Journal of Computer-Aided Molecular Design, **13:** 99–108, 1999.
KLUWER/ESCOM

Perspective

Scoring functions: A view from the bench

Jeremy R.H. Tame
Department of Chemistry, University of York, Heslington, York YO1 5DD, U.K.

Received 10 February 1998; Accepted 4 September 1998

Key words: binding, ligand, protein, thermodynamics

Abstract

Computational approaches to drug design are presently hindered by the complexity of the physical chemistry which underlies weak, non-covalent interactions between protein targets and small molecule ligands. Although a number of programs are now available for the design of novel potential ligands, it remains a key problem to rank these rapidly and reliably by estimated binding affinity. Such a step is necessary to select only the most promising candidates for synthesis and experimental characterisation. To calculate ligand affinity quickly and reliably is an extremely difficult problem, but it may well prove possible to estimate sufficiently accurately given an appropriate set of parameters to 'score' individual protein–ligand interactions. Improvements in the situation will require a wider set of thermodynamically characterised systems than is currently available.

Introduction

The past decade has seen considerable advances in computer-aided ligand design and computational studies are now making a large impact on the drug design process and helping to direct the search for new pharmaceuticals [1,2]. A variety of programs is now available to find new drug leads either by database searching or from fragment fitting. These methods can rapidly generate a large array of potential ligands, far too many for the chemist to synthesize or study by computer simulations requiring days of CPU time. A computational drug design programme therefore tends to arrive at a bottleneck at which it is necessary to select a limited number of molecules for further analysis. This problem has generated considerable interest in developing methods to calculate ligand affinity reliably for a widely diverse group of molecules binding to some target protein of known structure. Despite the large body of research currently directed towards improving our understanding of the relationship between the structure of macromolecule–ligand complexes and the affinity of interaction, it remains difficult to calculate ab initio the affinity from the structure. The field of affinity estimation is too large to be covered in this paper and current approaches to the problem have already been summarised elsewhere [2–8]. The most rigorous methods such as the free energy perturbation method (FEP) require molecular dynamics simulations and are therefore too time-consuming for initial screening, not merely due to the calculations but also the need for the intervention of a skilled chemist. In calculating the affinity of streptavidin for biotin, Miyamoto and Kollman ran five simulations, the first runs to determine which restraints to place on the system [9]. Even so, the ΔG of binding is overestimated because of the approximations in the treatment of conformational flexibility, the incomplete hydration and the restraints on hydrogen bonding. Despite the considerable computational power brought to bear by these methods the results of such simulations are not always consistent. For example, Cieplak and Kollman suggested that certain inhibitors of HIV-1 protease could be improved by replacing the main chain amides with hydrophobic groups [10], but Gustchina et al. found that the hydrogen bonds formed by these atoms contributed over half of the binding energy [11]. Although the FEP method has worked well in a number of cases comparing related ligands, it is unlikely to become useful in lead selection from

thousands of computer-generated test ligands. An automatic procedure is needed which requires neither human intervention nor dynamical simulation. In other words the problem must be reduced to one of Euclidean geometry, counting atomic contacts of different types between ligand and protein and multiplying by a suitable weight to give a summed free energy. Such weights are generally known as 'scoring functions'.

Scoring functions

Rather than calculate ligand affinity from first principles, a scoring function instead estimates the tightness of binding from structural parameters of the complex. One of the first attempts to make such an empirical connection between structure and energy was by Eisenberg and Maclachlan, who devised a simple additive scheme to estimate the solvation energy of different molecules from the exposed surface area of different atom types [12]. The most widely used scoring function for affinity estimation is LUDI, written by Böhm [2,13], which has inspired several other programs. The philosophy behind the program is relatively simple, the overall free energy of binding being broken down into contributions from hydrogen bonds, ionic interactions, apolar contacts and entropy penalties for fixing rotatable bonds and making one molecule from two. Such methods largely or completely ignore changes in the protein on ligand binding since only the structure of the complex is considered. They are therefore much more likely to work well with relatively rigid proteins. Obviously the determination of appropriate weights requires a database of reliable structures of the protein–ligand complexes and the affinity of each. Böhm used a dataset of 45 different protein–ligand complexes to determine appropriate weights for LUDI. More than half of these are proteases, several of which are represented four or more times, so there is clearly scope to increase the number and variety of proteins in the dataset presently available to modellers. Five rather different weighting schemes in LUDI were found to reproduce the basis set of affinity measurements with notable accuracy. This insensitivity to the scoring weights highlights the need for more thermodynamic data on a wide range of protein–ligand systems.

Three points are worth making about scoring functions. Firstly, they imply that each occurrence of a basic interaction being considered is equivalent. For example, all hydrogen bonds between neutral donor and acceptor atoms may score 4 kJ mol^{-1} or whatever, possibly with some correction made for the geometry of the groups involved. This may hold true on average over related structures but there are many instances where it does not. Cooperativity between different individual interactions can make it rather more difficult to estimate the overall free energy of binding. For example, the two triply hydrogen bonded systems shown in Figure 1 have very different affinities. Subtle electrostatic interactions make it much more favourable in this case to arrange all the hydrogen bond acceptors on one molecule and the donors on the other. DNA is another example of a multiply hydrogen bonded system. It has long been known to molecular biologists that they can calculate the melting temperature (T_M) in °C of short oligonucleotides by the Wallace rule: multiply the number of GC base pairs by 4 and add the number of AT pairs times 2 [14].

