



## Fundamentals of Multiprotic Ionization and Ionization Modeling

*Robert Fraczkiwicz  
Simulations Plus, Inc.*

*The Ionization Prediction Summit  
October-November, 2023*



# “You Must Unlearn What You Have Learned”:

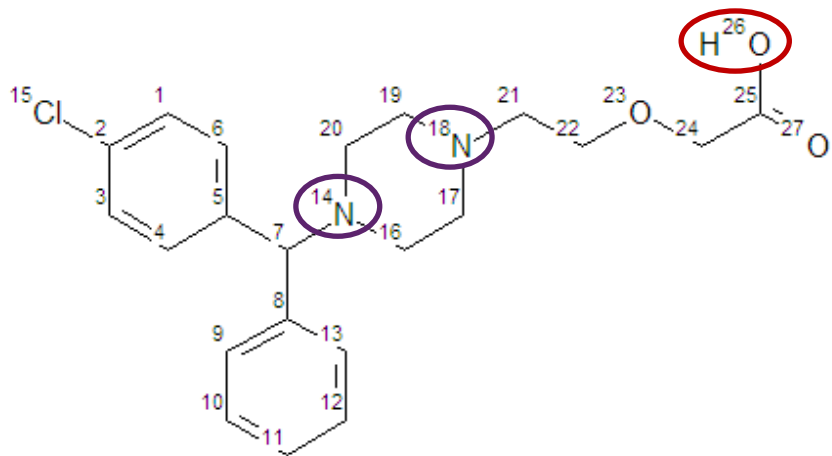


“No! No different. Only different in your mind... You must unlearn what you have learned.”

*Master Yoda to Luke Skywalker in the swamps of Dagobah.*

# Clearing Up Myths About Aqueous Ionization of Drugs

**Myth #1: apparent  $pK_a$  can always be “assigned” to a functional group**



2.10

3.01

8.17

Marosi A, Kovacs Z, Beni S,  
Kokosi J and Noszal B.  
European Journal of  
Pharmaceutical Sciences, 37:  
321-328, 2009.

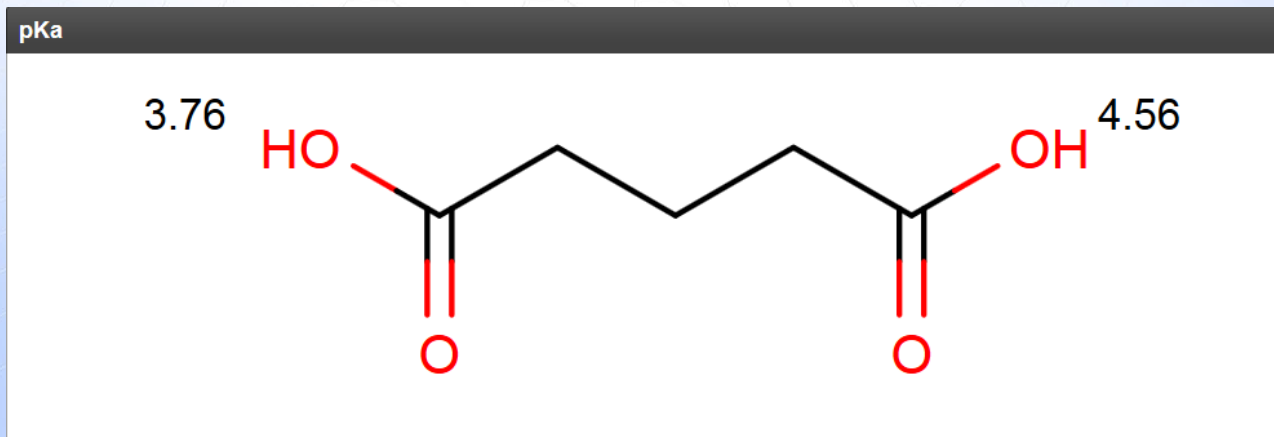
# Glutaric acid example

Measured apparent  $pK_a$ :

5.42  
4.35

German, W. & Vogel, A.,  
J. Am. Chem. Soc., 58, 1546 (1936)

How one popular program predicts and reports  $pK_a$  for glutaric acid:



# Futility of “assignments” – another example

Apparent  $pK_a$

1.40

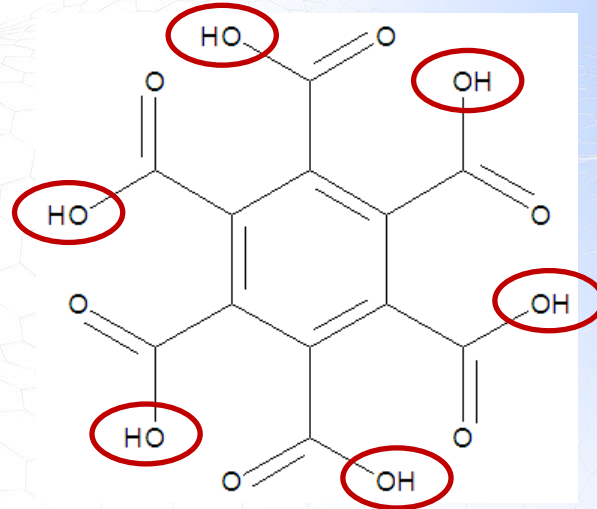
2.19

3.31

4.78

5.89

6.96

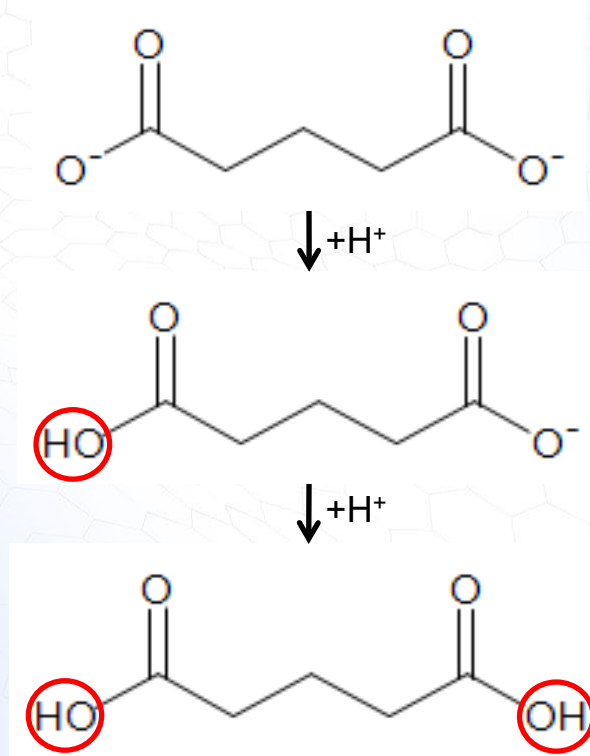


**“Assign”? How?**

Maxwell W & Partington J. *Trans Farad Soc.* **31**, 922 (1935)

