Essay

Mathematics, Thermodynamics, and Modeling to Address Ten Common Misconceptions about Protein Structure, Folding, and Stability

Srebrenka Robic

Agnes Scott College, Decatur, GA 30030

To fully understand the roles proteins play in cellular processes, students need to grasp complex ideas about protein structure, folding, and stability. Our current understanding of these topics is based on mathematical models and experimental data. However, protein structure, folding, and stability are often introduced as descriptive, qualitative phenomena in undergraduate classes. In the process of learning about these topics, students often form incorrect ideas. For example, by learning about protein folding in the context of protein synthesis, students may come to an incorrect conclusion that once synthesized on the ribosome, a protein spends its entire cellular life time in its fully folded native conformation. This is clearly not true; proteins are dynamic structures that undergo both local fluctuations and global unfolding events. To prevent and address such misconceptions, basic concepts of protein science can be introduced in the context of simple mathematical models and hands-on explorations of publicly available data sets. Ten common misconceptions about proteins are presented, along with suggestions for using equations, models, sequence, structure, and thermodynamic data to help students gain a deeper understanding of basic concepts relating to protein structure, folding, and stability.

INTRODUCTION

A typical undergraduate student's understanding of protein structure usually starts with learning that there is a complex three-dimensional structure beyond each textbook-style "blob" representing each protein. Although most students have a relatively easy time comprehending the differences among primary, secondary, tertiary, and quaternary levels of protein structures, they often struggle with understanding how protein structure relates to stability and activity of a protein. How can we help students develop a detailed understanding of protein structure and how it contributes to the protein function? These are complex ideas, and our "gut instincts" about protein structure are often wrong. Ten common misconceptions, along with suggestions for addressing these misconceptions in an undergraduate classroom, are discussed.

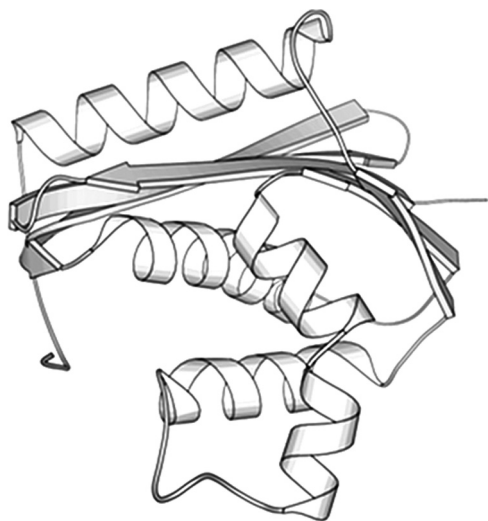## TEN COMMON MISCONCEPTIONS ABOUT PROTEINS

1. Crystal Structures Tell Us Exactly What a Protein Looks Like

Textbooks, journal articles, and the Protein Data Bank (PDB; RCSB Protein Data Bank, 2010) are full of three-dimensional structures of proteins. Most of these structures are derived from x-ray crystallography, although some are based on nuclear magnetic resonance (NMR). These structures provide valuable atomic-level details, but they also can be misleading in the eyes of a novice student. Does a crystal structure, such as that of ribonu-

Address correspondence to: Srebrenka Robic (srobic@agnesscott. edu).

Figure 1. Crystal structure of ribonuclease H (PDB access code 2RN2) as an example of a three-dimensional protein structure. Each structure is a data set of coordinates and other valuable information about the protein, such as the source organism, amino acid sequence, and relevant literature citations. Students can download PDB files from www.rcsb.org/pdb and use one of many free molecular visualization packages for exploration and quantitative analysis of protein structures.

cleave H (RNase H; Figure 1), imply that all of RNase H molecules look exactly as depicted in the structure diagram at all times? The answer is no, it is not that simple.

This topic can be introduced through exploration of a set of crystal structures for proteins for which multiple structures have been solved. For example, searching for human immunodeficiency virus protease in the PDB yields >200 hits. Students may explore this or a similar data set by asking the following questions:

1. Do the structures encompass the whole protein?
2. Are there any differences between multiple structures of the same protein?

