A Different Kind Of Stability

Energy of Protein Formation

Jeffrey M. Dick

School of Earth and Space Exploration
Arizona State University

April 5, 2011

Motivation

Background

Protein geochemistry is concerned with the occurrence and variation of proteins in all environments on Earth.

Example

Amino acid differences between mitochondrial and nuclear proteins

Motivation

Background

Protein geochemistry is concerned with the occurrence and variation of proteins in all environments on Earth.

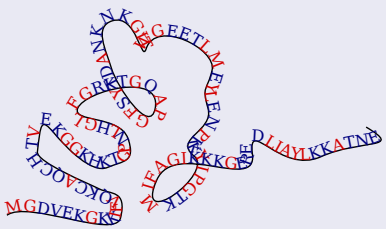
Hypothesis

Molecular evolution is a type of chemical reaction.

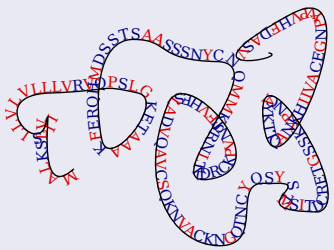
Outline

Folding Reactions

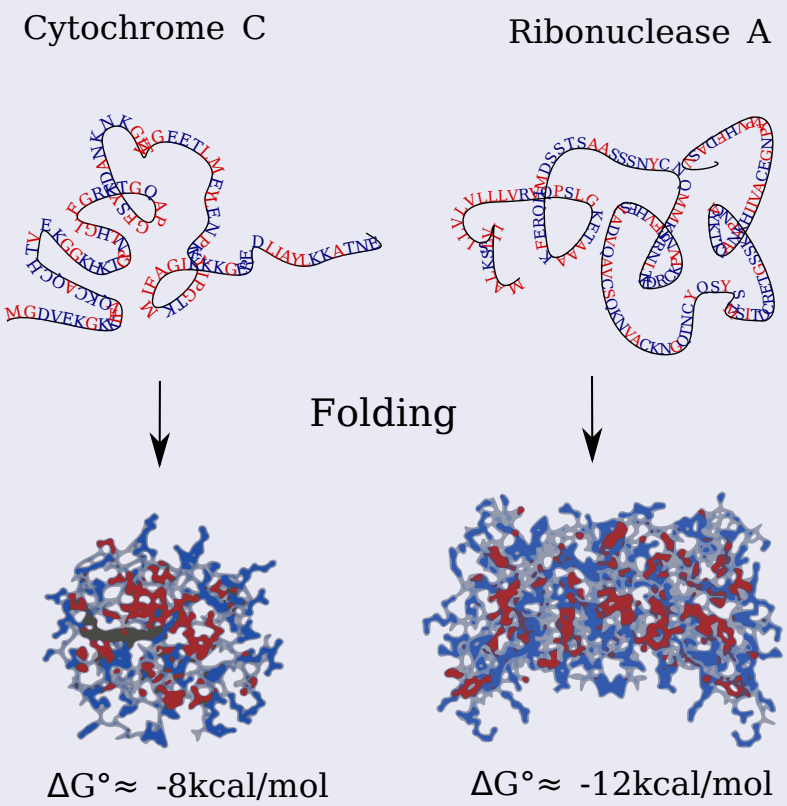
Cytochrome C



Ribonuclease A



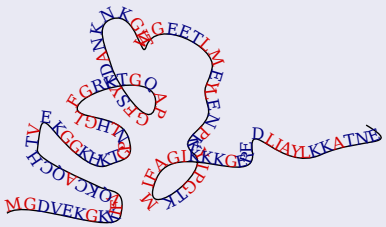
Folding Reactions



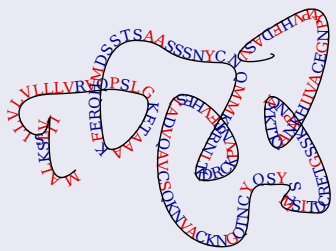
- Folding as a conformational process
- Stability referenced to unfolded protein
- Cellular/laboratory timescales

Formation Reactions

Cytochrome C



Ribonuclease A

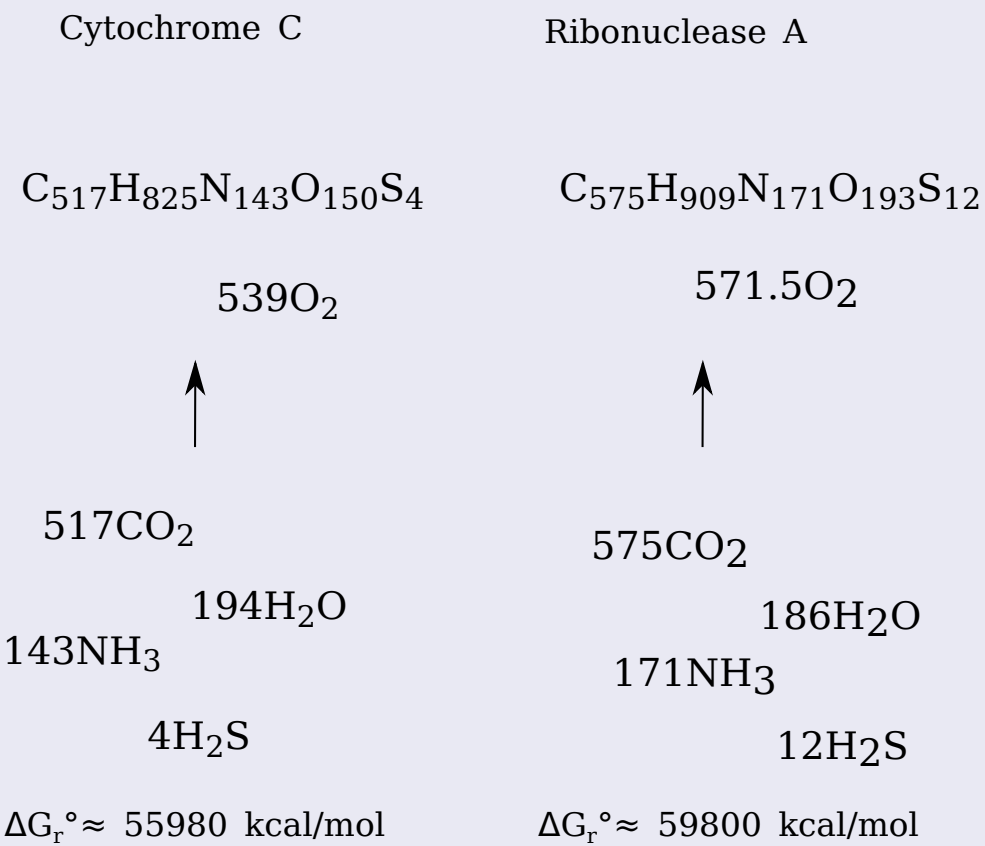


Formation Reactions

Cytochrome C	Ribonuclease A
$C_{517}H_{825}N_{143}O_{150}S_4$	$C_{575}H_{909}N_{171}O_{193}S_{12}$
$\Delta G_f^\circ \approx$ -3650 kcal/mol	$\Delta G_f^\circ \approx$ -4960 kcal/mol

