Supplementary Information

Structural and biochemical analysis of *Escherichia coli* ObgE, a central regulator of bacterial persistence

Sotirios Gkekas[‡§], Ranjan Kumar Singh[‡§], Alexander V. Shkumatov[‡§], Joris Messens[‡§], Maarten Fauvart[¶], Natalie Verstraeten[¶], Jan Michiels[¶] and Wim Versées[‡§]**

[‡]Structural Biology Brussels, Vrije Universiteit Brussel, 1050 Brussels, Belgium

[§]VIB-VUB Center for Structural Biology 1050 Brussels, Belgium

[¶]Centre of Microbial and Plant Genetics, KU Leuven - University of Leuven, 3001 Leuven, Belgium

[**]Department of Life Science Technologies, Smart Systems and Emerging Technologies Unit, imec, 3001 Leuven, Belgium

To whom correspondence should be addressed: Prof. Dr. ir. Wim Versées, Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium. Telephone: +32 2 629 18 49; FAX: +32 2 629 19 63; E-mail: wim.versees@vib-vub.be

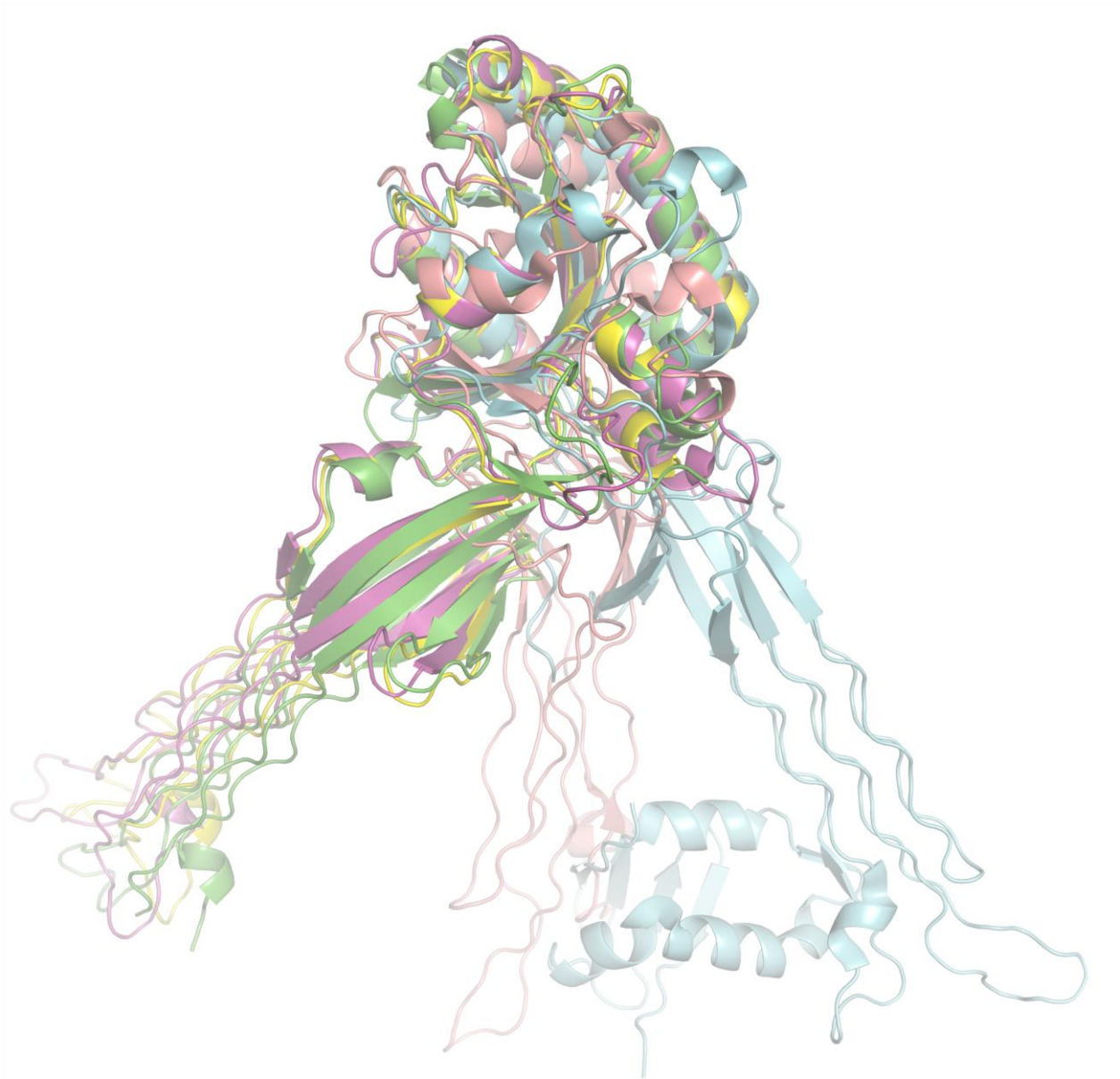

FIGURE S1. Superposition of *E. coli* ObgE_340 bound to GDP (green, this study), Obg from *B. subtilis* either in nucleotide free state (magenta) or bound to ppGpp (yellow) (PDB: 1LNZ), Obg from *T. thermophilus* in nucleotide-free state (cyan, PDB: 1UDX) and ObgE bound to the 50S ribosomal subunit (orange, PDB: 4CSU). All main chain atoms from the G domains are used for the superposition. This superposition shows a different orientation of the Obg domain vis-à-vis the G domain in *T. thermophilus* Obg compared to *E. coli* (this study) and *B. subtilis* Obg. In *E. coli* ObgE bound to the 50S ribosomal subunit, the Obg domain adopts an intermediate orientation.