The answer to the first question is easily available in the PDB entry page associated with each protein structure. This simple activity teaches a valuable concept that structures often do not reflect a whole protein but rather a fragment. The second question can be explored by using one of the freely available molecular visualization programs, such as Jmol (Jmol, 2010), PyMOL (DeLano, 2002), or SwissPDB Viewer (Guex *et al.*, 2008). Students can use these programs to overlay multiple structures and measure quantitative differences between them. The root mean square difference (RMSD) between coordinates of two structure files is a common measure of differences between structures. Structure viewing packages calculate RMSD between two sets of coordinates v and w using the following equation:

$$\text{RMSD}(v, w) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (v_{ix} - w_{ix})^2 + (v_{iy} - w_{iy})^2 + (v_{iz} - w_{iz})^2}$$

(1)

where n is the number of atoms considered (usually C_α atoms of protein structures), and x, y, and z are the coordinates of each of the considered atoms.
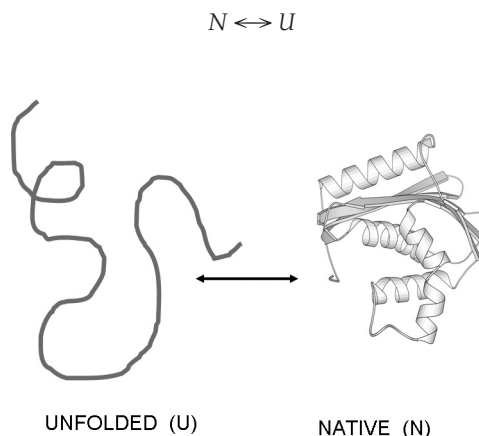
Such simple quantitative exploration will help students understand that protein structures depend on external conditions such as crystallization conditions (e.g., pH, buffer, and temperature), choice of the fragment, and presence or absence of various ligands and binding partners. In some cases, proteins have been crystallized in completely different conformations, such as the open and closed conformations of enzymes. This simple project will help the students develop an appreciation of protein structures as quantitative data sets that can be used as a basis for hypothesis testing and open-ended exploration of biological questions.

2. Once Synthesized on the Ribosome, Proteins Remain in Their Folded State

Beginning college biology students learn about protein folding as a process through which proteins attain their functional structure. Diagrams such as Figure 2 are often used to depict the process of protein folding. The concept of protein folding is typically first introduced in the context of protein synthesis, and students can easily get the impression that once a protein is synthesized and folded on the ribosome, it remains in that perfectly folded native state during its entire cellular life cycle. However, a protein exists in equilibrium between the native and unfolded state, and a number of folded proteins at some point unfold in the cellular environment.

To get students thinking about the folding process, it is helpful to introduce the idea of a quantifiable, measurable equilibrium between protein conformations. Under normal cellular conditions, proteins are found in an equilibrium between native and unfolded conformations. That means that a folded protein may occasionally unfold. The simplest model of folding, the two-state model, is easy to understand and can be introduced in introductory biology classes as a quantitative model of protein folding.

The two-state model states that under equilibrium conditions proteins are either in the fully folded native state (N) or unfolded state (U). This equilibrium,

$$N \longleftrightarrow U$$

Figure 2. Protein folding equilibrium. The simplest model of protein folding is the two-state model in which a protein is either fully unfolded (U) or fully folded to its native structure (N).

Table 1. Relationship between protein folding equilibrium constants (K_{unf}) and free energies of protein folding (ΔG_{unf})[a]

K_{unf}	ΔG_{unf} (kcal mol^{-1})
1	0.00
1/100	2.72
1/1,000	4.08
1/10,000	5.43
1/100,000	6.79
1/1,000,000	8.15
1/10,000,000	9.51

[a] Equation 2 in text was used to generate the table.

can be described using a simple equilibrium constant, $K_{unf} = \dfrac{[U]}{[N]}$, where [U] and [N] are concentrations of unfolded and native proteins, under a given set of conditions. Many introductory biology students are also taking introductory chemistry courses, in which they have learned about equilibrium constants in the context of simple chemical reactions.

This equilibrium constant can be related to the free energy of unfolding of a protein (ΔG_{unf}), also known as the thermodynamic stability of proteins, by the following equation:

$$\Delta G_{unf} = -RTlnK_{unf} \tag{2}$$

where R is the gas constant and T is temperature in Kelvin.

