Linear Free Energy Relationships and Pairwise Interactions in Supramolecular Chemistry

Hans-Jörg Schneider
FR Organische Chemie, Universität des Saarlandes, D 66041 Saarbrücken, Germany

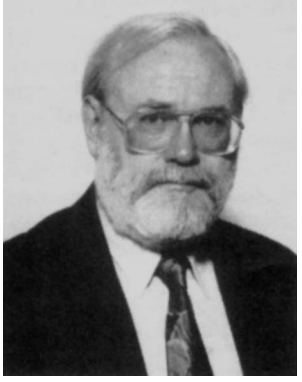
1 Introduction

In recent years the chemistry of host–guest complexes has become one of the most rapidly expanding fields of science.[1–3] Until now most of the activity has been directed towards the (frequently demanding) synthesis of host molecules, and recently also to functions of the supramolecular systems, *e.g.* as elements for sensors, molecular switches, *etc.* The essential basis of a multitude of possible applications is molecular recognition by non-covalent forces. Therefore in contrast to traditional organic chemistry, which most often concentrates on the development of suitable methods for the synthesis of new compounds, almost all studies in the supramolecular field involve physical measurements, in the first place of complexation constants. The literature already contains thousands of such data, which until now in most cases have only been used for qualitative *ad hoc* interpretations. Systematic analyses by physical organic methods should allow the development of an experimental basis for a comprehensive description of the essential non-covalent forces and for a rational approach to the design of new host–guest systems.

Linear free energy relationships have in the past provided the most useful way for the quantification of rates and equilibria in bond-breaking and -making reactions, for their dependence on microenvironment, for structure–activity correlations, and for obtaining insight into the reaction mechanisms.[4] The analysis of binding energies and conformations with host–guest complexes in solution has exactly the same aims for the chemistry of the non-covalent bond.[5] At the same time one can test and improve calculational models by properly designed experiments and systems, which – in contrast to biopolymers – allow focusing on any particular interaction, usually moreover in better defined conformational space.

The strategy of the approach discussed in this review is based on the summation of pairwise interactions in the form of empirical complexation free energy increments;[5] at the present time under the condition of sufficient geometrical matching between complementary binding sites, and/or negligible strain changes during complexation. For the control of both conditions, molecular mechanics calculations often play an essential role. NMR spectroscopy is of vital significance to check the geometric conditions in solution. Microenvironment effects can be quantified by, for example, empirical solvent hydrophobicity scales, or by Debye–Hückel correlations to ionic strength, *etc.*[5]

Hans-Jörg Schneider was born in Stuttgart, Germany in 1935. He studied chemistry in Tübingen, Munich, and Berlin and received his Ph.D. under Michael Hanack at the Tübingen Institut Walter Hückels in 1967. After postdoctoral research at the University of California, San Diego (1967—1969) he worked on his Habilitation and other research until 1971 in the laboratory of Walter Hückel. In 1971 he completed his Habilitation in Saarbrücken and since 1973 he has been Professor of Organic Chemistry at the Universität des Saarlandes. His areas of interest include: conformational analyses, quantitative structure–reactivity relationships for aliphatic compounds, including steroids, NMR spectroscopic methods with an emphasis on ^{13}C and ^{1}H NMR shielding parameters, selective functionalization of paraffins, and in the last few years, supramolecular receptor and enzyme models.



2 Pairwise Interactions in Seemingly Simple 1:1 Associations

The description of molecular complexes by pairwise interactions has a long tradition in theoretical as well as in empirical approaches. Molecular mechanics calculations, for example, inherently involve summation of interactions between the parts of supramolecular entities. For relatively simple systems they have already led to encouraging results, particularly since the advent of the free energy perturbation method,[6] although the latter is very time consuming. The limitations here are particularly in the applied potential functions and their parametrizations. Thus, almost all force fields neglect electron lone-pair directionality as well as polarizations, although these can play a major role (see Section 7).

The alternative approach of extracting parameters quantifying interactions of particular functionalities in organic molecules from appropriate series of measurements was first introduced to non-covalent systems by Gutmann[7] and by Drago[8] *et al.* They found that for many systems a single electron donor–acceptor parameter (*e.g.* C_D, C_A) for a particular function involved for example in a Lewis acid–base complex is sufficient to describe the equilibrium using multiplicative combination of C_A and C_B, or by an additive combination of the corresponding free energy or enthalpies (equation 1). Sherry and Purcell, Ioghansen, as well as Abraham, Raevsky, Taft, and others have shown that simple 1:1 hydrogen-bonded associations in aprotic solvents obey similar linear combinations of parameters which essentially reflect basicities and acidities of the participating functions.[9] In a number of cases additivity rules were also established on the basis of enthalpies ΔH instead of Gibbs values ΔG, as well as with mechanistic refinement of single parameters, reflecting for example charge-transfer, polarizibilities *etc.*[8,9] Raevsky *et al.* were able to describe over 900 associations, measured in carbon tetrachloride by linear correlations (equation 1), with coefficients above $r = 0.98$ on the basis of single donor and acceptor parameters ED and EA;[9a] later these proved to be ideal for the description of ionophores (see Section 8).

$$\Delta G = \text{const}^{*} C_A C_B \qquad (1)$$

All the associations discussed until now are based on molecules with single functions, generally assuming simple 1:1 complexes. The problems arising, however, from the occurrence of a multitude of other possible complexes (often including self-association) in these sterically mostly unbiased systems are evident from the conflicting results obtained for example with *N*-methylacetamide association in chloroform, which were expected to deliver cornerstone values for peptides;[10] they vary from 4 to 7 kJ/mol. Another serious limitation lies in ill-defined conformations, illustrated by the unresolved problem even of seemingly simple and intensively studied systems such as ammonia[11a] or benzene[11b] dimers. This makes it difficult to apply such values to larger molecular associations with more selective

orientations. It is for these reasons that more straightforward answers are in fact obtained with more complicated, yet better defined structures as we see them in synthetic host–guest complexes.