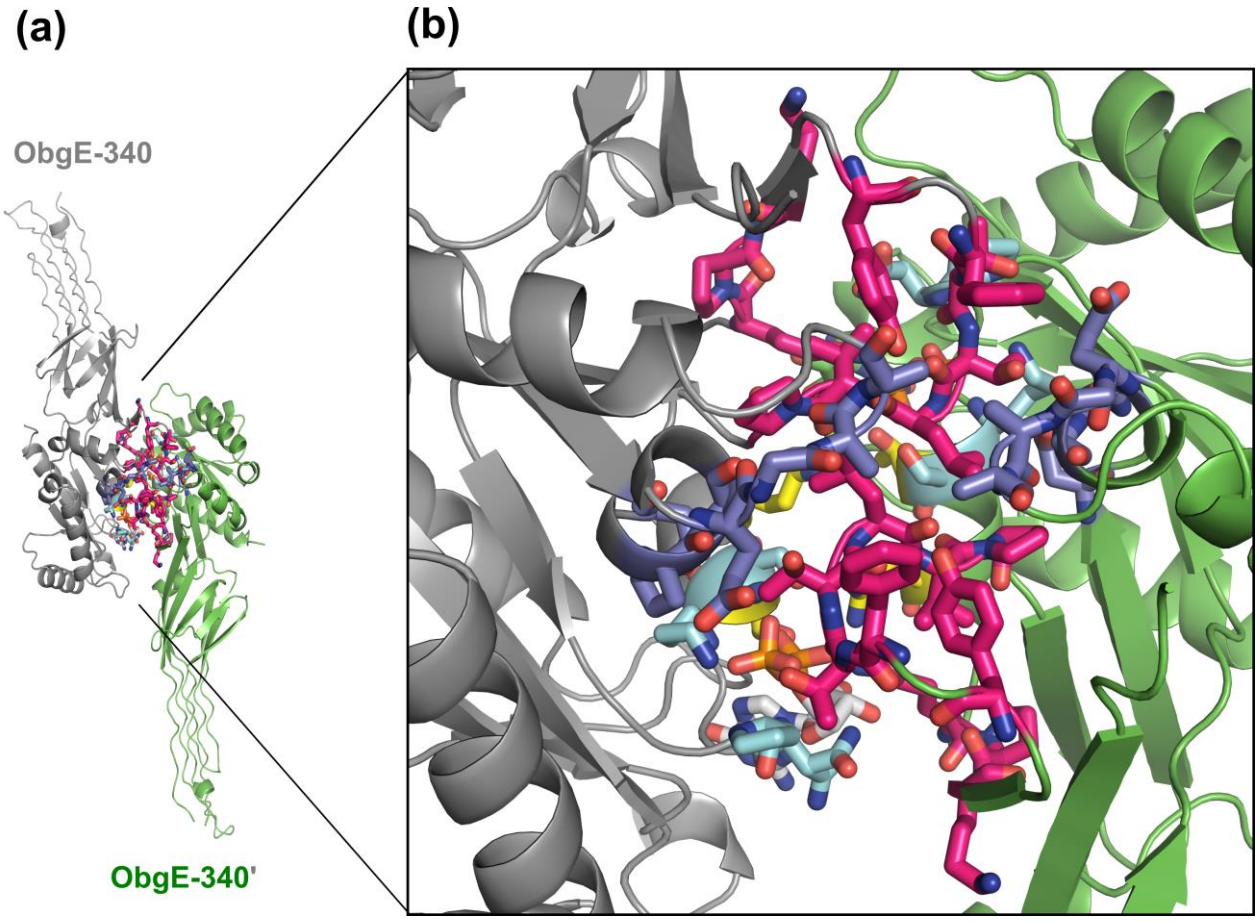

FIGURE S2. Detailed view on the potential dimer interface of ObgE_340 generated though crystal symmetry operations. (a) Potential dimer organization of ObgE_340 via a face-to-face interaction of its G domain with the G domain of a neighboring protein molecule in the crystal lattice. The two symmetry variants are colored grey and green, respectively, and the C-atoms of the bound GDP molecules are colored according to the protomer to which they belong. Amino acids involved in dimer contacts from the first α-helices of the G domain are colored in yellow, from the P loops in cyan, from the Switch I loops in magenta and from the Switch II loops in blue. (b) Close-up view on the potential dimer interface. The same colors as in (a) are used.