Most students also will have seen this equation in introductory chemistry courses. To demystify the numbers behind protein stability, students can use this equation to get a real sense of various equilibria. For example, the thermodynamic stability of chicken lysozyme is 10 kcal mol^{-1} (Ueda et al., 1993), whereas the thermodynamic stability of ubiquitin is 6.7 kcal mol^{-1} (Khorasanizadeh et al., 1993). What does this tell us about the proportion of unfolded molecules at room temperature?

We can encourage students to look up the thermodynamic stabilities of their favorite proteins and do the same calculations with the values they find in the literature or in the database of protein thermodynamic values (ProTherm, 2010). Students can use Excel (Microsoft, Redmond, WA) to generate a table exploring the relationship between equilibrium constants and free energy, based on Eq. 2 (Table 1).

This simple mathematical model gives meaning to thermodynamic stability measurements. The difference between a stability of 2.7 and 9.5 kcal mol^{-1} might not carry a lot of meaning to an undergraduate student. However, the difference between 1 in 100 molecules being unfolded versus 1 in 10 million molecules is much easier to grasp. Students realize that even proteins with high thermodynamic stabilities exist in equilibrium with a small fraction of unfolded proteins. This helps the student develop an appreciation of proteins as dynamic ensembles of conformations.

3. Unfolded Proteins Are Simply Stretched Out Polypeptide Chains

The unfolded state of proteins is even more difficult to grasp than the native state. Usually, the unfolded state is portrayed as a "polypeptide spaghetti" structure, similar to that shown in Figure 2. The unfolded state of a protein is an ensemble of many different conformations, some of which can be rather compact. Because of this diversity of conformations, structural characterization of the unfolded state is difficult. However, modern techniques, such as small-angle x-ray scattering, have helped us to learn more about the average properties of the unfolded state. One of these properties is the radius of gyration, the average distance from the center of gravity to each amino acid. Students can analyze the data in published primary research articles to learn that average unfolded protein structure is more compact than that of completely stretched out polypeptide (Kohn et al., 2004, McCarney et al., 2005). Molecular dynamics simulations also have helped us deepen our understanding of unfolded proteins, and they too provide a picture of diverse set of protein conformations, including relatively compact states (Snow et al., 2002). Students with interest in this topic can be directed to the Folding@home website that, in its Research Articles section, contains a deeper discussion of this topic at a level appropriate for undergraduate students (Pande, 2002, 2010). Students could use the interactive resources at Folding@home to explore topics such as forces that govern protein folding and unfolding, and the formation of secondary structural elements by viewing simulations and movies depicting the dynamic nature of proteins in both the folded and the unfolded state.

4. All Mutations Have a Drastic Effect on Protein Structure

Most students understand that sometimes a change as small as a single amino acid substitution can have a drastic effect on the function of a protein. Point mutations can result in a protein that can no longer bind to its binding partner, or a catalytically inactive enzyme. One possible explanation for this observation is a drastic change in protein structure. However, protein structure is rather robust, and in many cases mutations do not have a drastic effect on protein structure. In fact, some proteins can tolerate simultaneous substitutions of up to a quarter of their residues with different amino acids, without losing functionality (Besenmatter et al., 2007).

Rather than having a drastic effect on protein structure, most mutations have an effect on protein stability. Studies of individual proteins, such as staphylococcal nuclease and barnase, show that most individual mutations are destabilizing but do not change the overall protein structure drastically (Shortle et al., 1990; Green et al., 1992; Serrano et al., 1992). Furthermore, substituting an amino acid at a binding site, for example, can disrupt ionic quaternary interactions, rather than disrupting the protein fold.

To gain a deeper understanding of protein structure robustness, students can compare crystal structures of families of proteins, such as the globin family or the family of G protein-coupled receptors. They can ask questions such as: Which positions in the protein structure are more tolerant of mutations? Which types of amino acid substitutions will have more significant effects on protein structure? These types of questions can be explored using ConSurf (Landau et al., 2006). ConSurf is a web-based tool that helps in identification of functionally important, conserved protein regions

based on phylogenetic relationships between related sequence homologues. Students can use this application to identify, explore, and visualize both the regions that are conserved and the amino acid positions that are prone to variation.

