

A Study of the Gas Liquid Partition Coefficients of Eleven Normal, Branched and Cyclic Alkanes in Sixty Nine Common Organic Liquids: The Effect of Solute Structure

Won Jo Cheong

Department of Chemistry and Institute for Basic Research and Center for Advanced Bioseparation Technology,
Inha University, Incheon 402-751, Korea
Received September 10, 2001

Literature data measured by the author have been processed to report on the effect of solute structure on gas liquid partition coefficients of eleven normal, branched and cyclic alkanes ranging in carbon number from five to nine in sixty nine low molecular weight liquids. The alkane solutes are n-pentane(p), n-hexane(hx), n-heptane(hp), n-octane(o), n-nonane(n), 2-methylpentane(mp), 2,5-dimethylpentane(dp), 2,5-dimethylhexane(dh), 2,3,4-trimethylpentane(tp), cyclohexane(ch), and ethylcyclohexane(ec). The solvent set encompasses most of those studied by Rohrschneider as well as three homologous series of solvents (n-alkanes, 1-alcohols and 1-nitriles) and several perfluorinated alkanes and highly fluorinated alcohols. An excellent linear relationship was observed between $\ln K$ and the carbon number of n-alkanes. The effective carbon numbers of branched and cyclic alkanes were determined in a similar fashion to the method of Kovats index. We found that the logarithm of solute vapor pressure multiplied by solute molar volume was a perfect descriptor for the linear relationship with the median effective carbon number.

Keywords : Gas/liquid partition coefficients, Alkanes, Effective carbon number, Solute structure.

Introduction

The gas liquid partition coefficients of alkanes have been the subject of intense interest in many fields of chemistry, including chromatography,¹⁻⁵ chemical engineering^{6,7} and physical organic chemistry^{8,9} for many years. From a fundamental perspective, perhaps the greatest interest in solutions of alkanes in alkanes is related to their use as simple models of polymer mixtures.¹⁰⁻¹⁵ Due to their use as models of polymers and the fact that the chief methodology for the study of partition coefficients between gas and liquid at infinite dilution is the measurement of chromatographic retention, most of the systematic studies of the distribution of alkanes at infinite dilution have been limited to reasonably small solutes (butane through dodecane) in fairly large alkanes (hexadecane and above) usually at temperatures above 50 °C. Patterson and his collaborators carried out a very extensive set of studies of the mixing properties of a wide variety of alkanes in many different alkanes over a temperature range encompassing 25 °C.¹⁶⁻¹⁷ Their work has focused on excess enthalpies, heat capacities and volumes of mixing. Since most liquid chromatographic work is carried out at room temperature and solvents as large as hexadecane are of almost no interest as mobile phases, the above studies are not particularly pertinent to understanding mobile phase solute-solvent interactions in liquid chromatography. An automated head space gas chromatograph proved useful in measuring infinite dilution partition coefficients of volatile solutes in volatile solvents.¹⁸⁻¹⁹ This system had been used to measure partition coefficients of a variety of lower alkanes in a broad spectrum of solvents of general chemical interest at ambient rather than elevated temperatures.¹⁹ The gas

liquid partition coefficients used in this work are defined by equation 1.

$$K = [\text{solute}]_{\text{liquid}} / [\text{solute}]_{\text{gas}} \quad (1)$$

The free energy of solution of the solute corresponds to the process:

$$\text{solute (liquid, 1M)} = \text{solute (gas, 1M)} \quad (2)$$

and is given by:

$$\Delta G^\circ = -RT \ln K \quad (3)$$

Various models of the process of solution have been studied based on measurements of the gas liquid partition coefficient, solubility or the infinite dilution activity coefficient of alkanes. Conder and Young have summarized many of these studies and models.²⁰ Except for the geometrically simpler rare gases, the alkanes are the least complex probe solutes because the only possible attractive forces between an alkane and the solvent are dispersive and solvent dipole-solute induced dipole interactions. Dipole-dipole and specific chemical interactions (e.g. hydrogen bond formation) between the solute and the solvent are non-existent. Despite the above mentioned interest in these systems, most systematic studies of the free energy of solution of alkanes have been limited to rather non-volatile solvents.²⁰ There are a large number of studies of vapor liquid equilibria of a limited number of alkanes in relatively volatile solvents. However examination of extensive compilations²¹ indicates considerable imprecision and lack of agreement among data determined by different experimental methods, particularly in the critically important infinitely dilute solute concentration range.

The importance of detailed knowledge of the solution

properties of the alkanes in chromatography is exemplified in the well known Snyder solvent triangle and P' scale of solvent strength that have found widespread application in chromatography.^{22,23} Snyder's work is based in part on an intuitively appealing but certainly not particularly rigorous "normalization" of the gas liquid partition coefficients of the solutes studied by Rohrschneider²⁴ against the gas liquid partition coefficient of n-octane in the same solvent. Poppe has criticized the above normalization procedure,²⁵ and due to the paucity of data on alkanes in low molecular weight polar and protic solvents it has not been possible to validate Snyder's normalization method. Similarly, in an earlier study of linear solvation energy relationships, the partition coefficients of chemically interesting solutes were normalized by dividing them by the gas liquid partition coefficients of alkanes of roughly the same size.²⁶ These data were drawn from many diverse sources of often indeterminate reliability.

The gas liquid partition coefficients measured by HSGC (Head Space Gas Chromatography) and reported in the reference thesis¹⁹ were used in this study. The basic methodology is reviewed briefly here since the data was reported only in the thesis. In essence, a fully automated system was used to add a mixture of solutes to a known amount of solvent in a water jacketed thermostated cell (25 ± 0.01 °C). The vapor above the equilibrated liquid mixture was analyzed by gas chromatography. In general, three or four replicate samples of the vapor were measured and the liquid phase solute concentration was increased via a computer controlled dispenser. The vapor phase concentration was established via the use of a standard gas mixture of the solutes of interest contained in a temperature controlled gas "volumetric" flask. The gas liquid partition coefficient was computed from the average slope of a plot of the gas phase vs. liquid phase solute concentration based on a minimum of five equally spaced concentrations.

Four replicate runs on the gas phase standard were carried out immediately before and after each solvent was studied. In no case were data reported if more than a 3% change in the response factor was observed. All samples and standards were analyzed in terms of both peak area and peak height. The system was cleaned extensively after each run was completed by drawing air through the transfer lines to remove condensed or adsorbed solvent residue and impurities. We maintained an internal self-consistency of better than 2-3% on all of the data during the course of the six month data acquisition period. Results for the activity coefficient standard are summarized in Table 1 to give an indication of the system reliability. Note that we did not correct for gas phase non-ideality, and, thus, the activity coefficients were computed from the following equation:

$$\gamma = \frac{d_j RT}{K p^\circ M_j} \quad (4)$$

where d_j is the density of the solvent, K , the partition coefficient, p° , the vapor pressure of the pure solute, and M_j , the molecular weight of the solvent.

