Supporting information for

Desiccation tolerance- and globular- proteins adsorb similar amounts of water

Julia A. Brom[†] and Gary J. Pielak*,[†,‡,§,∥].

[†]Department of Chemistry, University of North Carolina at Chapel Hill (UNC-CH), [‡]Department of Biochemistry & Biophysics, UNC-CH, [§]Lineberger Cancer Center, UNC-CH, [∥]Integrative Program for Biological and Genome Sciences, UNC-CH, Chapel Hill, NC

Supplemental Tables

Table S1. Hydrated fraction of GB1 solvent-accessible surface area (SASA)

Time at 75% RH (hours)	Hydrated fraction of GB1 SASA[*]
0	0.55 ± 0.04
1	0.77 ± 0.05
2	0.8 ± 0.1
4	0.96 ± 0.06
6	1.05 ± 0.08
12	1.0 ± 0.1
24	1.1 ± 0.2
48	1.15 ± 0.04
72	1.16 ± 0.09

[*]Analysis assumes all water is bound to GB1. The values are calculated by multiplying the molar ratio of water to GB1 (as measured by thermogravimetric analysis) by the average amount of protein surface covered by a water molecule (20 $Å^2$),[1] and then dividing the result by the surface area of the native solution structure (3727 $Å^2$) as determined by the PyMOL get_area function using PDB 2QMT. Uncertainties are standard deviation of three independent measurements.

Table S2. Predicted equilibrium water content from sorptive groups

Protein	Predicted water content ~75% relative humidity (% weight)[*]
PvLEA4	21.4
GB1	18.8
CAHS D	19.1

[*]Predictions determined as per Leeder and Watt[2] using their values for degree of hydrophilicity of polar sorption sites in wool keratin.[2,3] We count the number of sorption sites by type in each protein assuming the proteins are completely accessible to vapor.[2] The moles of each type of sorption site per 100 g of protein are calculated by multiplying the moles of sorption site per protein by its inverse molecular weight and then by 100 g. We average Leeder and Watt's moles H_2O per mol sorption site at 65% and 80% relative humidity, assuming this procedure yields values comparable to 75% relative humidity.[2] The resulting moles of H_2O per type of sorption site are multiplied by the moles of sorption site per 100 g protein to yield g H_2O per 100 g protein. The results for each type of sorption site are summed, the sum divided by (100 g protein + summed g H_2O per 100 g protein) and then multiplied by 100 to yield weight percent water.

Table S3. Amino acid composition

Amino acid one-letter code	PvLEA4		GB1		CAHS D	
	#	% composition	#	% composition	#	% composition
A	13	9.1	6	10.7	26	11.5
C	3	2.1	0	0.0	0	0.0
D	19	13.3	5	8.9	9	4.0
E	14	9.8	5	8.9	33	14.5
F	1	0.7	2	3.6	4	1.8
G	8	5.6	4	7.1	12	5.3
H	1	0.7	1	1.8	8	3.5
I	1	0.7	6	10.7	10	4.4
K	26	18.2	3	5.4	18	7.9
L	6	4.2	1	1.8	10	4.4
M	5	3.5	3	5.4	6	2.6
N	4	2.8	2	3.6	7	3.1
P	0	0.0	0	0.0	5	2.2
Q	10	7.0	2	3.6	14	6.2
R	4	2.8	0	0.0	19	8.4
S	4	2.8	0	0.0	14	6.2
T	13	9.1	10	17.9	15	6.6
V	6	4.2	4	7.1	15	6.6
W	2	1.4	1	1.8	0	0.0
Y	3	2.1	3	5.4	2	0.9

Amino acid composition of each test protein was calculated using ProtParam on the ExPASy server.[4]

Table S4. Kyte-Doolittle hydropathy

Protein	Hydropathy*
PvLEA4	3.0
GB1	3.8
CAHS D	3.5

*Hydropathy is the 0-9 scaled Kyte-Doolittle hydropathy (9 most hydrophobic), calculated using CIDER.[5,6]

REFERENCES

1. Rupley JA, Gratton E, Careri G (1983) Water and globular proteins. Trends Biochem Sci 8**:** 18-22.
2. Leeder JD, Watt IC (1974) The stoichiometry of water sorption by proteins. J Colloid Interface Sci 48**:** 339-344.
3. Watt IC, Leeder JD (1968) 24—stoichiometric analysis of the wool-water isotherm. J Text Inst 59**:** 353-364.
4. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Ron D. Appel, Bairoch A (2005) Protein identification and analysis tools on the expasy server. In: *The proteomics protocols handbook.* (ed. JM Walker), pp. 571-607. Humana Press.
5. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157**:** 105-132.
6. Holehouse AS, Das RK, Ahad JN, Richardson MO, Pappu RV (2017) Cider: Resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. Biophys J 112**:** 16-21.