Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information

Integrative structural modeling of the multidomain Polo-like kinase 1

Hao Ruan,[a] Janna Kiselar,[d] Weilin Zhang,[a] Siyang Li,[b] Ruoyao Xiong,[a] Ying Liu,[a,b] Sichun Yang,[d*] Luhua Lai,[a,b,c*]

This SI includes:

Supplementary Figs. S1 to S8

Supplementary Tables S1 and S2

Figure S1

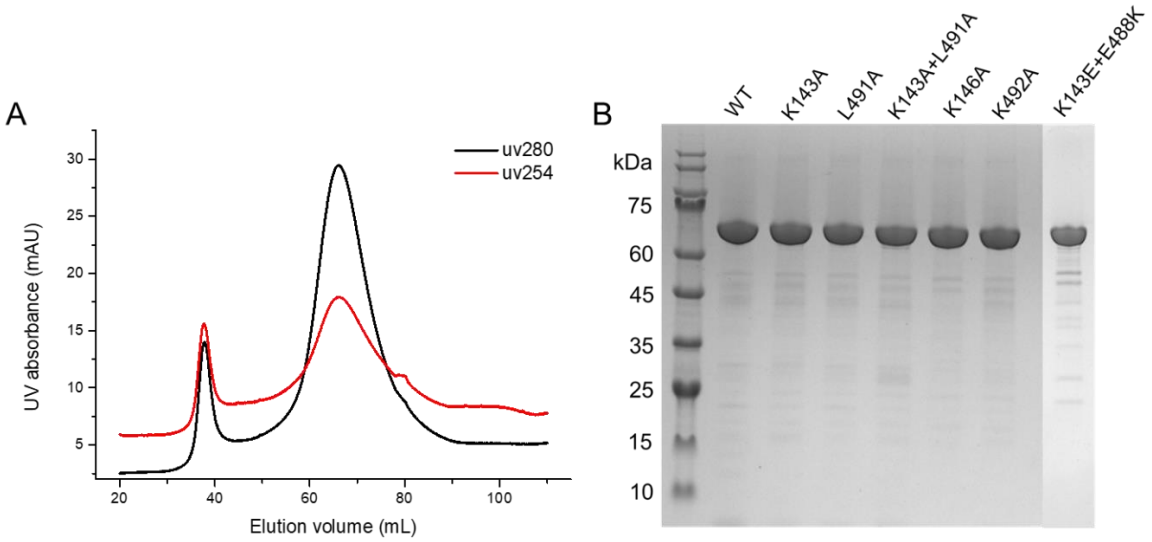

Fig. S1. Purification and characterization of human PLK1 and related mutants. (A) Elution profile of the final Sephacryl S-200 gel-filtration column. Human PLK1 was eluted as a single peak. Absorbance at 254 nm is shown in red and at 280 nm is shown in black. (B) SDS-PAGE analysis of hPLK1 and molecular mass markers run under reducing conditions.

