Growth temperature and relative stabilities of proteins

Jeffrey M. Dick

April 5, 2011

Abstract

This vignette demonstrates an analysis of the relative stabilities of 24 carboxylases from model organisms that grow optimally at different temperatures. In the chemical thermodynamic model, changes in both temperature and oxidation potential influence the relative stabilities of the proteins with an overall progression that fits with the growth temperatures of the organisms. The calculations supporting this study are presented using a series of equilibrium activity diagrams constructed using CHNOSZ. First, equilibrium activity diagrams as a function of either temperature or $\log a_{H_2}$ (at constant values of the other variable) show that both variables contribute significantly to the relative stabilities of the proteins. Second, an equilibrium predominance diagram is used to formulate an expression for $\log a_{H_2}$ as a function of temperature that crosses the protein stability fields in order of increasing growth temperature of the organism. Third, equilibrium activity and chemical species richness diagrams based on this function show pattern that mirrors the natural growth temperature of the organism. The final piece of code produces an animated series of rank-abundance diagrams for the proteins as both temperature and $\log a_{H_2}$ are increased.

1 Introduction

The motivation for this set of calculations is to explore how differences in the chemical stabilities of proteins might underlie adaptation of organisms to a range of environments that vary in temperature and chemical characteristics. The approach we take is to perform chemical thermodynamic calculations of the relative stabilities of proteins from different organisms. These calculations are carried out using the CHNOSZ package [5] for the R software environment [21].

A collection of model organisms and proteins is listed in Table 1. The organisms listed in Table 1 were chosen because the represent a cross-section of mesophilic and thermophilic organisms. This list was inspired by recognizing the great diversity of organisms reported in Yellowstone hot springs (see references in the table), but it also contains organisms from soils, ocean and submarine hot spring environments. Where possible, literature reference that identify types of organisms found in Yellowstone National Park are listed, as are references giving the optimal growth temperatures of the model organisms.

Carboxylases were identified as the model proteins in part because they are found in a wide range organisms and their sequences are often available, even for those organisms which have not whole genomes available. Homologs of the enzyme ribulose biphosphate carboxylase (RuBisCo) have been shown to be useful in exploring evolutionary relationships of organisms that have this enzyme [22, 1, 26]. Homologs of both RuBisCo and acetyl-coenzyme A carboxylase (AcCoACo is the abbreviation used here) are listed in Table 1. The AcCoACo homologs were included in order to expand the range of organisms considered here.

For each protein family there are twelve organisms listed in Table 1, in approximate order of increasing optimal growth temperature. For the purposes of the calculations described below, the first six proteins in each group are classified as coming from low-temperature environments, and the last six as coming from high-temperature environments.

2 Code

Load CHNOSZ.

```
> library(CHNOSZ)
```

Table 1: Model organisms and identities of model carboxylases used in this study. The protein IDs are those used in the SWISS-PROT/UNIPROT database [3]; if the ID is a SWISS-PROT name the accession number is given in parentheses. Abbreviations: A – archaea; B – bacteria; E – eukarya; T_{opt} – optimal growth temperature.

A/B/E	Organism	T_{opt}, °C	ID	Formula	Length
	Ribulose bisphosphate carboxylase (large subunit except *)				
B	*α-proteobacterium* (*B. japonicum*)	25	RBL_BRAJA (Q9ZI34)	$C_{2403}H_{3735}N_{663}O_{702}S_{21}$	486
B	*β- proteobacterium* (*T. denitrificans*) [8]	28-32 [16]	A6YF84_9PROT (*)	$C_{1061}H_{1670}N_{312}O_{299}S_{11}$	213
E	*Bracteacoccus giganteus* [20]	–	A1E8R4_9CHLO	$C_{1325}H_{2091}N_{383}O_{379}S_{14}$	269
B	*Mycobacterium* sp. DSM 3803 [25]	–	A8C9T6_9MYCO	$C_{2344}H_{3614}N_{640}O_{690}S_{19}$	476
B	*Leptospirillum* sp. Group II UBA [25]	∼37	A3EQE1_9BACT	$C_{1901}H_{3060}N_{530}O_{543}S_{23}$	392
B	*Cyanobium* sp. PCC 7009 [20]	∼45	A5CKC7_9CHRO	$C_{2345}H_{3606}N_{646}O_{687}S_{19}$	470
B	*Cyanobacterium* (*Synechococcus*) [1]	50 [1]	RBL_SYNJA (Q2JV67)	$C_{2366}H_{3652}N_{650}O_{683}S_{21}$	474
E	*Cyanidiales* [8, 25]	50–55 [4]	Q6JAIO_9RHOD	$C_{2029}H_{3169}N_{551}O_{575}S_{16}$	405
A	*Methanococcus jannaschii*	85 [13]	RBL_METJA (Q58632) (*)	$C_{2151}H_{3412}N_{574}O_{627}S_{16}$	425
A	*Staphylothermus marinus* [19]	85–92 [24, 7]	A3DND9_STAMF	$C_{2264}H_{3529}N_{603}O_{633}S_{13}$	443
A	*Thermofilum pendens* [19]	85–90 [24]	A1RZJ5_THEPD	$C_{2215}H_{3468}N_{598}O_{630}S_{13}$	443
A	*Pyrococcus horokoshii* [19]	98 [10]	RBL_PYRHO (O58677) (*)	$C_{2179}H_{3409}N_{587}O_{615}S_{18}$	430
	Acetyl-coenzyme A carboxylase (* carboxyl transferase subunit alpha)				
B	*γ-proteobacterium* EBAC31A08 [9, 23]	–	Q9F7M8_PRB01	$C_{2222}H_{3505}N_{607}O_{675}S_{19}$	447
B	*Deinococcus radiodurans*	25–30 [18]	ACCA_DEIRA (Q9RV16) (*)	$C_{1494}H_{2419}N_{453}O_{455}S_{8}$	316
B	*Planctomyces maris* [9]	30 [2]	A6CDM2_9PLAN	$C_{1611}H_{2578}N_{436}O_{477}S_{13}$	322
B	*Actinobacterium* PHSC20C1 [25]	∼30 [**?**]	A4AGS7_9ACTN	$C_{3108}H_{4983}N_{889}O_{981}S_{19}$	675
B	*α-proteobacterium* (*C. crescentus*)	30	ACCA_CAUCR (Q9A448) (*)	$C_{1517}H_{2440}N_{450}O_{464}S_{11}$	320
B	*δ-proteobacterium* (*D. vulgaris*) [9]	37 [17]	A1VC70_DESVV	$C_{3761}H_{5876}N_{1084}O_{1085}S_{29}$	751
A	*Methanococcus maripaludis*	35–40 [14]	A6VIX9_METM7	$C_{2443}H_{3887}N_{659}O_{729}S_{17}$	494
B	*Cyanobacterium* (*Synechococcus*) [1]	50 [1]	Q2JSS7_SYNJA (*)	$C_{1608}H_{2590}N_{464}O_{478}S_{9}$	327
B	*Chloroflexus aggregans* [9]	55 [11]	A0GZU2_9CHLR	$C_{2831}H_{4515}N_{821}O_{827}S_{17}$	590
B	*Hydrogenobaculum* sp. Y04AAS1 [23]	58 [6]	A7WGI1_9AQUI	$C_{2360}H_{3732}N_{622}O_{702}S_{13}$	472
B	*Hydrogenobacter thermophilus* [23]	70–75 [15]	Q05KD0_HYDTH	$C_{2218}H_{3568}N_{612}O_{641}S_{19}$	444
B	*Aquifex aeolicus* [23]	85 [12]	ACCA_AQUAE (O67260) (*)	$C_{1635}H_{2616}N_{440}O_{457}S_{14}$	323