The list of 69 solvents is given below:

Table 1. Long term stability of the activity coefficient standard solution

run #	Activity coefficients of				
	pentane	hexane	heptane	nonane	trimethylpentane
1	0.86	0.94	0.98	1.10	1.01
2	0.86	0.94	0.98	1.10	1.00
3	0.85	0.94	0.97	1.10	0.99
4	0.86	0.94	0.97	1.10	1.00
5	0.86	0.94	0.98	1.10	1.00
6	0.85	0.94	0.97	1.08	0.99
7	0.86	0.94	0.98	1.09	1.00
8	0.85	0.93	0.97	1.08	0.99
9	0.86	0.94	0.98	1.09	1.00
10	0.86	0.94	0.98	1.09	1.00
11	0.86	0.94	0.98	1.09	1.01
12	0.85	0.95	0.98	1.09	1.00
13	0.86	0.95	0.97	1.10	1.01
average	0.857	0.941	0.978	1.093	1.000
r.s.d. (%)	0.55	0.53	0.61	0.69	0.71

^aTemperature: 25 °C, solvent: n-octane. ^bSolute concentration: n-pentane, 0.04515 (mole fraction). ^cSolute concentration: n-hexane, 0.03484 (mole fraction). ^dSolute concentration: n-heptane, 0.03285 (mole fraction). ^eSolute concentration: n-nonane, 0.02583 (mole fraction). ^fSolute concentration: 2,3,4-trimethyl pentane, 0.02856 (mole fraction).

PENTANE, ISOCTANE, SQUALANE, P-XYLENE, HEXADECANE, BENZENE, N-HEXANE, HEPTANE, NONANE, DECANE, ISOPROPANOL, HEXAMETHYLPHOSPHORIC TRIAMIDE, TOLUENE, CYCLOHEXANE, ETHYL ETHER, ACETONE, ETHYL ACETATE, 1,4-DIOXANE, BUTYL ETHER, N-BUTANOL, PROPANOL, ISOPROPYL ETHER, CHLOROBENZENE, METHYLENE CHLORIDE, BROMOBENZENE, FLUOROBENZENE, CARBON TETRACHLORIDE, ETHYLENE CHLORIDE, CYCLOHEXANONE, OCTANOL, BUTYRONITRILE, NONANENITRILE, ACETOPHENONE, TRIETHYLAMINE, TETRAHYDROFURAN, CHLOROFORM, METHYL ETHYL KETONE, BENZONITRILE, T-BUTANOL, NITROBENZENE, CARBON DISULFIDE, PYRIDINE, ETHYL BROMIDE, ETHOXY BENZENE, BIS(2-ETHOXY-ETHYL)ETHER, 2-PICOLINE, PERFLUOROHEXANE, PERFLUORO-OCTANE, DIMETHYLACETAMIDE, DIMETHYL FORAMIDE, DIMETHYL SULFOXIDE, ACETONITRILE, NITROETHANE, ISOPENTANOL, ANILINE, METHYL-FORAMIDE, N-METHYL-2-PYROLIDONE, BUTYROLACTONE, BENZYL ALCOHOL, ETHANOL, METHOXYETANOL, TRIFLUOROETHANOL, PERFLUOROTRIBUTYLAMINE, HEXAFLUOROISOPROPANOL, METHYLENE IODIDE, PENTADECANITRILE, METHANOL, ACETIC ACID

Note that a very similar set of partition coefficient data of the same solutes in similar solvents as those used in this study is reported in the literature.²⁷⁻²⁸ The data of this study was collected in advance and reported in the PhD thesis,¹⁹ but some system problem was suspected, and publication of the data was delayed. Later, the system was improved and the same measurements were made. The old data was simply abandoned. The new data set is probably more reliable than

the old data set. However, the values of the old data are in general systematically shifted from those of the new data set, and there does not exist sufficient, reliable literature data for cross-checking. Furthermore the old data set includes some data on solvents that were not used in the new data set. Publishing this work independently seems to be desirable. The application of data in this study is different from application of the new data set. The new data set was used for a comparative study of semitheoretical models to predict infinite dilution activity coefficients of alkanes in organic solvents,²⁹ whereas the old data set was used to examine the effect of solute structure on $\ln K$ and ECN in this study.

Results and Discussion

The activity coefficients of the solutes in *p*-xylene were measured periodically by the solute addition procedure throughout the series of measurements. This had to be done whenever a change was made to the system and particularly after a very non-volatile solvent was studied, and, consequently, the system was cleaned by flushing it with a volatile solvent (pentane, ether) to remove the last traces of contaminant. These results are shown in Table 2. Evidently, the long term reproducibility of the system is 3% or better for all solutes, except for *n*-nonane.

The accuracy of the data can be tested by comparison with other literature data in several ways. Table 3 compares values of infinite dilution activity coefficients obtained in this work with a number of values obtained from measurements of retention in gas chromatography and by other methods.

Infinite dilution activity coefficient data obtained by different methods and/or from different sources are often in-

consistent.^{21,30} As an extreme example, the activity coefficients of nonpolar solutes in polar solvents, such as dimethylformamide, *N*-methylpyrrolidone, and acetonitrile (see Table 3), vary by up to 200%. A difference of 10% between different sources is frequently observed. Although we believe the accuracy of our method is in general better than 5%,¹⁹ it is not confirmed by other literature data which are inconsistent among themselves. We should note, however, there are some systematic differences in activity coefficients of alkane solutes in hexadecane between our data and the data collected by the dynamic GC method.³¹ The dynamic GC method is known to be well established for thermodynamic work as an excellent method of studying physicochemical properties. A 10% systematic difference is observed between our data and the dynamic GC data (see Table 3). Even though we found good agreement between our estimated activity coefficients for *n*-alkane solutes in themselves and estimated values from the literature vapor pressure data (see the discussions below), we must allow that a possible error was introduced in our system when a solvent of high boiling point was used. Adsorption/condensation of such solvents on the transfer system surface and consequent adsorption of solutes is possible. On the other hand, we observed that the difference does not vary with solute volatility and that our data is in rather good agreement with the literature data for squalane, which prompts us to point out the possibility of errors in the dynamic GC approach. A constant systematic error may be caused by the uncertainty in estimating the weight of the stationary phase. Adsorption of solutes onto the liquid surface and the solid support is another possible source of error.³² Whether the major error is in our system or in the dynamic GC system, the difference is not more than 10%, and the error is not random but systematic. The slope of the

Table 2. Long term stability of the activity coefficients^a

run #	solute ^b										
	p	mp	hx	dp	ch	hp	dh	tp	o	ec	n
1	1.36	1.47	1.41	1.51	1.32	1.39	1.52	1.46	1.41	1.34	1.62
2	1.39	1.49	1.41	1.54	1.33	1.41	1.53	1.46	1.39	1.33	1.46
3	1.36	1.47	1.38	1.49	1.29	1.36	1.50	1.42	1.36	1.29	1.39
4	1.38	1.47	1.40	1.50	1.32	1.43	1.55	1.48	1.42	1.37	1.57
5	1.34	1.43	1.35	1.54	1.29	1.34	1.52	1.48	1.42	1.40	1.63
6	1.33	1.42	1.34	1.52	1.29	1.38	1.52	1.48	1.40	1.39	1.60
7	1.42	1.49	1.41	1.56	1.33	1.40	1.54	1.48	1.43	1.37	1.56
8	1.40	1.47	1.38	1.55	1.32	1.41	1.54	1.49	1.45	1.38	1.60
9	1.40	1.49	1.41	1.58	1.35	1.44	1.58	1.52	1.47	1.42	1.45
10	1.38	1.46	1.38	1.53	1.31	1.39	1.50	1.44	1.39	1.35	1.50
11	1.33	1.40	1.33	1.48	1.26	1.36	1.48	1.43	1.38	1.33	1.49
12	1.36	1.44	1.36	1.54	1.30	1.39	1.53	1.47	1.45	1.40	1.56
13	1.34	1.46	1.38	1.55	1.31	1.41	1.55	1.49	1.45	1.40	1.62
14	1.34	1.40	1.33	1.48	1.27	1.34	1.48	1.42	1.40	1.36	1.53
15	1.34	1.43	1.35	1.52	1.30	1.39	1.53	1.48	1.44	1.39	-
average ^c	1.36	1.45	1.37	1.53	1.31	1.39	1.52	1.47	1.42	1.37	1.55
s.d. ^d	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.08

^aInfinite dilution activity coefficients in *p*-xylene and 25 °C. ^bThe solutes are: *n*-pentane (p), 2-methylpentane (mp), *n*-hexane (hx), 2,5-dimethylpentane (dp), cyclohexane (ch), *n*-heptane (hp), 2,5-dimethylhexane (dh), 2,3,4-trimethylpentane (tp), *n*-octane (o), ethylcyclohexane (ec), and *n*-nonane (n). ^cAverage for all measurements. ^dStandard deviation for all measurements.

