Supplementary information for

Dried protein structure revealed at the residue level by liquid-observed vapor exchange NMR

Candice J. Crilly[†], Julia A. Brom[†], Mark E. Kowalewski[†], Samantha Piszkiewicz[†], 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

Supplementary figures

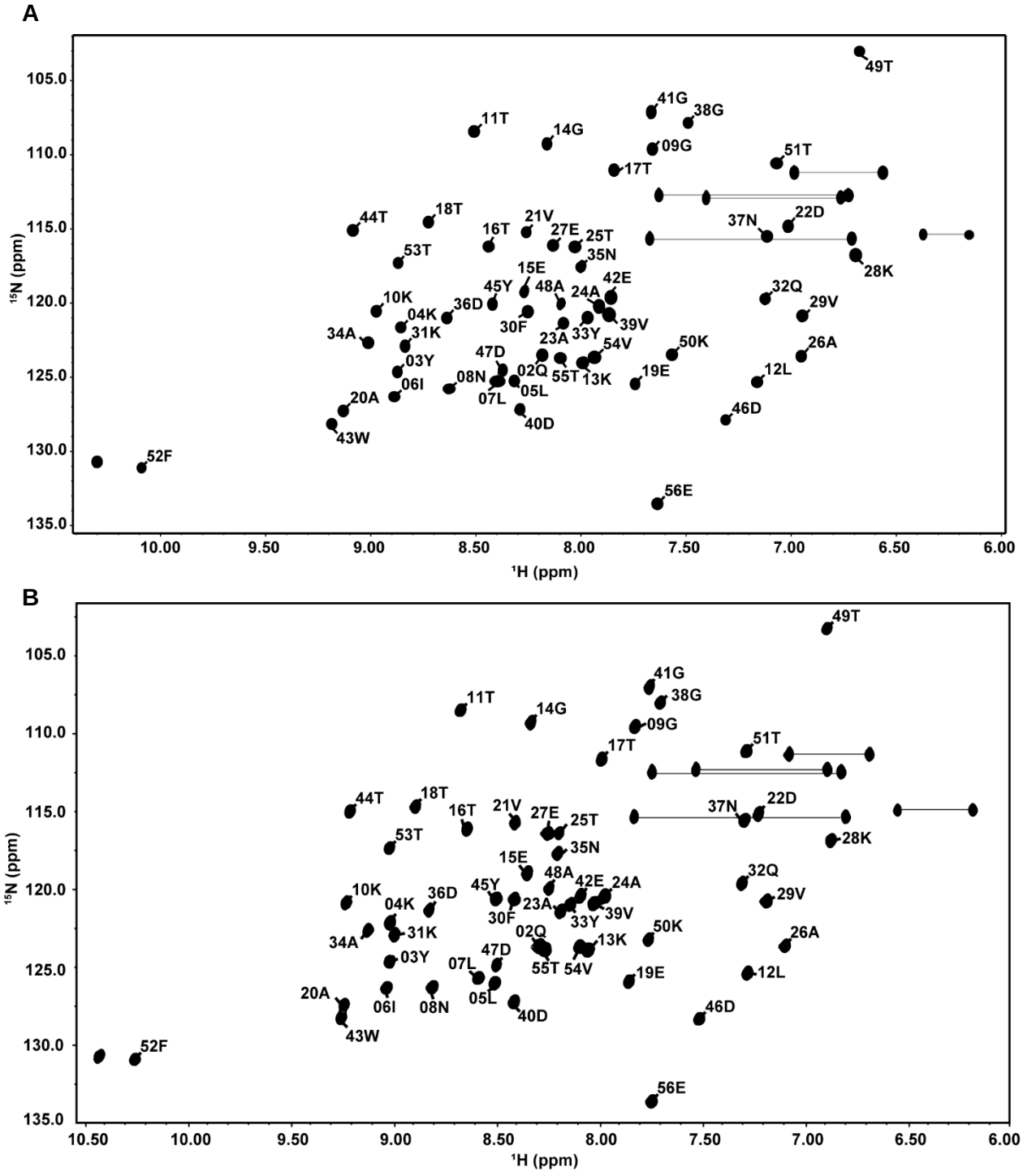

Figure S1. Backbone resonance assignments of GB1 T2Q at A) 4 °C in 100-mM citrate buffer (10% D_2O), pH 4.5 and B) 22 °C in 7.5 mM HEPES (10% D_2O) pH 7.5.

