Role of IscX in Iron-Sulfur Cluster Biogenesis in *Escherichia coli*

Jin Hae Kim,[a,§] Jameson R. Bothe,[b,§] Ronnie O. Frederick,[a] Johneisa C. Holder,[a] and

John L. Markley[a,b,*]

[a]Mitochondrial Protein Partnership, Center for Eukaryotic Structural Genomics, and [b]Department of Biochemistry, University of Wisconsin, 433 Babcock Drive, Madison, WI 53706

[§]These authors contributed equally to the work.

[*]To whom correspondence should be addressed. Phone: (608) 263-9349. Fax: (608) 262-3759.

Email: jmarkley@wisc.edu.

Supporting Information

Experimental Procedures

Production and purification of IscX

The *Escherichia coli* IscX gene was cloned into the BsaI and and XhoI (NEB) restriction sites of the pE-SUMO plasmid (Lifesensors) by the polymerase chain reaction (PCR) employing AccuPrime PCR Supermix^{TM} (Life Technologies). *E. coli* genomic DNA (Sigma-Aldrich, St. Louis, MO) was used as the template with gene-specific DNA primers (Biotech Center, University of Wisconsin-Madison). The PCR program used for gene isolation consisted of: initial denaturation at 95 °C for 5 min; extension reaction of 30 cycles of (95 °C for 20 s, 55 °C for 20 s, 68 °C for 30 s); and final fill-in step at 68 °C for 20 min. The IscX gene was ligated into the pE-SUMO plasmid by T4 DNA ligase under 800 cycles of (30 ºC for 10 s followed by 10 ºC for 10 s). *E. coli* 10G cells (Lucigen) were transformed with the ligated plasmids, plated onto YT plates (supplemented with 35 μg/mL kanamycin), and incubated at 37 ºC overnight. Colonies were picked and PCR screened for positive recombinants, and the expected sequence was confirmed by DNA sequencing.

Rosetta2(DE3) (Novagen) and BL21(DE3) CodonPlus RILP (Stratagene) *E. coli* host strains were used in small-scale (0.5 mL) screening for production of IscX protein from an auto-induction medium.[1,2] Transformed cells were incubated at 25 ºC overnight in YT medium (8 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl) supplemented with 1% glucose, 100 μg/mL kanamycin, and 35 μg/mL chloramphenicol. IscX protein production was induced by transferring the overnight culture into 0.5 mL Terrific Broth with glycerol (TB+g) auto-induction medium, and growing the cells at 25 ºC and shaking at 250 rpm for 24 h. The composition of the TB+g

auto-induction medium was: 12 g/L tryptone, 24 g/L yeast extract, 2.31 g/L KH_2PO_4, 12.54 g/L K_2HPO_4 , 10 g/L glycerol, 0.375% w/v sodium aspartate, 2 mM $MgSO_4$, 0.015% w/v glucose, 0.8% w/v glycerol, 0.5% w/v α–lactose, and a few drops of sterile antifoam.

The auto-induction method[1,2] was also used in the large-scale production of IscX. The expression strain BL21(DE3)CodonPlus RILP transformed with the IscX plasmid was used to inoculate 1 mL of YT medium supplemented with 1% glucose, 50 μg/mL kanamycin, and 35 μg/mL chloramphenicol; the inoculated medium was incubated for 3 h at 37 ºC with shaking at 250 rpm. This starter was used to inoculate 50 mL YT supplemented with 1% glucose, 50 μg/mL kanamycin, and 35 μg/mL chloramphenicol, which was incubated overnight at 25 ºC with shaking at 250 rpm. 12 mL aliquots of this growth were used to seed each of four 500 mL of TB+g auto-induction medium in disposable plastic drink bottles;[3] and these were incubated at 37 ºC with shaking at 250 rpm for 5 h. Subsequently, the temperature was lowered to 18 ºC, and the cultures were incubated for additional 24 h. Cells were harvested by centrifugation at 4,000 g for 30 min, and the cell pastes were stored at −80 ºC. Isotopically-labeled IscX was produced from an M9-based minimal medium, and with the identical scale-up method that was employed for expression of unlabeled proteins except that 0.2 mM isopropyl-1-thio-β-D-galactoside (IPTG) was used for induction. The M9 medium to produce [U-^{15}N]-IscX was supplemented by 1 g/L $^{15}NH_4Cl$, 4 g/L glucose, 1 mL/L of 1000x metal mixture (0.1 M $FeCl_3 \cdot 6H_2O$, 1 M $CaCl_2$, 1 M $MnCl_2 \cdot 4H_2O$, 1 M $ZnSO_4 \cdot 7H_2O$, 0.2 M $CoCl_2 \cdot 6H_2O$, 0.1 M $CuCl_2 \cdot 2H_2O$, 0.2 M $NiCl_2 \cdot 6H_2O$, 0.1 M $Na_2MoO_4 \cdot 5H_2O$, 0.1 M $Na_2SeO_3 \cdot 5H_2O$, and 0.1 M H_3BO_3), 1 mL/L of 1000x vitamin mixture (200 mM vitamin B_{12}, 200 mM nicotinic acid, 200 mM pyridoxine, 200 mM thiamine, 200 mM *p*-aminobenzoic acid, 5 mM folic acid, and 5 mM riboflavin), 30mg/L thiamine, 0.1 mM $CaCl_2$, and 2 mM $MgSO_4$ along with 50 μg/mL kanamycin and 35 μg/mL chloramphenicol.