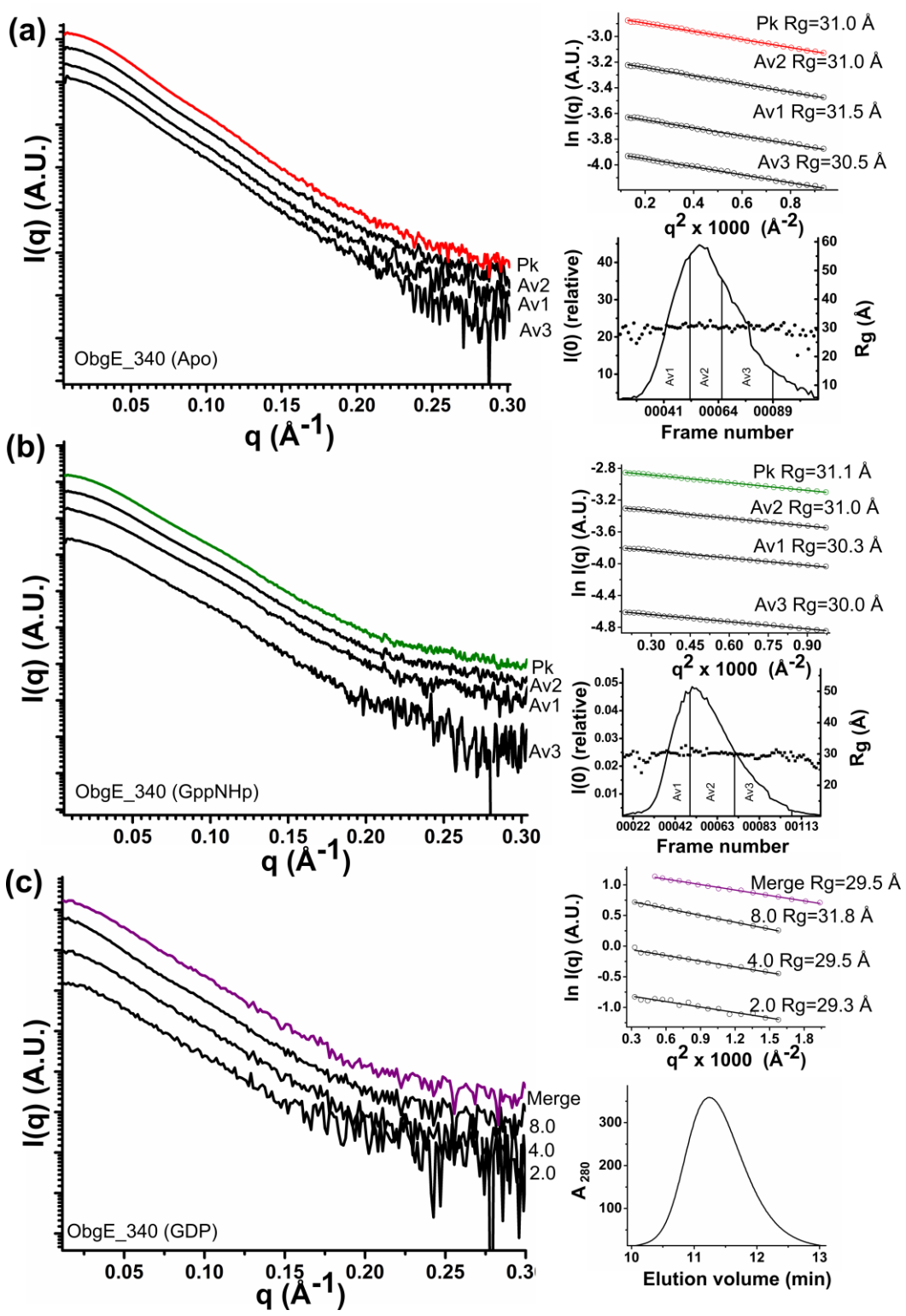

FIGURE S3. SEC-SAXS data for (a) ObgE_340 in nucleotide-free state and (b) ObgE_340 in GppNHp-bound state, and (c) batch SAXS data for ObgE_340 in GDP-bound state. In (a) and (b) **t**he lower right figure of each panel corresponds to the DATASW processing and represents the elution profiles of the SEC-SAXS experiment (I(0) vs. frame number) with each frame number corresponding to a scatter curve together with the corresponding R_g value. A stable R_g over the entire elution profile is observed. To further clarify

that no aggregation is taking place, averaged scatter curves of multiple frames (Av1, Av2, Av3) along the chromatogram were calculated (averaged regions are separated by vertical lines) which are shown on the figure on the left. The SAXS curve indicated by "Pk" represents the average over the entire peak of the elution profile and this curve is ultimately used for all further analyses. For clarity the latter curve is translated along the Y-axis. The top right figure in each panel represents the Guinier plot of the averaged SAXS curves, together with the calculated R_g value. In (c) scatter curves are shown collected in batch on a solution of 2 mg/ml, 4 mg/ml and 8 mg/ml of ObgE_340 in presence of GDP (indicated by 2.0, 4.0 and 8.0 respectively). The lower right panel shows the chromatogram of the (off-line) SEC that was performed just prior to preparing the samples for SAXS. The left panel shows the SAXS curves at each protein concentration together with the merged curve that was used for all analyses. For clarity the latter curve is translated along the Y-axis. The top right figure represents the Guinier plot of the SAXS curves, together with the calculated R_g value.

