Entropy–enthalpy compensation as deduced from measurements of temperature dependence

Athel Cornish-Bowden

Unité de Bioénergétique et Ingénierie des Protéines,
Centre National de la Recherche Scientifique, Marseilles, France

Correspondence: Dr. Cornish-Bowden, A., CNRS-BIP,
31 chemin Joseph-Aiguier, B.P. 71, 13402 Marseille Cedex 20, France
Email: acornish@ibsm.cnrs-mrs.fr; Fax: + 33 491 16 46 61

This is a preprint of a chapter in Protein–Ligand Interactions (2012, ed. H. Gohlke), Wiley–Blackwell, Weinheim, pp. 33-43

1 Introduction

The concept of entropy–enthalpy compensation has exercised a peculiar fascination for biochemists since it was first proposed.[1, 2] The idea is that if a particular rate constant or equation constant is measured for a series of sample of the same enzyme from different sources, the variations in enthalpies of activation or of reaction will largely cancelled out by variations in the corresponding entropies, cancelling out completely at a temperature known as the compensation temperature. Until recently, the evidence for this phenomenon came mainly from studies of temperature variation, more specifically from Arrhenius or van ’t Hoff plots, despite the fact that Exner [3] pointed out many years ago that the supposed correlation between the slopes (enthalpies) and ordinate intercepts (entropies) of these plots is almost wholly due to a statistical artefact, and means strictly nothing. In recent years there has been been an increasing tendency to base conclusions on direct calorimetric measurements of the heat produced or consumed in reactions [4–6], and it is important to note at the outset that nothing I say in the rest of this article has any bearing on calorimetric measurements, though Sharp [7], for example, has also criticized conclusions of compensation based on these. In addition, given a wide enough temperature range in which no complications such as changes in state intervene, say from 50 to 300 K (–223 to 27 °C), it is not impossible for an Arrhenius or a van ’t Hoff plot to give meaningful and approximately independent values of both entropy and enthalpy, but it is sufficiently unlikely for these conditions to be fulfilled in a biochemical context that it seems safe to ignore it.

If one surveys the recent literature, one can find at least four kinds of article. There are increasing numbers of papers based on calorimetry [4–6], and, as noted already, these fall outside the scope of the present article; there are a few — extremely few — that recognize the validity of Exner’s analysis, and conclude that no biological significance can be attached to estimates of thermodynamic parameters based on measurements of
temperature dependences [8]; there are others that admit the existence of problems but prefer to set them aside [9, 10] or argue that non-linear regression on abundant data [11], careful interpretation [12] or “extreme caution” [13] can avoid the problems inherent in the use of graphs; finally, there are papers that largely ignore the question of whether the evidence advanced in favour of entropy–enthalpy compensation is sufficient to justify regarding it as a real property, but instead discuss how it might be explained if it were real [14]. It will be evident that this last approach is essentially theological — let us study the attributes of the creator without raising the question of whether he exists — and in this article I shall be concerned with the third. I shall also refer for convenience only to kinetic data (Arrhenius analysis), though essentially everything applies equally to equilibrium data (van’ t Hoff analysis).

For this it is important to distinguish between the use of graphs to illustrate the results of an investigation, and their use to obtain them. Graphical analysis as such is less and less used in modern science, and some would like it to disappear altogether [15], but it remains an essential tool for purposes of illustration, and, most important, for making a visual assessment of the credibility of the results that emerge from a computer: the notion that graphs are obsolete is popular with non-statisticians, but if finds no support in the writings of statisticians [16]. It is important to emphasize, therefore, that Exner’s conclusions [3] were in no way dependent on an assumption that estimates of the entropy and enthalpy of activation are obtained directly from graphs. Likewise in my earlier paper [17] I used a graph (Fig. 2 therein) to illustrate the results of analysing 100 simulated experiments in which the value of a rate constant $\ln k_{18^\circ C}$ had a random value uniformly in the range $2.30 \pm 1.15$ and the Arrhenius activation energy $E_A$ was uniformly distributed in the range 30–160 kJ/mol. However, these results were not obtained by laboriously drawing 100 Arrhenius plots and estimating the the ordinate intercepts by inspection: instead the entire operation was carried out in the computer and no lack of care in drawing the plots or analysing them was involved. The point is that no amount of extreme caution or non-linear regression can reveal information that is not present in the data.