$$T_M(^oC) \;=\; 4 * (GC) + 2 * (AT)$$

This works very well for oligonucleotides between 8-mers and 18-mers (which form roughly one to two turns of double helical DNA with their complementary strand) but not for much longer pieces of DNA. The Wallace rule would predict a human chromosome to have a T_M of roughly 10^8–10^9 °C, so clearly some correction is required to extend it to very long DNA. Modellers should also note the results of Clackson and Wells [15] who showed that there is a marked 'hot spot' of binding energy on growth hormone receptor. Their results show that '*structural parameters such as buried surface area did not correlate well with the energetic importance of individual residues*'. This is also an example where the sum of the contributions of individual residues greatly exceeds the overall free energy of binding. This phenomenon of non-additivity was noted earlier by Jencks [16]:

'*It is not unusual to find that the binding of individual molecules A and B is weak or negligible, but AB binds well, so that the whole seems to be greater than the sum of the parts. It is frequently assumed that the observed Gibbs binding energies of two molecules A and B are additive in the molecule AB so that*

$$\Delta G_{AB} = \Delta G_A + \Delta G_B \qquad (1)$$

There is no basis for this assumption. The loss of entropy on combining A and B by a covalent bond can be as much as $-40\ cal\,mol^{-1}\,K^{-1}$ *and some unpredictable fraction of this difference will appear in the binding of AB compared to A and B.*'

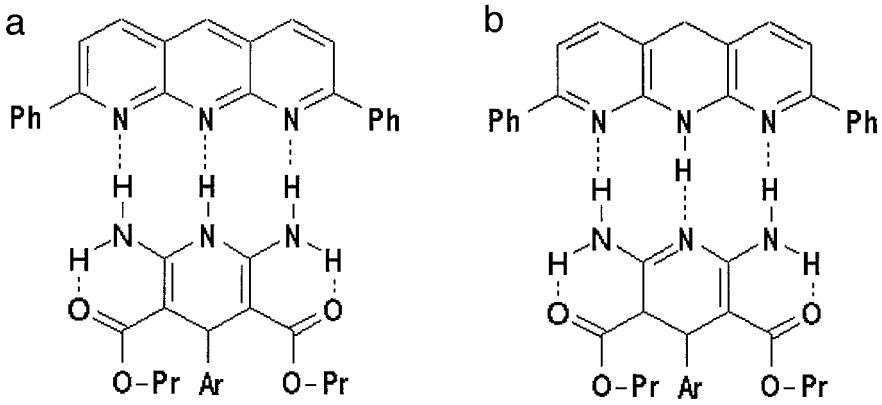

Figure 1. Multiply hydrogen bonded systems designed to test the contributions of hydrogen bonds to binding affinity. Despite the very similar chemistry of the two systems shown in (a) and (b), it is found that the affinity is approximately three orders of magnitude higher for (a) where one molecule carries all the hydrogen bond donors and the other all the acceptors (DDD:AAA). Unfavourable electrostatic interactions occur between the donor and acceptor atoms in (b) as these are arranged alternately (DAD:ADA), greatly lowering the affinity. This effect is probably rather greater in these model systems than in proteins as the experiments are conducted in organic solvent rather than water. Data taken from [62].

A clear hurdle in the construction of scoring functions is to justify the additivity on which they are based, an assumption which often passes without comment. The increase in binding energy of AB over A is dependent on how well the two parts of AB fit into their respective binding sites. If there is no strain in the link between A and B, and each part adopts its most favoured position, then this increase is equal to the intrinsic binding energy of B.

Secondly, in constructing a scoring function a modeller aims to reproduce the experimentally determined affinity or free energy of binding (ΔG) while ignoring the experimental conditions. In fact, the great majority of protein–ligand binding processes will involve a significant change in heat capacity (ΔC_p) and so both ΔH and ΔS will vary with temperature, not to mention pH and ionic strength. Fortunately due to enthalpy-entropy compensation, which is often highly pronounced for weak, non-covalent interactions such as those between a protein and its ligand, ΔG remains relatively unaffected as ΔH and ΔS alter. In building a function to score ($\Delta H - T\Delta S$) the modeller implicitly relies on this phenomenon. It is perfectly possible however for the energy of ligand binding to change markedly over a range of conditions in which the structure appears invariant. This is true of haemoglobin, whose oxygen affinity is strongly dependent on the presence of chloride ions or polyethylene glycol [17,18], although these molecules are not found in electron density maps. It should be noted that the overall observed enthalpy or entropy of binding cannot in general be used to judge whether association is driven by electrostatic or hydrophobic interactions. This is not only because of enthalpy-entropy compensation but also because such numbers include all processes on ligand binding, such as protonation or deprotonation of the buffer. Many ligand binding processes are sensitive to pH (hence the use of buffer in the reaction mixture), implying that protons are captured or released on ligand binding. The affinity may well be unchanged on replacing Tris buffer with phosphate (which has a much lower heat of protonation) but the observed ΔH may change considerably. This effect can be used to measure proton uptake, but is of little use to the computer modeller. Miyamoto and Kollman [9] speak of a conflict between data suggesting biotin binds streptavidin through hydrophobic interactions and the observed negative entropy of binding. There is no conflict in such data. Measurements of enthalpy or entropy can be compared usefully only between similar ligands binding to a given protein under the same experimental conditions.