- Formation as a chemical process
- ΔG_f°: Standard Gibbs energy of formation from the elements

Formation Reactions



- Formation as a chemical process
- Stability referenced to inorganic species
- ΔG_r°: Standard Gibbs energy of reaction
- Overall energy change is independent of mechanism

Residue Equivalents

Cytochrome C

$$\frac{1}{104}\,C_{517}H_{825}N_{143}O_{150}S_4 + 5.18O_2 \rightarrow 4.97CO_2 + 1.86H_2O + 1.38NH_3 + 0.04H_2S$$

$\Delta G_r^\circ \approx$ 538 kcal/mol

Ribonuclease A

$$\frac{1}{124}C_{575}H_{909}N_{171}O_{193}S_{12} + 4.61O_2 \rightarrow 4.64CO_2 + 1.50H_2O + 1.38NH_3 + 0.10H_2S$$

$\Delta G_r^\circ \approx$ 482 kcal/mol

- Reactions normalized by protein length
- Energetic meaning of reaction coefficients

Environment & Energy

Shift to lower O_2 potential more strongly decreases Gibbs energy of formation of CYC than RNAS1.

Residue Equivalents

Cytochrome C Ribonuclease A

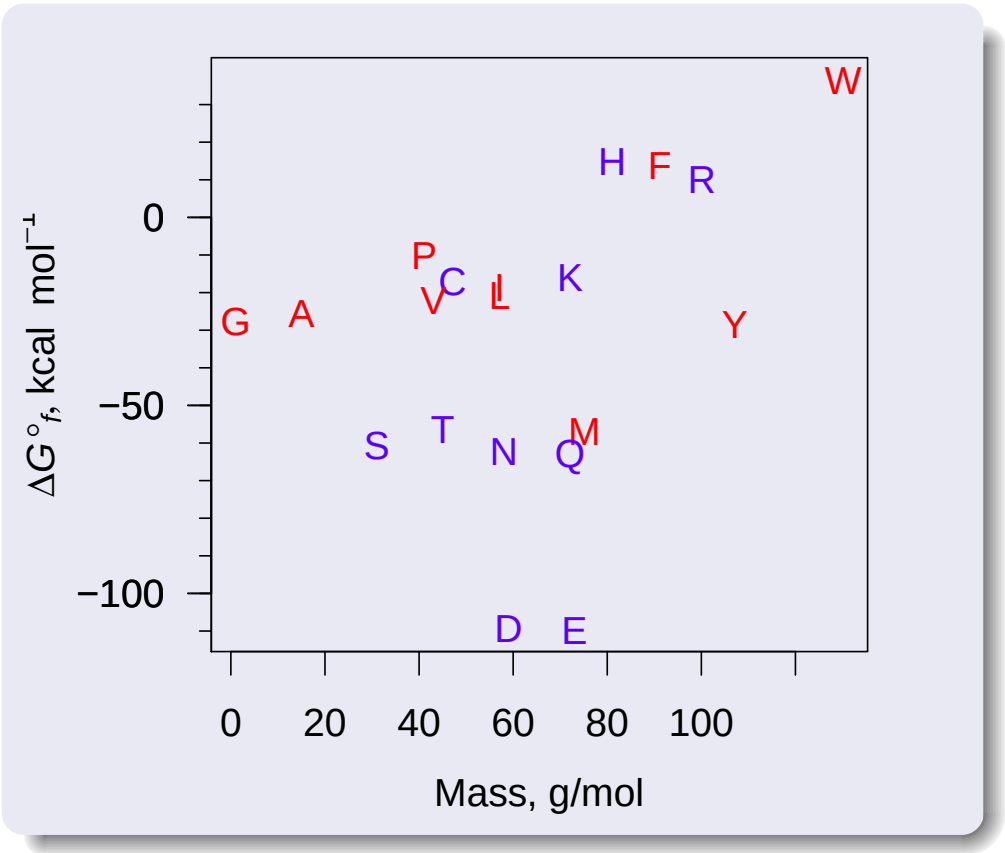
$$\frac{1}{104}\mathrm{C_{517}H_{825}N_{143}O_{150}S_4} + 0.004\mathrm{NH_3} + 0.57\mathrm{O_2} + 0.06\mathrm{H_2S} \rightleftharpoons \frac{1}{124}\mathrm{C_{575}H_{909}N_{171}O_{193}S_{12}} + 0.33\mathrm{CO_2} + 0.36\mathrm{H_2O}$$

- Reactions normalized by protein length
- Energetic meaning of reaction coefficients
- Transformation reaction; cellular to evolutionary timescales

Environment & Energy

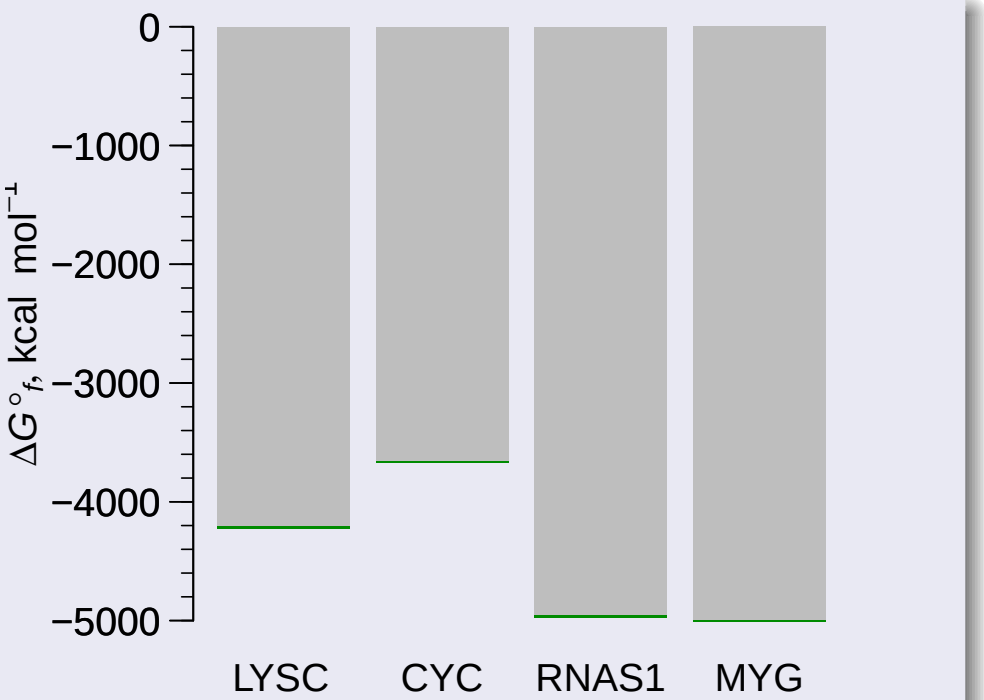
Shift to lower O_2 potential more strongly decreases Gibbs energy of formation of CYC than RNAS1.

