J Comput Aided Mol Des (2010) 24:613–616
DOI 10.1007/s10822-010-9325-9

# Tautomerism, Hammett $\sigma$, and QSAR

**Yvonne Connolly Martin**

Received: 23 February 2010 / Accepted: 5 March 2010 / Published online: 18 March 2010


**Abstract** A consideration of equilibrium model-based equations suggests that tautomeric equilibria do not markedly affect observed potency if the tautomer bound represents at least 50% of the compound in solution. Tautomeric equilibria can enhance or attenuate the correlation of potency with Hammett $\sigma$. Additionally, tautomeric equilibria can lead to a correlation of potency with $\sigma$ even in the absence of a correlation of binding with $\sigma$.

**Keywords** Tautomer · 2D QSAR · 3D QSAR · Hammett $\sigma$ · Hammett equation · Model-based equations · p$K_a$

**Abbreviations**

| | |
|---|---|
| log (1/$C$) | The relative concentration that produces a pre-established biological effect, the relative potency |
| $V_r$ | The volume of the macromolecular target |
| $V_a$ | The volume of the aqueous phase |
| $K_t$ | The equilibrium constant between the solution concentrations of the tautomer bound to the target and that of the other tautomer(s) |
| $K_r$ | The equilibrium constant between the concentration of the tautomer bound to the target and the solution concentration of the same tautomer |

## Introduction

How should one include tautomerism in an analysis of the quantitative relationship between the biological potency and physical properties of molecules? Are there circumstances for which one can ignore tautomerism? This communication will explore the insights that emerge from consideration of the simple equilibrium model shown in Fig. 1. This model assumes that compounds partition between two tautomeric states and that one tautomer binds to the macromolecular target. The approach follows the precedent of using compartment models to explore the effects of ionization on the relative observed potency of molecules [1–3].

Y. C. Martin (✉)
Martin Consulting, Waukegan, IL 60087, USA
e-mail: yvonnecmartin@comcast.net

## Derivation of the equation

If we assume that the total response measured, the relative potencies of the molecules, is proportional to the fraction of the total amount of compound that is bound, Eq. 1 results.

$$\log(1/C) = \log\left(\frac{V_r[M_r]}{V_r[M_r] + V_a([M_a] + [M_b])}\right). \quad (1)$$

In Eq. 1 $V_r$ is the volume of the biological target and $V_a$ is the volume of the aqueous compartment. Because the volume of the biological macromolecule is very small compared to the surrounding aqueous phase, a final assumption is that $V_r[M_r] \ll V_a([M_a] + [M_b])$. In other words, most of the added compound remains in the aqueous phase. Omitting that term produces Eq. 2:

$$\log(1/C) = \log\frac{1}{\frac{V_a}{V_r}\left(\frac{[M_a]}{[M_r]} + \frac{[M_b]}{[M_r]}\right)}. \quad (2)$$

The two equilibria shown in Fig. 1 are related to their respective equilibrium constants:

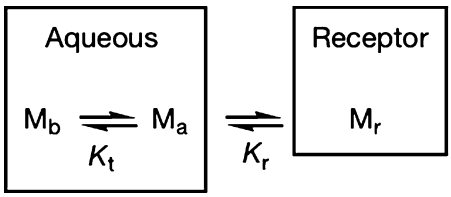


**Fig. 1** A model of the equilibrium of a molecule between two tautomeric states in water and one tautomer bound to the target biomolecule

$$K_t = \frac{[M_b]}{[M_a]} \tag{3}$$

$$K_r = \frac{[M_r]}{[M_b]}. \tag{4}$$

Substituting Eqs. 3 and 4 into Eq. 2 produces:

$$\log(1/C) = \log\frac{1}{\frac{V_a}{V_r}\left(\frac{K_t+1}{K_rK_t}\right)} = -\log\frac{V_a}{V_r} + \log K_r + \log K_t - \log(K_t+1). \tag{5}$$

Equation 5 shows that the effect of tautomerism on apparent affinity is independent of the affinity $K_r$ of the compound for the target: It depends only on the equilibrium constant for tautomerism. Figure 2 shows the relationship between log $K_t$ and log(1/$C$). As one might expect, if the bound tautomer is that which is favored in solution, the effect on log(1/$C$) is negligible. Only when less than 50% of the compound in solution is in the form that binds, log $K_t \leq 0$, is a correction needed.

