Journal of Computer-Aided Molecular Design (2005) 19: 445–451
DOI 10.1007/s10822-005-8483-7



# Scoring functions – the first 100 years

Jeremy R.H. Tame*
*Protein Design Laboratory, Yokohama City University, Suehiro 1-7-29, 230-0045, Tsurumi, Yokohama, Japan*

Received 15 April 2005; accepted 7 June 2005


*Key words:* scoring function, Einstein, thermodynamics, entropy, confusion

**Abstract**

The use of simple linear mathematical models to estimate chemical properties is not a new idea. Albert Einstein used very simple 'gravity-like' forces to explain the capillarity of different liquids in 1900–1901. Today such models are used in more complicated situations, and a great many have been developed to analyse interactions between proteins and their ligands. This is not surprising, since proteins are too complicated to model accurately without lengthy numerical analysis, and simple models often do at least as good a job in predicting binding constants as much more computationally expensive methods. One hundred years after Einstein's 'miraculous year' in which he transformed physics, it is instructive to recall some of his even earlier work. As approximations, 'scoring functions' are excellent, but it is dangerous to read too much into them. A few cautionary tales are presented for the beginner to the field of ligand affinity prediction by linear models.

## Introduction

Like all of Einstein's very early papers, his first recorded publication (in *Annalen der Physik* in 1901) concerned thermodynamics, and began with the Second Law [1]. It concerns the energy required to form surface area on liquids, which Einstein set out to calculate from the chemical formula. He wrote "*I proceeded from the simplest assumptions about the nature of molecular attraction forces and examined their consequences regarding their agreement with experiment. In this I was guided by the analogy with gravitational forces.*" Essentially Einstein assumed that the attraction between two atoms of types A and B was proportional to the product of constants $c_A$ and $c_B$ which were characteristic of the atoms, in exactly the same way that planets and stars attract each other according to their masses. Matter is considered homogenous, and no account is taken of molecules or chemical bonding, an explicitly noted approximation. Nevertheless, Einstein managed to reproduce an experimentally determined quantity (related to the surface tension) for 17 compounds of carbon, hydrogen and oxygen by assigning a single value to each atom type (Table 1). An additional 24 compounds also containing chlorine, bromine or iodine could also be fitted handsomely into this scheme, given a constant for each halogen. Typically, Einstein then rederived his data using a different phenomenon, in this case heating by compression, and showed an impressive agreement with the first results and a different set of experimental data. He concluded "*In summary, we may state that our basic assumption has stood the test: To each atom corresponds a molecular attraction field that is independent of the temperature and of the way in which the atom is chemically bound to other atoms.*" All this is quite impressive for a 22-year-old, especially one writing 5 years before Boltzmann committed suicide at his inability to persuade fellow chemists that atoms exist. Einstein used the same model in analysing

*To whom correspondence should be addressed, Fax: +81-(0)-45-508-7366; E-mail: jtame@tsurumi.yokohama-cu.ac.jp

*Table 1.* Experimental and calculated values of $\Sigma c_\alpha$ given by Einstein for various compounds.

| Compound | Formula | Experimental | Calculated |
|---|---|---|---|
| Limonene | $C_{10}H_{16}$ | 510 | 524 |
| Formic acid | $CO_2H_2$ | 140 | 145 |
| Acetic acid | $CO_2H_4$ | 193 | 197 |
| Propanoic acid | $C_3H_6O_2$ | 250 | 249 |
| Butyric acid | $C_3H_6O_2$ | 309 | 301 |
| Valerianic acid | $C_5H_{10}O_2$ | 365 | 352 |
| Acetic anhydride | $C_4H_6O_3$ | 350 | 350 |
| Ethyl oxalate | $C_6H_{10}O_4$ | 505 | 501 |
| Methyl benzoate | $C_8H_8O_2$ | 494 | 520 |
| Ethyl benzoate | $C_9H_{10}O_2$ | 553 | 562 |
| Ethyl acetoacetate | $C_6H_{10}O_3$ | 471 | 454 |
| Anisole | $C_7H_8O$ | 422 | 419 |
| Phenetole | $C_8H_{10}O$ | 479 | 470 |
| Dimethyl resorcinol | $C_8H_{10}O_2$ | 519 | 517 |
| Furfural | $C_5H_4O_2$ | 345 | 362 |
| Valeraldehyde | $C_5H_{10}O$ | 348 | 305 |
| d-carvone | $C_{10}H_{14}O$ | 587 | 574 |