5. The Stability of a Protein Can Be Determined by Considering Its Three-Dimensional Structure
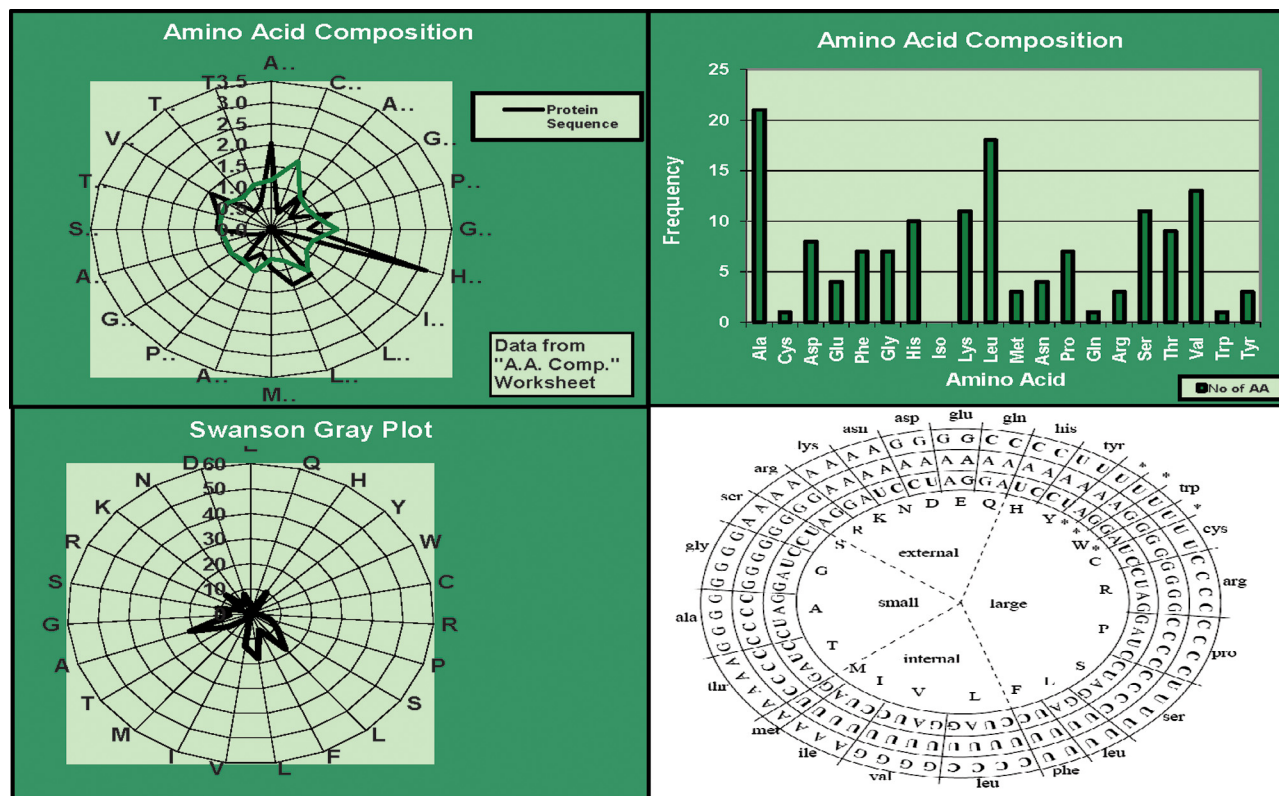
Introductory biology and biochemistry textbooks usually have a section describing forces that contribute to protein stability. Van der Waals interactions, hydrophobic interactions, hydrogen bonds, electrostatic interactions, salt bridges, and disulfide bridges are listed as factors that contribute to protein stability. I often encounter the perspective that you can compare the relative stabilities of two proteins by simply counting the numbers of various interactions, such as salt bridges or disulfides. Unfortunately, our understanding of these forces is not detailed enough to predict which protein is more stable simply by looking at its structure.

The relationship between structure and stability can be explored using protein structures and experimental data. A comparison of two homologous proteins from organisms with drastically different optimal growth temperatures can be used as a starting point for exploration of the relationship between structure and stability. An example is a pair of ribonucleases H, one from the mesophilic bacterium *Escherichia coli* (Figure 1)

and one from the thermophilic bacterium *Thermus thermophilus* (PDB code 1RIL), which has an optimal growth temperature of 66°C. These two proteins have almost identical structures, yet they are very different in their thermodynamic stability profiles.

