Electronic Supplementary Material (ESI) for CrystEngComm.
This journal is © The Royal Society of Chemistry 2018

Electronic supplementary information

Application of protein crystallization methodologies to enhance solubility, stability and mono-dispersity of proteins

Ren-Bin Zhou[†], Xiao-Li Lu[†], Chen Dong, Fiaz Ahmad, Chen-Yan Zhang, Da-Chuan Yin[*]

[†]R.B.Z and X.L.L contributed equally to this work.

Key Laboratory for Space Bioscience & Biotechnology; School of Life Sciences, Northwestern Polytechnical University, Xi'an, Shaanxi, PR China

Correspondence to: yindc@nwpu.edu.cn

Content:

S1. Construction of expression vectors

The gene sequence encoding RBM3 was synthesized by GenScript (Nanjin g, China) based on the amino acid sequence of human RBM3 from UniprotKB (P98179). Using full-length RBM3 as a template, we amplified the recombinant protein with upstream primer (5'- cgcGGATCCATGTCCTCTGAAGAAGGAAA GCT -3') and downstream primer (5'- ccgCTCGAGTCAGTTGTCATAATTGTC TCTGT -3'). BamHI (GGATCC) and XhoI (CTCGAG) restriction sites were in serted upstream and downstream of the sequence, respectively. The vectors pG EX-6P-1, pET28a(+) and pRSet-B were digested by BamHI and XhoI to constr uct the plasmids. The plasmids were then transformed into the expression strai ns (Rosetta, BL21(DE3), and BL21(DE3)pLysS), and transformants were verifie d by DNA sequence analysis by Sangon Biotech (Shang Hai, China). After scr eening, the best expresser is the pET28a(+) plasmid in expression strain BL21p LysS(DE3).

S2. Protein expression and purification

A positive single colony was inoculated into 5 ml of LB media containing 100 μg/ml ampicillin and grown overnight at 37℃. The overnight cultures were diluted into 0.5 L of LB media containing 100 μg/ml ampicillin. The cells were grown at 37 ℃ with 225 rpm shaking until the optical density at 600 nm (OD600) was between 0.6 and 0.8, after which protein expression was induced with 0.25 mM isopropyl-thio-galactoside (IPTG) for 14 h at 22℃. The expression level was analyzed by 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE). Cells were harvested by centrifugation (6000 rpm, 30 min, 4℃). The cell pellets from 1 L of culture were suspended in 40 ml lysis buffer (20 mM Tris-HCl, 0.5 M NaCl, pH 7.5). Cells were lysed on ice by sonication using 30% power with 96 cycles of 3 s on and 5 s off. The cell lysate was cleared by centrifugation at 12,000 rpm for 30 min at 4 ℃. The supernatant was filtered through a 0.45 mm nylon syringe filter (Fisher). The clarified

supernatant was loaded onto a 4 ml Ni sepharose™ 6 fast flow column (GE Healthcare) and pre-equilibrated with lysis buffer at a flow rate of 1 ml/min using a peristaltic pump. The column was washed with 24 column volumes of wash buffer (20 mM Tris-HCl, 0.5 M NaCl, 100 mM imidazole, pH 7.5) and protein was eluted with elution buffer (20 mM Tris-HCl, 0.5 M NaCl, 300 mM imidazole, pH 7.5). The purity of the eluted fusion protein was analyzed by SDS–PAGE with Coomassie brilliant blue R-250 staining.

S3. Dynamic light scattering

Dynamic light scattering was employed for analyzing and monitoring the purified RBM3 aggregation. The sample was prepared at a concentration of 1 mg/ml in elution buffer and passed through a 0.2 μm syringe filter before analysis. The hydrodynamic radius of each sample was determined based on the scattered light intensity using Zetasizer Software 6.0.