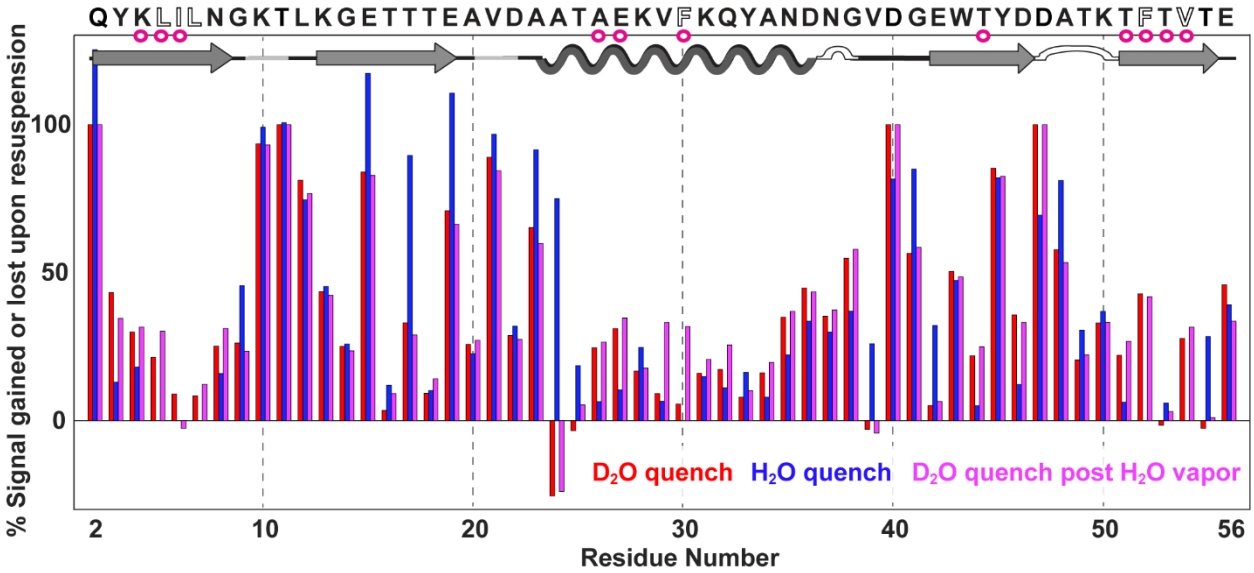

Figure S2. Quench-label profiles of GB1 exposed to different vapor and buffer conditions. The %quench-labelling experienced by a protonated (unexchanged) protein upon resuspension in D_2O-based quench buffer ("D_2O quench" values) is equivalent to the %signal lost upon resuspension, acquired and calculated as described in the main text. The %quench-labelling experienced by a sample of freeze-dried GB1 resuspended in D_2O quench buffer after 24-h incubation in a H_2O chamber at 75% relative humidity (achieved using a saturated NaCl solution – "D_2O quench post H_2O vapor" values) was determined in the same manner. To determine the %signal gained by a deuterium-exchanged protein upon resuspension in a H_2O-quench buffer ("H_2O quench" values), an aliquot of GB1 was twice exchanged into D_2O, lyophilized for 24 h, and resuspended in cold H_2O quench buffer before immediate HSQC spectrum acquisition; peak volumes were divided by the corresponding peak volumes of a fully protonated, non-exchanged sample (V_{T0}) and multiplied by 100%. The primary and secondary structures of GB1 are plotted at the top, with open letters indicating residues for which no signal is detected for the H_2O quenched sample. Measurements were made once for each condition.

