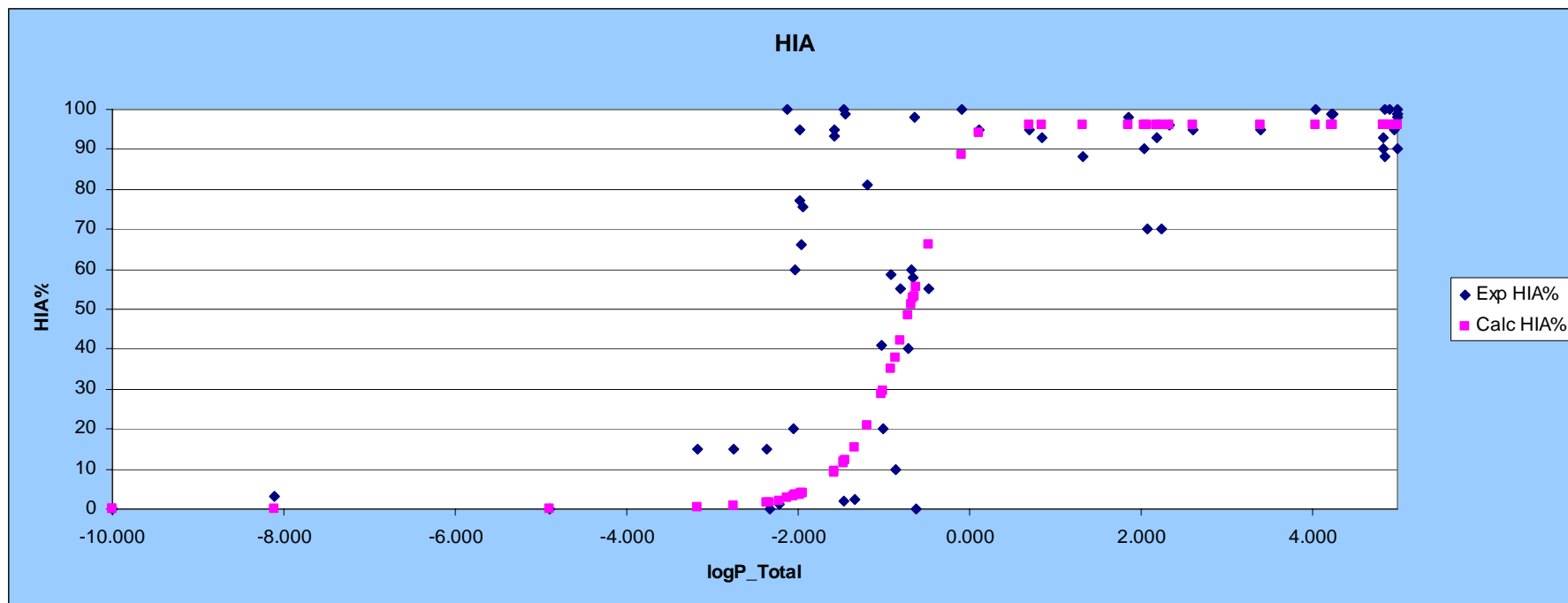
- **Paracellular:** Relative paracellular permeability of:
monoprotic-cations = 8, poly-anions = 0.2, all other species = 1