Table S1. The formulations of Morpheus

Tube	Conc	Ligands	Conc	Buffer	pH	Conc	Precipitant
B6	0.09 M	Halogens	0.1 M	Buffer system 2	7.5	50% v/v	Precipitant Mix 2
D7	0.12 M	Alcohols	0.1 M	Buffer system 2	7.5	50% v/v	Precipitant Mix 3
E2	0.12 M	Ethylene glycols	0.1 M	Buffer system 1	6.5	50% v/v	Precipitant Mix 2
G7	0.1 M	Carboxylic acids	0.1 M	Buffer system 2	7.5	50% v/v	Precipitant Mix 3
F1	0.12 M	Monosaccharides	0.1 M	Buffer System 3	6.5	50 % v/v	Precipitant Mix 1
F2	0.12 M	Monosaccharides	0.1 M	Buffer System 3	6.5	50 % v/v	Precipitant Mix 2
F3	0.12 M	Monosaccharides	0.1 M	Buffer System 3	6.5	50 % v/v	Precipitant Mix 3
F4	0.12 M	Monosaccharides	0.1 M	Buffer System 3	6.5	50 % v/v	Precipitant Mix 4
F5	0.12 M	Monosaccharides	0.1 M	Buffer System 3	7.5	50 % v/v	Precipitant Mix 1
F6	0.12 M	Monosaccharides	0.1 M	Buffer System 3	7.5	50 % v/v	Precipitant Mix 2
F7	0.12 M	Monosaccharides	0.1 M	Buffer System 3	7.5	50 % v/v	Precipitant Mix 3
F8	0.12 M	Monosaccharides	0.1 M	Buffer System 3	7.5	50 % v/v	Precipitant Mix 4

F9	0.12 M	Monosaccharides	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 1
F10	0.12 M	Monosaccharides	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 2
F11	0.12 M	Monosaccharides	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 3
F12	0.12 M	Monosaccharides	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 4

Table S2. Mixes of additives used in Morpheus

Mix name	Composition
Halogens	0.3M Sodium fluoride; 0.3M Sodium bromide; 0.3M Sodium iodide
Alcohols	0.2M 1,6-Hexanediol; 0.2M 1-Butanol 0.2M 1,2-Propanediol; 0.2M 2-Propanol; 0.2M 1,4-Butanediol; 0.2M 1,3-Propanediol
Ethylene glycols	0.3M Diethylene glycol; 0.3M Triethylene glycol; 0.3M Tetraethylene glycol; 0.3M Pentaethylene glycol
Carboxylic acids	0.2M Sodium formate; 0.2M Ammonium acetate; 0.2M Sodium citrate tribasic dihydrate; 0.2M Sodium potassium tartrate tetrahydrate; 0.2M Sodium oxamate
Monosaccharides	0.2M D-Glucose; 0.2M D-Mannose; 0.2M D-Galactose; 0.2M L-Fucose; 0.2M D-Xylose; 0.2M N-Acetyl-D-Glucosamine

Table S3. Buffer systems used in Morpheus

Mix name	Conc.	pH @ 20 ℃	Composition
Buffer System 1	1.0 M	6.5	Imidazole; MES monohydrate (acid)
Buffer System 1	1.0 M	7.5	Sodium HEPES; MOPS (acid)

Table S4. Mixes of Precipitants used in Morpheus

Mix name	Composition
Precipitant Mix 2	40% v/v Ethylene glycol; 20 % w/v PEG 8000
Precipitant Mix 2	40% v/v Glycerol; 20% w/v PEG 4000

Table S5. Examples of pharmaceutical protein formulations that are quite similar to the crystallization conditions by simply removing the precipitant

Pharmaceutical proteins	Pharmaceutical formulation	Crystallization condition
Human interferon-β 1a	AVONEX®: 11.6 mM sodium acetate and 150 mM arginine hydrochloride buffer, plus 0.005% polysorbate 20 (Tween@-20), pH 4.8	7% PEG 4000, 50 mM sodium acetate, and 0.1 M ammonium acetate, at pH 4.6 [1]
Human insulin	HUMALOG®: 3.47 mg/mL insulin, 30 mM metacresol, 13 mM Na_2HPO_4, 16 mg/mL glycerol, pH 7.0-7.8	12.1 mg/mL insulin, 41 mM phenol, 13 mM Na_2HPO_4, KH_2PO_4, and NaSCN, pH 7.01-8.2 [2]
Aprotinin	TRASYLOL®: 1.4 mg/mL aprotinin and 154 mM NaCl, pH 4.5-6.5	40mg/ml aprotinin and NaCl 2-3 M, pH 4.5[3]
Teripatide	FORTEO®: 250 ug/mL teriparatide, 0.41mg/ml glacial acetic acid, 0.1 mg/ml sodium acetate (anhydrous), 45.4 mg/ml mannitol, 3 mg Metacresol, pH 4	20 mg/mL teriparatide, 20% glycerol, at a 1:1 ratio (v/v), with a solution containing 2.5 M ammonium sulfate, 5% isopropanol, and 0.1 M sodium acetate buffer pH 4.5 [4]

