THE JOURNAL OF PHYSICAL CHEMISTRY B

Comment

pubs.acs.org/JPCB

Reply to "Comment on 'Hydration and Mobility of Trehalose in Aqueous Solution'"

Bertil Halle*

Department of Biophysical Chemistry, Lund University, SE-22100 Lund, Sweden

J. Phys. Chem. B **2012**, *116* (30), 9196−9207. DOI: 10.1021/jp304982c
J. Phys. Chem. B **2014**, *118*. DOI: 10.1021/jp508089t

Supporting Information

We recently presented a nuclear spin relaxation study of solvent and solute rotational dynamics in aqueous trehalose solution.[1] The results of this and other subsequent[2] studies are consistent with a solute-induced perturbation of water dynamics that is essentially confined to the first hydration shell. In contrast, THz absorption measurements by the Havenith group (HG) were taken to imply a "long-range" perturbation of solvent dynamics.[3] This discrepancy led us to critically examine the model used to interpret the THz data, with the conclusion that HG's THz approach does not yield a reliable estimate of the perturbation range. In their Comment,[4] HG reaffirm their analysis without convincingly addressing our principal objections. In this Reply, we attempt to clarify these issues by first explicitly stating the physical approximations in HG's model and then quantitatively demonstrating how and why the data analysis produces spurious results. In doing so we add new insights to this debate, including a vastly more efficient way of computing the overlap volume. To identify the physical basis of HG's puzzling THz results, it is helpful to contrast saccharides[3,5] with proteins.[6−10] Due to space limitations, we consider only two solutes: the disaccharide trehalose and the small globular protein ubiquitin. For proteins, the discrepancy with nuclear spin relaxation results[11,12] is even more striking, with the THz analysis suggesting significant perturbations of water dynamics out to ∼20 Å and more from the protein surface.[6,8−10]

According to the three-component model used by HG, the THz absorption coefficient α is a volume-based average of bulk water (W), hydration water (H), and solute (S) contributions

$$\alpha = \phi_W \alpha_W + \phi_H \alpha_H + \phi_S \alpha_S \tag{1}$$

The straight lines in Figure 1 of HG's Comment[4] correspond to the limiting cases of eq 1 with no hydration water ($\phi_H = 0$) or with no bulk water ($\phi_W = 0$). Equation 1 may appear innocuous, but a collective property like α cannot be neatly decomposed into additive and independent component contributions. In particular, the physical significance of α_H and α_S is not immediately obvious (see below).

The reported experiments measure α, typically averaged in the 2.1−2.8 THz window to reduce data scatter, as a function of solute volume fraction ϕ_S (or concentration $C = \phi_S/V_S$). Formally, this dependence can be expanded in a power series

$$\frac{\alpha}{\alpha_W} = 1 + B^{(1)}\phi_S + B^{(2)}\phi_S^2 + O(\phi_S^3) \tag{2}$$

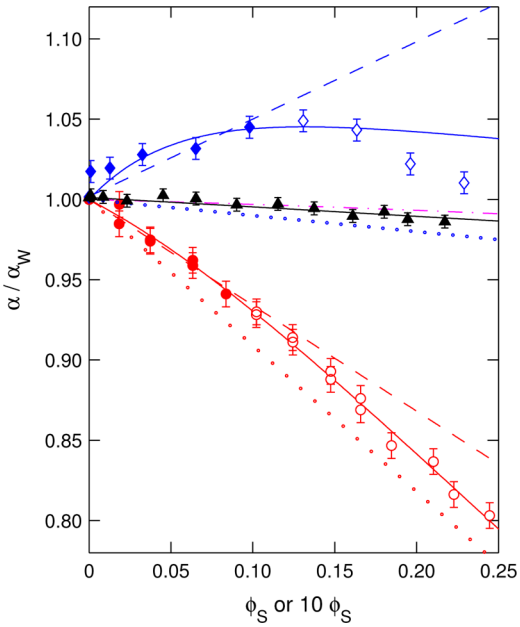

Figure 1. Relative THz absorption α/α_W versus volume fraction ϕ_S trehalose (red circles), ubiquitin at pH 4.8 (blue diamonds), and ubiquitin at pH 2 (black triangles). The data and error bars are from HG's work.[3,8] The red and blue solid curves are fits to the RP3C model (eqs 3−5) with the two highest concentrations omitted for ubiquitin at pH 4.8 (as in the original fit[8]). The solid (black) and dashed (red and blue) straight lines are fits to the linear model (eq 2 truncated after the linear term) using the data points with filled symbols. The dotted lines show the contribution from bulk water and solute (eq 1 with $\alpha_H = 0$ and α_S from the nonlinear fit). The magenta dash-dotted curve is the model prediction for ubiquitin at pH 4.8 with $\xi_H = 1.063$ and $h = 18$ Å (see text). For ubiquitin, the volume fraction on the abscissa has been multiplied by a factor of 10.

where the coefficients $B^{(n)}$ are independent of ϕ_S. All coefficients, except the first one, are model-dependent; in particular, they depend on the solute−solute interaction. The linear, model-independent coefficient has contributions from solute and hydration water

Received: August 18, 2014
Published: August 29, 2014

$$B^{(1)} = B_S^{(1)} + B_H^{(1)} = \left(\frac{\alpha_S}{\alpha_W} - 1\right) + \frac{V_W}{V_S}\nu_H(\xi_H - 1) \quad (3)$$

Here ν_H is the number of water molecules in the hydration domain (of arbitrary shape), and $\xi_H = \alpha_H/\alpha_W$ is the dynamic perturbation factor; i.e., $\xi_H - 1$ is the relative change in THz absorption between hydration and bulk water. Finally, V_S and V_W are the molar volumes of solute and of bulk (and hydration) water. At 20 °C, V_S/V_W = 11.46 for trehalose[13] and 355 for ubiquitin.

