Ionic liquids and protein folding—old tricks for new solvents

Ryota Wakayama¹ · Susumu Uchiyama¹ · Damien Hall²

Received: 30 January 2019 / Accepted: 15 February 2019 / Published online: 19 March 2019

© International Union for Pure and Applied Biophysics (IUPAB) and Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

One important aspect of the green chemistry revolution has been the use of ionic liquids as the solvent in liquid-phase enzymatic catalysis. An essential requirement for protein enzyme function is the correct folding of the polypeptide chain into its functional “native” state. Quantitative assessment of protein structure may be carried out either empirically, or by using model-based characterization procedures, in which the parameters are defined in terms of a standard reference state. In this short note, we briefly outline the nature of the parameters associated with different empirical and model-based characterization procedures and point out factors which affect their interpretation when using a base solvent different from water. This review principally describes arguments developed by Wakayama et al., *Protein Solubility and Amorphous Aggregation: From Academic Research to Applications in Drug Discovery and Bioindustry*, 2019, edited by Y. Kuroda and F. Arisaka; CMC Publishing House. Sections of that work are translated from the original Japanese and republished here with the full permission of CMC Publishing Corporation.

Keywords Protein folding · Ionic liquids · Data reduction · Thermodynamic characterization

The advent of ionic liquids¹ has provided a diverse array of new solvent media in which to carry out commercially important enzymatic processes (Itoh 2017; Meyer et al.

2018). From the biotechnology perspective, ionic liquids hold great promise as supporting solvents in protein chemistry due to their:

- I. Low volatility and wide temperature range of the liquid state (with some able to remain liquid from ~ -80 to $300\text{ }^{\circ}\text{C}$) (Aparicio et al. 2010; Kazakov et al. 2018).
- II. Ability to selectively stabilize/destabilize a protein's native state (Patel et al. 2014; Kumar et al. 2017)
- III. Tunable solubility towards a wide range of organic, inorganic, and metallic compounds (Zhao et al. 2005).
- IV. Tunable miscibility/immiscibility with a wide range of organic solvents, therefore providing the ability to establish single-phase and multi-phase reaction systems (Marsh et al. 2004).
- V. Potential for catalytic enhancement via solvent-induced stabilization of the substrate transition state (Welton 1999; Weingärtner 2008; Naushad et al. 2012).

¹ The term ionic liquid is typically reserved for that class of molten salts that are liquid below the boiling temperature of water (Sheldon 2001; Wilkes 2002; Koel 2009). The first room temperature ionic liquid (RTIL) to be prepared was ethyl ammonium nitrate in 1914 (Walden 1914). From the 1940s–1970s, new classes of RTILs based on N-alkylpyridinium chloroaluminate chemistry (Hurley and Wier 1951; Chum et al. 1975; Wilkes et al. 1982) were developed with the initial aim of using these RTILs as both the solvent and substrate in electroplating applications. The modern era of RTILs, in which they were purposefully employed as “active solvents” for facilitating chemical reactions, was marked by the creation of ionic liquids from dialkyl-imidazolium cations and anions derived from (i) chloroaluminate (Wilkes et al. 1982), (ii) tetrafluoroborate (Wilkes and Zaworodtko, 1992), and (iii) hexafluorophosphate (Suarez et al. 1997). Since then, a host of new ionic liquids have been constructed based on various substituted cations (including ammonium, imidazolium, phosphonium, piperidinium, pyrrolidinium, and pyridinium) and a plethora of substituted anions (including acetates, sulfonylimides, borates, bromides, chlorides, iodides, dicyanamides, phosphates, phosphonates, sulfates, sulfonates, and chloroaluminates) (Sheldon 2001; Vekariya 2017). A curated database of all ionic liquids and their properties is available at the NIST website developed by Kazakov and coworkers (Kazakov et al. 2018). An excellent introduction to the physical properties of a smaller set of commonly used ionic liquids is provided by Domańska (2009).

✉ Damien Hall
damien.hall@protein.osaka-u.ac.jp; damienhall30@gmail.com

¹ Department of Biotechnology, Graduate School of Engineering, Osaka University, Suita, 3-1-Yamada-oka, Osaka 565-0871, Japan

² Institute for Protein Research, Osaka University, 3-1-Yamada-oka, Suita, Osaka 565-0871, Japan

Used in conjunction with rational or directed protein design principles (Fersht 1999; Stevenson et al. 2008), exploitation of the favorable properties of ionic liquids holds significant potential for the development of a new range of super-catalysts. In turn, such optimized super enzyme/solvent pairings, presage a new future of biotechnology-based industrial development (Park and Kazlauskas 2003; Itoh 2017). However, irrespective of the nature of the solvent, an enzyme's catalytic power is principally determined by the fraction of time that it exists in its active folded state—i.e., its thermodynamic stability (Fersht 1999; Fields 2001). As such, prior measurement of protein stability is an essential prerequisite to the process of “solvent tuning,” i.e., the blending of two or more ionic liquids to produce a solvent environment capable of supporting maximum catalytic efficiency of the enzyme (Weingartner et al. 2012; Lesch et al. 2015). Our modern understanding of protein folding and stability (Dill 1985; Dill and MacCallum 2012) is built on the back of a great many experimental studies of the equilibrium unfolding of proteins (Anfinsen 1973) in response to the application of either heat (von Hippel and Wong 1965; McPhie et al. 2006), pressure (Mozhaev et al. 1996; Pandharipande and Makhatazde 2016), or the addition of denaturing solutes and/or denaturing compatible solvents (Pace 1986; Schellman 2002).² In these studies, the solvent reference state has typically been a neutral pH aqueous buffer of low ionic strength and the characteristic parameters describing a particular protein's stability are given in terms of this reference state (Tanford 1968, 1970; Pace 1986; Schellman 2002). Obviously, when using ionic liquids, these parameters will have a different meaning and to this end, it was judged worthwhile to present a brief review of some of the general ways in which we measure and assess protein folding (Pace 1986; Eftink and Ionescu 1997; Fersht 1999) and point out how use of an ionic liquid as the supporting solvent may complicate the interpretation of the characteristic parameters associated with the various data reduction schemes (Pace 1986; Fersht 1999).

Basic description of protein folding

The simplest entry point into any discussion of protein folding involves its consideration as a cooperative two-state process (Anfinsen 1973; Pace 1986; Fersht 1999) in which a folded

structure, F , is in equilibrium with an enthalpically degenerate ensemble of unfolded structures, U , (Eq. 1a). In this model, the extent of reversible conversion between F and U states is governed by a unit-less equilibrium constant, K_{FU} , defined in terms of the concentration of the two species, C_F and C_U (Eq. 1b). A single progress variable reflecting the fractional extent of total protein in unfolded form, f_U , can be specified in terms of the measured concentration of both species and/or the constant, K_{FU} , defining the equilibrium (Eq. 1c).

$$F \xrightleftharpoons{K_{FU}} U \quad (1a)$$

$$K_{FU} = \frac{C_U}{C_F} \quad (1b)$$

$$f_U = \frac{C_U}{[C_U + C_F]} = \frac{K_{FU}}{[1 + K_{FU}]} \quad (1c)$$

The equilibrium constant, K_{FU} , is defined with respect to independently variable environmental conditions, which under biochemistry-type laboratory conditions typically include the temperature, T , pressure, P , and solution composition, i.e., the number, n_j , of all of the Q different types of solution components, $\sum_{j=1}^Q n_j$. When a change is made to the system (by individual modification of any of the independent variables, i.e., T , P , or n_j) the fraction of unfolded protein can be induced to change. In a conventional protein folding experiment, we can systematically examine this dependence by plotting the measured value of f_U against χ —where χ denotes the changing value of a single independent variable³ (Fig. 1). In what follows, we examine some of the different ways for quantitatively reducing such f_U vs. χ (or $\Delta\chi$) data (previously baseline corrected) to a set of characteristic parameters and discuss what these parameters mean when the solvent is not water.

Empirical scheme 1: midpoint analysis

Depending on experimental design, one may either measure the dependence of f_U on the absolute value of the independent variable, χ , or on its absolute difference, $\Delta\chi$, between some reference value, χ_1 and the progressed value of that variable, χ_2 (Fig. 1a). Given that the means for affecting the protein-folded state typically involves altering system temperature, pressure, or composition, by far, the vast majority of published unfolding experiments are presented as either f_U vs. T (or ΔT), f_U vs. P (or ΔP) or f_U vs. C_D (or ΔC_D) (where C_D refers to the

² Other less commonly employed experiments include mechanical pulling (Oberhauser et al. 2001) and application of viscous shear (Jaspe and Hagen 2006).

³ In some cases (such as molar concentration or volume fraction), χ represents a composite independent variable, i.e., two components changing in a linked-dependent fashion, with the second component typically being solvent which exists in excess.

Fig. 1 Midpoint characterization scheme. **a** Schematic showing the dependence of f_U , vs. χ (or $\Delta\chi$) (i.e., fraction of unfolded protein upon the variable responsible for inducing change, e.g., temperature, pressure or denaturing solute, or miscible solvent). The domain of interest for analysis lies between lower ($\chi_{\min}$ or $\Delta\chi_{\min}$) and upper bounds ($\chi_{\max}$ or $\Delta\chi_{\max}$) which respectively define the operating range between $f_U \approx 0$ and $f_U \approx 1$. A single parameter χ_{mp} (or $\Delta\chi_{mp}$) at which f_U is equal to 0.5 is used to characterize the data. **b** Deficiency of the midpoint characterization scheme: three different unfolding curves (red, blue, and black) are all characterized by the same value of χ_{mp} . With only one parameter, the midpoint characterization scheme cannot describe the steepness of the transition between folded and unfolded states. (Translated and reprinted with full permission from Wakayama et al. (2019))

concentration of denaturant) (Eftink and Ionescu 1997; Fersht 1999). The simplest procedure for characterizing unfolding curves involves their reduction to a single parameter describing

their transition midpoint (mp), i.e., the value of χ_{mp} at which $f_U = 0.5$ (Fig. 1a). Such an empirical analysis has the advantage that it is entirely general and can, in the ideal case, be used to rank solvents according to their capacity for protein stabilization (Fersht 1999; Pfeil 2012; Kumar et al. 2017). However, such a procedure provides no information on the transition path adopted and can therefore potentially miss important additional aspects of protein stability associated with the thermodynamic pathway (Fig. 1b). One such matter is enzyme loss due to aggregation (Roberts 2007; Weingartner et al. 2012; Kamal et al. 2016). Unfolded proteins exhibit a strong tendency to aggregate in a concentration-dependent manner to form either amorphous or structured aggregates (Chi et al. 2003; Mezzenga and Fischer 2013) (with amyloid fibers being one prominent example of a structured aggregate class having relevance to disease and biotechnology (Stefani and Dobson 2003; Hall 2003; Hall and Hirota 2009; Hall and Huang 2012; Hall 2012; Hall and Edskes 2004, 2009, 2012)). In any multi-parameter optimization procedure, conditions are modified to achieve the best system performance and therefore care must be taken that the adjustable factors such as temperature, pressure, solvent composition, and enzyme concentration do not adversely affect enzyme stability and activity (Rogers and Bommarius 2010; Lima-Ramos et al. 2014). As such, knowing how far below the unfolding midpoint to set reaction conditions that minimize loss of protein due to protein aggregation but do not miss out on the catalytic benefits of an altered value of χ is vital to efficient process design (Fields 2001; Bommarius 2015). The inability of the midpoint analysis to provide information on the steepness and breadth of the unfolding transition can be rectified using a slightly more complex two-parameter empirical characterization procedure known as the *m*-value approach (Greene and Pace 1974; Pace 1986), which will be discussed in the next section.

