Cell, Volume *174*

Supplemental Information

Surface Properties Determining Passage Rates of Proteins through Nuclear Pores

Steffen Frey, Renate Rees, Jürgen Schünemann, Sheung Chun Ng, Kevser Fünfgeld, Trevor Huyton, and Dirk Görlich

# Plasmid	Permeation probe	Nsp1 FG hydrogel partition coefficient
pSF779	mCherry	0.40
pDG2049	EGFP wt	1.05
pDG2050	EGFP L7E	1.11
pDG2051	EGFP V11E	1.10
pDG2052	**EGFP Y39E**	**0.96**
pDG2053	**EGFP F99E**	**0.90**
pDG2055	**EGFP Y151R**	**0.83**
pDG2056	**EGFP M153E**	**0.87**
pDG2054	EGFP V176H	1.06
pDG2057	**EGFP Y182K**	**0.69**
pDG2058	EGFP I188K	1.10
pDG2059	EGFP F223R	1.16
pDG2060	**EGFP L231K, M233E, L236R, Y237D**	**0.84**
pSF1310	shGFP1 =EGFP+Y39E, F99E, Y151, M153R, Y182K, L231K, M233E, L236R, Y237D	**0.42**
pSF1438	shGFP2	**0.24**

Table S1: Mutations in EGFP diminishing FG-interactions, Related to Table 1
Gel: buffer partition coefficients for mCherry, EGFP, as well as of EGFP mutants where exposed hydrophobic residues had been changed to hydrophilic ones. Numbering of residues refers to the UniProt entry of wild type GFP (GFP_AEQVI).
Five single mutations (marked in bold) as well as the simultaneous mutation of four disordered residues at the C-terminus (pDG2060) reduced the partitioning into the FG hydrogel significantly. The data suggest that exposed hydrophobic residues promote an interaction with FG domains, even though these residues do not represent a dedicated FG-binding site. The combination of the effective mutations in shGFP1 (super-hydrophilic GFP 1) lowered the partition coefficient even more - to that of mCherry. The combined mutant effect approximates the product of the individual mutation effects, i.e. the effects are incremental.
The table also lists ineffective mutations (L7E and F223R), which illustrate early challenges in mutant design. The increase in FG interaction by the F223R mutation will be explained below by arginine also promoting FG interactions. Leu7 covers F114 in the GFP structure, and MD simulations suggest that the L7E mutation actually does not make GFP more hydrophilic; instead it makes F114 solvent-exposed.

# Plasmid	Permeation probe	Nsp1 FG hydrogel partition coefficient
pSF1526	EGFP wt	1.04
pSF1538	EGFP D36W	1.49
pSF1553	EGFP T38W	1.41
pSF1539	EGFP **K41W**	**2.89**
pSF1540	EGFP R73W	1.47
pSF1534	EGFP **T97W**	**3.73**
pSF1533	EGFP D117W	1.43
pSF1532	EGFP E132W	1.30
pSF1541	EGFP N149W	1.67
pSF1531	EGFP **N164W**	**3.23**
pSF1530	EGFP **D180W**	**2.02**
pSF1529	EGFP D190W	1.39
pSF1542	EGFP N198W	1.72
pSF1552	EGFP S202W	1.69
pSF1528	EGFP **Q204W**	**2.19**
pSF1554	EGFP A206K	0.99
pSF1527	EGFP K209W	1.40
pSF1550	EGFP K221W	1.60
pSF1551	EGFP T225W	1.50

Table S2: Position-specific effects of single tryptophans on FG-interaction of EGFP, Related to Table 1

Gel: buffer partition of EGFP mutants carrying single exchanges of surface residues to tryptophans. The most efficient mutations are highlighted in bold. Sorting is by amino acid position. A206K eliminates the weak dimerization tendency of GFP and was included as a control.

