Eur Biophys J (1998) 27: 409–410

BIOPHYSICS LETTER

Cathleen N. Early · B. Mark Britt

Sequence similarities of glyceraldehyde-3-phosphate dehydrogenases, phosphoglycerate kinases, and pyruvate kinases are species optimal temperature-dependent

Received: 22 September 1997 / Revised version: 2 February 1998 / Accepted: 6 February 1998

Abstract Data are presented that suggest enzyme sequence similarities among species are not solely a function of their evolutionary relationship. It is demonstrated that sequence similarities of glyceraldehyde-3-phosphate dehydrogenases, phosphoglycerate kinases, and pyruvate kinases from yeast, bacteria, mammals and a bird possess a significant species optimal thriving temperature dependence that crosses through conventional phylogenetic divisions. It is therefore suggested that species which are distantly related evolutionarily may possess some degree of enzyme sequence similarity if they happen to thrive at near the same optimal temperature; conversely, organisms which are closely related evolutionarily but function at radically different temperatures will possess a sequence dissimilarity that may mask the close relatedness.

Key words Sequence similarity · Glyceraldehyde-3-phosphate dehydrogenase · Phosphoglycerate kinase · Pyruvate kinase · Evolution · Phylogeny

Introduction

The assumption in the use of enzyme sequence similarities to establish evolutionary relationships between organisms is that greatly dissimilar sequences reflect a distant evolutionary relationship while similar sequences demonstrate closely related species. The purpose of this communication is to report that the sequence similarities of glyceraldehyde-3-phosphate dehydrogenases, phosphoglycerate kinases, and pyruvate kinases from a diverse collection of organisms possess a significant species optimal temperature dependence and to suggest that such a dependence may be characteristic of all enzymes. It is suggested that the cellular physical environment is a major determinant of the dynamical tertiary structure necessary for the functional enzyme. That is, a stable yet conformationally flexible enzyme must be found for each set of cellular physical conditions (Britt, 1993, 1997; Castro and Britt, 1998). Accordingly, the cellular environment also determines the enzyme primary structure. So, instead of being a function solely of evolutionary relatedness, sequence similarity may also be a function of optimal thriving temperature, external pressure, and pH. Therefore, enzymes from different organisms which function optimally at radically different temperatures, for example, will possess a marked sequence divergence that may mask a close evolutionary relatedness. Conversely, enzymes from organisms which are obviously distantly related (*E. coli* and humans, for example) will possess a degree of sequence similarity for the simple reason that they share a similar optimal temperature.

C. N. Early · B. M. Britt (✉)
Department of Chemistry, P.O. Box 97348,
Baylor University, Waco, TX 76798, USA

Materials and methods

Figure 1 shows the species optimal temperature dependences of sequence similarities relative to *Saccharomyces cerevisiae*, the organism with the lowest optimal temperature in this study. Data were obtained from enzymes from the following species (listed with their optimal thriving temperatures): *Saccharomyces cerevisiae* (27.5 °C), *Bacillus megaterium* (30 °C), *Schizosaccharomyces pombe* (31 °C), *Bacillus subtilis* (35 °C), *Haemophilus influenzae* (36 °C), *Mus musculus* (mouse) (36.5 °C), *Homo sapiens* (37 °C), *Emericella nidulans* (37 °C), *Salmonella typhimurium* (37 °C), *Escherichia coli* (37 °C), *Rattus norvegicus* (rat) (37.3 °C), *Equus caballus* (horse) (37.7 °C), *Felis catus* (domestic cat) (38.6 °C), *Oryctolagus cuniculus* (rabbit), (38.8 °C), *Sus scrofa (pig)* (39.3 °C), *Gallus gallus* (chicken) (41.7 °C), *Bacillus licheniformis* (42.5 °C), *Lactobacillus delbrueckii* (45 °C), *Lactobacillus lactis* (45 °C), *Bacillus stearothermophilus* (62 °C), *Thermus aquaticus* (71 °C), *Thermotoga maritima* (75 °C), *Methanothermus fervidus* (83 °C), *Sulfolobus solfactaricus* (87 °C), *Pyrococcus woesei* (100 °C). Bacterial optimal temperature data were obtained primarily from *Bergey's Manual of Deter-*