Standard Gibbs Energies



- Group additivity of aqueous species properties for amino acid residues [Dick et al., 2006]

Standard Gibbs Energies



- Group additivity of aqueous species properties for amino acid residues [Dick et al., 2006]
- Protein size dependence of standard Gibbs energies

LYSC	Lysozyme	129
CYC	Cytochrome C	104
RNAS1	Ribonuclease A	124
MYG	Myoglobin	153

- Gibbs energies of folding [Privalov and Khechinashvili, 1974] are 1% or less of energies of formation.

Computational Plan

- Calculate standard Gibbs energies of proteins from amino acid group additivity and equations of state for aqueous species.
- Proceed without accounting for energy of protein folding.
  - Per mole of protein, energy of folding is small compared to Gibbs energy of chemical formation reaction.
  - If all proteins are folded, energy of folding tends to cancel in relative stability calculations.

Metastable Equilibrium

- Chemical affinity is opposite of Gibbs energy change due to reaction progress,

- $dG = -SdT + VdP - Ad\xi$

G	Gibbs energy
S	Entropy
T	Temperature
V	Volume
P	Pressure
A	Chemical Affinity
ξ	Reaction Progress

Metastable Equilibrium

- Chemical affinity is opposite of Gibbs energy change due to reaction progress,
- computed as a function of standard Gibbs energy, reaction stoichiometry and chemical activities.

- $dG = -SdT + VdP - Ad\xi$
- $A = 2.303RT \log(K / Q)$

A	Chemical Affinity
K	Equilibrium Constant
Q	Activity Product
R	Gas Constant

Metastable Equilibrium

- Chemical affinity is opposite of Gibbs energy change due to reaction progress,
- computed as a function of standard Gibbs energy, reaction stoichiometry and chemical activities.

- $dG = -SdT + VdP - Ad\xi$
- $A = 2.303RT \log(K / Q)$

$$
\begin{aligned}
A &= -\Delta G_r \\
K &= 10\hat{}\left(-\Delta G_r^{\circ} / 2.303RT\right) \\
Q &= 10\hat{}\sum \nu \log a = \prod a^{\nu} \\
\nu &\quad \text{Reaction Coefficient} \\
a &\quad \text{Chemical Activity}
\end{aligned}
$$

Metastable Equilibrium

- Chemical affinity is opposite of Gibbs energy change due to reaction progress,
- computed as a function of standard Gibbs energy, reaction stoichiometry and chemical activities.
- Maxwell-Boltzmann distribution allows for a transformation between reference states.

- $dG = -SdT + VdP - Ad\xi$
- $A = 2.303RT \log(K / Q)$
- $\frac{a}{\sum a} = \frac{e^{A/RT}}{\sum e^{A/RT}}$

Equal-activity reference state

More stable: higher affinity (A)

Metastable Equilibrium

- Chemical affinity is opposite of Gibbs energy change due to reaction progress,
- computed as a function of standard Gibbs energy, reaction stoichiometry and chemical activities.
- Maxwell-Boltzmann distribution allows for a transformation between reference states.
- When the chemical affinities of the formation reactions are all equal, the proteins are in metastable equilibrium.

- $dG = -SdT + VdP - Ad\xi$
- $A = 2.303RT \log(K / Q)$
- $\frac{a}{\sum a} = \frac{e^{A/RT}}{\sum e^{A/RT}}$

Equal-activity reference state

More stable: higher affinity (A)

Equal-affinity reference state

More stable: higher activity (a)

Computational Plan

- Calculate standard Gibbs energies of proteins from amino acid group additivity and equations of state for aqueous species.
- Proceed without accounting for energy of protein folding.
- Calculate equilibrium activities of proteins.
  - Start with chemical affinities in equal-activity reference state.
  - Use reference state transformation to calculate equilibrium activities.

Formation of Ionized Unfolded CYC_BOVIN

$$4.97\mathrm{CO_2} + 1.38\mathrm{NH_3} + 0.04\mathrm{H_2S} + 1.86\mathrm{H_2O} + 0.08\mathrm{H^+} \rightarrow \frac{1}{104}\mathrm{C_{517}H_{825}N_{143}O_{15}S_4^{+0.08}} + 5.18\mathrm{O_2}$$

- Ionization of proteins using additivity

▸ more info?

Formation of Ionized Unfolded CYC_BOVIN

$$4.97\mathrm{CO_2} + 1.38\mathrm{NH_3} + 0.04\mathrm{H_2S} + 1.86\mathrm{H_2O} + 0.08\mathrm{H^+} \rightarrow \mathrm{C_{4.97}H_{7.93}N_{1.38}O_{0.14}S_{0.04}^{+0.08}} + 5.18\mathrm{O_2}$$

- Ionization of proteins using additivity
  ▸ more info?
- Write per-residue formulas.

Formation of Ionized Unfolded CYC_BOVIN

$$4.97\mathrm{CO_2} + 1.38\mathrm{NH_3} + 0.04\mathrm{H_2S} + 1.86\mathrm{H_2O} + 0.08\mathrm{H^+} \rightarrow \mathrm{C_{4.97}H_{7.93}N_{1.38}O_{0.14}S_{0.04}^{+0.08}} + 5.18\mathrm{O_2}$$

$$\log K = -\Delta G_r^\circ / 2.303RT$$

$$= -393.0$$

- Ionization of proteins using additivity
  - more info?
- Write per-residue formulas.

T	25 °C
P	1 bar

Formation of Ionized Unfolded CYC_BOVIN

$$4.97\mathrm{CO_2} + 1.38\mathrm{NH_3} + 0.04\mathrm{H_2S} + 1.86\mathrm{H_2O} + 0.08\mathrm{H^+} \rightarrow \mathrm{C_{4.97}H_{7.93}N_{1.38}O_{0.14}S_{0.04}^{+0.08}} + 5.18\mathrm{O_2}$$

$$\log Q = \log a_{\mathrm{C_{4.97}H_{7.93}N_{1.38}O_{0.14}S_{0.04}^{+0.08}}} + 5.18 \log f_{\mathrm{O_2}} - 4.97 \log a_{\mathrm{CO_2}} - 1.38 \log a_{\mathrm{NH_3}} - 0.04 \log a_{\mathrm{H_2S}} - 1.86 \log a_{\mathrm{H_2O}} - \log a_{\mathrm{H^+}}$$

$$= -393.4$$

- Ionization of proteins using additivity
  - more info?
- Write per-residue formulas.

$\log a_{\mathrm{residue}}$	0
$\log a_{\mathrm{CO_2}}$	-3
$\log a_{\mathrm{H_2O}}$	0
$\log a_{\mathrm{NH_3}}$	-4
$\log f_{\mathrm{O_2}}$	-80
$\log a_{\mathrm{H_2S}}$	-7
pH	7

Formation Properties

Values per Residue

Protein	$\log K$	$\log Q$	$\log a$	$A/2.303RT$
LYSC	-361.6	-357.4	0	-4.17
CYC	-393.0	-393.4	0	0.34
RNAS1	-352.6	-348.4	0	-4.27
MYG	-407.6	-408.6	0	0.96

- Equal activity reference state: MYG is more stable.

