Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Protein-complex stability in cells and *in vitro* under crowded conditions

Samantha S Stadmiller[1] and Gary J Pielak[1,2,3,4]

Biology is beginning to appreciate the effects of the crowded and complex intracellular environment on the equilibrium thermodynamics and kinetics of protein folding. The next logical step involves the interactions between proteins. We review quantitative, wet-experiment based efforts aimed at understanding how and why high concentrations of small molecules, synthetic polymers, biologically relevant cosolutes and the interior of living cells affect the energetics of protein-protein interactions. We then address popular theories used to explain the effects and suggest expeditious paths for a more methodical integration of experiment and simulation.

Addresses
[1] Department of Chemistry, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599-3290, USA

[2] Department of Biochemistry and Biophysics, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599, USA
[3] Lineberger Comprehensive Cancer Center, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599, USA
[4] Integrative Program for Biological and Genome Sciences, University of North Carolina-Chapel Hill, Chapel Hill, NC 27599, USA

Corresponding author: Pielak, Gary J (gary_pielak@unc.edu)

Current Opinion in Structural Biology 2021, **66**:183–192

This review comes from a themed issue on **Folding and binding**.

Edited by **Vic Arcus** and **Margaret Cheung**

https://doi.org/10.1016/j.sbi.2020.10.024

Introduction

'Biological macromolecules have evolved to function in the crowded conditions characteristic of intracellular milieu, so it is remarkable, not to say remiss, that most investigations of the properties of such macromolecules are still carried out in uncrowded buffers.' A great deal has changed since John Ellis wrote that sentence in his 2001 review on macromolecular crowding in *Current Opinions in Structural Biology* [1]. Here, we focus on one particular aspect of crowding, its effect on the energetics of protein-protein interactions (Figure 1). Beyond even Ellis' major concern, understanding these effects is important because almost two-thirds of disease-associated missense mutations perturb protein complexes [2].

> "Thermodynamics exhibits no curiosity, certain things are poured into its hopper, certain others emerge according to the laws of the machine." G. N. Lewis and M. Randall [3]

The reaction describing the association of proteins A and B to form the heterodimer complex A-B, is written as

$$A + B \rightleftharpoons A - B \tag{1}$$

For the first part of the Introduction, the reaction is considered to occur in dilute buffered aqueous solution near physiological pH, that is, modified standard state conditions. The equilibrium constant for dissociation, K_D, is written in terms of the molar concentrations, c or the rate constants for formation and dissociation, k_{on} and k_{off}, respectively, as follows:

$$K_D = \frac{C_A \cdot C_B}{C_{A-B}} = \frac{k_{off}}{k_{on}} \tag{2}$$

K_D has the units of concentration. The smaller its value, the more likely (i.e. the stronger) the interaction.

The most straightforward binding experiment involves fixing the concentration of protein A and varying the concentration of protein B (or vice versa). Such data are often plotted as a binding isotherm (Figure 2) with the fraction bound (f_b) on the y-axis, the concentration of B (or A), for heterodimerization, on the x-axis and the data fitted to the equation

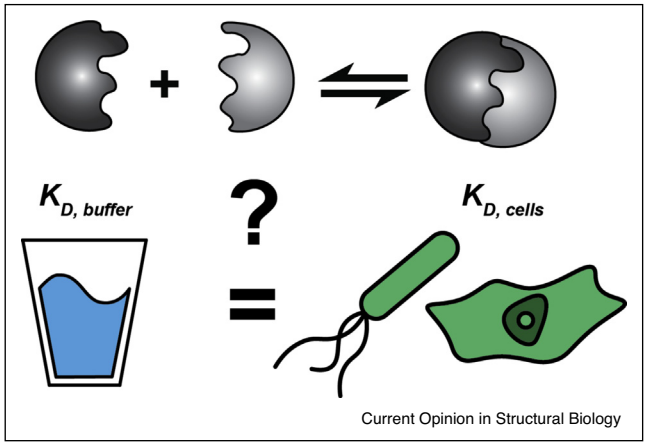
$$f_b = \frac{([A] + [B] + K_D) - \sqrt{([A] + [B] + K_D) - 4[A][B]}}{2[A]} \tag{3}$$

to yield the dissociation constant, K_D. When the product is a homodimer (i.e. A = B) and the total protein concentration is P_T, the data are fitted to the equation

$$f_b = \frac{4P_T + K_D - \sqrt{K_D + 8P_T K_D}}{4P_T} \tag{4}$$

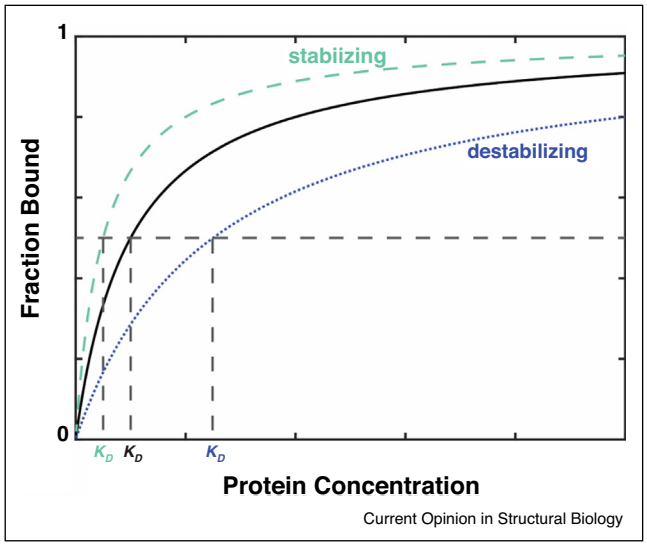
Stronger and weaker binding moves the isotherm to the left or right, respectively. K_D values can be converted to the free energies of dissociation($\Delta G^{\circ}_D{}'$) using the Gibbs equation,

Figure 1



The overarching question: are the energetics for dimerization the same in buffer as they are in living cells?

Figure 2



Simulated binding isotherms for a dimer in buffer (black, solid) and in destabilizing (blue, short dash) or stabilizing (green, long dash) conditions.

$$\Delta G^{\circ}_{D}\prime = -RT\ln(K_D) = -\Delta G^{\circ}_{A}\prime \tag{5}$$

where R is the gas constant, T is the absolute temperature, and $\Delta G^{\circ}_{A}\prime$ is the modified standard-state free energy of association.

Next, we consider the effect of adding a cosolute (i.e. a solute in addition to the protein(s) and buffer), at concentrations from tens to hundreds of g/L. Given that partial-specific volumes of cosolutes range from about 0.6 mL/g to 0.9 mL/g cosolute volume occupancies can reach 30% at 300 g/L. This occupancy is in the range of 20%–40% that is found in living cells [4].

