Supporting Information

Protein Solvent Shell Structure Provides Rapid Analysis of Hydration Dynamics

Jayangika N. Dahanayake,[1] Elaheh Shahryari,[1] Kirsten M. Roberts,[1] Micah E. Heikes,[1] Chandana Kasireddy,[2] Katie R. Mitchell-Koch[1]*

*Correspondence to: Katie R. Mitchell-Koch (E-mail: katie.mitchell-koch@wichita.edu)

[1] Department of Chemistry, Wichita State University, 1845 Fairmount Street, Wichita, KS, 67260-0051, United States.

[2] Current address: National Institute of Environmental Health Sciences, Bioinformatics111 TW Alexander Drive, Durham, NC, 27709, USA

Details of Simulation Protocols

The following simulation protocol was used for each protein studied. Each protein was centered in a cubic periodic box with a minimum distance of 1.0 nm between protein and any side of the box. Next, each protein was solvated with water, and the system was neutralized by adding Na^+ and Cl^- ions, replacing solvent molecules, to neutralize the systems at a 0.15 M salt concentration. Bond lengths were constrained using LINCS algorithm,[1] and Particle Mesh Ewald summation[2] was used for the electrostatic interaction calculations. A grid spacing of 0.12 nm combined with an interpolation order of 4 was used for long range interactions and van der Waals interactions were calculated with a cut-off of 1.4 nm. Energy minimization was carried out using the steepest descent algorithm,[3] in order to remove clashes between atoms that were too close. The system was gradually heated by increasing the temperature from 50 K to 300 K in 200 ps while position restraints were used on heavy atoms. For equilibration, a 20 ns trajectory at constant NPT was run. Equilibration was carried out at 300 K temperature using V-rescale thermostat[4] and at 1 bar pressure using Berendsen barostat.[5] Production runs were carried out with constant NVT

ensemble for 100 ns at 300 K, using Nosé-Hoover thermostat for a canonical ensemble.[6-7] In the RNase A protein simulation, the π-tautomeric (Nδ1-H) form of His12 and His119 was assigned, based on our recent evaluation of RNase A mechanistic experiments; elsewhere, the τ-tautomer (Nε2-H) of histidine was modeled.[8] For CALB and HEWL protein systems, two more trajectories were produced at 278 K and 288 K, to assess the validity of correlation between RDF and diffusive dynamics at different temperatures.

Details of Protocols for the Rapid Analysis of Hydration Dynamics

Protocol 1: Dividing protein regions

The secondary structure of protein was considered based on the secondary structure details of X-ray crystallographic structures provided in the RCSB Protein Data Bank website. The protein was divided into regions as individual secondary structures (α helices and β strands), and connectors between secondary structures. The solvent-accessible surface areas per residue were calculated for the initial crystal structure protein using the structure analysis tool of Chimera software package (areaSAS attribute calculator was used per residue). Buried/interior surface areas were identified using these calculations, and these buried/interior protein residues were excluded when dividing protein regions.

Protocol 2: Calculating RDF and determining first solvent shell cut-off

GROMACS software package was used to calculate the RDF of water around protein regions divided as described above. In this study, water center of mass RDF around protein side chains were calculated for different protein regions (divided according to Protocol 1). Using gmx make_ndx command, separate index files were generated for side chains of each protein region. Then, using gmx rdf command, water center of mass RDFs were calculated with respect to these side chain atoms of each protein region.

RDFs were visually inspected for a regular shape (see Fig. 3 in main text), and in order to determine the cut-off for the hydration layer. The lower edge of the first peak was considered as the cut off for the hydration layer (see Figure S1).

As described in the Methods section, regions with anomalous hydration shell structures according to RDF (Figure 3(b)), though rare, should be excluded when estimating hydration dynamics from water structure.

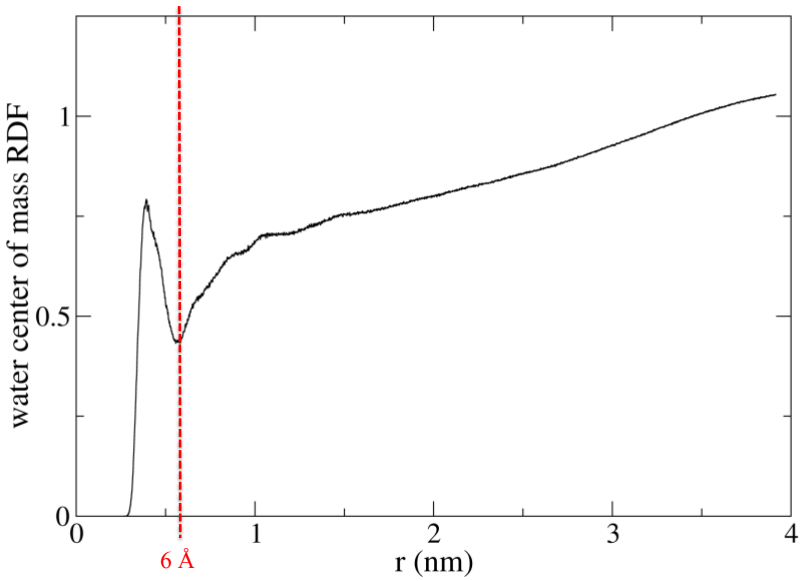


Figure S1. An example of selecting the first solvent shell cut-off using RDF of water center of mass around protein side chain atoms.

Protocol 3: RDF hydration layer integration for solvation layer around protein regions

MATLAB numerical program was used for the RDF peak integration. Using MATLAB numerical integration method, RDF peaks were integrated out to 6 Å, which is consistently the edge of the hydration layer.

Protocol 4: Identifying slow-diffusing/buried crystallographic water

Keeping all the crystallographic water identified in X-ray crystallographic structure, a 5 ns short MD-simulation was run. GROMACS software package was used for the calculation of diffusion coefficients for crystallographic water molecules. Using gmx make_ndx command,

separate index files were generated for these crystallographic water molecules. Next, using gmx msd command, the diffusion coefficients of these water molecules were calculated from mean square displacement using the Einstein relationship. By visually inspecting the mean square displacement graph, the best time interval over which $<r^2>$ was consistently linear was selected as the least square fitting time interval for a straight line.

Next, using these calculated diffusion coefficients as input and using the script developed and used in our previously published work,[9] an output with residue numbers of slowest waters (water molecules that have diffusion coefficient < 1/3 average diffusion coefficient, as suggested in our previously published work)[9] were obtained.

Visual Molecular Dynamics (VMD) program was used to identify the protein residues within 3 Å of these slow-diffusing crystallographic water. These protein residues were considered to be the locations where buried or long-residing water molecules are present. These regions should be excluded when estimating hydration dynamics from water structure.

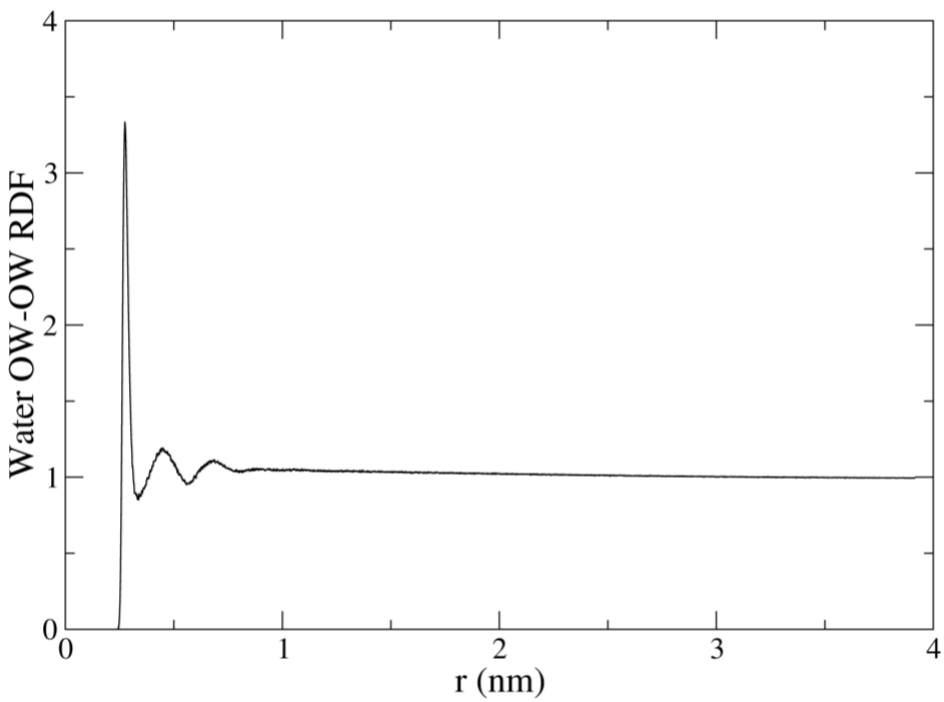


Figure S2. RDF for bulk water in the CALB protein simulation box, at a distance of 10 Å from the protein.

