Opportunities for Chemical Engineering Thermodynamics in Biotechnology: Some Examples

Siddharth S. Dey and John M. Prausnitz*

Department of Chemical Engineering, University of California, Berkeley, California 94720-1462

Because of its generality, thermodynamics is applicable to all substances, including biomacromolecules. To illustrate how thermodynamics can contribute to biotechnology, each of six examples gives a brief summary of pertinent, previously published research. Each example indicates that familiar concepts in chemical engineering thermodynamics can be applied to contribute toward solution of a practical problem. These examples are discussed here to encourage thermodynamically oriented chemical engineers to devote their talents toward helping to advance industrial biotechnology.

Introduction

Because thermodynamics has a wide range of applicability, it is one of the cornerstones of chemical engineering, for example, in the formulation of energy balances in plant design and for calculation of phase equilibria for separation operations. The last generation has produced an increasing extension of chemical engineering to include biotechnology. Can chemical engineering thermodynamics contribute to that extension? To illustrate that it can, this article presents six examples chosen from literature. While many more are available, these six examples serve to indicate that the thermodynamic concepts and tools developed for chemical engineering during the last 100 years are readily extended for aiding design and operation of biotechnological operations. In presenting these examples, no attempt is made to subject them to a critical analysis. They are given here primarily to encourage researchers in chemical engineering thermodynamics to turn their talents with confidence to assist the development of biotechnology.

It is appropriate to present these examples in the Festschrift dedicated to Stanley Sandler because he and his co-workers have been (and continue to be) pioneers in showing how chemical engineering thermodynamics can be applied to biomolecular engineering.

Example 1. Separation of Proteins and Viruses Using a Two-Phase Aqueous Micellar System

A common operation in biotechnology is the large-scale separation of desired proteins or viruses from the aqueous effluent of a bioreactor. Liquid–liquid extraction provides a well-established method that is easily scalable. However, because nonpolar solvents often denature biomolecules, conventional liquid–liquid extraction may not be suitable. Instead, as shown by Blankschtein and co-workers, separations of aqueous proteins or viruses can be achieved using a novel extraction system with two aqueous layers, one rich and the other poor in micelles.

In an aqueous solution of a surfactant, when the surfactant concentration exceeds its Critical Micellar Concentration (CMC), surfactant molecules self-assemble to form micelles that have hydrophobic tail groups forming the interior of the micelle and hydrophilic head groups in the aqueous solvent. Upon raising the temperature of a homogeneous micellar solution of nonionic surfactants, phase separation produces a two-phase aqueous system, with one phase rich and the other poor in micelles. To illustrate, Figure 1-1 shows the coexistence curve of an aqueous solution of the nonionic surfactant, C_{10}E_{4}, in the absence or presence of various proteins or viruses. Nonionic surfactants are preferred over ionic surfactants because they more readily exhibit phase separation and do not bind to proteins or cause denaturation. At low concentrations of proteins or viruses, the phase diagram in Figure 1-1 is not changed. For this case, Blankschtein et al.² presented a theoretical equation for the protein (or virus) partition coefficient, wherein it was assumed that the interactions between biomolecules and nonionic surfactant micelles are primarily repulsive due to their finite molecular size.

The partition coefficient, $K_z$, where subscript $z$ stands for the biomolecule (protein or virus), is defined as the ratio of the biomolecule concentration in the top ($C_{z,t}$, micelle-rich) phase to that in the bottom ($C_{z,b}$, micelle-poor) phase

$$K_z = \frac{C_{z,t}}{C_{z,b}} \quad (1-1)$$

For low biomolecule concentration and a nonionic surfactant, the partition coefficient is

$$K_z = \exp[-(\mu_{z,t}^0 - \mu_{z,b}^0) / k_B T] \quad (1-2)$$

where $k_B$ is the Boltzmann constant and $\mu_{z,t}^0$ and $\mu_{z,b}^0$ are standard-state chemical potentials of a biomolecule in the top and bottom phases, respectively. Because, by assumption, the biomolecules and the micelles interact only through repulsive excluded-volume interactions, the standard-state chemical potentials are a function of entropic factors only

$$\mu_{z,t}^0 - \mu_{z,b}^0 = -k_B T \ln \left( \frac{\Omega_t}{\Omega_b} \right) \quad (1-3)$$

where $\Omega_t$ and $\Omega_b$ represent the number of ways of placing a biomolecule in the top and bottom phases, respectively. Substitution of eq 1-3 into eq 1-2 gives

$$K_z = \left( \frac{\Omega_b}{\Omega_t} \right) \quad (1-4)$$

The number of ways to placing a biomolecule in each phase is proportional to the free volume available to the biomolecule;
this free volume can be estimated from molecular geometry. For two cases, the partition coefficient is then given by

\[ K_z = \exp \left[ -\frac{q_i - q_b}{1 + \frac{R}{R_b}} \right] \]

for spherocylindrical micelles

\[ K_z = \exp \left[ -\frac{q_i - q_b}{1 + \frac{R}{R_o}} \right] \]

for spherical micelles

where \( q_b \) and \( q_i \) are the volume fraction of surfactant in the bottom and top phases, respectively. \( R_z \) is the biomolecule radius, and \( R_b \) is the radius of a spherical micelle or the cross-sectional radius of a cylindrical micelle. These radii are obtained from crystallographic molecular-structure data.

Figure 1-2 shows experimental partition coefficients for different proteins as a function of temperature and protein radius; the experimental results are compared to those theoretically predicted based on the excluded-volume theory. Experimental and theoretical results are in good agreement. When \( K_z \) is less than 1, the protein preferentially partitions to the bottom (micelle-poor) phase. Figure 1-3 shows similar plots for the partition coefficients for different viruses. There are significant differences between the experimental and predicted results, and theory predicts separation much better than that observed.

Two possible reasons for this discrepancy are:

1. The theory does not take into account attractive interactions between the virus and the micelles. Because cellular membranes are lipid bilayers, similar to micelles, parts of the outer viral envelope may interact with micelles, increasing the viral concentration in the micelle-rich phase, giving a higher partition coefficient than that predicted. However, Blankschtein et al.\(^3\) showed that attractive interactions do not play a role in raising \( K_z \).

