SCALING OF ENERGY METABOLISM IN UNICELLULAR ORGANISMS: A RE-ANALYSIS*

JOHN PROTHERO
Department of Biological Structure, University of Washington, School of Medicine, Seattle, Washington 98195, U.S.A. Telephone: (206) 543-1860

(Received 21 May 1985)

Abstract—1. The database used by Hemmingsen (1960) to compute energy metabolism in unicellular organisms was reassembled and submitted to linear (log-log) analysis. As Hemmingsen noted, this data set includes marine zygotes, which are not unicellular organisms.

2. If no temperature correction factors are applied to the data the best-fit regression line has a slope of 0.698 ± 0.024.

3. Application of the temperature correction factors assumed to have been used by Hemmingsen gave a slope of 0.756 ± 0.021, identical to the value he reported. The correlation coefficient is 0.97. The mean scatter about the regression line exceeds 100%.

4. A revised set of temperature correction factors gave a slope of 0.730 ± 0.021, suggesting that the value of almost exactly three-quarters obtained by Hemmingsen was probably fortuitous.

5. The slope of the best-fit regression line is very sensitive to the inclusion of bacteria and flagellates. When the data points for these organisms are omitted from the calculation the slope decreases to 0.645 ± 0.045.

6. When the data points for bacteria, flagellates and marine zygotes are omitted, the slope drops to 0.608 ± 0.025. The correlation coefficient (0.97), compared to the best-fit line reported by Hemmingsen, is unaffected; the mean deviation about the regression line drops to 40% and the points are evenly distributed about the regression line.

7. Because of the small number of species for which measurements have been made, the existing database relating energy metabolism to cell size is not representative of unicellular organisms generally.

8. It is concluded that the case for a three-quarters power rule expressing energy metabolism as a function of size in unicellular organisms generally is not at all persuasive.

INTRODUCTION

There is general agreement that energy metabolism is the fundamental global measure of vital activity in all organisms. It has become widely accepted, following the work of Kleiber (1947) and Brody (1945), that standard energy metabolism in mammals varies as the three-quarters power of body weight. In two papers published in 1950 and 1960 Hemmingsen argued that the three-quarters power rule also applies to poikilotherms and unicellular organisms. As a result of these works, the three-quarters power rule has acquired the status of one of the more remarkable quantitative generalizations in comparative biology. Recent studies (Bartels, 1982; Prothero, 1984) have called into question the validity of this power law, as applied to standard energy metabolism in mammals. Both of these papers suggest that an empirically correct value for the slope of the regression line relating energy metabolism to body weight in mammals will be closer to two-thirds than three-quarters. In this paper I report the results of a careful analysis of the database utilized by Hemmingsen in his study of energy metabolism in unicellular organisms.
be fitted to the plot, even after the application of different trial values for various conversion factors, then it was rejected. Only after the database was reconstructed in this fashion were any regression calculations undertaken.

The interested reader can obtain a copy of the database and the computations at small cost by writing to the Documentation Center.

**Data analysis**

The data were expressed in logarithmic form and submitted to linear least squares regression analysis (Edwards, 1976). For each regression analysis a measure of the scatter about the regression line was computed in the form of a mean percent deviation (% dev.). Thus, the model for the data analysis and the mean percent deviation are given by:

\[
y' = ax^b
\]

\[
\% \text{ dev} = 100/N \sum_{i=1}^{N} |y_i - y'_i|/y_i
\]

where \((x_i, y_i)\) represent the coordinates of the \(i\)th datum point, \(y_i\) is the value of the ordinate predicted from the regression equation for the value \(x_i\), and \(N\) is the number of points. In addition to the mean percent deviation, the correlation coefficient and the total number of points lying above and below each regression line are reported.

It is convenient to adopt a measure of the weight range covered by scaling data. I will employ the notation \(pWR\) to denote the logarithm of the weight range, where \(pWR\) is defined by:

\[
pWR = \log_{10}(\text{maximum weight}/\text{minimum weight})
\]

It is also useful in scaling studies to have measures of taxonomic diversity. In this study I have identified the kingdoms from which the unicellular organisms are drawn for each regression line, as well as the number of species represented.