For protein purification, the cell paste harvested from 2-L medium was first resuspended in lysis buffer: 70 mL 1st IMAC buffer containing 10 μL Benzonase 25 U/μL (Novagen), 10 μL recombinant lysozyme 30 KU/μL (Novagen), 10 μL RNase A (10 mg/mL, Qiagen), and 0.2% NP-40 (Sigma). The 1st IMAC buffer consisted of 20 mM Tris pH 8, 500 mM NaCl, 10% glycerol, 5 mM imidazole, and 1 mM phenylmethylsulfonyl fluoride (PMSF). The resuspended cells were lysed by sonication (Misonix 3000) for 15 min at 4 ºC. Sonicated cells were centrifuged for 30 min at 25,000 rpm, and 0.1% w/v polyethylenimine (PEI, Fluka) was added and the spin was repeated. Ammonium sulfate was added up to 70% along with 2 mM dithiothreitol (DTT), and the sample was spun for 30 min at 25,000 rpm (JA-30.50 Ti rotor). The protein precipitate was resuspended in 50 mL of the 1st IMAC buffer, and then loaded onto 10 mL Qiagen Superflow IMAC resin. The IMAC column was washed with 10 column volumes of the 1st IMAC buffer followed by 10 column volumes of IMAC wash buffer (the 1st IMAC buffer plus 30 mM imidazole). IscX protein eluted with the IMAC elution buffer (the 1st IMAC buffer plus 250 mM imidazole). The eluted protein fractions were dialyzed overnight in a SUMO protease cleavage buffer (20 mM Tris pH 8, 100 mM NaCl and 2 mM DTT) in the presence of SUMO protease (0.5 mg, CESG cloned protease). The cleaved target protein was loaded onto the 2nd IMAC column that served to remove the His-tagged SUMO domain. The IscX protein (in the flow-through fractions) was collected and polished by gel filtration (HiLoad 26/60 Superdex 200 column, GE Healthcare) with a buffer containing 20 mM Tris pH 8, 150 mM NaCl, and 2 mM DTT. Protein purity at each stage was assessed by SDS-PAGE. A Sartorius Vivaspin 3 kDa concentrator was used to concentrate IscX, and the solution was frozen and stored at −80 ºC until used.

Table S1. Summary of the SAXS data.

Sample	IscX-IscS (5.0 mg·mL^{-1})	IscU-IscS (5.0 mg·mL^{-1})	Fdx-IscS (5.0 mg·mL^{-1})	IscX-IscU-IscS (5.0 mg·mL^{-1})
Data-collection parameters				
Instrument	APS – Sector 12-ID-B	APS – Sector 12-ID-B	APS – Sector 12-ID-B	APS – Sector 12-ID-B
Beam geometry	Synchrotron	Synchrotron	Synchrotron	Synchrotron
Wavelength (Å)	0.866 Å	0.866 Å	0.866 Å	0.866 Å
q range (Å^{-1})	0.004 – 0.522	0.004 – 0.522	0.004 – 0.522	0.004 – 0.522
Exposure time (s)	20	20	20	20
Concentration range (mgml^{-1})	1.3 -5.0	1.3 -5.0	1.3 -5.0	1.3 -5.0
Temperature (K)	298	298	298	298
Data-collection parameters				
I(0) [from P(r)]	0.2310 ± 0.0007	0.2882 ± 0.0009	0.2309 ± 0.0005	0.3062 ± 0.011
R_g (Å)[from P(r)]	31.8 ± 0.1	36.6 ± 0.1	31.8 ± 0.1	36.0 ± 0.2
I(0) (from Guinier)	0.2303 ± 0.0009	0.2861 ± .0001	0.2309 ± 0.0001	0.3030 ± 0.0015
R_g (Å) (from Guinier)	31.5 ± 0.3	35.8 ± 0.4	32.1 ± 0.3	35.1 ± 0.2
D_{max} (Å)	110 ± 5	128 ± 5	106 ± 5	130 ± 5
Porod volume estimate 10^3 (Å^3)	145 ± 10	147 ± 10	146 ± 10	166 ± 10
Dry volume calculated from sequence 10^3 (Å^3)	128	143	139	162
Molecular-mass determination				
Molecular mass [I(0)] (kDa)	88 ± 9	102 ± 10	91 ± 9	108 ± 11
Molecular mass [V_c] (kDa)	86 ± 9	92 ± 9	91± 9	100 ± 10
Calculated molecular mass from sequence (kDa)	106	118	115	133
Software Employed				
Primary data reduction	MATLAB	MATLAB	MATLAB	MATLAB
Data processing	PRIMUS/IGOR	PRIMUS/IGOR	PRIMUS/IGOR	PRIMUS/IGOR
Ab initio analysis	DAMMIF	DAMMIF	DAMMIF	DAMMIF
Validation and averaging	DAMAVER	DAMAVER	DAMAVER	DAMAVER
Rigid-body modeling	SASREF	SASREF	SASREF	SASREF
Computation of model intensities	CRYSOL/FoXS	CRYSOL/FoXS	CRYSOL/FoXS	CRYSOL/FoXS
Three-dimensional graphics representations	Chimera	Chimera	Chimera	Chimera