# Plasmid	Protein	MALS Mw	Oligomerisation state
Standard, hydrophilic and super-inert GFP monomers			
pSF1526	EGFP	28 kDa ±0.6%	1
pSF1438	shGFP2	29 kDa ±1.4%	1
pSF2893	sinGFP1	28 kDa ±1.5%	1
pDG2754	sinGFP4a	28 kDa ±0.5%	1
Superhydrophobic GFPs and GFPs testing specific amino acids at 8 pre-defined positions			
pSF2646	efGFP_0W	28 kDa ±1.6%	1
pSF2647	efGFP_3W	33 kDa ±0.5%	1.2
pSF2648	efGFP_5W	56 kDa ±0.3%	2
pSF2649	efGFP_8W	ND	≥4*
pSF2654	efGFP_8Y	ND	≥2*
pSF2653	efGFP_8F	170 kDa ±1.9%	6
pSF2651	efGFP_8L	59 kDa ±0.2%	2.2
pSF2650	efGFP_8I	56 kDa ±0.2%	2.1
pSF2652	efGFP_8M	48 kDa ±0.5%	1.7
pDG2939	efGFP_8V	52 kDa ±0.2%	1.8
pDG2931	efGFP_8C	37 kDa ±3.8%	1.2
pDG2940	efGFP_8A	28 kDa ±0.7%	1
pDG2934	efGFP_8H	38 kDa ±0.5%	1.3
pSF2892	efGFP_8R	30 kDa ±0.7%	1
pDG2936	efGFP_8Q	28 kDa ±0.5%	1
pDG2938	efGFP_8T	31 kDa ±0.5%	1
pDG2937	efGFP_8S	28 kDa ±0.5%	1
pDG2932	efGFP_8E	28 kDa ±0.4%	1
pDG2935	efGFP_8K	28 kDa ±0.5%	1
pDG2966	efGFP_8N	28 kDa ±0.5%	1
Superfast, Arginine-rich GFPs and R→K revertants			
pDG2805	sffrGFP4	29 kDa ±1.8%	1
pSF2884	sffrGFP4 18x R→K	28 kDa ±2.2%	1
pSF2885	sffrGFP4 25x R→K	27 kDa ±1.8%	1
Charge series based on sffrGFP4			
pDG2805	sffrGFP4	29 kDa ±0.3%	1
pDG2806	sffrGFP5	28 kDa ±1.0%	1
pDG2713	sffrGFP6	29 kDa ±1.6%	1
pDG2715	sffrGFP7	29 kDa ±1.4%	1
GFP^{NTR}- variants combining surface hydrophobicity with translocation-promoting arginines			
pSF2902	GFP_MaxR_3W	ND	≥1.5*
pSF2903	GFP_MaxR_5W	140 kDa ±0.4%	5
pSF2905	GFP_MaxR_8i	47 kDa ±1.6%	1.7
pDG2718	GFP^{NTR}_2B7	65 kDa ±1.7%	2.4
pDG2798	GFP^{NTR}_7B3	42 kDa ±2.0%	1.6
GFP^{NTR} tetramers			
pDG2721	3B1	114 kDa ±1.2%	4
pDG2722	3B7	110 kDa ±0.6%	4
pDG2779	3B7C	110 kDa ±0.5%	4
pDG2723	3B8	110 kDa ±0.5%	4
pDG2724	3B9	113 kDa ±0.7%	4

Table S3. MALS analysis of engineered GFP variants, Related to Table 1 and Figures 1-6

Native masses of indicated GFP variants were measured by multi-angle light scattering (MALS). Oligomeric states that deviate from whole numbers point to a oligomer mixture and association equilibrium at the measured concentration (~10 μM within the measured main peak).

Note that the superhydrophilic and super-inert GFPs are monomers. Their very slow NPC-passage can thus not be explained by an increase in size. The arginine-rich GFPs and the charge series represent also monomers. Superhydrophobic variants form (to varying degrees) oligomers – consistent with their aggregation propensities. The highly FG-specific GFP^{NTR} variants 3B1 – 3B9 are clean tetramers.

* efGFP_8W, efGFP_8Y, and GFP_MaxR_3W stick to Superdex and could therefore not be measured in the standard setup. Oligomeric states where therefore estimated at a ~10μM concentration by dynamic light scattering. At higher concentrations, larger complexes form.