Lumiliximab (Anti-CD23 human IgG1)	HUMIRA®: 100 mg/mL Adalimumab, 42 mg/mL mannitol, 0.1% PS80, pH 5.2	55.2 mg/mL IDEC-152, 0.7 M ammonium sulfate, pH 5.0 [5]

The precipitants in crystallization condition are highlighted with green color.

	1		2		3		4		5		6		7		8		9		10		11		12	
A	0.084	0.106 0.043	0.05	0.106 0.049	0.052	0.106 0.042	0.124	0.106 0.106	0.049	0.106 0.059	0.093	0.106 0.046	0.065	0.106 0.047	0.086	0.106 0.044	0.052	0.106 0.043	0.05	0.106 0.046	0.048	0.106 0.044	0.114	0.106 0.046
B	0.085	0.106 0.047	0.336	0.106 0.068	0.045	0.106 0.047	0.053	0.106 0.042	0.082	0.106 0.046	0.056	0.106 0.044	0.053	0.106 0.053	0.077	0.106 0.044	0.06	0.106 0.047	0.048	0.106 0.043	0.05	0.106 0.047	0.129	0.106 0.043
C	0.063	0.106 0.044	0.049	0.106 0.042	0.067	0.106 0.045	0.08	0.106 0.042	0.069	0.106 0.044	0.059	0.106 0.042	0.05	0.106 0.051	0.056	0.106 0.043	0.059	0.106 0.044	0.048	0.106 0.043	0.05	0.106 0.044	0.089	0.106 0.043
D	0.07	0.106 0.044	0.052	0.106 0.047	0.052	0.106 0.044	0.067	0.106 0.047	0.123	0.106 0.045	0.052	0.106 0.046	0.055	0.106 0.044	0.062	0.106 0.057	0.109	0.106 0.048	0.053	0.106 0.047	0.047	0.106 0.044	0.115	0.106 0.046
E	0.06	0.106 0.076	0.049	0.106 0.042	0.049	0.106 0.044	0.058	0.106 0.041	0.055	0.106 0.047	0.045	0.106 0.043	0.047	0.106 0.047	0.051	0.106 0.043	0.066	0.106 0.045	0.048	0.106 0.045	0.049	0.106 0.045	0.05	0.106 0.041
F	0.064	0.106 0.043	0.051	0.106 0.042	0.061	0.106 0.046	0.05	0.106 0.044	0.312	0.106 0.043	0.047	0.106 0.043	0.053	0.106 0.043	0.051	0.106 0.042	0.088	0.106 0.046	0.044	0.106 0.042	0.048	0.106 0.045	0.058	0.106 0.042
G	0.164	0.106 0.043	0.048	0.106 0.041	0.046	0.106 0.043	0.052	0.106 0.042	0.077	0.106 0.043	0.046	0.106 0.042	0.052	0.106 0.042	0.053	0.106 0.043	0.079	0.106 0.044	0.047	0.106 0.041	0.052	0.106 0.043	0.066	0.106 0.042
H	0.096	0.106 0.05	0.059	0.106 0.044	0.05	0.106 0.043	0.052	0.106 0.043	0.089	0.106 0.059	0.054	0.106 0.044	0.07	0.106 0.046	0.063	0.106 0.047	0.107	0.106 0.072	0.047	0.106 0.043	0.054	0.106 0.044	0.055	0.106 0.044

Figure S1. Assessment of trypsin aggregation (4 mg/ml) at 340nm absorbance by Epoch.

