Supporting Material: Prediction of hydrodynamic and other solution properties of rigid proteins from atomic and residue-level models

A. Ortega, D. Amorós, and J. García de la Torre

Departamento de Química Física, Facultad de Química,
Universidad de Murcia, 30071 Murcia, Spain,

June 23, 2011

Phone : Intnl - 34 968 367426 ; Fax: Intnl - 34 68 364148 ; e-mail: jgt@um.es

1 Sets of proteins

Table S-1 lists the sets of proteins used in this work. The GT set, first section, is completed with other (referenced) data compilations to construct the WHOLE set.

PDB	M (Da)	a_T (Å)	a_I (Å)	a_G (Å)	a_R (Å)
Set GT [1, 2]					
4PTI	6158	16.64			15.69
1RBX	13700	20.06	19.28	19.11	
6LYZ	14320	19.69	18.96	18.90	20.01
1MBO	17190	19.87	20.69	21.30	21.28
2CGA	25660	23.08	22.50	23.37	
1BEB	36730	27.44		27.89	27.79
1OVA	43500	26.96	28.90		
1CTS	97938	37.00	39.44	37.57	
4GPD	142868	42.92	42.76	41.44	
6LDH	145169	42.50	44.39	44.80	
1ADO	156000	48.23	48.10		
2MIN	220000	53.65			
1SVN	26700				22.88
1BVG	21580				23.36
1LKI	19100				24.32
6I1B	17400				22.88
1STN	15510				23.42
1HWA	14320				20.01
1WRT	11890				28.15
1BTA	10140				19.26
1UBQ	8540				17.34
1CLB	8430				16.79
2BCA	8430				17.01
1EGL	8150				18.16
1PIT	6160				16.20
1ZNF	2930				13.23
Data from ref. [3]					
1HRC	12400	18.50	17.53		
7RSA	13700	20.09	19.47		
1HFX	14200	20.27	19.32		
2CDS	14300	19.69	18.58		
1MBO	17200	20.54	20.70		
1AVU	20100	21.90	20.74		
1TPO	23300	23.08	22.54		
1TGN	24000	22.17	22.42		
4CHA	25200	21.04	22.89		
2CGA	25700	23.25	22.55		
2CAB	28800	24.14	24.50		
1ZAG	32600	33.02	29.57		
4PEP	34500	25.67	26.02		
1J6Z	43000	28.54	29.33		
6TAA	52500	29.12	30.17		
1AO6	66500	35.18	35.60		
1OVT	76000	36.37	35.78		
1LFG	77100	38.32	36.57		
2SOD	32500	25.95	25.72		
1BEB	36700	29.40	28.07		
4CHA	50400	29.60	32.19		
1GKB	51400	34.61	32.21		
8TIM	53200	31.75	31.63		
2AAI	61500	35.77	30.68		
1HHO	63200	31.65	30.28		
1ALK	94600	37.65	37.08		
1CTS	98000	37.00	39.45		
1FAJ	117300	37.65	42.06		
1ADO	157100	47.17	45.61		
4BLC	235700	52.34	52.63		
1BGL	465800	68.57	65.36		
Data from ref. [4]					
8RAT	13682	18.50	19.27		19.81
1A4V	15793	19.69			21.51
1DWR	17521	20.06			21.30
2CGA	25666	22.59	23.55		
1BEB	36608	27.34			28.19
1OVA	43157	27.76	30.14		27.23
1HCO	64557	31.10			
1GZX	64573	29.77	31.87		32.46
1CTS	97838	37.00	39.43		
2GD1	143540	42.92			
1GD1	146431	40.49	43.11		
5LDH	148636	42.41			
1ADO	157122		44.89		
4BLC	235762	52.34	52.63		
N		54	47	8	23

Table S-1: GT and WHOLE sets of proteins

2 Large proteins and macromolecular complexes

Figure S-1 displays the structures of some of the large proteins and macromolecular complexes:

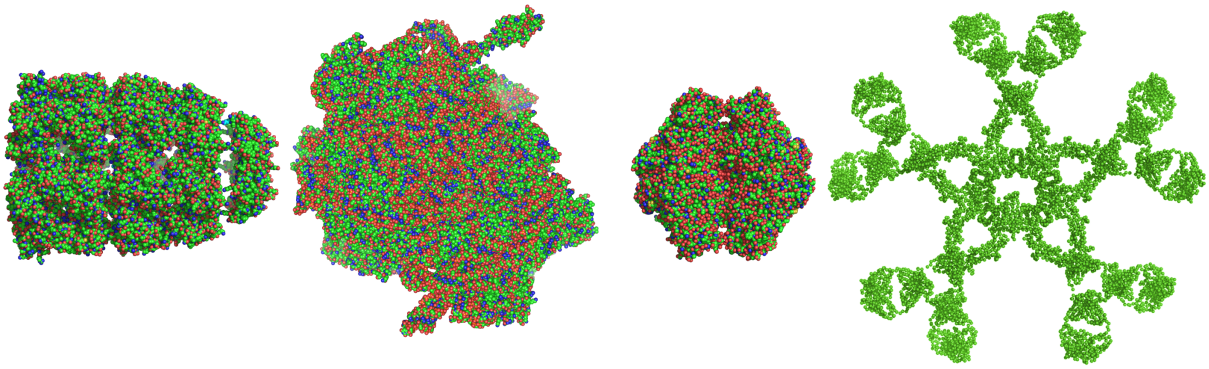

Figure S-1: From left to right: GroEL, Ribosome 70S, full urease (atomic models) and IgM antibody (Cα-only model)

Table S-2 contains the experimental data:

Protein	M (kDa)	PDB	$\bar{v}$ (cm^3/g)	Rg (Å)	Dt $\times 10^7$ (cm^2/s)	$[\eta]$ (cm^3/g)	s (S)
GroEL	802.6 [5]	2CGT [6]	0.747 [5]	67 [7]	2.59 [5]		22.13 [5]
Urease full	480.0 [8]	3LA4 [9]	0.73 [8]		3.46 [8]		18.6 [8]
Urease half	240.0 [10]	3LA4 [9]	0.73 [8]				11.5 [10]
Ig M	950.0 [11]	2RCJ [12]	0.722 [13]	121 [14]	1.82 [13]	13.4 [11]	17.5 [13]
Ribosome 30S	900.0 [15]	2AVY [16]	0.639 [17]	69 [15]	2.18 [18]	8.1 [15]	31.8 [15]
Ribosome 50S	1550.0 [15]	2AW4 [16]	0.639 [17]	77 [15]	1.90 [18]	5.6 [15]	50.2 [15]
Ribosome 70S	2500.0 [17]	2AVY & 2AW4 [16]	0.639 [17]	91.5 [19]	1.72 [17]	5.8 [20]	66.7 [17]

Table S-2: Structures and experimental data for the large proteins and macromolecular complexes

3 Comments of the intrinsic viscosity of ribosomal structures

It is necessary to comment that the differences observed in the intrinsic viscosity between the values calculated in this study from the crystalographic structures and the experimental data (which can be seen in Table II) in the particular case of *E. coli* 70S, 50S and 30S ribosomes arise from an anomalous value of the experimental property expected for this globular riboprotein, as previously commented by some authors [15, 17, 20]. As the authors themselves mention, due to the purification processes, since these particles may still contain some non-ribosomal RNA and protein, the reasons for such a difference might be attributed to a folding-in or release of some of this material. The evidence reported in those papers would indicate that this difference in viscosity is also, at least partially, due to dimers and higher aggregates being formed in the ribosomal samples.

4 CPU time benchmarks

HYDROPRO has been entirely rewritten in order to take advantage of advanced features of Fortran 90/95 compilers, such as dynamic memory allocation, which does not fix the size of arrays (and therefore the numbers of beads or minibeads), and remarkable efficiency in computation. The latest aspect is clearly illustrated by the following Table, which reports CPU times (t_{CPU}) in inexpensive conventional equipments (laptop, personal computer and workstations bought in 2009 or 2010) of the new HYDROPRO version 10, compared to those to the previously released version 7. (Runs were done under Windows XP^{TM} or Windows 7^{TM} , for a shell model calculation with 6 values of the number of minibeads, in the range $\approx$ 400-2000).

