**Supplementary material**

**The mitochondrial single-stranded DNA binding proteins from *S. cerevisiae*, Rim1, does not form stable homo-tetramers and binds DNA as a dimer of dimers.**

Saurabh P. Singh, Vandna Kukshal, Paolo De Bona, Edwin Antony and Roberto Galletto

**Protein purification.**

For purification of Rim1s, cells were harvested and resuspended in Buffer A (50 mM Tris–HCl pH 8.3, 400 mM NaCl, 1 mM EDTA, 15 mM spermidine, 10% (v/v) glycerol and 1 mM PMSF) and disrupted by sonication. 0.2 % (w/v) Polymin P was added slowly to the clarified supernatant. Following centrifugation, the pellet was gently extracted with 50mM Tris-HCl pH 8.3, 600mM NaCl, 20% (v/v) glycerol, 1mM EDTA. Following centrifugation, the supernatant was slowly precipitated by addition of 0.4 g/mL of (NH4)2SO4. After centrifugation, the pellet was dissolved in Buffer B (20 mM HEPES pH 7.4, 10 % (v/v) glycerol, 1 mM EDTA) + 100 mM NaCl. After extensive dialysis against Buffer B, Rim1 was loaded onto a strong anion exchange column (HighQ, Bio-Rad) and eluted with Buffer B using a salt gradient from 100 mM to 1 M NaCl. The fractions containing Rim1 protein were pooled and dialyzed against Buffer C (50 mM Tris-HCl pH 8.3, 300 mM NaCl, 10% (v/v) glycerol, 1 mM EDTA). Finally, Rim1 in Buffer C was loaded onto a single-strand DNA cellulose column, washed with Buffer C + 600 mM NaCl, and eluted with 50 mM Tris-HCl 8.3, 3 M NaCl, 20% (v/v) glycerol, 1 mM EDTA. The purified Rim1s (Fig. S1) were dialyzed in Storage Buffer (20 mM HEPES pH 7.4, 500 mM NaCl, 1 mM EDTA, 50% (v/v) glycerol) and stored at -20 °C.

For purification of *Hsmt*SSB, cells were opened in Buffer D (30 mM HEPES pH 7.4, 50 mM NaCl, 1 mM EDTA, 2% (v/v) glycerol and 1 mM PMSF). The clarified supernatant was loaded onto a weak anion exchange column (DEAE, BioRad) equilibrated in Buffer D, collected in the flow-through and directly loaded onto an Affi-Gel Blue column (BioRad) equilibrated in Buffer E (30 mM HEPES pH 7.4, 50 mM KCl, 1 mM EDTA, 1mM DTT). After washing with Buffer E + 800 mM KCl, *Hsmt*SSB was eluted with Buffer E + 1.5 M KSCN and was dialyzed in Buffer C. The protein was then loaded onto a ssDNA cellulose column and eluted as for Rim1. Finally, *Hsmt*SSB was dialyzed in Buffer B + 300 mM NaCl, diluted to 100 mM NaCl, loaded onto a POROS 50 HE heparin column (Applied Biosystems) and eluted with Buffer B using a salt gradient from 100 mM to 1 M NaCl. The fractions containing the purified *Hsmt*SSB (Fig. S1) were pooled and dialyzed in Storage Buffer (20 mM HEPES pH 7.4, 500 mM NaCl, 1 mM EDTA, 50% (v/v) glycerol) and stored at -20 °C.

**Structure determination and diffraction data processing.**

Diffraction data for the native and seleno-L-methionine labeled Rim1 crystals were collected under cryogenic conditions (100 K) at the peak wavelength of the selenium K absorption edge (0.97845 Å) at the Beamline ALS 4.2.2. Diffraction data were processed using XDS/XSCALE (49,50).

