Microscopic Binding Constants in Cyclodextrin Systems: Complexation of α-Cyclodextrin with Sym-1,4-Disubstituted Benzenes

Kenneth A. Connors* and David D. Pendergast

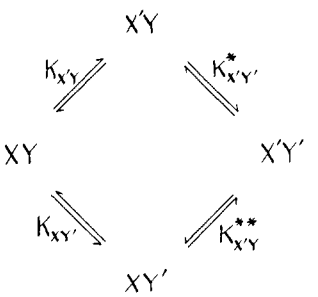
Contribution from the School of Pharmacy, University of Wisconsin, Madison, Wisconsin 53706.
Received May 18, 1984

Abstract: In a system consisting of a substrate XY having two binding sites and a ligand with one binding site (α-cyclodextrin in this study), there may exist two isomeric 1:1 complexes X′Y and XY′ and one 1:2 complex X′Y′. The observed stability constants are related to microscopic binding constants by $K_{11}^{XY} = K_{X'Y} + K_{XY'}$ and $K_{12}^{XY} = a_{XY}K_{X'Y}K_{XY'}/K_{11}^{XY}$, where a_{XY} is a site interaction parameter. If the two substrate sites are identical, as in XX, these equations become $K_{11}^{XX} = 2K_{X'X}$ and $K_{12}^{XX} = a_{XX}K_{11}^{XX}/4$, so $K_{X'X}$ and a_{XX} can be obtained. Ten XX-type substrates, sym-1,4-disubstituted benzenes, were studied in aqueous solution at 25 °C, allowing the evaluation of the microscopic binding constants $K_{X'X}$ and the interaction parameters a_{XX}. Empirical correlations of complex stability were found with (1) substrate solubility and binding site dipole moment and (2) substrate partition coefficient and heat of fusion. The microscopic binding constants $K_{X'X}$ correlate well with $K_{X'Y}$ values for the same binding site but different substituents, as estimated earlier, and give Hammett plots whose slopes correlate with complex strength according to $\rho_X = 0.324 \log K_{X'X} - 1.206$. These correlations allow $K_{X'Y}$ to be estimated for any X and Y. This permits the fractional isomeric solution composition of complexes X′Y and XY′ to be estimated. The interaction parameter is shown to be the equilibrium constant for the reaction $2X'X \rightleftharpoons X'X' + XX$; a_{XX} is analyzed in terms of three influences: (1) the electronic effect of cyclodextrin bound at site X′ on the nature of site X; (2) the repositioning effect resulting from displacement of the first cyclodextrin when the second cyclodextrin binds; and (3) the cyclodextrin–cyclodextrin interaction in the 1:2 complex. A free energy model of the binding process is developed, which gives an expression for K_{11}^{XY} in terms of K_{11}^{XX} and K_{11}^{YY}, and which shows that a_{XY} is not in general related to a_{XX} and a_{YY}. The energetics and forces of binding are discussed in terms of the relation $\Delta G°_{X'Y} = \Delta G°_{MM} + \Delta G°_{MS} + \Delta G°_{SS}$, where $\Delta G°_{MM}$ describes medium–medium interactions (the hydrophobic effect), $\Delta G°_{MS}$ medium–solute interactions (solvation effects), and $\Delta G°_{SS}$ solute–solute interactions. The cavity model of solvophobic effects applied to this system suggests that all inclusion complexes receive some stabilization via the solvophobic effect, but the extent of this is determined by solute–solute interactions as well as binding site hydrophobicity.

Cycloamyloses (cyclodextrins) are cyclic oligomers containing six or more D-glucose units linked 1–4. The six-, seven-, and eight-unit substances are called cyclohexaamylose (α-cyclodextrin), cycloheptaamylose (β-cyclodextrin), and cyclooctaamylose (γ-cyclodextrin), respectively. These molecules are doughnut shaped, and their possession of a cavity of fixed size and shape has led to considerable interest in their properties as host molecules for the formation of host–guest inclusion complexes. Many reviews are available.[1]

Most of the current work on cyclodextrins (and on modified cyclodextrins) is concerned with their catalytic properties, particularly as enzyme models and "artificial enzymes",[1c,d,2] and with their practical applications, primarily in pharmaceuticals and in chromatographic systems.[1e,f,g,3] Although many laboratories have described studies on the basic solution chemistry of cyclodextrins and their complexes, little of the current activity is directed toward such studies, and the present state of knowledge is unsatisfactory. Among the features of which we have, in a general sense, little understanding, and practically no capability for prediction are complex stoichiometries and structures, their thermodynamic stabilities, possible isomerism, and chemical and physical properties. Nor have the forces responsible for complex formation been unambiguously identified, though the published discussion on this subject is extensive.

Scheme I



It is evident that before we can claim to understand some of the interesting activities of the cyclodextrins, such as their catalytic properties and their roles as chromatographic stationary phases, and before we can hope to employ them in foods and pharmaceuticals, we must be able to describe their solution chemistry, in particular some of the features mentioned above. The present paper is one of a series from this laboratory describing systematic studies of substrate structure–complex stability relationships in cyclodextrin systems.

Binding Site Model. In this work the host (ligand, L) was α-cyclodextrin and the guest molecules (substrates, S) were 1,4-disubstituted benzenes. Such a host–guest system may be described as possessing two potential binding sites on the ligand (the two ends of the cavity) and two potential substrate binding sites. Then there may exist four possible isomeric 1:1 (SL) complexes, four 1:2 (SL_2) complexes, and four 2:1 (S_2L) complexes.[4] Since there appears to be no evidence for the significant presence of 2:1 complexes of organic substrates in α-cyclodextrin systems, we infer that only one end of the α-cyclodextrin cavity can be entered by the substrate. Consequently, the system can be described as having a single ligand binding site and two substrate binding sites. Two 1:1 complexes and a single 1:2 complex

(1) (a) French, D. *Adv. Carbohydr. Chem.* **1957**, *12*, 189. (b) Thoma, J. A.; Stewart, L. In "Starch: Chemistry and Technology", Whistler, R. L., Paschall, E. F., Eds.; Academic Press: New York, 1965; Vol. I, p 209. (c) Griffiths, D. W.; Bender, M. L. *Adv. Catal. Relat. Subj.*, **1973**, *23*, 209. (d) Bender, M. L.; Komiyama, M., "Cyclodextrin Chemistry"; Springer-Verlag: Berlin, 1978. (e) Saenger, W. *Angew. Chem., Int. Ed. Engl.* **1980**, *19*, 344. (f) Szejtli, J. "Cyclodextrins and their Inclusion Complexes"; Akadēmiai Kiadō: Budapest, 1982; D. Reidel Publishing Co., Dordrecht, Holland. (g) Pitha, J.; Szente, L.; Szejtli, J. In "Controlled Drug Delivery"; S. D. Brock, Ed.; CRC Press: Boca Raton, 1983; Vol. I, Chapter 5.

(2) (a) Breslow, R. *Adv. Chem. Ser.* **1980**, *191*, 1. (b) Tabushi, I. *Acc. Chem. Res.* **1982**, *15*, 66.

(3) Szejtli, J., Ed. "Proceedings of the First International Symposium on Cyclodextrins"; Akadēmiai Kiadō: Budapest, 1982; D. Reidel Publishing Co., Dordrecht, Holland.

(4) Rosanske, T. W.; Connors, K. A. *J. Pharm. Sci.* **1980**, *69*, 564.

0002-7863/84/1506-7607$01.50/0

may exist. The equilibria among these species are represented in Scheme I, where XY is the substrate with binding sites X and Y, and ligand binding at a site is represented by a superscript prime. In Scheme I, $K_{X'Y}$ is the microscopic binding constant for complex X'Y, and so on. The 1:2 complex X'Y' can be formed by adding a ligand molecule to either of the 1:1 complexes, and the existence of a 1:2 complex implies the existence of both isomeric 1:1 complexes.