Students can easily analyze the sequences of these two proteins to search for clues to a possible difference in stability. Is there a difference in overall percentage of charged residues? Is there a difference in the number of prolines, which might restrict the flexibility of the unfolded state? Is the thermophilic protein stabilized through increased compactness and shortening of loops as has been proposed for some other thermophilic proteins? (Kumar *et al.*, 2000). These and similar questions can be explored using the BioQUEST Esteem Module Protein Analysis (BioQUEST Curriculum Consortium, 2010). The module allows the students to input protein sequence data and visualize and quantify various parameters, such as amino acid and charge distributions (Figure 3).

In addition to analyzing sequence data, students also may want to explore three-dimensional protein structure data. Any of the already mentioned molecular visualization programs (Jmol, PyMOL, and SwissPDB Viewer) are suitable for quantitative analysis of molecular structures. Students can use these packages for visualization, as well as for quantitative measurements of size, distances, and differences between related structures. Through such hands-on

Figure 3. Amino acid distributions of a protein sequence is one of the outputs of Protein Analysis ESTEEM Module (BioQUEST, 2010). The module presents different ways of calculating and visualizing protein sequence parameters such as amino acid frequencies and charge distributions.

exploration of proteins with similar structures and different stability profiles, students will be able to better understand the complexities of the problem of protein stability.

6. A More Stable Protein Will Last Longer

The language of protein stability can be confusing even to a seasoned protein chemist, let alone an undergraduate student. Students often have a vague idea of protein stability having something to do with how resilient a protein is under certain conditions—or how long it can "last" under those conditions. Even biochemists often use the term stability to refer to various ideas ranging from resistance to various chemicals, to temperature, to enzyme activity under those conditions.

However, the thermodynamic stability of proteins is a precisely defined quantity that can be described with physical and mathematical models. The thermodynamic stability of a protein is defined as the difference in free energy between folded and unfolded conformations of the protein, i.e.,

$$\Delta G_{unf} = G_U - G_N \qquad (3)$$

where U is the unfolded state and N is the native conformation of a protein. This free energy of unfolding of a protein is related to the equilibrium constant describing the ratio of folded and unfolded molecules, as specified in Eq. 2.
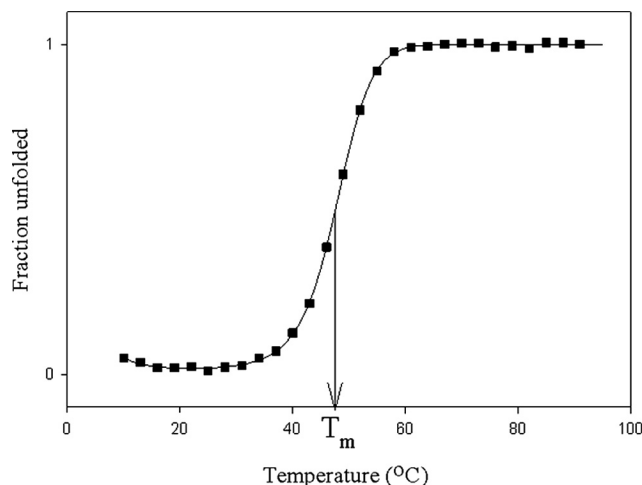
Once students grasp the idea of thermodynamic stability by exploring the relationship between free energy and ratios of folded to unfolded molecules, it becomes easier to explain why this concept is not related to how long proteins "last." This exploration can become a great starting point for discussing differences between thermodynamics and kinetics. Words such as "lasts" and "keeps" generally imply passage of time, which are indicative of kinetic properties and not at all related to thermodynamic properties, such as the thermodynamic stability of proteins.

7. Proteins with a Higher Melting Temperature Are Also Thermodynamically More Stable

One easy-to-understand, and easy-to-measure, parameter associated with protein stability is the so-called melting temperature, i.e., the midpoint of thermal denaturation curve (T_m). This parameter can be determined by observing a protein's structural signal, such as circular dichroism, fluorescence, or NMR signal, as a function of temperature. Experimental determination of T_m can be performed within a time frame of a typical undergraduate lab (Raabe and Gentile, 2008). T_m is simply the half point of the transition for a protein that folds in a cooperative two-state manner (Figure 4). Clearly, this value is related to protein stability, but does it actually tell us anything about the thermodynamic stability of proteins, as defined in the previous paragraph?