We begin with studies of cosolutes with molecular weights up to several hundred Daltons (Da) because such cosolutes represent natural osmolytes [5] and the monomers that comprise synthetic polymers, which are discussed later. We then shift our focus to cosolutes with molecule weights from a few kDa to MDa. Such cosolutes include synthetic polymers, individual proteins, collections of proteins in cell lysates, and those comprising the interior of intact living cells. We consider synthetic polymers because, although less physiologically relevant, they are often used to mimic the cellular interior and are useful for stabilizing protein based drugs [6]. Commonly used synthetic polymers include polyethylene glycol (PEG), polyvinylpyrrolidone (aka povidone), the sucrose polymer Ficoll, and the glucose polymer dextran.

The effects of cosolutes are quantified by comparing equilibrium-constants or rate-constants obtained in buffer alone to the constants acquired under crowded conditions. Equilibrium constants should be written in terms of thermodynamic activity, α, of each component, i:

$$\alpha_i = \gamma_i C_i \tag{6}$$

The effect of nonideal conditions is encoded in the unitless activity coefficient, γ. In Eq. 2 we assumed that γ is one and the activity equals the molar concentration, C, under the nearly ideal conditions in dilute buffer.

Crowding changes γ, such that the K_D under dilute solution conditions must be modified by the ratio of the activity coefficients to give the dissociation constant under crowded conditions.

$$K_{D,crowd} = K_D\left(\frac{\gamma_{A-B}}{\gamma_A\gamma_B}\right) \tag{7}$$

Thus, the crowding effects are contained in the ratio of $K_{D,crowd}$ to K_D. The effects can then be propagated into the changes in the equilibrium binding free-energy, enthalpy and entropy as well as activation parameters, if the kinetics are known [7].

Table 1 lists, in approximate chronological order, studies of protein-protein interactions studied at high cosolute concentrations. Most efforts quantify K_D using equilibrium techniques or by measuring k_{on} and k_{off}. In some instances, only one rate constant is determined, and some efforts focus on diffusion. We summarize the results from investigations using small molecules, synthetic polymers,

Table 1

Cosolute effects on protein complex stability near room temperature and neutral pH

Protein	Complex	Cosolute(s)	Method	Parameters
Modified apomyoglobin [11]	Globular homodimer	PEG, lysozyme, ribonuclease, β-lactoglobulin	Fluorescence polarization	+/− dimer
Cytochrome c/ cytochrome b_5 cytochrome c/ cytochrome *c* oxidase [12]	Globular/globular	Glycerol	Fluorescence spectroscopy	*K*
Cytochrome c/ cytochrome *c* oxidase [13]	Globular /globular	EG, glycerol, glucose, sucrose	Absorbance spectroscopy	*K*
Cytochrome *c*- cytochrome *c* peroxidase [14]	Globular/globular	Glucose, sucrose, stachyose	ITC	K, $H^{o'}$ $\Delta S^{o'}$
HyHEL-5/BWQL [15]	Antibody/globular	EG, glycerol, betaine	Fluorescence polarization, stopped-flow fluorescence spectroscopy, ITC	k_{on}, k_{off}, K
Concanavilin A [16]	Homodimer/tetramer	TMAO, betaine, proline, sarcosine, sorbitol, sucrose, trehalose, urea	CD	*K*
Barnase/barstar [17]	Globular/globular	Povidone-1300, sucrose	Stopped-flow fluorescence spectroscopy	k_{on}
apo-Mb [18]	Globular homodimer	RNase A, HSA	psec-resolved fluorescence anisotropy	apparent *K*
TEM-1/BLIP [19]	Globular/globular	EG, PEG- (200, 1000, 3500, 8000), Ficoll-70, Haemaccel	Stopped-flow fluorescence spectroscopy	k_{on}, k_{off}
TEM-1/BLIP [20]	Globular/globular	Glycerol, sucrose, EG, PEG (100, 600, 1000, 3350, 6000, 8000), Ficoll-70	Fluorescence anisotropy, FCS, stopped-flow fluorescence spectroscopy	k_a, D_t, D_r
Cytochrome *f*/ plastocyanin [21]	Soluble part of membrane protein protein/globular	Ficoll-70, dextran-70, glycerol, sucrose, ethane diol	Stopped-flow spectrophotometry	k_{on}
SOD/xanthine oxidase [22]	Globular/globular multi-enzyme complex	Glycerol, PEG-2000, PEG-10000, Ficoll-70, dextran-70	Intrinsic fluorescence spectroscopy & resonant mirror biosensors	k_a, k_d, K
SH3-peptide PDZ-peptide [23]	Globular/peptide	*E. coli* lysate	FRET and flow cytometry	*K*
α-chymotrypsin [24]	Globular homodimer	Sucrose, glucose, raffinose	AUC	*K*
MAPK Ste11, Ste7, Fus3/ Ste5 [25]	Globular/disordered	Yeast	FCCS	*K*
CDC42/N-WASp CDC42/CRIB CSC42/IRSp53 [26]	Globular/globular	CHO cells	SW-FCCS	*K*
CDC42/IQGAP1 [27]	Globular/globular	CHO cells and zebrafish embryos	SW-FCCS	*K*
TEM/BLIP barnase/barstar [28]	Globular/globular	Glucose, EG, PEG-(600, 1000, 8000), dextran-6	SPR & ITC	k_a, k_d, K
θ- and ε- subunits of polymerase III holoenzyme [29]	Globular/globular	Ficoll-70, dextran-(6, 40, 70, 100, 150)	Intrinsic tryptophan fluorescence spectroscopy	*K*
SOD/catalase [30]	Globular/globular multi-enzyme complex	dextran-70, Ficoll-70, PEG-2000	Intrinsic fluorescence spectroscopy	K, $H^{o'}$ $\Delta S^{o'}$
3CL peptidase [31]	Globular homodimer	PEG-(600, 6000) and BSA	Fluorescence enzyme activity assay	*K*
TEM-1/BLIP [32]	Globular/globular	HeLa cells lysate	FRET & FRAP	k_a, k_d, $t_{1/2}$
TEM-1/BLIP [33]	Globular/globular	dextran-40, PEG-20 000	Stopped-flow fluorescence spectroscopy	K, $H^{o'}$ $\Delta S^{o'}$
Calmodulin/target peptide [34]	Globular/peptide	sucrose, Ficoll-70, dextran-10	ITC, FCS, stopped-flow fluorescence spectroscopy	k_a, k_d, K

Table 1 (*Continued*)