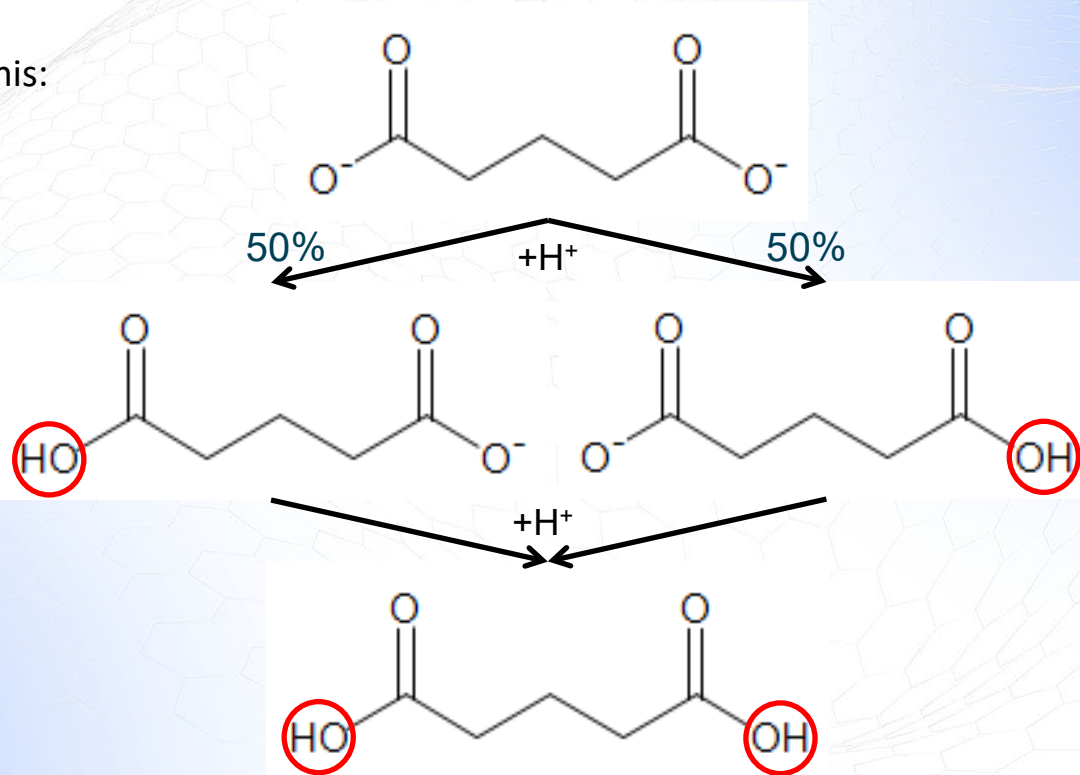
# How do polyprotic molecules protonate?

Like this? :



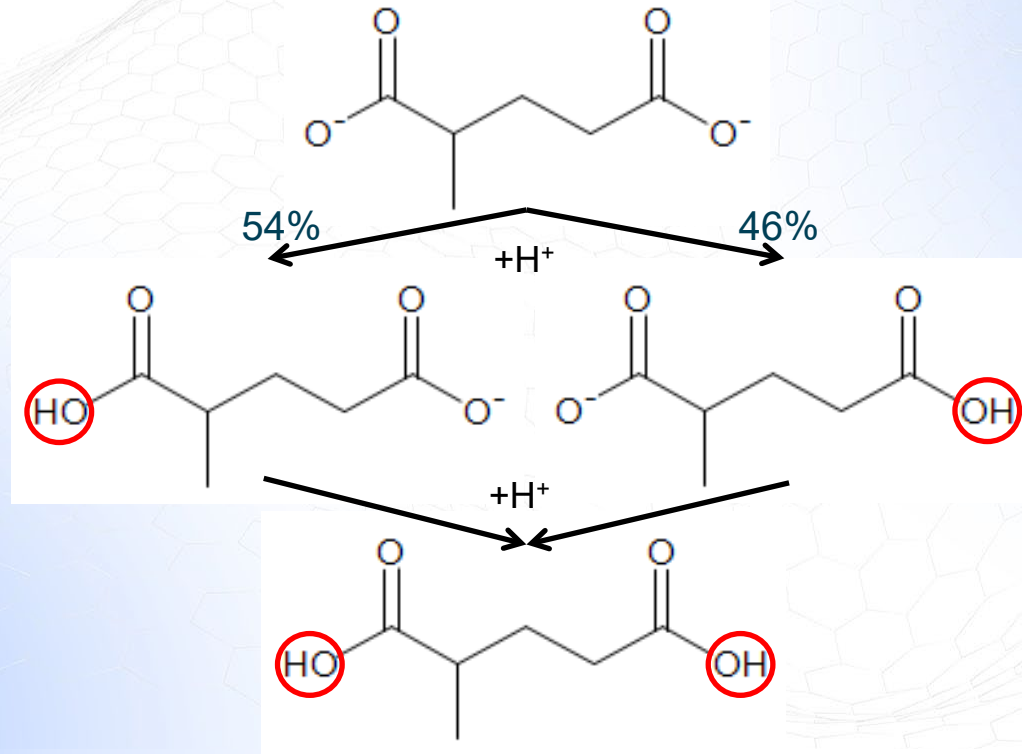
# How do polyprotic molecules protonate?

