Partition between phases of a solute that exists as two interconverting species

Michael H. Abraham[a] **and Albert J. Leo**[b]
[a] *The Department of Chemistry, University College London, 20 Gordon Street, London, UK WC1H 0AJ*
[b] *Medicinal Chemistry Project, Pomona College, Claremont, California 91711, USA*

The partition between water and other solvents of a solute that exists as two rapidly interconverting species is discussed with specific reference to acetylacetone. Partition coefficients for the keto form and the enol form have been obtained from the global observed partition coefficient and the equilibrium constant for the keto–enol equilibrium in the various solvents. It is shown that global descriptors can be assigned to acetylacetone that reproduce the observed partition coefficients to 0.125 log units, and which can be used to predict further such coefficients in other solvents. In a similar way, descriptors can be assigned to the separate keto and enol forms that can be used to predict the individual partition coefficients. It is further shown that the keto form is more dipolar/polarizable than the enol form, and is a stronger hydrogen-bond base as well. Neither the keto nor the enol form have any significant hydrogen-bond acidity, which indicates that the enol must form a very strong intramolecular hydrogen-bond that remains intact in all the solvents studied; this conclusion is in accord with previous findings of Emsley and Freeman.

Numerous compounds exist in solution as two or more rapidly interconverting forms, the most common being either conformers or tautomers. Almost any equilibrium between conformers will be solvent dependent; numerous studies of solvent effects on conformational equilibria have been reported.[1–3] Similarly, equilibria between tautomers will be solvent dependent, as shown by Masuda *et al.*[4] and by Emsley and Freeman,[5] for example. Since conformational and tautomeric equilibria are solvent dependent, it follows that partition coefficients between solvents, of compounds that exist as conformers or tautomers will depend in a rather complicated way on the Gibbs energies of the various species involved. However, there has been very little discussion of this, possibly because of lack of data. Leo,[6] however, has drawn attention to the problem of calculating partition coefficients for tautomeric compounds. He showed, for example, that the CLOGP program for calculating partition coefficients between water and octanol gave a log P value of -0.46 for the keto tautomer of acetylacetone, and an approximate value of $+0.66$ for the enol form, as compared to the observed value of $+0.40$ log units. However, this does not mean that the enol form predominates in both solvents. Indeed, Emsley and Freeman[5] showed that for acetylacetone the keto form predominates in water, but the enol form predominates in alcohols. For an understanding of solvent effects on partition coefficients of conformers or tautomers, a more detailed analysis is required.

We consider a solute that can exist in solution as two rapidly interconverting species. For example, these may be two conformers or may be two tautomers. We shall consider in detail the case of acetylacetone (pentane-2,4-dione) that exists as an equilibrium mixture of the ketone (C) and the enol (E) form, but the equations we formulate will be applicable to any solute that exists as two rapidly interconverting species. When such a solute is partitioned between water (aq) and some other solvent (s), the observed global partition coefficient, P_{obs}, is obtained from the formal concentration of the solute in the two solvents, eqn. (1).

$$P_{\mathrm{obs}} = (C_S + E_S)/(C_{\mathrm{aq}} + E_{\mathrm{aq}}) \qquad (1)$$

$$\begin{tikzcd} \text{Water} & \text{Solvent} & \\ C_{\mathrm{aq}} \arrow[r, "P_C"] \arrow[d, "K_{\mathrm{aq}}"] & C_S \arrow[d, "K_S"] & \text{Keto} \\ E_{\mathrm{aq}} \arrow[r, "P_E"] & E_S & \text{Enol} \end{tikzcd}$$

Fig. 1 Equilibria for the keto (C) and enol (E) forms

Let the partition coefficient of the keto form between water and the solvent, eqn. (2), be P_C, and the partition coefficient of the enol form between water and the solvent, eqn. (3) be P_E.

$$\text{Keto (aq)} \overset{P_C}{=} \text{Keto (s)} \qquad (2)$$

$$\text{Enol (aq)} \overset{P_E}{=} \text{Enol (s)} \qquad (3)$$

The equilibrium, eqn. (4), will be set up in both water and the solvent, with equilibrium constants we denote as K_{aq} and K_S respectively, defined by eqn. (5); see Fig. 1.

$$\text{Keto} = \text{Enol} \qquad (4)$$

$$K_{\mathrm{aq}} = E_{\mathrm{aq}}/C_{\mathrm{aq}} \quad \text{and} \quad K_S = E_S/C_S \qquad (5)$$