Seleno-L-methionine labeled Rim1 in crystal Form 2 diffracted to 2.8 Å resolution with 8 monomers in an asymmetric unit. The initial phase determination and model building were accomplished by the Single-wavelength Anomalous Dispersion (SAD) method with anomalous signals of 15 seleno-methionines. The Autosol module embedded in the PHENIX software package (PHENIX-1.11.1-2575) located the Selenium sites and provided the initial structural model and phasing. Iterative rounds of model building and refinement were then carried out using Coot (51) and phenix.refine (52) (PHENIX-1.11.1-2575), respectively. Multiple cycles of model building, refinement and phase improvement produced a model with extended loops which was subsequently refined ([Table S1](http://www.sciencedirect.com/science/article/pii/S0022283613001733#t0005)). The model for native Rim1 in crystal Form 2 was produced by refining the native data set with the model from seleno-L-methionine labeled Rim1 in which the seleno-methionines were replaced by methionine. The native Rim1 in crystal Form 1 with C2221 space group has 4 monomers in asymmetric unit and was solved by using molecular replacement with one monomer as model. Molecular replacement was performed using Phaser (54) within the PHENIX suite, with the Rim1 monomer as an initial model. Iterative rounds of model building and refinement were performed using Coot and phenix.refine. The quality of the final model was assessed using MolProbity (55), in addition to PROCHECK [(56)](http://www.sciencedirect.com/science/article/pii/S0022283613001733#bb0160) and SFCHECK (57) from the CCP4 suite (51). Figures for the manuscript were produced using PyMOL (58).

**Binding of a dimer of Rim1 to dTn-Cy3-T (n=15, 20, 28 nt) is an example of a limiting case.**

Based on an occluded site-size of 14 nt (Fig. 3B), no more than one Rim1 dimer (or two protomers within a single tetramer) should bind to ssDNA shorter than 28 nt. Figure S3A shows the fluorescence increase of dTn-Cy3-T (n = 15, 20 and 28 nt) as a function of the ratio of the total concentration of Rim1 monomers to the total concentration of DNA, in Buffer HK150 in the absence and presence of 5 mM MgCl2. The data strongly suggest that at saturation two Rim1 monomers (i.e. one dimer) bind to these lengths of ssDNA. Taken at face value, the data would argue that at the low DNA concentrations used in fluorescence experiments Rim1 binds to DNA exclusively as a dimer. However, binding of only a Rim1 dimer to DNA would be in apparent disagreement with the results from analytical equilibrium ultracentrifugation experiments in Figure 3D, indicating that a tetramer of Rim1 forms on dT15 and dT20 even for protein concentrations that are lower than the DNA concentration. While this discrepancy may be attributed to the overall higher concentration of DNA used in Figure 3D, we reasoned that one alternative explanation for the apparent stoichiometry of one Rim1 dimer per DNA observed in Figure 4A may reside in the fluorescent signal monitored. For example, while the equilibrium fluorescence DNA binding experiments in Figure 4A predominantly report on binding of a Rim1 dimer, they may be silent to binding of additional Rim1 molecules. This situation may arise under the limiting condition where the monitored signal is dominated by the high affinity binding of the first ligated species, followed by lower affinity binding of a second species that does not significantly contribute to the signal. In this limiting case the degree of binding as a function of the observed signal would report only on the stoichiometry of the first ligand bound (Fig. S3B). The observation of an additional phase in the fluorescence signal when Cy3 is moved from the 3’-end to the 5’-end (Fig. S3C), or when Alexa546 fluorophore is used (Fig. S3D), suggests that this is indeed the case. The DNA binding data presented in Figure 4A using carboxy-fluorescein labeled ssDNA (FAM-dT20) confirm this conclusion.

**Dimer-tetramer oligomerization model.**

(s1)

(s2)

(s3)

(s4)

where *Yobs* is either the observed sedimentation coefficient or molecular weight, and *YD* and *YT* are either sedimentation coefficient or the molecular weight of the dimer and tetramer, respectively. Fits and simulations were generated using Eqs. s1 and s4 with Scientist (Micromath), and the total Rim1 concentration plotted in monomer concentration.