Put the names of the model proteins in the variables `rubisco` and `accoaco`:

```
> rubisco <- c("RBL_BRAJA", "A6YF84_9PROT", "A1E8R4_9CHLO", "A8C9T6_9MYCO",
     "A3EQE1_9BACT", "A5CKC7_9CHRO", "RBL_SYNJA", "Q6JAI0_9RHOD",
     "RBL_METJA", "A3DND9_STAMF", "A1RZJ5_THEPD", "RBL_PYRHO")
> rubisco.organisms <- c("a-proteobacterium-R", "b-proteobacterium",
     "Bracteacoccus", "Mycobacterium", "Leptospirillum", "Cyanobium",
     "Synechococcus", "Cyanidiales", "Methanococcus-R", "Desulfurococcus",
     "Thermofilum", "Pyrococcus")
> accoaco <- c("Q9F7M8_PRB01", "ACCA_DEIRA", "A6CDM2_9PLAN", "A4AGS7_9ACTN",
     "ACCA_CAUCR", "A1VC70_DESVV", "A6VIX9_METM7", "Q2JSS7_SYNJA",
     "A0GZU2_9CHLR", "A7WGI1_9AQUI", "Q05KD0_HYDTH", "ACCA_AQUAE")
> accoaco.organisms <- c("g-proteobacterium", "Deinococcus", "Planctomyces",
     "Actinobacterium", "a-proteobacterium-A", "d-proteobacterium",
     "Methanococcus-A", "Synechococcus", "Chloroflexus", "Hydrogenobaculum",
     "Hydrogenobacter", "Aquifex")
```

Set up the basis species. The activities of basis species (except for $\log a_{\mathrm{H_2}}$, which is specified in the calculations below) are set to $\log a_{\mathrm{H_2O}} = 0$, $\log a_{\mathrm{CO_2}} = -3$, $\log a_{\mathrm{NH_3}} = -4$, $\log a_{\mathrm{H_2S}} = -7$ and pH = 7.

```
> basis(c("CO2", "H2O", "NH3", "H2", "H2S", "H+"), c("aq", "liq",
     "aq", "aq", "aq", "aq"), c(-3, 0, -4, -6, -7, -7))
```

```
basis: changed basis to CO2 H2O NH3 H2 H2S H+.
    C H N O S Z ispecies logact state
CO2 1 0 0 2 0 0       69     -3    aq
H2O 0 2 0 1 0 0        1      0   liq
NH3 0 3 1 0 0 0       68     -4    aq
H2  0 2 0 0 0 0       66     -6    aq
H2S 0 2 0 0 1 0       70     -7    aq
H+  0 1 0 0 0 1        3     -7    aq
```

Define a function for $\log a_{\mathrm{H_2}}$ as a function of temperature. This is the same function used in Ref. [**?**].

```
> get.logaH2 <- function(T) return(-11 + T * 3/40)
```

Write a function, named `plot.it`, to make the equilibrium activity diagrams, making use of the `affinity` and `diagram` functions in CHNOSZ. The function contains additional code for coloring the lines, adding legends to plots, and shading a rectangular area of the plot. `plot.it` also returns the result of the `diagram` function (the calculated equilibrium activities of the proteins) for use in further calculations. Instead of printing out the entire function here, only the argument list is shown.

```
> print(args(plot.it))
```

```
function (vars = "H2", lims = c(-15, 0), do.plot = TRUE, dolegend = NULL,
    ylim = c(-6, -2), rect = NULL, drawme = c("rubisco", "accoaco"))
NULL
```

Define the limits of temperature and $\log a_{\mathrm{H_2}}$ that we will use on the plots.

```
> Tlim <- c(0, 140)
> H2lim <- c(-10, -2)
> Tval <- seq(Tlim[1], Tlim[2], length.out = 128)
> H2val <- seq(H2lim[1], H2lim[2], length.out = 128)
```

Identify the low- and high-temperature model organisms.

```
> ilowT <- c(1:6, 13:18)
> ihighT <- c(7:12, 19:24)
```

Define the chemical activity cutoff for counting chemical species richness.