3 Additivity in Supramolecular Complexes

A further problem in the application of additive interactions arises from the restriction of translatory and rotatory freedom accompanying association. The entropy disadvantage has been estimated by Jencks, Page, and others[12] to amount to up to $T^*\Delta S = 60$—150 kJ/mol, part of which may be compensated by desolvation process. Supramolecular systems have the advantage of minimizing differential entropy effects since complex formation with the accompanying loss of mobility is already 'paid' to a large degree by one or few interactions so that additional pairwise interactions will cost less entropy sacrifice. Page and Jencks made use of this 'anchor' principle in 1971, applying it to ligand–biopolymer binding; however, only the result of exchanging one function was evaluated.[13]

The simultaneous pairwise action of many non-covalent forces is an important feature not only of synthetic, but also of natural receptor–effector complexes. Increasingly, the latter have also been used for deriving energetic factors of associations. Williams *et al.* have estimated that adverse $T\Delta S$ contributions can be anywhere between 9 and 45 kJ/mol, being stronger as expected for more exothermic complex formation.[14] Site-directed mutagenisis has been used by Fersht *et al.* to estimate binding contributions of different amino-acid residues.[15] Unfortunately, the numbers for an amide–amide hydrogen bond reported from the studies of some peptides and of proteins vary between 24 (Ref. 14) and 2 (Ref. 15) kJ/mol. The analyses of biological systems is hampered by their flexibility and complexity, leading to problems of identifying which groups are interacting in only approximately known geometries. Another reason for the observed large discrepancies must be seen in the again largely unknown involvement of water, ions, or other groups interfering at the ligand sites of biopolymers.

With synthetic host–guest complexes these problems can be largely overcome.[5,16a] Equation 2 summarizes the total complexation energy ΔG_t as the result of the thermodynamic circle containing, besides the energy ΔG_{HG} which results from direct interactions between host H and guest G, the desolvation energies ΔG_{ds} of H and G – which might imply only desolvation at the contact sites – as well as re-solvation ΔG_s of the complex. Again, the advantage of multifunctional host–guest complexes is that, for a given host, ΔG_{ds} of course remains constant, and that the desolvation ΔG_{ds} for guests varies essentially with the number of functions to be desolvated before complexation, and thus will become part of the increment $\Delta\Delta G$ per function which contributes to complexation. The latter assumption – like others – will be tested by success or failure of linear correlations between the experimentally observed ΔG_t and the sum $\sum\Delta G$ from all participating functions with the increment ΔG for each (equation 3). A further test is whether one obtains the same or different increments ΔG using different host structures with, of course, the same type of interactions and the same environment with respect to solvents and salts.

$$\Delta G_t = \Delta D_{HG} - \Delta G_{dsH} - \Delta G_{dsG} + \Delta G_{HGs} \quad (2)$$

$$\Delta G_t = l^*\Delta G_L + m^*\Delta G_M + n^*\Delta G_N + \ldots \quad (3)$$

A complex between multifunctional host and guest structures (Figure 1) may contain different numbers l, m, n *etc.* of different types of interaction characterized each by an increment ΔG_L, ΔG_M, ΔG_N *etc.* In some complexes like aromatic ion pairs (see Section 7) one function may exert *several* types of interactions, such as, for example, an ion acting on a benzene ring as well as on a counterion. In the most simple case – such as in ion pairs within aliphatic frameworks (see Section 4) – the summation according to equation 3 results in a linear correlation of ΔG_t against the number n of (for example) possible salt bridges. This indicates

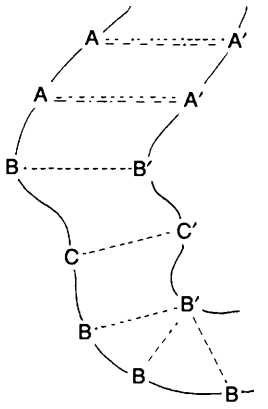
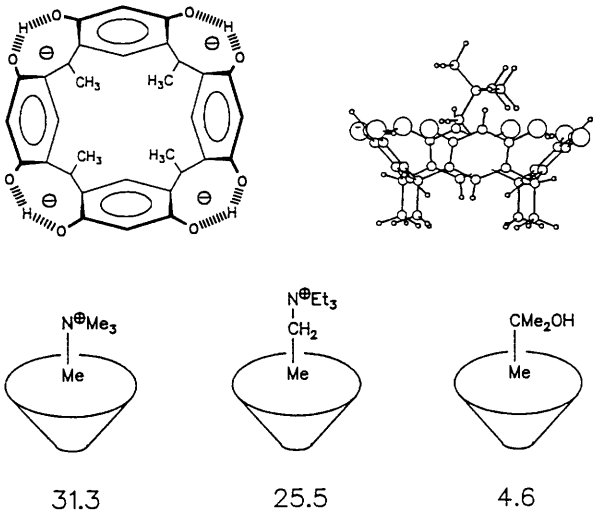


Figure 1 Additive interactions between host and guest. In this example there are 4 equivalent interactions between complementary B sites, 2 different ones (2 of kind l, 2 of kind m) between the A sites, and 1 between the C sites, hence $\Delta G_t = 4\Delta G_B + 2\Delta G_{A,l} + 2\Delta G_{A,m} + 1\Delta G_{GC}$

that one single increment is sufficient, which then can be obtained from the slope of the regression line. If more than one interaction type is involved than a multilinear regression will be required, or, better, increments for one type which have been determined independently will be used (see Sections 7, 8). A third, related approach is based on the use of relative parameters, characterizing, for example, electron donor–acceptor capacities of functions occurring in a complex. These may stem from entirely different measurements or calculations (Section 8).

4 Salt Bridges: From Simple Ion Pairs to DNA Groove Binding

The literature contains thermodynamic values for many inorganic ion pairs or, for example, tetramethylammonium salts, mostly in water as solvent, which often show large variations in entropy and enthalpy contributions.[17] Many of the organic host compounds which early on were proven to be very efficient binders for charged substrates, such as nucleotides,[18] choline *etc.*[19] rest essentially on Coulomb interactions. The latter can be identified if one compares the ΔG_t values of uncharged eqivalents, *i.e.* the macrocyclic tetraphenolate (1) complex with either tetramethylammonmium ions or with the isosteric electroneutral t-butyl alcohol (Scheme 1). This example[19] also demonstrates how, to a first approximation, the number of salt bridges



Scheme 1 The macrocyclic tetraphenolate, structure 1, and its complex with $^+NMe_4$. Selected complexation free energies ([kJ/mol], water, 25 °C, extrapolated to ionic strength = 0) (Ref. 19).

is counted independently of the charge delocalization, which in this particular tetra-anion is very strong