2 Measurement of the entropy and enthalpy of activation

According to the transition-state theory of reaction rates, the dependence of a rate constant $k$ on the absolute temperature $T$ is given (subject to various assumptions that I shall ignore) according to the following equation:

$$\ln k = \ln \left( \frac{RT}{Nh} \right) + \frac{\Delta S^\ddagger}{R} - \frac{\Delta H^\ddagger}{RT}$$

in which $R$ is the gas constant, $N$ is the Avogadro constant and $h$ is Planck’s constant. The variation due to the first term on the right-hand side is typically negligible compared with that due to the others, and is anyway calculable as it is composed of known quantities. It follows that to a good approximation $\ln k$ is a linear function of $1/T$, with $\Delta H^\ddagger$ determined by the derivative of $\ln k$ with respect to $1/T$, and $\Delta S^\ddagger$ by the value at infinite temperature, i.e. at $1/T = 0$. In biochemical experiments, however, the temperature range accessible to investigation is typically small, for example 0–18 °C in the experiments of Johnson and Goldspink [18] used by Gutfreund [19] as an illustration of entropy–enthalpy compensation. This corresponds to a range of $1000/T$ of 3.67–
Table 1. Illustrative data

<table>
<thead>
<tr>
<th>i</th>
<th>Temperature (°C)</th>
<th>1000/T (K⁻¹)</th>
<th>ln k</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.0</td>
<td>3.60</td>
<td>1.81</td>
</tr>
<tr>
<td>2</td>
<td>10.0</td>
<td>3.53</td>
<td>2.26</td>
</tr>
<tr>
<td>3</td>
<td>15.0</td>
<td>3.47</td>
<td>2.80</td>
</tr>
<tr>
<td>4</td>
<td>20.0</td>
<td>3.41</td>
<td>2.88</td>
</tr>
<tr>
<td>5</td>
<td>25.0</td>
<td>3.36</td>
<td>3.43</td>
</tr>
</tbody>
</table>

Table 2. Fitted parameters for the data in Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope</td>
<td>b</td>
<td>-6.40</td>
</tr>
<tr>
<td>Variance of the slope</td>
<td>σ²((\hat{b}))</td>
<td>0.436</td>
</tr>
<tr>
<td>Standard error of the slope</td>
<td>σ((\hat{b}))</td>
<td>0.66</td>
</tr>
<tr>
<td>Coefficient of variation of the slope</td>
<td>σ((\hat{b}))/</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>(\hat{a})</td>
<td>24.9</td>
</tr>
<tr>
<td>Variance of the intercept</td>
<td>σ²((\hat{a}))</td>
<td>5.26</td>
</tr>
<tr>
<td>Standard deviation of the intercept</td>
<td>σ((\hat{a}))</td>
<td>2.3</td>
</tr>
<tr>
<td>Coefficient of variation of the intercept</td>
<td>See text</td>
<td></td>
</tr>
<tr>
<td>Experimental variance</td>
<td>σ²((\ln k))</td>
<td>0.0159</td>
</tr>
<tr>
<td>Experimental standard error</td>
<td>σ((\ln k))</td>
<td>0.126</td>
</tr>
<tr>
<td>Covariance</td>
<td>cov((\hat{a}, \hat{b}))</td>
<td>-1.51</td>
</tr>
<tr>
<td>Correlation coefficient</td>
<td>ρ((\hat{a}, \hat{b}))</td>
<td>-0.9997</td>
</tr>
</tbody>
</table>

3.44 K⁻¹, so the range from the lower 1/T to 0 is about 15 times the range used in the experiments, and estimation of Δ𝑆^† involved a 15-fold extrapolation.

3 An example

In favourable cases one may be able to obtain usable data over a range as wide as 0–40 °C, but restricted ranges are quite common, and to illustrate the statistical problems I shall consider data in the range 5–25 °C, or 1000/T = 3.36–3.60 K⁻¹, as listed in Table 1. Linear regression of these values assuming equal weight for each value of \(\ln k\) gives the results in Table 2, in which the symbols used in the second column are defined at the left.