No! Like this:



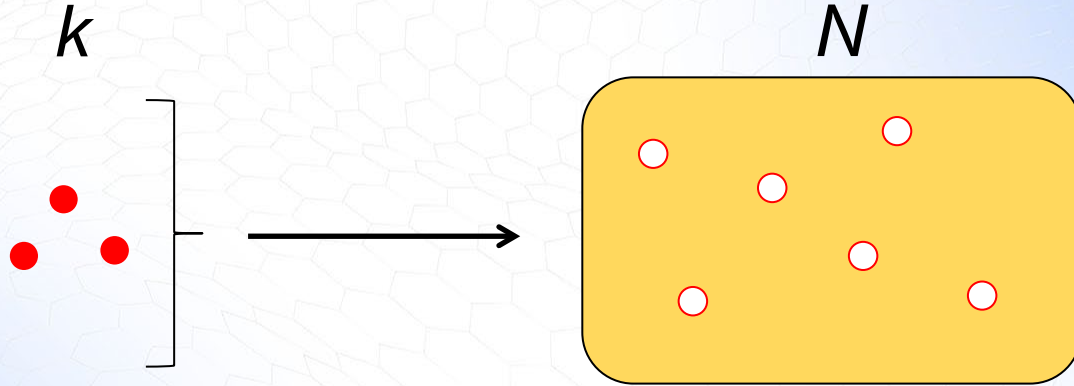


# How do polyprotic molecules protonate?



# It's a simple combinatorial problem

Distribute  $k$  protons among  $N$  sites;  $0 \leq k \leq N$



$$\binom{N}{k} \text{ distinct combinations}$$

**N=3**

**Macrostates:**

knowing there  
are  $k$  protons  
somewhere

**Microstates:**

knowing  
where the  $k$   
protons  
exactly are

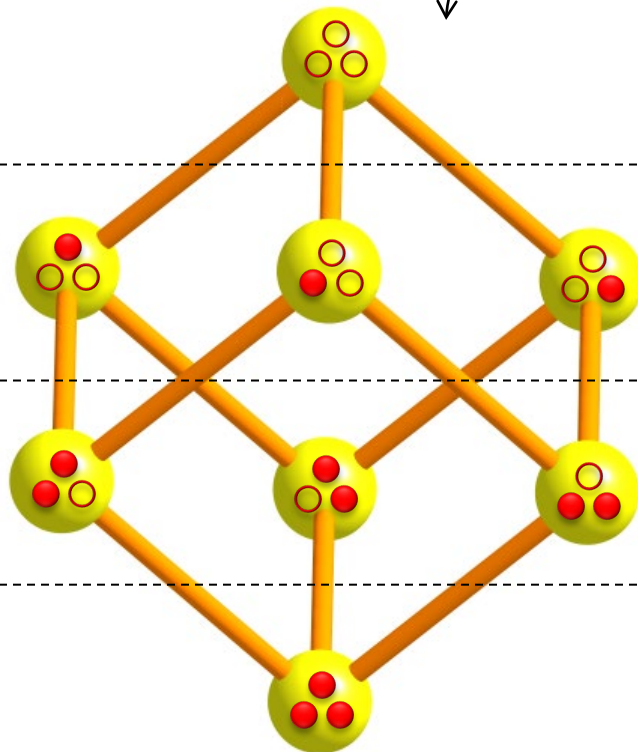
$k$

**0**

**1**

**2**

**3**



# Cetirizine, N=3

Macrostates

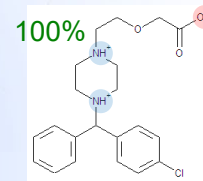
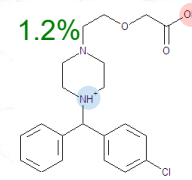
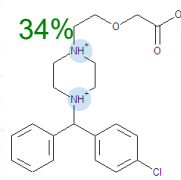
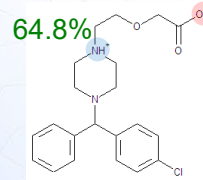
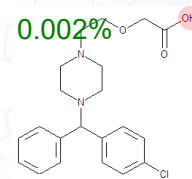
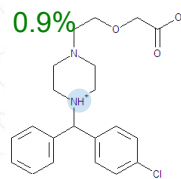
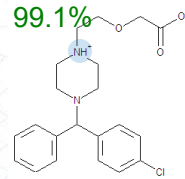
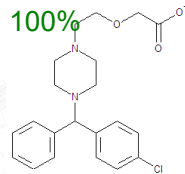
Microstates