**1:1 binding model.**

(s5)

(s6)

The relative change in fluorescence intensity (*F*) was expressed as:

(s7)

Because of negligible change in total intensity, the change in fluorescence anisotropy was expressed as:

(s8)

Where and are the fluorescence anisotropy of the DNA, free and protein-bound. Fits were generated using Eqs. S6 and s7 or s8 with Scientist (Micromath, Saint Louis), with the Rim1 concentration expressed either in dimers or tetramers.

**2:1 binding model.**

For analysis using a 2:1 DNA binding model the Rim1 concentration is expressed in dimers (P2).

(s9)

The data showed a change in total intensity that was accounted for by introducing the effect on fluorescence intensity of the first (F*P2S*) and second (F*P4S*) dimer bound, as follows:

(s10)  
 (s11)

(s12)

(s13)

(s14)

where , and are the fluorescence anisotropies of the free DNA, and DNA bound to one and two dimers. The data we analyzed using Eqs. s9-14 with Scientist (Micromath, Saint Louis). While is determined, there are six variables. , , Q1, and Q2 were fixed to values estimated by graphical inspection. Also, fitting for K and L is complicated by the fact that: 1) K is very high at the solution conditions tested; 2) K and L differ by at least two orders of magnitude. Therefore, we fitted the data at the lowest DNA concentration to better estimate K and obtain an initial estimate of L. Then, K was maintained constant for the higher DNA concentrations and L adjusted within a factor two if needed. In the main text we report Kx and Lx, where the subscript x indicates the length of the ssDNA in nt. The parameters used in the Figures in the Main Text are as follow:

Figure 4A (Rim1 binding to FAM-dT20):

[DNA] = 100nM: K = (1.6 ± 0.7) x 109 M-1, L = (3.2 ± 0.3) x 105 M-1

[DNA] = 200nM: K = 1.6 x 109 M-1, L = 2.8 x 105 M-1

[DNA] = 400nM: K = 1.6 x 109 M-1, L = 2 x 105 M-1

[DNA] = 800nM: K = 1.6 x 109 M-1, L = 1.5 x 105 M-1

Figure 5D (Rim1 binding to dT38-FL in Buffer HK150M5):

[DNA] = 50nM: K = (2.7 ± 1) x 109 M-1, L = (7.3 ± 1.3) x 107 M-1

[DNA] = 100nM: K = 2.7 x 109 M-1, L = 7.3 x 107 M-1

[DNA] = 200nM: K = 2.7 x 109 M-1, L = 7.3 x 107 M-1

[DNA] = 400nM: K = 2.7 x 109 M-1, L = 7.3 x 107 M-1

Figure 5E (Rim1 binding to dT38-FL in Buffer HK1000M5):

[DNA] = 100nM: K = (6.3 ± 2.3) x 108 M-1, L = (3.2 ± 0.3) x 106 M-1

[DNA] = 400nM: K = 6.3 x 109 M-1, L = 3 x 106 M-1

Figure 5F (Rim1Y85H binding to dT38-FL in Buffer HK150M5):

because of the higher [DNA], K was fixed at 3 x 109 M-1 and L was fitted for the lower [DNA].