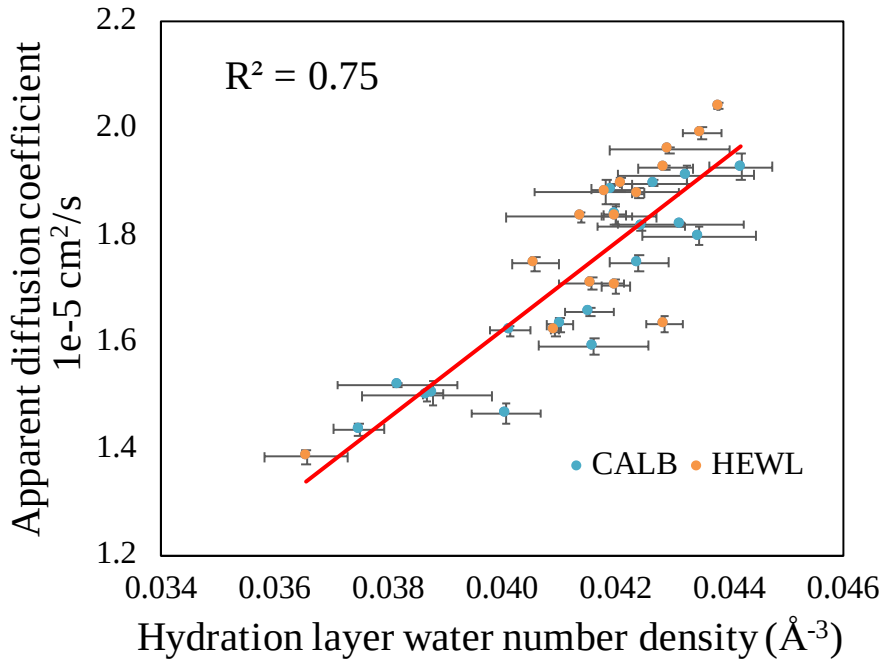


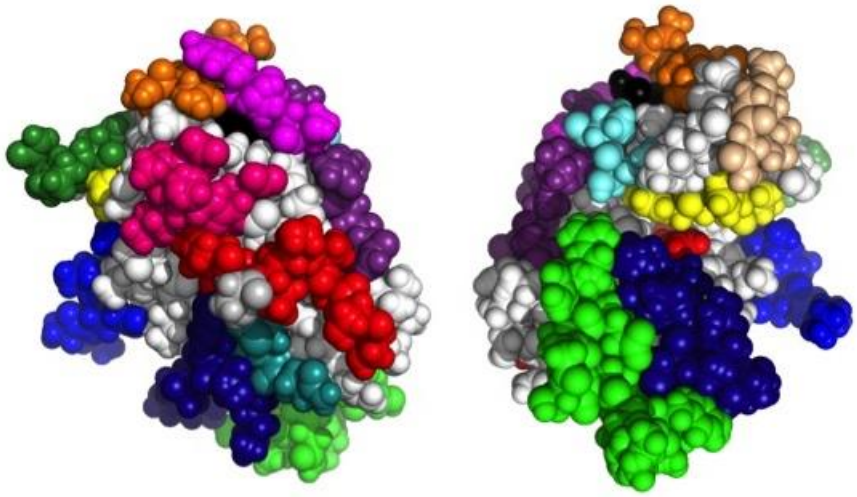
Figure S3. Linear correlation of local water density calculated using hydration shell excluded volume (details are provided in protocols), with local solvent shell water apparent diffusion coefficients.

<u>Protocol for the calculation of hydration layer water density using excluded volume</u>

Regional hydration shell excluded volumes were calculated using ProteinVolume free webserver[10] available on http://gmlab.bio.rpi.edu. Using these hydration shell excluded volumes and number of water molecules present at each regional hydration shell, water densities were calculated.

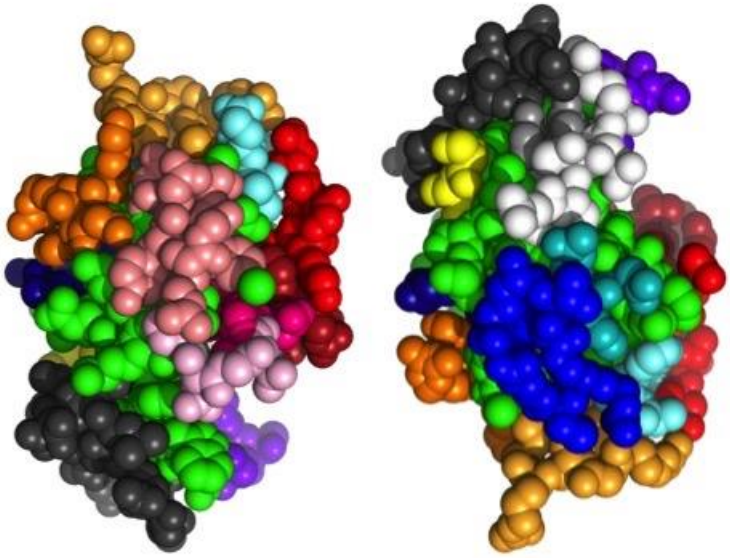
Figure S4. External, solvent-exposed regions of proteins used for analysis of solvent shell structure and regional hydration dynamics.

a) Protein: RNase A



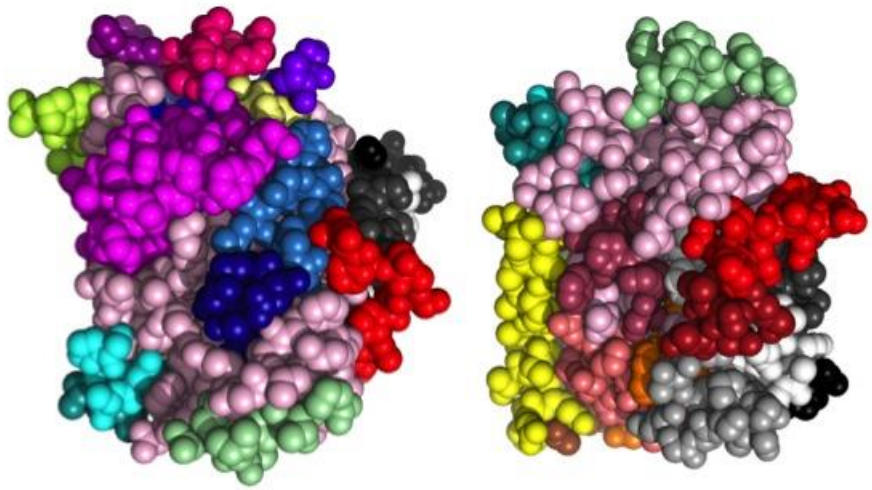
Region		Color	Residue
	1	blue	1 to 5
	2	red	12 to 22
	3	cyan	27 to 29
	4	dark blue	31 to 43
	5	dark pink	49 to 53
	6	orange	59 to 63
	7	wheat (beige)	67 to 71
	8	magenta	76 to 78
	9	green	85 to 96
	10	dark purple	100 to 105
	11	forest green	113 to 115
	12	yellow	116 to 120
	13	aquamarine	122 to 124

b) Protein: Lysozyme



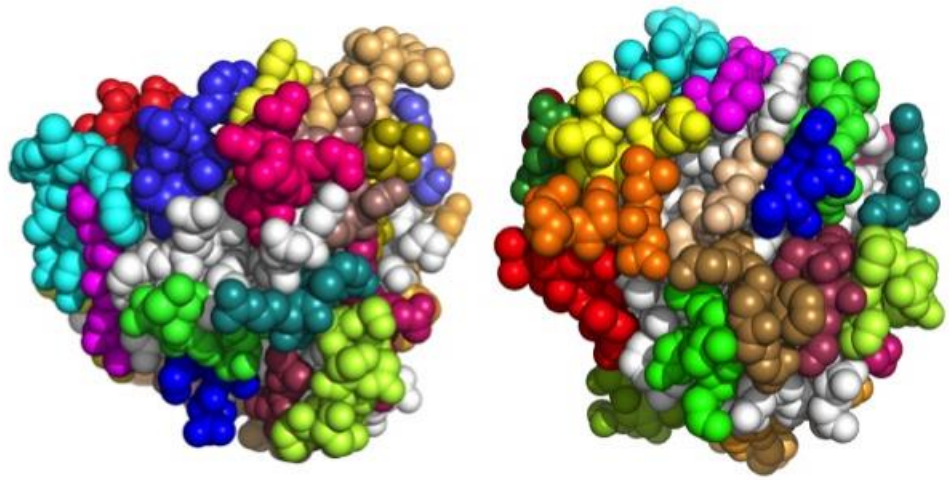
Region	Color	Residues
1	blue	1 to 7
2	orange	13 to 16
3	salmon	18 to 24
4	teal	37 to 39
5	white	41 to 45
6	purple	46 to 49
7	gray	51 to 53
8	dark gray	65 to 79
9	yellow	81 to 83
10	dark blue	87 to 89
11	pink	100 to 103
12	hot pink	104 to 108
13	brick red	109 to 113
14	red	114 to 119
15	light blue	120 to 123
16	light orange	124 to 129

c) Protein: *Candida antarctica* lipase B (CALB)



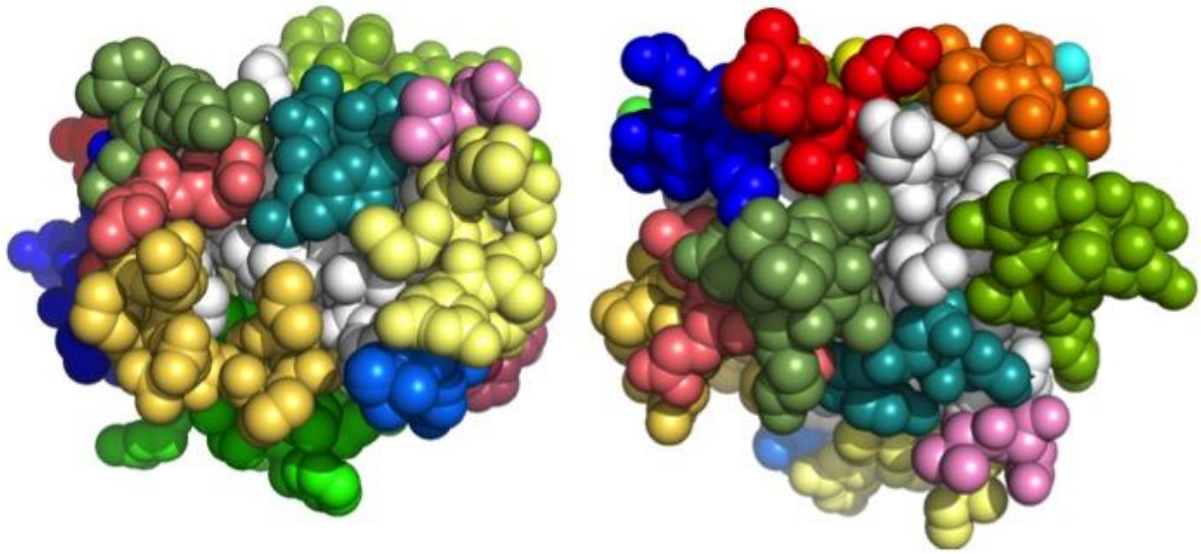
Region		Color	Residues
	1	red	1 to 12
	2	maroon	13 to 18 (alpha 1)
	3	light gray	20 to 32
	4	deep salmon	44 to 57 (alpha 2)
	5	orange	58 to 66
	6	raspberry	67 to 75
	7	white	76 to 93 (alpha 3)
	8	black	94 to 98
	9	dark gray	118 to 124
	10	light blue	142 to 146 (alpha 5)
	11	deep teal	147 to 151
	12	dark blue	162 to 169 (alpha 7)
	13	sky blue	170 to 178
	14	magenta	184 to 211
	15	limon	217 to 225
	16	blue	225-242 (alpha 9)
	17	purple blue	243 to 245
	18	pale yellow	246 to 248
	19	hot pink	249 to 253
	20	deep purple	254 to 262
	21	brown	264 to 267
	22	yellow	268 to 287 (alpha 10)
	23	pale green	309 to 317

d) Protein: subtilisin *Carlsberg* (SC)