0 Cet

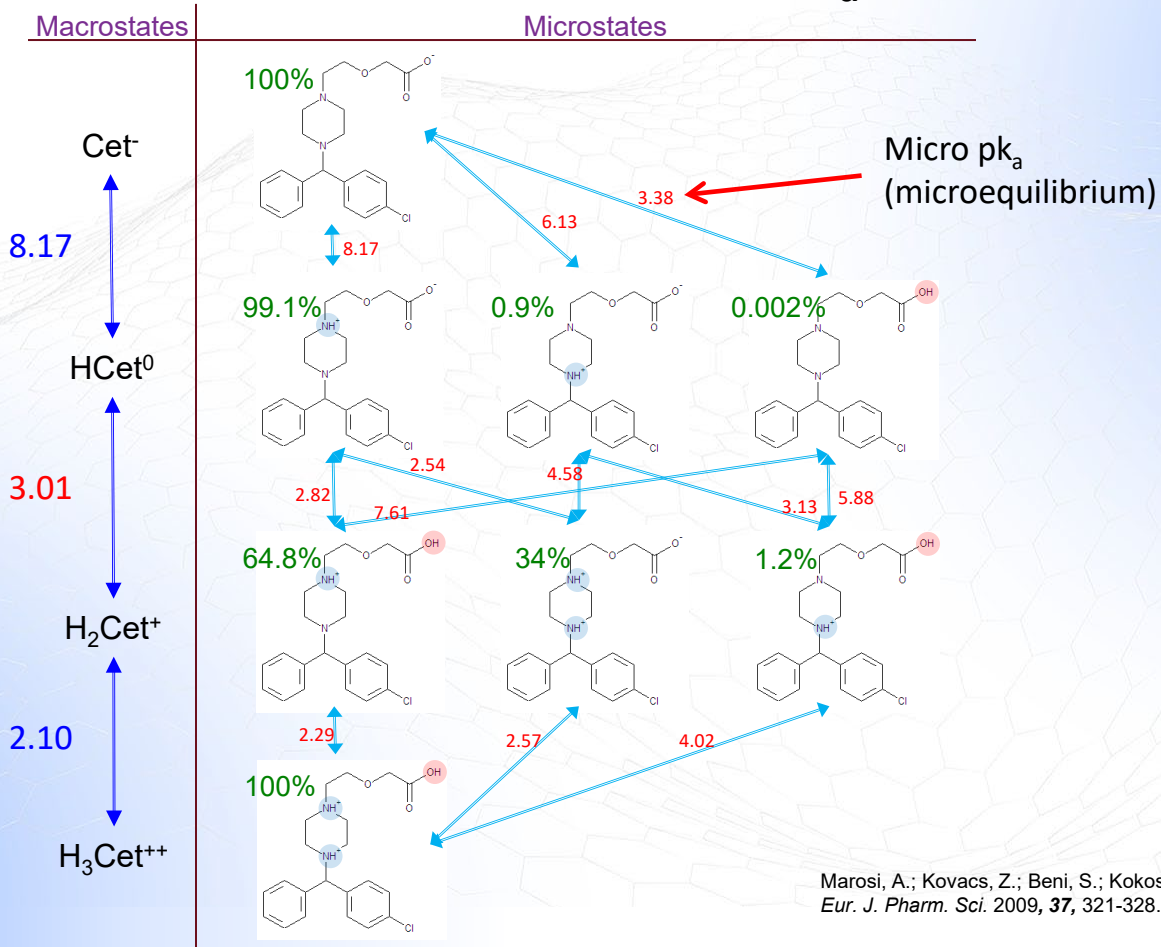
1 HCet<sup>0</sup>

2 H<sub>2</sub>Cet<sup>+</sup>

3 H<sub>3</sub>Cet<sup>++</sup>

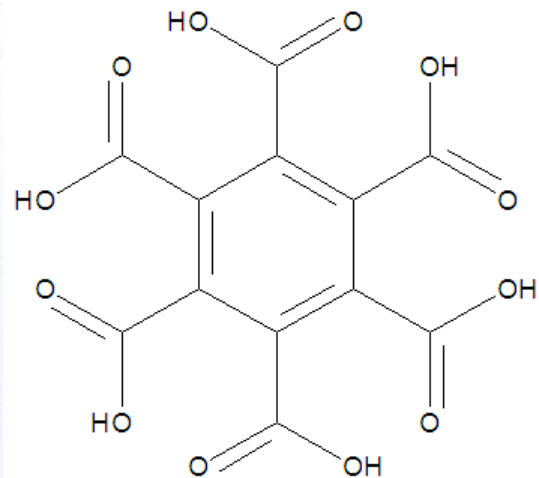


# Cetirizine, N=3, with pK<sub>a</sub>



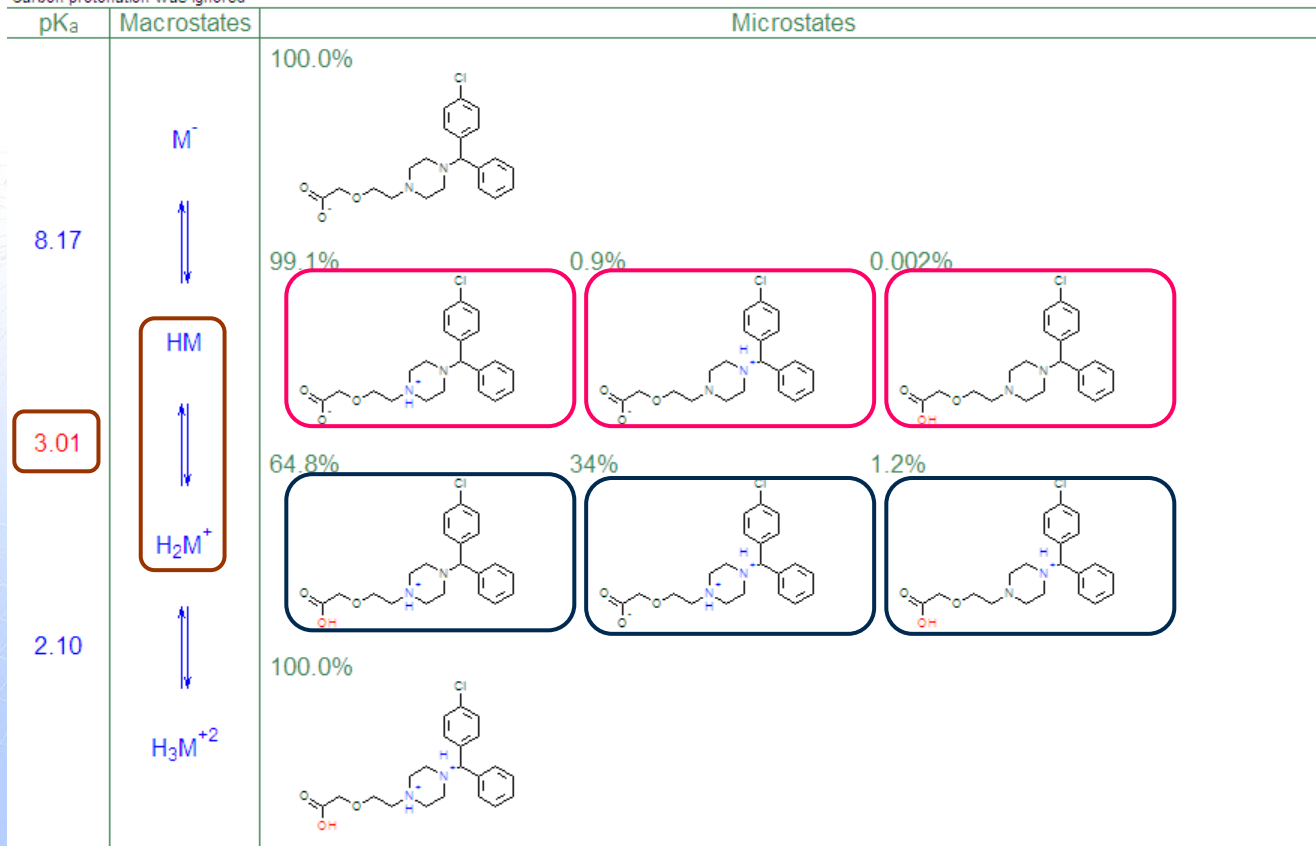
# Mellitic Acid, N=6

pK <sub>a</sub>	Macrostates	Microstates																									
6.66	M <sup>6-</sup>	100.0%																									
	HM <sup>5-</sup>	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%								
5.37	H <sub>2</sub> M <sup>4-</sup>	11.4%	11.4%	11.4%	7.4%	7.4%	7.4%	7.4%	7.4%	7.4%	3.5%	3.5%	3.5%	3.5%	3.5%	3.5%	3.5%	3.5%	3.5%								
	H <sub>3</sub> M <sup>3-</sup>	8.7%	8.7%	6.0%	6.0%	6.0%	6.0%	6.0%	6.0%	6.0%	6.0%	6.0%	6.0%	6.0%	6.0%	6.0%	6.0%	1.8%	1.8%	1.8%	1.8%	1.8%	1.8%	1.8%	1.8%	1.8%	1.8%
2.97	H <sub>4</sub> M <sup>2-</sup>	10.6%	10.6%	10.6%	7.7%	7.7%	7.7%	7.7%	7.7%	7.7%	7.7%	7.7%	7.7%	7.7%	7.7%	7.7%	7.7%	3.6%	3.6%	3.6%	3.6%	3.6%	3.6%	3.6%	3.6%	3.6%	3.6%
	H <sub>5</sub> M <sup>-</sup>	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%
1.56	H <sub>6</sub> M	100.0%																									

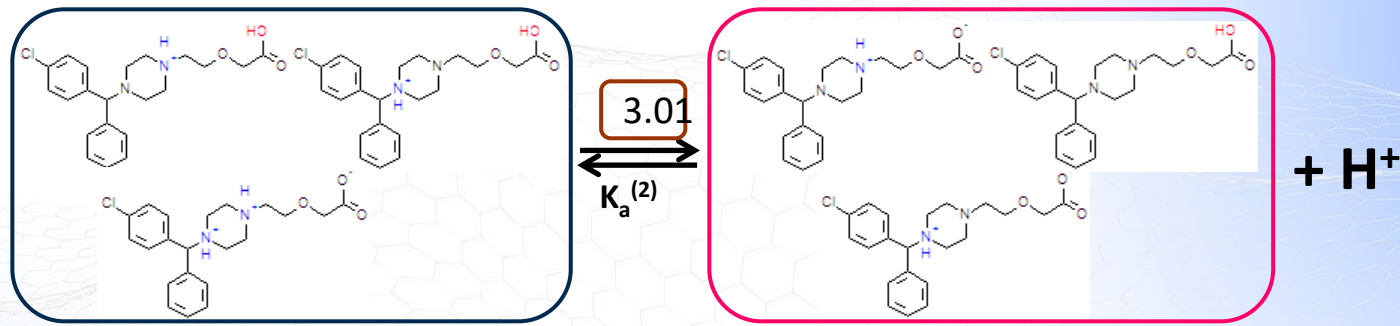


pKa Table for Cetirizine.mol  
 1 acidic atoms: 26(-OH)  
 2 basic atoms: 14(>N-)18(>N-)  
 Aliphatic -OH groups were ignored  
 Aliphatic amides were ignored  
 Carbon protonation was ignored