Figure S2

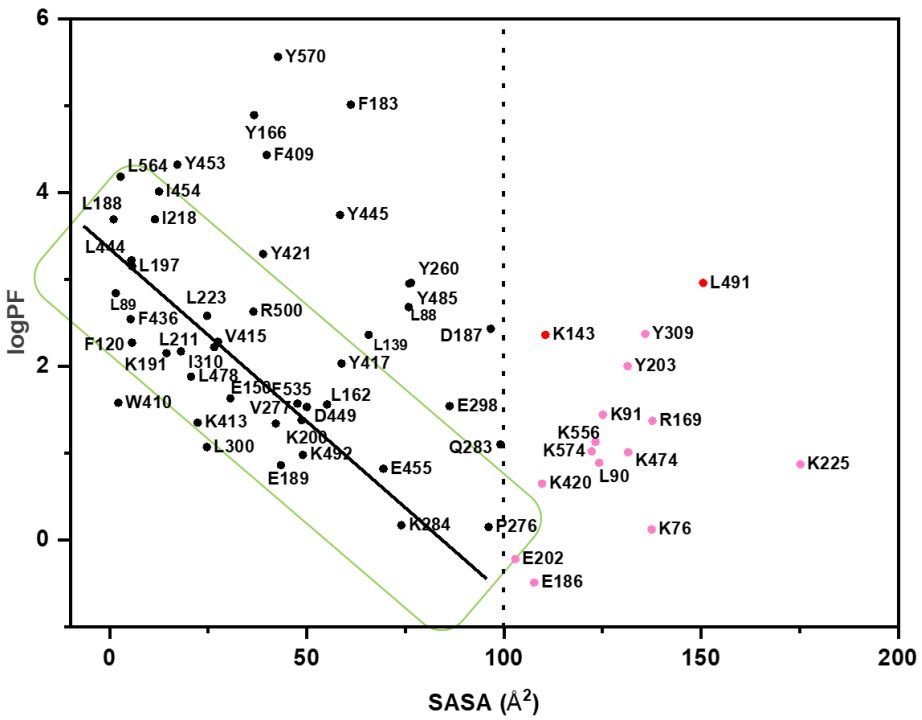

Fig. S2. Measured logPF values against the SASA values of individual domains of the KD and the PBD. Residues located in disordered regions are removed. The area of the greatest concentration is shown in green rectangle, where residues were selected for fitting between logPF values and the SASA values. Residues with SASA values more than 100 Å^2 are shown in red or pink dots. Among these residues, K143 and L491 have the highest signal-noise ratio (Y203, K225 and Y309 are located in the loop).

Figure S3

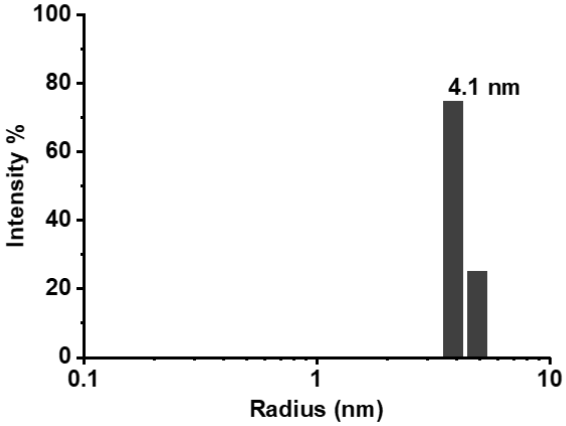

Fig. S3. Characterization of hPLK1 using Dynamic Light Scattering. The results indicate monodisperse of the hPLK1 sample under 20 mM sodium phosphate, 50 mM NaCl, 0.05 mM TCEP, pH 7.4 condition.