For this scheme eq 1 and 2 are readily obtained,[5] where a_{XY}

$$K_{11}^{XY} = K_{X'Y} + K_{XY'} \quad (1)$$

$$K_{12}^{XY} = \frac{a_{XY}K_{X'Y}K_{XY'}}{K_{11}^{XY}} \quad (2)$$

$= K_{X'Y'}^{*}/K_{XY'} = K_{X'Y'}^{**}/K_{X'Y}$, and K_{11}^{XY} and K_{12}^{XY} are stepwise experimental stability constants for 1:1 and 1:2 complex formation, defined[6] by $K_{11} = [SL]/[S][L]$ and $K_{12} = [SL_2]/[SL][L]$. The quantity a_{XY} is an interaction parameter that measures the extent of interaction between sites X and Y in 1:2 complex formation. If the sites are independent, $a_{XY} = 1$, but there is in general no restriction on the value of the interaction parameter.

Clearly an important goal in a study of such a system should be the evaluation of a_{XY}, $K_{X'Y}$, and $K_{XY'}$. This is not possible solely through eq 1 and 2.[7] If, however, the two substrate binding sites are identical, so that $K_{X'Y} = K_{XY'}$, then (letting the substrate be denoted XX), eq 1 and 2 give

$$K_{11}^{XX} = 2K_{X'X} \quad (3)$$

$$K_{12}^{XX} = a_{XX}K_{11}^{XX}/4 \quad (4)$$

For such a system the microscopic binding constant $K_{X'X}$ and the interaction parameter a_{XX} can be evaluated.

Substrates having two identical binding sites constitute a special case but not a trivial one, since they offer a route to the determination of these fundamental quantities. Before the work described in this paper was undertaken, for example, nothing was known about the range of values the interaction parameter might assume. Besides the role of the binding site model in providing a quantitative description of the detailed equilibrium system, it also offers a useful conceptual framework by focussing attention on potential binding sites within a substrate molecule.

The substrates chosen for this study are sym-1,4-disubstituted benzenes. The potential binding sites (obviously the sites of substitution) are of a size range that can fit into the α-cyclodextrin cavity. Because of the proximity of the two sites, and their electronic interaction, it seemed unlikely that they would behave independently. These 1,4-disubstituted benzenes were of additional interest for what they might reveal when compared with 4-substituted benzoic acids,[5] phenols,[8] and anilines[9] as substrates for α-cyclodextrin.

Experimental Section

Materials. α-Cyclodextrin (Sigma) was dried for 3 h at 105 °C. Methyl orange was recrystallized from water.[8] The substrates were from commercial sources (Aldrich or Eastman); most of them were recrystallized before use, and all melting points were consistent with literature values.

Apparatus. Spectral measurements were made with a Cary-Varian Model 2200 or a Perkin-Elmer Model 559 spectrophotometer equipped with thermostated cell compartments. Solubility measurements were carried out with a thermostated water bath fitted to rotate sample vials end over end at 32 rpm. All measurements were made at 25.0 ± 0.1 C.

Procedures. Solubility Method. A solubility study was carried out by measuring the apparent solubility S_t of the substrate as a function of total ligand concentration L_t. Usually 5 to 10 L_t values were studied. Solid substrate (in excess of its equilibrium solubility) was added to glass vials also containing the ligand solution. (The ligand solubility was never exceeded.) The vials were sealed with Teflon-lined screw caps and were rotated in the constant-temperature water bath for 24 h. The samples were filtered through 0.22-μm Teflon filters contained in 25-mm Millipore Swag-Lok filter assemblies attached to 10-cm³ disposable syringes. An accurately measured portion was immediately withdrawn and diluted volumetrically in 1:1 methanol:water; this solvent system and dilution served to dissociate the complexes and to solubilize the substrate. The diluted solutions were analyzed spectrophotometrically for total substrate concentration at the wavelength of maximum absorption.[10]

The medium for the solubility measurements was 0.10 M NaCl, except for terephthalic acid, which was studied in 0.10 M HCl to repress ionization. Dimethyl terephthalate was studied in both media.

The solubility data were treated according to eq 5,[11] where S_0 is the solubility when $L_t = 0$.

$$\frac{S_t - S_0}{L_t - 2(S_t - S_0)} = \alpha + \beta[L_t - 2(S_t - S_0)] \quad (5)$$

$$\alpha = \frac{K_{11}S_0}{1 - K_{11}S_0} \quad (6)$$

$$\beta = \frac{K_{11}K_{12}S_0}{(1 - K_{11}S_0)^2} \quad (7)$$

The stability constants were obtained from the intercept α and slope β of a weighted least-squares regression fit of the data to eq 5. The variances of S_t and S_0 were determined experimentally, and Deming's method[12] was used to incorporate weighting contributions from both variables. The uncertainties in the stability constants were obtained through a propagation of errors treatment.

Competitive Indicator Method. One substrate, 1,4-dimethoxybenzene, was also studied by a competitive spectrophotometric method in which the substrate perturbs an indicator–cyclodextrin equilibrium with a resultant spectral change. This technique has recently been improved to permit the extraction of both K_{11} and K_{12} values.[13] A variant of the method uses solutions that are saturated with respect to the substrate; this solubility competitive indicator method was also applied to dimethoxybenzene. The indicator was methyl orange, and the medium was 0.10 M HCl.

Table I. Stability Constants for α-Cyclodextrin Complexes with Sym-1,4-Disubstituted Benzenes, X-C_6H_4-X, at 25 °C in 0.10 M NaCl[a]

X	K_{11}^{XX}/M^{-1}	K_{12}^{XX}/M^{-1}	$10^4S_0/M$
OCH_3	55.8 (4.6)	193 (39)	55.3 (0.2)
OC_2H_5	128 (16)	326 (49)	4.56 (0.05)
I	5060 (940)	6255 (1270)	0.031 (0.002)
Br	913 (22)	397 (14)	0.59 (0.01)
Cl	232 (16)	90 (12)	3.99 (0.11)
CO_2H	1344 (17)	23.8 (0.6)	0.192[c] (0.002)
CO_2CH_3	464 (8)	109 (4.7)	1.69[d] (0.02)
$COCH_3$	10.2 (0.7)	*b*	38.9 (0.03)
CN	33.1 (1.0)	7.2 (0.7)	6.97 (0.08)
NO_2	35.8 (1.0)	4.6 (0.4)	2.35 (0.04)

[a] Standard deviations in parentheses. [b] $K_{12} = 0$ at the $P = 0.05$ level. [c] In 0.10 M HCl. [d] $10^4S_0 = 1.79$ (0.01) in 0.10 M HCl.

Results

Ten sym-1,4-disubstituted benzenes were studied as substrates for α-cyclodextrin in aqueous solution at 25 °C and 0.10 M ionic strength. The results are listed in Table I, which includes K_{11}^{XX}, K_{12}^{XX} and S_0 values, with their estimated uncertainties. The relatively large uncertainties in the stability constants for diiodobenzene are a consequence of the difficulty in measuring S_0 for this insoluble substrate. The values in Table I were all measured by the solubility method, and they all refer to 0.10 M NaCl as the solvent, except for terephthalic acid, which was studied

(5) Connors, K. A.; Lin, S.-F.; Wong, A. B. *J. Pharm. Sci.* **1982**, *71*, 217.
(6) The equilibrium concentrations denoted by brackets are molar concentrations. The reference state is taken to be the experimental solvent, namely water at 25 °C, ionic strength 0.10 M.
(7) This is a classical problem that has been most carefully studied for acid–base equilibria, the proton playing the role of L in the present terminology; see: Edsall, J. T.; Wyman, J. "Biophysical Chemistry"; Academic Press: New York, 1958; Vol. I, pp 477–485.
(8) Lin, S.-F.; Connors, K. A. *J. Pharm. Sci.* **1983**, *72*, 1333.
(9) Wong, A. B.; Lin, S.-F.; Connors, K. A. *J. Pharm. Sci.* **1983**, *72*, 388.
(10) Pendergast, D. D. Ph.D. Thesis, University of Wisconsin—Madison, 1983.
(11) Iga, K.; Hussain, A.; Kashihara, T. *J. Pharm. Sci.* **1981**, *70*, 108.
(12) Deming, W. E. "Statistical Adjustment of Data"; Wiley: New York, 1943.
(13) Pendergast, D. D.; Connors, K. A. *J. Pharm. Sci.*, accepted for publication.