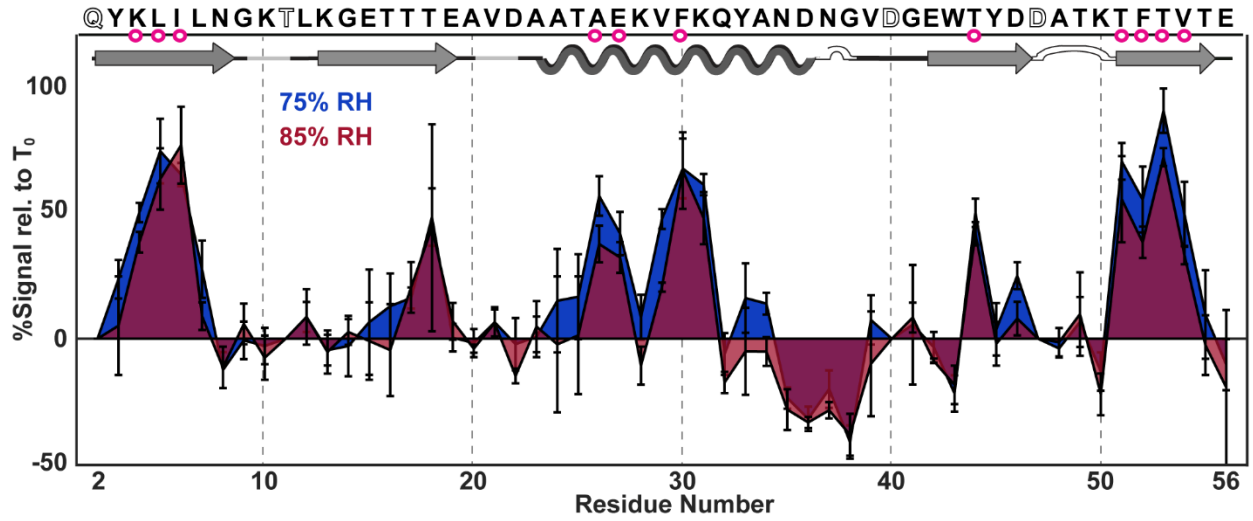

Fig. S3. % Signal remaining as a function of GB1 residue after exposure to 75% - and 85%- RH (D_2O) for 24 h. GB1 was lyophilized in 1.5 mM HEPES, pH 6.5 for 24 h and then placed at 75- or 85- % RH (D_2O) for 24 h before spectrum acquisition. Data were corrected for quench-labelling and normalized as described in Fig. 1. Primary and secondary structures are plotted at top, with pink circles indicating global unfolding residues and open letters indicating residues with undefined %signal remaining because they are 100% quench-labelled. Error bars represent uncertainty propagated from standard deviations of the mean from triplicate analysis.

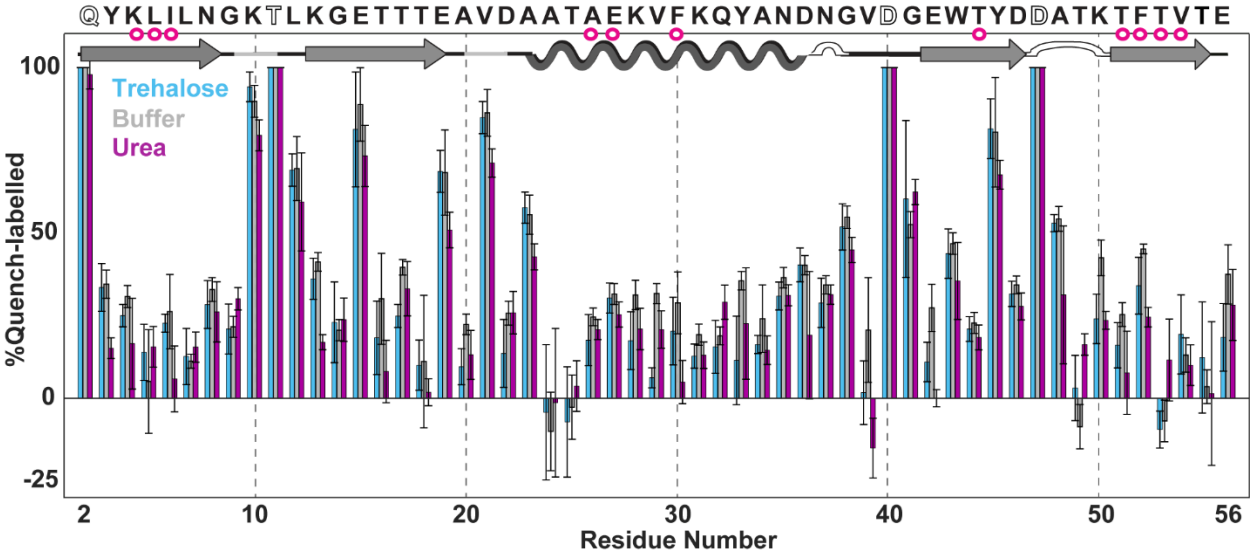

Figure S4. % Quench-labelled as a function of residue and cosolute plotted beneath a secondary structure map of GB1. ^{15}N-enriched GB1 was lyophilized in 1.5-mM HEPES + 20 g/L trehalose or urea and analyzed (Fig. 1). Percent quench-labelled was calculated by dividing the quench correction value for each residue by the respective T_0 volume (V_{T0}). Error bars represent uncertainties propagated from standard deviations of the mean from triplicate analysis. Other details are given in the caption to Figure 3.