Thirdly, the use of rigid protein and ligand must be accounted for. Fortunately it appears that freezing rotors seems to entail the same entropy penalty for each rotatable bond. Page and Jencks calculated an entropy change per internal rotation of 4–5 entropy units (cal $mol^{-1}K^{-1}$) for hydrocarbon cyclisation, equivalent to a free energy cost of 5–6 kJ mol^{-1} at room temperature [19]. Böhm employs a significantly smaller number in LUDI, 1.4 kJ mol^{-1}. The cost of fixing protein side chains in single conformations on protein

folding also seems to be reasonably constant for each residue type [20]. There is some disagreement in the literature about the best method of accounting for the loss of six degrees of freedom by the small ligand on binding to the protein (the same rotational and translational freedom familiar to the crystallographer who uses molecular replacement). Page and Jencks estimate on the basis of Trouton's rule that making one molecule from two will involve a considerable entropy cost of 50–70 kJ mol^{-1}, but this loss will be compensated to some extent by low frequency motions in the product [19]. Janin has used the Sackur–Tetrode equation to derive a similar figure, about 63 kJ mol^{-1}, with a small dependence on ligand size [3]. The use of this equation (developed for simple gases) has been disputed but given the rough agreement between this method and the estimate made by Page and Jencks it is probably better than ignoring the effect altogether. Erickson estimates the cost to be 7–11 kcal mol^{-1} [21]. This smaller figure may represent the entropy gain in the complex from vibrational modes absent in the free protein [22]. The entropy penalty used by LUDI is much smaller, 5.4 kJ mol^{-1}. If the term is considered to vary only very slightly with ligand size then clearly it plays little part in ranking different ligands, but in order to estimate the actual binding affinity this term must be considered. LUDI greatly overestimates the ligand affinity of the peptide binding protein OppA, probably largely as a result of the underestimate of the entropic factors involved [23; J. Bray and J.R.H.T., unpublished results].

To overcome the problem of deriving weights from experimental measurements, several groups have assigned free energy scores to interactions by estimating atomic contact energies from a statistical analysis of atom pairing frequencies, a technique first used to analyse protein folds [24–27]. These functions are generally known as 'knowledge-based potentials (of mean force)'. The use of these potentials has been critically discussed by BenNaim [28]. Since the base set of structures used by Delisi and co-workers to determine the contact energies are protein structures, it is an inherent assumption of their approach that the same energy function can be used for protein folding and ligand binding [26,27]. They neglect the contribution of van der Waals forces to binding since they assume it is the same for protein–ligand interactions as for protein–water interactions. The final calculated energy of binding or folding includes the contact energy, an electrostatics term and an entropy term. Computationally the Delisi method suffers from the disadvantage of requiring coordinates for hydrogen atoms, which are placed in the structures using energy minimisation, but otherwise it is undemanding in terms of CPU time. This method permitted the calculation of the affinity of nine endopeptidease–inhibitor complexes with reasonable accuracy, within 10%. Comparing this method with LUDI there are some notable differences in the weighting schemes; for example, the contact energy for a nitrogen-carbonyl bond is favourable for a backbone nitrogen but unfavourable for a side-chain nitrogen. LUDI scores such hydrogen bonds equally with a value of -4.7 kJ mol^{-1} [2,13]. This difference reflects the entropy cost of holding a side chain in place on protein folding, included implicitly in the contact energy derived by the Delisi group, but it is not clear that this is appropriate when considering a ligand fitting into a preformed binding site [29]. Verkhiver and colleagues have derived a knowledge-based potential from the crystal structures of HIV-1 protease–inhibitor complexes [25]. They considered only contacts between non-hydrogen atoms and calculated the desolvation energy using the empirical scale derived by Eisenberg and Maclachlan [12]. This model provided good correlation between the observed and calculated binding affinities.

Protein–protein association

Proteins associate by forming the same types of interactions that they form with small ligands. Horton and Lewis [30] derived an energy function from an analysis of 24 protein complexes and their dissociation constants, basing their function on the solvation parameters derived earlier by Eisenberg and Maclachlan [12]. The total free energy of binding is broken down into three terms, the solvation of polar and apolar groups, and the rotational-translational entropy cost. Least-squares fitting was used to derive the weights for the two solvation parameters and the entropy penalty of association. The energy function reproduced the basis dataset with a correlation of 96%, and predicted with acceptable accuracy the affinity of trypsin and subtilisin for inhibitors. The energy contribution of apolar surfaces was found to be close to the 24 cal mol^{-1} per $Å^2$ estimated much earlier by Chothia [31] and the rotational-translational entropy cost was found to be close to Erickson's estimate [21]. Hydrogen bonds were found to be much weaker than other estimates, averaging -0.24 kcal mol^{-1}, a third to a sixth of the value derived from side-directed mutational analysis by the group of Fersht [32]. An analysis of the dimer–

tetramer equilibrium of haemoglobin mutants by Vallone et al. also showed that the energy of association is well predicted by the Eisenberg and Maclachlan model, but they derived a different value for the free energy cost of solvating 1 $Å^2$ of apolar surface area, about 15 cal mol^{-1} [33].