[DNA] = 100nM: K = 3 x 109 M-1, L = (3.5 ± 0.4) x 106 M-1

[DNA] = 400nM: K = 3 x 109 M-1, L = 3 x 106 M-1

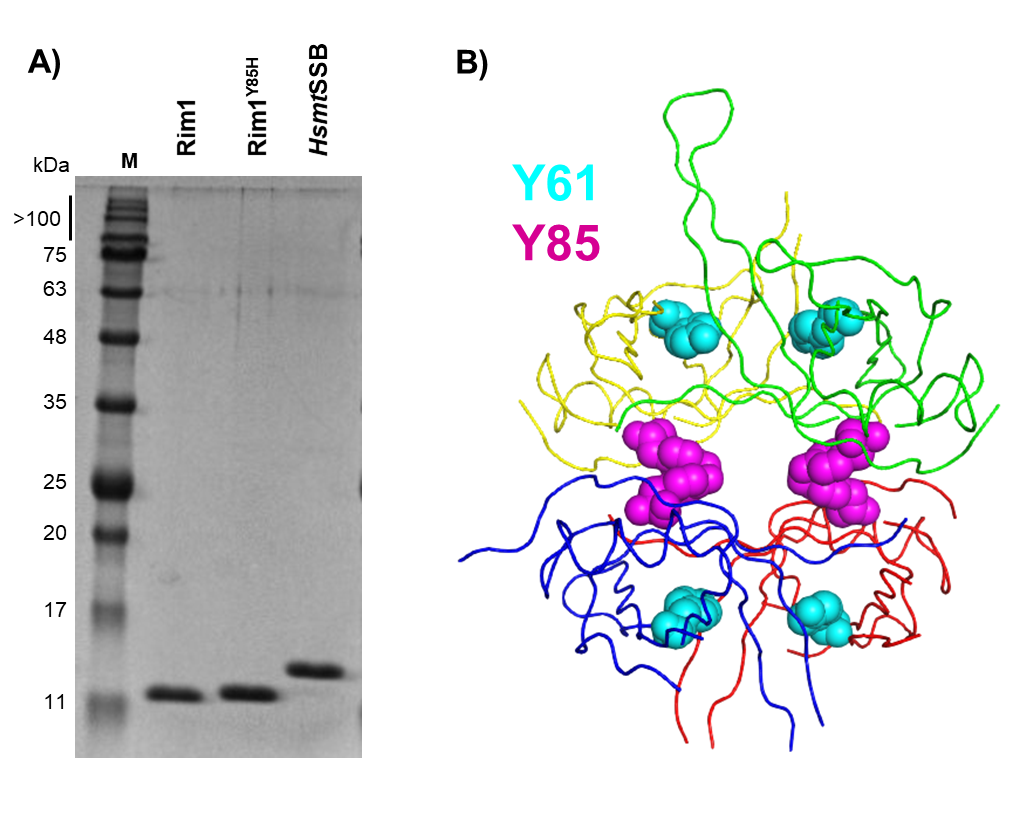
**Supplementary material references**

1. Kabsch, W. Xds. Acta Crystallogr. Sect. D Biol. Crystallogr. (2010); 66: 125–132.
2. Evans, P. Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr. (2006) Jan; 62(Pt 1):72-82.
3. CCP4: Acta Crystallogr., Sect. D: Biol. Crystallogr., 67 (2011), pp. 235–242
4. P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan,Features and development of Coot Acta crystallographica. Section D, Biol. Crystallogr., 66 (2010), pp. 486–501
5. Afonine P.V., Grosse-Kunstleve R.W., Echols N., Headd J.J., Moriarty N.W., Mustyakimov M., Terwilliger T.C., Urzhumtsev A., Zwart P.H., Adams P.D. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. Sect D: Biol. Crystallogr. (2012);68:352–367
6. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., & Read, R.J. Phaser crystallographic software. J. Appl. Cryst. (2007) 40, 658-674.
7. Davis IW, et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res (2007) 35:W375–83
8. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr (1993) 26:283–291.
9. Vaguine AA, Richelle J, Wodak SJ. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. Acta Crystallogr D Biol Crystallogr. 1999, 55(Pt 1):191-205.
10. The Pymol Molecular Graphics System (2002) [www.pymol.org](http://www.pymol.org)