	GFP^{NTR}_3B8 (pdb: 5MSE)
Wavelength	0.97 Å
Resolution range	46.28 - 1.664 (1.723 - 1.664)
Space group	P 1 21 1
Unit cell	73.401 73.398 93.927 90 93.734 90
Total reflections	780311 (77838)
Unique reflections	114871 (11195)
Multiplicity	6.8 (7.0)
Completeness (%)	98.38 (96.49)
Mean I/sigma(I)	12.47 (1.18)
Wilson B-factor	23.67
R-merge	0.1287 (1.91)
R-meas	0.1395 (2.062)
R-pim	0.05327 (0.7709)
CC1/2	0.998 (0.468)
CC*	0.999 (0.798)
Reflections used in refinement	114850 (11195)
Reflections used for R-free	5744 (560)
R-work	0.1852 (0.3318)
R-free	0.2170 (0.3398)
CC(work)	0.964 (0.637)
CC(free)	0.949 (0.550)
Number of non-hydrogen atoms	8180
macromolecules	7276
ligands	116
solvent	788
Protein residues	899
RMS(bonds)	0.027
RMS(angles)	2.19
Ramachandran favored (%)	98.29
Ramachandran allowed (%)	1.71
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	2.56
Clashscore	4.37
Average B-factor	23.97
macromolecules	22.41
ligands	19.63
solvent	39.07
Number of TLS groups	4

Table S4: Data collection and refinement statistics, related to Figure 6

Data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses.

# Plasmid	Target Protein	N-terminal tag	Source
pSF360	yNTF2	-	This study
pDG2121	rNTF2	-	This study
pSF779	mCherry	His14-Tev	This study
pSF1526	EGFP	His14-MBP-bdSUMO	This study
pDG2049	EGFP wt	His14-Tev	This study
pDG2050	EGFP L7E	His14-Tev	This study
pDG2051	EGFP V11E	His14-Tev	This study
pDG2052	EGFP Y39E	His14-Tev	This study
pDG2053	EGFP F99E	His14-Tev	This study
pDG2055	EGFP Y151R	His14-Tev	This study
pDG2056	EGFP M153E	His14-Tev	This study
pDG2054	EGFP V176H	His14-Tev	This study
pDG2057	EGFP Y182K	His14-Tev	This study
pDG2058	EGFP I188K	His14-Tev	This study
pDG2059	EGFP F223R	His14-Tev	This study
pDG2060	EGFP L231K, M233E, L236R, Y237D	His14-Tev	This study
pSF1310	shGFP1	His14-Tev	This study
pSF1438	shGFP2	His14-Tev	This study
pSF2893	sinGFP1	His14-bdSUMO	This study
pDG2754	sinGFP4a	His14-bdSUMO	This study
pSF1526	EGFP wt	His14-MBP-bdSUMO	This study
pSF1538	EGFP D36W	His14-MBP-bdSUMO	This study
pSF1553	EGFP T38W	His14-MBP-bdSUMO	This study
pSF1539	EGFP K41W	His14-MBP-bdSUMO	This study
pSF1540	EGFP R73W	His14-MBP-bdSUMO	This study
pSF1534	EGFP T97W	His14-MBP-bdSUMO	This study
pSF1533	EGFP D117W	His14-MBP-bdSUMO	This study
pSF1532	EGFP E132W	His14-MBP-bdSUMO	This study
pSF1541	EGFP N149W	His14-MBP-bdSUMO	This study
pSF1531	EGFP N164W	His14-MBP-bdSUMO	This study
pSF1530	EGFP D180W	His14-MBP-bdSUMO	This study
pSF1529	EGFP D190W	His14-MBP-bdSUMO	This study
pSF1542	EGFP N198W	His14-MBP-bdSUMO	This study
pSF1552	EGFP S202W	His14-MBP-bdSUMO	This study
pSF1528	EGFP Q204W	His14-MBP-bdSUMO	This study
pSF1554	EGFP A206K	His14-MBP-bdSUMO	This study
pSF1527	EGFP K209W	His14-MBP-bdSUMO	This study
pSF1550	EGFP K221W	His14-MBP-bdSUMO	This study
pSF1551	EGFP T225W	His14-MBP-bdSUMO	This study
pSF2646	efGFP_0W	His14-ZZ-SUMOstar	This study
pSF2647	efGFP_3W	His14-ZZ-SUMOstar	This study
pSF2648	efGFP_5W	His14-ZZ-SUMOstar	This study
pSF2649	efGFP_8W	His14-ZZ-SUMOstar	This study
pSF2654	efGFP_8Y	His14-ZZ-SUMOstar	This study
pSF2653	efGFP_8F	His14-ZZ-SUMOstar	This study
pSF2651	efGFP_8L	His14-ZZ-SUMOstar	This study
pSF2650	efGFP_8I	His14-ZZ-SUMOstar	This study
pSF2652	efGFP_8M	His14-ZZ-SUMOstar	This study
pSF2892	efGFP_8R	His14-ZZ-SUMOstar	This study
pDG2931	efGFP_8C	His14-bdSUMO	This study
pDG2932	efGFP_8E	His14-bdSUMO	This study
pDG2933	efGFP_8G	His14-bdSUMO	This study
pDG2934	efGFP_8H	His14-bdSUMO	This study
pDG2935	efGFP_8K	His14-bdSUMO	This study
pDG2936	efGFP_8Q	His14-bdSUMO	This study