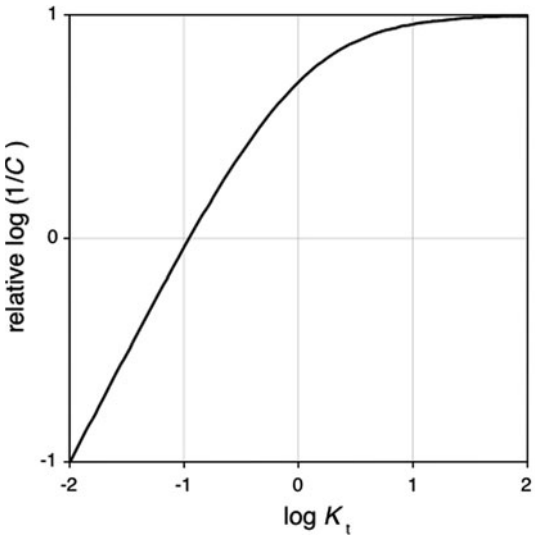


**Fig. 2** A plot of Eq. 5 that shows the influence of $K_t$ on the apparent potency of a molecule. The quantity $-\log\left(V_a/V_r\right) + \log K_r$ was set equal to 1.0

## Influence of Hammett $\sigma$

If one is exploring the structure–activity relationships in a set of molecules, one might not know the extent of tautomerism in each analogue. In spite of this, because tautomerism involves the competition of two sites on the molecule for a proton, it can be expressed as the competition between the p$K_a$'s of the two sites [4]. Thus it is not surprising that many investigations show that log $K_t$ is a function of some type of Hammett $\sigma$ (for example [4–9]).

$$\log K_t = a + b\sigma \tag{6}$$

However, frequently binding affinity is also a function of $\sigma$ [10, 11]:

$$\log K_r = c + d\sigma. \tag{7}$$

As a result, Eq. 5 can be expressed in terms of $\sigma$:

$$\log(1/C) = x + c + a + d\sigma + b\sigma - \log\left(10^a 10^{b\sigma} + 1\right). \tag{8}$$

Collecting constants and assuming for convenience that $a = 0$ yields:

$$\log(1/C) = X + d\sigma + b\sigma - \log\left(10^{b\sigma} + 1\right). \tag{9}$$

Although for every dataset the values of $a$, $b$, $c$, and $d$ will vary as will the type of $\sigma$ constants that are appropriate, example plots illustrate the complex relationships that result from Eq. 9. Figure 3 shows the relationship between the relative log(1/$C$) and $\sigma$ for cases in which the dependence of tautomeric equilibria on $\sigma$, the coefficient $b$, is varied from $-1$ to $+1$ while the dependence on binding is kept constant at $d = 1$. A dramatic difference is seen in the various curves. When $b = d = 1$, the dependence on $\sigma$ is greatest, spanning three orders of magnitude in log(1/$C$) when $\sigma$ is varied over its usual two log range. When the signs of the two relationships are opposite, ($b = -1$ and $d = +1$) the dependence on $\sigma$ is negative, reflecting the dominant effect of the tautomeric equilibrium. The heavy line in Fig. 3 shows that even if there is no $\sigma$ dependence of binding to the target, the relative affinity still shows a correlation with $\sigma$. In no case is the relationship strictly linear, but the slight curvatures might not be detected in any statistical fitting.

Figure 4 shows the relationship between the relative log(1/$C$) and $\sigma$ for cases in which the dependence of target binding equilibria on $\sigma$, the coefficient $d$, is varied from $-1$ to $+1$ while the dependence on the tautomeric equilibrium is kept constant at $b = 1$. A much less dramatic difference is seen in the various curves. In fact, when the signs of the two relationships are opposite, ($b = +1$ and $d = -1$), the dependence on $\sigma$ remains slightly positive, reflecting again the dominant effect of the tautomeric equilibrium. For the

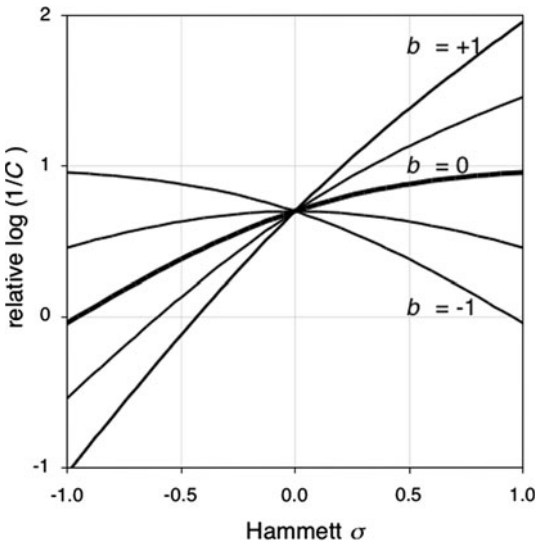


**Fig. 3** A plot of Eq. 9 that shows the influence of changing $\sigma$ on $K_t$. The lines for $b = 1.0, 0.5, 0.0, -0.5$, and $-1.0$ are shown. The values of $d$ and $X$ were set at 1.0