Empirical scheme 2: *m* value approach

In practice, an equilibrium constant reflecting the extent of protein unfolding can be determined for each set of conditions producing a different measured value of f_U (Fig. 2a). In a series of papers in the 1970s and 1980s Green and Pace (Greene and Pace 1974; Pace 1986; Pace and Shaw 2000) noted that the logarithm of the unfolding equilibrium constant frequently exhibited a linear dependence on $\Delta\chi$. In cases of non-linearity, an additional term (or terms), extending to a polynomial dependence, sufficed to provide a description of the system (Myers et al. 1995; Johnson and Fersht 1995; Eftink and Ionescu 1997; Nicholson and Scholtz 1996; O'Brien et al. 2009; this work). Although we will approach this point later in a more formal fashion, for now, we express this concept generally by noting that the equilibrium constant,

Fig. 2 Generalized m value approach. **a** Simulated dependence of f_U upon $\Delta\chi$ generated using the m value procedure (Eq. 5). This approach is able to account for the shape of the transition pathway based on just two parameters, the unfolding midpoint, χ_{mp} , and the m value (defined as $m = \left[\frac{d(\Delta g_{FU}^\circ)}{d\Delta\chi} \right]_{\Delta\chi=0}$). **b** Historically speaking, the m value approach was based on experimental observations of a linear dependence of $\log_e(K_{FU})$ upon the variable χ inducing the change. In this approach, values of K_{FU} were first evaluated from f_U (via Eq. 1c) at each point on the f_U vs. $\Delta\chi$ unfolding curve. Typically, noise within the data set only allowed points significantly removed from the extrema to be used in the estimation of K_{FU} , i.e., $0.1 < f_U < 0.9$. From a plot $\log_e(K_{FU})$ against $\Delta\chi$ a value of $\log_e(K_{FU})$ at some reference state χ_{min} could be determined by linear extrapolation. (Translated and reprinted with full permission from Wakayama et al. (2019))

for any reversible process, can be expressed in terms of the standard state molar free energy change between reactants and products, Δg_{FU}° , along with any additional free energy change, $d(\Delta g_{FU}^\circ(\Delta\chi))$, undergone by that system as a consequence of the change $\Delta\chi$ (Eq. 2).

$$(K_{FU})_{\Delta\chi} = \exp\left(\frac{-[\Delta g_{FU}^\circ + d(\Delta g_{FU}^\circ(\Delta\chi))]}{RT}\right) \quad (2a)$$

$$\log_e(K_{FU})_{\Delta\chi} = \log_e(K_{FU})_{\Delta\chi=0} - \left(\frac{1}{RT}\right)[d(\Delta g_{FU}^\circ(\Delta\chi))] \quad (2b)$$

The additional free energy term $d(\Delta g_{FU}^\circ)$ can be described as a function of $\Delta\chi$ using a Taylor expansion about the point ($\Delta\chi = 0$, Δg_{FU}°) (Eq. 3a). The often noted apparent linear dependence of a plot of $\log_e(K_{FU})_{\Delta\chi}$ upon $\Delta\chi$ (Auton and Bolen 2005; Myers et al. 1995; O'Brien et al. 2009) can be seen as a case of the adequacy of truncation of the Taylor expansion at the first term. This first derivative term was named as the m value by Pace and co-workers (Eq. 3b; Greene and Pace 1974).

$$d\Delta g_{FU}^\circ(\Delta\chi) = \left(\frac{d(\Delta g_{FU}^\circ)}{d\Delta\chi}\right)_{\Delta\chi=0} \cdot \Delta\chi + \frac{1}{2} \left(\frac{d^2(\Delta g_{FU}^\circ)}{d\Delta\chi^2}\right)_{\Delta\chi=0} \cdot (\Delta\chi)^2 + \dots \quad (3a)$$

$$m_\chi = \left(\frac{d(\Delta g_{FU}^\circ)}{d\Delta\chi}\right)_{\Delta\chi=0} \quad (3b)$$

For mathematical utility, the standard-state free energy is typically defined with regard to the midpoint of the unfolding transition (χ_{mp})—a point at which $\Delta\chi = 0$, $C_U = C_F$, $(K_{FU})_{\Delta\chi=0} = 1$, and $\Delta g_{FU}^\circ(\Delta\chi = 0) = 0$. Depending on the nature of the change (i.e., $\Delta\chi = \Delta P$, ΔT , or Δn_D (or as more commonly used ΔC_D)), the terms in the free energy derivative with respect to the change will differ. Here, we take a liberty in order to simplify the literature somewhat by employing the m value terminology coined originally for denaturant-induced unfolding (Greene and Pace 1974) for each of the three main methods for inducing protein unfolding, such that we have an m value for temperature, m_T (Eq. 4a), for pressure, m_P (Eq. 4b), and for denaturant, m_D (Eq. 4c).

$$m_T = \left(\frac{d(\Delta g_{FU}^\circ)}{dT}\right)_{\Delta T=0} \quad (4a)$$

$$m_P = \left(\frac{d(\Delta g_{FU}^\circ)}{dP}\right)_{\Delta P=0} \quad (4b)$$

$$m_D = \left(\frac{d(\Delta g_{FU}^\circ)}{dC_D}\right)_{\Delta C_D=0} \quad (4c)$$

Using this approach, protein unfolding curves may be minimally characterized using two parameters, the characteristic “ m value” and the transition midpoint χ_{mp} . Inserting the necessary terms for the description of K_{FU} at different values of χ into Eq. 1c yields Eq. 5 which can be used in non-linear regression analysis of f_U vs. χ data to provide estimates of the two relevant parameters χ_{mp} and m_χ .

$$f_U = \frac{\exp\left(\frac{-m_\chi \cdot (\chi - \chi_{mp})}{R.T}\right)}{1 + \exp\left(\frac{-m_\chi \cdot (\chi - \chi_{mp})}{R.T}\right)} \quad (5)$$

An important consequence of the m value procedure is that the difference in free energy between some additional reference state and the transition midpoint can be obtained via linear extrapolation allowing the intrinsic (unperturbed) stability of the protein to be estimated (Fig. 2b). For this reason, the m value procedure is sometimes called the linear extrapolation method (LEM) (Tanford 1970; Greene and Pace 1974; Pace 1986; Santoro and Bolen 1988; Auton and Bolen 2005). Despite the simplicity and obvious usefulness of the m value/LEM procedure, due to its empirical nature, it lacks mechanistic insight into the molecular causes of the unfolding reaction. In the following sections, we present some theory-based characterization procedures that rectify this problem.

Mechanistic approaches for characterizing protein folding based on classical thermodynamics

Noting the demonstrated success of the m value/LEM procedure described in the previous section, it will be instructive to examine the theoretical reasons as to (i) why such a linear extrapolation method is often observed to hold true in practice and (ii) what situations would lead to its invalidity. For the sake of making this section easily digestible, a general discussion of the thermodynamics of open systems containing components able to partake in reversible chemical equilibrium is provided as Appendix 1. Based on the material introduced therein, we develop our subsequent argument from a starting point provided by the three equations presented as Eq. 6. Briefly, Eq. 6a is concerned with the use of the chemical potential of the folded, μ_F , and unfolded, μ_U , protein species to define the condition for chemical equilibrium, Eq. 6b describes the method for incorporating change into the chemical potential upon a change in the reaction conditions $[T_1, P_1, \{n_j\}_1 \rightarrow T_2, P_2, \{n_j\}_2]$ and Eq. 6c describes the procedure for calculating the extent of change in chemical potential upon a change in conditions.⁴

$$(\mu_U)_{T_2, P_2, \{n_j\}_2} = (\mu_F)_{T_2, P_2, \{n_j\}_2} \quad (6a)$$

$$(\mu_i)_{T_2, P_2, \{n_j\}_2} = (\mu_i)_{T_1, P_1, \{n_j\}_1} + (d\mu_i)_{T_1, P_1, \{n_j\}_1 \rightarrow T_2, P_2, \{n_j\}_2} \quad (6b)$$

⁴ Note that Eq. 6c is written with the state (either 1 or 2) of the non-transitive independent variables unassigned to account for the fact that the changes may occur in any order. This is a required property for any state function. Furthermore, we have written the general case in which a change in composition of any component may potentially alter the chemical potential of the i th species without a specific interaction. Although not often recognized as such, this is the underlying formalism that provides the basis of the Tanford transfer free energy model, e.g., see Tanford (1970), Pace (1986), O'Brien et al. (2009), Horinek and Netz (2011). In this review, we will only consider the case of specific interactions thereby reducing Eq. 6c to $(d\mu_i)_{T_1, P_1, \{n_j\}_1 \rightarrow T_2, P_2, \{n_j\}_2} = (\partial\mu_i)_{T_1 \rightarrow T_2, P, \{n_j\}} + (\partial\mu_i)_{P_1 \rightarrow P_2, T, \{n_j\}} + (\partial\mu_i)_{(n_i)_1 \rightarrow (n_i)_2, T, P, \{n_{\neq i}\}}$.

$$\begin{aligned} (d\mu_i)_{T_1, P_1, \{n_j\}_1 \rightarrow T_2, P_2, \{n_j\}_2} &= (\partial\mu_i)_{T_1 \rightarrow T_2, P, \{n_j\}} \\ &+ (\partial\mu_i)_{P_1 \rightarrow P_2, T, \{n_j\}} + \sum_{j=1}^Q \left(\frac{\partial\mu_i}{\partial n_j} \right)_{(n_i)_1 \rightarrow (n_i)_2, T, P, \{n_{\neq i}\}} dn_j \end{aligned} \quad (6c)$$

On the back of these three equations, we derive analytical relations that describe the effect of change in temperature, pressure or concentration of added denaturant upon the unfolding equilibrium constant.