The plot of experimental free complexation energies ΔG_t in water *vs* the number *n* of possible salt bridges shows a linear correlation (Figure 2) with a regression value of 5 ± 1 kJ per mole and per salt bridge [5 16a] The correlation, which within the error passes through the origin, today comprises over 70 ion pairs These range from smaller to larger ions with low and high polarizibility, such as metal, sulfonium or ammonium ions, halides, sulfonates, phosphates, carboxylates, phenolates, *etc* [16 20] The 10 kJ/mol observed for zinc sulfate in water[16a] for example is close to the expected value for divalent ions from Bjerrum theory,[16 17 21] although the latter neglects entropy contributions and should hold only for small and spherical ions The surprisingly small variability of the observed $\Delta\Delta G$ increment reflects compensation effects well known from LFER studies In particular, a gain in ΔH, for example in hard–hard combinations of ions, will lead to less favourable ΔS values owing to the then stronger electrostriction of solvent molecules, the opposite will be true for interactions between soft ions Similarly, large and polarizable or soft ions will suffer from smaller Coulomb attractions but gain from charge polarization, *vice versa* arguments hold then for small and hard ions

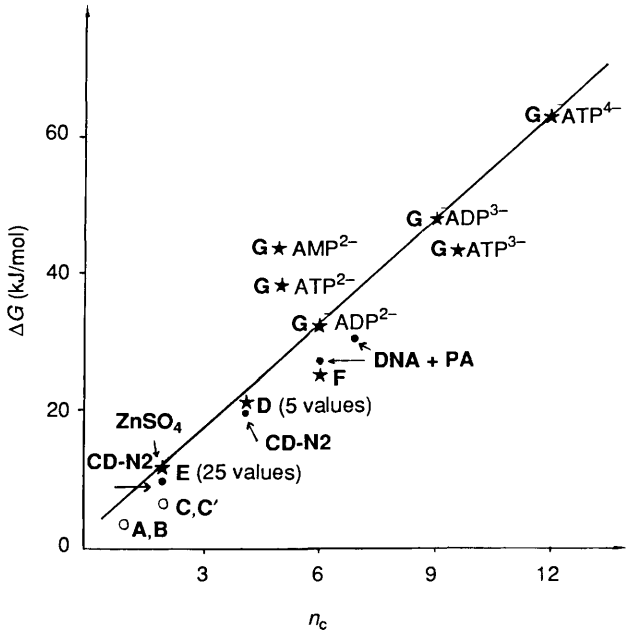


Figure 2 Experimental complexation free energies ΔG_t [kJ/mol] *vs* number *n* of salt bridges in ion pairs, in water For identification of the points see Ref 16*a* CD-N2 amino–cyclodextrin complexes (A V Eliseev and H -J Schneider, *Angew Chem Int Ed Engl* 1993, **32**, 1331), DNA + PA selection from Figure 3

The affinities of polyamines with DNA also are predictable with the 5 kJ increment per charge [22] In line with this, there is a fairly linear correlation of the affinities, as measured by a fluorescence assay, and the number of nitrogen atoms in the amines[23] (Figure 3) The linearity observed is remarkable for an inhomogeneous biopolymer, larger deviations occur only if the amine used bears additional functionalities It should be noted not only that natural amines such as spermine are on the line but that permethylation does not alter the affinities [22] This means that hydrogen bonds play a minor role in these polyamine associations, in agreement with measurements of complexes with either protonated or again permethylated azoniacyclophanes which, within + 1 kJ/mol, showed the same ΔG_t values with organic ions [24]

What are the problems involved in this rather simple analysis? The first step requires the assignment of the number of interactions responsible for the complex formation For simple structures such as (1) + choline (Scheme 1) or for small inorganic ion pairs one can assume contact, or secure it with the help of CPK or ball-and-stick models The latter are also helpful with more extended structures such as the azacrown ether derivatives complexing, for example, triphosphates [2 18] The uncertainties introduced here are simply, for example, whether one accounts for 10 or 12 salt bridges, and thus do not alter the correlation (Figure 2) substantially However, for complicated structures such as DNA complexes computer-aided molecular modelling is almost indispensible for localizing the possible interactions Such simulations reported in the literature[25] provided the basis of assigning an increment of 4—6 kJ mol and salt bridge to published DNA affinities of polyamines as well as of polyhistidine [22] Another complication arises if interactions other than salt bridges contribute in addition A possible way to handle this situation, which occurs in aromatic ion pairs, will be discussed in Section 7

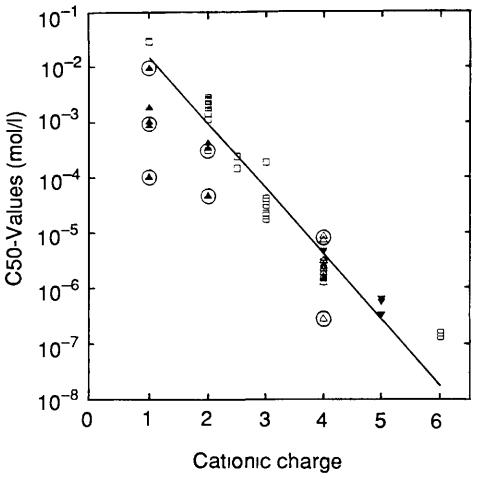


Figure 3 Affinities of polyamines (PA) to ds-DNA (measured as C_{50} values with ethidiumbromide, see text) against the number of positive charges in the amines Values for (1) aliphatic protonated PA (Ref 23), (2) permethylated non-cyclic PA, (3) permethylated azoniacyclophanes (Ref 22), (4), (5) different new PA (D Ruf, B Palm, unpublished results), (6) PA with additional affinities by naphthalene and/or amide residues (B Palm, unpublished) Linear correlation coefficient with (1),(2),(4) $r = 0\ 97$

5 Medium and Salt Effects on Ion Pairs

Ion pair stabilities usually increase with decreasing solvent polarities [17 21] With some aromatic ion pairs we observe an increase, by an order of magnitude, in 80% dioxane compared with pure water [16a] However, the simple dependencies (*e g* on the dielectric constant) assumed earlier do not generally hold For large polarizable host ions such as (1) (Scheme 1) an *opposite* stability increase with increasing water content in binary solvent mixtures has even been found [26] For the possible differentiation of tight and solvent-separated ion pairs which might blur the picture here the reader is referred to special texts [17 21]