Figure S4

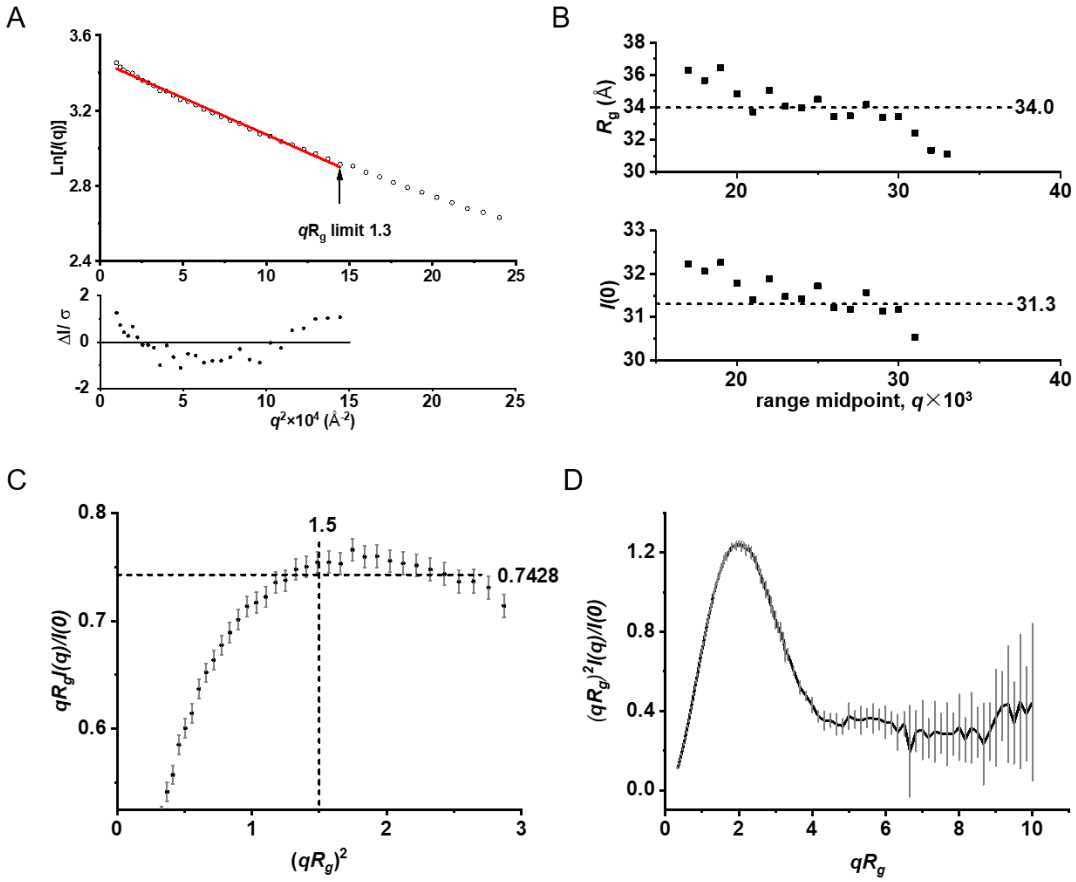

Fig. S4. SAXS data analysis by *Primus*.[1] (A) Human PLK1 had a linear Guinier region. (B) Estimates of R_g and $I(0)$ using 6 point q-ranges within the Guinier region. (C) The dimensionless GPA plot revealed that the peak of hPLK1 sample had a value of (1.7448, 0.7660), which was near to the theoretical position of (1.5, 0.7428). (D) Dimensionless Kratky plot of hPLK1.

Figure S5

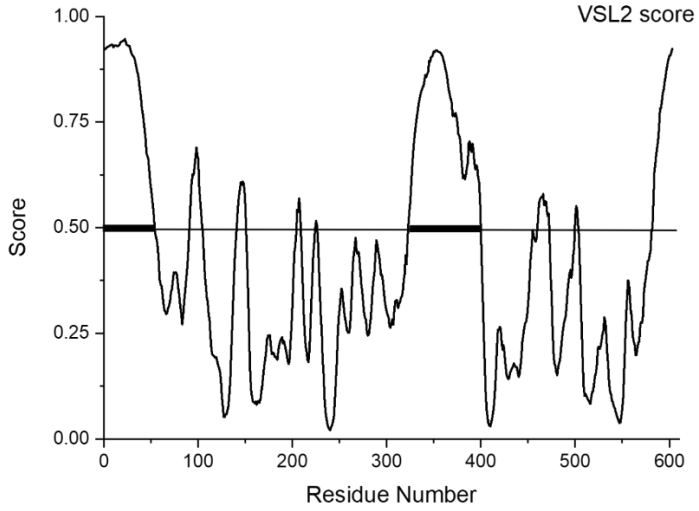

Fig. S5. Prediction of the N-terminal disordered region. The disorder scores using the VSL2 algorithm calculated by PONDR (Predictor of Natural Disordered Regions) [2, 3] are shown as function of residue indices.

Figure S6

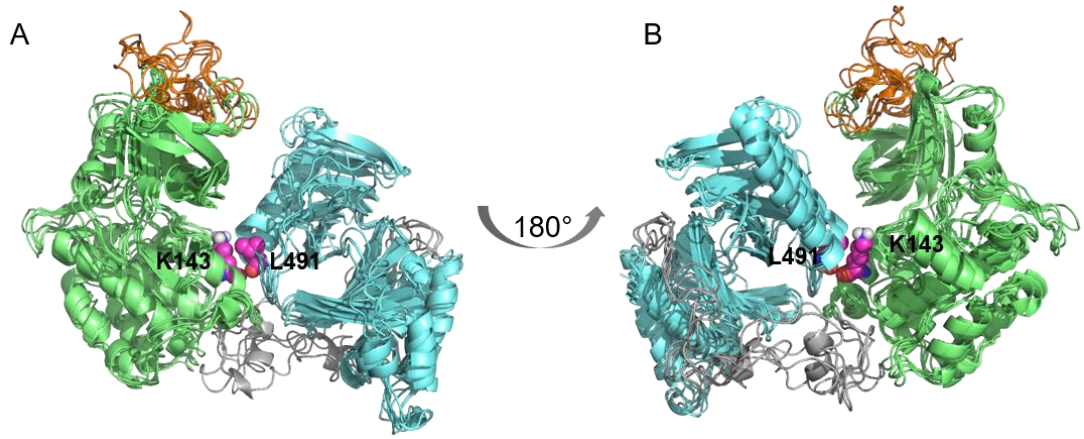

Fig. S6. Ensemble-structures of the hPLK1 complex with $\chi^2 = 2.4$, where the KD is in green and the PBD is in cyan, while the KD-PBD connecting linker in gray and the N-terminal disordered region in orange.