Two-state models of solvation dynamics, like the one considered here, are not new. For example, eqs 1−3 (with α replaced by the spin relaxation rate) have been used in innumerable NMR relaxation studies of hydration dynamics during the past half century.[14,15] Since NMR probes single-molecule (rather than collective) dynamics, eq 1 is rigorous (if volume fractions are replaced by mole fractions), and since only water is observed, there is no direct solute contribution. In NMR studies, one usually refrains from interpreting the model-dependent nonlinear concentration dependence. Instead, one determines the model-independent coefficient $B_H^{(1)}$ from the limiting slope (e.g., Figure 7 in ref 1). It is then not possible to separate the magnitude (ξ_H) from the range (ν_H) of the dynamic perturbation. It might be tempting to separate these factors by identifying the onset of nonlinearity with the concentration where hydration domains begin to overlap. For two good reasons, NMR spectroscopists tend to resist this temptation. First, to convert a solute concentration into a solute−solute separation one must know the spatial distribution of solutes. In particular, solute clustering can give rise to hydration domain overlap at solute concentrations much lower than would be the case if the solute−solute interaction only included hard-core repulsion. Second, in addition to this interaction-dependent geometric overlap effect, there is a dynamic overlap effect produced by the stronger dynamical perturbation of water confined between two or more solute molecules. Whereas the geometric overlap effect produces a weaker than linear concentration dependence, the dynamic overlap effect usually produces a stronger than linear concentration dependence. Typically, as for trehalose,[1] it is the dynamic overlap effect that is responsible for the observed deviation from linear concentration dependence.

Because $B^{(1)}$ is the limiting slope, all quantities appearing in eq 3 pertain to infinite dilution, with no overlap between hydration domains. As an alternative to the series expansion in eq 2, α can be cast on the pseudolinear form

$$\frac{\alpha}{\alpha_W} = 1 + B^{(1)}(\phi_S)\phi_S \quad (4)$$

$B^{(1)}(\phi_S)$ is still given by eq 3, where, however, $\alpha_S(\phi_S)$, $\nu_H(\phi_S)$, and $\xi_H(\phi_S)$ are now functions of the solute concentration ϕ_S. (We ignore the weak concentration dependence in the partial molar volume V_S.) The model used by HG, which we shall refer to as the restricted primitive three-component (RP3C) model, can be obtained from eq 4 by (1) neglecting dynamic cross correlations between different molecular components, (2) neglecting the spatial decay of the dynamic solvent perturbation, and (3) neglecting all solution nonideality beyond excluded volume. Failure of any of these approximations can seriously compromise the interpretation of THz absorption data. We now briefly discuss each of these three approximations.

The THz absorption coefficient α is a collective property that mainly reflects intermolecular vibrations. The component absorption coefficients α_H and α_S must therefore be regarded as effective quantities that contain contributions from dynamic solute−water cross correlations.[16] In general, the parameter ξ_H therefore cannot be regarded as a measure of intrinsic water dynamics. For the same reason, cross correlations complicate the extraction of hydration water dynamics from other collective observables, such as the dielectric permittivity[17,18] or the fluorescence emission frequency from a newly created chromophore dipole.[19,20] In addition, cross correlations between distinct solute molecules render α_S concentration dependent. In the RP3C model, these complications are ignored by treating α_H and α_S as fixed parameters, independent of ϕ_S. The effect of cross-correlations is likely to be exacerbated by attractive solute−solute interactions leading to solute clustering, a phenomenon which is not captured by single-solute simulations.[4]

Physical considerations suggest that $\alpha_H(r)$ gradually approaches the bulk value α_W with increasing distance r from the solute, mostly likely in an approximately exponential fashion. For the limiting slope in eq 3, this is not an issue because α_H can be regarded as a spatial average of $\alpha_H(r)$ over the hydration domain. However, to quantify the effect of hydration domain overlap, we must know the functional form of $\alpha_H(r)$. In the RP3C model, this is taken to be a step function. Hydration domain overlap then becomes a purely geometric problem, formalized by writing $\nu_H(\phi_S) = \nu_H[1 - \eta(\phi_S)]$. (Here, and in the following, quantities written without an explicit ϕ_S argument refer to infinite dilution, as in eq 3.) The overlap fraction $\eta(\phi_S)$ is the average fractional reduction of the hydration domain volume due to overlap. Note that $\eta(\phi_S)$ accounts for two kinds of overlap: (1) shared hydration, where water molecules simultaneously belong to the hydration domains of two or more solute molecules and (2) excluded hydration, where a solute molecule penetrates the hydration domain of one or more neighboring solute molecules.

The total solute−solute interaction in an aqueous solution results from several physical mechanisms, including (shape-dependent) hard-core repulsion (excluded volume), hydrogen bonds, and solvent-mediated Coulomb and hydrophobic interactions. For the purpose of computing the overlap fraction $\eta(\phi_S)$, the RP3C model ignores this complexity by treating the aqueous solution as a hard-sphere fluid. HG use a Monte Carlo procedure to determine the overlap fraction $\eta(\phi_S)$ "by randomly inserting solute molecules into water".[5] But no water and no interactions beyond the solute's hard core are involved. HG model the solute as a hard sphere or, for disaccharides, as two fused hard spheres, but $\eta(\phi_S)$ is practically the same for these geometries at the same core volume.[3,5]

The Monte Carlo procedure is inconvenient for data fitting since it must be repeated for each solute concentration (and each solute type). This hurdle can be avoided by recognizing that the bulk water volume fraction ϕ_W can be equated with the probability of successfully inserting, at an arbitrary point in a fluid of hard spheres of diameter σ and volume fraction ϕ_S, a test particle of radius h (where h is the thickness of the spherical hydration shell). A variety of simple analytical results, based on approximate equations of state, are available for this probability,[21,22] all of which are virtually exact at the solute volume fractions of interest here.[23] We adopt an approximation based on the Carnahan−Starling equation of state, which yields[22]

dx.doi.org/10.1021/jp5083467 | J. Phys. Chem. B 2014, 118, 10806−10812

Table 1. Results of RP3C Model Fits to the THz Absorption Data in Figure 1[a]

parameter	unit	trehalose	ubiquitin pH 4.8	ubiquitin pH 2
$100(\xi_H - 1)^b$	%	3.0 ± 2.2 (3.4)	6.3 ± 0.8 (2.4)	–
$\nu_H{}^b$	–	144 ± 126 (893)	[7.0 ± 2.2 (5.5)] × 10^4	–
h^d	Å	6.0 ± 2.8 (19.9)	66 ± 9 (20)	–
$\alpha_S{}^b$	cm^{-1}	38 ± 31 (39)	[0][e]	–
$\alpha_W{}^c$	cm^{-1}	420.0 ± 0.3 (0.6)	420.8 ± 0.3 (0.6)	420.5 ± 0.2 (0.3)
$B^{(1)c}$	–	−0.66 ± 0.06 (0.12)	4.8 ± 0.5 (1.1)	−0.6 ± 0.1 (0.2)