Table II. Stability Constants and Interaction Parameters for α-Cyclodextrin Complexes with *sym*-X-C_6H_4-X

X	K_{11}^{XX}/M^{-1}	K_{12}^{XX}/M^{-1}	a_{XX}[b]
NH_2[a]	2.3		
OCH_3	75.4	221	11.7 (1.0)
OC_2H_5	128	326	10.2 (1.3)
I	5060	6250	4.94 (0.95)
Br	913	397	1.74 (0.05)
Cl	232	90	1.55 (0.12)
CO_2H	1344	23.8	0.071 (0.001)
CO_2CH_3	454	106	0.93 (0.03)
$COCH_3$	10.2		
CN	33.1	7.2	0.87 (0.03)
NO_2	35.8	4.6	0.51 (0.02)

[a] Reference 9. [b] Standard deviation in parentheses.

in 0.10 M HCl. In order to learn if this difference in medium might have a significant effect on the parameters, dimethyl terephthalate was also studied in 0.10 M HCl. Its S_0 value was significantly different in 0.10 M NaCl and 0.10 M HCl (Table I), but the stability constants were not significantly different: K_{11} = 443 (13), K_{12} = 102 (2.7) in 0.10 M HCl (standard deviations in parentheses).

Because of their properties, in particular their low solubilities, most of these systems could not be studied by other techniques. 1,4-Dimethoxybenzene, however, could be studied by the methyl orange competitive indicator method. The standard technique[13] gave K_{11} = 55.3 (22), K_{12} = 208 (90); the combined solubility–competitive indicator method yielded K_{11} = 75.4 (2.6), K_{12} = 221 (8.6). These results are consistent with those in Table I.

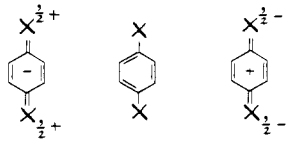
Table II gives our best current estimates of K_{11}^{XX} and K_{12}^{XX} for eleven sym-1,4-disubstituted benzenes and lists the a_{XX} values calculated by means of eq 4.

Discussion

Empirical Correlations. In the conventional sense of the term a "reaction series" is a set of compounds having a common reaction site and a variable substituent. In this sense, therefore, the substrates studied here do not constitute a reaction series because both the reaction site and the substituent are variable. There is nevertheless a familial relationship among these compounds as a consequence of their size, shape, and aromatic character. It is therefore reasonable to look for patterns in their complexing behavior.

In accounting for structure–stability relationships in series of 4-substituted benzoic acids,[5] phenols,[8] and anilines,[9] the following postulates have been valuable. It is proposed that complex stability (that is, at a binding site) is increased by an increase in site electron density, that it is increased by an increase in site polarizability, and that it is decreased by an increase in site polarity (in a polar solvent). These postulates will be applied to the present data.

One general pattern can be accounted for on the basis of the symmetry of this type of substrate, since its electronic structure can be represented as a combination of these valence bond representations: Thus very strongly electron-donating substituents



1 2 3

will lead to an electronic distribution heavily weighted with **1**, whereas powerfully electron-withdrawing groups will lead to a major contribution from **3**. For both of these extremes, binding site polarity should be high, hence we expect that substituents at the extremes of the Hammett σ scale should be complex destabilizing. This pattern is seen in the K_{11}^{XX} data in Table II.

It is fairly obvious from Table I that there is an inverse relationship between K_{11}^{XX} and S_0. For seven of these substrates a good linear free energy relationship was observed (see Figure 1). The equation of this line is

$$\log K_{11}^{XX} = -0.59 \log S_0 + 0.40 \quad (8)$$

with correlation coefficient r = 0.98. The three substrates failing to follow this correlation (X = NO_2, CN, $COCH_3$) all have high group (site) dipole moments,[14] which we anticipate will be complex destabilizing. A good correlation for all ten substrates was obtained by performing a multiple linear regression on log S_0 and μ (the group dipole moment), the regression equation (r = 0.94) being

$$\log K_{11}^{XX} = -0.636 \log S_0 - 0.231\mu + 0.524 \quad (9)$$

These correlations suggest that there is a similarity between inclusion complex formation and the dissolution process (or rather its reverse).

Figure 2 shows a plot of log K_{11}^{XX} against log P, where P is the *n*-octyl alcohol/water partition coefficient of the substrate.[15] For log $P > 3$ a reasonable linear relationship is seen, whereas the stability constants for substrates having log $P < 3$ fall significantly above the extension of this line. This correlation is greatly improved by making use of a free energy analysis of the dissolution process, eq 10, in which $\Delta G°_S$ is the standard molar free energy change upon dissolution; $\Delta G°_f = \Delta H°_f - T\Delta S°_f$ is the standard free energy of fusion; and $\Delta G°_t$ is the standard free energy change for the transfer of substrate from its melt to water, all at temperature T.

$$\Delta G°_S = \Delta H°_f - T\Delta S°_f + \Delta G°_t \quad (10)$$

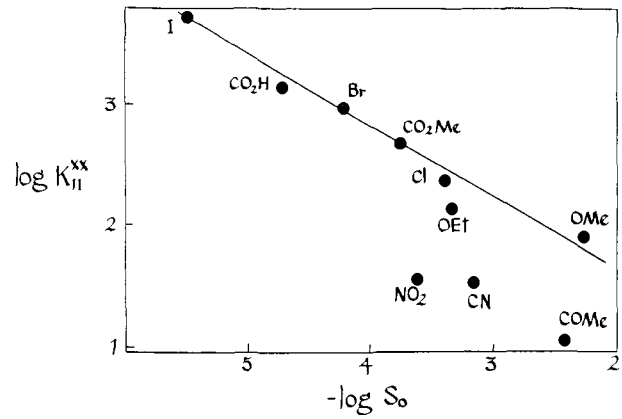


Figure 1. Plot of log K_{11}^{XX} against $-\log S_0$ for sym-1,4-disubstituted benzenes. The line, which was drawn with eq 8, is the regression line calculated with the seven highest K_{11}^{XX} values.

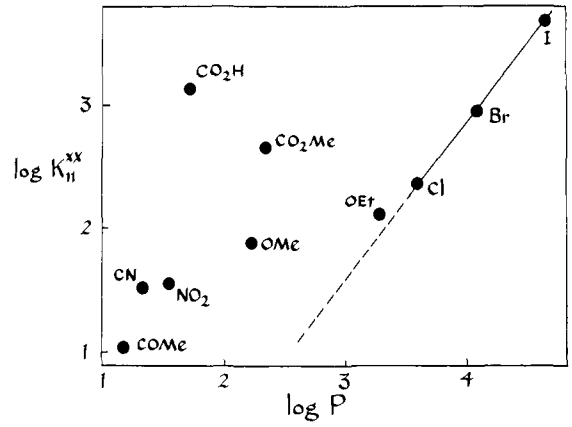


Figure 2. Plot of log K_{11}^{XX} against log P (*n*-octyl alcohol/water partition coefficient). The line is drawn through the halogenated substrate points.

(14) Lien, E.; Guo, Z.; Li, R.; Su, C. *J. Pharm. Sci.* **1982**, *71*, 641.
(15) log P values were computed by summing the appropriate hydrophobic fragmental constants: Nys, G. G.; Rekker, R. F. *Eur. J. Med. Chem.* **1974**, *9*, 361.