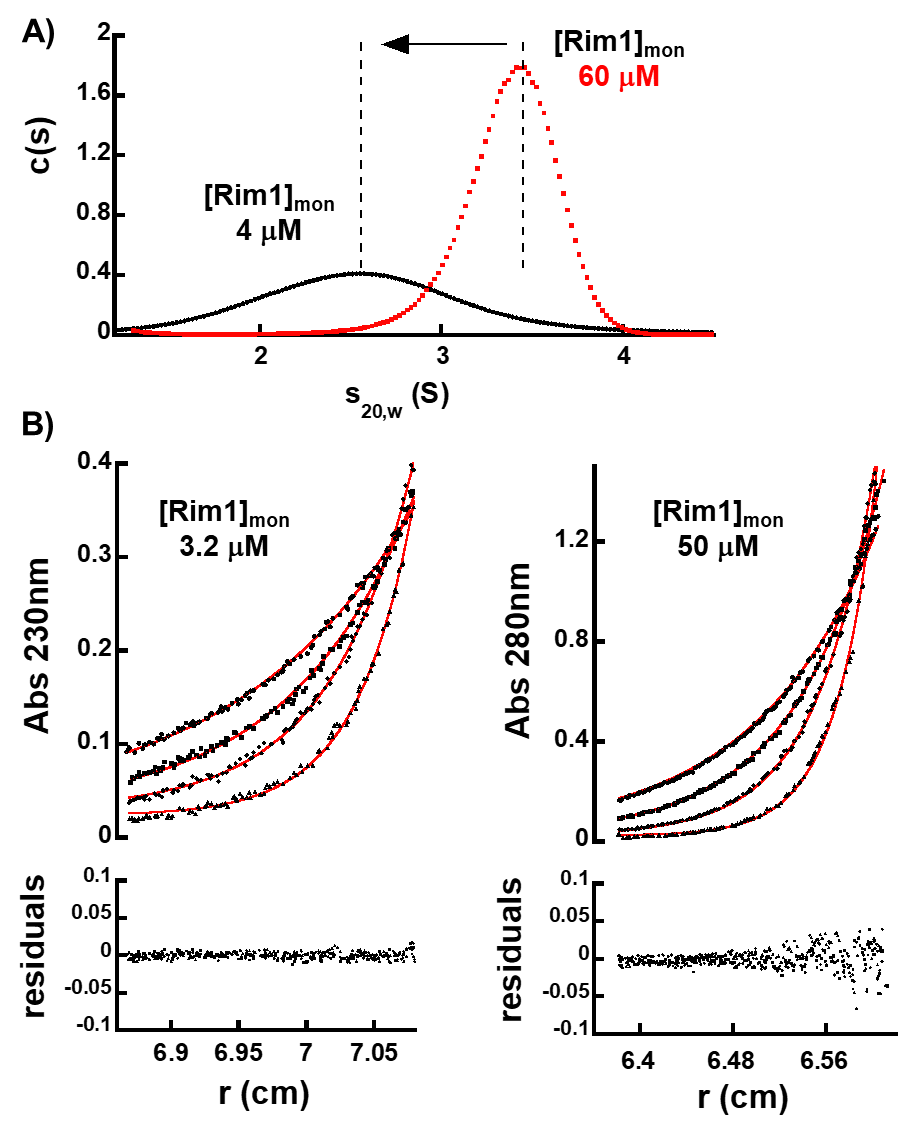
**Table S1. Structural parameters.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Data collection** | **P212121(Selmet)** | **P212121 (Native)** | **C2221 (Native)** |
| PDB | 6CQO | 6CQM | 6CQK |
| Space group | P212121 | P212121 | C2221 |
| Cell dimensions |  |  |  |
| a,b and c (Å) | 81.96,102.96,114.362 | 82.47,103.54,114.58 | 54.502,153.14,118.609 |
| α,β and γ (°) | 90.0,90.0,90.0 | 90.0,90.0,90.0 | 90.0,90.0,90.0 |
| Wavelength(Å) | 0.97625 | 0.97625 | 0.97625 |
| Resolution(Å) | 62.98-2.8 | 50.13-3.0 | 76.57-2.8 |
| Rmerge (%) | 14.0 | 13.9 | 3.8 |
| I/σI (outer shell) | 16(1.7) | 19(1.2) | 12(1.6) |
| Completeness (outer shell) (%) | 99.9(99.7) | 99.8(100) | 100(100) |
| Redundancy | 15.7 | 13.9 | 7.4 |
| Number of reflections observed | 47250 | 343304 | 93045 |
| Number of unique Reflections | 23661 | 20197 | 12631 |
| **Refinement** | | | |
| Number of atoms | 6113 | 5978 | 2896 |
| R work/Rfree (%) | 22.3/27.6 | 23.8/29.4 | 24.6/27.9 |
| r.m.s.d. | | | |
| Bond angle (°) | 0.417 | 0.37 | 1.789 |
| Bond Length(Å) | 0.002 | 0.001 | 0.0188 |
| Ramachandran analysis | | | |
| Most favored | 97.02% | 98.02% | 97.0% |
| Allowed | 2.84% | 1.84% | 2.4% |
| outlier | 0.14% | 0.14% | 0.6% |