Then if eqn. (5) is substituted into eqn. (1) we have eqn. (6).

$$P_{\mathrm{obs}} = (C_S + K_S C_S)/(C_{\mathrm{aq}} + K_{\mathrm{aq}} C_{\mathrm{aq}}) \qquad (6)$$

Rearranging eqn. (6) and inserting $P_C = C_S/C_{\mathrm{aq}}$ from eqn. (2) leads to eqn. (7).

$$P_C = P_{\mathrm{obs}}(1 + K_{\mathrm{aq}})/(1 + K_S) \qquad (7)$$

A similar derivation leads to the equivalent eqn. (8).

$$P_E = P_{\mathrm{obs}} K_S (1 + K_{\mathrm{aq}})/K_{\mathrm{aq}}(1 + K_S) \qquad (8)$$

Table 1 Coefficients in eqn. (10) for partition between water and solvents

Solvent	*c*	*r*	*s*	*a*	*b*	*v*
Pentan-1-ol	0.175	0.575	−0.787	0.020	−2.837	3.249
Hexan-1-ol	0.143	0.718	−0.980	0.145	−3.214	3.403
Octan-1-ol	0.088	0.562	−1.054	0.034	−3.460	3.814
Decan-1-ol	0.008	0.485	−0.974	0.015	−3.798	3.945
Dibutyl ether	0.177	0.980	−1.657	0.695	−5.150	4.699
Benzene	0.017	0.490	−0.604	−3.013	−4.628	4.587
Toluene	0.015	0.594	−0.781	−2.918	−4.571	4.533
Chlorobenzene	0.046	0.259	−0.466	−3.047	−4.819	4.660
Nitrobenzene	−0.181	0.576	0.003	−2.356	−4.420	4.263
CH_2Cl_2	0.326	0.097	−0.037	−3.312	−4.128	4.252
$CHCl_3$	0.125	0.118	−0.372	−3.390	−3.467	4.521
CCl_4	0.223	0.564	−1.151	−3.510	−4.536	4.501
CH_2ClCH_2Cl	0.161	0.124	−0.001	−3.047	−4.290	4.300
Cyclohexane	0.127	0.816	−1.731	−3.778	−4.905	4.646
Alkane	0.287	0.649	−1.657	−3.516	−4.818	4.282
n-Butyl acetate	−0.468	0.712	−0.397	0.010	−3.743	3.865

If both K_{aq} and K_S are very small, so that in both phases the keto form predominates, then $P_C = P_{obs}$ from eqn. (7), and if both K_{aq} and K_S are very large (but not necessarily equal) then $P_E = P_{obs}$ from eqn. (8). Of course for the trivial case that both K_{aq} and K_S are unity, $P_C = P_E = P_{obs}$. In the case of acetylacetone, for which K_{aq} is much less than unity, and K_S is usually much larger than unity; then $P_C < P_{obs}$, and $P_E > P_{obs}$.

Emsley and Freeman[5] measured equilibrium constants for the keto–enol equilibrium of acetylacetone at 298 K in 21 solvents. In water, $K_{aq} = 0.148$, but in most of the other solvents K_S was much larger than unity, being 17.5 in tetrachloromethane and 32.3 in cyclohexane. Emsley and Freeman showed that polar solvents favoured the keto form, and that there were good correlations of ΔG° for the equilibrium against various measures of solvent polarity. Abraham *et al.*[7] extended this work somewhat through a multiple regression equation, eqn. (9), using the Kamlet–Taft solvatochromic parameters, $\pi_1{}^*$ a

$$\Delta G^\circ/\text{kcal mol}^{-1} = -2.24 + 2.39\,\pi_1{}^* + 0.32\,\alpha_1 \quad (9)$$

measure of solvent polarity, and α_1 a measure of solvent hydrogen-bond acidity. This equation suggests that not only is the keto form more polar than the enol, but it is slightly more basic as well (because acidic solvents will interact with basic solutes).