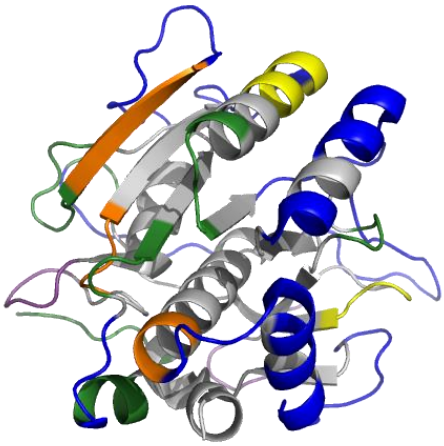
Region		Color	Residues
	1	red	1 to 6
	2	green	14-18
	3	sand	19-22
	4	blue	24-27
	5	yellow	36-43
	6	magenta	44-49
	7	cyan	50-62
	8	orange	75-80
	9	wheat	86-89
	10	tv red	96-103
	11	tv blue	104-109
	12	tv green	116-120
	13	tv yellow	125-132
	14	hot pink	133-138
	15	deep teal	143-146
	16	light orange	155-167
	17	dark salmon	170-173
	18	split pea	181-189
	19	olive	194-197
	20	firebrick	209-211
	21	forest	212-218
	22	raspberry	235-237
	23	limon	238-245
	24	warm pink	238-245
	25	bright orange	247-249
	26	slate	252-259

e) Protein: CheY

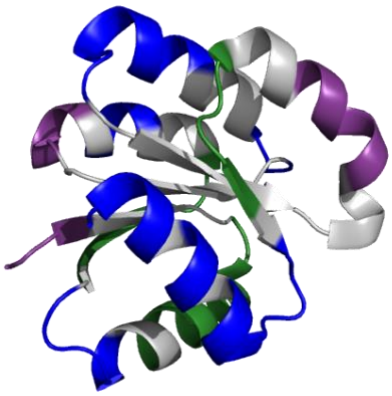


Region	Color	Residue
1	red	3 to 6
2	green	12 to 22
3	blue	25-29
4	yellow	31-38
5	magenta	40-41
6	cyan	43-45
7	orange	47-50
8	raspberry	59-64
9	splitpea	70-78
10	marine	87-88
11	pale yellow	90-95
12	deep teal	100-104
13	pink	97-98
14	yellow orange	106-113
15	deep salmon	115-117
16	smudge	119-125

a) Subtilisin



b) CheY



c) RNase A

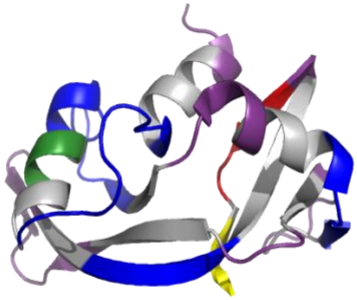


Figure S5. Diffusion maps for each protein. Protein regions are color coded according to diffusion coefficients (purple > 1.90 x 10^{-5} cm^2/s, blue 1.80-1.89 x 10^{-5} cm^2/s, green 1.70-1.79 x 10^{-5} cm^2/s, yellow 1.60-1.69 x 10^{-5} cm^2/s, orange 1.50-1.59 x 10^{-5} cm^2/s, red < 1.50 x 10^{-5} cm^2/s water diffusion coefficients).

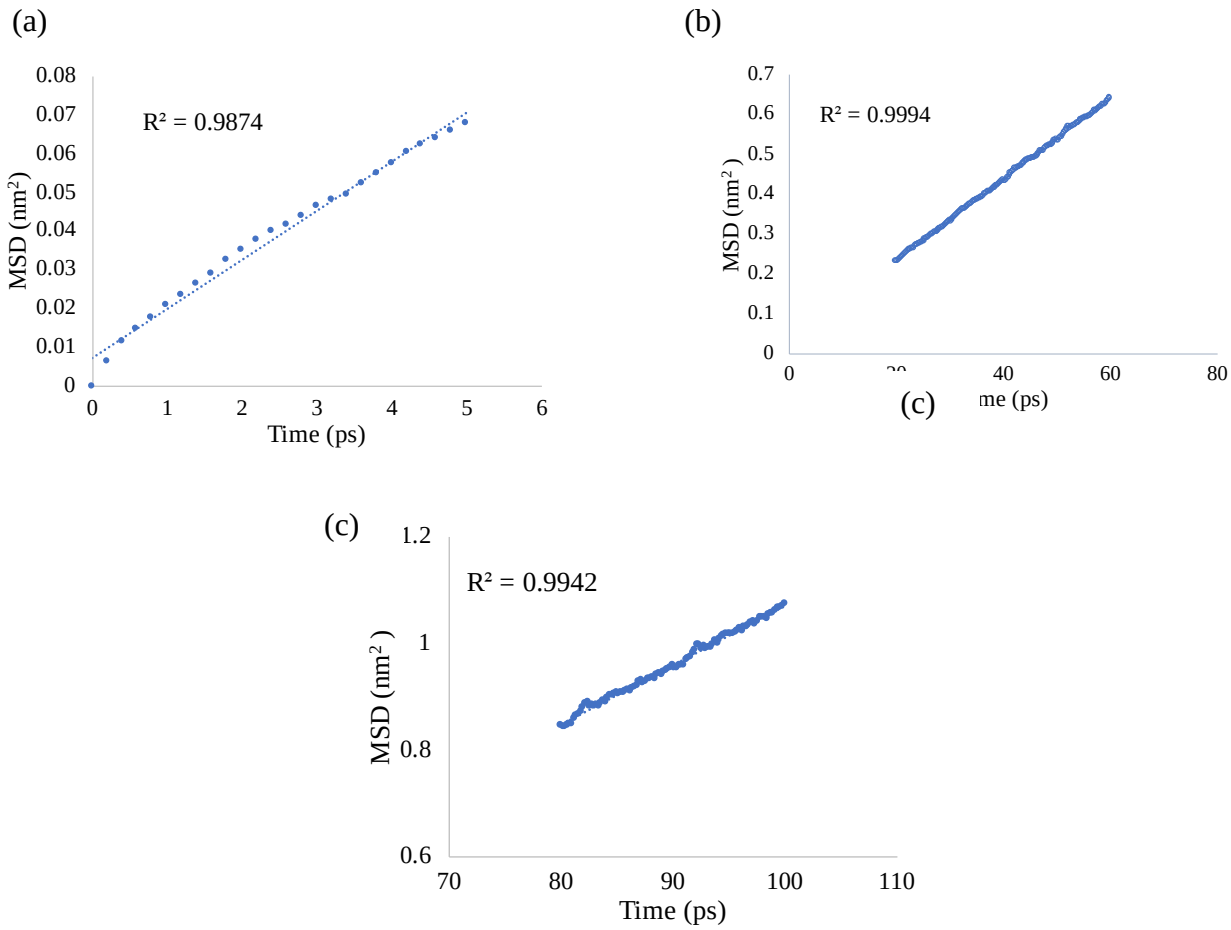


Figure S6. Mean square displacement linear fitting for different time intervals (a) 0-5 ps, (b) 20-60 ps, (c) 80-100 ps.

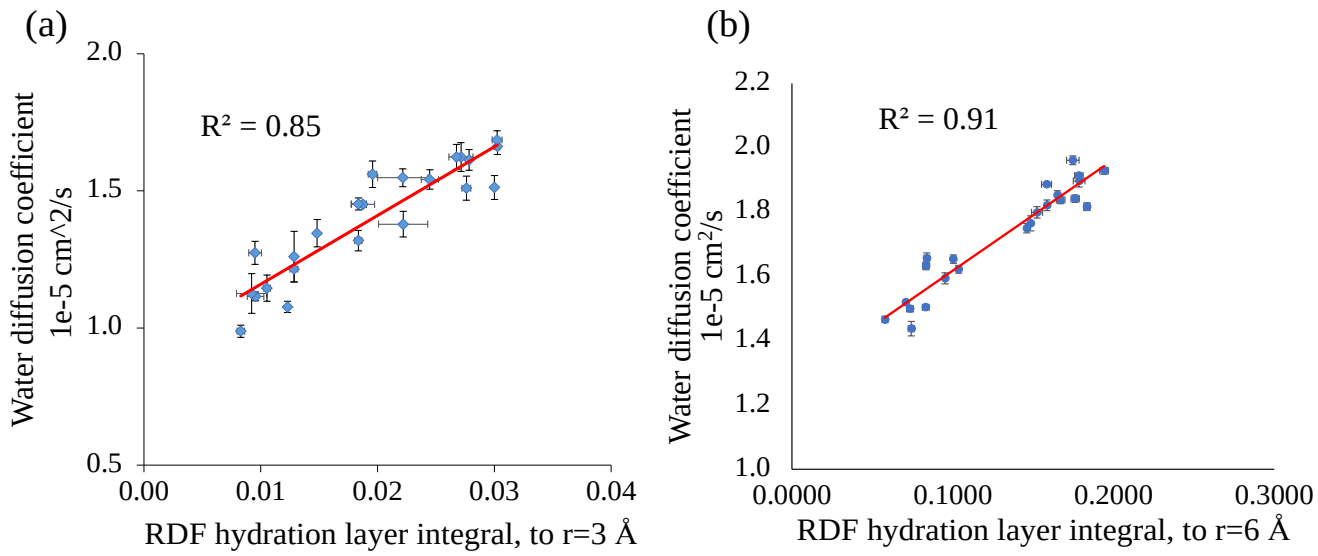


Figure S7. For CALB protein (a) relationship between apparent hydration layer diffusion coefficients and integration of **water oxygen** around protein side chain atom RDF by region (b) relationship between apparent hydration layer diffusion coefficients and integration of water **center of mass** RDF by region.