Table III. Microscopic Binding Constants $K_{X'Y}$ for 1,4-Disubstituted Benzene–α-Cyclodextrin Systems at 25 °C

X	Y	$K_{X'Y}/M^{-1}$	X	Y	$K_{X'Y}/M^{-1}$
NO_2	NO_2	17.9	Cl	NH_3^+	68.6
NO_2	COO^-	72.2[a]	Cl	Cl	116
NO_2	OH	245	Cl	OH	272
NO_2	NH_2	635	Cl	NH_2	251
NO_2	O^-	2408	Cl	O^-	488
Br	Br	457	I	I	2530
Br	OH	704	I	OH	2316
Br	O^-	1221	I	O^-	3955
CN	CN	16.6	COOH	COOH	672
CN	COO^-	71.5[a]	COOH	F	504
CN	OH	158	COOH	H	722
CN	NH_2	451	COOH	CH_3	1146[a]
CN	O^-	662	COOH	OCH_3	891[a]
			COOH	OH	1115[a]
			COOH	NH_2	1341
			COOH	$NHCH_3$	1301

[a] Mean of results in ref 5 and 10.

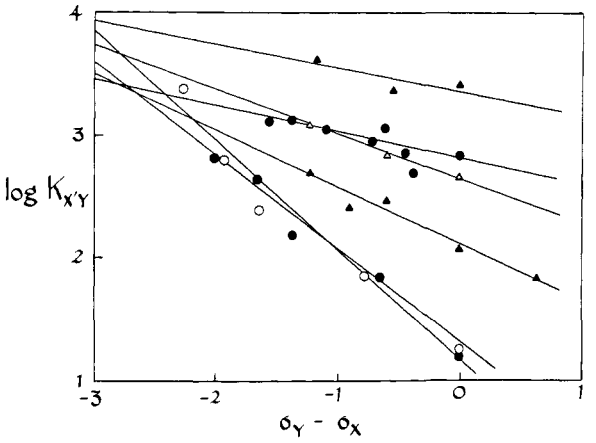


Figure 3. Hammett plots of microscopic binding constants in Table III. Identification of lines from top to bottom at far right: X = I, COOH, Br, Cl, CN, NO_2. Symbols: open circles, X = NO_2; filled circles, X = CN or COOH; open triangles, X = Br; filled triangles, X = Cl or I.

log P is taken as a measure of $\Delta G°_t$, and since $\Delta S°_f$ is nearly a constant for this type of compound,[16] a multiple linear regression could be carried out on the five substrates (X = Cl, Br, I, NO_2, CO_2CH_3) for which heats of fusion were available.[16,17] This gave eq 11 (r = 0.998). For the four additional substrates whose heats

$$\log K_{11}^{XX} = 0.428\Delta H°_f + 0.881 \log P - 2.665 \quad (11)$$

of fusion could be estimated from their melting points, eq 11 gave good agreement with experimental K_{11} values.[18]

Microscopic Binding Constants. Microscopic constants $K_{X'X}$ are available through eq 3 and the data in Table II, and the constants $K_{X'X'}$ are available through the relationship $K_{X'X'} = a_{XX}K_{X'X} = 2K_{12}^{XX}$.

The binding constants $K_{X'X}$ are, by themselves, of only moderate interest, but in combination with other data they provide very useful results. These other data are drawn from our studies on 4-substituted benzoic acids, phenols, and anilines.[5,8,9] By making use of the postulates on the effects of binding site electron density, polarizability, and polarity, it was possible to conclude that for some of these XY-type substrates the observed K_{11}^{XY} represented binding at solely one site (say X) and could therefore be interpreted, according to eq 1, as the microscopic binding constant $K_{X'Y}$.

(16) Martin, E.; Yalkowsky, S.; Wells, J. *J. Pharm. Sci.* **1979**, *68*, 565.
(17) Eliott, J. H.; Chris, M. D. *J. Chem. Eng. Data* **1968**, *13*, 475.
(18) For X = OCH_3 and OC_2H_5, $\Delta S_f°$ = 12.7 cal/K (mean for a large number of similar compounds[16]); for X = CN and $COCH_3$ the $\Delta S_f°$ = 15.0 cal/K of 1,4-dinitrobenzene seemed more appropriate. Terephthalic acid does not melt cleanly.

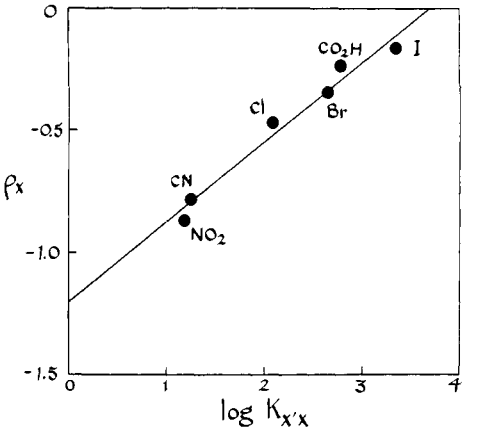


Figure 4. Linear correlation of ρ_X (slope of line in Figure 3) with log $K_{X'X}$.

Table III gives microscopic constants $K_{X'Y}$ for 1,4-disubstituted benzenes for systems in which we have constants for at least three Y substituents.[19]

Figure 3 is a Hammett plot of the binding constants in Table III against ($\sigma_Y - \sigma_X$). This abscissa scale places the values for all XX-type substrates at ($\sigma_Y - \sigma_X$) = 0. The trends are very clear, and the correlations are quite good.[20] One obvious conclusion is that the $K_{X'X}$ values, which are unambiguous, fall on the same lines as the $K_{X'Y}$, and this confirms the binding site assignments that were earlier made on the basis of structure–stability relationships.

The equations of these lines can be written

$$\log K_{X'Y} = \rho_X(\sigma_Y - \sigma_X) + \log K°_{X'Y} \quad (12)$$

where $K°_{X'Y} = K_{X'X}$ if the correlation is perfect. The slopes, ρ_X, clearly depend upon X, and a very simple relationship is observed as shown in Figure 4; the stronger the binding, the less susceptible it is to substituent effects. The equation of the line in Figure 4 is

$$\rho_X = 0.324 \log K_{X'X} - 1.206 \quad (13)$$

Equation 13 allows an interesting prediction to be made. According to the postulate that an increase in site electron density favors binding, ρ_X for binding at X cannot be positive, hence its maximum value is zero; alternatively, the observation in the preceding paragraph leads to the conclusion that when binding is maximal, ρ_X = 0. Placing this condition in eq 13 leads to an estimate of the maximum possible microscopic binding constant at a substituted benzene site. This value is log $(K_{X'X})_{max}$ = 3.72, or $(K_{X'X})_{max} = 5.3 \times 10^3$ M^{-1}. Since ρ_X = 0 under this condition, this is also the maximum value for $K_{X'Y}$.

No $K_{X'X}$ or $K_{X'Y}$ in Table III exceeds this estimate, but there is one $K_{X'X'}$ value, for diiodobenzene, that is calculated from Table II to be 12.5×10^3 M^{-1}. Since this may be significantly larger than $(K_{X'X})_{max}$, it can be inferred that this constant may not describe simple binding at a substituted benzene site; there may be an extra stabilizing effect, presumably an interaction between the two cyclodextrins in the 1:2 complex.

Equations 12 and 13 will be found useful later to estimate $K_{X'Y}$ values.

The Interaction Parameter. Since a_{XX} is a ratio of equilibrium constants ($a_{XX} = K_{X'X'}/K_{X'X}$), it is itself an equilibrium constant; a_{XX} is the equilibrium constant for the disproportionation:

$$2X'X \overset{a_{XX}}{\rightleftharpoons} X'X' + XX$$

Thus if the binding sites are independent a_{XX} = 1, and if $a_{XX} \neq 1$ the sites are not acting independently. Hence this parameter is a convenient measure with which to investigate the interaction between sites.

(19) For some of these compounds it is probable that there is a very small contribution to K_{11}^{XY} from binding at Y, but the effect is probably smaller than the experimental uncertainty in K_{11}^{XY}.

