Supporting Information

Probing the size of proteins with glass nanopores

Lorenz J. Steinbock, Swati Krishnan, Roman D. Bulushev, Sandrine Borgeaud, Melanie Blokesch, Lely Feletti and Aleksandra Radenovic*

Table S1: Overview of the utilized proteins and the homolog protein data bank (PDB) ID used for the representation of the proteins. The surface charge was calculated using the ProteinCalculator v3.4. The volume was calculated with the peptide property calculator of Northwestern University.

Protein	ComEA	GFP	FP	BSA	IgG	RNAp
PDB ID	2DUY	1EMA	20KY	4F5S	1IGT	3IYD
MW in kDa	12.2	26.9	54	66.5	145	480
Charges at pH 8	-3	-8	-15	-26	-9.2	-69
Volume (nm ³)	9.9	32.5	65.0	80.4	175.7	600.0
Diameter (nm)	2.6	3.96	5.0	5.35	6.9	10.46
Concentration (uM)	0.6	37	9	0.6	0.6	0.6
Charge/MW (e ⁻ /kDa)	-0.25	-0.30	-0.27	-0.39	-0.06	-0.30

Detailed explanation of PSCFP2-PAmCherry1 (FP):

The plasmid was transformed into the Bacterial strain BL21(DE3) pLysS (Stratagene, La Jolla, CA, USA) for high protein expression. Cultures were grown overnight in medium containing 100ug/mL ampicillin to a 0.4-0.5 optical density and then induced with 1mM IPTG at 37 °C for 4 hours. The cells were harvested from a 200 ml culture by centrifugation

and resuspended in 8 ml of binding buffer (500 mM NaCl, 50 mM NaH₂PO₄, pH 8). Then, 8 mg of lysozyme was added and the reaction was incubated 30 minutes on ice. The solution was sonicated on ice 6 times with a sonicator equipped with a microtip using six 10-second bursts at high intensity with a 10 second cooling period between each burst. The lysate was centrifuged at 3000 x g for 15 minutes to pellet cellular debris. The supernatant was transferred to a fresh tube and added to an equilibrated Invitrogen Ni-NTA purification Column. Binding was allowed for 1 hour with gentle agitation to keep the resin suspended in the lysate solution. The resin was allowed to settle by centrifugation (800 x g) and the supernatant was carefully aspirated. The column was washed with 8 ml of Native Wash Buffer with imidazole at low concentration (500 mM NaCl, 50 mM NaH₂PO₄, pH 8 and 20 mM imidazole). This last step was repeated three more times. The column was clamped in a vertical position and the cap was removed from the lower end in order to elute the protein with 8 ml native elution buffer (500 mM NaCl, 50 mM NaH₂PO₄, pH 8 and 250 mM imidazole). The high concentration of imidazole in the elution buffer allowed the PSCFP2-PAmCherry 1 proteins to detach from the Ni-NTA resin. After purification, the proteins were loaded on a SDS-Polyacrylamide electrophoresis gel stained with Comassie Brillant Blue staining.



Figure S1: Detailed representation of experiment presented in Figure 3 a, b and c. (a) Concatenated current trace while having *cis* filed with GFP and applying 0.5 V to the trans electrode causing events with decreases in the current. (b) After flushing the *cis* reservoir with 1 M KCl solution 0.5 V was again applied to the *trans* electrode as a control experiment. No translocations were observed visible in the stable current baseline. (c) FP was added to the *cis* chamber and translocated by applying 0.5 V to the *trans* electrode. Events manifested in the

concatenated current. (d) Again a control experiment was performed by washing the *cis* reservoir with a 1 M KCl solution and applying 0.5 V to the *trans* electrode. (e) λ -DNA was added to the *cis* side and translocated by applying 0.5 V to the *trans* electrode. Characteristic step-like events were recorded with clear quantization in the respective histogram.



Figure S2: Scheme showing the sizes of the proteins simplified as spheres and the corresponding sizes of the glass nanopore.



Figure S3: (a) The G_{Gauss} value, determined by Gauss fits in Figure 3 and 4, plotted against the molecular weight of the respective protein. The right histogram (b) presents the relative frequency count of the G* value shown in Figure 4 for the proteins GFP, FP, BSA, IgG and RNAp. The peaks were fitted with a Gauss function in order to determine the G_{Gauss} value. The error bars in (a) are the width of the fitted Gauss functions. As expected the overall trend shows an increase in the relative blocked conductance (G* value) for increasing MW values. The dashed line is calculated using equation (2) by assuming a membrane-like geometry. This causes higher values since the equation neglects the access resistance caused by the conical geometry of the glass nanopore. Moreover the low capture frequency and Bessel filter frequency of only 100 and 10 kHz, respectively, cause smaller peaks in the experimental data.^{1,2}

To compare the translocated proteins in one figure their G* value from the Gauss fit was plotted as a function of the molecular weight (Figure S3). The width of the Gauss fit was plotted as the error bars. G* increases for bigger molecular weights value except for the BSA (0.44 %) molecule and the ComEA experiment. The grey star at 12 kDa was generated from the difference in the G* value between 800 bp DNA and the 800 bp DNA-ComEA complex.