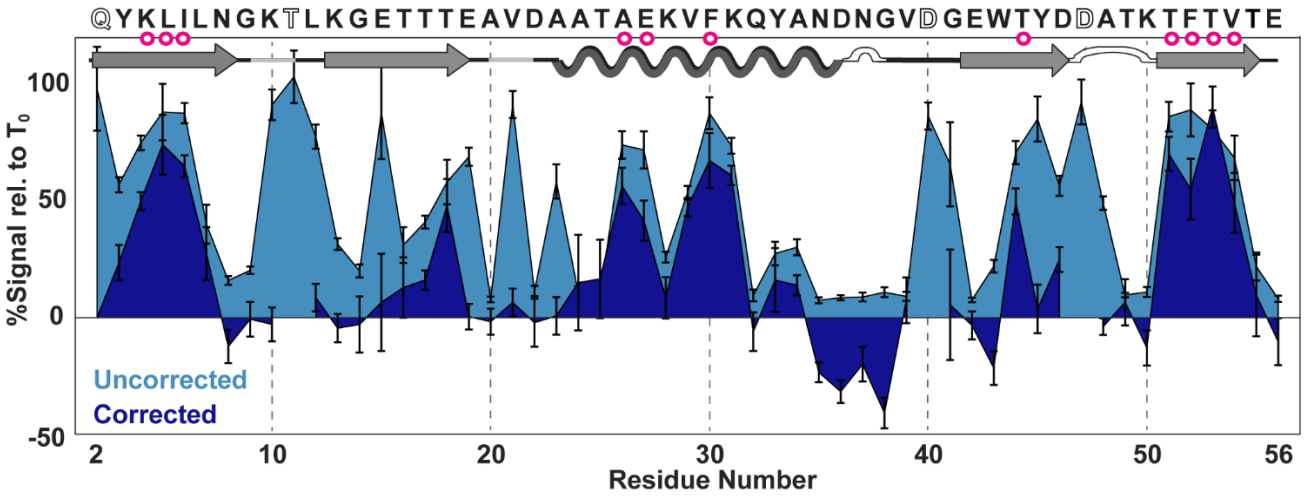

Figure S5. LOVE profiles of GB1 before and after applying the quench-correction, as described in the main text. Error bars represent uncertainty propagated from standard deviations of the mean from triplicate analysis. Other details are given in the caption to Figure 3.

Supplementary Tables

Table S1. ^{15}N and ^{1}H chemical shifts of GB1 backbone amides at pH 4.5, 4 °C (blue, left) and pH 7.5, 22 °C (pink, right).