2. Phase-separation equilibrium is not attained. Entrainment of small domains of the micelle-poor (virus-rich) phase in the micelle-rich (virus-poor) phase is responsible for the higher observed \( K_z \). Because entrained domains cause light scattering, the turbidity of the two phases was used to show the presence of entrainment that decreases with a declining volume ratio of the two phases.

While molecular-thermodynamic concepts suffice to explain the experimental results for proteins, an additional non-thermodynamic concept is needed to obtain a useful correlation of the experimental results. This example illustrates that in many practical cases in biotechnology, molecular thermodynamics is often helpful but not sufficient.

If \( x \) is the volume fraction of the micelle-poor domains in the micelle-rich phase and \( C_{v,rich}^E \) and \( C_{v,poor}^E \) are the concentration of virus in the micelle-rich and micelle-poor domain, respectively, the overall concentration of virus in the top micelle-rich phase is

\[ C_{v,t} = (1 - x)C_{v,rich}^E +xC_{v,poor}^E \]

Rearranging gives

\[ C_{v,t} = C_{v,rich}^E + x(C_{v,poor}^E - C_{v,rich}^E) \]

Assuming that the virus partitions between the micelle-rich and micelle-poor domain according to the excluded-volume theory, \( K_v^E = (C_{v,rich}^E)/(C_{v,poor}^E) \).

Figure 1-3 shows that \( K_v^E \approx 1 \) for viruses. Therefore

\[ C_{v,t} = C_{v,rich}^E + xC_{v,poor}^E \]

Because \( C_{v,poor}^E \approx C_{v,rich}^E \), even a small \( x \) can dramatically change the concentration of the virus in the top micelle-rich phase.

A similar analysis for the micelle-poor phase gives

\[ C_{v,b} = C_{v,poor}^E + y(C_{v,rich}^E - C_{v,poor}^E) \]

where \( y \) is the volume fraction of the micelle-rich domains in the micelle-poor phase.

Since \( y \) is small

\[ C_{v,b} = C_{v,poor}^E \]

The concentration of virus in the micelle-rich and that in the micelle-poor phase give a modified equation for the partition coefficient

\[ K_v = K_v^E + x(1 - K_v^E) \]
Figure 1-4 shows good agreement for the predicted and experimental partition coefficients for a virus.

Figures 1-3 and 1-4 show that, prior to interpreting data using thermodynamic analysis, it is of utmost importance to ensure that equilibrium has been attained; in the absence of equilibrium, semi-empirical corrections are necessary. Nevertheless, thermodynamic analysis can provide useful guidance; modified thermodynamic analysis of the two-phase aqueous micellar system described above may provide a useful method for designing a process for separating therapeutic proteins and, perhaps, for purification of viruses for large-scale production of gene delivery viral vectors.

This example shows that, in practice, thermodynamic analysis can sometimes provide only a partial explanation of the observed phenomena. For a more complete analysis, it may be necessary to consider other factors, in this case, entrainment.

Example 2. Isothermal Titration Calorimetry for AIDS Drug Development

Much effort has been directed toward finding drugs to help victims of the AIDS epidemic caused by the HIV virus. Thermodynamics makes a contribution to that effort through thermodynamic interpretation of experimental calorimetric titration data.

As shown in Figure 2-1, after the HIV virus infects a human cell, a series of steps produces replication of HIV. The infected cell expresses a polyprotein and an HIV protease; the role of
the protease is to cleave the polyprotein. The cleaved proteins assemble to form a new HIV virus.

An effective strategy to prevent formation of a new virus is provided by introducing a drug that deactivates the HIV protease. This drug, called a protease inhibitor, prevents cleavage because it has two rotating bonds, while drug 1 has only one, as indicated by dashed arrows. Asymmetry and flexibility provide the drug with additional conformations that can adapt to a mutant HIV binding site.

Deactivation of a protease by an inhibitor can be described by a traditional lock-and-key mechanism indicated in Figure 2-3. The inhibitor must have the correct shape to fit into the active site of the HIV protease; here, the inhibitor is the “key” that must fit into the protease “lock”.

However, because of mutations, the active site of an HIV protease can exist in a variety of forms, as shown schematically in Figure 2-3; the shape of the active site of the wild type (the lock) is represented by a cross, while for the mutant the shape of the lock is represented by a hexagon. We seek an inhibitor (the key) that is sufficiently flexible such that it “fits” into both types of locks. We need a drug (the key) that can bind not only to the active site of the wild-type protease but also to that of its mutant. Thermodynamics can help to identify the best drug candidates.

To illustrate, Figure 2-4 shows two possible drug candidates, 1 and 2. Drug 2 has better adaptability than drug 1 because, compared to drug 1, it has asymmetrical functionality; the toluene group has less symmetry than the tertiary butyl group, as indicated by oval circles. Further, drug 2 is more flexible because it has two rotating bonds, while drug 1 has only one, as indicated by dashed arrows. Asymmetry and flexibility provide the drug with additional conformations that can adapt to a mutant HIV binding site.

Toward obtaining a quantitative measure of drug efficacy, Ohtaka and Freire use thermodynamic analysis of data from isothermal titration calorimetry (ITC).

Let $A$ stand for the HIV protease and $B$ for the (drug) inhibitor. We define a dissociation constant $K_d$ and its reciprocal, association constant $K_a$, where subscript $a$ stands for association and subscript $d$ for dissociation.

$$K_d = \frac{[A][B]}{[AB]}; \quad K_a = \frac{[AB]}{[A][B]}; \quad K_a = (K_d)^{-1} \quad (2-1)$$

where $[\ ]$ stands for concentration in water. For good drug efficacy, we want $K_d$ to be small or, equivalently, $K_a$ to be large.

In the reaction $A + B \rightarrow AB$, the binding of $A$ (protease) and $B$ (inhibitor) is determined by standard-state enthalpy $\Delta H^o$ and entropy $\Delta S^o$ through the thermodynamic relation

$$-RT \ln K_a = \Delta G^o = \Delta H^o - T \Delta S^o \quad (2-2)$$

where superscript $o$ indicates a standard state.

Isothermal titration calorimetry (ITC) provides experimental results for $K_d$ (or $K_a$) and $\Delta H^o$. From these results, we calculate $\Delta G^o$ and $\Delta S^o$.

Figure 2-5 shows experimental results for $\Delta G^o$, $\Delta H^o$, and $-T \Delta S^o$ for a first-generation set of 12 drug candidates where the protease is of the wild-type.