Table 3. Comparison of measured activity coefficients to literature activity coefficients

Solute	Solvent	this work	lit. (t°C)	method	reference
pentane	hexadecane	0.757	0.850(25)	GC _{,dyn} ^a	(31)
hexane	hexadecane	0.817	0.897(25)	GC _{,dyn}	(31)
heptane	hexadecane	0.854	0.928(25)	GC _{,dyn}	(31)
octane	hexadecane	0.886	0.952(25)	GC _{,dyn}	(31)
nonane	hexadecane	0.941	0.974(25)	GC _{,dyn}	(31)
cyclohexane	hexadecane	0.737	0.790(25)	GC _{,dyn}	(31)
octane	hexadecane	0.886	0.93(25)	GC _{,dyn}	(19)
pentane	squalane	0.589	0.641(30)	GC _{,dyn}	(20)
hexane	squalane	0.648	0.640-0.649(30)	static ^b	(20)
hexane	squalane	0.648	0.640-0.687(30)	GC _{,dyn}	(20)
heptane	squalane	0.699	0.669(30)	static	(20)
heptane	squalane	0.699	0.681(30)	GC _{,dyn}	(20)
octane	squalane	0.74	0.70(30)	static	(19)
octane	benzene	2.159	2.03(25)	VLE ^c	(19)
octane	ethanol	15.8	15.9(25)	DILU ^d	(19)
cyclohexane	triethylamine	1.092	1.09(27.8)	EBUL ^e	(30)
hexane	toluene	1.624	1.74(20)	GC ^f	(30)
cyclohexane	toluene	1.434	1.59(20)	GC	(30)
hexane	benzene	2.072	2.21(20)	GC	(30)
pentane	p-xylene	1.385	1.48(20)	GC	(30)
hexane	p-xylene	1.409	1.44(20)	GC	(30)
cyclohexane	p-xylene	1.329	1.35(20)	GC	(30)
hexane	CCl ₄	1.226	1.33(20)	GC	(30)
cyclohexane	CCl ₄	0.995	1.12(20)	GC	(30)
Pentane	MEK ^g	3.322	3.66(20)	GC	(30)
hexane	anisole	3.359	3.94(20)	GC	(30)
heptane	octanol	2.679	3.06(25)	GC	(30)
octane	octanol	3.088	3.36(25)	GC	(30)
pentane	cyclohexanone	3.351	3.35(25)	GC	(30)
hexane	cyclohexanone	3.800	3.69(25)	GC	(30)
cyclohexane	cyclohexanone	2.708	2.50(25)	GC	(30)
pentane	butanol	4.248	4.06(25)	GC	(30)
hexane	butanol	5.251	5.00(20)	GC	(30)
cyclohexane	acetophenone	3.955	4.34(25)	GC	(30)
heptane	acetophenone	6.355	6.82(25)	GC	(30)
pentane	benzyl alcohol	10.46	11.80(25)	GC	(30)
hexane	benzyl alcohol	13.10	13.80(25)	GC	(30)
cyclohexane	benzyl alcohol	7.858	7.94(25)	GC	(30)
heptane	benzyl alcohol	16.63	16.80(25)	GC	(30)
hexane	DMF ^h	16.0	18-37(25)	Various ⁱ	(30)
cyclohexane	DMF	11.3	12-29(25)	Various	(30)
heptane	DMF	21.6	20-46(25)	Various	(30)
pentane	NMP ^j	9.351	10-14(25)	Various	(30)
hexane	NMP	11.85	12-23(25)	Various	(30)
cyclohexane	NMP	7.518	8-14(25)	Various	(30)
heptane	NMP	15.32	15-28(25)	Various	(30)
octane	NMP	19.72	17-34(25)	Various	(30)
hexane	acetonitrile	24.22	25-31(25)	Various	(30)
cyclohexane	acetonitrile	19.07	21-25(25)	Various	(30)
heptane	acetonitrile	34.50	33-43(25)	Various	(30)
octane	acetonitrile	49.35	52-61(25)	Various	(30)
EC ^k	acetonitrile	36.78	37-44(25)	Various	(30)
cyclohexane	aniline	10.50	11.9(25)	GC	(30)
hexane	nitromethane	45.56	58.0(20)	GC	(30)
hexane	DMSO ^l	64.01	77.4(25)	DILU	(30)
EC	DMSO	76.38	78.5(30)	GC	(30)
octane	DMSO	157.7	156.0(30)	GC	(30)
nonane	DMSO	240.9	229.0(30)	GC	(30)
pentane	heptane	0.918	1.00(20)	GC	(30)
hexane	heptane	0.975	1.00(20)	GC	(30)
cyclohexane	heptane	1.002	0.99(20)	GC	(30)

^aDynamic GC method with non-volatile solvents as the stationary phase. ^bEquilibrium still method. ^cExtrapolated from vapor-liquid equilibrium data.^dGas stripping dilutor technique. ^eEbulliometry. ^fDynamic GC method with somewhat volatile solvents as the stationary phase. ^gMethyl ethyl ketone.^hDimethyl formamide. ⁱData from various methods. ^jN-methylpyrrolidone. ^kEthylcyclohexane. ^lDimethyl sulfoxide.

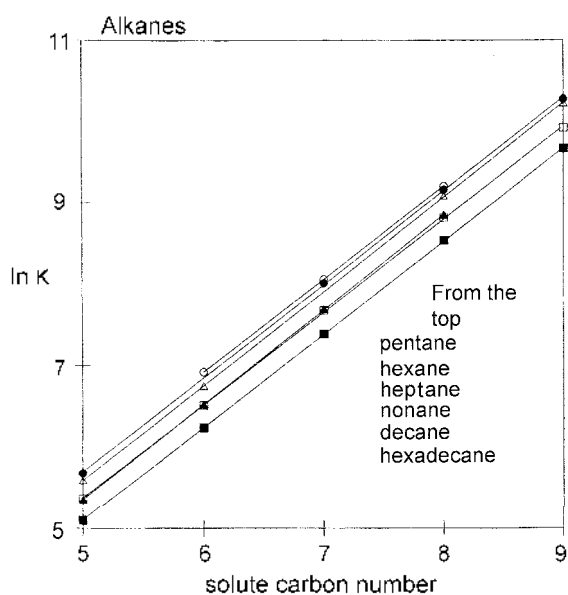


Figure 1. Plots of $\ln K$ vs. solute carbon number of n-alkane solutes in n-alkane solvents.

plot of $\ln K$ vs. solute carbon number from our data (1.15) is in a good agreement with the slope from the dynamic GC data (1.16). In spite of the disagreement in hexadecane, our system has proven to be reliable for solvents whose boiling points are lower than the transfer system temperature (170 °C).

