Enthalpy–entropy compensation: a phantom phenomenon

ATHEL CORNISH-BOWDEN

Bioénergétique et Ingénierie des Protéines, Institut Fédératif "Biologie Structurale et Microbiologie",
Centre National de la Recherche Scientifique, 31 chemin Joseph-Aiguier, B.P. 71, 13402 Marseille Cedex 20, France

(Fax, 33-491-16 46 61; Email, athel@ibsm.cnrs-mrs.fr)

Enthalpy–entropy compensation is the name given to the correlation sometimes observed between the estimates of the enthalpy and entropy of a reaction obtained from temperature-dependence data. Although the mainly artefactual nature of this correlation has been known for many years, the subject enjoys periodical revivals, in part because of the frequent excellence of the correlation. As with other cases of impossibly good correlation between two biological variables, the explanation is that what purports to be two variables are very largely the same variable looked at in two different ways.

1. Introduction

Enthalpy–entropy compensation enjoyed widespread popularity in the 1950s, even though it was soon shown (Exner 1964) to be largely derived from a statistical artefact of trying to extract two parameters from temperature-dependence data that could barely support one. Although it is now relatively rare in the literature, there have been occasional revivals: for example Gutfreund (1995) mentions it prominently in his book on biological kinetics, and it was the central theme in the book of Hochachka and Somero (1984) on biochemical adaptation. The basic observations are superficially straightforward and highly impressive: for example, the apparently excellent linear relationship between enthalpy and entropy of activation obtained by Johnson and Goldspink (1975) for the temperature dependences of the myofibrillar ATPases of various fishes, which provides the data for figure 7·5 of Gutfreund (1995). To avoid being misled into seeing a real biological relationship where none exists, it is therefore important to understand how the artefactual correlation arises.

First, however, it is important to emphasize that the major problem arises when the enthalpy and entropy values are estimated from the same temperature-dependence data. In principle, they can be determined independently by calorimetry, i.e. by directly measuring the heat transfer rather than just inferring it. This is sometimes also done in practice, and so it must be understood that when enthalpy–entropy compensation is deduced from calorimetric data it cannot be easily dismissed as a statistical artefact.

In deducing compensation from temperature-dependence data the essential idea is simple: an enzyme is isolated from various sources, typically fishes that live at significantly different temperatures, and for each enzyme the temperature dependence of a suitable kinetic parameter k is analysed by making Arrhenius plots of $\ln k$ against $1000/T$, where T is the temperature in kelvins. The slope and intercept of the resulting linear plot are then reprocessed in the light of transition-state theory (Gutfreund 1995; Cornish-Bowden 1995) to yield estimates of the enthalpy $\Delta H^{\ddagger}$ and entropy $\Delta S^{\ddagger}$ of activation. Plotting $\Delta H^{\ddagger}$ against $\Delta S^{\ddagger}$ yields a "compensation plot", typically a straight line whose slope has dimensions of temperature. The name refers to the idea that variations in $\Delta H^{\ddagger}$ that accompany variations in the temperature at which each enzyme is normally active are "compensated for" by variations in $\Delta S^{\ddagger}$. Significance is also sometimes attached to the fact that the slope of the plot, known as the