In Figure 2-5, the first generation of drugs contains molecules that are relatively rigid (not flexible). Within the active-site pocket of the protease, rigid molecules cannot interact strongly with the functional groups of the protease, and therefore, for these drugs, $\Delta H^o$ is not favorable; in some cases, while negative, the absolute value is small, and in other cases, most unfavorable, $\Delta H^o$ is positive. However, $-T \Delta S^o$ is favorable because upon
binding the hydrophobic drug releases hydrated water molecules with a subsequent desirable gain in entropy.

Because rigid drug molecules do not have as much adaptability to mutant proteases as flexible drug molecules, second-generation inhibitors are relatively flexible drugs. As shown in Figure 2-5, these drugs show a favorable (negative) $\Delta H^0$, but $-T\Delta S^0$ is now less favorable because of what is called the entropic penalty: when a flexible key enters a confining lock, it loses not only translational degrees of freedom (as does a rigid key) but also rotational degrees of freedom. These losses in freedom produce a decrease in entropy; because this decrease is unfavorable, flexible drugs exhibit a greater entropic penalty (smaller $-T\Delta S^0$) than rigid drugs. However, the favorable trade-off for the unfavorable entropic penalty is a stronger binding interaction (i.e., a large negative $\Delta H^0$). Figure 2-5 suggests that drugs 11 and 12 are best because $\Delta G^0$ and $\Delta H^0$ are both large and negative, as desired.

Because protease mutations are common, we seek an inhibitor that is not only effective for the wild-type protease but also adaptive to at least some of its mutants.

Figure 2-6 shows ITC results for a series of inhibitors interacting with the wild type and with a particular mutant. The ordinate, on a logarithmic scale, shows the dissociation constant $K_d$; we want $K_d$ to be small not only for the wild type but also for the mutant.

Figure 2-6 shows the importance of having a large, negative $\Delta H^0$. The more adaptive second-generation drugs (e.g., inhibitors 11 and 12) have the smallest dissociation constant $K_d$; we want $K_d$ to be small not only for the wild type but also for the mutant.

Figure 2-6 shows that,
while inhibitors 11 and 12 are good for the wild-type protease inhibitor 11 is superior for inhibiting the mutant protease. Figure 2-6 suggests that drug 11 is better than drug 12.

For a flexible inhibitor (as opposed to a rigid inhibitor), we must pay the price of a larger entropic penalty; this penalty is undesirable because it lowers the absolute value of a negative $\Delta G^\circ$. We want an inhibitor that is flexible enough to give a large negative $\Delta H^\circ$ and also useful for a protease mutant. The inhibitor should not be too flexible lest the entropic penalty become too large. This example shows that thermodynamic studies are useful for guidance toward identifying an optimum HIV inhibitor.

Example 3. Protein Purification: Effect of Salt on the Phase Diagram for an Aqueous Solution of Globular Proteins

In biotechnology, initial separation of a target protein from a bioreactor broth is often achieved by adding a salt to induce precipitation. To provide guidance for attaining the desired results, we require a pertinent phase diagram, i.e., a plot of temperature vs protein concentration (or its equivalent, protein density) designated by number density $\rho$. A common way to express protein density is provided by the protein packing fraction

$$\eta = \frac{\pi}{6} \rho \sigma_p^3$$

where $\sigma_p$ is the protein diameter.

While a bioreactor broth may contain several types of biomacromolecules in addition to other small solutes (e.g., a buffering salt), attention here is focused on a target protein, neglecting the presence of other biomacromolecules. The effect of other small solutes is included in the potential of mean force.

For fluid—solid equilibria, at constant temperature $T$, the essential equations of equilibrium are

$$\mu^F(\eta^F, T) = \mu^S(\eta^S, T)$$ (3-1)

$$p^F = p^S$$ (3-2)

where $\mu^F$ is the chemical potential of the protein in the fluid phase; $\mu^S$ is the chemical potential of the protein in the solid phase; and $p$ is pressure.

At constant temperature, pressure $p$ and chemical potential $\mu$ are related to density $\rho$ by well-known thermodynamics; both are derived from an expression for Helmholtz energy $A$ as a function of temperature and $\rho$ (or $\eta$), as discussed in numerous textbooks.

A convenient procedure for calculating $A^F$ for the fluid and $A^S$ for the solid is to use the first-order Barker–Henderson perturbation theory, one for the fluid phase and another for the solid phase, as discussed, for example, by Tavares and Prausnitz. For each phase, the chemical potential is obtained from two contributions, a hard-sphere ($hs$) term and a perturbation term ($P$)

$$\mu = \mu^{hs} + \mu^p$$ and $P = P^{hs} + P^p$

For the fluid phase, $\mu^{hs}$ is obtained from the Carnahan–Starling equation of state. For the solid phase, $\mu^{hs}$ is obtained from the equation of state proposed by Velasco and Mederos. To calculate $\mu^p$, we need an expression that provides information concerning the protein–protein interaction in the aqueous solution and in the solid precipitate. That expression is the potential of mean force $W(r)$ where $r$ is the center-to-center distance between two quasi-spherical (globular) protein particles. To a reasonable approximation, as suggested by colloid theory, this potential of mean force contains the following contributions

$$W(r) = W(r)^{hs} + W(r)^{el} + W(r)^{att} + W(r)^{ons}$$ (3-3)

where el refers to electrostatic forces and att refers to attractive (dispersion) forces.

$W^{hs}$ depends on the reduced distance $r/\sigma_p$. The potential of mean force depends on the concentration and nature of the precipitating salt.

When a salt is present in the aqueous protein solution, $W^{el}$ depends on $r/\sigma_p$, $\sigma_p$, $q_p$, $D$, and $T$ where $\sigma_p$ is the salt diameter, $q_p$ is the protein charge, $q_s$ is the salt charge, and $D$ is the dielectric constant of water. Protein charge $q_p$ depends on pH. $W^{att}$ depends on $r/\sigma_p$, on $H$, and on a new expression, $q'(ions)$. Here, $H$ is the Hamaker constant that reflects dispersion (van der Waals) forces between two aqueous protein particles in a salt-free medium; $W(r)^{ons}$ provides the contribution to protein–protein attractive forces due to dispersion forces arising from van der Waals interactions between the ions and between the ions and the protein particles. It is this function, $q'(ions)$, that can explain the well-known Hofmeister series. (The Hofmeister series lists ions in decreasing order in their ability to salt-out dissolved proteins.) While the Hofmeister series was discovered empirically more than 100 years ago, its quantitative (theoretical) formulation was derived only recently by Ninham and co-workers and by Tavares et al. In this formulation, the function $q'(ions)$ depends on the aqueous polarizabilities of the protein and the salt ions. An alternate explanation for the Hofmeister effect is based on selective hydration as discussed, for example, by Collins et al. and by Hummer and Garde et al.