Temperature-induced unfolding

A thermally induced change in the concentration of the i th component, occurring at constant pressure and concentration of other solution components, implies that the equation describing the change in chemical potential will only include two of the $Q+2$ possible transitions from Eq. 6c, $(\partial\mu_i)_{T_1 \rightarrow T_2, P, \{n_j\}}$ and $(\partial\mu_i)_{(n_i)_1 \rightarrow (n_i)_2, T, P, \{n_{\neq i}\}}$ (Eq. 7a). The chemical potential of the folded and unfolded protein forms at a new set of conditions (calculated according to Eq. 6b) can be respectively described by equations shown as Eq. 7b and c. For clarity, we have used curly brackets to contain the terms constituting the species chemical potential at the initial condition set $(T_1, P_1, \{n_j\}_1)$ and the $\partial\mu$ terms just for the temperature $(T_1 \rightarrow T_2)$ and concentration $((n_i)_1 \rightarrow (n_i)_2)$ transitions (as described in Eq. 6c). Small letters are used to denote the partial molar change in a quantity such that s_i and h_i respectively represent the partial molar change in system entropy and enthalpy for component i (i.e., $s_i = \left(\frac{\partial S}{\partial n_i} \right)_{T, P, \{n_{\neq i}\}}$ and $h_i = \left(\frac{\partial H}{\partial n_i} \right)_{T, P, \{n_{\neq i}\}}$)

$$(d\mu_i)_{T_1 \rightarrow T_2, (n_i)_1 \rightarrow (n_i)_2, P, \{n_{\neq i}\}} = (\partial\mu_i)_{T_1 \rightarrow T_2, P, \{n_j\}} + (\partial\mu_i)_{(n_i)_1 \rightarrow (n_i)_2, T, P, \{n_{\neq i}\}} \quad (7a)$$

$$\begin{aligned} (\mu_F)_{T_2, P, \{n_{\neq F}\}_2} &= \left\{ (\mu_F^\circ)_{T_1, P, \{n_{\neq F}\}_1} + RT_1 \log_e \left(\frac{(C_F)_1}{C_F^\circ} \right) \right\} \\ &+ \left\{ -\int_{T_1}^{T_2} s_F dT \right\} + \left\{ RT_1 \log_e \left(\frac{(C_F)_2}{(C_F)_1} \right) \right\} \end{aligned} \quad (7b)$$

$$\begin{aligned} (\mu_U)_{T_2, P, \{n_{\neq U}\}_2} &= \left\{ (\mu_U^\circ)_{T_1, P, \{n_{\neq U}\}_1} + RT_1 \log_e \left(\frac{(C_U)_1}{C_U^\circ} \right) \right\} \\ &+ \left\{ -\int_{T_1}^{T_2} s_U dT \right\} + \left\{ RT_1 \log_e \left(\frac{(C_U)_2}{(C_U)_1} \right) \right\} \end{aligned} \quad (7c)$$

Inserting these relations into the condition for a two-state equilibrium between folded and unfolded states (Eq. 6a) yields Eq. 8.

$$-R \log_e \left(\frac{C_{U\{2\}} \cdot C_F^\circ}{C_{F\{2\}} \cdot C_U^\circ} \right) = \left(\frac{1}{T_1} \right) \cdot \left(\Delta \mu_{FU}^\circ - \int_{T_1}^{T_2} \Delta s_{FU} \cdot dT \right) \quad (8a)$$

$$\Delta \mu_{FU}^\circ = (\mu_U^\circ)_{T_1, P_1, \{n_{j \neq U}\}_1} - (\mu_F^\circ)_{T_1, P_1, \{n_{j \neq F}\}_1} \quad (8b)$$

$$\Delta s_{FU} = (s_U - s_F) \quad (8c)$$

In order to calculate the change in entropy of the system at a fixed temperature with regard to change in the number of moles of either the folded or unfolded species, we note that the standard state chemical potential of component i appearing in Eq. A2c, μ_i° , can be expressed in terms of its component standard state enthalpy, h_i° , and entropy, s_i° , with this substitution allowing construction of Eq. 9.

$$\Delta \mu_{FU}^\circ = (h_U^\circ - h_F^\circ) - T \cdot (s_U^\circ - s_F^\circ) \quad (9)$$

If a set of conditions $[T_1, P_1, \{n_j\}_1]$ is selected such that $\Delta \mu_{FU}^\circ = 0$, Eq. 9 can be rearranged to Eq. 10.

$$\Delta s_{FU}^\circ = \frac{\Delta h_{FU}^\circ}{T_1} \quad (10)$$

We may calculate the value of Δs_{FU} at a different temperature by introducing the concept of heat capacity (Shiao et al. 1971; Privalov and Khechinashvili 1974; Sturtevant 1977; Becktel and Schellman 1987; Van Holde et al. 2006a). For small changes in temperature, the molar system enthalpy at a temperature T_2 can be determined from its value at a reference temperature, T_1 , using a linear approximation (Eq. 11a). Equation 11a can be further differentiated with respect to changes in composition allowing the partial molar enthalpy of unfolding, Δh_{FU} , at a second temperature to be determined using Eq. 11b.

$$H_{\text{system}}(T_2) = H_{\text{system}}^\circ(T_1) + \Psi \cdot dT \quad (11a)$$

$$(\Delta h_{FU})_{T_2, P_1, \{n_j\}_2} = (\Delta h_{FU})_{T_1, P_1, \{n_j\}_2} + \int_{T_1}^{T_2} (\Delta \sigma_{FU})_{P_1, \{n_j\}_2} \cdot dT \quad (11b)$$

Here, Ψ describes the heat capacity at constant pressure, i.e., the differential change in system enthalpy with change in temperature, i.e., $\psi = \left(\frac{\partial H^\circ}{\partial T} \right)_{P, \{n_j\}}$ and $\Delta \sigma_{FU}$ describes the difference in the partial derivatives of Ψ with regard to U and F species such that, $(\Delta \sigma_{FU})_{P, \{n_j\}} = \left(\frac{\partial \Psi}{\partial n_U} \right)_{P, \{n_{j \neq U}\}} - \left(\frac{\partial \Psi}{\partial n_F} \right)_{P, \{n_{j \neq F}\}}$. Using the formalism shown in Eqs. 10 and

11, the partial molar change in system entropy for the unfolding reaction at a different temperature, $\Delta s_{FU}(T)$, can be estimated by expanding around the exact result defined at T_1 in Eq. 12.⁵

$$\Delta s_{FU}(T_1) = \left(\frac{\Delta h_{FU}^\circ + \Delta \sigma_{FU} \cdot \Delta T}{T_1} \right) \quad (12)$$

Inserting Eq. 12 into Eq. 8a followed by integration and then multiplication through by T_2 yields Eq. 13a. Subsequent rearrangement and taking the anti-logarithm produces Eq. 13b.

$$-RT_2 \log_e \left(\frac{C_{U\{2\}} \cdot C_F^\circ}{C_{F\{2\}} \cdot C_U^\circ} \right) = \left(\frac{T_2}{T_1} \right) \cdot (\Delta \mu_{FU}^\circ)_{T_1, P_1, \{n_j\}_1} + \left[\Delta h_{FU}^\circ \cdot \left(1 - \frac{T_2}{T_1} \right) - \Delta \sigma_{FU} \cdot \left(T_2 \cdot \log_e \left(\frac{T_2}{T_1} \right) - (T_2 - T_1) \right) \right] \quad (13a)$$

$$(K_{FU})_{T_2, P_1, \{n_j\}_2} = [(K_{FU})_{T_1, P_1, \{n_j\}_1}] \cdot \exp \left\{ \frac{- \left[\Delta h_{FU}^\circ \cdot \left(1 - \frac{T_2}{T_1} \right) + \Delta \sigma_{FU} \cdot (T_2 - T_1) - \Delta \sigma_{FU} \cdot T_2 \cdot \log_e \left(\frac{T_2}{T_1} \right) \right]}{RT_2} \right\} \quad (13b)$$

One interesting aspect of Eq. 13 is associated with the definition of T_1 . This temperature must be set as the midpoint of the unfolding transition, i.e., $T_1 = T_{mp}$, to satisfy the standard state requirement in Eq. 10a and thus, by definition $(\Delta \mu_{FU}^\circ)_{T_1, P_1, \{n_j\}_1} = 0$; $C_F = C_U$, and $(K_{FU})_{T_1, P_1, \{n_j\}_1} = 1$. Inserting this description of K_{FU} into Eq. 1c provides a three parameter (T_{mp} , Δh_{FU}° , and $\Delta \sigma_{FU}$) expression (Eq. 14) useful for non-linear regression analysis of f_U vs. T data (Fig. 3a).

$$f_U = \frac{\exp \left\{ \frac{- \left[\Delta h_{FU}^\circ \cdot \left(1 - \frac{T_2}{T_{mp}} \right) + \Delta \sigma_{FU} \cdot (T_2 - T_{mp}) - \Delta \sigma_{FU} \cdot T_2 \cdot \log_e \left(\frac{T_2}{T_{mp}} \right) \right]}{RT_2} \right\}}{1 + \exp \left\{ \frac{- \left[\Delta h_{FU}^\circ \cdot \left(1 - \frac{T_2}{T_{mp}} \right) + \Delta \sigma_{FU} \cdot (T_2 - T_{mp}) - \Delta \sigma_{FU} \cdot T_2 \cdot \log_e \left(\frac{T_2}{T_{mp}} \right) \right]}{RT_2} \right\}} \quad (14)$$

In terms of explaining the dependence (linear or otherwise) of $\log_e(K_{FU})$ upon ΔT , we note that Eq. 13b

⁵ Noting that entropy is a function of both heat and temperature.

can be rearranged to Eq. 15a—a further simplification of which can be made (Eq. 15b) by noting⁶ that for

$$\log_e \left[(K_{FU})_{T_2, P_1, \{n_j\}_2} \right] = \left(\frac{1}{RT_2} \right) \left\{ - \left[\Delta h_{FU}^\circ \left(\frac{T_2 - T_{mp}}{T_{mp}} \right) + \Delta \sigma_{FU} \cdot (T_2 - T_{mp}) - \Delta \sigma_{FU} \cdot T_2 \cdot \log_e \left(\frac{T_2}{T_{mp}} \right) \right] \right\} \quad (15a)$$

$$\log_e \left[(K_{FU})_{T_2, P_1, \{n_j\}_2} \right] \approx \frac{-\Delta h_{FU}^\circ}{R \cdot T_{mp}^2} \cdot \Delta T \quad (15b)$$

The above two equations (Eq. 15a and b) are sufficient to understand why and under what regimes an m value-type description can represent a good approximation for thermal unfolding, i.e., Eq. 15b will hold for small changes in temperature but curvature should become apparent for larger temperature deviations or alternatively, when the system heat capacity is so sensitive to temperature change that it may not be described using a linear approximation. Importantly, Eq. 15 allows for the mechanistic interpretation of a $\log_e$ $\left[(K_{FU})_{T_2, P_1, \{n_j\}_2} \right]$ vs. T -type plot. Around the midpoint, the slope⁷ will be equal to $-\Delta h_{FU}/(RT_{mp}^2)$ (Fig. 3b) and the m_T value (as defined by Eq. 4a) will be approximately equal to $-\Delta h_{FU}/T_{mp}$. At positions farther away from the midpoint, the general slope of the $\log_e \left[(K_{FU})_{T_2, P_1, \{n_j\}_2} \right]$ vs. ΔT plot will be a complex non-constant function given by $\left(\frac{-1}{RT_2} \right)$

$\left[\Delta h_{FU}^\circ/T_{mp} + \Delta \sigma_{FU} - \frac{\Delta \sigma_{FU} \cdot T_2 \cdot \log_e \left(\frac{T_2}{T_{mp}} \right)}{(T_2 - T_{mp})} \right]$. Interestingly, the non-constant slope term contained within Eq. 15a provides the necessary thermodynamic formulation to quantitatively “understand” the phenomenon of low-temperature-induced protein unfolding—sometimes termed cold denaturation (Privalov 2007; Fig. 3). With a greater possible range of temperatures for the liquid state, such considerations will be important when interpreting the behavior of protein folding in ionic liquid media (Constantinescu et al. 2007; Constantinescu et al. 2010; Satish et al. 2016; Brogan and Hallett 2016).

Pressure-induced unfolding

The volume, V , of a multicomponent system can be expressed in terms of the sum of the number of moles of each component multiplied by their partial molar volume, $(v_i)_{T, P, \{n_j\}}$, where

⁶ Thereby effectively canceling out the heat capacity terms.

⁷ Note that we have derived the derivative form of the Gibbs-Helmholtz equation.

small deviations from the midpoint $\log_e(T_2/T_{mp}) \approx (T_2 - T_{mp})/T_{mp}$ (Fig. 3b).