The results in Table 2 repay careful study, as they illustrate several important points that are often overlooked in studies of entropy–enthalpy compensation. First of all, the small value of the experimental standard error (which can be understood as a measure of the precision of the value of \(\ln k\) at the centroid of the data) is not reflected in a small value of the standard error of the intercept, nearly 20 times larger: this is a direct consequence of the long extrapolation. It is not meaningful to express this as a coefficient
of variation, because the intercept is measured from an arbitrary standard state.\textsuperscript{1} However, an uncertainty of $\pm2.3$ in $\ln k$ implies an uncertainty of an order of magnitude in $e^b$, which is, of course, huge. The coefficient of variation of the slope is meaningful, however, as it does not depend of the units used for measuring the temperature, so one can say that values of $\ln k$ spread over a narrow range of temperature can give a reasonable estimate of the Arrhenius activation energy.

The next point to notice is that the covariance (often ignored, but crucial in this example) is almost exactly equal to the geometric mean of the two variances:

\[
\text{cov}(\hat{a}, \hat{b}) \approx |\sigma^2(\hat{a}) \cdot \sigma^2(\hat{b})|^{0.5}
\]

This is not an artefactual property of the specific numbers used, but is a universal feature of data sets in which the estimate of the intercept requires a very long extrapolation. It is responsible for the correlation coefficient between that two parameters that is barely distinguishable from $-1$. The practical meaning of this is that the data set is capable of defining one parameter only, not two. Artificial analysing this one parameter, the value of $\ln k$ at the centroid, which can be estimated quite precisely, into two, $\hat{b}$ and $\hat{a}$, is simply an arithmetical trick that does not create any new information that is not present in the data. As a result, any small variation in the estimate of $\hat{b}$ is transmitted almost perfectly to a corresponding variation in the estimate of $\hat{a}$. This is the basis of the illusion that entropy–enthalpy compensation can be detected from measurements over a restricted range of temperature.

Figure 2 illustrate the effect of making a long extrapolation on the possibilities for parameter estimation. A one should expect, the value of the dependent variable is very well defined at temperatures within the range of measurement, but increasingly less precisely as one moves away from this range. Notice that the standard error of a parameter defines, at best, about a 68\% confidence interval. A 95\% confidence interval corresponds to about two standard deviations, but this assumes a sample of infinite size, whereas Figure 2 is drawn for a sample of 5 observations. Moreover, we usually have no information in practice about whether the assumptions about error distribution needed for the calculation are valid. In this case giving equal weight to each value of $\ln k$ effectively assumes that the $k$ values are lognormally distributed, or approximately that they have a uniform coefficient of variation. The value of $\sigma(\ln k) = 0.126$ from Table 2 corresponds to a precision of about $\pm13\%$ in the values of $k$. This is reasonable, but we cannot be sure that is correct. It follows from all these considerations that it would be very optimistic to think that the data in Figure 2 allow the value of $k$ at $1/T = 0$ to be defined to as little as two orders of magnitude, and probably not as precisely as that.

It is important to emphasize that this conclusion is based entirely on computer analysis of the data. It does not depend in any way on how accurately one can draw an Arrhenius plot or draw conclusions from it. If one does use graphical analysis then inaccuracies in drawing and analysing the graph will of course add to the problems, but they are not the main sources of them.

The near impossibility of obtaining an accurate value of $\ln k$ at $1/T = 0$ is only the beginning of the difficulties. The virtually perfect correlation between estimates of $\hat{a}$ and $\hat{b}$, with $\rho(\hat{a}, \hat{b}) = -0.9997$ means that we cannot regard $\hat{a}$ and $\hat{b}$ as even approximately independent estimates of two different quantities. We cannot even regard them

\textsuperscript{1}For example, if we assume that the values of $k$ are measured in $s^{-1}$ but we convert them to $\text{min}^{-1}$ before taking logarithms, we would have $\hat{a} = 20.8 \pm 2.3$ instead of $\hat{a} = 24.9 \pm 2.3$, so the “coefficient of variation” would be changed from 9.2\% to 11.1\% with no change whatsoever in the physical reality.
as two estimates of the same quantity estimated in different ways: they are in reality measurements of one quantity measured in one way but subsequently transformed by a mathematical trick in two different ways.