- N=53 subset with $F^{\text{polyion}} > 0.1$ RMSE = 20.7 AAE = 9

Data Fitting 7: Permeability correction for ampholytes

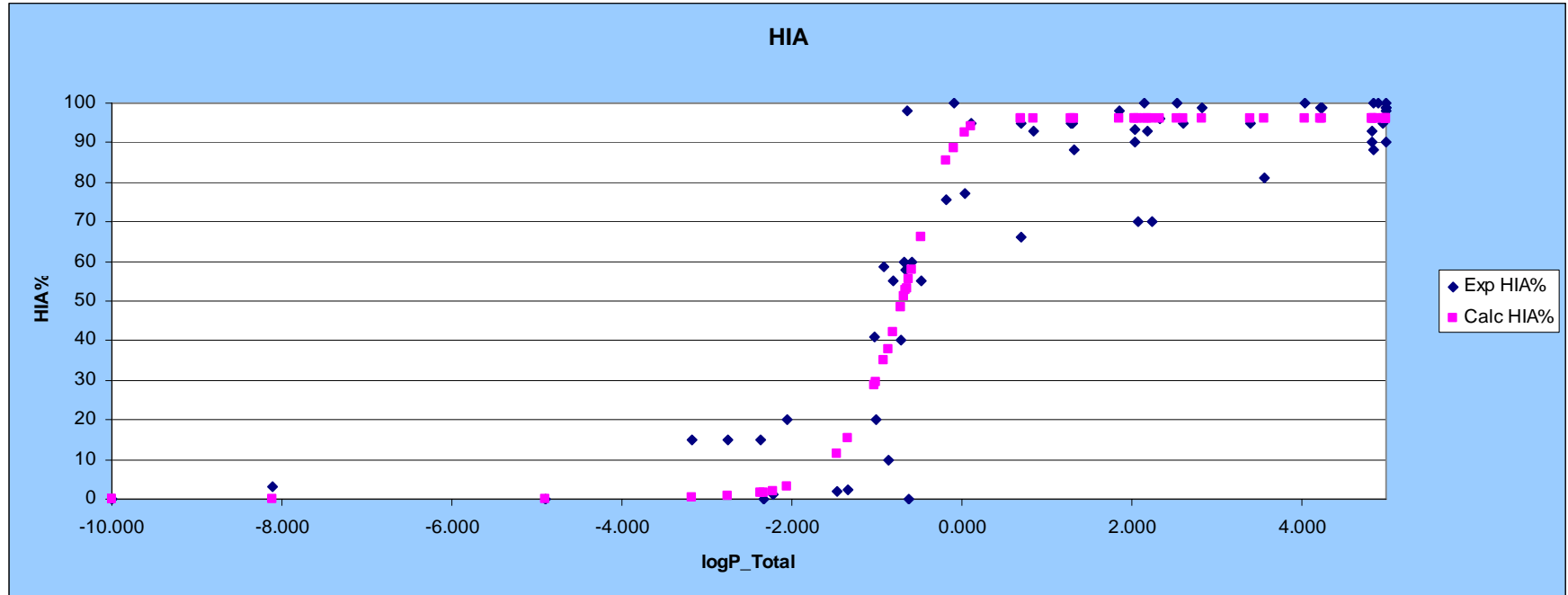


Pharma Algorithms, Inc.



- **Transcellular:** Ionic species assigned negligible permeability i.e. All $\Delta_i = -10$
- LogP of ampholytes (i.e. $F_{\pm} > 0.1$) reflects partition of the zwitterion at the isoelectric point and can be corrected with an indicator variable:
$$\text{Log } k_{\text{Trans}}^0 = \text{log } P_{\text{octanol}} - 0.9N_{\text{HD}} + 4.0I_{\text{ampholyte}} + 3.51$$
- **Paracellular:** Relative paracellular permeability of:
monoprotic-cations = 8, poly-anions = 0.2, all other species = 1
- N=61 Ampholytes: RMSE = 36 AAE = 26