Problems to be overcome in affinity estimation

While it appears relatively simple to derive scoring functions which reproduce the basis dataset from which they are derived, it has proved rather more difficult to find a predictive algorithm which handles novel ligands reliably. This is not due to the errors in atomic coordinates in well-refined X-ray crystallography and nuclear magnetic resonance (NMR) models. Certainly these errors are considerable on an energy landscape – even in a protein crystal structure solved to moderately high resolution the coordinate errors of the most accurately positioned non-hydrogen atoms may be 0.2 Å, so the distance between hydrogen bonded N and O atoms will have an even larger error, quite significant in terms of the stability of the hydrogen bond. Better structural data are unlikely however to make much difference in many cases – it is just as hard to estimate the affinity of an interaction whether the complex has been refined to 1.0 Å or 2.5 Å resolution. Although the errors in atomic positions may seem large, they are in fact smaller than the thermal vibration implied by the temperature factors in X-ray crystallographic structures. An atom with a temperature factor (B) of 20 $Å^2$ would be considered reasonably well ordered, but since

$$B = 8\pi^2\bar{u}^2 \quad (2)$$

where u is the mean deviation from the average position, its rms deviation is still around 0.5 Å. Thermal vibration will be no smaller in solution (the state of interest) than in a crystal. Since we are interested in the free energy change of ligand binding, which is a bulk property, it is appropriate to use the average positions of atoms anyway, which is what an X-ray or NMR structure represents. To rank computer-generated ligands correctly the affinity estimation program must be able to position novel ligands within the binding site and optimise their interactions with the target protein. The errors in this calculated position are unlikely to be smaller than those of the coordinates of the protein atoms, so the estimation of affinity must be able to accommodate small deviations in atom positions. Rather than the limited precision of structural methods, the problem of affinity estimation lies in our limited understanding of the physics and thermodynamics of ligand binding by biomolecules. The physical chemistry underlying non-covalent molecular association in water is highly complex, and it is clear that our understanding of the forces which drive it must improve if affinity estimation methods are to give quantitative agreement with experimental results.

The forces involved in ligand binding

The significant forces acting on molecules in aqueous solution are well known and generally broken down into four groups, hydrophobicity, van der Waals forces, hydrogen bonds and electrostatic interactions. The definitions of these forces are to some extent arbitrary. (Of the four forces known to physics only electromagnetism is relevant here.) Hydrogen bonding is electrostatic in origin, but usually considered in isolation because of its strongly directional nature and the difficulty in accurately assigning partial charges to the atoms involved. Considerable research has been devoted to this interaction, which is generally (but not universally) considered to favour ligand binding or protein folding, more so if one or both of the acceptor and donor atoms is/are charged. Interactions between charges or dipoles lead to an electrostatic energy which is given by the familiar Coulomb equation. Electrostatics has been an area of great interest in recent years, much attention focussing on the linearized Poisson–Boltzmann equation [34]. The calculated strength of charge-charge interactions is however strongly dependent on the chosen value or function used for the dielectric constant ϵ and a variety of distance-dependent expressions or fixed values are employed by different authors. Electrostatic effects in macromolecules have recently been reviewed by Warshel and Papazyan [35]. Van der Waals forces were originally invoked to explain the deviation of real gases from ideal behaviour. The term is now used to include induction energy (which arises from polarisation of a molecule in an applied electric field) and dispersion forces, which are quantum mechanical in nature but arise largely from induced-dipole:induced-dipole interactions. The energy of both types of interaction falls off approximately with the sixth power of the molecular separation. Hydrophobicity, the anomalously low solubility of hydrocarbons in water, may be defined in a number of almost equivalent ways which has led to considerable debate as to the cause of the effect. It may be defined in terms

of solubility or free energy of transfer from gas or liquid phase to water (ΔG). Traditionally it has been regarded as arising from the entropic penalty in solvating apolar surfaces, but 10 years ago Privalov and Gill suggested instead that '*the hydrophobic interaction is caused by the van der Waals interactions between the non-polar molecules*' [36]. This 'new view' of hydrophobicity is fraught with difficulties and generally dismissed. Computer modelling studies support the model put forward by Muller [37] in which water molecules in the bulk solvent form more but weaker hydrogen bonds than those in contact with apolar surfaces [38]. Thus, although the types of non-covalent interactions which occur between proteins and ligands are known, there has been a great deal of confusion in the literature about their relative importance and even the sign of their contribution to binding. This is true for each of the four forces listed above except van der Waals interactions between protein and ligand, which are generally considered favourable to binding but are ignored by some authors on the grounds that they are equally favourable with solvent. The divergence of views on some of these issues is shown by quotes taken from relatively recent publications:

'*If the anecdotal evidence so far is verified, the rather global implication is that electrostatic interactions tend to favour the unfolded states of macromolecules and the unbound states of complexes. This would imply that the thermodynamic driving force for most processes in aqueous solution results from 'non-polar' interactions such as the hydrophobic effect and close packing'* [34].

'*Just the opposite: the hydration of non-polar groups increases the solvation tendency of these groups in water and destabilises the compact protein structure which is in fact stabilised by van der Waals [forces] and hydrogen bonding of the tightly packed amino acid residues'* [36].

Obviously affinity estimation will be difficult if there is disagreement over the sign of the free energy contributions of different interactions.