I have explored the practical consequences of this high correlation between \( \hat{a} \) and \( \hat{b} \) elsewhere,[17] and here it will be sufficient simply to summarize the results. A random scatter of values of \( \ln k_{18^\circ C} \), the value of \( \ln k \) at 18 \( ^\circ \text{C} \), uniformly distributed in the range 2.3 \( \pm \) 1.15 (i.e. spread randomly over an order of magnitude) was generated, and for each of these a random value of the Arrhenius activation energy \( E_a \) uniformly distributed in the range 30–160 kJ/mol was generated. Each \((\ln k_{18^\circ C}, E_a)\) pair was then taken to define the true parameter values for simulating a temperature dependence with 19 observations in the range 0–18 \( ^\circ \text{C} \), normally distributed errors with standard deviation 0.05 being added to the calculated \( \ln k \) values. The conditions were chosen to correspond roughly to those of the experiments of Johnson and Goldspink [18], which have been used as a textbook illustration of entropy–enthalpy compensation.[19] It should be obvious that a completely random set of data of this kind cannot possibly provide evidence of a meaningful relationship between the entropy and enthalpy of activation. However, when the resulting dependences were analysed they showed a very good compensation, which became virtually perfect when only seven randomly chosen points were plotted instead of the full 100 generated. As the correlation was not present in the data it can only have been an artefact of the mathematical manipulation. Even when the scatter of \( \ln k_{18^\circ C} \) was set to 2.3 \( \pm \) 6.9 (i.e. a millionfold variation in \( k_{18} \)), the apparent correlation was very good, though no longer virtually perfect.
4 Evolutionary considerations

It is also worth asking what biological interpretation could be put on an enthalpy-entropy compensation if it truly existed and was not simply a statistical artefact. The fishes used by Johnson and Goldspink [18] as a source of ATPase are very different from one another: *Tilapia grahami*, for example, is a cichlid fish that lives in highly alkaline conditions at temperatures that can be as high as 43 °C, whereas *Notothenia rossii*, the “ice fish”, lives in Antarctic waters at temperatures as low as 0 °C. It would be quite surprising, therefore, if they showed any evolutionary similarities beyond the features common to all teleosts. Any genuine similarity in their thermodynamic parameters would be more easily explained as chance than as a result of natural selection. However, there is a further problem. Moreover, not only *Tilapia grahami*, but also some of the other fishes in the sample, lives at a much higher temperature than the upper limit of the range 0–18 °C considered in the analysis. It is therefore hard to see what relevance the properties of its ATPase below 18 °C could have to its evolutionary history.

5 Textbooks

Far from explaining the difficulty of estimating the entropy of activation from data of a limited temperature range, textbooks tend to exacerbate the problem. What is one to make, for example, of the graph in Figure 3? This is inspired by an illustration of a van ‘t Hoff plot in a recent and current textbook of thermodynamics,[20] but it is typical of what one can find in others. As no numerical values are shown on the $1/T$ scale, there is an obvious implication that the vertical axis was drawn at $1/T = 0$, so that only a short extrapolation is needed to estimate $\Delta S^\ddagger$. In fact this is more than an implication, because the labelling of the intercept as $\Delta S^\ddagger / R$ cannot be correct unless the vertical axis is at $1/T = 0$. 

![Figure 2](image-url)
Thus apparently the data span a 15-fold range of absolute temperature. Is this possible for a biological experiment? Strain 121 of Methanopyrus kandleri can grow at 121 °C,[21] so let us suppose that the upper range of temperature in the Arrhenius plot was 121 °C (394 K). If so, the lowest temperature was –247 °C, which appears impossible for a biological system. If we make the opposite assumption, that the low temperature limit corresponds to the lowest temperature at which organisms are known to grow, say –12 °C for Psychromonas ingrahamii [22], then the result is no more believable, because now the upper temperature must be 3640 °C. If we try to estimate the widest possible range of temperature in which an enzyme could be studied, and suppose that an enzyme can be active 10 °C hotter than any organism can grow, and 10 °C colder than any other organism can grow, we find a range of –22–131 °C, or 251–404 K, or 1.6-fold, very much less than the 15-fold range seen in Figure 3. It follows that not only does Figure 3 illustrate an invented experiment, it illustrates an impossible experiment that has no relevance to entropy measurements in the real world.

As anyone who has taught thermodynamics will know, textbooks of low quality, written by authors who do not understand the subject themselves, are very widespread, and if Figure 3 were based on an example from such a book one would find it regrettable but hardly surprising. In fact, however, Biological Thermodynamics [20] is not such a book, but is one of the best sources of thermodynamic information available for the modern student (especially for the more chemical topics). Kinetics for the Life Sciences [19] is older, but it is also one of the more authoritative books in its subject area. In the circumstances, therefore, we cannot be surprised if the dangers in concluding the reality of entropy–enthalpy compensation from measurements of temperature dependence continue to be poorly recognized near half a century after they were first pointed out. [3]
References