Sample	IscS (1.3 mg·mL^{-1})	CyaY-IscS (10.0 mg·mL^{-1})
Data-collection parameters		
Instrument	Bruker Nanostar	Bruker Nanostar
Beam geometry	2 pinholes (500 μm)	2 pinholes (500 μm)
Wavelength (Å)	1.5418 Å	1.5418 Å
q range (Å^{-1})	0.012 – 0.240	0.012 – 0.240
Exposure time (h)	4	4
Concentration range (mgml^{-1})	1.3 - 5	1.3 -10.0
Temperature (K)	298	298
Data-collection parameters		
I(0) [from *P*(r)]	1171 ± 5	12290 ±10
R_g (Å)[from *P*(r)]	32.0 ± 0.2	32.1 ± 0.1
I(0) (from Guinier)	1170 ± 7	12258 ± 14
R_g (Å) (from Guinier)	31.8 ± 0.3	31.9 ± 0.2
D_{max} (Å)	105 ± 5	111 ± 5
Porod volume estimate 10^3 (Å^3)	125 ± 10	143 ± 10
Dry volume calculated from sequence 10^3 (Å^3)	109	139
Molecular-mass determination		
Molecular mass [I(0)] (kDa)	88 ± 9	114 ± 11
Molecular mass [V_c] (kDa)	81 ± 8	88 ± 9
Calculated molecular mass from sequence (kDa)	90	115
Software Employed		
Primary data reduction	SAXS (Bruker)	SAXS (Bruker)
Data processing	PRIMUS	PRIMUS
Ab initio analysis	DAMMIF	DAMMIF
Validation and averaging	DAMAVER	DAMAVER
Rigid-body modeling	SASREF	SASREF
Computation of model intensities	CRYSOL/FoXS	CRYSOL/FoXS
Three-dimensional graphics representations	Chimera	Chimera

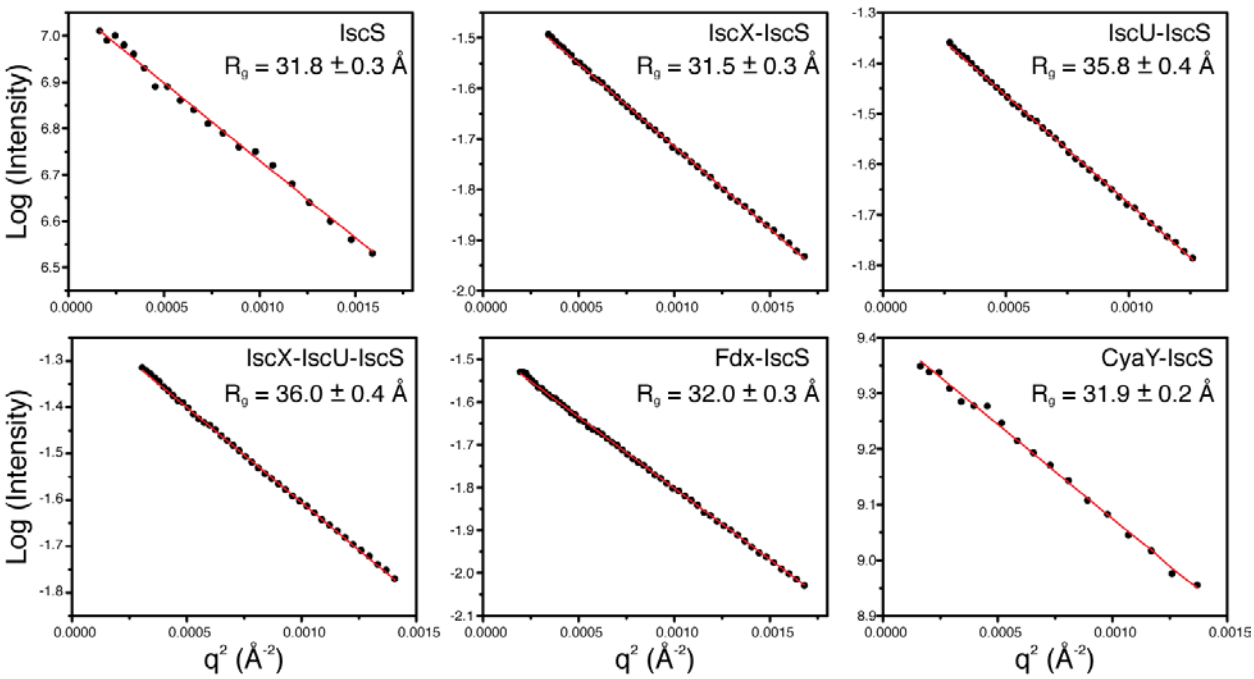

Figure S1. Guinier plots of experimental SAXS data used to determine structural parameters and carry out rigid body and molecular shape modeling.