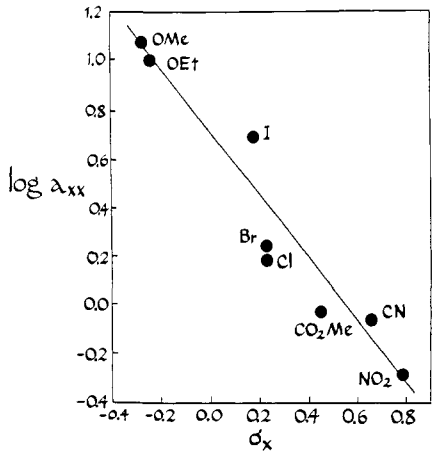


Figure 5. Hammett plot of a_{XX}.

We partition the influences on the magnitude of a_{XX} into the following three classes.

(i) *The electronic effect of L bound at site X' on the nature of site X*: If the sites in XX are electron deficient, as indicated by **1**, upon interaction of one of them with L to give X'X there will be a partial electron transfer from L to the binding site. This has the effect of increasing the charge density at site X in X'X relative to that at X in XX. Thus binding of the second ligand will be favored relative to that of the first one, and a_{XX} will be greater than unity through the operation of this effect. If the sites are electron rich the opposite drift of charge takes place, and a_{XX} will be less than unity. Thus a_{XX} may be expected to follow a Hammett plot with a negative slope.

(ii) *The repositioning effect*: In the 1:1 complex the relative position of ligand and binding site is optimal with regard to lowering the total free energy of the system. Formation of the 1:2 complex will result in adjustment of all three molecules to minimize the total free energy, since in the 1:2 complex X'X' the two bound sites are necessarily identical on average. This may require a repositioning of the substrate–ligand orientation that was reached in the 1:1 complex. Any such repositioning must therefore be destabilizing, since the orientation in the 1:1 complex is optimal, and will therefore lower a_{XX}.

(iii) *The ligand–ligand interaction effect*: In a 1:2 complex there is a posibility that the facing rims of the two cyclodextrin molecules may interact attractively (a sort of substrate-promoted ligand dimerization). Such an effect could only be manifested as 1:2 complex stabilizing (increasing a_{XX}), because any destabilizing repulsive interactions would be accounted for in terms of the repositioning effect.

Figure 5 is a Hammett plot of a_{XX} values. The correlation equation (r = 0.96) for eight substrates is

$$\log a_{XX} = -1.29\sigma_X + 0.68 \tag{14}$$

The slope is negative as expected. It is not clear whether or not this line provides a measure of the electronic effect uncontaminated by superimposed repositioning and ligand–ligand interaction effects, but it is quite evident that terephthalic acid (X = COOH) behaves atypically, the point for this system falling far below the correlation line. From the above arguments, this must be a consequence of a major repositioning effect, which suggests that the COOH group is very deeply inserted into the cyclodextrin cavity in the 1:1 complex.

It is worth noting that it does not necessarily follow that the sites are independent if a_{XX} = 1; it is quite possible for the several effects to combine so as to yield this result fortuitously.

A Free Energy Binding Model. From n symmetrical 2-site substrates XX or YY, a total of $n(n-1)/2$ distinguishable unsymmetrical 2-site substrates (XY) can be constructed. It is therefore of practical interest to investigate the possibility of predicting the complexing properties of XY from experimental studies on XX and YY. To do this we adapt a free energy model of Leffler and Grunwald[21] to the complexing system.

We define the standard molar free energy of XX as the sum of contributions from the two sites:

$$G^\circ_{XX} = 2G_X \tag{15}$$

and similarly for YY. In like manner the free energies of the 1:1 complexes X'X and Y'Y are written:

$$G^\circ_{X'X} = G_{X'} + G_X \tag{16}$$

where G_X in (16) is established by its definition in (15). Now the four determined quantities G_X, G_Y, $G_{X'}$, and $G_{Y'}$ are introduced into other free energy functions, which, however, in general require the addition of interaction terms to account for nonadditive effects, as follows:

$$G^\circ_{X'Y} = G_{X'} + G_Y + I_{X'Y} \tag{17}$$

$$G^\circ_{XY'} = G_X + G_{Y'} + I_{XY'} \tag{18}$$

$$G^\circ_{X'X'} = 2G_{X'} + I_{X'X'} \tag{19}$$

and so on.[22]

Formation of 1:1 complexes is represented XX $\rightleftharpoons$ X'X, XY $\rightleftharpoons$ X'Y, etc., and the standard molar free energy changes for these processes are

$$\Delta G^\circ_{X'X} = G^\circ_{X'X} - G^\circ_{XX} \tag{20}$$

$$\Delta G^\circ_{X'Y} = G^\circ_{X'Y} - G^\circ_{XY} \tag{21}$$

$$\Delta G^\circ_{XY'} = G^\circ_{XY'} - G^\circ_{XY} \tag{22}$$

Substitution from eq 15–19 into eq 20–22 gives

$$\Delta G^\circ_{X'Y} = \Delta G^\circ_{X'X} + (I_{X'Y} - I_{XY})$$

$$\Delta G^\circ_{XY'} = \Delta G^\circ_{Y'Y} + (I_{XY'} - I_{XY})$$

Writing $\Delta G^\circ_{X'X} = -RT \ln K_{X'X}$, etc., gives

$$\ln K_{X'Y} = \ln ({K_{11}}^{XX}/2) - (I_{X'Y} - I_{XY})/RT \tag{23}$$

$$\ln K_{XY'} = \ln ({K_{11}}^{YY}/2) - (I_{XY'} - I_{XY})/RT \tag{24}$$

Combining eq 1 with eq 23 and 24:

$${K_{11}}^{XY} = \frac{{K_{11}}^{XX}}{2B_{X'Y}} + \frac{{K_{11}}^{YY}}{2B_{XY'}} \tag{25}$$

where

$$RT \ln B_{X'Y} = (I_{X'Y} - I_{XY}) \tag{26}$$

$$RT \ln B_{XY'} = (I_{XY'} - I_{XY}) \tag{27}$$

Thus ${K_{11}}^{XY}$ is a linear combination of ${K_{11}}^{XX}$ and ${K_{11}}^{YY}$; however, the weighting factors $B_{X'Y}$ and $B_{XY'}$ are not obtainable from measurements solely on XX and YY.

Since a_{XX} is an equilibrium constant, this kind of treatment leads to eq 28 and 29.

$$RT \ln a_{XX} = -I_{X'X'} \tag{28}$$

$$RT \ln a_{YY} = -I_{Y'Y'} \tag{29}$$

In the same way it is seen that a_{XY} is the equilibrium constant for this reaction,

$$X'Y + XY' \overset{a_{XY}}{\rightleftharpoons} X'Y' + XY \tag{30}$$

and the free energy treatment applied to a_{XY} gives

$$RT \ln a_{XY} = (I_{X'Y} + I_{XY'}) - (I_{X'Y'} + I_{XY}) \tag{31}$$

Comparison of eq 28, 29, and 31 shows that a_{XX} and a_{YY} may

(20) σ for Y = O^- was taken to be −1.0; see: Berliner, E.; Monack, L. C. *J. Am. Chem. Soc.* **1952**, *74*, 1574. Hine, J. "Physical Organic Chemistry"; McGraw-Hill Book Co.: New York, 1956; p 72. Enhanced σ values were used for anilines and phenols. The lines in Figure 3 are unweighted linear least-squares regression lines.

(21) Leffler, J. E.; Grunwald, E. "Rates and Equilibria of Organic Reactions"; John Wiley and Sons: New York, 1963; pp 139–141.