"compensation temperature," is often found to be in the range at which most organisms live.
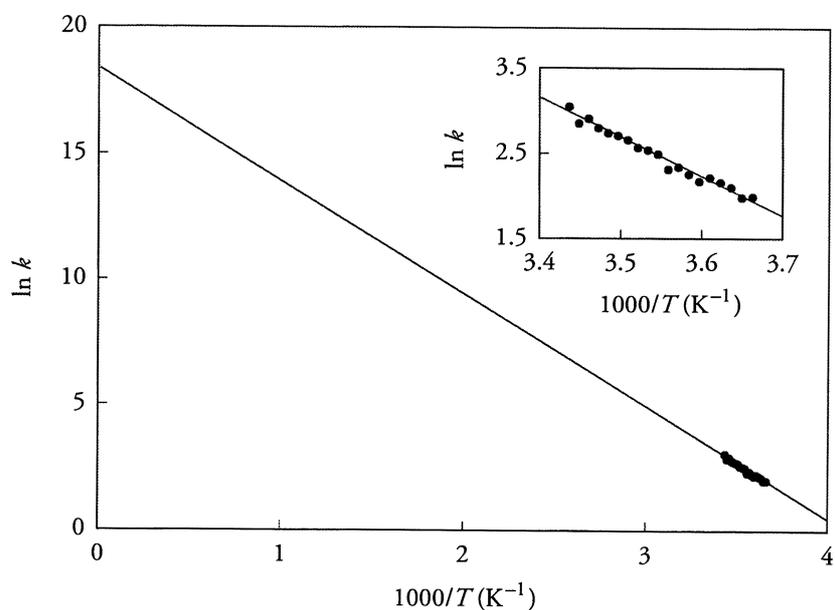
2.   Extrapolation of the Arrhenius plot

An Arrhenius plot for an enzyme parameter typically has the appearance shown in the inset to figure 1, apparently a well defined straight line with well defined slope and intercepts. It is important to notice, however, that the origin is almost never shown explicitly on an Arrhenius plot. So far as the abscissa axis is concerned this is of no importance, because its location is arbitrarily dependent on the units in which k is expressed before converting it to its logarithm – if the day had been adopted as the SI unit of time instead of the second, this axis would be shifted by 11·4, but the physical reality would be unchanged. However, the location of the ordinate axis is not arbitrary, and the fact that it typically lies very far from the range of experimental data has important consequences.

To indicate what analysis of an Arrhenius plot in terms of the transition-state theory actually involves one must use a scale with which the ordinate axis can be shown explicitly, as in the main body of figure 1. It is then evident that estimating $\Delta S^{\ddagger}$ by extrapolating the line to this axis – equivalent to extrapolating the value of k to

infinite temperature – involves an extrapolation of typically 8–20 times the range of the data. In figure 1, as in the experiments of Johnson and Goldspink (1975), the temperature range was 273–291 K (0–18°C), corresponding to a range of 3·44–3·66 K^{-1} in $1000/T$: the extrapolation is thus more than 15 times the range of the data. Even if the experimental range were as wide as 273–318 K (0–45°C) the extrapolation would still be seven times greater than the range of experimental $1000/T$ values.

To put this in terms more familiar to most biochemists, the extrapolation is similar to what would be involved in trying to estimate the limiting rate V of an enzyme from a Lineweaver–Burk plot of data in which the entire range of substrate concentrations extended from $0·083K_{m}$ to $0·089K_{m}$. (The parallel is not exact, however, because the location of the abscissa axis in an Arrhenius plot is arbitrary, as noted above, whereas in a Lineweaver–Burk plot it is not arbitrary.) Although badly designed experiments with inappropriate ranges of substrate concentrations are often used as the basis for published Lineweaver–Burk plots, it would be hard to find an example of one even remotely as badly designed as this!

One thing follows immediately from this analysis: even if a typical Arrhenius plot of biological data may be capable of giving a meaningful estimate of $\Delta H^{\ddagger}$, which is derived solely from the slope, it cannot give more than a

Figure 1.   A typical Arrhenius plot of biological data. Notice that the data extend over a very narrow range of temperature (here 0–18°C), so that extrapolating to infinite temperature to obtain the thermodynamic activation parameters is a highly uncertain procedure. The inset shows a more conventional representation of the same data, where the truncated scales permit an entirely misleading impression that the points lie on a well-defined straight line.

rough idea, at best, of $\Delta S^{\ddagger}$, because this requires knowledge of the ordinate intercept. However, although this warns us not to be too confident about attributing biological significance to observed variations in $\Delta S^{\ddagger}$, it is not obvious at first sight that it will generate a spurious correlation between $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$. To understand this it is useful to examine a more detailed example.