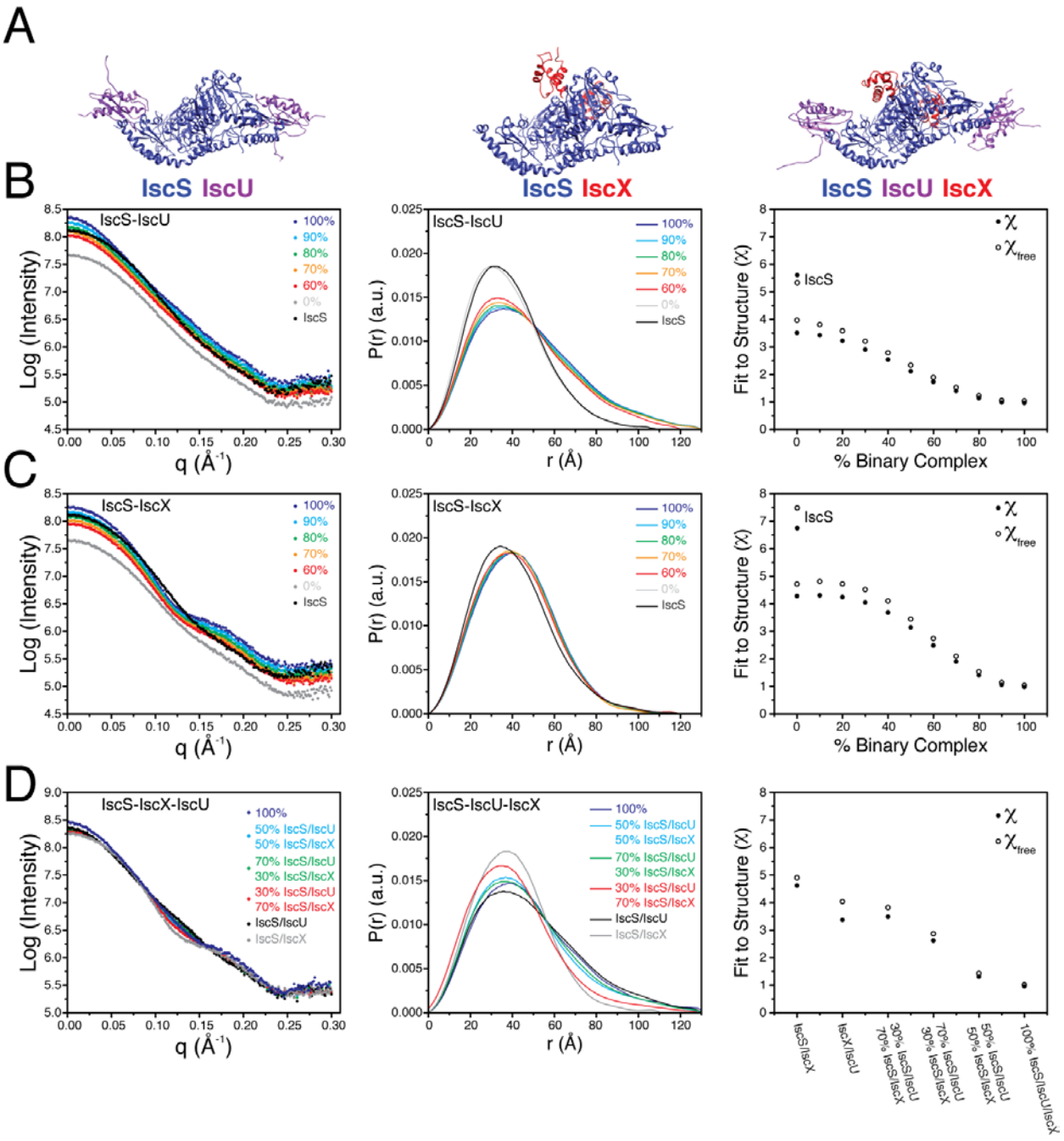

Figure S2. Simulation of the effect of weak complex formation on SAXS data for the IscU-IscS, IscX-IscS, and IscX-IscU-IscS complexes shown in (A). Free IscS was simulated as a homodimer. SAXS curves were generated by applying the CRYSOL[4] software package to structural models generated from rigid body modeling simulations using structural models from the Protein Data Bank (1P3W, 2L4X, 2BZT). The synthetic data sets were then noise corrupted (with errors consistent with the experimental data) by selecting a random number in a Gaussian distribution with a standard deviation proportional to experimental uncertainty in our measurements centered about each synthetic data point. The noise-corrupted synthetic data sets were then combined using concentration-weighted linear combinations of the scattering curves

from each molecular component present. (B) *Left*, noise-corrupted data sets of varying extents of IscS/IscU complex formation; *middle*, distance distribution functions calculated from each SAXS curve using GNOM[5] software; *right*, χ and χ_{free}[6] values resulting to the fit of each SAXS curve to the IscS/IscU structure depicted in A. (C) *Left*, noise-corrupted data sets of varying extents of IscX-IscS complex formation; *middle*, distance distribution functions calculated from each SAXS curve using GNOM;[5] *right*, χ and χ_{free}[6] values resulting to the fit of each SAXS curve to the IscS/IscX structure depicted in A. (D) *Left*, noise-corrupted data sets of varying extents of IscS/IscU/IscX complex formation; *middle*, distance distribution functions calculated from each SAXS curve using GNOM;[5] *right*, χ and χ_{free}[6] values for the fit of each SAXS curve to the IscX-IscU-IscS structure depicted in A.