Table IV. Calculation of Isomeric Fractional Composition in XY Systems

X	Y	$K_{X'Y}$[a]	$K_{XY'}$[b]	$f_{X'Y}$	$f_{XY'}$	a_{XY}[c]
Cl	Br	116	406	0.22	0.78	1.52
Br	I	475	1414	0.25	0.75	5.94
CN	COOH	25	431	0.05	0.95	1.06
NO_2	COOH	33	281	0.11	0.89	0.88
Cl	COOH	88	615	0.13	0.87	0.27

[a] Equation 34. [b] Equation 1. [c] Equation 2.

contain no information about a_{XY}. The arithmetic mean of ${K_{11}}^{XX}$ and ${K_{11}}^{YY}$ may be a reasonable approximation to ${K_{11}}^{XY}$, whereas the mean of a_{XX} and a_{YY} may have nothing to do with a_{XY}, in general. (When X and Y are very "similar," a_{XY} may approach a_{XX} and a_{YY} because when X = Y, I_{XY}, $I_{X'Y}$, and $I_{XY'}$ are all zero by definition.)[23]

Estimation of Isomeric Composition. In an α-cyclodextrin–XY system there may exist the isomeric 1:1 complexes X'Y and XY'. Since these will have different physical and chemical properties, it is of interest to be able to estimate the solution isomeric fractional composition, which can be expressed

$$f_{X'Y} = \frac{[X'Y]}{[X'Y] + [XY']} = \frac{K_{X'Y}}{{K_{11}}^{XY}} \quad (32)$$

$$f_{XY'} = K_{XY'}/{K_{11}}^{XY} \quad (33)$$

It is therefore necessary to estimate $K_{X'Y}$ and $K_{XY'}$. By relying solely on observations of XX and YY, we can use the correlations developed earlier to estimate these constants. Equation 12 is written

$$\log K_{X'Y} = \rho_X(\sigma_Y - \sigma_X) + \log K_{X'X} \quad (34)$$

$$\log K_{XY'} = \rho_Y(\sigma_X - \sigma_Y) + \log K_{Y'Y} \quad (35)$$

where $K_{X'X} = {K_{11}}^{XX}/2$, $K_{Y'Y} = {K_{11}}^{YY}/2$, and $\rho_X(\rho_Y)$ is calculated with eq 13. However, this method does not take advantage of data that may be available on XY. An alternative approach is to make use of measurements on the XY system. Define X and Y such that ${K_{11}}^{YY} > {K_{11}}^{XX}$. Then calculate $K_{X'Y}$ with eq 34, and finally calculate $K_{XY'}$ with eq 1, using the experimental ${K_{11}}^{XY}$. The calculation may equivalently be carried out in terms of eq 25 by estimating $B_{X'Y}$ with eq 36 and then calculating $B_{XY'}$ with eq 25.

$$\log B_{X'Y} = -\rho_X(\sigma_Y - \sigma_X) \quad (36)$$

When X and Y are defined in this way, $B_{XY'}$ is quite insensitive to substantial variation in $B_{X'Y}$. Table IV gives the results for five XY systems that were studied experimentally.[24] The interaction parameter a_{XY} was calculated with eq 2.

It appears that significant fractions of two isomeric complexes may exist in these systems. Several authors have commented on the possible existence of isomers, but they have interpreted their nuclear magnetic resonance data to indicate that isomers did not exist in their systems. For example, Bergeron et al.[25] studied numerous substituted nitrophenols and nitrophenolates, showing that only the nitro group penetrates the cyclodextrin cavity. This is consistent with our conclusion based on substituent effects;[8] these are essentially one-site substrates. Gelb et al.[26] examined many carboxylic acid and phenol substrates, finding no evidence for isomerism; 4-nitrobenzoic acid could not be studied owing to its low solubility. However, Table IV indicates that the ratio of nitro inserted to carboxylic acid inserted isomers is only about 1/9 for 4-nitrobenzoic acid, so it is a difficult substrate with which to detect isomerism. Better candidates for such a search could be designed by means of the data and calculational methods described above.

The a_{XY} values in Table IV, which might seem anomalous in the absence of theoretical guidance, are acceptable in light of the results of the free energy model. For chlorobromobenzene, in which the X and Y sites are very similar as measured by a_{XX} and a_{YY}, a_{XY} is very near to a_{XX} and a_{YY}. For the other substrates a_{XY} bears no obvious relationship to a_{XX} and a_{YY}, as expected from the model equations. These a_{XY} values, we recall, are equilibrium constants, and their relative values may provide information about the nature of the 1:2 complex. Thus the 4-substituted benzoic acids in Table IV constitute a reaction series whose a_{XY} values suggest the following. For cyanobenzoic and nitrobenzoic acids a_{XY} is approximately unity, far larger than the a_{YY} value of 0.071 for terephthalic acid. Apparently transfer of the second cyclodextrin molecule to the cyano or nitro site does not result in a destabilizing repositioning effect, as happens in terephthalic acid. In chlorobenzoic acid, however, there is a repositioning effect, with 1:2 destabilization. These effects are consistent with the stronger binding at each site in the chlorobenzoic acid. The picture suggested is that COOH site penetration is deep in the 1:2 complexes of cyanobenzoic and nitrobenzoic acids (similar to that in the 1:1 terephthalic acid complex), whereas penetration of COOH in the 1:2 chlorobenzoic acid complex is shallower, being more like that in the 1:2 terephthalic acid complex. The enhanced a_{XY} value for bromoiodobenzene suggests possible ligand–ligand interaction in its 1:2 complex, as described for diiodobenzene.

When studying the properties of cyclodextrin complexes the fractional isomeric composition is especially useful because the isomers will in general have different properties. Consider, as an example, the cyclodextrin-catalyzed hydrolysis of *m*-nitrophenyl acetate described by Van Etten et al.[27] and ascribed by them to binding of the nitro site in the cyclodextrin cavity with consequent positioning of the ester function near the secondary hydroxyls on the rim of the cavity. We use the calculational method given above, with ${K_{11}}^{XY} = 53\ M^{-1}$ as reported[27] and ${K_{11}}^{XX} = 22\ M^{-1}$, where X = NO_2 and Y = $OCOCH_3$.[28] The result is $K_{X'Y} = 21\ M^{-1}$, $K_{XY'} = 32\ M^{-1}$, $f_{X'Y} = 0.4$, $f_{XY'} = 0.6$. Thus only about 40% of the 1:1 complex exists as the catalytically productive form, and the actual rate enhancement is more than twice that calculated if isomerism is not taken into account.

Binding Forces. Many authors have discussed the nature of the binding forces in cyclodextrin complexes; Bender and Komiyama[1d] have reviewed this subject. There is no general agreement, and most of the possible modes of interaction have been suggested by one group or another. Besides the fundamental attractive forces, several chemical phenomena have been proposed as "driving forces" for complexation; these are the hydrophobic effect,[29,30] release of strain energy in the cyclodextrin ring,[31] release of cavity-bound "high-energy" water,[25] and hydration of the

(22) Interaction terms can be included in all of these equations, but some of them can be defined to be zero, as is done here for I_{XX}, I_{YY}, $I_{XX'}$, $I_{YY'}$. There is an advantage in retaining all the terms in a fuller development of the model, for then formal relationships among them may be more apparent.

(23) According to the separability postulate, the interaction terms can be factored, i.e., $I_{XY} = I_X I_Y$, etc.[21] This allows the right side of eq 30 to be written $-(I_{X'} - I_X)(I_{Y'} - I_Y)$. However, it does not provide a general relationship for a_{XY} in terms of a_{XX} and a_{YY}.

(24) Two of the 4-substituted benzoic acid results are the means of data reported in ref 5 and 10. 4-Chlorobenzoic acid, 4-chlorobromobenzene, and 4-bromoiodobenzene were studied by the solubility method[10] with these results (${K_{11}}^{XY}$, ${K_{12}}^{XY}$, units M^{-1}, standard deviations in parentheses): 4-chlorobenzoic acid, 703 (6.6), 21 (1.4); 4-chlorobromobenzene, 522 (26), 137 (12); 4-bromoiodobenzene, 1889 (64), 2111 (88).

(25) Bergeron, R.; Rowan, R. *Bioorg. Chem.* **1976**, *5*, 425. Bergeron, R.; Channing, M. A. *Bioorg. Chem.* **1976**, *5*, 437. Bergeron, R.; McPhie, P. *Bioorg. Chem.* **1977**, *6*, 465. Bergeron, R. J.; Channing, M. A.; Gibeily, G. J.; Pillor, D. M. *J. Am. Chem. Soc.* **1977**, *99*, 5146.