The contribution of hydrogen bonds to ligand binding has been studied by a number of groups. Fersht found that the loss of a typical hydrogen bond cost about -0.5 to -1.8 kcal mol^{-1} [32]. A much larger value of -24 kJ mol^{-1} was determined by Williams' group in their analysis of peptides binding to antibiotics [39], but this was later found to be spuriously high due to a hydrophobic interaction which had been overlooked [40]. Nevertheless, using the value of -24 kJ mol^{-1} per hydrogen bond, Williams and co-workers managed to account successfully for the binding affinities of two peptides to two antibiotics. This highlights the important point that self-consistency alone is not an adequate test of a set of free energy parameters. It also emphasises the difficulties in deriving such parameters from experimental measurements. A large number of papers have been published on the energy contribution of the hydrophobic effect, current estimates ranging from 4 to 32.5 cal mol^{-1}Å^{-2} [41]. Clearly we have yet to reach a consensus on the relative contributions of different interactions to ligand binding. A key problem is to collect sufficient suitable experimental data to assess the contributions of different interactions to the overall observed free energy change of binding. Unlike X-ray crystallography and NMR, the methods for studying the function of proteins have undergone relatively little development in recent years and several techniques used today have remained unchanged for decades. Thus, while the number of macromolecular structures held in the Brookhaven DataBank continues to grow exponentially and currently exceeds 6000, accurate affinity measurements have been made for a far smaller number.

Experimental determination of binding constants

How are affinity measurements made? Traditionally binding constants have been measured by techniques such as filter binding, equilibrium dialysis, and spectroscopic methods in which the ligand is titrated against the macromolecule. The concentration of ligand giving half-maximal occupancy of binding sites is determined graphically, for example using Scatchard plots. Filter binding assays and equilibrium dialysis require a radioactively labelled ligand and depend on the measurement of radioactivity counts. In recent years there has been a strong movement away from the use of radioisotopes, partly due to safety considerations and partly due to the difficulties in obtaining suitably labelled ligands. Spectroscopic methods such as the use of fluorescence to monitor ligand binding have been very successful in characterising biomolecular interactions. It is often possible to use the intrinsic fluorescence of tryptophan and tyrosine residues of proteins to monitor binding or conformational changes but this is not invariably true, and a chromophore or fluorophore may have to be attached to one of the partners in the interaction being studied in order to provide a measurable signal. The sensitivity of fluorescence measurements can allow K_d's as low as 10^{-11} M to

be measured in favourable cases. These methods have the advantage of using relatively little material, but since the dissociation constant is measured accurately only at roughly equivalent concentrations ([P]$\approx K_d$) the measured signal is often relatively small, introducing significant errors. Weber estimates that few measurements of K_d made by such techniques are accurate to better than 50% [42]. Analytical ultracentrifugation (AU) is a well-established method but fell into increasing disuse throughout the 1970s and 1980s. There has been some renewal of interest with the appearance of the Beckman XL-A and XL-I machines which are more user-friendly than their predecessors [43]. Equilibrium sedimentation experiments can yield affinity measurements in the range 10^{-3} M to 10^{-8} M. Importantly, AU requires no modification of the macromolecule or ligand under consideration. Two new techniques for affinity measurement have appeared in the last few years, microcalorimetry and surface plasmon resonance (SPR). Although calorimetry itself is an old method, the appearance of highly sensitive machines in the past 10 years has transformed its application to biomacromolecules. The use of isothermal titration calorimetry (ITC) to measure protein–ligand interactions has been reviewed recently [44]. The heat generated (or absorbed) as ligand is injected step-wise into a sample of the macromolecule yields the binding constant and molar enthalpy change of the interaction directly. In practice, ITC can measure dissociation constants in the range 10^{-3} to 10^{-9} M. Up to about 100-fold tighter binding can be measured by displacement techniques. Very low K_d's (high binding constants) can be measured using differential scanning calorimetry (DSC), for example K_d for soy bean trypsin inhibitor binding to trypsin was found to be 5×10^{-14} M by this method, and even higher K_d's should be obtainable [45].

ITC has the advantage over other techniques of affinity measurement that it measures the enthalpy change directly. Plotting the heat pulse for each injection against injection number gives a sigmoidal plot whose shape is determined by K_d. Knowing the concentration of the macromolecule, the ligand and the heat of dilution (which is subtracted from all injection heats) then K_d and ΔH may be obtained by a simple curve fitting algorithm.

Using the relationship

$$RTlnK_d = \Delta G^0 = \Delta H - T\Delta S \tag{3}$$

it is possible to calculate ΔS. A single experiment is therefore all that is needed in principle to provide a complete thermodynamic characterisation of a binding process at a given temperature.

If the experiment is performed over a range of temperatures then the change in heat capacity at constant pressure ΔC_p can be calculated. This is often constant within error over a temperature range within which mesophilic proteins are stable. Assuming ΔC_p is temperature-independent then

$$\Delta C_p = \frac{\Delta H_{T2} - \Delta H_{T1}}{T2 - T1} \tag{4}$$

where $T1$ and $T2$ are two different experimental temperatures (in Kelvin).

SPR has rapidly become a widespread method since the introduction of the first commercial instrument in 1990, despite its high cost. The physical principle of the method is discussed by Garland [46]. The macromolecule of interest is attached to a gold foil and the ligand is passed over it, binding being detected by a light beam reflected from the opposite side of the metal foil. Essentially the method measures changes in the refractive index of the solution close to the foil, which increases as ligand attaches to the immobilised protein. The method is very rapid, allowing fast screening of potential ligands for a protein. Since the binding and dissociation steps are monitored separately, the machine also permits kinetic analysis of the interaction. The accuracy of the kinetic data obtained in this way has been questioned [47] but such experiments may be useful if one wishes to screen for ligands with particular kinetic properties. (It is not always appreciated that the on and off rates may be as physiologically important as the overall ligand affinity. For example, it is believed that the dissociation rate of oxygen from haemoglobin must lie within a fairly narrow window if the protein is to transport the gas efficiently.) A large and growing number of papers have appeared in which affinity data are measured by this technique. Although the method essentially measures a change in density at the metal surface it has successfully been used with ligands as small as 180 Da.