Residue	δ ^{15}N (ppm)	δ ^{1}H (ppm)	δ ^{15}N (ppm)	δ ^{1}H (ppm)
02Q	123.5	8.2	123.6	8.3
03Y	124.6	8.9	124.6	9.0
04K	121.6	8.9	122.1	9.0
05L	125.2	8.3	126.0	8.5
06I	126.3	8.9	126.3	9.0
07L	125.3	8.4	125.7	8.6
08N	125.8	8.6	126.3	8.8
09G	109.6	7.7	109.5	7.8
10K	120.5	9.0	120.9	9.2
11T	108.4	8.5	108.5	8.7
12L	125.3	7.2	125.4	7.3
13K	124.0	8.0	123.9	8.0
14G	109.3	8.2	109.3	8.3
15E	119.2	8.3	119.0	8.3
16T	116.2	8.4	116.1	8.6
17T	111.0	7.8	111.6	8.0
18T	114.6	8.7	114.7	8.9
19E	125.4	7.7	125.9	7.9
20A	127.3	9.1	127.4	9.2
21V	115.2	8.3	115.7	8.4
22D	114.8	7.0	115.2	7.2
23A	121.4	8.1	121.4	8.2
24A	120.2	7.9	120.5	8.0
25T	116.2	8.0	116.4	8.2
26A	123.6	7.0	123.6	7.1
27E	116.1	8.1	116.4	8.2
28K	116.7	6.7	116.9	6.9
29V	120.8	6.9	120.8	7.2
30F	120.6	8.3	120.6	8.4
31K	122.9	8.8	122.9	9.0
32Q	119.7	7.1	119.6	7.3
33Y	121.0	8.0	121.0	8.1
34A	122.7	9.0	122.6	9.1
35N	117.6	8.0	117.7	8.2
36D	121.0	8.6	121.4	8.8
37N	115.5	7.1	115.6	7.3
38G	107.8	7.5	108.0	7.7
39V	120.8	7.9	120.9	8.0
40D	127.2	8.3	127.2	8.4
41G	107.1	7.7	107.0	7.8
42E	119.6	7.9	120.4	8.1
43W	128.2	9.2	128.3	9.2
44T	115.1	9.1	115.0	9.2
45Y	120.1	8.4	120.6	8.5
46D	127.9	7.3	128.3	7.5
47D	124.5	8.4	124.8	8.5
48A	120.0	8.1	120.0	8.2
49T	103.0	6.7	103.3	6.9
50K	123.5	7.6	123.2	7.8
51T	110.6	7.1	111.1	7.3
52F	131.1	10.1	130.9	10.3
53T	117.3	8.9	117.3	9.0
54V	123.7	7.9	123.7	8.1
55T	123.7	8.1	123.8	8.3
56E	133.5	7.6	133.6	7.7

Table S2. Signal change of T_{24} due to solution HDX during LOVE NMR spectrum acquisition*

Residue†	%Signal 0 h‡	%Signal 12 h‡	%Change in 12 h‡	%Signal change during expt§
N8	19	60	41	2
G9	27	78	51	6
K13	42	84	42	10
G14	17	51	34	2
T17	30	47	18	13
E19	83	96	12	24
D22	14	51	37	2
A23	71	92	21	18
A24	23	88	65	4
D36	12	35	23	1
N37	11	25	14	1
G38	12	47	35	2
G41	68	79	11	26
W43	31	85	54	4
A48	73	94	22	24
T49	13	56	43	2
E56	11	44	33	2

**Determined by taking serial HSQCs of T_{24} over 12 h period, at pH 4.5 and 4 °C. The majority of signal gain by T_{24} is accounted for by the quench correction, because D_0 witnesses a similar degree of signal loss due to solution HDX.*

†*Only residues that witness a change in signal >10% in 12 h are shown*

‡*Percent is defined relative to average T_0 signal for that residue, and is not corrected for quench-labelling*

§*Calculated from the experimentally-determined observed rate constant (k_{obs}) obtained by fitting normalized data to the equation $\%Signal = 100*(1-e^{-kobs * t})$, where t is time between resuspension and completion of spectrum acquisition (30 minutes for all LOVE NMR experiments performed for this manuscript)*

Table S3. Freeze-dried protein water content

Condition	Initial mass[†] (mg)	Final mass[†] (mg)	%H_2O at 0 h (w/w)	%(H_2O + D_2O) at 72 h (w/w)	%SASA covered by H_2O + D_2O at 72 h[§]
Buffer	2.6	2.8	10	19	40
+Urea	15.6	16.0	2	5	60
+Trehalose	16.1	17.4	5	12	170

[†]Assumes no sample lost during lyophilization or transfer to TGA instrument.

$$Mass = Total\ mass\ measured\ by\ TGA\ x \frac{Theoretical\ dry\ mass}{Measured\ dry\ mass}$$

[§]Calculated by multiplying the molar ratio of bound H_2O to protein by the average amount of protein surface covered by a water molecule (20 Å^2)[1] and dividing by the surface area of the native solution structure of GB1 (3727 Å^2), as determined by the PyMOL get_area function for PDB 2QMT.