We could not measure the K values of a specific solute in its own pure liquid. The data shown in Figure 1 are plots of $\ln K$ vs. solute carbon number for the n-alkane solutes in a series of n-alkane solvents. Based on the accuracy of the Kovats retention index scheme³³ and many models of the solution process,³⁴⁻⁴⁰ these plots are generally quite linear, thus, correlation coefficients are greater than 0.9995, and the average deviations are less than 0.05 in $\ln K$ units. Obviously, missing data points can be estimated by extrapolation and interpolation. These results are compared with literature data in Table 4. Since the Raoult's Law based activity coefficient is defined as unity for a solute in itself, equation (4) can be used to compute the K value based on the literature value of the vapor pressure (see Table 4). Apparently, there is excellent agreement between $\ln K$ obtained by extrapolation and interpolation of the data in Figure 1 and that computed

Table 4. Comparison of estimated $\ln K$ values to literature values^a

System	$\ln K_{lit}^b$	$\ln K_{est,solute}^c$	$\ln K_{est,solvent}^d$
n-pentane in n-pentane	5.744	5.722	5.677
n-hexane in n-hexane	6.839	6.849	6.800
n-heptane in n-heptane	7.923	7.908	7.909
n-octane in n-octane	9.010	–	9.000
n-nonane in n-nonane	10.10	10.026	10.07
n-decane in n-decane	11.12	11.203	–

^aAt 25 °C. ^bBased on K computed via equation (4) with gamma equal to 1.000 and the saturated vapor pressure given in Table 7. ^cFrom extrapolation of the data given in Figure 1 based on linearity between $\ln K$ and carbon number of the n-alkanes as solutes. ^dBased on data given in Figure 2.

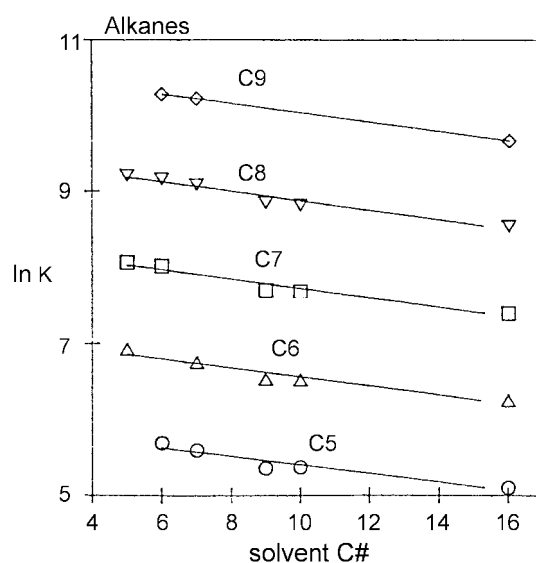


Figure 2. Plots of $\ln K$ vs. solvent carbon number of n-alkane solvents for n-alkane solutes.

from the literature. The results suggest that the $\ln K$ and $\ln \gamma$ are accurate to 0.03 units.

The existence of the linear relationship shown in Figure 1 is hardly surprising. The $\ln K$ values for each of the n-alkanes in a series of n-alkane solvents are plotted vs. the solvent carbon number in Figure 2. These plots are only roughly linear, and there is really no good reason to suppose that they should be precisely linear, but they are sufficiently linear and the variation in $\ln K$ with solvent is so small as to allow the estimation of missing data points. These data are given as the last column in Table 4. To get some idea of the self-consistency of the data between solvents a series of regressions between the $\ln K$ of each solute and that of an arbitrary but fixed reference solute was carried out. We chose this procedure because intuitively the best single correlator of the solvent induced change in the $\ln K$ of one alkane will be the variation in the $\ln K$ of a nearly identical alkane. We chose n-heptane as the fixed reference solute since it has intermediate volatility and should work reasonably well for all the alkanes. As can be seen in Table 5, the correlation is, in all cases better than 0.97. However, the quality of the fit improves when the test and reference alkane are most similar, as seen in the comparison of the regression of ethylcyclohexane against n-heptane (s.d. = 0.199) with the regression of ethylcyclohexane against cyclohexane (s.d. = 0.078). Similarly, the regressions of n-pentane and n-nonane vs. n-heptane are poorer than those of n-hexane and n-octane vs. n-heptane. Based on the average standard deviations of these fits in general we conclude that the total contribution of the random error to these fits is less than about 0.05 $\ln K$ units. Since there is no reason to believe that the reference solute is more precisely determined than the test solute, the average random error in each solute is likely to be about 0.03 units in $\ln K$ units.

To get some global idea of the physical characteristics of the solute that have the greatest effect on its partition co-

Table 5. Solute Intercorrelation^a

Test solute	Reference solute	Int. ^b	Slope ^c	r ^d	sd ^e	n ^f
p	hp	-0.479	0.756	0.9943	0.089	55
hx	hp	-0.226	0.877	0.9989	0.045	50
o	hp	0.275	1.112	0.9992	0.048	53
n	hp	0.493	1.223	0.9976	0.096	49
mp	hp	-0.318	0.843	0.9847	0.158	52
dp	hp	-0.213	0.918	0.9922	0.125	48
ch	hp	-0.019	1.043	0.9976	0.079	55
dh	hp	0.574	0.990	0.9989	0.052	54
tp	hp	1.041	0.782	0.9784	0.178	56
ec	hp	1.377	0.988	0.9846	0.199	48
mp	dp	-0.084	0.907	0.9996	0.027	56
ch	dp	0.342	1.120	0.9983	0.068	51
dh	dp	0.878	1.064	0.9967	0.092	53
tp	ec	0.231	1.237	0.9977	0.078	52

^aRegression of $\ln K$ for the indicated test solute vs. $\ln K$ for the indicated reference solute in all available solvents. The solutes are defined in Table 2. ^bLeast squares intercept of plot of $\ln K_{\text{test}}$ vs. $\ln K_{\text{ref}}$. ^cLeast squares slope of plot of $\ln K_{\text{test}}$ vs. $\ln K_{\text{ref}}$. ^dCorrelation coefficient of the above plot. ^eAverage residual for the above plot. ^fNumber of solvents included in the above plots.