Protein	Complex	Cosolute(s)	Method	Parameters
GB1 [35]	Side-by-side homodimer	Urea, EG, PEG-8000, sucrose, Ficoll-70, BSA, TMAO, *E. coli* cytosol	^{19}F NMR	*K*
GAPDH/PGK [36]	Globular /globular multi-enzyme complex	U-2 OS cells	FRET microscopy/cell volume perturbation	*K*
GB1 [37]	Side-by-side homodimer	BSA, lysozyme	^{19}F NMR	*K*
GB1 [38•]	Side-by-side homodimer	TMAO, alanine, trehalose, β-alanine, sarcosine, betaine, proline, sorbitol, sucrose, urea, glycerol	^{19}F NMR	*K*
GB1 [39•]	Side by side- & domain swapped- homodimers	EG, PEG-8000, Ficoll-70, sucrose, BSA, lysozyme	^{19}F NMR	*K*
PHD/methylated peptides [40]	Globular /peptide	Ficoll-70	^{1}H-^{15}N NMR	*K*
ACTR/NCBD [41••]	Disordered/disordered	EG, DEG, TEG, PEG-(200, 400, 1000, 2000, 4500, 6000, 35 000)	Confocal single-molecule FRET	k_a, k_d, *K*, D_t
XIAP [42]	Disordered/molten globule homodimer	Ficoll, HeLa cells, HeLa cell lysates	DEER spectroscopy	*K*

Abbreviations: ACTR, disordered activation domain of the steroid receptor coactivator 3; apoMb, apomyoglobin; AUC, analytical ultracentrifugation; BLIP, β-lactamase inhibitor protein; BSA, bovine serum albumin; BWQL, bobwhite quail lysozyme; CD, circular dichroism spectropolarimetry; CHO, Chinese hamster ovary; DEG, diethylene glycol; D_t, translational diffusion coefficient; D_r, rotational diffusion coefficient; EG, ethylene glycol; FCCS, FCS, fluorescence correlation spectroscopy; FRAP, fluorescence recovery after photobleaching; FRET, Förster resonance energy transfer; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; GB1, streptococcal B1 domain of protein G; HSA, human serum albumin; HyHEL-5, anti-hen egg lysozyme monoclonal antibody; IQGAP1, Ras GTPase-activating-like protein; IRSp53, insulin receptor substrate protein; ITC, isothermal titration calorimetry; *K*, equilibrium constant for association or dissociation; k_a, association rate constant; k_d, dissociation rate constant; MAPK, mitogen-activated protein kinase; NCBD, nuclear coactivator binding domain of CBP/p300; NMR, nuclear magnetic resonance spectroscopy; N-WASp, neural Wiskott-Aldrich syndrome protein; PHD, plant homeodomain; PEG, polyethylene glycol; PGK, phosphoglycerate kinase; RNase, ribonuclease; SH3, src-homology 3; SOD, superoxide dismutase; SPR, surface plasmon resonance; SW-FCCS, single wavelength fluorescence cross-correlation spectroscopy; TEG, triethylene glycol; TEM-1, β-lactamase; TMAO, trimethylamine N-oxide; $t_{1/2}$, half-time of fluorescence recovery; XIAP, X-chromosome-linked inhibitor of apoptosis.

more biologically relevant cosolutes and living cells. Values of $K_D/K_{D,crowd}$ are as large as 100, which translates to a $\Delta\Delta G^{o'}_D$ of less than 3 kcal/mol at physiological temperatures. Such changes seem small, but because biological macromolecules and their complexes are stabilized by cooperative interactions [8], small changes in free energy can have large biological effects. For instance, increasing the incubation temperature of alligator eggs by 4°C, corresponding to 0.01 kcal/mol of thermal energy, changes the sex of hatchlings from 100% female to 100% male [9] (See Ref. [10] for more examples). We then discuss the strengths and weaknesses of the methodologies used to acquire the data, the models used to interpret the results, and our ideas about what is needed to create a molecular level picture of crowding effects.

Stability in solutions of small cosolutes [12–17,19–21,24,34,35,38•,39•]

Most small molecule cosolutes, including naturally occurring osmolytes [5], are minimally perturbing or stabilizing, but there are exceptions. For instance, urea is always destabilizing, which is unsurprising given its well-known and well understood effect on protein stability [43]. Ethylene glycol can be stabilizing or destabilizing [15,19] depending on the complex. Destabilization of a protein complex by sugars and betaine (trimethyl glycine, [15]) is surprising because these cosolutes almost always stabilize proteins. As discussed later, the difference may arise from the amount and identity of the surface exposed upon denaturation and dissociation. Whatever their effect on stability, protein complexes tend to follow Stokes-Einstein behavior in solutions containing high g/L concentrations of small molecule cosolutes, which means the ratio of k_{on} values in cosolute and buffer alone is directly proportional to the ratio of the macroscopic viscosities of the two solutions. Stokes-Einstein behavior, however, is usually not observed in solutions containing high concentrations of synthetic polymers.

Stability in solutions of synthetic polymers [11,17,19–22,28–31,33–35,39•,40,41••,42]

We discuss these results in terms of three questions. The first question involves the relationship between k_{on} and viscosity. Does the protein and its complex experience the macroscopic viscosity? The other two questions deal with equilibrium thermodynamics. Is there a stabilizing macromolecular effect? That is, does the polymer have a larger stabilizing effect than its monomer? The third question involves polymer molecular weight. Specifically, does polymer molecular weight affect stability?

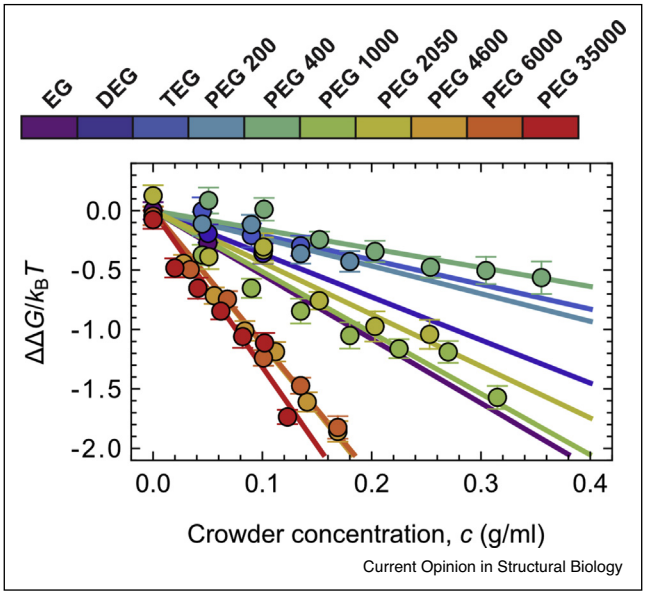
Does the protein and its complex experience the macroscopic viscosity? The relationship between k_{on} and the macroscopic viscosity in polymer solutions is the subject of many studies listed in Table 1. With one exception [19], the relationship is clear. At a fixed g/L concentration, the protein feels less of the macroscopic viscosity as the