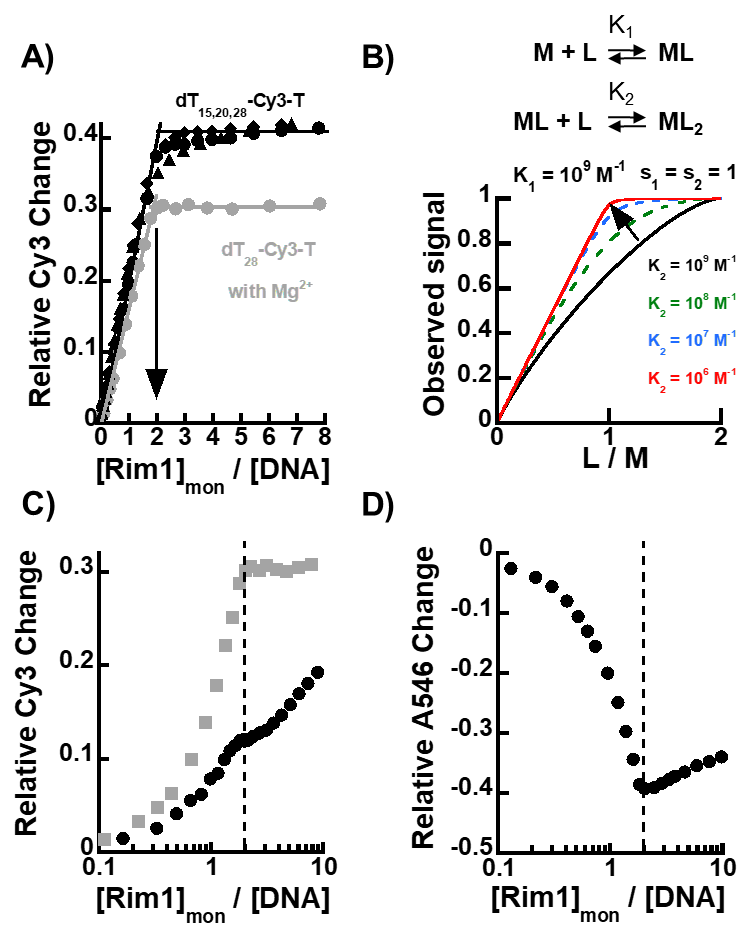
**Figure S1.** **A)** SDS-PAGE of purified proteins stained with Coomassie Blue. **B)** Model of the Rim1 tetramer showing the position of Y61 and Y85.