$(v_i)_{T, P, \{n_j\}} = \left(\frac{\partial V}{\partial n_i} \right)_{T, P, \{n_j\}}$ (Eq. 16). Assuming that the partial molar volume does not depend upon type and concentration of other components (Van Holde et al. 2006b) allows Eq. 16a to be approximated by Eq. 16b.

$$(V)_{T, P, \{n_j\}} = \sum_{i=1}^Q \int_0^{n_i} \left(\frac{\partial V}{\partial n_i} \right)_{T, P, \{n_j\}} \cdot dn_i \quad (16a)$$

$$(V)_{T, P, \{n_j\}} \approx \sum_{i=1}^Q n_i \cdot (v_i)_{T, P} \quad (16b)$$

At constant temperature but altered pressure, the chemical potentials for folded and unfolded protein forms at a new set of conditions $(T_1, P_2, \{n_j\}_2)$ corresponding to changes in P and $\{n_j\}$ are respectively given by Eq. 6b and c (resulting in Eq. 17a and b). Again, curly brackets are used to indicate terms arising from change in chemical potential, $d\mu$, (via Eq. 6c).

$$\begin{aligned} (\mu_F)_{T_1, P_2, \{n_j\}_2} &= \left\{ (\mu_F^\circ)_{T_1, P_1, \{n_{\neq F}\}_1} + RT_1 \ln \left(\frac{(C_F)_1}{C_F^\circ} \right) \right\} \\ &+ \left\{ \int_{P_1}^{P_2} \left(\frac{\partial V}{\partial n_F} \right)_{T_1, P, \{n_{\neq F}\}_2} \cdot dP \right\} + \left\{ RT_1 \ln \left(\frac{(C_F)_2}{(C_F)_1} \right) \right\} \end{aligned} \quad (17a)$$

$$\begin{aligned} (\mu_U)_{T_1, P_2, \{n_j\}_2} &= \left\{ (\mu_U^\circ)_{T_1, P_1, \{n_{\neq U}\}_1} + RT_1 \ln \left(\frac{(C_U)_1}{C_U^\circ} \right) \right\} \\ &+ \left\{ \int_{P_1}^{P_2} \left(\frac{\partial V}{\partial n_U} \right)_{T_1, P, \{n_{\neq U}\}_2} \cdot dP \right\} + \left\{ RT_1 \ln \left(\frac{(C_U)_2}{(C_U)_1} \right) \right\} \end{aligned} \quad (17b)$$

With the additional assumption that the partial molar volume for both folded, v_F , and unfolded forms, v_U , of protein are independent of pressure, the above expressions can be further simplified and inserted into the equilibrium relation shown by Eq. 6a to produce Eq. 18a which, with subsequent rearrangement, produces Eq. 18b.

$$\begin{aligned} (\mu_F^\circ)_{T_1, P_1, \{n_j\}_1} + \int_{P_1}^{P_2} v_F \cdot dP + RT_1 \ln \left(\frac{(C_F)_2}{C_F^\circ} \right) \\ \approx (\mu_U^\circ)_{T_1, P_1, \{n_j\}_1} + \int_{P_1}^{P_2} v_U \cdot dP + RT_1 \ln \left(\frac{(C_U)_2}{C_U^\circ} \right) \end{aligned} \quad (18a)$$

Fig. 3 Temperature-induced unfolding. **a** Simulated dependence of the fraction of unfolded protein upon temperature (f_U upon T) generated using the Gibbs-Helmholtz equation (Eq. 12). As formulated, the Gibbs-Helmholtz expression is based on three independent parameters, the heat-induced unfolding mid-point, T_{mp} , the change in molar enthalpy upon unfolding, Δh_{FU} , and the change in molar heat capacity upon unfolding, $\Delta \sigma_{FU}$. Note that the appearance of the cold induced unfolded region at low temperatures. **b** A plot of $\log_e(K_{FU})$ vs. T is typically not linear due to the additional molar heat capacity term ($\Delta \sigma_{FU}$) which accounts for changes in the enthalpy of unfolding at different temperatures. As such, the m_T value predicted in Eq. 4a will only describe the slope at the heat-induced transition midpoint, T_{mp} . [Parameter values used in the simulation were $\Delta h_{FU} = 6 \times 10^5 \text{ J mol}^{-1}$, $\Delta \sigma_{FU} = 2.4 \times 10^4 \text{ J K}^{-1} \text{ mol}^{-1}$ and $T_{mp} = 57.5^\circ \text{C}$]. (Translated and reprinted with full permission from Wakayama et al. (2019))

$$-RT_1 \ln \left(\frac{(C_U)_2 \cdot C_F^*}{(C_F)_2 \cdot C_U^*} \right) = (\mu_U^0)_{T_1, P_1, \{n_{j \neq U}\}_1} - (\mu_F^0)_{T_1, P_1, \{n_{j \neq F}\}_1} \quad (18b)$$

$$+ \int_{P_1}^{P_2} (v_U - v_F) \cdot dP$$

Integration of Eq. 18b yields a result for the equilibrium constant (Eq. 19a) which may be further simplified (Eq. 19b) (which is the analytical form most often used in the treatment of data; Mozhaev et al. 1996; Fersht 1999; Sasahara et al. 2001; Pandharipande and Makhatadze 2016).

$$(K_{FU})_{T_1, P_2, \{n_j\}_2} = \left(\frac{(C_U)_2 \cdot C_F^*}{(C_F)_2 \cdot C_U^*} \right)$$

$$= \exp \left(\frac{-[(\mu_U^0)_{T_1, P_1, \{n_{j \neq U}\}_1} - (\mu_F^0)_{T_1, P_1, \{n_{j \neq F}\}_1}] - \Delta v_{FU} \cdot [P_2 - P_1]}{RT_1} \right) \quad (19a)$$

$$(K_{FU})_{T_1, P_2, \{n_j\}_2} = [(K_{FU})_{T_1, P_1, \{n_j\}_1}] \cdot \exp \left[\frac{-\Delta v_{FU} \cdot [P_2 - P_1]}{RT_1} \right] \quad (19b)$$

If the initial state pressure, P_1 , is chosen as the midpoint of the unfolding transition, P_{mp} , then by definition, the equilibrium constant at this pressure is equal to 1, i.e., $(K_{FU})_{T_1, P_1, \{n_j\}_1} = 1$ and $C_F = C_U$ (Fig. 4a). Inserting this expression for K_{FU} into Eq. 1c provides an equation consisting of two parameters (P_{mp} and Δv_{FU}) that can be used directly in non-linear regression analysis of f_U vs. P unfolding data (Eq. 20a; Fig. 4a). Interestingly, as a non-colligative parameter, the change in partial molar volume, Δv_{FU} , will be dependent upon the molecular mass of the protein M_P (units kg/mol) and the change in partial specific volume upon unfolding, Δv_{FU} , with this latter quantity defined in terms of volume per weight concentration (units m^3/kg) such that $\Delta v_{FU} = M_P \Delta v_{FU}$. This feature introduces an underappreciated protein size dependence to the pressure-induced unfolding experiment (Fig. 4b).

$$f_U = \frac{\exp \left[\frac{-\Delta v_{FU} \cdot [P_2 - P_{mp}]}{RT_1} \right]}{1 + \exp \left[\frac{-\Delta v_{FU} \cdot [P_2 - P_{mp}]}{RT_1} \right]} \quad (20)$$

Taking the logarithm of Eq. 19b yields Eq. 21.⁸

$$\log_e [(K_{FU})_{T_1, P_2, \{n_j\}_2}] = \frac{-\Delta v_{FU}}{RT_1} \cdot [P_2 - P_1] \quad (21)$$

Equation 21 is sufficient to achieve an understanding as to why the linear dependence prescribed by the m value/LEM approach represents a good approximation for pressure-

⁸ With $(K_{FU})_{T_1, P_2, \{n_j\}_2}$ defined at the midpoint $\log_e((K_{FU})_{T_1, P_2, \{n_j\}_1}) = 0$.

induced unfolding (Fig. 4c). Equation 21 also allows for mechanistic interpretation of the meaning of the slope around the midpoint of a plot of $\log_e((K_{FU})_{T_1, P_2, \{n_j\}_2})$ vs. P , i.e., the slope will be equal to $\Delta v_{FU}/(RT_1)$ and the m_P value (as defined by Eq. 4b) will be equal to Δv_{FU} . Furthermore, from

Fig. 4 Pressure-induced unfolding. **a** Simulated dependence of the fraction of unfolded protein upon pressure (f_U upon P) generated using Eq. 18 which is based upon two parameters, the molar volume change upon unfolding, Δv_{FU} , and the pressure midpoint of the unfolding transition, P_{mp} . The pressure domain starts at standard atmospheric pressure, 101.3 kPa, and increases to 5 MPa. The red, blue, and black lines correspond to the pressure-induced unfolding transitions of three different proteins all jointly characterized by a mass $M_p = 50,000$ g/mol but respectively exhibiting three different values for Δv_{FU} of $-0.15, 0.1$ and 0.05 ml/g. **b** Simulated f_U upon P curves where red, blue, and black lines respectively describe the unfolding transitions of proteins possessing three different molecular masses $M_p = 50,000, 35,000$, and $20,000$ g/mol but all exhibiting a common value for Δv_{FU} of -0.15 ml/g. **c** A plot of $\log_e(K_{FU})$ vs. P will be linear if the assumptions made in the derivation of Eq. 18 hold. (Additional simulation parameters $P_{mp} = 2.53$ MPa, $T = 20$ °C]. (Translated and reprinted with full permission from Wakayama et al. (2019))

the explicit listing of the two approximations in its derivation (Eq. 16b— v_i is independent of composition and Eq. 18— v_i is independent of the applied pressure), we can understand which assumptions are being violated when the m value/LEM is seen to not to be applicable (Eftink and Ionescu 1997; Fersht 1998; Pandharipande and Makhatazde 2016). Although not treated here, potential solvent compressibility issues associated with the diverse range of ionic liquids and their blends, represent an open question in this field (Seddon et al. 2000; Jacquemin et al. 2007).

Denaturant-induced unfolding

At constant temperature and pressure, a protein can be induced to unfold by replacement of the solvent components with chemicals which differentially stabilize the unfolded vs. folded state(s) (Anfinsen 1973). Such denaturant-induced unfolding may involve either specific or non-specific interaction between the added denaturants and the protein (whereby the assignment of the designation specific or non-specific is made on the basis of an affinity for the denaturant at a particular chemical site on the protein (Pace 1986; Schellman 2002; Hall et al. 2018)). This basic difference in consideration of specific vs. non-specific binding of denaturant is the underlying cause of the conceptual schism that has led to different theoretical formulations of the denaturant induced unfolding process (e.g., Schellman 1956; Aune and Tanford 1969 vs. Schellman 2002; Smiatek 2017 vs. Tanford 1968; O'Brien et al. 2009). Although each of these different approaches are logically consistent (and, importantly, mutually compatible [see Pace 1986 for an easy introduction]) due to its ease of construction, in what follows, we adopt the quantitative formulation for the specific binding format (Schellman 1958; Aune and Tanford 1969; Hall et al. 2018). We will consider three different states of the system that exist at an identical fixed temperature, T_1 , and pressure, P_1 , but differ with respect to the concentration of denaturant and reversible components. These three states defined as follows—state 0 [T_1 ,

$P_1, \{n_j\}_0, (C_D = 0)$, state 1 $[T_1, P_1, \{n_j\}_1, (C_D)_1]$ and state 2 $[T_1, P_1, \{n_j\}_2, (C_D)_2]$.

In the specific binding model, the folded, F , and unfolded, U , forms of the protein are considered to respectively possess N_{FD} - and N_{UD} -specific independent and equivalent sites to which denaturant can bind, with site binding constants of k_{FD} and k_{UD} (Eq. 22).