In contrast to solvent effects, salt effects on supramolecular ion complexes show more consistent trends as well as the expected increase in ΔG_t with decreasing ionic strength [26 27] Even with large and quite anisotropic organic ions bearing multiple charges, surprisingly linear correlations with Debye–Huckel coefficients were observed[26 27] (Figure 4) However, larger polarizable ions, which are often used in organic buffers, may within a cavity of suitable charge and size lead to strong deviations owing to specific complexation [26]

6 Hydrogen Bonds in Peptide and Nucleobase Analogues

Hydrogen bonds of the amide-amide type dominate in the most efficient artificial hosts of high selectivity, developed in particu-

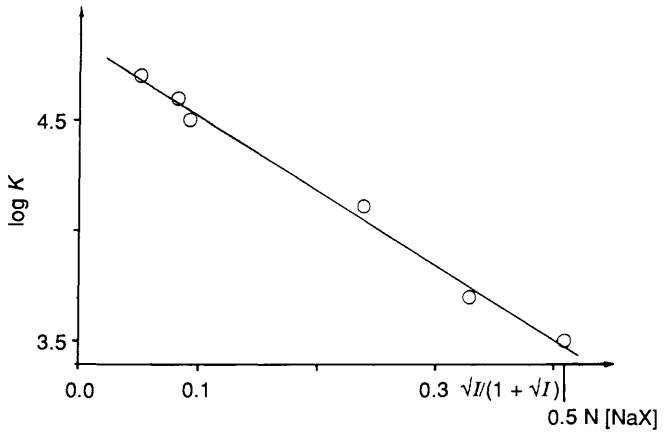
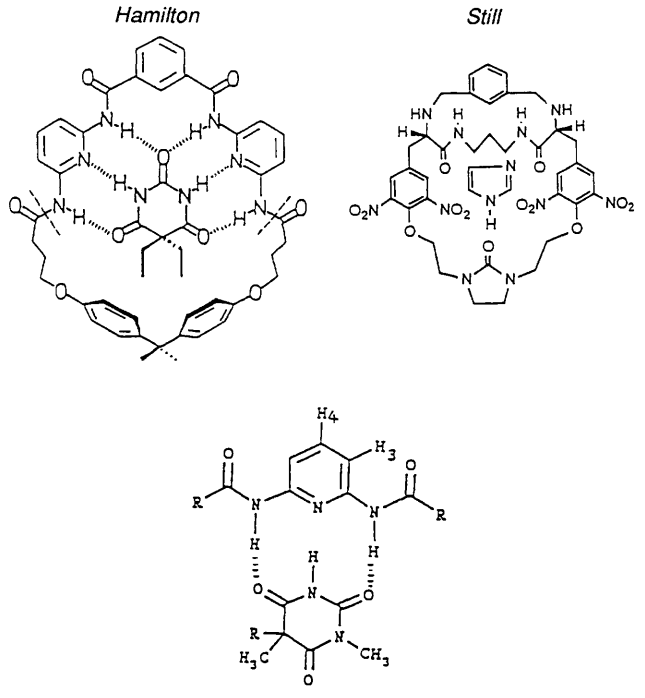


Figure 4 Correlation of ion pair lg K values with Debye–Hückel coefficients of ionic strength[26] for cyclophane structure (1) + Et_4NBr.



Scheme 2 Some examples for amide-type hydrogen bonds (see Refs. 5 and 28 for details).

lar by Hamilton, Rebek, Still, Zimmerman, and others (Scheme 2).[5,28] They contain a number n of well-tailored hydrogen bond donors and acceptors which again show a linear correlation with the measured complexation free energies[28] (Figure 5). Deviations from the correlation line are seen only for complexes which either suffer from particular entropy disadvantage as result of their acylic parts, or contain additional stacking interactions.

The strong influence of the medium on the complexation can be quantified by the hydrogen bonding capacity of the solvent which competes in high molarity for the solute. This leads to an increase in the constants K by factor of 10—20 if CCl_4 is used instead of $CHCl_3$.[28]

How does the structure and arrangement of donor and acceptor groups affect the results?[29,30] Although the measurements with corresponding model compounds are far from being accurate[29,31] we note that the Watson–Crick nucleobase pairs roughly follow the additivity scheme for amides (see above). For **A–T** with 2 hydrogen bonds one observes in chloroform $K = 10^2$ M^{-1}, for **G–C** with 3 bonds (Scheme 3) $K = 10^4$ M^{-1} or higher,[29,31] which is already more than expected. Other triply

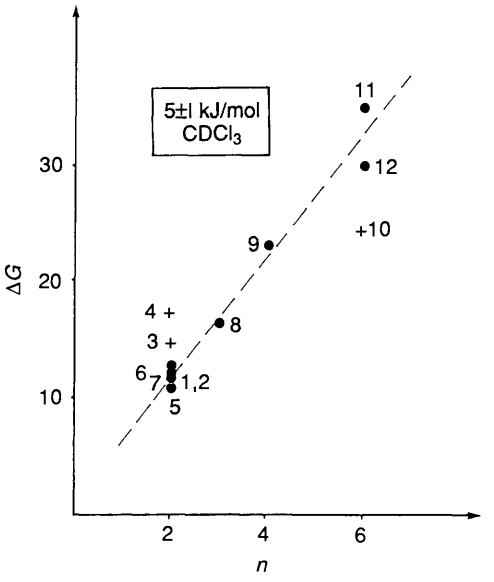
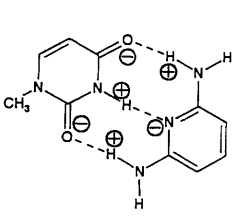


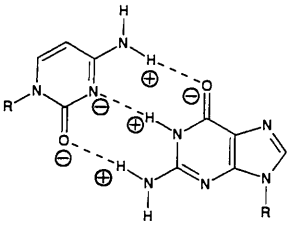
Figure 5 Free complexation energies of amide-type complexes measured in chloroform as a function of the number n of hydrogen bonds; for identification of points and explanation for the systems deviating from the regression see Ref. 28.

bound associations such as **U–DAP** (Scheme 3) show surprisingly low constants of around 10^2.[29]

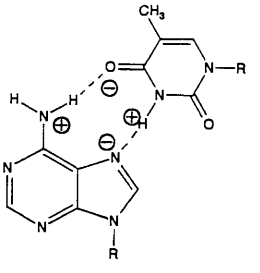
Jorgensen *et al.*, on the basis of molecular mechanics/MC calculations, have convincingly demonstrated that the reasons for these discrepancies are *secondary electrostatic interactions* between the adjacent donor and acceptor groups.[29] Besides the direct hydrogen bond, to which we assign an increment P in