Data Fitting 8a: Permeability correction for tetracyclines and PEPT1 substrates



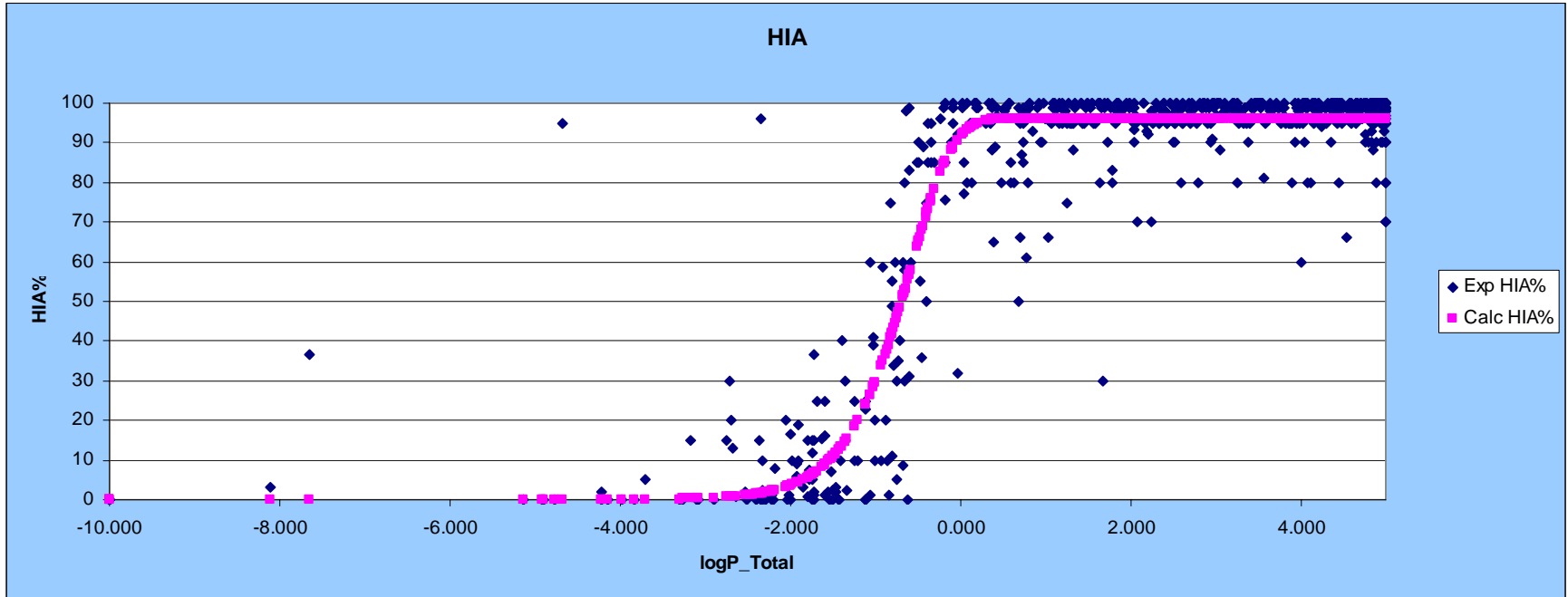
- **Transcellular:** Ionic species assigned negligible permeability i.e. All $\Delta_i = -10$

Additional indicator variables added for tetracyclines and PEPT1 substrates:

$$\text{Log } k_{\text{Trans}}^0 = \text{log } P_{\text{octanol}} - 0.9N_{\text{HD}} + 4.0I_{\text{amph}} + 6.0I_{\text{tet}} + 5.0I_{\text{PEPT1}} + 3.51$$

- **Paracellular:** Relative paracellular permeability of:
monoprotic-cations = 8, poly-anions = 0.2, all other species = 1
- N=61 Ampholytes: RMSE = 13.8 AAE = 11