For a certain loading of denaturant molecules attached to each protein state, an unfolding equilibrium constant can be defined based on an assumed stepwise equilibrium (Eq. 23).⁹

$$[.. \rightleftharpoons FD_2 \rightleftharpoons FD_1 \rightleftharpoons F] \rightleftharpoons [U \rightleftharpoons UD_1 \rightleftharpoons UD_2 \rightleftharpoons ..] \quad (23a)$$

$$(K_{FU})_{T_1, P_1, \{n_j\}_2, (C_D)_2} = \frac{\sum_{j=0}^{N_{UD}} C_{UDj}}{\sum_{i=0}^{N_{FD}} C_{FDi}} \quad (23b)$$

Expanding the stepwise equilibrium shown in Eq. 23 through inclusion of the concepts of statistical degeneracy of binding sites ((Van Holde et al. 2006b) described further in Appendix 1) allows for the enumeration of the binding polynomials for all unfolded (numerator) and folded (denominator) forms (Eq. 24).

$$(K_{FU})_{T_1, P_1, \{n_j\}_2, (C_D)_2} = \frac{C_U \left(\sum_{i=0}^{N_{UD}} A_i \cdot (k_{UD} C_D)^i \right)}{C_F \left(\sum_{j=0}^{N_{FD}} B_j \cdot (k_{FD} C_D)^j \right)} \quad (24a)$$

$$(K_{FU})_{T_1, P_1, \{n_j\}_2, (C_D)_2} = (K_{FU})_{T_1, P_1, \{n_j\}_0, (C_D=0)} \cdot \frac{\sum_{i=0}^{N_{UD}} A_i \cdot (k_{UD} C_D)^i}{\sum_{j=0}^{N_{FD}} B_j \cdot (k_{FD} C_D)^j} \quad (24b)$$

$$A_i = \left[\frac{N_{UD}!}{(N_{UD}-i)! \cdot i!} \right] \quad (24c)$$

$$B_j = \left[\frac{N_{FD}!}{(N_{FD}-j)! \cdot j!} \right] \quad (24d)$$

With some mathematical familiarity, one realizes that the bracketed summations shown in Eq. 24 represent a binomial series and as such, we may rewrite Eq. 24b as Eq. 25a. With the further assumption that the site-binding constants for denaturant binding to folded and unfolded states of the protein are equal, i.e., $k_{FD} = k_{UD}$

$= k_{PD}$, we may rewrite Eq. 23b in the form shown in Eq. 25b (Aune and Tanford 1969; Hall et al. 2018).

$$(K_{FU})_{T_1, P_1, \{n_j\}_2, (C_D)_2} = (K_{FU})_{T_1, P_1, \{n_j\}_0, (C_D=0)} \cdot \frac{\sum_{i=0}^{N_{UD}} A_i \cdot (k_{UD} C_D)^i}{\sum_{j=0}^{N_{FD}} B_j \cdot (k_{FD} C_D)^j} \quad (25a)$$

$$(K_{FU})_{T_1, P_1, \{n_j\}_2, (C_D)_2} = (K_{FU})_{T_1, P_1, \{n_j\}_0, (C_D=0)} (1 + k_{PD} C_D)^{\Delta N_D} \quad (25b)$$

This expression for K_{FU} can be inserted into Eq. 1c to provide an analytical expression based on three parameters $[(K_{FU})_{T_1, P_1, \{n_j\}_0, (C_D=0)}, k_{PD}, \Delta N_D]$ capable of being applied to f_U vs. C_D -type unfolding data through non-linear regression (Eq. 26; Fig. 5a).

$$f_U = \frac{(K_{FU})_{T_1, P_1, \{n_j\}_0, (C_D=0)} \cdot (1 + k_{PD} C_D)^{\Delta N_D}}{1 + (K_{FU})_{T_1, P_1, \{n_j\}_0, (C_D=0)} \cdot (1 + k_{PD} C_D)^{\Delta N_D}} \quad (26)$$

Searching for the causation of linearity in the standard m value-type plot of $\log_e(K_{FU})$ vs. C_D , we note that taking natural logarithms of both sides of Eq. 25b produces Eq. 27 (Fig. 5b).

$$\log_e(K_{FU})_{T_1, P_1, \{n_j\}_2, (C_D)_2} = \log_e(K_{FU})_{T_1, P_1, \{n_j\}_0, (C_D=0)} + \Delta N_D \cdot \log_e(1 + k_{PD} C_D) \quad (27)$$

For small values of x , the series expansion for $\log_e(1 + x) \approx \sum_{i=0}^{\infty} \frac{(-x)^i}{i!}$ allows one to write the approximate form at any concentration $(C_D)_2$ (Eq. 28a) and at a particular reference concentration $(C_D)_1$ (Eq. 28b).

$$\log_e(K_{FU})_{T_1, P_1, \{n_j\}_2, (C_D)_2} = \log_e(K_{FU})_{T_1, P_1, \{n_j\}_0, (C_D=0)} \quad (28a)$$

$$+ \Delta N_D \cdot \left[(k_{PD}(C_D)_2) - \frac{1}{2} \cdot (k_{PD}(C_D)_2)^2 + \dots \right]$$

$$\log_e(K_{FU})_{T_1, P_1, \{n_j\}_1, (C_D)_1} = \log_e(K_{FU})_{T_1, P_1, \{n_j\}_0, (C_D=0)}$$

$$+ \Delta N_D \cdot \left[(k_{PD}(C_D)_1) - \frac{1}{2} \cdot (k_{PD}(C_D)_1)^2 + \dots \right] \quad (28b)$$

If the reference concentration $(C_D)_1$ is set at the midpoint $(C_D)_{mp}$ then $(C_U)_{mp} = (C_F)_{mp}$ and $(K_{FU})_{T_1, P_1, \{n_j\}_1, (C_D)_1} = 1$. Subtracting Eq. 28b from 28a yields Eq. 29.

$$\begin{aligned} \log_e(K_{FU})_{T_1, P_1, \{n_j\}_2, (C_D)_2} &= \Delta N_D \cdot \left[\left(k_{PD} \{ (C_D)_2 - (C_D)_{mp} \} \right) - \frac{1}{2} \cdot \left\{ (k_{PD}(C_D)_2)^2 - (k_{PD}(C_D)_{mp})^2 \right\} + \dots \right] \end{aligned} \quad (29)$$

⁹ Although lacking mathematical correctness, to facilitate our exposition in this section, we will likewise specify the unfolding equilibrium constant with respect to the concentration of denaturant and number of moles of other solution components (including those involved in the equilibrium) using the following formalism, $(K_{FU})_{T_1, P_2, \{n_j\}_2, (C_D)_2}$.

Fig. 5 Denaturant-induced unfolding. **a** Simulated dependence of the fraction of unfolded protein upon denaturant concentration (f_U upon C_D) generated using Eq. 2 which is based upon three parameters, the equilibrium unfolding constant at zero denaturant concentration, $(K_{FU})_{T_1, P_1, \{n_j\}_0, (C_D=0)}$, the denaturant site-binding concentration, k_{PD} , and the difference in total number of denaturant sites available between the unfolded and folded forms of the protein, ΔN_D . Dotted lines indicate the transition midpoint. **b** A plot of $\log_e(K_{FU})$ vs. C_D will be linear around $C_D = 0$ but not necessarily linear in the transition region. Black-dashed lines are based on an extrapolation of values in the $f_U \in [0.1, 0.9]$ region with the extrapolated value at $C_D = 0$ different from that predicted by Eq. 26 (red circles). (Translated and reprinted with full permission from Wakayama et al. (2019))

Upon initial inspection, it would seem that Eq. 29 suggests an approximate value for the slope around the midpoint of a

plot of $\log_e \left[(K_{FU})_{T_1, P_1, \{n_j\}_2, (C_D)_2} \right]$ vs. ΔC_D , of $\Delta N_D \cdot k_{PD}$. For small perturbations, we may be tempted to define a corresponding m_D value at the midpoint which, according to Eq. 4c, should be given by $\Delta N_D \cdot R \cdot T_{mp} \cdot k_{PD}$. However, for a more complete interpretation, it is important to understand two points, (i) that to achieve the approximate forms (shown in Eq. 28), a Taylor expansion was performed around the point $C_D = 0$; therefore, the subtracted expression (shown as Eq. 29) involves the difference of two approximations which are themselves increasingly suspect beyond the point $C_D = 0$, and (ii) the expansion does not progress in ΔC_D but in C_D , i.e., note that the higher order z th terms involve $((C_D)_2)^z - ((C_D)_{mp})^z$ and not $((C_D)_2 - (C_D)_{mp})^z$. As such, a plot of $\log_e$

$\left[(K_{FU})_{T_1, P_1, \{n_j\}_2, (C_D)_2} \right]$ vs. C_D should be linear about $C_D = 0$ but not necessarily at the midpoint where this assumption has previously been used as the basis of extrapolation procedures (Pace 1986; Auton and Bolen 2005; Street et al. 2006). That linearity is often observed is due to the extremely small value of k_{PD} , $[1 \times 10^{-3} - 1 \times 10^{-1} \text{ M}^{-1}]$ measured for the binding of denaturants such as urea and guanidine hydrochloride in aqueous systems (Aune and Tanford 1968; Schellman 2002; Hall et al. 2018). It remains to be seen what magnitude k_{PD} will take when the denaturant effect arises from preferential solvation of protein by the ionic liquid or by denaturing solutes dissolved within them as opposed to the case for aqueous solvent (Canchi and García 2013; Smiatek 2017). Other complicating factors associated with this approach (such as differential binding of denaturant to folded and unfolded protein states) have been recently explored (Hall et al. 2018).

Conclusions

As discussed by Wakayama et al. (2019), the use of ionic liquids as the base solvent medium in the rational design of enzyme-based catalytic systems represents a point of profound divergence from the entire preceding historical record of enzymatic improvement (Park and Kazlauskas 2003; Itoh 2017). Whether occurring via natural evolution, rational design, or directed-evolution (Kuchner and Arnold 1997; Porter et al. 2016), improvements in enzyme function have previously taken place in aqueous solvent,¹⁰ which has constrained the reaction environment in terms of reactant/product solubility, reactant structure, reactant chemical state, and operative temperature range of the solution [typically ~ 0 – 100 °C]. With physical and chemical properties which can be vastly different from water (Seddon et al. 2000; Kazakov et al. 2018), the use of ionic liquids as the reaction solvent offers the potential to

¹⁰ There has been some development of modified enzymes for use in organic solvent (e.g., Reetz et al. 2010; Bommaris 2015).

alter the constraints associated with an aqueous reaction environment, thereby opening a door to previously un-thought of possibilities in the field of bio-catalysis (Brogan and Hallett 2016; Schröder 2017; Itoh 2017). Despite the excitement, the catalytic activity of enzymes is still predicated upon the polypeptide chain adopting the necessary three-dimensional structure capable of interacting with reactants to lower the relative free energy of their transition state.¹¹ As such, perhaps the most important factor in the use of ionic liquids to facilitate enzyme-based catalysis is their role in determining protein structure (Takekiyo et al. 2013; Patel et al. 2014; Reslan and Kayser 2018). In this short note, we have provided a brief overview of suitable empirical and theory-based strategies for the assessment of protein structure in response to changes made in either the temperature, pressure, or solvent composition and have attempted to explain the origin of the terms appearing within the analytical expressions.