3. Random temperature dependences

Figure 2a shows 100 points distributed at random in a space with the Arrhenius activation energy E_a as one dimension and $\ln k_{18°}$, the logarithm of the value of the parameter k at 18°C, as the other. The values of E_a are uniformly distributed in the range 25–160 kJ mol^{-1}, and those of $\ln k_{18°}$ are uniformly distributed over a range of 2·3 (so that the values of $k_{18°}$ extend over a ten-fold range). It is obvious from inspection that there is no relationship between the values of E_a and $\ln k_{18°}$, as any appropriate statistical test would confirm. Suppose now that a set of k values is generated at temperatures in the range 0–18°C for each of the 100 combinations of the two Arrhenius parameters. The data shown in figure 1 were one such set and are typical of the level of scatter

assumed through out. In general the conditions (range of temperatures, parameter values, numbers of observations in each experiment) in this simulation corresponds roughly with the data of Johnson and Goldspink (1975), but this is not important for the conclusions that follow, which would be essentially the same for any reasonable assumptions. The main difference is that whereas Johnson and Goldspink (1975) considered data for seven fishes the simulation here considers a much larger number of parameter combinations to eliminate the danger of spurious conclusions due to a small sample.

Each of the Arrhenius plots generated in this way can be analysed by fitting it to a straight line and estimating the slope $-E_a/R$, where R is the gas constant, and intercept $\ln A$ is on the ordinate. From this it is a simple matter to calculate $\Delta H^{\ddagger} = E_a - RT$ and $\Delta S^{\ddagger} = R \ln (AN_Ah/RT) - R$, where N_A is the Avogadro constant and h is Plank's constant, using the fact that RT/N_Ah has a value of $6·25 \times 10^{12}$ s^{-1} at 300 K. As the original parameter values (figure 2a) were uncorrelated, we might expect a corresponding random scatter of points in the compensation plot, but in reality the result is an excellent straight-line "relationship" (figure 2b), with a slope of 290 K (17°C), very close to the temperature at which the original $\ln k_{18°}$ values were defined.
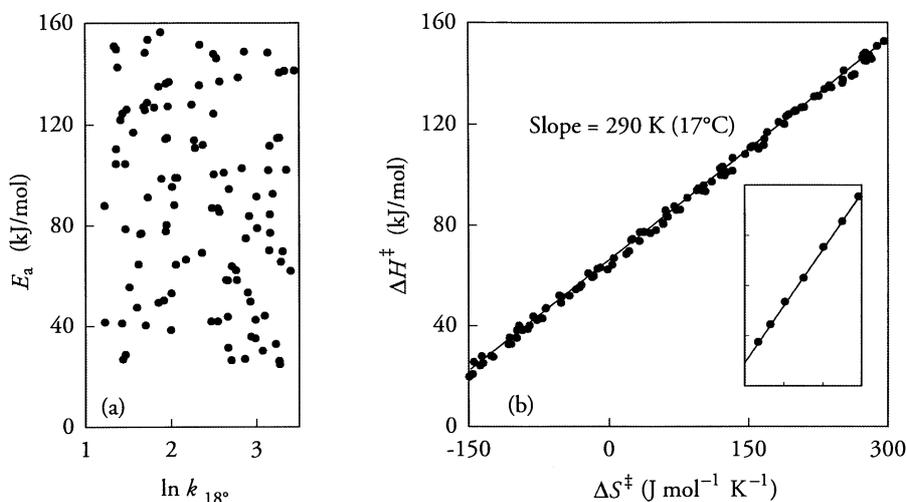
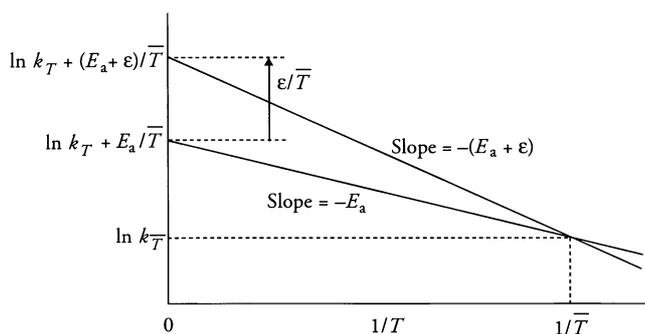**Figure 2.** (a) A random scatter of 100 points, with $\ln k_{18°}$ uniformly distributed in the range 2·3 ± 1·15 and E_a uniformly distributed in the range 30–160 kJ/mol. (b) Each of the points in (a) was used to provide the true parameter values for simulating a temperature dependence with 19 observations in the range 0–18°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 (1975). The resulting Arrhenius plots were then analysed to estimate the thermodynamic activation parameters, which were used to produce the compensation plot shown. The inset shows seven points, well spaced but otherwise selected at random from the original 100, plotted in the same way. This plot is unlabelled for clarity, but the range of values plotted in each axis was the same as in the main plot.