Data Fitting 8b: Full Training Set



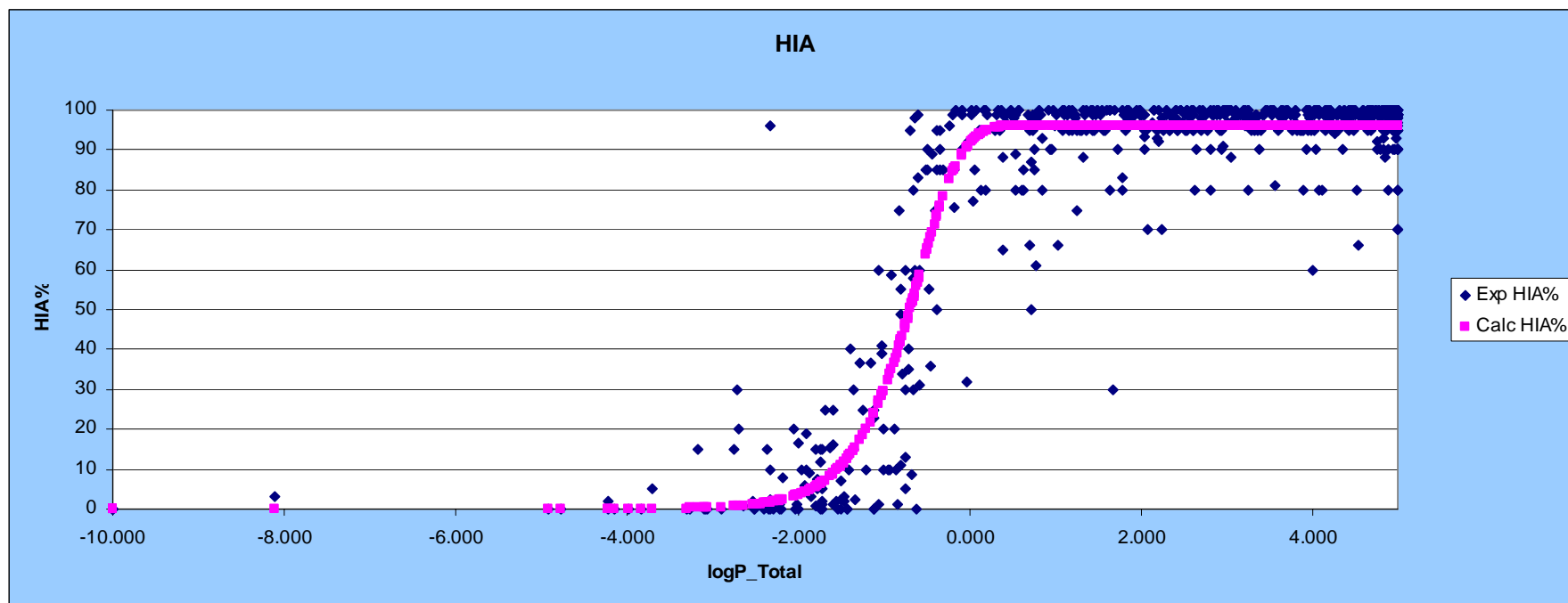
- **Transcellular:** Ionic species assigned negligible permeability i.e. All $\Delta_i = -10$
- $\text{Log } k_{\text{Trans}}^0 = \log P_{\text{octanol}} - 0.9N_{\text{HD}} + 4.0I_{\text{amph}} + 6.0I_{\text{tet}} + 5.0I_{\text{PEPTI}} + 3.51$
- **Paracellular:** Relative paracellular permeability of:
monoprotic-cations = 8, poly-anions = 0.2, all other species = 1
- N=702 Full training set: RMSE = 11.1 AAE = 6.2

Data Fitting 9: Final Optimisation

(Simultaneous gradient minimisation of all coefficients)



Pharma Algorithms, Inc.