[a]Fits were performed with the trust region reflective nonlinear optimization algorithm.[24] Parameter errors corresponding to a confidence interval of 68.3% (95%) were obtained with the Monte Carlo method[25] using 5000 synthetic data sets. [b]From nonlinear fit. [c]From linear fit. [d]Spherical hydration shell thickness h derived from fitted "hydration number" ν_H according to $\nu_H = (\pi h/v_W)[\sigma(\sigma + 2h) + 4h^2/3]$, where $v_W = 29.97$ Å^3 is the molecular volume of bulk (and hydration) water at 20 °C and σ is the hard-sphere diameter of the solute (8.7 Å for trehalose and 27.3 Å for ubiquitin). [e]Parameter value frozen in the fit.

$$\eta(\phi_S) = 1 - \frac{(1 - \phi_S)[1 - \exp(-\phi_S A)]}{\phi_S(\kappa^3 - 1)} \quad (5)$$

where $\kappa = 1 + 2h/\sigma$ and $A = a\kappa^3 + 3b\kappa^2 + 12c\kappa + d$ with coefficients $a = (1 + \phi_S)/(1 - \phi_S)^3$, $b = -\phi_S(3 + \phi_S)/[2(1 - \phi_S)^3]$, $c = \phi_S^2/[2(1 - \phi_S)^3]$, and $d = -(2 - 7\phi_S + 9\phi_S^2)[2(1 - \phi_S)^3]$.

Having defined the RP3C model and the underlying approximations, we shall now apply it to the THz absorption data (averaged over the interval 2.1−2.8 THz) measured by HG on aqueous solutions of trehalose[3] and ubiquitin.[8] The efficient computation of the overlap fraction by means of eq 5 allows us to rigorously determine parameter confidence intervals and covariances by nonlinear fits to 5000 synthetic data sets.[24,25] For the fits we use the experimental errors reported by HG,[3,8] except at $\phi_S = 0$ where the error was reduced by a factor of 10 since α_W was determined with higher precision.[4] The data and fits (solid curves) are shown in Figure 1, while the resulting values of the three RP3C model parameters (α_S, ν_H, and ξ_H) are collected in Table 1.

We first discuss the results for trehalose. As seen from Table 1, the parameter errors are large even at the 68% confidence level. For the other two saccharides examined by HG,[3,5] we find that the parameters are even less well-determined. For trehalose, the 68% confidence region for the "hydration number" ν_H ranges from less than a monolayer (the first hydration shell comprises ∼47 water molecules[1]) to about three water shells. All three parameters have large mutual covariances (Figures S1−S3, Supporting Information). Even if α_S is frozen in the fit, the remaining two parameters have a large covariance with a trade-off between the magnitude (ξ_H) and range (ν_H) of the dynamic perturbation. Because of data scatter, the linear regime is not well-defined. Using data at $\phi_S < 0.1$, we obtain a limiting slope of $B^{(1)} = -0.66 \pm 0.06$, but to isolate the hydration contribution $B_H^{(1)}$, α_S must be known.

When HG fitted the RP3C model to their trehalose data[3] they fixed the solute absorption coefficient to $\alpha_S = 53$ cm^{-1}. This value was reportedly[4] obtained from measurements on a "dry powder of trehalose", which presumably means the common crystal form trehalose dihydrate.[26] Similar measurements on glucose powders reveal that the absorption coefficient (averaged over the 2.1−2.8 THz interval) differs substantially between anhydrous glucose (15.7 cm^{-1}) and glucose monohydrate (9.0 cm^{-1}).[27] Even if we disregard phonon modes, which only contribute in the crystal, it is not clear what α_S value should be used for a solute surrounded by (and interacting strongly with) water molecules. Since α_S is dominated by intermolecular vibrations,[27] the value measured on a powder sample is almost entirely due to cross correlations that will be different for the solute in aqueous solution. This problem cannot be addressed by simulations of a single trehalose molecule in water.[4] For these reasons, it seems advisable to regard α_S as an adjustable effective parameter in the RP3C model, as we have done here.

The RP3C model assumes that the spatial distribution of solute molecules in a concentrated aqueous trehalose solution is the same as in a hard-sphere fluid. The ability of HG's THz approach to determine the range of the solute-induced perturbation of water dynamics (specified with sub-Å precision[3,5]) relies critically on the validity of this bold assumption. While HG believe that "the small deviation of the experimental data points from the best-fit curve supports the three-component model",[3] it is not surprising that the modest deviation from linearity can be captured by an additional parameter. For example, the trehalose data in Figure 1 are also very well described by eq 2 taken to second order. The decisive question here is not how well the model fits the data but what the model parameters mean. In the case of trehalose, there is strong experimental[1] and computational[2,28] evidence for transient solute clustering over the entire concentration range examined by HG.[3] This aggregation not only alters the magnitude and concentration dependence of the overlap fraction $\eta(\phi_S)$ from that computed for the hard-sphere fluid but also introduces a nonlinear ϕ_S dependence via the effect of cross correlations on α_H and α_S.

Our main criticism of HG's analysis of THz absorption data from trehalose and other saccharides[3,5] can be summed up in the following three points.

(1) HG state that the THz absorption coefficient α "*directly* probes solute-induced *retardation* of water dynamics" in aqueous saccharide solutions.[5] However, the link between α and water dynamics can hardly be described as "direct" when it relies heavily on dubious assumptions about solution structure and when the connection of the THz perturbation factor ξ_H to water dynamics is poorly understood, nonmonotonic, and strongly frequency-dependent. In fact, it is not even clear whether the inferred increased THz absorption coefficient ($\alpha_H > \alpha_W$ and thus $\xi_H > 1$) corresponds to slowing down or speeding up of hydration water dynamics.

(2) HG state that the THz absorption coefficient α reveals a "*long-range* influence of carbohydrates on the solvation dynamics of water".[3] However, according to our reanalysis of their trehalose data, the THz "hydration number" ν_H is consistent, even at the 68% confidence level, with monolayer hydration (Table 1). Nonetheless, given that α primarily probes *intermolecular* vibrations, the effect cannot be completely

dx.doi.org/10.1021/jp5083467 | J. Phys. Chem. B 2014, 118, 10806−10812

confined to the first water layer. Furthermore, the perturbation range is not necessarily the same for the collective dynamics probed by α as for the single-molecule rotation probed by NMR. However, in any case, the analysis of the THz data, even if the RP3C model is accepted, does not justify the description of the perturbation as "long-range".