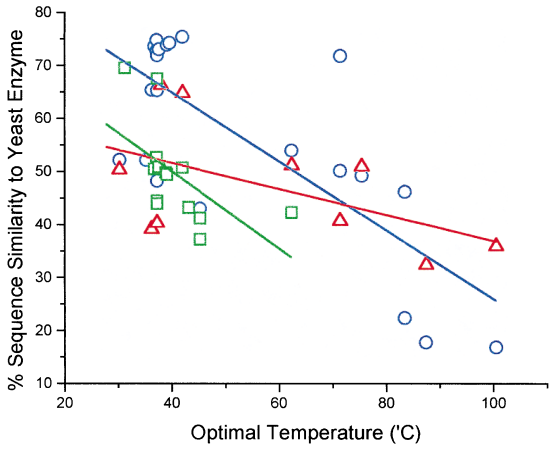

Fig. 1 Species optimal temperature dependences of sequence similarities of glyceraldehyde-3-phosphate dehydrogenases, phosphoglycerate kinases, and pyruvate kinases. Similarities in amino acid sequence are relative to *Saccharomyces cerevisiae*, the organism with the lowest optimal temperature in the study. Data are shown for glyceraldehyde-3-phosphate dehydrogenases (*blue circles*), phosphoglycerate kinases (*red triangles*) and pyruvate kinases (*olive squares*). Trends are indicated by linear fits to the data. The linear correlation coefficient, *r*, and the probability that there exists no linear correlation between percent sequence similarities and optimal thriving temperature for each enzyme are (Bevington 1969): glyceraldehyde-3-phosphate dehydrogenase (0.75, less than 1/1000); phosphoglycerated kinase (0.52, ~1/10); pyruvate kinase (0.57, less than 1/20)

minative Bacteriology (Holt et al. 1994), mammalian and avian temperature data are normal body temperatures (Rodbard, 1950), and yeast optimal temperatures are from Menendez-Arias and Argos (1989). Sequence similarities were calculated with the SWISS-PROT program SIM using the Blosum 30 comparison matrix and other parameters at their default values (Bairoch and Apweiler, 1996).

Results and discussion

The temperature dependence of sequence similarity of the enzymes diverges at an average rate of ~0.5%/°C. While there is a definite dependence upon optimal thriving temperature, it is significant that the trends observed are not strongly linearly dependent. Other factors in addition to temperature apparently are manifest in the divergences. These forces are likely to be pH and external pressure as well as classic evolutionary drift.

If the above trends are found to be characteristic of all enzymes it is then possible that traditional molecular phylogenetic approaches that correlate similar enzyme sequences with close evolutionary relationship may result in incorrect associations between species. For example, Hensel and coworkers constructed an evolutionary tree for a variety of organisms based on glyceraldehyde-3-phosphate dehydrogenase sequences which suggests that *E. coli* is more closely related to mammals than to the other eubacteria in the study (Hensel et al. 1989). The authors note that the association seems incongruous and others have suggested a gene transfer event from vertebrates to *E. coli* to explain the observation (Martin and Cerff, 1986). Interestingly, their evolutionary tree is composed of three branches which are loosely correlated with different optimal temperature ranges. As another example, Korn and co-workers note the striking similarities in human DNA polymerase α and replicative DNA polymerases from some phages, a yeast, a plasmid, and some viruses and speculate that these enzymes may have evolved from a common primordial gene (Wang et al. 1989). While this may be true, the data presented here suggests that the lack of divergence may be attributable to the common operating temperature of the enzymes (~37°C).

In conclusion, we reckon that the use of enzyme sequence homologies to establish evolutionary relationships will prove to be most fruitful when those substitutions which are present apparently only to maintain the enzyme in some other environment can somehow be 'subtracted out' to unveil the true evolutionary relationships between organisms.

Acknowledgements This work was supported by the Robert A. Welch Foundation, Baylor University, and the NSF Summer Undergraduate Research Program.

References

Bairoch A, Apweiler R (1996) The SWISS-PROT protein sequence data bank and its new supplement TREMBL. Nucleic Acids Res 24:21–25

Bevington PR (1969) Data reduction and error analysis for the physical sciences. McGraw-Hill, New York

Britt BM (1993) A shifting specificity model for enzyme catalysis. J Theor Biol 164:181–190

Britt BM (1997) For enzymes, bigger is better. Biophys Chem 69:63–70

Castro C, Britt BM (1998) Evidence for a low temperature transition state binding preference in bovine adenosine deaminase. Biophys Chem 70:87–92

Hensel R, Zwickl P, Fabry S, Lang J, Palm P (1989) Sequence comparison of glyceraldehyde-3-phosphate dehydrogenases from the three urkingdoms: evolutionary implication. Can J Microbiol 35:81–85

Holt JG, Krieg NR, Sneath PHA, Staley JT, Williams ST (1994) Bergey's manual of determinative bacteriology, 9th edn. Williams and Wilkins, Baltimore

Martin W, Cerff R (1986) Prokaryotic features of a nucleus-encoded enzyme: cDNA sequences for chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenases from mustard (*Sinapis alba*) Eur J Biochem 159:323–331

Menendez-Arias L, Argos P (1989) Engineering protein thermal stability: sequence statistics point to residue substitutions in α-helices. J Mol Biol 206:397–406

Rodbard S (1950) Weight, body temperature. Science 111:465–466

Wang T S-F, Wong SW, Korn D (1989) Human DNA polymerase α: predicted functional domains and relationships with viral DNA polymerases. FASEB J 3:14–21