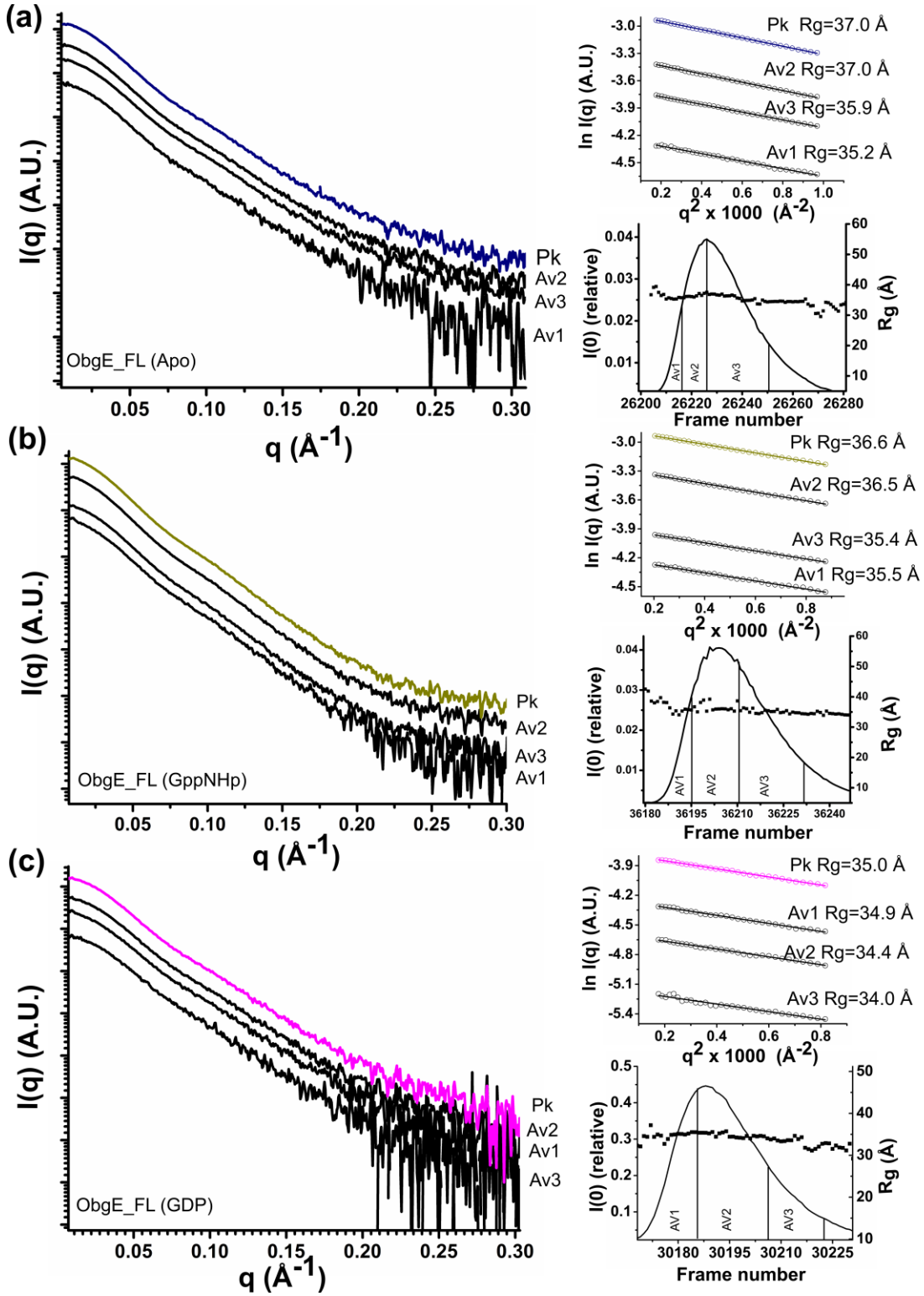

FIGURE S4. SEC-SAXS data for (a) ObgE_FL in nucleotide-free state, (b) ObgE_FL in GppNHp-bound state and (c) ObgE_FL in GDP-bound state. The lower right figure of each panel corresponds to the DATASW processing and represents the elution profiles of the SEC-SAXS experiment (I(0) vs. frame number) with each frame number corresponding to a scatter curve together with the corresponding R_g value. A stable R_g over the entire elution profile is observed. To further clarify that no aggregation is taking place,

averaged scatter curves of multiple frames (Av1, Av2, Av3) along the chromatogram were calculated (averaged regions are indicated by vertical lines) which are shown on the figure on the left. The SAXS curve indicated by "Pk" represents the average over the entire peak of the elution profile and this curve is ultimately used for all further analyses. For clarity the latter curve is translated along the Y-axis. The top right figure in each panel represents the Guinier plot of the averaged SAXS curves, together with the calculated R_g value.