No single method of protein–ligand affinity measurement has universal application, which partly explains why few systems have been measured by more than one method. This would probably be a useful check to show that the numbers produced agree. Certainly the errors in the methods are considerable, a value within 10% being the exception rather than the rule [42].

Heat capacity

The strongest correlation noted so far between the thermodynamics of biomacromolecular interactions and the structures of the molecules is between the surface area buried by complex formation and ΔC_p. The removal of hydrophobic surfaces from contact with water leads to a large negative ΔC_p first noted by Edsall in 1935. Simple models based solely on the polar and apolar surface area buried on ligand binding generally predict ΔC_p fairly well, although there are exceptions in which a large number of water molecules is found at the ligand binding site. Vibrational modes of the molecules under consideration, which are of interest from an entropic point of view, may also have a significant influence on ΔC_p [22].

The importance of ΔC_p lies in the fact that this parameter controls the variation of both ΔH and ΔS with temperature according to the following expressions:

$$\Delta H(T_1) = \Delta H(T_0) + \Delta C_p(T_1 - T_0) \quad (5)$$

$$S(T_1) = S(T_0) + C_p(\ln \frac{T_1}{T_0}) \quad (6)$$

Since ΔH and $T\Delta S$ move in the same direction as T varies their changes largely cancel, leaving ΔG relatively temperature-invariant (over the temperature range of interest). This effect, known as enthalpy-entropy compensation, is observed throughout chemistry but is particularly marked in weak interactions in aqueous solutions. Dunitz has suggested that it is strongest for interactions with roughly the strength of a hydrogen bond [48]. Entropy-enthalpy compensation and the temperature variation of ΔH and $T\Delta S$ make it impossible to consider polar or apolar interactions as purely enthalpic or entropic, respectively. Several authors appear to use these pairs of terms almost interchangeably. Interactions between charges may appear to be governed by the enthalpy term but there are many examples where this is not the case. For instance the ionisation of a carboxyl group is nearly isenthalpic ($\Delta H = 0$) under normal conditions due to the strong solvation of the carboxylate group, demonstrating the importance of water in the thermodynamics of proteins.

Water

Water is a deceptively complex liquid whose behaviour is dominated by its strong intermolecular hydrogen bonding [49]. This greatly complicates the estimation of ligand affinity since the different solvation of the AB complex and the free partners A and B must be taken into account. It is sometimes assumed that the exclusion of water from a binding site by a ligand is invariably favourable, due to the entropic contribution of the release of solvent molecules from the protein surface. This is not the case, however. Water molecules held in the binding site will give rise to a negative (unfavourable) change in entropy. The precise cost will depend on the interactions formed but it has been estimated from experiments on the water of hydration of crystal salts to be 10 to 30 J $mol^{-1}K^{-1}$ per water molecule [50]. The free energy ($T\Delta S$) contribution is therefore comparable to the enthalpy of formation of a hydrogen bond. Water molecules trapped in the binding interface between a protein and its ligand are not necessarily hindrances to ligand binding; the enthalpic contribution they provide through hydrogen bonding may outweigh the entropic cost of restricting their freedom of movement. This can be seen in some antibody–antigen complexes. Antibodies which have been selected on the basis of their affinity for a particular ligand are sometimes found to have very large numbers of waters at the ligand interface [51].

Water also plays an important role in a system being used at York to examine structure-energy relationships, the oligopeptide binding protein from Gram

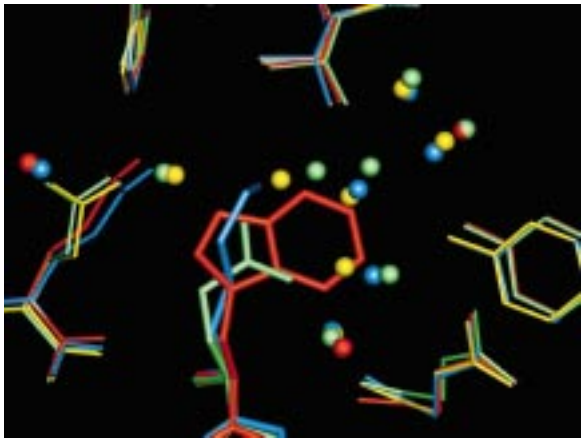

Figure 2. A view of the second side-chain pocket of the oligopeptide binding protein OppA, showing four overlapped structures. These are the complexes of the protein with Lys-Lys-Lys (blue), Lys-Trp-Lys (red), Lys-Glu-Lys (green) and Lys-Ala-Lys (yellow). The water molecules are shown as appropriately coloured spheres. It can be seen that different side-chains on the second residue of the ligand, shown in the centre of the figure, displace different numbers of water molecules from the pocket and form different interactions with the protein [23].

negative bacteria, OppA. This binds to small peptides up to five residues long with little regard to sequence [23,52]. Peptide binding is driven by the protein forming hydrogen bonds with the invariant main-chain atoms of the ligand. The side-chains of the ligand are accommodated in large hydrated pockets, small side-chains displacing fewer water molecules than large ones (Figure 2). The water molecules act as flexible adapters to allow the side-chain pockets to accept a variety of chemical groups. By challenging the protein with different (natural and non-natural) peptides it is intended to build a scoring function for the interactions observed in the crystal structures. For the small number of peptides tested there is no correlation between the size of the ligand side-chains and the affinity, so the expulsion of more water from the binding site does not improve the interaction. In other cases the expulsion of a water molecule from the binding site does contribute favourably to the affinity through an increase in entropy [53]. Water molecules are often found to mediate interactions between proteins and ligands, and are found to be well conserved in homologous structures [54]. Since they can clearly contribute to ligand binding they should not be ignored, and it has even been suggested that to engineer water binding sites may be a useful tool in drug design [55]. Attempts to improve drug binding by expelling water molecules from the binding site with hydrogen bonding groups are not always successful, an interesting example being the HIV protease inhibitors tested by Mikol et al. [56]. In spite of the evidence pointing to the importance of water and solvation in ligand binding by proteins in an aqueous environment it is still often neglected in many studies. The ligand scoring function of Jain [57] includes a solvation term which amounts to only 5% of the overall calculated free energy of binding. This is too small, but perhaps better than ignoring solvation altogether.