Machine	Processors	t_{CPU} v7 (s)	t_{CPU} v10 (s)
Hewlett Packard ^{TM} G62	One Intel Core 2 Duo T7500	190	30
DELL Optiplex ^{TM} 960	One Intel Core 2 Quad Q9550	141	13
DELL Optiplex ^{TM} 960	Two Intel Xeon x5660	110	8

Typical CPU times with version 7 (2007) in comparable computers of a few years ago was over 3 minutes, about 200 seconds. As indicated, the above mentioned benchmarks are for the HYDROPRO classical shell-model calculation. In the bead-model calculation of a residue-level model, CPU time depends roughly on the third power of the number of elements, i. e. residues. In this case, the calculation requires smaller CPU time than a shell-model calculation of 2000 minibeads when $N_r < 2000$, i. e. for proteins with M below 200 KDa. Beyond that limit, the shell-model calculation requires less time than a bead model calculations with one bead per residue.

5 Graphical user interfaces

The two following illustrations (Figures S-2 and S-3) show the new graphical user interface (GUI) of HYDROPRO, with the input data, before the calculation, and with the results, after the calculation.

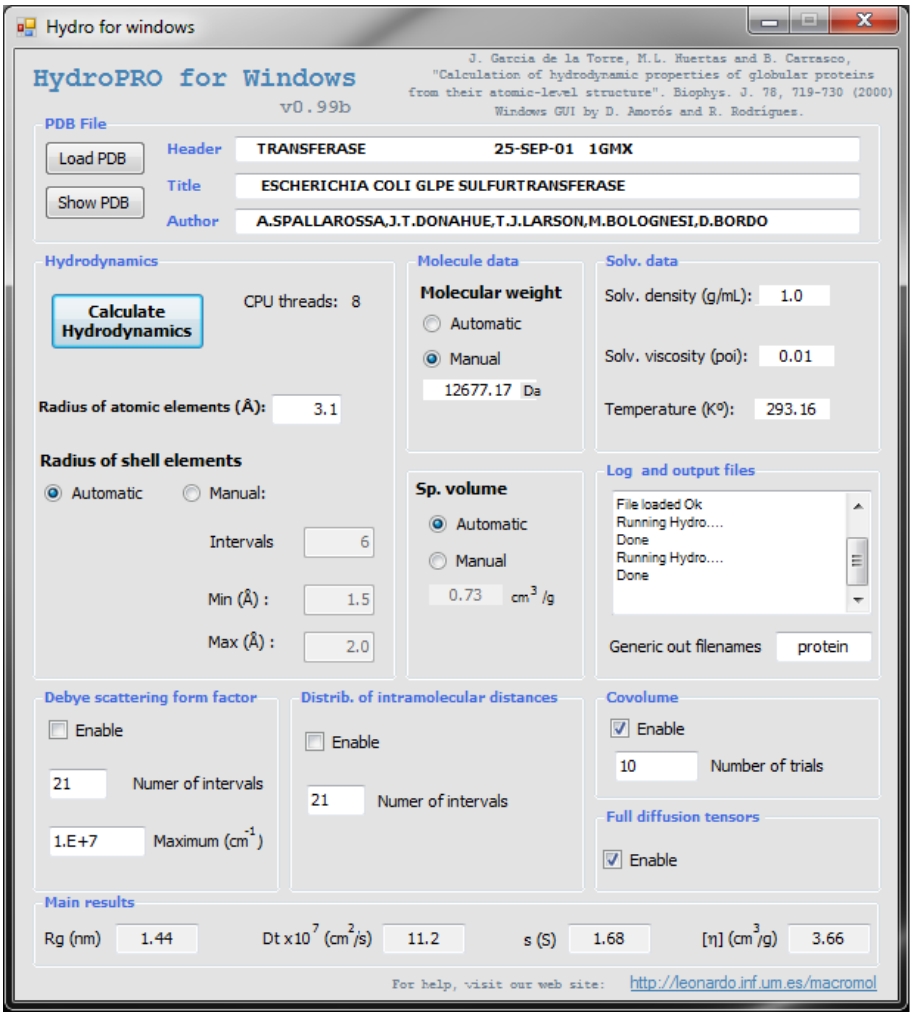

Figure S-2: HYDROPRO GUI with input data

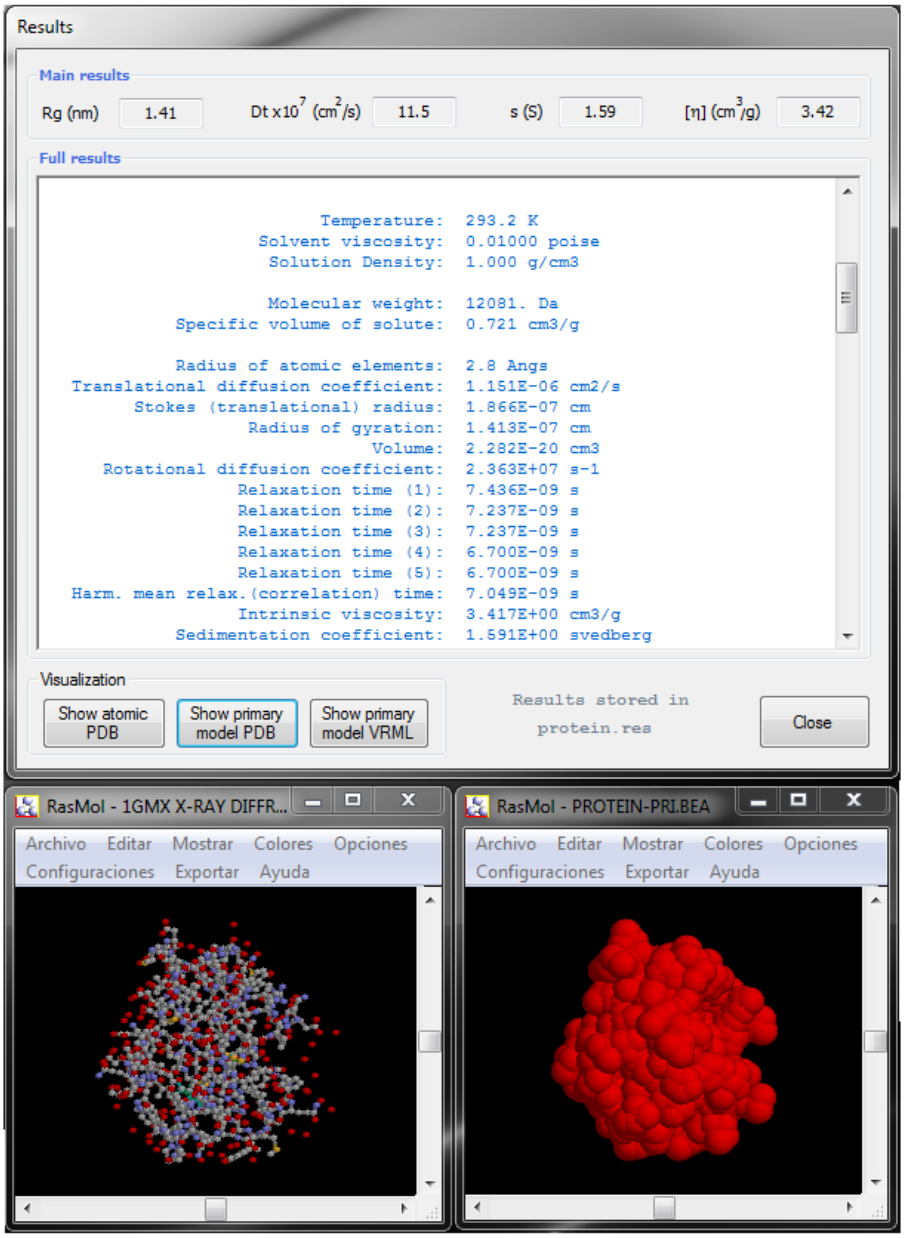

Figure S-3: HYDROPRO GUI. The results from a typical calculation are shown

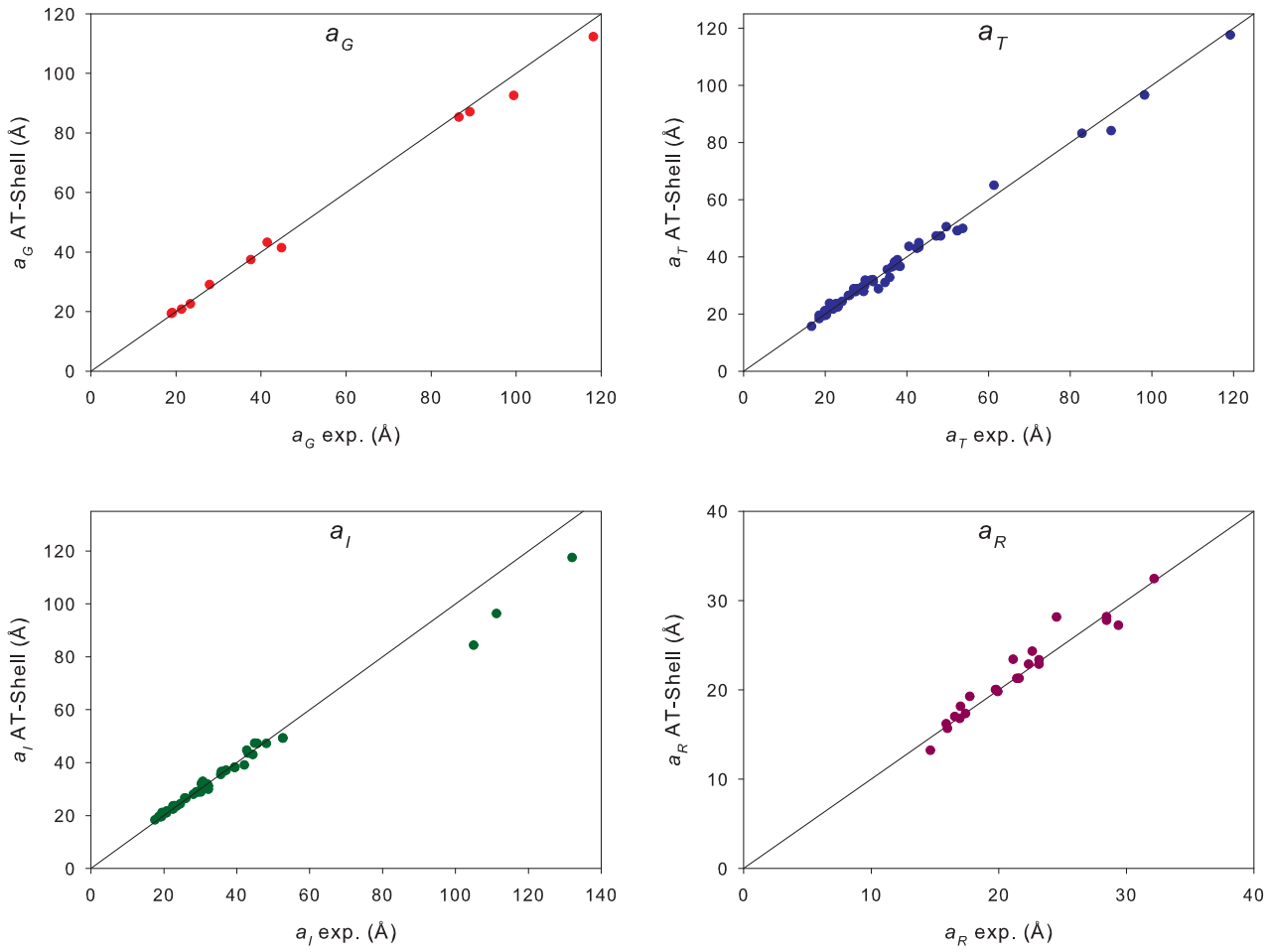

Figure S-4: Representation of the values calculated by HYDROPRO shell-model from the atomic coordinates (in y axis) vs. experimental values (x axis) of the four equivalent radii employed in this work, for the WHOLE set of proteins and including also the large proteins and macromolecular complexes