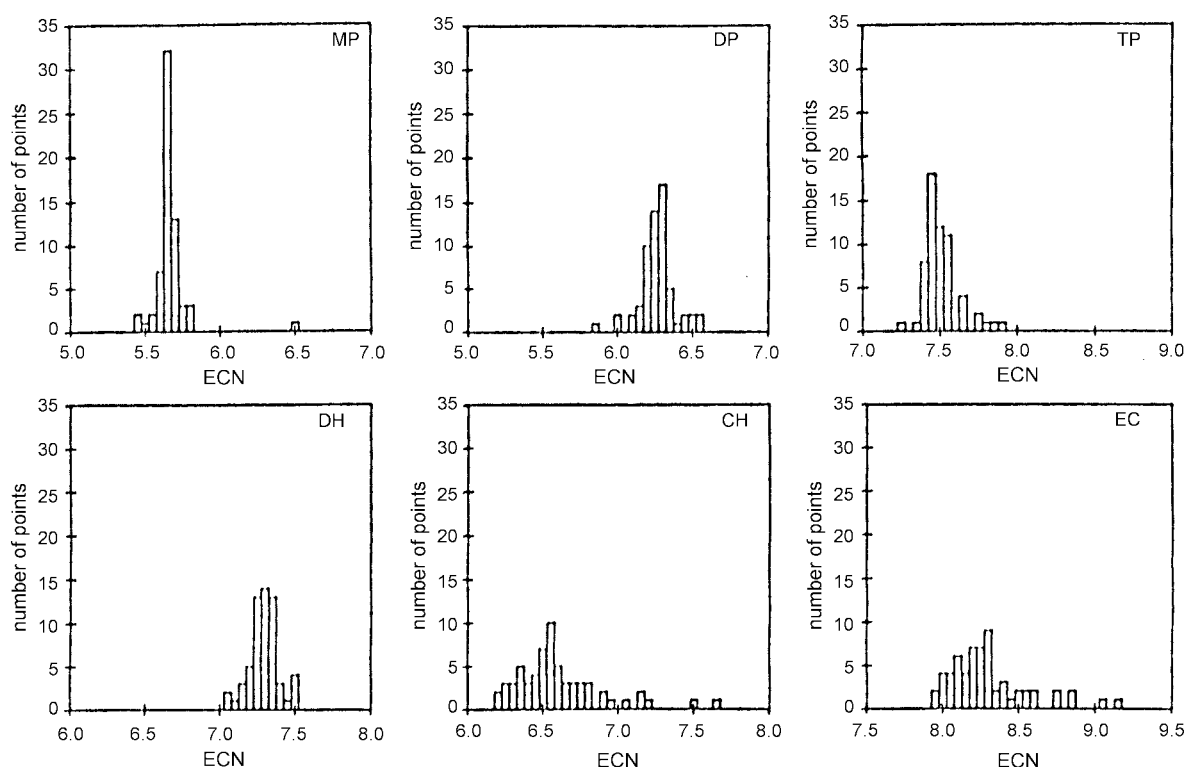
efficient we computed an "effective carbon number" (ECN). Conceptually, an ECN is similar to a Kovats index. This idea is receiving considerable attention in the chemical engineering literature for validating results and estimating compound properties.⁴¹ We compute the effective carbon number of a solute as follows: for each solvent the $\ln K$ for the n-alkanes are regressed against the number of carbons in the solute. In

Table 6. Effective carbon numbers for branched and cyclic alkanes

Basis for ECN	Solute					
	MP	DP	DH	TP	CH	EC
Average ^a	5.69	6.29	7.30	7.52	6.64	8.30
Median ^b	5.68	6.30	7.33	7.51	6.59	8.30
Alkanes ^c	5.72	6.33	7.33	7.44	6.33	8.03
Alcohols ^d	5.69	6.35	7.34	7.52	6.61	8.27
Dipolar solvents ^e	5.65	6.21	7.25	7.62	7.07	8.80
Perfluorinated ^f	5.84	6.54	7.53	7.81	6.31	8.13
Aromatic ^g	5.67	6.28	7.28	7.41	6.49	8.15
Chlorinated ^h	5.70	6.33	7.34	7.49	6.54	8.16
GC retention ⁱ	5.68	6.30	7.33	7.51	6.59	8.30
p ^{o,j}	5.73	6.37	7.35	7.45	6.37	8.00
H _{vap} ^k	5.67	6.28	7.28	7.24	6.28	7.84
sd ^l	0.12	0.11	0.12	0.12	0.29	0.26

^aAverage over all solvents. The average effective carbon numbers of the n-alkanes do not deviate from the defined value by more than 0.02 units which we believe is insignificant. ^bMedian value. ^cAverage over n-pentane through n-decane. ^dAverage of methanol, ethanol, n-propanol and iso-propanol. ^eAverage over the aprotic highly dipolar liquids dimethylsulfoxide, dimethylformamide, nitromethane and hexamethylphosphoramide. ^fAverage overperfluorohexane, perfluorooctaneand perfluorotributylamine. ^gAverage over benzene, toluene and p-xylene. ^hAverage over methylene chloride, chloroform and carbon tetrachloride. ⁱComputed from the measured retention times on OV-1701 at 150 °C. ^jBased on plot of $\ln p^o$ vs. carbon number for the n-alkanes. ^kBased on plot of ΔH_{vap} vs. carbon number for the n-alkanes. ^lStandard deviation of the entire data set.

general, the relationships are quite linear ($r > 0.998$). Nonetheless, we sometimes had reason to be suspicious of the data for n-pentane and n-nonane, the most volatile and non-volatile solutes among the solute set. Consequently, we also

**Figure 3.** Distribution histograms of the effective carbon numbers for branched and cyclic alkanes. See Table 3 for full solute names.

determined the least median line. As described by Rousseeuw and Massart,⁴²⁻⁴³ the least median line is a very robust estimator that can be very useful for small data sets.⁴⁴ In the context of the n-alkane data, two of the five points can be completely in error without biasing the least median line. The least median slope (m) and intercept (b) were used to compute the ECN for all eleven solutes in each solvent as per equation 5.

$$\text{ECN} = (\log K_{\text{solute}} - b)/m \quad (5)$$

We then establish an overall ECN for each solute by computing the mean and median ECN for the entire set of solvents. In addition, we computed averages for various classes of solvents. The results for the branched and cyclic alkanes are given in Table 6, and a set of histograms are presented in Figure 3. The ECN values of the n-alkanes (not given in Table 6) in nearly all solvents are very close to the expected integer, usually within 0.02 units.

The concept of an effective carbon number for the alkanes and whether or not the ECN is independent of solvent is a very significant idea. Suppose that the partition coefficients of all alkanes (normal, branched and cyclic) were governed by a single solute property (P_i) such that:

$$\ln K_{ij} = b_j + m_j P_i \quad (6)$$

where K_{ij} is the partition coefficient of solute i in solvent j , and b_j and m_j are the solute-independent, solvent-dependent

intercept and slope, respectively. The vast preponderance of our data for nearly seventy solvents indicates that a relationship such as equation 6 is valid for the n-alkanes if P_i is taken as the number of carbons. Any property of an n-alkane, such as its molar volume, hard core volume, hard core surface area, and molar refraction, that is collinear with carbon number will serve the same purpose. Thus, if ECN_j is solvent independent for all alkanes then some property (P_j) must exist.

The self-consistency of the ECNs within a subclass of solvents is typically 0.01-0.03 in ECN units (data not given). The data of Table 6 definitely indicate that the ECN depends upon the type of solvent used, but there is no clear relationship between the ECN and the type of solvent. For example, the ECN of a branched alkane in an alcohol is within 0.03 units of that in an alkane. In contrast, the ECNs of the branched alkanes in perfluorinated solvents are higher than average, whereas the ECNs of the two cyclic solutes are lower than average in this type of solvent. The solvent dependence of the ECN is supported by inspection of the histograms (see Figure 3).

The fact that $\ln K$ is linear vs. carbon number for n-alkanes indicates that these materials are very self-similar and the number of carbons is a good descriptor or encoder of the property that causes the differences between the n-alkanes. The fact that the ECNs for the branched and cyclic alkanes are at least somewhat constant indicates that their solvent dependent behavior is partially modeled by the n-alkanes

Table 7. Physical and Chemical Properties of the Alkanes.