Einstein derived an expression relating atomic attraction to the surface tension $\gamma$ of a liquid at its boiling point. Each atom type is given a single value, which can be simply summed using the molecular formula of a compound. The values of $c$ derived by Einstein were: $c_H = -1.6$, $c_C = 55.0$, $c_O = 46.8$. The sum of these values ($\Sigma c_\alpha$) should, according to Einsein, equal $\sqrt{\gamma - T\frac{d\gamma}{dT}}$. Agreement between the two columns of figures is very impressive, yet the model is extremely simplistic and ignores entirely any type of directional chemical bond.

the potential of metals and their salt solutions 2 years later before turning to general molecular theories of heat. Somehow though, the planetary model did not catch on in chemistry, and is no longer used to the best of my knowledge.

Interest in calculating the strength of interactions between molecules has grown much stronger however throughout the 20th century and into this one. Crystallography and NMR have allowed us to determine the structures of many biological molecules, but this is only half the story; Life is crucially dependent on molecular *binding* – to the right target, at the right time, in the right place, with the right affinity, and (sometimes) at the right speed. While methods in structural biology have been developing very quickly for some years, genomics appeared almost overnight. We now know (or can find out) what a protein looks like in most cases, and we can model many of the difficult ones. To understand protein function, what is needed is a way to predict the affinity of protein–ligand interactions – the same problem faced by chemists trying to design drugs. A variety of groups from different backgrounds have therefore come up with models to predict how tightly a given protein will bind to a given small molecule. These solutions vary widely in complexity, from simple linear models to full-scale molecular simulations requiring days of computer time. Since the drug design (or lead search) problem requires many molecules to be examined, the more lengthy routes are ruled out. At first it may seem unlikely that a very simple function can take into account the complexities of the problem; proteins are flexible polymers and the physical chemistry involved in ligand binding in water is not trivial. Like Einstein however, many groups have found that the approximations involved in formulating a linear model of few parameters need not prevent it from performing admirably when applied to static protein models. Basing a model on sound chemical principles probably helps, but it is best to follow Einstein and not interpret the model too deeply.

## The problem

The problem is thermodynamics. How does the system of interest attain maximum entropy given the boundary conditions (usually constant

temperature and pressure) in an experiment? The macroscopic properties of the system which we measure are the ones which correspond to the configuration of the system with maximum permutability. If we mix a protein and ligand in solution, at equilibrium the bound and free ligands will have the same free energy. The energy of the system goes where the entropy tells it to; a reaction may take in or give out heat, but it will always produce entropy if left to itself and not driven by outside influences. This is rather a problem for simplistic calculations. Energy is a conserved quantity, and rather intuitive. We can make models for it, calculate it from force-fields, and divide it into different categories called 'Heat' or 'Electric potential'. Energy is nice and easy; entropy is messy. If *only* we didn't have to bother with entropy (and by using rigid protein models we have already removed quite a lot). *But we do*.

Attempts to make ligand affinity predictions using only energy terms have not been successful, and there is no theoretical basis for assuming they might be. Recently for example Luque and Freire have attempted to find a simple parameterisation of binding enthalpy of small ligands binding to proteins [2]. They write "*Because the binding enthalpy is the term that predominantly reflects the strength of the interactions of the ligand with its target relative to those of the solvent, it is desirable to develop ways of predicting enthalpy changes from structural considerations*." In the same article they continue "*Because the magnitudes of the enthalpy and entropy changes reflect different underlying interactions, ligands that have been enthalpically or entropically optimised exhibit different responses to changes in the target or the environment, even if they have the same affinity under the initial set of conditions*." Neither statement is justified (by citation or theory), but the authors have at least clearly stated the premise on which they are working, rather than hide their assumption. The best way to show entropy cannot be ignored is perhaps with a simple example, of alkanes being dissolved in water. This is a simple example which nicely illustrates some of the problems we are dealing with. It has been known for almost 70 years that dissolving hydrocarbons in water results in a large increase in heat capacity. The underlying cause of this has been widely discussed, but is generally agreed to result from greater ordering of the solvent water around the solute. This causes a large decrease in entropy, and around room temperature it is found that the enthalpy change is close to zero. So the low solubility of alkanes in water is due to an entropic effect at room temperature. But, the heat capacity change is large, and by definition