**ANALYTICAL RESULTS**

The results of the regression analyses are summarized in Table 1. The measures of taxonomic diversity are given in columns two and three, whereas the regression characteristics per se are given in columns five to eleven. The nature of the data set employed in each specific regression calculation is indicated in column twelve. For example, in the second row the results are presented for the case in which no temperature corrections were applied to the data. The large percentage deviation and the non-uniform distribution of the points about the regression lines (see Table 1, columns ten and eleven) should be noted. Because of the large scatter, I have preferred to enclose the points by a minimal convex polygon (see Fig. 1), without showing a regression line (see below).

In the next several sections, the details of the individual regression calculations are discussed.

**Temperature corrections**

Hemmingsen corrected the measurements of energy metabolism to 20 C. apparently on the basis of the plant data of Kuijper (1910). I have attempted to derive as nearly as possible the temperature correction factors (tcf) which Hemmingsen employed by Hemmingsen in addition, I have derived a revised set of tcf for the same data. The details of these computations have been deposited with the Documentation Center.

We see from Table 1 (row 2) that the slope of the regression line fitted to the raw data (uncorrected for temperature) is 0.698 ± 0.024 (i.e. the slope significantly different from two-thirds). On the other hand, the slope of the regression line derived from the data corrected for temperature using the tcf assumed to have been used by Hemmingsen is 0.756 ± 0.021 (see Table 1, row 3). Hemmingsen reported an identical slope of 0.756 ± 0.021, an essentially identical intercept (log \(a\) (Hemmingsen, 1960).

Using a revised set of tcf, the slope is found to be 0.730 ± 0.021. Two points should be made. Firstly, it is evident that the temperature correction has a significant influence on the slope of the regression relating energy metabolism to body weight. Depending on the tcf, the slope varies by about 0.08. Secondly, the fact that Hemmingsen found a slope of 0.756 (i.e. almost exactly three-quarters) is regarded as fortuitous.

Whether it is at all reasonable to use data exclusively from plants to calculate tcf for diverse organisms, as Hemmingsen seems to have done, is questionable. It is true, however, that the more appropriate set of tcf probably would have yielded a slope of about 0.756, as the results given in Table 1 (row 4) very strongly support.

**Extreme values**

It is well known that, where the data are distributed, the slope of a regression line may be sensitive to the values at the extremes. That is the case here may be seen by examining the results given in Table 1 (row 5, Table 1). This effect cannot be attributed to a reduction in the number of points, since omission of the amoebae, which are at the other end of the range, has no effect on the slope (see Table 1, row 5).

In this case, the number of points is reduced by 14 points (19% of the total) corresponding to bacteria and flagellates. The slope then becomes 0.730 ± 0.021 to 0.645 ± 0.045 (compare row 5, Table 1). This effect cannot be attributed to a reduction in the number of points, since omission of the amoebae, which are at the other end of the range, has no effect on the slope (see Table 1, row 5).

Whether it is at all reasonable to use data exclusively from plants to calculate tcf for diverse organisms, as Hemmingsen seems to have done, is questionable. It is true, however, that the more appropriate set of tcf probably would have yielded a slope of about 0.756, as the results given in Table 1 (row 4) very strongly support.

**Sample size**

When we consider sample size in a scaling study, we are concerned with both species diversity and range. We can easily satisfy ourselves that the sample size employed by Hemmingsen is quite representative of the weight range exhibited by unicellular organisms. The smallest known unicellular organism is the rickettsia, with a minimum volume of about 1 cubic micra. The largest known unicellular organism is the amoeba, with volumes of about 15 cubic micra (see Poindexter, 1971; Stanie-
1970). Thus, pWR for unicellular organisms is about 10. The value of pWR for the data used by Hemmingsen is 8.6 (see Table 1, column five), or 86% of the total, on a logarithmic scale.