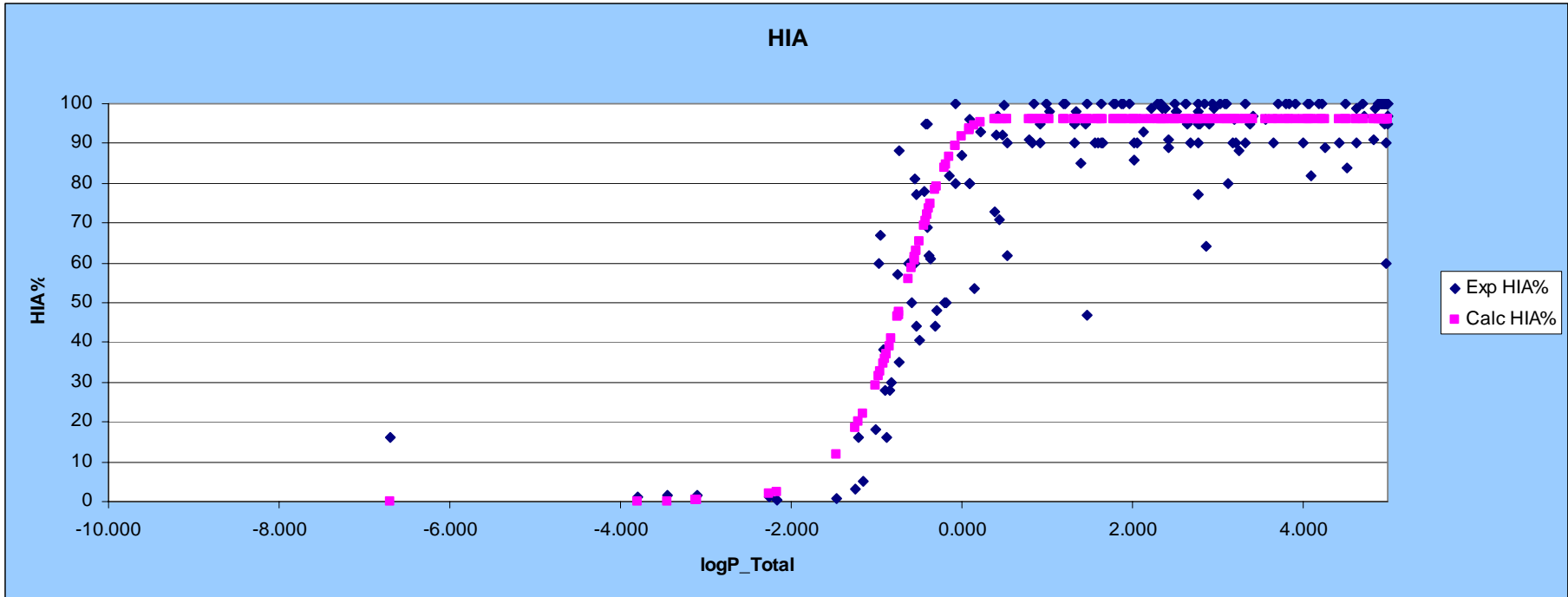


- **Transcellular:** $\log k_{\text{Trans}}^i = \log k_{\text{Trans}}^0 + \Delta_i$
Mono-ions: $\Delta_- = -5.5$, $\Delta_+ = -3.25$, $\Delta_{\pm} = -4$
Permanently charged cations have extra $\Delta_+ = -8.5$

$$\text{Log } k_{\text{Trans}}^0 = \log P_{\text{octanol}} - 0.9N_{\text{HD}} + 4.0I_{\text{amph}} + 6.0I_{\text{tet}} + 5.0I_{\text{PEPT1}} + 3.51$$

- **Paracellular:** Relative paracellular permeability of:
monoprotic-cations = 8, poly-anions = 0.2, all other species = 1
- N=702 Full Training Set: RMSE = 10.6 AAE = 6.1