U–DAP
3 P –4 S
$E_{calc.} = -46.5$
$E_{exp.(HCCl_3)} = -12.7$



C–G
P + 2 S –2 S
$E_{calc.} = -93.78$
$E_{exp.(HCCl_3)} = -17$ –25



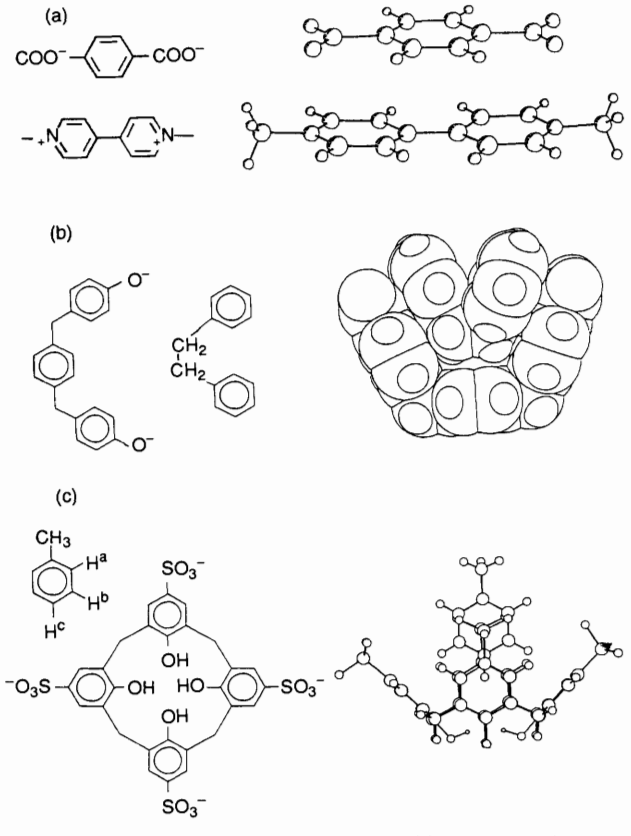
A–T
2 P –2S
$E_{calc.} = -44.4$
$E_{exp.(HCCl_3)} = -8.4$
$E_{exp.(CCl_4)} = -14.3$

Scheme 3 Hydrogen bonds in nucleobase analogues. Partial charges designated by +/− following Jorgensen *et al* (Ref. 29). Primary interactions (hydrogen bonds), P; secondary interactions, S; E_{calc}, ΔG_t calculated[29]; E_{exp}, ΔG_t in $CDCl_3$ (all in [kJ/mol], for data see Refs. 29, 31).

Scheme 3, the negative partial charges at the acceptor atoms O and N lead to additional attraction with a positively charged neighbouring proton, but to repulsion with neighbouring O and N atoms The combination **G–C** has two such attractive secondary minus/plus combinations, and two repulsive secondary combinations (assigned S in Scheme 3), which therefore cancel each other However, **U–DAP** has four repulsive plus/plus or minus/minus interactions, and **A–T** again two repulsive S increments, which explains the weak binding As Jorgensen *et al* [29] have pointed out, their calculational results could be represented by an additive scheme with 30 kJ/mol for each primary hydrogen bond, and 11 kJ/mol for each either repulsive or attractive secondary electrostatic interaction Such additional interactions between adjacent groups are reminiscent of related problems with the Hammett equation, which for quite similar reasons cannot usually be applied to arenes with groups in the *ortho*-position (See note added in proof, p 234)

7 Van der Waals Interactions/Hydrophobic Effects, Aromatic Ion Pairs, Associations with Porphyrins

Lipophilic and hydrophobic effects are known to play a major role in biological systems [15b] Earlier investigations of the aggregation behaviour of lipophilic model compounds[32] has been hampered by the ill-defined stoichiometries and conformations from stacked aromatic compounds as well as by measurement problems Application of the additivity principle to supramolecular complexes allows, for the first time, the analysis of such complexes by reliable methods The strategy is to provide host and if necessary also guest compounds with ionic groups (Scheme 4), besides securing water solubility these can produce salt bridges for which the corresponding increments are known from the independent measurements discussed in Section 4 After correction of the observed ΔG_t values with the 5 kJ/mol increments for each salt bridge one indeed observes linear correlations of, for example, the number m of phenyl rings occurring in aromatic ion pairs and ΔG_t (Scheme 4, Figure 6) [16a]



Scheme 4 Some aromatic ion pairs with additional van der Waals interactions (Ref 16)

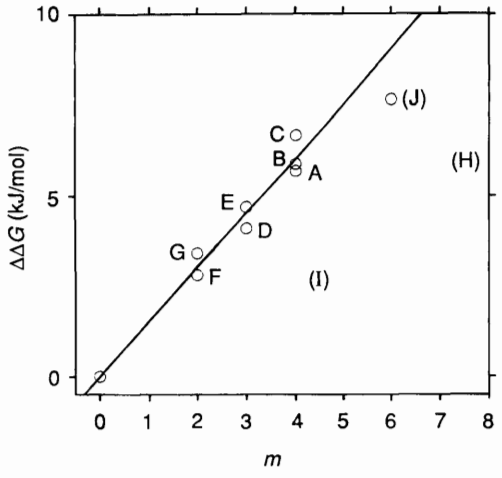


Figure 6 Correlation of non-ionic interactions energies with the number m of phenyl rings in aromatic ion pairs, after correlation for salt bridge contributions [see Ref 16*a*, also for explanation for deviating points (H) and (I)]

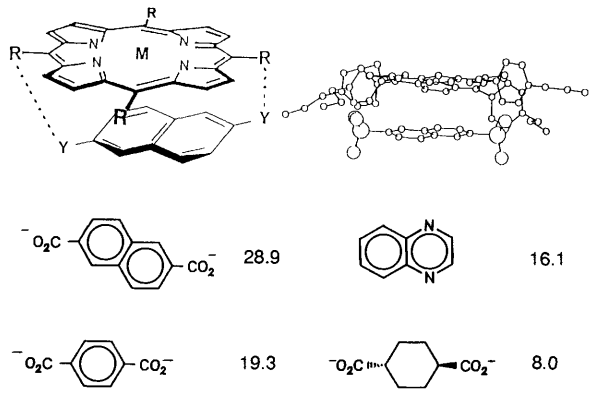
This corresponds to equation 2, with $\Delta G_{vdW} = \Delta G_t - n*5$, where n = number of salt bridges Deviations are only observed for systems in which for steric reasons the ion cannot approach the π-moiety of an opposing ring The increment derived from Figure 6 for this van der Waals interactions is around 2 kJ per mol and per phenyl moiety That this increment reflects largely the dipole induced in the arene by the ion is supported by measurements with corresponding cleft systems (Scheme 4) for which we observe similar values with anion – π instead of cation – π interaction [16b]