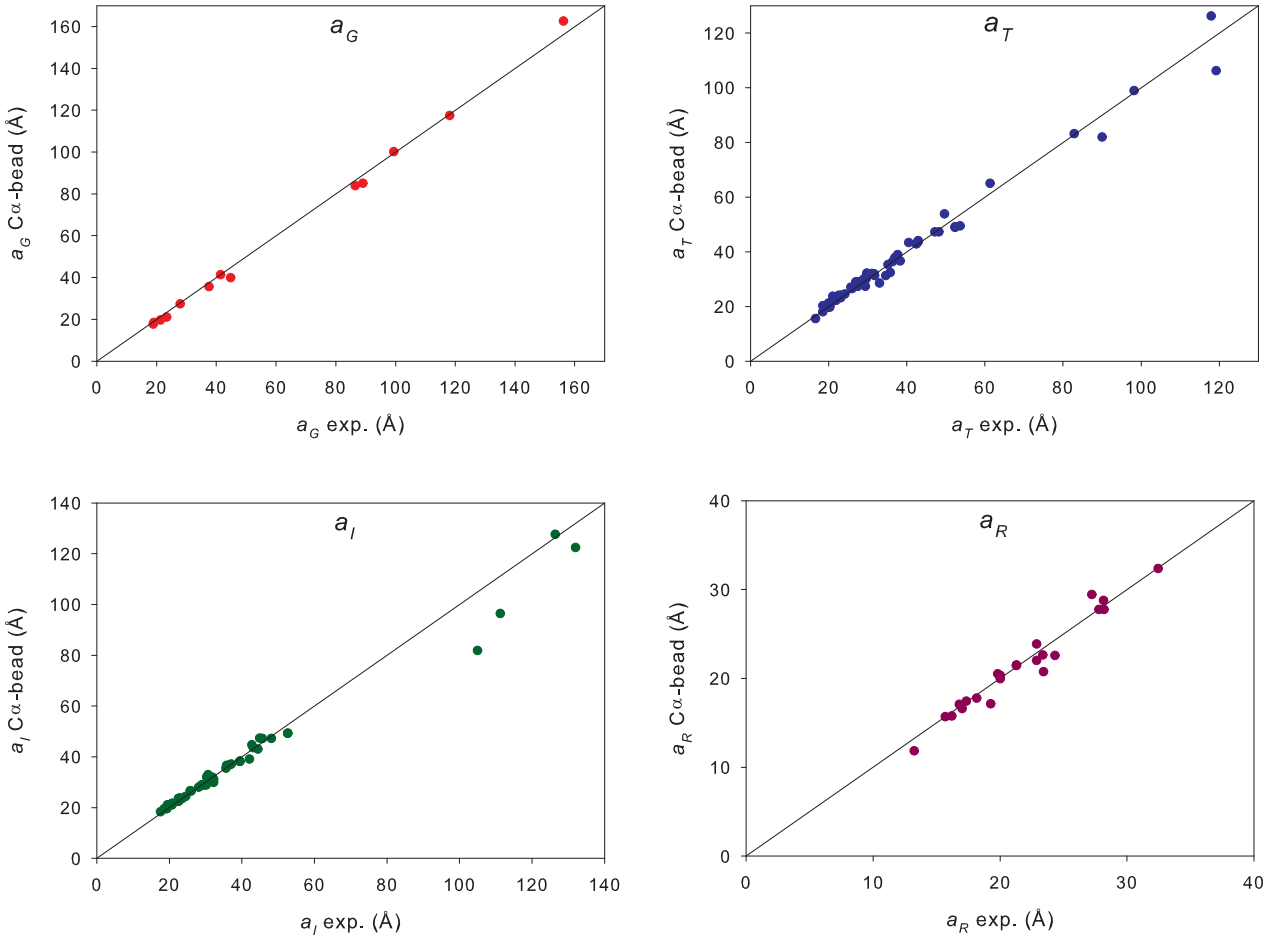

Figure S-5: Representation of the values calculated by HYDROPRO bead model from the C-α coordinates (y axis) vs. experimental values (x axis) of the four equivalent radii employed in this work, for the WHOLE set of proteins and including also the large proteins and macromolecular complexes

References

[1] J. García de la Torre, M.L. Huertas, and B. Carrasco. Calculation of hydrodynamic properties of globular proteins from their atomic-level structures. *Biophys. J.*, 78:719–730, 2000.

[2] J. García de la Torre. Hydration from hydrodynamics. general considerations and aplications of bead modelling to globular proteins. *Biophys. Chem.*, 93:150–159, 2001.

[3] D.K. Hahn and S.R. Aragon. Intrinsic viscosity of proteins and platonic solids by boundary element methods. *J. Chem. Theory Comput.*, 2:1416–1428, 2006.