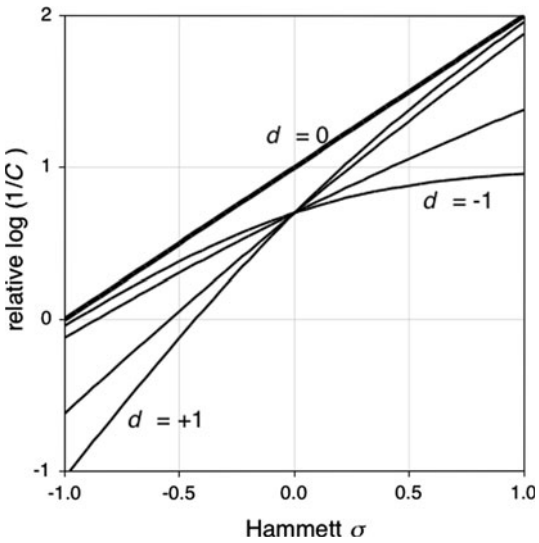


**Fig. 4** A plot of Eq. 9 that shows the influence of changing $\sigma$ on $K_r$. The lines for $d = 1.0, 0.5, 0.0, -0.5$, and $-1.0$ are shown. The values of $b$ and $X$ were set at 1.0

case for which the tautomeric equilibrium is independent of $\sigma$, there is the usual linear correlation of relative affinity with $\sigma$. Again, although the curves are not linear, this non-linearity would probably not be detected in statistical fitting.

## Discussion

The model shown was chosen to highlight the influence of tautomerism on QSAR. It usually would be necessary to also consider hydrophobicity or steric effects and additional compartments may also be required. None-the-less, the models clearly show that tautomerism affects the observed potency only if the minor tautomer is that which is bound. The complex relationships with Hammett $\sigma$ show that the apparent regression coefficient is a complex relationship between the regression coefficients for the individual equilibria.

To consider tautomerism in 3D QSAR one usually proposes, with some justification, the bioactive tautomer. The question then becomes how to account for the unknown magnitude of the tautomeric equilibrium that occurs in competition with binding. Because of their correlation with Hammett $\sigma$ constants [12, 13], the contribution of electrostatic fields in methods such as CoMFA might reflect tautomerism as well as target binding. The quality of the electrostatic field calculation would of course affect the quality of such an interpretation.

For other methods that provide a quantitative estimate of relative potency, a value for $K_t$ would be necessary. This can be estimated from the calculated p$K_a$'s of the various tautomers [14]. Alternatively, the value of $K_t$ can be estimated from quantum chemical (for example, [15–17]) or empirical calculations [18, 19]. The resulting $K_t$ would be substituted into the final two terms in Eq. 5 to correct the potency derived from the prediction method of interest.

## References

1. Martin YC, Hackbarth JJ (1976) J Med Chem 19:1033–1039
2. Martin YC (1978) Quantitative drug design. Dekker, New York Second edition in print, Taylor and Francis
3. Martin YC (1980) In: Yalkowsky SH, Sinkula AA, Valvani SC (eds) Physical chemical properties of drugs. Marcel Dekker, New York, pp 49–110
4. Katritzky A (1994) Lecture 6. Heterocyclic tautomerism. http://ufark12.chem.ufl.edu/Lectures/L6.pdf. Accessed February 22, 2010
5. Sorensen PE, McClelland RA, Gandour RD (1991) Acta Chem Scand 45:558–566
6. Elguero J, Katritzky AR, Denisko OV (2000) Adv Heterocycl Chem 76:1–84
7. Lazar L, Goblyos A, Martinek TA, Fulop F (2002) J Org Chem 67:4734–4741
8. Hetenyi A, Martinek TA, Lazar L, Zalan Z, Fulop F (2003) J Org Chem 68:5705–5712
9. Martiskainen O, Gawinecki R, Osmialowski B, Wiinamaki K, Pihlaja K (2009) Rapid Commun Mass Spectrom 23:1075–1084
10. Hansch C, Fujita T (1964) J Am Chem Soc 86:1616–1626
11. Hansch C, Leo A (1995) Exploring QSAR: fundamentals and applications in chemistry and biology. American Chemical Society, Washington DC
12. Kim KH, Martin YC (1991) J Med Chem 34:2056–2060
13. Kim KH, Martin YC (1991) J Org Chem 56:2723–2729
14. Szegezdi J, Csizmadia F (2007) Tautomer generation. p$K_a$ based dominance conditions for generating dominant tautomers.

American Chemical Society Fall meeting, Aug 19-23, 2007. http://www.chemaxon.com/conf/Tautomer_generation_A4.pdf. Accessed February 22, 2010
15. Dos Santos HF, De Almeida WB, Zerner MC (1998) J Pharm Sci 87:190–195
16. Jang YH, Goddard WA III, Noyes KT, Sowers LC, Hwang S, Chung DS (2002) Chem ResToxicol 15:1023–1035
17. Zeleny T, Hobza P, Kabelac M (2009) Phys Chem Chem Phys 11:3430–3435
18. Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M (2007) J Comp-Aid Mol Des 21:681–691
19. Milletti F, Storchi L, Sforna G, Cross S, Cruciani G (2009) J Chem Inf Model 49:68–75