A recent study of porphyrin associations allowed the normalization of such lipophilic interactions with respect to the number of π electrons involved, as well as separating them from solvophobic effects [33] The porphyrins used provide a π-surface of constant and large size interacting with a variety of smaller organic ligands, the observed ΔG_t values are remarkably independent of the number or position of nitrogen atoms within the ligands A regular difference of 5 kJ per mol salt bridge is observed again for the difference between electroneutral and charged ligands The ΔG values, after being corrected as before (equation 2) for ionic contributions, if applicable, correlate surprisingly well with the number of π-electrons involved in the associations (Figure 7) Copper ions in the porphyrin have negligible influence, whereas zinc leads to deviations from the correlation due to the known demand of axial coordination

Can one distinguish lipophilic from solvophobic contributions in such systems? That saturated ligands show no binding contribution beyond their salt bridges (Scheme 5) indicates that, even in pure water, hydrophobic effects on ligands of this size are negligible, in rough agreement with published[34] lower-end values of 0 1 kJ per mol and Å^2 In line with this, the observed increase of binding with the porphyrin associations with increasing water content in binary mixtures is much much smaller than observed, for example, for cyclodextrin complexes We have shown earlier[26] that the sensitivities of ligand complexation energies against water content can serve as a measure of hydrophobic binding contribution

8 Crown Ether and Cryptand Complexes – A Case for Multiple Regression Analysis?

Ionophore complexes differ from Figure 1 in that the guest atom has only one binding site, which, however, can use quite different donor functions of a host Consequently, the rich chemistry which has been developed for this oldest and, in terms of application, most important class of host–guest complexes has provided a large range of different donor functions such as R–O–R, R–NH–R, where R = alkyl, phenyl, heterocycle, *etc*



Scheme 5 Porphyrin complex, with typical ΔG_t values ([kJ/mol], water, 25 °C) for aromatic and saturated ligands (Ref. 33).

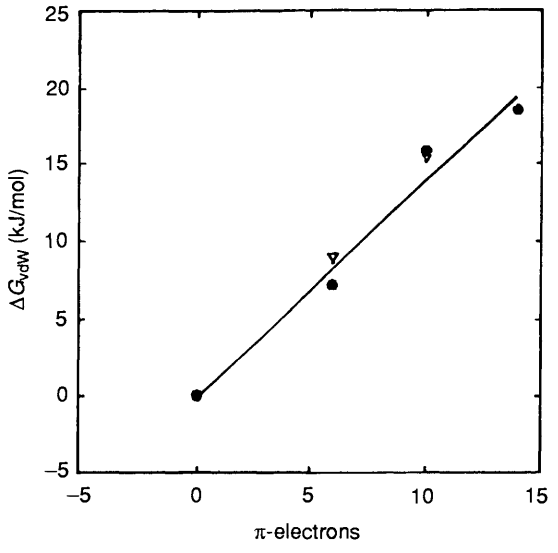


Figure 7 Correlation of non-ionic contributions to the binding with porphyrins with the number n of π electrons in the ligands; the point at $n = 6$ is the average value from 6 complexes, the one at $n = 10$ from 10 complexes (Ref. 33).

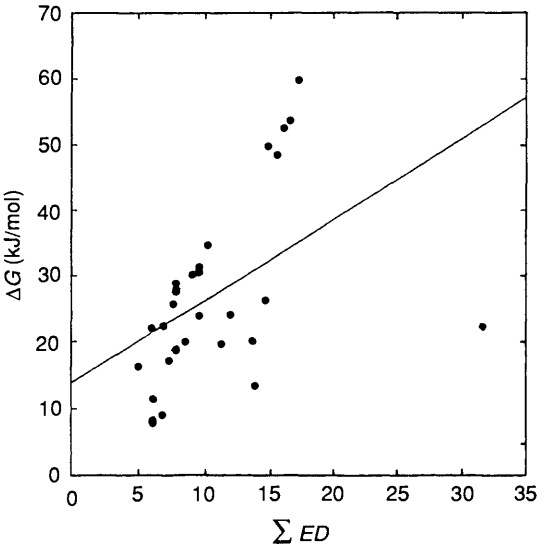


Figure 8 Correlation of ΔG_t for crown ether and cryptand complexes with K^+ (in [kJ/mol], methanol, 25 °C) with the sum of ED increments (Ref. 35); including ionophores which are too large (mismatch).

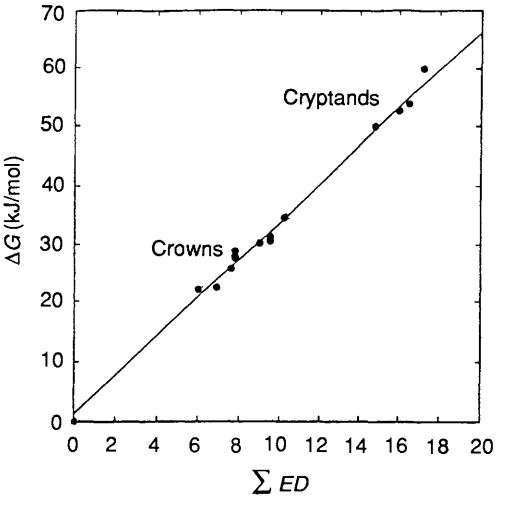


Figure 9 Correlation of ΔG_t for crown ether and cryptand complexes with K^+ (in [kJ/mol], methanol, 25 °C) with the sum of ED increments (Ref. 35). Different 18C6 and [2.2.2] cryptand complexes.