The relationship between temperature and thermodynamic stability of proteins can be modeled by the Gibbs–Helmholtz equation:

$$\Delta G_{unf} = \Delta H_0 - T\frac{\Delta H_0}{T_m} + \Delta C_P\left[T - T_m - T\ln\left(\frac{T}{T_m}\right)\right] \qquad (4)$$

Figure 4. Thermal denaturation of proteins can be monitored by probes such as circular dichroism or fluorescence. Many small proteins exhibit a cooperative, sigmoidal transition with temperature. The midpoint of this transition is the T_m, sometimes referred to as the melting point.

where T_m is the midpoint of thermal denaturation ("melting point"), ΔH_0 is the enthalpy at T_m, and ΔC_P is the change in heat capacity upon unfolding (Becktel and Schellman, 1987). Even a first glance at this equation shows that the relationship between ΔG_{unf} and T_m is not a simple linear dependence. By analyzing the equation, students can answer the question; Which other factors, in addition to T_m, will influence ΔG_{unf}?
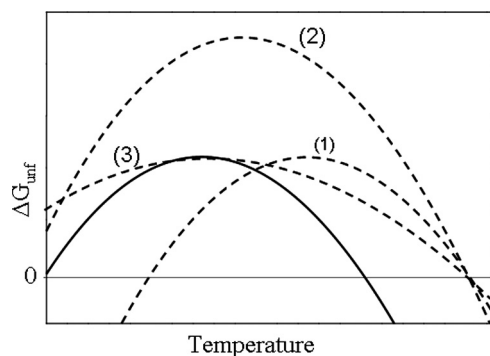
To better grasp this complex relationship, students may start by analyzing real data. Once again, the ProTherm database may be a valuable starting point, in addition to more involved literature searches (ProTherm, 2010). Is there a correlation between a high T_m and a high ΔG for all proteins?

This question has been of particular relevance in protein engineering. The method of protein engineering has been used to design proteins with increased stability. Although the method has succeeded in increasing the T_m in a few cases, such proteins do not have a desired higher ΔG_{unf} (Loladze et al., 1999). A protein with a high T_m does not necessarily also have a high thermodynamic stability.

Proteins can use several different thermodynamic strategies to achieve a higher T_m (Figure 5). One of these strategies is to have higher thermodynamic stabilities at all temperatures; however, a higher T_m also can be achieved by shifting the overall stability curve to the right, or by flattening the curve.

8. The Reason We Often Keep Proteins on Ice Is Because They Are More Stable on Ice Than at Room Temperature

Any student who has worked with enzymes knows to keep them on ice. From restriction enzymes in molecular biology, to enzymes and proteins they isolate from plants or bacteria, all students know they have to put the enzymes on ice. When asked why, a typical answer is "because proteins are

Figure 5. Protein stability curves depict the dependence of thermodynamic stability (ΔG_{unf}) on temperature. The stability curve of a mesophilic protein is depicted as a black continuous line. Three simple scenarios by which a thermophilic protein can achieve a higher T_m include shifting the stability curve to the right (1), raising it (2), or broadening the stability curve (3). Certain combinations of these three scenarios also could result in a T_m increase.

more stable on ice." Once again, the broad general use of the word "stability" can be confusing.

If by more stable, we simply mean that proteins will last longer, and lose less of their activity over time, we are not addressing the thermodynamic idea of protein stability introduced in the previous paragraph. So, how does temperature affect the thermodynamic stability of proteins? The answer can once again be found by examining the Gibbs–Helmhotz equation (Eq. 4). When plotted, the relationship between ΔG_{unf} and temperature has a parabola-like shape (Figure 5).

Students can explore the effect of change of various parameters on the shape and properties of the stability curves by entering experimentally determined parameters into a spreadsheet that plots the stability curve (see Stability Curve Excel Sheet in Supplemental Material). By varying the parameters, students will soon realize that the shape does not change and that the temperature of maximal stability does not shift very much regardless whether the proteins are of thermophilic or mesophilic origin (Rees and Robertson, 2001). In fact, for most proteins the thermodynamic stability of a protein is lower on ice at 4°C (277 K) compared with room temperature 25°C (298 K). What this means is that a larger proportion of proteins are unfolded on ice than at room temperature.