[4] E. Brookes, B. Demeler, and M. Rocco. Developments in the US-SOMO bead modeling suite: New features in the direct residue-to-bead method, improved grid routines and influence of accesible surface area screening. *Macromol. Biosci.*, 10:746–753, 2010.

[5] J. Behlke, O. Ristau, and H. J. Schönfeld. Nucleotide-dependent complex formation between the *Escherichia coli* GroEL and GroES studied under equilibrium conditions. *Biochemistry*, 36:5149–5156, 1997.

[6] D. K. Clare, P. J. Bakkes, H. Van Heerikhuizen, S. M. Van Der Vies, and H. R. Saibil. An expanded protein folding cage in the GroEL-gp31 complex. *J. Mol. Biol.*, 358:905–911, 2006.

[7] I. Tomonao, K. Takahashi, K. Maki, S. Enoki, K. Kamagata, A. Kadooka, M. Arai, and K. Kuwajima. Asymmetry of the GroEL-GroES complex under physiological conditions as revealed by small-angle X-Ray scattering. *Biophys. J.*, 94:1392–1402, 2008.

[8] J. B. Sumner, N. Gralén, and I. B. Eriksson-Quensel. The molecular weight of urease. *J. Biol. Chem.*, 125:37–44, 1932.

[9] A. Balasubramanian and K. Ponnuraj. Crystal structure of the first plant urease from jack bean: 83 years of journey from its first crystal to molecular structure. *J. Mol. Biol.*, 400:274–283, 2010.

[10] W. N. Fishbein, K. Nagarajan, and W. Scurzi. Urease catalysis and structure. IX. The half unit and hemipolymers of jack bean urease. *J. Biol. Chem.*, 248:7870–7877, 1973.

[11] J. K. Armstrong, R. B. Wenby, H. J. Meiselman, and T. C. Fisher. The hydrodynamic radii of macromolecules and their effect on red blood cell aggregation. *Biophys. J.*, 87:4259–4270, 2004.

[12] S. J. Perkins, A. S. Nealis, B. J. Sutton, and A. Feinstein. Solution structure of human and mouse immunoglobulin M by synchrotron X-ray scattering and molecular graphics modelling. a possible mechanism for complement activation. *J. Mol. Biol.*, 221:1345–1366, 1991.

[13] R. E. Schrohenloher and J. C. Bennett. Degradation of human IgM by pepsin: Characterization of a high molecular weight fragment. *J. Immunol.*, 107:870–880, 1971.

[14] P. Wilhelm, I. Pilz, W. Palm, and K. Bauer. Small-angle X-ray studies of a human immunoglobulin M. *Eur. J. Biochem.*, 84:457–463, 1978.

[15] W. E. Hill, G. P. Rossetini, and K. E. Van Holde. Physical studies of ribosomes from *Escherichia coli*. *J. Mol. Biol.*, 44:263–277, 1969.

[16] B. S. Schuwirth, M. A. Borovinskaya, C. W. Hau, W. Zhang, A. Vila-Sanjurjo, J. M. Holton, and J. H. D. Cate. Structures of the bacterial ribosome at 3.5 A resolution. *Science*, 310:827–834, 2005.

[17] V. B. Patel, C. C. Cunningham, and R. R. Hantgan. Physicochemical properties of rat liver mitochondrial ribosomes. *J. Biol. Chem.*, 276:6739–6746, 2001.

[18] D. E. Koppel. Study of *Escherichia coli* ribosomes by intensity fluctuation spectroscopy of scattered laser light. *Biochemistry*, 13:2712–2719, 1974.

[19] N. Burkhardt, G. Diedrich, K.H. Nierhaus, W. Meerwinck, H.B. Stuhrmann, J. Skov Pedersen, M.H.J. Koch, V.V. Volkov, M.B. Kozin, and D.I. Svergun. Architecture of the *E. coli* 70s ribosome. *Physica B*, 234-236:199–201, 1997.

[20] W. E. Hill, J. W. Anderegg, and K. E. van Holde. Effects of solvent environment and mode of preparation on the physical properties of ribosomes from *Escherichia coli*. *J. Mol. Biol.*, 53:107–110, 1970.