For typical protein solutions, a decrease in temperature may show two fluid phases because high temperature encourages random mixing. In some cases, these fluid phases may be metastable. At constant temperature, the compositions of these two phases (‘$\gamma$’ and ‘$\gamma''$) are related by:

$$\mu^{F}(\eta^{F'}, T) = \mu^{F}(\eta^{F''}, T)$$ (3-4)

$$\rho' = \rho''$$ (3-5)

At a fixed temperature, we have two unknowns, viz., the densities of the two equilibrium phases. They are obtained by simultaneous solution of eqs 3-1, 3-2, 3-4, and 3-5.

Figure 3-1 shows the important effect of the salt’s anion for lysozyme equilibria. NaCl and NaI show qualitatively similar behavior; we see a fluid—fluid region with an upper critical point on the left side and a fluid—solid region on the right. However, because the polarizability of the iodide ion is larger than that

![Figure 2-6. Dissociation constant for 10 inhibitors interacting with the wild-type HIV protease and one mutant form of the protease (data for inhibitor 8 omitted). Inhibitor 11 is superior to inhibitor 12 because $K_d$ is smaller for both the wild-type and the mutant protease. Adapted from ref 5.](image-url)
of the chloride ion and because higher polarizabilities increase attractive forces, the critical temperature for NaI is appreciably higher than that for NaCl.

For NaSCN, we obtain a phase diagram that is qualitatively different from those for NaCl and NaI because the polarizability of the SCN ion is larger than those for the iodide and chloride ions. For NaSCN, we do not have a stable fluid–fluid region; only fluid–solid equilibria are thermodynamically stable. However, in addition to fluid–solid equilibria, we also calculate metastable fluid–fluid equilibria, consistent with experiment. The calculations show a metastable region (dashed line where \( \mu^F = \mu^P \)), where superscript ‘ refers to a dilute aqueous lysozyme solution and superscript ” refers to a concentrated aqueous lysozyme solution.

For solutions containing NaCl or NaI, fluid–fluid equilibria are stable. However, for lysozyme solutions containing NaSCN, \( \mu^F \) and \( \mu^P \) are larger than the corresponding \( \mu^F \) at the same temperature. For this case, the calculated (and observed) fluid–fluid equilibria are metastable.

Because of the larger polarizability of the SCN ion, these ions are attracted more strongly to the positively charged lysozyme, reducing the repulsive forces between lysozyme particles. Thus, for lysozyme, at pH 4.3, NaSCN is a more effective precipitating salt than NaCl or NaI. The thermodynamic discussion outlined here may provide initial guidance for designing an effective precipitation process.

**Example 4. Liquid–Liquid Equilibria Suggest a Possible Method to Prevent Cataracts in the Human Eye**

Cataracts are formed by aggregation of proteins in the eye. Small assemblies of protein molecules produce mesoscopic particles that interfere with normal vision. To obtain a quantitative description of protein aggregation, it is useful to establish a phase diagram that gives the coexistence curve (binodal) for liquid–liquid equilibria; phase separation indicates aggregation. As shown by Benedek et al.,\(^{15}\) the coexistence curve can be modified by a small amount of an additive that lowers the upper critical solution temperature below body temperature. By preventing protein aggregation, it is possible to prevent formation of cataracts.

Condensation of proteins results in eye-lens turbidity that can be quantified by measuring the intensity of light scattered by the lens. As shown by Thurston et al.,\(^ {16}\) the intensity of light scattered (\( I_{\text{scatt}} \)) is given by

\[
I_{\text{scatt}} = I_0 \exp(t/\Delta Z)
\]

Figure 3-1. Three calculated phase diagrams for lysozyme at pH 4.3 where the lysozyme charge, \( q_y = +10 \). Each diagram is for an aqueous lysozyme solution containing a sodium salt with concentration 0.2 M. The protein diameter, \( a_y \), is 3.3 nm. \( T_0 = 298 \text{ K} \). In conventional theory, where the polarizabilities of protein and salt ions are neglected, these three phase diagrams are identical.

Figure 4-1. Intensity of light scattering from eye vs the age of a person. Straight solid line indicates experimental data. Dotted line shows a hypothetical case where the time constant \( \Delta Z \) is increased by 20%, possibly by the addition of a pharmacological agent. This rise in \( \Delta Z \) increases the age where 50% of the population would develop cataracts from 75 to 90 years. Adapted from ref 15.

where \( I_0 \) is the intensity of incident light; \( t \) is the age of a person; and \( \Delta Z \) is a time constant obtained from experimental data. Figure 4-1 shows that by addition of a cataract inhibitor, a 20% increase in \( \Delta Z \) can prevent the onset of cataracts. Thermodynamic studies can establish the effect of an inhibitor on the binodal curve.

Calf eye-lens proteins have an amino acid sequence similar to those in the human eye-lens. These proteins, called \( \gamma \)-Crystallins, have been studied extensively. They have been classified into two major groups based on their aqueous critical-solution temperatures (\( T_c \)):

1. High-\( T_c \) proteins: \( \gamma_{\text{IIa}} (\gamma_c) \) and \( \gamma_{\text{IVa}} (\gamma_b) \) Crystallins with \( T_c \approx 38 ^\circ \text{C} \).
2. Low-\( T_c \) proteins: \( \gamma_{\text{II}} (\gamma_b) \) and \( \gamma_{\text{INa}} (\gamma_b) \) Crystallins with \( T_c \approx 5 ^\circ \text{C} \).