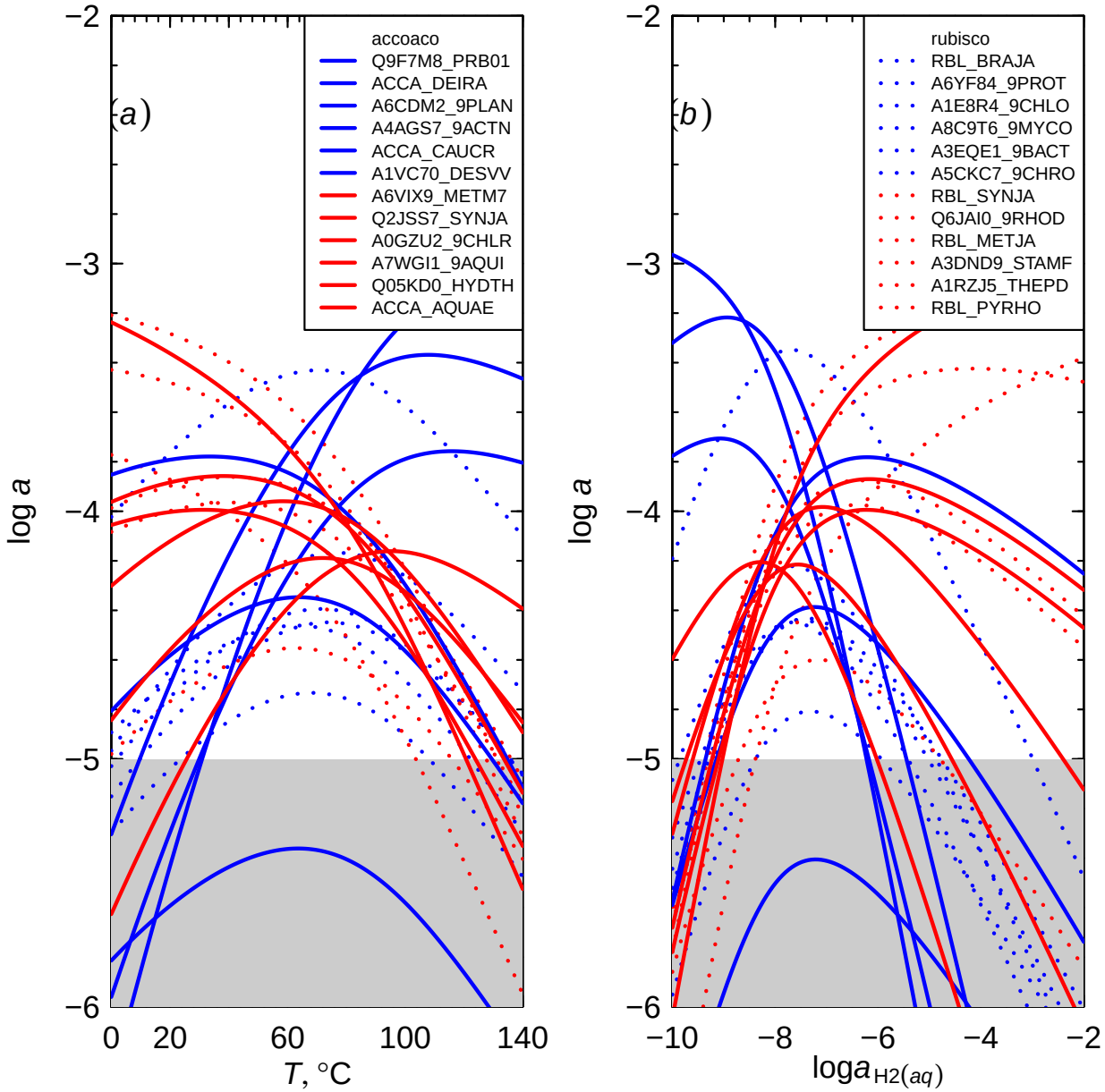
```
> loga.ref <- -5
```

3 Results

3.1 Changing a single variable

Fig. 3.1a shows the metastable equilibrium activities of the proteins as a function of temperature at $\log a_{H_2} = -6$, and Fig. 3.1b shows the metastable equilibrium activities of the proteins as a function of $\log a_{H_2}$ at $T = 25$ °C. The calculations were carried out for unit total activity of amino acid residues.

```
> par(mfrow = c(1, 2))
> par(cex = 1, lwd = 1, mar = c(2.5, 3, 1, 1))
> Tout <- plot.it("T", Tlim, dolegend = "accoaco", rect = c(Tlim[1],
      -6, Tlim[2], loga.ref))
> label.plot("a", xfrac = 0.1)
> H2out <- plot.it("H2", H2lim, dolegend = "rubisco", rect = c(H2lim[1],
      -6, H2lim[2], loga.ref))
> label.plot("b", xfrac = 0.1)
```