$$\Delta Cp = \frac{d\Delta H}{dT} = \frac{d}{dT}\frac{\Delta S}{T}$$

So as temperature changes, both the enthalpy and entropy changes change too. Since the overall free energy change (assuming constant temperature of reaction) is given by

$$\Delta G = \Delta H - T\Delta S$$

these changes largely cancel – the well-known enthalpy–entropy compensation – to leave $\Delta G$ relatively constant [3]. This does *not* mean the solubility of alkanes is independent of temperature. An equilibrium constant is related to $\Delta G$ by the equation

$$\Delta G = -RT\ln K$$

so the natural logarithm of the equilibrium constant varies with $\Delta G/T$ and not simply $\Delta G$. In fact it is fairly easy to show that $(d\Delta G/dT)$ equals $\Delta S$, so $\Delta G$ will be at a stationary value when the entropy is zero. More interestingly, the Gibbs–Helmholtz equation tells us that at constant pressure

$$\frac{d}{dT}\left(\frac{\Delta G}{T}\right) = -\frac{\Delta H}{T^2}$$

This can be rewritten as the van't Hoff equation:

$$\frac{d}{dT}(\ln K) = \frac{\Delta H}{RT^2}$$

As temperature is varied therefore, the equilibrium constant $K$ will pass through a maximum or minimum value at precisely the point $\Delta H$ drops to zero; whether we have a maximum or minimum depends on the sign of the heat capacity. In the case of alkanes dissolving in water, the forward reaction increases heat capacity so we have a minimum. At the temperature for which $\Delta H$ is zero, the solubility drops to a minimum. The heat capacity change itself is linear to a good approximation, over the temperature range of mammalian biochemistry anyway. The solubility of alkanes in water therefore describes a parabolic curve as temperature is increased from 0 to 40 °C and

beyond. (This gives rise to the half-truth sometimes still found in text-books that "Hydrophobic bonds become stronger at higher temperature"). This phenomenon explains a great deal of protein biochemistry, from cold denaturation to subunit dissociation at high or low temperatures. The hydrophobic effect is strongest around room temperature, though the temperature of minimum solubility is slightly different for aliphatics and aromatics, which are more polar.

The point I am trying to make is that it is not really possible to ascribe the insolubility of alkanes to the observed entropy or enthalpy since both will change with the conditions of the experiment while the *underlying phenomenon remains the same*. Hydrophobicity is not called different names for experiments at 20, 30 or 40 °C. Entropy–enthalpy compensation blurs the distinction between the two while largely preserving the overall free energy change. There has been enormous confusion in the literature concerning the hydrophobic effect, much of it arising in my view because of a confusion between $\Delta G$ and $\Delta G/T$. Essentially 'hydrophobicity' became associated with a large $\Delta G$ value in the minds of some authors, instead of a low solubility value (which is what it had meant and still means to everyone else). Since $\Delta G$ reaches a maximum when $\Delta S$ is zero, it was believed 'hydrophobicity' must arise from enthalpic considerations.

Another related mistake is also frequently found in the literature, that of assuming that an apolar ligand binding to an apolar pocket in a protein must bind through an increase in the system entropy – a positive $\Delta S$. But the chemistry involved is rather more complicated than a two-component oil–water mixture, and this assumption does not hold. Recently for example, Palencia and colleagues have found that proline-rich peptides bind to an SH3 domain with a favourable enthalpy change and an unfavourable entropy change [4]. They suggest this is 'inconsistent' with the known molecular structures which show the ligand and its binding site to be largely apolar. There is no inconsistency, just a growing list of papers whose authors puzzle unnecessarily over this frequently observed phenomenon.