Formation Properties

Values per Residue

Protein	$\log K$	$\log Q$	$\log a$	$A/2.303RT$
LYSC	-361.6	-362.0	-4.62	0.45
CYC	-393.0	-393.5	-0.10	0.45
RNAS1	-352.6	-353.1	-4.72	0.45
MYG	-407.6	-408.0	0.51	0.45

- Equal activity reference state: MYG is more stable.
- Equal affinity reference state: MYG is still more stable!

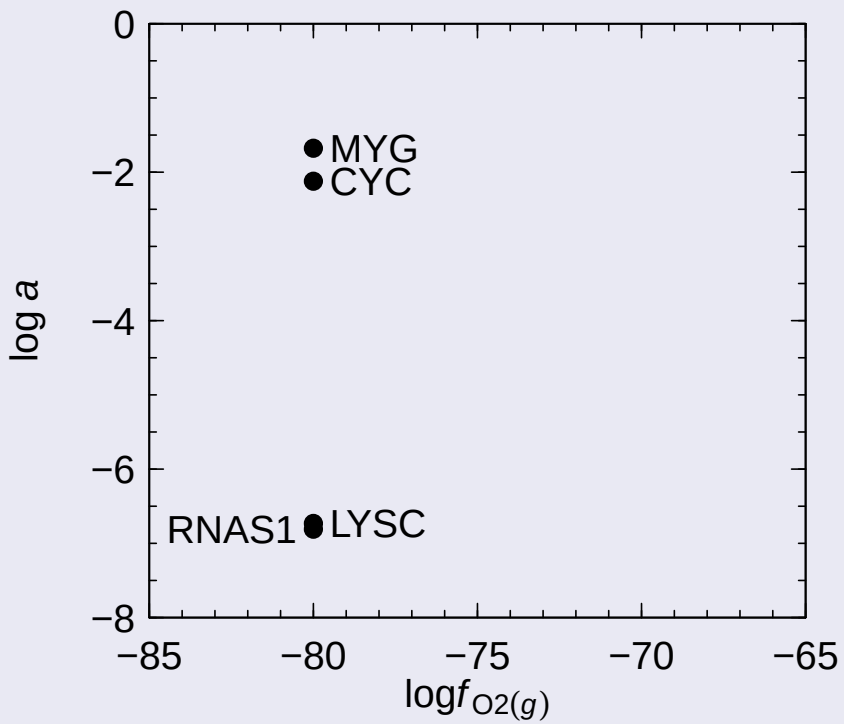
Formation Properties

Values per Residue

Protein	$\log K$	$\log Q$	$\log a$	$A/2.303RT$	$\log a_{\text{protein}}$
LYSC	-361.6	-362.0	-4.62	0.45	-6.73
CYC	-393.0	-393.5	-0.10	0.45	-2.12
RNAS1	-352.6	-353.1	-4.72	0.45	-6.81
MYG	-407.6	-408.0	0.51	0.45	-1.68

- Equal activity reference state: MYG is more stable.
- Equal affinity reference state: MYG is still more stable!
- Molality of residue = molality of protein * protein length
- Activity of residue = activity of protein * protein length (assuming ideality)

Equilibrium Activity Diagrams

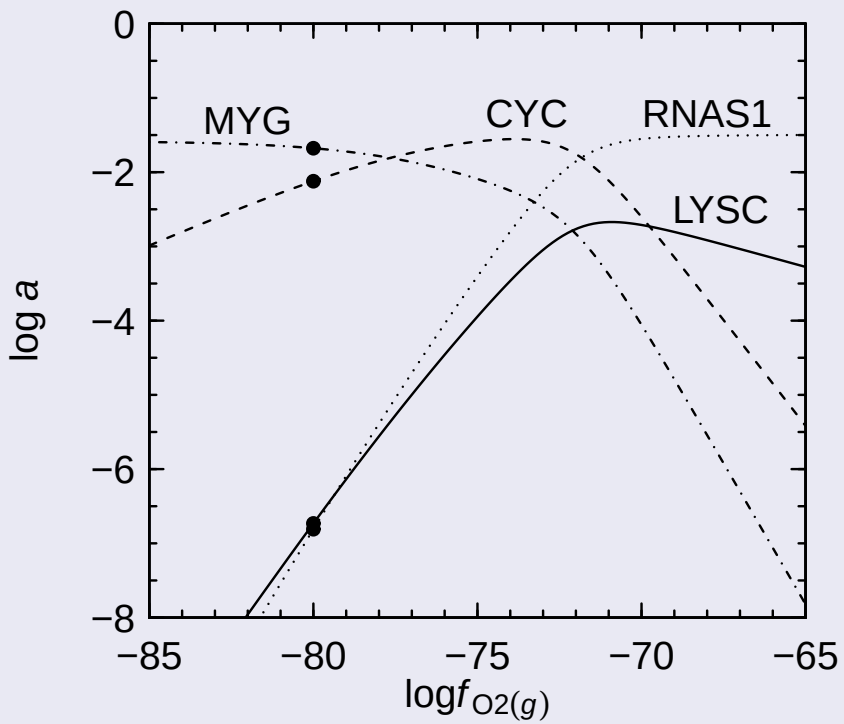


- Metastable equilibrium activities with total activity of residues $= 4$
- We are using $T = 25\ ^\circ$C and pH $= 7$

It Shows ...

MYG is relatively most stable at $\log f_{O_2} = -80$.

Equilibrium Activity Diagrams



- Metastable equilibrium activities as a function of $\log f_{O_2}$ with total activity of residues $= 4$
- $\log f_{O_2}$ can be converted to other measurements of oxidation-reduction potential. ▸ how?

