Atomistic Ensemble Modeling and Small-Angle Neutron Scattering of Intrinsically Disordered Protein Complexes: Applied to Minichromosome Maintenance

S. Krueger,[†]*J.-H. Shin,[§] S. Raghunandan,[†] J. E. Curtis,[†] and Z. Kelman[‡¶]

[†]National Institute of Standards and Technology Center for Neutron Research and [‡]Biochemical Sciences Division, National Institute of Standards and Technology, Gaithersburg, Maryland; [§]Division of Applied Biology and Chemistry, College of Agriculture and Life Sciences, Kyungpook National University, Daegu, Republic of Korea; and [¶]Institute for Bioscience and Biotechnology Research, Rockville, Maryland

Supporting Material

SASSIE Overlap Cutoff Comments

The value of 3 Å was chosen as a cutoff distance criterion for the “overlap” criterion based on both a survey of such inter-atomic distances in the protein data bank and via numerical studies which determined the optimal distance such that subsequent minimization steps can be completed without the existence of gross overlap. In our experience, this value should not be chosen to be less than 2.5 A. If one uses a value greater than 3 Å, then one would expect less overlap between side-chain atoms and, as larger values are used, perhaps less coverage of configurational space.

Figure S1

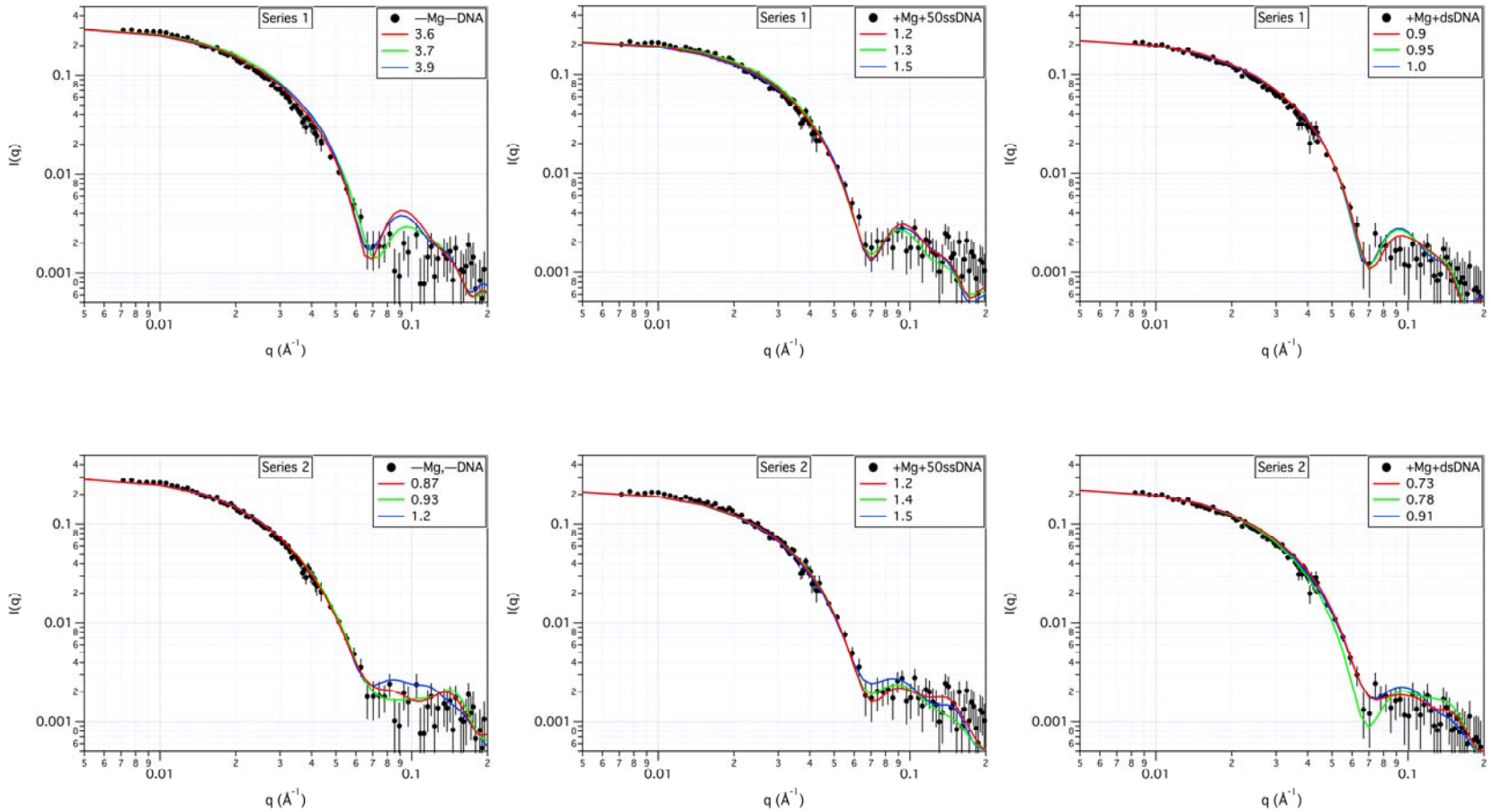


Figure S1. Representative Best Fit Model SANS Curves to SANS Data. The top row shows the log(I) vs log(q) plots of the measured SANS data for the (–Mg–DNA) (left), (+Mg+50ssDNA) (middle) and (+Mg+dsDNA) (right) samples, along with representative examples of best fit model SANS curves from the ensemble of Series 1 structures. Similar best fit model curves for the Series 2 structures are shown along with the measured SANS data in the bottom row. The best fit family of structures was chosen based on the lowest ≈ 20 % of the χ^2 values for each series. χ^2 values are listed in the legend. Error bars on the measured SANS data represent plus and minus the combined standard uncertainty of the data collection.

Figure S2a

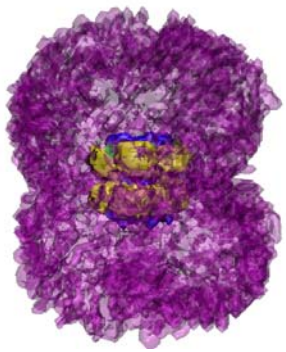
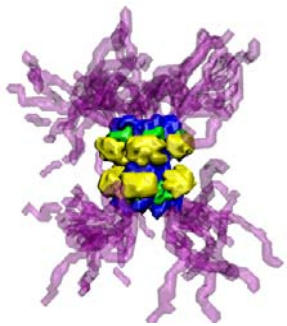
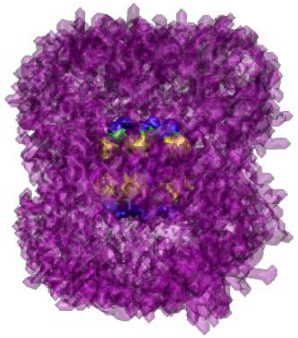
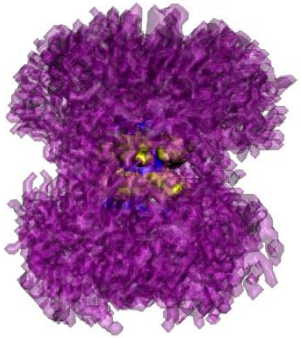
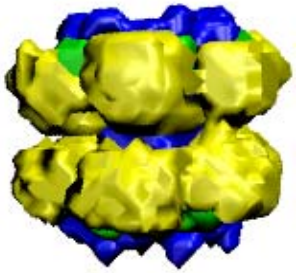
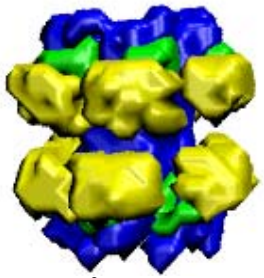
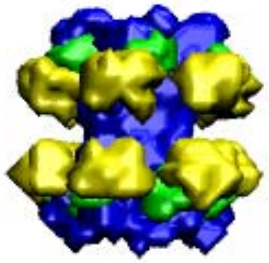
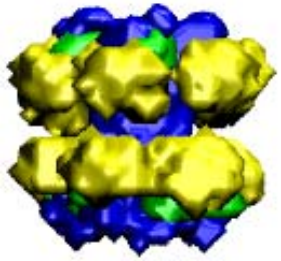
Series	All Accepted Structures	(–Mg–DNA)	(+Mg, +50ssDNA)	(+Mg+dsDNA)
1	 (20142)	 (6/20142 = 0.0003)	 (980/20142 = 0.049)	 (318/20142 = 0.016)
		 $\chi^2 \leq 4.0$	 $\chi^2 \leq 1.5$	 $\chi^2 \leq 1.1$