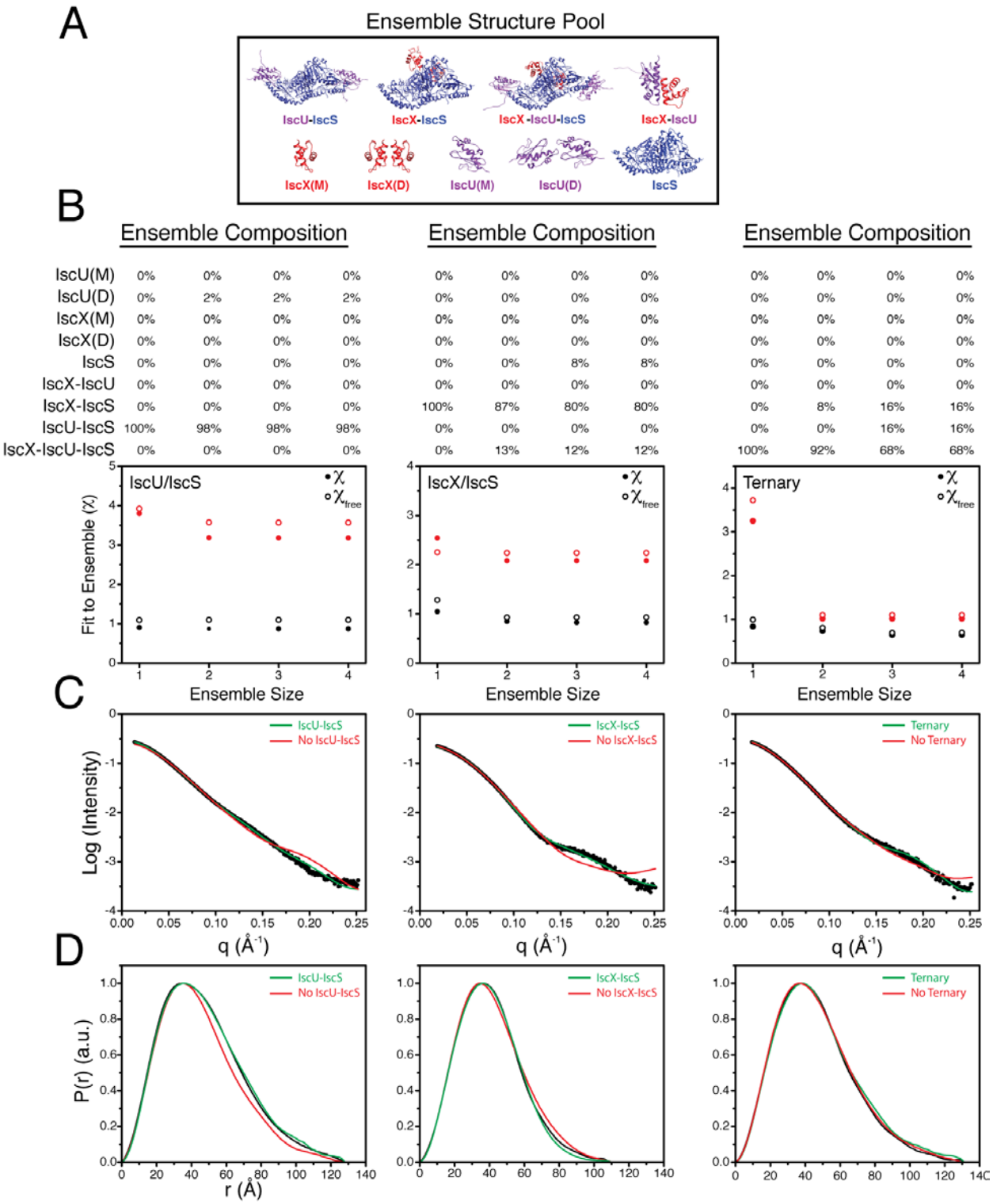

Figure S3. Use of the Minimal Ensemble Search (MES) algorithm[7] to fit experimental data from the IscU, IscX, IscS binary and ternary complexes. For the analysis we obtained the same results using the MES Linux program[7] and the FOXS web server.[8] (A) Pool of structures used in generating ensemble fits. (B) The selected ensemble compositions are shown for fits to the SAXS data for IscU/IscS, IscX/IscS, and IscX/IscU/IscS. Ensemble fits are shown for sizes up to

4 to show that minimal improvements are made to fitting the data when allowing for additional components. In some cases expanding the ensemble size beyond 2 did not improve the χ fit, so the MES program output assigned additional structural components a weight of zero. The graphs show χ (closed circles) and $\chi_{free}{}^{6}$ values (open circles) for fits (black) in which the full ensemble of species was considered and (red) fits in which the predominant complex (IscU-IscS, IscX-IscS, or Iscx-IscU-IscS, in the respective plots) was excluded from the structure ensemble pool. (C) SAXS profiles of the experimental data (dots) for mixtures of IscU/IscS, IscX/IscS, IscX/IscU/IscS compared to profiles for their respective complex structural model (green) and the ensemble fit when the predominant complex structural model was excluded from the structure pool (red). (D) Pairwise distance distribution functions (PDDF's) for mixtures of IscU/IscS, IscX/IscS, IscX/IscU/IscS (black) compared to the PDDF's of their respective complex structural model (green) and the ensemble fit when the predominant complex structural model was excluded from the structure pool (red).