molecular weight of the polymer increases. These positive deviations from Stokes-Einstein behavior [20] occur because the chains form a porous mesh, which means, crudely, that protein diffusion occurs mostly in the water enclosed by the chains. The effect is well described by Kozer and Schreiber [19], who also report the exception, Haemaccel. This 35-kDa colloid-forming collagen hydrolysate is not synthetic, but it is included here because, like synthetic polymers, this mixture of proteins lacks stable tertiary structure. In Haemaccel, k_{on} for the complex remains linearly related to the macroscopic viscosity as is observed for the small molecule cosolute ethylene glycol. As suggested by the authors, despite its similarity in molecular weight to the other polymers, Haemaccel does not form a mesh. Additional efforts comparing the effects of synthetic polymers to those of disordered protein crowders would be welcome.

Is there a stabilizing macromolecular effect? Many investigators associate macromolecular crowding with the stabilization of proteins and their complexes. A crude test of this idea is whether the polymer is more stabilizing than the same g/L concentration of the monomer. The test is crude because it ignores the end effect of the monomer [e.g. Ref. 44]. The results are few and mixed. Phillip *et al.* [28] report the opposite of the expected effect for PEG/ethylene glycol and dextran/glucose. Our lab reported a weak stabilizing effect for PEG and Ficoll/sucrose on a side-by-side dimer and a larger effect for a nearly-identical domain-swapped dimer made from the streptococcal B1 domain of protein G (GB1) [39•]. Zosel *et al.* [41••] report a strong macromolecular effect for PEG. Yang *et al.* [42] report that Ficoll decreases the stability of the homodimeric complex comprising the first domain of the X-chromosome-linked inhibitor of apoptosis (XIAP) compared to buffer alone. The results of more such tests are required to understand the basis of macromolecular crowding.

Does polymer molecular weight affect stability? There are few detailed investigations. Wilf and Minton [11] studied the effect of 6 kDa-PEG and 20 kDa-PEG. They observed no effect up to 250 g/L, the highest concentration studied. Zhou *et al.* [22] examined the effect of 2 kDa and 20 kDa-PEG on the interaction between superoxide dismutase to xanthine oxidase and observed an increase in stabilization with increased size at a common PEG concentration. Zosel *et al.* [41••] describe the most detailed effort by studying PEGs from the monomer up to 35 kDa on the stability of the dimer comprising an intrinsically-disordered protein, steroid receptor coactivator 3 (ACTR) and the molten globule-like nuclear coactivator binding domain of CBP/p300 (NCBD). They observe (Figure 3) that at a fixed g/L concentration, stability increases with increasing molecular weight. They also assessed the influence of PEG concentration and report that $\Delta\Delta G_D^{o'}/(g/L)$ is linear and its magnitude

Figure 3



Effect of PEG molecular weight and concentration on ACTR-NCBD dimer stability [41••]. k_B is the universal gas constant per molecule.

increases with increasing PEG molecular weight. Additional detailed systematic studies of size effects are absolutely required.

Stability under more physiologically relevant conditions [11,18,23,31,35,37,39•,45]

Studies with synthetic polymer crowders are important for understanding the mechanism of crowding and stabilizing protein-based drugs [6], but cells are mostly crowded with globular proteins. Shape and surface are obvious differences between synthetic polymers and globular proteins. In terms of shape, synthetic polymers at high concentrations form a porous mesh with a large macroscopic viscosity [46], but globular proteins, as their name implies, are compact objects that generally have a much smaller effect on macroscopic viscosity. Turning to surfaces, those of the usual synthetic polymers lack the variety of functional groups found on proteins, which are studded with groups capable of a variety of interactions. These groups include positive and negative charges at the termini and on the side chains of aspartic- and glutamic- acids, lysine, arginine and histidine, which affect a protein's isoelectric point (pI), as well as the hydrogen bond donors and acceptors. A key difference between protein crowders and common synthetic polymers is that proteins are charged.

Minton and Wilf were probably the first to assess protein complex formation in protein cosolutes [11]. Lysozyme (pI 10), ribonuclease A (pI 10), and β-lactoglobulin (pI 5) all brought about dimerization of fluorescently modified myoglobin (pI 9) at neutral pH. The ribonuclease effect

was confirmed [18] and, in the later study, human serum albumin (pI 4.7) did not promote self-association. For all except β-lactoglobulin, the results are consistent with the idea that charge-charge repulsions between myoglobin and the crowder protein promote dimerization. Phillip *et al.* [32] examined k_{on} and k_{off} of the TEM1 β-lactamase-β-lactamase inhibitor complex and several variants in human (HeLa) cell lysates (and cells, vide infra) but there is no combination of data that permit comparison of K_a values in lysate and buffer. The less compact side-by-side dimer and the more compact domain-swapped dimer made from the streptococcal B1 domain of protein G (pI 4.5) were examined in bovine serum albumin (pI 4.5) and lysozyme (pI 9.7) at pH 7.4 [35,39•]. Compared to buffer, the dimers are more stable in albumin and less stable in lysozyme. Additionally, the side-by-side dimer is more stable in freeze-dried *E. coli* cytosol than in buffer [35], which contains a majority of polyanionic proteins. These results are also consistent with the idea that crowder charge plays a key role.

We end this section with a word of caution about interpretations using enzyme activity. The SARS-CoV 3CL peptidase (pI 6.7) is only active as a homodimer [31]. Bovine serum albumin increases the activity of the enzyme, which suggests a role for repulsive electrostatics, but this cosolute does not stabilize the dimer.

In summary, the surface charge on protein-based crowders can control their effect on protein complex formation and stability. Examining protein dimer stability in protein cosolutes shows the effect of charge. Specifically, negatively charged protein crowders increase the stability of negatively charged protein complexes. Additional studies, including use of a larger variety of protein cosolutes in solutions at a variety of pH values, will provide a deeper understanding of the role of electrostatics.