(3) HG do not acknowledge the experimental[1] and computational[2,28] evidence for trehalose clustering in the concentration range of their THz study. As a result of this clustering, the dynamic overlap effect, which is neglected by HG, dominates over the geometric overlap effect.[1] Moreover, trehalose clustering invalidates the hard-sphere model used by HG to infer the range of the dynamic perturbation. HG ignore these complications because "additional parameters are not justified by the fittable features in the experimental data".[4] Another course of action would have been to follow in the footsteps of the many NMR spectroscopists who were content with determining the model-independent low-concentration slope, that is, $B^{(1)}$ in eq 3.

Turning now to the protein ubiquitin, we note, as did HG,[8] that the RP3C model can provide a reasonable fit only if the two data points at $\phi_S > 0.017$ are ignored (Figure 1). In contrast to the trehalose case, the (poorly defined) limiting slope $B^{(1)} = 4.8 \pm 0.5$ is now large and positive, indicating that $B_H^{(1)} \gg |B_S^{(1)}|$. The absorption of the solute is therefore relatively unimportant, and we follow HG in setting $\alpha_S = 0$. (For example, a fit with fixed $\alpha_S = 100$ yields $100(\xi_H - 1) = 5.9$ and $\nu_H = 7.6 \times 10^4$, well within the 68% confidence intervals in Table 1.) According to the fit, the relative solvent perturbation $(\xi_H - 1)$ is twice as large as for trehalose, and the dynamically perturbed hydration domain now comprises some 70 000 water molecules! For a spherical protein, this corresponds to a shell of thickness $h = 66$ Å.

HG also performed this fit (Figure 6 in ref 8), but they did not present the resulting parameter values. Instead they argued that "proteins are rather large molecules compared to disaccharides and at high enough concentration a large fraction of the solvation water lies in the hydration shells around two or more proteins" and that "the three-component model does not provide a completely quantitative description of the hydration layer thickness because it neglects that hydration water near multiple proteins contributes in a nontrivial way to the total THz absorption coefficient".[8] But in the examined concentration ranges, the overlap fraction $\eta(\phi_S)$, as determined with the aid of the RP3C model, is nearly the same for ubiquitin and trehalose (Figure 2).

In dealing with this perceived problem, HG did not attempt to improve the model but instead employed the same RP3C model in a less rigorous way, positing that "the precise measurement of the onset of nonlinearity allows us to estimate the average size of the dynamical hydration shell".[8] The data in Figure 1 do not define the "onset of nonlinearity" very precisely, but HG take the overlap concentration corresponding to this onset to be 1.5 mM or $\phi_S^* = 0.0096$. They then use their hard-sphere Monte Carlo procedure to compute "an average protein−protein distance", which we call D^*, at this concentration. Finally, they deduce the thickness of the spherical hydration shell as $h = L^*/2$, where $L^* = D^* - \sigma$ is the surface-to-surface separation corresponding to the center-to-center separation D^* and σ is the equivalent hard-sphere diameter for ubiquitin. In this way they arrive at $h = 18$ Å.[8] But if this result is correct, the overlap fraction is less than 10% even at the highest concentration (Figure 2), as compared to 75% for

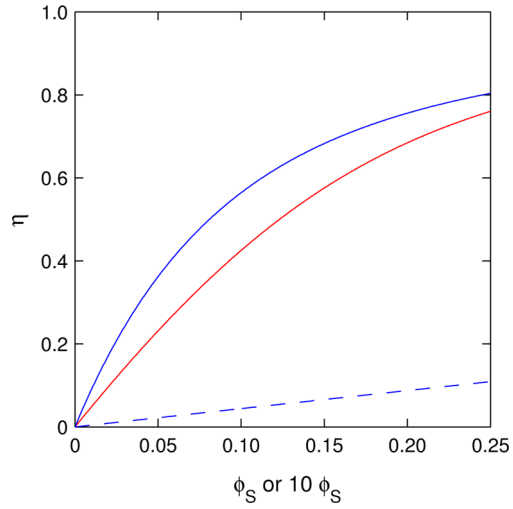

Figure 2. Hydration domain overlap fraction η versus solute volume fraction ϕ_S, computed with eq 5 and parameters for trehalose (red curve, $h/\sigma = 6.0/8.7$) and ubiquitin (blue curve, $h/\sigma = 66.0/27.3$). The dashed curve refers to ubiquitin with a smaller hydration domain ($h = 18$ Å). For ubiquitin, the volume fraction on the abscissa has been multiplied by a factor 10.

trehalose. The inability of the RP3C model to account for the ubiquitin data at the two highest concentrations can therefore not be due to extensive hydration overlap, as proposed by HG.[8]

HG do not define their "average protein−protein distance" D. If taken literally, it would be a macroscopic length. Presumably, they compute the mean *nearest-neighbor* distance, which we shall denote by L. For the hard-sphere fluid, L/σ is a universal function of ϕ_S readily computed to very high accuracy (without the need for simulations) by simple quadrature[22]

$$L = \sigma \int_1^\infty dx \, \exp[-\phi_S A_L(x)] \tag{6}$$

where $A_L(x) = 8a(x^3 - 1) + 12b(x^2 - 1) + 24c\,(x - 1)$ and the coefficients a, b, and c are given below eq 5. Figure 3 shows the mean nearest-neighbor surface separation L as a function of solute volume fraction for trehalose and ubiquitin. For ubiquitin

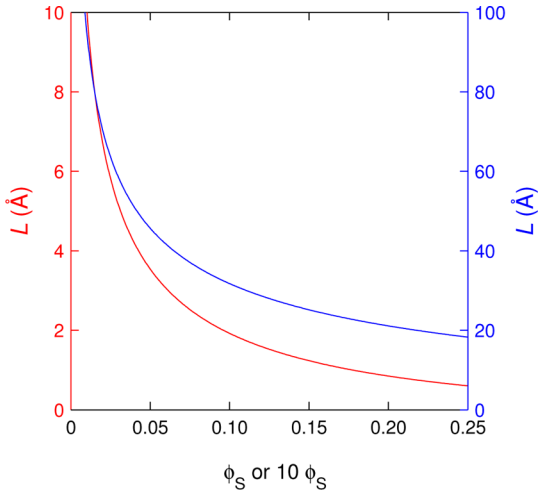

Figure 3. Mean nearest-neighbor surface separation L versus solute volume fraction ϕ_S for trehalose (red curve, left axis) and for ubiquitin (blue curve, right axis). For ubiquitin, the volume fraction on the abscissa has been multiplied by a factor 10.