Figure 4-2 shows binary aqueous phase diagrams of temperature (\( T \)) vs protein volume fraction (\( \phi \)) or concentration (\( C \)) for the four \( \gamma \)-Crystallins.\(^ {17}\) The region above the coexistence curve represents a single-phase region, and that below the coexistence curve represents a two-phase region. The critical temperature \( T_c \) provides a key parameter because it represents the maximum phase-separation temperature. Addition of a pharmacological agent that can lower \( T_c \) below the body temperature can reduce protein aggregation in the eye.

In the critical region (within 10 \(^\circ \text{C} \) of \( T_c \)), the coexistence curves can be fitted by the scaling relation

\[
\frac{C_c - C}{C_c} = J \left[ \frac{T_c - T}{T_c} \right]^f
\]

where

\[
J = \left( \frac{C_c - C}{C_c} \right) \left( \frac{T_c - T}{T_c} \right)^{-f}
\]

where \( C_c \) is the critical protein concentration; \( T_c \) is the absolute critical temperature; exponent \( \beta = 0.325 \) follows from the three-dimensional Ising model; and \( J \) is a parameter that characterizes the width of the coexistence curve. \( C_c \) and \( J \) are nearly identical for all four binary aqueous \( \gamma \)-Crystallin solutions with a \( C_c = 289 \pm 20 \) mg/mL and \( J = 2.6 \pm 0.1 \). The Ising exponent \( \beta \) gives a much better fit than that obtained by using the mean field (classical) exponent \( \beta = 0.5 \).

To understand interactions in a binary protein—water mixture, Taratuta et al.\(^{18} \) presented a molecular thermodynamic perturbation model where the reference system is an assembly of hard spheres dispersed in the continuous aqueous phase. The Gibbs free energy \( G \) for this system is given by

\[
\frac{G}{V} = \frac{G_o}{V} + \frac{k_B T}{\Omega_p} \left[ \phi \ln \phi - \frac{\phi - 6\phi^2 + 4\phi^3}{(1 - \phi)^2} \right] - \frac{k_B T}{\Omega_p} U \phi^2
\]

where \( V \) is the volume of the solution; \( G_o \) is the standard Gibbs free energy; \( \Omega_p \) is the volume of a protein molecule; \( k_B \) is Boltzmann’s constant; \( T \) is absolute temperature; \( \phi \) is the volume fraction of protein; and \( U \) is a dimensionless parameter that quantifies aqueous protein—protein, protein—water, and water—water interactions. \( G_o \) is linear in \( \phi \), and it is a measure of the Gibbs energy change of the solution when a single protein molecule is added to a pure solvent. The second term follows from the well-known Carnahan—Starling equation; it represents the entropy of mixing for hard spheres. The last term gives the mean-field approximation of the dissolved protein particles. The model shown in eq 4-3 for binary solutions predicts the critical volume fraction as 13%, while the mean experimental volume fraction for the four binary mixtures was 20.5%.

Because cataract formation in the eye is due to the interaction of several different lens proteins, a study of multicomponent aqueous—protein solutions is useful for understanding the parameters that influence phase separation. Liu et al.\(^{19} \) have extensively studied aqueous two-protein solutions to obtain the phase separation temperatures as a function of overall protein volume fraction and protein composition. Figure 4-3 shows the phase separation temperature as a function of the total protein volume fraction at different fixed compositions for the aqueous two-protein solution containing \( \gamma_{IVa} \) and \( \gamma_{IIIa} \) Crystallins. Liu et al.\(^{19} \) present a molecular-thermodynamic model for an aqueous mixture containing more than one protein. Liu’s study of such ternary systems has shown that a native \( \gamma \) Crystallin may play a significant role in maintaining transparency of the eye lens.\(^{20} \) Figure 4-4 shows that the critical phase-separation temperature of aqueous \( \gamma_{IVa} \) Crystallin in the presence of \( \gamma \) Crystallin is substantially reduced at high total volume fractions.

The purpose of a small-molecule cataract inhibitor is to reduce the energy of attraction between the eye-lens proteins. Several small-molecule inhibitors have been proposed in the literature that reduce condensation of the lens proteins. One such chemical inhibitor is oxidized glutathione (SSG) that forms a disulfide complex with \( \gamma_{IVa} \) Crystallin. Figure 4-5 shows that upon addition of SSG the phase-separation temperature of the \( \gamma_{IVa} \) Crystallin—SSG complex is significantly lower than that of \( \gamma_{IVa} \) Crystallin without SSG.\(^{15} \) Because SSG reduces attraction between Crystallin molecules, it decreases their tendency to aggregate.

This example illustrates how liquid—liquid phase-equilibrium thermodynamics can contribute to a possible therapy for preventing cataracts.

**Example 5. Calculation of Biomass Yield from Gibbs Energy Dissipation**

In a typical biochemical process, microbes are used to produce a desired product. Thermodynamics can sometimes help to
predict the yield of that product. In a typical application, we want to make biomass that may be produced by cells, or it may be the cells themselves, depending on whether the product is expelled or retained within the cell. As in any chemical reaction, the product yield is related to the Gibbs energy dissipation. Following the discussion by von Stockar et al.,21 this relationship is best shown by an example.

Methanobacterium thermoautotrophicum is an anaerobic bacterium that generates the energy required for biomass synthesis by methanogenesis, a process that converts $\text{H}_2$ and $\text{CO}_2$ to $\text{CH}_4$, $\text{H}_2\text{O}$, and biomass. In an aqueous broth, the growth of biomass is described by

$$\text{H}_2 + Y_{\text{CD}}\text{CO}_2 + Y_{\text{CD}}\text{NH}_3 \rightarrow Y_{\text{XD}}\text{CH}_4O_{0.39}N_{0.24} + Y_{\text{pD}}\text{CH}_4 + Y_{\text{wD}}\text{H}_2\text{O} \quad (5-1)$$

where D, C, N, X, P, and W stand for $\text{H}_2$, $\text{CO}_2$, $\text{NH}_3$, biomass, $\text{CH}_4$, and $\text{H}_2\text{O}$, respectively. $\text{NH}_3$ provides the microbes’ need for nitrogen. $Y_{\text{pD}}$ is called the yield coefficient for reactant or product i, defined as the ratio of the moles of component i to the moles of component D. In eq 5-1, yield coefficients are identical to stoichiometric coefficients. The elemental composition of the biomass is obtained experimentally.