Stability in cells [25–27,32,36,42,47,48]

These are the most technically challenging experiments and as shown in Table 1, most efforts use fluorescence detection to assess complex formation. Maeder *et al.* [25] made some of the first measurements of K_D in living cells using yeast. Unfortunately, the dilute solution values were not reported, and therefore the results cannot be used to assess the effect of the intracellular environment. Sudhaharan *et al.* [26] studied the interactions between the RhoGTPase Cdc42 and three of its effector proteins in Chinese hamster ovary cells. Comparison with *in vitro* data shows that the intracellular environment decreases K_D by about a factor of two. Shi *et al.* [27] quantified complex formation between a Cdc42 variant and an actin-binding scaffolding protein in zebra fish embryos. The K_D value of 100 nM in embryos is about fivefold larger than the value determined in buffer. Phillip *et al.* [32] identified a small decrease in k_a for the TEM1 β-lactamase-β-lactamase inhibitor complex in HeLa cells compared to buffer. They also showed that increasing the positive charge on one of the proteins decreases k_{on} but decreasing the positive charge has no effect. Considering the fact that the majority of proteins in most cells are polyanions [49], these data also suggest a role for charge. However, K_D in HeLa cells was not measured.

Sukenik *et al.* [36] assessed complex formation between the glycolytic enzymes, glyceraldehyde 3-phosphate dehydrogenase and phosphoglycerate kinase in human U-2 OS cells. The stoichiometry changes from 1:1 in buffer to 2:1 in cells. The K_D values, which, to make the units the same, were compared by taking the square root of the value determined in cells, decrease from 20 μM in buffer to 14 μM in cells. The most important conclusion, however, is that quinary interactions [50] can stabilize multi-enzyme complexes [51].

The K_a of the homodimeric XIAP complex has been quantified in HeLa cells [42]. The cellular interior destabilizes the protein compared to buffer. The crystal structure of the dimer shows that it is stabilized by a salt bridge. A possibility is that the higher ionic strength in cells breaks this interaction, but as discussed above, Ficoll, which is uncharged, also decreases its stability.

Progress has also been made in efforts to quantify protein-protein interactions in bacteria, but K_D values have yet to emerge. Our group has shown that the concentration of the side-by-side GB1 homodimer can be controlled and quantified in *Escherichia coli* cells [47,48].

In summary, most in-cell efforts exploit fluorescence as the detection method and most endeavors report either no effect or stabilization.

Measuring crowding effects

Although fluorescence-based methods can assess binding at or near physiologically relevant protein concentrations, detection often relies on large labels that might interfere with complex formation. This problem can be offset by comparing the results from cells to those obtained for the same constructs in dilute buffer. NMR-based detection involves less perturbation (i.e. isotopic enrichment or small labels like ^{19}F) but it lacks sensitivity and therefore, nonphysiologically large quantities of protein are required. The EPR-based DEER technique combines high sensitivity and rather small perturbations (i.e. spin labels), but detection requires nonphysiologically relevant low temperatures.

Bear in mind that K_D values from in-cell studies are often only 'apparent' because of competition between the labeled protein, which is required to make the measurements, and the natural version of the protein. A related challenge involves knowledge of the cell volume because the concentration of reactants, which is required to obtain

K_D (Figure 2), depends on knowing volume. Such values are often not measured directly but taken from the literature. Even when measured directly, however, the volume available to the complex is usually unknown (i.e. Should the volume occupied by organelles, ribosomes, etc. be subtracted?).

Models for interpreting binding data under crowded conditions

> "All models are wrong but some are useful" George E. P. Box [52]

Now, we turn to the molecular-level interpretations of how high concentrations of cosolutes from sugars to the complex mix of intracellular macromolecules affect protein complex stability. A more nuanced and complete discussion of many of these ideas is found in our contribution to *Annual Reviews of Biophysics and Biomolecular Structure* [53].

Preferential interactions

We begin with an analysis that is mostly applied to high concentrations of small molecule cosolutes. On a mass-per-molecule basis, protein complexes are hundreds to thousands of times heavier than the small cosolutes such as glycerol, glucose, betaine, and so on, listed in Table 1. Therefore, interactions between small molecules and proteins can reasonably be treated in terms of the free energy per unit area of cosolute functional groups. This idea is the basis of chemical-potential derivatives, as described by Timasheff [54]. The concept is referred to as preferential interaction if attractive, and preferential hydration repulsive. This model, as elegantly pursued by the Record group [43,55,56], works remarkably well for protein stability. Silvers and Myers [16] analyze their concanavilin A dimer-to-tetramer data in terms of the Tanford transfer free energy model [57] and the more sophisticated Record model [43]. The Record model fits better [44]. The difference in the quantity and quality of area buried in protein complex formation and upon folding may explain why sugars and osmolytes, which stabilize proteins can destabilize a protein complex [15].

The goal of several efforts in Table 1 is counting the water molecules absorbed or expelled on complex formation [12,13,15] using osmotic stress analysis [58]. It should be borne in mind that such interpretations are controversial and that using different techniques can lead to conflicting results [59]. Finally, Rydeen *et al.* [38•] suggest that naturally occurring osmolytes are differentiated from other cosolutes not by their stabilizing influences on protein tertiary structure but by their compatibility with the interactions between protein surfaces in complexes in cells.

Simple excluded volume theory [24,42]

Many of us first heard of excluded volume as a correction to the ideal gas law. Simply put, the analysis focuses on the fact that two molecules cannot access the same volume at the same time. As applied to crowded solutions, the complex, its constituent proteins and the crowding molecules, are all treated as hard (i.e. inert) spheres, and solvent water is treated as a featureless background. The volume excluded by cosolute equals the volume of a sphere with a radius that is the sum of radii of the protein and cosolute. $\Delta\Delta G_D^{\varphi'}$ is proportional to $-RT$ times the natural log of the difference in excluded volume of the free proteins and the complex. Crowding will stabilize a complex as long as the complex takes up less space than the sum of the individual proteins.

A requirement of hard particles is that they don't sense each other until they touch and, then, being impenetrable, there is a pure and infinite repulsion between them[5] . This requirement means pure hard particle effects are purely entropic. That is, they only depend on the arrangement of the particles. Simple excluded volume analysis is firmly grounded in classic solution theory [60,61] and can work well for small cosolutes [24]. There is also a clear prediction for simple excluded volume analyses: the effect decreases with increasing crowder size at a fixed crowder volume occupancy [53].

Scaled-particle theory [45,62]

This is another excluded volume-based analysis, but the shape of the components can be manipulated, and water can be treated specifically. Importantly, the analysis collapses to simple excluded volume analysis if water is ignored. It also leads to the same prediction about crowder size: a decreasing effect with increasing size at a fixed g/L concentration. A shortcoming is that scaled particle theory is not neatly tied to classic solution theory because water is treated explicitly rather than as a featureless background [53].