Where does the excellent correlation in figure 2a come from? Clearly it cannot reflect any biological relation between the actual parameter values, which, as seen in figure 2b, were unrelated. Moreover, a correlation as good as that in figure 2b ought to give rise to considerable suspicion, regardless of what sort of data are plotted, because few if any relationships between biological variables are as exact as this, and it is especially noteworthy that the quality of the transformed data appears to be much better than that of the primary data in the Arrhenius plots (of which figure 1 shows a typical example).

In general, an impossibly good correlation between two biological variables normally means that what purport to be two variables are, apart from a trivial amount of variation in a minor dimension, the same variable looked at in two different ways. So it proves in the present case. Consider, for example, the effect of displacing one of the experimental points in the Arrhenius plot by a small amount. The primary effect will be a small error in the slope accompanied by a negligible error in the coordinates of the centroid. In effect, therefore, we can expect the line to be rotated slightly around the centroid, which means in practice that an error of e in the estimate of E_a will translate to an error of $e/\overline{T}$, where $\overline{T}$ is the temperature at the centroid, in the estimate of the intercept on the ordinate (figure 3). To the extent, therefore, that the observed variations in $\Delta S^{\ddagger}$ primarily reflect experimental error we ought to expect a plot of $\Delta H^{\ddagger}$ against $\Delta S^{\ddagger}$ to have a slope of $\overline{T}$.

All of this suggests that the excellent correlation often observed between $\Delta S^{\ddagger}$ and $\Delta H^{\ddagger}$ mainly reflects the fact that both thermodynamic parameters are in reality two measures of the same thing, and that measuring a com-
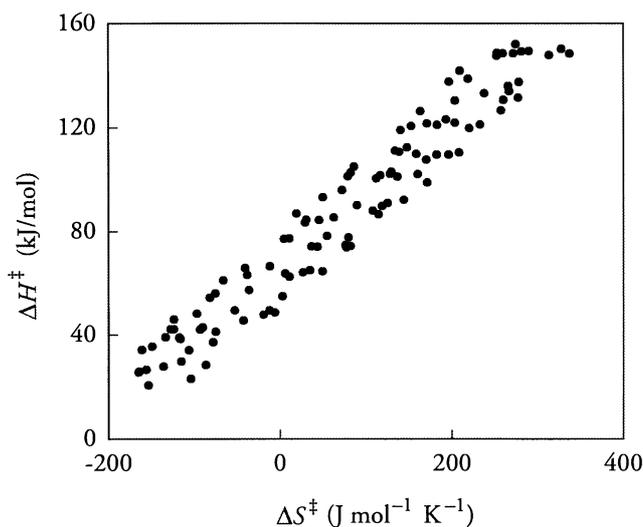
pensation temperature is just a rather indirect way of measuring the average temperature at which the experiments were carried out. As this temperature will often be in a range that the experimenter expects to have some biological significance, it is not surprising if the compensation temperature turns out to have a biologically suggestive value.