Solute	p^a	BP ^b	H _v ^c	δ^d	n^e	
n-Pentane (P)	512.5	36.065	6.316	7.021	1.35472	
n-Hexane (HX)	151.4	68.736	7.541	7.266	1.37226	
n-Heptane (HP)	45.74	98.424	8.736	7.431	1.38511	
n-Octane (O)	13.98	125.673	9.916	7.551	1.39505	
n-Nonane (N)	4.27	150.818	11.100	7.647	1.40311	
2-Methyl pentane (MP)	212.0	60.271	7.138	7.071	1.36873	
2,4-dimethyl pentane (DP)	98.5	80.500	7.861	6.963	1.37882	
2,5-dimethyl hexane (DH)	30.4	109.103	9.051	7.144	1.39004	
2,3,4-trimethyl pentane (TP)	27.0	113.467	9.014	7.260	1.40198	
Cyclohexane (CH)	97.7	80.730	7.861	8.195	1.42354	
Ethyl cyclohexane (EC)	12.79	131.995	9.720	7.965	1.43073	
Solute	d_{25}^f	V_{25}^g	V_B^h	V_P^i	A_P^j	π^k
P	.6214	116.11	58.03	57.141	81.088	-.08
HX	.6548	131.608	68.26	67.165	93.617	-.04
HP	.6795	147.468	77.35	77.186	107.03	-.02
O	.6985	163.543	87.58	87.231	119.561	.01
N	.7138	179.684	97.81	97.231	132.975	.02
MP	.6485	132.884	68.25	66.991	92.593	-.07*
DP	.6683	149.937	78.47	76.756	103.652	-.08*
DH	.6893	165.723	87.56	86.880	117.358	-.05*
TP	.7150	159.767	87.55	84.444/	106.292/	-.08*
				85.813	109.507	
CH	.7738	108.752	60.24	60.799	80.656	.00
EC	.7839	143.152	80.69	80.658	105.204	.028

^aVapor pressure (mmHg) of the pure liquid. All properties are at 25 °C. ^bNormal boiling point in °C. ^cHeat of vaporization (kcal/mole). ^dHildebrand solubility parameter (cal/cc)^{1/2}. ^eRefractive index. ^fDensity (gram/cc). ^gMolar volume (cc/mole). ^hHard core volume (cc/mole) computed by the method of Bondi. ⁱHard core volume (cc/mole) provided by Pearlman. ^jHard core surface area (cm²/mole*10⁸) provided by Pearlman. ^kSolvatochromic π values; values marked with * denote estimates from related compounds.

and there is likely one major effect that controls differences in $\ln K$ with these solutes, but it is definitely not the only effect.

The ECNs were also computed from the measured isothermal retention time of all of the alkanes on a OV-1701 column (50% phenyl 50% methyl silicone, 0.2 mm i.d., 30 m). The column dead time was estimated by finding the value that best linearized a plot of the corrected retention time vs. carbon number of the *n*-alkanes.⁴⁵ As can be seen from the data in Table 6, the ECNs obtained from the *K* values and those from the chromatographic retention are in excellent agreement. The ECNs based on the logarithm of the vapor pressure are in better agreement with those based on the *K* values than those based on the enthalpy of vaporization, which is for all solutes lower than that based on the vapor pressure. We note that $\ln K$ and $\ln p^0$ are free energy quantities and the heat of vaporization obviously is not.

A statistically clear observation is that the ECNs of the cyclic solutes are much more sensitive to the solvent type than the ECNs of the branched solutes. This is supported by the histograms (see Figure 3) and the overall standard deviation of the dataset for cyclohexane (s.d. 0.29) and ethylcyclohexane (s.d. 0.26) in comparison with any of the branched alkanes (s.d. 0.11 to 0.12). Clearly, the cyclic alkanes are not nearly as well modeled by the *n*-alkanes as the branched alkanes. The differential behaviour of the cyclic alkanes is due to their shape rather than the nature of

the interaction. This observation concerning the cyclic alkanes is in accord with many concepts found in the chemical engineering literature, where cyclic alkanes are often handled in a decidedly different fashion than the normal and branched alkanes and are classified as "naphthenes".⁴⁶ It is also interesting to note that various corrections to the Flory-Huggins athermal entropy of mixing specifically incorporate a correction for the number of rings in a compound.⁴⁷ In addition, the use of specific cyclic group factors is essential to properly estimating vapor liquid equilibria of cyclic compounds by the UNIFAC method of estimating activity coefficients.^{48,49}

Because the agreement between the median and GC based ECN's is extremely good, we chose the median value as the global representative estimate of the key solute property. Use of the median also circumvents the effect of the extreme solvents and the largest measurement errors on the data. The median ECNs are presented in Table 6 and other relevant physical properties of the solutes are given in Table 7.

Various models of solution have implicated the molar volume, hard core volume and molecular surface area as the key factor controlling solubility and other phase transfer properties. The dependence of the ECN on these factors are shown in Figure 4. The plots show the variation in ECN against the molar volume, hard core volume computed by the method of Bondi,⁵⁰ hard core volume computed by the molecular mechanics and molecular surface area from the

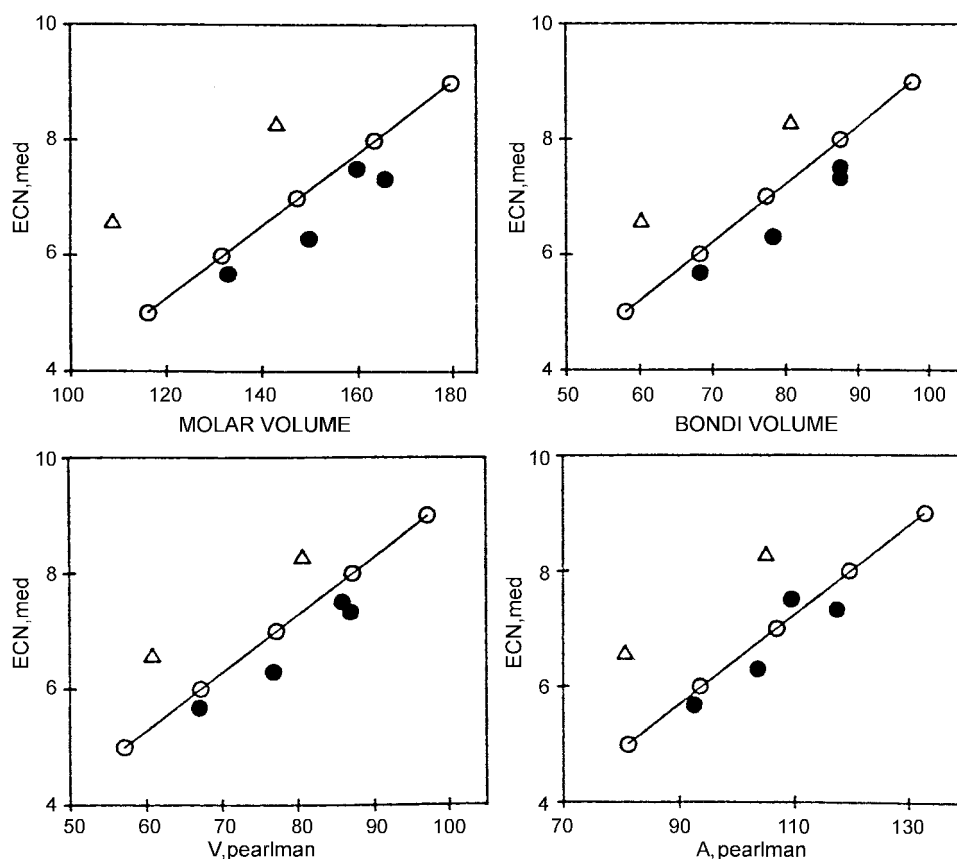


Figure 4. Plots of ECN vs. solute molar volume, Bondi volume, hard core volume, hard core surface area. Open circle (\circ); *n*-alkanes. Closed circles (\bullet), branched alkanes. Open triangle (\triangle); cyclic alkanes.