Depletion forces [41••,63]

This analysis also focuses on excluded volume and therefore hard interactions. The idea is based on osmotic pressure. When the proteins comprising the complex approach each other at a distance such that the crowder cannot fit between them, the crowder concentration between the proteins is less than the crowder concentration in the bulk solution. The difference in pressure

[5] The idea that excluded volume arises solely from hard-core repulsion is not shared by polymer chemists who accept the existence of negative excluded volume as a way to incorporate attractive chemical interactions between molecules. See Rubinstein M, Colby RH: *Polymer physics*: Oxford University Press; 2003.

draws the protein into the complex. Importantly, its pioneers, Asakura and Oosawa [63], not only considered hard particles but also synthetic polymers. Zosel *et al.* [41••] have applied this analysis to their data on protein complex formation in PEGs and showed that the expectations are opposite to those for simple excluded volume analysis and scaled particle theory. Namely, that the strength of the macromoleclar effect increases with polymer size at a fixed g/L concentration.

Potential of mean force [64]

The models discussed in the last three sections assume that the effect of high cosolute concentrations arises only from hard-core excluded volume. That is, crowding is purely entropic; there is no enthalpic effect. Sapir and Harries [64] have developed an elegant formulation that incorporates chemical interactions into depletion force analysis. We look forward to tests of this novel approach. Their ideas may explain the temperature dependence of crowding, which reveals enthalpic effects, meaning that cosolutes are attracted or repelled from protein surfaces [10]. These are the so-called soft interactions. Attractive soft interactions include hydrogen bonds, complementary electrostatic interactions, and so on. Repulsive soft interactions comprise those between like charges (i.e. they add to the hard-core interactions).

In summary, there are a plethora of analyses for interpreting crowding effects. The various explanations run the gamut from being based on chemical interactions alone, to hard-core repulsions alone, to combinations of the two. The various ideas also treat the macromolecular crowder in different ways: from spheres in a featureless sea of solvent, to spherical-based shapes in a spherical solvent, to treating crowders as polymeric chains. In the final section we turn to ways the application of these analyses can be used, abused, and perhaps improved.

> "When the going gets weird, the weird turn pro."
> Hunter S. Thompson [65]

Most importantly, we must not treat cosolutes as inert species and recognize that cosolutes can interact in an attractive or repulsive fashion with proteins. One need look no further than the work from Thomas Record's laboratory [44] for proof. Next, let's stop treating molecules of synthetic polymers in high concentration solutions as if they are in any sense spheres. High concentrations of hard spheres jam, whereas high concentrations of synthetic polymers form highly viscous semidilute solutions. An attractive feature of synthetic polymers especially PEGs and to a lesser extent dextrans, is that they are available in a range of molecular weights. Using this range will help reveal valid molecule level interpretations. Nevertheless, we must remember that cells are crowded with globular proteins (and nucleic acids and metabolites) [4], not synthetic polymers. Using synthetic polymers is key to understanding crowding and for preserving protein drugs, but the results may not be relevant to understanding biology.

The most important biologically relevant crowders are globular proteins. Unlike synthetic polymers, however, there is no protein family where increasing molecular weight leads to a smooth increase in size let alone smooth increase in surface properties. It might be possible to take a particularly stable and soluble protein (e.g. GB1) and alter the surface charge in a systematic manner and in this way assess the effect of surface. Such an effort is, however, challenging because the proteins would have to be produced in large quantities (i.e. it is much easier to buy lots of BSA and lysozyme).

As shown by scanning Table 1 from top to bottom, the rate of data accumulation on the equilibrium thermodynamics and kinetic effects of crowded conditions on protein complex formation under crowded conditions is increasing, but more systematic efforts, such as those from the Schuler lab [41••] are desperately required. Only sustained methodical efforts *in vitro*, in living cells and whole animals [66•] will provide the knowledge required to quantify, understand, explain, and exploit the effects of crowding on macromolecular interactions.

Gazing further, we anticipate that combining strong, well-designed wet-experiment based studies in cells with the theories discussed above and simulations of the cellular interior [67,68,69•] will result in the ability to predict the effect of the intracellular environmental on protein–protein interactions. The success of such syntheses will facilitate the modeling of metabolism and ultimately lead to cures for those diseases caused by missense mutations that perturb protein complexes [2].

Conflict of interest statement

Nothing declared.

Acknowledgements

Our research on macromolecular crowding is supported by the National Science Foundation (MCB-1410854) and the United States-Israel Binational Science Foundation (BSF 2017063). We thank Ben Schuler for Figure 3 and Elizabeth Pielak, Shannon Speer and Jan Spitzer for helpful comments.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Ellis RJ: **Macromolecular crowding: an important but neglected aspect of the intracellular environment**. *Curr Opin Struct Biol* 2001, **11**:114-119.

2. Sahni N, Yi S, Taipale M, Fuxman Bass Juan I, Coulombe-Huntington J, Yang F, Peng J, Weile J, Karras Georgios I, Wang Y *et al.*: **Widespread macromolecular interaction perturbations in human genetic disorders**. *Cell* 2015, **161**:647-660.

3. Lewis GN, Randall M: *Thermodynamics*. New York: McGraw-Hill; 1923.

4. Theillet F-X, Binolfi A, Frembgen-Kesner T, Hingorani K, Sarkar M, Kyne C, Li C, Crowley P, Gierasch L, Pielak GJ *et al.*: **Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs)**. *Chem Rev* 2014, **114**:6661-6714.

5. Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN: **Living with water stress: evolution of osmolyte systems**. *Science* 1982, **217**:1214-1222.

6. Piszkiewicz S, Pielak G: **Protecting enzymes from stress-induced inactivation**. *Biochemistry* 2019, **58**:3825-3833.

7. Stadmiller SS, Aguilar JS, Waudby C, Pielak GJ: **Rapid quantification of protein-ligand binding via ^{19}F NMR lineshape analysis**. *Biophys J* 2020, **118**:2337-2348.

8. Creighton TE: **An empirical approach to protein conformation stability and flexibility**. *Biopolymers* 1983, **22**:49-58.

9. Ferguson MWJ, Joanen T: **Temperature of egg incubation determines sex in *Alligator mississippiensis***. *Nature* 1982, **296**:850-853.

10. Guseman AJ, Pielak GJ: **Protein stability and weak intracellular interactions**. In *In-Cell NMR Spectroscopy: From Molecular Sciences to Cell Biology.* Edited by Ito Y, Dötsch V, Shirakawa M. The Royal Society of Chemistry; 2020:188-206.

11. Wilf J, Minton AP: **Evidence for protein self-association induced by excluded volume myoglobin in the presence of globular proteins**. *Biochim Biophys Acta* 1981, **670**:316-322.

12. Kornblatt JA, Kornblatt MJ, Hoa GH, Mauk AG: **Responses of two protein-protein complexes to solvent stress: does water play a role at the interface?** *Biophys J* 1993, **65**:1059-1065.