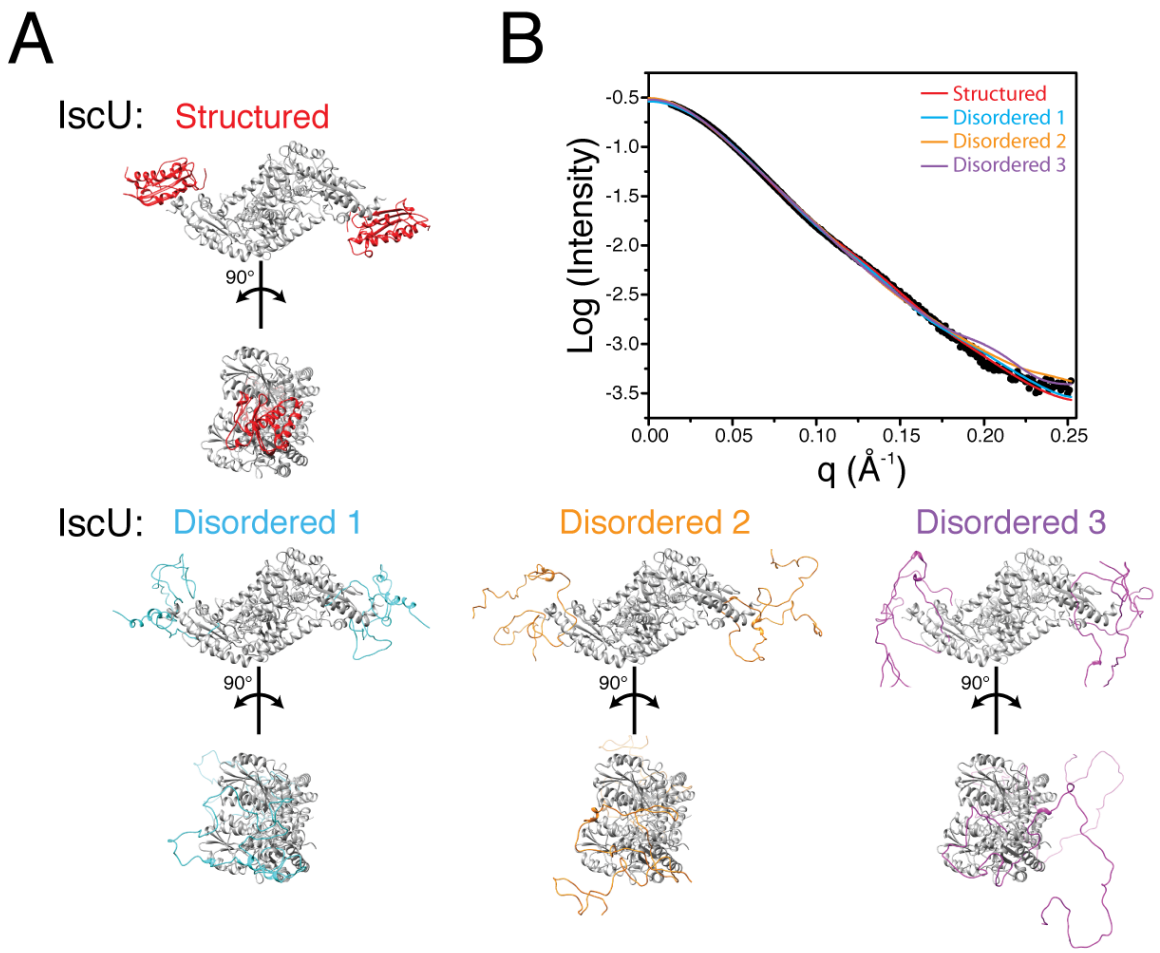

Figure S4. Experiment to determine whether SAXS data are sufficient to distinguish various conformations of IscU in the IscU-IscS complex. (A) Structural representations of the IscU-IscS complex were created with models of IscU in one structured and three disordered conformations. The models of the IscU-IscS complex were generated using SASREF[9] (as described in Experimental Procedures) with IscS (gray) and IscU (colored). The structured conformation of IscU (red) was generated from the NMR solution structure of the D39A IscU variant which populates the structured state 95 percent of the time. Disordered conformations of IscU (cyan, orange, purple) were randomly selected from a 100 ns molecular dynamics trajectory generated with Amber 12 with the structured conformation as the starting structure. The simulation was carried out using a generalized Born implicit solvent model[10] along with SHAKE[11] and Langevin dynamics with a collision frequency of 2 ps^{-1}.[12] A simulation temperature of 500K was found to yield unfolded structures. (B) SAXS curves (colored lines) generated from the various models are compared to the experimental data for the IscU-IscS complex (circles).