Data Fitting 10: Test Set

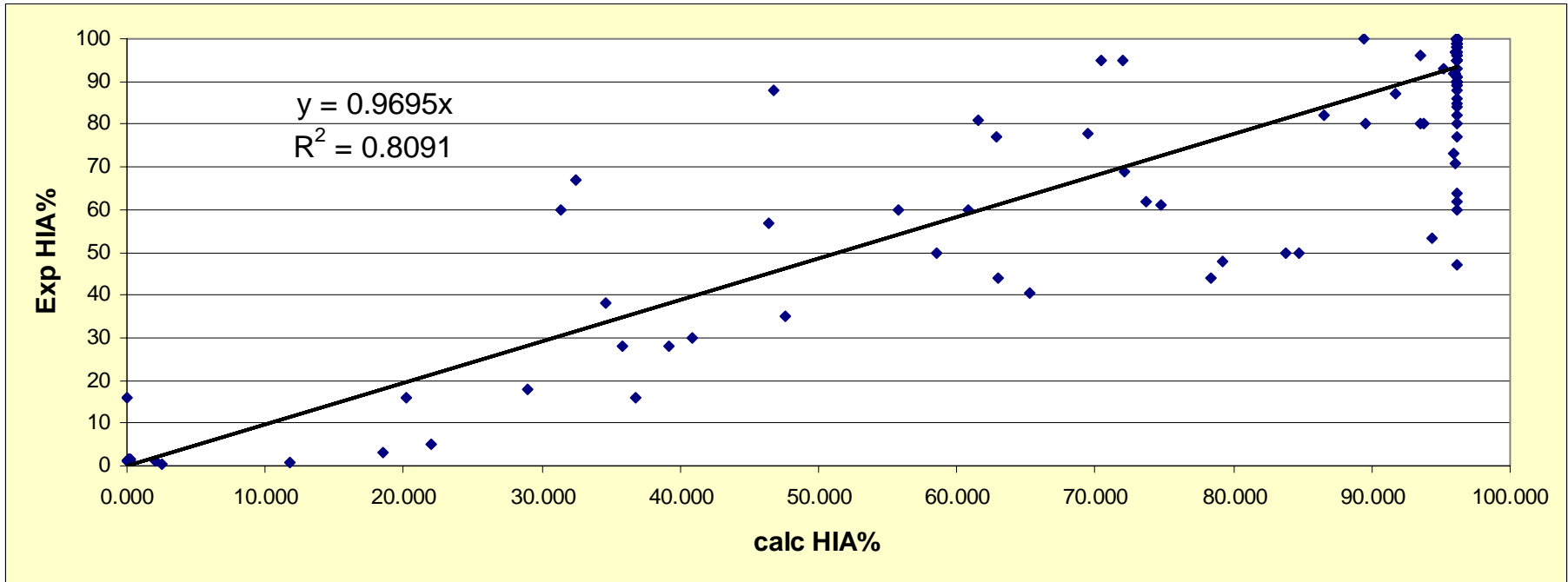


- **Transcellular:** $\log k_{\text{Trans}}^i = \log k_{\text{Trans}}^0 + \Delta_i$
Mono-ions: $\Delta_- = -5.5$, $\Delta_+ = -3.25$, $\Delta_{\pm} = -4$.

$$\text{Log } k_{\text{Trans}}^0 = \log P_{\text{octanol}} - 0.9N_{\text{HD}} + 4.0I_{\text{amph}} + 6.0I_{\text{tet}} + 5.0I_{\text{PEPTI}} + 3.51$$

- **Paracellular:** Relative paracellular permeability of:
monoprotic-cations = 8, poly-anions = 0.2, all other species = 1
- N=169 Abraham Test Set: RMSE = 12.4 AAE = 8.1

Data Fitting 10: Test Set



- Test Set this work: $N=169$, $r^2 = 0.81$, $RMSE = 12.4$
- For the same set Abraham [1] obtained his best statistics from a non-linear model:

$$\% HIA = 100 / [1 + 10^{-(1.02 + 0.062E + 0.098S - 0.60A - 0.68B + 0.45V)}]$$

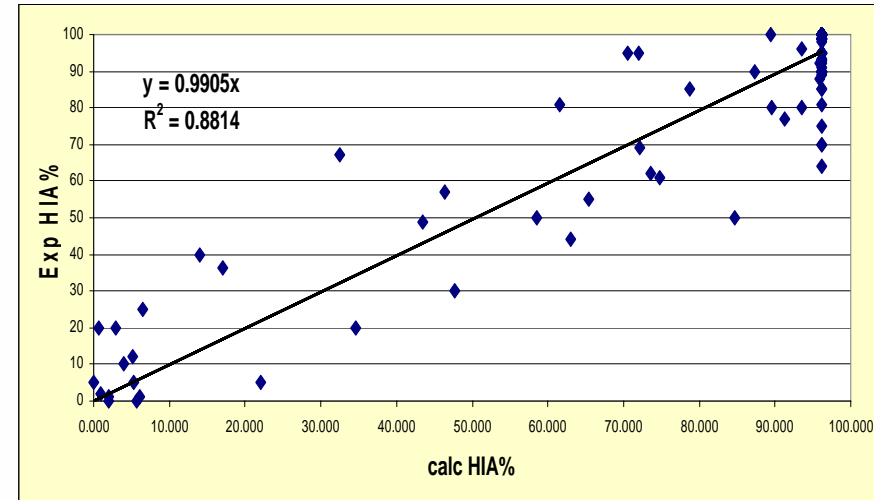
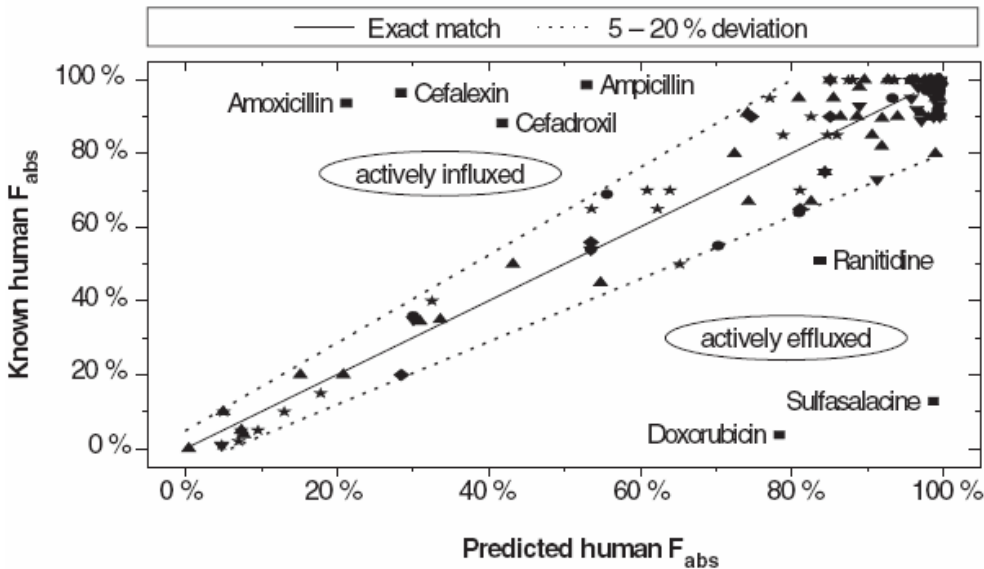
$N=169$, $r^2 = 0.79$, $RMSE = 13$

[1] Zhao, Y. H., M. H. Abraham, J. Le, A. Hersey, C. N. Luscombe, G. Beck, B. Sherborne and I. Cooper (2002) Pharm Res **19**(10): 1446-57.

Comparison with an Alternative Mechanistic Model



Pharma Algorithms, Inc.



- 126 compounds previously modelled by Willmann et. al.
- The Willmann model [1,2] is based on molecular weight and measured values for membrane retention. A term for transcellular and a term for paracellular absorption were included
- Doxorubicin and Sulfasalazine are highly ionised acidic molecules and are not outliers in our model

[1] Willmann, S., J. Lippert and C. Schmitt (2005). *Expert Opin. Drug Metab. Toxicol.* **1**(1): 159-168

[2] Willmann, S., W. Schmitt, J. Keldenich et. al. (2004). *J Med Chem* **47**(16): 4022-31

Outliers

(original predictions in green, modified descriptors and predictions in blue)



Pharma Algorithms, Inc.

	Name	Molecular Descriptor	Predicted %HIA	Observed %HIA
	Meropenem pKa1=3.2 pKa2=8.8	AABlogP=0.05	56%	0%
		Measured logP = -1.7 (Distribution coeff at iso-electric pH)	8%	
	Ethyl Biscoumacetate AABlogP=1.6 c.f. Warfarin pKa=4.8	pKa1=2.8 pKa2=5.8 F⁰=0.000	2%	96%
		F⁰=0.001	96%	

Better Descriptors?

– Transcellular pathway

- Octanol supports hydrogen bonding and is a poor model for passive diffusion across the lipid core of a membrane. Alkane/water partition should be better.
- N_{HD} is a crude correction factor. It does not take into account H-bond strength e.g. presence of intramolecular H-bonds
- Ampholytes are not well described. There are few examples where the pKa microconstants and the logP of the true neutral form (rather than the zwitterion/neutral mixture) been determined by experiment.