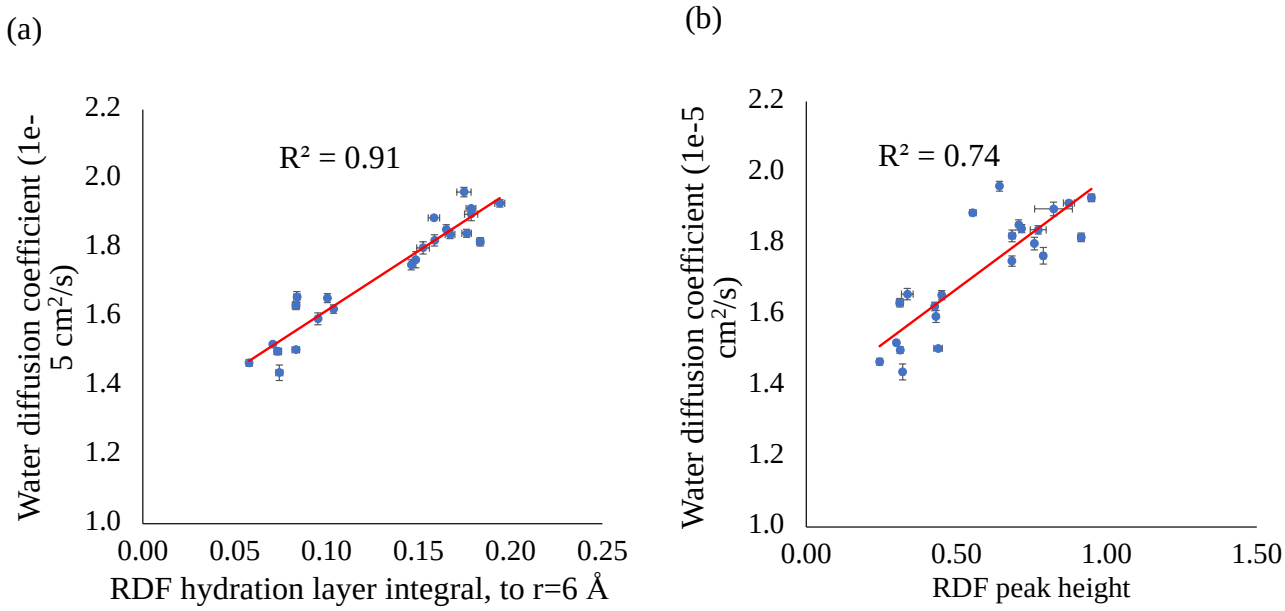


Figure S8. Correlation between local solvent shell water diffusion and water structure for CALB protein, considering (a) RDF hydration layer integration, (b) RDF peak height.

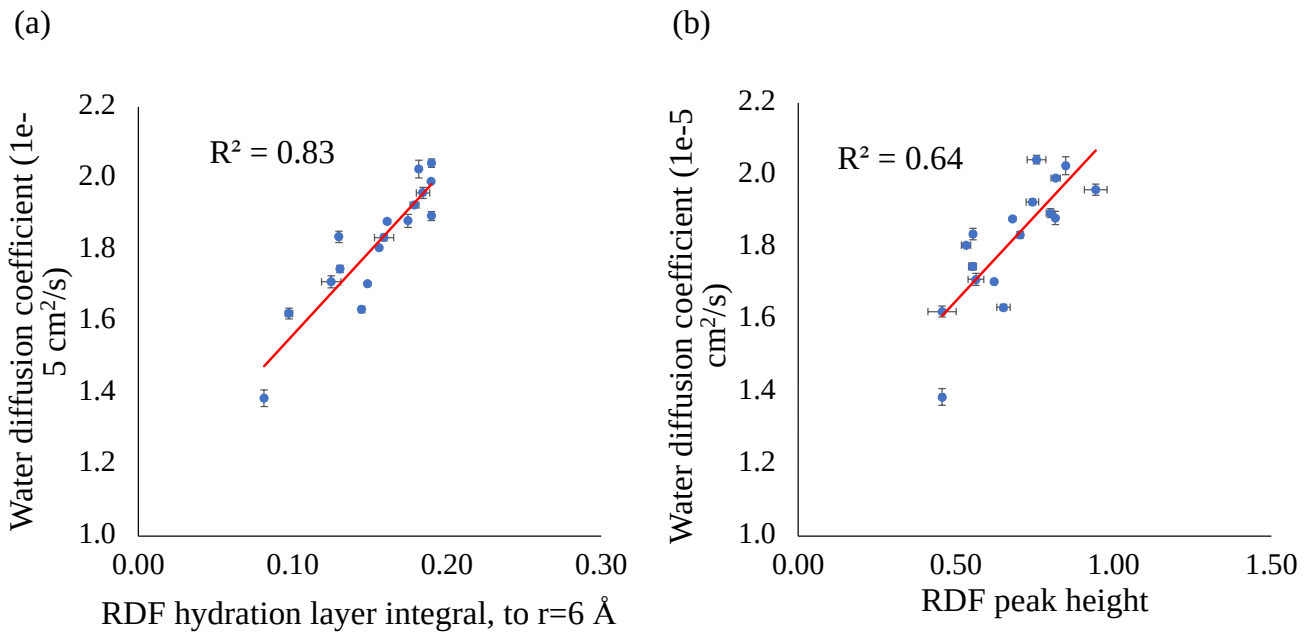


Figure S9. Correlation between local solvent shell water diffusion and water structure for HEWL protein, considering (a) RDF hydration layer integration, (b) RDF peak height.

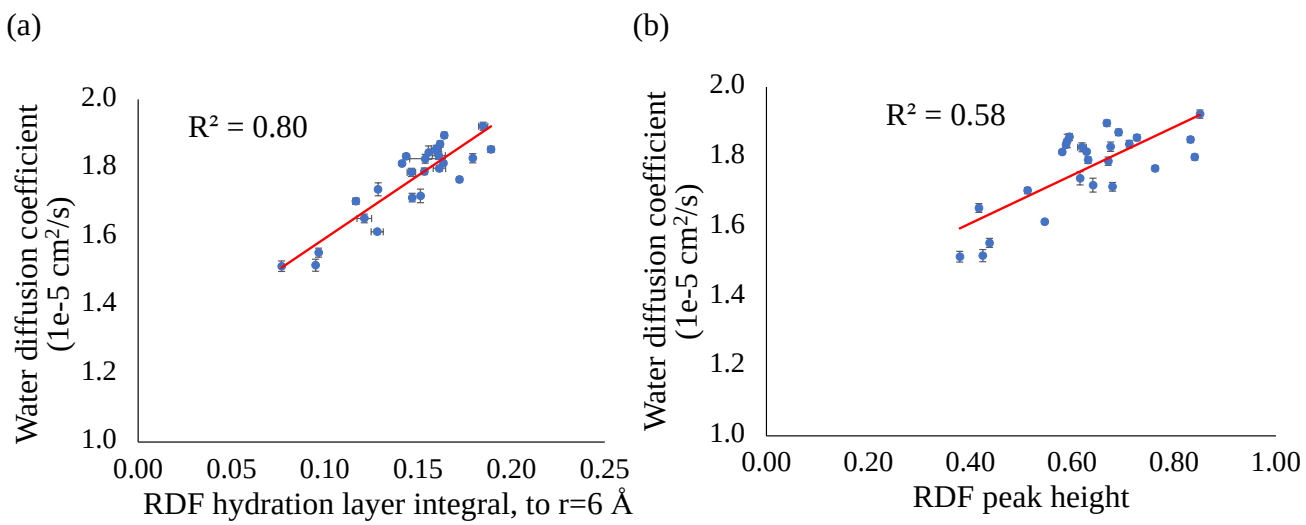


Figure S10. Correlation between local solvent shell water diffusion and water structure for Subtilisin Carlsberg (SC) protein, considering (a) RDF hydration layer integration, (b) RDF peak height.

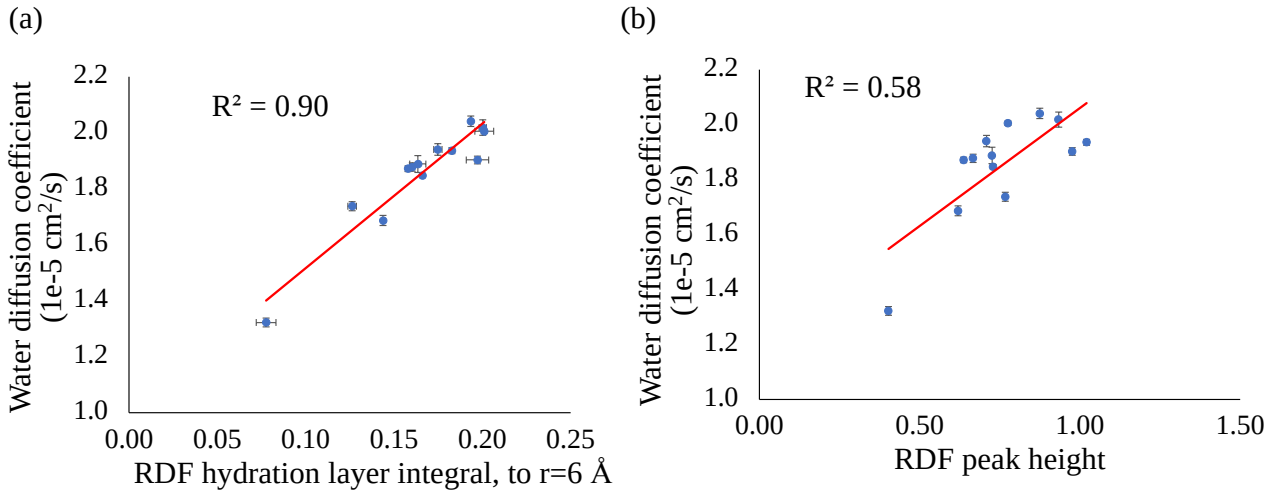


Figure S11. Correlation between local solvent shell water diffusion and water structure for RNase A protein, considering (a) RDF hydration layer integration, (b) RDF peak height.