	1		2		3		4		5		6		7		8		9		10		11		12	
A	0.087	0.106 0.043	0.064	0.106 0.049	0.063	0.106 0.042	0.084	0.106 0.106	0.09	0.106 0.059	0.073	0.106 0.046	0.062	0.106 0.047	0.363	0.106 0.044	0.077	0.106 0.043	0.06	0.106 0.046	0.058	0.106 0.044	0.22	0.106 0.046
B	0.083	0.106 0.047	0.063	0.106 0.068	0.061	0.106 0.047	0.098	0.106 0.042	0.083	0.106 0.046	0.061	0.106 0.044	0.059	0.106 0.053	0.085	0.106 0.044	0.08	0.106 0.047	0.082	0.106 0.043	0.068	0.106 0.047	0.086	0.106 0.043
C	0.079	0.106 0.044	0.059	0.106 0.042	0.058	0.106 0.045	0.094	0.106 0.042	0.08	0.106 0.044	0.073	0.106 0.042	0.069	0.106 0.051	0.071	0.106 0.043	0.139	0.106 0.044	0.063	0.106 0.043	0.057	0.106 0.044	0.098	0.106 0.043
D	0.083	0.106 0.044	0.067	0.106 0.047	0.062	0.106 0.044	0.094	0.106 0.047	0.084	0.106 0.045	0.077	0.106 0.046	0.06	0.106 0.044	0.085	0.106 0.057	0.082	0.106 0.048	0.07	0.106 0.047	0.059	0.106 0.044	0.093	0.106 0.046
E	0.088	0.106 0.076	0.067	0.106 0.042	0.061	0.106 0.044	0.102	0.106 0.041	0.097	0.106 0.047	0.083	0.106 0.043	0.15	0.106 0.047	0.069	0.106 0.043	0.065	0.106 0.045	0.068	0.106 0.045	0.058	0.106 0.045	0.094	0.106 0.041
F	0.087	0.106 0.043	0.065	0.106 0.042	0.06	0.106 0.046	0.11	0.106 0.044	0.068	0.106 0.043	0.071	0.106 0.043	0.06	0.106 0.043	0.074	0.106 0.042	0.074	0.106 0.046	0.081	0.106 0.042	0.06	0.106 0.045	0.088	0.106 0.042
G	0.09	0.106 0.043	0.067	0.106 0.041	0.055	0.106 0.043	0.087	0.106 0.042	0.074	0.106 0.043	0.069	0.106 0.042	0.061	0.106 0.042	0.072	0.106 0.043	0.088	0.106 0.044	0.064	0.106 0.041	0.058	0.106 0.043	0.089	0.106 0.042
H	0.112	0.106 0.05	0.068	0.106 0.044	0.07	0.106 0.043	0.107	0.106 0.043	0.106	0.106 0.059	0.07	0.106 0.044	0.069	0.106 0.046	0.086	0.106 0.047	0.126	0.106 0.072	0.072	0.106 0.043	0.07	0.106 0.044	0.102	0.106 0.044

Figure S2. Assessment of thaumatin aggregation (15 mg/ml) at 340nm absorbance by Epoch.