The biomass yield, $Y_{\text{XD}}$, is related to the Gibbs energy of dissipation per C-mol biomass produced, equivalent to $\Delta G_X$, the Gibbs energy of reaction per C-mol biomass produced

$$\Delta G_X = \frac{\Delta G_D}{Y_{\text{XD}}} = \frac{\Delta G_{D,\text{aq}}}{Y_{\text{XD}}} + \frac{\gamma_D}{Y_{\text{pD}}} (\Delta G_{\text{C,\text{aq}}} - \Delta G_{\text{C,\text{aq}}}) -$n\right) + \frac{\gamma_X}{Y_{\text{pD}}} (\Delta G_{\text{P,\text{aq}}} - \Delta G_{\text{P,\text{aq}}}) \quad (5-2)$$

A C-mol refers to the number of moles of carbon. Thus, there are six C-moles in one mole of benzene. From eq 5-2, $Y_{\text{XD}}$ can be calculated for a given $\Delta G_X$. In general, to estimate the biomass yield coefficient from eq 5-2, $\Delta G_X$ is obtained from correlations. Heijnen and van Dijken22 established a correlation for chemotrophs, that is, for microbes that obtain energy from a chemical source

$$-\Delta G_X^o = 200 + 18(6 - n) + 18 + \exp\{[(3.8 - \gamma_C)^2]^{0.16}(3.6 + 0.4n)\} \quad (5-3)$$

For chemotrophs with reverse electron transport, the following correlation was proposed

$$-\Delta G_X^o = 3500 \quad (5-4)$$

where $\Delta G_X^o$ is in kJ/C-mol, $\gamma_C$ and $n$ are the degree of reduction and the number of carbon atoms of the carbon source, respectively.

Liu et al.23 proposed a simple correlation for $\Delta G_X^o$

$$\Delta G_X^o = \begin{cases} -666.2 \gamma_D - 243.1 \quad (\gamma_D \leq 4.67) \\ -157\gamma_D + 339 \quad (\gamma_D > 4.67) \quad (5-5b) \end{cases}$$

$$-\Delta G_X^o = -3500 \quad (5-5c)$$

where $\Delta G_X^o$ is in kJ/C-mol and $\gamma_D$ is the degree of reduction of the electron donor.

The other terms in eq 5-2, defined below, can be calculated if the following data are obtained experimentally or from the literature:

1. Concentrations of $\text{H}_2$, $\text{CO}_2$, and $\text{CH}_4$ in the off-gas are obtained experimentally using gas chromatography.
2. Henry’s constants ($H_i$) for $\text{H}_2$, $\text{CH}_4$, and $\text{CO}_2$ in water.
3. Standard Gibbs energies of combustion in the gaseous phase for all components. [The Gibbs energy of combustion for biomass is obtained based on a correlation presented by Sandler et al.24]

Knowing the gas phase partial pressures $p_i$ and Henry’s constants $H_i$, the aqueous phase concentrations are obtained from

$$c_i = \frac{p_i}{H_i} \quad (i = \text{CH}_4, \text{H}_2, \text{CO}_2) \quad (5-6)$$

The cell wall of the bacterium is taken as the boundary for the system wherein $\text{H}_2$ and $\text{CH}_4$ are in a dilute aqueous solution. The Gibbs energies of combustion for $\text{H}_2$ and $\text{CH}_4$ in the aqueous state are given by

$$\Delta G_{\text{i,aq}} = \Delta G_{\text{i,aq}}^o + RT \ln c_i \quad (i = \text{CH}_4, \text{H}_2, \text{CO}_2) \quad (5-7)$$

where $\Delta G_{\text{i,aq}}^o$ is the standard aqueous phase Gibbs energy of combustion and $c_i$ is the concentration of component $i$ in the dilute aqueous solution. Equation 5-7 does not require an activity coefficient because at normal pressures the solubilities of $\text{H}_2$, $\text{CH}_4$, and $\text{CO}_2$ in the aqueous solution are very small.

Using standard thermodynamics, we obtain

$$\Delta G_{\text{i,q}} = \Delta G_{\text{i,\text{aq}}}^o + RT \ln H_i \quad (5-8)$$

Knowing the concentrations of $\text{H}_2$, $\text{CH}_4$, and $\text{CO}_2$ in the aqueous phase ($c_i$) and calculating the standard Gibbs energy of combustion in the aqueous phase ($\Delta G_{\text{i,\text{aq}}}$) from eq 5-8, the Gibbs energy of combustion in the aqueous phase ($\Delta G_{\text{\text{aq}}}$) can be calculated from eq 5-7. Because the standard Gibbs energies are generally much larger than the concentration-dependent term, this concentration-dependent term can often be neglected without significant loss of accuracy. Thus, the Gibbs energy of reaction (also called Gibbs energy of dissipation) for the reaction in eq 5-1, denoted by $\Delta G_r$, is

$$\Delta G_r = \Delta G_{\text{D,\text{aq}}} + Y_{\text{CD}}\Delta G_{\text{C,\text{aq}}}^o - (Y_{\text{X,D}}\Delta G_X^o + Y_{\text{pD}}\Delta G_{\text{P,\text{aq}}}^o) \quad (5-9)$$
where superscript * indicates that instead of N₂, NH₃ is chosen as the reference state for the Gibbs energies of combustion. To relate the yield coefficients in eq 5-9 to each other, we use an elemental mass balance and a degree of reductance balance. The degree of reductance (γ) for a compound is defined as the number of equivalents of available electrons per gram of atom of carbon of that substance. The number of equivalents of available electrons per gram of atom of carbon is then the degree of reductance for the compound. The number of equivalents of available electrons per gram of atom of carbon is then the number of equivalents of available electrons per gram of atom of carbon. The number of equivalents of available electrons per gram of atom of carbon of that substance is then the degree of reductance for the compound. The number of equivalents of available electrons per gram of atom of carbon of that substance is then the degree of reductance for the compound.

The degree of reductance balance is obtained from conservation of available electrons

\[ y_D + Y_{CD} y_C + Y_{ND} y_N = Y_{XD} y_X + Y_{PD} y_P + Y_{WD} y_W \]

\[ y_D = \frac{y_D}{y_P} = \frac{y_X}{y_P} Y_{XD} \]  

(5-11)

An elemental-carbon balance for the system gives

\[ Y_{CD} = Y_{XD} + Y_{PD} \]  

(5-12)

Combining eqs 5-9, 5-11, and 5-12 gives the required relation between the biomass yield and the Gibbs energy of dissipation per C-mol biomass produced, as shown in eq 5-2.