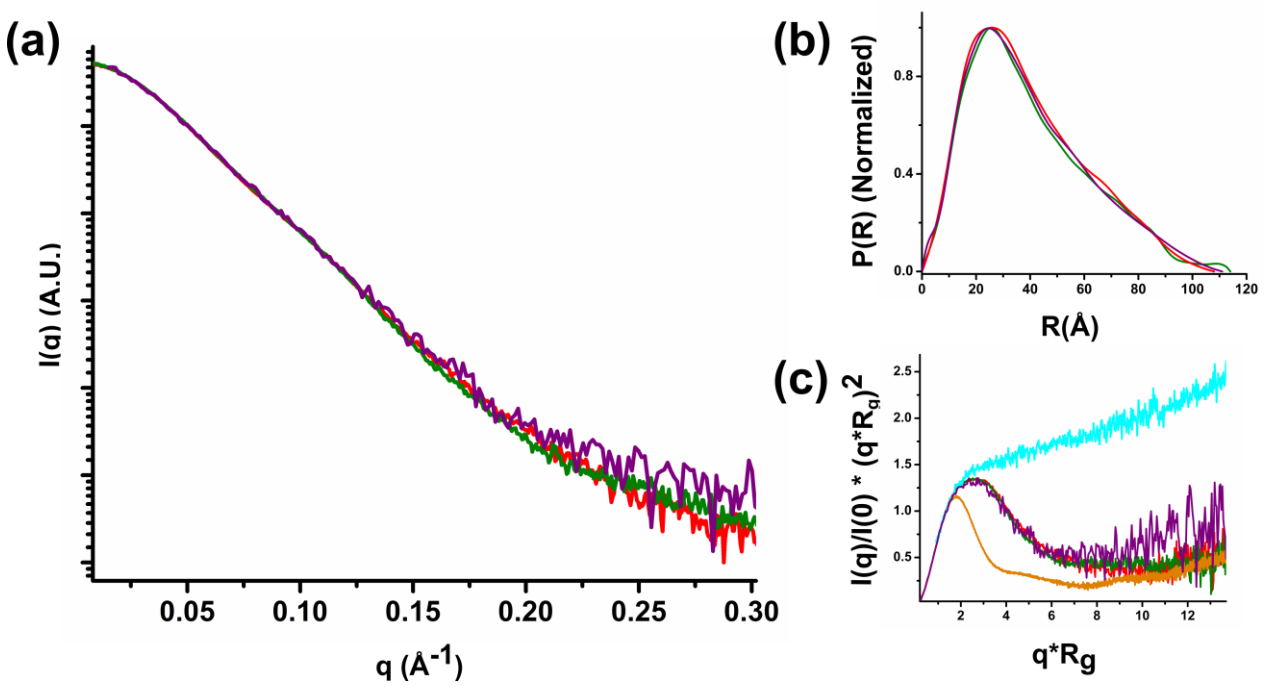

FIGURE S5. Comparison of SAXS profiles of ObgE_340 in nucleotide-free and nucleotide-bound state. (a) Averaged scattering profiles of ObgE_340 in the apo (red), GppNHp-bound (green) and GDP-bound (purple) state. (b) Normalized P(R) profiles of ObgE_340 in apo (red), GppNHp-bound (green) and GDP-bound (purple) state. (c) Dimensionless Kratky plot for ObgE_340 in apo (red), GppNHp-bound (green) and GDP-bound (purple) state. For comparison, similar plots of an intrinsically disordered protein (hTau40wt, cyan trace) and a globular protein (BSA, orange trace) are shown [71].

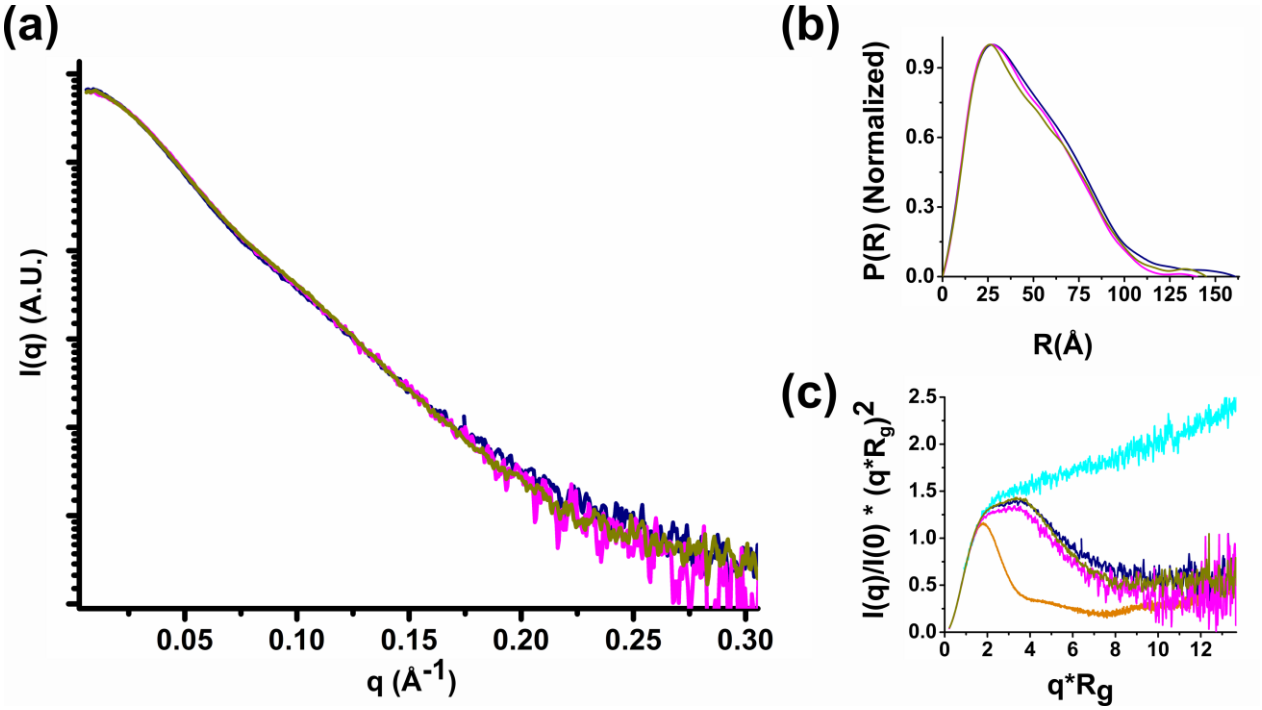

FIGURE S6. Comparison of SAXS profiles of ObgE_FL in nucleotide-free and nucleotide-bound state. (a) Averaged scattering profiles of ObgE_FL in the apo (blue), GppNHp-bound (dark yellow) and GDP-bound (magenta) state. (b) Normalized P(R) profiles of ObgE_FL in apo (blue), GppNHp-bound (dark yellow) and GDP-bound (magenta) state. (c) Dimensionless Kratky plot for ObgE_FL in apo (blue), GppNHp-bound (dark yellow) and GDP-bound (magenta) state. For comparison, similar plots of an intrinsically disordered protein (hTau40wt, cyan) and a globular protein (BSA, orange) are shown [71].