4. Non-artefactual aspects

It may be argued that even though the artefactual component in entropy–enthalpy compensation is very large it does not account for 100% of the effect. After all, the $\ln k_{18°}$ values in figure 2a are not distributed over an infinite range but confined to one decade in the values of $k_{18°}$. Thus the almost perfect correlation between $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ values can be interpreted as evidence that all of the enzymes have k values within an order of magnitude of one another at 18°C: this is a biological statement, not a statistical one. However, although the correlation deteriorates when the range of $\ln k_{18°}$ values is increased, it deteriorates very slowly, and is still quite good by the standards of most biological correlations when the $k_{18°}$ values are spread over six orders of magnitude (figure 4), and could still be described as excellent for three orders of magnitude (not shown). It follows, therefore, that even a very good correlation between $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ values tells us no more than that the k values do not differ from one another by more than a factor of 1000 at 18°C. To say this in the real world, however, is to say nothing, because

Figure 3. Origin of the correlation between entropy and enthalpy of activation. If the slope of an Arrhenius plot is decreased by an error e while maintaining the ordinate value constant at some temperature $\overline{T}$ (typically a temperature near the mean of those considered experimentally) then the rotation of the line will produce an error $e/\overline{T}$ in the intercept on the ordinate.

Figure 4. Enthalpy–entropy compensation with a million-fold range of $k_{18°}$ values. The points were obtained exactly as in figure 2b, except that the range of $\ln k_{18°}$ was $2\cdot3 \pm 6\cdot9$, corresponding to six orders of magnitude in the range of $k_{18°}$.

it is unlikely that an experimenter would select for study (or even detect) an enzyme with activity less than 0·1% of that of the more active enzymes considered.

Despite all this, comparison of figure 2b with figure 1 of Johnson and Goldspink (1975) or its representation as figure 7·5 of Gutfreund (1995) may lead to the observation that the simulated data do shown some deviations from a straight line, whereas the actual data for myofibrillar ATPase were, as far as the unaided eye can detect, perfect. This is probably mainly a small-sample effect, as a plot of seven well spaced but otherwise randomly selected points from the 100 in figure 2b yield a plot in which almost no deviation from linearity is evident (see the inset in figure 2b).

In fact, the values given by Johnson and Goldspink (1975) for V_{max} are spread over only a 2·3-fold at 18°C (assuming that the value of 0·045 μmol mg^{-1} min^{-1} given for *Tipalia nigra* is a typographical error for 0·45 μmol mg^{-1} min^{-1}), not a ten-fold range. As they did not record the molecular masses of the different enzymes one cannot convert this to a range of $k_{18°}$, but it is not unreasonable to guess that the seven enzymes had similar molecular masses, in which case the $k_{18°}$ values also spanned about a 2·3-fold range, and one could perhaps persuade oneself that the compensation plot demonstrated this.

The question then arises as to what biological interpretation one could put on an exact compensation at a particular temperature if such a compensation truly existed. Curiously, this point appears not to have been considered in any depth. What possible biological advantage would it be to *Tipalia grahami*, which spends its life in tropical hot springs at 25–38°C, and to *Notothenia rossii*, which lives in Antarctic waters at 0–2°C, to have the same myofibrillar ATPase activity at 18°C, a temperature that neither fish ever experiences in normal circumstances? What conceivable selective pressure, therefore, could have produced the virtually perfect compensation that the thermodynamic data suggest to the Panglossian observer? It seems best, therefore, not to be too assiduous in trying to explain any residual correlation that remains after statistical artefacts have been allowed for.

The story, however, does not end here, because even on the Panglossian interpretation the data do not show that all of the fishes have optimized myofibrillar ATPase activities. Arrhenius plots over a wider range of temperatures than 0–18°C show a discontinuity at 18·5°C (Johnson and Goldspink 1975), and it was "to overcome [this and other] complications" that only data in the lower range were used for the thermodynamic comparison. One may rephrase the above question, therefore: of what possible interest it can be to *T. grahami* if its myofibrillar ATPase has any particular activity when extracted from the fish and cooled to the other side of a discontinuity? And if the modern fish does not care, why should its

ancestors have been favoured by evolution if their myofibrillar ATPases tended towards a particular activity at 18°C?