Table 8. Correlation of effective carbon number vs. various solute properties

Property	Int.	Slope	r	sd	n
Molar volume ^a	0.837	0.0422	0.762	0.82	11
Bondi volume ^a	0.580	0.0825	0.861	0.64	11
Pearlman volume ^a	1.381	0.0439	0.701	0.90	11
Pearlman area ^b	1.55	0.0314	0.732	0.90	11
$L(n^2)^c$	-9.94	71.2	0.732	0.86	11
Molar Refraction ^d	0.463	0.188	0.901	0.55	11
$-\ln p_2^0$ ^e	-10.32	-0.855	0.9958	0.12	11
$-\ln(p_2^0 V_2)^f$	15.31	-0.9389	0.9990	0.057	11

^aIn cm³/mole at 25 °C. ^bIn (cm³/mole) × 10⁸. ^c $L(n^2)$ defined as $(n^2 - 1)/(2n^2 - 1)$ with the refractive index (n) at 25 °C. ^dDefined as the product of the molar volume and $L(n^2)$ at 25 °C. ^eVapor pressure of the pure solute in atm at 25 °C. ^f V_2 in gram/cc at 25 °C.

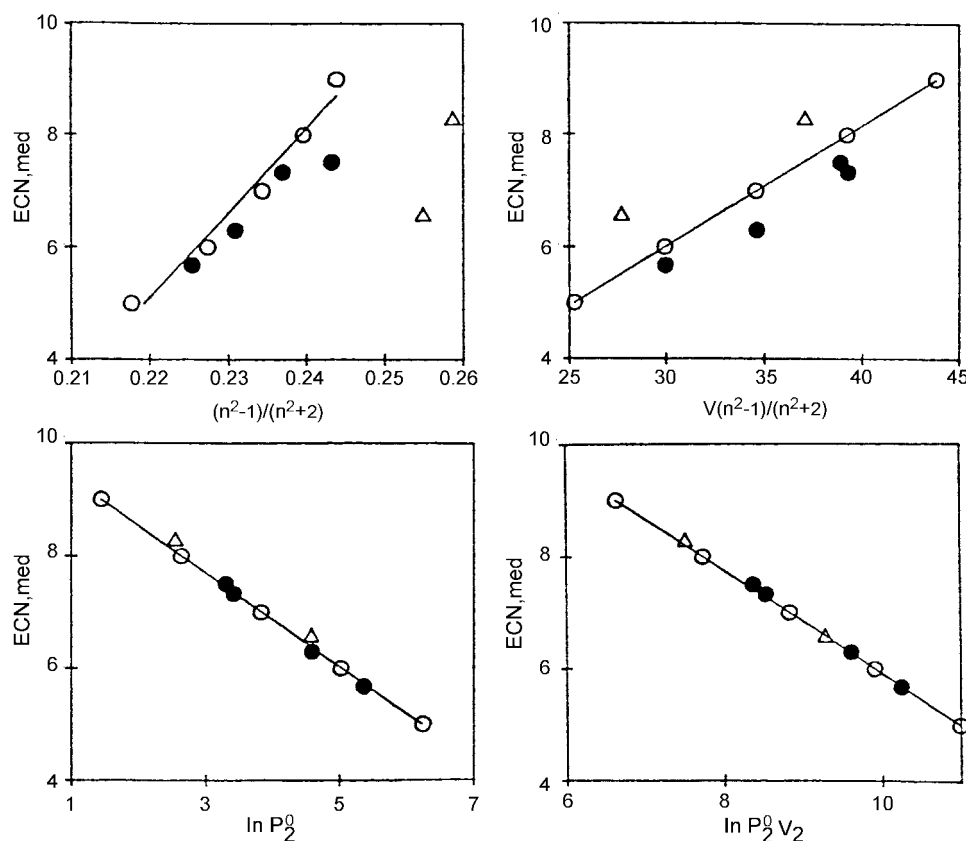
same approach. The statistics of the fits are summarized in Table 8.

Figure 4 makes it clear that none of the above measures of size come anywhere near correlating the ECN. Indeed, the average deviations shown in Table 8 are well outside the standard deviations for the ECN given in Table 6. The deviations are extremely systematic. The n-alkanes form a very straight line, the branched alkanes and cyclic alkanes fall below and above the line, respectively. If one plots the ECNs in specific solvents, for example hexadecane and

methanol, a virtually identical pattern is observed. Indeed, if the raw data, the $\ln K$ in a specific solvent, is plotted as in Figure 4, the identical pattern is observed. This rules out the supposition that the use of the overall ECN procedure inherently forces the n-alkanes to fall on a straight line. The fact that the plots are so similar in virtually all solvents indicates that the lack of fit is not due to the relative size of the solutes and the solvents, or the configurational entropy of mixing. The deviations are much larger than one would calculate based on the Flory-Huggins or the Staverman-Guggenheim modification of Flory-Huggins mixing entropy.^{51,52}

Molecular polarizability and molar refraction are also common measures of the interactions of alkanes with solvents. As shown in Figure 5, these correlations also show significant classwise lack of fit, although the molar refraction is beginning to show more than a crude correlation ($r = 0.90$).

We note that there is good correspondence between the overall median ECN for each solute and the ECN based on the vapor pressure of the pure solute (see Table 6). Based on the above discussion of the relationship between K and the saturated liquid vapor pressure of the pure solute it is reasonable to expect that the solute vapor pressure per se will be a good correlator of the ECNs of the alkanes. This should certainly be the case for the alkanes in non-polar solvents. This result is shown in Figure 5 and in Table 8. The vapor pressure is in fact an excellent correlator of the ECN ($r = 0.9968$), but classwise deviations still persist.

**Figure 5.** Plots of ECN vs. Onsage function, molar refraction, the logarithm of vapor pressure, and the logarithm of vapor pressure times molar volume. Symbols are the same as in Figure 4.

The K value can be thought of as a measure of the escaping tendency of a solute from a dilute solution to a gas phase. At infinite dilution, all solutes experience the same environment, both energetically and in terms of the total volume, since they are all in the same liquid. In contrast, the vapor pressure of the pure liquid solute can be imagined in terms of an isotopically labelled dilute solute escaping from a liquid of the "same" but unlabelled substance. When one uses the solute vapor pressure as a correlator of the escaping tendency, the environment is varied with variation of solute, including the energetics and the total volume of liquid, as reflected in the molar volume of the pure substance. The above argument is similar to that developed by Ben Naim^{53,56} in his discussion of the various standard states used to compute transfer free energies. Thus, by analogy to equation (4), which converts the Henry's Law constant to the partition coefficient, we regressed the ECN against $\ln(p_2^\circ V_2)$. The resulting correlation is excellent ($r = 0.9993$, $sd = 0.053$).