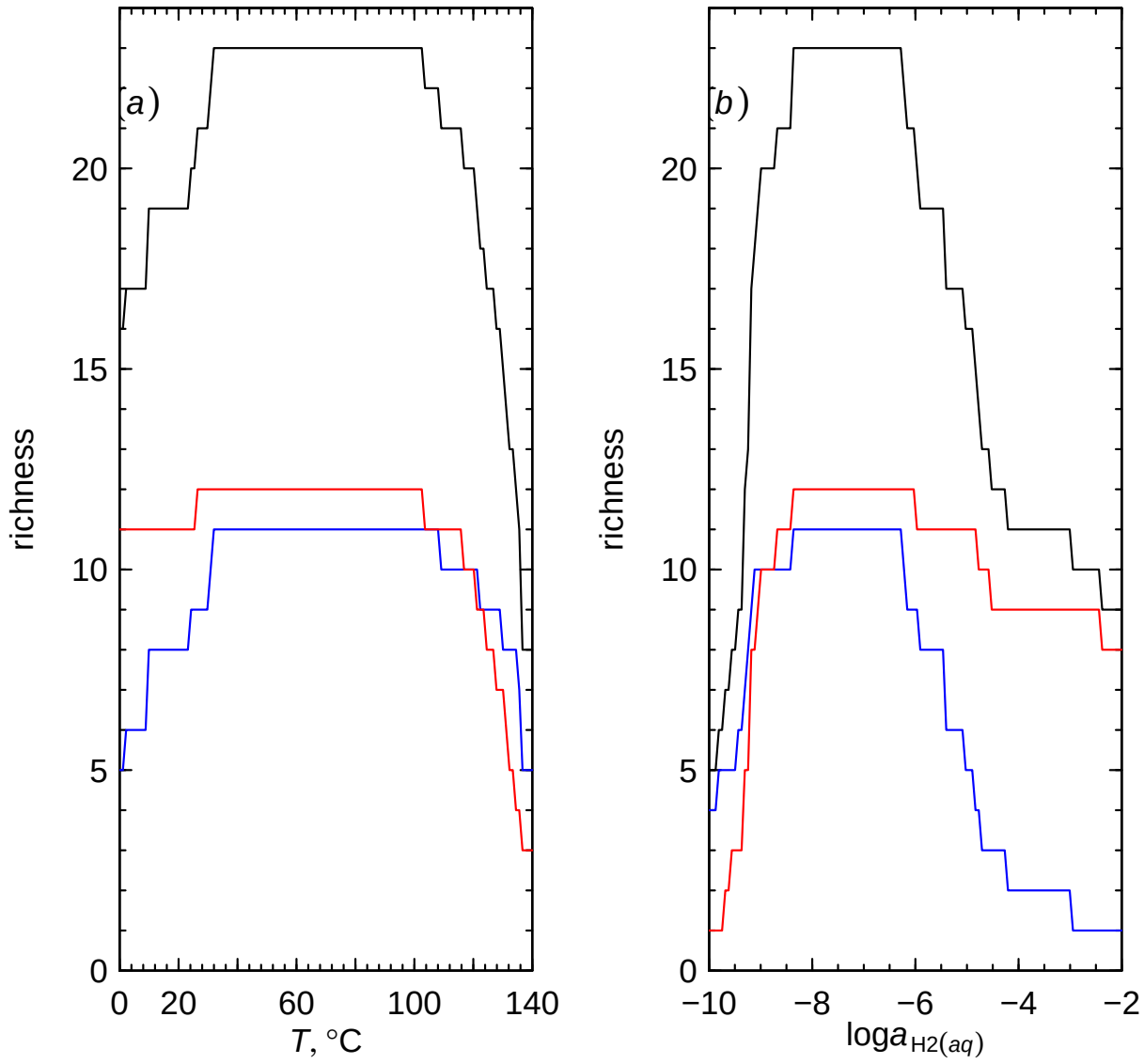


The figures each have all 24 proteins. The legend is split between the two panes, with AcCoACo listed on the left (solid lines) and RuBisCO on the right (dotted lines). The colors of the lines indicate the growth temperatures of the organisms taken from Table 1. The two figures are approximately mirror images of each other; increasing $\log a_{H_2}$ at constant temperature (right) has a similar effect as decreasing temperature at constant $\log a_{H_2}$ (left). According to the figures, either of these changes tends to stabilize the proteins from the organisms adapted to growth at high temperature; that is, there are more red lines higher on the diagram at either low temperature or high $\log a_{H_2}$.

3.2 Counting species

If we count the numbers of red and blue lines that are above the shaded area in Fig. 3.1, we get the diagrams shown in Fig. 3.2. These figures, constructed using the `revisit` function in CHNOSZ, show predicted richness, i.e. counts of chemical species whose activities are above a certain value (−3.4 in this case). The red line, which indicates the sum of proteins from organisms adapted to growth at high temperature, moves to higher values at low temperature and/or high $\log a_{H_2}$.

```
> par(mfrow = c(1, 2))
> par(cex = 1, lwd = 1, mar = c(2.5, 3, 1, 1))
> revisit(Tout$d, "richness", loga.ref = loga.ref, yline = 2, plot.ext = FALSE)
> revisit(Tout$d, "richness", loga.ref = loga.ref, ispecies = ilowT,
     col = "blue", add = TRUE, plot.ext = FALSE)
> revisit(Tout$d, "richness", loga.ref = loga.ref, ispecies = ihighT,
     col = "red", add = TRUE, plot.ext = FALSE)
> label.plot("a", xfrac = 0.1)
> revisit(H2out$d, "richness", loga.ref = loga.ref, yline = 2,
     plot.ext = FALSE)
> revisit(H2out$d, "richness", loga.ref = loga.ref, ispecies = ilowT,
     col = "blue", add = TRUE, plot.ext = FALSE)
> revisit(H2out$d, "richness", loga.ref = loga.ref, ispecies = ihighT,
     col = "red", add = TRUE, plot.ext = FALSE)
> label.plot("b", xfrac = 0.1)
```