– Paracellular pathway

- McGowan's volume is used to estimate molecular radius. It does not take into account shape. Thin flexible molecules (e.g. PEG's) have been reported [1] to have higher than expected paracellular permeability
- Alternative equations have been proposed [2] to model the size dependence of paracellular permeation

[1] Artursson, P., A. L. Ungell and J. E. Lofroth (1993). *Pharm Res* **10**(8): 1123-9.

[2] Willmann, S., W. Schmitt, J. Keldenich et. al. (2004). *J Med Chem* **47**(16): 4022-31



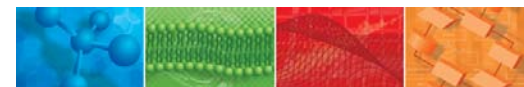
Better Data?

– Available Data is ‘Skewed’

- There are many BCS type 2 compounds (i.e. lipophilic and low solubility) and relatively few ‘parameter critical’ compounds.
- Not enough diverse lipophilicity/pKa combinations e.g. only one compound existing largely as a double zwitterion (2+2-) i.e. Lisinopril
- Only two compounds dominate the final optimisation that defines the eventual values for Δ_i for anions and cations i.e. Carbenoxolone and Azithromycin

$$\log k_{\text{Trans}}^i = \log k_{\text{Trans}}^o + \Delta_i$$

Lipophilic ions (negligible absorption of neutral form predicted)



Pharma Algorithms, Inc.

	Name (physchem data)	Model parameter value	Predicted %HIA	Observed %HIA
	Carbenoxolone AABlogP=5.6 pKa1=4.3 pKa2=5.0 At pH7 $F^{-1}=0.01$ $F^{-2}=0.99$	$\Delta_- = -10$	0%	95%
		$\Delta_- = -5.5$	85%	
	Azithromycin AABlogP= 1.98 pKa1=9.0 pKa2=8.3 At pH7 $F^{+1}=0.04$ $F^{+2}=0.96$	$\Delta_+ = -10$	0%	36.5%
		$\Delta_+ = -3.25$	35%	

Measured Data?

- **The current model is built on predicted values for pKa and logP**
 - We have database values of $\log P_{\text{octanol}}$ for about half of the compounds
 - There is a good correspondence between measured logP and AABlogP
 - Reported experimental values can differ by over 1 log unit (173 compounds have 2 or more values available and for 17 of these the difference is over 1 log unit)
 - For compounds containing acidic and basic groups an accurate estimate of neutral form fraction requires determination of the pKa microconstants

Is this a true passive absorption model?



- Clinical absorption values can be affected by other mechanisms e.g. active uptake, P-glycoprotein efflux, solubility and formulation dependent factors
- PAMPA models support the idea of negligible transcellular diffusion of ions
- The pH dependent permeability of cationic drugs in Caco-2 monolayers has been analysed and compared to PAMPA [1]. The evidence supported values for permeability of the cations ca.1000 to 10,000 fold lower than that of the neutral molecules by a mechanism other than paracellular diffusion. The mechanism was not elucidated but active transport was suspected.
- Parameters in our model may well be influenced by other ways in which ions can cross membranes e.g. ion channels and ion transporters.

[1] Avdeef, A., P. Artursson, S. Neuhoff, L. Lazorova, J. Grasjo and S. Tavelin (2005). "Caco-2 permeability of weakly basic drugs predicted with the double-sink PAMPA pKa(flux) method." Eur J Pharm Sci **24**(4): 333-49.

Summary



- The HIA model uses five descriptors calculated from structure:
Molecular size, N_{HD} , $\log P$, pK_a (Acid), pK_a (Base)
- Model outputs can include:
 - An overall intestinal absorption rate constant (k_a) for the clinical situation
 - An estimate of maximum passive absorption and the relative contributions of the transcellular and paracellular routes

Conclusion



They may not be in the Lipinski Rule!

- MWt is over 500
- Log P is over 5
- More than 5 H-bond donors
- More than 10 H-bond acceptors

but

Ionisation and pKa are important