## A microscopic/thermodynamic view of cetirizine ionization



Apparent  $pK_a$  is NOT a property of a single ionizable group



$H_2M^+$

$HM$

$$K_a^{(2)} = \frac{\left[ \begin{array}{c} \text{Cl-Ph-N(CH}_2\text{)}_2\text{-N}^-\text{H-CH}_2\text{CH}_2\text{O-CH}_2\text{COO}^- \\ \text{Ph} \end{array} \right] + \left[ \begin{array}{c} \text{Cl-Ph-N(CH}_2\text{)}_2\text{-N}^-\text{H-CH}_2\text{CH}_2\text{O-CH}_2\text{COOH} \\ \text{Ph} \end{array} \right] + \left[ \begin{array}{c} \text{Cl-Ph-N(CH}_2\text{)}_2\text{-N}^-\text{H-CH}_2\text{CH}_2\text{O-CH}_2\text{COO}^- \\ \text{Ph} \end{array} \right]}{\left[ \begin{array}{c} \text{Cl-Ph-N(CH}_2\text{)}_2\text{-N}^-\text{H-CH}_2\text{CH}_2\text{O-CH}_2\text{COOH} \\ \text{Ph} \end{array} \right] + \left[ \begin{array}{c} \text{Cl-Ph-N(CH}_2\text{)}_2\text{-N}^-\text{H-CH}_2\text{CH}_2\text{O-CH}_2\text{COOH} \\ \text{Ph} \end{array} \right] + \left[ \begin{array}{c} \text{Cl-Ph-N(CH}_2\text{)}_2\text{-N}^-\text{H-CH}_2\text{CH}_2\text{O-CH}_2\text{COO}^- \\ \text{Ph} \end{array} \right]} [H^+]$$



# Clearing Up Myths About Aqueous Ionization of Drugs

**Myth #2: apparent  $pK_a$  can unambiguously be labeled as either "acidic", or "basic" every time**

pK <sub>a</sub>	Macrostates
	M <sup>-</sup>
9.01	↕
	HM
2.44	↕
	H <sub>2</sub> M <sup>+</sup>