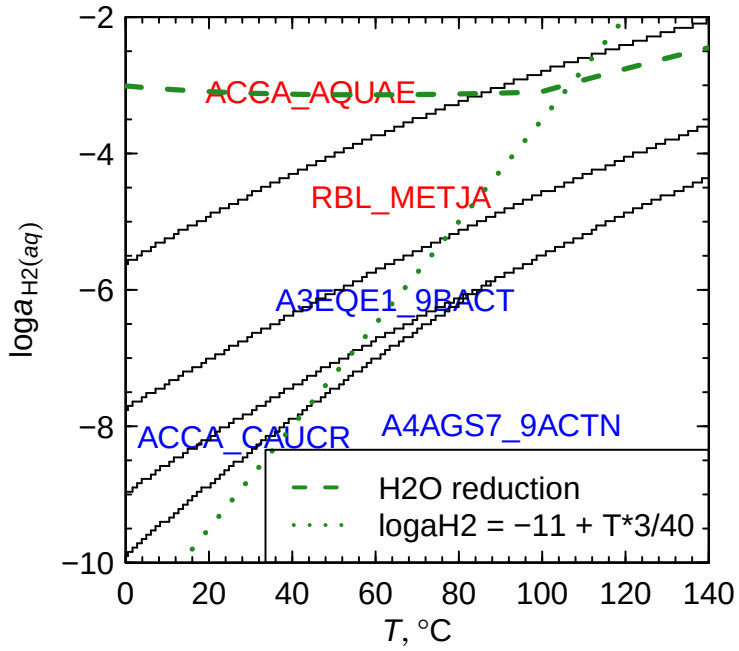
If you're like me, you wonder why low temperatures apparently favor formation of the proteins from no-so-hot organisms. Conversely, high $\log a_{H_2}$ apparently favors the hot organisms, which might make sense if hot environments are also relatively reducing.

3.3 As a function of two variables

Let's look at what happens when we vary temperature and $\log a_{H_2}$ independently. The following diagram is an equilibrium predominance diagram with all 24 proteins as candidates; only a few most stable proteins appear (they correspond to the highest lines in Fig. 3.1).

```
> par(cex = 1.5, lwd = 1.5)
> logaH2.line <- get.logaH2(Tval)
> a <- affinity(T = Tlim, H2 = H2lim)
> col <- rep(c(rep("blue", 6), rep("red", 6)), 2)
> diagram(a, residue = TRUE, col.names = col, color = NULL)
```

```
> s <- subcrt(c("H2", "H2"), c("gas", "aq"), c(-1, 1), T = Tval)
> logaH2 <- s$out$logK + log10(s$out$P)
> lines(Tval, logaH2, lwd = 4, lty = 2, col = "forestgreen")
> lines(Tval, logaH2.line, lwd = 4, lty = 3, col = "forestgreen")
> legend("bottomright", legend = c("H2O reduction", "logaH2 = -11 + T*3/40"),
    lty = c(2, 3), lwd = 3, col = "forestgreen")
```



The color code used for the protein names is the same as above: red for organisms from higher temperatures (the last six organisms in each part of Table 1) and blue for the others. The dashed green line indicates the reduction stability limit of H_2O, i.e. the value of $\log a_{H_2}$ taken from

$$H_{2(g)} \rightleftarrows H_{2(aq)}\,,$$

using $\log f_{H_{2(g)}} = 1$ below 100 °C and $\log f_{H_{2(g)}} = P_{\rm sat}$, or pressure of vapor-liquid equilibrium of H_2O, at or above 100 °C. The dotted green line is a hypothetical trajectory of $\log a_{H_2}$ given by