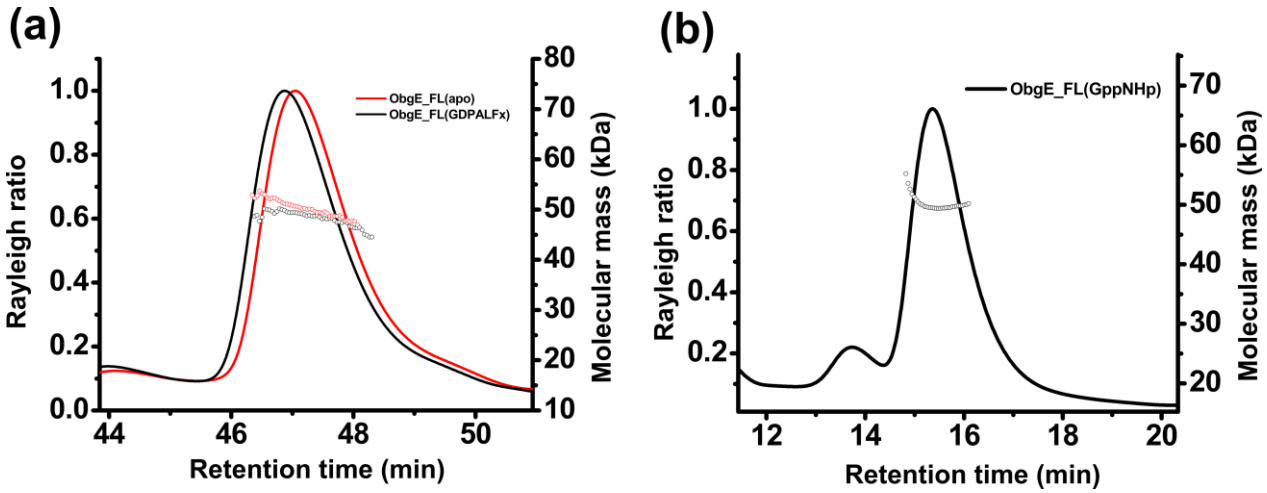

FIGURE S7. Determination of molecular mass of ObgE_FL in nucleotide-free, GppNHp-bound and GDP-AlF_x-bound state using Multi Angle Light Scattering coupled to size exclusion chromatography (SEC-MALS). (a) SEC-MALS of ObgE_FL in the nucleotide-free (red) and GDP-AlF_x-bound (black) state loaded on a Shodex KW-800 column. (b) SEC-MALS of ObgE_FL in the GppNHp-bound state (black) loaded on an Agilent Bio-SEC 3 column.

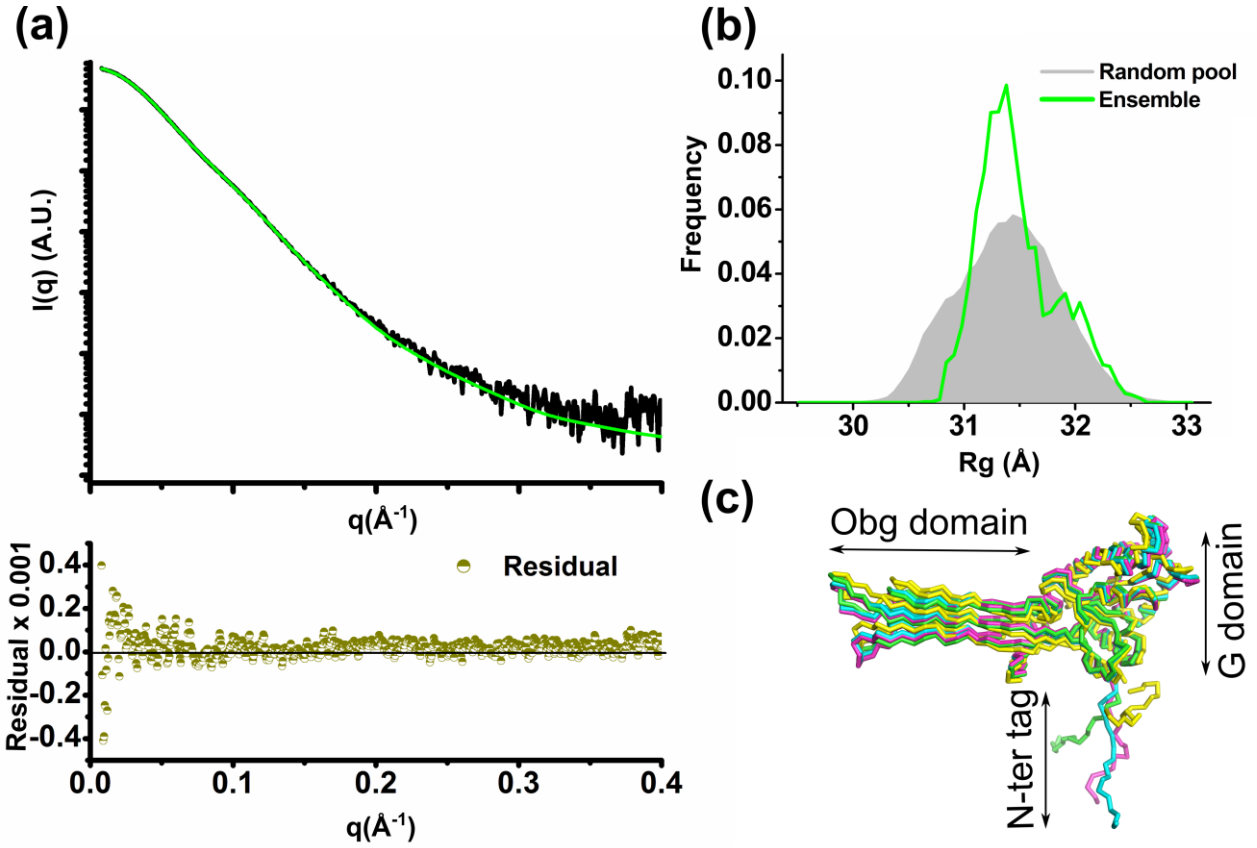

FIGURE S8. Comparison of experimental SAXS data for ObgE_340 with the theoretical scattering curve of the ObgE_340 crystal structure taking into account the flexibility of the N-terminal purification tag. (a) Ensemble fit of ObgE_340 (green line) to the experimental ObgE_340 scattering profile (black trace). The residuals are shown at the bottom of the fit (dark green). (b) The R_g distribution of the ensemble (green curve) compared to the R_g distribution of a random pool of models (grey filled area) generated using EOM coupled to all-atom modeling, model validation and NMA-refinement. (c) Ribbon representation of 4 ensemble models, selected by the genetic algorithm, which give an average theoretical curve that fits the experimental ObgE_340 SAXS profile as shown in (a).