Acid

Base

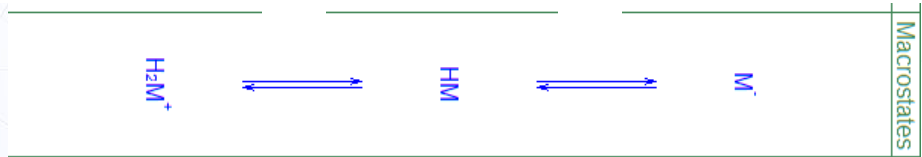
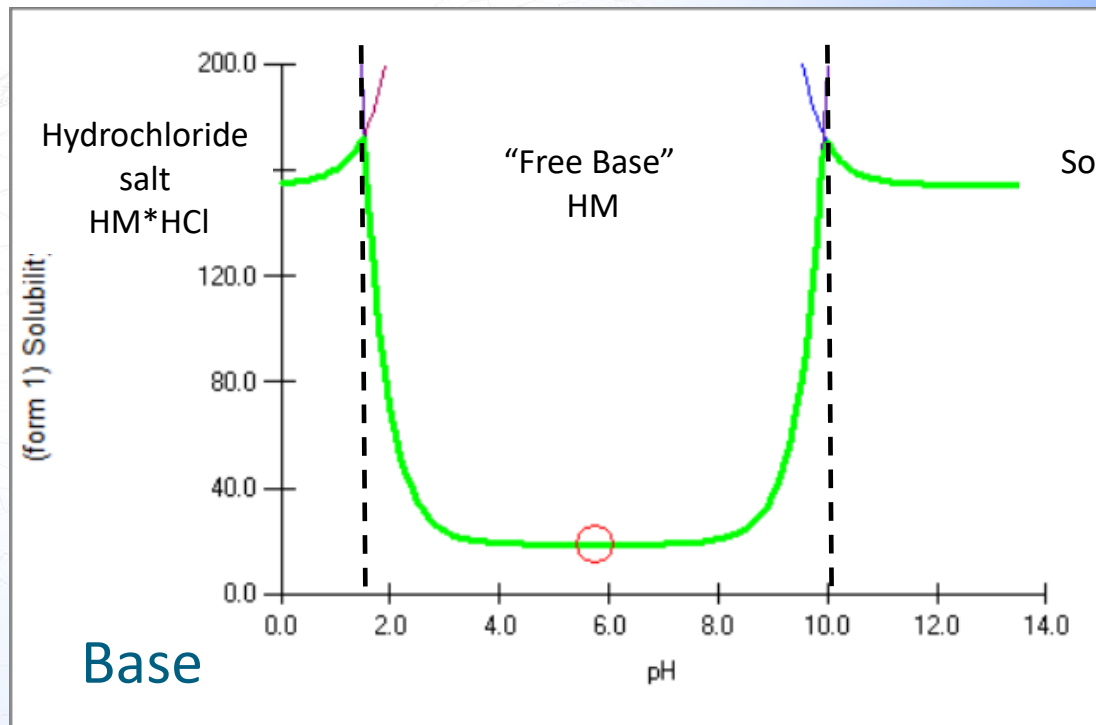
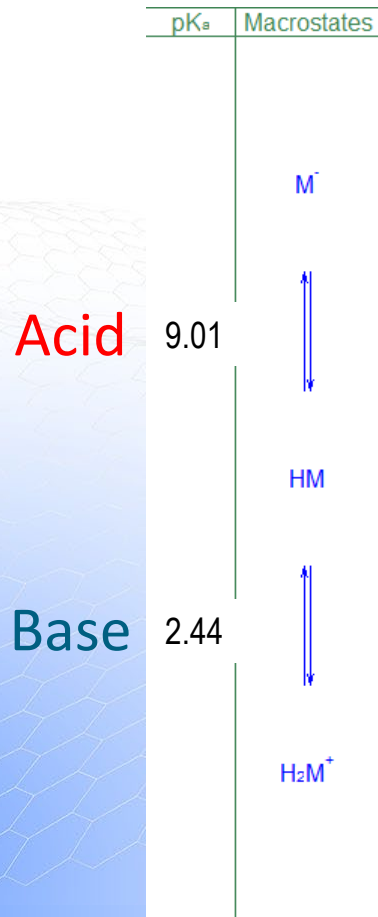
A "physical chemist" point of view: focus on the whole molecule and its charge states (macroscopic)

Proton loss = Acid

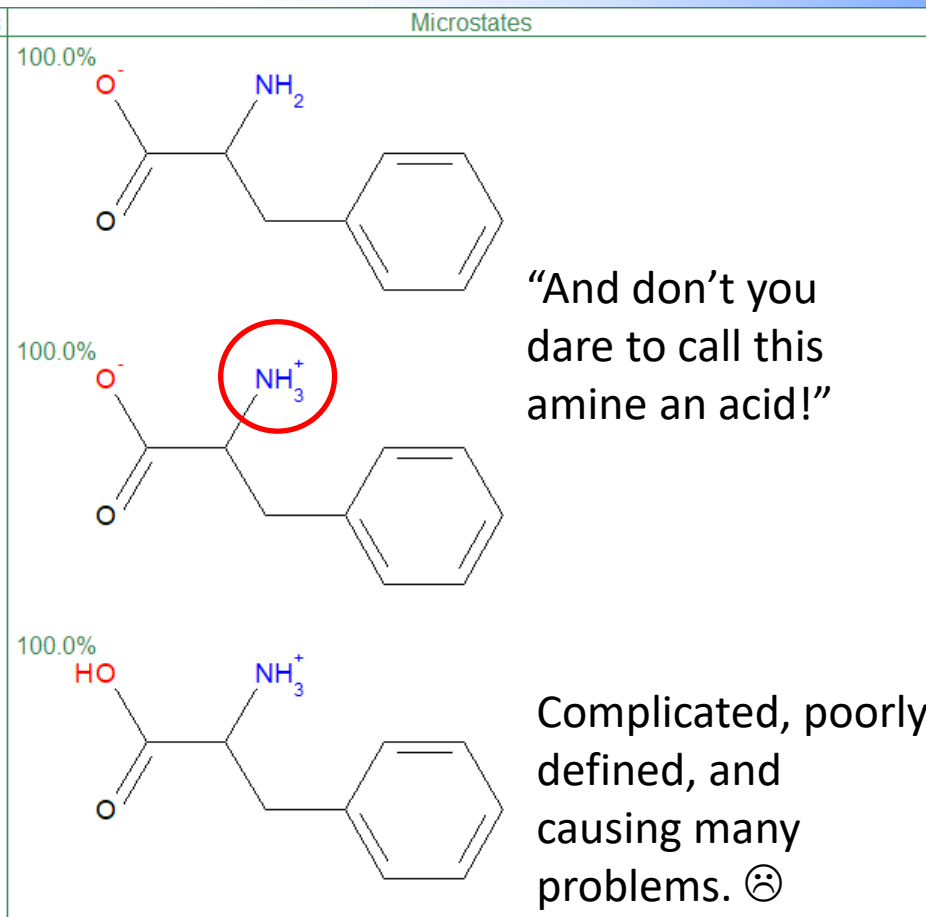
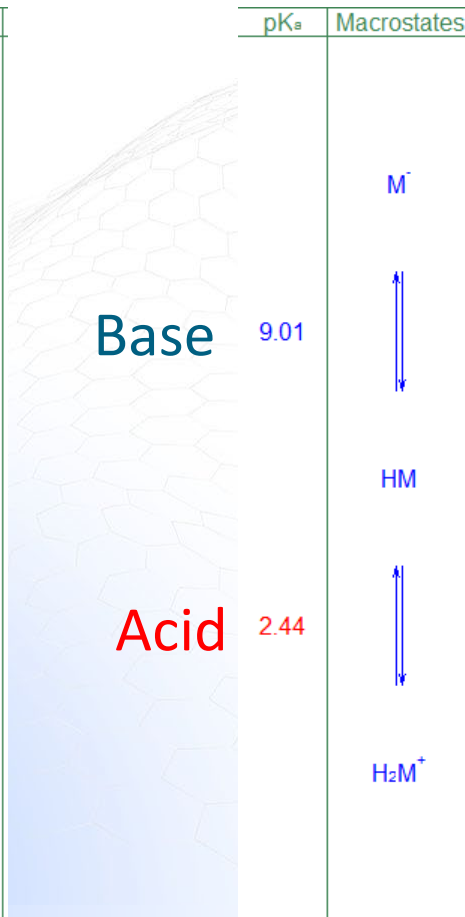
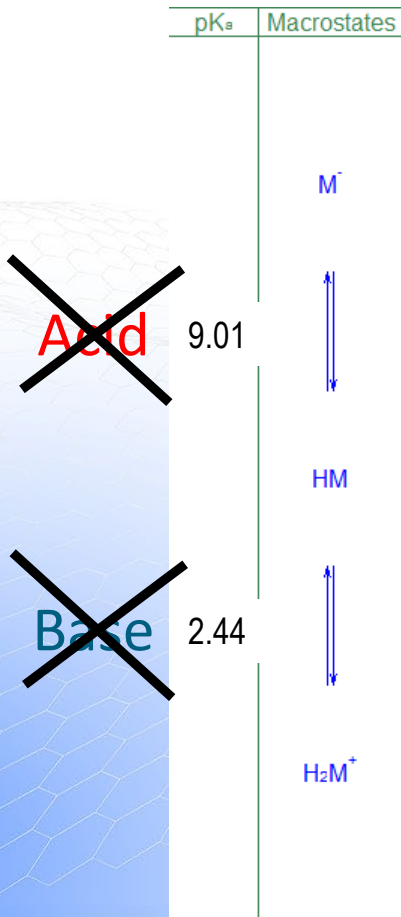
Proton gain = Base