	1		2		3		4		5		6		7		8		9		10		11		12	
A	0.322	0.146 0.043	0.098	0.146 0.049	0.335	0.146 0.042	0.133	0.146 0.106	0.335	0.146 0.059	0.133	0.146 0.046	0.065	0.146 0.047	0.381	0.146 0.044	0.387	0.146 0.043	0.131	0.146 0.046	0.075	0.146 0.044	0.382	0.146 0.046
B	0.184	0.146 0.047	0.245	0.146 0.068	0.091	0.146 0.047	0.107	0.146 0.042	0.091	0.146 0.046	0.107	0.146 0.044	0.053	0.146 0.053	0.153	0.146 0.044	0.229	0.146 0.047	0.138	0.146 0.043	0.113	0.146 0.047	0.215	0.146 0.043
C	0.256	0.146 0.044	0.101	0.146 0.042	0.259	0.146 0.045	0.104	0.146 0.042	0.259	0.146 0.044	0.104	0.146 0.042	0.05	0.146 0.051	0.273	0.146 0.043	0.169	0.146 0.044	0.105	0.146 0.043	0.079	0.146 0.044	0.241	0.146 0.043
D	0.337	0.146 0.044	0.119	0.146 0.047	0.22	0.146 0.044	0.112	0.146 0.047	0.22	0.146 0.045	0.112	0.146 0.046	0.055	0.146 0.044	0.166	0.146 0.057	0.173	0.146 0.048	0.108	0.146 0.047	0.083	0.146 0.044	0.239	0.146 0.046
E	0.312	0.146 0.076	0.137	0.146 0.042	0.264	0.146 0.044	0.116	0.146 0.041	0.264	0.146 0.047	0.116	0.146 0.043	0.047	0.146 0.047	0.178	0.146 0.043	0.372	0.146 0.045	0.109	0.146 0.045	0.077	0.146 0.045	0.189	0.146 0.041
F	0.451	0.146 0.043	0.111	0.146 0.042	0.231	0.146 0.046	0.223	0.146 0.044	0.231	0.146 0.043	0.223	0.146 0.043	0.053	0.146 0.043	0.157	0.146 0.042	0.313	0.146 0.046	0.109	0.146 0.042	0.088	0.146 0.045	0.161	0.146 0.042
G	0.242	0.146 0.043	0.106	0.146 0.041	0.184	0.146 0.043	0.116	0.146 0.042	0.184	0.146 0.043	0.116	0.146 0.042	0.052	0.146 0.042	0.15	0.146 0.043	0.172	0.146 0.044	0.11	0.146 0.041	0.089	0.146 0.043	0.157	0.146 0.042
H	0.667	0.146 0.05	0.143	0.146 0.044	0.332	0.146 0.043	0.127	0.146 0.043	0.332	0.146 0.059	0.127	0.146 0.044	0.07	0.146 0.046	0.166	0.146 0.047	0.285	0.146 0.072	0.119	0.146 0.043	0.094	0.146 0.044	0.164	0.146 0.044

Figure S3. Assessment of catalase aggregation (10 mg/ml) at 340nm absorbance by Epoch.

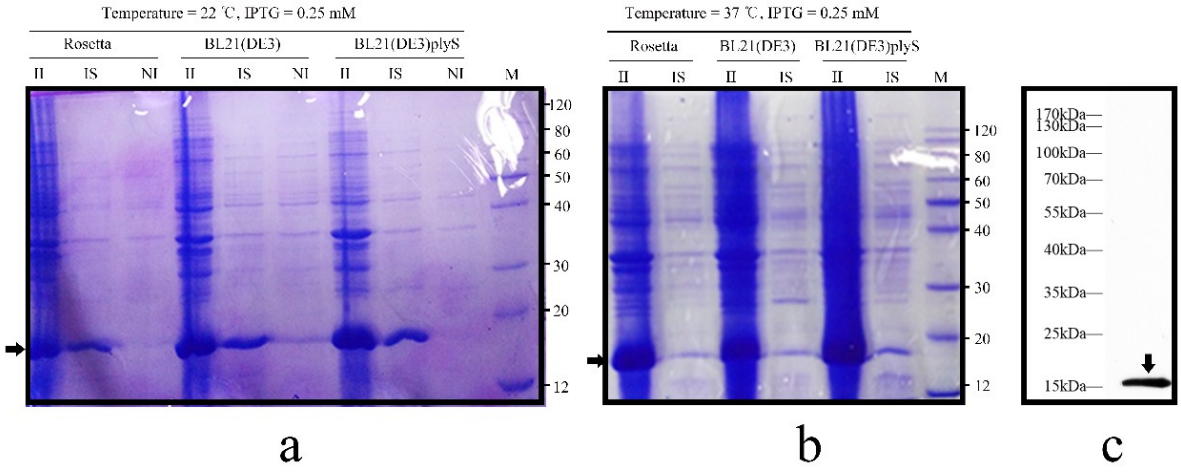

Figure S4. Screening of expression conditions for recombinant RBM3 induced with 0.25 mM IPTG at 22 °C (**a**) and 37 °C (**b**) in different *E. coli* host strains. Each condition was separated by SDS-PAGE, and then stained with Coomassie Brilliant Blue (CBB). Lane II, induced insoluble fraction; Lane IS, induced soluble fraction; Lane NI, non-induced fraction; Lane M, molecular marker (shown on the right)(**C**). The optimize expression condition of recombinant RBM3 is induced with 0.25 mM IPTG at 22 °C in BL21(DE3)plyS strain and confirmed by western blotting with an anti-RBM polyclonal antibody. The arrow shows the size of RBM3.