Figure 5-1 shows the calculated relation between the Gibbs energy of dissipation and the desired yield coefficient \( Y_{XD} \). For a biomass yield, \( Y_{XD} \sim 0.015 \text{ mol/mol of H}_2 \), and a mass balance gives the other yields in units of mol/mol H₂:

\[ Y_{ND} = 3.6 \times 10^{-3} \]
\[ Y_{PD} = 0.2572 \]
\[ Y_{WD} = 0.2422 \]
\[ Y_{PD} = 0.5085 \]

Thus, for every mole of H₂ consumed, the microbe produces 0.2422 mol of CH₄. Figure 5-1 shows a frequently observed result that the experimental biomass yields are closely clustered together in the middle, rather than at the extremes. A qualitative explanation for this result is provided by splitting eq 5-1 into two reactions—catabolism and anabolism.

Catabolism

\[ H_2 + aCO_2 \rightarrow bCH_4 + cH_2O \quad \Delta G_{cat} \]  

(5-13a)

Anabolism

\[ aCH_4 + eH_2O + Y_{ND}NH_3 \rightarrow Y_{XD}X + fCO_2 \quad \Delta G = Y_{XD} \Delta G_{ana} \]  

(5-13b)

where, for eqs 5-13a and 5-13b, \( \Delta G_{cat} \) is the Gibbs energy of reaction per C-mol of H₂ for catabolism and \( \Delta G_{ana} \) is the Gibbs energy of reaction per C-mol biomass produced for anabolism. Here, \( a, b, c, f \) are stoichiometric coefficients for the catabolism and anabolism reactions; \( \Delta G_{cat} \) and \( \Delta G_{ana} \) are related to the overall Gibbs energy of reaction (or Gibbs energy of dissipation) by

\[ \Delta G_D = \Delta G_{cat} + Y_{XD} \Delta G_{ana} \]  

(5-14)

Catabolism is an energy-producing reaction, with a negative \( \Delta G_{cat} \), and anabolism is an endothermic reaction with a positive \( \Delta G_{ana} \). In an inefficient growth system (where \( |\Delta G_{cat}| > |\Delta G_{ana}| \)), the Gibbs energy of dissipation is also a large negative quantity, and eq 5-14 tells us that the biomass yield is small and not desirable for the microbe. At the other extreme, an efficient growth system (where \( |\Delta G_{ana}| > |\Delta G_{cat}| \)), the biomass yields are optimal; however, the Gibbs energy of dissipation is small and hence the rate of the overall reaction is likely to be slow. A slow reaction rate is undesirable because under these conditions the microbe will be outgrown by competitors. Hence, microbes have possibly evolved to have a \( \Delta G_X \) and \( Y_{XD} \) somewhere in between the two extremes of inefficient growth (very high \( \Delta G_{cat} \) and very low \( Y_{XD} \)) and efficient growth (very low \( \Delta G_{ana} \) and very high \( Y_{XD} \)), as indicated in Figure 5-1.

von Stockar et al. and others have presented analyses similar to that above for a variety of relatively simple bioreactions using a large number of aerobic and anaerobic microbes. These analyses indicate that thermodynamics may provide good estimates of the biomass yield for simple microbial systems, but further developments are required to apply these methods to complex mammalian cell cultures.

Example 6: Prevention of Liquid–Liquid Phase Separation in Freeze-Concentrated Formulations for Protein Lyophilization

Freeze-drying (lyophilization) is a process for removing water from a solution in the form of ice; the advantage of this process is that it produces a nearly anhydrous solute. Freeze-drying is a popular method of drying in the pharmaceutical industry because it increases the shelf life of a drug. Freeze-drying...
enhances the chemical stability of biomolecules. Aqueous biomolecules are susceptible to numerous unwanted reactions like hydrolysis, cross-linking, oxidation, aggregation, and disulfide rearrangements. In any drying process, it is necessary to prevent protein denaturing. To increase the stability of a protein in solution, it is common practice to add excipients to the solution. Excipients are also used to promote stabilization of freeze-dried proteins. However, the formation of ice during freeze-drying increases the concentration of excipients in the liquid. Unlike dilute solutions, the high concentration of excipients during freeze-drying increases the tendency for undesired liquid—liquid phase separation. Phase separation of these excipients into two phases can denature the protein. Because phase separation must be avoided, it is useful to determine those conditions when such separation is likely. As shown by Randolph and co-workers, thermodynamics provides guidance toward determining those conditions.27

For a dilute or semidilute aqueous system with two polymeric excipients, denoted by subscripts 2 and 3, and water denoted by subscript 1, the chemical potentials (μ) for the excipients are given by

\[
\mu_2 - \mu_2^0 = RT(\ln m_2 + a_{22}m_2 + a_{23}m_3) \quad (6-1)
\]

\[
\mu_3 - \mu_3^0 = RT(\ln m_3 + a_{32}m_2 + a_{33}m_3) \quad (6-2)
\]

where \(m_i\) is the molar concentration of polymer \(i\) and superscript 0 indicates the (infinite-dilution) standard state chemical potential; \(a_{22}, a_{33},\) and \(a_{23}\) are coefficients that characterize 2–2, 3–3, and 2–3 interactions. The chemical potential of water is calculated by combining eqs 6-1 and 6-2 with the Gibbs–Duhem equation

\[
\mu_i - \mu_i^0 = -RTV_i\rho_i(m_i + m_1 + a_{22}(m_2)^2 + a_{33}(m_3)^2 + a_{23}m_2m_3) \quad (6-3)
\]

where \(V_i\) and \(\rho_i\) are the molar volume and mass density of pure water, respectively, and \(\mu_i^0\) is the standard-state chemical potential of pure water at temperature \(T\). The interaction parameters are obtained from second virial coefficients \(B_i\) that can be obtained from osmotic-pressure or laser-light scattering data

\[
2B_{22} = 1000\frac{a_{22}}{(M_2)^2} \quad (6-4a)
\]

\[
2B_{33} = 1000\frac{a_{33}}{(M_3)^2} \quad (6-4b)
\]

\[
2B_{23} = 1000\frac{a_{23}}{M_2M_3} \quad (6-4c)
\]

where \(M_i\) is the molecular weight of polymer \(i\).