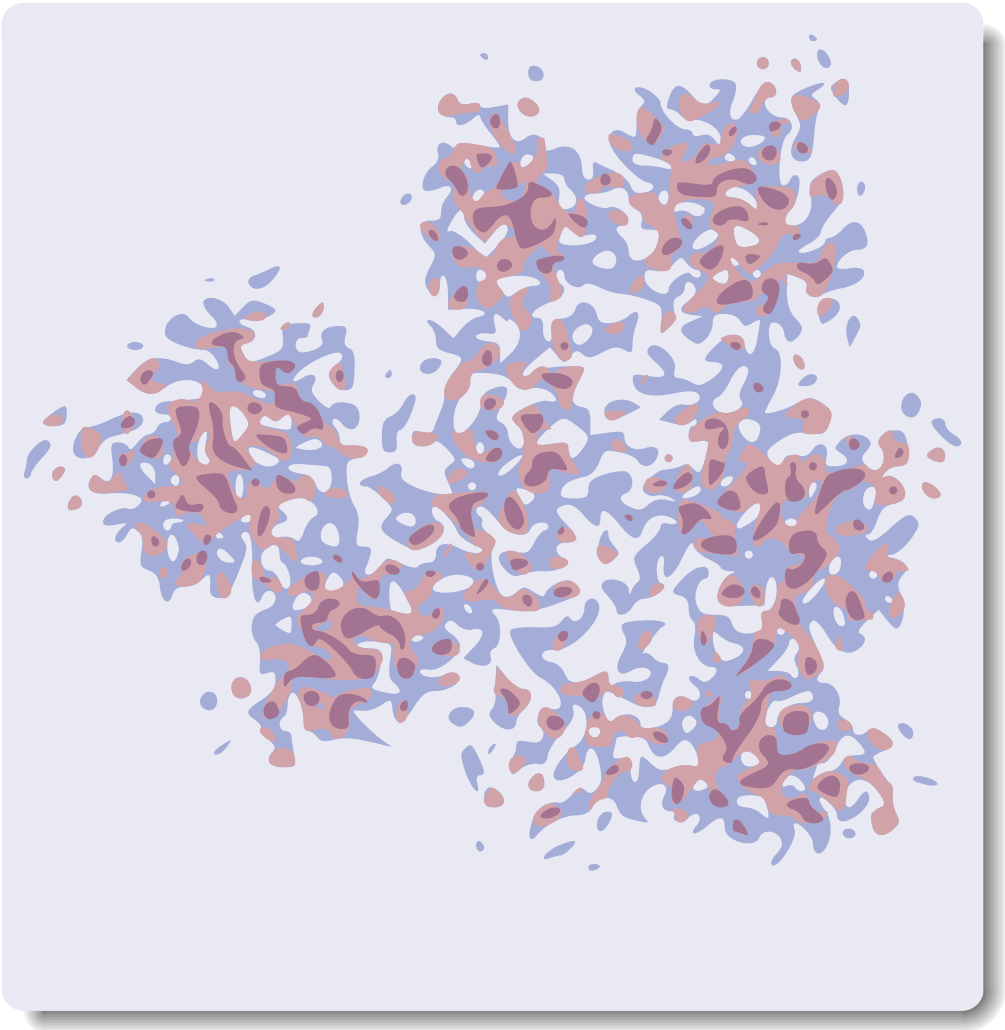
It Shows ...

Relative stability is sensitive to oxidation potential.

Computational Plan

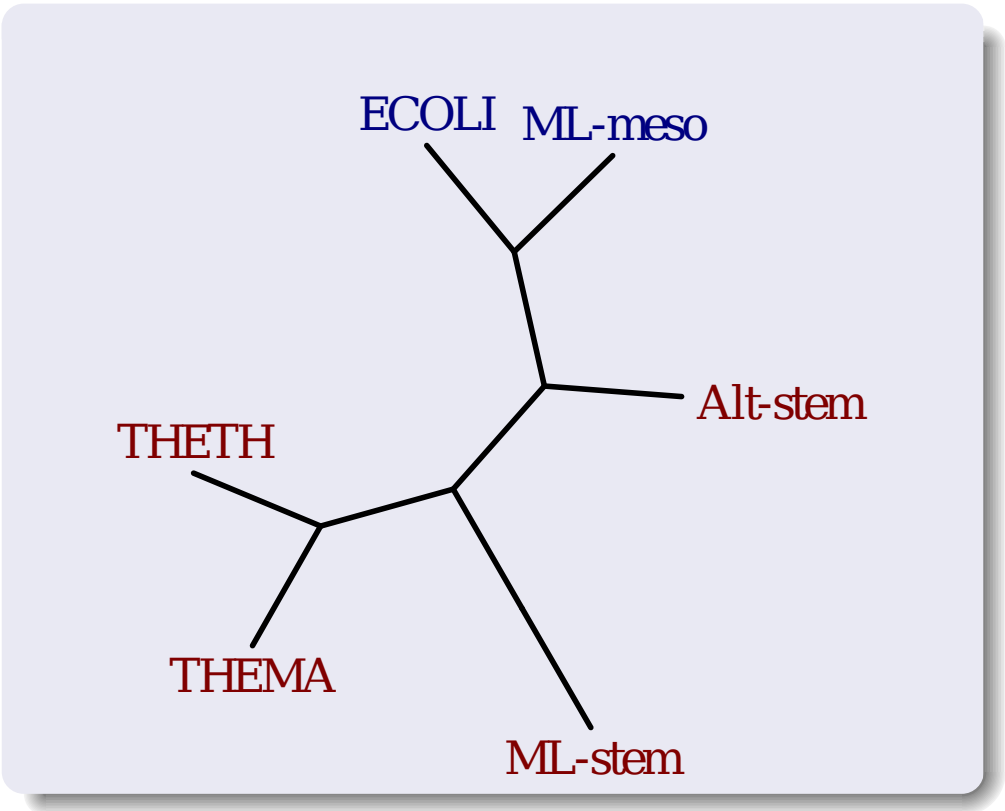
- Calculate standard Gibbs energies of proteins from amino acid group additivity and equations of state for aqueous species.
- Proceed without accounting for energy of protein folding.
- Calculate equilibrium activities of proteins.
  - Relative stabilities depend on the species (chemical compositions, standard Gibbs energies).
  - Relative stabilities depend on the environment (temperature, pressure, activities/fugacities of basis species).