$$\log a_{H_2} = -11 + T * 3/40\,.$$

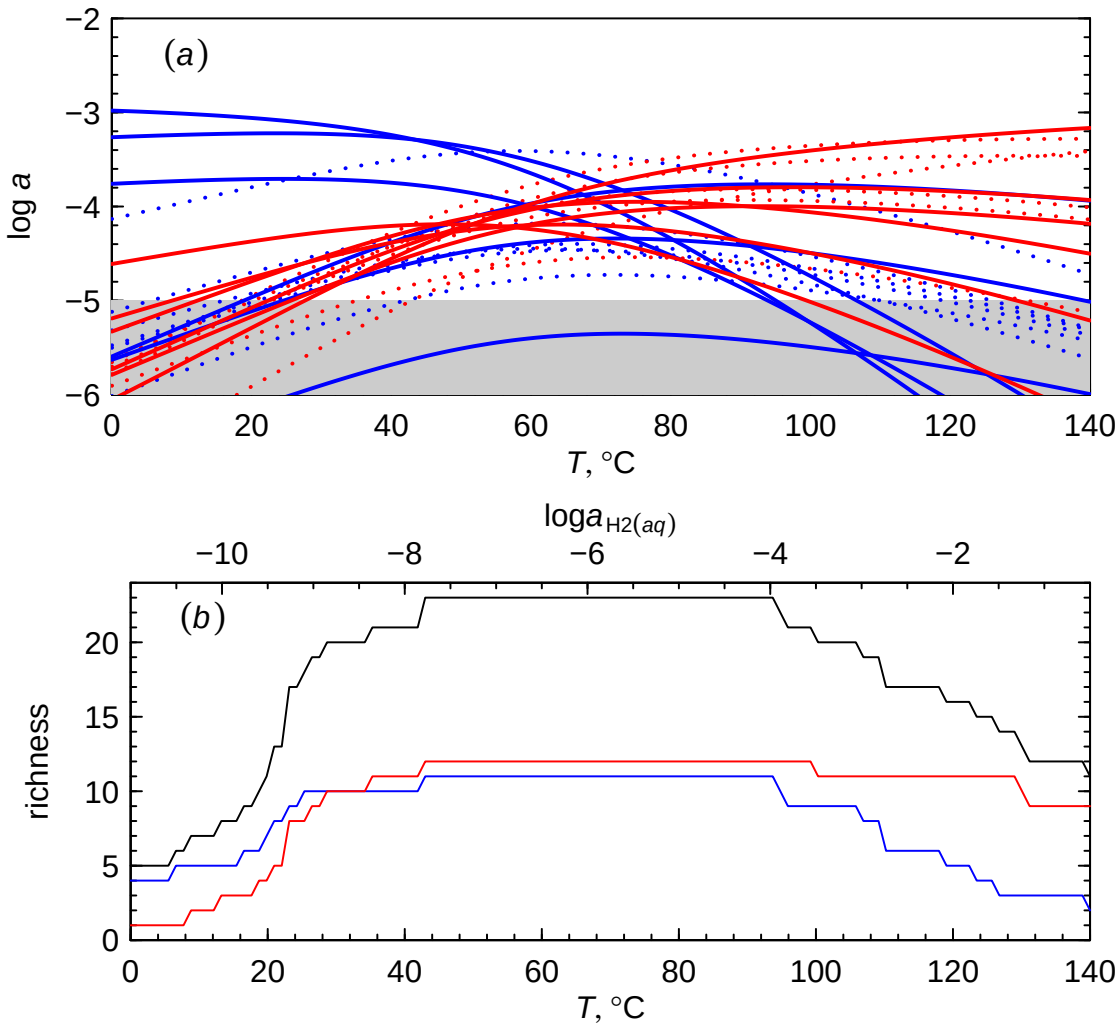
This equation is simply a proposal for a line cutting across the predominance diagram that lands in the low-temperature-protein fields at low temperatures and the high-temperature-protein fields at high temperatures.

3.4 Along the temperature-oxidation trajectory

Now let's make a chemical activity diagram showing the equilibrium activities of the proteins when temperature and $\log a_{H_2}$ are varied simultaneously according to the equation above. This ends up using the provision in `affinity` for specifying arbitrary increments of variables.

```
> par(mfrow = c(2, 1))
> mylogaH2 <- get.logaH2(Tval)
> out <- plot.it(c("T", "H2"), list(Tval, mylogaH2), rect = c(Tlim[1],
    -6, Tlim[2], loga.ref))
> label.plot("a", xfrac = 0.1)
```

```
> revisit(out$d, "richness", loga.ref = loga.ref, yline = 2, side = c(1,
    2, 4), plot.ext = FALSE)
> revisit(out$d, "richness", loga.ref = loga.ref, ispecies = ilowT,
    col = "blue", add = TRUE, plot.ext = FALSE)
> revisit(out$d, "richness", loga.ref = loga.ref, ispecies = ihighT,
    col = "red", add = TRUE, plot.ext = FALSE)
> label.plot("b", xfrac = 0.1)
> usr <- par("usr")
> usr[1:2] <- range(mylogaH2)
> opar <- par(usr = usr)
> thermo.axis(axis.label("H2"), side = 3, line = par("mgp")[1],
    cex = par("cex"))
```

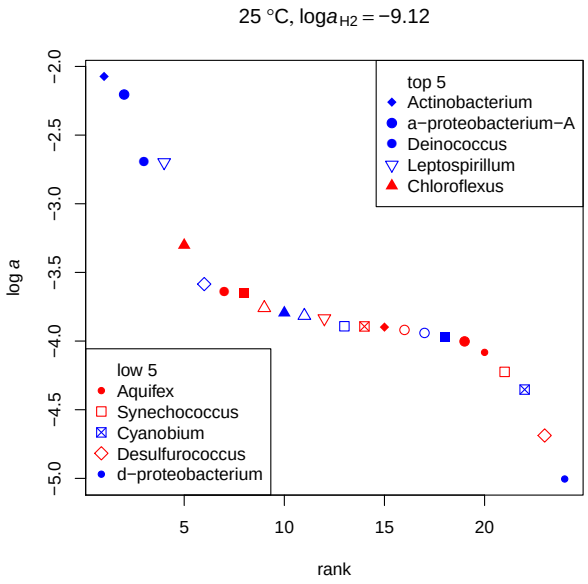


Now, we see that higher temperatures coupled with higher values of $\log a_{H_2}$ stabilize the proteins from organisms adapted to growth at higher temperatures.

3.5 Toward species abundance distributions

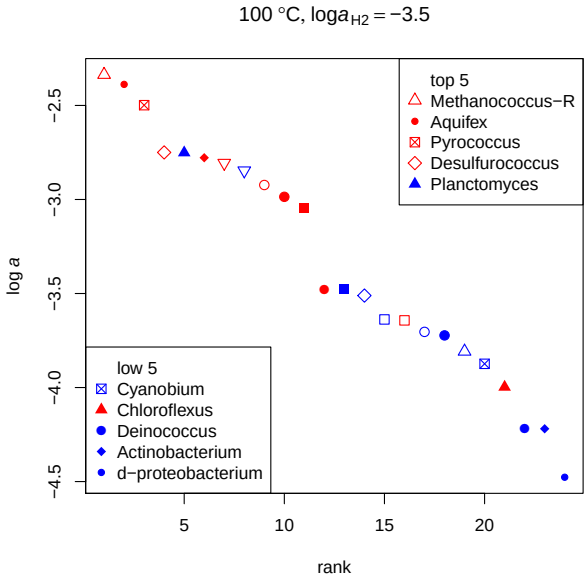
Another way of portraying the chemical activities shown in Fig. 3.4 above is with a rank-activity diagram.

```
> plot.rank(T = 25)
```