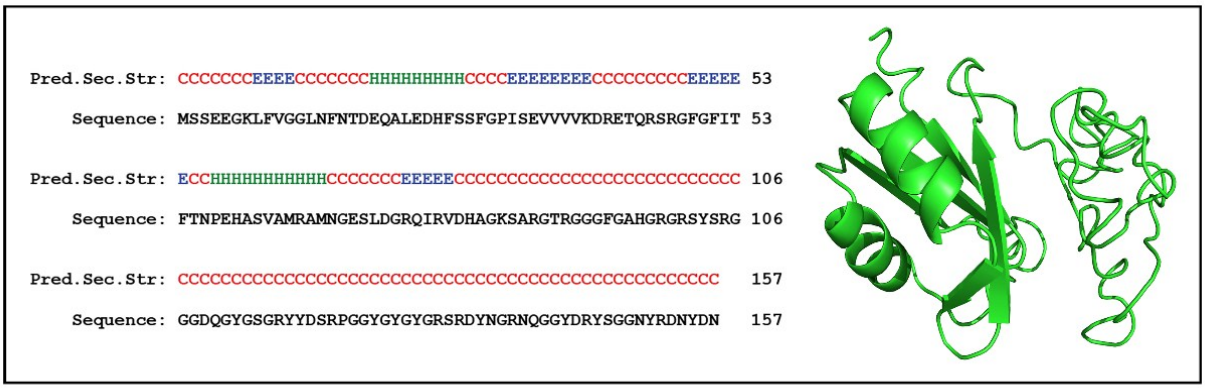

Figure S5. The intrinsically disordered region of RBM3 in the C-terminal. Left, the second structure of RBM3 predicted by PSIPRED server (H stands for α-helix, E stands for β-sheet and C stands for coil).[6] Right, the tertiary structure of RBM3 modeled I-TASSER server.[7]