Why is it then that we keep proteins on ice? The question is of relevance to anybody in experimental biochemistry or molecular biology, but the answer is not directly related to the thermodynamic stabilities of individual proteins. By keeping enzymes at lower temperatures, we decrease the rate at which harmful contaminants, such as oxidizing agents and proteases, destroy and deactivate our enzymes.

9. The More Stable a Protein Is, the Better It Will Function

Clearly, thermodynamic stability is important for protein function. It is easy to jump to the conclusion that higher stability might be beneficial to proteins. This is true to a certain extent. Stability is essential for maintaining proteins

in their native state, in a precise conformation needed for the function of that protein. However, stability of proteins needs to be balanced with flexibility. Proteins are dynamic molecules, and slight conformational changes within active sites, binding sites, and other regions of the protein are necessary for their functions. Interestingly, comparisons of homologous proteins from organisms that live at different temperatures show that these two properties, stability and flexibility, are finely balanced. In fact, regardless of the temperature at which host organisms live, homologous proteins from hosts as diverse as thermophiles and mesophiles tend to have similar thermodynamic stabilities at the hosts' optimal growth temperatures (Hollien and Marqusee, 1999).

Students can explore the scope and relevance of protein flexibility by modeling and characterizing the motions of proteins, by examining crystal structures of proteins captured in different conformational states (Gerstein and Echols, 2004). Protein visualization packages, mentioned above, can be used to calculate the overall RMSD differences between conformations. Students also can ask questions such as: What is the maximal displacement of any individual amino acid between two conformations? What is the average displacement of any mobile regions of the protein?

Students interested in visualization and modeling of molecular movements can be directed to the Database of Molecular Movements, a valuable resource for mapping and visualizing molecular movements (The Yale Morph Server, 2010).

10. Proteins Can Be Destabilized by Ligand Binding

Many proteins bind either small molecules or other protein binding partners. Ligand binding is a crucial step in the life cycle of many cellular and extracellular proteins. Ligand binding is important for cellular signaling cascades and regulatory pathways. Because protein interactions are so relevant in cell biology, it is important to understand the basic thermodynamics of ligand interactions.

How does ligand binding affect protein stability? Most proteins bind their binding partners in the native conformation. If ligand-bound native state of a protein were less stable than the ligand-free (apo) protein state, the ligand simply would not bind. The free energy (ΔG) of the following reaction,

$$P + L \rightarrow PL$$

(where P is the protein of consideration, L is the ligand, and PL is the protein–ligand complex) would be positive, and the reaction would not occur. Destabilization of protein by a ligand is possible only when the ligand binds to nonnative conformation. Considering this binding reaction in terms of equilibrium constants and free energy provides a framework for understanding that all natively folded ligand-bound proteins are more thermodynamically stable compared with the apo form.

CONCLUSIONS

Students often have a difficult time comprehending abstract ideas about protein structure, folding, and stability when they learn about them from textbooks and lectures only.

Modern research in protein science involves combining experimentation with computational modeling and analysis of complex data sets. Rather than introducing proteins only in terms of known facts and theories, we can better engage the students by providing them with opportunities to explore and investigate the same types of questions that drive the researchers. Experimental data, equations, and models can be introduced as early as introductory biology courses in college.

Exploration of relevant questions, and working with real data, bring excitement back to the classroom. One of the most rewarding aspects of using such approaches in teaching is when students take the given assignments to the next level, by raising their own questions and making their own discoveries. This type of learning helps students overcome misconceptions and gain a deeper understanding about proteins, and biological processes in general. In addition, the explorative learning experiences prepare the students for graduate school and work experiences, where they will face new questions in the ever-more interdisciplinary and increasingly more quantitative world of biological science research.

REFERENCES

Becktel, W. J., and Schellman, J. A. (1987). Protein stability curves. Biopolymers 26, 1859–1877.

Besenmatter, W., Kast, P., and Hilvert, D. (2007). Relative tolerance of mesostable and thermostable protein homologs to extensive mutation. Proteins 66, 500–506.