To my mind enthalpy–entropy compensation is no difficulty to ligand modelling as long as one tries to predict binding *affinity*. Predicting enthalpy and entropy separately seems to me an extremely difficult task, and for the purpose of drug design not really a necessary one. Luque and Freire [2] note that it is possible to parameterise the enthalpy of denaturation of a protein from its structure, but these parameters prove unsuccessful in predicting the enthalpy of ligand binding. There are several issues here. As the authors note, the dataset used is rather small, and the errors in the measured enthalpies of ligand binding are comparatively large since the values are smaller than the enthalpy of unfolding. The same group has attempted this sort of parameterisation before, but no mention of the previous scheme [5] is given in the latest paper. It is always possible to divide up the measured affinity and enthalpy into components, but the values derived for the worth of a hydrogen bond or a square Angstrom of apolar surface area are not always found to be reliable outside the dataset used to calculate them. In the same way, it is simple to multiply two large prime numbers together, but very difficult to factorise the result.

Many different groups have produced different schemes for estimating binding affinity on the basis of X-ray structures of various protein–ligand complexes. In general these work fairly well in good cases (where the rigid body assumption is valid), but the scoring weights for hydrogen bonds and so on show considerable variation [6]. Very good agreement within the base dataset is a poor indicator of how well the scheme will cope with new complexes. (From a simple statistical viewpoint, it is often worthwhile omitting a small portion of the data for use as a test-set to give an unbiased estimate of how well the model derived from the 'working' data truly model the whole data-set. This type of cross-correlation analysis revolutionised the way X-ray crystallographers regarded traditional R-factors of their atomic models). Enthalpy–entropy compensation seems to improve significantly the accuracy in predicting energy, whereas estimates of enthalpy (or entropy) alone do not have this advantage. Clearly the physical phenomena underlying ligand binding to proteins and protein folding are the same, and if we understood them perfectly then we could no doubt predict the thermodynamics of both processes. But scoring functions are inherently approximations, and very different assumptions are made concerning the protein entropy change in the two cases. For ligand binding it is generally assumed to be zero (or very close), a wholly

inappropriate assumption for the case of protein folding. So it is not altogether surprising to find that other groups have noted systematic differences in the weights derived from ligand binding and protein folding [7]. The very large entropy term in the protein folding case is absorbed into the derived parameters. Theoretical calculations suggest that a small molecule binding to a protein will also incur a large entropy penalty, but to date no experimental data support such a view.

A number of methods rely on the calculation of buried surface areas. In itself this is not a complicated idea, but it is important to be careful what is meant. There are at least three different types of 'surface area' used in molecular modelling [8], and even more programs to calculate them. The following extract from an email to PDB (the Protein Databank, now held at Rutgers University) in 2001 is fairly typical:

*Does anybody have a list of molecular surface areas of all residues in extended conformations? Or knows where to find it? I'm trying to calculate surface areas and I get different numbers from different programs. For example, Leu has area of 115* $Å^2$ *from one program and 300* $Å^2$ *from another. Which one is closer to the truth?*

While the larger value may represent the main-chain as well as the side-chain, the smaller value is still much smaller than the 180 $Å^2$ reported by Miller and colleagues [9]. So there is huge variation, and not all programs give a clear description of what they are doing. The program WHATIF is claimed to reproduce the accessible molecular surface area to within 5%, the difference being that it does not take the re-entrant surface into account. Personally I think an error of 5% in this term pretty good, though one referee of this paper begs to differ on this point.

## Protein flexibility

As an example of the difficulties in predicting binding affinity from an atomic model of the protein complex, Figure 1 shows the data obtained from an isothermal titration calorimeter on adding aliquots of tryptophan to the tryptophan attenuation RNA-binding protein, TRAP [10]. TRAP binds to the *trp* mRNA of *Bacillus subtilis* in a tryptophan-dependent manner. The protein is an 11-mer ring with identical subunits, the tryptophan