There are a number of current examples of the use of ionic liquids for the enhancement of the enzyme-catalyzed production of commercially important chemicals (De Diego et al. 2005; Dang et al. 2007; Dabirmanesh et al. 2012; Kumar et al. 2017; Itoh 2017; Meyer et al. 2018). From a standard process engineering viewpoint, the target function for optimization in any commercial venture is the unit cost of production of the chemical product. For the process-design engineer, the alterable variables to achieve this enhancement in production cost are typically the system temperature and pressure, the solvent composition, the reactant concentrations, and the type of catalyst along with its chemical and structural state (Bommarius 2015). Use of ionic liquids as the solvent in enzyme-catalyzed reactions opens new areas of search-space with regard to all of the system variables, including the design and optimization of the enzyme catalyst itself and exploration of the effect of added cosolvents in regulating enzyme behavior (Hall 2002, 2006; Hall and Dobson 2006). Despite the existence of these new horizons, a proper understanding of the enzyme's protein structure along with its sensitivity to, and competition between, denaturation and aggregation (Constatinescu et al. 2010; Hall 2006) will be a vital component of intelligent scientific and business-based decision-making.

In our attempts to review this field, we were impressed with the large volume of published work. Indeed, the sentiment expressed in the prefatory comments of a recent monograph on ionic liquids (Koel 2009),

Ex sale et sole existunt omnia – From the salt and sun all things come forth,

seems to be apt in at least a few different ways. This review is different from the many others that relate to the use of ionic liquids in biological applications (e.g., see articles within the recent special issue edited by Benedetto and Galla 2018) in that, it has specifically attempted to fill the information gap necessary for the field to move beyond the commonly adopted practice of ranking the stability of proteins in ionic liquids in terms of conditions defining the unfolding midpoint. In this aim, the present review shares a common purpose with the review by Smiatek (2017) although the theoretical approaches are geared to different audiences in the sense that the review by Smiatek is developed using theory largely useful to computational scientists while the current review deals in observables and equations useful for “wet” experimentalists. We hope that the detailed and thermodynamically consistent description of existing assessment strategies presented in this review will be beneficial to those crossing into this bold new frontier of ionic-liquid-enabled enzyme design.

Acknowledgements DH acknowledges Dr. Nicholas Kanizaj, Dr. Bradley Stevenson, and Dr. Christoph Nitsche from the Research School of Chemistry at the ANU for helpful comments on an earlier version of this manuscript and Dr. Allen Minton, of the NIH, for conversations relating to the derivation of the Gibbs-Helmholtz equation.

Compliance with ethical standards

Funding information The work of DH was jointly supported by funds and resources associated with an Australian National University (ANU) Senior Research Fellowship, a Guest Associate Professor position at the Institute for Protein Research, Osaka University and a short term fellowship at the Laboratory of Biochemistry and Genetics of the NIH (NIDDK) in August of 2018. The receipt of a TOBITATE exchange program grant funded a 1-year exchange for RW to ANU. TOBITATE was funded by the Japan National Exchange Program part of the Ministry of Education, Culture Science and Technology (MEXT) of Japan.

Conflict of interest Ryota Wakayama declares that he has no conflict of interest. Susumu Uchiyama declares that he has no conflict of interest. Damien Hall declares that he has no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Appendix 1: Thermodynamics of open systems

For an open system (for which we may independently change either the applied pressure, the temperature, or the number/concentration of any of the non-reacting solution components), the difference in free energy between two equilibrium states is given in Eq. A1 (Van Holde et al. 2006a).

$$dG = \left(\frac{\partial G}{\partial T} \right)_{P, \{n_i\}} dT + \left(\frac{\partial G}{\partial P} \right)_{T, \{n_i\}} dP + \sum_{i=1}^Q \left(\frac{\partial G}{\partial n_i} \right)_{T, P, \{n_{\neq i}\}} dn_i \quad (\text{A1})$$

¹¹ The transition state is the structural and chemical state adopted by reactants which represents the major energetic obstacle on the reaction pathway defining their conversion to products.

Each of the partial derivative terms can be evaluated to produce the standard identities shown in Eq. A2,

$$\left(\frac{\partial G}{\partial T}\right)_{P,\{n_j\}} = -S_{\text{system}} \quad (\text{A2a})$$

$$\left(\frac{\partial G}{\partial P}\right)_{T,\{n_j\}} = V_{\text{system}} \quad (\text{A2b})$$

$$\left(\frac{\partial G}{\partial n_i}\right)_{T,P,\{n_{j \neq i}\}} = \mu_i = \mu_i^\circ + RT \ln \left(\frac{C_i}{C_i^\circ}\right) \quad (\text{A2c})$$

In Eq. A2, S and V are respectively the entropy and volume of the system and μ_i is the change in free energy of the system per fractional change in mole composition of a particular i th type of chemical component (and furthermore μ_i° is the change in free energy per fractional change in mole composition of the i th type of chemical component when the concentration of that component is at its standard state concentration C_i° typically a concentration of 1 M). Upon substitution of these identities into Eq. A1, we obtain Eq. A3,

$$dG = -S_{\text{system}}.dT + V_{\text{system}}.dP + \sum_{j=1}^Q \mu_j.dn_j \quad (\text{A3})$$

Importantly, Eq. A3 describes a change in free energy in terms of partial differentials with $Q + 2$ independent variables (where Q is the number of independently variable chemical components). If a change is made to the system, the overall change in free energy, dG , can be calculated through a basic book-keeping exercise according to Eq. A3 with each partial derivative evaluated in turn with independent variables updated as required for evaluation of a state variable. For a situation in which the temperature, pressure, and number of particles of a component can be reversibly altered, the derivative of the difference in free energy of the system with respect to the molar composition of the i th component is given by Eq. A4a. Upon realizing that the order of differentiation does not matter, we may re-express, $\frac{d(dG)}{dn_i}$, as the incremental change in chemical potential, $d\mu_i$, and likewise denote the various component terms in Eq. A4a as representing constrained partial changes of the chemical potential, $\partial\mu_i$, under the condition that $dn_j = 0$ for all components $j \neq i$ (Eq. A4b). The result described by Eq. A4b can be used to calculate the chemical potential of component i at a new set of conditions according to Eq. A4c.

$$\begin{aligned} \frac{d(dG)}{dn_i}_{T_1,P_1,\{n_i\}_1 \rightarrow T_2,P_2,\{n_i\}_2} &= -\left(\frac{\partial S}{\partial n_i}\right)_{T,P,\{n_j\}}.dT \\ &+ \left(\frac{\partial V}{\partial n_i}\right)_{T,P,\{n_j\}}.dP + \sum_{j=1}^Q \left(\frac{\partial \mu_j}{\partial n_i}\right)_{(n_i)_1 \rightarrow (n_i)_2; T,P,\{n_{j \neq i}\}}.dn_j \end{aligned} \quad (\text{A4a})$$

$$\begin{aligned} (d\mu_i)_{T_1,P_1,\{n_i\}_1 \rightarrow T_2,P_2,\{n_i\}_2} &= (\partial\mu_i)_{T_1 \rightarrow T_2; P,\{n_{j \neq i}\}} \\ &+ (\partial\mu_i)_{P_1 \rightarrow P_2; T,\{n_{j \neq i}\}} \\ &+ (\partial\mu_i)_{(n_i)_1 \rightarrow (n_i)_2; T,P,\{n_{j \neq i}\}} \end{aligned} \quad (\text{A4b})$$

$$\begin{aligned} (\mu_i)_{T_2,P_2,\{n_j\}_2} &= (\mu_i)_{T_1,P_1,\{n_j\}_1} \\ &+ (d\mu_i)_{T_1,P_1,\{n_j\}_1 \rightarrow T_2,P_2,\{n_j\}_2} \end{aligned} \quad (\text{A4c})$$

Systems at equilibrium and the definition of equilibrium constants

At constant temperature and pressure, Eq. A3 collapses to A5a. We may simplify Eq. A5a on the condition of chemical components participating in a chemical equilibrium. If components A and B exist in chemical equilibrium, such that, $A \rightleftharpoons B$ ¹² then we may express change in terms of a reaction progress variable ξ , such that $d\xi = |dn_A| = |dn_B|$ (Eq. A5b). Differentiation with respect to the reaction progress variable produces Eq. A5c.

$$dG = \sum_{i=1}^Q \mu_i.dn_i \quad (\text{A5a})$$

$$dG = \mu_B.d\xi - \mu_A.d\xi \quad (\text{A5b})$$

$$\frac{dG}{d\xi} = \Delta\mu_{AB} \quad (\text{A5c})$$

Equation A5c defines the condition for equilibrium in the sense that the reaction will proceed until the differential described in Eq. A5c is located at a minima, i.e., $\frac{dG}{d\xi} = 0$, from which Eq. A6 follows,

$$\Delta\mu_{AB} = 0 \quad (\text{A6a})$$

$$(\mu_A)_{T_1,P_1,\{n_{j \neq A}\}_1} = (\mu_B)_{T_1,P_1,\{n_{j \neq B}\}_1} \quad (\text{A6b})$$

Inserting the equation for the definition of the chemical potential from Eq. A2c yields Eq. A7.

$$(\mu_A^\circ)_{T_1,P_1,\{n_{j \neq A}\}_1} + RT_1 \ln \left(\frac{(C_A)_1}{C_A^\circ}\right) = (\mu_B^\circ)_{T_1,P_1,\{n_{j \neq B}\}_1} + RT_1 \ln \left(\frac{(C_B)_1}{C_B^\circ}\right) \quad (\text{A7})$$

From which it follows,

$$-RT_1 \ln \left(\frac{(C_B)_1.C_A^\circ}{(C_A)_1.C_B^\circ}\right) = (\mu_B^\circ)_{T_1,P_1,\{n_{j \neq B}\}_1} - (\mu_A^\circ)_{T_1,P_1,\{n_{j \neq A}\}_1} \quad (\text{A8a})$$

¹² We will revise this later for the case of denaturant induced unfolding.

$$\begin{aligned}
 (K_{AB})_{T_1, P_1, \{n_j\}_1} &= \left(\frac{(C_B)_1 \cdot C_A^*}{(C_A)_1 \cdot C_B^*} \right) \\
 &= \exp \left(\frac{- \left[(\mu_B^{\circ})_{T_1, P_1, \{n_j \neq B\}_1} - (\mu_A^{\circ})_{T_1, P_1, \{n_j \neq A\}_1} \right]}{RT_1} \right) \quad (\text{A8b})
 \end{aligned}$$

Calculating equilibrium constants at a new set of conditions

To understand how the equilibrium constant may change with changing conditions, it is important to realize that at any altered set of conditions (say $T_2, P_2, \{n_j\}_2$), we may express the equilibrium condition at constant temperature and pressure in an identical manner to Eq. A8b (Eq. A9).

$$\begin{aligned}
 (K_{AB})_{T_2, P_2, \{n_j\}_2} &= \left(\frac{(C_B)_2 \cdot C_A^*}{(C_A)_2 \cdot C_B^*} \right) \\
 &= \exp \left(\frac{- \left[(\mu_B^{\circ})_{T_2, P_2, \{n_j \neq B\}_2} - (\mu_A^{\circ})_{T_2, P_2, \{n_j \neq A\}_2} \right]}{RT_2} \right) \quad (\text{A9})
 \end{aligned}$$

We can relate the chemical potential of components A and B defined at a new set of conditions ($T_2, P_2, \{n_j\}_2$) to their chemical potentials defined under a previous set of conditions ($T_1, P_1, \{n_j\}_1$) by use of the equation for the total differential given in A4b (Eq. A10).

$$\begin{aligned}
 (\mu_A)_{T_2, P_2, \{n_j \neq A\}_2} &= (\mu_A)_{T_1, P_1, \{n_j \neq A\}_1} \\
 &\quad + (d\mu_A)_{T_1, P_1, \{n_j\}_1 \rightarrow T_2, P_2, \{n_j\}_2} \quad (\text{A10a})
 \end{aligned}$$

$$\begin{aligned}
 (\mu_B)_{T_2, P_2, \{n_j \neq B\}_2} &= (\mu_B)_{T_1, P_1, \{n_j \neq B\}_1} \\
 &\quad + (d\mu_B)_{T_1, P_1, \{n_j\}_1 \rightarrow T_2, P_2, \{n_j\}_2} \quad (\text{A10b})
 \end{aligned}$$

Inserting the result for Eq. A2c into Eq. A10, we have the condition for equilibrium after some change is made in temperature, pressure, or composition to a system at equilibrium (Eq. A11).

$$\begin{aligned}
 &(\mu_A^{\circ})_{T_1, P_1, \{n_j \neq A\}_1} + RT_1 \ln \left(\frac{(C_A)_1}{C_A^*} \right) + (d\mu_A)_{T_1, P_1, \{n_j\}_1 \rightarrow T_2, P_2, \{n_j\}_2} \\
 &= (\mu_B^{\circ})_{T_1, P_1, \{n_j \neq B\}_1} + RT_1 \ln \left(\frac{(C_B)_1}{C_B^*} \right) + (d\mu_B)_{T_1, P_1, \{n_j\}_1 \rightarrow T_2, P_2, \{n_j\}_2} \quad (\text{A11})
 \end{aligned}$$

In the main paper, we utilize Eq. A11 to derive analytical relations for the change in extent of protein unfolding with changes in environmental conditions.