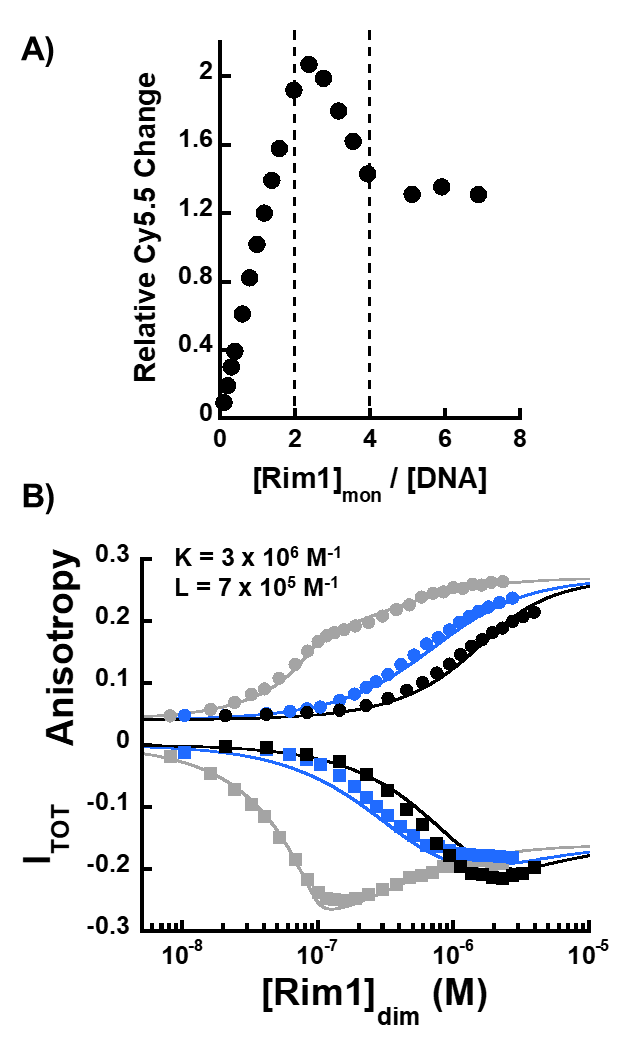
**Figure S2.** **A)** Distributions of sedimentation coefficients for Rim1 at two loading concentrations in Buffer HK150. **B)** Analysis with a single-species model of sedimentation equilibrium absorbance profiles determined at different rotor speeds (16000, 19000, 23000, 28000 rpm) and at two Rim1 concentrations in Buffer HK150.



**Figure S3. A)** Change in fluorescence intensity of dTn-Cy3-T as a function of the ratio of Rim1 concentration to the concentration of the DNA, in Buffer HK150 for n = 15 (▲) at 800 nM, n = 20 (◆) and n = 28 (⚫) at 200 nM, and in Buffer HK150M5 for n = 28 (⚫) at 200 nM. The solid lines are for reference. **B)** Simulation of the dependence of an observed signal as a function of the degree of binding (L/M) for a 2:1 binding model where binding of the second ligand does not contribute to the signal (s1=s2). As the difference in K1 and K2 increases the stoichiometry shifts from the correct 2:1 to an apparent stoichiometry of 1:1. **C)** Change in fluorescence intensity of Cy3-dT24 (⚫) as a function of the ratio of Rim1 concentration to the concentration of the DNA (400 nM), in Buffer HK150. In grey are the data for dT28-Cy3-T (200 nM) as a reference. **D)** Same as in B but for dT28-Alexa546



**Figure S4. A)** Change in fluorescence intensity of Cy5.5-dT40-Cy3-T, monitoring the change in Cy5.5 fluorescence intensity upon excitation of Cy3, as a function of the ratio of Rim1 concentration to the concentration of the DNA (200 nM), in Buffer HK150. **B)** Change in fluorescence anisotropy (circles) and total intensity (squares) of FAM-dT20 at different concentrations (100 nM blue, 850 nM black), as a function of Rim1 concentration (dimer), in Buffer HK1000M5. The solid lines are the fits with a 2:1 binding model. For reference, in grey are the data with 100 nM FAM-dT20 in Buffer HK150M5.



**Figure S5.** Change in fluorescence intensity of dT15-Cy3-T at 200 nM (squares) and 800 nM (circles) as a function of Rim1 concentration (dimer), in either Buffer HK150 (black) or HK150M5 (grey). Assuming that Rim1 binds exclusively as a dimer the isotherms can be fitted with a simple a 1:1 binding model (solid lines).