Elongation Factor Tu



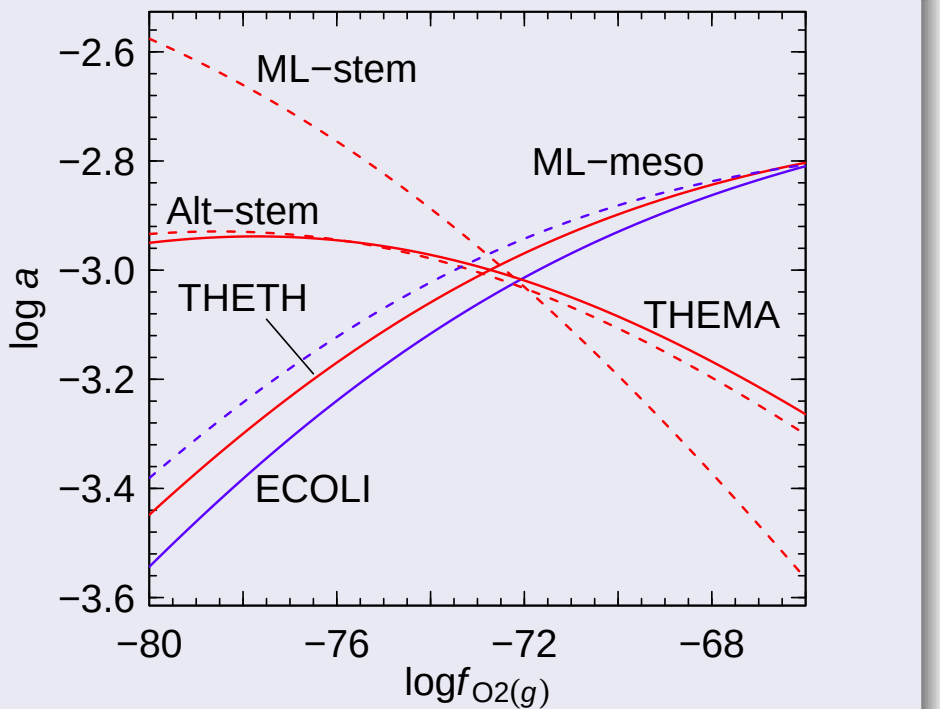
- EF-Tu from *Escherichia coli*

Elongation Factor Tu



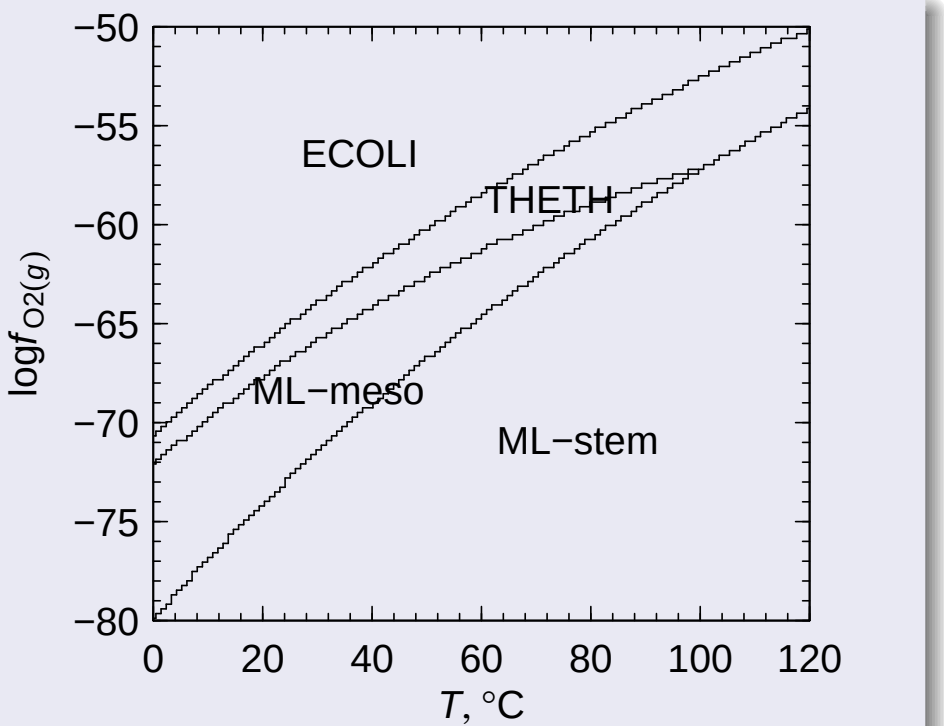
- EF-Tu's from *Escherichia coli*, *Thermotoga maritima*, *Thermus thermophilus*
- and reconstructed by maximum likelihood (**ML**) **stem** of bacterial tree, stem of **meso**philic bacteria, and **Alt**ernative tree topology [Gaucher et al., 2003]
- This tree built using parsimony (PHYLIP software), 394 aligned amino acids.

Equilibrium Activity Diagrams



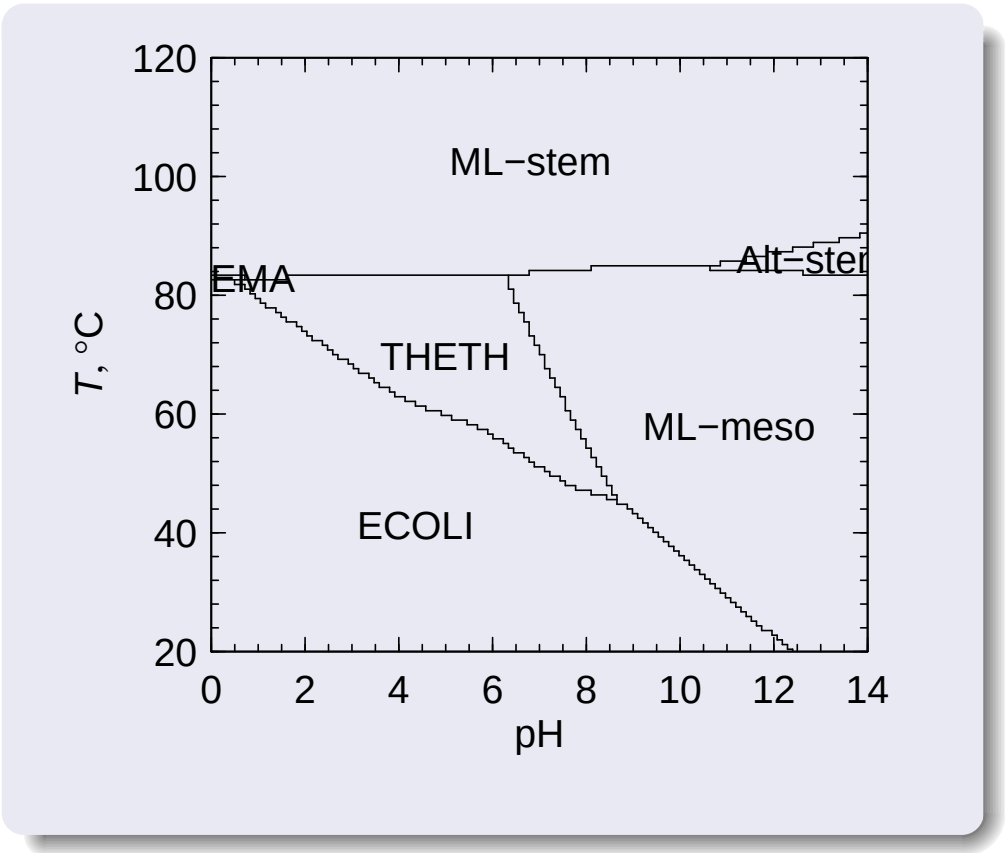
- Equilibrium activity diagram
- In most reactions, proteins from higher temperature favored by lower $\log f_{\mathrm{O}_{2(g)}}$.