Definition of equilibrium constants for the denaturant binding mechanism

To derive a set of equations which account for successive denaturant binding to various “sites” on the protein, we first develop the case for the reversible binding of denaturant to a specific site and then later relax the requirement for site specificity, to include the statistically degenerate case in which each site is considered equivalent. For interaction between a specific protein site and denaturant molecule ($P_{\text{site}} + D \rightleftharpoons P_{\text{site}}D$), the condition for the site-binding equilibrium and the site-binding equilibrium constant, $(k_{P_{\text{site}}D})_{T_1, P_1, \{n_j\}_1, (C_D)_1}$, are respectively given in Eq. A12a and b.

$$\mu_{P_{\text{site}}} + \mu_D = \mu_{P_{\text{site}}D} \quad (\text{A12a})$$

$$\begin{aligned}
 (k_{P_{\text{site}}D})_{T_1, P_1, \{n_j\}_1, (C_D)_1} &= \exp \left(\frac{- \left[\mu_{P_{\text{site}}D}^{\circ} - (\mu_{P_{\text{site}}}^{\circ} + \mu_D^{\circ}) \right]}{RT_1} \right) \quad (\text{A12b}) \\
 &= \frac{C_{PD}}{C_{P_{\text{site}}} \cdot (C_D)_1}
 \end{aligned}$$

If there are multiple ways of configuring a particular j th loading state of denaturant on the protein molecule, C_{PD_j} , then the site-binding constant must be modified by a statistical factor, F_{PD_j} , which accounts for all of the different possible ways that the PD_j state can be constructed, Eq. A13a. Incorporating this degeneracy into the equilibrium defines the relationship between the operative equilibrium constant, $(K_{PD_j})_{T_1, P_1, \{n_j\}_1, (C_D)_1}$, and the site-binding equilibrium constant, $(k_{P_{\text{site}}D})_{T_1, P_1, \{n_j\}_1, (C_D)_1}$, Eq. A13b (Hall et al. 2018).

$$C_{PD_j} = F_{PD_j} \cdot C_{P_{\text{site}}D} \quad (\text{A13a})$$

$$\begin{aligned}
 (K_{PD_j})_{T_1, P_1, \{n_j\}_1, (C_D)_1} &= \frac{C_{PD_j}}{C_{PD_{j-1}} \cdot (C_D)_1} \\
 &= \frac{F_{PD_j} \cdot C_{P_{\text{site}}D}}{F_{PD_{j-1}} \cdot C_{P_{\text{site}}} \cdot (C_D)_1} \\
 &= \left(\frac{F_{PD_j}}{F_{PD_{j-1}}} \right) \cdot (k_{P_{\text{site}}D})_{T_1, P_1, \{n_j\}_1, (C_D)_1} \quad (\text{A13b})
 \end{aligned}$$

For j denaturant molecules loaded onto a protein with a total of N_{PD} equivalent and indistinguishable sites, the statistical factor, F_{PD_j} , is given by Eq. A14 (which readers will note is functionally equivalent to Eq. 24c and d provided in the main text).

$$F_{PD_j} = \frac{N_{PD}!}{(N_{PD}-j)! \cdot j!} \quad (\text{A14})$$

Summation of all forms of protein-denaturant complex yields Eq. A15a. Using Eq. A13b and Eq. A15a may be expressed in an equivalent form known as the binding polynomial, Eq. A15b (Van Holde et al. 2006b).

$$(C_P)_{\text{TOTAL}} = \sum_{i=0}^{\infty} C_{PD_i} \quad (\text{A15a})$$

$$(C_P)_{\text{TOTAL}} = (C_P)_1 \cdot \sum_{i=0}^{\infty} F_{PD_i} \cdot (k_{P_{\text{site}}} D \cdot (C_D)_1)^i \quad (\text{A15b})$$

If additional consideration is given to the two different F and U forms of the protein, we may use the result given in Eq. A15b to produce an operative equilibrium constant of a type identical to that given in Eq. 25.

References

- Anfinsen CB (1973) Principles that govern the folding of protein chains. *Science* 181:223–230
- Aparicio S, Atilhan M, Karadas F (2010) Thermophysical properties of pure ionic liquids: review of present situation. *Ind Eng Chem Res* 49(20):9580–9595
- Aune KC, Tanford C (1969) Thermodynamics of the denaturation of lysozyme by guanidine hydrochloride. II. Dependence on denaturant concentration at 25. *Biochemistry* 8(11):4586–4590
- Auton M, Bolen DW (2005) Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proc Natl Acad Sci* 102:15065–15068
- Becktel WJ, Schellman JA (1987) Protein stability curves. *Biopolymers: Original Research on Biomolecules* 26(11):1859–1877
- Benedetto A, Galla HJ (2018) Editorial of the “ionic liquids and biomolecules” special issue. *Biophys Rev* 9(3):1–4
- Bommarius AS (2015) Biocatalysis: a status report. *Ann Rev Chem Biomol Eng* 6:319–345
- Brogan AP, Hallett JP (2016) Solubilizing and stabilizing proteins in anhydrous ionic liquids through formation of protein–polymer surfactant nanoconstructs. *J Am Chem Soc* 138(13):4494–4501
- Canchi DR, Garcia AE (2013) Cosolvent effects on protein stability. *Annu Rev Phys Chem* 64:273–293
- Chi EY, Krishnan S, Randolph TW, Carpenter JF (2003) Physical stability of proteins in aqueous solution: mechanism and driving forces in nonnative protein aggregation. *Pharm Res* 20(9):1325–1336
- Chum HL, Koch VR, Miller LL, Osteryoung RA (1975) Electrochemical scrutiny of organometallic iron complexes and hexamethylbenzene in a room temperature molten salt. *J Am Chem Soc* 97(11):3264–3265
- Constantinescu D, Weingärtner H, Herrmann C (2007) Protein denaturation by ionic liquids and the Hofmeister series: a case study of aqueous solutions of ribonuclease a. *Angew Chem Int Ed* 46(46):8887–8889
- Constantinescu D, Herrmann C, Weingärtner H (2010) Patterns of protein unfolding and protein aggregation in ionic liquids. *Phys Chem Chem Phys* 12(8):1756–1763
- Dabirmanesh B, Khajeh K, Ranjbar B, Ghazi F, Heydari A (2012) Inhibition mediated stabilization effect of imidazolium based ionic liquids on alcohol dehydrogenase. *J Mol Liq* 170:66–71
- Dang DT, Ha SH, Lee SM, Chang WJ, Koo YM (2007) Enhanced activity and stability of ionic liquid-pretreated lipase. *J Mol Catal B Enzym* 45(3–4):118–121
- De Diego T, Lozano P, Gmouh S, Vaultier M, Iborra JL (2005) Understanding structure–stability relationships of *Candida antarctica* Lipase B in ionic liquids. *Biomacromolecules* 6(3):1457–1464
- Dill KA (1985) Theory for the folding and stability of globular proteins. *Biochemistry* 24(6):1501–1509
- Dill KA, MacCallum JL (2012) The protein-folding problem, 50 years on. *Science* 338(6110):1042–1046
- Domańska U (2009). General review of ionic liquids and their properties. Chapter 1 of “Analytical Chemistry: ionic liquids in chemical analysis” (Ed. Koel, M), p. 1–71
- Eftink MR, Ionescu R (1997) Thermodynamics of protein unfolding: questions pertinent to testing the validity of the two-state model. *Biophys Chem* 64(1–3):175–197
- Fersht A (1999). A guide to enzyme catalysis and protein folding. *Structure and Mechanism in Protein Science*, “Protein Stability”, New York, W.H. Freeman Press, pp. 508–539 Chapter 17
- Fields PA (2001) Protein function at thermal extremes: balancing stability and flexibility. *Comp Biochem Physiol A Mol Integr Physiol* 129(2–3):417–431
- Greene RF, Pace CN (1974) Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, α -chymotrypsin, and β -lactoglobulin. *J Biol Chem* 249:5388–5393
- Hall D (2002) On the role of the macromolecular phase transitions in biology in response to change in solution volume or macromolecular composition: action as an entropy buffer. *Biophys Chem* 98(3):233–248
- Hall D (2003) The effects of tubulin denaturation on the characterization of its polymerization behavior. *Biophys Chem* 104(3):655–682
- Hall D (2006) Protein self-association in the cell: a mechanism for fine tuning the level of macromolecular crowding? *Eur Biophys J* 35(3):276–280
- Hall D (2012) Semi-automated methods for simulation and measurement of amyloid fiber distributions obtained from transmission electron microscopy experiments. *Anal Biochem* 421(1):262–277
- Hall D, Dobson CM (2006) Expanding to fill the gap: a possible role for inert biopolymers in regulating the extent of the ‘macromolecular crowding’ effect. *FEBS Lett* 580(11):2584–2590
- Hall D, Edskes H (2004) Silent prions lying in wait: a two-hit model of prion/amyloid formation and infection. *J Mol Biol* 336(3):775–786
- Hall D, Edskes H (2009) A model of amyloid’s role in disease based on fibril fracture. *Biophys Chem* 145(1):17–28
- Hall D, Edskes H (2012) Computational modeling of the relationship between amyloid and disease. *Biophys Rev* 4(3):205–222
- Hall D, Hirota N (2009) Multi-scale modelling of amyloid formation from unfolded proteins using a set of theory derived rate constants. *Biophys Chem* 140(1–3):122–128
- Hall D, Huang L (2012) On the use of size exclusion chromatography for the resolution of mixed amyloid aggregate distributions: I. Equilibrium partition models. *Anal Biochem* 426(1):69–85
- Hall D, Kinjo A, Goto Y (2018) A new look at an old view of denaturant induced protein unfolding. *Anal Biochem* 542:40–57
- Horinek D, Netz RR (2011) Can simulations quantitatively predict peptide transfer free energies to urea solutions? Thermodynamic concepts and force field limitations. *J Phys Chem A* 115(23):6125–6136
- Hurley FH, Wier TP (1951) The electrodeposition of aluminium from non-aqueous solutions at room temperature. *J Electrochem Soc* 98:207–212
- Itoh T (2017) Ionic liquids as tool to improve enzymatic organic synthesis. *Chem Rev* 117(15):10567–10607
- Jacquemin J, Husson P, Mayer V, Cibulka I (2007) High-pressure volumetric properties of imidazolium-based ionic liquids: effect of the anion. *J Chem Eng Data* 52(6):2204–2211
- Jaspe J, Hagen SJ (2006) Do protein molecules unfold in a simple shear flow? *Biophys J* 91(9):3415–3424
- Johnson CM, Fersht AR (1995) Protein stability as a function of denaturant concentration: the thermal stability of barnase in the presence of urea. *Biochemistry* 34(20):6795–6804
- Kamal MZ, Kumar V, Satyamurthi K, Das KK, Rao NM (2016) Mutational probing of protein aggregates to design aggregation-resistant proteins. *FEBS open bio* 6(2):126–134
- Kazakov, A, Magee JW, Chirico RD, Paulechka E, Diky V, Muzny CD, Kroenlein K, and Frenkel M (2018) “NIST standard reference database 147: NIST ionic liquids database - (ILThermo)”, Version 2.0, National Institute of Standards and Technology, Gaithersburg MD, 20899, <http://ilthermo.boulder.nist.gov>.