Conclusions

Ligand affinity estimation remains a complicated problem, despite the growing number of complex structures available and ever-increasing computer power. Several scoring functions have been described which appear to reproduce well the binding energy of the basis dataset and related protein–ligand systems. Since the reported errors are on a par with the expected experimental error it appears superficially that the problem of rapid affinity estimation is solved. Unfortunately the predictive power of the methods falls short of their statistical performance [4] and different scoring functions show wide variation in the relative weight given to different interactions, so there is still considerable room for improvement in the understanding of the ligand binding process [58]. Personally I prefer attempts to derive the free energy of different effects from established chemical thermodynamics, for example the entropy of phase transitions or cyclisation [16,40,58], to those which use regression alone to find these parameters. In fitting weights to experimental data it is important to make sure that the values arrived at are chemically reasonable, not least because it is possible to derive a number of schemes which predict ΔG to within 10–15% of the experimental value, especially if using a large number of parameters. This is probably the closest a general scoring function can be expected to come to the 'true' value, given the inherent assumptions in the method. At present there is little consensus however on what a hydrogen bond or a square Å of apolar surface area is 'worth'. Direct comparison of different schemes with the same protein–ligand systems is necessary in order to assess their relative merits.

It is unlikely that binding processes involving flexible proteins or large conformational changes will be modelled well from consideration of the complex alone. This is shown by the fact that scoring functions predict very similar binding energy for trypsin and trypsinogen binding to BPTI [60]. Both form very similar contacts with the inhibitor, but trypsinogen binds more weakly by orders of magnitude. It has been suggested that this is due to a conformational change of trypsinogen on ligand binding but may also be partly due to the greater flexibility of the zymogen compared to the mature protease. Where sufficient structural and experimental data have been collected for a particular protein or class of proteins it may prove possible to construct a tailor-made scoring function, as has been done with HIV protease [25,59]. This protein is known to undergo conformational changes on ligand binding and it may well not behave according to a model assuming rigidity. In the case of OppA the protein adopts the same closed conformation for every ligand, each of which presumably pays the same entropic cost for restricting motion about the hinge; if so then scoring functions should still manage to rank ligands correctly. Gilson et al. have briefly described a new class of models for estimating affinity which aim to be faster than molecular dynamics but which take account of conformational change [61]. They point out that scoring functions are highly simplistic, in their

view too much so. Clearly scoring functions should be made as simple as possible, but not simpler. How simple that is remains debatable, and may well depend on whether the scoring function is intended for lead selection or lead optimisation, speed being less important in the latter case. A universal scoring function may well prove an impossible dream but there is a lot to be learned in trying to build one. The differences in the various methods described need to be addressed by detailed studies on new protein–ligand systems. The current set of data available is too small and limited, and the published affinity data is measured over a variety of experimental conditions. An increasing pool of examples of protein–ligand complexes, both of known affinity and accurately characterised by X-ray crystallography, will clearly be an essential element in improving current scoring methods.