The open symbols represent RuBisCo and the closed symbols AcCoACo. Note that the proteins with the highest activity in this figure are the same as those with the highest activities at $T = 25$ °C in Fig. 3.4. Next up are the 100 °C results.

```
> plot.rank(T = 100)
```



Going from low to high temperature (along with the associated changes in $\log a_{H_2}$) there is clearly a major shift in the structure of the metastable assemblage of proteins. Not only do the most dominant (in metastable equilibrium) proteins change, but the overall shape goes from something that resembles a lognormal distribution to one that is more even.

The plotting was actually done using a function defined here called `plot.rank`. The function is too long to print out here, so check it out in the Sweave source for this vignette. The function might seem long and intimidating, but a lot of it is there to add the title and legends and to deal with sequences of plots

(i.e., increasing the point sizes for proteins that changed rank in the previous few frames). For example, the following bit of code produces a bunch of png files along a temperature/$\log a_{H_2}$ profile.

```
> png(filename = "Rplot%03d.png", width = 420, height = 320)
> plot.rank(T = 25:125, lcex = 0.8)
> dev.off()
```

The png files can be converted to an animated GIF using the ImageMagick (`http://imagemagick.org`) `convert` program (e.g., `convert -loop 0 -delay 15 Rplot*.png movie.gif`). A movie produced using this method (`carboxylase.gif`) is available online at the CHNOSZ website (`http://chnosz.net/vignettes`).

4 Conclusions

The calculations described above summarize a simple chemical thermodynamic model that uses protein reactions to study the adaptation of organisms to different growth temperatures. Using the parameters in the model, changing temperature alone produces results that do not align with the actual growth temperatures of organisms. When a temperature-dependent value of oxidation potential (expressed as activity of dissolved hydrogen) is imposed on the system, the predicted stabilities of the proteins fit better with the growth temperatures of the organisms.

References

[1] J. P. Allewalt, M. M. Bateson, N. P. Revsbech, K. Slack, and D. M. Ward. Effect of temperature and light on growth of and photosynthesis by *Synechococcus* isolates typical of those predominating in the octopus spring microbial mat community of Yellowstone National Park. *Appl. Environ. Microbiol.*, 72(1): 544 – 550, JAN 2006. ISSN 0099-2240. URL `http://dx.doi.org/10.1128/AEM.72.1.544-550.2006`.

[2] J. Bauld and J. T. Staley. *Planctomyces maris* sp. nov.: Marine isolate of *Planctomyces blastocaulis* group of budding bacteria. *J. Gen. Microbiol.*, 97:45 – 55, 1976. ISSN 0022-1287. URL `http://mic.sgmjournals.org/cgi/content/abstract/97/1/45`.

[3] B. Boeckmann, A. Bairoch, R. Apweiler, M.-C. Blatter, A. Estreicher, E. Gasteiger, M. J. Martin, K. Michoud, C. O'Donovan, I. Phan, S. Pilbout, and M. Schneider. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.*, 31(1):365 – 370, JAN 1 2003. URL `http://dx.doi.org/10.1093/nar/gkg095`.

[4] C. Ciniglia, H. S. Yoon, A. Pollio, G. Pinto, and D. Bhattacharya. Hidden biodiversity of the extremophilic Cyanidiales red algae. *Mol. Ecol.*, 13(7):1827 – 1838, JUL 2004. ISSN 0962-1083. URL `http://dx.doi.org/10.1111/j.1365-294X.2004.02180.x`.

[5] J. M. Dick. Calculation of the relative metastabilities of proteins using the CHNOSZ software package. *Geochem. Trans.*, 9:10, 2008. doi: 10.1186/1467-4866-9-10. URL `http://www.geochemicaltransactions.com/content/9/1/10`.

[6] I. Ferrera, S. Longhorn, A. B. Banta, Y. Liu, D. Preston, and A.-L. Reysenbach. Diversity of 16S rRNA gene, ITS region and *aclB* gene of the *Aquificales*. *Extremophiles*, 11(1):57 – 64, JAN 2007. ISSN 1431-0651. URL `http://dx.doi.org/10.1007/s00792-006-0009-2`.

[7] G. Fiala, K. O. Stetter, H. W. Jannasch, T. A. Langworthy, and J. Madon. *Staphylothermus marinus* sp. nov. represents a novel genus of extremely thermophilic submarine heterotrophic archaebacteria. *Syst. Appl. Microbiol.*, 8:106 – 113, JUL 1986. ISSN 0723-2020.

[8] C. B. Fliermans and T. D. Brock. Ecology of sulfur-oxidizing bacteria in hot acid soils. *J. Bacteriol.*, 111 (2):343 – 350, 1972. ISSN 0021-9193. URL `http://jb.asm.org/cgi/content/abstract/111/2/343`.

[9] B. W. Fouke, G. T. Bonheyo, B. Sanzenbacher, and J. Frias-Lopez. Partitioning of bacterial communities between travertine depositional facies at Mammoth Hot Springs, Yellowstone National Park, USA. *Can. J. Earth Sci.*, 40(11):1531 – 1548, NOV 2003. ISSN 0008-4077. URL http://dx.doi.org/10.1139/E03-067.