In the case of acetylacetone we have a considerable amount of information on values of log P_{obs}, K_{aq} and K_S. However, we still do not know if it is possible to estimate the global partition of acetylacetone, as log P_{obs}, in other solvents, or if it is possible to estimate the partition of the individual species. Our method of analysis uses the general solvation equation of Abraham.[8–10]

In eqn. (10) the dependent variable is log SP, where SP is some

$$\log \text{SP} = c + rR_2 + s\pi_2{}^H + a\Sigma\alpha_2{}^H + b\Sigma\beta_2{}^H + vV_x \quad (10)$$

property of a series of solutes in a given solvent system. The independent variables are solute properties or descriptors as follows:[8,9] R_2 is an excess molar refraction that can be determined simply from a knowledge of the compound refractive index or can be easily estimated from fragments or substructures;[11] $\pi_2{}^H$ is the solute dipolarity/polarizability, it being not possible to devise descriptors for these separately; $\Sigma\alpha_2{}^H$ is the solute overall hydrogen-bond acidity; $\Sigma\beta_2{}^H$ is the solute overall hydrogen-bond basicity; V_x is the McGowan characteristic volume,[12] calculated from molecular structure. For certain particular classes of solute such as sulfoxides, anilines and alkylpyridines, the relative basicity appears to alter with the solvent system, and the basicity parameter $\Sigma\beta_2{}^H$ is replaced by the alternative descriptor $\Sigma\beta_2{}^O$ for partitions between water and organic solvents in which water is quite soluble.[10] These solvents include butyl acetate, diethyl ether, isobutanol and octanol. It should be noted that in eqn. (9) and eqn. (10) the subscript '1' is used to indicate a solvent property or descriptor, and subscript '2' denotes a solute property.

Eqn. (10) has been applied to numerous partitions between water and other solvents, and a collection[10,13–15] of the coefficients in eqn. (10) for these partitions is in Table 1. Now if log P_{obs} values are known for acetylacetone between water and the solvents in Table 1, it should be possible to determine the solute descriptors for acetylacetone that best reproduce the observed log *P* values. The calculation is considerably aided in that the descriptors R_2 and V_x are known (0.414 and 0.8445 respectively, with the latter in units of $\text{cm}^3\ \text{mol}^{-1}/100$), so that only the three descriptors $\pi_2{}^H$, $\Sigma\alpha_2{}^H$ and $\Sigma\beta_2{}^H$ need to be assigned. Leo[16] has collected numerous partition coefficients between water and other solvents for acetylacetone; we give in Table 2 those for solvents for which we have applied eqn. (10). Of the given partition coefficients, that between water and n-butyl acetate seems out of line. From the remaining 15 observed partition coefficients we assigned descriptors for acetylacetone as $\pi_2{}^H = 0.81$, $\Sigma\alpha_2{}^H = 0.00$ and $\Sigma\beta_2{}^H = 0.63$ units. The calculated and observed partition coefficients are in Table 3, where the standard deviation, sd, between calculated and observed values is 0.125 log units. In several cases, replicate partition coefficients have been determined by different workers.[16] For chloroform, log $P = 1.40 \pm 0.06$ ($n = 7$), for tetrachloromethane log $P = 0.51 \pm 0.04$ ($n = 8$), for benzene log $P = 0.72 \pm 0.11$ ($n = 9$), for toluene log $P = 0.62 \pm 0.02$ ($n = 5$), for nitrobenzene log $P = 0.85 \pm 0.05$ ($n = 4$), and for the alkanes log $P = -0.02 \pm 0.13$ ($n = 11$). Hence our sd value of 0.125 log units is not more than twice the average deviations in the observed values.

The above analysis shows that it is possible to assign 'global' or 'average' descriptors to acetylacetone that will reproduce log P_{obs} values, and can be used to estimate values in other systems. We can, however, use the K_{aq} and K_S values of Emsley and Freeman in order to calculate log P_C and log P_E values for the individual keto and enol species. Unfortunately, out of the 16 solvents in Table 2, only (water), trichloromethane, tetrachloromethane, and cyclohexane were studied by Emsley and Freeman.[5] We therefore used eqn. (9) to calculate ΔG° and thence log K_S for eqn. (2), taking $\pi_1{}^*$ and α_1 values for the aprotic solvents from Abraham *et al.*[7] The alcoholic solvents cannot be dealt with so easily, because the partitions refer to the

Table 2 Partition coefficients and K_S values for acetylacetone

Solvent	K_S	log P_{obs}	log P_E	log P_C
Water	0.148			
Pentan-1-ol	2.21	0.43	1.16	−0.02
Hexan-1-ol	2.62	0.35	1.10	−0.15
Octan-1-ol	3.49	0.40	1.18	−0.19
Decan-1-ol	3.80	0.26	1.05	−0.36
Dibutyl ether	16.75	0.04	0.90	−1.15
Benzene	4.06	0.72	1.51	0.08
Toluene	4.97	0.62	1.43	−0.10
Chlorobenzene	2.49	0.77	1.51	0.29
Nitrobenzene	0.75	0.85	1.37	0.67
CH_2Cl_2	1.36	1.33	1.98	1.02
$CHCl_3$	4.75	1.40	2.21	0.70
CCl_4	17.52	0.51	1.38	−0.70
CH_2ClCH_2Cl	1.66	1.08	1.76	0.72
Cyclohexane	32.33	−0.02	0.86	−1.48
Alkane	43.80	−0.02	0.86	−1.61
n-Butyl acetate	5.88	0.79	1.61	0.01