BioQUEST Curriculum Consortium (2010). The Biological ESTEEM Collection: Excel Simulations and Tools for Exploratory, Experiential Mathematics. http://bioquest.org/esteem/ (accessed March 2010).

DeLano, W. L. (2002). The PyMOL Molecular Graphics System. www.pymol.org (accessed March 2010).

Gerstein, M., and Echols, N. (2004). Exploring the range of protein flexibility, from a structural proteomics perspective. Curr. Opin. Chem. Biol. 8, 14–19.

Green, S. M., Meeker, A. K., and Shortle, D. (1992). Contributions of the polar, uncharged amino acids to the stability of staphylococcal nuclease: evidence for mutational effects on the free energy of the denatured state. Biochemistry 31, 5717–5728.

Guex, N., Diemand, A., Peitsch, M. C., and Schwede, T. (2008). Swiss-PdbViewer DeepView v4.0. www.expasy.org/spdbv (accessed March 2010).

Hollien, J., and Marqusee, S. (1999). A thermodynamic comparison of mesophilic and thermophilic ribonucleases H. Biochemistry 38, 3831–3836.

Jmol (2010). Jmol: an open-source Java viewer for chemical structures in 3D. www.jmol.org (accessed March 2010).

Khorasanizadeh, S., Peters, I. D., Butt, T. R., and Roder, H. (1993). Folding and stability of a tryptophan-containing mutant of ubiquitin. Biochemistry 32, 7054–7063.

Kohn, J. E. et al. (2004). Random-coil behavior and the dimensions of chemically unfolded proteins. Proc. Natl. Acad. Sci. USA 101, 12491–12496.

Kumar, S., Tsai, C. J., and Nussinov, R. (2000). Factors enhancing protein thermostability. Protein Eng. 13, 179–191.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2006). The ConSurf Server—Server for the Identification of Functional Regions in Proteins 2006. http://consurf.tau.ac.il/index.html (accessed March 2010).

Loladze, V. V., Ibarra-Molero, B., Sanchez-Ruiz, J. M., and Makhatadze, G. I. (1999). Engineering a thermostable protein via optimization of charge-charge interactions on the protein surface. Biochemistry 38, 16419–16423.

McCarney, E. R., Kohn, J. E., and Plaxco, K. W. (2005). Is there or isn't there? The case for (and against) residual structure in chemically denatured proteins. Crit. Rev. Biochem. Mol. Biol. 40, 181–189.

Pande, V. (2002). Folding@Home. http://folding.stanford.edu (accessed March 2010).

Pande, V. (2010). Summaries of Selected Journal Articles. www.stanford.edu/group/pandegroup/folding/education/papers/JMB02.html (accessed March 2010).

ProTherm (2010). ProTherm Thermodynamic Database for Proteins and Mutants. http://gibk26.bse.kyutech.ac.jp/jouhou/Protherm/protherm.html (accessed March 2010).

RCSB Protein Data Bank (2010). PDB Home Page. www.rcsb.org (accessed March 2010).

Rees, D. C., and Robertson, A. D. (2001). Some thermodynamic implications for the thermostability of proteins. Protein Sci. 10, 1187–1194.

Raabe, R., and Gentile, L. (2008). Thermal and chemical denaturation of Bacillus circulans xylanase. Biochem. Mol. Biol. Educ. 36, 428–432.

Serrano, L., Kellis, J. T., Jr., Cann, P., Matouschek, A., and Fersht, A. R. (1992). The folding of an enzyme. II. Substructure of barnase and the contribution of different interactions to protein stability. J. Mol. Biol. 224, 783–804.

Shortle, D., Stites, W. E., and Meeker, A. K. (1990). Contributions of the large hydrophobic amino acids to the stability of staphylococcal nuclease. Biochemistry 29, 8033–8041.

Snow, C. D., Zagrovic, B., and Pande, V. S. (2002). The trp cage: folding kinetics and unfolded state topology via molecular dynamics simulations. J. Am. Chem. Soc. 124, 14548–14549.

Ueda, T., Tamura, T., Maeda, Y., Hashimoto, Y., Miki, T., Yamada, H., and Imoto, T. (1993). Stabilization of lysozyme by the introduction of Gly-Pro sequence. Protein Eng. 6, 183–187.

The Yale Morph Server (2010). Molecular Movement Database. http://molmovdb.org (accessed March 2010).