References

1. Verlinde, C.L.M.J. and Hol, W.G.J., Structure, 2 (1994) 577.
2. Böhm, H.-J., Prog. Biophys. Mol. Biol., 66 (1996) 197.
3. Janin, J., Proteins, 21 (1995) 30.
4. Knegtel, R.M.A. and Grootenhuis, P.D.J., Perspect. Drug Discov. Des., 9/10/11 (1998) 99.
5. Kollman, P., Chem. Rev., 93 (1993) 2395.
6. Ajay and Murcko, M.A., J. Med. Chem., 38 (1996) 4953.
7. Gilson, M.K., Given, J.A., Bush, B.L. and McCammon, J.A., Biophys. J., 72 (1997) 1047.
8. Hubbard, R.E., Curr. Opin. Struct. Biotechnol., 8 (1997) 696.
9. Miyamoto, S. and Kollman, P., Proteins, 16 (1993) 226.
10. Cieplak, P. and Kollman, P.A., J. Comput.-Aided Mol. Design, 7 (1993) 291.
11. Gustchina, A., Samson, C., Prevost, M., Richelle, J., Wodak, S.Y., Wlodawer, A. and Weber, I.T., Protein Eng., 7 (1994) 309.
12. Eisenberg, D. and Maclachlan, A.D., Nature, 319 (1986) 199.
13. Böhm, H.-J., J. Comput.-Aided Mol. Design, 8 (1994) 243.
14. Wallace, R.B., Johnson, M.J., Hirose, T., Miyake, T., Kawashima, E.H. and Itakura, K., Nucleic Acids Res., 9 (1981) 879.
15. Clackson, T. and Wells, J.A., Science, 267 (1995) 383.
16. Jencks, W.P., Proc. Natl. Acad. Sci. USA, 78 (1981) 4046.
17. Perutz, M.F., Shih, D.T.B. and Williamson, D., J. Mol. Biol., 239 (1994) 555.
18. Colombo, M.F., Rau, D.C. and Parsegian, V.A., Science, 256 (1992) 655.
19. Page, M. and Jencks, W.P., Proc. Natl. Acad. Sci. USA, 68 (1971) 1678.
20. Doig, A.J. and Sternberg, M.J.E., Protein Sci., 4 (1995) 2247.
21. Erickson, H.P., J. Mol. Biol., 206 (1989) 465.
22. Sturtevant, J.M., Proc. Natl. Acad. Sci. USA, 74 (1977) 2236.
23. Tame, J.R.H., Sleigh, S.H., Wilkinson, A.J. and Ladbury, J.E., Nat. Struct. Biol., 3 (1996) 998.
24. Sippl, M.J., Ortner, M., Jaritz, M., Lackner, P. and Flockner, H., Folding Des., 1 (1996) 289.
25. Verkhiver, G., Appelt, K., Freer, S.T. and Villafranca, J.E., Protein Eng., 8 (1995) 677.
26. Wallquist, A., Jernigan, R.L. and Covell, D.G., Protein Sci., 4 (1995) 1881.
27. Zhang, C., Vasmatzis, G., Cornette, J.L. and Delisi, C., J. Mol. Biol., 267 (1997) 707.
28. BenNaim, A., J. Chem. Phys., 107 (1997) 3698.
29. Xu, D., Lin, S.L. and Nussinov, R., J. Mol. Biol., 265 (1997) 68.
30. Horton, N. and Lewis, M., Protein Sci., 1 (1992) 169.
31. Chothia, C., Nature, 248 (1974) 338.
32. Fersht. A.R., Trends Biochem. Sci., 12 (1987) 301.
33. Vallone, B., Miele, A.E., Vecchini, P., Chiancone, E. and Brunori, M., Proc. Natl. Acad. Sci. USA, 95 (1998) 6103.
34. Honig, B. and Nicholls, A., Science, 268 (1995) 1144.
35. Warshel, A. and Papazyan, A., Curr. Opin. Struct. Biol., 8 (1998) 211.
36. Privalov, P.L. and Gill, S.J., Adv. Protein Chem., 40 (1988) 191.
37. Muller, N., Acc. Chem. Res., 23 (1990) 23.
38. Laidig, K.E. and Daggett, V., J. Phys. Chem., 100 (1996) 5616.
39. Williams, D.H., Cox, J.P.L., Doig, A.J., Gardner, M., Gerhard, U., Kaye, P.T., Lal, A.R., Nicholls, I.A., Salter, C.J. and Mitchell, R.G., J. Am. Chem. Soc., 113 (1991) 7020.
40. Searle, M.S., Williams, D.H. and Gerhard, U., J. Am. Chem. Soc., 114 (1992) 10697.
41. Juffer, A.H., Eisenhaber, F., Hubbard, S.J., Walter, D. and Argos, P., Protein Sci., 4 (1995) 2499.
42. Weber, G., Protein Interactions, Chapman and Hall, New York, NY, 1992.
43. Hensley, P., Structure, 4 (1996) 367.
44. Ladbury, J.E. and Chowdhry, B.Z., Chem. Biol., 3 (1996) 791.
45. Brandts, J.F. and Lin, L.N., Biochemistry, 29 (1990) 6927.
46. Garland, P.B., Quart. Rev. Biophys., 29 (1997) 91.
47. Schuck, P. and Minton, A.P., Trends Biochem. Sci., 21 (1996) 458.
48. Dunitz, J.D., Chem. Biol., 2 (1995) 709.
49. Franks, F., Water, Royal Society of Chemistry, London, 1984.
50. Dunitz, J.D., Science, 264 (1994) 670.
51. Bhat, T.N. Bertley, G.A., Boulot, G., Greene, M.I., Tello, D., Dallacqua, W., Souchon, H., Schwarz, F.P., Mariuzza, R.A. and Poljak, R.J., Proc. Natl. Acad. Sci. USA, 91 (1994) 1089.
52. Tame, J.R.H., Murshudov, G.N., Dodson, E.J., Neil, T.K., Dodson, G.G., Higgins, C.F. and Wilkinson, A.J., Science, 264 (1994) 1578.
53. Pearlman, D.A. and Connelly, P.R., J. Mol. Biol., 248 (1995) 696.
54. Poornima, C.S. and Dean, P.M., J. Comput.-Aided Mol. Design, 9 (1995) 513.
55. Ladbury, J.E., Chem. Biol., 3 (1996) 973.
56. Mikol, V., Papageorgiou, C. and Borer, X., J. Med. Chem., 38 (1995) 3361.
57. Jain, A.N., J. Comput.-Aided Mol. Design, 10 (1996) 427.
58. Searle, M.S. and Williams, D.H., J. Am. Chem. Soc., 114 (1992) 10690.
59. Oprea, T.I. and Marshall, G.R., Perspect. Drug Discov. Des., 9/10/11 (1998) 35.
60. Krystek, S., Stouch, T. and Novotny, J., J. Mol. Biol., 234 (1993) 661.
61. Gilson, M.K., Given, J.A. and Head, M.S., Chem. Biol., 4 (1997) 87.
62. Zimmerman, S.C. and Murray, T.J., Phil. Trans. R. Soc. London, A 345 (1993) 49.