Molecular modelling of the mass density of single proteins

Meike Hutt, Tobias Kulschewski, Jürgen Pleiss

Supplementary Material

Table SI: Apparent specific volumes of all amino acid residues at 298 K (20)

Amino acid residue	Apparent specific volume [cm³/g]
Alanine	0.74
Arginine	0.70
Asparagine	0.62
Aspartic acid	0.60
Cysteine	0.63
Glutamic acid	0.66
Glutamine	0.67
Glycine	0.64
Histidine	0.67
Isoleucine	0.90
Leucine	0.90
Lysine	0.82
Methionine	0.75
Phenylalanine	0.77
Proline	0.76
Serine	0.63
Threonine	0.70
Tryptophan	0.74
Tyrosine	0.71
Valine	0.86

Table SII: Sensitivity of the simulated density towards changes in the simulation parameters (temperature coupling, pressure coupling, water model, or protein force field)

Parameters	Simulated density ρ_{sim} [g/cm³]	Difference to the simulation with the method used in this work in %	Difference to experimental density in %
Method used in this work	1.381	-	-1.4
Temperature coupling: v-rescale	1.364(0)	-1.2	-2.6
Pressure coupling: Parrinello-Rahman	1.364(3)	-1.2	-2.6
Water model: TIP4P	1.383	0.17	-1.2
Protein force field: GROMOS	1.387	0.41	-0.95

Figure S1: Plot of the solution density of 8 mg/cm³ *Candida antarctica* lipase B (1 *Candida antarctica* lipase B in a water box) at 298 K against the simulated time

Figure S2: Plot of the solution density against the protein concentration for *Candida antarctica* lipase B at 298 K for one protein per box (■; ...; $R^2=0.9999$) and for eight proteins per box (◆; —; $R^2=0.9999$)

Figure S3: Plot of the solution density against the protein concentration for ubiquitin at 298 K for the structure without crystal water (♦; —; $R^2=0.9999$) and for the structure with crystal water (■; ...; $R^2=1$)

Figure S4: Plot of the solution density against the protein concentration for ribonuclease-A at 293 K for the protein (◆; $R^2=0.9998$) and for the electrolyte solution (■)

Figure S5: Plot of the solution density against the protein concentration for lysozyme at 298 K for the protein (♦; $R^2=1$) and for the electrolyte solution (■)

Figure S6: Plot of the solution density against the protein concentration for adenylate kinase at 293 K for the protein (♦; $R^2=0.9999$) and for the electrolyte solution (■)

Figure S7: Plot of the solution density against the protein concentration for *Pseudomonas glumae* lipase at 298 K ($R^2=0.9999$)

Figure S8: Plot of the solution density against the protein concentration for malate dehydrogenase at 298 K for the monomer (♦; —; $R^2=0.9999$) and dimer (■; ...; $R^2=0.9999$)

Figure S9: Plot of the thermal expansion coefficient against the protein density at 298 K

Figure S10: Plot of the thermal expansion coefficient against the secondary structure content (♦ helix, □ strand, Δ helix + strand)

Figure S11: Plot of the secondary structure content against the temperature for scorpion toxin variant-3 (◆), ribonuclease-A (■), lysozyme (▲), *Pseudomonas glumae* lipase (○), and *Candida antarctica* lipase B (□)

Figure S12: Plot of the thermal expansion coefficient against the median B-factor at 298 K