dx.doi.org/10.1021/jp5083467 | J. Phys. Chem. B 2014, 118, 10806−10812

we thus obtain $L^* = L(\phi_S^*) = 32.4$ Å and $h = 16.2$ Å. (HG obtained a slightly larger value, $h = 18$ Å, because they used a 35% too large protein volume, deduced from the radius of gyration.[8]) But for any of these h values the hydration overlap at ϕ_S^* is only a few % (Figure 2) so the expansion in eq 2 can be truncated after the quadratic term. According to the RP3C model, the coefficient of this term is

$$B^{(2)} = \frac{1}{2}(\xi_H - 1)[2 - 9\kappa^2 - \kappa^3(\kappa^3 - 8)] \quad (7)$$

which is negative for $\xi_H > 1$. The magenta dash-dotted (nearly linear) curve in Figure 1 is the RP3C model prediction for $\xi_H = 1.063$ (as obtained from the nonlinear fit) and $h = 18$ Å (as inferred by HG from the overlap separation L^*). This prediction does not bear any semblance to the data. In particular, the observed strong nonlinearity with an α maximum at $\phi_S \approx 0.10-0.15$ cannot be explained. This internal inconsistency exposes a fundamental flaw in the procedure used by HG to infer the hydration range h from the mean nearest-neighbor separation at the overlap concentration. The large discrepancy between the RP3C model fit (Table 1) and the h value deduced by HG from the "overlap concentration" ϕ_S^* arises because, in a hard-sphere fluid, the mean nearest-neighbor separation L is much smaller than the typical neighbor separation, defined as the mean separation of the neighbors responsible for the first peak of the pair correlation function. If L had been a typical neighbor separation, then we would expect $L^* = 12$ Å for trehalose (Table 1). The onset of nonlinearity should then have occurred at $\phi_S^* = 0.007$ (Figure 3), an order of magnitude below the observed (weak) deviation from linearity ($\phi_S \approx 0.08$) in Figure 1.

In the RP3C model, a pronounced nonlinearity in $\alpha(\phi_S)$ requires substantial hydration overlap. To account for the strong nonlinearity observed for ubiquitin at volume fractions of a few % (Figure 1), the RP3C model requires each protein molecule to perturb on the order of 10^5 water molecules (Table 1). Such a scenario is incompatible with a large body of experimental and computational work on protein hydration dynamics,[11,29] indicating that the dynamic perturbation is essentially confined to the first water monolayer (~450 water molecules for ubiquitin). Moreover, for a hydration range h as large as 66 Å (or even 18 Å), the assumption in the RP3C model of a step function profile for $\alpha_H(r)$ is highly unphysical. This assumption is reasonable if the hydration domain comprises one or two water shells, but for the much larger h values deduced from the ubiquitin fit there must be a substantial variation of $\alpha_H(r)$ within the hydration domain.

The decay length λ of $\alpha_H(r)$ is not known, but a recent classical MD study[30] presented a spatially resolved analysis of the vibrational density of states, a readily computed proxy for α, for water near the lambda repressor fragment λ_{6-85}^*. It was thus found that the water−water autocorrelation and protein−water cross correlation both have decay lengths of 3−4 Å in the frequency range of interest here.[30] For an exponential decay with $\lambda = 3.5$ Å, the dynamic perturbation $\alpha_H(r) - \alpha_W$ would vary by 8 orders of magnitude across a 66 Å thick hydration shell. Even a substantial overlap of such weak tails of the perturbation profile would not be detectable, but the RP3C model assumes that removal of the outermost hydration waters (66 Å from the protein surface) has the same effect as removal of water in contact with the protein surface. We note that the dynamic heterogeneity discussed here is a property of the isolated (without overlap) hydration domain. A related, but distinct, problem is the interference of perturbations from two nearby solutes. The latter problem is likely to be most important at high volume fractions (as in the trehalose study), where it is intimately connected to the cross-correlation problem.

Another indication that HG's analysis is flawed comes from the observation that the extensive hydration shell of proteins (comprising ~10^5 water molecules) mysteriously disappears when the protein is altered by nondenaturing point mutations or pH variations.[7−9] As an example, Figure 1 shows HG's THz data for ubiquitin at pH 2.[8] According to the RP3C model, the slightly negative slope of the linearly decreasing data (Figure 1) implies that a pH drop from 4.8 to 2 is accompanied by a 30-fold reduction of either the magnitude ($\xi_H - 1$) or the extent (ν_H) of the dynamic perturbation (Table 1). HG attribute this drastic change to protein denaturation,[8,9] but without providing independent evidence for any structural change. Ubiquitin is an exceptionally stable protein that has been repeatedly shown[31−33] to remain in its native folded state over the pH range 1−13. In fact, even the hydration dynamics of ubiquitin has been studied down to pH 2, with no sign of denaturation and only a modest effect on the hydration dynamics, consistent with the protonation of ubiquitin's 12 carboxylate groups.[35,36] If ubiquitin remains folded down to pH 2, what is then the origin of the drastic change in the THz absorption? As for trehalose, solute−solute interactions and the associated cross correlations may be largely responsible for the nonlinear concentration dependence at pH 4.8. Indeed, ubiquitin undergoes reversible dimerization with dissociation constant $K_D = 4.9$ mM at pH 6.0 and 30 °C.[34] Only 1.7 carboxyl groups are protonated on going from pH 6.0 to 4.8, but at pH 2 essentially all carboxylate groups are protonated, yielding a net charge of +13 that likely prevents dimerization.

An even more striking pH effect was reported for the lambda repressor fragment λ_{6-85}^*, where a nonlinear concentration dependence similar to that seen for ubiquitin at pH 4.8 was found at pH 7.3, whereas at pH 5.0 no significant hydration effect was detected (i.e., $\alpha/\alpha_W \approx 1 - \phi_S$).[7] Using the same flawed analysis as for ubiquitin, HG inferred an extensive dynamic perturbation ($h \approx 22$ Å) at the higher pH.[6,9] Protein aggregation was reportedly[7,9] "ruled out" by a SAXS study,[37] which, however, examined the protein in a 45:55 v/v glycol/water mixture at −28 °C. As for ubiquitin, the abolished hydration effect at the lower pH was attributed[7,9] to protein unfolding even though the CD data showed no sign of denaturation at pH 5.0.[7]

Our main criticism of HG's analysis of THz absorption data from ubiquitin and other proteins[6−10] can be summed up in the following four points.