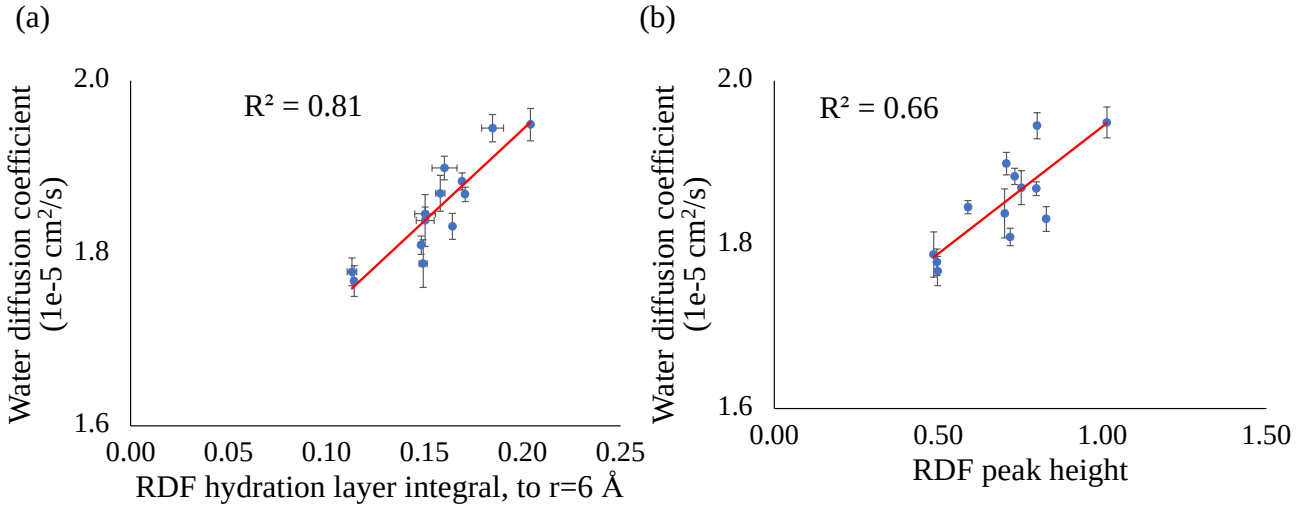


Figure S12. Correlation between local solvent shell water diffusion and water structure for CheY protein, considering (a) RDF hydration layer integration, (b) RDF peak height.

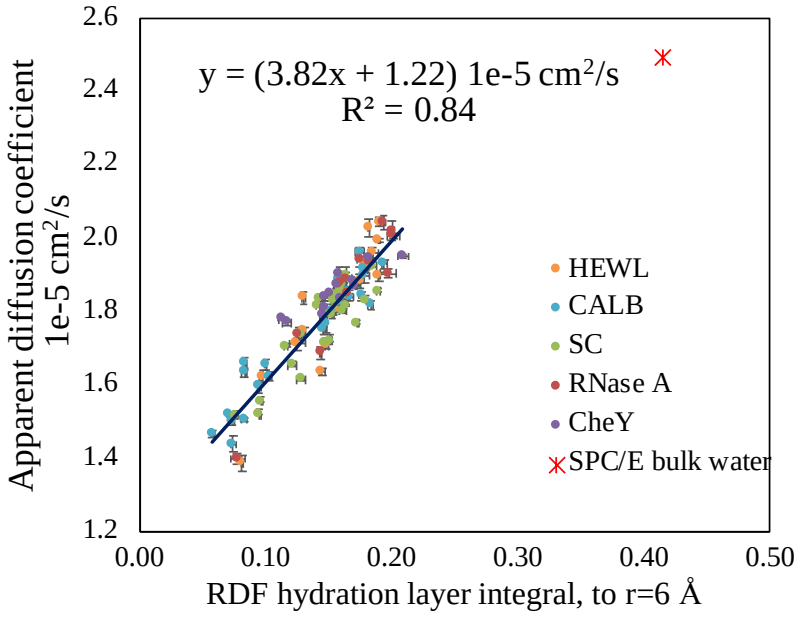


Figure S13. Linear correlation of local water structure obtained using integration of the *hydration layer* RDF of water center of mass around protein side chain atoms, with local solvent shell water diffusion coefficients across a set of five proteins, with the data point correspondence to SPC/E bulk water (red star).

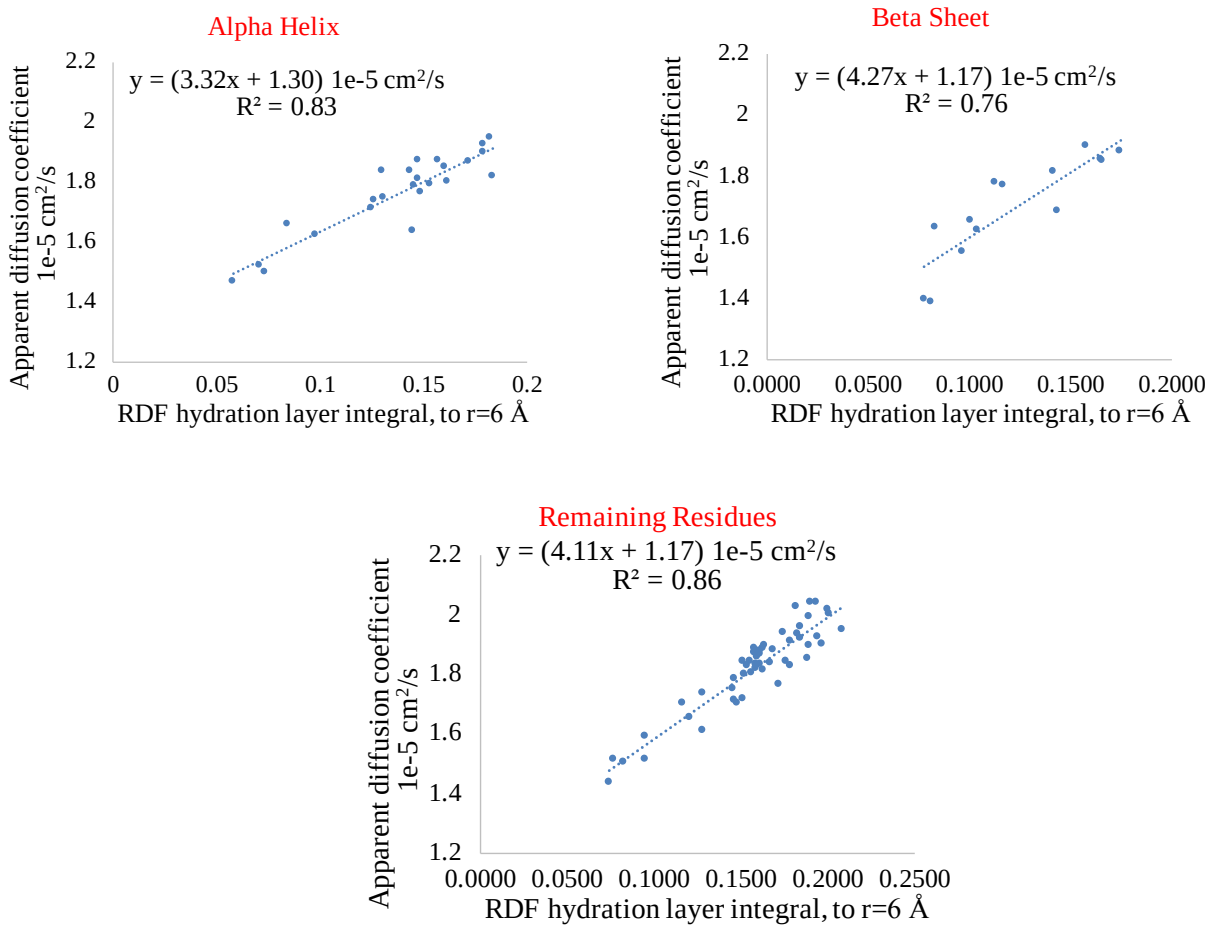


Figure S14. Linear correlation of local water structure/density obtained using integration of the *hydration layer* RDF of water center of mass around protein side chain atoms, with local solvent shell water diffusion coefficients for alpha helices, beta sheets and remining residues (loop/turn regions).

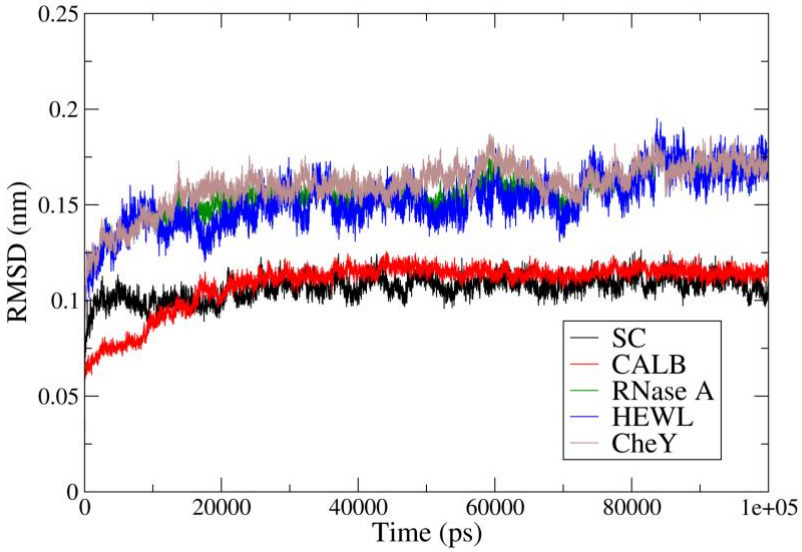


Figure S15. Root mean square deviation (RMSD) for each protein (black-SC, red-CALB, green-RNaseA, blue-HEWL, brown-Chey).

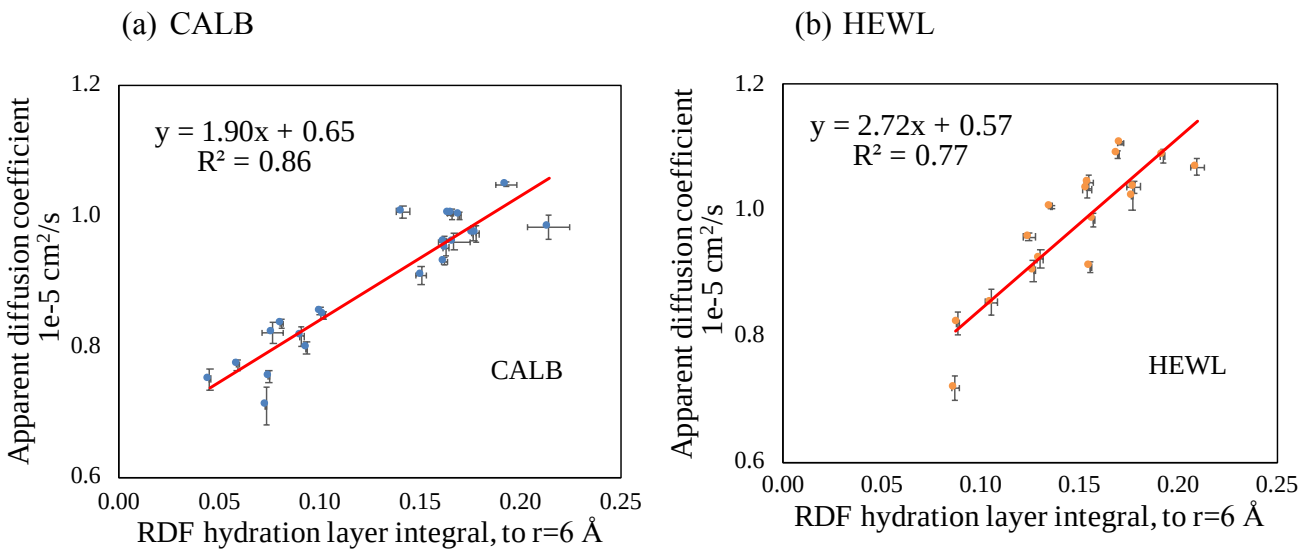


Figure S16. Linear correlation of local water structure obtained using integration of the RDF, with local hydration layer diffusive dynamics for (a) CALB protein and (b) HEWL protein at T=278 K.