These functions must be characterized by a whole range of separate parameters. In principle one could try to identify such parameters directly from the correlations, as outlined in the previous sections. Here one would need multilinear regression as it is to a lesser degree the number, but more the kind of interactions which determines the observed large variation of complex stability. However, although the complexation energies reported, *e.g.* for the potassium complexes of crown ether and cryptand type, range over 40 kJ/mol, even if only hosts of appropriate size are taken into consideration the topological variation in view of the obvious geometric restriction for ionophores seems rather small. Quite in analogy to classical LFER, where substituent constants are taken from other reactions than the one to be analysed, we therefore decided to use electron donor factors ED^{9a} which were obtained from measurements of simple 1:1 hydrogen bond equilibria in carbon tetrachloride (*cf.* Section 1).

For the analysis,[35] based on equation 3, all available stability constants were used except from those crown ethers or cryptands which are either too small or too large for a given cation, and therefore cannot materialize all possible interactions simultaneously. If the latter are included one observes poor correlations (Figure 8), very much in analogy to LFER in the case of steric distortions. Otherwise, the use of additive increments leads for the first time to a comprehensive and rather accurate description of crown ether and cryptand complexation energy on a *common* scale (Figure 9).

On the basis of only 11 ED group (substituent) parameters which were *not* treated as adjustible but stem from independent hydrogen bond measurements, we could in this way calculate over 120 complex stabilities.[35] This also included complexes with R^+NH_3 cations, without taking into account special factors for hydrogen bonding. This suggests that the binding here is of similar, largely electrostatic nature to that of metal cations. As with related LFER it is necessary to keep the solvent constant, as well as the anion. The sensitivities (slopes) for different cations again seem to be a simple linear function of their hydration energies.[35]

9 Additive Interactions in Protein–Ligand Complexes

The identification of functional group contributions in ligand–protein complexes was started over 20 years ago by Page and Jencks.[13] Through the availability of many data, in particular on equilibrium constants of enzyme inhibitors, the fast progress of computer-aided structure simulation, and owing to the

medicinal and industrial significance of rational drug design, such studies have gained considerable momentum In particular Andrews,[36] Goodford,[37] and others[38] have extended the approach to the analysis of the many interactions acting simultaneously in a protein–ligand complex More recently, Bohm has combined computer-aided automated searches for the identification of possible binding sites in drug receptors with scoring functions for obtaining free energy increments [38] The results of such a multilinear regression[38c] (Figure 10) demonstrate the promise of the additivity principle in this field It should be noted that the regression in Figure 10[38c] is obtained with adjustable increments The scatter demonstrates the problems involved in the extraction of single increments from biopolymer studies, for the reasons discussed in Section 3

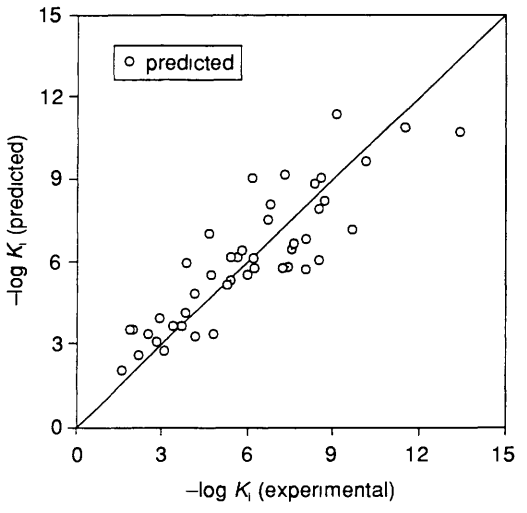


Figure 10 Estimated *vs* experimental binding constants K_i for 45 protein–ligand complexes, after H -J Bohm (Ref 38*c*)

10 Conclusions

Numerical descriptions of chemical systems will show their practical usefulness by systematical comparison between calculated and experimental data, which is the reason why we have confronted the reader with so many linear correlations It is desirable to arrive at descriptions which will predict properties with few, simple, and chemically meaningful rules and parameters, rather than to calculate them afterwards The supramolecular complex is not a black box, but a challenge for physical organic chemistry We hope to have shown that quantitative predictions are already possible within certain limits for complexes comprising many non-covalent bonds This should allow the experimental chemist to design new supramolecular systems, as well as to learn from experiment – in our case especially from stabilities and conformations of host–guest complexes – and thus to improve theoretical models

Why, and for which cases, does one see such simple linear descriptions of supramolecular complexes? First, the reader is reminded that there is no rigorous theoretical foundation for any LFER, the possible compensational factors, for example with respect to entropy/enthalpy, hard/soft behaviour *etc* have been mentioned in Section 4 A significant factor in the obvious tolerance of many supramolecular complexes to smaller steric distortions lies in the dominance of electrostatic forces, which fall off with distance very slowly, another one is the geometric tolerance in solvophobic interactions Future efforts will be directed towards refinement of the additivity approach by broadening the experimental basis as well as by taking into account (i) secondary interactions, such as in nucleobase pairs discussed in Section 6, (ii) unsufficient geometric matching, and (iii) steric distortions with concommitant strain-energy changes