Figure S7

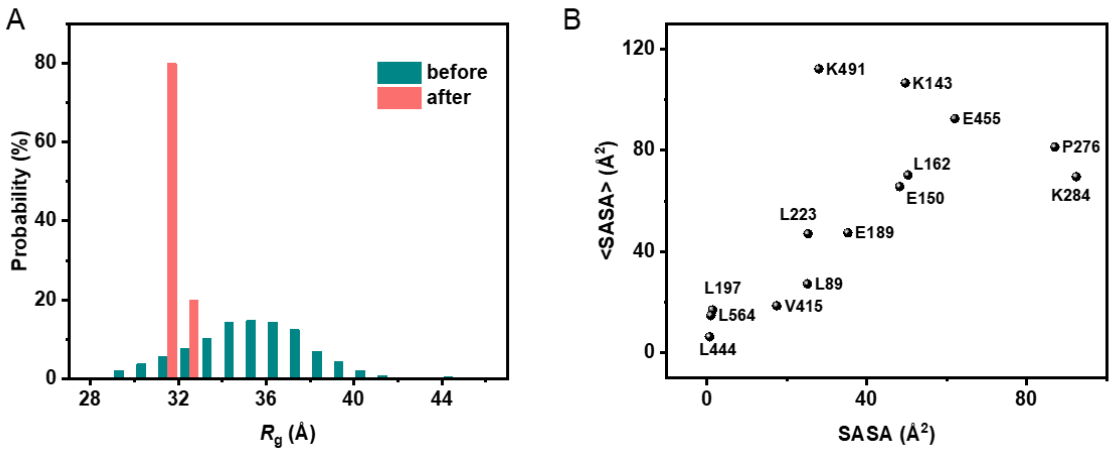

Fig. S7. Structural fitting by integrative selection against SAXS and footprinting data. (A) Density of R_g of PLK1 ensemble before (blue) and after (red) integrative fitting. (B) The ensemble-averages of solvent accessible surface area values of 14 footprinting-probed residues before (<SASA>) and after (SASA) integrative fitting.

Figure S8

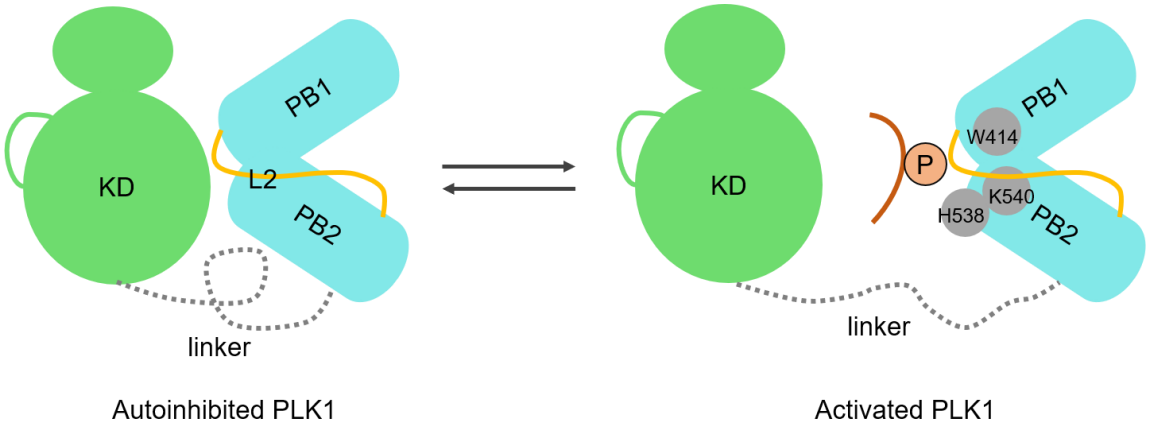

Fig. S8. Structural model of PLK1 regulation. In the autoinhibited state, the PBD primarily interacts with the C-lobe of the KD in PLK1 via the N-terminal end of the linker region (L2, shown in orange) between the PB1 and PB2 motifs. The phosphopeptide binding may weaken or break the KD-PBD interaction, thus activating PLK1.

Table S1. Related to Figure 2. The measured footprinting rates, protection factors and fitting results of full-length hPLK1. There are two groups of residues: the control group contains 12 residues for a linear regression and the other group contains 2 residues with large changes in logPF and SASA values. Residues located in disordered regions were removed at this step. SASA values were calculated by VMD in individual domains. The $SASA^{fit}$ values were derived from the linear fit of measured logPF values and calculated SASA values.