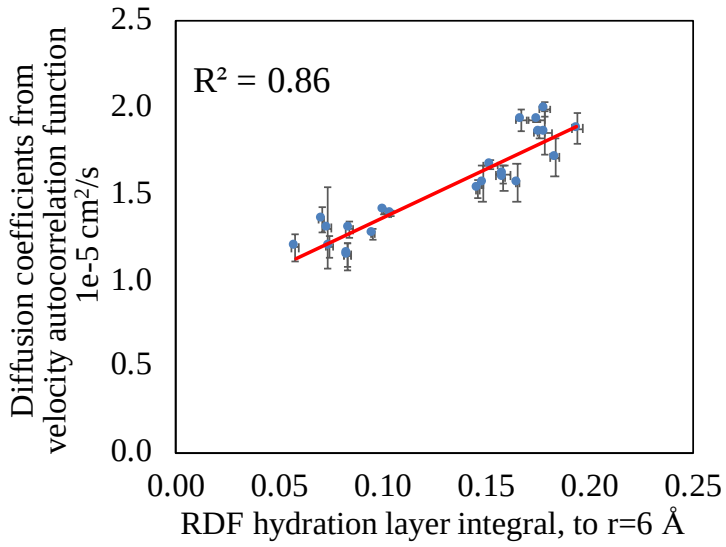


Figure S17. Linear correlation of local water structure/density obtained using integration of the RDF, with local solvent shell water diffusion coefficients calculated using velocity autocorrelation function for CALB protein.

Table S1. Diffusion coefficients of CALB regional solvent shell water molecules calculated from MD simulations.

	Ensemble1				Ensemble2				Ensemble3			
	0-25ns	25-50ns	50-75ns	75-100ns	0-25ns	25-50ns	50-75ns	75-100ns	0-25ns	25-50ns	50-75ns	75-100ns
alpha1	1.88E-05	1.89E-05	1.88E-05	1.88E-05	1.85E-05	1.94E-05	1.93E-05	1.94E-05	1.85E-05	1.92E-05	1.94E-05	1.85E-05
alpha2	1.51E-05	1.53E-05	1.52E-05	1.51E-05	1.49E-05	1.48E-05	1.47E-05	1.51E-05	1.49E-05	1.49E-05	1.48E-05	1.51E-05
alpha3	1.53E-05	1.52E-05	1.53E-05	1.53E-05	1.51E-05	1.51E-05	1.53E-05	1.52E-05	1.51E-05	1.52E-05	1.52E-05	1.52E-05
alpha5	1.82E-05	1.82E-05	1.82E-05	1.82E-05	1.84E-05	1.78E-05	1.86E-05	1.83E-05	1.78E-05	1.79E-05	1.82E-05	1.82E-05
alpha7	1.66E-05	1.65E-05	1.65E-05	1.66E-05	1.61E-05	1.69E-05	1.65E-05	1.67E-05	1.62E-05	1.73E-05	1.66E-05	1.64E-05
alpha9	1.49E-05	1.50E-05	1.49E-05	1.49E-05	1.46E-05	1.45E-05	1.45E-05	1.46E-05	1.46E-05	1.44E-05	1.47E-05	1.45E-05
alpha10	1.77E-05	1.77E-05	1.77E-05	1.76E-05	1.72E-05	1.70E-05	1.70E-05	1.73E-05	1.81E-05	1.81E-05	1.80E-05	1.84E-05
1-12	1.87E-05	1.87E-05	1.86E-05	1.87E-05	1.82E-05	1.83E-05	1.83E-05	1.82E-05	1.84E-05	1.83E-05	1.81E-05	1.83E-05
309-317	2.00E-05	1.95E-05	2.00E-05	1.98E-05	1.97E-05	1.97E-05	1.95E-05	1.92E-05	1.99E-05	1.95E-05	1.94E-05	1.92E-05
20-32	1.88E-05	1.90E-05	1.89E-05	1.89E-05	1.83E-05	1.84E-05	1.84E-05	1.83E-05	1.83E-05	1.85E-05	1.82E-05	1.83E-05
58-66	1.64E-05	1.65E-05	1.65E-05	1.64E-05	1.61E-05	1.62E-05	1.66E-05	1.60E-05	1.61E-05	1.64E-05	1.66E-05	1.60E-05
67-76	1.51E-05	1.50E-05	1.50E-05	1.50E-05	1.49E-05	1.48E-05	1.51E-05	1.52E-05	1.50E-05	1.50E-05	1.51E-05	1.52E-05
94-98	1.94E-05	1.93E-05	1.93E-05	1.93E-05	1.90E-05	1.90E-05	1.95E-05	1.95E-05	1.91E-05	1.91E-05	1.95E-05	1.95E-05
118-124	1.64E-05	1.66E-05	1.65E-05	1.65E-05	1.61E-05	1.59E-05	1.63E-05	1.60E-05	1.62E-05	1.60E-05	1.60E-05	1.61E-05
147-151	1.89E-05	1.88E-05	1.88E-05	1.89E-05	1.86E-05	1.89E-05	1.90E-05	1.89E-05	1.88E-05	1.89E-05	1.89E-05	1.89E-05
170-178	1.62E-05	1.65E-05	1.63E-05	1.63E-05	1.57E-05	1.56E-05	1.58E-05	1.59E-05	1.59E-05	1.54E-05	1.60E-05	1.58E-05
184-211	1.68E-05	1.68E-05	1.69E-05	1.69E-05	1.63E-05	1.64E-05	1.63E-05	1.64E-05	1.63E-05	1.66E-05	1.63E-05	1.65E-05
217-225	1.76E-05	1.74E-05	1.75E-05	1.75E-05	1.72E-05	1.74E-05	1.71E-05	1.78E-05	1.71E-05	1.76E-05	1.79E-05	1.79E-05
243-245	1.76E-05	1.76E-05	1.75E-05	1.76E-05	1.82E-05	1.83E-05	1.83E-05	1.81E-05	1.83E-05	1.81E-05	1.84E-05	1.79E-05
246-248	1.41E-05	1.45E-05	1.43E-05	1.43E-05	1.45E-05	1.39E-05	1.46E-05	1.42E-05	1.40E-05	1.55E-05	1.46E-05	1.39E-05
249-253	1.93E-05	1.92E-05	1.93E-05	1.92E-05	1.91E-05	1.89E-05	1.92E-05	1.91E-05	1.90E-05	1.91E-05	1.93E-05	1.90E-05
254-262	1.85E-05	1.85E-05	1.86E-05	1.86E-05	1.85E-05	1.80E-05	1.81E-05	1.78E-05	1.81E-05	1.83E-05	1.79E-05	1.78E-05
264-267	1.83E-05	1.86E-05	1.84E-05	1.85E-05	1.81E-05	1.85E-05	1.80E-05	1.83E-05	1.87E-05	1.87E-05	1.82E-05	1.83E-05

Table S2. Diffusion coefficients of HEWL regional solvent shell water molecules calculated from MD simulations.

	Ensemble1				Ensemble2				Ensemble3			
	0-25ns	25-50ns	50-75ns	75-100ns	0-25ns	25-50ns	50-75ns	75-100ns	0-25ns	25-50ns	50-75ns	75-100ns
1-7	1.87E-05	1.86E-05	1.87E-05	1.88E-05	1.89E-05	1.89E-05	1.87E-05	1.88E-05	1.88E-05	1.90E-05	1.87E-05	1.88E-05
13-16	1.94E-05	1.92E-05	1.92E-05	1.96E-05	1.93E-05	1.92E-05	1.92E-05	1.93E-05	1.91E-05	1.90E-05	1.93E-05	1.93E-05
18-24	2.00E-05	1.99E-05	1.98E-05	1.98E-05	1.99E-05	1.99E-05	2.00E-05	2.01E-05	2.00E-05	1.99E-05	1.98E-05	2.00E-05
33-35	1.66E-05	1.66E-05	1.62E-05	1.63E-05	1.66E-05	1.61E-05	1.63E-05	1.64E-05	1.63E-05	1.62E-05	1.63E-05	1.62E-05
37-39	1.86E-05	1.83E-05	1.83E-05	1.83E-05	1.86E-05	1.83E-05	1.82E-05	1.85E-05	1.83E-05	1.86E-05	1.81E-05	1.82E-05
41-45	1.88E-05	1.87E-05	1.88E-05	1.92E-05	1.88E-05	1.85E-05	1.80E-05	1.87E-05	1.94E-05	1.88E-05	1.88E-05	1.92E-05
46-49	1.87E-05	1.89E-05	1.92E-05	1.87E-05	1.88E-05	1.91E-05	1.91E-05	1.91E-05	1.92E-05	1.85E-05	1.89E-05	1.92E-05
51-53	1.47E-05	1.39E-05	1.41E-05	1.45E-05	1.39E-05	1.40E-05	1.36E-05	1.31E-05	1.32E-05	1.37E-05	1.38E-05	1.38E-05
65-79	2.05E-05	2.05E-05	2.05E-05	2.05E-05	2.05E-05	2.05E-05	2.07E-05	2.09E-05	1.97E-05	1.96E-05	1.97E-05	1.96E-05
81-83	1.73E-05	1.72E-05	1.72E-05	1.73E-05	1.78E-05	1.74E-05	1.75E-05	1.77E-05	1.76E-05	1.77E-05	1.76E-05	1.74E-05
87-89	1.71E-05	1.71E-05	1.71E-05	1.71E-05	1.70E-05	1.71E-05	1.69E-05	1.73E-05	1.71E-05	1.69E-05	1.71E-05	1.69E-05
100-103	1.97E-05	1.99E-05	1.96E-05	1.94E-05	1.96E-05	1.96E-05	1.91E-05	1.99E-05	1.98E-05	1.95E-05	1.99E-05	1.90E-05
104-108	1.55E-05	1.63E-05	1.64E-05	1.60E-05	1.63E-05	1.64E-05	1.65E-05	1.64E-05	1.61E-05	1.64E-05	1.61E-05	1.63E-05
109-113	1.73E-05	1.71E-05	1.65E-05	1.65E-05	1.74E-05	1.71E-05	1.73E-05	1.73E-05	1.73E-05	1.75E-05	1.69E-05	1.73E-05
114-119	2.08E-05	2.02E-05	2.01E-05	2.01E-05	2.07E-05	2.03E-05	2.05E-05	2.07E-05	2.05E-05	2.05E-05	2.05E-05	2.03E-05
120-123	1.80E-05	1.79E-05	1.80E-05	1.80E-05	1.87E-05	1.85E-05	1.85E-05	1.87E-05	1.86E-05	1.83E-05	1.85E-05	1.86E-05
124-129	1.81E-05	1.80E-05	1.80E-05	1.80E-05	1.81E-05	1.81E-05	1.81E-05	1.81E-05	1.80E-05	1.81E-05	1.79E-05	1.83E-05

Table S3. Diffusion coefficients of SC regional solvent shell water molecules calculated from MD simulations.