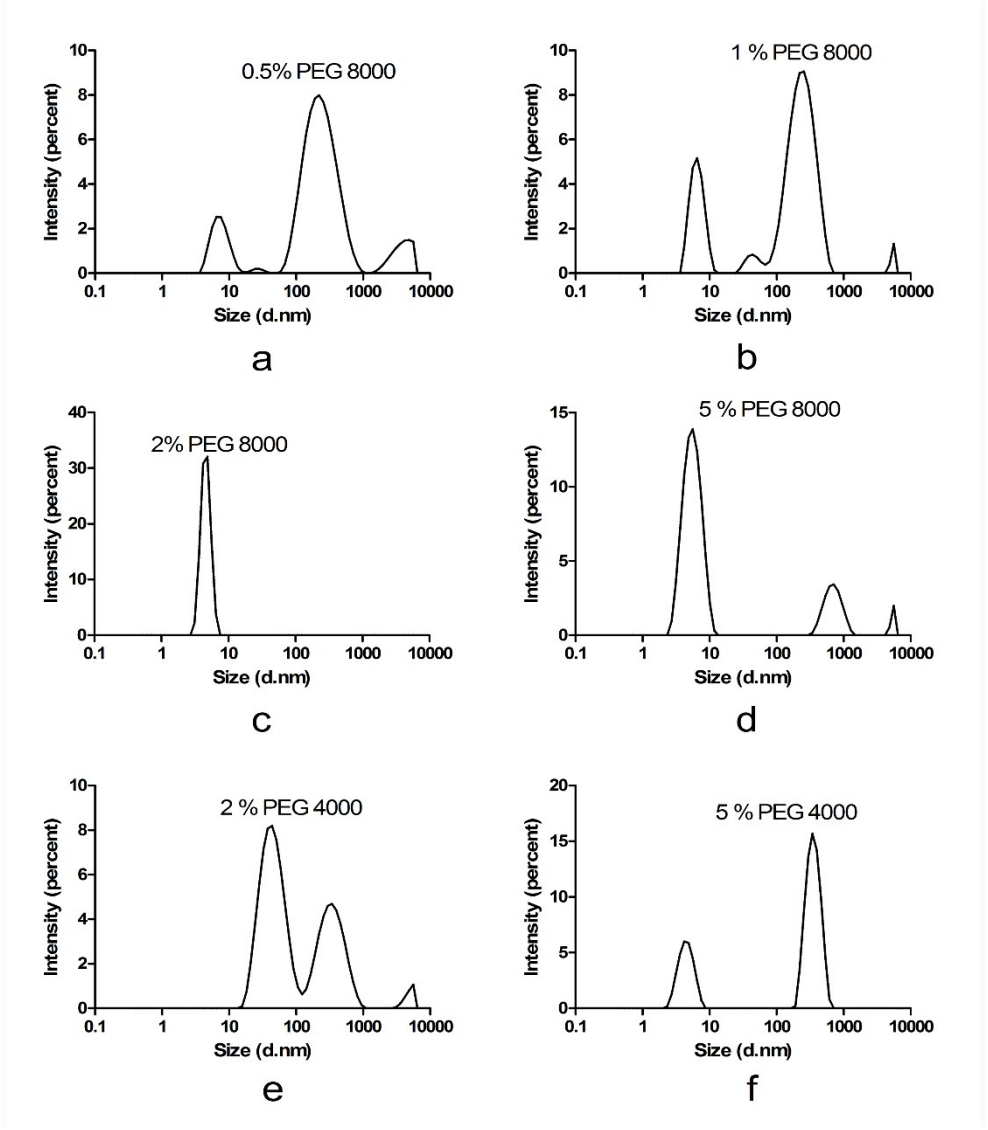

Figure S6. The size distribution of RBM3 by Intensity by dynamic light scattering. (a) The size distribution of RBM3 is measured in the presence of 0.05% PEG 8000 in the purification buffer; (b) The size distribution of RBM3 is measured in the presence of 1% PEG 8000 in the purification buffer; (c) The size distribution of RBM3 is measured in the presence of 2% PEG 8000 in the purification buffer; (d) The size distribution of RBM3 is measured in the presence of 5% PEG 8000 in the purification buffer; (e) The size distribution of RBM3 is measured in the presence of 2% PEG 4000 in the

purification buffer; (f) The size distribution of RBM3 is measured in the presence of 5% PEG 4000 in the purification buffer.

1 M. Karpusas, M. Nolte, C. B. Benton, W. Meier, W. N. Lipscomb and S. Goelz, *Proceedings of the National Academy of Sciences*, 1997, **94**, 11813-11818.

2 F. Karavassili, A. E. Giannopoulou, E. Kotsiliti, L. Knight, M. Norrman, G. Schluckebier, L. Drube, A. N. Fitch, J. P. Wright and I. Margiolaki, *Acta Crystallographica Section D: Biological Crystallography*, 2012, **68**, 1632-1641.

3 S. Tanaka, M. Ataka, K. Onuma, J.-P. Astier and S. Veesler, *Journal of crystal growth*, 2002, **237**, 289-294.

4 L. Jin, S. L. Briggs, S. Chandrasekhar, N. Y. Chirgadze, D. K. Clawson, R. W. Schevitz, D. L. Smiley, A. H. Tashjian and F. Zhang, *J Biol Chem*, 2000, **275**, 27238-27244.

5 R. A. Lewus, P. A. Darcy, A. M. Lenhoff and S. I. Sandler, *Biotechnology progress*, 2011, **27**, 280-289.

6 L. J. McGuffin, K. Bryson and D. T. Jones, *Bioinformatics*, 2000, **16**, 404-405.

7 J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson and Y. Zhang, *Nature methods*, 2015, **12**, 7-8.