13. Sato W, Hitaoka S, Uchida T, Shinzawa-Itoh K, Yoshizawa K, Yoshikawa S, Ishimori K: **Osmotic pressure effects identify dehydration upon cytochrome *c* - cytochrome *c* oxidase complex formation contributing to a specific electron pathway formation**. *Biochem J* 2020.

14. Morar AS, Wang X, Pielak GJ: **Effects of crowding by mono-, di-, and tetrasaccharides on cytochrome *c*-cytochrome *c* peroxidase binding: comparing experiment to theory**. *Biochemistry* 2001, **40**:281-285.

15. Xavier KA, Shick KA, Smith-Gil SJ, Willson RC: **Involvement of water molecules in the association of monoclonal antibody HyHEL-5 with bobwhite quail lysozyme**. *Biophys J* 1997, **73**:2116-2125.

16. Silvers TR, Myers JK: **Osmolyte effects on the self-association of Concanavalin A: testing theoretical models**. *Biochemistry* 2013, **52**:9367-9374.

17. Ladurner AG, Fersht AR: **Upper limit of the time scale for diffusion and chain collapse in chymotrypsin inhibitor 2**. *Nat Struct Biol* 1999, **6**:28-31.

18. Zorrilla S, Rivas G, Acuña AU, Lillo MP: **Protein self-association in crowded protein solutions: A time-resolved fluorescence polarization study**. *Protein Sci* 2004, **13**:2960-2969.

19. Kozer N, Schreiber G: **Effect of crowding on protein-protein association rates: Fundamental differences between low and high mass crowding agents**. *J Mol Biol* 2004, **336**:763-774.

20. Kozer N, Kuttner YY, Haran G, Schreiber G: **Protein-protein association in polymer solutions: From dilute to semidilute to concentrated**. *Biophys J* 2007, **92**:2139-2149.

21. Schlarb-Ridley BG, Mi HL, Teale WD, Meyer VS, Howe CJ, Bendall DS: **Implications of the effects of viscosity, macromolecular crowding, and temperature for the transient interaction between cytochrome *f* and plastocyanin from the cyanobacterium *Phormidium laminosum***. *Biochemistry* 2005, **44**:6232-6238.

22. Zhou YL, Liao JM, Chen J, Liang Y: **Macromolecular crowding enhances the binding of superoxide dismutase to xanthine oxidase: implications for protein-protein interactions in intracellular environments**. *Int J Biochem Cell Biol* 2006, **38**:1986-1994.

23. You X, Nguyen AW, Jabaiah A, Sheff MA, Thorn KS, Daugherty PS: **Intracellular protein interaction mapping with FRET hybrids**. *Proc Natl Acad Sci U S A* 2006, **103**:18458-18463.

24. Winzor DJ, Patel CN, Pielak GJ: **Reconsideration of sedimentation equilibrium distributions reflecting the effects of small inert cosolutes on the dimerization of α-chymotrypsin**. *Biophys Chem* 2007, **130**:89-92.

25. Maeder CI, Hink MA, Kinkhabwala A, Mayr R, Bastiaens PIH, Knop M: **Spatial regulation of fus3 map kinase activity through a reaction-diffusion mechanism in yeast pheromone signalling**. *Nat Cell Biol* 2007, **9**:1319-1326.

26. Sudhaharan T, Liu P, Foo YH, Bu W, Lim KB, Wohland T, Ahmed S: **Determination of *in vivo* dissociation constant, K_D, of cdc42-effector complexes in live mammalian cells using single wavelength fluorescence cross-correlation spectroscopy**. *J Biol Chem* 2009, **284**:13602-13609.

27. Shi X, Foo YH, Sudhaharan T, Chong S-W, Korzh V, Ahmed S, Wohland T: **Determination of dissociation constants in living zebrafish embryos with single wavelength fluorescence cross-correlation spectroscopy**. *Biophys J* 2009, **97**:678-686.

28. Phillip Y, Sherman E, Haran G, Schreiber G: **Common crowding agents have only a small effect on protein-protein interactions**. *Biophys J* 2009, **97**:875-885.

29. Batra J, Xu K, Qin S, Zhou H-X: **Effect of macromolecular crowding on protein binding stability: Modest stabilization and significant biological consequences**. *Biophys J* 2009, **97**:906-911.

30. Jiao M, Li HT, Chen J, Minton AP, Liang Y: **Attractive protein-polymer interactions markedly alter the effect of macromolecular crowding on protein association equilibria**. *Biophys J* 2010, **99**:914-923.

31. Okamoto DN, Oliveira LC, Kondo MY, Cezari MH, Szeltner Z, Juhasz T, Juliano MA, Polgar L, Juliano L, Gouvea IE: **Increase of SARS-CoV 3CL peptidase activity due to macromolecular crowding effects in the milieu composition**. *Biol Chem* 2010, **391**:1461-1468.

32. Phillip Y, Kiss V, Schreiber G: **Protein-binding dynamics imaged in a living cell**. *Proc Natl Acad Sci U S A* 2012, **109**:1461-1466.

33. Phillip Y, Harel M, Khait R, Qin SB, Zhou HX, Schreiber G: **Contrasting factors on the kinetic path to protein complex formation diminish the effects of crowding agents**. *Biophys J* 2012, **103**:1011-1019.

34. Hoffman L, Wang X, Sanabria H, Cheung MS, Putkey JA, Waxham MN: **Relative cosolute size influences the kinetics of protein-protein interactions**. *Biophys J* 2015, **109**:510-520.

35. Guseman AJ, Pielak GJ: **Cosolute and crowding effects on a side-by-side protein dimer**. *Biochemistry* 2017, **56**:971-976.

36. Sukenik S, Ren P, Gruebele M: **Weak protein-protein interactions in live cells are quantified by cell-volume modulation**. *Proc Natl Acad Sci U S A* 2017, **114**:6776-6781.

37. Guseman AJ, Speer SL, Perez Goncalves GM, Pielak GJ: **Surface charge modulates protein-protein interactions in physiologically relevant environments**. *Biochemistry* 2018, **57**:1681-1684.

38. • Rydeen AE, Brustad EM, Pielak GJ: **Osmolytes and protein−protein interactions**. *J Am Chem Soc* 2018, **140**:7441-7444.

Osmolytes are differentiated from other cosolutes by their compatibility with interactions between protein surfaces in cells, not by their effect on protein and protein complex stability.

39. • Guseman AJ, Perez Goncalves GM, Speer SL, Young GB, Pielak GJ: **Protein shape modulates crowding effects**. *Proc Natl Acad Sci U S A* 2018, **115**:10965-10970.