(1) HG attribute the inability of the RP3C model to account for the THz data from ubiquitin (and other proteins) to the large fraction of overlapping hydration shells.[8] However, for the hydration shell thickness deduced by HG ($h = 18$ Å), the overlap fraction is actually an order of magnitude *smaller* (Figure 2) than in the saccharide solutions, where the RP3C model *can* fit the data.

(2) Abandoning the RP3C model, HG deduce the average thickness (h) of the perturbed hydration shell from the protein concentration (ϕ_S^*) corresponding to the "onset of nonlinearity" in $\alpha(\phi_S)$. In doing so, they incorrectly identify the typical protein−protein separation with the much smaller mean nearest-neighbor separation. For the shell thickness thus deduced ($h = 18$ Å), the $\alpha(\phi_S)$ curve predicted by the RP3C

dx.doi.org/10.1021/jp5083467 | J. Phys. Chem. B 2014, 118, 10806−10812

model differs drastically and qualitatively from the observed curve, even at the lowest protein concentrations (Figure 1). To account for the observed maximum in $\alpha(\phi_S)$, the RP3C model requires a much thicker hydration shell (h = 66 Å).

(3) In either case (h = 18 or 66 Å), the result is based on the unphysical assumption that a water molecule at a distance h from the protein surface is as much perturbed as a water molecule in contact with the surface. In fact, MD simulations yield a decay length of 3−4 Å for the protein's perturbation of single-molecule water rotation as well as of collective THz water dynamics, consistent with NMR results but grossly inconsistent with the thick protein hydration layers inferred by HG.

(4) HG invoke protein unfolding to rationalize the nearly complete disappearance of the THz hydration effect upon pH reduction, but without verifying that the protein actually is unfolded. In fact, the available evidence strongly indicates that the proteins remain folded at the lower pH. On the other hand, significant protein self-association is indicated at the higher pH. As in the case of trehalose, the nonlinear $\alpha(\phi_S)$ dependence thus appears to be caused by solute−solute interactions.

Like HG's original work, this Reply has focused on the issue of the *range* of the perturbation, as reflected in the parameter h or ν_H. In closing, we shall briefly comment on the physical significance of the *magnitude* of the perturbation, as described by the parameter ξ_H. The NMR-derived dynamic perturbation factor ξ_H is the ratio of correlation times for water rotation in the hydration domain and in bulk water. The corresponding THz-derived quantity $\xi_H = \alpha_H/\alpha_W$ is not a ratio of characteristic time scales and therefore cannot be directly interpreted in terms of slowing down or speeding up of water dynamics. When the water−oxygen vibrational density of states is computed from classical MD simulations, one finds a slight blue-shift of the broad intermolecular vibration band for protein hydration water.[9] As a result of this speeding up of intermolecular vibrations in the first hydration shell, $\xi_H > 1$ in the frequency window (2.1−2.8 THz) commonly probed by HG, whereas $\xi_H < 1$ at 1.0 THz. Moreover, at an intermediate frequency of ∼1.5 THz one finds $\xi_H = 1$,[38] showing that $\alpha_H = \alpha_W$ is not necessarily an indication of an unperturbed hydration shell. A 10% shift of the intermolecular vibration frequency corresponds to an energetic perturbation of $0.025k_BT$. Such a small perturbation is not likely to have a significant impact on any molecular process. HG argue in their Comment[4] that the H-bond shift (56 cm^{-1}) of the intramolecular O−H stretch vibration frequency (3657 cm^{-1}) is likewise a small percentage,[39] but in absolute terms it is an order of magnitude larger than the intermolecular shift.

In conclusion, we have elaborated our earlier concerns[1,40] about HG's interpretation of their THz absorption data from saccharides and proteins. Whereas HG argue that "the extent to which the solute influences the solvent has been underestimated by other experimental techniques",[5] the analysis presented here indicates that the "long-range" hydration effect inferred by HG is an interpretational artifact. It is to be hoped that this issue will be further clarified by systematic experimental and theoretical studies, preferably in several laboratories, of the effect of independently monitored solute aggregation on the THz absorption coefficient.

■ ASSOCIATED CONTENT

Supporting Information

Three figures showing covariances of RP3C model parameters. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: bertil.halle@bpc.lu.se.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Swedish Research Council.

■ REFERENCES

(1) Winther, L. R.; Qvist, J.; Halle, B. Hydration and Mobility of Trehalose in Aqueous Solution. *J. Phys. Chem. B* **2012**, *116*, 9196−9207.

(2) Lupi, L.; Comez, L.; Paolantoni, M.; Perticaroli, S.; Sassi, P.; Morresi, A.; Ladanyi, B. M.; Fioretto, D. Hydration and Aggregation in Mono- and Disaccharide Aqueous Solutions by Gigahertz-to-Terahertz Light Scattering and Molecular Dynamics Simulations. *J. Phys. Chem. B* **2012**, *116*, 14760−14767.

(3) Heyden, M.; Bründermann, E.; Heugen, U.; Niehues, G.; Leitner, D. M.; Havenith, M. Long-Range Influence of Carbohydrates on the Solvation Dynamics of Water - Answers from Terahertz Absorption Measurements and Molecular Modeling Simulations. *J. Am. Chem. Soc.* **2008**, *130*, 5773−5779.

(4) Heyden, M.; Schwaab, G.; Havenith, M. Comment on "Hydration and Mobility of Trehalose in Aqueous Solution". *J. Phys. Chem. B* **2014**, DOI: 10.1021/jp508089t.

(5) Heugen, U.; Schwaab, G.; Bründermann, E.; Heyden, M.; Yu, X.; Leitner, D. M.; Havenith, M. Solute-Induced Retardation of Water Dynamics Probed Directly by Terahertz Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12301−12306.

(6) Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M. An Extended Dynamical Hydration Shell around Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20749−20752.

(7) Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Gruebele, M.; Leitner, D. M.; Havenith, M. Protein Sequence- and pH-Dependent Hydration Probed by Terahertz Spectroscopy. *J. Am. Chem. Soc.* **2008**, *130*, 2374−2375.

(8) Born, B.; Kim, S. J.; Ebbinghaus, S.; Gruebele, M.; Havenith, M. The Terahertz Dance of Water with the Proteins: The Effect of Protein Flexibility on the Dynamical Hydration Shell of Ubiquitin. *Faraday Discuss.* **2009**, *141*, 161−173.