Table S4. Average %Protected values of GB1 lyophilized in buffer only or with cosolutes

Residue	%Protected ± SEM†		
	Buffer	+ 20 g/L trehalose	+ 20 g/L urea
02Q	N/A	N/A	N/A
03Y	20 ± 10	*66 ± 7	17 ± 3
04K	50 ± 4	*69 ± 6	20 ± 10
05L	70 ± 10	*100 ± 30	16 ± 7
06I	65 ± 5	*80 ± 20	30 ± 10
07L	30 ± 10	82 ± 5	13 ± 5
08N	-12 ± 7	-9 ± 4	-11 ± 9
09G	-1 ± 7	1 ± 4	-11 ± 4
10K	-3 ± 7	-3 ± 8	11 ± 4
11T	N/A	N/A	N/A
12L	9 ± 6	7 ± 10	20 ± 20
13K	-5 ± 6	-11 ± 4	19 ± 2
14G	0 ± 10	-1 ± 4	-2 ± 7
15E	10 ± 20	0 ± 10	30 ± 10
16T	10 ± 10	0 ± 12	10 ± 10
17T	16 ± 4	2 ± 5	-13 ± 8
18T	50 ± 10	*80 ± 30	27 ± 4
19E	0 ± 5	2 ± 7	20 ± 20
20A	-2 ± 6	-13 ± 4	-4 ± 8
21V	7 ± 6	9 ± 7	24 ± 4
22D	0 ± 10	-13 ± 5	-14 ± 6
23A	1 ± 8	4 ± 6	16 ± 4
24A	20 ± 20	*30 ± 10	20 ± 20
25T	20 ± 20	15 ± 6	4 ± 8
26A	60 ± 10	*84 ± 5	14 ± 3
27E	40 ± 10	*70 ± 5	10 ± 4
28K	9 ± 9	56 ± 7	-2 ± 7
29V	47 ± 4	*66 ± 6	‡3 ± 6
30F	70 ± 10	*70 ± 10	37 ± 7
31K	61 ± 5	*85 ± 7	22 ± 4
32Q	-6 ± 8	46 ± 6	-23 ± 5
33Y	20 ± 10	*61 ± 5	0 ± 20
34A	14 ± 4	50 ± 20	-1 ± 4
35N	-23 ± 4	-7 ± 9	-24 ± 4
36D	-32 ± 5	-30 ± 4	0 ± 20
37N	-20 ± 8	-15 ± 5	-22 ± 3
38G	-41 ± 7	-41 ± 4	-34 ± 4
39V	10 ± 10	50 ± 30	27 ± 9
40D	N/A	N/A	N/A
41G	0 ± 20	9 ± 5	-30 ± 3
42E	-3 ± 6	20 ± 10	8 ± 3
43W	-22 ± 7	-22 ± 5	-10 ± 10
44T	50 ± 6	*77 ± 6	15 ± 4
45Y	0 ± 10	0 ± 10	0 ± 4
46D	25 ± 5	*60 ± 6	3 ± 4
47D	N/A	N/A	N/A
48A	-4 ± 4	-6 ± 5	30 ± 20
49T	10 ± 10	22 ± 6	3 ± 5
50K	-13 ± 8	-25 ± 3	-14 ± 3
51T	70 ± 9	*78 ± 7	30 ± 10
52F	60 ± 10	*56 ± 3	9 ± 3
53T	90 ± 10	*107 ± 8	30 ± 10
54V	50 ± 10	*85 ± 6	21 ± 7
55T	10 ± 20	*98 ± 9	30 ± 20
56E	-10 ± 10	-7 ± 8	-20 ± 10

*$V_{T24} = V_{T0}$ (i.e. vapor exchange was not observable).

†Standard error of the mean propagated from standard errors of the mean from triplicate analysis of each dataset used to calculate %Protected (T_0, D_0, and T_{24}).

Table S5. Average free energies of opening in solution ($\Delta G_{op}°'$) [§]