	Ensemble 1				Ensemble 2				Ensemble 3			
	0-25 ns	25-50 ns	50-75 ns	75-100 ns	0-25 ns	25-50 ns	50-75 ns	75-100 ns	0-25 ns	25-50 ns	50-75 ns	75-100 ns
1-6	1.74E-05	1.76E-05	1.79E-05	1.79E-05	1.76E-05	1.80E-05	1.78E-05	1.82E-05	1.80E-05	1.79E-05	1.82E-05	1.79E-05
14-18	1.79E-05	1.75E-05	1.79E-05	1.78E-05	1.76E-05	1.80E-05	1.79E-05	1.82E-05	1.79E-05	1.80E-05	1.81E-05	1.79E-05
19-22	1.81E-05	1.81E-05	1.83E-05	1.80E-05	1.80E-05	1.82E-05	1.81E-05	1.83E-05	1.81E-05	1.81E-05	1.82E-05	1.83E-05
104-109	1.68E-05	1.67E-05	1.67E-05	1.68E-05	1.66E-05	1.68E-05	1.61E-05	1.63E-05	1.66E-05	1.63E-05	1.62E-05	1.65E-05
116-120	1.73E-05	1.73E-05	1.74E-05	1.73E-05	1.72E-05	1.73E-05	1.68E-05	1.69E-05	1.72E-05	1.73E-05	1.67E-05	1.69E-05
125-132	1.81E-05	1.81E-05	1.80E-05	1.81E-05	1.83E-05	1.85E-05	1.86E-05	1.81E-05	1.83E-05	1.86E-05	1.86E-05	1.80E-05
133-138	1.87E-05	1.83E-05	1.85E-05	1.86E-05	1.83E-05	1.84E-05	1.86E-05	1.86E-05	1.82E-05	1.86E-05	1.85E-05	1.87E-05
143-146	1.88E-05	1.87E-05	1.88E-05	1.85E-05	1.85E-05	1.83E-05	1.87E-05	1.84E-05	1.84E-05	1.86E-05	1.86E-05	1.83E-05
155-167	1.83E-05	1.85E-05	1.88E-05	1.84E-05	1.85E-05	1.86E-05	1.87E-05	1.86E-05	1.84E-05	1.86E-05	1.90E-05	1.84E-05
170-173	1.73E-05	1.73E-05	1.76E-05	1.71E-05	1.75E-05	1.67E-05	1.76E-05	1.67E-05	1.74E-05	1.67E-05	1.67E-05	1.76E-05
181-189	1.90E-05	1.92E-05	1.92E-05	1.90E-05	1.90E-05	1.87E-05	1.91E-05	1.88E-05	1.89E-05	1.89E-05	1.90E-05	1.87E-05
194-197	1.62E-05	1.62E-05	1.62E-05	1.63E-05	1.61E-05	1.61E-05	1.59E-05	1.62E-05	1.60E-05	1.60E-05	1.61E-05	1.63E-05
209-211	1.75E-05	1.70E-05	1.70E-05	1.78E-05	1.71E-05	1.71E-05	1.72E-05	1.77E-05	1.72E-05	1.80E-05	1.71E-05	1.78E-05
212-218	1.82E-05	1.82E-05	1.82E-05	1.81E-05	1.80E-05	1.80E-05	1.84E-05	1.81E-05	1.80E-05	1.81E-05	1.83E-05	1.81E-05
235-237	1.55E-05	1.51E-05	1.52E-05	1.53E-05	1.49E-05	1.49E-05	1.53E-05	1.47E-05	1.48E-05	1.48E-05	1.56E-05	1.55E-05
238-245	1.87E-05	1.83E-05	1.88E-05	1.83E-05	1.85E-05	1.83E-05	1.81E-05	1.83E-05	1.85E-05	1.82E-05	1.82E-05	1.82E-05
24-27	1.77E-05	1.73E-05	1.77E-05	1.77E-05	1.75E-05	1.78E-05	1.76E-05	1.78E-05	1.76E-05	1.78E-05	1.76E-05	1.78E-05
247-249	1.79E-05	1.78E-05	1.81E-05	1.78E-05	1.82E-05	1.81E-05	1.78E-05	1.81E-05	1.82E-05	1.79E-05	1.80E-05	1.80E-05
252-259	1.86E-05	1.86E-05	1.82E-05	1.80E-05	1.82E-05	1.81E-05	1.88E-05	1.89E-05	1.82E-05	1.89E-05	1.80E-05	1.88E-05
260-263	1.81E-05	1.81E-05	1.84E-05	1.81E-05	1.85E-05	1.85E-05	1.83E-05	1.84E-05	1.84E-05	1.85E-05	1.85E-05	1.83E-05
36-43	1.73E-05	1.66E-05	1.71E-05	1.69E-05	1.69E-05	1.70E-05	1.72E-05	1.70E-05	1.70E-05	1.71E-05	1.72E-05	1.70E-05
44-49	1.54E-05	1.60E-05	1.58E-05	1.56E-05	1.51E-05	1.57E-05	1.57E-05	1.52E-05	1.53E-05	1.57E-05	1.54E-05	1.54E-05
50-62	1.88E-05	1.84E-05	1.87E-05	1.88E-05	1.87E-05	1.90E-05	1.88E-05	1.84E-05	1.88E-05	1.88E-05	1.86E-05	1.85E-05
75-80	1.92E-05	1.98E-05	1.93E-05	1.91E-05	1.91E-05	1.90E-05	1.91E-05	1.93E-05	1.91E-05	1.90E-05	1.93E-05	1.93E-05
86-89	1.52E-05	1.49E-05	1.54E-05	1.47E-05	1.51E-05	1.55E-05	1.46E-05	1.55E-05	1.53E-05	1.47E-05	1.55E-05	1.54E-05
96-103	1.82E-05	1.89E-05	1.82E-05	1.82E-05	1.83E-05	1.86E-05	1.82E-05	1.80E-05	1.84E-05	1.83E-05	1.82E-05	1.79E-05

Table S4. Diffusion coefficients of RNase A regional solvent shell water molecules calculated from MD simulations.

	Ensemble 1				Ensemble 2				Ensemble 3			
	0-25 ns	25-50 ns	50-75 ns	75-100 ns	0-25 ns	25-50 ns	50-75 ns	75-100 ns	0-25 ns	25-50 ns	50-75 ns	75-100 ns
1-5	1.88E-05	1.95E-05	1.92E-05	1.90E-05	1.92E-05	1.89E-05	1.87E-05	1.94E-05	1.89E-05	1.92E-05	1.88E-05	1.88E-05
113-115	2.07E-05	1.98E-05	2.06E-05	1.98E-05	2.13E-05	2.03E-05	1.97E-05	1.99E-05	2.05E-05	2.05E-05	1.97E-05	1.95E-05
12-22	1.88E-05	1.85E-05	1.88E-05	1.87E-05	1.90E-05	1.89E-05	1.85E-05	1.84E-05	1.87E-05	1.90E-05	1.87E-05	1.86E-05
27-29	1.72E-05	1.78E-05	1.73E-05	1.71E-05	1.70E-05	1.72E-05	1.76E-05	1.70E-05	1.72E-05	1.78E-05	1.78E-05	1.75E-05
49-53	1.94E-05	1.92E-05	1.96E-05	1.96E-05	1.96E-05	1.92E-05	1.92E-05	1.91E-05	1.96E-05	1.92E-05	1.91E-05	1.94E-05
67-71	2.00E-05	2.06E-05	2.06E-05	2.05E-05	1.94E-05	2.04E-05	2.07E-05	2.06E-05	2.06E-05	2.07E-05	2.04E-05	2.05E-05
76-78	2.01E-05	2.01E-05	1.99E-05	2.00E-05	2.01E-05	2.02E-05	2.01E-05	2.01E-05	1.99E-05	2.01E-05	2.02E-05	1.97E-05
100-105	1.85E-05	1.85E-05	1.85E-05	1.85E-05	1.83E-05	1.85E-05	1.86E-05	1.81E-05	1.87E-05	1.84E-05	1.84E-05	1.86E-05
116-120	1.44E-05	1.42E-05	1.38E-05	1.43E-05	1.39E-05	1.35E-05	1.42E-05	1.40E-05	1.37E-05	1.40E-05	1.41E-05	1.35E-05
122-124	1.66E-05	1.69E-05	1.68E-05	1.67E-05	1.75E-05	1.74E-05	1.70E-05	1.71E-05	1.65E-05	1.68E-05	1.68E-05	1.63E-05
31-43	1.88E-05	1.85E-05	1.92E-05	1.90E-05	1.82E-05	1.89E-05	1.87E-05	1.84E-05	1.89E-05	1.92E-05	1.88E-05	1.88E-05
59-63	1.98E-05	1.97E-05	1.92E-05	1.94E-05	1.89E-05	1.91E-05	1.87E-05	1.85E-05	1.83E-05	1.82E-05	1.84E-05	1.83E-05
85-96	1.96E-05	1.99E-05	1.96E-05	1.99E-05	1.94E-05	1.89E-05	1.90E-05	1.87E-05	1.96E-05	1.97E-05	1.96E-05	1.91E-05

Table S5. Diffusion coefficients of CheY regional solvent shell water molecules calculated from MD simulations.