Group	**Residue**	k_{fp} (s^{-1})	**logPF**	**SASA (Å^2)**	**$SASA^{fit}$ (Å^2)**
Control (12)	L89	0.258 ± 0.0098	2.84 ± 0.04	1.55	28.23
	E150	0.135 ± 0.0031	1.63 ± 0.02	30.66	40.96
	L162	0.92 ± 0.0085	1.56 ± 0.01	55.17	45.86
	E189	0.293 ± 0.017	0.86 ± 0.06	43.46	77.45
	L197	0.189 ± 0.0094	3.15 ± 0.05	5.69	1.83
	L223	0.333 ± 0.015	2.58 ± 0.05	24.71	22.83
	P276	0.86 ± 0.029	0.15 ± 0.03	96.15	105.32
	K284	1.86 ± 0.085	0.17 ± 0.05	73.99	86.69
	V415	0.195 ± 0.013	2.28 ± 0.07	27.47	41.13
	L444	0.176 ± 0.0044	3.22 ± 0.03	5.55	0
	E455	0.305 ± 0.011	0.82 ± 0.04	69.43	61.01
	L564	0.067 ± 0.00054	4.18 ± 0.01	2.8	2.46
Prediction (2)	K143	0.207 ± 0.0081	2.36 ± 0.04	110.55	49.31
	L491	0.228 ± 0.013	2.96 ± 0.06	146.56	18.65
Removed (25)	V25	2.13 ± 0.12	-0.11 ± 0.06	-	-
	P28	3.26 ± 0.21	-1.18 ± 0.06	-	-
	P35	0.256 ± 0.012	1.36 ± 0.05	-	-
	P36	4.28 ± 0.25	-1.45 ± 0.06	-	-
	V43	1.62 ± 0.052	0.16 ± 0.03	-	-
	V45	1.69 ± 0.050	0.12 ± 0.03	-	-
	D46	0.81 ± 0.033	-0.66 ± 0.04	-	-
	P47	0.654 ± 0.023	0.42 ± 0.04	-	-
	R48	1.63 ± 0.039	0.58 ± 0.02	-	-
	F325	0.645 ± 0.019	2.85 ± 0.03	-	-
	I327	0.437 ± 0.025	2.31 ± 0.06	-	-
	P329	1.53 ± 0.049	-0.43 ± 0.03	-	-
	L332	3.86 ± 0.14	0.13 ± 0.04	-	-
	P334	0.633 ± 0.027	0.46 ± 0.04	-	-
	R337	1.02 ± 0.038	1.04 ± 0.04	-	-
	K338	1.38 ± 0.033	0.47 ± 0.02	-	-
	L340	2.62 ± 0.12	0.52 ± 0.05	-	-
	L343	1.19 ± 0.046	1.31 ± 0.04	-	-
	K345	1.67 ± 0.11	0.28 ± 0.07	-	-
	V342	4.25 ± 0.25	-0.81 ± 0.06	-	-
	L347	1.17 ± 0.086	1.32 ± 0.07	-	-
	P350	0.259 ± 0.0091	1.35 ± 0.04	-	-
	L351	0.18 ± 0.011	3.20 ± 0.06	-	-
	E353	0.507 ± 0.018	0.31 ± 0.04	-	-
	P355	0.20 ± 0.039	1.61 ± 0.20	-	-
Others (48)	K76	1.95 ± 0.084	0.12 ± 0.04	137.48	-
	L88	0.302 ± 0.013	2.68 ± 0.04	75.84	-
	L90	1.80 ± 0.070	0.89 ± 0.04	124.15	-
	K91	0.522 ± 0.032	1.44 ± 0.06	125.13	-
	F120	1.16 ± 0.082	2.27 ± 0.07	5.65	-
	L139	0.414 ± 0.012	2.36 ± 0.03	65.70	-
	K143	0.207 ± 0.0081	2.36 ± 0.04	110.55	-
	Y166	0.090 ± 0.0020	4.89 ± 0.02	36.67	-
	R169	0.738 ± 0.020	1.37 ± 0.03	137.70	-