It therefore does not represent the translocation of a genuine free ComEA protein and cannot be directly compared to the other protein translocation experiments. Moreover, the difference can also occur from multiple ComEA binding to the same DNA strand as shown in AFM experiments by Seitz *et al.* and in similar nanopore experiments with the DNA-binding protein RecA.^{3,4}

Regarding the G* value for the FP, which is only 0.03 % bigger than the G* value for the BSA protein, the difference is small and the error bars have a wide distribution of the corresponding peaks (Figure S3 b). This might be due to the oval shape of the molecule (Figure 1 e). Another factor is the use of different glass nanopores causing changes in the fractional conductance blockade value. Further, looking at the histogram the red BSA distribution shows long tail reaching higher G* values than the orange FP distribution. This is in accordance with the bigger size of the BSA protein. Another feature in the histogram is the distinct peak for the FP and the GFP protein, in comparison with the broad tail of BSA, IgG and RNAp. One reason is the composition of the proteins. While GFP and FP are constituted by one amino acid chain, RNAp and IgG are built up by several amino acid chains connected only by relative weak covalent bonds such as di-sulfide interactions or even non-covalent bonds. Such proteins can disintegrate in the electrical field of the glass nanopore, causing it to fall apart into smaller fragments, which in return broadens the peak.⁵ Bigger proteins are also less compact, which is especially visible in the IgG molecule with its Y-structure (Figure 1 f), leading to different levels of blocked currents depending on the translocating orientation.² Further since the experiments are done at pH 8 the glass is negatively charged generating an electro-osmotic flow pointing in the opposite direction than the translocation direction of the used proteins here.^{6,7} This flow could obstruct larger molecules and increase their translocation time explaining different amount of blocked current.¹

Theories to predict the fractional conductance blocked were proposed in the early days of the resistive pulse technique by DeBlois and Bean.⁸ It was further refined by Ito *et al.* on

particles⁹, applied to proteins by Han *et al.*¹⁰ and tested lately on the enzymes ProtK and RNase by Larkin *et al.*¹¹ We used an adapted model to extract the effective sensing length of our glass nanopore by fitting the fractional conductance blockade values (G^*) as a function of the diameter of the proteins:¹¹

$$G^*(d_p) \cong \frac{d_m^3}{(h_{\text{eff}} + 0.8d_p)d_p^2} , \qquad (1)$$

where d_m is the diameter of the molecule, d_p the diameter of the glass nanopore and h_{eff} the effective length of the nanopore. The fit can be seen in Figure 6 as the dashed line with effective nanopore length as the fitting parameter, giving a value of 174 nm for h_{eff} . This value is higher than the ones used for nanopores in silicon nitride and HfO₂ membranes ranging around 7 nm, which are very thin due to the fabrication process giving rise to fractional current blockade values of around 40 %.¹¹ Taking into account the conical shape of our glass nanopores and the relative low fractional current blockade values between 0.2 and 3 % a higher effective nanopore length is comprehensible.

References

- 1. D. Pedone, M. Firnkes, and U. Rant, Anal. Chem., 2009, 81, 9689–94.
- 2. W. Li, N. A. W. Bell, S. Hernández-Ainsa, V. V Thacker, A. M. Thackray, R. Bujdoso, and U. F. Keyser, *ACS Nano*, 2013, **7**, 4129–34.
- 3. P. Seitz, H. Pezeshgi Modarres, S. Borgeaud, R. D. Bulushev, L. J. Steinbock, A. Radenovic, M. Dal Peraro, and M. Blokesch, *PLoS Genet.*, 2014, **10**, e1004066.
- 4. R. M. M. Smeets, S. W. Kowalczyk, a R. Hall, N. H. Dekker, and C. Dekker, *Nano Lett.*, 2009, **9**, 3089–96.
- 5. C. Raillon, P. Cousin, F. Traversi, E. Garcia-Cordero, N. Hernandez, and A. Radenovic, *Nano Lett.*, 2012, **12**, 1157–64.
- 6. M. Firnkes, D. Pedone, J. Knezevic, M. Döblinger, and U. Rant, *Nano Lett.*, 2010, **10**, 2162–7.
- 7. S. Umehara, N. Pourmand, C. D. Webb, R. W. Davis, K. Yasuda, and M. Karhanek, *Nano Lett.*, 2006, **6**, 2486–92.

- 8. R. W. DeBlois and C. P. Bean, *Rev. Sci. Instrum.*, 1970, **41**, 909.
- 9. T. Ito, R. M. Crooks, and L. Sun, Anal. Chem., 2003, 75, 2399–2406.
- 10. A. Han, G. Schürmann, G. Mondin, R. A. Bitterli, N. G. Hegelbach, N. F. de Rooij, and U. Staufer, *Appl. Phys. Lett.*, 2006, **88**, 093901–3.
- 11. J. Larkin, R. Y. Henley, M. Muthukumar, J. K. Rosenstein, and M. Wanunu, *Biophys. J.*, 2014, **106**, 696–704.