In a two-phase system at equilibrium, the chemical potential of each component in phase ‘\(\gamma\)’ is equal to that in phase ‘\(\gamma\)’

\[
\mu_i^\gamma = \mu_i^{\gamma'}; \quad \mu_2^\gamma = \mu_2^{\gamma'}; \quad \mu_3^\gamma = \mu_3^{\gamma'} \quad (6-5)
\]

Coupled with mass balances, eqs 6-1, 6-2, 6-3, and 6-5 give the binodal curve. For practical calculations, it is more convenient to obtain the spinodal curve given by Gibbs’ stability criterion

\[
\left(\frac{\partial \mu_2}{\partial m_2}\right)_{T,p,n_1,n_2} \left(\frac{\partial \mu_3}{\partial m_3}\right)_{T,p,n_1,n_2} - \left(\frac{\partial \mu_2}{\partial m_2}\right)_{T,p,n_1,n_2}^2 = 0 \quad (6-6)
\]

where \(n_i\) is the number of moles of component \(i\).

In freeze-drying, when the solution is cooled, ice is formed, and the aqueous-phase concentration of the excipients increases. At temperature \(T\), when equilibrium is attained, the chemical potential of ice is equal to that of liquid water

\[
\mu_{\text{ice}}(T) = \mu_i(T, m_2, m_3) \quad (6-7)
\]

Equation 6-7, along with those given above, can be used to calculate the concentration of the excipients when phase separation occurs at temperature \(T\). At this temperature, the change in chemical potential in going from pure liquid water to ice is given by

\[
\mu_{\text{ice}}(T) - \mu_i^0(T) = -\frac{\Delta H}{T_i} (T - T_i) - (C_{\text{ice}} - C_{\text{water}}) \cdot \left[ T \ln \left( \frac{T}{T_i} \right) - (T - T_i) \right] \quad (6-8)
\]

where \(T_i\) is the freezing temperature of pure water (273 K) and \(C_{\text{ice}}\) and \(C_{\text{water}}\) are the heat capacities of ice and pure supercooled liquid water, respectively; \(\mu_i^0\) is the standard state chemical potential of pure water at \(T_i\); and \(\Delta H\) is the enthalpy of fusion.

The equations above represent the thermodynamic limit. However, in practice, because the supercooled phase is very viscous, components cannot move easily from one phase to another. Phase separation is restrained by the slow rate of diffusion of water molecules from one phase to another, as shown by a curve of transition temperature as a function of temperature on the phase diagram.

Figure 6-1 shows the spinodal curve as a function of weight fraction of commonly used excipients, PEG3350 and dextran T500. Upon freeze-drying, Randolph et al.27 show that the concentrations of the excipients rise, but the weight ratio of the two excipients remains the same. Figure 6-2 shows a plot of temperature vs the weight ratios of PEG/dextran. The transition temperature at any weight ratio of PEG and dextran determines the minimum temperature for phase separation, although thermodynamics favors phase splitting at lower temperatures. Thus, the shaded region in Figure 6-2 represents the region where undesired phase separation is likely to occur. Figure 6-3 shows a phase diagram for a similar system containing the excipients PEG and albumin. Because the transition curve is above the spinodal curve for all considered weight ratios, Figure 6-3 implies that phase separation does not occur for this system.
Because we want to avoid phase separation, Figure 6-3 suggests that excipients PEG and albumin are preferable to PEG and dextran.

Randolph et al.\textsuperscript{27} obtained Scanning Electron Microscopy (SEM) images for the PEG/dextran system for different weight ratios. SEM images showed that for a weight ratio of 1 two condensed phases are present, suggesting that phase separation has occurred. However, at lower or higher weight ratios, only one phase is observed.

To determine the effect of phase separation on the stability of protein structure, Randolph et al.\textsuperscript{27} used infrared spectroscopy to obtain the loss of native structure of hemoglobin. Figure 6-4 shows how the width of a characteristic peak of hemoglobin at 1655 cm\textsuperscript{-1} changes with loss of structure. In this case, freeze-drying of hemoglobin was studied for different weight ratios of PEG and dextran as excipients. Prior to drying, the system was annealed at $-7^\circ C$ for 1 h in one case and 12 h in the other. The loss of native structure of the protein is shown in Figure 6-5 for different weight ratios. For those weight ratios where phase separation occurs, we expect a change in the structure of the protein; therefore, we anticipate significant differences in the width of the band obtained for 1 h compared to 12 h annealing. This difference is verified by results shown in Figure 6-5 that show, for weight ratios close to 1, there are significant differences in the width of the band. Also, SEM images show that this is the weight ratio where phase separation occurs. The experimental results indicate that phase separation is correlated with protein denaturation as predicted in Figure 6-2. This figure shows that at $-7^\circ C$ weight ratios close to 1 fall within the phase separation envelope that is detrimental to protein structure.

This example shows that thermodynamic analysis can provide a useful tool for predicting the region where undesired phase separation may occur during freeze-drying. Thermodynamic analysis can help find the optimal conditions where phase separation can be avoided. Thermodynamic analysis can reduce
the amount of experimental work that would be required in the absence of thermodynamic calculations.

**Conclusion**

The examples shown here illustrate how chemical thermodynamics provides a useful tool for biotechnology. In each example, thermodynamics provides a theoretical framework for contributing toward an answer to a biotechnical problem. Coupled with insights from molecular physics and with limited suitable experimental data, the framework can lead to useful results.

The generality and versatility of thermodynamics allow application to all materials, no matter how complex. Those who seek new applications of thermodynamics need not exclude biomacromolecules; they too are subject to the laws of thermodynamics.

Nearly 90 years ago, Gilbert Lewis and Merle Randall published a pioneering book (Thermodynamics and the Free Energy of Substances) that showed how the abstract chemical thermodynamics of Gibbs can be applied to real solids and fluids. Once again, we can be inspired by the famous introduction:

> *Let this book be dedicated to the chemists of the newer generation, who will not wish to reject all inferences from conjecture or surmise, but who will not care to speculate concerning that which may be surely known. The fascination of a growing science lies in the work of the pioneers at the very borderland of the unknown, but to reach this frontier one must pass over well-traveled roads; of these one of the safest and surest is the broad highway of thermodynamics.*

**Literature Cited**