pDG2937	efGFP_8S	His14-bdSUMO	This study
pDG2938	efGFP_8T	His14-bdSUMO	This study
pDG2939	efGFP_8V	His14-bdSUMO	This study
pDG2940	efGFP_8A	His14-bdSUMO	This study
pDG2966	efGFP_8N	His14-bdSUMO	This study
pDG2805	sffrGFP4	His14-bdSUMO	This study
pSF2884	sffrGFP4 18xR→K	His14-bdSUMO	This study
pSF2885	sffrGFP4 25xR→K	His14-bdSUMO	This study
pDG2806	sffrGFP5	His14-bdSUMO	This study
pDG2713	sffrGFP6	His14-bdSUMO	This study
pDG2715	sffrGFP7	His14-bdSUMO	This study
pSF2902	GFP_MaxR_3W	His14-bdSUMO	This study
pSF2903	GFP_MaxR_5W	His14-bdSUMO	This study
pSF2905	GFP_MaxR_8i	His14-bdSUMO	This study
pDG2718	GFP^{NTR}_2B7	His14-bdSUMO	This study
pDG2798	GFP^{NTR}_7B3	His14-bdSUMO	This study
pDG2804	Sin_tCherry2	His14-bdSUMO	This study
pDG2721	GFP^{NTR}_3B1	His14-bdSUMO	This study
pDG2722	GFP^{NTR}_3B7	His14-bdSUMO	This study
pDG2723	GFP^{NTR}_3B8	His14-bdSUMO	This study
pDG2724	GFP^{NTR}_3B9	His14-bdSUMO	This study
pDG2779	GFP^{NTR}_3B7C	His14-bdSUMO	This study
pDG2781	GFP^{NTR}_3B7C covalent tetramer	His14-bdSUMO	This study
pDG2725	GFP NTR 4B1	His14-bdSUMO	This study
pSF1864	MBP[G261C]	His14-bdSUMO	This study
pSF2895	MBP[G261C] K→R	His14-bdSUMO	This study
pSF2911	Importin β 1-493	His14-bdSUMO	This study
pSF2912	Importin β 1-493_R→K	His14-bdSUMO	This study
pSF2913	Importin β 1-493_K→R	His14-bdSUMO	This study
pDG2894	IBB sinGFP4A	His14-bdSUMO	This study
pDG2895	IBB EGFP	His14-bdSUMO	This study
pDG2896	IBB EGFP T97W	His14-bdSUMO	This study
pDG2898	IBB sffrGFP4	His14-bdSUMO	This study
pDG2899	IBB sffrGFP7	His14-bdSUMO	This study
pDG2910	IBB-3x sinGFP4a	His14-bdSUMO	This study
pDG2911	IBB-3x EGFP	His14-bdSUMO	This study
pDG2915	Human importin β	His-bdNEDD8	This study
pDG2960	Human Ran	His-ZZ-scSUMO	This study
pSF590	Tev-Protease	MBP-Tev-His14	Frey and Görlich, 2014a
pSF1390	bdSENP1 protease	His14-Tev	Frey and Görlich, 2014a
pDG2582	bdNEDP1 protease	His14-MBP-bdSUMO	Pleiner et al., 2018
pHBS418	Tth MacNup98A FG domain (1-666)	His18	Schmidt and Görlich, 2015
pHBS698	Sc Nup116 FG domain (1-736)	His18	Schmidt and Görlich, 2015
pSF1020	Sc Nsp1 FG domain (2-601)	His14-Tev	This study

Table S5: Vectors for recombinant protein expression in *E. coli*, related to STAR Methods

Plasmid numbers are unique identifies. Not listed are plasmids of >200 GFP variants that were intermediates in GFP^{NTR} evolution experiments.