Figure S2b

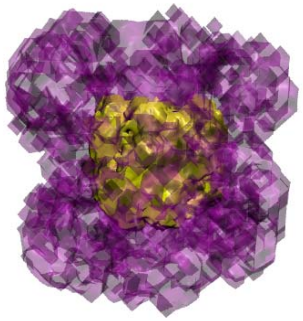
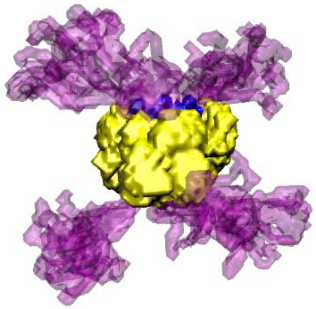
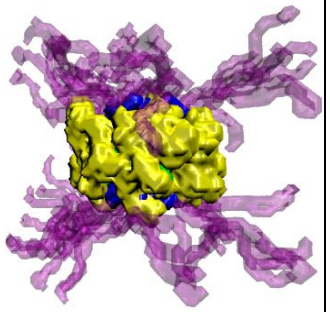
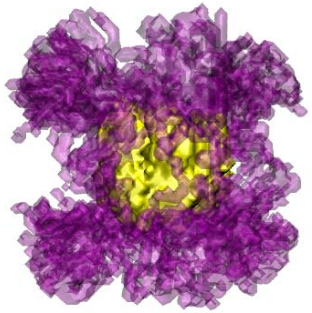
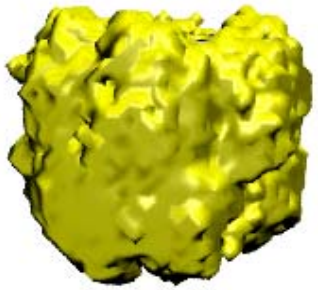
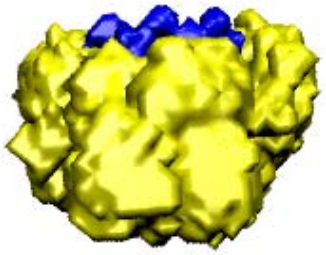
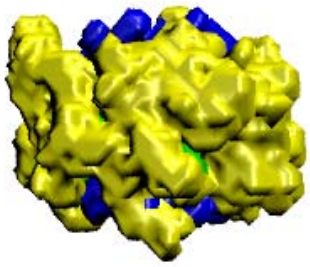
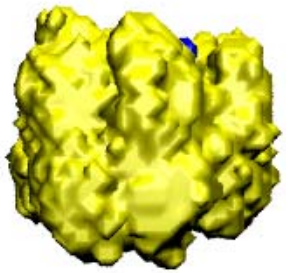
Series	All Accepted Structures	(–Mg–DNA)	(+Mg, +50ssDNA)	(+Mg+dsDNA)
2	 (7050)	 (88/7050 = 0.012)	 (10/7050 = 0.0014)	 (326/7050 = 0.046)
		 $\chi^2 \leq 1.1$	 $\chi^2 \leq 1.4$	 $\chi^2 \leq 1.1$

Figure S2. Structure Density Plots. Isodensity surface plots for a) Series 1 and b) Series 2 structures showing the conformation space occupied by all of the accepted structures in each ensemble and for the best fit structures of each ensemble as compared to the SANS data for the (–Mg–DNA), (+Mg+50ssDNA) and (+Mg+dsDNA) samples. The first row shows the density plots for the full molecule and the second row shows the same plots for the molecule without the C-terminal unstructured region (residues 244 to 286). The range of χ^2 values in each ensemble is stated and the fraction of structures in each ensemble, compared to the total number of accepted structure, is shown in parenthesis. Color coding is the same as for Figure 1. For the Series 1 structures in a), only residues 244 to 286 (purple) are allowed to be flexible. Residues 89 to 108 (green) are positioned as in the crystal structure. For Series 2 structures in b), residues 89 to 108 (green) and 244 to 286 (purple) are both allowed to be flexible.