(26) Gelb, R. L.; Schwartz, L. M.; Cardelino, B.; Fuhrman, H. S.; Johnson, R. F.; Laufer, D. A. *J. Am. Chem. Soc.* **1981**, *103*, 1750.

(27) Van Etten, R. L.; Sebastian, J. F.; Clowes, G. A.; Bender, M. L. *J. Am. Chem. Soc.* **1967**, *89*, 3242.

(28) 1,3-Dinitrobenzene is XX here; it was studied by the solubility method.[10] 1,3-Disubstituted benzenes are actually 3-site substrates, but in the compounds considered in this example the 5-position is deactivated and binding at this site should be negligible. These calculations are obviously more approximate than in the case of the para-substituted substrates.

(29) Wishnia, A.; Lappi, S. J. *J. Mol. Biol.* **1974**, *82*, 77.

(30) Komiyama, M.; Bender, M. L. *J. Am. Chem. Soc.* **1978**, *100*, 2259. But this conclusion, based on the thermodynamics of binding of 1-adamantanecarboxylate, is disputed by Gelb et al.: Gelb, R. I.; Schwartz, L. M.; Laufer, D. A. *Bioorg. Chem.* **1980**, *9*, 450.

(31) Manor, P. C.; Saenger, W. *Nature (London)* **1972**, *237*, 392.

complex.[32] Note that most of the "driving forces" mentioned above are effects of the solvent. In keeping with the main point of this paper, we focus attention on the binding site, and its corresponding microscopic stability constant, rather than on the substrate molecule as a whole. In order to provide a framework for discussion, we employ eq 37. Here $\Delta G°_{X'Y}$ has the meaning

$$\Delta G°_{X'Y} = \Delta G°_{MM} + \Delta G°_{MS} + \Delta G°_{SS} \qquad (37)$$

given earlier; that is, $\Delta G°_{X'Y} = -RT \ln K_{X'Y}$, where $K_{X'Y}$ is the microscopic binding constant for binding to site X. Equation 37 partitions this free energy change into three components. $\Delta G°_{MM}$ is the contribution arising from *medium–medium* interactions, and it embodies the hydrophobic effect (solvophobic effect in general); $\Delta G°_{MS}$ includes all *medium–solute* interactions, that is, all solvation phenomena; and $\Delta G°_{SS}$ includes all *solute–solute* interactions, which would be those between ligand and binding site X in 1:1 complex X'Y, but would include L–X, L–Y, and possibly L–L interactions in 1:2 complex X'Y. All three of the free energy components in eq 37 owe their existence ultimately to the basic forces of interaction. Equation 37 thus provides a convenient way to differentiate between the roles of the intermolecular forces and of the "driving forces" (the $\Delta G°$ terms themselves) in determining $\Delta G°_{X'Y}$. The three free energy terms are not independent.

The hydrophobic contribution $\Delta G°_{MM}$ can be discussed in terms of the cavity model of the solvophobic effect. (In the present paragraph the word "cavity" refers to the cavity in the solvent that must be created when a solute molecule is inserted into the solvent. The cyclodextrin intramolecular space is called its "interior".) The high surface tension of a polar solvent like water leads to a minimization of solvent surface area. The driving force for association of two solute molecules is the reduction in free energy resulting from the reduction in surface area as the two cavities containing the separated solute species coalesce into a single cavity containing the complex. Theory shows that, to a first approximation, the solvophobic contribution to the free energy of binding is proportional to the product of the decrease in cavity surface area and the solvent surface tension,[33] and these dependencies have been found experimentally for stacking interactions of planar molecules in aqueous media.[34,35] The nature of these dependencies for the insertion of a substrate binding site into the cyclodextrin interior is an interesting point that requires experimental study, but we can examine the qualitative reasonableness of this model by making use of the available data. These show that a good estimate of K_{11} for molecular complex formation in water at 25 °C, on the average over many types of compounds, is given by

$$\log (K_{11}/M^{-1}) = 0.077A - 1.74 \qquad (38)$$

where A is the area of overlap, in angstroms squared, between the two molecules in the complex.[36] In an earlier section we estimated that the maximum value of log K_{11} for binding between α-cyclodextrin and a substituted phenyl binding site is about 3.72. Using this quantity in eq 38 leads to an estimate of 71 Å^2 as the overlap area between the binding site and the cyclodextrin interior. This seems reasonable since the entire area of the cyclodextrin interior is about 95 Å^2 (the interior approximated as a cylinder of diameter 4.5 Å and depth 6.7 Å).[1d] This calculation is presented not for its quantitative significance but to show that the direction suggested may be a fruitful one. Note that, in this view of the process, the hydrophobic contribution to complex stability is a function of depth of penetration of the binding site into the cyclodextrin interior, for this determines the overlap area. The depth of penetration is, in turn, determined by several factors, including the MS and SS terms in eq 37, and it is in this sense that $\Delta G°_{MM}$, $\Delta G°_{MS}$, and $\Delta G°_{SS}$ are not independent. The solvophobic contribution can only be stabilizing if the solvent is more polar than the interior of the cyclodextrin. Any factors inhibiting the hydrophobic effect will appear as destabilizing effects elsewhere.

The $\Delta G°_{MS}$ quantity represents all solvation phenomena, and it can be either stabilizing or destabilizing. The polarity of the substrate binding site is of major interest in controlling this term, since usually substrate solvation will be lost or much reduced upon insertion of the site into the cyclodextrin (although deep penetration may permit recovery of site solvation on the "far" side of the cyclodextrin, as may occur with cinnamate anion complexes[37]).

$\Delta G°_{SS}$ may receive contributions from electrostatic, induction, and dispersion forces, and from hydrogen bonding. Binding site charge density and polarizability are important factors. Binding site polarity can contribute to stabilization through electrostatic and induction forces via $\Delta G°_{SS}$, whereas this same property may be destabilizing via $\Delta G°_{MS}$. $\Delta G°_{SS}$ can only make a stabilizing contribution to $\Delta G°_{X'Y}$.

We can now discuss the present experimental results in these terms. The substrates in Table II represent a wide range in properties, as exemplified by their solubilities and partition coefficients. The octyl alcohol/water partition coefficient P is widely considered to be a measure of hydrophobicity.[38] In a cyclodextrin substrate we are interested in the binding site properties, and because of the symmetry of the substrates in Table II we can take the P value for the compound to be representative of the hydrophobicity of the binding site. Figure 2 is a plot of complex stability against log P. Two features are of special interest, namely the apparently linear correlation for highly hydrophobic sites and the positive deviations for less hydrophobic sites. The linear correlation in the halogen series cannot be unambiguously assigned to either the $\Delta G°_{MM}$ or the $\Delta G°_{SS}$ term in eq 37, and the slope of the line is probably influenced by both phenomena. Its extrapolation is therefore uncertain, and it may underestimate the contribution of the hydrophobic effect at low P values; other lines could be drawn. Nevertheless, it is evident that more than one effect may contribute to complex stability, and this demonstration is a useful consequence of using a set of substrates without a common reaction site. The $\Delta G°_{SS}$ term may have a major role in the terephthalic acid and dimethyl terephthalate complexes. However, the hydrophobic effect is not determined solely by the inherent hydrophobicity of the binding site, but rather by the depth of penetration into the cyclodextrin interior, according to the picture developed above. Thus it is conceivable that in the terephthalic acid complex (which on the basis of a_{XX} values has already been identified as unusual), a highly specific site-ligand interaction, perhaps hydrogen bonding, produces the deep penetration already inferred, and which consequently results in a major stabilization from the hydrophobic effect. In this sense the hydrophobic contribution may truly be an effect, not a cause, of the association. Solutes of high inherent hydrophobicity will enjoy little solvation stabilization as separated species, and the solvophobic cavity effect will drive them together, with $\Delta G°_{MM}$ making the principal contribution to $\Delta G°_{X'Y}$. Solutes of low hydrophobicity may have large positive $\Delta G°_{MS}$ values because of strong solvation of the separated species, and hence little tendency to complex. However, there may be substrates

(32) Gerasimowicz, W. V.; Wojcik, J. F. *Bioorg. Chem.* **1982**, *11*, 420.