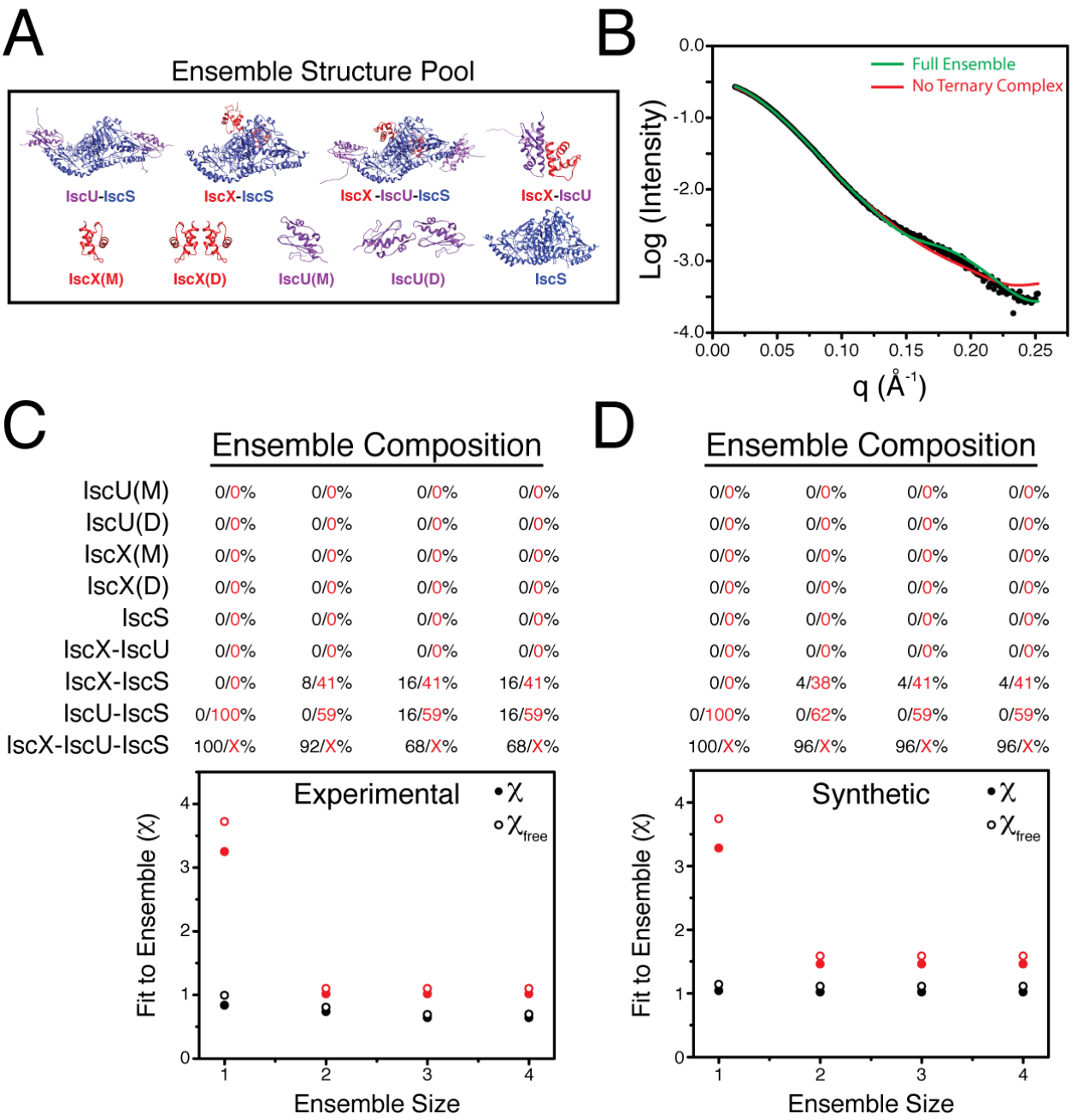

Figure S5. Use of the Minimal Ensemble Search (MES) algorithm[7] to analyze the effect of weak ternary complex formation. Synthetic data were generated by using the structure of the ternary complex, and the resulting data points were corrupted to match the level of error in the experimental data. (A) Pool of structures used to generate the ensemble fits. (B) Best ensemble fits to experimental data when a ternary complex was included in the structure pool (green) or when excluded from the structure pool (red). (C) Ensemble compositions and fits of experimental data to varying ensemble sizes determined from MES where χ and χ_{free}[6] values are shown as closed and open circles, respectively. Black, the full ensemble was allowed for selection. Red, the ternary complex was excluded from the structure pool as indicated by "X". Fits of ensemble sizes up to 6 gave identical results to those obtained for ensemble sizes 3 and 4 (data not shown). (D) Ensemble compositions and fits of synthetic data to varying ensemble

sizes determined from MES where χ and χ_{free}[6] values are shown as closed and open circles, respectively. Black, the full ensemble was allowed for selection. Red, the ternary complex was excluded from the structure pool as indicated by "X". Fits of ensemble sizes up to 6 gave results identical to those obtained for ensemble sizes 3 and 4 (data not shown).