(9) Heyden, M.; Havenith, M. Combining THz Spectroscopy and MD Simulations to Study Protein-Hydration Coupling. *Methods* **2010**, *52*, 74−83.

(10) Meister, K.; Ebbinghaus, S.; Xu, Y.; Duman, J. G.; DeVries, A.; Gruebele, M.; Leitner, D. M.; Havenith, M. Long-Range Protein-Water Dynamics in Hyperactive Insect Antifreeze Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 1617−1622.

(11) Mattea, C.; Qvist, J.; Halle, B. Dynamics at the Protein-Water Interface from ^{17}O Spin Relaxation in Deeply Supercooled Solutions. *Biophys. J.* **2008**, *95*, 2951−2963.

(12) Modig, K.; Qvist, J.; Marshall, C. B.; Davies, P. L.; Halle, B. High Water Mobility on the Ice-Binding Surface of a Hyperactive Antifreeze Protein. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10189−10197.

(13) Banipal, P. K.; Banipal, T. S.; Lark, B. S.; Ahluwalia, J. C. Partial Molar Heat Capacities and Volumes of Some Mono-, Di- and Tri-Saccharides in Water at 298.15, 308.15 and 318.15 K. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 81−87.

(14) Daszkiewicz, O. K.; Hennel, J. W.; Lubas, B. Proton Magnetic Relaxation and Protein Hydration. *Nature* **1963**, *200*, 1006−1107.

(15) Fister, F.; Hertz, H. G. O^{17}-NMR Study of Aqueous Electrolyte and Non-electrolyte Solutions. *Ber. Bunsenges. Phys. Chem.* **1967**, *71*, 1032−1040.
(16) Bernard, W.; Callen, H. B. Irreversible Thermodynamics of Nonlinear Processes and Noise in Driven Systems. *Rev. Mod. Phys.* **1959**, *31*, 1017−1044.
(17) Weingärtner, H.; Knocks, A.; Boresch, S.; Höchtl, P.; Steinhauser, O. Dielectric Spectroscopy in Aqueous Solutions of Oligosaccharides: Experiment Meets Simulation. *J. Chem. Phys.* **2001**, *115*, 1463−1472.
(18) Rudas, T.; Schröder, C.; Boresch, S.; Steinhauser, O. Simulation Studies of the Protein-Water Interface. II. Properties at the Mesoscopic Resolution. *J. Chem. Phys.* **2006**, *124* (234908), 1−19.
(19) Nilsson, L.; Halle, B. Molecular Origin of Time-Dependent Fluorescence Shifts in Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13867−13872.
(20) Halle, B.; Nilsson, L. Does the Dynamic Stokes Shift Report on Slow Protein Hydration Dynamics? *J. Phys. Chem. B* **2009**, *113*, 8210−8213.
(21) Reiss, H.; Frisch, H. L.; Lebowitz, J. L. Statistical Mechanics of Rigid Spheres. *J. Chem. Phys.* **1959**, *31*, 369−380.
(22) Torquato, S.; Lu, B.; Rubinstein, J. Nearest-Neighbor Distribution Functions in Many-Body Systems. *Phys. Rev. A* **1990**, *41*, 2059−2075.
(23) Torquato, S.; Lee, S. B. Computer Simulations of Nearest-Neighbor Distribution Functions and Related Quantities for Hard-Sphere Systems. *Phys. A* **1990**, *167*, 361−383.
(24) Coleman, T. F.; Li, Y. An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM J. Optim.* **1996**, *6*, 418−445.
(25) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes*, 3rd ed.; Cambridge University Press: New York, 1986.
(26) Kilburn, D.; Townrow, S.; Meunier, V.; Richardson, R.; Alam, A.; Ubbink, J. Organization and Mobility of Water in Amorphous and Crystalline Trehalose. *Nat. Mater.* **2006**, *5*, 632−635.
(27) Fischer, B. M.; Franz, M.; Abbott, D. THz Spectroscopy as a Versatile Tool for Investigating Crystalline Structures. In *Infrared and Millimeter Waves*, Proceedings of the 31st International Conference, Shanghai, China, Sept 18−22, 2006.
(28) Sapir, L.; Harries, D. Linking Trehalose Self-Association with Binary Aqueous Solution Equation of State. *J. Phys. Chem. B* **2011**, *115*, 624−634.
(29) Halle, B. Protein Hydration Dynamics in Solution: A Critical Survey. *Philos. Trans. R. Soc. London B* **2004**, *359*, 1207−1224.
(30) Heyden, M.; Tobias, D. J. Spatial Dependence of Protein-Water Collective Hydrogen-Bond Dynamics. *Phys. Rev. Lett.* **2013**, *111* (218101), 1−5.
(31) Lenkinski, R. E.; Chen, D. M.; Glickson, J. D.; Goldstein, G. Nuclear Magnetic Resonance Studies of the Denaturation of Ubiquitin. *Biochim. Biophys. Acta* **1977**, *494*, 126−130.
(32) Cary, P. D.; King, D. S.; Crane-Robinson, C.; Bradbury, E. M.; Rabbani, A.; Goodwin, G. H.; Johns, E. W. Structural Studies on Two High-Mobility-Group Proteins from Calf Thymus, HMG-14 and HMG-20 (Ubiquitin), and Their Interaction with DNA. *Eur. J. Biochem.* **1980**, *112*, 577−580.
(33) Woolfson, D. N.; Cooper, A.; Harding, M. M.; Williams, D. H.; Evans, P. A. Protein Folding in the Absence of the Solvent Ordering Contribution to the Hydrophobic Interaction. *J. Mol. Biol.* **1993**, *229*, 502−511.
(34) Liu, Z.; Zhang, W.-P.; Xing, Q.; Ren, X.; Liu, M.; Tang, C. Noncovalent Dimerization of Ubiquitin. *Angew. Chem., Int. Ed.* **2012**, *51*, 469−472.
(35) Denisov, V. P.; Halle, B. Protein Hydration Dynamics in Aqueous Solution: A Comparison of Bovine Pancreatic Trypsin Inhibitor and Ubiquitin by Oxygen-17 Spin Relaxation Dispersion. *J. Mol. Biol.* **1995**, *245*, 682−697.
(36) Halle, B. Water in Biological Systems: The NMR Picture. In *Hydration Processes in Biology: Theoretical and Experimental Approaches*; Bellisent-Funel, M.-C., Ed.; IOS Press: Dordrecht, 1999; pp 233−249.
(37) Kim, S. J.; Dumont, C.; Gruebele, M. Simulation-Based Fitting of Protein-Protein Interaction Potentials to SAXS Experiments. *Biophys. J.* **2008**, *94*, 4924−4931.
(38) Xu, J.; Plaxco, K. W.; Allen, S. J. Probing the Collective Vibrational Dynamics of a Protein in Liquid Water by Terahertz Absorption Spectroscopy. *Protein Sci.* **2006**, *15*, 1175−1181.
(39) Buck, U.; Huisken, F. Infrared Spectroscopy of Size-Selected Water and Methanol Clusters. *Chem. Rev.* **2000**, *100*, 3863−3890.
(40) Halle, B.; Havenith, M. General Discussion. *Faraday Discuss.* **2009**, *141*, 200−202.