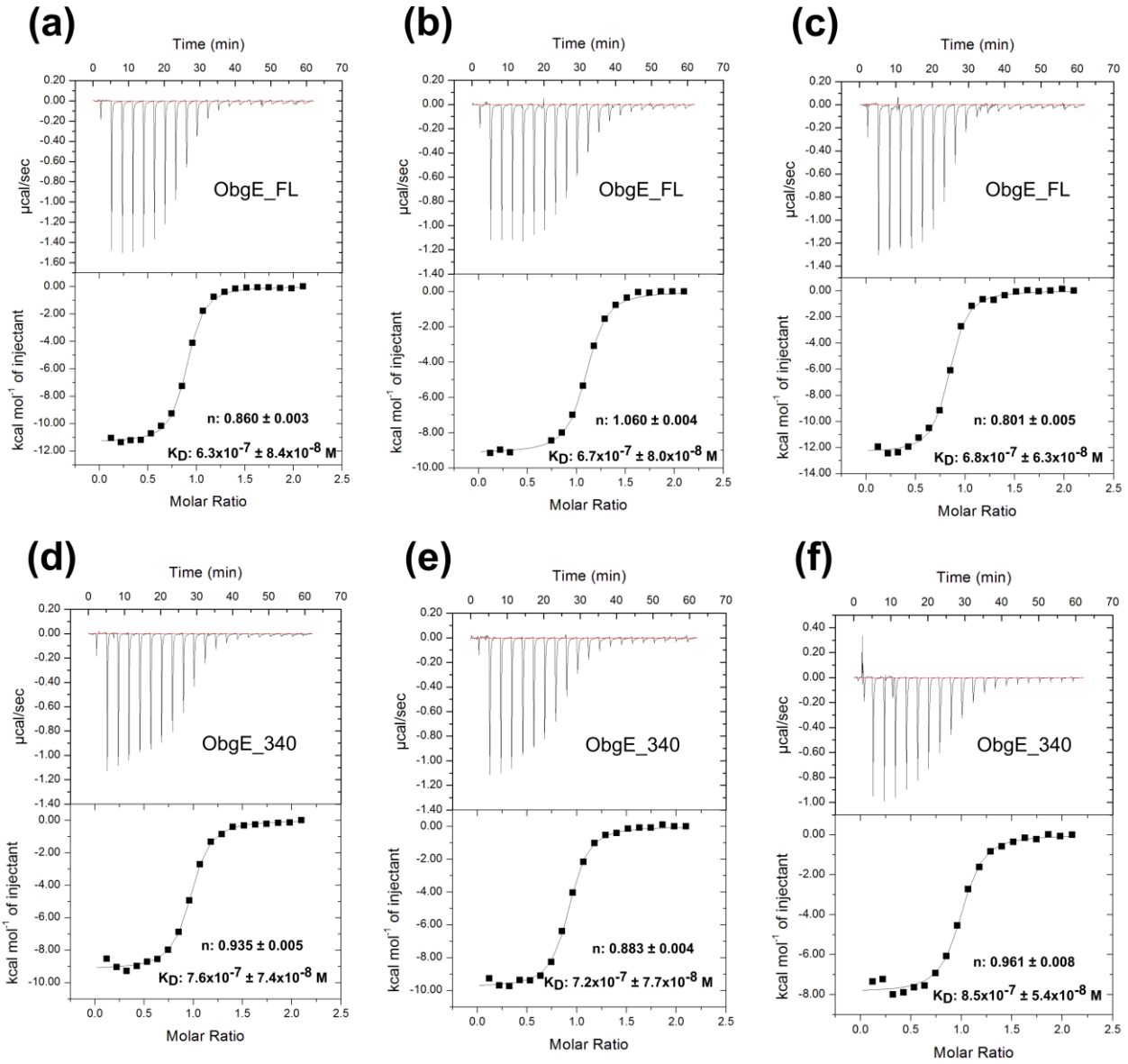

FIGURE S9. Isothermal titration calorimetry (ITC) experiments, performed in triplicate, for binding of ppGpp to ObgE_FL (a, b, c) and ObgE_340 (d, e, f) respectively. Experiments were performed at 25°C by titrating ppGpp from a stock solution of 750 µM into a solution of 75 µM nucleotide-free ObgE. Fitting on a single-site binding model yields the reported equilibrium dissociation constants (K_D) and binding stoichiometry (n) (± error of the fit).