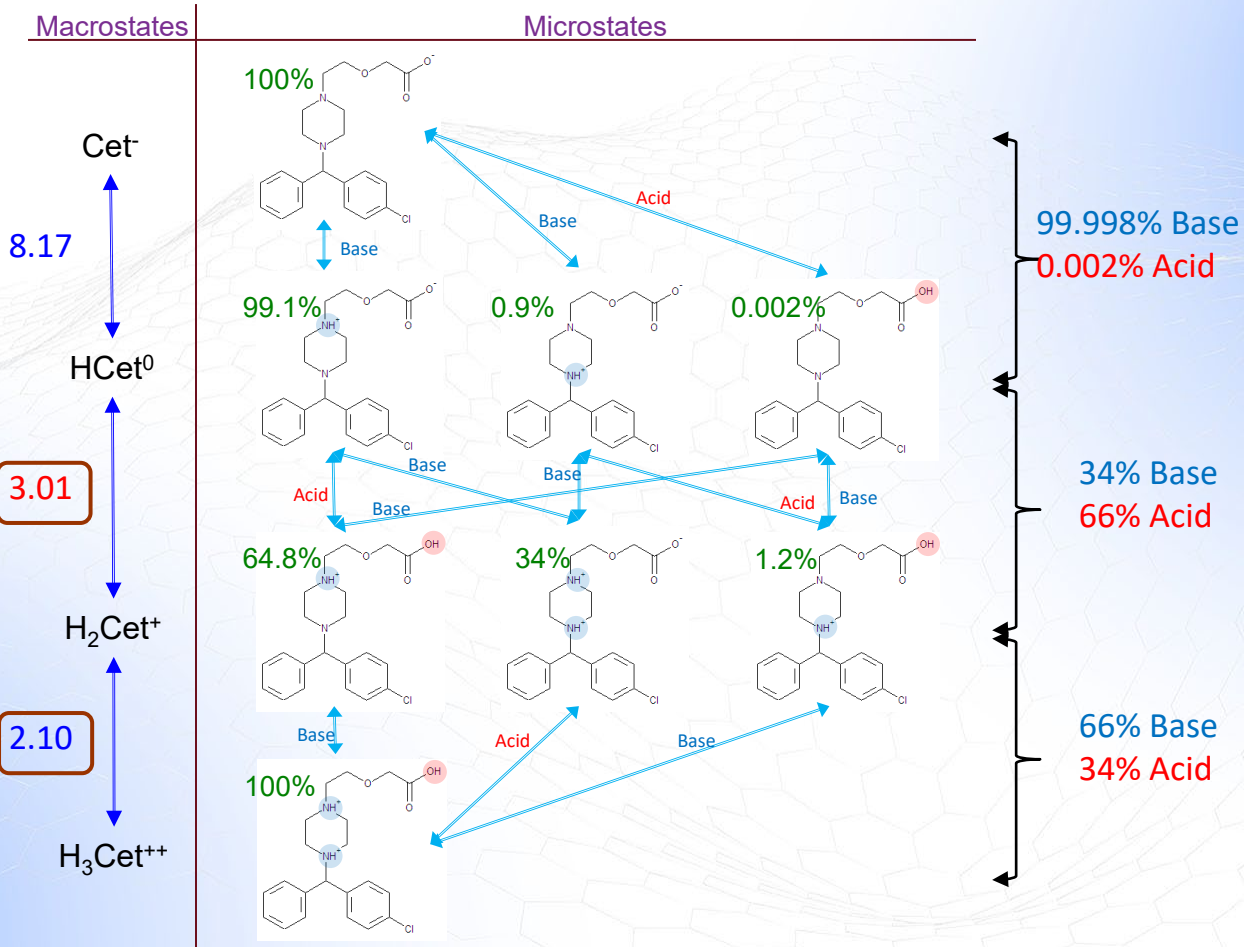
Simple, unambiguous, and accurate!