On the other hand, the diversity of unicellular species in the sample available to Hemmingsen (or current workers) is altogether unrepresentative of the variety of species of unicellular organisms thought to exist (see Table 2). In comparing the size of the sample used by Hemmingsen with all species of unicellular organisms, I have not included the marine zygotes, since these are not unicellular organisms. It is fair to note that reliable energy metabolism data are available for only a few species of unicellular organisms. With the exception of specialized cells, such as red blood cells, there has been very little interest in cellular energy metabolism and cell size during the past quarter century. Thus, the potential database for this kind of study has not increased significantly since Hemmingsen carried out his analysis. The results of a recent multivariate analysis of energy metabolism in unicellular organisms are not strictly comparable to the results obtained in the bivariate analysis reported here (Robinson et al., 1983).

Subsets of the data

Thirty-three of the data points, or 44% of the total, come from a single laboratory, that of Scholander et al. (1952). These workers fitted a regression line to their data, for which they reported a slope of 0.5. This compares well with the slope of 0.545 obtained here [note that these data were obtained from a graph (Scholander et al., 1952) by digitization (see Mannar et al., 1982)]. The mean percent deviation about the regression line is 25% (see Table 1, row 7).

It has already been observed that the bacteria and flagellates exert a disproportionate influence on the slope of the regression line relating energy metabolism to cell weight. It is also the case that zygotes (small embryos) cannot be regarded as unicellular organisms. It is, therefore, of interest to calculate a regression line with these data omitted. The results of this analysis are given in Table 1 (row 8). The slope of the regression line is 0.608 ± 0.025, and the mean deviation around the line is 40% (see Fig. 2). The data points are not very uniformly distributed along the weight range. In particular, the points corresponding to fungi are off by themselves at the lower end of the weight range. However, when we omit the two points corresponding to fungi, the regression characteristics are essentially unchanged (see Table row 9).

Aggregated data

An obvious objection to the validity of the regression lines fitted to the data by Hemmingsen is the large mean deviation (116–146%). It is often useful to...
Scaling of energy metabolism in unicellular organisms

Fungi
Amoebae
Ciliates

slope = 0.608

Fig. 2. Energy metabolism as a function of cell size in fungi, amoebae and ciliates (see Table 1, row 8).

such cases to divide the data into small groups, where each group extends over a narrow weight range. An average is then calculated for each group and a regression line is fitted to the averages. Often the effect of such aggregation will be to smooth out the fluctuations in the data. However, when this experiment was carried out with the data set of Hemmingsen, using logarithmic averages, there was no significant effect on the mean deviation of the aggregated points from the best-fit line (see Table 1, row 10). A possible implication is that the scatter reflects a qualitative dissimilarity in energy metabolism in these different forms.

Substituted data
Zygotes and small embryos are not unicellular organisms, as Hemmingsen noted (Hemmingsen, 1960). If one wished to include data for germ cells, it would be more consistent with the purposes of the study to use the data for unfertilized marine eggs, which even though haploid (in effect like bacteria) are at least unicellular. In the majority of the cases cited by Hemmingsen, the effect of fertilization is to increase the rate of oxygen consumption. It was possible to obtain data on energy metabolism in unfertilized eggs in eleven of the fourteen cases for which data were available for marine zygotes. The regression characteristics for this instance are given in Table 1 (row 11). The slope of the regression line is, unsurprisingly, essentially unchanged, since the points are close to the center of the regression line. There is, however, a large increase in the scatter; to 189%. It was because of this large scatter that Hemmingsen chose to use the data for marine zygotes (Hemmingsen, 1960).

DISCUSSION
It has been possible to reconstruct a data set which gives essentially identical regression parameters to those reported by Hemmingsen. Various features of this data set have been analyzed. It has been shown that the deviation between three-quarters and two-thirds in the slope of the regression line fitted to the data set employed by Hemmingsen is, in part, a function of the temperature correction factors applied. A more fastidious analysis of the temperature data employed by Hemmingsen suggests that 0.73 is a better value for the slope of the regression line than the value of 0.756 obtained by Hemmingsen. But the temperature correction factors are based on data drawn from the garden pea, yellow lupin and wheat. Whether it makes sense at all to apply these results to unicellular organisms as diverse as bacteria, flagellates, fungi, marine zygotes, ciliates and amoebae, as Hemmingsen seems to have done, is open to doubt.