Group	**Residue**	k_{fp} (s^{-1})	**logPF**	**SASA (Å^2)**	**SASA^{fit} (Å^2)**
	F183	0.0748 ± 0.0021	5.01 ± 0.03	61.17	-
	E186	1.13 ± 0.040	-0.49 ± 0.04	107.68	-
	D187	0.037 ± 0.0022	2.43 ± 0.06	96.67	-
	L188	0.11 ± 0.011	3.69 ± 0.10	1.01	-
	K191	0.257 ± 0.012	2.15 ± 0.05	14.45	-
	K200	0.551 ± 0.020	1.38 ± 0.04	48.69	-
	E202	0.858 ± 0.022	-0.22 ± 0.03	102.89	-
	Y203	1.62 ± 0.077	2.00 ± 0.05	131.38	-
	L211	0.50 ± 0.015	2.17 ± 0.03	18.10	-
	I218	0.11 ± 0.0061	3.69 ± 0.06	11.49	-
	K225	0.92 ± 0.081	0.87 ± 0.09	175.21	-
	Y260	0.621 ± 0.024	2.96 ± 0.04	76.44	-
	V277	0.498 ± 0.016	1.34 ± 0.03	42.14	-
	Q283	0.219 ± 0.0079	1.10 ± 0.04	99.09	-
	E298	0.148 ± 0.010	1.54 ± 0.07	86.22	-
	L300	1.51 ± 0.084	1.07 ± 0.06	24.67	-
	Y309	1.12 ± 0.055	2.37 ± 0.05	135.88	-
	I310	0.48 ± 0.020	2.22 ± 0.04	26.55	-
	F409	0.133 ± 0.0016	4.43 ± 0.01	39.87	-
	W410	3.58 ± 0.066	1.58 ± 0.02	2.17	-
	K413	0.570 ± 0.021	1.35 ± 0.04	22.32	-
	Y417	1.58 ± 0.085	2.03 ± 0.05	58.83	-
	K420	1.15 ± 0.021	0.65 ± 0.02	109.68	-
	Y421	0.449 ± 0.017	3.29 ± 0.04	38.95	-
	F436	0.881 ± 0.023	2.54 ± 0.03	5.36	-
	Y445	0.285 ± 0.0067	3.74 ± 0.02	58.46	-
	D449	0.091 ± 0.0032	1.53 ± 0.04	50.04	-
	Y453	0.159 ± 0.0033	4.32 ± 0.02	17.17	-
	I454	0.080 ± 0.0014	4.01 ± 0.02	12.55	-
	K474	0.805 ± 0.067	1.01 ± 0.08	131.52	-
	L478	0.671 ± 0.021	1.88 ± 0.03	20.66	-
	Y485	0.625 ± 0.066	2.95 ± 0.11	76.02	-
	L491	0.228 ± 0.013	2.96 ± 0.06	150.56	-
	K492	0.829 ± 0.054	0.98 ± 0.07	49.03	-
	R500	0.209 ± 0.0096	2.63 ± 0.05	36.49	-
	F535	2.34 ± 0.15	1.57 ± 0.06	47.61	-
	K556	0.712 ± 0.082	1.13 ± 0.12	123.21	-
	Y570	0.046 ± 0.0034	5.56 ± 0.07	42.70	-
	K574	0.79 ± 0.021	1.02 ± 0.03	122.27	-

*PF is calculated using the webserver at http://www.theyanglab.org/protection.html

Table S2. Molecular weight estimation from SAXS

Software ATSAS/Primus Molecular Weight Wizard		
Q_p	**MoW**	**Vc**
q_{max} [Å^{-1}] 0.205	q_{max} [Å^{-1}] 0.300	q_{max} [Å^{-1}] 0.300
MW [Da] 82483	V [Å^3] 94369	Vc 539
	MW [Da] 77858	MW [Da] 69214
Bayesian Inference		
MW Estimate [Da] 76350		
MW Probability [%] 47.54		
Credibility Interval [Da] [71450, 79600]		
Credibility Interval Probability [%] 94.66		
Calculated Mw from sequence [Da] 68267		

Supplementary References

1 P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch and D. I. Svergun, *J. Appl. Crystallogr.*, 2003, **36**, 1277-1282.

2 X. Li, P. Romero, M. Rani, A. K. Dunker and Z. Obradovic, *Genome Inform Ser Workshop Genome Inform*, 1999, **10**, 30-40.

3 A. K. Dunker, J. D. Lawson, C. J. Brown, R. M. Williams, P. Romero, J. S. Oh, C. J. Oldfield, A. M. Campen, C. M. Ratliff, K. W. Hipps, J. Ausio, M. S. Nissen, R. Reeves, C. Kang, C. R. Kissinger, R. W. Bailey, M. D. Griswold, W. Chiu, E. C. Garner and Z. Obradovic, *J. Mol. Graph. Model.*, 2001, **19**, 26-59.