The physical chemist point of view is exactly reflected in, e.g., pH dependence of solubility



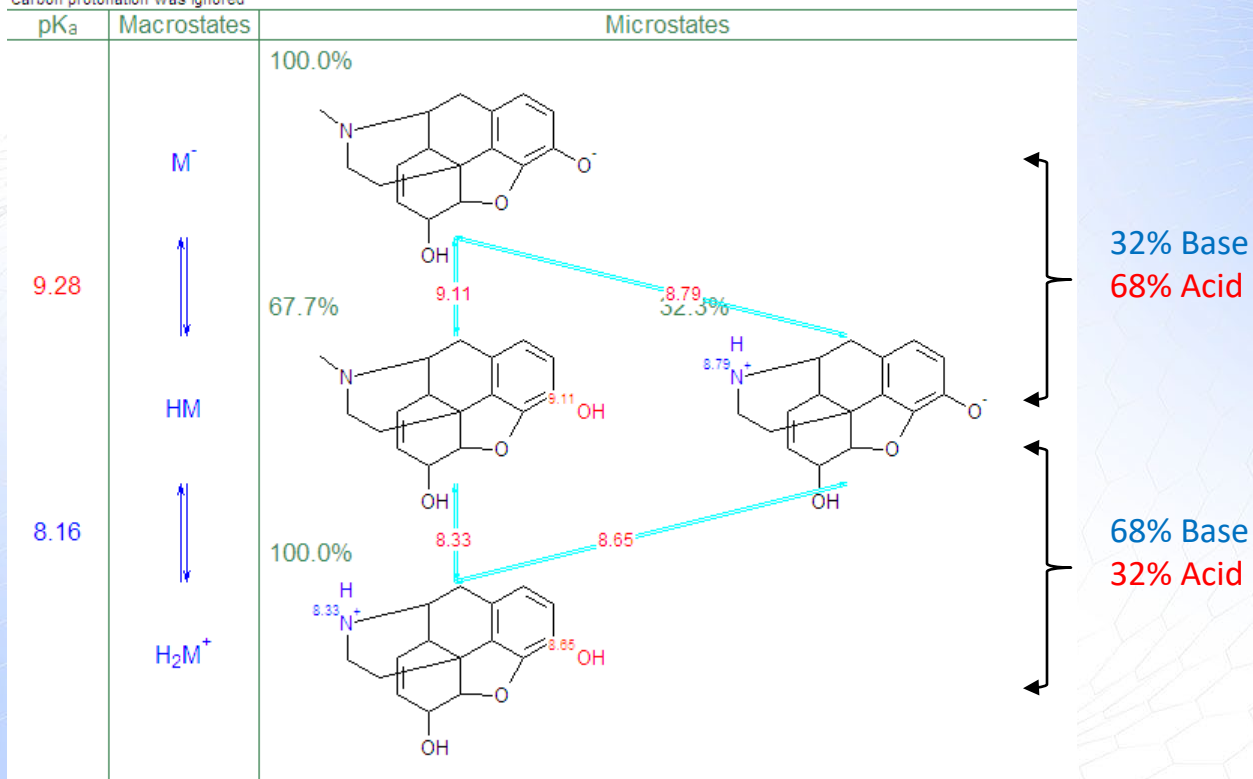
An "organic chemist" point of view: focus on individual functional groups (microscopic)





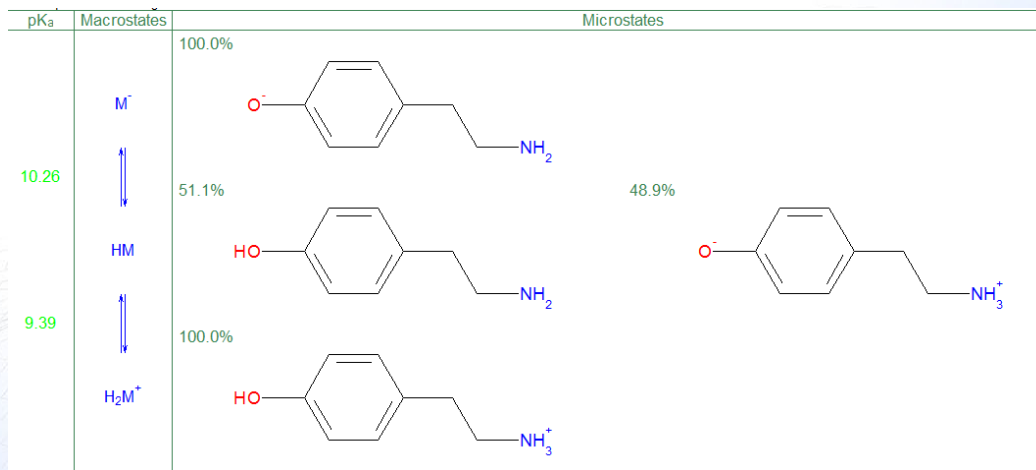
# Apparent and microscopic $pK_a$ for morphine

pKa Table for Morphine.mol  
 1 acidic atoms: 20(-OH)  
 1 basic atoms: 2(>N-)  
 Aliphatic -OH groups were ignored  
 Aliphatic amides were ignored  
 Carbon protonation was ignored



[Experimental data from Mazák K, Noszál B. Poster presented at the LogP2009 Symposium, Zürich, Switzerland, 2009]

# What is mixed pK<sub>a</sub>?



49% Base  
51% Acid

51% Base  
49% Acid

ADMET Predictor(TM) : Tyramine.mol (f:\pka\knownmicroconstants\)

File Batch Edit Calculate View Tools Help

Basic Modeler Settings Adv. Modeler Settings Ensemble Statistics Model Export

Molecular Data Prop./Desc. Histograms Prop./Desc. Correlations 4D Data Mining

Molecular Record Spreadsheet

MolFile	*molname	Orig.Order	S+Acidic_pKa	S+Mixed_pKa	S+Basic_pKa	DiffCo
Tyramine.mol 	Tyramine	1	None	10.26; 9.39	None	1.11

All User Inputs PChemBio Metabolism Toxicity Simulation Descriptors User Models ADMET Risk Global

Operation completed successfully. 138 properties and 2 user data columns. 1 records 341 descriptors 1:17 PM



**Questions?**