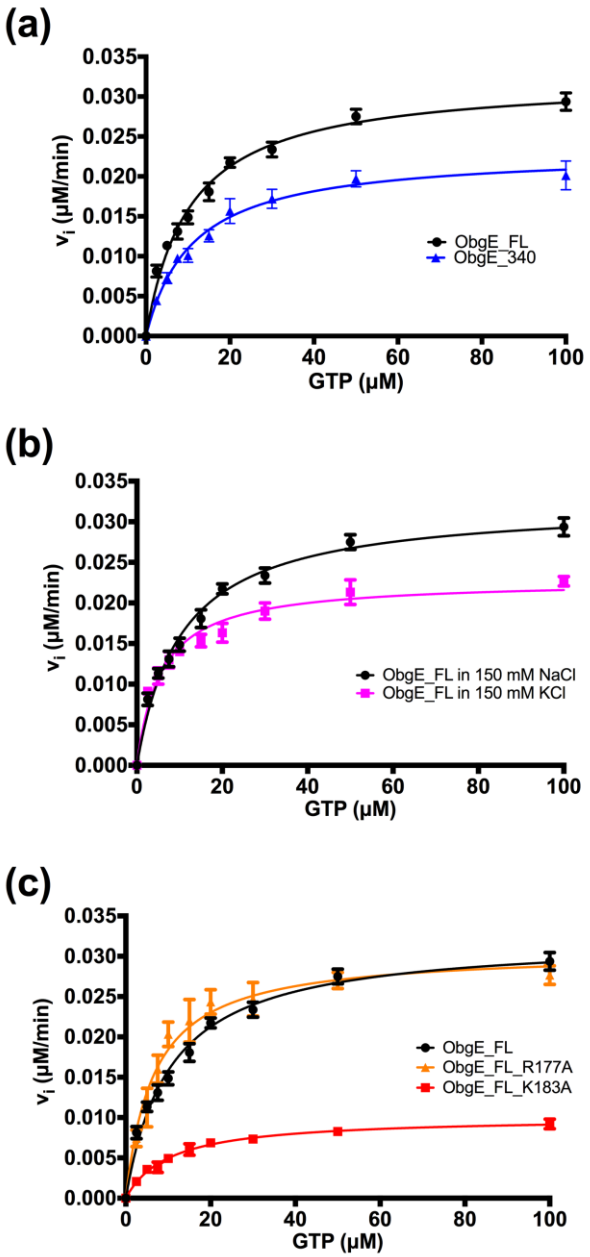

FIGURE S10. Initial rate steady-state kinetics (Michaelis-Menten) for GTP hydrolysis by ObgE. (a) for ObgE_FL and ObgE_340, (b) for ObgE_FL in presence of 150 mM NaCl or 150 mM KCl, (c) for ObgE_FL wild type and the R177A and K183A mutants. Each data point represents the average ± s.d. of 3 independent measurements. See Table 2 for steady-state kinetic parameters resulting from the fit on the Michaelis-Menten equation.

TABLE S1: SAXS data collection statistics and derived parameters.

	ObgE_FL Apo	**ObgE_FL GDP**	**ObgE_FL GppNHp**	**ObgE_340 Apo**	**ObgE_340 GDP**	**ObgE_ 340 GppNHp**
Data collection parameters						
Beamline	Swing, Soleil	Swing, Soleil	Swing, Soleil	Swing, Soleil	BioSAXS-2000	Swing, Soleil
Wavelength (Å)	1.03	1.03	1.03	1.03	1.54	1.03
q range (Å^{-1})*	0.0061-0.553	0.0061-0.553	0.0061-0.553	0.0062-0.614	0.008-0.683	0.0062-0.614
Conc(mg/ml)/Vol (µl)	10/70	5/70	10/70	7.0/70	2-8/70	7.0/70
Temperature (K)	283	283	283	283	283	283
Structural parameters:						
I(0) (relative) (from *p(r)*)	0.03888±0.2E^{-4}	0.01552±0.1E^{-4}	0.032460±0.1E^{-4}	0.036840± 0.1E^{-4}	1.03 ±0.01	0.03857±0.11E^{-4}
R_g (Å) (from *p(r)*)	39.19±0.05	36.33±0.05	37.64±0.02	31.45±0.01	31.7±0.35	31.81±0.02
I(0) (cm^{-1}) (from Guinier)	0.03830±2.7E^{-5}	0.01539±1.2E^{-5}	0.03233±1.1E^{-5}	0.036814±2.3E^{-5}	1.01±0.01	0.038495±2.1E^{-5}
R_g (Å) (from Guinier)	37.00±0.05	35.00±0.06	36.60±0.02	30.90±0.04	29.5±0.4	31.10±0.04
D_{max} (Å)	160	144.5	145	108	110	114
Porod volume V_p (Å^3)	102101	94673	90194	69682	70673	71951
Excluded volume V_{ex} (Å^3) (Dammin/P1)	94218±552	93280±1289	97257±809	69900±506	74677±675	76173±446
Molecular mass determination (kDa)						
From Porod volume (V_p/1.7)	60.0	55.7	53.0	41.0	41.5	42.0
From excluded volume (V_{ex}/1.7)	55.4 ±0.1	54.8 ±0.7	57.2±0.1	41.1±0.1	43.9±0.4	44.8±0.1
From Q_R (q=0.25Å^{-1})	44.8	47.3	45.1	36.0	34.0	37.7
From sequence (including tag)	45.4	45.4	45.4	39.1	39.1	39.1
Modeling parameters						
Shape reconstruction	DAMMIN	DAMMIN	DAMMIN	DAMMIN	DAMMIN	DAMMIN
Symmetry	P1	P1	P1	P1	P1	P1
NSD (var) /# models	0.729(0.028)/10	0.653(0.015)/10	0.707(0.029)/10	0.639(0.018)/10	0.578(0.009)/10	0.595(0.015)/10
Flexibility assessment	FULCHER			FULCHER		
χ^2	11.4			6.5		
SAXSDB codes	SASDBS6	SASDBB9	SASDBT6	SASDBU6	SASDBC9	SASDBV6
Software employed						
Primary data reduction	FOXTROT					
Data processing	SCATTER GNOM DATASW PRIMUS					
Computation of model intensities	CRYSOL					
Model representations	PYMOL					

Abbreviations: *I(0)*, extrapolated scattering intensity at zero angle; R_g, radius of gyration calculated using either Guinier approximation (from Guinier) or the indirect Fourier transform package GNOM [from *p(r)*]; M_r, molecular mass; D_{max}, maximal particle dimension; V_p, Porod volume; V_{ex}, particle excluded volume. NA: experiment not performed.

*Momentum transfer $|q| = 4\pi\sin(\theta)/\lambda$.

Conc = concentration in mg/ml; Vol= volume in µl