The accuracy of the product of the solute vapor pressure multiplied by its molar volume as a correlator of the solute partition coefficient can be assessed by regressing the $\ln K$ values against this factor. Details of the effect of solvent on K will be the subject of future work. For the moment we choose a particularly simple solvent namely hexadecane.

$$\ln K_{\text{hexadecane}} = 16.78 - 1.06(+/- 0.013)\ln(p_2^\circ V_2) \quad (7)$$

$$r = 0.9993, \text{ sd} = 0.053, n = 11$$

The slope of $\ln K$ vs. $\ln(p_2^\circ V_2)$ is very close to unity. The direction of the deviation is in the anticipated direction.

Acknowledgment. This work was supported by the fund of Center for Advanced Bioseparation Technology and 21 Century Foundation of Hanjin Group.

References

- Ashworth, A. J. *J. Chem. Soc. Faraday II* **1973**, *69*, 459.
- Ashworth, A. J.; Everett, D. H. *Trans. Faraday Soc.* **1960**, *56*, 1609.
- Cruikshank, A. J. B.; Windsor, M. L.; Young, C. L. *Proc. Roy. Soc. A* **1966**, *295*, 271.
- Conder, J. R.; Purnell, J. H. *Trans. Faraday Soc.* **1963**, *59*, 1655.
- Martire, D. E.; Pescok, R. L.; Purnell, J. H. *Trans. Faraday Soc.* **1965**, *61*, 2496.
- Prausnitz, J. M.; Lichtenthaler, R. M.; de Azevedo, E. G. *Molecular Thermodynamics of Fluid Phase Equilibria*, 2nd Ed.; Prentice Hall: Englewood Cliffs, 1986.
- Walas, S. M. *Phase Equilibria in Chemical Engineering*. Butterworth Publishers: Boston, 1985.
- Abraham, A. H.; Liszi, J. *J. Chem. Soc. Faraday Soc.* **1978**, *74*, 1604.
- Abraham, M. H. *J. Am. Chem. Soc.* **1979**, *101*, 5477.
- Huggins, M. L. *J. Chem. Phys.* **1941**, *9*, 440.
- Flory, P. J. *J. Chem. Phys.* **1941**, *9*, 660.
- Flory, P. J. *Disc. Faraday Soc.* **1970**, *49*, 7.
- Flory, P. J. *J. Chem. Phys.* **1942**, *10*, 51.
- Flory, P. J. *Principles of Polymer Chemistry*. Cornell University Press: Ithaca, New York, 1953.
- Kauzman, W.; Eyring, H. *J. Am. Chem. Soc.* **1940**, *62*, 3113.
- Rodriguez, T.; Patterson, D. *J. Chem. Soc. Faraday Trans. II* **1982**, *78*, 501.
- Rodriguez, T.; Patterson, D. *J. Chem. Soc. Faraday Trans. II* **1982**, *78*, 491.
- Park, J. H.; Hussam, A.; Couasnon, P.; Fritz, D.; Carr, P. W. *Anal. Chem.* **1987**, *59*, 1970.
- Cheong, W. J. *PhD Thesis*, Univ. of Minnesota: Minneapolis, 1989.
- Conder, J. R.; Young, C. L. *Physicochemical Measurements by Gas Chromatography*. Wiley-Interscience: New York, 1978.
- Thomas, E. R.; Eckert, C. A. *Ind. Eng. Chem. Process Des. Dev.* **1984**, *23*, 194.
- Snyder, L. R. *J. Chromatogr. Sci.* **1978**, *16*, 223.
- Snyder, L. R. *J. Chromatogr.* **1974**, *92*, 223.
- Rohrschneider, L. *Anal. Chem.* **1973**, *45*, 1241.
- Poppe, H.; Slaats, E. H. *Chromatographia* **1981**, *14*, 89.
- Kamlet, M. J.; Taft, R. W.; Carr, P. W.; Abraham, M. H. *J. Chem. Soc. Faraday I* **1982**, *78*, 1689.
- Castells, C. B.; Eikens, D. I.; Carr, P. W. *J. Chem. Eng. Data* **2000**, *45*, 369.
- Castells, C. B.; Eikens, D. I.; Carr, P. W. *J. Chem. Eng. Data* **2000**, *45*, 376.
- Castells, C. B.; Carr, P. W.; Eikens, D. I.; Bush, D.; Eckert, C. A. *Ind. Eng. Chem. Res.* **1999**, *38*, 4104.
- Tiegs, D.; Gmehling, J.; Medina, A.; Soares, M.; Bastos, J.; Alessi, P.; Kikic, I. *Activity Coefficients at Infinite Dilution Part I & II*. Schon & Wetzel GmbH: Frankfurt/Main, F. R. Germany, 1986.
- Abraham, M. H.; Grellier, P. L.; McGill, R. A. *J. Chem. Soc. Perkin Trans. II* **1987**, *1987*, 797.
- Santacesaria, E.; Berlendis, D.; Carra, S. *Fluid Phase Equilib.* **1979**, *3*, 167.
- Kovats, E. *Adv. Chromatogr.* **1965**, *1*, 229.
- Hildebrand, J. H.; Scott, R. L. *Regular Solutions*. Prentice-Hall: Englewood Cliff, New Jersey, 1962.
- Scatchard, G. *Chem. Rev.* **1931**, *8*, 321.
- Williamson, A. G. *Am. Rev. Phys. Chem.* **1964**, *15*, 63.
- Cruikshank, A. J. B.; Windsor, M. L.; Young, C. L. *Proc. Roy. Soc. A* **1966**, *295*, 271.
- Guggerheim, E. A. *Proc. Roy. Soc. A* **1944**, *183*, 213.
- McGlashan, M. L.; Morcom, K. W.; Williamson, A. G. *Trans. Faraday Soc.* **1961**, *59*, 601.
- Luckhurst, G. R.; Martire, D. E. *Trans. Faraday Soc.* **1969**, *65*, 1248.
- Dickson, J. N.; Daubert, T. E. *Ind. Eng. Chem. Res.* **1988**, *27*, 523.
- Massart, D. L.; Kaufman, L.; Rousseeuw, P. J.; Leroy, A. *Anal. Chim. Acta* **1986**, *187*, 171.
- Rousseeuw, P. J. *J. Am. Stat. Assoc.* **1984**, *79*, 871.
- Rousseeuw, P. J.; Leory, A. *Robust Regression and Outlier Detection*. Wiley-Science: New York, 1986.
- Gold, H. J. *Anal. Chem.* **1962**, *34*, 174.
- Kelhaian, H. V.; Renon, H. *Measurement, Evaluation and Prediction of Phase Equilibria*; Elsevier, Amsterdam, 1986.
- Sayegh, S. G.; Vera, J. H. *Chem. Eng. J.* **1980**, *19*, 1.
- Fredenslund, A.; Jones, R. L.; Prausnitz, J. M. *AIChE J.* **1975**, *21*, 1086.
- Gmehling, J.; Rasmussen, P.; Fredenslund, A. *Ind. Eng. Chem. Process Des. Dev.* **1982**, *21*, 118.
- Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.
- Staverman, A. *J. Rec. Trav. Chim. Pays-Bas* **1950**, *69*, 163.
- Guggenheim, E. A. *Mixtures*. Clarendon Press: Oxford, 1952.
- Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *80*, 4438.
- Ben-Naim, A. *J. Phys. Chem.* **1978**, *82*, 792.
- Ben-Naim, A. *Hydrophobic Interactions*. Plenum Press: New York, 1980.
- Ben-Naim, A. *Biopolymers* **1975**, *14*, 1337.