5. Discussion

Gutfreund and Chock (1991) remarked in another context that "this is not Voltaire's best of all possible worlds, it is the best world we have got." This is true, of course, and it should not be forgotten, but it should not blind us to the existence of genuine optimization principles in biochemistry, several of which have been carefully documented by Meléndez-Hevia (1993). Even in temperature dependence studies there may be genuinely interesting (albeit imperfect) correlations that can easily be overlooked in an excessive zeal to find perfect but meaningless compensation effects. Consider, for example, the Arrhenius activation energies and the temperatures at which the various fishes live that are given in table 1 of Johnson and Goldspink (1975). Although not plotted by these authors, these values do appear to be genuinely correlated, as may be seen in figure 5. One can imagine various possible explanations, not necessarily directly related to evolution: for example, an increased activation energy might follow naturally from having a protein structure stable at a higher temperature. But that is not the main point here, which is that genuine but imperfect correlations are biologically more interesting than meaningless perfect ones.

Recently Sharp (2001) made a more detailed analysis of several supposed cases of enthalpy–entropy compen-
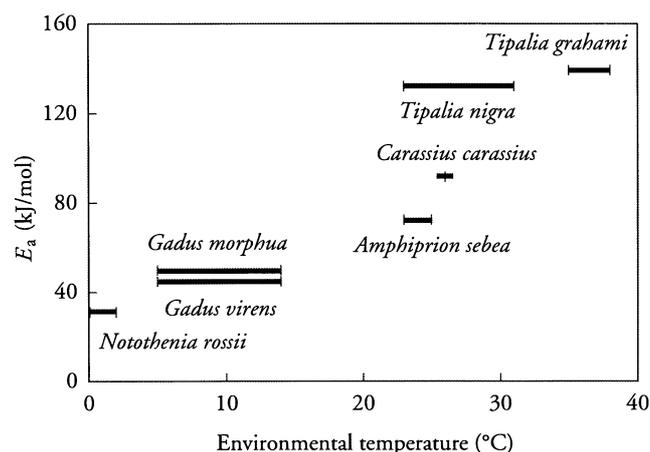
Figure 5. Correlation of activation energy with environmental temperature for seven fishes. The plot shows data for myofibrillar ATPase from table 1 of Johnson and Goldspink (1975). In the case of *Carassius carassius* (a domestic fish) the width of the bar is arbitrary, the midpoint being the temperature of domestication.

sation in the recent literature. These included one derived from calorimetric data (Privalov and Gill 1988), in principle less susceptible to statistical artefacts than the more usual deduction of compensation from temperature-dependence measurements. He concluded that all of the examples examined, including the one based on calorimetric data, could be "shown to be better explained by other causes". Unfortunately the history of the past half-century suggests that this will need to be shown again and again until it finally becomes accepted that compensation, if it is still postulated, needs far stronger evidence than is normally adduced.

References

Cornish-Bowden A 1995 *Fundamentals of enzyme kinetics* 2nd edition (London: Portland Press) pp 14–16

Exner O 1964 On the enthalpy–entropy relationship; *Coll. Czech. Chem. Comm.* **26** 1094–1113

Gutfreund H 1995 *Kinetics for the life sciences* (Cambridge: Cambridge University Press) pp 246–248

Gutfreund H and Chock P B 1991 Substrate channeling among glycolytic enzymes: fact or fiction?; *J. Theor. Biol.* **152** 117–121

Hochachka P W and Somero G N 1984 *Biochemical adaptation* (Princeton: Princeton University Press)

Johnson I A and Goldspink G 1975 Thermodynamic activation parameters of fish myofibrillar ATPase enzyme and evolutionary adaptations to temperature; *Nature* (*London*) **257** 620–622

Meléndez-Hevia E 1993 *La evolución del metabolismo: hacia la simplicidad* (Madrid: Eudema)

Privalov P L and Gill S J 1988 Stability of protein structure and hydrophobic interaction; *Adv. Prot. Chem.* **39** 191–234

Sharp K 2001 Entropy–enthalpy compensation: fact or artifact?; *Prot. Sci.* **10** 661–667