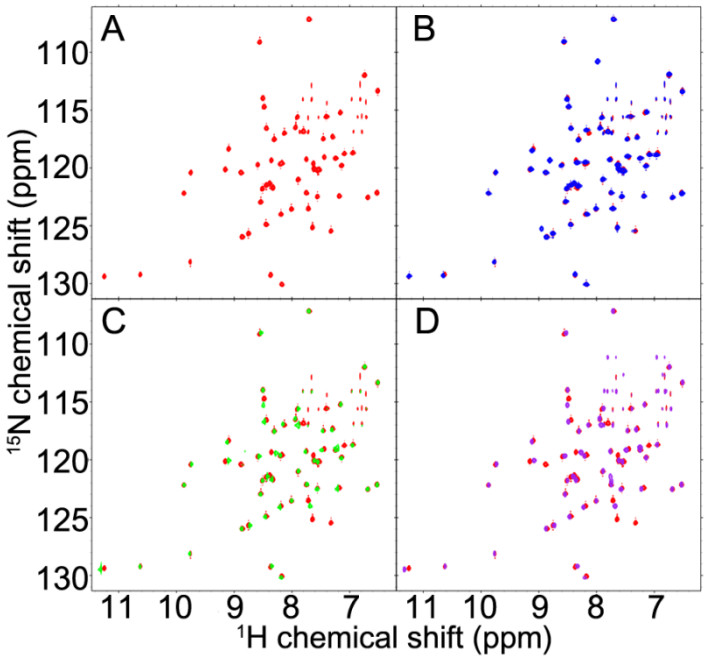

Figure S6. 2D ^{15}N-TROSY-HSQC spectra of 0.5 mM [U-^{15}N]-IscX showing that Fe^{2+} from added ferrous ammonium sulfate binds IscX much more avidly than does Fe^{3+} from added ferric ammonium citrate. (A) Spectrum of IscX (red). (B) Spectrum of [U-^{15}N]-IscX in the presence of 3-fold molar excess Fe^{3+} (blue). (C) Spectrum of [U-^{15}N]-IscX with 3-fold molar excess Fe^{3+} following the addition of excess DTT to reduce the iron (green). (D) Spectrum of [U-^{15}N]-IscX in the presence of 3-fold molar excess Fe^{2+} (purple). For comparison, the spectrum of [U-^{15}N]-IscX alone (red) is shown in panels B−D and the NMR signal perturbation profiles by Fe^{2+}/Fe^{3+} are shown in Figure 6. All the experiments were conducted anaerobically. The addition of Fe^{3+} led to minimal perturbation of the spectrum of [U-^{15}N]-IscX. The minor perturbations that were observed do not correspond to the Fe^{2+}-binding site (the largest perturbation by Fe^{3+} was at H33) and thus appear to result from non-specific binding. However, when the Fe^{3+} was reduced to Fe^{2+}, the spectrum showed large chemical shift perturbations indicative of metal binding. Equivalent perturbations occurred upon direct addition of Fe^{2+}.

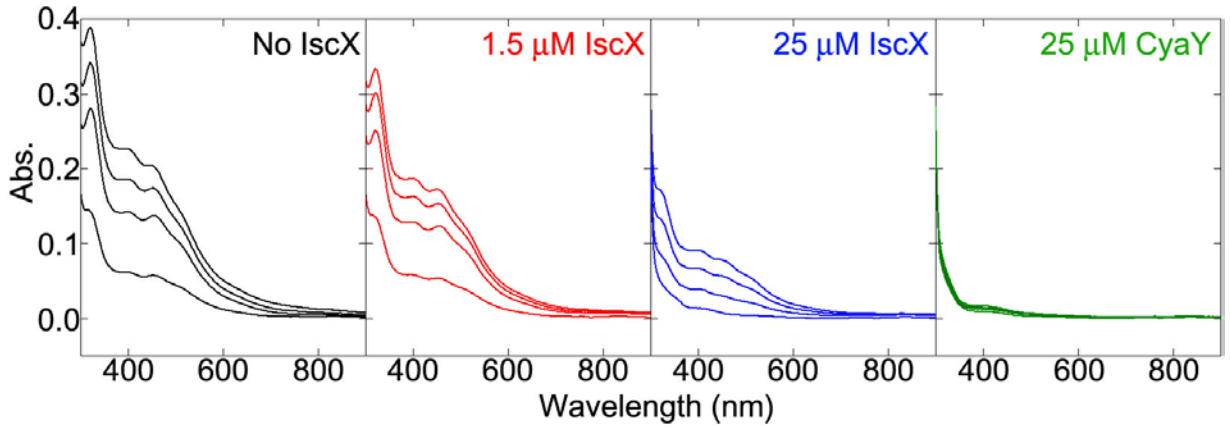

Figure S7. The uv/vis spectra (measured from 900 nm to 300 nm) of Fe-S cluster assembly reactions in the presence of IscX or CyaY. All the reaction mixtures contained 25 μM IscU, 0.5 μM IscS, 5 mM DTT, 125 μM L-cysteine, and 125 μM ferrous ammonium sulfate. The color codes used here are the same used in Figure 9 (main text), which shows the time course of the absorbance at 456 nm: (black) reaction mixture without IscX or CyaY; (red) reaction mixture containing 1.5 μM IscX; (blue) reaction mixture containing 25 μM IscX; (green) reaction mixture containing 25 μM CyaY. In each panel, the spectrum with lowest absorbance was taken immediately after preparing the reaction mixture, and subsequent spectra were taken at 10-min intervals.

References

(1) Studier, F. W. *Protein Expr. Purif.* **2005**, *41*, 207.
(2) Blommel, P. G.; Becker, K. J.; Duvnjak, P.; Fox, B. G. *Biotechnol. Prog.* **2007**, *23*, 585.
(3) Sanville, M. C.; Stols, L.; Quartey, P.; Kim, Y.; Dementieva, I.; Donnelly, M. I. *Protein Expr. Purif.* **2003**, *29*, 311.
(4) Svergun, D.; Barberato, C.; Koch, M. H. J. *J. Appl. Cryst.* **1995**, *28*, 768.
(5) Svergun, D. I. *J. Appl. Cryatallog.* **1992**, *25*, 495.
(6) Rambo, R. P.; Tainer, J. A. *Nature* **2013**, *496*, 477.
(7) Pelikan, M.; Hura, G. L.; Hammel, M. *Gen. Physiol. Biophys.* **2009**, *28, 174.*
(8) Schneidman-Duhovny D.; Hammel, M.; Tainer J. A.; Sali, A. *Biophys. J.* **2013**, *105, 962.*
(9) Petoukhov, M. V.; Svergun, D. I. *Biophys. J.* **2005**, *89*, 1237.
(10) Mongan, J.; Simmerling, C.; McCammon, J. A.; Case, D. A.; Onufriev, A. *J. Chem. Theory Comput.* **2007**, *3*, 156.
(11) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(12) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. *Biopolymers* **1992**, *32*, 523.