Table 3 Calculations of partitions between water and various solvents for acetylacetone (global)

Solvent	log P_{obs}	log P_{calc}
Pentan-1-ol	0.43	0.731
Hexan-1-ol	0.35	0.494
Octan-1-ol	0.40	0.514
Decan-1-ol	0.26	0.358
Dibutyl ether	0.04	−0.038
Benzene	0.72	0.688
Toluene	0.62	0.576
Chlorobenzene	0.77	0.675
Nitrobenzene	0.85	0.874
CH_2Cl_2	1.33	1.326
$CHCl_3$	1.40	1.506
CCl_4	0.51	0.466
CH_2ClCH_2Cl	1.08	1.140
Cyclohexane	−0.02	−0.125
Alkane	−0.02	−0.220
n-Butyl acetate	0.79	0.410

Table 4 Calculations of partitions between water and various solvents for the enol form of acetylacetone

Solvent	log P_{Eobs}	log P_{Ecalc}
Pentan-1-ol	1.16	1.275
Hexan-1-ol	1.10	1.123
Octan-1-ol	1.18	1.188
Decan-1-ol	1.05	1.073
Dibutyl ether	0.90	0.980
Benzene	1.51	1.473
Toluene	1.43	1.378
Chlorobenzene	1.51	1.467
Nitrobenzene	1.37	1.537
CH_2Cl_2	1.98	1.951
$CHCl_3$	2.21	2.082
CCl_4	1.38	1.320
CH_2ClCH_2Cl	1.76	1.760
Cyclohexane	0.86	0.883
Alkane	0.86	0.760
n-Butyl acetate	1.61	1.031

water saturated alcohols and not to the dry alcohols, and so we used the π_1* and α_1 values given by Marcus[17] for the wet alcohols. In Table 2 are the log K_S values we thus calculated, together with the corresponding log P_E and log P_C values obtained through eqn. (7) and eqn. (8). As suggested above, log P_E is always larger than log P_{obs}, and log P_C is always less.

Table 5 Calculations of partitions between water and various solvents for the keto form of acetylacetone

Solvent	log P_{Cobs}	log P_{Ccalc}
Pentan-1-ol	−0.02	0.188
Hexan-1-ol	−0.15	−0.168
Octan-1-ol	−0.19	−0.227
Decan-1-ol	−0.36	−0.321
Dibutyl ether	−1.15	−1.119
Benzene	0.08	0.010
Toluene	−0.10	−0.204
Chlorobenzene	0.29	0.076
Nitrobenzene	0.67	0.623
CH_2Cl_2	1.02	0.989
$CHCl_3$	0.70	0.969
CCl_4	−0.70	−0.584
CH_2ClCH_2Cl	0.72	0.840
Cyclohexane	−1.48	−1.601
Alkane	−1.61	−1.609
n-Butyl acetate	1.01	0.089

Table 6 Descriptors for acetylacetone

Species	R_2	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2^H$	V_x
Global	0.412	0.81	0.00	0.63	0.8445
Enol form	0.412	0.66	0.00	0.48	0.8445
Keto form	0.412	1.43	0.07	0.65	0.8445
Keto form	0.412	1.36	0.00	0.70	0.8445

If we take R_2 and V_x as before, we can assign values to the other three descriptors that best reproduce the log P_E values. We find that with $\pi_2^H = 0.66$, $\Sigma\alpha_2^H = 0.00$ and $\Sigma\beta_2^H = 0.48$ units, the log P_E values can be reproduced to around 0.08 log units, see Table 4. But now, the descriptors refer explicitly to the enol form of acetylacetone. We can do exactly the same for the keto form, now using the log P_C values. Details are in Table 5; the three descriptors calculated are now $\pi_2^H = 1.43$, $\Sigma\alpha_2^H = 0.07$ and $\Sigma\beta_2^H = 0.65$ units respectively, and the standard deviation in log P_C observed and calculated is 0.127 log units.