Supporting Information

for

Reply to "Comment on 'Hydration and Mobility of Trehalose in Aqueous Solution' "

Bertil Halle*

Department of Biophysical Chemistry, Lund University, SE-22100 Lund, Sweden

E-mail: bertil.halle@bpc.lu.se

*To whom correspondence should be addressed

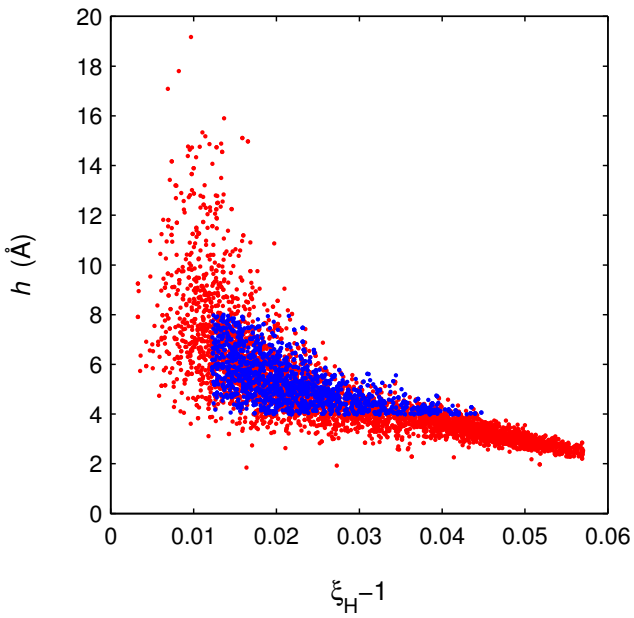

Figure S1: Results for the RP3C model parameters $\xi_{\mathrm{H}} - 1$ and h obtained from nonlinear fits to 5000 synthetic data sets generated with the trehalose parameters in Table 1 and a Gaussian error distribution consistent with the experimental error bars in Fig. 1. Red (blue) dots correspond to 95 % (68.3 %) confidence interval. The fits were performed with the trust region reflective nonlinear optimization algorithm, restricting the parameters α_{S} and h to physically admissible (nonnegative) values.

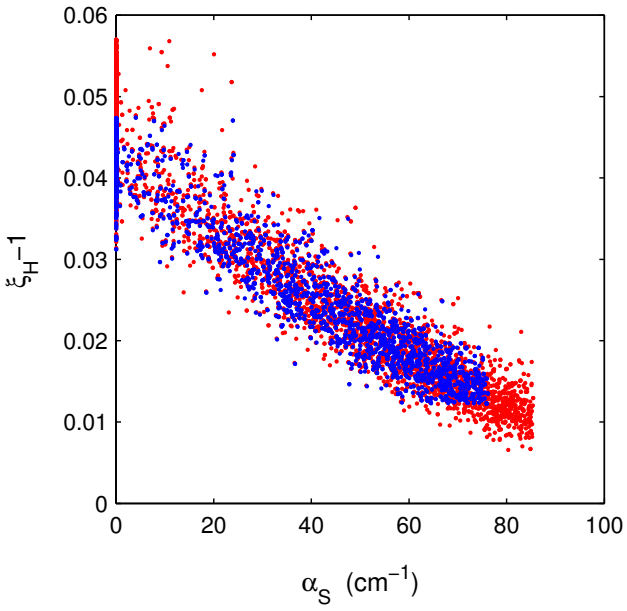

Figure S2: Results for the RP3C model parameters α_S and $\xi_H - 1$ obtained from nonlinear fits to 5000 synthetic data sets generated with the trehalose parameters in Table 1 and a Gaussian error distribution consistent with the experimental error bars in Fig. 1. Red (blue) dots correspond to 95 % (68.3 %) confidence interval. The fits were performed with the trust region reflective nonlinear optimization algorithm, restricting the parameters α_S and h to physically admissible (nonnegative) values.

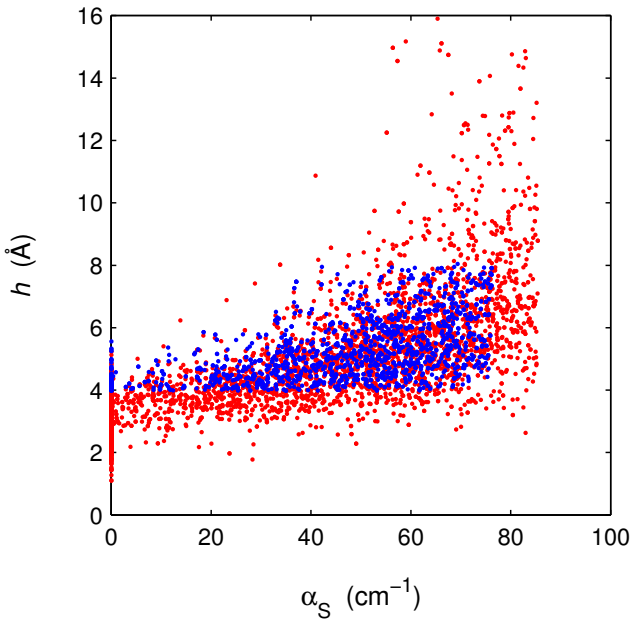

Figure S3: Results for the RP3C model parameters α_S and h obtained from nonlinear fits to 5000 synthetic data sets generated with the trehalose parameters in Table 1 and a Gaussian error distribution consistent with the experimental error bars in Fig. 1. Red (blue) dots correspond to 95 % (68.3 %) confidence interval. The fits were performed with the trust region reflective nonlinear optimization algorithm, restricting the parameters α_S and h to physically admissible (nonnegative) values.