Equilibrium Activity Diagrams



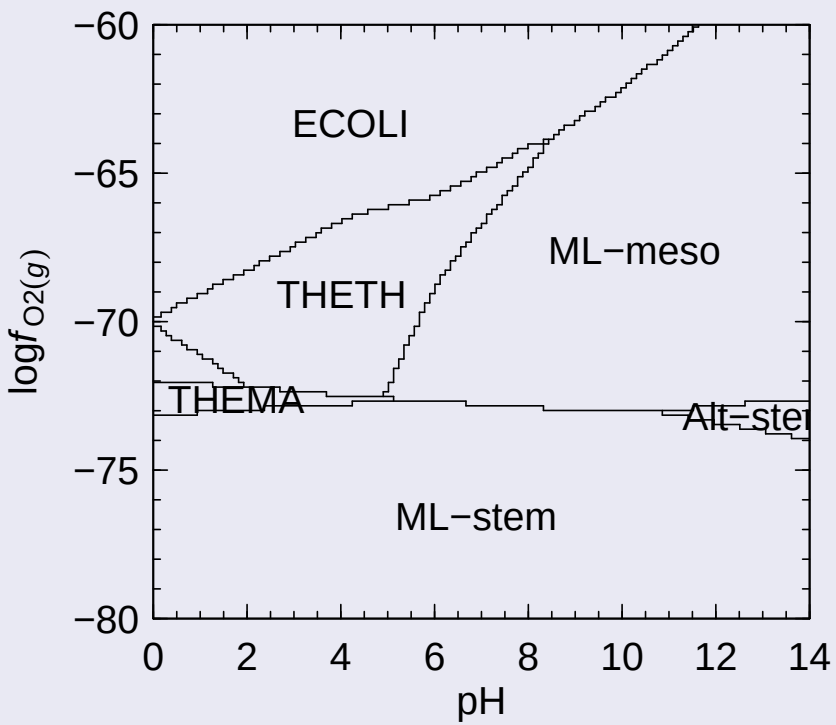
- Equilibrium predominance diagram
- At constant $f_{\mathrm{O}_{2(g)}}$, increasing T tends to favor "ML-stem".

Equilibrium Activity Diagrams



- Equilibrium predominance diagram
- $\log f_{O_2} = -60$

Equilibrium Activity Diagrams



- Equilibrium predominance diagram
- $T = 25\ ^\circ\mathrm{C}$

Multidimensionality

What about $\log a_{\mathrm{H_2O}}$, $\log a_{\mathrm{CO_2}}$, etc.?

Computational Plan

- Calculate standard Gibbs energies of proteins from amino acid group additivity and equations of state for aqueous species.
- Proceed without accounting for energy of protein folding.
- Calculate equilibrium activities of proteins.
- Investigate temperature, oxidation potential, pH, other chemical potentials.
  - Are reducing conditions associated with hotter environments?
  - The system is multidimensional; could also vary the chemical potentials of carbon, nitrogen, sulfur.

Computational Plan

- Calculate standard Gibbs energies of proteins from amino acid group additivity and equations of state for aqueous species.
- Proceed without accounting for energy of protein folding.
- Calculate equilibrium activities of proteins.
- Investigate temperature, oxidation potential, pH, other chemical potentials.
- Explore the protein universe using model systems.
- CHNOSZ is the software package used for the preceding calculations.

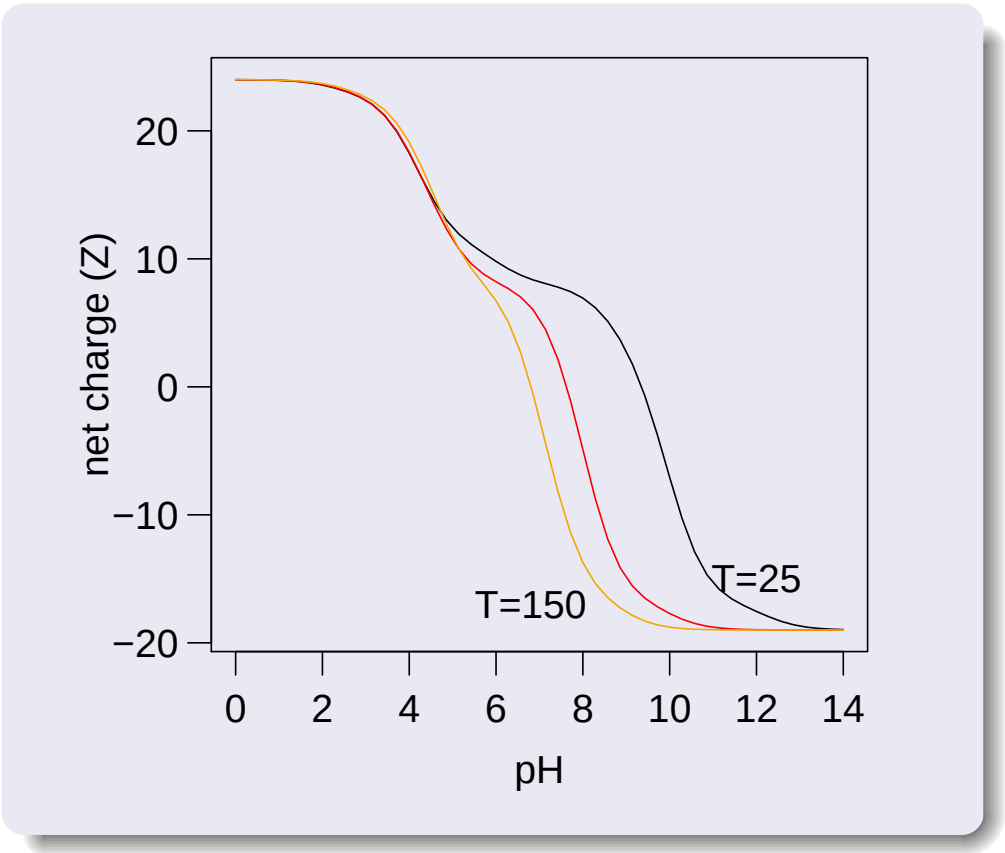
References

J. M. Dick, D. E. LaRowe, and H. C. Helgeson. Temperature, pressure, and electrochemical constraints on protein speciation: Group additivity calculation of the standard molal thermodynamic properties of ionized unfolded proteins. *Biogeosciences*, 3(3):311 – 336, 2006. URL http://www.biogeosciences.net/3/311/2006/.

E. A. Gaucher, J. M. Thomson, M. F. Burgan, and S. A. Benner. Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. *Nature*, 425(6955):285 – 288, SEP 18 2003. ISSN 0028-0836. URL http://dx.doi.org/10.1038/nature01977.