(33) (a) Sinanoglu, O.; Abdulnur, S. *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **1965**, *24*, *Suppl. 15*, S-12. (b) Sinanoglu, O. In "Molecular Associations in Biology"; Pullman, B., Ed.; Academic Press: New York, 1968; pp 427–445. (c) Halicioglu, T.; Sinanoglu, O. *Ann. N.Y. Acad. Sci.* **1969**, *158*, 308.

(34) Connors, K. A.; Sun, S. *J. Am. Chem. Soc.* **1971**, *93*, 7239.

(35) (a) Connors, K. A.; Infeld, M. H.; Kline, B. J. *J. Am. Chem. Soc.* **1969**, *91*, 3597, 5697. (b) Cohen, J. L.; Connors, K. A. *J. Pharm. Sci.* **1970**, *59*, 1271.

(36) This is the origin of eq 38: The standard unitary free energy change is a linear function of A passing through the origin,[35b] and it is proportional to γ, the surface tension,[34] or $\Delta G°_{\mu} = k\gamma A$, where k is a dimensional constant. We also have $\Delta G°_{\mu} = -RT \ln K_x$, where K_x is the stability constant on the mole fraction scale. K_x and K_{11} are related by $K_x = K_{11}M^*\rho$, where M^* is the number of moles of solvent contained in 1000 g of solvent and ρ is the solvent density. Combining these gives $\log K_{11} = \gamma A/3.15\ T - \log (M^*\rho)$, or for water at 25 °C, eq 38. In this equation A is the overlap area; the change in cavity surface area is $2A$, and the absence of this factor of 2 in the slope of eq 38 probably is fortuitously compensated for by the effect of surface curvature at molecular dimensions on the surface tension.

(37) Bergeron, R. J.; Channing, M. A.; McGovern, K. A.; Roberts, W. P. *Bioorg. Chem.* **1979**, *8*, 263.

(38) Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525.

(sites) of low hydrophobicity that are able to seek out features within the cyclodextrin interior and thus to achieve stabilizing contributions from $\Delta G°_{SS}$, with consequent penetration of the cyclodextrin interior to a depth determined by the particular interactions; this penetration further stabilizes the complex through a $\Delta G°_{MM}$ contribution. According to this viewpoint every inclusion complex is at least partially stabilized by a contribution from the $\Delta G°_{MM}$ term (in solvents that are more polar than the interior of the host).

Acknowledgment. This work was supported by a grant from The Upjohn Company.

Stereospecific Synthesis of Substituted *cis*-Hydrindan-5-ones and Their Regiospecific Enolization and Functionalization: Synthetic Intermediates for Reserpine

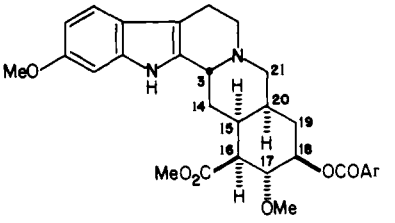
Michael E. Jung*[1] and Lynn A. Light[†]

Contribution from the Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90024. *Received March 7, 1984. Revised Manuscript Received July 11, 1984*

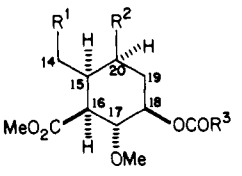
Abstract: An approach to reserpine (**1**) is described which is based on an anionic oxy-Cope rearrangement to establish three of the five contiguous asymmetric centers of the E ring and a regiospecific enol silylation coupled with hydroboration–oxidation to introduce the last two centers. The key intermediate **8b** for the anionic oxy-Cope rearrangement is prepared from the readily available 7,7-dimethoxynorbornenone **6** in 50% overall yield by addition of propargyl alcohol dianion, hydride reduction, and TBS formation. Rearrangement of **8b** followed by hydrogenation affords the desired 4-[(silyloxy)methyl]-1,1-dimethoxytetrahydroindan-6-one **10** in which all the protons were cis in nearly 80% yield. Force-field calculations on the two regioisomeric methyl enol ethers derived from **10** indicate a strong preference (96:4 at 25 °C) for enolization toward the ring juncture. Indeed treatment of **10** under thermodynamic conditions (BMDA, Me_3SiCl) produced a single (>95%) enol ether assigned the desired Δ^{17} (reserpine numbering) structure **11** by a combination of spectroscopic and chemical means. Conversion of **11** into its TBS enol ether **15** followed by hydroboration–oxidation introduced the last two asymmetric centers in an E-ring precursor **5** of reserpine (93% for the last step). Finally the unusual base-catalyzed hydrolysis of the dimethyl ketal which occurs in this process is rationalized by strong intramolecular complexation of a methoxyl group with the boron atom leading to ultimate displacement of this methoxyl group with hydroxide.

Reserpine (**1**), a *Rauwolfia* alkaloid with good antihypertensive properties, has been synthesized several times since Woodward's classical synthesis in the late 1950s.[2] This first synthesis prepared the E-ring derivative **2a** in which the five contiguous asymmetric centers were correctly established and then coupled it to 6-methoxytryptamine to give, after reduction, the lactam **3a** and thence reserpine in several further steps. Pearlman[3] prepared an analogous E-ring intermediate **2b** from a different precursor and then followed the Woodward route to reserpine. In Wender's synthesis,[4] the cyclic amine **4** was prepared by a completely different approach and taken onto reserpine. We have begun an approach to reserpine in which the E-ring target molecule is of lower functionality, namely, the aldehyde tosylate **2c**, which on reaction with 6-methoxytryptamine would give directly the desired immonium salt necessary for cyclization to reserpine.[5] We now report the anionic oxy-Cope rearrangement of the readily available substrate **8b** to produce a *cis*-hydrindenone with three of the five contiguous asymmetric centers established and the regiospecific enolization of a derivative of this hydrindenone which permits the introduction of the final two centers with the correct stereochemistry. In this manner, the hydrindanone **5**, with the correct stereochemistry at the five contiguous asymmetric centers, has been prepared in eight or nine steps in good overall yield (12–14%).

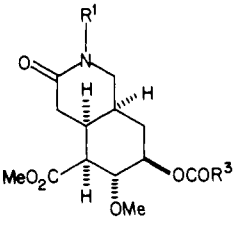
In our earlier work on the rearrangement of substituted allylic alcohols derived from the bicyclic enone **6**, we showed that the stereochemistry of the tertiary hydrogens about the periphery of the cyclohexanone system was all cis due to the geometric demands of the transition state.[6] We reasoned that the use of a simple *trans*-3-alkoxypropenyl-substituted alcohol, e.g., **8b**, would permit the easy preparation of the hydrindenone **9** with the all-cis arrangement of the three contiguous asymmetric centers. This was put into practice as follows: Reaction of the enone **6** (Scheme I) with the dilithium salt of propargyl alcohol dianion gave the endo adduct **7a** in 83% crude yield (62% recrystallized, mp 93–94



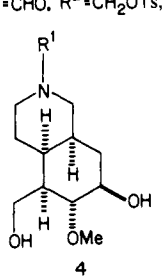
1, Ar = 3,4,5-$(MeO)_3C_6H_2$



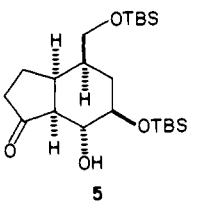
2a, R^1=CO_2Me, R^2=CHO, R^3=Me
b, R^1=CO_2Me, R^2=CHO, R^3=Ar
c, R^1=CHO, R^2=CH_2OTs, R^3=Ar



3a, R^3=Me, R^1=6-methoxytrytophyl
b, R^3=Ar, R^1=6-methoxytrytophyl



4



5

[†] Deceased September 20, 1984.

0002-7863/84/1506-7614$01.50/0