	Ensemble 1				Ensemble 2				Ensemble 3			
	0-25 ns	25-50 ns	50-75 ns	75-100 ns	0-25 ns	25-50 ns	50-75 ns	75-100 ns	0-25 ns	25-50 ns	50-75 ns	75-100 ns
3-6	1.85E-05	1.86E-05	1.93E-05	1.94E-05	1.85E-05	1.85E-05	1.93E-05	1.93E-05	1.93E-05	1.92E-05	1.88E-05	1.91E-05
12-22	1.78E-05	1.80E-05	1.80E-05	1.78E-05	1.80E-05	1.79E-05	1.81E-05	1.78E-05	1.78E-05	1.78E-05	1.78E-05	1.78E-05
25-29	1.91E-05	1.89E-05	1.87E-05	1.89E-05	1.88E-05	1.88E-05	1.89E-05	1.89E-05	1.87E-05	1.90E-05	1.87E-05	1.86E-05
31-38	1.78E-05	1.76E-05	1.76E-05	1.79E-05	1.78E-05	1.78E-05	1.78E-05	1.79E-05	1.78E-05	1.78E-05	1.78E-05	1.78E-05
43-45	1.82E-05	1.79E-05	1.81E-05	1.79E-05	1.80E-05	1.81E-05	1.81E-05	1.79E-05	1.83E-05	1.82E-05	1.81E-05	1.83E-05
47-50	1.95E-05	1.95E-05	1.94E-05	1.95E-05	1.94E-05	1.95E-05	1.96E-05	1.93E-05	1.96E-05	1.97E-05	1.94E-05	1.95E-05
59-64	1.83E-05	1.86E-05	1.84E-05	1.84E-05	1.87E-05	1.84E-05	1.84E-05	1.85E-05	1.85E-05	1.86E-05	1.82E-05	1.85E-05
70-78	1.87E-05	1.87E-05	1.87E-05	1.86E-05	1.87E-05	1.86E-05	1.87E-05	1.87E-05	1.87E-05	1.87E-05	1.87E-05	1.86E-05
90-95	1.94E-05	1.96E-05	1.95E-05	1.96E-05	1.95E-05	1.95E-05	1.95E-05	1.92E-05	1.92E-05	1.95E-05	1.94E-05	1.95E-05
100-104	1.77E-05	1.76E-05	1.80E-05	1.76E-05	1.76E-05	1.78E-05	1.79E-05	1.76E-05	1.75E-05	1.78E-05	1.78E-05	1.73E-05
106-113	1.82E-05	1.84E-05	1.83E-05	1.82E-05	1.83E-05	1.84E-05	1.83E-05	1.82E-05	1.84E-05	1.82E-05	1.85E-05	1.84E-05
115-117	1.85E-05	1.86E-05	1.84E-05	1.82E-05	1.84E-05	1.86E-05	1.85E-05	1.82E-05	1.83E-05	1.82E-05	1.84E-05	1.83E-05
119-123	1.87E-05	1.87E-05	1.88E-05	1.86E-05	1.86E-05	1.88E-05	1.87E-05	1.86E-05	1.87E-05	1.87E-05	1.87E-05	1.87E-05

Table S6. Comparison of protein regional solvent shell water diffusion coefficients diffusion coefficients produced by the developed linear regression model using RDF hydration layer integral values, compared with diffusion coefficients calculated directly from MD simulations. Values are in units 10^{-5} cm^2/s.

Protein	Regions	Diffusion from MD	Diffusion from Model	Difference
HEWL	1-7	1.8791	1.8368	0.04224
	13-16	1.9256	1.9038	0.02179
	18-24	1.9916	1.9448	0.04679
	33-35	1.6336	1.7730	0.13940
	37-39	1.8346	1.8284	0.00620
	41-45	1.8814	1.8878	0.00639
	46-49	1.8948	1.9458	0.05101
	51-53	1.3855	1.5314	0.14589
	65-79	2.0261	1.9145	0.11152
	81-83	1.7470	1.7192	0.02773
	87-89	1.7049	1.7876	0.08264
	100-103	1.9596	1.9248	0.03473
	104-108	1.6221	1.5929	0.02919
	109-113	1.7110	1.6974	0.01361
	114-119	2.0429	1.9461	0.09678
	120-123	1.8365	1.7166	0.11988
	124-129	1.8059	1.8164	0.01049
CALB	alpha1	1.8970	1.9023	0.00534
	alpha2	1.4992	1.5005	0.00127
	alpha3	1.5194	1.4900	0.02939
	alpha5	1.8169	1.9209	0.10400
	alpha7	1.6567	1.5403	0.11638
	alpha9	1.4660	1.4409	0.02507
	alpha10	1.7648	1.7871	0.02229
	1-12	1.8415	1.8928	0.05128
	309-317	1.9609	1.8872	0.07371
	20-32	1.8521	1.8505	0.00164
	58-66	1.6325	1.5382	0.09429
	67-75	1.5037	1.5379	0.03421
	94-98	1.9283	1.9617	0.03346
	118-124	1.6219	1.6163	0.00561
	147-151	1.8862	1.8249	0.06131
	170-178	1.5938	1.5837	0.01012
	184-211	1.6536	1.6035	0.05006
	217-225	1.7500	1.7778	0.02783
	243-245	1.7991	1.8023	0.00314
	246-248	1.4374	1.5036	0.06625
	249-253	1.9136	1.9017	0.01193
	254-262	1.8210	1.8260	0.00502
	264-267	1.8380	1.8584	0.02045
Subtilisin	1 6	1.7872	1.7792	0.00799
	14-18	1.7903	1.8056	0.01531

	19-22	1.8145	1.8443	0.02984
	104-109	1.6527	1.6828	0.03012
	116-120	1.7136	1.7806	0.06702
	125-132	1.8269	1.8076	0.01927
	133-138	1.8487	1.8315	0.01719
	143-146	1.8543	1.9420	0.08770
	155-167	1.8563	1.8299	0.02637
	170-173	1.7185	1.7977	0.07915
	181-189	1.8959	1.8466	0.04936
	194-197	1.6137	1.7092	0.09547
	209-211	1.7379	1.7111	0.02676
	212-218	1.8133	1.7599	0.05340
	235-237	1.5133	1.5135	0.00016
	238-245	1.8358	1.8353	0.00052
	24-27	1.7665	1.8777	0.11122
	247-249	1.7988	1.8364	0.03768
	252-259	1.8449	1.8142	0.03072
	260-263	1.8339	1.7681	0.06576
	36-43	1.7029	1.6655	0.03741
	44-49	1.5529	1.5893	0.03637
	50-62	1.8696	1.8380	0.03162
	75-80	1.9218	1.9261	0.00426
	86-89	1.5164	1.5831	0.06667
	96-103	1.8285	1.9047	0.07619
RNase A	1-6	1.9025	1.9736	0.07105
	113-115	2.0184	1.9850	0.03333
	12-22	1.8715	1.8232	0.04830
	27-29	1.7375	1.7025	0.03501
	49-53	1.9358	1.9183	0.01747
	67-71	2.0410	1.9592	0.08178
	76-78	2.0052	1.9882	0.01697
	100-105	1.8472	1.8544	0.00720
	116-120	1.3971	1.5164	0.11930
	122-124	1.6866	1.7693	0.08275
	31-43	1.8775	1.8325	0.04504
	59-63	1.8881	1.8446	0.04348
	85-96	1.9403	1.8879	0.05242
CheY	3-6	1.8988	1.8236	0.07522
	12-22	1.7880	1.7762	0.01178
	25-29	1.8832	1.8667	0.01650
	31-38	1.7785	1.6513	0.12720
	43-45	1.8093	1.7831	0.02628
	47-50	1.9491	2.0180	0.06890
	59-64	1.8456	1.7983	0.04726
	70-78	1.8682	1.8747	0.00652
	90-95	1.9450	1.9160	0.02902
	100-104	1.7678	1.6669	0.10084
	106-113	1.8312	1.8297	0.00152
	115-117	1.8381	1.7827	0.05538
	119-123	1.8695	1.8182	0.05132

Table S7. Protein simulation system details.

Protein	Box Size (nm^3)	Number of Water Molecules	Number of protein atoms (after adding hydrogen)	Time required for the equilibration
CheY	6.56 x 6.56 x 6.56	8792	2308	5 ns
CALB	7.91 x 7.91 x 7.91	15043	4624	8 ns
HEWL	7.40 x 7.40 x 7.40	12840	1961	5 ns
SC	7.27 x 7.27 x 7.27	11615	3805	2 ns
RNaseA	7.04 x 7.04 x 7.04	11021	1856	5 ns

REFERENCES

1. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A linear constraint solver for molecular simulations. *J Comput Chem* **1997,** *18* (12), 1463-1472.
2. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *J Chem Phys* **1993,** *98* (12), 10089-10092.
3. Arfken, G. B.; Weber, H. J., Mathematical methods for physicists. AAPT: 1999.
4. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J Chem Phys* **2007,** *126* (1).
5. Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., Molecular-Dynamics with Coupling to an External Bath. *J Chem Phys* **1984,** *81* (8), 3684-3690.
6. Nose, S., A Unified Formulation of the Constant Temperature Molecular-Dynamics Methods. *J Chem Phys* **1984,** *81* (1), 511-519.
7. Hoover, W. G., Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A Gen Phys* **1985,** *31* (3), 1695-1697.
8. Kasireddy, C.; Ellis, J. M.; Bann, J. G.; Mitchell-Koch, K. R., Tautomeric stabilities of 4-fluorohistidine shed new light on mechanistic experiments with labeled ribonuclease A. *Chem Phys Lett* **2016,** *666*, 58-61.
9. Dahanayake, J. N.; Gautam, D. N.; Verma, R.; Mitchell-Koch, K. R., To Keep or Not to Keep? The Question of Crystallographic Waters for Enzyme Simulations in Organic Solvent. *Mol Simul* **2016,** *42* (12), 1001-1013.
10. Chen, C. R.; Makhatadze, G. I., ProteinVolume: calculating molecular van der Waals and void volumes in proteins. *BMC Bioinformatics* **2015,** *16*, 101.