A summary of the descriptors that we have obtained is in Table 6. As suggested by Emsley and Freeman[5] the keto form is much more polar than is the enol form, but also, as is implied by eqn. (9), the keto form is a stronger hydrogen-bond base than is the enol. Perhaps surprisingly, the keto form is a weak hydrogen-bond acid whereas the enol has no hydrogen-bond acidity at all. However, it is possible to correlate the log P_C values to within 0.157 log units (rather than to within 0.127 log units) if the keto form is assigned zero acidity, rather than $\Sigma\alpha_2^H = 0.07$ units. We do not regard this difference in goodness-of-fit as very significant, and leave open the question as to whether or not the keto form is a weak acid; in Table 6 are given the two alternative sets of descriptors for the keto form. The fact that the enol form has zero hydrogen-bond acidity confirms the suggestion of Emsley and Freeman[5] that the cyclic internal hydrogen-bond remains intact in all the solvents that they studied; these include water, methanol, DMF, DMSO and formic acid. This implies a very strong intramolecular hydrogen-bond in the enol form, so that the hydroxy hydrogen atom is not free to take part in any intermolecular hydrogen-bonding.

The CLOGP calculation of −0.46 for the partition of the keto form between water and octanol may be compared with our deduced value of −0.19 through eqn. (7) and our calculated value of −0.23 in Table 5. The corresponding approximate CLOGP calculation for the enol form of 0.66 is some way short of the values in Table 4, 1.18 and 1.19, but the CLOGP calculated value would be higher if an especially strong hydrogen bond was formed.[6]

Conclusions

In conclusion, the system we have studied is a reasonably general one; the ratio of the two forms varies markedly from solvent to solvent, and neither form completely predominates over the other within the solvent range studied. Yet even in this quite complicated system we have shown that the global log P_{obs} values can be treated as though acetylacetone were a single species with the global descriptors given in Table 6. These descriptors will enable log P_{obs} values to be predicted for any system for which the coefficients in eqn. (10) are known. Of course, for a detailed analysis of this system, or any related system, values of K_{aq} and K_S must be known in order to calculate the partition coefficients of the two individual species. However, if this can be done for a few selected partition systems, it is possible to determine descriptors for the two individual species, and then to be able to predict the individual partition coefficients in numerous other systems. More usually, the required information will not be known, but considerable headway can still be made, as we have shown. The problem of solutes that exist as two or more rapidly interconverting species in solution is thus not as intractable as might have been thought.

References

1 E. L. Eliel, N. L. Allinger, S. J. Angyal and G. A. Morrison, *Conformational Analysis*, Wiley, New York, 1965.
2 R. J. Abraham and E. Bretschneider, in *Internal Rotation in Molecules*, ed. W. J. Orville-Thomas, Academic Press, New York, 1974.
3 M. H. Abraham, R. J. Abraham, P. Leonard, N. S. True and C. Suarez, *J. Chem. Soc.*, *Perkin Trans. 2*, 1991, 463.
4 S. Masuda, M. Tanaka and T. Ota, *Makromol. Chem.*, 1986, **181**, 1087.
5 J. Emsley and N. J. Freeman, *J. Mol. Struct.*, 1987, **161**, 193.
6 A. J. Leo, *Chem. Rev.*, 1993, **93**, 1281, in press.
7 M. H. Abraham, P. L. Grellier, J.-L. M. Abboud, R. M. Doherty and R. W. Taft, *Can. J. Chem.*, 1988, **66**, 2673.
8 M. H. Abraham, *Chem. Soc. Rev.*, 1993, **22**, 73.
9 M. H. Abraham, *Pure Appl. Chem.*, 1993, **65**, 2503.
10 M. H. Abraham, *J. Phys. Org. Chem.*, 1993, **6**, 660.
11 M. H. Abraham, H. S. Chadha and R. C. Mitchell, *J. Pharm. Sci.*, 1994, **83**, 1257.
12 M. H. Abraham and J. C. McGowan, *Chromatographia*, 1987, **23**, 243.
13 M. H. Abraham, H. S. Chadha, G. S. Whiting and R. C. Mitchell, *J. Pharm. Sci.*, 1994, **83**, 1085.
14 M. H. Abraham, H. S. Chadha, J. P. Dixon and A. J. Leo, *J. Phys. Org. Chem.*, 1994, **7**, 712.
15 M. H. Abraham, unpublished work.
16 A. Leo, *The Pomona Medicinal Chemistry Project.*
17 Y. Marcus, *Chem. Soc. Rev.*, 1993, **22**, 409.

Paper 5/02386F
Received 13*th April* 1995
Accepted 26*th May* 1995