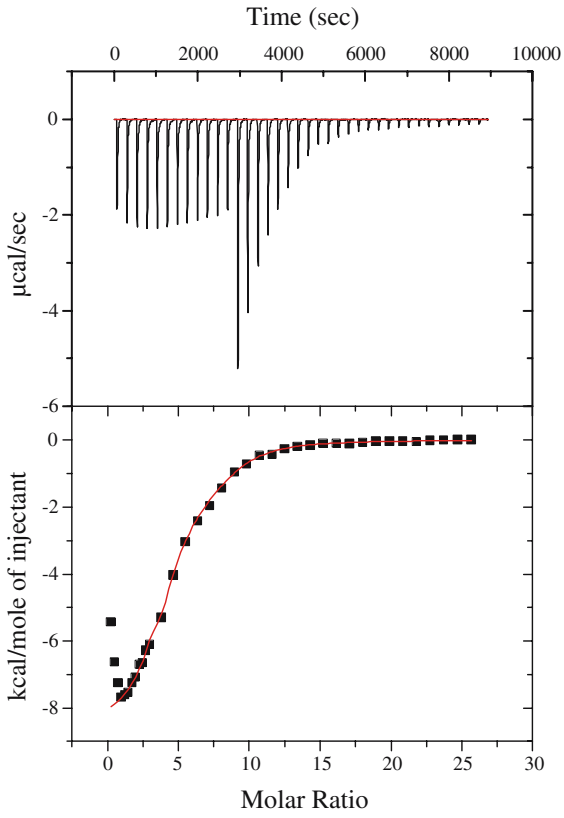


*Figure 1.* The isothermal titration calorimetry trace obtained on injecting L-tryptophan ligand into TRAP in 50 mM sodium phosphate pH 7.2 at 25 °C. The ligand concentration was 2.4 mM and the initial protein concentration 21.8 µM. A total of 37 injections were made, the first 12 of 3 µl and the remainder 10 µl. The downward deflection in the trace of the power supplied to the sample cell (shown in the upper panel) indicates an exothermic reaction occurs on mixing ligand and protein. The integrated heats are shown in the lower panel as black squares, and the line indicates the heat per mole expected given a model of two sites. The first three data points show an increase in the heat released per mole of tryptophan injected. Often in ITC the first data point proves unreliable due to diffusion during the thermal equilibration of the instrument immediately prior to an experiment. The initial data point has not been omitted from any of the experiments presented here. Small injections were used at first in order to observe the highly reproducible increase in exothermicity, seen at all temperatures studied. This effect cannot be modelled assuming independent binding sites.

molecules binding at 11 symmetry-related sites at the monomer interfaces. From the calorimeter trace it can be seen that the initial tryptophans bind exothermically, and increasingly so over the first few injections (the system has positive cooperativity). The experiment shown in Figure 1 was performed at 25 °C. Repeating the same experiment

at 12 °C (shown in Figure 2), $\Delta H$ rises rapidly to become endothermic for the last few ligands. This very marked change in the enthalpy of binding (of the same ligand to identical binding sites) is incompatible with the view that $\Delta H$ must reflect the molecular contacts between a protein and ligand. Despite its rigid appearance, the TRAP ring is clearly capable of cooperative behaviour, through a mechanism much more subtle than that of haemoglobin. Further studies are underway on this system, but these early results are already a conclusive demonstration of the difficulties in affinity prediction, let alone enthalpy prediction, by simple linear models. This problem is not confined to proteins with a single ligand binding site. A protein has considerable flexibility and we currently have no simple means of estimating how much its intrinsic entropy will be changed by ligand binding. There are also plenty of examples of proteins whose binding sites substantially alter on ligand binding. It is an inherent assumption of simple scoring functions that these effects are small, though as shown above it is quite possible for the enthalpy and entropy terms to change sign.