Residue	$\Delta G_{op}°'$ ± SEM[‡] (kcal/mol) Buffer	+ 100 g/L trehalose	+ 100 g/L urea
02Q	N/A	N/A	N/A
03Y	7.59 ± 0.01	7.83 ± 0.02	6.46 ± 0.01
04K	8.19 ± 0.01	8.52 ± 0.01	N/A
05L	8.13 ± 0.00	8.48 ± 0.01	6.61 ± 0.01
06I	8.11 ± 0.02	8.21 ± 0.02	6.46 ± 0.03
07L	6.16 ± 0.04	6.26 ± 0.03	5.68 ± 0.01
08N	5.25 ± 0.01	5.23 ± 0.02	4.74 ± 0.01
09G	4.84 ± 0.03	4.80 ± 0.02**	N/A
10K	N/A	N/A	N/A
11T	N/A	N/A	N/A
12L	N/A	N/A	N/A
13K	N/A	N/A	N/A
14G	4.55 ± 0.03	4.61 ± 0.02**	N/A
15E	N/A	N/A	N/A
16T	5.34 ± 0.02	5.28 ± 0.08	4.627 ± 0.004
17T	N/A	N/A	N/A
18T	7.00 ± 0.04	7.15 ± 0.02	6.52 ± 0.01
19E	N/A	N/A	N/A
20A	4.81 ± 0.01	4.79 ± 0.04	3.96 ± .02**
21V	N/A	N/A	N/A
22D	N/A	N/A	N/A
23A	N/A	N/A	N/A
24A	N/A	N/A	N/A
25T	4.46 ± 0.02	4.46 ± 0.01	N/A
26A	8.62 ± 0.01	8.78 ± 0.00	7.42 ± 0.01
27E	8.43 ± 0.01	8.61 ± 0.00	6.97 ± 0.01
28K	5.89 ± 0.03	5.96 ± 0.03	5.72 ± 0.01
29V	5.84 ± 0.03	5.93 ± 0.04	5.74 ± 0.01
30F	8.53 ± 0.01	8.75 ± 0.00	6.96 ± 0.01
31K	8.65 ± 0.00	8.90 ± 0.01	6.99 ± 0.01
32Q	5.55 ± 0.04	5.63 ± 0.03	5.43 ± 0.00
33Y	6.61 ± 0.04	6.67 ± 0.02	6.29 ± 0.01
34A	7.36 ± 0.02	7.526 ± 0.003	6.78 ± 0.01
35N	6.97 ± 0.04	7.03 ± 0.05	6.55 ± 0.01
36D	5.60 ± 0.04	5.65 ± 0.02	5.39 ± 0.02
37N	6.30 ± 0.03	6.37 ± 0.03	6.07 ± 0.02
38G	4.91*	4.99 ± 0.04**	N/A
39V	6.11 ± 0.03	6.23 ± 0.02	5.89 ± 0.01
40D	3.47 ± 0.03	3.49 ± 0.03**	N/A
41G	N/A	N/A	N/A
42E	5.45 ± 0.04	5.54 ± 0.03	5.36 ± 0.01
43W	N/A	N/A	N/A
44T	8.755 ± 0.003	8.89 ± 0.01	7.32 ± 0.01
45Y	N/A	N/A	N/A
46D	7.60 ± 0.01	7.69 ± 0.01	7.00 ± 0.01
47D	N/A	N/A	N/A
48A	N/A	N/A	N/A
49T	N/A	N/A	N/A
50K	5.70 ± 0.03	5.76 ± 0.03	5.58 ± 0.01
51T	8.64 ± 0.01	8.93 ± 0.00	7.20 ± 0.01
52F	8.64 ± 0.01	8.95 ± 0.01	7.15 ± 0.01
53T	8.61 ± 0.00	8.90 ± 0.01	7.05 ± 0.01
54V	8.33 ± 0.00	8.66 ± 0.00	6.87 ± 0.02
55T	6.83 ± 0.04	6.96 ± 0.03	6.48 ± 0.01
56E	2.90 ± 0.03	2.96 ± 0.03	2.88 ± 0.01

Footnotes

[§] $\Delta G_{op}°'$ values from NMR-detected solution H-D exchange at 22 °C, pH 7.5.

[‡] Standard error of the mean from triplicate analysis. Stars mark residues for which only *one or **two measurements were made.

Table S6. Correlations between residue-specific predictors and Δ%Protected$_{buffer→cosolute}$