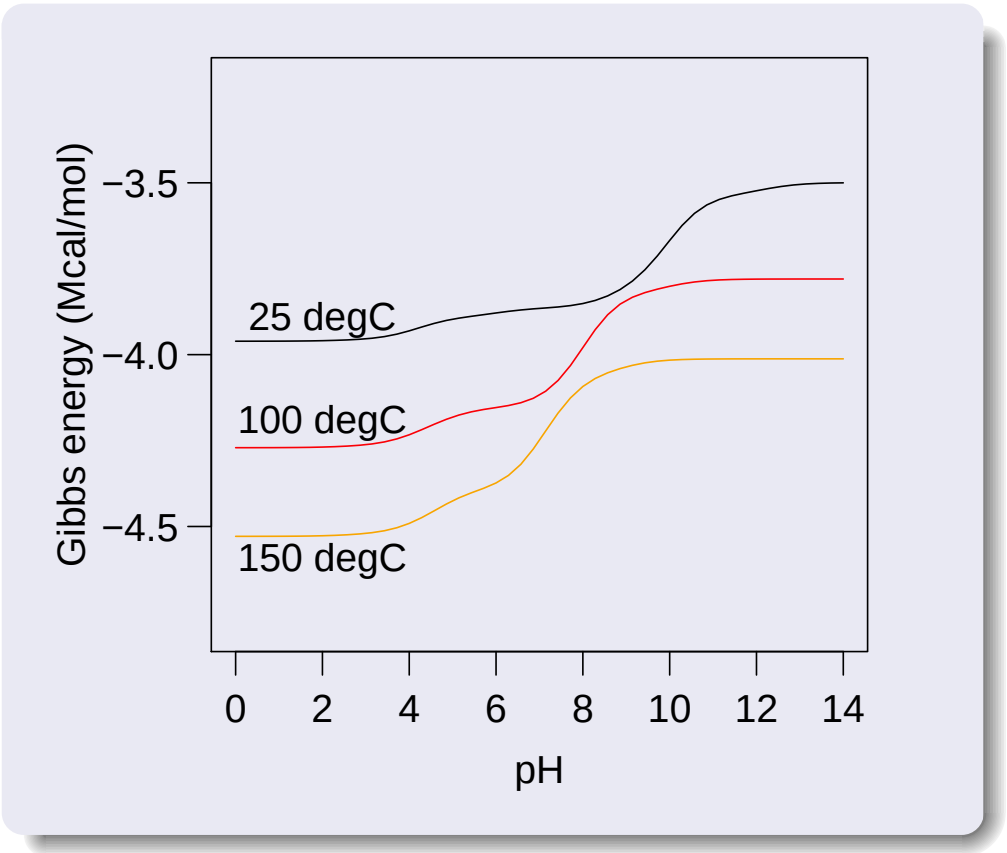
P. L. Privalov and N. N. Khechinashvili. A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. *J. Mol. Biol.*, 86(3):665 – 684, 1974. ISSN 0022-2836. URL http://dx.doi.org/10.1016/0022-2836(74)90188-0.

Protein Ionization (CYC_BOVIN)



- Net charges computed additively [Dick et al., 2006] using temperature-dependent sidechain p*K*a values

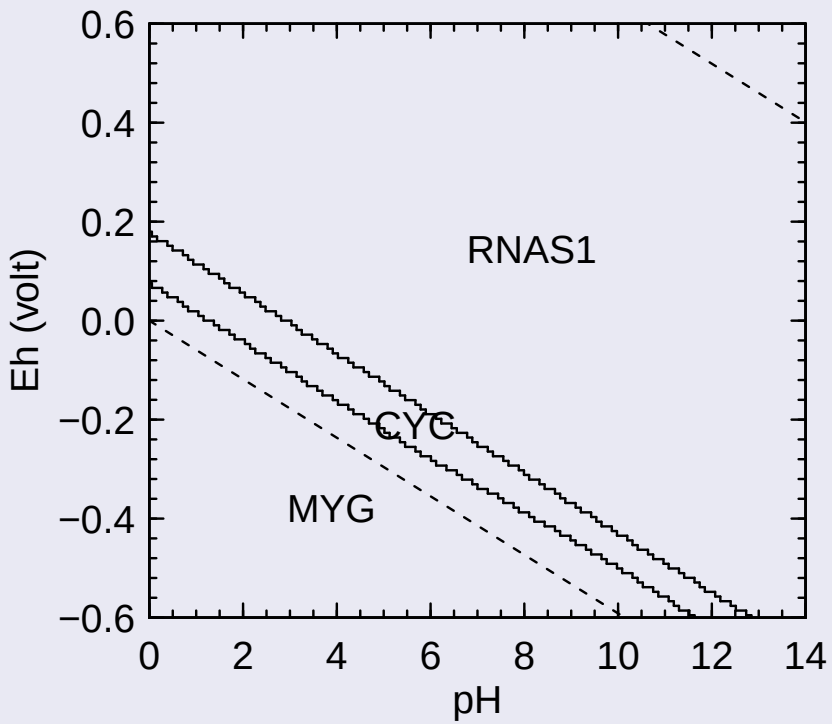
Protein Ionization (CYC_BOVIN)



- Net charges computed additively [Dick et al., 2006] using temperature-dependent sidechain p*K*a values
- Also affects standard Gibbs energies of the ionized proteins

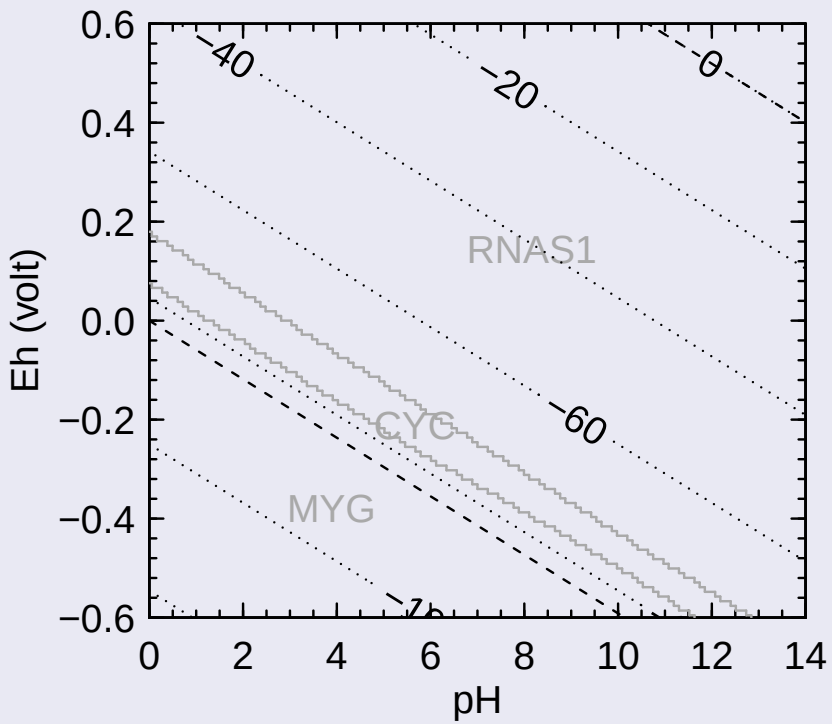
return

Oxygen Fugacity



- Eh-pH diagram for proteins
- Dashed lines indicate stability limits of $\mathrm{H_2O}$: $\log f_{\mathrm{O_2}} = 0$ (upper), $\log f_{\mathrm{O_2}} = -83.1$ (lower)

Oxygen Fugacity



- Eh-pH diagram for proteins
- Convert between Eh and $\log f_{O_2}$ using law of mass action for $H_2O \rightleftharpoons \frac{1}{2}O_2 + 2H^+ + 2e^-$
- $Eh = \frac{RT}{F} pe = -\frac{RT}{F} \log a_{e^-}$

return