A point which is perhaps not sufficiently appreciated in scaling studies is that the correlation coefficient, by itself, is an inadequate measure of whether a linear regression model is appropriate for a given set of log-log transformed data. A high correlation coefficient is perfectly compatible with large systematic deviations. In addition to the correlation coefficient, it is useful to look at the mean percent deviation about the regression line and the distribution of points about the line.

The large deviation—greater than 100%—associated with the regression lines fitted to the whole data set argues that energy metabolism in these organisms is not well described by a linear model. We know that bacteria lack mitochondria and other
membrane-bound organelles possessed by eukaryotes (Davis et al., 1980). Furthermore, at least some of the flagellates (e.g. trypanosomes) are thought not to ingest solid matter. Thus it is possible that energy metabolism in these forms, at least, is qualitatively dissimilar to that in other unicellular organisms. When we omit these forms, and zygoles, which are not unicellular organisms in any case, we find that the mean deviation drops to the more reasonable value of 40%. In this case (see Table 1, row 8), the correlation coefficient (0.97) is identical to the value for Hemmingsen’s data set (Table 1, rows 3 and 4) and the points are evenly distributed about the regression line. By these objective criteria the slope of the regression line which is least open to objections in this data set has a slope of 0.608 ± 0.025 (see Table 1, row 8). The fact that the mean deviation about the regression line does not decrease when we aggregate the data also argues that the complete data set is internally inconsistent. Hemmingsen omitted the data for fertilized insect eggs from his calculations, apparently on the grounds that they did not fit the unicellular line (Hemmingsen, 1960). No biological case is made for this omission. Inclusion of these data would have increased the slope, and more especially the scatter, considerably.

At the time the measurements were made which Hemmingsen draws upon, there was no method available to measure the oxygen concentration in the culture medium. In some cases, at least, the oxygen concentration may have been sufficiently reduced locally to adversely affect oxygen consumption. Many other difficulties with the data base are apparent. Measurements were made at different temperatures and under different conditions. Some of the measurements were made on single animals, whereas others were made on many. Some animals were starved for long periods (e.g., weeks) and others were not.

Hemmingsen noted (Hemmingsen, 1950) that it is difficult to define standard conditions for bacteria, but it seems likely that this remark applies to unicellular organisms generally, since some are photosynthetic and some not, some are motile and some not, some ingest solid materials and some do not. These limitations of the data set do not constitute a criticism of Hemmingsen, who brought together the best data available. The key question is whether the data, and his analysis of the data, supports his conclusions.

CONCLUSIONS

Hemmingsen’s suggestion that energy metabolism in unicellular organisms varies as the three-quarters power of cell size is both novel and bold. If the hypothesis ultimately proves to be correct, it will be of fundamental significance for the understanding of cellular energetics. But the available evidence and Hemmingsen’s analysis of that evidence in support of a three-quarters power rule are not persuasive. A more careful application of temperature correction factors leads to a slope of 0.73, suggesting that the value close to the simple fraction of three-quarters found by Hemmingsen was fortuitous. Omission of the data on bacteria and flagellates drops the slope to 0.645; but even then the scatter around the best-fit regression line is substantial. Omission of the data on marine zygotes leads to a slope of 0.608 ± 0.025, with a mean deviation of 40%.

There are reasonable grounds for believing that energy metabolism in bacteria may be qualitatively different from that in eukaryotes. If this is the case there is no good reason to expect that energy metabolism in unicellular organisms can be adequately represented by a single linear regression line. Until we have a better understanding of energy metabolism in unicellular organisms, it may be best to represent the aggregate body of information by a convex polygon. And until energy metabolism has been determined under standard conditions, assuming that is possible in a much wider variety of species, it is rather doubtful that any two-parameter representation of the data is warranted.

Acknowledgements—I am grateful to Dr. K. D. Jürgens for his constructive comments upon a draft of this paper, thank Ms. Doris Ringer for preparation of the manuscript, and this work was supported in part by a USPHS grant RO1-HL-29761 to J. Prothero.

REFERENCES