[10] J. M. Gonzalez, Y. Masuchi, F. T. Robb, J. W. Ammerman, D. L. Maeder, M. Yanagibayashi, J. Tamaoka, and C. Kato. *Pyrococcus horikoshii* sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. *Extremophiles*, 2(2):123 – 130, MAY 1998. ISSN 1431-0651. URL http://dx.doi.org/10.1007/s007920050051.

[11] S. Hanada, A. Hirasshi, K. Shimada, and K. Matsuura. *Chloroflexus aggregans* sp. nov., a filamentous phototrophic bacterium which forms dense cell aggregates by active gliding movement. *Int. J. Syst. Bacteriol.*, 45(4):676 – 681, OCT 1995. ISSN 0020-7713. URL http://dx.doi.org/10.1099/00207713-45-4-676.

[12] R. Huber, T. Wilharm, D. Huber, A. Trincone, S. Burggraf, H. König, R. Rachel, I. Rockinger, H. Fricke, and K. O. Stetter. *Aquifex pyrophilus* gen. nov. sp. nov., represents a novel group of marine hyperthermophilic hydrogen-oxidizing bacteria. *Syst. Appl. Microbiol.*, 15(3):340 – 351, AUG 1992.

[13] W. J. Jones, J. A. Leigh, F. Mayer, C. R. Woese, and R. S. Wolfe. *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. *Arch. Microbiol.*, 136(4): 254 – 261, 1983. doi: 10.1007/BF00425213. URL http://dx.doi.org/10.1007/BF00425213.

[14] W. J. Jones, M. J. B. Paynter, and R. Gupta. Characterization of *Methanococcus maripaludis* sp. nov., a new methanogen isolated from salt-marsh sediment. *Arch. Microbiol.*, 135(2):91 – 97, 1983. ISSN 0302-8933. URL http://dx.doi.org/10.1007/BF00408015.

[15] T. Kawasumi, Y. Igarashi, T. Kodama, and Y. Minoda. *Hydrogenobacter thermophilus* gen. nov., sp. nov., an extremely thermophilic, aerobic, hydrogen-oxidizing bacterium. *Int. J. Syst. Bacteriol.*, 34(1): 5 – 10, 1984. ISSN 0020-7713. URL http://dx.doi.org/10.1099/00207713-34-1-5.

[16] D. P. Kelly and A. P. Wood. Confirmation of *Thiobacillus denitrificans* as a species of the genus *Thiobacillus*, in the β-subclass of the *Proteobacteria*, with strain NCIMB 9548 as the type strain. *Int. J. Syst. Evol. Microbiol.*, 50:547 – 550, MAR 2000. ISSN 1466-5026. URL http://ijs.sgmjournals.org/cgi/content/abstract/50/2/547.

[17] A. López-Cortés, M.-L. Fardeau, G. Fauque, C. Joulian, and B. Ollivier. Reclassification of the sulfate- and nitrate-reducing bacterium *Desulfovibrio vulgaris* subsp. *oxamicus* as *Desulfovibrio oxamicus* sp nov., comb. nov. *Int. J. Syst. Evol. Microbiol.*, 56:1495 – 1499, JUL 2006. ISSN 1466-5026. URL http://dx.doi.org/10.1099/ijs.0.64074-0.

[18] J. H. McDonald. Patterns of temperature adaptation in proteins from the bacteria *Deinococcus radiodurans* and *Thermus thermophilus*. *Mol. Biol. Evol.*, 18(5):741 – 749, MAY 2001. ISSN 0737-4038. URL http://mbe.oxfordjournals.org/content/18/5/741.abstract.

[19] D. R. Meyer-Dombard, E. L. Shock, and J. P. Amend. Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. *Geobiology*, 3(3):211 – 227, JUL 2005 2005. ISSN 1472-4677. URL http://dx.doi.org/10.1111/j.1472-4669.2005.00052.x.

[20] T. B. Norris and R. W. Castenholz. Endolithic photosynthetic communities within ancient and recent travertine deposits in Yellowstone National Park. *FEMS Microbiol. Ecol.*, 57(3):470 – 483, SEP 2006. ISSN 0168-6496. URL http://dx.doi.org/10.1111/j.1574-6941.2006.00134.x.

[21] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2010. URL http://www.R-project.org.

[22] J. Raymond. The evolution of biological carbon and nitrogen cycling - a genomic perspective. In *Molecular Geomicrobiology*, volume 59 of *Reviews in Mineralogy and Geochemistry*, pages 211–231. 2005. URL http://dx.doi.org/10.2138/rmg.2005.59.9.

[23] J. R. Spear, J. J. Walker, T. M. McCollom, and N. R. Pace. Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proc. Natl. Acad. Sci. U. S. A.*, 102(7):2555 – 2560, FEB 15 2005. URL http://dx.doi.org/10.1073.pnas.0409574102.

[24] K. O. Stetter. Hyperthermophilic procaryotes. *FEMS Microbiol. Rev.*, 18:149 – 158, MAY 1996. ISSN 0168-6445. URL http://dx.doi.org/10.1016/0168-6445(96)00008-3.

[25] J. J. Walker, J. R. Spear, and N. R. Pace. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. *Nature*, 434(7036):1011 – 1014, APR 21 2005. ISSN 0028-0836. URL http://dx.doi.org/10.1038/nature03447.

[26] D. Wu, J. Raymond, M. Wu, S. Chatterji, Q. Ren, J. E. Graham, D. A. Bryant, F. Robb, A. Colman, L. J. Tallon, J. H. Badger, R. Madupu, N. L. Ward, and J. A. Eisen. Complete genome sequence of the aerobic co-oxidizing thermophile *Thermomicrobium roseum*. *PLoS One*, 4(1):e4207, 2009. URL http://dx.doi.org/10.1371/journal.pone.0004207.