Crowding stabilizes more compact- better than less- compact complexes, but soft interactions can be even more important.

40. Palacios A, Blanco FJ: **Macromolecular crowding increases the affinity of the PHD of ING4 for the histone H3K4me3 mark**. *Biomolecules* 2020, **10**:234.

41. •• Zosel F, Soranno A, Nettels D, Schuler B: **Depletion interactions modulate coupled folding and binding in crowded environments**. *Proc Natl Acad Sci U S A* 2020, **117**:13480-13489.
A well designed and executed systematic study showing the effect of crowder size and concentration on the stability and diffusion of a protein complex. The systematic nature of the effort allowed the first detailed comparison of expectations from scaled particle theory and depletion force analysis.

42. Yang Y, Chen S-N, Yang F, Li X-Y, Feintuch A, Su X-C, Goldfarb D: **In-cell destabilization of a homo-dimeric protein complex detected by DEER spectroscopy**. *Proc Natl Acad Sci U S A* 2020, **117**:20566-20575.

43. Guinn EJ, Pegram LM, Capp MW, Pollock MN, Record MT: **Quantifying why urea is a protein denaturant, whereas glycine betaine is a protein stabilizer**. *Proc Natl Acad Sci U S A* 2011, **108**:16932-16937.

44. Knowles DB, Shkel IA, Phan NM, Sternke M, Lingeman E, Cheng X, Cheng L, O'Connor K, Record MT: **Chemical interactions of polyethylene glycols (PEGs) and glycerol with protein functional groups: applications to effects of PEG and glycerol on protein processes**. *Biochemistry* 2015, **54**:3528-3542.

45. Graziano G: **Shape effect on non-covalent dimer stability using classic scaled particle theory**. *Chem Phys Lett* 2020, **743** 137176.

46. Rubinstein M, Colby RH: *Polymer physics*. Oxford University Press; 2003.

47. Speer SL, Guseman AJ, Patteson JB, Ehrmann BM, Pielak GJ: **Controlling and quantifying protein concentration in *Escherichia coli***. *Protein Sci* 2019, **28**:1307-1311.

48. Chu IT, Speer SL, Pielak GJ: **Rheostatic control of protein expression using tuner cells**. *Biochemistry* 2020, **59**:733-735.

49. Kiraga J, Mackiewicz P, Mackiewicz D, Kowalczuk M, Biecek P, Polak N, Smolarczyk K, Dudek MR, Cebrat S: **The relationships between the isoelectric point and: Length of proteins, taxonomy and ecology of organisms**. *BMC Genomics* 2007, **8**:163.

50. Cohen RD, Pielak GJ: **A cell is more than the sum of its (dilute) parts: a brief history of quinary structure**. *Protein Sci* 2017, **26**:403-413.

51. Srere PA: **Macromolecular interactions: tracing the roots**. *Trends Biochem Sci* 2000, **25**:150-153.

52. Box GEP: **Robustness in the strategy of scientific model building**. In *Robustness in statistics.* Edited by Launer RL, Wilkinson GN. Academic; 1979:201-236.

53. Davis-Searles PR, Saunders AJ, Erie DA, Winzor DJ, Pielak GJ: **Interpreting the effects of small uncharged solutes on protein-folding equilibria**. *Annu Rev Biophys Biomol Struct* 2001, **30**:271-306.

54. Timasheff SN: **The control of protein stability and association by weak interactions with water: how do solvents control these processes?** *Annu Rev Biophys Biomol Struct* 1993, **22**:67-97.

55. Cheng X, Shkel IA, Connor KO, Henrich J, Molzahn C, Lambert D, Record MT: **Experimental atom-by-atom dissection of amide-amide and amide-hydrocarbon interactions in H_2O**. *J Am Chem Soc* 2017, **139**:9885-9894.

56. Cheng X, Shkel IA, O'Connor K, Record MT: **Experimentally-determined strengths of favorable and unfavorable interactions of amide atoms involved in protein self-assembly in water**. *Proc Natl Acad Sci U S A* 2020, **117**:27339-27345.

57. Tanford C: **Protein denaturation: Part C. Theoretical models for the mechanism of denaturation**. *Adv Protein Chem* 1970, **24**:1-95.

58. Parsegian VA, Rand RP, Rau DC: **Osmotic stress, preferential hydration, and binding: a comparison of perspectives**. *Proc Natl Acad Sci U S A* 2000, **97**:3987-3992.

59. Kornblatt JA, Kornblatt MJ, Rajotte I, Hoa GHB, Kahn PC: **Thermodynamic volume cycles for electron transfer in the cytochrome *c* oxidase and for the binding of cytochrome *c* to cytochrome *c* oxidase**. *Biophys J* 1998, **75**:435-444.

60. McMillan WG, Mayer JE: **The statistical thermodynamics of multicomponent systems**. *J Chem Phys* 1945, **13**:276-305.

61. Kirkwood JG, Buff FP: **The statistical mechanical theory of solutions. I**. *J Chem Phys* 1951, **19**:774-777.

62. Berg OG: **The influence of macromolecular crowding on thermodynamic activity: Solubility and dimerization constants for spherical and dumbbell-shaped molecules in a hard-sphere mixture**. *Biopolymers* 1990, **30**:1027-1037.

63. Asakura S, Oosawa F: **On interaction between two bodies immersed in a solution of macromolecules**. *J Chem Phys* 1954, **22**:1255-1256.

64. Sapir L, Harries D: **Macromolecular stabilization by excluded cosolutes: Mean field theory of crowded solutions**. *J Chem Theory Comput* 2015, **11**:3478-3490.

65. Thompson HS, Steadman R: *Fear and Loathing in Las Vegas: A Savage Journey to the Heart of the American Dream*. Random House; 1971.

66. • Feng R, Gruebele M, Davis CM: **Quantifying protein dynamics and stability in a living organism**. *Nat Commun* 2019, **10**:1179.
A peak at the future where protein complexes are quantitatively studied in whole organisms.

67. Ostrowska N, Feig M, Trylska J: **Modeling crowded environment in molecular simulations**. *Front Mol Biosci* 2019, **6**:B6.

68. Rickard MM, Zhang Y, Gruebele M, Pogorelov TV: **In-cell protein–protein contacts: Transient interactions in the crowd**. *J Phys Chem Lett* 2019, **10**:5667-5673.

69. • von Bülow S, Siggel M, Linke M, Hummer G: **Dynamic cluster formation determines viscosity and diffusion in dense protein solutions**. *Proc Natl Acad Sci U S A* 2019, **116**:9843.
An important effort to explain why rotational diffusion is slowed more than translational diffusion under physiologically relevant crowded conditions.