11 References

1 D J Cram, *Angew Chem Int Ed Engl*, 1986, **25**, 1039, D J Cram *Science* 1988, **240**, 760
2 J -M Lehn, *Angew Chem Int Ed Engl*, 1988, **27**, 89, J M Lehn, in 'Frontiers in Supramolecular Organic Chemistry and Photochemistry', ed H -J Schneider and H Durr, VCH, Weinheim 1991, p 1, J -M Lehn, *Angew Chem*, 1990, **102**, 1347, *Angew Chem Int Ed Engl*, 1990, **29**, 1304
3 J F Stoddart, *Annu Rep Progr Chem Sect B*, 1988, **85**, 353, 'Host–Guest Molecular Interactions From Chemistry to Biology', Wiley, Chichester, 1991
4 For a recent review highlighting the role of LFER/QSAR for the understanding of chemistry life interactions with physical organic chemistry "on the verge of a golden age" see C Hansch, *Acc Chem Res*, 1993, **26**, 147
5 See also H -J Schneider, *Angew Chem Int Ed Engl*, 1991, **30**, 1417
6 For references on the computational work of Jorgensen, Kollman, Wipff, and others see (*a*) 'Computational Approaches in Supramolecular Chemistry', ed G Wipff, Kluwer Academic Publishers, 1994, (*b*) H -J Schneider, *Rec Trav Chim Pays-Bas*, 1993, **112**, 412, (*c*) Ref 5
7 V Gutmann, 'The Donor Acceptor Approach to Molecular Interactions', Plenum Press, New York, 1978
8 R S Drago, in 'Structure and Bonding', Springer, Heidelberg, 1973, p 73ff
9 See recent reviews (*a*) O A Raevsky, *Russ Chem Rev*, 1990, **59**, 219, (*b*) M H Abraham, *Chem Soc Rev*, 1993, **22**, 73, and references cited therein
10 G E Schulz and R H Schirmer, 'Principles of Protein Structure', Springer, New York, 1979, J N Spencer, R C Garrtett, F J Mayer, J E Merkle, C R Powell, M T Tran, and S K Berger, *Can J Chem*, 1980, **58**, 1372, and references cited therein
11 See (*a*) R J Saykally and J G Loeser, *J Chem Phys*, 1992, **97**, 4727, (*b*) S L Price and A J Stone, *J Phys Chem*, 1987, **86**, 2859, for recent references see G Klebe and F Diederich, *Phil Trans R Soc Lond A*, 1993, **345**, 37
12 See M I Page, *Angew Chem Int Ed Engl*, 1977, **16**, 449, W P Jencks, *Proc Natl Acad Sci USA*, 1981, **78**, 4046
13 M I Page and W P Jencks, *Proc Nat Acad Sci USA*, 1971, **68**, 1678
14 (*a*) M S Searle and D H Williams, *J Am Chem Soc*, 1992, **114**, 10690, (*b*) M S Searle, D H Williams, and U Gerhard, *J Am Chem Soc*, 1992, **114**, 10697
15 (*a*) A R Fersht *et al*, *Nature*, 1985, **314**, 235, see also (*b*) S K Burley and G A Petsko, *Adv Protein Chem*, 1988, **39**, 125
16 (*a*)H -J Schneider, T Schiestel, and P Zimmermann, *J Am Chem Soc*, 1992, **114**, 7698, see also (*b*) H -J Schneider, F Werner, and T Blatter *J Phys Org Chem*, 1993, **6**, 590, and references cited therein
17 K Burger, 'Solvation, Ionic and Complex Formation Reactions', Elsevier, Amsterdam, 1983, U Meyer, *Coord Chem Rev*, 1976, **21**, 159
18 E Kimura, *Top Curr Chem*, 1985, **128**, 131, 141, E Kimura, in 'Crown Ethers and Analogous Compounds', ed M Hiraoka, 'Studies in Organic Chemistry' Vol 45, Tokyo, 1992, p 381, M W Hosseini, J -M Lehn, and M P Mertes, *Helv Chim Acta*, 1983, **66**, 2454, M P Mertes, K Bowman, and K B Mertes, *Acc Chem Res*, 1990, **23**, 413
19 H -J Schneider, D Guttes, and U Schneider, *J Am Chem Soc*, 1988, **110**, 6449
20 H -J Schneider and I Theis, *Angew Chem Int Ed Engl*, 1989, **28**, 753
21 A Marcus, 'Ion Solvation', Wiley, Chichester, 1986, Y Marcus, *Chem Rev*, 1988, **88**, 1475, J A Gordon, in 'The Organic Chemistry of Electrolyte Solutions', Wiley, New York, 1975
22 H -J Schneider and T Blatter, *Angew Chem Int Ed Engl*, 1992, **31**, 1207
23 K D Stewart and T A Gray, *J Phys Org Chem*, 1992, **5**, 461
24 H -J Schneider, T Blatter, B Palm, U Pfingstag, V Rudiger, and I Theis, *J Am Chem Soc*, 1992, **114**, 7704, and references cited therein
25 See *e g* L Stekowski, D B Harden, R L Wydra, K D Stewart, and W D Wilson, *J Mol Recog*, 1989, **2**, 158
26 H -J Schneider, R Kramer, S Simova, and U Schneider, *J Am Chem Soc*, 1988, **110**, 6442
27 H -J Schneider and I Theis, *J Org Chem*, 1992, **57**, 3066
28 H -J Schneider, R K Juneja, and S Simova, *Chem Ber*, 1989, **112**, 1211

29 For arguments, data, and references see: J. Pranata, S. G. Wierschke, and W. L. Jorgensen, *J. Am. Chem. Soc.*, 1991, **113**, 2810.
30 See also T. J. Murray and S. C. Zimmermann, *J. Am. Chem. Soc.*, 1992, **114**, 4010, and references cited therein.
31 H.-J. Schneider, H.-D. Junker, and J. Sartorius, unpublished results.
32 See K. A. Connors, A. Paulson, and D. Toledo-Velasquez, *J. Org. Chem.*, 1988, **53**, 2033, and earlier references.
33 (*a*) H.-J. Schneider and M. Wang, *J. Chem. Soc., Chem. Commun.*, 1994, 413; (*b*) unpublished results; see also (*c*) Ref. 16*b*.
34 F. M. Richards, *Annu. Rev. Biophys. Bioeng.*, 1977, **6**, 151; a 100% higher value has also been estimated: K. A. Sharp, A. Nicholls, R. Friedman, and B. Honig, *Biochemistry*, 1991, **30**, 9686.

Note Added in Proof

A recent analysis shows that association free energies (in chloroform) of about 40 nucleobase analogues can be described by single increments of 7.9 kJ/mol for primary hydrogen bond interactions, and of 2.9 kJ/mol for secondary interactions (H.-J. Schneider, J. Sartorius).

35 H.-J. Schneider, V. Rüdiger, and O. A. Raevsky *J. Org. Chem.*, 1993, **58**, 3648.
36 P. R. Andrews, D. J. Craik, and J. L. Martin, *J. Med. Chem.*, 1984, **27**, 1648.
37 P. J. Goodford, *J. Med. Chem.*, 1985, **28**, 849; D. N. A. Boobyer, P. J. Goodford, P. M. McWhinnie, and R. C. Wade, *J. Med. Chem.*, 1989, **32**, 1083. S. H. Rotstein and M. A. Murcko, *J. Med. Chem.*, 1993, **36**, 1700; E. C. Meng, B. K. Shoichet, and I. D. Kuntz, *J. Comp. Chem.*, 1992, **13**, 505. For related factorizations see *e.g.* B. Tidar and M. Karplus *Biochemistry*, 1991, **30**, 3217 and references cited therein.
38 (*a*) H.-J. Böhm, *J. Comp. Chem.*, 1992, **6**, 61; (*b*) 1992, **6**, 593; (*c*) 1994, in press (I thank Dr. Böhm for a preprint).