- Koel M (2009). Introduction to 'Analytical Chemistry: ionic liquids in chemical analysis' (Ed. Koel, M). p. xxvii–xxxii
- Kuchner O, Arnold FH (1997) Directed evolution of enzyme catalysts. *Trends Biotechnol* 15(12):523–530
- Kumar PK, Jha I, Venkatesu P, Bahadur I, Ebenso EE (2017) A comparative study of the stability of stem bromelain based on the variation of anions of imidazolium-based ionic liquids. *J Mol Liq* 246:178–186
- Lesch V, Heuer A, Tatsis VA, Holm C, Smiatek J (2015) Peptides in the presence of aqueous ionic liquids: tunable co-solutes as denaturants or protectants? *Phys Chem Chem Phys* 17(39):26049–26053
- Lima-Ramos J, Neto W, Woodley JM (2014) Engineering of biocatalysts and biocatalytic processes. *Top Catal* 57(5):301–320
- Marsh KN, Boxall JA, Lichtenthaler R (2004) Room temperature ionic liquids and their mixtures—a review. *Fluid Phase Equilib* 219(1): 93–98
- McPhie P, Ni YS, Minton AP (2006) Macromolecular crowding stabilizes the molten globule form of apomyoglobin with respect to both cold and heat unfolding. *J Mol Biol* 361(1):7–10
- Meyer LE, von Langermann J, Kragl U (2018) Recent developments in biocatalysis in multiphasic ionic liquid reaction systems. *Biophys Rev* 10:1–10
- Mezzenga R, Fischer P (2013) The self-assembly, aggregation and phase transitions of food protein systems in one, two and three dimensions. *Rep Prog Phys* 76(4):046601
- Mozhaev VV, Heremans K, Frank J, Masson P, Balny C (1996) High pressure effects on protein structure and function. *Proteins-Struct Funct Genet* 24(1):81–91
- Myers JK, Pace CN, Scholtz JM (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci* 4(10):2138–2148
- Naushad M, Al-Othman ZA, Khan AB, Ali M (2012) Effect of ionic liquid on activity, stability, and structure of enzymes: a review. *Int J Biol Macromol* 51(4):555–560
- Nicholson EM, Scholtz JM (1996) Conformational stability of the Escherichia coli HPr protein: test of the linear extrapolation method and a thermodynamic characterization of cold denaturation. *Biochemistry* 35(35):11369–11378
- O'Brien EP, Brooks BR, Thirumalai D (2009) Molecular origin of constant m-values, denatured state collapse, and residue-dependent transition midpoints in globular proteins. *Biochemistry* 48(17):3743–3754
- Oberhauser AF, Hansma PK, Carrion-Vazquez M, Fernandez JM (2001) Stepwise unfolding of titin under force-clamp atomic force microscopy. *Proc Natl Acad Sci* 98(2):468–472
- Pace CN, (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. In *Methods in enzymology* (Vol. 131, pp. 266–280). Academic Press
- Pace CN, Shaw KL (2000) Linear extrapolation method of analyzing solvent denaturation curves. *Proteins: Structure, Function, and Bioinformatics* 41(S4):1–7
- Pandharipande PP, Makhatadze GI (2016) Applications of pressure perturbation calorimetry to study factors contributing to the volume changes upon protein unfolding. *Biochimica Biophysica Acta (BBA)-General Subj* 1860(5):1036–1042
- Park S, Kazlauskas RJ (2003) Biocatalysis in ionic liquids—advantages beyond green technology. *Curr Opin Biotechnol* 14(4):432–437
- Patel R, Kumari M, Khan AB (2014) Recent advances in the applications of ionic liquids in protein stability and activity: a review. *Appl Biochem Biotechnol* 172(8):3701–3720
- Pfeil W, (2012). Protein stability and folding supplement 1: a collection of thermodynamic data. Springer Science & Business Media
- Porter JL, Rusli RA, Ollis DL (2016) Directed evolution of enzymes for industrial biocatalysis. *ChemBioChem* 17(3):197–203
- Privalov PL (2007) Thermodynamic problems in structural molecular biology. *Pure Appl Chem* 79(8):1445–1462
- Privalov PL, Khechinashvili NN (1974) A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *J Mol Biol* 86(3):665–684
- Reetz MT, Soni P, Fernandez L, Gumulya Y, Carballeira JD (2010) Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method. *Chem Commun* 46(45):8657–8658
- Reslan, M. and Kayser, V., (2018). Ionic liquids as biocompatible stabilizers of proteins. *Biophys Rev* pp1–13
- Roberts CJ (2007) Non-native protein aggregation kinetics. *Biotechnol Bioeng* 98(5):927–938
- Rogers TA, Bommarius AS (2010) Utilizing simple biochemical measurements to predict lifetime output of biocatalysts in continuous isothermal processes. *Chem Eng Sci* 65(6):2118–2124
- Santoro MM, Bolen DW (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl. alpha-chymotrypsin using different denaturants. *Biochemistry* 27(21):8063–8068
- Sasahara K, Sakurai M, Nitta K (2001) Pressure effect on denaturant-induced unfolding of hen egg white lysozyme. *Proteins Struct Funct Bioinforma* 44(3):180–187
- Satish L, Rana S, Arakha M, Rout L, Ekka B, Jha S, Dash P, Sahoo H (2016) Impact of imidazolium-based ionic liquids on the structure and stability of lysozyme. *Spectrosc Lett* 49(6):383–390
- Schellman JA (1956) The stability of hydrogen-bonded peptide structures in aqueous solution. *C R Trav Lab Carlsb Ser Chim* 29:230–259
- Schellman JA (2002) Fifty years of solvent denaturation. *Biophys Chem* 96(2–3):91–101
- Schröder C (2017) Proteins in ionic liquids: current status of experiments and simulations. *Top Curr Chem* 375(2):25
- Seddon KR, Stark A, Torres MJ (2000) Influence of chloride, water, and organic solvents on the physical properties of ionic liquids. *Pure Appl Chem* 72(12):2275–2287
- Sheldon R (2001) Catalytic reactions in ionic liquids. *Chem Commun* 23: 2399–2407
- Shiao DF, Lumry R, Fahey J (1971) Chymotrypsinogen family of proteins. XI. Heat-capacity changes accompanying reversible thermal unfolding of proteins. *J Am Chem Soc* 93(8):2024–2035
- Smiatek J (2017) Aqueous ionic liquids and their effects on protein structures: an overview on recent theoretical and experimental results. *J Phys Condens Matter* 29:233001
- Stefani M, Dobson CM (2003) Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J Mol Med* 81(11):678–699
- Stevenson BJ, Liu JW, Ollis DL (2008) Directed evolution of yeast pyruvate decarboxylase 1 for attenuated regulation and increased stability. *Biochemistry* 47(9):3013–3025
- Street TO, Bolen DW, Rose GD (2006) A molecular mechanism for osmolyte-induced protein stability. *Proc Natl Acad Sci* 103: 13997–14002
- Sturtevant JM (1977) Heat capacity and entropy changes in processes involving proteins. *Proc Natl Acad Sci* 74(6):2236–2240
- Suarez PA, Selbach VM, Dullius JE, Einloft S, Piatnicki CM, Azambuja DS, de Souza RF, Dupont J (1997) Enlarged electrochemical window in dialkyl-imidazolium cation based room-temperature air and water-stable molten salts. *Electrochim Acta* 42(16):2533–2535
- Takekiyo T, Koyama Y, Yamazaki K, Abe H, Yoshimura Y (2013) Ionic liquid-induced formation of the α -helical structure of β -lactoglobulin. *J Phys Chem B* 117(35):10142–10148
- Tanford C, (1968). Protein denaturation. In *Advances in protein chemistry* (Vol. 23, pp. 121–282). Academic Press
- Tanford C, (1970). Protein denaturation: part C.* Theoretical models for the mechanism of denaturation. In *Advances in protein chemistry* (Vol. 24, pp. 1–95). Academic Press
- Van Holde KE, Johnson WC and Ho PS, (2006a). Principles of physical biochemistry. Chapter 2: thermodynamics and biochemistry

- Van Holde KE, Johnson WC and Ho PS, (2006b). Principles of physical biochemistry. Chapter 13. Macromolecules in solution: thermodynamics and equilibria
- Vekariya RL (2017) A review of ionic liquids: applications towards catalytic organic transformations. *J Mol Liq* 227:44–60
- Von Hippel PH, Wong KY (1965) On the conformational stability of globular proteins—the effects of various electrolytes and nonelectrolytes on the thermal ribonuclease transition. *J Biol Chem* 240(10):3909–3923
- Wakayama R, Uchiyama S, and Hall D. (2019). Ion solutions and protein folding - old strategies for new solvents. Chapter 5 of ‘Protein aggregation and solubility’ edited by Y. Kuroda and F. Arisaka. CMC Publishing House
- Walden P, (1914) Ueber die Molekulargröße und elektrische Leitfähigkeit einiger geschmolzenen Salze. *Bulletin de l’Académie Imperiale des Sciences de St. Petersburg*, pp. 405–422
- Weingärtner H (2008) Understanding ionic liquids at the molecular level: facts, problems, and controversies. *Angew Chem Int Ed* 47(4):654–670
- Weingärtner H, Cabrele C, Herrmann C (2012) How ionic liquids can help to stabilize native proteins. *Phys Chem Chem Phys* 14:415–426
- Welton T (1999) Room-temperature ionic liquids. Solvents for synthesis and catalysis. *Chem Rev* 99(8):2071–2084
- Wilkes JS (2002) A short history of ionic liquids – from molten salts to neoteric solvents. *Green Chem* 4:73–80
- Wilkes JS, Zaworodtko MJ (1992) Air and water stable 1-ethyl-3-methylimidazolium based ionic liquids. *J Chem Soc Chem Commun* 0:965–967
- Wilkes JS, Levinsky JA, Wilson RA, Hussey CA (1982) Dialkyl imidazolium chloroaluminate melts: a new class of room-temperature ionic liquids for electrochemistry, spectroscopy, and synthesis. *Inorg Chem* 21:1263–1264
- Zhao H, Xia S, Ma P (2005) Use of ionic liquids as ‘green’ solvents for extractions. *J Chem Technol Biotechnol: International Research in Process, Environmental & Clean Technology* 80(10):1089–1096

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.