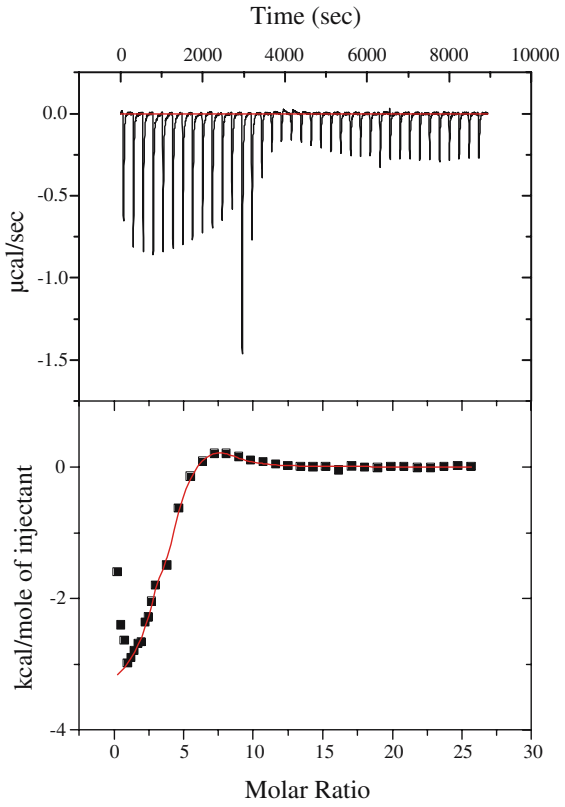


*Figure 2.* The same experiment shown in Figure 1, but at 12 °C. In this case the last ligand molecules to bind do so endothermically.

## Conclusion

There has been considerable progress in ligand scoring in the 5 years since I wrote my previous review of scoring functions [6], and computational methods are contributing more and more to drug discovery and design [11, 12]. In a 2002 review of ligand docking and scoring functions, Nussinov and colleagues cited nearly 300 papers [13]. This has been aided by the growing number of well-characterised systems in which both accurate thermodynamic data and a high resolution X-ray structure of the protein–ligand complex are available. More groups are focussing on the problem of flexibility of both protein and ligand, which clearly must be addressed if present scoring functions are to improve significantly in accuracy and precision [14]. Fitting hydrophobic ligands into an apolar pocket in a computer often gives a very broad potential energy well as the ligand can slide or rotate while maintaining hydrophobic contacts. Crystallography suggests however that even almost entirely hydrophobic drug molecules can adopt one highly preferred conformation in their binding sites, indicative of a much steeper potential energy well than the modelling studies suggest. To be fair, in some cases this may result from low resolution X-ray data, or inappropriate weighting of geometrical restraints. Nevertheless, simple scoring functions are clearly capable of rapid screening with appreciable success, and compensating to some extent for inaccurate docking. Kontoyianni and colleagues have recently presented an interesting comparison of 10 scoring functions judging ligands placed by four different docking engines, and show that the scoring functions can discriminate successfully between accurate and inaccurate ligand poses [15]. While scoring functions have already demonstrated great potential in rapid lead selection, it remains questionable whether or not the inherent assumptions on which they are based will permit reliable, general affinity estimation. Simply associating a protein–ligand complex with a measured $K_d$ value ignores a great many variables including temperature, pH, and the nature of the biophysical technique employed in affinity

measurement. But the next 100 years may have a few more surprises in store.

**Acknowledgements**

I would like to thank Prof. Adrian Elcock for helpful comments on the manuscript, and the enthusiastic but anonymous reviewer who has travelled the same road before me!

**References**

1. Einstein, A., Annalen der Physik, 4 (1901) 513.
2. Luque, I. and Freire, E., Proteins, 49 (2002) 181.
3. Sharp, K., Protein Sci, 10 (2001) 661.
4. Palencia, A., Cobos, E.S., Mateo, P.L., Martinez, J.C. and Luque, I., J. Mol. Biol., 336 (2004) 527.
5. Luque, I. and Freire, E., Methods Enzymol., 295 (1998) 100.
6. Tame, J.R.H., J. Comput. Aided Mol. Des., 13 (1999) 99.
7. Xu, D., Lin, S.L. and Nussinov, R., J. Mol. Biol., 265 (1997) 68.
8. Greer, J. and Bush, B.L., Proc. Natl. Acad. Sci. U.S.A, 75 (1978) 303.
9. Miller, S., Janin, J., Lesk, A.M. and Chothia, C., J. Mol. Biol., 196 (1987) 641.
10. Babitzke, P., Curr. Opin. Microbiol., 7 (2004) 132.
11. Ferrara, P., Gohlke, H., Price, D.J., Klebe, G. and Brooks, C.L.I., J. Med. Chem., 47 (2004) 3032.
12. Ma, X.H., Wang, C.X., Li, C.H. and Chen, W.Z., Prot. Eng., 15 (2002) 677.
13. Halperin, I., Ma, B., Wolfson, H. and Nussinov, R., Proteins, 47 (2002) 409.
14. Stone, M.J., Acc. Chem. Res., 34 (2001) 379.
15. Kontoyianni, M., Sokol, G.S. and McClellan, L.M., J. Comput. Chem., 26 (2005) 11.