Predictor[†]	Δ%Protected$_{buffer→trehalose}$		Δ%Protected$_{buffer→urea}$	
	Coeff.	R^2	*Coeff.*	R^2
$SASA_{backbone}$	-	0.15	+	0.25
$CASA_{backbone}$	-	0.14	+	0.25
Δ $SASA_{U-->F}$	-	0.08	+	0.25
# H-bonds	+	0.07	-	0.14
# H-bonds with water	-	0.06	+	0.08
# Weak H-bonds with water	-	0.06	+	0.01
k_{int} (pH 4.5)	-	0.05	+	0.03
# Highly frustrated contacts	+	0.04	-	0.00
# Ionic interactions	-	0.03	+	0.00
# Minimally frustrated contacts	+	0.03	-	0.03
# Aromatic contacts	-	0.02	-	0.03
# Neutral frustrated contacts	+	0.02	-	0.02
$CASA_{apolar}$	-	0.02	+	0.12
k_{int} (pH 7.0)	-	0.02	+	0.00
# Weak H-bonds	+	0.02	-	0.08
$CASA_{total}$	-	0.02	+	0.22
$SASA_{total}$	-	0.02	+	0.21
$\Delta G°_{tr, Octanol --> H2O}$	+	0.02	-	0.06
$SASA_{polar}$	-	0.01	+	0.16
# Carbonyl interactions	+	0.01	-	0.18
SASA ratio (polar:nonpolar)	-	0.01	+	0.03
$SASA_{apolar}$	-	0.01	+	0.11
$\Delta G°_{tr, Cyclohexane --> H2O}$	+	0.01	-	<.01
$\Delta G°_{tr, NMA --> H2O}$ (normalized)	-	0.01	+	<.01
$\Delta G°_{tr, EtOH --> H2O}$ (normalized)	-	< 0.01	+	<.01
$\Delta G°_{tr, EtOH --> H2O}$	+	< 0.01	-	0.07
$\Delta G°_{tr, vapor --> H2O}$ (normalized)	+	< 0.01	+	<.01
$\Delta G°_{tr, NMA --> H2O}$	+	< 0.01	-	0.03
Net frustration parameter	+	< 0.01	-	0.02
# Hydrophobic contacts	+	< 0.01	-	0.09
$\Delta G°_{tr, Cyclohexane --> H2O}$ (normalized)	-	< 0.01	-	0.04
$CASA_{polar}$	-	< 0.01	+	0.15
$CASA_{sidechain}$	-	< 0.01	+	0.10
$\Delta G°_{tr, vapor --> H2O}$	-	< 0.01	+	0.01
$\Delta G°_{tr, Octanol --> H2O}$ (normalized)	+	< 0.01	+	<.01
$SASA_{sidechain}$	+	< 0.01	+	0.09

[†]Predictor values are calculated using the native structure of GB1 (PDB 2QMT), with the exception of non-normalized transfer free energies ($\Delta G°_{tr}$). The specific types of intra- and inter-molecular contacts (H-bonds, carbonyl interactions, etc.) were obtained from the Arpeggio server.[2] Solvent-accessible surface areas (SASA), cosolute-accessible surface areas (CASA), and change in surface area upon folding ($\Delta SASA_{U-->F}$) were obtained from the ProtSA server using a probe radius of 1.40-, 4.00-, and 1.85- Å for water, trehalose, and urea, respectively.[3] Frustration parameters are from the Frustratometer server.[4] Transfer free energies of amino acid side chains are from.[5-6] Normalized transfer free energies were calculated by multiplying literature values by the fraction of sidechain area accessible to solvent in the native structure.

References

(1) Rupley, J. A., Gratton, E., and Careri, G. (1983) Water and globular proteins, *Trends Biochem. Sci. 8*, 18-22.
(2) Jubb, H. C., Higueruelo, A. P., Ochoa-Montano, B., Pitt, W. R., Ascher, D. B., and Blundell, T. L. (2017) Arpeggio: A web server for calculating and visualising interatomic interactions in protein structures, *J. Mol. Biol. 429*, 365-371.
(3) Estrada, J., Bernado, P., Blackledge, M., and Sancho, J. (2009) ProtSA: A web application for calculating sequence specific protein solvent accessibilities in the unfolded ensemble, *BMC Bioinf. 10*, 104.
(4) Jenik, M., Parra, R. G., Radusky, L. G., Turjanski, A., Wolynes, P. G., and Ferreiro, D. U. (2012) Protein frustratometer: A tool to localize energetic frustration in protein molecules, *Nucleic Acids Res. 40*, W348-351.
(5) Damodaran, S., and Song, K. B. (1986) The role of solvent polarity in the free energy of transfer of amino acid side chains from water to organic solvents, *J. Biol. Chem. 261*, 7220-7222.
(6) Radzicka, A., and Wolfenden, R. (1988) Comparing the polarities of the amino acids: Side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution, *Biochemistry 27*, 1664-1670.